

Geo-Seas

Pan-European infrastructure for management of marine and ocean geological and geophysical data



D11.4: List and definitions of DTM products with guidelines for description and representation based on user requirements

Organisation name for lead contractor for this deliverable: University College Cork

Project acronym: Geo-Seas

Project full title: Pan-European infrastructure for management of marine and ocean geological and geophysical data

Grant Agreement Number: 238952

Start date of project: 1st May 2009

Co-ordinator: Helen Glaves



Deliverable number	Short Title
11.4	List and definition of DTM products
Long Title	
List and definitions of DTM products with guidelines for description and representation based on user requirements	
Short Description	
This document provides a guide to the issues, sourcing, specification and production of prototype bathymetric digital terrain models. It also provides insights into the key factors that are involved in the generation and viewing of such terrain models which have been generated by trans-national data providers for various communities of users.	
Keywords	
DTM ; prototype ; fusion ; product ; bathymetry	
Authors / Organisation(s)	Editor / Organisation
Thierry SCHMITT (SHOM)	H. Glaves (NERC-BGS)
File name	
GS_WP11_D11.4_DTM_Guidelines_v5.docx	
Deliverable due date	Deliverable submitted date
Month 29	Month 41
Comments	

History				
Version	Author(s)	Status	Date	Comments
0.1	Thierry SCHMITT (SHOM)	DRAFT	30 June 2010	Editing
0.2	Eric MOUSSAT (IFREMER)	DRAFT	09 July 2010	Specification of DTM variables, processing, layers, etc.
0.3	Thierry SCHMITT (SHOM)	DRAFT	21 August 2010	Integration of preliminary comments
0.4	Thierry SCHMITT (SHOM)	DRAFT	7 September 2010	Editing
0.6	Gerry Sutton (UCC)	DRAFT	26 October 2010	Editing
0.7	Thierry SCHMITT (SHOM)	DRAFT	8 November 2010	Corrections and adding gridding methodology

0.8	Eric MOUSSAT (IFREMER)	DRAFT	24 February 2011	Integration of analysis of existing data sources. Additional analyses and specifications
0.9	Eric MOUSSAT (IFREMER) Thierry SCHMITT (SHOM)	DRAFT	03 March 2011	Editing
1.0	Gerry Sutton (UCC)	FINAL	11 April 2011	Editing including syntax and formatting.
1.1	Thierry SCHMITT (SHOM)	FINAL	10 June 2011	Updates
1.2	Thierry SCHMITT (SHOM) Gerry Sutton (UCC)	FINAL	05 December 2012	Reformatting with new template. Extensive syntax and grammatical edits.
2.0	Helen Glaves (NERC-BGS)	FINAL	13 June 2013	Final editing
3.0	Peter Miles	FINAL	18 July 2013	Proof reading
4.0	Helen Glaves (NERC-BGS)	FINAL	18 July 2013	Sign-off

Dissemination level		
PU	Public	
CO	Confidential, for project partners and the European Commission only	X

Executive Summary

The overall objective of the Geo-Seas project is to create an e-infrastructure for the sharing of marine geological and geophysical data. Bathymetric data are of fundamental importance within such an infrastructure as it is a foundation layer of information which supports many other layers as well as multiple end uses. However bathymetric data are collected and distributed by public or private organisations for a variety of applications giving rise to a diverse range of technical attributes (density, sampling method, accuracy, dataset size, formats etc.) and distribution policies. This complicates the process of extracting information which is appropriate to the needs of a range of end users. Work package 11.2 has focussed on developing a conceptual approach to enabling end-users to better discover the datasets that are available and to analyse multiple bathymetric datasets shared through a distributed infrastructure prior to gaining access to them.

This document lists the various requirements for bathymetric products as identified by the user community and the characteristics of existing data sources. It then discusses the technical (horizontal and vertical references, level of processing of the datasets, etc.) and administrative difficulties often related to the level of resolution of the dataset (data distribution policy) which constrain the broader distribution of bathymetric data. Concepts and specifications are provided with which bathymetric grids can be derived that will help to address a number of issues related to data providers data distribution policies (dataset resolution) whilst at the same time optimising grid sampling to satisfy a maximum number of potential uses of the data. Specifications for accompanying visualisation and downloading tools are also described in this document.

Geo-Seas deliverables D11.5 *DTM Prototype delivery for evaluation by users* and D11.5A *3D DTM viewer- end user manual* outline the prototypes and describe which have demonstrated that the concepts and specifications proposed in this document can fulfil most of the proposed data uses whilst respecting the use-policy requirements for the majority of data providers.

Content

1. Introduction	7
2. Definition of requirements	8
2.1. End-user requirements by category	8
2.2. Additional requirements	12
2.2.1. Content requirement (parameters)	12
2.2.2. Common Reference System (CRS) and datum requirements	12
2.2.3. Spatial and time coverage requirements	12
2.2.4. Gridding processing requirements	12
2.2.5. Services requirements	12
2.2.6. Data organisation	12
2.2.7. DTM metadata requirements	12
3. Existing data sources	14
3.1. Hydrographic organisations	14
3.1.1. Context	14
3.1.2. Data set content	14
3.1.3. Resolution	15
3.1.4. Vertical Datum	15
3.1.5. Temporal aspects of coverage / age of soundings	16
3.1.6. Delivery of Datasets	16
3.2. Non-hydrographic organisations	17
3.2.1. Context	17
3.2.2. Dataset content	18
3.2.3. Dataset delivery	18
4. Specification for the distribution of bathymetric data via a distributed database for purposes of data fusion	20
4.1. Introduction	20
4.2. Dataset harmonisation specifications	21
4.2.1. Systems of reference	21
4.2.2. Grid geometry and resolution	23
4.2.3. Content specification	27
4.3. Source metadata specifications	28
4.4. Complex DEM processing workflow	31



4.4.1. Gridding method.....	31
4.4.2. Product metadata	35
4.4.3. Quality evaluation.....	36
5. Product handling services specifications	37
5.1. Viewing services	37
5.2. Downloading services	41
6. Prototype development	44
Appendix A.....	47
Appendix B.....	48
Appendix C Terminology.....	49

1. Introduction

The overall objective of the Geo-Seas project is to create an e-infrastructure for sharing marine geological and geophysical data. Bathymetric data sets (composed of arrays of depth soundings) are of primary importance for such an infrastructure as they combine to form one of the fundamental layers of information which can and do support many different end use functions across a range of themes and topics (see WP10 and Section 2.1 of this document). However the actual bathymetric data sets that are collected and distributed by public or private organisations often possess technical characteristics and properties (e.g. sounding density, sampling method, accuracy, dataset size, format etc.), particular to the circumstances or specific operational requirements at the time of collection. These datasets are also frequently the subject of one or more disconnected dissemination policies. Individually or a combination these factors often impede the efficient extraction and re-use of information appropriate to a wide range of end-user needs and scenarios.

To overcome these limitations WP11 has developed new services and products to facilitate access to the data by generating harmonized geophysical and geological products and services. Within WP10 the focus is on the development of 1D and 2D geological standard products and mapping services. The challenge for WP11 is to explore the options for providing more complex services in a distributed infrastructure of multi-disciplinary data centres. This allows end-users to evaluate the usefulness of products such as a DTM (Digital Terrain Model) and access them in a common simple and understandable manner.

Grid systems are powerful tools for harmonisation and reduction of the complexity of spatial datasets, and for communication of spatial information [1]. The main objectives of the current WP11.2 subtask are:

- to define the characteristics of bathymetric grids (DTM) required by end-users according to the results of the survey carried out by the WP10,
- to specify procedures which allow the combination of different sources and features of data in a target DTM
- to develop a pilot DTM as a case study to evaluate the feasibility of generating such grids and services in accordance with end user requirements and recognising the constraints resulting from the use of multiple of datasets from multiple data sources (data being made available by Geo-Seas partners).

In this document we aim to:

- Synthesize user requirements for bathymetric DTM products according to the results of the WP10 questionnaire
- In the context of a distributed infrastructure, document a tested methodology to create DTMs across international borders using multiple data sources
- Define common viewing and downloading services enabling users to better discover the data available and to obtain data products fitting their needs
- Develop a pilot DTM according to data source availability and partners capabilities
- Propose guidelines for the computation and the visualisation of DTM products from web based services

2. Definition of requirements

The end-user requirements for the DTM products being delivered by WP11 have been developed using the responses provided to the WP10.1 questionnaire by users. These requirements have also been supplemented with the expert knowledge of the relevant project partners. However, the needs identified by the end-users which relate specifically to maritime navigation are considered beyond the immediate scope of this work as these activities lie within the formal remit of official bodies such as national hydrographic organisations.

2.1. End-user requirements by category

From the WP10 questionnaire responses it is possible to group end-users needs into broad categories that share similar DTM specifications. The following list outlines the major communities of users and the main DTM characteristics (i.e. scale/resolution, vertical datum, coordinate system and specific processing requirement such as interpolation) that corresponds to each. These are classified by relative importance from the results of the WP10 questionnaire and are summarised in Table 1.

Geology/Sedimentology/Morphology

Generally this community aims at building:

1. **regional models** at scales of tenths of kilometres. Sedimentology and morphology (e.g. for geo-hazard work) require working scales of hundreds of meters (typically the scale of intermediate bedforms)
2. In order to obtain a reference surface that is continuous across the land-sea interface, this community prefers to use Mean Sea Level (MSL) as reference vertical datum. In terms of horizontal datum WGS84 is generally preferred along with projected coordinates (allowing measurements of morphological geometries in metric units). For this community the data products are mainly used in research and development contexts, and to a lesser extent, for teaching or outreach.

Oceanography/Hydrodynamic/Climate change

This community requires DTM products as a fundamental pre-requisite for building numerical models which use non-linear equations to predict the physical characteristics and parameters for currents, waves, sea level and sediment dynamics.

Note that the output of these models are very widely used in forecasting, hind-casting and operational prediction, as well as being instrumental in helping to develop advanced understanding of all marine environments in terms of structure, function and dynamics.

