DATA ADEQUACY REPORT
Deliverable 14.2.1

Lead contractor for this deliverable: Ifremer

Work Package 14 : Second Data Adequacy Report – Vol. 1	
Author(s)	**Organisation**
Jacques Populus	Ifremer
Eric Moussat	Ifremer
Mickaël Vasquez	Ifremer
Erwann Quimbert	Ifremer
Frédérique Blanc	CLS/ARGOS
Laurent Soudarin	CLS/ARGOS

A project funded by:

EUROPEAN COMMISSION, DIRECTORATE-GENERAL FOR MARITIME AFFAIRS AND FISHERIES,
MARITIME POLICY ATLANTIC, OUTERMOST REGIONS AND ARCTIC

Work Package number	**14**
Work Package title	**Data Adequacy Report n°2**
Deliverable number	**D14.2.1**
Title	
Short description	This data adequacy report gives an account of the marine data landscape by analysing the contribution (or the lack) of data sets to the creation of products in the frame of 11 thematic challenges. The methodology, based on metadata catalogues and specifically quality measures derived from ISO, builds a quantitative comparison between the specifications of the products, their actual results and how suitably the existing data sets contributed to them. The results assess data from three standpoints: marine thematic domains (challenges), marine characteristics (parameters) and key EU providers, and make recommendation to improve the situation in three areas: by additional surveys, by strengthening data assembly, by improving data services.
Keywords	Checkpoint, data fitness for use, availability, ISO, metadata, catalogue, data sets
Editor / Organisation	Jacques Populus / Ifremer
Deliverable due date	M33
Comments	

History				
Version	**Author(s) / organisation**	**Status**	**Date**	**Comments**
1.0	**Jacques Populus, Eric Moussat and Mickaël Vasquez - Ifremer**	**Submitted**	**26/05/18**	
1.1	**Jacques Populus, Eric Moussat and Mickaël Vasquez - Ifremer**	**Final**	**17/07/18**	

Dissemination level		
PU	Public	
CO	Confidential for project partners and EU Commission only	**x**

Disclaimer

"The information and views set out in this [report/study/article/publication…] are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein."

Table of contents

Glossary

AORA	Atlantic Ocean Research Alliance
BODC	British Oceanographic Data Centre
CFP	Common Fisheries Policy
Characteristic	Distinguishing feature [ISO 9000:2005]
Copernicus	European Program for establishing European capacity for Earth Observation
CDI	Common Data Index
CMEMS	Copernicus Marine Environment Monitoring Service
CSW	Catalogue Service for Web
DAR	Data Adequacy Report
DCF	Data Collection Framework
DPS	Data Product Specification (ISO 19131)
DOI	Digital Object Identifier
EBSA	Ecologically or Biologically Significant Marine Areas
EBV	Essentiel Biodiversity Variable
EDMED	Directory of Marine Environmental Data
EDMERP	European Directory of Marine Environmental Research Projects
EDMO	European Directory of Marine Organizations
EEA	European Environmental Agency
EIONet	European Environment Information and Observation Network
EMSA	European Maritime Safety Agency
ESDI	European Spatial Data Infrastructure
ESI	Environmental Sensitivity Index
EOV	Essential Ocean Variables
EEZ	Economic Exclusive Zone
FAO	Food and Agriculture Organization
GCMD	Global Change Master Directory
GEBCO	General Bathymetric Chart of the Oceans
GES	Good Environmental Status
GMES	Global Monitoring for Environment and Security
GIS	Geographical Information System
GRDC	Global Runoff Data Base
HFR	High Frequency Radar
IBA	Important Bird Area
ICES	International council for the exploration of the sea
INSPIRE	Infrastructure for Spatial Information in the European Community
ISO	International Organization for Standardization
ISO IEC	ISO International Electrotechnical Commission
ISO NP	ISO New Proposal
ISO NP TS	ISO NP Technical Specification
JRC	Joint Research Centre
MBO	Management By Objectives
MESA	Monitoring for Environment and Security in Africa
MS	Member States
MSFD	Marine Strategy Framework Directive
MSPD	Marine Spatial Planning Directive
MSP	Maritime Spatial Planning

NIS	Non-indigenous species
NODC	National Ocenaographic Data Centre
OSPAR	The Oslo Paris Convention
P01	BODC Parameter Usage Vocabulary
P02	SeaDataNet Parameter Discovery Vocabulary
P03	SeaDataNet Agreed Parameter Groups
PSMSL	Permanent Service for Mean Sea Level
RSC	Regional sea convention
SDM	Species Distribution Modelling
SEO	Search Engine Optimization
SDN	SeaDataNet
STECF	Scientific, Technical and Economic Committee for Fisheries
TAC	Thematic Assembly Centre
TDP	Targeted Data Product
TRL	Technical Readiness Level
UD	Upstream Data
VME	Vulnerable Marine Ecosystems
VMS	Vessel Monitoring System

Executive summary

The work presented in this report follows on from the first Data Adequacy Report which presented data availability issues from a catalogue of data sets identified by the challenges before making their products. This second DAR gives an account of data appropriateness, the other aspect of adequacy.

At the time the 53 products across the 11 challenges were specified, a catalogue of products was implemented in Sextant and populated with the products specifications expressed by way of 8 quantitative quality measures derived from ISO 19157. Similarly upon completion, the achieved products were assessed using the same measures, which permitted the computation of the discrepancies between the two sets of quality measures. All these measures were illustrated by bar charts for each quality measure showing at a glance where the gaps are.

The data sets contributing to the challenges were also assessed along the same indicators and reported in spreadsheet form under the characteristic they refer to.

Data analysis was conducted from three angles:

- The challenges, specifically which ones had least performed and why;
- The characteristics (or variables), by assessing for each inadequate product what were their shortcomings using quality measures;
- The main EU data providers (EMODnet, CMEMS and the DCF).

In the synthesis a few key characteristics were selected according to their higher relevance to the challenges and recommendations were formulated in three areas of potential action:

- Data assembly when data exist but need to be pulled together;
- Data availability when data exist but for some reasons are too difficult to use;
- Gap bridging by surveys when either coverage or resolution is lacking.

From the Challenges web page a table gives access to the specifications of the products, their scores, their metadata including appropriateness and use limitations, download facilities via a DOI and a view of the geospatial layers in the Web GIS.

It is reminded that the Checkpoint was intended not only to our commissioners but also to the providers community. This was the reason for us to develop objective and quantitative tools enabling the providers to search their characteristics of interest by application and discover the specifications and related data shortcomings as a basis for improvement of their data quality management.

1 Introduction

When dealing with data, views on data quality from the users and from the providers may differ to a large extent. In spite of referring to users committees to orientate their policy, it may happen that organizations of providers restrain their views to only certain domains or fail to consider upcoming issues.

Based on this assumption, after a time when observations of the sea have been made for specific purposes, e.g. for specific national purposes or to demonstrate a technological capability, the European Commission has now moved to a new paradigm where the leading edge is the users' view and where data are collected once and used for as many purposes as possible. This means relying preferably on users rather than on producers to assess existing data sets and data sources and promote recommendations for a better fulfilment of their needs.

There is no comprehensive overview of gaps and duplications at EU basins level, let alone at whole EU marine domain level. Consequently we lack an overall view of what the priorities are for further data assembly or collection. The Checkpoints were therefore implemented as data stress tests using challenges as benchmarks to assess the marine data landscape, in the present case for the Atlantic Ocean. This sea basin has peculiar aspects:

- On one hand it's quite big, from East to West and from the Equator to the Polar Circle, with a great variety of shores and adjacent countries, which makes it very challenging to grasp so much diversity in data landscape;
- On the other hand it is bordered by Western Europe and North America, places where marine data are rather well catered for, especially with e.g. a strong regional convention (OSPAR) and a very active scientific organisation (ICES) that both have a remit in managing marine data.

The Atlantic Checkpoint has been implemented as a stress test to assess data adequacy with a view to obtain:

- A set of products, generally as digital map outputs but also as times series, possibly along with their confidence;
- A list of the data sources used and their data providers. This should include primary data producers as well as intermediaries (e.g. Copernicus, EMODnet) and the analysis should say whether the latter provide useful layers or whether it is necessary to go back to the original data;
- An analysis of the usefulness of each data source in terms of identification (attributes, quality features), delivery and usability and in particular whether some data sets were too complicated to access or use;
- An identification of how well the present data collection, monitoring and survey programs meet users needs and simultaneously an identification of gaps in data sets;
- Recommendations for data collection or assembly to address these gaps, including options that could be implemented to increase the links between different monitoring, observation and data collection programs;
- Finally, an opinion from the Project as to whether the availability of data are improving or worsening.

The checkpoints are requested to contribute in the first place to the identification of priorities in terms of making existing data more available and usable. Unveiling existing data is the most costly task as it only needs to resolve political and technical issues. Only when this has been thrashed out can recommendations for data collection be considered, a much more demanding and costly issue.

2 Structure of the document

The results in the form of tables are shown in the core text of this report (D14.2.1) as the list of products scores (Table 3) and the main findings of the data analysis (Tables 5 to 7).

All other tables can be found in two separate Annexes.

D14.2.2 contains Annexes 1 to 4 as follows:

- Annex 1 provides the comprehensive table of products components. As a support to Table 3 in the core text, while using the same colour code, it gives a more detailed view of the proportion of components which could not be made or were produced in a limited or inadequate way, along with their detailed lineage (P02 and P01);
- The bar charts in Annex 2 give a quantitative view of the quality measurements for each characteristic (parameter). An example of a set of bar charts is given in the introduction of section 3 (Data analysis), so the reader is guided to properly interpret them. For all the categories of characteristics (P02), placed in alphabetical order, the reader is invited to consult this Annex 2;
- Annex 3 contains the list of the 53 products listed per challenge;
- Annex 4 contains the list of the used datasets classified in providers' names alphabetical order.

D14.2.3 provides additional information in the form of bar charts for those wishing to know more about related datasets, contains all a series of spreadsheets exported from the catalogues in 3 separate annexes. Annex 5 describes P02s (characteristics) for which the components were not covered for absence of data or for data not available. Annex 6 contains all P02s not meeting the products requirements whose appropriateness was assessed. For each characteristic these tables (spreadsheets) enumerate all the concerned components and give a list of the contributing data sets. These three documents enables the reader to see the P02 analysis and follow on to the corresponding bar charts and further to the related spreadsheets if desired.

Note: For the sake of simplicity, in the whole document challenges with be called by their name in capitals: For example "The Bathymetry challenge" will just read "Bathymetry".

3 Methodological framework

It is considered important to refer here to the methodological basis and vocabulary of the Atlantic Checkpoint used in the Literature Survey report and DAR 1 to ensure common understanding throughout the document.

The main principles of the methodology were to implement:

- An objective, quantitative and reproducible assessment of data adequacy in solving challenge issues;
- The use of established vocabularies, in keeping with current initiatives in marine data management, ensuring common understanding and enabling straightforward replication of methods;
- An illustration of the marine data landscape by key quality indicators and bar charts helping set priorities between variables but also between basins (primarily those using the same method);
- Tools enabling challenge experts and users alike to assess the marine data landscape and provide their own feedback (catalogues, browser etc.);
- Services to provide specifications for corrective actions to whom it may concern (data providers, decision makers etc.

3.1 Terminology

The definitions of the vocabulary below have a key role in understanding the Checkpoint assessments.

- **Characteristic:** an attribute of a distinguishing feature that refers either to a variable derived from the observation, the measurement or the numerical modeling of a phenomenon or of an object in the environment, or to the geographical representation of an object on a map by a set of vectors (polygon, curve, point), e.g. “coastline”. The SeaDataNet classification offers three different levels of granularity to group characteristics from the finer to the coarser: the SDN BODC Parameter Usage Vocabulary list P01 for characteristics (when existing), the SDN Parameter Discovery P02 list for categories of characteristics and the SDN Agreed Parameter Groups P03 list for group of categories. The vocabularies and definitions are available on-line at: http://www.seadatanet.org/Standards-Software/Common-Vocabularies;
- **Component:** see “quality unit”
- **Environmental matrices:** The environments where characteristics are measured **or** computed: Air, Ice, Fresh water, Marine water, Biota/Biology, Riverbed/Seabed and Human activities;
- **Data:** reinterpretable representation of information in a formalised manner suitable for communication, interpretation or processing (ISO 19115);
- **Dataset:** an identifiable collection of data (ISO 19115). It can be a time series, a lithological description of a marine sample, a gridded dataset such as a DTM, an hydrodynamic model output, a GIS dataset or a feature layer of a GIS dataset, a data base or a table of values in a publication. A data set can be made of several files (e.g. the set of seismic data files recorded along the same line);

- **Collection of datasets:** A collection of datasets is a set of datasets. A collection of datasets sharing the same specifications of production is a data set series;
- **Upstream Data (UD):** a set of data used which serves as input to a challenge data product. An upstream dataset corresponds to a unique characteristic and is identified by the category of the characteristic, the characteristic, the name of the data provider of the dataset and the name of the dataset;
- **Data Product Specification (DPS):** a precise technical description to build the desired product in terms of the requirements that it will or may fulfill (ISO 19131). The DPS contains both the specifications of the product and of its quality evaluation;
- **Data Product (TDP):** a dataset created according to a data product specification
- **Data adequacy:** can be defined as the fitness-for-use of the data for a particular user or for a variety of users**.** Since different applications require different properties associated with the data themselves, adequacy should be defined objectively using standardized quality nomenclature and methods;
- **Quality :** totality of characteristics of a product that bear on its ability to satisfy stated and implied needs (ISO19115 -1);
- **Quality unit** (alias **"component"**)**:** assessment unit defined by a combination of a subset of a dataset and the selected quality measures applied to it;
- **Indicator:** information that is measurable, accurate, reliable and usable to implement corrective actions when performance is not in conformity with the objectives (ISO 9004).

3.2 A normative framework

EMODnet is a network of organisations collecting and providing data through thematic portals to support the EU marine strategy under the aegis of the DG MARE. To assess its capacity to achieve its objectives, an ISO 9004-like quality management process to improve the network (Figure 1) has been set up by DG MARE and expressed by the EMODnet Checkpoint concept in their calls for tenders.

Checkpoints can be regarded as overarching observing systems. What they observe is the whole realm of marine data distributed among a great number of organizations, people and places and in a variety of systems addressing many different purposes. The measurement and analysis process is a series of stress-tests called "challenges" carried out by users for which a series of products (maps, time series, tables) making use of existing data must be delivered according to specific objectives derived from the call.

The results must provide the information necessary for the evaluation of the performance of the existing data collections to meet the user needs. They are assembled in Data Adequacy Reports for effective decision making to improve the overall data management system and are designed to monitor its progress.

By "Data adequacy", the call for tender includes quality aspects not only related to data *sensu stricto* "How much existing data meet users' needs for the challenge products?", but also to their conditions of availability: "How are data made available to the Challenges"? These two strands of adequacy were initially called "Appropriateness" and "Availability" by the MedSea Checkpoint and taken forward by the Atlantic Checkpoint. They are covered by the quality concept of ISO 9000 as the "Degree to which a set of inherent characteristics fulfils requirements".

