Bayesian downscaling of global bathymetry using spaceborne lidar

M. A. E. Lindsay

Technische Universiteit Delft

TU Delft

Van Oord

Marine ingenuity

Bayesian downscaling of global bathymetry using spaceborne lidar

by

M. A. E. Lindsay

to obtain the degree of Master of Science
at the Delft University of Technology,

Student number:	5243610	
Project duration:	March, 2022 – October, 2022	
Thesis committee:	Prof. dr. ir. S. Aarninkhof,	TU Delft, Chair
	Dr. ir. B. Wouters,	TU Delft, Supervisor
	Dr. ir. A. Gijón-Mancheño,	TU Delft
	Dr. ir. G. J. de Boer,	Van Oord, Supervisor

To be defended publicly on:
28 October, 2022
13:30 - 15:00

Van Oord
Marine ingenuity

TU Delft

Preface

This is the end of a long journey to complete my Master's of Science in Civil Engineering, with a specialization in Coastal Engineering. It has been 2 years of work under sometimes less-than-ideal conditions caused by a global pandemic. But I am lucky to have written this thesis at the TU Delft and with Van Oord. Moving to a foreign country during a global pandemic sometimes feels almost reckless in retrospect, but I have grown and learned so much from all the people I have met here at the TU that I'm incredibly glad I made the leap.

I would like to thank my graduation committee members for their input and guidance throughout this project. Especially Gerben de Boer, for choosing me to work with him and his team at the Datalab, and for some very nerdy conversations about our shared love of Docker. Your enthusiasm, sharp questions, and constant ideas were integral to the final product. Thank you for contributing your time and your team's resources to my project. Thanks to Alejandra Gijón Mancheño, for your help shaping the early version of my project, and your flexibility as the scope changed. Your feedback, and our conversations about how to approach things really helped me develop as a student/researcher. To Bert Wouters, thank you for your feedback on my writing and providing answers to my ICESat-2 questions. And last but not least, thank you to my committee chair Stefan Aarninkhof for your encouragement and enthusiasm during committee meetings, and your suggestions which gave the project shape and form.

Finally, thank you to my parents and my entire family for your unending support — emotional, financial and in more other ways than I can articulate. None of this would have been possible without your love and support. Finally, thank you to my friends: new friends here in Delft who kept me sane in the darkest days of the pandemic, and longtime friends on the East coast and all over the world, who I still am close to despite living on another continent. I cannot wait to see what adventures our futures hold. Thank you so much, for everything.

M. A. E. Lindsay
Delft, October 2022

Abstract

Nearshore bathymetry needs Knowledge of nearshore bathymetry is crucial for every aspect of the blue economy. Currently, it is expensive to obtain nearshore bathymetry at regional scales, and therefore this data is not available for many coasts. The data scarcity occurs especially in the global south and big ocean states that are at the highest risk from climate change. Currently the only global bathymetric dataset, GEBCO, provides depth data at approximately a 500 m resolution. This data is of limited accuracy in the nearshore zone because the source data is sparse in nearshore areas. Recent research has found that NASA's ICESat-2, launched in 2018 to study the cryosphere, can incidentally capture bathymetric data. Some studies have been done in small sites, but the potential for a global product has not yet been investigated. This thesis proposes a method of using GEBCO data as a starting point and incorporating ICESat-2 data via a Kalman filter. This results in a product with a downscaled spatial resolution and improved accuracy without requiring any in-situ data. If other local bathymetry data is available, it can be added as input to the Kalman filter or used for validation of the method.
Lidar satellite-derived bathymetry Recent studies in lidar remote sensing have shown that ICESat-2 can capture nearshore bathymetry at depths of up to 20 meters in tracks 0–3 km apart, provided atmospheric conditions are good and the water is sufficiently clear. Hence, this data could provide a source of high-resolution bathymetric depth profiles in tropical areas and possibly at higher latitudes. This research project proposes an automated processing chain, written in python, for extracting bathymetry points from the lidar data based on the density of the photon returns in the underwater zone. This method can reliably identify points containing bathymetric signal.
Data assimilation via Kalman Updating To downscale the data using a Kalman filter, first global data from GEBCO is clipped to the area of interest, and then resampled bilinearly to 50 m resolution. Then, the ICESat-2 photon data for the area is processed to generate point estimates of bathymetric depth. To fill in the gaps between these point estimates, the bathymetric points are subsampled and interpolated to the same resolution as the GEBCO data using a universal kriging interpolator. This interpolator results in a raster of the estimated depth, and a raster of the estimated uncertainty. To update the interpolated GEBCO grid, the Kalman gain is calculated for each raster cell and using the Kalman state equation a new bathymetry grid is produced. If other data is available, the process can be applied recursively with other depth and uncertainty grids, allowing the Bayesian combination of any number of bathymetry datasets for the site.
Validation To validate the method, the improvement in RMS error is calculated between the resulting bathymetry grid and previously validated, high-accuracy survey data. The validation has been applied at several global test sites to verify that the method is generalizable to other regions. Validation sites are chosen based on the availability of validation data from either USGS survey, Dutch Jarkus Data, or survey data from Van Oord projects. The approach was found to improve RMS error compared to validation data by up to 34% in some test sites. However, other sites showed marginal improvement, or increased RMS error.
Outlook The results of this research could allow easier methods of characterizing nearshore bathymetry in remote areas. It could also be extended to include temporal variation – as more bathymetric data becomes available, it could be used to measure the dynamic changes of coastal systems. Data with this temporal dimension could provide valuable validation data for coastal dynamics models. However, the method is limited by strict water clarity requirements and missing data due to atmospheric conditions. Despite these limitations, it could potentially be expanded to create a downscaled global bathymetry dataset of clearwater areas.

Summary

Bathymetric data is valuable because it is essential for nearly every aspect of coastal management and monitoring but is difficult to obtain, especially in the nearshore zone. Traditional techniques like acoustic sounding provide reasonably high-accuracy data, but survey campaigns are expensive and the survey vessels cannot operate in the shallowest areas of the nearshore zone. Airborne lidar surveying provides high resolution, high accuracy data in the shallowest waters, but the associated survey campaigns are extremely expensive and are very sensitive to environmental conditions. Because of the limitations of conventional survey techniques in the nearshore zone, there is much interest in ways of estimating bathymetry using only spaceborne remote sensing data. Spaceborne remote sensing data is advantageous because of wide global coverage and because the data is often publicly available.

Traditionally, there are two main techniques for extracting bathymetric estimates from satellite data. One approach, called optical satellite derived bathymetry (SDB), estimates the bathymetry based on the relationship between the attenuation of different parts of the visible light spectrum in the water column. These techniques require in-situ data to calibrate, and require the site to have water that is clear down to the seabed. The second approach is to use satellite imagery to estimate the wave field in a region, and then back-calculate the bathymetry based on the evolution of the wave field in space. This is called the wave-kinematic approach. This approach provides less accurate and lower resolution bathymetry estimates than optical SDB, and the deepest depth that can be estimated is limited by the maximum wavelength at the time of the satellite data acquisition. The advantage of wave-kinematic approaches is that the deepest estimation depth is not limited by water clarity.

NASA's ICESat-2 satellite is a lidar satellite that provides high-accuracy point elevations along linear transects all over the world. It was originally intended to study the elevation of ice in the polar regions, but researchers have discovered that it can sometimes also capture the elevation of the seabed in coasts with very clear water. This has led to a number of papers that extract bathymetric point data from ICESat-2 and use it in combination with optical SDB methods to produce a bathymetric estimate without requiring any in-situ data.

Most of the prior studies that extract bathymetric point estimates from ICESat-2 look at a limited number of satellite passes (3–10), and either use manual extraction of the seabed signal based on the researcher's judgement, or use other semi-automated techniques that estimate the seabed based on the density of photons in the vertical direction that are manually checked before use. The goals of this project are to find ways to automate the signal extraction from ICESat-2 data so that it can be implemented at a larger scale (up to hundreds of transects), to evaluate how these sparse point measurements can be interpolated into a continuous gridded product, and to evaluate if this interpolated data can be combined with existing global datasets to add additional value by increasing the resolution and decreasing the error.

The methodology proposed is split into 3 major parts. First, the data processing and signal extraction from the satellite data: For a given area of interest, the ICESat-2 data is downloaded and processed to extract only photons that are located in the nearshore zone. It is then corrected for the effect of light refraction in the water column. Then, points that are likely to be sea surface signal are extracted based on the density of photons in the vertical direction, estimated by applying a Kernel Density Estimation (KDE) function along a moving window of 100 adjacent photons. The second part of the method is to interpolate the resulting points into a continuous bathymetric surface via universal kriging, a geostatistical approach that takes into account the elevation at each point and the distance between them to produce a gridded output of the estimated interpolated surface and the confidence in the interpolation. The third step is to combine the interpolated surface with a prior estimate, in this case the global bathymetry dataset GEBCO, via a Kalman filter. The Kalman filter is a bayesian approach that takes into account the new estimate and the uncertainty. Areas of the interpolated data with high certainty (i.e., nearer to ICESat-2 point measurements) will have a larger impact on the prior estimate,

whereas areas with a lower certainty will have a correspondingly lower magnitude of impact on the prior depth estimate.

To test the feasibility of the kriging and Kalman updating approach, a synthetic experiment is created using a one of validation data sets. Random point samples from the validation data are taken, and these point samples are used as input to the kriging and kalman updating process. For the study site evaluated here, it is found that by kriging these input points and then combining them with GEBCO data, the error of the combined product is lower than either the error of GEBCO or the kriging surface. The Kriging + Kalman updating approach is also validated using a subset of the JarKus dataset, a series of point measurements of bathymetry surveyed annually along the entire Dutch coast. These data points were used as input to the kriging process, and the output of the kriging process was combined with the prior estimate of the bathymetry (i.e, GEBCO) using the Kalman filter step. The combination of Kriging and Bayesian combination with existing data improves the estimate than either dataset alone, showing that the Bayesian combination approach can add value compared to existing data.

The entire processing chain is then tested at 4 test sites. The proposed signal extraction technique produced bathymetric point estimates with an RMSE between 0.54 m and 9.53 m at the various sites. The sites with the largest magnitude errors are due to errors in the photon filtering approach, which could be corrected by improving the land mask data. If the subsurface filtering errors are removed, the highest RMSE at a site is 2.42 m. After extracting the bathymetry points and applying the kriging and kalman updating techniques to produce the final output, a gridded product at higher resolution, some sites show a reduction in RMSE of up to 34% compared to GEBCO. However, a few sites also show an increase in RMSE. This is due primarily to uneven and anisotropic distributions of ICESat-2 point bathymetry estimates at some sites; kriging depends on estimating trends in data, and if the points are not well-distributed in the site it results in a lower quality of the universal kriging interpolation surface.

To implement this approach at a global scale, the computational efficiency of the processing chain must be maximized. One way to increase the efficiency is to apply the KDE signal extraction step only to ICESat-2 transects that are the most likely to contain good quality bathymetric signal. Therefore, several ways of pre-selecting the most likely transects to provide good bathymetry signal are investigated. Because high water clarity is required for the lidar signal reach and be returned by the sea bed, the estimated Secchi depth along the ICESat-2 transect was found and compared to the results. No clear relationship between the Secchi depth and the RMSE of the lidar data along a transect was found. Several other transect metadata variables were also checked for their relationship with the transect RMSE; the mean fraction of fully saturated photons, the density of photons along the transect, and the percent of ocean surface photons classified as 'high-confidence' in the default photon classification provided by NASA. However, none of the variables were found to provide a clear way to filter out high-error transects while retaining the low-error transects.

The proposed processing chain has the potential to increase the horizontal resolution of GEBCO data from 500 m to 50 m while also decreasing the vertical RMS error by up to 35% without requiring any manually-surveyed data. This result was achieved using the exact same input parameters to the processing chain for every study site; it is likely that the parameter choice could be further optimized to produce better results. The computational bottlenecks in the approach are the KDE signal finding step and the Kriging step. Currently, the approach is practical to implement on consumer computers with a study site of up to 500 km^2, but finding ways to pre-select the best transects could allow for significant efficiency gains in the KDE signal finding process. The approach is also limited by the need for exceptionally clear water to extract lidar bathymetric signal. Due to these characteristics, the processing chain proposed here could provide the basis for a global coral reef bathymetry data set.

List of Figures

List of Tables

Contents

1 Introduction

Thirty-nine percent of humans live within 100 km of a coast, and many of the world's largest cities are located in coastal areas (Magdalena et al. 2021). Further, 10% of all human beings live in zones less than 10 m above sea level, an area known as the Low Elevation Coastal Zone (LECZ) (B. Neumann et al. 2015; Lichter et al. 2010). These economically and socially important areas are under significant stress due to anthropogenic climate change and the consequent sea level rise (SLR). Changing climate patterns will likely create more intense storms, which combined with a higher sea level boundary condition due to SLR, puts these populations and infrastructure investments at significant risk. Additionally, climate change can increase water stress and result in increases in groundwater abstraction, leading to subsidence of coastal areas. Climate change presents severe, even existential risks for coastal zones and those who inhabit them.

To mitigate these risks economically, substantial research into coastal protection strategies and climate adaptation measures is required. Scientific research will help fully quantify the risks, allow better prioritization of mitigation efforts, and help develop strategies for climate adaptation and coastal protection that work in harmony with nature. A critical variable for this research is nearshore bathymetric data (Holman et al. 2013), which allows mapping of the spatial extents of flooding and better quantification of the wave transformation on the coasts.

However, this critical data is often missing in most of the world, due to the complication and expense of surveying it. This project aims to capitalize on recent advances in remote sensing technology to explore ways of improving current bathymetry estimates without requiring any in-situ data.

1.1. Motivation and Relevance

1.1.1. Importance of Nearshore Bathymetric Data

Bathymetric data is essential for many aspects of the blue economy, including aquaculture, marine energy, submarine cables, dredging operations, design of sea defenses, navigation, ecosystem preservation, and as a boundary condition for numerical studies of wave transformation. (Cesbron et al. 2021; Ashphaq et al. 2021). However, the nearshore zone is notoriously difficult to survey. There is currently a global lack of data in the 5–10 m zone (Albright and Glennie 2021) and up to 50% of the world's shallow coastal zones remain unsurveyed (IHO/OHI 2022). Where surveys do exist, they can be decades out of date, especially in the 40% of the world's coasts that are sandy and highly dynamic (Almar et al. 2021).

There have been attempts at global bathymetric datasets, most notably the General Bathymetric Chart of the Ocean (GEBCO), which is an annually-updated global bathymetric and topographic dataset. The bathymetry in GEBCO grids is derived by assimilation of many different types of historic data, including acoustic soundings provided by ships and gravimetric bathymetry measurements (Cesbron et al. 2021). While it provides useful information about deep oceans, GEBCO accuracy in the nearshore

zone is limited because sonar data is limited in many shallow nearshore zones. (Monteys et al. 2015). Additionally, the horizontal and vertical resolution is not sufficient for accurate numerical studies of wave modeling. Bathymetric data with sufficient resolution for coastal wave modeling is available in most of the United States from NOAA, and most of the European Union via EMODnet. Outside these regions, there is little to no high-resolution bathymetric data.

1.1.2. Satellite-derived bathymetry

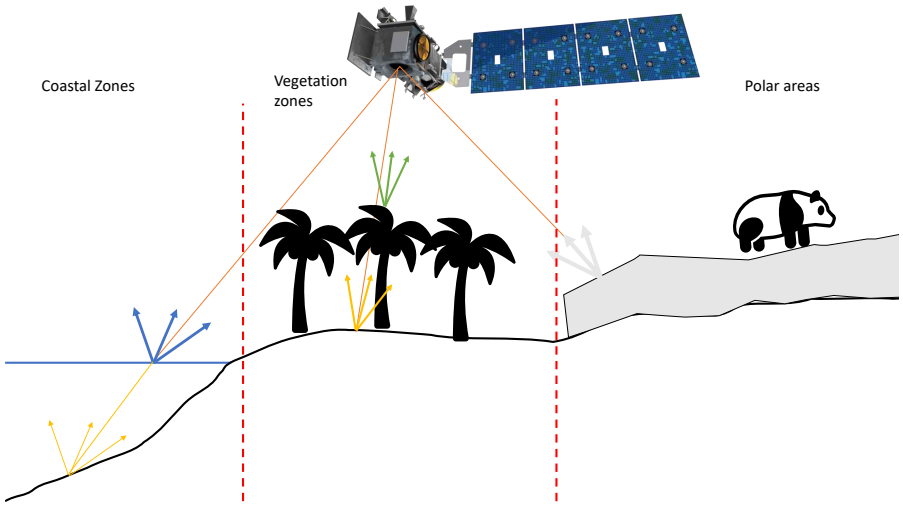
Because of the difficulty of directly measuring bathymetry, there have been many attempts to find ways to estimate bathymetry based on remote sensing satellite data. There are two broad types of Satellite-Derived bathymetry (SDB): Wave-kinematic, which estimates the wave field parameters for a site based on their satellite signature, then approximates the bathymetry based on the linear wave transform. These methods do not require water penetration. The other main method is optical approaches, which use the optical qualities of the water column to estimate the bathymetry. These require optically shallow water so that the sunlight can reach the seabed to be able to derive an accurate reading.

1.1.3. Spaceborne lidar

NASA launched the Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) mission in September 2018. The primary purpose of the mission is to study the height of the glaciers and ice sheets (i.e., the cryosphere), with a secondary purpose of collecting altimetry data of land, vegetation, and the sea surface (Markus et al. 2017). The satellite carries a lidar altimeter called the Advanced Topographic Laser Altimeter (ATLAS) which counts individual photon returns. The ability to count photons allows the instrument to continuously take high-density lidar measurements along 3 tracks approximately 3 km apart, while having power requirements low enough to be carried on a satellite (Popescu et al. 2018).

The geodetic location of each detected photon relative to the WGS-84 ellipsoid is calculated based on the roundtrip time of the photon's path, the GPS location of the satellite, and the pointing vector. Because the photon detector is very sensitive, significant noise photons from the sun also are detected. During processing, each photon is assigned a confidence of being signal or noise. This data is released as a level 2A product called ATL03 (T. A. Neumann, Martino, et al. 2019). This geolocated photon data is released for public download via the National Snow and Ice Data Center (NSIDC) Earthdata website.

Bathymetric measurements were not part of the mission design goals. However, after the start of the mission, researchers found ICESat-2 laser pulses can penetrate the water surface and returns from the seabed can be detected in optically shallow water (Parrish et al. 2019). This has provided a new source of high-accuracy bathymetric data. However, most coastal transects globally data do not contain any data from the seabed. The ability of the laser pulses to reach the seabed is affected by turbidity, atmospheric conditions, and instrument conditions. The best ways of identifying bathymetric data within the photon returns, and how to use these sparse point estimates to improve bathymetric datasets are still an area of active research.



1.2. Research Question

The primary question that this project intends to answer is:

How can spaceborne remote sensing data be combined with existing global datasets to improve estimates of nearshore bathymetry?

To answer this question, the following subquestions will be investigated:

1. How can ICESat-2 transects that contain bathymetry be identified algorithmically?
2. Once transects with bathymetry are found, how can nearshore subsurface photon returns that are located in the nearshore zone be extracted?
3. How can lidar photon return locations reflecting the seafloor be separated from background noise?
4. What is the potential to scale up bathymetry detection to a global scale to produce high-level processed bathymetry product using ICESat-2 data?
5. How can spaceborne remote sensing sources be used to improve existing global bathymetry datasets?
6. Under what conditions can remotely-sensed lidar data provide useful improvement on bathymetric data estimates?

2

Background

2.1. Remote sensing of bathymetry

There are several established methods for calculating bathymetric data from passive optical and SAR satellite data, and recent advances in cloud computing capabilities like Google Earth Engine (GEE) (Gorelick et al. 2017) make large catalogs of remote sensing data more accessible (Pike et al. 2019; Turner et al. 2021). The approaches can be broadly classified into *wave-kinematic* and *optical inversion* techniques.

2.1.1. Wave-Kinematic bathymetry

This approach uses the hydrodynamic properties of ocean surface to produce an estimate of the bathymetry in a 2D grid. Hydrodynamic variables such as wave celerity and wavenumber can be estimated from either optical or SAR satellite data. The bathymetry can then be calculated using the wave dispersion relation $c^2 = \frac{g}{k} \tanh kh$, which relates the wave celerity c to the wavenumber k and the depth h (Almar et al. 2021). The major advantage of this method is that it does not require light to penetrate the water column. Therefore, it is not limited by water column properties such as turbidity, which can be a significant limitation to optical SDB methods along many coastlines. The disadvantages of this approach are that the horizontal resolution is limited compared to optical methods, and that the maximum depths where this approach can be applied is limited by the wavelength: longer waves feel the bottom earlier, so bathymetry estimates in deeper areas are only possible on coasts exposed to swell waves, which have a longer wavelength (Almar et al. 2021).

2.1.2. Bathymetry from optical remote sensing

Optical remote sensing is a passive technique, as it detects light from the sun reflected by the earth. Since the 1970s many methods of estimating bathymetry based on the optical quality of the water have been found, all based on the physical principle that water attenuates light. Optical methods require that the water is *optically shallow*, or clear enough that light can reach the sea floor. This is a significant restriction, but in places where it is applicable it can provide continuous estimates of bathymetry with a resolution up to the resolution of the optical imagery. This means bathymetry grids at 10 m horizontal resolution using publicly available Sentinel-2 imagery, up to 0.5 m if commercial imagery is purchased (Babbel et al. 2021; Le Quilleuc et al. 2022; Pike et al. 2019). There are two broad types of algorithms for extracting bathymetric data from optical satellite imagery, analytical and empirical. Empirical models link the amount of attenuation of each pixel to in-situ depth measurements and derive a relationship between color and depth. One advantage is that empirical approaches are generally computationally inexpensive. Analytical or physics-based approaches require corrections for atmospheric and subsurface factors (Turner et al. 2021) and require more sophisticated computational capabilities to apply.

Empirical optical methods require some in-situ data to establish the relationship between the optical

properties and the depth in the area. Parrish et al. 2019 conclude that the most effective way to use ICESat-2 data for bathymetric estimation is to combine it with optical techniques. Because the ICESat-2 data provides relatively high accuracy point estimates, and optical methods allow the estimation of a 2D bathymetry grid but require in-situ point measurements for calibration, combining the two techniques provides a synergistic fusion of the strengths of both. There have since been several studies that employ this approach, and evaluate different techniques for correcting refraction, identifying bathymetry signal, and combining the lidar data with optical techniques (For examples of studies combining ICESat-2 with optical SDB methods, see Geyman and Maloof 2019; Pike et al. 2019; Ma et al. 2020; Z. Lee et al. 2021; Albright and Glennie 2021; Ranndal et al. 2021; Gleason et al. 2021; Thomas, Pertiwi, et al. 2021; Babbel et al. 2021; Hsu et al. 2021; Cao et al. 2021; Xie et al. 2021; Surisetty et al. 2022; Zhong et al. 2022; Zheng et al. 2022; Daly et al. 2022; Xu et al. 2022; L. A. Magruder et al. 2020; Thomas, B. Lee, et al. 2022; Le Quilleuc et al. 2022). Suffice it to say this topic has seen significant interest from researchers in recent years.

2.2. Lidar bathymetric Surveying

The earliest attempts to use Light Detecting and Ranging (lidar) to survey the coastal zone date back to the late 1960s. (Bailly et al. 2016). The technology has matured significantly since then and currently airborne lidar using a strong 532nm laser beam is a common technique for high accuracy bathymetric and topographic surveying. The downside of this technique is that it does not scale well to large areas because it requires expensive equipment, and extensive post-processing work is needed after the survey to produce a final product.

Recent advances in lidar technology have allowed the development of the photon-counting lidar, which requires significantly less energy to detect a return signal. These have allowed the practical application of constant lidar data collection in satellites. The use of spaceborne lidar is a more recent area of research, but some early results have shown that spaceborne lidar can find depths as deep as 38 m (Parrish et al. 2019).

The potential for bathymetric mapping using spaceborne laser observations has been noted since before the advent of the ICESat-2 mission. The predecessor mission carried a lidar instrument called the Geoscience Laser Altimeter System (GLAS). GLAS was a green-light laser intended for measuring atmospheric aerosols (Abshire et al. 2005). However, because of the laser architecture, GLAS was not able to penetrate the water column (Forfinski-Sarkozi and Parrish 2016). However, a prototype of ATLAS, called the Multiple Altimeter Beam Experimental lidar (MABEL) instrument was tested with high-altitude aircraft missions, allowing a simulation of the data that would be provided by ATLAS (Mcgill et al. 2013). Early experiments with MABEL showed good agreement between bathymetric measurements from MABEL and high-quality airborne reference data (Jasinski et al. 2016; Forfinski-Sarkozi and Parrish 2016).

To detect the relatively weak signal reflected off the seafloor, very clear water and ideal atmospheric and oceanic conditions are required. Therefore, only a fraction of transects contain bathymetric signal even for ideal sites. One of the most important considerations to use underwater lidar data is the method of finding bathymetric signal while rejecting noise. One known property of the bathymetric signal is that a relatively high density of returns in the vertical (Z) direction.

2.3. Geodetic reference systems for oceanography

The ICESat-2 photon heights are by default referenced to the WGS84 ellipsoid. To convert these to an approximate mean sea level, some geodetic correction is required. The *geoid* is defined as a level surface that is normal to the gravitational acceleration at every point. It is not just the surface of equal gravitational potential, but it also includes the centrifugal force due to a fixed local reference frame. The geoid can be defined as a level surface in a theoretical potential field W such that $g = -\nabla W$. More intuitively, it can also be defined as the surface where the ocean surface would settle if it was in static equilibrium. (Hughes and Bingham 2008).

The contours of the geoid can be estimated using satellite gravimetry, and from these measurements the shape of the geoid can be mathematically approximated by a sum of a series of spherical harmonic functions.

The surface of the actual geoid changes with over time due to deformation of the earth induced by the gravitational forcing from the Sun and the Moon. Because the gravitational forcing of both the sun and the moon has both periodic and non-periodic components, the time average of their effect is non-zero. This time-averaged deformation is called the *permanent tide*. For satellite altimetry, a reference system that changes with time can be inconvenient, so measured heights are often reported in a coordinate system that is referenced to the *tide-free geoid*. The tide-free geoid is a theoretical reference surface that removes both the periodic effect and the permanent tide (Mäkinen and Ihde 2009). It can be thought of as the gravitational potential of the Earth if the distance to the sun and moon was increased to infinity.

2.3.1. Inverted barometer effect and dynamic atmosphere correction

The instantaneous sea surface elevation is affected by the local weather and atmospheric conditions. The Inverted barometer (IB) effect refers to the effect of air pressure on the ocean height, or the local dynamic sea-surface topography (Robbins et al. 2022). Spatial gradients in air pressure due to local high and low pressure systems are equalized by gradients in the water surface height. When atmospheric pressure is low, the ocean level rises, and when atmospheric pressure increases it pushes the ocean surface down.

In the simplest formulation, the drop in ocean level is approximately 1 cm per additional mbar of atmospheric pressure: $D_h \approx 0.99484(P_i - P_{ref})$, where P_i is the instantaneous pressure in mbar and P_{ref} is an assumed average pressure (1013.3 mbar) (T. Neumann, Brenner, Hancock, Robbins, Saba, and Harbeck 2019)

In addition to surface height changes due to atmospheric pressure forcing, the local wind stress field also creates gradients in the sea surface height due to water mass momentum forcing. These variations are higher frequency compared to the pressure forcing that is captured by the IB correction.

The combined effect of the IB effect and local wind forcing can be modeled using a global wave model that is forced using weather data. The combined correction factor is called the Dynamic Atmospheric Correction (DAC). The DAC correction provided in the ATL03 data product is sourced from the MOG2D model (Le Provost et al. 1994), provided by AVISO. It is a global, barotropic, non-linear model of the world's oceans. It has a variable grid resolution to provide higher resolution data in coastal areas and enclosed seas. The MOG2D model is forced using the wind velocity and atmospheric pressure at 10 m height from the European Centre for Medium-Range Weather Forecasts (ECMWF) global weather model. The MOG2D model provides the DAC correction factor with a 6 hour temporal resolution and a 0.25° horizontal resolution. These DAC values are interpolated in space and time to the nearest reference photon for inclusion in the ATL03 data product. The magnitude of the DAC correction factor is $\pm 50cm$ in the vertical direction.

2.4. ICESat-2

The ICESat-2 mission is intended to gather high resolution topographic data on a global scale. The satellite carries the Advanced Topographic Laser Altimeter System (ATLAS). ATLAS is a highly sensitive photon-counting, green-light lidar. The satellite instrument points at Reference Ground Tracks (RGT) along the earth's surface, and returns with a repeat time of 91 days. Along the reference track, there are 3 beams, one pointing directly at the reference track, and two that are offset by approximately 3 km on either side. Each of these three beams is further split into a weak and strong. The layout of the beams relative to the RGT are shown in Figure 2.1, and the orientation of the spacecraft in 3D is shown in Figure 2.2

Each of the 3 beams is split into a strong and weak beam, with the strong beam being approximately 4x more powerful (T. A. Neumann, Martino, et al. 2019). With each laser pulse the instrument emits 10^{14} photons which travel to the earth where they are reflected. For more reflective surfaces like snow and ice, up to 10 make it back to the sensor and are detected. For less reflective surfaces such as the

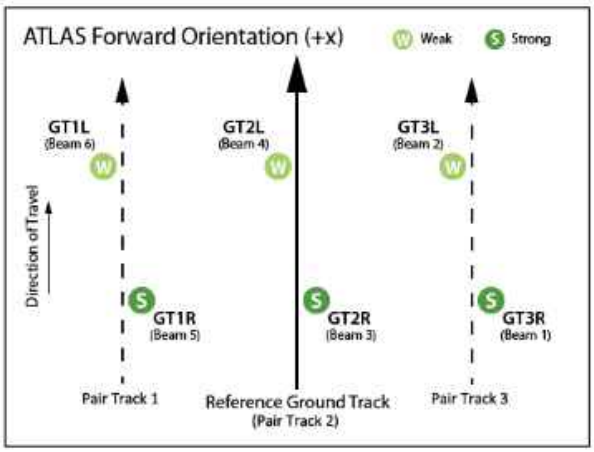


Figure 2.1: Layout of the ICESat-2 beams relative to the Reference Ground Track

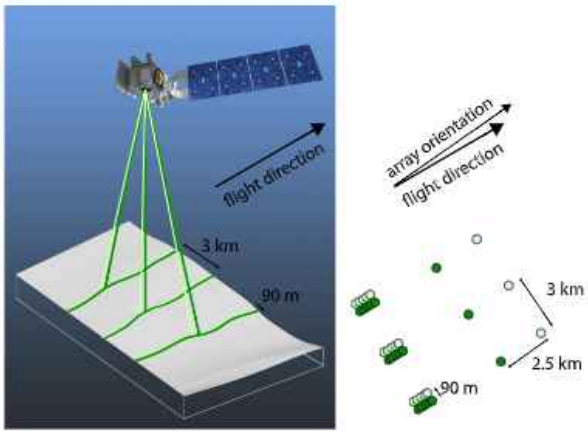


Figure 2.2: The layout of the ICESat-2 beams in 3D space. From T. A. Neumann, Martino, et al. 2019

open ocean, only 0–4 photons are detected at the satellite (T. A. Neumann, Martino, et al. 2019). The exact number of emitted photons that return to the sensor and are detected at the sensor depends on the local atmospheric conditions and the reflectivity of the surface (T. Neumann, Brenner, Hancock, Robbins, Saba, and Harbeck 2019). The highly sensitive instrument also receives significant noise, due to atmospheric scattering and photons from the sun. In clear water, some photons can travel to the bottom of the water and return. In water with a high suspended sediment concentration, these photons are more likely to be reflected higher in the water column (Ranndal et al. 2021).

The main mission of the satellite is to gather data about mass and elevation changes in ice sheets and glaciers, and to study global canopy height (Markus et al. 2017). As part of the vegetation height mission, the observatory is sometimes pointed away from the reference ground track when flying over land. This increases the spatial density of the observations of vegetation height (Markus et al. 2017). This operation, called *offpointing*, begins before the satellite begins to record data over land, so the nearshore coastal area is also included, and therefore bathymetric and hydrological applications also benefit from increased spatial coverage (L. Magruder et al. 2021).

To locate the position of each photon in 3D space, the time of flight of the photon is calculated with a precision of 800 ps (T. A. Neumann, Martino, et al. 2019). The location of the center of mass of the instrument is found using Global Positioning System (GPS) systems onboard the satellite. By combining the measured time of flight and satellite position and attitude, the geolocation of each returning photon is calculated (T. A. Neumann, Martino, et al. 2019).

Figure 2.3 shows the various reference systems used in calculating the height of an individual photon in the ICESat-2 data production. The location of the satellite is resolved in ellipsoidal coordinates using the Precision Orbit Determination (POD) module. Then, using the travel time between the photon and the satellite, the location of the photon is calculated using the speed of light and corrections for atmospheric delays, instrument bias, and wave height. The ellipsoidal height of a photon located directly below the instrument nadir is calculated by the difference:

$$H_{satellite_{surface}} = c(t_T - t_R)/2 \tag{2.1}$$

$$H_p = H_{satellite_{ellipsoid}} - H_{satellite_{surface}} \tag{2.2}$$

Where:

- H_{photon} is the height of the photon above the ellipsoid
- $H_{satellite_{ellipsoid}}$ is the height of the satellite above the ellipsoid (the brown line in Fig. 2.3)
- $H_{satellite_{surface}}$ is the height of the satellite above the surface (the green line in 2.3)

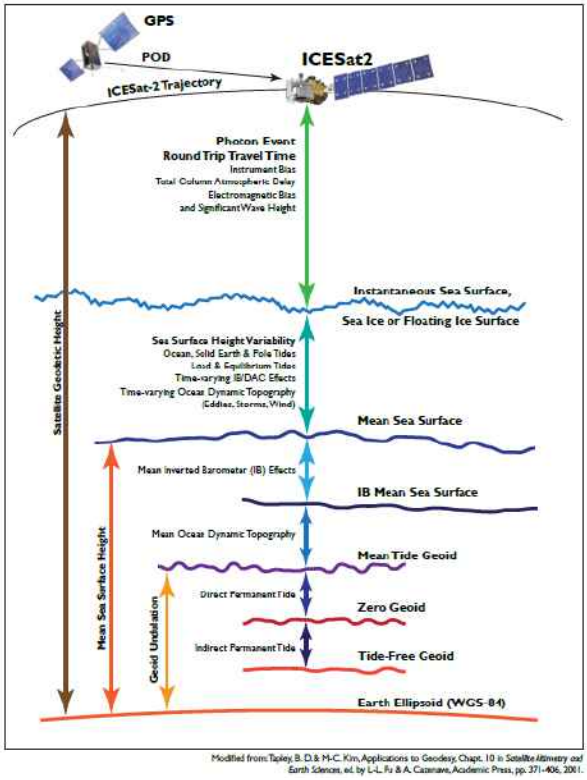


Figure 2.3: Relative levels of the reference planes

- c is the speed of light
- t_T and t_R are the photon transmit and receive times respectively

This gives the height of a photon directly at the nadir of the satellite; since photon are actually collected along 6 different beams at various angles, the actual X,Y,Z location of the photon is triangulated based on the angle of the beam.

The ATL03 data product reports the photon heights relative to the WGS84 reference ellipsoid. These ellipsoidal heights already include corrections for the solid earth tides, ocean loading, ocean pole tides, and atmospheric delays.

The height provided in the ATL03 data product is calculated by equation:

$$H_{GC} = H_P - H_{OPT} - H_{OL} - H_{SEPT} - H_{SET} - H_{TCA} \tag{2.3}$$

Where:

- H_{GC} is the geophysically corrected photon height above the WGS84 ellipsoid
- H_P is the uncorrected photon height above the WGS84 ellipsoid, from equation 2.1
- H_{OPT} is the height of the ocean pole tide correction, which is the deformation at the poles induces by centrifugal effects of variations in the earth's rotational axis.
- H_{OL} is the height of the ocean loading tide, which is the deformation of earth's crust due to the weight of ocean tides
- H_{SEPT} is the height of the solid earth pole tide

- H_{SET} is the solid earth tide, the deformation of the earth's crust due to the gravitational effects of the sun and the moon
- H_{TCA} is the height of the total column atmospheric delay, which is the effect of atmospheric conditions the photon travel time.

H_{GC} is the height above the WGS84 ellipsoid. To convert this ellipsoidal height into a value that can be used for computing the bathymetry, it needs to be converted into an instantaneous sea surface height. The equation that relates these is:

$$H_{inst} = H_{GC} - H_{geoid} - H_{free2mean} - H_{DAC} - H_{tide} \tag{2.4}$$

Where:

- H_{geoid} is the height of the tide-free geoid above the reference ellipsoid
- $H_{free2mean}$ is the difference between the tide-free and the mean-tide geoid
- H_{DAC} is the dynamic atmospheric correction factor which includes periodic effects like local wind forcing

2.4.1. ICESat-2 data processing levels

ICESat-2 data is provided to the public at a number of different processing levels. NASA provides a number of processing chains that produce higher-level products for different uses. The raw data transmitted from the satellite is considered to be Level 0. The NASA NSIDC decompresses this level 0 data, orders it in time and formats it in the HDF5 file format, and it is made available as a level 1 product, the lowest-level data that is available to the public.

In the level 2 products, the photon flight times are corrected for atmospheric effects, and the actual photon locations are resolved, then geodetically transformed to coordinates in WGS-84 ellipsoidal units.

In accordance with the scientific mission of the satellite, the level 2 product is then transformed into a number of higher-level products, each for a specific use case. Level 3 is divided into two sublevels, 3A and 3B. 3A products are point elevations of Land Ice, Sea Ice, Land vegetation, inland water elevation and ocean surface elevation. These point measurement products are then further transformed into level 3B or gridded products, which aggregate the point measurements over time to a grid.

2.4.2. Weak vs. strong beams

The beams are divided into weak and strong signals to enhance the radiometric dynamic range. The strong beams are expected to provide better signal-noise ratios over low-reflectivity surfaces, like the ocean and seafloor, while the weak beams are better for capturing very high reflectivity surfaces like ice, which might otherwise saturate the sensor and do not provide usable measurements (T. A. Neumann, Martino, et al. 2019). The strong beams have been found to provide the better data for lidar bathymetry measurements, but the weak beams can still contain useful bathymetry data (Hsu et al. 2021).

2.4.3. Refraction correction

The locations calculated by the data products from the satellite do not correct for the refraction induced by the water column. The subsurface photons are geolocated as if they are not underwater. However, because the speed at which light travels is different in water than the atmosphere. This effect introduces both a horizontal and vertical error in the photon location, as shown in Figure 2.4.

When the instrument is pointed directly at the reference ground track, the laser beams point nearly directly at the satellite's nadir. When directly on-nadir, the additional horizontal error induced by

refraction is ~ 9 cm (Parrish et al. 2019), which for bathymetric purposes is negligible. Some studies of bathymetry extraction from ICESat-2 data only use data that is collected when the instrument is pointing on-nadir, and therefore only correct for the vertical error using Snell's law. This is referred to as *first order* refraction correction in the table summarizing previous research (Table 2.1).

However, by design ATLAS is capable of pointing up to 5° off-nadir, or ~ 43 km away from the RGT (L. Magruder et al. 2021). During offpointing, the horizontal error is more significant and must be corrected for accurate bathymetry (Parrish et al. 2019). Parrish et al. propose a method to correct for both horizontal and vertical error that is widely cited in later papers. The Parrish method assumes a flat water surface, but other studies have extended their method to include the effect of water slope or wave action on the refraction error (Ma et al. 2020; Zhang et al. 2022).

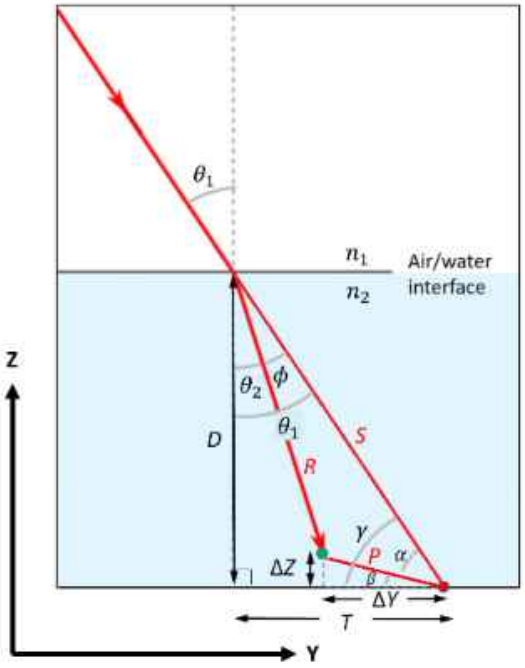


Figure 2.4: The geometry of the error induced by refraction at the air-water interface. From Parrish et al. 2019

The method proposed by Parrish et al. is outlined below. Figure 2.4 shows some of the geometric variables used in the equations. First, the angle of refraction θ_2 is calculated based on the relative speed of light in the air and water. This first step, shown in equation 2.5, is the application of Snell's law.

$$\theta_2 = \sin^{-1} \frac{n_1 \sin \theta_1}{n_2} \tag{2.5}$$

The relative length of R and S are described by the ratio:

$$R = \frac{S n_1}{n_2} \tag{2.6}$$

Then the horizontal and vertical distance length corrections ΔY and ΔZ can be calculated by the law of sines:

$$S = \frac{D}{\cos \theta_1} \tag{2.7}$$

$$\gamma = \frac{\pi}{2} - \theta_1 \tag{2.8}$$

$$\alpha = \sin^{-1} \frac{R \sin \phi}{P} \tag{2.9}$$

$$P = \sqrt{R^2 + S^2 - 2RS\cos(\theta_1 - \theta_2)} \tag{2.10}$$

$$\beta = \gamma - \alpha \tag{2.11}$$

The horizontal correction ΔY is along the direction of the laser pointing vector projected on to the earth's surface. To get the distance in local, rectangular coordinates that can used to correct the easting and northing coordinates in a local UTM system, ΔY is projected onto the local coordinate system using the azimuth of the laser pointing vector κ.

$$\Delta Y = P\cos\beta \tag{2.12}$$
$$\Delta Z = P\sin\beta \tag{2.13}$$
$$\Delta E = \Delta Y\sin\kappa \tag{2.14}$$
$$\Delta N = \Delta Y\cos\kappa \tag{2.15}$$

The curvature of the earth can also affect the accuracy of the refraction correction. For longer transects this effect can be corrected with an additional correction suggested by Parrish et al. The equation for this correction factor is

$$\delta\theta_{EC} = \arctan\frac{H\tan\theta_1}{R_e} \tag{2.16}$$

2.4.4. Extraction of bathymetric signal from photon data

The ATL03 data product includes a calculated confidence that a given photon return is signal or noise. The probability is assigned for each of the 5 surface types. Because bathymetric survey was not part of the original mission scope, there is no pre-calculated classification for subsurface returns. Therefore within the default classification bathymetric photons are often classified as noise. However, the classification of ocean surface classification is mostly reliable and can be used to filter out points near or above the sea surface (Ranndal et al. 2021).

