Bathymetry

Matthew Poti[1,2], Brian Kinlan[1,2,3], and Charles Menza[1]

2.1 SUMMARY

A new bathymetric model with spatially-explicit uncertainty estimates was developed for the New York study area (Figure 1.1). The model builds on previous predictive bathymetric modeling approaches in the region (e.g., Calder, 2006), provides a continuous gridded bathymetric surface for the study area, and allows users to view and explore spatial variation in the vertical accuracy of depth predictions. The spatial resolution of the model is identical to the National Oceanic and Atmospheric Administration's (NOAA) Coastal Relief Model (CRM; horizontal resolution approximately 83.8 m) in the study area and was built from the same database of hydrographic survey points. Unlike the CRM, the new geostatistical model provides estimates of prediction certainty, which can be used to prioritize locations for new bathymetric surveys and better understand the reliability of depth predictions and derived spatial layers (e.g., benthic habitats, positions of depth contours).

2.2 BACKGROUND

Bathymetry (also called seafloor topography) is an important base environmental layer for spatial planning since it influences both planning of human activities (e.g., construction, shipping) and many physical, chemical and ecological processes. For instance, reliable bathymetric information can simultaneously improve habitat conservation and energy development by supporting the identification of:

- Unique or vulnerable benthic habitats
- Distributions of rare or endangered species
- Efficient corridors for transmission lines
- Suitable sites for turbine platforms, and
- Potential construction hazards

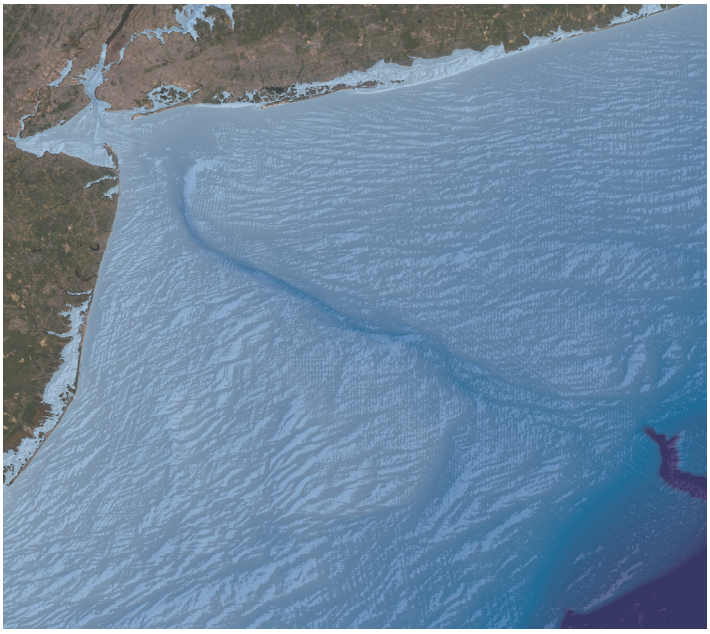

Image 2.1. An example of a bathymetric surface in the New York Bight, showing the change in depth with distance from shore and the complexity of the seafloor across the shelf. The Hudson Shelf Valley is prominently visible in the center of the model as the area of darker blue extending from New York harbor (top left) towards the shelf edge (bottom right). Coastal managers and engineers use bathymetric surfaces to assess shipping lanes, identify fish habitats, lay undersea cables and find sand and gravel resources. The bathymetric surface shown here is the the NOAA Coastal Relief Model (CRM), draped over a derived hillshade layer to highlight bathymetric variation. Terrestrial imagery is the ArcGIS Online World User Imagery layer (ESRI Online).

Bathymetry can be measured by a range of instruments, which determine the precision, spatial resolution, extent and cost of bathymetric information and nautical charts. Until the latter half of the 20th century, lead lines dropped from a ship were used to estimate depths (Calder, 2006) and were compiled on charts to give a coarse-scale representation of the seafloor and identify navigation hazards. Lead lines were eventually replaced by more accurate vertical beam echosounders (VBES) and subsequently by multibeam echosounders (MBES). Modern MBES can collect millions of precise soundings efficiently and quickly, making possible high-resolution bathymetric maps that reveal fine-scale features of the seafloor (Calder, 2006). Horizontal positioning technologies have also advanced over the years from sextant-based navigation to modern GPS.

When combined with backscatter information and validation samples, MBES data offers an unprecedented view of the composition and morphology of the seafloor at multiple spatial scales (Kostylev et al., 2001; Gardner et al., 2003). The States of Oregon and California recently collected new data to take advantage of insights

[1] Center for Coastal Monitoring and Assessment, National Centers for Coastal Ocean Science, National Ocean Service, National Oceanic and Atmospheric Administration
[2] Consolidated Safety Services, Inc., under NOAA Contract No. DG133C07NC0616
[3] Corresponding author: Brian.Kinlan@noaa.gov

provided by MBES and these data have been integrated into state spatial planning initiatives, specifically the Oregon Territorial Sea Plan and California Marine Life Protection Act Initiative. On the East Coast, fewer states have comprehensive MBES coverage; this may be a reflection of the increased costs involved in surveying comparatively wide, shallow continental shelves.

About 20% of the New York planning area is covered by MBES surveys (Figure 2.1) which have been collected by the United States Geological Survey (USGS), NOAA and the Woods Hole Oceanographic Institution (Schwab et al., 1997a and 1997b; Butman et al., 1998; Goff et al., 1999; Butman et al., 2006). The corresponding data have helped researchers map benthic habitats and identify physical features on the seafloor within the footprints of surveys.

Unfortunately, the incomplete distribution of multibeam surveys limits their usefulness for understanding the relative distribution of habitats, features, processes and species over the entire planning area, a critical component of integrated marine spatial planning. The U.S. Coastal Relief Model developed by the National Geophysical Data Center (NGDC; http://www.ngdc.noaa.gov/mgg/coastal/model.html) offers a 3-arc second continuous bathymetric model that covers the majority of the study area (including all of the continental shelf and slope). The CRM is derived from the largest single compilation of bathymetric soundings for US coastal waters and has been used effectively to inform spatial planning on the East Coast as part of the Massachusetts

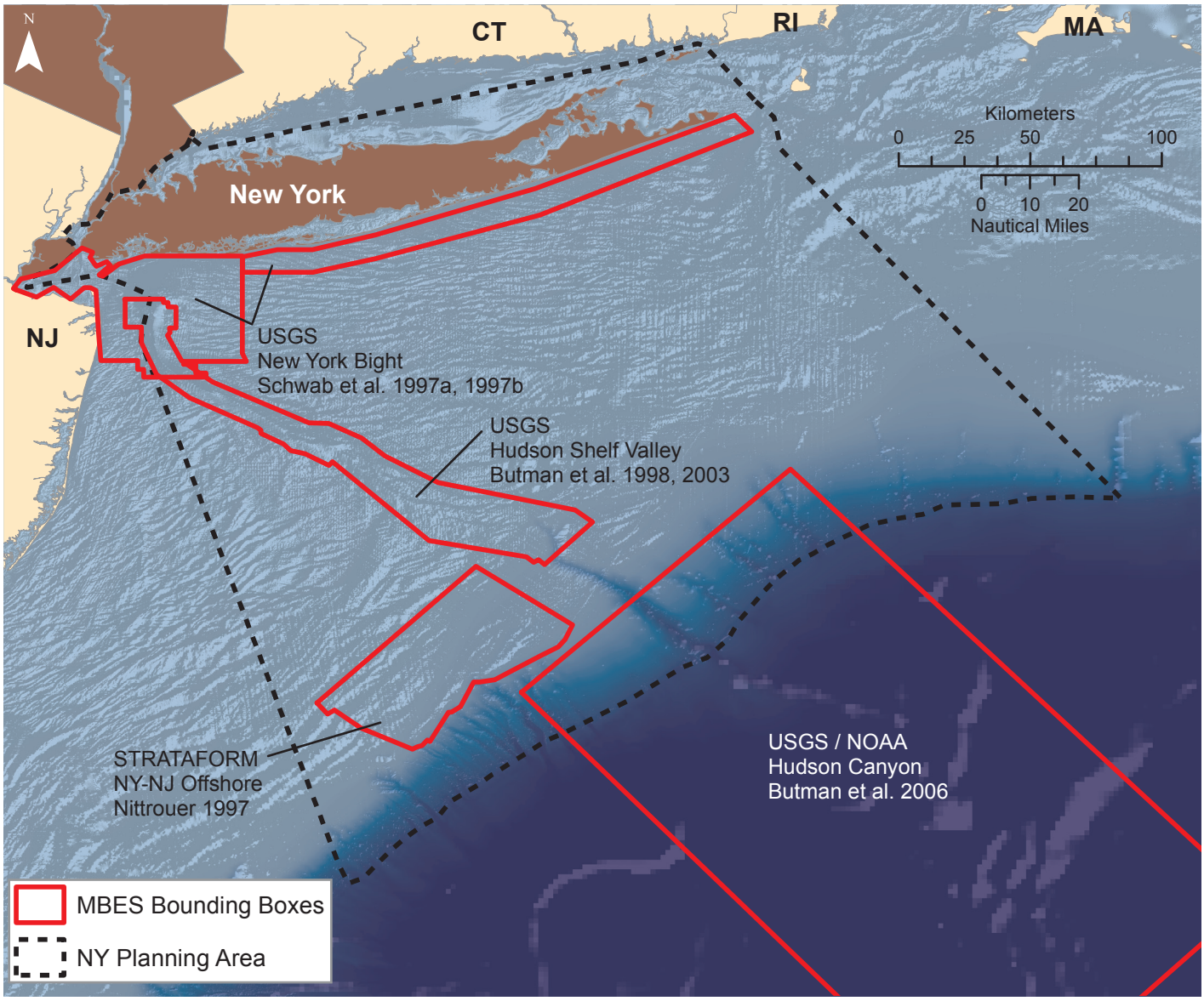

Figure 2.1. Spatial extent of selected multibeam and sidescan sonar surveys in the study area. Survey boundaries are overlaid on bathymetry data from the Coastal Relief Model blended with the ETOPO1 Global Relief Model (Amante and Eakins 2009).

Ocean Plan and Rhode Island Special Area Management Plan. Although coarser than multibeam data, the ~84 m horizontal resolution of the CRM is still sufficient to resolve general features of interest for marine spatial planning (e.g., canyons, ridges, sand waves, bathymetric contours).

The portion of the CRM that overlaps the New York planning area was produced in 1999 and is a compilation of historical hydrographic surveys, collected using VBES and MBES from various data sources, including NOAA, USGS, the U.S. Army Corps of Engineers, and various academic institutions. Although compiled surveys are brought together under a common spatial framework, they possess different spatial footprints and resolutions, and were collected using different instruments. Newer surveys commonly overlap, adjoin and supersede older surveys.

Generally, the CRM is used in resource management applications assuming the depth measurements and predictions to be accurate, but significant uncertainty in model depth estimates arises from measurement error in hydrographic surveys, methods used to interpolate between survey points and data processing (discussed in detail in Calder, 2006 and references therein). These errors are variable over the study area (Figure 2.2) due to various factors, including survey age, processing guidelines, and distances amongst soundings. Since bathymetric errors in the CRM are not quantified, users cannot know whether depth predictions at a given location are likely to deviate from the true value by a few centimeters or hundreds of meters. Disregarding uncertainty might be acceptable for some analyses conducted at coarse spatial resolutions, but is problematic for finer resolution analyses and when precise measurements are needed. Knowing where bathymetric predictions are precise and where they are not provides managers with information to define and manage risk associated with decisions relying on bathymetry or derived products (e.g., defining benthic habitats, estimating construction costs, placing shipping lanes).

