EMODnet Black Sea Checkpoint
Data Adequacy Report

EMODnet
European Marine
Observation and
Data Network

Black Sea Checkpoint

Black Sea
First Adequacy Report
REVISED VERSION

Total number of pages: 75

Workpackage:	15	15.2 Final revised version of the First Data Adequacy Report	
Author(s):	Nadia Pinardi		CMCC
	Vladyslav Lyubartsev		CMCC
	Giuseppe Manzella		CMCC
	Atanas Palasov		IO-BAS
	Violeta Slabakova		IO-BAS
	Luminita Buga		NIMRD
	George Kallos		NKUA
	George Zodiatis		ORION
	Stavros Stylianou		ORION
	Frederique Blanc		CLS
	Claudia Cesarini		CLU

Document Log

Date	Author	Changes	Version	Status
29.07.2016	Nadia Pinardi	Chapter first draft	1	completed
02.10.2016	Nadia Pinardi, Giuseppe Manzella and Luminita Buga	First, second and third Chapters draft	2	completed
04.10.2016	Nadia Pinardi, Vladyslav Lyubartsev	Chapter 4 and Annex 2	3	completed
05.10.2016	Nadia Pinardi, Vladyslav Lyubartsev, Giuseppe Manzella	Update of Chapter 4 and Chapter 5 first version	4	completed
07.10.2016	F.Blanc, G.Manzella, L.Buga	Corrections and Annex 1	5	completed
08.10.2016	V.Lyubartsev N.Pinardi	Corrections to Annex 2 and section 4 First version of conclusions	6	completed
10.10.2016	N.Pinardi, C.Cesarini	Title page and executive summary	7	completed
13.10.2016	N.Pinardi, external consultant	English revisions	8	completed
15.10.2016	N.Pinardi, L.Buga and A.Palazov	Final version	9	completed
15.12.2016	N.Pinardi, G.Manzella, L.Buga	Revised version after consolidated comments by EASME	10	draft

Table of Contents

Glossary

2AR	Second Report on the Adequacy of the Global Climate Observing System for Climate
ADCP	Acoustic Doppler current profiler
AG FOMLR	Advisory Group on Environmental Aspects of Management of Fisheries and Other Marine Living Resources
AREG	Adriatic Sea regional model
ASAR	Advanced Synthetic Aperture Radar
ASCII	American Standard Code for Information Interchange
AQUASTAT	FAO's global information system on water and agriculture
ARGO	The broad-scale global array of temperature/salinity profiling floats
BGODC	Bulgarian National Oceanographic Data Centre
BfG	German Federal Institute of Hydrology (Bundesanstalt für Gewässerkunde), Germany
BODC	British Oceanographic Data Centre
BSC	Black Sea Commission
BSERP	Black Sea Ecosystem Recovery Project
BSIS	Black Sea Information System
CAF	Committee of Administration and Finance
CARLIT	Cartography of littoral rocky-shore communities
CAQ	Committee on Aquaculture
CDI	Common Data Index
CDOM	Colored dissolved organic matter
CDS	Catalogue of Data Sources
CFP	Common Fisheries Policy
CFRI	Central Fisheries Research Institute, Trabzon, Turkey
CH	Challenge – Check point application area
Characteristic	Distinguishing feature[1]
CI	Citation
Class	Description of a set of objects that share the same attributes, operations, methods, relationships, and semantics [UML Semantics] NOTE: A class does not always have an associated geometry (e.g. the metadata class).
CLS	Collecte Localisation Satellites (FR)
CLU	CLU s.r.l. (IT)
CMEMS	Copernicus Marine Environment Monitoring Service
CMS	Content management System
CoC	Compliance Committee
Copernicus	European Programme for the establishment of a European capacity for Earth Observation
Coverage:	a feature that has multiple values for each attribute type, where each direct position within the geometric representation of the feature has a single value for each attribute type[2]. Coverage is an abstraction of continuous real world phenomena[3].
COST	Cooperation in Science and Technology
CSR	Cruise Summary Report
CSW	Catalogue Service for Web
CTD	Conductivity, Temperature, Depth
DAC	Data Assembly Center
DAR	Data Adequacy Report
Data	Re-interpretable representation of information in a formalized manner suitable for communication, interpretation, or processing[4]
DCF	Data Collection Framework
DCR	Data Collection Regulation

---

[1] ISO 9000:2005 Quality management systems. Fundamentals and vocabulary
[2] Quality/FDIS 19123 2005 Geographic information – Schema for coverage geometry and functions
[3] S. Nativi, J.Caron, B.Domenico and L.Bigagli, 2008. Unidata's Common Data model mapping to the ISO 19123 Data Model, Earth Sc. Informatics, Vomule 1,Issue 2, pp 59–78
[4] ISO/IEC 2382-1:1993 Information technology – Vocabulary – Part 1: Fundamental terms

DCRF	Data Collection Reference Framework
DG-MARE	Directorate-General for Maritime Affairs and Fisheries
DEM	Digital Elevation models
DO	Dissolved oxygen
DPS	Data Product Specification[5]
DQ	Data quality
DTM	Digital Terrain Model
EAFM	Ecosystem-based approach to fisheries management
EC	European Commission
ECMWF	European Centre for Medium-Range Weather Forecast
ECVs	Essential Climate Variables
EDIOS	European Directory of Oceanographic Observing Systems
EDMED	European Directory of Marine Environmental Data
EDMERP	European Directory of Marine Environmental Research Projects
EDMO	European Directory of Marine Organisations
EEA	European Environmental Agency
EEC	European Economic Community
EEZs	Exclusive Economic Zones
EIONet	European Environment Information and Observation Network
EMBRC	European Marine Biological Resource Centre
EMODnet	European Marine Observation and Data Network
EMSA	European Maritime Safety Agency
EMSO	European Multidisciplinary Seafloor and water-column Observatory
ERIC	European Research Infrastructure Consortium
ESA	European Space Agency
ESFRI	European Strategy Forum on Research Infrastructures
ETC	European Topic Centre
EU	European Union
EUMETNET	European National Meteorological Services
EUNIS	European Nature Information System
EUROGOOS	European Global Ocean Observing System
FAO	Food and Agriculture Organization
Feature	Abstraction of real world phenomena. Discrete world phenomena are conceived as (discrete) features while continuous phenomena are conceived as features that "acts as a function to return values from its range for any direct position within its spatial, temporal or spatiotemporal domain (e.g., grids or images) named coverages[2]
Feature's attribute	Characteristic of a feature
FixO3	Fixed point Open Ocean Observatory network
Fondazione CMCC	Foundation Euro-Mediterranean Center for Climate Change (IT)
GEBCO	General Bathymetric Chart of the Oceans
GEMET	General Multilingual Environmental Thesaurus[6]
GEMS	Global Environment Monitoring System
GES	Good Environmental Status
GEO	Group on Earth Observation
Geoportal	Type of web portal used to find and access geographical information
GEOSS	Global Earth Observation System of Systems
GeoTIFF	Public domain metadata standard
GFCM	General Fisheries Commission for the Mediterranean
GIS	Geographic information system
GMES	Global Monitoring for Environment and Security
GOOS	Global Ocean Observing System
GPRS	General Packet Radio Service
GPS	Global Positioning System
GRDC	Global Runoff Data Centre

---

[5] ISO 19131:2007/Amd 1:2011 Requirements relating to the inclusion of an application schema and feature catalogue and the treatment of coverages in an application schema

[6] Marine Metadata Interoperability Project - GEMET - GEneral Multilingual Environmental Thesaurus http://marinemetadata.org/references/gemet

GTS	Greenwich Time Signal
ICZM	Integrated Coastal Zone Management
IEC	International Electrotechnical Commission
IFR	Institute of Fishing Resources, Varna, Bulgaria
IFREMER	Institut Français de Recherche pour l'Exploitation de la Mer (FR)
IHO	International Hydrographic Organization
IMO	International Maritime Organization
IMP	Integrated Maritime Policy
IMS	Middle East Technical University Institute of Marine Sciences (TR)
Information	Knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning[4]
INSPIRE	Infrastructure for Spatial Information in the European Community[7]
IO-BAS	Institute of oceanology, Bulgarian Academy of Sciences (BG)
IOC	Intergovernmental Oceanographic Commission
IPCC	Intergovernmental Panel on Climate Change
IR	Infrared
ISO	International Organization for Standardization
ISO IEC	ISO International Electrotechnical Commission
ISO NP	ISO New Proposal
ISO NP TS	ISO NP Technical Specification
IT	Information Technology
IUU	Illegal, Unreported and Unregulated
JCOMM	Joint WMO-IOC Commission on Marine Meteorology
JECMAP	Joint European Coastal Mapping Programme
JRC	Joint Research Centre
KTU-MSF	Black Sea Technical University, Marine Science Faculty, Trabzon, Turkey
LAT	Lowest Astronomical Tide
LE	Lineage extended
LI	Lineage
LWN	Normalized water leaving radiance
LiDAR	3D laser scanning
MARBOUND	Maritime Boundaries Geodatabase
MD	Metadata
MedSea	Mediterranean Sea
MERCATOR	French center for analysis and forecasting of the global ocean
MMI	Marine Metadata Initiative/Marine Metadata Interoperability[8]
MPA	Marine protected areas
MRE	Marine renewable energy
MS	Member States
MSFD	Marine Strategy Framework Directive
MSP	Maritime Spatial Planning
MSSD	Mediterranean Strategy for Sustainable Development
MyOcean	Series of projects granted by the European Commission within the GMES Program (Seventh Framework Program)
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NEBS	North-Eastern Black Sea
NetCDF	Network Common Data Form
NeXOS	Next Generation Web-Enabled Sensors for the Monitoring of a Changing Ocean
NGO	Non-governmental organization
NIC	National Ice Center
NOAA	National Oceanic and Atmospheric Administration
NRT	Near Real Time

---

[7] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE): http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32007L0002

[8] https://marinemetadata.org/

NKUA	National and Kapodistrian University of Athens (GR)
NIMRD	National Institute for Marine Research and Development "Grigore Antipa"(RO)
NWBS	North-Western Black Sea
Object	Entity with a well-defined boundary and identity that encapsulates state and behaviour [UML Semantics]. NOTE: An object is an instance of a class
ODV	Ocean Data View
OGC	Open Geospatial Consortium
ORION	Joint research and development centre (CY)
OSSE	Observing System Simulation Experiments
OSE	Observing System Experiment
OWF	Offshore Wind Farms
pH	Logarithmic measure of hydrogen ion concentration
P01	BODC Parameter Usage Vocabulary
P02	SeaDataNet Parameter Discovery Vocabulary
P03	SeaDataNet Agreed Parameter Groups
Package	Grouping of a set of classes, relationships, and even other packages with a view to organizing the model into more abstract structures
PNG	Portable network graphics
POMOS	Port Operational Marine Observing System
PR	Pre-eutrophication
PS	Post-eutrophication
PSMSL	Permanent Service for Mean Sea Level
PSU	Practical Salinity Units
QC	Quality Control
QE	Data quality extended
Quality	Degree to which a set of inherent characteristics fulfils requirements[1]
Requirement	Need or expectation that is stated, generally implied or obligatory[1]
RES	Renewable Energy Systems Limited (UK)
RFMO	Regional fisheries management organization
RivDIS	Global River Discharge data set
ROOS	Regional operational system
ROV	Remotely operated underwater vehicle
SAC	Scientific Advisory Committee
SAGE	Systems Approach to Geomorphic Engineering
SAR	Synthetic aperture radar
S-AWS	Ship-borne Automated Weather Stations
SCMR	SC Marine Research SRL (RO)
SeaDataNet/SDN	Pan-European infrastructure to ease the access to marine data measured by the countries bordering the European seas
SeaVoX	Combined SeaDataNet and MarineXML Vocabulary Content Governance Group
SID	Source identifier
SIO-RAS	P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences (RU)
Specification scope	Part of the data content of a product sharing the same specifications
SPLASHCOS	Submerged Prehistoric Archaeology and Landscapes of the Continental Shelf
SRTM	Shuttle Radar Topography Mission
SST	Sea Surface Temperature
SPOT	Commercial high-resolution optical imaging Earth observation satellite system operating from space
STAG	Scientific and technical advisory group
STECF	Technical and Economic Committee for Fisheries
SWH	Significant wave heights
TSU	Ivane Javakhishvili Tbilisi State University (GE)
TAC	Total Allowable Catch
UkrSCES	Ukrainian Scientific Centre of Ecology of the Sea (UA)
UPL	Plymouth University (UK)
UN	United Nations
UNCLOS	United Nations Convention on the Law of the Sea
UNEP	United Nations Environment Programme

UNIDATA	Data Services and Tools for Geosciences[9]
UNESCO	United Nations Educational, Scientific and Cultural Organization
URL	Uniform Resource Locator
USE CASE	Exemplary literature case related to Challenge Targeted products
USOF	University of Sofia (BG)
UV	Ultraviolet
WISE	Water Information System for Europe
WBS	Western Black Sea
WFD	Water Framework Directive
WGBS	Working Group for the Black Sea
WMO	World Meteorological Organisation
VLIZ	Vlaams Instituut voor de Zee, Belgium
VMS	Vessel Monitoring System
VOS	Voluntary Observing Ships
XML	eXtensible Markup Language
YugNIRO	Southern Research Institute of Sea Fisheries and Oceanography, Kerch (RU)

---

[9] http://www.unidata.ucar.edu/

"The information and views set out in this report are those of the authors 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."

Executive Summary

This document illustrates the first Data Adequacy Report (DAR) of the EMODnet Black Sea Checkpoint project. The aim of the DAR is to assess the basin scale monitoring systems on the basis of input data sets for 11 prescribed Challenges. Nothing equivalent exists in the world. In fact, a completely new methodology has been developed based upon ISO[10] standards, INSPIRE principles and a set of quality indicators.

The document contains seven sections, the first three overviewing the assessment framework and the geoportal infrastructure build to implement such an assessment (Sections 1,2,3) and the other three showing the results of the assessment (Sections 4,5,6). A conclusive chapter discusses the major findings.

The first step in this process involved the definition of a "Data Adequacy Framework", which was derived from the ISO 9004:2009 standards. Data Adequacy is essentially defined as the fitness for use of the monitoring data required by the Challenges. The CheckPoint adequacy relates to both the requirements as well as the needs of the Challenges and was developed considering the ISO 9001 Quality Management System. The quality assessment is subdivided into two major "territories": "appropriateness" and "availability", the first answering the questions "*what* is the quality of the monitoring data for the Challenge products" and the second "*how* is the data made available to the Challenges". In this first DAR only the "availability" indicators are explored and analyzed. Eight indicators are defined that are:

Availability indicators Short name	Short definitions	Explanation
AV-VI-1	**Easily found**	Can the datasets be found easily?
AV-VI-2	**INSPIRE Catalogue service**	Is the dataset referenced by an EU catalogue service?
AV-AC-1	**Policy visibility**	Visibility on data policy adopted by data providers
AV-AC-2	**Delivery mechanism**	What are the services available to access data
AV-AC-3	**Data Policy**	Data policy (open and free, open but at cost, etc.)
AV-AC-4	**Pricing**	Cost basis / price policy
AV-AC-5	**Readiness**	Format for use
AV-PE-1	**Responsiveness**	How responsive is the delivery service?

The indicator preset values are explained in details in Table 5.1 at page 34-35.

The second step in the analysis is to set up a metadatabase containing standardized information

---

[10] All acronyms and specific terms are explained in the Glossary and, when required, referenced to a bibliographic source in a footnote.

about the input datasets potentially usable by the Challenges to generate their products (to be generated in the second phase of the project). The Black Sea Checkpoint metadatabase currently contains information describing input data to the Challenges that are uniquely identified as a combination of categories of characteristics, providers and other descriptors, together with the availability indicators set. The metadatabase is at the back-end of an INSPIRE Web and GIS platform, known as Sextant, and uses the SeaDataNet common vocabulary to identify the categories of characteristics (see definition in Section 2.1) needed by the Challenges and to analyze the statistics of indicators.

This DAR includes the consolidated analysis of the needed characteristics required for the Challenge products and their input data sets. The Black Sea Checkpoint metadatabase now contains 452 input data sets describing 40 different characteristics categories. About half of these categories are used by more than one Challenge, and several input datasets are potentially capable of describing them for the Challenge products. Most of the Challenges involve several characteristics categories which will be potentially used to generate the required products. For each characteristics category there is large number of input data sets (50-100) mainly due to routine choices made by the Challenge experts.

This DAR contains the first assessment of the Black Sea monitoring system on the basis of the analysis of the availability indicators across all Challenges for the 452 input data sets and the 40 characteristic categories. The 8 availability indicators are classified based upon a three value range color system: "red" meaning "not adequate", "yellow" "partly adequate" and "green" "fully adequate". The analysis shows that for most of the indicators half are "not adequate" and the other half, are "adequate". The single most negative score is for the "INSPIRE catalogue service" (AV-VI-2) indicator, which is generally not adequate. Furthermore, the "Pricing" (AV-VI-4) indicator is split in half between "not well documented pricing policy" and "open and free data policy". In summary, however, the majority of the scores for the availability indicators are positive, meaning that most of the data sets to be used by the Challenges are "adequate" in the terms of how they are made available to the Challenges.

Another result was obtained by mapping the availability indicators of each data set as a function of the 40 different characteristics required by the Challenges. Twelve characteristic categories score "not adequate" or "partly adequate" and they are:
1. TSED/Concentration of suspended particulate material in the water column
2. CRYS/Snow and Ice mass, thickness and extent
3. CAPH/pressure (measured variable)
4. CDTA/Air temperature and density
5. CHUM/Atmospheric humidity
6. CNTX/Phytoplankton generic biomass in water bodies
7. GP079/Zooplankton wet weight biomass
8. EWSB/Wind strength and direction
9. MBAN/Bathymetry and elevation
10. PSAL/Salinity in the water column
11. WVSP/Spectral wave data parameters
12. WVST/Wave height and period statistics

This might already indicate a gap in the basin scale monitoring for the specified characteristics, at least from the point of view of adequacy in terms of availability.

Finally the EMODnet Thematic Portals and the Copernicus Marine Environment Service (CMEMS) have been evaluated in terms of availability indicators. Results show that with a few exceptions the scores are "totally adequate".

In conclusion, it is worth mentioning that the Black Sea Checkpoint service, based upon the metadatabase and the GIS web portal, is coordinated with the Mediterranean Sea and the Atlantic Checkpoint so that the availability indicator analysis will be carried out in the same way in the three basins. This will allow users to differentiate in a near future between the data adequacy of the three basin-scale monitoring systems.

1. Introduction

The concept of the EMODnet Checkpoint originated from the Marine Knowledge 2020 Communication11, which foresaw the development of new European Instruments such as Thematic Assembly centers (the EMODnet Portals). These centers are responsible for assembling data and re-distributing them after quality control and harmonization. They are a network repository infrastructure of private and public organizations that are designed to become an operational service.

In order to achieve the sustainable operation of marine observations systems and assembly centers, critical gaps in these systems need to be identified, based upon an integrated viewpoint at a sea-basin level. To ensure this, the Commission is exploring how a framework of sea-basin Checkpoints might work by setting up pilots in the period 2011-2015. Checkpoints would independently check data layers from each thematic assembly group, ensure that the data from different groups are mutually compatible and define priorities for further observations based on interaction with local stakeholders. These checkpoints should act on behalf of all users of marine data within that sea-basin and cover all EU and national initiatives on marine data – EMODnet, GMES, Data Collection Framework, etc.

The first two Pilot Checkpoint initiatives started in 2013 in the Mediterranean Sea and North Sea, followed two years later by the Black Sea, the Arctic basin, the Baltic Sea and the Atlantic Ocean. The sea basin approach is also being followed by the other large European initiative of Copernicus Marine Environment Service (CMEMS) for space and ocean forecasting and has proven to be effective in pooling the expertise together in order to collect, assemble and assess the monitoring data.

Fig. 1 The implementation areas of the Basin Checkpoints activites

During the implementation of Checkpoints, the concept of Data Adequacy was formulated, which consisted in the assessment of "usability" of monitoring data sets in view of 11 Challenges:

CH1- Windfarm Siting, CH2- Marine Protected Areas, CH3- Oil Platform Leak, CH4- Climate, CH5- Coasts, CH6- Fishery Management, CH7- Fishery Impacts, CH8- Eutrophication, CH9- River Inputs, CH10- Bathymetry, CH11- Alien species.

