FILLING THE BATHYMETRIC GAPS: TOWARD QUANTIFYING AND CHARACTERIZING UNCERTAINTY IN INTERPOLATED BATHYMETRY

By

ELIAS OPEYEMI ADEDIRAN

ND (Distinction) in Surveying and Geoinformatics, Yaba College of Technology, Nigeria, 2013

BS (First Class) in Surveying and Geoinformatics, University of Lagos, Nigeria, 2018

THESIS

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

in

Ocean Engineering: Ocean Mapping

May, 2024

This thesis has been examined and approved in partial fulfillment of the requirements for the degree of Master of Science in Ocean Engineering: Ocean Mapping by:

Kim Lowell (Thesis co-advisor and chair), Research Scientist,
University of New Hampshire.

Christos Kastrisios (Thesis co-advisor), Research Associate Professor,
University of New Hampshire.

Glen Rice, Physical Scientist,
Hydrographic Systems and Technology Branch, National Oceanic and Atmospheric Administration.

On February 8th, 2024

Original approval signatures are on file with the University of New Hampshire Graduate School.

DEDICATION

All gratitude is due to Allah, the Lord of all worlds (Q1:2) and everything within it (Q20:6). He fashioned me from dust (Q23:12-14), brought me into the world from my mother's womb when I was unaware (Q16:78), and initiated it all with the first word “Read” (Q96:1). Through the pen, He imparted knowledge to me (Q96:4) about the ocean (Q24:40; Q25:53; Q52:6; Q55:19-20) that I was unaware of (Q96:5).

To my beloveds, Faatimah and Ahmad, my parents, Tpl. Ganiyu Adediran and Late Basirat Adediran (may Allah be pleased with her), and all who supported me on this journey, this dedication is for you. I am eternally grateful for your unwavering support.

ACKNOWLEDGEMENTS

The successful completion of this challenging project owes its accomplishment to the strength, knowledge, and support systems bestowed upon me by my Creator. All praise is due to Him, "Alhamdulillah Robbil ’Aalameen."

I extend my gratitude to NOAA for funding this work through grant NA20NOS4000196 and to the Center for Coastal and Ocean Mapping (CCOM) for finding me worthy of the graduate research assistantship. The numerous opportunities provided by the Center, not limited to research cruises, conferences, outreach events, and knowledge and exposure gained from renowned scholars, have been invaluable. My time at CCOM/UNH has undoubtedly been transformative and enriching, and I take pride in being part of the CCOM/UNH family.

I express sincere thanks to my committee members Kim Lowell, Christos Kastrisios, and Glen Rice for their guidance, teaching, support, inspiration, encouragement, and insightful comments throughout the research. Being under the supervision of such brilliant minds is a privilege, and I eagerly anticipate engaging in more exciting and groundbreaking endeavors with all of you.

Acknowledgment goes to Katrina Wyllie of the National Bathymetric Source program for her support. Special thanks to Qi Zhang for his statistical insights during the final phase of this work. Appreciation extends to all UNH and CCOM colleagues, friends, GEBCO Scholars of Years 18, 19, and 20, and Nigerian families in the US, without whom this journey would not have been possible. The list is long, so I refrain from listing names, trusting that you all recognize your invaluable contributions.

Heartfelt thanks to every family member for their unwavering commitment and support. A special acknowledgment to my father for his unflinching care and unflagging support physically, spiritually, and morally towards my academic success. Thank you for always believing in me.

Lastly, my deepest gratitude goes to my beloved wife, whose untainted love, care, and support have been the cornerstone of my achievements. Thank you for your patience, cooperation, and sacrifices. A virtuous wife is a treasure beyond measure, and I feel blessed to have you by my side. “Together to the top and Jannatul Firdaous bi idhnillah”

Chapter 2, in part, has been submitted for publication of the material as it may appear in the Journal of Marine Geodesy, 2024, Adediran *et al.*, 2024. The dissertation author was the primary researcher and author of this paper.

Chapter 3, in part, is currently being prepared for submission for publication of the material. Adediran *et al*. The dissertation author was the primary researcher and author of this material.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The oceans remain one of Earth's most enigmatic frontiers, with approximately 75% of the world's oceans still unmapped to modern standards. To overcome this, interpolation serves as the primary method for creating seamless coverage from incomplete coverage hydrographic data sets and is essential for creating a seamless digital bathymetric model (DBM) compiled from sparse bathymetric datasets and set-line spacing surveys. While methods for quantifying the uncertainty in depth measurements are well-researched, interpolation introduces unqualified depth uncertainties. This study aims to estimate and characterize these uncertainties which are essential to nautical charting, and navigational safety, and important in many other fields.

Organized into two source data scenarios, the research focuses on uncertainties arising from randomly sampled data and set-line surveys. It employs three widely recognized deterministic interpolation methods—Linear, Inverse Distance Weighting (IDW), and Spline— across five testbeds that vary in slope and roughness. The goal is to identify the interpolation method with the lowest uncertainty and unravel the relationships between interpolation uncertainty and three ancillary parameters (distance to the nearest measurement, slope, and roughness) for estimating interpolation uncertainty.

By sampling complete seafloor coverage sonar depth data at different densities and line spacings, the study interpolates across entire testbed areas using the chosen methods. Uncertainty is calculated by comparing interpolated depths against the true depths for independent points. The resulting uncertainties are analyzed statistically and spatially to assess consistency across interpolation methods and determine the interpolation method that yields the least interpolation uncertainty. Linear regression and machine learning techniques (neural networks and random

forest) are used to model the relationship between these uncertainties and ancillary parameters to estimate uncertainty.

Evaluation across the five testbeds, encompassing both random sampling and set-line spacing scenarios, reveals the following: 1) Spline performs better than Linear and IDW in estimating depths from a purely scientific perspective; however, 2) differences among the interpolation methods are not statistically significant and minimal from an operational standpoint; 3) sampling density, line spacing, and spatial scales impact uncertainty; 4) roughness is the most important parameter and distance the least important; 5) relationships between ancillary parameters and uncertainty are weak though statistically significant.

The findings of this work suggest the presence of unaccounted-for factors shaping uncertainty or indicate a strong random component within interpolation uncertainty yet lay a foundational understanding for improving the estimate of uncertainty in DBMs within operational settings. Future research recommendations involve exploring supplementary predictors to enhance the predictive capacity of ancillary parameters. Additionally, innovative approaches such as spectral analysis for uncertainty estimation hold promise in advancing methodologies within this domain.

CHAPTER 1 : INTRODUCTION

The Earth's oceans, constituting approximately 71% of its surface (Weatherall *et al.* 2015), remain one of our planet's last great unknowns. Essential for sustaining life, climate regulation, and economic activities, oceans remain a vast reservoir of resources and wealth (Mayer *et al.* 2018). Surprisingly though, the surfaces of Mars, Venus, and the Earth's Moon are mapped at a higher spatial resolution than the seafloor (Smith 2004), where individual depth measurements can exhibit data gaps spanning hundreds of kilometers (Smith and Sandwell 1997). Digital bathymetric models (DBMs) are continuous representations of the seafloor that are derived from depth measurements and are commonly stored in a raster data format comprising a matrix of same-sized square cells, with each cell representing the average depth of the area contained within that cell (Amante and Eakins 2016, Jakobsson *et al.* 2019). Other depth representations (shallow, deep, or nodal depths) are also employed based on application needs.

With 75% of the world's oceans unmapped to modern standards (Seabed 2030 2023), 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 largely known at high resolution. Although incorporating data derived from the traditional sounding techniques of the lead line, single-beam echo sounders (SBES), and modern high-resolution multibeam echo sounders (MBES), seabed models largely rely on interpolation and altimetry-derived data (Weatherall *et al.* 2015).

The National Oceanic and Atmospheric Administration's (NOAA) National Bathymetric Source (NBS) project is currently creating DBMs for the United States from the best available data, aimed at the creation of next-generation nautical charts while also providing support for modeling, industry, science, regulation, and public curiosity (Rice *et al.* 2023). Consequently, interpolation becomes an important step, often required to fill data gaps among sparse bathymetric datasets, creating a comprehensive nationwide model of the seafloor. Data gaps can be the result of survey design or survey oversight and can span meters to kilometers between survey measurements. Given this reliance on interpolation, the precision of 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 ensuring 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, and even dangerous. This underlines the importance of bathymetry quality, forming a vital part of hydrography's Greenaway stool alongside depth and its source (Figure 1) (NOAA 2020). A precise estimate of uncertainty enhances navigation safety by mitigating the challenge of assigning a representative and useful category zone of confidence (CATZOC) (IHO S-57 2014) to areas without full seafloor bathymetric coverage. It also aids in optimizing shipping routes and enhances survey planning with associated cost savings.

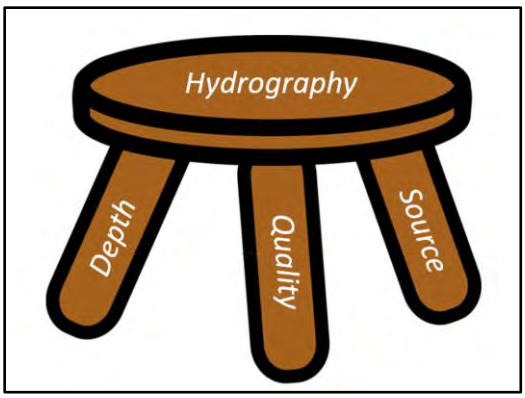

Figure 1: Greenaway stool of Hydrography (NOAA 2020)

Despite the widespread use of DBMs, the nature of the inherent uncertainty associated with interpolation in these models remains largely unexplored, holding broad implications across academic, commercial, and potentially life-saving domains. Estimating DBM uncertainty contributes 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. Hazard modeling, encompassing tsunami inundation and hurricane storm-surge inundation, relies on DBMs 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 future human losses.

Despite the critical importance of comprehending the nature of uncertainty in interpolated bathymetric models, the ocean mapping field lacks comprehensive research on estimating and characterizing interpolation uncertainty. In light of this gap, this study aims to identify the deterministic interpolation method that produces the lowest interpolation uncertainty and, importantly, accurately estimate and characterize interpolation uncertainties in DBMs. Moreover, this will be done within the context of operational settings. The investigation involves assessing the correlation between interpolation uncertainties and various ancillary parameters — distance to

the nearest known/actual measurement, seabed slope, and roughness — across multiple geomorphological testbeds. Building upon the groundwork of (Amante & Eakins 2016), this research delves into evaluating the strength of the relationship of these three parameters to interpolated bathymetry uncertainty. Furthermore, it explores the practicality of leveraging such relationships for predicting uncertainty during survey execution. The study adopts a comprehensive approach, examining the impact of seafloor morphology, data scarcity, spatial resolution, and spatial scales on uncertainty. This holistic exploration aims to provide a nuanced understanding of the multifaceted dynamics involved in characterizing uncertainty in interpolated bathymetry.

The aim of this work is in part to establish the deterministic interpolation method that produces the lowest interpolation uncertainty but, more importantly, to explore methods for estimating uncertainty from these methods within bathymetric datasets in operational settings. Therefore, the focus is on characterizing uncertainty from diverse interpolation methods to offer optimal insights in operational settings. The study delves into three common deterministic interpolation techniques that can be employed operationally, *i.e.*, Linear, Spline, and inverse distance weighting (IDW) with a comprehensive explanation provided in the methodology section.

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, Precision Navigation, and Office of Coast Survey Hydrographic Health Model that drives Survey Planning and Prioritization which

are the cornerstones of safer navigation, resilient coastal communities and ecosystems, and a stronger blue economy.

For terminology consistency and clarity in this study, we designate depth measurements from the NOAA BlueTopo's (NOAA OCS BlueTopo 2023) non-interpolated surface as "true depth", despite any uncertainty associated with these measurements. Three terms, interpolation uncertainty, residual, or interpolation deviation, can describe the difference between interpolated depths and measured depths. While the difference among them is acknowledged, the term "interpolation uncertainty" is adopted herein. Furthermore, in the context of this work, "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 remainder of this work is organized in the following manner:

Chapter 2 discusses the problems, questions, and objectives of this research work.

Chapter 3 provides general background information for this research work including bathymetric data acquisition methods, hydrographic survey design techniques, related works, nautical charting uncertainty standards, and the study area.

Chapter 4 explores the estimation and characterization of uncertainty in sparse hydrographic datasets,

Chapter 5 focuses on similar aspects in the context of set-line spacing surveys, and lastly,

Chapter 6 encapsulates conclusions and provides recommendations for future studies and implementations.

CHAPTER 2 : RESEARCH QUESTIONS AND OBJECTIVES

In light of the extensive portion of the ocean yet to be mapped to modern standards, as discussed in the Introduction, interpolation is usually required to create a DBM of the seafloor from the available datasets derived from traditional sounding techniques of the lead line, single-beam echo sounders (SBES), and modern high-resolution MBES set-line spacing surveys, and full seabed coverage surveys. With the widespread application of DBMs derived from interpolation methods across scientific and industrial domains and recognizing the significance of comprehending the intrinsic uncertainty in these DBMs, this study investigates the estimation and characterization of interpolation uncertainty in interpolated bathymetry and is structured into two primary components: 1) sparse bathymetric datasets addressed in chapter 4 and 2) set-line spacing hydrographic surveys addressed in chapter 5.

2.1 Research Questions

Based on this background, the key questions addressed in this study are outlined as follows:

(1) Which deterministic interpolation method is most accurate for bathymetry, yielding the lowest interpolation uncertainty?

(2) How efficiently can these ancillary parameters, *i.e.,* distance to the nearest measurement, seabed slope, and roughness, be used to estimate and characterize interpolation uncertainty in bathymetric models in an operational setting?

(3) To what extent can we improve the estimated uncertainty by employing Machine Learning techniques to combine these parameters?

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

2.2 Research Objectives

To answer the research questions, the following objectives are defined in this study:

(1) Determine the deterministic interpolation method that produces the lowest interpolation uncertainty in bathymetry.

(2) Investigate the relationship between interpolation uncertainty and ancillary parameters, such as distance to the nearest measurement, seabed slope, and roughness, for estimating and characterize uncertainty in bathymetric models.

(3) Explore the potential improvements in uncertainty estimation by integrating Machine Learning techniques to combine these parameters.

(4) Characterize interpolation uncertainties to examine the impact of seafloor morphology, data scarcity, spatial resolution, and spatial scales on uncertainty estimation.

CHAPTER 3 : BACKGROUND INFORMATION

3.1 Data Acquisition/Bathymetric Methods

Rapid technological advancements are transforming the landscape of ocean mapping, empowering users with unprecedented capabilities through modern hydrographic systems. Traditionally, lead-lines were employed for depth collection, where lead weights attached to a calibrated line were thrown from the side of a boat until it reached the bottom (Bongiovanni 2018). The depth would then be read from the line and the depth and location would be recorded on the survey sheet. This method, however, was time-consuming, tedious, and susceptible to locational errors caused by wind and currents (van der Wal and Pye 2003). To account for these errors, depths were typically rounded down (making them shoaler) to the nearest fathom (van der Wal and Pye 2003, Calder 2006).

Beyond measurement inaccuracies, lead-line techniques suffered from limited measurement frequency and density, capturing only individual points along the vessel track and leaving the intervening seafloor unmeasured and unknown. Despite these limitations, lead-lines persisted as the primary depth measurement method until the 1930s when sonic echosounders were introduced, revolutionizing depth collection with faster and continuous methods (Hawley 1931, Adams 1942).

The mid-1900s witnessed advancements in geospatial positioning, transitioning from sextant measurements to electronic positioning in the 1950s and satellite positioning in the 1990s. These improvements helped shape the performances of both the deep-water multibeam systems in the 1980s and the shallow-water multibeam systems in the 1990s (Wong *et al.* 2007). The profiling echosounder significantly increased hydrographic surveying capabilities by collecting a constant data stream of depths recorded directly under a boat as it moves along a track or course. The

invention of side-scan sonar (SSS) and multibeam echo sounders (MBES) led to different types of survey designs.

3.2 Hydrographic Survey Design Techniques

According to the NOAA Hydrographic Survey Specifications and Deliverables (HSSD) 2022, there are four classifications of coverage: Object Detection Coverage, Complete Coverage, Set Line Spacing, and Track line (transit and reconnaissance) (NOAA OCS 2022):

(1) The object detection coverage is assigned for critical under keel clearance areas and may be accomplished with either a 100% bathymetric bottom coverage with multibeam sonars with object detection multibeam developments (*i.e.*, 50 cm grid resolution in 0-20 m depth range) of contacts and features or a 200% side scan sonar coverage with concurrent multibeam bathymetry collection.

(2) Complete Coverage may be accomplished with either a 100% bathymetric bottom coverage with multibeam sonars with complete coverage multibeam developments (*i.e.*, 1 m grid resolution in 0-20 m depth range) of contacts and features or 100% side scan sonar coverage with concurrent multibeam bathymetry collection with complete coverage multibeam developments (*i.e.*, 1 m grid resolution in 0-20 m depth range) of contacts and features.

(3) Set Line Spacing is assigned when acquiring bathymetric data in areas too shallow for efficient full bottom coverage bathymetry or too hazardous for use of equipment. Set line spacing may be accomplished with a single beam or multibeam and it is highlighted that this technique is used when complete coverage cannot be achieved due to associated costs of survey.

(4) Track line survey operations can be classified as either Transit, which is intended to be used simply as an opportunity to collect data while a vessel transits from location A to location B; or Reconnaissance, which is intended to be used when the intended survey products will require a higher level of accuracy than Transit specifications will produce, but a traditional survey consisting of systematic line spacing or full bottom coverage is not required.

Chapter 5 of this thesis focuses on quantifying the uncertainty within the interpolated areas resulting from gaps in bathymetric data caused by the design of set-line spacing. An expanded methodology, known as the skunk stripping technique, involves the simultaneous deployment of MBES or SBES and SSS. While this technique also leads to gaps in bathymetric coverage, it effectively ensures a comprehensive search across the entire area and facilitates the detection of features in alignment with the CATZOC requirements using SSS. It should be noted that a SSS collects data from larger swaths than MBES but is not capable of measuring depths. The exceptions to this rule are phase-measuring bathymetric sonars (PMBSs) that concurrently collect bathymetric and side-scan imaging from one system. However, PMBS systems are not frequently used in hydrography since their extremely large raw datasets require extensive manual filtering and have greater potential for errors. Traditional SSSs are more prevalent in the hydrographic community and are primarily used to identify possible dangers to navigation that require additional investigation.

One of the advantages of the set-line spacing and/or skunk stripping technique(s) lies in their ability to survey expansive areas efficiently by reducing time and costs without compromising navigational safety. They are particularly effective in shallow water and hazardous environments. Importantly, the set-line spacing approach is commonly applied in areas with relatively flat

seafloors, low navigational risk, and existing satellite-derived/lidar bathymetry data where bottom detection is challenging due to water clarity or extinction depth (Neff and Wilson 2018, NOAA OCS 2022). This underscores the necessity for interpolation and, more crucially, the accurate quantification of uncertainty in these interpolated regions to provide users with confidence in the data.

Additionally, these hydrographic survey designs are based on line planning that meets the desired accuracy criterion of the survey. According to the NOAA Field Procedure Manual (NOAA OCS 2021), line plans for data acquisition fall into various categories, including Mainscheme, Holidays, Crosslines, Developments, Bathymetric Splits, Target Files, and Special Circumstances. In the context of this study, the applicable and defined categories are Mainscheme, referring to the primary survey data acquired during a survey project, and Crosslines, used to identify systematic data problems by comparing them with Mainscheme data.

3.3 Previous Works

3.3.1 Interpolation methods

Interpolation is a mathematical process of predicting the values of unknown locations based on surrounding measured values (Burrough and McDonnell 1998). Interpolation requires some basic assumptions about the surface: that the surface is continuous and smooth, the sampled points are representative of the overall characteristics of the surface and that measured values at neighboring data points are highly correlated with the value at the unknown point (Liu *et al.* 2007). There are numerous interpolation methods, e.g., IDW, Spline, Linear, Natural Neighbor, and Kriging, and all are based on the same assumption that bathymetry sampled at disparate points is positively spatially autocorrelated. The notion of spatial autocorrelation is largely attributed to Tobler's 1st law of geography, *i.e.,* "Everything is related to everything else, but near things are more related

than distant things" (Tobler 1970). That is, the depth at one location is more similar to depths nearby than the depths far away.

Interpolation methods can be classified as any combination of geostatistical or deterministic, local or global, and exact or inexact (Li and Heap 2008) based on the assumptions and features used to estimate the depths of unknown areas using known measurements. Depending on the method selected, the different mathematical algorithms used in each interpolator produce divergent DEMs, even when developed from the same source data (Aguilar *et al.* 2005, Erdogan 2009). Geostatistical methods, such as Kriging, use both mathematical and statistical functions to estimate depths, while deterministic methods, such as Spline, IDW, and triangulated irregular network (TIN), use the measurements directly and mathematical functions only to predict unknown values (Childs 2004). Geostatistical methods are typically more computationally and time-intensive in order to accurately quantify the statistical relationship among depths compared to deterministic methods, which have simpler parameters and are computationally faster (Grayson and Blöschl 2000, Castiglioni *et al.* 2009).

In the context of uncertainty estimation, the geostatistical and deterministic interpolation techniques have an important distinction. The geostatistical Kriging method utilizes the semivariogram to estimate unknown elevations and to also predict their uncertainty (*i.e.*, variance). A semivariogram captures the spatial autocorrelation of the terrain by plotting the elevation variance of each pair of measurements as a function of the distance among the measurements, and then a mathematical model (e.g., Linear, spherical, exponential) is fit to the semivariogram. Conversely, the deterministic interpolation techniques predict depth values using the data directly rather than mathematical functions, but, notably, they provide no estimates of their vertical uncertainty.

Local interpolation methods use a subset of measurements surrounding the location to be predicted and global methods use all available measurements (Burrough and McDonnell 1998). Local methods are preferred when the seafloor structure is driven by local variation, and global methods should be used when the seafloor structure is driven by a trend over a larger area. Lastly, exact interpolators respect the known measurements by creating a surface that retains the values at the measured locations of measurements. This contrasts with inexact interpolators, which are not constrained by the depth measurement values at those locations (Burrough and McDonnell 1998). Inexact interpolators can add additional uncertainty to the DBM by creating surfaces that diverge from the measurements (Hare *et al.* 2011), but can be useful when there is already large measurement uncertainty. For nautical charting purposes, exact interpolators seem to be a more fitting choice as they ensure that depths are not overestimated at known locations. This is of utmost importance for navigational safety.

The interpolation method employed in a particular situation is chosen based on the data quality, sampling distribution, terrain characteristics, computational resources, and application requirements. Each interpolation method has particular 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, these studies exhibited no consensus regarding which interpolation method performs the best in generating bathymetric surfaces (Wu *et al.* 2019). This is no surprise since the performance of an interpolation method depends on factors like datasets and their properties, seafloor characteristics, study areas, etc. This study is an attempt to determine if there is an optimal, global

interpolation method for generating DBMs. The deterministic interpolators align well with the goal of this research which is to generate data-driven products from large national datasets in operational settings. Hence, the three deterministic spatial interpolation methods that are investigated in this work for improved uncertainty quantification are Spline, Linear, and Inverse Distance Weighting (IDW).

IDW interpolation is a spatial interpolation technique used to estimate values at unsampled locations based on the known values at sampled locations (Liu *et al.* 2007). In IDW, each sampled point's influence on the interpolation at a target location is determined by its distance from the target location raised to a selected power — closer points exert more influence than distant ones (Caruso and Quarta 1998). The method assigns weights inversely proportionally to distances from the target location, with the power parameter controlling the rate of influence decrease. A higher power yields a quicker decline, resulting in a less smooth surface with more detail, while a lower power favors points farther away, creating a smoother surface with less detail. Control over the interpolated surface's characteristics is achieved through fixed or variable search radii, limiting input points for each interpolated cell calculation (Guo *et al.* 2010).

Spline interpolation, a mathematical method, estimates values between known data points by fitting a piecewise-defined polynomial function, typically a cubic Spline, to the dataset (Bojanov *et al.* 1993). This technique employs a series of polynomial functions, or "Splines," within distinct intervals, creating a smoother and more flexible curve that seamlessly passes through all given data points. The resulting curve is continuous, with both its first and second derivatives matching at the points where the polynomial segments connect. This characteristic ensures a smooth transition between intervals, effectively capturing more variations in the data.

Linear interpolation is a simple method used to estimate values between two known data points determined to be immediate neighbors by assuming a linear relationship between them. This technique operates under the assumption that the change in the variable being interpolated is constant across the interval between two adjacent data points regardless of the distance between them. The process involves determining the equation of the straight line connecting the known points and then calculating the interpolated value based on the proportion of the distance along this line of the point for which a value is desired. Essentially, Linear interpolation assumes a linear trend between data points and fills in values within the given range by drawing a straight line between neighboring points.

3.3.2 Interpolation Uncertainty Estimation

This research builds on previous studies 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). Studies have found that the accuracy of all interpolation techniques is related to the sampling density and distribution of measurements (Aguilar *et al*., 2005; Amante, 2018; Amante & Eakins, 2016; Anderson *et al*., 2005; Chaplot *et al.*, 2006; Erdogan, 2009; Erdoğan, 2010; Guo *et al*., 2010; MacEachren & Davidson, 1987). In these previous studies, sampling density referred to 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) or 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 final definition, cell sampling density, is the terminology and approach this study will adopt. Previous studies have also found that the accuracy of all interpolation techniques is related to

terrain characteristics (Aguilar *et al.* 2005, Erdogan 2009, 2010, Guo *et al.* 2010, Amoroso *et al.* 2023). This will also be a focus of this study.

Calder (2006) and Bongiovanni et al. (2018) used the geostatistical Kriging interpolation method which gives an estimate of uncertainty; however, Kriging requires that underlying assumptions such as stationarity, and isotropy be satisfied. Furthermore, due to the complex, highly computational, and time-consuming nature of Kriging, the method is not ideal for the intended purpose of generating data-driven products from large national datasets in operational settings.

Amante (2018) and Amante & Eakins (2016) investigated the accuracy of interpolated DBMs using different deterministic methods of interpolation (Spline, IDW, and TIN) by examining the relationship between interpolation deviations from measured depths, sample density, and distance to the nearest depth measurement. Their predictive models of the cell-level uncertainty were derived as a function of cell sampling density and interpolation distance on a testbed of varying terrain. Their equations used only a single parameter (distance to the nearest depth measurement) and did not incorporate terrain characteristics such as slope and curvature. This constrained the applicability of their results to areas with similar terrain as the study area, to avoid under- or over-estimating the interpolation uncertainty. The research undertaken herein is unique because it is tailored towards an operational setting and will advance these previous studies by incorporating terrain characteristics, specifically slope, and roughness, into the cell-level uncertainty equation. In addition to using machine learning techniques to better explore the relationship between interpolation uncertainty and the ancillary parameters, this investigation is based on five testbeds. This is an effort to improve the characterization of uncertainty in interpolated bathymetric datasets.

3.4 Nautical Charting Uncertainty Standards

The International Hydrographic Organization (IHO) established the S-44 IHO Standards for Hydrographic Surveys (IHO S-44 2020) in 1968 to define data quality requirements for charting purposes, with subsequent updates reflecting technological advancements (IHO S-44 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 S-57 2014) and the Quality of Bathymetric Data (QoBD) in the new S-101 ENC Product Specification (IHO S-101 2022) outline these parameters.

Every chart is a mosaic of polygons / sectors, 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 S-57 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 aims to facilitate the CATZOC classification in areas with incomplete bathymetric seafloor coverage. This includes the *a posteriori* CATZOC classification of interpolated datasets from sparse and set-line spacing bathymetry as well as during set-line spacing surveys targeted to meet CATZOC level within the constraints of existing resources. Specifically, this study addresses the uncertainty metric associated with depth accuracy in CATZOC, excluding considerations for seafloor feature detection and factors like expected feature size or seabed undulations. Thus, 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 SSS. In regions like the United States, where skunk stripping is employed to attain seafloor coverage and target detection,

the attainment of CATZOC A1 or A2 is hindered by the depth accuracy aspect of the CATZOC. This study also aims to provide insights to inform CATZOC classification under such circumstances.

3.5 Study Area/Dataset

The study investigated five distinct testbeds located within U.S. waters, each representing a unique seafloor morphology (Figure 2). The choice of testbeds was made through a dual assessment approach — qualitative evaluation via visual inspection and quantitative analysis detailed in the subsequent section. These testbed models encompass varying combinations of slope and roughness characteristics, facilitating a thorough analysis of factors that may impact interpolation uncertainty. The categorized testbeds are outlined as follows:

- Testbed 1: Characterized by low slope and low roughness,
- Testbed 2: Characterized by low slope and high roughness,
- Testbed 3: Characterized by high slope and low roughness,
- Testbed 4: Characterized by high slope and high roughness, and
- Testbed 5: Shares high slope and high roughness attributes with Testbed 4 but is unique due to its high spatial data resolution. Testbed 5 is considered "special" in the context of this study and is used to explore the impact of spatial data resolution on uncertainty estimation.

For a detailed overview of the testbeds, including their respective locations and key attributes, refer to Table 2.

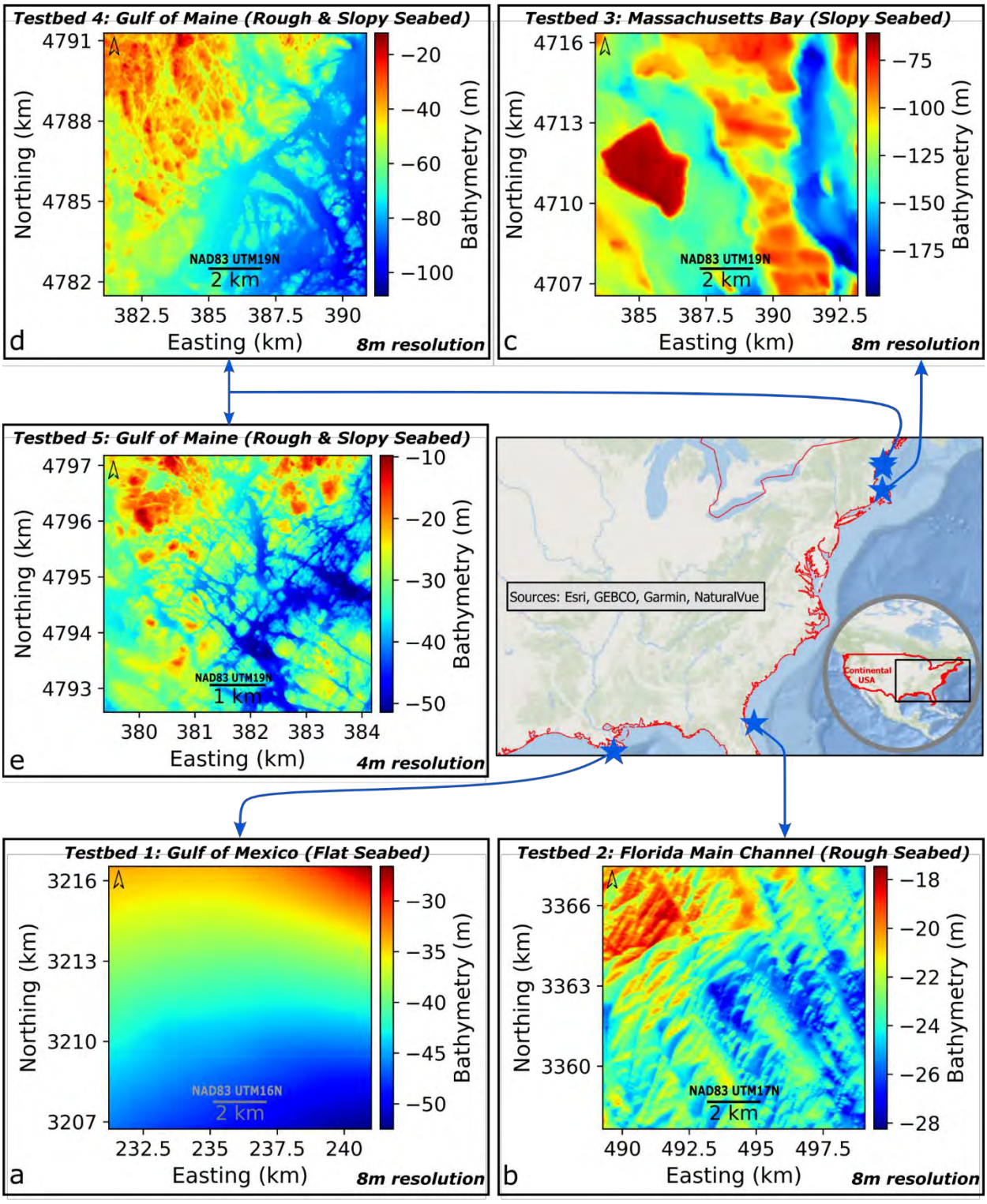

Figure 2: Testbeds. See text and Table 2 for details.

