| | |
|---|---|
| Project | AtlantOS – 633211 |
| Deliverable number | D8.9 |
| Deliverable title | Report on AtlantOS fitness for Atlantic Albacore: Assessment of the observing system fitness for tuna catch estimates. |
| Description | The model SEAPODYM simulates the spatial dynamics of tuna populations and fisheries under the influence of both fishing and environmental effects. The model is coupled off-line to temperature, horizontal currents, primary production and dissolved oxygen concentration. It is proposed to implement an operational forecast system for the Atlantic albacore tuna [D8.3] using available forcing variables from COPERNICUS CMEMS. This report evaluates the interest of the AtlantOS fitness for this operational production of tuna stock distributions. |
| Work Package number | WP8 |
| Work Package title | Societal benefits from observing/information systems |
| Lead beneficiary | Collecte Localisation Satellites |
| Lead authors | Patrick Lehodey |
| Contributors | |
| Submission date | 21/12/2018 |
| Due date | 31/12/2018 |
| Comments | |

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement nº 633211.

Stakeholder engagement relating to this task*

| | |
|---|---|
| WHO are your most important stakeholders? | ☐ Private company<br>If yes, is it an SME ☐ or a large company ☐?<br>☐ National governmental body<br>X International organization<br>☐ NGO<br>☐ others<br>Please give the name(s) of the stakeholder(s):<br>ICCAT |
| WHERE is/are the company(ies) or organization(s) from? | ☐ Your own country<br>X Another country in the EU<br>☐ Another country outside the EU<br>Please name the country(ies):<br>Spain |
| Is this deliverable a success story? If yes, why?<br>If not, why? | X Yes, because it is one of the first demonstration of operational fisheries oceanography linking COPERNICUS MEMS products with a spatially-explicit ecosystem and fish population dynamics.<br><br>☐ No, because ….. |
| Will this deliverable be used?<br>If yes, who will use it?<br>If not, why will it not be used? | X Yes, we will maintain and improve the model production to be used for management and fishing monitoring applications.<br><br>☐ No, because ….. |

NOTE: This information is being collected for the following purposes:

1. To make a list of all companies/organizations with which AtlantOS partners have had contact. This is important to demonstrate the extent of industry and public-sector collaboration in the obs community. Please note that we will only publish one aggregated list of companies and not mention specific partnerships.
2. To better report success stories from the AtlantOS community on how observing delivers concrete value to society.

*For ideas about relations with stakeholders you are invited to consult D10.5 Best Practices in Stakeholder Engagement, Data Dissemination and Exploitation.

## *Table of Content*

# 1. Summary

Within the AtlantOS WP8 “Societal benefits from observing/information systems”, the use case 8.7 will demonstrate the interest of improved oceanographic variables to develop the operational modelling of Atlantic albacore tuna dynamics to simulate in real time the distribution of the species under the influence of both fishing and environmental variability. This model called Spatial Ecosystem And Population Dynamics Model (SEAPODYM) estimates stock dynamics (standard stock assessment modelling). It also predicts the spatial distribution of fish density (by cohort) and can distinguish between fishing impact and natural variability (environment and climate). Simulation outputs for the historical period of fishing have been evaluated based on the current knowledge of this species, the statistical fit to all fishing data and the comparison with other population dynamic model estimates. The downscaling phase to the ¼° resolution operational forcing uses the free run version of the Copernicus reanalysis GLORYS2v4 due to detected biases in the equatorial circulation. The quality, limitation and further needs of data used for this application are discussed in this report.

# 2. Introduction

The AtlantOS WP8 “Societal benefits from observing/information systems” aims at providing new information products in several GEOSS societal benefit areas (i.e. climate, disasters, ecosystems, health and fresh water, increased safety for offshore activities and coastal communities). The present report describes AtlantOS fitness for the operational modelling of micronekton and tuna with the ecosystem and population dynamics model SEAPODYM. This task required i) to simulate micronekton functional groups to provide prey fields of tuna, ii) rebuild the history of the fish population and its fisheries over the recent decades and iii) to downscale the optimal parameterisation estimated through maximum likelihood approach with a coarse scale resolution hindcast simulation to the CMEMS products used for the operational modelling (AtlantOS deliverable D8.3).

Two different input datasets are needed to simulate and optimize the model over long historical series (1980-2010) and then to run higher resolution simulation over the recent period to prepare the operational production. The model is sensitive to its forcing variables. This sensitivity can be observed through the results achieved with the Maximum Likelihood Estimation approach to estimate the best parameter values allowing to fit historical catch data. This report evaluates the interest of the AtlantOS fitness for this operational production of tuna stock distributions.

