Needle in a haystack

R.H. Righolt, J. Schaap, L.L. Dorst, E.M. Vos
Hydrographic Service of the Royal Netherlands Navy

Abstract

The Hydrographic Service of the Royal Netherlands Navy (NLHO) is responsible for charting the Dutch sea areas in Europe and the Caribbean. Products of the NLHO include official Electronic Navigational Charts (ENC's), Paper Charts and other Nautical Publications.
To meet quality requirements for both military and civilian purposes, the Royal Netherlands Navy deploys two survey vessels. HNLMS Snellius (2003) and HNLMS Luymes (2004). Both vessels are equipped with state of the art hydrographic survey systems. Through various causes, large multibeam data files may contain erroneous points.

In this paper we focus on the need to filter the 'needle' from the 'haystack' of multibeam data and how to ensure that only validated data is used in this process. Automated processes should be used where possible to reduce operator workload and maintain consistency. To that end both algorithms and visualization techniques should be improved to deal with noisy point clouds.
It is anticipated that developments in water column imaging, backscatter based bottom classification and artificial aperture sonar data processing will further increase the quantity of data collected. This will require more data storage, more processing power and improved visualization techniques to cope with this flood of data.
In this complex environment with many innovations in the fields of survey, production technology, policy and management it is important for the NLHO to realize its main objective, 'to serve the mariner: Ex Usu Nautae'.

1. Introduction

On behalf of the Netherlands Government, the Hydrographic Service of the Royal Netherlands Navy is the official producer of electronic and paper navigational charts and nautical publications. The products cover the North Sea, the Dutch Antilles, Aruba, Surinam and adjacent ports.
Primary objective is safe navigation for SOLAS-shipping (SOLAS stands for the treaty on Safety of Live at Sea). The products are composed in accordance with rules laid down within the International Maritime Organization (IMO) and International Hydrographic Organization (IHO).

In order to obtain the necessary information, two hydrographic vessels are continuously conducting surveys to register and record sea bottom changes. Not only changes in depth are registered but also obstructions and wrecks. Also data is exchanged with the Directory of Transport and Public works (Rijkswaterstaat), who mainly survey the inland waterways of the Netherlands and hold responsibility for buoyage.
At the shore based Hydrographic Office the collected data together with other relevant information such as buoys, beacons, lights etc. is stored in databases. The content of the databases is the foundation for all nautical products.
In addition to producing charts and nautical publications, the NLHO makes its nautical knowledge and extensive experience available to support Naval operations in the fields of hydrography, oceanography, meteorology, positioning, tides and marine geodesy.

The aim of this paper is to give the reader an insight into the problems that are encountered in the processing of large amounts of data. Hereto we will give a brief introduction on the NLHO and its data now and in the past (section 2). We will describe the validation techniques in use (section 3), possible ways to overcome the various artifacts that may appear in multibeam soundings (section 4) and (future) ways to identify or diminish these errors (section 5). In relation to finding the needle in the big haystack we discuss the current practice at the NLHO and the role that academia and industry could play (section 6).

2. Historic and current use of data

Starting after the fairsheet era, digital storage media of various natures like tape cartridges, floppy disks CDs and the like were used. Although at the time we believed that we were dealing with large quantities of data, present tools have no difficulty dealing with it.
Since the introduction of the MultiBeam Echo Sounder (MBES) the quantity of raw data has increased significantly. In the beginning automatic pruning processes were implemented but it was found that the algorithms were too basic and important data could be lost. Now all data is labeled, evaluated and verified on board then zipped & stored on a portable hard disk for transport to the Hydrographic Office.

Developments in hardware have significantly improved storage devices. Currently all raw data from a survey can be stored on a single Terabyte sized portable hard disk. With two Hydrographic Survey Vessels active in the North Sea, some 35 of these portable hard disks are received by the Hydrographic Office each year. Subsequently all data is binned (5 m x 3 m bin size) and copied to safe storage.

