

Developing an Informational Web Portal for Coastal Data in Ireland: Data Issues in the Marine Irish Digital Atlas

L. O'Dea, V. Cummins, N. Dwyer
Coastal and Marine Resources Centre, ERI, University College Cork

Abstract

Data accessibility is a problem for many GIS users. In Ireland, improving accessibility to coastal and marine data is a real challenge. The Marine Irish Digital Atlas (MIDA) project aims to improve access to coastal data by offering a web GIS-based portal where data providers can display and advertise their coastal data holdings. The MIDA will also serve as an educational tool for the public and professionals alike, presenting information and links pertaining to the various elements of the marine environment, technology and management contained in the atlas. By providing a single source for locating coastal data, the MIDA will allow users to see what data exist, where data quality should be improved, and where gaps in data exist. This paper addresses data issues encountered in the first half of the project.



1. Introduction

In November 2002, Ireland began the process of developing the Irish Spatial Data Infrastructure (ISDI), which will provide a standard that government agencies will be required to use in managing their geospatial data, and that other non-governmental organisations can adopt (McCormack, 2003). This initiative will have a positive impact on the development of integrated information systems, capable of handling data from a wide variety of sources. It is envisioned that this will facilitate developments in coastal and marine area management by paving the way for protocols dealing with data harmonisation and metadata standards. By specifying standards for metadata, the ISDI will have positive implications for the vast quantities of digital coastal and marine data generated within Ireland, including data from the current the Irish Geological Survey's seabed mapping project.

Nevertheless, problems associated with data accessibility, ownership, quality, and awareness of data resources are likely to persist in preventing optimum use of coastal and marine data in Ireland. These problems are compounded by management structures for the coastal and marine environment, which are characterised by a plethora of government departments and agencies with sectoral responsibilities (e.g. fisheries, hydrocarbons, conservation, coastal engineering, research and development). As a result of the fragmented approach to management, relevant data are collected and held by 18 different agencies in six different government departments (Cummins *et al.*, 2003). Commercial organisations, universities, and non-governmental organisations are also involved in data acquisition and management.

Various other barriers exist that hinder access to, and use of, Irish coastal and marine data. Similar problems prevail in the UK (DEFRA, 2002). In order to determine what data exist and who owns them, users must devote a significant amount of time to contacting organisations and asking questions about what is

available. This results in a time-wasting exercise for acquiring even the most basic data layers, such as bathymetry or coastline. Resources are also wasted in creating datasets that may already exist due to the lack of awareness of, or restricted access to, what have already been created. In addition, data owners may not recognise the value of their data beyond the scope of their intended purpose, and thus those data are not publicised to potential users in other study areas (DEFRA, 2002). In order to improve the quality of marine research and development, users need better resources to quickly and easily identify, locate, access, and exchange coastal data and information (Cahill *et al.*, 1999). Efforts need to be made to overcome the barriers of information management to better enable access to coastal and marine data.

As an online gateway to coastal data and information, the Marine Irish Digital Atlas (MIDA) aims to address current challenges and to improve coastal data accessibility in Ireland. While data is a vital component of the MIDA, the atlas has a broader objective, which is to provide an information resource where maps can be visualised and coastal information are easily accessible (Dwyer *et al.*, 2003a). In order to reach a diverse audience and to provide a resource that is beneficial to everyone, it is important to develop a gateway to Irish coastal information that contains tools to view and query data, as well as interactive pages that educate users about coastal resources and issues relevant to Ireland and its coastal communities.

The MIDA project, which has been underway since September 2002, has provided valuable insight into issues involved with integrating disparate datasets in a web-based GIS framework. This paper looks at some of the lessons learned to date by focusing on the geospatial data component of the atlas, including issues of displaying data and their related information, data acquisition and proprietary data, as well as the management of data within the system.

2. The MIDA Framework

Web-Based Mapping System

Using a web-based mapping system as its core, the MIDA makes data available for users to visualise as layers in the map. Users control what they see by selecting layers and zooming to areas of interest. Map features can be queried, and the data owner who supplies each dataset determines the level of detail provided. Downloading datasets in GIS format will also be possible, where the data owners provide consent.

