B-12 Edition 2.0.3

Guidance on Crowdsourced Bathymetry

IHO

International Hydrographic Organization

Published by the
International Hydrographic Organization
4b quai Antoine 1er
Principauté de Monaco
Tel: (377) 93.10.81.00
Fax: (377) 93.10.81.40
info@iho.int
www.iho.in

© Copyright International Hydrographic Organization 2020

This work is copyright. Apart from any use permitted in accordance with the Berne Convention for the Protection of Literary and Artistic Works (1886), and except in the circumstances described below, no part may be translated, reproduced by any process, adapted, communicated or commercially exploited without prior written permission from the International Hydrographic Organization (IHO) Secretariat. Copyright in some of the material in this publication may be owned by another party, and permission for the translation and/or reproduction of that material must be obtained from the owner.

This document, or partial material from this document, may be translated, reproduced or distributed for general information, on no more than a cost recovery basis. Copies may not be sold or distributed for profit or gain without prior written agreement of the IHO Secretariat and any other copyright holders.

In the event that this document or partial material from this document is reproduced, translated or distributed under the terms described above, the following statements are to be included:

"Material from IHO publication B-12 Edition 1.0.0 is reproduced with the permission of the IHO Secretariat (Permission No ……./…) acting for the International Hydrographic Organization (IHO), which does not accept responsibility for the correctness of the material as reproduced: in case of doubt, the IHO's authentic text shall prevail. The incorporation of material sourced from IHO shall not be construed as constituting an endorsement by IHO of this product."

"This publication is a translation of IHO B-12 Edition 1.0.0 – IHO Guidance on Crowdsourced Bathymetry. The IHO has not checked this translation and therefore takes no responsibility for its accuracy. In case of doubt the source version of IHO publication B-12 Edition 1.0.0 in English should be consulted."

The IHO Logo or other identifiers shall not be used in any derived product without prior written permission from the IHO Secretariat.

Contributors

IHO Statement on Crowdsourced Bathymetry

Crowdsourced bathymetry (CSB) is the collection of depth measurements from vessels, using standard navigation instruments, while engaged in routine maritime operations. The International Hydrographic Organization (IHO) has a long history of encouraging the collection of crowdsourced bathymetry, to help improve mankind's understanding of the shape and depth of the seafloor.

The General Bathymetric Chart of the Ocean (GEBCO) project was initiated in 1903 by Prince Albert I of Monaco to provide the most authoritative, publicly-available bathymetry (depth maps) of the world's oceans. Over the years, the GEBCO project, now jointly overseen by the IHO and the Intergovernmental Oceanographic Commission (IOC) of UNESCO, has produced maps of the ocean floor from depth measurements collected by vessels as they journeyed across the oceans. These "passage soundings" have enabled the creation of progressively more-detailed seafloor maps and digital data grids. More recently, systematic surveys have also been used to improve the maps and grids.

Unfortunately, despite the multitude of data that have been collected since 1903, less than fifteen percent of the world's ocean depths have been measured; the rest of the data used to compile seafloor maps are estimated depths. These estimated depths are largely derived from satellite gravity measurements, which can miss significant features and provide only coarse-resolution depictions of the largest seamounts, ridges and canyons. Progress in mapping coastal waters is only marginally better. IHO publication C-55, Status of Surveying and Charting Worldwide, indicates that about fifty percent of the world's coastal waters shallower than 200 metres remain unsurveyed.

While the hydrographic and scientific community lament this lack of data, the world's interest in seas, oceans and waterways continues to increase. The concept of a blue economy is firmly established, along with an ever-growing public awareness of mankind's dependence upon, and vulnerability to, the sea. Several high-level global initiatives are now in place that seek to address ocean issues, including the United Nation's 2030 Agenda for Sustainable Development Goals, the Paris Agreement under the United Nations Framework Convention on Climate Change, and the Sendai Framework for Disaster Risk Reduction 2015-2030. In this context, the shortfall in bathymetric data is even more significant, as it is now recognised that knowledge of the depth and shape of the seafloor underpins the safe, sustainable, and cost-effective execution of almost every human activity on, or beneath, the sea.

In 2014, the IHO, at its Fifth Extraordinary International Hydrographic Conference (EIHC-5), determined to improve this situation by progressing actions to improve the collection, quality and availability of hydrographic data worldwide. One of these actions, Proposal 4, concerned crowdsourced bathymetry.

The EIHC-5, considering Proposal 4, and the comments made during the Conference, decided, in Decision 8, to task the Inter-Regional Coordination Committee (IRCC) to establish a working group to prepare a new IHO publication on policy for crowdsourced bathymetry.

The International Maritime Organization (IMO) Safety of Life at Sea (SOLAS) carriage requirements oblige all commercial vessels to be equipped with certified echo-sounders and satellite-based navigation systems. As a result, the world's commercial fleet represents a significant, untapped source of potential depth measurements. While CSB data may not meet accuracy requirements for charting areas of critical under-keel clearance, it holds limitless potential for myriad other uses. If vessels collect and donate depth information while on passage, the data can be used to identify uncharted features, to assist in verifying charted information, and to help confirm that existing charts are appropriate for the latest traffic patterns. Crowdsourced bathymetry can also provide vital information to support national and regional development activities, and scientific studies in areas where little or no other data exists.

Recognizing the relevance of bathymetry to international maritime policy and the blue economy and noting that crowdsourced bathymetry may be useful for many potential users of the world's seas, oceans and waterways, the IHO has developed this guidance document to state its policy towards, and provide best practices for collecting, crowdsourced bathymetry. It is hoped that this document will provide volunteer data collectors and interested parties with guidelines for gathering and assessing the quality of CSB data.

This document provides technical guidelines only that in no way supersede national or international laws and regulations.

Secretary-General
20 January 2020

Contents

Contributors 2
IHO Statement on Crowdsourced Bathymetry 3
List of Figures 7
List of Tables 8
Introduction 9
I. Purpose and Scope 9
II. Target Audience 9
III. Document Structure 9
1. Data Contribution 11
1.1 IHO Data Centre for Digital Bathymetry 11
1.2 The Trusted Node Model 11
1.2.1 Transmission Protocol 12
1.2.2 Authentication Method 12
1.2.3 Data and Metadata Formats 13
1.3 Overview of CSB Data Flow 13
1.3.1 Submitting CSB data to the DCDB 13
1.3.2 Accessing CSB data 14
2. Data Collection 16
2.1 Systems and Sensors 16
2.1.1 Echo-sounders 16
2.1.1.1 Single Beam Echo-sounders 16
2.1.1.2 Multibeam Echo-sounders 16
2.1.2 Positioning Systems 17
2.1.3 Motion Sensors 17
2.2 Hardware and Software 17
2.2.1 Data Loggers 17
2.2.2 Understanding NMEA 0183 18
2.2.2.1 GNSS NMEA Sentences 18
2.2.2.2 Echo-sounder NMEA Sentences 18
2.2.2.3 NMEA Data Logging 18
2.2.2.4 Computer Time 19

2.2.3 Onboard Data Storage 19
2.2.4 Data Transfer 19
2.2.5 Continuity of Electrical Power 19
2.3 Vessel and Sensor Measurements 20
2.3.1 Sensor Offsets 20
2.3.2 Variations in Draft 21
3. Data and Metadata 23
3.1 Data vs. Metadata 23
3.2 The Importance of Metadata 23
3.3 Metadata and Data Formats 23
3.3.1 Required Data 24
3.3.2 Optional Metadata 24
3.3.3 Required Metadata from Trusted Nodes 28
4. Uncertainty 30
4.1 Introduction to Uncertainty 30
4.2 Meaning, Sources, and Consequences of Uncertainty 31
4.2.1 The Meaning of Uncertainty 31
4.2.2 Categorisation of Uncertainty 31
4.2.3 Estimation and Expression of Uncertainty 33
4.2.4 Uncertainty for Trusted Nodes and Data Users 34
4.2.4.1 Effects of Sensor Integration on Data Capture 34
4.2.4.2 Modelling Uncertainty 36
4.2.4.3 Consequences of Uncertainty 37
4.3 Uncertainty Guidance for User Groups 39
4.3.1 Data Corrections and Depth Calibration 39
4.3.2 Uncertainty Budget 40
4.3.3 Uncertainty for Trusted Nodes 42
4.3.4 Database Users 43
5. Additional Considerations 45
Annex A – Abbreviations 47
Annex B – Glossary 48
Annex C – GeoJSON Data Contribution Format 50

List of Figures

Figure 1. Data flow from vessels, through Trusted Nodes, to the DCDB. 12
Figure 2. A schematic of a filtered flow of CSB data based on the response provided by a coastal state.... 13
Figure 3. A schematic of the flow of CSB data from the Trusted Node, to the IHO DCDB, to the public.... 14
Figure 4. The IHO CSB Data Viewer, which enables discovery of crowdsourced bathymetry. 15
Figure 5. How to measure offsets between GNSS antenna and echo-sounder transducer. 21
Figure 6. How to measure the depth of the transducer below the waterline. 22
Figure 7. Example of estimating depth with a simple echo-sounder (left), and illustration (right) of the potential for blunders (e.g., the echo-sounder detecting the depth of a school of fish, rather than the seafloor). 30
Figure 8. Effects of accuracy and precision (variance and bias) of measurements on the ability of a system to measure. 32
Figure 9. Example of depth measurements, from the four quadrants of Figure 7. 32
Figure 10. Effects of accurate, but not precise (mostly random) uncertainty on a depth sounding. Here, on average the depth measured (red line) is correct, but point to point it varies from the true depth (yellow line). 33
Figure 11. Examples of the effects of not correcting for vertical offsets. Here, not correcting for the offset of the transducer from the waterline leads to a measurement (red line) that differs significantly from reality (yellow line). This gives a bias (systematic) uncertainty to the measurements. 35
Figure 12. Effects of not correcting for horizontal offsets. Here, not measuring the horizontal offset between the GNSS receiver position and the echo-sounder results in along-track offsets of seafloor features. Red line: measured; yellow line: reality. 36
Figure 13. Examples of shoal-clearance depths for different probabilities of excession, based on the same basic uncertainty estimate of 12.0 ± 0.3m (95% CI). Assuming a 12.0m clearance is only true 50% of the time (left); a 5% probability of being shallower requires the depth to be reduced to 11.74m (middle); a 1:1000 chance of being shallower requires a clearance depth of 11.34m (right). 38
Figure 14. Difference between crowdsourced observations and a reference grid model (data courtesy of SHOM). Errors in the crowdsourced observations are clearly seen in plan view (left), and are reflected in the bimodal distribution of differences (right). The uncertainty associated with the crowdsourced observations might not reflect these differences if the observer's metadata was incomplete. 39
Figure 15. Example of problems that can occur when predicting uncertainty from sparse data, where all objects are not captured in the dataset. From the data (top diagram), geostatistical techniques might predict an uncertainty that the user, without further data or reference, might assume to be the outer limits of the true depth. With objects not captured by the sparse data (bottom diagram), however, there could in reality be discrepancies not captured in the interpolation, and outside of the implied bounds predicted by the interpolation method. 44

List of Tables

Table 1. Required Information 24
Table 2. Optional Metadata 25
Table 3. Trusted Node Metadata 28
Table 4. Sample uncertainty budget for a shallow-water echo-sounder and modern GNSS system. 40

Introduction

I. Purpose and Scope

The purpose of this document is to provide guidance to mariners to help them collect and contribute crowdsourced bathymetric data in a format that is useful to the broadest possible audience. It is hoped that this document will help mariners optimise data collection, and will provide them with information about devices, techniques and formats that are recommended by the International Hydrographic Organization for gathering and contributing CSB data.

This document also provides information about data uncertainty, to help data collectors and data users better understand quality, completeness, and accuracy issues with crowdsourced bathymetry. Additional considerations related to crowdsourced bathymetric data logging and data sharing are also briefly explored.

This document is not intended to provide definitive guidance on how best to use crowdsourced data, as the scope of CSB is far-reaching and has many potential future applications.

II. Target Audience

The IHO seeks to inform and guide collectors of crowdsourced bathymetry data. Organizations (also referred to as 'Trusted Nodes') interested in serving as liaisons between data collectors and the IHO may also find the information helpful. Users of crowdsourced bathymetry data may find this document informative, as well, although they are not the primary audience.

III. Document Structure

This document addresses several topics related to crowdsourced bathymetry. Chapter One, "*Data Contribution*," provides information about how to send crowdsourced bathymetry to the IHO Data Centre for Digital Bathymetry (DCDB), via Trusted Nodes.

Chapter Two, "*Data Collection*," provides a high-level overview of the sensors necessary for collecting crowdsourced bathymetry, as well as best practices and recommendations for collecting CSB. Chapter Three, "*Data and Metadata*," describes the importance of data and metadata, and details the information that is required for submitting CSB to the DCDB, as well as additional information that should be collected whenever possible.