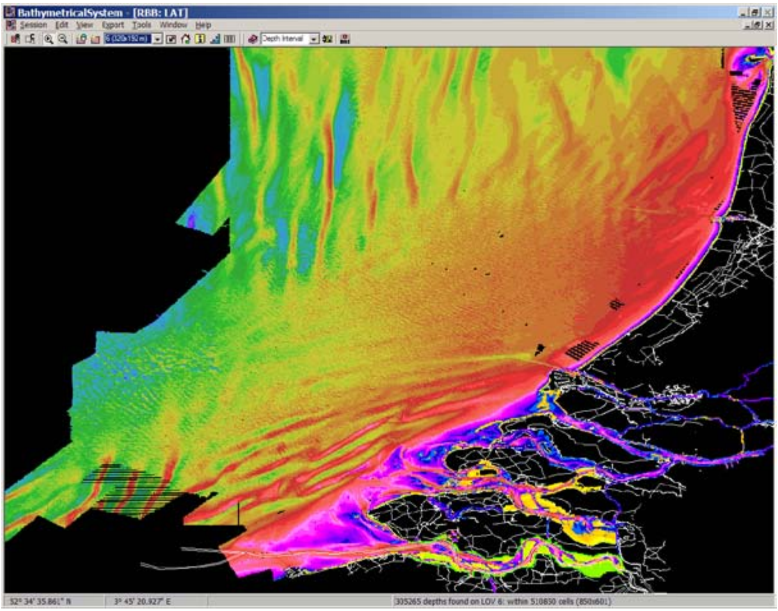

Figure 1. Screen dump of Bathymetric Archive System (BASRBB) on Level of Visualisation (LOV) 6 showing the Southern North Sea, coloured by depth. The area is characterized by sand waves and mega ripples. To better understand the dynamic behaviour of these sand waves studies are underway by various groups at Dutch universities and government institutions.

In the nineties the NLHO developed a software package (Bathymetric Archive System – BAS) that can evaluate, store and retrieve all binned survey data along with its associated meta-data. All data is binned to a worldwide grid, reducing the size of the data files (the cell size is 5 m in latitude by 5 m in longitude at the equator, which equates to 5 m times 3 m at our latitudes.).
A number of tools allow the user to query the database and export selected data in a variety of formats and densities. Due to restraints of the system, we currently only store the Minimum Depth per bin. Other values that describe the quality of a sounding like the Total Propagated Uncertainty (TPU), Standard Deviation (SD), Hits per Bin, Mean & Maximum Depth are used in the evaluation process and subsequently discarded.

Survey data is also received from Rijkswaterstaat (RWS). Their operations are focused on the inland waterways, river deltas and offshore areas out to the 10 m depth contour. Their data is binned to comply with our formats and stored in the BAS database. At present only Minimum Depth and associated meta-data is received. However since the acquisition systems of RWS and the RN navy are essentially the same, it would be possible to devise a storage format that retains all relevant data in a similar fashion.
Within the NLHO these BAS exports are subsequently used by the data processing department to compose and update the bathymetry for the topographical database (TLDB). For the production of Paper Charts we utilize the CARIS Editor, a hydrographic software package that enables us to draw contour lines and select soundings. This package is also used to maintain the TLDB.

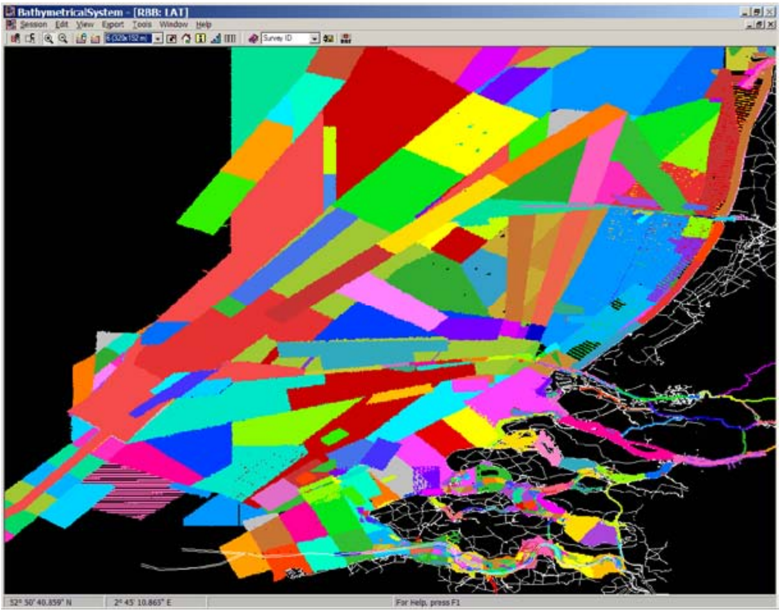

Figure 2. Screen dump of BASRBB LOV6 showing the Southern North Sea, colored by survey. More than 500 surveys are processed by the NLHO each year.

