On the Estimation of Errors in Sparse Bathymetric Data Sets

Martin Jakobsson, Brian Calder, and Larry Mayer
Center for Coastal and Ocean Mapping and Joint Hydrography Center,
University of New Hampshire, Durham, New Hampshire

Short title: ESTIMATION OF ERRORS IN SPARSE BATHYMETRIC DATA
Abstract. We address the problem of compiling bathymetric data sets with heterogeneous coverage and a range of data measurement accuracies. To generate the required regularly spaced grid, we are obliged to interpolate sparse data; our objective here is to augment this product with an estimate of confidence in the interpolated bathymetry based on our knowledge of the component error sources. Using a direct simulation Monte Carlo method, we use data from the International Bathymetric Chart of the Arctic Ocean (IBCAO) database to develop a suitable methodology for assessment of the standard deviations of depths in the interpolated grid. Our assessment of errors in each dataset are heuristic but realistic, and are based on available metadata from the data providers. We show that a confidence grid can be built using this method, and that this product can be used to assess reliability of data sources used in the compilation. The methodology is applied here to bathymetric data, but is equally applicable to other interpolated data sets, such as gravity and magnetic data.
1. Introduction

There is a growing demand in the geophysical community for better regional representations of the world ocean’s bathymetry. Researchers dealing with sea-level change, ocean circulation, sediment transport, seafloor spreading, and modeling of ice sheets all require information on the shape of the sea floor. While multibeam sonar systems are filling hydrographic databases with relatively accurate bathymetric data sets from specific target areas, the quantity of observations being collected in most of the deep oceans (particularly remote areas) remains static and relatively small. The accuracy and resolution achievable with modern sonars is far better than that obtained with older single beam systems particularly considering the recent improvements in ship positioning. However, given the vastness of the oceans and the relatively limited coverage of even the most modern mapping systems, it is likely that many of the older data sets will remain part of our cumulative database for several more decades. Given the reality, regional bathymetrical compilations that are based on a mixture of historic and contemporary data sets will remain the standard for the production of bathymetric charts. This raises the problem of assembling such bathymetric compilations and utilizing data sets with both a heterogeneous cover and a wide range of accuracies [Bernardel, 1997; Macnab and Jakobsson, 2000].

We address the issue of compiling bathymetric data sets with heterogeneous cover and a range of accuracies in the context of generating regularly spaced grids. For generating the grid we are often forced to use a complex interpolation scheme due to the
sparseness and irregularity of the input data points. Consequently, we are faced with the difficult task of assessing the confidence that we can assign to the final grid product, a task that is not usually addressed in most bathymetric compilations. Traditionally the hydrographic community has considered each sounding equally accurate and there has been no error evaluation of the bathymetric end product. This has important implications for use of the gridded bathymetry, especially when it is used for generating further scientific interpretations. The method we describe is equally valid for the compilations of any other gridded data that are based on multiple sources with varying densities and accuracies (e.g. magnetic data, gravity data, etc.). We approach the problem of assessing the confidence of the final bathymetry gridded product via a direct-simulation Monte Carlo method. We start with a small subset of data from the International Bathymetric Chart of the Arctic Ocean (IBCAO) grid model [Jakobsson et al., 2000]. This grid is compiled from a mixture of data sources ranging from single beam soundings with available metadata, to spot soundings with no available metadata, to digitized contours; the test dataset shows examples of all of these types. From this database, we assign a priori error variances based on available meta-data, and when this is not available, based on a worst-case scenario in an essentially heuristic manner. We then generate a number of synthetic datasets by randomly perturbing the base data using normally distributed random variates, scaled according to the predicted error model. These datasets are next re-gridded using the same methodology as the original product, generating a set of plausible grid models of the regional bathymetry that we can use for standard deviation estimates. Finally, we repeat the entire random estimation
process and analyze each run's standard deviation grids in order to examine sampling bias and standard error in the predictions. The final products of the estimation are a collection of standard deviation grids, which we combine with the source data density in order to create a grid that contains information about the bathymetric model's reliability.

