Exploring Ancillary Parameters for Quantifying Interpolation Uncertainty in Digital Bathymetric Models

Elias Adediran[a]*, Kim Lowell[a], Christos Kastrisios[a], Glen Rice[b] and Qi Zhang[c]

[a] *Center for Coastal and Ocean Mapping/UNH-NOAA Joint Hydrographic Center, University of New Hampshire, Durham, NH, USA;* [b] *U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Office of Coast Survey, Hydrographic Systems and Technology Branch, USA;* [c] *Department of Mathematics and Statistics, University of New Hampshire, NH, USA.*

*elias.adediran@unh.edu

Elias Adediran is an Ocean Engineering: Ocean Mapping master's student and a Research Assistant at the Center for Coastal and Ocean Mapping / UNH-NOAA Joint Hydrographic Center, University of New Hampshire. His research focuses on utilizing a multidisciplinary approach to characterizing and estimating the uncertainty in interpolated bathymetry.

Kim Lowell is a Research Scientist at the Center for Coastal and Ocean Mapping (CCOM). His primary focus at CCOM is the application of machine learning, deep learning, and other data analytics techniques to improve the accuracy of bathymetric charts.

Christos Kastrisios is a Research Associate Professor at the University of New Hampshire Center for Coastal & Ocean Mapping. His research at CCOM focuses on data generalization, visualization, and topology on nautical charts.

Glen Rice is a Physical Scientist within the National Oceanic and Atmospheric Administration's (NOAA) Office of Coast Survey's Hydrographic Systems and Technology Branch and the technical team lead for the National Bathymetric Source Program. Mr. Rice has a B.S. in Physics and an M.S. in Ocean Engineering from the University of New Hampshire.

Qi Zhang is an Associate Professor in the department of Mathematics and Statistics at the University of New Hampshire. His research interests include developing statistical models for high dimensional data and causal inference.

Exploring Ancillary Parameters for Quantifying Interpolation Uncertainty in Digital Bathymetric Models

The oceans remain one of Earth's most enigmatic frontiers, with approximately 75% of the world's oceans still unmapped. To create a seamless digital bathymetric model (DBM) from sparse bathymetric datasets, interpolation is employed, but this introduces uncertainties of unknown nature. This study aims to estimate and characterize these uncertainties, which is important in many fields, particularly nautical charting, and navigational safety. Complete seafloor coverage sonar depth data for five testbeds that varied in slope and roughness are sampled at a range of densities (1% to 50%) and interpolated across an entire area using spline, inverse distance weighting (IDW), and linear interpolation. The resulting uncertainties are evaluated from both scientific and operational perspectives. Employing linear regression and machine learning techniques, the relationships between these uncertainties and ancillary parameters (distance to the nearest measurement, seabed roughness, and slope) are examined for quantifying and characterising the interpolation uncertainty. Results indicate insignificant operational differences among the three interpolators in depth estimation, as well as the statistical significance of the examined uncertainty predictors. Additionally, findings suggest the potential presence of unaccounted-for factors shaping uncertainty, yet this work lays a foundational understanding for improving the estimate of uncertainty in DBMs.

Keywords: hydrography; interpolation-based digital bathymetric models; uncertainty estimation; geospatial analysis; machine learning; nautical charting

Introduction

Essential for sustaining life, climate regulation, and economic activities, the Earth's oceans, that constitute approximately 71% of our planet's surface (Weatherall *et al.* 2015), are a vast reservoir of resources and wealth (Mayer *et al.* 2018). Digital bathymetric models (DBMs), derived from depth measurements, are commonly used for the representation of the seafloor. DBMs are stored in a raster data format with each cell representing the average depth of the area contained within that cell (Amante and Eakins 2016, Jakobsson *et al.* 2019) but other depth representations (shallow, deep, or nodal depths) are also employed based on application needs.

A cursory glance at the available global DBMs (e.g., The General Bathymetric Chart of the Oceans (GEBCO) (Mayer *et al.* 2018), the Global Multi-Resolution Topography (GMRT) (Ryan *et al.* 2009)) and regional DBMs (e.g., the International Bathymetric Chart of the Arctic Ocean (IBCAO) (Jakobsson *et al.* 2020), the European Marine Observation and Data Network (EMODnet) (Schaap and Schmitt 2020)) may provide the false impression that the seafloor bathymetry of the oceans is well known. However, these models, which incorporate data derived from the traditional sounding technique of lead line, single-beam echosounders (SBES) and modern high-resolution multibeam echo sounders (MBES), largely rely on interpolated and altimetry-derived data (Weatherall *et al.* 2015) as over 75% of the seabed remains unmapped to modern standards (Nippon Foundation-GEBCO 2023).

In creating a comprehensive model of the seafloor, interpolation is often required to fill data gaps spanning meters to kilometers between survey depth measurements (Smith and Sandwell 1997). Given this reliance on interpolation, satisfying the precision needs for applications utilizing bathymetry hinges significantly on the accuracy and associated uncertainties introduced by the interpolation method. Indeed, while various

sources contribute to the overall uncertainty, the interpolation process emerges as a predominant factor, potentially exerting the most substantial influence on the accuracy of diverse applications.

As with any scientific measurement, the uncertainty associated with interpolated bathymetry is crucial for many applications, but particularly in nautical charting and to ensure the safety of navigation. This becomes crucial in the context of maritime accidents, as highlighted by Kastrisios and Ware (2022). Knowing how deep an area is without the knowledge of quality and source of the reported depth makes the depth information incomplete and, consequently, less useful, or even dangerous. This underlines the importance of bathymetry quality, forming a vital component of the pillars of hydrography alongside depth and its source Click or tap here to enter text.(Rice et al. 2023). An accurate and precise estimate of uncertainty helps hydrographic offices globally address the challenge of assigning a representative and useful category zone of confidence (CATZOC) (IHO 2014) to areas without full seafloor bathymetric coverage. Consequently, this enhances navigational safety, optimizes shipping routes, and improves survey planning, leading to associated cost savings.

Despite the widespread use of DBMs, the inherent interpolation uncertainty in these models remains largely unexplored, holding broad implications across academic, commercial, and potentially life-saving domains. Estimating DBM uncertainty may contribute to various fields, including research in climate change, stream-flow, sediment transport, soil-landscape modeling, and commercial interests such as shipping route optimization and resource extraction. For example, hazard modeling, encompassing tsunami inundation and hurricane storm-surge inundation, relies on bathymetric-topographic digital elevation models developed through interpolation methods (Eakins and Taylor 2010), contributing to the creation of flood maps and evacuation routes.

Consequently, estimating and effectively communicating the uncertainty inherent in these models and flood maps to the public can potentially mitigate the risk for human losses.

Despite the critical importance of understanding uncertainty in interpolated bathymetric models, the field of ocean mapping lacks comprehensive research on estimating interpolation uncertainty. In response to this gap, this study aims to identify the optimal deterministic interpolation methods that produce the lowest interpolation uncertainty from both scientific and operational perspective and, importantly, explore the relationships between interpolation uncertainty and ancillary parameters for accurately estimating interpolation uncertainties in DBMs within operational settings. In this context, "operational" refers to the efficient generation of bathymetric products through a data-driven workflow from extensive national datasets, incorporating accurate hydrographic quality metrics of uncertainty in a timely manner.

The study delves into three commonly employed deterministic interpolation techniques—Linear, Spline, and Inverse Distance Weighting (IDW)—with a detailed explanation provided in the Background and Related Work Section.

This research aims to address the following questions:

(1) Which deterministic interpolation method is most effective for bathymetry, yielding the lowest interpolation uncertainty?
(2) How efficiently can the relationship between interpolation uncertainty and a suite of ancillary parameters, including distance to the nearest measurement and terrain characteristics, specifically slope and roughness, be leveraged to accurately estimate uncertainty in bathymetric models in an operational setting?
(3) Which parameter serves as the most important predictor of uncertainty, and to what extent can we improve the estimated uncertainty by employing machine learning techniques to combine these parameters?

(4) How adeptly can we characterize these uncertainties to investigate the impact of seafloor morphology, data paucity, spatial resolution, and spatial scales?

Addressing these research gaps of improving the estimate and characterization of interpolation uncertainty and identifying the optimal interpolation method are not only particularly important for better filling the bathymetric gaps but also significant to the entire hydrographic community at large for the safety of navigation, shipping route optimization, and survey planning and associated cost savings. It is also of great relevance to the many ongoing NOAA data-driven projects such as the National Bathymetric Source Program, aimed at building a DBM for the U.S. (Rice *et al.* 2023). It will also support Precision Navigation and Office of Coast Survey Hydrographic Health Model that drives Survey Planning and Prioritization through the NBS project which are the cornerstones of safer navigation, resilient coastal communities and ecosystems, and a stronger blue economy i.e., sustainable use and management of ocean resources for economic growth.

The subsequent sections of this paper are structured as follows: Background and Related Works offers background information and delves into related work, Materials and Methods details the datasets and methods employed, Results presents the obtained results on the five testbeds, Discussion engages in a comprehensive discussion of the findings, and Conclusion offers concluding remarks.

Background and Related Works

Interpolation methods

Interpolation is a mathematical process that predicts unknown values based on surrounding measured values, assuming that the surface is continuous and smooth, and that the measured values at neighboring data points are highly correlated with the value at the unknown point (Burrough and McDonnell 1998, Liu *et al.* 2007). There are various interpolation methods, all of which assume that bathymetry is positively spatially autocorrelated, based on Tobler's 1st law of geography (Tobler 1970). Interpolation methods can be classified as geostatistical or deterministic, local or global, and exact or inexact based on the assumptions and features used. Different methods produce different results, even when developed from the same source data (Aguilar *et al.* 2005, Erdogan 2009).

Geostatistical methods, such as kriging, leverage mathematical and statistical functions to estimate depths by incorporating the concepts of randomness and uncertainty (Childs 2004, Negreiros *et al.* 2010, Gunarathna *et al.* 2016). On the other hand, deterministic methods like Spline, Inverse Distance Weighting (IDW), and Triangulated Irregular Network (TIN) use measurements directly and mathematical functions solely for predicting unknown values, without offering estimates of vertical uncertainty (Childs 2004, Negreiros *et al.* 2008). Beyond the crucial distinction of providing uncertainty estimates, it is noteworthy that geostatistical methods come with the trade-off of being more computationally— and time—intensive. Given the limitations of Kriging, deterministic interpolators are better suited to the goals of this research, which centers on data-driven operational hydrography with the objective of generating products from extensive national datasets while assigning the appropriate hydrographic quality metric

of uncertainty. Consequently, this study focuses on investigating deterministic spatial interpolation methods for uncertainty quantification.

Exact interpolation methods employ known data points to derive a surface that precisely passes through all the data points (Özdamar *et al.* 1999), while inexact interpolation methods construct interpolated surfaces that approximate, but do not have to pass through known points (Lam 1983). For nautical charting purposes, exact interpolators seem to be a more fitting choice as they ensure that depth in DBMs at the known locations remains as measured. This is of utmost importance for navigational safety.

Global interpolation methods utilize a single function or model to fit the entire dataset, enabling the estimation of values at any point within the data range (Arun 2013). In contrast, local interpolation methods employ a different model for each point or local neighborhood of points. These methods leverage nearby data points to estimate the value of an unknown point (Li and Heap 2014, Wu *et al.* 2019).

The interpolation method is chosen based on data quality, sampling distribution, terrain characteristics, computational resources, and application requirements, with each method having specific mathematical constraints for predicting unknown values (Amante 2012). Bathymetry interpolation methods have been extensively studied, including both deterministic and geostatistical techniques (Legleiter and Kyriakidis 2006, Merwade *et al.* 2006, Merwade 2009, Vetter *et al.* 2011, Šiljeg *et al.* 2014, Curtarelli *et al.* 2015, Amante and Eakins 2016, Panhalakr and Jarag 2016, Chowdhury *et al.* 2017, Henrico 2021). However, there is no consensus on the best method (Wu *et al.* 2019), as performance depends on factors like dataset specifics, seafloor characteristics, and study areas.

This research in part aims to identify the optimal global deterministic interpolation method for generating bathymetric models for different seafloor types. Among the three deterministic, local, and exact spatial interpolation methods examined in this study—IDW, Spline, and Linear—IDW employs weights inversely proportional to distances, with a chosen power parameter controlling the rate of influence decrease (Caruso and Quarta 1998, Liu *et al.* 2007, Guo *et al.* 2010). A power value of 2 identified by (Amante and Eakins, 2016) as optimal, is employed in this work. Spline interpolation, on the other hand, fits a piecewise-defined polynomial function, typically a cubic spline, to the dataset, creating a smooth curve passing through all data points (Bojanov *et al.* 1993). Linear interpolation estimates values by assuming a constant linear relationship among known data points. These three methods are investigated for completeness and robustness of results.

Interpolation Uncertainty Estimation

This study builds upon prior research that investigated the estimation of uncertainty in interpolated bathymetric models (Jakobsson *et al.* 2002, Elmore *et al.* 2012, Amante and Eakins 2016, Amante 2018, Bongiovanni 2018). Other works (e.g., Calder 2006 and, Bongiovanni et al. 2018) used the geostatistical Kriging interpolation method as this gives an estimate of uncertainty. However, besides the underlying associated assumptions of Kriging (e.g., stationarity and isotropy), the method is not ideal for data-driven operational hydrography because it is computationally— and time— intensive.