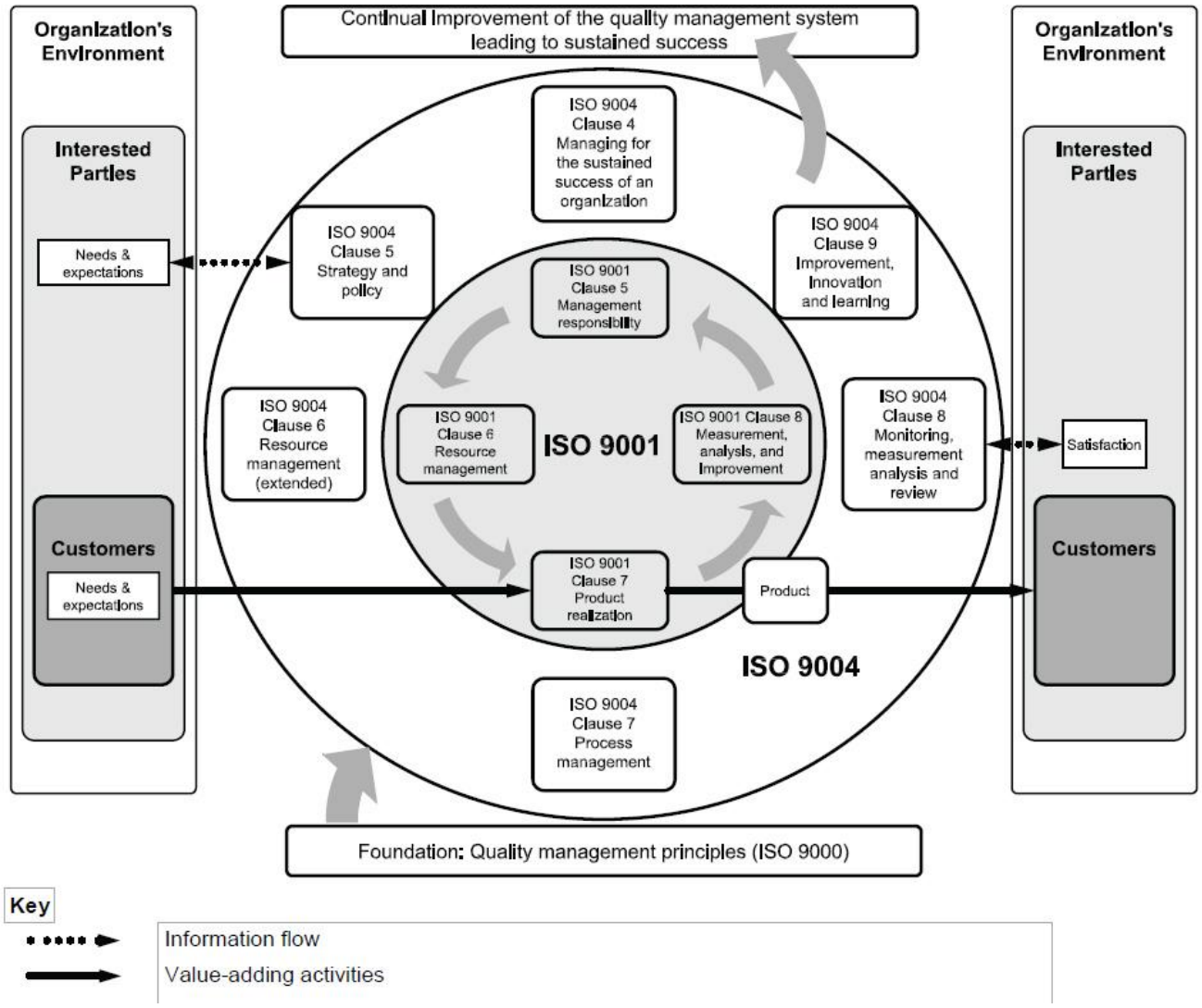


Figure 1: ISO 9004 extended model of a process-based quality management

The selection of the monitoring methodology and key performance indicators to determine the data adequacy is critical for the success of the measurement and analysis process. It should be appropriate to the nature of the activities and the context of the organization's environment. The ISO standards for geographical information are designed for this purpose and are applicable to the environmental data needed by the challenge applications (see the Literature Survey). ISO19115 is the well-known metadata norm which is being used by many NODCs such as Ifremer indata catalogues. ISO 19131 relates to "Data product specifications", while ISO 19157 "Data Quality" is used to assess the quality of the data sets against their specifications.

Above all, they provide a common framework for the evaluation both from the producer's point of view i.e: "How well does a data set reflect its universe of discourse as defined in the data product specifications" and from the user's point of view i.e: "How well is a dataset able to satisfy the requirements of the user's application", which DG MARE refers to in its call.

While the ISO 19157 principles of data quality initially described in ISO19113 have been applied by spatial data providers for a long time (R. Devillers et al., Thirty years of Research on Spatial Data Quality : achievements, Failures, and Opportunities, Transactions in GIS,

2010, 14(4) :387-400), their application at sea basin scale to define, collate and report data adequacy assessments from data end-users had no equivalent at the time of project inception.

However, the other ISO standards were already used by a wide panel of data providers and NODCs such as Ifremer for spatial data discovery, viewing and downloading services in application of the EU INSPIRE Directive. The recent implementation of ISO19157 in the Geonetwork open source metadata platform made its use much easier in cataloguing services. This context determined the methodological choice of the MedSea checkpoint taken forward by the Atlantic checkpoint.

3.3 Quality metrics

The quality metrics were defined according to the ISO 9004 requirements, i.e. “focused on user satisfaction” (ISO 9004) and based “on factual evidence”. They must be “SMART” to provide indicators usable to implement corrective actions when performance is not in conformity with objectives. The SMART principles (which come from the “Management By Objectives” - MBO) specify that indicators must be:

- Specific (or significant): target a specific area for improvement;
- Measurable: quantitative for progress monitoring (and thus reproducible);
- Assignable (or actionable): agreed upon;
- Realistic i.e. achievable given the available resources;
- Time-bound;

3.3.1 Availability conditions

The ISO19157 Data Quality standard describes quality classes relevant for quantitative evaluation of data *sensu stricto* and provides guidelines and examples of measure for its implementation (next section). However we only found the «Technical guidelines for the implementation of INSPIRE dataset and service metadata based on ISO19139» with (a few) examples of application to services. This is oriented towards the technical evaluation of the conformance of services to the INSPIRE directive, a task out of the scope of the challenge.

No obvious realistic measurement could be defined by the challenge users except for responsiveness of data downloading services. In DAR1, we adopted six non-quantitative criteria with scores from low to high adequacy was assigned to the datasets identified in the Literature Survey as potential sources of data for the challenge applications.

Information was compiled from dataset metadata, data and provider portals, literature or request to data providers or through social networks. The assessment was based on over 650 data sets. In spite of this high number, it may have been slightly biased because at this early time in the process, before making the challenge products, only partial attempts to download the datasets were carried out for scrutiny. However their value should not be underestimated: the uncertainty of the assessment reflects the limits of observations not correctly informed or when standardisation of web portals is lacking (e.g. data policy buried deep-down in a web site). This has been duly reported in DAR1.

DAR2 provides a complementary view on the 104 datasets actually downloaded and used to make the products. A few of them with limited availability hampering the making of products in whole or in part (referred to as “components not covered”) were reported by the challenges in Sextant using the adequate descriptive fields. DAR2 grouped them in function of the indicator targets, either Data policy, Responsiveness or Data formats (readiness). But other

recurrent limiting factors such as data being scatter and lack of information on data (quality metadata) not evaluated in DAR1 have also been reported here.

Table 1: Indicators of availability

Targets	Indicators
Visibility of data and availibility conditions	Ease to find
	Visibility of data policy
Conditions of access	Delivery (services to discover, view and load data)
	Data policy
	Data format (readiness for use)
Performance	Responsiveness (from request to delivery including delays due to policy procedure)

Since DAR1 delivery, the challenge experts have used the experience on real access to data to inform on availability limitations. This information can be found in section 4.2.6 of this DAR2 report (more comprehensive than the information shown in DAR1) and lead to the final recommendations for data adequacy.

3.3.2 Appropriateness

ISO19157 defines five main classes of data quality determining data adequacy: Completeness, Logical consistency, Thematic accuracy, Temporal quality and Positional accuracy, themselves split into 15 subclasses (detailed in DAR1). Positional accuracy was not used by the challenges owing to the coarse spatial scale they worked at.

Ten quality metrics were defined by the checkpoint. They are quickly summarized below and described in Table 2.

Table 2 : Indicators of appropriateness

ISO quality element	Metric name	Definition	Unit
Completeness	Horizontal coverage	Surface area covered	km²
	Vertical coverage	Vertical depth covered	m
	Temporal coverage	Time span covered	day
	Number of items	Object type (country, species etc.)	Occurrence
Consistency	Number of characteristics	-	-

Accuracy	Horizontal resolution	Mean horizontal interval	m
	Vertical resolution	Mean vertical interval	m
	Temporal resolution	Time lag	day
	Thematic accuracy	Percentage	%
Temporal quality	Temporal validity	Data freshness (time since last update)	day

The indicators are the differences between the obtained and specified values of the metrics selected to assess the adequacy of the input data for a given product. They are expressed as percentages of the specified value and signed in such a way that positive values mean better than specified while negative ones mean less good than specified. Differences greater than 100 % in absolute value are truncated to 100 for display. In addition, a field is associated to each metric to comment the result (especially to specify the object type when determining a “number of items”).

The usability class designed by ISO aggregates the results to report quality aspects. Aspects that could not be measured were collated as expert opinion on the capacity of the data to satisfy the challenge requirements. A usability score scaled from 1 (inadequate) to 5 (excellent) integrating the quantitative and non-quantitative results of the assessment is attributed by the expert to sort out the assessment results.

3.4 Assessment process

The assessment is a two-step process (Figure 2) to determine:

- How each product meets the objectives of its production specified with the metrics and values selected by the challenge expert;
- The causes of eventual inadequacy by looking at the contribution of each input data set used.

3.4.1 Definition of the assessment units

The assessment is carried out by “quality units” alias “components” which make part of the product specifications. A component is a “combination of a scope and data quality elements” (ISO19157: see § 3.2.2 for quality elements or classes). The scope “specifies the extent, spatial and/or temporal, and/or common characteristic(s)[1] that identify the data on which data quality is to be evaluated”. In other words, a component is a view of a subset of a dataset to which selected quality measures are applied. A scope usually corresponds to one of the characteristics needed to create the product (within the meaning of the Checkpoint terminology). Several of them are sometimes specified when they share the same requirements (e.g. due to sampling location consistency).

To be able to evaluate and report the specifics in relation to data sources depending on geographical areas or periods of time, several components were defined with different spatial or temporal extents (e.g. Europe/North America/Africa or 10/50/100 years).

[1] The term « characteristic » refers here to any property which characterizes the scope of the subset

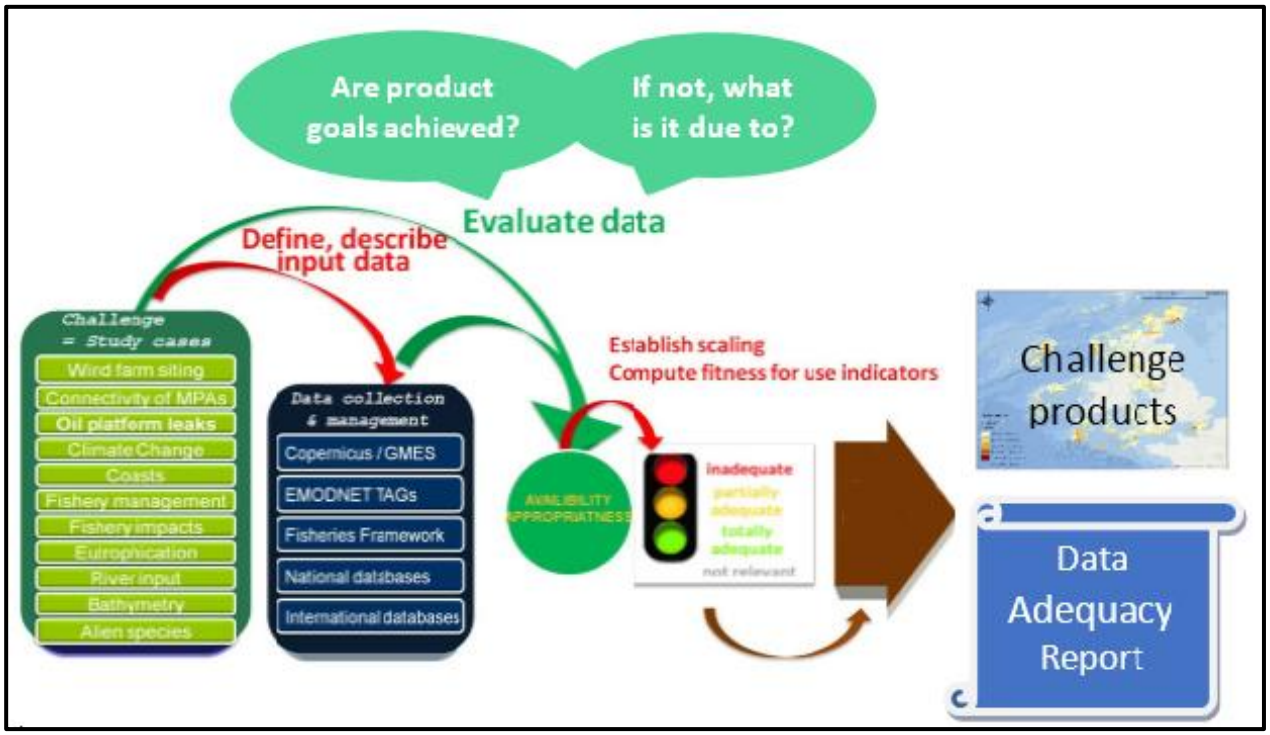


Figure 2: Assessment process

3.4.2 Organisation of the assessment information

The key assets of the assessment work are the catalogues.

The catalogue of Products contains:

- the Data Product Specification (DPS) split into one or several components (138 for 53 products in the Atlantic checkpoint);
- the achieved Data Product (TDP) description which is linked to the DPS. It is split in as many components as in the DPS, each one containing the assessment results unless issues in getting data prevented doing it (next section). The product components are linked to the Upstream Datasets (UDs) selected to provide the required characteristic(s).

The catalogue of UDs contains the description of the datasets which the challenges attempted to use as input data to their product components. Each such UD is evaluated against the specifications of this component and the assessment results are kept in an eponymous component of the UD description. So a given UD may contribute to several components and depending on the specifications of this component, be appropriate or not. As an example, a 100m depth DTM can be adequate for a component of the Windfarm challenge while inadequate for the Bathymetry challenge dealing with new navigation fairways.

Figure 3 illustrates the way Upstream Data and Products are linked. Starting from the left, a catalogue of data sets stores all data sets anticipated for use in the literature survey and assessed for their availability in the first phase of work (DAR1). 669 data sets were primarily identified at that time. There is a “N to one” relation between these 669 data sets and the 82

characteristics deemed necessary to make the challenge products because of course for a given characteristic, many data sets can be found.

At the right side are challenge products, 53 of them in the Atlantic case. These products have components related to one or more characteristics sharing the same quality requirements.

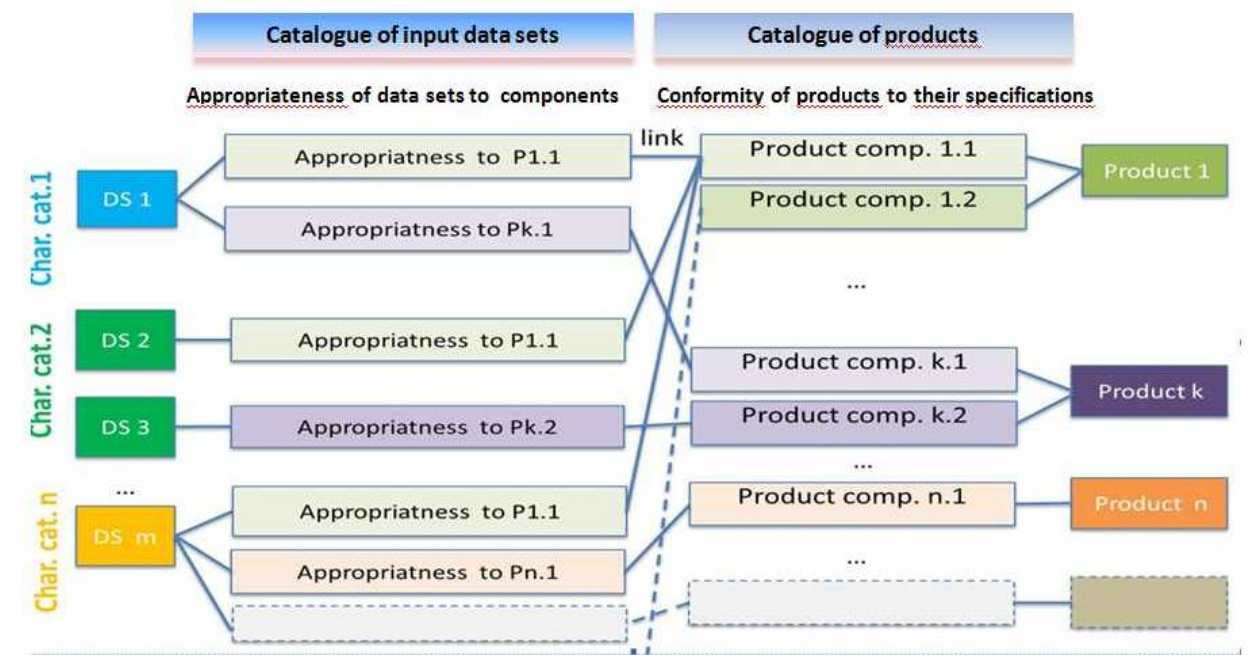


Figure 3: From data sets to products

Each dataset (DS) refers to a unique characteristic allowing to group datasets by categories of characteristics (P02). This organisation allows to search assessment results and specifications either by challenge application (and product component) or by category of characteristics.