Their specific requirements are:

1. **Nested models** of different scales corresponding to resolutions which vary with depth (i.e. the shallower the water, the higher the resolution). Required resolutions vary from 1 degree down to 10 m which can be summarized as follows :
 - tidal to sub-tidal areas (a few meters down to a sub-metric density)
 - coastal areas (depths ranging from 10 to 30 meters)
 - upper shelves (depths ranging from 30 to 100 m)

- lower shelves to break of the slope (depths ranging from 100 to 500m)
- 2. Vertical datum required as a geopotential surface (i.e. geoid). The Lowest Astronomical Tide (LAT) surface which is the datum recommended for data distribution by the International Hydrographic Office (IHO) and in use by most of the HO is sufficient according to the results of the questionnaire. MSL can also be appropriate in many contexts. In terms of horizontal datum the WGS84 is preferred, with geographic coordinates.
- 3. Voids in the spatial coverage should be limited, and artefacts must be limited or virtually non-existent to ensure **spatial continuity** is preserved in order to limit the propagation of errors throughout the hydrodynamic model.

Ecology/Habitat mapping/Fisheries

This community uses DTMs to define the spatial distribution of marine environments (habitats) as a function of parameters either derived from the morphology such as the slope, aspect, roughness, BPI (Bathymetric Profile Index) or dependent on depth such as the penetration of light, the dynamic of the environment, the nature of the substrate, the time of immersion (in relation with tides).

DTM requirements are as follows:

1. Very high resolution in the order of 1 (tidal zone) to 100 m (shelf).
2. Vertical datum specifications are not clearly stated in the questionnaire. However considering the strong links between these themes and the processes affecting the coastal/terrestrial domain it appears that the MSL level can be considered as appropriate.
3. For the calculation of the parameters deriving from depth, preference is given to **projected coordinates** in metric units (horizontal datum WGS84 is generally suggested).
4. Generation of the DTM must prevent over-interpolation along with minimal artefacts within the datasets. Spatial continuity is highly desirable (for usages related to geomorphological classification).

Dredging disposal/Mineral/Coastal and marine engineering

The major needs of this community are to be able to compare time series of bathymetric surveys and to be able to measure volumes of dredging or erosion / deposition or bed form migration.

1. The resolution (metric scale) and the accuracy of the bathymetric data must be displayed at the highest level possible.
2. A lowest level of interpolation will make sure that the level of accuracy of the data will be preserved. Again the presence of holes is permitted as the users prefer accurate data over continuous coverage.
3. The preference of this community for a specified vertical datum is not clearly defined. This is due to one part of the community being interested in issues relating to the safety of navigation (dredging disposal) and the other part of the community having strong relations with conventional, terrestrial engineering (coastal and marine engineering). This lack of consensus regarding vertical datum suggests that this community will require a



selection of tools and mechanisms to ensure consistent accurate shifting between specific vertical datums across wide geographical areas. Note that this is considered to be beyond the scope of the current project.

Community of users	Scale/ Resolution	Vertical datum	Horizontal datum and horizontal coordinates	Interpolation characteristics	By products	Depth value	Special requests
Geology Sedimentology Morphology	Regional (10s of km) Local (100s of m)	MSL	WGS84 projected coordinates	Smoothing should be limited	Seamless sea – land interface	mean	
Oceanography Hydrodynamic Climate change	1°, 1km, 100s m, 10s m	Geoid or LAT	WGS84 geographic coordinates	Smoothing and hole filling allowed		mean	Nested models with consistency between the different resolution
Ecology Habitat mapping Fisheries	1 to 100 m	MSL preferred	WGS84 projected coordinates	Smoothing must be limited. Residual (grid height – sounding height) must be minimal.	Geomorphology (slope, aspect, roughness, BPI) and dependant physical variables	mean	
Dredging disposal Mineral Coastal and marine engineering	Several meters	Not clearly defined	WGS84 no preference	Holes permitted. Minimal residual			Availability of time series for comparison

Table 1 User requirements according to thematic community

2.2. Additional requirements

Additional technical requirements which are not specific to the above categories have been determined from the results of the WP10 questionnaire and a subsequent specific follow-up WP11.2 questionnaire undertaken.

2.2.1. Content requirement (parameters)

- Depth value must be representative of the cell surface.
- In addition to the average, relevant values are often required (observed minimum depth, observed maximum, nearest sounding to cell centre).
- Non shoal biased depth is required unless no other value can be delivered.

2.2.2. Common Reference System (CRS) and datum requirements

- Refer to existing common conventions and initiatives in Europe (such as EMODNET Hydrography, Eurofleets, MyOcean, etc).

2.2.3. Spatial and time coverage requirements

- A qualification of the time validity of the product (e.g. for high energy areas where mobile sediments can modify the topography) is required.
- The potential to include **time as an additional dimension** to allow for the elaboration of DTM time series is required.

2.2.4. Gridding processing requirements

- Multiple levels (nesting) of resolution / scale of DTM must be consistent with regards to the usage (e.g. consistency of the volume of the water column at different resolutions).

2.2.5. Services requirements

- Making viewing services available to enable zooming and panning within the various levels of complex product.
- Providing methods for creating and visualising derived parameters such as morphological indices (slope, aspect, Benthic Position Index (BPI) etc.) and derived products such as contours on the fly.
- The ability to download the entire region or a subset of the created products (either the bathymetric grid or the any derived morphological products) with an OPeNDAP server.

2.2.6. Data organisation

- Grids must be downloaded in a common, INSPIRE compliant format (NetCDF-CF: Network Common Data Format with the Climate and Forecast convention).

Note: It is recognised that a range of common data formats for grids are in regular use in different end user communities, however these are not being considered in the current document in order to ensure interoperability (within the constraints of the Geo-Seas data transport format policy[13] [14]).

2.2.7. DTM metadata requirements

- The use of standard metadata to document the level of processing of the source data is important (and specifically the state of the tidal corrections and associated vertical datum).

-
- A qualification of the period of time over which the product remains valid (the shelf life) is necessary. This will be dependent on the source of the data, for example, for high energy areas with mobile sediments (from the tidal zone to shallow waters of a few tens of meters) the period of validity will be much shorter than that for more stable environments.
 - Standardised metadata appropriate to the method of generation and the quality of the composite products is required. A methodology for tracing the steps involved in the product creation is needed in order to identify and understand the lineage of the DTM product. (i.e. the concept of metadata related to a data/product that subsequently becomes incorporated into a global product).

3. Existing data sources

This section deals with the existing sources of data which have been identified for the development of advanced DTM products and services and how these relate to the end-user requirements identified as part of the WP10 user survey.

There are two major public sources of bathymetric data in terms of acquisition, processing and usage in Europe. The following description is based on experiences of data exchange and compilation both between the project partners and these organisations and also within the framework of a number of other relevant projects.

3.1. Hydrographic organisations

3.1.1. Context

Official bodies such as hydrographic offices (“HO”s), Port & Harbour Authorities as well as state agencies and other bodies in charge of coastal planning collect bathymetric data for the purposes of ensuring the safety of navigation.

In many countries the national nautical chart series produced by these bodies are often the largest scale publications available which show the configuration of the seabed offshore. National hydrographic offices are generally responsible for the delivery of bathymetric information for their national waters and it is usually the only source of bathymetric information that has a legal value.

The equipment used for hydrographic surveys has evolved through time from the lead-lines used up until World War II, through vertical single beam echo sounders used during the 70s, to the oblique single or multi-beam echo sounders used today. The standards for conducting hydrographic surveys for the purposes of data collection have similarly evolved to take into account the use of the data for purposes other than for the security of navigation, such as protection of the environment [2]

From the extent and coverage of the nautical charts currently available it may appear that the seafloor has been fully surveyed by the national hydrographic offices. However this is generally not the case because the density of the source data and the frequency of repeat surveying of an area are functions of the navigational requirements for a given nautical area [2].

3.1.2. Data set content

To produce navigational charts, soundings are filtered for legibility and security (shallowest sounding in a given neighbourhood). Unless explicitly specified, digital products are usually biased in the same way.

Due to technical constraints (e.g. size of database, confidentiality) and also as a result of common hydrographic office practices, digital data are frequently delivered as a set of selected soundings in tiles (e.g. of 1°x1°) or as digital terrain models, both of which having been derived from multiple surveys. These data sets are usually produced according to the principles used for the representation of the bathymetric information on nautical charts and are therefore biased (selection of the shallowest soundings in a given area e.g. 25 m in water depth less than 100m) to ensure the safety of navigation. This explains those cases where the observed water depth is significantly greater than that indicated on charts or in data sets made available by national hydrographic offices (set of soundings, isobaths or DTM) and especially in high energy areas or areas with frequent dredging.

3.1.3. Resolution

This spatial density of soundings varies from full coverage in areas of water depth less than 40 m such as harbours and shipping channels (IHO special order and order 1a [2]) to a discrete coverage with line spacing varying as a function of the water-depth from a maximum of 25m in areas shallower than 100m (order 1b [2]) to a spacing representing 4 times the average depth in areas generally deeper than 100m where a general description of the sea floor is considered adequate (order 2 [2]).

Note that some areas even in shallow waters may have almost no data because no shipping is expected to transit in these areas. This is also the case for the foreshore.

The resolution of soundings made available as digital products by hydrographic offices and similar bodies depends on:

- The density of sounding available in their respective data holdings [2]
- The filtering applied when extracting data which is also dependant on the data policy.

3.1.4. Vertical Datum

The value of each sounding is given as the distance either above or below a conventional vertical datum known as “Chart Datum”. Technical Resolution A2.5 [3] of the IHO states that LAT (Lowest Astronomical Tide) shall be adopted as “Chart Datum” where tides have an appreciable effect on water level. Chart datum is established based on local water level measurements at discrete locations and is selected as a surface that is so low that the tide will not frequently fall below it but not so low as to be unrealistic and only gradually varying between adjacent datums.

Most nations, if not already using approximate LAT, are moving towards its use [4]. There are noticeable exceptions in EU where the tide is practically unobservable or minimal:

- In the Baltic Sea where countries use the Mean Sea Level (MSL)
- In Mediterranean Sea where other levels are used such as the Mean Lower Low Water (MLLW) in Greece.

