Regional Coordination Group
on Economic Issues

REPORT
OF THE 3rd ISSG WORKSHOP ON AN ALTERNATIVE APPROACH TO THE SEGMENTATION OF FISHING FLEETS

Chaired by Erik Sulanke and 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

3 – 4 May 2023

Online Workshop

Introduction

The third ISSG workshop on an alternative approach to the segmentation of fishing fleets was held online from 3rd to 4th of May 2023. The workshop was held online using Cisco Webex and hosted by the Thünen Institute for Sea, Fisheries, part of the Johann Heinrich von Thünen Institute, Germany's Federal Research Institute for rural areas, forestry, and fisheries.

16 experts representing 10 nations and the JRC participated in the workshop. Eight 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 a second workshop was organized, aiming at harmonization of the data preparation and development of a standardized protocol, evaluation of the new elements and definition of regionally consistent fleet segments over multiple member states operating in the same fishing regions.

The third workshop mainly dealt with suggestions for improvement developed during the 2nd workshop. In particular, the suggestion of pre-segmentation was pursued. Participants were asked to perform varying pre-segmentation-methods. Approaches based on supra-region, length, and gear were regarded conducive. In the previous workshop the length class threshold of the DCF segmentation were regarded too close in several cases, so that a reduction of length classes appeared reasonable. On the other hand, the fishing technique classification of the DCF appeared too low in resolution. Hence, it was suggested to use the gear instead, where applicable. Furthermore, it was suggested to use the métier (level 4) for pre-segmentation, which is mainly defined by the gear, but some gears are grouped (e.g. FPO and FYK).

Participants applied the routine, using different pre-segmentation approaches and trying to assess a reasonable level of clustering. It has to be borne in mind that the clusters are not the final segments, but serve as a basis for fleet segmentation, since they usually have to be aggregated. There are cases of directed fisheries with a specific catch profile and also very little change of gear over the year. In these cases, it is mostly straightforward to derive fleet segments from the clustering results. On the other hand, however, there are cases in which the catch composition is quite diverse over the year, and so might be the gear used. These groups of vessels are regarded polyvalent. These vessels are also highlighted by the clustering procedure.

Using the clustering procedure, participants could derive fleet segments which appeared homogeneous and suitable for representing groups of vessels which perform similar fisheries. On the other hand, groups of polyvalent vessels were found, thus hampering the formation of

segments. This is, however, not a limitation of the clustering approach, but only the reflection of the diversity of the fishing fleet.

The same applies to the fact that catch-based clusters are sometimes very small – containing one or very few vessels. This is as well the reflection of reality and not a drawback of the approach as such: Any realistic segmentation approach would highlight the same small groups. Small segments are only a problem in data reporting when it comes to confidentiality. Grouping small segments for the purpose of publishing data is always a challenge. The catch-based clustering approach provides a useful background for this step.

Four principles as laid down in the 2nd workshop have been repeated:

- Connection to specific fisheries (high priority)
- Cost structure (high priority)
- Feasibility (high priority)
- Compatibility (lower priority)

These principles should be taken into account when stipulating criteria for comparing different segmentation approaches.
Following this list, an exemplary analysis of the link between alternative segments and stock exploitation was presented for the German fleet. It could be clearly shown that the number of segments being involved in the exploitation of important stocks is, in most cases, lower for alternative segments than for DCF segments.
Following up on the 2nd workshop, the results of an Artificial Intelligence procedure were shown for the German fleet. After introducing a learning dataset, the system was able to classify segments with an accuracy of about 99%. This procedure appears worth being tested on further MS fleets.

Recommendation from RCGECON 2022 and Terms of Reference

RCGECON recommended that "*The alternative approach should be further developed to achieve a consistent and comprehensive procedure which can be tested against the existing procedure with respect to specific criteria, in particular the link to stocks and the homogeneity in cost structure.*". Moreover, some follow-up actions were specified:

Develop pre-segmentation protocols:

- *introducing vessel characteristics indicator (e.g. size, power, equipment)*
- *grouping gears (regionally) to replace "main gear" as currently used*
- *introducing reasonable target assemblages or métier groups by region*

Continue describing regional, consistent fisheries
Check effects on cost structure
Draft set of rules

The workshop focused on the evaluation of pre-segmentation effects as observed by participants. Another focus was set on quantitative comparison of segmentation approaches. Moreover, the potential of Artificial Intelligence to automatize the segmentation procedure was addressed.

In the end the workshop was too short to address the last three points as there is still need for implementing pre-segmentation steps as well as further approaches to derive segments from the clusters.

Background

The current DCF 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 established, but has some drawbacks, as it does not adequately represent target fisheries. Vessels with similar technical parameters are often active in different fisheries, targeting different stocks and having different catch composition, fishing activity, and cost structure. However, fisheries or, more specific, stocks are one key reference for fisheries management.

In the two preceding workshops a clustering procedure was applied to several fleets. The procedure is using multivariate statistics analysing the catch composition with reference to specific stocks. The results of this procedure are the basis for generating segments using expert knowledge. The results were promising, and suggestions for further improvement have been derived.

The 3rd workshop is intended for testing these suggestions and further amending the procedure.

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 preceding workshops on fleet segmentation (https://datacollection.jrc.ec.europa.eu/docs/other-meetings).

MS results – general description

Greece: Pre-segmentation as "PS-LSF" and "DTS (mainly-LSF)".

Denmark: The Danish dataset was analysed using two different pre-segmentation approaches – one based on the DCF fishing technique (tab. 8, EU 2021/1167) plus a length threshold for DFN at 10m, the other using métier level 4 gear type (corresponding to annex XI, EU 404/2011) or combination of gears in case they represent similar fisheries.

Spain: Exemplary analysis is focused on the Supra-region Med/BS. Partial pre-segmentation by gear (DFN, DRB, FPO, HOK, HOK-LLD, PS,) but also gear classes (DTS, PMP). No further pre-segmentation with respect to length class was performed as the number of SSF vessels was low for some gear and, moreover, there is no evidence for size-dependent differences in activities.

The Netherlands: three different pre-segmentation approaches – Ia. combining DTS and TBB, Ib. combining all other techniques, II. No pre-segmentation, but analysis of fleet as whole, III. Analyzing TBB and DTS separately.

Cyprus: Three gear groups were analysed – PG (325 vessels), PGP (36), and DTS (5). No further pre-segmentation was applied.

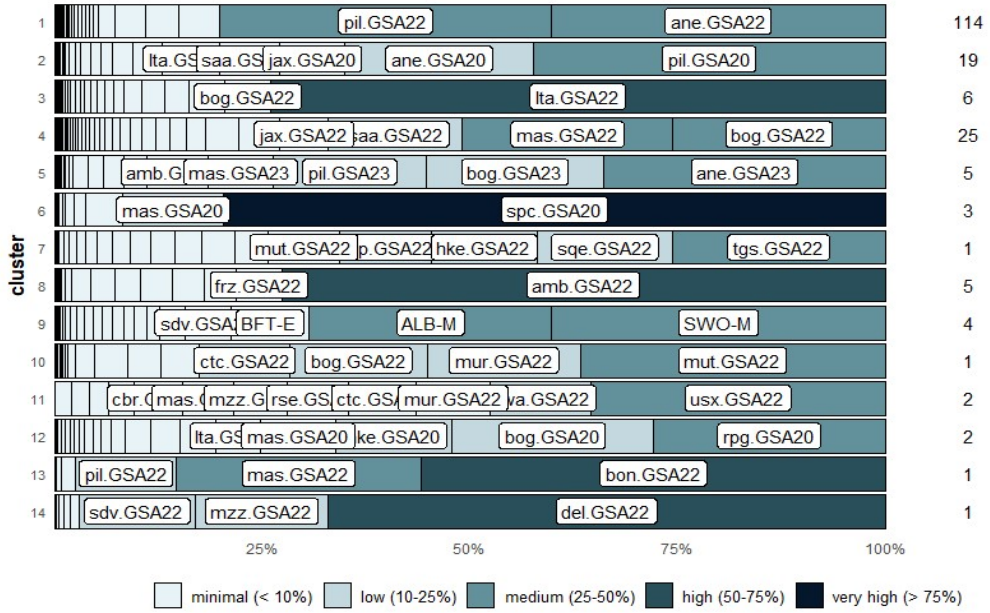
Finland: The Finnish dataset was analysed while splitting TM into single trawl (OTM, 26 vessels) and pair trawl (PTM, 17) on the one hand, and 1005 DFN vessels from the PG gear group on the other hand.

France: For the French case the vessels from the supra-region Atlantic were analysed. Pre-segmentation was applied using DCF gear class (DTS), DTS with three length classes (0-12m, 12-24m, >24m) and a self-defined (Ifremer) group of exclusive demersal trawlers (TB "bottom trawls, not specified"). Moreover, the analysis was performed on the basis of métier level 5 (gear+target assemblage). Also, value instead of weight was used for one set of analyses.

Ireland: The segmentation of the Irish fleet was performed only after the workshop, but for illustration purposes some examples are included in this report. The analysis was performed on HOK, DTS, DRB, DFN, FPO, TBB, and TM. For DTS and FPO pre-segmentation by size (SSCF/LSF) was introduced.

Results with "good fit"

Greece: Purse seiners could be grouped into four distinct segments (Figure 1), separated by groups of targeted species "Small pelagics" (163 vessels), tuna/tuna-like" (11), "other fish" (14), and "crustaceans" (1).



