Automatic calculation of bathymetry for coastal hydrodynamic models

P. Bailly Du Bois

Automatic calculation of bathymetry for coastal hydrodynamic models

BAILLY DU BOIS P.

IRSN/DEI/SECRE/LRC - Institut de Radioprotection et de Sûreté Nucléaire, Direction de l'Environnement et de l'Intervention, Laboratoire de Radioécologie de Cherbourg - Octeville, rue Max Pol Fouchet, B.P. 10, 50130 Octeville – France.

Correspondence:

Pascal Bailly du Bois

Laboratoire de Radioécologie de Cherbourg - Octeville (LRC)

IRSN-LRC

rue Max Pol Fouchet

B.P. 10

50130 CHERBOURG-OCTEVILLE

Tel. direct line: (33) 02 33 01 41 05

Tel. secretary: (33) 02 33 01 41 00

FAX: (33) 02 33 01 41 30

Email: pascal.bailly-du-bois@irsn.fr

Abstract

A methodology is proposed to facilitate the construction of gridded bathymetry data for the use of hydrodynamic models on the continental shelf.

It relies on the carrying out of three successive tasks:

– Automatic selection of records of better quality among multiple sets of overlapping data;

– Elimination of data points located on land;

– Taking into account of the shoreline as the terrestrial limit.

Algorithms are proposed to perform sorting of the records as well as masking by the coastline.

The suggested method facilitates the updating of bathymetry data and optimizes their use. It enables automatic execution of all the tasks and building of new digital bathymetry models in a few hours, without action from the operator.

The method has been tested many times on the continental shelf of North-West Europe. The bathymetric data so generated do not require corrections, and the hydrodynamic models on which these data are based could be very precise.

Keywords: modelling, bathymetry, continental shelf, sounding, quality

1. INTRODUCTION

Great progress has been made in the hydrodynamic modelling of water-mass movements, and advances in computing have enabled the simulation of short- or long-term physical transport and dispersion of waters on the scale of continental shelves. These models make it possible to predict the consequences of man-made pollutant discharges into the marine environment under realistic conditions of release, weather or tide. They are increasingly vital tools for marine ecosystem management, for both managed and accidental releases.

Computer programs have reached a sufficient level of sophistication so they can be implemented in any zone without modification of the program, requiring only the adjustment of some parameters (the time step used in calculations, for example). The demand for the development of new models is continually growing, along with the aim of adapting as well as possible the precision and spatial coverage of studies to the available computing resources. Tools such ROMSTOOLs can provide oceanic models with the initial data sets necessary to carry out calculations, in particular the bottom topography (Penven et al. 2007). For coastal models, whose calculation grid-cell size is less than 1 km, one of the main obstacles to the development of new models remains the construction of digital representations of bottom topography which are suitable for new coverages and resolutions.

For each new model it is necessary:

− To collect sufficiently precise bathymetric data, over the whole domain of the simulation;

− To adjust the vertical reference of the data points so that it corresponds to that used in the models;

− Sorting and selecting the best data when dealing with multiple sources;

− To deal with overlaps between data sets/series;

− To interpolate individual data points to calculate the depths in each calculation grid of the model;

− To check the digital bathymetry model so that there are no aberrant data at sea or on land

− To make sure that the shoreline is well accounted with realistic slope between soundings and the coastline. Jones and Davies (2005) pointed that "the exact choice of water depth at the coast and in the near coastal area are crucial in determining the model's accuracy in the coastal region"

Carrying out all these tasks can involve a considerable investment in human resources, with manual checking lasting several days or weeks.

Being regularly confronted by such obstacles, we developed a methodology along with algorithms and a computer program allowing us to carry out the tasks automatically and perform an automated and rapid setting up of new models.

Some of the stages presented here are described elsewhere and are not developed further, such as, for example, the techniques of interpolation of soundings. The aim of this study is to present a tested methodology, based on the computerized processing of bathymetric data, for calculating digital bathymetry models that can be immediately utilized by hydrodynamic models.

The methodology and computer programs presented here were implemented for models with regular grids having a constant step in x and y coordinates. However, since the algorithms are generic, they can be adapted to models with irregular grids by modifying the grid-cell size used. Gorman et al (2006) proposed a method for optimisation of bathymetry and representation of shoreline more adapted for unstructured mesh models.

The study presented here is organized as follows:

− Correction of the collected bathymetric data, in particular with reference to mean sea level;

− Selection of higher-quality records according to their origin or acquisition methods;

− Elimination of onland records/data points;

− Taking into account of the shoreline to represent the land/sea interface;

− Interpolation of individual records within the coverage of the models;

− Checking;

− Discussion;

− Conclusion.

2. CORRECTION OF BATHYMETRIC DATA

The precision of a hydrodynamic model depends directly on the precision and distribution of the bathymetric measurements that are used to build it. According to the area and coverage of the model, several data sets can be used, each having different characteristics in terms of vertical and horizontal precision.