For each layer, information is also available and easily accessible from the main atlas window (Figure 1). From the Information Pane, users can view metadata related to layers they select in order to learn more about each dataset, including the contact details of data providers. There are also links to interactive information pages related to the thematic layers displayed in the atlas. These pages aim to improve awareness of the coastal

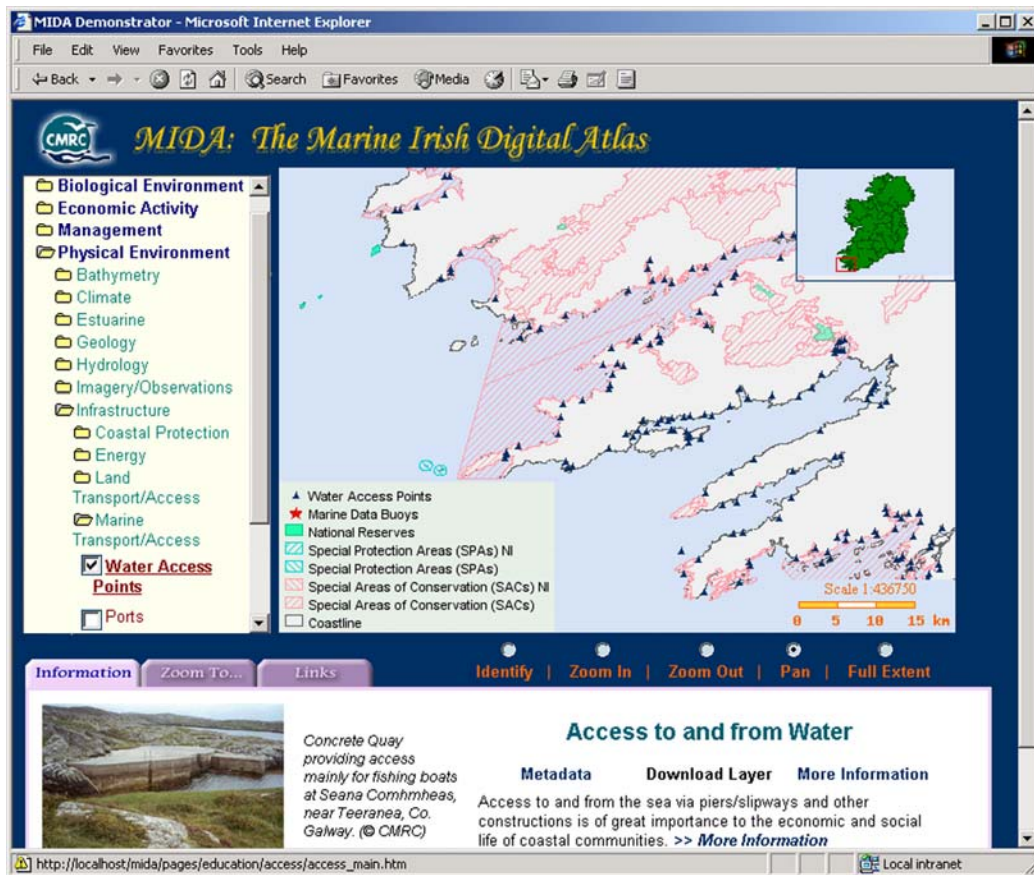


Figure 1: The main page of the MIDA Prototype. The three main components of the atlas page are the Map Pane, the Layers Pane, and the Information Pane.

environment for the general public, as well as to provide resources at varying levels of detail to accommodate a range of audiences (from students doing research to professionals looking for links to organisations in a specific sector). Links to outside resources are provided for additional information. For example, the atlas contains a layer showing the Marine Institute and UK Hydrographic Office's data buoys, which collect weather and oceanographic information from various locations off the Irish coast. When this layer is activated in the MIDA, the atlas gives the user background information to the data buoys, as well as a link to the near real time data that is displayed on the Marine Institute's web site (Figure 2).

