

Interactive comment on “The role of EMODnet Chemistry in the European challenge for Good Environmental Status” by Matteo Vinci et al.

Matteo Vinci et al.

mvinci@ogs.trieste.it

Received and published: 28 October 2016

Interactive comment on “The role of EMODnet Chemistry in the European challenge for Good Environmental Status” by Matteo Vinci et al. Anonymous Referee #2 Received and published: 11 September 2016 Review of the manuscript: The role of EMODnet chemistry in the European challenge for good environmental status. M. Vinci, A.Giorgetti, M. Lipizer Submitted to Natural Hazards and Earth System Sciences. REf: NHESS226-2016.

The paper describes the effort carried out by the EMODnet Chemistry community to make available, to the marine sciences audience, a reliable and effectively usable dataset for chemical and biogeochemical properties for the European Seas, aiming to contribute the information and knowledge base necessary for the EU-MSFD ob-

C1

jective of “good environmental status” for the European Seas. It is an important and highly valuable effort. The present manuscript does not attempt to extract scientific information from the data collected, but tries to provide a description of the structure and the quality of the data that EMODnet chemistry is going to make available to the marine scientists communities. Therefore, the manuscript, even if it cannot precisely be considered as a scientific paper providing original findings, deserves (in principle) publication. Unfortunately due to a series of formal problems it cannot be published in the present form, as it looks like hastily written, without taking much care in clarity and ordered structure. A revision of the formal structure of the manuscript is absolutely mandatory. Many concepts and information are taken for granted and a general reader might therefore find the manuscript rather confusing. Figures (see below) are not correctly referenced in the text and often the relative captions are very, very sloppy.

The authors acknowledge the comments and suggestions of the reviewer and the need of a revision of the manuscript with the objective to improve the clarity of the described topic. The manuscript has been integrated explaining better the concepts (that sometimes were taken for granted). Presentation of figures and tables were improved following the reviewer’s suggestions. Following these actions we believe that the manuscript has already improved and been made clearer. If the editor considers it necessary, we can also reorganize the structure of text, add a glossary to facilitate understanding of several specific terms used in the manuscript and the text can be revised by a native English speaker. Below the replies to the specific comments received. A reviewed version of the manuscript will be included in the reply supplement to clarify the improvements. In red are highlighted the changes done following the comment n.1 while in orange the changes done following the comment n.2.

Below some specific remarks that I hope might help the Authors to improve the manuscript.

Line 50: please explain better how a data management system could achieve interoperability and resilience. The explanations given are still a bit “obscure”.

C2

We suggest the following improvement.

In the field of marine research, during the last decades several oceanographic data management initiatives faced the challenge of data availability, interoperability and resilience at Pan-European level. (EU MAST MTP II MATER 1996-1999, EU MAST-INCO MEDAR 1999-2001, FP6 SeaDataNet 2006-2011, FP7 SeaDataNet2 2011-2015). Interoperability is defined as "the ability of a system to work with or use the parts of another system", while resilience is defined as "the ability of a system to cope with change". The translation of these principles in the oceanographic data management consists in the development of a long life system able to easily interact with other systems. As example the adoption of common formats for data and metadata and a system of common vocabularies ensure that the network of involved persons is working in a homogeneous environment from the syntactic and semantic point of view (speaking a common language). The resilience is safeguarded by metadata and quality flags that provide clear knowledge of which kind of information the users are handling even long time after the data measurement (e.g. use of historical data for time series studies).

Line 111: MSFD and not MSDF

Ok

Line163. explain better the criteria for data restriction

Following the comment to the manuscript received by the previous reviewer this part has been eliminated. The text describing the data policies is available from line 175 until 186 as follow:

Data access is regulated by a data policy (defined in agreement with data originators) which aims to establish a balance between the right of the originator to get proper acknowledgment for data acquisition, and the need for open access through free and unrestricted exchange of data, meta-data and data products. The analysis of data policies Ffor EMODnet Chemistry data shows differences between data access restrictions

C3

for nutrients and contaminants (Fig. 3).