2. Data Description

2.1. Implementation of the IBCAO Grid Model

The General Bathymetric Chart of the Oceans (GEBCO) fifth-edition Sheet 5.17 \cite{Canadian_Hydrographic_Service_1979}, portraying the sea floor above 64°N, has been considered the standard map of the Arctic Ocean for over two decades. While this contour map provided a general description of major bathymetric features, evidence was growing to indicate that many of the smaller and scientifically significant features were poorly defined. This is not surprising as the Arctic poses particular problems for data collection due to its remoteness and especially to its permanent sea-ice cover. The creation of Sheet 5.17 was possible only because a few profiles collected by US nuclear submarines had been released to supplement data from drifting ice islands, scattered spot soundings and a few tracks from icebreakers. These were all the direct data available at the time of compilation of Sheet 5.17. Since then, there has been much anecdotal evidence about the inadequacy of Sheet 5.17 and previously published bathymetric maps, particularly in the deep waters of the central basins.
Given the poor state of the regional Arctic bathymetry maps, the International Bathymetric Chart of the Arctic Ocean (IBCAO) was initiated during 1997 in St Petersburg, with the goal of collecting all available data north of 64°N [Macnab and Nielsen, 1999]. One of the major goals was to compile a regular grid model from the data collected within IBCAO. The IBCAO data consisted of digital information that was obtained during recent icebreaker and SCICEX submarine cruises and older digital information that consisted of recently-declassified soundings collected between 1957 and 1988 by submarines of the US and UK Navies, and of observations obtained from the public-domain archives of world and national data centers. In addition hydrographic charts and compilation maps, portraying depth in the form of point soundings and hand-drawn contours, published by the Russian Federation Navy [Head Department of Navigation and Oceanography, 1999], by the US Naval Research Laboratory [Perry et al., 1986; Cherkis et al., 1991; Matishev et al., 1995], and by other agencies, were digitized using heads up digitizing techniques (500-1000 m between digitized points on contours) to supplement the original bathymetric measurements in the IBCAO data base.

The IBCAO grid model also contains topography which was derived mainly from the USGS GTOPO30 topographic model [U.S. Geological Survey, 1997], with the exception of Greenland where the topographic model developed by KMS, the Danish National Survey and Cadastre, was used [Ekholm, 1996]. In order to constrain the coastline the World Vector Shoreline (WVS) was used in all areas except Greenland and northern Ellesmere Island, where an updated coastline was made available by KMS.
Initially, the original bathymetric soundings were corrected for sound speed in water using Carter’s tables, or CTD profiles where available. After sound speed corrections, a suite of tools and statistical routines based upon the Helical-Hyperspatial (HH) scheme for data encoding [Varma et al., 1990] was used to flag data as unusable if they were found to not statistically conform to nearby data. After this initial statistical cleaning all data (digitized bathymetric contours, land and marine relief grids, point, profile and swath observations, and vector shorelines) were imported into Intergraph’s MGE (Modular GIS Environment) with projection parameters set to polar stereographic on the WGS 1984 ellipsoid, true scale at 75°N. The observations along ship tracks were sub-sampled to maintain a minimum of 500–1000 m between every point in each track. Soundings were color-coded according to depth to facilitate a visual inspection of the statistical cleaning results. Outliers, cross-track errors, and the fit between isobaths and original observations were checked during this process. Further suspicious soundings were flagged, and where contours showed major discrepancies with soundings, the contours were adjusted manually to fit the new bathymetric track line data.

After editing the entire Arctic Ocean bathymetry data set, the mixture of track and digitized contour values were used to construct a grid with a cell size of 2.5 × 2.5 km. The variable density of the different components in the compilation led us to consider an interpolated gridding algorithm, namely the continuous curvature spline in tension algorithm of Smith & Wessel [Smith and Wessel, 1990], as implemented in the GMT package [Wessel and Smith, 1991]. Prior to gridding the data was preprocessed by applying a block median filter with a block size equal to the final grid cell spacing of
2.5 × 2.5 km. This filtering serves the main purpose of preventing spatial aliasing. Finally, the GMT continuous spline-in-tension algorithm was used with the tension (T) parameter set to 0.35 in order to avoid overshooting in the interpolated regularly sampled surface. The resulting grid was inspected visually and problems identified. For example, a common problem was in narrow fjords without bathymetric data points where the gridding algorithm assigned 0 m values; the same value as the coastline. This was controlled by manually inserting control contours typically representing a depth of a few meters, near the coastline, which conditioned the interpolation. A shaded relief of the entire IBCAO grid is shown in Figure 1. Further description about the IBCAO grid model can be found in [Jakobsson et al., 2000; Macnab and Jakobsson, 2000].