Chapter Four, "*Uncertainty*", delves into data quality issues, and discusses how mariners and end users can better understand the impact of various factors on the reliability of a dataset. Chapter Five,

"*Additional Considerations*", discusses issues that collectors and Trusted Nodes may wish to consider before engaging in CSB activities.

Additional detail and further reading are provided via Annexes and external links. This guidance document is intended to be a living document and will be updated in light of further experience and feedback from data collectors and data users.

1. Data Contribution

This chapter details the process for contributing crowdsourced bathymetry (CSB) data to the IHO Data Centre for Digital Bathymetry (DCDB), via Trusted Nodes, and specifies required data formats. CSB data collectors and Trusted Nodes are strongly encouraged to provide their data to the DCDB to help fill gaps in global bathymetric coverage. These data will in turn be made freely and publicly available through the DCDB CSB web portal.

Hydrographic and bathymetric data collected by the National Antarctic Programmes' ships and others linked with their activity in the Antarctic, should be forwarded by the National Antarctic Programmes, or by other means, to the national hydrographic services using the IHO Collection and Rendering of Hydrographic Data Form available via the IHO website[1]. This process and form, which are safety-of-navigation and charting oriented, do not prevent ships of opportunity from applying guidelines given in B-12, which are multi-purpose oriented as far as bathymetric data are concerned.

1.1 IHO Data Centre for Digital Bathymetry

The DCDB was established by the IHO in 1988 to steward the worldwide collection of open bathymetric data. The Centre archives and shares depth data contributed by mariners and others from across the world. The DCDB is hosted by the US National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information (NCEI) in Boulder, Colorado. All data hosted by the DCDB is accessible online via interactive web map services.

1.2 The Trusted Node Model

The DCDB currently accepts crowdsourced bathymetry (CSB) contributions through a network of Trusted Nodes, which are organizations or individuals that serve as data liaisons between mariners (data collectors) and the DCDB (Figure 1). Trusted Nodes may assist the mariner by supplying data logging equipment, providing technical support to vessels, downloading data from data loggers, and providing the information to the DCDB. The DCDB works with each Trusted Node to standardize metadata and data formats and define data delivery requirements. This model normalises data contributions and minimizes the requirements and effort for mariners.

At present, individual data contributors are encouraged to join an existing Trusted Node. In the future, the DCDB may expand its capability to support individual contributors or other models of contribution.

[1] https://iho.int/en/hpwg

Parties interested in becoming a Trusted Node should contact the DCDB at mb.info@noaa.gov.

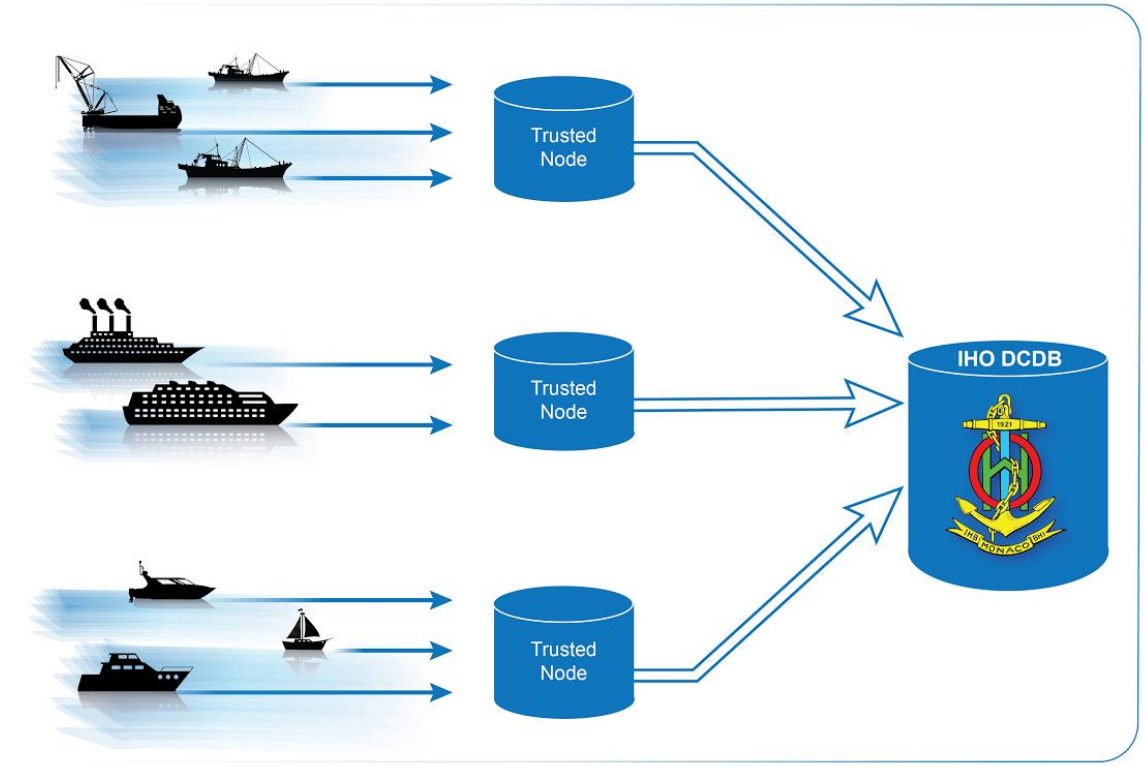

Figure 1. Data flow from vessels, through Trusted Nodes, to the DCDB.

1.2.1 Transmission Protocol

Trusted Nodes have the option of making CSB available for collection by the DCDB using a standard network protocol such as FTP, or via an HTTP post. There are no DCDB requirements for frequency or size of data submissions.

1.2.2 Authentication Method

The DCDB needs to ensure the integrity of incoming data streams, so a unique key is assigned to each Trusted Node to authenticate the provider. The unique key is submitted with the HTTP post and identifies the validity of the data stream in the post. If the unique key is not submitted, or is unknown, the data submission is rejected, and an HTTP 401 error code is returned to the provider. The unique key is only used for the submission process and is not tied to the data files, which can provide a degree of anonymity to the data provider, if desired.

1.2.3 Data and Metadata Formats

All data contributions should conform to the data format and metadata standards described in the Data and Metadata chapter of this document, unless separately and specifically agreed otherwise by the Director, IHO DCDB.

1.3 Overview of CSB Data Flow

CSB data, identified as belonging to the high seas (according to UNCLOS as "the area"), will be ingested into the DCDB database without restrictions on its further reuse. Figure 2 is designed to show future filter processes, including a geographical location checking process and any subsequent masking actions, which will be applied to any incoming CSB data, collected within waters of national jurisdiction, before ingestion into the DCDB database. The flow diagram is of generic nature. The filtered flow of CSB data will be based on the information received by the IHO Secretariat from individual coastal states on request. Further details of which coastal states support CSB activities within their waters of national jurisdiction, along with any caveats they have articulated, will available from the IHO website. CSB data, collected within waters of national jurisdiction of coastal states who did not notify their respective CSB support towards the IHO Secretariat, will not be ingested. This data will be stored and only made available at such time as authorization is received by the IHO from the respective coastal state.

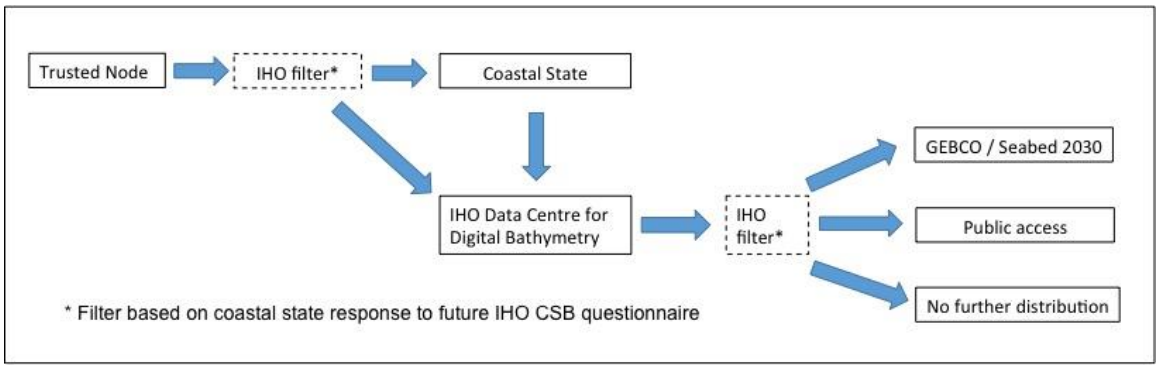

Figure 2. A schematic of a filtered flow of CSB data based on the response provided by a coastal state to a future IHO CSB questionnaire.

1.3.1 Submitting CSB data to the DCDB

CSB data submitted to the IHO DCDB are automatically verified upon receipt. The verification confirms that the data are from a trusted source, and that the submission contains valid file types. The files are then logged in a tracking system at the DCDB. Ingest scripts convert data to GeoJSON if necessary, store the files for user access and archiving, and populate a metadata catalogue. An Extract, Transform, and Load (ETL) process then creates file geometries and populates a spatial database with the geometries and

a subset of the metadata. Figure 3 illustrates the flow of CSB data from the mariner, to the Trusted Node, the IHO DCDB, and finally, to the public.

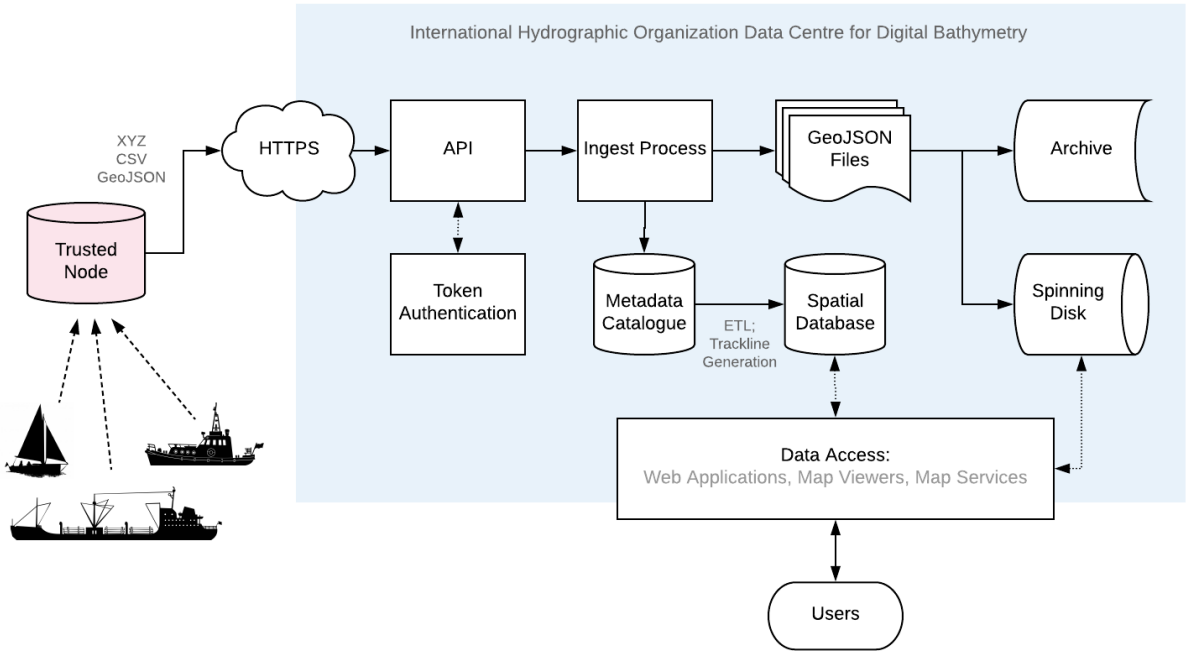

Figure 3. A schematic of the flow of CSB data from the Trusted Node, to the IHO DCDB, and to the public.

1.3.2 Accessing CSB data

The spatial database feeds a map viewer at https://maps.ngdc.noaa.gov/viewers/csb/index.html that enables data discovery (Figure 4). The map viewer is an online tool where users can search for, identify, and obtain CSB data. To help users search for specific data that they are looking for, the map viewer contains filters that correspond to a specified time range or vessel (unless the vessel chooses to remain anonymous). Users can also identify data files geographically, using the Identify tool, which allows users to click on a single point, draw a rectangle or polygon, or input geographic bounds.

Once a selection has been made, a pop-up window shows the corresponding files. Clicking on a file name yields additional information about the file. By selecting "Extract Data," a data request is made, and the user is taken to the Data Access page, where they can edit or finalize their order. The application then sends this data request, along with the requestor's email, to the data delivery system, which verifies that the request is well-formed and then queues the work in the processing system. When data retrieval and preparation are complete, the user is notified via email, and is provided with a URL where they can retrieve the data package.

