# External Expert Report

# Review of Oman s data collection system and statistics and retrospective data analysis for 2014-2023

**Muscat, 10th November 2024**

# Contents

## 1. Purpose of the document

In the last meeting of the 26th Session of the Working Party on Tropical Tunas Data Preparatory Meeting, Oman reported that "*is internally reviewing its sampling protocol, with adjustments to data from 2014 where catches may have been underestimated"*, in particular, in relation to yellowfin catches, as was included in the Minutes of the Meeting.

During the last months and until now, the Department of Fisheries Statistics and Information of the Ministry of Agriculture, Fisheries and Water Resources, Directorate General of Fisheries Research, (hereinafter, the "MAFWR"), has been working on this task with a view to present a full report to the WP of Data Collection and Statistics to take place in Cape Town by the end of November 2024.

In this respect, an Omani Delegation of the MAFWR travelled to the IOTC headquarters at end of July 2024 and held working sessions with the IOTC Data Officers.

As a next step MAFWR hired the services of an external data expert, Dr Constantine Stamatopoulos, a senior fisheries consultant in fisheries data and statistics (hereinafter, the "External Expert"), whose CV is attached. His terms of reference were to prepare and present a Preliminary Report at the 26th Session of the WPTT (Seychelles, 28 October – 2 November 2024), conduct a retrospective data analysis for 2014-2023 and present a full Report at the WPDCS and the Scientific Committee (end November - early December 2024).

This document has been prepared by the External Expert with the support of the MAFWR, with a view of presenting the results of a review of Oman's fisheries statistical programme with specific focus on artisanal fisheries. The review started in August 2024. It evaluated the current data collection system and verified its compliance with regional and international standards. It also evaluated the related Oracle database and the statistical reports resulting from the collected information and data. Based on these findings a catch/effort analysis has been conducted, followed by a retrospective revision of catch/effort figures for 2014-2023.

## 2. Background

With an extensive coastline stretching for 3,165 kilometers, Oman is uniquely positioned to capitalize on its marine resources. The fisheries sector stands out as a cornerstone of the Omani economy, not only as an important source of income but also as a cultural mainstay. For many individuals fishing is more than just a livelihood; it represents a way of life and sustains a rich maritime heritage. Consequently, Omani authorities have always shown great interest in sustaining and conserving this activity.

Statistical fisheries data collection in Oman began in 1984, as part of a joint Omani-American committee established during that period to foster cooperation between the two countries. These early statistics laid the foundation for the methodologies based on a sampling approach to estimate total fish production by month, region, and fish species, with specific focus on artisanal and coastal fisheries This development involved financial resources, human capabilities, and technical systems for data collection, storage and analysis. Fisheries statistics personnel received specialized training both within Oman and abroad.

Over time the fisheries statistical monitoring programme underwent multiple stages of development and improvement. At present the programme relies on a robust infrastructure involving human resources for data collection, processing and analysis, advanced data collection systems, and well-structured databases. Additionally, statistical tools and techniques have been introduced to assess the current statistical and computational practices and verify statistical results through parallel methodologies. Automatic diagnostic procedures have been introduced to provide regular indicators related to data consistency and reliability.

Intensive training programmes have also been implemented, focusing on practical marine statistics to address current needs for accuracy, inspection, and changes. Several of these programs were executed: Athens in July-August 2010, University of Reading in England from May 28 to June 6, 2013, and in the Netherlands.

Oman is witnessing the rise of an emerging industry within the fishing sector. This includes the development of processing and canning factories, which add value to the raw marine products harvested from the sea, and aquaculture projects, which aim to cultivate fish and seafood in a controlled environment, contributing to food security and diversification of marine resources.

In line with these developments, Omani fishermen have become increasingly professional, adopting more sophisticated fishing techniques and practices. There is a concurrent push to modernize the fishing fleet to meet the demands of a growing market and enhance the efficiency of operations. However, Oman is committed to balancing modernization with sustainability. The principle of sustainable fishing activity is paramount, ensuring that marine ecosystems are preserved for future generations while allowing the current population to benefit from the ocean's bounty.

In this transformative period, Oman's approach is to harness the potential of its fisheries sector in a way that is both environmentally responsible and economically viable, aiming to create a legacy that supports both the nation's prosperity and the conservation of its natural resources.

## 3. Present state of the data collection system

Most of the fleet comprises artisanal vessels (98%), some smaller coastal (1%) and commercial (0.8%) vessels (as per 2015 fleet census). The artisanal vessels are made up of approximately 13,000 skiffs undertaking day trips, and about 600 dhows (9-24m length) undertaking longer trips of 1-10 days.

To date, onboard observer and logbook schemes are in the early stage of implementation. Plans have been made for a stepwise implementation of logbooks for certain fleet segments. At present data collection on artisanal and coastal fisheries is entirely based on sampling. Data for the industrial fleet concerns a few longliners and 3 industrial purse seiners. This information is based on real data from VMS and log-books, including e-logbooks on board the purse seiners.

The port-based sampling system involves 42 data collectors who cover over 156 landing sites. Data collection records fish species and weights as they are landed and includes first-sale prices, as well as average weight of individuals. An electronic data recording device is used that has an inbuilt catalogue of species to be used for species identification; this practice is reported to have improved in recent years the detail of information reported.

Data collection on fishing effort is performed independently by recording the fishing activities of fishermen on the previous day (1 if fished or 0 if otherwise). For some fleet segments, notably dhows, the effort query refers to the total days worked during the reference month.

Sampling operations are conducted according to well-prepared guidelines and protocols and are supervised effectively. The quality of landings data submitted to the database is generally good. An important point concerns the trip duration which is missing in some places, on the assumption that it will automatically be set to 1 by the system. This practice can impede the occasional use of trip durations that are shorter than one day (case of sampling multiple landings of the same gear during the day).

With regards to effort samples, the quality of data submitted to the database, is also good. However, there is an occasional mix of monthly effort with boat proportions. Such a mix is not arithmetically wrong, but it affects the calculation of accuracy for effort samples.