Table 2: Summary of 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 BF2G72KP_20230505	16N	8
2	Rough	High roughness and low slope	Florida Main Channel, FL	-17 – -28	BH4ST58G_20230607 BH4ST58H_20230607 BH4SV58G_20221125 BH4SV58H_20221125	17N	8
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 BF2JK2MG_20230418	19N	8
5	Special	High slope and high roughness	Gulf of Maine, ME	-10 – -51	BF2JK2MD_20230614	19N	4, 8 & 16

The NOAA BlueTopo website served as the primary source for bathymetric data. The BlueTopo repository offers both collected/measured and interpolated bathymetry. Great care was taken to exclusively utilize data that had not been subjected to interpolation in this study. Our approach involved a thorough review of NOAA National Centers for Environmental Information bathymetric attributed grids (BAGs), with only multibeam echo sounder data. These grids were scrutinized to identify areas of interest based on specific morphologies. Subsequently, the selected regions intersecting with BlueTopo tiles were downloaded, ensuring the use of non-interpolated

data with sufficient size, depth, and relatively high spatial resolution (4m and 8m) tailored to this study's analytical requirements. The non-interpolated BlueTopo tiles are grids of appropriate spatial resolution using the average depth derived from the raw high resolution multibeam echo-sounder data. As noted in the introductory section of this thesis, these non-interpolated BlueTopo depths are employed as the “true depth”.

Testbed 1, featuring a flat seafloor, spans 10km-by-10km and is located approximately 36km south of Barataria Pass, Louisiana, in the Gulf of Mexico, as depicted in Figure 2d. The water depth ranges from 26m to 52m, with geographical coordinates between UTM eastings 231111m and 241111m and UTM northings 32066338m and 3216633m (NAD83 UTM Zone 16N). The Gulf of Mexico has a complex geological history shaped by various tectonic, sedimentary, and hydrological processes spanning millions of years.

Testbed 2, characterized by a rough seafloor, spans 10km-by-10km and is located approximately 33km east of Jacksonville in the Florida main channel, as presented in Figure 2c. The water depth ranges from 17m to 28m and is situated between UTM eastings 489167m and 499167m and UTM northings 3357558m and 3367558m (NAD83 UTM Zone 17N). The general geological setting of northeastern Florida, including Jacksonville, is marked by a combination of volcanic activity and marine sedimentation during the early Ordovician Period (Lane, 1994).

Testbed 3, featuring a Slopy seafloor, spans 10km-by-10km in Massachusetts Bay, approximately 28km east of Eastern Point and 7km south of Jefferys Ledge, as illustrated in Figure 2b. It ranges from 60m to 200m in water depth, with geographic coordinates between UTM eastings 383256m and 393256m and UTM northings 4706455m and 4716455m (NAD83 UTM Zone 19N). Jeffreys Ledge and its surroundings likely owe their origin and morphologic features

to a combination of late Neogene fluvial erosion, Quaternary glaciations, and late-Pleistocene and Holocene marine processes (Uchupi and Bolmer 2008).

Testbed 4, characterized by a Rough and Slopy seafloor, spans a 10km-by-10km area in the Gulf of Maine, approximately 12km offshore in the vicinity of Bigelow Bight, as depicted in Figure 2a. Its depth range is from 12m to 102m, and it is geographically situated between UTM eastings 380978m and 390978m and UTM northings 4781382m and 4791382m (NAD83 UTM Zone 19N). The Gulf of Maine is a geologically complex area with diverse bathymetric structures resulting from a complex interplay of marine deposition, subsequent river-based deposition, and alterations by glacial erosion and deposition (Backus and Bourne 1987).

Testbed 5, also featuring a rough and Slopy seafloor but with higher resolution than other testbeds, spans 5km-by-5km in the Gulf of Maine, approximately 8km offshore in the vicinity of Bigelow Bight, as shown in Figure 2e. The water depth ranges from 10m to 51m, with geographic coordinates between UTM eastings 379270m and 384266m and UTM northings 4792478m and 4797274m (NAD83 UTM Zone 19N). Testbed 5 is unique in providing resolutions of 4m, 8m, and 16m, enabling an exploration of the impact of spatial resolution on uncertainty quantification. Notably, Testbed 5 shares the same geology as Testbed 4.

3.6 Terrain Characterization

Aguilar *et al*. (2005) identified seabed morphology as the paramount factor influencing interpolation quality. Likewise, Guo *et al*., 2010 associated interpolation uncertainty with the variability in terrain elevation. In the context of this study, terrain characteristics, specifically slope and roughness, assume centrality in assessing the impact of seabed morphology on uncertainty estimation. Although curvature and aspect were also initially considered, they are omitted from this work as preliminary analysis indicated that they exhibited no 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.

3.6.1 Slope

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

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

To calculate slope (or other terrain parameter such as roughness) for each pixel, an analysis window is effectively moved across the raster DBM surface such that each pixel in turn becomes the central or subject pixel on which calculations are based (refer to Figure 3). The resulting calculations are still reported at the original pixel size; it is merely the window size or ground area considered in the analysis which varies. This generalization allows the parameter to be analyzed at a range of scales (different odd values of $n \geq 3$) (Wilson *et al.* 2007). This study investigated three window sizes — 3 pixels by 3 pixels, 5 pixels by 5 pixels, and 7 pixels by 7 pixels, see the spatial scale section of this chapter. Figure 3 presents a raster grid, showing a numbering system for cells in an analysis window where Z is the depth of a given raster cell. The central cell is the origin of the local coordinate system (x, y) and the positions relative to this are denoted by subscripts. To simplify notation, we use $N = (n - 1)/2$ for any $n \times n$ analysis window where n may be any odd integer. These are shown in full for a 3×3 window. Larger values of n indicate that more cells (larger area) are considered in the analysis.

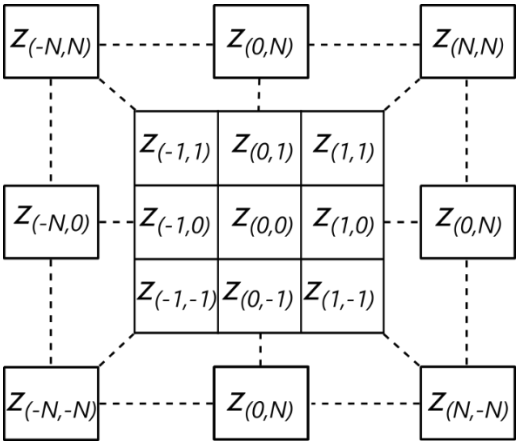

Figure 3: An n x n analysis window for a raster grid.

In this study, Horn's finite difference method (Horn, 1981) was employed to calculate fx and fy of the central point of the analysis window (x, y = 0). Horn's method utilizes convolution kernels tailored to each moving analysis window to estimate a value for a subject pixel. By calculating the weighted sum of neighboring elevation values, Horn's method provides slope approximations in both the east-west (X) and north-south (Y) directions of the analysis window. Using the notation in Figure 3 for a 3×3 analysis window, equations 1 and 2 provide Horn's formulae for calculating fx and fy:

$$fx = \frac{(Z_{-1,1} - Z_{1,-1} + 2(Z_{-1,0} - Z_{1,0}) + Z_{-1,1} - Z_{1,1})}{8L} \quad (2)$$

$$fy = \frac{(Z_{1,1} - z_{1,-1} + 2(Z_{0,1} - Z_{0,-1}) + Z_{-1,1} - Z_{-1,-1})}{8L} \quad (3)$$

Here, L represents the data cell size or data resolution, which is crucial for normalizing the gradients with respect to the grid spacing, ensuring that the calculated slope values are consistent and independent of the specific resolution or scale of the depth data.

3.6.2 Roughness

Roughness is defined as the variability or irregularity in elevation—capturing both highs and lows—within a sampled terrain unit. It is crucial to precisely define the scale or level at which terrain roughness is considered, along with the unit used to measure across a 'window' of scale, as various attributes become relevant at different scales (Smith 2014). Consequently, the definition of terrain surface roughness is often ambiguous (Fan 2022a). Roughness indices usually rely on a quantitative description of specific terrain characteristics changes, such as the degree of local undulation, the degree of local folds, or the degree of local abrupt changes(Fan 2022b). In this study, roughness is calculated using Equation 4 as defined by Wilson *et al*. (2007), which computes the largest inter-cell difference of a central pixel and its surrounding cells in an n x n rectangular neighborhood (refer to Figure 3).

$$R(n) = Bmax - Bmin \quad (4)$$

Bmax(n) = maximum Z in n × n window

Bmin(n) = minimum Z in n × n window

To ensure the accurate representation of their respective seabed morphologies, the testbeds underwent both quantitative assessments for slope and roughness and qualitative assessments through visual inspection. The results of this evaluation are presented in Figure 4.

Figures 4a & 4b showcase the slope and roughness histograms of Testbed 1, respectively, while 4c & 4d depict the slope and roughness histograms of Testbed 2. In a similar fashion, 4e & 4f represent the slope and roughness histograms of Testbed 3, 4g & 4h illustrate the slope and roughness histograms of Testbed 4, and finally, Figures 4i & 4j portray the slope and roughness histograms of Testbed 5. It is important to note that panels 4e to 4j highlight the 99th percentile of

data, plotted to improve visualization and facilitate effective comparisons between the seabed morphologies, while panels 4a to 4f 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 as desired to achieve the goals of this thesis.

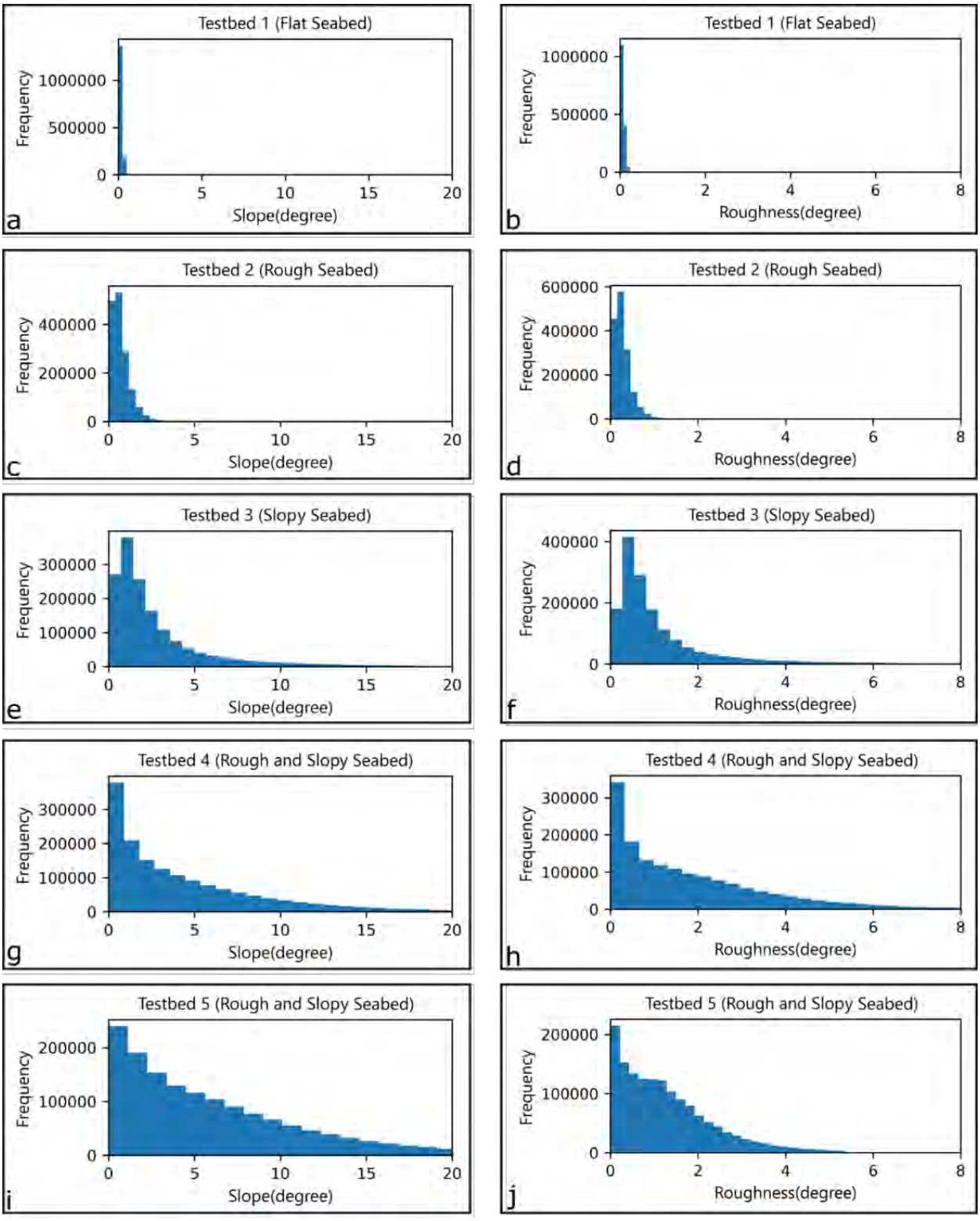

Figure 4: Histograms depicting the slope and roughness characteristics of the original testbeds' bathymetric datasets.

3.7 Spatial Scales

Spatial scales, in the context of this study which only applies to Chapter 4, refer to the various resolutions and sizes at which bathymetric data and terrain characteristics are analyzed (Gustafson 1998). Terrain attributes vary with scale (Chorley n.d.). Thus, their computation does not result in only one true, real fixed value, but in a range of possible values that depend on the resolution of the data and the extent of the analysis window (Hengl 2006). The consideration of spatial scales is pivotal in understanding how different features manifest at different levels of granularity and how this impacts the estimation of interpolation uncertainty. This investigation navigates spatial scales by considering analysis window sizes applied in computing terrain characteristics, specifically slope and roughness, and by exploring different datasets' spatial resolutions. The study implements three window sizes—3 pixels by 3 pixels, 5 pixels by 5 pixels, and 7 pixels by 7 pixels — utilizing the developed algorithms for roughness and Horn's slope calculation in Python (Horn 1981). While the 3-by-3 analysis window was 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 4. The impact of spatial resolution on interpolation uncertainty was investigated focusing on only Testbed 5. Analyzing bathymetric data at various scales provides insights into the influence of local and regional factors on interpolation uncertainty, allowing for a better comprehension of how the interaction between interpolation uncertainty and ancillary parameters varies with changes in spatial scale.

CHAPTER 4 : ESTIMATING AND CHARACTERIZING INTERPOLATION UNCERTAINTY IN SPARSE BATHYMETRIC DERIVED DIGITAL BATHYMETRIC MODELS USING ANCILLARY PARAMETERS

4.1 Introduction

This chapter investigates the characterization and estimation of the uncertainty in DBMs derived from sparse archived bathymetric datasets. The interest in interpolating across numerous cells arises from the need to fill gaps in sparse hydrographic soundings collected using lead lines and single-beam sonar technology. This interpolation between sparse bathymetric soundings is typically required for many applications including nautical charting, tsunami propagation, and inundation modeling (Hare *et al.* 2011), etc. Hydrographic offices worldwide face the challenge of designating a CATZOC classification for areas having sparse soundings/depths from SBES, older technology of MBES, wire drags, and lead line surveys without full seabed coverage.

4.2 Methods

The split-sample methodology is employed to simulate a sparse bathymetric data set by randomly sampling the testbeds' depth measurements based on sampling density before interpolation. 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 chapter 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 5). The split-sample methodology undergoes 10 iterations for each interpolation technique, at each of the five sampling densities, and on the five testbeds. The deterministic interpolation techniques employed are IDW, Spline, and Linear. The 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; 10 iterations were deemed sufficient to fairly represent bathymetric variability in a 10km-by-10km area.

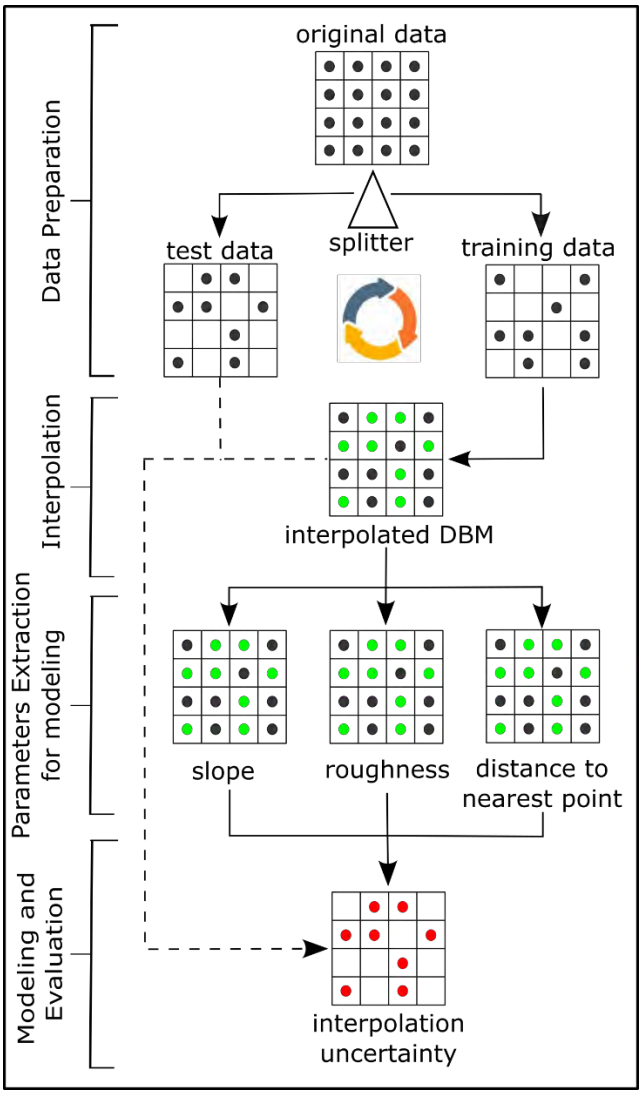

Figure 5: Workflow for quantifying interpolation uncertainty in sparse bathymetric datasets.

During each split-sample iteration, training depths were gridded using the specified interpolation technique. The resulting interpolated raster was then compared, on a cell-by-cell basis, to the test depths to quantify interpolation uncertainty. To clarify, interpolation uncertainty was determined by taking the absolute value of the difference between the interpolated depths and the measured depths. Ancillary parameters, such as the Euclidean distance to the nearest measurement for each test cell was generated from the closest cell in the training data set and the slope and roughness raster surfaces, were generated from the interpolated DBM after each split-sample routine (Figure 6). These parameters, compiled from each split-sample routine, along with their respective interpolation uncertainties, constituted the datasets prepared for modeling to discern the relationship between the parameters and interpolation uncertainties. Notably, the analysis to identify the interpolator producing the lowest uncertainty using box and whisker plots and descriptive statistics was based on the median of the statistics from the 10 raster surfaces. The only exception, which is the spatial distribution of the interpolation uncertainty was based on a randomly selected surface from the 10 raster surfaces.

Figure 5 provides a visual representation of the entire process, illustrating the ***data preparation*** phase, which involves the random splitting of original data into training and test datasets using 50% sampling density. The subsequent steps include ***interpolation***, where various interpolation techniques are applied, ***parameter extraction for modeling***, encompassing the calculation of slope and roughness from the interpolated DBM and the determination of the distance to the nearest measurement from the training data. The final step, ***modeling and evaluation***, compares interpolated and test data to quantify interpolation uncertainty, and models the relationship between interpolation uncertainty and ancillary parameters, both individually and in combination. The model's performance is then evaluated.

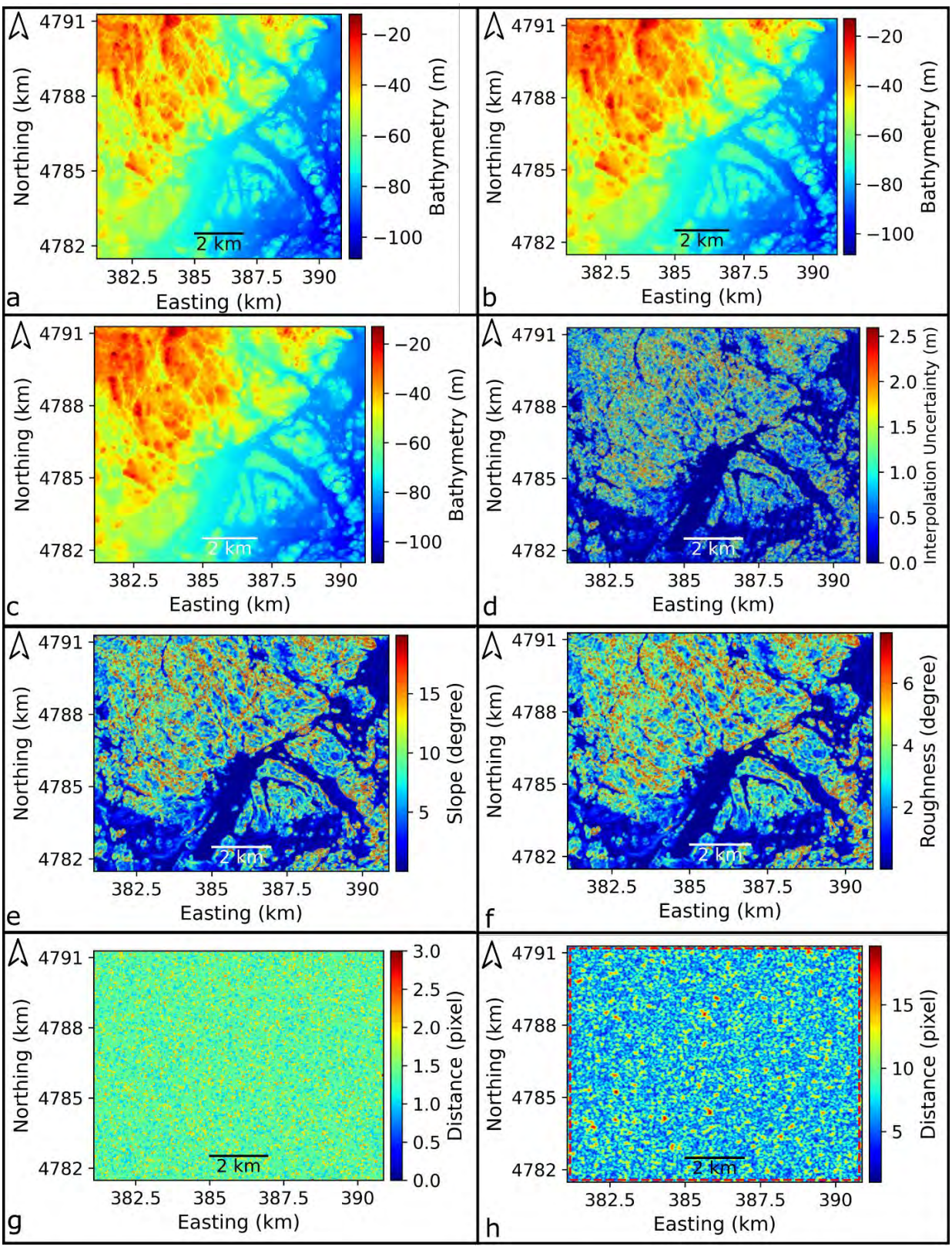

Figure 6: Testbed 4 measured depth with 50% sampling density (a), training depths (b), IDW interpolated depths (c), interpolation uncertainty clipped to 99 percentile (d), generated slope from

interpolated depths (e), generated roughness from interpolated depths (f), distance to the nearest depth measurement raster (g), and distance to the nearest depth measurement raster (1% sampling density), the inner dashed red lines shows the data "buffer" along the border to guide the interpolation (h).

To mitigate potential bias in the analysis, precautionary measures were implemented. Specifically, while depth measurements were interpolated across the entire area, the analysis was confined to a subset reduced by 12 cells (96m) on each side. This approach aimed to address potential edge effects by ensuring that interpolated values near the boundaries of the study area were influenced by actual measured depths only, thereby minimizing inaccuracies that could arise from extrapolation. This ensured that the edge of the area of analysis had a range of distances to the nearest measurement without biasing toward depths along the outermost border of the study area (refer to Figure 6h). It also guaranteed accurate slope and roughness values by maintaining appropriate window size for calculation.

Notably, the algorithms for these interpolators were developed in Python such that minimal adjustment of interpolation parameters was required to undertake the analysis. This deliberate approach aimed 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. The Python libraries were Numpy, Pandas, Matplotlib, GDAL, SciPy, Scikit-learn, Keras, and Statsmodel. Linear and Spline interpolations are implemented using the Python SciPy griddata library with “method” set to linear and cubic respectively. IDW interpolation is implemented using Scikit Learn Neighbors.

4.3 Modeling 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 employs machine learning techniques to capture non-linearities, interactions, and hidden relationships among the combined parameters. While multivariate regression (MVR) (see Jobson 1991) and random forest (RF) (see Breiman, 2001) techniques were initially explored, only Artificial Neural Network (ANN) (see Agatonovic-Kustrin and Beresford 2000) is utilized in this work due to its improved accuracy. For assessing importance of individual variables, this study recognized the computational efficiency of RF's feature importance compared to ANN's SHAP (SHapley Additive exPlanations) (see Lundberg and Lee 2017) that can be used to identify the most important predictor of interpolation uncertainty. The MVRs that included all 2-way and single 3-way multiplicative interactions of the parameters were also explored. This investigation revealed that the enhanced performance of ANN was attributable to its adept handling of both interactions and non-linearities inherent in the parameters.

The ANN was implemented in Python using the Keras library with a TensorFlow backend. The dataset was split into 70% training data and 30% test data, and standard scaling normalization was applied to both sets. The model architecture included an input layer with 10 neurons, a hidden layer with five neurons using the Rectified Linear Units (ReLU) activation function (Fred Agarap, 2018), and an output layer with one neuron employing linear activation for regression tasks. The training involved specifying mean squared error (MSE) as the loss function, using the Adam optimizer for gradient descent (Kingma and Ba 2014), and monitoring mean absolute error (MAE)

during five epochs of training. The model predicted interpolation uncertainty for the test data, and performance was assessed using adjusted R^2 and RMSE.

The RF implementation in Python utilized the Scikit Learn Random Forest Regressor. Data preparation included splitting them into training and testing sets and normalizing the features. An RF model with 10 decision trees was created and trained on the standardized training data. Performance evaluation metrics included adjusted R^2 and RMSE. The importance of each parameter in the 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. For computational efficiency and analytical accuracy, a substantial sample of the data, yielding an average of 120,000 data points, underwent bootstrapping 500 times. Subsequently, the RF model is fitted to each bootstrapped dataset iteratively, totaling 500 repetitions. The variable importance values generated by each RF model were aggregated to compute the mean importance and establish the 95% confidence intervals. The overall mean importance and 95% confidence intervals of each variable were computed by averaging across all sampling densities and interpolation methods. It is important to note that a sample of data was used for computational efficiency. Using all the data would cause any significant differences to be more significant because frequency distributions would narrow, and the degrees of freedom would increase.

4.4 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 (see Figure 5), 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, MAE and R^2 are used in this study in conjunction with other descriptive statistics such as minimum, maximum, and median error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Z_i^{interpolated} - Z_i^{true})^2}{N}} \quad (5)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{n}\left|Z_i^{interpolated} - Z_i^{true}\right| \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Z_i^{true} - Z_i^{interpolated})^2}{\sum_{i=1}^{n}(Z_i^{true} - \overline{Z_i^{true}})^2} \quad (7)$$

where $z^{interpolated}$ is the predicted depth for the *i*-th pixel,

z^{true} is the true depth from BlueTopo for the *i*-th pixel, and

N is the total number of pixels.

Pairwise t-tests (Ross and Willson 2017) were employed to determine the statistical significance of differences in interpolation methods, following an investigation into both parametric (ANOVA) and non-parametric (Kruskal-Wallis) techniques. These initial methods faced challenges as our datasets violated underlying assumptions, such as data independence, and variance homoscedasticity. The pairwise t-tests focused on evaluating differences between pairwise interpolation methods' uncertainties and comparing them with zero. The pairwise difference distributions were normally distributed and are used to corroborate the findings of t-

tests. This approach effectively addressed the aforementioned challenges, ensuring a thorough and reliable assessment of differences among the interpolation techniques.

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.

4.5 Results

The presentation of results is structured based on testbeds as outlined in the subsequent subsections.

4.5.1 Testbed 1 (Flat Seabed)

4.5.1.1 Interpolation Methods

Figure 7 displays the box and whisker plots illustrating the performance of interpolation methods across all sampling densities for Testbed 1. Complementing these visualizations, Table 3 provides descriptive statistics of interpolation uncertainties at the 99th percentile confidence interval for each sampling density and interpolation method. Notably, the interpolation uncertainties of the interpolation methods are relatively similar at the same sampling density with Linear interpolation performing slightly better than IDW and Spline. It is highlighted that the interpolation uncertainties here are in the order of centimeters, attributed to the less complex morphology of Testbed 1. The simpler morphology is also a contributing factor to the Linear interpolation method exhibiting the lowest uncertainty compared to the other two methods. Table 4 and Figure 8 collectively demonstrate that there are no statistically significant differences in the performance of interpolation methods at each sampling density. This assertion is supported by the t-test results, and the distributions of pair-wise interpolation uncertainty differences centered around zero.

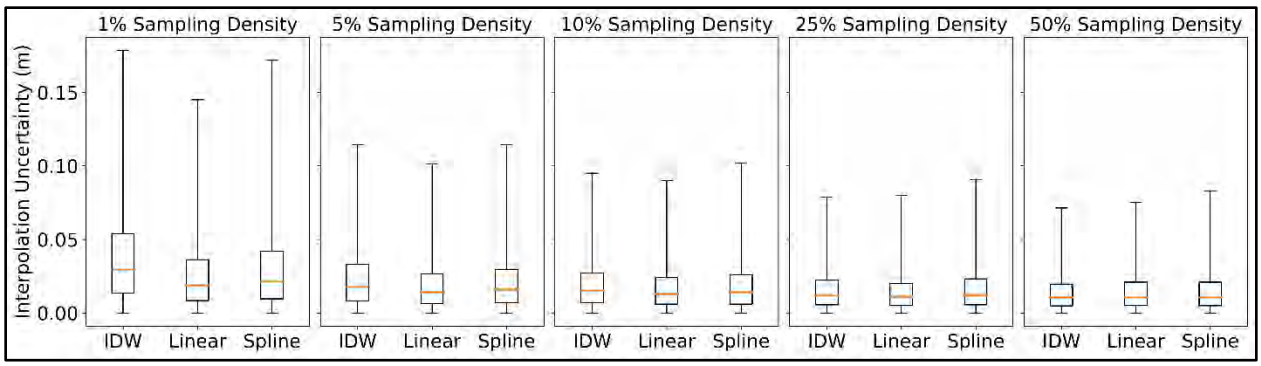

Figure 7: Testbed 1 interpolation methods uncertainty comparison at various sampling densities using box and whisker, plotted with 99% percentile of data.

Table 3: Testbed 1 interpolation uncertainty descriptive statistics at various sampling densities.

S/N	Sampling density (%)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error (m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	1	IDW	0.000	0.179	0.029	0.038	0.051
2	1	Linear	0.000	0.145	0.019	0.027	0.037
3	1	Spline	0.000	0.172	0.021	0.031	0.043
4	5	IDW	0.000	0.114	0.018	0.024	0.032
5	5	Linear	0.000	0.101	0.014	0.019	0.026
6	5	Spline	0.000	0.115	0.016	0.022	0.029
7	10	IDW	0.000	0.095	0.015	0.020	0.026
8	10	Linear	0.000	0.090	0.013	0.017	0.023
9	10	Spline	0.000	0.102	0.014	0.019	0.026
10	25	IDW	0.000	0.079	0.012	0.017	0.021
11	25	Linear	0.000	0.080	0.011	0.016	0.021
12	25	Spline	0.000	0.091	0.012	0.017	0.023
13	50	IDW	0.000	0.071	0.011	0.014	0.019
14	50	Linear	0.000	0.075	0.010	0.015	0.020
15	50	Spline	0.000	0.083	0.011	0.015	0.021

Table 4: Testbed 1 statistics for pairwise interpolation methods comparison at various sampling densities.