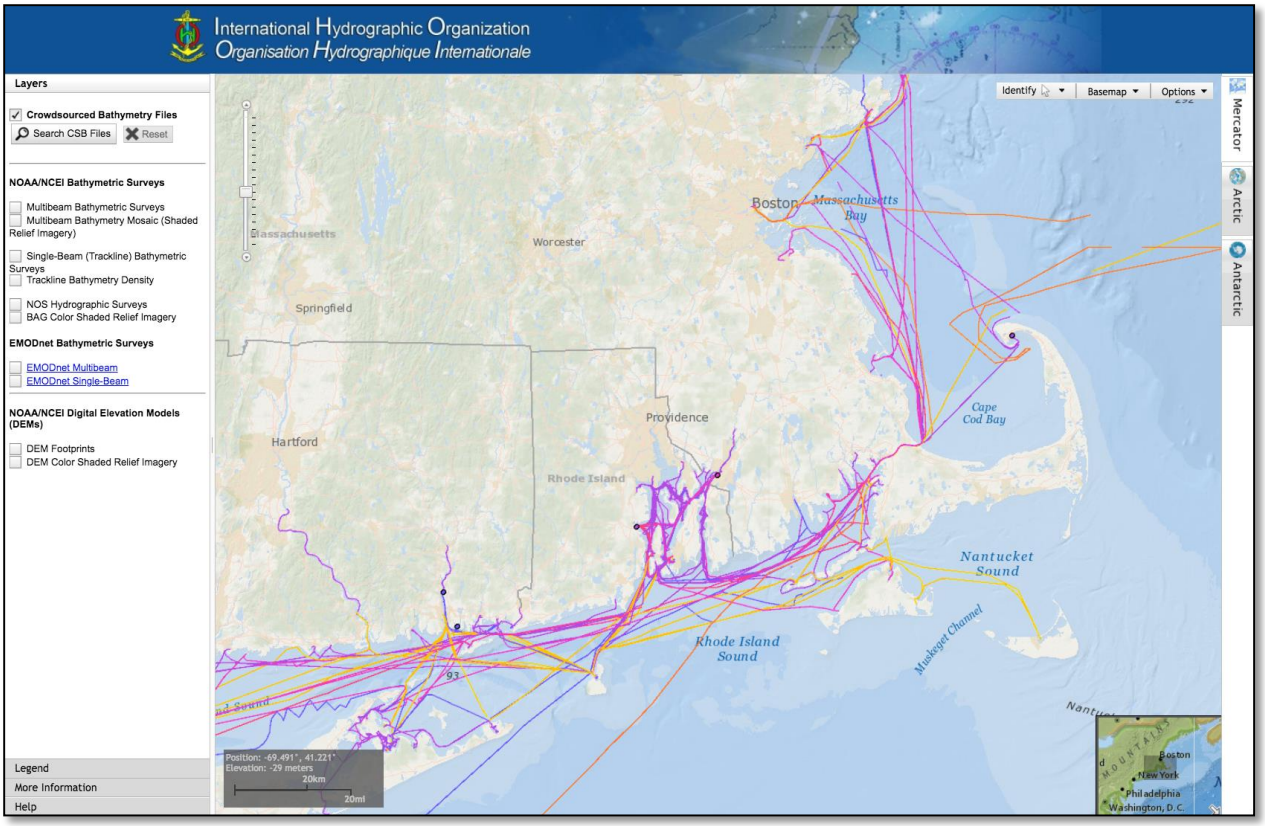

Figure 4. The IHO CSB Data Viewer, which enables discovery of, and access to, crowdsourced bathymetry.

2. Data Collection

2.1 Systems and Sensors

Many vessels already possess the minimum equipment needed to collect CSB, and only need to install a data logger, or enable logging software, to begin collecting CSB. The following sections provide basic information about sensors, as well as best practices and recommendations for collecting CSB. For more in-depth information about systems and sensors, please refer to the IHO publication C-13, Manual on Hydrography (Chapters 2 and 3).

2.1.1 Echo-sounders

Echo-sounders, or depth sounders, determine the water depth by transmitting sound pulses from a transducer, and recording the time it takes for the sensor to receive the return echo from the seafloor. Transducers are usually mounted on the hull of a vessel, but can be mounted on other platforms, as well. There are two main types of echo-sounders: single beam and multibeam. Either of these echo-sounders can be used to collect crowdsourced bathymetry, however the guidance developed in this document will focus on single beam CSB, as the Trusted Node/DCDB model is currently equipped to receive and process those data. Vessels using multibeam for safety of navigation purposes and wishing to contribute their single profiles should submit their data to the appropriate coastal state national Hydrographic Office under cover of a Hydrographic Note.

2.1.1.1 Single Beam Echo-sounders

Single beam echo-sounders collect a single depth measurement from a relatively narrow beam of sound focussed on the seafloor directly under the transducer. Many vessels are equipped with single beam echo-sounders, as they provide sufficient under-keel clearance information for safe navigation. The Trusted Node model is currently designed for donating single beam echo-sounder data to the DCDB.

2.1.1.2 Multibeam Echo-sounders

Multibeam echo-sounders collect depth measurements by forming many receive beams in a wide arc below (or in the case of forward-looking navigation sonar, in front of) the vessel. Multibeam echo-sounders provide a much more detailed representation of the seafloor than single beam depth sounders, and thus can provide additional information about hazards or objects on the seafloor. Multibeam echo-sounders are often found on research vessels, as well as some commercial vessels, expedition cruise ships, and recreational vessels. Vessels equipped with multibeam echo-sounders that wish to contribute data to the DCDB's established multibeam pipeline should contact the DCDB directly at mb.info@noaa.gov.

2.1.2 Positioning Systems

Positioning systems help mariners determine their location on the Earth's surface and provide vital information for CSB. Without accurate location information, CSB has little value. Most vessels carry a Global Navigation Satellite System (GNSS), such as GPS, GLONASS or any other constellation, which obtain position fixes automatically. GNSS positions are typically provided once per second and are accompanied by a time and date stamp. CSB data collection systems should provide a position and timestamp with every depth reading. This allows data users to accurately position depth measurements and apply corrections to the data if needed. The GNSS can also output information about the quality of the signal and interruptions in service, and these data should also be logged, if possible.

2.1.3 Motion Sensors

Some vessels may be equipped with motion sensors. Motion sensors measure the movement of a vessel caused by the waves and swell. Motion sensor data is not required data, and it is acknowledged that most vessels will not be equipped with this technology. For single beam echo-sounders, however, motion sensors capture vertical movement, and are used to correct depth measurements for a vessel's heave. For multibeam echo-sounders, motion sensors measure a vessel's movement in three dimensions, so that corrections can be applied to the data to account for the heave, pitch, and roll of the vessel. Vessels that are equipped with a motion sensor should include motion sensor data at the time of data collection in the dataset they send to their Trusted Node, as it can greatly improve the quality of the final dataset.

2.2 Hardware and Software

In addition to depth, positioning, and motion sensors, there are several hardware and software variables that mariners should consider, when collecting CSB data.

2.2.1 Data Loggers

Crowdsourced bathymetry data loggers are electronic devices or software that connect to a vessel's echo-sounder and positioning system and record the sensor outputs. They write to files in a format designated by the designer of the data logger or software, such as NMEA 0183. The recorded data is then relayed to a Trusted Node, who prepares the data for contribution to the DCDB. Software-based data loggers may be available in an ECDIS or electronic chart plotter that already incorporates input from the echo-sounder and the GNSS. Vessels that do not possess a suitable navigation system, or data logging software, will need to install a standalone logger. Current hardware-based data loggers typically require the installation of a simple, small electronic component that connects to the echo-sounder and GNSS and records their output. Trusted Nodes can provide mariners with data loggers, as well as installation guidance and assistance.

2.2.2 Understanding NMEA 0183

It is helpful for mariners to understand the raw data that is being output by their sensors. Many marine sensors, such as GNSS positioning systems or echo-sounder transducers, transmit data in accordance with standards developed by the National Marine Electronics Association (NMEA). The NMEA 0183 standard data format, or 'sentences,' are both human- and machine-readable, and provide information about the sensor measurement and status. All NMEA sentences begin with a $, and each field is comma-delimited. There are many different types of NMEA sentences. The following sections describe a few that may be useful for CSB data collection.

2.2.2.1 GNSS NMEA Sentences

RMC, GGA, and GLL sentences provide output from the GNSS sensor. Each sentence type provides slightly different information. A 'GLL' NMEA 0183 sentence provides position and time information, and may look like this: **$GPGLL,0424.9921,N,11359.7734,E,012636.21,A,D,*5E**. In this sentence, the 'GLL' designator is followed by the latitude and longitude (with hemispheres), and the time (but not date), in UTC hhmmss.ss format.

A GNSS 'GGA' sentence provides time, position, and fix information, and may look like this: **$GPGGA,071953.00,0424.9862,N,11359.7661,E,1,9,1.8,21,M,,M,,*68**. In this example, the 'GGA' designator is followed by the time (in UTC), the latitude, longitude, and information about the accuracy of the GNSS position fix.

The 'RMC' sentence output from a GNSS contains the recommended minimum navigation information, and provides position, velocity, track made good, date, time, and magnetic variation information. It may look like this: **$GPRMC,102318.23,A,4537.0226,N,03243.0262,E,015.3,186.3,211217,007.2,W*6A**. In this sentence, the 'RMC' designator is followed by the time, in UTC (hhmmss.ss), the latitude and longitude, vessel speed (in knots), track made good (in degrees true), date (ddmmyy), and magnetic variation (degrees and E/W).

2.2.2.2 Echo-sounder NMEA Sentences

NMEA 'DBT' (depth below transducer) sentences for echo-sounders provide depth measurements in several units, and may look like this: **$SDDBT,0006.0,f,0001.828,M,0001.0,F*3A.** The depth, in feet, metres, and fathoms are visible in each of the comma-delimited fields, separated by their unit of measure. For this depth measurement, vertical correctors, such as the draft of the vessel and water level, must be added to the DBT depth to derive the full depth of the water column.

2.2.2.3 NMEA Data Logging

Stripping data from an NMEA sentence is not recommended. Saving the data in its original format will help validate sensor readings and troubleshoot potential anomalies in the data. While the IHO DCDB only

accepts GeoJSON, CSV, or XYZT (longitude, latitude, depth, time) data in one standard format, logging the full NMEA string and submitting it to the Trusted Node is highly recommended. Many data loggers provided by Trusted Nodes already preserve the entire NMEA sentence.

2.2.2.4 Computer Time

A computer's internal clock typically drifts several seconds per week. To maintain the most correct timestamp possible, which will help to maintain the best position information for the depth data, the data logger, or logging software, should instead be set to incorporate the time provided by the GNSS GGA, GLL, or RMC sentence. If it is necessary to rely on the computer clock for the date, document this, if possible, and investigate how well the computer's internal clock will keep accurate time after a long period without system power.

2.2.3 Onboard Data Storage

Vessel owners and operators should ensure that they have adequate onboard data storage capabilities to log depth and positioning data until they can transfer the data to a Trusted Node. Conducting one or two days of trial data logging may help the mariner identify the average file sizes logged by their unique systems and derive an estimate of data storage requirements for longer voyages. If a vessel is installing a hardware-based data logger, the mariner should consult with the Trusted Node to determine the logger's data storage limits. If additional storage is needed, the mariner should ask the Trusted Node if it is possible to transfer data from the logger to ancillary storage (such as an external hard drive) while underway.

2.2.4 Data Transfer

After the CSB data is logged, it should be transmitted to a Trusted Node. Logging and transmitting processes should be as simple and automated as possible to encourage continued contribution of data. Each Trusted Node or data aggregator will provide mariners with the appropriate procedure for CSB data delivery. Sending and receiving data at sea is challenging, and communication systems and bandwidth may be limited or expensive. Because of this, it is important to note that CSB data are not normally time-sensitive; the most important factor is ensuring that the data are shared. Some mariners may wish to leverage communications systems to transfer data while still underway; however, the method of data transmission could also be as simple as mailing a USB storage device to the Trusted Node. Mariners are encouraged to work with their Trusted Node or data logger supplier to identify the preferred method for data transfer.

2.2.5 Continuity of Electrical Power

Continuous power aboard vessels is never a guarantee. Some vessels invest in, or are required to carry, an Uninterruptable Power Supply (UPS) to provide power to navigation equipment in the event of a loss of vessel power. However, not all vessels have a UPS, and even with a UPS, there are times when the

transition from shore power to a generator causes a momentary loss in power. When this happens, data loggers and instruments must reboot and recover. Consider using a data logger that will recover automatically if there is a power interruption, or one that has a back-up battery.

