General Guidance Notes on the collection and use of Crowdsourced Bathymetry

Foreword

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 governed by the IHO and the Intergovernmental Oceanographic Commission (IOC) of UNESCO, has produced maps of the ocean floor, based on depth measurements made by vessels as they journeyed across the oceans. These "passage soundings" have progressively enabled more detailed maps (and now digital data grids) to be created. More recently, systematic surveys are also used to improve the maps and grids when they become available. The GEBCO maps, and now digital data grids, are available to the public to use as they wish. Most of the ocean areas shown in Google Earth and most of the world's maps of the ocean are based on GEBCO data. For the most part, the underlying GEBCO data is kept in the IHO Data Centre for Digital Bathymetry, from where it can also be downloaded by the public.

Unfortunately, despite all the data that has been provided since 1903, less than 15% of the depth of the world's ocean area has actually been measured - all the rest remains as an estimated depth, mostly using satellite gravity measurements, which can miss significant features and sometimes give misleading results.

Progress in measuring the depth in coastal waters is only marginally better than in the deeper ocean. According the IHO's figures in its publication C-55 – Status of Surveying and Charting Worldwide, about 50% of the world's coastal waters down to a depth of 200m remain unsurveyed.

This means that in the 21st century, we have better resolution maps of the Moon and Mars than we do of most of the Earth's seas, oceans and waterways.

In reviewing the situation in 2014, the IHO, at its 5th Extraordinary International Hydrographic Conference determined to improve the situation by progressing actions to improve the collection, quality and availability of hydrographic (depth data) worldwide. One of these actions concerned crowdsourced bathymetry - the collection of depth measurements from vessels, mostly using their standard navigation instruments, engaged in their normal operations of sailing from one place to another.

As a result, the IHO DCDB is undergoing progressive enhancements to make it easier for crowdsourced bathymetry (both passage soundings and systematic depth measurements conducted for purposes other than nautical chart improvement)) collected by vessels to be uploaded and thereby increase the data available to everyone, in this world bathymetric data base. Enhancements are also being made to the DCDB to enable all the available data for any area in the database to be downloaded by the public to use as they see fit. Wherever possible, the data in the data base will be associated with metadata to allow users to decide whether the data they download is good enough for their intended purpose.

The data in the DCDB is available on an as is, and user beware basis. The data in the DCDB, by its very nature of collection, is of variable quality depending on the type of equipment used and the way that the equipment was set up and used. Nevertheless, against a background where there is little or no other data available, the data in the DCDB will always be useful. Even for nautical chart improvement, crowdsourced bathymetry can be useful for such things as identifying otherwise unknown navigational hazards, for validating the depths shown on existing charts, for identifying changes in depths on published charts, and for confirming that nautical charts are covering the most used routes. For other uses, even depths with a relatively high value of uncertainty may be acceptable - particularly where no other data exists.

This IHO document is intended to provide both prospective data observers and data users with appropriate guidance on the various matters that they should consider when observing and providing crowdsourced bathymetry data to the DCDB and when they extract data from the DCDB to use for their particular purposes.

XXXX
Secretary-General
date

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 The Need for Crowdsourced Bathymetry

Seventy-one percent of the Earth's surface is covered with water, yet only ten percent of the seafloor has been mapped to a resolution of one arc-second or better. Apart from a comparatively small percentage of systematic hydrographic surveys, our current view of the shape of the ocean floor is pieced together from satellite measurements of the sea surface, global gravity models and passage soundings collected by commercial or scientific vessels.

Detailed knowledge of global bathymetry can provide vital insight into the behaviour of global earth systems and their impact on our world. The morphology of ocean basins, undersea ridges and seamounts influence the flow of ocean currents that nourish fisheries, sustain ecosystems, and impact weather and climate. Detailed maps of the seafloor would help scientists and policymakers to better understand the global marine ecosystem, and enable them to respond to pollutants or contaminants that are carried by ocean currents. Seafloor characteristics also influence the transmission of energy from undersea seismic events, and accurate bathymetric models could help to predict and mitigate tsunami or storm surge impact in coastal areas. Timely information about shifting shoals or submerged hazards could also provide national hydrographic offices with critical data for nautical chart updates.

The urgent need for comprehensive bathymetric coverage will not be met by government and hydrographic office efforts alone. If we harness the collective reach of private and commercial vessels and empower seagoers to 'map the gaps' in seafloor data in a coordinated way, we can greatly increase our understanding of the seafloor, and its influence on the world around us. Crowdsourced bathymetry can play a vital role in creating the global seafloor map of the future.

1.2 Purpose and Scope of this Document

The purpose of this document is to provide guidance to seagoers to help them collect and contribute bathymetric data in a format that is useful to the broadest possible audience. It is hoped that this document will help seagoers to optimise data that is already being collected on vessels equipped with common commercial depth sounders and Global Navigation Satellite System (GNSS) receivers, and will provide them with information about devices, techniques and formats that are recommended by the International Hydrographic Organization (IHO) for crowdsourced bathymetry (CSB) data sharing.

This document also provides guidance on position and depth accuracy and data uncertainty, to help both the seagoer collecting the data and potential data users to better understand some of the considerations and limitations of crowdsourced bathymetry, as well as the feasibility of using the data for certain applications. The legal considerations of bathymetric data logging and crowdsourced bathymetry sharing are also briefly explored.

This document is not intended to provide definitive guidance on how best to use crowdsourced data by an end user, although it is acknowledged that the scope of CSB is far-reaching and has many potential future applications.

1.3 Target Audience

This guidance document seeks to inform and guide potential individual data collectors and contributors of crowdsourced bathymetry data. In addition, organizations (also referred to as 'trusted nodes') interested in serving as liaison organizations between the contributing seagoers and the IHO may find the information helpful. Users of crowdsourced bathymetry data may also find this document informative, although they are not the primary audience.

1.4 Document Structure

This document addresses six topics related to crowdsourced bathymetry. The first chapter introduces the concept of crowdsourced bathymetry and discusses its potential benefits. The second chapter, "*Overview of Systems and Sensors*", provides basic information about systems, sensors and concepts that are necessary for collecting bathymetric data.

Chapter three, "*Metadata*", details a standard metadata structure for crowdsourced bathymetry datasets which facilitates the efficient exchange and use of the data. This chapter delineates required and recommended metadata fields, as well as the importance of each metadata field. Chapter four, "*Data Collection*", outlines hardware and software considerations for logging CSB information, and provides recommendations for best practices for shipboard data collection.

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

Chapter six, "*Data Contribution*", focuses on methods for contributing data to the global database of bathymetric data. The Trusted Node model is explained, and includes information regarding the various aspects of the data submissions that are agreed upon by a trusted node and the IHO Data Centre for Digital Bathymetry (DCDB).

Chapter seven, "*Legal Considerations*", discusses several legal considerations related to crowdsourced bathymetry data 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.

2. Overview of Systems and Sensors

2.1 Depth sounders

Depth sounders determine water depth by transmitting sound pulses from a transducer mounted in the hull of a ship or boat and recording the time taken to receive the return echo from the seafloor.

The speed of a sound wave traveling through water is influenced by the temperature, salinity, and depth of the water. For most seagoers, setting a depth sounder to an average sound velocity constant of 1500 m/s (in saltwater) is adequate for general navigation. When a very precise understanding of seafloor depths is desirable (such as for coastal hydrographic surveys) scientists may deploy sensors to directly measure the composition of the water column. These measurements are then used to more accurately calculate the speed of sound in the water and thereby determine a more accurate water depth.

2.2 Single Beam Depth Sounders

Single beam depth sounders collect a single depth measurement from a relatively narrow beam of sound focussed on the seafloor directly under the transducer. Most commercial and recreational vessels are equipped with single beam depth sounders that operate using a standard setting for the speed of sound in water.

2.3 Multibeam Depth Sounders

Multibeam depth sounders collect depth measurements by emitting a large number of focussed beams of sound in an athwartship arc below the hull. Multibeam depth sounders provide a much more detailed representation of the seafloor than single beam depth sounders. Multibeam depth sounders may be found on research vessels, and some specialist commercial and recreational vessels.

2.4 Positioning Systems

Positioning systems allow the seagoer to determine their location on the Earth's surface. Without a precise location, bathymetric data is of limited value for many end users. Most ships are required to carry a GNSS to obtain position fixes automatically. GNSS position fixes are typically provided once per second (one Hertz). GNSS fitted in ships can provide information about the quality of the signal, and interruptions in service.

2.5 Time and Date

Time is very important when collecting depth data. Knowing the time when the depth data was collected enables the data to be adjusted for tides and other corrections. Time is one of the outputs from GNSS, and should be recorded in Universal Time Coordinated, or UTC.