In view of the numerous sources of bathymetry data, their collection is not discussed further here. Once the data is gathered, we need to check the correspondence between the geographical reference frame of the models and the datum used for the bathymetry data (generally, WGS84). In the case of tidal seas, the correction for the tides should be validated.

2.1. CALCULATION OF MEAN SEA LEVEL

The models require soundings that are referenced to mean sea level, whereas the collected data, generally used for the drawing up of nautical charts, are referenced to the Lowest level of the Astronomical Tide – LAT - (chart datum convention of the Hydrographic and Oceanographic Department of the French Navy - SHOM - adopted by the Member States of the International Hydrographic Organization – IHO 1996).

The bathymetric data measured with respect to the chart datum must then be corrected for the maximum half-amplitude of the tide to obtain the water depths with respect to mean sea level. This correction requires a knowledge of the maximum amplitude of the tide at all points of the coverage of the models. We can use several sources of tide-gauge measurements and/or tidal components:

For French ports, the values of the maximum amplitudes of the Highest Astronomical Tide are available on the Internet site of the SHOM (http://www.SHOM.fr/) ©SHOM-2008. The data is extracted from SONEL (System for Observing Coastal Sea Level, www.sonel.org).

For English ports, observations of the tide along the English coasts since 1912 are accessible on the BODC site (http://www.bodc.ac.uk/data/online_delivery/ntslf/processed/), with an observation frequency of 15 min over all or part of the period (1990 – 2008). The maximum observed tidal range can be calculated directly from the measurements.

The data collected within the framework of the GLOSS programme (Global Sea Level Observing System, www.gloss-sealevel.org) can also be used.

At the planetary scale the sea-level data derived from satellite observations could be used (FES2004, Lyard et al. 2006). These data are accessible on the LEGOS site: ftp.legos.obs-mip.fr/pub/soa/maree/tide_model/global_solution/ Copyright: 1992-2005 Legos/CNES/CLS. The values of amplitude and phase of 14 components of the tide are calculated over the whole of the globe with a horizontal precision of $1/8^{th}$ of a degree. From these harmonics, the maximum amplitudes can be obtained by searching for the maxima over a duration of 18 years (astronomical return period).

The values of the periods of the various tide waves can be taken from Lefevre (2000).

For North-West Europe, a data set gathering together all the specific values of maximum amplitude was compiled from the French, English and satellite data. This data compilation was interpolated by linear kriging over the whole of the European continental shelf, which allowed us to determine the range of maximum tide at all points (Fig. 1). We used this data set to calculate the water depths with respect to mean sea level for each sounding.
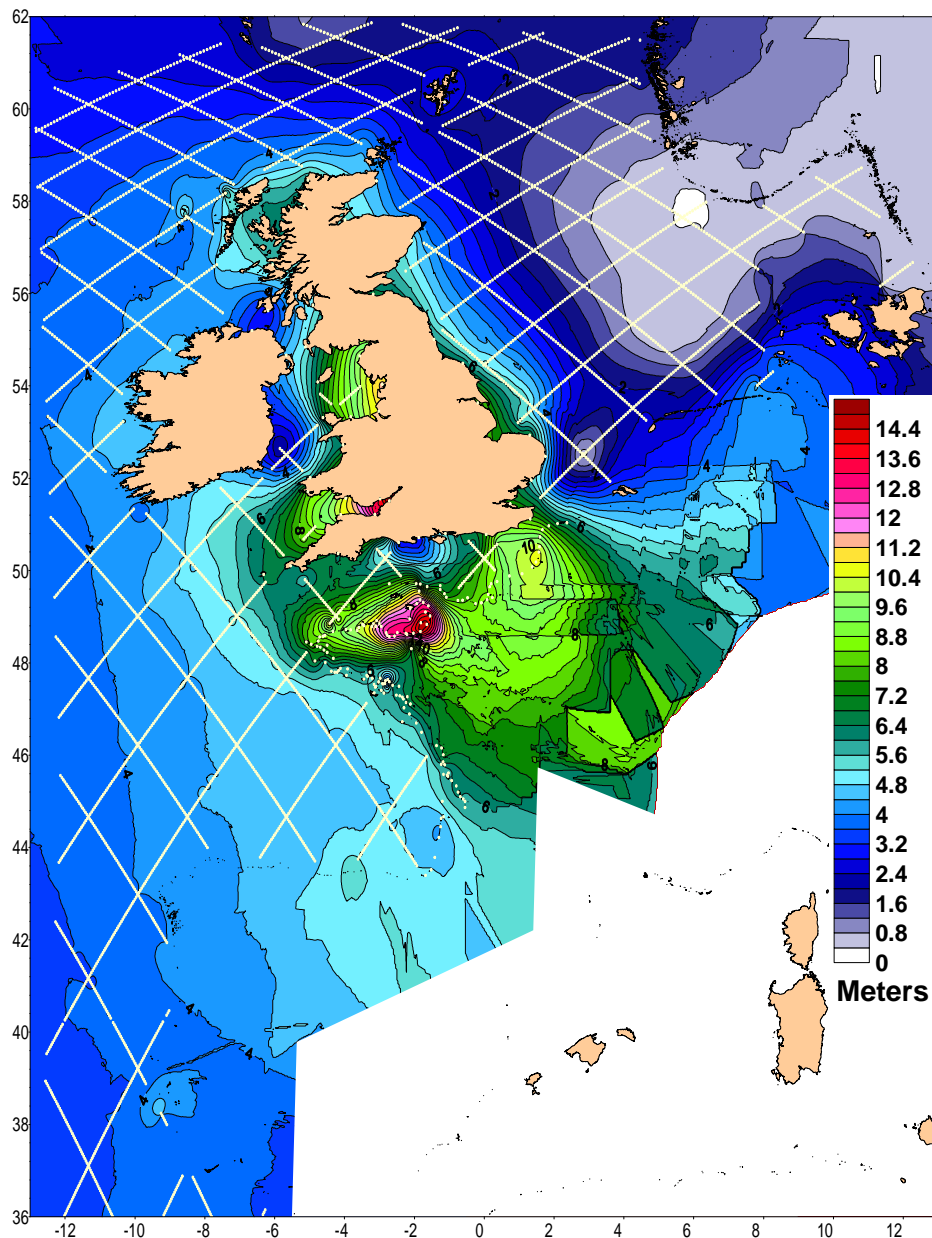
Fig. 1 ***Maximum amplitude of the astronomical tide measured on the European continental shelf (yellow dots: tide-gauge and TOPEX measurements used for interpolation).***