2.3 METHODS

2.3.1 General Modeling Approach

A geostatistical modeling approach was used to predict a continuous, gridded bathymetric surface from scattered sounding points and to generate corresponding spatially-explicit uncertainty estimates. Geostatistical methods are based on the premise that neighboring samples are more similar than samples farther away (Tobler, 1970), a phenomenon known as spatial autocorrelation. Spatial autocorrelation can be detected, quantified and modeled by semivariogram analysis, and used to make predictions at locations that have not been measured

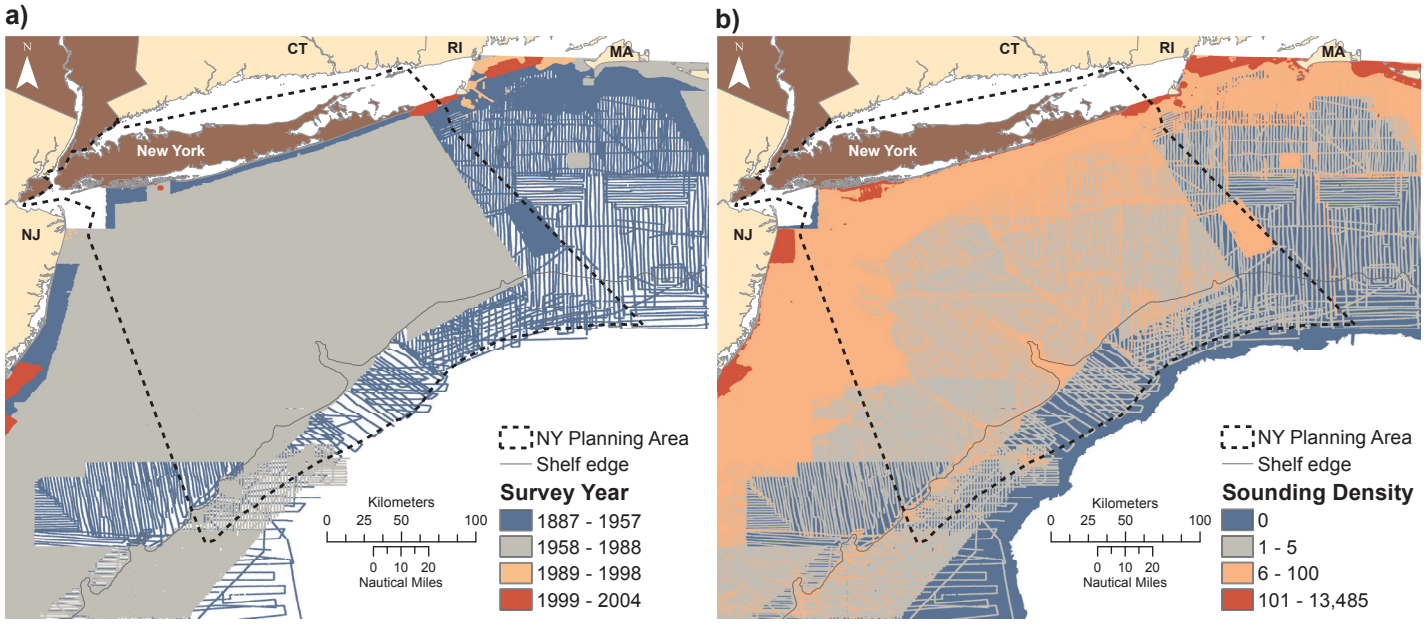

Figure 2.2. (a) Most recent survey year for soundings within 1 km rectangular neighborhoods. Survey year classes generally correspond with the evolution of horizontal positioning and vertical sounding technologies. More recent soundings tend to be more precise. (b) The number of bathymetric soundings per square kilometer. The shelf edge corresponds to the 200 m depth contour.

(Cressie, 1993; Chiles and Delfiner, 1999). In addition, the same spatial model used to develop predictions can be used to model uncertainty (i.e., expected precision) of predictions.

The geostatistical modeling approach used here follows Cressie (1993), where estimates of depth for a given location, Z(x,y), are modeled as a linear combination (sum) of components representing a deterministic mean trend, μ(x,y), a spatially structured random process, δ(x,y), and non-spatially structured error, ε.

$$Z(x,y) = \mu(x,y) + \delta(x,y) + \varepsilon \qquad \text{(Equation 2.1)}$$

Equation 2.1 defines the workflow used to arrive at Z(x,y). The deterministic mean trend and spatially structured random process with error term are modeled separately and then combined by summation (see Figure 2.3 for schematic representation of work flow). The deterministic mean trend, μ(x,y), is estimated using a suitable smoothing function and residuals of original data from this smoothing function are computed at the data positions by subtracting the trend prediction from the observed data value. The spatially structured random process, δ(x,y), and error term, ε, are then estimated by fitting a suitable semivariogram model to the empirical semivariogram of the residuals. The error term is defined by the semivariogram nugget and represents error that is not spatially correlated, which includes both measurement error and variability occurring at spatial scales shorter than the sampling resolution (Cressie, 1993).

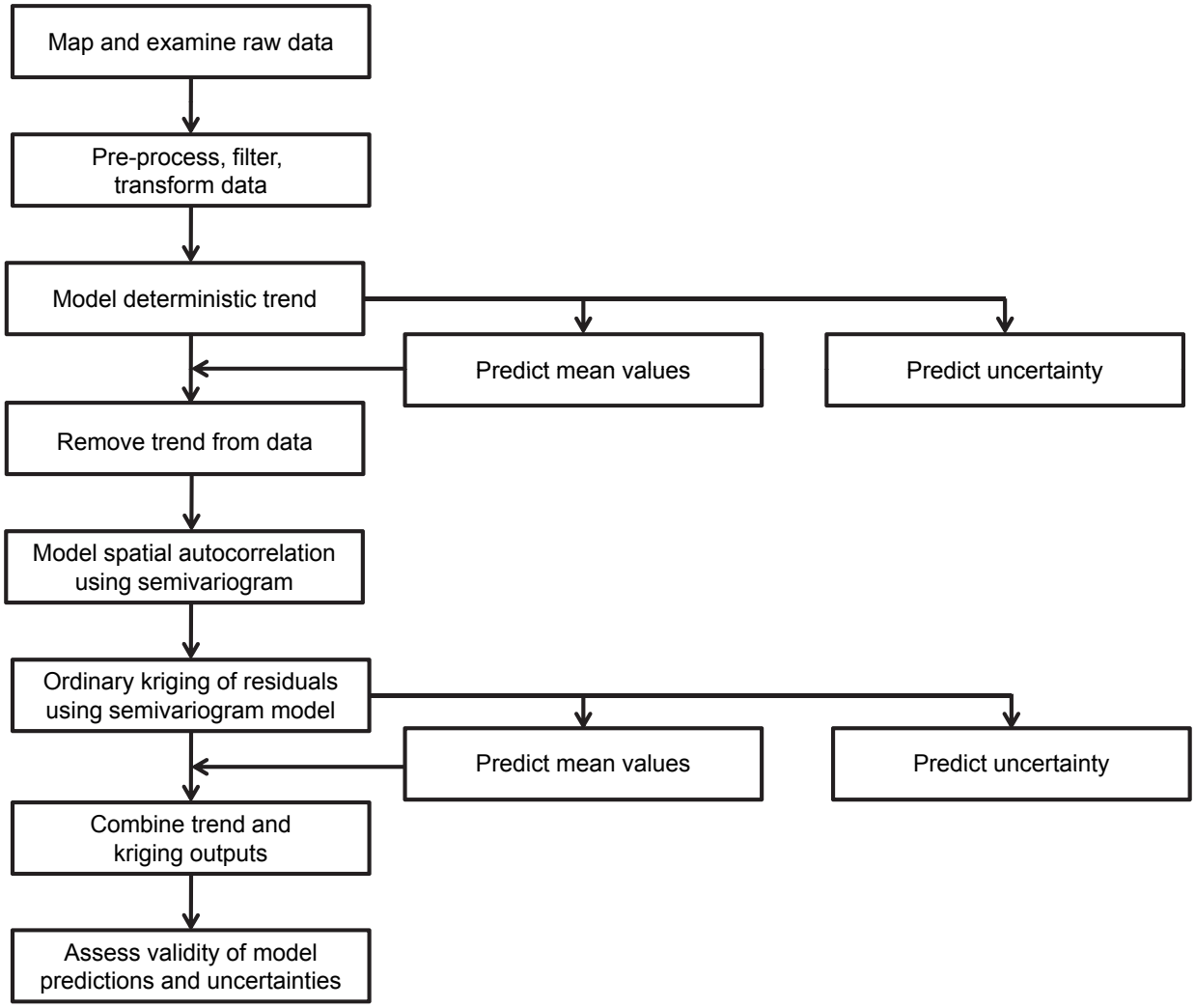

Figure 2.3. General workflow describing the geostatistical approach used to develop the predictive model for bathymetry (see sections 2.3.3 and 2.3.4 for a more detailed description of the methods).

Geostatistical models involve a number of statistical assumptions (for detailed discussions see Cressie, 1993; Chiles and Delfiner, 1999). The accuracy of model predictions and uncertainty bounds depends to varying degrees on these assumptions being met. Thus, an important part of any geostatistical analysis is model validation, which is usually done by cross-validation, a process in which some data are left out for purposes of model fitting and model predictions are tested against those data. Model predictions can also be tested against high-precision "ground-truth" datasets where such datasets are available. We use both methods to validate model predictions in this chapter (see Section 2.3.4).

2.3.2 Data Acquisition and Preparation

Raw Sounding Data

To develop a new geostatistical bathymetric model for the NY Bight, all available National Ocean Science (NOS) Hydrographic Survey Data overlapping the study area were downloaded from the National Geophysical Data Center Hydrographic Survey Database (http://www.ngdc.noaa.gov/mgg/bathymetry/hydro.html) on April 21, 2011. Survey measurements and metadata were extracted from raw HYD93 formatted files and exported into plain ASCII text tabular files using custom parsing scripts. Depths are represented as positive numbers with increasing depth below sea level defined by the Mean Lower Low Water (MLLW) vertical datum. Sounding data were merged with survey metadata, so that information detailing when and how each sounding was collected was retained. Sounding locations, originally in decimal degrees (NAD83 datum), were projected into a Universal Transverse Mercator projection (UTM 18N), since subsequent processing requires measurement of point-to-point distances in a Cartesian coordinate system. UTM 18N has its central meridian at 75ºW and thus allows calculation of distances in our study area with negligible distortion relative to grid resolution (83.8 m).

Hydrographic soundings in the study area come from a multitude of surveys distributed between 1887 and 2004. Surveys used an assortment of positioning and sounding technologies, resulting in a patchwork of overlapping soundings collected with variable sample spacing and different precisions (Figure 2.2). In addition, survey data was processed using varying methods which created varying post-processing errors (see Calder, 2006 for a full discussion). These errors can propagate to the final model creating distortions that do not correspond to changes on the seafloor. While steps have been taken to partially correct for and reduce the impact of these data quality issues, it is important to understand that they cannot be entirely eliminated. No bathymetric model based on historical hydrographic sounding data will be completely free from such considerations.

In general, the vast majority of soundings were retained to maximize data density. However, some soundings were corrected or eliminated prior to modeling. First, based on Calder (2006), we applied a +1.48 m correction to all soundings collected by lead line to correct for the observed bias of lead line soundings compared to multibeam sonar measurements. Second, lead line and VBES surveys were identified that showed evidence of quantization due to rounding to the nearest whole fathom (resulting in an error of up to 1.8 m). Data from these surveys were eliminated when they were located within the footprint of more accurate surveys (i.e., surveys that did not exhibit quantization). Survey footprints were hand digitized at 1:50,000 in ArcGIS 10 (ESRI, 2011). Different types of rounding and conversions created surveys with varying degrees of quantization, but only those surveys with the most severe fathom-rounding quantization of 1.8 m were eliminated from analysis, and those only when more recent information was present. Other types of quantization are expected to result in errors less than 1 m. Other sources of error in raw soundings, including un-accounted for changes in vertical and tidal datums, are expected to be small (on the order of 10's of centimeters) and are discussed in detail in Calder (2006).