2.6 Motion Sensors

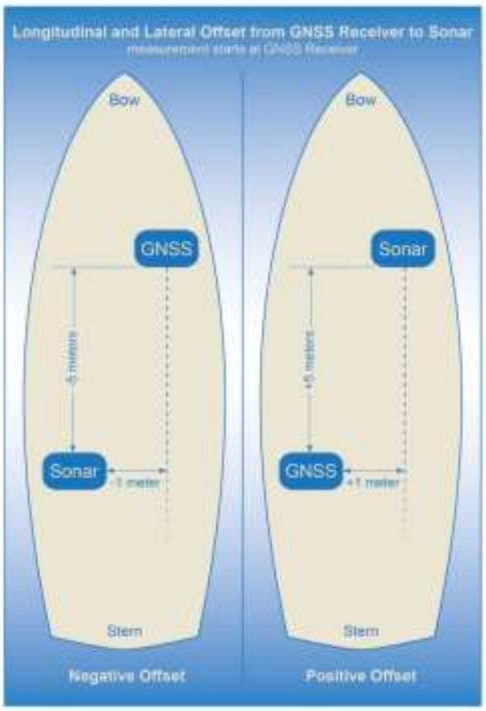
Motion sensors measure the movement of a vessel caused by the waves and swell. For single beam depth sounders, motion sensors identify the amount of vertical movement, and are used to correct depth measurements for a vessel's heave caused by wave action. For multibeam depth sounders, motion sensors measure a vessel's movement in all three-dimensions so that corrections to the multibeam depth measurements can be made to account for the heave, pitch and roll of the vessel.

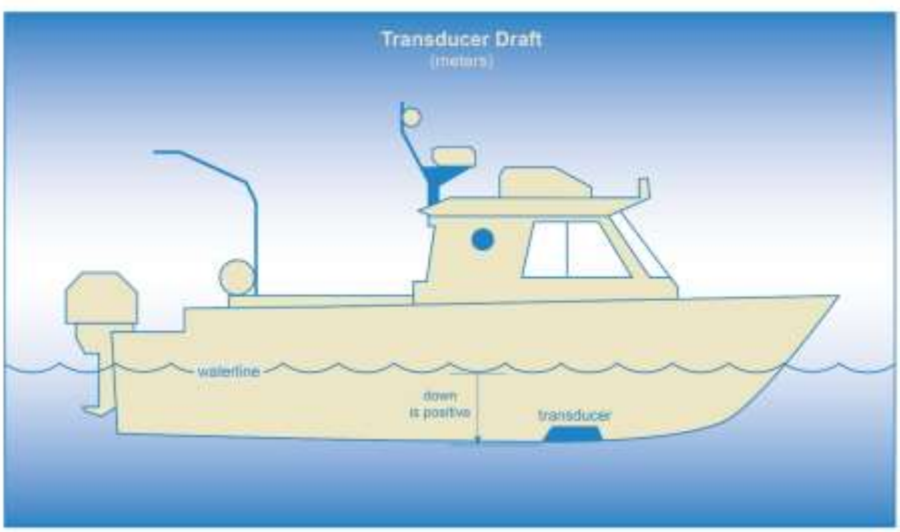
2.7 Data Loggers

Data loggers used in crowdsourced bathymetry are electronic devices or software used to record the sensor data over time. They record position, depth, date and time. The loggers connect to existing ship's depth sounders and satellite positioning systems. They write to files in a format defined by the designer of the data logger. The data recorded by the loggers can then be relayed to a Trusted Node, or data aggregator.

2.8 Sensor Offsets

Sensor offsets are the distances between the positioning equipment and the depth sounder transducer in a vessel. The offset measurements are required to correct the position obtained by the positioning equipment so that it is the same as the position of the depth sounder transducer. This greatly improves the positional accuracy of the depth data.


Longitudinal and Lateral Offset from GNSS Receiver to Sonar
measurement starts at GNSS Receiver
Bow
GNSS
Sonar
-1 meter
Stern
Negative Offset
Bow
Sonar
+5 meters
GNSS
+1 meter
Stern
Positive Offset


Transducer Draft
(meters)
waterline
down
is positive
transducer

2.9 Integrated Systems

[Suggest diagram of integrated system, i.e. depth sounder, positional system, motion sensor, and data logger – and the data flow within the system.]

3. **Metadata**

3.1 Introduction to Metadata

3.1.1 What is Metadata?

It is important to understand the difference between data and metadata. Simply put, data is information, and metadata is information about the data. In the context of crowdsourced bathymetry, the data are the individual point observations from a vessel (consisting of a depth, date, time, and geographic position), whereas the metadata provides the user of the data with additional facts about the dataset. For example, metadata can provide information about where the depth sounder transducer is mounted on a ship's keel, and what type of vessel collected the data.

3.1.2 Why is Metadata Necessary?

Crowdsourced bathymetry is the collection of depth data from a 'crowd' of seagoers, using a wide variety of sensors and techniques. For the data to be useful, there must be some unifying information that describes the factors influencing the data collection. Metadata provides that information, and ensures that all crowdsourced depth data has consistent collection criteria, nomenclature, and structure.

The metadata associated with crowdsourced bathymetry is vital, because it allows future end-users to make informed decisions about the quality and potential applications of the dataset and to apply enhancements or corrections if necessary. For example, documentation about the time and date when a depth measurement is collected allows a future data user to apply tidal corrections to data. Similarly, information about a transducer's vertical offset from the waterline, or its horizontal offset from a GPS receiver, allows a user to apply vessel draft and horizontal positioning corrections to the data. By applying corrections based on information in the metadata, end data users can greatly improve the accuracy and value, of the bathymetric data for research, industry or other applications.

A uniform metadata structure ensures that data is consistent and accessible, regardless of the platform that collected the information. The subsequent section provides guidelines for data and metadata structures that will allow seagoers to collect depth data in a format that enables easy data-sharing and crowdsourcing.

3.2 Metadata Descriptions

3.2.1 Introduction

This section provides guidance to the seagoer and Trusted Nodes about what metadata is required. A minimum of metadata is required for the seagoer to deliver data to the Trusted Node, however it is better if the seagoer can provide additional metadata, as each bit of information potentially increases the uses for the data. The tables below are based on the CSB GeoJSON data format 2.0.

3.2.2 Required Data

The seagoer should collect at least a minimum set of data to ensure that the depth data can be ingested by a Trusted Node, and then catalogued by the Data Centre for Digital Bathymetry (DCDB). The table below lists the required metadata fields.

Table 1. Required Data

Data Field	Description	Example
Longitude	Describes the longitude value of the horizontal geographic position, in WGS84, Decimal Degrees to six decimal places. Normally derived from the GPS NMEA GGA String; Positive = East; Negative = West	19.005236
Latitude	Describes the latitude value of the horizontal geographic position, in WGS84, Decimal Degrees to six decimal places. Normally derived from the GPS NMEA GGA String; Positive = North; Negative = South. Request as many decimals as can be provided.	40.914812
Depth	Describes the measured distance to the sea floor. Depth is always a positive value in *metres with accuracy of tenths of metres.* Normally derived from NMEA DPT data string.	7.3
Depth Units	Designates the depth measurement units. Always defined as metres.	Meters
Date & Time	ISO 8601/UTC Time Stamp of the depth measurement. Normally derived from NMEA GGA string .	2015-08-06T22:00:00Z
Time Units	Designates the time measurement units. Always defined as ISO 8601/UTC.	ISO 8601

3.2.3 Requested Metadata

Additional information about the vessel, sensor models, and sensor installation measurements allows an end user of the data to undertake such things as an assessment of the quality of the data, or to consider applying certain corrections. This increases the potential applications of the data for oceanographic research, scientific study, commercial applications and other uses. The table below lists metadata that the seagoer should provide whenever possible.

Table 2. Requested Metadata

Metadata Field	Description	Example
Platform Type	The type of vessel collecting the data.	Fishing, sailing, recreational, passenger, cargo, tanker, research vessel, other
Platform Name	The name of the vessel (open string)	White Rose of Drachs
Platform Length	Length of the vessel; a positive value in meters, with accuracy to the nearest metre.	65
Platform Length Unit of Measure	Always in metres.	Metres
ID Type	This designates the ID number provided. Vessels can choose only one type. Currently, only two types of vessel ID numbers are available: MMSI and IMO. If these IDs are unknown or not assigned, a None value is accepted.	Options for ID Type: IMO, MMSI, None
ID Number	Provides the input value for the ID Type chosen above.	369958000
Sensor Type Sounder	Defines the sensor type for depth sounders. This must always be defined as: "Sounder" (not an optional field that users can change).	Sounder
Sounder Make	Free text. In the future, a list of sounder makes will be provided to contributors (Trusted Nodes) via web API.	Sperry Marine (L3 ELAC)
Sounder Model	Free text. In the future, a list of sounder models will be provided to contributors (Trusted Nodes) via web API.	ES155100-2
Sounder Transducer	Free text. In the future, a list of depth sounder transducer options will be provided to contributors (trusted nodes) via web API.	Dual Freq 200/400 kHz

