Error estimation of bathymetric grid models derived from historic and contemporary datasets

Martin Jakobsson, Brian Calder, Larry Mayer and Andy Armstrong
Center for Coastal and Ocean Mapping,
University of New Hampshire
Durham, NH 03824

1 Introduction

The past century has seen remarkable advances in technologies associated with positioning and the measurement of depth. Lead lines have given way to single beam echo sounders, which in turn are being replaced by multibeam sonars and other means of remotely and rapidly collecting dense bathymetric datasets. Sextants were replaced by radio navigation, then transit satellite, GPS and now differential GPS. With each new advance comes tremendous improvement in the accuracy and resolution of the data we collect. Given these changes and given the vastness of the ocean areas we must map, the charts we produce are mainly compilations of multiple data sets collected over many years and representing a range of technologies. Yet despite our knowledge that the accuracy of the various technologies differs, our compilations have traditionally treated each sounding with equal weight. We address these issues in the context of generating regularly spaced grids containing bathymetric values. Gridded products are required for a number of earth sciences studies and 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.

In this paper 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 [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 or wrongly defined.

Given the poor situation with the regional Arctic bathymetry maps, the International Bathymetric Chart of the Arctic Ocean 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 et al., 1999], by the US Naval Research Laboratory [Perry et al., 1986; Cherkis et al., 1991; Matishov et al., 1995], and by other agencies, were digitized using heads up digitizing techniques 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 [US 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 velocity using Carter's tables, or CTD profiles where available. Subsequently, 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 data 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 x 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 [1990], as implemented in the GMT package [Wessel & 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 x 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 [2000] and Macnab & Jakobsson [2000].

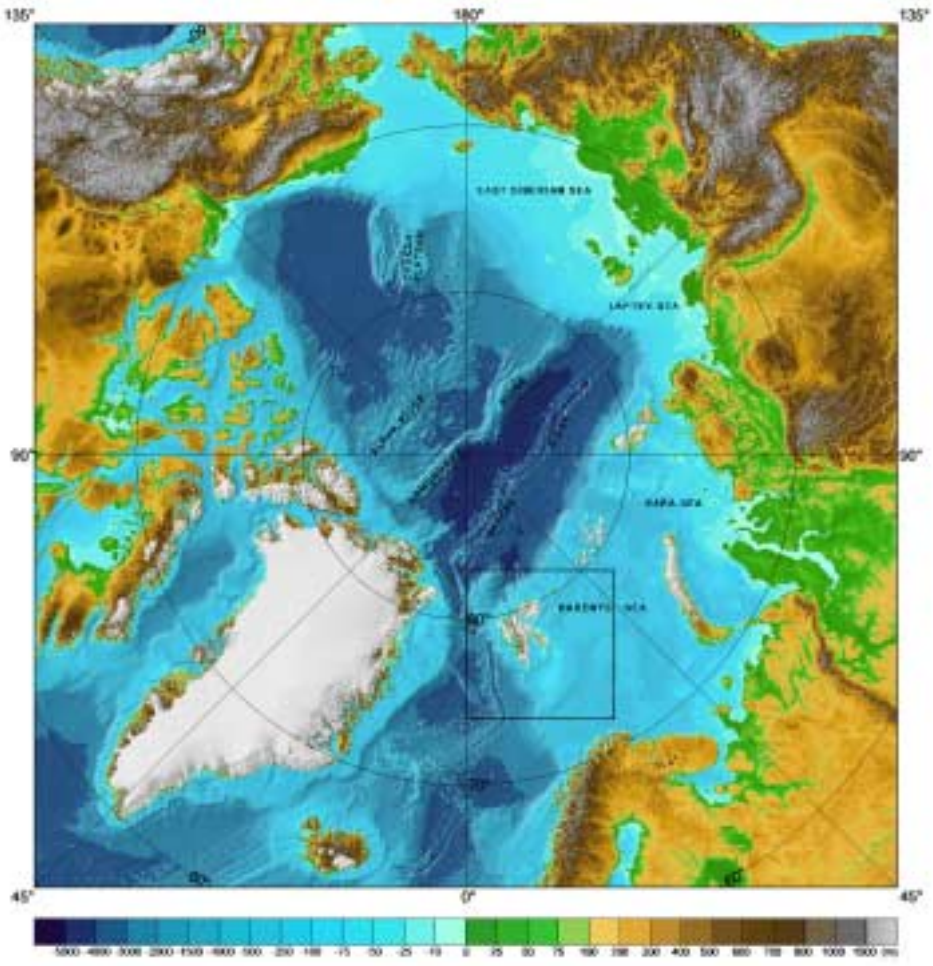

Figure 1. A color-coded shaded relief portraying bathymetry and topography of the Arctic region created from the IBCAO grid model with a nominal resolution at 2500m [Jakobsson et al., 2000]. The area subjected to our error modeling experiment is indicated by a bold rectangle. Projection: Polar Stereographic with true scale at 75°N. Datum: WGS-84.