Previous studies have found that the accuracy of all interpolation techniques is related to the terrain characteristics, sampling density, and distribution of measurements (MacEachren and Davidson 1987, Aguilar *et al.* 2005, Anderson *et al.* 2005, Chaplot *et al.* 2006, Erdogan 2009, 2010, Guo *et al.* 2010, Amante and Eakins 2016, Amante 2018,

Amoroso *et al.* 2023). These factors will be explored in this study to estimate interpolation uncertainty.

Amante (2018) and Amante and Eakins (2016) investigated the accuracy of interpolated DBMs through deterministic interpolation methods. They analyzed the relationship between interpolation deviations, sample density, distance to the nearest depth measurement, and terrain characteristics on one testbed. However, their regression models relied on a single parameter (distance to the nearest depth measurement) without considering the terrain characteristics, thus limiting the generalizability of their findings to similar terrain areas.

Amoroso et al. (2023) conducted a statistical assessment of DBMs generated from both geostatistical and deterministic interpolation methods. They considered factors such as seabed morphological complexity and sampling density across six different sectors with different seabed morphologies. Nevertheless, the limitation of sector size, fixed at 100m by 100m, poses challenges from an operational perspective because the area investigated is small. Importantly, it is worth noting that this study did not delve into the estimation of interpolation uncertainty.

The research undertaken herein is unique for its comprehensive evaluation of interpolation-based DBMs, considering both scientific and operational perspectives. This investigation is based on four different types of seabed morphologies, each covering an expansive 10km by 10km area. Advancing beyond prior research, this study incorporates terrain characteristics—specifically slope and roughness—into the cell-level uncertainty model. It also leverages machine learning techniques to explore the hidden and non-linear relationships between interpolation uncertainty and the ancillary parameters. This concerted effort aims to provide a thorough characterization of uncertainty in interpolated bathymetric datasets.

Nautical Charting Uncertainty Standards

The International Hydrographic Organization (IHO) established the S-44 IHO Standards for Hydrographic Surveys (IHO 2020) in 1968 to define data quality requirements for charting purposes, with subsequent updates reflecting technological advancements (IHO 2020). Recent shifts towards electronic navigational products, notably electronic navigational charts (ENCs), have prompted a re-evaluation of how data uncertainty and quality are communicated to mariners. The Category of Zones of Confidence (CATZOC) levels in the S-57 IHO Transfer Standard for Digital Hydrographic Data (IHO 2014) and the Quality of Bathymetric Data (QoBD) in the new S-101 ENC Product Specification (IHO 2022) outline these parameters.

Every chart is a mosaic of polygons / sectors (Ware and Kastrisios 2022), each assigned a CATZOC that represents the vertical and horizontal uncertainty and completeness of the collected data of the underlying survey (see Table 1). The CATZOC concept offers a consistent methodology of assessing data quality by end users.

Table 1: The International Hydrographic Organization (IHO) S-57 Category of Zones of Confidence (CATZOC) Levels. (IHO 2014).

CATZOC LEVEL	POSITIONAL ACCURACY	DEPTH ACCURACY	SEAFLOOR COVERAGE
A1	+/-5m + 5% depth	0.5m + 1% depth	Full area search undertaken. Significant seafloor features detected and depths measured.
A2	+/-20m	1m + 2% depth	Full area search undertaken. Significant seafloor features detected and depths measured.
B	+/-50m	1m + 2% depth	Full area search not achieved; uncharted features hazardous surface navigation are not expected but may exist.
C	+/-500m	2m + 5% depth	Full area search not achieved; depth anomalies may be expected.
D	Worse than CATZOC C	Worse than CATZOC C	Full area search not achieved, large depth anomalies may be expected.
U	Unassessed - The quality of data has yet to be assessed		

This study, as part of our overarching effort in estimating uncertainty in interpolated DBM, aims to facilitate the *a posteriori* CATZOC classification of interpolated datasets in areas where complete bathymetric seafloor coverage was not achieved. Specifically, the overarching research effort aims to address the uncertainty metric associated with depth accuracy in CATZOC, excluding, however, considerations for seafloor feature detection and factors like expected feature size or seabed undulations. Thus, possible CATZOC classification improvement through better estimation of interpolation uncertainty is generally limited to CATZOC B unless hydrographic offices' full seabed coverage and feature detection requirements for CATZOC A1 and A2 are met using side-scan sonar.

Materials and Methods

This section provides a description of the study areas under investigation and the datasets employed. Furthermore, it defines the parameters and outlines the methodologies employed to accomplish the specified objectives.

Testbeds and Datasets

The study investigated five distinct testbeds located within U.S. waters, each representing a unique seafloor morphology (Figure 1).

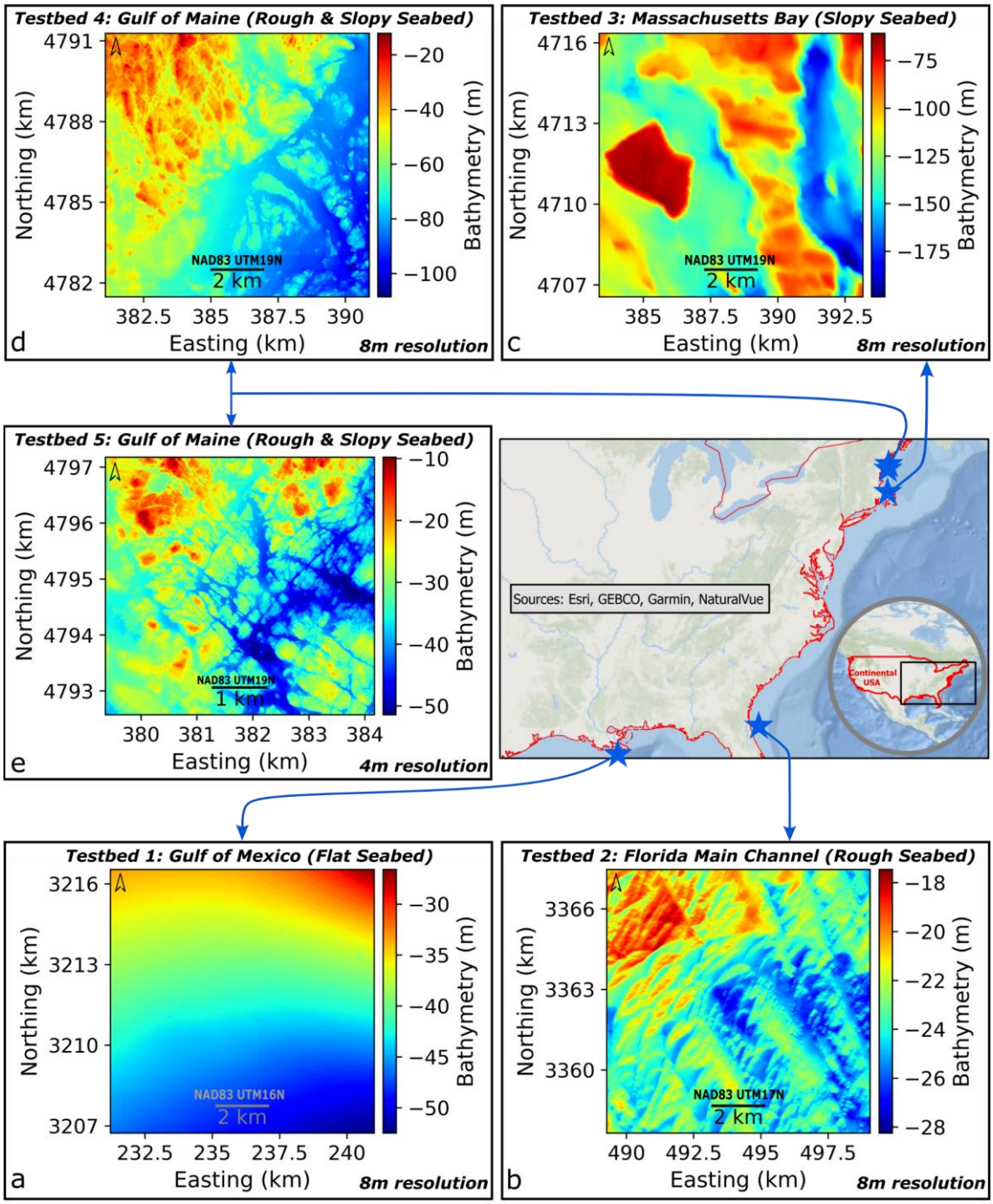

Figure 1: Five testbeds of various seabed characteristics for testing the proposed methodology in: a) Gulf of Mexico (Flat seabed), b) Florida Main Channel (Rough seabed), c) Massachusetts Bay (Slopy seabed), d) Gulf of Maine (Rough and Slopy seabed) and e) Gulf of Maine (Rough and Slopy seabed with multi-resolution).

These testbeds encompass varying combinations of slope and roughness characteristics, to facilitate a thorough analysis of factors that may impact interpolation

uncertainty. For a detailed overview of the testbeds, including their names, morphology, respective locations, and key attributes, refer to Table 2. Testbed 5 is considered "special" in the study and is used to explore the impact of variable spatial data resolution on uncertainty estimation. Furthermore, in contrast to the other testbeds, which are 10km by 10km in size, Testbed 5 has dimensions of 5km by 5km.

Table 2: Summary of metadata for the five testbeds. NAD83 - North American Datum of 1983 and UTM - Universal Transverse Mercator.

Test bed	**Name**	**Morpho-logy**	**Locality**	**Depth Range (m)**	**BlueTopo Tiles**	**NAD83 UTM Zone**	**Resolution (m)**
1	Flat	Low roughness and low slope	Gulf of Mexico, LA	-26 – -52	BF2G62KP_20230505	16N	8
					BF2G72KP_20230505		
2	Rough	High roughness and low slope	Florida Main Channel, FL	-17 – -28	BH4ST58G_20230607	17N	8
					BH4ST58H_20230607		
					BH4SV58G_20221125		
					BH4SV58H_20221125		
3	Slopy	High slope and low roughness	Massa-chusetts Bay, MA	-60 – -200	BF2JK2MD_20230614	19N	8
4	Rough and Slopy	High slope and high roughness	Gulf of Maine, ME	-12 – -102	BF2JK2MH_20230626	19N	8
					BF2JK2MG_20230418		
5	Special	High slope and high roughness	Gulf of Maine, ME	-10 – -51	BF2JK2MD_20230614	19N	4, 8 & 16

The bathymetric datasets for these testbeds shown in Figure 1 were sourced from NOAA's BlueTopo repository, horizontally referenced to their respective North American Datum of 1983 (NAD83) and Universal Transverse Mercator (UTM) zone (Table 2), and vertically referenced to North American Vertical Datum 1988 (NADV88), positive upward. These non-interpolated BlueTopo depths are designated as the "true depth" despite any uncertainty associated with these measurements. The selection of

testbeds involved a dual assessment, including qualitative visual inspection and quantitative analysis to ensure an accurate representation of respective seabed morphologies. The evaluation in the form of histograms of slope and roughness are presented in Figure 2.