2.2. Experimental subset of IBCAO

Our experiment is based on a subsection of the data used to construct the IBCAO grid, as shown in Figure 1. We have chosen the area around Svalbard since it contains a cross-section of the various régimes within the IBCAO source data, including GTOPO30 land data, near-shore regions, dense single-beam data, transect lines, bathymetric contours and control contours. There is also a significant depth range due to the relatively shallow areas of the Barents Sea around Svalbard and the contrasting deep of the Greenland Sea in the western part of the area where the Knipovich Ridge, the northernmost part of the North Atlantic Spreading Ridge, is coming through. We have followed the exact same methodology as used for the construction of the IBCAO grid model. The data stored in the compilation database have been inspected, and cleaned
relative to the original data, and, if not previously adjusted, the depths have been
corrected for sound speed using Carter’s tables. The resulting data set shows a distinct
data density gradient from very dense survey data near Svalbard and in the south-east
of the region to poorly constrained re-digitized contours in the north and north-east
(Figure 2).

2.3. Contours and Associated Problems

The contour is still the traditional means of displaying bathymetry (or any
geophysical parameter). In theory, a bathymetric contour represents exactly one depth
along its entire extension. However, in reality the contours are typically interpolations
based on underlying sparse ship track data and, thus, the depth is only true where
the contour crosses a trackline. Without the source data and its associated meta-data
available, it is not possible to judge the accuracy of the bathymetric contours. In addition
a manually derived contour map will inherit a style from the cartographer/geophysicist
who drew it. All this makes it difficult to produce an error estimation of a bathymetric
contour map. When contours are used as source data for creating a gridded surface these
problems go along with it. In addition, a gridding algorithm suitable for interpolating
a regular grid from the contours is required [Smith and Wessel, 1990]. For example,
terracing is an artifact in grid models that arises from gridding source data that consists
primarily, or in parts, of digitized contours. The terracing is due to a bias towards the
contour values. If artificial illumination is applied to a gridded surface suffering from
terracing the surface shows steps between terraces at the contour values. The IBCAO
grid shows some terracing along the central Arctic Ocean continental slopes where
digitized contours dominate. A histogram of depths of the gridded surface's nodes
clearly reveals a bias towards the digitized contours (Figure 3).

3. Error Model

3.1. Methodology of Monte Carlo Simulation

In principal, estimation of errors associated with the grid is a relatively simple
matter. We need to gather a number of datasets for the same area (keeping track of
the error sources in each), estimate depths in the area concerned, and then look at the
variability in the depth estimates. However, in the regional case the vastness of the area
and the difficulty and expense of collecting the data precludes repeated surveys. As an
alternative, we must consider whether we can approximate the error estimates required
based on the best available data, our knowledge of the likely errors involved, and a
simulation method.

The Monte Carlo method [Hammersley and Handscomb, 1964; Gentle, 1998] is
a numerical technique for evaluation of difficult integrals. In many cases, integrals
cannot be solved analytically, and in some cases not even by the usual quadrature
methods [Press et al., 1986]. This is particularly true where the integral is over multiple
dimensions as often happens in computational probability and physics [Brooks, 1998;
Binder and Heerman, 1988]. In these cases, the simulation methodology of the Monte
Carlo method is preferred.
At its simplest, the Monte Carlo method is very straightforward. For an integral \( I = \int_0^1 f(x)dx \) (to which we can reduce any integral in one dimension with suitable transformations), the Monte Carlo method proceeds by refactorizing the integrand \( f(x) = g(x)p(x) \), and then observing that if we interpret \( p(x) \) as a probability density function with support on \([0, 1]\), then \( I = \int_0^1 g(x)p(x)dx \) is simply the expected value of \( g(x) \), \( \mathbb{E}[g(x)] \). Given a set of \( N \) samples from the probability density we can approximate this expectation by the sample mean,

\[
\mathbb{E}[g(x)] \approx \frac{1}{N} \sum_{i=1}^{N} g(X_i)
\]

in the usual way. Consequently, as long as we can factorize the integrand and generate samples from the resultant pdf, we can approximate the integral with an error proportional to the number of samples used (and hence indirectly proportional to the time that we are willing to expend on the process). In the simplest case, we can choose the uniform distribution for \( p(x) \), and hence \( g(x) = f(x) \); the estimation is therefore:

\[
I \approx \frac{1}{N} \sum_{i=1}^{N} f(X_i) \quad \text{where} \quad X_i \sim \mathcal{U}[0, 1]
\]

A small subtlety of the method is that, since these integral estimates are based on a random dataset, they are themselves subject to variation (i.e., are random variables). As usual, we must provide an estimate of this variance to complement the base estimate. However, since we do not normally know the error in each estimate, we are forced to apply the same technique again and estimate the Monte Carlo error by multiple repeated runs of the whole simulation. It is important to distinguish carefully between estimated standard deviation of the computed bathymetry (the target of the simulation)
and the summary of the sampling distribution of the standard deviation grids (i.e., the standard error of the standard deviation estimate), computed between different simulation runs. Although computed in a very similar fashion, the former exhibit true variation corresponding to the problem under investigation, while the magnitude of the latter grids is an artifact of the sampling approach to estimation. The estimation of the Monte Carlo error is simply a quality assurance check and calibration.