Sounder Draft	Vertical distance in metres from the waterline to a vessel's transducer. Draft value is always a positive value in meters with accuracy of tenths of metres.	4.6
Sounder Draft Unit of Measure	Always defined as metres.	Metres
Sounder Draft Applied	Boolean - true or false - designation for reporting whether or not the Depth values reported in the file have been vertically corrected for the Draft offset.	False
Sensor Type GNSS	Defines the sensor type for GNSS receivers. This must always be defined as: "GNSS" (not an optional field that users can change).	GNSS
GNSS Make	Free text. In the future, a list of GNSS receiver makes will be provided to contributors (trusted nodes) via web API.	Litton Marine Systems
GNSS Model	Free text. In the future, a list of GNSS receiver models will be provided to contributors (trusted nodes) via web API.	LMX420
Longitudinal Offset from GNSS to Sounder	Longitudinal offset from GNSS receiver to sounder. Values are in metres, positive moving from the stern to bow; i.e. when the GNSS receiver is aft of the sounder, the value is positive; and when the GNSS receiver is forward of the sounder, the value is negative. Accuracy given to the hundredths of metres.	3.52
Longitudinal Offset Unit of Measure	Always defined as meters.	metres
Lateral Offset from GNSS to Sounder	Lateral offset from GNSS receiver to sounder. Values are in metres, positive moving from port to starboard; i.e. when the GNSS receiver is on the port side of the sounder, the value is positive. Accuracy given to the hundredths of metres.	-0.76

Lateral Offset Unit of Measure	Always defined as metres.	Metres
Position Offsets Applied	Boolean - true or false - designation for reporting whether the final Position (Lon, Lat) reported in the file has been corrected for the lateral and longitudinal offsets between the GNSS receiver and the sounder.	False
Sound Speed	Value describes the assumed speed of sound used in the depth sounder to calculate the distance to the sea floor. Value is reported in metres per second (m/s). The average speed of sound in seawater is around 1500 m/s, and this is normally the default value used by depth sounder processing units. If left blank, users will assume 1500 m/s, unless they use an overriding sound speed environmental model.	1500
Sound Speed Unit of Measure	Always defined as metres per second (m/s).	m/s
Motion Sensor Information	Information from motion sensor, measuring the heave, pitch and roll of the vessel.	

3.2.3 Required Trusted Node Metadata

When a Trusted Node receives data from a seagoer, there are some required metadata that should be assigned by the Trusted Node before it delivers the data to the IHO DCDB. The table below lists metadata that a Trusted Node should provide.

Table 3. Trusted Node Metadata

Provider Contact Point Organization Name (orgName)	Trusted Node Name.	Sea-ID
Provider Email	Trusted Node email address. Used as	support@sea-id.org

	contact point when users of the data want more information.	
Unique Vessel ID “platform.uniqueID”	Generated by the Trusted Node, this number identifies the Trusted Node source, and also uniquely identifies the contributing vessel; the first five characters designate the Trusted Node, the sixth character is a hyphen (-), and the remaining characters are generated using a UUID. The UUID is consistent for each contributing vessel throughout the life of service of the vessel. However, **if the vessel chooses to remain anonymous to the end data users, the Trusted Node does not need to publish the vessel name in association with the UUID**.	SEAID-*UUID* Use *ROSEP* for Rose Point Navigation contributors.
Convention	Describes what CSB JSON format version is used.	CSB 2.0
Provider Logger	Name of software program or hardware logger used.	Rose Point ECS
Provider Logger Version	Version of the software or hardware logger used.	1.0

4. **Data Collection**

4.1 Introduction

Vessels range in size and capability, and may be outfitted with a wide variety of CSB data logging software, hardware, and data collection sensors. These variable equipment configurations require consideration when logging CSB data.

For instance, some smaller vessels may have limited space below decks, and may not have Very Small Aperture Terminal (VSAT) internet connectivity. They may be better served by data logging hardware that easily survives a power cycle and stores its data on an external USB drive that can be unplugged when the day is over. Conversely, a larger vessel may be able to support an automated data logging solution that can upload its data through a VSAT internet link every day. Pilot projects are currently underway to identify solutions to the limitations and develop systems that are as universal as possible.

~~Diversity of sources can also create a challenge in data exchange formats and project requests and requirements. It is hoped that Trusted Nodes will remain actively involved with the IHO Crowd Sourced Bathymetry Working Group (CSBWG), which works to standardize these processes. Whichever Trusted Node the mariner chooses to contribute data through, if the project is involved in the CSBWG, the data will be usable.~~

4.2 Crowdsourced Bathymetry Data Collection Process

There are three steps in the CSB gathering and transmittal process:

1. data collection;
2. storage; and
3. exchange.

4.2.1 Data Collection

Some CSB data logging methods require only minimal software to be active. For example, in an ECDIS or an ECS that already takes inputs from the depth sounder and the GNSS, the ECDIS or ECS manufacturer could easily record that data as CSB and forward it to a Trusted Node. However, most CSB data logging processes are currently not so integrated.

Typically, a seagoer will need to connect a separate data logger to the two minimum inputs: (1) input from the depth sounder, and (2) input from the GNSS.

Most navigation aids on a bridge transmit data in accordance with standards developed by the NMEA (National Marine Electronics Association). The NMEA 0183 standard is the standard format that most instruments use.

Most instruments have an output port incorporating a pair of copper wires on which "sentences" are broadcast. These data or "sentences" in the NMEA format is both human and machine readable.

An actual NMEA sentence from a GNSS unit may look like this, where latitude and longitude are visible:

$GPGLL,0424.99,N,11359.77,E,012636.21,A,D,*5E

The same GNSS may also provide a sentence that looks like this:

$GPGGA,071953.00,0424.9862,N,11359.7661,E,1,9,1.8,21,M,,M,,*68

Again, the position can be derived from this sentence. It also includes information about the accuracy, altitude, and time the GNSS unit acquired the position.

The same logic is used by the depth instrument, which gets input from the transducer. The depth instrument outputs a sentence like this:

$SDDBT,0006.0,f,0001.828,M,0001.0,F*3A

In this sentence, DBT stands for depth below transducer. The depth in feet, metres, and fathoms are also visible.

Integrated instruments on a bridge can read inputs from various sources, such as wind, water temperature, depth and position, and perhaps even information from the autopilot. These integrated instruments produce different NMEA sentences.

In order to start logging depth data, instruments must be connected to the data logger via the copper cables that carry the NMEA "sentences": the cable(s) that carry the DBT string (depth from transducer), and the GNSS cables that carry either the GGA (position, time, satellites used, and HDOP) or GLL (position, time, satellites, HDOP and status) string. The time record included in GLL and GGA strings is critical, because it allows an end data user to apply tidal corrections if desired.

4.2.2 Storing the Data

Single beam depth sounders provide a simple, relatively narrow beam depth reading directly below the transducer to the seafloor. Common single beam depth sounders may output about two megabytes of data or less (which is less than the size of a picture taken by an iPhone). Multibeam depth sounders can be found on survey vessels and some commercial, oceanographic, and other vessels. They provide a continuous swath of detailed depth information beneath the ship, and may output gigabytes (or more) of data per day. Vessels owners and operators contemplating taking part in CSB should ensure that they have adequate onboard data storage space to log data until they can transfer the data to a Trusted Node.

4.2.3 Data Exchange

After the CSB data is collected and stored onboard, it needs to be transmitted to a Trusted Node. Sending and receiving data at sea is challenging. It is possible to receive a 3G/4G signal on a phone or through a 3G/4G router on board. However, once a vessel is beyond sight of land, it may be necessary to subscribe to a satellite service in order to communicate with any party ashore.

AIS (Automatic Information System), which is commonly used to send certain navigational messages between vessels, can be used for certain bathymetric applications. However, transmitting whole data log files of depth soundings via AIS is inefficient.

Depending on the CSB project or Trusted node, the CSB data logging equipment will have a particular method of transferring data. Some methods of data transmission will be simple, such as sending a USB stick via mail to the data centre, connecting to a smartphone via Bluetooth to upload files, or directly plugging into a VSAT modem. Other methods may be more expensive. Inmarsat SAT-C or FBB (Fleet Broadband) is billed per-megabyte, so transmitting the data on land via Wi-Fi may be cost-saving. Note that CSB data is not normally time sensitive: as long as data is consistently sent to a data centre, it is a valuable contribution.

4.3 Best Practices and Recommendations

Many of the existing CSB projects have developed "best practices," outlining what should and should not be done in order to be a successful contributor. These might be primarily of interest to developers, but anyone interested in CSB can learn from these.

4.3.1 Keep the Data in the Original Format

Stripping data from an NMEA sentence and only saving parts of it is not recommended. Saving the data in its original format will help validate data recordings and troubleshoot potential anomalies in the data. For example, if only *depth in metres* from the DBT string is saved, then a strange reading cannot be compared to the depth in feet. Or, if only the latitude and longitude for position are saved, the detailed information in the GGA sentence, such as a quality assessment and the timestamp on which the fix was taken, will be missing.

While the IHO Data Centre for Digital Bathymetry only accepts GeoJSON or XYZT data, in which the depth and position are only noted in one format, having the additional data available is highly recommended.

4.3.2 GPS Latency and Quality

As mentioned above, GGA provides more information than GLL. If both sentences are available, save GGA. If the GPS unit provides even more information (such as latency), save that also. Only a few units do.

4.3.3 Real-Time Clock

The internal clock of a computer typically runs 'off' by several seconds per week and synchronizing the clock by NTP (Network Time Protocol) is only possible if there is a network connection.

The GGA and GLL sentences from a GPS unit provide all the information needed. Set the system clock to the correct date and time (you need the date anyway) for logging and debugging, but use the time provided by the GPS and set the internal clock to that input if you can.

Some systems use internal counters (first recording, second recording, and so on) but this is not recommended. If it is necessary to rely on the system clock for the date, document (and save) the process of setting this, and investigate how it will behave after a long period without power to the system.