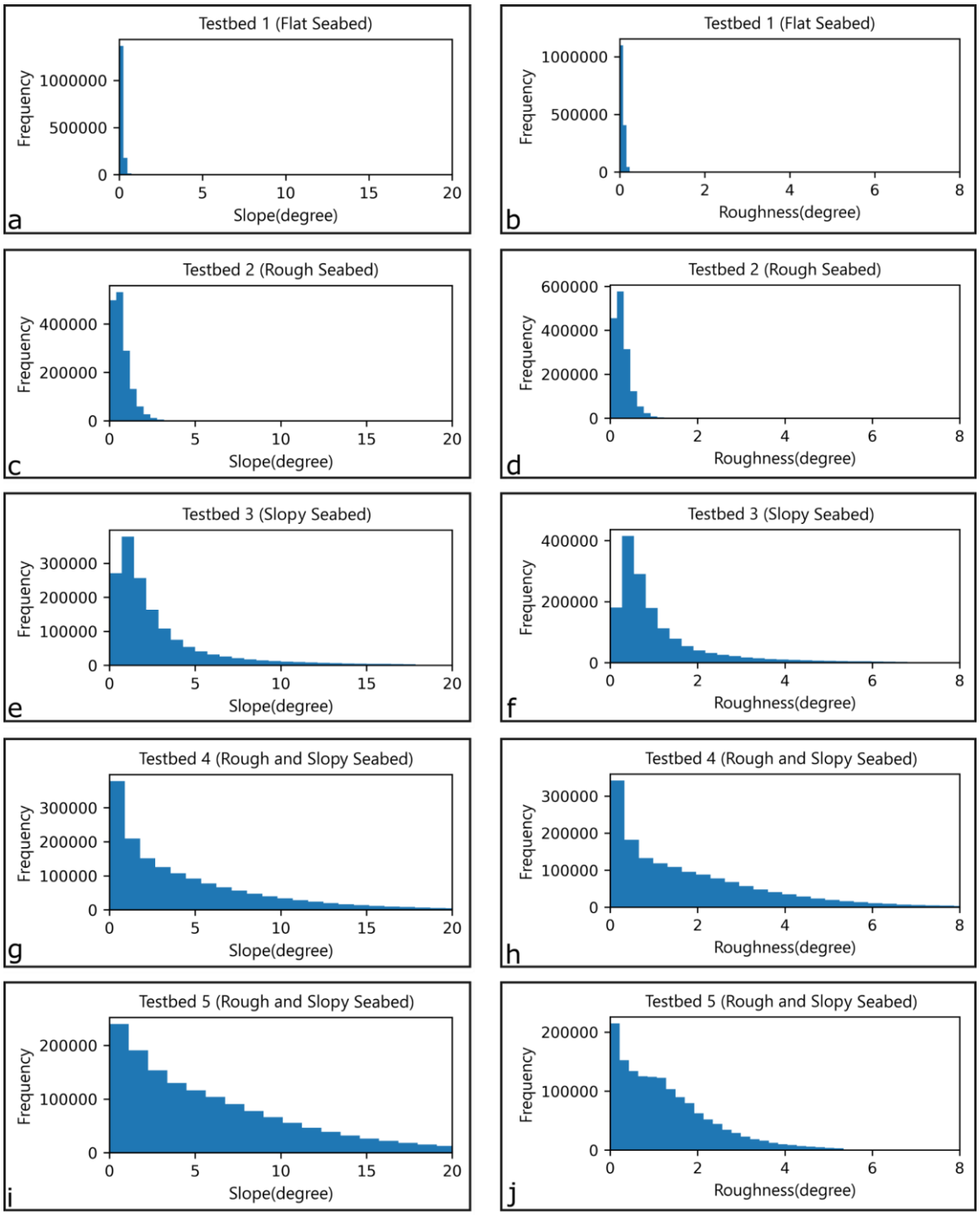

Figure 2: Histograms depicting the slope and roughness characteristics of the original testbeds' bathymetric datasets.

Panels e to j highlight the 99th percentile of data, plotted to improve visualization and facilitate effective comparisons between the seabed morphologies, while Panels a to f were generated using the entire dataset. Comparing the slope and roughness values across the testbeds, the histograms demonstrate that the testbeds capture their respective morphologies. Terrain characteristics of slope and roughness in the context of this work are defined in the following section.

Terrain Characterization

In this study, terrain characteristics, specifically Slope and Roughness, assume centrality in assessing the impact of seabed morphology on uncertainty estimation. While curvature and aspect were also initially considered, they are omitted from this work as preliminary analysis indicated that they exhibited very low correlation with interpolation uncertainty. The following sections give an in-depth explanation of what Slope and Roughness are and how they have been calculated in this work.

Slope

Slope denotes the maximum rate of depth change within a moving analysis window (Burrough and McDonnell, 1998), calculated as a function of gradients in the X direction (fx) and Y direction (fy) (equation 1):

$$\text{Slope} = \arctan\left(\sqrt{(fx)^2 + (fy)^2}\right) \quad (1)$$

To compute slope (or other terrain parameters such as roughness), an analysis window traverses the raster DBM surface, with each pixel serving as the central point for calculations (Figure 3).

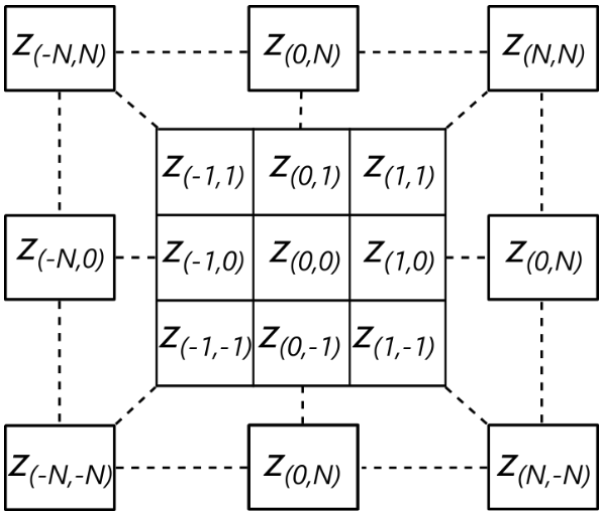

Figure 3: An n x n analysis window for a raster grid.

This generalization allows the parameter to be analysed at a range of scales (different values of $n \geq 3$) (Wilson *et al.* 2007). The Spatial Scales section delves into the discussion of the different window sizes employed in this study. Horn's finite difference method (Horn 1981) is employed to calculate fx and fy of the central point of the analysis window using appropriate convolution kernels. Horn's method utilizes convolution kernels tailored to each moving analysis window to estimate a value for a subject pixel (Horn 1981).

Roughness

Roughness, a measure of elevation variability within a sampled terrain unit, captures both highs and lows. The definition of terrain surface roughness is dependent on the scale considered, and the unit of measurement across a scale window is crucial. In this study, roughness is calculated as defined by Wilson et al. (2007), which is the largest inter-cell difference of a central pixel and its surrounding cells in an n x n rectangular neighbourhood (refer to Figure 3).

Spatial Scales

Spatial scales in this study refer to the resolutions and sizes at which bathymetric data

and terrain characteristics are analysed. The computation of the terrain attributes, slope and roughness, results in a range of values depending on data resolution and analysis window size. This consideration is crucial for understanding how different features manifest at various levels of granularity and the impact of spatial scale on interpolation uncertainty estimation. The study navigates spatial scales by employing 3-by-3, 5-by-5, and 7-by-7 pixels window sizes. While the 3-by-3 window is used for most parts of the study (since it is standard window in GIS software), the other two sizes are selectively employed to investigate the influence of window sizes on uncertainty estimation, specifically on Testbed 1. Additionally, the impact of spatial resolution on interpolation uncertainty is investigated with the support of Testbed 5 due to its availability of multiple spatial resolutions.

Interpolation Techniques

This study quantified uncertainty in interpolated bathymetric datasets generated from three common deterministic interpolation techniques, i.e., IDW, Spline, and Linear, across five distinct test beds in the United States. The algorithms for these interpolators are developed in Python such that it requires minimal adjustment of interpolation parameters. This deliberate approach aims to streamline the implementation process and enhance the applicability of the developed algorithms thereby eliminating dependence on specific GIS software. Consequently, the optimization of interpolation parameters was not within the scope of this work, and the optimal IDW parameters identified by Amante and Eakins (2016) were employed.

Implementation of Split-Sample Methodology

The split-sample methodology is employed to randomly sample the testbeds' depth measurements based on sampling density. Sampling density in prior studies had varied

definitions, including a percentage of original measurements (MacEachren and Davidson 1987, Aguilar *et al.* 2005, Anderson *et al.* 2005, Guo *et al.* 2010, Alcaras *et al.* 2022), a count of measurements per area (Chaplot *et al.* 2006, Erdogan 2009, 2010), or a percentage of DBM grid cells constrained by depth measurements (Amante and Eakins 2016, Amante 2018). This study adopts the latter definition utilizing five sampling densities of 50%, 25%, 10%, 5%, and 1%.

Each testbed's dataset serves as the environmental canvas and is divided into training (depths), earmarked for interpolation, and test data, earmarked for uncertainty quantification and error analysis (Figure 4). The split-sample methodology undergoes 10 iterations for each interpolation technique, at each of the five cell sampling densities, and on the five test beds. This 10 times repetition aims to capture the bathymetric variability of each testbed morphology and to prevent bias in the estimation of interpolation uncertainty. The choice of 10 iterations considers both algorithm processing time and memory usage, deeming it sufficient to fairly represent bathymetric variability in a 10 km-by-10 km area.

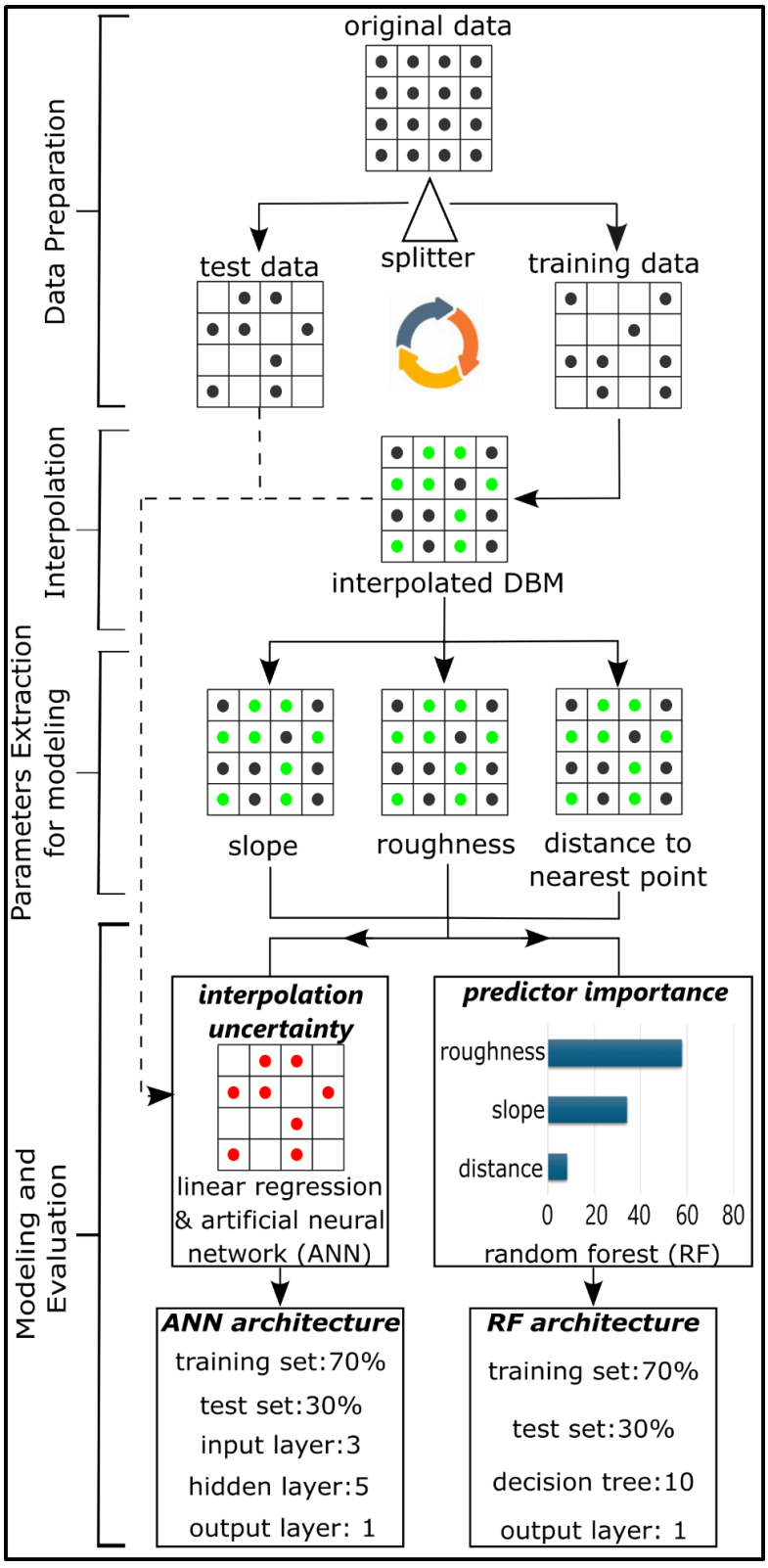

Figure 4: Workflow for quantifying interpolation uncertainty, including the machine learning techniques' architectures.

The process, illustrated in Figure 4, involves random splitting of original data during each iteration and training depths are gridded using the specified interpolation technique. The resulting interpolated raster is then compared, on a cell-by-cell basis, to the test depths to quantify interpolation uncertainty. Ancillary parameters, such as the Euclidean distance from the nearest measurement (i.e., distance of a cell from a cell of known depth) are generated from the training depths, while slope and roughness raster surfaces are generated from the interpolated DBM after each split-sample routine. These parameters, compiled from each split-sample routine, along with their respective interpolation uncertainties, constitute the datasets prepared for modelling to discern the relationship between the parameters and interpolation uncertainties, and determine the most important predictor of interpolation uncertainty.

To mitigate potential bias in the analysis, precautionary measures are implemented. Specifically, while depth measurements are interpolated across the entire area, the analysis is confined to a subset reduced by 12 cells (96m) on each side. This approach aims to address edge effects by ensuring that interpolated values near the boundaries of the study area are influenced by actual measured depths only, thereby minimizing inaccuracies that could arise from extrapolation. This ensures that the edge of the area of analysis has a range of distances to the nearest measurement without biasing toward depths along the outermost border of the study area. It also guarantees accurate slope and roughness values by maintaining appropriate window size for calculation.

Modelling and Analysis

Linear regression is used to model the relationship between interpolation uncertainty and individual parameters, producing the targeted estimated uncertainty. For each testbed, linear regression models are fitted for each of the three ancillary parameters – distance to

the nearest measurement, slope, and roughness – individually for all sampling densities and interpolation techniques. The accuracy of these models is evaluated using the metrics adjusted coefficient of determination R^2 and root-mean-square error (RMSE) in m.