3. SELECTION OF SOUNDINGS

Data series are commonly found to overlap, while the spacing of the soundings and the acquisition techniques vary from one data series to another. The compilation of a homogeneous set of records involves sorting between these series. As an example, Fig. 2 shows the overlapping of data gathered by the SHOM in 1922 and records obtained during oceanographic campaigns in 2005. The surveys of 2005 have a better vertical precision than those of 1922, but their distribution makes it difficult to carry out a selection with regard to the 1922 data. While this sorting is a lengthy and cumbersome task for two sets of data, it can become insurmountable when dealing with several data series with frequent updates. We were confronted with this problem when compiling a set of records that would allow us to build nested models for the whole of the continental shelf and ranging down to a local scale (Fig. 3). We have at our disposal 187 files including 27 million records extracted from many different sources (satellite data, digitization of paper charts, oceanographic cruises, hydrographic surveys, lateral sonar, etc.). The selection of the most representative surveys within this data base required the development of the procedure described below.
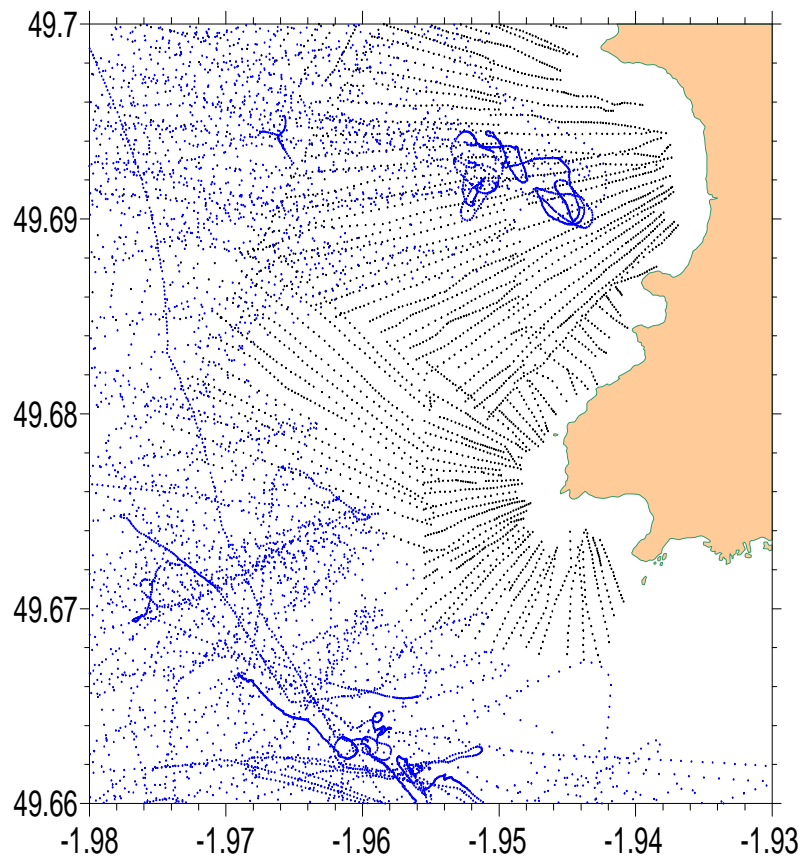
Fig. 2 *Example of overlapping bathymetric surveys. Black dots: 1922 data; blue dots: 2005 data; green line: shoreline according to the SHOM.*
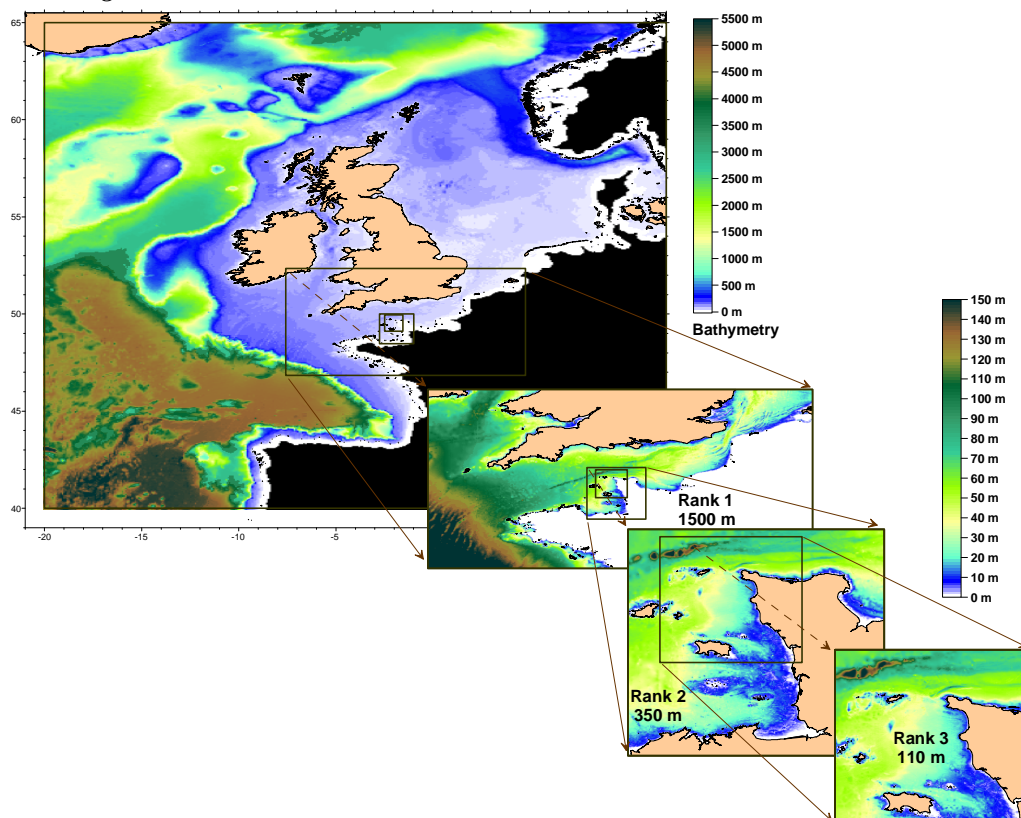