The chief difficulty with this simple method is that this Monte Carlo error reduces only slowly with the number of component estimates, \( N \), and hence the method is not very efficient as stated. Most improvements to the method are targeted at making the error shrink more quickly [Gentle, 1998].

In the case of estimating uncertainty in the bathymetric grid, however, the task is sufficiently simple that we can use a direct simulation, paying a penalty in time for simplicity in theory and implementation. This simplified method is analogous to taking repeated measurements in the same area; given the best data available, we generate synthetic (but realistic) pseudo-datasets based on our knowledge of the likely error sources and their magnitudes, and we then run the experiment assuming that the datasets are truly independent. Given our assumptions, summary statistics generated from the simulations are valid estimates of the true values of error, in the same way as above.

The first principal assumption made here is that the datasets and the measurements within them are independent of each other and that they are free of any systematic bias. In this case, we can use the data points given as a basis for all of the pseudo-datasets,
perturbing about the values supplied. In effect, we assume that the points recorded are unbiased estimates of the mean bathymetry and position. Our other principal assumption is that the errors in location and depth are normally distributed, independent of each other and of each data point. This assumption is more weakly justifiable, since we may have some systematic bias in navigation (e.g., a miss-navigated submarine track 10 km from the true location), or we may have some correlation between the two horizontal offsets. However, such fine detail is essentially unknown and unknowable in the datasets we are considering, and we are forced, reluctantly, to accept this assumption in order to carry out the analysis.

In a similar vein, we note in passing that this analysis does not give us any more insight into the error budget for the grid than a full formal error analysis would. However, it does provide a very simple way to carry out what would otherwise be a very complex computation. Pragmatically, we trade off accuracy for tractability.

### 3.2. Estimation of Errors in the IBCAO source data

Our error modeling approach is based on an assumption of normally distributed random errors in the source data. In the case of bathymetric data this may be subdivided into errors in determining position (xy) and errors in measuring depths (z). For recently collected survey data an estimate of the random errors and possible constant errors (which then can be corrected) may be available from those who collected the data. However, in the case of the IBCAO source data, the majority of the data sets are historic and, thus, only the meta-data is available to make a realistic initial
random error assumption. If there is no meta-data available the assigned errors must be based on a worst-case scenario in order to highlight this uncertainty. The most critical information in the meta-data is: type of positioning/navigational instrumentation, bathymetric instrumentation and year. Lack of other information like geodetic datum and sound speed correction would also contribute to the initial error assignment. From the meta-data the random error is estimated at a selected confidence interval, 95%. This means that 95% of the normally distributed positions should fall within a circle with a radius of the assigned error. It is not a simple and straightforward task to assign an error simply based on the meta-data but it is the only approach possible for historic data sets. For example if it is found that the positions were acquired using a GPS system during year 1990 the random error may be in the order of (±80 m [Wells et al., 1986]) whereas if the positions were acquired using Loran C it gets more complicated since the accuracy of Loran C varies to a greater extent with time and geographic location [Maloney, 1985]. Again, the worst-case scenario is the easiest and safest approach.

Constant errors are more problematic to account for in historic data sets, although they may possibly be distinguished through crossing track lines.

Some of the data sets within the IBCAO source data have no meta-data associated with them. However, in our error modeling experiment, which is focused on developing the modeling approach rather than producing an accurate estimate of the errors in the subset of the IBCAO grid around Svalbard, we have assigned somewhat arbitrary errors to all data sets based on the rough classification described in Table 1. Contours are assumed to have the largest errors whereas the data from the Norwegian sources,
which mainly consist of survey data, and from the Oden icebreaker collected using GPS positioning system is considered to have the highest accuracies. The data collected using submarines are considered fairly inaccurately positioned (±5000 – 10000 m) due to the use of inertial navigational system for long periods between surface fixes. The data collected with the Ymer icebreaker was positioned using fixes from a Magnavox one-channel satellite navigational system [Eldholm et al., 1982], which we have assigned an accuracy on the order of ±1 nm. The data retrieved from the NGDC data center is assigned common positional and depth errors. A full error estimation of the entire IBCAO grid is one of our future goals. This will require a large amount of time-consuming “data detective” work in order to find meta-data for many of the older data sets. It will then be possible to assign errors that are more closely related to the “true errors”.