In the Challenge groups of each Checkpoint initiative, products will be developed following specifications given by DGMARE and the requirements established by the experts. The Challenge products are practical examples of how the monitoring data might be used by Member States governmental organizations and/or private companies for developing a renewable energy strategy, and protecting the sea from natural and man-made hazards and impacts (climate, oil spills, alien

---

[11] European Commission – Maritime Affairs – from seabed mapping to ocean forecasting
http://ec.europa.eu/maritimeaffairs/documentation/publications/documents/marine-knowledge-2020-green-paper_en.pdf

species, etc.). They were selected by DGMARE stakeholders and hopefully they will increase to cover all possible usages of marine data for the Blue Economy goals.

1.1 The overall Checkpoint methodology

The Checkpoint methodology was developed following two basic principles:

1) to develop a service to facilitate access to upstream data for the Challenges and make available the Challenge products as open data following the recommendations of INSPIRE;

2) to use the EEA core set of indicator concepts (EEA, 2005[12]) and develop indicators to assess the data adequacy following ISO-adapted standards.

The general flow of information from the upstream input data through the Challenge products to the Data Adequacy Reports and the Checkpoint service is illustrated in Fig. 2.
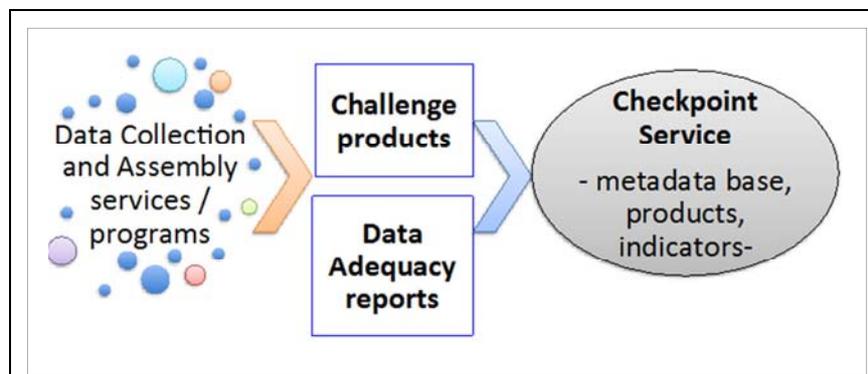
Fig. 2 The Checkpoint Framework, from Upstream Data to Checkpoint Service

The Checkpoint Service is an open data service and it will provide:

1) a metadatabase which describes the Upstream Data to be used for the Challenges;

2) the Challenge products with explanations about their quality and the methodology for their construction;

3) the indicators on the basis of which the Data Adequacy Assessment will be carried out.

The service works on a platform, called Sextant[13], that is entirely interoperable in adopting the ISO rules and the OGC (Open Geospatial Consortium) standards to get access to data. A Sextant metadatabase has been organized for the Black Sea EMODnet Checkpoint which contains information about the upstream data and will contain the indicators for the Data Adequacy assessment. This first DAR was constructed with the data contained in the Sextant Black Sea metadatabase.

1.2 The Literature Survey

The first step in the Checkpoint assessment led to the Black Sea Checkpoint Literature Survey[14] (LS), which mapped for the first time the main input data sets for the Challenges.

In the LS, an initial list of characteristics (nomenclature explained in Section 2 of this document) was produced with an overall analysis of the most requested input data sets and the potential data providers for the Challenges. A total of 48 different characteristics are required by the Challenges and over 400 data sets are in principle available to create the products. A list of 56 data providers

---

[12] http://www.eea.europa.eu/publications/technical_report_2005_1
[13] http://www.ifremer.fr/sextant
[14] To be found here: http://emodnet-blacksea.eu/practical-outputs/

was identified from international, European, EU Member State plus Institutions and projects in Russia and the USA.

An overall survey of the observational capacities and monitoring systems of the Black Sea was discussed which showed that there is a comprehensive overview of the current monitoring systems. It emerged that European Projects contribute to the development of the Black Sea large scale monitoring system while in the past ten years national measurement platforms have been put in place for national waters.

In order to have an initial understanding of the adequacy of the basin scale monitoring systems, the LS analysed several Use Cases from the literature that described similar products to the Checkpoint products. These Use Cases were chosen on the basis of the fact that they presented and evaluated the quality and availability of the input data sets required for the Use Case.

The 24 literature Use Cases were found to use data adequacy elements such as: 1) spatial resolution and area coverage; 2) temporal resolution and extent. Only one Use Case referred to the accuracy of the input data set, which was for the river runoff data. For all 24 Use cases, data availability was generally high except for accessibility. In a large number of Use Cases the data are completely restricted and/or access to the input data requires specific agreements with the data owners. Thus, the answers to the questions are:

Q1. Are there references in the literature to goals not achieved because of inadequate data (appropriateness)?

A1. Yes, main concerning data coverage in space and time and the resolution of the input data sets.

Q2 Is the inadequacy due to the reluctance of the data-owners to release data, the time taken to obtain data, a lack of measurements, lack of accuracy or lack of precision (availability)?

A2. Generally, data availability is thought to be high, with some reservations regarding the accessibility and responsiveness. There is no mention of accuracy except for one Use Case.

Q3. Are there any statements made as to fitness for purpose of data?

A3. The fitness for purpose of data, which in this document we call fitness for use, is related to the previously mentioned appropriateness elements.

The results of the LS will be consolidated in this first DAR since the Sextant metadatabase has been upgraded and completed with all the information.

1.3 Structure of the document

This document is structured around five major sections:

> 1) The Data Adequacy Assessment framework and methodology (Section 2);

> 2) The structure of the metadatabase and its content (Section 3);

> 3) The consolidated LS and the revised results (Section 4);

> 4) The first results on Data Adequacy with the availability indicators (Section 5) and the specific analysis of the indicators for the CMEMS and EMODnet portals (Section 6);

Section 7 concludes with a synthesis of findings.

2. Data Adequacy framework and methodology

An objective assessment of the existing data quality and their 'usability' for Challenge products should be based on selected ISO standards which provide the methodology, definitions and quality elements used to establish indicators.

The methodology is derived from ISO9004-2009 standards, which are part of the ISO9000 series. These are based on principles that can be adapted to Checkpoints:

- Customer focus: understand current and future customer needs, should meet customer requirements and strive to exceed customer expectations.
- Factual approach to decision making: effective decisions are based on the analysis of data and information.
- Continual improvements: assessment of existing monitoring systems, improvements and overall performance should be a permanent objective of decision makers
- Mutually beneficial supplier relationships: data suppliers and data users are interdependent and a mutually beneficial relationship enhances the ability of both to create value.

ISO 9004:2009[15] provides organizations with a model for "sustained success" in today's complex, demanding, and ever-changing environment. This can be used to assess the quality of the products of the Challenges and the existing service delivery to stakeholders by:

- Benchmarking their level of quality
- Identify their strengths and weaknesses
- Identify opportunities for either improvements or innovation, or both.

ISO/IEC 25010 standard introduces concepts that are also contained in the other ISO standards, such as 'Usability' (ISO/IEC 25010 Product Quality) and 'Effectiveness and Context Coverage' (ISO/IEC 25010 Quality in Use).

The general framework for the assessment of the quality of products of Challenges and their input data sets has been provided by the "Methodology to assess and communicate the economic benefits of consensus-based standards"[16] developed by ISO. The quality elements that allow the objective assessment are provided mainly by the ISO 19157:2013(E) Data Quality and ISO 19115:2014(E) Metadata.

ISO19157:2013(E) provides the quality element that is called "usability", i.e. the extent to which data sets or data set series can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. In other words, ISO 19157 standards contain elements for assessing 'how' and 'how much' data meets requirements in order to enhance user satisfaction.

The greatest effort in the Black Sea Checkpoint has been the development of an objective assessment of the existing data quality and their 'usability' for challenges products.

---

[15] ISO9004-2009. Managing for the sustained success of an organization. A quality management approach
[16] Assessing economic benefits of consensus-based standards – The ISO methodology.
http://www.iso.org/iso/home/standards/benefitsofstandards/benefits-detail.htm?emid=6

Note that datasets or dataset series could be used with intentions different from the one for which they were originally created.

2.1 Key definitions

In the Literature Review for the Black Sea Checkpoint, following the Mediterranean Checkpoint work, important efforts were made to provide definitions based on ISO standards.

- **Characteristic:** a distinguishing feature which refers :

    - either to a variable derived from the observation, the measurement or the numerical model output of a phenomenon or of an object property in the environment;
    - or to the geographical representation of an object on a map (ie a layer such as a protected area, a coastline or wrecks) by a set of vectors (polygon, curve, point) or a raster (a spatial data model that defines space as an array of equally sized cells such as a grid or an image).

- **Environmental matrices:** The environments where characteristics are measured or computed:
    - Ice,
    - Air,
    - Fresh water,
    - Marine water,
    - Biota/Biology,
    - Riverbed/Seabed,
    - 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 dataset can be constituted of several files (e.g. the set of seismic data files recorded along the same line**).**
- **Collection of datasets:** a set of datasets.
- **Dataset series:** a collection of datasets sharing the same specifications of production. This is the concept in use on the INSPIRE Geoportal.
- **Input Dataset:** the collection of existing data to be input to the Challenges
- **Assessment criteria:** the criteria are focused on two questions: "what" is made available to the challenges and 'how'. Appropriateness (what) and availability (how) indicators were defined using ISO 19113 and ISO 19157 standards.
- **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 itself, 'adequacy' should be defined objectively using standardized nomenclature and methods. In an EC Report[17] adequacy was defined as an assessment of the reported information to meet the objectives of the Marine Strategy Framework Directive (MSFD) and its

---

[17] The first phase of implementation of the Marine Strategy Framework Directive (2008/56/EC) - The European Commission's assessment and guidance. CELEX_52014SC0049_EN_TXT

technical requirements listed in MSFD Articles 8, 9 and 10[18]. The CheckPoint adequacy is close to this definition but focused on several Challenges. In other words **adequacy** is here intended as 'sufficient to satisfy a requirement or meet a need'[19]. From this definition, 'adequacy' relates to meeting both requirements as well as needs and is normally applied within the framework of an ISO 9001 based Quality Management System.

2.2 Assessment criteria

The aim of the Checkpoints is to assess the adequacy of the monitoring or data collection strategy at the European basin scale level under specified operational conditions (represented by the Challenges). 'Adequacy' thus relates to meeting both requirements as well as needs of the users and is normally defined within the framework of an ISO 9001 based Quality Management System.

The "Methodology to assess and communicate the economic benefits of consensus-based standards"[20] developed by ISO proposes a general framework for the objective assessment of the monitoring system quality and allows sharing of results, supporting benchmarking and identification of general trends.

ISO 19157 standards introduce the elements for assessing 'how' and 'how much' data meets requirements in order to enhance user satisfaction. The application of the ISO methodology provides a set of standards that objectively measure the 'data adequacy' and it was chosen by the CheckPoint as the basic methodology.

The assessment criteria are subdivided into two 'Territories' that need to be evaluated in terms of Challenge requirements. They are:

Territory 1: Availability
How is the data made available to the Challenges
Territory 2: Appropriateness
What is the quality of the monitoring data for the challenge products

Data Adequacy will be assessed in two tiers. This first DAR will analyze the 'Availability' territory for each Challenge input dataset. Adequacy assessments will be based on the transformation of the input metadata into an ordered set of indicators.

2.3 Territory 1: Availability indicators

'Availability' measures the extent to which datasets are ready for use and are obtainable. To obtain datasets, information is needed on the data provider (visibility), how to access them (accessibility), and long it will take to obtain them (performance). The 8 Availability indicators chosen by the project are listed in Table 2.1.

---

[18] Adequacy does not necessarily mean, for instance, that if the defined data is adequate, this automatically means that the quality of the marine waters Is acceptable

[19] Random House Unabridged Dictionary, Random House Inc, 2006

[20] Assessing economic benefits of consensus-based standards – The ISO methodology.
http://www.iso.org/iso/home/standards/benefitsofstandards/benefits-detail.htm?emid=6

Name of Availability indicators	Definitions
AV-VI-1	Easily found
AV-VI-2	INSPIRE Catalogue service
AV-AC-1	Policy visibility
AV-AC-2	Delivery mechanism
AV-AC-3	Data Policy
AV-AC-4	Pricing
AV-AC-5	Readiness
AV-PE-1	Responsiveness

Table 2.1 Availability indicators and their short description

The availability indicators (AV) provide an understanding of the readiness and service performance of the infrastructure providing access to data. The availability indicators are subdivided into three parts:

1) Visibility (VI), i.e. the possibility of identifying and quickly accessing the appropriate site for the required data sets;
2) Accessibility (AC) i.e. the possibility, for non expert users, to understand the retrieval model status;
3) Performance (PE) i.e. the ability of a system to keep operating over time and to meet real time operational conditions. This is related to service performance.

The following sections provide details of each indicator.

2.3.1 Visibility indicators

"Visibility" is the ability to identify and quickly access the appropriate site delivering the desired data sets. In other words it is the ability for all users, including non-experts, to perform data sourcing through an EU Inspire catalogue. Two indicators have been defined for the visibility element, i.e.:

AV-VI-1 Easily found	Can the data sets or series of data sets be found easily?
AV-VI-2 INSPIRE catalogue service	Is the dataset referenced by a EU catalogue service (INSPIRE) or by other bodies (private or public, national or international non EU services[21])

By referring to the INSPIRE Directive, this AV-VI-1 indicator provides information on visibility of data in catalogues. The AV-VI-2 indicator informs users whether the characteristic can be searched for by a catalogue service, such as EMODnet Thematic Portals, Copernicus core services, EEA services, DG MARE services, INSPIRE Geoportal, etc. Both indicators are identified as part of the INSPIRE Metadata Implementing Rules B 1.4 and the technical guidelines are based on EN ISO 19115 and EN ISO 19119.

---

[21] The non-EU services could be advanced services, but not following the guidelines defined in INSPIRE and its technical annexes.

2.3.2 Accessibility indicators

Accessibility is the ability of all users, including non-experts, to understand the retrieval model status and its appropriateness. ISO 19115 provides a general mechanism for documenting different categories of constraints applicable to the resource (or its metadata). The constraints could be legal and/or security constraints.

The INSPIRE Implementing Rules defines the metadata concepts for limitations on public access in part B 8.2 that apply to access constraints in order to ensure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the resource. In relation to constraints classes, there may be three scenarios according to the INSPIRE rules:

- There might be no limitation on public access;
- There might be only a classification property when expressing a security constraint;
- There might be one or more instances of the access constraints property, possibly associated with one or more instances of other restrictions property (e.g., Legal Constraints).

There are five indicators devised for accessibility:

AV-AC-1 Policy visibility	Visibility on data policy adopted by data providers.
AV-AC-2 Delivery mechanism	Data delivery mechanisms, i.e. the services available to the user to access data
AV-AC-3 Data Policy	Data policy (open and free, open but at cost, etc.)
AV-AC-4 Pricing	Cost basis / price policy
AV-AC-5 Readiness	Format for use

In the framework of the "blue growth" and for the specific indicator on Data policy, the exact meaning 'open' has not been established. Among the many definitions of 'open', one or more of these can be adopted:

- Accessible to all; unrestricted to participants
- Free from limitations, boundaries, or restrictions
- Usable by registered users

The indicator will classify all of these under the same score value.

2.3.3 Performance indicators

The performance indicators indicate the ability of a system to keep operating over time and to meet real time operational conditions. It is related to service performance. Only one indicator is defined for performance:

AV-PE-1 Responsiveness	How responsive is the delivery service for the available data?

2.4 Territory 2: Appropriateness indicators

Appropriateness indicators focus on "what" is available for the Challenges and include the spatial extent and resolution, time extent and resolution, purpose, lineage, usage, completeness, consistency, and accuracy. Appropriateness will consist in the assessment of fitness for use of input data sets for making Challenge products.

The assessment with appropriateness indicators will be carried out in the second DAR by comparing the Challenge products with specifications defined by the user requirements. The EMODnet Black Sea Checkpoint has not yet developed the products.

2.5 Indicator evaluation scale

Indicators provide both an overview of the situation at a high level of aggregation as well as detailed information about trends and links. The difficult taks is to find an appropriate balance between simplification and completeness. However indicators offer an objective way to assess a problem without directly accessing all the metadata.

The indicators are built from descriptors in several stages. In this DAR we build the indicators from the raw descriptors without prior aggregation of the information by the Challenges, releasing a neutral and basic status of indicators. The indicators of the first DAR will be time-bound and will be updated every six months up until the second DAR.

To provide an easy and straightforward visual representation of indicators, allowing a non-expert to easily assess the fitness for use without looking at the metadata and reports, it was decided to use three colours:

- Red: urgent actions are required to provide datasets and services fitting for use – totally inadequate
- Yellow: limited actions are required to provide datasets and services fitting for use –partly adequate
- Green: actions and services are fit for use and should be maintained –fully adequate

3. The Checkpoint metadatabase

As part of the Literature Survey, information was collected and analysed for upstream data classification, which is an iterative process that has now been consolidated for this first DAR. The information is managed by an INSPIRE Web and GIS platform, the Sextant, and some of its features are described below.

3.1 From Literature Survey to Data Adequacy Report

In order to organize information on the characteristics and Challenge products and to manage the checkpoint information, a metadatabase was set up. In the Literature Survey, the nomenclature and the assessment criteria were defined, while in this first DAR the metadatabase was consolidated and the indicators extracted from the statistical analysis of the metadatabase.

Checkpoint information describes the upstream data used as input data sets for the creation of the challenge products requested by the call. Their descriptor set (metadata set) will first be used for discovery purposes (inventory of data collections and sets needed for each Challenges), then to record the quality information resulting from the fitness for use evaluations. The latter will be carried out when the challenge uses them to create the product specified in the call.

3.2 Sextant and GIS Technology

To meet INSPIRE Directive requirements, scope of the web, checkpoint needs and dissemination of challenge products, the Consortium decided to use the Sextant data infrastructure for marine environments. It provides access to various geographical data via web services using standards defined by the Open Geospatial Consortium (OGC) and the ISO Technical Committee ISO/TC 210, Geographic information/Geomatics.

Sextant uses Geonetwork to set up the Catalogue Services for the Web and is used by several EU projects, such as EMODnet, GeoSeas, SeaDataNet and services, such as CMEMS. It is compliant with ISO 119115 and ISO19139 metadata standards and fulfills INSPIRE Directive requirements.

For the purposes of the EMODnet Black Sea Checkpoint, Sextant and the CMS suite of tools are used to render a number of metadata services as presented in Fig. 3.1 They are thus used to describe the upstream data and to handle the information needed for the establishment of the fitness for use indicators.
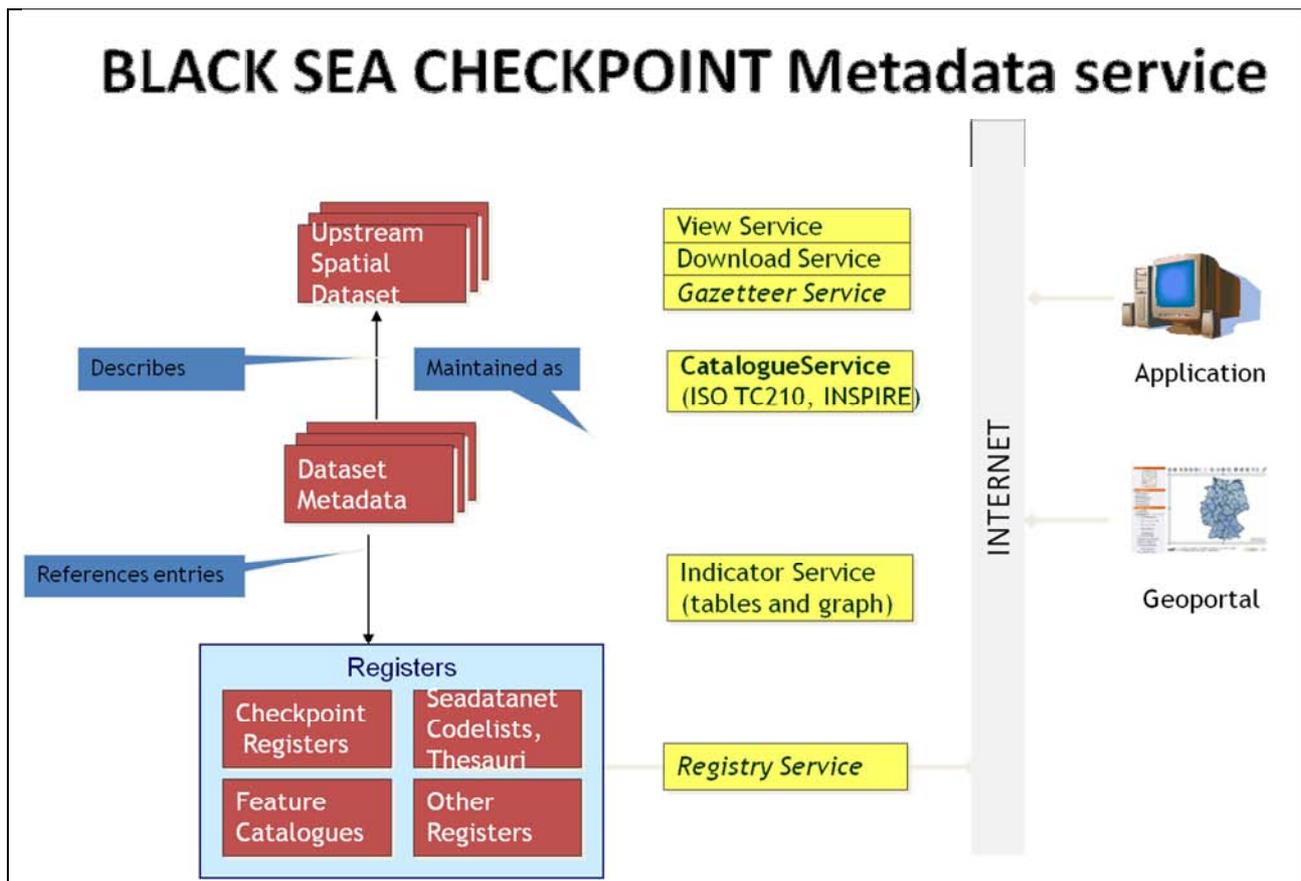
Fig. 3.1 The Black sea checkpoint metadata service model

3.3 Metadatabase structure

The Black Sea CheckPoint metadatabase contains information describing input data to the Challenges that are uniquely identified as a combination of categories of characteristics, providers and other descriptors. The descriptors are split into seven groups:

1. Characteristics (= What)
2. Data source (=Who)
3. Overview (= What for)
4. Spatial coverage (= Where)
5. Time coverage (= When)
6. Availability (= How)
7. Management metadata

These descriptors make up the elements of the territory assessment criteria previously described in Section 2. Descriptor groups 1, 2 and 3 are required to uniquely identify input datasets needed by the Challenges as well as to identify the catalogues (visibility) and documentations where metadata describing data can be found. The Checkpoint metadata structure for one characteristic is given as an example in Annex 1.