# 3. Overall product quality score with respect to scope (fitness)

The input variables required to run SEAPODYM micronekton and tuna applications are temperature and horizontal currents in the surface - ~1000m depth vertical layer, the total primary production, the euphotic depth and the dissolved oxygen concentration. The Euphotic depth is used to define the boundaries of the vertical layers. Temperature and currents are averaged at the resolution of the model.

## 3.1. Hindcast 1° x month forcing (1972-2011)

Tuna, especially albacore tuna, are long-living species requiring running population dynamics models over long time series to limit the influence of initial conditions of the population structure on the final parameterization achieved using maximum likelihood approaches. Due to exploitation, tuna stocks are reduced by > 60-70 % of their theoretical unfished biomass. Therefore, including the historical fishing is another key condition to get a good estimate of model parameters. For such reasons the fish population model needs to cover the last few decades, ideally since 1950, i.e. the development of industrial fisheries. The computational needs to estimate the model parameters over this long period also limits the spatial and temporal resolution that can be used. Given that most fishing data are not available at better resolution than 1 degree square and one month, this resolution has been selected for the model parameter optimization phase.

Given the list of forcing variables and the availability of satellite ocean colour data only since 1998, we used a hindcast simulation (1972- 2011) from the NEMO ocean model (www.nemo-ocean.eu/), forced by the ERA40-INTERIM atmospheric reanalysis (atmospheric temperature, zonal and meridional wind speeds, radiative heat fluxes, relative humidity, and precipitation) corrected with satellite data. The Nucleus for European Modelling of the Ocean or NEMO model is also the core model of COPERNICUS Marine Environment Monitoring Service (CMEMS) operational products. NEMO was coupled to the biogeochemical model PISCES (Pelagic Interaction Scheme for Carbon and Ecosystem Studies; Aumont *et al.* 2015[1]). Since the beginning of the project, update of PISCES hindcast have been added to the COPERNICUS catalogue, but this latest release has not been tested yet.

All forcing variables were interpolated on a regular grid and similar time step prior to their use in the SEAPODYM simulations.

## 3.2. Pre-operational ¼° x week forcing (1998-2016)

The pre-operational system uses physical fields and satellite ocean colour data from Copernicus CMEMS, with the objective of providing the most realistic environmental conditions to simulate the fish population dynamics. There is a lack of historical synoptic datasets available for the ocean colour (SeaWiFS) prior to 1998. Ocean reanalyses with satellite derived primary production are therefore unavailable to simulate albacore tuna dynamics with SEAPODYM before 1998. Unlike the hindcast simulation driven only by atmospheric conditions, an ocean reanalysis assimilates oceanic variable observations derived from satellite (Sea Level Anomalies, Sea Ice Concentration and Sea Surface Temperature) or *in-situ* (temperature and salinity profiles) measurements, to provide a more realistic prediction.

In its final operational configuration, it is envisaged that the SEAPODYM albacore model will use both the reanalysis (GLORYS) to produce the initial conditions at the start of the operational chain of production and then its operational version, i.e. the Mercator-Ocean (PSY4) model 1/12° version available on CMEMS, degraded first to 1° and then if possible, to ¼° to improve the model results in areas of complex and shallow bathymetry. Primary production and associated euphotic depth used in this configuration are derived from ocean color data using the VGPM model of Behrenfeld and Falkowski (1997)[2], while a climatology from the World Ocean Atlas (Garcia *et al.*, 2010)[3] is used for the dissolved oxygen concentration.

---

[1] Aumont O., Ethé C., Tagliabue A., Bopp L., Gehlen M., (2015). PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geosci. Model Dev.*, 8, 2465–2513.

[2] Behrenfeld, M. J. and Falkowski, P. G. (1997). Photosynthetic rates derived from satellite based chlorophyll concentration. Limnology and oceanography, 42(1):1–20.

[3] Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Johnson, D.R., 2010. World Ocean Atlas 2009. In: Levitus,S. (Ed.), Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation, vol.3. NOAA Atlas NESDIS70, U.S. Government Printing Office, Washington, D.C.(344pp.).

The fitness for use of the hindcast and operational products are summarized in tables 1 and 2 and discussed in the following sections.