Fig. 3 *Example of overlap of models implemented on the European continental shelf.*

3.1. DATA EDITING

All the data sets used here comply with the same format; each record is represented by a line in ASCII code with X, Y and Z coordinates. X and Y represent longitude and the latitude in decimal degrees in the WGS84

reference system, Z is the underwater depth in metres counted positively towards the bottom, while the depths on land are negative. An identifier and a quality index are attributed to each data file (ID and QF, Table 1).

As an illustrative example, Table 1 presents 4 of the 187 data sets/series that we used to build the models for the continental shelf of North-West Europe.

Origin	ID	QF	Number of records	Longitude min	Longitude max	Latitude min	Latitude max	Maximum Depth
Shoreline, 25-m resolution with height of maximum tide amplitude * www.ngdc.noaa.gov/mgg/ shorelines/dated/gshhs/version1.10/ www.shom.fr	1005	4	5594584	21.0 W	15.5 E	36.0 N	65.5 N	0
Côtes de la Manche* ftp://njord.dt.insu.cnrs.fr/pub/daufin/ this study	2050	6	1927402	7.6 W	3.2 E	43.3 N	54.7 N	150
Zodiac, Cap de La Hague* this study	3030	8	33104	2.0 W	- 1.9 E	49.6 N	49.7 N	243
Istolith SHOM www.shom.fr	4000	10	8449521	4.0 W	2.5 E	43.4 N	51.5 N	1333
Digitization SHOM paper charts	10000	30	426307	34.2 W	15.8 E	34.9 N	64.8 N	6324
Gravimetric anomalies data from TOPEX depth greater than 200 m Smith and Sandwell, 1997	40000	80	790815	20.2 W	16.0 E	32.2 N	72.0 N	3939

