VALMER

TOPIC PAPER 2

Social and economic data management

The need for high-quality data relating to the marine environment has never been greater. National and international legislation and directives depend upon best available evidence to designate marine protected areas or report on the environmental status of our seas. The marine sector has made excellent progress in collating and standardising environmental marine data however, an understanding of human activities and pressures on the marine environment is also central to successful marine planning and conservation activities. These activities and pressures data are often omitted or, if collected, managed by disparate organisations and therefore difficult to collate. Until very recently, large EU data initiatives such as SeaDataNet[1] and EMODNet[2] have largely ignored marine social and economic data, yet many of the challenges faced are common.These relate to effective description of data, its availability, format and exchange.

Principles of data reuse - Collect once use many times

The concept of open data involves a move towards the principle of 'collect once, use many times.' Only a little extra effort is required to collect data in a way in which it can be reused, whereas trying to describe or standardise previously created datasets can be very time intensive and frustrating. It may also result in the loss of good datasets from analyses because they lack vital information.

[1] SeaDataNet (www.seadatanet.org) has developed an efficient distributed Marine Data Management Infrastructure for the management of large and diverse sets of data deriving from in situ and remote observation of the seas and oceans and links 45 national oceanographic data centres and marine data centres from 35 countries across all European seas.

[2] The European Marine Observation and Data Network (EMODnet) is a consortium of organisations within Europe that assembles marine data, data products and metadata from diverse sources in a uniform way. Presently, there are six sub-portals in operation that provide access to marine data from the following themes: bathymetry, geology, physics, chemistry, biology, and seabed habitats . One further portal covering human activities is currently under construction.

Describe your data well – Metadata matters

Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called "data about data" or "information about data" (National Information Standards Organization).

- **What is your data about?** – Ensure that you have clearly recorded the 'why, what, when and where' associated with your project.
- **Methodology** – Understanding how your data were collected is critical for people wanting to reuse it. Reference standard protocols where possible and if you are using a unique method, ensure you supply a detailed explanation of how data were collected and processed.
- **Ownership and access** - Make it clear who owns your data, how people can access your data and any restrictions to access and use of data.
- **Quality information** - make clear any caveats or limitations to using the data and include any steps taken to Quality Control and Quality Assure your data.
- **Keywords and common vocabularies** - Help people find your data by using keywords which describe the main themes of your data, and where possible, add terms from common vocabularies to tag your data.

Within the EU, the INSPIRE Directive has developed a common metadata standard to be used by member states. Through the VALMER project, a set of common vocabularies for describing social and economic data have been created and maintained.

Creating better data

Some useful points to consider when creating a dataset.

- Are there standards for the type of data you are collecting already?
- Create, document and use Quality Controls and Quality Assurance to ensure your data are as accurate as possible.
- Can your data stand alone? – Include relevant background data if your outputs cannot be understood without them.
- Describe each field properly. It makes sense to use short names for fields in spreadsheets and databases but make sure you define and document these so other people can understand them.
- Describe your data processing clearly.

Storing data

Social and economic data are most often collated and managed by organisations without direct experience of the marine sector, since data tend to focus on societal impacts and the land-based results of marine activities (energy generation, values of catches/harvests, etc). As such there are established data centres including the Office of National Statistics and the UK Data Service (formally ESDS) as the French INSEE (Institut national de la statistique et des études économiques).

However, there is limited engagement with the marine data community and no common standards for data storage or onward dissemination. The UK and French marine data community are actively working to improve this situation, with the VALMER project acting as an exemplar.

Sharing data

Data sharing is, very simply, the practice of making data available. Under INSPIRE, data sharing specifically relates to "establishing an infrastructure for spatial information in Europe to support community environmental policies, and policies or activities which may have an impact on the environment".

While transparency and openness are considered part of the scientific method, there still remains a significant hole in the amount of research data which are made available to other researchers and sectors. This is particularly notable within the biological and social sciences.

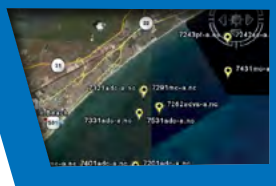
The Open Geospatial Consortium (OGC) content and services exists to encourage development and implementation of open standards for geospatial content and services, GIS data processing, and data sharing. Within VALMER we have been working to establish a data sharing infrastructure to ensure the longevity of the VALMER data and outputs beyond the projects close using such services.

Data is stored in the powerful open source object relational database system PostgreSQL. Spatial data is made "readable" by the database using the open source PostGIS programme.

Views are created in the database using SQL scripting language. These tell the database which columns of data to make available for viewing and downloading.

The view is then published from the database to GeoServer, a java-based server software that allows users to view, edit and "SHARE" their spatial information.

Data is shared using a URL on a specified website. The format of the data can be tailored to the end users needs: tabulated files (csv), geo-spatial files (shp, kml), web coded files (open layers, GeoJSON).



Allowing the user to analyse in statistical packages, GIS software and web-based applications.

UK

North Devon

N

Exeter

Dorchester

Poole Harbour

Plymouth

Truro

English Channel

Plymouth Sound - Fowey

La Manche

Cherbourg

Golfe Normand-Breton

Caen

PNMI

Granville

Saint-Malo

Brest

Saint-Brieuc

FRANCE

Golfe du Morbihan

Vannes

Creation / Adaptation : VALMER - 10/2014 - Photo Credits: Laurie-Anne Le Grégam, Yannis Turpin / Agence des aires marines protégées - Art of sun, Ethan Daniels, Florida Stock / Shutterstock.com

VALMER

The VALMER project was selected under the European cross-border cooperation programme INTERREG IV A France (Channel) - England, co-funded by the ERDF.