3. General aspects of data validation

The raw data is validated, evaluated and stored. The validation process comprises a number of steps. Its prime purpose is to ensure that the data complies with the accuracy requirements as defined by the IHO Special Publication Nr. 44.
The first step is a comparison with previous surveys of that particular area. In case of discrepancies the cause must be identified. The North Sea bottom is characterized by dynamic bedforms of different spatial scales, each changing at different temporal scales. These changes in the morphology can be explained and are not due to measurement errors. Some other changes may indicate that shoaling has taken place and a dangerous situation has arisen for the mariner. In that case a follow up with a Notice to Mariners (NtM) may be called for.

Secondly, we use the IVS Fledermaus 3D visualization package to inspect and identify artifacts in the point cloud. These artifacts are either introduced by the Multibeam Echo Sounder (MBES) itself or they could be real artificial or natural physical features. In order to be able to recognize these artifacts for what they are, the spatial resolution must be adequate for the intended purpose. A bin size of 1m or smaller is best to achieve this.
Since modern Hydrographic Survey Vessels (HOV) routinely cover large areas, the resulting data files are substantial. A typical survey area of a few square km may contain in excess of 100 million data points. Considering that for every sounding a position, a depth value and some statistically relevant information about the quality of the point must be logged (TPU), the size of

a data files quickly grows to many Gigabytes. Visualizing this survey data as a single image is desirable. It is only when the whole survey is visible, that some of the artifacts become clear. It is then easy to zoom in and pan to an area of interest. For a detailed inspection of the sea bottom full resolution must be used. The IHO S44 rules prescribe that objects of certain size are found and - if possible - identified. At present there is no alternative but to zoom in and pan around the whole work area. The operator must make a judgment as to the nature of the observed features and mark those that warrant further investigation. This is a time consuming, but necessary task!

Manipulating these data files starts to become a problem when time series of the same area are available and need to be retrieved for inspection. Since all data resides on a dedicated office server and are accessed via a local network, this takes some time.
A good visualization allows the analyst to asses the quality of the dataset. Some software will allow the user to toggle between the rendering of the point cloud and the rendering of a Triangulated Irregular Network (TIN) or Digital Elevation Model (DEM) generated from it.
The NLHO utilizes the IVS Fledermaus suite of software. This provides a set of interactive 3D visualization tools for analysis and presentation. It allows us to visually identify and delineate morphologically distinct areas as well as wrecks, pipelines, obstructions and objects. The current version of this software package theoretically supports visualization of up to 100 million points in a rectangular area. Since many survey areas have odd shapes, this means in practice that the point cloud can have far fewer points. Unfortunately the larger surveys need to be split into a number of parts that can be evaluated individually.

The Fledermaus viewer enables the user to rotate the picture and change the viewpoint. It allows the operator to inspect an object from all angles and with a different color palette. Even when viewed at full density and at the perfect viewing angle, the identification of many objects still presents a challenge. Interpretation is highly subjective and depends on operator experience & acquired skills.
For charting purposes it is important to obtain an accurate figure for the safe clearance depth over an object. The nature of the MBES is such that the bottom-track features of the processing software may preclude the proper detection of a protruding object. It is interpreted as a spike in the system because the surrounding data does not confirm it and is edited out. Also the applied weighting parameter should be set to a low value because there maybe very few hits on the shallowest point and averaging can introduce errors. Developments in the area of water-column imaging may alleviate some of these problems (See section 5 Future Developments).

4. Detection of artifacts

Artifacts that were introduced as a result of the MBES and ancillary equipment must be quantified whenever possible. Some of these errors point to a problem in the MBES that can be resolved or minimized by recalibration and careful adjustments of the MBES operating parameters. Many classes of problems cause their own recognizable artifacts. A few are discussed here.
Tidal correction problems manifest themselves as ‘swaths’ that do not match with neighboring tracks. Reprocessing all data with better refined tidal information may be required or in the worst case the survey must be rejected.

Another significant source of errors is the use of an incorrect sound velocity profile. With the use of a Moving Vessel Profiler (MVP) a sound velocity profile is obtained that is subsequently applied to all observations. However due to infrequent sampling and local variations in temperature or salinity, these errors can increase rapidly. Smile and frown artifacts will be the result. With the benefit of hindsight, small improvements can be made by applying the obtained velocity profile retroactively during post processing. Often residual artifacts remain. A software solution that tries to minimize these ‘smilies’ by estimating the best fitting velocity of propagation is