ID: Identification; QF: quality Factor; *: available on request

Table 1   ***Principal bathymetric data files used.***

The principle of selection adopted here is to preserve, in a given zone, only those records having the best quality (QF, Table 1). To prevent loss of information during interpolation, the size of the zone, or grid cell, must be less than a quarter of the smallest calculation grid-cell used by the model concerned. The sorting procedure initially consists of using this resolution to divide up the area of interest. Three two-dimensional tables are drawn up for QF, Nb and Depth. The index of the columns corresponds to longitude, while the index of the lines corresponds to latitude. For each grid, the QF table contains the best QF values observed, the Nb table contains the number of records with the best QF and the Depth table contains the summation of depths having the best QF. When all the records have been scanned, the average depth is assigned to Depth by calculating Depth/Nb if Nb is higher than 1.

We adopted a quality index that increases in value from best to worse quality. Table 1 gives some examples of the selected values. A QF quality index is attributed to each file or type of bathymetric survey. Table 1 is an extract of a table summarising the characteristics of all the files/data files. All the records of the bathymetric data files are then processed in order of decreasing quality (with increasing QF) to find if a record already exists within the grid cell containing the data point.

If no such record exists, the following values are attributed to the data point :

− QF corresponding to the value of the file being read.

− Depth corresponding to the measured value of the record.

− Nb value of 1.

If a record already exists in this grid and its QF is the equal to the record being read, the measurement of depth is added to that of the grid and Nb is incremented by 1.

If the QF of the grid cell is better (lower in value) that that of the tested record, nothing is modified (the record is not taken into account).

This procedure is repeated for all the files that are to be compiled. The average depths are then calculated for all the grid cells where there are several records of the same quality by dividing the summed depths by the total number of records.

The results file is built up by creating a record for each cell where there are one or more records. This file is then used as a basis for interpolating the values in each grid cell of the model.

During the sorting, it is possible to register each individual record kept in a single file, by adding the identifier of the file (ID) and the number of the line in the original file. The file containing all the filtered individual records can be used to search for aberrant values and extract them from the original data.

The following parameters are required by the subroutine to select records according to their quality:

Results file name;

Characteristics of the files to be compiled. A tables is filled in beforehand by a routine that enters this information into a file (example   Table 1: Name, ID, QF, Number of records, Longitude min, Longitude max, Latitude min, Latitude max, Maximum Depth);

LoMin, LoMax, LaMin, LaMax: Limits of the area of the future models in longitude and latitude;

Size of the desired resolution of the sorting procedure.

A Boolean value indicates whether to generate a file containing all the individual selected values.

4.   ELIMINATION OF ONLAND RECORDS

During the collection of data from various origins, it is possible to observe abnormal records that are located on land. The elimination of these records is essential to ensure the coherence of the digital bathymetry model and the resulting hydrodynamic calculations. We developed a filtering procedure using data files representing the shoreline.

The procedure consists of creating a table with a grid representing the study area with the best possible resolution (at least that adopted in the preceding paragraph). This table is initialized to zero, which represents an indefinite value. The shoreline is represented in this table by assigning a value of 1 to each grid located on the coastline. The grid cells of the marine zone are then assigned a value of 2 using a filling procedure described further below. The table is then used as a test chart by checking that each record selected in the preceding paragraph lies within a grid located at sea (value 2).

Worldwide shoreline data have been made available for several years by the Global Self-consistent, Hierarchical, High resolution Shoreline Database (Wessel and Smith 1996), updated in 2009.

These data have a resolution of 100 to 250 m, sufficient for the majority of applications. In addition, improved resolution is available on the Total Prototype Shoreline Data (Satellite Derived High Water Line Data) http://shoreline.noaa.gov/data/datasheets/pgs.html).

Since each country can distribute its own shoreline data with a better precision, it is advisable to obtain information from the organizations concerned.

During the transfer of the shoreline points to the test table, we need to ensure that all the grid cells located between two points are filled. The equation of the straight line connecting the two points is calculated, and the grids located between the two points are filled using the parameters of this equation.

A filling algorithm was developed to search for zones located on land. The general principle is as follows:

A pointer is initialized by positioning it on a cell located at sea. The algorithm searches among neighbouring cells for those which do not form part of the shoreline and which are not marked as marine. The co-ordinates of these cells are recorded in a Holes table. The NbHoles register is incremented with each new cell added to the Holes table. The pointer is then positioned on one of these cells, NbHoles is decreased by one increment and a value is assigned to the cell representing the sea. The test is then reiterated on the surrounding cells, and so on. When all the cells located around the pointer have been assigned to the sea or the shoreline, the pointer is positioned on the last empty cell recorded in the Holes table and the procedure starts again. When all the cells of the Holes table have been scanned (NbHoles=0), the filling is finished.