Fig. 3: Data policy for nutrients and contaminant data

Data requests from registered users are handled by NODCs through a data policy management system. Unrestricted data are freely available while restricted ones need negotiation with data originators. This kind of filter on data access is an effective way to establish contacts and trust between data originators and data management centres, ensuring correct acknowledgement, which ultimately encourages data sharing.

Line 195 and followings. Explain better (for the general reader) the meaning of codes such as P01 vocabulary and P35 vocabulary.

We prefer a more general explanation of the vocabularies involved in our workflow to avoid too specific or technical descriptions. Following this we suggest the following improvement.

... Data aggregation is done with the objective to unify the various analytic terms into a unique aggregated term with conversion to a unique measurement unit.. The ODV software has a built-in aggregation procedure applying a number of business rules like possible units conversions. (Lowry R. et al., 2013) The main goal of this activity is to obtain a harmonized dataset (e.g. a unique dataset of phosphate concentration in the water column starting from different datasets of phosphate concentration expressed with different units) that could be used to generate homogeneous data products. The results of the regional quality control are sent to the data collators (NODCs) to correct errors or anomalies in the original copy of the data available in the EMODnet infrastructure. This feedback loop guarantees data quality upgrade (Fig.5).

Section 6. Spend more words to illustrate the procedure for data mapping (DIVA protocol)!!!!

We suggest the following improvement.

The interpolated maps have been produced with the variational inverse method (VIM;

C4

Brasseur et al., 1996), using the software DIVA (Data-Interpolating Variational Analysis; Troupin et al., 2010). DIVA is an appropriate numerical implementation of VIM suitable for oceanographic data spatial analysis as it is designed to obtain a gridded field from the availability of non-uniformly distributed observations (Barth et al., 2010; Troupin et al., 2012).

Section 6 Validation loop must be described better. Just putting a (not referenced) figure with a sloppy caption is not enough!!!!

Figure 6 has been moved to section “5 Data Quality” where the “Validation loop” was described in a quite detailed way in the text from line 202 to 227, and the figure is now cited in the text. Now the figure is in the correct position and cited to link the description to the image (that in our opinion clarify in a simple but efficient way the workflow).

5. Data Quality The quality of the data is a key issue when merging heterogeneous data coming from different sources, periods and geographic areas. Within EMODnet chemistry community, commonly agreed and standardized data quality control (QC) protocols have been defined (Holdsworth, 2010) to guarantee consistency among comprehensive databases which include data from different and/or unknown origin and covering long time periods. As a first step, the data are checked and completed by collators with a standard set of metadata that provide the basic information necessary for their long term use. Afterwards, data undergo a validation loop which consists in several validation steps. The first is done by data collators, prior to the inclusion in the decentralized infrastructure and the second step, which consists in regional quality control, is performed at regional scale on aggregated datasets. The first quality controls (QC) ensure that position and time of data are realistic and compare measurements with broad ranges and specific regional ranges. Whenever available, data are also compared with climatology. As a result of the first QC step, all data are archived with a quality flag value that provides information about their reliability. At this point, data aggregation and regional quality control are performed at regional scale, following a common protocol. Data aggregation is done with the objective to unify the various an-

C5

alytic terms into a unique aggregated term with conversion to a unique measurement unit. The ODV software has a built-in aggregation procedure applying a number of business rules like possible units conversions. (Lowry R. et al., 2013) The main goal of this activity is to obtain a harmonized dataset (e.g. a unique dataset of phosphate concentration in the water column starting from different datasets of phosphate concentration expressed with different units) that could be used to generate homogeneous data products. The results of the regional quality control are sent to the data collators (NODCs) to correct errors or anomalies in the original copy of the data available in the EMODnet infrastructure. This feedback loop guarantees data quality upgrade (Fig.5).
Fig. 5: Data validation loop

To improve and homogenize the quality control procedures and standards adopted (at least at regional level), a quality control survey has been carried out within EMODnet Chemistry community, in order to collect the best practices in data validation and highlight gaps of the different institutes involved (Vinci et al., 2015).