Moving beyond individual parameters, the study employed artificial neural network (ANN) techniques (see Agatonovic-Kustrin and Beresford 2000) to capture non-linearities, interactions, and hidden relationships within the combined parameters. To assess the importance of uncertainty predictors, this study utilized the random forest (RF) technique because of its computational efficiency. Both ANN and RF techniques were implemented in Python and their architectures are presented in Figure 4. The evaluation of the ANN model involved assessing its predicted uncertainty for the test data through adjusted R^2 and RMSE. The importance of each parameter in the RF model's prediction was assessed, considering the number of trees in which a variable appeared.

To determine the significance of variables' contributions, a bootstrap approach was used. The RF model underwent fitting to 500 bootstrapped datasets. The resulting variable importance values from each RF model were aggregated to calculate the mean importance and establish 95% confidence intervals. By averaging across all sampling densities and interpolation methods, we computed the overall mean importance and corresponding 95% confidence intervals for each variable.

Validation Technique/Accuracy Assessment

Two established validation methods for evaluating interpolation accuracy are cross-validation (Davis 1987, Tomczak 1998, Erdogan 2009, Amante and Eakins 2016) and the split-sample method (Voltz and Webster 1990, Declercq 1996, Lloyd and Atkinson 2002, Amante and Eakins 2016). This study adopts the split-sample method, the commonly used method to assess changes in the accuracy of an interpolation technique when using

various sampling densities. The split-sample method involves dividing the dataset into training and test subsets, using the former for interpolation and evaluating performance with the latter. Common statistical measures such as RMSE, MAE, bias, and coefficient of determination (R^2) are widely employed for evaluation (Isaaks and Srivastava 1989, Zar 1999, Li and Heap 2008). RMSE and R^2 are used in this study.

To evaluate the statistical significance of differences in interpolation methods, we utilized pairwise t-tests. These tests specifically aimed at evaluating differences between pairwise interpolation methods uncertainties and comparing them with zero. The pairwise difference distributions were normally distributed, corroborating the findings of t-tests. In addition to the statistical assessment, the accuracy of interpolation methods was spatially evaluated based on visual inspection to comprehend the spatial distribution of interpolation uncertainty. This is an effort to characterize the interpolation uncertainties qualitatively.

Results

Interpolation Methods

For each testbed, the performance of interpolation methods at each sampling density is assessed through box and whisker plots. In general, Spline demonstrated superior performance compared to IDW and Linear with the exception of Testbed 1 where Linear outperformed. IDW consistently exhibited the worst performance. For a representation of the interpolation methods' performance across all sampling densities for Testbeds 1 and 4, refer to Figures 5 and 6. These specific testbeds were selected due to their notable variation in uncertainty magnitude, encompassing the range of uncertainties observed in the other testbeds.

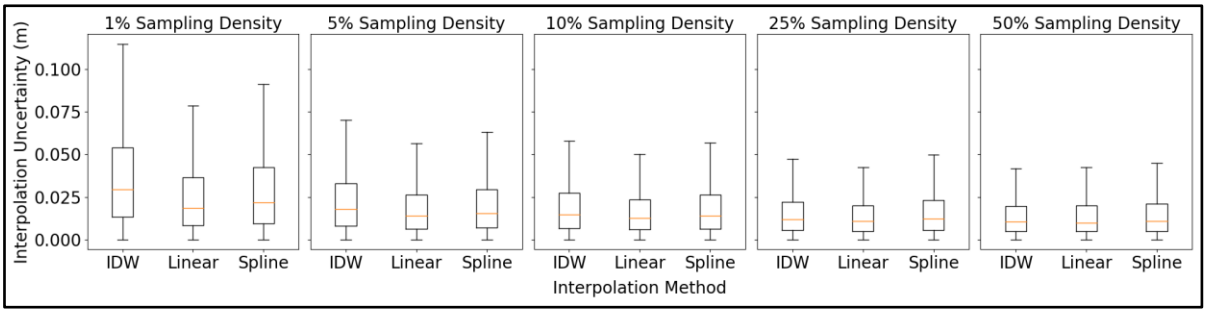

Figure 5: Interpolation methods uncertainty comparison on Testbed 1(Flat) using box and whisker, plotted with 99% percentile of data.

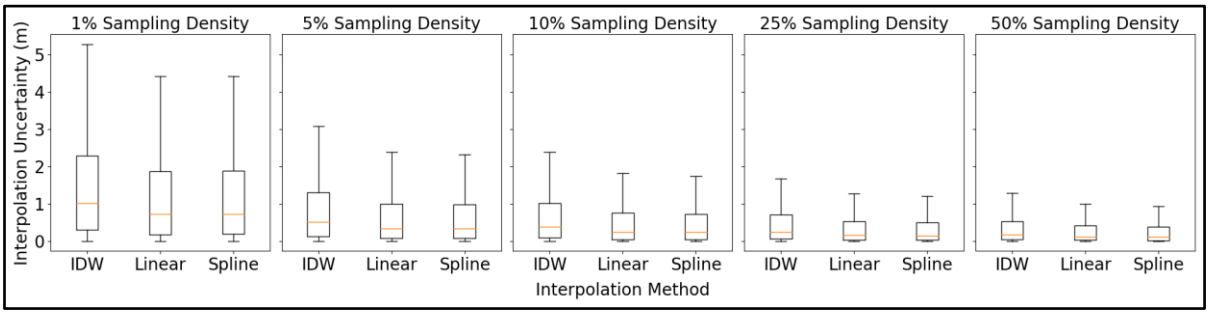

Figure 6: Interpolation methods uncertainty comparison on Testbed 4 (Rough and Slopy) using box and whisker, plotted with 99% percentile of data.

However, it is essential to note that the distinctions in interpolation methods at each sampling density across all testbeds are not statistically significant; Figure 7 and Table 3 showcase pairwise interpolation uncertainty difference distribution and t-test results for Testbed 4, representative of other testbeds.

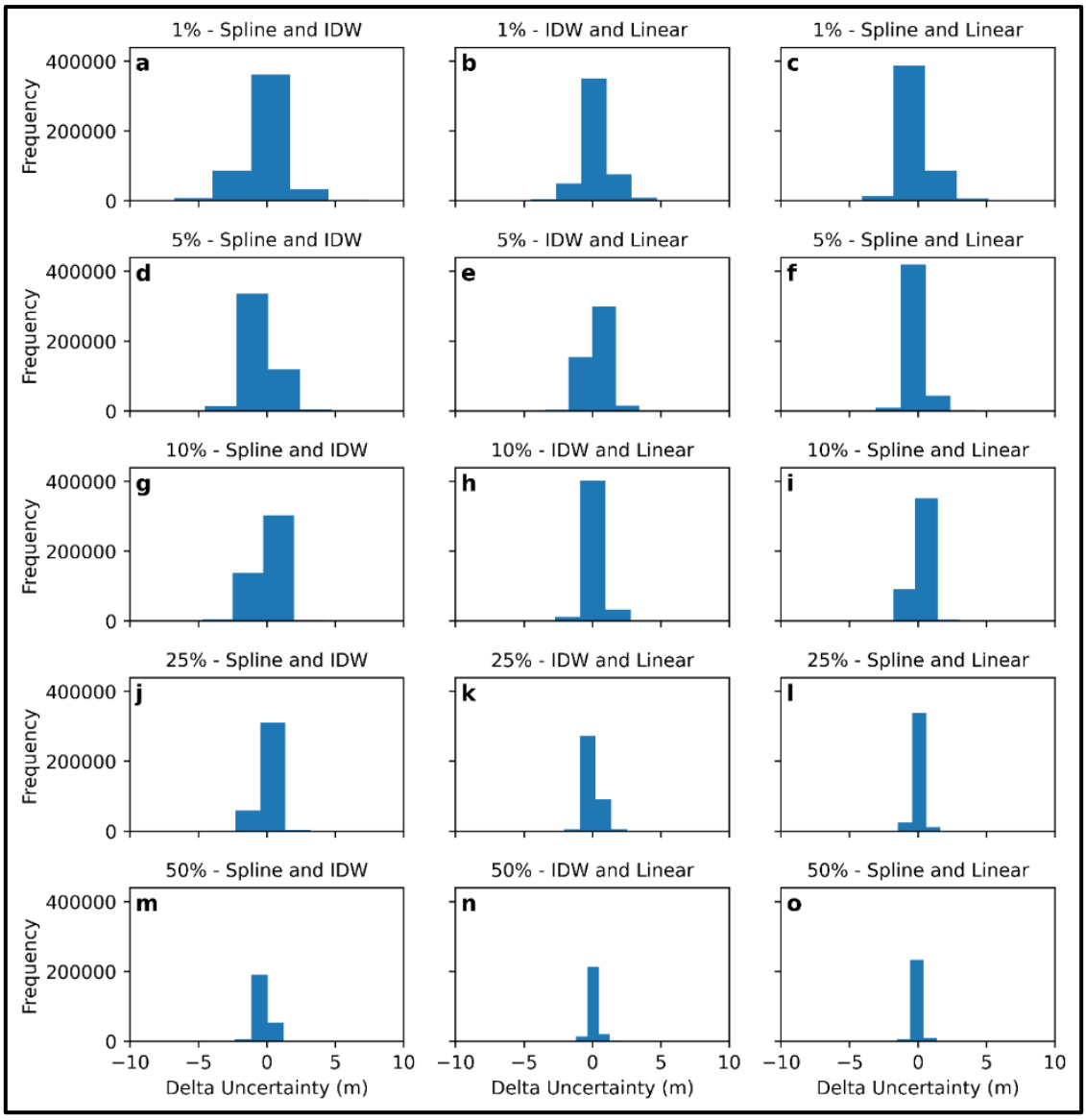

Figure 7: Histograms showing pair-wise differences of interpolation methods' uncertainties for Testbed 4 (Rough and Slopy) at various sampling densities.

Table 3: Testbed 4 (Rough and Slopy) Statistics for pair-wise interpolation methods comparison at various sampling densities.

Sampling Density (%)	Interpolation Methods	t-statistics	p-value
1	Spline and IDW	-91.9	0
	IDW and Linear	158.9	0
	Spline and Linear	36.1	0
5	Spline and IDW	-141.0	0

	IDW and Linear	184.3	0
	Spline and Linear	-2.8	< 0.01
10	Spline and IDW	-150.9	0
	IDW and Linear	183.6	0
	Spline and Linear	-9.4	0
25	Spline and IDW	-148.7	0
	IDW and Linear	161.3	0
	Spline and Linear	-31.5	0
50	Spline and IDW	-115.8	0
	IDW and Linear	111.1	0
	Spline and Linear	-35.0	0

The magnitude of interpolation uncertainties is in the order of centimetres for Testbeds 1 (Figure 5) and 2, while it is in the order of meters for the more complex Testbeds 3 and 4 (Figure 6). Moreover, differences between uncertainties are relatively higher at lower sampling densities across all testbeds, while at higher sampling densities, uncertainties tend to become the same.

Spatial Pattern of Interpolation Uncertainties

The spatial distribution of interpolation uncertainties is visually examined across Testbeds 1 to 4, excluding Testbed 5, which is specifically designated for assessing the influence of spatial resolutions on uncertainty estimation. The interpolation uncertainties follow specific patterns and are not random at each of the sampling densities across all testbeds. For Testbed 1, uncertainties are concentrated on the eastern side due to

multibeam artifacts in the original datasets. Testbed 2 shows that higher uncertainties correlate with areas of high roughness values. In Testbed 3, higher uncertainties are related to high slope values. Testbed 4 exhibits non-random uncertainties, with higher values correlating with both high slope and roughness values. Given the higher uncertainty magnitude in Testbed 4 compared to others, and the peculiarity of the Testbed 1 dataset with multibeam artifacts, their spatial uncertainty plots are provided in Figure 8 and Figure 9, respectively. To facilitate meaningful comparison, the colour bar (key) has been standardized across sampling densities and interpolation methods. The differences in the subplots at identical sampling densities, i.e., horizontally, are not apparent due to the comparable performance of the interpolation methods.