This procedure is fast and simple to program; it requires large-sized tables in the case of wide coverages with a precise resolution. For example, with a resolution of 25 m for a coverage of 1000x1000 km, we need a table of $1.6.10^9$ cells. Therefore, it is advantageous to divide up the area in such a way that several files of records can be prepared for the main zones and resolutions used in the calculations: a file of records with a resolution of 500 m for continental-shelf models, and some files with a resolution of 25 m or less for coastal models. This adaptation is also useful to facilitate interpolation calculations. Such calculations would be unnecessarily long if we used a sounding every 25 m to calculate the bathymetry of a continental-shelf model having a grid cell size of 1 km.

The subroutine for elimination of onland records requires the following parameters:

Shoreline data file name covering the zone to be masked;

Bathymetry file name to be corrected by the mask;

Output file name of the corrected bathymetry;

LoMin, LoMax, LaMin, LaMax: coverage of the masking zone;

Cell size (resolution in meters) of the masking table;

LoInit, LaInit: coordinates of a point located in the marine zone in order to initiate filling.

5.   TAKING INTO ACCOUNT OF THE SHORELINE

In general, bathymetric surveys comprise soundings acquired at sea in zones accessible by oceanographic vessels. As a result, areas of shallow water are poorly represented, in particular the intertidal zone (foreshore

between high and low tide levels). If the land/sea boundary is imprecisely known, the procedures for interpolation of bathymetry records are likely to give poor results.

This problem is made worse when the shoreline is advancing towards the sea, because soundings located on either side of the newly advancing coast will give depths by interpolation at points actually situated on land. One way of avoiding this problem involves adding points on land at each location where the shoreline is advancing seaward. This method is cumbersome and leads to interpolation errors by creating unrealistic slopes. **Fig. 4** illustrates this phenomenon, with a over-representation of the onland zones between the shoreline and the sea.
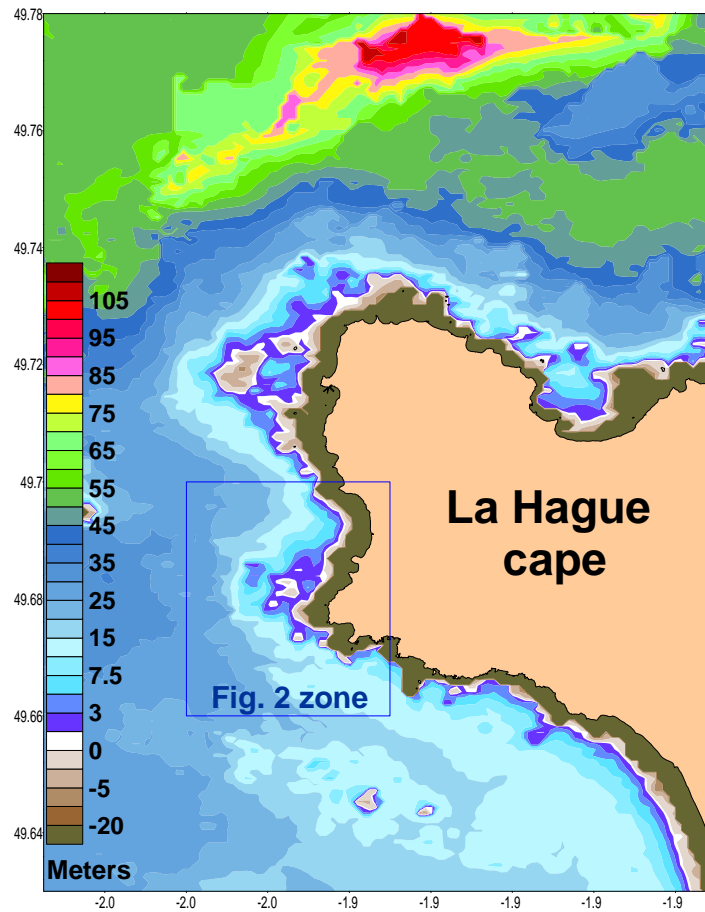
Fig. 4 ***Detail of bathymetry initially present in rank-3 model (grid cell size of 110 m). The shoreline-delimited area is represented in light brown.***
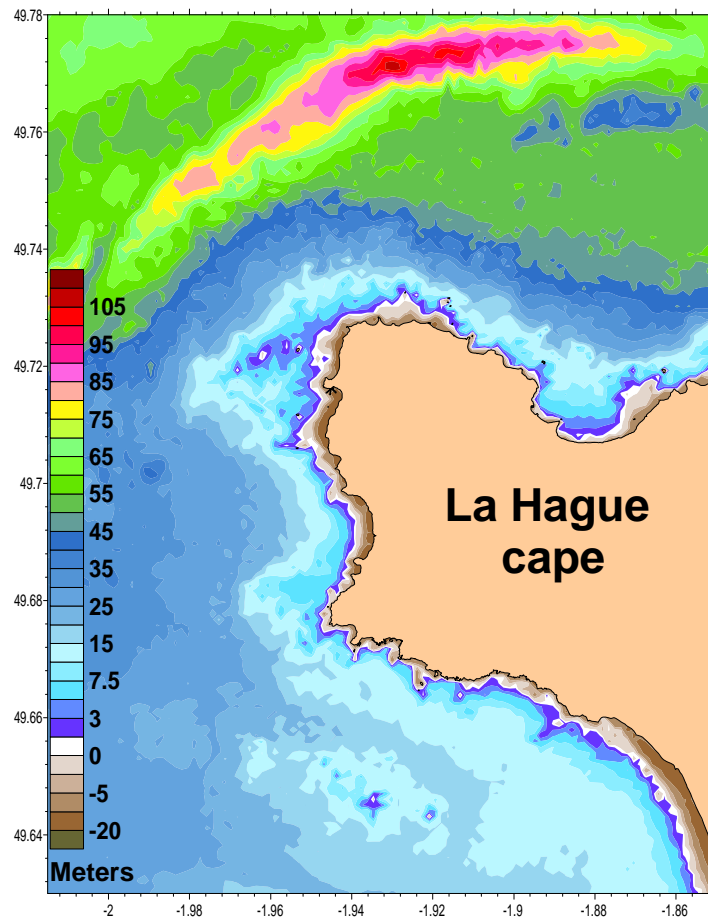
Fig. 5 ***Detail of the bathymetry of rank-3 model after accounting with several sets of data and shoreline used as bathymetry.***