**Table 1: Fitness for use of forcing variables from the NEMO-PISCES hindcast simulation for historical reconstruction.**

| | Variable | | | | |
|---|---|---|---|---|---|
| | Temperature | Horizontal velocity | Total Primary Production | Euphotic depth | Dissolved oxygen |
| **Origine** | NEMO | NEMO | PISCES | PISCES | PISCES |
| Timeliness: Is the data available for use in the time frame in which it is expected? | Yes | Yes | Limited | Limited | Limited |
| Time coverage: Are data values missing or unusable for the expected time series? | No | No | No | No | No |
| Spatial Coverage: Is the data available for use in which it is expected in the spatial domain? | Yes | Yes | Yes | Yes | Yes |
| Conformity: Does the data meet expected format? | Yes | Yes | Yes | No | Yes |
| Accuracy: Do data accurately represent the real-world as expected? | Yes | Yes | Reasonably | Reasonably | Reasonably |

**Table 2: Fitness for use of forcing variables from the GLORYS reanalysis and satellite product for pre-operational model.**

| | Variable | | | | |
|---|---|---|---|---|---|
| | Temperature | Horizontal velocity | Total Primary Production | Euphotic depth | Dissolved oxygen |
| **Origine** | GLORYS | GLORYS | Satellite-derived | Satellite-derived | - |
| Timeliness: Is the data available for use in the time frame in which it is expected? | Yes | Yes | Yes | Yes | No |
| Time coverage: Are data values missing or unusable for the expected time series? | No | No | Yes Generated from existing product | Yes Generated from existing product | Yes Replaced by WOA |
| Spatial Coverage: Is the data available for use in which it is expected in the spatial domain? | Yes | Yes | No | No | Yes |
| Conformity: Does the data meet expected format? | Yes | Yes | Yes | Yes | Yes |
| Accuracy: Do data accurately represent the real-world as expected? | Reasonably Needs correction | Reasonably Needs correction | Reasonably at basin scale (problem at the coast) | Reasonably | average only (climatology) |

The fitness for purpose of the albacore Atlantic model cannot be achieved in its operational mode with the current version of higher resolution satellite driven forcing variables (Table 3). Despite the downscaling technique that has been used to rescale the optimized parameters to the different environmental forcing, there are still too strong biases especially in the satellite derived primary fields to achieve a correct parameterisation. This is illustrated with the Figure 1: Comparison of primary production from biogeochemical model (PISCES-BIORYS simulation) and satellite derived (VGPM) product used in this study, at the peak of production in the first week of May 1998.Figure 1 showing primary production fields at the seasonal peak of production in the North East Atlantic. This bias is due to the use of the “oceanic case” for the empirical model estimating primary production from ocean colour at global scale. To solve this issue, and thus still rely on satellite data to get the most realistic forcing, it is needed to adapt the approach. A practical approach could try merging the oceanic case algorithm with specific algorithms developed in different coastal regions. Another alternative would be to use outputs from biogeochemical models that assimilate the satellite ocean colour data and thus become more accurate in their predictions of the real world. This type of products is expected to become available in the coming year on the CMEMS catalogue and will have to be tested.

**Table 3: Fitness for purpose of the pre-operational model for Atlantic tuna.**

| **Based on forcing:** | NEMO-PISCES coarse resolution hindcast | GLORYS- Satellite derived Primary production |
|---|---|---|
| Timeliness: Is the data available for use in the time frame in which it is expected? | Limited | Yes |
| Time coverage: Are data values missing or unusable for the expected time series? | No | No |
| Spatial Coverage: Is the data available for use in which it is expected in the spatial domain? | Yes | Yes |
| Conformity: Does the data meet expected format? | Yes | Yes |
| Accuracy: Do model products accurately represent the real-world as expected? | Reasonably well for the open ocean. Not so well in the NE Atlantic (e.g., Bay of Biscay) | No |

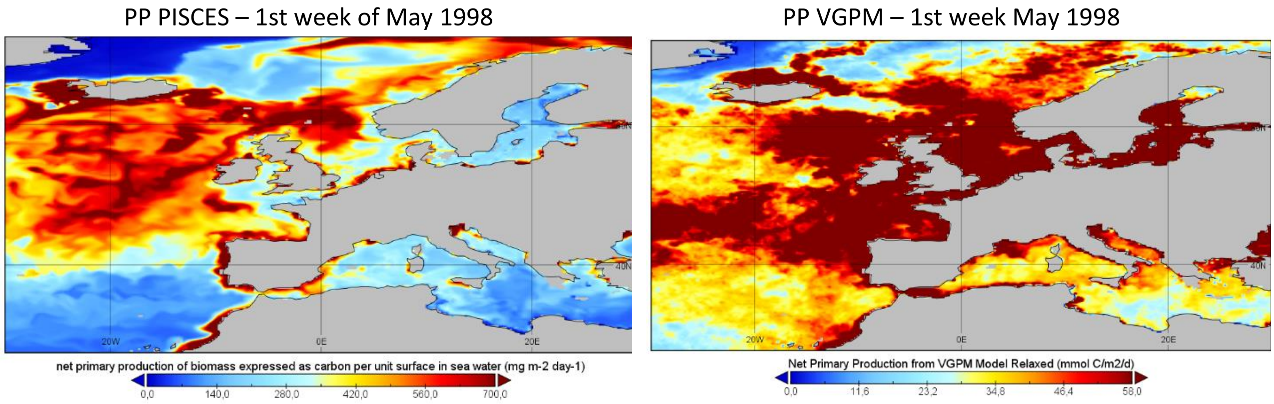


**Figure 1: Comparison of primary production from biogeochemical model (PISCES-BIORYS simulation) and satellite derived (VGPM) product used in this study, at the peak of production in the first week of May 1998.**

