Hydrographic processing considerations in the “Big Data” age: An overview of technology trends in ocean and coastal surveys

M Holland¹, A Hoggarth¹ and J Nicholson¹
¹CARIS, 115 Waggoners Lane, Fredericton, NB, Canada E3B 2L4

E-mail: matthew.holland@caris.com

Abstract. The quantity of information generated by survey sensors for ocean and coastal zone mapping has reached the “Big Data” age. This is influenced by the number of survey sensors available to conduct a survey, high data resolution, commercial availability, as well as an increased use of autonomous platforms. The number of users of sophisticated survey information is also growing with the increase in data volume. This is leading to a greater demand and broader use of the processed results, which includes marine archeology, disaster response, and many other applications. Data processing and exchange techniques are evolving to ensure this increased accuracy in acquired data meets the user demand, and leads to an improved understanding of the ocean environment. This includes the use of automated processing, models that maintain the best possible representation of varying resolution data to reduce duplication, as well as data plug-ins and interoperability standards. Through the adoption of interoperable standards, data can be exchanged between stakeholders and used many times in any GIS to support an even wider range of activities. The growing importance of Marine Spatial Data Infrastructure (MSDI) is also contributing to the increased access of marine information to support sustainable use of ocean and coastal environments. This paper offers an industry perspective on trends in hydrographic surveying and processing, and the increased use of marine spatial data.

1. Introduction
The quantity of information generated by survey sensors for ocean and coastal zone mapping has reached the “Big Data” age. Modern equipment used for seabed mapping activities may generate more than 100 GB of data per hour [1]; this high volume information needs to be adequately handled and analysed. The number of users of sophisticated survey information is also growing with the increase in data volume and resolution. This is leading to a greater demand and broader use of the processed results. Data processing and exchange techniques are also evolving to ensure that user demand can be met, and that an improved understanding of the ocean environment can be obtained. By applying both new technologies and new methodologies to acquire, process and discover Big Data, sustainable economic development and resource management for the marine environment can be achieved.

Economic development is one of many marine activities that are described by the International Hydrographic Organization (IHO) as being supported by hydrography, and studies show that an investment in hydrography by a nation provides a positive return [2]. Yet, what equipment and investments are needed to conduct accurate and efficient ocean and coastal mapping activities?
2. Sensors and hardware
High accuracy sensors that are available to modern day surveyors can be grouped into three categories: light, sound and location. Light Detection and Ranging (LiDAR) and laser scan technology have an established history in topographic mapping, but are also used in coastal and shallow water environments. This includes vessel-mounted laser scanners to survey coastline and port infrastructure, as well as airborne bathymetric LiDAR systems, which can achieve survey coverage for large areas in relatively short time frame. Acoustic multibeam and side scan sonars have been standard tools in hydrographic surveys for more than 25 years, and the achievable data resolution and quality of these sensors continues to improve with next generation systems. Synthetic Aperture Sonar (SAS) also provides an ultra-high resolution solution by synthesizing the size of a transducer array in software, which results in images with 3 cm by 3 cm pixels over the entire sonar range (figure 1). This type of coverage and image quality is applicable to security and defence, deep ocean searches (e.g., for downed aircraft), and more recently subsea pipeline inspection. In all cases the chosen sensor(s) need to be coupled with a positioning source to ensure the data is spatially referenced for integration with other data sources and ready for use in a Geographic Information System (GIS). Accurate positioning may be obtained using a combination of equipment, such as Global Navigation Satellite System (GNSS), Inertial Navigation System (INS) and Ultra-Short BaseLine (USBL), depending on the survey platform type (e.g., airborne, sea surface or subsea vehicle).