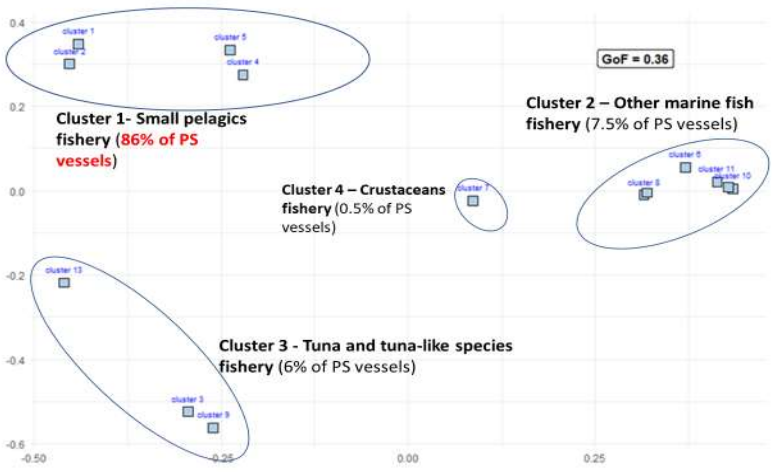
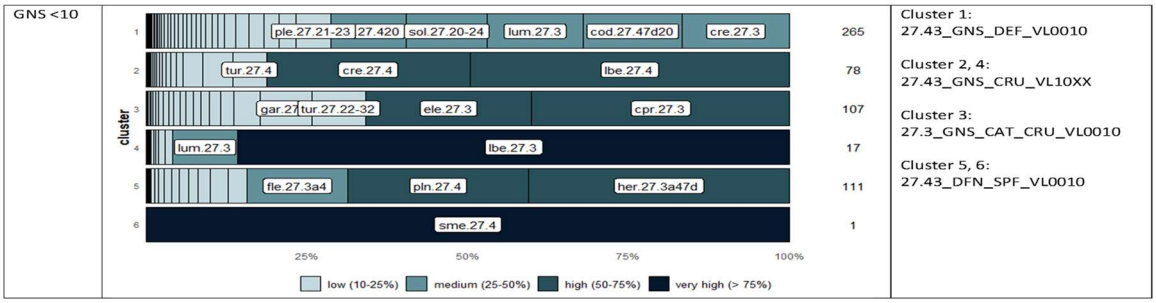


Figure 1: Greek cluster analysis for PS, catch composition and grouping to clusters

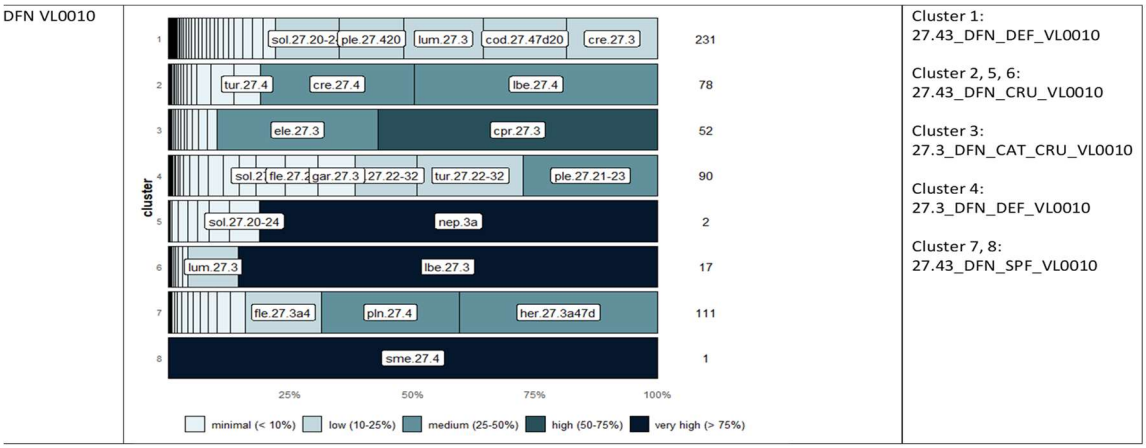
Denmark: each of the two pre-segmentation approaches applied to the entire fleet resulted in 22 (25) well-defined clusters, some of which were quite small (1-8 vessels), but all of them performing a fishing pattern which can be distinguished from others. In some cases both approaches delivered identical results, whereas some other segments differ in name or vessel number. Figure 2 illustrates the catch profile of the Danish GNS métier for vessels <10m, the corresponding clusters, and, based on these shares, segments as clusters, combined by expert knowledge.



North Sea/Skagerrak Gillnet vessels <10 m oal demersal fish	27.43_GNS_DEF_VL0010	Gillnet vessels less than 10 m targeting demersal fish (cod, plaice, lumpsucker, crab) in the North Sea/Skagerrak	265
Baltic and North Sea Gillnet vessels <10 m oal crustaceans	27.43_GNS_CRU_VL0010	Gillnet vessels less than 10 m targeting crustaceans (lobster, crab, nephrops) in the North Sea and area 3a	95
Baltic Sea Gillnet vessels <10 m oal eel and prawn	27.3_GNS_CAT_CRU_VL0010	Gillnet vessels less than 10 m targeting eel and prawn in the Baltic Sea	107
Gillnet vessels <10 m oal small pelagic fish	27.43_GNS_SPF_VL0010	Gillnet vessels less than 10 m targeting herring	112

Figure 2: Danish GNS<10m metier, segmentation ("cluster") with support of catch profile and description of the corresponding segments

For the Fishing_Tech approach a slightly different segmentation was derived (Figure 3)



North Sea/Skagerrak Gillnet vessels <10 m oal demersal fish	27.43_DFN_DEF_VL0010	Gillnet vessels less than 10 m targeting demersal fish (cod, plaice, lumpsucker) in the North Sea/Skagerrak	231
Baltic and North Sea Gillnet vessels <10 m oal crustaceans	27.43_DFN_CRU_VL0010	Gillnet vessels less than 10 m targeting crustaceans (lobster, crab, nephrops) in the North Sea and area 3a	97
Baltic Sea Gillnet vessels <10 m oal eel and prawn	27.3_DFN_CAT_CRU_VL0010	Gillnet vessels less than 10 m targeting eel and prawn in the Baltic Sea	52
Baltic Sea Gillnet vessels <10 m oal demersal fish	27.3_DFN_DEF_VL0010	Gillnet vessels less than 10 m targeting demersal fish (plaice, turbot, cod) in the Baltic Sea	90
Gillnet vessels <10 m oal small pelagic fish	27.43_DFN_SPF_VL0010	Gillnet vessels less than 10 m targeting herring	112

Figure 3: Danish DFN gear class <10m, segmentation ("cluster") with support of catch profile and description of the corresponding segments

The Netherlands:

The Dutch investigation focused on the variation in resolution during pre-segmentation and on the effect of changes in number of clusters. Figure 4 shows the results for clustering TBB&DTS together versus separate clustering. Some distinct clusters (raj, ers, "guu-mur-mac") are detected regardless of pre-segmentation, whereas results for flatfish fishery (ple, sol, fle) indicate a quite variable composition. The decision for a certain version might be made based on the importance of certain fisheries. If the total catch of a segment is high, then a separation might be worth considering. The DTS cluster "csh-nep-ple" might represent a mixed fishery as these species are

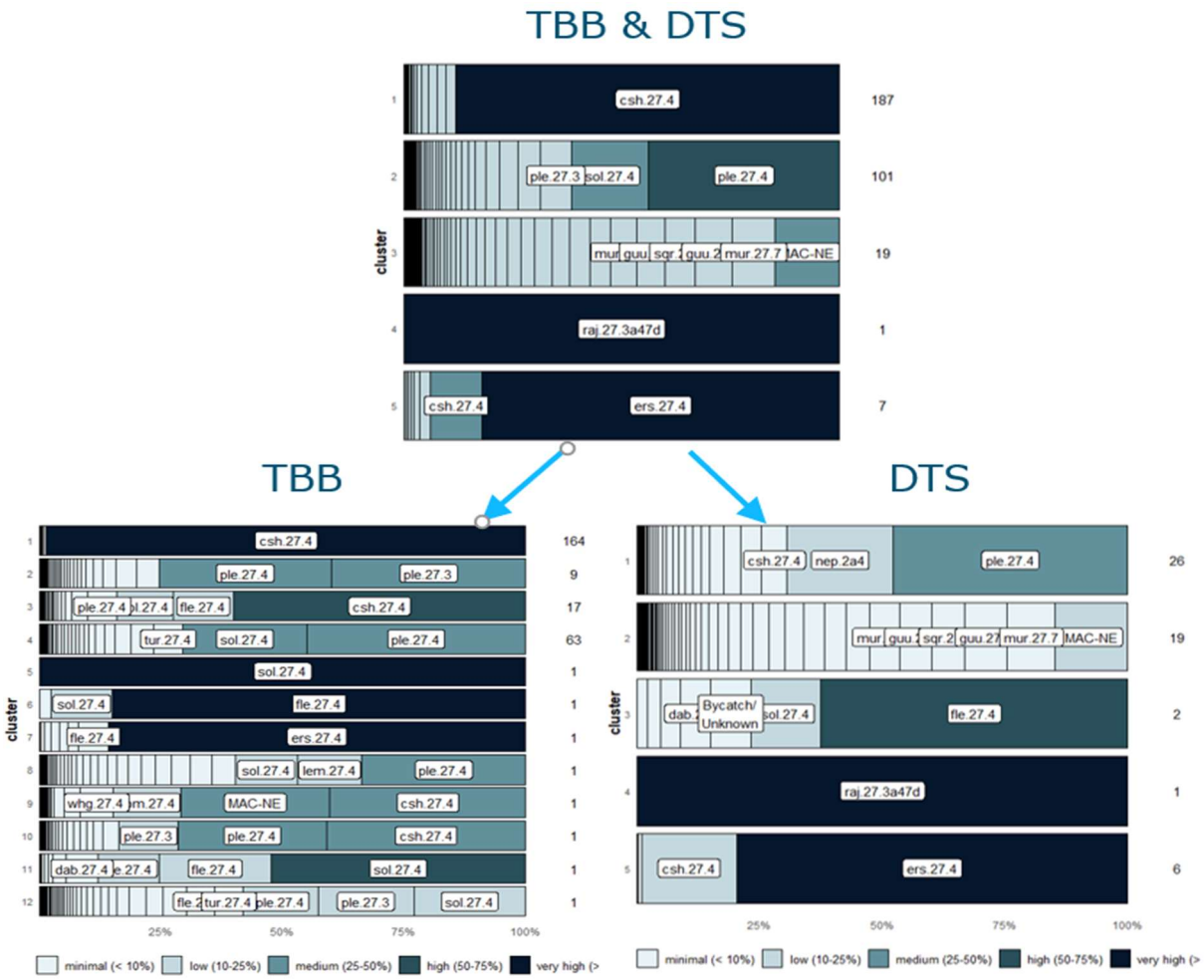


Figure 4: Comparison of results with different gear separation during pre-segmentation

Spain:

For the group of dredgers the analysis of catch composition resulted in two segments which quite distinct catch profiles (Figure 5). For the Med the Spanish fleet consists of three DCF segments, two of which are very small, and hence all three are grouped to one DCF cluster. The alternative, catch-based approach results in two segments which can be clearly distinguished by their catch composition – one targeting mainly smooth callista and tuberculate cockle, the other striped venus and truncate donax.

The Spanish analysis provides further examples for good fits, also for fisheries which remain undetected under the current DCF segmentation (e.g. bluefin tuna). There are, however, also cases for which the catch composition alone does not give clear evidence for segmentation.

