

SEAMLESS HANDLING OF LARGE MULTIBEAM DATASETS

Seamless Multibeam Data Analysis and Management

How to tackle problems traditionally associated with the management of multibeam data? A management tool has been developed which offers concrete operative solutions by eliminating lengthy manual editing stages and considering the entirety of survey swaths as a single dataset whilst maintaining control of each single sounding.

The revolution created by the introduction of remote sensing in ways of perceiving, studying, interpreting, representing and operating in a land environment has been mirrored in the marine, river and lake environment by the introduction of multibeam soundings. There is an important difference between the two technologies: systematic checking on the grounds of multibeam data is impossible. Both technologies come up, however, against exactly the same type of problems - signals crossing an anisotropic physical medium, the exploration of increasingly large areas in an increasingly extreme degree of detail, huge amounts of data to be analysed and managed and acquisition instruments produced by different manufacturers.

The current technology that has been used, is based on a specific conceptual model and provides software instruments which can homogenise/structure the sources of information available and analyse/ manage huge amounts of data without a continuity solution making their breakdown (into swaths, cells, sub-areas, etc.) useless and even damaging. Such an approach is useful in support of the techniques which first pinpoint anomalous data and then allow for the separation of "noise" (spikes) from the information. The surveyor perceives the dataset or datasets in his possession as a whole and the instruments at his disposal allow him either to carry out in-depth analyses beam by beam or to evaluate the overall quality of the survey.

Data Problems and Characteristics

Acquisition procedures provide a huge quantity of measurements (soundings) divided into swaths, which in turn are broken down into fans and beams. This breakdown is due to the nature of the acquisition process and has nothing to do with the "reality to be modelled". It must be noted that the swaths may partially overlap, intersect each other or become denser in particular zones. Furthermore, there is often the need for a comparison between surveys carried out at different times in adjoining and/or partially or completely overlapping geographical areas.

Conceptual Model and Computerised Geometry

The philosophy behind the "Multibeam Manager" software technology, developed by an Italian company, aims at providing instruments for data analysis and interpretation regardless of the means of acquisition or the quantity of data surveyed. There is no formulation of an "a priori" hypothesis to describe either the characteristics of the data or the morphological features of the area under study.

The entire multibeam dataset is seen as a whole and is represented (using probability theory techniques) with uncertain variables which depend on geographic parameters.

In fact, the probability theory provides a corpus of axioms and theorems with which correctly to "model the reality", even in the presence of huge amounts of data which may be locally incorrect, unreliable or even completely absent.

Effective implementation of the mathematical model of the data has necessitated, in addition, the development of innovative computerised geometry techniques in order to tackle the difficult problem of organisation of and optimum access to huge amounts of georeferenced data.

It has therefore become possible to:

- Overcome the restraints imposed by the initial structuring into swaths, fans and beams
- Achieve a unitary and homogenous treatment of the entire dataset
- Avoid unnatural fragmentation of geographic data into sub-areas, with the introduction of artificial sub-divisions producing troublesome "junction and/or connection effects" between different parts and leading to loss of overall "view" of the dataset

Mathematical Basis

The approach adopted has required the use of sophisticated mathematical techniques so as to ensure that the algorithms have a simultaneous high degree of reliability, flexibility and power. Particular reference is here made to "Analysis of Conditioning" and "Hilbert Spaces".

The concept of conditioning was introduced by the British mathematician Alan Turing to evaluate and check the stability properties of an

algorithm. It has been widely used in all fields of numerical analysis, both in theoretical questions and in practical applications, from the theory of linear systems to the solution of differential equations. In the field of "Multibeam Manager"™, all the algorithms used operate with a careful analysis/check of their conditioning properties in order to guarantee a maximum degree of reliability/quality, particularly in the field of modelling techniques.

Hilbert's Spaces, dating from the beginning of the 20th century, are a natural generalisation of the plane and of ordinary space. They have become a fundamental instrument in the study of all those statistical methods which fall under the general name of "Variance analysis"™, "Least-squares Method"™ and similar, as well as for the description of refined stochastic processes such as Brownian motion (random motion of a particle subject to collision), Kalman-Bucy filters, etc. In the multibeam data managing software technology they have been used together with statistic-type analysis with a number of parameters not "a priori"™ known.

Operational Procedures

Despite the complexity of the elements involved, the software technology is flexible and easy to use.

It foresees the following steps bench-marked by the survey department of the Italian Hydrographic Office, using a set of data taken in the area of Monte di Portofino (Genova).

1. Homogenisation of data. This consists of the extraction of data from the original manufacturer, the contextual application of any necessary corrections (tide, sound velocity profile) and the automatic creation of metadata (detailed statistical reports and geographic index of the survey showing both the occupation of each single swath and the presence of any areas not covered). This stage is essentially "batch"™ and requires minimum human intervention.
2. Creation of EPSILON model. This consists of the reorganisation of the entire dataset and produces a first comprehensive model of the sounding, giving a geographic-statistical synthesis of the information contained in the dataset. This stage is batch too and requires minimum human intervention.
3. Data analysis. The user is given integrated instruments for analysing and classifying data, allowing him to analyse from the statistical point of view and geographically represent: the level of confidence (in accordance with specification S-44), overlapping of swaths, density of the dataset, the local spread of data and the automatic identification and classification of each single item of "anomalous data"™ present in the dataset. Finally, data analysis allows for in-depth sessions of object detection. This stage foresees direct interaction between the user and the system
4. Creation of GAMMA model. This produces an overall, smooth three-dimensional model of the sounding. A high-quality, "continuous"™ representation is obtained - free of the interpolation errors which normally affect models based on TIN. The GAMMA model is stable with regard to noise present in the data and the representation is highly reliable. GAMMA modelling is completely automatic (no form of manual editing is required) once the user has selected the anomalous data identified from one of the analyses carried out at the preceding stage
5. Creation of GAMMA PRODUCTS. From a "GAMMA model"™ the user can obtain, automatically and parametrically: bathymetry, fair sheet, morphological analyses (maximums, minima and saddles) and can also carry out hydrographic soundings selection

Performance Analysis

The following summary report shows the results obtained by installing "Multibeam Manager"™ in a PC costing a few thousand Euros (processor P4 2,0 GHz, 1Gb of RAM, 160 Gb HD).

It allows to change the hydrographic rule of thumb which states that: "for every day dedicated to multibeam data acquisition you need approximately one day to process the relative data"™.

Survey summary report

- Survey Area: 21.5 x 15.8km
- Overall Survey time: 11 days (launch), + 2 day (ship)
- Surveyed soundings: 132,000,000
- Depth range: 0 to 120m
- Mbeam Echosounder: Simrad EM 300 (ship) and EM 3,000 (launch)

Preliminary activities

- Homogenisation of data
- Creation of EPSILON model

Evaluation (including processing of data and evaluation of the results):

- Level of confidence
- Spread of data
- Anomalous data analysis

Products

- Creation of GAMMA model
- Sounding selection
- Creation of bathymetry
- Fusion with Land detailed topography data
- Creation of the Nautical Chart (scale 1:10,000) according to IHB standards

Overall time required: 24 hours

The already cited one day survey of Portofino (32,000,000 of soundings) requires four hours to get the same results.

Conclusions

It is worth remembering that all that is needed to install and use the software technology “Multibeam Manager”™ is a PC with enough disk space for the volume of the dataset to be processed.

<https://www.hydro-international.com/content/article/seamless-multibeam-data-analysis-and-management>
