Regional Coordination Group
on Economic Issues

REPORT
OF THE SECOND WORKSHOP ON AN ALTERNATIVE APPROACH TO THE SEGMENTATION OF FISHING FLEETS

Edited by Erik Sulanke and Dr. Jörg Berkenhagen

Workshop organized and hosted by:

Thünen Institute of Sea Fisheries, Bremerhaven, Germany, part of

Johann Heinrich von Thünen Institute

Federal Research Institute for

Rural Areas, Forestry and Fisheries

28 – 30 March 2022

Online Workshop

Introduction

Our second workshop on a newly developed, alternative approach to the segmentation of fishing fleets was held from 28th to 30th of March 2022. Due to travel restrictions caused by the global Covid-19 pandemic, the workshop took place online. Cisco Webex provided the online platform.

The workshop was hosted by the Thünen Institute for Sea, Fisheries, part of the Johann Heinrich von Thünen Institute, Germany's Federal Research Insitute for rural areas, forestry, and fisheries.

Thirty-seven experts representing 16 nations, the ICES, and DG MARE, participated in the workshop. Fifteen national fisheries data sets were analyzed. The list of participants can be found in Annex 1. Erik Sulanke and Jörg Berkenhagen chaired the workshop. The agenda can be found in Annex 2.

Executive summary

A new approach to the segmentation of fishing fleets was developed in a DCF pilot project at the Thünen Institute of Sea Fisheries and transferred to an R package, referred to as "FS-package" in the following. In March 2021, the first workshop on the novel approach and the FS-package was held with 34 experts representing 15 nations, and major progress in improving the package was made. After implementing the suggestions made by the attendants of the first workshop and identifying the next urgent steps for further developing the novel approach, the creators of the approach organized a second workshop and formulated the following ToRs:

1. Harmonize the data preparation and develop a standardized protocol.
2. Clarify the application of the fleet segmentation approach and evaluate newly developed tools.
3. Establish regionally consistent fleet segments over multiple member states operating in the same fishing regions.

Besides the mentioned ToRs, various technical issues needed to be addressed before the workshop, the most significant being the publication of the FS-package in the publically available repository GitHub. Regarding ToR 1, newly established, standardized thresholds in a key function assigning stocks to catch data as well as a novel function for automatic data preparation were well received by the participants. However, an appropriate way to pre-separate fleet data to improve the clustering result was not agreed upon. The remaining issues include the grouping of demersal seiners and demersal trawlers in the DCF, no clear agreement on appropriate length pre-segmentation, and the possibility of a regional pre-segmentation, e.g., ICES and non-ICES area fleets.

The workshop participants gave overall positive feedback on the additional new functions of the FS-package. Especially the new functions making use of target assemblages were highlighted, as they have the potential to foster comparability between economic and metiér-based biological fleet classification methods. The regional group work, which was specified in ToR3, comprised the major work conducted during the workshop. A full list of the analyzed regions can be found in Annex 4. Due to the composition of workshop participants, some regions were more well-represented than others, i.e., a comprehensive analysis could be conducted for, e.g., the North Sea and the Northeast-Atlantic, but not, e.g., for the Black Sea. Despite this imbalance in the analysis, the overall feedback of the regional group analyses was positive, and regionally consistent fleet segments could be identified in all adequately represented regions. Nevertheless, more detailed, extensive, and inclusive work in separate regional sessions is necessary for an adequate regional analysis of fishing fleets.

In conclusion, we highlight the necessity of a standardized workflow to classify and name fleet segments. In addition, an artificial intelligence analysis approach developed in cooperation with the technical university of Kaiserlautern significantly improves the applicability of the FS-package. A corresponding publication is in preparation, and the AI approach needs to be included in the terms of reference of a future fleet segmentation workshop. All considerations and discussions of reforming the DCF fleet segmentation procedure need to take into account the criteria of the suitability of a fleet segmentation procedure. We identified a close connection of segments to fisheries, a homogenous cost structure of the resulting fleet segments, and the feasibility of the approach as the most important.

In the following, we present the Terms of Reference.

Terms of Reference

The overarching aim of our workshop was to test the amendments made to the newly developed approach to the segmentation of fishing fleets. This approach was created in a DFC pilot project of the Thünen Institute between September 2019 and December 2020 and transferred to an R package. After testing the approach in cooperation with various national partners, a workshop on the novel approach was held from March 29th to 31st, 2021. The results of this first workshop led to the following ToRs:

1. Harmonize the data preparation and develop a standardized protocol.
2. Clarify the application of the fleet segmentation approach and evaluate newly developed tools.
3. Establish regionally consistent fleet segments over multiple member states operating in the same fishing regions.

Background

Under the Data Collection Framework (DCF), economic data of European fishing fleets have to be provided by fleet segment. The current fleet segmentation scheme is based on the vessel length class and the main fishing gear, which are both technical parameters of the vessels. This segmentation method is well defined and easily applicable, but it does not adequately represent target fisheries. In addition, the length classes used for the classification often fail to group vessels of matching operational scales correctly. Vessels with similar technical parameters are often active in varying fisheries that differ in terms of catch composition, fishing activity, and cost structure. To improve reporting with respect to individual target fisheries, a transferable, systematic approach based on multivariate statistics methods was developed in the pilot project 'Fleet Segmentation' at the Thünen Institute for Sea Fisheries in Bremerhaven, Germany, from September 2019 to December 2020. The statistical framework was transferred to a user-friendly function package for the R statistics software, which was tested by multiple partners in the developmental process. From the 29th to the 31st of March 2021, the first workshop on the novel approach to fleet segmentation was held. The workshop was well attended by national fisheries experts, and based on the discussions, feedback, and suggestions gathered, the ToRs for the second fleet segmentation workshop were formulated.

Statistical Framework and Technical Amendments

A detailed description of the statistical background of the alternative fleet segmentation approach can be found in the report of the first fleet segmentation workshop. For an overview, please see figure 1. This section will focus on the amendments made to the approach prior to the workshop. First and foremost, the R package containing the novel approach was made available in a public programming repository, GitHub. This is in accordance with one major conclusion of the first workshop. Publishing the package makes it available to all potential users, simplifies the implementation of updates, and creates a citable source for reports and publications. On the technical side, firstly, the standard distance measure used in the clustering algorithm was changed from the Euclidean distance to a metric conversion of the Bray-Curtis dissimilarity. This distance measure is less sensitive to pairs of zeros in the data. It, therefore, creates more balanced clusters and helps reduce the abundance of single-ship clusters, which tend to be a problem, especially in large, heterogenic data sets.

As suggested in the first workshop, a function allowing standardized data preparation was created. The user no longer needs to create multiple specific data sets, only one comprehensive data set is needed from which all necessary subsets of data are created. This data preparation function and possibilities for improving it further will be discussed in the next section regarding ToR 1; additional newly developed or updated functions aiding the applicability of the package will be discussed in the section of ToR 2.

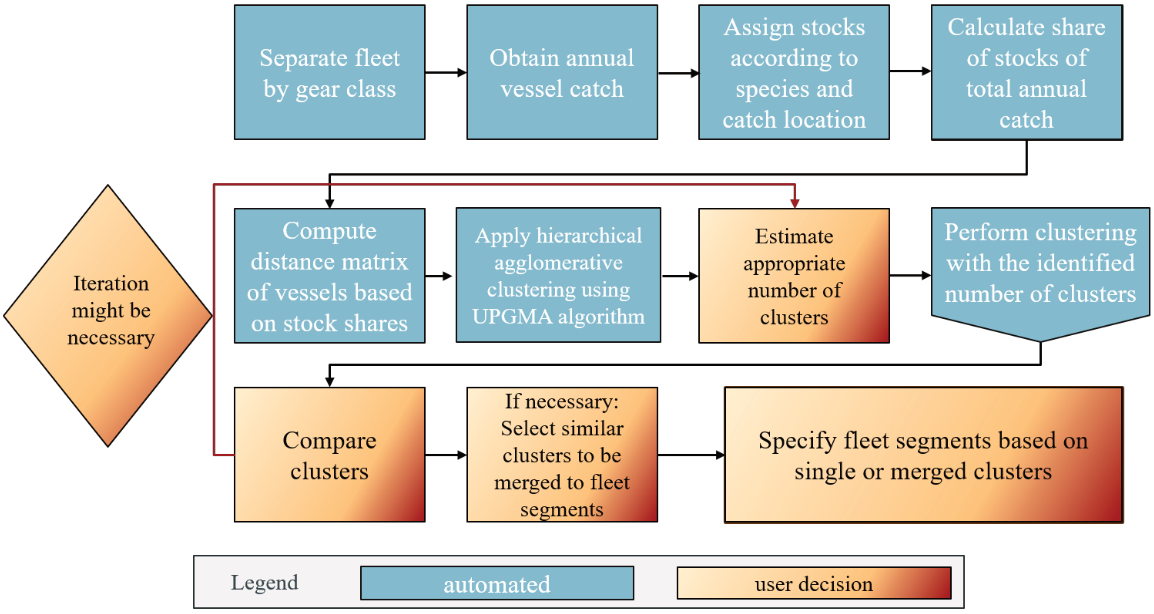


Figure 1: Updated stepwise flowchart of the newly developed fleet segmentation approach.

ToR 1. Harmonize the data preparation and develop a standardized protocol

The hurdles in installing the FleetSegmentation R package from a .zip file that were observed in the first fleet segmentation workshop illustrated the necessity of having the package publicly available. Hence, the FleetSegmentation R package was publically released on the GitHub repository in August 2021. As a consequence, updating the package has become considerably easier. However, installing a package from a non-quality checked repository like GitHub can be an obstacle for inexperienced users. Therefore, publishing the FleetSegmentation package in the official, quality-checked R repository CRAN (The **C**omprehensive **R** **A**rchive **N**etwork) is the next step in improving availability and ensuring a high code quality that meets scientific standards.

During the first fleet segmentation workshop, the ICES stock database used to create the distance matrices for the clustering algorithm was extended by stock lists of the ICES and the GFCM. This and the commonly agreed thresholds for automatic creation of species-area combinations considerably improved the clustering result. Nevertheless, especially the stock list of the Mediterranean Sea appears to be still incomplete and requires validation from experts of Mediterranean Fisheries Management.

The function to automatically assign stocks to the national data sets was nested in a new function that was created to harmonize the data preparation procedure. This new function requires only one comprehensive data frame, including all necessary fleet variables (vessel ID, main gear, species, area, catch weight), and computes all data frames for subsequent analysis per gear class. While this step reduces the preparatory work and the potential for errors, the pre-segmentation by main gear class requires further elaboration. In the second fleet segmentation workshop, most participants used the DCF main gear, as an alternative gear classification was not yet agreed upon. This DCF main gear is assigned based on gear use time, and any gear being used in more than 50% of the fishing time is considered the main gear. This not only conceals patterns of polyvalent gear use but also classifies some very heterogenic fishing gears, specifically demersal trawlers and demersal seiners (DTS). Based on our observations of fishing patterns and cost structure, we highly recommend separating demersal trawlers and demersal seiners in the DCF just like pelagic trawlers (TM) and pelagic purse seiners (PS). Regarding polyvalency, we emphasize the necessity to consider polyvalent gear classes like polyvalent active gears (MGP), a mix of active and passive gears (PMP). Thresholds for assigning polyvalent gear use based on time of use need to be agreed upon by the experts of the STECF.

Especially the large, heterogenic fleets of major fishing nations analyzed in the workshop elucidated the necessity of additional pre-segmentation measures. In these cases, analyzing vessel groups based only on the gear classes might lead to the grouping of vessels with very different technical parameters and cost structures, especially when these vessels are targeting widely distributed fish stocks. Participants tested various strategies of pre-separation during the workshop, e.g., size-based pre-separation of passive geares over and under 12m vessel length or area-based pre-separation between vessels that operate in non-ICES areas and those that do not. The different measures of pre-separation are currently collected from the participants, and surveys on which of those measures appear to be most promising will be held. It might be advisable that member states with large, diverse fleets introduce additional characteristics to further segment the fleet without initiating confidentiality issues.

Conversely, it was suggested to group different active gears or passive gears where necessary in a way that there is evidence for a comparable cost structure within that group. One idea was to test grouping all passive gears, all active gears, and all dredgers. However, this has to be further analyzed. Moreover, it was suggested to apply a pre-segmentation based on the technical equipment of the vessels, e.g., net drums, beams, net hauler, line hauler, etc. It is intended to include several pre-separation protocols in an updated version of the FleetSegmentation package.

ToR 2. Clarify the application of the fleet segmentation approach and evaluate newly developed tools

Both the newly developed and the updated functions were generally reviewed positively by the participants, especially by those who were familiar with the FleetSegmentation package from the first workshop. As suggested in the first workshop, the bar plot displaying the stock-based catch composition of the clusters (*clustering_stockshares_plot*) was added with the option to display the number of vessels in the corresponding clusters. This feature helps identify major fishing strategies and supports the user in decision-making on which clusters should be joined. The generic plots displaying cluster catch composition and technical variables (*clustering_stockshares_plot* and *clustering_plotgrid,* as well as all sub-plot functions included in the latter) were updated with an option to subset the plot to a selection of clusters to give additional advice for cluster joining and segment identification. This helps the user not only in getting an overview of the cluster features but also in saving results in a more structured and clear way. In addition to those updates, the target assemblage was introduced as a new feature for analysis. It is now possible to create a bar plot (*clustering_assemblageshares_plot*) and a multi-dimensional scaling (*clustering_assemblage_MDS*) of the target assemblages linked to the species caught by vessels. The assemblages help identify vessels with similar fishing strategies (e.g., vessels targeting small pelagic fish), especially if the user is not familiar with all species codes and stock units. Since ICES alone considers 269 stock units and fishing fleets generally target several hundred species, this is quite common. Including the assemblages in the segmentation procedure has the additional advantage that resulting fleet segments are more comparable to the métier-based biological analysis, where the target assemblage is included as a feature of the six-level métier definition. A complete list of the target assemblages can be found in Annex 3.

ToR 3. Establish regionally consistent fleet segments over multiple member states operating in the same fishing regions

The primary focus of the second fleet segmentation workshop was the regional group work to identify regionally consistent fishing strategies and, ultimately, also fleet segments between multiple fishing nations. Major fishing regions considered relevant for the participant nations were the Baltic Sea, the Black Sea, the Eastern Arctic, the Mediterranean, the Northeast Atlantic, the North Sea, and Other fishing regions (e.g., NAFO areas or Southern Atlantic). Due to the fact that not all European fishing nations were represented in the workshop, some regions were more diversely analyzed than others. All available regional plot sheets and participant case studies can be found in Annexes 5.1-5.4. The regional group work in the Mediterranean was mainly characterized by the comparison of the Maltese and Greek fishing fleets. For the Baltic Sea analysis, fleets from Germany, Finland, Sweden, Denmark, Lithuania, and Estonia were compared. In the North Sea region, analysis was carried out for fishing fleets from Germany, Denmark, the Netherlands, Sweden, and France. The Northeast Atlantic fishing fleets hailed from Portugal, Spain, France, Ireland, and Germany. In all fishing regions, consistent fleet segments could be, to some extent, identified among fishing nations. Of course, identifying targeted and distinct fishing patterns like, e.g., the Brown Shrimp fishery in the North Sea is less challenging than comparing diverse, heterogenic segments like mixed demersal fisheries. Yet, the regional group work highlighted that different fishing nations are in many cases engaged in comparable, regionally consistent fishing activities and that those activities can be identified using the novel fleet segmentation approach. We, therefore, emphasize the necessity for additional, detailed analysis of regional fishing patterns, ideally in regional subgroups. This analysis will help not only to identify major European fishing fleets but also to finalize the application protocol of our alternative approach to the segmentation of fishing fleets.

Overall workflow perspective

The alternative segmentation package is, in general, a necessary step but not entirely sufficient for a regular segmentation routine. The main purpose of the segmentation package is to provide support in defining and describing alternative, fisheries-related segments. This step is a manual one that requires expert knowledge. Expert knowledge is needed for pre-segmenting the fleet and for checks if the clusters generated are appropriate or can/need to be combined or split.

Thus far, participants were, in many cases, able to derive alternative fleet segments from the output of the segmentation package more or less intuitively. However, general principles for these steps still need to be elaborated. The characteristics of alternative segments should be clearly documented, e.g., in a table. The description should include

- Segment name
- Number of vessels
- Main gears
- Main stocks or stock assemblages targeted
- Vessel size range
- Engine power range
- DCF segments included
- Effort and/or landings measures were regarded as informative
- …

Once the fleet has been segmented for one year, this procedure should be repeated for 2-4 more years. The results should be checked for consistency. Where necessary, segment definitions and descriptions can be adjusted. For transparency reasons, it is advisable to cross-check the result of this step within regional working groups.

As soon as appropriate segments are characterized, the raw data will be used in an Artificial Intelligence (AI) package, which can identify patterns in vessel, logbook, and landings data and then assign vessels to alternative segments automatically that were defined in the aforementioned way. The AI package was developed by Verena Dully and Prof. Thorsten Stoeck of the Technical University of Kaiserslautern. The results of this tool are highly convincing, and a corresponding publication is in preparation. The AI tool will be part of a future workshop.

Further considerations – criteria for evaluating the suitability of segmentation principles

The existing segmentation approach, which is based on the dominant gear class and vessel length, has been in place for several years, with an introduction of two additional length class thresholds (10m and 18m) and a split of one gear class (PTS to TM and PS) in 2008 and a shift in length threshold from 10m to 8m in the Baltic in 2022. Length classes are easy to derive from the fleet register, and the main gear is usually determined using logbook information, where available. This approach has been used for many years and has proven to be feasible for all fleets. Still, the only application of fleet economics data with the same resolution as provided in the DCF has been the Annual Economic Report. Hence, the same resolution is used in the "Balance Report".

However, fisheries management often requires fleet economic data at a different resolution. One way to address this mismatch is to re-allocate economic data proportional to, e.g., effort or landings. In order to overcome this auxiliary step, the existing segmentation principle should be reviewed with a focus on linking segments more closely to stocks or target assemblages.

It appears desirable to aim at keeping the total number of segments similar to the DCF segmentation. If this is the case, the approach with a closer link to stock is to be preferred. If deemed necessary, the number of segments can be increased, though. Even in that case, the sampling effort might not need to be increased as alternative segments can be less heterogeneous in cost structure and thus require lower sampling rates per segment. In order to compare an alternative segmentation approach with the existing (DCF) one, several criteria should be taken into consideration. The following criteria should be taken into account but do not claim to be complete:

Connection to specific fisheries (high priority)

Any alternative approach should aim for a closer link of segments to stocks or groups of stocks. This aim can compete with the requirements of data confidentiality and potentially with cost-effectiveness. The closer the link to specific fisheries, the higher the number of segments and the smaller the number of vessels per segment. If segments get too small, data might become confidential.

Cost structure (high priority)

The need for an amendment of the current DCF fleet segmentation is based on the observation that DCF segments, in some cases, combine vessels with different cost structures. Hence, any alternative approach should result in segments with a more homogeneous cost structure. Like the DCF segmentation, any segmentation will combine vessels with a broad range of landings and effort. These characteristics are therefore not in the focus of an analysis of homogeneity. Likewise, cost items should not be analyzed based on absolute values but rather on indicators or proxies. E.g., if a group of shrimp beam trawlers has a broad range of fuel cost per year, this will be mainly due to a broad range of fishing days and probably engine power. Therefore, the indicator fuel cost/kwday might be appropriate to describe homogeneity.

Feasibility (high priority)

The segmentation procedure has to be clear, doable without excessive extra burden, and repeatable.

Compatibility (lower priority)

It is desirable if the segmentation is compatible with any existing time series. However, a time series is only as good as the information that comes with it. The link between the current DCF segmentation and particular fisheries can be quite loose. In this case, the time series is not very informative in terms of fisheries management. If the link between the DCF segment and certain fisheries is close, then the link to the alternative segmentation is close as well, and the time series is more or less stable. It has to be borne in mind that the recent EUMAP legislation has introduced a length class threshold at 8m instead of 10m for the Baltic Sea, thus introducing a break in time series anyway.

Annex

1. List of participants
2. Workshop agenda
3. List of target assemblages
4. List of analyzed regions
5. Case studies
    5.1. Finland
    5.2. Greece
    5.3. Romania
    5.4. France
    5.5. Germany
    5.6. Portugal

Annex 1. List of participants

Table 1: Particpants of the second fleet segmentation workshop in alphabetic order of first name.

Name	Institution	Nation
Adelbert De Clercq	Instituut voor Landbouw-, Visserij- en Voedingsonderzoek, Flanders Research Institute for Agriculture, Fisheries and Food, Visserij & Aquacultuur Biologie	Belgium
Alvaro Minguez Velasco	S.G. de Sostenibilidad Económica y Asuntos Sociales, D.G de Ordenación Pesquera y Acuicultura, Ministerio de Agricultura, Pesca y Alimentación	Spain
Ana Cláudia Fernandes	Portuguese Institute for the Sea and Atmosphere - IPMA Division of Modelling and Management of Fisheries Resources	Portugal
Andrew Sciberras	Department of Fisheries and Aquaculture, Malta	Malta
Antonios Papoutsis	Fisheries Research Institute (FRI)	Greece
Antonios Stamoulis	European Comission	EU
Antti Sykkö	Natural Resources Institute Finland (LUKE)	Finland
Brian Paul Zammit	Department of Fisheries and Aquaculture, Malta	Malta
Christina Stefanou	Department of Fisheries and Marine Research	Cyprus
Grigorias Daniel	National Institute for Marine Research and Development “Grigore Antipa”	Romania
Hanna Swahnberg	The Swedish Agency for Marine and Water Management	Sweden
Hans van Oostenbrugge	Wageningen University & Research	Netherlands
Irene Tzouramani	Agricultural Economics Research Institute, Hellenic Agricultural Organization-DEMETER	Greece
Irina Davidjuka	Institut of Food Safety, Animal Health and Environment	Latvia

Irina Jakovleva	Žuvininkystės įsipareigojimų vykdymo skyriaus	Lithuania
Jamal Roskam	Wageningen University & Research	Netherlands
Janek Lees	Estonian Marine Institute, University of Tartu	Estonia
Joonas Valve	Natural Resources Institute Finland (LUKE)	Finland
Josefine Egekvist	DTU AQUA National Institute of Aquatic Resources Section for Monitoring and Data	Denmark
Julie Kellner	International Council for the Exploration of the Sea (ICES)	ICES
Jurgen Mifsud	Department of Fisheries and Aquaculture, Malta	Malta
Katell Hamon	Wageningen University & Research	Netherlands
Lazaros Tsiridis	Fisheries Research Institute	Greece
Maria Valiante Viana	S.G. de Sostenibilidad Económica y Asuntos Sociales, D.G de Ordenación Pesquera y Acuicultura, Ministerio de Agricultura, Pesca y Alimentación	Spain
Marie-Dominique Minne	Ministère De L'agriculture Et De L'alimentation	France
Miriam Gambin	Department of Fisheries and Aquaculture, Malta	Malta
Monika Sterczewska	European Comission	EU
Myrto Iouannou	Department of Fisheries and Marine Research	Cyprus
Olivier Guyader	Ifremer (Institut français de recherche pour l'exploitation de la mer)	France
Paoloa Arcadia	Fisheries and Aquaculture Economic Research (NISEA)	Italy
Paun Catalin	National Institute for Marine Research and Development "Grigore Antipa"	Romania
Richard Meitern	Estonian Marine Institute, University of Tartu	Estonia
Rosaria Sabatella	Fisheries and Aquaculture Economic Research (NISEA)	Italy
Sébastien Demaneche	Ifremer (Institut français de recherche pour l'exploitation de la mer)	France
Stamatis Mantziaris	Agricultural Economics Research Institute	Greece
Stefanos Kavadas	Hellenic Centre for Marine Research (HCMR)	Greece
Suzana Faria Cano	Direcção Geral de Recursos Naturais, Segurança e Serviços Marítimos	Portugal

Second workshop on a novel approach to the segmentation of fishing fleets

THÜNEN

Agenda*

** all times are in **CEST** and might be changed in the progression of the workshop*

Monday

10:00	Welcome and housekeeping
10:30	Brief introduction to the current fleet segmentation system and its issues (Jörg Berkenhagen)
11:00	Summary of the last fleet segmentation workshop and ammendments made to the fleet segmentation approach (Erik Sulanke)
11:30	Data preparation seminar and getting started
13:00	Lunch break
14:00	Regional subgroups – forming and running national analyses, if missing
17:00	End of the first day

Tuesday

09:30	Regional subgroup work
13:00	Lunch break
13:00	Regional subgroup work cont.
16:30	Impressions from data work
17:00	End of the second day

Wednesday

10:00	Summary and accomplishments of data work day (Erik Sulanke)
10:30	Regional subgroups present their results
13:00	Lunch break
14:00	Experience presentation, pt. II
15:00	Key results and implications for the future
16:00	End of the workshop

Annex 3. Target assemblages

Table 2: Target assemblages used in the FS-package. Three-letter code and example species given.

Target assemblage - code	Target assemblage	Species (example.)
LPF	Large pelagic fish	Albacore *(Thunnus alalunga)*
DEF	Demersal fish	Cod *(Gadus morhua)*
SPF	Small pelagic fish	Herring *(Clupea harengus)*
DWS	Deep-water species	Roundnose grenadier *(Coryphaenoides rupestris)*
FWS	Freshwater species	Northern pike *(Esox lucius)*
MOL	Molluscs	Blue mussel *(Mytilus edulis)*
CRU	Crustaceans	Brown crab *(Cancer pagurus)*
CAT	Catadromous species	European eel *(Anguilla Anguilla)*
ANA	Anadromous species	Salmon *(Salmo salar)*
CEP	Cephalopods	European squid (*Loligo vulgaris)*
GAS	Gastropods	Periwinkles (*Littorina* spp.)

Annex 4. Analyzed regions

Region	Extent / Remarks
Baltic Sea	Including Kattegat (ICES area 27.3.a.20)
Black Sea	-
Eastern Arctic	-
Mediterranean	-
Northeast Atlantic	-
North Sea	Including Skagerrak (ICES area 27.3.a.21)
Other regions	e.g., overseas fishing regions like FAO area 34 (Eastern Central Atlantic)

Annex 5.1. Case study - Finland

Alternative Approach to Fleet Segmentation

Workshop report - Finland

Antti Sykkö, Natural Resources Institute Finland (antti.sykko@luke.fi (mailto:antti.sykko@luke.fi)), Joonas Valve, Natural Resources Institute Finland (joonas.valve@luke.fi (mailto:joonas.valve@luke.fi))

2022-04-20

Experiments with The FIN 2020 catch data

The following clustering analysis was carried out via `fleetSegmentation()` -package and with the 2020 catch data by Finnish trawlers ($n = 43$) and vessels which used passive gears ($n = 1245$). The average vessel lengths in catch data were 25 and 6 meters for trawlers and vessels using passive gears, respectively (Figure 1).

The pelagic trawlers gear class TM consist of trawlers, which mostly target Baltic herring (Clupea harengus membras) in the Northern Baltic Sea, with additional target species being sprat (Sprattus sprattus), smelt (Osmerus eperlanus) and vendace (Coregonus albula). Other species are caught in minor amounts and, such as Herring, are mostly sold as fish fodder. All the vessels fish in the Baltic Sea, (27.d), with majority of the fishing occurring in sub-areas 27.d.32,27.d.31, 27.d.29 and 27.d.30.

The passive gear class PG involves small coastal vessels targeting their fishing to, e.g., Baltic herring (Clupea harengus), atlantic salmon (Salmo salar), pikeperch (Stizostedion lucioperca) and European perch (Perca fluviatilis), which are economically the most important target species for the coastal fisheries. The vessels using passive gears fish in the Northern Baltic Sea, mostly in Gulf of Bothnia and Gulf of Finland at the coast of Finland.

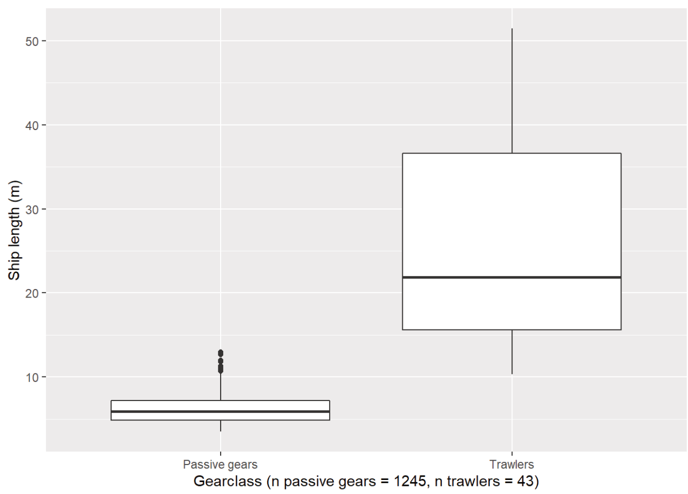


Figure 1: The length of ships in Finnish fleet in 2020 in two gear classes - Passive Gears (PG) and Pelagic Trawlers (TM).

Pelagic trawlers - TM

We utilized the `segmentation_datapreparation()` -function for trawler data preprocessing. The amount of catch in kilograms was chosen as the catch measurement. We exploited `numberclust_table()` and `numberclust_plot()` to receive statistics related to the optimal k, and visualized the suggested clusters via `numberclust_dendrogram()` and `numberclust_clustree()` .

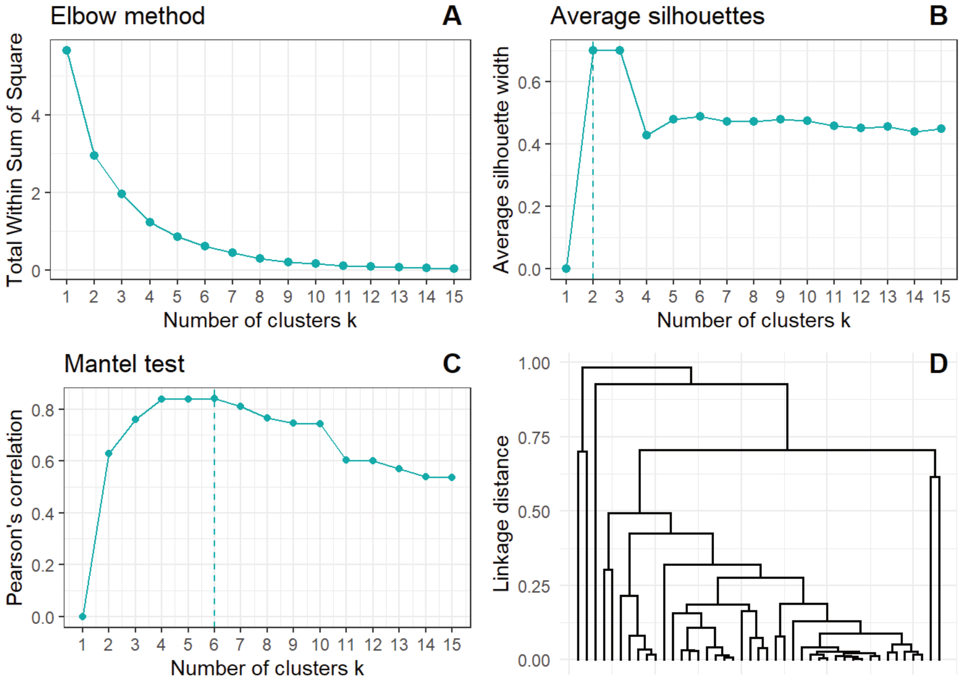


Figure 2: Visual diagnostics for optimal number of clusters in TM catch data.

The initial analysis were performed in order to find the optimal number of clusters k. Based on statistics (Avg. Silhouette score $= 0.699$ for $k = 2$, Mantel score $= 0.84$ for $k = 6$), the optimal choice would have been either $k = 2$ or $k = 6$. The visual diagnostics (Figure 2) show that elbow suggests to place k somewhere between four and eight, silhouette proposes two or three while Mantel encourages to try between four and seven. However, since we only had 43 vessels and we knew beforehand that the vast majority of herring would be fished, we wanted to examine the clustering with a sufficiently large k. Thus, we chose $k = 8$ for further analysis.

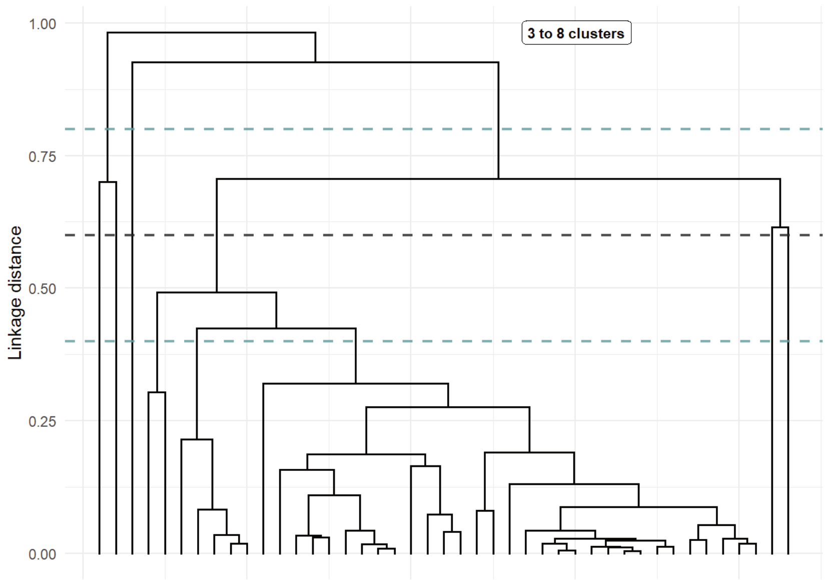


Figure 3: Clustering dendogram with range-method for pelagic trawlers

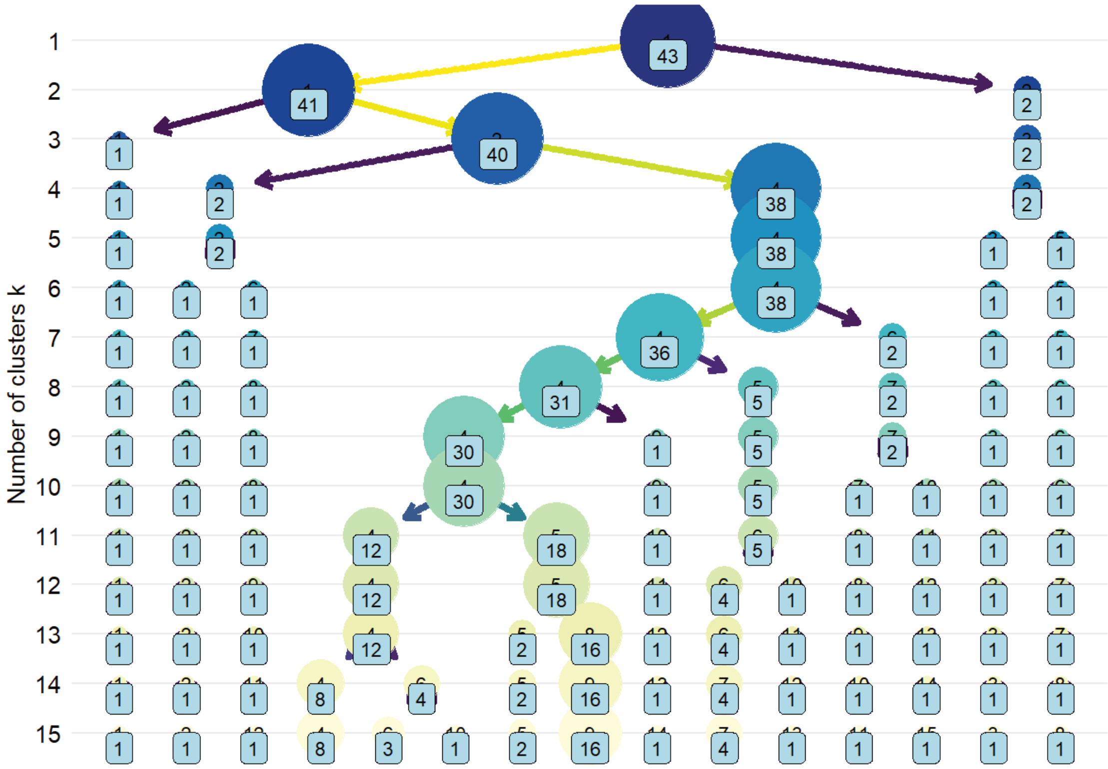


Figure 4: Clustering tree for TM trawler data

TM with $k = 8$

The clustering analysis was performed with number of clusters $k = 8$. We deployed `segmentation_clustering()` with thee catch data of trawlers and examined the visualizations provided by `clustering_stockshares_plot()`, `clustering_assemblageshares_plot()`, `cluster_assemblages_MDS()` and `clustering_plotgrid()`. The results were in line with our initial thoughts as $31/43 \approx 70\%$ of the vessels were placed into one, herring-weighted, cluster.