4.3.4 Time Synchronization of Sensor Input

The NMEA sentences, a first stream from the GPS unit and a second from the depth sounder, will come at intervals dictated by the unit's capacities. The GPS might send a location sentence every second; the depth sounder might send one every three seconds.

At its simplest form, it is necessary to 'couple' both bits of information as well as possible. At its most complex, calculate 'where precisely' the depth sounder was at the time it took its reading-meaning, calculate the location based on the timestamp of the depth measurement, and find the spot in between the two closest position measurements.

It is essential to store all measurements with an accurate timestamp 'at the time of measure' and then allow for the complex calculations to happen in post processing (even in the data centre if the collected data is not used onboard). Saving timestamps with every reading allows data to be re-processed if changes or improvements in interpolation methods occur in the future later. This approach is strongly recommended.

4.3.5 Varying Draft, Keel Depth

As described above, the draft and position underwater of the depth sounder can be transmitted with the collected data. If the vessel collecting data has a varying draft (for example, because it takes a lot of fuel on or offloads goods), it is important to collect this information and connect it to the series of depth readings.

It can be as simple as storing, with a timestamp, the current draft of the vessel. So, every time the draft changes significantly, record it in the stream of collected data. This will allow for adjustments during post-processing.

4.3.6 Compliance with NMEA Specification & Specifically with Opto-isolation

As mentioned above, NMEA has been developing and promoting its standard through which, for many years, messages have been sent between bridge equipment. NMEA also makes sure that bridge hardware is developed according to the same standards.

Hardware interoperability is critical. In short, one piece of hardware could negatively affect the other. On an integrated bridge, where everything is connected, it would be difficult to identify and locate one piece of equipment that is malfunctioning. Isolating the signal through optometry, away from the small current on a copper wire, is one way to ensure this.

Commented [REW1]: what does this mean?

It is recommended that this potential issue is at least considered on larger vessels and in certain data loggers. Vessels that can only fit type-approved hardware will request it. Hardware developers considering setting up a CSB project should keep this in mind. Seagoers selecting a Trusted Node to work with, should consult a professional bridge equipment installer on this issue, and ensure the NMEA standard is respected throughout.

4.3.7 Continuity of Electrical Power

Continuous power aboard ships is never a guarantee. Some vessels invest or are required to carry a large and well-maintained Universal Power Supply (UPS) for all the bridge equipment. However, there are still times when the transition from shore power to a generator causes a momentary loss in power and data loggers must reboot and recover. Consider investing in a small, serviceable, built-in battery for the CSB logger to ensure smooth operation.

4.3.8 Hands-free Operation

The best results from any CSB project occur when the user is passive and measurements are purely based on the technology doing its job. For example, when Google developed an algorithm for calculating average speed on roads, they found that once it started tracking phones in cars, rather than user-reported data, the algorithms became far more accurate. Relying on a user to 'report' data is rarely consistent, and supports only a certain kind of report (such as slow-downs, accidents, etc.).

There are times, however, where a user's interaction is valuable and even required; for example, when the transducer depth or vessel draft changes, or the sensor configuration is changed, such as the GNSS antenna or the position of a sensor. A data logger and its setup should allow the user to record such changes at the time that they occur and after the occurrence, in order to allow for retrospective changes, should an required entry be temporarily forgotten or time did not allow the change to be made at the time.

5. **Uncertainty**

Crowdsourced bathymetric data are subject to a number of quality issues that could cause depth sounder measurements to differ from the true depth of the seafloor. For example, a depth sounder relies on an accurate measurement of time, which is then converted to depth based on an assumption about the speed of sound in water. If the sound velocity estimate is incorrect, then the depth will be incorrect, as well. Similarly, if the sound wave reflects off a fish in the water column, or if the depth sounder captures acoustic noise from other boats in the area, errors could creep into the data.

These errors, and others, could lead to uncertainty in the validity and position of a depth measurement. This uncertainty in the quality of the data should be taken into account when the data is processed, stored and used. This chapter presents features of uncertainty that are likely to be of general interest, as well as issues relevant to individual observers, trusted nodes, and end users of the crowdsourced bathymetry database.

5.1 Meaning, Sources, and Consequences of Uncertainty

5.1.1 The Meaning of Uncertainty

In a scientific context, "uncertainty" is a measure of how significantly different a measurement could be from its true value. The best way to calculate this uncertainty would be to directly compare a collected measurement to its actual real-world value. Unfortunately, it is usually impossible to physically verify the true value. The best we can do is estimate the scale of error in the measurement, and express it as a degree of uncertainty. Estimating the uncertainty of a depth measurement allows the end user to judge whether the data is suitable for a given purpose, and allows for the selection of appropriate processing techniques.

5.1.2 Sources of Uncertainty

All of the measurements that are made to support bathymetric mapping are heavily composite, meaning that a number of different measurements are combined to construct the depth estimate. In order to maintain some level of control over the complexity that can ensue, it is common to categorise the different types of uncertainties that can occur, and then estimate their magnitudes before combining them.

The most common method for categorizing uncertainty is to estimate the precision (or variance) and accuracy (or bias) of the observations. Figure 1 and 2 show examples for observing systems with high or low precisions and/or accuracy. Ideally, all observations would be accurate and precise, but random variations in measurements can result in an observation that is on average correct, but which varies a lot about the correct value (accurate, but not precise). Well-calibrated depth estimates are often of this kind (Figure 3).

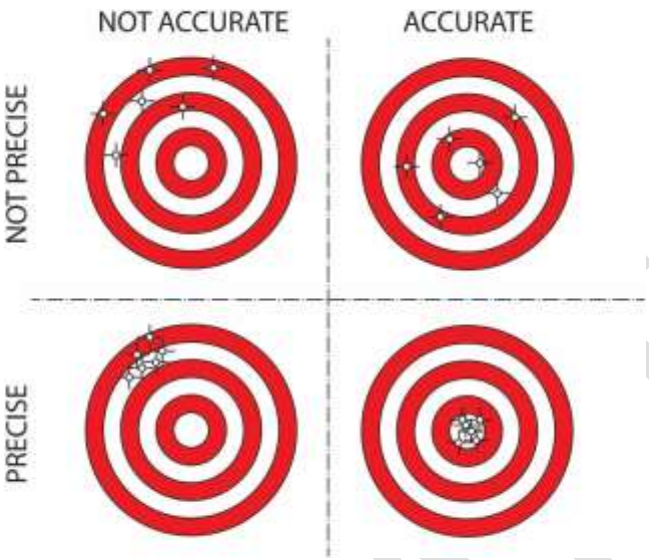


Figure 1. *Effects of accuracy and precision (variance and bias) of measurements on the ability of a system to measure.*

The precise, but not accurate, situation can occur when 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, rather than being measured, observations will be offset from the true value of depth, even though they are in relatively close agreement about the apparent depth. In most cases, a correction could be applied to improve this situation; however, this might not be practical or time-efficient. It might be more pragmatic to assess the level of bias that is believed to exist, and consider it an uncertainty.

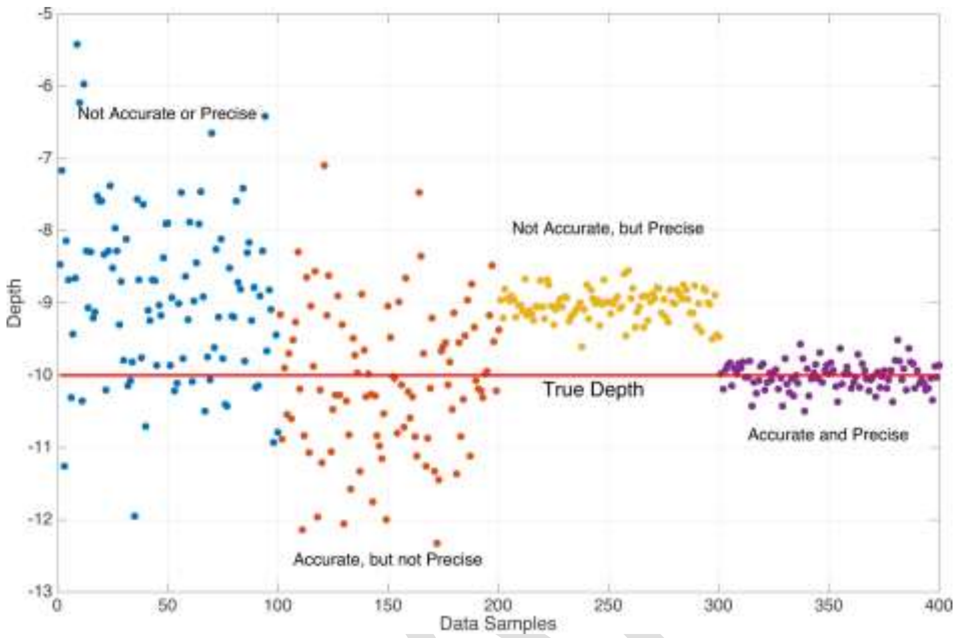


Figure 2. *Example of depth measurements from the four quadrants of Figure 1.*

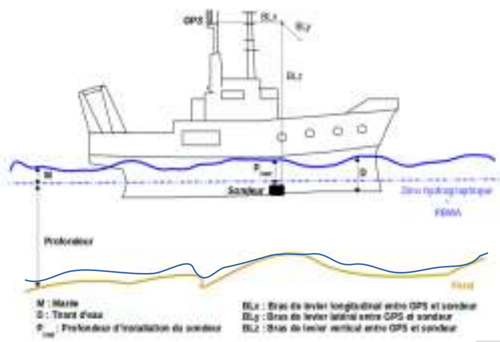


Figure 3. *Effects of accurate, but not precise (i.e., mostly random) uncertainty on depth sounding. Here, on average the depth measured (blue line) is correct, but point to point it varies from the true depth (yellow line).*

In addition to the uncertainty of the measurements, there are two further uncertainty components that are important for crowdsourced bathymetry:

- **Integration uncertainty**. This is an uncertainty associated with failure to install an instrument correctly, or for failing to adequately document the installations that were done. This could be considered a type of systematic uncertainty, since the behaviours of the uncertainty are very similar, but is often better considered separately, since it is something that can reliably be done for all systems. The uncertainty caused by not measuring the offset of the depth sounder transducer from the waterline (Figure 4) or the offset between the GNSS and the depth sounder (Figure 5) is this type of uncertainty.