Despite the apparent homogeneity in the definition of the chart datum (in particular if we consider that chart datum should not differ significantly from LAT where tides have minimal effect on the sea level), it does not represent a seamless reference surface as it varies from location to location. Offsets between datums are in the order of several decimetres. In the Baltic Sea it can be expected that the maximum differences of chart datum levels between neighbouring countries are not more than one decimetre [5]. In countries such as the UK and France where tides have an appreciable effect on sea level, local chart datums corresponding to tide areas have been defined along the coasts (e.g. there are 16 local chart datums between Belgium, Spain and France).

Chart datums have also changed through time since the 19th century when they were first established. These changes are a function of the improvement of the technology used to define them and also take into account relative sea level rise (e.g. at Brest, France, the reference level was changed by 50 cm in 1996) or continental uplift (e.g. 6 to 7 mm per year in Gulf of Bothnia).

Several initiatives to harmonise chart datums at national level (eg. UK, France) or at regional level are currently in existence (e.g. the BLAST project for the North Sea; the

ChartdatumWG of the Baltic Sea Hydrographic Commission) but during the transition phase data collected before the implementation of the new reference system, are still in use.

At the present time, depths are not related to a common reference level in Europe.

Note :

Chart datums are connected to the national height vertical datum. Height vertical datums vary from one country to the next reflecting errors of levelling, different methods of data reduction, differences in the epoch definition of the national levelling networks. Transformation parameters of national height reference systems to the United European Levelling Network [6] see Figure 1) show that offsets between height datums of 2 neighbouring countries in Western Europe is generally less than 20 centimetres but significant offsets do exist e.g. between France and UK (> 40 cm), France and Belgium (>180 cm), Belgium and Netherland (230cm)



Figure 1. Transformation parameters from national heights to UELN height in Europe (values in cm).

3.1.5. Temporal aspects of coverage / age of soundings

One of the noticeable aspects of European hydrographic offices is their age. For example, the French and the UK hydrographic organisations were established in 1720 and 1795 respectively. As a result data sets such as depth sounding data from surveys conducted in the 19th century using lead-lines are still made available, and there are currently very large areas which are covered by surveys that are in excess 50 years old.

3.1.6. Delivery of Datasets

The conditions for data access vary from one organisation to another and range from free access to the bathymetric data (e.g. Ireland) to the more commonly used license at cost. In addition, hydrographic offices are often attached to national defence organisations and this generally means that they have very conservative data policies.

Despite the implementation of the INSPIRE Directive it is still difficult to access the data held by these hydrographic organisations. The conditions of access and use applied to the data often prevent users from accessing the full set of soundings. The data are commonly

delivered either as a subset of the soundings sampled at various resolutions (e.g. 25 m in coastal waters of France) or as a DTM derived either from survey data or nautical charts.

International Hydrographic Organisation (IHO) Publication S-57 [6] defines the IHO Transfer Standard for Digital Hydrographic data. This standard was designed for the exchange of digital data between hydrographic organisations and for the distribution of hydrographic data to end-users (e.g. environmental management organisations). However, to date, S-57 has almost exclusively been used for encoding Electronic Navigational Charts (ENCs). One of the reasons for this is that S-57 is not a modern standard that is widely accepted in the GIS domain. As a result the IHO is developing a new standard (S-100) that will comply with the ISO 19115 series of geographic information standards. This new standard will provide the basis for the development of a wide range of digital products and transfer standards using hydrographic data. These will include imagery, gridded data, 3-D and time-varying data (x,y,z and time) as well as applications that go beyond the scope of traditional hydrography according to the IHO [7]. The S-100 standard itself does not mandate a particular encoding format. This means that the developers of product specifications can decide on the most suitable encoding standard for their applications. However, the S-100 does encourage the use of a generic code format in order to promote data exchange between potential users. For imagery and gridded data (regular or not), a suitable generic format would be using XML for describing the metadata and an appropriate value element encoding mechanism such as the Hierarchical Data Format (HDF version 5) for the data itself. HDF is an object oriented format suitable for all types of coverage data which includes sets of point and TIN triangles. This format is the basis of NetCDF. S-100, version 1.0.0, which was published in January 2010.

Up to now the formats of the datasets made available to end-users have varied according to the provider:

- Soundings are often delivered in ASCII (x,y,z)
- DTM are delivered in ASCII (x,y,z), ESRI grid, GMT GRD, GeoTIFF, USGS DEM (Norway) as well as in proprietary formats.
- Shape files for use with web mapping services (WMS)

3.2. Non-hydrographic organisations

3.2.1. Context

This category of organisations includes oceanographic institutions, universities, geological surveys, geophysical observatories, councils, and other state and semi-state bodies. These organisations primarily collect and use bathymetric data for various applications such as the exploration of the ocean and seas, the exploitation of marine resources, the study of climate change impacts, operational oceanography, marine environmental monitoring and management of the coastal seas.

Bathymetric surveying started in the 1960s using single beam echo sounders. Since then the spatial coverage of this type of data has expanded exponentially mainly due to the development of the multi-beam echo sounder (MBES) and improvements in positioning systems during the 1980s. The areas surveyed by organisations using the MBES represent hundreds of thousands of square kilometres of the European margins. Several of these institutions have also been responsible for the surveying and mapping of their national exclusive economic zone (EEZ) and continental shelf as defined under the United Nations Convention on the Law of the Sea (UNCLOS).

Recently, the number of surveys in shallow waters has increased in response to user needs for seabed habitat mapping, operational oceanography and coastal planning. In addition to single and multi-beam echo sounders, LIDAR and interferometric sonars are also being used by some of these organisations.

3.2.2. Dataset content

Unlike the hydrographic offices, these organisations do not share common standards for the collection, processing and delivery of data. The level of processing of the datasets which are made available varies from one organisation to another and from one application to another. The data being made available ranges from raw through to fully processed survey data, or may alternatively be provided as DTMs.

3.2.3. Dataset delivery

The conditions of delivery of the data are influenced by the available resources and the technical environment (e.g. processing software and database available) of the organisation. Two main factors determine the conditions of access to the data:

- the evolution of survey technologies which have multiple high resolution sensors producing large volumes of data that are often several terabytes in size
- the development of data policies, which are often poorly documented, that dictate the nature of the datasets that are made available to external end-users. These policies are generally determined by: 1) the restrictions applied to the data, 2) the will of the organisation to exploit the data they have collected, and 3) the need for the organisations to develop collaborative relationships with other institutions in order to support their surveying activities.

Table 2 below shows the formats that are commonly used by these organisations for data delivery. The metadata for these survey datasets are not systematically made available and the mode of access varies from one institution to another with several web portals currently delivering ISO 19115 compliant descriptions.

Survey Type	Format
Multi-beam/LIDAR (single survey)	<ul style="list-style-type: none"> - binary (either delivered directly from the acquisition system or from processing software (MBLDEOIH, or NetCDF) - XYZ ASCII files - High resolution grid for shallow waters either in ASCII or binary format from processing software (easy handling and preserving the high level of resolution of the sensors) - Low resolution grid in binary or ASCII format (partially alter the resolution to conform to the policy of diffusion) - Generally binary gridded datasets for LIDAR surveys (LAS format)
Single beam (single survey)	<ul style="list-style-type: none"> - rarely binary format - standardised (MGD77) or tailor made XYZ ASCII files
Combined datasets (data from a combination of sources – different acquisition systems and different level of processing on which an interpolation algorithm has been applied)	<p>Regular grid (either in geographical or projected coordinates) in one of the various format (generally dependent on the community of users)</p> <ul style="list-style-type: none"> - formats designed for spatial analysis and viewing with GIS applications : ESRI ASCII, ESRI GRID, shape, WMS/WCS, GeoTIFF - formats designed for computation such as hydrodynamic modelling (ie. not limited to viewing): GMT grid, USGS DEM (Norway Digital Web site), NetCDF (Ifremer) etc. - Various XYZ formats <p>Note that metadata related to the source of the node, the value assigned to a node or the quality of the estimation is seldom given.</p>

Table 2. Typical data formats used for the distribution of bathymetric data

4. Specification for the distribution of bathymetric data via a distributed database for purposes of data fusion

4.1. Introduction

The product developments being undertaken by WP10 are focused on 1D and 2D standard geological products and mapping services. Within WP11 the objective is to explore the possibilities for providing more complex services within a distributed infrastructure of multi-disciplinary data centres which will allow the end-user to evaluate the usefulness of products such as DTMs and then to access them in a common, simple and understandable manner.

In response to the user requirements identified in the user survey carried out by WP10 and the subsequent expert/user consultations, Work Package 11.2 aims to develop the concepts required to build a prototype harmonised DTM and associated services. This will improve the usefulness of DTMs for a variety of end-users (in particular the interoperability in trans-boundary contexts) and also facilitate accessing these DTMs in a distributed data infrastructure such as that being used for the Geo-Seas project.