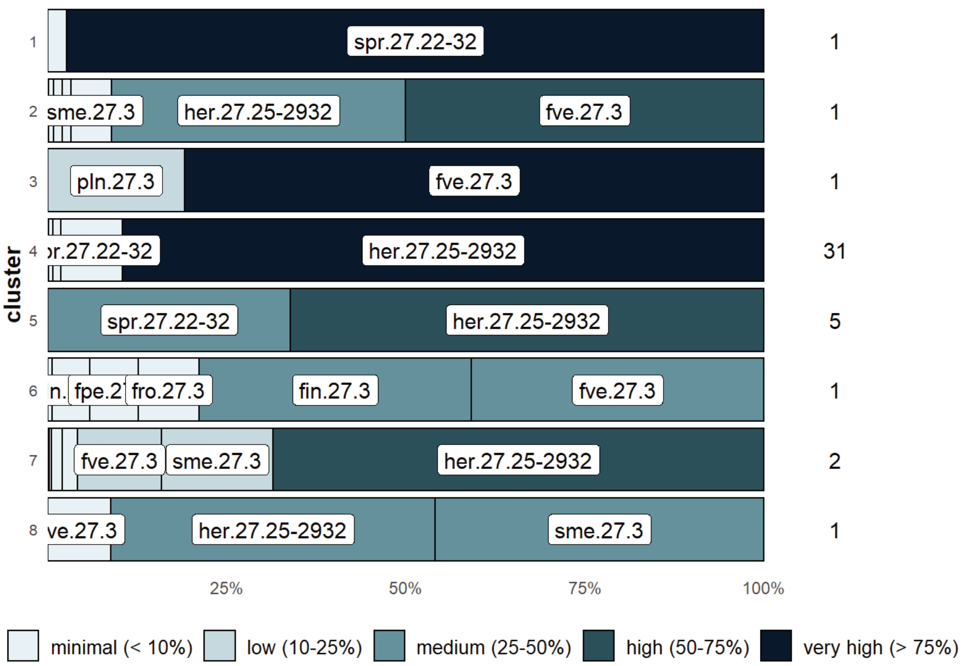


Figure 5: Stockshare plot of TM catch data (n=43)

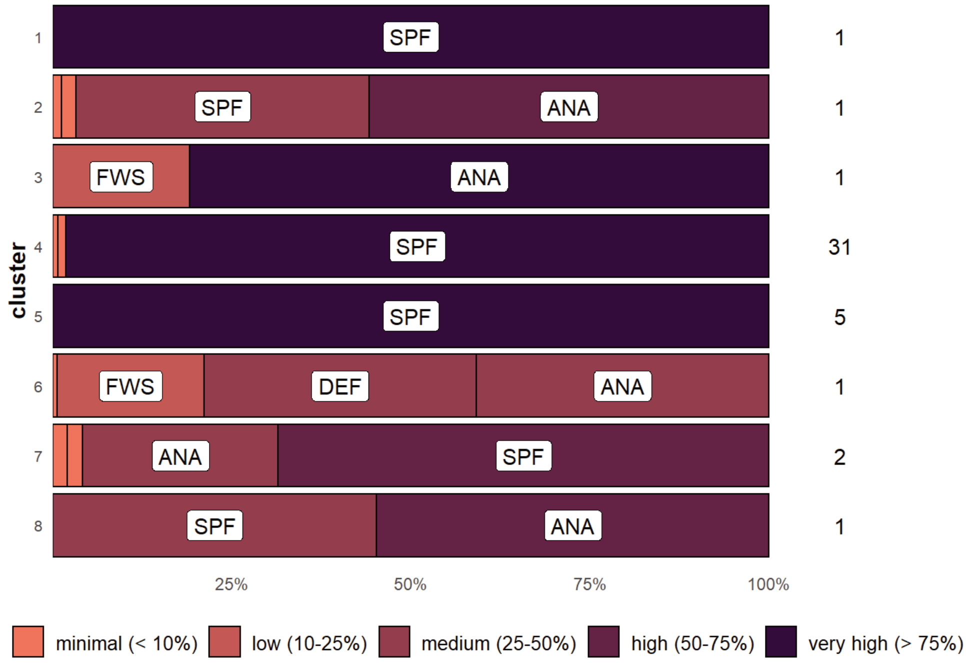


Figure 6: Assemblageshares plot of TM catch data (n=43)

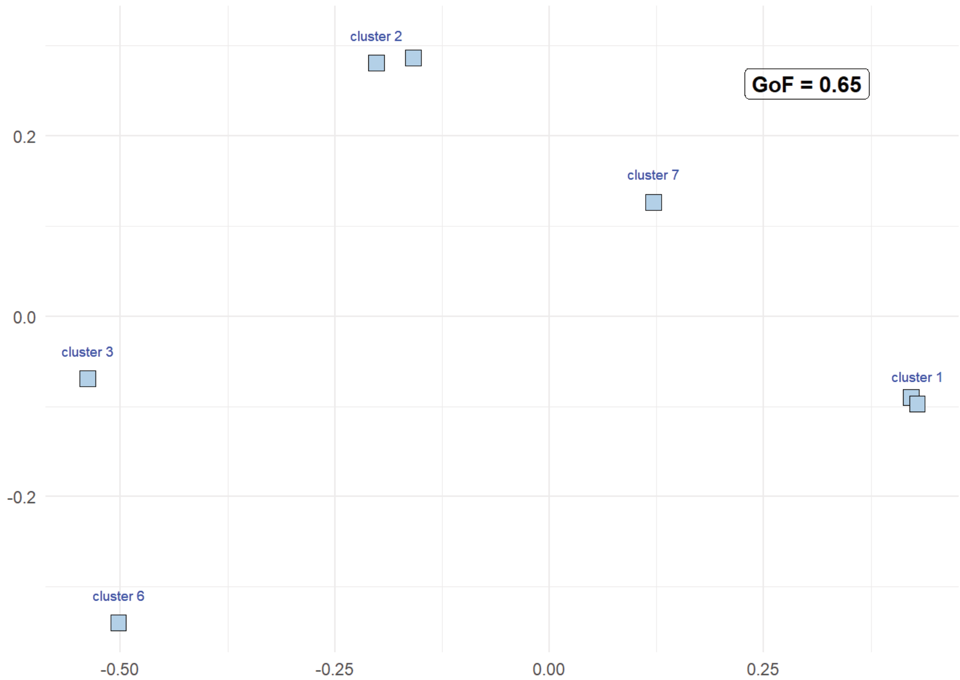


Figure 7: Assemblages MD5 plot of TM catch data (n=43)

We looked the properties of the vessels in each cluster (figure 8) before nailing down the final clustering. While the clusters 1,2,3,6,7 and 8 are not similar in terms of catch composition (figure 7), they are very similar in terms of vessel diagnostics. We can see that with respect to vessel length, annual catch per ship and annual catch per cluster, our choice to combine the previously mentioned group of clusters is well justified.

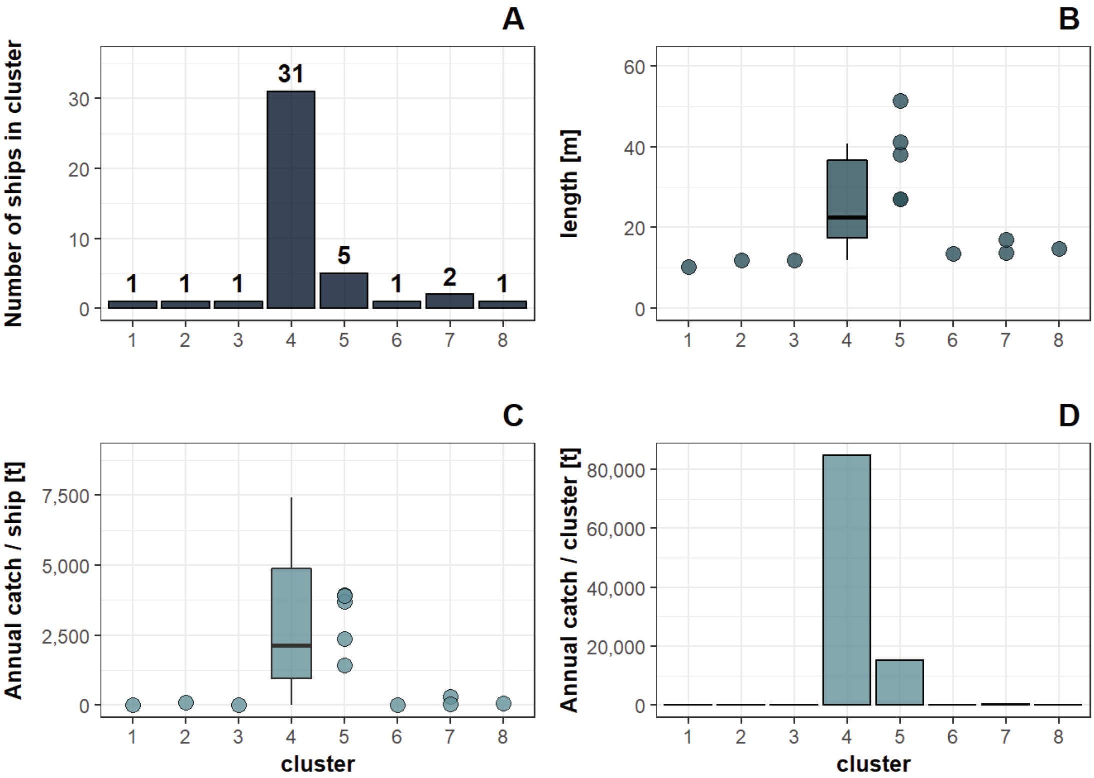


Figure 8: Diagnostics and statistics of TM catch data vessels (n=43)

TM Final clustering

Final clustering for pelagic trawlers was formed by combining the initial clusters 1,2,3,6 and 7 together and leaving clusters 4 and 5 as they were suggested by the algorithm. We identified three different clusters.

- **Herring cluster** having 31 trawlers of length between 20-40 meters and targeting their fishing heavily on herring (initial cluster 4, figure 9).
- **Herring & Sprat cluster**, which contains 5 relatively large vessels targeting on herring and sprat (initial cluster 5, figure 10).
- **Mixed cluster** holding 6 vessels with length less than 20 meter. These vessels fish a quite of a number of different species, none of which seems to surpass the others (initial clusters 1,2,3,6,7 , figure 11).

It is interesting to notice that these three final clusters are in practice identical when comparing the final clusters defined in these analysis to the final clustering determined in previous workshop (2021) for $k = 10$.

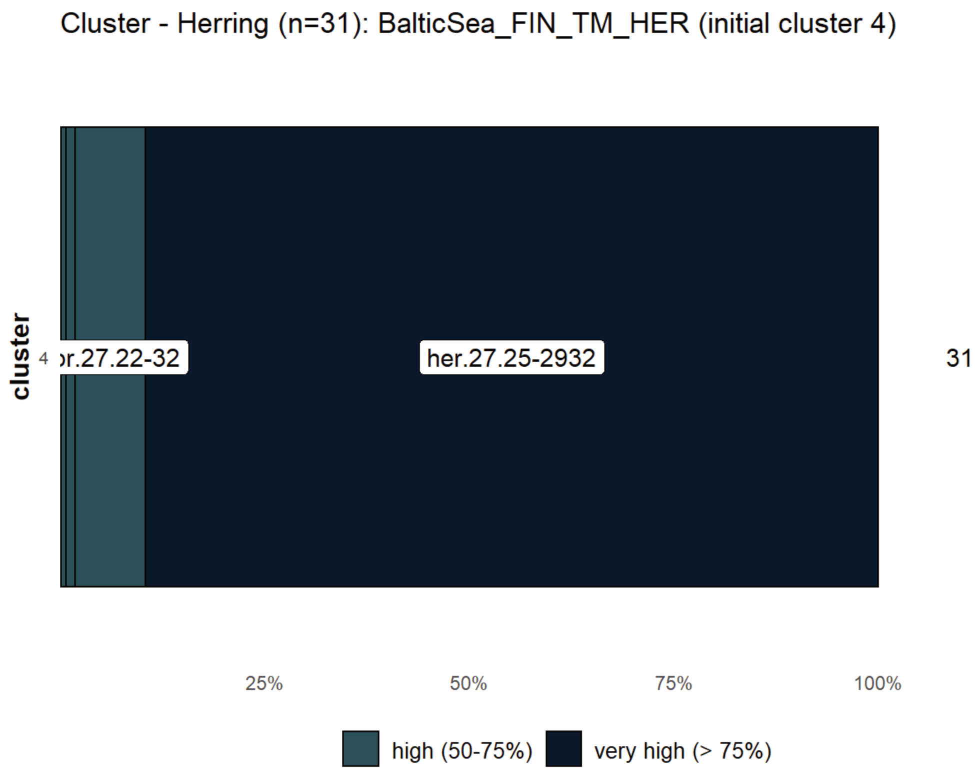


Figure 9: TM Cluster - Herring

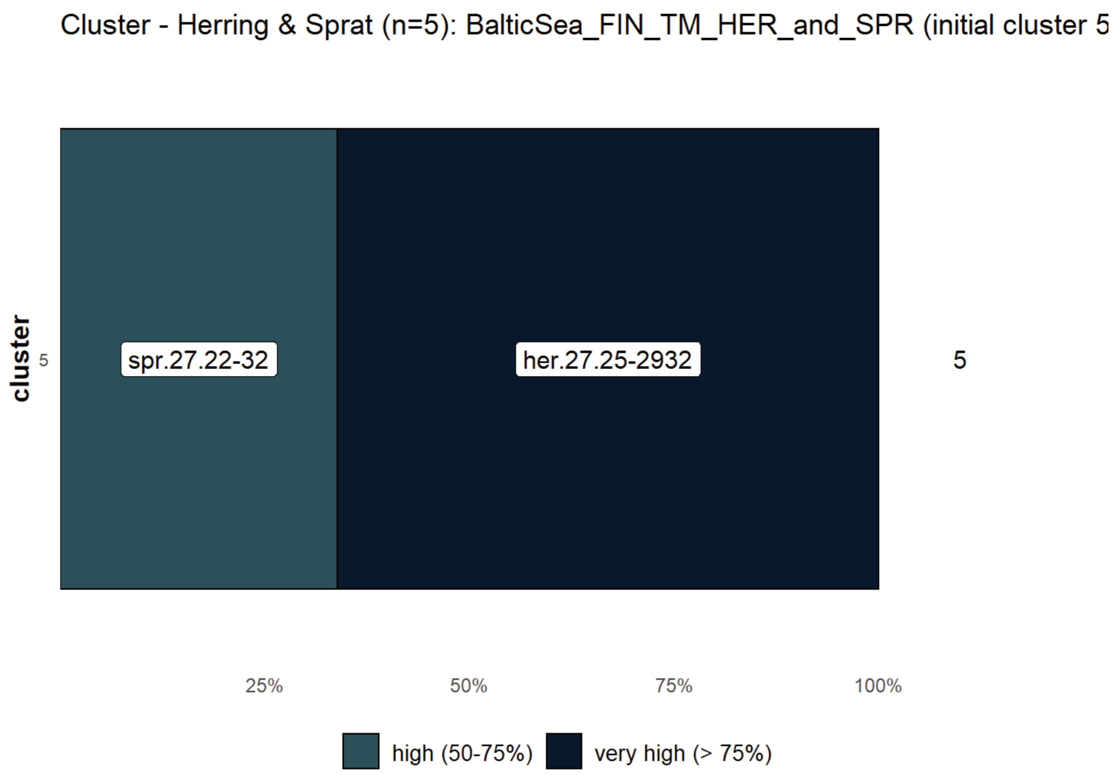


Figure 10: TM Cluster - Herring + sprat

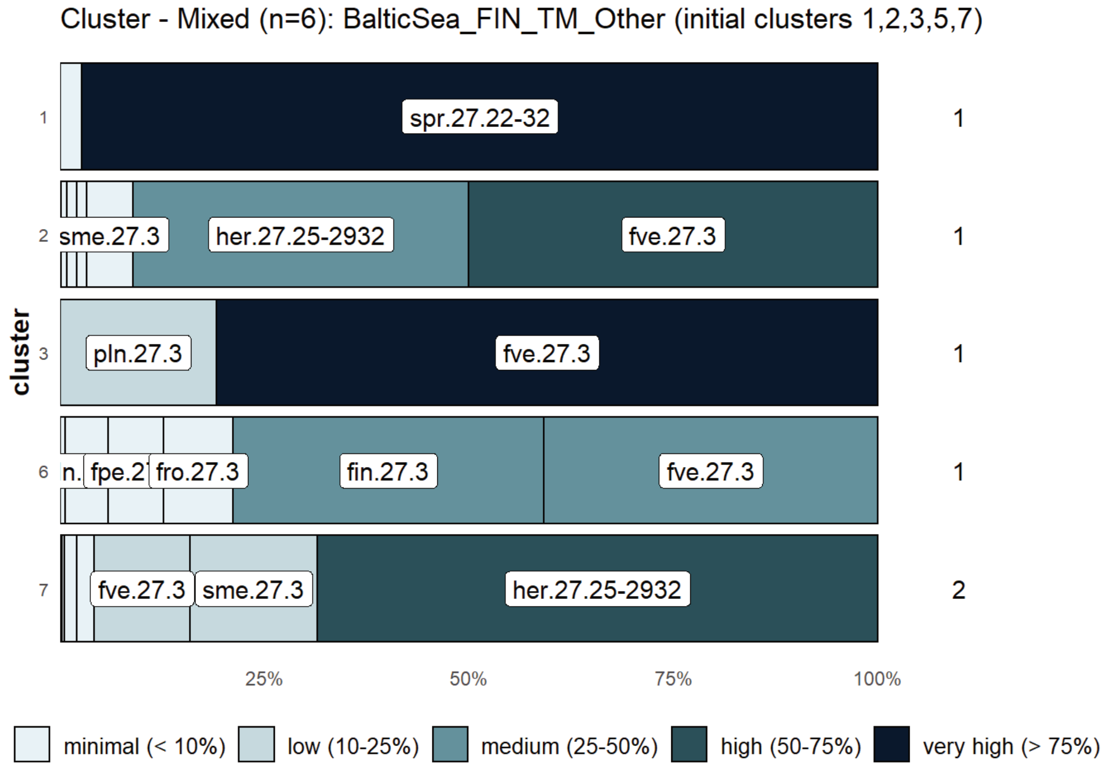


Figure 11: TM Cluster - Mixed

Passive Gears - PG

The passive gear class analysis was implemented by first running the `segmentation_datapreparation()` -function with PG data ($n = 1245$) followed by the actual clustering analysis. We used the amount of catch in kilogram as a cath measurement. The diagnostics and statistics were then examined in order to choose k for further analysis. Based on elbow and silhouette means (score ≈ 0.388 for $k = 6$ and $k = 7$), we can see that, overall, the optimal number of clusters seems to be either $k = 6$ or $k = 7$ (figure 1). However, Mantel test makes an exception suggesting to try clustering with $k = 15$ clusters (highest score 0.749 for $k = 15$). Moreover, we also deployed `numberclust_clustree()` -function to create a clustering tree to provide additional visualizations.

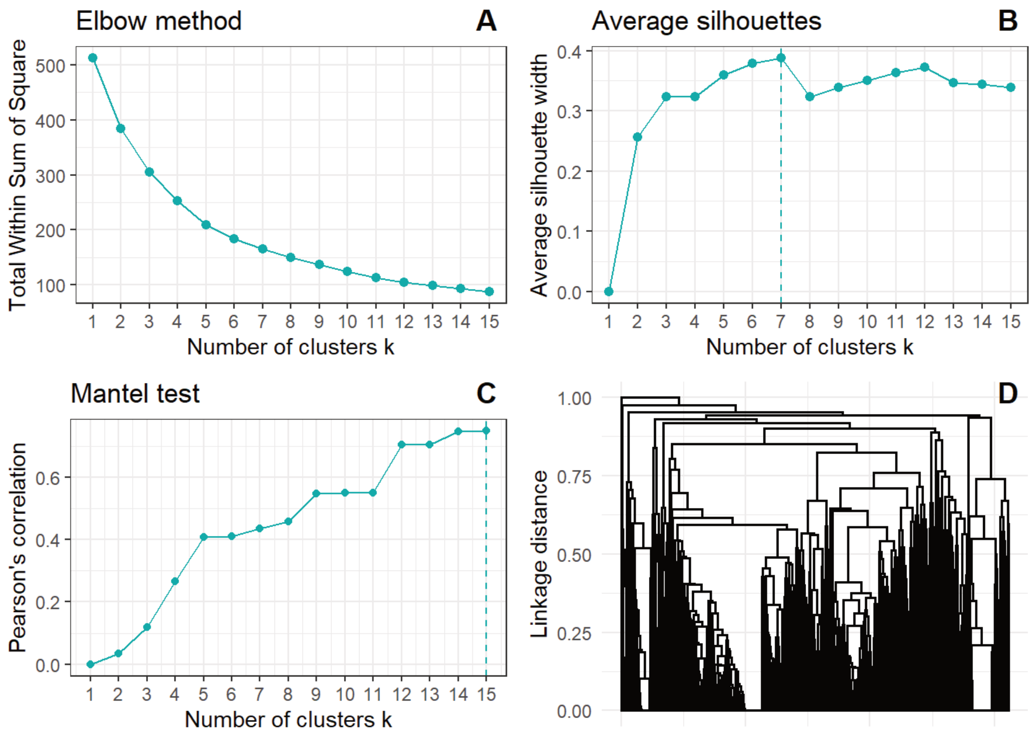


Figure 12: Visualization - Estimated number of clusters for passive gear (PG) class.

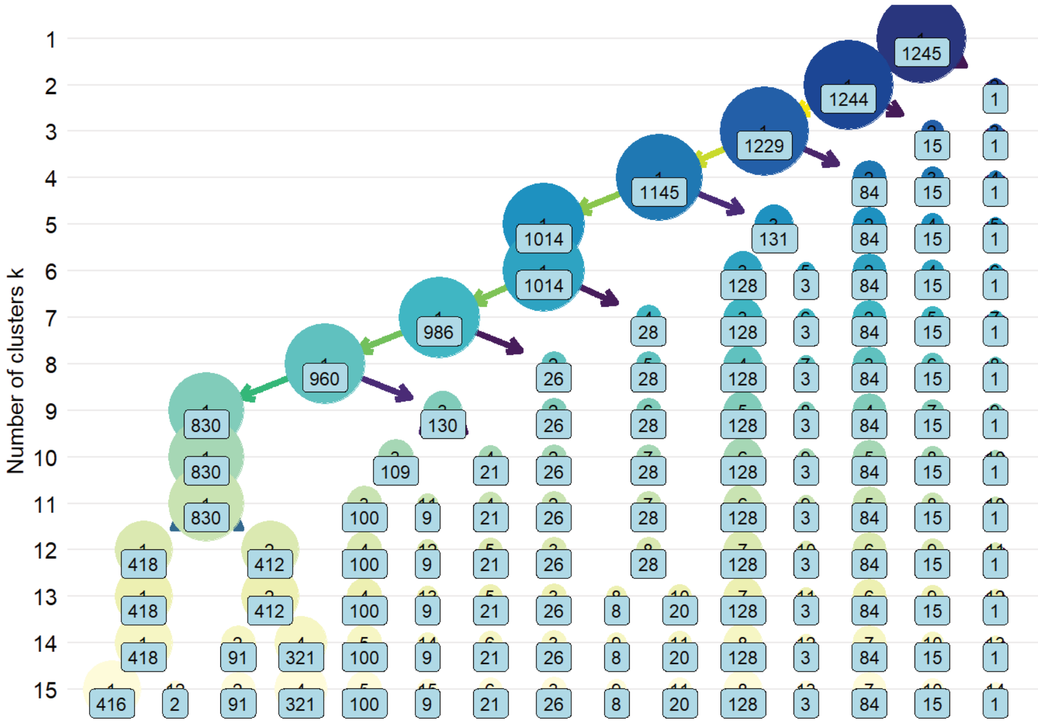


Figure 13: Clustering three with PG data

PG with k = 7

We chose k = 7 clusters for further analysis. We utilized passive gear catch data and performed clustering with `segmentation_clustering()` -function. We started our experiments by looking the stock shares and assemblage shares over the seven clusters. The initial results were found quite promising, as we can see (figure 14) that there is relatively clear distinctive specie representation among the clusters.

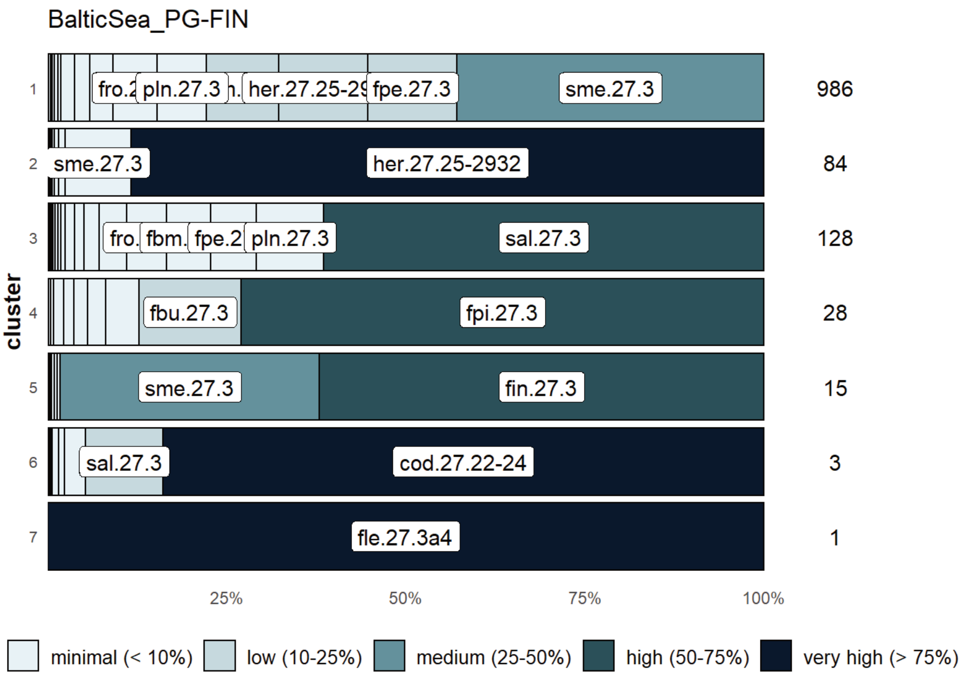


Figure 14: Stockshares of Passive Gear data with n=1245 and k=7

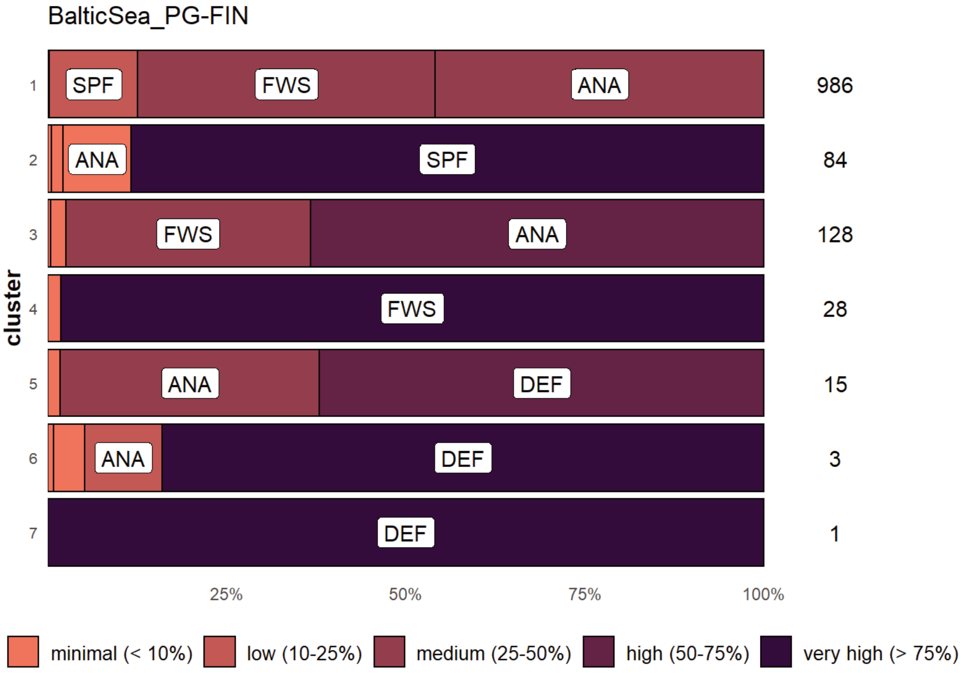


Figure 15: Assemblage shares plot of PG data

We utilized `cluster_assemblages_MDS()` and `clustering_MDS()` functions to gain more information of the cluster candidates. From the left hand side (of figure 16) can be seen that cluster 2 stands out from the crowd, but clusters 5,6,7 and 1,3,4 appear to form relatively similar groups. The same phenomena can be seen from the right hand side (figure 16)). That is, it looks like there is a major overlap among vessels. However, if we look at the three dimensional plot (figure 17) and rotate it appropriately, we can notice that the ostensible overlap is just a result of visualizing the ships in two dimensional space.

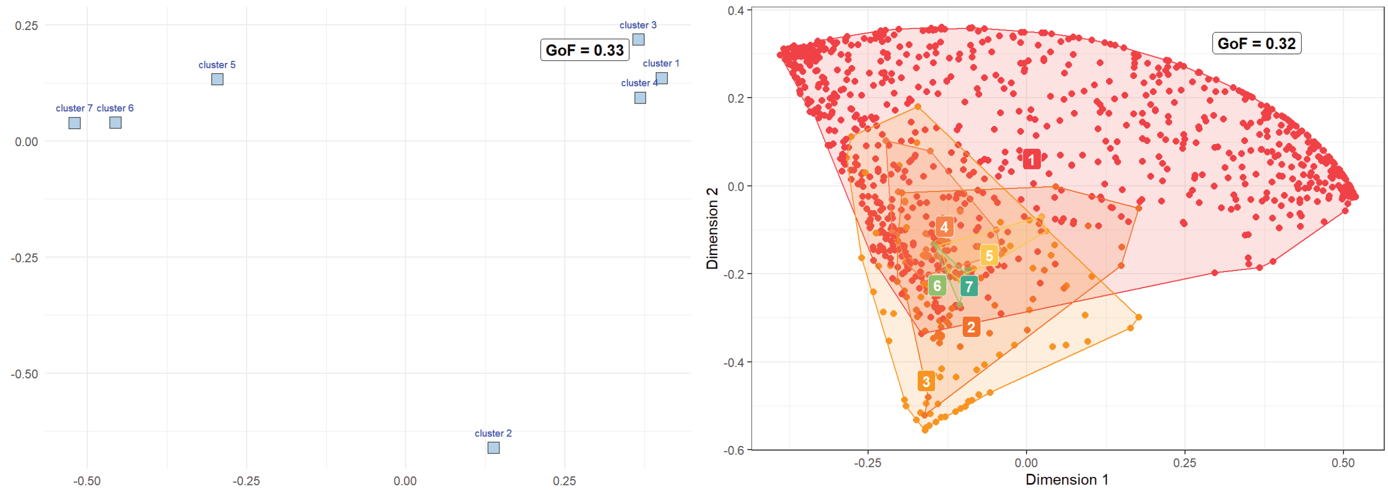


Figure 16: Cluster assemblages and clustering MD5 of PG gear class - Multidimensional scaling to two dimensions.

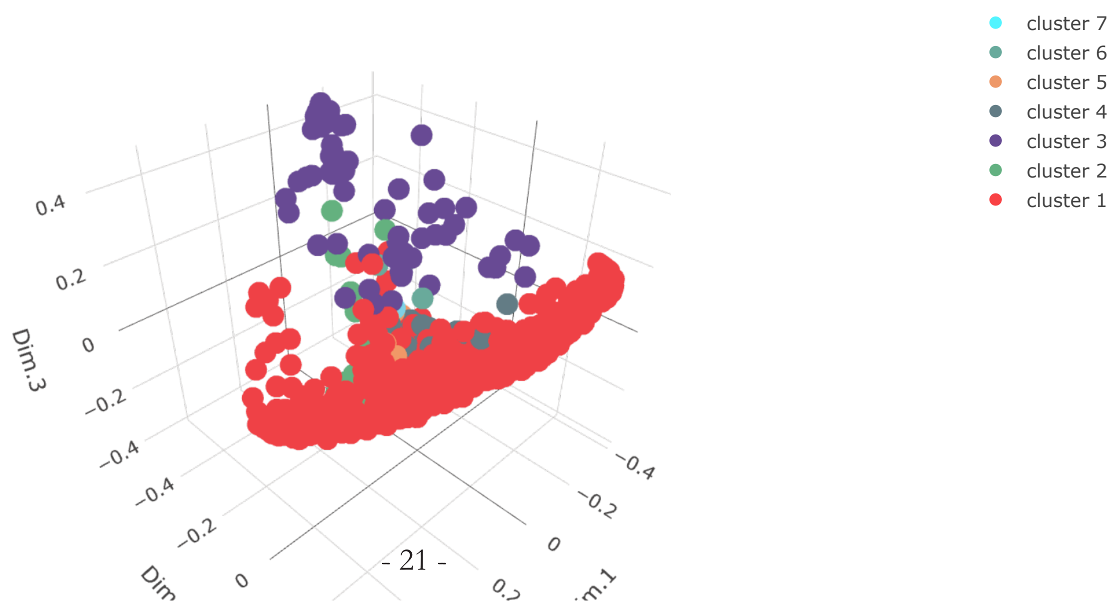


Figure 17: Clustering MD5 with PG catch data - Multidimensional scaling to three dimensional space

Then we looked at the properties of the clusters in terms of the vessel length, cluster size and annual catches (figure 18). From the annual catch point of view, it is interesting to see that the biggest share of catch is landed by cluster 2 (figure 18) and the vessels in this cluster are the longest. This notable amount of catch is natural, as the vessels in cluster 2 are targeting herring, and herring forms undoubtedly the largest share of annual landings among Finnish fisheries.

The vessels in clusters 1,3,4 and 5 have approximately equal lengths, but the annual catch in cluster 1 is considerably larger. However, this is clearly explained by the number of vessels in cluster 1 versus other clusters. Furthermore, it is interesting to see that there are some outliers in clusters 1 and 3 in terms of ship length. Thus, we can notice that, from the similarity of fishing activities point of view, the vessel length is not a one classification feature above the others yet it is an important factor in many extent.

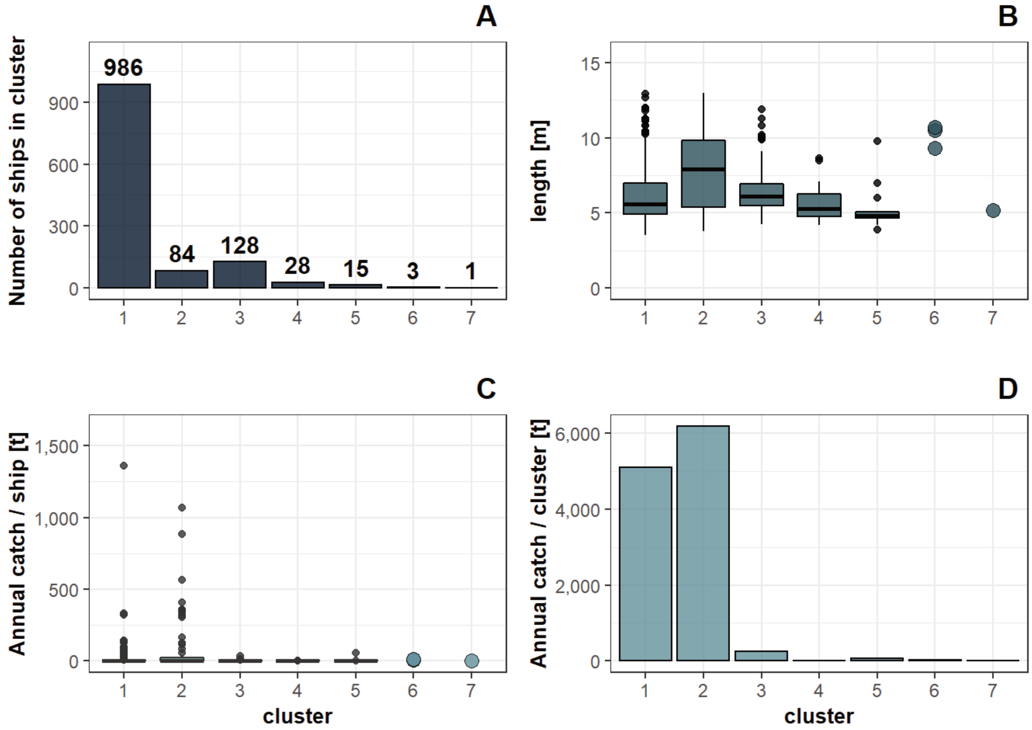


Figure 18: PG Vessel properties over different clusters

PG Final clustering

We introduce the final PG clustering in this section. Based on the previous diagnostics and analysis, the seven clusters provided by `fleetSegmentation` were categorized as follows:

- **Herring cluster** holding in total of 84 medium size coastal vessels (initial cluster 2, figure 19), which use passive gears and target their fishing heavily on herring.
- **Salmon cluster** having 128 fishing small coast ships (initial cluster 3, figure 20) targeting mostly salmon but other species, e.g, european perch, beam and roach, to some extent as well.
- **Northern pike cluster** consists of 28 small coastal vessels (initial cluster 4, figure 21) targeting their fishing primarily to northern pike and secondary to burbot.
- **Smelt & Mixed cluster** is by far the largest cluster with 986 vessels (initial cluster 1, figure 22) of which most are small coast ships. These vessels catch mostly smelt, but also fish other species (bream, perch, roach) on side.
- **Mixed cluster** involves 19 vessels (initial clusters 5,6,7, figure 23) of which 15 of them are targeting mostly unknown species and smelt. Three vessels are catching cod heavily. Furthermore, this cluster has one ship fishing solely flounder.