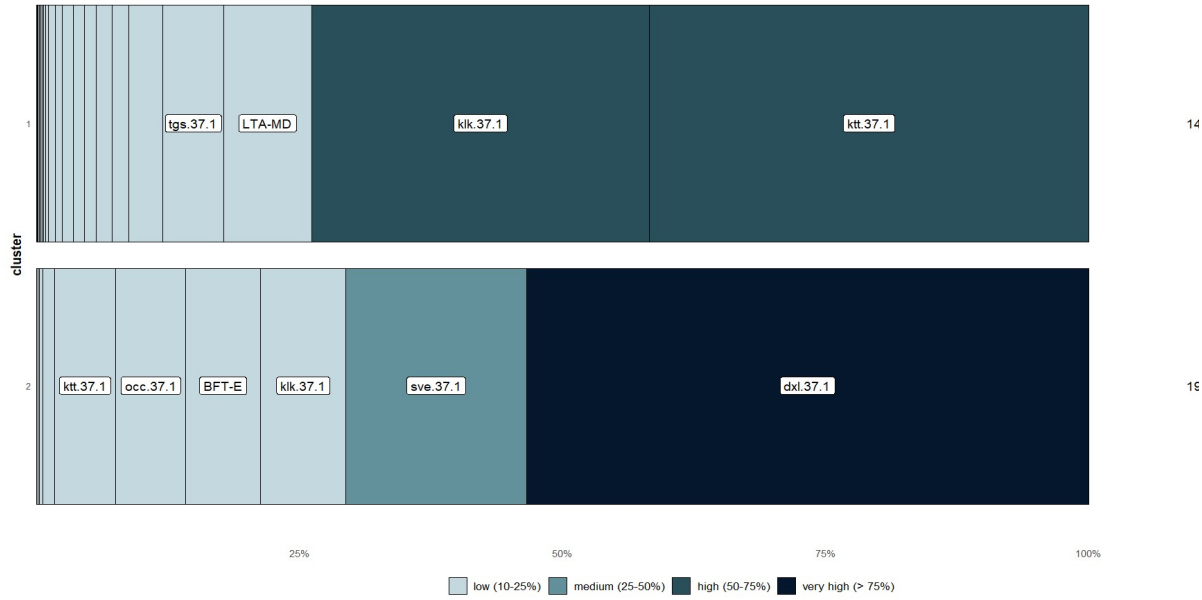
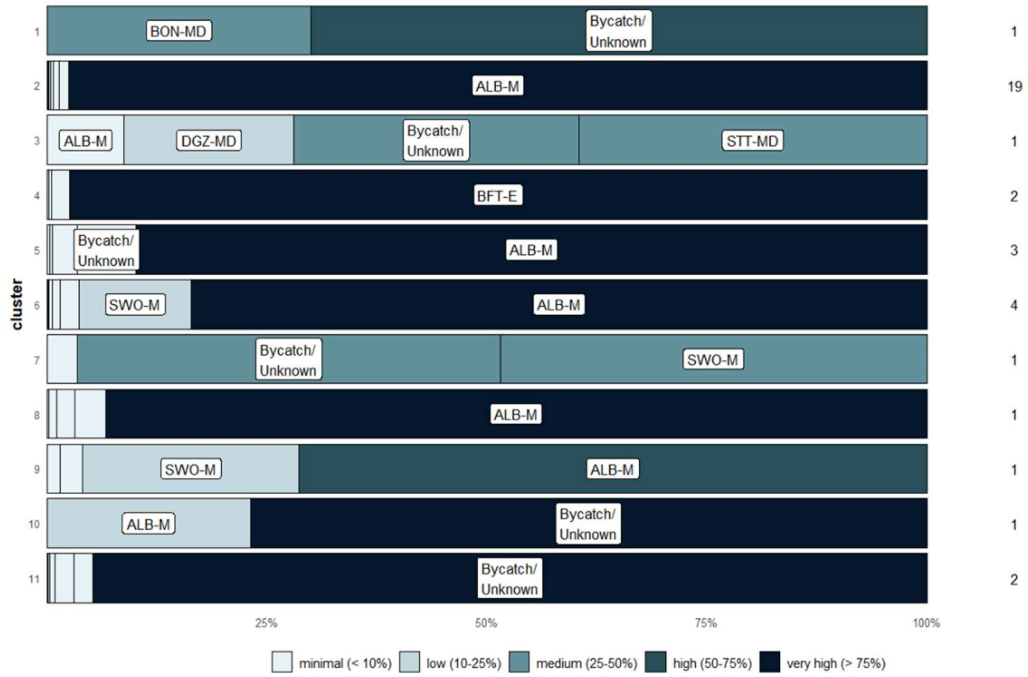


Figure 5: catch profile of Spanish dredgers in the Mediterranean Sea. Three DCF segments are transferred into two alternative segments with distinct differences in the catch profile.

Cyprus:

For the Cyprus PGP fleet, consisting of 36 vessels, the threshold was set at 11 clusters (Figure 6). Most clusters contain few vessels only, several clusters have low catches on a broad range of species, which are therefore classified as "bycatch/unknown" and are grouped into one alternative segment. Yet, the classification of "bycatch/unknown" represents an artefact of a threshold contained in the first version of the FS package used by the MS. This could be easily adjusted by setting the threshold lower. Three segments represent rather targeted fisheries, on bluefin tuna, albacore or albacore/swordfish. The bluefin tuna segment contains two vessels only and might therefore be further aggregated. However, the analysis indicates that this specific fishery exists.



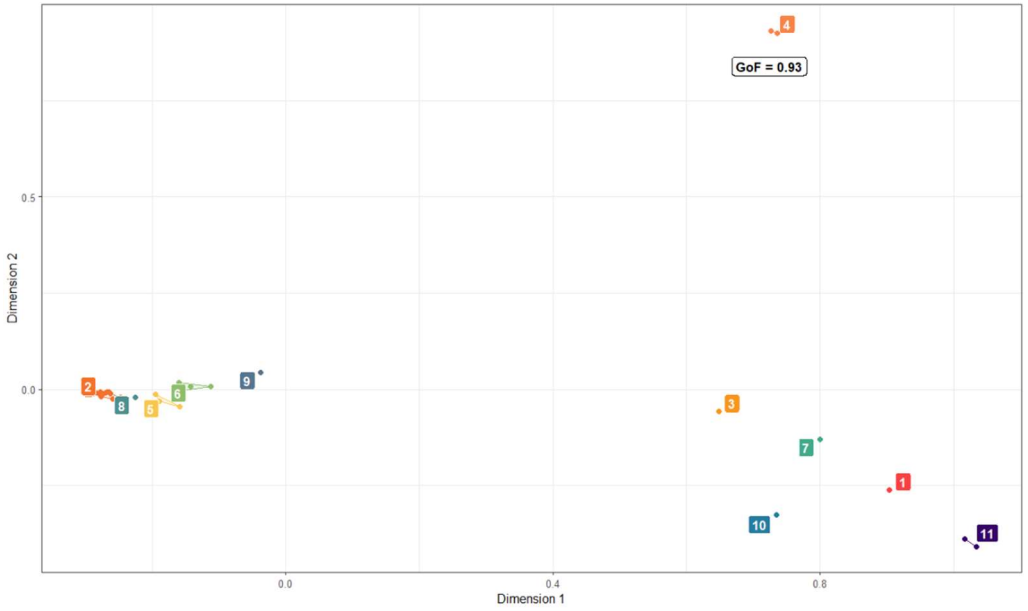


Figure 6: Results for the Cyprus fleet: 11 catch-related clusters were chosen, which are then grouped to 4 alternative segments, based on MDS analysis: I = 1/3/7/10/11, low annual catch, high share of diverse species, grouped as "bycatch"; II = 2/5/8, targeting mainly albacore; III = 4, two vessels targeting mainly bluefin tuna; IV = 6/9, targeting mainly albacore and swordfish

Finland:

The large (1005 vessels) group of Finnish fixed netters from the PG gear group (<12m) was analysed with respect to the catch composition. The version with 11 clusters was presented (). The MDS analysis indicates three cluster groups which might be combined as alternative segments.

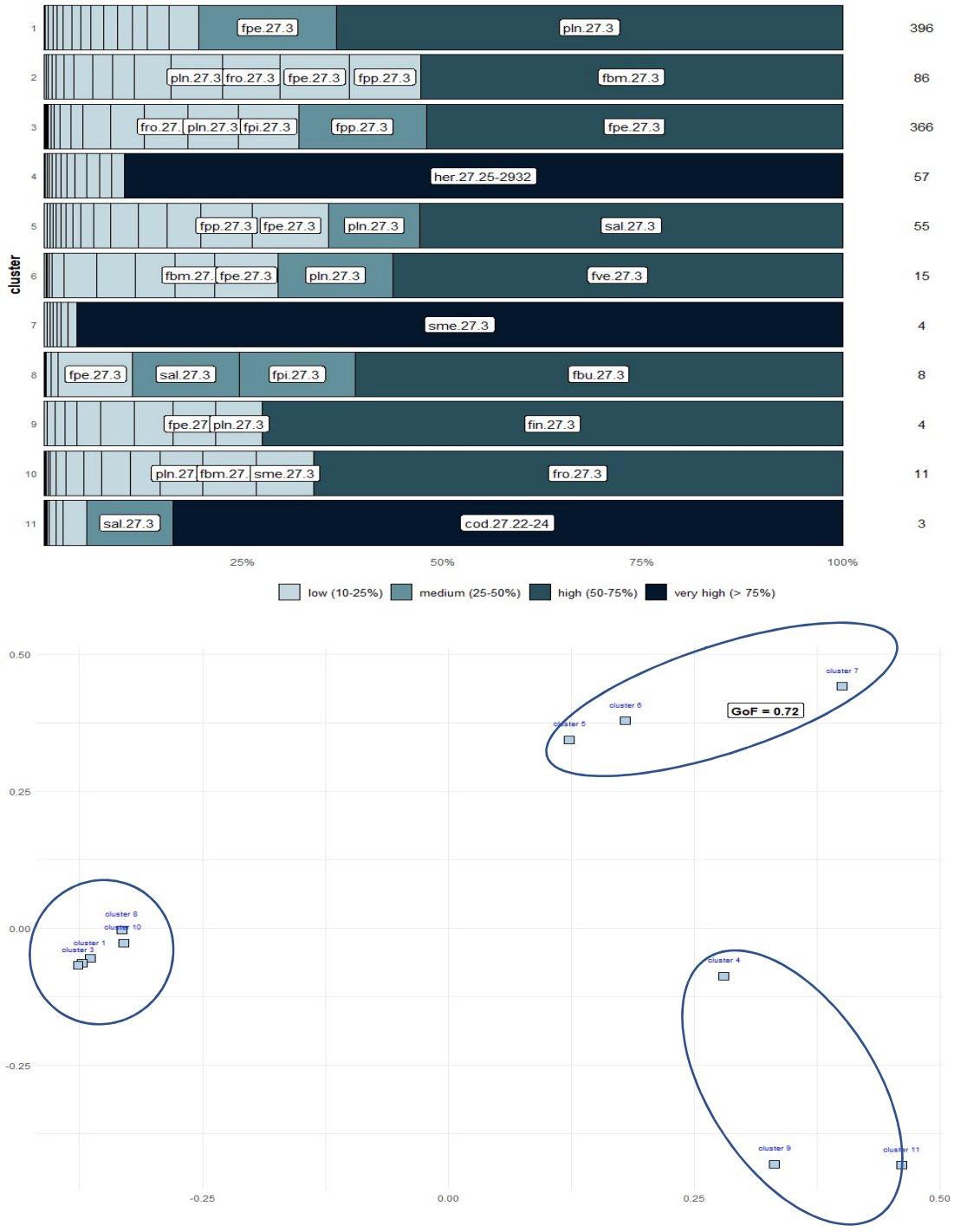


Figure 7: Results for the Finnish PG-DFN fleet (1005 vessels). The MDS results suggest grouping of the catch-based clusters into three alternative segments. Overall, the catch composition is quite heterogeneous, most vessels target freshwater or diadromous fish, while only few catch mainly marine fish