To illustrate this, the Oil leak challenge that has two distinct strands of work:

- the oil spill drift component using dynamic real time but rather low resolution data;
- the impact at the coast using static high resolution data with longer temporal validity.

To simplify the data base and avoid heavy metadata capture work, components were created for each specific strand of a product. A product can have several components (in this case 138 for 53 products) and each component uses one or more data sets (components and data sets are in a “M to N” relationship).

So a given data set may contribute to several components and, depending on the specifications of this component, be appropriate or not.

3.4.3 Typology of issues

The components may not have been covered (i.e. created) depending of the existence or availability of data. Three situations may occur:

- The component is covered, which means data could be used to create part of the component, although of course there may be gaps in it. In this case appropriateness metrics are informed and the indicators are computed. A component “usability” score of from 1 to 5 (inadequate to excellent) is given by the challenge, along with explanatory comments taking into account factors that could not be measured;

- The component could not be covered to a satisfactory level of thoroughness because most of the necessary data were not available. In this case appropriateness is not assessed, but the reason for the lack of availability is given: either a policy, responsiveness or readiness issue or any other factor such as data scattering made collation impossible in the time frame of the project;
- The component could not be covered to a satisfactory level of thoroughness because most of the necessary data did not exist: none have been observed or measured (meaning: with the appropriate level resolution).

In the latter two cases, no usability score was produced at the component level but a description of the causes of the failure was provided by the expert.

3.5- Reporting

3.5.1 Challenge reports

Challenge leaders were requested to produce challenge reports. Although not contractual, these reports provide the following information:

- The outline of the challenge and its products;
- The challenge and products scores;
- A synthesis per characteristic giving the recommendation for improvements as well as the contributing data sets and their limitations.

In particular, all the comments produced by the challenges in the Sextant catalogues are repeated her in a more synthetic way by providing:

- The expert opinion justifying product scores;
- The “usability” of the components, also justifying their scores and;
- The usability of the data sets, i.e. how effectively they contributed to a component.

In Annex 5 to 7, all characteristics are listed, along with the relevant components and the availability and appropriateness of the input data sets. Data sources are also listed, with a focus on those having hampered the making of the product.

3.5.2 Assessment by product

This is a quick reminder of the structure of the checkpoint into 11 challenges (see DAR1), each of them featuring a number of products in keeping with the terms of reference. The Atlantic checkpoint made 53 products, whose majority are GIS products and a few are data spreadsheets in cases where the spatial component was irrelevant or not representative.

Some challenges preferred to show individual basic products that perhaps would need to be combined later rather than attempt at producing a too complicated and probably worthless result. What was important in the end was to assess the contributing data sets whichever way the products were achieved.

The number of products may differ from the expected one when reading the tender. This is due to the choice of each challenge leader to make more meaningful elementary products rather than complex ones, especially in areas where science still lacks to make meaningful integrated products (e.g. eutrophication).

- In the case of MPA, three issues were raised by the tender but five products were deemed necessary to best render them;
- Climate and coasts had statistics to produce for various time spans and lags and chose to make one product per time specification;

- The same was done for rivers where each characteristic of fresh river inputs was assigned to a product;
- Eutrophication used both a global and a local example for chlorophyll assessment based on models as well as separate products using *in situ* observations.

The table below gives the number of products per challenge.

Windfarm	MPA	Oil leak	Climate	Coasts	Fisheries Managt
4	5	2	8	9	3
Fisheries impact	**Eutrophication**	**Rivers**	**Bathymetry**	**Alien species**	
2	7	9	3	1	

To report the conformance of the products to the challenges goals, a summary of the scores of the covered components and the number of components not covered is made to produce an overall product score (also going from inadequate to excellent) along with the associated "expert opinion" at the product level. For example in the Eutrophication challenge, there were products that could not be achieved on the ground that "There was not enough *in situ* data to meet the requirements of spatial and temporal resolution in the study area. The lack of seasonal measurements in most of the study area, in addition to the lack of coordination between measurements was a limitation to generate data products for eutrophication'".

3.5.3 Assessment reporting by data characteristic

The data adequacy assessment is based on data investigations made by the eleven challenges, from both the data availability and the data appropriateness standpoints, with availability aiming to address the question: "How are data collated to make products?" and for appropriateness: "What are challenges going to do with them and how properly"?

From the initial 672 data sets identified in the literature survey, only 104 were really used to make the products, an approximate 15%. This figure is quite in line with other checkpoints (18 and 34%). Data availability had been assessed in DAR1 by looking at all these data sets. Further assessment of data sets showed that a number of them were discarded for a variety of reasons:

- Most importantly the literature survey and its identification of data sets were made prior to writing products specifications, which lead to overestimate the number of related data sets;
- Data sets selected in the first place may have been deemed out of scope of the challenge upon more focused examination;
- The highly patchy nature of some data sets, the redundancies between data sources (original data sets versus assembled ones);
- Finally the format of some data sets made them improper (e.g. reports in pdf form).

As an example, from the 41 data sets identified for Alien species, some of them were later pooled together to form one unique reference, resulting in the reduction in the number of source data sets eventually used.

Ten indicators of appropriateness were selected to properly assess the data sets. These were computed only when the components were actually “covered”; in the case of “not covered components” appropriateness was not assessed, but in cases where data existed availability was assessed. Then the results of the appropriateness assessment were grouped to provide an overview of the existing data adequacy by category of characteristics (P02), by quality indicators, by level of satisfaction using graphical representations allowed by the standardization of the information.

3.6 Checkpoint services

The Checkpoint service allows end-users to search, discover, display and download input datasets and products through different tools and guides producers and users to develop best practice and synergies. A sustainable infrastructure was set up during this project by using the Sextant platform, a Spatial Data Infrastructure for marine environment based on three main modules which have been designed according to the European INSPIRE directive for interoperability and according to the OGC and ISO TC 211 standards

3.6.1 The metadata catalogue service

The catalogue service is part of the checkpoint service and it is accessible from the web portal main page: http://www.emodnet-atlantic.eu/Checkpoint-service/Browser

From this catalogue service end-users can easily access upstream data (UDs) metadata and products (specifications and actual products). The catalogue uses the Geonetwork software to set up Catalogue Services for the Web (CSW) and to design the form used to edit the appropriate metadata. Metadata are recorded in this database using the standards defined by the Open Geospatial Consortium (OGC) and the Technical Committee ISO TC 211, Geographic information and Geomatics (ISO 19115 and ISO 19157).

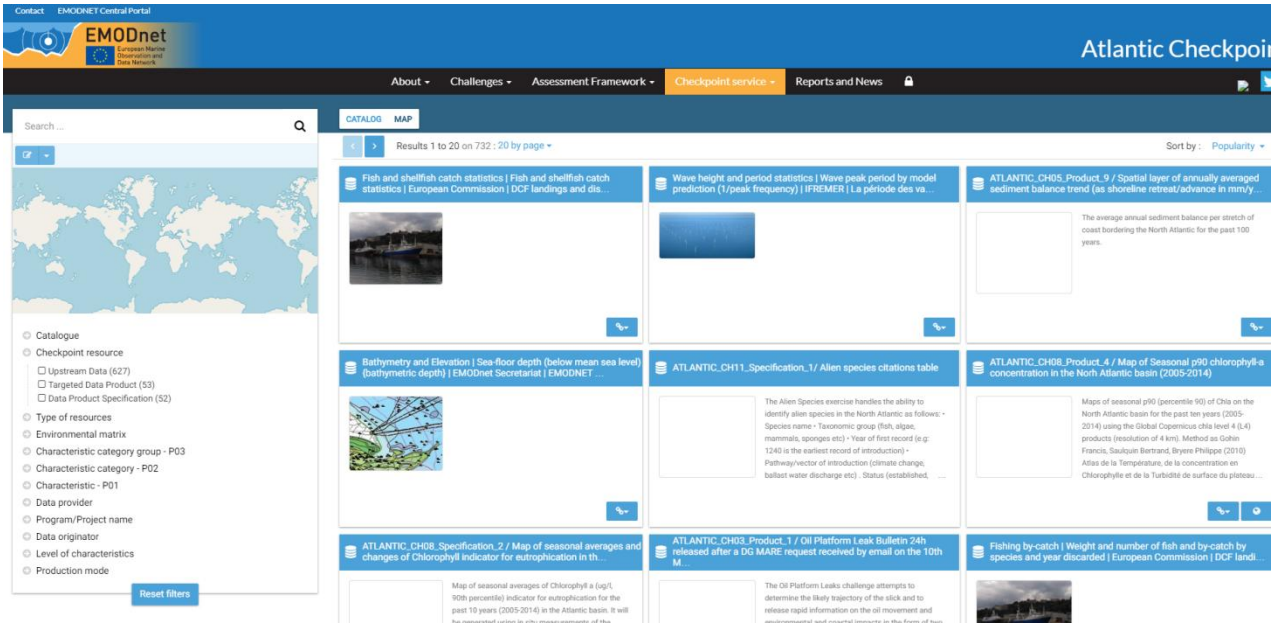


Figure 4: Metadata catalogue

3.6.2 The Viewer

A Web GIS has been implemented to display layers and products created by the challenges. These products can be viewed in the Web GIS and can be downloaded from the challenge web pages. In the Web GIS users can consult and explore products (navigation, zoom tools, query layers, export map, etc). All the products are displayed through a specific Web Map Service: http://www.ifremer.fr/services/wms/atlantic checkpoint (for use only in GIS software).

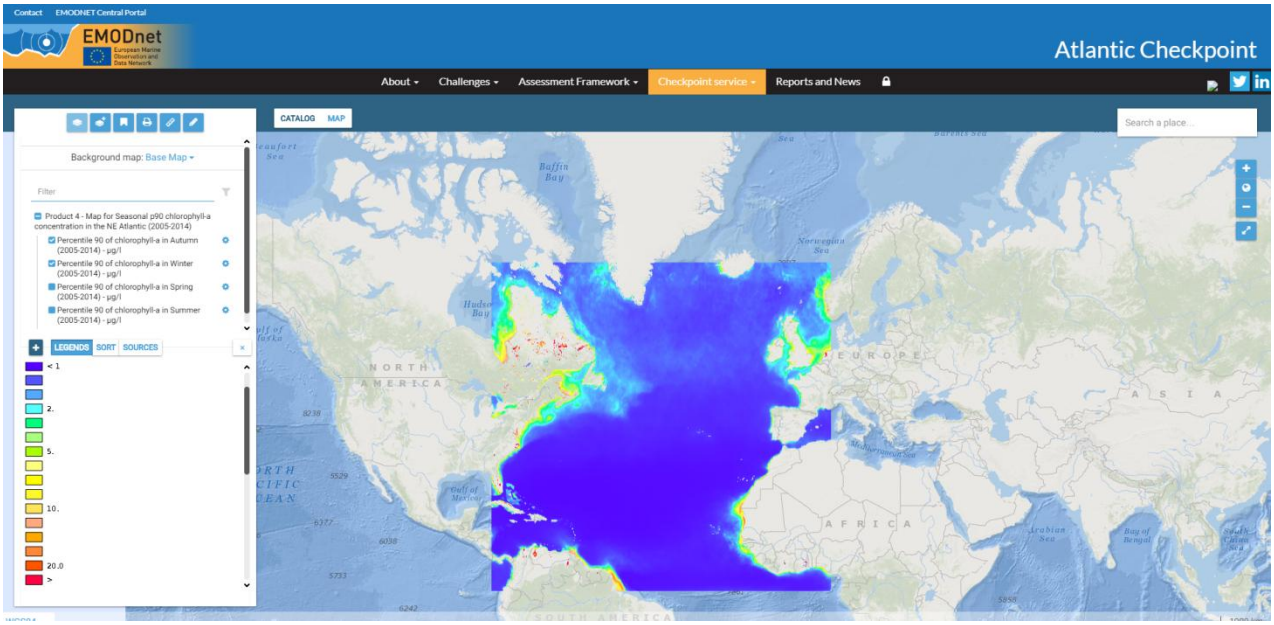


Figure 5: Web GIS platform – Map for Seasonal p90 chlorophyll-a concentration in the NE Atlantic in Autumn (2005-2014) - µg/l

3.6.3 The online services

The availability indicator assessments are available in the web portal:

http://www.emodnet-atlantic.eu/Checkpoint-service/Availability-assessment

The online services use Kibana open sources web tools and presents the indicators automatically produced from the metadata catalogue content. It aggregates values and allows interactive filtering on challenges. The colors illustrate the degree of satisfaction determined by comparing actual conditions of availability to the expected ones (user requirements) with the following general meaning:

- Red: actions are required to provide fit for use datasets and services;
- Green: services are fit for use and must be maintained.

It displays dynamic graphics allowing non-expert public to assess the fitness for use without spending a lot of time looking at metadata and checkpoints reports

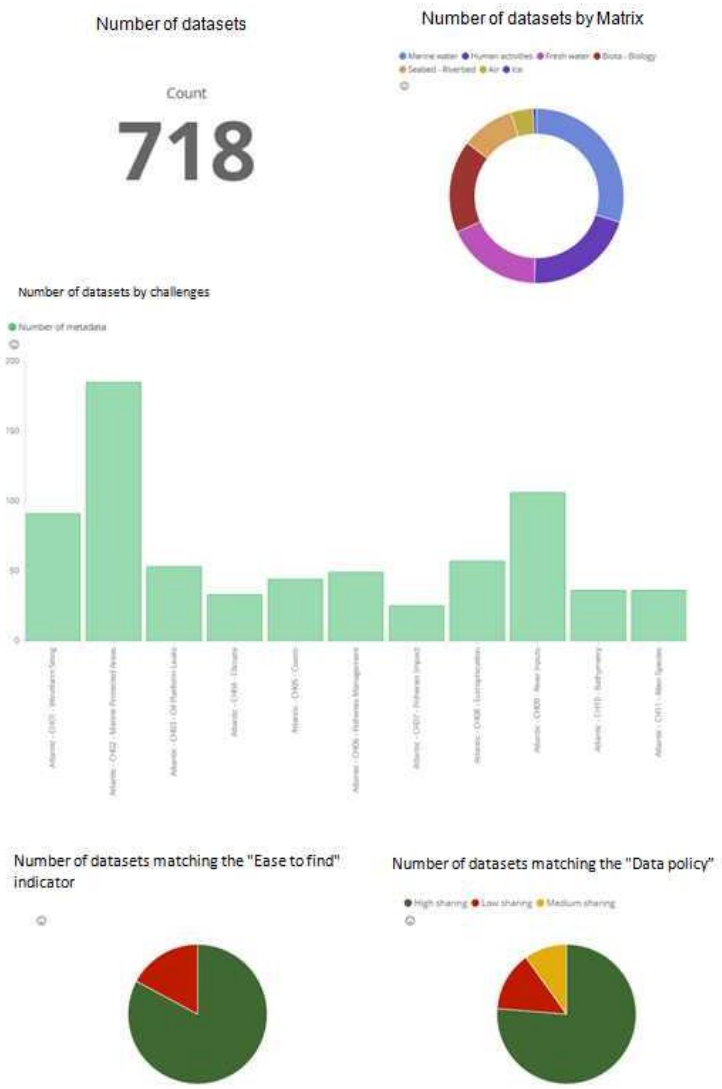


Figure 6: The online services

3.6.4 Checkpoint services from the web portal

From the Challenge web page a table gives access to the following (figure 7):

- Specifications of products;
- Results: products metadata including appropriateness and use limitations;
- Download: direct access via a DOI to the landing page of the product which contains a product overview and a download service (figure 8);
- View map: geospatial layers in the Web GIS;
- Product score.