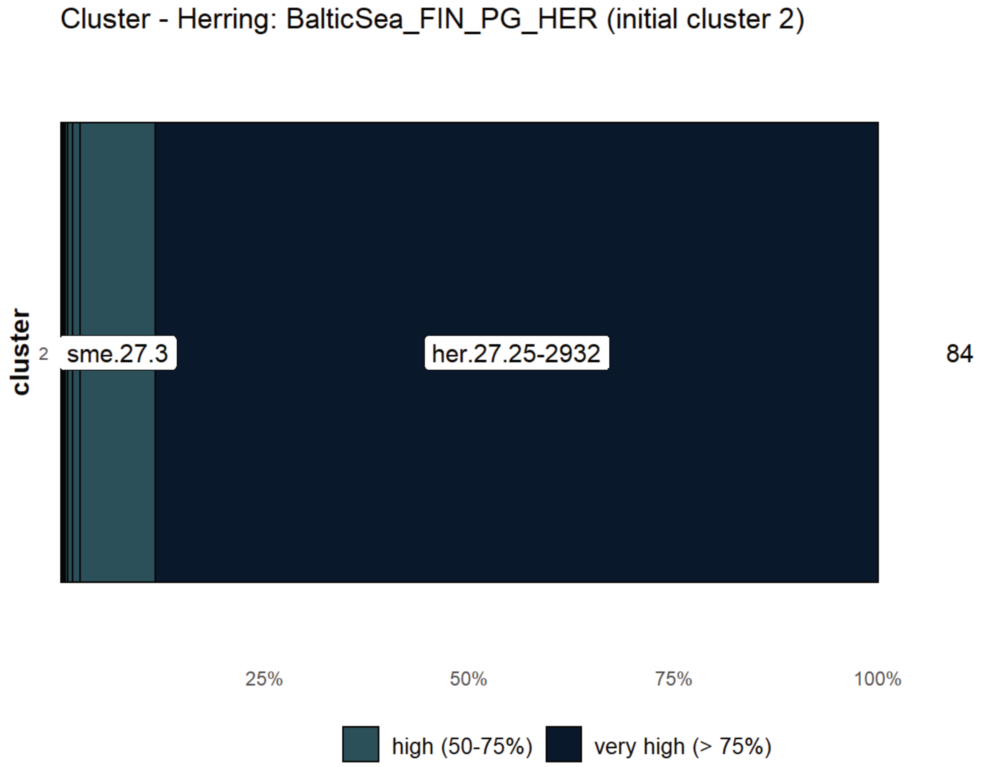


Figure 19: PG Cluster - Herring

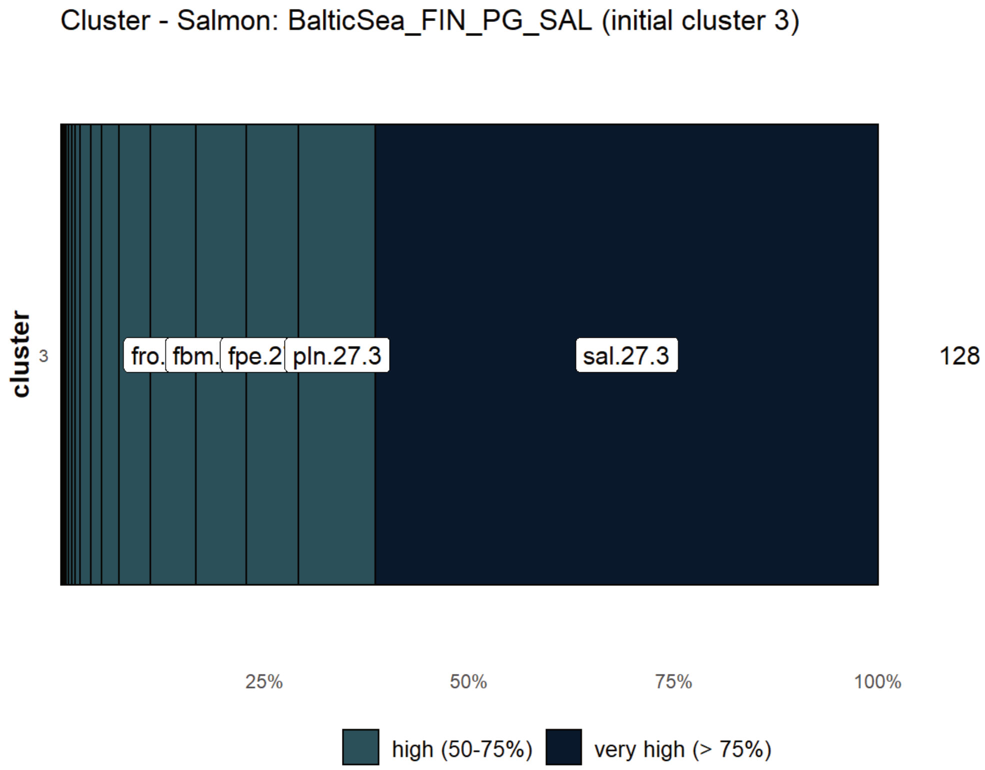


Figure 20: PG Cluster - Salmon

Cluster - Northern pike: BalticSea_FIN_PG_FPI (initial cluster 4)

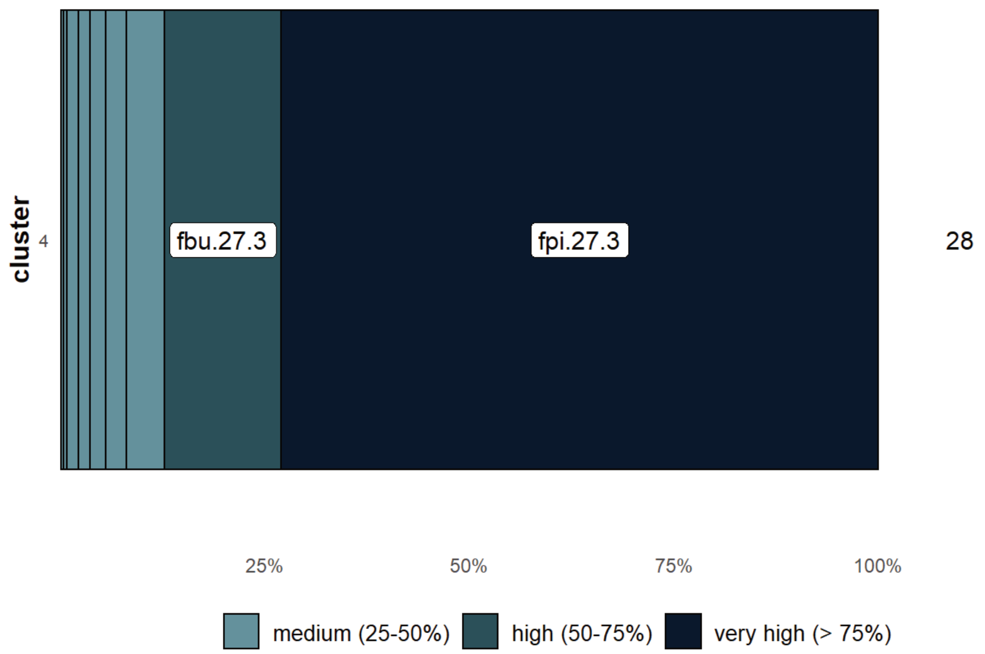


Figure 21: PG Cluster - Northern pike

Cluster - Smelt & Mixed : BalticSea_FIN_PG_SME_and_Mixed (initial cluster 1)

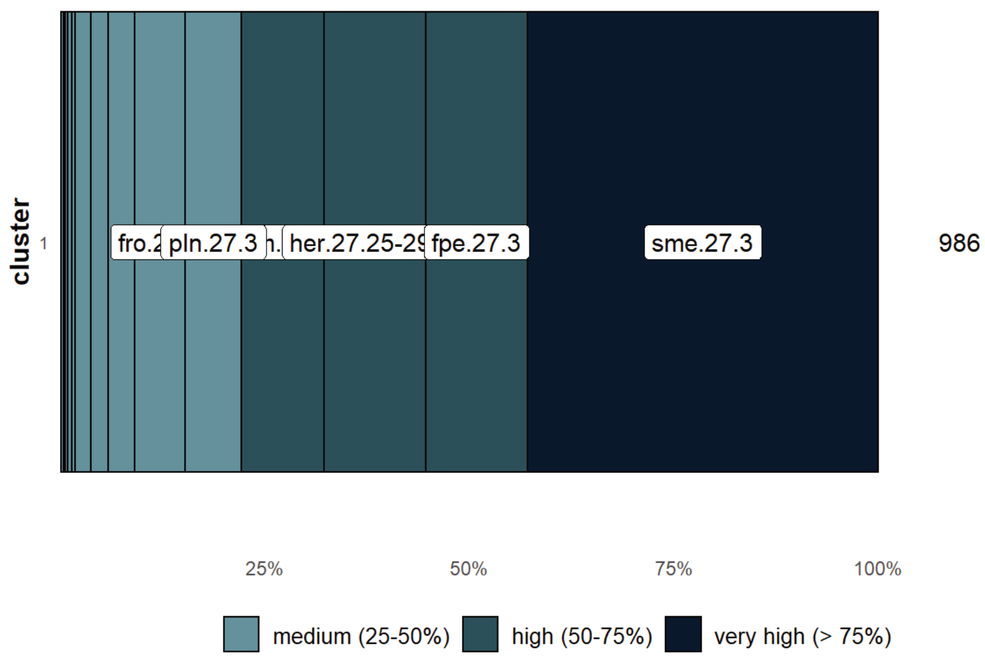


Figure 22: PG Cluster - Smelt & Mixed

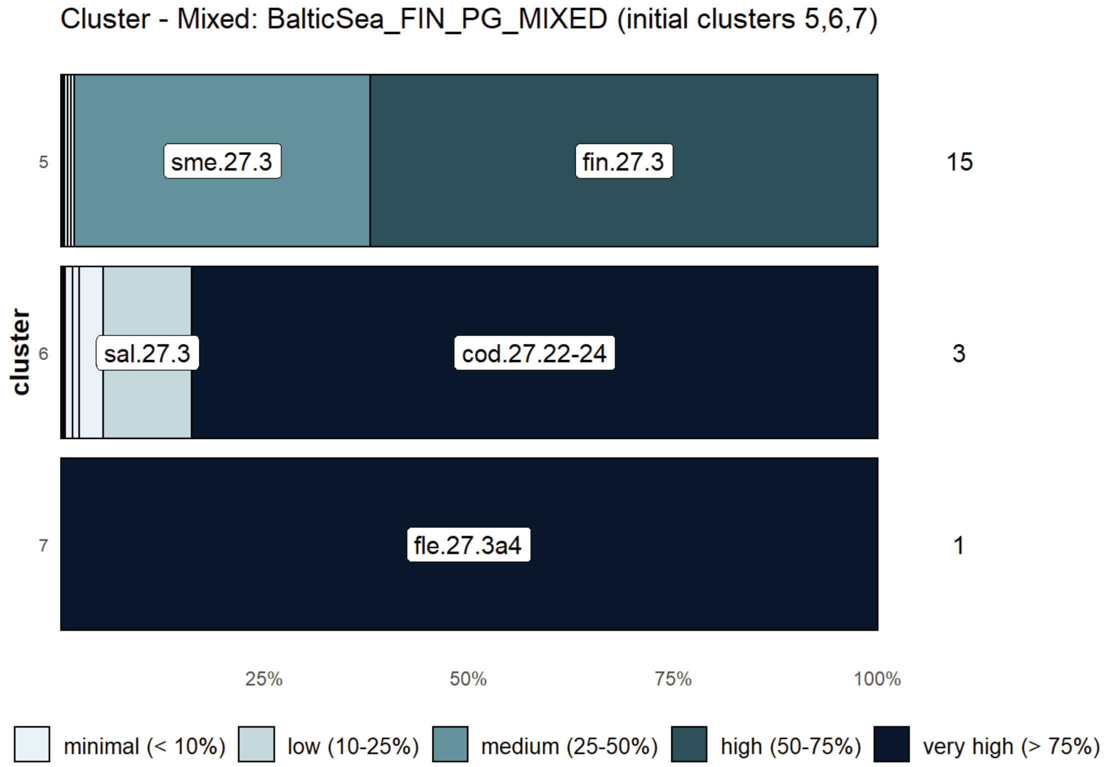


Figure 23: PG Cluster - Mixed

Discussion

We found the `fleetSegmentation` -package and the results provided by the package very interesting. The problems related to package's wider usage in official reporting and applications are most likely to be related to small cluster sizes (privacy and applicability of the analysis). In spite of this, we encourage the further development and usage of the package. We think that utilizing this tool should not be restricted only to official reporting. Instead, this approach could have a potential analysis usage at those stages where, e.g., number of vessels is invisible to end users. One example could be to examine the performance in imputation with, e.g., economical data - the traditional way to form stratum versus clustering approach.

Some keynotes that were not considered in these analysis due to time limitations but good to take account further.

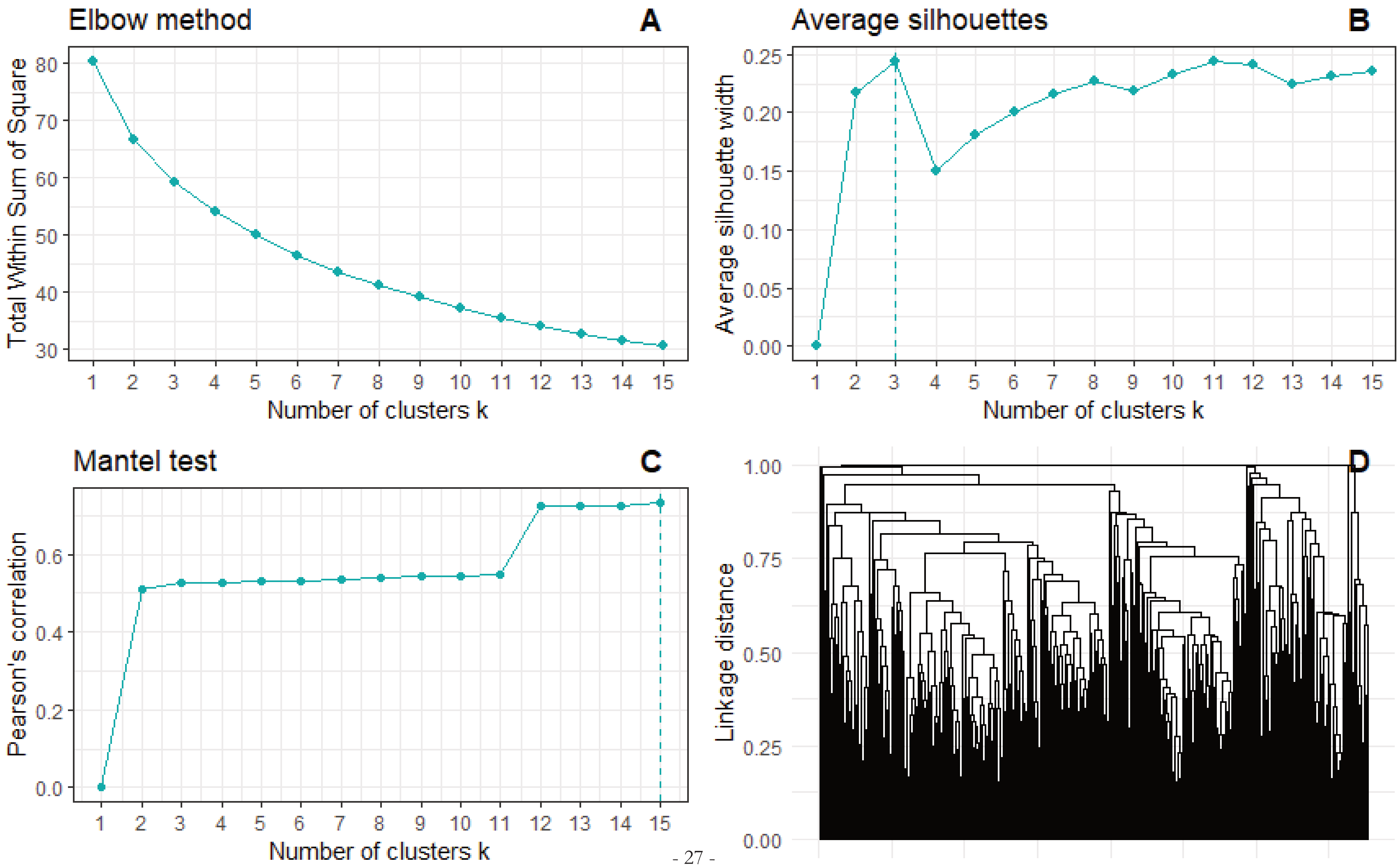
- **Value of landings** Would we see more variety if we include the value of landings into the analysis. This would take into account whether herring is used to feed or food production, which can lead to drastic differences in price and value. We are aware that value-option is already available in the package, but reliable results would require some data processing in terms of value calculations. Unfortunately we had no time to experiment this during the workshop.
- **Twin-trawlers** It would be interesting to test if twin-trawlers could be identified as an independent cluster.
- **Merging PG data** We know there are few small trawlers that report their catch with coastal landings report and are omitted from the data used here. They should be included.
- **PG's targeting Cod** It would have been nice to separate cod fisheries into their own cluster, but unfortunately the amount of vessels targeting cod is relatively small.
- **PG fishing on ice** We have an additional PG data containing the fishing without a ship ($n = 373$). This is mostly fishing on ice, but these were not analyzed in the workshop because of time limitations.

Finally, some suggestions for the future development.

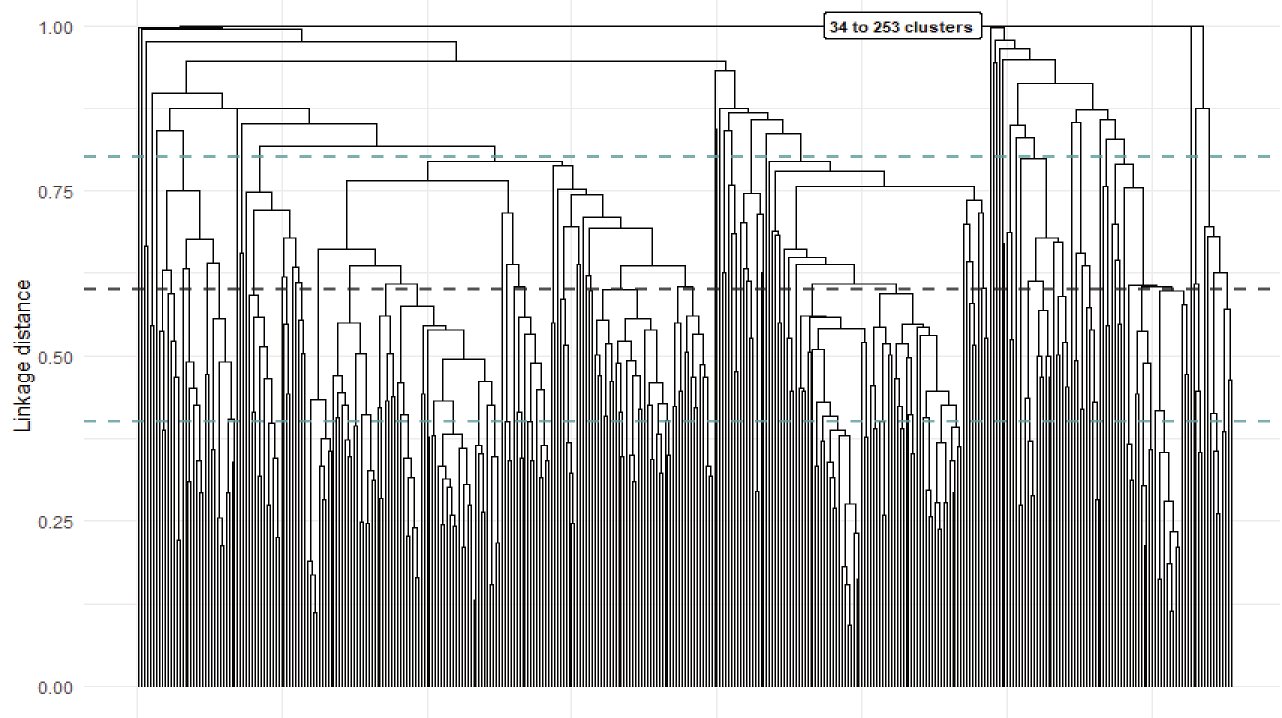
- **Clustering features** Some alternative clustering variables would be nice to include in the future. Our understanding is that currently the package supports a fixed list of clustering variables. Would be interesting to see what type of results we would get by using, e.g., a combination of horse power and catch measurement.
- **Combining the clusters** The strategy in combining clusters likely involves using the vessel properties at least to some extent. An additional function, which would leverage the cluster combining automatically w.r.t. vessel properties would be an interesting extension to the package.

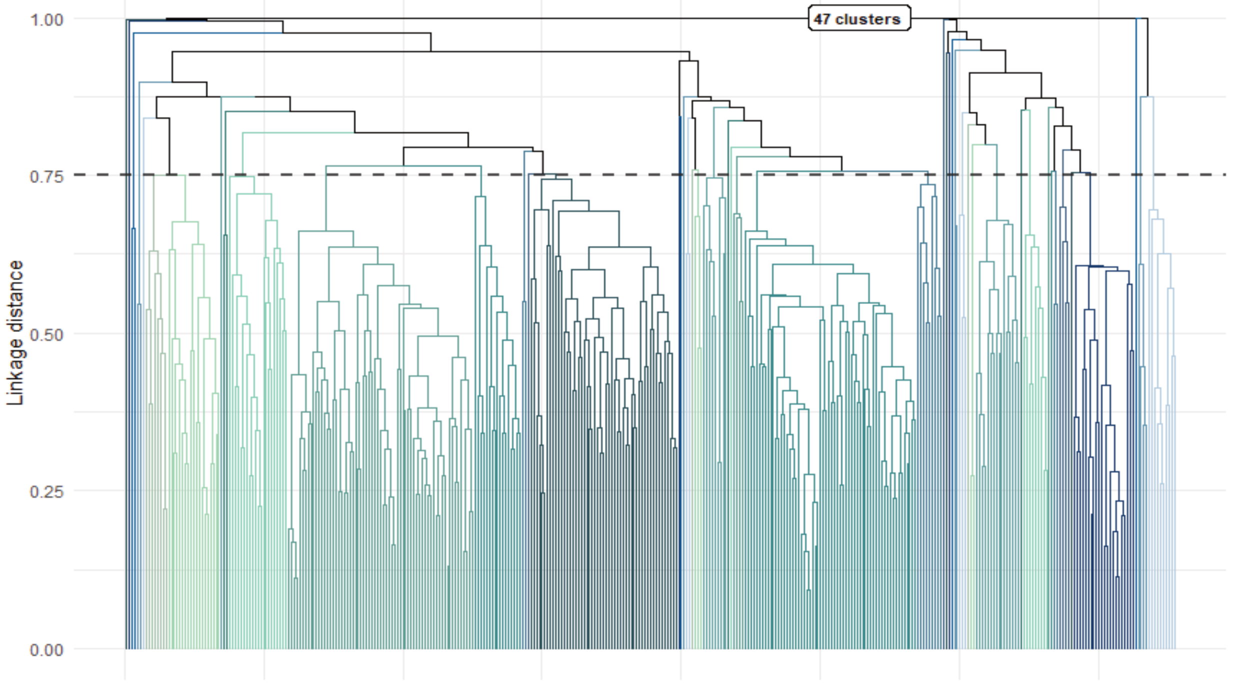
GR-DTS (377 vessels, GSA20, GSA 22, GSA23)

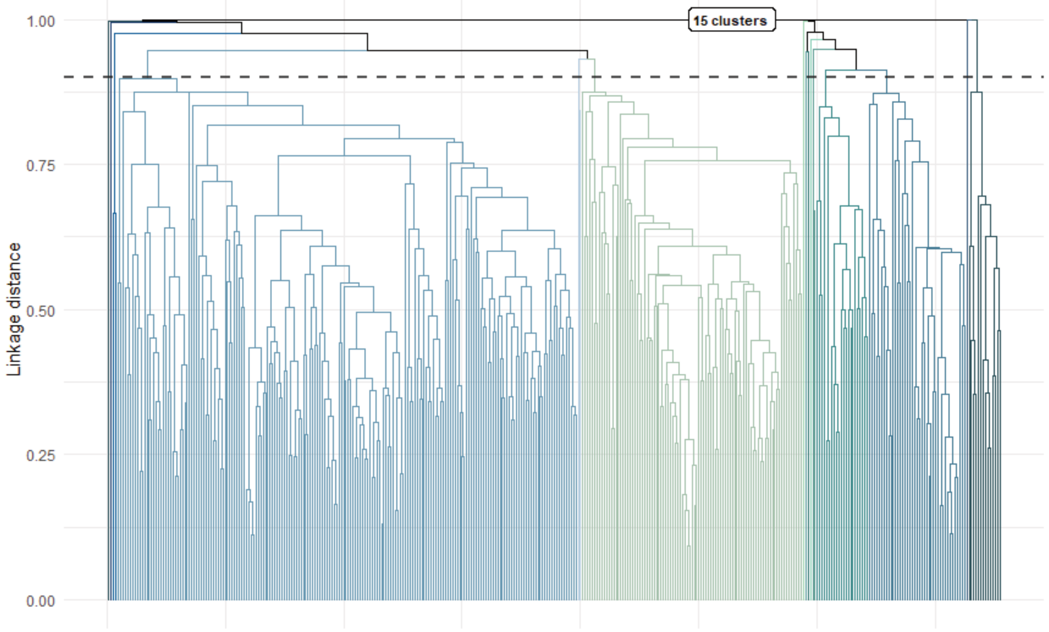
Agricultural Economics Research Institute (Greece)

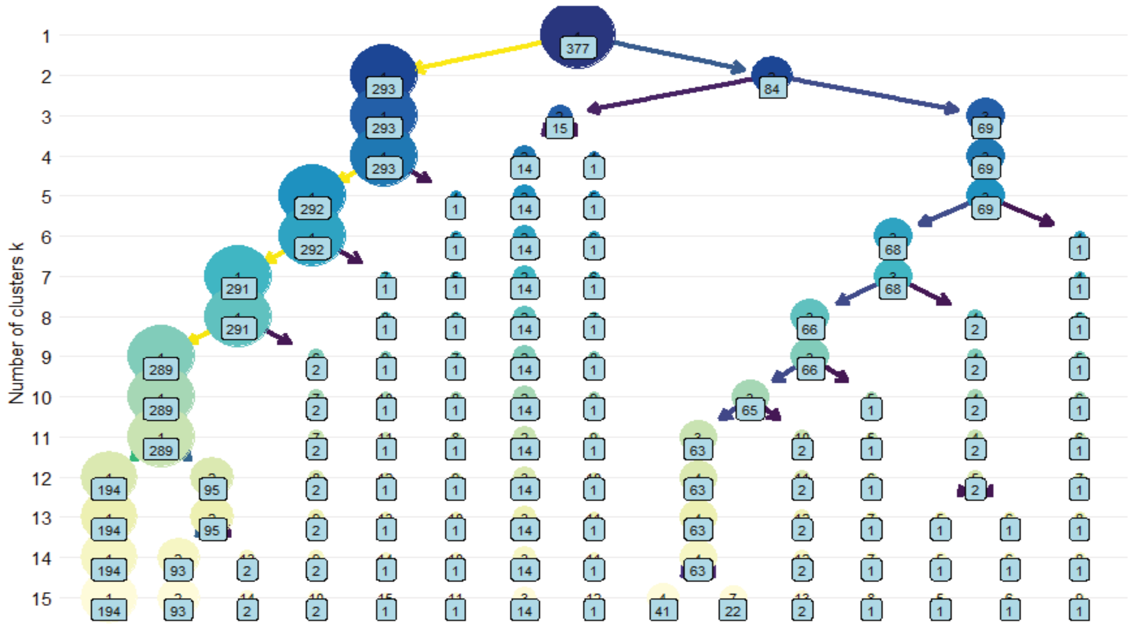


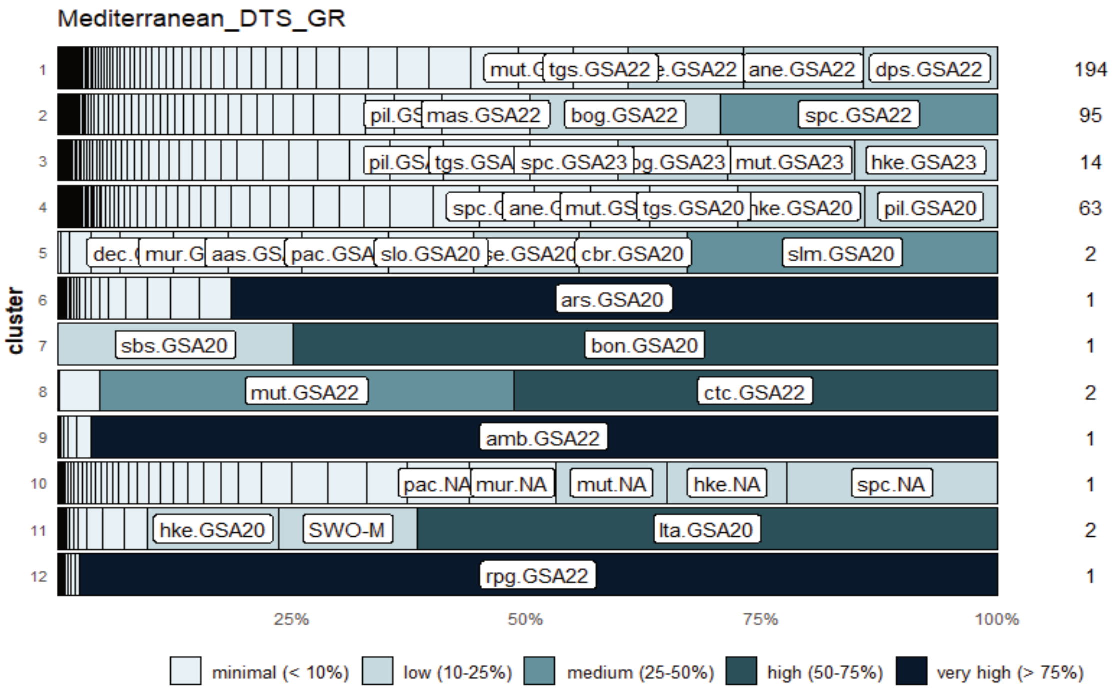
Elbow method
A
Total Within Sum of Square
Number of clusters k
Average silhouettes
B
Average silhouette width
Number of clusters k
Mantel test
C
Pearson's correlation
Number of clusters k
D
Linkage distance

