

The relationship between the number of trawls and long line sets at different spatiotemporal resolutions

by

Mike Bergh, OLSPS Marine Silvermine House Steenberg Office Park Tokai 7945

11 November 2022

1 Summary

HAKE/BG1 argues that a number of scientific issues need to be settled before a comprehensive assessment of the impact of a proposed 52.6% increase (see Government Gazette, 2021) in the allocation of the TAC to hake longlining can be said to have been carried out. One of these is the possibility that there is competition for space to fish at sea, between trawling and hake longlining operations, and that the outcome of this competition is to compromise trawl CPUE levels. Whether there is indeed such a process in effect must demonstrate a number of features in the spatiotemporal pattern of longline and trawl fishing, including the following:

- 1. An overlap in space of trawls and longline sets.
- 2. That at a certain operationally meaningful spatio-temporal scale there is a negative relationship between the number of trawls and the number of longline sets.

HAKE/P3 shows that there has, over the last 7 years, been a marked increase in the amount of hake longlining effort on the West Coast at depths and in locations typically fished by the deep-sea trawl sector. The demonstration in HAKE/P3 that this has occurred confirms that Condition 1 above is met. This document goes further and demonstrates that Condition 2 is met as well.

Results presented here show that (see [Figure 1\)](#page-6-0) hake trawl and longline operations have over the last 5-7 years taken place in the same locations when viewing these data over large temporal scales, in this case 2015 to 2020. Further results are presented (se[e Figure 6](#page-11-0) and [Table 1\)](#page-5-0) which demonstrate that when the spatio-temporal resolution is attenuated to the spatial scale of 10 to 60 km, and to a temporal scale of a single day, then there is a negative relationship between the number of trawls and the number of longline sets. In aggregate, the analytical results presented here are consistent with a scenario in which on any day there is competition for access to some spatial areas (represented as a K-means cluster), resulting in a negative relationship between the number of trawls and sets, as shown in [Figure 6](#page-11-0) and [Table 1.](#page-5-0)

2 Background

HAKE/P3 showed that there was a shift in longline effort onto the West Coast in 2014/2015, coupled with an increase in the depth and set length of hake longline fishing. The main question which follows is "Is there a negative interaction between the spatial distribution of longline sets and the spatial distribution of hake trawls". This document focusses on this point.

3 Methods

3.1 Datasets

OLSPS Marine has received logbook data for both the hake longline fishery and the hake offshore trawl fishery. The former covers the period 1994 to 2022, with only partial data available for 2022. The latter covers the period 1983 to 2020, but only data from 2000 to 2020 was imported for the data preprocessing and quality checking purposes. This document is concerned only with the period following the escalation of hake longline fishing on the West Coast in depths comparable to those where offshore trawling is taking place. For the purpose of analysis results presented here, that period has been taken to be calendar years 2015 to 2020 inclusive.

Extensive data audits have already been carried out on the hake and longline data provided to OLSPS Marine by DFFE to ascertain the quality of the data, especially w.r.t. the completeness of date, latitude and longitude and fishing depth information. 2.2% of the hake longline data had to be discarded because of quality considerations, but 24.8% of the trawl data were unusable, expressed as % of trawl drags. In terms of Hake green weight, 0.3% of hake longline catches and 19.5% of hake trawl catches had to be discarded. The reasons for this relate, as just stated, to the completeness of date, longitude and latitude and fishing depth information. The year level data quality w.r.t. date, GPS and fishing depth information is described in Appendix A. That appendix shows that for the time period which is the focus of this study, 2015 to 2020, the data quality is close to 100%, so that problems of missing data is not a serious concern here.

3.2 Spatial location of fishing event

There are two types of fishing events, either a hake longline set, or an offshore trawl. A simple depiction of these might be a line segment spanning the start and end location of these two events. Both may be more complex than a straight line. However, for the purpose of this document, the location of an event is represented as a point, calculated as the mid-point of the line segment just described.

3.3 Basic Proposition

It is proposed that if indeed there is some sort of spatial interaction between trawling and longline operations, then there would most likely be a relatively high spatio-temporal resolution at which this would manifest as a negative correlation between the number of trawls and number of sets.