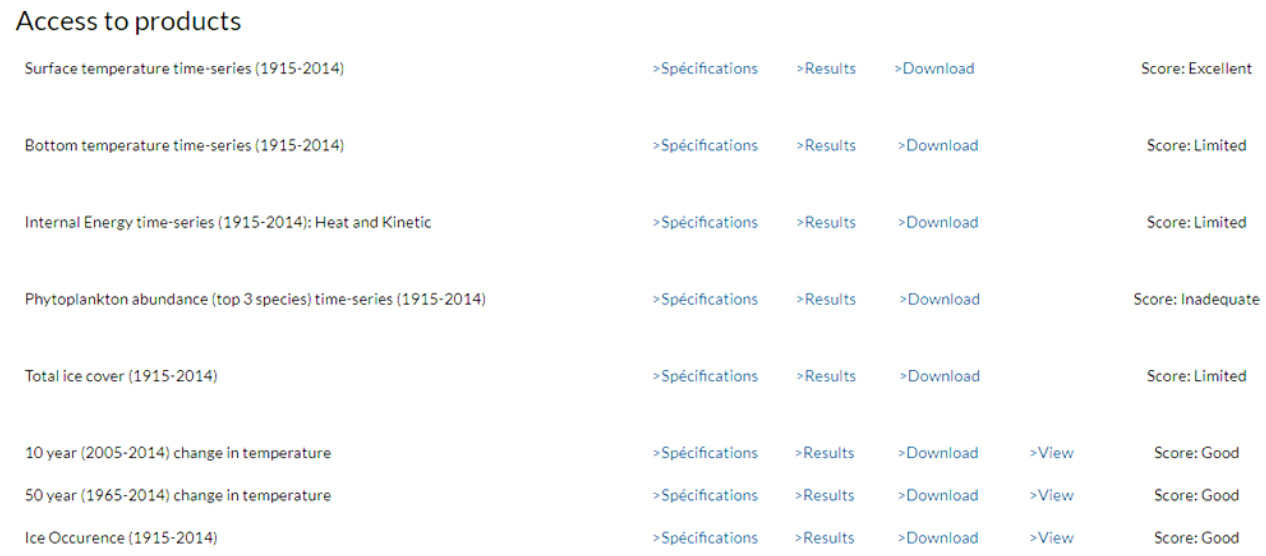


Figure 7: Example from Climate

The online visualisation of appropriateness indicators could not be implemented during the project time span (nor was it in the the tender specifications). However to navigate among the thousands of measures and indicators, a dynamic web interface allowing to select and view the adequacy assessment and products specifications in graphical form is going to be developed in the next few months. This is an innovative area of work whose results are intended to be presented at the IMDIS 2018 conference.

ATLANTIC_CH05_Product_5 / Spatial layer of relative annually averaged sea level trend (in mm/year) for period of 50 years

Access to data and metadata

Link to the data services and to the full metadataset

Date(s) 2018-01-10 (Creation)

Author(s): Irish Marine Institute

Point of contact(s): EuroGOOS AISBL, Irish Marine Institute

Publisher(s): EMODnet Atlantic Checkpoint

DOI 10.12770/49c6dab4-c95f-4ff7-be56-739905157745

Abstract This product attempt to follow up on the sea level rise per stretch of coast of the North Atlantic, over 50 years as follows:
• Characterization of absolute sea level trend at annual resolution, along the coasts of EU Member States (including Outermost Regions), Canada, Faroes, Greenland, Iceland, Mexico, Morocco, Norway and USA;
The stretchs or coast are defined by the administrative regions of the Atlantic Coast:
• from NUTS3** administrative division for EU countries (see Eurostat), and
• from GADM*** administrative divisions for non-EU countries.
** Third level of Nomenclature of Territorial Units for Statistics
*** Global Administrative Areas

For relative sea level trend for 50 years we extract the information from coastal tide gauges data available at each stretch of coast, if there is not a tide gauge there is a data gap. The product is provided in tabular form and as a map layer.

Keywords Atlantic - CH05 - Coasts, Observation (raw, QC), Delayed, Oceanographic geographical features

Utilisation Limited: The product fails in most aspects to meet the objectives but meet some.
It was only possible to produce an average annual relative (to the land) sea-level rise trend for 50 years for only some stretch of coast required thanks to the existence of some long-term sea level tide gauges with enough temporal resolution. Input data is available in a well-documented format. The limitation is that there is not a tide gauge for all the stretches of coast (i.e. administrative divisions), and some of the time series are shorter than 50 years. Moreover, it has been assumed the value of a single tide gauge to all the length of the stretch of coast.
This EMODnet Atlantic Checkpoint data product (the "Product") is the result of a "challenge" stress test specified by the call for tenders No MARE/2014/11. Irish Marine Institute (the "Producer") was commissioned to perform a number of tests to evaluate marine data according to set user requirements. This product and any resultant data produced is a consequence of these tests and provided for interest only. The Product is provided on an "as is" basis, and the producer disclaims any and all warranties, conditions, or representations (express, implied, oral or written), relating to the Product or any part thereof, including, without limitation, any and all implied warranties of quality, performance, suitability or fitness for a particular purpose to the full extent permitted by law. The Producer accepts no liability for loss suffered or incurred as a result of reliance on the Product or any output of its use. The Producer shall not be liable for any economic loss, loss of profit, loss of business or contract, depletion of goodwill or otherwise, in each case whether direct, or indirect or any indirect loss whatsoever (howsoever caused) including loss or damage suffered as a result of an action brought by any person which arise out of or in connection with the use of the Product or inability to use the Product.

Temporal Extent 1965-01-01 - 2014-12-31

Associated resources

Access to the website

How to cite

Irish Marine Institute (2018). ATLANTIC_CH05_Product_5 / Spatial layer of relative annually averaged sea level trend (in mm/year) for period of 50 years. EMODnet Atlantic Checkpoint.
http://doi.org/10.12770/49c6dab4-c95f-4ff7-be56-739905157745

Figure 8: http://doi.org/10.12770/49c6dab4-c95f-4ff7-be56-739905157745

4 Data analysis

4.1 Assessment of challenges

The challenges were split in 53 products. Table 3a gives the percentages of products fkr each score. Scores go from 1 to 5, resp. inadequate to excellent, with the colour coding shown below. The medium mark 3 (yellow) means “good”, in other words products that were achieved to at least 50% of the requirements.

Table 3a: Percentages of challenge products according to their scores

1	Inadequate		21%
2	Limited		43%
3	Good		23%
4	Very good		13%
5	Excellent		-

From this table the average score is 2.28 and 64% of the products were deemed of limited value by their creators. Individual products can be seen with their score in table 3b.

These products were themselves split into a total of 206 components or “quality units” (see components spreadsheet in Annex 1), which allows to assess adequacy with a finer “grain size”. 44 components (21%) were not covered and 162 were covered. The reason for components not being covered is either because of sheer lack of data or because of data too difficult to get and use. Components not covered were not given a score, however challenge experts gave a comment of the reason for this.

The ones covered were given a score reflecting how thoroughly they had been produced, based on the comparison of the quality measures values between the DPS (Data Product Specifications) and the TDP (achieved data product) in the database. The average component score derived from the 206 components in Annex 1 is 2.74, which means that on average these 162 components meet more than 50% of the objectives specified by the challenge.

The discrepancy between the two scores is easily explained by the fact that the challenge experts, upon scoring their products, took into account the components not covered, which of course brought down their overall score.

It is noteworthy that when a component was said to be inadequate or limited, this does not mean that the data sets are of bad quality. It only means they were not appropriate for the specific product targeted by the challenge, but may well be sufficient for another purpose with other types of requirements.

Challenges that at first glance appear to have most difficulties in correctly achieving their products are Eutrophication, Rivers, Coasts or Oil leak, because data did not meet the requirements. Reasons for these difficulties are various but they have a lot to do with the questions asked to the challenges : some of them have a narrow scope whether others have a much broader one. It is expected that the latter had to overcome stronger data issues than the former. So these scores shown in Table 3 should be regarded in that context. .

Table 3b: Individual products and their scores

Windfarm	MPA	Oil leak	Climate	Coasts	Fisheries Management
Product_1	Product_1	Product_1	Product_1	Product_1	Product_1
Product_2	Product_2	Product_2	Product_2	Product_2	Product_2
Product_3	Product_3		Product_3	Product_3	Product_3
Product_4	Product_4		Product_4	Product_4	
	Product_5		Product_5	Product_5	
			Product_6	Product_6	
			Product_7	Product_7	
			Product_8	Product_8	
				Product_9	

Fisheries Impact	Eutrophication	Rivers	Bathymetry	Alien species
Product_1	Product_1	Product_1	Product_1	Product_1
Product_2	Product_2	Product_2	Product_2	
	Product_3	Product_3	Product_3	
	Product_4	Product_4		
	Product_5	Product_5		
	Product_6	Product_6		
	Product_7	Product_7		
		Product_8		
		Product_9		

4.2 Assessment of characteristics

4.2.1 Analysis process

Among the 206 components, only the 162 “covered” ones were subject to appropriateness assessment. From these 162, only the ones scoring from 1 to 3 (inadequate to good, resp. 1 to 3) were kept for analysis, a number of 126.

These 126 components were assessed in terms of their “usability”, a synthesis of the 8 quantitative quality measures consisting in a score reflecting these measures (1 to 5, resp. inadequate to excellent).

The challenges specified 43 P02s. For 23 of them the anticipated data sets could not be found, either because they did not exist or were not available (Annex 5 and 6).

27 P02s did contribute to products, were assessed for appropriateness but did not meet the specifications. These can be seen in table 4 along with their respective quality measures.

Table 4: Categories of characteristics and related quality measures

ADMINISTRATIVE UNITS
NUMBER OF ITEMS
BATHYMETRY AND ELEVATION
TEMPORAL RESOLUTION
NUMBER OF ITEMS
CONCENTRATION OF SUSPENDED PARTICULATE MATERIAL IN THE WATER COLUMN
TEMPORAL COVERAGE
NUMBER OF ITEMS
FISH ABUNDANCE IN WATER BODIES
TEMPORAL COVERAGE
NUMBER OF ITEMS
FISHERY CHARACTERISATION
NUMBER OF ITEMS
FISHING BY-CATCH
NUMBER OF ITEMS
FISHING EFFORT
TEMPORAL COVERAGE
HORIZONTAL RESOLUTION
NUMBER OF ITEMS
HABITAT CHARACTERIZATION
NUMBER OF ITEMS
HABITAT EXTENT
NUMBER OF ITEMS
TEMPORAL VALIDITY
HEAT FLUXES BETWEEN THE WATER COLUMN AND THE ATMOSPHERE
TEMPORAL COVERAGE
HORIZONTAL SPATIAL CO-ORDINATES
HORIZONTAL COVERAGE
HORIZONTAL VELOCITY OF THE WATER COLUMN (CURRENTS)
TEMPORAL COVERAGE
TEMPORAL RESOLUTION
VERTICAL RESOLUTION
HORIZONTAL RESOLUTION
INDUSTRIAL ACTIVITY
NUMBER OF ITEMS
INVASIVE SPECIES MONITORING PARAMETERS
NUMBER OF ITEMS
LITHOLOGY
HORIZONTAL COVERAGE
HORIZONTAL RESOLUTION

We only retained P02 featuring scores from 1-3, i.e. inadequate, limited or good, the latter featuring cases where components were achieved to 50% of the expectations;

- For components “not covered”, although appropriateness was not assessed (meaning there were no quality measures), challenge leaders were requested to inform the reason for not making the product. This is of course of paramount importance because components not achieved mean very serious data issues;
- Any time a P02 was deemed of low quality, bar charts were made (Annex 3) to enable the readers to see in quantitative terms across the challenges what their specifications were for this P02 and how far they were from meeting them. Of course it may happen that the products of a given challenge were feasible with a certain dataset while

another one had more stringent specifications and could not meet them when using the very same dataset.

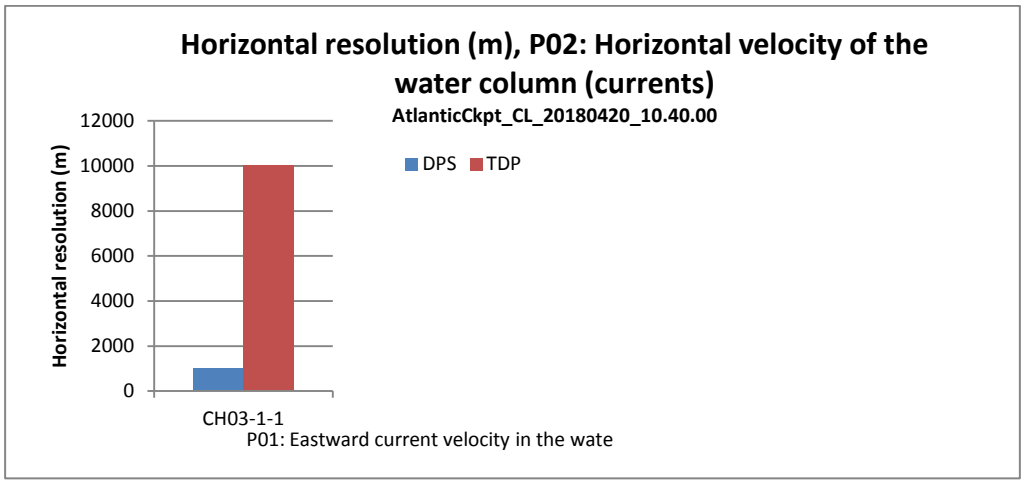
- Additionally when more explanations were needed we used the challenge reports where usability was again reported and the failing characteristics enhanced, along with the data sets at stake.

4.2.2 P02 bar charts

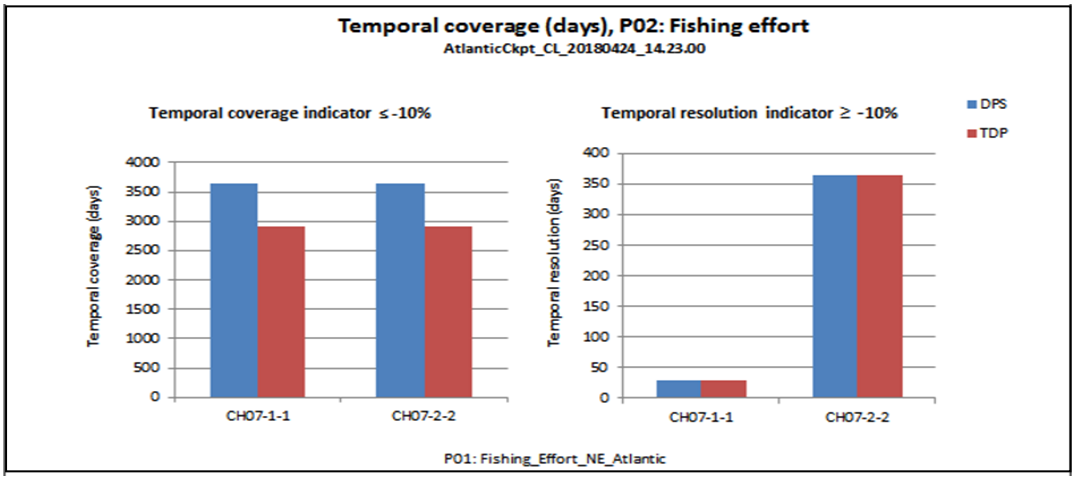
Bar charts illustrate, for each P02 and through the components having this P02 in their lineage, the quality measures per indicator specified (DPS in blue) and achieved (TDP in red). The P01 is specified at the bottom of the chart (beware that some texts being too long are truncated) and the component identifier in abscissa in noted for "CH03-1-1". The component full names can be found in Annexes 5 and 6 tables using a P02 entry.

The rationale was based on the assumption that a low resolution (in either dimension) characteristic, no matter its coverage, is improper to address an issue. To help understand these different cases, examples of the categories of characteristics "Horizontal velocity of the water column" and "Fishing effort" are given below:

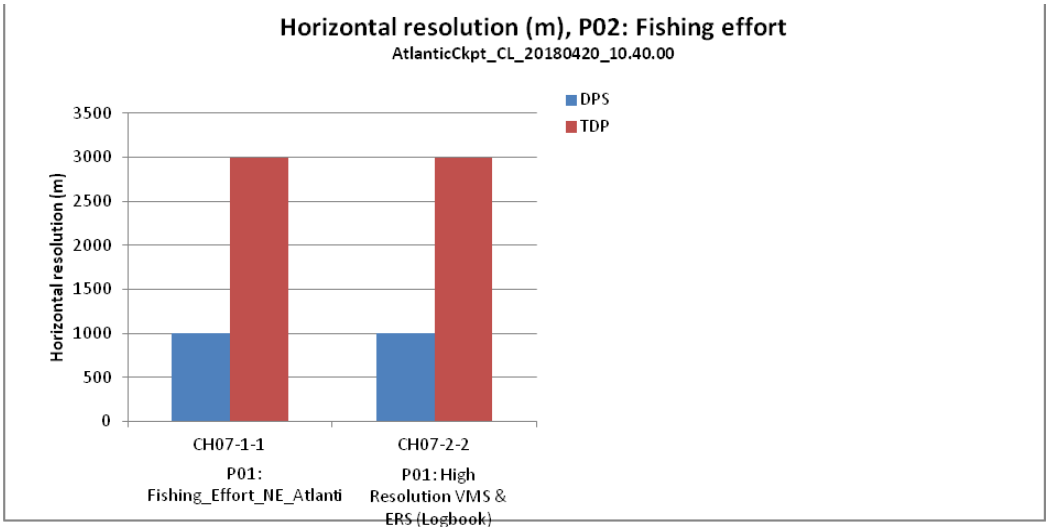
- If resolution - in either dimension, horizontal, vertical or temporal) - was no good, coverage was disregarded. This was the case for currents, only used in component 1, product 1 of the Oil leak challenge (CH03_1_1) in the figure where horizontal resolution appears at 10% of that specified (10km in TDP instead of 1km in DPS). It was therefore worthless to look at coverage and the unique message should then be: "resolution to be improved", along with implicit full coverage of the Atlantic.



