INTERNATIONAL BATHYMETRIC CHART OF THE NORTH ATLANTIC

Information for Data Providers: Metadata and Data Format

Version 1.0, 2006-03-31

Benjamin Hell

benjamin.hell@geo.su.se

Department of Geology and Geochemistry, Stockholm University, 10691 Stockholm, Sweden

Data analysis and processing for the International Bathymetric Chart of the North Atlantic (IBCNA) bathymetry compilation project should ideally to some extent be based on metadata accompanying the raw data (single beam and multibeam soundings). This document describes a "wish list" of metadata information which will be useful during data analysis, and which data providers are urged to provide together with the data as far as possible. This list is a best case scenario for very well documented data, and we expect that in reality most of the time not all information will be available. The metadata information provided with the data has not necessarily to be organized as it is in this document, though this of course would substantially facilitate the data integration process.

The metadata format used for the IBCNA project is compliant to the FGCD Content Standard for Digital Geospatial Metadata (version 2.0), FGDC-STD-001-1998 [1]. FGDC-STD-001-1998 standard compliant metadata can be converted in order to meet the ISO 19115:2003 standard [2]. Metadata will be stored on a per-survey basis; i.e. there will be generally one metadata record for every survey. The only metadata stored on a per-sounding basis will be the date and time when the sounding measurement was carried out.

All raw data for the bathymetry compilation will be converted to single sounding points (single beam data) and integrated in an Oracle™ 10g database with the Oracle Spatial extension. Data should preferably be provided as simple ASCII text files. Software available at the Department of Geology and Geochemistry, Stockholm University for reading raw data includes Geographic Information Systems (Intergraph GeoMedia, ESRI ArcGIS), bathymetric data processing software (CARIS SIPS and HIPS), mapping software (Generic Mapping Tools) and CAD software (Bentley Systems MicroStation). If data cannot be provided in a format readable by this software or human-readable ASCII text, detailed information about the data format is necessary. This information should be as detailed as possible, preferably from a programmer's perspective, given the case that it is necessary to program an appropriate data integration tool. Section 6 of the metadata is an appropriate place for this information.

Important metadata information

Five sections of the FGDC standard are especially important for data providers:

- Identification Information (section 1)
- Data Quality Information (section 2)
- Spatial Data Organization Information (section 3)
- Spatial Reference Information (section 4)
- Distribution Information (section 6)

Apart from that, some additional metadata not covered by FGDC standard would be appreciated. The following description is based on the FGDC standard.

Section 1: Identification Information

Basic information about the dataset and its identity.

Originator The name of an organization or individual that developed the data set.

Publication Date The date when the data set is published or otherwise made available for release.

Title The name by which the data set is known.

Supplemental Information Other descriptive information about the data set identity.

Access Constraints Restrictions and legal prerequisites for accessing the data set. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the data set.

Use Constraints Restrictions and legal prerequisites for using the data set after access is granted. These include any use constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on using the data set.

Point of Contact Contact information for an individual or organization that is knowledgeable about the data set.

Native Data Set Environment A description of the data set in the producer's processing environment, including items such as the name of the software (including version), the computer operating system, file name (including host-, path-, and filenames), and the data set size. Also includes information such as ship name, cruise and leg identification as well as project and survey identification.

Section 2: Data Quality Information

A general assessment of the quality of the data set.

Attribute Accuracy Report An explanation of the accuracy of the identification of the entities and assignments of values in the data set and a description of the tests used. This metadata record holds information about the sounding quality (depth or travel time) and includes (but is not limited to): Bathymetry measurement instrumentation and bathymetry measurement accuracy.

Logical Consistency Report An explanation of the fidelity of relationships in the data set and tests used.

Completeness Report Information about omissions, selection criteria, generalization, definitions used and other rules used to derive the data set.

Horizontal Positional Accuracy Report An estimate of accuracy of the horizontal positions of the soundings. Including (but not limited to): Navigation instrumentation, attitude determination instrumentation, horizontal navigation accuracy, attitude measurement accuracy and completeness of navigation data.

Vertical Positional Accuracy Report An estimate of accuracy of the vertical positions in the data set. Includes (but not limited to): Navigation instrumentation, vertical navigation accuracy, settlement and squat.

Process Description An explanation of processing steps and related parameters or tolerances. Especially important is any information about the travel time to depth conversion: Sound velocity, CTDs, Matthew's zones, Carter's tables etc.

Section 3: Spatial Data Organization Information

The mechanism used to represent spatial information in the data set.

Direct Spatial Reference Method The system of objects used to represent space in the data set: point or raster, with necessary parameters such as resolution, row count, etc.

Section 4: Spatial Reference Information

The description of the reference frame for, and the means to encode, coordinates in the data set.

Horizontal Coordinate System Definition Geographic coordinates or projection, projection parameters if applicable.

Geodetic Datum Information about the geodetic datum used and ellipsoid parameters if necessary.

Bathymetric Datum Information about the bathymetric datum used.

Section 6: Distribution Information

Information about the data format.

Format Name The name of the data transfer format.

Format Information Content Detailed description of the data transfer format, unless the data is directly readable by the software listed above.

Technical Prerequisites Description of any technical capabilities that the consumer must have to use the data set in the form(s) provided by the distributor.

Further metadata apart from FGDC standard

Especially for single beam data, timestamp information of the single soundings would be very useful for data fidelity checks.

References

[1] Federal Geographic Data Committee (1998): Content Standard for Digital Geospatial Metadata (version 2.0), FGDC-STD-001-1998. Online at http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/base-metadata/v2_0698.pdf

[2] ISO Technical Committee ISO/TC 211 (2003): International Standard ISO/FDIS 19115:2003(E) Geographical Information – Metadata