Given the significant increase in production over the period 2019–2021, as reported to the IOTC, and the high importance of strategic species such as tuna and Indian oil sardine, the Ministry may consider conducting limited census-based data collection for such fisheries. For the other fisheries, it should simplify the effort scheme by generalizing the monthly effort approach; an action that would increase sampling accuracy and decrease data collection costs.

## 4. The Oracle database

The structure and contents of the database are of good quality. There are some minor problems with the system referentials (tables) where some descriptions are misspelled; in other cases, different codes correspond to the same description (duplication of names). Data integrity and data security are issues that should be paid particular attention; additional backup functions performed locally would reduce the database dependence on the data security services provided by the server team.

With respect to the reliability of effort and catch samples, it is recalled that the Statistics Directorate has been aiming at a sampling accuracy level of 95%. This has proved to be a little too ambitious for the size and frequency of sampling operations; nevertheless the achieved accuracy level of 90% over the period 2017-2022 is quite acceptable and conforms to regional and international standards.

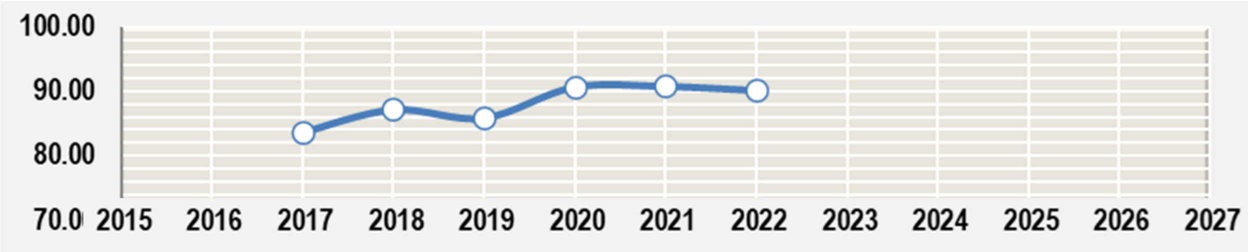



**Figure 1. Overall sampling accuracy for 2017-2022.**

In examining the Oracle script that extrapolates sample fishing effort, the review observed that the estimator uses an almost unrestricted approach whereby only two days of reduced or zero fishing activity (weekends, bad weather, holidays, market days, etc.) are excluded. Such a practice tends to over-estimate fishing effort; this overestimation is not dramatic but it nevertheless introduces a constant positive bias in the estimation. To be noted that small variations of fishing effort can result in visible production increases in fisheries with high catch rates, such as small pelagic nets and beach seine. The Statistics Group may consider the highly recommended alternative of generalising the monthly effort scheme (at present used only partially), so as to permanently remedy the parametrisation problem in effort estimation.

The above considerations only concern the absolute size of total catch and effort and not the relative differences between the peak years 2020-2021 and the rest. A more detailed analysis is given in the coming sections.

## 5. Verification process

The objective of the verification process undertaken by the External Expert was to confirm the compliance of the Oracle estimator with recognized generic approaches for estimating fishing effort and catch. For this purpose the review made use of the standard FAO utility ArtWeb and of a programmed Excel workbook prepared specifically for the verification exercise. Both tools were parametrized to operate under the same conditions and assumptions built into the Oracle estimator.

The verification exercise was conducted by means of three parallel applications on the same datasets for 2017-2022. As shown in Figure 2 the resulting estimates are almost identical. This fact permitted the utilization of FAO ArtWeb as a versatile instrument for the revision of effort and catch figures, thus facilitating the review process.

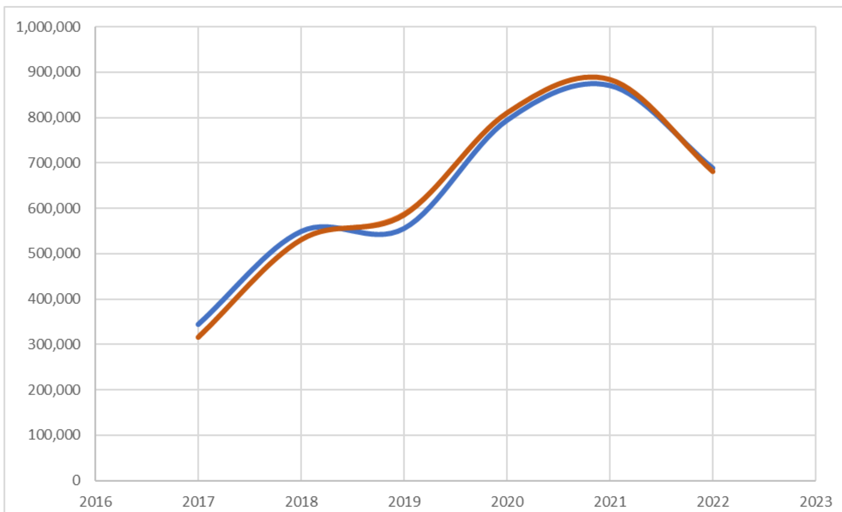



**Figure 2. Results of the verification process for 2017-2022. The blue line represents total catch estimates resulting from the Oracle database. The red line represents the results of FAO ArtWeb and Excel.**

## 6. Observations on the production figures of 2017-2022

Total production amounted to around 348,000 Tons in 2017 with an increase of approximately 24% compared to 2016. In 2021 production reached record levels (Figure 2 – blue line), representing a significant increase compared to reported catches of the previous year and continuing a rising trend observed since 2017. These figures were obtained by the original Oracle estimator and verified against the verification utilities FAO ArtWeb and Excel.

Detailed within the overall production are the species catches. Indian oil sardine escalated from 82,654 tons in 2015 to 430,243 tons in 2020, with a pronounced 56% increase from 2019.

Similarly, yellowfin catches surged from 14,957 tons to 68,815 tons in the same timeframe, marking an 86% jump from the preceding year.

There is a clear increase in the production figures for 2020 and 2021 which raises the following two basic questions:

(a) Whether the relative differences between 2020-2021 and the rest of annual productions are indeed real and could be explained statistically.
(b) Whether the absolute figures for 2017-2022 could be revised and still maintain the relative differences mentioned above.