2.2 Experiment 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* (Figure 2 and Table 1). There is also a significant depth range to be considered, due to the relatively shallow areas of the Barents Sea around Svalbard and the contrasting deep North Atlantic in the western part of the area where also 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, and if not previously adjusted

the depths have been corrected for sound velocity 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).

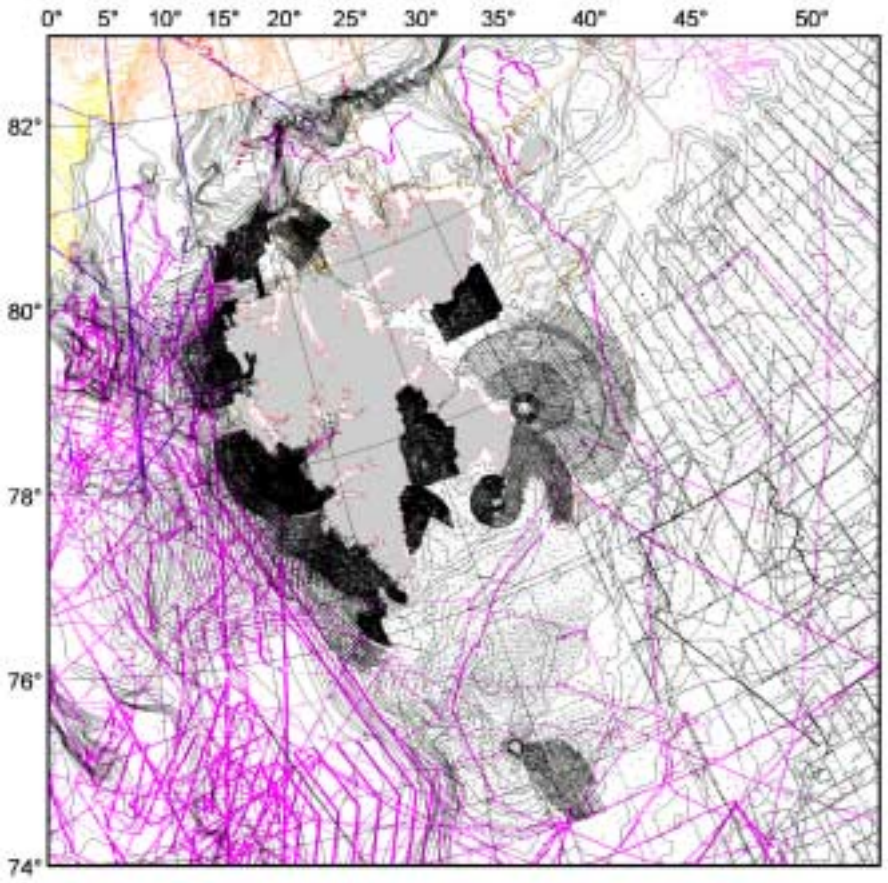

Figure 2. Data from the IBCAO construction database used for the error modeling. This includes all data that falls within the bounds indicated in figure 1 and covers almost all of the component data sets used in the entire IBCAO grid compilation. Projection parameters as figure 1. The key to the used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

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 & 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., 1995]. 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 $g(x)$, $\mathsf{E}[g(x)]$. Given a set of *N* samples from the probability density, then, we can approximate this expectation by the sample mean,

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

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:

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

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 rather more weakly justifiable, since we may have some systematic bias in navigation (e.g., a mis-navigated submarine track 10km 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 velocity 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, they may possible 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 arbitrarily 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 the *Oden* icebreaker data 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, which requires a large amount of time-consuming "data detective" work in order to find meta-data for many of the older data sets. It is then possible to assign errors that are more closely related to the "true errors".

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

Source data	**xy σ error (m)**	**z σ error (% depth)**
Digitized contours		
Contours drawn during the IBCAO project (Yellow)	12000	5
Bathymetry of the Franz Josef Land Area [Matishov et al., 1995] (Magenta)	12000	5
Bathymetry of the Barents and Kara Seas [Cherkis et al., 1995] (Black)	12000	5
Bottom relief of the Arctic Ocean [Head Department of Navigation and Oceanography et al., 1999] (Orange)	12000	5
Soundings		
Swedish icebreaker Oden, 1991 and 1996 (Magenta)	100	5
Swedish icebreaker Ymer, 1980 (Red)	1852	5
US and British Royal Navies submarines, 1958-1988 (Lila)	10000	5
Data collected during SCICEX by USS Hawkbill, 1999 (Lila)	5000	5
Data from Norwegian sources (Black)	200	2
Soundings obtained from the US National Geophysical Data Center (NGDC) (Magenta)	1000	5
Land and support data		
World Vector Shoreline (Black)	0	0
Control contours (Red)	0	0
GTOPO30 (Gray)	0	0