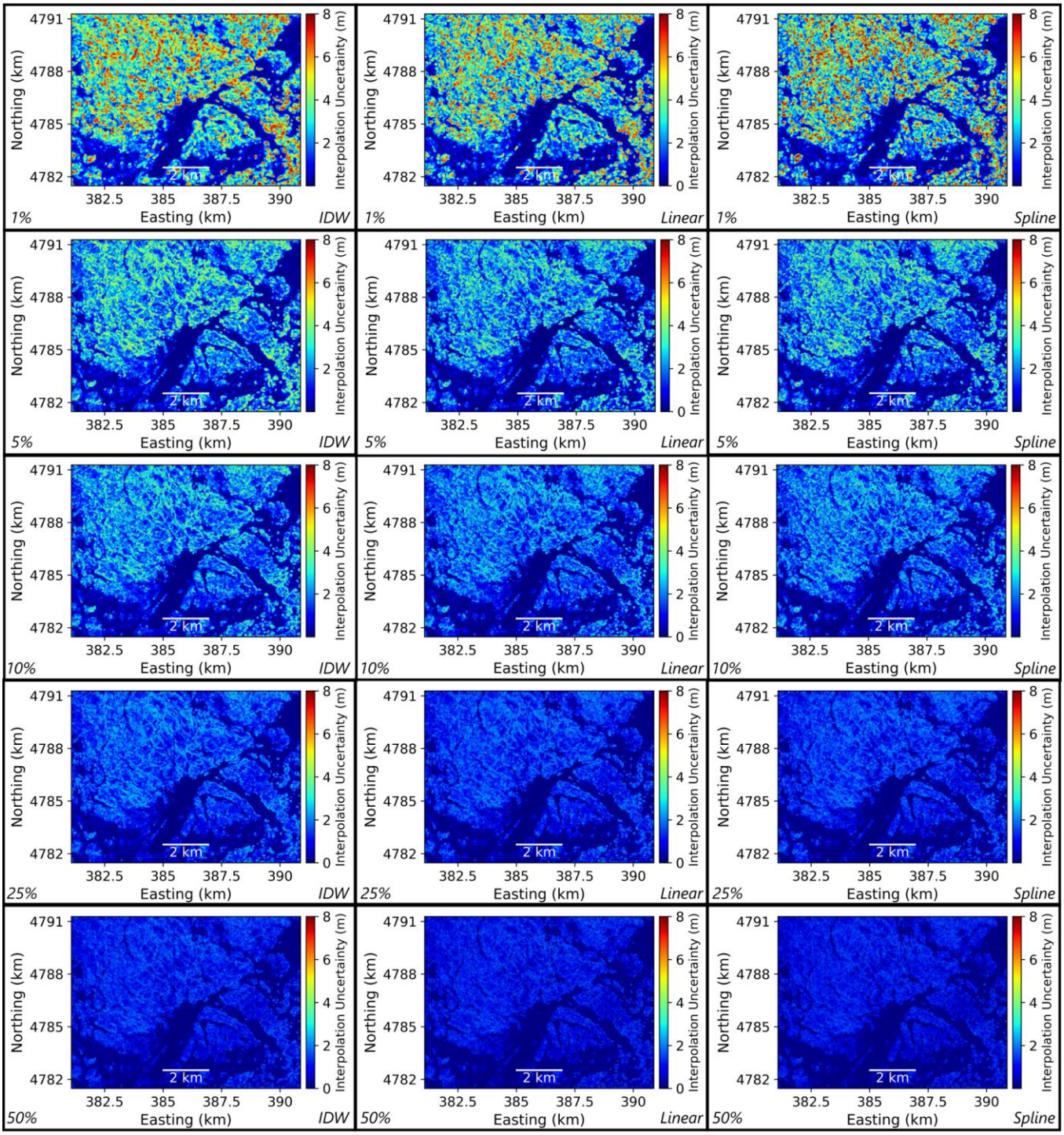

Figure 8: Interpolation uncertainty on Testbed 4 (Rough and Slopy) across all sampling densities (columns) and interpolation methods (rows), using 99th percentile of data.

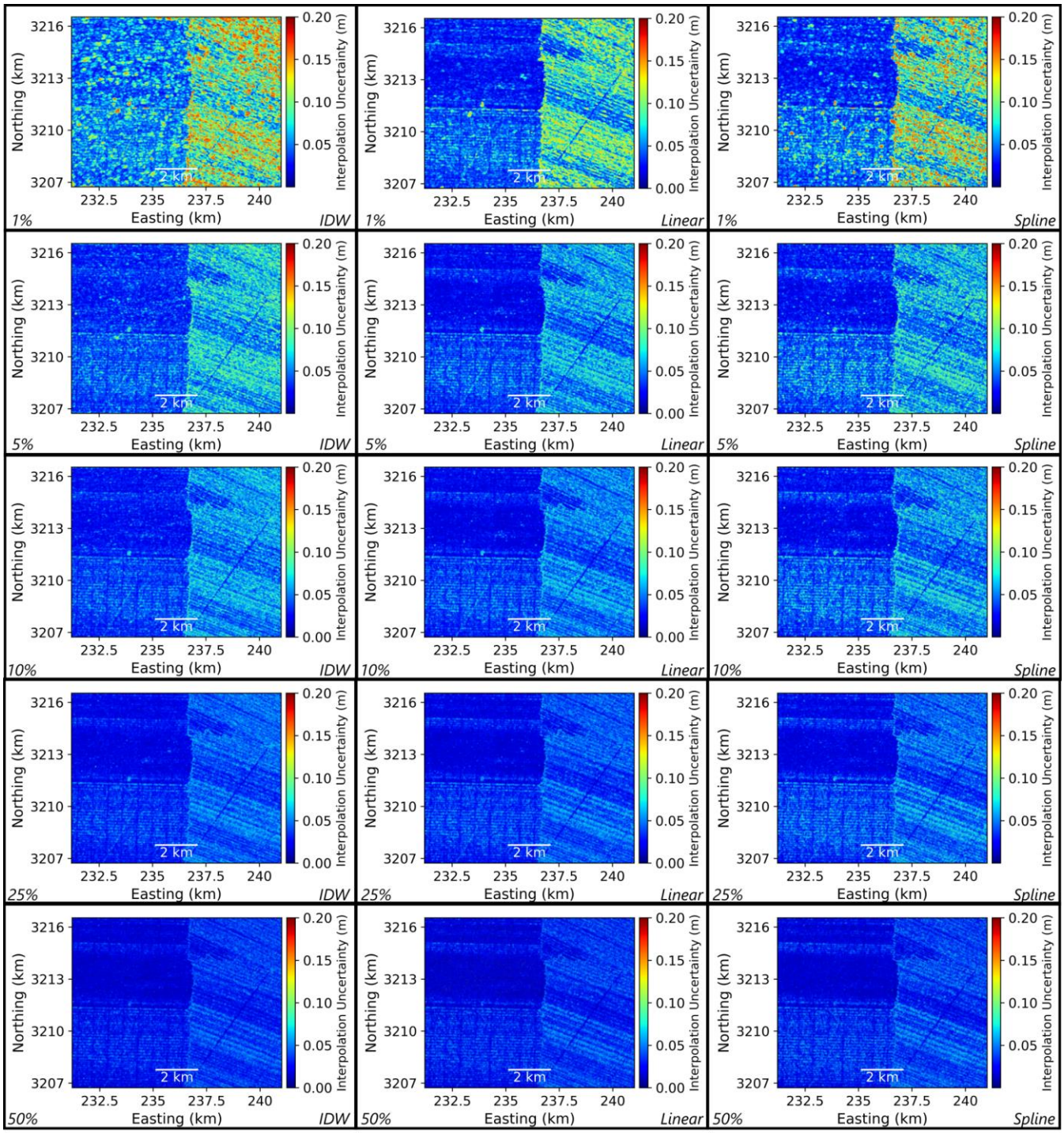

Figure 9: Interpolation uncertainty on Testbed 1 (Flat) across all sampling densities (columns) and interpolation methods, using 99th percentile of data.

Predictive Models of Interpolation Uncertainties

Generally, across the testbeds, the linear regression relationships between interpolation uncertainty and individual ancillary parameters (i.e., distance to the nearest measurement,

slope, and roughness) were weak for all interpolation methods. To investigate the non-linearities as well as hidden and interactive relationships between the combined parameters and the estimated uncertainty, we employed ANN. This slightly improved the adjusted R^2, although the accuracy of these combined parameters' predictive models for estimating uncertainty varies significantly across the testbeds, with the best performance observed in Testbed 4, followed by 3, 2, and 1.

The adjusted R^2 of IDW's predictive uncertainty model generally outperformed others except on Testbed 1 where Spline performed better. Nevertheless, the differences between the interpolation predictive uncertainty models are only in the order of centimetres. Performance improvements were observed at higher sampling densities across all testbeds, with disparities becoming negligible beyond a 10% sampling density. To illustrate the above, Figure 10 and Figure 11 respectively present the individual and combined parameters' interpolation predictive uncertainty models for Testbed 4, the best-performing testbed.

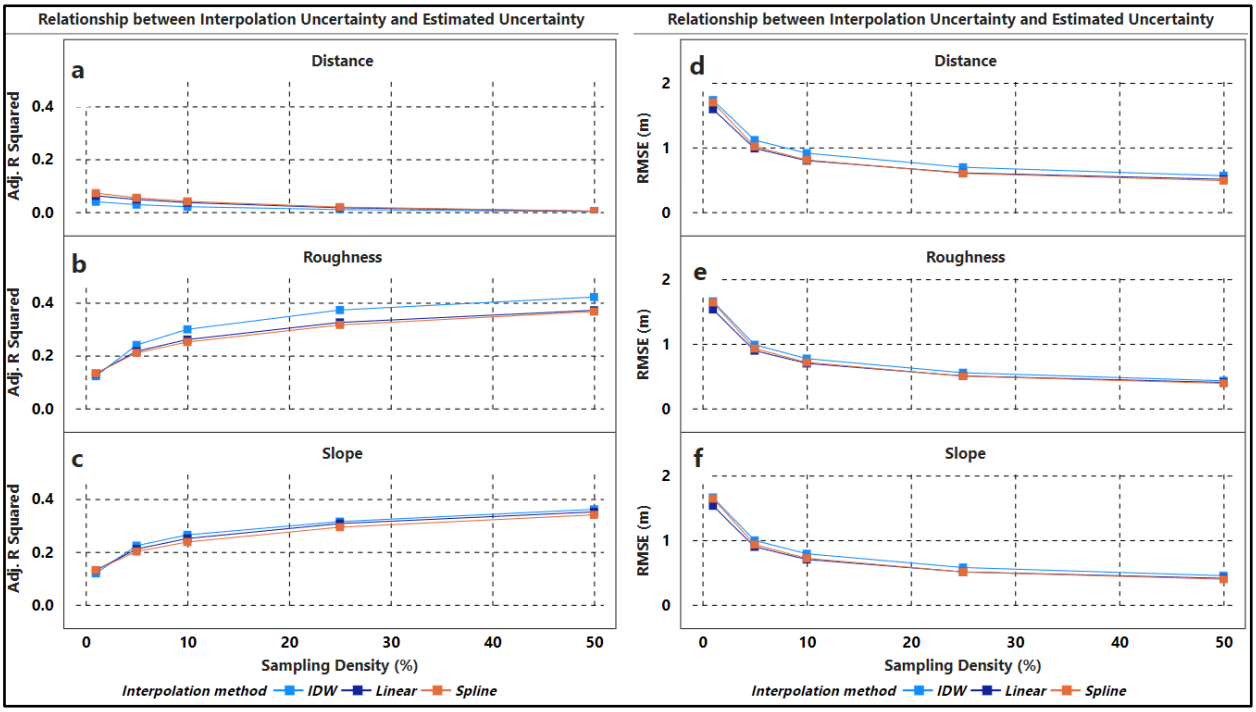

Figure 10: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to the nearest measurement, roughness, and slope respectively for Testbed 4 (Rough and Slopy).

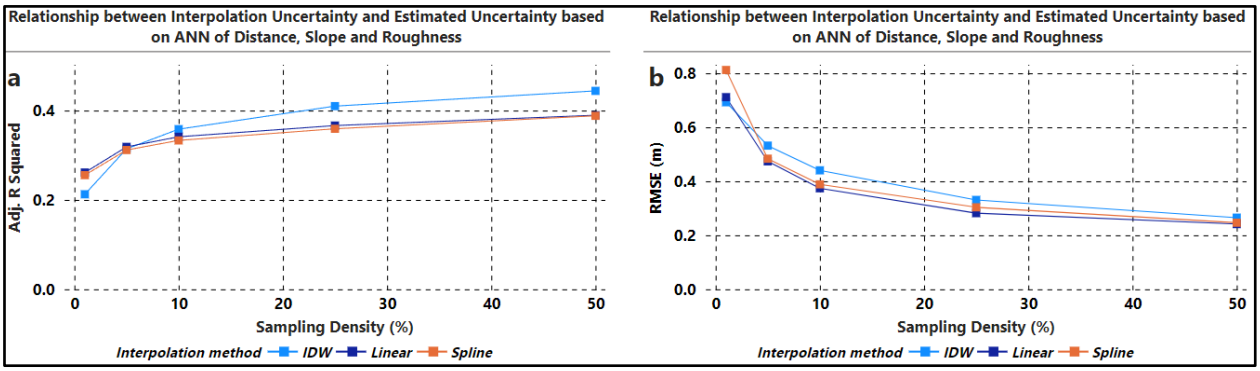

Figure 11: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to the nearest measurement, roughness, and slope combined for Testbed 4 (Rough and Slopy).

Results from the RF analysis coupled with the bootstrap statistical technique, presented in Table 4, highlighted roughness as the best predictor of uncertainty, followed by slope and distance to the nearest measurement.