Sampling Density (%)	Interpolation Methods	t-statistics	p-value
1	Spline and IDW	-142.8	0
	IDW and Linear	258.0	0
	Spline and Linear	161.7	0
5	Spline and IDW	-59.6	0
	IDW and Linear	169.2	0
	Spline and Linear	125.3	0
10	Spline and IDW	-3.8	< 0.01
	IDW and Linear	115.4	0
	Spline and Linear	137.6	0
25	Spline and IDW	43.3	0
	IDW and Linear	50.3	0
	Spline and Linear	135.9	0
50	Spline and IDW	44.5	0
	IDW and Linear	17.7	< 0.01
	Spline and Linear	99.7	0

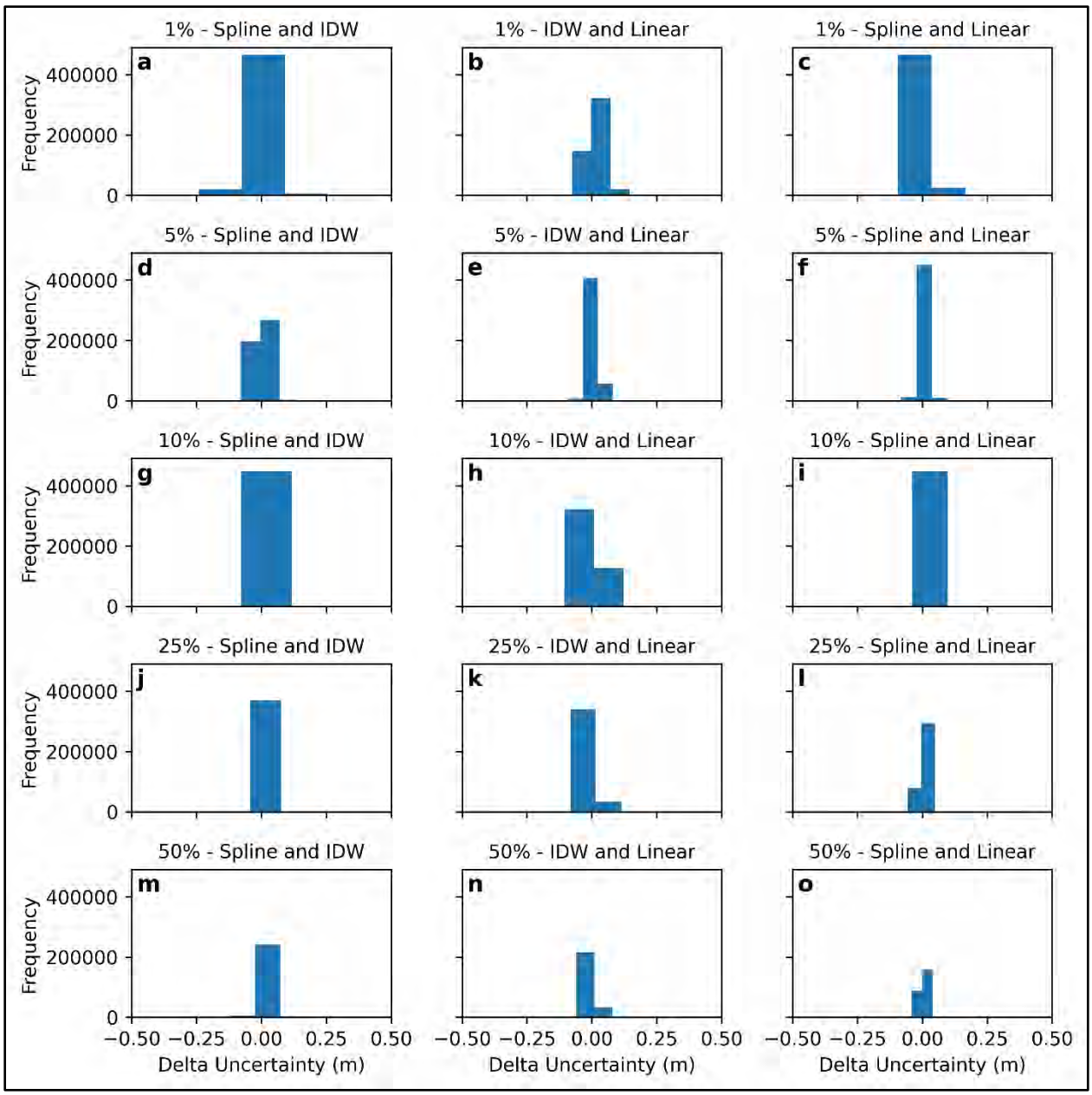

Figure 8: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 1.

4.5.1.2 Spatial Pattern of Interpolation Uncertainties

Figure 9 illustrates the spatial distribution of interpolation uncertainties of the interpolation methods across all sampling densities for Testbed 1. To facilitate comparison, the color bar has been standardized across sampling densities and interpolation methods. An intriguing observation is that the interpolation uncertainties do not exhibit a random pattern across the testbed with all

sampling densities. Instead, they are notably concentrated on the eastern side of the plots, attributed to multibeam artifacts (strips) present in the original datasets from BlueTopo (see figure 2a).

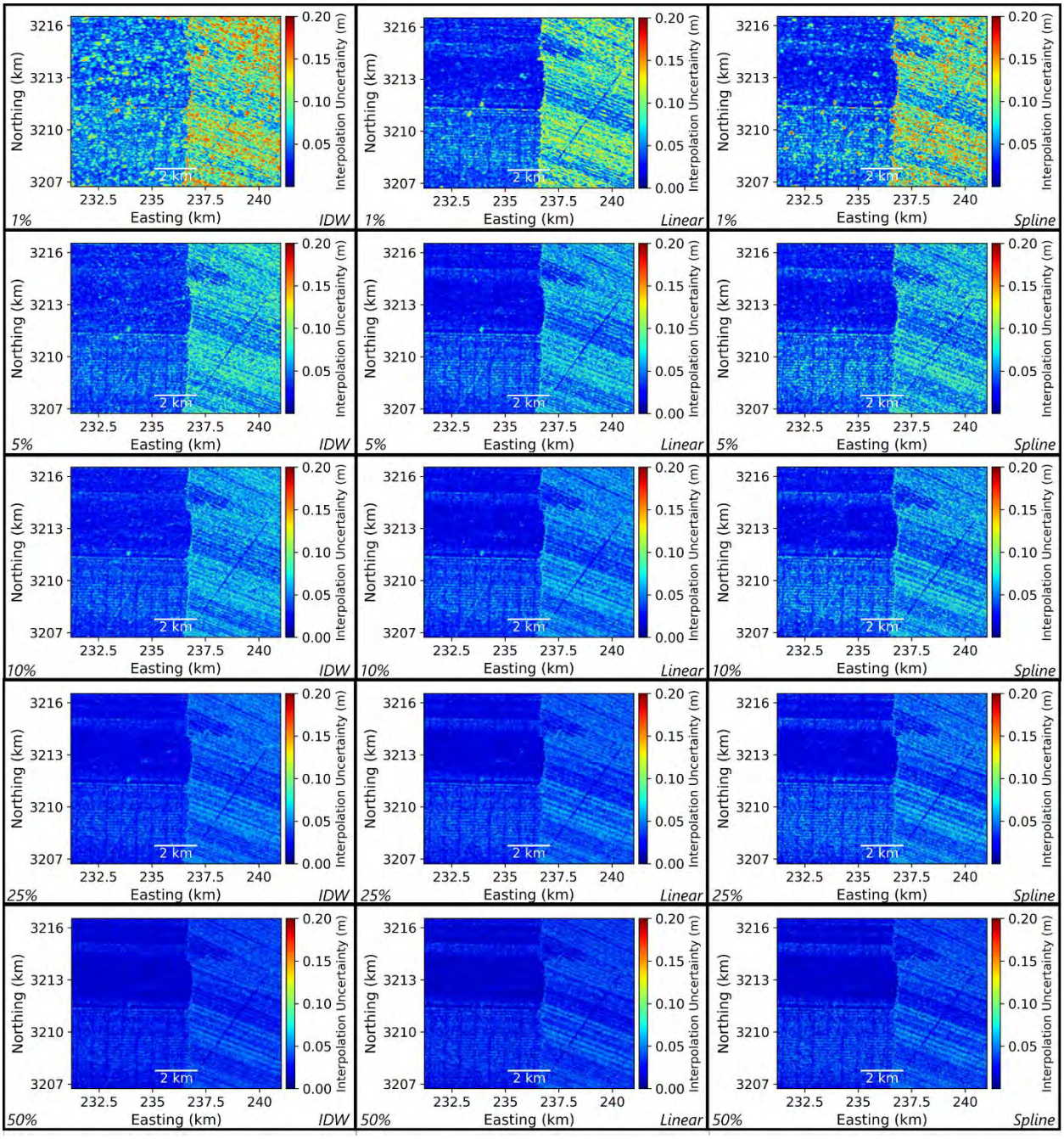

Figure 9: Testbed 1 interpolation uncertainty across all sampling densities (columns) and interpolation methods (rows), using 99th percentile of data.

4.5.1.3 Predictive Models of Interpolation Uncertainties

Figures 10a&d present the adjusted R^2 and RMSE values for the estimated uncertainty for Testbed 1, assessed in relation to the distance to the nearest measurement across various cell sampling densities (1%, 5%, 10%, 25%, and 50%). The adjusted R^2 values indicate a weak linear relationship between the distance to the nearest measurement and interpolation uncertainty, with all interpolation methods performing relatively the same. Additionally, higher sampling densities correspond to decreased RMSE values, with Linear interpolation outperforming IDW and Spline.

Figures 10b&e (Testbed 1) present the adjusted R^2 and RMSE values for estimated uncertainty based on roughness. A weak relationship intensifies with increased sampling density. Spline interpolation attains the highest adjusted R^2 but demonstrates the poorest performance in RMSE. IDW follows, ranking second based on both R^2 and RMSE, while Linear interpolation performs the worst in adjusted R^2 but performs the best in RMSE.

Figures 10c&f (Testbed 1) present the adjusted R^2 and RMSE values for estimated uncertainty based on slope. Similar to roughness, a weak relationship intensifies with increasing sampling density. In this case, Spline interpolation again performs best in adjusted R^2, while Linear interpolation performs the worst in RMSE.

To investigate the hidden non-linear relationships between the combined parameters and the estimated uncertainty, we employed an Artificial Neural Network (ANN). Results indicate that combining distance to the nearest measurement, slope, and roughness in a model improves the single variable model's performance, albeit with marginal enhancements in both R^2 and RMSE (Figure 11). Among the interpolation techniques, Spline performs best in R^2 but performs the worst in RMSE which is corroborated by findings in Figure 10. IDW follows Spline in adjusted R^2 and

outperforms it in RMSE. Linear interpolation exhibits the lowest adjusted R^2 but performs best in RMSE.

In general, the differences in the interpolation uncertainty models' RMSEs are not substantial from an operational perspective because they are in the order of centimeters. Additionally, the performance of the uncertainty models improves at higher sampling densities, as evidenced by both individual parameter and combined parameters' interpolation uncertainty models. However, the disparities in model performance become negligible beyond a 10% sampling density. Therefore, 10% emerges as the optimal operational sampling density, representing a 'sweet spot' for Testbed 1. This sweet spot denotes the sampling density beyond which higher values do not significantly contribute to uncertainty estimation.

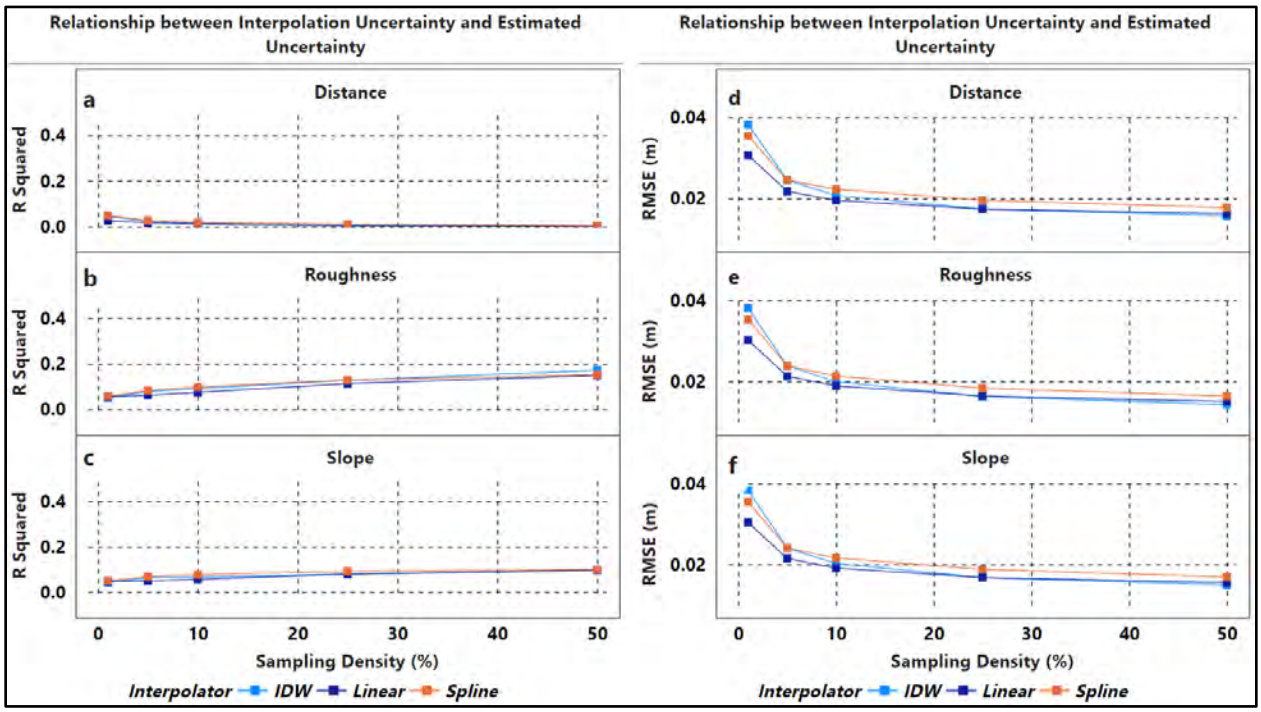

Figure 10: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 1.

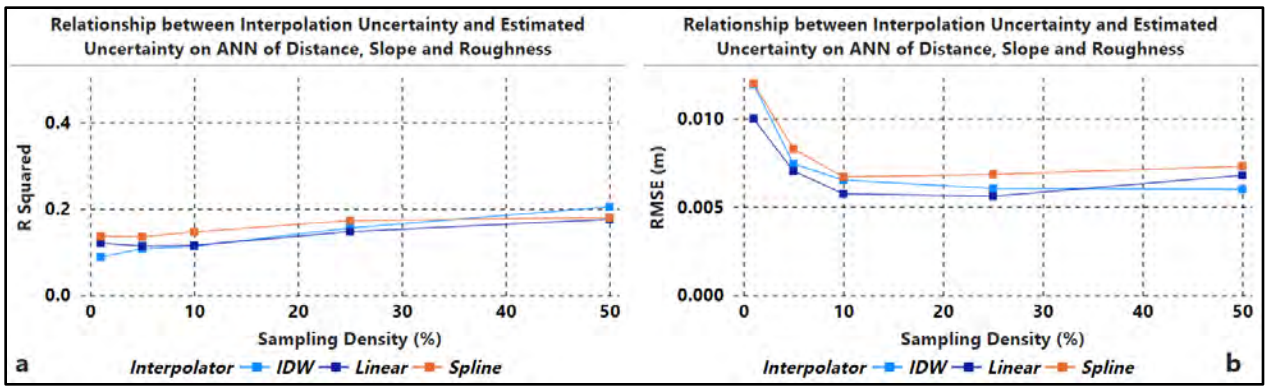

Figure 11: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 1.

Furthermore, the application of the RF machine learning algorithm alongside the bootstrap statistical technique reveals that roughness is the most influential predictor ranking as the most important in 10 out of 15 instances, with a mean importance level of 45.3% (Table 5). These instances stem from the combination of three interpolation methods and five sampling densities. Following closely is slope, identified as the most important in five out of 15 instances with a mean importance of 46.5%. Conversely, distance to nearest measurement is the least important predictor, not appearing as most important in any instance and holding a mean importance of 8.2%. Crucially, the bootstrap technique confirms the statistically significant differences in these parameters.

Table 5: Statistics of the importance of predictors of uncertainty for Testbed 1.

Statistics/Parameter	Distance	Slope	Roughness
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)

4.5.2 Testbed 2 (Rough Seabed)

4.5.2.1 Interpolation Methods

Figure 12 displays the box and whisker plots illustrating the performance of interpolation methods across all sampling densities for Testbed 2. Complementing these visual representations, Table 5 provides descriptive statistics of interpolation uncertainties at the 99th percentile confidence interval for each sampling density and interpolation method. Similar to Testbed 1, the interpolation uncertainties of the interpolation methods are relatively the same across the same sampling density. However, on Testbed 2, the Spline interpolation performs slightly better than Linear and IDW. Furthermore, the interpolation uncertainties here are in the order of centimeters, attributed to the less complex morphology of Testbed 2 relative to Testbeds 3 and 4. Table 7 and Figure 13 collectively show that there are no statistically significant differences in the interpolation methods at each of the sampling densities.

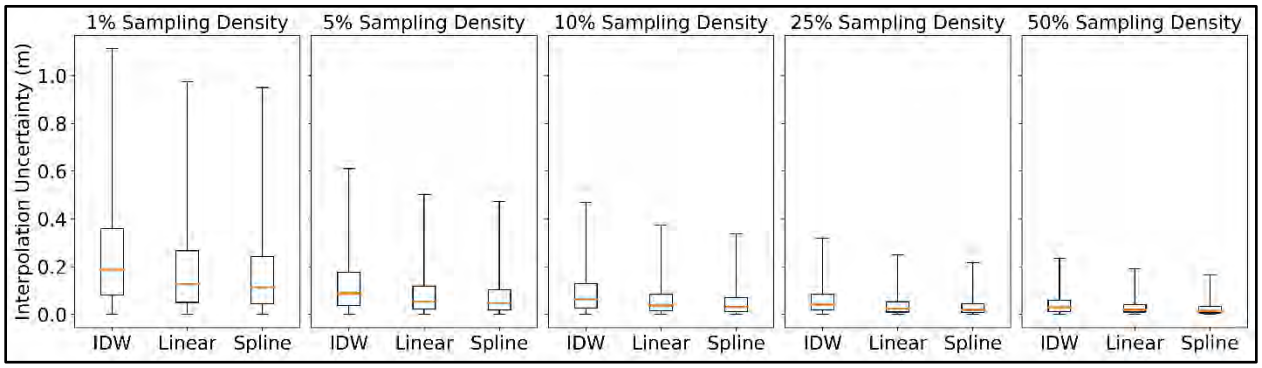

Figure 12: Testbed 2 interpolation methods uncertainty comparison at various sampling densities using box and whisker, plotted with 99% percentile of data.

Table 6: Testbed 2 interpolation uncertainty descriptive statistics at various sampling densities.

S/N	Sampling density (%)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error (m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	1	IDW	0.000	1.113	0.187	0.251	0.335
2	1	Linear	0.000	0.975	0.126	0.188	0.264
3	1	Spline	0.000	0.950	0.113	0.173	0.247
4	5	IDW	0.000	0.611	0.089	0.125	0.170
5	5	Linear	0.000	0.503	0.054	0.086	0.125
6	5	Spline	0.000	0.474	0.046	0.076	0.113
7	10	IDW	0.000	0.469	0.064	0.092	0.126
8	10	Linear	0.000	0.375	0.038	0.061	0.090
9	10	Spline	0.000	0.337	0.031	0.052	0.078
10	25	IDW	0.000	0.319	0.041	0.060	0.084
11	25	Linear	0.000	0.248	0.025	0.040	0.059
12	25	Spline	0.000	0.217	0.019	0.033	0.050
13	50	IDW	0.000	0.234	0.029	0.043	0.060
14	50	Linear	0.000	0.190	0.020	0.031	0.045
15	50	Spline	0.000	0.165	0.015	0.025	0.038

Table 7: Testbed 2 statistics for pairwise interpolation methods comparison at various sampling densities.

Sampling Density (%)	Interpolation Methods	t-statistics	p-value
1	Spline and IDW	-142.8	0
	IDW and Linear	258.0	0
	Spline and Linear	161.7	0
5	Spline and IDW	-59.6	0
	IDW and Linear	169.2	0
	Spline and Linear	125.3	0
10	Spline and IDW	-3.8	< 0.01
	IDW and Linear	115.4	0
	Spline and Linear	137.6	0
25	Spline and IDW	43.3	0
	IDW and Linear	50.3	0
	Spline and Linear	135.9	0
50	Spline and IDW	44.5	0
	IDW and Linear	17.7	< 0.01
	Spline and Linear	99.7	0

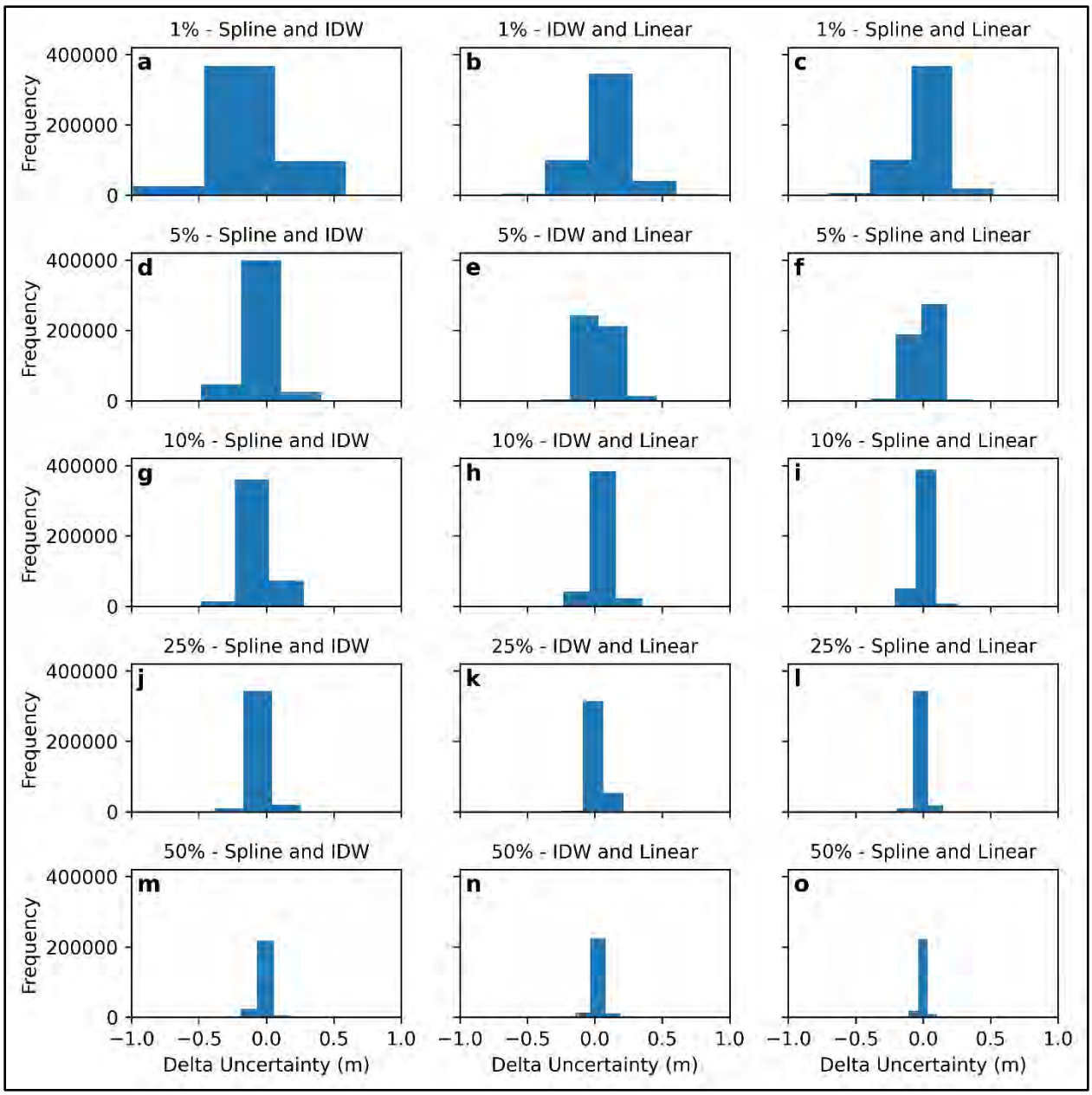

Figure 13: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 2.

4.5.2.2 Spatial Pattern of Interpolation Uncertainties

Figure 14 shows that the spatial distribution of interpolation uncertainties on Testbed 2 is not random. The color bar, standardized for ease of comparison, highlights that areas with higher interpolation uncertainties are correlated with high roughness values (see Figure 2b for comparison). This correlation points to the influence of seabed roughness on uncertainty patterns.

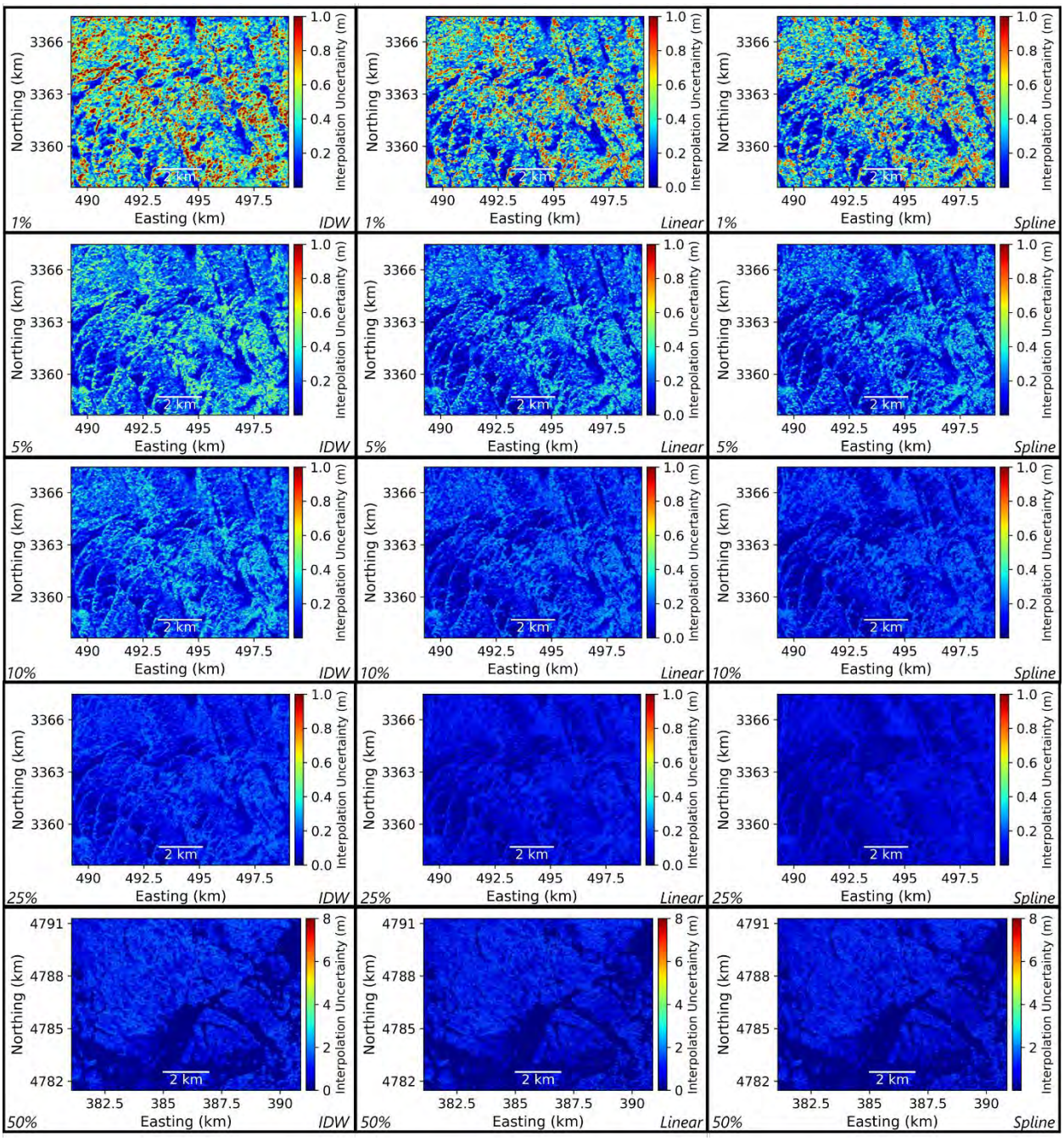

Figure 14: Testbed 2 interpolation uncertainty across all sampling densities (columns) and interpolation methods (rows), using 99th percentile of data.

4.5.2.3 Predictive Models of Interpolation Uncertainties

Figures 15a&d present the adjusted R^2 and RMSE values for estimated uncertainty based on the distance to the nearest measurement for Testbed 2. Surprisingly, a very weak relationship is

observed, which diminishes with increasing sampling density. Spline interpolation consistently outperforms Linear and IDW interpolation in both metrics.

Figures 15b&e (Testbed 2) present the adjusted R^2 and RMSE values for estimated uncertainty based on roughness. A weak relationship emerges, intensifying with increased sampling density. IDW achieves the highest adjusted R^2 but performs the worst in RMSE. Linear and Spline interpolations exhibit similar performance levels.

Figures 15c&f (Testbed 2) depict the adjusted R^2 and RMSE values for estimated uncertainty based on slope. Similar to roughness, a weak relationship strengthens with increasing sampling density. In this case, IDW interpolation again performs best in adjusted R^2 and performs the worst in RMSE, while Linear and Spline interpolations exhibit similar performance levels.

To investigate the hidden non-linear relationships, we employed an ANN, revealing marginal improvements in both R^2 and RMSE (Figure 16). Among interpolation techniques, IDW performs the best in R^2 but performs the worst in RMSE. Linear and Spline interpolations exhibit similar performance levels.

Much like Testbed 1, the distinctions in RMSE among the interpolation uncertainty models are not operationally significant, given that they are on the order of centimeters. Additionally, the variations in the performance of uncertainty models become inconsequential beyond a 10% sampling density, thereby designating the 10% sampling density as the optimal sampling density for Testbed 2.

Results from the RF analysis coupled with the bootstrap statistical technique reveal that roughness emerges as the most influential predictor ranking as the most important in 10 out of 15 instances, with a mean importance level of 46.1% (Table 6). It is followed by slope with mean importance

of 43.9% and appearing as the most important in five of 15 instances. Distance to the nearest measurement is the least important predictor, with a mean importance of 10%. Importantly, the bootstrap technique confirms the statistical significance of the differences in these parameters.

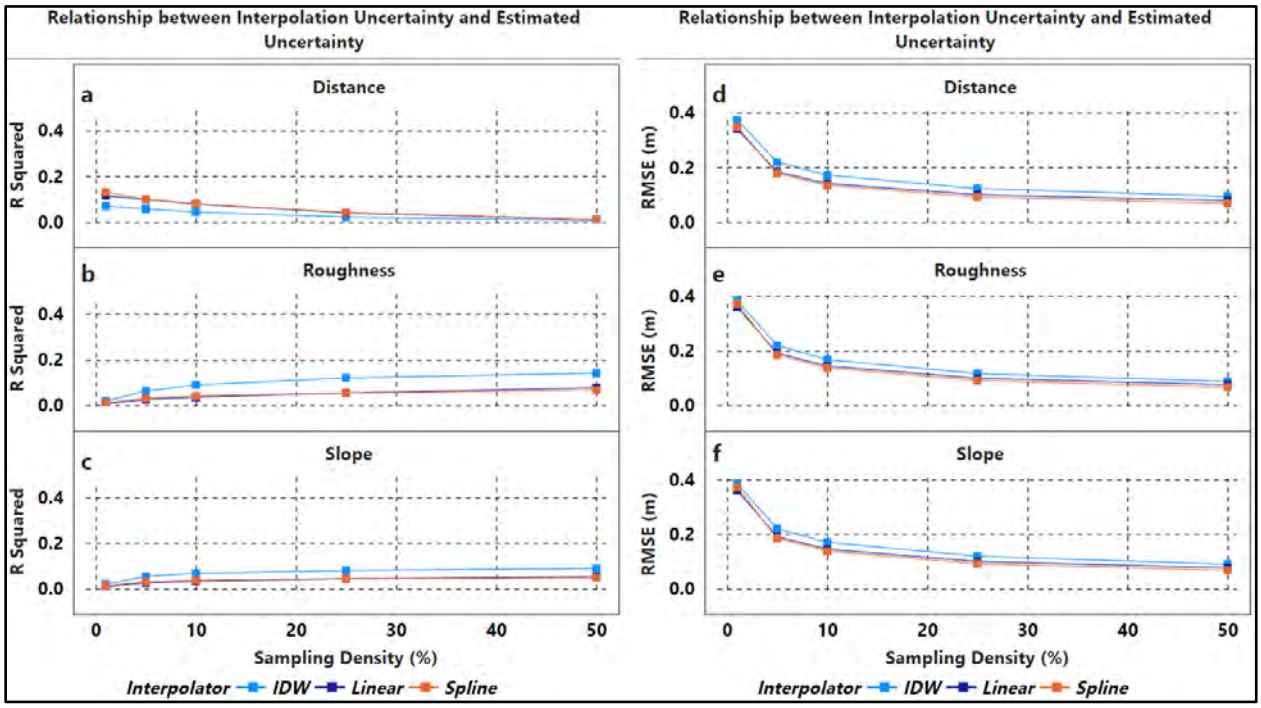

Figure 15: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 2.