Implementation

In general, open source software is used for the implementation of the MIDA system. An independent server running Linux at University College Cork stores all the data and web elements contained in the MIDA. Users can access the atlas over the Internet from their web browsers. When a user makes a request of the data, processing is done on the server and then returned to the user over the web (Dwyer *et al.*, 2003a). The MIDA's web mapping software is the University of Minnesota's MapServer. While it is not a

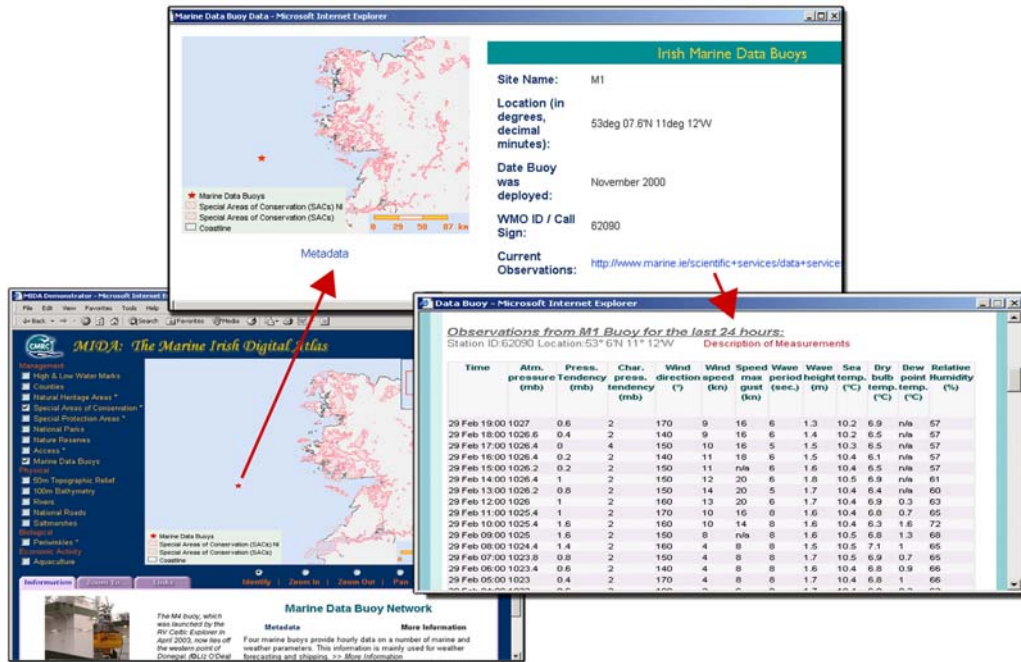


Figure 2: When available, the MIDA links to near real time data, such as the observations that are collected every hour by the four marine data buoys off of the Irish coast and displayed on the Marine Institute's web site.

fully functional web GIS system, it provides enough data viewing and querying functionality to meet the basic needs of the MIDA.

MapServer is served over the Internet using Apache Web Server. Both of these programmes are freely downloadable from the Internet and are open to customisation. Various open source databases, such as PostGRES and XML, are currently being investigated to assist in managing the vast amounts of information and data that will ultimately be stored on the server and displayed in the atlas.

Metadata framework

Metadata is collected using an in-house-designed metadata collection wizard, which stores the metadata in XML format and then displays it using XSL. An easily accessible, searchable metadata database within the atlas allows users to determine if a dataset suits their needs by providing information such as the age, resolution, and origin of the data. Metadata is displayed in a three-tier method: Abstract, which consists of a simple description; Discovery, which provides more detail and serves as the core of the searchable metadata database; and Full, which is the metadata that the data owners provide and thus in some cases may not exist.

In order to enhance data sharing, a subset of 55 elements was selected from the ISO 19115 geospatial metadata standard for the MIDA's Discovery metadata. ISO was favoured because of its recognition within SDIs worldwide. The XML tags used for storing the MIDA's metadata elements match the ISO

tags, which will enable easy access by metadata search engines as they develop in the future (Dwyer *et al.*, 2003a).

3. Data Collection

In the initial planning efforts, all possible themes and layers that could potentially be included in the atlas were identified and classified according to a data management structure developed for the MIDA (Dwyer *et al.*, 2002). This structure determines how the datasets can be accessed in the atlas through the user interface (see Figure 1). A data classification structure, with four main categories, evolved following widespread consultation with experts in the marine field. The main categories are: Management, Physical Environment, Biological Environment, and Socio-Economic Activity. For the prototype, some base layers (e.g. coastline, counties, rivers, and bathymetry) as well as easily accessible in-house data (e.g. shoreline access points) were added to demonstrate the atlas functionality.

In order to focus data collection efforts, and bearing in mind the ambitious scope of the MIDA project, decisions were taken to prioritise the datasets according to user preferences. The data were divided into Priority 1, Priority 2 and Priority 3 levels. The Priority 1 list contains a balance of datasets that are most sought after by professionals and highly desired by general users, including high and low water marks, territorial limits, protected areas, and bathymetry. Focus is placed on datasets with complete coverage of the island of Ireland. Time constraints and future project support will determine to what extent the atlas will progress to hold more regionalised data.