Table 4: Statistics of the Importance of Predictors of Uncertainty across testbeds

Testbed	Statistics/Parameter	Distance	Slope	Roughness
1	Mean Importance (%)	8.2	46.5	45.3
	# times of most Important	0	5	10
	95% bootstrap percentile CI of the Importance	(7.3,9.6)	(42.0,49.9)	(41.8,48.9)
2	Mean Importance (%)	10	43.9	46.1
	# times of most Important	0	5	10
	95% bootstrap percentile CI of the Importance	(9.6,10.3)	(43.3,44.6)	(45.4,46.7)
3	Mean Importance (%)	8.2	34.200	57.6
	# times of most Important	0	0	15
	95% bootstrap percentile CI of the Importance	(7.4,9.3)	(29.7,40.6)	(51,62.1)
4	Mean Importance (%)	9.5	33.1	57.4
	# times of most Important	0	0	15
	95% bootstrap percentile CI of the Importance	(8.9,10.2)	(31.4,35.6)	(54.9,59.1)

Window Sizes

The investigation of the impact of window sizes on uncertainty estimation focused on Testbed 4, the best-performing testbed, with an 8m pixel size resolution. The R^2 of the ANN relationship between estimated uncertainty and combined parameters shown in Figure 12 revealed that the 3-by-3 window initially demonstrated the weakest performance compared to other windows from 1% to 10% sampling density but outperformed them at 25% and 50% sampling density. The 5-by-5 window consistently secured the second position, while the 7-by-7 window performed optimally from 1% to 10% sampling density but exhibited the lowest performance at 25% and 50% sampling density. These trends of improved estimates from a larger window size were consistent

across all interpolation methods for Testbed 4.

Examining RMSE, the 3-by-3 window outperformed other window sizes at all sampling densities except 50% where it secured the second position with IDW. The 5-by-5 and 7-by-7 windows demonstrated similar performance across various sampling densities for IDW. For Linear and Spline interpolations, all three window sizes exhibited comparable performance levels across different sampling densities.

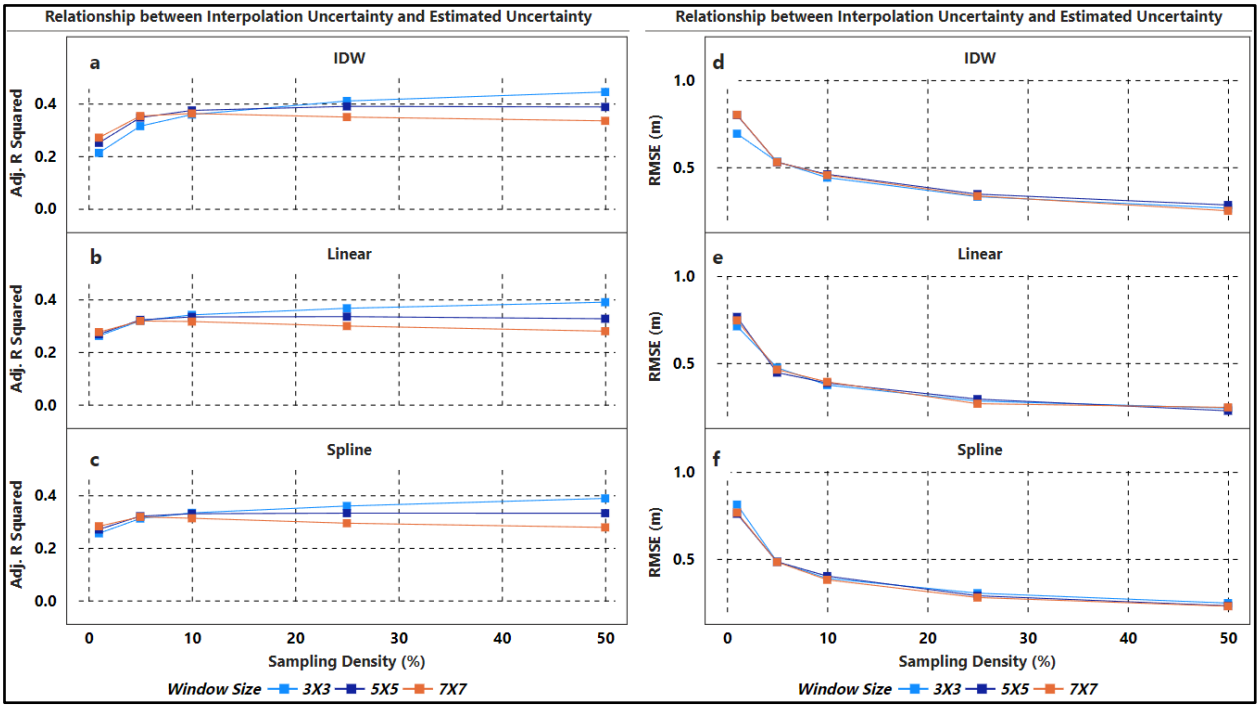

Figure 12: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on window sizes using Testbed 4 (Rough and Slopy).Data Resolutions

Figure 13 presents the outcomes of our investigation into the impact of data resolutions on interpolation uncertainty, focusing on Testbed 5, the special Testbed with 4m, 8m, and 16m resolutions. Utilizing a standard window size of 3-by-3 and employing ANN to assess the relationship between estimated uncertainty and combined parameters, the analysis based on R^2 revealed that the 4m resolution performed the best, with increasing performance observed with higher sampling densities across all interpolation methods.

All three interpolation methods exhibited comparable performance for all resolutions and the 8m resolution performed next to the 4m resolution, followed by 16m resolution. Examining RMSE, the 4m resolution outperformed the 8m and 16m resolutions, with decreasing RMSE as sampling density increased.

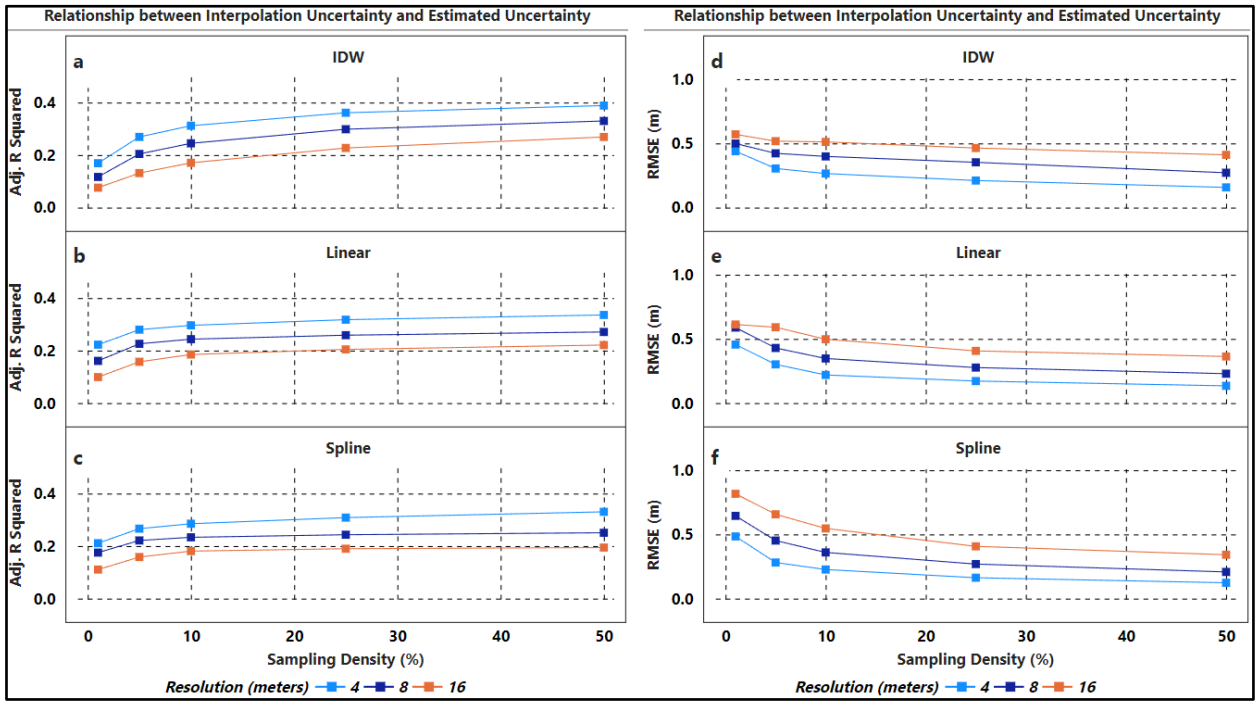

Figure 13: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on data resolution using Testbed 5.

Discussion

Unravelling the Contextual Performance of Interpolation Methods

The analysis of results revealed that interpolation methods when examined from a scientific perspective i.e., to determine the interpolation method that produces the lowest uncertainty. Linear interpolation proved most effective for Testbed 1 (Flat), while Spline is preferred for Testbeds 2 (Rough), 3 (Slopy), and 4 (Rough and Slopy). However, these methods demonstrated comparable performance at the same sampling density on each testbed, particularly when assessed from an operational standpoint (i.e., differences of a

few centimetres, that may not affect the CATZOC allocation due to the depth total vertical uncertainty shown in Table 1). The absence of statistically significant differences in their performance suggests operational equivalence among the interpolation methods.

Testbeds 1 and 2 exhibited uncertainties in the order of centimetres due to their relatively less complex morphologies. In contrast, Testbeds 3 and 4, featuring more complex morphologies, show uncertainties in the order of meters. To facilitate analysis and enhance result comprehension in this study, interpolation uncertainties are presented as absolute values. Importantly, it is worth noting that the same outcomes arise when expressing uncertainty as a percentage of depth, a crucial consideration for CATZOC values in nautical charting. The observed variability in uncertainties indicates that seabed complexity significantly influences interpolation uncertainties and highlights that Testbeds 1 and 2 will yield more accurate results compared to Testbeds 3 and 4 from an operational standpoint, particularly in the context of constructing DBM for nautical charting.

The impact of sampling density is conspicuous, particularly in Testbeds 3 and 4, where uncertainties become relatively similar at higher sampling densities. In summary, interpolation method performance varies across Testbeds, influenced by seabed morphology and sampling density. While nuanced differences exist, the overall insights suggest that selecting an interpolation method should be tailored to the specific characteristics of the seabed under consideration.

Moreover, the spatial analysis of interpolation uncertainties across Testbeds reveals non-random patterns. Concentrations of uncertainties in specific regions, such as the eastern side of Testbed 1 (see Figure 9) and areas correlated with high slope and/or high roughness values in Testbeds 2, 3, and 4, underscore the impact of underlying terrain characteristics. The presence of multibeam artifacts in Testbed 1 highlights the transfer

of uncertainties from original data to the interpolation process. Even though survey data can be within the IHO specifications it was targeted for, they can still be affected by multibeam artifact (Hughes Clarke *et al.* 1996). These findings emphasize the relationship among data quality, seabed morphology complexity, and the resulting interpolation uncertainties.

Relationship between Parameters, Sampling Density, and Interpolation Uncertainty

The observed weak association among roughness, slope, and distance with interpolation uncertainty using ANN highlights the intricacy involved in estimating uncertainty in interpolated bathymetry. Through varied sampling densities and across diverse test beds, roughness emerged as the most important predictor of uncertainty, followed by slope and distance to the nearest measurement. It is important to note that machine learning improves the predictive accuracy of the model but only in a small way. Roughly only 40% of the variability in the data is explained with the combined predictors at 50% sampling density, as Figure 12 illustrated.

Distance to the nearest measurement, the least important predictor of uncertainty, makes the minimum contribution to the overall estimation as indicated by the linear regression result and further confirmed by RF analysis. A previous study by Henrico (2021b) corroborated the findings of this work, reporting a weak correlation in using distance to the nearest measurement for depth estimation at sampling densities below 100%. Conversely, Amante and Eakins (2016) noted a robust correlation between distance and interpolation uncertainty, irrespective of sampling density.

Notably, the introduction of slope and roughness from interpolated depths as uncertainty estimators is a novel contribution, without comparative studies. The extraction of the terrain parameters, i.e., slope and roughness, from the interpolated

surfaces is deliberate and justifiable because it reflects the reality of real-world operational settings where bathymetric products are generated from gap-filled datasets. Furthermore, terrain parameters are heavily scale dependent and this work attempted to represent these parameters because formally, 'true' slope or roughness does not exist. Despite their importance, their combined explanatory power is limited, suggesting the presence of unaccounted-for factors influencing uncertainty or indicating a strong random component within interpolation uncertainty. The high correlation between slope and roughness might contribute to their marginal improvement of the uncertainty estimation. This weak relationship underscores the complexity of the problem, signalling the potential necessity for alternative approaches beyond those applied in our study. While additional parameters were considered, such as morphological aspects and curvature, they exhibited no discernible relationship with interpolation uncertainty and, therefore, are not included in this study.

The diminishing effectiveness of the model at sparser sampling densities underscores the importance of a well-distributed sample network. Higher sampling density corresponds to decreased interpolation uncertainty, and vice-versa. Additionally, as sampling density decreases, the uncertainty model struggles to capture the subtle variations in seabed morphology, leading to a reduction in predictive accuracy. However, the disparities in model performance become negligible beyond a 10% sampling density, suggesting that 10% is an adequate sampling density for estimating interpolation uncertainty. This finding accentuates the significance of strategic sampling designs, especially in areas with sparse data. This has implications for real-world hydrographic surveys that will have a linear sampling pattern rather than the spatially random sample employed here to increase fundamental understanding of various data and methodological issues.