Figure 2 below summarises the different steps in the development of a demonstrator including:

1. Harmonisation of data sources through dataset merging and data conformity checks
2. Standard composite multi-resolution grid processing (including metadata and statistics)
3. Advanced visualisation services: multi-scale 2D and 3D viewing for analysis of the DTM combined with derived parameters or layers from other sensors (including time-dependent ones)
4. down loading services (e.g. data extraction) and processing services

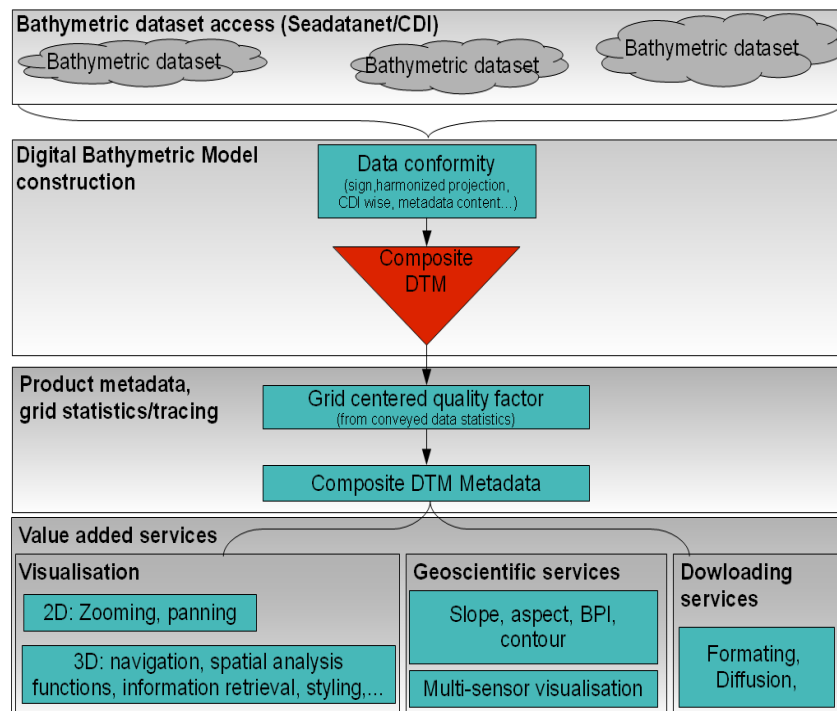


Figure 2. Conceptual workflow for product development from distributed data and associated services

4.2. Dataset harmonisation specifications

4.2.1. Systems of reference

Coordinate reference systems (CRS)

The choice of the coordinate (geographic angular Latitude, Longitude or projected metric) system affects two different steps of the processing flow:

- Data integration: compilation of data delivered as set of soundings and regular DTM
- Viewing and down loading services.

While the choice for the second step (i.e. equal angle, equal distance, equal area projection or no projection) is mainly driven by considerations related to the end-usage, the choice for the first step is driven by statistical considerations related to the sampling process. (Note: the depth expected by end-users who completed the WP10 questionnaire is the average value assigned to the surface unit of the sea-floor represented by the cell area of the DTM). All providers and end-users working with data from European seas should be using the same definitions for cell dimension and location to avoid discrepancies in the estimation of the depth and its associated accuracy. This implies that a common definition of the CRS and of the spatial relationship of the grid attached to it is required.

Geographic (angular) coordinates are very convenient for data handling. However data delivered in regular grids that are defined in geographic (angular) coordinates will have cells

of unequal area makes it difficult to compare values and statistics of cells at different latitudes.

This is an issue which has already been considered by the INSPIRE working group, however recommendations concerning depth information are still rudimentary. Currently these are based on the recommendations of the 1st European Workshop on Reference Grids for spatial reference system and map projections for statistical analysis [8] and the INSPIRE Specification on Coordinate Reference Systems [9].

It has been proposed that the Lambert Azimuthal Equal Area ETRS89 – LAEA (see extent in Figure 3 below) is adopted. This projection is expected to be available in INSPIRE transformation services (but at the time of the recommendation, LAEA was only used on a sphere in some GIS software). However ETRS89 – LAEA is not recommended for large scale mapping. This has raised the question of the integration of grids derived from multiple sources designed for use on different scales.



Figure 3. Lambert Azimuthal Equal Area (ETRS89-LAEA) projection extent [9]

Different INSPIRE themes or applications [10] may recommend the use of other appropriate map projections but they are not available at the present time.

It is evident from the results of the WP10 questionnaire that most end-users and organisations making surveys at sea collect their data and store them using the WGS84¹ (Lat/Lon) coordinate system. In addition, geographic coordinate grids are also widely used as this avoids distortion and conversion of coordinates which in turn simplifies exporting and importing data.

¹ Note that in principle, ETRS89 is more appropriate but no conversion of WGS84 to ETRS89 is carried out by data providers at present because the expected level of horizontal accuracy does not usually require correction of the difference between the ITRS (precise WGS84) coordinates of a point and the ETRS89 coordinates. However using ETRS 89 is recommended for high accuracy surveying as the difference between the ITRS and the ETRS89 coordinates is about 25cm in 2000 and is increasing by about 2.5 cm per year (in relation to plate tectonics) ie. by 1m by 2030

To avoid difficulties in sourcing datasets and also inconsistencies in delivery from different data providers, the use of a regular grid with angular coordinates in WGS84¹ is a temporary alternative which will be adopted for demonstration purposes e.g. for the GEBCO 30" arc grid (See cell geometry and origin for compatibility). However, it is recognised that for the reasons outlined above this solution is not ideal for the European seas in high latitudes.

Depth datum

According to the results of the WP10 questionnaire, MSL is the most popular vertical datum for various applications including the requirements of hydrodynamic modelling. MSL is the vertical datum recommended for data DTM production.

Reciprocal conversion between the various datums and MSL is a task which requires the use of conversion grids together with their associated uncertainty values which are beyond the scope of the Geo-Seas project. However it is recognised that this issue will need to be resolved at the international level. In the absence of an appropriate conversion tool and the lack of availability of the requisite parameters for all datasets, the use of LAT is considered as a viable alternative.

LAT is the datum recommended by IHO [3] Technical Resolution A2.5 of the IHO for navigational charts. Where MSL is still in use (no or negligible tide), LAT would not differ significantly from MSL. Also data acquired from the main data providers operating in shallow water areas (hydrographic offices, harbours etc.) usually refers to "Chart Datum". The adoption of LAT will enable soundings to be used to generate a consistent surface (provided variations less than 50cm are acceptable) without conversion of datum.

Time dimension

There may be a requirement to be able to generate a series of products at various time intervals (4D) from the same geographic area, but there are currently no identified use cases or appropriate datasets to fulfil this requirement. In order to enable future interoperability, the rule for data exchange is to use UTC and not local time. This point will be considered in the specification of the format as well as in the metadata.

4.2.2. Grid geometry and resolution

Grid geometry

For statistical purposes, a reference grid must be regular and easy for users to manipulate [10]. Regular grids are in general use in the ocean community (e.g. GEBCO [11] and EMODNET-Hydrography) and can be handled by all current software and applications.

Regular grids appear a de facto standard for data exchange which is appropriate for the purpose of the deliverable D11.4.

Grid resolution

The DTM resolution should be considered as a function of:

- The depth range of interest for the end-user applications which determines the definition of depth expected by the end user i.e. the pixel size (resolution) of the seafloor required for a specific application:
- The density and resolution of data collected by sensors in these depth ranges
- The mechanisms used to bring the compiled data sets to the level of resolution expected by the end-users.

The requirements vary as a function of the depth range for applications such as seabed habitat mapping and hydrodynamic modelling: the greater the depth, the lower the resolution. According to the results described in section 2.1, the typology of a DTM in terms of resolution can be summarized as follows:

Environment	Depth range	Horizontal Resolution
Coastal	10 -20 m	Few tens of centimetres to few meters
Upper shelf	Below 100 m	Ten to few tens of meters
Shelf	Below to 200 m	Hundred to few hundreds of meters
Slope and abyssal plains	Above 200 m	Few hundreds of meters to kilometers

Table 3. Typology of DTM resolution with respect to water depth.

However the performance of sounders has recently increased (see table below of possible DTM resolution as a function of water depth, speed, beam resolution and spacing for various MBES. Source : Ifremer, H. Bisquay). Current and future systems should allow even higher resolutions than those specified in the following table.

	Depth	Speed	Swath width	E/R cycle	Beam resolution along	Beam resolution across	Sounding spacing along	Sounding spacing across	Grid resolution
	m	knots	m	s	m	m	m	m	m
Reson Seabat 7125 (ROV Victor)	20	1	80	0,1	0,35	0,17	0,05	0,16	0,2
Kongsberg/Si mrad EM710	20	8	100	0,2	0,17	0,35	0,41	0,25	0,4
Kongsberg/Si mrad ME70	100	8	350	0,5	5,2	5,2	2,1	1,8	2

Kongsberg/Si mrad EM302	500	8	2500	2	8,7	17,5	4,1	5,8	5
Reson Seabat 7150HF	1000	8	5000	5	8,7	8,7	5,1	5,7	5
Kongsberg/Si mrad EM122	4000	8	20000	15	70	140	31	46	50

Table 4 Modern acquisition systems resolution capacities

In addition, the facility for multiscale viewing is required (e.g. for geomorphology) and must be taken into account in the definition of the grid system (e.g. for ascertaining spatial relationships between features at the metric scale such as pockmarks, mud volcanoes and hectometric to kilometric geological features such as fault and folds). This also affects the viewing and display functions that may be integral to the GIS software, and also how consistency is maintained in the computation of depths moving from one level of resolution to another, especially when merging datasets of different source and resolution. To facilitate the combination and merging of source data with different resolutions, the 1st European Workshop on the European reference grid has recommended adopting a hierarchical grid system.

Based on this recommendation and on the levels of resolution specified by end-users, as well as on the performance of measurement tools, a hierarchy of resolution is proposed (Table 5), which will be tested as part of the prototype.

Data providers that prefer to deliver data as grids (rather than data points at the full density) will have to select the level of resolution corresponding to the local delivery policy.

If no projection is adopted (as preferred above):

Resolution level	Mesh size in fraction of minute of arc (multiple of 2)	Corresponding value in meter
7	1	1852,00
6	4	463,00
5	16	115,75
4	64	28,94
3	256	7,23
2	1024	1,81
1	4096	0,45

Table 5. Proposed level of resolution (no projection)

If a projection such as LAEA is adopted :

Resolution level	Mesh size meter
7	2096
6	512
5	128
4	32
3	8
2	2
1	0,50

Table 6. Proposed levels of resolution (with a projection)

It should also be noted that a hierarchy based on a power of 10 (i.e. 1, 10, 100, 1000 etc.) is not fully appropriate as intermediate levels are required.

Cell geometry and origin

In order to achieve consistent and comparable grid cells, the grid system should have a clear and simple relationship to the coordinate reference system in order to facilitate data delivery and exchange.

The grid will be cell based i.e. the thematic attributes (depth etc.) are geometrically associated with the centre of a grid cell. In order to avoid problems of offset between grids of the same resolution or of a resolution being a multiple of the first level, a convention must be adopted.

The origin of the grid coverage is the SW corner of the bounding box.