To find bathymetric signal, a separate algorithm specifically calibrated to distinguish bathymetric signal from noise photons is required. There are several different techniques proposed in the literature. Some early research on small sites used manual classification (Forfinski-Sarkozi and Parrish 2016; Thomas, Pertiwi, et al. 2021; Babbel et al. 2021; Albright and Glennie 2021).

Other researchers have used implementations of the density-based spatial clustering for applications with Noise (DBSCAN) algorithm (Ester et al. 1996), with parameters that are set adaptively based on the local density of returns, called *adaptive DBSCAN* in the literature (Ma et al. 2020; Xie et al. 2021; Le Quilleuc et al. 2022; Liu et al. 2021). Thomas, B. Lee, et al. created an algorithm called C-SHELPh, which uses a binned histogram approach to predict the location of the sea surface. Datta and Wouters 2021 proposed a method based on a kernel density estimation function. All of these automated classification approaches rely on the fact that when there is bathymetric signal present, the density of returns in the vertical direction is higher than in the surrounding area (Neuenschwander and L. A. Magruder 2019).

2.4.5. Summary of Prior Research on ICESat-2 derived bathymetry

Previous studies that have used bathymetry data extracted from ICESat-2 data are summarized in table 2.1.

Paper	Year	Refraction Correction Method	S/N Classification method
Parrish et al.	2019	Parrish method	Manual
Ma et al.	2020	Parrish + sloping sea surface	Adaptive DBSCAN
Thomas, Pertiwi, et al.	2020	Parrish Method	Manual
Albright and Glennie	2021	First-order	Manual
Babbel et al.	2021	Parrish	Manual
Xie et al.	2021	Parrish Method	Adaptive DBSCAN
Cao et al.	2021	First-order depth correction	A-DRAGANN
Z. Lee et al.	2021	Not specified	Not specified
Liu et al.	2021	Liu method	Adaptive DBSCAN
Coveney et al.	2021	First-order	Manual
Datta and Wouters	2021	Parrish	Watta
Le Quilleuc et al.	2022	Parrish	DBSCAN
Thomas, B. Lee, et al.	2022	Parrish	C-SHELPh

Table 2.1: Summary of prior studies that extract bathymetric data from ICESat-2 photons

3

Data Sources and Methods

This chapter lays out the approach to the project, and the data sources that were used as input.

3.1. Data Sources

The following section details the data sources were used as input to the processing chain.

3.1.1. NASA ATL03 global geolocated photon data

The main source of data is NASA's ATL03 V005 data product (T. Neumann, Brenner, Hancock, Robbins, Saba, Harbeck, et al. 2021). ATL03 is a level 1 data product that consists of the precise latitude, longitude, and elevation for each received photon. As a level 1 data product, it has already undergone some processing by the Atlas Science Algorithm Software (ASAS) to correct for instrument errors, to classify photons as likely signal or noise for different surface types, and to correct for some geophysical effects including earth tides to provide measurements relative to the WGS-84 ellipsoid.

Additionally, the data includes variables that allow for further corrections and adjustments to change the height reference from the ellipsoidal height. These additional variables include correction factors for the tide, ocean surface depression due to atmospheric pressure, and factors to convert the ellipsoidal elevation to a height relative to the tide-free geoid.

The data is available via the National Snow and Ice Data Center Distributed Active Archive Center (NSIDC DAAC) via an Application Programming interface (API). The data is available divided into *granules* that include several seconds worth of observations, or approximately 1/14th of an orbit Lu et al. 2021. The API allows searching for data based on spatial location and date of acquisition. Once the search has been completed, the resulting granules of ATL03 data can be downloaded. Before downloading, the data be further subset by spatial location, so that only photons within a geographic area are included. The granules can also be subset by variables, so that only specific variables of interest are included in the resulting download. This can significantly reduce the required download size if only a few variables are needed.

3.1.2. GEBCO Global Grid 2021

The General Bathymetric Chart of the Ocean (GEBCO) is a global grid of topography and bathymetry at a 30 arc-second mesh resolution (GEBCO Bathymetric Compilation Group 2021 2021). GEBCO is assembled by compiling many different data sources, including multibeam sonar data, nautical charts, and satellite gravimetric measurements for deep-ocean bathymetry (IHO 2019). The elevation data is referenced to a vaguely-defined 'mean sea level'. The various data sets included in GEBCO are all

assumed to be referenced to MSL, but some datasets referenced to chart datum are included.

GEBCO has limited accuracy and resolution, but it is the only available data source in many places in the world, so it is sometimes used as the best-estimate in very data-poor sites. However, the accuracy of GEBCO varies depending on the input data sources. In this project, the GEBCO elevation is used both to filter locations that *may* contain valid bathymetry, and used as a prior guess to the Bayesian updating approach.

The GEBCO 2021 data includes a grid of the elevation values, and an additional metadata grid called the Type IDentifier grid (TID) that shows the source of the GEBCO input data (e.g., singlebeam sonar, multibeam sonar, satellite altimetry, nautical charts) for each grid cell. Both datasets were downloaded in NetCDF4 format.

3.1.3. GlobColour Daily Secchi disk depth data

To investigate the relationship between the water clarity and the availability and quality of the bathymetric data from the spaceborne lidar, ocean color data data from Garnesson et al.(2019) is also linked to each transect. This dataset estimates the Secchi disk depth every day at a 4 km grid resolution across the entire world's ocean. Because this data is derived from optical remote sensing, the temporal resolution is sometimes limited by cloud cover over the area of interest on a certain day. When direct observations are not available, the ocean color is interpolated to create a daily product (Garnesson et al. 2019).

The data is accessed via the OPeNDAP protocol on a server hosted by the Copernicus Marine Service. The Xarray python package (Hoyer, Roos, et al. 2022; Hoyer and Hamman 2017) is used to subset and download the data.

3.2. Methodology

To reach the end goal of incorporating ICESat-2 into GEBCO grids, first the lidar photon data for the area of interest is downloaded, processed into geoidal heights, subset to only include subsurface photons, find bathymetric signal within the subsurface data, interpolate into a 2D grid, then finally combine the interpolated ICESat-2 data with the resampled GEBCO grid within the area of interest. Then, for each test site the change in root mean square error (RMSE), Mean Absolute Error (MAE), and mean error (ME) when compared to the raw GEBCO bathymetry is calculated.

3.2.1. Processing ATL03

To download data for a site, a geographic Area of Interest (AOI) is defined first. This area is passed to the NSIDC download API to obtain only the data within the requested spatial subset in HDF5 format. By spatial subsetting of the data the download side is significantly reduced compared to downloading an entire granule. The NSIDC API also allows subsetting by the variable name, so only the ATL03 variables that are relevant for this research are downloaded, which further reduces the file size for practical download and storage of the ATL03 data.

The three variables that define the 3D location of each photon in the WGS-84 ellipsoidal reference frame are *h_ph*, *lat_ph*, and *lon_ph*. They are located in the */heights/* group within the ATL03 data structure. To use these variables for the purpose of bathymetric measurement, several other variables are required for processing. To transform the ellipsoidal elevation to the geoidal elevation, two additive factors *geoid* and *geoid_free2mean* are included in the download. For ocean surface correction, the *dac* and *tide_ocean* are downloaded from the *geophys_corr* group.

These correction factors are not provided for every photon but are provided for each 20 m segment because they vary at scales longer than the nominal 0.7 m between each photon. To find the correct adjustment factors for each photon, we need to match the segment-rate variables to the photon rate variables. The spatial relationship between the reference photons is shown in Figure 3.2. All blue photons within the black box around the reference photon are assigned the value from the reference

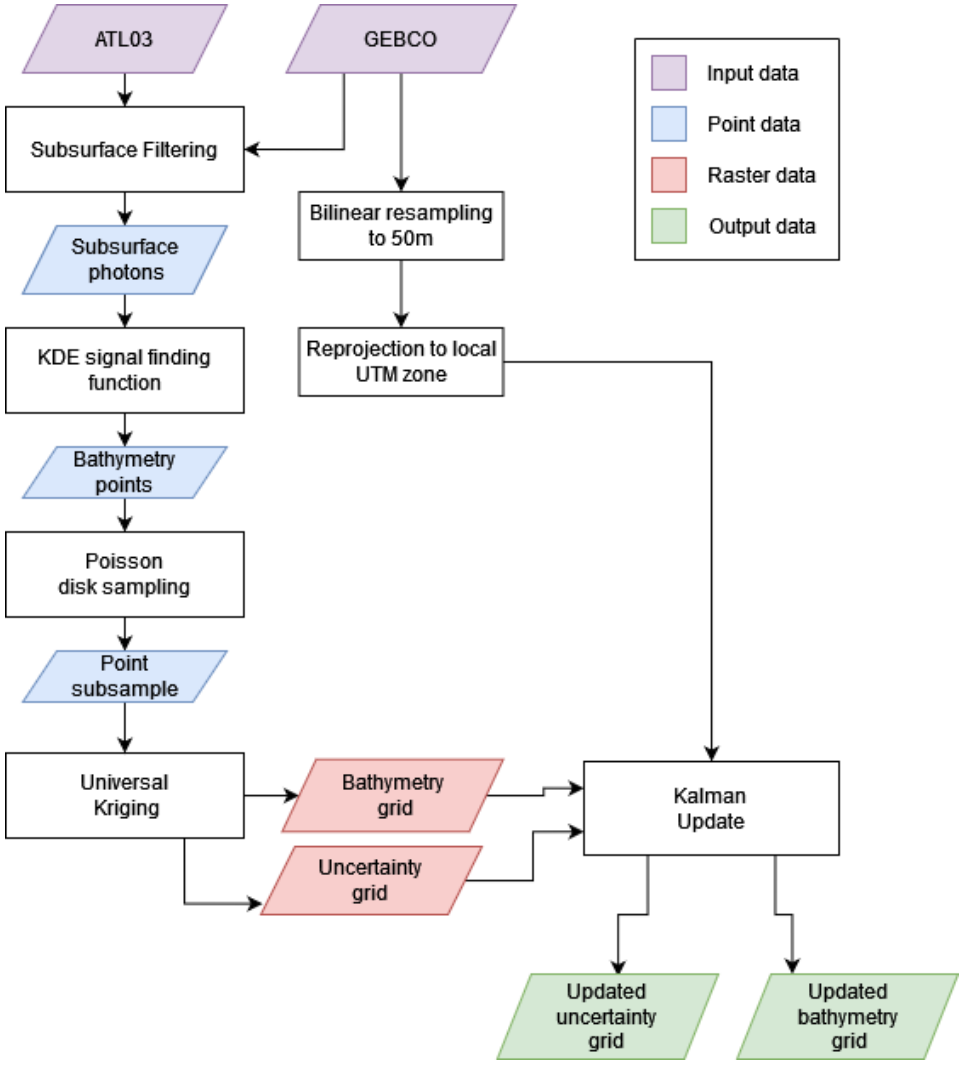


Figure 3.1: Overview of the basic methodology

photon. This matching is implemented using the python package Pandas (Reback et al. 2022; McKinney 2010) which has functions for joining time series data. For each regular photon, the appropriate value of the segment-rate variable is determined using the Pandas DataFrame *.asof()* method to find the closest segment rate variables in time to each photon.

3.2.2. Filtering ATL03 to subsurface returns

The bathymetric signal that we are seeking to find is located in the shallow-water nearshore zone. Therefore, photons outside that zone need to be removed to reduce processing time and eliminate false positives as much as possible. To reduce the downloaded transect data within the area of interest, the following filtering steps are applied to each transect:

1. First, horizontal filtering is applied to areas along the transect to remove photons in areas that are too deep to contain bathymetric signal or are located on land. For every photon, the GEBCO elevation at the latitude and longitude of the photon is found. Any photons with a GEBCO elevation between -40 m and 3 m are selected, and those outside of this region are culled from the data set. Photons that are outside this range are assumed to be deeper than the maximum known depth detectable by ICESat-2 (38 m per Parrish et al.), or assumed to be on land. The results of this step can be seen in Figure 3.3. This figure shows a transect that crosses Marathon Key. The island

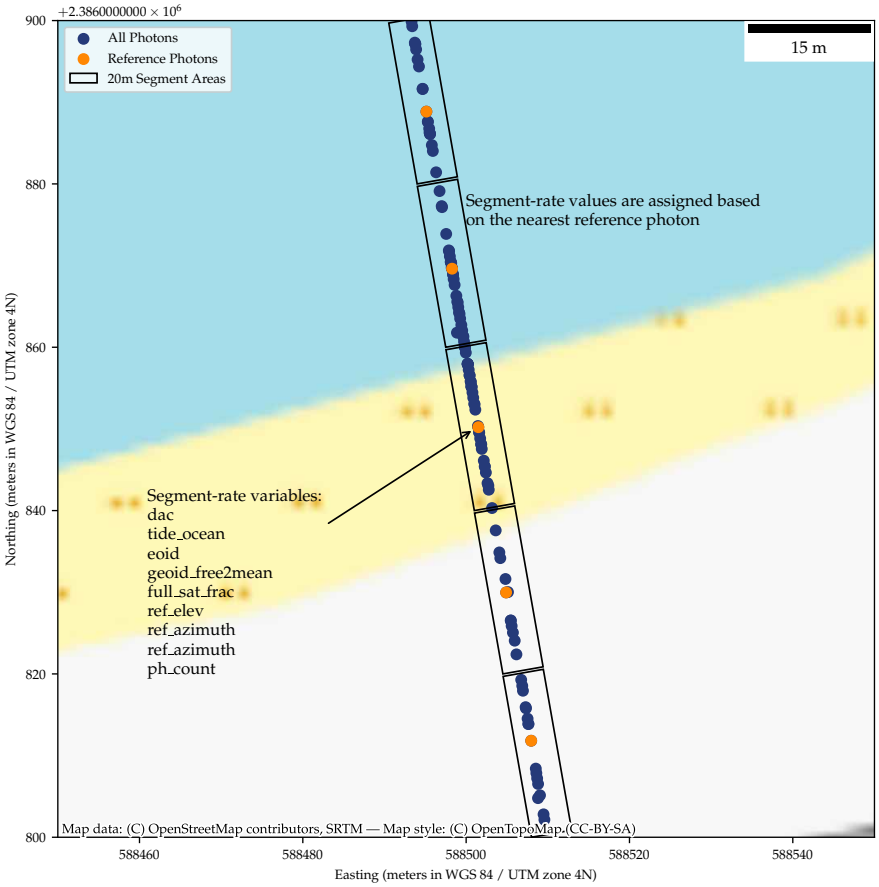


Figure 3.2: Relation between the regular photons and the reference photon locations

is located in the middle of the transect, and the points over the island are removed by this step.

2. The local sea-surface elevation h_{sea} is calculated by taking the median elevation of the high-confidence sea surface photons in the transect. The water depth for each photon is then calculated. The standard deviation of the elevation high-confidence photons $\sigma_{h_{sea}}$ is also calculated. This is used as a proxy to estimate the magnitude of the wave height at the time of the observation.

3. Any photons with a water depth greater than 40 meters or a geoidal height less than -40 m are removed, based on the same assumption that they are too deep to be bathymetric photons.

4. To remove the sea surface signal and any noise points above it, any photons that are higher than $h_{sea} - \max(2.5\sigma_{h_{sea}}, 1\ m)$ are removed. This removes the sea surface signal and any high noise photons.

   The results of steps 2–4 are shown in Figure 3.4

5. As an additional check, any photons with a geoid height greater than 10 m are removed. This is justifiable because it does make physical sense to find bathymetric photons that high above the geoid surface. This additional step is required because it was found that in Oahu there are some photons classified as high-confidence ocean signal with an elevation in the range 50-100 m. This is likely due to misclassification of cloud photons as ocean surface signal.

After these filtering steps, the resulting subsurface photons for this example transect are shown in Figure 3.5. The bathymetric signal can be seen clearly throughout the entire transect.

An overview of the entire filtering chain is shown in Figure 3.6

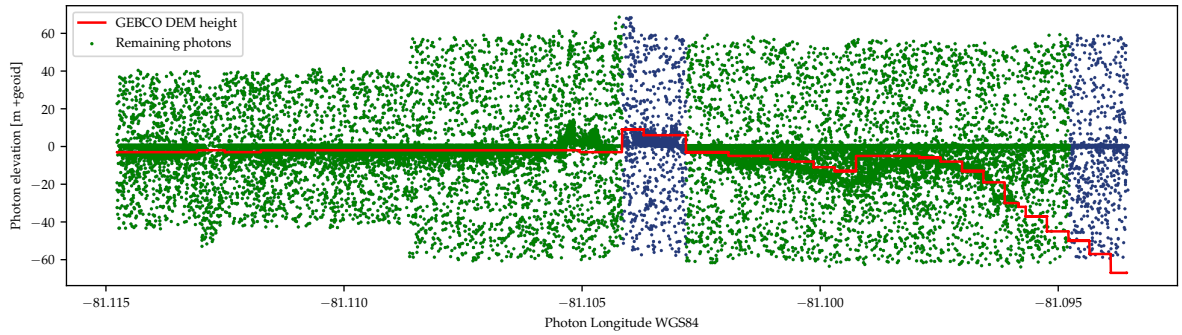


Figure 3.3: The GEBCO data for the example transect, and the photons that are removed due to the GEBCO depth

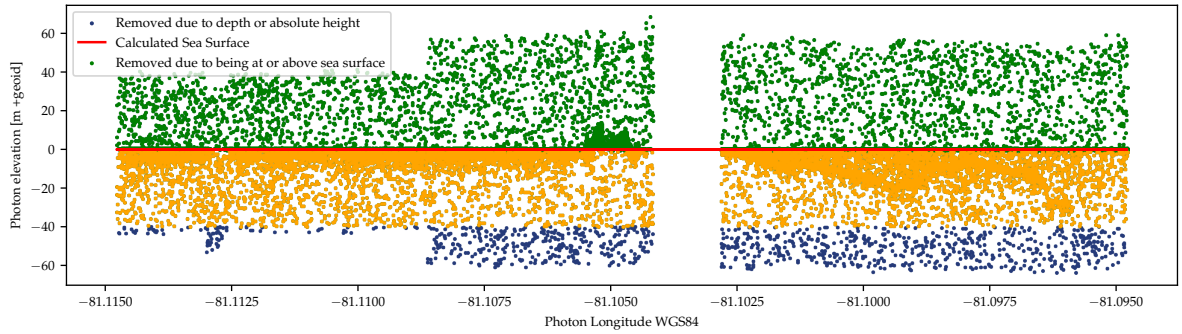


Figure 3.4: Vertical point filtering based on the local sea surface elevation

Finally, before the signal-finding is applied, the Parrish method of refraction correction, including the correction factor for curvature of the earth, is applied to all subsurface photons. Figure 3.7 shows the effect of the refraction correction step on the subsurface photons.

3.2.3. Bathymetric signal extraction

The filtering steps reduce the dataset to only photons that have been geolocated to the subsurface zone. However, there might not be any usable photon data in that zone. To determine if there is bathymetric signal present, further processing is required. Some proposed methods for separating bathymetric signal photons from noise are explained in section 2.4.4. These approaches are all based on the same principle, namely that an increased density of photon returns in an area means that it is more likely to be bathymetry signal. For this project, a method is proposed based on a gaussian Kernel Density Estimation (KDE) function, a non-parametric estimator of the local probability density.

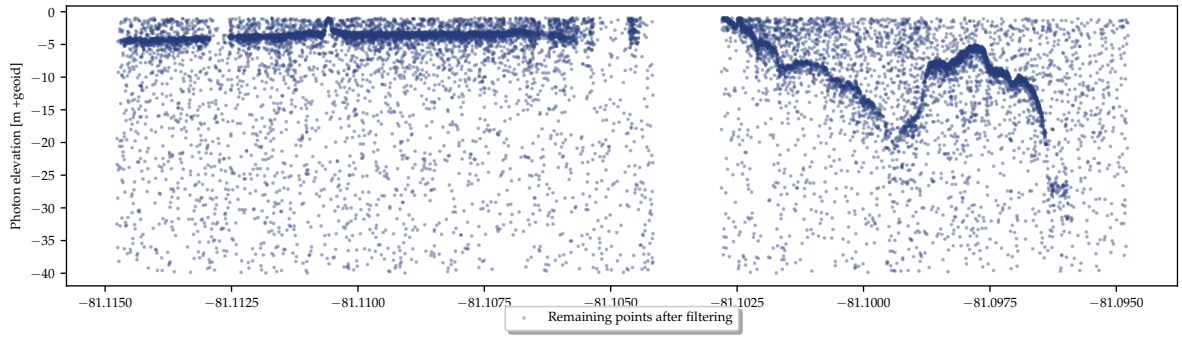


Figure 3.5: Subsurface photons found resulting after the filtering process

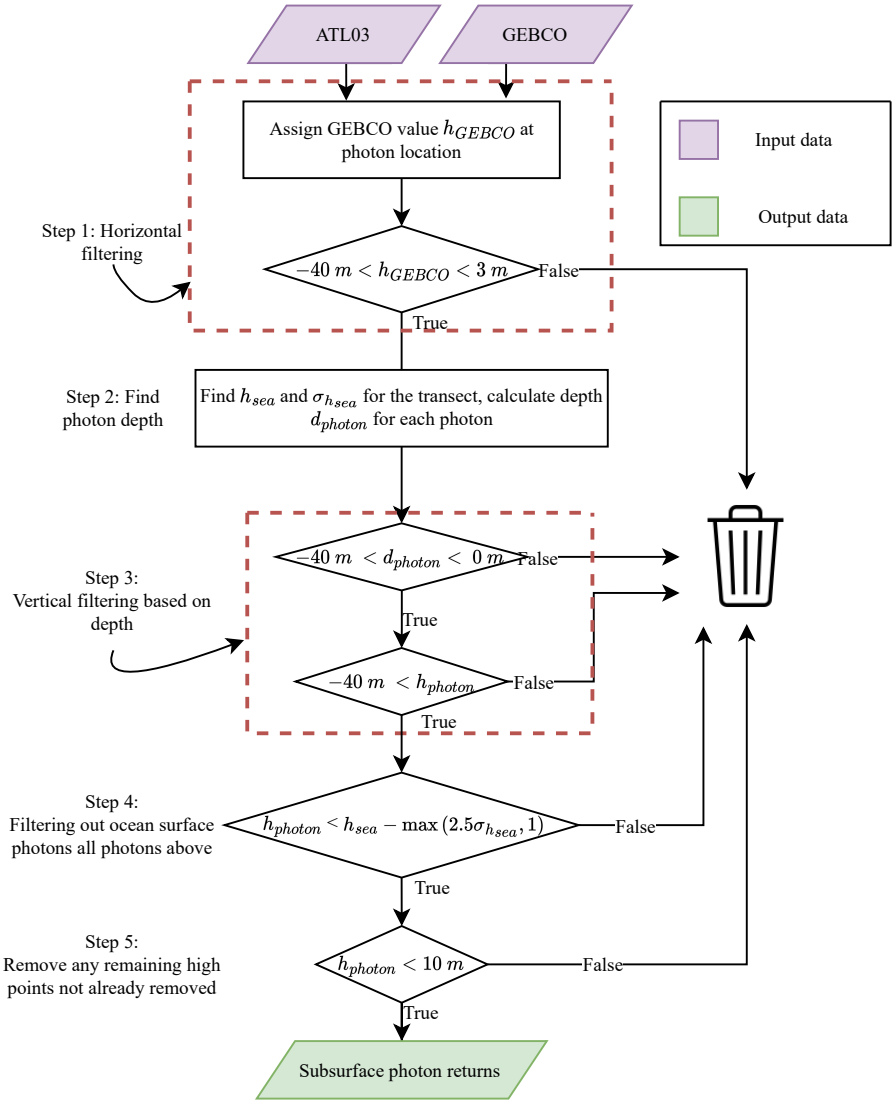


Figure 3.6: Subsurface filtering overview

The KDE function is highly influenced by the *bandwidth* parameter. For this implementation, the Scott method (Scott 2012) is used to estimate the KDE bandwidth:

$$n^{\frac{-1}{d+4}}$$

where n is the number of data points and d is the number of dimensions of the data.

A function is created that calculates the maximum kernel density, and the Z location at which it occurs.

$$f(\hat{z}_{window}) \rightarrow kde_{max}, z_{kde_{max}}$$

This function is then applied across a rolling window of the transect. The rolling window is applied across the time dimension to 100 adjacent points. Figure 3.8 shows the KDE function as applied to a single example window, and the resulting kernel density plot of that window.

This kernel density function gives an indication of the strength and location of the density peak for every photon in the transect. To reject locations where the signal is weak, a threshold (i.e., a minimum KDE density) needs to be established to distinguish between signal and noise. Because there can be significant variation in the density value between different transects, this is set adaptively based on either the median kde value kde_{50}, or kde_{min}, whichever is greater. The absolute minimum threshold of

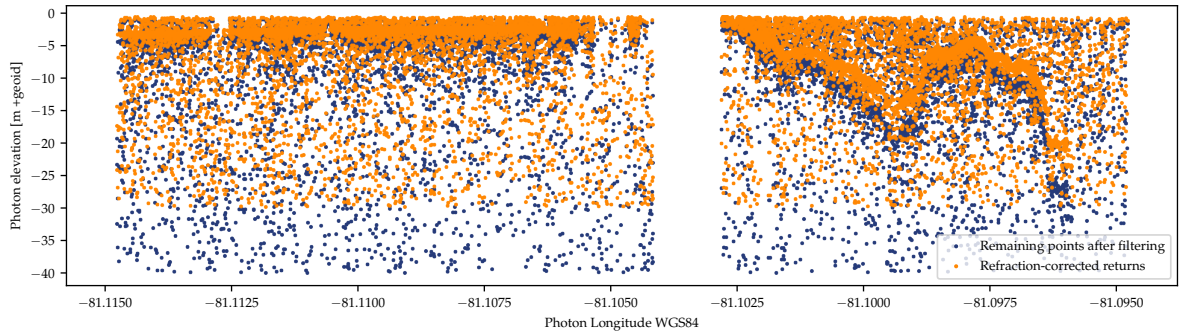


Figure 3.7: The refraction correction applied to the remaining photons

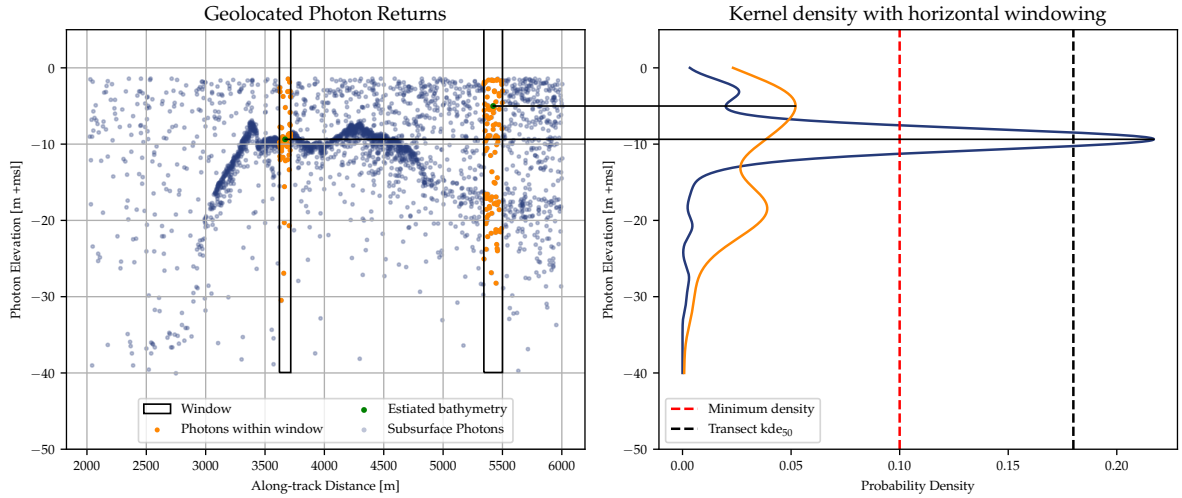


Figure 3.8: KDE function as applied to single window

kde_{min} is required because some transects contain no useful bathymetric signal. Setting an absolute minimum ensures these transects are more likely to be rejected as noise.

Any photons with a KDE value less than the threshold are assigned a NaN value and are dropped from the analysis and the remaining photons are saved for further use.

To summarize, the input parameters to the signal finding process are:

1. The size of the window in number of adjacent points
2. the minimum density value for the kernel density required for point to be considered signal

Based on empirical analysis of the test sites, a window of 100 adjacent photons and a minimum density of 0.1 were found to provide good results when compared to validation data.

3.2.4. Interpolation to a 2D grid

After the bathymetric signal points are identified using the approach described in section 3.2.3, the resulting bathymetry points are densely spaced along satellite tracklines, but are absent between them. A number of interpolation techniques have been applied to create bathymetry grids from point data. Commonly, inverse-distance weighting (IDW), tension splines, or loess interpolators have been used (IHO 2019; Ferreira et al. 2017). These approaches are relatively straightforward and easy to implement. However, the disadvantage of these simple approaches is that there is not a clear methodology of establishing the uncertainty of the resulting interpolation. Intuitively, areas with a higher point density

will lead to an interpolation with a lower uncertainty that areas with very sparse points. Kriging is a geostatistical technique that allows interpolating points into a grid, while also giving an indication of the uncertainty in the neighborhood. By knowing both our estimated sea floor depth and the uncertainty in the estimate, we can use a Bayesian approach to update our initial guess of the seafloor location.

Subsampling of bathymetric points using Poisson disk sampling

The bathymetric points are extremely densely spaced in some areas and transects. This can present an issue to kriging, which is computationally expensive and points too close together can make the algorithm even slower. To reduce the number of points fed into the algorithm, a subsample of the points is taken using the Poisson disk sampling technique. Poisson disk sampling is a strategy that generates random samples of a distribution that are a minimum distance apart. When applied to a point cloud, in this case the point cloud of all the bathymetric points, the sampling function will only return points that are a minimum distance apart in 3D space. This ensures that the sampling is as representative as possible of the underlying surface.

The implementation used is from the PDAL software library (Butler et al. 2022), which provides an iterative Poisson disk sampling function based on McCool and Fiume (1992). The disk sampling is applied iteratively with a decreasing radius until the desired number of points is reached. In this case, it was found that the maximum number of points that can practically be used in the kriging step is approximately 2000 points. Therefore, this was the number of points used to allow the maximum possible point density while using a consumer-grade laptop to run the algorithm.

Kriging interpolation

The subsample of points obtained from the Poisson disk sampling is then interpolated into a grid format using a universal kriging approach. This geostatistical technique results in a grid of the estimated depth as well as the estimated uncertainty.

The universal kriging interpolator used is a 2D interpolator, and both the estimated surface and its associated uncertainty are generated for the entire site. A single slice of the output, taken along an ICESat-2 transect, is shown in Figure 3.9. Note that the uncertainty estimate for the interpolated GEBCO surface is constant, and the uncertainty estimate for the kriging interpolator increases with increased distance from the input points.

A spherical variogram with a range of 10 km, a nugget variance of 0.7 m^2, and a sill variance of 25 m^2 was found to produce good results in most sites. The python package Pykrige (Murphy et al. 2021) was used to implement the universal kriging approach.

3.2.5. Bayesian data assimilation using Kalman update equation

The Kalman Filter is a mathematical technique to predict the state of systems based on uncertain measurements. It consists of a loop of two steps, a *time update* step which updates the position based on a measurement and a known measurement uncertainty, and a *measurement update* step which predicts the state based on the dynamic equations of the system. The Kalman filter equations are used to estimate the state of a linear process for a state x_k and a vector of measurements of the state z_k (Welch 2021). It is assumed that the state of a system evolves according to the difference equation

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \tag{3.1}$$

Where A is a matrix that relates the state at time step k to the state at the previous time step and w_k represents normally distributed process noise with a variance Q.

Measurements of the state are assumed to have the form:

$$z_k = Hx_k + v_k \tag{3.2}$$

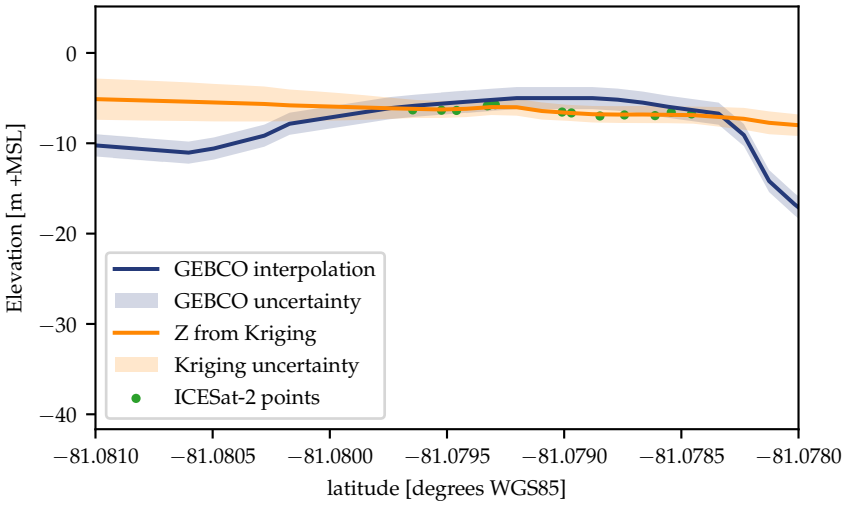


Figure 3.9: Example of the subsampled points, and the results of the universal kriging interpolation. In blue, the result of GEBCO is shown for comparison. Note that this figure shows a single 2D slice of a 3D field.

Where H is a matrix that relates the state to the measurement, and v_k is measurement noise which is assumed to be normally distributed with a variance R.

At each time step k, the estimates of the state can be updated according to the following two steps:

Time Update:

$$\hat{x}_{\bar{k}} = A\hat{x}_{k-1} + B\hat{u}_{k-1} \tag{3.3}$$

$$P_{\bar{k}} = AP_{k-1}A^T + Q \tag{3.4}$$

Measurement Update:

$$K = P_{\bar{k}}H^T(HP_{\bar{k}}H^T + R)^{-1} \tag{3.5}$$

$$\hat{x}_k = \hat{x}_{\bar{k}} + K(\hat{z}_k - H\hat{x}_{\bar{k}}) \tag{3.6}$$

$$P_k = (I - KH)P_{\bar{k}} \tag{3.7}$$

Where:

- $\hat{x}_k$ is the current state of the system
- $\bar{k}$ is the measurement matrix
- P_k is the error covariance matrix
- K is the Kalman gain

To apply this to this bathymetry estimation problem, some simplifying assumptions are made. The nearshore zone is a highly dynamic system. However, for the purposes of this project it is assumed that the temporal variations over the time scale being studied are within the margin of error of the measurements, so the bathymetry of the nearshore zone is assumed to be a static system and the time update equations 3.3 and 3.4.

Equation 3.5 is the Kalman gain, an estimation of the strength of the estimation. For this case, it is assumed that the matrix H is the identity matrix. Equations 3.5, 3.6, and 3.7 can be simplified to

$$K = \frac{P_k}{P_k + R}$$

$$\hat{x}_k = \hat{x}_{\bar{k}} + K(\hat{z}_k - \hat{x}_{\bar{k}})$$

$$P_k = (1 - K)P_{\bar{k}}$$

It is also assumed that all estimates are measurements of the same underlying physical depth, and that differences between measurements are due to normally distributed measurement error, with magnitude of the error varying depending on the method. To combine multiple measurements, the *measurement update* step is applied recursively for each available measurement to produce an updated estimate of the bathymetry grid. This is a Bayesian update, in that it considers the uncertainty of each measurement when combining them.

In this application, the starting point of estimate is based on a bilinear resampling of the GEBCO dataset, with an assumed variance of 1.5 m^2. The value was found empirically to give good results in most test sites. Then, the simplified Kalman equations above are applied. At each raster grid cell the kriging uncertainty is used as R, and the kriged estimate of the seafloor is used as $\hat{z}_k$. The result is a new estimate of the bathymetry that includes both the kriged lidar surface and a priori depth of GEBCO. Figure 3.10 shows both outputs of the kriging process, including the uncertainty bars for each. The plot on the left side shows the estimated probability density functions of the bathymetry surface along the black vertical line. The green PDF curve, labeled *new estimate* shows the resulting *a posteriori* estimate of the sea surface that is calculated using the Kalman update equation.

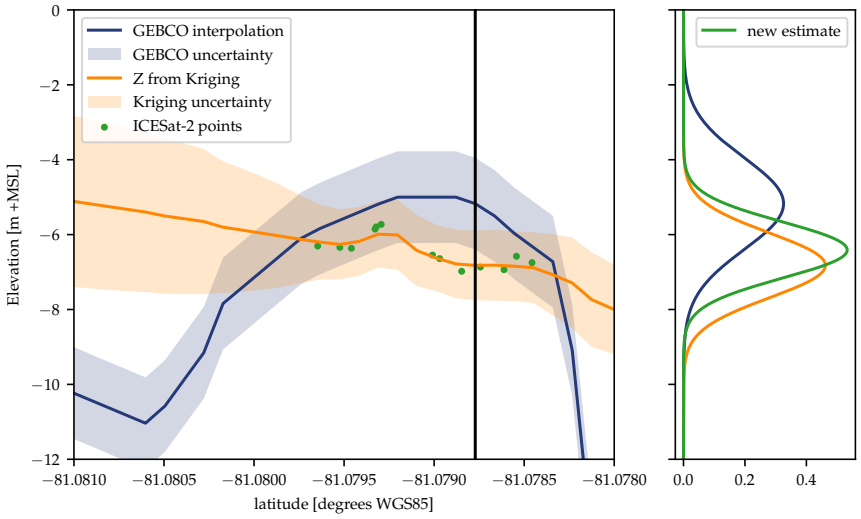


Figure 3.10: Using a section of the same transect in Figure 3.9, the results of the Kalman measurement update is shown. The right side of the plot shows the plots along the black section line on the left side of the plot

3.2.6. Per-track Secchi disk depth estimation

To evaluate the relationship between the Secchi disk depth at the location and the time of the transect, and the quality of the bathymetry estimate, the Secchi depth value from the GlobColour dataset (Garnesson et al. 2019) is found at points distributed along each ICESat-2 transect.

The GlobColour Daily Secchi Depth dataset is provided at a 4 km horizontal resolution, and a 1 day temporal resolution. To calculate the Secchi depth in water areas at the time of the satellite pass, the ATL03 tracklines are divided into points every 4 km starting at the end of the transect. For each of these points, the closest ocean color data in time and space is calculated using Xarray.

One issue is that this approach generates some points over land. The GlobColour data does include a land mask, but because of the 4 km horizontal resolution of the ocean color product, the land mask is not reliable along coasts. However, the GEBCO dataset has a much higher horizontal resolution than the GlobColour data (450 m). Therefore, in this case the GEBCO data can be used as a land mask. For each Secchi disk depth point, the corresponding GEBCO elevation is found, and if it is greater than 0 the point is assumed to be on land and is rejected.

This provides a dataset of the water clarity (as estimated by Secchi disk depth) data at the location and day of the satellite pass. Ocean color statistics can then be calculated for each unique trackline, or for a site as a whole.

3.2.7. Error evaluation

To compare the results to the validation data, several different error metrics are considered. Both the error between the validation data and the lidar data is checked to evaluate the performance of ICESat-2 measurements, and the signal finding algorithm, and the total decrease in error between GEBCO and the version of GEBCO that includes the interpolated ICESat-2 data.

The error metrics that are evaluated for both types of error are the root mean square error (RMSE) and the mean absolute error (MAE). The RMSE is calculated by taking the mean of the squared difference between the true value and the estimated value, and then taking the square root of this mean.

$$\text{RMSE} = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(\frac{x_i - y_i}{\sigma_i}\right)^2}$$

By squaring the error first, larger magnitude errors are given relatively more weight in the metric. Mean absolute error also gives an idea of the average deviation, but gives equal weight to all errors.

$$\text{MAE} = \sum_{i=1}^{D} |x_i - y_i|$$

To evaluate these errors between the Kalman-updated bathymetry grid and the validation data reprojection and resampling is required. The validation data has a horizontal resolution on the order of 1 m, while GEBCO has a horizontal resolution on the order of 450 m. Before taking the error, the data being compared is first reprojected into the native coordinate system and resolution of the validation data using GDAL (Rouault et al. 2022). Then, the RMSE and MAE can be calculated between each cell of the raster grid.

The coefficient of determination R^2 can evaluated between the ICESat-2 bathymetry estimates and the true elevation at the corresponding point. It gives an indication of how well a ICESat-2 bathymetric photon value without a known ground-truth elevation would be predicted by the approach. This metric is given by the formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \bar{y})^2} \tag{3.8}$$

3.3. Possible limitations

This section introduces some possible sources inherent limitations or error in the method, and how some of them might be mitigated.

3.3.1. Limitations in Photon Geolocation Accuracy

The pointing determination algorithm used by the satellite has a high vertical accuracy, but there is an inherent limitation on the horizontal accuracy. The current best estimate of the vertical accuracy is 0.17 cm, and the estimate of the x and y position uncertainty is 5 m (T. A. Neumann, Brenner, et al. 2019).

This uncertainty, and the horizontal refraction, are more likely second order effects. Because the kriging is used to create a product of 50 m resolution, any uncertainty introduced by this will be masked by the interpolation to a 50 m grid.

3.3.2. Errors in refraction correction

The refraction correction method used accounts for the additional horizontal error that is introduced by off-pointing. However, there are several other second-order effects that are not considered by the methodology. One of these is the estimation of the refractive index; the temperature and salinity affect the speed at which the water transmits light. By assuming a default value, it could introduce some error in sites that deviate significantly from the assumed default. This effect was deemed likely a second order effect and was not considered here. However, this could be corrected by either looking up the appropriate values for each site beforehand, or by connecting the algorithm to an API that can provide a temperature and salinity value for any site and time.

Another potential source of error is the slope of the water surface. Since there is a slope to the water surface, this affects the bounce angle of the photon. This can be corrected for and some papers that investigate ICESat-2 bathymetry have attempted to correct for it. For this project the magnitude of error due to slope was considered small enough to be within the margin of error of the method.