Examining Disparities in Test Bed Predictive Performance of Interpolation Methods

The significant variations in predictive performance observed across test beds underscore the context-specific nuances of interpolation uncertainty. Notably, Testbed 4 (Rough and Slopy) exhibited the highest predictive accuracy, likely attributed to the heightened spatial variability inherent in Rough and Slopy seabeds, contributing to a more robust model fit. The incorporation of slope and roughness as parameters may have further influenced the superior performance observed in Testbed 4.

Surprisingly, Testbed 3 (Slopy) followed closely in performance, challenging expectations based on the identified significance of roughness as the primary predictor of uncertainty throughout the study. This divergence suggests that other unexplored factors may contribute to the observed results, indicating the complexity of the relationship between parameters and predictive performance.

Contrary to expectations, Testbed 2 (Rough) did not exhibit a performance level proximate to Testbed 4 (Rough and Slopy), highlighting a potential dissociation between seabed characteristics and predictive accuracy. This incongruity emphasizes the multifaceted nature of the relationship, surpassing a simple parameter-bed type association.

Finally, Testbed 1 (Flat) displayed a lower predictive performance, hinting at the challenge of capturing variability in less complex terrains. This finding underscores the importance of considering the specific characteristics of the seabed when developing and applying interpolation models, and recognizing the intricate interplay of factors influencing predictive accuracy in diverse seabeds.

The findings from predictive accuracy of the uncertainty models are that IDW provides a more accurate quantification of interpolation uncertainty than Spline and

Linear interpolation methods for most testbeds. This superior performance of IDW's uncertainty model could have been influenced by distance to the nearest measurement being a predictor of uncertainty. It is also plausible that while Spline and Linear methods are good for interpolation, their efficacy in quantifying the associated uncertainty might be comparatively limited.

Importance of Predictors

An in-depth analysis of RF, complemented by the bootstrap statistical technique, revealed that roughness holds the highest predictive importance, followed by slope and distance to the nearest measurement. Across all test beds, the consistently low importance of distance to the nearest measurement implies its limited contribution to the predictive models. In contrast, the dominance of roughness, followed by slope, underscores the significance of terrain characteristics in driving interpolation uncertainties.

Notably, on Testbeds 1 and 2, slope appeared to compete with roughness, potentially due to the less complex morphology of these seabeds, occurring as the most important in five out of 15 instances. Additionally, the roughness and slope values for these test beds are relatively low. Conversely, on Testbeds 3 and 4, roughness took dominance, possibly owing to their complex nature. The application of the bootstrap statistical technique reinforces the reliability of these findings, affirming the statistical significance of observed differences among predictors.

These results collectively contribute to a comprehensive understanding of the influential factors in predictive models of interpolation uncertainties in marine environments. Understanding the hierarchical importance of predictors provides valuable insights into how uncertainty in interpolated bathymetry can be quantified.

Impact of Window Size on Interpolation Uncertainty

In the context of sampling density, the choice of spatial scale is crucial, as different window sizes perform differently depending on the density of measurement points and unique seabed characteristics. Both R^2 and RMSE trends highlight the sensitivity of interpolation uncertainty to spatial scales, emphasizing the importance of careful window size selection. The 3-by-3 window shows better performance at higher sampling densities, suggesting nuanced smaller-scale seabed features in Testbed 4. However, the effectiveness of smaller window sizes depends on the interplay between seabed morphology and sampling density. On the other hand, the 7-by-7 window performs better at lower sampling densities but falters at higher densities, indicating its proficiency in capturing features effectively with fewer measurements. This effectiveness diminishes as sampling density increases, revealing the larger window's decline in performance due to increased variability. The interrelation among window size, sampling density, and seabed complexity highlights the need to carefully choose an optimal window size, considering the multifaceted factors at play. Larger windows capture broader trends in sparser data scenarios but become less effective in denser data environments where smaller-scale features and variations dominate. This suggests that a larger window size is preferable for interpolating low sampling density data in Testbed 4. However, the generalizability of this conclusion to other testbeds requires further investigation.

Impact of Data Resolutions on Interpolation Uncertainty

Expectedly, our examination of the impact of data resolutions on interpolation uncertainty underscores the significance of higher resolutions in enhancing uncertainty estimation, providing a more detailed depiction of the seabed. Beyond solely quantifying the effect of data resolution on estimating interpolation uncertainty, this investigation is pertinent

to the broader context of quantifying uncertainty in hydrographic databases that encompass datasets with variable resolutions. This insightful exploration enriches our comprehension of the complexities inherent in estimating uncertainty in DBMs.

Limitations of the Study

Our study focuses on the estimation of interpolation uncertainty in operational settings, where data-driven products are derived from extensive datasets. Consequently, geostatistical interpolation methods such as Kriging, which provide uncertainty estimates, were omitted due to their computationally intensive nature and demand for substantial memory resources.

Furthermore, our study utilized general interpolation parameters derived from existing literature, without optimizing them for the specific testbed characteristics. Although optimization could potentially improve accuracy, such fine-tuning was outside the scope of this research.

Future Research Directions

Subsequent research endeavours could explore the incorporation of supplementary predictors, such as the morphological variation index introduced by Alcaras *et al*. (2022), to enhance the predictive capabilities of interpolation models. Additionally, integrating advanced machine learning techniques and developing hybrid models offer promising avenues for achieving heightened predictive performance. The exploration of sophisticated alternatives, such as spectral analysis for uncertainty estimation, may unveil innovative approaches.

Within the scope of this study, we operated under the assumption that the depth data obtained from BlueTopo were free of measurement uncertainty. It is crucial to acknowledge, however, that depth measurements are usually associated with uncertainty

attributed to various factors (refer to Hare et al. 2011). Future research endeavours should consider integrating measurement uncertainty with interpolation uncertainty, a practice that would facilitate a more reliable assignment of CATZOC values for use in nautical charting.

Another compelling area for future research involves comparing the actual interpolation uncertainty with Kriging uncertainty. This is especially pertinent given that interpolation uncertainty statistics, as observed in this study, deviate from assumptions e.g., homogeneity of variance – assumptions that are fundamental to the Kriging method.

Addressing the accurate quantification of uncertainty in setline spacing hydrographic surveys is another crucial aspect of future research. Understanding how this uncertainty can be leveraged to optimize setline spacing hydrographic survey design for meeting desired CATZOC standards could prove invaluable. Such optimization has the potential to enhance efficiency in terms of both time and cost while ensuring essential accuracy standards are maintained. These multidimensional explorations hold the promise of revolutionizing the methodologies employed in hydrographic surveys.

Conclusions

This study in part sought to identify the optimal bathymetric gap-filling interpolation method that produces the lowest interpolation uncertainty from both scientific and operational perspectives. Importantly, it also aimed to quantify and characterize uncertainty in interpolated bathymetry within operational settings. The investigation explored the hidden and non-linear relationship(s) between interpolation uncertainty and a suite of ancillary parameters – distance to the nearest measurement, slope, and roughness – for quantifying interpolation uncertainty across five testbeds in the United States, using various sampling densities. Additionally, the impact of seabed morphology, data paucity, and spatial scales on uncertainty estimation was examined.

Results revealed Spline as the interpolation method that had the lowest uncertainty for Testbeds 2 (Rough), 3 (Slopy) and 4 (Rough and Slopy), followed by Linear and IDW. However, the difference among them was minimal from an operational standpoint. For Testbed 1, Linear interpolation resulted in the least interpolation uncertainty, followed by Spline and IDW. The relationship between the ancillary parameters and interpolation uncertainty was weak individually using separate Linear Regression models for each and only slightly improved when combined into a single ANN model. Notably, roughness emerged as the most important predictor, followed by slope and distance to the nearest measurement across the testbeds. Additionally, IDW provides a more accurate quantification of interpolation uncertainty than Spline and Linear interpolation methods for most testbeds.

The impact of cell sampling density on the uncertainty model's explanatory power is evident, diminishing effectiveness with smaller sampling densities and increasing distances to the nearest measurement point. However, 10% sampling density was identified as the optimal sampling density for bathymetric interpolation uncertainty

estimation (as higher sampling densities do not significantly improve the predictive capabilities of the model).

Moreover, the estimation of interpolation uncertainty varied across testbeds, with Testbed 4 (Rough and Slopy seabed) yielding the best results (R^2 of 0.44 at 50% sampling density), followed by Testbed 3 (Slopy seabed), Testbed 2 (Rough seabed), and Testbed 1 (Flat seabed). These insights highlight the presence of unaccounted-for factors influencing uncertainty, accentuating that higher data resolution (smaller pixel size) enhances uncertainty estimation, while the optimal window size depends on sampling density.

While the study focuses on deterministic interpolation methods, the decision not to optimize interpolation parameters for different testbeds aligns with an operational setting's data-driven workflow, prioritizing moderate processing time and minimal interpolation parameter tweaking. This research advances our understanding of how measurable factors contribute to uncertainty estimates in bathymetric models, offering valuable perspectives for uncertainty estimation, hydrographic survey planning, and future research and applications in this domain.

Funding

The work was supported by the National Oceanic and Atmospheric Administration under [grant number NA20NOS4000196]

Acknowledgements

We express our gratitude to the National Bathymetric Source team at the National Oceanic and Atmospheric Administration’s Office of the Coast Survey for their invaluable support throughout this study. Additionally, we extend our appreciation to Ioannis Kornaros for engaging in constructive discussions related to data analysis.

Disclosure statement

The authors do not report any conflicts of interest. The views and opinions expressed in this work are those of the authors and do not necessarily reflect the views or positions of any entities they represent.

Data availability

The data and code supporting the findings of this study are available in figshare at https://figshare.com/s/56fa7877c787ddb26975

References

Agatonovic-Kustrin, S. and Beresford, R., 2000. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, 22 (5), 717–727.

Aguilar, F.J., Agüera, F., Aguilar, M.A., and Carvajal, F., 2005. Effects of Terrain Morphology, Sampling Density, and Interpolation Methods on Grid DEM Accuracy. *undefined*, 71 (7), 805–816.

Alcaras, E., Amoroso, P.P., and Parente, C., 2022. The Influence of Interpolated Point Location and Density on 3D Bathymetric Models Generated by Kriging Methods: An Application on the Giglio Island Seabed (Italy). *Geosciences*, 12 (2), 62.

Amante, C.J., 2012. Accuracy of Interpolated Bathymetric Digital Elevation Models. University of Colorado Boulder.

Amante, C.J., 2018. Estimating Coastal Digital Elevation Model Uncertainty. *Journal of Coastal Research*, 34 (6), 1382–1397.

Amante, C.J. and Eakins, B.W., 2016. Accuracy of interpolated bathymetry in digital elevation models. *Journal of Coastal Research*, 76 (sp1), 123–133.

Amoroso, P.P., Aguilar, F.J., Parente, C., and Aguilar, M.A., 2023. Statistical Assessment of Some Interpolation Methods for Building Grid Format Digital Bathymetric Models. *Remote Sensing*, 15 (8), 2072.

Anderson, E.S., Thompson, J.A., and Austin, R.E., 2005. LIDAR density and linear interpolator effects on elevation estimates. *International Journal of Remote Sensing*, 26 (18), 3889–3900.

Arun, P.V., 2013. A comparative analysis of different DEM interpolation methods. *The Egyptian Journal of Remote Sensing and Space Science*, 16 (2), 133–139.

Bojanov, B.D., Hakopian, H.A., and Sahakian, A.A., 1993. *Spline Functions and Multivariate Interpolations*. Dordrecht: Springer Netherlands.

Bongiovanni, C., 2018. Quantifying Vertical Uncertainty and the Temporal Variability of the Seafloor to Inform Hydrographic Survey Priorities. Master of Science. University of New Hampshire, Durham.

Bongiovanni, C., Armstrong, A., Calder, B., and Lippman, T., 2018. IDENTIFYING FUTURE HYDROGRAPHIC SURVEY PRIORITIES: A QUANTITATIVE UNCERTAINTY BASED APPROACH. *International Hydrographic Review*.

Burrough, P.A. and McDonnell, R.A., 1998. *Principles of Geographical Information Systems*. Oxford: Oxford University Press.

Calder, B., 2006. On the Uncertainty of Archive Hydrographic Data Sets. *IEEE Journal of Oceanic Engineering*, 31 (2), 249–265.

Caruso, C. and Quarta, F., 1998. Interpolation methods comparison. *Computers & Mathematics with Applications*, 35 (12), 109–126.

Chaplot, V., Darboux, F., Bourennane, H., Leguédois, S., Silvera, N., and Phachomphon, K., 2006. Accuracy of interpolation techniques for the derivation of

digital elevation models in relation to landform types and data density. *Geomorphology*, 77 (1–2), 126–141.

Childs, C., 2004. Interpolating Surfaces in ArcGIS Spatial Analyst. *In*: *ArcUser*. Redlands: ESRI Press.

Chowdhury, E., Hassan, Q., Achari, G., and Gupta, A., 2017. Use of Bathymetric and LiDAR Data in Generating Digital Elevation Model over the Lower Athabasca River Watershed in Alberta, Canada. *Water*, 9 (1), 19.