Depth Stratification

The resulting hydrographic sounding database was divided into four strata based on depth thresholds (Table 2.1). Thresholds were chosen based on depth ranges that correspond to different maximum uncertainty specifications under International Hydrographic Organization (IHO) standards (S.44 Order 1 and 2, IHO 1998) and on coarse-scale changes in geomorphology (e.g., the continental shelf break). Neighboring strata overlap slightly to facilitate merging of outputs from individual strata into a continuous surface.

Table 2.1. Depths used to stratify hydrographic soundings and the corresponding number of soundings within each stratum.

DEPTH STRATUM	NUMBER OF SOUNDINGS (EXCLUDING OVERLAP)	PERCENTAGE OF SOUNDINGS	ADDITIONAL SOUNDINGS FROM OVERLAP	PERCENTAGE OF SOUNDINGS FROM OVERLAP
0-30 m	2,077,055	83.9%	0	0%
30-100 m	337,238	13.6%	176,843	34.4%
100-200 m	21,932	0.9%	9,931	31.2%
200-2,000 m	40,196	1.6%	3,517	8.0%
Total	2,476,421	100%	190,291	7.1%

Transformation

Prior to statistical modeling, depth values were transformed using the following logarithmic function to normalize error distributions:

$$Z_{transform} = \log(Z*b+a) \qquad \text{(Equation 2.2)}$$

where Z is depth in meters (with positive numbers representing depth below sea surface) and the transformation parameters a and b are taken from the appropriate error model for each depth stratum identified by IHO standards (a = 0.5 m, b = 0.013 when Z < 100 m [IHO S.44 Order 1] and a = 1.0 m, b = 0.023 when Z >= 100 m [IHO S.44 Order 2]) (IHO, 1998). This transformation was based on a standard bathymetric error model formulation (IHO, 1998; Calder, 2006), and improves homogeneity of conditional error variances within local regression and kriging neighborhoods, a desirable statistical property.

2.3.3 Development of the Bathymetric Model

The deterministic mean bathymetric trend surface was estimated using LOESS, a semi-parametric local regression technique (Cleveland and Devlin, 1988). LOESS estimates a smooth trend surface using weighted least-squares regression in local neighborhoods defined by a fixed number of points closest to each prediction location (the span, measured as a percentage of the total number of data points). Specifically, quadratic LOESS was used with a span of 1%, corresponding to an average neighborhood width of between 3 and 12 km depending on point density.

LOESS was implemented in Matlab version 7.13 (R2011b) with the Curve Fitting toolbox (The MathWorks Inc., Natick, MA). The standard Matlab toolbox function (curvefit/curvefit/+curvefit/LowessFit.m) was modified to reduce processing times and increase matrix stability. Execution speed was improved by using k-dimensional search trees (KD-Tree for MATLAB, Tagliasacchi, 2011) to identify and sort soundings in each local neighborhood. Under certain conditions local regression methods such as LOESS can exhibit instability due to limits on the precision of matrix calculations. X and Y coordinates were centered and re-scaled to minimize the possibility of matrix stability problems. The condition number of each local regression design matrix was also evaluated to diagnose areas where matrix precision might affect the accuracy of regression fits.

Local regression matrix stability was problematic when points that were very close to each other had very different values, which occasionally occurred in areas of high sounding density and resulted in gaps in the LOESS prediction surface. To eliminate gaps, soundings within a horizontal distance of ± 10 cm were identified, grouped and then dispersed with a small random nudge. Coordinates for the first occurrence of a sounding in a group were retained and subsequent coordinates were shifted by adding a uniform random number in the range ± (0.5,1.5 m). Displacements were only accepted if they did not create conflicts (within 10 cm) with other soundings. A total of 700 soundings were modified (<0.03% of all data). Although this dispersion adds some positional error to each sounding, the displacement is negligible in relation to other sources of positioning error caused by geographic positioning systems or ship heave/pitch/roll.

A similar displacement procedure was applied to soundings that fell within 10 cm of the prediction grid coordinates. The purpose of this was to ensure that measurement and micro-scale error were filtered out of the prediction surface, because at the precise locations of original data, the kriging prediction surface exhibits

spikes due to the nugget effect. Filtering the nugget effect in this way is similar to the maximum *a posteriori* resampling technique proposed for filtering noisy bathymetric data by Goff et al. (2006).

The parametric standard error of the mean trend was estimated using a Monte Carlo approach. Specifically, the approximation method of Durban, et al. (1999) was used to estimate the variance-covariance matrix of the estimated local regression coefficients, $Var(\hat{\beta})$. The scale of the variance-covariance matrix was estimated as the sum of squares of the residuals for the whole model (i.e., the residuals of the original data from the LOESS fit at all data points), divided by (N – λ), where N is the number of observations and λ is the effective number of parameters, estimated as, λ = 2*(1+[N/(N*span)]).

Regression coefficient vectors were simulated by 1,000 draws from a multivariate normal distribution defined by mean vector $\hat{\beta}$ and covariance matrix $Var(\hat{\beta})$, and the LOESS prediction was re-calculated for each simulated regression coefficient vector. The standard error was estimated as the standard deviation of the simulated LOESS predictions at each location. The condition number of the design matrix of each local regression was also recorded as an additional diagnostic measure.

Residuals were obtained by subtracting the LOESS trend surface prediction at each data location from the observed data value. Semivariograms of residuals were then calculated and modeled in ArcGIS 10 with the Geostatistical Analyst extension (ESRI, 2011). A separate anisotropic semivariogram model was fit independently for each depth stratum (Figure 2.4, Table 2.2). The nugget effect was adjusted manually based on visual inspection and prior expectations from measurement error models (see below). The rest of the model parameters, including anisotropy ranges and direction, were fit automatically using non-linear weighted least-squares (ESRI, 2011).

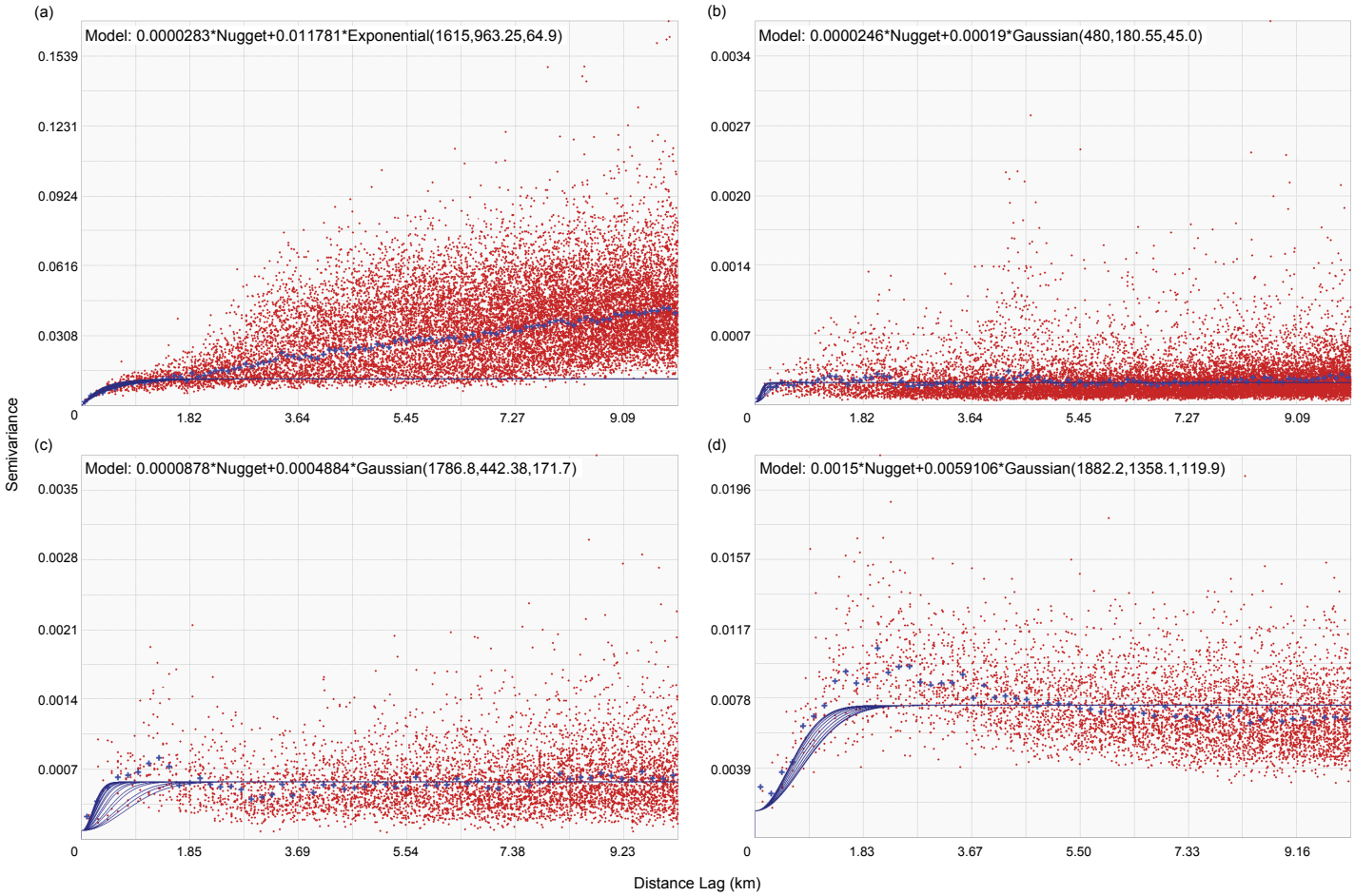

Figure 2.4. Estimated semivariograms of residuals and fitted semivariogram models for each depth stratum. (a) 0-30 m, (b) 30-100 m, (c) 100-200 m, (d) 200-2,000 m. Red dots represent sample semivariance values, blue crosses represent averaged semivariance values, solid blue lines represent directional semivariogram model fits.

Table 2.2. Semivariogram parameters for each depth stratum.

DEPTH STRATUM (m)	n	TYPE OF VARIO-GRAM MODEL	NO. OF LAGS	LAG SIZE (m)	NUGGET $x10^{-5}$(m)	MAJOR RANGE (m)	MINOR RANGE (m)	DIREC-TION (°)*	PARTIAL SILL $x10^{-3}$	% OF THE SILL DUE TO THE NUGGET
0-30	2,077,055	Exp	100	100	2.83	1,615	963.25	64.86	11.8	0.2
30-100	514,081	Gau	100	100	2.46	480	180.55	45.00	0.19	11.5
100-200	31,863	Gau	175	58	8.78	1,787	442.38	171.74	0.488	15.2
200-2,000	43,713	Gau	180	56	150.0	1,882	1,358.12	119.88	5.91	20.2

Exp= exponential; Gau= Gaussian; *Clockwise from North

Nugget selection was guided by a Monte Carlo simulation of measurement error expected for each depth stratum based on maximum measurement error models defined by IHO standards (IHO S.44 Orders 1 and 2). We simulated depth observations with measurement error across the depth range of each stratum (n=100,000 points), log-transformed the simulated depths using Equation 2.2, calculated residuals from the recorded depth, and calculated the depth-averaged measurement error for each stratum in the log-transformed space. This served as a lower bound on the nugget for semivariogram fitting, which was then adjusted higher if necessary based on the best fit to empirical semivariogram plots. The rationale for this approach is that the minimum value of the nugget is equal to measurement error. Small-scale spatial features not resolved by the sample spacing (so-called “micro-scale structures”) can add to this error and raise the value of the nugget, but not lower it.

