



SEA Wave: Strategic Environmental Assessment of Wave energy technologies

Deliverable Report D3.5

Guidelines on data collection, structuring and metadata



This project has been co-funded by the European Maritime and Fisheries Fund (EMFF) of the European Union. The contents of this document reflect only the (author's view. EASME is not responsible for any use that may be made of the information it contains.





Revision

Revision	Date	Description	Originated by	Reviewed by	Approved by
1.0	29/4/2019	First version of the document	José Chambel Leitão		
2.0	29/4/2019	Review 1 st draft	José Chambel Leitão	David Darbinyan	

Project Information

Project title	Strategic Environmental Assessment of Wave energy technologies
Project acronym	SEA Wave
Grant agreement number	EASME/EMFF/2017/1.2.1.1/01/SI2.787660
Project start date	01/11/2018
Project duration	36 months
Project lead	The European Marine Energy Centre (EMEC) Ltd
Project website	www.seawave-emff.eu



Contents

1	Context	1
2	Data Produced within SEA Wave	1
3	Environmental data	2
4	Metadata	2
4.1	General concepts	2
4.2	Reference documents	4
5	Storage and uploading of data	4



List of Tables

Table 1 - Summary of data from survey campaigns	1
Table 2: Guidelines on ICES data types.....	4



1 Context

Data measured by project partners must be structured for dissemination purposes and to ensure transferability with existing data platforms EMODNet, SeaDataNet and Copernicus Marine Services.

As the development of the Data Platform will run in parallel with data campaigns, a set of standards (in line the above-listed existing platforms) on data structure must be in place. The present document includes guidelines on:

- What data are expected to be disseminated and how will they be classified according to format, size or other characteristics;
- What standards will be used and what metadata (compliant with INSPIRE Directive) will be produced for each dataset;
- Where each type of data will be stored;
- How each dataset will be uploaded and by whom.

This document summarizes the concept of metadata that will be adopted by SEA WAVE data platform, following the INSPIRE data specification template in its relevant parts, i.e., dataset-level, services metadata and data quality.

2 Data Produced within SEA Wave

A summary of the data that will be gathered in survey campaigns during the project by the universities of Exeter and Plymouth is presented in Table 1. Also data on device activity at the time of monitoring will be collected.

An example of summary data table (Table 1) is given bellow. Data will be created following each survey campaign as *.csv file, using these suggested field headers.

Device	Date & Time	Location	Equipm-ent Type	Data description	Survey Duration / scale	Geometry
Towed camera	2019-08-11-1206	Lat/Lon	Bowtech Camera	Species abundance & diversity	20 minutes / 200 metres	Lines 2D on seabed
Baited camera	2019-08-11-1206	Lat/Lon	BRUV systems	Species abundance & diversity	30 minutes	Points 2D on seabed
Fisheries acoustics	2019-08-11-1206	Lat/Lon	Simrad EK80	NASC	Instantaneous at 30-sec interval	Points 3D
Ambient acoustics	2019-08-11-1206	Lat/Lon	Jasco AMAR-G2	Third Octave Band	1-hr interval averages	Time-series at fixed location

Table 1 - Summary of data from survey campaigns



Some videos may be published in YouTube using deep linking to highlight the most relevant features obtained in the campaign. These videos would be used primarily as a dissemination tool for the project.

In a later stage of the project, at least the following items must be addressed, thereby resulting in a revised deliverable 3.5:

- Additional details on data classification according to format, size and characteristics;
- Exactly where each data set will be uploaded (project's database or a SeaDataNet Centre, for example);
- Who will be responsible for the upload and acceptance of the data.

3 Environmental data

Data on marine weather conditions during the campaigns will help interpret the results. This means that measured or modelled data on atmospheric and hydrodynamic parameters will be added to the data platform. During the project, this data will be selected in terms of its significance for the project and its availability.

4 Metadata

4.1 General concepts

Metadata refers to the description of datasets and services in a compliant form as it has been defined by the Directive 2007/2/EC (INSPIRE) and Commission Regulation No 1205/2008.

Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is essential for understanding stored information. It describes the content, quality, condition, and other characteristics of a data set or the capabilities of a service. Creating metadata or data documentation for geospatial datasets is crucial to the data development process. Metadata is a valuable part of a dataset and can be used to:

- **Organize** data holdings (Do you know what you have?).
- Provide **information about** data holdings (Can you describe to someone else what you have?).
- Provide information **to data users** (Can they figure out if your data are useful to them?).
- **Maintain the value** of your data (Can they figure out if your data are useful 20 years from now?).