4. Data pre-processing and simulation

4.1. Pre-processing

We used the bounds indicated in Figure 1 to extract the relevant data from the IBCAO compilation. Data are represented as flat-file (x, y, z) triples using projected coordinates and corrected depths.

4.2. Dataset Simulation

The experimental estimation of standard deviations on the grids consists of a number of repeated simulations. In particular, we have to consider M sets of N grids.
We therefore subscript all variables $A_{nm}$ or $A_m$ as appropriate, where upper case bold letters indicate matrices (grids) and lower case bold letters indicate (column) vectors. Operations on grids are always taken pointwise (so $A = \mathcal{F}(B)$ for some operator $\mathcal{F}(\cdot)$ means $A_{ij} = \mathcal{F}(B_{ij}) \forall i, j$ on the domain). Sets of variables are indicated by sans serif letters, e.g., $Y = \{Y(1), \ldots, Y(k)\}$; when appropriate, we refer to components of a set indexed over $\mathbb{N}$ (the set of natural numbers) with an essentially arbitrary, but fixed, indexing scheme.

Given the collection of cleaned datasets $X = \{X(1), \ldots, X(s)\}$ and a corresponding error model, $E = \{e(1), \ldots, e(s)\}$, $e(i) = [\sigma_x^2(i), \sigma_y^2(i), \alpha_0^2(i)]^T$, we generate a pseudo-dataset $X_{nm}$ by perturbing each sounding with a random vector as follows:

\[
X_{nm} = \{X_{nm}(1), \ldots, X_{nm}(s)\}
\]

\[
X_{nm}(i) = X(i) + \mathcal{E}(i)
\]

\[
\mathcal{E}(i) = [e_1(i), \ldots, e_{J(i)}(i)] \quad \text{where} \quad J(i) = |X(i)|
\]

\[
e_j(i) \sim \mathcal{M}(0, \text{diag}(1, 1, \frac{1}{z_j^2(i)})e(i))
\]

Gaussian variates are generated using the Box-Muller equations driven by a non-linear congruential generator that is known to have sequence length of at least $2^{35}$ and produces equally random bits in all sections of the output word. The uniform variates are scaled to $[0, 1)$ before conversion to Gaussian distributions.

The basic component of the simulation is a block of $N = 100$ pseudo-datasets, from which we construct a set of $N$ grids using the same algorithm as the IBCAO compilation, using a mask prepared from GTOPO30 topography to constrain the standard deviation.
estimate to be zero on land. We then compute the expectation $B_m = E[B_{nm}]$ and the standard deviation $\epsilon_m = \sqrt{E[(B_{nm} - B_m)^2]}$ of this set, the latter estimating our confidence in the former's depth prediction. Computations are done directly on the grids using the GMT grid calculator; this avoids any conversion errors or approximations. The standard deviation estimate $\epsilon_m$ is the primary outcome of the simulation.

To estimate the Monte Carlo error, we repeat the above basic simulation $M = 20$ times. We then compute the expectation $\bar{\epsilon} = E[\epsilon_m]$ and standard error estimate $\epsilon = \sqrt{E[(\epsilon_m - \bar{\epsilon})^2]}$ of the individual standard deviation grids, providing us with a spatially localized estimate of the variability of standard deviation at each estimation grid point.

5. Results

5.1. Standard Deviations in Gridding and Gridding Stability

Standard deviation grids for a single run of the algorithm (i.e., $\epsilon_m$) are shown in Figure 4a and 4b. We can visualize the error in two ways: either as true meters, or as a percentage of the depth estimated from the unperturbed data, $X$. Based on the assumption that we are more interested in relative errors, especially in the near-shore region, we will concentrate mainly on the percentage error grid, Figure 4b.

On first examination, the results appear to agree with intuition. In regions where there have been rigorous hydrographic surveys (e.g., 77°N 22°30'E), the estimated error is significantly lower than regions where only a single trackline is used to constrain the
grid (e.g., at 82°N 5°E). We also see that where there are track lines, the error is lower, and that inshore the error is proportionately higher.

However, comparing the grids to the source data density grid, we see some anomalies. For example, the region near 79°30′N 37°E has suspiciously low error given the scarcity of data in the area. We attribute this to the smoothing interpolative nature of the gridding algorithm and the fact that the source data in this region predominantly derives from contours. That is, in flat regions with little data enclosed by contours, what we see is a smooth approximation between the contour limits, rather than a realistic error estimate.

Consequently, we chose to remove from consideration areas that are equal to or smaller than 7500 × 7500 m that contain no soundings, to prevent too many spurious removals. The resulting reduced grid is shown in Figure 5, and by comparison with Figure 4b, we can see that the anomalous area described above is completely removed. We note in passing that some variant of a combination of sounding density and different resolution grids may be a way to approach a prediction of the required gridding density for any particular data set, a topic we are currently investigating further.