France:

The French group of demersal trawlers and seiners was separated into four main clusters see Figure 8: "1218 exclusive trawlers" targeting mainly Norway lobster, albacore, monkfish, cuttlefish and hake (cluster 1), "1824 exclusive trawlers", targeting scallops and cuttlefish (2), "1224 mixed trawlers", targeting herring, mackerel and whiting (4), and "1224 trawler dredgers", targeting monkfish, haddock, whiting and megrim (6).. Clusters 3,5, and 7 contain only one vessel each and target a very specific assemblage of species. In this particular case, the catch profiles correspond with length classes.

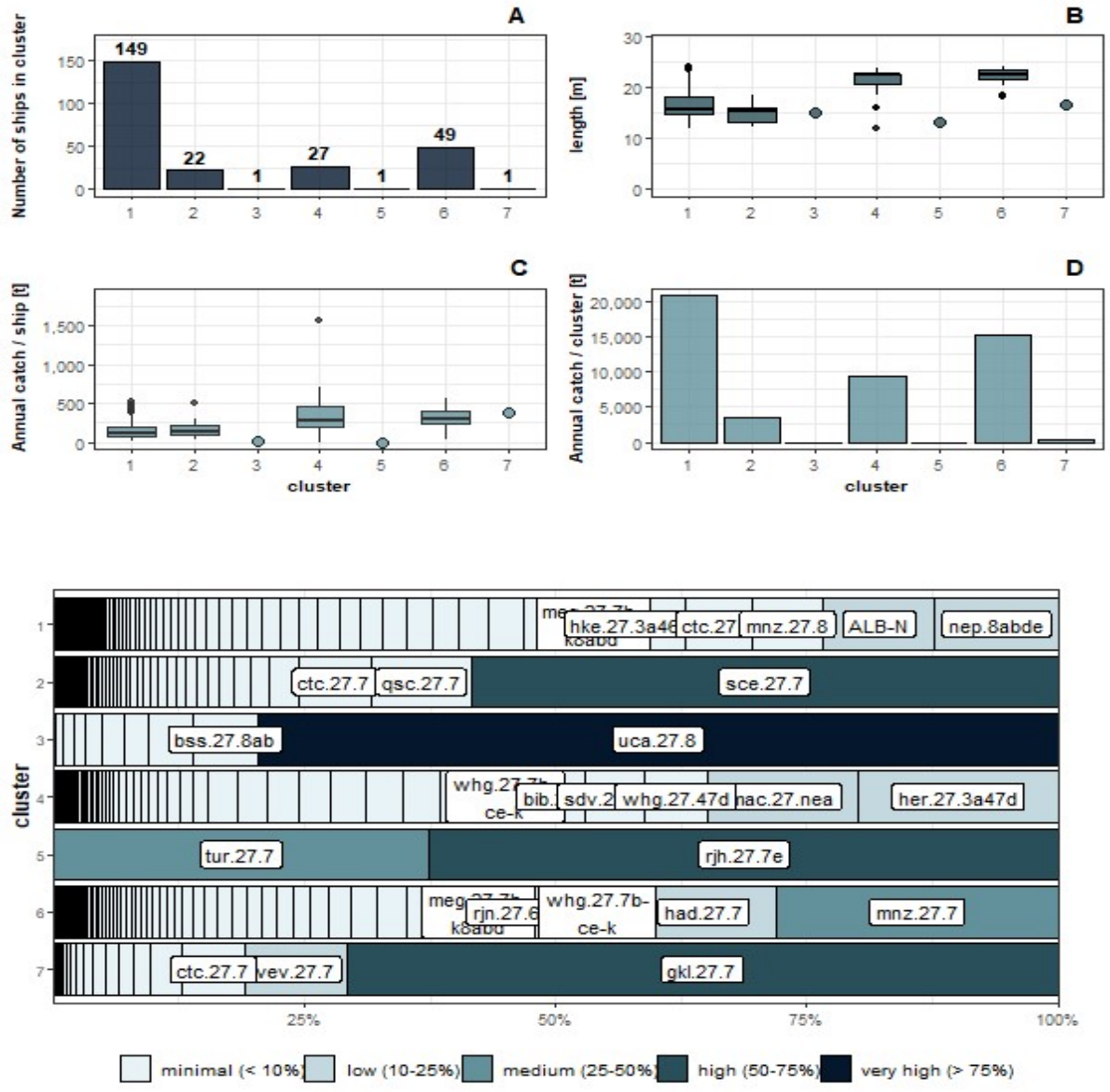


Figure 8: French DTS 1224 vessels (Atlantique) – four main segments were derived from the cluster analysis,, three clusters have one vessel only and target a specific assemblage of species

Ireland:

The analysis of Irish beam trawlers indicates two clusters which clearly differ in catch composition (Figure 9). Some species are usually not caught with beam trawls (e.g. sprat, haddock, hake). Hence, it is quite likely that beam trawling is the dominant gear over the year, but other techniques were in use as well. As both clusters are rather small, it was suggested to merge them into one alternative segment. In this case, the result is the same as under the DCF segmentation where two length classes had to be clustered. However, the length distribution plot in Figure 9 (B) indicates that the catch profiles do not coincide with length classes.

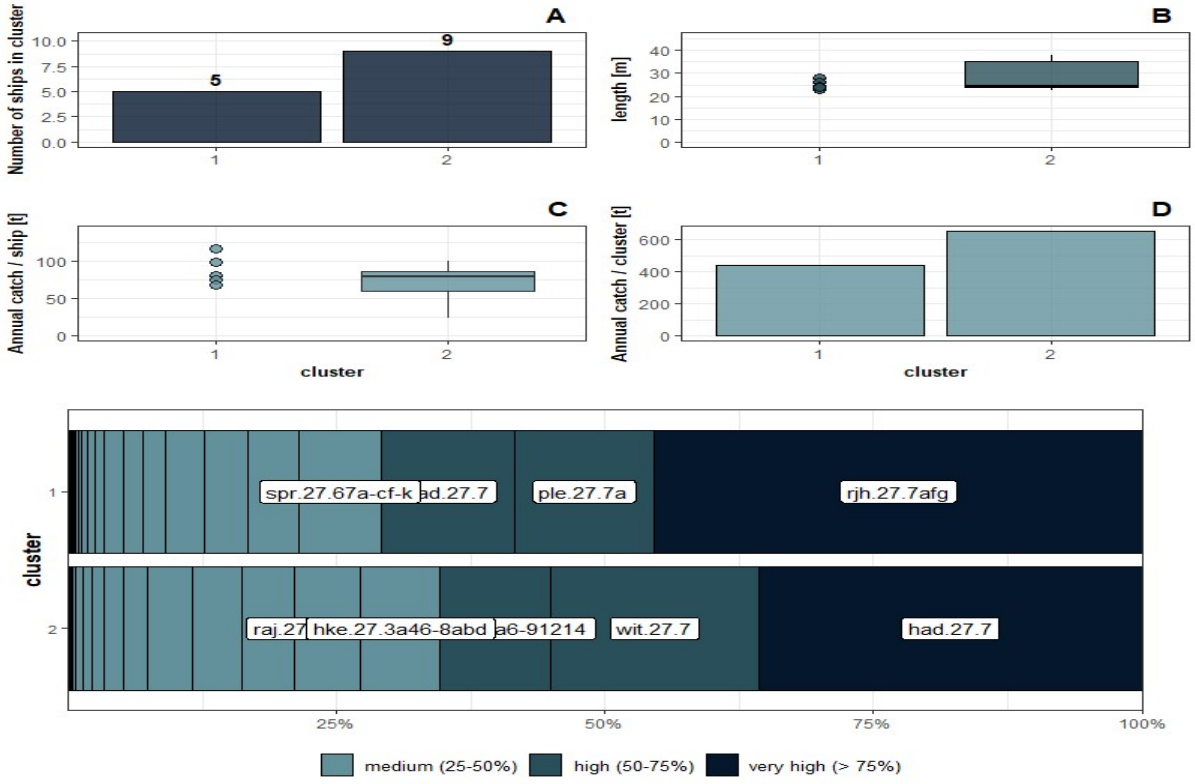


Figure 9: Irish beam trawlers – Within a group of 14 vessels two different catch profiles could be derived, with some species that are not typically caught with beam trawls (sprat, haddock, hake), suggesting that the vessels might change gear oer the year.

Another insightful example from the Irish analysis is that of FPO-LSF (Figure 10). While four clusters target whelk, brown crab and spider crab, one cluster indicates a herring profile. Most likely this fish is caught by nets which are classified under the gear group "traps" (Dec. 1167/2021) and thus are coded "FPO". This brings up the idea of reconsidering the gear group classification as net-based traps might have different biological and economic characteristics than pots. Overall, it appears reasonable not to merge the herring cluster with the CRU/MOL clusters, even though it consists of two vessels only.