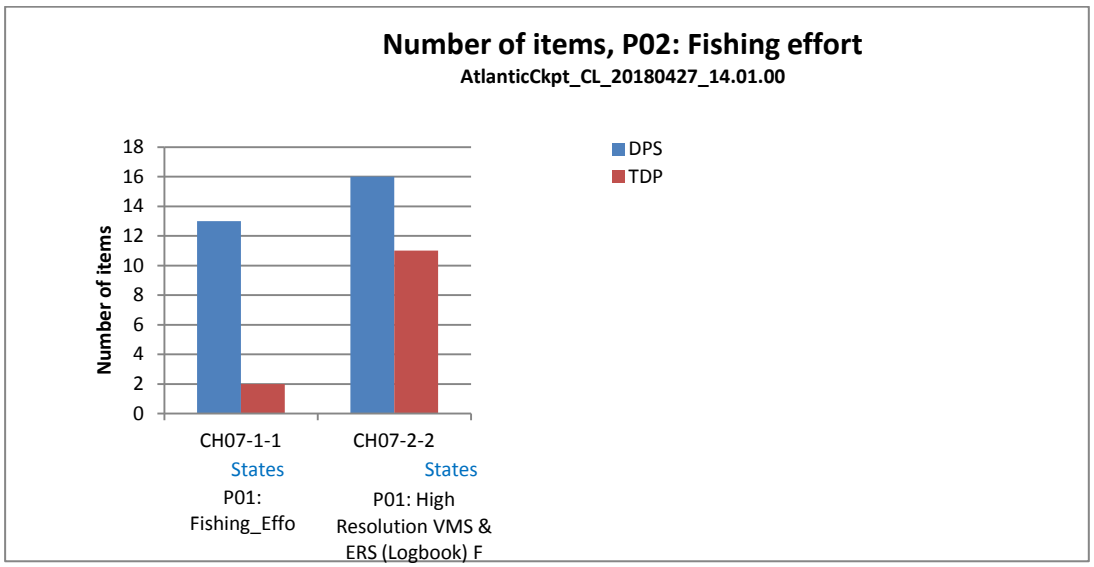
- If resolution was satisfactory, then coverage was investigated, as illustrated below for P02 "Fishing effort" used in 2 components of challenge 7 Fisheries Impact. Temporal resolution (right plot) was adequate for both component 1 of product 1 (CH07_1_1) and component 2 of product 2 (CH07_2_2), albeit at very different values, resp. one month and one year, as shown by the identical bars for the specifications (DPS) and the product (TDP). Then we looked at temporal coverage (left plot) and identified a slight shortage of 20%, i.e. 8 years achieved instead of 10 years specified. The recommendation would then be to try and extend the time period to 10 years.



Additionally fishing effort has a horizontal component and only 3km resolution was achieved where 1km was required. Note that although referring to the same P02, two different P01 in two different components were created for this challenge.



On top of this, there was a gap in fishing effort resulting from not all countries reporting movements of their fleets. In order to catch this information, the measure “”Number of items” was assigned to the representation of Countries – named States in the plots.



Also note that the measure "Number of items" was used in many components to reflect either the coverage for characteristics not amenable to a surface area in km² (e.g. number of tidal gauges, rivers etc.) or any type of reporting entities such as large spatial units used to report Salmon or Eels in the River output challenge.

"Temporal validity", a quality measure mostly applying to static data, was hardly used at all by challenges mainly because many challenges were asked time-specific question (e.g. real-time products) where the temporal validity is covered by the coverage element. This may be due to challenges specifications (DPS) and how they dealt with data sets that were not up to date enough to inform their product's metadata (TDP). In some cases, they found no need to inform on temporal validity. It may also be that for a number of data sets on human activities, the age of the data was not specified.

Finally, "Precision" was not requested from the challenges for this round in spite of its relevance. We justify this by the fact that precision is seldom assessed, let alone informed in metadata, so it would have been mostly impossible for challenges to get it. Where challenges mentioned it in their challenge reports, we added the information to this text.

4.2.3 Categories of adequacy issues

As was mentioned above this report intends to focus on data appropriateness, however as we analysed data we found out that a number of data availability issues were raised by the challenges. Additional issues were identified that do not really belong to either type, such as assembly needs.

Adequacy issues encountered were classified into the three following types, which gives the architecture of this section:

- Data gaps which would need new data collection: this concerns the horizontal, vertical of temporal domains. Gaps may be geographic (some areas are simply not covered), or resolution gaps when coverage is continuous but mesh size insufficient. These are reported in 4.2.4.
- Assembly needs which would need to be covered by the creation of Thematic Assembly Centres (TAC were primarily introduced by GMES). This concerns either data

scattered but taken care of by various providers or even data disregarded, i.e. object of no collation at all. These are analysed in section 4.2.5. TAC

- Finally Availability needs (called “Services issues” in 4.2.5), which can resort to:
  - Either policy in case of restricted data;
  - Or the status of quality metadata in cases where completeness of characteristics is not informed (i.e. data centres not applying yet the ISO19157 rules);
  - Or other technical availability issues such as responsiveness or readiness.

Another frequently reported issue that can be said to resort to availability is metadata completeness. Note that we distinguish in the above “standard metadata” (ISO 19115) and “quality metadata” (ISO 19157).

4.2.4 Gaps in characteristics (P02s)

Table 5 gives a view of the characteristics featuring gaps that: a) either prevented challenges from making their products or b) only enabled them to produce inadequate or limited products. The table entry is by Matrix (broad environmental categories), followed by P02s and P01s, the latter giving more details when P02s are very generic such as Administrative limits. Note that P02s are established SeaDataNet categories of characteristics while P01s are chosen by their creators, which means a broad variety of P01s can be found, or event sometimes no P01 at all.

The needs for surveys are split into their two major strands, i.e. resolution needs and coverage needs. However resolution and coverage are sometimes quite intricate and seamless. When addressing a characteristic produced by what is referred to as a “coverage tool” (typically remote sensing or modelling), the spatial (or time) coverage quite often is rather global due to the nature of these techniques while resolution may be too coarse: this is the case or such variables as wind, currents or waves. When addressing *in situ* data measurements, resolution in no longer meaningful and is superseded by “density of measurements”. This then resorts to coverage rather than resolution albeit being of the same essence. This is why the column “coverage needs” is highly populated by variables belonging to such domains as Rivers, Eutrophication (needing a variety of co-located *in situ* measurements) or Sea level requiring a denser tidal gauges network.

Coverage usually makes the distinction between three parts of the Atlantic basin, i.e. waters of the EU, North America or the rest of the basin. For example broad-scale seabed habitats maps are available for the whole EU marine domain while this is not the case in North America.

In the discussion below Table 5 is further commented by matrix for each P02. Bar charts supporting this discussion are available in D14.2.2, Annex 3.

Air matrix

The air matrix only contains “Wind speed and direction”. These are modeled for the whole globe but lack vertical coverage for renewable energy or oil spill management applications. Adequate resolution would be at 1 km horizontal and 10m vertical.

Biota-Biology matrix

- Fish abundance in water bodies is ill-known, especially for Eel and Salmon, two key species for river environments. There is a need of instrumenting a higher number of rivers and increasing the measurement frequency;

- Phytoplankton generic abundance in water bodies is not known;
- Invasive species monitoring parameters are not reliable enough because a) they would need higher observation density, b) measurement standards do not exist.

Fresh water matrix

Too few rivers are instrumented for a number of parameters needed to model the fate of coastal marine waters: water flow, temperature, salinity, oxygen, concentration of nitrates, phosphates etc.

Marine water matrix

- Horizontal velocity of the water column (currents) is generally needed at at least 1 km resolution to properly address phenomena at the coast, whether for renewable energies or larvae dispersion computations;
- Sea level change at the coast cannot be properly assessed due to the lack of density of tidal gauges long-term time series (up to 100 years). Many tide gauges are already near a GPS station, but co-location is still an issue for most of them: there is not always a geodetic connection between the tide gauge and the GPS, usually not at the tide gauge itself, or this is not done periodically. Also, even if this is solved, getting adequate information of land movement and ellipsoidal height at the tide gauge is not easy to find now for an external user. It is crucial to strengthen the collaboration with the geodetic community: e.g. SONEL, recognized by GLOSS (Global Sea Level Observing System, IOC, UNESCO) as the formal data bank for GPS data collection of national institutions around the world).
- *In situ* chlorophyll pigments concentration, dissolved oxygen, salinity and temperature, along with chemicals such as nitrates and phosphates, are all suffering from a low density and lack of co-located measurements in the coastal zone, which severely limits the ability to obtain data products to assess eutrophication.

Seabed-riverbed matrix

- The composition of the seabed is ill-known. In terms of sediment and lithology, a scale of 1: 250 000 would be necessary to address windfarm siting while the EU broad-scale seabed habitats map should be extended to the whole basin, with particular focus on the coastal zone where challenges are more acute.
- Bathymetry is generally well catered for, except for new needs that appear with increasing maritime traffic and new routes. New surveys are needed for approaches to harbours and new routes along with higher sounding accuracy.
- Lack of enough sediment balance data to compute shoreline advance or retreat.

Table 5: P02 requiring surveys

Matrix	P02 characteric category	Characteristic	Needs for surveys		
			Resolution needs	Coverage needs	Additional
Air	Wind speed and direction	Wind velocity (10m along y-axis) in the atmosphere by model prediction	Horizontal : 1 km Vertical: 10m	Vertical: 150m	
Biota-Biology	Administrative units	Biodiversity critical areas		Whole EU	
Biota-Biology	Fish abundance in water bodies	Eel and salmon	Time: yearly	More EMUs (management units)	
Biota-Biology	Habitat extent	Essential Fish habitat extent		Except USA	
Biota-Biology	Habitat extent	OSPAR Distribution and extent of threatened and/or declining habitats			Lack of reporting comprehensiveness by Contracting Parties
Biota-Biology	Invasive species monitoring parameters	Alien species citation date, coordinates, vector, status, impact		Higher observation density	More (comprehensive program with harmonised protocols)
Biota - Biology	Phytoplankton generic abundance in water bodies	Abundance of phytoplankton per unit volume of the water body by optical microscopy		More CPR (Continuous Plankton Recorders) across Atlantic Ocean	
Fresh water	Phosphates	Concentration of phosphates [mg/l P]		More rivers to be monitored	
Fresh water	Particulate total and organic nitrogen	Concentration of total nitrogen [mg/l N]		More rivers to be monitored	
Fresh water	River flow and discharge	Flow rate [1000m3/d]		More rivers to be monitored	
Human activities	Administrative units	Fishing areas (leisure)		Whole Atlantic except USA	
Marine water	Horizontal velocity of the water column (currents)	Eastward current velocity in the water body	Horizontal: 1km		

Marine water	Temperature of the water column	Temperature of the water body			Historic temporal coverage limited
Marine water	Wave height and period statistics	Significant height of waves {Hs} on the water body		More time coverage needed	
Marine water	Sea level	Height above sea level		Higher density of tidal gauges	
Marine water	Fishing by-catch	Weight and number of fish and by-catch by species and year discarded		Discards data need to be collected in Atlantic Ocean	
Marine water	Chlorophyll pigment concentrations in water bodies	Concentration of chlorophyll per unit volume of the water body		Higher density of in situ measurements	
Marine water	Dissolved oxygen parameters in the water column	Concentration of Oxygen per unit mass of the water body			
Marine water	Salinity of the water column	Salinity of the water body			
Marine water	Temperature of the water column	Temperature of the water body			
Marine water	Concentration of total nitrogen [mg/l N]	Temperature of the water body			
Seabed - Riverbed	Bathymetry and Elevation	Height above LAT around shoreline	Space: high Time: Every few years	Whole EU	
Seabed - Riverbed	Habitat characterisation	Predicted broad-scale EUNIS Habitats		Except EU	
Seabed-Riverbed	Bathymetry and Elevation	Height above LAT around shoreline	100m	Approaches to harbours, new routes	Accuracy issue (see Bathymetry challenge GIS)
Seabed - Riverbed	Lithology	Spatial distribution of seabed substrate types	Horizontal scale: 1 : 100000	Whole Atlantic basin	

4.2.5 P02 needing assembly

Besides true data gaps, a lot of data suffer from a lack of assembly, i.e. data exist but are either quite scattered or not catered for or not homogeneous. This is true in the EU, let alone for other parts of the Atlantic basin. In Table 6 assembly needs are split into:

- Sheer assembly needs, i.e. P02 with no currents services needing the creation of a Thematic Assembly Centres (TAC);
- Update needs, i.e. existing TACs which do not manage to keep abreast of increasing data;
- Specific cases where international cooperation would improve the situation.

The comments below are given per matrix. In cases where data were not existing, although appropriateness was not assessed and no bar charts produced, challenges gave their expert opinion. In cases where data featured low availability, appropriateness was not informed either but then availability criteria were informed.

4.2.5.1 - Need for creation of TACs

For the MPA challenge, two requirements resorting to “Administrative units” are mentioned:

- Biodiversity critical areas (EBSA, under the aegis of the United Nations) which are missing in North East Atlantic. EBSAs are delineated based on a number of biological characteristics of threatened species (Birds, Cetaceans, Fish, Reptiles, Seals) which are reported;
- Another recommendation is for fostering the advent of EBVs (Essential Biodiversity Variable), a minimal set of biological variables to be measured to capture the main dimensions of the state and dynamics of biodiversity on its different levels of organization; one of the findings reported is that observation systems are not adapted to EBVs today and must be cross-aanlysed to reach them;
- MPA observatories to monitor essential physical, chemical and biological markers of climate change. A TAC keeping track of the implementation of these observatories in Atlantic MPAs is needed;

4.2.5.2 - Need for TAC updates

Human activities matrix

- There is a need tor strengthened activity of the EMODnet Human activity lot for a number of Administrative units, whose data generally exist but are still very scattered or need to be rescued. This is the case for:
    - Pipe-lines and cables
    - Dredge spoil dumping areas
    - Military activities areas
    - Munition dumping areas
    - Scientific activities areas
    - Aquaculture activities sites
    - Offshore installations and renewable energies sites
    - Industrial activities
    - Leisure activities
- In the field of fisheries, Fishing by-catch in numbers and discards in weight should be measured in standardized ways and be made available to the public.

- For oil spill contingency planning, Environmental Sensitivity Indexes (ESI) should be elaborated and made available in digital form for the whole Atlantic basin;

Marine water matrix

Many data for marine water characteristics are available but are assembled by only a few EU countries, so before considering enhanced data collection (as above-mentioned in 4.2.4), more efforts should be undertaken in collating and integrating data on the following:

- chlorophyll pigment concentration
- dissolved oxygen
- salinity and temperature
- chemicals such as nitrates and phosphates

This is clearly within the remit of EMODnet and CMEMS.

4.2.5.3 - International cooperation

The emphasis is placed by the challenges on the Human activities matrix perhaps because it is an area where joint efforts are not as frequent as in natural sciences.

The primary needs are about adjacent third party countries whose maritime activities have an impact on EU waters, e.g. Russia. Active collaboration with these countries should be sought.

Another request is for extending assembly efforts across the Atlantic and a couple of initiatives are mentioned (AORA, MESA) that perhaps could host such endeavours.