To perform ordinary kriging of residuals, semivariogram model parameters fitted in ArcGIS 10 (ESRI, 2011) were input into the KT3D module of GSLIB (Geostatistical Software Library, Deutsch and Journel, 1992). KT3D was used instead of ArcGIS 10 because of the prohibitively slow computational speed of the ArcGIS kriging implementation. KT3D was run with ordinary kriging, 8-sector search neighborhoods, a minimum search radius necessary for a gap-free kriging prediction, and a maximum search radius equal to the minimum radius times the anisotropy ratio. At least 1 and no more than 80 points (a maximum of 10 from each sector) were used to produce each kriging prediction. Table 2.3 provides more detailed information for search neighborhood parameters by stratum.

Table 2.3. Search neighborhood parameters by depth stratum.

DEPTH STRATUM (m)	MINIMUM SEARCH RADIUS (km)	MAXIMUM SEARCH RADIUS (km)	SEARCH ELLIPSOID ANGLE (°)*
0-30	3.353	2.0	64.86
30-100	5.317	2.0	45.00
100-200	12.117	3.0	171.74
200-2,000	6.929	5.0	119.88

*Clockwise from North

At each grid location for which sufficient data existed to produce a kriging prediction, the LOESS trend surface was evaluated and estimates of LOESS prediction standard error and condition number were produced. The kriging prediction, kriging variance, LOESS prediction, LOESS variance, and LOESS condition number were exported from Matlab and GSLIB formats to ESRI GRID format for post-processing using the Spatial Analyst extension in ArcGIS 10 (ESRI, 2011).

The model surface representing predicted depth for each stratum was then calculated as the sum of the LOESS and kriging prediction surfaces (see Equation 2.1). The corresponding prediction variance surface was calculated as the square root of the sum of the LOESS variance and kriging variance estimates. This calculation of the total variance assumes that the spatially structured random error component (δ and ε in Equation 2.1) is uncorrelated with the mean component (μ). The prediction variance was used to construct ±1 standard error and 95% confidence interval surfaces (using the standard normal distribution critical value of 1.96).

Prediction, ±1 standard error, and 95% confidence interval surfaces were back-transformed using the equation:

$$(\mathrm{Exp}(Z_{transform}) - a) / b \qquad \text{(Equation 2.3)}$$

where $Z_{transform}$ is the depth prediction in transformed units and a and b are the error model parameters described for Equation 2.2.

Finally, the separate surfaces representing predicted depth and uncertainty for all four strata were mosaicked to generate seamless surfaces covering the whole study area. At locations with more than one prediction (i.e., where strata overlap), values for the mean (or variance) were calculated by a weighted average, where weights corresponded to the inverse of prediction variance (normalized by the sum of the weights for all the depth strata).

2.3.4 Model Validation and Accuracy Assessment

Cross-validation

A cross-validation exercise was carried out for each depth stratum to assess the accuracy of the geostatistical modeling approach. For purposes of this exercise, 50% of the data points in each stratum were selected at random for inclusion as "training data", with the remaining points held out as "validation data." Models were developed following the methods above applied only to the training data, and predictions were evaluated at the validation data locations. The values of the mosaicked prediction and final mosaicked prediction ± 1 standard error surfaces built from the training dataset were extracted at the validation point locations and cross-validation error statistics (Mean Average Error [MAE], Mean Average Percentage Error [MAPE], and Root Mean Square Error [RMSE]) were calculated. Since the final model was produced using the entire dataset, two times larger than the training dataset, these cross-validation statistics represent a conservative upper bound on the error statistics of the final model.

Independent Accuracy Assessment

In addition to cross-validation, geostatistical model predictions were also compared to depth predictions of the CRM (described in Section 2.1) and to a multibeam dataset, hereafter referred to as the STRATAFORM survey. The STRATAFORM survey collected soundings offshore of New York and New Jersey around 39° 12'N 72°50'W as part of the STRATAFORM project using an EM1000 MBES (Mayer et al., 1999; Nittrouer, 1999; Goff et al., 1999). The STRATAFORM survey covered 2,500 km^2 of seafloor in water depths ranging from 20 to 400 m and provides depth estimates for a contiguous surface at 10 m gridded resolution. Although MBES data does contain some error, for our purposes we consider it to represent the "ground truth" since accuracy and resolution of MBES surveys is much better than interpolated and compiled archival hydrographic surveys. To facilitate comparison, our 83.8 m model grid was overlaid on the 10 m STRATAFORM grid and the mean STRATAFORM values in each model grid cell were calculated. The geostatistical model and CRM, were compared to the STRATAFORM survey within the STRATAFORM survey footprint by calculating mean difference (bias), MAE, MAPE and RMSE. Comparison statistics were calculated for the entire area of overlap and for depth strata within that area (30-100 m, 100-200 m).

2.4 RESULTS AND DISCUSSION

2.4.1 Bathymetry Model Predictions

The new bathymetric model extends over the continental shelf and across the shelf slope, covering the majority of the planning area (Figure 2.5). Depth predictions ranged from 0 m at the shore to around 2,100 m on the shelf slope. Some nearshore areas, like the approach to New York Harbor, were not modeled due to processing limitations arising from extremely high data density. A few small patches in the nearshore areas off southern New Jersey and western Long Island did not have model predictions because the geostatistical model was unable to produce predictions where soundings with distinctly different measured depths occurred at virtually the same location. This occurred where older, less accurate surveys coincided with more recent, more accurate surveys.

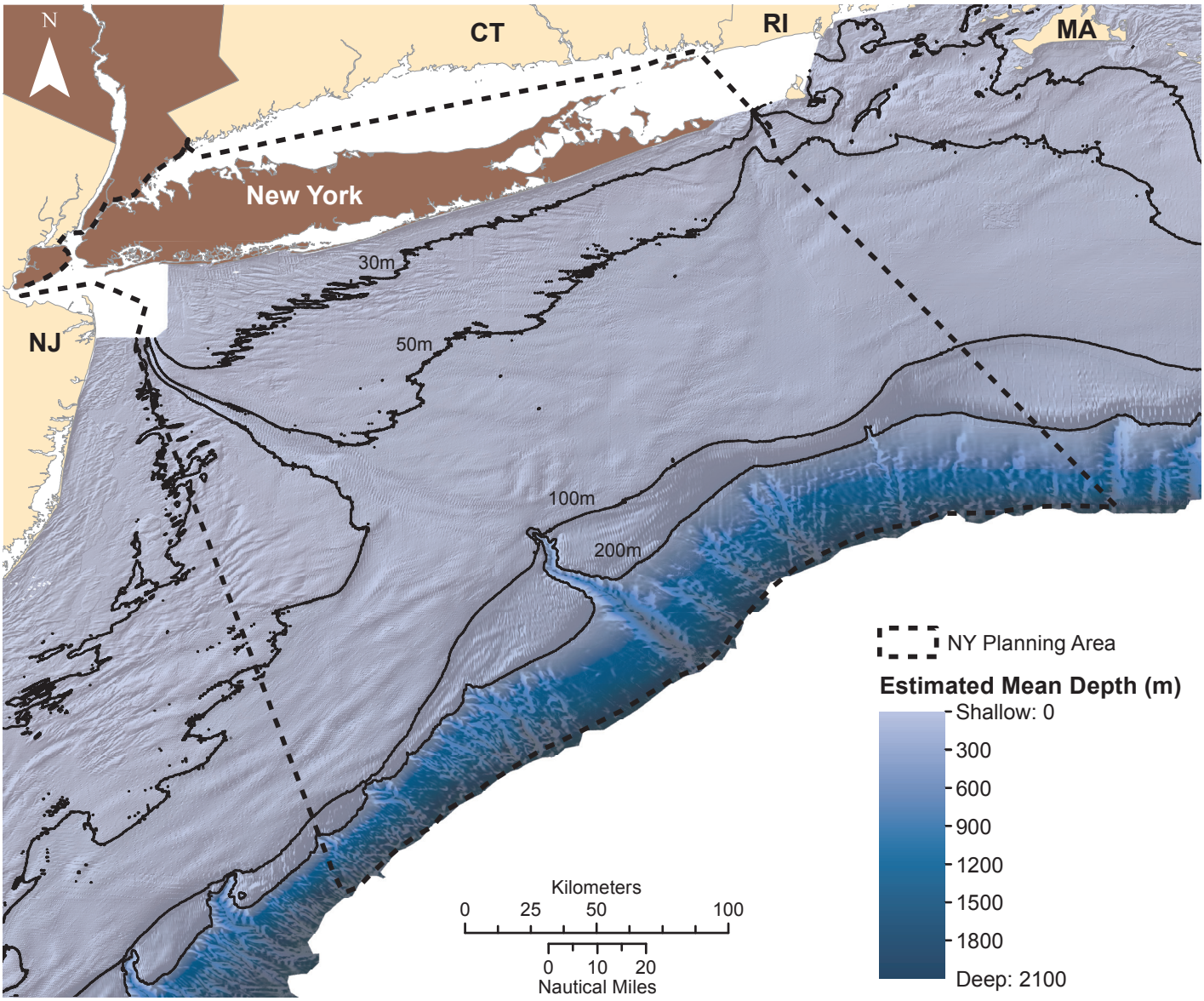

Figure 2.5. Bathymetric surface developed from the geostatistical model, draped over separate hillshade layers derived for the continental shelf (0-200 m) and the shelf slope (>200 m). Solid black lines depict depth contours derived from the bathymetric surface.

Duane et al. (1972) found that sand ridges were a dominant geomorphologic feature on most of the northeast U.S. Atlantic continental shelf. These features were evident in the bathymetry model, particularly to the west of Hudson Canyon and in the northeast of the study area. Submarine canyons, like Hudson Canyon, and shallower networks of gullies were also evident in the model along the shelf slope.

2.4.2 Bathymetry Model Uncertainty

Model standard error ranged from 0.026 m to almost 200 m over the study area (Figure 2.6). In general, model standard error was less than 5 m at depths shallower than 30 m. In this depth stratum, standard error was relatively higher (2-10 m) in areas where surveys occurred prior to 1958 (Figure 2.2). At depths from 30 m to 100 m, standard error was typically less than 2 m, but reached as much as 5 m in some areas where depths approached 100 m. Standard error typically ranged from 2-5 m for depths between 100 and 200 m but was higher (5-10 m) in some areas where depths approached 200 m at the shelf edge. Along the shelf slope, standard error increased from 10-20 m at depths closer to 200 m to greater than 50 m in areas deeper than 500 m. Although standard error generally increased with sounding depth beyond the 30 m depth contour, error was also dependent on distance between surveys. As expected, within each depth stratum, error was generally lower along survey transects where distance between soundings was shortest (lines clearly