2.3 Vessel and Sensor Measurements

The horizontal and vertical measurements between the GNSS and the echo-sounder, and between the waterline and the transducer, are key components of the quality and accuracy of the data. Some systems are programmed to incorporate these offsets when the sensors are installed. If they do not, mariners should record these measurements as best as possible, and include them in their metadata. The following sections provide information about these measurements, and best practices for collecting and recording them.

2.3.1 Sensor Offsets

Sensor offsets refer to the fore-and-aft and port-and-starboard distances from a vessel's GNSS antenna and the transducer. When measuring offsets, it is important to record the axial directions of positive and negative values, as these conventions can vary. The graphic below (Figure 5) shows an example where measurements are taken from the GNSS antenna to the sonar transducer, with positive values towards the bow and starboard. In some systems, the GNSS antenna offset is already incorporated into the echo-sounder's measurements. If this offset is not automatically integrated, mariners should record their sensor offsets, and relay that information to their Trusted Node. These offset measurements help correct the bathymetric data so that the position indicated by the GNSS is the same as the position of the transducer. This greatly improves the positional accuracy of the depth data.

If the depth information is not corrected with an offset from the GNSS antenna, the depth data may appear to be in a different location than it is. On very large vessels, where the offset between the GNSS antenna and the transducer could be greater, the error could increase.

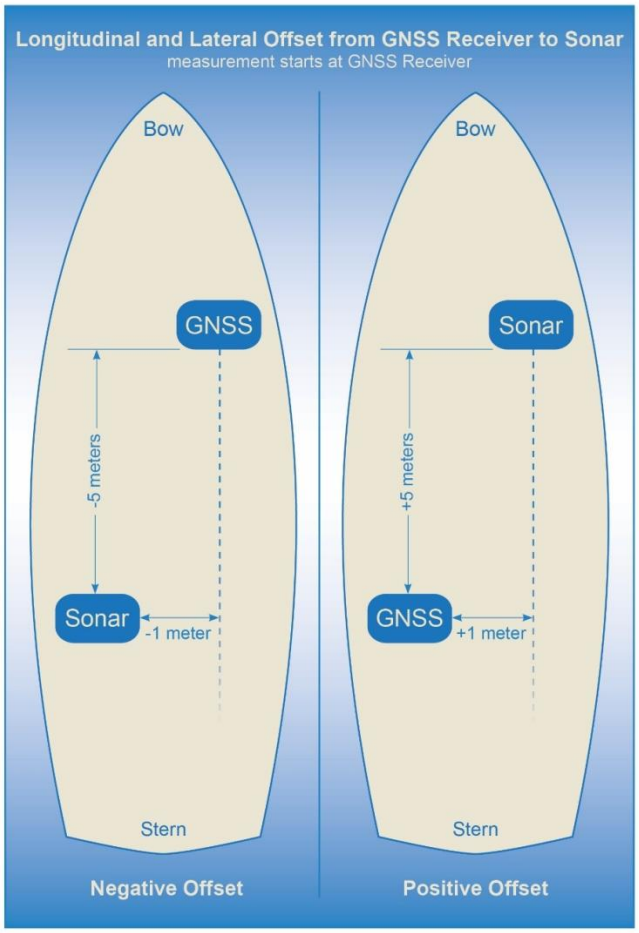

Figure 5. How to measure offsets between GNSS antenna and echo-sounder transducer.

2.3.2 Variations in Draft

If a vessel takes on cargo, fuel, or supplies, the draft of the vessel will vary, which changes the depth of the echo-sounder transducer below the waterline. This change in depth can make the transducer record measurements that are deeper or shallower than reality. As with the sensor offsets, it is important for the mariner to record this information, so that vertical adjustments can be made to the data during post-processing. This can be accomplished by recording the draft of the vessel, together with the time and date, at the beginning and end of a voyage, and providing that information to the Trusted Node (Figure 6).

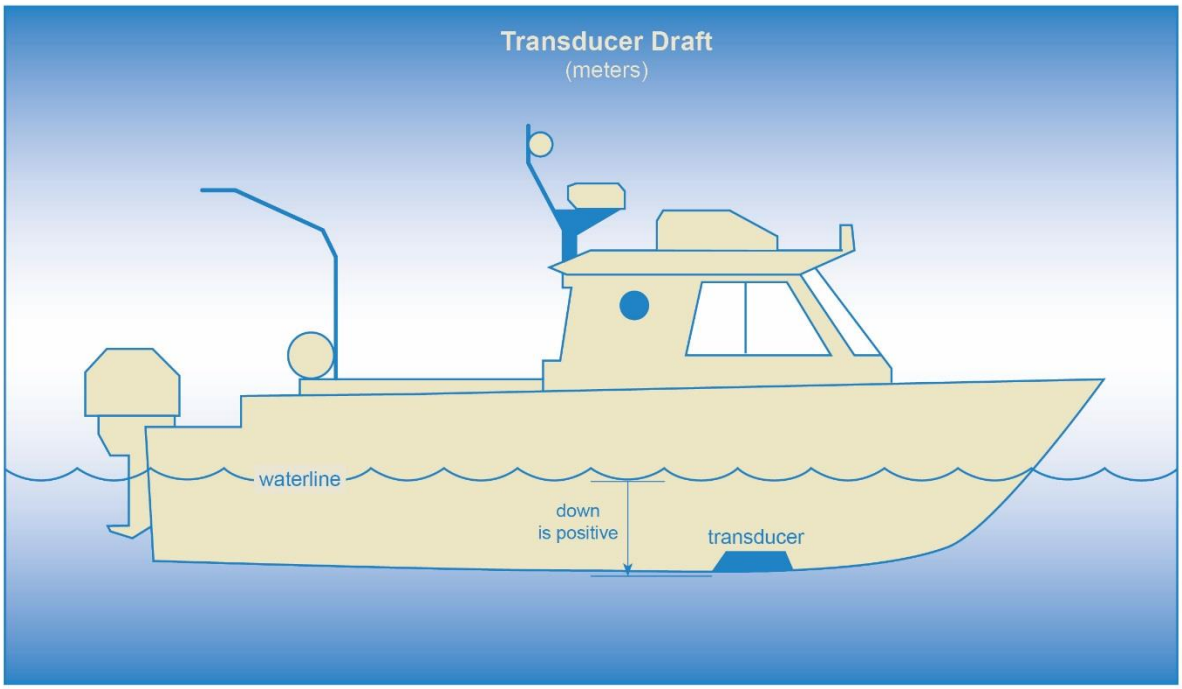

Figure 6. How to measure the depth of the transducer below the waterline.

3. Data and Metadata

3.1 Data vs. Metadata

It is important to understand the difference between data and metadata. Data is the core information, and metadata describes the data. For crowdsourced bathymetry, the data are the depths and geographic positions collected by a vessel, along with the date and time that they were collected. The metadata provides additional, supporting information about the data, such as the make and model of the echo-sounder and GNSS, the vessel's draft, where the sensors were installed on the vessel, and so forth.

3.2 The Importance of Metadata

Metadata provides information to data users that helps them determine the quality of the data, and therefore use the data for more applications than would be possible with depth and position information alone. If the metadata are also consistent, it is easier to incorporate the data into a database, and for users to manipulate the data for their own purposes.

3.2.1 Tidal Information

Crowdsourced bathymetry that is submitted to the IHO DCDB should not have tidal corrections applied. This keeps the data in a standard format. If the data collector provides information about the time and date when a depth measurement was collected, it allows future data users to apply tidal corrections to the data, if they so choose.

3.2.2 Sensor Offsets

Information about a transducer's vertical offset from the waterline, or its horizontal offset from a GNSS receiver, allow users to apply vessel draft and horizontal positioning corrections to the data.

By applying corrections based on information in the metadata, data users can greatly improve the accuracy and value of the bathymetric data for research, industry, or other applications. Refer to section 2.3.1 for more information about sensor offsets.

3.3 Metadata and Data Formats

This section provides guidance to data collectors and Trusted Nodes about the standard metadata that is required for submitting data to the DCDB. In addition, it provides information about additional metadata that would enhance the value of the data for end users. Mariners should collect and forward this information whenever possible.

3.3.1 Required Data

A minimum of information is required, to enable Trusted Nodes to receive and process crowdsourced bathymetry for delivery to the DCDB. Table 1 lists the required information.

Table 1. Required Information

Data Field	Description	Example
Longitude	The vessel's longitudinal geographic position, in WGS84 decimal degrees, to a precision of six decimal places. This can be extracted from the NMEA RMC, GGL or GGA String. Negative values are in the western hemisphere; positive values are in the eastern hemisphere.	-19.005236
Latitude	The vessel's latitudinal geographic position, in WGS84 decimal degrees, to a precision of six decimal places. This can be extracted from the NMEA GGA, GLL or RMC String. Negative values are in the southern hemisphere; positive values are in the northern hemisphere.	40.914812
Depth	The distance from the echo-sounder to the seafloor. Should be collected as a positive value, in metres, with decimetre precision. This can be extracted from the NMEA DBT, DBK or DBS data string. DBT will require additional information about the draft of the vessel. DBS incorporates the draft of the vessel into the overall depth measurement.	7.3
Date & Timestamp	The date and UTC time stamp for the depth measurement. This can be extracted from the NMEA RMC string.	2015-08-06T22:00:00Z

3.3.2 Optional Metadata

Additional information about the vessel, sensors, and sensor installation allows data users to assess the quality of the data, and apply corrections, if necessary. This greatly increases the potential applications

of the data for oceanographic research, scientific and feasibility studies and other uses. Table 2 lists metadata that mariners should provide whenever possible.

Table 2. Optional Metadata

Metadata Field	Description	Example
Vessel Type	The type of vessel collecting the data, such as a cargo ship, fishing vessel, private vessel, research vessel, etc.	Private vessel
Vessel Name	The name of the vessel, in open string format.	White Rose of Drachs
Vessel Length	The length overall (LOA) of the vessel, expressed as a positive value, in metres, to the nearest metre.	65
ID Type	ID numbers used to uniquely identify vessels. Currently, only two types are available: Maritime Mobile Service Identity (MMSI) or International Maritime Organization (IMO) number. The MMSI number is used to uniquely identify a vessel through services such as AIS. The IMO number is linked to a vessel for its lifetime, regardless of change in flag or ownership. Contributors may select only one ID Type.	MMSI
ID Number	The value for the ID Type. MMSI numbers are often nine digits, while IMO numbers are the letters "IMO," followed by a seven-digit number.	369958000
Sensor Type Sounder	This indicates the type of echo-sounder. The only current option is 'Sounder.'' 'Sounders' are simple single beam echo-sounders. In the future, 'Multibeam' may be added as an option. 'Multibeam' refers to vessels equipped with swath sonar	Sounder

	systems.	
Sounder Make	The make of the echo-sounder system. This information may be obtained from a list provided by a Trusted Node.	Sperry Marine (L3 ELAC)
Sounder Model	A free-text value, which provides information about the echo-sounder model. In the future, a list of sounder models may be provided through Trusted Nodes.	ES155100-2
Sounder Frequency	A free-text value, which provides information about the operating frequency of the echo-sounder. In the future, a list of transducer frequencies may be provided through Trusted Nodes.	Dual Freq 200/400 kHz
Sounder Draft	The vertical distance, in metres, from the waterline to the echo-sounder's transducer. The draft should be expressed as a positive value, in metres, with decimetre or better precision if possible. For vessels that operate with a range of drafts, recommended to put in the summer load line.	4.6
Uncertainty of Sounder Draft	The data contributor's estimate of the uncertainty of the echo-sounder's draft measurement, expressed as metres. Vessel draft may be affected by cargo, fuel, or other factors. It is helpful for the data contributor to provide an estimate of how these factors may have affected the transducer's normal depth below waterline, at the time of data collection. Refer to chapter on Uncertainty for more information about how to calculate this value.	1.0
Sounder Draft Applied	Some echo-sounder systems apply vessel	False

	draft in real-time. This field allows the data contributor to state whether draft corrections were applied during data collection ('True') or if they were not ('False').	
Sound Speed Applied	Some systems may have the ability to provide sound speed data and correct the sounding. This field allows the data contributor to state whether sound speed corrections were applied during data collection ('True') or if they were not ('False')	False
Reference point for Depth	The reference point is the location on the vessel to which all echo-sounder depths are referenced. Echo-sounder depths can be referenced to the waterline, the vessel's keel, the echo-sounder transducer, or the GNSS receiver. Information about the reference point helps data users standardise the depth data to a common water level.	Transducer
Sensor Type GNSS	This field defines the sensor type for GNSS receivers. This must always be defined as: "GNSS," and is not a value that data contributors can change.	GNSS
GNSS Make	The make of the vessel's GNSS receiver, which may be selected from a list provided by a Trusted Node.	Litton Marine Systems
GNSS Model	The model of the vessel's GNSS receiver, which may be selected from a list provided by a Trusted Node.	LMX420
Longitudinal Offset from GNSS to Sounder	This is the longitudinal (fore-and-aft) measurement (offset) between the GNSS receiver and the echo-sounder's transducer. This value should be expressed	3.52

	in metres, with centimetre precision. If the GNSS receiver is aft of the sounder, the measurement value is positive. If the GNSS receiver is forward of the sounder, the measurement value is negative.	
Lateral Offset from GNSS to Sounder	This is the lateral (athwartships) measurement from the GNSS receiver to the echo-sounder. This value should be expressed in metres, with centimetre precision. If the GNSS receiver is on the port side of the echo-sounder, the value is positive. If the GNSS is on the starboard side of the echo-sounder, the value is negative.	-0.76
Position Offsets Applied	This field describes whether the final vessel position (longitude and latitude) has been corrected for the lateral and longitudinal offsets between the GNSS receiver and the echo-sounder transducer ("True"), or if they were not ("False").	False
Contributor comments	If the contributor believes there were any problems or events that may have degraded the quality of the position or depth measurements, they can enter that information in this free-text field.	On 3/8/2018, at 20:30 UTC, the echo-sounder lost bottom tracking after the vessel crossed another vessel's wake.