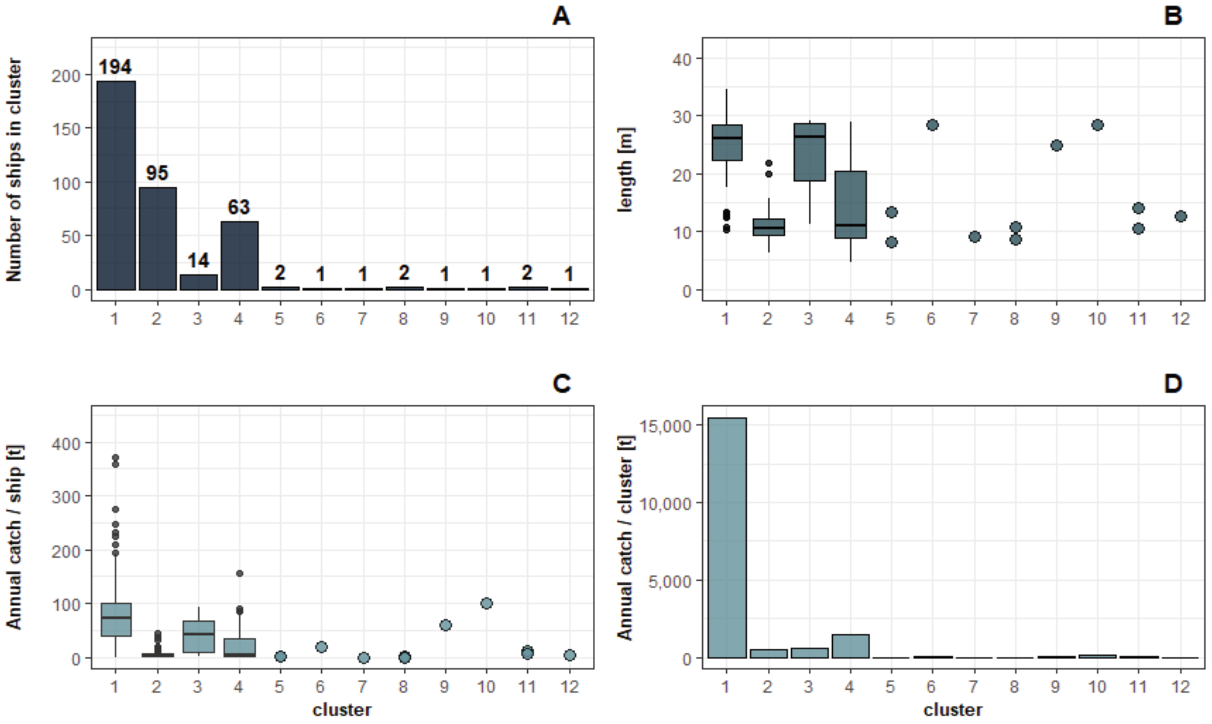

34 to 253 clusters
Linkage distance
1.00
0.75
0.50
0.25
0.00

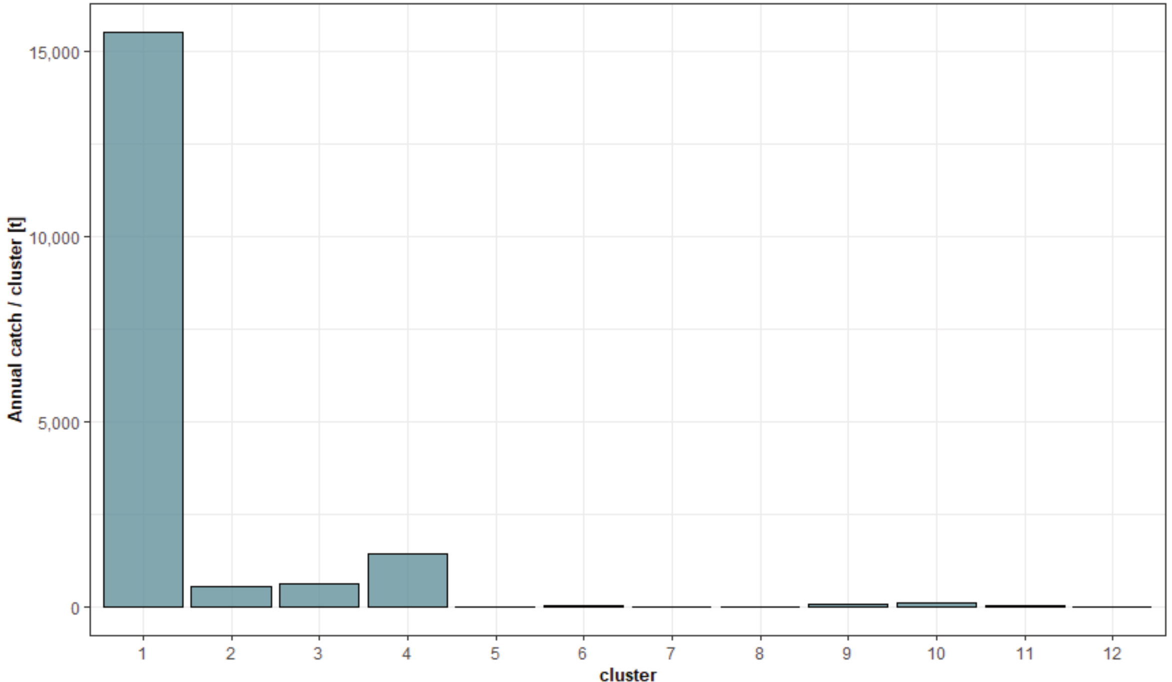

47 clusters
Linkage distance
1.00
0.75
0.50
0.25
0.00

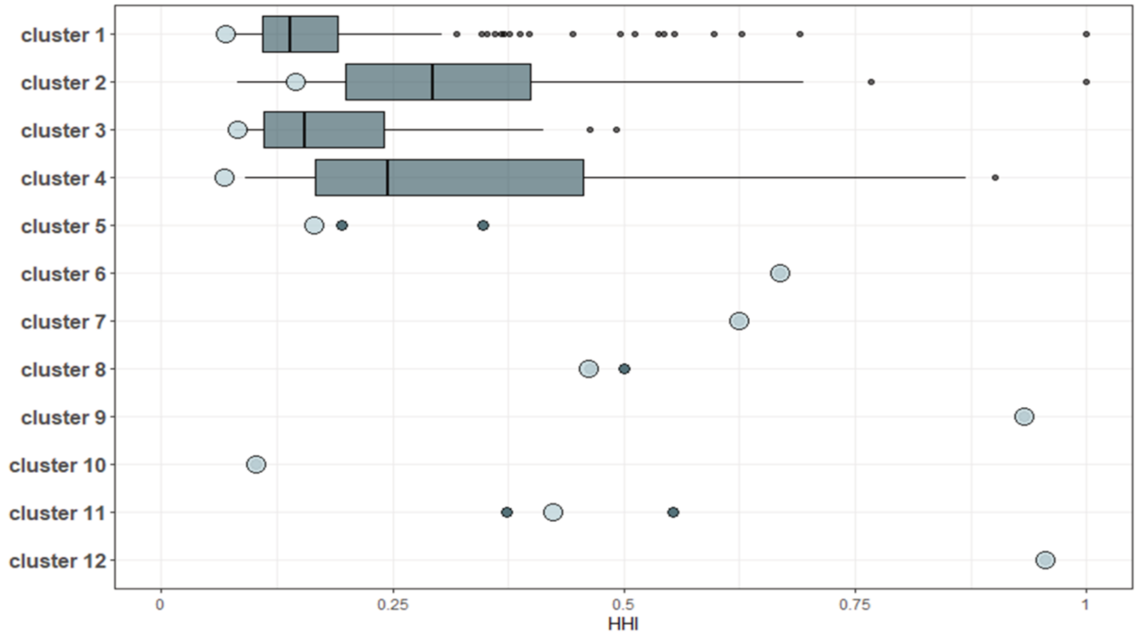

15 clusters
Linkage distance
1.00
0.75
0.50
0.25
0.00

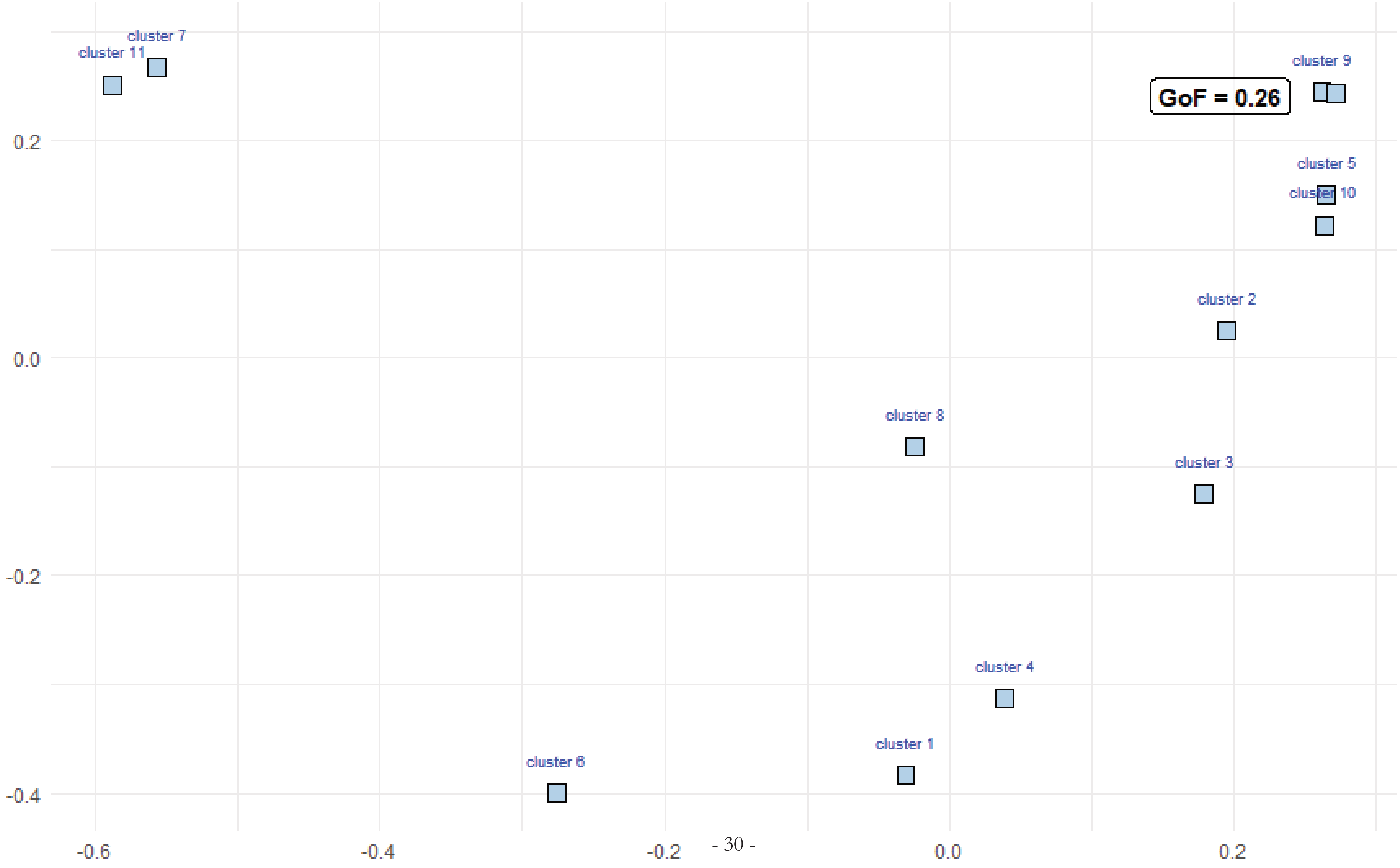

Number of clusters k
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15


Mediterranean_DTS_GR
cluster
1 2 3 4 5 6 7 8 9 10 11 12
194 95 14 63 2 1 1 2 1 1 2 1
mut.G tgs.GSA22 e.GSA22 ane.GSA22 dps.GSA22
pil.GS mas.GSA22 bog.GSA22 spc.GSA22
pil.GS tgs.GSA spc.GSA23 og.GSA23 mut.GSA23 hke.GSA23
spc.G ane.G mut.GS tgs.GSA20 hke.GSA20 pil.GSA20
dec.G mur.G aas.GS pac.GSA slo.GSA20 se.GSA20 cbr.GSA20 slm.GSA20
ars.GSA20
sbs.GSA20 bon.GSA20
mut.GSA22 ctc.GSA22
amb.GSA22
pac.NA mur.NA mut.NA hke.NA spc.NA
hke.GSA20 SWO-M lta.GSA20
rpg.GSA22
25% 50% 75% 100%
minimal (< 10%) low (10-25%) medium (25-50%) high (50-75%) very high (> 75%)


A
Number of ships in cluster
194 95 14 63 2 1 1 2 1 1 2 1
B
length [m]
C
Annual catch / ship [t]
cluster
D
Annual catch / cluster [t]
cluster

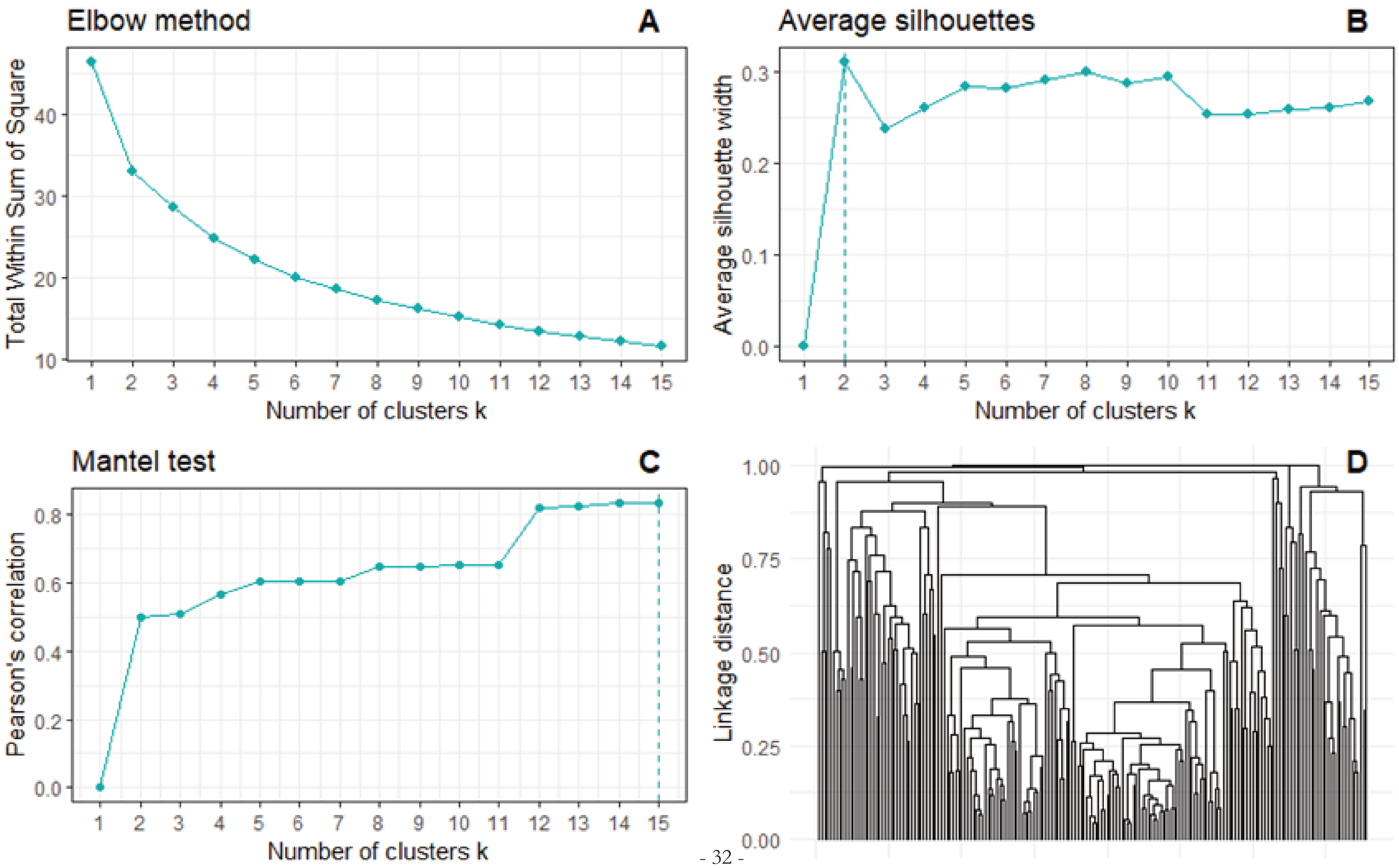

Annual catch / cluster [t]
15,000
10,000
5,000
0
1 2 3 4 5 6 7 8 9 10 11 12
cluster


cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
cluster 7
cluster 8
cluster 9
cluster 10
cluster 11
cluster 12
0 0.25 0.5 0.75 1
HHI

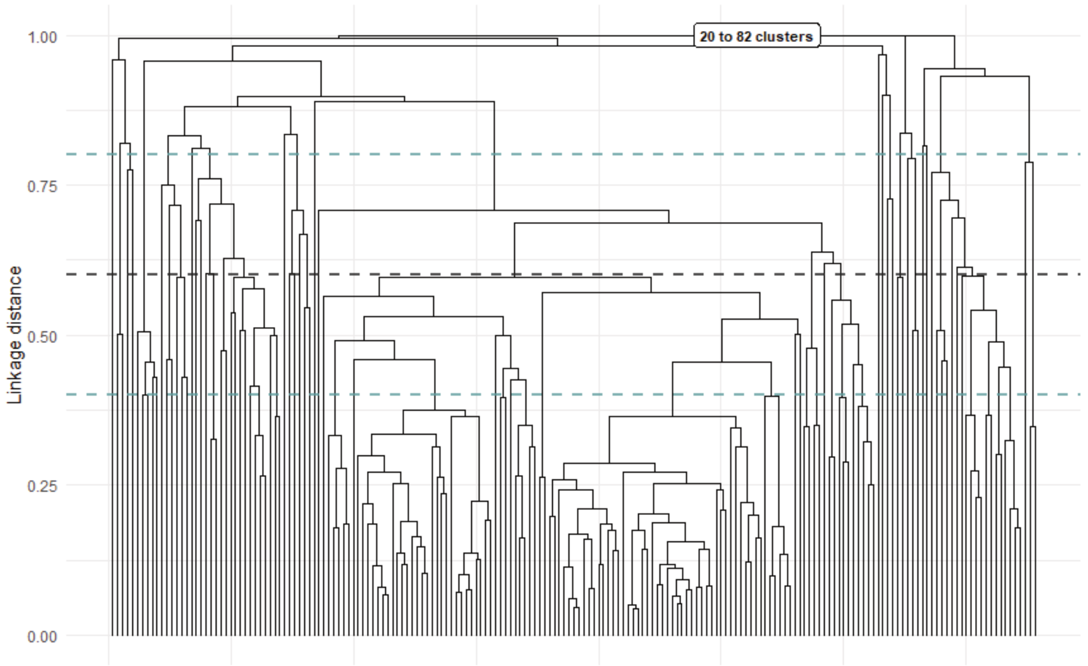

cluster 7
cluster 11
cluster 9
GoF = 0.26
cluster 5
cluster 10
cluster 2
cluster 8
cluster 3
cluster 4
cluster 1
cluster 6
0.2
0.0
-0.2
-0.4
-0.6
-0.4
-0.2
0.0
0.2

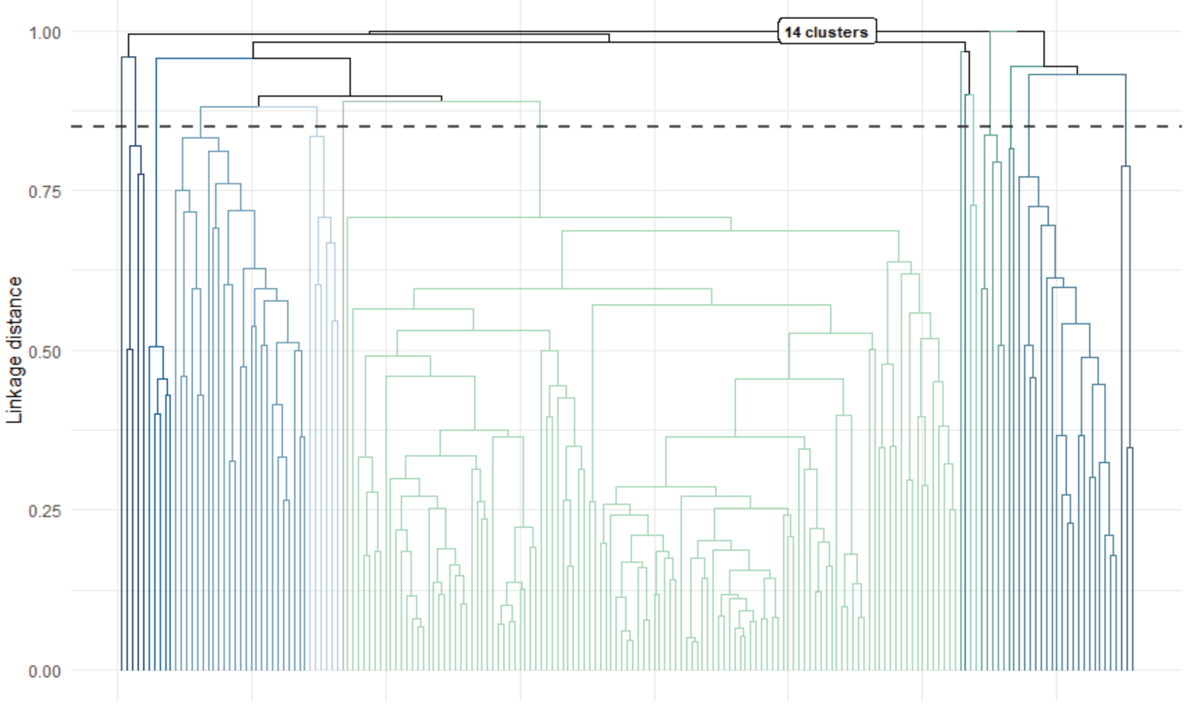
GR-PS (189 vessels, GSA20, GSA 22, GSA23)

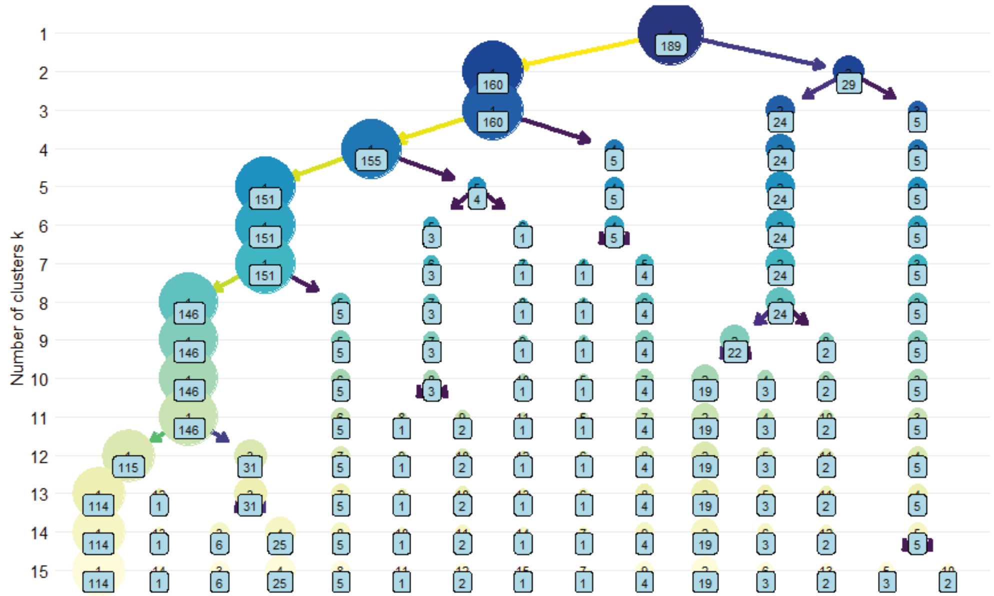
Agricultural Economics Research Institute (Greece)

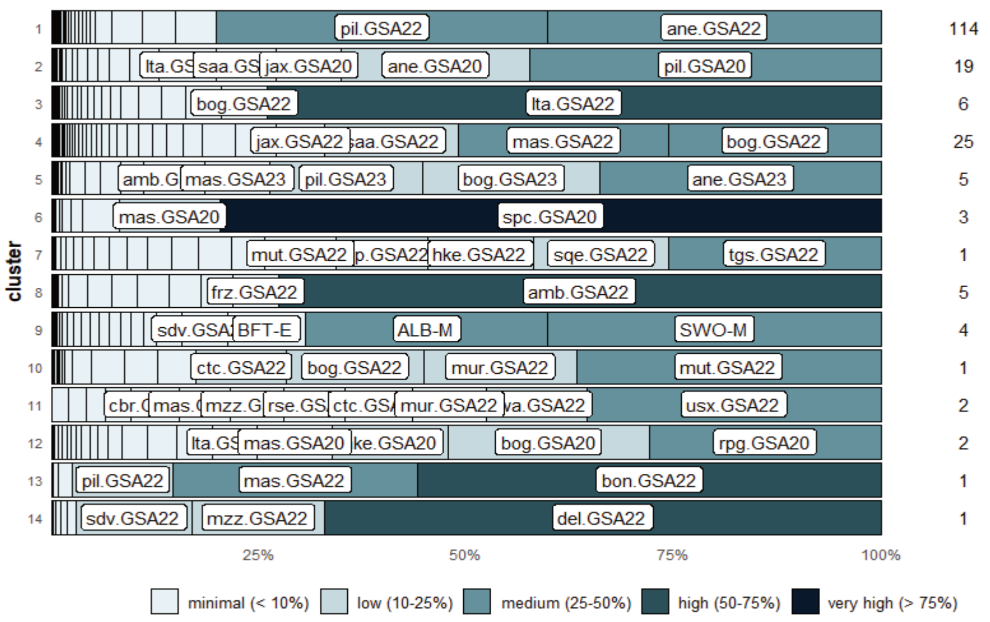
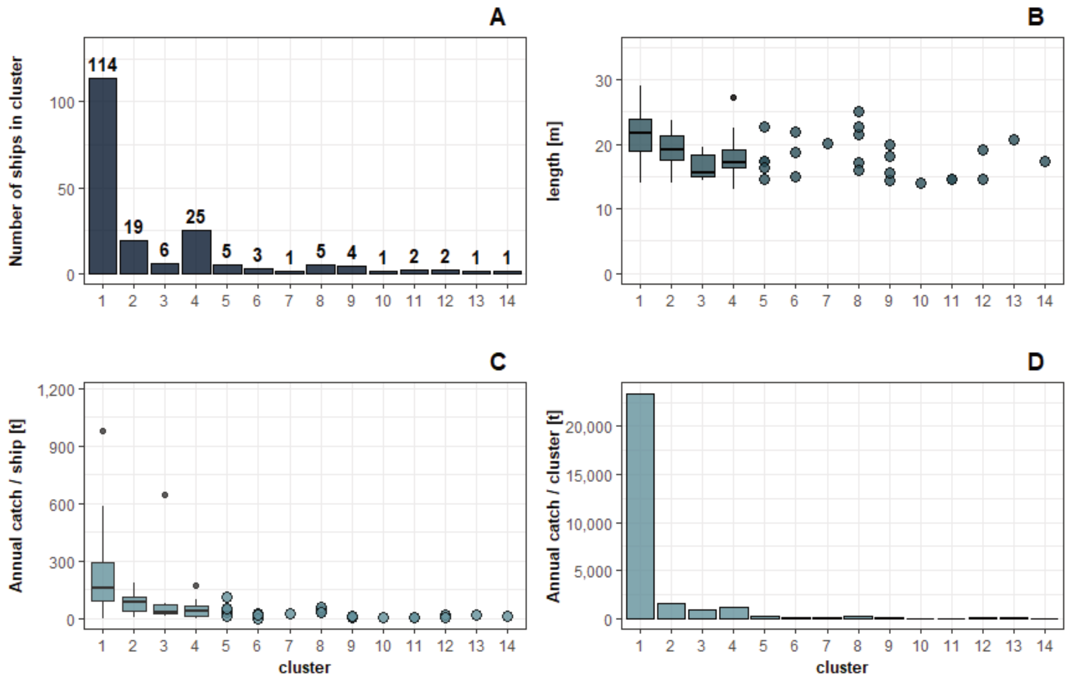
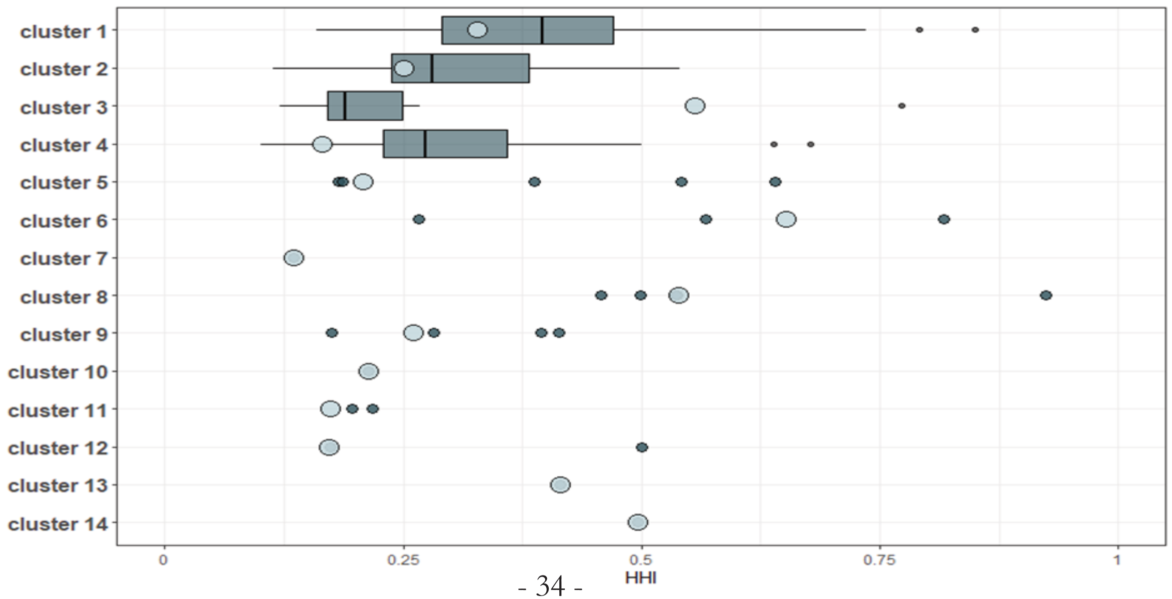


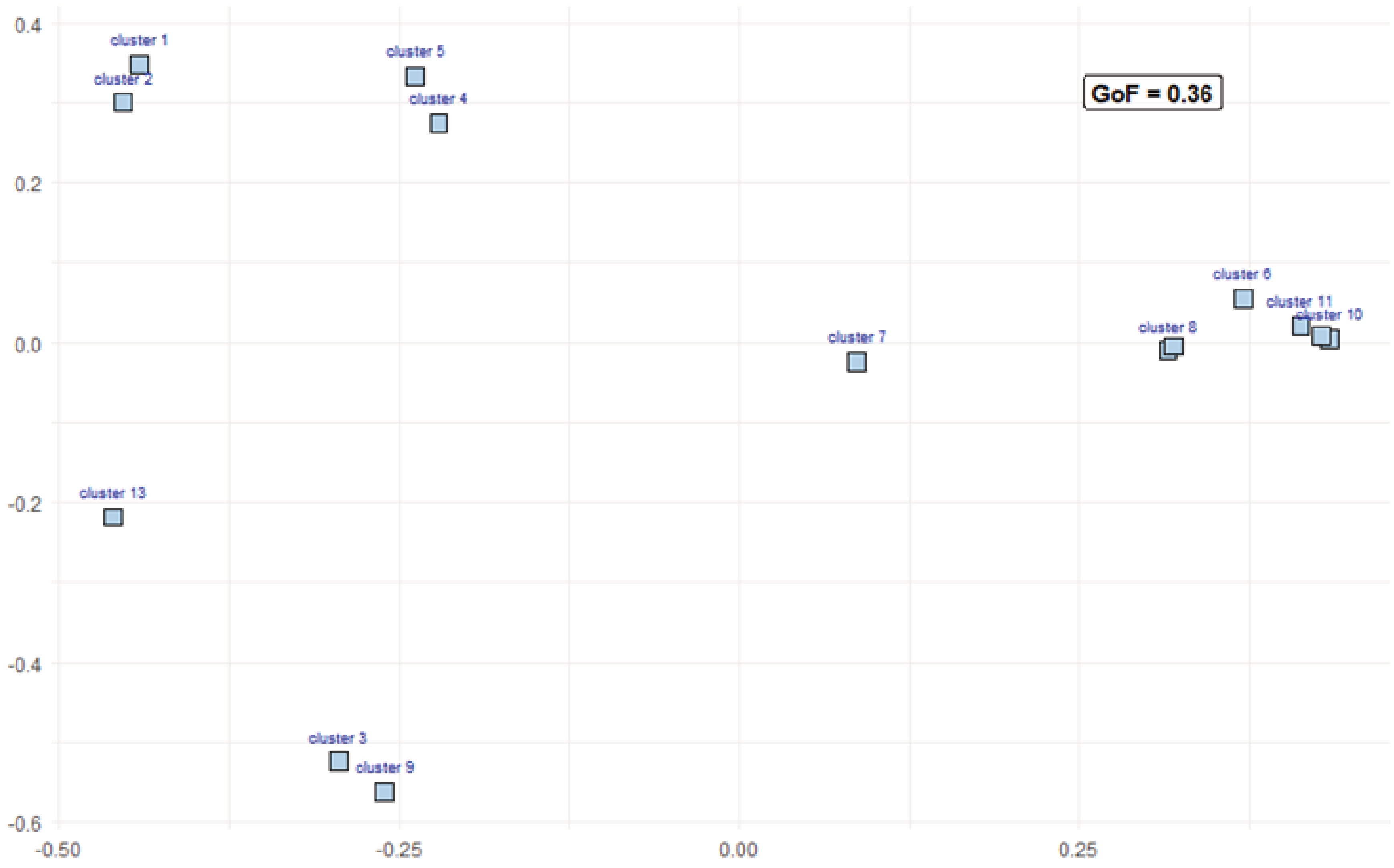
Elbow method
A
Total Within Sum of Square
Number of clusters k
Average silhouettes
B
Average silhouette width
Number of clusters k
Mantel test
C
Pearson's correlation
Number of clusters k
D
Linkage distance


20 to 82 clusters
Linkage distance
1.00
0.75
0.50
0.25
0.00


14 clusters
Linkage distance
1.00
0.75
0.50
0.25
0.00


Number of clusters k
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15


cluster
pil.GSA22
ane.GSA22
114
lta.GS
saa.GS
jax.GSA20
ane.GSA20
pil.GSA20
19
bog.GSA22
lta.GSA22
6
jax.GSA22
saa.GSA22
mas.GSA22
bog.GSA22
25
amb.G
mas.GSA23
pil.GSA23
bog.GSA23
ane.GSA23
5
mas.GSA20
spc.GSA20
3
mut.GSA22
p.GSA22
hke.GSA22
sqe.GSA22
tgs.GSA22
1
frz.GSA22
amb.GSA22
5
sdv.GSA
BFT-E
ALB-M
SWO-M
4
ctc.GSA22
bog.GSA22
mur.GSA22
mut.GSA22
1
cbr.
mas.
mzz.G
rse.GS
ctc.GS
mur.GSA22
va.GSA22
usx.GSA22
2
lta.GS
mas.GSA20
ke.GSA20
bog.GSA20
rpg.GSA20
2
pil.GSA22
mas.GSA22
bon.GSA22
1
sdv.GSA22
mzz.GSA22
del.GSA22
1
25%
50%
75%
100%
minimal (< 10%)
low (10-25%)
medium (25-50%)
high (50-75%)
very high (> 75%)

A
Number of ships in cluster
114
19
6
25
5
3
1
5
4
1
2
2
1
1
B
length [m]
C
Annual catch / ship [t]
cluster
D
Annual catch / cluster [t]
cluster

cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
cluster 7
cluster 8
cluster 9
cluster 10
cluster 11
cluster 12
cluster 13
cluster 14
0
0.25
0.5
0.75
1
HHI



GoF = 0.36
cluster 1
cluster 2
cluster 5
cluster 4
cluster 6
cluster 11
cluster 10
cluster 8
cluster 7
cluster 13
cluster 3
cluster 9
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.50
-0.25
0.00
0.25

Fleet Segmentation Romania BlackSea(GSA29)

Fishing techniques: PMP & PG

Daniel GRIGORAȘ

Cătălin PĂUN

Fleet of Romania in 2021

- The Romanian fishing vessel fleet was analyzed in 2021 and composed of 130 active vessels which are divided into two fishing techniques respectively PMP with 52 vessels and PG with 78 vessels.
- A vessel is assigned to the PMP fishing technique when using multiple gears (both active gear and passive gear or only active gear), and a vessel is classified as PG when using one or more passive gears.
- We have the following length segments: 24-40 m PMP, 18-24 m PMP, 12-18 m PMP, 06-12 m PMP, 06-12 m PG, 00-06 m PG.