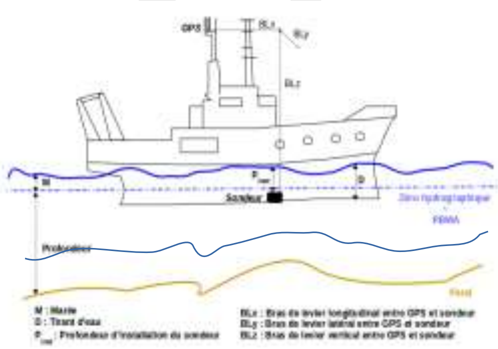


Figure 4. *Example of the effects of not correcting for vertical offsets. Here, not correcting for the offset of the depth sounder from the waterline leads to a measurement (blue line) that differs significantly from reality (yellow line). This gives a bias (systematic) uncertainty to the measurements.*

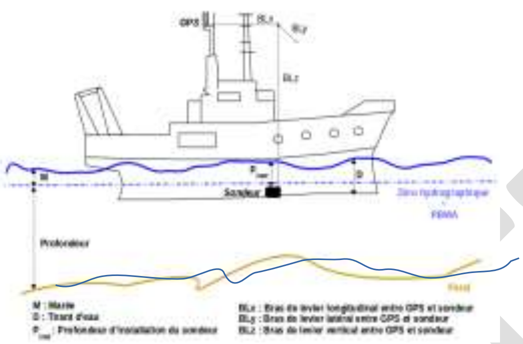


Figure 5. *Effects of not correcting for horizontal offsets. Here, not measuring the horizontal offset between the GNSS receiver position and the depth sounder results in along-track offsets of seafloor features. Blue line: measured; yellow line: reality.*

- **Modeling uncertainty.** Every measurement is an abstraction of the real world. We know that the seafloor is complex, but often model it as a continuous mathematical surface and interpolate depths where there were no actual measurements. Any interpolated depth will only be as accurate as the model is valid. Since the only alternative is to develop a model with increasingly more data, such that eventually no interpolation is required, we often have to accept the approximations based on interpolation, but include a component of uncertainty to reflect the fact. This is the most difficult of the uncertainties to estimate, and is often ignored.

Many datasets do not contain sufficient data to completely specify the measurements being reported, or the products which are subsequently constructed. For example, if a dataset consists of measurements of depth that are more than 50m apart, it is impossible to assess the shape, location, or presence of objects smaller than 100m - and often significantly larger. It is, of course, possible to interpolate the data to an arbitrary resolution - generating, say, a 1m grid. However, the information in this grid at the smaller scales is mostly an artefact of the assumptions built in to the interpolation scheme, rather than what is actually present in the real world.

Consequences of not understanding this interpolation is that data may appear to be accurate, but does not reflect the actual environment. Gridded data, in particular, can be very visually persuasive, which can result in the erroneous belief that the data are better than they actually are.

5.1.3 Estimation & Expression of Uncertainty

Types of uncertainty are often estimated separately, and are then combined into an overall assessment. This works well when there is sufficient information or metadata available to help with the calculation. Unfortunately, that is not always the case.

The most common method for estimating uncertainty is by collecting the same observation multiple times, and then assessing the degree to which the measurement changes between different observations. Taking all of the measurements together, it is possible to estimate the average depth returned, and the degree of variability of the depths about this value.

Uncertainty can be expressed as a range of values in which the true value of the measurement is expected to lie. So, a depth could be specified as being "between 12.3 and 14.2m, 95% of the time". Where the range is either known or assumed to be symmetric, the mean value and spread might be given, so that the depth might be specified as "13.25 ± 0.95m, 95% of the time". Whichever method is used, it is important to clearly identify the limits of the estimate.

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 on a scale of 1 to 5 based on a subjective assessment how the data was collected, or by comparison with coincident data.

5.1.4 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 the 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 (95% CI), it would be unwise to assume that a ship had at least 12.0m clearance in this depth area: the uncertainty principle being used here says that this is true only half of the time, Figure 6(a), which is surely lower odds than any prudent seagoer would allow for a navigation decision. A value of 11.74m would be a better choice, Figure 6(b), but if the seagoer wanted there to be less than a 1:1000 chance of the depth being shallower than the declared value, a depth of 11.34m would need to be used, Figure 6(c). Clearly, the "safe" depth depends on the user's need, and it would be incorrect, and unwise, to report simply the mean depth.

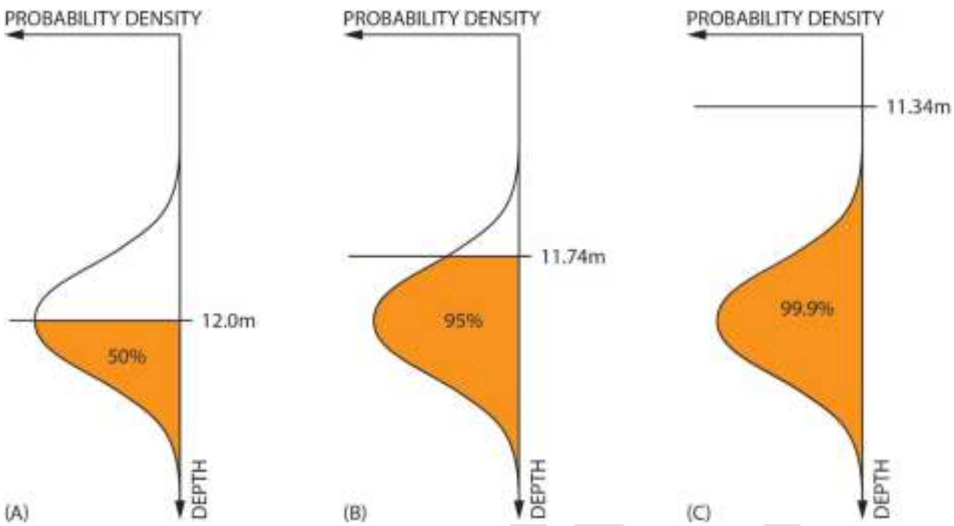


Figure 6: *Examples of shoal-clearance depths for different probabilities of excession, based on the same basic uncertainty estimate of 12.0 ± 0.3m (95% CI). Assuming a 12.0m clearance is only true 50% of the time (left); a 5% probability of being shallower requires the depth to be reduced to 11.74m (middle); a 1:1000 chance of being shallower requires a clearance depth of 11.34m (right).*

Like depths, uncertainties are only estimates. They are a best guess, based on assumptions. It is possible for an observation to be provided with an uncertainty estimate that does not reflect the difference between the declared depth and the true depth. Data users should be aware of this possibility, and take appropriate steps to rectify the situation if it occurs.

Consider the data in Figure 7. Here, the data from crowdsourced observations have been compared to high-resolution, authoritative data. There are significant differences between the two in some areas. The error 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 identify errors or uncertainty in the data.

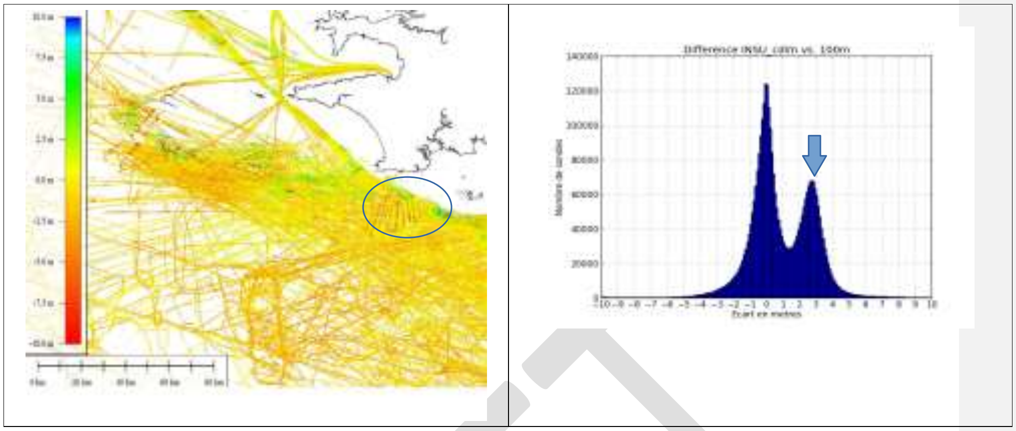


Figure 7. *Difference between crowdsourced observations and a reference grid model (data courtesy of SHOM). Errors in the crowdsourced observations are clearly seen in plan view (left), and are reflected in the bimodal distribution of differences (right). The uncertainty associated with the crowdsourced observations might not reflect these differences if the observer's metadata was incomplete.*

5.2 Uncertainty Guidance for User Groups

5.2.1 Uncertainty for Individual Observers

End data users need to know if corrections (such as vessel draft or tidal offsets) need to be applied to crowdsourced datasets before use. You can make it easier for potential users of the data to assess this information by providing metadata about which corrections were made by the observer, and how they were applied. The more corrections that a potential user is able to make, the more useful the data becomes.