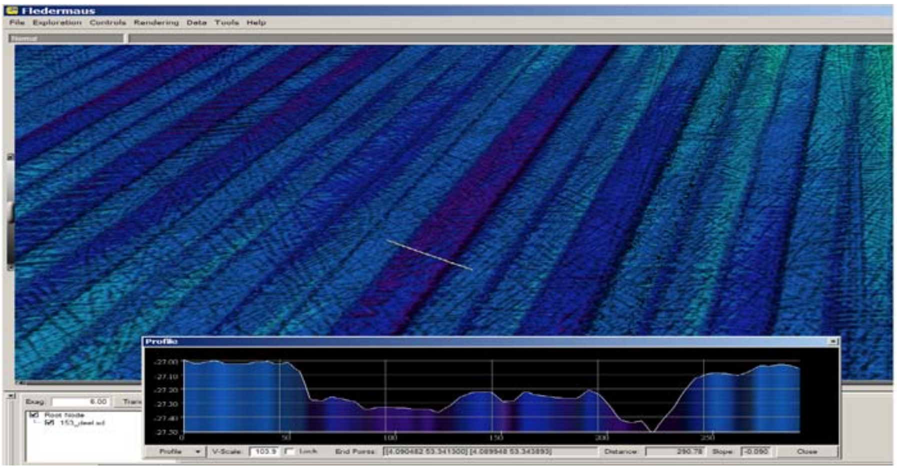

Figure 3. Example of tidal correction error artifacts.

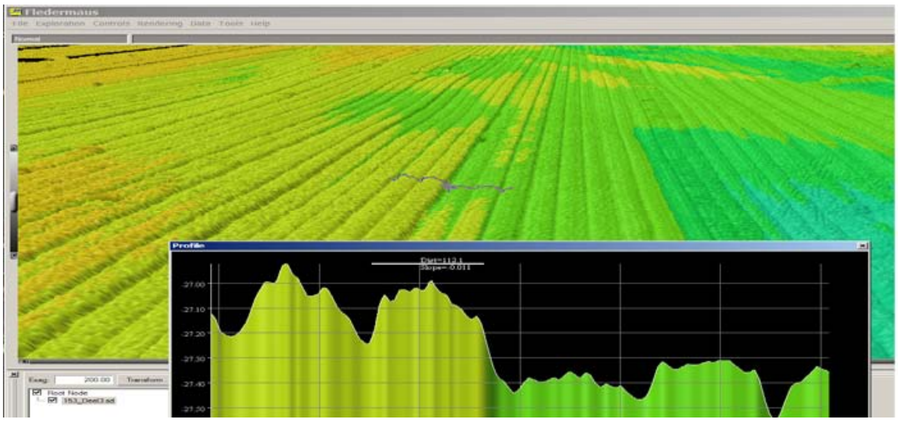

Figure 4. Example of velocity profile error artifacts (smile/frown).

on the horizon but much more work needs to be done to prove its merits before it can be accepted by in the survey industry as a valid tool.

Poor weather conditions may cause heave and roll artifacts to be introduced. The surveyor in charge must decide if the results are still within tolerance. Since many of these errors can be recognized as an artifact, it should be possible to identify the source and subsequently make a correction for it. Software algorithms that operate on the point cloud rather than the swath data can be developed to assist in this process.

Again the IHO S44 document provides guidelines and limits. It is therefore necessary to put numbers to the observed artifacts. Summaries with explanations of error diagrams and examples are available, but often these errors mix and it is not always possible to recognize these artifacts for what they are. It is our wish and hope that some time in the future quality control software will become available that can automatically identify and quantify many of these errors (Hughes Clark, 2003).

Various artifacts that are due to calibration errors or hardware faults may also show up and require attention. Some of these errors can be corrected during post-processing, others must be spotted in the acquisition phase and remedied before the survey commences, such as the misalignment of dual heads of a MBES (Figure 6).