To position the bounding box in a unique way, it is proposed to adopt the following convention related to :

- the Greenwich meridian (0°) in longitude
- the Equator (0°) in latitude

If the grid were extended to this point the origin of the cell should have (0,0) and the coordinates of the cell would be (LX/2, LY /2), LX and LY being the dimensions of the cell in X and Y.

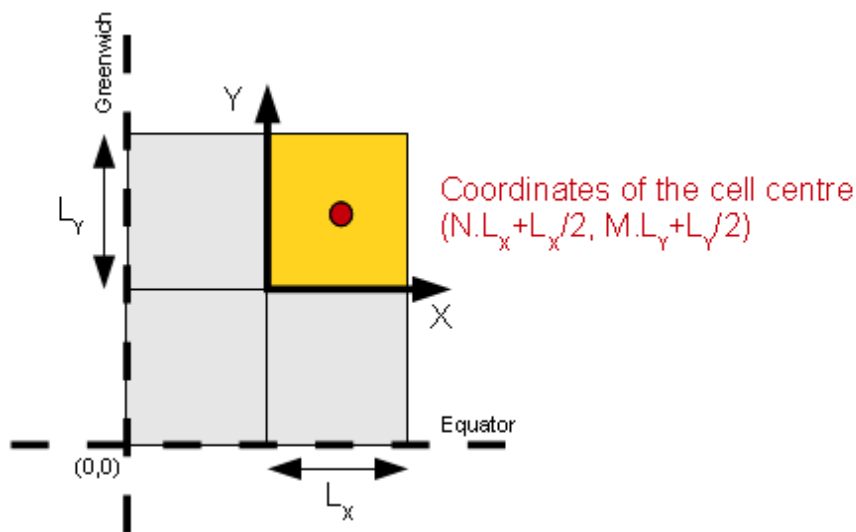


Figure 4 Cell/grid geometry

It should be noted that adopting this convention makes cells of level 7 resolution compatible with the global 1 minute GEBCO grid.

4.2.3. Content specification

Depth

Different algorithms can be applied as a function of the estimated depth (product depth) expected by the end-user when data are gridded. Depth can be the nearest sounding value, the median value observed in the cell, the average, etc. This impacts upon the production of models at lower resolution because often the values resulting from these processes are neither consistent nor representative of the cell surface. For technical and (more frequently) policy reasons, bathymetric data are more and more frequently distributed in the form of a DTM. As a consequence a common definition of the estimated depth is required to enable inter-comparisons and to increase the reusability of DTMs (in particular in merging datasets), irrespective of which estimator of this depth was chosen by the originator of the dataset. The following specification is in line with the findings of the WP 10 survey:

- A depth value must be representative of the cell surface i.e. the “mean of the cell surface”
- A depth value is required that is not shoal-biased unless no other value can be delivered (in which case, this should be documented in the DTM metadata). It should also be noted that complex data products will not satisfy the requirement of safety of navigation (which is solely the responsibility of Hydrographic Offices).

Other layers

Other information will be assigned at each node of the grids and will be delivered as additional layers. Distinction must be made between:

- **Statistics:** this layer gives the basic statistics allowing aggregation with other sources delivered in the same way. This layer will be mandatory as part of the gridded file. For each cell :
 - mean,
 - min,
 - max,
 - standard deviation of the soundings,
 - number of soundings
 - **Metadata:** number of source datasets, (main) dataset identifier (represented by their CDI identifier to facilitate a link with the Geo-Seas data discovery services). This will allow the computation of additional layers using the CDI_ID such as the age distribution of the coverage of datasets or the horizontal uncertainty of the positioning system of the main source. Note that interpolated depth in a cell where no observed data are available must be identifiable. This layer will be available as part of the gridded file.
-
- **Derived parameters such as morphological indices:** slope, aspect, Benthic Position Index (BPI) which are derivatives that can be computed from the input bathymetric data by an appropriate tool on the client side (e.g. GIS) and which are scale-dependant i.e. dependant on the specific end-user need. Values computed for these parameters will not be archived with the corresponding bathymetric grids.

4.3. Source metadata specifications

Table 7 below refers to the metadata that have been identified to be useful for filtering the source data prior to the development of products. Of these the geographical extent of the data has been identified as being of primary importance, but other variables are required such as those related to the time at which the data have been collected or provided (this is particularly important for the use of data sets from high energy areas with mobile sediments or where there is a need to develop 4D products which are essentially 3D products created for different points in time for the same area). Most of the required metadata are included in the CDI discovery metadata which is being captured for the Geo-Seas and SeaDataNet projects. There are also other metadata which might be useful to include and these are also shown in Table 7 below.

Variable	Typical values/comments	Mapping with CDI discovery metadata of SDN	Discussion points
ID	A unique local identifier for the dataset at the distributing data centre to access metadata and dataset.	CDI ID	
Data reference provider	EDMO identifier	Originator or distributing data centre	
Instrument	Depth: Single beam, multibeam, lidar, other. Position : Indirectly determines the accuracy of the positioning system	Instrument type used to collect the data and positioning system	
Creation of the dataset	Date or year of the survey or gridding	Creation/revision date	
Revision of the dataset	Date or the year	Creation/revision date	
Start Date	Start date of the survey (single survey) or of the oldest survey used to produce the DTM	Start Date	
End date	End date of the survey (single survey) or of the most recent survey used to produce the DTM	End date	
Bounding	Bounding box, curve (track line) or surface (seafloor coverage) polygon	idem	
Horizontal CRS	CRS (preferably the EPSG code, see SDN list)	Datum coordinate system	
Vertical CRS	Vertical datum used for depth (e.g. Mean sea level, Chart datum, Lowest astronomical tide, not applicable, ellipsoid)	Vertical datum	Note: additional information would be necessary to be able to convert from one datum to another. This can either be included in the description of the dataset or in the usage metadata
Parameters	Usage parameters defining the depth values and attributes	No appropriate field in CDI	Can be included in the datasets (ODV and NetCDF) but would be useful at a usage level Tbd: P011 list of BODC parameters?
Shoal biasness	Flag used to state that data have been sampled in order to ensure the security of navigation. (1 is true 0 is false)	No appropriate field in CDI (philosophy initially designed for observation data : is added in the abstract (“what”) for single survey DTM	Can be added in the abstract field Alternatively, could be in the definition of the depth parameter Tbd: P011 list of BODC parameters?

Sampling and gridding method	Sampling and interpolation method used with processing parameters	No appropriate field in CDI (initially designed for observation data : is added in the abstract (“what”) for single survey DTM)	Tbd: methods and processing parameters in a procedure field
Resolution	Spatial resolution Time resolution	“Track resolution”: term not appropriate in CDI but field appropriate for discovery of spatial resolution. Idem for the time dimension	Tbd; as a function of the type of grid in relation to cell geometry
File format	In the framework of Geo-Seas/SDN, datasets used for the demonstrator must be transformed to the standard format specified by WP4 ie : ODV or NetCDF and will be indexed according to the rules of the infrastructure	Distributed format	

Table 7. Mandatory and useful source metadata

4.4. Complex DEM processing workflow

4.4.1. Gridding method

Context

Historically, algorithms of various complexity have been used to compute gridded products. However it is recognised that there are major differences between the terrestrial and marine environments in terms of both the spatial distribution of data points, and the equipment used to acquire them. While it is now commonplace for terrestrial survey areas to be regularly, densely and frequently sampled, the converse is true in the marine situation, where levels of resolution, survey frequency and sampling density (from sparse to “completely ensonified”, i.e. the seabed has been fully sampled with respect to the IHO S44 standards) vary widely.

Four main classes of gridding algorithms can be recognised as shown in table 8.

Method	Description	Advantages	Disadvantages
Triangular Irregular Network (TIN)	Irregular mesh of triangles	Acknowledges each of the data points	Is not smooth, can not be derivable, is computer intensive to compute. Not all applications can handle TINs.
Binning/linear method	Fill a regular grid using an estimator (average, median, mean, weighted mean)	Conceptually easy to handle Particularly well adapted for large amount of data (Multibeam, LIDAR) as it is moderately computer intensive.	Results depend on the spatial distribution of the raw data. Generally biased towards extreme data (spurious data)
Kriging/geostatistics	Fill a regular grid using an estimator built from a careful study of the spatial correlation in between the soundings	Optimal weighting scheme for gridded data Naturally provide an estimate of the uncertainty related to the interpolation	Needs operators with a complete understanding of the technique and the data (must respect mathematical assumptions, elaboration of the variogram) Computer intensive
Polynomial and spline functions	Fits locally polynomial bending surfaces through the data (with or without minimisation of the distance between the interpolated surface and the data sources)	Provides smooth surfaces particularly well adapted to derivations.	Is generally controlled by a smoothing factor. Does not honour all the source data

Table 8. Common gridding methods

It is generally accepted that that there is no superior interpolating method when analysing the average error of the resulting DTM in relation to a reference dataset. Generally the quality of an interpolation is dependent on striking a balance between the following attributes:

- The “desired” smoothing effect
- The proximity effect (honouring the source data points).

For practical reasons, in WP11.2 will also consider:

- The practicality of the output data structure (regular grid)

- Implementation of the algorithm in an open source solution
- Proven examples of uses of the algorithms for marine data.

Binning/linear methods have some clear advantages. Although some drawbacks have been identified, their effects will be limited if a large enough volume of data is available. Moreover, the EMODNET-Hydrography initiative has already gained some experience using a basic experimental algorithm which data providers can readily implement. This is based on the averaging of bathymetric data at 1/16th arc-minute for data transport toward the integrator (i.e. whoever is responsible for the production of the target composite DTM), which is then averaged to the target ¼ arc-minute as shown in Figure 5.