For the high spatio-temporal resolution referred to (implying a relatively short time scale and a relatively small spatial scale), day as the time resolution seems to be appropriate, since any displacement of trawling by the presence of static gear is linked to longline lines which are freshly reset each day and their potential impact on the location of trawls on that day. As regards an appropriate spatial resolution, this is more difficult to decide. For this document a number of different options were explored:

1. Clustering the merged longline and trawl datasets using either 200, 300 or 500 clusters produced by a Kmeans clustering algorithm. Intuitively one expects that the spatial scale at which negative interactions manifest is related to the scale of fishing events, where sets are typically and at present in the order of 20 km in length. A degree of longitude or latitude is roughly 100 km in length at Cape Town latitude, so given the length of coastline involved, 200, 300 to 500 clusters seems appropriate.

2. Using latitude and longitude bins of fixed width to define a grid of spatial locations. Bin lengths between 0.03 degrees and 0,2 degrees were explored. At the latitudes typical of these data 1 degree of either longitude or latitude corresponds to roughly 100 km in length.

The following data preprocessing and analytical steps were implemented:

- 1. Append the hake longlining dataset to the hake offshore trawl dataset. Provide a common field name for midpoints by latitude and longitude, day designation, depth and catch.
- 2. Select data for years 2015 to 2020 inclusive, to ensure a complete year of data for both longlining and trawling.
- 3. Eliminate the few depth outliers above 900 metres.
- 4. Execute K means with either 200, 300 or 500 clusters.
- 5. Aggregate the catch for each day into cluster level across all years 2015 to 2020 and both fishing methods. Sort by cluster from the cluster with the most catch to the least catch. Select the clusters for which the aggregate catch encompasses 95% of the total catch. Designate these clusters as Selection A.
- 6. Select the day/cluster/fishing method level data by application of Selection A.
- 7. Arrange the data so that each day/cluster is adjacent to its corresponding number of trawls and number of longline sets, making sure to include records where, if this is the case, the number of fishing events is 0 for either one or the other fishing method.
- 8. For the data in the format described above, examine the relationship between the number of trawls and sets, in particular to ascertain if there is a negative relationship. This possibility was also explored using GLM techniques.

4 Results

A series of outputs have been produced illustrating the relationship between the number of trawls (the terms **trawls** or **drags** are used here but refer to the same fishing activity) and the number of hake longline sets. These are as presented in the following subsections.

4.1 Spatial overlap of trawls and longline sets; 2015 to 2020

[Figure 1. A scatter plot of the location of offshore trawls and hake longline sets according to the mid-point of fishing](#page-6-0) [events \(either sets in red or drags in light blue\), for 2015 to 2020 only, following a run of the K-means algorithm with](#page-6-0) 500 clusters, **and trimming away clusters [which all contribute less than any other clusters and collectively less than](#page-6-0) [5% of the total cumulative hake catch 2015](#page-6-0) to 2020**. The colour shown here distinguishes only sets from trawls and [are not intended to depict individual clusters.](#page-6-0) This figure shows the extent of overlap between hake trawling operations and hake longline operations.

[Figure 2. A scatter plot of the location of offshore trawls and hake longline sets by the mid-point of the fishing](#page-7-0) [event, for 2015 to 2020 only, following a run of the K-means algorithm with 500 clusters, without any trimming of](#page-7-0) the data. The colours [shown here distinguish only sets from trawls and are not intended to show clusters.](#page-7-0)

4.2 Depiction of K-means Clusters

[Figure 3. Plot of the location of fishing sets and trawls, 2015 to 2020, colour coded by K-means cluster, for the 200](#page-8-0) [cluster case. For this case 50 clusters encompass 95% of the gross catch and so only 50 clusters are represented in](#page-8-0) [this figure.](#page-8-0)

4.3 GLM Results

[Table 1. GLM based estimates of the slope of the covariate](#page-5-0) 'number of sets (n(SETS))' when the dependent variable [is the number of drags \(n\(DRAGS\)\). For these analyses a record represents a single Cluster_Day](#page-5-0) or Bin_Day