The technique adopted to solve this problem consists of using the shoreline as a series of bathymetric records whose elevation corresponds to the mean sea level or the highest tide level, according to the origin of the shoreline. The slope close to the coast will be more realistic, even if the intermediate data on the foreshore are not available.

To represent the terrestrial limit of the model, it is necessary to assign the most realistic possible coordinates to the points defining the shoreline. According to the characteristics of the shoreline communicated by the provider of the data, the elevation of each point on the shoreline represents either the mean sea level, or the level of the highest tide. Insofar as hydrodynamic models are likely be used to simulate extreme weather events, the values of the maximum amplitudes of the tide should be applied to the shoreline. The highest astronomical tide can be calculated from the tidal harmonics by searching for the maximum over a return period of 18 years (Ballay 2002). This assumes that the position of the shoreline is not modified during maximum tides. The tidal observation data presented in section 2.1 can be used for this purpose. We add 0.5 m in height to the maximum values of amplitude to make sure that the shoreline is never crossed. Particularly for studying the incursion of the sea towards the land during extreme events, it will be necessary to supplement the bathymetry by the onland topography of the zone considered.

Once the shoreline data are transformed into a series of bathymetric records, we need to ensure that the shoreline forms an impassable limit when carrying out the interpolation. In general, the shoreline is made up only of significant points. Even if the resolution of representation is 100 m, certain points can be spaced at a distance of several hundreds of metres in the case of a straight shoreline. The shoreline points must then be supplemented by adding intermediate points, which are spaced at a shorter distance than the minimum resolution of the model (25 m for a calculation grid-cell size of 50 m, for example). This over-representation guarantees that the process of interpolation will not generate any landward incursion of the sea.

6. INTERPOLATION

By converting the shoreline data into points that are spaced sufficiently closer together, we can add them to the selected bathymetric records. This file is then ready to be interpolated. The purpose of the interpolation is to calculate the most realistic depth in each model calculation cell. While many methods are available, their discussion falls outside the scope of this study. One of most robust methods is triangulation, which involves a linear interpolation between records surrounding the calculation grid cell. This method does not carry out any smoothing, and can lead to unrealistic changes of slope. We generally prefer to use linear kriging as performed for instance by

the Surfer® software. It should be checked whether the search area is sufficiently extensive to cover zones where the density of the bathymetric surveys is low. The Surfer® software is of interest because it can be activated from the program used for treatment of the bathymetric records (using the COM objects), while facilitating the visualization of results. In this way, a single program can carry out all the tasks.

7. CHECKING

During the incorporation of a new data file, it is advisable to check that there are no errors subsequent to interpolation. In general when the sets of data are very large, only the worst errors can be identified. By calculating the gradient between two points, we can detect unrealistic gradients, while a comparison with previous data can also be useful. The positive depths interpolated by error on land can be filtered using the procedure described above for the elimination of onland data.

In certain cases, the models display inconsistencies when the bathymetry near the limits is too complex or when bays have dimensions incompatible with the calculation grid. It is possible to eliminate these zones artificially by creating masks made up of a series of points that can be treated in a similar way to the shoreline.