Table 6: P02 requiring assembly

Recommended actions				
Matrix	**P02 characteric category**	**Characteristic**	**Creation of TAC**	**Suggested TAC**
Human activities	Administrative units	Biodiversity critical areas , e.g. EBSA distribution (Biologically or Ecologically Significant Areas)	Specific TAC for MPAs	EBSA (UN)
Human activities	Administrative units	Biological markers monitoring		
Human activities	Administrative units	Chemical markers monitoring		
Human activity	Administrative units	Physical markers monitoring		
Human activities	Transport activity	Marine Traffic : Vessel Identification and Positioning, Traffic density	Vessel identification and position and derived information e.g. traffic density (AIS, VMS): open and free	EMODnet Human activities
Matrix	**P02 characteric category**	**Characteristic**	**Update of TAC**	**Suggested TAC**
Human activities	Administrative units	Biodiversity critical areas	EU coverage	EMODnet Biology or Human activities
Human activities	Administrative units	Fishing areas	Fishing ships under 12m length	EMODnet Human activities or DCF
Human activities	Administrative units	Limits of Important Bird Areas (IBA) sites		EMODnet Biology or Human activities
Human activities	Administrative units	Military activities		EMODnet Human activities
Human activities	Administrative units	Munition dumping areas		
Human activities	Administrative units	Scientific activity area		
Human activities	Fishery characterisation	Aquaculture	Data from countries other than Ireland, Spain and UK	EMODnet Human activities or DCF
Human activities	Fishing by-catch	Weight and number of fish and by-catch by species and year discarded	By-catch data in the DCF	DCF
Human activities	Industrial activity	Offshore activities		EMODnet Human activities
Human activities	Hazards to navigation	Cables and pipelines	a) Actual cable paths	

			b) Onshore power grid network connection to submarine cables c) Suspected gap in completeness: need for data rescue	
Human activities	Marine environment leisure usage	Leisure Activity (Diving, Snorkeling)		
Human activities	Pollution event	Environmental Sensitivity Index	ESI digital atlas	
Marine water	Dissolved inorganic nitrogen concentration	Concentration of nitrate+nitrite {NO3+NO2} per unit mass of the water body	Data from countries other than Spain, France and Portugal	EMODnet Physics or Chemistry
Marine water	Dissolved oxygen parameters	Concentration of Oxygen per unit mass of the water body		
Marine water	Phosphate concentration parameters	Concentration of phosphate {PO43- CAS 14265-44-2} per unit volume of the water body [unknown phase]		
Marine water	Salinity	Salinity of the water body,		
Marine water	Chlorophyll pigment concentrations	Concentration of chlorophyll per unit volume of the water body		
Marine water	Temperature	Temperature of the water body		
Seabed-Riverbed	Bathymetry and Elevation	Sea-floor height in the water body	Shoreline change	EMODnet Geology (with contrib. of Bathymetry)
Matrix	**P02 characteric category**	**Characteristic**	**International cooperation**	**Suggested framework**
Human activities	Administrative units	Waste disposal area	Extend EMODnet to whole Atlantic	AORA for North Atlantic / MESA for Africa
Human activities	Industrial activity	Aggregate and hydrocarbon extraction		
Human activities	Litter abundance and type	Waste disposal area		
Human activities	Pollution events			
Human activities	Man-made structures	Offshore Installations and Renewable energy offshore facilities		

4.2.6 P02s with limited services

In DAR2, only 104 data sets were actually downloaded and used to make the products. Those with limited availability may have hampered the making of products (referred to as “components not covered”), and were then reported in Sextant by the challenges in the component “expert opinion” and summarised in Table 7.

Limited services are split here into several categories:

- Availability issues as reported in DAR1: policy restrictions, low readiness or performance;
- Missing quality metadata;
- Particular technical issues that were not reported elsewhere.

It may happen that for some P02 these gaps in existing data services may also add up to assembly needs mentioned in section 4.2.5. These will be recapped and synthetised in the recommendations.

4.2.6.1 - Data policy

The main issue is Human activities, specifically transport and navigation data, namely high resolution VMS and ERS (logbook) which are missing from most countries due to policy issues. This is a big issue for the assessment of fisheries impact on the seabed, but also for other applications where ship movements are needed.

4.2.6.2 - Quality metadata

- Several areas in Human activities are reported to lack metadata, which makes the identification of useful data sets difficult. Generally the question is whether data sets are complete or do gaps mean absence of data. Metadata should include a description of data completeness.
- Technical metadata gaps are also reported for bathymetry, which limits the ability to work on confidence;

4.2.6.3 - Common standards

- Issues are reported in the way biological observations of threatened and declining species are made and data analysed. Homogeneous surveying protocols and analysis standards are needed to ensure comparability and smooth data assembly.
- Regarding water key constituents, marine and fresh water alike, more coordination is needed from EMODnet Physics and Chemistry TACs to provide users with co-located *in situ* observations. Surveying protocols and analysis standards are needed to ensure comparability and smooth data assembly

4.2.6.4 - Responsiveness or format readiness

There are several important issues with either responsiveness or format readiness:

- Formats are not always digital, which is a strong impediment to using them.
- Some formats are not self-descriptive or are binary, making it difficult to view or assess their appropriateness. The remedy would be to offer cloud computing services.
- The lack of standard identifiers raised by some challenges (e.g. Bathymetry) is a general issue for data traceability. In the UD catalogue some data sets even do not have a proper name, making it difficult to refer to them. Harmonised DOIs would be a solution.

- TACs or programs delivering data should have the capacity to be maintained as operational structures, which is not always the case, for example in the case of EBSA.

Table 7: P02 lacking availability

Availibility issues			
Matrix	**P02**	**P01**	**Data Policy**
Human activities	Fishing effort	Fishing_Effort_NE_Atlantic	From DCF: data not made available except by UK and Estonia, which creates underestimation of fishing effort as a proportion of total effort
		High Resolution VMS & ERS (Logbook) Fishing Intensity Data	Missing VMS and Logbook ICES data from the Faroe Islands, Russia, Iceland and Spain
			Data from VMS or Logbook to form a high resolution fishing effort analysis of the region is not available for areas from Canada to Central America
			Missing High Resolution VMS & ERS (Logbook) for West Africa down Equator
	Transport activity	Aids to Navigation, Marine vessel traffic	Make full resolution navigation data (AIS) open and free
Matrix	**P02**	**P01**	**Information on data quality (metadata)**
Human activities	Administrative units	Dredge spoil dumping areas	Needs information on data completeness
		Natura2000 spatial extent	
	Hazards to navigation	Submarine cables	Needs complete datasets descriptions with information on completeness
Seabed - Riverbed	Bathymetry and Elevation	Sea-floor depth (below mean sea level)	Needs complete datasets descriptions (survey date, instrument…) with information on accuracy and completeness
			Technical issues
Matrix	**P02**	**P01**	**Sampling and measurement protocols**
Biota - Biology	Bird counts	Threatened and/or declining species	Harmonisation of analysis protocols and their descriptions
	Cetacean abundance		
	Fish taxonomy-related abundance		
	Reptile abundance		
	Seal abundance		
Marine and fresh water	Chlorophyll pigment concentrations	Concentration of chlorophyll per unit volume of the water body	Harmonisation of measurement protocols to get consistent spatial and temporal sampling between the different variables

	Dissolved inorganic nitrogen concentration	Concentration of nitrate+nitrite {NO3+NO2	Coordination needed between EMODnet Physics and Chemistry
	Dissolved oxygen	Concentration of oxygen per unit mass of the water body	
	Phosphate concentration	Concentration of phosphate	
	Temperature	Temperature of the water body	
	Salinity	Salinity of the water body	
Matrix	**P02**	**P01**	**Responsiveness or format readiness**
Seabed - Riverbed	Bathymetry and Elevation	Sea-floor depth (below mean sea level)	Needs standard identifiers (survey, product...) from data providers
Human activities	Administrative units	Fishing areas	Digital format required with georeferenced data
	Pollution event	Environmental Sensitivity Index	Digital format required
	Administrative units	Dredge spoil dumping areas	Polygons of sites more appropriate than point locations
		Munition dumping areas	
	Industrial activity	Aggregate Extraction locations	
		Dredging sites and locations.	
		Offshore activities	
Marine and fresh water	Dissolved inorganic nitrogen concentration	Concentration of nitrate+nitrite {NO3+NO2}	Readiness: Viewing services required for data delivered in binary format Responsiveness: Data access and download from some TACs portals was challenging: a service level agreement (SLA) ensuring high responsiveness is required. Such difficulties have been reported for other P02s requiring TAC.
	Chlorophyll pigment concentrations	Concentration of chlorophyll per unit volume of the water body	
	Dissolved oxygen parameters	Concentration of Oxygen per unit mass of the water body	
	Phosphate concentration	Concentration of phosphate	
	Temperature	Temperature of the water body	
	Salinity	Salinity of the water body	
Marine water	Horizontal velocity of the water column (currents)	Northward current velocity in the water body	Extreme large files or compilations. Require cloud computing services (for viewing, downloading and processing) to face the evolution of the required coverage and resolution. Data sources concerned : Copernicus and EMODnet bathymetry
Seabed - Riverbed	Bathymetry and Elevation	Sea-floor depth (below mean sea level)	

4.3 Assessment of EMODnet, Copernicus and the DCF

Here we identify how well the datasets provided by three key European marine data platforms, namely CMEMS, the DCF and EMODnet, met the challenge product requirements. This assessment only applied to North-East Atlantic European waters and left aside EU West Indies territories.

For each data platform a table indicates the usability of one or several dataset(s) with respect to one or more challenge product requirements. In this table usability is classified in 'Low', 'Medium' or 'High'. The column 'Expert judgment' provides further information on dataset usability. Some recommendations from the challenge experts are also given for each data platform.

The need for more surveys obviously represents a large part of the issues identified by the challenges when attempting to use EMODnet TACs, but it is not within their remit. Thus recommendations on that point are being left to other sections of this report.

These recommendations are to be forwarded to the EMODnet portals by the Secretariat for potential corrective action.

4.3.1 Copernicus Marine Environment Monitoring Service

Assessment

Nine datasets from CMEMS were used by the challenges Windfarm siting, Oil leak, Climate, Coast and Eutrophication (table 8).

Table 8: Assessment of CMEMS

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 3-Windfarm siting map at Portugal / Spain Southern boundary 4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Wind	Global ocean wind observations climatology (monthly means) - reprocessed	Low	Coarse spatial resolution (25km while 10km required) and available only for 10m height
Windfarm Siting	4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Wave	Atlantic-Iberian Biscay Irish- Ocean Wave Analysis and Forecast	Low	Coverage falls outside the study zone
Windfarm Siting	4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Wave	Global Ocean Waves Analysis and Forecast updated Daily	Low	Spatial resolution (10km) and temporal coverage not appropriate

Oil Platform Leaks	1-Oil Platform Leak Bulletin (24h)	Currents Temperature (water)	Global Ocean 1/12 Physics Analysis and Forecast updated Daily	**Low**	Horizontal and vertical resolutions are resp. only 10 and 20% of needed ones, which drastically reduces the ability of the model to predict at the coast.
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 3-Windfarm siting map at Portugal / Spain Southern boundary	Wave	Atlantic-Iberian Biscay Irish- Ocean Wave Analysis and Forecast	**Medium**	Temporal coverage limited but deemed fit-for-purpose
Climate	3-Internal Energy time-series (1915-2014): Heat and Kinetic	Temperature (water)	GLOBAL OCEAN PHYSICS REANALYSIS GLORYS2V4	**Medium**	Big data issue: Attempt to calculate the mean kinetic energy of the basin. Good appropriateness but due to extreme size of data files it was only possible to download and process 1 year of the 20 year dataset
Coasts	1-Map of absolute annually averaged sea level trend (10-year period)	Sea level	AVISO ODES (ONLINE DATA EXTRACTION SERVICE)	**High**	Meets requirements except that the satellite altimetry data has coarse resolution and had to be extrapolated to assign a value to each stretch of coast, therefore losing some accuracy
Eutrophication	4-Map of Seasonal p90 chlorophyll-a concentration - North Atlantic basin (2005-2014) 6-Maps of Nitrogen and Phosphorus inputs from rivers correlation with Chl (Loire-Vilaine estuary, 2005-2014)	Chlorophyll	Global Ocean Chlorophyll (Copernicus-GlobColour) from Satellite observations Monthly, 8-days, Daily-Interpolated (Reprocessed from 1997)	**High**	Whole period covered and the area covered except over 65°N in winter. The daily products with a resolution of 4km allows a good accuracy to make an eutrophication indicator over the North Atlantic Ocean.
Eutrophication	5-Maps of Seasonal p90 chlorophyll-a concentration - Loire-Vilaine estuary (2005-2014)	Chlorophyll	North Atlantic Chlorophyll (Copernicus-GlobColour) from Satellite Observations: Daily Interpolated	**High**	Whole area and period is covered with an adequate temporal and spatial resolution for this local area

			(Reprocessed from 1997)		

Recommendations

The wind speed data used by Windfarm lacks the horizontal and vertical resolutions required for a regional scale exercise. Therefore, the production of new data series to bridge these gaps is required. It should be noted that the portal has just issued a new global data product (GLOBAL_ANALYSIS_FORECAST_WAV_001_027) with improved horizontal resolution (10km), but still limited to the sea surface.

It is also recommended that wave (Windfarm) and currents (Oil leaks) data be improved with respect to horizontal resolution in the Atlantic.

For large files (e.g. GLOBAL OCEAN PHYSICS REANALYSIS GLORYS2V4 used by Climate) the feasibility of a cloud facility both for working storage and data processing might be considered. This challenge also mentions that Copernicus Global 027 dataset was removed in 2017 and superseded by a different data set Global 047. The initial method was developed by the challenge to use this dataset, and for the new one a new method would have been required. Where datasets are superseded secondary users need clear notification or legacy access.

The following additional points can be made:

- Wave/current-induced water motion, salinity, temperature, oxygen, and nutrients are acknowledged as the main factors driving the spatial distribution of marine species. As species distribution modelling (SDM) currently is among the most promising activities in marine science, improvement in model horizontal resolution down to 200-300m for all European basins would be invaluable. Further, SDM initiatives may need to provide confidence in their products at any location. Therefore, provision of spatially-explicit layers describing the uncertainty of the CMEMS product values is also recommended.
- We also recommend that for variables such as salinity, temperature, oxygen and wave/current energy, sea-bottom values be calculated and made available, as is done for a few products (e.g. product BLKSEA_ANALYSIS_FORECAST_BIO_007_009, which provides a layer with dissolved oxygen values at the seabed).
- CMEMS products mostly make available their variables in the form of monthly or daily averaged values. One may need other statistics over time such as percentiles, and currently the only way to obtain such layers would be to compute them in-house by downloading the full data series, which may be critical to some users. Therefore, a web service that would enable users to compute their own data layer online from the archives, with a time window and various statistics of their choice, would be invaluable.

4.3.2 Data Collection Framework

Assessment

The DCF (Data Collection Framework) is an EU framework for the collection and management of fisheries data. Under this framework Member States collect, manage and make available a wide range of fisheries data needed for scientific advice. Part of this data are uploaded in databases managed by the JRC which in turn makes the data available to the Scientific, Technical and Economic Committee for Fisheries (STECF) working groups for analysis and provision of scientific reports. The data are further disseminated in aggregated

form to a target audience of experts for further use in scientific analyses and policy. Only a very small part of the data is made available to the public trough the JRC website, as is demonstrated by the assessment below (table 9).

Table 9: Assessment of the DCF DCF publicly available data on discards

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Fisheries Management	2-Discards by species and year, in mass and number	fisheries	Discards in mass	**Low**	Data available to the general public in a well documented format including metadata, but only for the very few stocks subject to special effort regimes (e.g under recovery plans)
Fisheries Management	2-Discards by species and year, in mass and number	fisheries	Discards in number	**Not available**	Discards in number is only an input in the scientific process and is not usually made available for the general public

Recommendations

Unlike discards in mass, which is commonly used to support fisheries management, discards in numbers, usually associated with age or size data, are only useful for scientific purposes. No recommendations can be made on making these estimated data (when they exist, which is seldom the case) available for management purposes.

In the EU context, datasets on discards in mass were found to have low usability, and bycatch data are not binding under the DCF. Discard and bycatch levels have interest for the society in general and as such they should be made available to the interested parties in the future. However, EU scientists have many concerns on the quality of the discards estimates[2]. To date, it is still doubtful that discard estimates for all stocks being sampled will be disseminated due to concerns on the quality of the estimates.

Therefore, no particular recommendations can be made on improvements on the availability of these data for the public. Recommendations are to address the reliability of fisheries bycatch and discard data, which needs to be improved by way of adequate sampling protocols based on further research, and their implementation in the DCF. Notice that these recommendations are made in the EU context only since no other region in the geographical scope of this study make discards data available to the public.

4.3.3 EMODnet Bathymetry

Assessment

The DTM provided by EMODnet Bathymetry was used by Windfarm and Bathymetry. Bathymetry also attempted to use original soundings (table 10).

[2] STECF (2017). Fisheries Dependent Information Classic (STECF-17-09). JRC science for policy report.