The table of indices resulting

- PMP

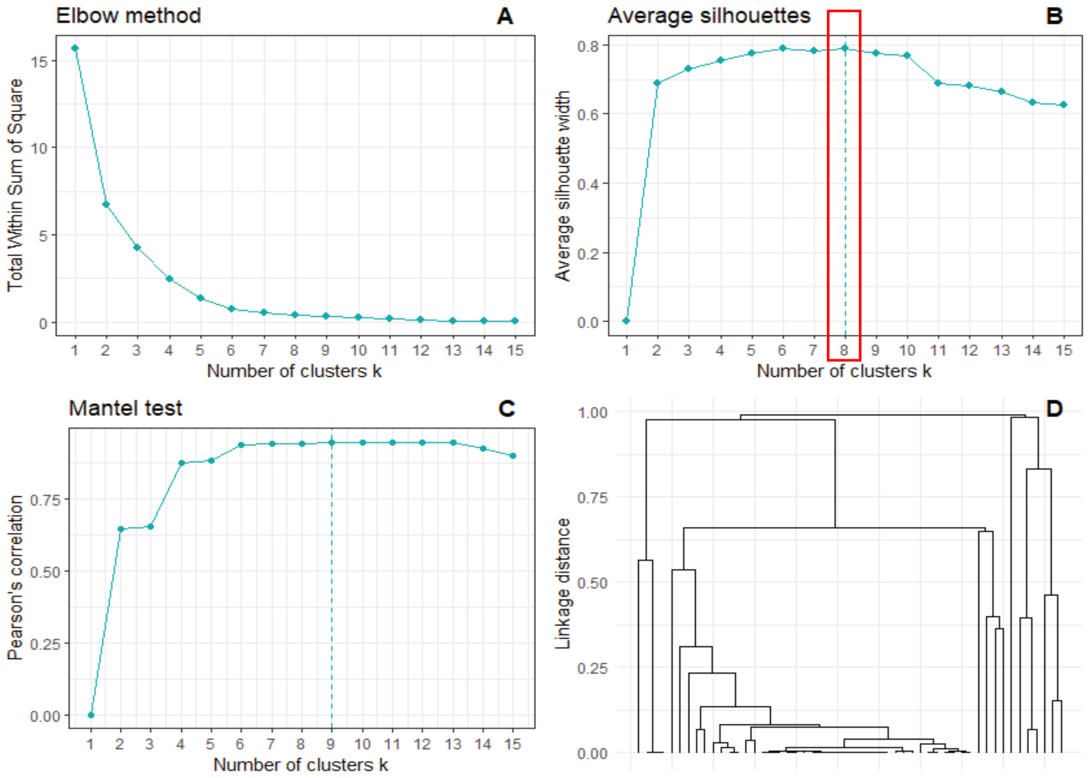
	Indices	Optimal number of clusters	Index value
1	Average silhouettes	8	0.788
2	Mantel test	9	0.949
3	Davis-Bouldin index	14	0.088
4	SD index	5	2.884
5	Calinski-Harabasz index	15	1129.278

- PG

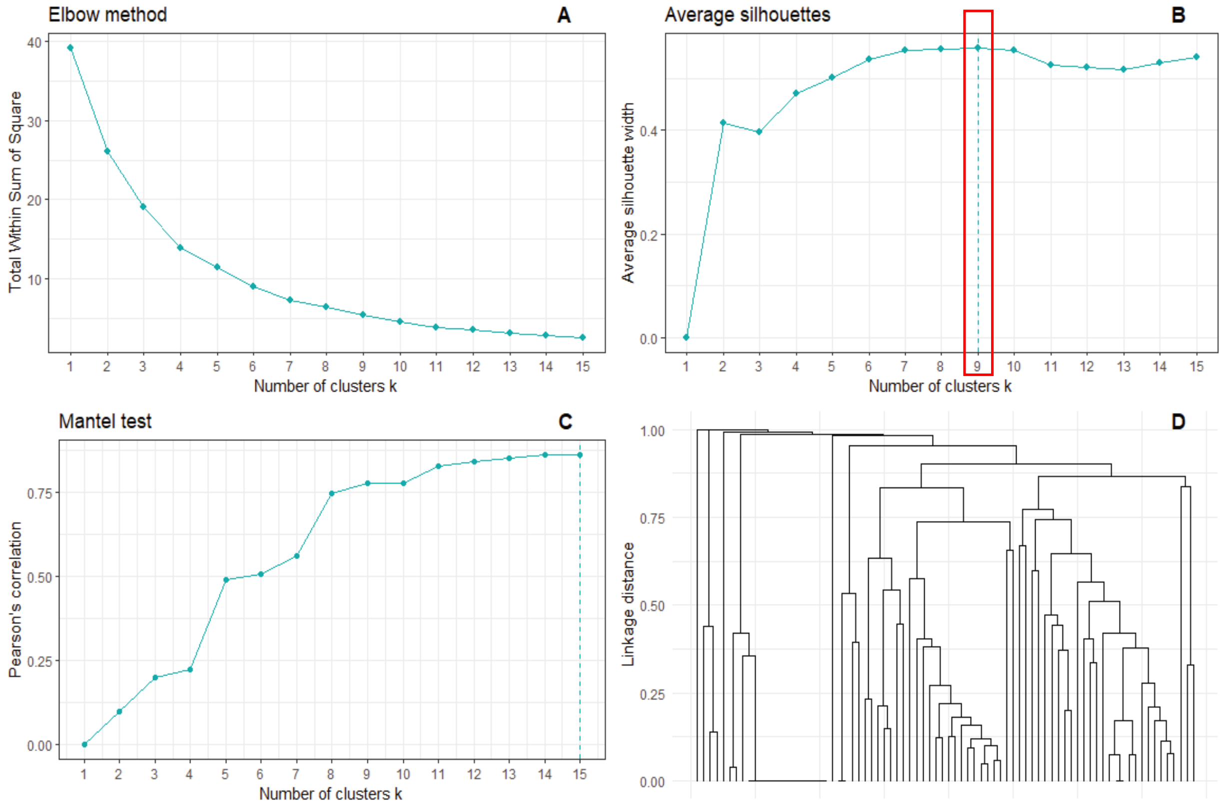
	Indices	Optimal number of clusters	Index value
1	Average silhouettes	9	0.559
2	Mantel test	15	0.861
3	Davis-Bouldin index	11	0.531
4	SD index	6	2.110
5	Calinski-Harabasz index	14	44.184

Explore the number of clusters with plots

- PMP

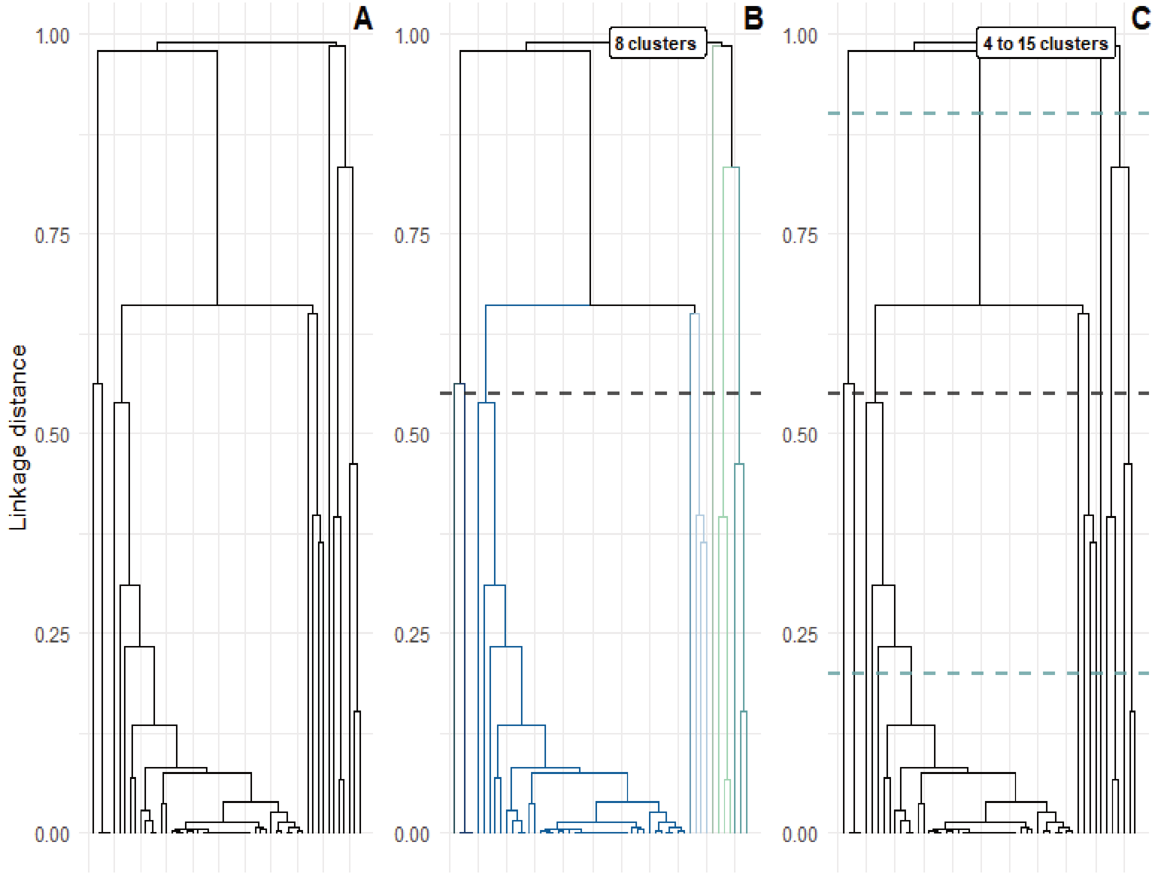


- PG

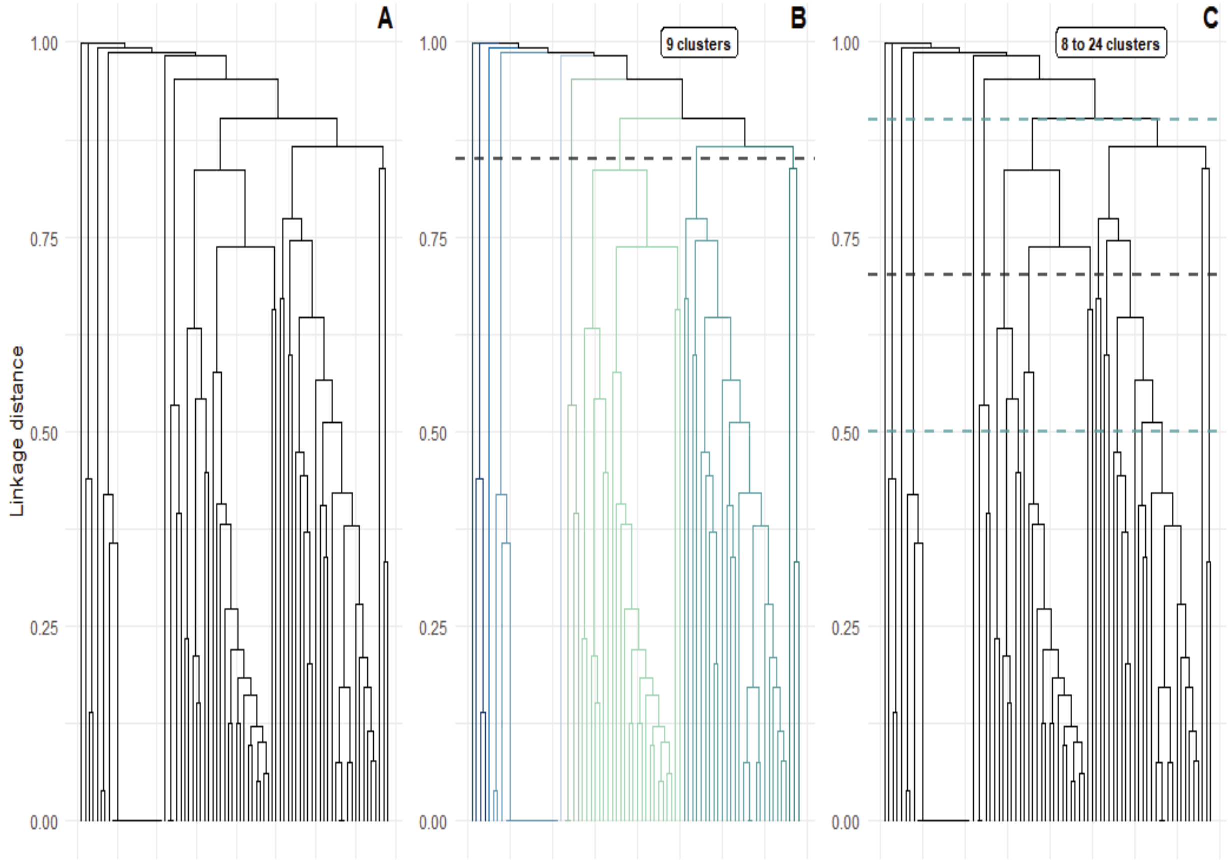


Analyzing and diagnostics of clusters

- PMP

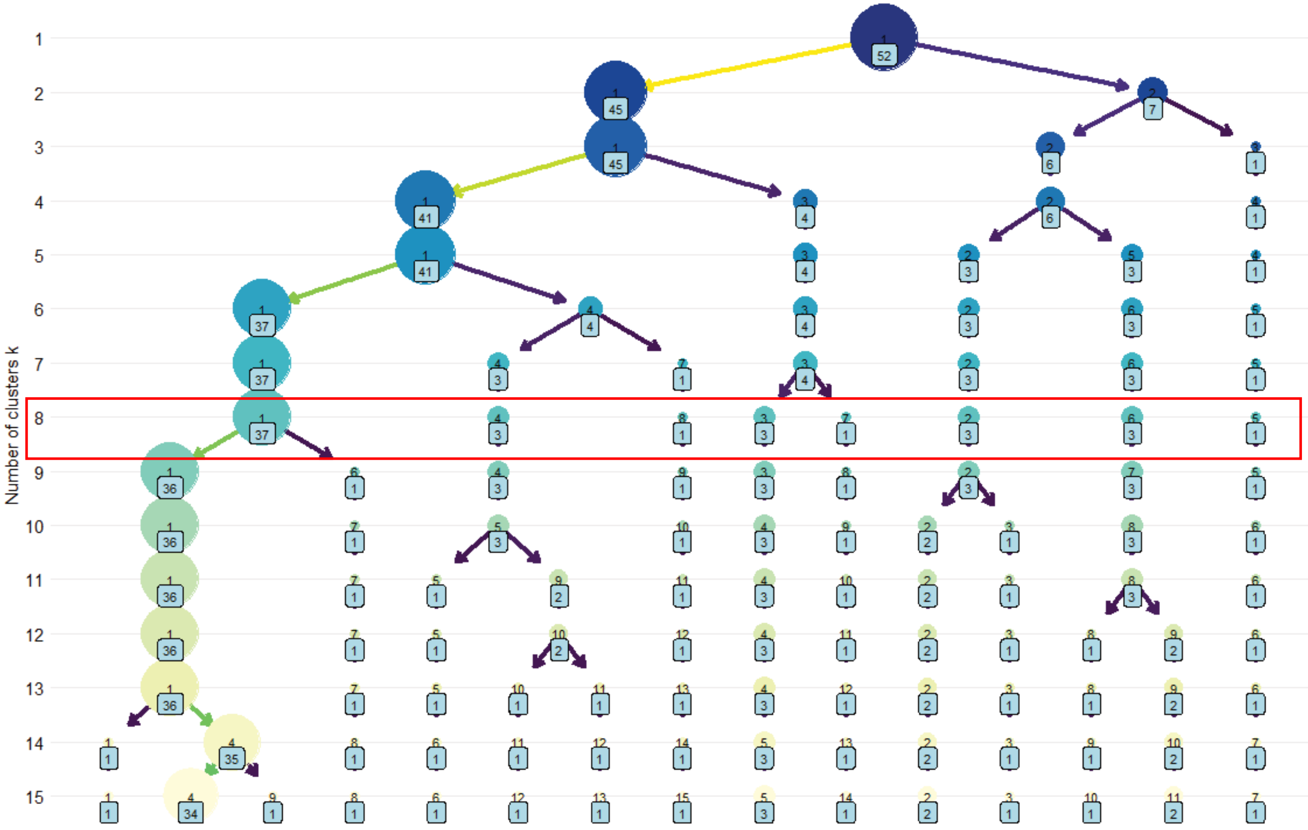


- PG

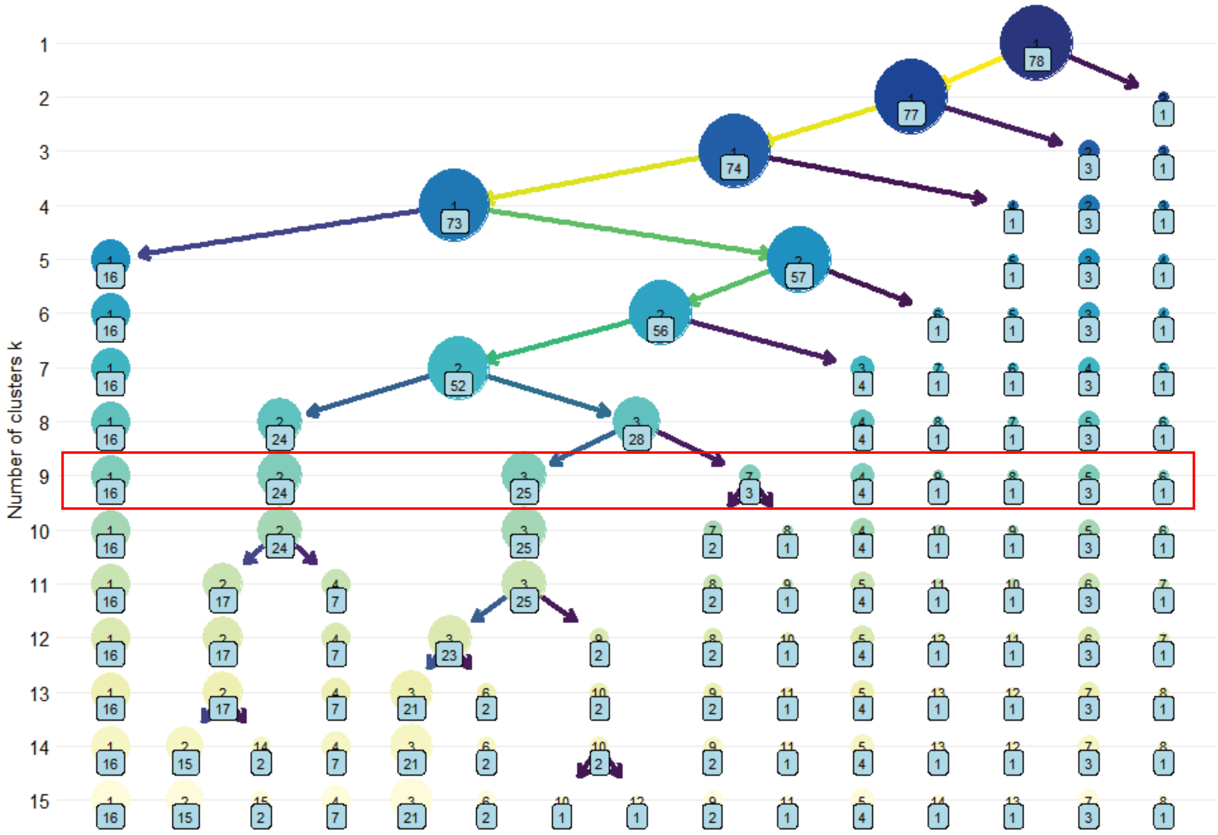


Analyzing and identifying the suitable cluster

- PMP

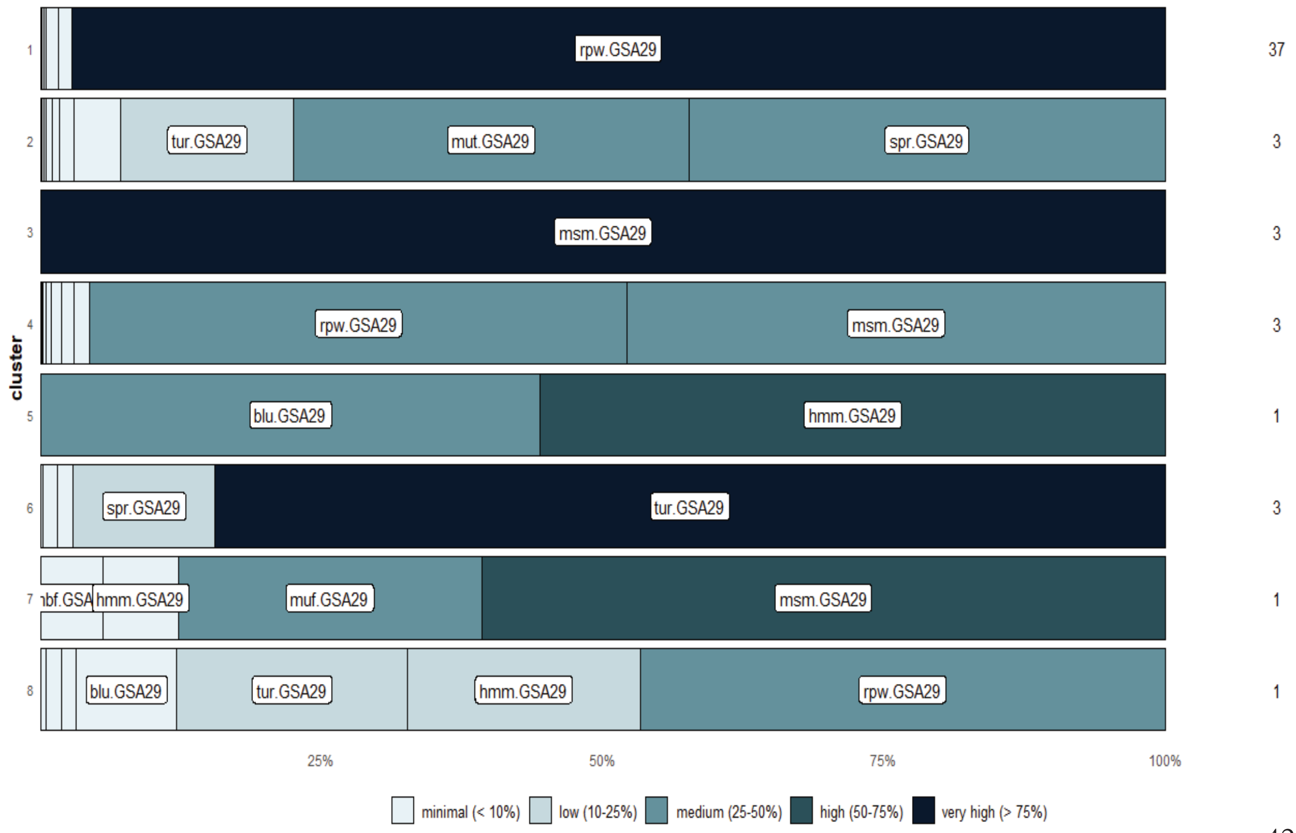


- PG

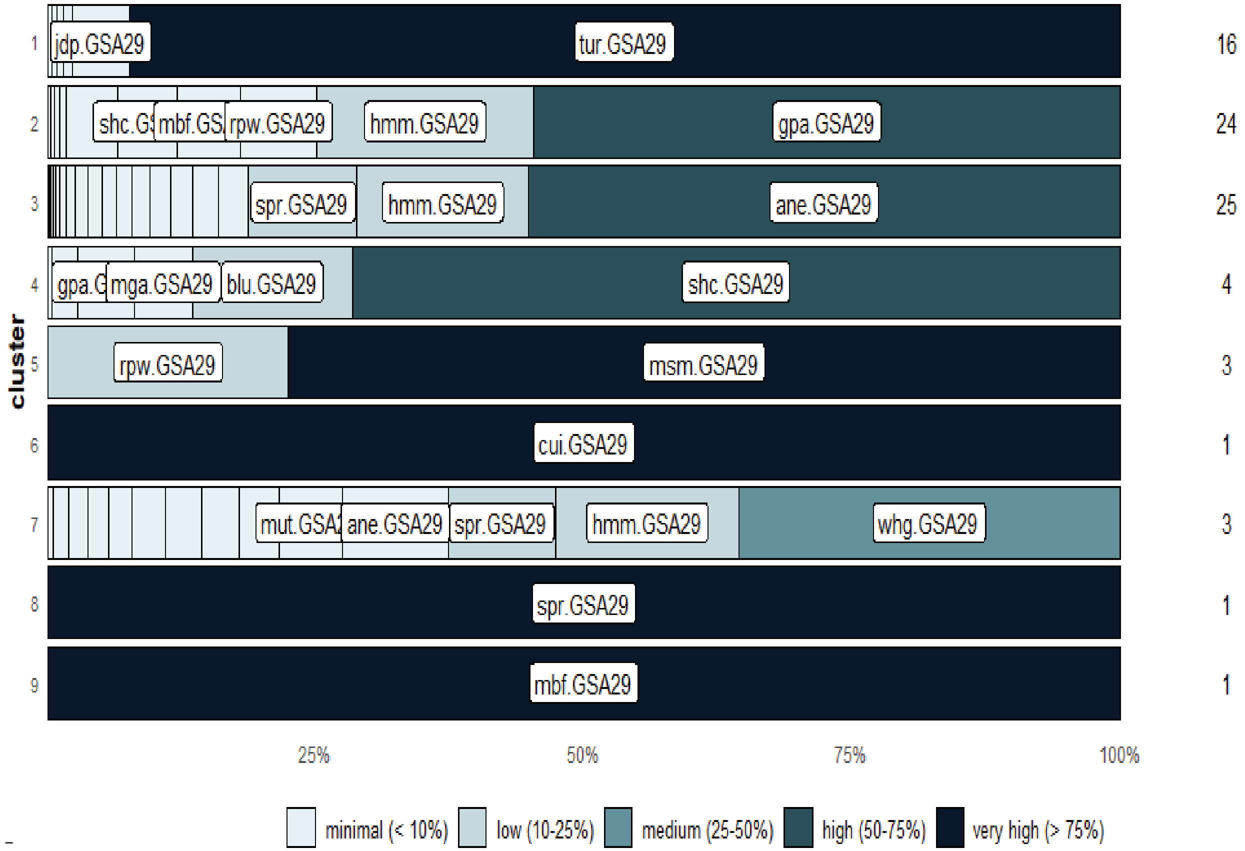


Catch composition of clusters

- PMP

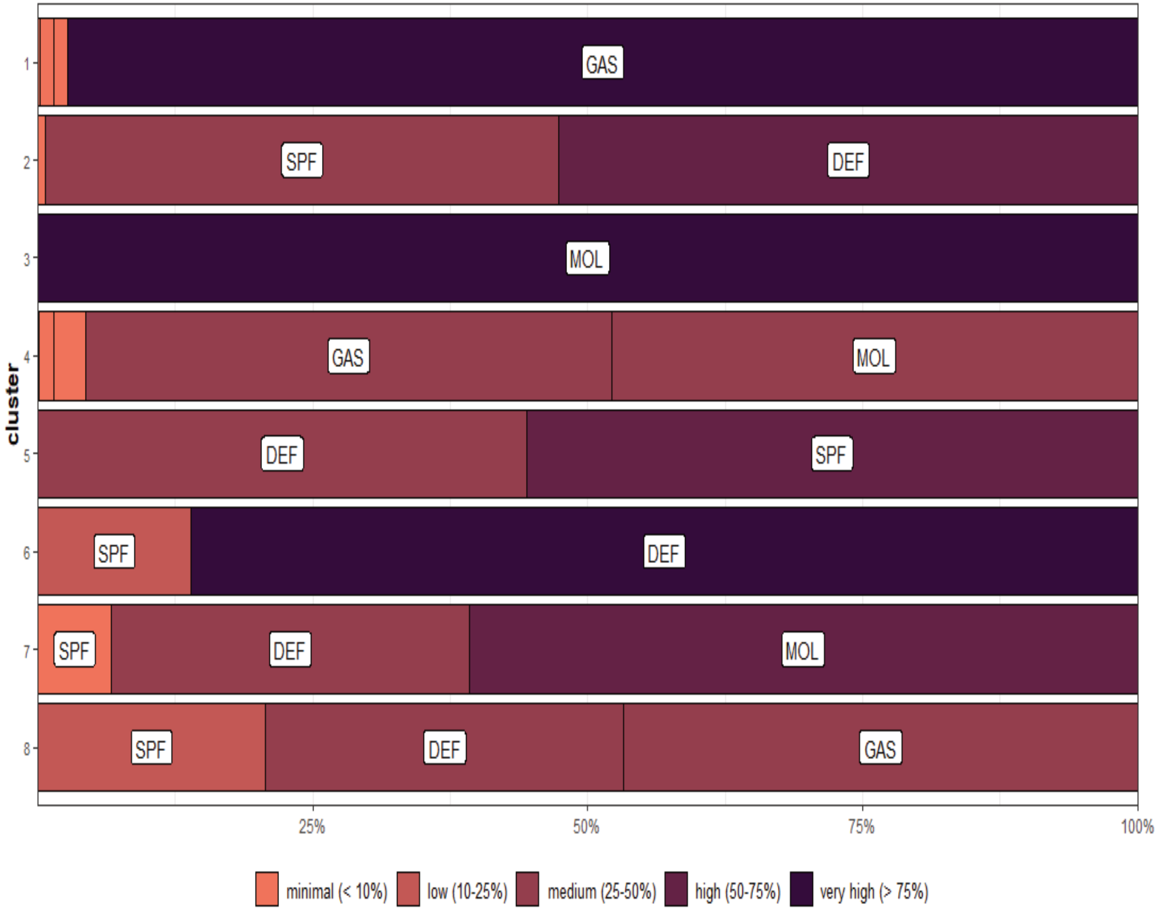


- PG

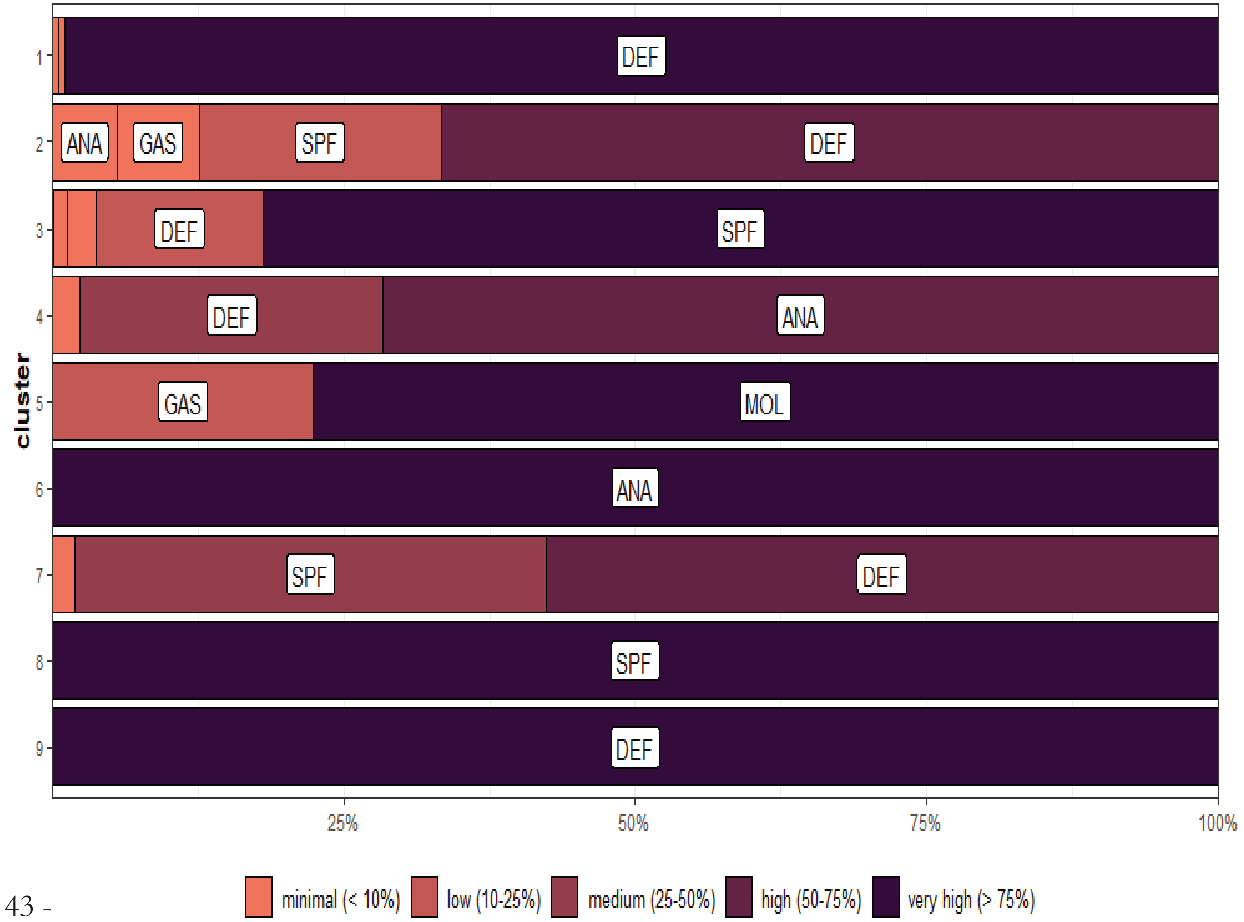


Target assemblage clustering

- PMP

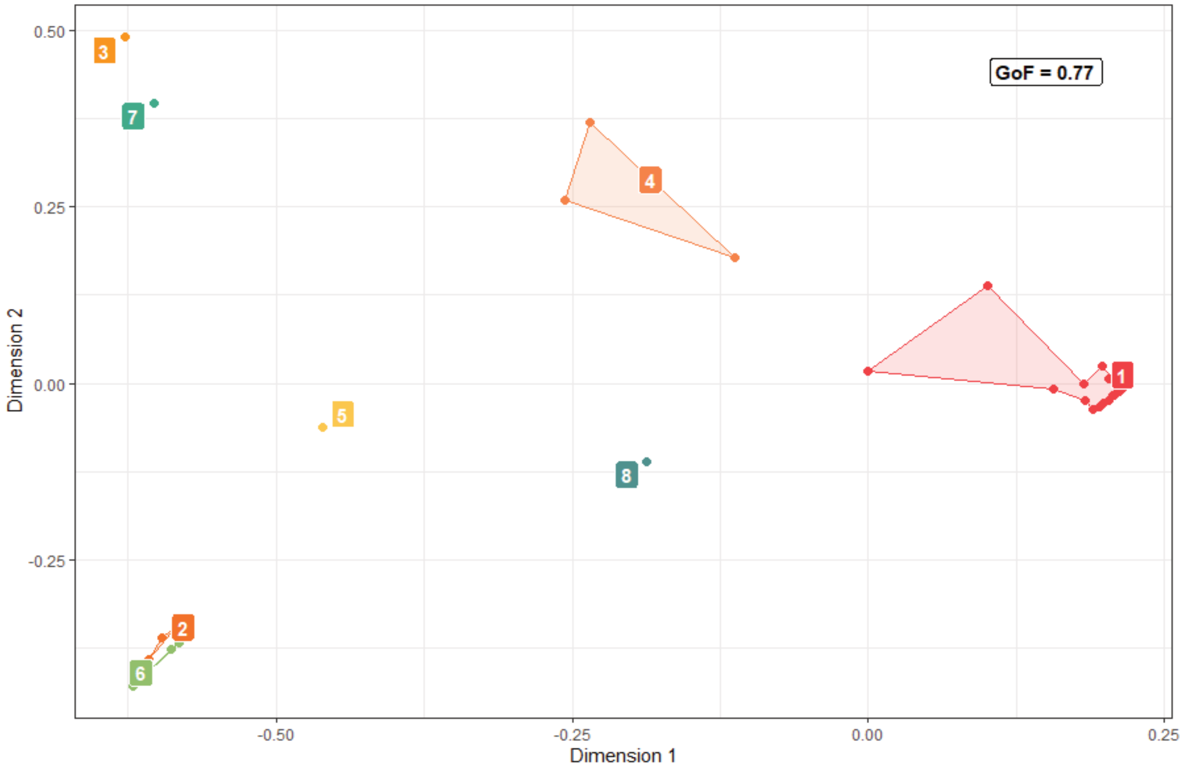


- PG

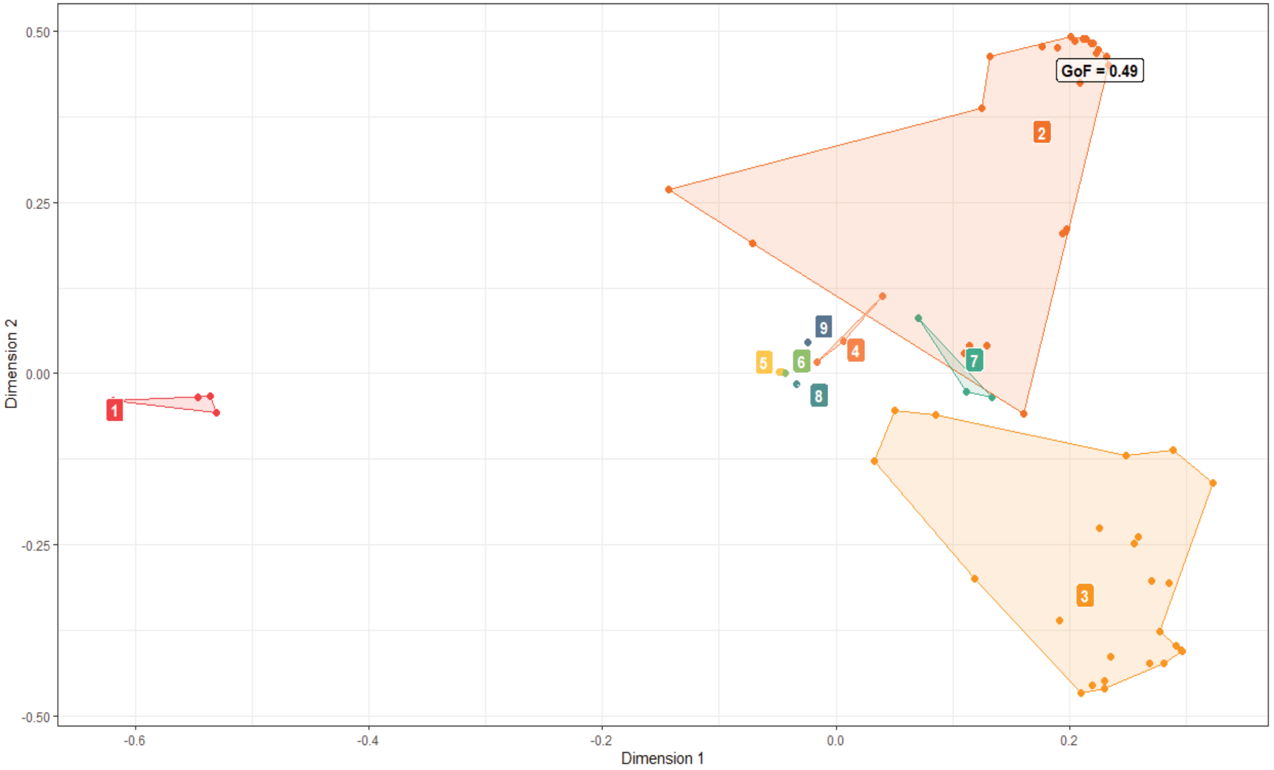


Relations in the catch composition and GoF (goodness of fit): plotting a 2-dimensional ordination