# 4. Most important characteristics for the Targeted Product quality

For the coarse resolution configuration needed for the reconstruction of the fish population history, the length of the time series is a first criterion that is sought to cover multiple fish generations. This is critical to estimate the parameters of the stock recruitment relationship, while it allows reducing the influence of initial conditions of the population structure on the parameter estimation. The full domain of the species habitat (60°N-60°S) needs to be covered. Then ocean circulation, temperature and primary production are expected to provide reasonable large-scale estimate of mean state and seasonal and interannual variability. This is essential to expect a good fit to fishing observations (spatially-explicit catch data) used for parameter optimisation with the Maximum likelihood Estimation approach.

Once the optimal solution is achieved, the initial conditions of the species population (density by cohorts and space at a given date) are used to seed the operational global model. The change of forcing models, and resolution, requires applying a downscaling approach for rescaling optimal parameters to the new environmental model configuration. This step is conducted in the MLE framework of the model with the objective to estimate habitat distributions like those achieved with the first configuration. However, if the two forcings show important discrepancies, the downscaling approach may not converge to a satisfactory solution. Therefore, for the forcing fields used in the pre-operational model, it is expected to observe a good mean state close to the mean state of the hindcast simulation, and better accuracy in the spatial and temporal variability, especially when moving to higher resolution including mesoscale activity. The temperature and primary production fields need to be coherent and consistent with the mesoscale patterns (i.e., fronts and eddies).

# 5. Quality elements of the most important characteristics that affect the Targeted Product quality

The length of the time series is an issue for the hindcast simulation that is limited by the availability of the atmospheric reanalysis ERA-Interim. It is a global atmospheric reanalysis starting from 1979 produced by ECMWF, continuously updated, once per month with a delay of two months to allow for quality assurance. This is enough for the optimisation tasks. However, reanalysis back to the sixties or fifties would be useful to run ocean hindcast simulations, that even with a lower accuracy would help to decrease the influence of initial conditions. It is noted with interest that ECMWF should release the ERA5 reanalysis that could eventually replace the ERA-Interim reanalysis and would cover the historical period back to 1950.

The spatial coverage becomes an issue only for satellite derived products, i.e. the primary production and euphotic depth used for the operational model. There is no satellite chlorophyll data in latitudes higher than ~45° in winter season (Figure 2). This is due to sunglint, cloud coverage and low light levels.  One option to solve this problem is to replace missing data with those of the operational biogeochemical model, after verification that there is small differences in the mean state and variability of both datasets. This is the case for high latitudes in winter where the primary production is very low. A future alternative could be to use directly the primary production from operational biogeochemical models that assimilate ocean colour data to provide the most possible realistic maps of primary production.

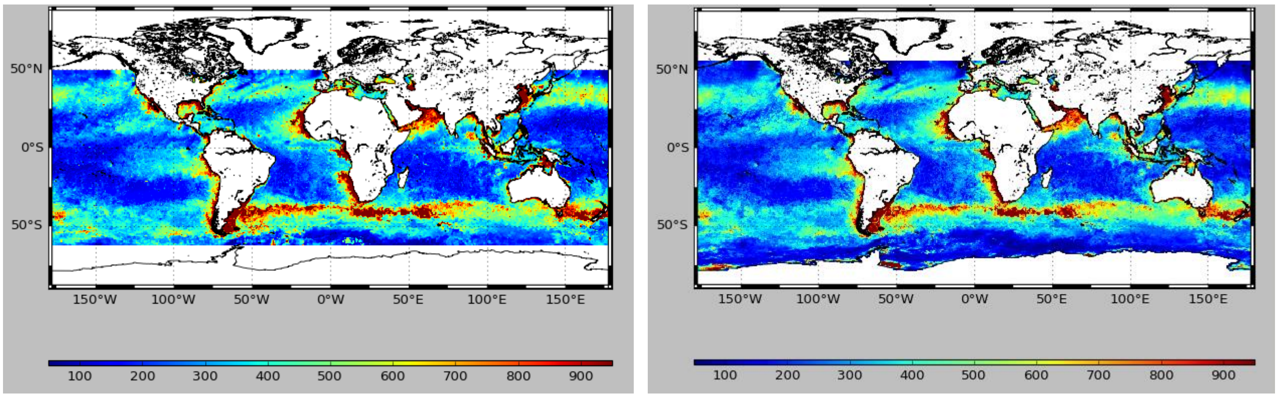


**Figure 2: Change in spatial coverage between old and new CMEMS product for the global satellite chlorophyll a (GLOBCOLOR). New product covers the high latitudes but only during spring - summer seasons.**