Concerning point (a) the review only made use of factors that stem from the statistical system itself and do not constitute assumptions. In other words, the review did not consider other possible factors affecting production, such as increased abundance, new markets, decreased fishing pressure in neighbouring waters, etc.

The first statistical factor related to point (a) concerns intensified fishing activities, i.e. more boats out fishing, while the number of boats is unaltered. This is shown in Figure 3.

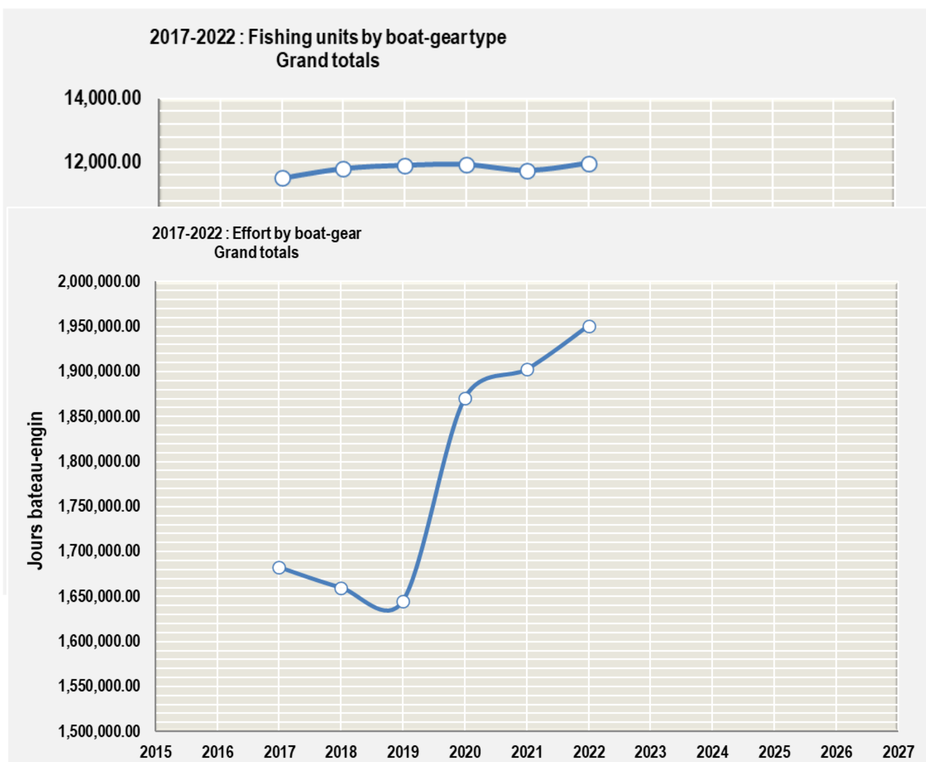



**Figure 3. The upper plot shows the mobilization of fishing vessels during 2017-2022, which is about constant. The lower plot clearly highlights the intensification of fishing activities over 2020-2022.**

The second factor related to point (a) concerns higher catch rates. Field agents have reported that during the peak years daily trips lasted longer, thus increasing the daily yields. Figure 4 illustrates this fact. To be sure there may exist other external factors that have contributed to the higher catch rates in 2020 and 2021. One example is the period of the COVID crisis during which the government allowed the fishery sector to continue operating while other sectors were totally closed (such as construction). This action brought more fishermen at sea.

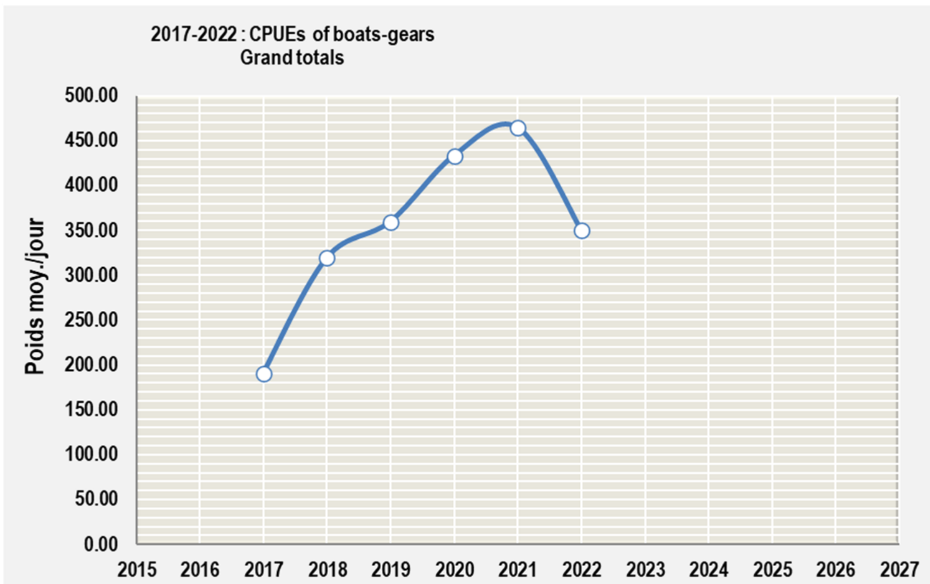



**Figure 4. Overall catch rates (all boats-gears and species) for 2017-2022. The plot shows a clear increase over 2020-2021.**

High CPUE values (in kg/day) are not due to “statistical noise”. There is consistency in the appearance of high catch rates in the samples as shown in Figure 5.

**MAX-MIN values : 12 / 2017**

Records with zero or without value: 0 / 1243

| Sample no. | Site / Port | Boat-Gear type | Species | CPUE |
|---|---|---|---|---|
| 1016 | Taqah | Beach Seine | | 19025 |
| 0906 | Taqah | Beach Seine | | 12972 |
| 1017 | Taqah | Beach Seine | | 12972 |
| 0281 | Daba Al Baiya | Beach Seine | | 10500 |
| 0489 | Daba Al Baiya | Beach Seine | | 10150 |
| 0280 | Daba Al Baiya | Beach Seine | | 9975 |

**Figure 5. Maximum CPUE values in December 2017.**

At the moment it is not certain whether the peak figures of 2020 and 2021 are indeed part of a trend or are circumstantial. Estimates for 2022 and and 2023 suggest that after a decline in 2022, production resumes an upward trend.