- PMP

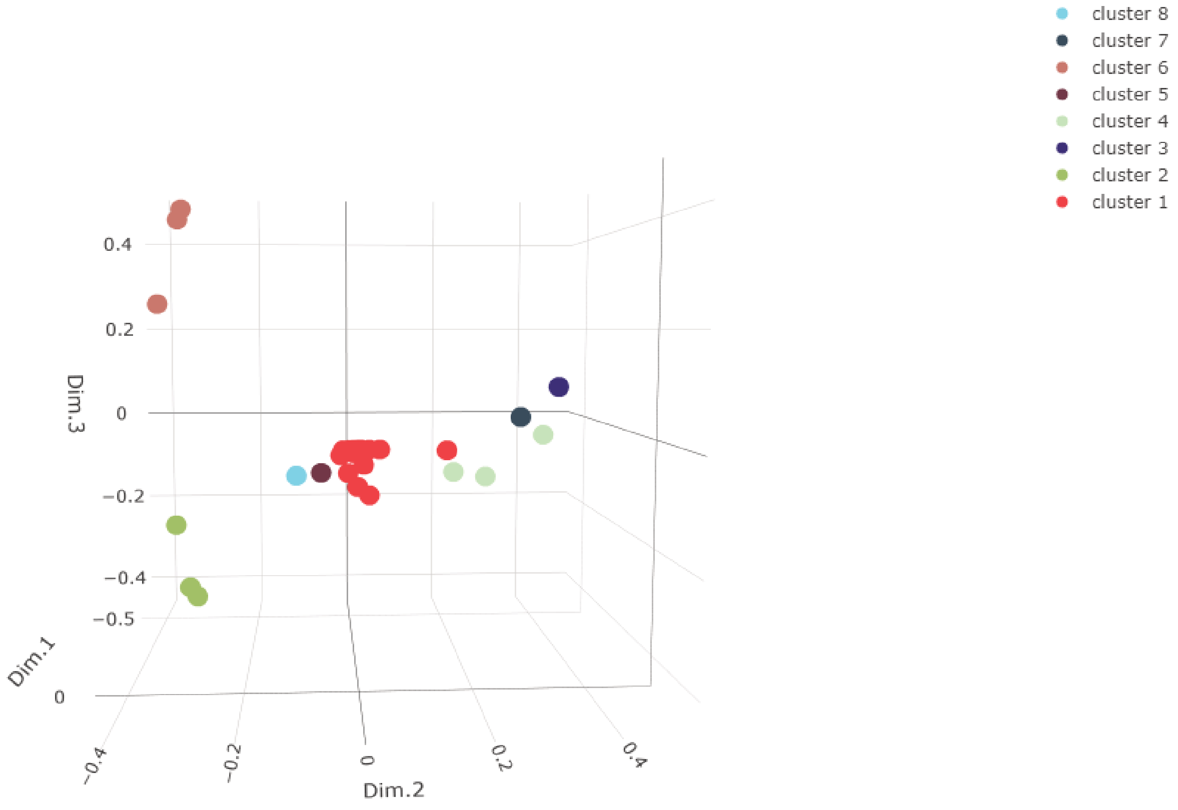


- PG

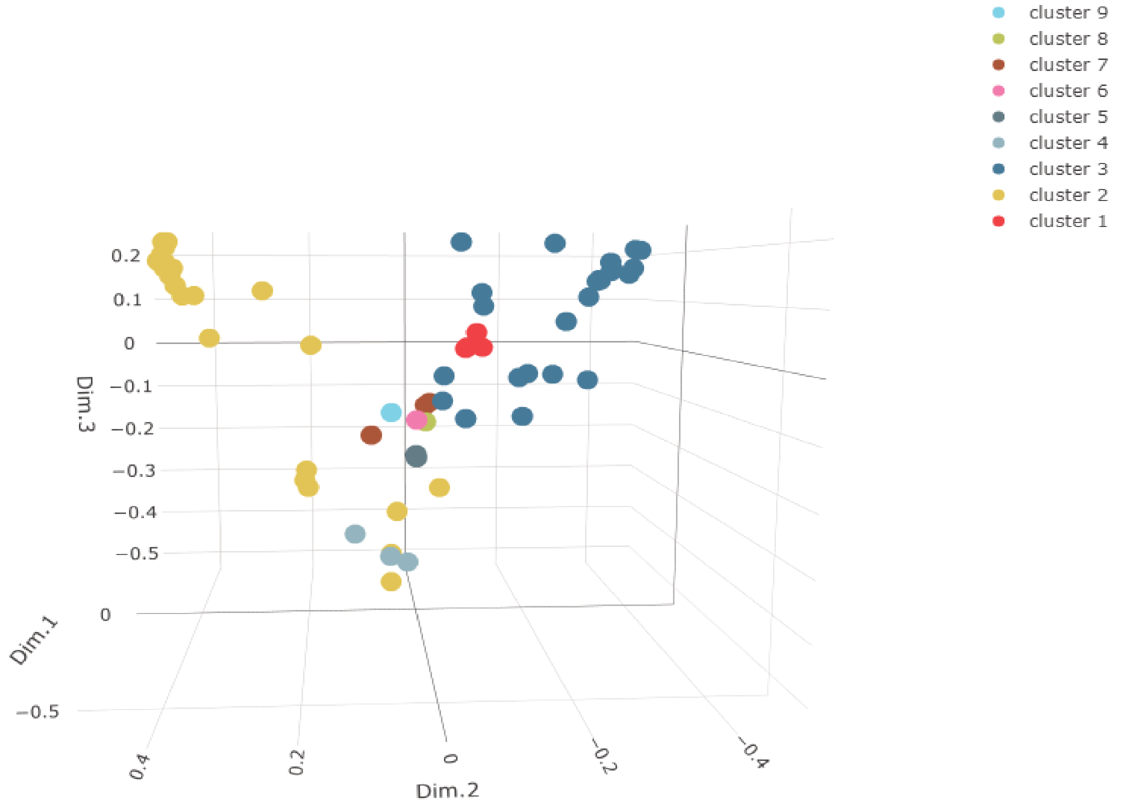


Relations in the catch composition and GoF (goodness of fit): plotting a 3-dimensional ordination

PMP

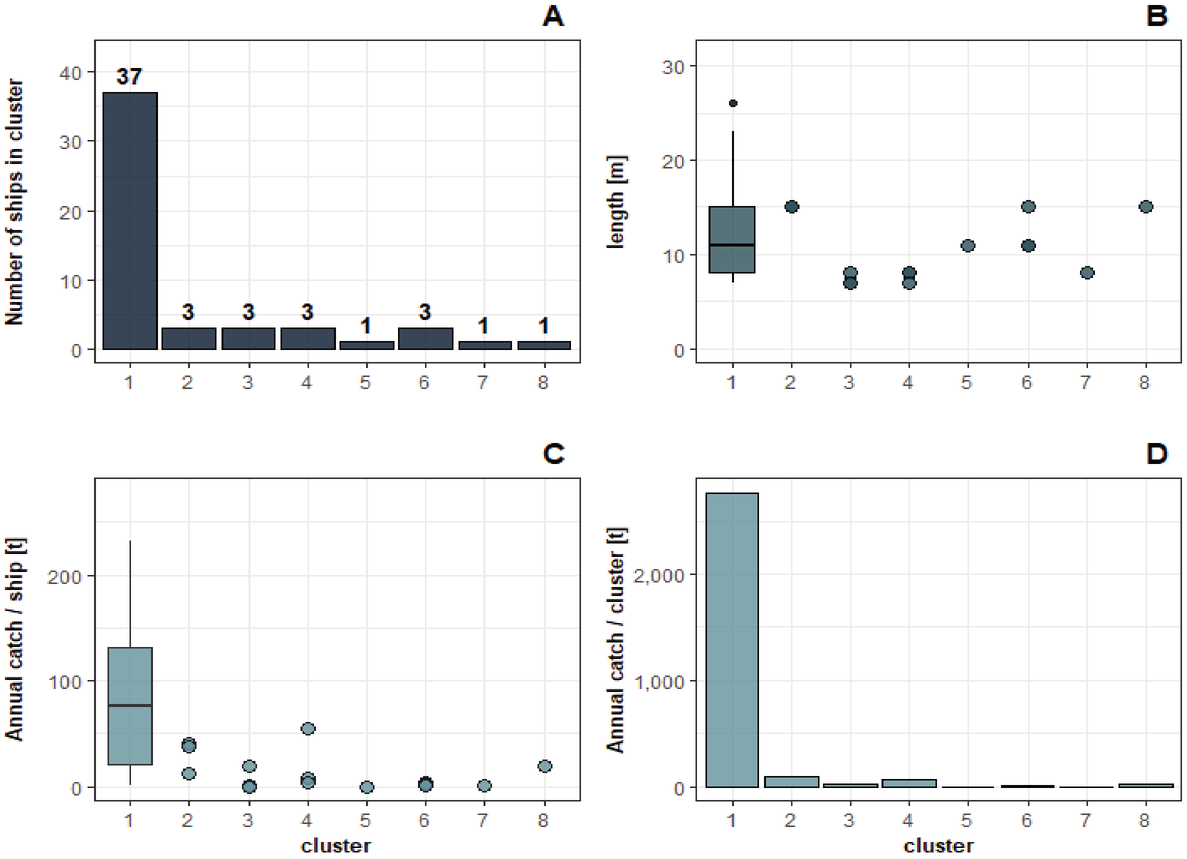


- PG

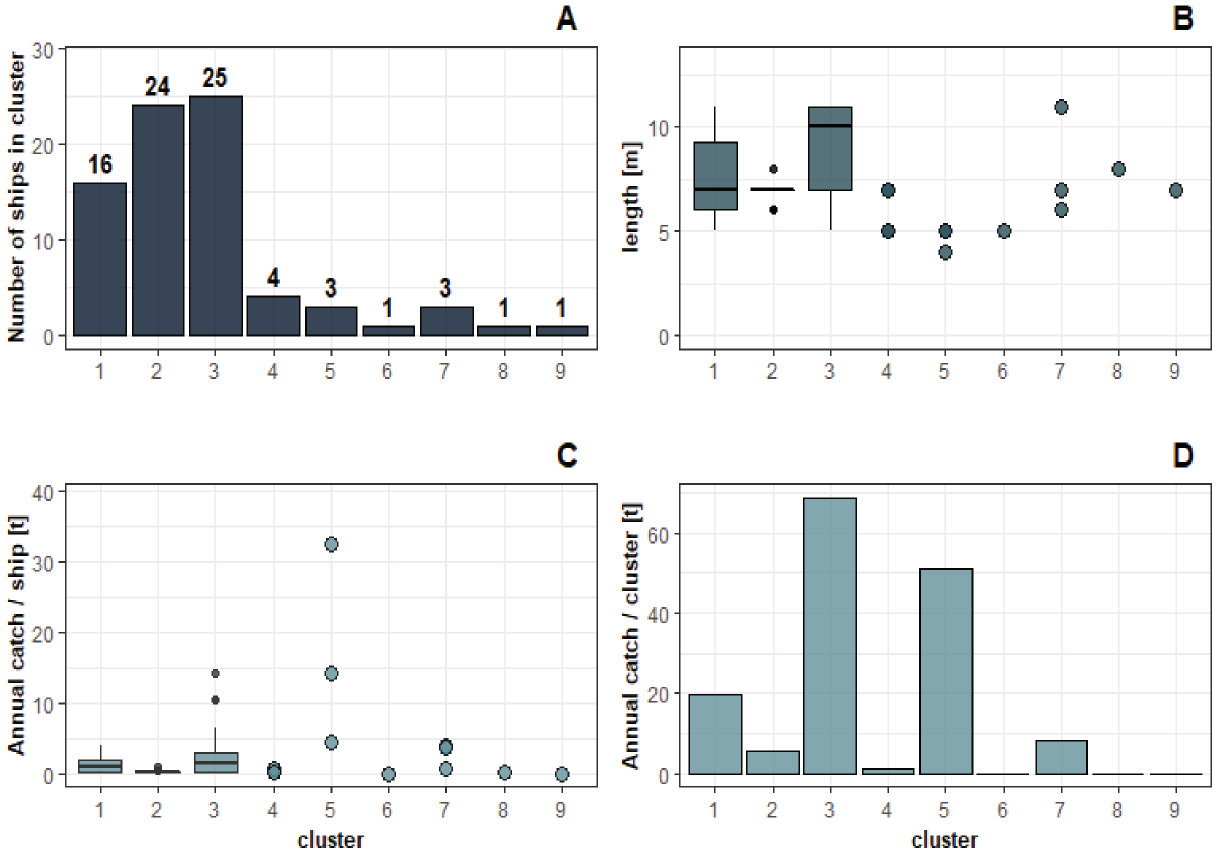


Vessel and cluster properties

- PMP

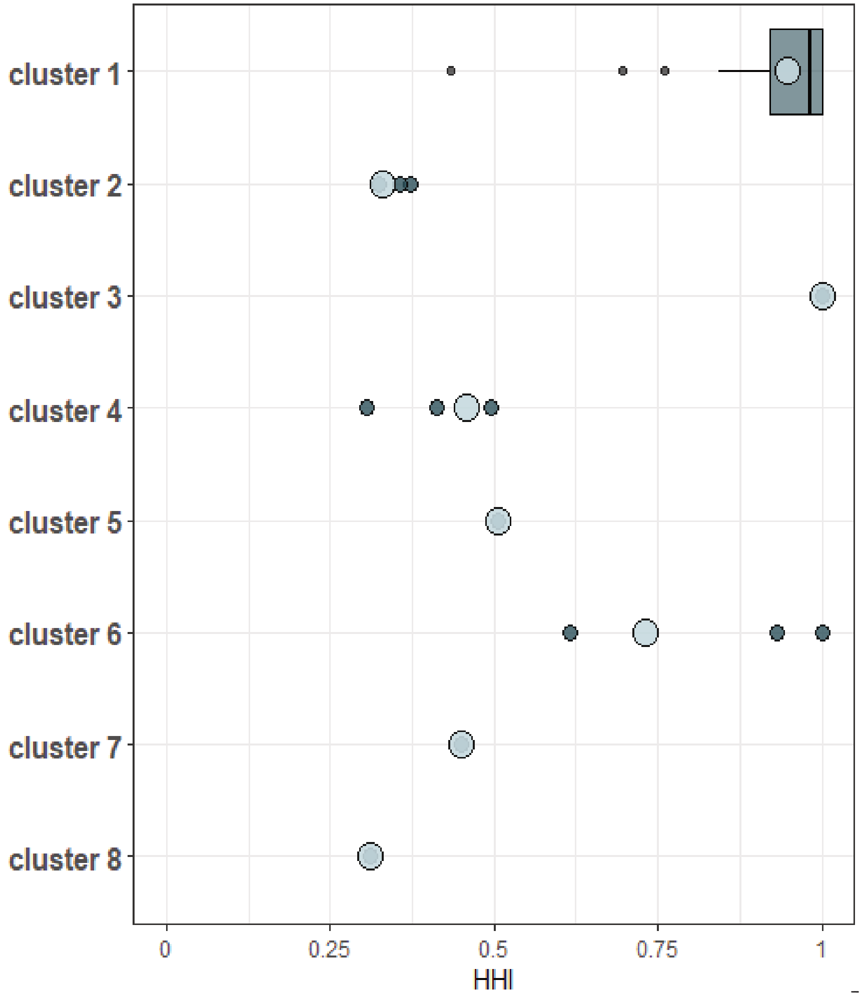


- PG

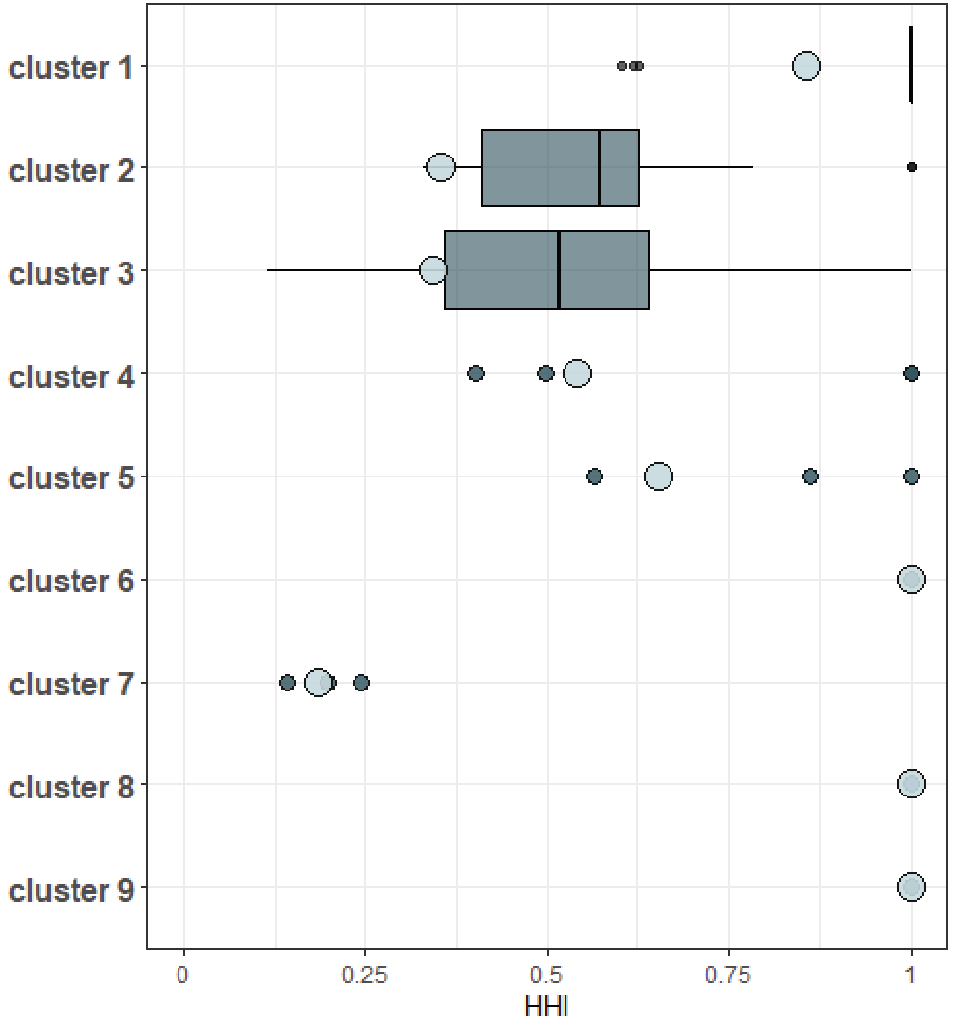


HHI (Herfindahl-Hirschman-Index) of the clusters

- PMP



- PG



5 Applications of the alternative segmentation tool on the French national fleets operating in the supra-region Atlantic*

5.1 The French fleet operating in the supra-region Atlantic

As mentioned above, the application exercise of the alternative segmentation tool is applicable only to the French fleet operating in the supra-region Atlantic (FAO area 27) based on the vessel (individual) landings data available in the SACROIS data-set (see before for details). The application exercise is neither feasible for the Mediterranean fleet operating in FAO 37 nor for the Outermost fleet operating in Indian Ocean (area 51) and Western Atlantic (FAO area 31 & 41).

In 2020, French fishing fleet operating in the supra-region Atlantic consisted of 2 900 vessels; 187 being inactive. The 2 713 active vessels presented a high variability in term of vessel length from less than 4 meters to more than 90 meters vessels (see next figure). The majority (52%) were less than 10 meters vessels (1495 vessels) when the more than 24 meters vessels represented less than 5% of the total fleet (125 vessels).

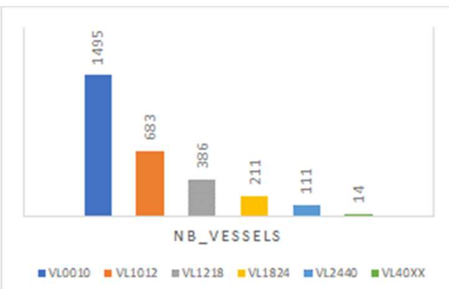


Figure 13: Number of vessels per vessel length ranges (Atlantic area 27) in 2020

In 2020, The Atlantic fleet was distributed as follows according to the EU-MAP and Ifremer-FIS segmentation.

DCF Fleet segment		**VL0010**	**VL1012**	**VL1224**	**VL2440**	**VL40XX**	**Nb_vessels**	**%**
Active gears	**Beam trawlers (TBB)**		1	1			2	0%
	Demersal trawlers and/or demersal seiners (DTS)	83	169	268	56	10	586	22%
	Pelagic trawlers (OTM)	1	5	26	1	4	37	1%
	Purse seiners (PS_)		3	27			30	1%
	Dredgers (DRB)	73	81	87	1		242	9%
	Vessels using other active gears (MGO)	171	8				179	7%
	Vessels using polyvalent active gears only (MGP)	12	56	48	5		121	4%
Passive gears	**Drift and/or fixed netters (DFN)**	309	151	93	24		577	21%
	Vessels using pots and/or traps (FPO)	296	80	19	1		396	15%
	Vessels using hooks (HOK)	216	49	3	20		288	11%
	Vessels using other passive gears (PGO)	98	4	1			103	4%
	Vessels using polyvalent passive gears only (PGP)	59	12	1			72	3%
Active/Passive gears	**Vessels using active and passive gears (PMP)**	38	37	5			80	3%
							2713	
Non Active vessels		139	27	18	3		**187**	

Table 20: Number of vessels per EU-MAP fleet segment (Atlantic area 27) in 2020

* Please note that this case study is an extract of the original Ifremer document:
Demaneche Sebastien, Guyader Olivier, La Grand Christelle, Merzereaud Mathieu, Vigneau Joel, Quentin Laurent (2022). Alternative approaches to the segmentation of EU fishing fleets. Workshop II - 28-30th March 2022. Previous experiences, tests for application in the French context and recommendations. PDG-RBE-HISSEO, PDGRBE-EM, PDG-RBE-HMMN-LRHPB.
https://doi.org.10.13155/89336

Ifremer Fleet segment		VL0010	VL1012	VL1224	VL2440	VL40XX	Nb_vessels	%
Active gears eventually combined with passive gears	Demersal trawlers exclusive	21	54	183	45	10	313	12%
	Pelagic trawlers exclusive		3	8	1	4	16	1%
	Mixed Trawlers exclusive	1	8	58	7		74	3%
	Trawlers Dredgers	47	173	135	1		356	13%
	Trawlers Glass eel fishing	30	22				52	2%
	Trawlers Passive gears	10	18	2			30	1%
	Demersal seiners		1	19	9		29	1%
	Purse seiners		4	27			31	1%
	Dredgers exclusive	43	34	29			106	4%
	Dredgers Passive gears	92	69	6			167	6%
	Glass eel fishing exclusive	86	4				90	3%
	Glass eel fishing Passive gears	165	9				174	6%
Passive gears exclusive	Netters exclusive	145	84	82	24		335	12%
	Netters Potters/Traps	204	57	9			270	10%
	Netters Hooks	63	28	3			94	3%
	Potters/Traps exclusive	156	44	14	1		215	8%
	Potters/Traps Hooks	77	16				93	3%
	Hooks exclusive	136	20	3	20		179	7%
	Other passive gears	80	8	1			89	3%
							2713	
Non Active vessels		139	27	18	3		187	

Table 21: Number of vessels per Ifremer-FIS fleet segment (Atlantic area 27) in 2020

A high diversity of the fishing gears in used by these vessels was observed which lead to a high distributed fleet by fleet segment. The Ifremer segmentation allows to assess the exclusive or non-exclusive nature of fishing strategies of the vessels highlighting that the combination of two *(or more)* fishing gears during the year is very common. This should be considered to carry out a fleet segmentation of interest.

5.2 Methodology to tests the alternative segmentation tool

Different tests of the alternative segmentation tool (clustering approach) were applied on the French fleets operating in the supra-region Atlantic based on the “Fleet Segmentation manual”[35]. In a last step, results of the clustering approach obtained by EU-MAP fleet segment were also compared with other alternative pre-existing fleet segmentations using intra vs inter variances indicators and stability indicators over the years.

Two different methods were tested to pre-segment the full dataset by:

1) EU-MAP fleet segment,
2) Ifremer fleet segment

Then the clustering approach *(R-package provided in the context of the workshop)* was tested based on the following metrics:

- Catch composition profiles in weight *(Ldgs/species*sect)* and value *(val/species*sect)*
- Total landings by “fishing gear” *(métier DCF level4)* in weight *(Ldgs/metDCF4)* and value *(val/metDCF4)*

[35] Fleet Segmentation - Package Manual. Erik Sulanke Thuenen-Institute for Sea Fisheries, Bremerhaven, Germany. https://rdrr.io/github/ESulanke/FleetSegmentation/f/vignettes/FleetSegmentation_vignette.Rmd

- Total landings by “métier” *(métier DCF level5)* in weight *(Ldgs/metDCF5)* and value *(val/metDCF5)*
- Total landings by “métier” * “ICES division” in weight *(Ldgs/metDCF5*sect)* and value *(val/metDCF5*sect)*

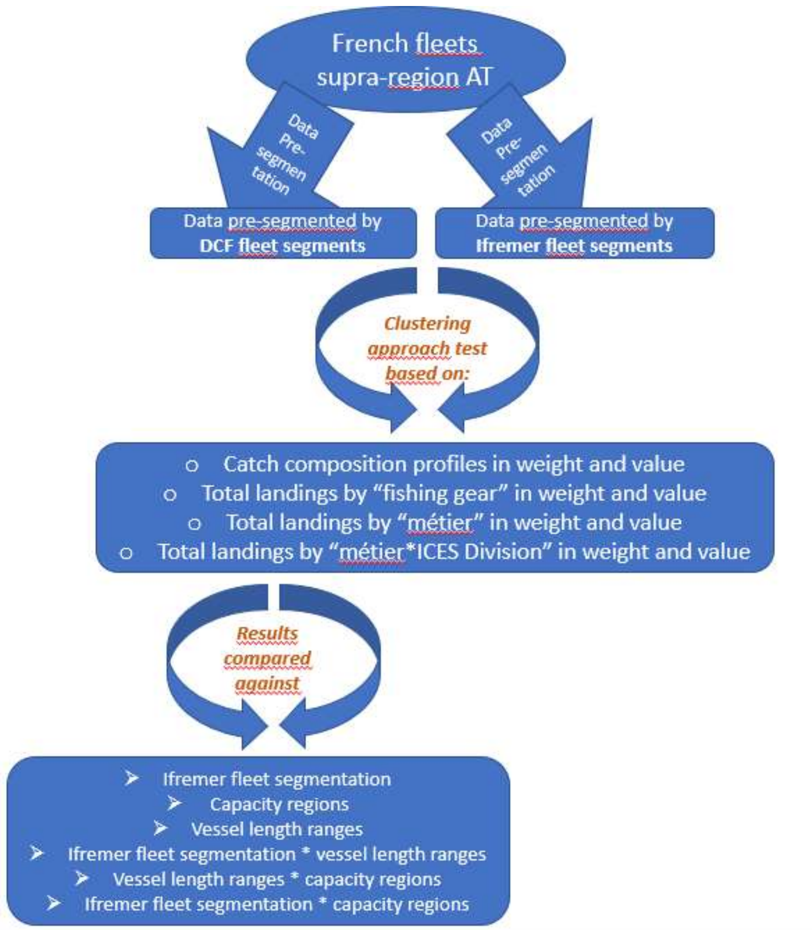


Figure 34: Methodology used to compare different segmentation and different metrics

Finally, the clustering approaches results obtained by EU-MAP fleet segment were compared with the following alternative pre-existing fleet segmentations:

- Ifremer-FIS fleet segmentation *(FLEET_IFR)*,
- Capacity regions of the vessels *(REG_CAP)*,
- Vessel length ranges *(VSL_LGTH)*,
- Ifremer-FIS fleet segmentation * vessel length ranges *(VSL_LGTH/FLEET_IFR)*,
- Vessel length ranges * capacity regions *(VSL_LGTH/REG_CAP)*,
- Ifremer-FIS fleet segmentation * capacity regions * vessel length ranges *(VSL_LGTH/REG_CAP/FLT_IFR)*

At the end, 14 different fleet segmentations were compared.

5.2.1 "Demersal trawlers and/or demersal seiners (DTS)" EU-MAP fleet segment, metric = catch composition profiles in weight

First application exercise of the alternative fleet segmentation tool was applied by EU-MAP fleet segment on catch composition profiles in weight *(default approach proposed by the tool)*. The result obtained for "Demersal trawlers and seiners (DTS)" EU-MAP fleet segment is briefly presented hereafter. The results of the other EU-MAP fleet segments are available in *Annex I*.

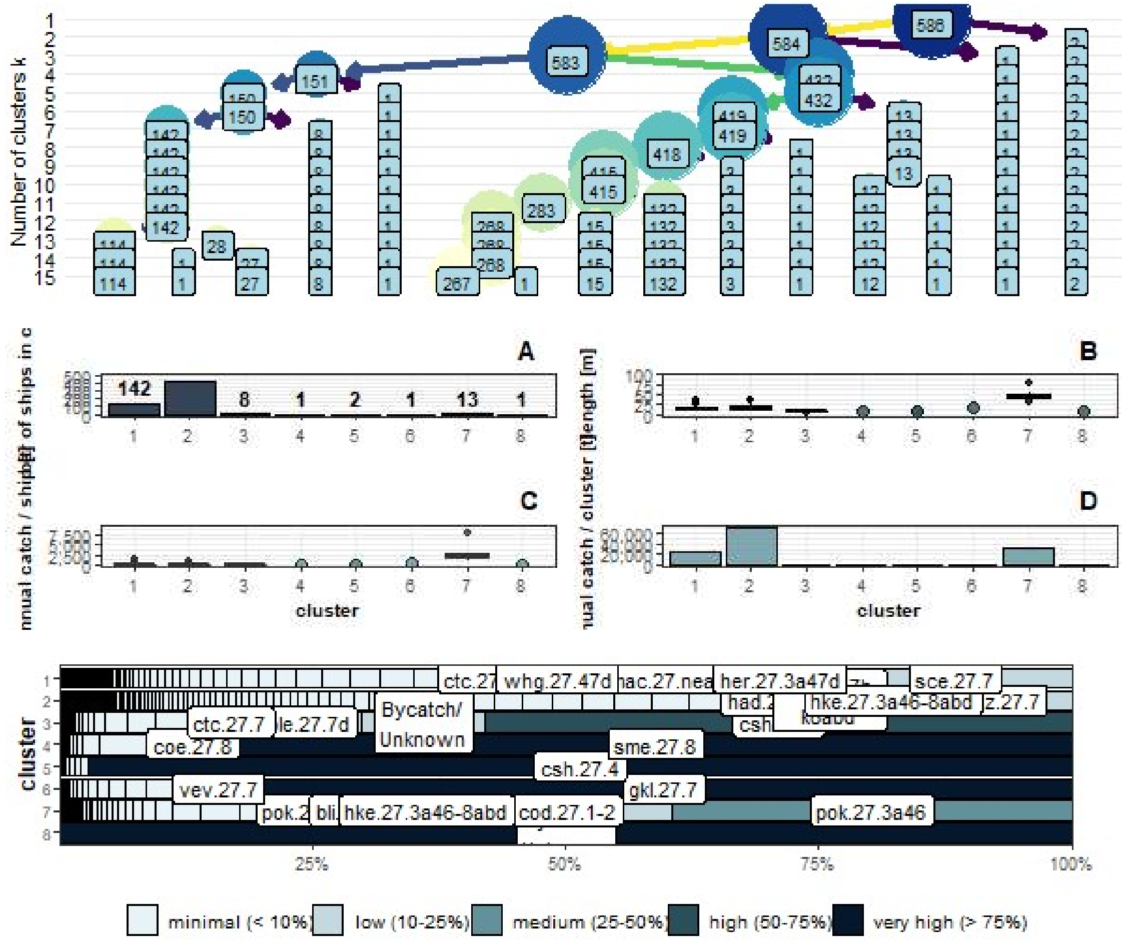


Figure 22: Results for the Demersal trawlers and/or demersal seiners (DTS) EU-MAP segment, metric = catch composition profiles in weight.

A first conclusion of this exercise is that the alternative fleet segmentation tool seems not well adapted to the specificity and the diversity of the French fishing fleets. One of the major issues from this application is that the tool tends to highlight some very specific/specialized vessels designing fishing segments with less than 5 to 10 vessels and keeping the majority of the other fishing vessels in 2 to 3 large diverse groups where the principal stocks landed are grouped. The segmentation carried out with the proposed clustering R-package failed to achieve the objective.

This seems to be linked to the high diversity observed in the different French EU-MAP fleet segments. Pre-segmentation of the data before applying the approach is therefore a key issue to consider. To try to avoid the issue linked with the polyvalence of the vessels belonging to the same EU-MAP fleet segment, same application exercise was carried out on the pre-segmented data set by Ifremer fleet segment *(segmentation based on gear or combination of gears practiced)*.

5.2.2 "Demersal trawlers exclusive" Ifremer-FIS fleet segment. metric = catch composition profiles in weight

The results achieved for the application of the alternative fleet segmentation tool by Ifremer-FIS fleet segment for "Demersal trawlers exclusive" Ifremer-FIS fleet segment is briefly presented hereafter, other segments could be found in *Annex II*.

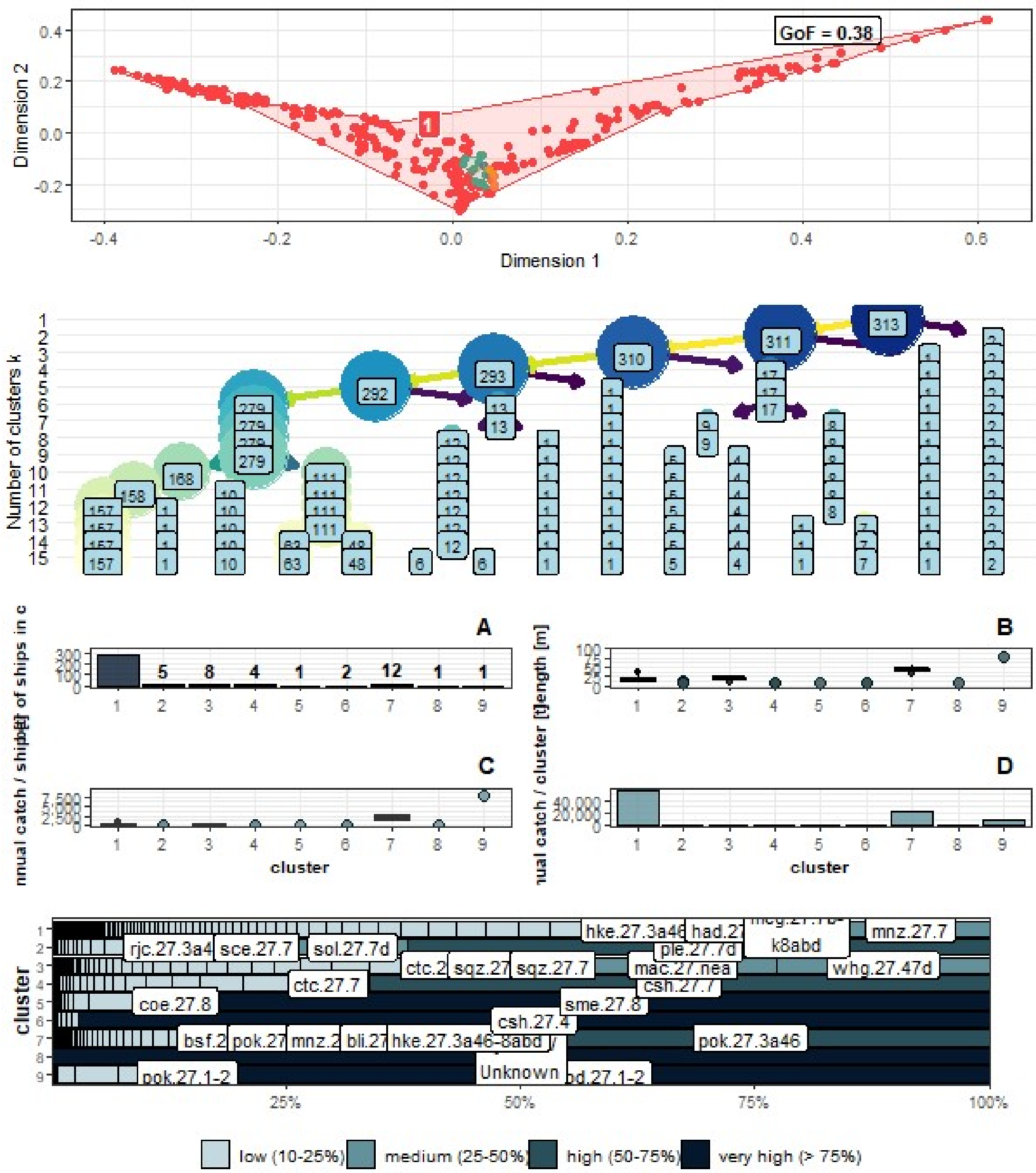


Figure 23: Results for Demersal trawlers exclusive Ifremer-FIS segment, metric = catch composition profiles in weight.

The application exercise gives better results on Ifremer-FIS Fleet segmentations *(especially for specialized vessels e.g. "Demersal trawlers exclusive" or "Netters exclusive")* but the tool seems to have still difficulties to identify a segmentation adapted to the polyvalent/diversified vessels considered and tends to continue to group specific/specialized vessels into small groups and to keep other fishing vessels in 2 to 3 large diverse groups.