[combination which is a record because there was at least one drag or one set in that Cluster_Day or Bin_Day. For](#page-5-0) either of the Cluser Day only, or Cluster then Bin Day data preparation and spatial definitions, the data are first reduced to the [highest ranked \(by green weight catch\) clusters which make up 95% of the total hake catch over the](#page-5-0) [period 2015 to 2020 inclusive. This involves 121\(or 74 or 50\) clusters when the K-means algorithm used a](#page-5-0) predefined and fixed number of 500 (or 300 or [200\) clusters. GLMs were run using either the 'identity' or 'log'](#page-5-0)-link [functions, and the explanatory variable is either \(a\) n\(SETS\) and K-means clusters or \(b\) only n\(SETS\). Two bin width](#page-5-0) [options were explored when implementing the Bin_Days](#page-5-0) step, either 0.2 degrees or 0.05 degrees (where a degree at [the applicable latitude is approximately 100 km\), but only for the 200 and 500 cluster options.](#page-5-0)

4.4 Number of fishing events per Cluster_Day and per Day

[Figure 4. Histogram of the number of events on a Cluster_Day, where a Cluster_Day is only included](#page-9-0) in this histogram if there was at least one [trawl or one longline set, for the 200 cluster case, but for the 50 clusters which in this case](#page-9-0) [encompass 95% of the catch.](#page-9-0)

Figure 5 [Histogram of the number of events on a Day, where a Day is only included in this histogram if there was at](#page-10-0) least 1 trawl or one set [\(for the 200 cluster case, but for the 50 clusters which in this case encompass 95% of the](#page-10-0) [catch\).](#page-10-0)

4.5 Visualisation of negative relationships between n(DRAGS) and n(SETS), at Cluster_Day resolution

Figure 6. For the 200, 300 [and the 500 cluster cases, a plot of the number of drags \(n\(DRAGS\)\) vs the number of sets](#page-11-0) [\(n\(SETS\)\) at the resolution of Cluster_Days. The top panels show the average number of drags on the y axis for given](#page-11-0) [numbers of sets to clarify the average relationship. The bottom panels show all cases, except that larger dots](#page-11-0) represent more than one value [\(where a value is a Cluster_Day\).](#page-11-0)

4.6 Whether the evidence of spatial competition is an artefact

There is perhaps a valid concern that the foregoing results which demonstrate spatial competition between trawl and long line effort could be an artefact. To test this, an analysis based purely on the trawl drag records, which comprise about 93% of the fishing events in the combined trawl + longlining dataset, was carried out. 7% of these trawl records were randomly designated as 'longline sets', and the calculations carried out to produce the results shown in [Figure 6](#page-11-0) were repeated with this new semi-artificial dataset (although the calculation of clusters and their definitions were not updated). The plots are shown in [Figure 7.](#page-12-0) These show a positive relationship between the number of artificial sets (which are actually drags) and the number of drags, an opposite result to the result obtained using true sets. The indication from this is that the evidence demonstrating spatial competition (in e.g. [Figure 6\)](#page-11-0) is not an artefact, since the way that trawl locations relate to other trawl locations is very different to the way that trawl locations relate to real longline set locations.

5 Comments

[Figure 1](#page-6-0) provides a visual illustration that hake trawl and longline operations take place in the same locations when viewing these data over large spatial and temporal scales, in this case 2015 to 2020, and the entire fishing area.

However, when the spatio-temporal resolution is attenuated sufficiently, to the spatial scale of, say, 10 to 60 km, and to a temporal scale of a single day, then there is evidence of a negative relationship between the number of trawls and the number of longline sets – see e.g. [Figure 6](#page-11-0) which presents scatterplots that back this up, see also the GLM model results reported i[n Table 1.](#page-5-0) This shows that the presence of one type of fishing operation reduces the incidence of another kind of fishing operations in proximal locations. This is consistent with a scenario in which in any day there is competition for access to some clusters, resulting in the negative relationship between the number of trawls and sets evident in [Figure 6](#page-11-0) and [Table 1.](#page-5-0) This process is central to concerns being expressed by the deepsea trawl sector about a further 52.6% escalation in the amount of hake longline effort (see e.g. HAKE/BG1).

6 Acknowledgements

To DFFE for the provision of the hake longline and hake deep-sea trawl data.

7 References

- MARAM/IWS/2022/HAKE/BG1: Bergh, M., 2022a. Economic and other impacts of proposed changes to hake sectoral allocations. FISHERIES/2022/APR/SWG-DEM/07. 64 pp.
- MARAM/IWS/2022/HAKE/P3: Bergh, M., 2022b. Trends in the spatial distribution of hake long line fishing effort. FISHERIES/2022/AUG/SWG-DEM/10REV. 19 pp.
- Government Gazette, 2021. Draft Policy on the Allocation and Management of Hake Longline Fishery: 2021 (September).