3.3.3 Required Metadata from Trusted Nodes

Trusted Nodes should assign additional metadata to crowdsourced bathymetry before they deliver data to the DCDB. Table 3 lists metadata that Trusted Nodes should provide.

Table 3. Trusted Node Metadata

Provider Contact Point Organization Name	The Trusted Node's name, in free-text format.	Sea-ID

Provider Email	A free-text field for the Trusted Node's email address, so that data users can contact the Trusted Node with questions about the data.	support@sea-id.org/
Unique Vessel ID	Generated by the Trusted Node, this number identifies the Trusted Node and uniquely identifies the contributing vessel. The first five characters identify the Trusted Node, followed by a hyphen (-), and then the vessel's unique identifier. The UUID assigned by the Trusted Node is consistent for each contributing vessel, throughout the life of service of the vessel. However, **if the vessel chooses to remain anonymous to data users, the Trusted Node does not need to publish the vessel name in association with the UUID**.	SEAID-*UUID*
Convention	This field describes the format and version for the data and metadata, such as CSB 2.0, CSV, or XYZT	CSB 2.0
Provider Logger	The software program or hardware logger used to log the data.	Rose Point ECS
Provider Logger Version	The software or hardware logger version.	1.0

4. Uncertainty

4.1 Introduction to Uncertainty

There are many variables that could cause echo-sounder measurements to differ from the true depth of the seafloor. For example, an echo-sounder measures the time it takes for an acoustic pulse to reflect off the seafloor and return to the transducer. That measurement is then converted into a depth measurement, based on an assumption about the speed of sound in water (v) (Figure 7). If the sound speed estimate is incorrect, then the depth (D) will also be incorrect. Similarly, if the sound wave reflects off fish in the water column (Figure 6), or if the echo-sounder captures acoustic noise from other boats in the area, errors (often distinguished as "blunders") could be introduced into the data.

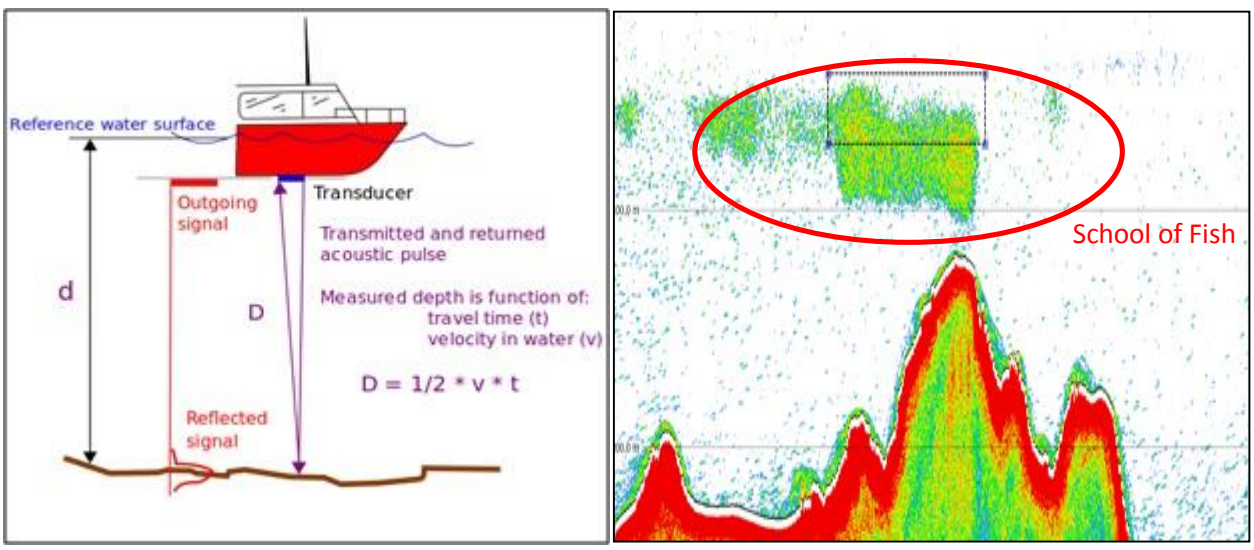

Figure 7. Example of estimating depth with a simple echo-sounder (left), and illustration (right) of the potential for blunders (e.g., the echo-sounder detecting the depth of a school of fish, rather than the seafloor).

These errors, and others, could lead to uncertainty in the accuracy of a depth measurement, which should be considered when the data is processed, stored, and used. This chapter presents features of uncertainty that may be of interest to data collectors, trusted nodes, and end users of the crowdsourced bathymetry database.

The topic of uncertainty can become quite involved. This document provides an overview of the topic, but IHO Special Publication S-44 (Standards for Hydrographic Surveys, 5ed, 2008), IHO Publication C-13 (Manual on Hydrography, 2010), and the ISO Guide to Uncertainty in Measurements (ISO, 1995) contain additional background material, and may be consulted for further details.

4.2 Meaning, Sources, and Consequences of Uncertainty

4.2.1 The Meaning of Uncertainty

In a scientific context, "error" is the difference between the measured and the true value of the thing being measured. Unfortunately, it is usually impossible to directly, physically verify the true value, and therefore the actual error is unknown, and unknowable. Instead, we can estimate the likely amount of error in the measurement, which is called the "uncertainty", and report it with the measurement. A quantified uncertainty is essential in understanding and qualifying a measurement for use. For example, an estimate of the uncertainty of a depth measurement allows data users to determine the data's suitability for a given purpose, and to apply appropriate processing techniques.

4.2.2 Categorisation of Uncertainty

Many different measurements are combined to create a depth estimate. As a result, there are many potential sources for error and therefore uncertainty. It is helpful to categorise the different types of uncertainties that could affect these measurements, and then estimate their individual magnitudes, before combining them into a general estimate of uncertainty. This is typically done in the form of an uncertainty budget (see Section 4.3.2 for a comprehensive example).

For each source of uncertainty, the most common method for categorising the uncertainty type is to estimate the precision (or variance) and accuracy (or bias) of observations. All observations have the potential for both types of uncertainty, although any given observation might be dominated more by one or the other type (which can make estimation simpler). Ideally, estimates of precision and accuracy would be tracked separately for each observation, until all sources of uncertainty are combined.

Figures 8 and 9 show illustrative examples of precision and accuracy. By preference, all depth observations would be both accurate and precise, but random variations in measurements can result in an observation that is accurate, but not precise. Well-calibrated depth estimates often fall into this category (Figures 8-10).

Alternately, depth measurements may be precise, but not accurate, if there is an offset that could be corrected but for some reason is not. For example, if the speed of sound is assumed to be some fixed value, and is not actually measured, depth measurements will be offset from the true depth (i.e., of poor accuracy), even though consecutive measurements appear similar (i.e., of high precision). A correction could be applied to improve the data, but this might not be practical or time-efficient. It might therefore be more pragmatic just to estimate the level of bias and consider it an uncertainty.

Ultimately, it may be difficult to conduct a full analysis of the uncertainty for each observation. So long as what was done is documented, however, any information provided is still valuable.

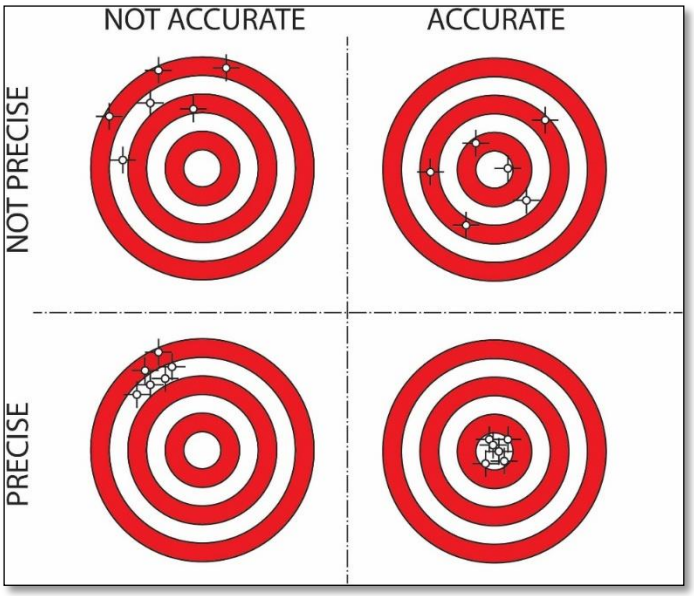

Figure 8. Effects of accuracy and precision (bias and variance) of measurements on the ability of a system to measure.

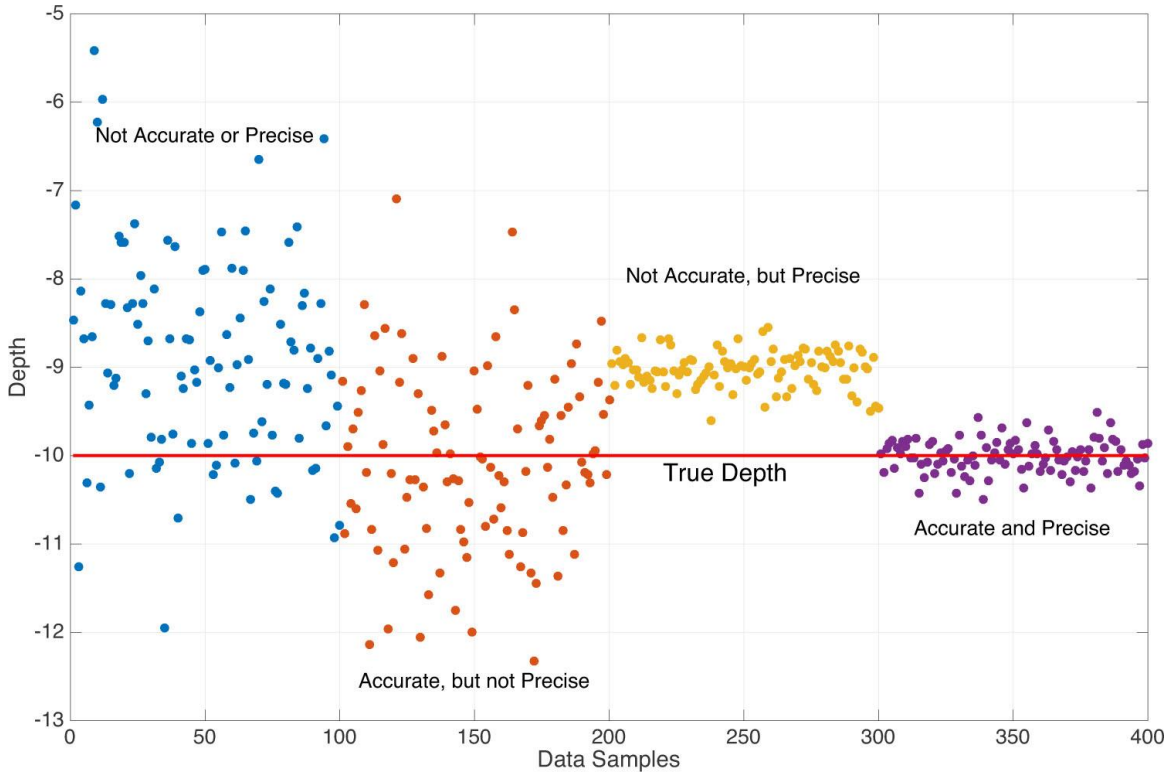

Figure 9. Example of depth measurements, from the four quadrants of Figure 8.