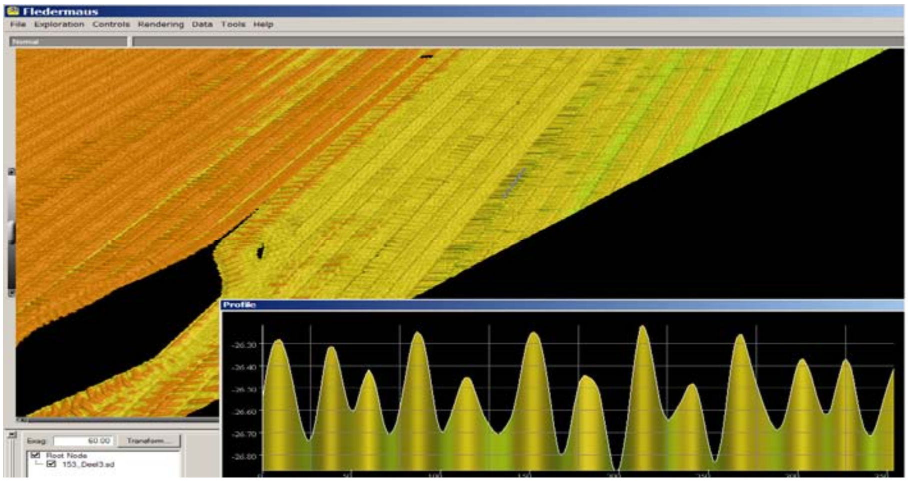

Figure 5. Example of heave artifacts.

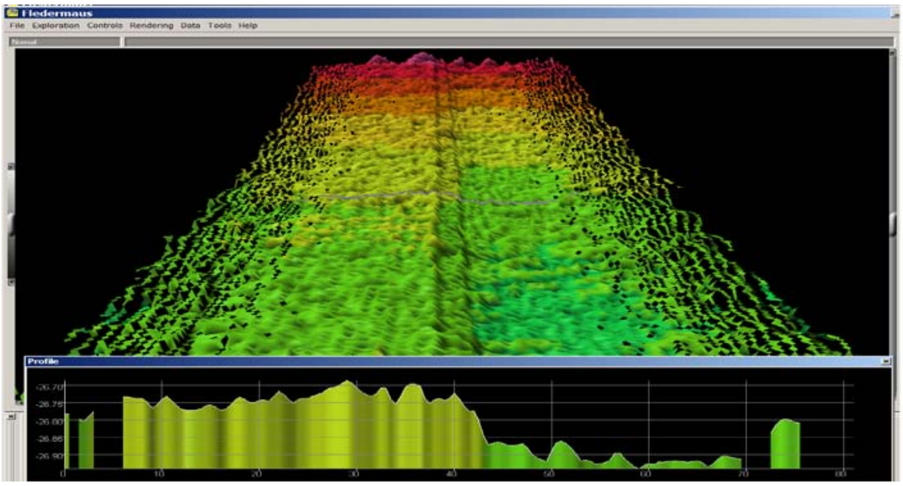

Figure 6. Example of dual-head transducer misalignment error artefact.

5. Future developments

Considering that the current hard- and software can barely meet our need for visualization and manipulation, we have to prepare ourselves for future developments. It is expected that the next generation of acquisition systems will generate an even denser point cloud.

To name just a few systems that are close to reaching acceptance in the survey industry the following developments can be considered.

- The use of multiple Autonomous Underwater Vehicles (AUV) per survey vessel can improve the efficiency of the survey operation with existing technology. No new MBES technology needs to be developed and tested (Kongsberg Huggin – Remus).
- Interferometric sonar systems are close to meeting nautical charting requirements. Advances in electronics and phase deconvolution algorithms promises high-resolution wide-swath bathymetry. Initial results from observations by NOAA indicate that these tools are capable of improving survey efficiency. A swath width of 10 – 15 times water depth seems achievable (Gostnell, 2006).
- Water column imaging is already with us. The data is acquired with the existing MBES equipment, but processed differently. The software to visualize the results still need to be further optimized. Although IVS has published results of their new mid-water visualization tool,

it needs to be proven that it can reliably detect the least-depth of a wreck or object, in which case the need to use valuable boat time for bar-sweep operations is diminished. Possibly algorithms and migration routines may be borrowed from the seismic industry, where 3D-binning and stacking are routine operations. The treatment of data in the seismic industry and water column imaging seem very similar, only the voxel (Volumetric pixel) size is an order of magnitude smaller (Hughes Clark, 2006).
- The Hydrographic Service of the Royal Netherlands Navy is developing a method to extract information about seabed dynamics from bathymetric archives. Software that can make such an intelligent estimate of Seabed dynamics can lead to a review of resurvey frequency. Boat time can be allocated based on need, rather than traffic intensity or perceived dangers. Again such software operates on large quantities of data and hence requires powerful computer hardware (Dorst, 2009).

A treatise on the merits of various visualization techniques and the terminology involved can be found on the internet. A good visualization not only allows the analyst to quickly assess the quality of a dataset, but also enables the planning and control of different processing schemes and ultimately will provide the presentation of the final product.