The SeaDataNet common vocabularies (http://www.seadatanet.org) are used to classify the characteristics, associate them to the Challenges and to the environmental matrices.

The European Directory of Marine Organisations - EDMO (http://www.seadatanet.org/Metadata/EDMO-Organisations) and the European Directory of Marine

23

Environmental Research Projects EDMERP (http://www.seadatanet.org/Metadata/EDMERP-Projects) are used for the data providers.

The data provider here is intended both as primary provider, i.e. data originator, and secondary data provider, i.e. data distributor. The 2014 version of ISO19115 provides a list of the different categories of data providers:

1. Originator : "party that created the resource " (i.e. dataset)
2. Distributor : "party that distributes the resources"
3. Custodian "party that accepts accountability and responsibility for the resource and ensures appropriate care and maintenance of the resource"

The data sources are characterized by 13 items to fully describe (when possible) the data sources, the provider, originating program of dataset or dataset series including identification in catalogues or directories. However no characterization in terms of primary or secondary provider is given in the metadatabase since in some cases the data are not published by the primary sources but only by the secondary sources and it is difficult then to identify the primary data sources.

The metadatabase is enabled by a dedicated CSW technology based on an ISO XML schema embedding:

- ISO 19115/139 for contextual metadata - classical reference, for catalogue
- ISO 19157 for CheckPoint indicators - such as fitness for purpose and "used by" any application (use cases).

In the Literature Survey the following ISO standards were analysed to construct the metadatabase which in turn helps in constructing 'adequacy' indicators:

1. ISO/NP (New Proposal) 19157 Geographic information - Data quality
2. ISO/NP TS (Technical Specification) 19158 Geographic information - Quality assurance of data supply

The metadatabase follows a well-defined process to edit, validate and consolidate its content. The metadatabase will be visible and accessible through three interfaces:

✓ Checkpoint Browser - A public access function to search and explore the input datasets and the Challenge products;
✓ Checkpoint GIS - A function under public and restricted access to visualize and manage Checkpoint input datasets and the Challenge products;
✓ Checkpoint Dashboard - A function under restricted access to compute and display indicators, directly built from metadatabase descriptors.

3.4 Vocabularies

Data processing and interoperability is possible only if common vocabularies are used. The same variable sometimes has different names depending on the repositories or the applications. Common vocabularies have been developed in many international initiatives, such as GEMET (promoted by INSPIRE as a multilingual thesaurus), UNIDATA, SeaDataNet, and the Marine Metadata Initiative (MMI).

When a vocabulary is formally managed, it becomes a controlled vocabulary. In this case, "managed" means the terms are stored and maintained using agreed- upon procedures. Procedures should exist for adding terms, modifying terms and, more rarely, rejecting terms from a controlled vocabulary.

The Sextant monitoring data catalogue uses the SeaDataNet Common Vocabularies.

"Use of common vocabularies in all metadatabases and data formats is an important prerequisite towards consistency and interoperability. Common vocabularies consist of lists of standardised terms that cover a broad spectrum of disciplines of relevance to the oceanographic and wider community. Using standardised sets of terms solves the problem of ambiguities associated with data markup and also enables records to be interpreted by computers". (http://www.seadatanet.org/)

The SeaDataNet is a controlled vocabulary containing terms that are:

1. Accepted: Each term adheres to community practices;
2. Defined: The terms are characterized precisely. Typically, this means the terms have rigorous definitions;
3. Managed: There is a body of experts that create and maintain the controlled vocabulary. The controlled vocabulary maintenance involves periodic review, addition of new terms, and modification of terms.

The Black Sea Checkpoint uses the SeaDataNet parameter vocabularies and INSPIRE spatial themes. The SeaDataNet vocabulary adopts a hierarchical approach for the classification of terms:

1. Agreed Parameter Groups (P03) - Terms agreed within the EU SeaDataNet community to describe coarse-grained groupings of related measurement phenomena,
2. Parameter Discovery Vocabulary (P02) Terms describing fine-grained related groups of measurement phenomena designed to be used in dataset discovery interfaces,
3. Parameter Usage Vocabulary (P01) - Parameter semantic model designed to describe individual measured phenomena.

This hierarchy goes from a coarser (P03) to a finer classification of a given dataset (P01).
The INSPIRE spatial themes are managed/governed by the Inspire geographic community of practices: web service GEMET (General Multilingual Environmental Thesaurus), 33 spatial themes described in

4. GEMET - INSPIRE themes (P22)   - Groupings of spatial data according to Annex I, II and III of the INSPIRE Directive [DS-D2.5]

In the Literature Review and the DAR we use P01,P02, P03 and P22 parameter nomenclature.

4. Updated Literature Survey Results

In this section, we present a statistical analysis of the metadatabase content, in particular the number of characteristics and input data sets required by the Challenges. The Literature Survey already presented results but the metadatabase had not been built at the time, and several iterations were made afterwards to consolidate the content and correct the mistakes, a process will continue until the second DAR.

4.1 Identification and classification of characteristics and input data sets for Challenges

Table 4.1 summarizes the characteristics and input datasets required by the Challenges. The changes with respect to the Literature Survey results are as follows:

- the total number of input data sets has changed from 436 to 452;
- the number of environmental matrices now includes "ice";
- the characteristics identified by P03, P02, and P22 have been reduced as a result of removing duplicates.

	Previous value from Literature Survey	Present value	Challenge										
			01	02	03	04	05	06	07	08	09	10	11
Number of input data sets identified	**436**	**452**	47	37	19	116	45	3	7	45	71	38	24
Environmental matrices identified	**6**	**7**	3	4	5	3	2	1	2	1	2	1	1
Numbers of different P03 identified	**26**	**23**	6	10	8	3	2	1	1	6	5	1	1
Numbers of different P02 identified	**48**	**40**	12	12	9	5	2	2	3	10	7	1	2
Numbers of different INSPIRE spatial themes identified (P22)	**14**	**12**	4	5	6	2	3	1	1	1	3	1	1

Table 4.1: Two first columns: Updated number of input data sets aggregated with the three different vocabulariy nomenclatures and environmental matrices, from the Literature Survey and the present metadatabase. Column 4 to 14: The number of input datasets by Challenge and the enrivonmental matrices, P02, P03 and P22 characteritstics by Challenge. (P02, P03, P22 numbers do not match with the overall "Numbers of different P02, P03 and P22 identified" because the same characteristic is requested by more than one Challenge).

At present, there are 452 input data set descriptions in the Black Sea EMODnet Checkpoint metadatabase. Note that these are only potentially useful input data sets for the Challenge products, which will be selected in the second part of the project when the Challenge outputs are produced.

In Figure 4.1 shows the number of input data sets as a function of the Challenges. CH04 "Climate" shows the maximum number of input datasets, 116, while CH06 "Fish management" shows the minimum, 3 datasets.
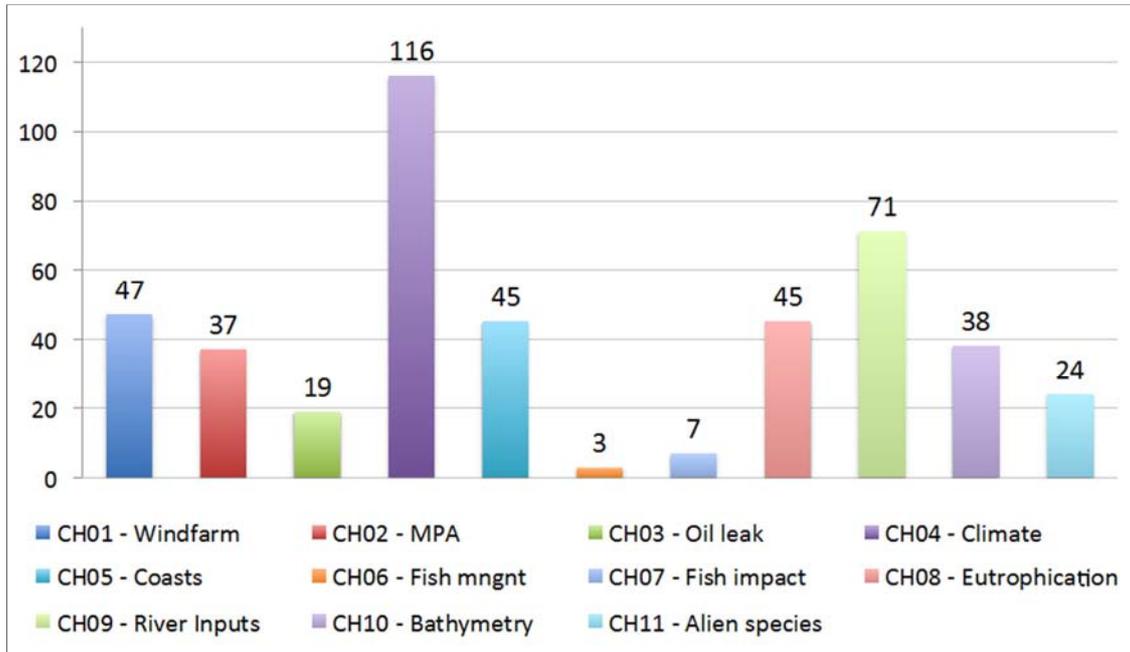
Figure 4.1: Number of input data sets for each challenge (visual representation of Table 4.1)

Figure 4.2 shows the number of characteristics mapped by the SeaDataNet parameter vocabularies P02, P03, and P22, as a function of Challenges. As CH10 "Bathymetry", is a "data assembly" Challenge, it only has one characteristic identified. CH01, CH02, CH08 and CH09 have the largest number of characteristics required.
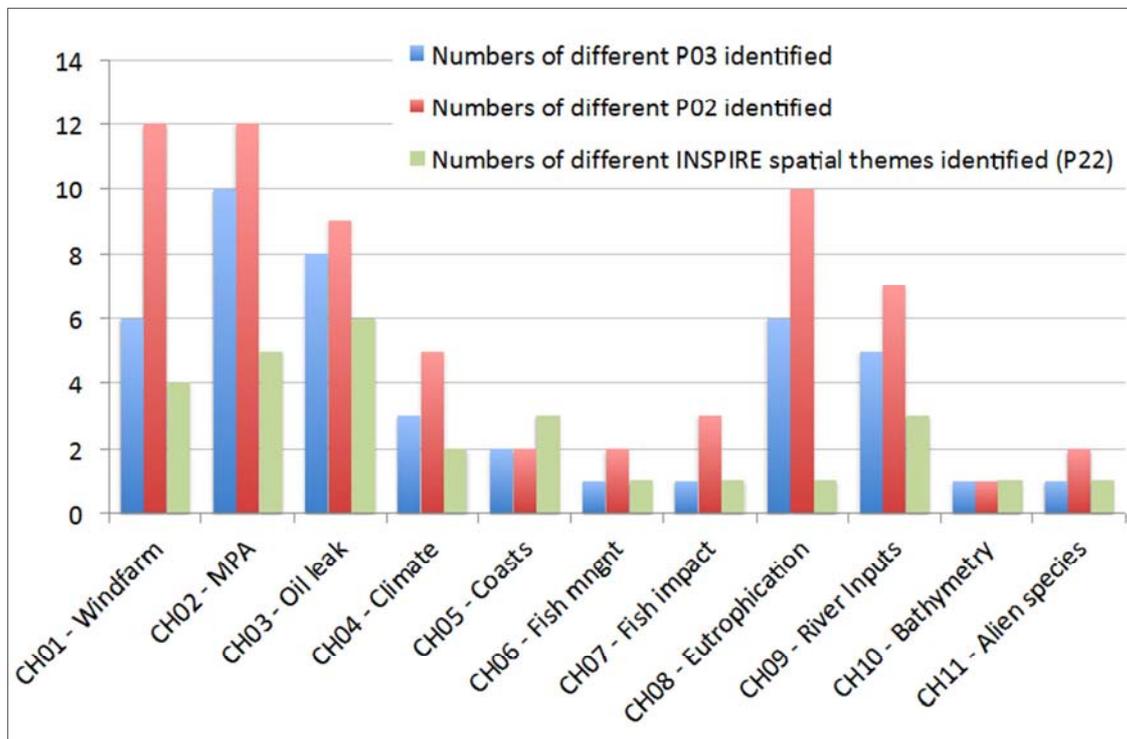

**Figure 4.2: Number of Input Characteristic categories identified by P02, P03 and P22 as a function of Challenges
(visual representation of Table 4.1)**

Environmental Matrix	All	Challenge										
		01	02	03	04	05	06	07	08	09	10	11
Air	15	11		4								
Biota/Biology	49		15	1	9							24
Fresh water	34					2				32		
Human activities	5		1	1				3				
Ice	17				17							
Marine water	284	32	19	9	90	43	3	4	45	39		
Seabed/Riverbed	48	4	2	4							38	

Table 4.2: Number of input data sets mapped as a function of Challenges and Environmental Matrix Elements

Each challenge has its own spectrum of monitoring needs appropriate to its targeted product. Table 4.2 lists the input datasets in terms of environmental matrices for each Challenge and Figure 4.3 presents the results. Challenge 10 (Bathymetry) shows only one environmental matrix identified, since it is monothematic. Challenges 6 and 7 (Fishery management and Fishery impact) have the least number of input data sets as the data collection by the governmental authorities follows standardized protocols. Challenge 3 (Oil leak) uses data from 5 of the 7 environmental matrices.

Table 4.2 highlights that marine water, fresh water, riverbed/seabed and biota/biology are the main matrices where the input data sets are potentially numerous. Human activities is the matrix covered by the least number of data sets.
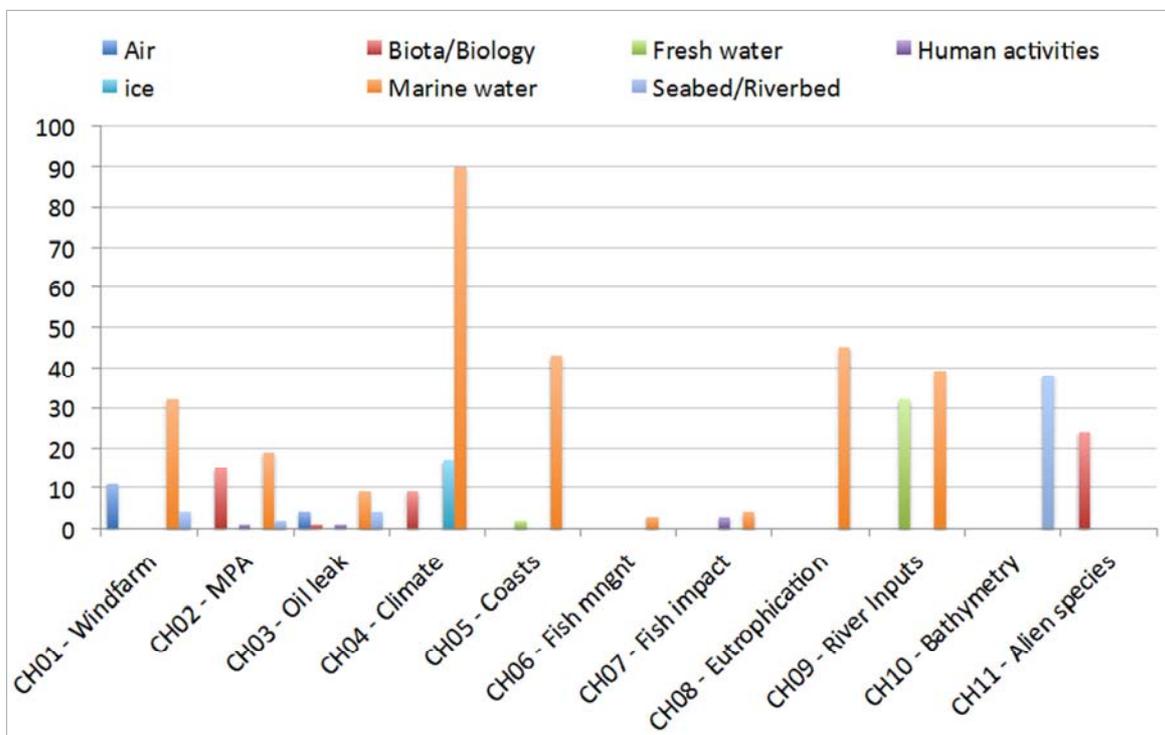
Figure 4.3: Number of input data sets as a function of environmental matrix and Challenges

(visual representation of Table 4.2)

4.2 Inventory of Challenge characteristic categories

This section contains a detailed list of Characteristics for each Challenge.

4.2.1 Characteristics for Challenge 1: Windfarm siting

Challenge 1 partners have identified 47 input datasets which correspond to 12 categories of characteristics (P02) and 3 environmental matrices as follows:

Air-   pressure exerted by the atmosphere, air temperature and density, atmospheric humidity, wind strength and direction;

Marine water-   temperature of the water column, salinity of the water column, horizontal velocity of the water column (currents), sea level, wave direction, spectral wave data parameters, wave height and period statistics;

Seabed/ Riverbed-   bathymetry and elevation.

4.2.2 Characteristics for Challenge 2: Protected marine areas

Challenge 2 partners have identified 37 input datasets which correspond to 12 categories of characteristics (P02) and 4 environmental matrices as follows:

Marine water-   skin temperature of the water column, temperature of the water column, horizontal velocity of the water column (currents), sea level;

Seabed/ Riverbed-   bathymetry and elevation;

Biota/biology-   bird taxonomy-related counts, cetacean abundance, chlorophyll pigment concentrations in water bodies, fauna abundance per unit area of the bed, habitat extent, horizontal spatial co-ordinates;

Human activities-   fishery characterization.

4.2.3 Characteristics for Challenge 3: Oil leak

Challenge 3 partners have identified 19 input datasets which correspond to 9 categories of characteristics and 5 environmental matrices as follows:

Air-   wind stress and direction;

Marine water-   temperature of the water column, horizontal velocity of the water column (currents), wave direction, wave height and period statistics;

Seabed/ Riverbed-   bathymetry and elevation, terrestrial mapping, coastal geomorphology;

Biota/biology-   habitat extent;

Human activities-   geological sample density.

4.2.4 Characteristics for Challenge 4: Climate

Challenge 4 partners have identified 116 input datasets which correspond to 5 categories of characteristics and 3 environmental matrices as follows:

Ice-   snow and ice mass, thickness and extent;

Marine water-   skin temperature of the water column, temperature of the water column;

Biota/Biology-   phytoplankton generic biomass in water bodies, phytoplankton generic abundance in water bodies.