Supplying information about corrections is only part of the story. Each correction has an effect on the overall uncertainty of the depth measurements. Recording how corrections were determined and applied, or an estimate of the uncertainty, is also very important. There is no "right" value of uncertainty, although underestimating the uncertainty is more problematic than overestimating it. If you feel that there was a degree of uncertainty in a correction that you applied to the data, be sure to indicate that in your metadata.

The biggest contribution that individual observers can make to uncertainty estimation is to provide data that will assist in the independent assessment of uncertainty. The simplest method is to provide data that approximates repeated measurements of the same depth, so that statistical methods can be used to estimate the variability of the measurements directly. This can mean keeping the depth sounder running while the vessel is stationary, but is most effective when done over a known depth so that biases in measurement can also be assessed.

Known depths, sometimes called calibration surfaces, are sometimes established by hydrographic agencies or harbour authorities on prominent markers (for example, the channel buoy, the fuel dock) or in well-trafficked areas. Collecting data over these areas makes a dataset significantly more valuable. Occasionally doing a cross-check (in other words, sounding the same area on a track orthogonal to the previous one) can also be useful in identifying internal dataset inconsistencies.

How often calibrations need to be done depends on the rate of environmental changes around the observing ship. Island areas where there is significant riverine freshwater discharge, there may be differences in reported depths that are actually due to changes in the water's salinity. A thorough calibration for a depth sounder system can be quite complex, but in many cases is not required. Details on how to do a full calibration can be found in IHO publication M.13 (Manual of Hydrography), but even something as simple as recording data over a known marker could be enough.

Individual data collectors can summarise all of the uncertainties associated with their depth observations in a table, known as an "uncertainty budget." An example table is shown in Table 1. Observers may not be able to fill in all of the details, but the more information that you can be supplied, the more valuable the observations become.

Sources of Uncertainty	**Applied YES/NO**	**Example of assessed standard uncertainties values at 50 m (m)**	**REMARKS**
Static draught Setting		0.1	Specify if the draught value was set in the depth sounder.
Variation of draught		0.05	Change of draught due to variation in loading condition. Average draught to be assessed from full loaded and ballast condition.
Sound speed		0.07	Measurement is based on the equipment. It depends on temperature, salinity and depth.
Depth sounder instrumental uncertainty		0.1	Not to be confused with the resolution of the instrument. It varies with the type of equipment
Motion sensor		0.05	Measurement depends of the equipment.

Dynamic draft, settlement and squat		0.1	Effects data primarily in shallow water. Settlement depends on speed of vessel and draught.
Tide measurement		0.06	Tide is the variation in the sea level and depends on the station from where the measurement is done. Not applicable for depths more than 200m.
Sensor offset		0.01 – 0.1	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 type of differential correction being applied.

Table 1. *Sample uncertainty budget for a shallow water depth sounder and modern GPS system.*

Creating a complete uncertainty estimate can be time consuming. Some of the uncertainties are more important than others, depending on where a vessel is observing. For example, in shallow water, recording draught, squat, and water level is essential. However, in deeper water, sound speed information is more important. In most cases, motion effects are likely to have relatively small impact on uncertainty. Understanding these differences can reduce the complexity of uncertainty estimation.

5.2.2 Uncertainty for Trusted Nodes

Trusted Nodes are in an ideal position, as aggregators of observations from multiple sources, to generate uncertainty estimates for the data they transmit to the DCDB. They have the opportunity to cross-check between observers, remove data biases, calculate the uncertainty associated with individual observers, and potentially correct for them. Such activities can greatly increase the value of the observations being passed to DCDB.

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. Data aggregators may also choose to collaborate with harbour authorities to establish areas of known depth where individual users can calibrate their depth sounder measurements. Such calibration sites would facilitate efforts to calculate uncertainty in individual datasets.

Analysis of multiple user observations within the same area could also be used to establish baseline uncertainties for observers, and to flag dubious data. Trusted Nodes could then establish a calibration and uncertainty history for each observer, which could be contributed to the DCDB as part of the dataset's metadata.

Analysis of multiple observers might also allow Trusted Nodes to implement cross-calibration, by using data collected by an observer with well-established uncertainty and calibration to determine the installation or measurement uncertainty of other observers in the area. Ideally, the known observer would be an authoritative source, but could also be an observer which has been tracked for some time, and has proven reliable in calibrations against authoritative sources. Metadata of this kind can help database users establish confidence in individual observers.

Trusted Nodes also have the opportunity to make dataset corrections that individual observers cannot. For example, it may be difficult for many observers to establish an uncertainty associated with squat corrections or 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.

Trusted Nodes will have a more direct relationship with the individual observers than the DCDB or database users, and as a result they are well-placed to vet 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 in an ideal position to encourage individual observers to improve the metadata that they provide and to attempt data corrections. They might offer individual observers feedback on areas for improvement, which would improve their deliverables and increase end user's confidence in their data quality.

5.2.3 Uncertainty for Database Users

As end-users of the crowdsourced data, database users must interpret the uncertainty information provided with the dataset, and generate new uncertainty estimates for the end results of the work that they wish to do.

Database users should be aware that the uncertainties provided by individual observers, or assessed by Trusted Nodes, might not be consistent. The uncertainties could be assessed differently by different observers, and may not have been verified against authoritative sources of depth information. Low levels of uncertainty, in particular, should be treated with caution.

The DCDB provides no guarantee of the correctness of crowdsourced bathymetry observations. If some of the observations have uncorrected vertical offsets, and are inconsistent with other observers or authoritative source, there is no guarantee that this will be flagged in the metadata. Higher *Confidence of Reporting* assessments for an observer may increase dataset confidence, and some Trusted Nodes might provide stronger guarantees for data that they aggregate. The database user, however, must assume that residual blunders might exist that are difficult to capture in conventional uncertainty statistics.

End data users should consider the combination of uncertainties associated with a dataset before using the data.

Database users should be cautious to avoid over-confidence in uncertainty values when using interpolation methods that estimate their uncertainties from the geostatistics of the observations (e.g., kriging). Interpolating sparse datasets can be misleading, since the data density may not support accurate results. Figure 8 provides a diagrammatic example of problems that can arise from applying geostatistical interpolation to sparse datasets.

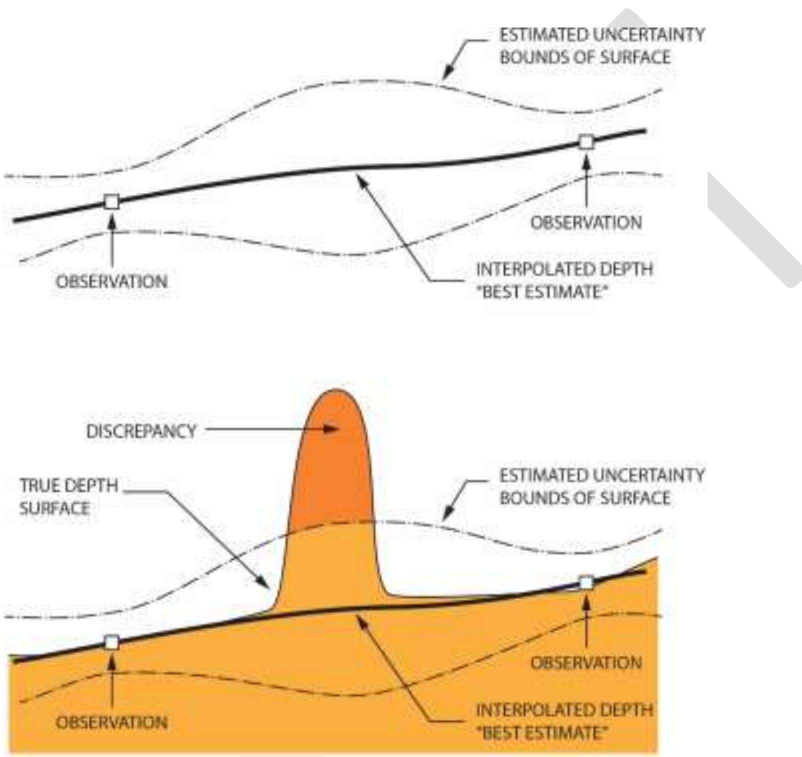


Figure 8. *Example of problems that can occur when predicting uncertainty from sparse data, where all objects are not captured in the dataset. From the data (top diagram), geostatistical techniques might predict an uncertainty that the user, without further data or reference, might assume to be the outer 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.*

One problem with using geostatistical interpolation to predict depths from sparse datasets is that the assumptions made by the model (that all significant variability is captured by the geostatistics) are not valid for the real world. Database users should be aware of this, and should identify how they compensated for sparse data in the dataset.