## 7. Revision of statistical data for 2014-2023: Methodology

As mentioned earlier the Oracle system uses an almost unrestricted extrapolation approach for fishing effort in which days of reduced or zero fishing activities are not accounted for sufficiently. As a result the current effort estimation procedure generally over-estimates fishing effort.

A more detailed discussion on this point is given below.

It is recalled that total catch is computed by means of the formula:

Estimated Catch = (Estimated Effort) x (Sample CPUE)

When fishing effort is over-estimated then, due to high CPUE values in some fisheries (such as the small pelagic net and beach), the associated production figures will be inflated.

Here we examine the formula for estimating fishing effort:

Estimated Effort = PBA x B x A

where:

PBA = Probability Boat Active (from samples)
B = Number of boats (fishing units) in fleet census.
A = Active Days.

A is calculated from the calendar days of the reference month by taking out days of zero activity.

Figure 6 explains schematically the impact of the temporal extrapolation factor A on effort estimation.

Days: 30

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | | 1 | 1 | 1 | 1 | 1 | 0 | | 1 | 1 | | 0 | 1 | 0 | | 1 | 0 | 1 | 0 | 1 | 0 | | 1 | 0 | 1 | 1 | 0 | 1 |
| 2 | 1 | 1 | | 0 | 1 | 0 | 0 | 1 | 0 | | 1 | 1 | | 1 | 1 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 1 | 1 | 0 | 0 |
| 3 | 1 | 0 | | 1 | 1 | 1 | 1 | 1 | 0 | | 0 | 0 | | 1 | 0 | 0 | | 1 | 0 | 1 | 0 | 1 | 0 | | 1 | 0 | 0 | 1 | 0 | 1 |
| 4 | 1 | 1 | | 0 | 0 | 0 | 0 | 0 | 0 | | 1 | 1 | | 1 | 1 | 0 | | 0 | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 1 | 0 | 0 | 0 |
| 5 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 1 | | 0 | 0 | | 0 | 0 | 1 | | 0 | 1 | 0 | 1 | 0 | 1 | | 0 | 1 | 0 | 0 | 1 | 0 |
| 6 | 0 | 1 | | 1 | 0 | 1 | 0 | 0 | 0 | | 1 | 1 | | 0 | 1 | 0 | | 1 | 0 | 1 | 0 | 1 | 0 | | 0 | 0 | 1 | 0 | 0 | 1 |
| 7 | 1 | 1 | | 0 | 0 | 1 | 1 | 0 | 1 | | 1 | 1 | | 1 | 1 | 1 | | 0 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 0 | 1 | 1 |
| 8 | 0 | 0 | | 0 | 0 | 0 | 1 | 0 | 1 | | 0 | 0 | | 0 | 0 | 1 | | 0 | 1 | 0 | 1 | 0 | 1 | | 1 | 1 | 0 | 0 | 1 | 0 |
| 9 | 1 | 0 | | 1 | 1 | 0 | 0 | 1 | 1 | | 0 | 0 | | 1 | 0 | 1 | | 1 | 1 | 0 | 1 | 0 | 1 | | 0 | 1 | 0 | 1 | 1 | 0 |
| 10 | 0 | 0 | | 1 | 1 | 1 | 1 | 1 | 1 | | 0 | 0 | | 0 | 0 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 0 | 1 | 1 | 1 |

**Figure 6. A theoretical example of effort estimation from sample boat activities.**

In the theoretical example illustrated in Figure 6 there are 10 boats that operated during April 2024. Here the number of boats (B) is 10 and the calendar days (D) are 30. Each cell of the 10 x 30 matrix represents the status of a boat on that day: 1 if active or 0 if otherwise.

There are four Fridays (coloured yellow and grey) and one day of bad weather (in brown). During these five days fishing activities are zero.

By summing all cells containing 1 we find that the total effort is 125 days.

Sampling of boat activities occurred on days 2, 5, 8, 15, 19, 22, 26, 29 (columns marked in black).

The resulting sample PBA is 0.5, since the sum of boats found active is 40 and that of boats examined is 80.

Recall that there are four Fridays (in yellow and grey) and one day of bad weather (in brown).

The correct estimation of effort E uses PBA=0.5, B=10 and **A**=30-5=**25** active days.

$$E = 0.5 \times 10 \times 25 = 125 \text{ days.}$$

If instead the factor A is set to 28 days, the effort will be over-estimated:

$$E = 0.5 \times 10 \times 28 = 140 \text{ days.}$$

Setting up the active days A requires some extent of extra data collection. Data collectors specify active days empirically, based on the knowledge of local conditions in sites and wilayats. From the locally specified active days, a compound temporal factor A at regional (=estimation) level is computed using the numbers of boats in the fleet census as weighting factors.

<u>Compound A at regional level</u>

A = SUM (Boats x locally determined A) / SUM(Boats)

To be noted that the current system does not collect such data for determining A.

In the effort estimation formula it determines A as D - 2:

where D represents calendar days (31, 30, 29, 28) in the reference month.

Estimated Effort = PBA x B x (D-2)

Likewise the revision approach was unable to determine compound temporal extrapolation factors due to the lack of locally collected data for A. However it made use of a more conservative approach in which calendar days were reduced by 5 days instead of 2. These 5 days accounted for 4 Fridays and 1 day of bad weather, holiday, etc. The choice of D-5 was based on the following considerations:

- Since a month has 4 weeks, it would make sense to count 4 Fridays instead of 2, and adding just one day of general zero activity;
- There is some statistical evidence supporting the choice of a more conservative factor A. Using occasional “mixed” effort approaches in which monthly effort samples were collected along with boat activities, and by separating the related PBA’s, it was found that those for monthly effort were (as expected) systematically lower to a ratio of about 0.83. It is recalled that the ratio (D-2) / (D-5) is 0.89.

The revised estimation process resulted in lower effort and catch, as shown in the coming sections.