4.2.5 Characteristics for Challenge 5: Coasts

Challenge 5 partners have identified 45 input datasets which correspond to 2 categories of characteristics and 2 environmental matrices as follows:

Fresh water-   river flow and discharge;

Marine water-   sea level.

4.2.6 Characteristicsfor Challenge 6: Fishery management

Challenge 6 partners have identified 3 input datasets which correspond to 2 categories of characteristics and 1 environmental matrix as follows:

Marine water-   fish taxonomy-related counts, fish and shellfish catch statistics.

4.2.7 Characteristics for Challenge 7: Fishery impact

Challenge 7 partners have identified 7 input datasets which correspond to 3 categories of characteristics for 2 environmental matrices as follows:

Marine water-   fish and shellfish catch statistics

Human activities-   fish biomass in water bodies, fishery characterization

4.2.8 Characteristics for Challenge 8: Eutrophication

Challenge 8 partners have identified 45 input datasets which correspond to 10 categories of characteristics and 1 environmental matrix as follows:

Marine water-   temperature of the water column, nitrate concentration parameters in the water column, phosphate concentration parameters in the water column, dissolved total and organic nitrogen concentrations in the water column, dissolved total or organic phosphorus concentration in the water column, dissolved oxygen parameters in the water column, phytoplankton taxonomic surface area in water bodies, phytoplankton generic biomass in water bodies, phytoplankton generic abundance in water bodies, chlorophyll pigment concentrations in water bodies.

4.2.9 Characteristics for Challenge 9: River inputs

Challenge 9 partners have identified 71 input datasets, which correspond to 7 categories of characteristics and 2 environmental matrices as follows:

Fresh water-   dissolved total and organic nitrogen concentrations in the water column, particulate total and organic phosphorus concentrations in the water column,

concentration of suspended particulate material in the water column, temperature of the water column, river flow and discharge;

Marine water-   nitrate concentration parameters in the water column, phosphate concentration parameters in the water column.

4.2.10 Characteristics for Challenge 10: Bathymetry

Challenge 10 partners have identified 38 input datasets, which correspond to 1 category of characteristics and 1 environmental matrix as follows:

Seabed/ Riverbed -   bathymetry and elevation.

4.2.11 Characteristics for Challenge 11: Alien species

Challenge 11 partners have identified 24 input datasets, which correspond to 2 categories of characteristics for 1 environmental matrix as follows:

Biota/Biology-   zooplankton wet weight biomass, zooplankton taxonomy-related abundance per unit volume of the water column.

4.3 Most requested categories of characteristics

The aim of the DAR is to assess the adequacy of the basin scale monitoring system by estimating the fitness for use criteria for the input data sets explained in Section 2. The input data sets previously described have been selected by the project experts to match the Challenge product specifications in terms of necessary input characteristics. For example, for the CH04-Climate temperature and sea level data sets are required to estimate trends and different data sets have been identified, from in situ sensors, satellites, and numerical models.

Table 4.1 presents the 40 different characteristics requested by all the Challenges and Table 4.5 lists the number of different characteristics requested more than once. The number of such different characteristics is 17 out of 40. It is immediately clear that

1. TEMP/temperature of the water column and
2. MBAN/bathymetry and elevation

are the most requested monitoring characteristics among all the Challenges, followed by

3. ASLV/Sea level
4. NTRA/Nitrate concentration parameters in the water column
5. PHOS/Phosphate concentration parameters in the water column
6. TDNT/Dissolved total and organic nitrogen concentrations in the water column
7. RFVL/Horizontal velocity of the water column (currents)

These are the characteristics that require a focused assessment since they cover the needs of at least 3 Challenges.

P02 / Characteristic category	Challenges using the Characteristic
TEMP/Temperature of the water column	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH03 - Oil Platform Leaks CH04 - Climate CH08 - Eutrophication CH09 - River Inputs
ASLV/Sea level	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH05 – Coasts
MBAN/Bathymetry and Elevation	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH03 - Oil Platform Leaks CH10 – Bathymetry
NTRA/Nitrate concentration parameters in the water column	CH09 - River Inputs CH08 - Eutrophication CH09 - River Inputs
PHOS/Phosphate concentration parameters in the water column	CH09 - River Inputs CH08 - Eutrophication CH09 - River Inputs
RVDS/River flow and discharge	CH05 - Coasts CH09 - River Inputs
PSST/Skin temperature of the water column	CH02 - Marine Protected Areas CH04 – Climate
RFVL/Horizontal velocity of the water column (currents)	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH03 - Oil Platform Leaks
EWSB/Wind strength and direction	CH01 - Windfarm Siting CH03 - Oil Platform Leaks
WVST/Wave height and period statistics	CH01 - Windfarm Siting CH03 - Oil Platform Leaks
CPWC/Chlorophyll pigment concentrations in water bodies	CH02 - Marine Protected Areas CH08 – Eutrophication
TDNT/Dissolved total and organic nitrogen concentrations in the water column	CH09 - River Inputs CH08 - Eutrophication CH09 - River Inputs
CNTX/Phytoplankton generic biomass in water bodies	CH04 - Climate CH08 – Eutrophication
PNTX/Phytoplankton generic abundance in water bodies	CH04 - Climate CH08 – Eutrophication
FCST/Fish and shellfish catch statistics	CH07 - Fisheries Impact CH06 - Fisheries Management
GP087/Fishery characterisation	CH02 - Marine Protected Areas CH07 - Fisheries Impact
HBEX/Habitat extent	CH02 - Marine Protected Areas CH03 - Oil Platform Leaks

Table 4.5: Characteristic categories being requested more than once by the Challenges.

Table 4.6 lists the number of input datasets requested for the 40 different characteristic categories. Further analysis of the Characteristics and input data sets is given in Annex 2.

P02	Input data sets
AATX/Phytoplankton taxonomic surface area in water bodies	1
ALAT/Horizontal spatial co-ordinates	2
ASLV/Sea level	54
BRDA/Bird taxonomy-related counts	4
CAPH/Pressure (measured variable)	2
CDTA/Air temperature and density	3
CETA/Cetacean abundance	2
CHUM/Atmospheric humidity	2
CNTX/Phytoplankton generic biomass in water bodies	6
COAS/Terrestrial mapping	1
COGE/Coastal geomorphology	1
CPWC/Chlorophyll pigment concentrations in water bodies	8
CRYS/Snow and ice mass, thickness and extent	17
DBDX/Geological sample density	1
DOXY/Dissolved oxygen parameters in the water column	2
EWSB/Wind strength and direction	8
FABD/Fauna abundance per unit area of the bed	5
FCNT/Fish taxonomy-related counts	2
FCST/Fish and shellfish catch statistics	4
FIBM/Fish biomass in water bodies	2
GP079/Zooplankton wet weight biomass	11
GP087/Fishery characterisation	3
GWDR/Wave direction	1
HBEX/Habitat extent	3
MBAN/Bathymetry and Elevation	46
NTRA/Nitrate concentration parameters in the water column	29
PHOS/Phosphate concentration parameters in the water column	21
PNTX/Phytoplankton generic abundance in water bodies	5
PSAL/Salinity of the water column	2
PSST/Skin temperature of the water column	11
RFVL/Horizontal velocity of the water column (currents)	9
RVDS/River flow and discharge	16
TDNT/Dissolved total and organic nitrogen concentrations in the water column	6
TDPX/Dissolved total or organic phosphorus concentration in the water column	2
TEMP/Temperature of the water column	120
TPHS/Particulate total and organic phosphorus concentrations in the water column	5
TSED/Concentration of suspended particulate material in the water column	12
WVSP/Spectral wave data parameters	2
WVST/Wave height and period statistics	8
ZATX/Zooplankton taxonomy-related abundance per unit volume of the water column	13

Table 4.6: The 40 different characteristic categories and their respective number of input data sets

5. Data Adequacy: availability indicators

The availability indicators were described in Section 2. Here we discuss the results of the statistical analysis of the indicators for the 452 input data sets inserted in the metadatabase and potentially required by the Challenges.

Table 5.1 lists each of the availability indicators together with the meaning of the scores. For each indicator there are three different scores (red, yellow, and green), plus the "unknown" status which highlights a problem in the metadatabase that was not possible to eliminate for the first DAR. The number of unknowns is less than 5% of the total number of datasets, and thus is not believed to affect the validity of the statistical analysis. In general "red" means inadequate, "yellow", partly adequate and "green" fully adequate, as discussed in section 2.5.

The scores were identified in the MedSea Checkpoint First DAR and here they are reconsidered with the same meaning, thereby providing the basis for an across-basin assessment. For each indicator the meaning of the three classes of scores is different and is reported both in Table 5.1 and in all the Tables for each indicator.

Table 5.1 Indicator acronym, short description and preset values subdivided into three scores: red, yellow and green

Name of indicator	Indicator values and meaning
AV-VI-1- Easily found	**"Red":** *"Cited in peer reviewed paper or grey literature but no info on how to access" or "Information retrieved upon specific request to the data source* **"Yellow":** *"Use of social network, community of practices sharing information, Portals of organization where no search is organized by an engine"* **"Green":** *"Use of open search engines, searching by the name of either the data provider or the characteristics" or "Search via a reference catalogue is possible"*
AV-VI-2-INSPIRE Catalogue service	**"Red":** *"Data sets are not referenced in a catalogue or are referenced in a non public catalogue"* **"Yellow":**"The datasets are referenced in a public national catalogue, in an international catalogue service " **"Green":** *"The datasets provide a full EU INSPIRE catalogue service "*
AV-AC-1- Policy visibility	**"Red":** *"There is no information at all on data policy adopted by data providers"* **"Yellow":** *"There is information, but details are available only on request"* **"Green":** *"There is detailed information provided to understand data policy"*
AV-AC-2- Delivery mechanism	**"Red":** *"No information was found on data delivery mechanisms" or "Order form/invoice is requested"* **"Yellow":** *"Online downloading services "* **"Green":** *"Online discovery and downloading services" or "Online discovery + downloading + viewing services (Advanced services)"*

Table 5.1 (cont.)

Name of indicator	Indicator values and meaning
AV-AC-3- Data Policy	**"Red":** " Not or not well documented" **"Yellow":** "Restricted" **"Green":** "Unrestricted"
AV-AC-4- Pricing	**"Red":** "Not or not well documented" or "Commercial cost charge" **"Yellow":** "Distribution charge" or "Collection charge" or "Free of charge for academic institutions and uses" **"Green":** "Open and Free, No charge"
AV-AC-5- Readiness	**"Red":** "Format not or not well documented" or "Proprietary format and not well documented " or "Not proprietary format but content not clearly specified " **"Yellow":** "Proprietary format but content clearly specified " **"Green":** "Not proprietary format and content clearly specified (eg auto-descriptive eg ODV, NetCDF CF) or at least with appropriate documentation"
AV-PE-1- Responsiveness	**"Red":** "No information is found on response time" or "More than 1 week for release" **"Yellow":** "Less or equal to 1 week for release" **"Green":** "Online downloading (i.e. a few hours or less) for release"

In the following sections, we describe the score for each Challenge input data set for each availability indicator and the overall score across the Challenges.

5.1 AV-VI-1: Easily Found

Table 5.2: Scores for the AV-VI-1 "Esily found" indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	Score value	meaning	Challenge 1	2	3	4	5	6	7	8	9	10	11	
AV-VI-1 Easily found	Low visibility	Choice 1: Red												
		"Cited in peer reviewed paper or grey literature but no info on how to access"	8	1	4	8								21
		Choice 2 : Red												
		"Information retrieved upon specific request to the data source "	27	9		18	3							57
	Medium visibility	Choice 3: Yellow												
		"Use of social network, community of practices sharing information, portals of organization where no search is organized by an engine"	5			1					1			7
	Good visibility	Choice 4: Green												
		"Use of open search engines, searching by name either the data provider or the characteristics"		18	1	89	6	3	6	30	51	1	21	226
		Choice 5: Green												
		"Search via reference catalogue (e.g. Copernicus, GEOSS Geoportal…)"	6	7	5		36			15	19	37		125
		Unknown	1	2	9				1				3	16
		Total input data sets	47	37	19	116	45	3	7	45	71	38	24	452

The AV-VI-1 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.2:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
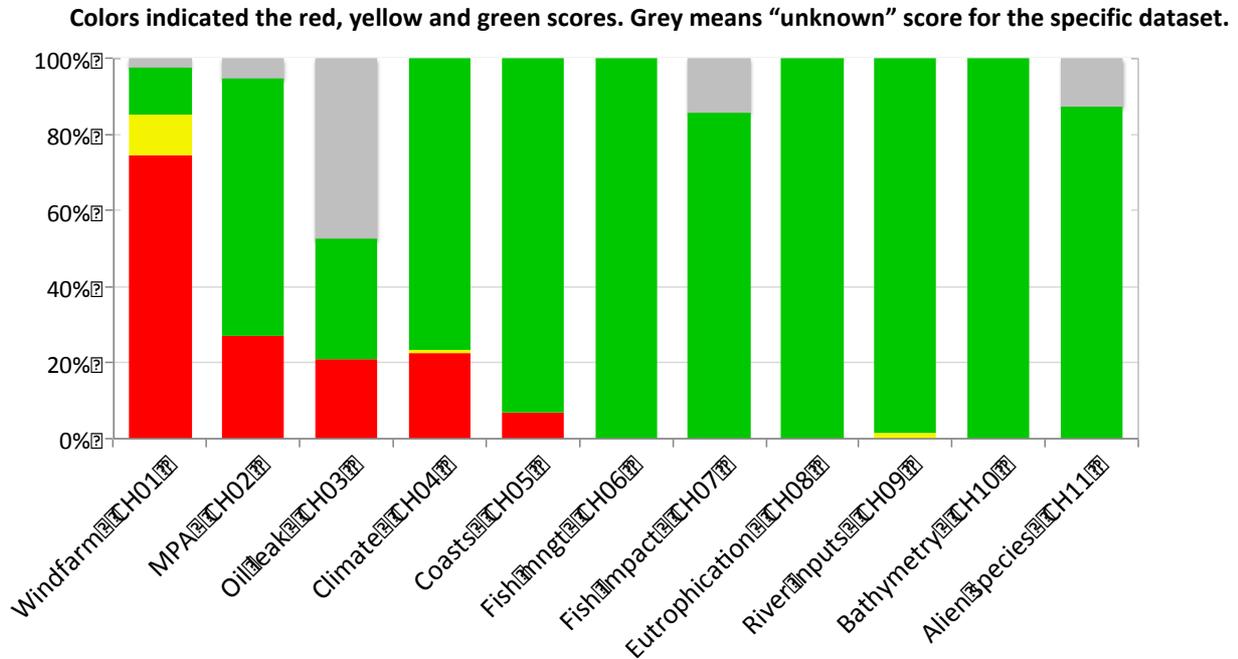
The Easily Found indicator for all challenges is summarised in the following table and figure.

Red	78
Yellow	7
Green	351
Unknown	16
Total	452
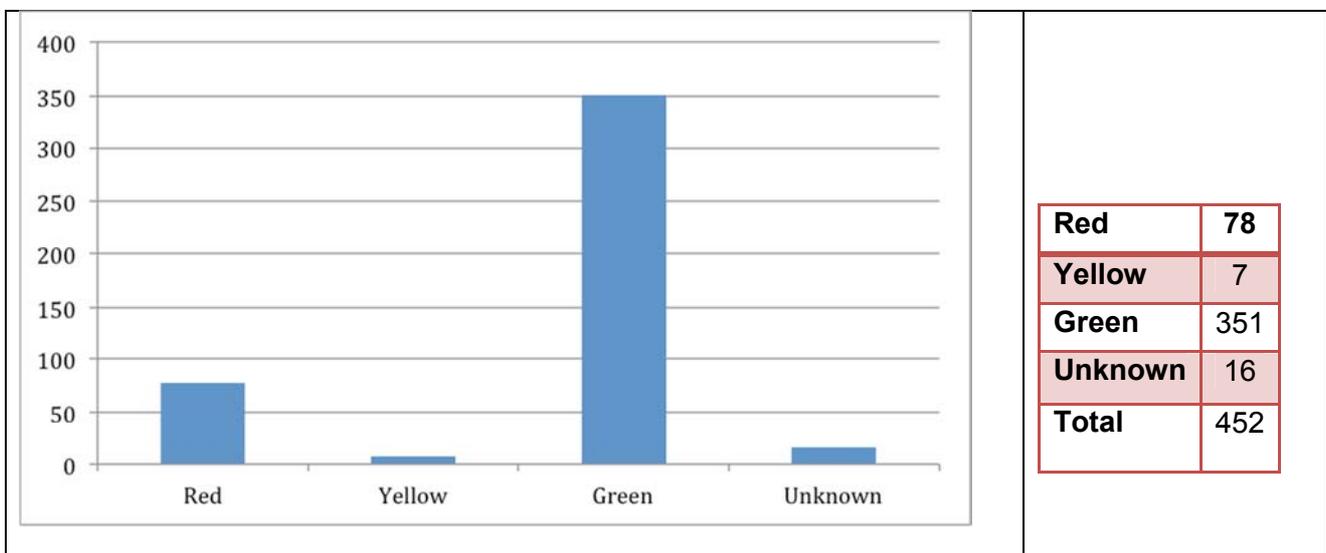
Easily Found indicator across all Challenges

For all challenges, about 78% of characteristics can be found easily. However, there are important differences among the Challenges, the most important of which are for Challenge 1 (Wind Farm Siting) where about 74% of the characteristics has a low visibility.

5.2 AV-VI-2: INSPIRE Catalogue Service

Table 5.3: Scores for the AV-VI-2 INSPIRE Catalogue Service indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	Score value	Meaning	Challenge											
			1	2	3	4	5	6	7	8	9	10	11	
AV-VI-2 INSPIRE catalogue service	Inadequate	Choice 1: Red												
		"Data sets are not referenced in a catalogue or are referenced in a non public catalogue"	34		14	9	3			20	4	1		85
	Partly adequate	Choice 2: Yellow												
		"The datasets are referenced in a public national catalogue, in an international catalogue service "	12	37	4	107	42	3	6	25	67	37	24	364
	totally adequate	Choice 3: Green *"The datasets provide a full EU Inspire catalogue service "*												
		unknown	1		1				1					6
		Total input data sets	47	37	19	116	45	3	7	45	71	38	24	452

The AV-VI-2 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.3:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
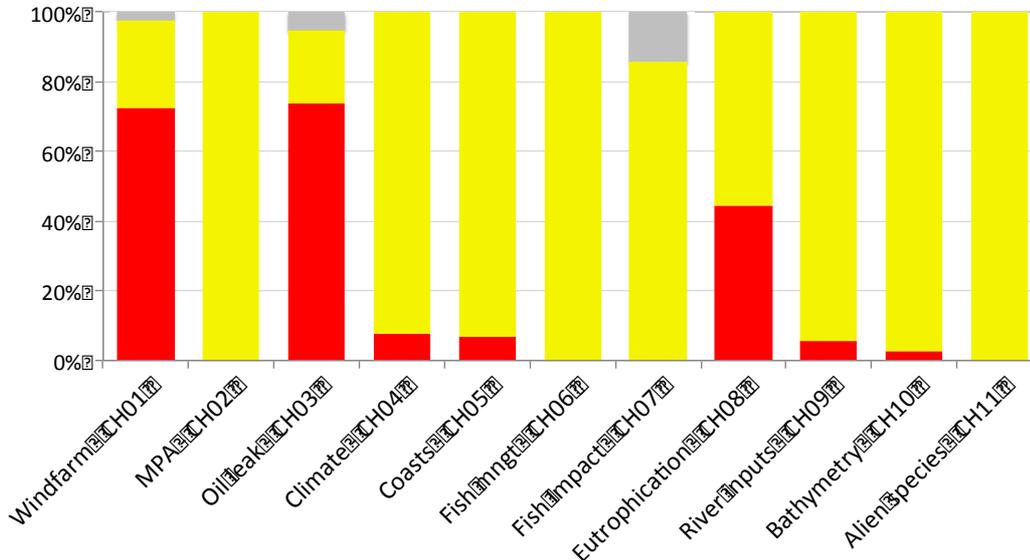
The INSPIRE Catalogue Service indicator for all challenges is summarised in the following table and figure.