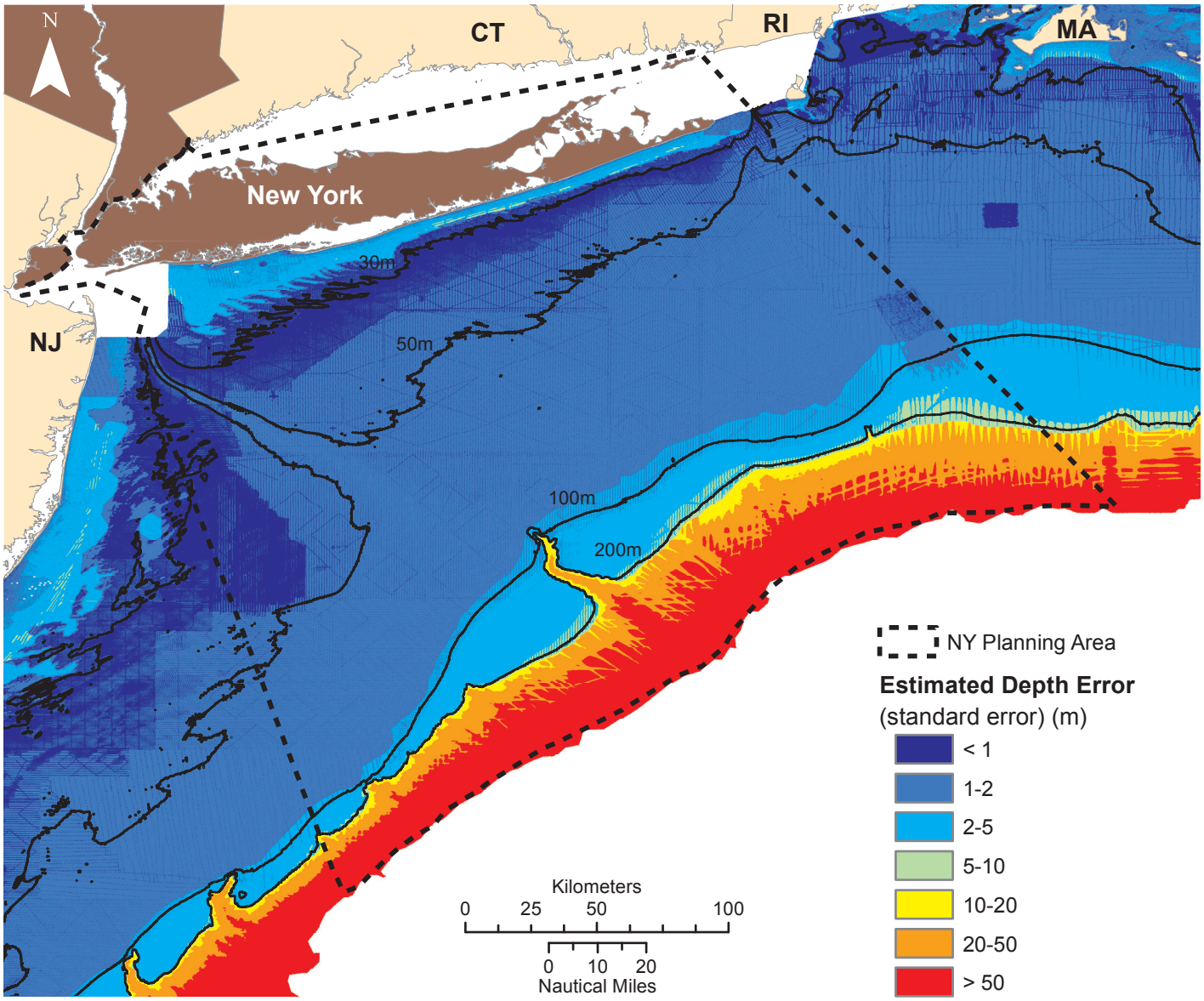

Figure 2.6. Estimated error of predictions (standard error) from the bathymetric model. Model standard error provides an indication of prediction uncertainty.

distinguishable in Figure 2.6). There were two primary reasons for higher error along the shelf slope when compared to shallow areas. First, the absolute precision of sounding instruments generally decreases with depth, and second, soundings become sparser farther offshore. There were several areas south of Hudson Canyon where sounding tracks are greater than 8 kilometers apart (Figure 2.2).

The prediction standard error surface indicates the uncertainty associated with the model prediction at each location, assuming that statistical assumptions of the model are met. Local regressions can be inaccurate as the limits of matrix precision are approached, as indicated by high condition numbers (Figure 2.7). The condition number is a diagnostic that indicates the stability of the local regression trend model at each location. Higher spatial condition number values indicate that the regression solution is less stable at that location, such that small variations in the input data (e.g., uncertainty due to measurement error) can result in large variations in the prediction. For second order polynomials, the critical spatial condition number threshold value is approximately 100, meaning that predictions should be considered with caution at locations where the spatial condition number is close to 100 and should be considered unreliable where it is greater than 100 (Golub and Van Loan, 1996). Under these conditions, the standard error surface may underestimate actual error. This occurs only in a narrow band along the southern coast of Long Island.

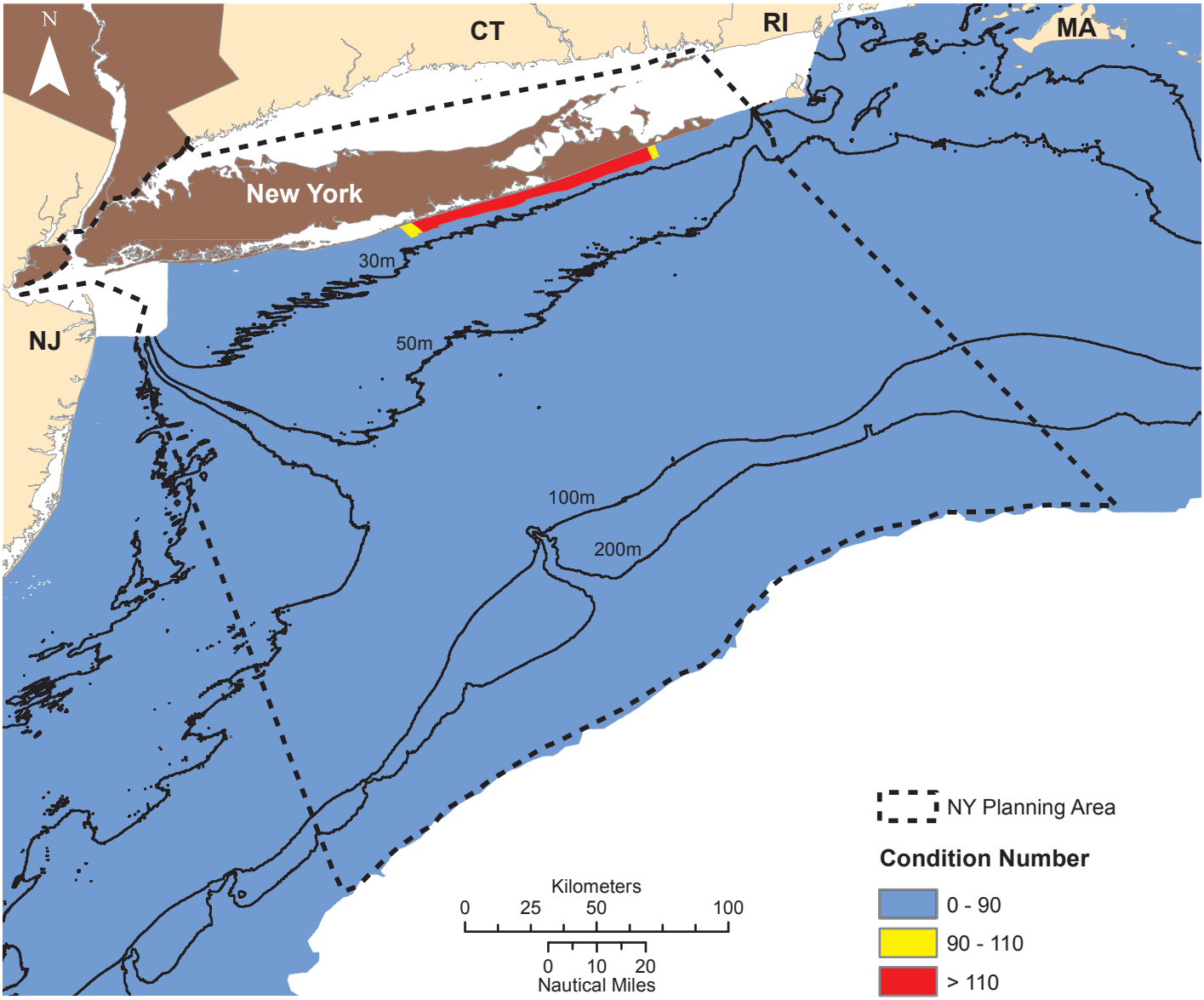

Figure 2.7. Spatial condition number from the LOESS trend model. Condition number classes reflect the threshold at which standard error predictions should be considered unreliable.

Model uncertainty was also depicted using theoretical 95% confidence intervals of the depth predictions (predicted depth ± 1.96*standard error) along two hypothetical transects (Transect 1 and Transect 2) that spanned from the shoreline to the shelf slope (Figure 2.8). Transect 2 differed from Transect 1 in that it cut across a submarine canyon (Hudson Canyon) at the shelf edge. Maximum and minimum depth values representing the upper and lower bounds of the 95% confidence intervals were extracted at 100 m intervals along each transect. For both transects, the width of the 95% confidence intervals generally increased with distance from shore and with depth. However, model uncertainty was greater and more variable in the 0-30 m depth stratum than it was in the 30-100 m depth stratum (Figure 2.9, Figure 2.10). At depths greater than 200 m, the 95% confidence interval widths increased dramatically with distance from shore and had an average vertical width of almost 0.25 km.

To explore how uncertainty in depth predictions may translate into horizontal uncertainty and how this uncertainty could impact management decisions (e.g., the siting of a wind farm), depth contours (30 m, 50 m, 100 m, 200 m) derived from the model prediction and from the upper and lower limits of the theoretical 95% confidence intervals (±1.96*standard error) were overlaid on the boundaries of theoretical wind farm areas within the NY study area (Figure 2.11). These theoretical wind farm areas correspond to areas outside of shipping lanes and within current depth constraints for wind farm structures. These depth contours were mapped to produce a rough estimate of the horizontal uncertainty associated with the model predictions. Within the potential

wind farm areas, transects were drawn between the 50 m depth contours derived from confidence interval limits. The transects were drawn approximately perpendicular to the 50 m depth contour derived from the model prediction at intervals of approximately 5 km.

In this region, the mean distance between the depth contours derived from the confidence interval limits was approximately 8 km with a standard deviation of almost 3 km. While this measure only represents a rough estimate of horizontal uncertainty, it suggests that depth predictions from this model and other models developed using similar data should be used with caution when high positional precision is needed.

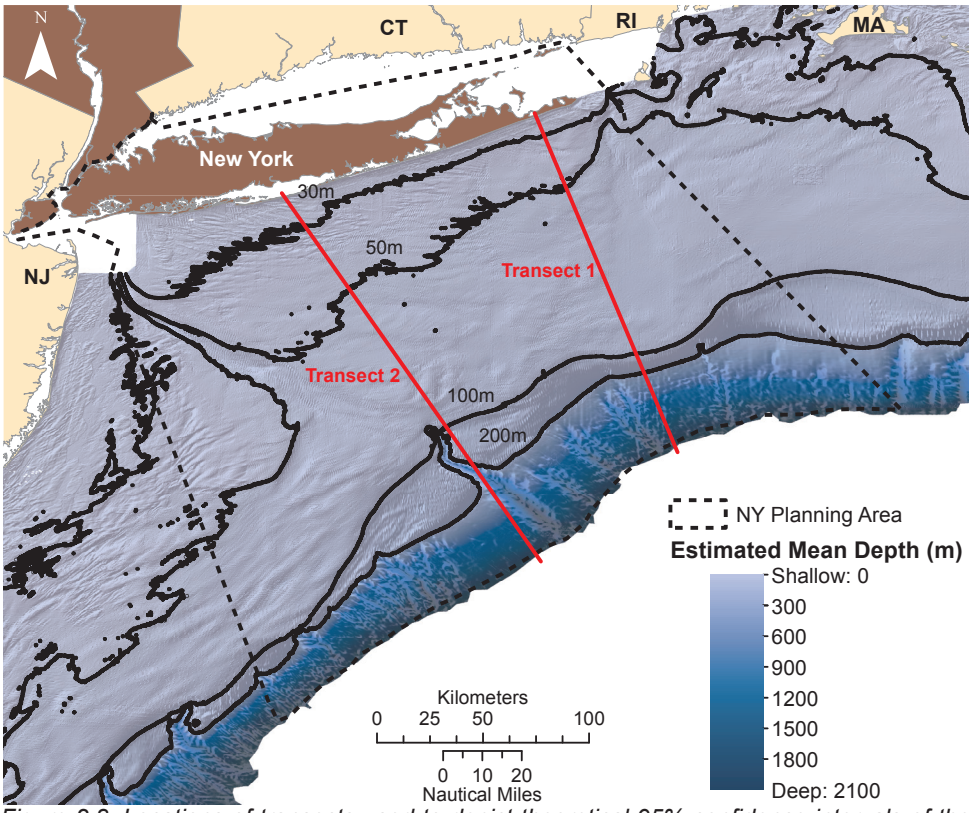

Figure 2.8. Locations of transects used to depict theoretical 95% confidence intervals of the geostatistical model predictions as a function of distance from score and local geomorphology.