To be noted that the revision approach did not have to re-estimate all months and years in the Oracle database. It simply adjusted existing effort and catch estimates using the ratio:

R = (D-2) / (D-5)

In fact, the current effort estimate is:

Ec = PBA x B x (D-2)

The (more) conservative estimate is:

Ea = PBA x B x (D-5)

By dividing the second expression by the first we obtain:

Adjusted effort Ea = (Current effort estimate Ec) x (D-2) / (D-5)

This adjustment has applied on the entire period 2014-2023. For 2017-2023 only the downward adjustment applied. The right part of the plot in Figure 7 presents schematically this revision.

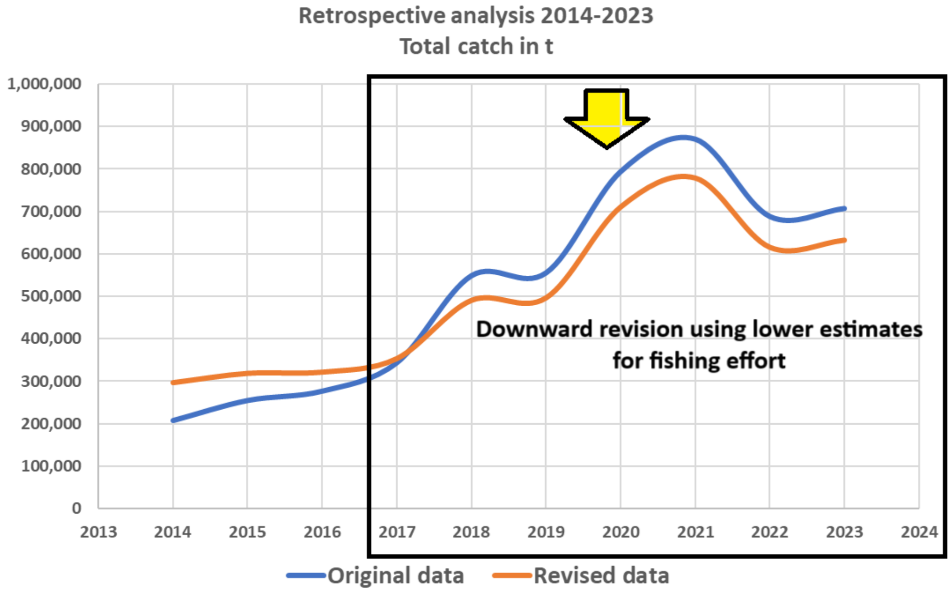



**Figure 7. Adjustment of production figures for 2017-2023 (right plot). The blue line describes the original Oracle data, whereas the red line refers to revised data.**

For the years before 2017 the retrospective revision followed a mixed pattern of two revision types.

The first (downward) revision corresponds to the conservative effort estimation described earlier and which applied for the entire period of 2014-2023.

The second (and upward) adjustment made use of two positive factors. The first factor derived from hindsight knowledge regarding the fishing fleet. This factor increased the number of outdated fishing units during 2014-2016, thus increasing fishing effort and production.

The second upward factor derived from information concerning statistical gaps in the early operational steps of the current system, during which certain regions and/or boats-gears were not covered appropriately. The information was made available by the estimation logs that signaled gaps in statistical coverage during the calculation of fishing effort and catch.

The left part of the plot in Figure 8 presents schematically this mixed approach. The prevailing factor is positive (green arrow).