Red	85
Yellow	364
Green	0
Unknown	3
Total	452
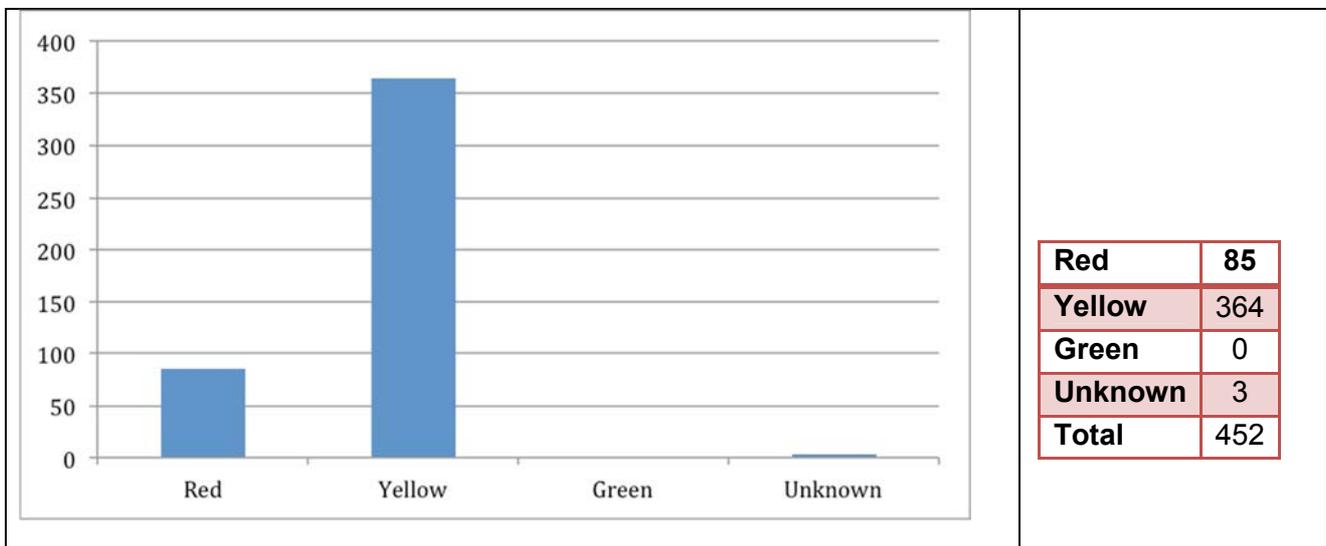
INSPIRE Catalogue Service indicator across all Challenges

For the INSPIRE catalogue indicator the dominant value is "yellow". This means that the input data sets do not have a fully adequate INSPIRE service, i.e., they are referenced in public catalogues but without all the required INSPIRE functionalities, such as discovering the data, viewing and transforming them. The 'red' score is mainly due to Challenges 1, 3 and 8. INSPIRE is the only indicator that has no green values.

5.3 AV-AC-1: Policy visibility

Table 5.4: Scores for the AV-AC-1 Policy visibility indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	Score	Meaning	Challenge 1	2	3	4	5	6	7	8	9	10	11	
AV-AC-1 Policy visibility	Low transparency	Choice 1: Red												
		"There is no information at all on data policy adopted by data providers"	9			24	2			20	19		4	78
	Medium transparency	Choice 2: Yellow												
		"There is information, but details are available only on request"	19	1	9	20	8					37	16	110
	High transparency	Choice 3: Green												
		"There is detailed information provided to understand data policy"	18	36	10	72	35	3	6	25	52	1	4	262
		unknown	1	0	0	0	0	0	1	0	0	0	0	5
		Total input data sets	47	37	19	116	45	3	7	45	71	38	24	452

The AV-AC-1 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.4:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
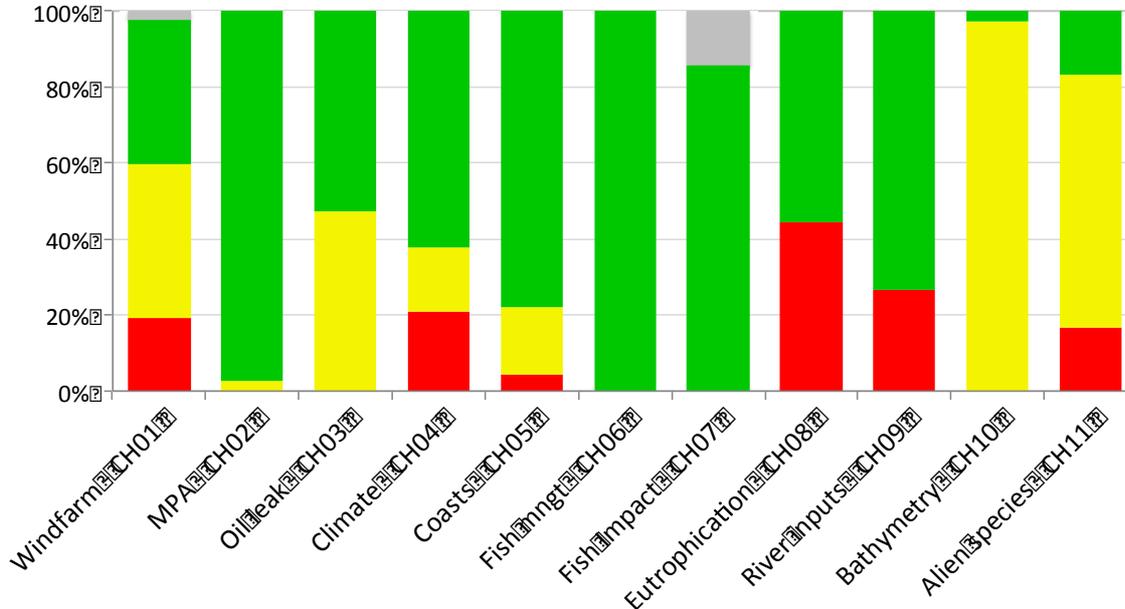
The Policy Visibility indicator for all challenges is summarised in the following table and figure.

Red	78
Yellow	110
Green	262
Unknown	2
Total	452
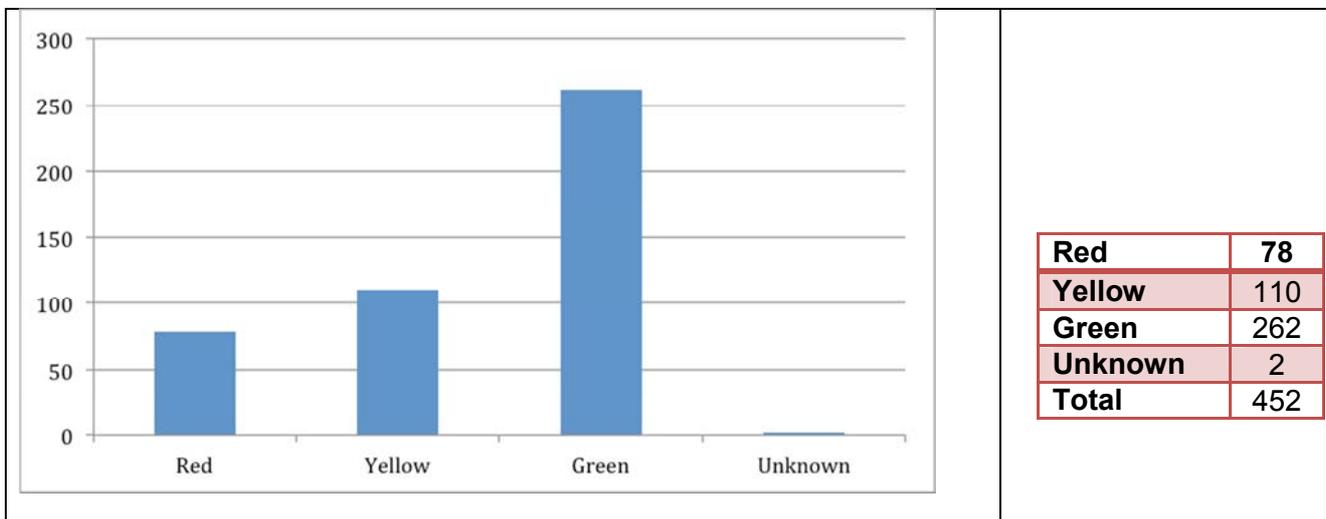
Policy visibility indicator across all Challenges

Policy Visibility is generally good, with the exceptions of Challenge 1 (Wind Farm Siting) where 'yellow' is of the same order as 'green' and Challenge 11 (Alien species).

5.4 AV-AC-2: Delivery Mechanism

Table 5.5: Scores for the AV-AC-2 Delivery mechanism indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	Score	Meaning	Challenge											
			1	2	3	4	5	6	7	8	9	10	11	
	No information	Choice 1: Red												
		"No information was found on data delivery mechanisms"	8			22	1						16	47
	Manual	Choice 2: Red												
		"Order form/invoice is requested"	17			9	6							32
	Partial Inspire function	Choice 3: Yellow												
AV-AC-2 Delivery mechanism		*"Online downloading services "*		19	11	5	3				2	1	2	43
	Full Inspire function	Choice 4: Green												
		"Online discovery and downloading services"	6	12	1	40	35		5	28	45	37	6	215
		Choice 5: Green												
		"Online discovery + downloading + viewing services(Advanced services)"	12	6	7	31		3	1	17	24			101
		unknown	4			9			1					21
		Total input data sets	47	47	19	116	45	3	7	45	71	38	24	452

![EMODnet logo – European Marine Observation and Data Network]

**Sea Basin Checkpoint
Lot 4: Black Sea**

D 15.2
Version: V10
Date: 15/12/2016

The AV-AC-2 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.5:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
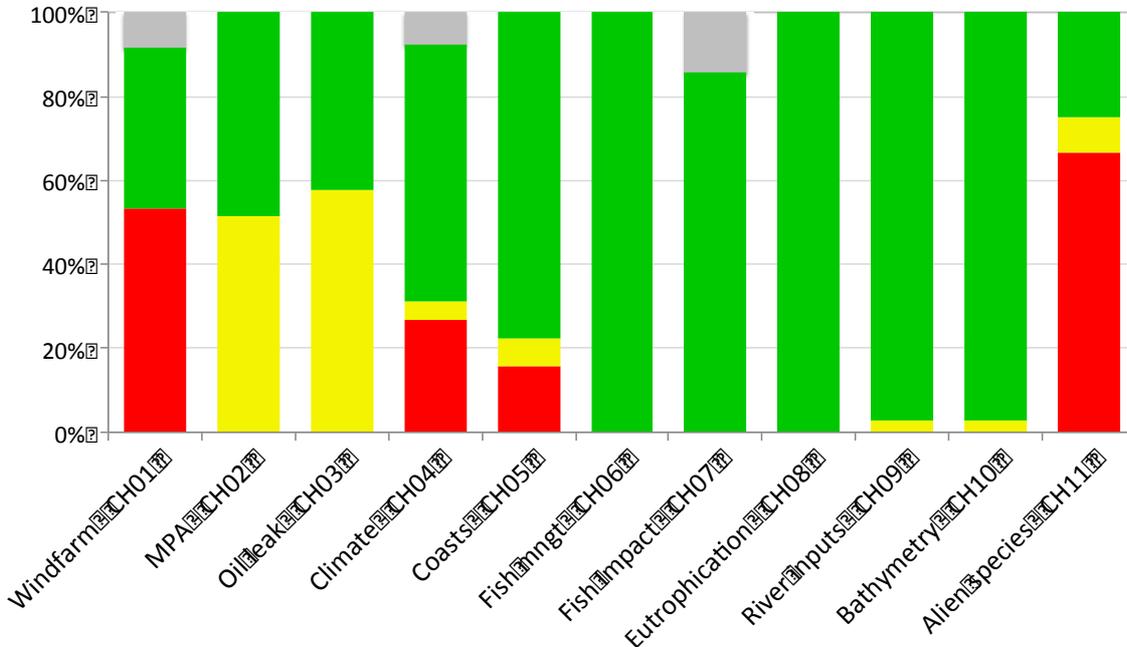
The Delivery Mechanism indicator for all challenges is summarised in the following table and figure.

Red	79
Yellow	43
Green	316
Unknown	14
Total	452
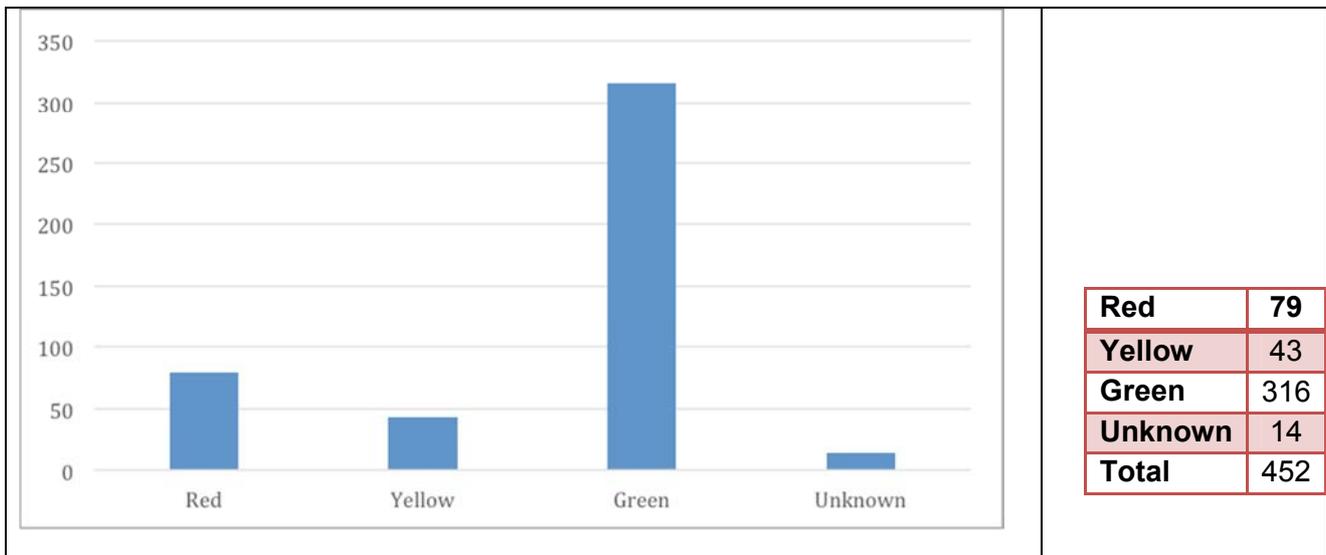
Delivery mechanism indicator across all Challenges

The dominant score in the delivery mechanism is "green" meaning that online downloading is possible, with or without additional functionalities. However, Challenges 2 (MPAs) and 3 (Oil Leaks) have more 'yellow' characteristics than 'green'. Challenge 4 has an equal amount of the three scores and Challenge 11 has the majority of characteristics in 'red'.

D 15.2
Version: V10
Date: 15/12/2016

Sea Basin Checkpoint
Lot 4: Black Sea

EMODnet
European Marine
Observation and
Data Network

5.5 AV-AC-3: Data Policy

Table 5.6: Scores for the AV-AC-3 Data Policy indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	score	Meaning	Challenge											
			1	2	3	4	5	6	7	8	9	10	11	
AV-AC-3 Data policy	Not documented	Choice 1:Red												
		"Not or not well documented"		1		11	2					37		51
	Restricted	Choice 2: Red												
		"Restricted"	1	18		27	5			8	15		16	90
	Partially restricted	Choice 3: Yellow												
		"Accessible under moratorium"	27		11	17	3			2	2		4	66
	Unrestricted	Choice 4: Green												
		"Unrestricted"	17	21	8	61	34	3	6	35	54	1	4	244
			.											
		unknown	1			1			1					3
		Total input data sets	47	37	19	116	45	3	7	45	71	38	24	452

EMODnet
European Marine
Observation and
Data Network

**Sea Basin Checkpoint
Lot 4: Black Sea**

D 15.2
Version: V10
Date: 15/12/2016

The AV-AC-3 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.6:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
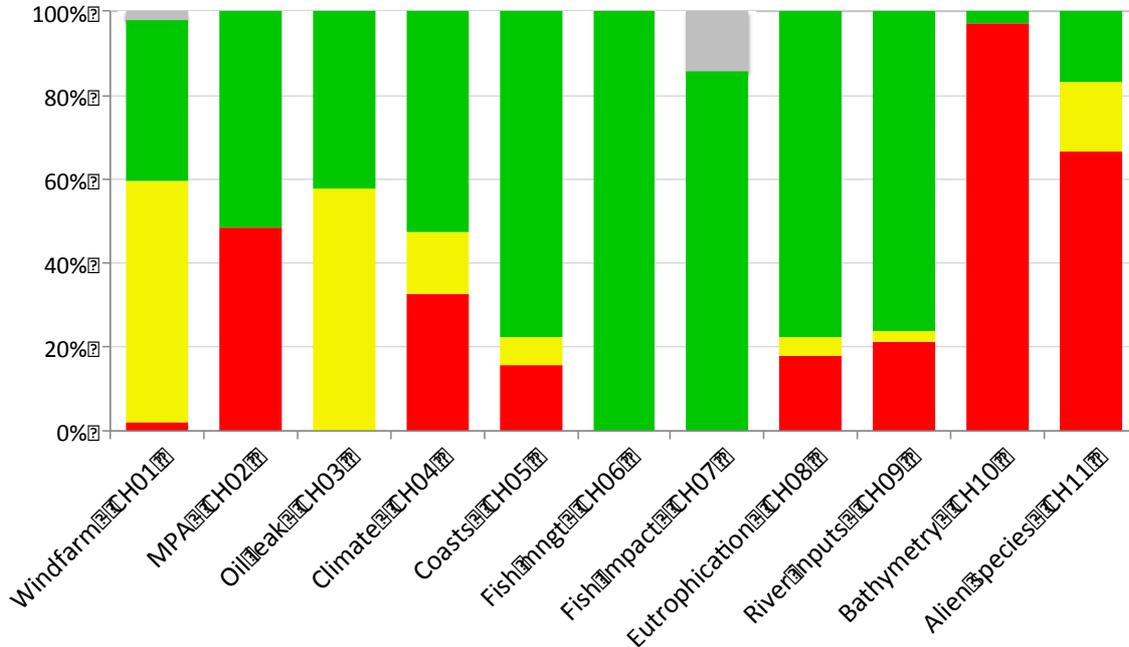
The Data Policy indicator for all challenges is summarised in the following table and figure.

Red	**140**
Yellow	66
Green	244
Unknown	2
Total	452
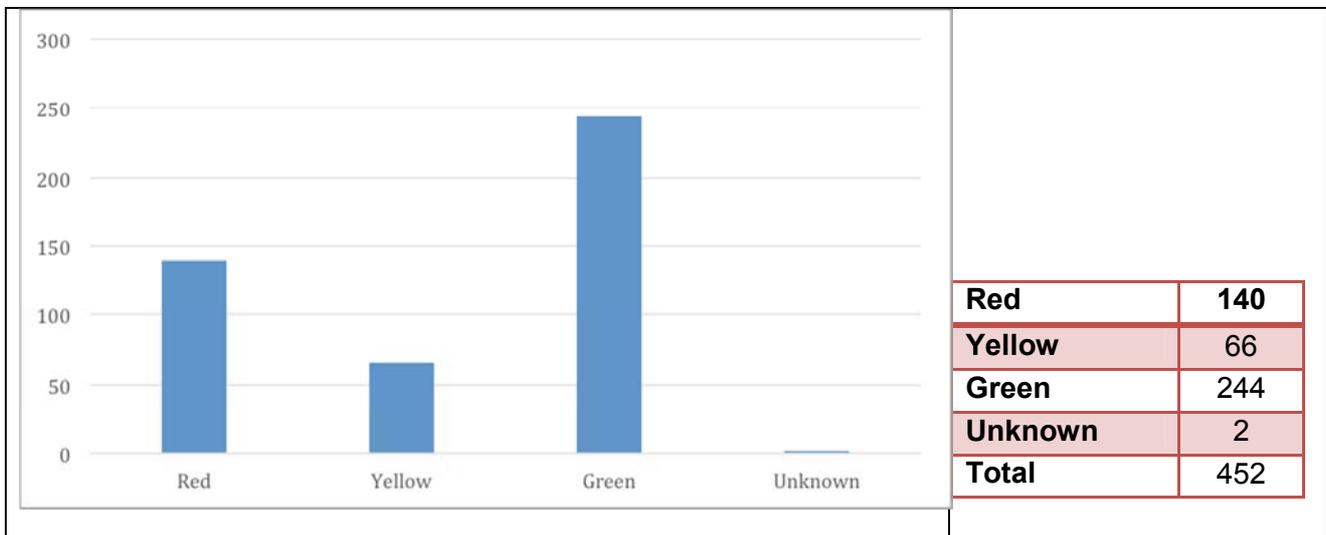
Data policy indicator across all Challenges

Data policy is an issue in the Black Sea, showing a significant proportion of the input data sets as "restricted" or "not well documented". There are a significant number of 'red' scores in Challenges 2 (MPA) and 4 (Climate), while for Challenge 10 (Bathymetry) almost all the data sets are restricted.