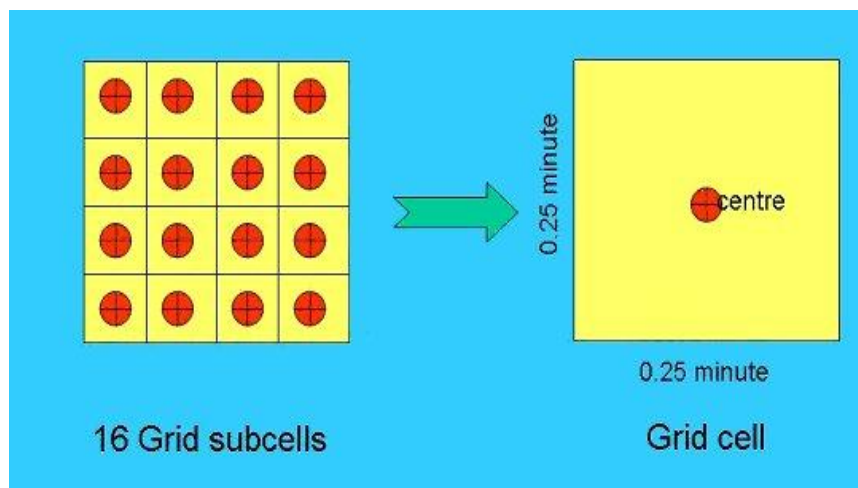


Figure 5. EMODNET-Hydrography gridding scheme

The advantages of this mechanism are that:

- It allows data providers to fulfil the requirements of their distribution policies by decimating source data in a simple and homogeneous way prior to delivery.
- It reduces the volume of data that needs to be handled
- It tends to balance the weight of the sources in the target surface grid product.

Once the data is merged, a second step can be undertaken to interpolate the missing values with an interpolator such as a spline function which gives a continuous and smoothed coverage. Various parameters can then be derived, the first of which being the mean depth of the surface of the cells of the target surface product at a ¼ arc-minute. The main advantage of this solution is that the associated essential statistics (mean, min, max standard deviation, number of soundings, CDI IDs etc) can be easily propagated from one level of resolution to another which allows the target grid to be easily updated when new datasets are made available.

Creating a complex DTM can be broken down into two distinct operations. The first part relates to the availability of data from the provider, and the second part is the amalgamation of the formatted data into a single DTM from interconnected sources that are distributed across a network.

Dataset Provision

A data provider will need to deliver data in one of two forms for purposes of creating a DTM:

- as raw source data (a point cloud)
- in a gridded format where:
 - the resolution corresponds to one listed in Table 5 which provides a relevant data density at a resolution which conforms to their data dissemination policy
 - the depth value is considered to be the mean of the soundings in the corresponding cell
 - grid and cell geometry have the characteristics listed in 4.2.2
 - depth and statistics identified in 4.2.3 will have to be calculated

Note: whatever the format the data provider has selected, correlated “source metadata” will have to be made available.

Merging grid cells

Considering two sources of data to be merged, the statistics for a given cell (same coordinates, same geometry) are as follows:

	Source 1	Source 2	Merged grid cell
Soundings number	N_1	N_2	$N = N_1 + N_2$
Average depth	$M_1 = \frac{\sum X_1}{N_1}$	$M_2 = \frac{\sum X_2}{N_2}$	$M = \frac{N_1 \cdot M_1 + N_2 \cdot M_2}{N_1 + N_2}$
Standard deviation	$\sigma_1^2 = \frac{\sum X_1^2}{N_1} - M_1^2$	$\sigma_2^2 = \frac{\sum X_2^2}{N_2} - M_2^2$	$\sigma = \sqrt{\sum_i \sigma_i^2 + \sigma_2^2}$
Maximum depth	$\max(X_1)$	$\max(X_2)$	$\max(\max(X_1), \max(X_2))$
Minimum depth	$\min(X_1)$	$\min(X_2)$	$\min(\min(X_1), \min(X_2))$

Table 9 Merging grid cells

Multiple resolution aggregation

The formulas presented in Table 9, can be generalised to a finite number of sources. These generalised formulas can be used to aggregate 16 cells of the resolution level L-1 to 1 cell in the resolution level L, with the only assumption being that the 16 cells (L-1) are totally included in the target cell (L), hence the importance of the hierarchy of Section 4.2.2.

Average depth product

As mentioned in Section 2, the end-user requires a meaningful estimate of the water depth. The “statistical” average water depth computed as mentioned in Table 9, may be useful (it will be given as part of the statistics layer associated with the final product). However this could result in a product that contains holes or gaps in the coverage and therefore presents a sub-optimal representation of the seafloor.

To provide a meaningful estimate of the average depth for an end-user specified resolution (L) it is proposed that:

1. A spline function (with experimentation on the tension values) should be used for the aggregated cells, at a resolution below the end user requested resolution (L-1) to use a spline function (with tension values to be experimented) to smooth and fill any gaps.
2. The value of the spline function at the cell centre is taken as the estimate of the water depth
3. All the (16) estimates of water depth at the resolution L-1 are averaged out to construct the grid cell at the end-user specified resolution.

Note that using a spline function induces a level of subjective interpretation (choice of the tension factor). However, the statistics provided with the final product will allow users to re-compute a surface.

Product grid processing

The following process is proposed in order to compute a product grid from formatted/non-formatted data sources (see Figure 6):

1. The user will select the available metadata for each dataset (based primarily on geographic variables but also on any other variables available in Table 7)
2. The Geo-Seas discovery and access service will request the data which fulfils the specified criteria from the relevant data centres.
3. The data centres will provide access to the data (source or gridded)
4. Grids will be generated at the highest resolution permitted for the data provided as source soundings
5. Merging/aggregation of source grids (with corresponding statistics)
6. Interpolation and smoothing using a spline function.
7. Product metadata generation

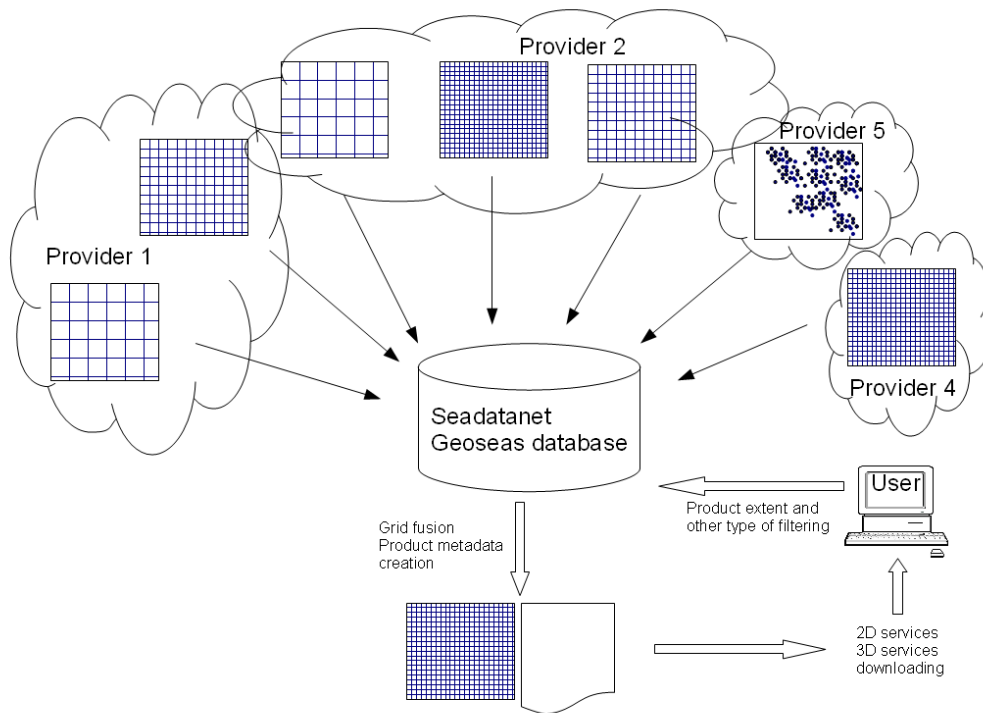


Figure 6. Selecting and merging data from distributed databases. (Note that the grids are scalable)

4.4.2. Product metadata

At the file level

Product metadata will need to include:

- A unique product identifier
- The lineage of the data included in the product
- Metadata information related to the individual source datasets
- Filtering rules used by the data provider for data selection
- Any ISO 19115 compliant metadata used to describe gridded products

At the cell level

For the purpose of the fusion/integration of the datasets it will be particularly important to keep a link to the source data after they have been blended. Becker et al. [12], defines a unique source identification number (SID) for each of the source data. By this means any attribution on a data point in the grid can be traced back to the original cruise and eventually back to the original data source. The EMODNET Hydrography preparatory action makes use of the same unique Common Data Index identifier developed by the European SeaDataNet project for this purpose. At present only information relating to the main data provider is preserved in the metadata information (per cell of the grid).

Using a similar approach for WP11, the relationship between the relevant datasets and the associated metadata will be recorded using a unique identifier, preferably that used in the CDI in line with the methodology adopted for the EMODNET Hydrography project.

4.4.3. Quality evaluation

4.4.3.1. Quantification of interpolated elevations.

The quantification of interpolated elevations has been outlined in the GEBCO user guide for the creation of bathymetric grids (in preparation).

Various methods are described but these are difficult to apply to compiled datasets especially DTM's :

- because error budgets are rarely made available by agencies responsible for the survey quality
- because these error budgets have a limited shelf life especially in areas of high energy (dynamic seabed)
- it is dependent on assumptions that have been made regarding the spatial distribution of the underlying data or on the reference datasets used to evaluate the interpolated elevations.

In addition, evaluation of the interpolated elevations must be easy to implement and carry out by the data originator who supplies the DTM. It must also be easy for end-users to understand and to compare between datasets.

- Calculating basic statistics such as the number of observed values and the average and standard deviation of measurements available in each cell is a first step. This provides a source data statistics layer which can be used for the purposes of evaluating the DTM.
- A second approach may be to compute the average difference between the source datasets involved in the computation of the DTM product.
- A third approach could be to calculate the difference between the DTM product and known references such as global grids. However, it should be noted that these grids are often at a relatively coarse resolution and may not be reliable

4.4.3.2. Managing biases between dataset intersections

Managing biases between dataset intersections can only be done at the level of the data provider, who is solely able to make a judgement about applying corrections or the identification of corrections that might potentially need to be applied. A mechanism to detect and report bias issues is the only way to improve product data quality.