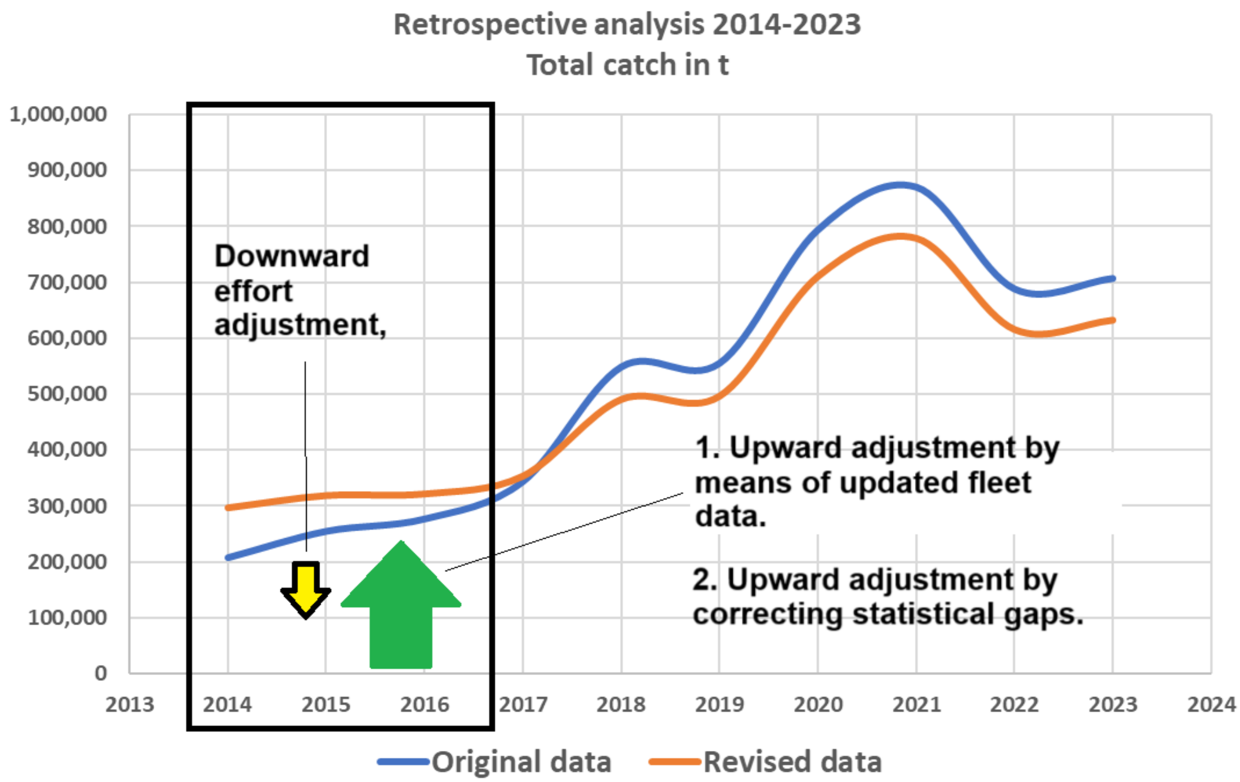


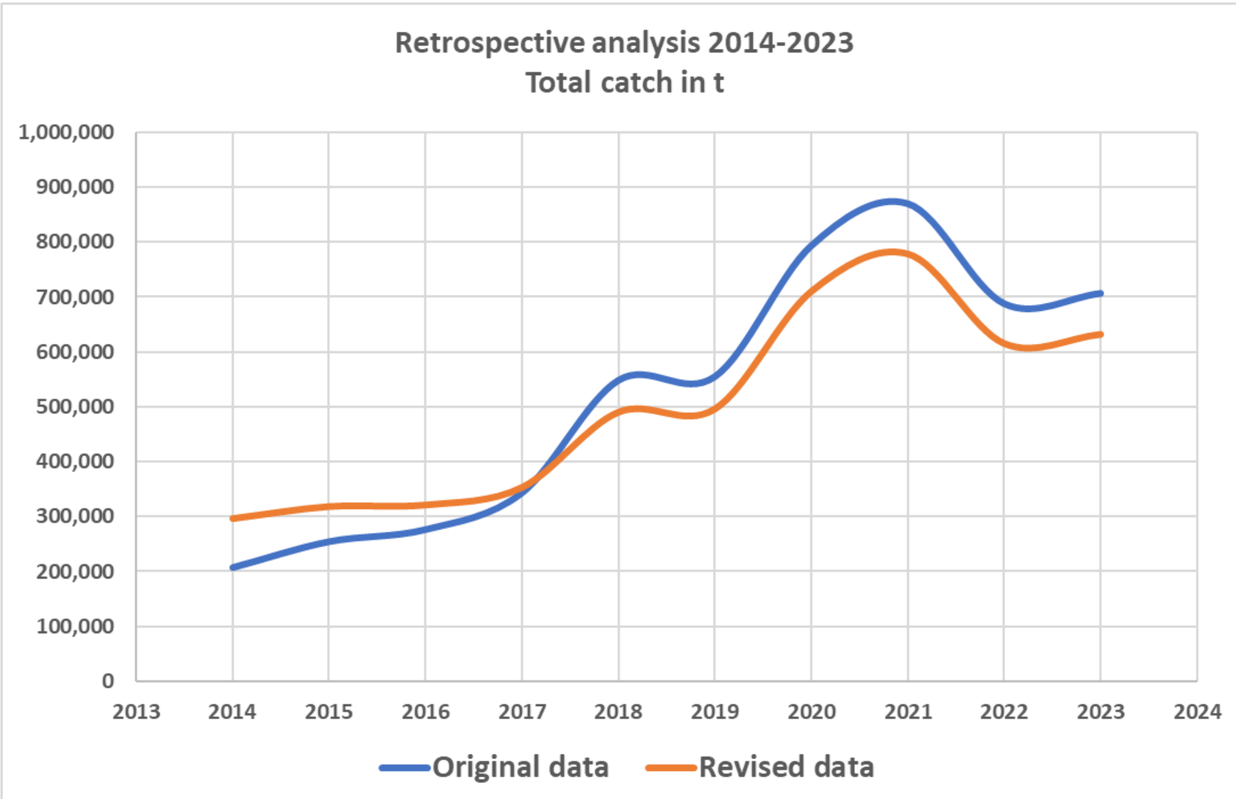

**Figure 8. Adjustment of production figures for 2014-2016 (left plot).**

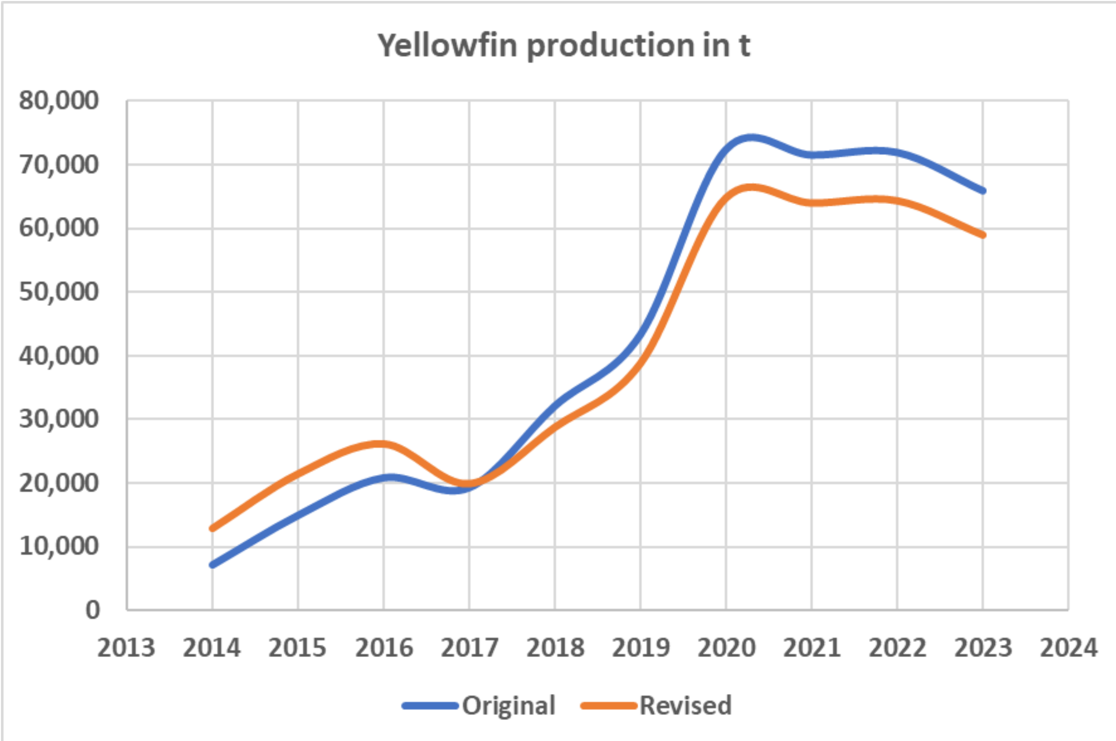
Data revision concerns the artisanal and coastal data.