4. Importance of User Feedback

The MIDA's success, in part, depends on its appeal and relevance to a wide range of users. In order for it to achieve the goal of being the gateway to Irish coastal information, people must find the MIDA intuitive, user friendly, and informative. Experience from other mapping projects shows that involving users in the development of a tool from the beginning is vital. Developing a successful GIS tool for coastal information must be led by the demands of the users, and not by the technology or the data behind the tool (DEFRA, 2002).

To meet this challenge, a MIDA User Group Strategy was developed. The MIDA team met with representatives from some of the key data holders and users in Dublin in November 2003, and in Coleraine in January 2004. The initial prototype, completed in November 2003, was demonstrated and attendees were provided with an opportunity to interact with the atlas. Overall, feedback was positive towards the concept of the atlas, although there were some concerns over whether existing challenges could be surmounted, particularly in relation to the timeframe available for implementing the project. Suggestions were made in relation to priority requirements for specific data layers, improvements to the user interface, the development of enhanced search functions, and mechanisms for guiding different levels of users to their desired points of interest within the atlas. In general, attendees found the interface intuitive, the metadata useful and well-represented, and the layout

clear and concise. Much valuable discussion was generated which assisted the team in determining where to focus their energies.

Beginning in summer 2004, the MIDA development team will host a series of focus user group meetings around the country to obtain input on the current prototype. This will engage representatives from each of the potential user groups, including government agencies, research facilities, educational institutions, local authorities, commercial organisations, non-governmental organisations, and the general public. These user groups will be called upon to provide feedback on the atlas' ease of use and relevance to their needs and interests.

5. GIS Data Visualisation

The ability to visualise geospatial data via a web-based GIS is beneficial to professionals and the general public alike. Without tools to visualise such data over the web, access to that data is limited to expert users with the knowledge of what to do with the data once it is acquired. The MIDA's web based user interface eliminates the need for a desktop GIS to view the data, as well as the technical knowledge associated with operating proprietary GIS software packages, therefore making data available to a wider audience. For professionals working with coastal data, the ability to view the quality of the data, visually as well as through the metadata, can save time and money for both the user and the data supplier. Data value can increase when users have the ability to see its potential when displayed with other available atlas layers.

The benefits of making data available in this way, however, can be negated by issues that occur with increased use of GIS datasets by non-expert users. Providing a place where users can view datasets from numerous sources and of variable quality opens the door to misinterpretation of those datasets. This can be overcome to an extent in the way data are presented in the MIDA. For example, MapServer allows the atlas developer to set scale limits at which the data are displayed, so that study areas drawn at a 1:10,000 scale are not displayed with a coastline generated at 1:250,000, which could potentially place a land-based study under 20 metres of water!

6. Data Issues

Data Collection and Processing

As data collection for the atlas has progressed, numerous predicted and unexpected challenges have been encountered with regards to acquiring, preparing, and managing data. The most time consuming component of the implementation process relates to identifying what data exists, followed by time allocated for data collection. Acquisition of a typical dataset can involve several phone calls to multiple agencies, over a period of many months. Once obtained, time is then required to prepare the data for the atlas.

Due to the variety of GIS packages that exist (e.g. ArcGIS 8, MapInfo, and CARIS GIS), data comes in various GIS formats and may need to be converted

into ArcView Shapefiles or GeoTiff images, which are required for layer display in MapServer. Some information is not readily available in GIS format, which means that a certain amount of data capture and data processing is necessary to transform data from paper or alternative digital formats for inclusion in a GIS framework. For example, existing data from seabird and cetacean surveys in Irish waters currently resides in various databases in the Coastal and Marine Resources Centre. Because of complex data relationships, time must be spent determining the most informative way of displaying that data spatially, and GIS layers must then be created. Matching layers in different data projections and datums is another issue when preparing acquired data for display in the atlas. Layers must be converted into Irish National Grid, which is the chosen projection for data display in the atlas.