Database users should always consider providing an uncertainty estimate with any end product that they generate from using the data. There are multiple methods by which uncertainty can be specified. If a user creates an interpolated depth surface, the uncertainty could be provided in terms of the standard deviation of depth expected at each point, or at the 95% confidence interval, or by other methods. There is no universally accepted best practice for the statement of uncertainty, although the 95% confidence interval is very common. What is essential is that the type of uncertainty being reported is well documented, and that this documentation is embedded in the product's metadata. Without such documentation, the value of the uncertainty statement is greatly diminished.

As the translators of observations into products, database users are ideally placed to identify problems with individual observers or datasets. Database users who identify outliers or anomalous observers, are encouraged to communicate this information to the DCDB.

6. **Data Contribution**

This section details the process for contributing data to the IHO Data Centre for Digital Bathymetry, and specifies the required data formats. Data collectors of CSB are strongly encouraged to provide their data to the DCDB to help fill gaps in global bathymetric coverage. The resulting data in the database will be available to the general public to use as they see fit. Individual contributors can join an existing Trusted Node, or work with the DCDB to become a Trusted Node. Providing depth data through a Trusted Node will result in consistent formats with sufficient metadata to allow end users to assess the data quality.

6.1 IHO Data Centre for Digital Bathymetry

The IHO DCDB was established in 1988 to steward worldwide bathymetric data on behalf of the IHO Member States. The Centre archives and shares depth datasets contributed by seagoers 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. NCEI stewards global bathymetric data to international standards and provides it to the public freely and without restrictions. All data hosted by the DCDB is accessible online via interactive web map services. Spatial footprints of bathymetric data that is not publically available or hosted on other sites are provided with metadata for easy search and discovery.

6.2 The Trusted Node Model

The DCDB currently supports contribution of CSB data via a network of 'Trusted Nodes'; organizations or individuals that serve as liaisons between a like group of seagoers and the data centre. This model ensures clarity of requirements and data consistency, while minimizing the effort on individual seagoers to participate in contributing data to the public cloud. In the future, the DCDB plans to expand its capability to support other models, including individual seagoer contributions.

For submissions to the DCDB, a Trusted Node's procedure will need to be established through agreement with the DCDB. DCDB works with each Trusted Node on an individual basis to customize a streamlined process for contributing data. For example, file format, transmission protocol, minimum metadata fields, authentication method, data logging, and data aggregation techniques are weighed to select the best option, with the expectation that these can be updated to respond to changes in technology and software. Minimally, the Trusted Node needs to assign each vessel a unique identifier and adhere to the metadata and data format guidelines outlined in Chapter 3. Parties interested in becoming a Trusted Node should contact the DCDB, dcdb@iho.tint.

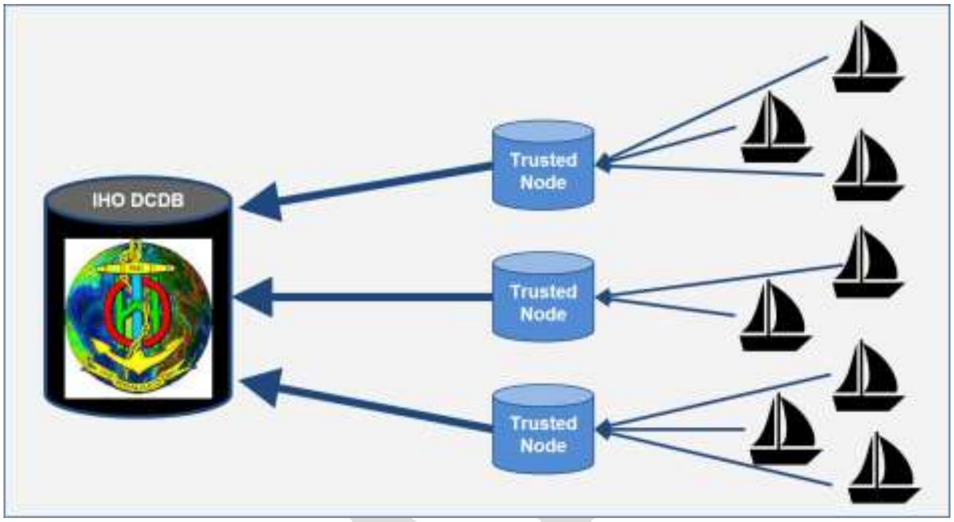


Figure Caption: Illustration of the flow of data acquired from vessels and packaged in a consistent manner and format for by the Trusted Nodes for submission to the DCDB.

6.2.1 Transmission Protocol

For submission to the DCDB, Trusted Nodes have the option of hosting the data and making it available for retrieval by the DCDB using a standard network protocol (such as: FTP, HTTP) or by pushing the data to the DCDB via an HTTP post. Multiple files or a single file can be submitted, depending on the data and metadata format the Trusted Node is using and the frequency of submission.

6.2.2 Data Aggregation

Crowdsourced bathymetry data is collected using different approaches depending on the platform and instrumentation. This results in a variety of files and file sizes. Files should be kept to a minimum for data transmissions. This requires that the aggregation method considers upper limits when determining file number and size limitations. It is a best practice for the Trusted Node to produce a new file every time new logging starts or at a regular time or file size limit.

6.2.3 Data Logging Rate

Where possible, it is suggested that data logging occur at the maximum available rate that the instrumentation can handle, or one hertz when a vessel is in motion. It is acceptable to lower the logging rate when a vessel is anchored or in port.

6.2.4 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.

6.2.5 Data and Metadata Formats

All data contributions should conform to the data format and metadata standards described in the "Metadata" section of this document (Chapter 3), unless separately and specifically agreed otherwise by the Director, IHO DCDB.

Data are currently being accepted in two formats: GeoJSON and CSV with a JSON metadata header. Data should be delivered as point soundings with the required fields of latitude, longitude, depth, and time for each sounding. Trusted Nodes may accept other data formats from their contributors if they can convert the data to the specified formats for delivery to the DCDB.

6.3 Overview of CSB Data Flow through the IHO DCDB

6.3.1 Submitting CSB data

CSB data submitted to the IHO DCDB are automatically verified upon receipt. The verification confirms that the data are from a trusted node and that the submission contains valid file types. The files are then logged in a submission tracking system at the DCDB. Ingest scripts convert formats as necessary, store the GeoJSON files for user access and long term archive, and populate a metadata catalogue. An extract, transform, and load process then creates file geometries and populates a spatial database with the geometries and a subset of the metadata.

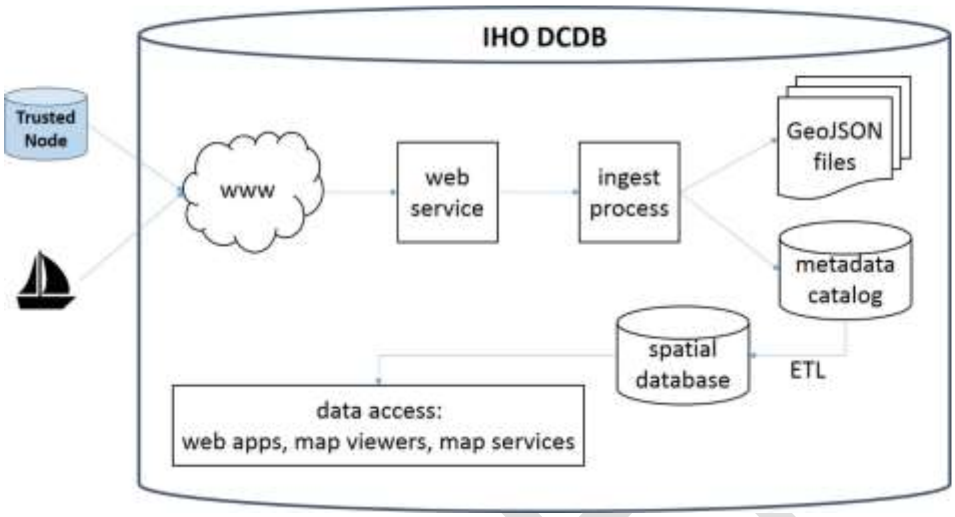


Figure Caption: A schematic of the flow of CSB data from the mariner to the IHO DCDB to the public

6.3.2 Accessing CSB data

The spatial database feeds a map viewer (http://maps.ngdc.noaa.gov/viewers/csb/index.html) allowing for data discovery. The map viewer is an online tool where users can search for, identify, and eventually obtain CSB data. To help users find the specific data that they're looking for, the map viewer contains filter capabilities that will match a specified time range and/or vessel. Users can then identify data files geographically using the Identify tool which provides selection options 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 files identified. Further clicking on a file name yields additional information about the file. Current work will add download options to the map viewer so these data may be freely obtained and reused for industry or research needs.