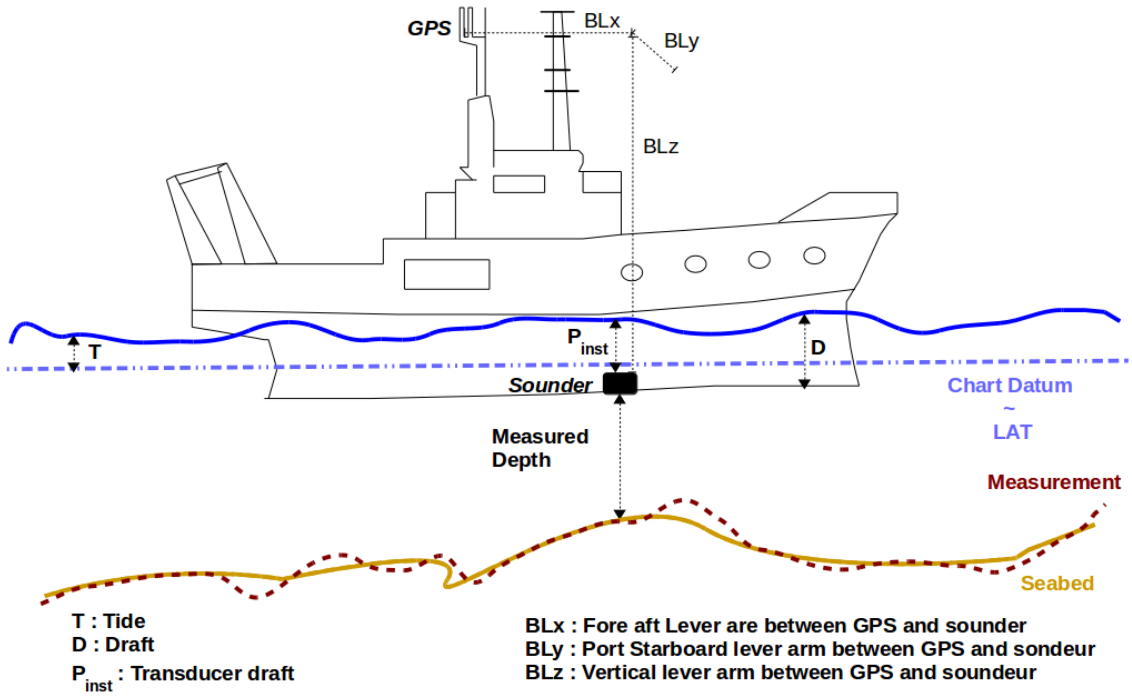

Figure 10. Effects of accurate, but not precise uncertainty on a depth sounding. Here, on average the depth measured (red line) is correct but point to point it varies from the true depth (yellow line).

4.2.3 Estimation and Expression of Uncertainty

Different types of uncertainty can be estimated separately, and then combined into an overall value. This works well when there is sufficient metadata available to help with the calculation, which is why it is so important for crowdsourced bathymetry data collectors to provide as much information as possible about a dataset. Unfortunately, this information is not always available, or provided.

Uncertainty can be expressed as a range of values, within which the true value of the measurement is expected to lie. For example, a depth could be specified as being "between 12.3 and 14.2 metres, 95% of the time." Where the range is known, or assumed to be symmetric, the mean value and spread might be given, so that the depth might be specified as "13.25 ± 0.95 metres, 95% of the time." Whichever method is used, it is important to clearly state what method has been used (i.e., specify explicitly which confidence limits, or other measures, are being used).

Where available, the simplest method for assessing uncertainty is to collect redundant measurements and compute an empirical estimate of the standard deviation (or other range) by considering the

variability in the measurements, which are assumed or known to be of the same thing (e.g., the depth). In practice, this is often done by considering lines of soundings that cross each other, or visit the same area, although this only measures the precision of the observations, and not necessarily the accuracy.

Although statistical descriptions of uncertainty are preferred, there might not always be sufficient information to provide a complete description of uncertainty. Under these circumstances, data might be described as “Poor”, “Medium,” or “Good” quality, or rated as an index (e.g., in the range [1,5] with a one end of the scale defined as “best”) based on a subjective assessment of how the data was collected, or by comparing the data with other datasets. The Category Zone of Confidence (CATZOC) characteristic of the S-57 Electronic Navigational Chart (ENC) specification is an example of this type of subjective assessment.

4.2.4 Uncertainty for Trusted Nodes and Data Users

There are additional uncertainty components that Trusted Nodes and data users should understand when dealing with crowdsourced bathymetry. These uncertainty concerns are mostly about validation and processing of the data, and therefore can only truly be assessed at the Trusted Node, or from the DCDB as a whole. It is expected that data collectors are unlikely to be able to assess these types of uncertainty, although they may find the information of interest.

4.2.4.1 Effects of Sensor Integration on Data Capture

Integration uncertainty becomes an issue when an instrument is installed incorrectly, or when the installation is poorly documented. For example, if the offset between the echo-sounder transducer and the waterline (Figure 11), or between the GNSS receiver and the transducer (Figure 12), are not measured, or are measured incorrectly, an additional uncertainty will affect the depth estimates. Where possible, therefore, Trusted Nodes should attempt to assess this uncertainty, and provide this estimate with the data to DCDB as metadata. In many cases, the limit of this assessment might be a binary flag to indicate whether such measurements exist or not, and if they have been applied already. Even this level of information in invaluable for end users to determine fitness for purpose.

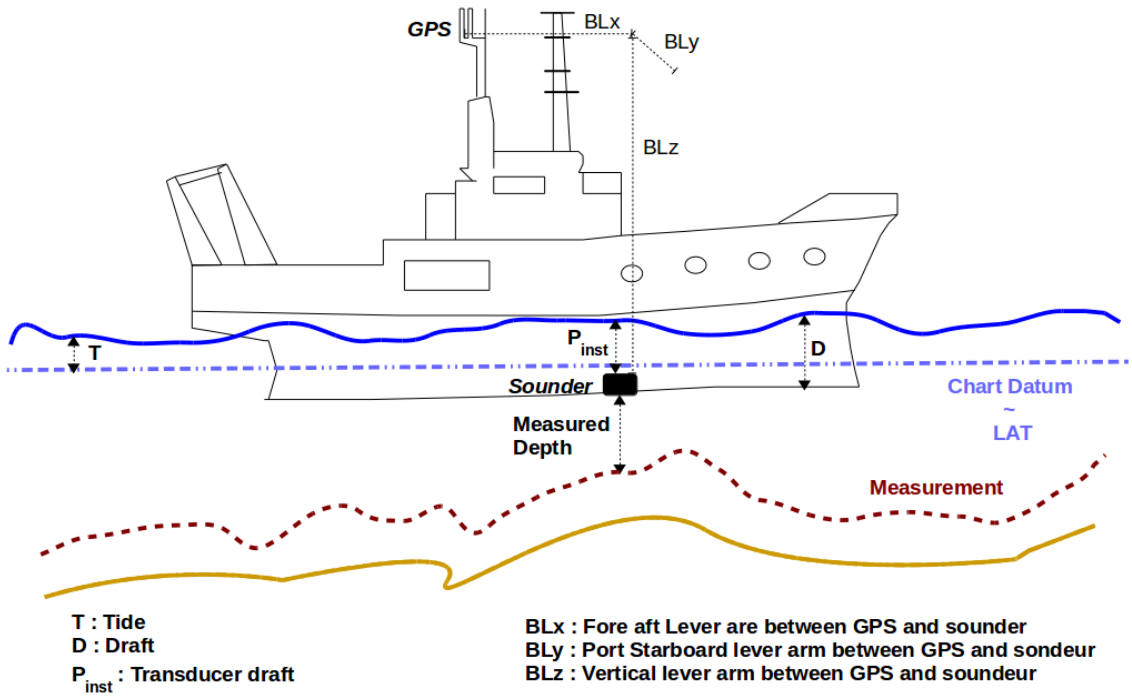

Figure 11. Examples of the effects of not correcting for vertical offsets. Here, not correcting for the offset of the transducer from the waterline leads to a measurement (red line) that differs significantly from reality (yellow line). This gives a bias (systematic) uncertainty to the measurements.

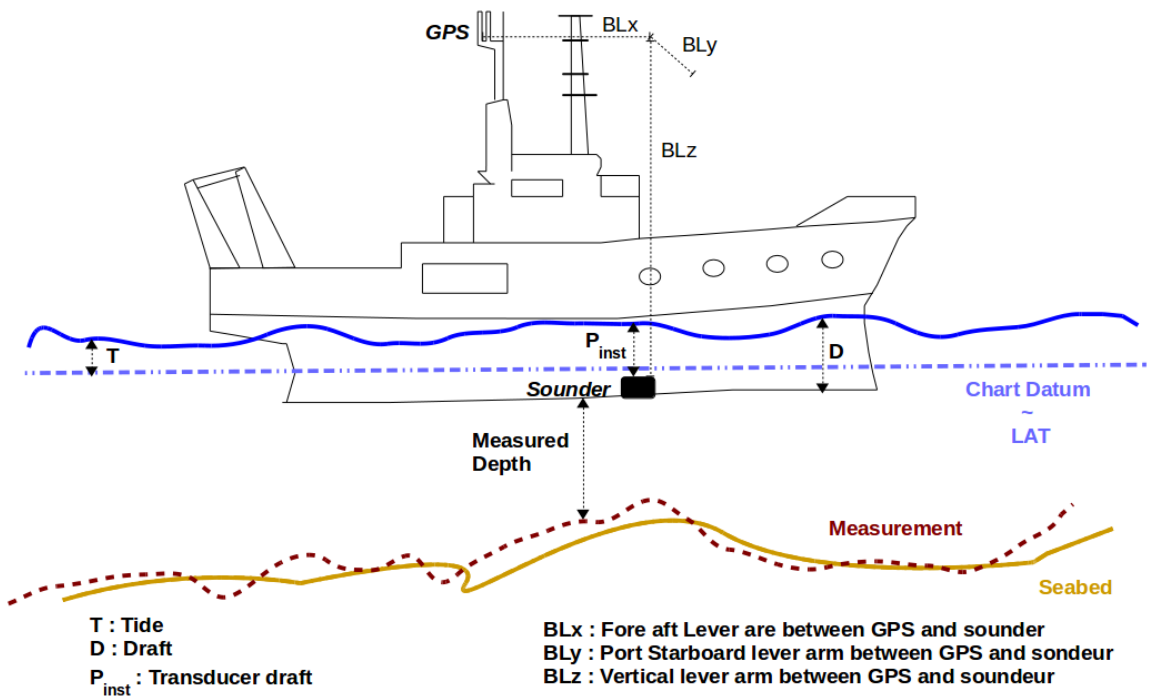

Figure 12. Effects of not correcting for horizontal offsets. Here, not measuring the horizontal offset between the GNSS receiver position and the echo-sounder results in along-track offsets of seafloor features. Red line: measured; yellow line: reality.

4.2.4.2 Modelling Uncertainty

The seafloor is complex, and most of the seafloor is unsurveyed, but it is often modelled as a continuous mathematical surface with interpolated depths where no observations exist. The assumptions involved in constructing such a model will obviously affect the uncertainty of the resulting depths. This is the most difficult of the uncertainties to estimate, and is often ignored.

Many datasets do not contain sufficient data to allow a model to be built that completely describes the seafloor being reported, or for users to determine the resulting quality. For example, if a model was constructed from depth measurements that are more than 50m apart, it is impossible to assess the shape, location, or presence of objects smaller than 100m. It is possible (although not recommended) to interpolate any data, no matter how sparse, to an arbitrary resolution, such as a 1m grid. However, most

of the information in this grid would be an artefact of the interpolation scheme, and would not reliably represent the real world.

If data users do not understand these issues, models may appear to be accurate when they are actually heavily, or even mostly, interpolated. Gridded data can be very visually persuasive, which can result in the erroneous belief that the data are better than they are. Those who construct models like this should carefully document the procedures used in order to inform the potential end user. The metadata is an appropriate venue for this.

4.2.4.3 Consequences of Uncertainty

Although the use of uncertainty models and budgets have been a part of modern hydrographic practice since the late 1990s, uncertainties are often computed as part of data processing, but then either forgotten or dropped when the data are presented or interpreted. This is a mistake.

For example, if a depth is reported as 12.0 ± 0.3m (at a 95% confidence interval), it would be unwise to assume that a vessel has at least 12m clearance in this depth area; with the usual probabilistic assumptions of the distribution of the uncertainty this is true only half of the time (Figure 13(a)), which is surely lower odds than any prudent mariner would allow for a navigation decision. A value of 11.74m would be a better choice (Figure 13(b)), but if a mariner wanted a less than 1:1000 chance of the depth being shallower than the declared value, they should use a depth of 11.34m (Figure 13(c)). Clearly, the "safe" depth depends on the user's needs, and it would be incorrect, and unwise, to report simply the mean depth.