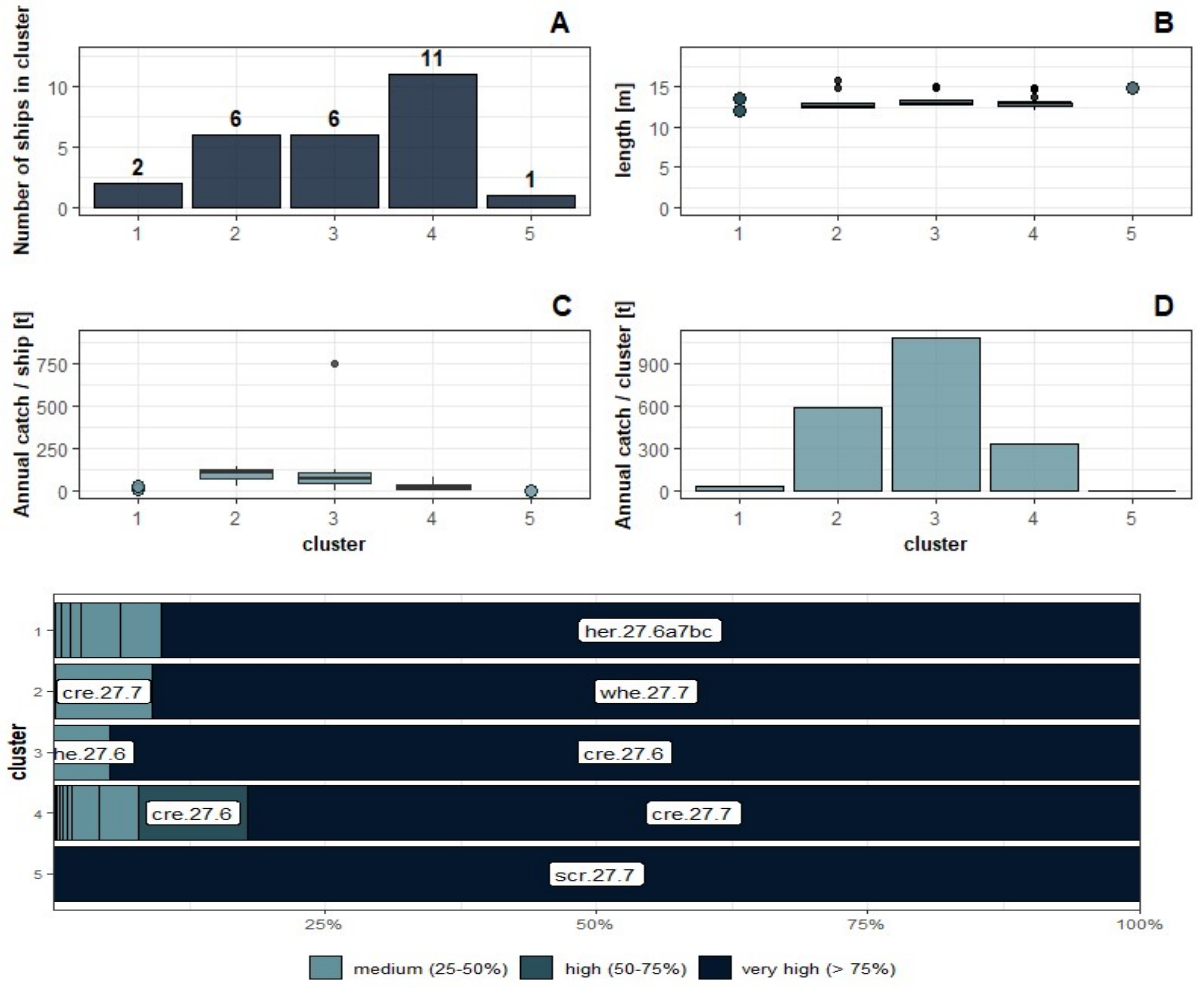


Figure 10: Irish pots and traps large scale group – the group is rather small (26 vessels), most clusters are similar in terms of catch composition (whelk, brown crab), the herring cluster stands out, though.

Results with "not so good fit"

Greece:

Demersal trawlers and seiners showed a rather diverse range of catch profiles. Species distinction by GSA results in (too) many different clusters which might better be treated as unit. Clusters 5-12 are small in numbers (1 or 2 vessels) and thus have to be grouped for confidentiality reasons. Catch and length profiles give some support for the grouping decision.

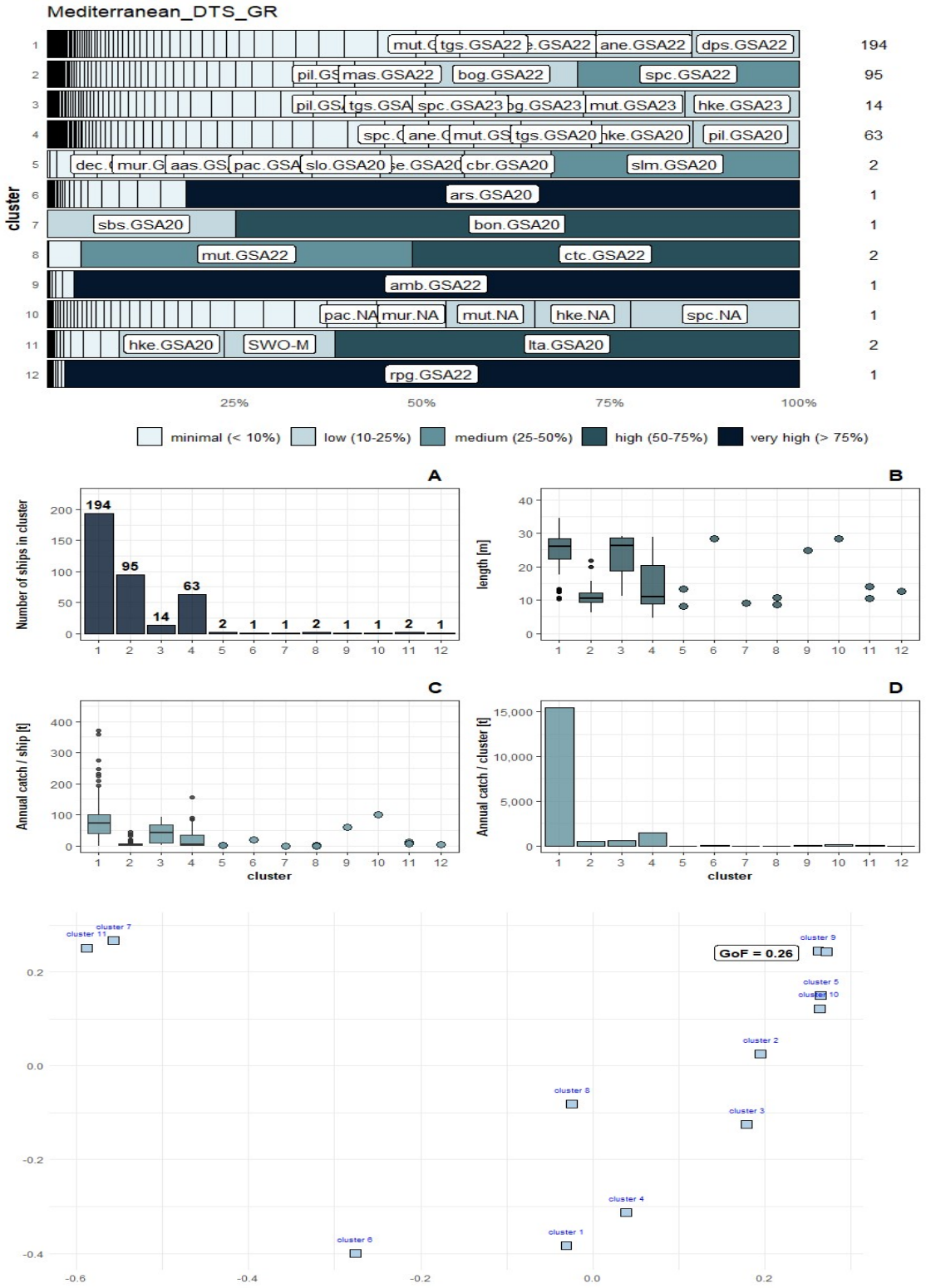


Figure 11: Greek cluster analysis for DTS, deriving segments from catch clusters is not evident

Spain:

The Spanish DTS group, representing four DCF segments, is analysed altogether, without further length or gear pre-segmentation. The group targets a wide range of species, thus resulting in clusters with more or less diverse catch profiles. This changes typically with changing number of clusters. Therefore, the clusters appear quite heterogeneous. It is suggested to include effort information (days at sea) to separate extreme cases the catch profile of which is more random, but would not represent a typical fishery. Further consideration regards overcoming the "500 species problem" in the Med by using species assemblage information rather than a stock-based clustering approach.

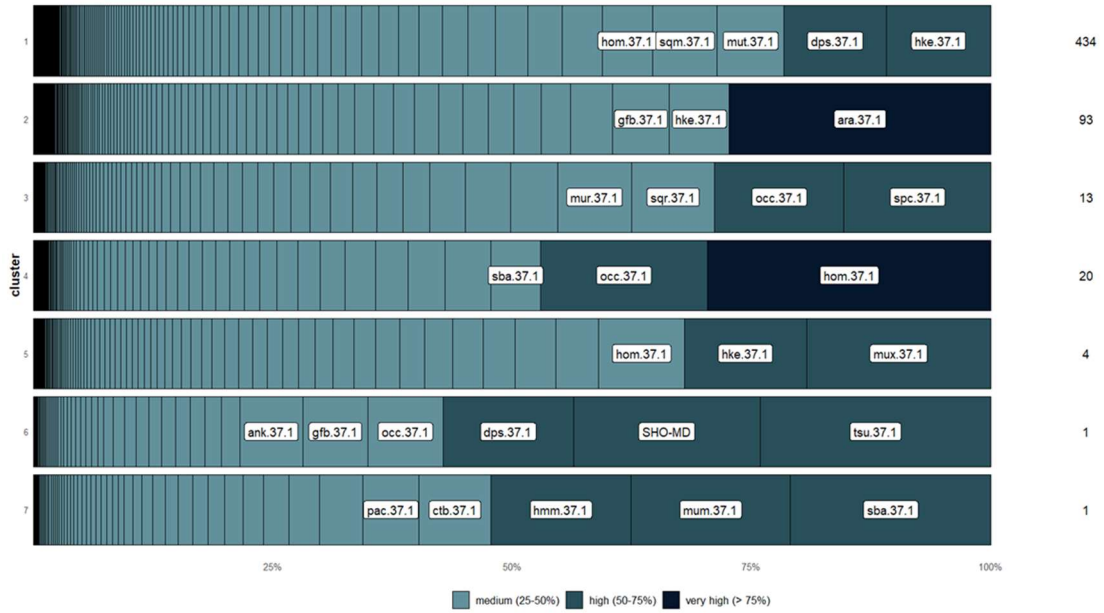


Figure 12: Spanish DTS group. When choosing seven clusters, three clusters with very few vessels are generated. These might have to be further grouped. The vessels range over four DCF length classes, from 6 to 40m.