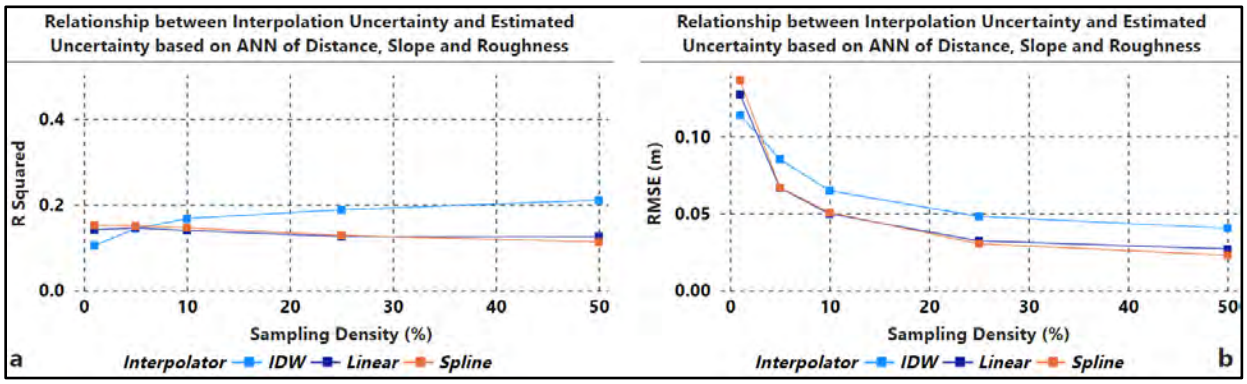

Figure 16: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 2.

Table 8: Statistics of the importance of predictors of uncertainty for Testbed 2.

Statistics/Parameter	Distance	Slope	Roughness
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)

4.5.3 Testbed 3 (Slopy Seabed)

4.5.3.1 Interpolation Methods

Figure 17 displays the box and whisker plots illustrating the performance of interpolation methods across all sampling densities for Testbed 3. To complement these visualizations, Table 9 presents the descriptive statistics of interpolation uncertainties at the 99th percentile confidence interval for each sampling density and interpolation method. The interpolation uncertainties of the interpolation methods vary at lower sampling densities with Spline being the best interpolation method followed by Linear and IDW. The interpolation uncertainties become relatively the same at higher sampling densities, 25% and 50%. At the lower sampling densities, the Spline interpolation method performs better than Linear and IDW; and at higher sampling densities, Linear interpolation performs the best. Compared to Testbed 1 and 2, the interpolation uncertainties here are in the order of meters, attributed to the more complex morphology of Testbed 3 compared to Testbeds 1 and 2. Importantly, Figure 18 and Table 10 affirm that the observed differences in the descriptive statistics and box and whisker plots are not statistically significant.

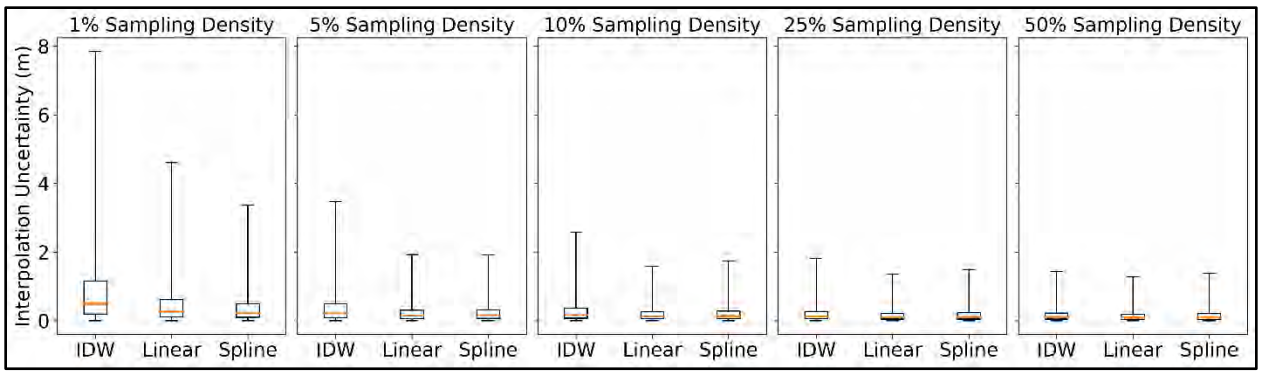

Figure 17: Testbed 3 interpolation methods uncertainty comparison at various sampling densities using box and whisker, plotted with 99% percentile of data.

Table 9: Testbed 3 interpolation uncertainty descriptive statistics at various sampling densities.

S/N	Sampling density (%)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error (m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	1	IDW	0.000	7.852	0.496	0.946	1.547
2	1	Linear	0.000	4.605	0.267	0.512	0.854
3	1	Spline	0.000	3.376	0.232	0.402	0.638
4	5	IDW	0.000	3.470	0.223	0.406	0.656
5	5	Linear	0.000	1.930	0.148	0.244	0.376
6	5	Spline	0.000	1.918	0.156	0.250	0.380
7	10	IDW	0.000	2.571	0.173	0.304	0.486
8	10	Linear	0.000	1.585	0.126	0.204	0.312
9	10	Spline	0.000	1.736	0.137	0.221	0.338
10	25	IDW	0.000	1.825	0.131	0.222	0.349
11	25	Linear	0.000	1.360	0.105	0.170	0.262
12	25	Spline	0.000	1.496	0.112	0.184	0.284
13	50	IDW	0.000	1.436	0.108	0.179	0.277
14	50	Linear	0.000	1.273	0.093	0.153	0.238
15	50	Spline	0.000	1.379	0.099	0.165	0.257

Table 10: Testbed 3 statistics for pairwise interpolation methods comparison at various sampling densities.

Sampling Density (%)	Interpolation Methods	t-statistics	p-value
1	Spline and IDW	-305.2	0
	IDW and Linear	269.4	0
	Spline and Linear	-145.3	0
5	Spline and IDW	-205.6	0
	IDW and Linear	228.5	< 0.01
	Spline and Linear	20.6	0
10	Spline and IDW	-143.4	0
	IDW and Linear	190.1	0
	Spline and Linear	71.0	0
25	Spline and IDW	-82.0	0
	IDW and Linear	129.5	0
	Spline and Linear	73.5	0
50	Spline and IDW	-27.9	< 0.01
	IDW and Linear	69.9	0
	Spline and Linear	63.2	0

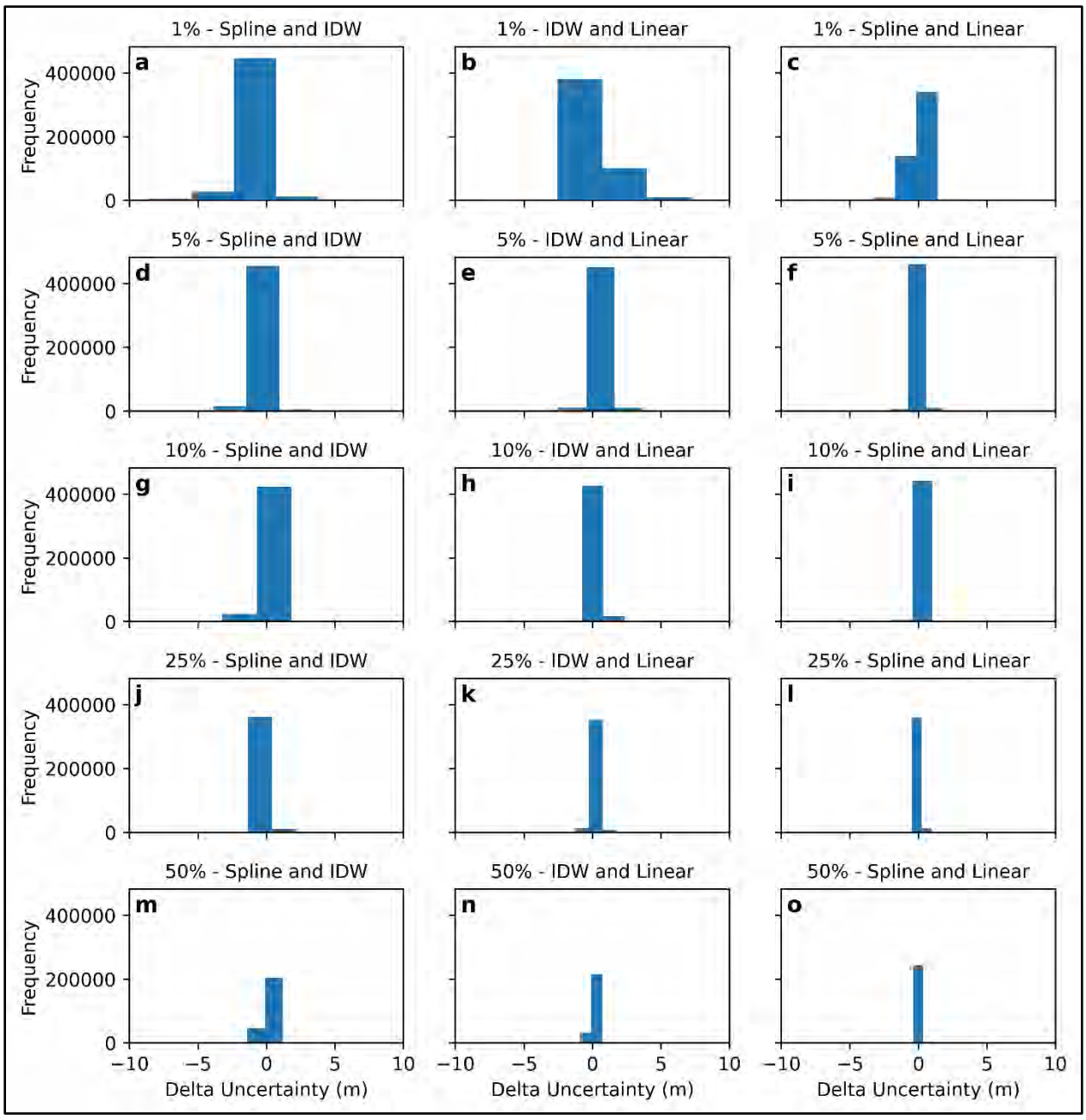

Figure 18: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 3.

4.5.3.2 Spatial Pattern of Interpolation Uncertainties

Examining Figure 19, the spatial distribution of uncertainties in Testbed 3 reveals a non-random pattern. Higher uncertainties are visually correlated with areas of high slope values (see Figure 2c

for comparison). This observation underscores the impact of slope characteristics on interpolation uncertainty.

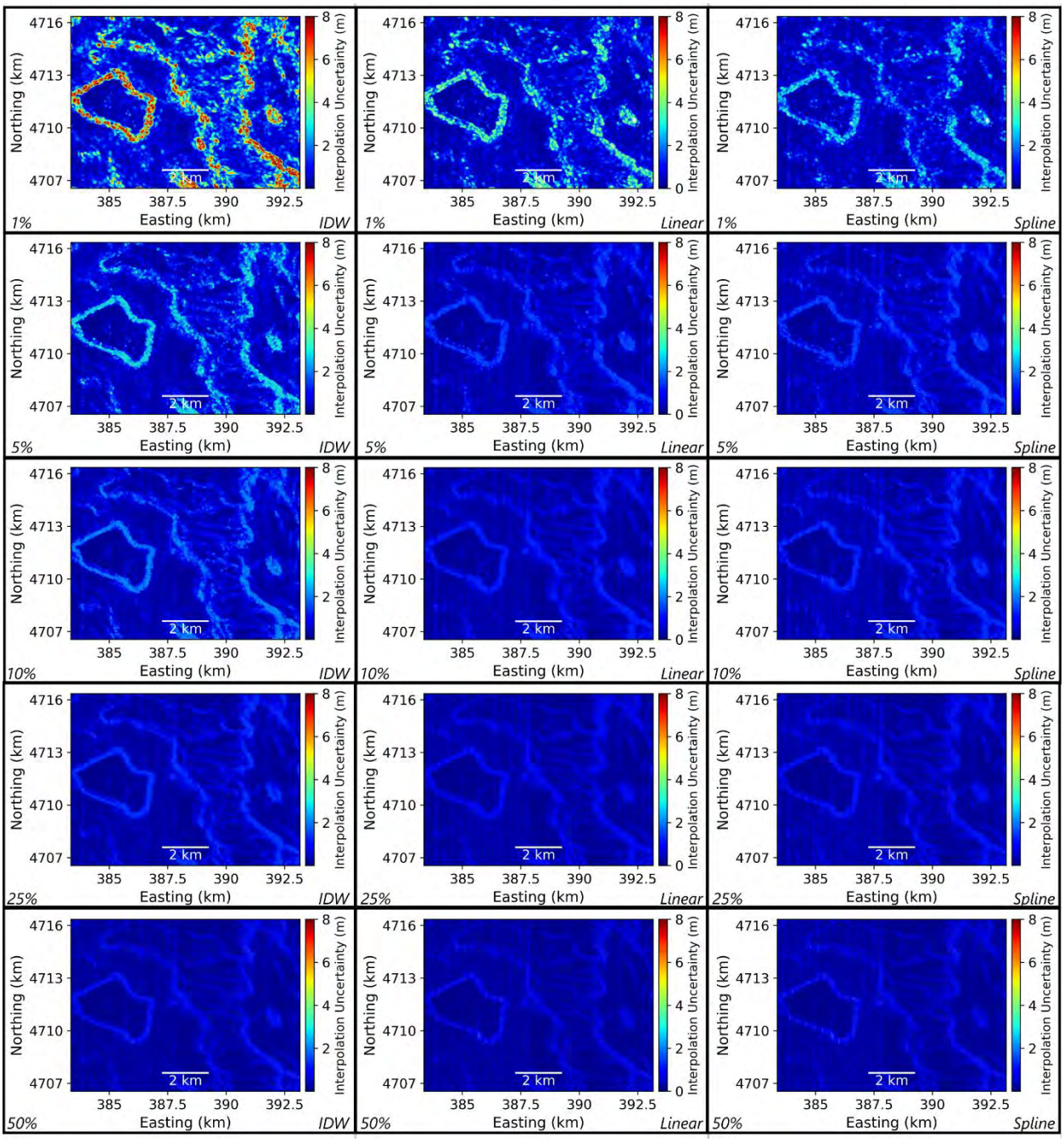

Figure 19: Testbed 3 interpolation uncertainty across all sampling densities (columns) and interpolation methods (rows), using 99th percentile of data.

4.5.3.3 Predictive Models of Interpolation Uncertainties

Figures 20a&d present the adjusted R^2 and RMSE values for estimated uncertainty based on the distance to the nearest measurement for Testbed 3. Surprisingly, a very weak relationship diminishes with increasing sampling density. Spline interpolation consistently outperforms Linear and IDW interpolation in both metrics.

Figures 20b&e (Testbed 3) present the adjusted R^2 and RMSE values for estimated uncertainty based on roughness. A moderate relationship intensifies with increased sampling density, showing an asymptotic trend at 25%. IDW achieves the highest adjusted R^2 but performs the worst in RMSE. Linear and Spline interpolations exhibit comparable performance levels.

Figures 20c&f (Testbed 3) depicts adjusted R^2 and RMSE values for estimated uncertainty based on slope. Similar to roughness, a moderate relationship intensifies with increasing sampling density, showing an asymptotic trend at 25%. In this case, IDW interpolation performs best in adjusted R^2, while Linear and Spline interpolations exhibit similar performance levels.

To investigate the hidden non-linear relationships, we employed an ANN, revealing marginal improvements in both R^2 and RMSE (Figure 21). Among interpolation techniques, IDW performs best in R^2 but the worst in RMSE. Linear and Spline interpolations exhibit similar performance levels.

Compared to Testbed 1 and 2, the distinctions in RMSE among the interpolation uncertainty models might be operationally important, given that they are on the order of meters. In furtherance with Testbed 1 and 2, the 10% sampling density appears to be the sweet spot for Testbed 3.

Results from the RF analysis coupled with the double bootstrap statistical technique reveal that roughness emerges as the most influential predictor ranking as the most important in all 15

instances, with a mean importance level of 59% (Table 8). Following in importance are slope and distance to the nearest measurement, with mean importance levels of 34% and 7%, respectively. Notably, distance to the nearest measurement consistently emerges as the least important factor across all 15 instances. The observed variations in the results are statistically significant, indicating robust differences among the predictors.

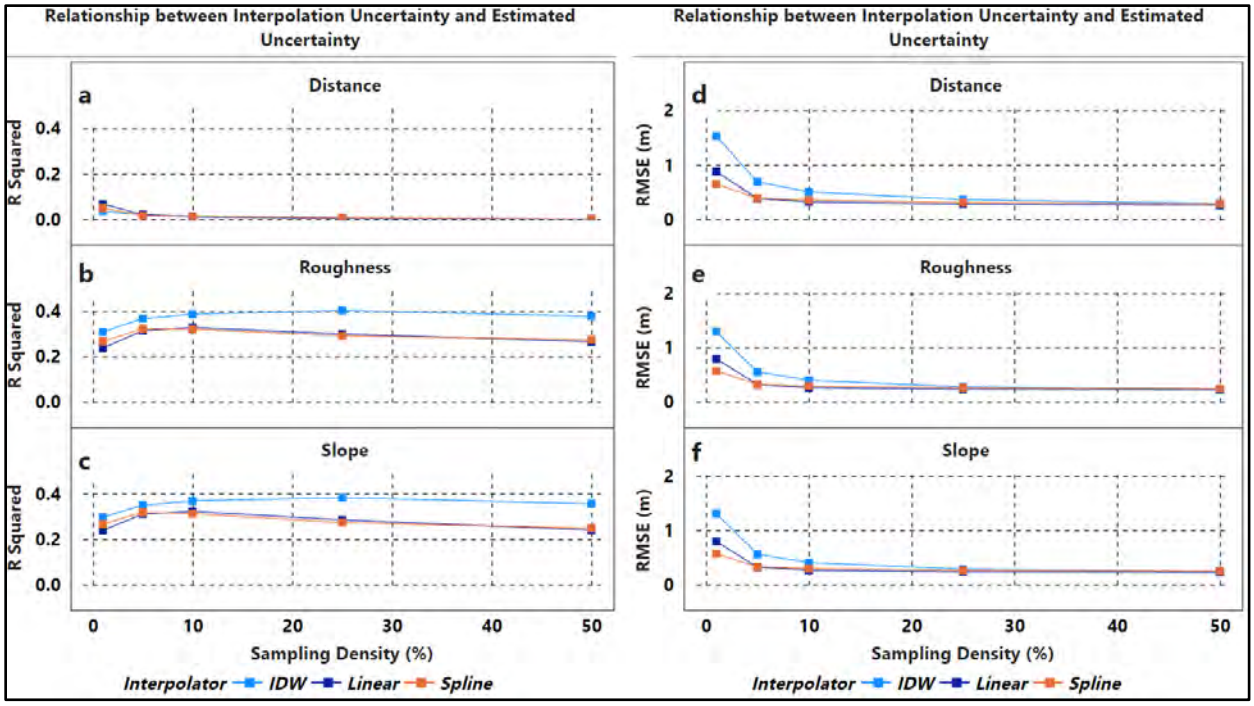

Figure 20: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 3.

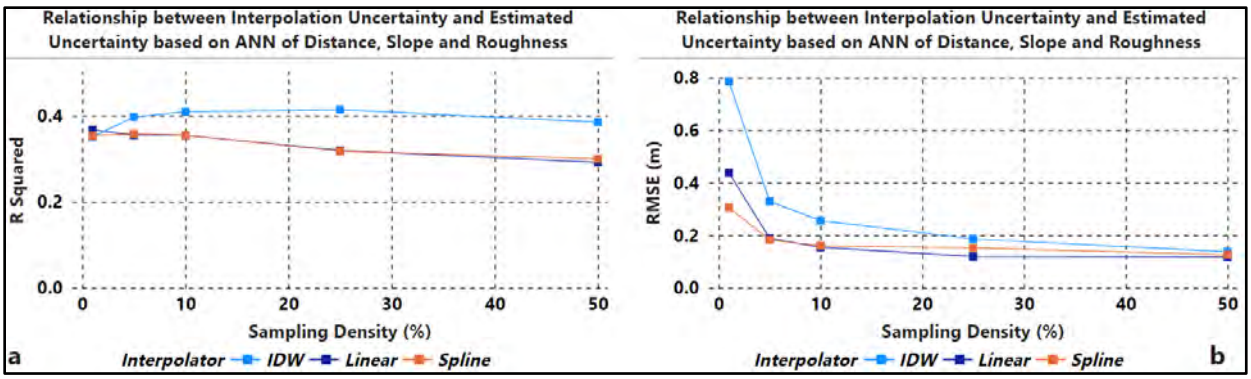

Figure 21: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 3.

Table 11: Statistics of the importance of predictors of uncertainty for Testbed 3.

Statistics/Parameter	Distance	Slope	Roughness
Mean Importance (%)	8.2	34.2	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)

4.5.4 Testbed 4 Results (Rough and Slopy Seabed)

4.5.4.1 Interpolation Methods

Figure 22 displays the box and whisker plots illustrating the performance of interpolation methods across all sampling densities for Testbed 4, complemented by the descriptive statistics presented in Table 12. The interpolation uncertainties of the interpolation methods vary at lower sampling densities with Linear as the best interpolation method followed by Spline and IDW. The interpolation uncertainties become relatively the same at higher sampling densities, 25% and 50%. At the higher sampling densities, Spline performs the best. Much like Testbed 3, the interpolation uncertainties here are in the order of meters, attributed to the relatively complex morphology of Testbed 4. Importantly, Figure 23 and Table 13 assert that the differences seen in the box and whisker plots and descriptive statistics are not statistically significant.

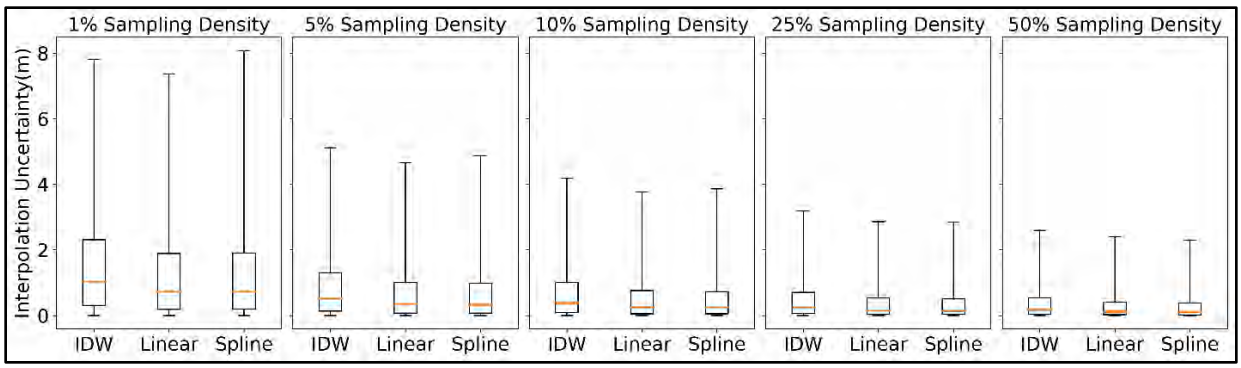

Figure 22: Testbed 4 interpolation methods uncertainty comparison at various sampling densities using box and whisker, plotted with 99% percentile of data.

Table 12: Testbed 4 interpolation uncertainty descriptive statistics at various sampling densities.

S/N	Sampling density (%)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error (m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	1	IDW	0.000	7.818	0.520	1.554	2.218
2	1	Linear	0.000	7.371	0.347	1.286	1.937
3	1	Spline	0.000	8.080	0.335	1.329	2.036
4	5	IDW	0.000	5.115	1.025	0.887	1.329
5	5	Linear	0.000	4.659	0.738	0.700	1.116
6	5	Spline	0.000	4.879	0.734	0.699	1.131
7	10	IDW	0.000	4.184	0.379	0.689	1.056
8	10	Linear	0.000	3.764	0.248	0.536	0.873
9	10	Spline	0.000	3.863	0.236	0.526	0.869
10	25	IDW	0.000	3.190	0.246	0.490	0.773
11	25	Linear	0.000	2.860	0.160	0.383	0.640
12	25	Spline	0.000	2.849	0.149	0.366	0.619
13	50	IDW	0.000	2.594	0.177	0.378	0.609
14	50	Linear	0.000	2.400	0.120	0.306	0.522
15	50	Spline	0.000	2.304	0.107	0.285	0.491

Table 13: Testbed 4 statistics for pairwise 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

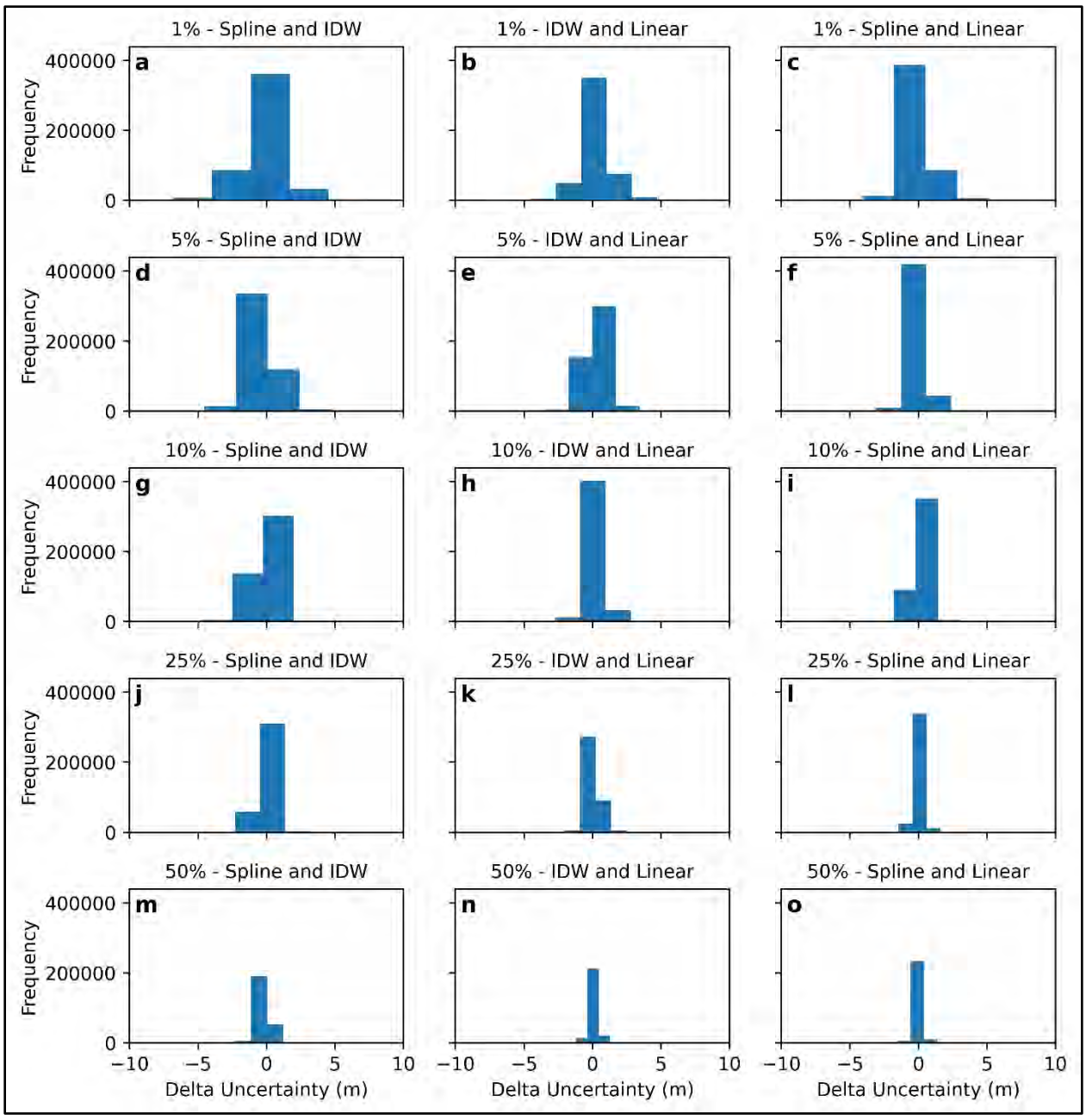

Figure 23: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 4.

4.5.4.2 Spatial Pattern of Interpolation Uncertainties

In Figure 24, the non-random pattern of uncertainties in Testbed 4, similar to other testbeds, is evident. Visual examination indicates that areas with higher uncertainties correlate with both high

slope and roughness values (see Figure 2d for comparison). This dual correlation emphasizes the combined influence of slope and roughness on interpolation uncertainties.

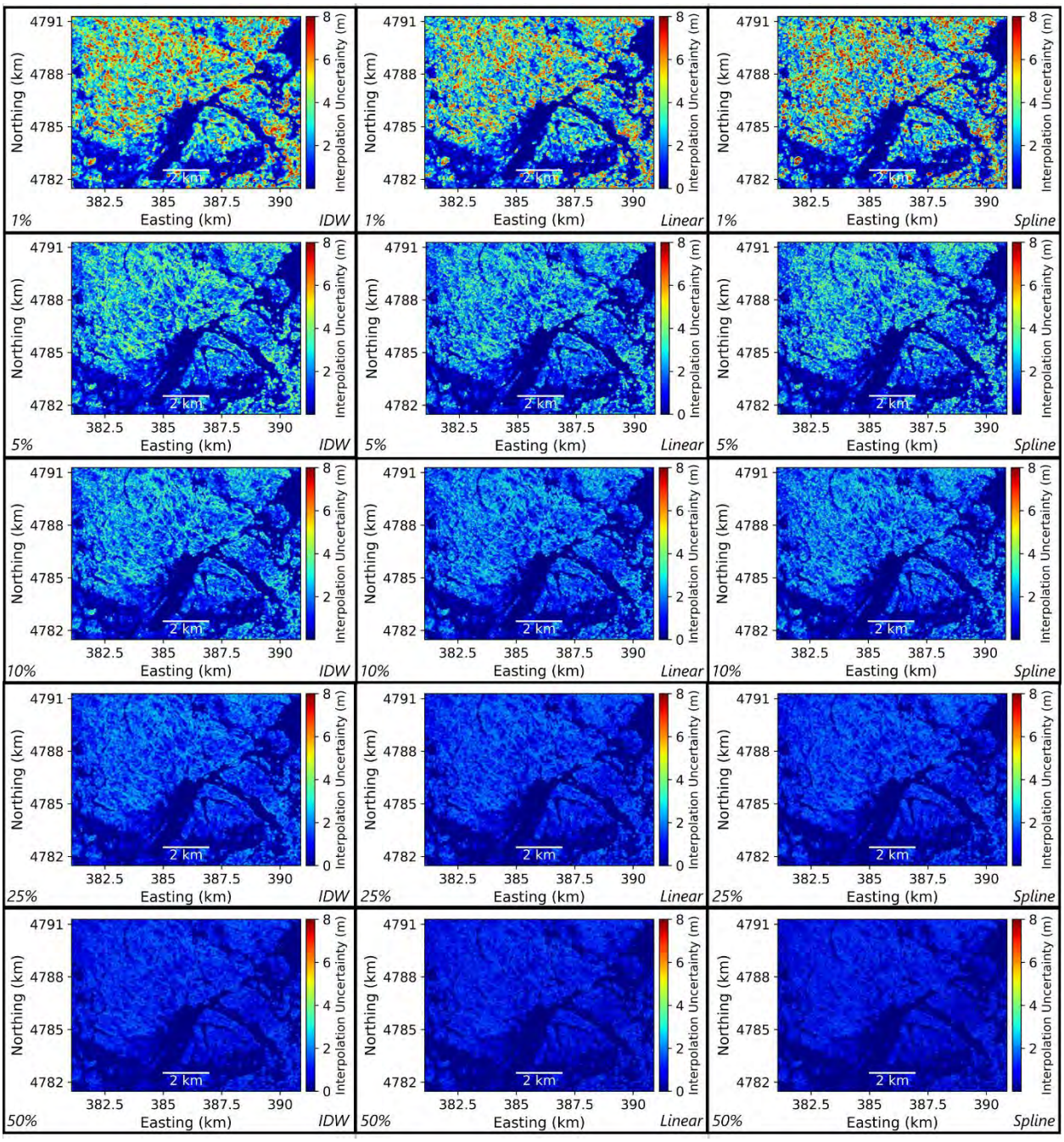

Figure 24: Testbed 4 interpolation uncertainty across all sampling densities (columns) and interpolation methods (rows), using 99th percentile of data.