Water temperature in the model is critical to simulate the dynamics of micronekton with a relationship between time of development and ambient temperature. It is also essential for the fish model, both in the definition of the optimal spawning temperature and thermal feeding habitat that largely control the distribution and dynamics. The quality of predicted water temperature is good enough for both the hindcast or the operational products in regard of the model sensitivity to this parameter.

The model is much more sensitive to the horizontal currents. By integrating high frequency variability when building micronekton biomass over time and space, the model provides a useful feedback to identify potential inconsistencies in this variable. A simple comparison of model outputs allows to quickly detect divergences between simulations and to identify the region and depth where they occur, thus providing helpful direction in the analysis of physical fields. For instance, the micronekton produced from the GLORYS2v4 (G2v4) reanalysis and its free run equivalent have been compared (Figure 3). The main differences appear in the equatorial region, particularly in the Pacific. Model outputs were compared to climatological observations in this region (Figure 4). G2V4_free-run simulates reasonably well the South Equatorial Current (SEC: westward) which starts decreasing from 170°W. But the SEC is predicted too intense in GLORYS2v4 and HYCOM. GLORYS2v4-free seems to provide the best compromise for the equatorial circulation. However, in the higher latitudes, the products using data assimilation perform better. This is particularly clear with the position of the Kuroshio extension that is not well represented in GLORYS2V4-free.

The coarse resolution simulates the general circulation correctly in average, including the interannual variability (e.g., ENSO). The resolution is not sufficient however to correctly simulate intense and narrow jets, e.g. the Kuroshio and Gulf Stream.