![Figure 1. SAS imagery (3 cm) of a sunken dory, with image data draped over co-registered bathymetry (25 cm) (Image credit: Kraken Sonar)](image)

Over the past 20 years, the marine industry has observed a reduction in price for some sophisticated sensors, while the cost of other sensors has remained approximately the same. Multibeam sonar equipment is one example where costs have been reduced by approximately 50 per cent for shallow water systems, but overall the costs for deep water systems have not decreased. This same trend can also be seen with bathymetric LiDAR, where less powerful and more portable units on the market are approximately 50 per cent of the cost of larger systems. However, the price for systems with greater achievable depth is still roughly the same as 20 years ago. This suggests that more innovation, and in turn cost savings, is occurring with smaller and more portable systems.

This trend is even more evident when looking at computing hardware. In 1995, a Toshiba laptop with 16MB RAM and 150MHz Intel Processor cost approximately $3,000. Whereas today you can get a Raspberry Pi 2 computer, which is the size of a credit card, with 1GB RAM and 900MHz Quad-Core CPU for about $50. The Raspberry Pi uses a low-powered ARM processor with a Linux-based operating system (figure 2), which is an attractive computing configuration for lowering power consumption and increasing endurance in autonomous survey systems.
3. Force multiplication

The number of sensors available to conduct surveys, the commercial availability of sensors, and the volume of data generated are all factors that have pushed ocean mapping into the Big Data age. Another important contributor in recent years is the increased use of autonomous platforms. These include Autonomous Underwater Vehicles (AUVs) and Autonomous Surface Vehicles (ASVs). AUVs, for example, offer a number of survey benefits, such as rapid deployment and recovery, and the ability to work close to an intended target (figure 3). Also, depending on the size and power endurance of the vehicle, it may be configured with multiple payloads to simultaneously acquire data from sonars, cameras, scanners, or other instruments.

Surveys in the marine environment have been traditionally carried out by staffed survey vessels and launches. Organizations can dramatically increase survey capacity and achieve force multiplication by combining traditional platforms with multiple autonomous vehicles configured with various high accuracy sensors. However, a current limitation of autonomous operations is that data is stored internally on the vehicle, and due to limited communications, is not accessible during operation. Full resolution data is only accessed for processing after the mission is complete and the AUV has been recovered. This creates a bottleneck in the survey workflow. The use of real-time processing and automated techniques will be essential for organizations conducting autonomous surveys to keep up with the volume of acquired data. By processing data on board the survey platform, and monitoring

![Figures 2 and 3](image-url)
the quality and coverage remotely during the mission, the time between data acquisition and the processed result can be drastically reduced. Results are available for quality control and decision support immediately following data recovery, which allows for the increased data volumes to be managed more efficiently. This approach also allows for multi-sensor data to be processed in a standardized way, and for available human resources to focus on other tasks during survey operations.

4. Modelling various data resolutions
Combining multiple sensors in mapping operations can provide efficiencies and allow complete coverage to be obtained with fewer survey passes (e.g., SAS combined with gap filling multibeam sonar for pipeline inspection). However, because sensor resolutions often differ, the modelling of data with varying resolutions can present challenges during processing. These same challenges also exist when modelling data over changing depth ranges or when combining topographic and bathymetric data.

Current workflows involve the creation of numerous single-resolution surfaces in order to model data with varying densities and to represent the sensor data at the best possible resolution. This leads to duplication of data, increased human intervention in the creation of surfaces and products (e.g., contours), and workflow inefficiencies. However, a Variable Resolution (VR) surface, which supports data of multiple resolutions in a single seamless surface, offers a more robust method to model an area with varying resolution data sources [3]. This method maintains the best possible representation of the processed data, and reduces duplication and disk storage requirements, which are fundamental requirements for efficiently handling large datasets. The VR surface can also be contoured in single pass so that the time-consuming process of stitching contours together is removed from the product creation workflow. The ability to maintain appropriate data resolutions in a single surface has applications for ocean mapping and modern navigation concepts, such as Under Keel Clearance (UKC) analysis, as well as the accurate modelling of land and sea data in a coastal zone (figure 4).