4.5.4.3 Predictive Models of Interpolation Uncertainties

Figures 25a&d present adjusted R^2 and RMSE values for estimated uncertainty based on the distance to the nearest measurement for Testbed 4. The adjusted R^2 values indicate a weak relationship between the distance to the nearest measurement and interpolation uncertainty, which surprisingly diminishes as the sampling density increases. Additionally, the RMSE decreases with higher sampling density. Notably, Spline interpolation consistently outperforms Linear and IDW interpolation in both metrics

In Figures 25b&e (Testbed 4), the adjusted R^2 and RMSE values are presented for the estimated uncertainty based on roughness. A moderate relationship emerged, intensifying with increased sampling density. IDW interpolation achieves the highest adjusted R^2 but performs the worst in RMSE. Linear and Spline interpolations exhibit comparable performance levels.

Figures 25c&f (Testbed 4) depicts the adjusted R^2 and RMSE values for the estimated uncertainty based on slope. Similar to roughness, a moderate relationship intensifies with rising sampling density. In this case, IDW interpolation excels in adjusted R^2, but Linear interpolation emerges as the best performer in RMSE. Spline interpolation consistently lags behind in both adjusted R^2 and RMSE.

To investigate non-linear relationships, we employed an ANN, revealing marginal improvements in both R^2 and RMSE (Figure 26). Among the interpolation techniques, IDW remains superior in R^2, followed by Linear interpolation, while Spline interpolation consistently ranks lowest.

Similar to Testbed 3, the distinctions in RMSE among the interpolation uncertainty models might be operationally significant, given that they are on the order of meters. In furtherance with the preceding testbeds, the 10% sampling density appears to be the “sweet spot”.

Results from the Random Forest analysis coupled with the bootstrap statistical technique closely follow Testbed 3 with 57.4%, 33.1% and 9.5% mean importance of roughness, slope, and distance to nearest measurement, respectively (Table 14)

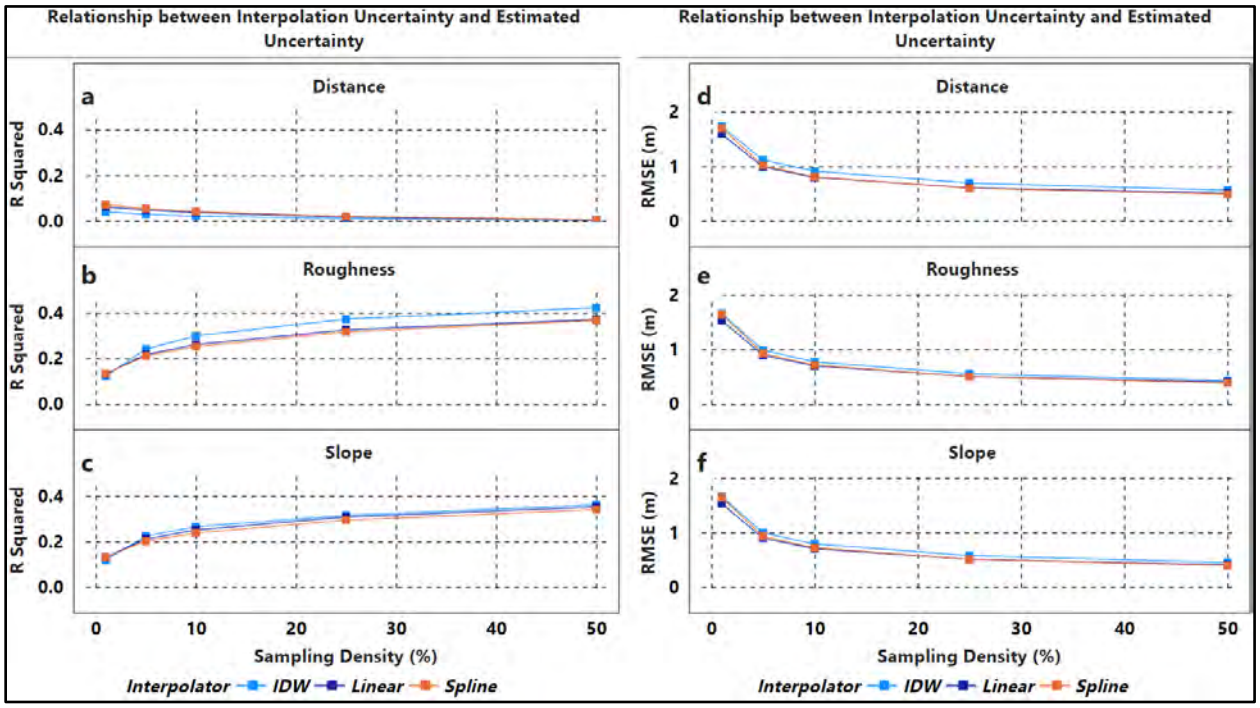

Figure 25: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 4.

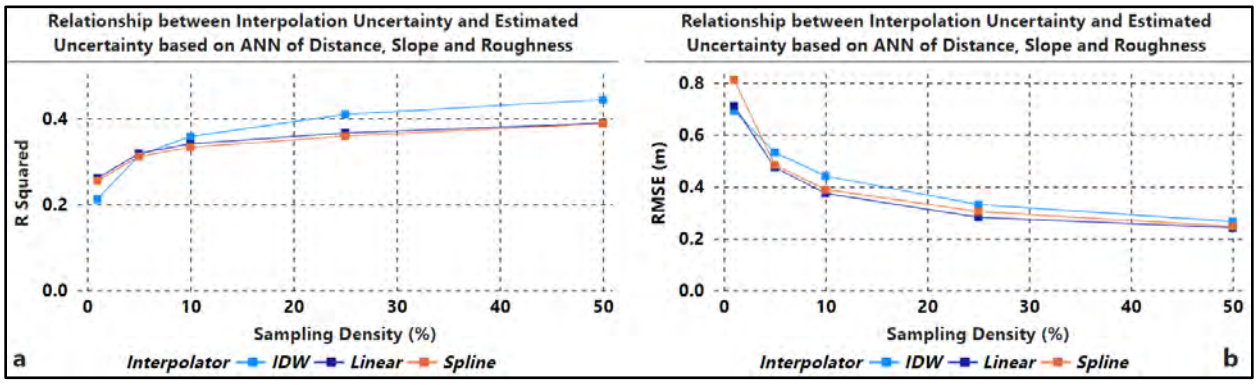

Figure 26: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 4.

Table 14: Statistics of the importance of predictors of uncertainty for Testbed 4.

Statistics/Parameter	Distance	Slope	Roughness
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)

4.5.5 Testbed 5 Results (Rough and Slopy Seabed with high spatial resolution)

Testbed 5 is a "special" testbed that was used to investigate the impact of data resolutions on interpolation uncertainty. This result is presented in the subsequent Spatial Scales section.

4.5.6 Spatial Scales

4.5.6.1 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. Testbed 4 was chosen because it had the highest R^2, offering the best potential to reveal the substantial impact of window size on uncertainty estimation. The R^2 of the ANN relationship between estimated uncertainty and combined parameters shown in Figure 27 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.

These findings suggest that a larger window is preferable for interpolating low sampling density data in Testbed 4. While this conclusion may not universally apply to other testbeds, it remains a noteworthy and insightful discovery.

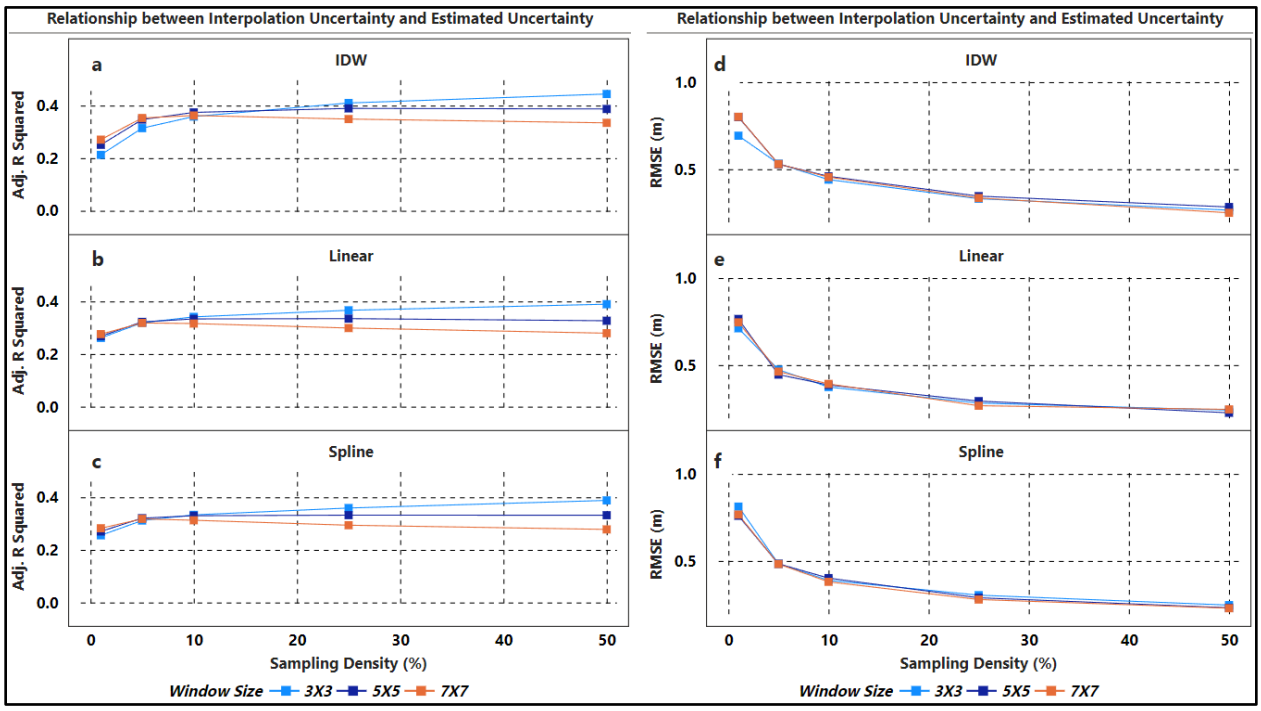

Figure 27: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on window sizes using Testbed 4.

4.5.6.2 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. The IDW error was the minimum compared to Linear and Spline interpolations, with Spline interpolation showing the highest error.

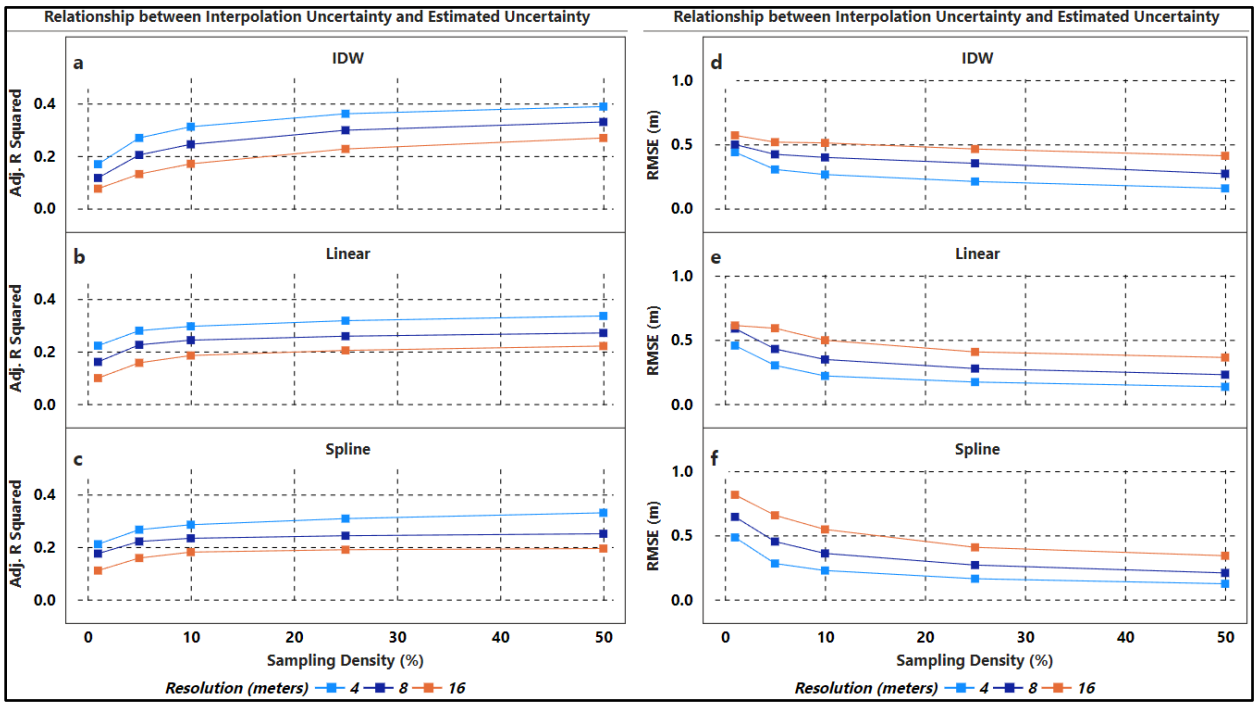

Figure 28: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on data resolution using Testbed 5.

4.6 Discussion

The results of our study offer valuable insights into the optimal interpolation method and complex interplay between interpolation uncertainty and a suite of ancillary parameters including distance to the nearest measurement, roughness, slope, using various sampling densities across diverse seabed testbeds. The identification of significant predictors and the evaluation of their impacts on interpolation uncertainty contribute valuable perspectives toward improving uncertainty estimation in interpolated bathymetry. The nuanced characterization of this uncertainty, achieved through an exploration of diverse seabed testbeds, data scarcity, data resolution, and spatial scales, has further advanced our comprehension of the subtleties inherent in estimating interpolation uncertainty. In the following sections, we delve into key findings, and avenues for further exploration.

4.6.1 Unraveling 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 best for Testbed 1 (Flat), while Spline proved best 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 (*et*, differences of a few centimeters, 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 centimeters due to their relatively less complex morphologies. In contrast, Testbeds 3 and 4, with more complex morphologies, showed 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. This insight suggests that it is best to interpolate over simpler morphologies, as the magnitude of uncertainty incurred from the interpolation process can reach up to 8 meters or more in complex morphologies (refer to Figure 24). Nevertheless, since interpolation is not limited to simpler morphologies in practice, this implies that while sampling density can be reduced over simple morphologies to save cost and time, it must be increased over complex morphologies.

Expectedly, 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.

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 they were 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.

4.6.2 Relationship among 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 testbeds, 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. 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, signaling 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 when there is a limited budget for hydrographic surveys. 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. This is addressed in the next chapter.

4.6.3 Examining Disparities in Testbed Predictive Performance of Interpolation Methods

The significant variations in predictive performance observed across testbeds 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 comparable 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, affirming 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 the 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.

4.6.4 Importance of Predictors

An in-depth analysis of RF results, 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 testbeds, 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 testbeds 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.

4.6.5 Impact of Window Size on Interpolation Uncertainty

Spatial scale becomes especially important in the context of sampling density. The performance of different window sizes varies, contingent upon both the density of measurement points and the unique characteristics of the seabed. The observed trends in both R^2 and RMSE underscore the sensitivity of interpolation uncertainty to spatial scales, emphasizing the need for careful consideration when selecting window sizes.

The trend of the 3-by-3 window, particularly excelling at higher sampling densities, hints at the nuanced presence of smaller-scale seabed features in Testbed 4. This nuanced observation implies that the efficacy of smaller window sizes is not universally applicable; rather, it hinges on the interplay between seabed morphology and sampling density.

Concurrently, the performance of the 7-by-7 window, excelling at smaller sampling densities but faltering at higher densities, suggests its adeptness at capturing and representing features on Testbed 4 effectively with fewer measurements. However, this effectiveness diminishes as the sampling density increases, indicative of the larger window size incorporating more variability and experiencing a subsequent decline in performance.

This dynamic interrelation among window size, sampling density, and seabed complexity underscores the need to carefully determine an optimal window size, considering the multifaceted factors at play. While larger windows excel at capturing broader trends in sparser data scenarios, their effectiveness diminishes in denser data environments where smaller-scale features and variations take precedence.

Understanding this intricate relationship facilitates the selection of an appropriate spatial scale tailored to specific sampling conditions. This nuanced comprehension significantly enhances our overarching understanding of characterizing uncertainty in interpolated bathymetry.

4.6.6 Impact of Data Resolutions on Interpolation Uncertainty

The examination of the impact of data resolutions on interpolation uncertainty, illustrated in Figure 28, underscores the significance of higher data resolutions in improving uncertainty estimation and providing a more detailed depiction of the seabed. In our exploration, we employed Testbed 5 as the experimental ground, featuring data resolutions of 4m, 8m, and 16m, providing a comprehensive understanding of the dynamics involved. This insight adds depth to the understanding of the complexities involved in characterizing uncertainty in bathymetric interpolations.

4.7 Summary

This analysis identifies Spline as the interpolation method that produced the least interpolation uncertainty among those examined for Testbeds 2 (Rough), 3 (Slopy), and 4 (Rough and Slopy), followed by Linear and IDW from a scientific perspective. For Testbed 1 (Flat), Linear interpolation produced the least interpolation uncertainty, followed by Spline and IDW. However, the differences in the performances of the interpolation methods are not substantial from an operational standpoint. Additionally, the analysis underscores the importance of roughness, slope, and distance as predictors of interpolation uncertainty across various cell sampling densities and diverse testbeds. It is crucial to note, however, that these predictors of uncertainty individually exhibit a weak relationship with interpolation uncertainty and the relationship slightly improved when predictors were combined into a single ANN model. The performance of uncertainty estimation varies significantly across the testbeds, with Testbed 4 demonstrating the best performance with the highest adjusted R^2 and lowest RMSE, followed by Testbed 3, Testbed 2, and Testbed 1. While the predictive accuracy, based on adjusted R^2, of the uncertainty model generated through IDW interpolation slightly outperformed those produced by Spline and Linear interpolation methods across the majority of testbeds, the RMSE of the IDW uncertainty model was the least favorable. Notably, the impact of cell sampling density on the model's explanatory power is evident, diminishing effectiveness with smaller sampling densities and increasing distances to the nearest measurement point. Furthermore, the result of the effects of spatial scales showed that interpolation uncertainty can be better estimated with a higher data resolution (smaller pixel size) and not necessarily at lower analysis window sizes.

4.8 Conclusion

This study aimed to identify the best bathymetric gap-filling interpolation method – that minimizes interpolation uncertainty – and accurately quantify and characterize uncertainty in interpolated bathymetry within operational settings. It sought to establish the relationship between interpolation uncertainty and a suite of ancillary parameters – distance to the nearest measurement, slope, and roughness – across five testbeds in the United States, using various sampling densities. The impact of seabed morphology, data paucity, and spatial scales on uncertainty estimation was also explored.

The findings reveal that Spline is the best interpolation method for Testbeds 2, 3 and 4, followed by Linear and IDW. For Testbed 1, Linear interpolation resulted in the least interpolation uncertainty, followed by Spline and IDW. The ancillary parameters exhibited a weak individual and combined relationship with interpolation uncertainty. Notably, roughness emerged as the most important predictor, followed by slope and distance to the nearest measurement across the testbeds. Additionally, IDW provided 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.

Moreover, the precision 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 substantially 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.

CHAPTER 5 : ESTIMATING AND CHARACTERIZING INTERPOLATION UNCERTAINTY IN SET-LINE SPACING HYDROGRAPHIC SURVEYS

5.1 Introduction

This chapter investigates the estimation and characterization of interpolation uncertainty in set-line spacing hydrographic surveys executed with modern survey techniques. Modern hydrographic surveys are conducted systematically, guided by survey methodology, and are fundamentally distinct from the random sampling approach investigated in the preceding chapter. Set-line spacing surveys are used to survey quickly because full seafloor bathymetric surveys are expensive and time-consuming, especially in shallow water. Set-line spacing bathymetric surveys when performed in tandem with Side Scan Sonar, popularly known as Skunk Stripping in the US as explained in the introduction section have numerous benefits without compromising the safety of navigation. One of these advantages is that they survey a relatively large area in less time and at a reduced cost compared to full seabed coverage bathymetric surveys. They are also efficient in shallow water and hazardous environments. It is paramount to also note that the set-line spacing approach is commonly used in areas with relatively flat seafloor, low navigational risk, and existing data from satellite-derived/lidar bathymetry where bottom detection was not achieved due to water clarity or extinction depth (Neff and Wilson 2018, NOAA OCS 2021). However, it leaves us with bathymetric coverage gaps that can span meters to kilometers between survey measurements and thus calls for a need to interpolate and more importantly, necessitates accurately quantifying the uncertainty in the interpolated regions to inform users, especially mariners, of the confidence in the data. This uncertainty estimation will help to improve hydrographic offices' CATZOC classification of surveys and is not limited to CATZOC B as full seabed coverage and feature detection requirements for CATZOC A1 and A2 would be met using SSS. Additionally, characterizing interpolation uncertainty can help to optimize set-line spacing hydrographic survey

designs to meet the desired uncertainty of the grid within the constraints of existing resources and their impact on field data collection efforts.

5.2 Methods

The methodology utilized herein is the simulation of set-line spacing surveys from the testbeds' complete coverage depth measurement (Figure 29) and investigation of uncertainty estimation associated with interpolation within this known area. The main line spacings used are 16m, 32m, 64m, 128m, 256m, and 512m. These line spacings are chosen thoughtfully to understand how interpolation uncertainty will behave in relatively small and wide line spacings. Since main lines are usually supplemented by crosslines to verify and evaluate the internal consistency of hydrographic surveys, the set-line spacing surveys' crossline spacings were calculated using 9% of main line mileage according to the (NOAA OCS 2021) and distributed geographically across the Testbeds (Figure 29). The set-line survey was simulated using these main line spacings and their appropriate cross line spacings.

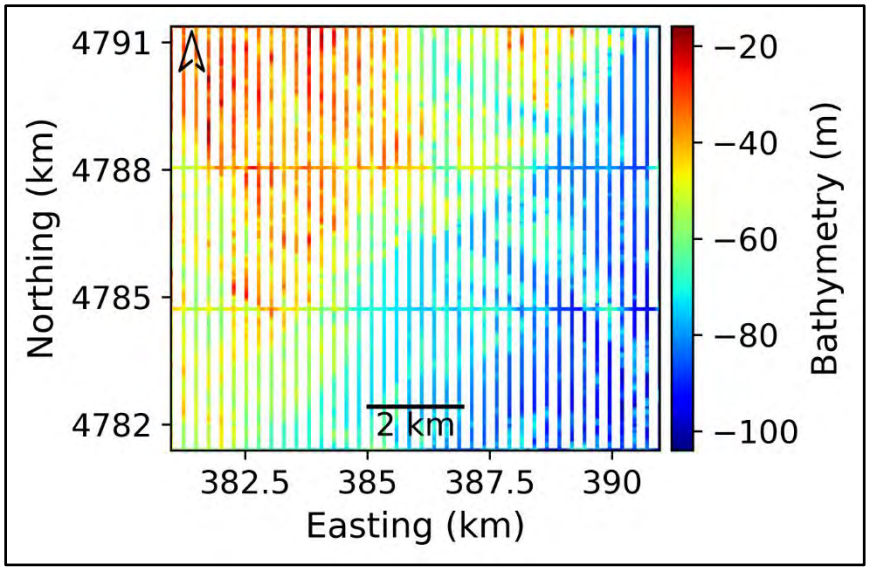

Figure 29: 256-meters main lines spacings and equivalent cross lines spacings training depths on Testbed 4.

The testbeds' selected main lines and cross lines depths were the training depths, earmarked for interpolation, and the unselected depths which are the test depths, earmarked for uncertainty quantification and error analysis (refer to Figure 30).

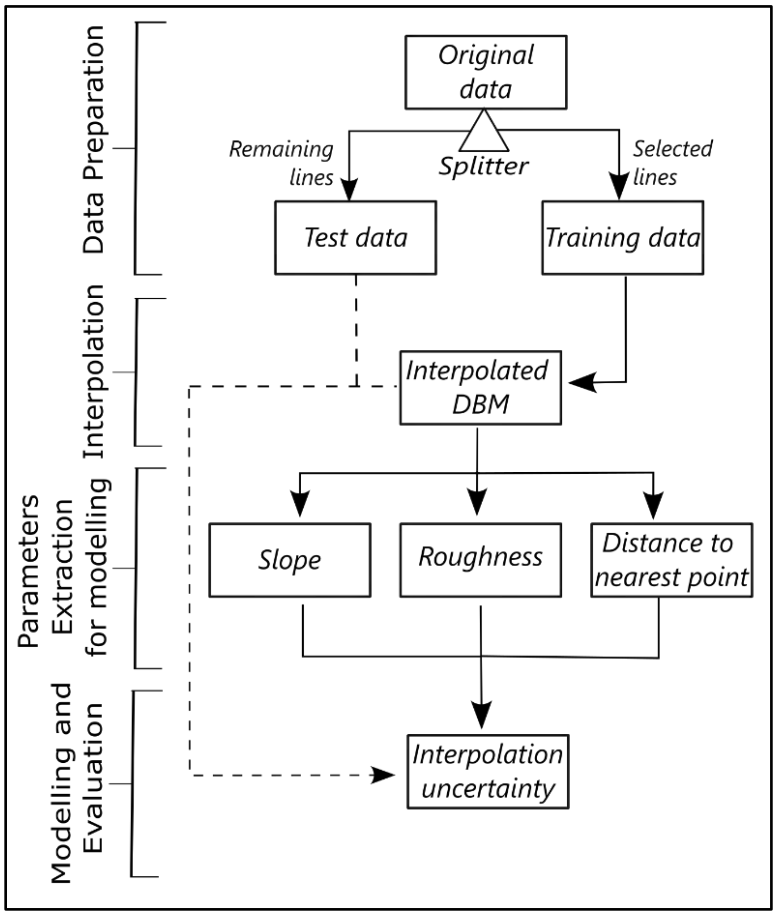

Figure 30: Workflow for quantifying interpolation uncertainty in set-line spacing surveys.

The training depths were interpolated using the specified interpolation technique. The deterministic interpolation techniques employed are IDW, Spline, and Linear. The resulting interpolated raster was then compared, on a cell-by-cell basis, to the test depths to quantify interpolation uncertainty. To clarify, interpolation uncertainty was determined by taking the absolute value of the difference between the interpolated depths and the measured depths for the cells that were not within one of the set lines. Ancillary parameters, such as the Euclidean distance

to the nearest measurement, were generated from the training depths; the slope and roughness raster surfaces, were generated from the interpolated DBM (refer to Figure 31). These parameters, along with their respective interpolation uncertainties, constituted the datasets prepared for modeling to discern the relationship between the parameters and interpolation uncertainties. It should be noted that there were no iterations in this procedure compared to the random sampling methodology in chapter 4, because there is no randomness associated with set-line spacing surveys *i.e.,* set-line spacing surveys are systematic sampling or survey techniques.

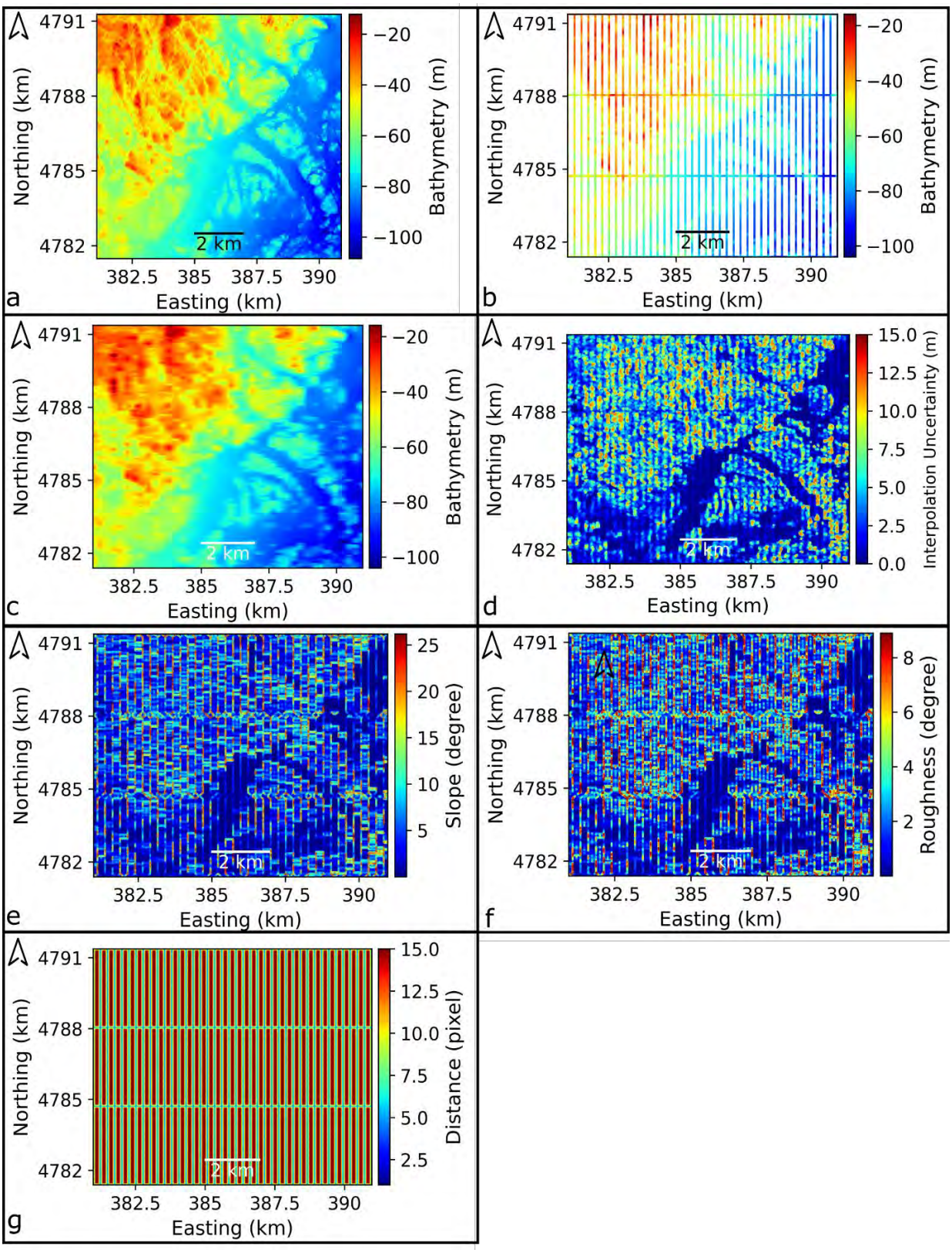

Figure 31: Testbed 4 measured depth (a), 256m main lines spacings and equivalent cross lines spacings training depths (b), Linear interpolated depths (c), interpolation uncertainty clipped to 99

percentile (d), generated slope from interpolated depths (e), generated roughness from interpolated depths (f), and distance to the nearest depth measurement raster (g).

To mitigate bias in the analysis, depth measurements along the border of study areas were part of the data used for interpolation. This approach minimized potential edge effects by ensuring that the interpolated values near the study area's boundaries were influenced by actual measured depths, reducing potential inaccuracies that could arise from extrapolation.

5.3 Modeling and Analysis

The approach used in Chapter 4's Modeling and Analysis section is replicated here. Below is a summary of the steps:

- Linear regression modeling of the relationship between interpolation uncertainty and individual parameters.
- ANN implementation to capture the non-linearities, interactions, and hidden relationships within combined parameters.
- RF's feature importance together with bootstrap statistical technique to determine the importance of individual parameters and their significance.

5.4 Validation Technique/Accuracy Assessment

The Validation Technique/Accuracy Assessment section (Chapter 4) is replicated with the following summarized steps:

- Split-sample method to divide the dataset into training and test subsets, using the former for interpolation and evaluating performance with the latter.
- Utilize common statistical metrics (RMSE, MAE, R^2) to evaluate the accuracy of the interpolation methods.
- Conduct pair-wise t-tests to statistically assess differences in interpolation methods.
- Spatial evaluation of interpolation uncertainty through visual inspection.

5.5 Results

The presentation of results is organized based on testbeds, outlined in the following subsections.