3.3.3. Misclassified photons in ATL03 data

Correct filtering of sea surface returns is very important to the accuracy of the bathymetry signal finding, because the sea surface signal is often several orders of magnitude more dense than the bathymetric signal, and the location of the sea surface is also used to calculate the depth of a photon for refraction correction. The default photon classification provided in the ATL03 data is used to identify the sea surface within the filtering algorithm. Errors in the default ocean signal classification can result in sea surface signal being inadvertently included in the subsurface data and biasing the KDE results.

The default classification is often reliable, but when there is a large area where the bathymetric surface is shallow and nearly parallel to the sea surface, there can be misclassifications. An example of this is shown in Figure 3.11. In this example, actual sea surface is not classified as a high confidence ocean surface return, and some areas that appear to be bathymetric signal are classified as ocean surface. This causes the bathymetric signal to be thrown out because it is incorrectly considered to be the sea surface.

This could potentially be mitigated by a different filtering strategy that calculates the local sea surface based on the local geoid. This could be very feasible in microtidal areas where the tidal signal has a smaller impact.

3.3.4. Limited spatial coverage in some islands

In support of the vegetation mission of ICESat-2, the instrument is sometimes pointed as much as several degrees to the side of the reference ground tracks when the satellite passes over land. This increases the spatial density of points at the expense of the temporal resolution. For bathymetric purposes the

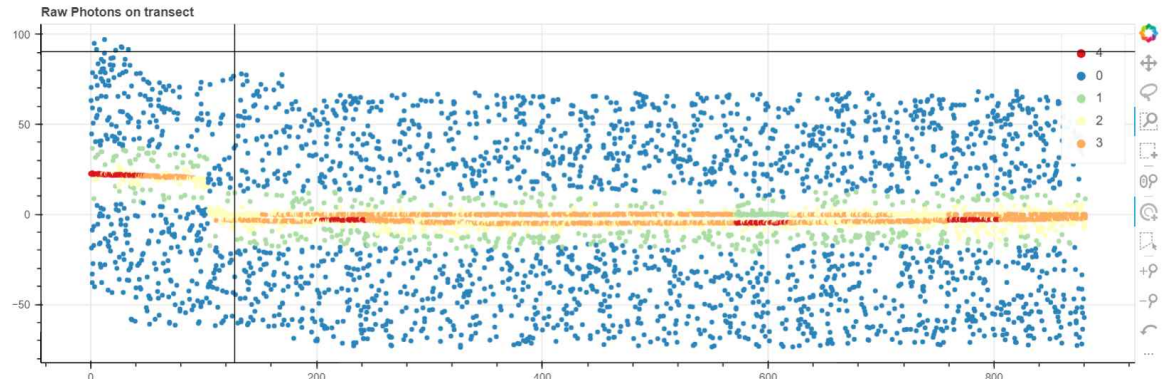


Figure 3.11: Classification of photons from 2021-07-19, Beam gt3r, reference ground track 396. This transect is located off the coast of Egypt. The two parallel straight lines from 200 to 800 are the sea surface and the bathymetric signal. The NASA photon classification algorithm misclassifies the bathymetric points as ocean surface returns

increased spatial resolution gives a more even coverage of nearshore zone bathymetry.

However, the land mask that is used by NASA to determine the off-pointing strategy has a limited resolution, and therefore some island nations do not benefit from the increased spatial density. This was noted when trying to collect data from Fiji and the Maldives. Due to apparently being located within the off-pointing zones, both of the aforementioned islands only have tracks which are 3 km apart. They can still potentially collect bathymetry data if conditions are otherwise good, but the further reduction in spatial coverage limits the accuracy of the kriging method. This is unfortunate because many of the states that are at the highest need of detailed bathymetry for numerical studies are big ocean island nations. The tradeoff for this scenario is that the temporal resolution is significantly better, so the spaceborne lidar could be useful for studying the changes over time.

3.3.5. Inherent uncertainty of KDE Method

There are a number of input parameters to the filtering and the density-based bathymetry finding methods. These parameters can be optimized for each site to reduce the RMS error as much as possible if there is some validation data available. However since the end goal of the project is to be able to improve estimates without using any in situ data, ideally there would be no need for optimization based on the site.

Currently the globally-set parameters are sufficient to extract bathymetry without any tuning for all of the case studies that are investigated. However, the inability to tune in advance is a limitation.

One possible future step would be to gather even more validation sites, and explore which other variables might influence the best parameter setting. It is possible that there are certain site variables which predict the optimal parameter options. Even so, upscaling of validation sites would allow better insight into which variables predict the presence of valid data.

3.3.6. ATL03 data quality issues

There are a number of known issues with the ICESat-2 data. They are either due to atmospheric and environmental conditions, or due to limitations of the instrument. Many of them can be detected in advance, and then the effected granule data can be thrown out or the issue otherwise corrected for. However, there might be some edge cases related to these issues that cause either false bathymetric signal points, or cause the algorithm to miss valid bathymetric data. These data issues could present an issue for the scaling up the signal finding without any manual intervention. Currently the process is run without any intervention, but the sites are small enough to manually check several of the transects.

The following known data issues could affect the results of the KDE signal finding algorithm:

Clouds

The presence of some clouds along a single granule can cause the loss of data that might otherwise be valid

Clouds reflect sunlight which causes a higher background photon rate, and this can create issues with the telemetry bands and cause the telemetry bands to not include the surface. Even if the actual earth surface is included in the telemetry band, the clouds can affect the travel times and create inaccurate readings (NASA NSIDC 2022).

During processing from L0 to L1, if the elevation from NASA reference DEM is not within the telemetry bands, no photons will be classified as signal. Therefore, if the entire granule is affected by this issue, there will be no sea surface found and therefore the entire granule will be filtered out. This can cause a significant loss of data but it is an issue inherent to nearly any remote-sensing based approach. One possible way to mitigate this would be to combine the ICESat-2 bathymetry data with synthetic aperture radar (SAR) remote sensing data. SAR remote sensing data can penetrate clouds because it uses radiometry outside of the visible spectrum. Although the spectrum used in SAR sensing cannot directly penetrate even clear water, the data can be used to estimate wave conditions, and the bathymetry can be estimated based on the transformation of the wave (called *wave kinematic satellite-derived bathymetry*). The bathymetry estimates from SAR-derived wave kinematic SDB could be incorporated into the Kalman filtering step.

In situations where the photons are able to pass through clouds, the changes to travel time through the clouds can affect the accuracy. This could potentially be something that is hard to detect and affects the accuracy of the bathymetric points if any are found.

Multiple Telemetry Bands

If the signal detection on board the satellite cannot determine where the primary surface is located, it will open another telemetry band to try to collect more signal. This can create other areas of photons that are significantly above or below the surface.

4

Results

This section details the results of investigations into how point estimates of bathymetry can be extracted from ICESat-2 data, and then how these point estimates can be interpolated and integrated into existing bathymetry grids to produce an upscaled product. Sections 4.1 and 4.2 validate the principle of interpolating point estimates and combining them with GEBCO. In section 4.1 random point samples of the validation data for a region are taken, then those points are used as synthetic input data to the kriging + Kalman updating approach. Then, in section 4.2, point estimates of bathymetry from JarKus are used as input to the Kriging Kalman updating process. The rest of the chapter shows the results of the full methodology, including ICESat-2 signal finding and the interpolation/updating of the GEBCO data. This full cycle is applied at 4 test sites: Marathon Key (Florida, USA), St. Croix (US Virgin Islands), St. Thomas and St. John (US Virgin Islands), and the Island of Oahu (Hawai'i, USA). Then, the full set of ICESat-2 data points extracted from all test sites are aggregated to evaluate if there are any variables which predict the quality or availability of bathymetric signal in the ICESat-2 data.

The global distribution of the test sites where the full cycle is implemented is shown in Figure 4.1

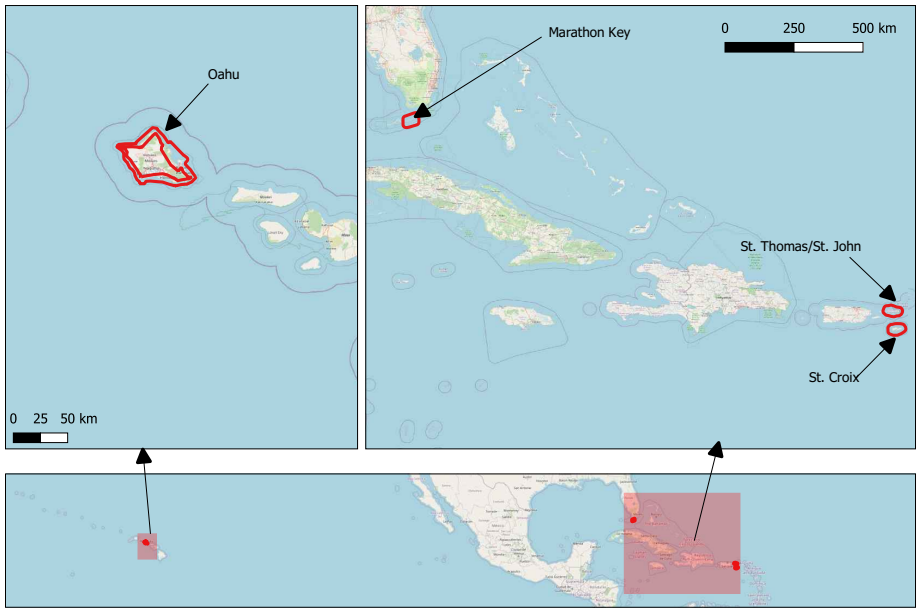


Figure 4.1: All test sites where the full methodology was implemented and validated

4.1. Kriging and Kalman update validation by random sampling

The basic approach behind both the kriging interpolation method and Bayesian updating of GEBCO is validated using a synthetic experiment. Instead of using point data from ICESat-2 as an input to the kriging step, random point samples of the validation data are taken throughout a single area of interest. This provides a source of exact bathymetry point measurements with an RMSE of 0 m. Two types of random sampling are considered, randomly across the surface of the validation data, and colinear samples that are taken only along the corresponding ICESat-2 tracks. The latter simulates the availability of theoretical perfect measurements from ICESat-2 bathymetry data, and can be compared to the random sampling case to investigate the effect of sparse, irregular spacing of data points induced by the satellite tracklines. Any available ICESat-2 bathymetric data will occur along colinear track lines, and therefore there will always be some spatial gaps in the data even if perfect ICESat-2 data were present along every single tract. By comparing the randomly sampled points to colinear points sampled along the ICESat tracks, we can get an estimate of how much the accuracy is degraded by the data being available only along certain lines.

A small section of the St. Croix validation data is used to implement this trial. First, the raster of the validation data is sampled randomly until we have a set of 10^4 random data points. From this set a subsample of 2000 points is taken using the Poisson disk sampling method. This subsample is used as input to the kriging algorithm.

Figure 4.2 shows both sets of points that were used as input to the kriging process for the two cases under evaluation. Note that in the bottom figure there are significant spatial gaps in coverage. This is due to the pattern of ICESat-2 data collection. Even in this theoretical scenario with exact bathymetry values available across the entire track, the spatial anisotropy of the availability of the points can potentially limit the accuracy of the interpolation.

Because we know the bathymetry points used for this section are exactly the same as the validation data at each point, the *nugget* of the variogram for the Kriging interpolation is set to 0, so areas that are close known points are given a variance of 0 by the kriging process.

		RMSE [m]	MAE [m]
	Truth vs GEBCO	4.48	3.31
Random sampling	Truth vs Kriging	3.13	1.18
	Truth vs Kalman updated	1.38	0.89
Colinear sampling	Truth Vs Kriging	4.49	2.12
	Truth Vs Kalman updated	1.97	1.32

Table 4.1: Comparison of error between the two different sampling strategies

Table 4.1 shows the results of the two sampling strategies. It can be seen that the more homogenous distribution of bathymetric points in the randomly-sampled data produces a better kriging estimate (RMSE of 3.13 m and 4.49 m respectively).

Also of note is that for both cases, the Bayesian combination of the kriging surface with GEBCO produces a better estimate than either on their own. This indicates that there is value in starting from GEBCO and using a Bayesian combination approach when compared to either a kriging interpolation of point data or using GEBCO as a data source.

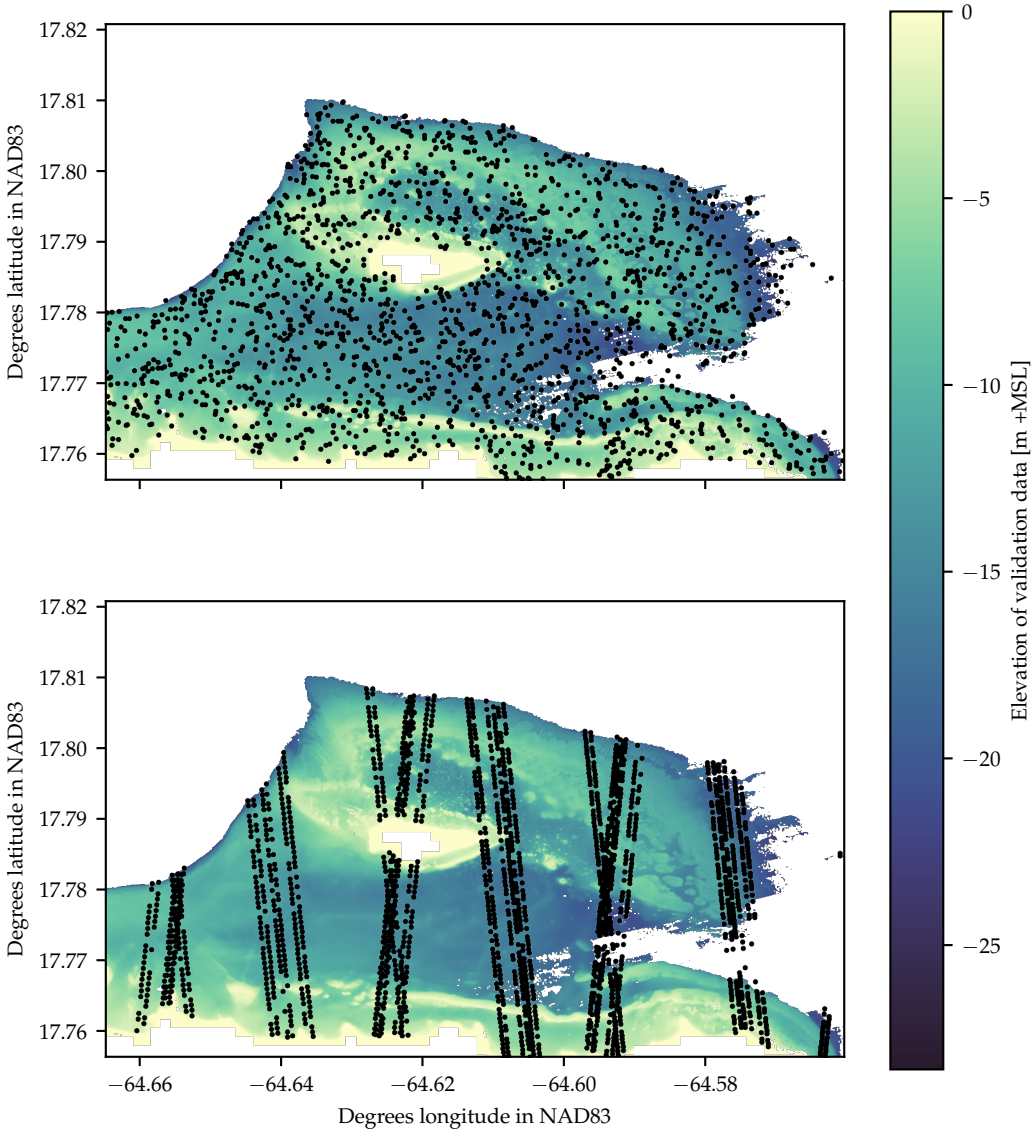


Figure 4.2: Top: Random sampling across the area of interest
Bottom: Random sampling only along the ICESat-2 tracks for the site

4.2. Petten Test Site

Most research results show that spaceborne lidar is not effective at retrieving bathymetry from very turbid waters (Daly et al. 2022; Coveney et al. 2021). To test the upper limits of this, ICESat-2 data for the coast of Petten, NL was downloaded. As expected, there was no bathymetric signal found in the ICESat-2 data, likely due to the inability of the laser to penetrate the turbid water.

However the site provides a useful case to validate the approach of the combination of universal kriging of point bathymetry and Bayesian combination of bathymetry grids. The Dutch government provides a dataset called JarKus which consists of annual surveyed elevations of the coastal zone (Minneboo 1995). These point measurements provide an interesting case to apply the same method but using the Jarkus point measurements as an input instead of ICESat-2 point measurements.

The location of Petten, The Netherlands, located on the Noord-Holland coast, was chosen based on the availability of high-resolution ground-truth bathymetry provided by Van Oord. The survey was performed in 2021, and therefore the 2021 Jarkus points were selected from the dataset. All Jarkus

points that overlap the validation data raster were selected. Like the ICESat-2 data points, the Jarkus points are available only along colinear lines. This is due to the measurement method that is based on surveying the dune and seabed starting from a series of fixed measurement poles and then surveying points normal to the shore starting at the reference pole.

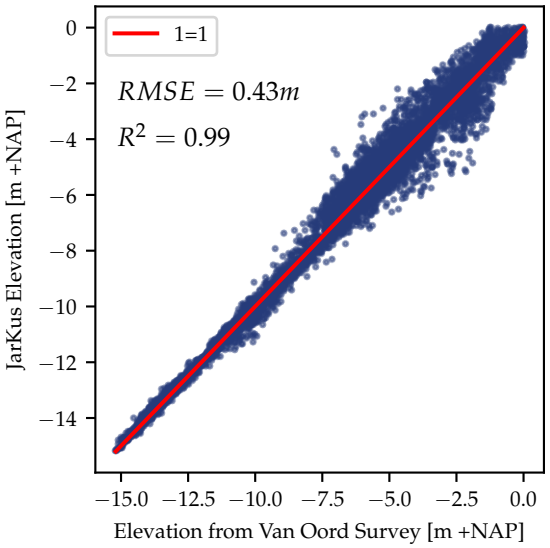


Figure 4.3: Petten Test site: Agreement between Jarkus and Van Oord multibeam survey data

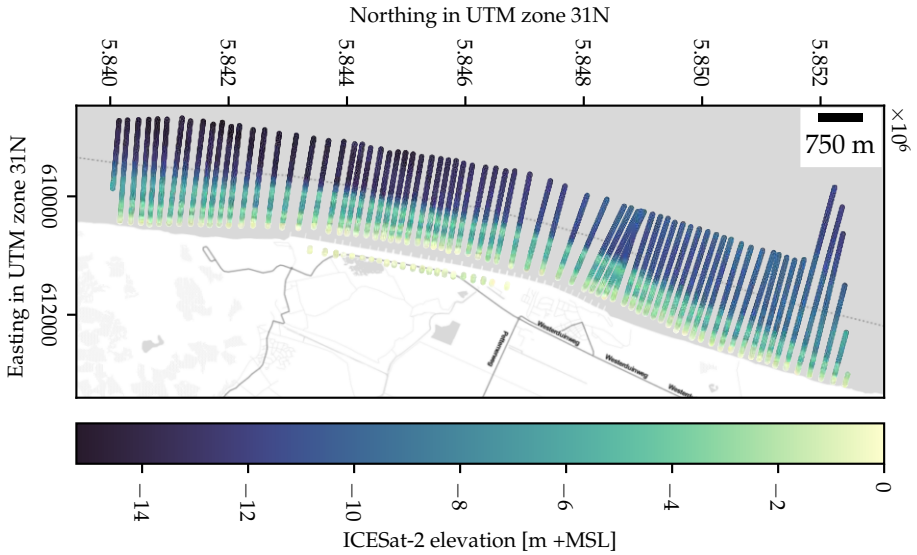


Figure 4.4: Jarkus 2021 measurement points within the site

The Jarkus point measurements were checked for the same error metrics as the ICESat-2 points. The bias plot is shown in Figure 4.3. It is notable that the deviation between the Jarkus data and the survey data is greatest in the in the shallowest part of the nearshore zone; this is likely due to some morphological changes due to wave-driven sediment exchange between the nearshore and the offshore. Both datasets are from 2021, but this area is highly morphologically active. Also, in the deeper part of the study area, the error between the two datasets decreases significantly. This is likely due to the depth of closure for most profiles in the study area being located between -5 and -10 m.

The Jarkus points are used as input to the kriging process instead of ICESat-2 points to validate the

universal kriging and Kalman updating process at this site.

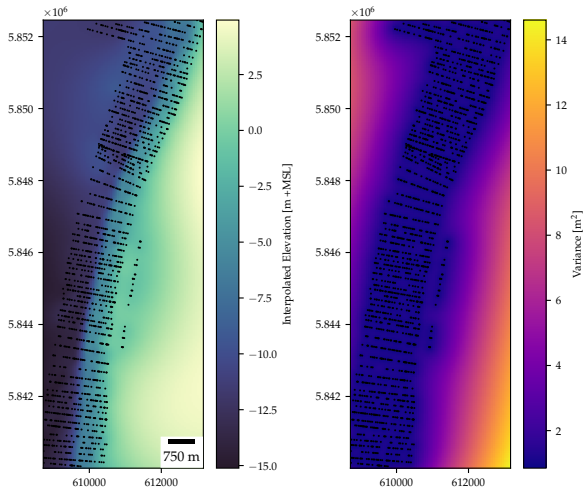


Figure 4.5: Petten site: result of kriging interpolation

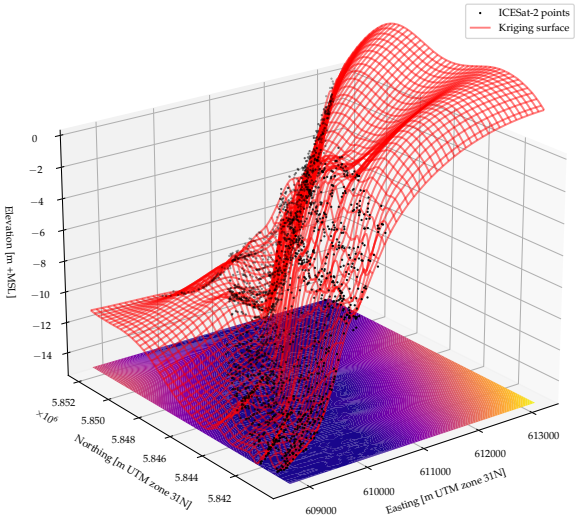


Figure 4.6: 3D results at Petten site

The kriging output is shown in Figures 4.5 and 4.6. The Jarkus points show an even distribution of measurement points in the cross-shore and alongshore directions in the entire study site, so the RMSE between the kriging output surface is very low. This site is also very well-suited to interpolation, since the bathymetry of the site is relatively homogenous and isotropic - the slopes and elevations are all relatively constant. Because this nearshore area has good quality point data and also is easy to interpolate, the kriged bathymetry surface was found to have a lower RMSE and MAE error than GEBCO.

It is notable that in this site, the error between GEBCO and the high resolution bathymetry is significantly lower than other test sites. This could be due to higher quality or more recent input data into the GEBCO grid as compared to the other sites. Even with a higher accuracy in the GEBCO data it was found that the Bayesian combination of both of these datasets has the highest accuracy — indicating that the Bayesian combination of these two datasets increases the value compared to either dataset individually.

	RMSE [m]	MAE [m]	ME [m]
GEBCO	1.48	1.31	-1.01
Kriging Surface	1.22	0.70	-0.31
Kalman Output	1.00	0.76	-0.50

Table 4.2: Petten Test site: Improvement in error metrics after applying Kalman updating of kriged data

4.3. Florida Keys test site

The area surrounding Marathon Key in the Florida Keys archipelago in Florida, USA was used as one site to test the entire process, including the ICESat-2 signal finding in additional to the kriging and Kalman updating. The area has a wide and relatively shallow ($\sim -10m + MSL$) shelf, a microtidal tidal environment, and very clear water, so it is an ideal site for remote sensing of bathymetry. The clear and shallow water provides a relatively strong subsurface signal within ICESat-2 transects. The site was found to have many transects of ICESat-2 data with distinct bathymetric signal. The KDE signal finding approach was applied to all transects within the study area and resulting distribution of the identified signal points is shown in Figure 4.7.

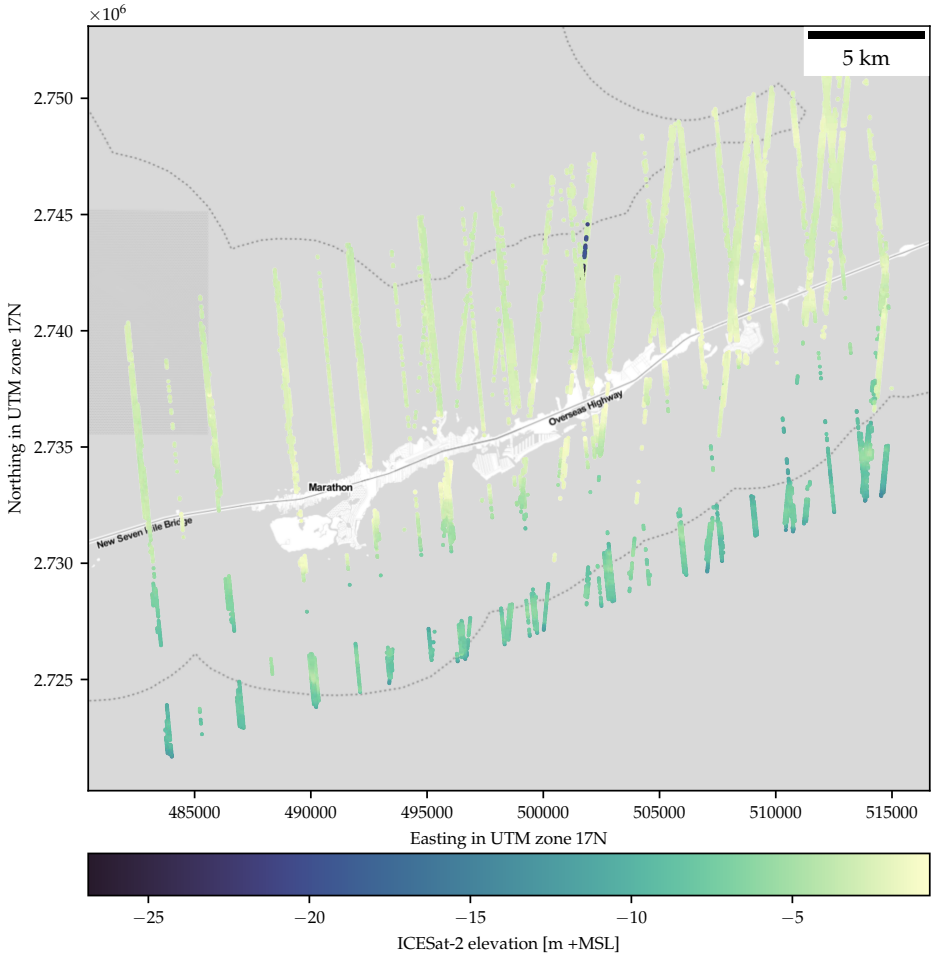


Figure 4.7: Marathon Key Site: Location and depth of bathymetric signal points found using KDE signal finding method

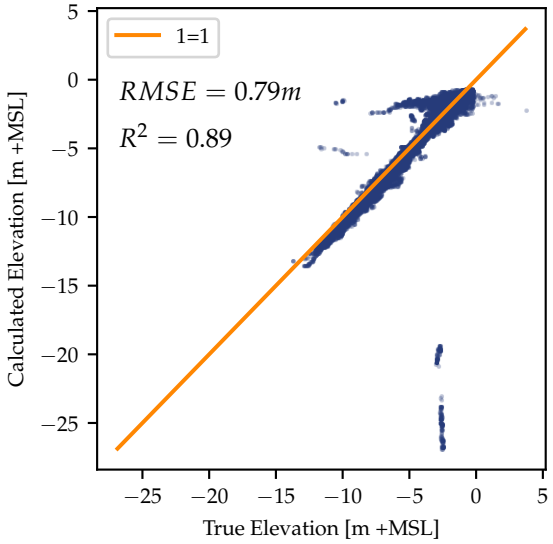


Figure 4.8: Marathon Key site: Bias plot showing the results of the KDE signal finding algorithm

Figure 4.8 shows how the error is distributed within the Florida Keys test site. This site has many

transects containing data, but some errors in the signal finding can be seen in this bias plot. There is one cluster of points with a real elevation of -2.5 m, while the KDE finding output identifies the seafloor elevation as between -20 and -27 m. This is caused by data being missing along this transect in one area. For this location, all photons above -15 m MSL are completely missing from the data. This is likely due to data quality issues, such as the onboard computer not being able to correctly identify the elevation of the primary surface return and opening another telemetry band. The figure also shows a few clusters of points where the seabed location was overestimated slightly. These errors are due to some places where the filtering step did not fully remove the ocean surface signal, which can bias the KDE signal finding algorithm by adding strong signal near the surface, which causes the algorithm to find the density peak higher in the water column than it should appear.

However, the kriging process can account for an expected degree of error, and can be set up such that the input points are not considered exact measurements but are considered to have their own uncertainty at each point. This is controlled by the *nugget* variogram parameter.

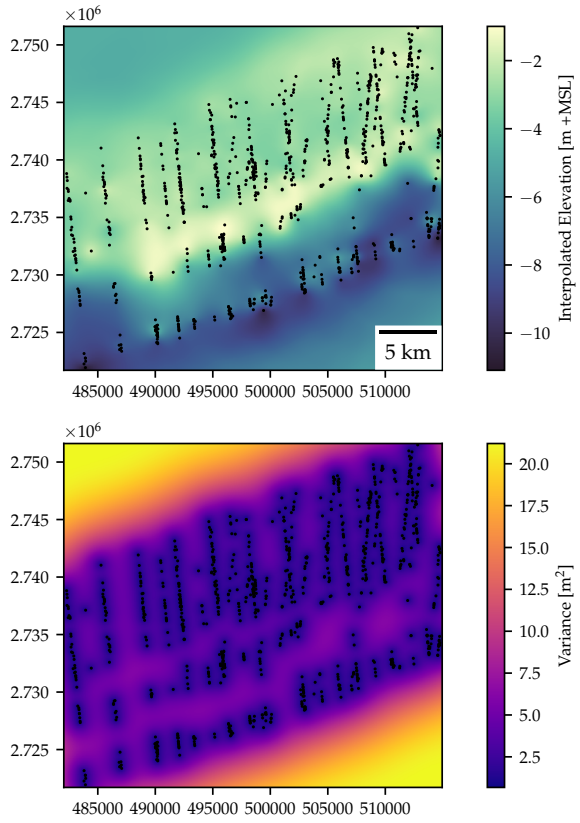


Figure 4.9: Output of the kriging process for the marathon key test site

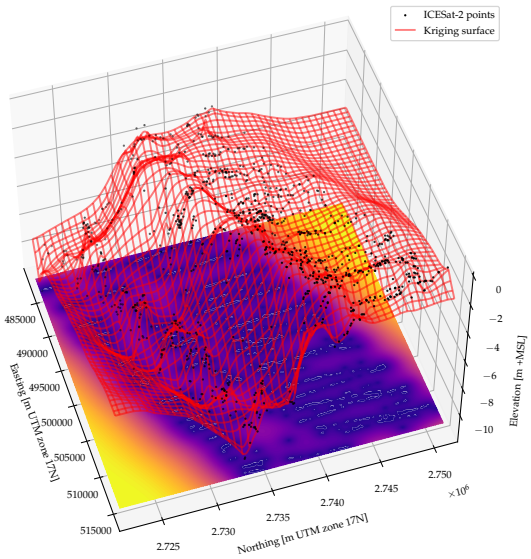


Figure 4.10: The Results of the kriging process in 3D

The Kriging interpolation results are shown in Figures 4.10 and 4.9.

The change in error after the Kalman updating step is shown in table 4.3, and the spatial distribution of this improvement is shown in Figure 4.11. At this site the largest improvements actually happen on the edges of the available spaceborne lidar data. The kriging surface that has been fit through these points happens to match the validation data well along these points. Also, notably this improvement occurs along a change in the GEBCO source data for this site.

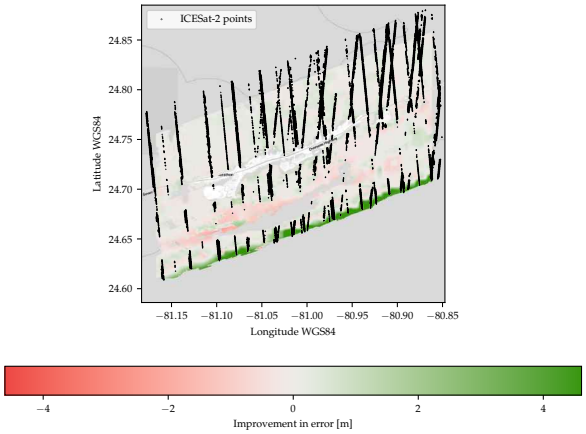


Figure 4.11: Improvement in error after application of Kalman updating

	RMSE [m]	MAE [m]	ME [m]
GEBCO	2.21	1.11	0.04
Kriging Surface	2.28	1.31	-0.60
Kalman Output	1.45	0.86	-0.14

Table 4.3: Improvement in error metrics after applying Kalman updating of kriged data

4.4. St. Croix test site

Another site where the entire processing chain was implemented was in the Caribbean island of St. Croix. The site was chosen based on the availability of recent high-resolution bathymetry validation data, and also on the basis of the clear water. The site provides some interesting contrasts: the south edge of the island is a relatively gently sloping shelf, while the north edge of the island has an extremely steep shoreline. On the northeast edge there is a bank between the main island and a smaller island.

The validation data used is provided by a USGS-sponsored survey of the area using a Riegl VQ-880-G II lidar sensor between January and June 2019. The lidar point data was then post-processed into a 1 m raster DEM product with a bathymetric vertical accuracy of 12.1 cm when compared to the survey control points (NOAA 2020). This DEM was used for validation. Note that the areas to the Northwest of the island has a very steep slope, and is therefore missing in the validation data - the slope is steep enough that the airborne lidar sensor could not get sufficient signal in that area.

The bathymetric point measurements obtained from the KDE signal finding algorithm are shown in 4.12. This site showed excellent agreement between the ICESat-2 data and the validation data, and the KDE signal finding algorithm reliably identified bathymetric points as deep as -20 m +MSL. The overall RMSE at the site was the lowest of any sites tested, at 0.54 m. Figure 4.13 shows the spread of the ICESat-2 points vs the corresponding point in the validation data. Just as in the validation data, the north west slope of the island does not contain any signal.

The results of the kriging output for this site are shown in Figures 4.14 and 4.15. Because the north end of this site has a very steep slope, there is no bathymetric signal found in this area and the resulting uncertainty is very high in this region. The south end of the island has a more gentle slope and there is a significant amount of bathymetric data available, so it has a correspondingly higher confidence in the kriging estimate.

After implementing the Kriging/Kalman updating process for the St. Croix site it is found that the combination of the kriged ICESat-2 raster with the GEBCO data produces a resulting product that has a higher accuracy than either of the input datasets. Particularly notable here is the inaccuracy of the GEBCO data at this site — while the overall RMSE of the resulting raster is high at this site compared to some others, in this case it is limited by the accuracy of the starting GEBCO data, and the accuracy of the kriging raster. Table 4.4 shows the error metrics for the site for the three bathymetric grids, and the spatial distribution of the improvement in error is shown in Figure 4.16. The improvement in error is also well-distributed within areas that contain bathymetric signal, with the largest improvements in error having a magnitude of about 6 meters. In some areas, there was a minor increase in absolute error. These areas are mostly in places where the bathymetric points are not evenly spatially distributed.

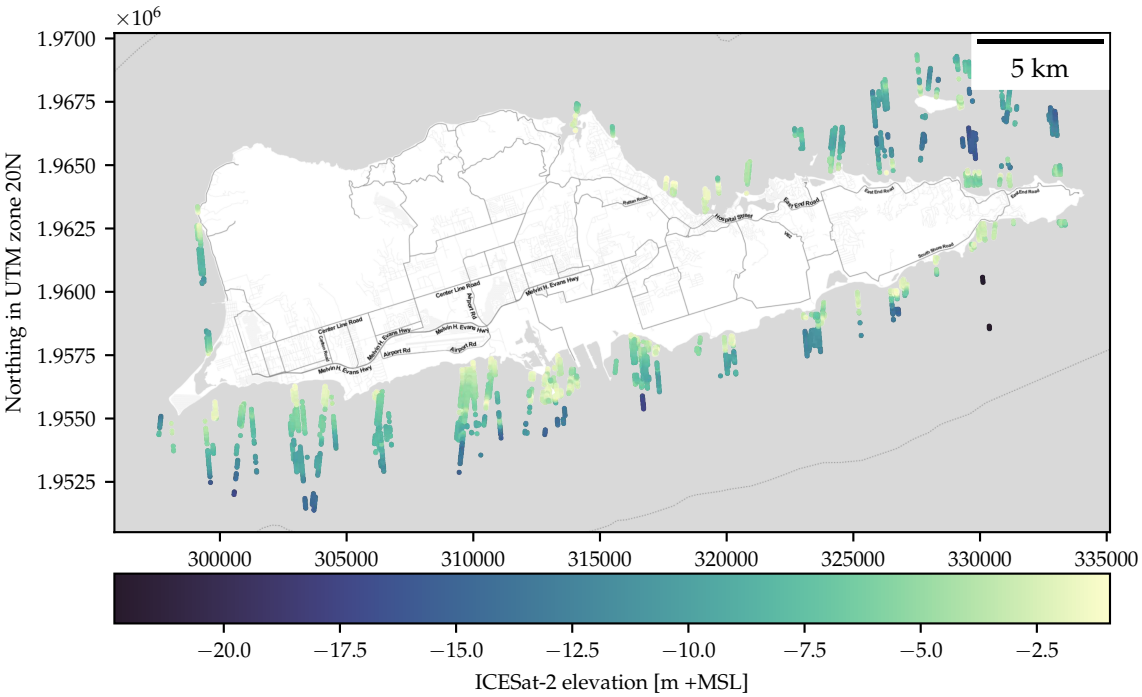


Figure 4.12: St. Croix site: Identified photons and their depth

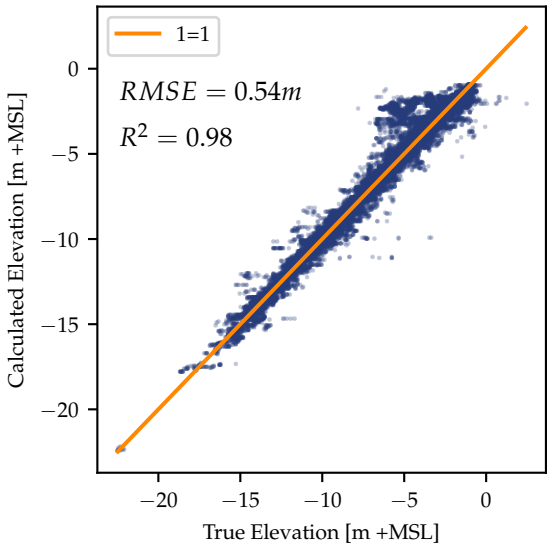


Figure 4.13: St. Croix site: Error Bias plot

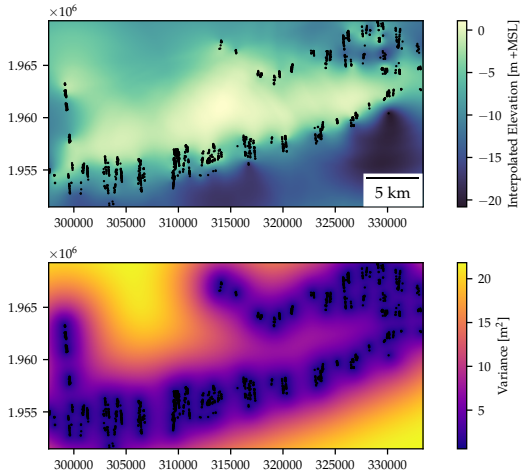


Figure 4.14: Results of Universal Kriging interpolation in the St. Croix Test site

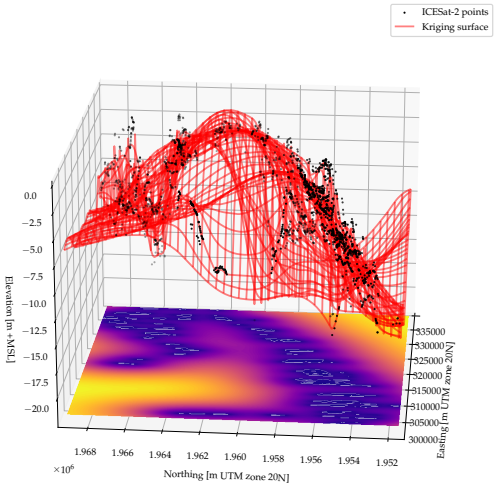


Figure 4.15: 3D Plot of the kriging results (red wireframe) with the resulting uncertainty shown on the bottom of the plot. Note that the scale of the uncertainty (plotted on the bottom) is the same used in the lower half of 4.14

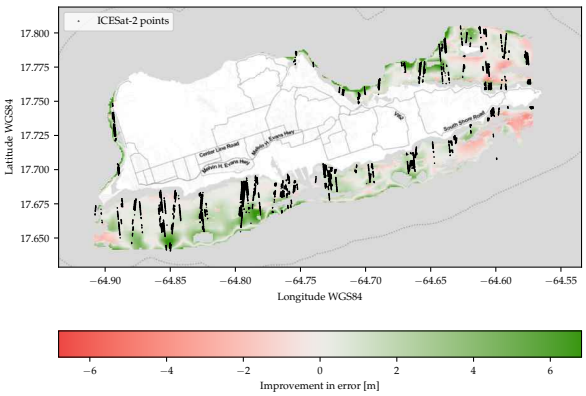


Figure 4.16: Improvement in absolute error in meters after combining the interpolated bathymetry grid via Kalman updating. The improvement at this site occurs mostly at the middle depths. A larger version of this figure is shown in C.15

	RMSE [m]	MAE [m]	ME [m]
GEBCO	6.43	4.29	-1.01
Kriging Surface	6.54	3.90	-2.67
Kalman Output	4.62	3.15	-1.05

Table 4.4: Improvement in error metrics after applying Kalman updating of kriged data

4.5. St. Thomas and St. John test site

Another test site used to evaluate the ICESat-2 bathymetry signal finding steps is area around the islands of St. Thomas and St. John, located in the U.S. Virgin Islands. The islands are ideal for spaceborne bathymetry for the same reason that many Caribbean islands are: they have naturally clear water that is easily penetrated by lidar signal. Additionally, the area has recent, high-resolution bathymetric data available from NOAA to allow validation of the approach. The validation data for the site is from the same surveying project described in Section 4.4. Because the bathymetric DEM is continuous between the two islands, the two islands were considered as a single test site.