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 $\mathbf{A}_{nm}$ or $\mathbf{A}_{m}$ 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 $\mathbf{A} = \mathsf{F}(\mathbf{B})$ for some operator $\mathsf{F}(\cdot)$ means $A_{ij} = \mathsf{F}(B_{ij})\ \forall i, j$ on the domain). Sets of variables are indicated by sans serif letters, e.g., $\mathsf{Y} = \{\mathbf{Y}(1), \ldots, \mathbf{Y}(k)\}$; when appropriate, we refer to components of a set indexed over N (the set of natural numbers) with an essentially arbitrary, but fixed, indexing scheme.

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

$$\mathsf{X}_{nm} = \{\mathbf{X}_{nm}(1), \ldots, \mathbf{X}_{nm}(s)\}$$

$$\mathbf{X}_{nm}(i) = \mathbf{X}(i) + \mathsf{E}(i)$$

$$\mathsf{E}(i) = \left[\mathbf{e}_1(i), \cdots, \mathbf{e}_{J(i)}(i)\right] \text{ where } J(i) = \left|\mathbf{X}(i)\right|$$

$$\mathbf{e}_j(i) \sim \mathsf{M}\left(\mathbf{0}, \mathrm{diag}\left(1, 1, z_j^2(i)\right)\mathbf{e}(i)\right)$$

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 $\mathbf{B}_m = \mathsf{E}\left[\mathbf{B}_{nm}\right]$ and the standard deviation $\boldsymbol{\varepsilon}_m^2 = \mathsf{E}\left[\left(\mathbf{B}_{nm} - \mathbf{B}_m\right)^2\right]$ 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 $\boldsymbol{\varepsilon}_m$ 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 $\bar{\boldsymbol{\varepsilon}} = \mathsf{E}\left[\boldsymbol{\varepsilon}_m\right]$ and standard error estimate $\boldsymbol{\varepsilon}^2 = \mathsf{E}\left[\left(\boldsymbol{\varepsilon}_m - \bar{\boldsymbol{\varepsilon}}\right)^2\right]$ 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 and Discussion

5.1 Standard Deviations in Gridding

The standard deviation grid for a single run of the algorithm (i.e., $\boldsymbol{\varepsilon}_m$) is shown in Figure 3. 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 3(B).