Cyprus:

The Cyprus fleet has a large PG group. Their catch is highly diverse, and as a consequence, the highest share of catch is assigned to "bycatch/unknown". This is determined by settings of the evaluation procedure and could be easily adjusted. The GoF analysis would suggest some clusters, but this appears to be not reasonable. It is suggested to treat the entire PG group as one segment. It is stated that a length separation, as prescribed under the DCF, is not regarded reasonable.

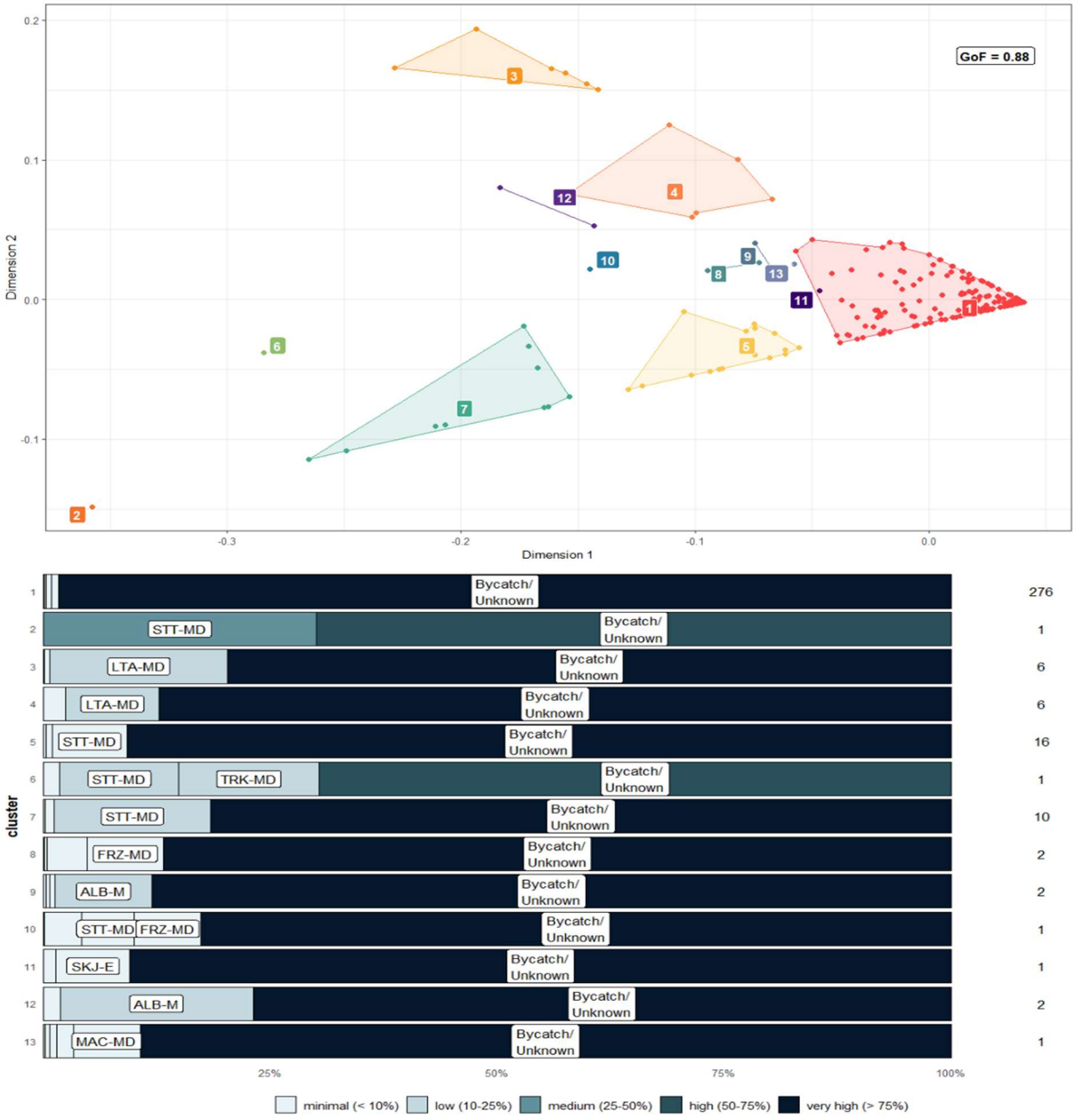


Figure 13: Cyprus PG group; due to the low amounts, most catch for this group was assigned to "bycatch/unknown". A distinction by catch composition is not regarded useful, even though the GoF analysis would suggest some grouping.

France

The French group of demersal trawlers and seiners >24m consists of 64 vessels, representing a broad range of catch profiles (Figure 14). The number cluster implies that a merger of these clusters is necessary to drive clusters with more vessels. Even though the catch composition is very diverse, the diagrams support this task, e.g. clusters 7, 12, and 13 are the largest vessels in the group and have some similarities in catch composition. In addition, their catch profile is rather homogeneous, and they might be segmented as "large demersal roundfish trawlers".

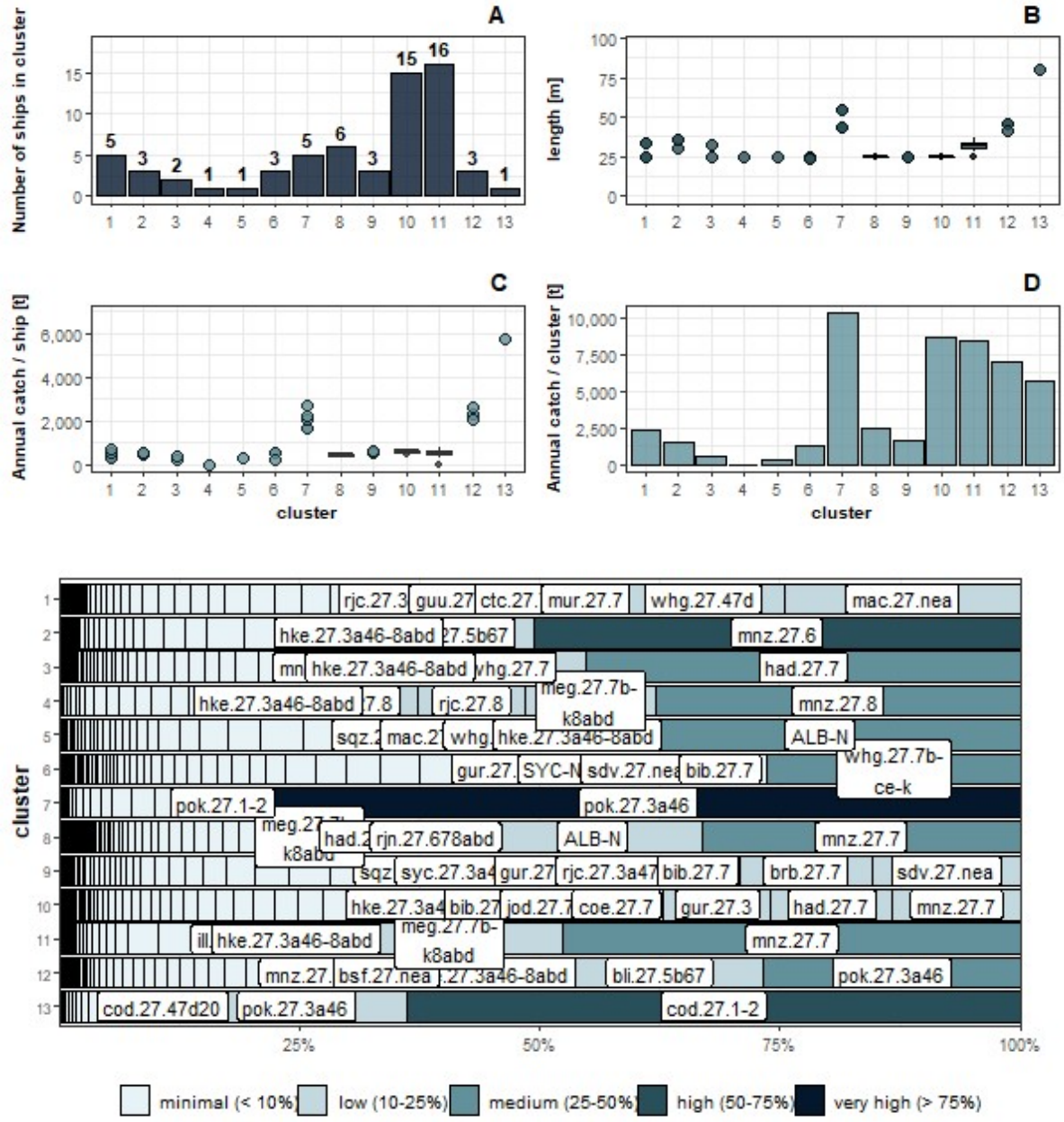


Figure 14: French demersal trawlers and seiners >24m operating in the Atlantic. Clusters with a broad range of catch profiles were achieved, reflecting the diversity of the fleet. Most clusters contain 6 vessels or less.

Ireland

The DTS-LSF group targets a broad range of stocks, thus representing a mixed fishery. Most of the stocks are demersal roundfish, but also some flatfish and small pelagics, like mackerel, sprat, or herring. Moreover, some lobster, brown cab, and nephrops are caught (Figure 15). The catch patterns are such that whenever the number of clusters is increased, only a small new cluster is generated. It appears worth trying to overcome this fragmented structure by grouping stocks or deriving a set of principal components from the overall catch composition. If typical species compositions can be observed from catch data by haul or trip, then this could give some guidance on how to group stocks.