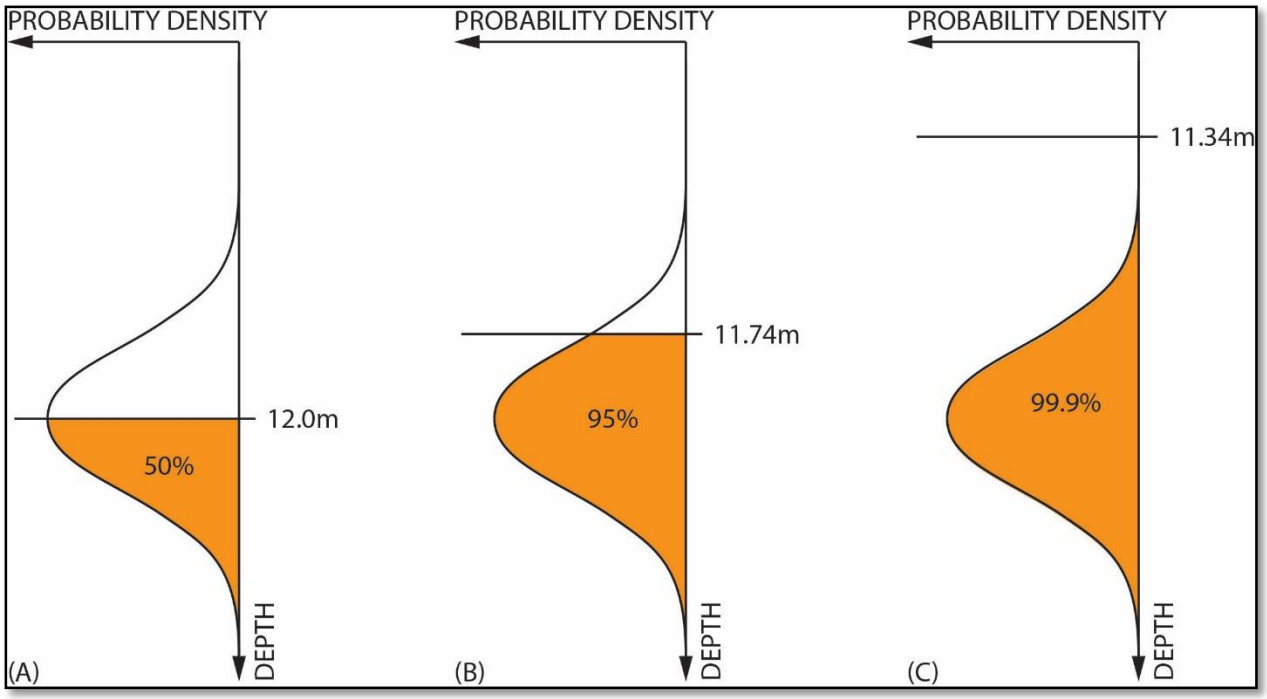

Figure 13. Examples of shoal-clearance depths for different probabilities of excession, based on the same basic uncertainty estimate of 12.0 ± 0.3m (95% CI). Assuming a 12.0m clearance is only true 50% of the time (left); a 5% probability of being shallower requires the depth to be reduced to 11.74m (middle); a 1:1000 chance of being shallower requires a clearance depth of 11.34m (right).

Like depths, uncertainties are only estimates: a best guess, based on what the provider assumes to be the behaviour of the data collection system. Hence, it is possible for an observation to have an uncertainty estimate that does not actually reflect the difference between the measurement and the true depth.

Consider, for example, the data in Figure 14. Here, the data from crowdsourced observations have been compared to high-resolution, authoritative data, which shows significant differences between the two in some areas. The mistake here is that vertical offsets (such as tidal corrections) have not been appropriately applied to the crowdsourced observations. This error would not be apparent to individual data contributors, who do not have access to the comparison data. One of the benefits of donating data to the DCDB through a Trusted Node is that these data aggregators can compare individual datasets to other sources and can identify errors or uncertainty in the data.

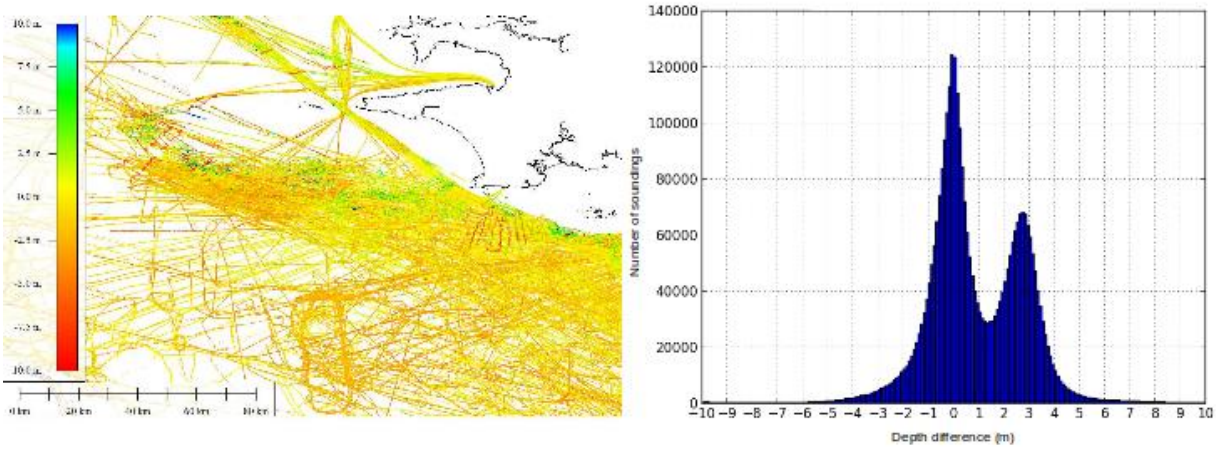

Figure 14. Difference between crowdsourced observations and a reference grid model (data courtesy of SHOM). Errors in the crowdsourced observations are clearly seen in plan view (left) and are reflected in the bimodal distribution of differences (right). The uncertainty associated with the crowdsourced observations might not reflect these differences if the observer's metadata was incomplete.

Note that a 10% uncertainty in depth would be very important to known about, but a 10% uncertainty in the uncertainty (i.e., that it is in the range 9-11%) is probably not as important. Therefore, so long as the uncertainty estimate is plausible, and free from blunders as outlined above, the requirements for estimating the uncertainty are not as stringent. This idea can be used to rationalise the effort required to estimate uncertainties to a reasonable level.

4.3 Uncertainty Guidance for User Groups

4.3.1 Data Corrections and Depth Calibration

Data users need to know if corrections, such as vessel draft or tidal offsets, should be applied to crowdsourced datasets before use. Metadata (Chapter 3) provides the key information that lets data users determine what corrections are needed: the more information that the users have at their disposal, the more corrections can be applied, and the more useful the data then becomes.

Determining which corrections are necessary is only part of the story. Each correction influences the overall uncertainty of the depth measurements, so recording how corrections were determined and applied is also very important. If there was a degree of uncertainty in a correction applied to the data, then that should be indicated in the metadata.

Areas of known depth, also known as calibration surfaces, are sometimes established by hydrographic agencies or harbour authorities on prominent markers such as channel buoys, fuel docks, or well-trafficked areas. Collecting data over these areas makes a dataset significantly more valuable; collecting many observations while stationary (or very slowly drifting) in such an area also allows the measurement

uncertainty of the echosounder to be estimated in some cases. If the data collector also conducts a cross-check, by collecting depths perpendicular to a previous track, that information can be useful for identifying internal dataset inconsistencies.

Environmental changes around a vessel can significantly impact depth measurements and may necessitate more frequent calibrations. In coastal areas where there is significant riverine freshwater discharge, for example, changes in the salinity of the water that affect the speed of sound can cause the echo-sounder to register incorrect depths. Details on how to do a full echo-sounder calibration can be found in IHO publication C-13, *Manual of Hydrography.*

4.3.2 Uncertainty Budget

Data collectors can summarise uncertainties associated with their depth observations in a table known as an uncertainty budget. Some components of the uncertainty vary with the depth being measured, others are fixed. An example of an uncertainty budget from a professional survey is shown in Table 4. Volunteer data measurements will probably not be this precise, or provide all of this metadata, but the more information that is gathered and provided, the more valuable the depth measurements become.

Table 4. Sample uncertainty budget for a shallow-water echo-sounder and modern GNSS system.

Sources of Uncertainty	**Applied (Yes/No)**	**Example of assessed standard uncertainties (95%) values at 50 m**	**Remarks**
Static draft setting		±0.1 m	The static value for draft that was set in the echo-sounder.
Variation of draft		±0.05 m	Change of draft due to variation in loading condition. Average draft to be assessed from full loaded and ballast condition.
Sound speed		±0.2 m/s	Measurement is based on the equipment. It depends on temperature, salinity and depth.

Echo-sounder instrumental uncertainty		±0.1 m	Not to be confused with the resolution of the instrument, this varies with the type of equipment.
Motion sensor Roll/Pitch		±0.05 deg.	This measurement depends on the sensor.
Heading		±0.05 degrees	This measurement depends on the sensor.
Heave		±0.05 m	This measurement depends on the sensor.
Dynamic draft, settlement and squat		±0.1 m	Effects data primarily in shallow water. Settlement depends on speed of vessel and draft.
Tide measurement		±0.06 m	Tide is the variation in sea level and depends on the location at which the tidal measurement is calculated or observed. This location is not always the same place as the depth measurement. Tide measurement not applicable to depths more than 200m.
Sensor offset		±0.01 – 0.1 m	Offset needs to be measured as accurately as possible. Measure of uncertainty depends on how offset was measured.
Position		±2 – 10 m	Measurement depends on the equipment and whether any GNSS sensor offsets have been applied.
Time Synchronization		<1 ms	Measurement depends on the equipment.

Creating a complete uncertainty estimate can be time-consuming, but uncertainty variables can be prioritized, based on the vessel's operating environment. For example, in shallow water, recording draft and water level is particularly important, as variations in these values greatly impact the depth measurement when it is reduced to a charting datum. In deeper water, sound speed information is more important than other factors. In most cases, vessel pitch and roll has a relatively small impact on uncertainty for the data considered here.

4.3.3 Uncertainty for Trusted Nodes

Trusted Nodes are in an ideal position to generate uncertainty estimates for data they transmit to the DCDB. They can cross-check between datasets, remove data biases, calculate the uncertainty associated with data collectors and depth measurements, and potentially correct for them. This can greatly increase the value of crowdsourced bathymetry sent to the DCDB, and is recommended for all Trusted Nodes.

Trusted Nodes can apply corrections to the data that individual observers cannot. They can compare data with authoritative datasets or evaluate data for internal consistency. Trusted Nodes may also choose to collaborate with harbour authorities to establish areas of known depth where individual users can calibrate their echo-sounder measurements. Similarly, it may be difficult for many collectors to establish an uncertainty associated with water level offsets. A Trusted Node, however, might be able to establish, from data taken *en masse*, a plausible buffer to add to the uncertainty budget to represent those corrections.

Analysis of multiple datasets within the same area could also be used to establish baseline uncertainties for data collectors, and to identify data quality issues. Trusted Nodes could then establish a calibration and uncertainty history for each data collector, which could be contributed to the DCDB as part of the metadata supplied with each dataset. A history of user behaviour could also be used to help identify changes in instrumentation.

Trusted Nodes could cross-calibrate data, by using data collected by a vessel with well-established uncertainty and calibration values to determine the installation or measurement uncertainty of other data collectors in the same area. Metadata of this kind can help database users establish confidence in individual data collectors.

Trusted Nodes will have a more direct relationship with data collectors than the DCDB or database users, and as a result they are well-placed to evaluate the metadata and resolve missing, corrupted, or ambiguous information. This can improve the uncertainty associated with each observation, and the end user's confidence in the data.

Trusted Nodes are also in an ideal position to encourage data collectors to improve the metadata that they provide and to attempt data corrections. They might also provide data collectors with feedback on areas for improvement.

4.3.4 Database Users

Database users should interpret the uncertainty information provided with a dataset and generate new uncertainty estimates for their own work. In doing so, they should be aware that the uncertainties provided by data collectors, or assessed by Trusted Nodes, might not be consistent: the uncertainties assessed by data collectors could be subjective and may not have been verified against authoritative sources of depth information. Very low uncertainty estimates should be treated with caution. There is no universally accepted best practice for the statement of uncertainty, although the 95% confidence interval is very common. The type of uncertainty reported should be well-documented, and embedded in the product's metadata.

Users Beware. The DCDB provides no guarantee of the correctness of crowdsourced bathymetry observations. However, some Trusted Nodes might provide stronger guarantees for data that they aggregate. The database user must assume that residual blunders might exist that are difficult to capture in conventional uncertainty statistics.