Curtarelli, M., Leão, J., Ogashawara, I., Lorenzzetti, J., and Stech, J., 2015. Assessment of Spatial Interpolation Methods to Map the Bathymetry of an Amazonian Hydroelectric Reservoir to Aid in Decision Making for Water Management. *ISPRS International Journal of Geo-Information*, 4 (1), 220–235.

Davis, B.M., 1987. Uses and abuses of cross-validation in geostatistics. *Mathematical Geology*, 19 (3), 241–248.

Declercq, F.A.N., 1996. Interpolation Methods for Scattered Sample Data: Accuracy, Spatial Patterns, Processing Time. *Cartography and Geographic Information Systems*, 23 (3), 128–144.

Eakins, B.W. and Taylor, L.A., 2010. Seamlessly integrating bathymetric and topographic data to support tsunami modeling and forecasting efforts. *In*: J. Breman, ed. *Ocean Globe*. Redlands: ESRI Press, 37–56.

Elmore, P.A., Fabre, D.H., Sawyer, R.T., Ladner, R.W., Elmore, P.A., Fabre, D.H., Sawyer, R.T., and Ladner, R.W., 2012. Uncertainty estimation for databased bathymetry using a Bayesian network approach. *Geochemistry, Geophysics, Geosystems*, 13 (9), 9011.

Erdogan, S., 2009. A comparision of interpolation methods for producing digital elevation models at the field scale. *Earth Surface Processes and Landforms*, 34 (3), 366–376.

Erdogan, S., 2010. Modelling the spatial distribution of DEM error with geographically weighted regression: An experimental study. *Computers & Geosciences*, 36 (1), 34–43.

Gunarathna, M.H.J.P., Nirmanee, K.G.S., and Kumari, M.K.N., 2016. Are Geostatistical Interpolation Methods Better than Deterministic Interpolation Methods in Mapping Salinity of Groundwater? *International Journal of Research and Innovations in Earth Science*, 3 (3).

Guo, Q., Li, W., Yu, H., and Alvarez, O., 2010. Effects of Topographic Variability and Lidar Sampling Density on Several DEM Interpolation Methods. *Photogrammetric Engineering & Remote Sensing*, 76 (6), 701–712.

Hare, R., Eakins, B., and Amante, C., 2011. Modelling Bathymetric Uncertainty. *International Hydrographic Review*, 31–42.

Henrico, I., 2021. Optimal interpolation method to predict the bathymetry of Saldanha Bay. *Transactions in GIS*, 25, 1991–2009.

Horn, B.K.P., 1981. Hill shading and the reflectance map. *Proceedings of the IEEE*, 69 (1), 14–47.

Hughes Clarke, J.E., Mayer, L.A., and Wells, D.E., 1996. Shallow-water imaging multibeam sonars: A new tool for investigating seafloor processes in the coastal zone and on the continental shelf. *Marine Geophysical Researches*, 18 (6), 607–629.

IHO, 2014. IHO Transfer Standard for Digital Hydrographic Data Publication S-57.

IHO, 2020. International Hydrographic Organization Standards for Hydrographic Surveys S-44 Edition 6.0.0.

IHO, 2022. International Hydrographic Organization S-101 Annex A, Data Classification and Encoding Guide.

Isaaks, E.H. and Srivastava, R.M., 1989. *An Introduction to Applied Geostatistics*. New York: Oxford University Press.

Jakobsson, M., Calder, B., and Mayer, L., 2002. On the effect of random errors in gridded bathymetric compilations. *Journal of Geophysical Research: Solid Earth*, 107 (B12), ETG 14-1-ETG 14-11.

Jakobsson, M., Mayer, L.A., Bringensparr, C., Castro, C.F., Mohammad, R., Johnson, P., Ketter, T., Accettella, D., Amblas, D., An, L., Arndt, J.E., Canals, M., Casamor, J.L., Chauché, N., Coakley, B., Danielson, S., Demarte, M., Dickson, M.-L., Dorschel, B., Dowdeswell, J.A., Dreutter, S., Fremand, A.C., Gallant, D., Hall, J.K., Hehemann, L., Hodnesdal, H., Hong, J., Ivaldi, R., Kane, E., Klaucke, I., Krawczyk, D.W., Kristoffersen, Y., Kuipers, B.R., Millan, R., Masetti, G., Morlighem, M., Noormets, R., Prescott, M.M., Rebesco, M., Rignot, E., Semiletov, I., Tate, A.J., Travaglini, P., Velicogna, I., Weatherall, P., Weinrebe, W., Willis, J.K., Wood, M., Zarayskaya, Y., Zhang, T., Zimmermann, M., and Zinglersen, K.B., 2020. The International Bathymetric Chart of the Arctic Ocean Version 4.0. *Scientific Data*, 7 (1), 176.

Jakobsson, M., Stranne, C., O'Regan, M., Greenwood, S.L., Gustafsson, B., Humborg, C., and Weidner, E., 2019. Bathymetric properties of the Baltic Sea. *Ocean Science*, 15 (4), 905–924.

Kastrisios, C. and Ware, C., 2022. Textures for coding bathymetric data quality sectors on electronic navigational chart displays: design and evaluation. *Cartography and Geographic Information Science*.

Kruskal, W.H. and Wallis, W.A., 1952. Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47 (260), 583–621.

Lam, N.S.-N., 1983. Spatial Interpolation Methods: A Review. *The American Cartographer*, 10 (2), 129–150.

Legleiter, C.J. and Kyriakidis, P.C., 2006. Forward and Inverse Transformations between Cartesian and Channel-fitted Coordinate Systems for Meandering Rivers. *Mathematical Geology*, 38 (8), 927–958.

Li, J. and Heap, A.D., 2008. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Canberra: Geoscience Australia.

Li, J. and Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, 53, 173–189.

Liu, X., Zhang, Z., Peterson, J., and Chandra, S., 2007. LiDAR-derived high quality ground control information and DEM for image orthorectification. *GeoInformatica*, 11 (1), 37–53.

Lloyd, C.D. and Atkinson, P.M., 2002. Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing*, 23 (12), 2519–2524.

MacEachren, A.M. and Davidson, J. V., 1987. Sampling and Isometric Mapping of Continuous Geographic Surfaces. *The American Cartographer*, 14 (4), 299–320.

Mayer, L., Jakobsson, M., Allen, G., Dorschel, B., Falconer, R., Ferrini, V., Lamarche, G., Snaith, H., and Weatherall, P., 2018. The Nippon Foundation—GEBCO Seabed 2030 Project: The Quest to See the World's Oceans Completely Mapped by 2030. *Geosciences*, 8 (2), 63.

Merwade, V., 2009. Effect of spatial trends on interpolation of river bathymetry. *Journal of Hydrology*, 371 (1–4), 169–181.

Merwade, V.M., Maidment, D.R., and Goff, J.A., 2006. Anisotropic considerations while interpolating river channel bathymetry. *Journal of Hydrology*, 331 (3–4), 731–741.

Negreiros, J., Painho, M., Aguilar, F., and Aguilar, M., 2010. Geographical Information Systems Principles of Ordinary Kriging Interpolator. *Journal of Applied Sciences*, 10 (11), 852–867.

Negreiros, J., Painho, M., Aguilar, M.A., and Aguilar, F.J., 2008. Spatial Error and Interpolation Uncertainty Appraisal Within Geographic Information Systems. *Research Journal of Applied Sciences*, 3, 471–479.

Nippon Foundation-GEBCO, 2023. Our Mission — Seabed 2030 [online]. Available from: https://seabed2030.org/our-mission/ [Accessed 8 Dec 2023].

NOAA, 2020. The National Bathymetric Source Presentation [online]. Available from: https://www.weather.gov/media/watercommunity/Webinar/20Jan21/Jan21_CC_COP_Wyllie.pdf [Accessed 15 Oct 2023].

Özdamar, L., Demirhan, M., and Özpınar, A., 1999. A comparison of spatial interpolation methods and a fuzzy areal evaluation scheme in environmental site characterization. *Computers, Environment and Urban Systems*, 23 (5), 399–422.

Panhalakr, S.S. and Jarag, A.P., 2016. Assessment of Spatial Interpolation Techniques for River Bathymetry Generation of Panchganga River Basin Using Geoinformatic Techniques. *Asian Journal of Geoinformatics*, 15 (3), 9–5.

Rice, G., Wyllie, K., Gallagher, B., and Geleg, P., 2023. The National Bathymetric Source. *In*: *OCEANS 2023 - MTS/IEEE U.S. Gulf Coast*. IEEE, 1–7.

Ryan, W.B.F., Carbotte, S.M., Coplan, J.O., O'Hara, S., Melkonian, A., Arko, R., Weissel, R.A., Ferrini, V., Goodwillie, A., Nitsche, F., Bonczkowski, J., and Zemsky, R., 2009. Global Multi-Resolution Topography synthesis. *Geochemistry, Geophysics, Geosystems*, 10 (3).

Schaap, D.M.A. and Schmitt, T., 2020. EMODnet Bathymetry – further developing a high resolution digital bathymetry for European seas. *EGU2020*.

Šiljeg, A., Lozić, S., and Šiljeg, S., 2014. A Comparison of Interpolation Methods on the Basis of Data Obtained from A Bathymetric Survey of Lake Vrana, Croatia. *Hydrology & Earth System Sciences Discussions* .

Smith, W.H.F. and Sandwell, D.T., 1997. Global sea floor topography from satellite altimetry and ship depth soundings. *Science*, 277 (5334), 1956–1962.

Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46, 234–240.

Tomczak, M., 1998. Spatial Interpolation and its Uncertainty Using Automated Anisotropic Inverse Distance Weighting (IDW) - Cross-Validation/Jackknife Approach.

Vetter, M., Höfle, B., Mandlburger, G., and Rutzinger, M., 2011. Estimating changes of riverine landscapes and riverbeds by using airborne LiDAR data and river cross-sections. *Zeitschrift für Geomorphologie, Supplementary Issues*, 55 (2), 51–65.

Voltz, M. and Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *Journal of Soil Science*, 41 (3), 473–490.

Ware, C. and Kastrisios, C., 2022. Evaluating Countable Texture Elements to Represent Bathymetric Uncertainty. *In*: *EuroVis*. Rome, Italy: The Eurographics Association.

Weatherall, P., Marks, K.M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J.E., Rovere, M., Chayes, D., Ferrini, V., and Wigley, R., 2015. A new digital bathymetric model of the world's oceans. *Earth and Space Science*, 2 (8), 331–345.

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., and Grehan, A.J., 2007. Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. *Marine Geodesy*, 30 (1–2), 3–35.

Wu, C.Y., Mossa, J., Mao, L., and Almulla, M., 2019. Comparison of different spatial interpolation methods for historical hydrographic data of the lowermost Mississippi River. *Annals of GIS*, 25 (2), 133–151.

Zar, J.H., 1999. *Biostatistical analysis*. Edgewood Cliffs, New Jersey: Prentice Hall.

Figures and Tables

Figure 1: Five testbeds of various seabed characteristics for testing the proposed methodology in: a) Gulf of Mexico (Flat seabed), b) Florida Main Channel (Rough seabed), c) Massachusetts Bay (Slopy seabed), d) Gulf of Maine (Rough and Slopy seabed) and e) Gulf of Maine (Rough and Slopy seabed with multi-resolution).

Figure 2: Histograms depicting the slope and roughness characteristics of the original testbeds' bathymetric datasets.

Figure 3: An n x n analysis window for a raster grid.

Figure 4: Workflow for quantifying interpolation uncertainty, including the machine learning techniques' architectures.

Figure 5: Interpolation methods uncertainty comparison on Testbed 1(Flat) using box and whisker, plotted with 99% percentile of data.

Figure 6: Interpolation methods uncertainty comparison on Testbed 4 (Rough and Slopy) using box and whisker, plotted with 99% percentile of data.
Figure 7: Histograms showing pair-wise differences of interpolation methods' uncertainties for Testbed 4 (Rough and Slopy) at various sampling densities.
Figure 8: Interpolation uncertainty on Testbed 4 (Rough and Slopy) across all sampling densities (columns) and interpolation methods (rows), using 99th percentile of data.
Figure 9: Interpolation uncertainty on Testbed 1 (Flat) across all sampling densities (columns) and interpolation methods, using 99th percentile of data.
Figure 10: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to the nearest measurement, roughness, and slope respectively for Testbed 4 (Rough and Slopy).
Figure 11: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to the nearest measurement, roughness, and slope combined for Testbed 4 (Rough and Slopy).
Figure 12: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on window sizes using Testbed 4 (Rough and Slopy).
Figure 13: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on data resolution using Testbed 5.

Table 1: The International Hydrographic Organization (IHO) S-57 Category of Zones of Confidence (CATZOC) Levels. (IHO 2014).
Table 2: Summary of metadata for the five testbeds. NAD83 - North American Datum of 1983 and UTM - Universal Transverse Mercator

Table 3: Testbed 4 (Rough and Slopy) Statistics for pair-wise interpolation methods comparison at various sampling densities.

Table 4: Statistics of the Importance of Predictors of Uncertainty across testbeds