With the empty grid cells removed, we can interpret the results with more confidence. The major feature of the grid is the significantly increased error in regions of higher slope (Figure 5 and Figure 6). This is principally due to problems of excessive horizontal error, which cause a significantly increased depth error as they shift slopes from place to place. This is also the case for the ridge of high error running from Bjornoyna (74°30′N 19°E) to Svalbard, although the ridge is not as obvious in the
bathymetry.

Perhaps more surprising than the regions with significant errors is the remarkable uniformity of error over much of the shelf area to the south-east of Svalbard. This is due principally to the relatively even distribution of soundings over the entire area (which are mainly derived from Norwegian hydrographic survey data), in combination with the general lack of large relief in the area. If the region being studied was essentially flat on a suitably large horizontal scale, then it is immaterial how much horizontal error is found in the data; all soundings should indicate the same true depth, and hence the principal error source observed will come from vertical error in the sounding system, rather than from geographical variability coupled into the soundings through mislocation.

The potential of this approach is demonstrated by the identification of the error associated with the seamounts at 82°N 24°E (Figure 4a, 4b and 6). The error estimate shows clearly that significant error is associated with the whole seamount, rather than just the slopes. Indeed, a profile of the gridded bathymetry across the feature shows a relief of approximately 1500–1700 m from the surrounding seafloor, and the estimated standard deviation is over 1000 m on the tops of the seamounts, suggesting that we would not be able to tell the seamounts from the background with any significant confidence (in the sense of a statistical test of the null hypothesis that there is no difference between the background and the tops of the seamounts). In fact, these seamounts have been found to be generated by sounding errors from a malfunctioning echo-sounder aboard the R/V Oden, which is the only data to delineate the features. The ARK-XV cruise with the R/V Polarstern (not in the database) showed that these
'seamounts' are in fact not present (Hans-Werner Schenke and Wilfried Jokat pers. comm.).

5.2. Monte Carlo Estimation Errors

The estimate of Monte Carlo error, \( e^2 \), is shown in Figure 7. Recalling that the purpose of this grid is to show regions of the target area where our prediction of standard deviation is more variable, we can see that most areas are estimated in a stable fashion, but that there is more variability in the deeper areas. We conjecture that this is simply due to the higher vertical error in deeper water, resulting in higher variability in the grids, and hence in both the standard deviation estimate and variability of that estimate.

In practice (in this dataset), the relative error in the estimation is slight. We could reduce the Monte Carlo error by taking more samples (i.e., increasing \( N \)), or by agglomerating two or more of the independent runs. However, this only reduces the error slowly, and in this case is probably not warranted.

6. Discussion

6.1. Assessment of methodology

We observe immediately that our numerical approximation to error estimates does not give us any more information than a formal error propagation analysis of the gridding algorithm would provide. However, we also observe that such a formal analysis would be very difficult (or impossible) to carry out under the circumstances, and hence
that the approximation here is justified. Given the assumptions inherent in the method, our results show that the technique generates results in line with intuition and common sense. However, the assumptions that we have made do limit the applicability of the method as it stands.

We have assumed that only random errors need to be modeled. This is implicit in the use of normally distributed random variates to generate the pseudo-independent data sets used in the Monte Carlo estimation. Although we believe that this is the most likely error mode for the data once it is cleaned and entered into the database, it is not the only error mode possible. In particular, we could encounter a survey line that is very badly navigated, for example due to long-term drift in inertial navigation systems. In this case, a systematic bias has been introduced into the data and our assumption that the recorded data is a valid estimate of the population mean is violated. However, it is impossible to identify these types of errors without further cross checks (such as the repeated observations which highlighted the problem with the pseudo-seamounts in this work), and we consider the identification of such errors beyond the scope of our statistical error assessment. A cross track analysis and correction of data sets that are offset due to poor navigation should be performed prior to the gridding; there are several possible approaches to this problem (e.g., [Caress and Chayes, 2000]).