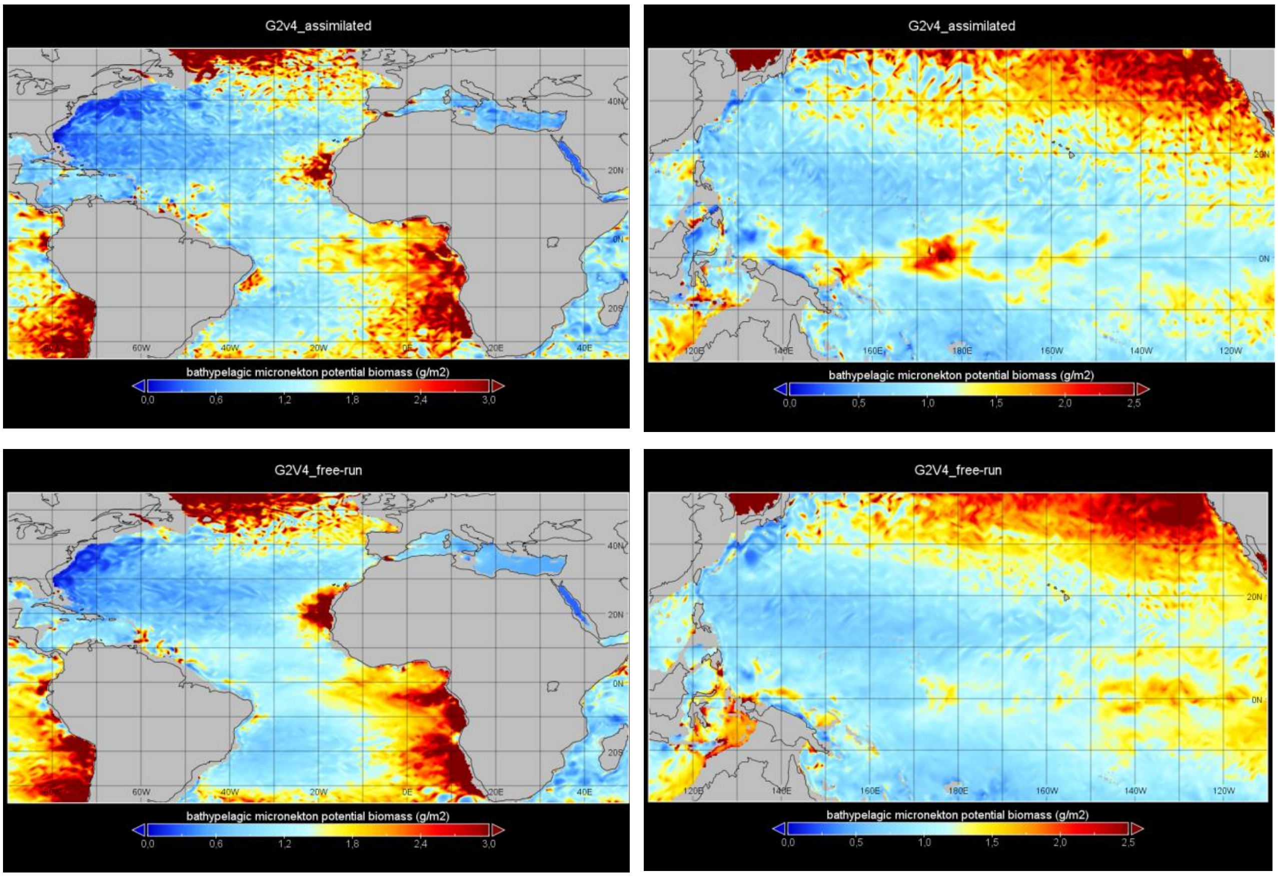


**Figure 3: Lower mesopelagic micronekton biomass distribution (1st week of June 2015) simulated with G2V4 (top) and G2V4-free_run (bottom) in the Atlantic and Pacific Ocean.**

The primary production derived from the satellite ocean colour is useful to start from realistic distribution. It should be noted that this product is not directly available from the CMEMS catalogue but has been prepared from the surface chlorophyll concentration of the COPERNICUS GLOBCOLOUR (L4) product using the empirical model VGPM (Behrenfeld and Falkowski 1997) and the Photosynthetically Active Radiation (PAR) from ERA-INTERIM (ECMWF) reanalysis. The known issues for this kind of product are mainly i) the difference in the signal observed in oceanic and coastal waters due to possible reflectance of sea bottom in shallow waters and the presence of sediments and other suspended matters and ii) the uncertainty in the empirical approach to model the total primary production based on surface data. These sources of uncertainty impact the absolute biomass of micronekton. However, the spatial dynamics is strongly correlated to the mesoscale physical processes and remain unchanged between the different products, while it is coherent with outputs from ocean circulation model assimilating SST and altimetry data.