The ICESat-2 data was downloaded for an area of interest that includes extent of the validation data around both islands. Figure 4.17 shows the signal that was identified from the ICESat-2 transects at the site. Figure 4.18 shows the error between the ICESat-2 points and the validation data. There is a cluster of several points on the right side of the plot that shows points with a true elevation of between 10–20 m. This is due to several land points being incorrectly identified as bathymetric signal, which significantly increases the RMS error at the site.

The lidar bathymetric points in this site are moderately well distributed, but each identified cluster of signal is smaller. This is due to the very steep slope of the underlying bathymetry. This steep slope also affects the quality of the bathymetric data around the site. The offshore extent of the bathymetric data is relatively small, because in most directions the depth increases to beyond the maximum depth of the airborne lidar very quickly.

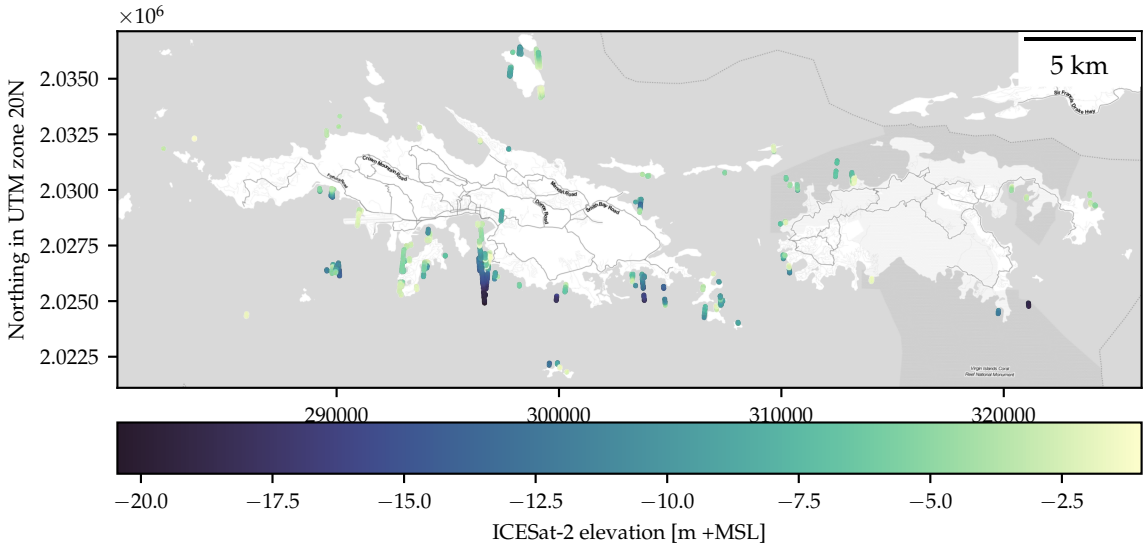


Figure 4.17: St. Thomas and St. John test site: Location and depth of bathymetric signal points found using KDE signal finding method

The bathymetric points found via the KDE signal finding process were then interpolated using universal kriging interpolation. The interpolated surface is shown in the upper half of Figure 4.19 and in the wireframe surface shown in Figure 4.20. This site shows many high-accuracy point estimates in deeper areas of the site (as can be seen in Fig. 4.18). However, these points are relatively isolated so confidence in the kriged bathymetry estimate falls off quickly. Also since the clusters of signal have large gaps between them, the accuracy of the fitted curves is likely reduced since trends cannot be estimated as accurately as they would be if there was more intermediate points between the clusters.

After the Kalman updating of the GEBCO bathymetry data, the RMS error was reduced from 7.35 m to 5.78 m. The spatial distribution of the improvement is shown in Figure 4.21. In some areas, the absolute error is improved by up to 7.5 m compared to the GEBCO data.

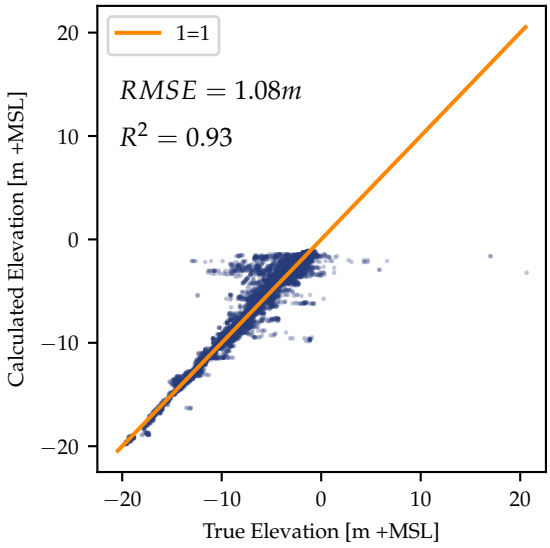


Figure 4.18: St. Thomas and St. John Test site: Bias plot showing the error of ICESat-2 bathymetric signal

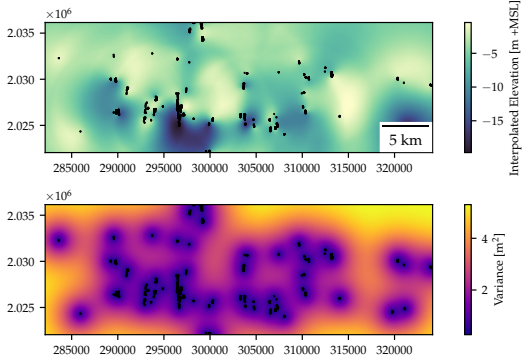


Figure 4.19: St. Thomas and St. John Test site: result of Kriging interpolation

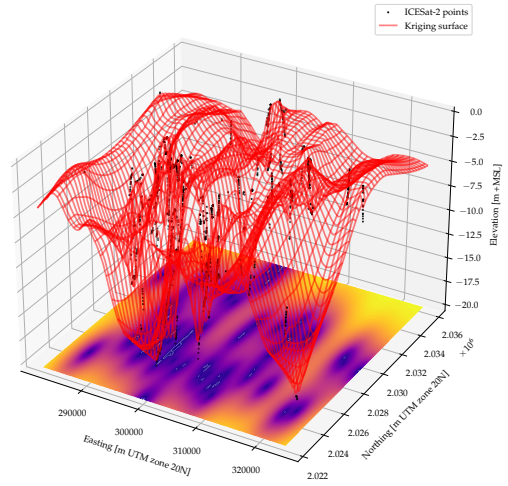


Figure 4.20: St. Thomas and St. John Test site: 3D results of the kriging interpolation

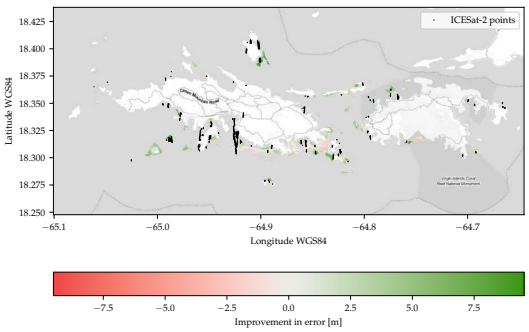


Figure 4.21: St. Thomas and St. John Test site: Improvement in absolute error [m] after Kalman update. Note the relatively limited offshore extent of the validation data

	RMSE [m]	MAE [m]	ME [m]
GEBCO	7.35	5.39	0.92
Kriging Surface	7.53	5.91	-3.61
Kalman Output	5.78	4.32	-0.87

Table 4.5: Improvement in error metrics after applying Kalman updating of kriged data

4.6. Oahu Test sites

For another validation site that is outside of the Caribbean sea, the island of Oahu in the US state of Hawai'i was chosen. The validation data used is a lidar survey completed in 2013. The nearshore zone was surveyed by the United States Army Corps of Engineers (USACE) using the Coastal Zone Mapping and Imaging Lidar (CZMIL). The survey data was collected by aircraft from September to November 2013. The raw lidar point cloud data was then further processed into a 1 m DEM via TIN interpolation. The resulting 1 m DEM is referenced to a horizontal datum of NAD83(PA11), and the vertical reference system is the local mean sea level. The data has a vertical accuracy at the 95% confidence interval of $\sqrt{0.20^2 + (0.013d)^2}$ m, where d is the depth in meters. This translates to a 95% confidence vertical accuracy of ±28 cm at a depth of 15 m.

Because the island of Oahu is substantially larger than the other test sites ($\approx 4x$ the area), the coast of the island is first divided into 8 smaller subsites to allow for easier data downloading and processing. The layout of these sites was chosen based on approximately similar coastline characteristics, and to try to minimize any interpolation over convex points in the coast. Only some selected sites are highlighted here. Other sections Figure 4.22 shows the selected layout of the subsites along the coast of Oahu.

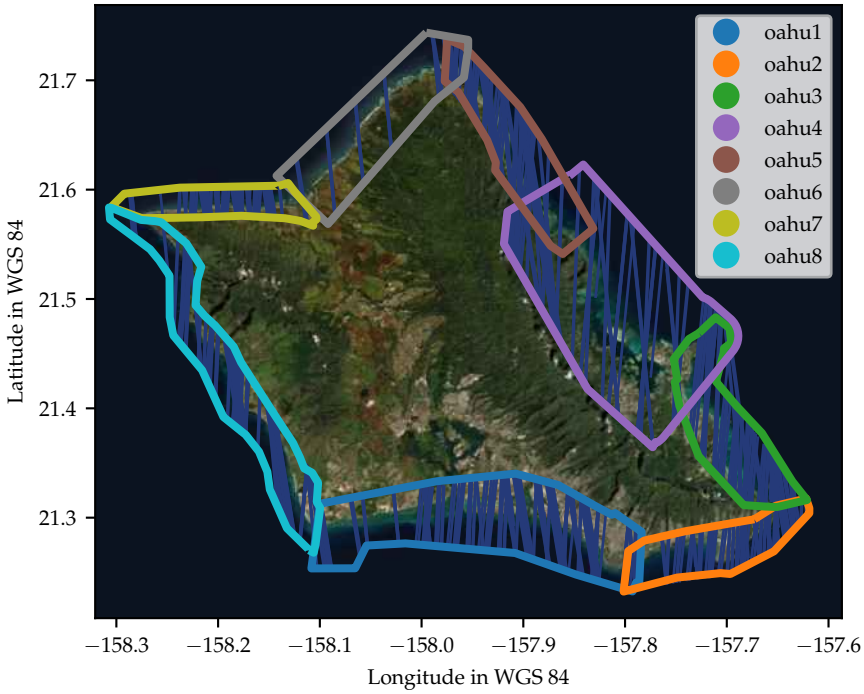


Figure 4.22: Oahu test site: The availability of ICESat-2 transects within the selected subsites. Base layer data provided by: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

The different subsites also exhibit different hydraulic characteristics. The Northern edge of the island is exposed to longer swells and larger wave heights, while the south is relatively sheltered (Vitousek and Fletcher 2008). Another aspect is that distribution of ICESat-2 tracks is not even across all sites. As seen in Figure 4.22, sites 4 and 6 do not contain any transects going in the other direction with good data, so distribution of the data will be more uneven and might reduce the quality of the kriging raster.

The accuracy of the ICESat-2 bathymetry points from the KDE signal finding algorithm, grouped by subsite, are shown in Table 4.7. It can be seen that site 2 has an anomalously high RMS error.

The cause of this is a weakness in the subsurface filtering process. In This case, there is steep seaside cliff that is not captured by the low-resolution of the GEBCO grid. As this area of the transect occurs on land, it ideally should have been removed during the horizontal filtering step. Due to this issue, points are inadvertently included in the subsurface photon dataset for the transect, and the KDE signal finding

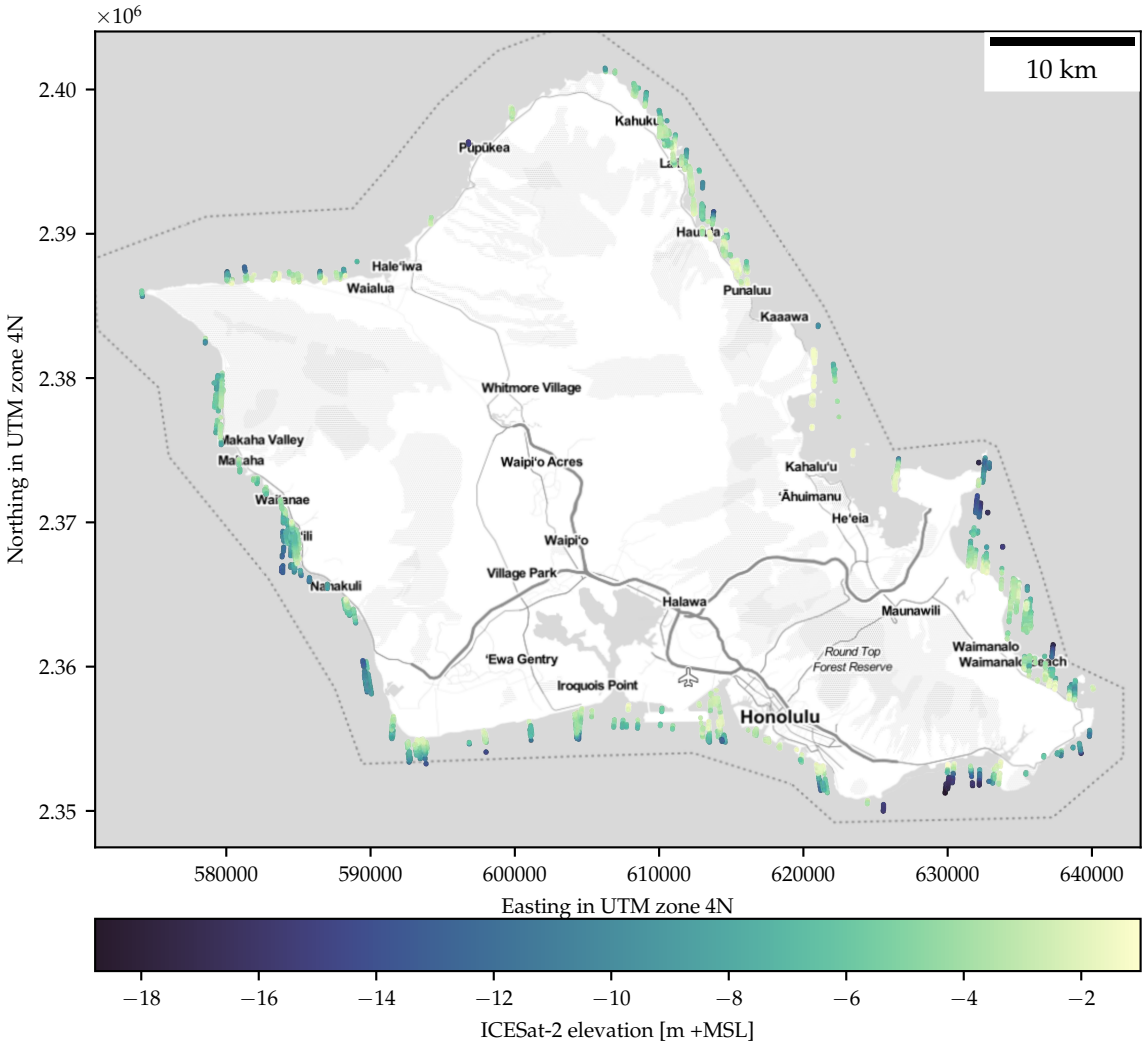


Figure 4.23: Bathymetry point estimates around Oahu from KDE signal finding algorithm

Subsite	RMSE Change	MAE Change	ME Change
1	28.83%	30.53%	73.06%
2	31.62%	29.61%	51.10%
3	25.18%	22.68%	75.06%
4	-1.76%	-1.01%	123.78%
5	1.94%	4.52%	116.88%
6	-18.14%	-16.12%	-14.68%
7	-22.61%	-15.04%	-16.58%
8	-25.15%	-21.05%	-544.94%

Table 4.6: Percent reduction in error metrics via the Kalman updating approach. Positive values indicate reduced error, negative ones indicate increased error

Subsite	RMSE [m]	MAE [m]
1	1.16	0.77
2	10.60	1.45
3	1.24	0.46
4	0.61	0.39
5	0.73	0.50
6	2.42	1.76
7	1.11	0.72
8	0.67	0.52

Table 4.7: Error metrics between ICESat-2 and validation data for all sites in Oahu

algorithm incorrectly identifies these points as signal. Because the actual elevation of the validation data is about 80–140 m in this area, and the output of the KDE signal finding is approximately -5 m in this area. Because of this, the RMSE metric for each of these points is anomalously high. With this in mind, these points have been removed from some subsequent analyses of the data to better show the distribution of errors in other points. Figures 4.24 and 4.25 are combined bias plots that show the ICESat-2 bathymetry points of all subsites. The incorrectly classified mountain points can be seen on the right edge of Figure 4.24. Figure 4.25 is a plot of the error excluding these mountain points.

The Kalman update step was also applied individually to each subsite. The Kalman updating step for the various Oahu subsites shows a more mixed result than in other test sites. A summary of the changes in the error metrics for each subsite is provided in Table 4.6. In some sites, the Kalman update does not improve the estimate, and in some cases has an even higher error than the a-priori GEBCO estimate. This is an indication that the parameters are not set correctly for the site — using Bayesian estimation, the parameters should be able to be set such that areas within the grid that have lower confidence are not changed significantly.

The integration of the ICESat-2 to GEBCO via Kalman updating does not improve the bathymetry estimates in subsites 6–8. For site 6, this is likely due to the quality of the ICESat-2 data — it has a very high RMSE and very few bathymetry points in this subsite that were identified by the KDE signal-finding algorithm. There are several possible reasons for this. For one, the site has fewer transects available from the NSIDC portal, so there is less data to start with. Additionally, subsite 6 is located along the north edge of the island which experiences higher wave energy conditions than other subsites (Vitousek and Fletcher 2008). The higher energy wave environment could degrade the quality ICESat-2 bathymetry data in two ways: increasing the slope of the local sea surface and therefore increasing the refraction error, and the foam due to breaking waves at site could affect the ability of laser photons to penetrate to the seabed. The breaking waves can be seen in satellite imagery of the subsite in Figure C.41.

However, the effect cannot be fully explained by the quality and quantity of ICESat-2 bathymetry points. Sites 7 and 8 both have significant amounts of ICESat-2 bathymetry points available with relatively low RMSE values (1.11 m and 0.67 m respectively). Understanding why the Kalman updating approach did not improve the results requires examining the resulting bathymetric surfaces. On the west side of Oahu, the validation dataset includes bathymetry as deep as -60 m in some areas. However, the deepest ICESat-2 photons in the subsite are only at a depth of -13 m. Because the deepest depths that are input to the kriging algorithm are -13 m, the deepest parts of the kriged bathymetric surface are approximately the same depth. What is happening here is that the initial guess of approximately -60 m is being affected by the -13 m kriging output. The strength of the update is small, but because the difference between -60 m and -13 m is so large it has a large impact on the RMS error. This could be improved by either experimenting with different variogram parameters setting to decrease the confidence of the kriging output in these areas, or simply by masking out any values deeper than the deepest ICESat-2 bathymetry point from the final product.

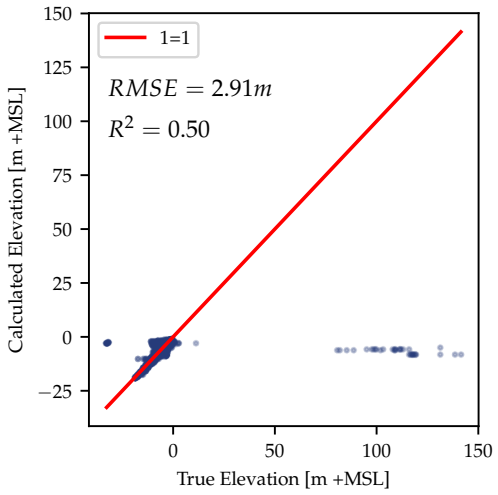


Figure 4.24: Bias plot of all points within all Oahu Subsites

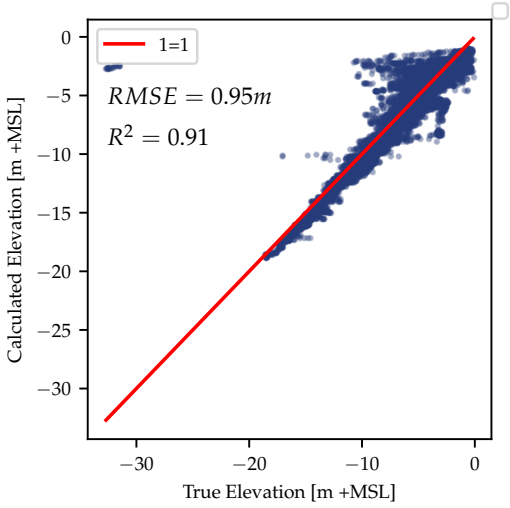


Figure 4.25: Bias plot with the mountain points removed to better show the actual distribution of the error

4.7. Prediction of Lidar error

One of the challenges of extracting bathymetry from ICESat-2 data is finding robust, automated ways of distinguishing between signal and noise. For the KDE approach utilized, a minimum kernel density threshold of $\max(kde_{50}, 0.10)$ was found to provide a good balance between filtering out the most severe errors and retaining as much data as possible. Increasing the minimum parameter to 0.15 was found to decrease RMS error of the ICESat-2 point bathymetry, but at the expense of throwing out significant amount of high quality data at the same time. Therefore, increasing the threshold was found to lower the error for the point data while simultaneously *increasing* the error in the resulting kriging surface.

Therefore, better methods are required to predict the quality of the bathymetry data in coastlines without in-situ data. Two ways of approaching this are explored: analyzing which parameters predict the error at the level of an individual photon, and analyzing if certain variables predict the aggregate error of each transect. Improved filtering by transect would allow better discrimination of signal/noise, improving the quality of the output, while finding variables that indicate if a transect contains any bathymetry data would allow pre-filtering of the best transects, allowing the KDE signal finding approach to be scale to larger sites more efficiently.

4.7.1. Per-photon error

Several variables that are provided on a per-photon basis which might have some impact on the error at each photon. Two variables that were checked for their relationship to the error of a single photon are the count of photons in the 20 m segment (*count_ph* in the ATL03 data). This is the count of photons that are in the 20 m segment around the reference photon. An indication of the total count of photons could be useful to predict the presence of bathymetric signal, since fewer photons present in a segment might indicate data issues (e.g., interference from clouds) in the surrounding area.

To evaluate the effect, every single bathymetry point from all sites was combined. The comparison between the absolute error and the count of photons in the local 20 m segment is shown in Figure 4.26. The photons from Oahu subsite 2 with an error in the range 80–150 m have been removed from the data before plotting. Because the reason for the anomalously high error in these points is known, they will not examined any further here (For the reason for these errors see section 4.6).

Looking at the other photons, clustered near the bottom of Figure 4.26, there is indeed a relationship between the photon count variable and the absolute error; the photons with higher absolute error (from 20–40 m) are all clustered near the left side of the graph, indicating that many of the photons with a

large error occur in transects with a photon count from 0–200 photons. However, there is a large density of photons with a relatively low error value that are also in this range. Therefore, it presents the same tradeoff as increasing the kernel density threshold - increasing it will increase the quality of the data, but also requires throwing out a very significant amount of good data at the same time.

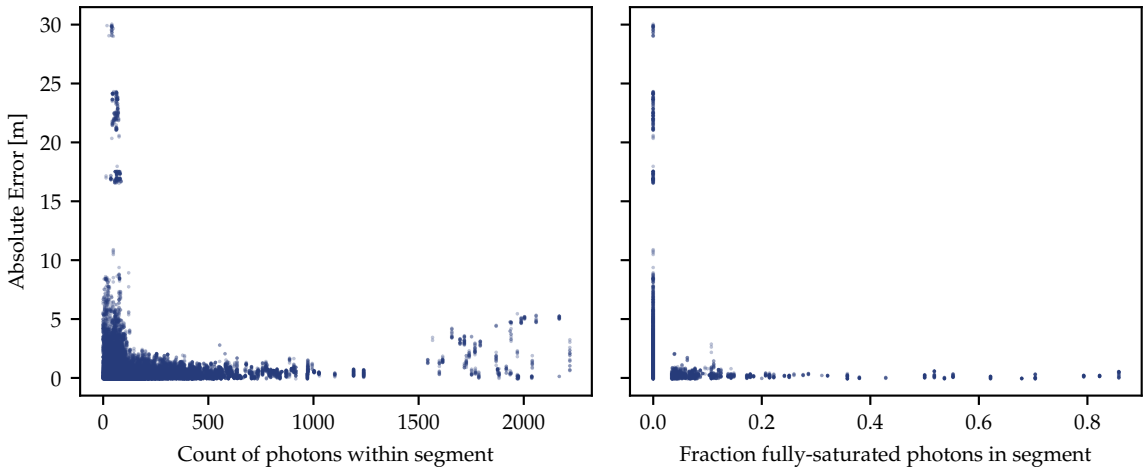


Figure 4.26: Left: Relationship between Photon count within 20 m segment and absolute error
Right: The number of full saturated photons in a segment vs the error at that point

Another variable that was checked at the photon level is the percent of pulses within the segment that are determined to be fully saturated. This can also be an indication of data issues in the segment. However this variable is also not a good predictor of error in an individual photon, as seen in the left side of Figure 4.26. Part of the reason for this is that nearly all of the segments included here have a full saturation fraction of 0.0 (99.1% of photons in all study sites).

4.7.2. Per-transect error

Another useful metric is understanding which transect-level variables could indicate high-quality bathymetry data within a transect. If some transect metadata variables are a strong indication of the quality of the bathymetric signal in the transect, this would allow filtering of transects before running the KDE algorithm, allowing a significant reduction of required computational resources. To calculate the transect-level statistics, all the signal photons are aggregated by the unique combination of date, site, and beam id. The unique combination of these three identifies a single transect at each site. The aggregate statistics are calculated to find the RMS error, the median Secchi disk depth, the average count of fully saturated photons, the photon density per meter, and the percent of high-confidence ocean photons.

One variable that was checked at the transect level is the Secchi disk depth. The median Secchi disk depth was calculated for each unique transect, and then compared to the aggregate RMSE for that transect. The upper left side of Figure 4.27 shows the resulting plot. For this purpose, the large magnitude errors from Oahu subsite 2 were removed, since they are caused by a known issue that is not related to the optical depth. The RMS error shows an even distribution across the entire range of Secchi depth values at the transect. Therefore there is no clear relationship between the median Secchi disk depth of a trackline and the RMS error of the ICESat-2 bathymetry data contained in that trackline, so this variable would not be useful to determine if a transect is likely to contain bathymetric signal.

The results of the photon-level analysis indicates the fraction of fully-saturated photons in the 20 m segment around a photon does not provide a good way to classify an individual photon as signal or noise. However, if many segments within a certain transect have a large fraction of saturated pulses, this indicates possible instrument errors along this transect. The average of the fraction of fully-saturated pulses for each unique transect was calculated. The upper right subfigure of Figure 4.27 shows the

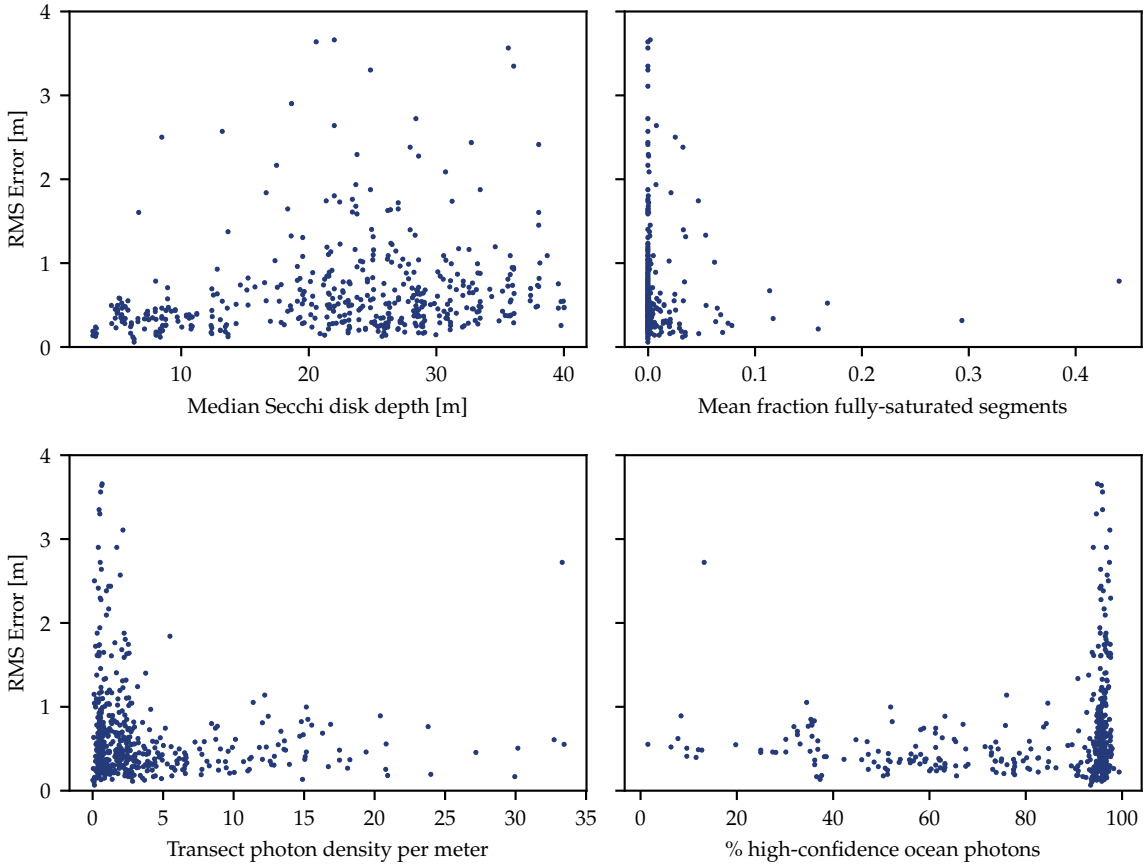


Figure 4.27: Relationship between transect-level statistics and RMS error of ICESat-2 data. None of the investigated variables can serve as a good way to predict the presence or quality of bathymetric signal.

relationship between this mean fraction of fully-saturated pulses and the RMS error in the transect. No significant relationship is found.

It was also hypothesized that the number of photons per meter of along-transect distance could provide a way to filter for high-quality data in a transect. It is inexpensive to calculate and indicates transects which consistently contain a high number of photons across their length with no significant horizontal gaps, which also likely indicates ideal data capture conditions since large horizontal gaps in signal are often induced by poor atmospheric conditions such as clouds or fog (NASA NSIDC 2022). The density of photons per meter length for each transect was calculated. Its relationship with the transect RMS error is shown in the lower right of Figure 4.27. The plot shows that transects with a relatively high density relative to their length tend to have a lower RMS error. However, it does not serve as an effective way to filter: the majority of transects fall within the 0–5 m range, including those with very low RMSE and relatively high RMSE. Therefore using this value as a threshold would require filtering out many good transects to remove a handful of low-quality transects.

Another variable which might provide a an indication of the presence of bathymetry signal in a given transect is the percentage of high-confidence ocean signal photons. This metadata variable is provided in *qa_perc_signal_conf_ph_high* within the */quality_assesment/* group in the ATL03 data. Generally, the ATL03 classification does not classify bathymetry signal points *themselves* as high confidence ocean-surface photons (often, bathymetric signal points are classified as noise). However, there could be an indirect relationship between the percentage of high confidence ocean photons and the bathymetric signal quality because the percentage of high confidence ocean photons could be a proxy for the transect data quality in general (i.e., good atmospheric/weather conditions, no instrument issues, etc). The lower right side of Figure 4.27 shows the relationship between the percentage of high confidence ocean signal photons and the RMS error. On the right side of the graph, it can be seen there are a large

number of transects with a more than 95% high confidence signal but still showing significant RMS error. Therefore, this variable does not provide a good way to determine which transects contain quality signal; any filtering based on the percent of high confidence ocean signal photons would remove significant amounts of high-quality transects, while leaving many of the lower quality transects.

Previous research on ICESat-2 bathymetry has found that the beam strength can affect the quality of the bathymetry data obtained. To investigate this relationship, the transects were grouped by this binary variable to find the RMSE of all strong beams and all weak beams. For this approach, it was found that the strong beams did have a slightly higher RMS error, but also were able to extract many more photons. Table 4.8 shows the error metrics, and the total count of points found for each type of transect.

Beam type	RMSE [m]	MAE [m]	ME [m]	Count of signal points
Strong	0.81	0.34	0.19	360,556
Weak	0.66	0.38	0.09	56,484

Table 4.8: Error metrics per-transect based on beam type

5

Discussion

5.1. A-priori identification of transects with bathymetric signal

One of the goals of the project is to find a way to predict nearshore zones which are most likely to have ICESat-2 bathymetry available before downloading the photon-level data and running the KDE signal finding algorithm. A-priori identification of promising sites is useful to minimize the computing resources required when scaling up the approach to a global level.

When looking at the transect data as a whole, there was no single variable which predicts either the availability or the quality (measured by RMSE) of bathymetric data contained in a given transect. It is well-established that lidar and optical bathymetry are highly sensitive to turbidity levels, since they require light to be able to reach the seafloor. It was hypothesized that the Secchi depth, as estimated by optical satellite methods, would be a reliable indicator of bathymetric quality and quantity for a given transect. However, no clear relationship was found between the Secchi depth and the quality and quantity of bathymetric signal photons. In fact, the Florida Keys test site has the lowest median Secchi depth (8.3 m) of any of the test sites investigated, while also providing some of the highest quality data (0.87 m RMSE), and the largest percent of transects containing valid signal of any test site.

The most likely explanation for this seeming paradox is that the characteristics that make a site ideal for ICESat-2 bathymetry (i.e., a wide, shallow shelf with clear water) could affect the accuracy of optical ocean color measurements. The area of interest selected in the Florida Keys site has large areas with a depth of less than 10 m, and exceptionally clear water. The Secchi depth determination is based on passive optical remote sensing response of the site. It is possible that because in this site there is a very significant amount of sunlight reflecting from the seabed, the Secchi depth algorithms used in the GlobColour product have limited accuracy. This is further supported by the fact that the Secchi depth uncertainty metric (median 56% uncertainty) is much higher for the Florida Keys site than any other site. Table 5.1 shows the median Secchi depth and light diffusion coefficient for each subsite, as well as the RMSE and MAE error between the ICESat-2 point data and the validation data.

Site Name	Zsd_{50}[m]	$\sigma_{Zsd_{50}}$	RMSE [m]	MAE [m]	ME [m]
St. Thomas/St. John	23.52	23.92	1.08	0.56	-0.17
Florida Keys	8.29	56.01	0.79	0.31	0.24
Oahu 1	29.70	29.52	1.16	0.77	0.54
Oahu 2	31.59	28.97	10.60	1.45	1.41
Oahu 3	28.43	50.92	1.24	0.46	0.30
Oahu 4	28.43	27.14	0.61	0.39	0.26
Oahu 5	29.84	25.48	0.73	0.50	0.27
Oahu 6	28.93	31.82	2.42	1.76	-1.43
Oahu 7	32.76	25.54	1.11	0.72	-0.02
Oahu 8	33.26	23.75	0.67	0.52	0.35
St. Croix	26.89	41.20	0.54	0.30	-0.11

Table 5.1: Secchi depth and error in bathymetry point measurements for each site

5.2. Bathymetry extraction from ICESat-2 lidar

Compared to other signal finding methods proposed in the literature, the implementation used here results in a similar order of RMSE for most of the test sites. The KDE signal finding method proposed was able to identify bathymetric photons with an RMSE of between 0.54 m and 10.6 m. This is a very wide range of reliability, and would require further refinement before scaling the method to more sites. The site with the largest error was in an area with steep cliffs on the south end of Oahu, Hawai'i. The source of the error was a failure in the horizontal filtering approach. The error was caused by some points with a true elevation of 118 m being labeled as subsurface photons. Because they are in a land area, they should have been removed by the GEBCO filtering step. However, because of the limited horizontal resolution of GEBCO data they were not removed from the photon cloud during subsurface photon filtering.

The best way to mitigate this would be to find a better land mask dataset. This is likely possible for sites in the global north that have good data availability, but higher resolution land masks might not be available in the entire world. It should be noted that the ATL03 dataset does include a surface mask flag as one of the variables. However, it is not suitable for this purpose because there is an intentional overlap between the boundaries of these different masks. Therefore, a photon return in a coastal zone can be classified as being both in the ocean, land, and inland water zones. This is done because it is assumed there is a high uncertainty in the exact boundaries between surface types.

One of the challenges of the proposed KDE signal finding method is that it is very sensitive to the input parameters, especially the threshold KDE value to distinguish signal from noise. The threshold used here was set empirically, and it was the greater of either a.) the median KDE of the transect, or b.) 0.1. By making this criterion stricter (increasing the minimum to 0.15) the RMS error at each site can be reduced significantly. However, this presents a tradeoff. While a few large magnitude errors are filtered out by increasing this threshold, there are many times more high-quality points that are also excluded. This can be seen in Figure 5.1, where the inset plot shows the relative density of low-error photons which have a KDE value less than 0.15. By excluding these points, it often reduces the overall spatial distribution of the resulting point data (the KDE is correlated for a given transect, so by increasing the threshold, often entire transects worth of points are removed). Therefore, increasing this threshold is quite a blunt instrument that removes significant amounts of high quality data. It was found that using a lower density threshold, and therefore allowing some large errors to remain in the point data, the overall quality of the kriging interpolation and the resulting Kalman-updated output product was improved.

One potential avenue for improving the discretion of the algorithm would be considering the local horizontal density of photon returns. A rolling window based on distance rather than a set number of adjacent photons could potentially improve the quality of the KDE magnitude estimation. Another approach might be to normalize the KDE value based on the transect length, and see if that improves the ability to throw out low-quality data while retaining high-quality data.

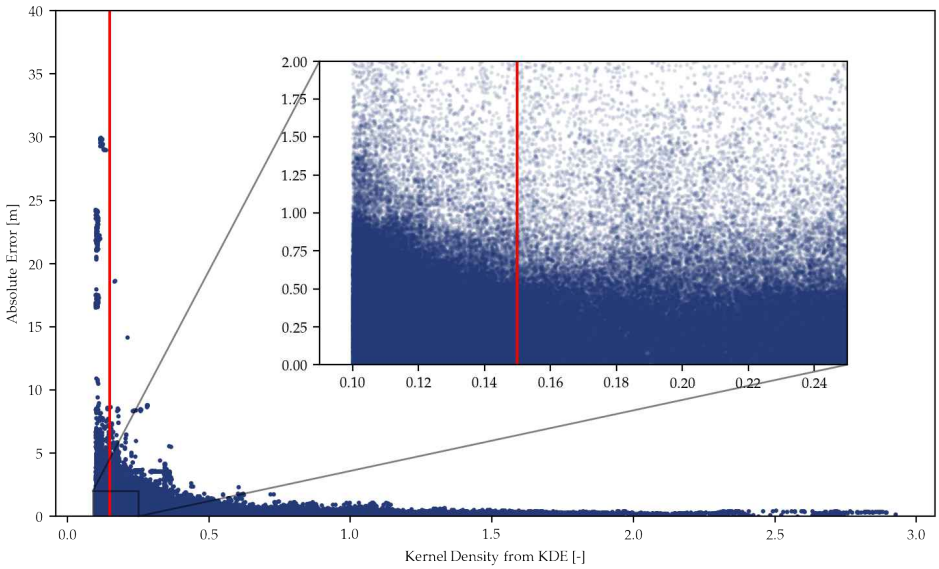


Figure 5.1: Effect of minimum density threshold parameter choice. The red line shows a threshold of 0.15. If this was used as the signal classification threshold, this removes almost all the most severe errors but also eliminates many times more high quality points, resulting in a tradeoff between the quantity and quality of points

Another approach for discriminating between signal and noise photons could be to set a threshold based on the depth of the photon. In the bias plots of the lidar error for each site, the spread around the 1=1 line is greater for shallower parts of the nearshore zone. Because the attenuation of light energy increases with depth, one would normally expect less strong signal, and therefore lower accuracy, in deeper waters. However, this error could be an artifact of the KDE algorithm; sometimes photons are reflected in the first few meters of the water column, therefore the first few meters of the water column contain noise that is slightly biased toward the top of the water column. This noise is less dense than a seafloor return, but might sometimes be dense enough to reach the minimal density threshold and be misclassified as signal.

It is possible that the results could be further improved by building a larger dataset of validation sites, and trying to find globally optimum parameters that provide a good balance of excluding low accuracy points and maintaining as much quality data as possible.

5.3. 2D interpolation via universal kriging

The universal kriging approach proved effective to transform the point estimates into a gridded estimate of seabed elevation and uncertainty.

The largest practical limitation to the Bayesian updating method is the requirement to interpolate the data. The Kriging interpolater is ideal because they are robust to outliers and provide an uncertainty estimate as well as a depth estimate. The downside of the kriging interpolator is both the computational complexity and the requirement that points are not too close together. If points are too close to one another, the kriging matrix is not soluble, and the algorithm has a complexity of $O(n^4)$ where n is the number of points being interpolated. This means that there is a relatively strict practical limitation to the number of points that can be used as input to the interpolator. The upper limit with a laptop with 32GB RAM was found to be approximately 2000 points without exceeding the available memory. One way to deal with this is to use a tiling strategy, and repeat the process using tiles which contain 2000 bathymetric points or less. This would need to be done adaptively since the bathymetric points are not evenly distributed. Using a larger number of smaller sub-sites would take longer to process, but it would at least be feasible using a consumer-grade computer.