MARAM/IWS/2022/HAKE/P4

Table 1. GLM based estimates of the slope of the covariate 'number of sets (n(SETS))' when the dependent variable is the number of drags (n(DRAGS)). For these analyses a record represents a single Cluster_Day or Bin_Day combination which is a record because there was at least one drag or one set in that Cluster Day or Bin Day. For either of the Cluser Day only, or Cluster then Bin Day data preparation and spatial definitions, the data are first reduced to the highest ranked (by green weight catch) clusters which make up 95% of the total hake catch over the period 2015 to 2020 inclusive. This involves 121(or 74 or 50) clusters when the K-means algorithm used a predefined and fixed number of 500 (or 300 or 200) clusters. GLMs were run using either the 'identity' or 'log'-link functions, and the explanatory variable is either (a) n(SETS) and K-means clusters or (b) only n(SETS). Two bin width options were explored when implementing the Bin_Days step, either 0.2 degrees or 0.05 degrees (where a degree at the applicable latitude is approximately 100 km), but only for the 200 and 500 cluster options.

Figure 1. A scatter plot of the location of offshore trawls and hake longline sets according to the mid-point of fishing events (either sets in red or drags in light blue), for 2015 to 2020 only, following a run of the K-means algorithm with 500 clusters, **and trimming away clusters which all contribute less than any other clusters and collectively less than 5% of the total cumulative hake catch 2015 to 2020**. The colour shown here distinguishes only sets from trawls and are not intended to depict individual clusters.

Figure 2. A scatter plot of the location of offshore trawls and hake longline sets by the mid-point of the fishing event, for 2015 to 2020 only, following a run of the K-means algorithm with 500 clusters, without any trimming of the data. The colours shown here distinguish only sets from trawls and are not intended to show clusters

Figure 3. Plot of the location of fishing sets and trawls, 2015 to 2020, colour coded by K-means cluster, for the 200 cluster case. For this case 50 clusters encompass 95% of the gross catch and so only 50 clusters are represented in this figure.

Figure 4. Histogram of the number of events on a Cluster_Day, where a Cluster_Day is only included in this histogram if there was at least one trawl or one longline set, for the 200 cluster case, but for the 50 clusters which in this case encompass 95% of the catch.

Figure 5 Histogram of the number of events on a Day, where a Day is only included in this histogram if there was at least 1 trawl or one set (for the 200 cluster case, but for the 50 clusters which in this case encompass 95% of the catch).

Figure 6. For the 200, 300 and the 500 cluster cases, a plot of the number of drags (n(DRAGS)) vs the number of sets (n(SETS)) at the resolution of Cluster_Days. The top panels show the average number of drags on the y axis for given numbers of sets to clarify the average relationship. The bottom panels show all cases, except that larger dots represent more than one value (where a value is a Cluster Day).

Figure 7. For the 200, 300 and the 500 cluster cases, a plot of the number of drags (n(DRAGS)) vs the number of sets (n(SETS)) at the resolution of Cluster_Days, **where for this case the sets are a random selection of 7% of the trawls and the true set level information has been excluded from the calculations**. The top panels show the average number of drags on the y axis for given 'numbers of sets' to clarify the average relationship. The bottom panels show all cases, except that larger dots represent more than one value (where a value is a Cluster_Day) as indicated by the legend.

Appendix A. An assessment of the quality of the trawl and longline data made available

Table 2. The amount of catch i.t.o green weight that is available in the trawl dataset 'Raw data', followed by the amount of catch for which different combinations of the drag date, fishing depth and position (GPS) information is available.

Table 3 . The number of records that are available in the trawl dataset 'Raw data', followed by the number of records for which different combinations of the drag date, fishing depth and position (GPS) information is available.

MARAM/IWS/2022/HAKE/P4

Table 4 . The amount of catch i.t.o green weight that is available in the longline dataset 'Raw data', followed by the amount of catch for which different combinations of the set date, fishing depth and position (GPS) information is available.

MARAM/IWS/2022/HAKE/P4

Table 5 . The number of records that are available in the longline dataset 'Raw data', followed by the number of records for which different combinations of the set date, fishing depth and position (GPS) information is available.