Table 10: Assessment of EMODnet Bathymetry

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Bathymetry	1-Sample areas of digital bathymetry covering representative areas of the North Atlantic 2-Digital bathymetry uncertainty covering the North Atlantic and a representative sample area	Bathymetry	EMODNET Bathymetry	Medium	Big data issue: Individual surveys could be downloaded from SeaDataNet, but this was not feasible as such a large area was being considered. Supporting files could not be opened/analysed by standard GIS software: 4.5GB CSV file too big for 4GB RAM and softwares; EMO file formats too big/complex ; SD files not feasible given the software available ; GeoTiffs proved unusable as they included only RGB values
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 3-Windfarm siting map at Portugal / Spain Southern boundary 4-Windfarm siting map off S. Miguel Island	Bathymetry	250m Digital terrain model	High	This dataset is performing far beyond the needs.

Recommendations

As reported by Bathymetry, the EMODnet Bathymetry product exhibits a number of usability issues compared with the other sources:

- An important part of this challenge tasks was to understand the source information and metadata for each data layer. Whilst it would have been preferable to download the 'source references' layer from this portal also, this was not possible, so source data could not be used adequately due to large file and lack of computer memory. Parts of this information were transcribed by clicking on individual points, but this was not practical for the whole area of interest. Lack of visibility of information such as vertical datums and horizontal/vertical accuracy is reported.
- Issues with the various formats in which the DTM is made available are also reported (see table 10). Data files were too big (csv format) or not ready-for-use (EMO, SD)

formats) or unusable (Geotiffs) while binary format such as NetCDF were not usable with the user software.
It is recommended that these points be addressed in future iterations if this has not happened already. In particular, the USGS product is highlighted as a positive example. Whilst EMODnet and USGS might not be a like-for-like comparison, it is recommended that the USGS source be reviewed with a view to replicating the positive experience in accessing and using that data source where appropriate.

4.3.4 EMODnet Biology

Assessment

Only the MPA challenge used data from EMODnet Biology (table 11).

Table 11: Assessment of EMODnet Biology

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Marine Protected Areas	2-Quantitative analysis of MPA coherency	Birds, mammals, reptiles	- Abundance of Threatened and/or Declining Cetacean Species - Abundance of Threatened and/or Declining Seal Species - Bird taxonomy-related abundanc - Fish abundance - Pagophila eburnean - Reptile	Low	Data was found for all the 85 indicator species considered in the IUCN categories Unfortunately there is a lack of observation density across Atlantic to perform the task in an appropriate way

Recommendations

In the Atlantic checkpoint, EMODnet Biology data was not much used because of a lack of spatial density of observations hampering the production of abundance maps.

As an example, Climate had to identify the 3 most abundant phytoplankton species in the North Atlantic and then track their abundance over the last century. The conclusion is that some datasets are available, but these have too limited spatial and temporal coverage. Even at genus rather than species level there is too little data to make a product reflecting the basin situation.

The MPA Challenge conclusions are in line with this: for birds, mammals and reptiles some datasets are available for the 85 required species but due to poor observation density are not appropriate for studies at basin scale.

Therefore, since shortcomings with biology data are not due to EMODnet Biology platform but to spatial gaps in the available data, no recommendations can be made on improvements of the platform.

Additionally, the Alien species challenge points to the need for a central regularly updated alien species information system at Atlantic scale and adjacent basins. EMODnet Biology would be an appropriate European platform in this respect.

4.3.5 EMODnet Chemistry

Assessment

Table 12: Assessment of EMODnet Chemistry

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Marine Protected Areas	5-Chemical parameters network monitoring systems within MPAs	Chemicals	Alkalinity, acidity and pH of the water column pH	Low	Most datasets did not cover the 50-year temporal coverage required for climate change effect assessments
Eutrophication	1-Map of seasonal averages and changes of dissolved oxygen indicator for eutrophication (2005-2014) 2-Map of seasonal averages and changes of Chlorophyll indicator for eutrophication (2005-2014) 3-Map of seasonal averages and changes of Dissolved inorganic nitrogen indicator for eutrophication (2005-2014)	Chemicals	Chlorophyll-a concentrations Dissolved oxygen Nitrite Nitrate	Low	Requirements for spatial and temporal coverage were not met: 1) most of the in-situ data gaps reflected past monitoring design, 2) very few locations where a seasonal temporal resolution was available, 3) most of the North Atlantic basin with no data except for certain areas of the European coast, 4) spatial coverage not appropriate where eutrophication is likely to occur (i.e. coastal and estuary areas)
Eutrophication	7-Map of potential for eutrophication due to commercial activities off the west coast of Ireland	Chemicals	Chlorophyll-a concentrations Dissolved oxygen Nitrite Nitrate	Low	The Celtic sea, which encompass the study area, is not appropriately covered

Recommendations

The challenge Eutrophication points that EMODnet Chemistry provides good overall accessibility services, visibility and well-documented metadata. However the challenge mentions that it is necessary to enable users to select time range and resolution options

more easily, particularly for seasonal periods. Users should also be able to download more than 500 results.

Another issue is the lack of standardization in measurements, which has a bearing on assembly processing tools. Feedback is needed from EMODnet Chemistry to the entities collecting data.

Improvements are also recommended in terms of consistency between data providers and portals such as EMODNet, ICES, and OSPAR (ODIMS) to inform users of data sets availability and clarify if there is any overlap or duplication with other portals. If the same data sets are available from these portals (and the national portals) there is a need to clarify and guide users on overlaps and differences. Ideally users should obtain the data sets from one unique portal with certainty they are not missing any other data entries.

Regarding appropriateness, although the portal provides the largest and most comprehensive datasets in terms of spatial and temporal coverage for the North Atlantic basin, important spatial and temporal gaps against requirements were mentioned for products for Eutrophication and climate change impact on MPAs.

4.3.6 EMODnet Physics

Assessment

Table 13: Assessment of EMODnet Physics

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Marine Protected Areas	5-Physical parameters network monitoring systems within MPAs	Temperature (air) Suspended particulate material Salinity	- Atmospheric temperature - Light Attenuation/ Absorption /Fluorescence/ Back Scattering - Salinity - Water Temperature - Waves - Winds	Low	Most datasets did not cover the 50-year temporal coverage required for climate change effect assessments
Coasts	4-Map of relative annually averaged sea level trend (10-year period)	Sea level	Near real time sea level data	Medium	Rather complete but not possible to have a tide gauge for each stretch of coast in the Atlantic

Recommendations

With respect to the assessment of climate change impact on MPAs, EMODnet physics datasets suffers from the same issues of temporal coverage than those of EMODnet Chemistry. Coast points out some issues but these are not due to EMODnet Physics.

River Inputs mentioned that efforts have been allocated to Emodnet Physics to gather all flow rate information at European level since 2017. That initiative should be maintained and could be consolidated by producing more intercalibration between Member States methods to estimate flow rate from water levels.

4.3.7 EMODnet Geology

Assessment

EMODnet Geology seabed substrate layers were used exclusively by Windfarm (table 14).

Table 14: Assessment of EMODnet Geology

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 3-Windfarm siting map at Portugal / Spain Southern boundary 4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Lithology	1:1,000,000 Seabed substrate map of the European marine areas	Low	Scale not appropriate (1:250,000 suitable) Spatial coverage gaps : Product 1 : 65% not covered Product 2 : 30% not covered, especially in the deep-sea Product 3 : 30% not covered, especially in the deep-sea Product 4 : >99% not covered
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Lithology	1:250,000 Seabed substrate map of the European marine areas	Low	Spatial coverage gaps : Product 1 : 90% not covered Product 2 : 80% not covered, especially in the deep-sea Product 3 : 30% not covered, especially in the deep-sea Product 4 : almost not covered Very poor spatial coverage (90% not covered) for the study area
Windfarm Siting	3-Windfarm siting map at Portugal / Spain Southern boundary	Lithology	1:250,000 Seabed substrate map of the European marine areas	Medium	Spatial coverage acceptable although 30% not covered, but essentially in the deep sea Suitable scale

Recommendations

As for example the EU seabed substrate at scale 1 : 250 000 map only covers about 20% of the EEZ, recommendations are for more spatial resolution and coverage, so in an implicit way new surveys are called for.

4.3.8 EMODnet Human activities

Assessment

Table 15: Assessment of EMODnet Human Activities

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 3-Windfarm siting map at Portugal / Spain Southern boundary 4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Human activity	Cables	Low	We did not use the dataset due to poor completeness in number of features
Windfarm Siting	3-Windfarm siting map at Portugal / Spain Southern boundary 4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Human activity	3 datasets : Aggregate Extraction Offshore Installations Dumped munitions	Low	Portugal did not deliver data
Windfarm Siting	3-Windfarm siting map at Portugal / Spain Southern boundary	Human activity	Dredging	Low	Portugal did not deliver data
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet	Administr. units	Dumped munitions	Low	Ireland and UK did not deliver data
Marine Protected Areas	1-MPA Atlantic network classified in IUCN 3-Distribution of No-Take zones	Human activity	Status of Bathing Waters	Low	Proxy for leisure activities (diving, surfing, sailing, tourist beaches). Low spatial completeness.
Oil Platform Leaks	1-Oil Platform Leak Bulletin (72h)	Human activity	Status of Bathing Waters	Low	Proxy for beach spatial distribution. Highly spatially incomplete
Windfarm Siting	1-Windfarm siting map where waters of FR, IR	Human activity	Aggregate Extraction	Medium	All required countries delivered data, however

	and UK meet 2-Windfarm siting map where waters of France and Spain meet				medium score because polygon features would be more appropriate than points
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 3-Windfarm siting map at Portugal / Spain Southern boundary 4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Administr. units	Dredge spoil dumping	Medium	All required countries deliver data but medium score because polygon features would be more appropriate than points
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet 4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Human activity	Dredging	Medium	
Windfarm Siting	2-Windfarm siting map where waters of France and Spain meet	Administr. units	Dumped munitions	Medium	
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France and Spain meet	Human activity	Offshore Installations	Medium	No data for France but it is assumed it is no data gap (probably no offshore installation in France) Medium score because polygon features would be more appropriate than points
Marine Protected Areas	1-MPA Atlantic network classified in IUCN 3-Distribution of No-Take zones	Human activity	- Aggregate Extraction - Dredging - Dumped Munitions - Offshore Installations	Medium	Portugal do not deliver data
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet 2-Windfarm siting map where waters of France	Human activity	Main ports	High	All required European countries deliver data

	and Spain meet 3-Windfarm siting map at Portugal / Spain Southern boundary 4-Windfarm siting map off S. Miguel Island (Azores archipelago)				
Marine Protected Areas	1-MPA Atlantic network classified in IUCN 3-Distribution of No-Take zones	Human activity	- Boreholes - Finfish farming sites - Hydrocarbon Extraction (Active Licenses) - Ocean Energy Facilities - Shellfish Production - Wind Farms (Points)	High	All required European countries delivered data
Oil Platform Leaks	1-Oil Platform Leak Bulletin (72h)	Fisheries	- Finfish production - Hydrocarbon Extraction	High	

Recommendations

Key characteristics for several challenges could not be found:

- marine traffic (VMS, AIS)
- fishing activity
- leisure activities (diving, surfing, sailing, tourist beaches)
- recreational fisheries
- land power grid network
- marine bird corridors
- marine Important Bird and Biodiversity Areas (IBAs)

We recommend these data layers be collated (or developed, if need be) by EMODnet Human activities, with the highest priority to the first three ones.

Some datasets contain only point features. This type of geographic representation is not appropriate for windfarm siting or oil leak impact assessment purposes. Where the data exists in polygon form, they should be made available.

Most datasets lack appropriate metadata information on contributing countries. An example is offshore installations, for which there is no occurrence in French waters, but from the accompanying metadata it is impossible to know if it is due absence of data or actual absence of offshore installations. Any information that would help the user evaluate the spatial completeness of a dataset should be mentioned in the metadata, including contributing countries.

Windfarm also revealed that many datasets do not have records in Portugal waters.

An Environmental Sensitivity Index (ESI) map summarising sensitivity data into a comprehensive index has long proved to be essential in oil spill contingency planning. While available for the USA coasts (albeit in paper form), Europe save France still lacks a digital atlas and a dedicated Thematic Assembly Centre (TAC) would be appropriate under Human activities in close collaboration with the EMSA.

4.3.9 EMODnet Seabed habitat

Assessment

The data made available by EMODnet Seabed Habitat (table 16) were used by challenges MPA and Windfarm, Oil platform leak and Fisheries impact.

Table 16: Assessment of EMODnet Seabed habitats

Challenge	Product	Caract.	Dataset	Usability	Expert judgment
Marine Protected Areas	2-Quantitative analysis of MPA coherency	Habitat	Predicted broad-scale EUNIS habitats - Atlantic area (updated 9 December 2013)	Low	This map covers a quite extensive area in Europe but does not have the appropriate thematic accuracy : only 3 (Rock, Cymododea meadows, Posidonia meadows) of the 26 required habitats (e.g. Carbonate mounds, Cold seeps, Cold water coral, kelp forest) is addressed.
Windfarm Siting	4-Windfarm siting map off S. Miguel Island (Azores archipelago)	Habitat	Medium scale habitat maps	Medium	This dataset is a thematic map for habitat mapping purposes. Its use with confidence for other applications requiring only substrate classes was not found in the metadata.
Windfarm Siting	1-Windfarm siting map where waters of FR, IR and UK meet	Habitat	EUNIS habitat maps from surveys - Medium scale	Medium	Contains a good EUNIS physical classification of benthic habitats with abundant information on substrate types. For the Azores it was used, in a great extent, as a substitute data to the EMODnet geology substrate datasets that has very low spatial coverage in that area.
Oil Platform Leaks	1-Oil Platform Leak Bulletin (24h)	Habitat	Predicted broad-scale EUNIS habitats - Atlantic area	Medium	

Fisheries Impact	2-Damage to seafloor to both living and non-living components	Habitat	Broad-scale habitat map (EUSeaMap) including classified habitat descriptors and confidence (updated 15th June 2017)	High	The Habitat map is spatially complete in terms of area of 1000m in depth, and the scale is adequate

Recommendations

For the quantitative analysis of MPA coherency, the challenge required data on the spatial distribution of OSPAR habitats, the ICES List of VME habitat types and other habitats such as carbon sink habitats (seagrass meadows, mangrove swamp and saltmarshes), coral reefs and rock.

Data on the spatial distribution of these habitats are either missing or scattered. It is within the remit of EMODnet Seabed habitats, with the help of EMODnet Biology, to collate exiting polygon or point data across Europe and make them available in compiled layers.

4.3.10 Conclusions

CMEMS and EMODnet have collated, indexed, stantardised and made available millions of data records that used to be scattered among isolated data portals or even not available at all. EMODnet has also developed from these observation data spatial products covering extensive areas (e.g. the 250m depth DTM, the broad-scale seabed habitat map or the seabed substrate map). Thanks to these two initiatives the situation has dramatically improved in a relatively short time period. However there are still are a few shortcomings regarding data availability or portal facilities that were reported here. We have no doubt that these will rapidly be fixed. There are also some data gaps in spatial and temporal coverage/resolution.

As mentioned above, these gaps are due to several reasons:

- Some datasets may have not been collated yet;
- Some characteristics may not have been addressed yet by the EMODnet lots (e.g. marine traffic, vulnerable habitats);
- Data may not exist due to the lack of suitable resolution or spatial coverage.

The former two reasons are being addressed by the ongoing phase 3 of EMODnet and initiatives such as the Data Ingestion Project.

For the latter, beyond the need for new surveys providing *in situ* data for assimilation, some suggestions have been made to Copernicus for improving the spatial resolution of models. Such improved data sets would be a breakthrough in marine science by enabling the development of a new generation of high-level spatially-explicit products based on Species Distribution Modelling (SDM) or artificial intelligence techniques, with considerably better accuracy and more extensive coverage than today.

The experiences faced by the Climate and Bathymetry challenges in terms of extreme file size, similar to data types such as acoustic imagery reflect a general trend in environmental data. There is a need to offer to end-users cloud computing services for data viewing and processing to tackle the increase in resolution, spatial or temporal extent and metadata for identification and quality checks. It is no longer possible to ask end-users to download data

through the internet or other devices without knowing if the requested datasets meet their needs. Cloud computing services will have to be proposed to end-users on request.

5 Synthesis and recommendations

5.1- Recommendations on challenges

It is very difficult to give recommendations about the relevance of the challenges because everyone would have different priorities and opinions. The request for some of the challenges may have seemed unclear to the Project. For example, the tender asked for Alien Species a citation table that is of little use to determine the appropriate quality measures.