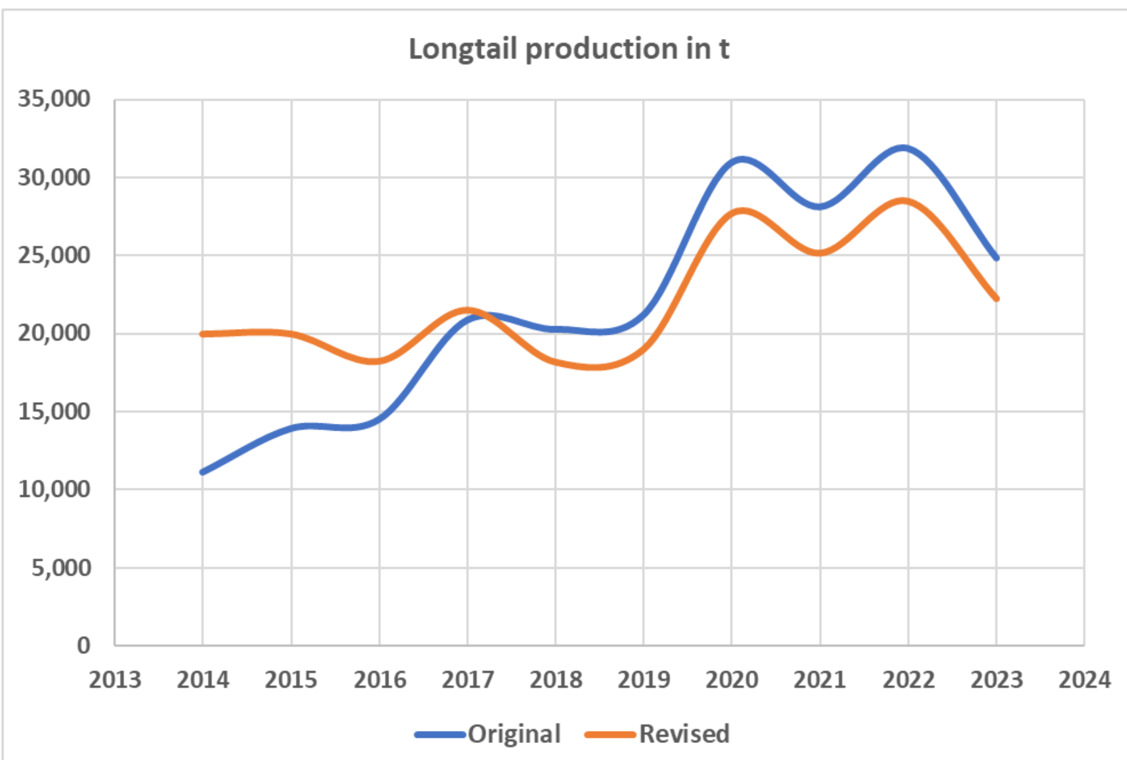
Catch revisions apply to total effort and catch by region and boat-gear category. Species catches are revised accordingly, based on species proportions within each region and boat/gear combination.

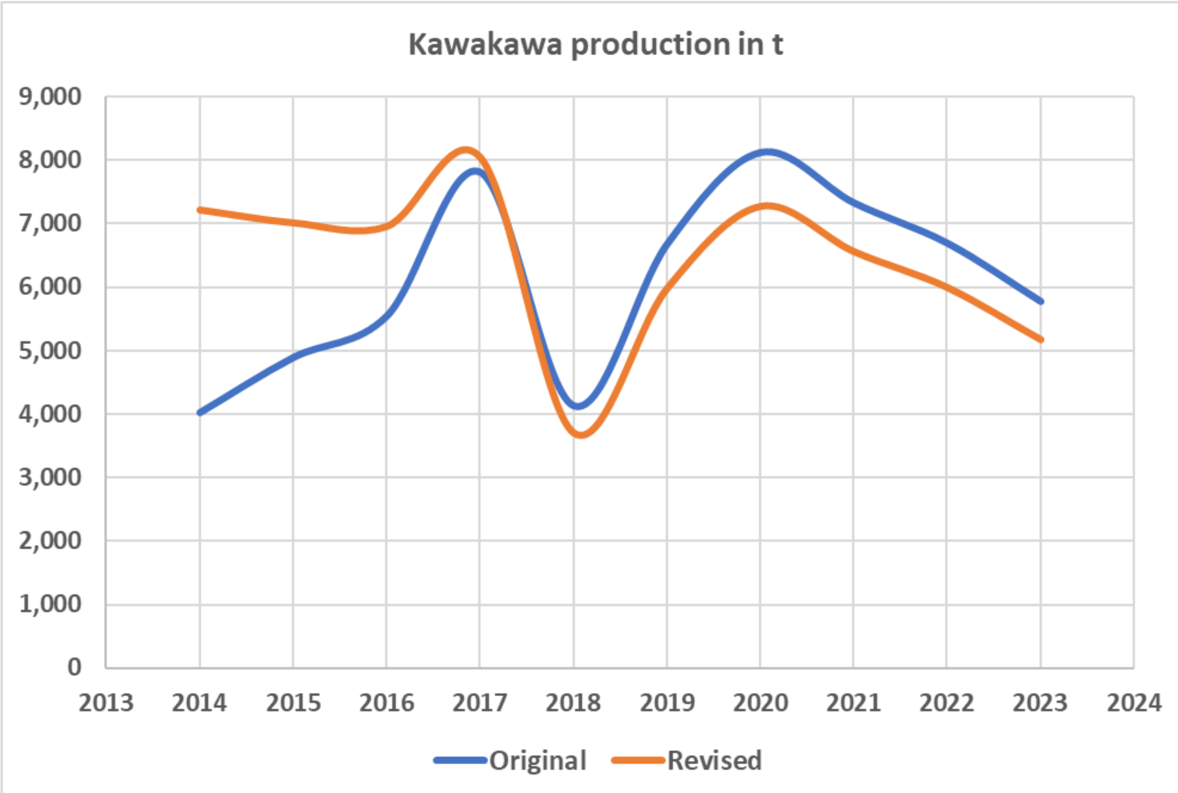
In addition to re-estimated catches of artisanal fisheries the revised official tables that will be submitted to IOTC and FAO, will include real catches declared by the industrial fleet in recent years, and which are based on log-book declarations.