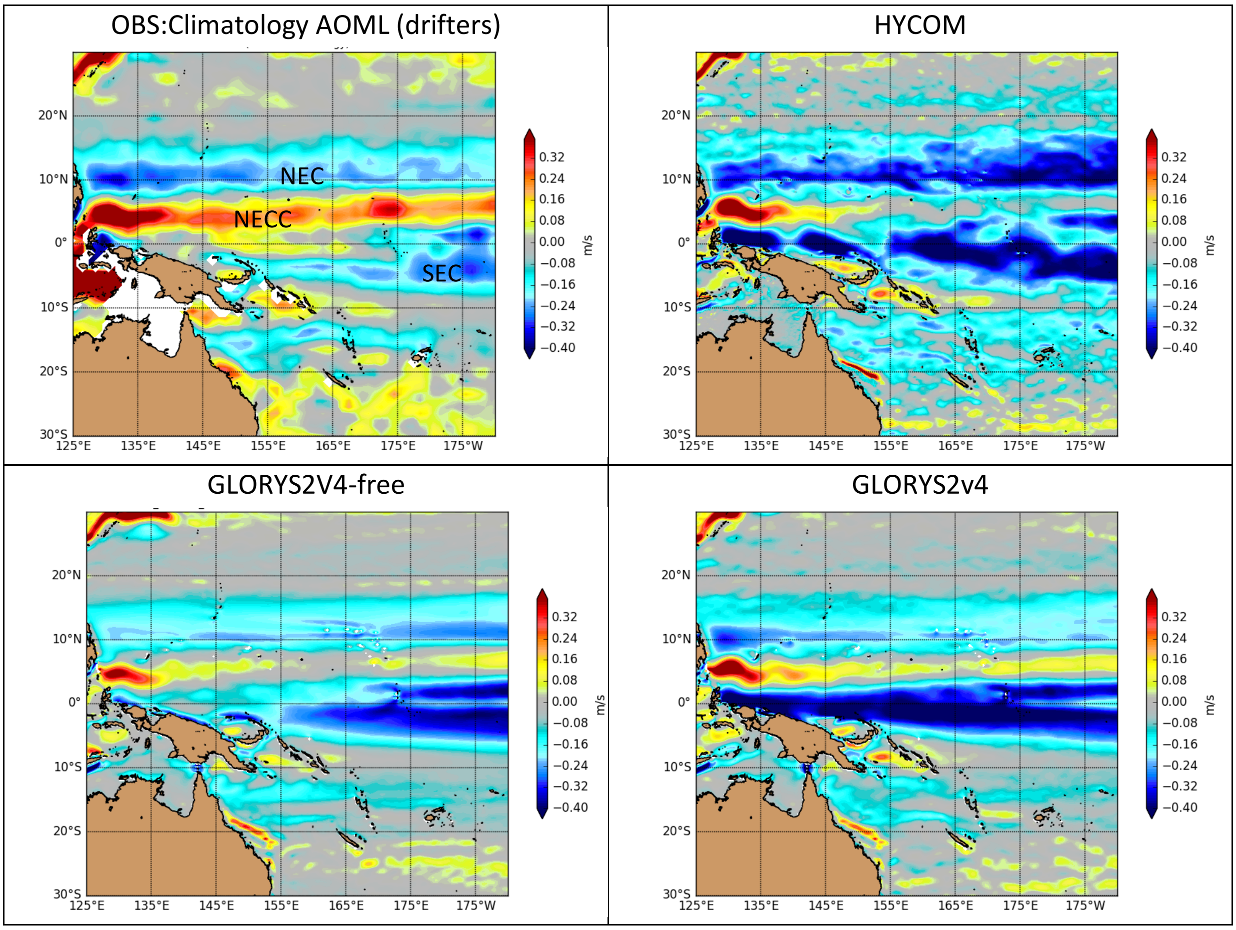


**Figure 4: Zonal surface currents from observations and different reanalyses/operational models**

## 6. Limitations of the quality of the targeted products due to the input data set used

**A first limitation of the quality of the targeted product (i.e., the biomass distribution of Atlantic albacore tuna and then the predicted catch) can be linked to the hindcast simulation at coarse resolution that is used to estimate the fish population parameters. For instance, it has been noted in a different application in the**

**Pacific Ocean that the biogeochemical model underestimated the bloom of phytoplankton in the eastern Pacific during the onset of La Niña following an El Niño event, with the consequence of underestimating the larval survival and subsequent recruitment in the fish stock. In the Atlantic Ocean, the main issue concerns the Gulf of Biscay that is a key area for the fisheries during the species feeding migration in summer, and that is not well represented at the resolution of 1°. The coarse resolution in this region may lead to errors in the model optimal estimation. Despite the proposed downscaling method to re-estimate fish population dynamics parameters when moving from the hindcast configuration to the pre-operational (reanalysis) configuration, the result achieved at this higher resolution is not satisfactory. There are too strong regional differences between the primary production from the hindcast simulation using the biogeochemical model and the one generated from the satellite ocean colour, especially in coastal and upwelling regions, e.g., the Bay of Biscay, the North Sea and the Canary current upwelling system (**

**Figure 5). Primary production has a strong impact in predicted abundance and distribution of fish due both to the larvae survival mechanisms and adult distributions driven by the prey (micronekton groups) distributions, themselves strongly dependent on primary production. In absence of correction of these biases, it is preferable to use the coarse resolution configuration that is well fitted to the historical fishing data (**
Figure 5). It can be used to investigate basin scale management questions. However, this model configuration cannot address real time monitoring of fisheries nor regional management issues, e.g., testing management scenarios for the large European albacore fishery in the Bay of Biscay.