5. Coastal and infrastructure mapping
Hydrographers, land surveyors, GIS analysts and coastal engineers also encounter the variable resolution challenge when analyzing overlapping topographic and bathymetric data sources in a coastal zone. The issue is further complicated because of different Coordinate Reference Systems

![Figure 4. Variable Resolution (VR) surface (Image credit: CARIS)](image-url)
(CRS) for land and sea data. For example, the terrain models may be prepared by a national mapping agency and referenced to a vertical datum, such as Mean Sea Level (MSL); whereas the near shore bathymetry is provided by a hydrographic authority and is referenced to Lowest Astronomical Tide (LAT). It is important to accurately resolve the datum difference and shift one or both of the data sets to a common CRS. Once both datasets have a common CRS they can then be combined in a VR surface to support coastal planning, engineering and research (figure 5).

Figure 5. Seamless land and sea model with contours (Image credit: CARIS)

Ports, harbours and seaways are a crucial part of the coastal environment and national economy. By maintaining a sufficient channel depth, ports can attract cruise ships for tourism, and encourage container ship traffic. This maintained depth is important for a port to remain competitive with others in the region, as mere centimetres in the channel depth can significantly impact the amount of cargo that can be safely carried. Accurate models of the channel bottom can be used in UKC analysis to maximise ship draft and economic return. But models are also needed for more than just the seafloor in a port environment.

Quay walls and jetties also require regular inspection and maintenance. The inspection process itself can be time consuming and costly, and has traditionally been performed by divers or by Remotely Operated Vehicles (ROVs) equipped with cameras. This visual inspection process to identify problems and changes in vertical or near-vertical walls, is however subjective. It also does not allow for repeatable measurements or robust analysis.

The combination of laser scan and multibeam technology provides an alternative inspection method and allows detailed 3D point clouds to be generated for the structure both above and below the waterline. This approach enables the acquired data to be processed and visualized in a GIS, but few computations can be performed on point clouds to objectively detect changes. However, the creation of models for vertical and inclined features from the 3D point clouds provides decision makers with a
GIS product. The vertical feature models support repeatable computations and can be used for detailed deformation analysis to enhance infrastructure maintenance operations (figure 6).

6. Data standards and interoperability

The ongoing collection and processing of hydrographic data is essential to expanding our knowledge of the ocean environment. But to maximize the use of Big Data, the marine community needs to ensure that data is interoperable to support effective decision making. This can be achieved through the implementation and use of interoperability standards that allow various systems to be connected and for data to be available in an open and consistent way.

Examples of available technologies that support data access, sharing and system extension include the following: data plug-ins, such as the Geospatial and Point Data Abstraction Libraries (GDAL and PDAL), that can be embedded in any GIS for raster and point cloud format support; Open Geospatial Consortium (OGC) web services for data discovery; and Application Programming Interfaces (APIs) to extend capabilities of existing solutions. In addition to the maximized use of data, the adoption of interoperable standards offers lower implementation cost, when compared to proprietary standards, and leads to long-term operation and maintenance cost savings [4]. The IHO and International Association of Oil and Gas Producers (IOGP) are two organizations involved in the definition of data models and exchange formats used by the marine community for data delivery and sharing. Members of the IHO and IOGP, and their survey contractors, collect data and produce products that describe bathymetry, subsea infrastructure, seabed type, and other information to support safety of navigation or Exploration and Production (E&P) activities (figure 7). The data is used by each organization for a different purpose, yet much of the data is the same. However, by developing non-proprietary data exchange formats, the information is interoperable, and it can be used many times in any GIS. The IHO S-100 – Universal Hydrographic Data Model provides a framework for products, such as Electronic Navigational Charts (ENCs), Bathymetric Surface Products, and others to be defined for wider use of digital hydrographic data. The IOGP has also developed SeabedML, which is an open exchange format for the Seabed Survey Data Model (SSDM) used by the E&P industry for survey data. These standards provide interoperability between the two industries, as well as data exchange with other stakeholders for greater use.
7. Increase in utilization
Modern survey equipment is being effectively used to improve our understanding of the world’s oceans, and there are many applications for the data beyond maritime trade, E&P and other economic development activities. This includes the field of marine archaeology where sensors are employed to locate and investigate wreck sites, and other artifacts. The density of the data collected during the archaeological investigations also supports the creation of products, which can be used to conduct further research, to plan additional explorations, and to educate the public. The discovery of HMS Erebus, which was part Sir John Franklin’s 1845 expedition to find the North West Passage, is a recent example. The ship wreckage was mapped in great detail using high resolution multibeam and long-range laser scanner technology. Processed results were then exported to produce high precision models of the wreck and surrounding features using 3D printer technology. The physical 3D models allow the data to be presented in a way that historians and scientists can use, and the general public can appreciate as part of museum exhibitions and other events (figure 8).