5.5.1 Testbed 1 (Flat Seabed)

5.5.1.1 Interpolation Methods

Figure 32 displays the box and whisker plots illustrating the performance of interpolation methods across all mainline spacings (16m, 32m, 64m, 128m, 256m, 512m) for Testbed 1. Complementing these visualizations, Table 15 provides descriptive statistics of interpolation uncertainties at the 99th percentile confidence interval for each line spacing and interpolation method. Notably, the uncertainties of the different interpolation methods are relatively the same at the same mainline spacing with Linear interpolation performing slightly better than IDW and Spline. It is highlighted that the interpolation uncertainties here are in the order of centimeters, attributed to the simpler morphology of Testbed 1. This simpler morphology is also a contributing factor to the Linear interpolation method exhibiting the lowest uncertainty compared to the other two methods. Table 16 and Figure 33 show that there are no statistically significant differences in the interpolation methods at each of the mainline spacings. This is indicated by the t-test result and the distributions of each pair-wise interpolation uncertainty difference centered around zero.

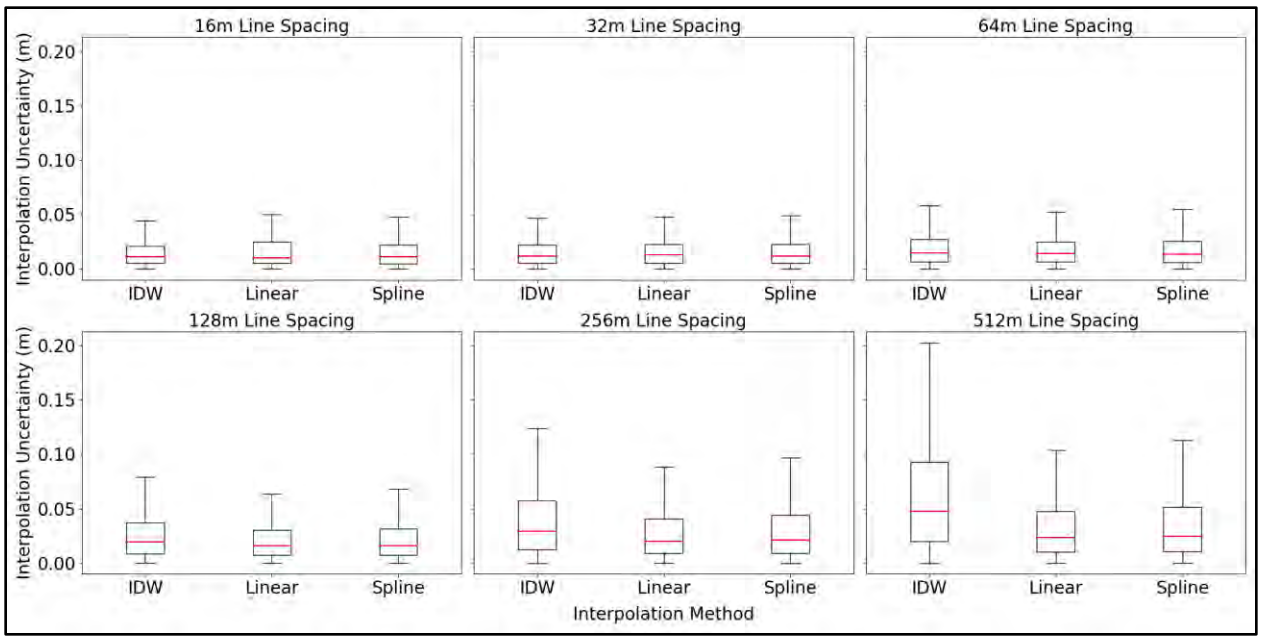

Figure 32: Testbed 1 interpolation methods uncertainty comparison at various line spacings using box and whisker, plotted with 99% percentile of data.

Table 15: Testbed 1 interpolation uncertainty descriptive statistics at various line spacings.

S/N	Line Spacing (m)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error (m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	16	IDW	0.000	0.070	0.011	0.017	0.020
2	16	Linear	0.000	0.080	0.010	0.015	0.022
3	16	Spline	0.000	0.090	0.011	0.016	0.022
4	32	IDW	0.000	0.080	0.012	0.017	0.021
5	32	Linear	0.000	0.090	0.012	0.016	0.023
6	32	Spline	0.000	0.090	0.012	0.016	0.023
7	64	IDW	0.000	0.090	0.015	0.018	0.026
8	64	Linear	0.000	0.100	0.014	0.019	0.025
9	64	Spline	0.000	0.100	0.013	0.018	0.025
10	128	IDW	0.000	0.140	0.019	0.022	0.036
11	128	Linear	0.000	0.120	0.016	0.027	0.031
12	128	Spline	0.000	0.120	0.016	0.023	0.032
13	256	IDW	0.000	0.220	0.029	0.030	0.057
14	256	Linear	0.000	0.160	0.020	0.041	0.043
15	256	Spline	0.000	0.180	0.021	0.033	0.047
16	512	IDW	0.000	0.350	0.047	0.035	0.091
17	512	Linear	0.000	0.190	0.023	0.066	0.050
18	512	Spline	0.000	0.230	0.024	0.039	0.056

Table 16: Testbed 1 statistics for pairwise interpolation methods comparison at various line spacings.

Line Spacing (m)	Interpolation Methods	t-statistics	p-value
16	Spline and IDW	96	0
	IDW and Linear	-9.8	0
	Spline and Linear	220.3	0
32	Spline and IDW	82.4	0
	IDW and Linear	-0.8	0
	Spline and Linear	209.2	0
64	Spline and IDW	-64	0
	IDW and Linear	128.8	0
	Spline and Linear	137.2	0
128	Spline and IDW	-184.7	0
	IDW and Linear	238.8	0
	Spline and Linear	103.9	0
256	Spline and IDW	-260.7	0
	IDW and Linear	362.8	0
	Spline and Linear	266.9	0
512	Spline and IDW	-536.4	0
	IDW and Linear	612.5	0
	Spline and Linear	212.3	0

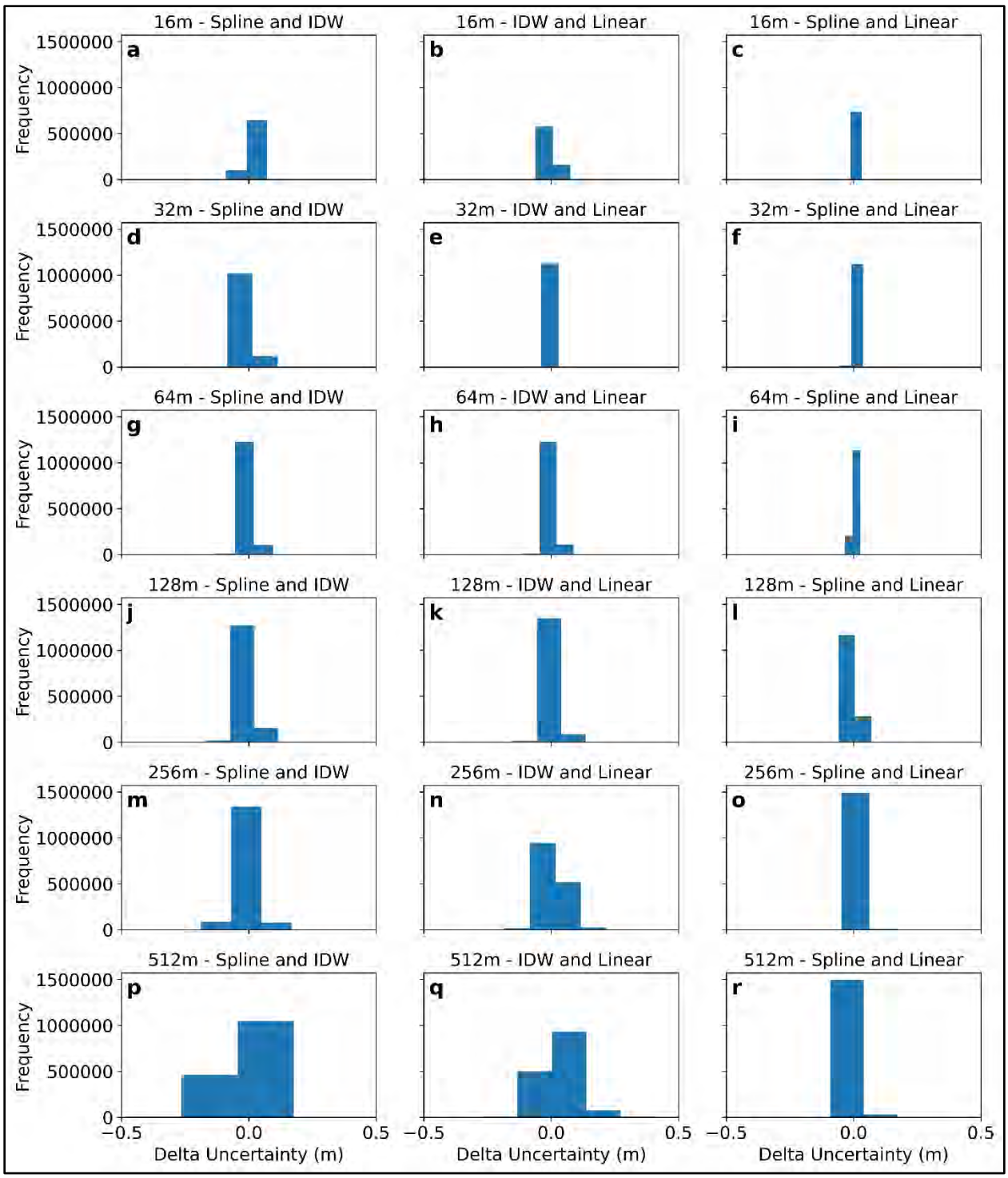

Figure 33: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 1.

5.5.1.2 Spatial Pattern of Interpolation Uncertainties

Figure 34 illustrates the spatial distribution of interpolation uncertainties of the interpolation methods across all line spacings for Testbed 1. To facilitate comparison, the color bar has been standardized across line spacings and interpolation methods. An intriguing observation is that the interpolation uncertainties do not exhibit a random pattern across the testbed with all line spacings. Instead, they are notably concentrated on the eastern side of the plots, attributed to multibeam artifacts (strips) present in the original datasets from BlueTopo. Furthermore, the interpolation uncertainties are pronounced in areas farthest from the set-lines sampled, *i.e.*, the wider the line spacing, the higher the interpolation uncertainty.

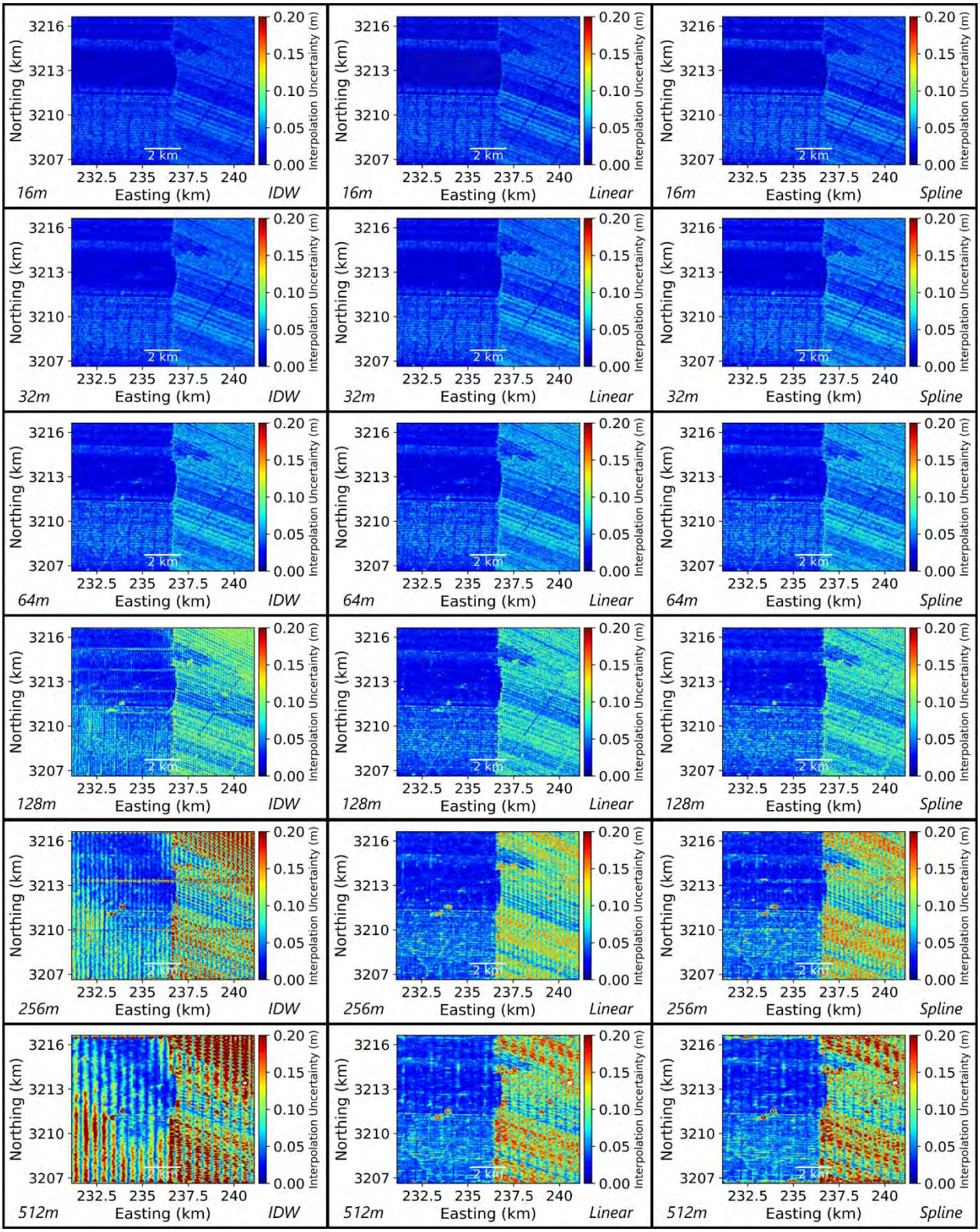

Figure 34: Testbed 1 interpolation uncertainty across all line spacings (columns) and interpolation methods (rows), using 99th percentile of data.

5.5.1.3 Predictive Models of Interpolation Uncertainties

Figures 35a & 35d present the adjusted R^2 and RMSE values for the estimated uncertainty for Testbed 1, assessed in relation to the distance to the nearest measurement across the line spacings (16m, 32m, 64m, 128m, 256m, 512m). The adjusted R^2 values indicate a weak relationship between the distance to the nearest measurement and interpolation uncertainty, with IDW outperforming the two other interpolation methods. Additionally, wider line spacings correspond to increased RMSE, with Linear interpolation outperforming IDW and Spline.

Figures 35b & 35e depict adjusted R^2 and RMSE values for the estimated uncertainty based on roughness for Testbed 1. A weak relationship diminishes with increased line spacing. Spline outperformed Linear and IDW. Similar to the distance to the nearest measurement, wider line spacings correspond to increased RMSE, with Linear interpolation outperforming Spline and IDW.

Figures 35c & 35f present the adjusted R^2 and RMSE values for the estimated uncertainty based on slope for Testbed 1. A relatively same weak relationship was observed among the interpolation methods with increased line spacing, with Spline outperforming Linear and IDW. Similar to the distance to the nearest measurement and roughness, wider line spacings correspond to increased RMSE, with Linear interpolation outperforming Spline and IDW.

To investigate the hidden non-linear relationships between the combined parameters and the estimated uncertainty, we employed an Artificial Neural Network (ANN). Results indicate that combining distance to the nearest measurement, slope, and roughness in a model improves the model's performance, albeit with marginal enhancements in both R^2 and RMSE (Figure 36). Among the interpolation techniques, IDW performs best in R^2 but performs the worst in RMSE which is corroborated by findings in Figure 35. Linear interpolation follows Spline in adjusted R^2 and outperforms it in RMSE.

In general, the differences in the interpolation uncertainty models' RMSE are not noteworthy from an operational perspective because they are in the order of centimeters. Surprisingly, based on R^2, the performance of the uncertainty models improves at wider line spacings, as evidenced in the combined parameters' interpolation uncertainty models.

Furthermore, the application of the RF machine learning algorithm alongside the bootstrap statistical technique reveals that slope is the most influential predictor ranking as the most important in 13 out of 18 instances, with a mean importance level of 52.9% (Table 17). These instances stem from the combination of three interpolation methods and six line spacings. Followed by slope is roughness, identified as the most important in 5 out of 18 instances, particularly dominant with Spline across line spacings except 16m, with a mean importance of 38.7%. Conversely, distance to the nearest measurement is the least important predictor, not appearing as the most important in any instance and holding a mean importance of 8.4%. Crucially, the bootstrap technique confirms the statistically significant differences in these parameters.

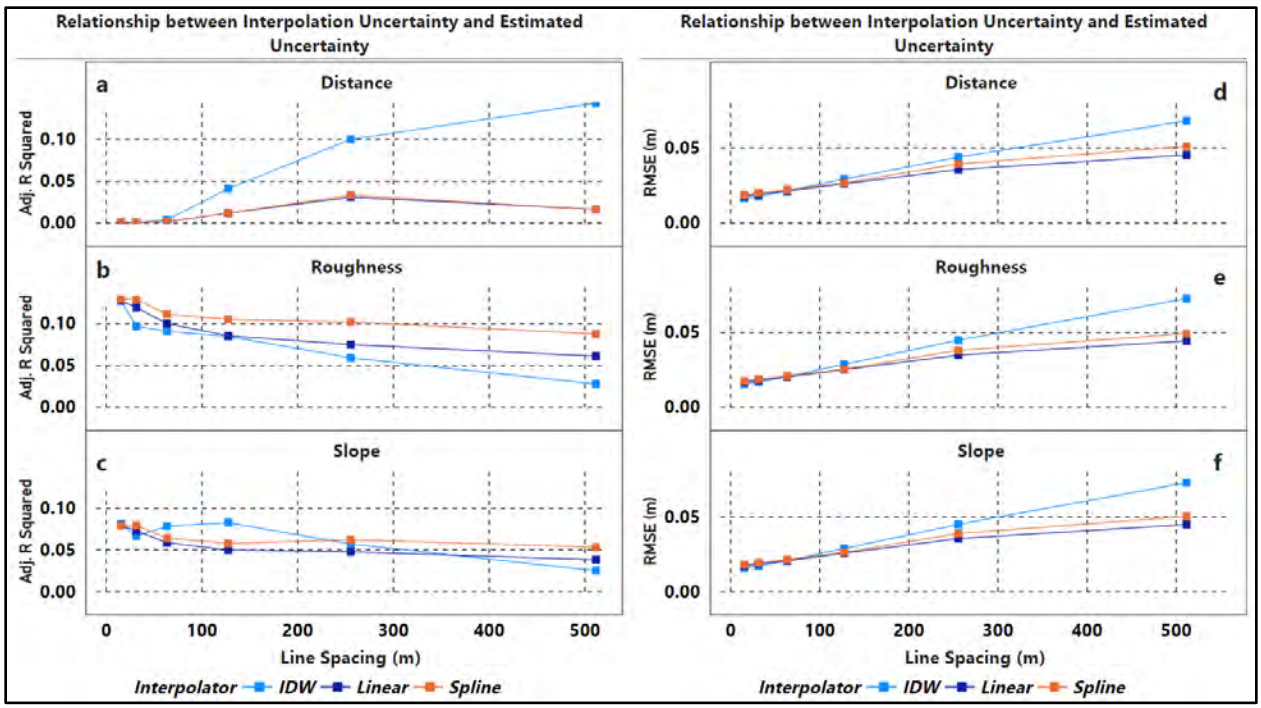

Figure 35: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 1.

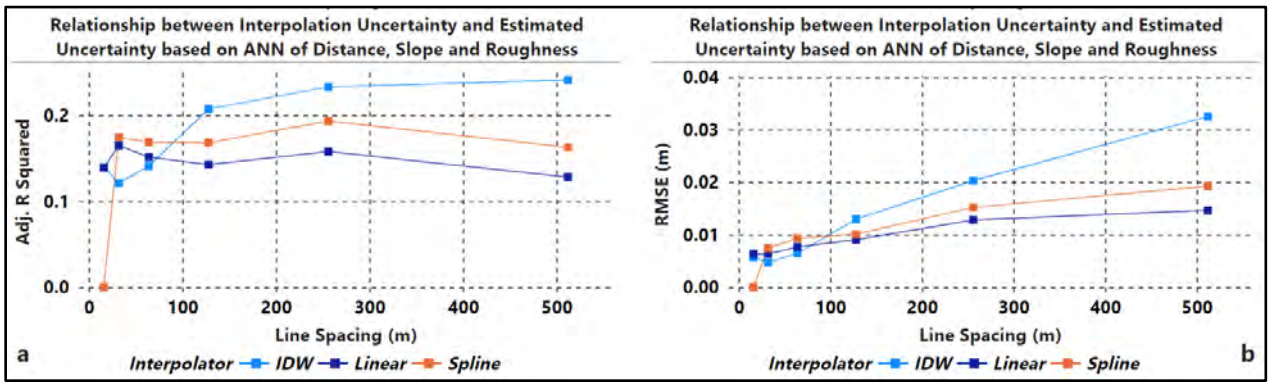

Figure 36: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 1.

Table 17: Statistics of the importance of predictors of uncertainty for Testbed 1

Statistics/Parameter	Distance	Slope	Roughness
Mean Importance (%)	8.4	52.8	38.7
# times of most Important	0	13	5
95% bootstrap percentile CI of the Importance	(8.0,8.8)	(51.4,54.3)	(37.3,40.1)

5.5.2 Testbed 2 (Rough Seabed)

5.5.2.1 Interpolation Methods

Figure 37 displays the box and whisker plots illustrating the performance of interpolation methods across all mainline spacings for Testbed 2. Similar to Testbed 1, the interpolation uncertainties of the interpolation methods are relatively the same across the same line spacing. However, on Testbed 2, the Spline interpolation performs slightly better than Linear and IDW as shown in Table 18. Additionally, the interpolation uncertainties here are in the order of centimeters, attributed to the relatively simple morphology of Testbed 2. Table 19 and Figure 38 also show that there are no statistically significant differences in the interpolation methods at each of the line spacings.

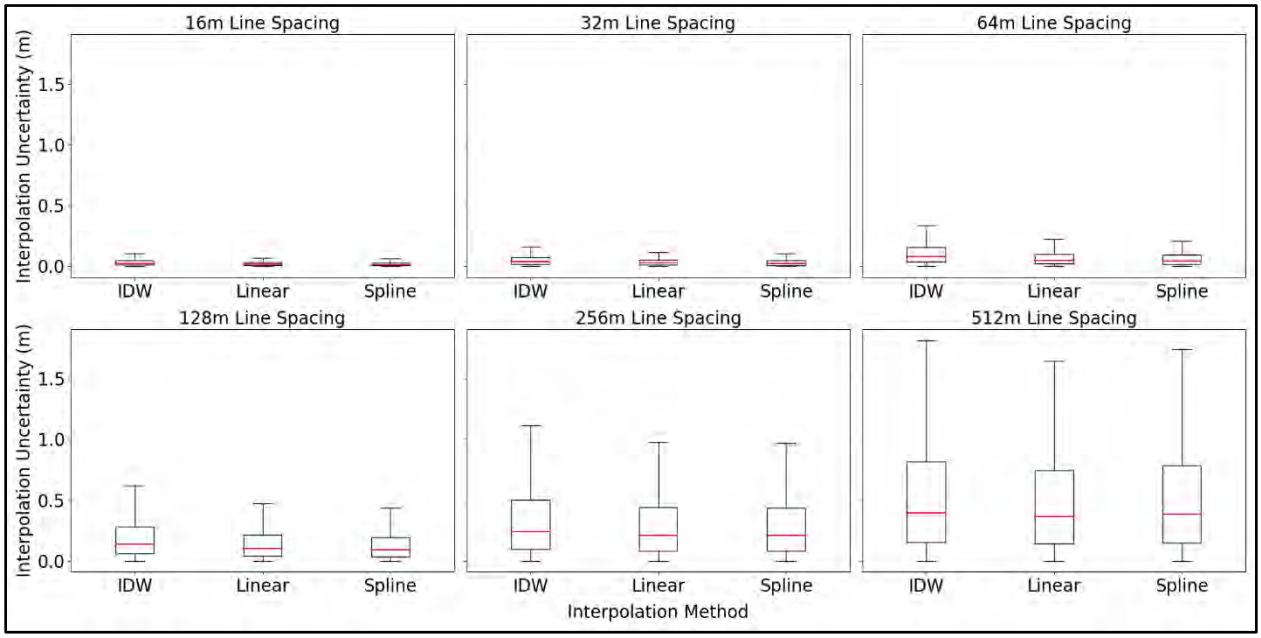

Figure 37: Testbed 2 interpolation methods uncertainty comparison at various line spacings using box and whisker, plotted with 99% percentile of data.

Table 18: Testbed 2 interpolation uncertainty descriptive statistics at various line spacings.

S/N	Line Spacing (m)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error(m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	16	IDW	0.000	0.185	0.024	0.024	0.049
2	16	Linear	0.000	0.150	0.015	0.035	0.035
3	16	Spline	0.000	0.140	0.013	0.022	0.033
4	32	IDW	0.000	0.275	0.037	0.039	0.073
5	32	Linear	0.000	0.237	0.025	0.053	0.057
6	32	Spline	0.000	0.223	0.021	0.035	0.052
7	64	IDW	0.000	0.560	0.079	0.074	0.152
8	64	Linear	0.000	0.423	0.048	0.111	0.106
9	64	Spline	0.000	0.397	0.042	0.067	0.098
10	128	IDW	0.000	1.030	0.138	0.152	0.281
11	128	Linear	0.000	0.794	0.100	0.202	0.214
12	128	Spline	0.000	0.769	0.091	0.141	0.201
13	256	IDW	0.000	1.721	0.240	0.304	0.492
14	256	Linear	0.000	1.408	0.211	0.353	0.419
15	256	Spline	0.000	1.453	0.209	0.305	0.423
16	512	IDW	0.000	2.575	0.394	0.508	0.774
17	512	Linear	0.000	2.233	0.365	0.563	0.690
18	512	Spline	0.000	2.397	0.383	0.540	0.738

Table 19: Testbed 2 statistics for pairwise interpolation methods comparison at various line spacings.

Line Spacing (m)	Interpolation Methods	t-statistics	p-value
16	Spline and IDW	-341.3	0
	IDW and Linear	356.3	0
	Spline and Linear	-63.4	0
32	Spline and IDW	-390.6	0
	IDW and Linear	366.3	0
	Spline and Linear	-147.9	0
64	Spline and IDW	-501.8	0
	IDW and Linear	452.5	0
	Spline and Linear	-158.9	0
128	Spline and IDW	-409.8	0
	IDW and Linear	351.3	0
	Spline and Linear	-142.3	0
256	Spline and IDW	-215.1	0
	IDW and Linear	226.5	0
	Spline and Linear	12.9	0
512	Spline and IDW	-78.9	0
	IDW and Linear	190.5	0
	Spline and Linear	210.8	0

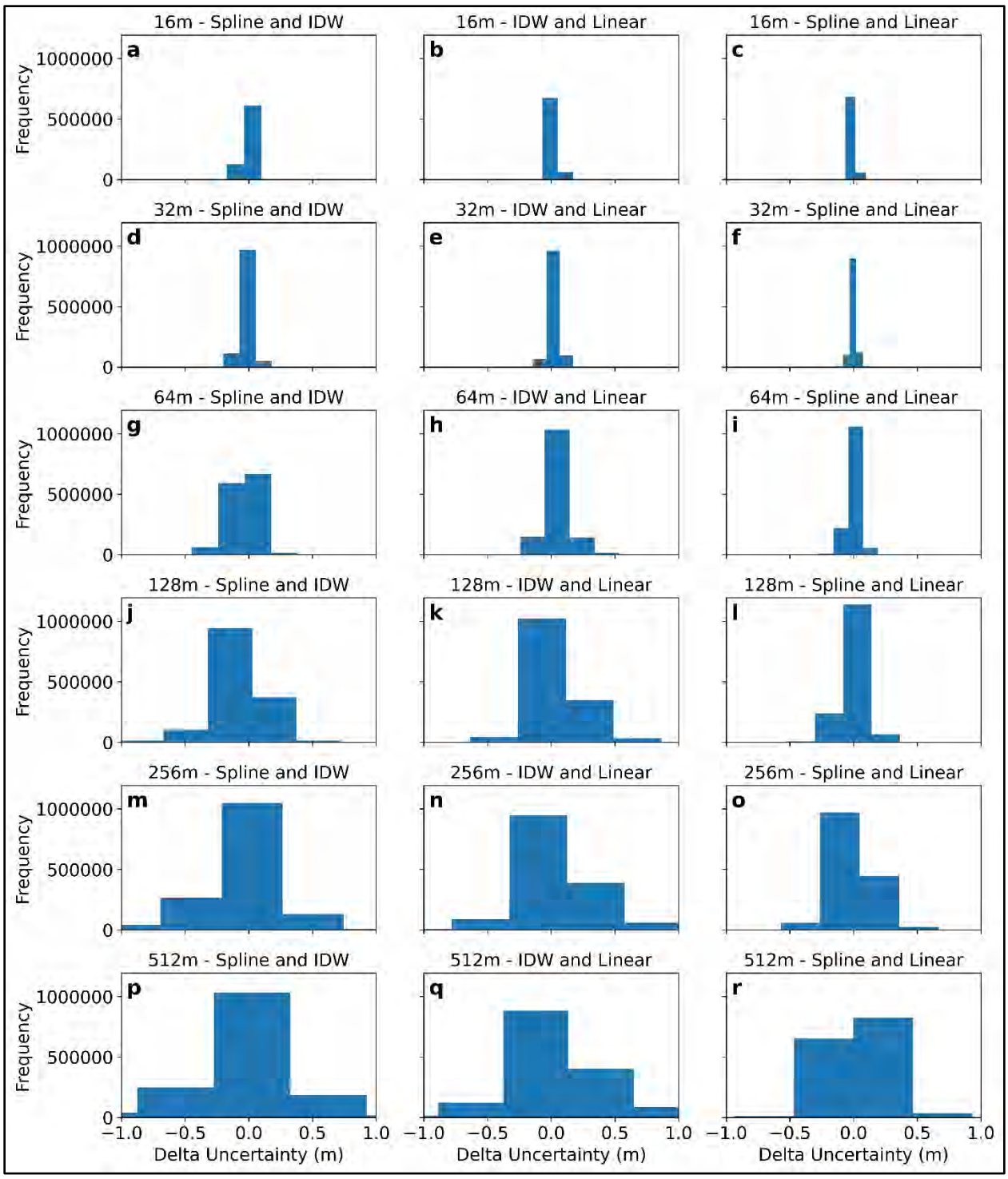

Figure 38: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 2.

5.5.2.2 Spatial Pattern of Interpolation Uncertainties

Figure 39 showcases the spatial distribution of interpolation uncertainties in Testbed 2. The color bar, standardized for ease of comparison, highlights the interpolation uncertainties do not exhibit a random pattern across the testbed. It also indicates that areas with higher interpolation uncertainties are the farthest from a measurement of known depth and are also visually related to high roughness values. This underscores the influence of distance to the nearest measurement and seabed roughness on uncertainty patterns.