One disadvantage of ICESat-2 data is the uneven spatial distribution of the resulting bathymetric

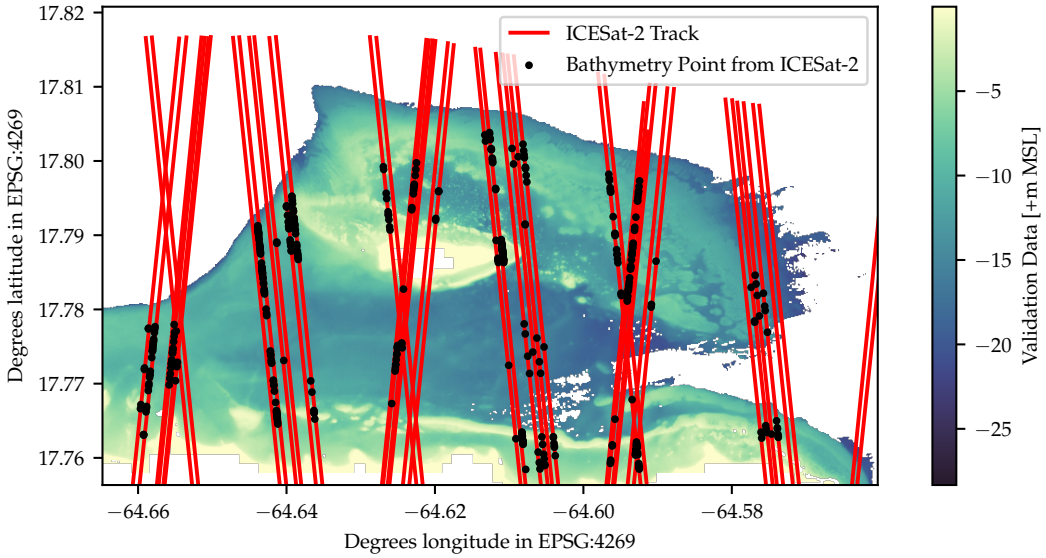


Figure 5.2: Spatial distribution of bathymetry points in St. Croix north end, and the true bathymetry. The red lines show transects. The north end of the transects in this figure are cut off by the request polygon. The black dots show the locations of bathymetric signal as identified by the KDE signal finding algorithm.

points. Due to the orbit of ICESat-2, the ATL03 data is available along transects oriented ±6° relative to north (see Figure 5.2). Because of this pattern, even if there was perfect bathymetry data along every single ICESat-2 transect within the study area, the spatial distribution of the points would be anisotropic. Given that there are gaps in these transects where the bathymetric signal is weaker, the spatial distribution of the data can vary significantly by site. An example of the gaps that can occur along transects can be seen in Figure 5.2. The gaps of weaker signal can be caused by areas that are too deep or too turbid for the lidar photon to reach the seabed, or due to instrument/atmospheric issues (see sec. 3.3.6).

However, handling this uneven spatial distribution is one of the advantages of using a geostatisical approach. Areas that are distant from any input points have a higher uncertainty in the output grid, and therefore there is little or no change from the prior estimate, since the Kalman update step accounts for the estimate uncertainty. Therefore, in locations that are further away from valid data, there is little or no change from the prior estimate. A limitation of this is that any accuracy gains are might not be evenly distributed across the site.

5.4. Kalman updating and improvement over GEBCO

It was found that in most sites, the Bayesian combination of bilinearly resampled GEBCO data and the interpolated ICESat-2 data increased the accuracy compared to the validation data than either product on their own. The decrease in RMSE achieved was up to 34% in some sites, but in other sites actually showed an increase. This could likely be mitigated by optimizing the assumptions behind the measurement uncertainties.

The optimal decrease in RMS error was found by assuming a GEBCO uncertainty of $\sigma = 1.5\ m$ standard deviation. This produced the best results for the test sites under consideration, however the actual GEBCO uncertainty depends on myriad factors including the original data source, survey method, and how old it is. One improvement that could be further explored is how the uncertainty of GEBCO varies with the Type IDentifier (TID) grid. The TID is a GEBCO data product that indicates the original source of the data. It is likely that the original data source affects the error between the GEBCO grid and the validation data.

One limitation is that the temporal resolution of the ICESat-2 data is not sufficient to take full

Site name	RMSE Change	MAE Change	ME Change
Florida Keys	34.41%	22.83%	456.68%
St. Croix	28.15%	26.59%	-4.74%
St. Thomas and St. John	21.41%	19.88%	194.37%
Oahu 1	28.43%	30.75%	73.46%
Oahu 2	31.62%	29.61%	51.10%
Oahu 3	25.18%	22.68%	75.06%
Oahu 4	-1.76%	-1.01%	123.78%
Oahu 5	1.94%	4.52%	116.88%
Oahu 6	-18.14%	-16.12%	-14.68%
Oahu 7	-22.61%	-15.04%	-16.58%
Oahu 8	-25.15%	-21.05%	-544.94%

Table 5.2: Percent reduction in error metrics via the Kalman updating approach. Positive values indicate reduced error, negative ones indicate increased error

advantage of the Kalman filtering approach, which normally includes updates both in space and time. The satellite repeat time for a given reference ground track is 91 days, and the data is available starting in 2018. Due to off-pointing over most land areas, there is often greater than 91 days between passes of the *exact same* track. Since coasts are among the most dynamic morphological environments, there might be substantial changes between repeat passes. The proposed method assumes that all ICESat-2 bathymetry data is contemporaneous. This is an inherent limitation to the method, since it cannot account for morphological changes that might occur during the 4-year time window in which the data was collected. However, the ICESat-2 data is very likely more recent than the input data to the GEBCO gridding process, which in some places can come from digitized nautical maps that can be hundreds of years old. Therefore, all data was assumed to be contemporaneous, and the time update step of the Kalman Filter was not used.

One potential downside of the Bayesian update approach is that it could create artifacts in the resulting output data, even if the RMS error is decreased relative to GEBCO. If there is an area with a high density of poor quality point measurements, the resulting uncertainty data might have a high confidence in that local area, leading to a significant update to the prior estimate which does not reflect a physical seabed feature. This effect could create artificial shoals or shallow areas. Because the RMS error is decreasing relative to the validation data, we know that the *net* effect of the Bayesian update is an improvement. However, the addition of phantom shoals could make the unsuitable for numerical modeling studies, since introducing an artificial high point offshore could change the wave transformation results significantly. However, the decreased average error of the downscaled product could have utility for estimating dredging quantities, which are less sensitive to an individual high point.

6

Conclusions

This section summarizes the findings after implementing the toolchain as described in section 3 as described.

6.1. Research Question

To answer the primary research question *How can spaceborne remote sensing data be combined with existing global datasets to improve estimates of nearshore bathymetry?*, a number of subquestions were pursued:

1. **How can ICESat-2 transects that contain bathymetry be identified algorithmically?**

   Currently, the KDE signal finding method employed can reliably identify transects which do not contain any bathymetric signal. However, it requires the processing chain to be fully run for the transect.

   For spaceborne lidar bathymetry to be practical at scale, there must be a way to identify coastal zones which are likely to have bathymetric signal to limit the amount of data downloading and processing that is required. Therefore further investigation of which metadata and ATL03 variables predict the availability of useful bathymetry would allow better filtering of transects before downloading.

   The NSIDC download API allows variable subsetting before download, so if there are a limited number of known variables that correlate strongly with good bathymetry signal in a transect, these few variables could be easily downloaded for large regions. Then, based on the properties of these variables, specific areas and transects that are the best candidates could be then downloaded in full.

2. **Once transects with bathymetry are found, how can nearshore subsurface photon returns that are located in the nearshore zone be extracted?**

   The filtering methodology employed allowed the generally reliable identification of subsurface returns, with a few notable edge cases. By first applying a horizontal filter based on a maximum and minimum GEBCO depth, photons that are likely located in the deep sea or on land can be removed from the dataset. Then, the probable sea surface for the transect is calculated based on the average elevation of the high-confidence ocean surface photons, and any photons with an elevation greater than one standard deviation below the sea surface are removed to eliminate the ocean surface signal. Then using the calculated sea surface, the depth of each remaining photon is calculated, and photons with a depth of greater than 40 m or a geoidal elevation of less than - removed.

   Any remaining photons that are greater than 5 m above the geoid are also removed, since that is above the tidal range for most of the world, and any remaining photons in this zone are not likely

to be located in the nearshore zone.

Once the filtering method is applied all remaining photons are assumed to be subsurface returns in the nearshore zone. Then, the refraction correction methodology from Parrish et al.(2019) is applied, using the calculated depth and the satellite orbit data as an input.

The filtering strategy based on GEBCO elevation did result in some issues in areas of steep topography, where the GEBCO resolution is sufficiently high to capture some steep mountains and sea cliffs, and land areas that should be masked out of the transect are inadvertently included in the subsurface photon set. This could be improved by using another source of land mask data to determine the inland limit of horizontal filtering, and using GEBCO data only for setting the offshore limit.

3. **How can lidar photon return locations reflecting the seafloor be separated from background noise?**

   The developed toolchain implements a method of isolating bathymetric data based on the local density of photons in the vertical direction. A rolling window function is applied longitudinally along the transect. For each window a kernel density estimator is used to find the density of the Z values within that window. Both the magnitude of the peak density and the Z value where the peak value occurs are recorded.

   Once the rolling window has been applied to the entire transect, the mean kernel density for all photons in the transect is calculated, and this is used as a threshold. Any points with a kernel density greater than the transect mean are considered to be bathymetric signal, and for these signal points, the Z value of the peak of magnitude is assigned to each photon as the uncorrected seafloor location.

   At the various test sites, the KDE method was able to estimate point bathymetry with an RMS error of between 0.49 m to 9 m, depending on site conditions. The sites with a higher error are due to issues in the filtering approach, and not directly from the KDE function.

4. **What is the potential to scale up bathymetry detection to a global scale to produce high-level processed bathymetry product using ICESat-2 data?**

   To be able to practically expand this data to create either a point product (like NASA L3A) or a gridded product (such as NASA L3B), an important prerequisite is finding ways of predicting which transects will contain bathymetric data. One easy way to do this is to only select from regions that are known to have very clear water. However, even for sites with clear water, many transects do not contain useful data. By finding granule-level metadata or atmospheric parameters which predict the presence of bathymetric data, it would significantly reduce the number of transects to be input into the KDE signal finding function. The ocean color of a site or transect was not found to be a good indicator of the presence of useful bathymetry.

5. **How can spaceborne remote sensing sources be used to improve existing global bathymetry datasets?**

   A kriging interpolator is used to resample the point estimates of bathymetry from the KDE algorithm to a continuous bathymetry grid, and a grid of the uncertainty. Areas that contain many bathymetry points in the local area have a lower uncertainty, whereas grid cells that are further away from any measured points with have a relatively higher uncertainty.

   To create an upscaled version of the global GEBCO data that incorporates the lidar data, the GEBCO data is first subset to the area of interest and then is resampled bilinearly to a grid in the local UTM coordinate system with the same resolution as the kriging output. Then, the bilinear data is updated using the Kalman update equation for each grid cell.

   By incorporating the lidar data, the RMSE between the sites can be reduced by up to approximately 30% at the sites tested.

6. **Under what conditions can remotely-sensed lidar data provide useful improvement on bathymetric data estimates?**

   By applying the method to the test sites where lidar data was available, the RMS error between the raw GEBCO data and high-accuracy validation data can be reduced by up to approximately

30% in clear water sites. This is a non-trivial decrease in the overall error. However, the constraint of requiring very clear water makes the method most useful for tropical waters, where satellite-derived SDB methods are also often applicable.

Based on the subquestions evaluated, ICESat-2 data can provide reliable bathymetry estimates along satellite transects in areas with sufficiently clear water. One method of using this data to upscale low-resolution global bathymetry data is to interpolate the points data along these tracks, and then produce a new, upscaled version of the low resolution data by combining the interpolated lidar data with a Kalman filter. This method can produce significant reductions in RMSE when compared to high resolution validation data.

However, the spatial distribution of ATL03 data can be a significant hinderance to the accuracy of the interpolated data. While this method is practical at the scale of a single site ($\sim$ 500 km^2) there are practical constraints on scaling up the approach to the regional or global scale. One is that downloading and finding the bathymetric signal in the L2 product requires non-trivial amounts of computing resources. Another is the practical computational expense of the kriging algorithm.

6.2. Recommendations

The dire need for bathymetry data is well-established in the literature. ICESat-2 has been shown to provide an extremely valuable source of accurate but sparse bathymetry data points. However, a significant bottleneck is the processing required to download and analyze the data. One of the best ways to promote further use of this data would be for NASA to provide a L3A and L3B data product of the bathymetric data.

In addition to the signal finding method proposed here, there have been a number of other methods in the literature to isolate bathymetric signal from ATL03 data. These methods are all based on the idea that increased density in the vertical direction indicates possible bathymetric signal. These algorithms are all fundamentally similar to the approaches developed by NASA for other L3A data products, like the global vegetation height data (ATL08), ocean elevation (ATL12), or inland water height (ATL13). NASA could facilitate collaborations between researchers working on on how to improve and scale these algorithms, and develop an Algorithm Theoretical Basis Document (ATBD) for bathymetry points in the same way that they already have for other higher level products. Such a collaborative effort between scientists studying this issue would likely lead to improvements in the signal extraction quality. This would also be practical to implement, since it would make use of the existing processing infrastructure already used to produce the other L3A and L3B products. The global surface area of shallow water zones is a small fraction of the earth's surface compared to ocean, forests, and ice cover, therefore the marginal increase in computational complexity by adding bathymetry processing would likely be vanishingly small. However, the benefit to geoscientists, physical oceanographers, and others would be revolutionary. A global analysis would likely reveal thousands of sites with bathymetric signal that have not been identified yet, even by the myriad papers that look at a single site at a time. Having such a global dataset would be an extremely valuable validation data source for many areas of coastal research.

Bibliography

Abshire, J. B., X. Sun, H. Riris, J. M. Sirota, J. F. McGarry, S. Palm, D. Yi, and P. Liiva (Nov. 2005). "Geoscience Laser Altimeter System (GLAS) on the ICESat Mission: On-orbit measurement performance". In: *Geophysical Research Letters* 32.21, pp. 1–4. ISSN: 1944-8007. DOI: 10.1029/2005GL024028. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2005GL024028.

Albright, A. and C. Glennie (May 2021). "Nearshore Bathymetry from Fusion of Sentinel-2 and ICESat-2 Observations". In: *IEEE Geoscience and Remote Sensing Letters* 18.5, pp. 900–904. ISSN: 15580571. DOI: 10.1109/LGRS.2020.2987778.

Almar, R., E. W. Bergsma, G. Thoumyre, M. W. Baba, G. Cesbron, C. Daly, T. Garlan, and A. Lifermann (Nov. 2021). "Global Satellite-Based Coastal Bathymetry from Waves". In: *Remote Sensing 2021, Vol. 13, Page 4628* 13.22, p. 4628. ISSN: 2072-4292. DOI: 10.3390/RS13224628. URL: https://www.mdpi.com/2072-4292/13/22/4628/htm.

Ashphaq, M., P. K. Srivastava, and D. Mitra (Dec. 2021). "Review of near-shore satellite derived bathymetry: Classification and account of five decades of coastal bathymetry research". In: *Journal of Ocean Engineering and Science* 6.4, pp. 340–359. ISSN: 2468-0133. DOI: 10.1016/J.JOES.2021.02.006.

Babbel, B. J., C. E. Parrish, and L. A. Magruder (Mar. 2021). "ICESat-2 Elevation Retrievals in Support of Satellite-Derived Bathymetry for Global Science Applications". en. In: *Geophysical Research Letters* 48.5, pp. 1–9. ISSN: 19448007. DOI: 10.1029/2020GL090629. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2020GL090629.

Bailly, J.-S., M. Montes-Hugo, Y. Pastol, and N. Baghdadi (Jan. 2016). "LiDAR Measurements and Applications in Coastal and Continental Waters". In: *Land Surface Remote Sensing in Urban and Coastal Areas*. Elsevier, pp. 185–229. DOI: 10.1016/B978-1-78548-160-4.50005-4. URL: https://linkinghub.elsevier.com/retrieve/pii/B9781785481604500054.

Butler, H., A. Bell, M. P. Gerlek, chambbj, P. Gadomski, C. Manning, M. Łoskot, B. Couwenberg, P. Ramsey, N. Barker, N. Chaulet, C. Foster, E. Rouault, M. Rosen, S. Lewis, Grigory, G. Villemin, K. McKelvey, K. Evers, V. Vergara, A. Yonas, V. Petras, xantares, D. Brookes, M. Visser, R. Pals, and É. Lemoine (Mar. 2022). *PDAL/PDAL: 2.4.0RC1*. Version 2.4.0RC1. DOI: 10.5281/zenodo.6369164. URL: https://doi.org/10.5281/zenodo.6369164.

Cao, B., Y. Fang, L. Gao, H. Hu, Z. Jiang, B. Sun, and L. Lou (Apr. 2021). "An active-passive fusion strategy and accuracy evaluation for shallow water bathymetry based on ICESat-2 ATLAS laser point cloud and satellite remote sensing imagery". In: *International Journal of Remote Sensing* 42.8, pp. 2783–2806. ISSN: 0143-1161. DOI: 10.1080/01431161.2020.1862441. URL: https://www.tandfonline.com/doi/full/10.1080/01431161.2020.1862441.

Cesbron, G., A. Melet, R. Almar, A. Lifermann, D. Tullot, and L. Crosnier (2021). "Pan-European Satellite-Derived Coastal Bathymetry-Review, User Needs and Future Services". In: DOI: 10.3389/fmars.2021.740830. URL: www.frontiersin.org.

Coveney, S., X. Monteys, J. D. Hedley, Y. Castillo-Campo, and B. Kelleher (Oct. 2021). "ICESat-2 Marine Bathymetry: Extraction, Refraction Adjustment and Vertical Accuracy as a Function of Depth in Mid-Latitude Temperate Contexts". In: *Remote Sensing 2021, Vol. 13, Page 4352* 13.21, p. 4352. ISSN: 2072-4292. DOI: 10.3390/rs13214352. URL: https://www.mdpi.com/2072-4292/13/21/4352.

Daly, C., W. Baba, E. Bergsma, G. Thoumyre, R. Almar, and T. Garlan (Sept. 2022). "The new era of regional coastal bathymetry from space: A showcase for West Africa using optical Sentinel-2 imagery". In: *Remote Sensing of Environment* 278, p. 113084. ISSN: 0034-4257. DOI: 10.1016/J.RSE.2022.113084.

Datta, R. T. and B. Wouters (2021). "Supraglacial lake bathymetry automatically derived from ICESat-2 constraining lake depth estimates from multi-source satellite imagery". In: *The Cryosphere* 15, pp. 5115–5132. DOI: 10.5194/tc-15-5115-2021. URL: https://doi.org/10.5194/tc-15-5115-2021.

Ester, M., H.-P. Kriegel, J. Sander, and X. Xu (1996). "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: *KDD-96 Proceedings*. URL: www.aaai.org.

Ferreira, I. O., D. D. Rodrigues, G. R. dos Santos, and L. M. F. Rosa (Sept. 2017). "IN BATHYMETRIC SURFACES: IDW OR KRIGING?" In: *Boletim de Ciências Geodésicas* 23.3, pp. 493–508. ISSN: 1982-2170. DOI: 10.1590/s1982-21702017000300033. URL: http://www.scielo.br/scielo.php?script=sci%7B%5C_%7Darttext%7B%5C&%7Dpid=S1982-21702017000300493%7B%5C&%7Dlng=en%7B%5C&%7Dtlng=en.

Forfinski-Sarkozi, N. and C. E. Parrish (Sept. 2016). "Analysis of MABEL Bathymetry in Keweenaw Bay and Implications for ICESat-2 ATLAS". In: *Remote Sensing* 8.9, p. 772. DOI: 10.3390/rs8090772. URL: http://www.mdpi.com/2072-4292/8/9/772.

Garnesson, P., A. Mangin, O. F. D'Andon, J. Demaria, and M. Bretagnon (June 2019). "The CMEMS GlobColour chlorophyll a product based on satellite observation: Multi-sensor merging and flagging strategies". In: *Ocean Science* 15.3, pp. 819–830. ISSN: 18120792. DOI: 10.5194/OS-15-819-2019.

GEBCO Bathymetric Compilation Group 2021 (July 2021). *The GEBCO_2021 Grid - a continuous terrain model of the global oceans and land*. DOI: 10.5285/c6612cbe-50b3-0cff-e053-6c86abc09f8f. URL: https://www.bodc.ac.uk/data/published%7B%5C_%7Ddata%7B%5C_%7Dlibrary/catalogue/10.5285/c6612cbe-50b3-0cff-e053-6c86abc09f8f/.

Geyman, E. C. and A. C. Maloof (Mar. 2019). "A Simple Method for Extracting Water Depth From Multispectral Satellite Imagery in Regions of Variable Bottom Type". In: *Earth and Space Science* 6.3, pp. 527–537. ISSN: 2333-5084. DOI: 10.1029/2018EA000539. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2018EA000539%20https://onlinelibrary.wiley.com/doi/abs/10.1029/2018EA000539%20https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2018EA000539.

Gleason, A. C., R. Smith, S. J. Purkis, K. Goodrich, A. Dempsey, and A. Mantero (Oct. 2021). "The Prospect of Global Coral Reef Bathymetry by Combining Ice, Cloud, and Land Elevation Satellite-2 Altimetry With Multispectral Satellite Imagery". In: *Frontiers in Marine Science* 8, p. 1529. ISSN: 22967745. DOI: 10.3389/fmars.2021.694783.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore (Dec. 2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone". en. In: *Remote Sensing of Environment*. Big Remotely Sensed Data: tools, applications and experiences 202, pp. 18–27. ISSN: 00344257. DOI: 10.1016/j.rse.2017.06.031. URL: https://www.sciencedirect.com/science/article/pii/S0034425717302900.

Holman, R., N. Plant, and T. Holland (May 2013). "cBathy: A robust algorithm for estimating nearshore bathymetry". In: *Journal of Geophysical Research: Oceans* 118.5, pp. 2595–2609. ISSN: 2169-9291. DOI: 10.1002/JGRC.20199. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/jgrc.20199.

Hoyer, S. and J. Hamman (2017). "xarray: N-D labeled arrays and datasets in Python". In: *Journal of Open Research Software* 5.1. DOI: 10.5334/jors.148. URL: https://doi.org/10.5334/jors.148.

Hoyer, S., M. Roos, H. Joseph, J. Magin, D. Cherian, C. Fitzgerald, M. Hauser, K. Fujii, F. Maussion, G. Imperiale, S. Clark, A. Kleeman, T. Nicholas, T. Kluyver, J. Westling, J. Munroe, A. Amici, A. Barghini, A. Banihirwe, R. Bell, Z. Hatfield-Dodds, R. Abernathey, B. Bovy, J. Omotani, K. Mühlbauer, M. K. Roszko, and P. J. Wolfram (Mar. 2022). *xarray*. Version v2022.03.0. If you use this software, please cite it as below. DOI: 10.5281/zenodo.6323468. URL: https://doi.org/10.5281/zenodo.6323468.

Hsu, H. J., C. Y. Huang, M. Jasinski, Y. Li, H. Gao, T. Yamanokuchi, C. G. Wang, T. M. Chang, H. Ren, C. Y. Kuo, and K. H. Tseng (Aug. 2021). "A semi-empirical scheme for bathymetric mapping in shallow water by ICESat-2 and Sentinel-2: A case study in the South China Sea". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 178, pp. 1–19. ISSN: 0924-2716. DOI: 10.1016/J.ISPRSJPRS.2021.05.012.

Hughes, C. W. and R. J. Bingham (2008). "An Oceanographer's Guide to GOCE and the Geoid". In: *Ocean Sci* 4, pp. 15–29. URL: www.ocean-sci.net/4/15/2008/.

IHO (2019). *The IHO-IOC GEBCO Cook Book*. Tech. rep. Monaco: International Hydrographic Organization, Intergovernmental Oceanographic Commission. URL: http://www.ngdc.noaa.gov.

IHO/OHI (2022). *INTERNATIONAL HYDROGRAPHIC ORGANIZATION ORGANISATION HYDROGRAPHIQUE INTERNATIONALE ORGANIZACION HIDROGRAFICA INTERNACIONAL IHO/OHI Publication C-55 Status of Hydrographic Surveying and Charting Worldwide Etat des levés hydrographiques et de la cartographi*. Tech. rep. IHO.

Jasinski, M. F., J. D. Stoll, W. B. Cook, M. Ondrusek, E. Stengel, and K. Brunt (Dec. 2016). "Inland and Near-Shore Water Profiles Derived from the High-Altitude Multiple Altimeter Beam Experimental Lidar

(MABEL)". In: *Journal of Coastal Research* 76.sp1, pp. 44–55. ISSN: 0749-0208. DOI: 10.2112/SI76-005. URL: http://www.bioone.org/doi/10.2112/SI76-005.

Le Provost, C., M. L. Genco, F. Lyard, P. Vincent, and P. Canceil (1994). "Spectroscopy of the world ocean tides from a finite element hydrodynamic model". In: *Journal of Geophysical Research* 99.C12. ISSN: 01480227. DOI: 10.1029/94JC01381.

Le Quilleuc, A., A. Collin, M. F. Jasinski, and R. Devillers (Jan. 2022). "Very high-resolution satellite-derived bathymetry and habitat mapping using pleiades-1 and icesat-2". en. In: *Remote Sensing* 14.1, p. 133. ISSN: 20724292. DOI: 10.3390/rs14010133. URL: https://www.mdpi.com/2072-4292/14/1/133.

Lee, Z., M. Shangguan, R. A. Garcia, W. Lai, X. Lu, J. Wang, and X. Yan (Apr. 2021). "Confidence Measure of the Shallow-Water Bathymetry Map Obtained through the Fusion of Lidar and Multiband Image Data". In: *Journal of Remote Sensing* 2021, pp. 1–16. DOI: 10.34133/2021/9841804.

Lichter, M., A. T. Vafeidis, and R. J. Nicholls (Oct. 2010). "Exploring Data-Related Uncertainties in Analyses of Land Area and Population in the "Low-Elevation Coastal Zone" (LECZ)". In: *Journal of Coastal Research* 27.4, p. 757. ISSN: 0749-0208. DOI: 10.2112/JCOASTRES-D-10-00072.1. URL: https://bioone.org/journals/journal-of-coastal-research/volume-27/issue-4/JCOASTRES-D-10-00072.1/Exploring-Data-Related-Uncertainties-in-Analyses-of-Land-Area-and/10.2112/JCOASTRES-D-10-00072.1.full.

Liu, C., J. Qi, J. Li, Q. Tang, W. Xu, X. Zhou, and W. Meng (Oct. 2021). "Accurate Refraction Correction—Assisted Bathymetric Inversion Using ICESat-2 and Multispectral Data". In: *Remote Sensing 2021, Vol. 13, Page 4355* 13.21, p. 4355. ISSN: 2072-4292. DOI: 10.3390/RS13214355. URL: https://www.mdpi.com/2072-4292/13/21/4355/htm.

Lu, X., Y. Hu, Y. Yang, T. Neumann, A. Omar, R. Baize, M. Vaughan, S. Rodier, B. Getzewich, P. Lucker, C. Trepte, C. Hostetler, and D. Winker (Oct. 2021). "New Ocean Subsurface Optical Properties From Space Lidars: CALIOP/CALIPSO and ATLAS/ICESat-2". In: *Earth and Space Science* 8.10, e2021EA001839. ISSN: 2333-5084. DOI: 10.1029/2021EA001839. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2021EA001839.

Ma, Y., N. Xu, Z. Liu, B. Yang, F. Yang, X. H. Wang, and S. Li (Dec. 2020). "Satellite-derived bathymetry using the ICESat-2 lidar and Sentinel-2 imagery datasets". en. In: *Remote Sensing of Environment* 250, p. 112047. URL: https://www.sciencedirect.com/science/article/pii/S003442572030417X.

Magdalena, I., V. Kusnowo, M. I. Azis, and Widowati (June 2021). "1D–2D Numerical Model for Wave Attenuation by Mangroves as a Porous Structure". In: *Computation 2021, Vol. 9, Page 66* 9.6, p. 66. ISSN: 2079-3197. DOI: 10.3390/COMPUTATION9060066. URL: https://www.mdpi.com/2079-3197/9/6/66/htm.

Magruder, L., T. A. Neumann, and N. Kurtz (May 2021). "ICESat-2 Early Mission Synopsis and Observatory Performance". In: *Earth and Space Science* 8.5. DOI: 10.1029/2020EA001555.

Magruder, L. A., K. M. Brunt, and M. Alonzo (Nov. 2020). "Early ICESat-2 on-orbit Geolocation Validation Using Ground-Based Corner Cube Retro-Reflectors". In: *Remote Sensing 2020, Vol. 12, Page 3653* 12.21, p. 3653. ISSN: 2072-4292. DOI: 10.3390/RS12213653. URL: https://www.mdpi.com/2072-4292/12/21/3653/htm%20https://www.mdpi.com/2072-4292/12/21/3653.

Mäkinen, J. and J. Ihde (2009). "The Permanent Tide In Height Systems". In: *International Association of Geodesy Symposia* 133, pp. 81–87. ISSN: 09399585. DOI: 10.1007/978-3-540-85426-5_10. URL: https://link.springer.com/chapter/10.1007/978-3-540-85426-5%7B%5C_%7D10.

Markus, T., T. Neumann, A. Martino, W. Abdalati, K. Brunt, B. Csatho, S. Farrell, H. Fricker, A. Gardner, D. Harding, M. Jasinski, R. Kwok, L. Magruder, D. Lubin, S. Luthcke, J. Morison, R. Nelson, A. Neuenschwander, S. Palm, S. Popescu, C. K. Shum, B. E. Schutz, B. Smith, Y. Yang, and J. Zwally (Mar. 2017). "The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation". In: *Remote Sensing of Environment* 190, pp. 260–273. ISSN: 00344257. DOI: 10.1016/j.rse.2016.12.029.

McCool, M. and E. Fiume (1992). "Hierarchical Poisson disk sampling distributions". In: *Graphics Interface*. URL: http://www.dgp.toronto.edu/%7B~%7Delf/.misc/poissondisk.pdf.

Mcgill, M., T. Markus, S. S. Scott, and T. Neumann (Feb. 2013). "The Multiple Altimeter Beam Experimental Lidar (MABEL): An Airborne Simulator for the ICESat-2 Mission". In: *Journal of Atmospheric and Oceanic Technology* 30.2, pp. 345–352. ISSN: 0739-0572. DOI: 10.1175/JTECH-D-12-00076.1. URL: https://journals.ametsoc.org/view/journals/atot/30/2/jtech-d-12-00076%7B%5C_%7D1.xml.

McKinney, W. (2010). "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by S. van der Walt and J. Millman, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

Minneboo, F. (1995). "Jaarlijkse kustmetingen. Richtlijnen voor de inwinning, bewerking en opslag van gegevens van jaarlijkse kustmetingen. Rapport RIKZ-95.022". In: RIKZ-95.022, p. 22. URL: http://resolver.tudelft.nl/uuid:76f2634d-f3c4-4609-aa4c-44d641da28f1.

Monteys, X., P. Harris, S. Caloca, C. Cahalane, X. Li, M. Koch, and P. S. Thenkabail (Oct. 2015). "Spatial Prediction of Coastal Bathymetry Based on Multispectral Satellite Imagery and Multibeam Data". In: *Remote Sensing 2015, Vol. 7, Pages 13782-13806* 7.10, pp. 13782–13806. ISSN: 2072-4292. DOI: 10.3390/RS71013782. URL: https://www.mdpi.com/2072-4292/7/10/13782/htm%20https://www.mdpi.com/2072-4292/7/10/13782.

Murphy, B., S. Müller, and R. Yurchak (Sept. 2021). *GeoStat-Framework/PyKrige: v1.6.1*. Version v1.6.1. DOI: 10.5281/zenodo.5380342. URL: https://doi.org/10.5281/zenodo.5380342.

NASA NSIDC (2022). *ATL03 Known Issues*. Tech. rep.

Neuenschwander, A. L. and L. A. Magruder (July 2019). "Canopy and Terrain Height Retrievals with ICESat-2: A First Look". In: *Remote Sensing 2019, Vol. 11, Page 1721* 11.14, p. 1721. ISSN: 2072-4292. DOI: 10.3390/RS11141721. URL: https://www.mdpi.com/2072-4292/11/14/1721/htm.

Neumann, B., A. T. Vafeidis, J. Zimmermann, and R. J. Nicholls (Mar. 2015). "Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment". In: *PLOS ONE* 10.3, e0118571. ISSN: 1932-6203. DOI: 10.1371/JOURNAL.PONE.0118571. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118571.

Neumann, T., A. Brenner, D. Hancock, J. Robbins, J. Saba, K. Harbeck, A. Gibbons, J. Lee, S. Luthcke, and T. Rebold (2021). *ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 5. ATL03*. Boulder, Colorado USA. DOI: 10.5067/ATLAS/ATL03.005.

Neumann, T. A., A. Brenner, D. Hancock, J. Robbins, J. Saba, K. Harbeck, A. Gibbons, J. Lee, S. B. Luthcke, K. Harbeck, J. Lee, A. Gibbons, J. Saba, and K. Brunt (2019). "ATLAS/ICESat-2 L2A Global Geolocated Photon Data User Guide". en. In: DOI: 10.5067/ATLAS/ATL03.005. URL: http://nsidc.org/data/atl03/versions/1%20https://doi.org/10.5067/ATLAS/ATL03.005.

Neumann, T. A., A. J. Martino, T. Markus, S. Bae, M. R. Bock, A. C. Brenner, K. M. Brunt, J. Cavanaugh, S. T. Fernandes, D. W. Hancock, K. Harbeck, J. Lee, N. T. Kurtz, P. J. Luers, S. B. Luthcke, L. Magruder, T. A. Pennington, L. Ramos-Izquierdo, T. Rebold, J. Skoog, and T. C. Thomas (Nov. 2019). "The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System". In: *Remote Sensing of Environment* 233, p. 111325. ISSN: 00344257. DOI: 10.1016/j.rse.2019.111325. URL: https://linkinghub.elsevier.com/retrieve/pii/S003442571930344X.

Neumann, T., A. Brenner, D. Hancock, J. Robbins, J. Saba, and K. Harbeck (2019). *Ice, Cloud, and Land Elevation Satellite–2 (ICESat-2) Project: Algorithm Theoretical Basis Document (ATBD) for Global Geolocated Photons (ATL03)*. Tech. rep., p. 207.

NOAA (2020). *Puerto Rico and US Virgin Islands Topobathy Final Report of Survey Produced for National Oceanic and Atmospheric Administration, Office for Coastal Management*. Tech. rep.

Parrish, C. E., L. A. Magruder, A. L. Neuenschwander, N. Forfinski-Sarkozi, M. Alonzo, and M. Jasinski (July 2019). "Validation of ICESat-2 ATLAS Bathymetry and Analysis of ATLAS's Bathymetric Mapping Performance". en. In: *Remote Sensing* 11.14, p. 1634. ISSN: 2072-4292. DOI: 10.3390/rs11141634. URL: https://www.mdpi.com/2072-4292/11/14/1634.

Pike, S., D. Traganos, D. Poursanidis, J. Williams, K. Medcalf, P. Reinartz, and N. Chrysoulakis (2019). "Leveraging commercial high-resolution multispectral satellite and multibeam sonar data to estimate bathymetry: The case study of the Caribbean Sea". In: *Remote Sensing* 11.15, pp. 1–16. ISSN: 20724292. DOI: 10.3390/rs11151830.

Popescu, S. C., T. Zhou, R. Nelson, A. Neuenschwander, R. Sheridan, L. Narine, and K. M. Walsh (Apr. 2018). "Photon counting LiDAR: An adaptive ground and canopy height retrieval algorithm for ICESat-2 data". In: *Remote Sensing of Environment* 208, pp. 154–170. ISSN: 0034-4257. DOI: 10.1016/J.RSE.2018.02.019.

Randal, H., P. S. Christiansen, P. Kliving, O. B. Andersen, K. Nielsen, P. Sigaard Christiansen, P. Kliving, O. Baltazar Andersen, K. Nielsen, P. S. Christiansen, P. Kliving, O. B. Andersen, and K. Nielsen (Sept. 2021). "Evaluation of a Statistical Approach for Extracting Shallow Water Bathymetry Signals

from ICESat-2 ATL03 Photon Data". en. In: *Remote Sensing* 13.17, p. 3548. ISSN: 2072-4292. DOI: 10.3390/rs13173548. URL: https://www.mdpi.com/2072-4292/13/17/3548.

Reback, J., jbrockmendel, W. McKinney, J. V. den Bossche, T. Augspurger, M. Roeschke, S. Hawkins, P. Cloud, gfyoung, Sinhrks, P. Hoefler, A. Klein, T. Petersen, J. Tratner, C. She, W. Ayd, S. Naveh, J. Darbyshire, M. Garcia, R. Shadrach, J. Schendel, A. Hayden, D. Saxton, M. E. Gorelli, F. Li, M. Zeitlin, V. Jancauskas, A. McMaster, T. Wörtwein, and P. Battiston (Apr. 2022). *pandas-dev/pandas: Pandas 1.4.2*. Version v1.4.2. DOI: 10.5281/zenodo.6408044. URL: https://doi.org/10.5281/zenodo.6408044.

Robbins, J., T. Neumann, N. Kurtz, K. Brunt, M. Bagnardi, D. Hancock, and J. Lee (2022). *ICESat-2 Data Comparison User's Guide Release 5 ICESat-2 Data Comparison User's Guide for Rel005*. Tech. rep.

Rouault, E., F. Warmerdam, K. Schwehr, A. Kiselev, H. Butler, M. Łoskot, T. Szekeres, E. Tourigny, M. Landa, I. Miara, B. Elliston, C. Kumar, L. Plesea, D. Morissette, A. Jolma, and N. Dawson (Mar. 2022). *GDAL*. Version v3.4.2. DOI: 10.5281/zenodo.6352176. URL: https://doi.org/10.5281/zenodo.6352176.

Scott, D. W. (2012). *Multivariate density estimation and visualization*. John Wiley & Sons, pp. 549–569. ISBN: 9783642215513. DOI: 10.1007/978-3-642-21551-3__19.

Surisetty, V. V. K., P. Rajput, R. Ramakrishnan, and C. Venkateswarlu (2022). "Synergistic Fusion of ICESat-2 Lidar and Sentinel-2 Data to Leverage Potential Mapping of Bathymetry in Remote Islands Using SVR". In: *Journal of the Indian Society of Remote Sensing*. ISSN: 09743006. DOI: 10.1007/s12524-022-01537-4.

Thomas, N., A. P. Pertiwi, D. Traganos, D. Lagomasino, D. Poursanidis, S. Moreno, and L. Fatoyinbo (Mar. 2021). "Space-Borne Cloud-Native Satellite-Derived Bathymetry (SDB) Models Using ICESat-2 And Sentinel-2". en. In: *Geophysical Research Letters* 48.6, e2020GL092170. ISSN: 1944-8007. DOI: 10.1029/2020GL092170. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL092170.

Thomas, N., B. Lee, O. Coutts, P. Bunting, D. Lagomasino, and L. Fatoyinbo (2022). "A Purely Spaceborne Open Source Approach for Regional Bathymetry Mapping". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–9. ISSN: 0196-2892. DOI: 10.1109/TGRS.2022.3192825. URL: https://ieeexplore.ieee.org/document/9834969/.

Turner, I. L., M. D. Harley, R. Almar, and E. W. Bergsma (Aug. 2021). "Satellite optical imagery in Coastal Engineering". In: *Coastal Engineering* 167, p. 103919. ISSN: 0378-3839. DOI: 10.1016/J.COASTALENG.2021.103919.

Vitousek, S. and C. H. Fletcher (Oct. 2008). "Maximum Annually Recurring Wave Heights in Hawai'i1". In: *https://doi.org/10.2984/1534-6188(2008)62[541:MARWHI]2.0.CO;2* 62.4, pp. 541–553. ISSN: 0030-8870. DOI: 10.2984/1534-6188(2008)62. URL: https://bioone.org/journals/pacific-science/volume-62/issue-4/1534-6188%7B%5C_%7D2008%7B%5C_%7D62%7B%5C_%7D541%7B%5C_%7DMARWHI%7B%5C_%7D2.0.CO%7B%5C_%7D2/Maximum-Annually-Recurring-Wave-Heights-in-Hawaii1/10.2984/1534-6188(2008)62[541:MARWHI]2.0.CO;2.full.

Welch, G. F. (2021). "Kalman Filter". In: DOI: 10.1007/978-3-030-03243-2_716-1. URL: https://doi.org/10.1007/978-3-030-03243-2%7B%5C_%7D716-1.

Xie, C., P. Chen, D. Pan, C. Zhong, and Z. Zhang (Oct. 2021). "Improved Filtering of ICESat-2 Lidar Data for Nearshore Bathymetry Estimation Using Sentinel-2 Imagery". In: *Remote Sensing 2021, Vol. 13, Page 4303* 13.21, p. 4303. ISSN: 2072-4292. DOI: 10.3390/RS13214303. URL: https://www.mdpi.com/2072-4292/13/21/4303.

Xu, Q., H. Xie, Y. Sun, X. Liu, Y. Guo, P. Huang, B. Li, and X. Tong (May 2022). "AUTO-ADAPTIVE MULTI-LEVEL SEAFLOOR RECOGNITION AND LAND SEA CLASSIFICATION (AMSRLC) IN REEF-ISLAND ZONES USING ICESAT-2 LASER ALTIMETRY". In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLIII-B2-2, pp. 309–314. ISSN: 2194-9034. DOI: 10.5194/ISPRS-ARCHIVES-XLIII-B2-2022-309-2022. URL: https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLIII-B2-2022/309/2022/.

Zhang, D., Y. Chen, Y. Le, Y. Dong, G. Dai, and L. Wang (Apr. 2022). "Refraction and coordinate correction with the JONSWAP model for ICESat-2 bathymetry". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 186, pp. 285–300. ISSN: 0924-2716. DOI: 10.1016/J.ISPRSJPRS.2022.02.020.

Zheng, H., Y. Ma, Y. Ma, J. Huang, J. Yang, D. Su, F. Yang, and X. H. Wang (Aug. 2022). "Deriving vertical profiles of chlorophyll-a concentration in the upper layer of seawaters using ICESat-2 photon-counting lidar". In: *Optics Express, Vol. 30, Issue 18, pp. 33320-33336* 30.18, pp. 33320–33336. ISSN: 1094-4087. DOI: 10.1364/OE.463622. URL: https://opg.optica.org/viewmedia.cfm?uri=oe-30-18-33320%7B%5C&%7Dseq=0%7B%5C&%7Dhtml=true.

Zhong, J., J. Sun, Z. Lai, and Y. Song (Aug. 2022). "Nearshore Bathymetry from ICESat-2 LiDAR and Sentinel-2 Imagery Datasets Using Deep Learning Approach". In: *Remote Sensing 2022, Vol. 14, Page 4229* 14.17, p. 4229. ISSN: 2072-4292. DOI: 10.3390/RS14174229. URL: https://www.mdpi.com/2072-4292/14/17/4229/htm.