2.4.3 Cross-validation of the Training Dataset

Cross-validation results indicated that the geostatistical model performed extremely well in the 0-30 m and 30-100 m depth strata (mean absolute errors 0.60 m and 0.55 m, respectively). The model performed reasonably well in the 100-200 m depth stratum (mean absolute error of 2.1 m, or 1.40%), but model accuracy was considerably degraded in the 200-2,000 m depth stratum (mean absolute error 25.76 m, or 3.44%) (Table 2.4).

Cross-validation was also used to assess the accuracy of confidence intervals. The theoretical 68% confidence intervals (model prediction ± standard error) are somewhat conservative for all depth strata (Table 2.5). For depths below 100 m, the theoretical 95% confidence intervals (model prediction ± 1.96*standard error) are slightly conservative, but the geostatistical model underestimates error at depths greater than 100 m, especially at depths greater than 200 m. For example, in the 200-2,000 m depth range model standard errors should be multiplied by a factor of 2.46, rather than the theoretical value of 1.96, to produce true 95% confidence intervals (Table 2.5).

Table 2.4. Cross-validation statistics for the geostatistical model built from the training dataset. Negative bias indicates a deep bias while positive bias indicates a shallow bias. MAE = Mean Absolute Error, MAPE = Mean Absolute Percentage Error, RMSE = Root Mean Square Error.

DEPTH STRATUM	COMPARISON STATISTIC	CROSS-VALIDATION ERROR
Overall (0-2,000 m)	Bias	-0.08 m
	MAE	1.04 m
	MAPE	6.53%
	RMSE	5.53 m
0-30 m	Bias	-0.17 m
	MAE	0.60 m
	MAPE	7.72%
	RMSE	1.24 m
30-100 m	Bias	0.02 m
	MAE	0.55 m
	MAPE	1.42%
	RMSE	0.96 m
100-200 m	Bias	0.06 m
	MAE	2.14 m
	MAPE	1.40%
	RMSE	5.48 m
200-2,000 m	Bias	2.97 m
	MAE	25.76 m
	MAPE	3.44%
	RMSE	41.23 m

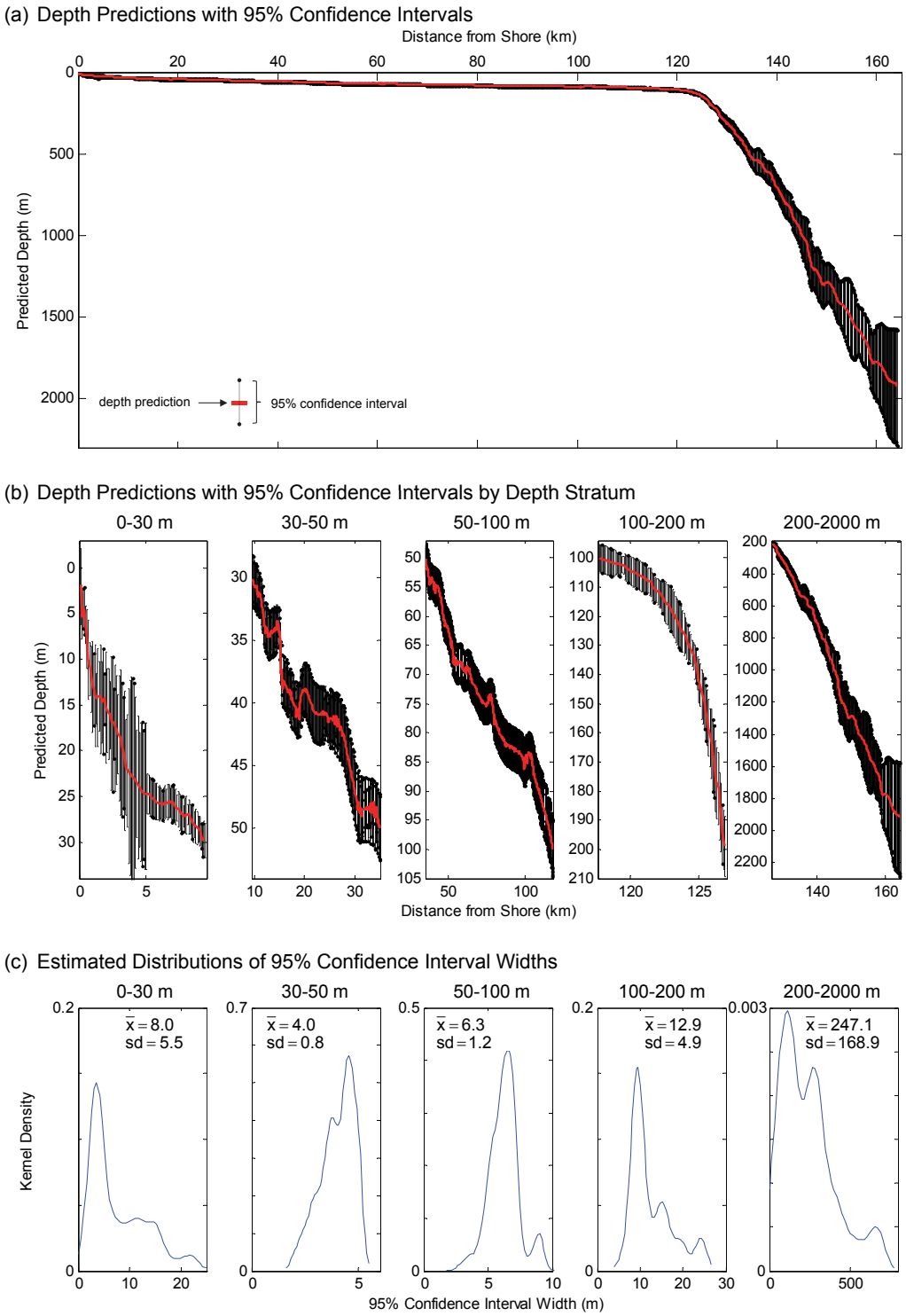

*Figure 2.9. (a) Predicted depth (m) with theoretical 95% confidence intervals (± 1.96*standard error) vs. distance from shore (km) along Transect 1. (b) Predicted depth (m) with 95% confidence intervals by depth stratum. (c) Distribution of 95% confidence interval widths with the mean ($\bar{x}$) and standard deviation (sd) for each depth stratum. The probability density of the confidence interval widths was estimated using a kernel density function (Särkkä 1999).*

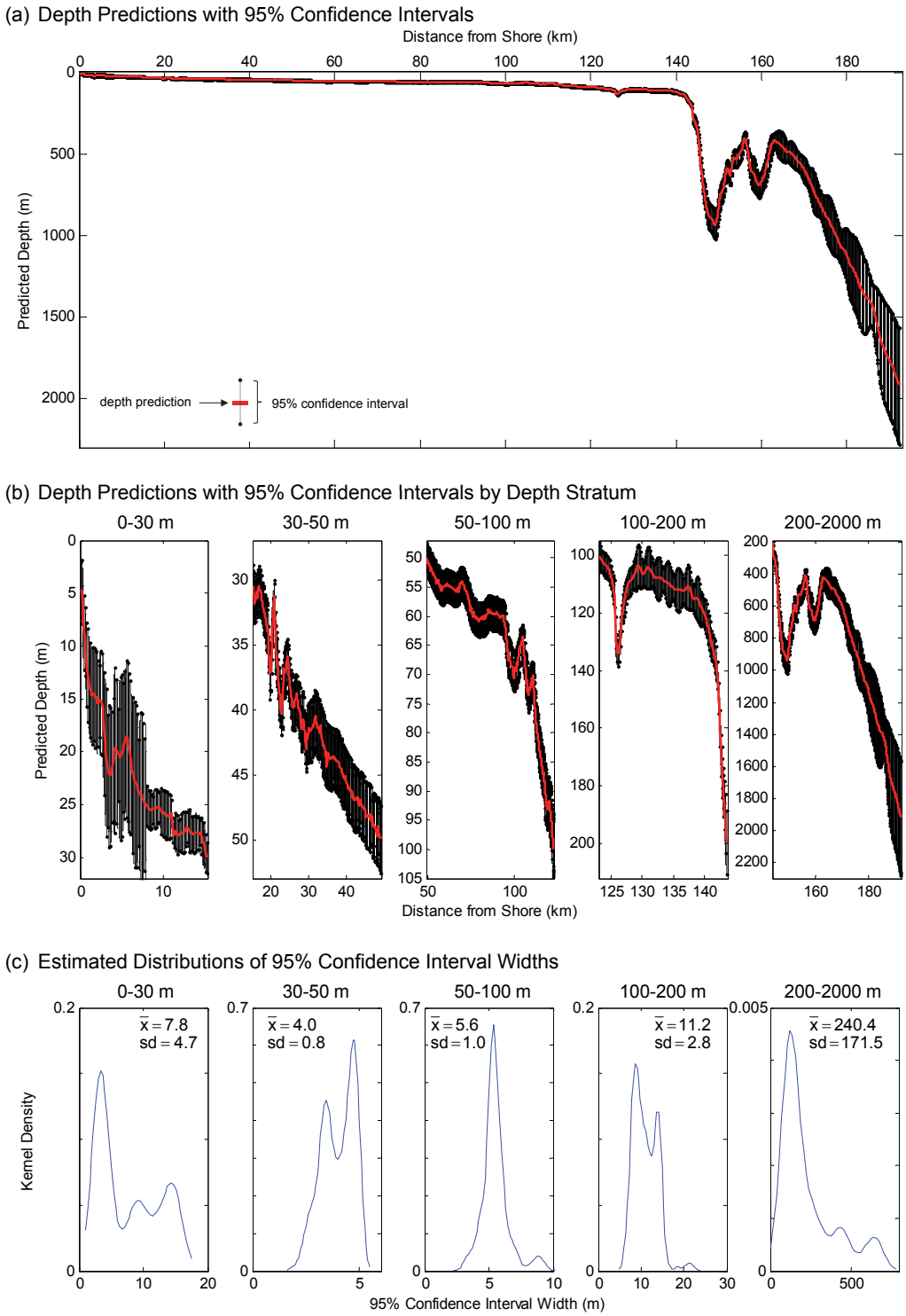

*Figure 2.10. (a) Predicted depth (m) with theoretical 95% confidence intervals (± 1.96*standard error) vs. distance from shore (km) along Transect 2. (b) Predicted depth (m) with 95% confidence intervals by depth stratum. (c) Distribution of 95% confidence interval widths with the mean ($\bar{x}$) and standard deviation (sd) for each depth stratum. The probability density of the confidence interval widths was estimated using a kernel density function (Särkkä 1999).*