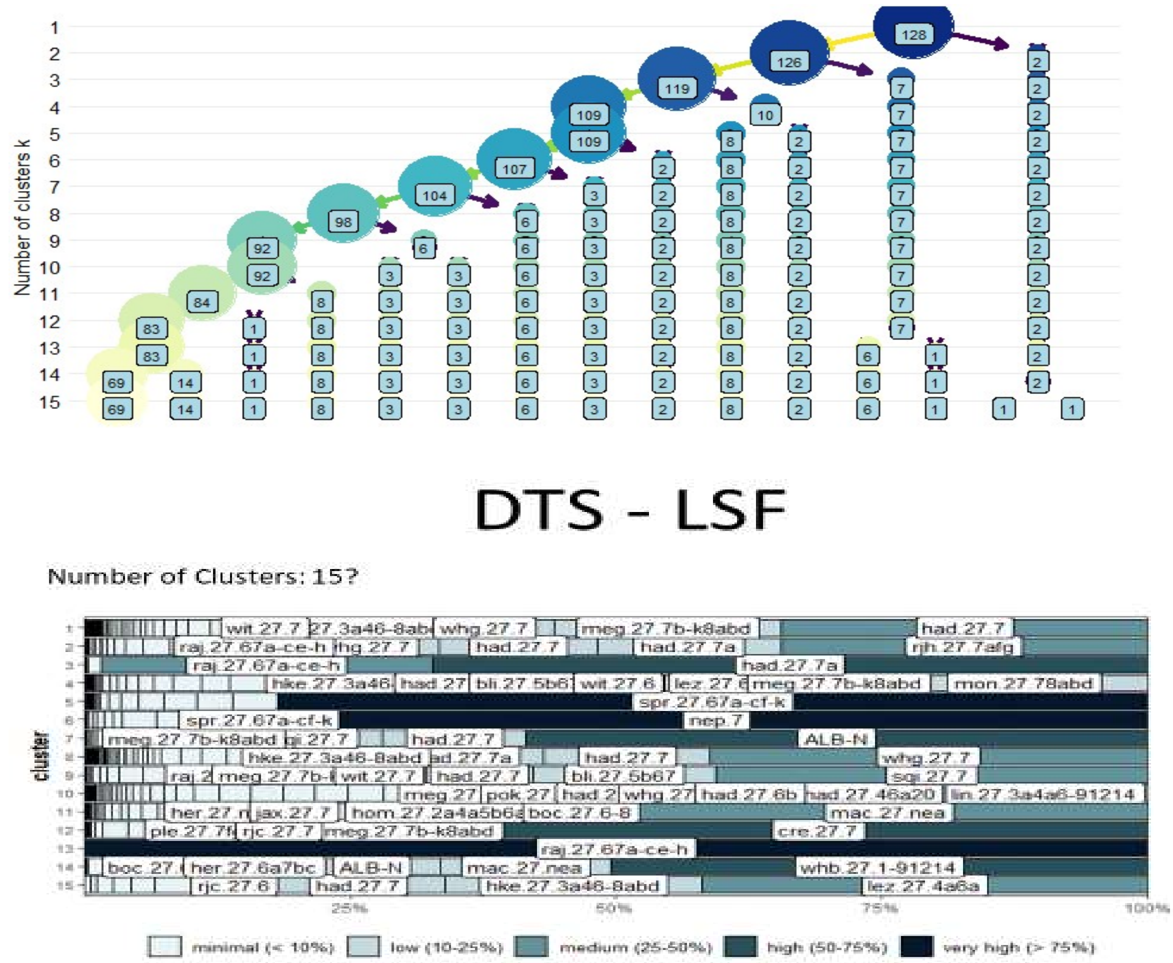


Figure 15: Clustering tree and catch composition for the Irish DTS-LSF fleet. A large number of stocks ans species is targeted, hence grouping clusters to segments is not straightforward.

Feedback from participants during the workshop

Participants applied the R tool to their fleet, partly introducing the suggested pre-segmentation steps. All participants managed to run the tool and to generate the support figures. The decision on the number of clusters remains challenging, but most participants managed to end up with a setting regarded suitable.

Some feedback topics from participants were:

- Future research should analyse year to year changes in gear reclassification of polyvalent vessels
- Time series might be compromised when altering the segmentation
- Question whether segmentation is stable across years
- Altering the segmentation would also affect legislation on fisheries management
- Main gear might be taken from fleet register for vessels without logbooks
- Bycatch threshold can be an issue in case of multispecies fishery
- Adjacent GSAs in the Med might have to be merged for same species
- A reduction of dimensionality in catch data sets via principal component analyses might aid applying the approach to large, diverse fleet data sets
- Activity level measure (e.g. days at sea) might be advisable to be included to separate low active vessels in the clustering procedure
- Cost structure is important aspect – how can the issue be thoroughly analysed given the fact that raw data is limited and sometimes problematic?
- Pre-segmentation should account for separating gears for which cost structure is known to be different
- Not adapted to fleets with incomplete individual-vessel declarative landings data
- Definition of groups often too large or too small
- Approach should better represent the vessel's operating strategy over the year
- 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.

Principles for alternative segmentation

The initial idea of considering a segmentation approach as alternative to the existing DCF segmentation that the DCF segmentation is often not sufficiently linked to management measures. As management measures are often related with stocks it appears advisable to focus on a closer link to stocks. This has been the main principle for the work on an alternative segmentation. It is, however, evident, that the number and composition of stocks exploited by vessels is often quite diverse. Therefore, the catch composition itself can only be one aspect of the segmentation. Further aspects include vessel size range and gear.

During the first two workshops it was concluded that DCF length classes appear to narrow in several cases while the fishing technique is sometimes not detailed enough. Therefore, a pre-segmentation, based on these characteristics, was recommended as topic for the third workshop. The ISSG on metiers had suggested using metiers for pre-segmentation, this approach was pursued as well. An advantage of the pre-segmentation step is that it is to some extent compatible with the existing DCF segmentation. Even though not stipulated in legislation, the

recent JRC data calls contain the"fishery" as an optional dimension of segment definition which would allow specifying typical catch composition of segments.

Targeted fishery vs. polyvalent fishery

Almost all analyses resulted partly in targeted fisheries focusing on specific stocks or stock assemblages and using specific gear, while other fisheries are polyvalent, targeting a broad range of species and changing the gear frequently. This reflects the reality within the fleet, and any segmentation approach should reveal these characteristics. Grouping vessels that perform targeted fisheries is rather straightforward and should raise little doubt. However, defining the transition from targeted to polyvalent fishery is challenging, no matter which method is being applied. Aspects to consider for this task could be, e.g. total landings or homogeneity in cost structure. A reduction of dimensionality in large catch data sets, e.g., via principal component analysis, might simplify the statistical analysis.

Segment or cluster size for assessing a segmentation approach

It has frequently been stated that the alternative segmentation approach based on catch composition would deliver clusters with too few vessels. In fact, the size of a cluster alone cannot be a criterion for assessing a segmentation approach. If a fleet is structured in a manner that individual vessels follow a unique fishing pattern, then it is quite obvious that any reasonable segmentation approach should highlight these individual vessels. This feature can also be observed in the current DCF segmentation, and it occurs in particular in the context of fisheries dependent information (FDI) data. However, small clusters/segments as such are not per se a problem – the main issue is that of confidentiality.

Small segments require a concept how to combine them to units for which the data can be published. This is independent of the segmentation procedure. The only existing concept for merging small segment is from SGECA 09/02, stating

> *"…MS should distinguish between segments considered for clustering as follows:*
>
> 1. *Important segments with distinct characteristics*
> 2. *Segments similar to other segments*
> 3. *Non-important segments with distinct characteristics*
>
> *Importance of fleet segments should be assessed in terms of landings (value and volume) and/or effort. Similarity should be demonstrated using expert knowledge on fishing patterns or on available data on landings and/or effort."*

Similarity in that sense has never been clearly defined, i.e. expert knowledge on fishing patterns has always been part of the approach.

The alternative segmentation approach, based on the landings composition, can, in fact, be a supporting tool to find vessels with similar catch characteristics.

Criteria for comparison of segmentation approaches

It has been occasionally stated that the alternative approach, based on landings, would not be applicable to certain fleets, referring not to technical aspects of the procedure, but to the results. The main drawback is said to be the number of clusters with few vessels. This issue is further discussed in the preceding paragraph.

However, it is quite obvious that criteria are needed to compare different segmentation approaches with respect to their fitness for purpose. In the report of the 2nd workshop some criteria for evaluating the suitability of a segmentation approach were defined:

- Connection to specific fisheries (high priority)
  Segmentation should aim for a closer link of segments to stocks or groups of stocks.
- Cost structure (high priority)
  Segments should combine vessels with homogeneous cost structure (reflected by indicators or proxies)
- Feasibility (high priority)
  The segmentation procedure has to be clear, doable without excessive extra burden, and repeatable.
- Compatibility (lower priority)
  It is desirable that the segmentation is compatible with an existing time series.

In order to address the Connection to specific fisheries, a comparative analysis was performed on the link between stocks and DCF versus alternative segments. This approach was made available to participants, but further results could not be added, given the time restrictions.

The analysis of cost structure is a challenging task and should be further elaborated. Main difficulties are seen in the data availability and variability. Small samples with high variability will quite likely either not fulfil the requirements for statistical testing or not indicate significant differences. Cost items should not be analyzed based on absolute values but rather on indicators or proxies, e.g. fuel cost/kwday or crew cost/value of landings. It should be further elaborated which indicators and cost items appear most relevant for comparison.

Feasibility is an aspect which is hard to quantify. Any more elaborate segmentation procedure will require some programming in the beginning, and qualified staff will be necessary.

Compatibility with the existing segmentation is desirable, but it is evident that any new segmentation approach will cause a break in time series as soon as the new approach is not just a propagation of principles of the existing approach.

First approach to analyze the suitability of segmentation principles

For the German case an analytical approach for analyzing the link between segments and stocks (Figure 16) was presented. Both graphs clearly indicate that substantially, as fewer segments are involved in the exploitation of certain stocks when using the alternative approach compared to the current DCF approach. In a pragmatic manner, the mean number of segments necessary to cover the entire catch can be compared: for the stocks selected in the left diagram, the mean was 5.36 for the DCF segmentation and 3.79 for the alternative segmentation. Basically, the same information is given by the right diagram: The curves are steeper for the alternative segments, hence, less alternative segments than DCF segments are required to represent the total catch of selected stocks.

The R code for creating these figures was shared with all participants.