5.6 AV-AC-4: Pricing

Table 5.7: Scores for the AV-AC-4 Pricing indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	score	Meaning	Challenge											
			1	2	3	4	5	6	7	8	9	10	11	
AV-AC-4 Pricing	Not documented	Choice 1: Red												
		"Not or not well documented"		4		41	35			9	15	37	18	159
	Commercial cost Charge	Choice 2: Red												
		"Commercial cost charge"					2					1		3
	Distribution limitations	Choice 3: Yellow "Distribution charge"												
		"Collection charge"												
		"Free of charge for academic institutions and uses"	19	4		1	5						2	31
	Free	Choice 4: Green												
		"Open and Free, No charge"	26	29	19	74	3	3	6	36	56		2	252
		unknown	2						1				2	8
		Total input data sets	47	37	19	116	45	3	7	45	71	38	24	452

European Marine Observation and Data Network

Sea Basin Checkpoint
Lot 4: Black Sea

D 15.2
Version: V10
Date: 15/12/2016

The AV-AC-4 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.7:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
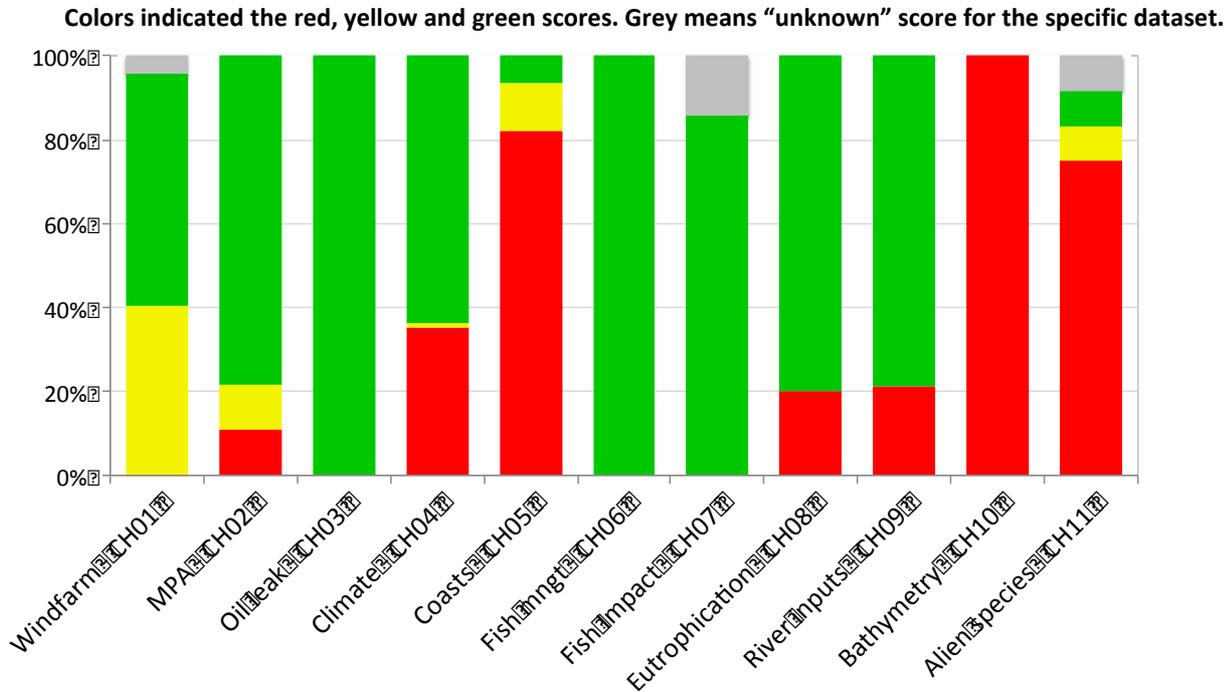
The Pricing indicator for all challenges is summarised in the following table and figure.

Red	**162**
Yellow	31
Green	254
Unknown	5
	452
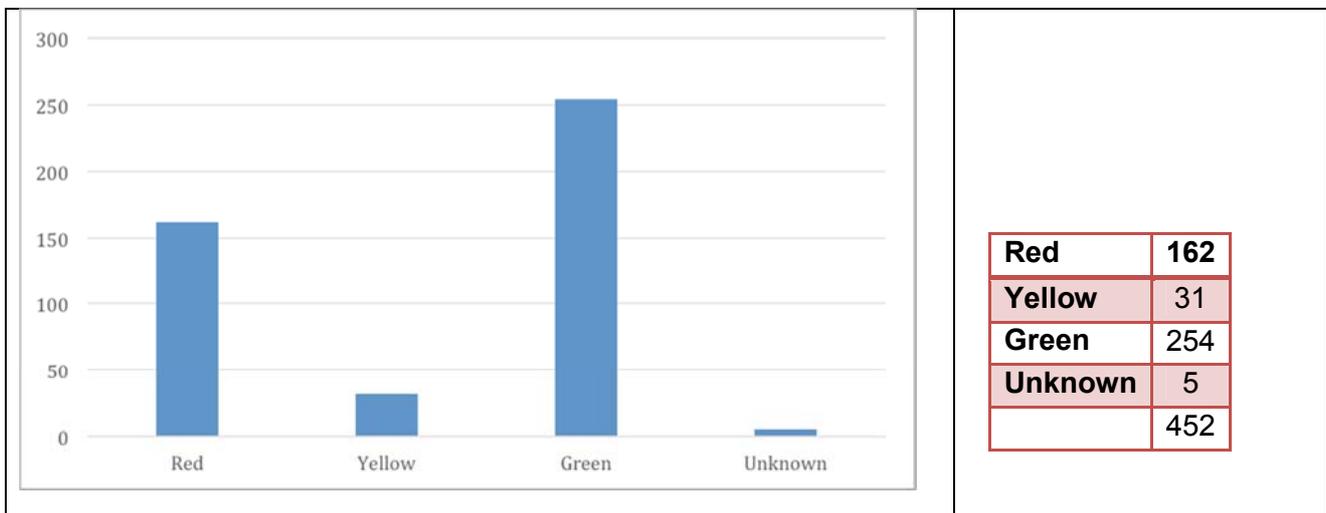
Pricing indicator across all Challenges

Pricing is an indicator with dominant dipolar values across the Challenges. The 'red' score is essentially due to Challenges 5, 11 and 4. For the other challenges the indicator is positive.

5.7 AV-AC-5: Readiness

Table 5.8: Scores for the AV-AC-5 Readiness indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	score	Meaning	Challenge											
			1	2	3	4	5	6	7	8	9	10	11	
AV-AC-5 Readiness	Not documented	Choice 1: Red												
		"Format not or not well documented"	8			46	3						12	69
	Not ready to be consumed	Choice 2: Red												
		"Proprietary format and not well documented "												
		Choice 3: Red												
		"Not proprietary format but content not clearly specified"				14					1			15
	Can be processed to be consumed	Choice 4: Yellow												
		"Proprietary format but content clearly specified "		4		1	7	3	2			37	2	56
	Ready to be consumed	Choice 5: Green												
		"Not proprietary format and content clearly specified (eg auto-descriptive eg ODV, NetCDF CF) or at least with appropriate documentation	38	32	15	53	35		4	45	70		10	302
		unknown	1	1	4	2			1		1			13
		Total input data sets	47	37	19	116	45	3	7	45	71	38	24	452

The AV-AC-5 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.8:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
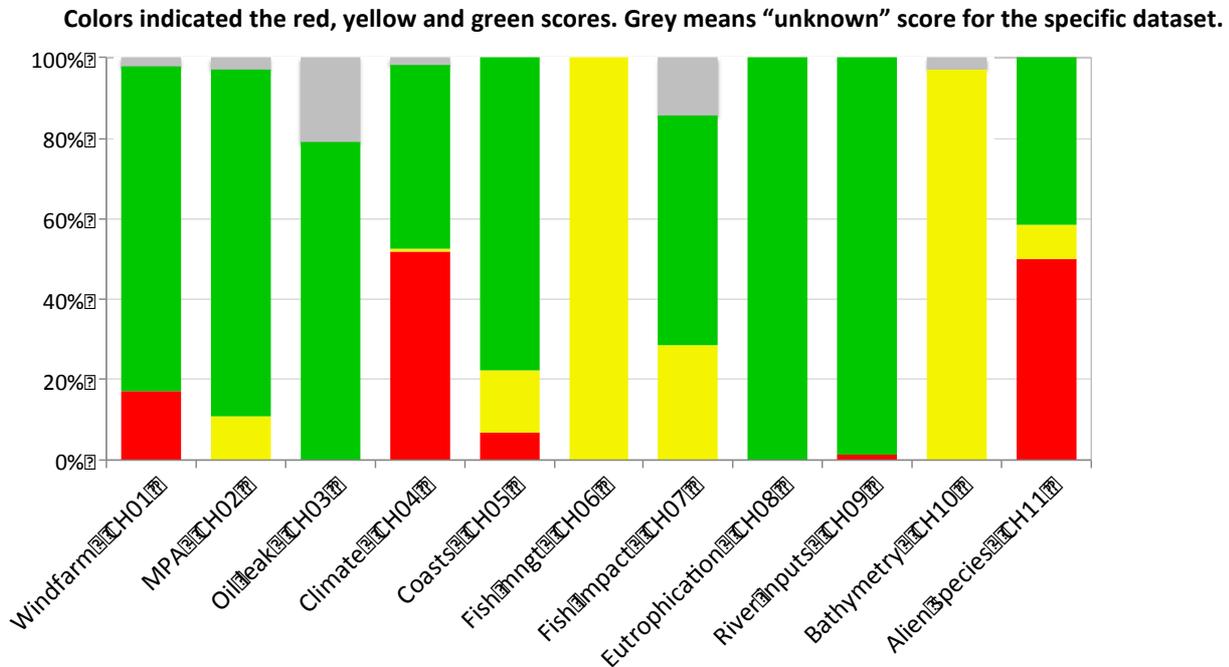
The Readiness indicator for all challenges is summarised in the following table and figure.

Red	84
Yellow	56
Green	302
Unknown	10
Total	452
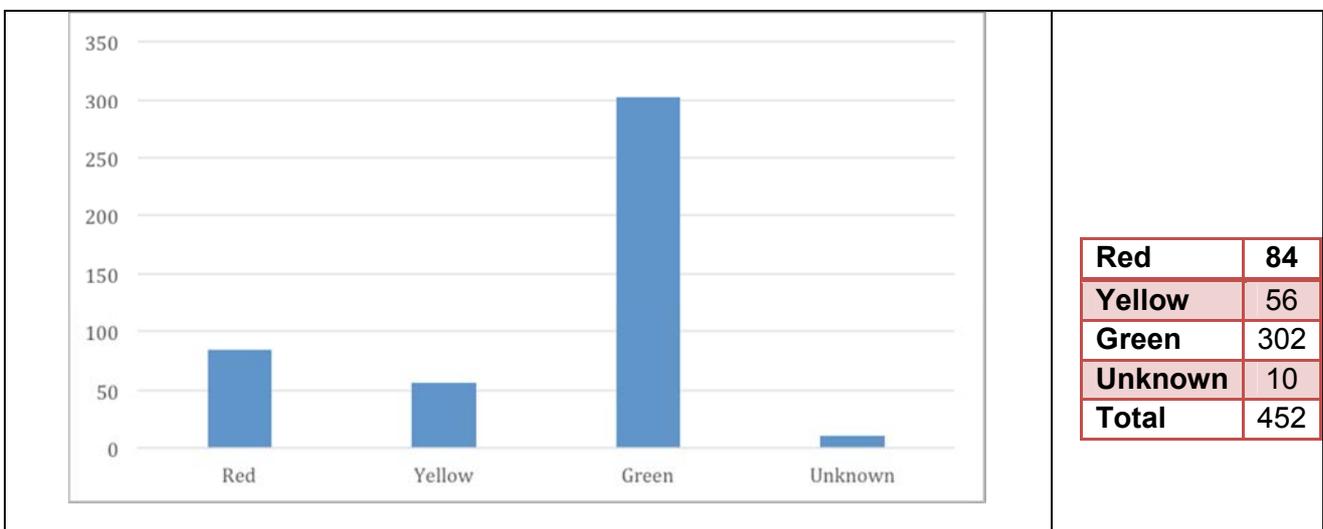
Readiness indicator across all Challenges

From the readiness indicator across all Challenges, it is clear that the input data sets are available for use, with the exception of Challenges 4 (climate) and 11 (alien species) which have half of their data sets with a 'red' score. Challenge 10 (bathymetry) is instead predominantly composed of input data sets with 'yellow' scores, i.e. their content is not adequately explained.

5.8 AV-PE-1: Responsiveness

Table 5.9: Scores for the AV-PE-1 Responsiveness indicator as a function of Challenges for all input data sets. The last column indicates the score across all challenges

Indicator name	score	Meaning	Challenge 1	2	3	4	5	6	7	8	9	10	11	
AV-PE-1 Responsiveness	Low response	Choice 1: Red *"No information is found on response time"*	8	1		44	7			29	18		20	127
		Choice 2: Red												
		"More than 1 week for release"	1				1							2
	Medium response	Choice 3: Yellow												
		"Less or equal to 1 week for release"	19				1						2	22
	High response	Choice 4: Green *"Online downloading (i.e. a few hours or less) for release"*	18	36	19	71	36	3	6	16	53		2	260
		Unkown	1			1			1			38		41
		Total input data sets	47	37	19	116	45	3	7	45	71	38	24	452

The AV-PE-1 indicator scores for each Challenge are visualized as a function of the percentage of input data sets with the specific score, i.e., a graphic visualization of the contents of Table 5.9:

Colors indicated the red, yellow and green scores. Grey means "unknown" score for the specific dataset.
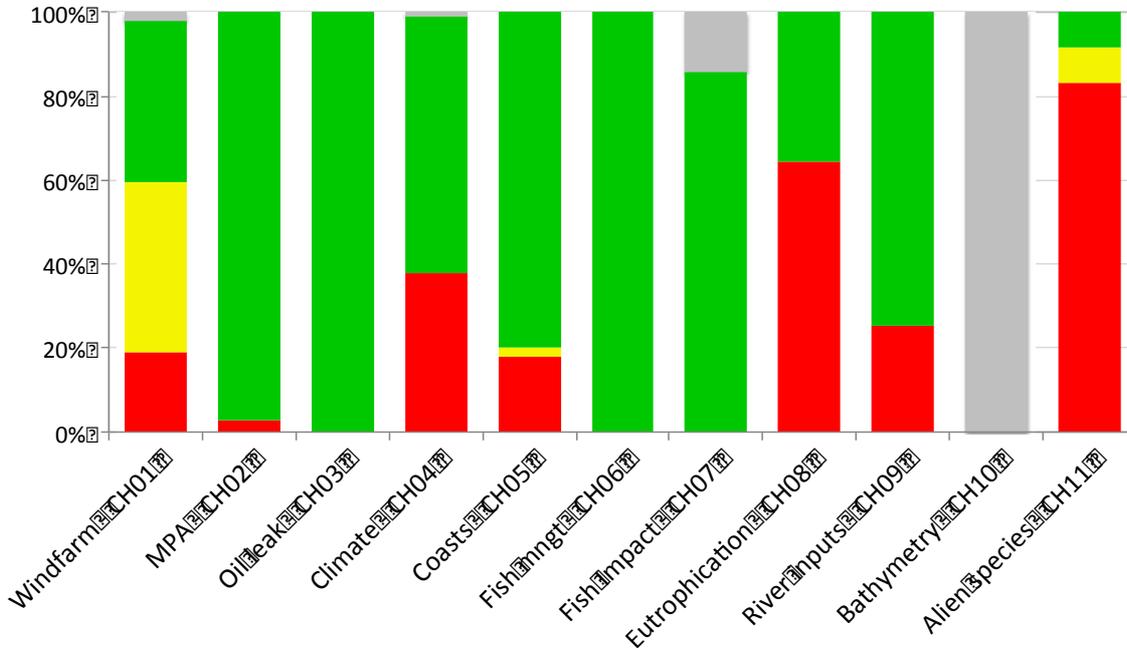
The Responsiveness indicators for all challenges are summarised in the following Table and figure.

Red	**129**
Yellow	22
Green	260
Unknown	41
Total	452
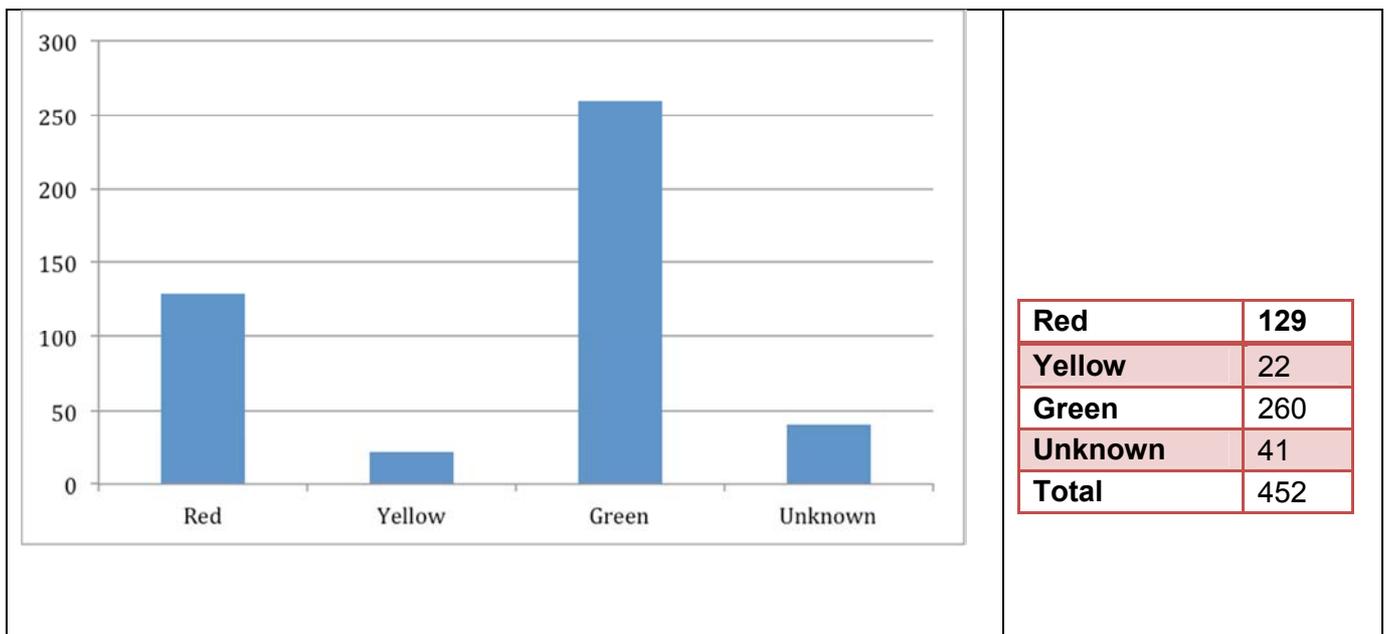
Responsiveness indicator across all Challenges

Responsiveness is another bipolar indicator: either the service has on-line downloading (green) or it has a downloading response time greater than a week. Challenge 6 and 7 (Fishery and Management) data sets are completely "green" while all the others have almost an equal amount of potentially useful input data sets with low and optimal quality downloading.

5.9 Summary of the results for availability indicators

The above indicators will now be described per Challenge and per Characteristics.

Starting with the Challenges, results are shown synthetically in Fig. 5.1. It is evident that the Challenge 06-Fisheries and Challenge 07-Fishery impact have the best scores, probably due to the fact that source of data is well defined and available. The worse cases are for Challenge 01-wind farm siting, Challenge04-Climate, Challenge10-Bathymetry and Challenge11-Alien Species where the red and green values are in equal proportion or red-yellow is dominant.