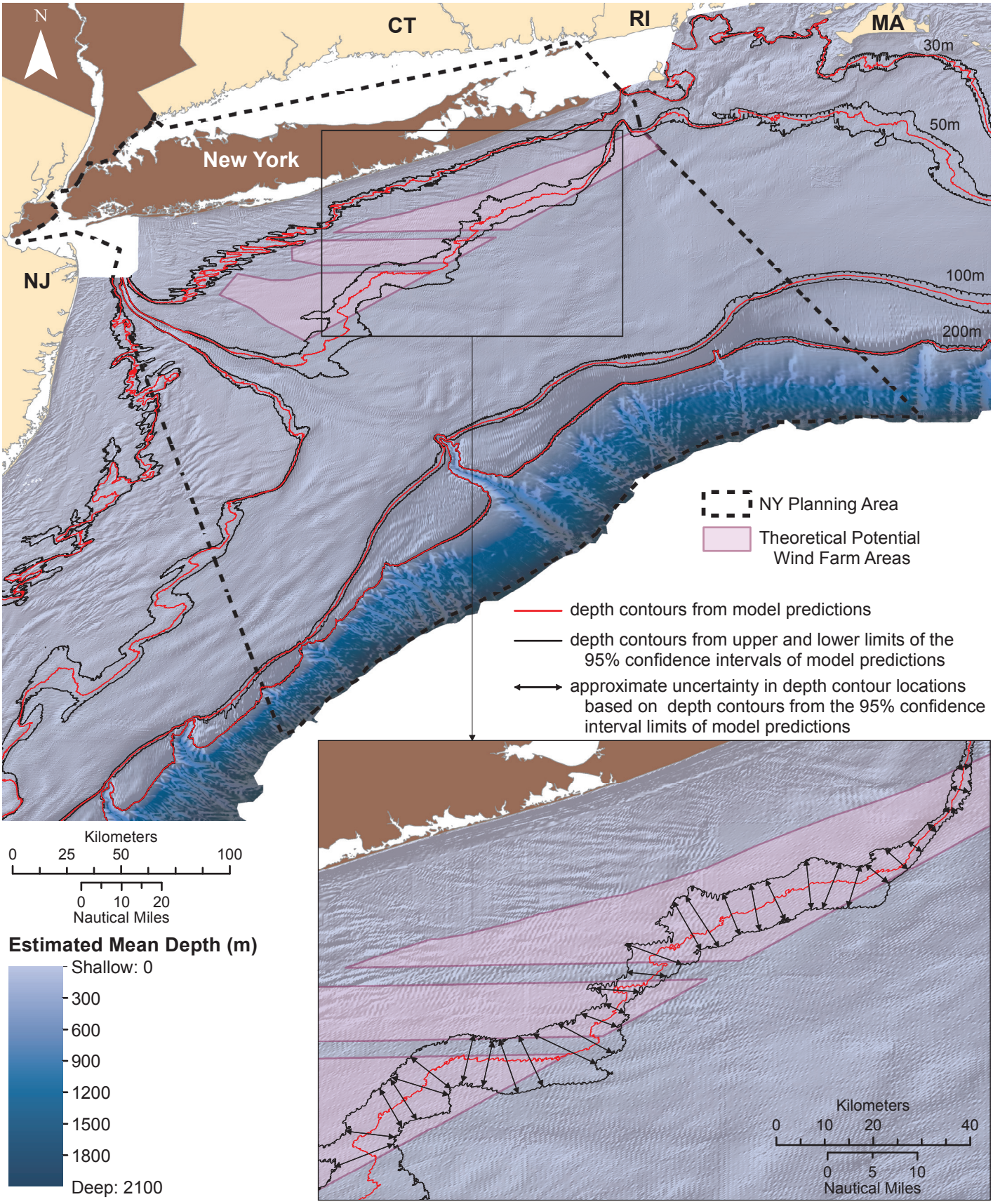

*Figure 2.11. Depth contours derived from the predicted bathymetric surface and from the upper and lower limits of the theoretical 95% confidence intervals (±1.96*standard error) overlaid on the predicted bathymetric surface and theoretical potential wind farm areas for New York. Within the inset map, approximate uncertainty in depth contour locations was estimated using transects approximately perpendicular to the depth contour derived from the predicted bathymetric surface.*

Table 2.5. Performance of theoretical 68% and 95% confidence intervals.

DEPTH STRATUM	PERCENTAGE OF VALIDATION DATA WITHIN THEORETICAL 68% CONFIDENCE INTERVAL	STANDARD ERROR MULTIPLIER FOR 68% CONFIDENCE INTERVAL (COMPARED TO 1.0)	PERCENTAGE OF VALIDATION DATA WITHIN THEORETICAL 95% CONFIDENCE INTERVAL	STANDARD ERROR MULTIPLIER FOR 68% CONFIDENCE INTERVAL (COMPARED TO 1.96)
Overall (0-2,000 m)	87.50%	0.42	95.50%	1.84
0-30 m	87.60%	0.40	95.40%	1.86
30-100 m	87.00%	0.49	95.90%	1.77
100-200 m	88.30%	0.35	94.70%	2.04
200-2,000 m	80.80%	0.62	92.70%	2.46

2.4.4 Independent Accuracy Assessment

The STRATAFORM dataset is ideally located for an independent accuracy assessment of the geostatistical model because it extends across multiple depth strata and overlaps NOS hydrographic soundings collected using multiple sounding and positioning methods.

As a benchmark, we evaluated our model performance in comparison to the CRM in the STRATAFORM area. This may not be an entirely fair comparison, since it is possible (though not verifiable) that the CRM included some version of the STRATAFORM data; however, we proceed anyway with the caution that the CRM error statistics may be significantly better in this region than in other parts of the study area.

Overall we found that both the geostatistical model and CRM are excellent models in the 30-100 m depth range for the STRATAFORM area (mean absolute error [MAE] < 1 m, mean absolute percent error [MAPE] < 1.5%), but accuracy of both models degraded with depth in areas deeper than 100 m (Table 2.6). We found the geostatistical model did not improve upon the CRM in terms of accuracy (when comparing MAE, MAPE, and root mean-square-error [RMSE]) (Table 2.6), but it was able to provide reliable estimates of uncertainty, at least for depths less than 200 m (see Figure 2.6), and obtaining these estimates was the principal reason for undertaking a geostatistical model in the first place. For the 30-100 m and 100-200 m depth strata the percent correct within theoretical 95% confidence intervals were 97.89% and 97.06%, respectively, indicating that the confidence intervals were accurate.

Table 2.6. Results from an accuracy assessment of the geostatistical model and the coastal relief model (CRM). The new geostatistical model and the CRM are compared against the STRATAFORM data. Negative bias indicates a deep bias while positive bias indicates a shallow bias. MAE = Mean Absolute Error, MAPE = Mean Absolute Percentage Error, RMSE = Root Mean Square Error.

DEPTH STRATUM	COMPARISON STATISTIC	INDEPENDENT ACCURACY ASSESSMENT ERROR, GEOSTATISTICAL MODEL	INDEPENDENT ACCURACY ASSESSMENT ERROR, COASTAL RELIEF MODEL
Overall (30-200 m)	Bias	-0.23 m	0.21 m
	MAE	1.17 m	1.15 m
	MAPE	1.18%	1.12%
	RMSE	3.27 m	3.39 m
30-100 m	Bias	0.05 m	0.55 m
	MAE	0.90 m	0.84 m
	MAPE	1.20%	1.10%
	RMSE	1.27 m	1.15 m
100-200 m	Bias	-0.68 m	-0.35 m
	MAE	1.60 m	1.65 m
	MAPE	1.14%	1.15%
	RMSE	5.06 m	5.32 m

Results on bias were mixed. In the 30-100 m depth stratum, the geostatistical model was approximately unbiased, whereas the CRM exhibited a slight shallow bias (+0.55). However, in the 100-200 m depth stratum the geostatistical model exhibited more of a deep bias (-0.68 m) than the CRM (-0.35 m). These biases are small and of the magnitude expected due to known sources of error (e.g., quantization due to rounding of measurement units, changes in tidal references and vertical datums; Calder, 2006).

Taking all comparative data together, the CRM may be the best model if the average value of depth is the primary variable of interest. However, when certainty in depth estimates needs to be accounted for, then the geostatistical model should be preferred, particularly when the depths of interest are shallower than 200 m.

Examples of situations where estimates of bathymetric uncertainty may be useful include measuring the amount of habitat area falling into a given depth range, or identifying suitable construction zones for wind farms based on a depth limit (Figure 2.11). In the latter case, uncertainty can be used to define risk of additional development costs and be used to target the best areas to build within potential construction zones.

2.5 LIMITATIONS TO INTERPRETATION AND FUTURE DIRECTIONS

The geostatistical approach we have employed to create a gridded, interpolated bathymetry surface is an improvement over previous bathymetry models in that it generates a spatially explicit error map to accompany the predicted surface. The cross-validation and independent accuracy assessments show that the model performs similar to the NOAA Coastal Relief Model with the advantage of providing reliable uncertainty estimates. However, several limitations and potential improvements to our approach should be noted here to support interpretation of our models and development of future efforts. Noted limitations will also apply to the CRM and other modeling techniques. Principal limitations of our geostatistical models arise from three general factors:

1) data quality: integrating diverse soundings collected over time and using different methodologies results in a variety of potential distortions in the final surface,
2) resolution: the spatial resolution of original sample data and of the model output grid limit the minimum scale of features that can be resolved, and,
3) model assumptions: geostatistical models involve a number of simplifying assumptions that do not fully capture the complexity of underlying geomorphological patterns.

To help users better understand limitations and appropriate uses of this model, and to guide development of future models, we provide brief explanations and examples of these limitations below and suggest some potential improvements.

2.5.1 Data Quality

Hydrographic soundings in the study area came from a multitude of surveys spanning more than a century (1887-2004). Surveys used an assortment of positioning and sounding technologies, resulting in a patchwork of overlapping soundings collected at varying sample spacings and with different precisions. These errors can propagate to the final model creating distortions that do not correspond to changes on the seafloor.

We have not dealt explicitly with horizontal positioning error. The impact of horizontal positioning error will show up in our models as an increase in the nugget effect over the actual instrument measurement error. Some studies have integrated estimates of positioning uncertainty explicitly into spatial models (Kielland and Tubman, 1994; Jakobsson et al., 2002). Kielland and Tubman (1994) used pseudo-points about the nominal location to combine ship position uncertainties with modeling uncertainties. Jakobsson et al. (2002) used a direct simulation Monte Carlo method in which an ensemble of possible data configurations were drawn assuming a distribution of positioning errors. These approaches could be used to improve the precision of our estimates of bathymetric uncertainty by accounting for differences in positioning certainty between older and newer data.

We also did not explicitly account for differences and possible systematic biases in vertical accuracy of survey data. Archival NOS Hydrographic Survey data has been processed using varying methods over the years which have created some systematic post-processing errors (see Calder, 2006 for a full discussion). Briefly, Calder (2006) reported that archival lead line soundings (common prior to 1978) are systematically shallow-biased because of “hydrographic rounding” (a tendency to round down to the next shallowest whole fathom). Generally, the more recent VBES data appears approximately unbiased (but see below), and modern multibeam surveys offer the most precise information. Calder’s findings suggest that some older VBES soundings are also biased

because they were digitized from paper charts for which data were first rounded to the next shallowest foot or fathom (listed in the metadata as “smooth sheets digitized for NOS under contract”). It may be impossible to correct for biases in these data because the precise procedures followed were not recorded. More recent data entered directly into the database after collection by digital instruments are less likely to have systematic rounding error.

For our purposes, we applied a correction factor to lead line surveys because these were found to have a predictable, systematic error in our study area (Calder, 2006). However, we were unable to correct for probable systematic biases in other sounding methods (e.g., VBES data that went through a smooth sheet digitization). Survey metadata indicates that approximately 18% of soundings were non-lead line data that were digitized using smooth sheets, and therefore would be improved by some bias correction. We attempted to reduce the effect of these potentially biased surveys by eliminating data from those surveys where they fell within the footprint of more modern surveys known to be unbiased (direct digitally ingested VBES and MBES). However, unfortunately, about 40% of the study area was only covered by archival lead line data and/or VBES data digitized from paper charts. These areas are less reliable and may contain systematic biases (typically shallow-biased by <2 m) that are not fully reflected in our model uncertainty estimates. In places where only older, less reliable data are available, we suggest using maps of survey age (Figure 2.2) and/or estimated survey measurement error based on the technique used for sounding (Figure 2.12) to supplement model-based uncertainty maps. These maps viewed alongside geostatistical model errors can help identify unreliable areas and areas where additional bathymetric information would improve future planning decisions.