All the challenges are recognized essential although some of them are interdependent:

- Eutrophication cannot be addressed without knowing more about river run-off. Coasts needs to handle sediment transport in the coastal zone, a large part of which is provided by river discharge. Therefore one would tend to give the River Inputs some increased priority;
- Shoreline change in Coasts is closely linked to climate change, so if we knew more about climate change drivers, modelling shoreline change would probably be easier;
- Fisheries management and Fisheries impact are obviously closely linked, although in the way the challenges were specified, very different characteristics were used for both.
- Finally Oil leak (which also concerns oil spills from ships) is linked to Bathymetry because, as new routes open up, good charting is *a sine qua non* condition to avoid ship groundings.

So we did not give recommendations about the relative importance of the challenges themselves other than those for data characteristics below.

5.2- Most urgent recommendations on characteristics

This section, by building on the data analysis made in section 3 tries to provide the most essential recommendations to orientate future action.

It has been recognized that for the providers community (and for funding institutions), trying in the first place to thrash out either assembly or availability issues would be more effective in terms of financial resources than collecting new data, so this is the way we are going to formulate recommendations. As has been shown above a lot of progress still lies in assembling data highly scattered by nature, as is the case for e.g. human activities or biology data in general. This may mean going to Member States and giving them incentives or obligations (for example through directives such as the MSFD or MSPD) to collate and make available their data to EMODnet and to develop collaborations with countries having significant activities in EU waters of interest such as fisheries.

Availability issues may also be at stake, as has been described in DAR1 and specified in the present data analysis, mostly for data suffering from policy, readiness or performance restrictions. These latter issues should be possible to solve at a reduced cost over effort ratio, so this is why we recommend to primarily address them. Only in a second set of recommendations do we summarize areas where new surveys would really be necessary.

We also ranked the 43 characteristics according to how often they had been used by challenges. In table 17, P02s are classified in decreasing order of contribution to multiple challenges. The most pressing recommendations will primarily concern categories of characteristics that contributed to many challenges.

Finally, the dependence between the challenges explained above but also their environmental or economic importance also came into play. For example, given the utmost relevance of wind data for economic reasons linked to Blue Growth, Wind speed was given some precedence over other oceanographic characteristics.

The final order of priority among P02 was established as follows:

- *Vessel traffic – Fishing effort*
- *Administrative units – Cables - Dredging - Aggregate extraction*
- *River flow - Sediment load - Dissolved oxygen - Nitrate - Phosphate*
- *Wind speed - Currents - Temperature - Salinity*
- *Habitat extent*
- *Bathymetry*

5.2.1 Vessel traffic – Fishing effort

This category is the one missing most as it is currently being used by five challenges. Additionally it could be of relevance for Alien Species, which has to deal with with vectors of dissemination.

VMS or AIS data as well as and ERS (logbook), including smaller fishing boats data (VMS are not yet compulsory for vessels under 12 metres) are missing from most countries due to policy issues but there are also technical issues about the way the products are made. In this particular case users may want to get access to raw data and compute their products in-house. Promising attempts to use marine traffic to determine currents should also be mentioned. Given the variety of applications of these data, VMS/AIS data providers should be careful to fetch as many specifications from potential users as possible.

This is essentially an availability/assembly issue which requires no data collection. A breakthrough for these data are expected in the near future as an agreement has been passed between EMODnet and EMSA and derived products technical specifications are underway.

5.2.2 Administrative units – Industrial activities - Hazards (Cables and Pipe-lines) - Pollution

Administrative units (“areas where authorities have or exercise specific rights or obligation”) of all sorts are of primary importance for two challenges but they will be more and more required in the frame of Marine Spatial Planning. There availability would rapidly gain from assembly efforts.

- Many types of administrative units are still lacking assembly, among which for example EBSAs, fishing, waste disposal, ammunition dumping, military activity or scientific activity areas) but also beyond these, other types of human activities such as Aquaculture, Leisure Activities, Industrial activities (offshore) and installations. Within Europe, the issue is mostly their lacking assembly, which is a huge task, in spite of EMODnet Human Activities being very active in this field. Other issues mentioned are the lack of metadata thoroughness and the fact that activities being conducted within certain sizeable areas are often represented by point objects instead of polygons. However in this case the most sensible way to proceed would be to provide point data and allow users to aggregate them in the most appropriate way for their issue.
- The paper form of some datasets also limits their use, as is the case with Environmental Sensitivity Index maps, which need to be digitally assembled in a TAC;

- Finally, what is true for Europe is even more true for other regions of the Atlantic Ocean, in Europe but more acutely for the rest of the Atlantic Ocean, where international cooperation should be fostered.

5.2.3 Wind speed - Currents - Temperature- Salinity

These four categories have been grouped together because they are frequently associated in making products. A distinction should be made on one hand between *in situ* data and modeled data, and on the other hand between the high seas and the coast, where needs may be dramatically different.

Wind, for which no data above 10m height are available while measures up to 150m are required with a 10m vertical sampling interval, is at the top of the list because of its economic importance in windfarm suitability studies.

For applications needing in situ data, there is clearly a twofold issue of surveys and assembly. A higher density of multi-parameter monitoring stations is necessary alongside a stronger capacity to route the data towards assembly centres.

For applications using models, in the high seas multi-kilometric resolution is sufficient while in the coastal zone, an improvement of an order of magnitude is necessary. The best illustration is probably Oil leak, a user of these four categories of characteristics. This challenge would not have performed well should the spill have occurred near the coast. Additionally fine resolution models along with better data on nutrients from rivers would be paramount for eutrophication, a largely coastal problem strongly heterogeneous in both temporal and spatial scales.

Note: As had been mentioned in the literature survey, High Frequency Radar (HFR) have become a key tool for operational oceanography for monitoring the coastal surface currents, waves and winds. An appropriate data description complying with accepted standards is crucial for ensuring discovery and access. EMODnet Physics and the other major European integrators and infrastructures (CMEMS, SeaDataNet) are supporting and promoting the EuroGOOS HFR Task Team activities towards this integration..

5.2.4 River flow – Concentration in particulate material (Sediment load) - Dissolved oxygen - Nitrate - Phosphate

The knowledge of river inputs to the sea is critical for modeling the fate of coastal waters, especially with regard to Eutrophication and shoreline change (Coasts). Freshwater inputs to the coastal zone are still largely unknown, especially for smaller tributaries. There is a need to implement river monitoring stations to record frequently and in a synchronous way a number of physical and biological parameters.

In coastal marine waters, the same recommendation applies because we have far too few recording stations of *in situ* physical, chemical but also biological parameters (e.g. chlorophyll content) which are requested by Eutrophication.

Not only is there a need for more sampling density but also an issue about the harmonization of collection protocols and calibration/processing standards.

Table 17: P02s in decreasing order of multiple contributions to the challenges

	Used by Challenge											
P02 name (Comments)	**CH1**	**CH2**	**CH3**	**CH4**	**CH5**	**CH6**	**CH7**	**CH8**	**CH9**	**CH10**	**CH11**	
Transport activity (Vessel traffic)	1	1	1				1			1		5
Temperature (including Fresh Waters)		1	1	1				1	1			5
Bathymetry	1	1	1		1					1		5
Habitat extent	1	1	1				1					4
Industrial activity (Offshore activities and installations)	1	1	1							1		4
Wind speed	1	1	1									3
Transport activity (Ports)	1		1							1		3
Habitat characterisation		1	1				1					3
Currents - Horizontal velocity		1	1	1								3
Salinity		1	1					1				3
PHOS – Phosphate		1						1	1			3
Administrative units (MPA, Biodiversity critical areas, Fishing areas, Waste disposal, Ammunition dumping areas, Military activity, Scientific activity areas)	1	1										2
Fishery characterisation (aquaculture sites)		1	1									2
Hazard to navigation (submarine cables and pipelines)	1	1										2
Industrial activity (Dredging sites, aggregate extraction, Hydrocarbon extraction)	1	1										2
Leisure/recreational activities		1	1									2
Pollution event (including sensitivity inde1)		1	1									2
Terrestrial mapping (coastline)	1		1									2
Fish (abundance, behavior, reproduction, threatened & declining Species, Eel and Salmon markers)		1							1			2

Phytoplankton generic abundance in water bodies		1		1								2
Chlorophyll pigment concentration in water bodies		1						1				2
NTOT - Nitrate		1						1				2
Dissolved oxygen		1						1				2
RVDS - River flow		1							1			2
Concentration of suspended particulate material		1							1			2
Sea level		1			1							2
Wave height and period statistics	1	1										2
Air temperature		1										1
Snow and ice (extent)				1								1
Fishing statistics (landings)						1						1
Fishing by-catch						1						1
Fishing discards						1						1
Fishing effort							1					1
Bird (abundance, behavior, reproduction, threatened/declining Species)		1										1
Mammals (Cetaceans, Seals : abundance, behavior, reproduction, threatened & declining Species)		1										1
Reptile (abundance, behavior, reproduction, threatened/declining Species)		1										1
Primary production		1										1
Benthic primary production		1										1
Invasive species monitoring parameters											1	1
Alkalinity, acidity and pH of the water column		1										1
Heat fluxes between the water column and the atmosphere				1								1
Lithology	1											1
Coastal geomorphology		1										1

5.2.5 Bathymetry

Bathymetry is used by many challenges, mostly for regional studies. Although being EMODnet's flagship in providing a unified high-resolution DTM fit for use in many applications, bathymetry still suffers from two important drawbacks:

- Accuracy: in spite of efforts made by the Bathymetry lot, bathymetry is still lacking accuracy attached to each cell of the DTM. Further to this, in order to address mapping new maritime routes or modified approaches to harbours, complete metadata of the primary datasets need to be made available to users;
- Big volumes: when giving access to raw bathymetry, users are faced with extremely large files and big data viewing and processing issues that need to be addressed in the cloud;
- New data: Bathymetry highlighted areas needing new surveys for navigation purposes;
- to address shoreline erosion, regularly updated high resolution bathymetry is necessary. The Surveying strategy is within the remit of EMODnet Bathymetry, while the thematic area itself resorts to the Geology lot, so active collaboration between the two of them is vital for this topic.

5.2.6 Habitat extent

Habitats are being used by Fisheries Impact and Oil Leak

- EUSeaMap, the EU broad-scale habitat map produced under the aegis of EMODnet Seabed habitats is a product that meets the requirements where it is available, i.e. on a large fraction of European waters, however there are still a few uncovered places where e.g. the impact of the fisheries on the seabed habitats cannot be assessed.
- Priority habitats (not described by the above broad-scale map) are still collected on an opportunity basis and would need to be more comprehensively mapped. In European waters, the contribution of OSPAR Contracting Parties in providing updates of priority habitats and species is irregular and patchy, resulting in many geographic gaps.

5.3 - How the Checkpoint have addressed the tender issues

A set of products with their confidence

The Checkpoint products were scored from inadequate to excellent by the challenges. This reflected their overall expert's judgment based, for each of the 8 quality measures applied to the various components of the products, on the discrepancy between quantitative specified values and actual achieved ones. So this score indicates whether the product is fit for use, which is what the tender asked for.

A list of the data sources used and data providers

Data sources are listed in Annex 4, along with the challenge name, the category of characteristic (P02), the characteristic (P01), the description of the data set and its provider. Should more information be desired about data sets (UDs), users can refer to the spreadsheets in Annex 5 that give, in a P02 ordered list, all quality measures of the UDs and in cases the product was not made, the ones in Annex 6 and 7 for products not covered where some UDs, although not used, were assessed in terms of availability only.

Usefulness of each data source in terms of identification, delivery and usability

A great number of data sources had primarily been identified in DAR1 but their availability assessment was probably biased because they had not yet been effectively used in products. This was also a consequence of product description in the tender without a defined purpose, e.g. the total energy in the ocean in the Climate challenge. In this report all data sources effectively used were quantitatively assessed for their 8 quality measures and the results of each of these can be found in Annexes 5, 6 and 7.

Identification of gaps in data sets

This is covered in sections 4.2.4 where surveys needs (resolution, coverage and number of items) are detailed for P02s suffering from gaps. The reporting of resolution gaps used numbers as percentage of discrepancy between specified and achieved. The reporting of either geographic gaps or time gaps also used numbers (resp. km² or days) and resulted in a percentage, however it would have been much more efficient to produce either maps (or calendars) of completeness showing for a spatial/temporal view of gaps. It is noteworthy that ISO 19157 makes provision for such maps so it could be an issue to address in a future checkpoint update.

Data collation, assembly and synergy between providers

There was a whole section about improving data assembly (4.2.5) where recommendations were given about the need for more links between various TACs, even within EMODnet itself, when dealing with joint needs of physical/chemical data, biological/habitat data, bathymetry/elevation and erosion issues (Bathymetry and Geology lots) for which consistent spatial sampling and measurement protocols should be applied to get harmonized and colocated data, but also between institutions such as EMODnet, OSPAR, the EMSA or ICES that would benefit from working in a closer relationship.

Is data availability worsening or improving?

During the last twenty years people were focused primarily on discovering data. The INSPIRE Directive had made discovery tools an obligation for data providers, with the support of ISO 19115 for metadata and ISO 19139 for their implementation. With the need to have a Global Operationnal Observing System for sustainable Blue Growth and the Open data movement, the trend is towards data qualification both to allow end-users to determine by themselves the usabilility of the data they need and providers to protect themselves against any misuse of their data and liability costs. This trend is reinforced by the increase in size of the available data making mandatory the usability assessment of data before downloading, especially when it is at cost. Completeness is the first concern of users which such queries as: “In this place what can I find?” and if something is found “Is data resolution good enough?” Only when these two queries have been answered do users ask about data accuracy or temporal validity. All these quality items (developed in this Checkpoint) are being implemented by spatial data providers such as NODCs under ISO 19157 in line with the Inspire Directive. The process is still at a very early stage. As a conclusion the answer to this tender issue definitely is that the data landscape is steadily improving.

6 Conclusion

The terms of reference of the contract were to provide an adequacy report that:

- *"Look at the needs of user"", which is the challenge standpoint;*
- *"Looks at data on a parameter basis" which is the characteristic standpoint;*
- *"Provides a view of the monitoring efforts in the basin".*

The first item was primarily reported in the challenge score table (Table 3) which give an overview of the success of the products composing the challenges. Some of the products performed inadequately, others were limited, another large number were only good, which in fact means medium. Upon identifying the failing products the reader was directed to the comprehensive components table (Annex 1) to be able to identify the categories of characteristics (P02) that contributed to these low scores.

The characteristics (P02) were then analysed to answer the second and major term of reference. P02 may be short of the requirements for several reasons, either gaps (object of the appropriateness assessment and qualified by several quality measurements) or availability issues. Most of the latter were reported in detail in DAR 1 and the reader in search of specific details can refer to it. Another issue - assembly needs emerging from the scattered character of data - was the most tricky to thrash out because it did not really belong to either "availability" or "appropriateness" and was not so steadily reported. In fact for challenges working across multiple characteristics this was one of the biggest problems. It may be more an artefact of the way data are collected rather than how they are made available.

A view of the monitoring efforts is given by the number and types of UDs which are listed in Annex 4. By scrutinising the spreadsheets in Annex 5 to 7 (listed in P02 alphabetical order), or looking at bar charts in Annex 3 the reader can see how much coverage is fulfilled for each P02 in the Atlantic basin or in situations where full coverage is achieved, what kind of resolution is available. However the dispersion of data, which occurs for many characteristics, is not easy to render but the section on "assembly needs" to a large extent gives a statistical account of it by way of the coverage indicator.

The main recommendations were focused on a series of P02 that were recognised as most useful for the Checkpoint's challenges. These recommendations should contribute to a better definition of EOVs, especially for the human activities matrix where EOVs are still weak.

In the follow-on report delivered in Feb. 2018 the way this work could be taken forward has been described. A Thematic Checkpoint Service could be set up within EMOdnet for regular quantitative adequacy assessment and progress monitoring. The value of the present work is with its high Technical Readiness Level (TRL), an indicator estimating the technology maturity of acquisition & processing system. Based on a scale from 1 to 9, TRL enables consistent and uniform discussions of technical maturity across different types of technology. The Atlantic Checkpoint TRL is estimated at 8 because its technology is implemented, the full-scale prototype is built and integrated into an intended operating system.

The data and products catalogues are currently maintained by Ifremer, the French NODC. They provide a reference for future assessments and as an evidence have just been used by the H2020 Atlantos Project for its own purposes. The online display tools enabling users to

query the catalogues for products and datasets alike are still under development but they should be ready by the contractual end of the project.

Future prospects include a highly expected functionality enabling users to get a spatial representation of the quality indicators, a functionality that would imply additional resources not initially planned in the present Checkpoint.