Fig. 5.1 The eight availability indicators for each Challenge

P02 characteristics category	AV-VI-1 Easily found	AV-VI-2 INSPIRE catalogue service	AV-AC-1 Visibility of Data policy	AC-AC-3 Data Policy	AV-AC-4 Pricing	AV-AC-2 Data delivery mechanisms	AV-PE-1 Responsiveness	AV-AC-5 Readiness	Comment
AATX/Phytoplankton taxonomic surface area in water bodies									
ALAT/Horizontal spatial co-ordinates									
ASLV/Sea level									
BRDA/Bird taxonomy-related counts									
CAPH/Pressure (measured variable)									
CDTA/Air temperature and density									
CETA/Cetacean abundance									
CHUM/Atmospheric humidity									
CNTX/Phytoplankton generic biomass in water bodies									
COAS/Terrestrial mapping									
COGE/Coastal geomorphology									
CPWC/Chlorophyll pigment concentrations in water bodies									
CRYS/Snow and ice mass, thickness and extent									Obs.
CRYS/Snow and ice mass, thickness and extent									Models
DBDX/Geological sample density									
DOXY/Dissolved oxygen parameters in the water column									
EWSB/Wind strength and direction									CH 1
EWSB/Wind strength and direction									CH 3
FABD/Fauna abundance per unit area of the bed									
FCNT/Fish taxonomy-related counts									
FCST/Fish and shellfish catch statistics									
FIBM/Fish biomass in water bodies									
GP079/Zooplankton wet weight biomass									
GP087/Fishery characterisation									
GWDR/Wave direction									
HBEX/Habitat extent									
MBAN/Bathymetry and Elevation									CH 10
MBAN/Bathymetry and Elevation									CH 2
NTRA/Nitrate concentration parameters in the water column									CH 8
NTRA/Nitrate concentration parameters in the water column									CH 9
PHOS/Phosphate concentration parameters in the water column									CH 8
PHOS/Phosphate concentration parameters in the water column									CH 9
PNTX/Phytoplankton generic abundance in water bodies									
PSAL/Salinity of the water column									
PSST/Skin temperature of the water column									
RFVL/Horizontal velocity of the water column (currents)									CH 1
RFVL/Horizontal velocity of the water column (currents)									CH3 mod.
RFVL/Horizontal velocity of the water column (currents)									CH3 obs.
RVDS/River flow and discharge									
TDNT/Dissolved total and organic nitrogen concentrations in the water column									
TDPX/Dissolved total or organic phosphorus concentration in the water column									
TEMP/Temperature of the water column									Obs.
TEMP/Temperature of the water column									Models
TPHS/Particulate total and organic phosphorus concentrations in the water column									
TSED/Concentration of suspended particulate material in the water column									
WVSP/Spectral wave data parameters									
WVST/Wave height and period statistics									
ZATX/Zooplankton taxonomy-related abundance per unit volume of the water column									

Table 5.10 Characteristics categories and availability indicators across Challenges

Furthermore the indicators were summarized for each P02 characteristic and for the Challenges. Table 5.10 shows the 40 different characteristics categories together with their scores across all Challenges. The number of P02 characteristics in the Table 5.10 is different from 40 (Table 4.1) because the Challenges refer to different data sets for the same characteristics. This highlights a gap in the sharing of information across disciplines in the marine community. While this could generate low quality products in one Challenge that uses multidisciplinary data sets, it might also be an indication of a duplication of efforts in the monitoring system.

In Table 5.10 the indicator value for each characteristics is based upon more than one data set. The value reported in the Table is the most frequent one among the different data sets that refers to the specific characteristic.

Table 5.10 clearly shows that most "red" and "yellow" scores are related to the AV-VI and AV-AC-1 indicators. This highlights the lack of an adequate data management infrastructure at the Black Sea basin scale level.

The outstanding specific characteristic category that are not adequately covered by input data set availability adequacy indicators are:

1) TSED/Concentration of suspended particulate material in the water column

2) CRYS/Snow and Ice mass, thickness and extent

3) CAPH/pressure (measured variable)

4) CDTA/Air temperature and density

5) CHUM/Atmospheric humidity

6) CNTX/Phytoplankton generic biomass in water bodies

7) GP079/Zooplankton wet weight biomass

There are also other characteristics whose sum of "yellow" and "red" is greater than the "green" scores and they are:

8) EWSB/wind strength and direction

9) MBAN/Bathymetry and elevation

10) PSAL/salinity in the water column

11) WVSP/Spectral wave data parameters

12) WVST/Wave height and period statistics

In addition, whenever the characteristic category is monitored by "observations" instead of "model" outputs, the availability indicators for the data sets are less adequate, highlighting the lack of adequate availability of the observational data sets with respect to the model ones.

An interesting result of this analysis is that input data sets regarding atmospheric characteristics are not adequate to meet the needs of the Challenges, which is probably due to the lack of an open data policy at the level of meteorological organizations.

6. Data Adequacy: EMODnet and Copernicus evaluation

EMODnet and Marine Copernicus are important providers of data and products for the Challenges. One of their objectives is to support the EU blue growth strategy and thus data availability has been at the center of their development.

6.1 CMEMS assessment

At present, the CMEMS catalogue lists the following products for the Black Sea:

- SLA [res. 7km; 1 level (surface)]   L3 along swath mono sensors
- SLA [res. 1/8°; 1 level (surface)]   L4 along swath mono sensors
- CHL [res. 1km; 1 level (surface)]   L3 gridded merged sensors
- CHL [res. 1km, 4km; 1 level (surface)]   L4 gridded merged sensors
- OPT [res. 1km; 1 level (surface)]   L3 gridded merged sensors
- SST [0.04°- 4km; 1 level (surface)]   L4
- T S   In Situ   L2 discrete mono platforms

Legend:
SLA = Sea Level Anomalies
CHL = Chlorophyll
SST = Sea Surface Temperature
T S = Temperature – Salinity
OPT = Optical properties of the water column

The CMEMS availability assessment will then be carried out for these Black Sea input data sets, used by several Challenges.

6.1.1 AV-VI-1: Easily Found.

The indicator comments on the visibility of data in the catalogues and all CMEMS products are "Searchable via reference catalogue (e.g. Copernicus, GEOSS Geoportal…)". Using the previous color scale, the indicator is 'GREEN'.

6.1.2 AV-VI-2: EU INSPIRE catalogue service

The characteristics can be searched by a catalogue service and CMEMS was defined, by default, as one of the ISO and INSPIRE catalogues compliant with OGC services. CMEMS is thus classified as "The datasets providing a full INSPIRE catalogue service " and the indicator is 'GREEN'.

6.1.3 AV-AC-1: Policy visibility

The indicator comments on the information given on the data policy adopted by data providers. CMEMS provides clear information on services and the adopted policy, by providing the initial information on the web pages 'About Us – About Your Copernicus Marine Service' together with a very detailed service policy in 'Service Portfolio – Service Commitments and Licence'. Thus CMEMS is classified as " Detailed information is provided to clarify data policy" and the indicator is again 'GREEN'.

6.1.4 AV-AC-2: Delivery mechanism

The indicator reports the type of services available to the user to access data. CMEMS offers advanced services with "Online discovery + downloading + viewing services" and the indicator is 'GREEN'.

6.1.5 AV-AC-3: Data policy

The indicator comments on the status of information provided for data policy. The information is provided on the same web pages as 'Policy visibility', products are 'Unrestricted', although a password is requested. On the basis of the classification adopted, the CMEMS data policy indicator is 'GREEN'.

6.1.6 AV-AC-4: Pricing

The indicator comments on the provision of information related to cost basis. In the web pages 'About Us – About Your Copernicus Marine Service' it is stated that the access to products is 'open and free'. The user must be registered to access the products, and from the above definitions the CMEMS products are "Open and Free. No charge". The indicator is 'GREEN'.

6.1.7 AV-AC-5: Readiness

The indicator comments on the information provided for distribution format. The CMEMS web pages do not clearly indicate the format of the products. For this information the users must go in the 'Services portfolio – access to products' and then select a product. In the 'Overview' there is a description of the product, however to find the format, the user has to click on 'Full Overview'. In conclusion the format is "Not proprietary and content clearly specified (eg auto-descriptive eg ODV, NetCDF CF) or at least with appropriate document" and thus the indicator is 'GREEN' on the basis of this definition. However, the number of clicks for new users is quite high and it could take a long time to understand what to do. The simplest solution for a new user is to send a message to the Service Desk.

6.1.8 AV-PE-1: Responsiveness

The indicator comments on the information provided for the timeliness or ability to process a request in a reproducible and acceptable amount of time and for CMEMS, there is an "Online downloading (i.e. a few hours or less) for release" and thus the indicator is 'GREEN'.

6.1.9 Summary of CMEMS availability

AV-VI-1 Easily found	🟢
AV-VI-2 EU Inspire catalogue service	🟢

AV-AC-1 Policy visibility	🟢
AV-AC-2 Delivery	🟢

AV-AC-3 Data Policy	🟢	
AV-AC-4 Pricing	🟢	
AV-AC-5 Readiness	🟢	But with some difficulties to find the information about formats in the web pages

AV-PE-1 Responsiveness	🟢	

6.2 EMODnet assessment

There are 8 thematic portals.

- Bathymetry: Data on bathymetry (water depth), coastlines, and geographical location of underwater features: wrecks.
- Geology: Data on seabed substrate, sea-floor geology, coastal behaviour, geological events, and minerals.
- Seabed habitats: Data on modelled seabed habitats based on seadbed substrate, energy , biological zone, and salinity.
- Chemistry: Data on the concentration of nutrients, organic matter, pesticides, heavy metals, radionuclides and antifoulants in water, sediment and biota.
- Biology: Data on temporal and spatial distribution of species abundance and biomass from several taxa.
- Physics: Data on salinity, temperature, waves, currents, sea-level, light attenuation, and FerryBoxes.
- Human activities: Data on the intensity and spatial extent of human activities at sea.
- Coastal mapping: Building a Joint European Coastal Mapping Programme.

The last portal was not evaluated since 'It will soon provide services for discovery and access to coastlines, baselines, bathymetric data, aerial photography and additional layers from EMODNet Bathymetry, Geology and Habitats portals.'
Note that the indicators are constructed using the portal metadata. The assessment is carried out as a non expert user, who needs to be able to get information in an easy way.

6.2.1 AV-VI-1: Easily Found.

- EMODnet Bathymetry: provides bathymetry and very good information on QA/QC. The data is visible and simply "Searchable via reference catalogue (e.g. Copernicus, GEOSS Geoportal…)". This indicator is 'GREEN'.
- EMODnet Geology: good information on data and products, also in this case the indicator is 'GREEN'
- EMODnet Seabed habitats: interactive maps and metadata catalogue are visible, 'GREEN'.
- EMODnet Chemistry: data access service is based on SeaDataNet and the catalogue is easily visible with a couple of clicks. 'GREEN'.
- EMODnet biology: the catalogue is accessible from the home page. 'GREEN'
- EMODnet physics: data are presented on the web pages and a catalogue based on Sextant is also available. All contributors are acknowledged in a visible way. 'GREEN'.

- EMODnet human activities: the activity started quite recently, however a list of products and a search page have been constructed. 'GREEN'

6.2.2 AV-VI-2: INSPIRE catalogue service

- EMODnet Bathymetry: the catalogue in this case is very simple and is ISO and INSPIRE compliant with OGC services. It can thus be classified with "The datasets providing a full EU INSPIRE catalogue service " and is 'GREEN'.
- EMODnet Geology: the catalogue is well done and the indicator is 'GREEN'.
- EMODnet Seabed habitats: the catalogue is the ICES one build using Geonetwork which uses all the ISO and INSPIRE elements. 'GREEN'.
- EMODnet Chemistry: also in this case the underlying SeaDataNet and Sextant assure a good catalogue service. 'GREEN'.
- EMODnet biology: the indicator is 'GREEN'.
- EMODnet Physics: the whole portal is INSPIRE compliant. 'GREEN'.
- EMODnet human activities: the service is under construction, but INSPIRE compliance is ensured. 'GREEN'.

6.2.3 AV-AC-1: Policy visibility

- EMODnet Bathymetry: policy is described on 'Data products' web page and is again 'GREEN'.
- EMODnet Geology: this is one of the few cases where "There is information, but details are available only on request". This is due to the fact that the data reside in different servers and could have different management systems. For this reason the indicator is 'YELLOW'.
- EMODnet Seabed habitats: policy is visible on the first web page. 'GREEN'
- EMODnet Chemistry: the information is not directly provided in the portal, but the user has to go to SeaDataNet. The indicator is 'YELLOW' ("There is information, but details are available only on request")' since the details can be found in another portal.
- EMODnet biology: the policy is described on the home page and the indicator is 'GREEN'
- EMODnet Physics: the data policy is based on Copernicus, EuroGOOS and SeadataNet conditions. 'GREEN'
- EMODnet human activities: the policy is described on the 'About' page. 'GREEN'.

6.2.4 AV-AC-2: Delivery mechanism

- EMODnet Bathymetry: provides an advanced services with "Online discovery + downloading + viewing services" and is 'GREEN'.
- EMODnet Geology: provides an advanced services with "Online discovery + downloading + viewing services" and is 'GREEN'.
- EMODnet Seabed habitats: also in this case there is an advanced service. 'GREEN'
- EMODnet Chemistry: provides an advanced service. 'GREEN'.
- EMODnet biology: the data are downloadable with an advanced service and the indicator is 'GREEN'
- EMODnet Physics: provides an advanced service. 'GREEN'.
- EMODnet human activities: the mechanism is well stated as 'web services' - 'GREEN'.

6.2.5 AV-AC-3: Data policy

- EMODnet Bathymetry: The information is provided on the same web pages as 'Policy visibility'. On the basis of the classification adopted, the data policy is 'GREEN'.
- EMODnet Geology: the portal provides a service (VMS) and according to the 'Terms of use', the data are accessible under moratorium. The indicator is 'YELLOW'.
- EMODnet Seabed habitats: the policy is clearly visible on the first web page. 'GREEN'
- EMODnet Chemistry: is based on the SeaDataNet data policy. Data policy may vary and the web interface enables data to be selected on the basis of their policy. Overall, the indicator is 'GREEN' but the issues regarding the partners' policy are still open within the SeaDataNet community (although the partners are strongly encouraged to make data open and free).
- EMODnet biology: the indicator is 'GREEN'
- EMODnet Physics: states that the data policy is based on Copernicus, EuroGOOS and SeadataNet conditions. 'GREEN'
- EMODnet human activities: the information is provided in 'About', according to which there are "data with as few restrictions as possible". Thus the data are 'Partially restricted' – 'YELLOW'.

6.2.6 AV-AC-4: Pricing

- EMODnet Bathymetry: The user must be registered to access data, and from the above definitions the products are "Open and Free. No charge", the indicator is 'GREEN'.
- EMODnet Geology: again from the 'Terms of use' there are no clear indications regarding pricing. This is probably due to the fact that data reside in different partner servers, and each partner has data 'under moratorium' (free, but the availability depends on how ready the data are for downloading). In strict agreement with the indicator definition it is thus classified as " Not or not well documented", i.e., 'RED'.
- EMODnet Seabed habitats: defined on the first web page as free. 'GREEN'.
- EMODnet Chemistry: normally the data are free. 'GREEN'.
- EMODnet biology: the indicator is 'GREEN'
- EMODnet Physics: states that access is based on Copernicus, EuroGOOS and SeadataNet conditions. 'GREEN'
- EMODnet human activities: in 'Help' it is clearly stated that the products are for free. 'GREEN'.

6.2.7 AV-AC-5: Readiness

- EMODnet Bathymetry: As for CMEMS, on the web pages there is no clear statement of the format, although the user can understand that the data are in ASCII. This is clearly stated in a 'DTM format documentation' accessible through the web pages 'Data products'. The format is "Not proprietary and content clearly specified (eg auto-descriptive eg ODV, NetCDF CF) or at least with appropriate documentation". The indicator is 'GREEN' on the basis of this definition, however the user is obliged to read the documentation to understand the format. For the maps, the user must click on the map itself to find the different formats for downloading the data. EMODnet Geology: the portal does not state

whether the original data can be accessed, but provides a VMS / OGC service. Regarding the VMS service the indicator is GREEN.

- EMODnet Seabed habitats: there are good services to view and download maps and data. 'GREEN'
- EMODnet Chemistry: the data are ready for use, the indicator is 'GREEN'.
- EMODnet biology: the data are ready to be used, the indicator is 'GREEN'
- EMODnet Physics: the data are ready to be used, the indicator is 'GREEN'
- EMODnet human activities: products ready for use. 'GREEN'

6.2.8 AV-PE-1: Responsiveness

- EMODnet Bathymetry: for the bathymetry products there is "Online downloading (i.e. a few hours or less) for release" and thus the indicator is 'GREEN'.
- EMODnet Geology: the portal offers a very good VMS service and maps can be easily viewed and downloaded: the indicator is "GREEN".
- EMODnet Seabed habitats: products are on the fly. The indicator is 'GREEN'
- EMODnet Chemistry: the machinery depends on partners' responsibility, however during the assessment the response was very high: the indicator is 'GREEN'.
- EMODnet biology: very fast response and the indicator is 'GREEN'
- EMODnet Physics: for the real time data the response is fast and the indicator is 'GREEN'. For the delayed mode, the responsiveness depends on partner accessibility. During the assessment, the response was very high and the indicator is 'GREEN'.
- EMODnet human activities: fast response for products. The indicator is 'GREEN'

6.3 EMODnet availability in synthesis

AV-VI-1 Easily found	
Bathymetry	🟢
Geology	🟢
Seabed Habitats	🟢
Chemistry	🟢
Biology	🟢
Physics	🟢
Human Activities	🟢

AV-VI-2 INSPIRE catalogue service	
Bathymetry	🟢
Geology	🟢
Seabed Habitats	🟢
Chemistry	🟢
Biology	🟢
Physics	🟢

Human Activities	🟢

AV-AC-1 Policy visibility	
Bathymetry	🟢
Geology	🟡 See text
Seabed Habitats	🟢
Chemistry	🟢 🟡 See text
Biology	🟢
Physics	🟢
Human Activities	🟢

AV-AC-2 Delivery Mechanism	
Bathymetry	🟢
Geology	🟢
Seabed Habitats	🟢
Chemistry	🟢
Biology	🟢
Physics	🟢
Human Activities	🟡 See text

AV-AC-3 Data Policy	
Bathymetry	🟢
Geology	🟢 🟡 see text
Seabed Habitats	🟢
Chemistry	🟢
Biology	🟢
Physics	🟢
Human Activities	🟢

AV-AC-4 Pricing	
Bathymetry	🟢
Geology	🔴 Not well documented
Seabed Habitats	🟢
Chemistry	🟢
Biology	🟢

Physics	🟢
Human Activities	🟢

AV-AC-5 Readiness	
Bathymetry	🟢
Geology	🟢
Seabed Habitats	🟢
Chemistry	🟢
Biology	🟢
Physics	🟢
Human Activities	🟢

AV-PE-1 Responsiveness	
Bathymetry	🟢
Geology	🟢
Seabed Habitats	🟢
Chemistry	🟢
Biology	🟢
Physics	🟢
Human Activities	🟢

7. Discussion and Conclusions

In the first Data Adequacy Report for the EMODnet Black Sea Checkpoint we described the initial assessment of the monitoring system using the potential input data sets for the Challenge products and availability indicators.

The process to set up the Checkpoint assessment is long, but timely and necessitates an extended period of tool construction and harmonization of metadata information. The Black Sea EMODnet Checkpoint came after the Mediterranean Sea Checkpoint project which enabled considerably more results to be obtained than the first Mediterranean Sea DAR using the same framework already developed in the Mediterranean Sea.

The Checkpoint methodology is based on two main goals: the building of an INSPIRE catalogue of the input data sets for the Challenges, including a standard nomenclature and vocabulary for the marine monitoring characteristics, and secondly to use "indicators" to demonstrate the fitness for use of the monitoring data sets for all Challenges.

This methodology is new and has never been developed for the marine environment. The sustainability of the blue economy through the MSFD, the WFD and the Integrated Maritime Policy requires a standardized means of assessing the environmental status of European marine waters. It is clear that environmental monitoring, across the many stakeholders of the maritime economic sectors, requires robust and authoritative methods to assess the gaps and duplications in order to reduce costs and increase efficiency.

This is why the Checkpoint concentrates on an innovative concept of "assessment through products" at the European basin scales. The general framework of environmental monitoring is encapsulated into three components which constitute the marine environmental data information flow:
1) the upstream data collection;
2) the assembly of data and forecasting products;
3) the downstream, customized services using outputs from 2).

EMODnet Checkpoint Challenges aim to generate products belonging to the third step, where multidisciplinary data are merged together to produce a final product that is close to showing societal benefits. The EMODnet Checkpoints try to answer the question: how can I use the Challenge products in order to judge the adequacy of upstream data collection, assembly and forecasting information?

On route to answering this complicated question, the EMODnet Black Sea Checkpoint has developed:
1) an INSPIRE metadatabase containing information on the input data sets that are potentially important for the Challenge products, using the SeaDataNet common vocabulary;
2) two main assessment criteria, adapted from the ISO framework, containing elements that will objectively assess the "adequacy" of the input data sets for the Challenges.