We have assumed that random errors are distributed as Gaussian random variates, and that the variance is the same for all points in each component data set. Lacking a more formal error assessment of the errors associated with each data set, this is probably unavoidable. Indeed, with historical data sets, it may be difficult to determine
reasonable errors with any certainty since the meta-data describing the instrumentation package may not be adequate (or extant). As we have done here, future work will have to assign subjective (but informed) error bounds, making sure to err on the side of caution and assess slightly exaggerated errors. Since the principle goal of the research is to inform users about how much trust they can have in the generated grid data, too high an error assessment is not as important as underestimation. The assignment of the initial errors bounds to the data sets, which is the backbone for “real use” of our modeling approach, will be one subject for our further research. During the past century, positioning and depth measurements have gone from lead lines to multibeam sonars and from sextants to differential GPS. Each advance in technology is associated with a tremendous improvement in the accuracy and resolution. Considering the technology used for bathymetric data collection, a thorough classification in terms random errors may be done. However, there are also other, more difficult to discern factors, that may add to the random errors of the collected data, including weather conditions and human.

Finally, we have assumed that the IBCAO gridding algorithm is the right way to compute the gridded product. In terms of comparing the error assessment with the original data, this is unavoidable, but it is certainly a limiting factor to the application of the methodology to other data sets. However, nothing in the theoretical foundation of the technique requires that the IBCAO gridding model is used, and substitutions may be made on a black-box basis. The principal shortcoming of the methodology outlined is that it is difficult to discriminate between low errors arising from accurate data, and low errors arising from gridding artifacts where there is low data density. This
is typically associated with areas defined principally through contour data, which also causes significant terracing in the bathymetry. Although we can avoid this problem by removing from consideration areas with low data density, it would be better to avoid contours whenever possible and instead use the original source data. The need to rely on contour-only derived data is probably most acute in the Arctic where the strategic sensitivity of data sources has resulted in a general reluctance to distribute original sounding data.

Related to the problem of density is one of resolution, which we have not really considered in the present work. That is, at what grid resolution should we work, and should this be the same across the entire grid? Determination of a reasonable resolution at which to work is essential, particularly to ensure that we are making interpretations correctly when we look at error estimates. By preference, we should work at a resolution determined by the spatial frequencies associated with the bathymetry in question, and have data-driven grid cells. In practice, the spatial frequency spectrum of the bathymetry is not known a priori, and we have to resort to some approximation, such as reducing the resolution and repeating the error estimation until our standard errors stabilize. We would then be able to use the stabilization resolution as a local working resolution for further work. Assessment of stability of estimation also informs us about the fineness of interpretation that is reasonable for the data in question - and this is not necessarily stable across the grid when we have many different vintages of data sets. We consider that the issue of gridding resolution is the most significant outstanding issue with the current methodology, and are actively pursuing improvements in this area.
One of our initial goals was to create a final output error assessment in the same format as the bathymetry grid. The reason for this is to prevent over-interpretation of the bathymetry by facilitating the visualization of the errors associated with the gridded data set in modern GIS tools and other grid analysis software. The standard deviation error associated with each depth in a grid cell may be represented as an attribute to the bathymetric depth. In Figure 6 the form of the 3D surface visualizes the morphology of the sea bottom while the color-coding represents the standard deviation error (% depth) in each grid cell. In this way the relationship between the uncertainty of the bathymetry and the regional morphology can clearly be seen. Such an approach allows for the intuitive interpretation of potential sources of uncertainty. We have limited our error modeling to bathymetry, however, this approach may be applicable to other geophysical data sets such as gravity and magnetics, which exist as gridded data compiled using similar approaches as we used for the IBCAO bathymetry grid [Verhoef et al., 1996; Laxon and McAdoo, 1994].

7. Conclusions

We conclude that the prediction of accuracies in final gridded bathymetric products is possible using a Monte Carlo approach, with the added advantage that the output error assessment is in the same form as the original gridded product. Our predictions on the test dataset agree with common-sense, and provide important caveats on the use of gridded data products. Our experiments clearly show areas of high certainty associated with the more accurate data in the dataset, and regions which are less reliable, typically
associated with contour-based data. We also discovered areas of unexpectedly high uncertainty, which we subsequently found to be associated with erroneous data from a single track line.

We have outlined a methodology for the assessment of errors in gridded bathymetric datasets. Although the method has been applied to the problem of sparse datasets requiring interpolation, we believe that it can also be applied to other types of data. However, there may be alternatives when data is denser, for example when dealing with full rate multibeam echo-sounder data. The methodology relies on essentially subjective assessment of the errors associated with the datasets used, even if survey meta-data is available. It would be preferable to assign errors based on a more formal model of the data gathering equipment, but these are often not available. In many cases, the subjective but conservative estimate of an experienced observer may be more useful than the alternatives.

We have essentially neglected the issue of resolution in this analysis, although it is probably a significant constraint, especially in sparse data. The resolution at which data is analyzed should be data-driven, rather than specified a priori. However, implementation and theoretical difficulties involved in this are not trivial, and are currently the subject of further research.