5. Product handling services specifications

In response to user requirements identified in the WP10 questionnaire the portal being used to access the computed DTM will offer both viewing and downloading facilities.

5.1. Viewing services

2 D viewing needs

In line with the standards chosen for the Geo-Seas project, 2D viewing will be based on OGC Web services which provide standard GIS functions such as:

- zoom and extension with corresponding refinement or decimation
- transparency and/or texturing (in order to display several layers)
- cartographic layout/ GeoTIFF export
- allow other OGC services to be imported

Web mapping services (WMS) will allow end-users to view the DTM and derive parameters for use in a local GIS. Additional WFS/WCS will also allow users to identify source data for those DTMs derived from multiple surveys DTM.

3D viewing needs

In line with one of the user requirements identified through the WP10 questionnaire, one of the objectives of WP11 is to explore the potential to offer high resolution 3D viewing services for the purposes of:

- analyzing interactions between processes in the seabed or in the water column and topographic features by including additional geological data (e.g. geological samples, sediment properties etc.).
- in-depth evaluation of dataset content to assess the fitness for purpose of a DTM (resolution, coverage, quality) before downloading it from the Geo-Seas distributed data infrastructure.

To fulfil these general requirements, the system must provide:

- **Interactive, real time visualization in 3D** to give a relatively realistic portrayal (Virtual Reality) of the physical realities: topographical features at various resolutions, properties of the seafloor (derived parameters from DTM or additional properties) or of the water column or location of observations (e.g. marine samples)
- **Navigation** with a virtual camera to explore the 3D scene or to search for specific features.
- **Spatial analysis functionalities.** Basic functionalities include dimension measurements such as distance.
- **Information retrieval** allowing the use of visualisation as an interface to the source data.
- **Styling** for visual analytics e.g. colouring a surface as a function of value attributes and simulating environmental effects such as sun shading, as well as harmonised visualisation of combined scenes.
- In addition, the system must be able to generate animations.

Since the objective is to facilitate access to data held by Geo-Seas data centres, a suitable solution should have the following key properties.

- The application must be accessible and simple to use without the need for specific user training.
- It must be useable with standard input devices such as a keyboard and mouse.
- It should provide access to data for users that lack the financial resources for primary data acquisition.

It is also advantageous to limit the amount of effort that is required for data providers to prepare data. This means that the preferred solutions are those that facilitate viewing without the data provider having to produce an additional format beyond the common data transport format specified by the Geo-Seas partners. Further more detailed specifications will be developed in consultation with expert end-users.

Existing technologies

Visualisation is a four-stage process which includes:

- selection of non-graphical object representations in a repository
- generation of graphical representations (mapping stage)
- rendering of the generated elements to an image
- presentation of graphical objects using a display device.

In client-server applications, the lower level components are usually installed on one or more servers while the remaining tasks are handled by the client. According to their complexity clients are classified into thick, medium and thin clients (figure 7 and table 10, below):

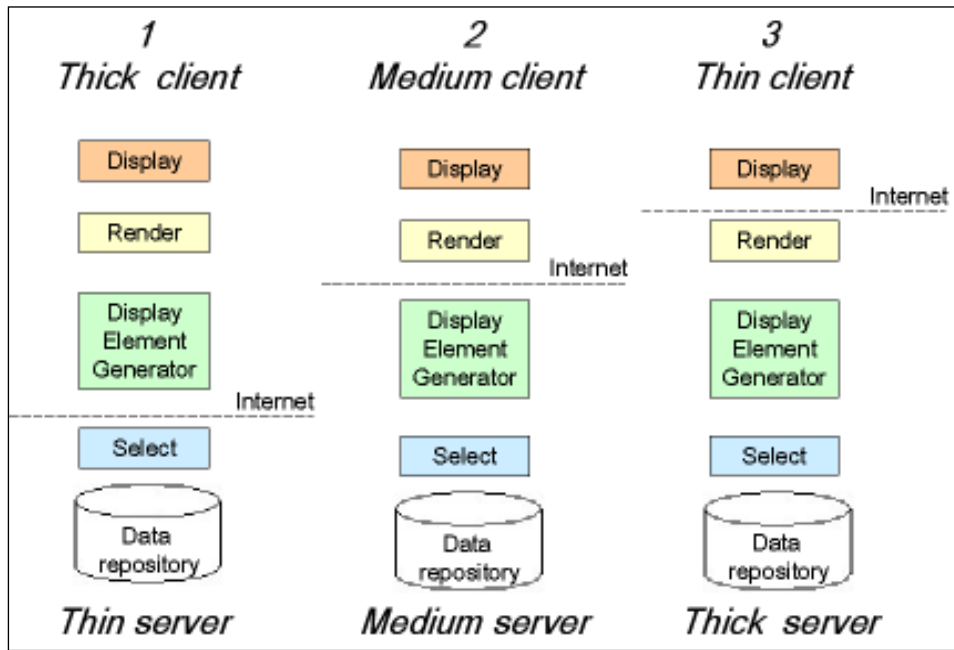


Figure 7. Grid rendering (and associated services) client-server schemas

	Thick client	Medium client	Thin client
Details	Thick clients communicate on the feature level (e.g. WFS) with the server.	Medium clients download display elements which can be easily processed by the graphics hardware	Only transfer rendered multi-layer images from the server to the client.
Advantage	The client is free to realise any interaction and analytic functions The server working load is minimal	This solution has the advantage that it avoids end-users having to install dedicated software before visualisation.	No special hardware is required on client-side. The client side processing load is minimal.
Disadvantage	High computing memory and bandwidth requirements for processing 3D data which requires specific hardware on client side (3D graphics card mandatory).	Transmission load depends on the size of the 3D model and on required quality. The visual quality depends on the client's computing and rendering capabilities.	Interaction on the client side is reduced (e.g. perspective view based on the camera position and viewing direction as specified in the request). Processing load is essentially on the server side.
Technical solutions	Applications such as GIS with 3D extensions or virtual globes such as NASA World Wind or Google Earth. This involves making datasets available in a data repository in pre-processed 3D graphics formats such as VRML (Virtual Reality Modelling Language), X3D (the ISO standard XML-based file format, the successor to VRML), KML (for Virtual Globes) or GeoTIFF Or specific software developed on one or a combination of Java 3D, NASA World Wind library, PHP/VRML, DirectX/ActiveX ... Note that this last solution allows for the inclusion of a pre-processing stage to transform the source data format such as NetCDF to the appropriate graphic format which avoids the requirement for an additional format for data providers.	KML can be used by virtual globes such as NASA World Wind or Google Earth. 3D browser plug-ins are available for X3D. W3DS is a web view service OGC standard under discussion (OGC 09-166r2). Graphical representations are transmitted in VRML, X3D or COLLADA. In connection with 3D navigators in development (Xnavigator, Firefox, IE, GE etc.), this service could be a future alternative.	The WVS is a server side approach for visualisation, analysis, navigation and information retrieval proposed by the OGC. ((OGC 09-166r2) still in discussion). It will be designed to overcome the restricted visualisation and interaction capabilities of prior services (Web Terrain Service/Web Perspective View Service). It provides additional geometrical and thematic data, such as depth information and object identity information, which are encoded in the retrieved multi-layer images that can be easily retrieved and visualised (allows measurements functionalities and enhanced navigation support). The WVS can be used by resource-limited clients and under low bandwidth conditions. In connection with a web browser with 3D graphics, this solution would have the advantage to minimize the effort of end-users to view data set. However 3D browsers are still in development.

Table 10. Product visualisation and associated services technical solutions

Proposed solutions

The objective of the Geo-Seas e-infrastructure is to provide easier access to harmonised data and metadata held by the data centres without creating a significant additional work load either for the end-user or the data centre.

The delivery of data in a graphical format is not appropriate for many applications which require DTMs. Data centres cannot be expected to deliver data in an appropriate graphical format in addition to the agreed data transport format. As a result the solution which is developed for the viewing services must be based on specific viewer technology which is able to process the common data transport formats currently available from the data centres. This requirement essentially constrains the available options, leaving a thick client with software which can handle the data transport formats and provide the required 3D viewing services as the primary mechanism.

Several different solutions can be considered including: Java 3D, NASA World Wind etc. For example the ncBrowse tool is a suitable software tool based on an open-source Java application that provides flexible, interactive graphical displays including 3D visualisation of data and attributes from a wide range of NetCDF data file conventions. Linked with an Opendap library, ncBrowse would allow the user to directly access remote NetCDF files using this application. This solution would require an Opendap server to be made available at each data centre but it would also offer a common approach for accessing data for both 2D and 3D viewing services. However, the ncBrowse documentation, and in particular that for the 3D functionalities, is limited (<http://www.epic.noaa.gov/java/ncBrowse/help.htm>). Other solutions might be considered such as the NetCDF 3D viewer or the use of the NASA World Wind library but these options will need to be evaluated.

5.2. Downloading services

Data format

Among the relevant harmonisation requirements identified by the WP 10 survey, a significant requirement was the need to access grids in common international, INSPIRE compliant formats. OGC standards will be adopted by Geo-Seas for quick viewing and downloading services. In particular, the description of the data will use the INSPIRE metadata implementing rules, while the downloading of the data should be based on the INSPIRE download service implementing rules.

In spite of the limited number of user survey responses related to standards (25), it is notable that the IHO and NetCDF standard formats were the download formats most frequently cited.

The new IHO S-100 hydrographic geospatial standard for marine data and information published in 2010 (http://www.iho-ohi.net/iho_pubs/standard/S-100) is based on the ISO 19100 series of geographic standards. S-100 itself does not mandate a particular encoding format. This means that the developers of product specifications can decide on the suitable encoding standard for their applications. However there is some guidance provided regarding the use of neutral encoding with S-100 in order to promote compatible data exchange. "For imagery and gridded data (regular or not), a neutral encoding would consist of the use of an XML encoding to describe metadata aspects and an appropriate value element encoding mechanism citing the example of the Hierarchical Data Format (HDF version 5) which is object oriented and suitable for all types of coverage data which is the basis of NetCDF". However it will be some time before the tools currently used to ingest and deliver appropriate data formats includes the application of the S-100 standard.