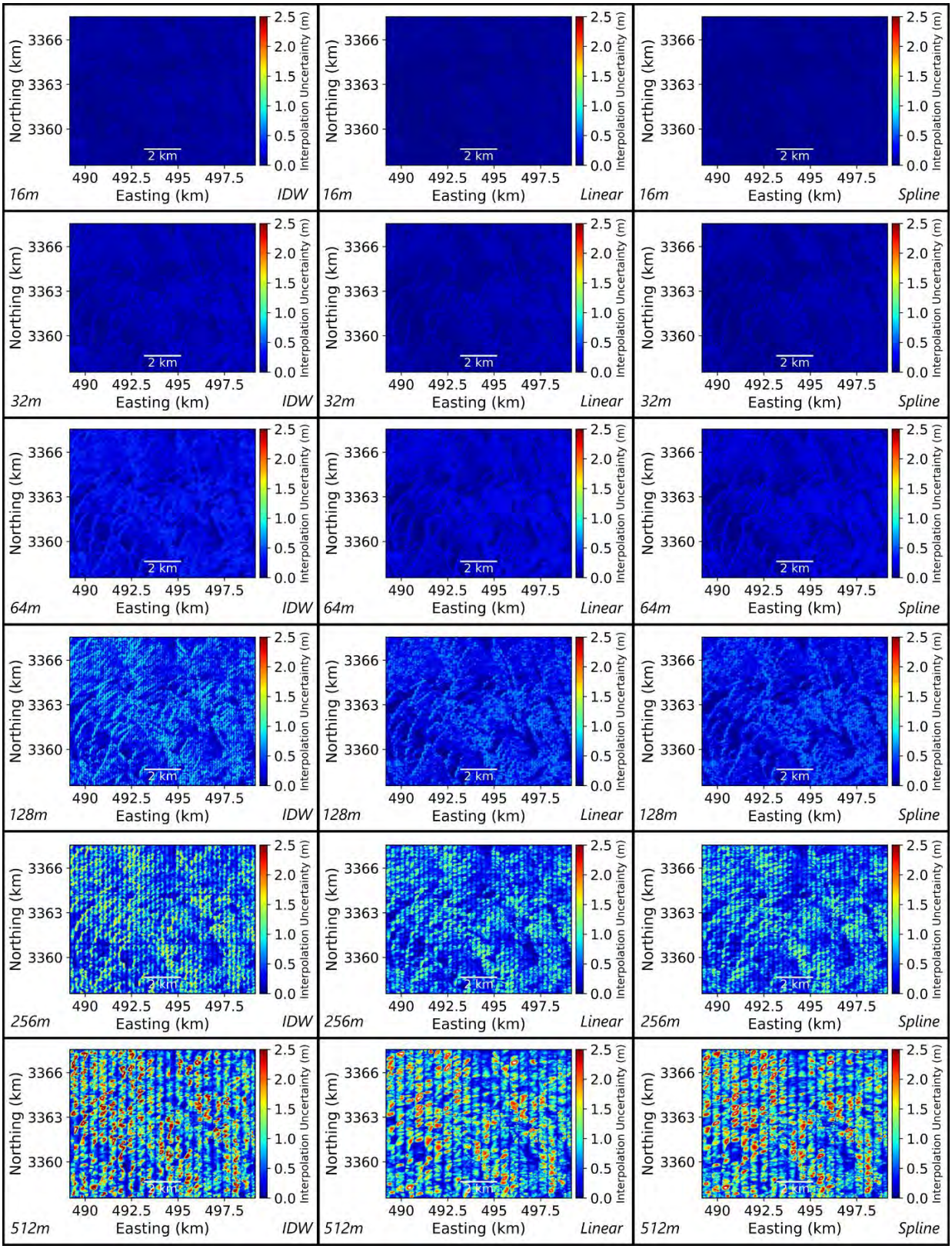

Figure 39: Testbed 2 interpolation uncertainty across all line spacings (columns) and interpolation methods (rows), using 99th percentile of data.

5.5.2.3 Predictive Models of Interpolation Uncertainties

Figures 40a & 40d present the adjusted R^2 and RMSE values for estimated uncertainty based on the distance to the nearest measurement for Testbed 2. A weak relationship is observed which improves with increasing line spacing. IDW outperforms Linear and Spline based on R^2 and performs the worst in RMSE. Linear and Spline perform comparably the same in both metrics.

Figures 40b & 40e (Testbed 2) present the adjusted R^2 and RMSE values for estimated uncertainty based on roughness. A weak relationship is observed which weakens with increasing line spacing. IDW performs the best based on R^2, followed by Spline. Based on RMSE, IDW performs the worst in RMSE while Linear and Spline exhibit similar performances in both metrics.

Figures 40c & 40f (Testbed 2) depict the adjusted R^2 and RMSE values for estimated uncertainty based on slope. The trends observed in roughness are also seen here.

To investigate the hidden non-linear relationships, we employed an ANN, revealing marginal improvements in both R^2 and RMSE (Figure 41). Among interpolation techniques, IDW performs the best in R^2 and performs the least in RMSE. Linear and Spline interpolations exhibit similar performance levels in both metrics. The distinctions in RMSE among the interpolation uncertainty models are not operationally concerning.

Results from the RF analysis coupled with the bootstrap statistical technique reveal that slope emerges as the most influential predictor ranking as the most important in 12 out of 18 instances, with a mean importance level of 49.7% (Table 20). It is followed by slope with a mean importance of 38.8% and appearing as the most important in six of 18 instances. Distance to the nearest measurement is the least important predictor, with a mean importance of 11.5%. Importantly, the bootstrap technique confirms the statistical significance of the differences in these parameters.

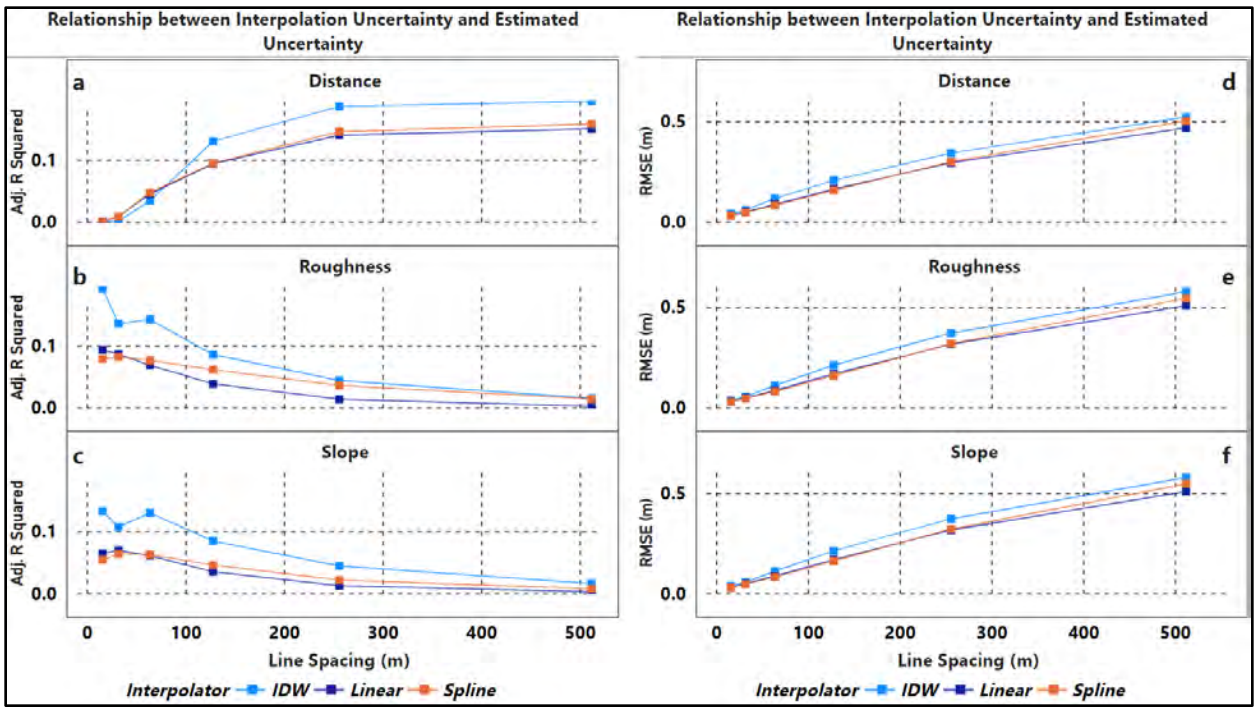

Figure 40: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 2.

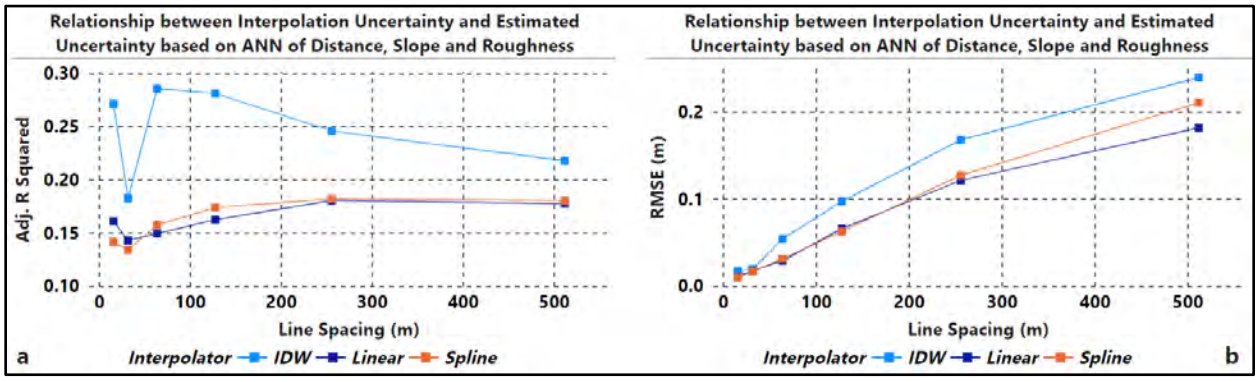

Figure 41: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 2.

Table 20: Statistics of the importance of predictors of uncertainty for Testbed 2

Statistics/Parameter	Distance	Slope	Roughness
Mean Importance (%)	11.5	49.7	38.8
# times of most Important	0	12	6
95% bootstrap percentile CI of the Importance	(11.1,11.8)	(48.9,50.5)	(38.0,39.7)

5.5.3 Testbed 3 (Slopy Seabed)

5.5.3.1 Interpolation Methods

Figure 42 displays the box and whisker plots illustrating the performance of interpolation methods across all line spacings for Testbed 3. These visualizations are complemented by the descriptive statistics provided in Table 21. The interpolation uncertainties of the interpolation methods vary at wider line spacings and become relatively the same at the tighter line spacings of 16m, 32m and 64m. The Spline interpolation method performs better than Linear and IDW at wider line spacings (64m, 128m, 256m, and 512m) while Linear performed the best at 16m and 32m line spacings. Compared to Testbeds 1 and 2, the interpolation uncertainties here are in the order of meters, attributed to the complex morphology of Testbed 3. Importantly, Table 22 and Figure 43 show that the differences seen among the interpolation methods in the box and whisker plots are not statistically significant.

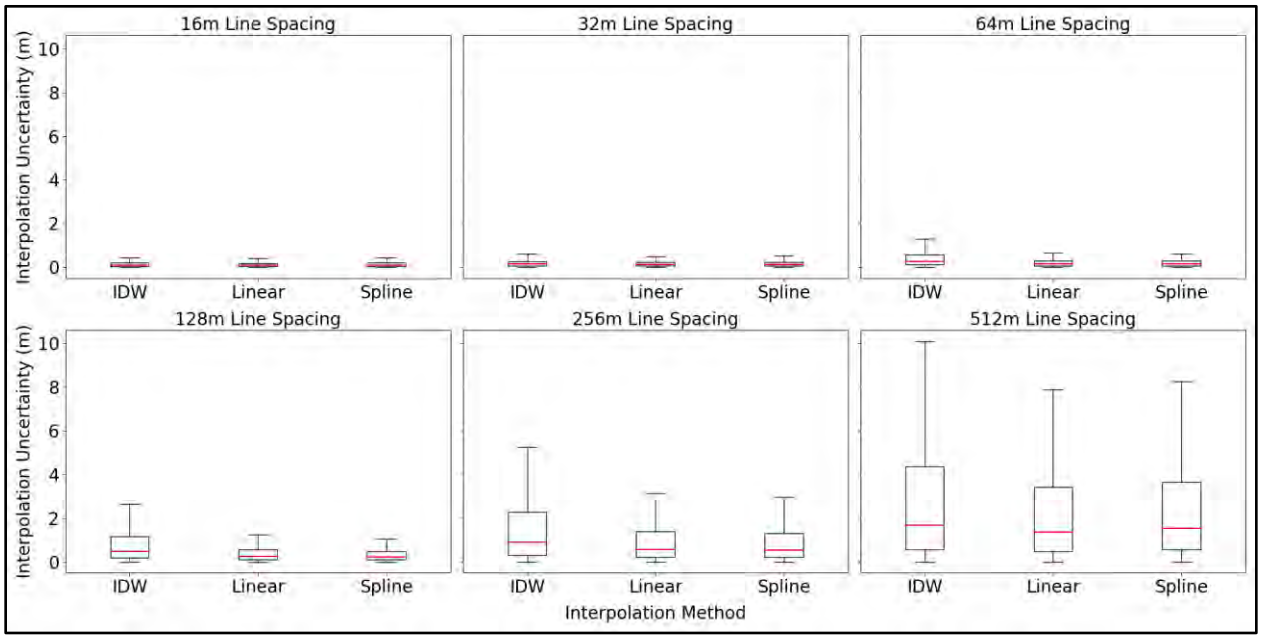

Figure 42: Testbed 3 interpolation methods uncertainty comparison at various line spacings using box and whisker, plotted with 99% percentile of data.

Table 21: Testbed 3 interpolation uncertainty descriptive statistics at various line spacings.

S/N	Line Spacing (m)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error (m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	16	IDW	0.000	1.264	0.101	0.152	0.250
2	16	Linear	0.000	1.285	0.090	0.163	0.240
3	16	Spline	0.000	1.289	0.091	0.155	0.244
4	32	IDW	0.000	1.704	0.135	0.186	0.342
5	32	Linear	0.000	1.465	0.115	0.222	0.286
6	32	Spline	0.000	1.522	0.116	0.190	0.294
7	64	IDW	0.000	4.423	0.258	0.245	0.828
8	64	Linear	0.000	1.889	0.150	0.496	0.378
9	64	Spline	0.000	1.711	0.143	0.227	0.344
10	128	IDW	0.000	9.361	0.472	0.471	1.730
11	128	Linear	0.000	3.925	0.253	0.998	0.766
12	128	Spline	0.000	3.056	0.220	0.387	0.611
13	256	IDW	0.000	18.339	0.897	1.148	3.464
14	256	Linear	0.000	9.955	0.564	1.973	1.927
15	256	Spline	0.000	8.200	0.548	1.049	1.694
16	512	IDW	0.000	33.851	1.695	2.799	6.622
17	512	Linear	0.000	21.378	1.357	3.768	4.635
18	512	Spline	0.000	20.569	1.526	2.817	4.450

Table 22: Testbed 3 statistics for pairwise interpolation methods comparison at various line spacings.

Line Spacing (m)	Interpolation Methods	t-statistics	p-value
16	Spline and IDW	-36.3	0
	IDW and Linear	68.6	0
	Spline and Linear	69.5	0
32	Spline and IDW	-135.4	0
	IDW and Linear	173.5	0
	Spline and Linear	64.4	0
64	Spline and IDW	-475.9	0
	IDW and Linear	453.1	0
	Spline and Linear	-114.1	0
128	Spline and IDW	-529	0
	IDW and Linear	475.8	0
	Spline and Linear	-230.5	0
256	Spline and IDW	-438.3	0
	IDW and Linear	402.3	0
	Spline and Linear	-132	0
512	Spline and IDW	-273.6	0
	IDW and Linear	281	0
	Spline and Linear	6.7	0

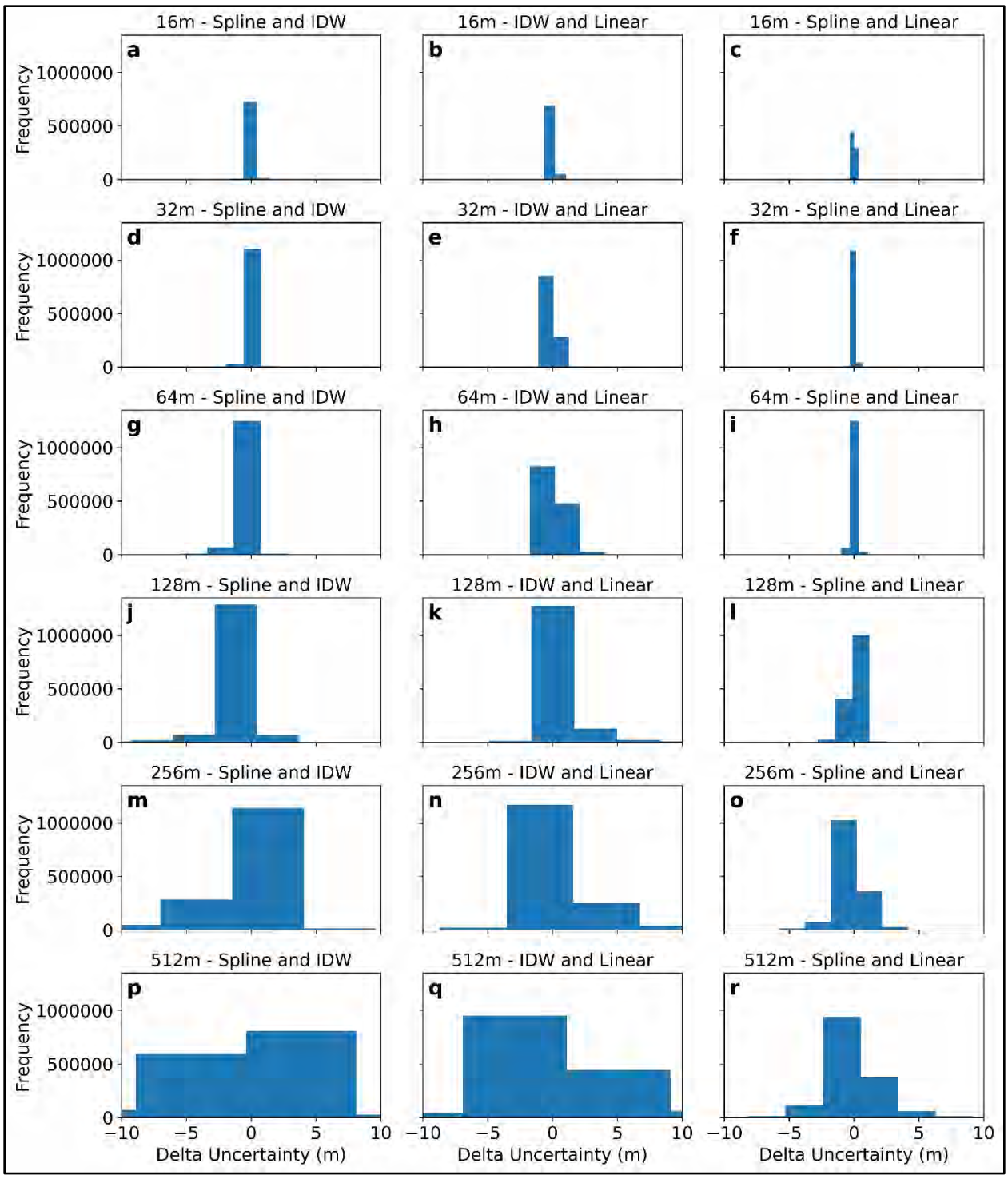

Figure 43: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 3.

5.5.3.2 Spatial Pattern of Interpolation Uncertainties

Examining Figure 44, the spatial distribution of uncertainties in Testbed 3 reveals a non-random pattern. Higher uncertainties are visually related with areas of high distance to nearest measurement and slope values. This observation underscores the impact of distance to the nearest measurement and slope characteristics on interpolation uncertainty.

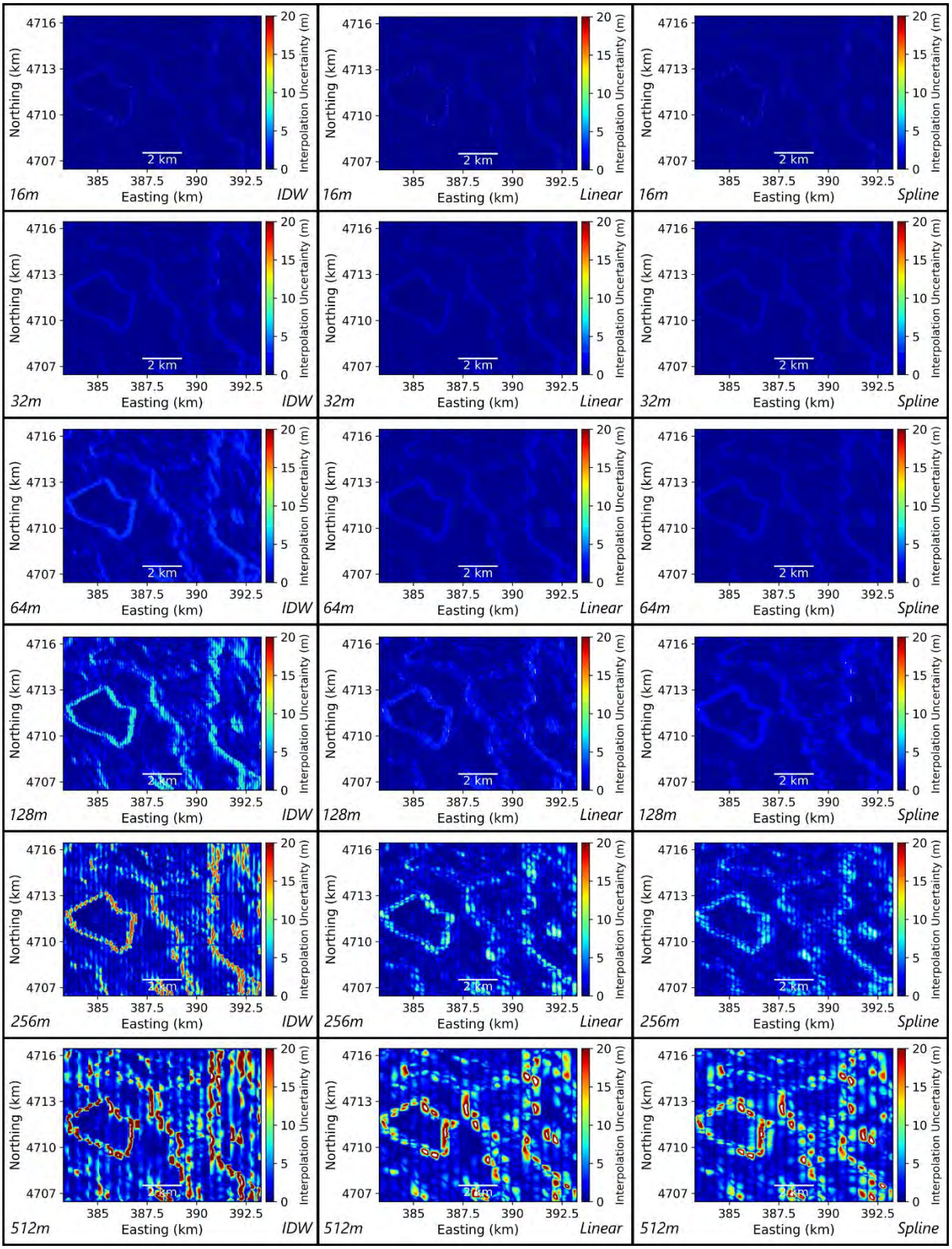

Figure 44: Testbed 3 interpolation uncertainty across all line spacings (columns) and interpolation methods (rows), using 99th percentile of data.

5.5.3.3 Predictive Models of Interpolation Uncertainties

Figures 45a & 45d present the adjusted R^2 and RMSE values for estimated uncertainty based on the distance to the nearest measurement for Testbed 3. A weak relationship is observed which improves with increasing line spacing. The interpolation methods perform comparably the same based on R^2 and Spline outperforms Linear and IDW in RMSE.

Figures 45b & 45e (Testbed 3) present the adjusted R^2 and RMSE values for estimated uncertainty based on roughness. A weak relationship emerges with Spline interpolation consistently outperforming Linear and IDW interpolation in both metrics.

Figures 45c & 45f (Testbed 3) depict the adjusted R^2 and RMSE values for estimated uncertainty based on slope. The relationship is also weak and, in this case, IDW interpolation performs the worst in both metrics, while Linear and Spline interpolations exhibit similar performance levels.

To investigate the hidden non-linear relationships, we employed an ANN, revealing marginal improvements in both R^2 and RMSE (Figure 46). Among interpolation techniques, IDW performs the best in R^2 at tighter line spacing but performs the worst at wider line spacings. IDW also performs the least in RMSE. Linear and Spline interpolations exhibit similar performance levels in both metrics. The distinctions in RMSE among the interpolation uncertainty models are not operationally concerning, with the exception of IDW given that it differs from others by meters. There seems to be an anomaly in the uncertainty model results of IDW (refer to Figure 46a). This anomaly is unique to IDW and is suspected to be caused by the influence of distance to the nearest measurement as a predictor, based on our investigation.

Results from the RF analysis coupled with the bootstrap statistical technique reveal that roughness emerges as the most influential predictor ranking as the most important in 15 out of 18

instances, with a mean importance level of 53% (Table 23). Following in importance are slope and distance to the nearest measurement, with mean importance levels of 37.7% and 9.3%, respectively. Notably, distance to the nearest measurement consistently emerges as the least important factor across all 18 instances. The observed variations in the results are statistically significant, indicating robust differences among the predictors.

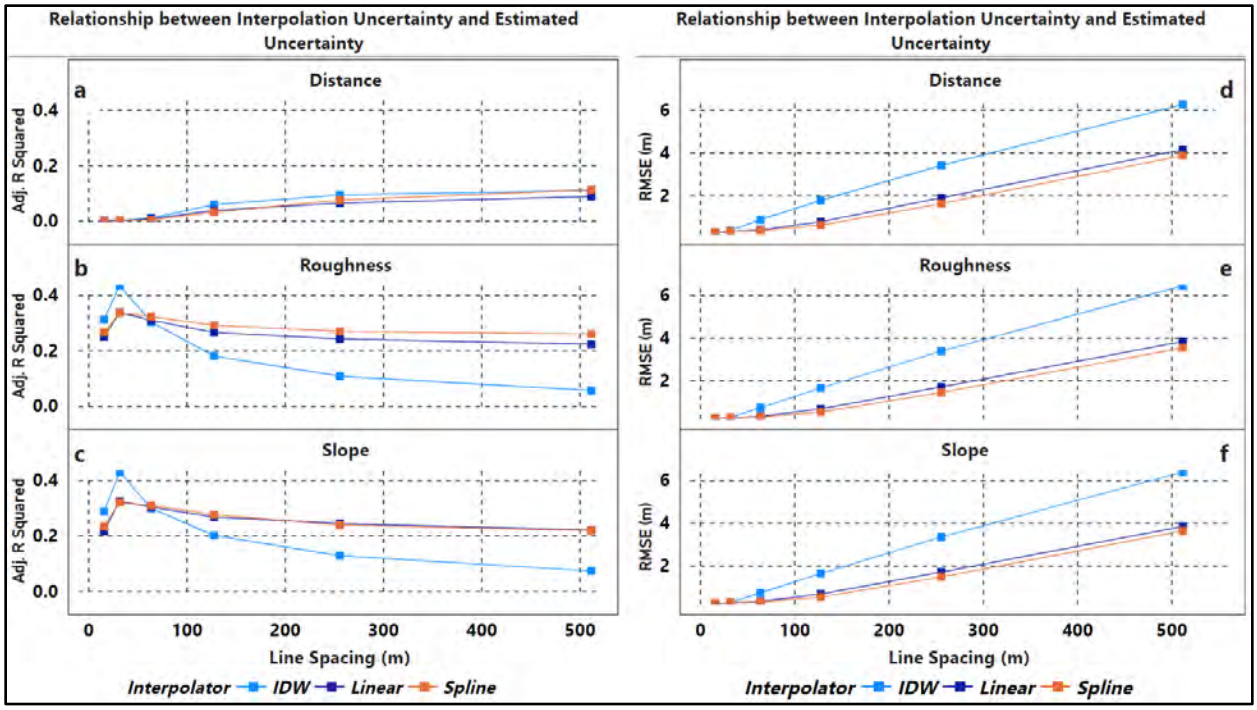

Figure 45: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 3.

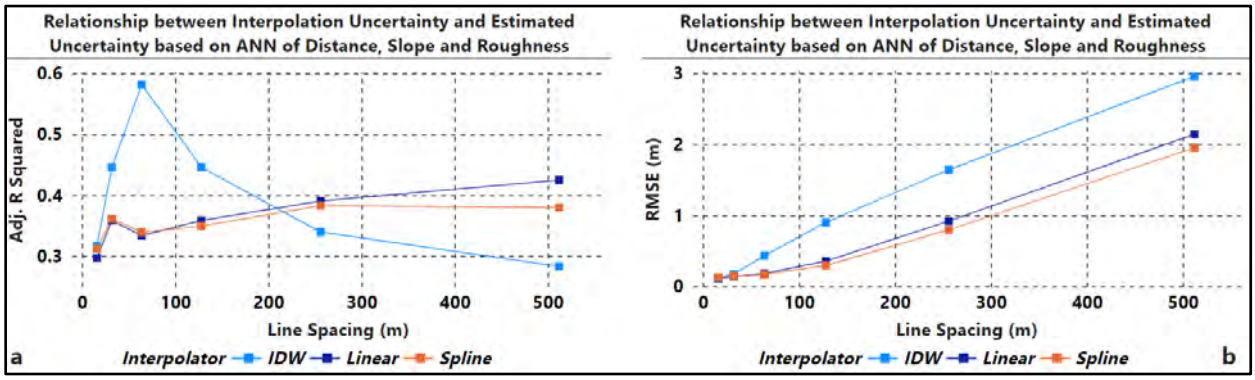

Figure 46: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 3.

Table 23: Statistics of the importance of predictors of uncertainty for Testbed 3

Statistics/Parameter	Distance	Slope	Roughness
Mean Importance (%)	9.3	37.7	53.0
# times of most Important	0	3	15
95% bootstrap percentile CI of the Importance	(8.9,9.7)	(36.2,40.4)	(50.3,54.5)

5.5.4 Testbed 4 (Rough and Slopy Seabed)

5.5.4.1 Interpolation Methods

Figure 47 displays the box and whisker plots illustrating the performance of interpolation methods across all line spacings for Testbed 4. Table 24 complements these visualizations with descriptive statistics of the interpolation uncertainties at the 99th percentile confidence interval for each line spacing and interpolation method. The interpolation uncertainties of the interpolation methods vary at wider line spacings and become relatively the same at the tighter line spacings of 16m, 32m and 64m. The Spline interpolation method performs better than Linear and IDW at both tighter and wider line spacings. Like Testbed 3, the interpolation uncertainties here are in the order of meters, attributed to the complex morphology of Testbed 4. Importantly, Table 25 and Figure 48 show that the differences seen in the box and whisker plots and descriptive statistics for each line spacing are not statistically significant.

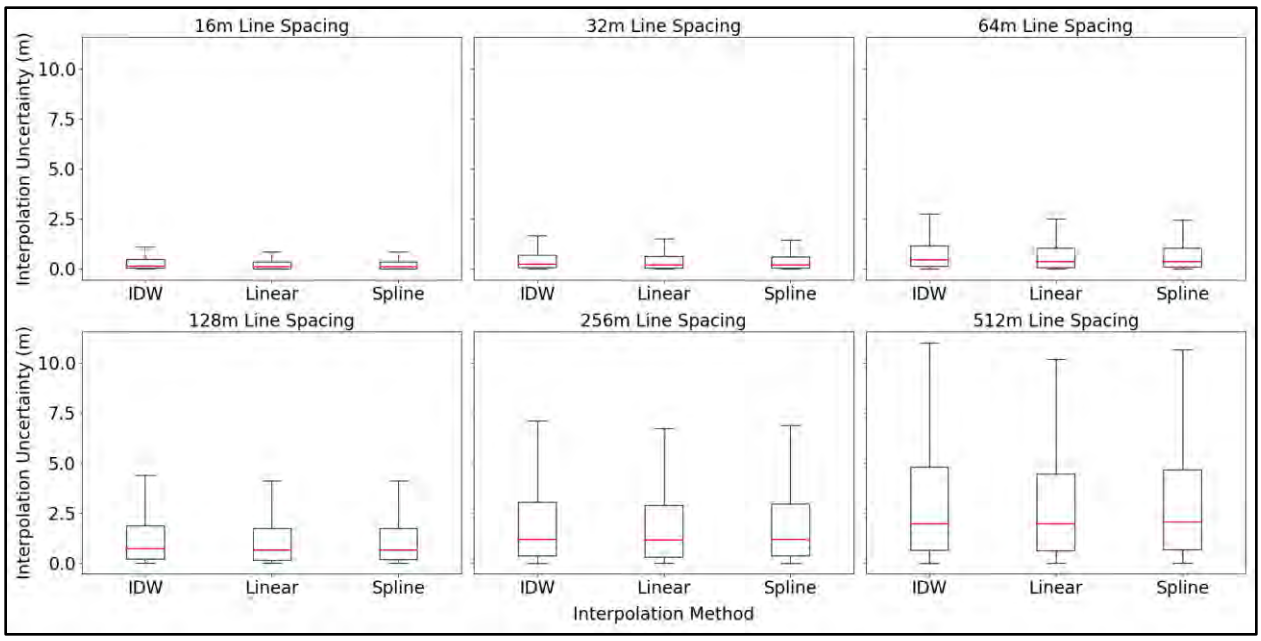

Figure 47: Testbed 4 interpolation methods uncertainty comparison at various line spacings using box and whisker, plotted with 99% percentile of data.

Table 24: Testbed 4 interpolation uncertainty descriptive statistics at various line spacings.