6. Finding that needle in a haystack

Many survey vessels are active in North Sea and Dutch river estuaries, thus acquiring large quantities of soundings. The Hydrographic Office receives validated soundings from their own surveys vessels and 3rd parties. Of all these millions of soundings only a few thousand can ultimately appear on the charts, making the selection process a critical one.

An automated process that has knowledge of relevant morphological and topographical features can aid this selection process. In this process a massive point cloud need to be searched to identify a small group of soundings that accurately represents the large group whilst emphasizing the nautical hazards. Reduction of the point cloud is currently achieved by a binning process that is implemented as part of the BASRBB software package. Depending on the desired chart compilation scale, this package can further reduce the number of soundings by exporting from a different Level of Visualization (LOV) layer. Every level up will reduce the number of soundings by three quarters. The actual selection is then based solely on minimum depth of the four neighboring bins.

Considering that there are many rules that govern the way in which a sounding is posted there is little room for maneuver once a selection is made (posting a sounding is the process of assigning a sounding to a particular area or feature). Just by inserting an extra sounding one of these rules is easily violated and more corrective action needs to be taken. Our experience with the current generation algorithms is that it is often better and faster to do the sounding selection by hand. It would save time if better algorithms were developed which can take many of these rules into account. Such an algorithm needs to have knowledge of the generalized contour lines, the location of buoys and wrecks, morphological features and a host of other details that determine the posting of a sounding.

Examples of potential hazards are shoals, lost containers, wellheads or wrecks. The mariner needs to be warned about newly identified hazards by issuing a Notice to Mariners (NtM). Note that worldwide up to 10,000 containers are lost at sea in a single year! Due to shifting sands, wrecks can move and tumble over time and thereby possibly changing the minimum safe depth. Whenever a wreck is identified in the point cloud, the minimum depth clearance should be established and deviations from the published values should be investigated further.

We would recommend improving the current practice in the following ways.

Analyzing the remainder of the haystack
To analyze the remainder of the haystack it is essential to be able to distinguish between artifacts and real world features. This pleads for the storage of detailed quality attributes throughout the processing chain.

Error analysis on different levels
Error analysis on different levels of aggregation improves the traceability of both artifacts and real world features hidden in the data. Storage of detailed quality attributes does increase the size of files however.

Visualization
The ability to recognize artifacts depends to a large extent on the kind and quality of the visualization. It is here that improvements can still be made. Also there are algorithms that can remove or mask certain artifacts but by their nature introduce uncertainties with regard to accuracy. In that case it cannot be used for charting purposes, but can still be very helpful to visually identify objects or features that would otherwise remain hidden. For instance, severe heave or roll artifacts can easily obscure features like small wrecks, containers, pipelines, wellheads or mines. The small object detectability as is also described in the IHO S-44 document is thus improved. More investigations should be done to determine appropriate ways for visualizing massive point clouds that represent a marine environment.

In order to reach these goals the knowledge of academia and industry is needed.

7. Conclusion

The existing hard- and software is barely adequate to work with current quantities of data, but new acquisition and processing systems are on the horizon. The expectation is that these new systems will generate ever larger quantities of data that need to be manipulated and evaluated. To that end both algorithms and visualization techniques should be improved to deal with these noisy massive point clouds in an automated way. In this complex environment with many innovations in the fields of computer hardware and visualization techniques, the NLHO is dedicated to be an active player in the field and continues to 'serve the mariner': Ex Usu Nautae!

8. References

Brooke Ocean Technology Ltd. Moving Vessel Profiler, MVP-100, www.brook-ocean.com.

CARIS Editor. http://www.caris.com.

Dorst, Leendert L. (2009). Estimating Seafloor Dynamics in the Southern North Sea to improve bathymetric survey planning.

Gostnell, Caleb and Yoos, Jake and Brodet, Steve (2006), Test and Evaluation of Interferometric Sonar Technology. http://www.noaa.gov.

Hughes Clark, John E. (2003). Dynamic Motion Residuals in Swath Sonar Data: Ironing out the Creases, International Hydrographic Review.

Hughes Clark, John E (2006). Multibeam Water Column Imaging: Improved Wreck Least-Depth Determination.

International Hydrographic Organization (2009). IHO Standards for Hydrographic Surveys, Special Publication Nr. 44. http://www.iho.int.

Kongsberg Maritime AS, EM3002D Multibeam. http://www.km.kongsberg.com.

Sang Yun Lee, Kwang-Wu Lee, Ulrich Neuman. Interactive Visualization of Oil Reservoir Data. http://www.usc.edu.

United Kingdom Hydrographic Office UKHO Training Manual, Sounding selection.