This issue should possibly be linked to the technical statistical parameters considered in the tool as the "distance" or the "segmentation method" *(e.g. hierarchical agglomerative cluster analysis (HAC))*. Maybe should be valuable to propose in the tool an alternative choice 1) for the distance as a "denormalized distance", 2) for the segmentation method as a "k-means clustering method" or 3) to parameter a "minimum cluster size control". Furthermore, it is not obvious if the classification tool considers the "absolute value" or recalculate the data in "percentage". The two different possibilities should be possibly allowed and tested.

5.2.3 "Demersal trawlers exclusive" (Ifremer-FIS-segment)- catch composition profiles in value

The metric to perform the segmentation could be questioned especially considering the landings weight vs the landings value. To classify the vessels into fleet segment, like to define the metier and for the same reasons, it seems that value landed should be a better metric to consider. Actually, same considerations apply that the ones approved in the DCF WK Métier Workshop[36]:

> *"However, it is the recommendation of this group that if target assemblage is defined as describing the fisher intent then value is the metric that should be used, as fisheries are conducted for economic gain. Likewise, when species with a low weight relative the value is the real target, then value is a better metric. Finally, the use of value as the metric for target assemblage would help to avoid the complication created by the implementation of the landings obligation, where potentially large weights of low economic value could affect any post classification system based solely on weight, resulting in incorrect definition of fishing intention. Despite this, there might be some cases where a combination of value and weight should be used. For example, purse seiners targeting small pelagic fish can catch a school of the target species but if some other valuable species are caught in less weight the output of the trip can be conditioned by the more valuable species although it was not the original target. A combination of the two criteria should be used in these cases."*

In order to test this assumption same application exercise was applied on the French fishing fleet considering the total value landed by species/stocks rather than the weight. See hereafter, an example of the results obtained on the Ifremer fleet segment *"Demersal trawlers exclusive"* using the value to be compared with the previous plot presenting the results for the same vessels with the metric in weight.

[36] Anonymous report: DCF Métier Workshop: Sub-group of the RCGs - North Sea and Eastern Arctic and North Atlantic. 22 - 26 January 2018. DTU Aqua, Lyngby, Denmark. https://webgate.ec.europa.eu/regdel/web/meetings/507/documents/1697

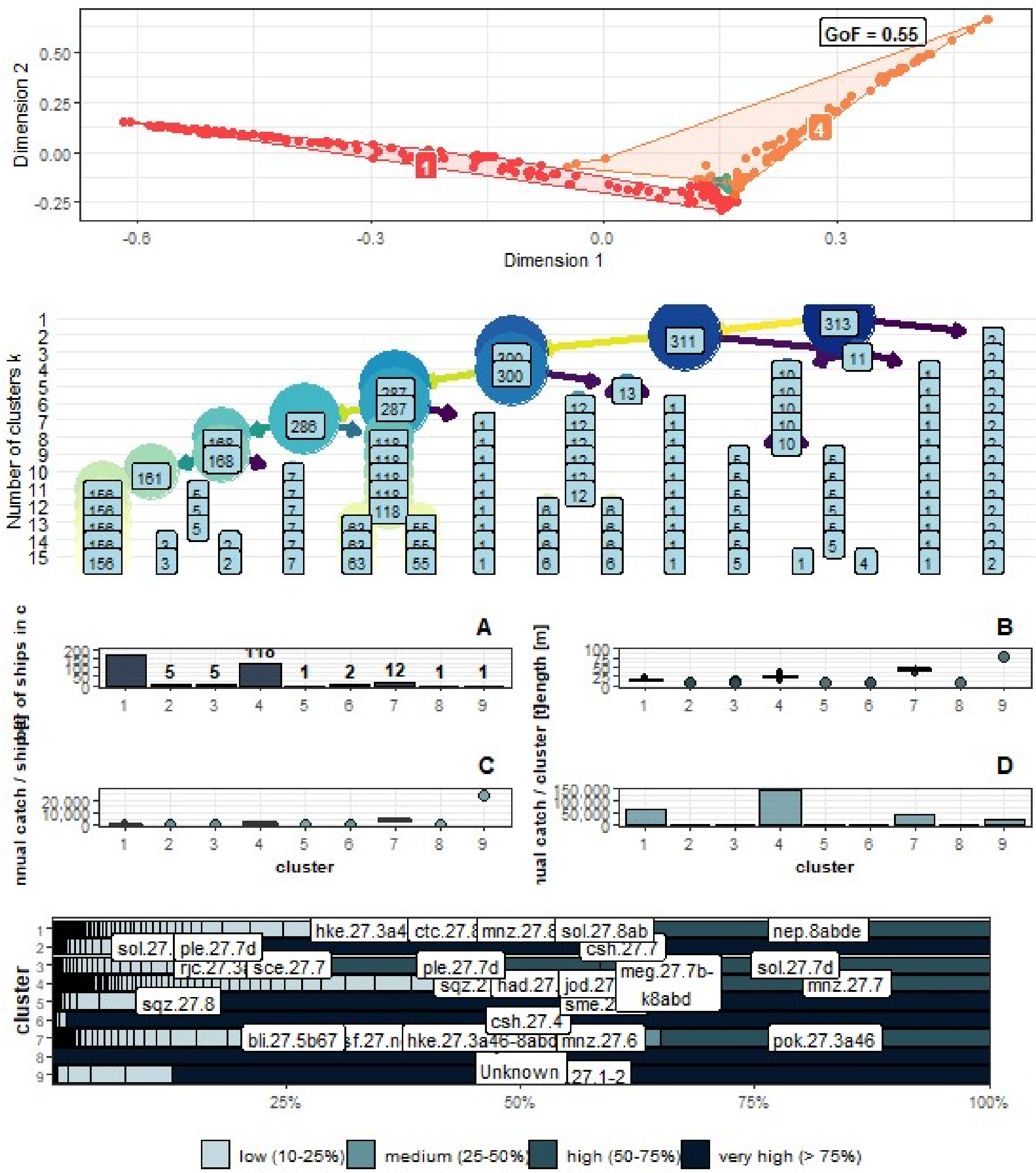


Figure 24: Results for Demersal trawlers exclusive Ifremer-FIS segment, metric = catch composition profiles in value.

Considering the value of species landed allow to better segment the fleet especially into two different 'big' groups and presents a better GoF (0.55 vs 0.38).

However, the two principal groups defined remain relatively big *(168 and 118 vessels, groups 1 & 4)* and the tool still seems to focus on very specific/specialized vessels regrouping them in small groups *(groups 2, 3, 5, 6, 7, 8 & 9)*. For example, the group 9 concerns only one vessel with landings declared in the 27.1.2 which is very specific when the group 1 aggregate 168 vessels with important landings of nephrops (NEP), sole (SOL), anglerfish (MNZ), hake (HKE) and cuttlefish (CTC); all of them being "structuring species" for the exclusive Demersal trawlers fleet operating in the supra-region Atlantic.

5.2.4 *Catch composition in weight vs total landings in value by fishing gear (metier DCF level4) for polyvalent fleets (example of Netters Potters/Traps)*

In order to develop a proposal flowchart to be applied to segment the fishing fleets. Another test was done on the “Netters Potters/Traps” Ifremer fleet segment *(polyvalent fleet)* comparing the results obtained by the tool directly *(based on catch composition)* from another approach considering the vessel’ total landings in value by fishing gear practiced during the year *(in order to better take into consideration the polyvalent nature of the vessels considered)*.

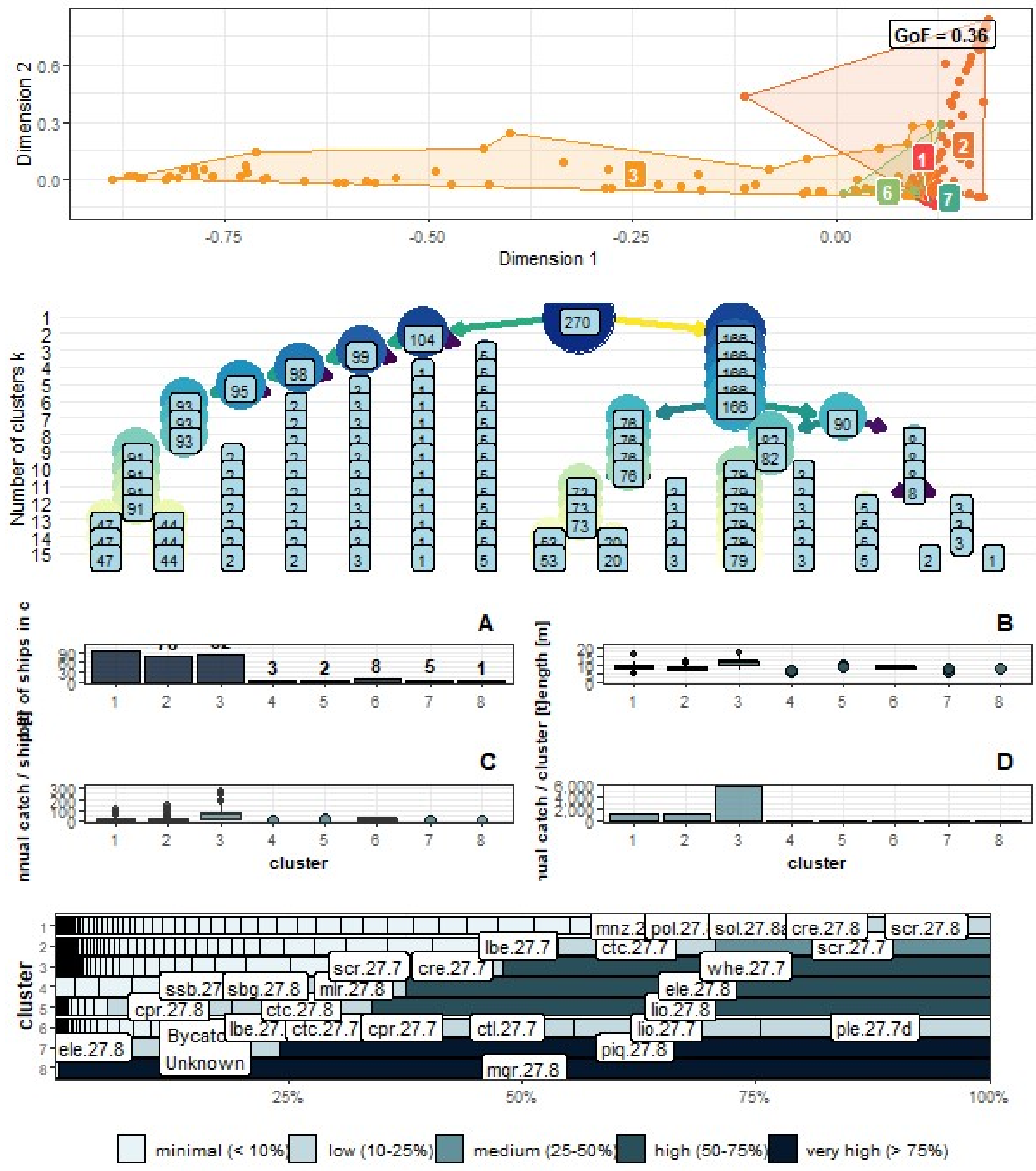


Figure 25: Results for Netters Potters/Traps – Metric: catch composition profile in weight.

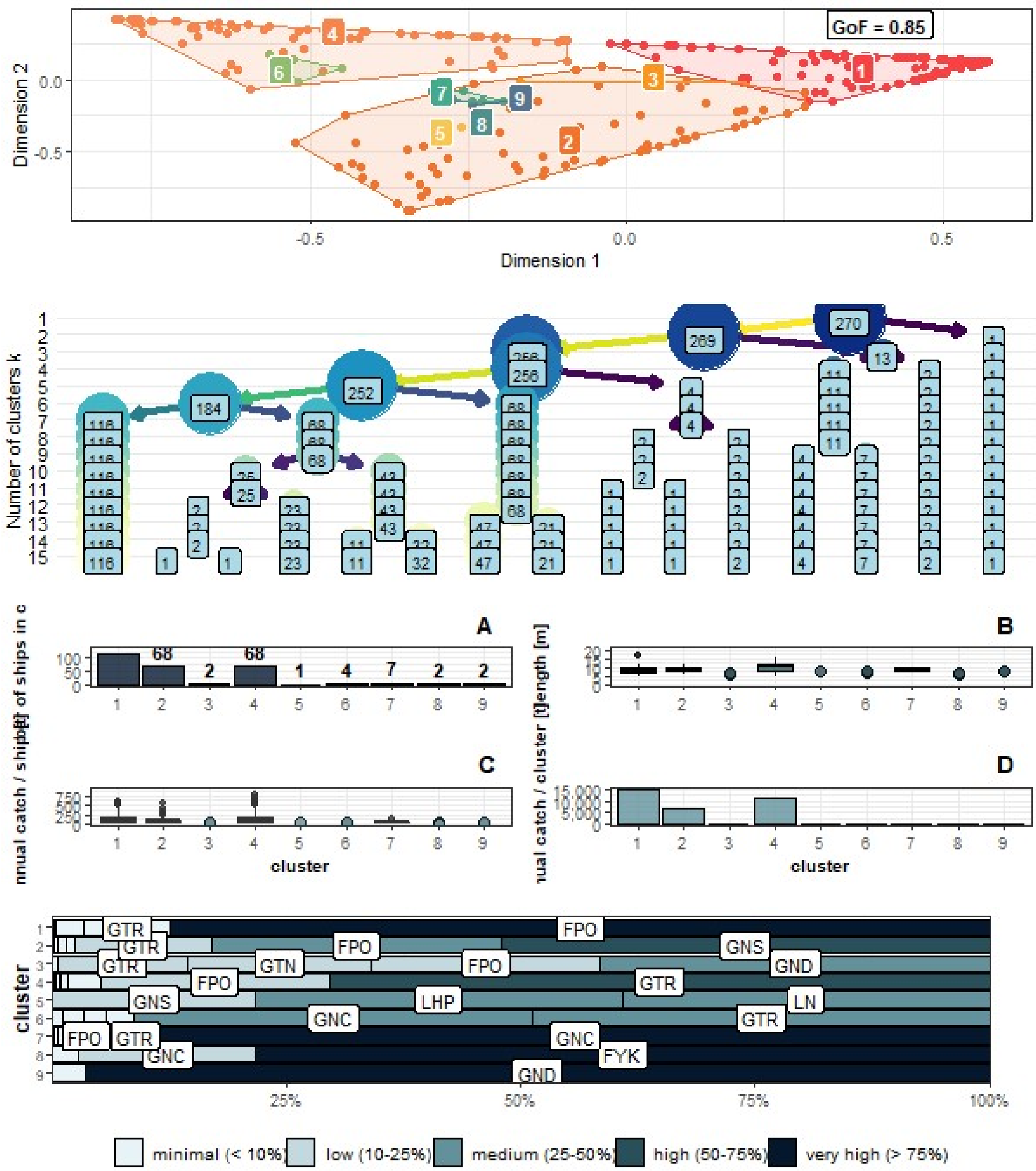


Figure 26: Results for Netters Potters/Traps – Metric: total landings in value by fishing gear (metier DCF level4)

Although the groups are similarly balanced in both analysis with three big diverse groups constituted and other groups being relatively small; it seems that the "big" diverse group are more heterogeneous in the first process *(see dimension1 * dimension 2 maps)*. Therefore, a first step based on the combination of gears used by the vessels seems to better structure the fleet *(pre-segment the dataset)* before getting one step further regarding the group of species targeted *(i.e. the metiers practiced during the year)* and finally the species/stocks composition. For example, here it should be useful to distinguish vessel combining *"nets and pots metiers"* vs *"nets and fike nets (traps)"*.

5.2.5 "Demersal trawlers exclusive" Ifremer-FIS segment, metric catch composition in weight vs total landings in value by "métier" (metier DCF level5)

In the same way and also to test the proposal flowchart presented hereunder, a test was applied on the "Demersal trawlers exclusive" Ifremer-FIS fleet segment *(exclusive fleet)* comparing the results obtained by the tool directly *(based on catch composition)* from another approach considering the vessel' total landings in value by métier DCF level5 during the year.

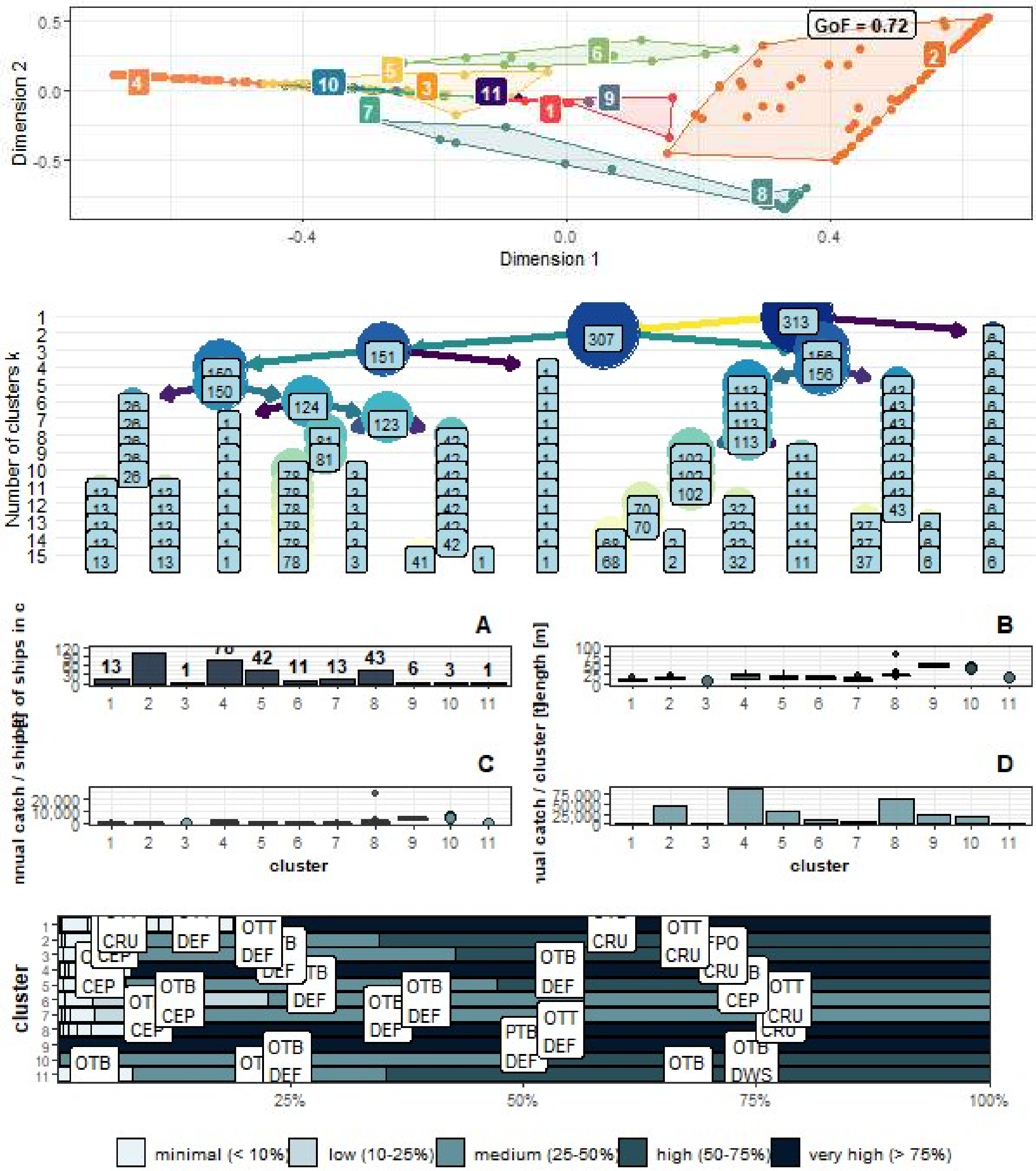


Figure 27: Results for Demersal trawlers exclusive Ifremer-FIS segment, metric = total landings in value by métier DCF level5

The tool *(even if issue stated before remains)* applied on the basis of total landings in value by metier DCF level5 *(i.e. by group of species targeted)* allow here to better divide the fleet into groups more balanced than the groups obtained applying the tool on catch composition profile. For example, considering the metiers for this exclusive fleet allow to divide the *"exclusive demersal trawlers"* between *"exclusive demersal trawlers targeting crustaceans", "targeting cephalopods"* or *"targeting demersal fishes"*. This highlight that further steps should be then required to segment the fleets regarding the species/stocks' catch profile composition i.e. taking into consideration the different métiers operating by the vessels during the years *(their operating strategy)*.

Indeed, métiers *(regrouping fishing activity based on gear type, mesh size & target species/fish stocks)* have been defined to picture the fishing strategies of the vessels and regroup fishing trips according to similar exploitation patterns. See following conclusions stated during the DCF WK Métier Workshop[37]:

"Recently, the recast of the EU-MAP Regulation reaffirms the métier as an important domain of interest. Today fleet and métiers are commonly employed in European fisheries to form the building blocks which describe the heterogeneity of fishing activity in both biological and economic terms. These building blocks allow the partitioning of landings and effort into 'sensible' sized units representing the fishing activities within them (ICES, 2003). The functionality of métiers is evident in the number of groups (i.e. DCF, ICES, RCG, GFCM, RFMO, ...) who now use them for a variety of programs, such as the pre or post stratification/aggregation of national sampling programs, bio-economic modelling (e.g. Ulrich, Reeves, Vermard, Holmes, & Vanhee, 2011) and management strategy evaluations (e.g. Vermard et al., 2008). Ultimately, well-defined métiers provide the building blocks of more effective management (Davie & Lordan, 2011) and constitute a potent tool to improve biological and bio-economic expertise, to move towards an ecosystem-based approach and to better estimate PETS bycatch data. The use of métiers makes it possible to describe the fishing behaviour/fishing practices of fishermen and constitute a sound basis for the typological classifications of vessels by fleet segment, which forms the basis of economic data collection."

Furthermore, the matrix *"metier*fleet"* developed since the inception of the DCF has been defined to consider the fact that fishing activities on a yearly basis *(vessel' operating strategy)* affects the economic performance and the fishing activity at the trip level defines the exploitation pattern to sample, the metiers making the connection between the economic and biologic parts. Finally, métier were harmonized between countries in order that one metier can be used in a region to describe the same types of fishing activities across nations which reinforce the importance and the needs to consider the métiers and eventual combination of in the process of fleet segmentation.

[37] Anonymous report: DCF Métier Workshop: Sub-group of the RCGs - North Sea and Eastern Arctic and North Atlantic. 22 - 26 January 2018. DTU Aqua, Lyngby, Denmark. https://webgate.ec.europa.eu/regdel/web/meetings/507/documents/1697

5.3 Comparison of results achieved through the clustering approaches based on different metrics with the alternative pre-existing fleet segmentations

One of the main objectives of the fleet segmentation is to build homogeneous groups of vessels stable over the years to improve the accuracy and precision of the calculated estimates especially in terms of revenue and cost structure and other related indicators. To test and compare the results obtained from the tool *(clustering approach based on different metrics in weight and value)* against other alternative pre-existing fleet segmentation, inter-stratum variance of key indicators were calculated and analysed. The problem of small size clusters was also considered because of the confidentiality issues at the stage of the estimate's restitution. Finally, the issue of the stability of the clusters was considered in a second step.

5.3.1 Inter and intra variance for different segmentations and metrics

Following graphical outputs were edited by EU-MAP fleet segment *(example for the EU-MAP "Drift and/or fixed netters (DFN)" fleet segment is presented hereafter; other results are available in Annex III).* Each column of the graph illustrates one of the tested segmentations, and each row refers to a particular indicator. The first row presents an indicator assessing the importance of small-size clusters in the result obtained. It describes the number of clusters aggregating only one vessel, aggregating 2 to 4 vessels or aggregating more than 5 vessels. The second row completes the first one presenting the number of vessels allocated by groups with less than 5 vessels.

The other rows present the inter-stratum *(green)* and intra-stratum *(red)* variances calculated by fleet segmentation for some key indicators-metrics: fishing days, days at sea, hours at sea, total landed weights and total landed values[38]. One of the principals aims of a fleet segmentation being to maximize the inter-stratum variance and minimize the intra-stratum variance of economic indicators.

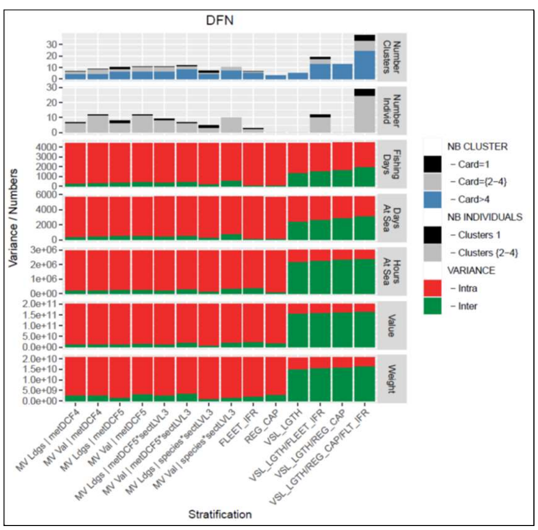


Figure 28: DFN EU-MAP Fleet segment – Comparison inter/intra stratum and small size clusters between results of alternative segmentation tools and pre-existing fleet segmentations

[38] At this stage, cost indicators were not included in the analysis.

Regarding the inter/intra-stratum variance analysis, the graph highlights the importance to first segment the dataset by vessel length ranges which concentrate a lot of the variability *(strong contribution to the inter-stratum variance)* observed between vessels and present better results than all the other segmentations tested, whatever the variable considered. At this stage, the clustering approach *(whatever the metrics considered)* failed to propose a fleet segmentation which explain more variability than first separate vessels by vessel length ranges.

Following that, pre-segment the data set by vessel length ranges could be a way to produce better results and to improve the homogeneity of the fleets segments obtained. Same analysis was carried out by EU-MAP fleet segment * vessels length ranges. Example for the DFN EU-MAP Fleet segment and the VL1012 vessel length range follows:

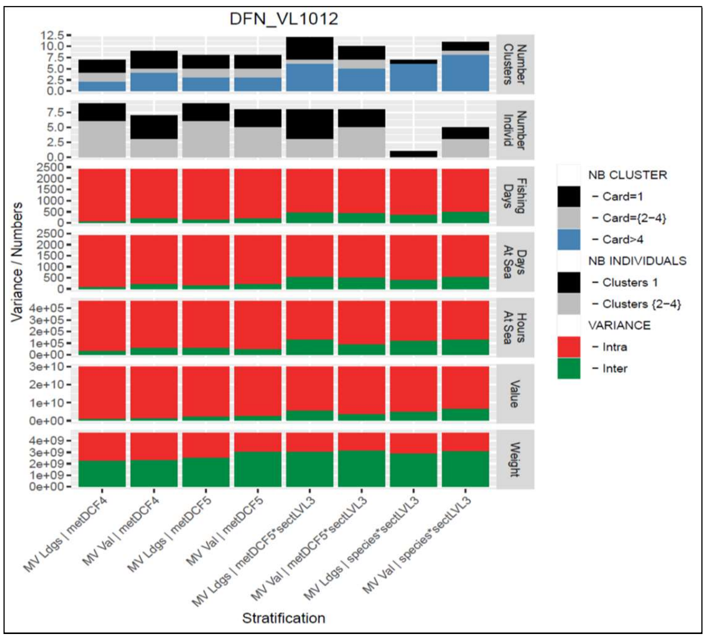


Figure 29: DFN EU-MAP Fleet segment – VL1012 – Comparison inter/intra stratum and small size clusters between results of alternative segmentation tools and pre-existing fleet segmentations

As expected, pre-segmenting the EU-MAP fleet segment first by vessel length ranges allow to decrease the total variance i.e. the process benefit from this preliminary stratification mitigating the negative effects of too much overall heterogeneity in the population considered. This should be regarded also considering the possible threshold issue linked with the predefine vessels length limit to define vessels length ranges and also the usefulness to aggregate vessels from different vessel length ranges presenting similar fishing strategies.

5.3.2 Stability of the clustering approaches

Another issue of the approach is the requirement to compare fleet segments across years. The following graph compare the results obtained for two different years (2018 vs 2019) on the same fleet segment and highlight the instability of the results obtained year to year. The graph presents information about the stability of the results obtained in the two years from the different tested segmentations i.e. segmentation provided by the tool *(clustering approach based on different* *metrics in weight and value)* and pre-existing fleet segmentation. Each square present a tested segmentation and compare the results obtained in the two years (2018 & 2019).

For DFN EU-MAP fleet segment (presented as an example in figure 30 hereunder), it highlights a high instability when using the clustering approach which is less the case regarding the pre-existing fleet segmentation. Similar results have been observed for the other EU-MAP fleet segments. It concludes that clustering approaches are a good statistical mean to analyse/explore the fishing fleets studied but it is crucial from their results to define stabilized set of rules which could be applied all along the period, year after year.

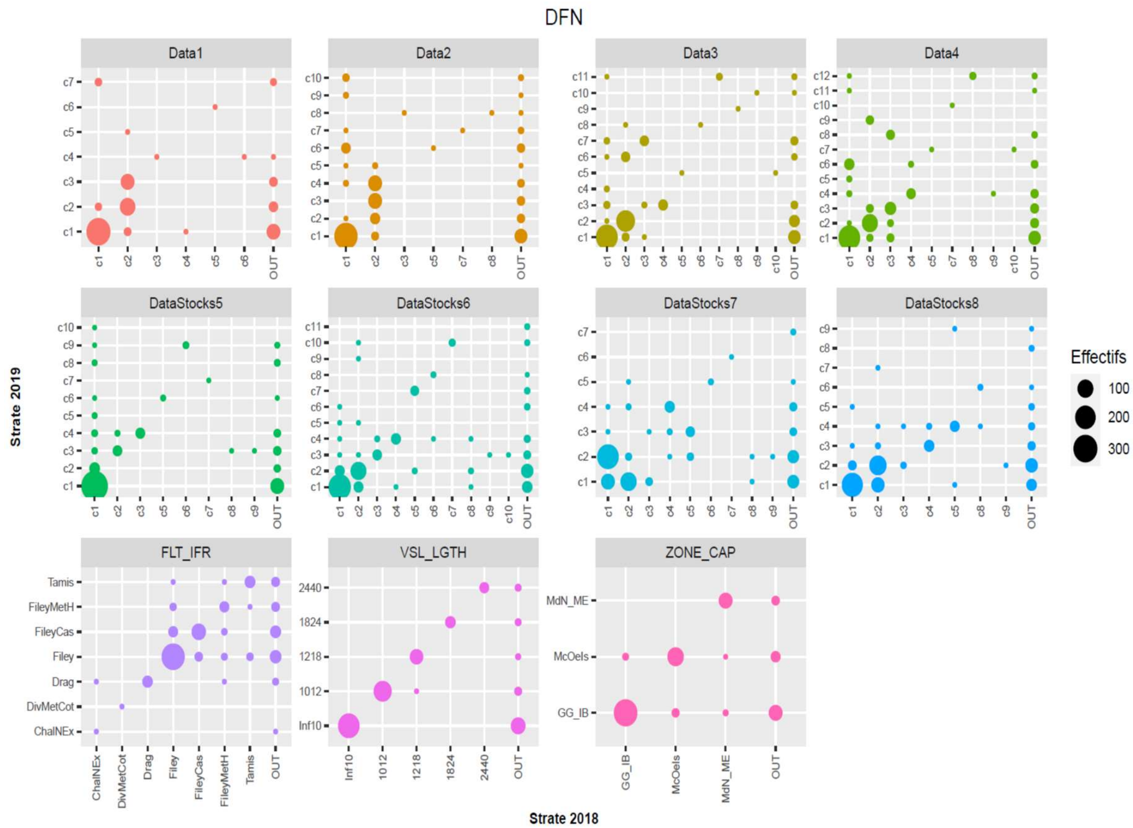


Figure 30: DFN EU-MAP Fleet segment – Comparison of the results obtained in 2018 and 2019 for the different tested fleet segmentations. Indicator of stability/instability of the results obtained.

5.4 Conclusions and recommendations

Based on the application of the proposed clustering R-Package to the French fleet operating in the supra region Atlantic, our first conclusion is that the use of Principal Component Analysis and clustering approaches is not appropriate to define fleet segmentation. If the tool is very interesting for a preliminary understanding of the fishing fleets and activities which are complex by nature, one of the issues with PCA analyses is that it is difficult to control how the groups are formed. Moreover, PCA analysis do not produce stable groups/segments over time and across countries and may even change historical perspectives upon addition of new years in the dataset. PCA results can also lead to the definition of groups that are often too large or too small, which is also a pitfall to be avoided *(small groups)* for statistical and confidential reasons.

The metric proposed by the tool is the "Catch composition profile in weight" because it is supposed *"to better represent the fishing strategies of the vessels, the stocks used and how mixed or targeted a fishery is"*. Based on our analysis, the landings in value per species or stock seems to be a better metric than weight for the majority of the fleets. Reasons for that are similar as what it was approved in the DCF WK on metier issues. Based on our results, it seems also crucial to better consider the polyvalent/non-exclusive nature of the fleets in terms of fishing gears and métiers. The specific methodology we developed for the analysis of inter/intra-stratum variance also highlights the importance to first segment the dataset by vessel length ranges which concentrate a lot of the variability. Indeed, vessel length ranges present better results than all the other segmentations tested, whatever the variable considered.

A flowchart proposal for the segmentation of the EU fleet[39]

The following flowchart tries to synthetize the step-by-step approach proposed with the objective to define a set of agreed and objective rules for the improvement of the EU fleet segmentation.

1. Necessary criteria for fleet segmentation

First of all, it is crucial to consider that necessary criteria for fleet segmentation should be i) stabilized and easily replicable over time and ii) harmonized and standardized between member states and fisheries ecoregion. Any segmentation should be stable: This means that segmentation rules cannot be changed every year or too regularly. Obviously, if the segments change regularly, the basis for calculating indicators evolves over time and it is therefore not possible to monitor the economic performance or other indicators related to the vessels and fleets over time and across countries. Furthermore, even if specific fishing activities may be operated in each member state, the same easily identified set rules must be applied in each MS or/and each fishing ecoregion for different member states. Moreover, the segmentation should not be too fine. Indeed, when the segmentation is too fine, vessels can migrate from one fleet/segment to another too easily even with minor changes in their fishing and production strategy. This can result in an instability of the groups, which is also not desirable for the analysis of series. Another consideration is the compatibility with previous DCF time series.

[39] Complementary to fleet segmentation, it is fundamental to keep in mind the "metier * fleet" matrix which gives the possibility to connect vessels to species and stocks through metiers.

Finally, it is imperative on the one hand to respect the rules relating to confidentiality *(for example at least 5 vessels per segment)* and on the other hand to have segments of sufficiently large size to be able to get economic samples of acceptable *size (in other terms limiting also the number of segments)*.

2. Design of the set of rules

There are different options as soon as step 0: Either to reconsider completely the fleet segmentation at global EU fleet level or to develop a new sub-segmentation of the ongoing EU-MAP fleet segments. Vessels considered could also be regionalized by fishing ecoregion (e.g. Bay of Biscay and Iberian waters, North Sea and Eastern Channel, …). Before any further investigation, vessel length ranges, as key parameter, should be considered and discussed for improvement.

Then, whatever the option adopted, the first step (step 1) is to segment the fleet by fishing gear or combination of fishing gears used by the vessels, then (step 2) by metier DCF level5 *(i.e. principal group of species targeted)* or combination of and then (step 3) by catch composition in value by species/stocks. The benefit of such an approach would be to better consider the different dependencies to species and contribution of fishing mortality to fish stocks as well as the polyvalence of vessels. Based on this analysis, the set of rules need then to be codified to be easily replicable each year *(threshold to be developed)*.

All of that, lead to consider the following fleet segmentation flowchart proposal. It takes first (step1) into consideration the fishing gears or combination of, used by the vessels (exclusive or polyvalent vessels), separating vessels by their exclusive (e.g. exclusive trawlers, exclusive netters, …) or non-exclusive/polyvalent nature (e.g. trawlers-dredgers, netters-potters, …). Next step (step2) considers the metiers or combination of metiers practiced (i.e. the principal target species or combination of targeted) for example separating exclusive trawlers vessels between exclusive pelagic trawlers, exclusive demersal trawlers or mixed exclusive trawlers targeting demersal and pelagic fishes. Finally, in a third round (step3) some specificities regarding the catch composition in species/stocks of the vessels considered could be highlighted for example separating exclusive demersal trawlers between Nephrops specialized exclusive demersal trawlers and non-specialized exclusive demersal trawlers. This last step allows to better define/divide the groups established. This should be associate with considering the vessel length ranges and fishing areas. In this method, the alternative fleet segmentation tool developed will be useful as a statistical mean to analyse the dataset and define the set of rules to be applied in application of the flowchart.