In the geographical domain we can have a description of spatial data (**spatial data** metadata), a service (**service** metadata) or a special analysis process



(**process** metadata). Most for the standardization work is done for data metadata, however service and process metadata become increasingly important. Metadata is used in discovery mechanisms to bring spatial information providers and users together. The following mechanisms are recognized:

- **Discovery**: which data source contains the information that I am looking for?
- **Exploration (or evaluation)**: do I find within the data sources the right information to suit my information needs?
- **Exploitation (use and access)**: how can I obtain and use the data sources?

Each mechanism has its own use of metadata. The selected standards should fulfil the needs to carry out services using these mechanisms. Metadata is required to provide information about an organisation's data holdings. Data resources are a major asset, and information of what datasets exist within different organisations, particularly in the public sector, is required to improve efficiencies and reduce data duplication. Data catalogues and data discovery services enable potential users to find, evaluate and use that data, thereby increasing its value. In addition, metadata received from an external source may require further information supplied to metadata to allow easy process and interpretation.

In this context for all types of data the following information is required (SeaDataNet, 2010):

- **Where** the data were collected: location (preferably as latitude and longitude) and depth/height;
- **When** the data were collected (date and time in UTC or clearly specified local time zone);
- **How** the data were collected (e.g., sampling methods, instrument types, analytical techniques). How do we organize the data (e.g., in terms of station numbers, cast numbers);
- **Who** collected the data, including name and institution of the data originator(s) and the principal investigator;
- **What** has been done to the data (e.g., details of processing and calibrations applied, algorithms used to compute derived parameters);
- **Watch** points for other users of the data (e.g., problems encountered and comments on data quality).

The ICES¹ Working Group on Data and Information Management (WGDIM) has developed a number of data type guidelines which itemize these elements that are required for thirteen different data types (see table below). These Data Type Guidelines have been developed using the expertise of the oceanographic data centres of ICES Member Countries. They have been designed to describe the elements of data and metadata considered as important to the ocean research community. These guidelines are targeted towards most physical-chemical-biological data types collected on oceanographic research vessel cruises. Each guideline addresses the data and metadata requirements of a specific data type. This covers three main areas:

¹ International Council for the Exploration of the Sea (www.ices.dk)



- What the data collector should provide to the data centre (e.g., collection information, processing, etc.);
- How the data centre handles data supplied (e.g., value added, quality control, etc.);
- What the data centre can provide in terms of data, referral services and expertise back to the data collector. A selection of these guidelines, in particular for those data types that are not yet dealt with in detail here, are included in Appendix 1 of this document.

ICES Data Type Guidelines		
CTD	Moored ADCP	Moored Current Meter
Shipborne ADCP	Seasoar (Batfish)	Surface (Underway)
Water Level	XBT	Net Tow (Plankton)
Surface Drifting Buoy	Profiling Float and Drifting Buoy	Discrete water sample
Multibeam echosounder data		

Table 2: Guidelines on ICES data types

Additional information on this topic can be seen at “SeaDataNet - Data Quality control Procedures, Version 2.0, May 2010”²

4.2 Reference documents

Two reference documents cited above must be used, along with the present Deliverable, in the course of the project:

1. COMMISSION REGULATION (EC) No 1205/2008 of 3 December 2008; “implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata”;
2. SeaDataNet - Data Quality control Procedures, Version 2.0, May 2010

These documents are available in the project’s SharePoint.

5 Storage and uploading of data

Data storage locations, for data referred in sections 2 and 3, will be decided during the project.

Metadata will be generated by partners who also generate the data. To help on this task, a software tool will be made available by Hidromod. In this tool, metadata information will also be validated against the relevant standards. This tool will be based on GeoNetwork opensource³.

² [https://www.seadatanet.org/content/download/596/3118/file/SeaDataNet_QC_procedures_V2_\(May_2010\).pdf?version=1](https://www.seadatanet.org/content/download/596/3118/file/SeaDataNet_QC_procedures_V2_(May_2010).pdf?version=1)

³ GeoNetwork is a catalog application to manage spatially referenced resources. It provides powerful metadata editing and search functions as well as an interactive web map viewer. It is currently used in numerous Spatial Data Infrastructure initiatives across the world (<https://geonetwork-opensource.org/>)



This project has been co-funded by the European Maritime and Fisheries Fund (EMFF) of the European Union. The contents of this document reflect only the author's view. EASME is not responsible for any use that may be made of the information it contains.

www.seawave-emff.eu