NetCDF is an open-source format that can be used for a variety of scientific data and metadata from a wide range of disciplines. Developed and maintained by UCAR's Unidata program since 1989 (<http://www.unidata.ucar.edu/software/netcdf/>), NetCDF enables quick software development in widely used programming languages. NetCDF is extensible and earlier versions are compatible with current versions. The format is platform independent with issues such as 'endianness' being addressed in the software libraries making this format well adapted for large binary files. NetCDF can handle various data structures. This format is also well designed for use with multidimensional arrays (in particular time and depth) and allows efficient sub-setting. Among one of the significant strengths of NetCDF is that it is "self-describing". This means that each individual data file contains a header which describes the layout of the rest of file, in particular the data arrays, as well as the associated file metadata in the form of name/value attributes. This facilitates the use of conventions such as the Climate and Forecast (CF <http://cf-pcmdi.llnl.gov/documents/cf-conventions/1.0/cf-conventions.htm>) which provide a definitive and sharable description of what each variable represents and the spatial and temporal properties of the data. This enables users of data from different sources to decide which values are comparable, and facilitates the building of applications with powerful extraction, re-gridding, and display capabilities. NetCDF has become a 'de facto' standard in the atmosphere, climate and ocean modelling communities and its use for observational data types is growing. OGC and UCAR believe that establishing CT-NetCDF as a standard for binary encoding will make it possible to incorporate standard delivery of data in binary form via OGC protocols, e.g. WCS, WFS and SOS (Draft CF-NetCDF specification 0.1.0, 2009, OGC 09-122). CF-NetCDF is the format of the General Bathymetric Chart of the Oceans GEBCO-08 grid (http://www.gebco.net/data_and_products/gridded_bathymetry_data/).

It is recognised that a range of other "common" data formats for grids, are in regular use in different end-user communities, however these are not being considered as options for the activities of WP11 because the data transport formats are restricted to those agreed by the Geo-Seas partners [13] [14].

For the purpose of optimising interoperability, Geo-Seas specifies the use of NetCDF (Network Common Data Form) together with the Climate and Forecast (CF) convention (<http://cf-pcmdi.llnl.gov>) as the data transport format for swath and gridded data. The WP4 approach is to assign the Standard Name of the CF convention. The option to converge with the new S-100 standard is being considered however this is not currently a formal requirement.

Downloading mechanisms

An INSPIRE compliant downloading tool has been implemented for use with datasets held by the data centres which are part of the in the Geo-Seas e-infrastructure [15]. However this service will also need additional functionality for use with DTMs to allow for 'on-the-fly' sub-setting of the large gridded data files.

A combination of NetCDF with an Opendap server could enable this sub-setting functionality.

DTM product metadata specifications

The Geo-Seas Common Data Index (CDI) is designed for observational data and is not appropriate for products such as composite DTMs. Interoperability requires referencing and describing the target DTM in a product catalogue compliant with ISO19115 standard metadata and consistent with CDI metadata. This requires:

- Assigning a unique identifier for the product at each data centre in accordance with the Geo-Seas indexing mechanism.
- A common CRS and datum
- Having a time validity indicator (shelf life) of the product (e.g. the date of the oldest data and the date of the most recent data used to create the product which is important for high energy areas with mobile sediments). However these dates are similar to the start and end date of the CDI.
- A description of the spatial representation: dimensions of the grid (X,Y,Z,T etc.), associated resolution and cell geometry, the maximum scale of use
- A robust documentation of the source data used in terms of:
 - any corrections applied
 - parameters and the lineage of the processing steps used for DTM generation
 - a quality indicator for the composite product
 - the suitability and limitations that may apply to the product (e.g. if depths are shoal biased).

Some of the metadata are important in order to ascertain whether the product may be used for a specific application or may be integrated into another product. It is recognised that some of these metadata may be considered too specific at the discovery level of a catalogue of products e.g. parameters of user specific projection of a CRS or cell geometry. In order to deal with this, a solution based on the extension model (tools/formats O&M and SensorML) currently under development by WP 4 to deal with metadata requirements for seismic data could be adopted.

Other options under consideration are :

- To describe the demonstrator DTM in an ISO19115 compliant product catalogue such as that used in the MyOcean project, (<http://operation.myocean.eu/web/24-catalogue.php>) and its CAMIOON interface.
- To use free text fields to capture these specific metadata requirements, ie. in the abstract of the data identification information, as an interim measure until the full O&M extension model (under development in WP4), is implemented.

6. Prototype development

Chapters 3 and 4 proposed a concept for a possible specification for the creation of DTMs from a distributed database of bathymetric data. To test and validate some of these concepts, it will be necessary to develop a series of prototypes. It is expected that prototype testing will mainly provide practical feedback and evidence on the following issues, both at the conceptual level, and from an operational perspective:

- Gridding and data provision (data provider side)
- Efficacy of the scales selected, both grid products and visualisation
- Feasibility of filling the source metadata corresponding to the datasets (data provider side)
- Filtering the available data based on the source metadata provided (product elaboration side)
- Gridding and fusion of multiple datasets from multiple data providers, along with relevant grid statistics and grid cell metadata (product development)
- Product metadata elaboration and creation (product development)
- Dissemination and enhanced visualisation

In its current form the Geo-Seas infrastructure is still under development and lacks the capacity, at least in the immediate future, to provide suitable data. There is therefore a requirement to source suitable data elsewhere. The following potential areas have been identified based on the experience of partners (Table 11).

Location	Countries	Pros	Cons	Data providers
West Med	France and Spain	Data existing for Spain and France	Availability of Spanish data questionable	IEO Sp
Bay of Biscay	France and Spain	Good quality data exists	AZTI not a partner	AZTI, IFREMER, SHOM
Channel	France and UK	Good quality data exists	Uncertainty regarding data availability	UK HO, MCA, SHOM, IFREMER
Gulf of Lyons	France	Good data available	Limited size of grid Controlled areas, restricted areas	IFREMER, SHOM
N Sea	France Belgium	Good quality exists	Uncertainty regarding availability e.g. Dunkirk harbour (could be solved by cooperation with SHOM/IFREMER)	IFREMER, SHOM, Dunkirk harbour, and MUMM
Western Approaches/ Goban Spur	France Ireland and UK	Good data exists e.g. to support UNCLOS		GSI, IFREMER, SHOM and NOCS

Table 11. Proposed test areas for the development of prototype DTMs.

A number of additional tools will be required to enable data providers, users and the emerging system to manage key processes:

For gridding: Generic Mapping Tools (GMT) offer a number of benefits:

- free and easily scriptable
- provides a flavour of NetCDF
- provides a blockmean function

It has also been recognised that the tools which have been developed by IFREMER as part of the EMODNET project may provide a suitable solution, and discussions are on-going regarding the viability of this option.

For product metadata development it is proposed to use a MyOcean-like catalogue for data product and its CAMIOON interface to index and describe the target DTM.

For visualisation the current preferred option (at the time of writing) is to utilise NASA World Wind libraries coupled with other dedicated software, however this practical solution remains under discussion.

Appendix A

References

- [1] Digital elevation model technologies and applications: The DEM Users Manual, Maune D. F. et al., 2nd Edition, 2007
- [2] IHO standards for hydrography surveys”, Special Publication S-44, 5th Edition, February 2008
- [3] Resolutions of the International Hydrographic Organisation, Publication M-3 (2nd edition), November 2010.
- [4] Guide on the development of a vertical reference surface for Hydrography, FIG Publication N°37, September 2006
- [5] Harmonization of Vertical References within the Baltic Sea, FIG Working week 2008, Stockholm, June 2008
- [6] A unified global height reference system as a basis for IGOOS, J. Ihde, L. Sanchez, 2005, Journal of Geodynamics, V. 40, issues 4-5, pp400-413
- [7] IHO Transfer standard for digital Hydrographic data, Special Publication S-57, Edition 3.1, November 2000.
- [8] IHO S-100: The New Hydrographic Geospatial Standard for Marine Data and Information, Ward, R.W., Alexander, L., Greenslade, B., Pharaoh, A. presented at the Canadian Hydrographic Conference, May 2008
- [9] 1st Workshop on European Reference Grids, Short Proceedings, Ispra, October, 2004
- [10] D2.8.I.1 Inspire Specification on Coordinate Reference Systems – Guideline, Thematic Working Group Coordinate Reference Systems and Geographical Grid Systems, 2010
- [11] User guide to the GEBCO one minute grid – centenary edition. BODC, 2003
- [12] The Global Bathymetry and Elevation Data at 30 Arc seconds resolution: SRMT30_Plus. Becker, J. J., Sandwell, D. T., Smith, W. H. F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S-H., Ladner, R., Marks, K., Nelson, S., Pharaoh, A., Trimmer, R., Von Rosenberg, J., Wallace, G., Weatherall, P. Marine Geodesy, 32 : 355-371, 2009.
- [13] Geo-Seas geophysical CDI metadata XML standard, and possible CDI modification list and international conversion protocols, Geo-Seas deliverable D4.2
- [14] Geophysical data transport format, Geo-Seas deliverable D4.5
- [15] Adapted and tested Geo-Seas components for entry, production and exchange, Geo-Seas deliverable D8.1

Appendix C Terminology

Term	Definition
BPI	Benthic Position Index
CRS	Coordinate Reference System
CDI	Common Data Index
DTM	Digital Terrain Model
EEZ	Exclusive Economic Zone
ENC	Electronic Nautical Chart
GIS	Geographic Information System
IHO	International Hydrographic Office
LAEA	Lambert Azimuthal Equal Area
LAT	Lowest Astronomical Tide
MBES	Multi-Beam Echo Sounder
MLLW	Mean Lower Low Water
MSL	Mean Sea Level
OGC	Open Geospatial Consortium
SDN	SeaDataNet
SID	Source Identifier Number
TIN	Triangular Irregular Network
UELN	United European Levelling Network
UNCLOS	United Nations Convention on the Law of the Sea
UTC	Universal Time Convention
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service