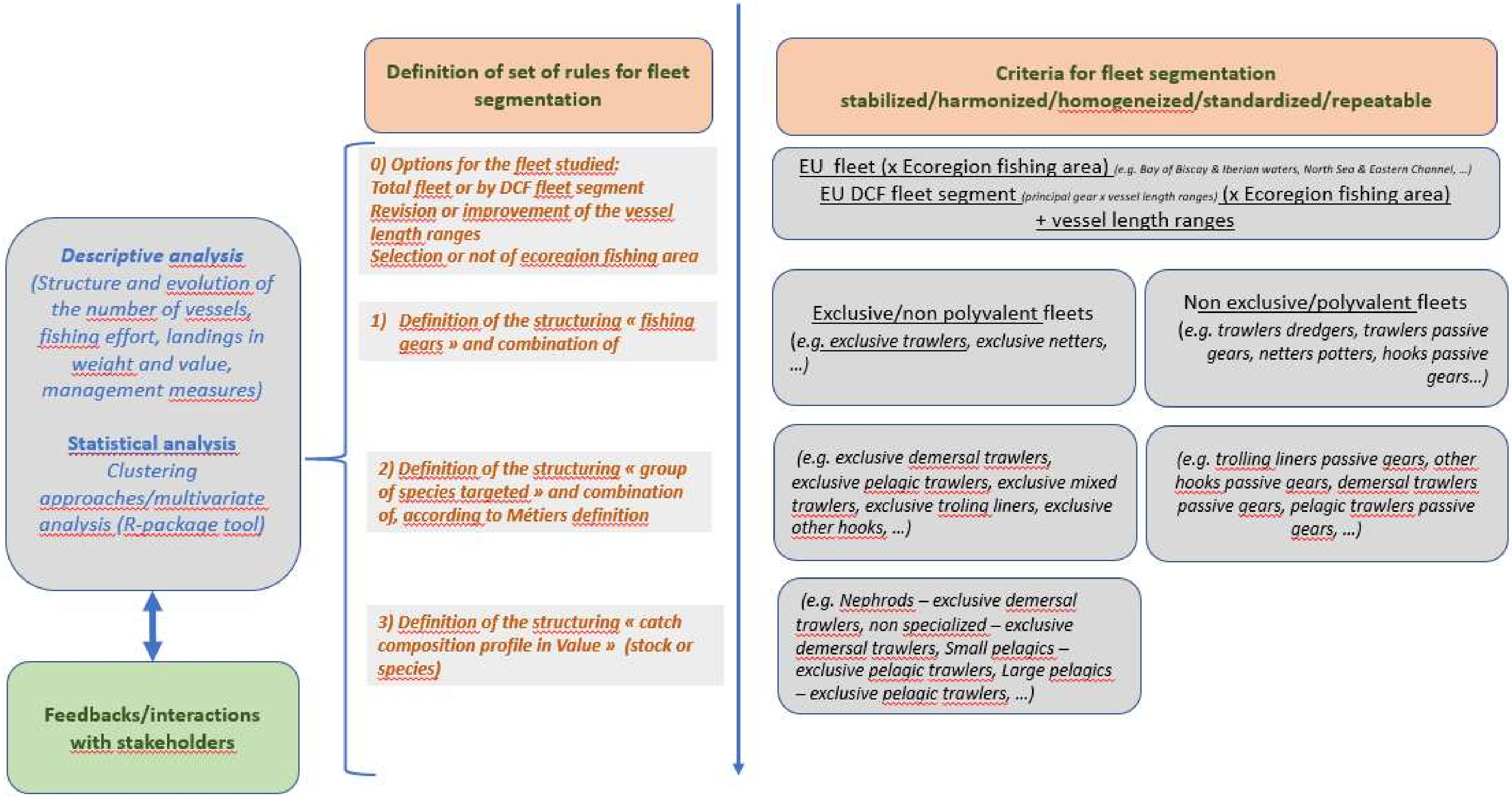


Figure 31 Flowchart proposal to define a set of rules to segment the EU fleet

3. Methodology

To define the set of rules, preliminary analyses of fishing fleets *(by length category including the evolution of the number of vessels, fishing effort and landings by species in weight and value)* including the regulatory contexts should be developed[40]. PCA tools and clustering approaches are very interesting for a comprehension of the fleets and fishing activities which are complex by nature. Based on these different approaches, the results should be translated into stabilized/standardized and harmonized set of decision rules shared between member state, easily reproducible year by year in order a vessel will be allocated to one fleet segment in the same way in each MS for the fishing ecoregion considered**.** Exchange and discussion with stakeholders could be also useful at this stage.

4. Data availability issues and small-scale fleets

As mentioned above, the application exercise of the alternative segmentation tool is only applicable to fleets where vessel (individual) landings data (landings per species and stocks) are available. However, the lack and incompleteness of reliable data at vessel level has been reported in many contexts especially for small-scale fleet. This situation may jeopardise the capacity to carry out alternative segmentation approaches but there is no valid reason not to apply an alternative segmentation as these fleets are economically and socially important and may also be affected by management measures or more broadly by management plans. Because small-scale fleets present regularly a greater diversity in term of fishing gears used than the large-scale fleets, it considered as inadequate to allocate them into one unique heterogeneous PGP (*Vessels using polyvalent "passive" gears only)* Fleet segment. For these fleet, complementary data collection scheme as the vessel fishing activity calendar census survey (VFACCS) is considered as an appropriate approach as soon as declarative data are assessed as incomplete or insufficient to meet the end-users needs.

[40] See the fisheries overviews as a first step to follow: https://www.ices.dk/advice/Fisheries-overviews/Pages/fisheries-overviews.aspx

Thünen Institute of Sea Fisheries, Bremerhaven, Germany
30.03.2021

Annex 5.5. Case study - Germany

Applying the alternative fleet segmentation approach to the German fishing fleet

Erik Sulanke & Jörg Berkenhagen

The newly developed, alternative fleet segmentation approach is based on the catch composition of the considered fishing vessels within one reference year. We chose data of the German fishing fleet from the reporting year 2021 as the basis of our analysis. The specific indices, tests, and visual validation methods of the package were applied to determine the optimal number of clusters. The procedure was finalized by a post-hoc validation of the clustering result to identify the actual fleet segments. Fifteen fleet segments were created from the underlying data, representing 919 vessels of seven different gear classes. This result represents an effective reduction of fleet segments compared to the twenty DCF segments assigned to the fleet (see Figure 1). The vessels ranged from 4m to 140m in length and operated in various management areas, from Baltic inshore fisheries to distant overseas territories. We detected mixed fisheries, especially on diverse demersal fish assemblages, as well as target fisheries on demersal and pelagic fish, crustacean, and bivalve species. The detected fleet segments were in accordance with expert predictions on the German fishing fleet structure, which were made prior to the analysis. Major changes from the DCF to the alternative segmentation included the formation of the Brown shrimp fishery as a segment, unifying 173 vessels from five different DCF segments, and the split of the DTS2440 segment. The novel segmentation approach revealed that this DTS2440 segment includes vessels operating in four different fisheries with very heterogenic areas of operation, and target stocks. Furthermore, a first check on raw data suggests sensible differences in cost structure. A comprehensive analysis of the cost structure of the respective vessels is a necessary next step to provide solid evidence that fleet segments formed using the novel approach to fleet segmentation are actually more homogenous in cost structure. We therefore propose an additional workshop with special emphasis on cost structure analysis, focusing on well-documented fleets with a solid cost database. In addition, we want to emphasize the necessity to revise the main gear classification in the DCF. Polyvalent vessels using multiple gear over the course of a year are also present in the German data and hinder the formation of consistent fleet segments, as described in ToR 3 of the main report.

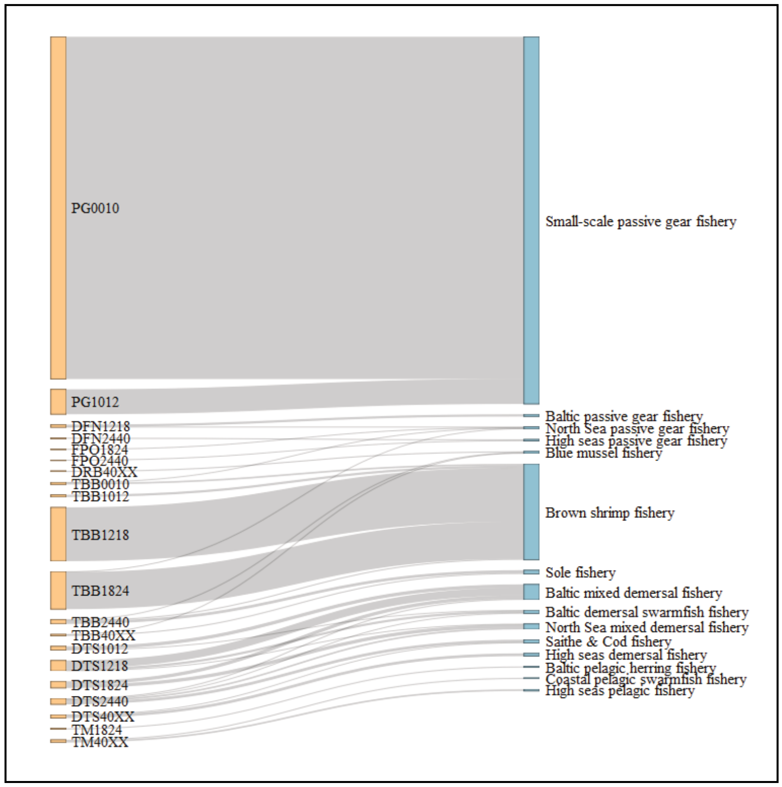


Figure 1: Sankey flowchart of changes in fleet segmentation. DCF segmentation is on the left (orange), alternative segmentation is on the right (blue). German fleet, reference year 2021.

Annex 5.6. Case study - Portugal

Segmentation of the Demersal Trawlers (DTS) operating in Iberian Waters (ICES Divisions 27.9.a and 27.8)

Authors: Ana Cláudia Fernandes, Suzana Faria Cano

Description of analyzed population

The Portuguese Demersal Trawlers (DTS) have annual fixed licenses to operate with bottom otter trawl and, considering the mesh sizes used, they can target mainly demersal species (65-69mm or ≥70mm) or crustacean species (55-59mm and ≥70mm). The areas considered for this group reflect their fishing behavior: it is known that some vessels may operate in ICES Div 27.8.c (Spanish waters) part of the year. The data used in the analysis is from 2021 and includes 90 vessels with vessel length range between 9-34m (mean=23 m; median=24m).

In Portugal there are fisheries targeting species that are not assessed by ICES but are very important at national level (e.g. Atlantic chub mackerel). In this work we removed the ICES stock assignment limitation to enable the classification of all 'species *versus* area' as a new 'stock'. The high number of species*area present in the data (276) reflects the multispecificity of this fleet and the possible difficulty in assigning the target assemblages to some of the vessels (or fishing trips).

The objective of this case study is to analyze the differences observed in the results, if we consider different number of clusters in the data.

Analysis of the data

The analysis was performed using the *'FleetSegmentation'* package.

The clustering diagnosis used to analyze the optimal number of clusters were the average silhouettes, the Mantel test, the Davis-Bouldin index, the SD index and the Calinski-Harabasz index (Figure 1). The optimal number of clusters varies between 2 and 15, depending on the indices considered. Also the clustering dendrograms presented in Figure 2 support that 1 to 24 distinct clusters are present in the data. These results reflect well the diversity of species assemblages observed in this fleet. In our study we will start to consider the optimal cluster number obtained from the average silhouette (6 clusters).

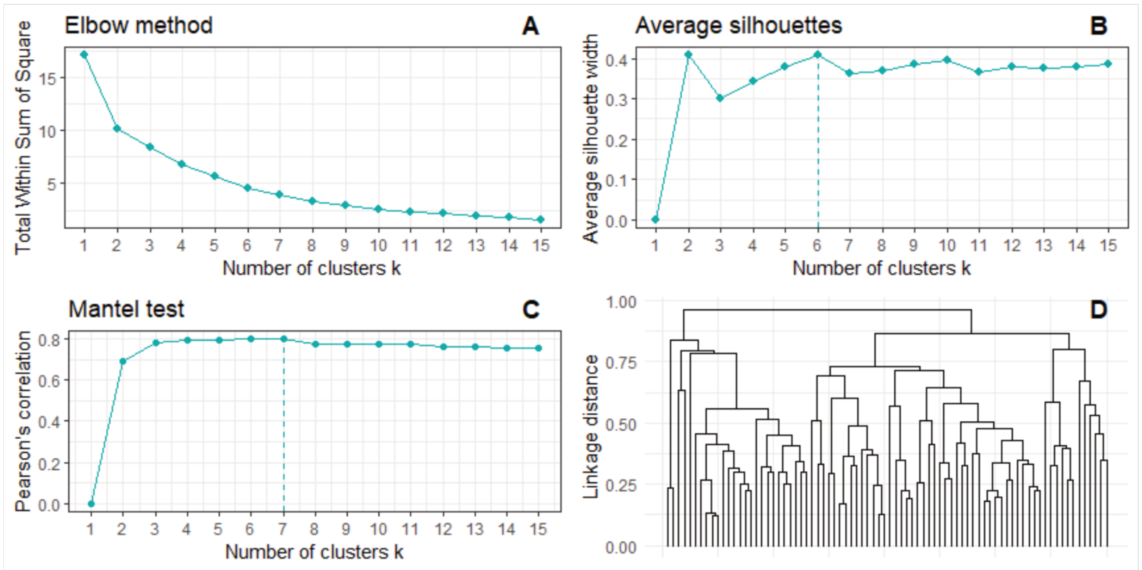


Figure 1 – Clustering diagnosis plots. A) Screen plot showing the total within sums of squares of the clustering vs. the number of clusters used. B) Average silhouette width of the clusters vs. the number of clusters. C) Mantel test, i.e. the Pearson correlation between the clustering and the original distance matrix vs. the number of clusters. In B) and C), the best score is marked with a vertical line. D) dendrogram of the HAC procedure with the linkage distance of the clusters on the y-axis.

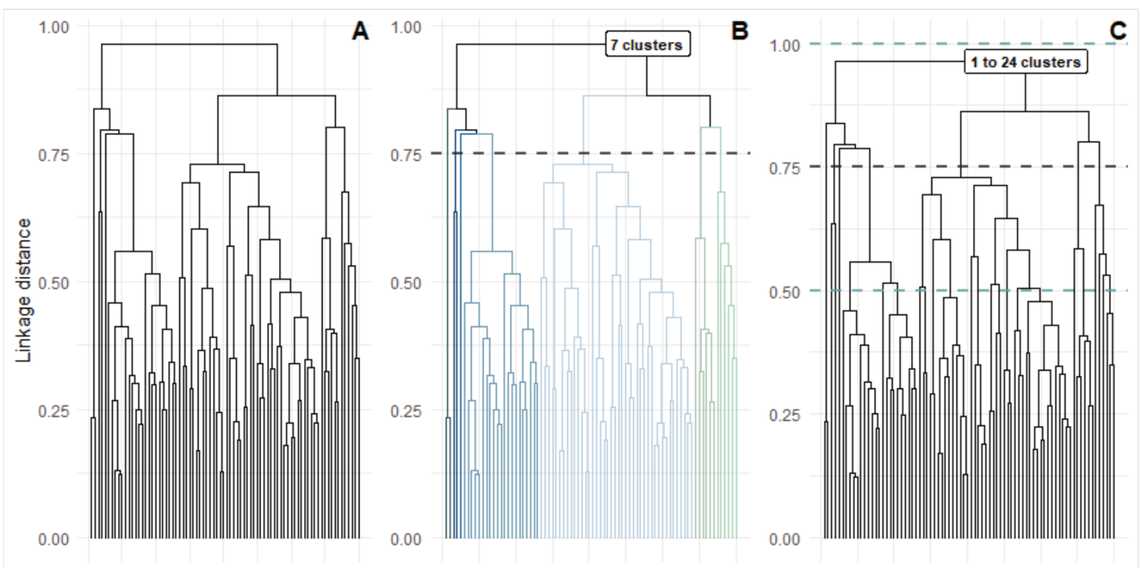


Figure 2 - Grided plot of three dendrograms of the HAC procedure for the Portuguese DTS fleet. The y-axis shows the linkage distance of clusters. A) is unmodified, B) is cut at a linkage distance of 0.75 and resulting branches are individually colored and the resulting number of clusters is labelled, C) has cutting lines at linkages distance 0.5, 0.75 and 1 and the range of cluster numbers resulting from those cutting heights is labelled.

The diagnosis using the clustering tree (Figure 3) showed that the selection of the 6 clusters resulted in groups with very low number of vessels in some of the clusters (2 vessels in 2 of the clusters). Due to confidentiality issues and in order to overcome this low number of vessels, the possible way forward is to consider a maximum number of clusters that are composed by a number of vessels that don't compromise confidentiality of the data, and this is achieved if we consider the data separated in 3 clusters.

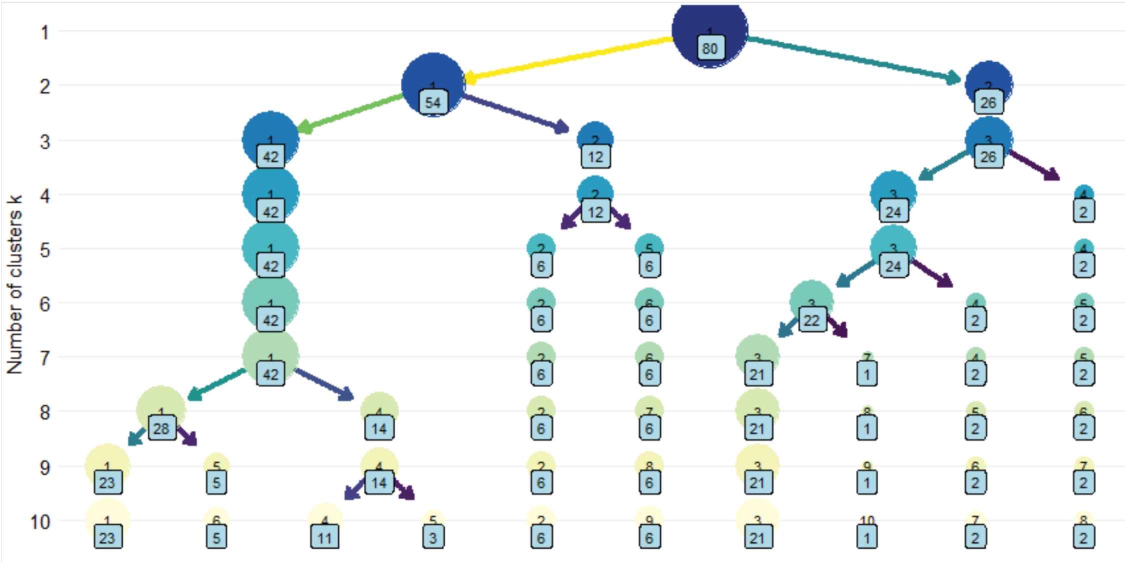


Figure 3 - Treeplot of the HAC-procedure for the Portuguese DTS fleet. The circle size indicates the cluster size, which is also labelled. The number in the circle shows the number of the cluster and in the squares is the number of vessels. The number of clusters in the respective step is depicted on the y-axis.

Cluster characterization and validation

In this part of the work, we present and compare the results obtained in the two approaches: 6-clusters (6CL) and 3-clusters (3CL).

The stockshares of each approach are presented in Figure 4 (A and B - species*area; C and D - target assemblages). The results for the 6CL approach (Figure 4, A and C; Figure 5, A) show that there are 3 clusters that present the fishing behavior of the vessels with the OTB_CRU licenses (clusters 3, 4 and 5), targeting crustaceans and blue whiting (using mesh size ≥70mm), and for the vessels with the OTB_DEF licenses it presents 2 clusters for a mixed target fishery (cluster 2 and 6) and a cluster mainly targeting horse mackerel and mackerel (cluster 1). The results obtained for the 3CL approach (Figure 4, B and D; Figure 5, B) reflect the grouping of the main fisheries obtained in the 6CL approach: cluster 1 seems to be the same in both approaches; cluster 2 in the 3CL results from the grouping of clusters 2 and 6 from the 6CL; cluster 3 in the 3CL is a combination of the clusters 3, 4 and 5 from the 6CL approach.

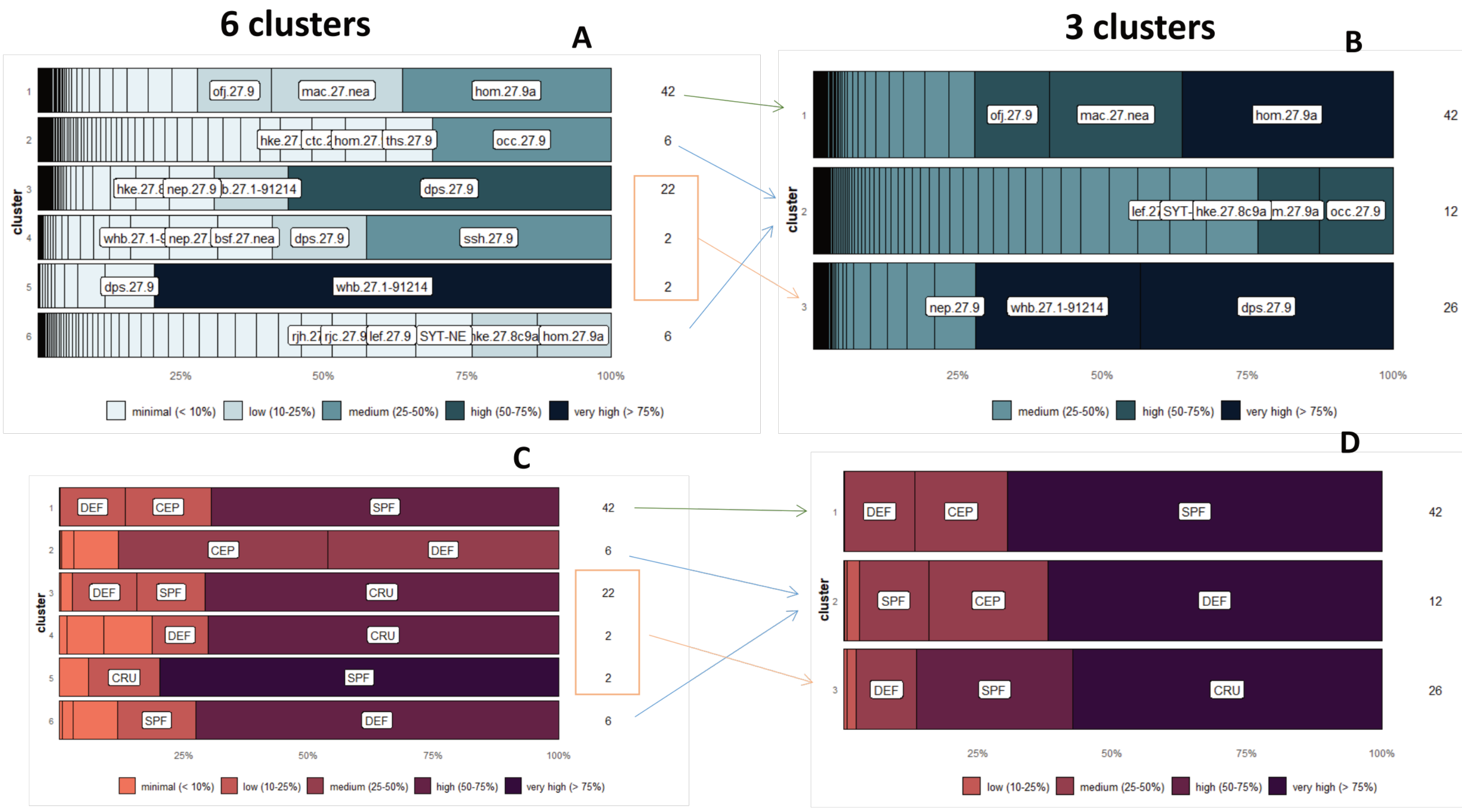


Figure 4 - Barplots of the average percentage of each stock (species*area) and of the target assemblages on the total landings of each cluster of the Portuguese DTS fleet, considering 6 clusters (panels A and C) and 3 clusters (panels B and D) in the data. The average percentage is depicted on the x-axis, the clusters on the y-axis. The colors indicate the magnitude of the stocks contribution to the total average landings. The stock names are indicated with labels; stocks are ordered by their average percentage contribution. The size of each cluster is presented on the right side of each plot.

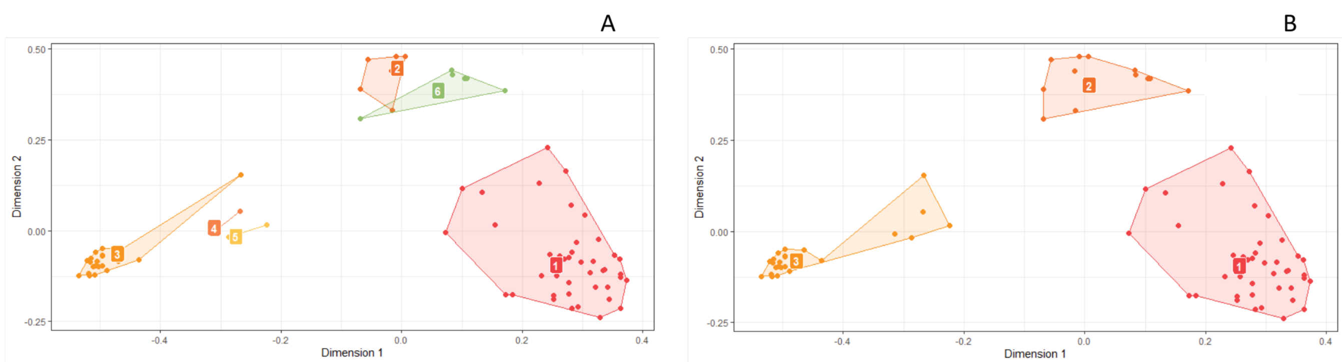


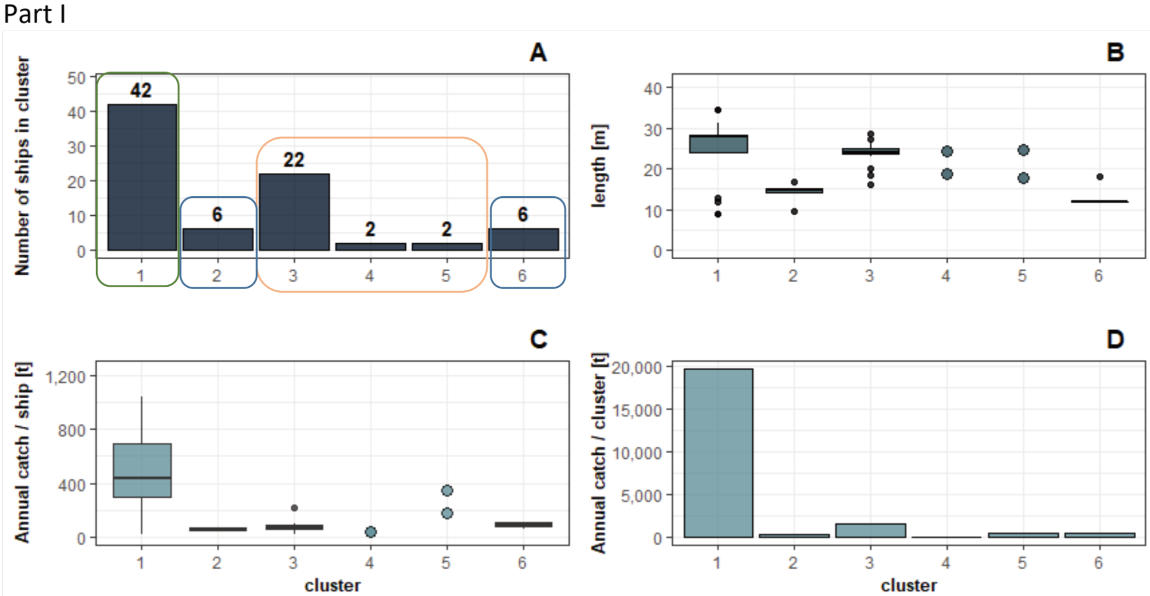
Figure 5 - Multidimensional scaling (MDS) of the transformed vessel catch data for the Portuguese DTS fleet, for the 6CL (panel A) and 3CL (panel B) approaches. Points represent individual vessels, colored according to their cluster affiliation. Clusters are labelled, labels are colored similar to points.

Vessel and cluster properties

The vessels characteristics, considering the two approaches, are presented in Figure 6. It shows that the first cluster in both approaches is composed by the same number of vessels, with exactly the same characteristics. It is also easy to identify the clusters grouped in the 6CL to obtain the clusters identified in the 3CL approach.

Considering the results obtained for the 3CL approach, the Figure 6 (Part II) shows that clusters 1 and 3 are linked to the Portuguese vessels with fishing licenses for mesh sizes 65-69mm and ≥70mm. In cluster 1 they are targeting mainly horse mackerel (27.9.a) and mackerel (27.8.c). The cluster 1 is the bigger cluster of the fleet (42 vessels) and presents also the higher values for the catch per vessel (Figure 6, C) and annual catches (Figure 6, D). The Herfindahl-Hirschman Index (HHI) presented in Figure 7 also gives the indication of a more targeted fishery, when compared to cluster 3 (from the 3CL approach). In what concerns to cluster 3, the main differences observed from the cluster 1, besides the lower number of vessels, is the fact that vessels have smaller size range (median ≈ 13m), thus having a more coastal fishery and presenting lower catches (by vessel and annual), when compared to the vessels in cluster 1. They can be classified as a multispecific fishery (Figure 7, B), targeting a high variety of species (mainly demersal) that may also be conditioned by seasonal factors (e.g. cephalopods) and interannual variability of species distribution. The cluster 2 is composed by vessels with fishing licenses for the mesh sizes 55-59mm and ≥70mm that target mainly crustacean species but also blue whiting. These vessels are part of the other main fishing segment of the Portuguese trawl fleet (OTB_CRU), composed by 26 vessels. Although presenting the lower catches (by vessel and annual) from all the clusters, if value was used for the clustering (instead of the weight) it would highlight better the importance of this fishery at national level. The Figure 6 also shows that the grouping of clusters 3, 4 and 5 from the 6CL approach into the cluster 2 from the 3CL approach showed no relevant impact on the mean values observed for the different variables analyzed. Figure 7 (panel B) supports the fact that cluster 2 represents a high targeted fishery, with the highest cluster overall value of HHI among the 3 clusters.

Part I



Part II

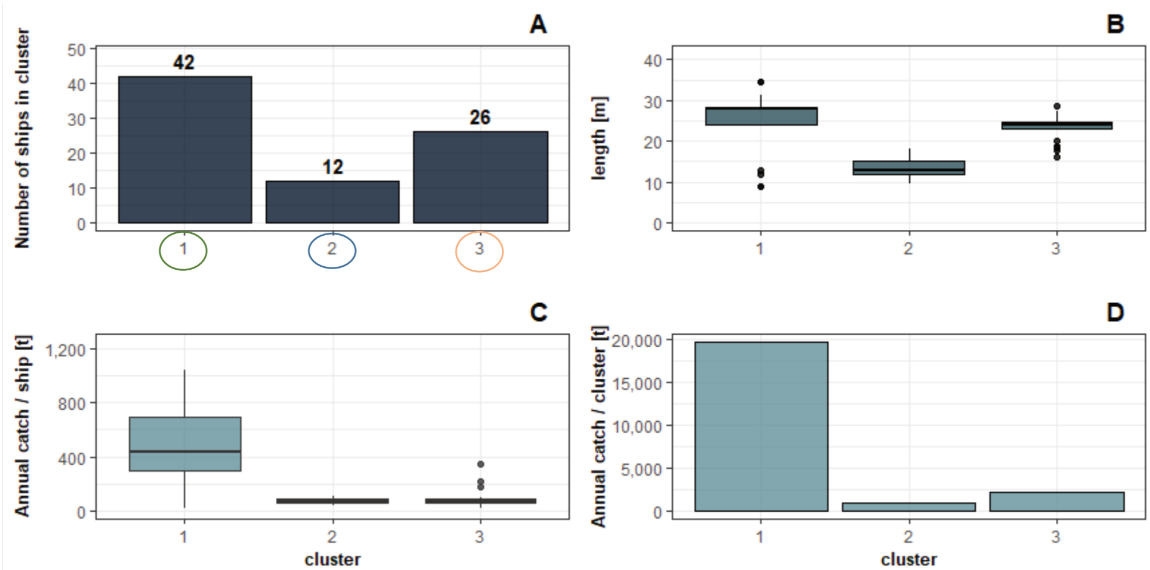


Figure 6 - Grid of mixed plot types displaying vessel characteristics for the Portuguese DTS fleet (**Block I** – 6CL approach, **Block II** – 3CL approach). **A)** Barplot of cluster size with the number of ships in the cluster over each bar. **B)** Mixed box- and dotplot of the length of the vessels in meters, each cluster. **C)** Mixed boxplot of the annual catch of individual vessels in tons. **D)** Barplot of the total annual catch of all vessels in the clusters. (For B) and C), clusters containing more than 5 vessels are depicted with a boxplot, for smaller clusters each vessel catch is plotted with a point).

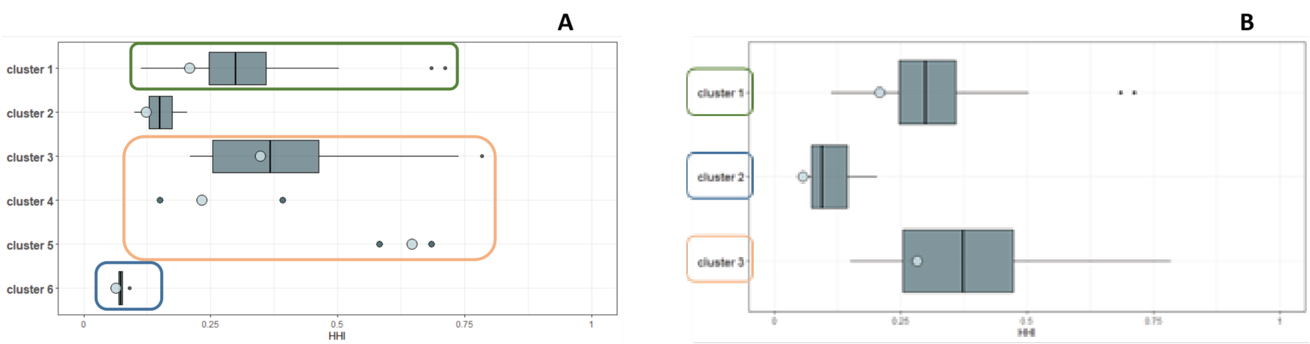


Figure 7 – Herfindahl-Hirschman Index (HHI) of the clusters from the 6CL (panel A) and 3CL (panel B) approaches. Circles are the values for the catch of single vessels (dark blue) and clusters overall (light blue). High score means a targeted fishery and low score a mixed fishery.

Main outcomes

The clustering of the Portuguese DTS fleet data, operating in ICES Divisions 27.9.a and 27.8.c., resulted in 6 different clusters where two of them had only 2 vessels each. Because there is a confidentiality agreement that doesn't allow these type of data (low number of units, with traceable information) to be published, we also considered the highest number of clusters in the data that didn't compromise the rule of the low number of vessels (3-cluster approach) in our study, and compared the results obtained with the two approaches.

In summary, the results showed that:

- The main fishing segments of this fleet can be identified using the 3-cluster approach (OTB_MPD – Cluster 1; OTB_DEF – Cluster 2; OTB_CRU – Cluster3)

- The grouping of the clusters doesn't compromise the information related to the vessels characteristics;

- In relation to the species composition of the landings/target assemblages, we think that the divergence between some of the previous clusters (6CL) and the ones from 3CL is minimal and may also be linked to the annual variability of the species catches and/or fishing behavior. In the case of cluster 3 (OTB_CRU), it also includes vessels targeting blue whiting with crustaceans (cluster 5 from the 6CL approach) but this is also a known fishing behavior of this fishery – some trips/vessels are classified in the present as OTB_MCD, within the OTB_CRU, because of the fishing license they have (55-59mm and ≥70mm).

- The assumption of the low importance of some of the clusters grouped is that those clusters represent a minor part of the vessels of the fleet and they also have a low contribution to the overall catches.