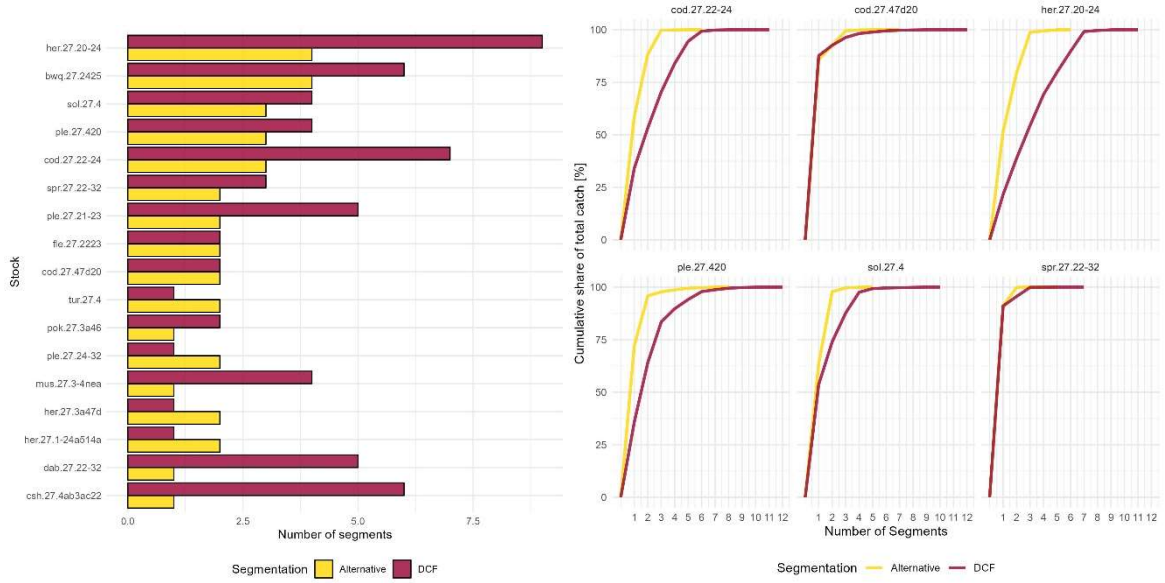


Figure 16: Link between segments and stocks. Left: number of segments involved in exploitation of selected, important stocks; right: cumulative catch for selected stocks with respect to the number of MS segments involved in the exploitation

Artificial Intelligence in the application of the approach

As indicated in the preceding workshop, an artificial intelligence (AI) tool was developed and submitted for publication. Some results of this tool when applying it to the German dataset were presented during the workshop. A learning dataset was generated. That dataset contained the alternative segment name which was assigned to vessels using the described procedure as well as some explanatory variables. It was subjected to a random forest (RF) algorithm. Part of the dataset was used for learning, the other part for testing. The algorithm correctly assigned alternative segments by more than 99% of the test dataset, which is regarded as "perfect agreement" between reference and prediction. Misclassifications were only present in segments with overlapping catch compositions and/or vessels switching segments between years.

This procedure is intended to be tested also on other fleets. Regardless of the principles of segmentation, it appears to be a very powerful tool for the automating the segmentation procedure, once an appropriate learning dataset has been provided.

Next steps

Following the discussion during the 3rd workshop, the alternative segmentation approach, based on landings, is functional, yet there are some aspects which require further clarification and standardization.

Concerning the gear, there are two main aspects to be considered, the level of gear resolution (e.g. métier level 4, or groups of gears) and the annual gear use pattern. This requires to be dealt with specifically since there are extreme cases like single gear being used throughout the year on the one hand and frequently altered gear on the other. These patterns need to be transferred into manageable groups. It is evident that vessels with polyvalent gear patterns are also more heterogeneous in catch composition and, therefore, more challenging to analyze using the alternative segmentation approach.

Another aspect that requires further standardization is the definition of length classes. In the context of pre-segmentation it was agreed that the DCF length thresholds are in most cases not appropriate, but some classification is deemed necessary, especially with regards to small-scale fisheries. However, an overall reduction of length classes remains a stated objective of the approach.

Moreover, the final step of the alternative segmentation, where segments are formed from clusters, is largely based on expert knowledge even though diagnostic plots and variables included in the FS-package support the expert. The expert decisions as observed during the 3rd workshop as well as in the workshops before are following similar principles, regardless of the fleet and member state. These principles should be transferred into a standardized decision support system included in the package.

Further progress is expected from a systematic aggregation of stocks, thus representing typical exploitation patterns. For this process, aid from regional experts, especially with regard to the Mediterranean, is highly desirable.

Annexes

1. List of participants
2. Workshop agenda
3. WS preparation document

Further documents are available on the RCG web repository.

Annex 1. List of participants

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

Name	Institution	Nation
Afentra Theocharous	Department of Fisheries and Marine Research	Cyprus
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
Erik Sulanke	Thünen Institute of Sea Fisheries	Germany
Irene Tzouramani	Agricultural Economics Research Institute, Hellenic Agricultural Organization-DEMETER	Greece
Jamal Roskam	Wageningen University & Research	Netherlands
Jane Burmanje	Bord Iascaigh Mhara	Ireland
Jarno Virtanen	Joint Research Centre	EU COM
Joonas Valve	Natural Resources Institute Finland (LUKE)	Finland
Jörg Berkenhagen	Thünen Institute of Sea Fisheries	Germany
Josefine Egekvist	DTU AQUA National Institute of Aquatic Resources Section for Monitoring and Data	Denmark
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
Myrto Iouannou	Department of Fisheries and Marine Research	Cyprus
Sarah Perry	Bord Iascaigh Mhara	Ireland
Sébastien Demaneche	Ifremer (Institut français de recherche pour l'exploitation de la mer)	France
Stamatis Mantziaris	Agricultural Economics Research Institute	Greece
Suzana Faria Cano	Direcção Geral de Recursos Naturais, Segurança e Serviços Marítimos	Portugal

Annex 2. Workshop agenda

Wednesday, 3.5.
13:00 Welcome and housekeeping
13:30 DEU presentation, stock analysis tool, AI tool
14:30 GRC presentation
15:15 Break
15:30 DNK presentation
16:15 NLD presentation
17:00 End of the first day

Thursday, 4.5.
13:00 ESP, CYP, FIN presentation
15:15 Break
15:30 FRA presentation
15:30 Compilation of results, next steps
17:00 End

Invitation / advice / request for preparation of the III. workshop on segmentation

We kindly ask you to run some analyses as described below which will serve as the basis for the III. workshop on segmentation (scheduled for the 1st week in May (2 half days, online).

In the context of the last WS it was concluded that a more detailed pre-segmentation of the MS can simplify the application of the novel segmentation approach and lead to more precise results of the clustering procedure. Therefore, it is suggested to perform this step and apply the clustering procedure using the updated, pre-segmented data set.

1) Perform a pre-segmentation of your fleet, using criteria which appear most appropriate for your fleet. In the last workshops we mostly used only the main gear for pre-segmentation, as defined in the current DCF legislation. For the third workshop, we propose a slightly more detailed, stepwise pre-segmentation.
   - As first criterion the supra-region can be introduced. This might be relevant for some fleets.
   - According to the discussion held at the previous workshop, a vessel size indicator can be included. The main aim is to separate small-scale, large scale and high seas vessels. These groups can be defined, e.g. by length thresholds at 12 and 40m, but this depends on the individual fleet. It could as well be defined by other length thresholds or by GT thresholds. In accordance with the DCF nomenclature, we suggest using SSF for small-scale fishery, LSF for large-scale fishery and DSF for distant water fishery.
   - A third criterion which advisable to be included is the gear aspect. If regarded sufficient, the DCF classification can be used, as in the preceding workshops. If the fleet composition is such that a more detailed classification is preferable, then this can be introduced. One possibility is to refer to the „Gear“ dimension of the segment dimension as provided for the AER (https://stecf.jrc.ec.europa.eu/dd/fleet - „Metadata“, paragraph 3.2.2.). A reference to Reg. 404/2011 (Annex XI) is suggested.

The principles of pre-segmentation are grossly based on your expert knowledge. If you think the pre-segmentation as used in the last WS (based on dominant gear only) is sufficient, then that version can be kept. This is all optional and just an opportunity to implement the suggestions provided by some participants in the last round and at the same time an attempt to be nearer to DCF nomenclature. The pre-segmented groups should, whenever feasible, be named in a way that the specifications can be derived more or less intuitively (e.g. „Med_LSF_PS“).

Please document the rationale behind each of these steps of extended pre-segmentation using the template provided and please get prepared for a brief presentation of your approach during the WS III.

2) Run the clustering procedure and define segments

Whoever has participated in the previous workshop will most likely be familiar with the amended R package and its application. Be aware that the manual has not been updated for the latest amendments, but Erik's presentation of March 2022 („WSII_Sulanke_Amendments_Presentation.pdf"),will give you further insight. It is to a large extent an implementation of features suggested by WS participants.

The next step again requires expert knowledge – it is the analysis of plots of catch composition (see slide 3 of the presentation) to define clusters. If clusters have very few vessels only, it should be checked if there are segments with similar catch profiles, e.g. with respect to the target assemblage. This might reduce the linkage between cluster and stock, but could be a compromise between that linkage and data protection for very small clusters.

This step can be double-checked with respect to the composition of technical characteristics within the clusters. Such characteristics, as described in slide 10 (length, catch volume and value) provide further insight into whether clusters can be regarded as similar and therefore could be merged, when necessary, for data protection purposes.

A good general characteristic to decide whether clusters should be merged into fleet segments is whether they represent directed or rather mixed fisheries. This aspect can be checked via the Herfindahl-Hirschman-Index (HHI) of the cluster catch composition plot. Unfortunately, the function for generating this plot is not described in the manual, but a description of its structure can be assessed in R directly via running *"HHI_plot"* from the console. The higher the HHI within a cluster, the more homogeneous is the fishery in terms of catch composition.

The experience of previous workshop showed that mixed fisheries are considerably more difficult to cluster than directed fisheries. If a pre-segmented data set contains a large fraction of vessels operating in mixed fisheries, it is advised to run multiple trials of the clustering procedure with different numbers of clusters to make sure that no distinct fishing strategy is overlooked.

At the end of the analysis you would ideally have a list of segments with distinct names and with the underlying principles being documented. Segments can be identical to the clusters determined, a merger of clusters or derived as a split of clusters. In order to remain within the proposed DCF-alike nomenclature, we propose either species names (e.g., Med_LSF_PS_PIL for large-scale purse sein pilchard fishery in the Mediterranean) or distinct identifiers for mixed fisheries (e.g., 27_DWF_OTB_MIX for mixed-demersal distant water bottom trawl fishery in FAO area 27).

<u>Shared drive</u>

Performing this analysis requires some expertise in using the R software. During the first two workshops substantial advice has been provided, and if you have participated you will be well prepared. You can find comprehensive documentation in the WS I and WS II folders under

https://cloud.thuenen.de/index.php/s/MFoafQGtqdrTHXC

The password is …

There is also a folder “WS III Fleet Segmentation” with subfolders where you find the templates for documenting your findings of both the pre-segmentation and the segmentation exercise. Please upload your files in the respective folders, indicating your MS in the file name.

Dates

Please provide the results using the templates by March 23, 2023.

Technical support

The fleet segmentation package is available for download in the GitHub directory.

The most convenient way to install it is by running

library(devtools) and then install_github("ESulanke/FleetSegmentation", build_manual = T, build_vignettes = T) from your console in R Studio.

Please remember to install devtools and rtools40 or rtools42 prior to the package installation.

The package contains an updated example data frame called example_fleetdata, which will be automatically loaded once you activated the package from your library.

You can use this data frame as an example for the variables your national fleet data set should contain.

If you have further questions or encounter errors, please contact erik.sulanke@thuenen.de

Annex 4. …