A

Data availability

All code required to reproduce the analysis can be found at https://github.com/mlinds/thesis

B

Raster error metrics for all Oahu subsites

The main body includes only a summary of the changes in the error metrics due to the Kalman updating approach.

sitename	Error Type	RMSE [m]	MAE [m]	ME [m]
1	GEBCO	5.00	3.52	2.64
	Kriging Surface	10.37	6.19	-4.94
	Kalman Output	3.56	2.45	0.71
2	GEBCO	5.94	4.11	3.92
	Kriging Surface	17.42	11.74	-11.38
	Kalman Output	4.06	2.89	1.92
3	GEBCO	3.73	2.38	1.61
	Kriging Surface	17.41	11.61	-11.29
	Kalman Output	2.79	1.84	0.40
4	GEBCO	3.95	2.72	0.75
	Kriging Surface	17.35	11.56	-10.98
	Kalman Output	4.02	2.75	-0.18
5	GEBCO	3.42	2.50	1.50
	Kriging Surface	17.63	11.91	-11.66
	Kalman Output	3.36	2.38	-0.25
6	GEBCO	5.24	4.19	-3.03
	Kriging Surface	17.67	12.02	-11.72
	Kalman Output	6.19	4.87	-3.47
7	GEBCO	4.65	3.40	-2.85
	Kriging Surface	17.81	12.11	-11.94
	Kalman Output	5.70	3.91	-3.32
8	GEBCO	4.73	3.34	-0.36
	Kriging Surface	16.99	11.39	-11.06
	Kalman Output	5.92	4.04	-2.29

Table B.1: Error metrics between ICESat-2 and ground-truth data for all sites in Oahu

C

Detailed output by test site

This appendix provides larger copy of the output images for all sites, as well as maps of the sites with aerial imagery and the locations of the ICESat-2 tracklines that were analyzed for each site.

C.1. St. Thomas/St. John test site

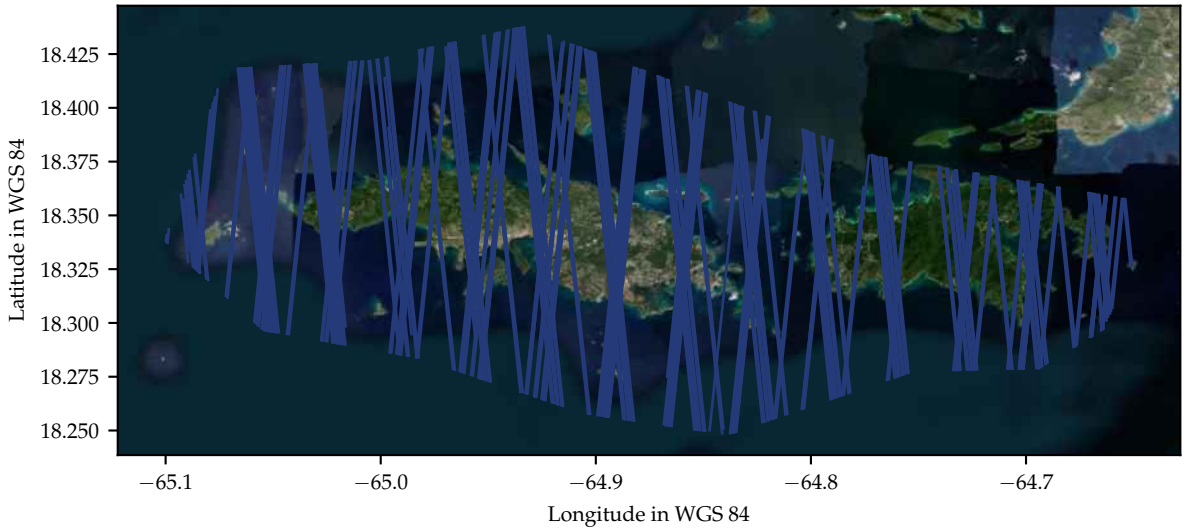


Figure C.1: St. Thomas/St. John test site: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

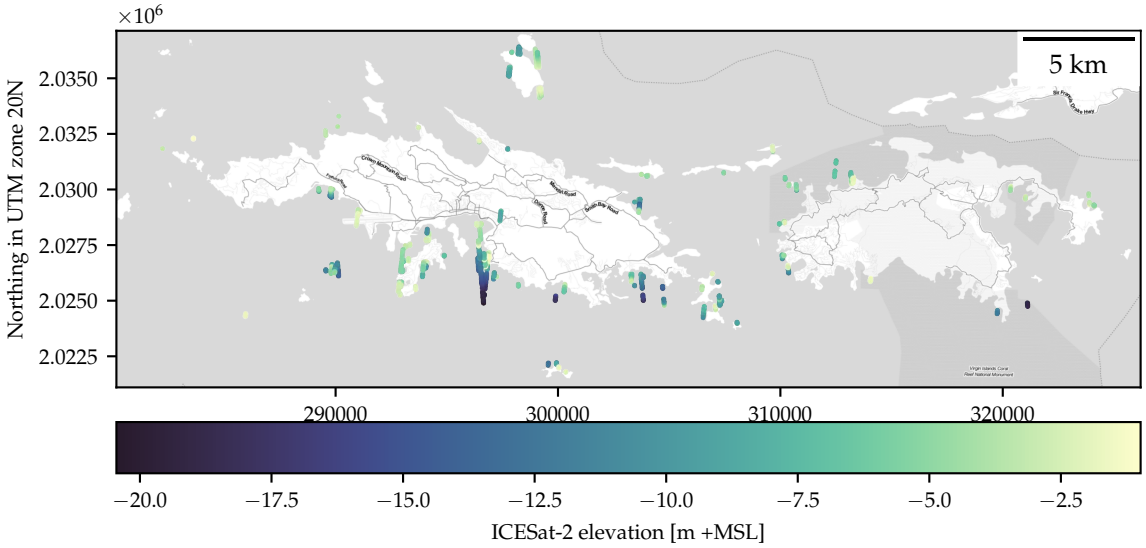


Figure C.2: St. Thomas/St. John test site: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

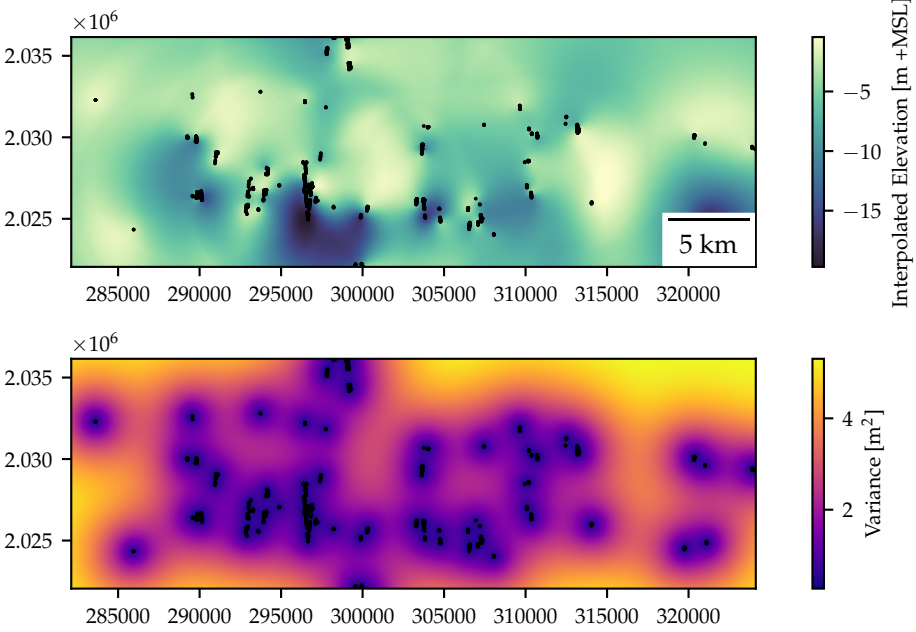


Figure C.3: St. Thomas/St. John test site: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

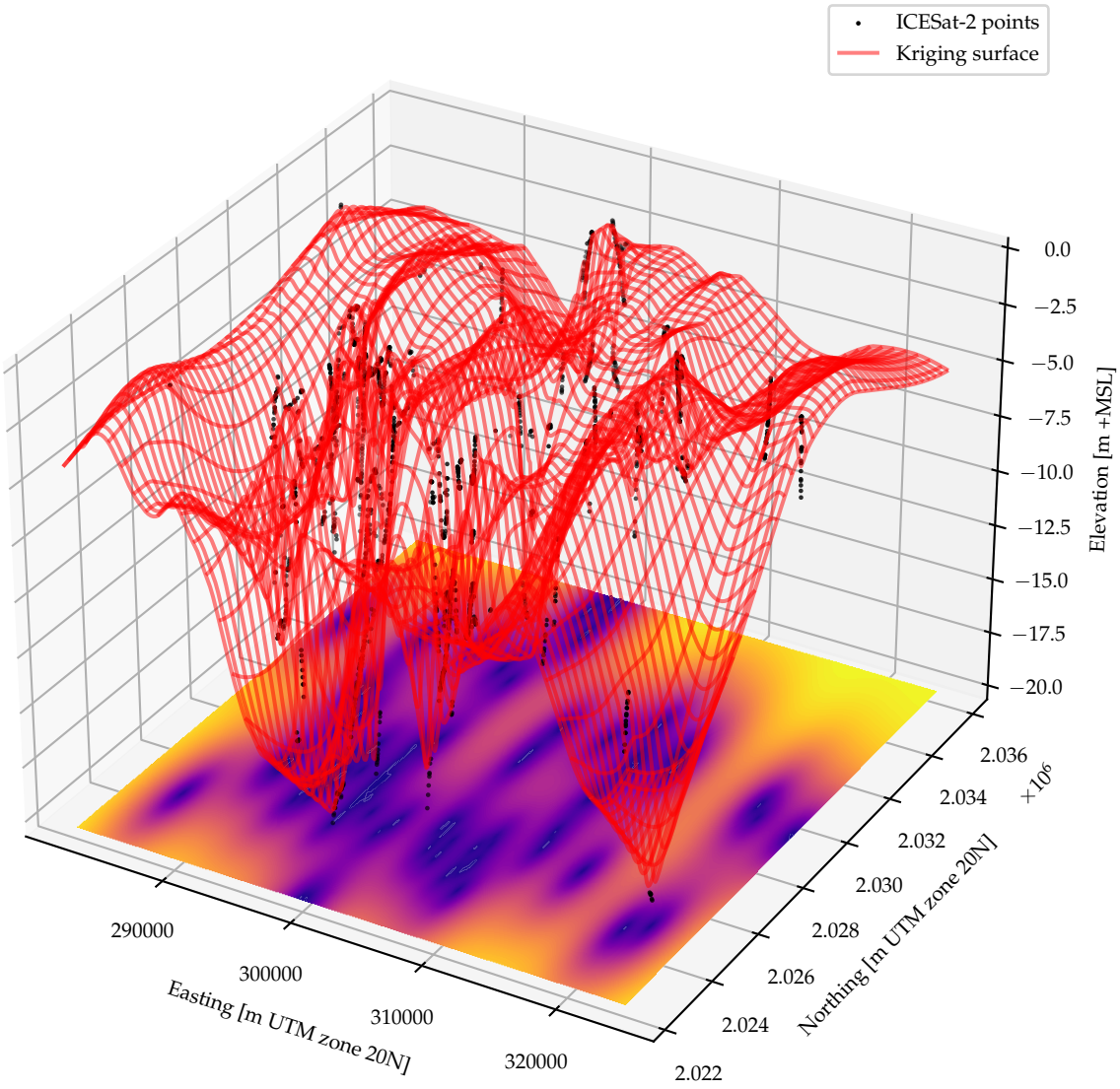


Figure C.4: St. Thomas/St. John test site: bathymetry points from KDE signal finding. Basemap data:

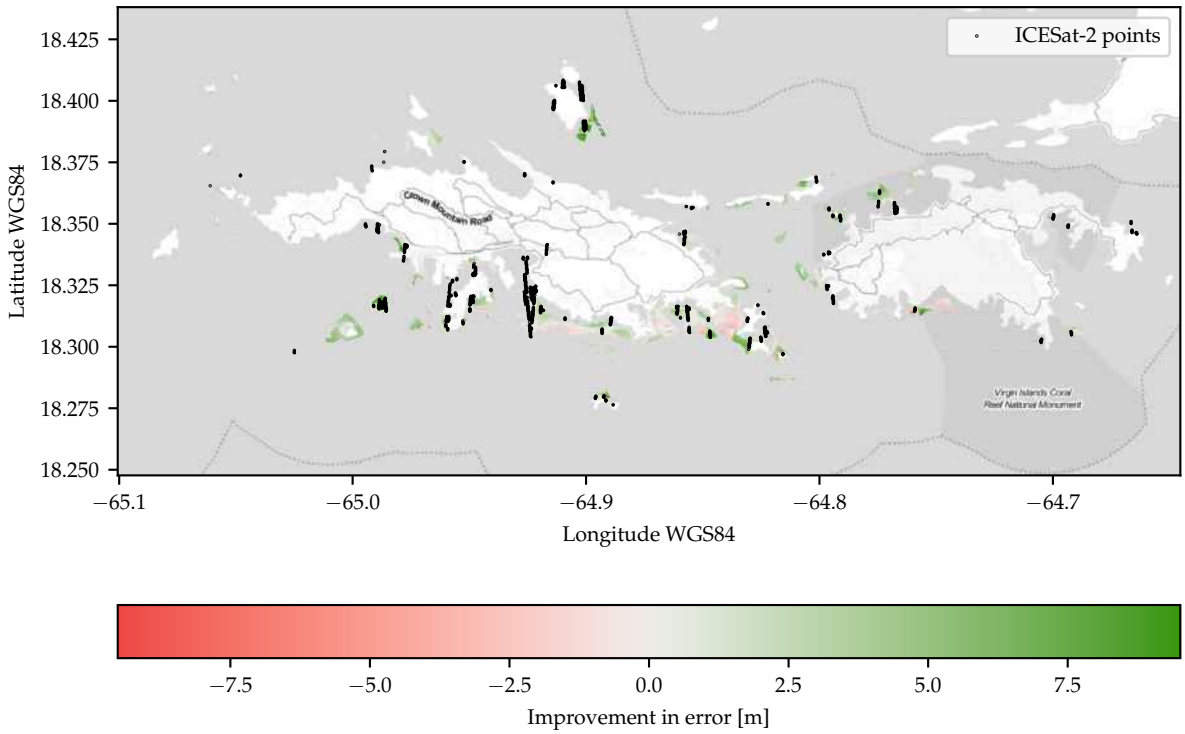


Figure C.5: Error Improvement in St. Thomas/St. John test site.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.2. Florida Keys test site

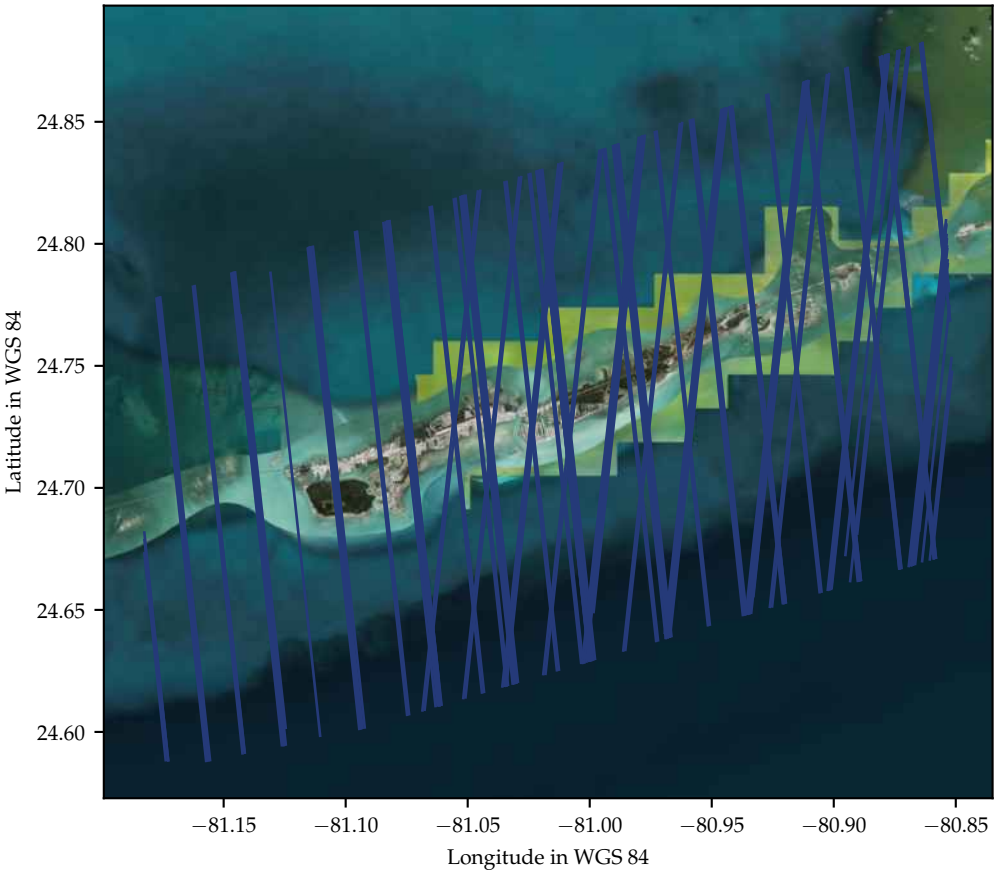


Figure C.6: Florida Keys test site: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

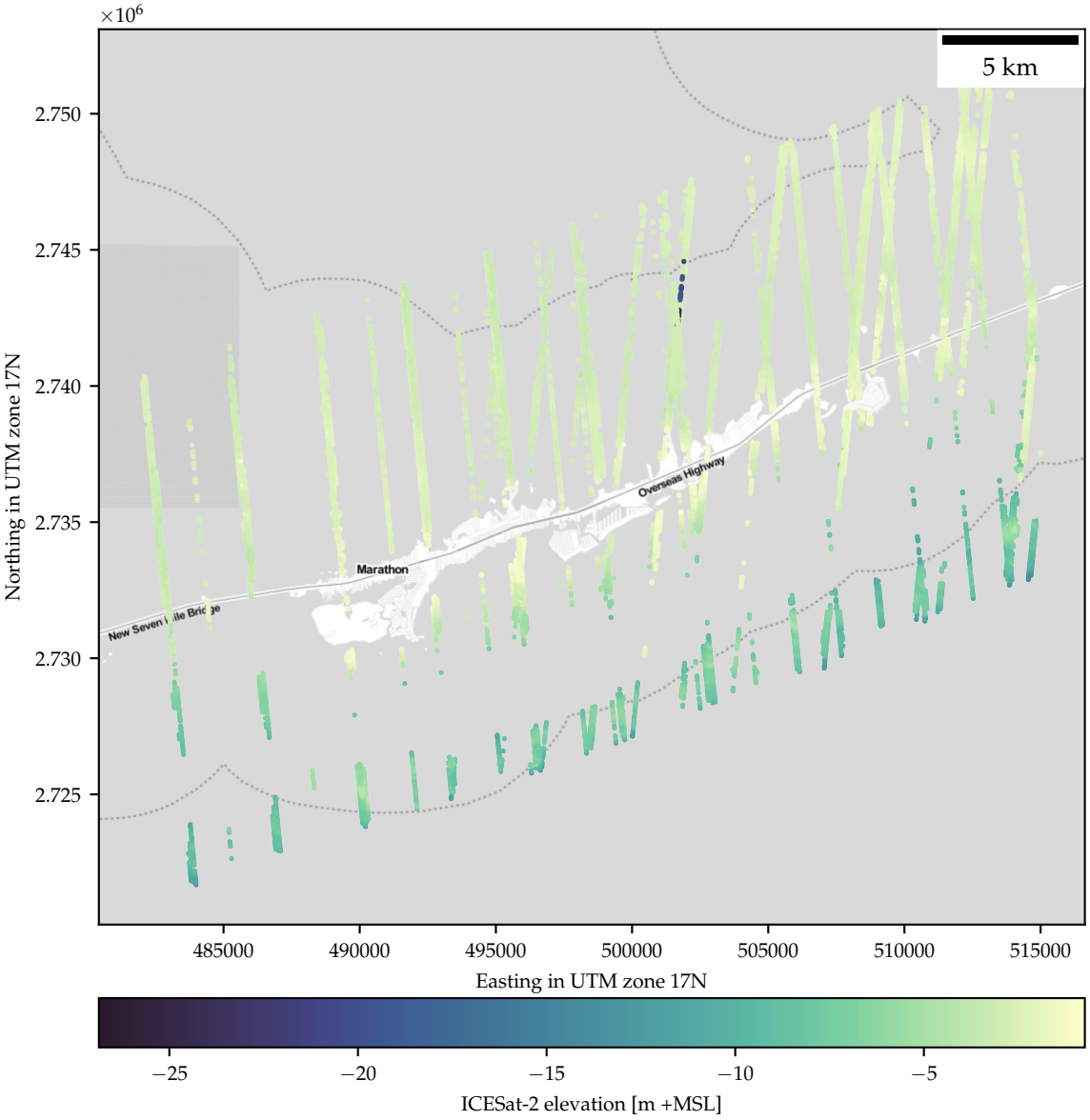


Figure C.7: Florida Keys test site: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

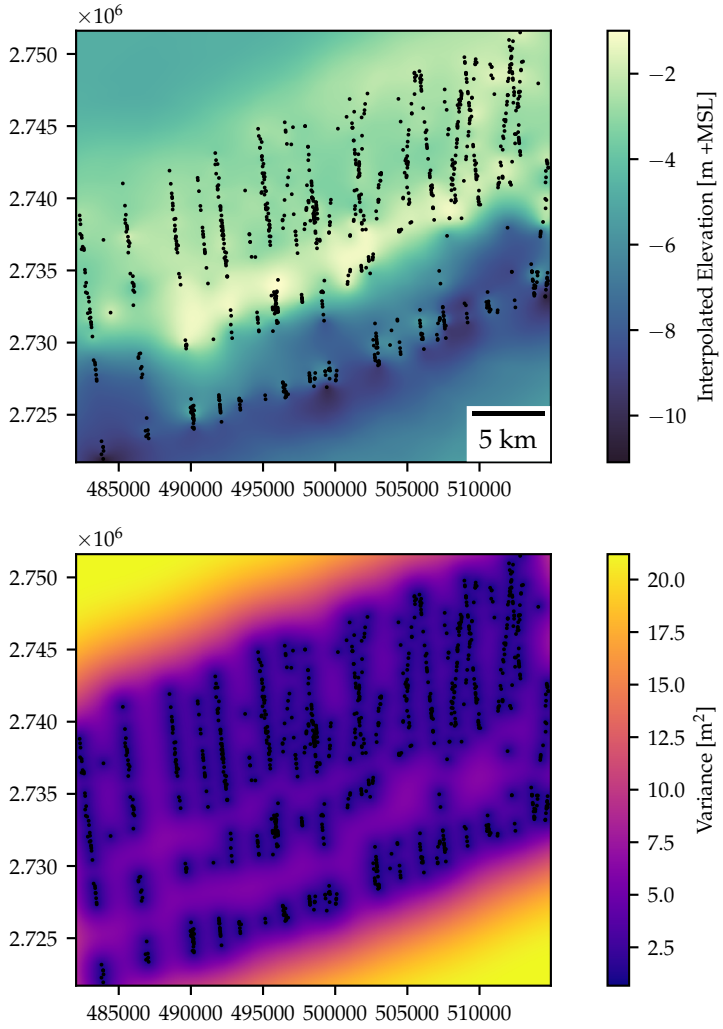


Figure C.8: Florida Keys test site: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

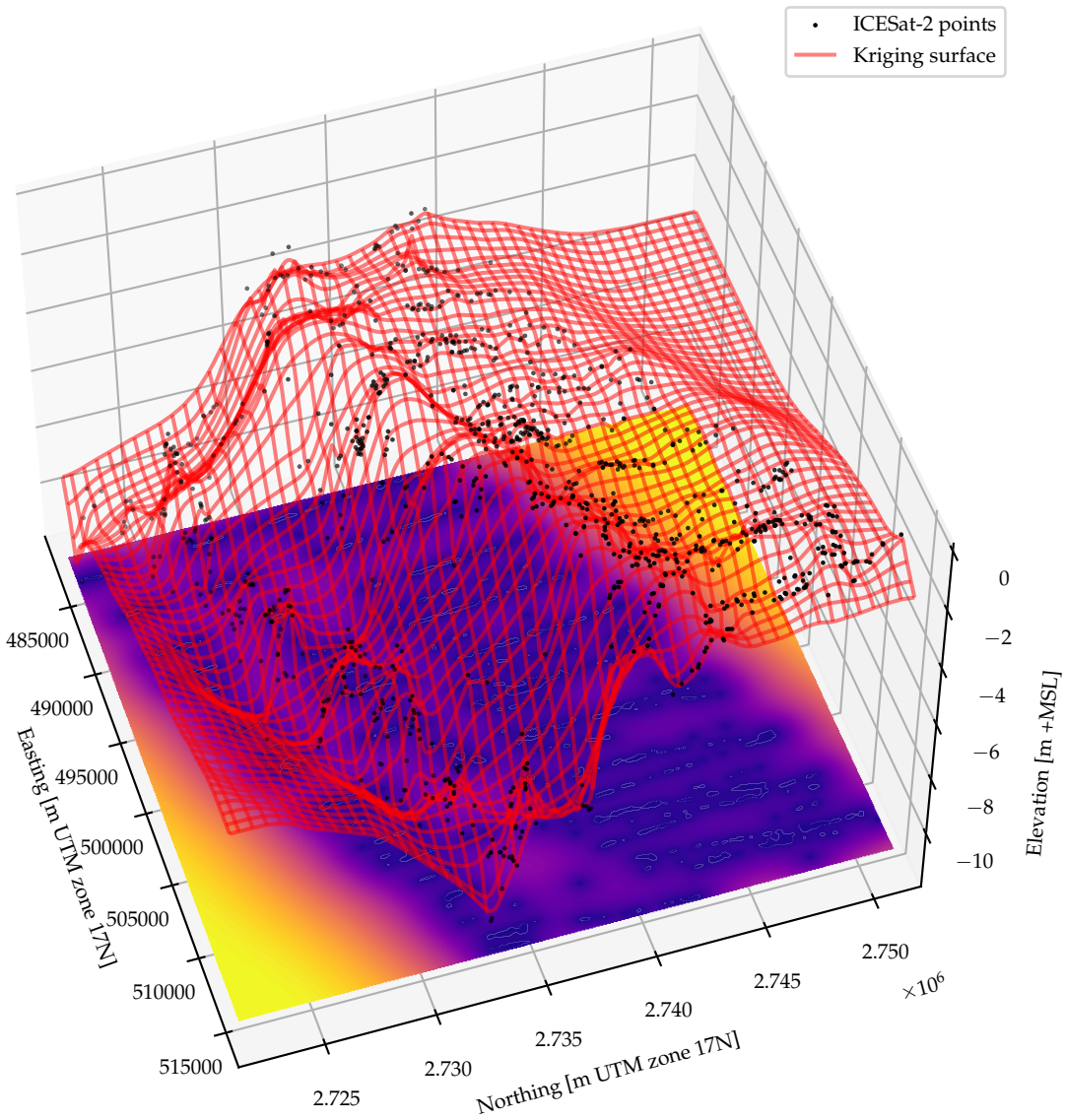


Figure C.9: Florida Keys test site: bathymetry points from KDE signal finding.
Basemap data:

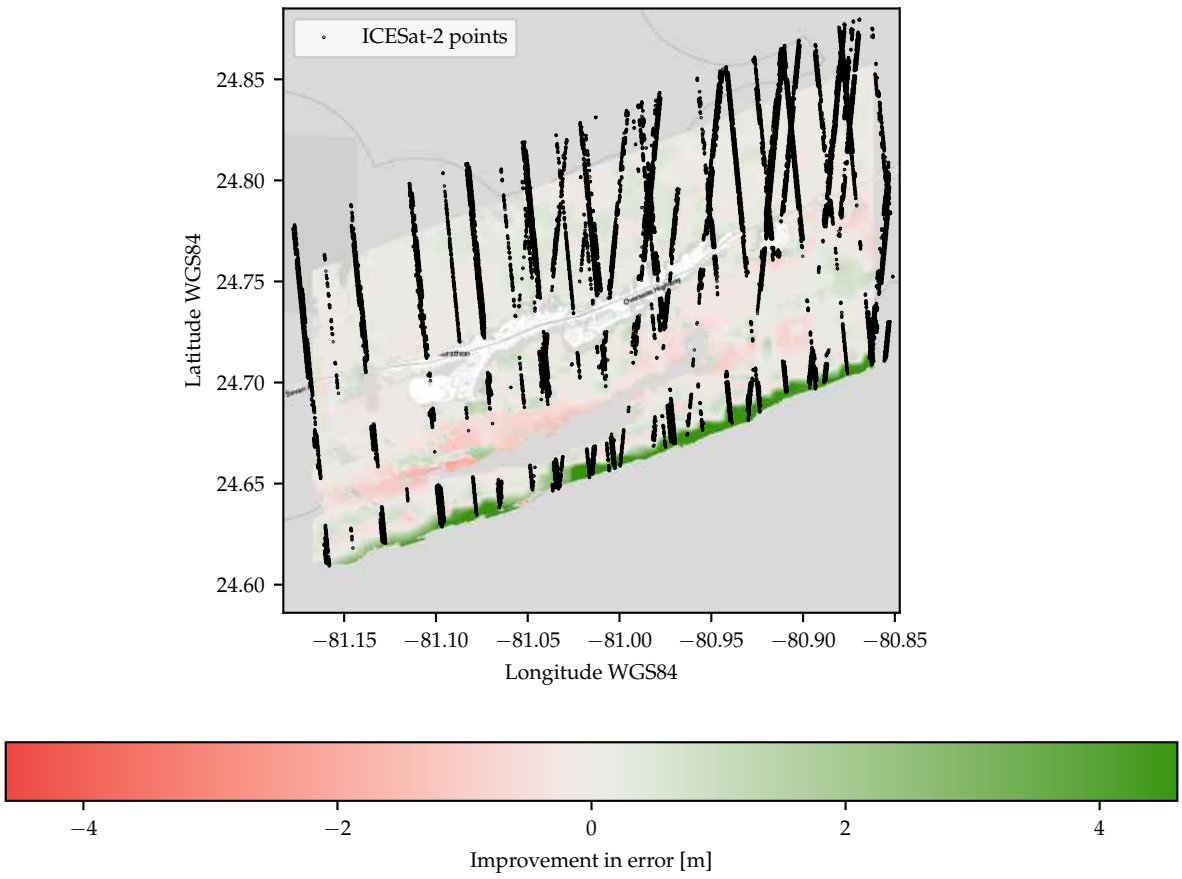


Figure C.10: Error Improvement in Florida Keys test site.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.3. St. Croix test site

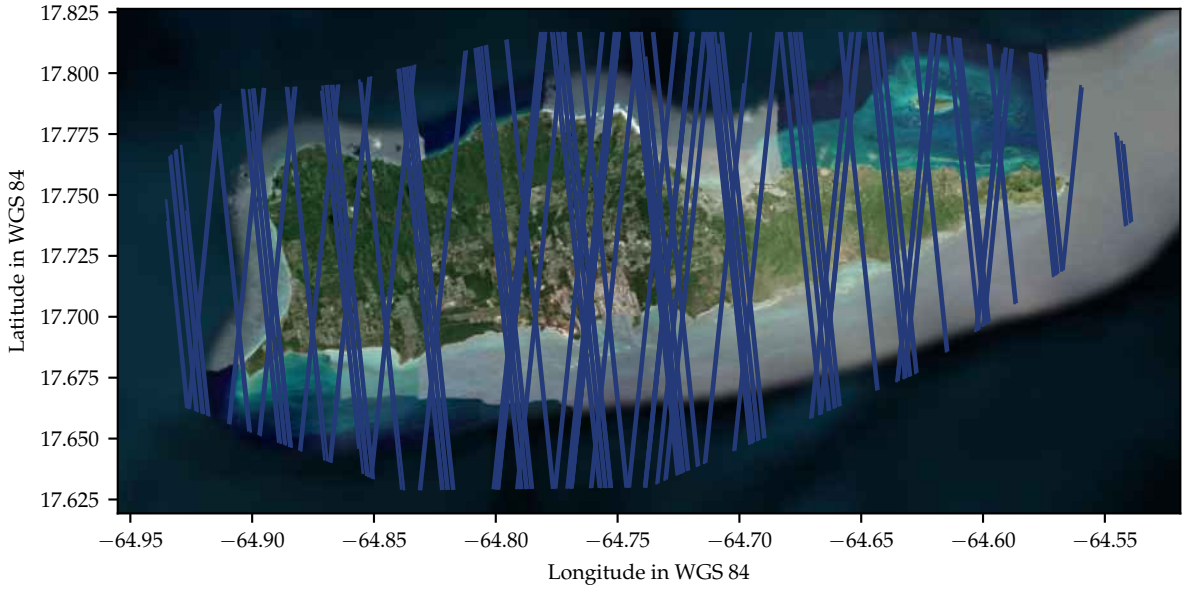


Figure C.11: St. Croix test site: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

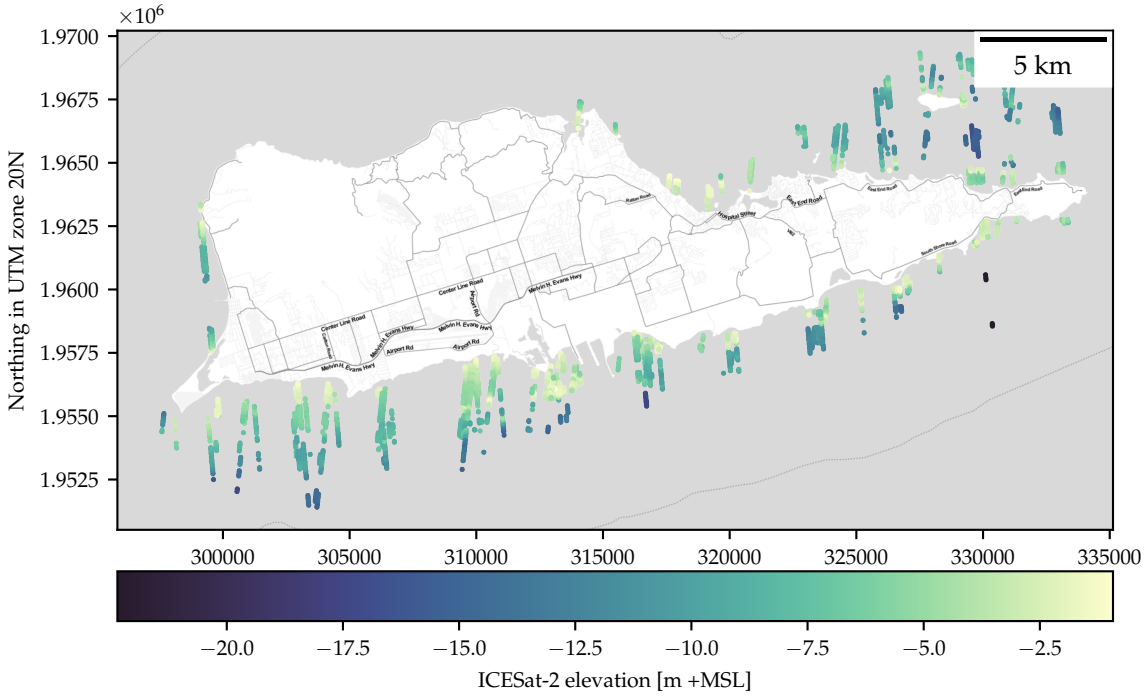


Figure C.12: St. Croix test site: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

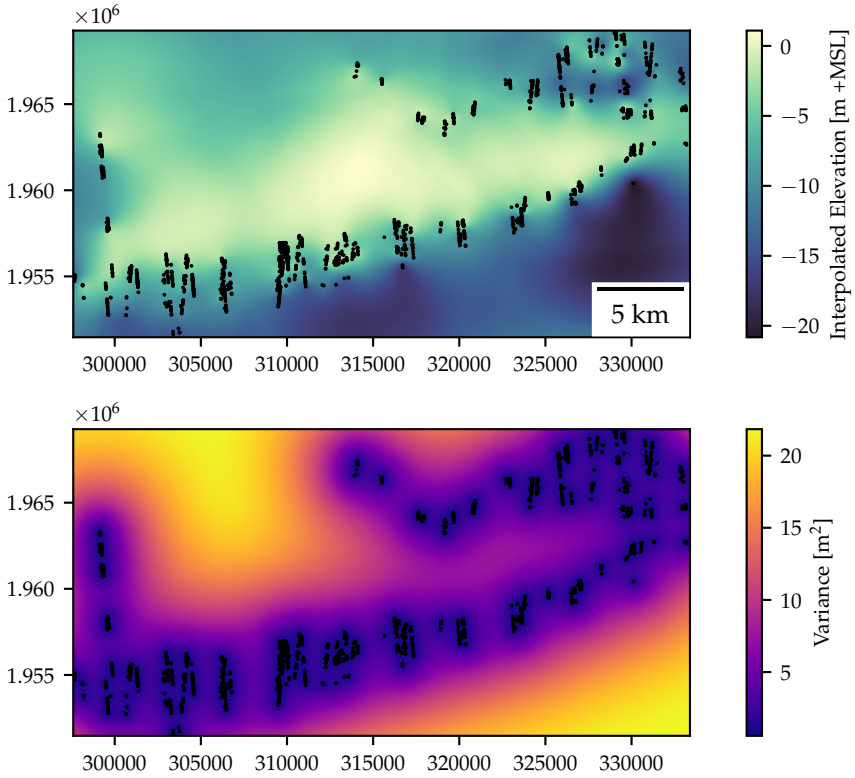


Figure C.13: St. Croix test site: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

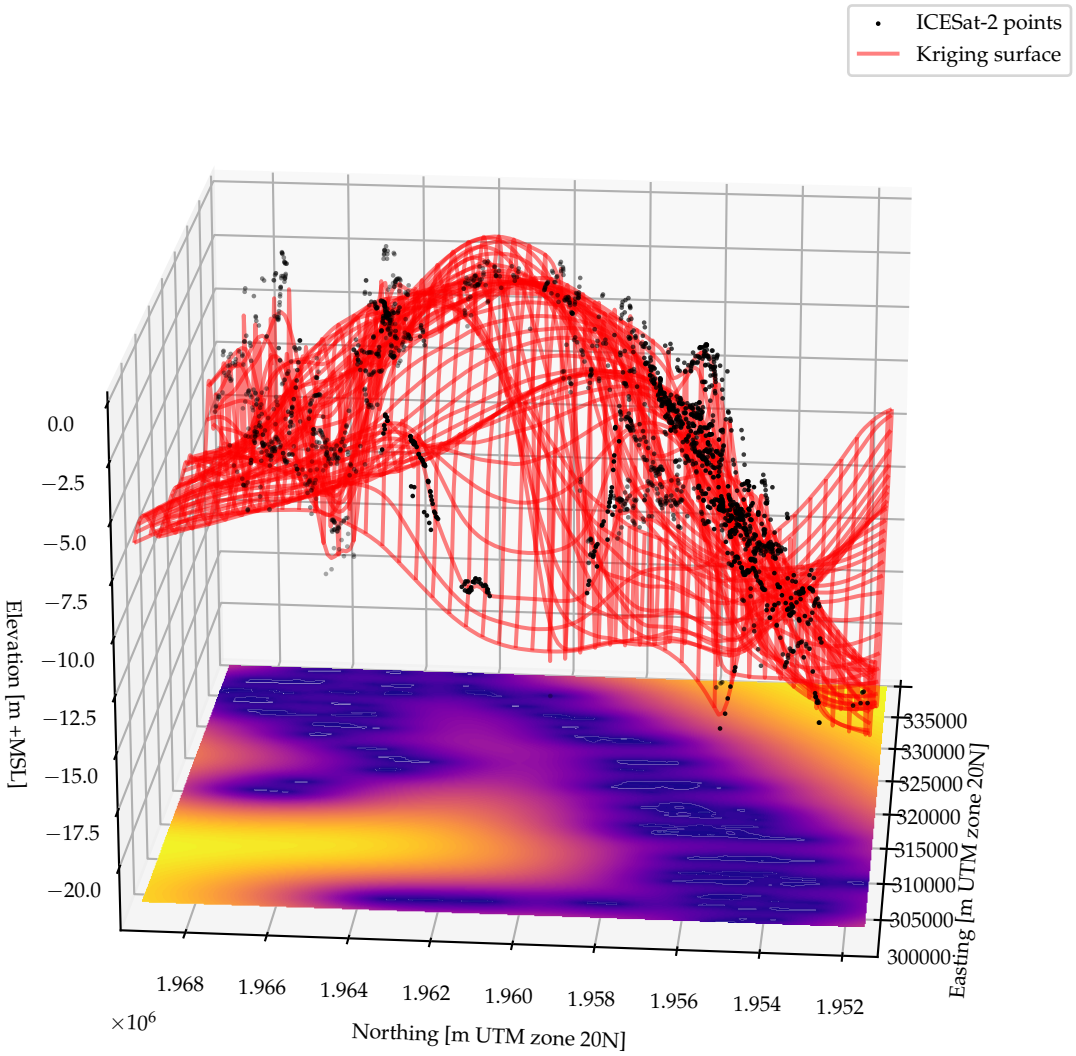


Figure C.14: St. Croix test site: bathymetry points from KDE signal finding.
Basemap data:

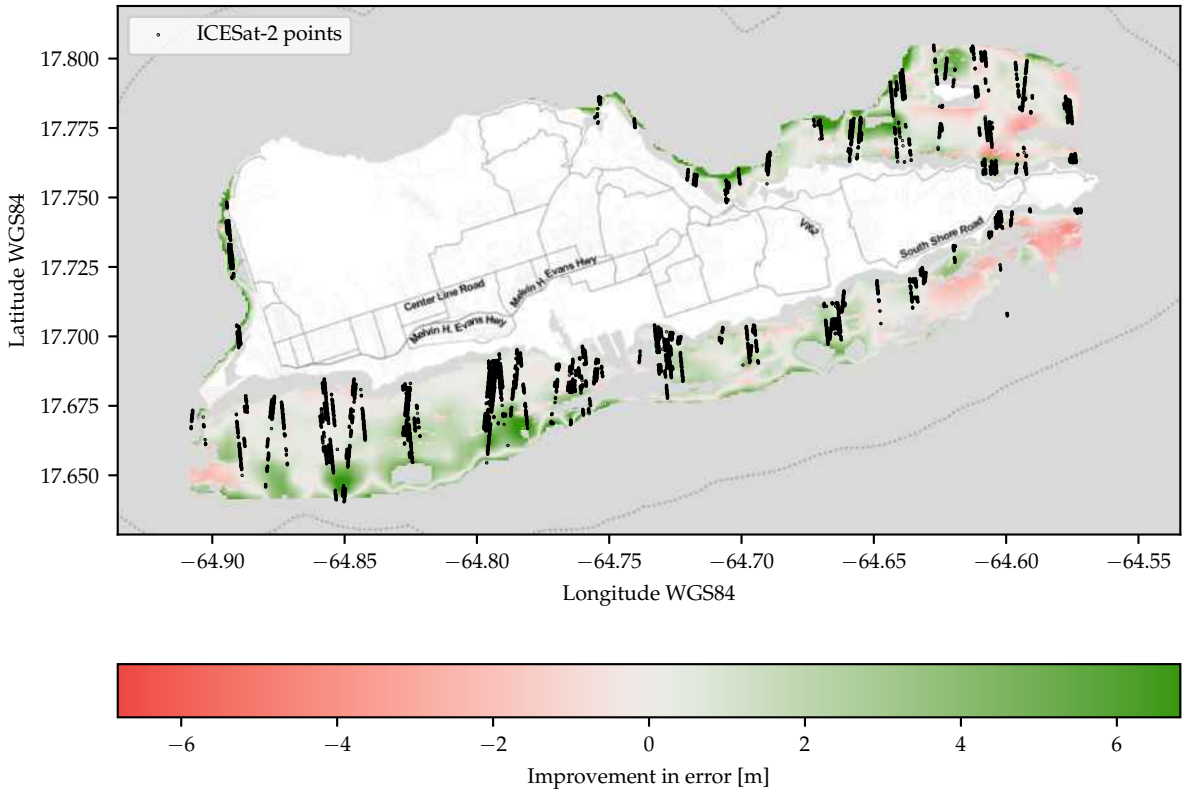


Figure C.15: Error Improvement in St. Croix test site.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.4. Oahu Subsite 1

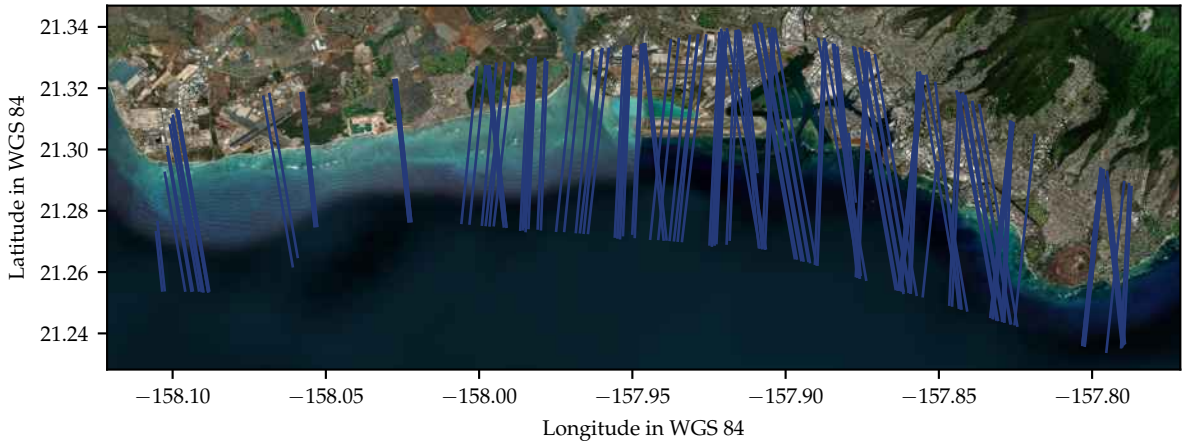


Figure C.16: Oahu Subsite 1: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

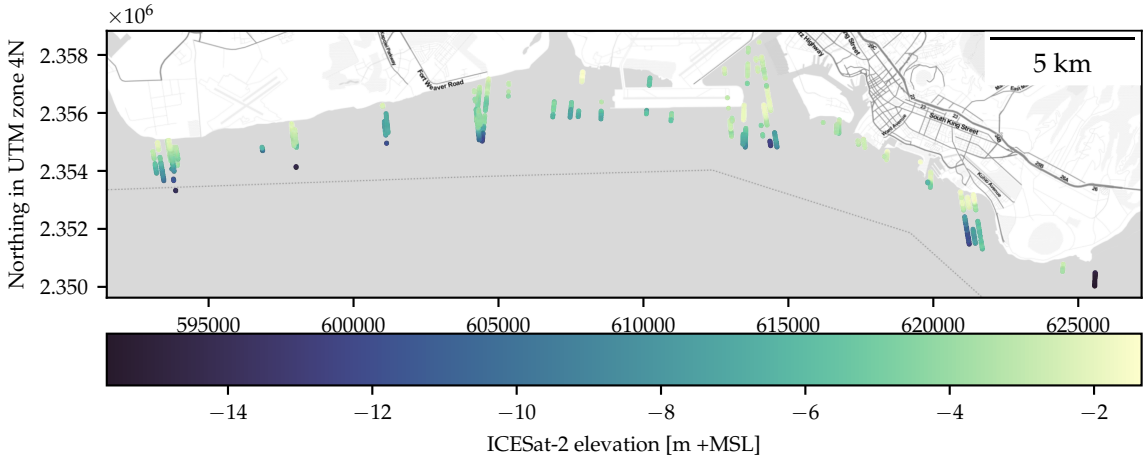


Figure C.17: Oahu Subsite 1: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

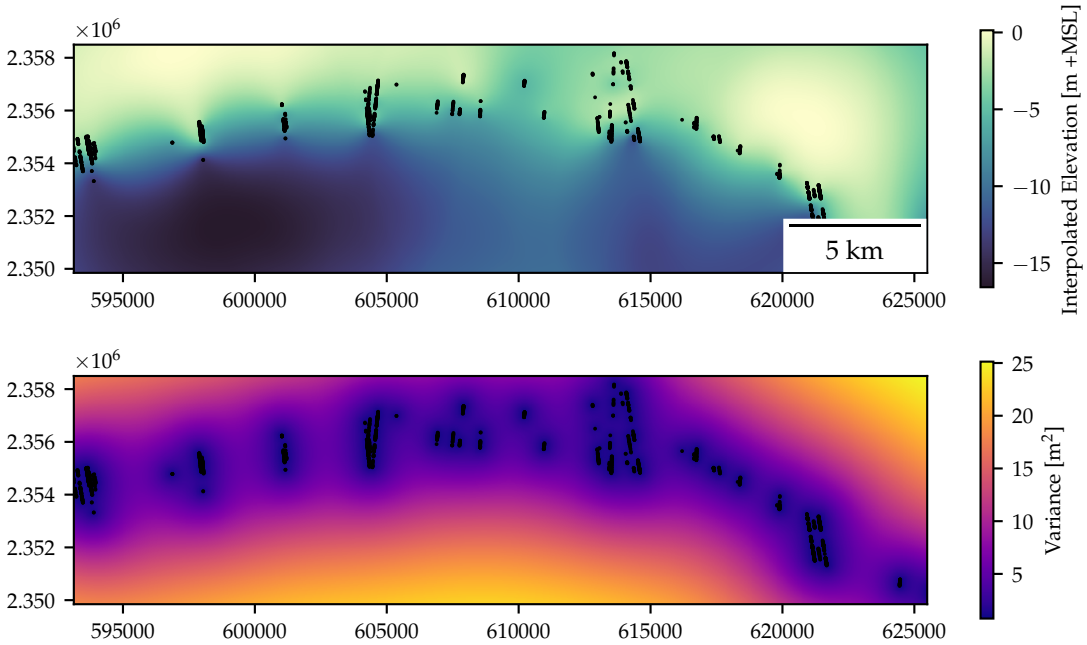


Figure C.18: Oahu Subsite 1: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

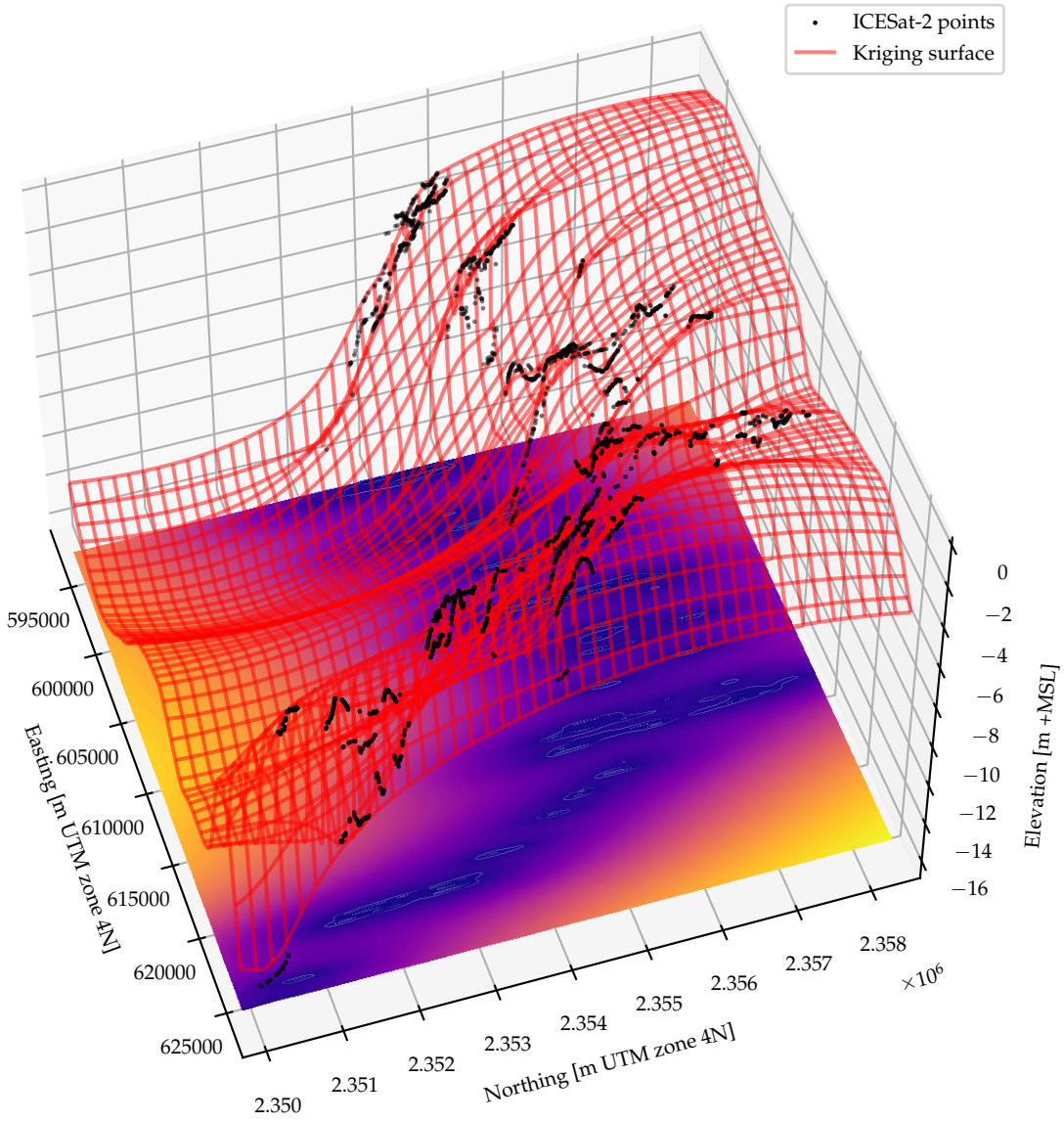


Figure C.19: Oahu Subsite 1: bathymetry points from KDE signal finding.
Basemap data:

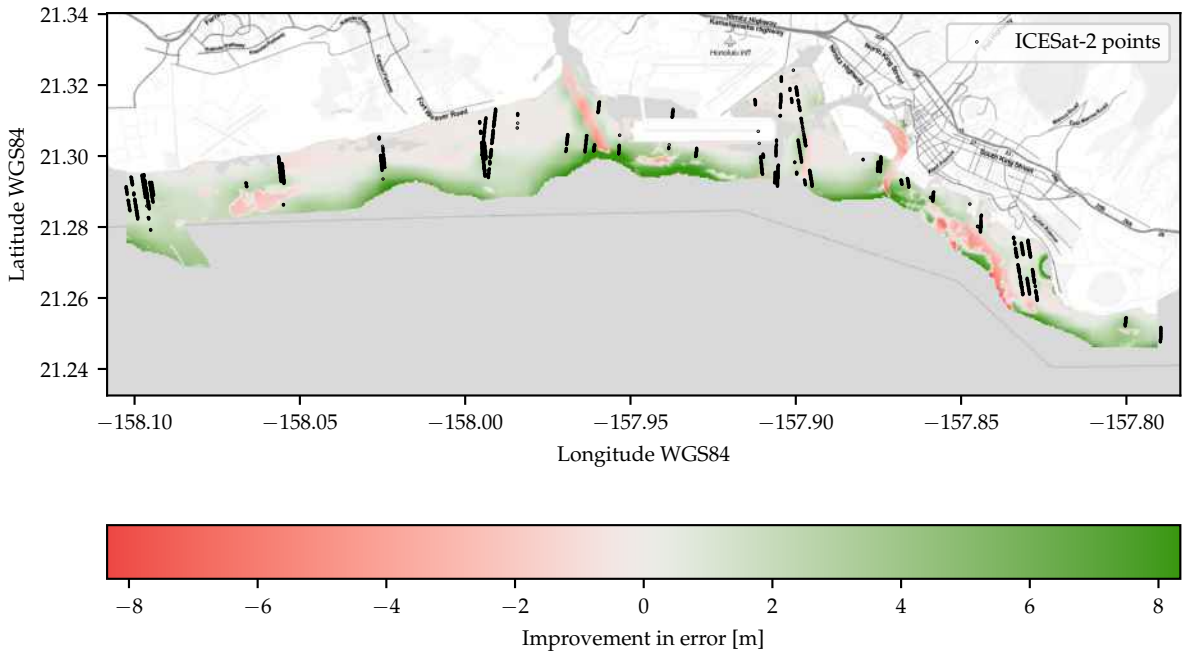


Figure C.20: Error Improvement in Oahu Subsite 1.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.5. Oahu Subsite 2

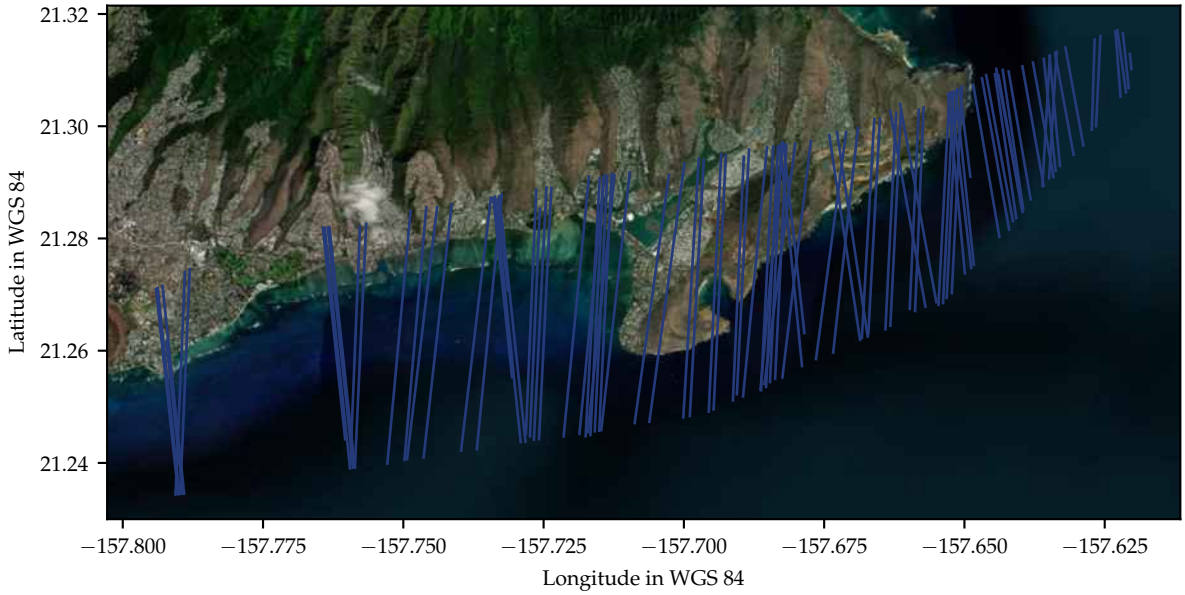


Figure C.21: Oahu Subsite 2: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

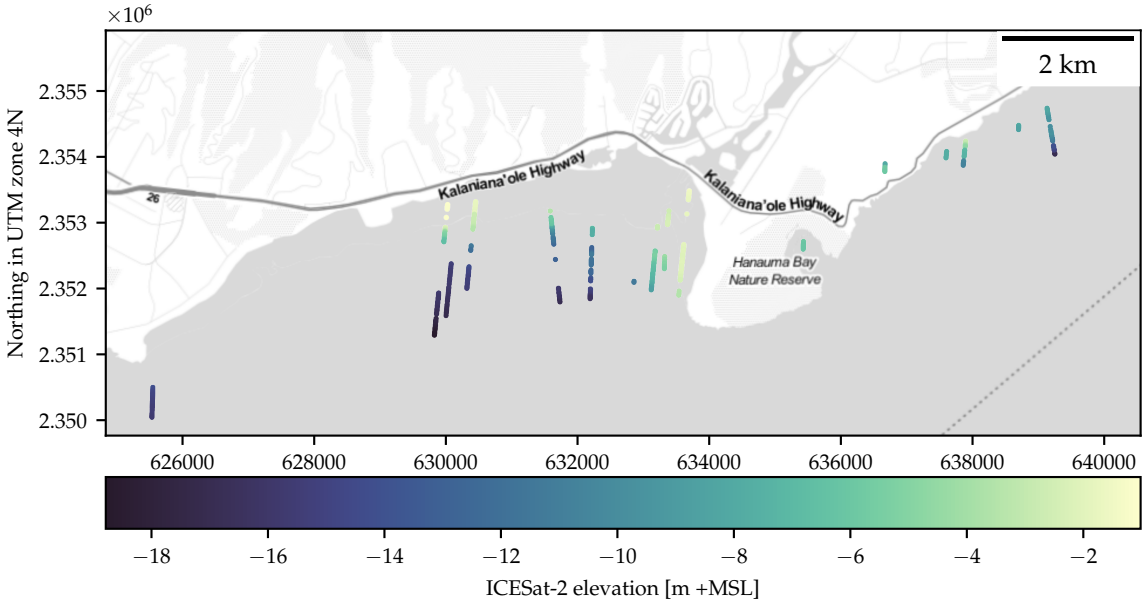


Figure C.22: Oahu Subsite 2: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

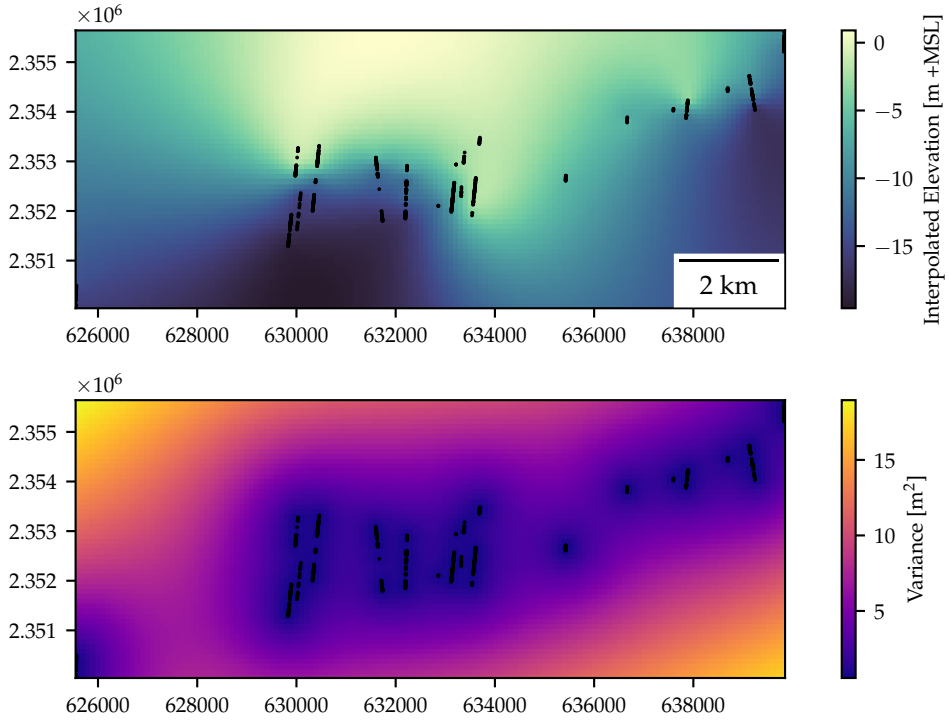


Figure C.23: Oahu Subsite 2: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

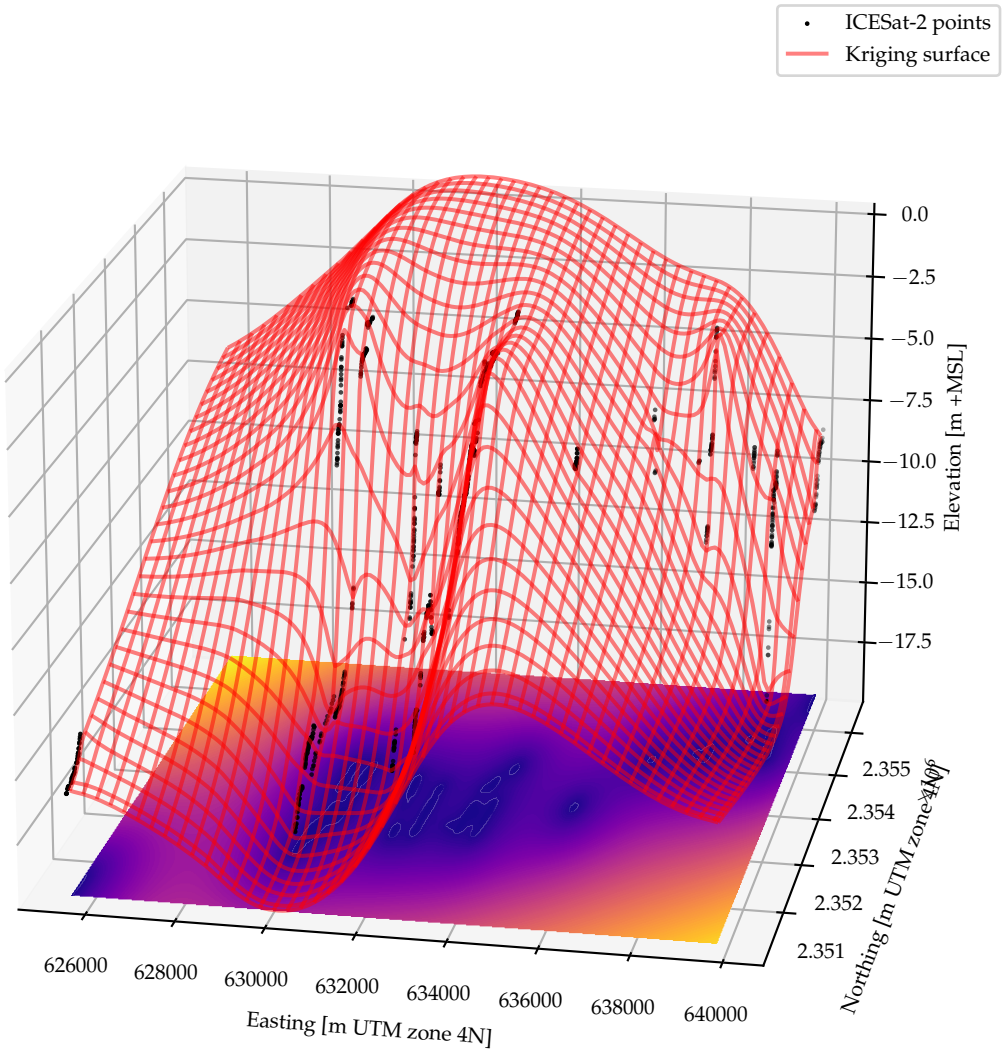


Figure C.24: Oahu Subsite 2: bathymetry points from KDE signal finding.
Basemap data:

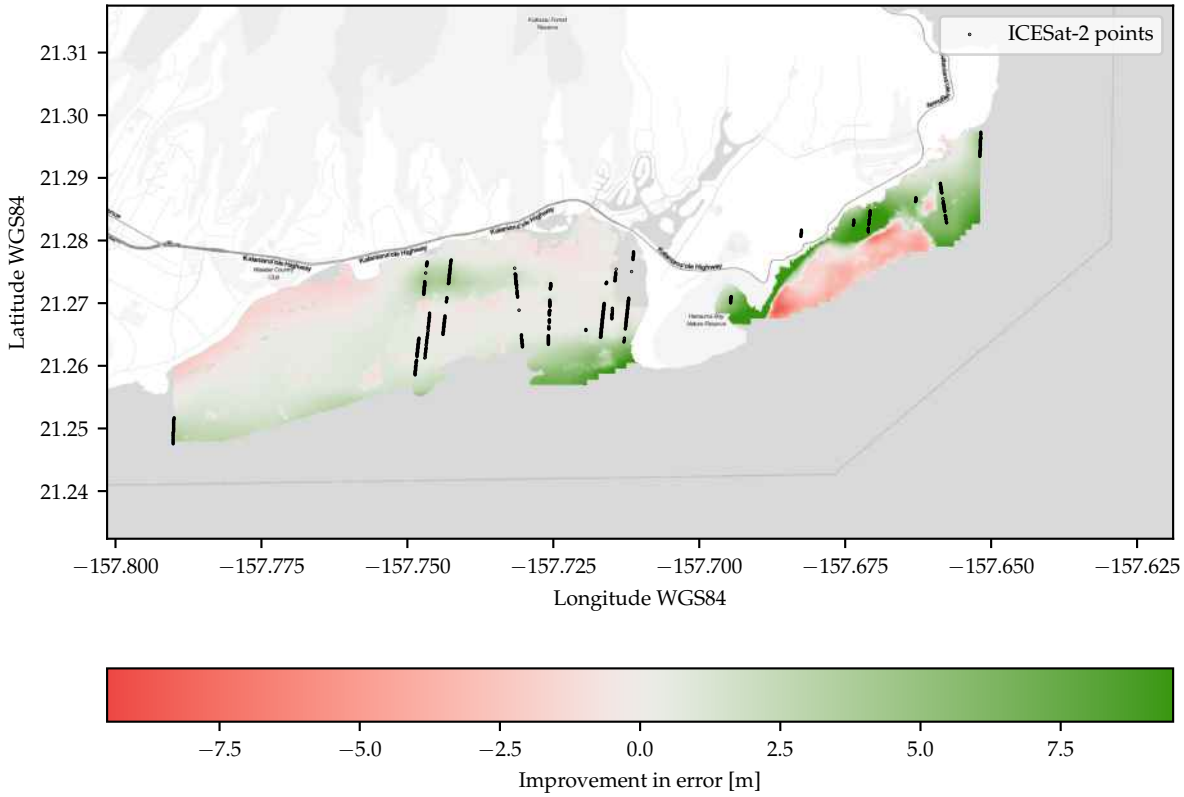


Figure C.25: Error Improvement in Oahu Subsite 2.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.6. Oahu Subsite 3

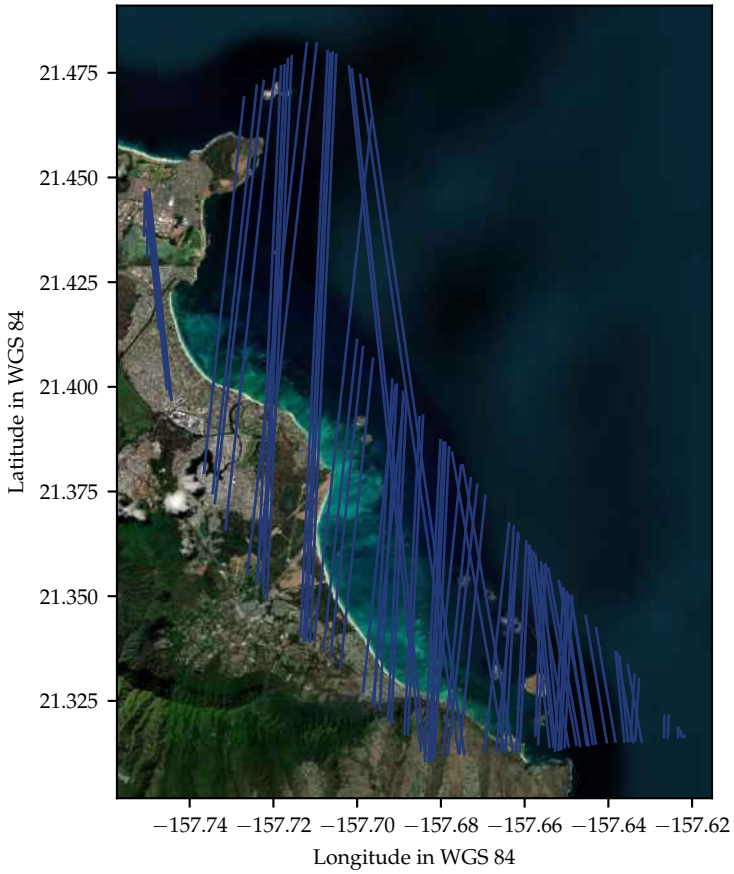


Figure C.26: Oahu Subsite 3: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

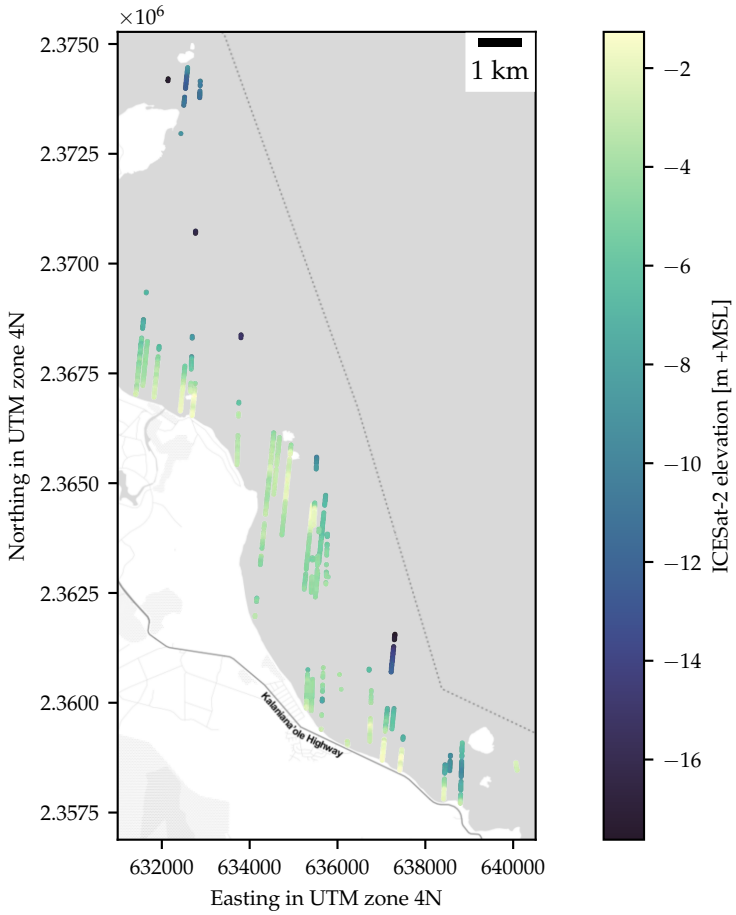


Figure C.27: Oahu Subsite 3: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

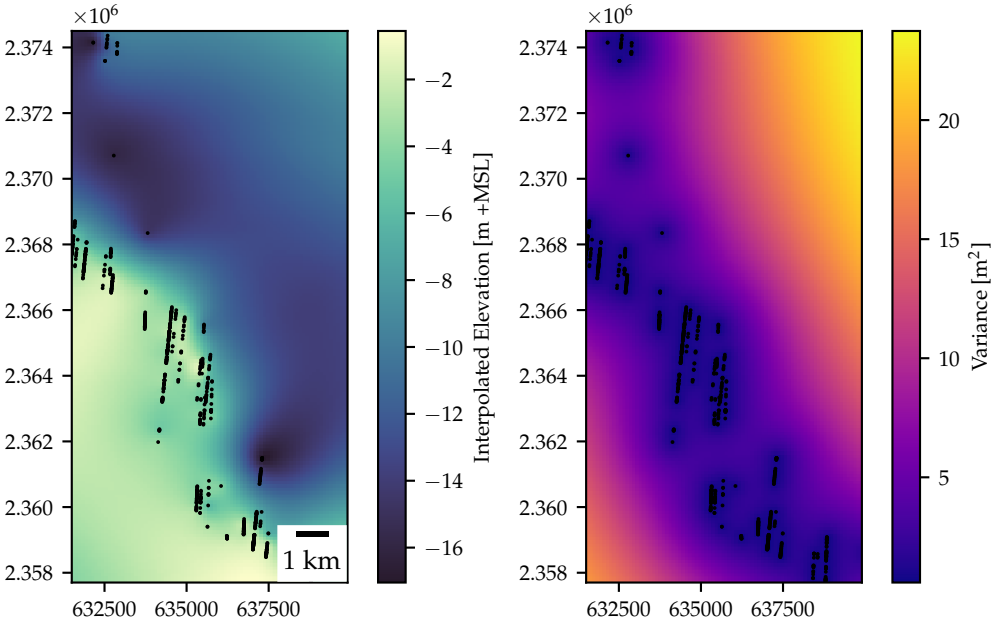


Figure C.28: Oahu Subsite 3: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

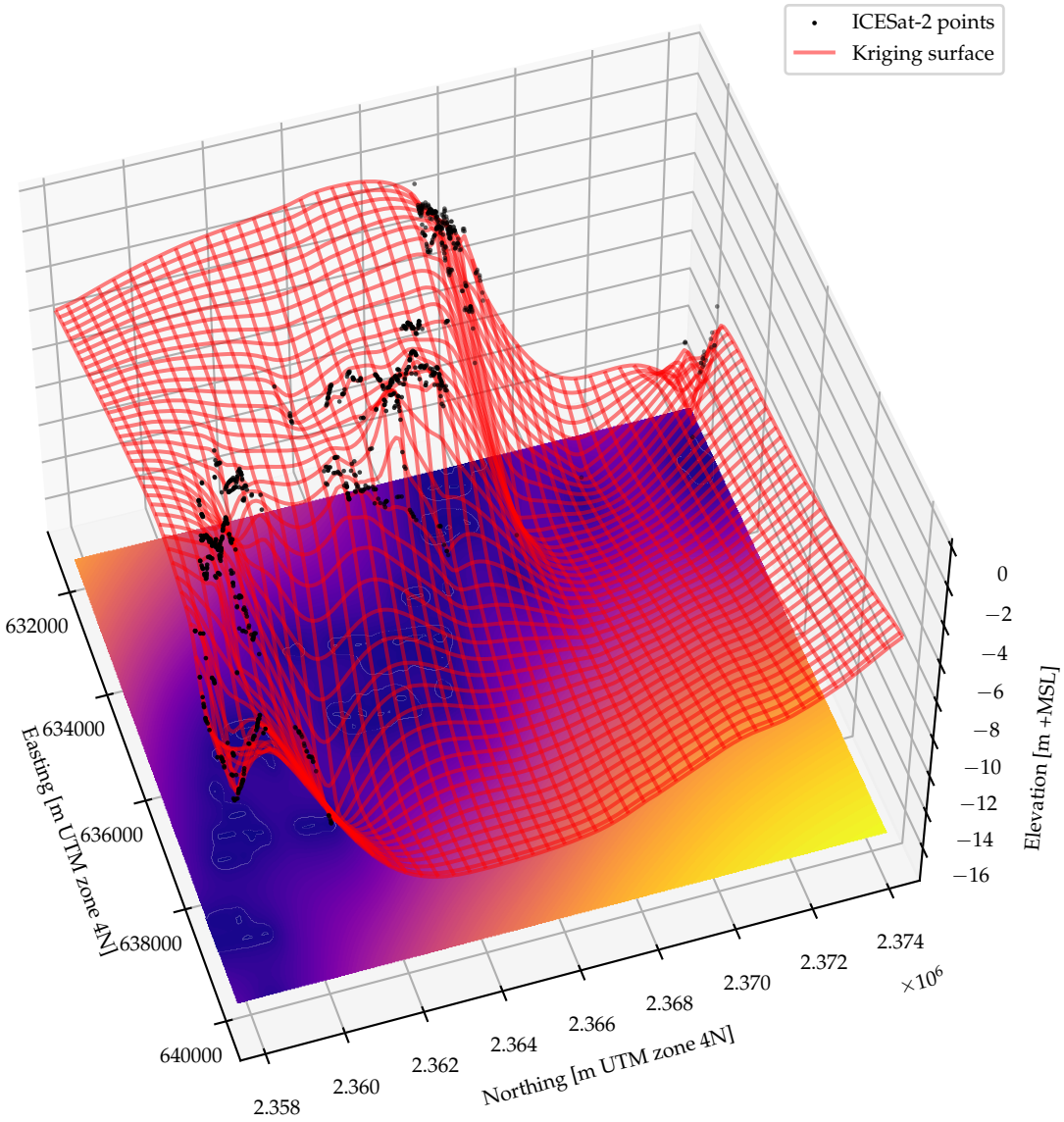


Figure C.29: Oahu Subsite 3: bathymetry points from KDE signal finding.
Basemap data:

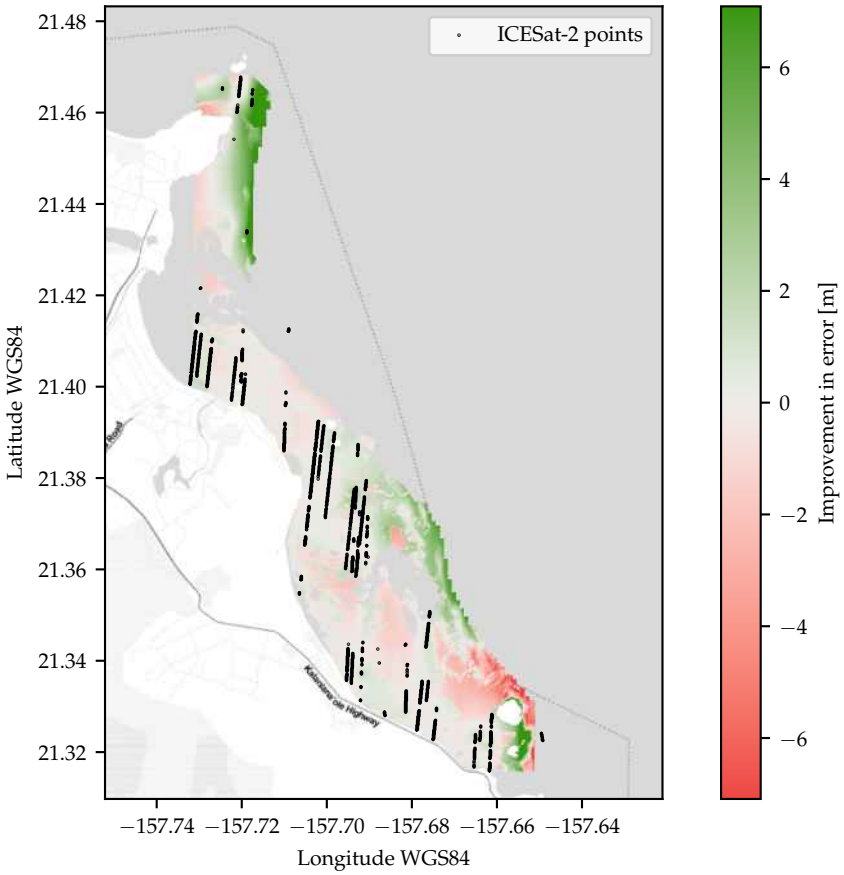


Figure C.30: Error Improvement in Oahu Subsite 3.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.7. Oahu Subsite 4

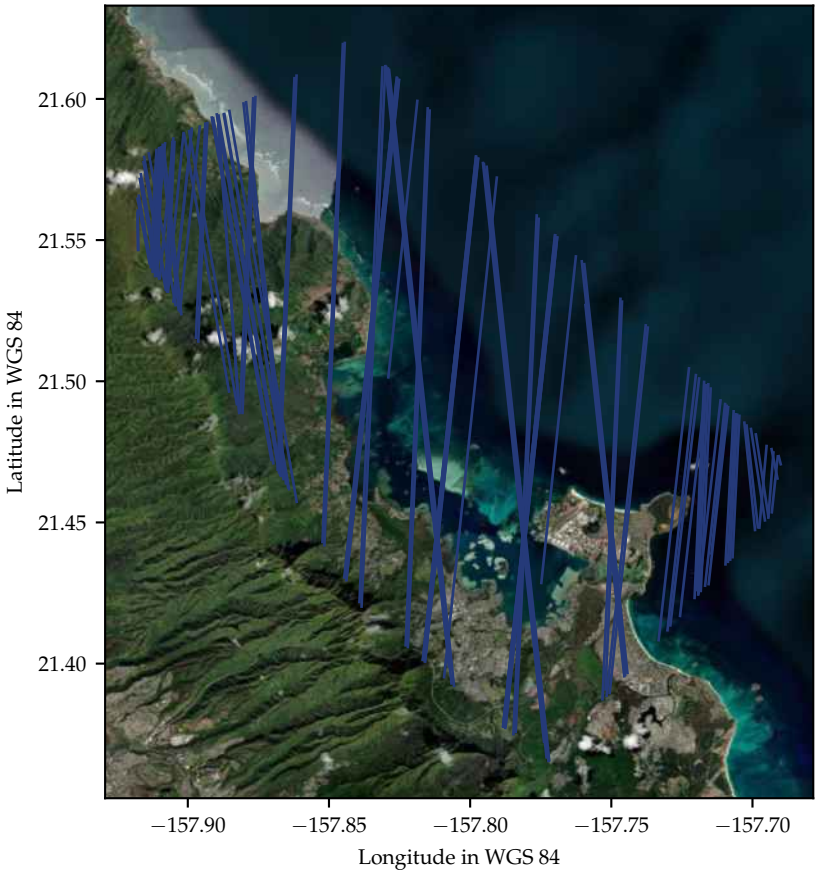


Figure C.31: Oahu Subsite 4: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