Database users should avoid over-confidence in uncertainty values when using interpolation methods that estimate their uncertainties from the geostatistics of the observations (e.g., kriging), since the data density may be insufficient for the purpose. In practice, the assumption that all significant variability is captured by the geostatistics may not be valid for the real world. Database users should be aware of this, and should identify how they will compensate for sparse data in the dataset. Figure 15 provides a diagrammatic example of problems that can arise from applying geostatistical interpolation to sparse datasets.

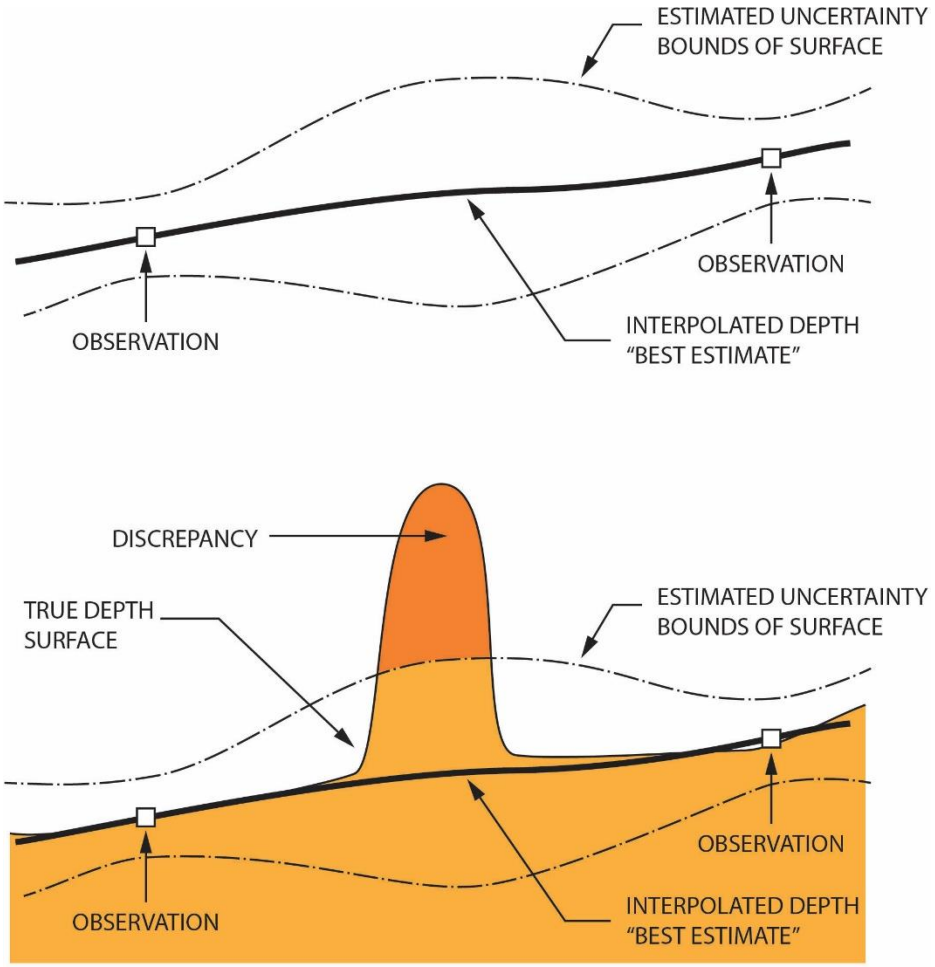

Figure 15. Example of problems that can occur when predicting uncertainty from sparse data, where all objects are not captured in the dataset. From the data (top diagram), geostatistical techniques might predict an uncertainty that the user, without further data or reference, might assume to be the outer limits of the true depth. With objects not captured by the sparse data (bottom diagram), however, there could be discrepancies not captured in the interpolation, outside of the implied bounds predicted by the interpolation method.

Database users are ideally placed to identify problems with individual observers or datasets. Users who identify outliers, or anomalous observers, are encouraged to communicate this information to the DCDB.

5. Additional Considerations

The following notes, which are not exhaustive, are intended for information only.

The principles of the IHO crowdsourced bathymetry (CSB) programme are similar to many other initiatives where environmental data and information are collected on a voluntary basis by users and the public, and provided under an open data licensing infrastructure in the interests of the common good. In particular, the collection and forwarding of bathymetric data by mariners as part of "passage sounding" in support of global initiatives such as the GEBCO project has been taking place for more than a century without issue. Since charts were first produced, mariners have noted and highlighted any inconsistencies with published data, identified during their passage, in the form of a Hydrographic Note. This CSB guidance document is designed to assist the mariner with improving the quality and consistency of any information they may wish to contribute to the public domain.

When considering participation in the IHO CSB programme, mariners should consider the following:

- Mariners proposing to collect bathymetric data as a "*passage sounding"* activity need to be aware of conditions that may be associated with collecting such environmental information within waters of national jurisdiction;
- Those involved in the IHO CSB programme, whether as a data collector, a trusted node or a user, need to be aware of the conditions of the licensing regime under which the bathymetric data will be made available;
- Those using data obtained from the IHO DCDB need to consider the nature and the uncertainty of the data and whether it is fit for the purposes intended.

In order for users to be clear on their rights and obligations while using these data, the IHO CSB Programme has selected a set of licenses from the Creative Commons. The IHO CSB Programme operates under the Creative Commons licensing framework (www.creativecommons.org). Data supplied to the IHO DCDB by vessels, either directly or through Trusted Nodes, is licenced in accordance with the "*Attribution 4.0 International*" license (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0/), and the "*Attribution 3.0 IGO*" license (CC BY 3.0 IGO) (https://creativecommons.org/licenses/by/3.0/igo/). The IHO may, in the future, update its selected licenses as the versions and terms of the Creative Commons licenses change. However, the IHO will maintain at least the rights currently provided by the CC BY 4.0 and the CC BY-IGO 3.0 licenses.

CSB collectors, including Trusted Nodes and the DCDB, are expected to acknowledge that by providing their data for inclusion in the IHO DCDB database, they are doing so in good faith and for the purpose of increasing bathymetric knowledge of the world's seas, oceans and waterways. If the bathymetric data is provided to the IHO by a CSB collector, then the free-use of the data granted by the data collector should

apply. They also acknowledge that the IHO may allow anyone to copy and redistribute the data that they supply to the IHO DCDB in any medium or format and may remix, transform, and build upon the data for any purpose. CSB collectors cannot revoke these freedoms so long as users of their data follow the licensing terms stated above.

The IHO will also make it clear that that data is being made available on a “user-beware” basis; in particular, emphasizing that the user must carefully consider the nature and the uncertainty of the data being used in relation to any use proposed by the user.

In granting its licence to data users, it should be noted that the IHO, as an intergovernmental organization, enjoys certain rights and privileges, which include immunity from the jurisdiction of national courts.

Annex A – Abbreviations

AIS - Automatic Identification System

CI - Confidence interval

CSB - Crowdsourced bathymetry

CSV - Comma separated values

DBT - Depth Below Transducer (NMEA sentence)

DCDB - IHO Data Centre for Digital Bathymetry

ECDIS - Electronic Chart Display and Information System

FTP - File Transfer Protocol

GGA – position fix information (NMEA sentence)

GEBCO - General Bathymetric Chart of the Ocean

GLL - Geographic position, latitude / longitude (NMEA sentence)

GNSS - Global Navigation Satellite System

GPS - Global Positioning System

HTTP(S) - Hypertext Transfer Protocol (Secure)

IHO - International Hydrographic Organization

IOC - Intergovernmental Oceanographic Commission of UNESCO

IMO - International Maritime Organization

MMSI - Maritime Mobile Service Identity

NCEI - National Centers for Environmental Information

NMEA - National Marine Electronics Association

NOAA - National Oceanic and Atmospheric Administration

RMC – Recommended minimum data for GPS (NMEA sentence)

UNESCO - United Nations Education Scientific and Cultural Organization

UPS - Uninterruptable Power Supply

UTC - Coordinated Universal Time

UUID - Unique Uniform Identification

Annex B – Glossary

Automatic Identification System (AIS). A tracking system that broadcasts, via VHF, the position, course and speed of a vessel to other vessels in the vicinity, to reduce the risk of collisions.

Confidence Interval (CI). A range of values so defined that there is a specified probability that the value of a parameter lies within it.

Crowdsourced bathymetry (CSB). The collection of depth measurements from vessels, using standard navigation instruments, while engaged in routine maritime operations.

Electronic Chart Display and Information System (ECDIS). A computer-based navigation system that complies with IMO requirements and can be used for navigation instead of paper navigation charts.

General Bathymetric Chart of the Ocean (GEBCO). Publicly-available bathymetric map, and associated products, of the world's oceans. GEBCO is an IHO and IOC joint Project that relies largely on the voluntary efforts of an international collaborating community of scientists and hydrographers with the support of their parent organizations..

Global Navigation Satellite System (GNSS). A satellite navigation system with global coverage, such as the United States' NAVSTAR Global Positioning System (GPS), the Russian Federation's GLONASS, and the European Union's Galileo.

International Hydrographic Organization (IHO). The IHO is the intergovernmental consultative and technical organization that was established in 1921 to support safety of navigation and the protection of the marine environment. The principal aim of the IHO is to ensure that all the world's seas, oceans and navigable waters are surveyed and charted.

International Maritime Organization (IMO). The IMO is the United Nations specialized agency with responsibility for the safety and security of shipping and the prevention of marine pollution by ships. It is the global standard-setting authority for the safety, security and environmental performance of international shipping. Its main role is to create a regulatory framework for the shipping industry that is fair and effective, universally adopted and universally implemented.

Intergovernmental Oceanographic Commission of UNESCO (IOC-UNESCO). The IOC is the United Nation's competent body for marine science. The IOC's role is to promote international cooperation and to coordinate programmes in research, services and capacity-building, in order to learn more about the nature and resources of the ocean and coastal areas and to apply that knowledge for the improvement of management, sustainable development, the protection of the marine environment, and the decision-making processes of its Member States.

National Marine Electronics Association (NMEA). The US-based marine electronics trade organisation setting standards of communication between marine electronics

Trusted Nodes. Organizations or individuals that serve as data liaisons between mariners (data collectors) and the DCDB. Can provide mariners with data loggers, installation and data download assistance, recommendations on best practices for collecting CSB, and providing the information to the DCDB.

Annex C – GeoJSON Data Contribution Format

Crowdsourced Bathymetry GeoJSON

Format Version 2.0
Last Update: April 4, 2018


```
"type": "FeatureCollection",
"crs": {
        "type": "name",
        "properties": {
                "name": "EPSG:4326"
        }
},
"properties": {
        "convention": "CSB 2.0",
        "platform": {
                "uniqueID": "SEAID-e8c469f8-df38-11e5-b86d-9a79f06e9478",
                "type": "Ship",
                "name": "White Rose of Drachs",
                "length": 65,
                "lengthUnitOfMeasure": "meters",
                "IDType": "IMO",
                "IDNumber": "1008140",
                "sensors": [{
                        "type": "Sounder",
                        "make": "Sperry Marine (L3 ELAC)",
                        "model": "ES155100-02",
                        "transducer": ""
                }, {
                        "type": "GNSS",
                        "make": "Litton Marine Systems",
                        "model": "LMX420"
                }],
                "correctors": {
                        "sounderDraft": 4.6,
                        "sounderDraftUnitOfMeasure": "meters",
                        "sounderDraftApplied": false,

                        "longitudinalOffsetFromGNSStoSounder": 3.52,
                        "longitudinalOffsetUnitOfMeasure": "meters",
                        "lateralOffsetFromGNSSStoSounder": -0.76,
                        "lateralOffsetUnitOfMeasure": "meters",
                        "positionOffsetsApplied": false,

                        "soundSpeed": 1500,
                        "soundSpeedUnitOfMeasure": "m/s"
                }
        },
```

```
                "providerContactPoint": {
                        "orgName": "Sea-ID",
                        "email": "support@sea-id.org",
                        "logger": "MarkIII",
                        "loggerVersion": "1.0"
                },
                "depthUnits": "meters",
                "timeUnits": "ISO 8601"
        },
        "features": [{
                        "type": "Feature",
                        "geometry": {
                                "type": "Point",
                                "coordinates": [41.914832, 18.005296]
                        },
                        "properties": {
                                "depth": 15.8,
                                "time": "2016-03-03T18:41:49Z"
                        }
                },
                {
                        "type": "Feature",
                        "geometry": {
                                "type": "Point",
                                "coordinates": [40.914789, 19.005552]
                        },
                        "properties": {
                                "depth": 15.2,
                                "time": "2016-03-03T18:41:50Z"
                        }
                }
        ]
}
```