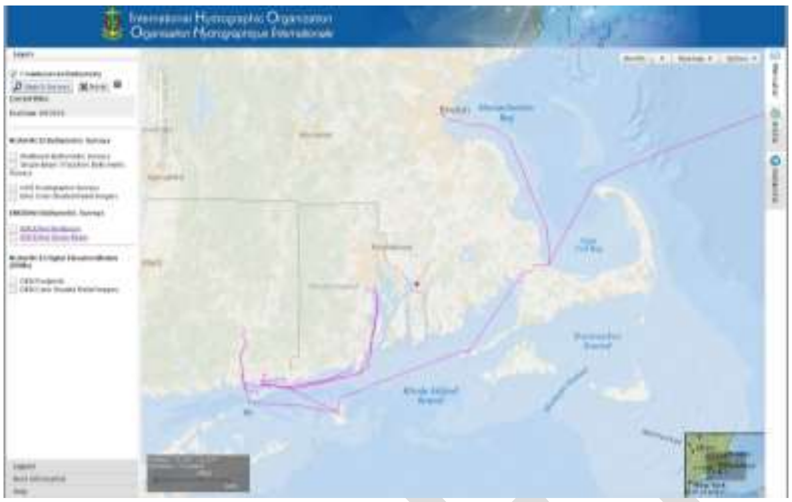


Figure Caption: Screen grab of IHO CSB Data Viewer which provides data discovery of CSB data.

7. Legal Considerations

7.1 Introduction

The following notes, which are not exhaustive, are intended for information only, and do not constitute legal advice.

Those considering taking part in the IHO crowdsourced bathymetry programme should be aware of the following potential legal considerations:

- Seagoers proposing to collect bathymetric data as a "*passage sounding*" activity should 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, an intermediary or a user, should 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 crowdsourced bathymetry database should be aware that they may expose themselves to liability if they provide value-added or other services that are based on the data.

7.2 Maritime Jurisdiction

Under international law, as reflected in the 1982 United Nations Convention on the Law of the Sea (UNCLOS), Coastal States have the right to impose certain restrictions within waters under their jurisdiction. Coastal States may have differing views on whether collecting bathymetric data on passage and providing it to the IHO DCDB for the common good is considered acceptable within the framework of the restrictions they impose under UNCLOS. In this context, the collecting of bathymetric data on passage and providing it to the IHO DCDB without permission could affect such things as the rights of innocent passage within territorial seas or contravene the requirements set by a country to seek permission to collect data in areas extending beyond their territorial sea. Potential CSB collectors should, therefore, ensure that the collection of passage soundings is acceptable in the waters through which they are voyaging.

To assist potential CSB collectors, the IHO publishes a list of its Member States that have indicated that the collection of CSB as part of passage sounding is acceptable within the waters under their jurisdiction. This information is kept up to data as much as possible and is available on the IHO website - *www.iho.int*.

7.3 Rights and Responsibilities

The principles of the IHO CSB programme are similar to many other initiatives where environmental data and information are collected on a voluntary basis by 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 seagoers 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.

However, recent advances in technology, which assist Volunteered Geographic Information (VGI), including CSB, to be collected, have moved *passage sounding* from a predominantly manual/manuscript data collection method to one that can be achieved efficiently and cost effectively using digital technologies. At the same time, the growth in VGI, in support of such things as urban mapping and the identification of hazards during natural disasters has also given rise to concerns over security and legal liability. The legal frameworks in relation to VGI are, as yet, not well developed.

It is therefore important that all parties participating in the IHO CSB programme carefully consider their rights and responsibilities in relation to the various legal jurisdictions under which they are operating.

The following notes are provided in relation to several of the potential participant communities in the IHO CSB programme with regard to their rights and responsibilities. The notes, which are not exhaustive, are intended for information only and do not constitute legal advice.

7.3.1 Licensing Regime for IHO CSB Data

The IHO CSB Programme operates under the Creative Commons licensing framework (*www.creativecommons.org*), currently using the "*Attribution 4.0 International*" license (CC BY 4.0) (*https://creativecommons.org/licenses/by/4.0/*) for CSB collectors and intermediaries such as Trusted Nodes, and the "*Attribution 3.0 IGO*" license (CC BY 3.0 IGO) (https://creativecommons.org/licenses/by/3.0/igo/) for the IHO DCDB. 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.

7.3.2 Crowdsourced Bathymetry Collectors

CSB collectors are expected to acknowledge that in providing their data for inclusion in the IHO DCDB database, they are doing so in good faith and for the purpose of improving bathymetric knowledge of the world's seas, oceans and waterways. 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, even commercially. CSB collectors cannot revoke these freedoms as long as users of their data follow the licensing terms.

CSB collectors should also consider their position in relation to the subsequent use of their observations. In all cases, the negligent contribution of erroneous data may expose the contributor to legal consequences in the event of an incident when the data is used subsequently. However, the position of a supplier of potential aid or assistance (in this case CSB data, provided in good faith for the common good) may be viewed in different ways in different countries.

CSB collectors should carefully consider their responsibilities in relation to the various legal jurisdictions under which they are operating. As indicated earlier in the notes, to assist potential CSB collectors, the IHO publishes a list of its Member States that have indicated that the collection of CSB as part of passage sounding is acceptable within the waters under their jurisdiction. Wherever possible, the IHO will also provide information on the "good Samaritan" protection that the laws may provide in each country.

7.3.3 Trusted Nodes

If the bathymetric data provided to the IHO by a CSB collector is passed to the IHO DCDB through a Trusted Node or a similar intermediary arrangement, then the free-use of the data provided under the Creative Commons license granted by the data collector should continue to apply. This may require a Trusted Node or a similar intermediary arrangement to harmonize any other applicable licencing arrangements with the CSB collector in order to preserve these principles.

If the Trusted Node or similar intermediary arrangement was to adjust, manipulate, or alter the data, the Trusted Node or similar intermediary arrangement takes full responsibility for any potential consequences resulting from that action.

7.3.4 IHO Data Centre for Digital Bathymetry

The IHO will observe the licensing conditions granted by the CSB collector and will grant an open license for any user to use the CSB data in the DCDB. In doing so, 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. This means that claims for liability against the suppliers of the data obtained through the IHO DCDB programme cannot easily be pursued through national courts.

Anyone may use the CSB data in the DCDB for their purposes, taking note that the data user must carefully consider the nature and the uncertainty of the data and whether it is fit for the purposes intended by the user.

Any data user (including a Trusted Node) that provides a value-added or other service based on the CSB data, should be aware that they take full responsibility for any potential consequences resulting from that action.

Users of the CSB data must give appropriate credit, provide a link to the license, and indicate if any changes were made. This may be done in any reasonable manner, but not in any way that suggests that the IHO endorses the user or the use that the user has made of the data.

References:

1. Division for Ocean Affairs and the Law of the Sea (DOALOS), 1997, United Nations Convention on the Law of the Sea, New York: United Nations. 294 pp. (available at http://www.un.org/Depts/los/convention_agreements/texts/unclos/closindx.htm)
2. Argo Ref: http://www.argo.ucsd.edu/About_Argo.html

3. Mateos and Gorina-Ysern, “Climate Change and Guidelines for Argo Profiling Float Deployment on the High Seas.”
4. J. A. Roach, “Defining Scientific Research: Marine Data Collection
5. Beckman and Davenport, “The EEZ Regime: Reflections after 30 Years” 2012

Appendix A – Data Contribution Format

GeoJSON validator

Crowdsourced Bathymetry JSON
Format Version 1.0
Last Update: October 21st, 2015


```
{
  "type": "FeatureCollection",
  "crs": {
    "type": "name",
    "properties": {
        "name": "EPSG:4326"
    }
  },
  "properties": {
    "convention": "CSB 1.0",
    "platform": {
        "type": "Ship",
        "name": "White Rose of Drachs",
        "ImoNumber": "1008140",
        "platformStatus": "new",
        "draft": {
                "value": 4.6[1] ,
        "uom": "m",
          "offsetApplied": false
        },
        "sensors": [
                {
                "type": "Sounder",
                "make": "Sperry Marine (L3 ELAC)",
                "model": "ES155100-02",
                "serialNumber": "136",

                "longitudinalOffsetFromGPStoSonar[2] ": {
                        "value": 3.52,
                        "uom": "m"
                },
                "lateralOffsetFromGPStoSonar[3] ": {
                        "value": -0.76,
                        "uom": "m"
                },
                "velocity": {
                        "value": 1500,
```

"uom": "m/s"
}
},
{
"type": "GPS",
"make": "Litton Marine Systems",
"model": "LMX420",
“offSetApplied”: “false”
}
]
},
"providerContact[1] Point": {
"orgName": "Sea-ID",
"hasEmail": "support@sea-id.org",
"orgUrl": "http://www.sea-id.org"[4] ,
“logger”: “MarkIII”,
“loggerVersion”: “1.0”
},
"processorContactPoint": {
"hasEmail": "support@sea-id.org"
},
"ownerContactPoint": {
"hasEmail": "support@sea-id.org"
},
"depthUnits": "meters",
"timeUnits": "ISO 8601"
},
"features": [
{
"type": "Feature",
"geometry": {
"type": "Point",
"coordinates": [
40.9148,
19.0052
]
},
"properties": {
"depth": 15.8[5] ,
"time": 2015-08-06T22:00:00Z[6]
}
}
]
}