Figures and tables are not correctly referenced in the text. Please reference them correctly. Just writing (for instance) “with the following data policy distribution” and putting below a figure (or a table) is not OK. Moreover figure captions need to be rewritten in order to be more consistent with the pertinent text and with the figures themselves In particular: Figs 1, 2, 3, 5, 6, 7 are not referenced in the text Tab 1. Is not referenced in the text and the caption needs rewriting Fig. 4 needs a better caption.

Authors agree on the improvement of figures, tables, captions and references that have been already updated in text.

Fig.1 is cited now at line 150 Fig.2 is cited now at line 168 Fig.3 is cited now at line 178 Fig.4 and Tab.1 are cited now at line 189 Fig.5 is cited now at line 221 Fig.6 is cited now at line 242 Fig.7 is cited now at line 252

Please also note the supplement to this comment:

C6

C7

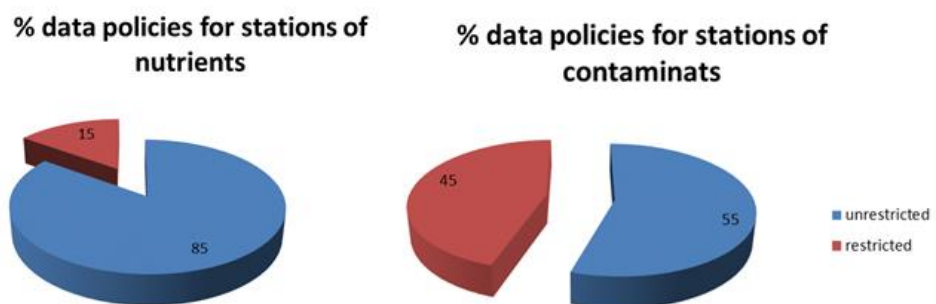


Fig. 1.

C8

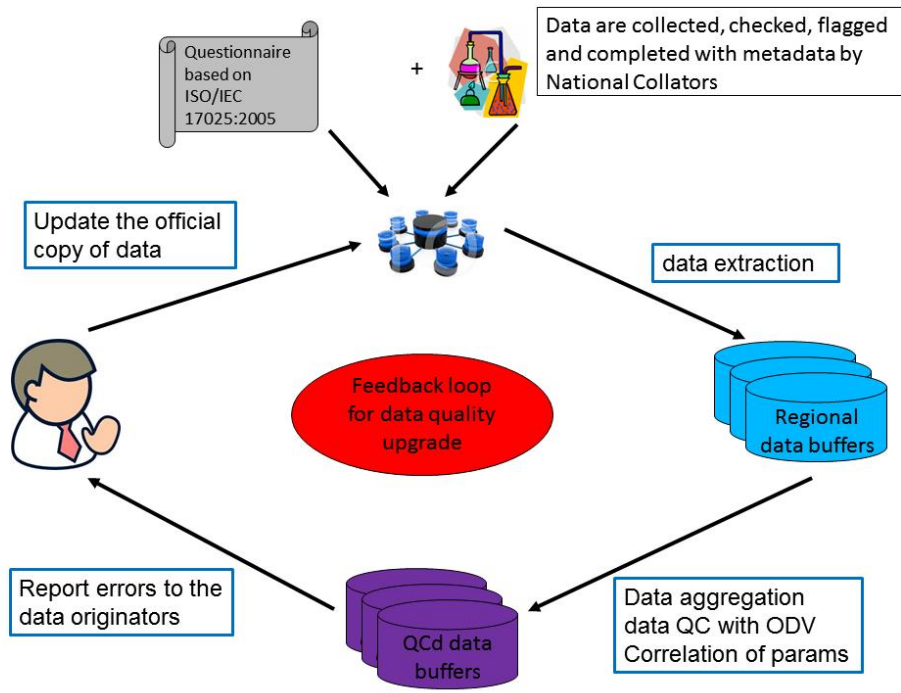


Fig. 2.