The results shown in this DAR start with an updated Literature Survey statistical analysis of the input data sets for the Black Sea. Project partners have documented a total of 452 input data sets as potentially suitable for generating the 11 Challenge products. These input data sets contain monitoring information on 40 different characteristics that are either used only by one Challenge or by more than one Challenge. The results are shown in Table 4.6 and 4.5, the latter showing that 17 of these characteristics are used by more than one Challenge. Each challenge has its own spectrum of monitoring needs that are appropriate for its targeted product. Characteristics are

D 15.2
Version: V10
Date: 15/12/2016

Sea Basin Checkpoint
Lot 4: Black Sea

EMODnet
European Marine
Observation and
Data Network

unequivocally identified through the environmental matrix where they are monitored. All seven environmental matrices were found to be important for the Black Sea Challenges (air, ice, marine water, fresh water, biota/biology, seabed/riverbed, human activities). Figure 4.3 shows that Challenge 10 (Bathymetry) has only one environmental matrix identified, since it is monothematic. Challenges 6 and 7 (Fishery management and Fishery impact) have the least number of input datasets as the data collection is structured by the governmental authorities following standardized protocols. Challenge 3 (Oil leak) uses data from 5 of the seven environmental matrices.

Table 4.5 lists the number of different characteristics requested more than once. The seven most requested characteristics are:
1. TEMP/temperature of the water column and
2. MBAN/bathymetry and elevation

are the most requested monitoring characteristics among all the Challenges, followed by
3. ASLV/Sea level
4. NTRA/Nitrate concentration parameters in the water column
5. PHOS/Phosphate concentration parameters in the water column
6. TDNT/Dissolved total and organic nitrogen concentrations in the water column
7. RFVL/Horizontal velocity of the water column (currents).

These are the characteristics that require a focused assessment since they cover the needs of at least three Challenges. The number of data sets for these different categories is high, and some of the Challenges use different input data sets for the same characteristic probably because of sectorial and disciplinary habits. A gap is already emerging from this initial analysis: exchange across disciplines and Challenges is lacking, generating multiple data sets for the same characteristic category but probably not with the same quality. This might prevent innovation in Challenge products and lower the quality of products.

In Table 7.1 we summarizes the number of data sets documented for the 7 most used Characteristics by the different Challenges. The DAR then continues with an initial evaluation of the availability indicators for the 452 input data sets. The 8 availability indicators are shown in Table 5.1 along with their evaluation scores. These scores are subdivided into three color ranges: "red" meaning "not adequate", "yellow" meaning "partly adequate" and "green" fully adequate.

The overall results across all Challenges are shown in Table 7.2. For most of the indicators the situation is bipolar, half "not adequate" and the other half "adequate. The single most negative indicator is the "INSPIRE catalogue service" (AV-VI-2) visibility indicator which is generally not adequate. Furthemore the "Pricing" (AV-VI-4) indicator is subdivided in half between "not well documented pricing policy" and "open and free data policy". The results however show that there is a net positive score, most of the data sets used by the Challenges are "adequate" in terms of the availability criteria.

Most used Characteristics among Challenges	Number of input data sets potentially useful	Challenges using the characteristics
1. TEMP/temperature of the water column	120	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH03 - Oil Platform Leaks CH04 - Climate CH08 - Eutrophication CH09 - River Inputs
2. MBAN/bathymetry and elevation	46	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH03 - Oil Platform Leaks CH10 - Bathymetry
3. ASLV/Sea level	54	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH05 - Coasts
4. NTRA/Nitrate concentration parameters in the water column	29	CH09 - River Inputs CH08 - Eutrophication CH09 - River Inputs
5. PHOS/Phosphate concentration parameters in the water column	21	CH09 - River Inputs CH08 - Eutrophication CH09 - River Inputs
6. TDNT/Dissolved total and organic nitrogen concentrations in the water column	6	CH09 - River Inputs CH08 - Eutrophication CH09 - River Inputs
7. RFVL/Horizontal velocity of the water column (currents	9	CH01 - Windfarm Siting CH02 - Marine Protected Areas CH03 - Oil Platform Leaks

Table 7.1 Most requested characteristics by different Challenges and number of data sets listed as potentially interesting
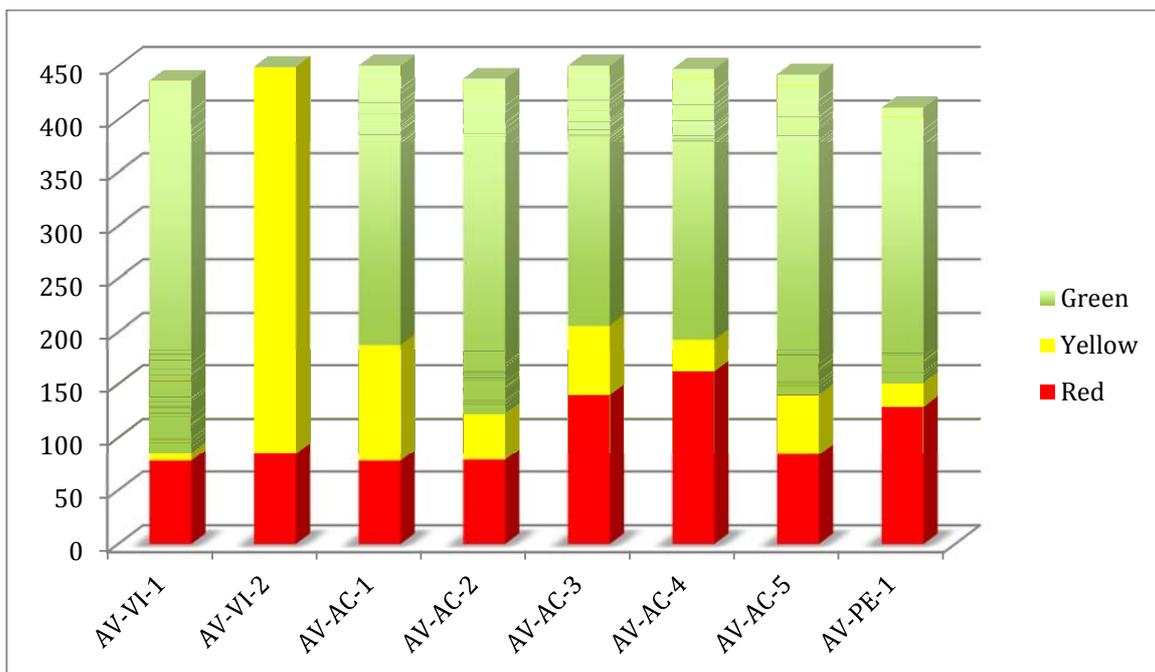

Table 7.2 Overview of availability indicator scores across Challenges

If we now examine the availability indicators for input data sets related to the 40 different characteristics used by all Challenges, 12 characteristics categories have related datasets that are partly or not adequate in terms of availability. They are:

1) the "TSED/Concentration of suspended particulate material in the water column".
2) the CRYS/Snow and Ice mass, thickness and extent
3) The CAPH/pressure (measured variable)
4) CDTA/Air temperature and density
5) CHUM/Atmospheric humidity
6) CNTX/Phytoplankton generic biomass in water bodies
7) GP079/Zooplankton wet weight biomass
8) EWSB/Wind strength and direction
9) MBAN/Bathymetry and elevation
10) PSAL/Salinity in the water column
11) WVSP/Spectral wave data parameters
12) WVST/Wave height and period statistics

For these characteristics the input data sets are "difficult to find", they do not have an "INSPIRE catalogue service", most of the time the "data policy is not stated" and/or the "data are proprietary and/or there is no indication on how to get them". Furthermore, the accessibility is reduced, either because the "format of the data is not appropriate" for the re-use of the data or because the "data delivery mechanism" is slow.

Looking in detail at the CMEMS and the 7 EMODnet Thematic Portals, not considering "coastal mapping" which is still under construction, Section 7 shows that with a few exceptions the availability scores are "green".

The Portal that shows the greatest number of weaknesses is the "Geology" portal, i.e.:
1) AV-AC-1 Policy Visibility: for geology this is one of the few cases in which "There is information, but details are available only on request". This is due to the fact that the data reside in different servers and may have different management systems. For this reason it is 'YELLOW'.
2) AV-AC-3 Data Policy: for geology the portal provides a service (VMS) and according to the data from 'Terms of use', the data may be accessible under moratorium. The indicator is 'YELLOW'.
3) AV-AC-4 Pricing: for EMODnet Geology from the 'Terms of use' there are no clear indications on pricing. This is probably due to the fact that data reside in different partner servers, and each partner has data 'under moratorium' (free, but the availability depends on how ready are the data for downloading). Within the strict interpretation of the indicator definition, it is classified as " Not or not well documented", i.e., 'RED'.

For the other two Portals there are another two partially negative indicators.
1) For EMODnet Chemistry and the AV-AC-2 Policy visibility indicator: the information is not directly provided in the portal, but the user has to go to SeaDataNet. The indicator is 'YELLOW' ("There is information, but details are available only on request")' since the details can be found in another portal.
2) For EMODnet Human Activities and the AV-AC-3 Data Policy: the information is given in 'About', where it is stated that "data with as few restrictions as possible". For this reason it seems that the data are 'Partially restricted' – 'YELLOW'.
For CMEMS and EMODnet Thematic Portals, we can conclude that the systems are fully adequate in terms of the availability of the data they produce and disseminate.

In conclusion, the first DAR highlights that for the EMODnet Black Sea Checkpoint, a large number of data sets are available for the Challenges products (452) but their adequacy varies greatly from Challenge to Challenge and in terms of the different characteristics. The method developed in the Black Sea EMODnet Checkpoint seems useful for an objective assessment of the adequacy of the monitoring system at the basin scales. The second DAR will couple the availability criteria scores with the appropriateness, thus producing a complete assessment of monitoring systems at the Black Sea basin scale level.

Annex 1 Metadata base structure

An example of the metadatabase content for the Characteristics "Nitrate concentration" is given below.

Nitrate concentration parameters in the water column | Concentration of nitrate {NO3} per unit volume of the water body [unknown phase] | EMODNet Chemistry 2 | Black Sea DIVA 4D analysis of Water_body_nitrate - Winter

CHARACTERISTICS

Environmental matrix	Marine water
Group of Category of characteristics (P03)	Carbon, nitrogen and phosphorus
Category of characteristics (P02)	Nitrate concentration parameters in the water column
Characteristic name (P01)	Concentration of nitrate {NO3} per unit volume of the water body [unknown phase]
Characteristic name	Concentration of nitrate {NO3- CAS 14797-55-8} per unit volume of the water body [unknown phase]
INSPIRE themes	Environmental monitoring facilities
Processing level of characteristics	High level analyzed
Production mode	Delayed

DATA SOURCES

Program name/project name

Program title	European Marine Observation and Data Network (EMODNET) - Chemistry
Program acronym	EMODNet Chemistry 2
Program/project web site URL	http://www.emodnet-chemistry.eu/

Program name/project name

Program title	European Marine Observation and Data Network (EMODNET) - Chemistry
Program acronym	EMODNet Chemistry 2
Program/project web site URL	http://www.emodnet-chemistry.eu/

Data provider

Organisation name	EMODNet Chemistry 2
Data provider website	http://www.emodnet-chemistry.eu/
Hierarchy data level	Dataset
Catalogue name from provider	Portal for Chemistry - Sextant Catalogue services
Catalog URL	http://sextant.ifremer.fr/en/web/emodnet_chemistry/catalogue
Data collection name or data set name giv…	Black Sea DIVA 4D analysis of Water_body_nitrate - Winter
Data collection name or data set name	Black Sea DIVA 4D analysis of Water_body_nitrate - Winter
Dataset identifier in the catalog	bcedaa30-1a56-11e5-a8e5-fc4dd4d3ff8e

Dataset URL	http://sextant.ifremer.fr/record/bcedaa30-1a56-11e5-a8e5-fc4dd4d3ff8e/

OVERVIEW

Purpose of characteristics production	Gridded product for visualization of Water body nitrate in the Black Sea generated by DIVA 4.6.9 using all EMODNET data from 1975-01-01 to 2013-12-31. Every year of the time dimension corresponds to a 10-year centred average for the Winter season (December-February).
Production and quality assessment specification reference	http://www.emodnet-chemistry.eu/documents/EMD2chem_QCreport_V8-072015.pdf
Used by challenges	BlackSea - CH09 - River Inputs
Intended use description: objective, process, description, output data	Seasonal estimates of nutrients in marine water in the area close to the river inputs into the sea

SPATIAL COVERAGE
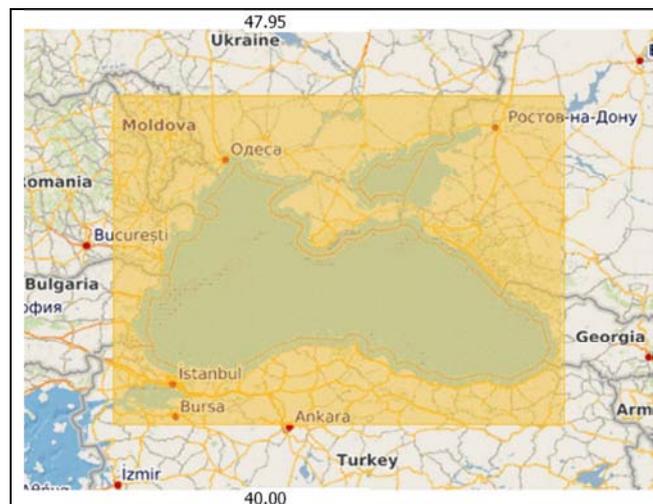
Horizontal resolution (meters)	1
Min depth (meters)	0
Max depth (meters)	150

TIME COVERAGE

Temporal extent

Time resolution	10-year centred average
Maintenance and update	Unknown

frequency

AVAILABILITY

Visibility of dataset

Easily found

EU catalogue service

Search via reference catalogue (e.g. MyOcean, GEOSS Geoportal…)
The datasets are referenced in a public national catalogue, in an international
catalogue service

Accessibility of dataset

Visibility of data policy

Data delivery mechanisms

Data policy

Pricing

Readiness

There is detailed information provided to understand data policy
Online discovery + downloading + viewing services (Advanced services)
Unrestricted
Open and Free. No charge
Format not proprietary and content clearly specified (e.g. autodescriptive like ODV,
NetCDF CF) or at least with appropriate document describing the content

Data formats and conventions
Format

NetCDF

Performance of dataset

Responsiveness

Online downloading (i.e. a few hours or less) for release

METADATA INFORMATION

Title

Nitrate concentration parameters in the water column | Concentration of nitrate
{NO3} per unit volume of the water body [unknown phase] | EMODNet Chemistry 2 |
Black Sea DIVA 4D analysis of Water_body_nitrate - Winter

Contact for the metadata
Date info (Revision)

2016-08-17T10:42:17

European Marine Observation and Data Network

Sea Basin Checkpoint
Lot 4: Black Sea

D 15.2
Version: V10
Date: 15/12/2016

Annex 2 Metadata base contents

The combination of the environmental matrices and the Seadatanet common vocabulary enables the Challenge needs to be analysed easily and comprehensively.

We can see in detail which category of characteristics has the largest number of input data sets indicated by the Challenges. The statistics are given in Table A.1. The largest number of datasets is listed for the characteristic category "TEMP/Temperature of the water column". "ASLV/Sea level" follows, listed as monitored by 54 data sets, "MBAN/Bathymetry and Elevation" by 46 data sets. Other categories of characteristics have a number of input data sets well below 30.

P02	Input data sets	01	02	03	04	05	06	07	08	09	10	11
AATX/Phytoplankton taxonomic surface area in water bodies	1								1			
ALAT/Horizontal spatial co-ordinates	2		2									
ASLV/Sea level	54	3	8			43						
BRDA/Bird taxonomy-related counts	4		4									
CAPH/Pressure (measured variable)	2	2										
CDTA/Air temperature and density	3	3										
CETA/Cetacean abundance	2		2									
CHUM/Atmospheric humidity	2	2										
CNTX/Phytoplankton generic biomass in water bodies	6				5				1			
COAS/Terrestrial mapping	1		1									
COGE/Coastal geomorphology	1		1									
CPWC/Chlorophyll pigment concentrations in water bodies	8		3						5			
CRYS/Snow and ice mass, thickness and extent	17				17							
DBDX/Geological sample density	1		1									
DOXY/Dissolved oxygen parameters in the water column	2								2			
EWSB/Wind strength and direction	8	4	4									
FABD/Fauna abundance per unit area of the bed	5		5									
FCNT/Fish taxonomy-related	2						2					

counts										
FCST/Fish and shellfish catch statistics	4						1	3		
FIBM/Fish biomass in water bodies	2							2		
GP079/Zooplankton wet weight biomass	11									11
GP087/Fishery characterisation	3		1					2		
GWDR/Wave direction	1	1								
HBEX/Habitat extent	3		2	1						
MBAN/Bathymetry and Elevation	46	4	2	2						38
NTRA/Nitrate concentration parameters in the water column	29							14	15	
PHOS/Phosphate concentration parameters in the water column	21							10	11	
PNTX/Phytoplankton generic abundance in water bodies	5				4			1		
PSAL/Salinity of the water column	2	2								
PSST/Skin temperature of the water column	11		4		7					
RFVL/Horizontal velocity of the water column (currents)	9	4	1	4						
RVDS/River flow and discharge	16					2			14	
TDNT/Dissolved total and organic nitrogen concentrations in the water column	6							2	4	
TDPX/Dissolved total or organic phosphorus concentration in the water column	2							2		
TEMP/Temperature of the water column	120	15	3	2	83			7	10	
TPHS/Particulate total and organic phosphorus concentrations in the water column	5								5	
TSED/Concentration of suspended particulate material in the water column	12								12	
WVSP/Spectral wave data	2	2								

	Input data sets	01	02	03	04	05	06	07	08	09	10	11
parameters												
WVST/Wave height and period statistics	8	5		3								
ZATX/Zooplankton taxonomy-related abundance per unit volume of the water column	13											13

Table A.1: P02 classified characteristics as a function of Challenges and number of input data sets

If we use the P22 INSPIRE themes (Table A.2), the largest number of data sets is from category "28/Oceanographic geographical features" with 161 data sets. These data sets are distributed as follows: 107 data sets belong to CH04 Climate, followed by CH01 Windfarm Siting (16 data sets), CH02 - Marine Protected Areas (11 datasets) and so on, in descending order. After this, the category of "20/Environmental monitoring facilities" follows, covering 110 data sets.

P22	Input data sets	01	02	03	04	05	06	07	08	09	10	11
10/Elevation	47		10	3		34						
13/Geology	1			1								
158/Protected sites	1			1								
163/Hydrography	58	4				2				14	38	
20/Environmental monitoring facilities	110	17							45	48		
24/Area management /restriction/regulation zones and reporting units	2		2									
26/Atmospheric conditions	10	10										
27/Meteorological geographical features	4			4								
28/Oceanographic geographical features	161	16	11	9	107	9				9		
31/Habitats and biotopes	2		2									
32/Species distribution	55		12		9		3	7				24
34/Mineral resources	1			1								

Table A.2: P22 classified characteristics as a function of Challenges and number of input data sets

If we use the P03 Agreed Parameter Groups (Table A.3), the parameter group with the most input datasets that are potentially useful for the Challenge products is "D025/Water column temperature and salinity" with 133 data sets. The majority of them belongs to CH04 Climate (90 data sets). After this, the parameter group "D032/Sea level", "C040/Nutrients" and "G05/Gravity, magnetics and bathymetry" follow with 54, 49 and 42 data sets. After this, the remaining parameter groups include less than 25 input data sets.

P03	Input data sets	01	02	03	04	05	06	07	08	09	10	11
B015/Birds, mammals and reptiles	4		4									
B030/Phytoplankton and microphytobenthos	12				9				3			
B035/Pigments	8		3						5			
B045/Zooplankton	24											24
B050/Habitat	3		2	1								
B070/Biota abundance, biomass and diversity	7		7									
C005/Carbon, nitrogen and phosphorus	14								4	10		
C015/Dissolved gases	2								2			
C025/Hydrocarbons	1			1								
C040/Nutrients	49								24	25		
C025/Hydrocarbons	1			1								
D025/Water column temperature and salinity	133	17	7	2	90				7	10		
D030/Currents	8	3	1	4								
D032/Sea level	54	3	8			43						
D034/Waves	12	9		3								
G005/Gravity, magnetics and bathymetry	4		2	2								
G015/Suspended particulate material	12									12		
G05/Gravity, magnetics and bathymetry	42	4									38	
H004/Fisheries	11		1			3	7					
M010/Meteorology	15	11		4								
M015/Cryosphere	17				17							
O005/Fluxes	16					2				14		
T001/Terrestrial	2		2									
Z005/Administration and dimensions	2		2									

Table A.3: P03 parameter groups as a function of Challenge and number of input data sets