Data Quality

Dealing with data quality issues involves checking for errors in the datasets and eliminating inconsistencies in the way data are presented in order to provide a greater confidence in the data. It is important that end users see data exactly as the owner provides them so that the end users can make their own judgements on the data quality. It is not the MIDA policy to correct errors for this reason; on the other hand it is necessary to eliminate inconsistencies in the way data are presented, to ensure the optimum visual output for the end user. For example, a river layer obtained for the atlas consisted of rivers that were broken into hundreds of line segments, for no obvious reason. If a user were to select a river in the atlas, only a small fraction of that river would be highlighted. These line segments had to be joined together to improve the ease of access for users in the atlas.

Data Display

The extent of attribute information to display from any one dataset must also be taken into consideration. In most cases this must be decided with the data provider, particularly if a dataset contains sensitive (e.g. names and addresses of fish farmers) or proprietary information. Data can also be generalised spatially, as discussed later.

In terms of the cartographic display of thematic layers, the sheer number of individual layers in the atlas makes it challenging to select a range of colours and feature representations that are appealing as well as nonconflicting with each other. For example, polygon colours of Special Areas of Conservation, Special Protection Areas, and National Heritage Areas must complement each other, as well as be semitransparent in order to allow users to see where different protected areas overlap. Therefore, time must be devoted to take into consideration how the layers may be viewed together when deciding how features are displayed.

Layers generated at very different scales cannot appropriately be viewed together without the possibility of misinterpretation. Scale factor limits must be declared in the atlas display, so that when a user zooms in on an area, the

data viewed are at a scale appropriate to the level of detail required (von Meyer *et al.*, 1999).

Decisions must be made to determine how best to manage data themes that come in varying levels of quality from multiple sources. For example, the coastline layer can be acquired from the Ordnance Survey Ireland, the UK Hydrographic Office, or a low-resolution version from the GEBCO product, among others. In order to provide options to users that wish to acquire the data, links to the various versions must be made available.

Metadata issues

Acquiring adequate metadata, assuming it exists in the first place, can be difficult in Ireland. The lack of quality metadata can render specific datasets virtually useless due to the uncertainty of data quality, content and reliability (Bartlett, 1999). For datasets without sufficient metadata documentation, that information must be chased up so that the Discovery metadata can be completed.

Finally, an update strategy needs to be developed in order to keep the atlas current, particularly with datasets that change often. This is the case with the aquaculture data held by the Department of Communications, Marine and Natural Resources, which is updated monthly. If for some reason the datasets cannot be updated regularly, the publication date of the displayed dataset and the update frequency are prominently displayed in the Discovery metadata.

7. Data Sharing and Intellectual Property

Data cost and licensing can prohibit what data can actually be displayed in the atlas, and solutions must be developed that both the data owners and the MIDA development team can agree upon. The solution varies for each dataset, and time is required for decision-making and negotiation.

The strong protectionist attitude of data owners in Ireland has merit, be it for financial or intellectual property's sake, however a cultural shift is needed to improve data access. In Ireland, as is the case in the United Kingdom, much of the available data reside with public bodies who must balance their commercial need to profit on their investment with their public duty to provide data (DEFRA, 2002). This cost, particularly for base datasets, is a hindrance to national research and development by limiting advances to those who can afford the data required. It is challenging to balance these financial needs with the less immediately tangible benefits that data sharing provides to research and development, which extend beyond the costs of creating the data (McCormack, 2003). Indeed, access to data is necessary to support vital coastal components such as commerce, environmental protection, health, and safety (Lockwood and Fowler, 1999). Data sharing also reduces redundant creation of datasets, increases efficiency, and can improve decision-making (DEFRA, 2002; Millard and Sayers, 2000).

By providing datasets to the MIDA, data owners are not asked to relinquish full control of their data because they decide how their data will be represented. Their data can be displayed in full or in part, with options to remove sensitive or commercial data within the data table. Modifications can also be made to reduce spatial data quality. For example, the resolution of vector or raster data can be altered using a GIS system so that the level of detail is reduced. For sensitive point data, locations can be generalised by drawing buffer zones around study areas. On the other hand, those data providers who would like a means of making their data freely downloadable from Internet will have that option within the MIDA.

All data owners are encouraged to submit some form of spatial data with its associated metadata, as the MIDA is substantially more than just a metadata database and is centred around a map display. Data owners are also encouraged to provide data to the MIDA by recognising the benefits that the atlas can bring to their organisation, in terms of the time saved in dealing with data queries and the potential for promoting the availability of certain datasets.