Acknowledgments.

We thank Magnus Mörth for coming up with the idea that the Monte Carlo method might be useful for our error modeling of bathymetric data.
References


Head Department of Navigation and Oceanography, *Bottom Relief of the Arctic Ocean; map*
1:5,000,000., Head Department of Navigation and Oceanography, All Russian Research Institute of Sciences, St. Petersburg, Russia, 1999.


Macnab, R., and M. Jakobsson, Something old, something new: compiling historic and contemporary data to construct regional bathymetric maps, with the arctic ocean as a case study, Int. Hydro. Review, 1, 2–16, 2000.


Martin Jakobsson, Brian Calder and Larry Mayer, Center for Coastal and Ocean Mapping, University of New Hampshire, Durham, NH 03824. (e-mail: martin.jakobsson@unh.edu, brc@ccom.unh.edu, larry.mayer@unh.edu)

Received ____________________

This manuscript was prepared with AGU's \LaTeX macros v5, with the extension package 'AGU++' by P. W. Daly, version 1.6b from 1999/08/19.
Figure Captions

Plate 1. A color-coded shaded relief portraying bathymetry and topography of the Arctic region created from the IBCAO grid model with a nominal resolution at 2500m. The area subjected to our error modeling experiment is indicated by a bold rectangle. Projection: Polar Stereographic with true scale at 75°N. Datum: WGS-84.

Plate 2. Data from the IBCAO construction database used for the error modeling. This includes all data that falls within the bounds indicated in Figure 1 and covers almost all of the component data sets used in the entire IBCAO grid compilation. Projection parameters as Figure 1. The key to the used color-coding of the source data is found in table 1.

Plate 3. Histogram of the grid node depths in the subsection of the IBCAO model showed in Figure 1. The spikes are caused by a bias towards the depths as represented by contours in the source data.

Plate 4a. Estimated standard deviation of gridded depth based on $N = 100$ Monte Carlo simulation runs. Standard deviation is depth in meters.

Plate 4b. Estimated standard deviation of gridded depth based on $N = 100$ Monte Carlo simulation runs. Standard deviation is shown as a percentage of the depth estimated on the unperturbed grid. The percentage grid gives a better feel for the errors involved, and is the preferred grid for interpretation.
Plate 5. Standard error grid (percentage of estimated depth) with empty grid cells removed. Interpretation of standard error in the grid where there is no data is essentially a function of the interpolation surface used in gridding, rather than actual errors caused by variability of data. We remove empty cells to avoid over-interpretation in sparse data.

Plate 6. 3D-image, created using the software Fledermaus, showing the standard deviation as a percentage of the depth estimated on the unperturbed grid and draped on the IBCAO bathymetry. The error estimate shows clearly that significant error is associated with the indicated seamounts later revealed not to exist by a recent survey with R/V Polarstern.

Plate 7. Estimate of Monte Carlo error associated with the standard deviation estimates. This grid is estimated as the sampler error associated with the standard deviation estimates, computed over $M = 20$ example standard deviation grids.
### Tables

**Table 1.** Classification of the source data shown in Figure 2 and initial assignment of standard deviation of errors at 95% confidence interval.

<table>
<thead>
<tr>
<th>Source data</th>
<th>Horizontal error (m)</th>
<th>Vertical error (% depth)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Digitized contours</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contours drawn during the IBCAO project</td>
<td>12000</td>
<td>5</td>
</tr>
<tr>
<td>Bathymetry of the Franz Josef Land Area [Matishev et al., 1995]</td>
<td>12000</td>
<td>5</td>
</tr>
<tr>
<td>Bathymetry of the Barents and Kara Seas [Cherkis et al., 1991]</td>
<td>12000</td>
<td>5</td>
</tr>
<tr>
<td>Bottom relief of the Arctic Ocean [Head Department of Navigation and Oceanography, 1999]</td>
<td>12000</td>
<td>5</td>
</tr>
<tr>
<td><strong>Soundings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swedish icebreaker Oden, 1991 and 1996</td>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>Swedish icebreaker Ymer, 1980</td>
<td>1852</td>
<td>5</td>
</tr>
<tr>
<td>Data collected during SCICEX by USS Hawkbill, 1999</td>
<td>5000</td>
<td>5</td>
</tr>
<tr>
<td>Data from Norwegian sources</td>
<td>200</td>
<td>2</td>
</tr>
<tr>
<td>Soundings obtained from the US National Geophysical Data Center (NGDC)</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td><strong>Land and support data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>World Vector Shoreline</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Control contours</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GTOPO30</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>