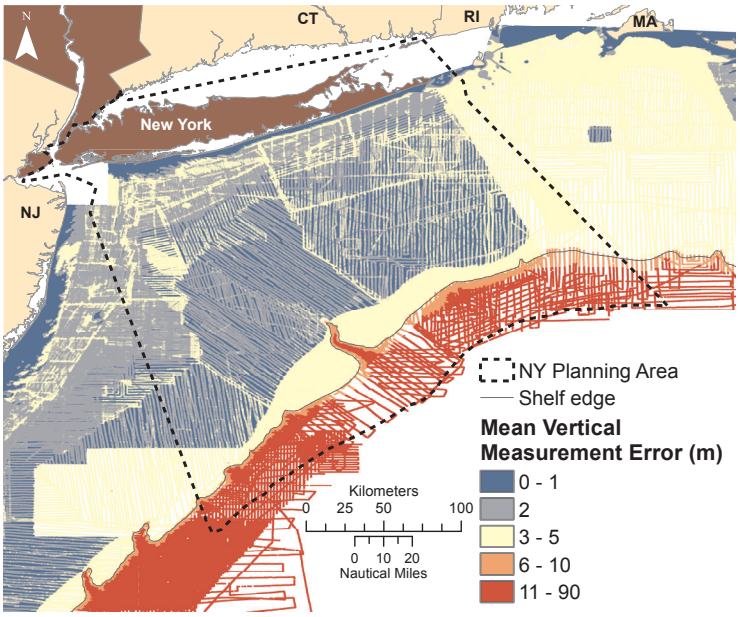

Figure 2.12. Mean estimated vertical measurement error for soundings within 1 km rectangular neighborhoods. Vertical measurement errors were estimated based on vertical sounding technology used. For some surveys this was inferred from survey age. The shelf edge corresponds to the 200 m depth contour.

We have purposefully neglected consideration of changes to the seafloor occurring over time. Temporal changes may or may not be reflected in spatially-explicit model error estimates depending on the ages of nearby surveys. We expect positional error attributed to change over time may be substantial in some areas, especially in highly dynamic areas, such as where tidal and riverine influences are great.

Finally, we note that recently developed geostatistical algorithms could be used in the future to account for heterogeneous measurement error among methods (Christensen, 2011) to improve accuracy and more appropriately weight higher quality data.

2.5.2 Resolution

The distance between soundings was not uniform across the study area (Figure 2.2). The length scale of features that can be resolved will be shorter (higher resolution) in areas with greater sounding density. Moreover, it is possible that our model-based uncertainties will underestimate true error in areas with sparse soundings, especially when very high amplitude, high frequency features are present (e.g., high frequency, short-wavelength sand waves). The chances of this and other problems arising from interactions of sample spacing and high-frequency features (e.g., aliasing) are greater when samples are both sparse and very regularly spaced. Fortunately, in our cross-validation and independent accuracy assessment we did not find evidence for significant overall underestimation of uncertainty, but localized impacts are still possible in areas with high frequency features relative to sample spacing (e.g., sand waves).

In general, fine-scale or very sharply defined features, such as erratics, deep sea reefs or man-made artifacts, will not be reliably resolved in our model. The spatial scale at which these features become visible is dependent on the relative distribution of soundings, the methods used to model spatial structure and the output resolution of maps. Although the sounding density would support resolution approaching 10 m in some limited areas (see Figure 2.2 where sounding density exceeds 100 km^{-1}), the output model grid size was ~85 m. The minimum scale of resolved length scales is twice the output resolution, or ~170 m. For the vast majority of the study area data density was much sparser and could only detect features at scales on the order of 10^2 m and in some areas 10^3 m. In general, new MBES and/or sidescan sonar surveys are needed if greater detectability at short spatial scales is required. In some cases, modern VBES surveys acquired with co-registered sidescan information may be used to identify some missed features. Additionally, Calder (2006) presents a unique method of integrating a variance term corresponding to “hydrographic oversight” of smaller features, but the term requires a very good understanding of the data and geomorphology.

2.5.3. Model Assumptions

A full discussion of the statistical assumptions inherent in the LOESS local regression and geostatistical approaches used here is beyond the scope of this chapter; the reader is referred to texts on the subjects (e.g., Cleveland and Devlin, 1988; Cressie, 1993; Chiles and Delfiner, 1999). However, it is important to note here that geostatistical models are not capable of reproducing the full complexity of geomorphological patterns (e.g., alluvial fans, sand waves) unless data are very dense. This is because geostatistical models describe spatial correlation as a simple function of distance between points, allowing only for very simple geometric anisotropy. Complex multi-point erosional and depositional patterns can’t be resolved unless they are densely sampled. Texture-mapping approaches could possibly improve prediction of complex geomorphology (e.g., Boucher, 2009). Ultimately, however, collection of new multibeam bathymetry is preferable to any attempt to statistically reconstruct fine-level details in archival hydrographic surveys.

2.6 ACKNOWLEDGMENTS

We thank Brian Calder and Larry Mayer (The Center for Coastal and Ocean Mapping & NOAA/UNH Joint Hydrographic Center) for providing the gridded STRATAFORM multibeam data used in the accuracy assessment. Bryan Costa, Will Sautter, Tim Battista provided helpful discussions in support of this work and also assisted with data acquisition/analysis for bathymetry and vector shorelines. We are grateful to Brian Calder and John Goff for detailed technical reviews and for discussions that substantially improved this work; however, they are not responsible for any errors or omissions in the final product.

2.7 REFERENCES

Amante, C. and B.W. Eakins. 2009. ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. 19 pp.

Boucher, A. 2009. Considering complex training images with search tree partitioning. Computers and Geosciences 35:1151-1158.

Butman, B., W.W. Danforth, W.C. Schwab, and M.R. Buchholtz ten Brink. 1998. Multibeambathymetric and backscatter maps of the Upper Hudson Shelf Valley and adjacent shelf, offshore of New York. U. S. Geological Survey Open-File Report 98-616.

Butman, B., T.J. Middleton, E.R. Thieler, and W.C. Schwab. 2003. Topography, shaded relief, and backscatter intensity of the Hudson Shelf Valley, offshore of New York. U. S. Geological Survey Open-File Report 03-372.

Butman, B., D.C. Twichell, P.A. Roma, B.E. Tucholke, T.J. Middleton, and J.M. Robb. 2006. Sea floor topography and backscatter intensity of the Hudson Canyon region offshore of New York and New Jersey. U.S. Geological Survey Open-File Report 2004-1441.

Calder, B.R. 2006. On the uncertainty of archive hydrographic datasets. IEEE Journal of Oceanic Engineering 31(2): 249-265.

Calderbank, B. 2001. Radio positioning accuracies. Lighthouse 59:12-15.

Chiles, J. P. and P. Delfiner. 1999. Geostatististics: modelling spatial uncertainty. New York: Wiley-Interscience.

Christensen, W.F. 2011. Filtered kriging for spatial data with heterogeneous measurement error variances. Biometrics. DOI: 10.1111/j.1541-0420.2011.01563.x.

Cleveland, W.S. and S.J. Devlin. 1988. Locally weighted regression: an approach to regression analysis by local fitting. Journal of the American Statistical Association 83(403):596-610.

Cressie, N.A.C. 1993. Statistics for spatial data (revised ed.). New York: John Wiley & Sons, Inc.

Curry, R.G. 1996. HydroBase: a database of hydrographic stations and tools for climatological analysis. Woods Hole Oceanographic Institution Technical Report WHOI 96-01, 44 pp.

Deutsch, C.V. and A.G. Journel. 1992. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press. 340 pp.

Duane, D.B., M.E. Field, E.P. Meisburger, D.J.P. Swift, and S.J. Williams. 1972. Linear shoals on the Atlantic Inner Continental Shelf, Florida to Long Island. In: Swift, D.J.P., D.B. Duane, and O.H. Pilkey (eds.). Shelf sediment transport: process and pattern. Stroudsburg, PA: Dowden, Hutchinson and Ross.

Durban, M., C.A. Hackett, and I.D. Currie. 1999. Approximate standard errors in semiparametric models. Biometrics 55(3):699-703.

Environmental Systems Research Institute, Inc. (ESRI). 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

ESRI Online. 2011. ArcGIS Online World Imagery service from ESRI ArcGIS Online and data partners, including imagery from agencies supplied via the Content Sharing Program. Available from http://www.arcgis.com/home/, Accessed Dec 14, 2011.

Gardner, J.V., P. Dartnell, L.A. Mayer, and J.E. Hughes-Clarke. 2003. Geomorphology, acoustic backscatter, and processes in Santa Monica Bay from multibeam mapping. Marine Environmental Research 56:15-46.

Goff, J.A., D.J.P. Swift, C.S. Duncan, L.A. Mayer, and J. Hughes-Clarke. 1999. High resolution swath sonar investigation of sand ridge, dune and ribbon morphology in the offshore environment of the New Jersey Margin. Marine Geology 161:309–339.

Goff, J.A., C.J. Jenkins, and B. Calder. 2006. Maximum likelihood resampling of noisy, spatially correlated data. Geochemistry, Geophysics, and Geosystems 7, Q08003, DOI:10.1029/2006GC001297.

Golub, G.H. and C.F. Van Loan. 1996. Matrix Computations, 3rd edition. Johns Hopkins University Press.

International Hydrographic Organization. 1998. IHO Standards for Hydrographic Surveys, Special Publication No. 44, 4th Edition.

Jakobsson, M., B. Calder, and L. Mayer. 2002. On the effect of random errors in gridded bathymetric compilations. Journal of Geophysical Research – Solid Earth 107(B12).

Kielland, P. and T. Tubman. 1994. On estimating map model errors and GPS position errors: applying more science to the art of navigation. Navigation 41:479-499.

Kostylev, V.E., B.J. Todd, G.B.J. Fader, R.C. Courtney, G.D.M. Cameron, and R.A. Pickrill. 2001. Benthic habitat mapping on the Scotian shelf, based on multibeam bathymetry, surficial geology and sea floor photographs. Marine Ecology Progress Series 219:121-137.

The MathWorks Inc. 2003. MATLAB version 7.13 (R2011b) Natick, Massachusetts.

Mayer, L., L. Fonseca, M. Pacheco, S. Galway, J. Martinez, and T. Hou. 1999. The STRATAFORM GIS, U.S. Office of Naval Research (ONR), St. Andrews, Canada.

Nittrouer, C. STRATAFORM: overview of its design and synthesis of its results. Marine Geology 154(1-4):3-12.

Särkkä, S. 1999. Kernel1 function (for MATLAB). Downloaded November 2011.

Schwab, W.C., W. Corso, M.A. Alison, J.F. Denny, L. Lotto, W.W. Danforth, D.S. Foster, T.F. O'Brien, D.A. Nichols, B.J. Irwin, and K.F. Parolski. 1997a. Mapping the sea floor geology offshore of the New York-New Jersey metropolitan area using sidescan sonar: preliminary report. U.S. Geological Survey Open-File Report 97-61.

Schwab, W.C., M.A. Allison, W. Corso, L.L. Lotto, B. Butman, M. Buchholtz ten Brink, J. Denny, W.W. Danforth, and D.S. Foster. 1997b. Initial results of high-resolution sea-floor mapping offshore of the New York - New Jersey metropolitan area using sidescan sonar. Northeastern Geology and Environmental Sciences 19(4):243-262.

Tagliasacchi, A. 2011. KD-Tree (Library for MATLAB). Downloaded August 2011 from http://www.mathworks.com/matlabcentral/fileexchange/21512-kd-tree-for-matlab.

Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46:234–240.

This page intentionally left blank.