8. Conclusions

Response to date has indicated a strong level of support for the Marine Irish Digital Atlas. While welcoming the atlas, some users have expressed concerns that the scale of the project is overly ambitious. A step-by-step approach is necessary to ensure the successful implementation of the project within its planned timeframe. The project, which is on schedule to date, is due for completion in September 2005. It is envisioned that the final system will include the priority datasets identified by the end users, with the most highly desired technical features incorporated. The sustainability and future expansion of the atlas will be dependent on future funding. Thus it is imperative to demonstrate the value of the atlas for promoting education and awareness as well as better management of the coastal and marine environment.

While providing a web-based informational resource for all things coastal in Ireland, the MIDA will aid in resolving some of the data access issues that are common among professionals. By improving data access and providing information and links to key organisations, the MIDA can enable better management of coastal resources for the entire island of Ireland. Data owners will benefit from the promotion of their organisations and data holdings, and anyone involved in working with coastal data will profit from the time saved in searching for datasets relevant to their work. The atlas can also be of benefit to the educational community and general public by providing a resource that raises awareness of coastal issues. Ultimately the MIDA's success depends on the cooperation of data owners in supplying data and information, and the voices of the coastal community in providing valuable feedback on how the MIDA can best meet their needs.

*The MIDA Prototype 2 is due to **go live** on the web in **Autumn 2004**. The current prototype is available for viewing by appointment. Please contact Liz O'Dea (l.odea@ucc.ie) or Ned Dwyer (n.dwyer@ucc.ie) for further information. To keep*

updated on the progress of the Marine Irish Data Atlas, please refer to <http://mida.ucc.ie>.



This work is currently funded for the period from September 2002 – August 2005 by the Higher Education Authority (HEA) of Ireland under the PRTL 3 Programme, as part of the National Development Plan 2002 – 2006.



Bibliography

- Bartlett, D.J. 1999. Working on the Frontiers of Science: Applying GIS to the Coastal Zone. In *Marine and Coastal Geographical Systems*, edited by Wright, D.J. and Bartlett, D.J. Taylor & Francis: London, pp.11-24.
- Cahill, B., Kennedy, T., and Ní Cheileachair, Ó. 1999. Managing Marine and Coastal Data: A National Oceanographic Data Centre Perspective on GIS. In *Marine and Coastal Geographical Systems*, edited by Wright, D.J. and Bartlett, D.J. Taylor & Francis: London, pp.255-260.
- Cummins, V., O Mahony, C. and Connolly, N. (In press). Integrated Coastal Management: Review of Best Practice for Ireland. Heritage Council.
- Department of Environment, Food, and Rural Affairs (DEFRA). 2002. *Delivering Integrated Marine Mapping for the UK: Report of DEFRA-funded workshop held at Church House, London, 11 September 2002*, compiled by F.L. Franklin, DEFRA, Burnham Laboratory.
- Dwyer, N., O'Dea, L., and Cummins, V. 2002. Work Package 1: Data Model, Hardware, and Software Requirements for the Marine Irish Digital Atlas.
- Dwyer, N., O'Dea, L., and Cummins, V. 2003a. The Marine Irish Digital Atlas: a Web Portal to Coastal and Marine Data in Ireland, paper presented at CoastGIS 2003, Genova, Italy. <http://mida.ucc.ie/assets/documents/MIDA_CoastGIS2003.pdf>
- Dwyer, N., O'Dea, L., and Cummins, V. 2003b. Work Package 2: Metadata Review and Recommendations.
- Lockwood, M. and Fowler, C. 1999. Significance of Coastal and Marine Data within the Context of the United States National Spatial Data Infrastructure. In *Marine and Coastal Geographical Systems*, edited by Wright, D.J. and Bartlett, D.J. Taylor & Francis: London, pp.11-24.
- McCormack, B. October 2003. *Irish Spatial Data Infrastructure*, paper presented at Irish Organization for Geographic Information (IRLOGI) 2003 Conference. <http://www.irlogi.ie/pdf/ISDI_IRLOGI_2003.pdf>
- Millard, K. and Sayers, P. 2000. *Maximising the use and exchange of coastal data: a guide to best practice*. CIRIA: London.
- Von Meyer, N., Foote, K.E., and Huebner, D.J. 1999. Information Quality Considerations for Coastal Data. In *Marine and Coastal Geographical Systems*, edited by Wright, D.J. and Bartlett, D.J. Taylor & Francis: London, pp.295-308.