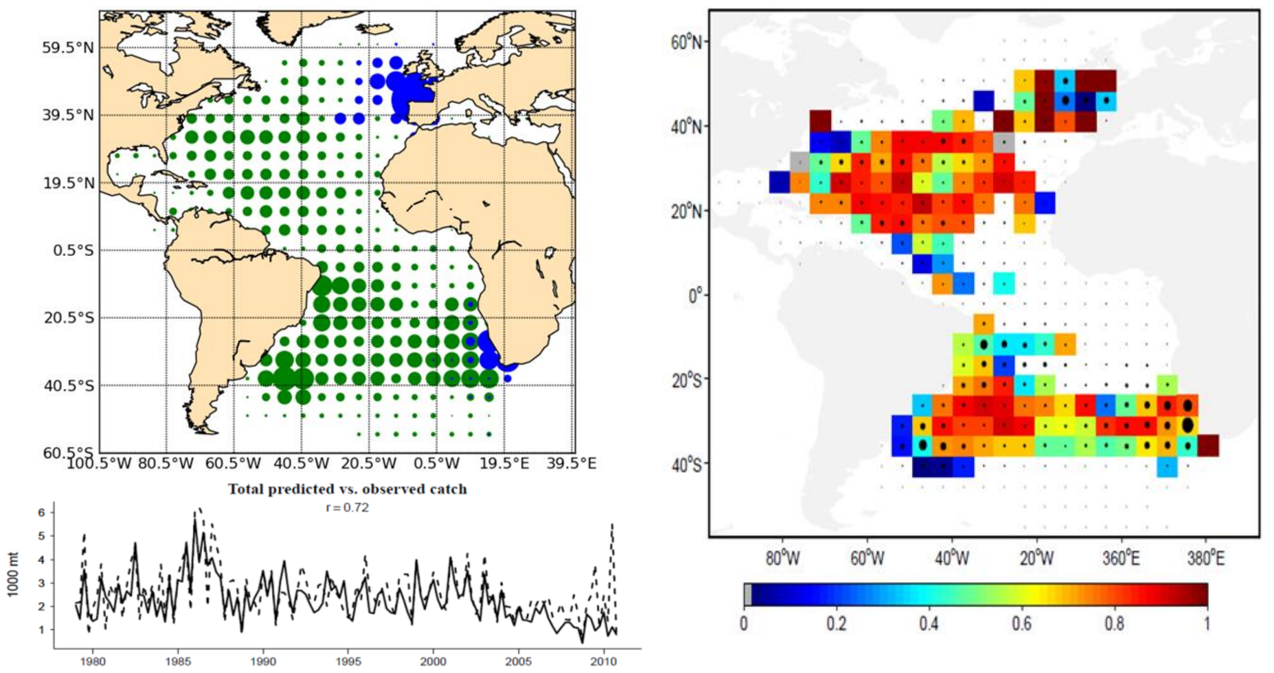


**Figure 5: Left: Geographic distribution (5°x5°) of albacore mean annual catch by major gears for 1987-2014 (longline in green and surface gears in blue). Right: Fit ($R^2$ goodness of fit) between observed and predicted historical catch data 1979-2010 with the hindcast simulation at coarse resolution. Bottom: time series of predicted vs observed total catch of Atlantic albacore.**

## 7. Characteristics that most fail to meet the scope of the Targeted Product

Due to the essential phase of parameter estimation with the maximum likelihood approach, it is critical to use the most realistic environmental dataset (temperature, horizontal currents and primary production) to predict the spatial distributions of fish density and then catch based on observed fishing effort. Thus, theoretically, the best approach would be to use ocean reanalyses and primary production derived from satellite ocean colour with mesoscale structure coherent with the physics thanks to assimilation in the circulation model. The high computational requirement for the model parameter estimation would be solved by degrading the resolution of the environmental variables, while the parameter optimisation would remain much coherent with the higher resolution when moving back to the original forcing for the operational model. However, the optimization procedure requires reanalyses going back to at least the 1970s. They are not available, while global ocean colour data are only available since 1998 (SeaWiFS).

Therefore, the alternative is to use coupled physical-biogeochemical model outputs. A physical reanalysis can be coupled (offline) with a biogeochemical model allowing going back to 1970 (ERA interim), but the assimilation of altimetry data seem to introduce a bias in the equatorial circulation. The best compromise was to use a hindcast simulation to achieve the model optimisation over the longest possible historical period, and then to use the GLORYS2v4-free_run reanalysis for the pre-operational development to avoid problems due to the bias in the equatorial circulation. However, this is not fully satisfactory as the lack of assimilation introduce errors in mesoscale patterns at mid to high latitudes and local inconsistencies with satellite derived primary production.

# 8. Expert Judgement of the most important Gaps in the input datasets

Despite impressive progress in the last decade, the development of operational fisheries oceanography still misses long historical reanalyses (i.e., since ~ 1950) of the ocean state with corrected biases. Given that ocean colour data are not available before 1998, the primary production can be only provided from biogeochemical model for long historical analyses. The quality of biogeochemical model are becoming sufficient at basin scale to be used in ecosystem and fish models, but errors and limitations exist that need to be well known from the users. It can be expected that the quality of these model outputs will benefit of the rapid development of ocean colour data assimilation. Alternatively, a downscaling approach can be proposed to export results (i.e., initial conditions) and parameterization achieved with the long term hindcast to a realistic high resolution and real time forcing. However, if satellite derived primary production should be used, it requires correction in regions of shallow waters.

For practical applications in operational fisheries oceanography, it would be useful to propose a long ocean reanalysis corrected when possible from its most obvious biases (e.g., equatorial circulation), with the associated biogeochemical variables generated by a simulation driven by this corrected reanalysis and assimilating ocean colour data. The same corrected physical-biogeochemical variables could be also proposed at different resolutions from the highest (e.g., 1/12°x d) to one or two selected lower resolutions (1° x month) since the users (marine biologists and fishery scientists) do not have the expertise to do this properly.

--------