S/N	Line Spacing (m)	Interpolator	99th percentile Min (m)	99th percentile Max (m)	99th percentile Median Error (m)	99th percentile MAE (m)	99th percentile RMSE (m)
1	16	IDW	0.000	2.278	0.145	0.273	0.533
2	16	Linear	0.000	2.255	0.100	0.327	0.477
3	16	Spline	0.000	2.203	0.095	0.264	0.461
4	32	IDW	0.000	3.160	0.238	0.442	0.763
5	32	Linear	0.000	3.195	0.190	0.481	0.731
6	32	Spline	0.000	3.219	0.192	0.433	0.718
7	64	IDW	0.000	5.315	0.434	0.732	1.288
8	64	Linear	0.000	4.820	0.360	0.827	1.167
9	64	Spline	0.000	4.923	0.362	0.731	1.170
10	128	IDW	0.000	8.361	0.703	1.195	2.061
11	128	Linear	0.000	7.169	0.646	1.330	1.838
12	128	Spline	0.000	7.427	0.656	1.222	1.886
13	256	IDW	0.000	12.217	1.171	1.947	3.216
14	256	Linear	0.000	10.381	1.151	2.125	2.887
15	256	Spline	0.000	11.026	1.180	2.025	3.009
16	512	IDW	0.000	15.922	1.961	2.971	4.657
17	512	Linear	0.000	14.098	1.973	3.218	4.212
18	512	Spline	0.000	14.875	2.046	3.132	4.441

Table 25: Testbed 4 statistics for pairwise interpolation methods comparison at various line spacings.

Line Spacing (m)	Interpolation Methods	t-statistics	p-value
16	Spline and IDW	-160	0
	IDW and Linear	176.8	0
	Spline and Linear	-20.3	0
32	Spline and IDW	-106.7	0
	IDW and Linear	113.4	0
	Spline and Linear	-22.5	0
64	Spline and IDW	-139.8	0
	IDW and Linear	153.4	0
	Spline and Linear	6	0
128	Spline and IDW	-111.8	0
	IDW and Linear	147.4	0
	Spline and Linear	59.6	0
256	Spline and IDW	-76.6	0
	IDW and Linear	140.4	0
	Spline and Linear	118.3	0
512	Spline and IDW	-53.9	0
	IDW and Linear	151.8	0
	Spline and Linear	171.9	0

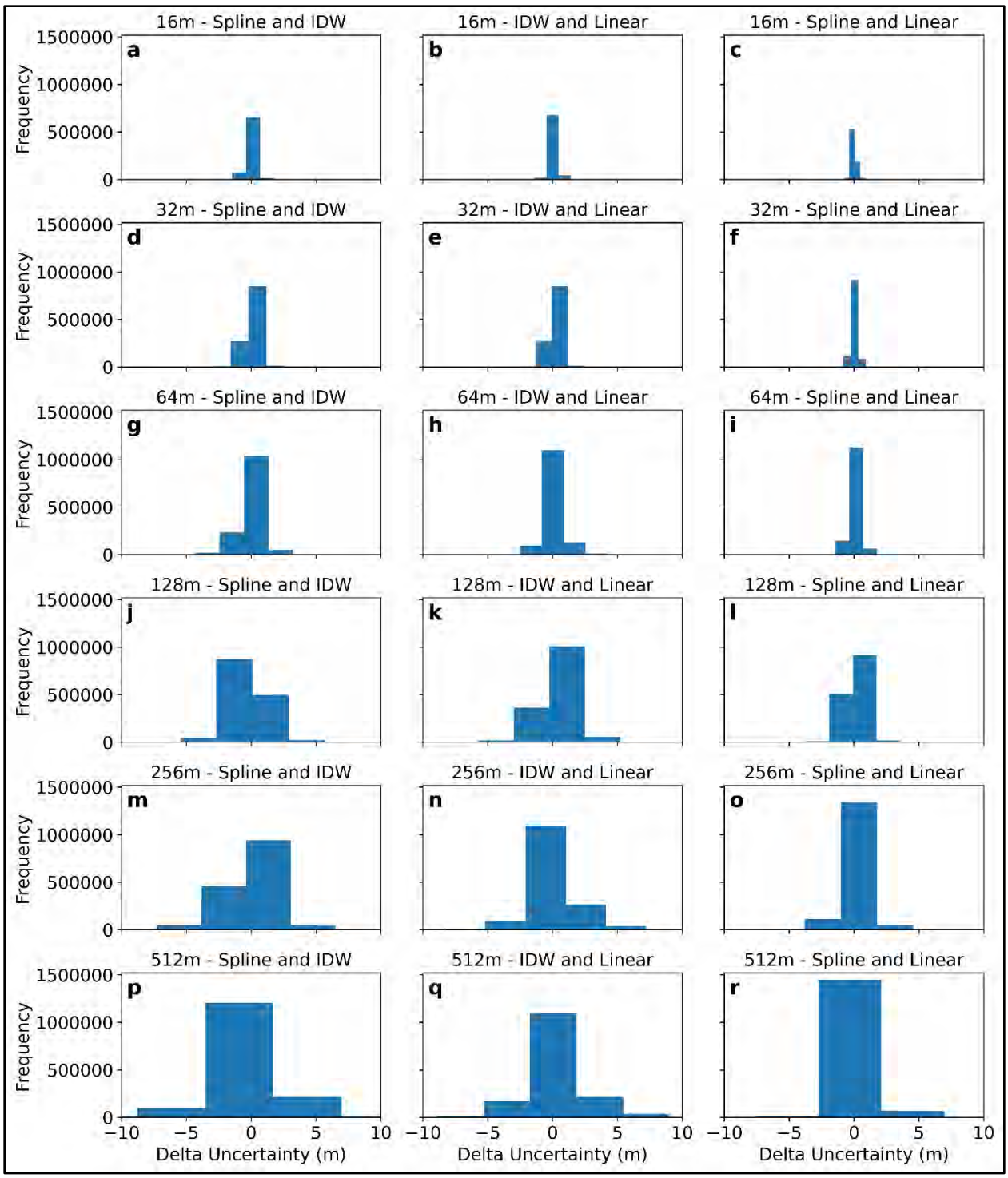

Figure 48: Histograms showing pair-wise differences of interpolation methods uncertainties for Testbed 4.

5.5.4.2 Spatial Pattern of Interpolation Uncertainties

In Figure 49, the non-random pattern of uncertainties in Testbed 4, similar to other testbeds, is evident. Visual examination indicates that areas with higher uncertainties correlate with high distance to nearest measurement and both high slope and roughness values. This relationship emphasizes the combined influence of distance to the nearest measurement, slope, and roughness on interpolation uncertainties.

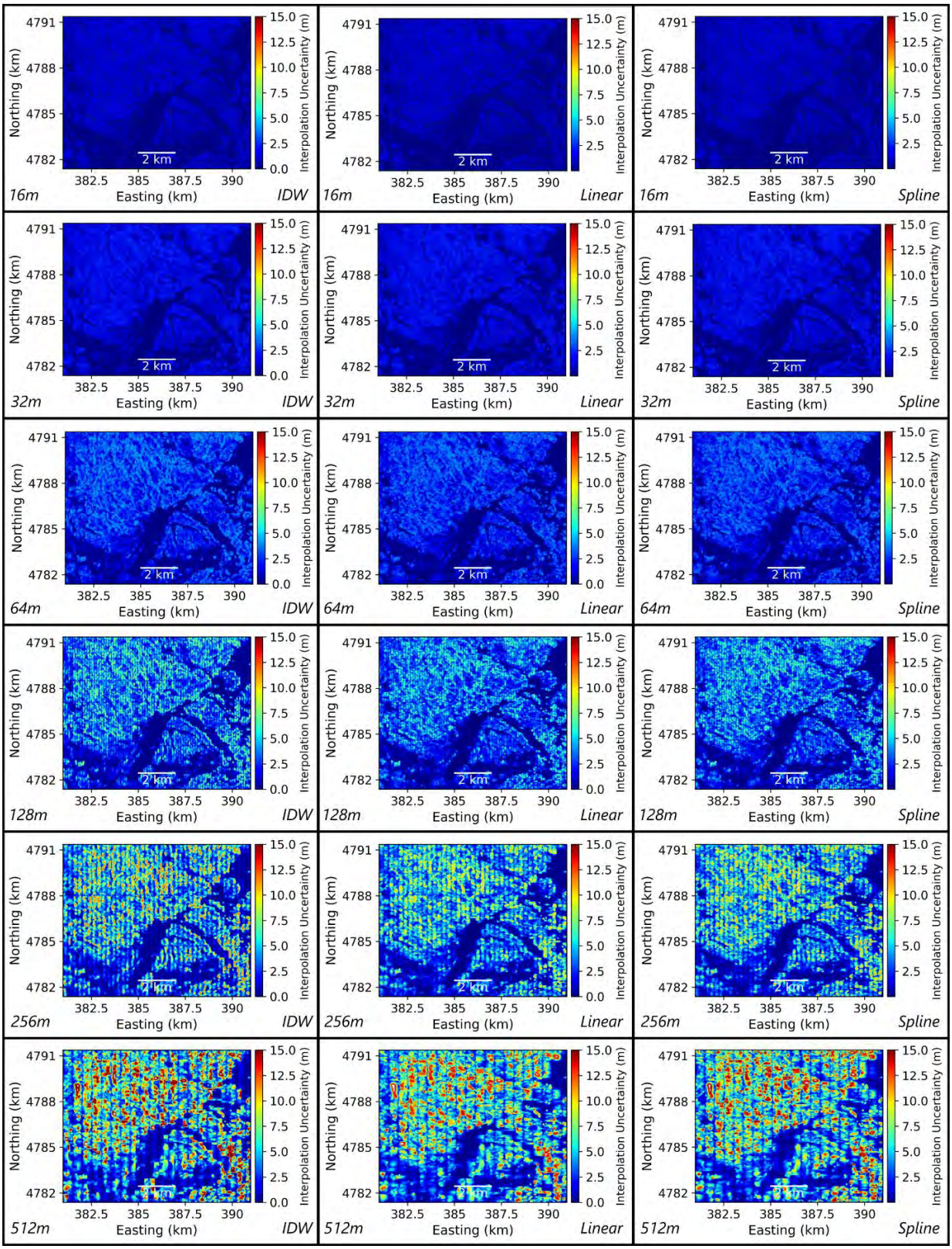

Figure 49: Testbed 4 interpolation uncertainty across all line spacings (columns) and interpolation methods (rows), using 99th percentile of data.

5.5.4.3 Predictive Models of Interpolation Uncertainties

Figures 50a & 50d present adjusted R^2 and RMSE values for estimated uncertainty based on the distance to the nearest measurement for Testbed 4. A very weak relationship improves with increasing line spacing and the interpolation methods exhibit relatively similar performance levels in both metrics.

Figures 50b & 50e (Testbed 4) present the adjusted R^2 and RMSE values for estimated uncertainty based on roughness. A moderately strong relationship weakens with increasing line spacing, and the interpolation methods exhibit relatively similar performance levels in both metrics.

Figures 50c & 50f (Testbed 4) depict adjusted R^2 and RMSE values for estimated uncertainty based on slope. Similar to roughness, a moderate relationship intensifies with rising line spacing, and the interpolation methods exhibit relatively comparable performance levels in both metrics.

To investigate the hidden non-linear relationships, we employed an ANN, revealing marginal improvements in both R^2 and RMSE (Figure 51). Among interpolation techniques, IDW performs best in R^2 but worst in RMSE. Linear and Spline interpolations exhibit similar performance levels. Much like Testbed 1 and Testbed 2, the differences in RMSE among the interpolation uncertainty models are not operationally significant, given that they are in the order of centimeters.

Results from the RF analysis coupled with the bootstrap statistical technique reveal that roughness emerges as the most influential predictor ranking as the most important in 17 out of 18 instances, with a mean importance level of 52.9% (Table 26). Following in importance are slope and distance to the nearest measurement, with mean importance levels of 38.3% and 8.8%,

respectively. Notably, distance to the nearest measurement consistently emerges as the least important factor across all 18 instances. The observed variations in the results are statistically significant, indicating robust differences among the predictors.

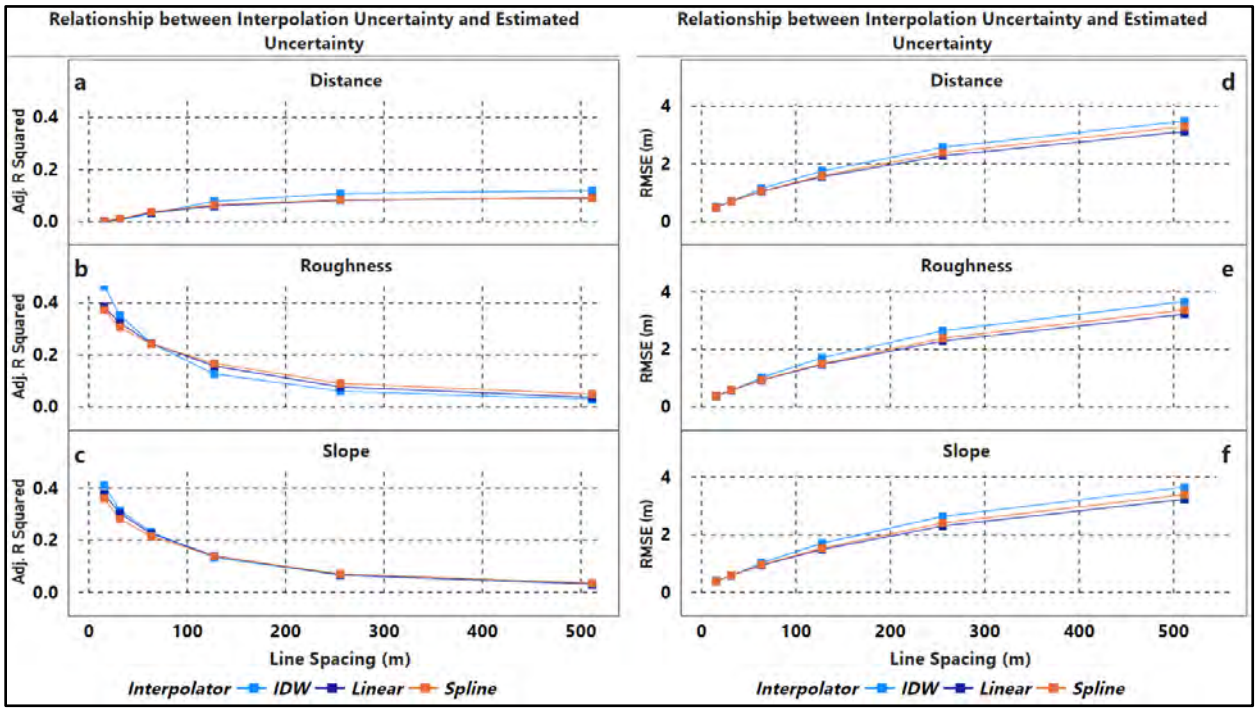

Figure 50: Adjusted R^2 (a-c) and RMSE (d-f) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope respectively for Testbed 4.

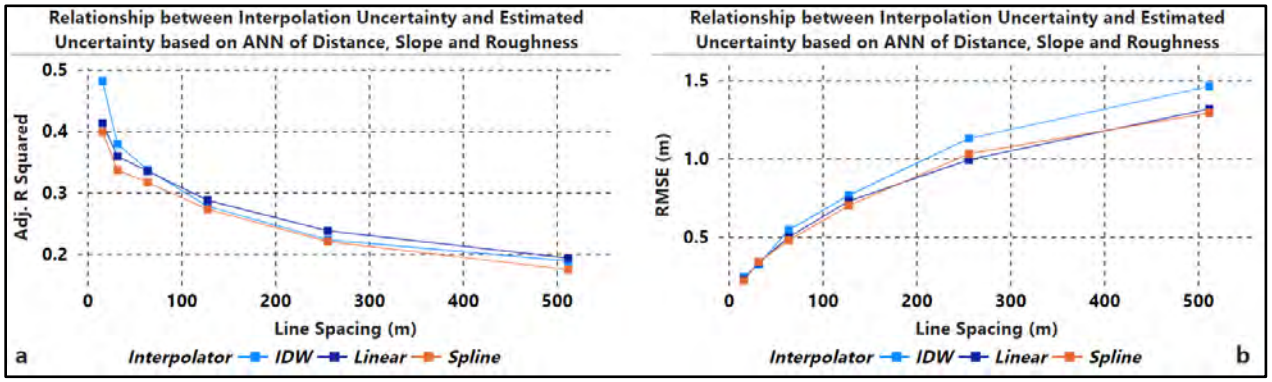

Figure 51: Adjusted R^2 (a) and RMSE (b) of the relationship between interpolated uncertainty and estimated uncertainty based on distance to nearest measurement, roughness, and slope combined for Testbed 4.

Table 26: Statistics of the importance of predictors of uncertainty for Testbed 4.

Statistics/Parameter	Distance	Slope	Roughness
Mean Importance (%)	8.8	38.3	52.9
# times of most Important	0	1	17
95% bootstrap percentile CI of the Importance	(8.5,9.0)	(37.5,39.1)	(52.2,53.7)

5.6 Discussion

5.6.1 Unraveling 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 best for Testbed 1 (Flat), while Spline proved best for Testbeds 2 (Rough), 3 (Slopy), and 4 (Rough and Slopy). IDW consistently performed the worst for all testbeds. However, these interpolation methods demonstrated comparable performance at the same line spacing on each testbed, particularly when assessed from an operational standpoint (*i.e.*, differences of a few centimeters, that may not affect the CATZOC designation due to the depth total vertical uncertainty shown in Table 1). The absence of statistically significant differences in their performance suggests operational equivalence among all three interpolation methods.

Testbeds 1 and 2 exhibited uncertainties in the order of centimeters due to their relatively less complex morphologies. In contrast, Testbeds 3 and 4, featuring more complex morphologies, show uncertainties in the order of meters particularly at wider line spacings (256m and 512m). 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, particularly in the context of generating DBMs from set-line spacing surveys and field survey operations. This insight indicates that seabed morphology is one of the variables that should drive set-line spacing survey density. It also suggests that it is best to use a set-line spacing survey approach on simpler morphologies, as the magnitude of uncertainty incurred from the interpolation process can reach up to 20 meters or more in complex morphologies (refer to Figure 44).

The impact of line spacing is conspicuous, particularly in Testbeds 3 and 4, where uncertainties become relatively similar and smaller at tighter line spacing. This suggests that denser line spacing is a way of overcoming geomorphologically related uncertainty on Testbeds 3 and 4. In summary, interpolation method performance varies across Testbeds, influenced by seabed morphology and line spacing. 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 distance to the nearest measurement and high slope and/or high roughness values in Testbeds 2, 3, and 4, underscore the impact of underlying terrain characteristics and line spacings. The presence of multibeam artifacts in Testbed 1 highlights the transfer of uncertainties from the original data to the interpolation process. Even though survey data can be within the IHO specifications they were targeted for, they can still be affected by multibeam artifacts (Hughes Clarke *et al*. 1996). These findings emphasize the

relationship among data quality, line spacing density, seabed morphology complexity, and the resulting interpolation uncertainties.

In summary, the choice of interpolation method does not significantly impact results, and differences observed in uncertainties are generally not statistically significant. Performance varies across testbeds, suggesting that morphological complexity should drive set-line spacing density.

5.6.2 Relationship among Parameters, Line Spacing, 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 line spacings, roughness emerged as the most important predictor of uncertainty for Testbeds 3 and 4, followed by slope and distance to the nearest measurement, while slope emerged as the most important predictor for Testbeds 1 and 2, followed by roughness 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 16m line spacing sampling density, as Figure 51 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. Distance was observed to correlate well with interpolation uncertainty at wider line spacings.

Notably, the introduction of slope and roughness from interpolated depths to estimate interpolation uncertainty is a novel contribution, without the existence of comparative studies to the best of our knowledge. 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 between interpolation uncertainty and combined ancillary parameters underscores the complexity of the problem, signaling the potential necessity for alternative approaches beyond those applied in our study.

The diminishing effectiveness of the model at tighter line spacings across the testbeds underscores the importance of optimal line spacings. Wider line spacings correspond to increased interpolation uncertainty, and vice-versa. Additionally, as line spacing increases, the uncertainty model struggles to capture the subtle variations in seabed morphology, leading to a reduction in predictive accuracy. This finding accentuates the significance of strategic set-line survey designs, where appropriate line spacings can enhance the estimation of interpolation uncertainty. This has implications for real-world hydrographic surveys targeted at a particular CATZOC level to maximize survey efficiencies and cost and time.

5.6.3 Examining Disparities in Testbed Predictive Performance of Interpolation Methods

The significant variations in predictive performance observed across testbeds underscore the context-specific nuances of interpolation uncertainty. Notably, Testbed 4 (Rough and Slopy) exhibited the highest predictive accuracy (but lowest interpolation 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, with roughness emerging as the most important predictor of uncertainty. This suggests that other unexplored factors may

contribute to the observed results, indicating the complexity of the relationship between parameters and predictive performance.

Testbed 2 (Rough) performed slightly worse than Testbed 3 with slope as the most important predictor of uncertainty. This indicates a potential dissociation between seabed characteristics and predictive accuracy, emphasizing the multifaceted nature of this relationship.

Finally, Testbed 1 (Flat) displayed a lower predictive performance, affirming the challenge of capturing variability in less complex terrains.

These findings underscore 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.

5.6.4 Importance of Predictors

An in-depth analysis of RF results, complemented by the bootstrap statistical technique, revealed that roughness holds the highest predictive importance, followed by slope and distance to the nearest measurement for Testbeds 3 and 4 while slope emerged as the most important predictor for Testbeds 1 and 2. Across all testbeds, the consistently low importance of distance to the nearest measurement implies its limited contribution to the predictive models. In contrast, the dominance of roughness and slope underscores the significance of terrain characteristics in driving interpolation uncertainties.

Notably, on Testbeds 1 and 2, slope emerged as the most important predictor while competing with roughness, potentially due to the less complex morphology of these seabeds. 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 set-line spacing surveys. Understanding the hierarchical importance of predictors provides valuable insights into how uncertainty in interpolated bathymetry from set-line spacing surveys can be quantified.

5.7 Summary

The examination of interpolation methods within set-line spacing surveys revealed that Linear produced the lowest interpolation uncertainty on Testbed 1 (Flat), followed by Spline and IDW. On the other hand, Spline produced the lowest uncertainty for Testbeds 2 (Rough), 3 (Slopy), and 4 (Rough and Slopy). However, the differences in the performance of the interpolation methods are operationally insignificant, *i.e.*, their impact on the CATZOC allocations for charting is the same. Additionally, seabed complexity significantly influences uncertainties, with Testbeds 1 and 2 yielding more accurate results compared to Testbeds 3 and 4. Interpolations are best performed on less complex morphologies and the choice of interpolation method should be tailored to specific seabed characteristics. The predictors of interpolation uncertainty, roughness, slope, and distance to the nearest measurement, are important from a statistical perspective. However, the weak relationship between interpolation uncertainty and roughness, slope, and distance to the nearest measurement, individually and when combined highlighted the complexity of estimating uncertainty in interpolated bathymetry. Combining parameters with ANN slightly improves the predictive accuracy of the uncertainty model in comparison to the linear regression model of the interpolation uncertainty with each individual parameter. Roughness held the highest predictive importance, followed by slope and distance to the nearest measurement in Testbeds 3 and 4. Bootstrap analysis confirms the statistical significance of observed differences in the importance of predictors. On Testbeds 1 and 2, slope competes with roughness to emerge as the most important

predictor. The low importance of distance to the nearest measurement across all testbeds emphasized the significance of terrain characteristics in driving interpolation uncertainties. The performance of uncertainty estimation varies significantly across the testbeds, with Testbed 4 (Rough and Slopy) demonstrating the best performance with highest adjusted R^2 and lowest RMSE, followed closely by Testbed 3 (Slopy), then Testbed 2 (Rough), and Testbed 1 (Flat). The diminishing effectiveness of the model at tighter line spacings underscores the importance of optimal survey designs in interpolation uncertainty quantification. While the predictive accuracy, based on adjusted R^2, of the uncertainty model generated through IDW interpolation slightly outperformed those produced by Spline and Linear interpolation methods across the majority of testbeds, the RMSE of the IDW uncertainty model was the least favorable. Summarily, the choice of interpolation method does not significantly impact results, and differences observed in uncertainties are generally not statistically significant. Performance varies across testbeds, influenced by morphological complexity. Most importantly, it does not seem feasible to accurately estimate interpolation uncertainty solely based on geomorphology and set-line spacing.

5.8 Conclusion

This study aimed to identify the best bathymetric gap-filling interpolation method – that minimizes interpolation uncertainty – and accurately quantify and characterize uncertainty in set-line spacing surveys. It sought to establish the relationship between interpolation uncertainty and a suite of ancillary parameters – distance to the nearest measurement, slope, and roughness – across four testbeds in the United States, using tighter and wider line spacings.

The findings revealed that Spline is the best interpolation method for Testbeds 2 (Rough), 3 (Slopy), and 4 (Rough and Slopy), followed by Linear and IDW. For Testbed 1, Linear

interpolation resulted in the least interpolation uncertainty, followed by Spline and IDW. However, the study also showed the operational equivalence of accuracies produced by different interpolation methods, as the differences observed are generally negligible from an operational standpoint, reflecting uncertainties in centimeters that may not practically impact CATZOC designations.

The weak relationship between the interpolation uncertainty and the parameters, roughness, slope, and distance to the nearest measurement highlighted the challenges in estimating uncertainty in set-line spacing surveys. This in turn impacts the ability to establish operational survey plans that will maximize accuracy for a minimum cost. The introduction of machine learning techniques to combine the parameters into a single ANN model provides marginal improvements, while revealing the dominance of roughness and slope as the most important predictors, across the testbeds. Additionally, IDW captured better the variability of interpolation uncertainty than Spline and Linear methods for most testbeds. The impact of appropriate line spacing on an uncertainty model's explanatory power is evident, diminishing effectiveness with tighter line spacings and wider line spacings.

Moreover, the estimation of interpolation uncertainty varied across testbeds, with Testbed 4 (Rough and Slopy seabed) yielding the best results (R^2 of 0.4 at 16m line spacing), followed by Testbed 3 (Slopy seabed), Testbed 2 (Rough seabed), and Testbed 1 (Flat seabed). These insights highlighted the presence of unaccounted-for factors influencing uncertainty or a strong random component within interpolation uncertainty.

While the study focused 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 set-line spacing surveys, offering valuable perspectives for uncertainty estimation, hydrographic survey planning, and future research and applications in this domain.

CHAPTER 6 : CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS

6.1 Conclusions

This study aimed to quantify and characterize the interpolation uncertainty in both random sampling and set-line spacing survey datasets and also identify the bathymetric interpolation method that produces the lowest interpolation uncertainty. Across five testbeds in the United States, encompassing varied seabed morphologies and employing different sampling densities and line spacings, the research investigated the relationship between interpolation uncertainty and ancillary parameters: distance to the nearest measurement, slope, and roughness.

In both Chapter 3 (set-line spacing approach) and Chapter 2 (random sampling approach), interpolation methods exhibited comparable performance, with Linear performing best for Testbed 1 (Flat seabed) and Spline performing best for Testbeds 2 (Rough seabed), 3 (Slopy seabed) and 4 (Rough and Slopy seabed). However, the differences in the performance of interpolation methods are not operationally significant, translating to uncertainties in centimeters that insignificantly affect CATZOC allocations.

The magnitude of uncertainties in both chapters was comparable for Testbeds 1 and 2, whereas, for Testbeds 3 and 4, Chapter 3 demonstrated higher interpolation uncertainty. The spatial distribution of interpolation uncertainty in both chapters follows non-random patterns, strongly influenced by morphological complexity. Notably, in Chapter 3, the distance to the nearest measurement also impacted the spatial distribution of uncertainty significantly.

Interpolation uncertainty estimation using geomorphological variables and distance from a point of known depth demonstrated comparable performance across testbeds in both chapters, with Testbed 4 exhibiting the best results, followed by Testbeds 3, 2, and 1. Despite weak relationships between interpolation uncertainty and ancillary parameters, the parameters examined were

statistically significant. In Chapter 2, roughness emerged as the most important predictor for all testbeds, while in Chapter 3, slope dominated for Testbeds 1 and 2, and roughness for Testbeds 3 and 4. The impact of sampling density and line spacing on uncertainty models was evident, identifying optimal operational conditions at 10% sampling density and appropriate line spacing.

This research underscores the complexity of estimating uncertainty in interpolated bathymetric datasets, indicating either the presence of unaccounted-for factors driving uncertainty, or a random component within interpolation uncertainty. Nevertheless, this study 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.

6.2 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, in which uncertainty estimates are inherent, 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 characteristics of the testbeds. While optimization could potentially enhance the accuracy of interpolation uncertainty estimates, such fine-tuning was beyond the scope of this research.

Additionally, the work focused on the uncertainty metric associated with depth accuracy in CATZOC, excluding the horizontal accuracy part and considerations for seafloor feature detection and factors like expected feature size or seabed undulations.

Finally, while this study attempts to consider morphologies that are widely known and prevalent within the US waters, it is important to note that the broader applicability of the findings herein might not be applicable universally, geographically, morphologically, and/or analytically.

6.3 Future Research Directions

The results of this study indicated that ancillary variables considered do not provide high accurate interpolation uncertainty estimation. Subsequent research endeavors 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 associated with uncertainty attributed to various factors (refer to Hare *et al*. 2011). Future research endeavors should consider integrating measurement uncertainty with interpolation uncertainty, a practice that would facilitate the 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 methodological data assumptions – e.g., homogeneity of variance – assumptions that are fundamental to the Kriging method.

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

REFERENCES

Adams, K.T., 1942. Hydrographic Manual. *U.S. Department of Commerce, Coast and Geodetic Survey, Special Publication*, (143).

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.

Backus, R. and Bourne, D., 1987. Georges Bank. *MIT Press*.

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.

Castiglioni, S., Castellarin, A., and Montanari, A., 2009. Prediction of low-flow indices in ungauged basins through physiographical space-based interpolation. *Journal of Hydrology*, 378 (3–4), 272–280.

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.

Chorley, R.J., n.d. Spatial analysis in geomorphology.

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.

Fan, L., 2022a. Comparisons of five indices for estimating local terrain surface roughness using LiDAR point clouds. *In*: *2022 29th International Conference on Geoinformatics*. IEEE, 1–6.

Fan, L., 2022b. Comparisons of five indices for estimating local terrain surface roughness using LiDAR point clouds. *In*: *2022 29th International Conference on Geoinformatics*. IEEE, 1–6.

Fred Agarap, A.M., 2018. Deep Learning using Rectified Linear Units (ReLU).

Grayson, R. and Blöschl, G., 2000. Spatial patterns in catchment hydrology: observations and modelling. *Spatial patterns in catchment hydrology: observations and modelling.*

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.

Gustafson, E.J., 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems*, 1 (2), 143–156.

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

Hawley, J.H., 1931. Hydrographic Manual. *U.S. Department of Commerce, U.S. Coast and Geodetic Survey, Special Publication*, (143).

Hengl, T., 2006. Finding the right pixel size. *Computers & Geosciences*, 32 (9), 1283–1298.

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

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

IHO S-101, 2022. International Hydrographic Organization S-101 Annex A, Data Classification and Encoding Guide, (1.0.2).

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.

Jobson, J.D., 1991. Multiple Linear Regression. 219–398.

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*.

Kingma, D.P. and Ba, J., 2014. Adam: A Method for Stochastic Optimization.

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.

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.

Lundberg, S. and Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions.

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.

Neff, D. and Wilson, M., 2018. ETRAC'S EVALUATION OF QIMERA: Accomplishing the NOAA Workflow. *In*: *Canadian Hydrographic Conference*.

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].

NOAA OCS, 2022. Hydrographic Survey Specifications and Deliverables.

NOAA OCS, 2021. Field Procedures Manual.

NOAA OCS BlueTopo, 2023. BlueTopo [online]. Available from: https://www.nauticalcharts.noaa.gov/data/bluetopo.html [Accessed 24 Jan 2023].

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.

Ross, A. and Willson, V.L., 2017. Paired Samples T-Test. *In*: *Basic and Advanced Statistical Tests*. Rotterdam: SensePublishers, 17–19.

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*.

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

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, M., 2014. Roughness in the Earth Sciences.

Smith, W.H.F., 2004. Introduction to this special issue on bathymetry from space. *Oceanography*, 17 (SPL.ISS. 1), 6–7.

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.

Uchupi, E. and Bolmer, S.T., 2008. Geologic evolution of the Gulf of Maine region. *Earth Science Reviews*, 91, 27–76.

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.

van der Wal, D. and Pye, K., 2003. The use of historical bathymetric charts in a GIS to assess morphological change in estuaries. *The Geographical Journal*, 169 (1), 21–31.

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.

Wong, A.M., Campagnoli, J.G., and Cole, M.A., 2007. Assessing 155 years of hydrographic survey data for high resolution bathymetry grids. *Oceans Conference Record (IEEE)*.

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.