It remains to format the depth data in a grid that is readable by the model.

8. DISCUSSION

The methodology presented here was first developed in 2004, but has been subsequently automated and implemented many times to generate new models on the European continental shelf. The setting up and generation of a new digital bathymetry model takes approximately three hours. This process includes the incorporation of new bathymetric data, selection of the best records, the elimination of onland data points and the taking into account of the shoreline. Most of the processing time is taken up by the interpolation of values in each cell of the model, followed by displaying the results and compiling files that are usable by models. These tasks are carried out automatically for all the nested models required in the more detailed type of model.

The main advantage of this methodology is its robustness. Apart from the exclusion of aberrant record values, we never found it necessary to modify manually the bathymetric data calculated for the models. This robustness allows us to develop automatic zoom tools for specific needs such as simulating the consequences of accidental releases. For example we have developed a specific models to improve the work organization during oceanographic campaigns.

The precision of dispersion models remains strongly dependant on the number and quality of bathymetric records in the zone covered by the calculations. The results obtained from our methodology shows that it makes good use of the available data. The representativity of the currents and simulated dispersion gives evidences (Bailly du Bois et al.; 2005a and 2005b; Blanpain et al.; 2009).

9. CONCLUSION

The main advantage of the method proposed is its robustness. Apart from the exclusion of aberrant record values, we never found it necessary to modify manually the bathymetric data calculated for hydrodynamic models. The method improves significantly the bathymetry used and allows the automatic definition and updating of the bathymetry. In our case, setting up of a new model with specific coverage and resolution takes only two hours on a desktop computer.

The methods applied here can be transposed to other hydrodynamic models concerning other areas of the continental shelf. The emerged/uncovered zones are represented by taking the shoreline into account, which should facilitate the automatic definition and updating of the bathymetry used for hydrodynamic models in other parts of the World using GSHHS shoreline data and FES2004 tide data.

We wish to continue the approach proposed here in collaboration with other interested organizations to build a complete and homogeneous bathymetric data base. In this way, it will be possible to implement - quickly and at less time expense - realistic hydrodynamic models in any area.

10. ACKNOWLEDGMENTS

11. REFERENCES

IHO (1996) Report of "groupe de travail de l'OHI sur les marées" Monaco, 22-25 april 1996.

Bailly du Bois P, Dumas F (2005a) Fast hydrodynamic model for of medium- and long-term dispersion in seawater in the English Channel and southern North Sea, qualitative and quantitative validation by radionuclide tracers. Ocean Model 9/2:169-210.

Bailly du Bois P, Dumas F, Solier L (2005b) Hydrodynamic modelling of short-term dispersion in a macro-tidal sea, validation by high-resolution radionuclide tracer measurements. Radioprotection, Proc ECORAD 2004 "The scientific basis for environment protection against radioactivity" Aix-en-Provence (France), 6-10 september, 2004. Edited by: J.C. Barescut, J.C. Gariel and J.M. Péres 40:563-568.

Blanpain O, Bailly du Bois P, Cugier P, Lafite R, Lunven M, Dupont J, Le Gall E, Legrand J, Pichavant P (2009) Dynamic sediment profile imaging (DySPI): a new field method for the study of dynamic processes at the sediment-water interface. Limnol Oceanogr: Methods 7:8-20.

Ballay A (2002) Cartographie des niveaux extrêmes de pleine et basse mer atteints après une période de retour donnée. Service Hydrographique et Océanographique de la Marine, Section Géodésie-Géophysique, febr 2002

Gorman G, Piggott M, Pain C, de Oliveira C, Umpleby A, Goddard A (2006) Optimisation based bathymetry approximation through constrained unstructured mesh adaptivity. Ocean Modelling 12:436-452

Jones JE, Davies AM (2005) An intercomparison between finite difference and finite element (TELEMAC) approaches to modelling west coast of Britain tides. Ocean Dynam 55: 178–198

Lefevre F (2000) Modélisation des marées océaniques à l'échelle globale : assimilation de données in-situ et altimétriques. Thesys Université Toulouse III - Paul Sabatier 22 sept 2000, 280p http://fabien.lefevre.free.fr/these.htm accessed 30 april 2010

Lyard F, Lefèvre F, Letellier T, Francis O (2006) Modelling the global ocean tides: a modern insight from FES2004. Ocean Dynamics 56:394-415

Penven P, Marchesiello P, Debreu L, Lefèvre L (2008) Software tools for pre- and post-processing of oceanic regional simulations. Environ Model Softw 23:660 -662

Wessel, P., and W. H. F. Smith (1996), A global, self-consistent, hierarchical, high-resolution shoreline database, J. Geophys. Res., 101(B4), 8741–8743. http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html accessed 30 april 2010