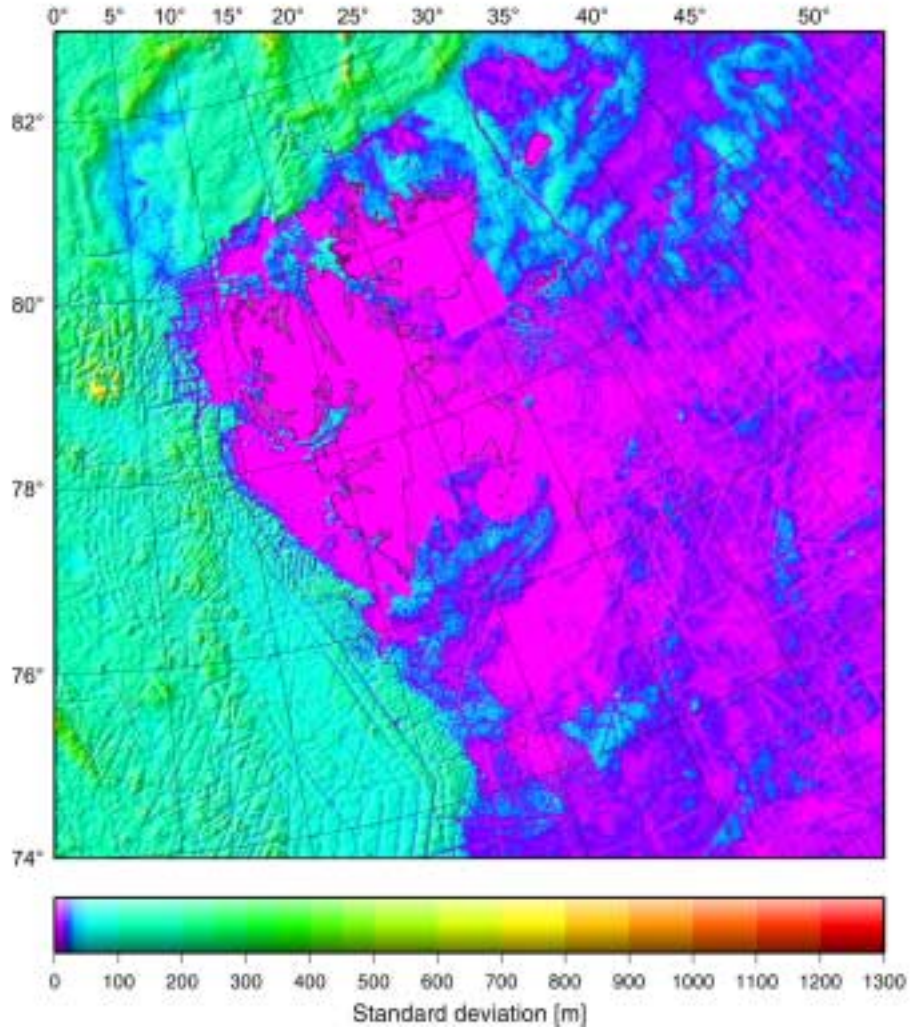

Figure 3A: Estimated standard deviation of gridded depth based on *N*=100 Monte Carlo simulation runs. Standard deviation is depth in meters.

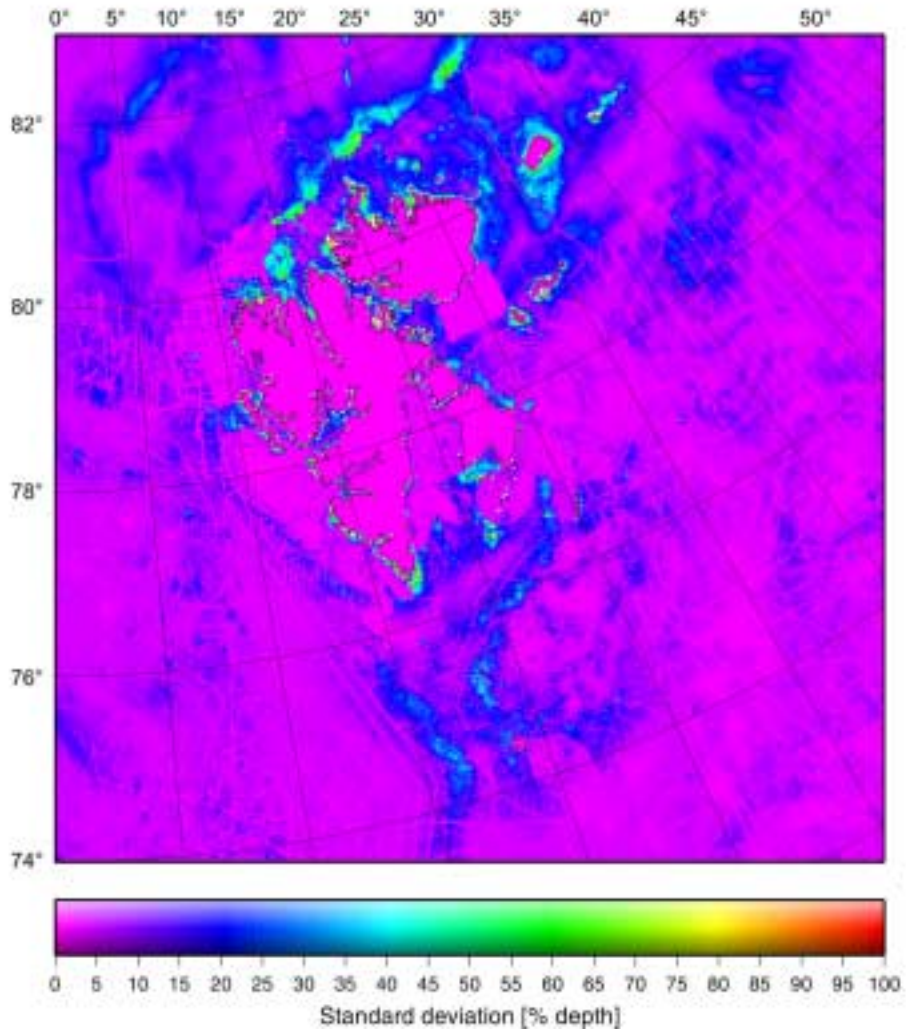

Figure 3B. Estimated standard deviation of gridded depth based on *N*=100 Monte Carlo simulation runs. Standard deviation is shown as a percentage of the depth estimated on the unperturbed grid. The percentage grid gives a better feel for the errors involved, and is the preferred grid for interpretation.

On first examination, the results appear to agree with intuition. In regions where there have been rigorous hydrographic surveys (e.g., 77ºN 22º 30'E), the estimated error is significantly lower than regions where only a single trackline is used to constrain the grid (e.g., at 82ºN 5ºE). 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, we see some anomalies. For example, the region near 79º 30'N 37ºE 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 x 7500 m that contain no soundings to prevent too many spurious

removals. The resulting reduced grid is shown in Figure 4, and by comparison with Figure 3(B), 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.