The growing importance of Marine Spatial Data Infrastructure (MSDI) is also contributing to the increased access and use of bathymetry and other marine information sources. MSDI is a recognized...
way to facilitate data access and is based on a collection of policies and technologies [5]. While some nations still face policy challenges to establish MSDI, many of the technical issues have been solved through the use of interoperable standards. There are now many data sharing portals that provide access to hydrographic data. The use of OGC web services as part of MSDI allow agencies to establish data portals for viewing, searches, queries and download of data. All of this can lead to better management of coastal and offshore areas, as well as improved operations in case of disaster. The integration of spatial data with other information sources in a GIS can form a Common Operating Picture (COP). A COP offers a single platform and comprehensive tool to improve awareness and decision support for emergency responders. The COP for Oil Spill Response is a real world example of an MSDI that demonstrates how hydrographic information and open data standards (e.g., defined by IHO, IOGP, OGC) can be implemented and used to improve the response to an oil spill incident. This includes the ability to minimize impact and restoration costs, and protect the ocean environment [6].

8. Conclusion
The increased number of high resolution sensors, the growing volumes of processed results, and a survey force that has been multiplied through the use of autonomous survey platforms are all factors contributing to the Big Data age for ocean mapping. Real-time processing will be essential to keep pace with the amounts of acquired data, and the efficient creation of detailed models is needed to support in depth analysis for maritime trade, coastal engineering, and resource management. The greater detail and volume of data from modern surveys is helping to better the understanding of our oceans, but scientists, engineers, emergency responders and other stakeholders require access to the marine information. Through the adoption of interoperable standards data duplication can be reduced, information can be reused, economies can be developed, history can be discovered, and Big Data can be leveraged for sustainable use of the ocean and coastal environments.

9. References
[1] Shea, D 2015 Latest developments in seabed mapping using commercial Interferometric Synthetic Aperture Sonar (INSAS) Presentation at the US Hydro 2015 Conference, 38 p http://www.hypack.com/ushydro/2015/papers/pdf/US%20HYDRO%202015%20- %20Kraken%20Sonar%20Inc%20r02.pdf
[2] International Hydrographic Organization 2013 World hydrography day underpinning the blue economy - http://www.iho.int/mtg_docs/WHD/2013/WHDBulletin_Final.pdf
[3] Masry, M and Lamey B 2015 Variable resolution surfaces in practice Proceedings of Shallow Survey 2015 Conference
[4] Open Geospatial Consortium, The International Organization for Standardization, Technical Committee 211, Geographic information and International Hydrographic Organization 2014 A guide to the role of standards in geospatial information management
[5] International Hydrographic Organization 2011 Spatial Data Infrastructures- The Marine Dimension Publication C-17 – Edition 1.1.0 2011 https://www.iho.int/iho_pubs/CB/C-17_e1.1.0_2011_EN.pdf
[6] OGC-IPIECA 2015 Recommended practice for common operating picture architecture for oil spill response – Final Report, 103 p https://portal.opengeospatial.org/files/?artifact_id=6333