## 8. Selected examples of the retrospective analysis




Yellowfin production in t
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024
Original
Revised


Longtail production in t
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024
Original
Revised




## 9. Some facts from figures (original data)

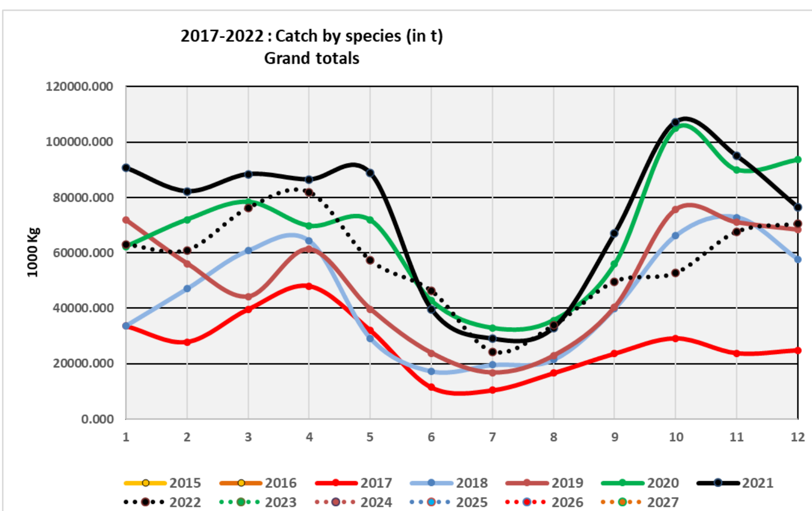



**Figure 9. Monthly fluctuations are U-shaped with low values in June-July and high values in January-May and October-December (2017-2022 data).**

| 2017-2022 Totals over all years : Ranking and cumulative percentages | | | |
|---|---|---|---|
| FG(Small Pelagic Net) | 1,197,845 | 31.4 % | **31.4 %** |
| FG(HL+TL) | 867,744 | 22.7 % | **54.1 %** |
| FG(NET) | 848,428 | 22.2 % | **76.3 %** |
| Beach Seine | 309,578 | 8.1 % | **84.4 %** |
| Launch(Net) | 237,487 | 6.2 % | **90.6 %** |
| FG(FT) | 155,258 | 4.1 % | **94.7 %** |
| FG(CN) | 121,983 | 3.2 % | **97.8 %** |
| FG(CuttleFish+Squid) | 49,146 | 1.3 % | **99.1 %** |
| Launch(HL+TL) | 13,253 | 0.3 % | **99.5 %** |
| Launch(FT) | 7,018 | 0.2 % | **99.7 %** |
| FG(Shrimps-Trowler) | 5,652 | 0.1 % | **99.8 %** |
| FG(Lobster) | 4,485 | 0.1 % | **99.9 %** |
| Launch(Beach Seine) | 2,641 | 0.1 % | **100.0 %** |

**Figure 10. Predominant fishing unit is the small pelagic net (31.4% of total production). Four boat-gear types account for about 85% of the production: Small pelagic net, handline and trawl, nets and, beach seine.**

| 2017-2022 Totals over all years : Ranking and cumulative percentages | | | |
|---|---|---|---|
| Al-Wusta Governorate | 1,962,950 | 51.4 % | **51.4 %** |
| Al Sharqiyah South Governorate | 706,830 | 18.5 % | **69.9 %** |
| Al Batinah North & South Governorate | 430,993 | 11.3 % | **81.2 %** |
| Dhofar Governorate | 410,329 | 10.7 % | **91.9 %** |
| Musandam Governorate | 207,552 | 5.4 % | **97.3 %** |
| Muscat Governorate | 101,863 | 2.7 % | **100.0 %** |

**Figure 11. Al Wusta is by far the most productive region. It alone accounts for more than 50% of the national production.**

## 10. Basic conclusions

<u>Production figures</u>

- 2014-2016: Rather significantly underestimated due to outdated fleet data and statistical gaps in the early stages of the present (upgraded in 2015) system. There is some good contrast between the original and the revised (upwards) figures.

- 2017-2023: Overestimated to some extent due to some assumptions favouring sharper extrapolation of fishing effort. The revised figures show lower production, as a consequence of a more conservative effort estimation approach.

- Peak production in 2020-2021. Result of effort intensification and higher catch rates. There are no grounds for suspecting the collected data and/or the Oracle estimator. A full and rigorous verification approach was implemented during August-November.

<u>Data Collection and estimations</u>

- Sample data on fishing effort can be significantly improved if the present data collection scheme, which is mostly based on the "boat approach", is replaced by the "monthly effort" approach. This scheme is already operational for certain fleet segments (dhows). Replacing the current boat activity approach offers several advantages regarding reliability of resulting effort estimates.

- With respect to landings and given the strategic importance of certain fisheries such as tunas and Indian oil sardinella, specific attention ought to be given to these fisheries; such as intensifying sampling and cross-checking operations.

- The Oracle estimation script should be made totally transparent and free of built-in conditions. If the need arises for exceptional data treatment, this should be achieved by means of externally supplied (hence transparent) parameters.

- Due to very high costs fleet censuses are conducted only every ten years with the result that boat/gear numbers that are used to extrapolate sample effort are often outdated and constitute a constant source of concern. A possible alternative is the implementation of annual "rapid" surveys, aiming simply at numbers of fishing units. Such data can be communicated by third parties via phone, SMS, email, etc.