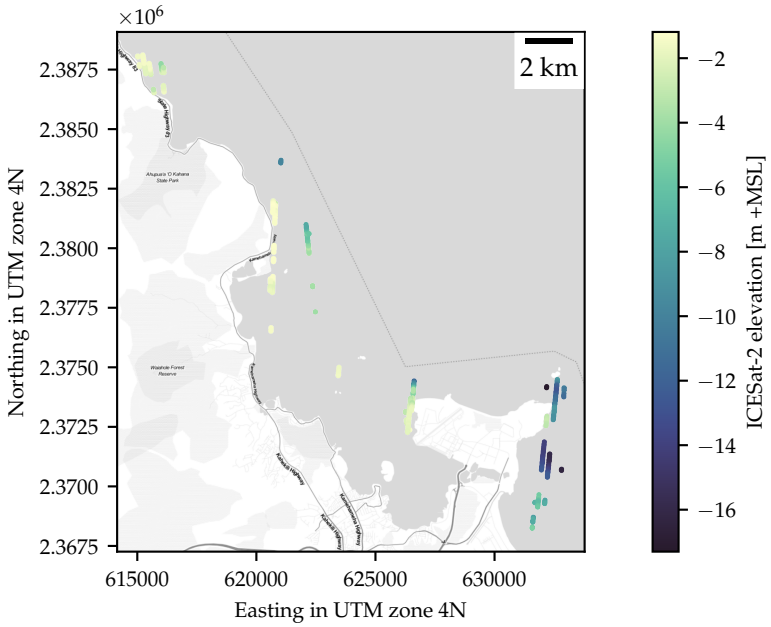


Figure C.32: Oahu Subsite 4: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

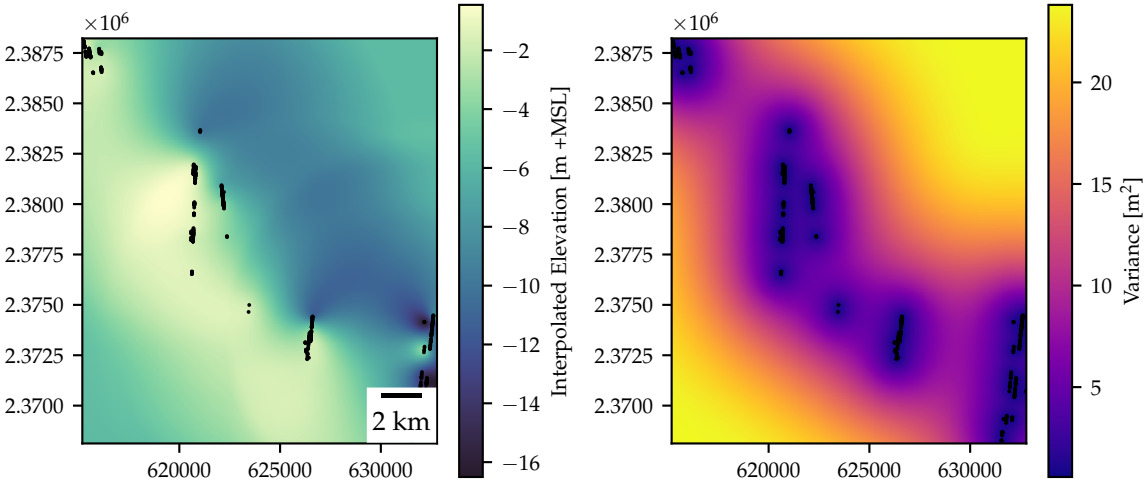


Figure C.33: Oahu Subsite 4: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

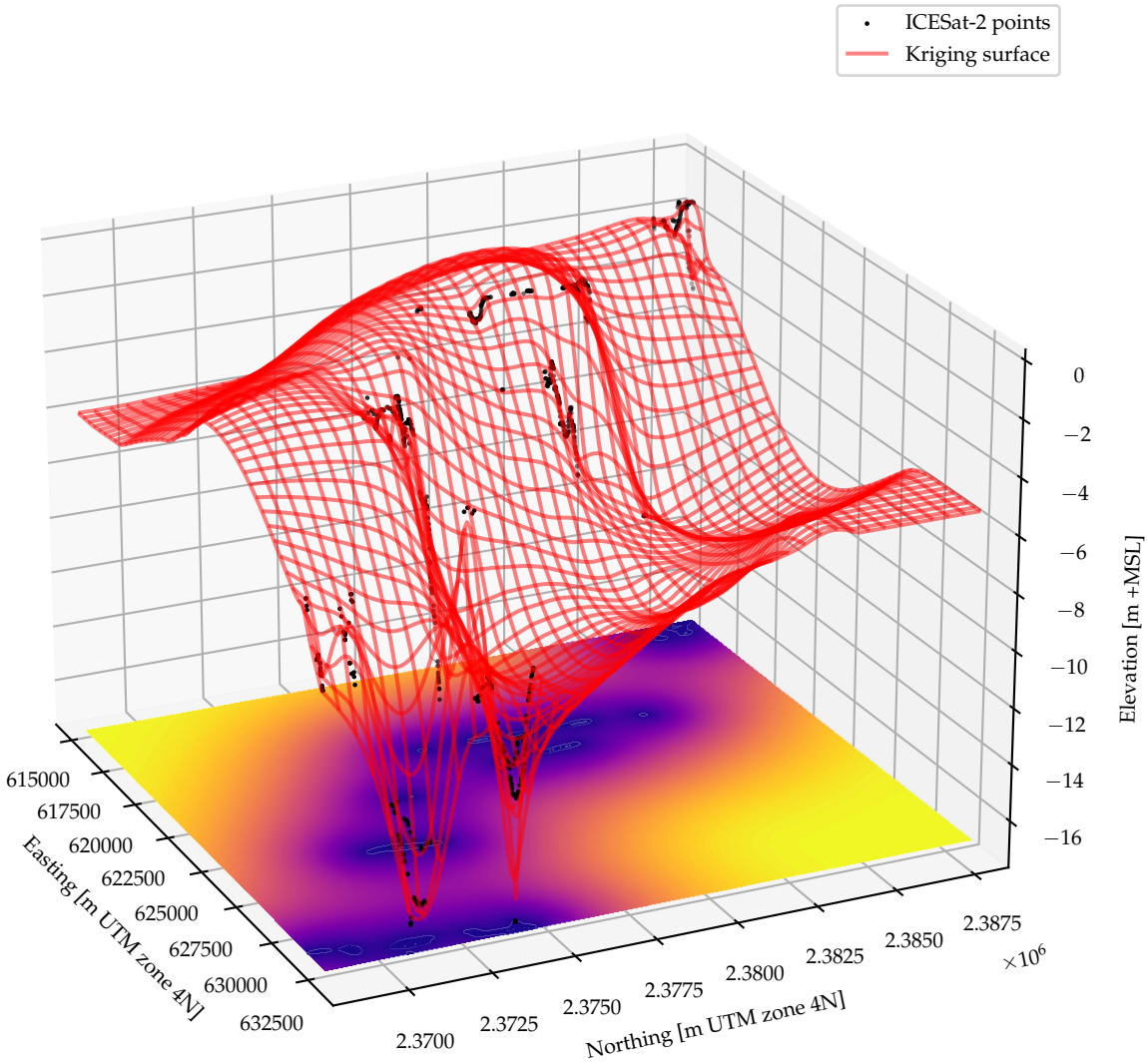


Figure C.34: Oahu Subsite 4: bathymetry points from KDE signal finding.
Basemap data:

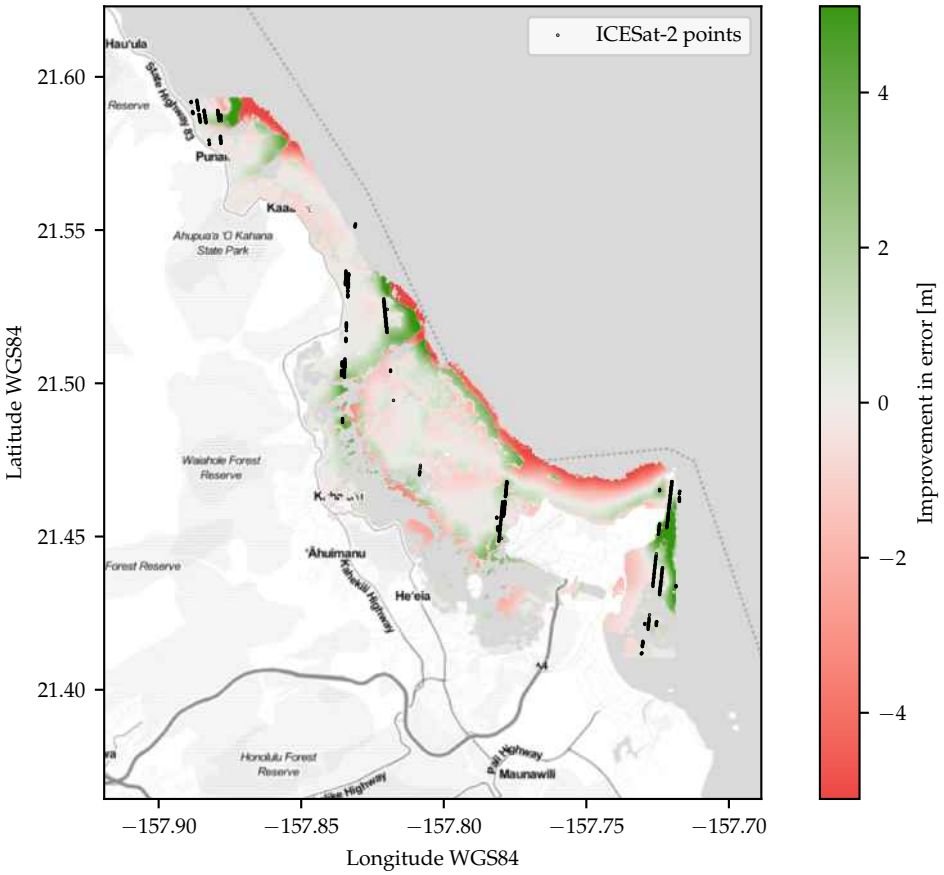


Figure C.35: Error Improvement in Oahu Subsite 4.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.8. Oahu Subsite 5

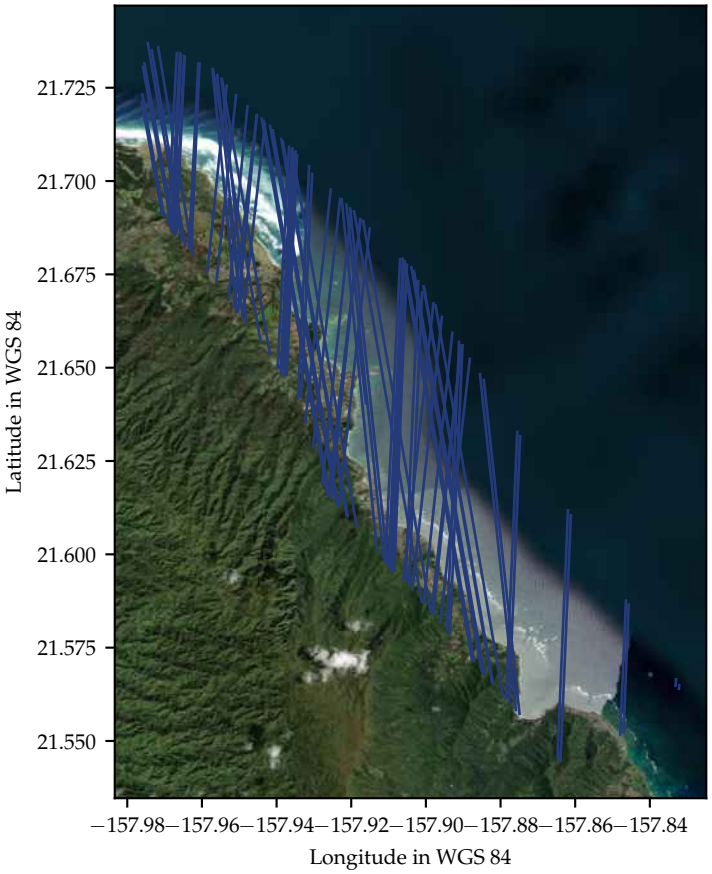


Figure C.36: Oahu Subsite 5: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

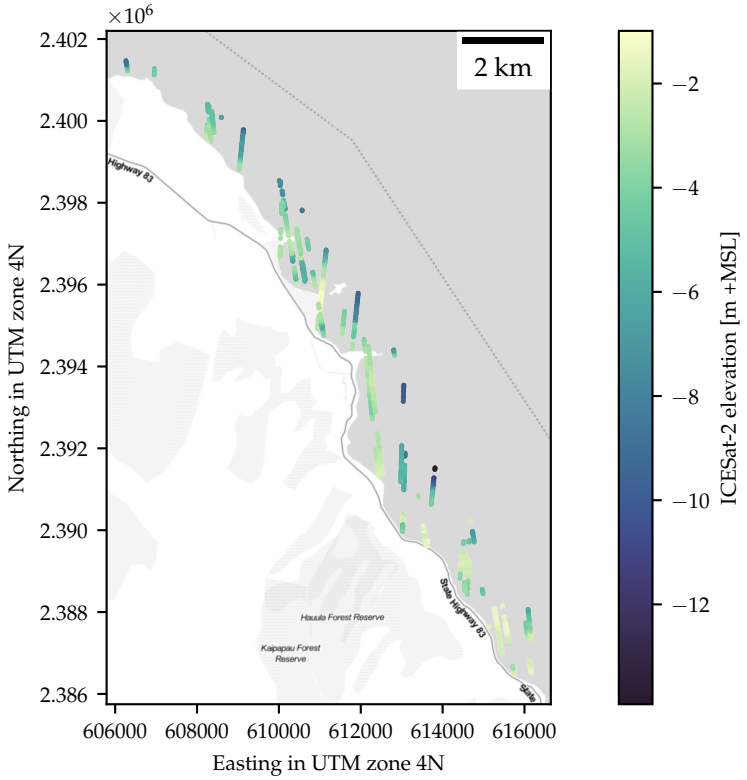


Figure C.37: Oahu Subsite 5: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

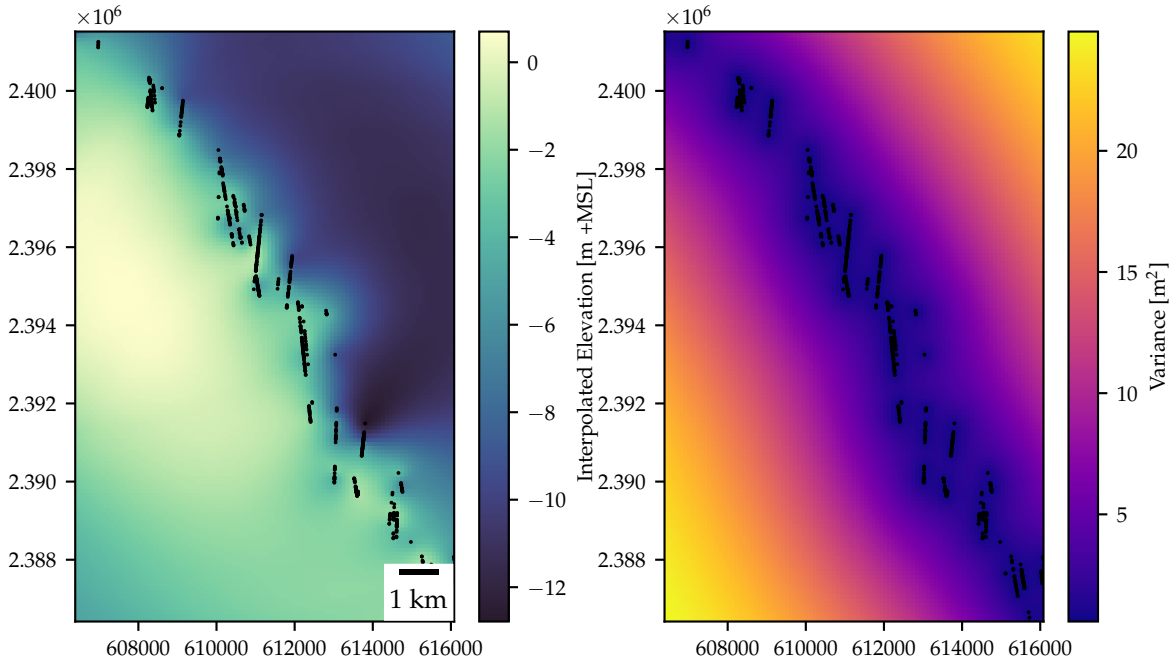


Figure C.38: Oahu Subsite 5: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

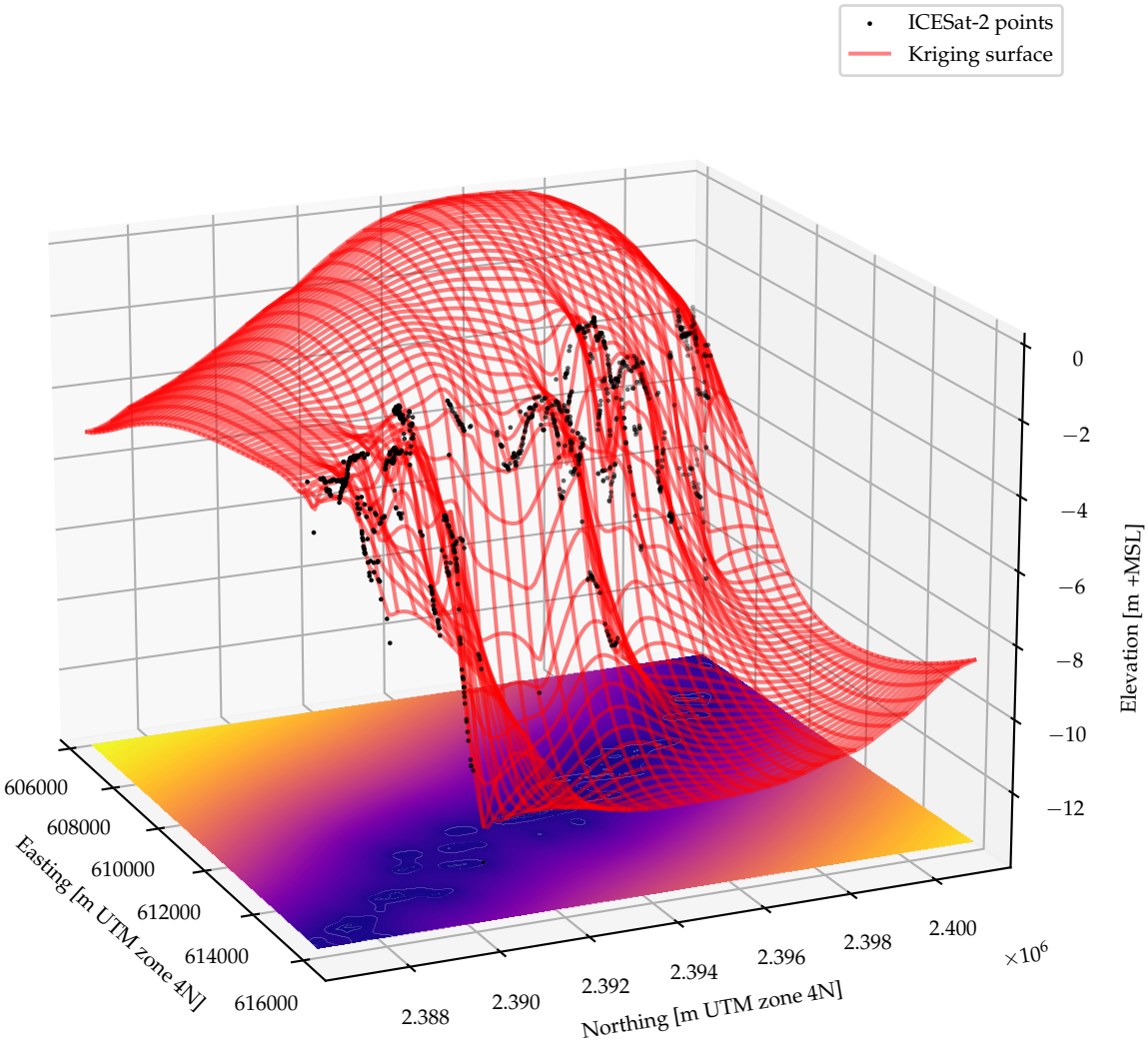


Figure C.39: Oahu Subsite 5: bathymetry points from KDE signal finding.
Basemap data:

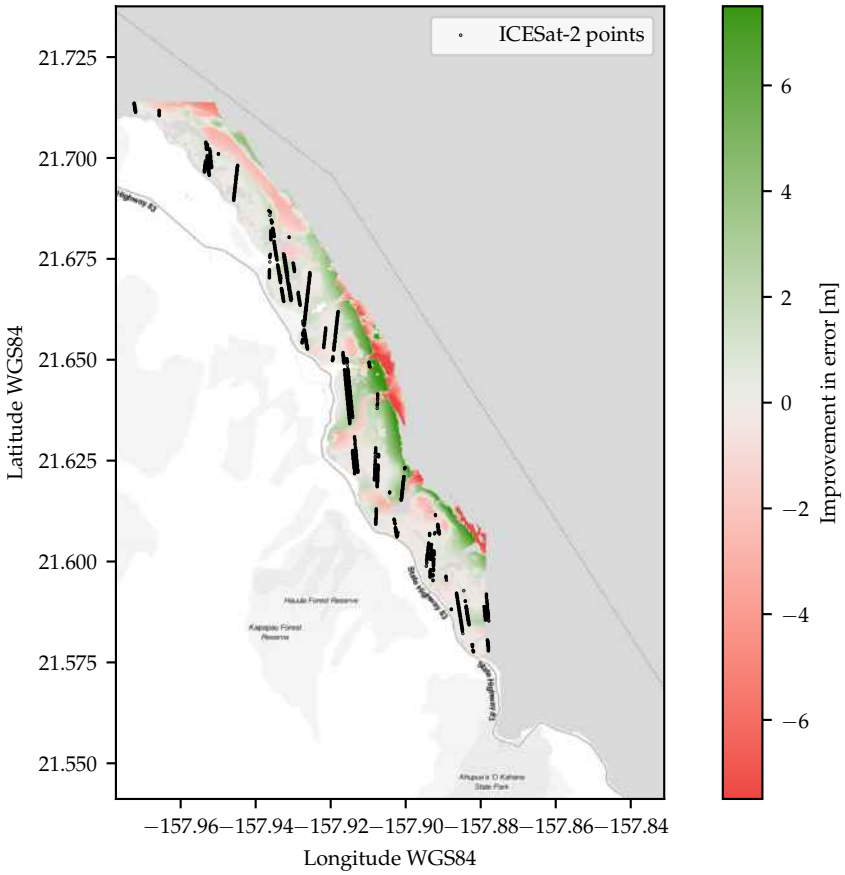


Figure C.40: Error Improvement in Oahu Subsite 5.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.9. Oahu Subsite 6

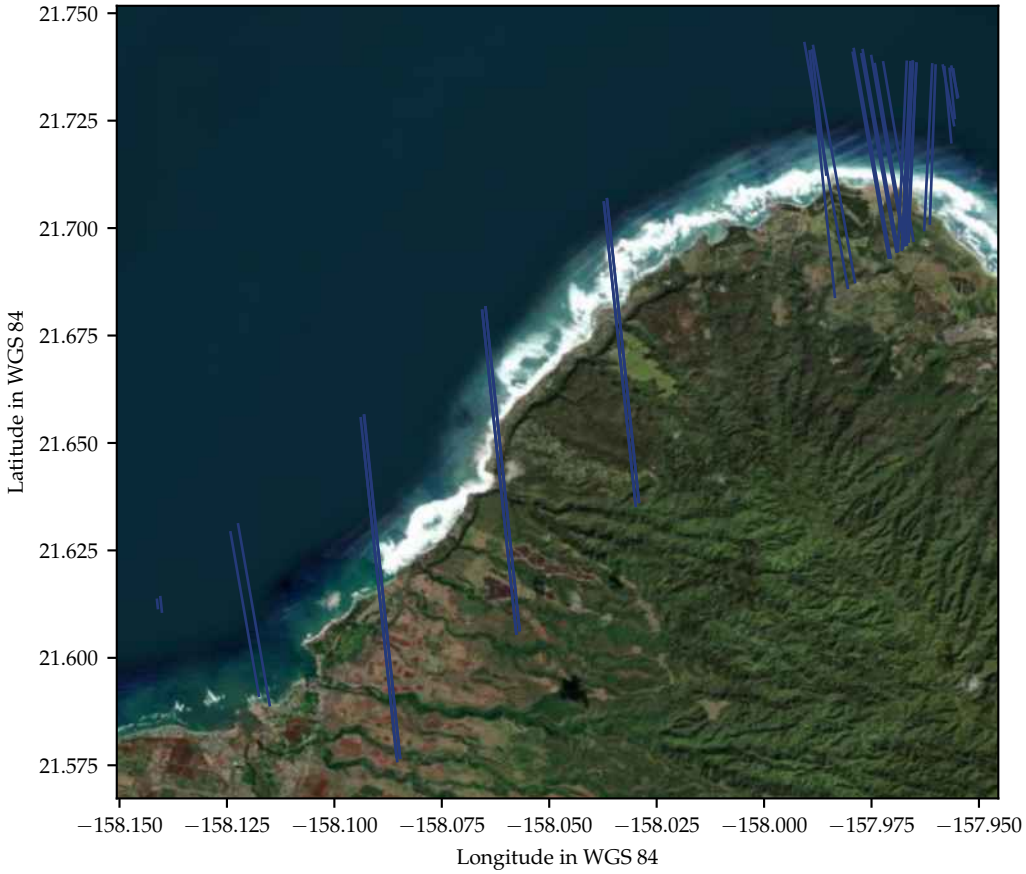


Figure C.41: Oahu Subsite 6: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

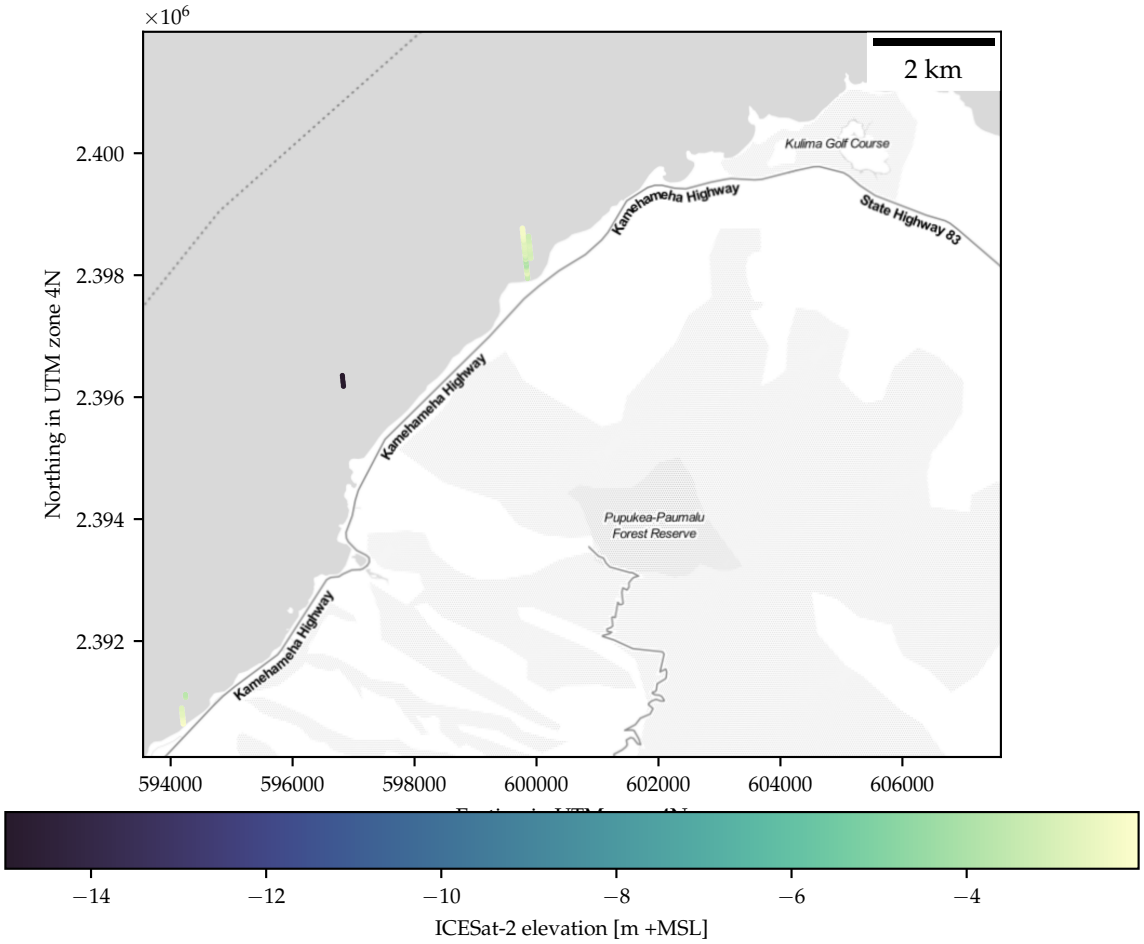


Figure C.42: Oahu Subsite 6: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

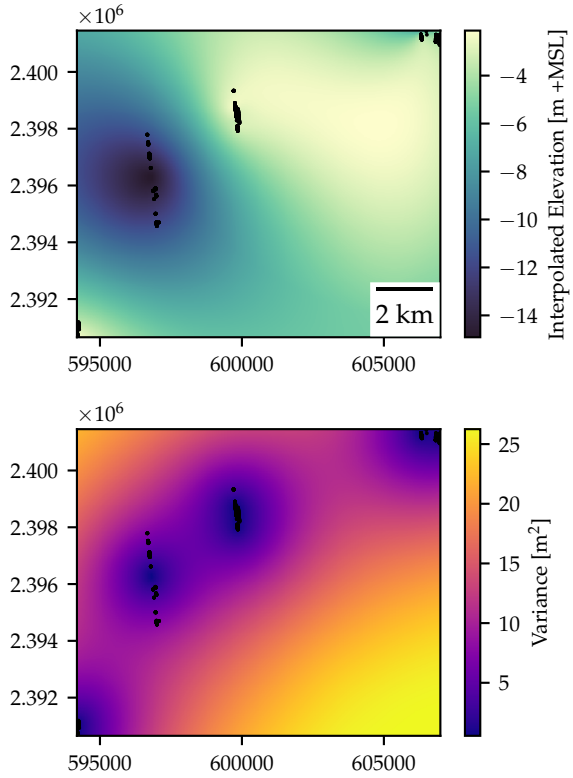


Figure C.43: Oahu Subsite 6: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

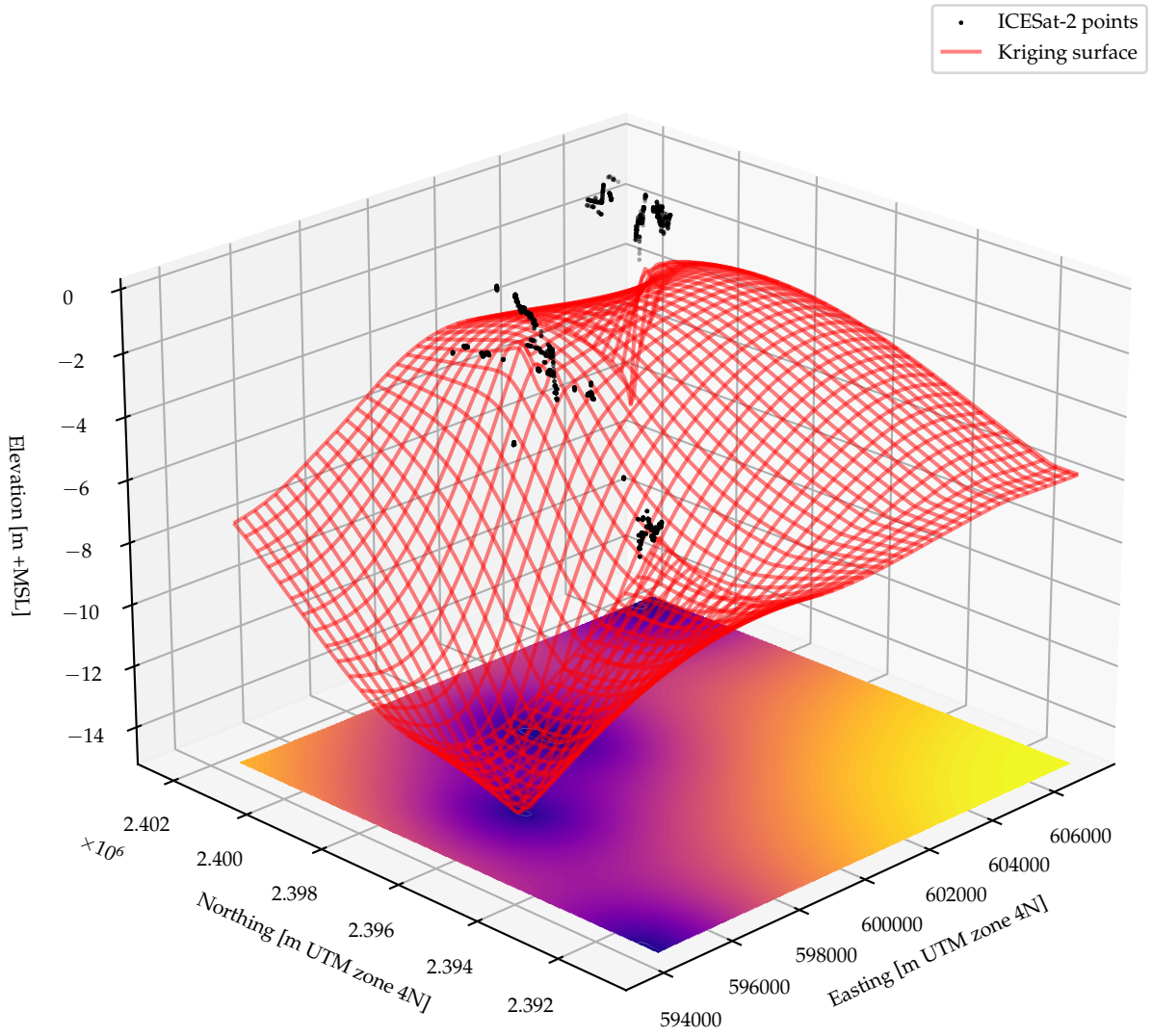


Figure C.44: Oahu Subsite 6: bathymetry points from KDE signal finding.
Basemap data:

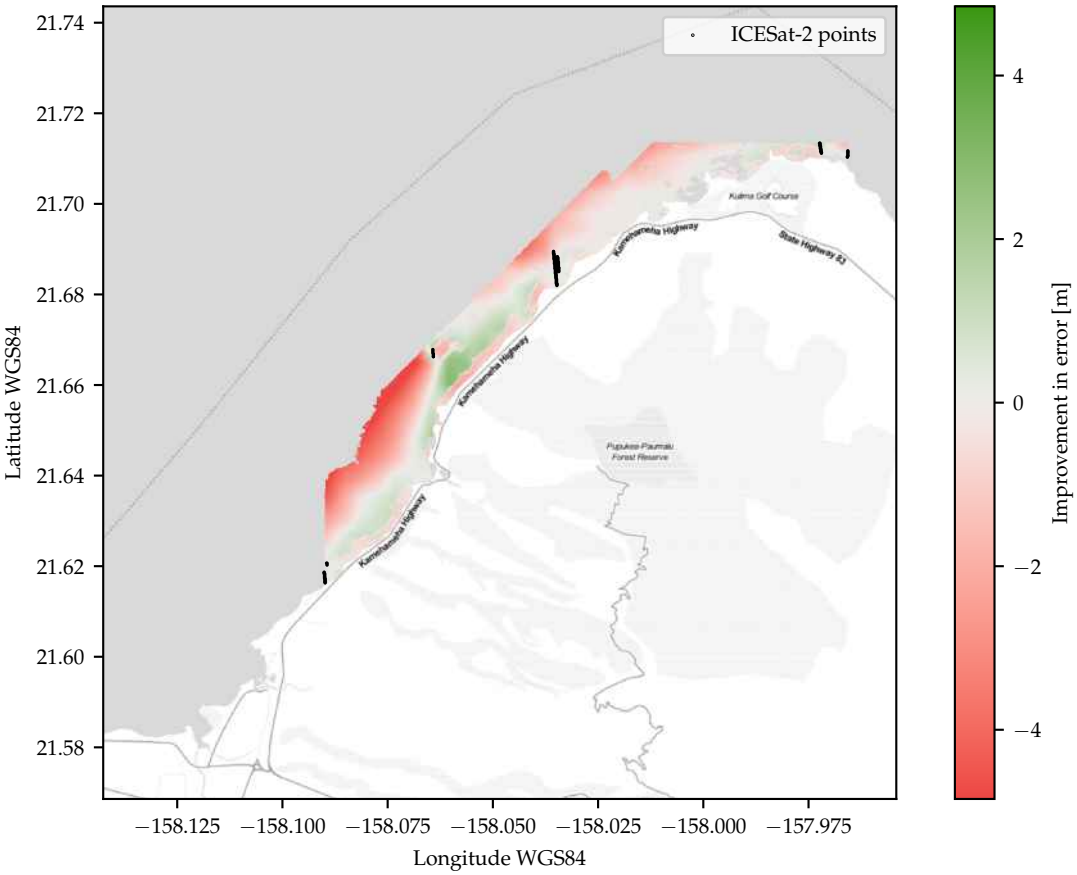


Figure C.45: Error Improvement in Oahu Subsite 6.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.10. Oahu Subsite 7

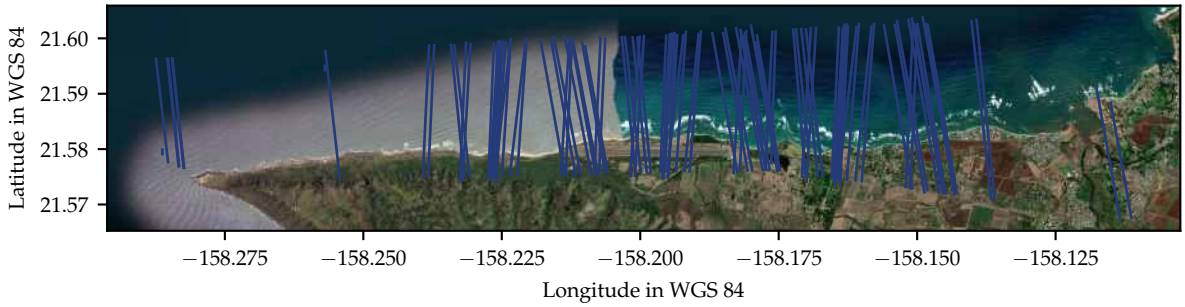


Figure C.46: Oahu Subsite 7: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

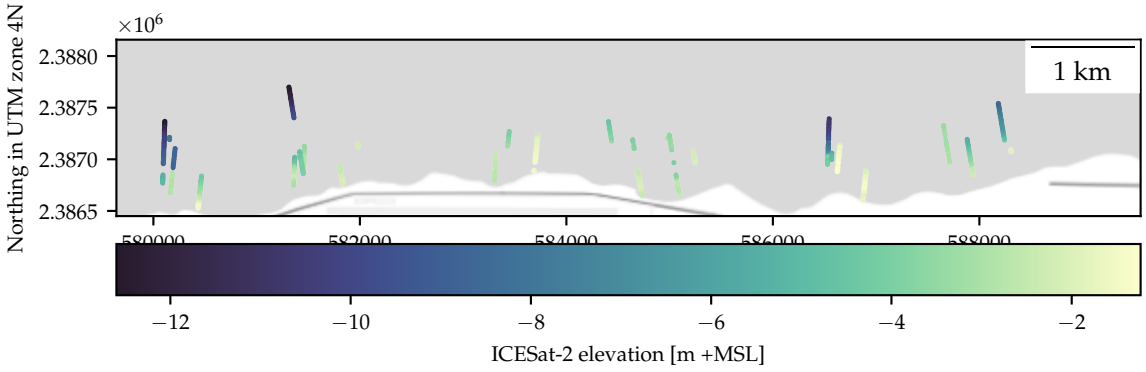


Figure C.47: Oahu Subsite 7: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

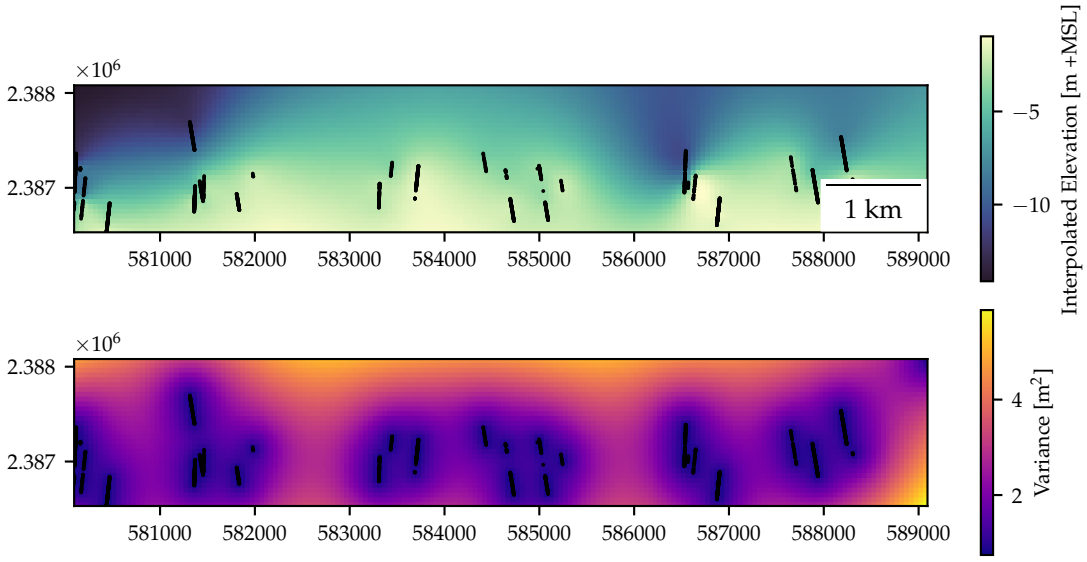


Figure C.48: Oahu Subsite 7: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

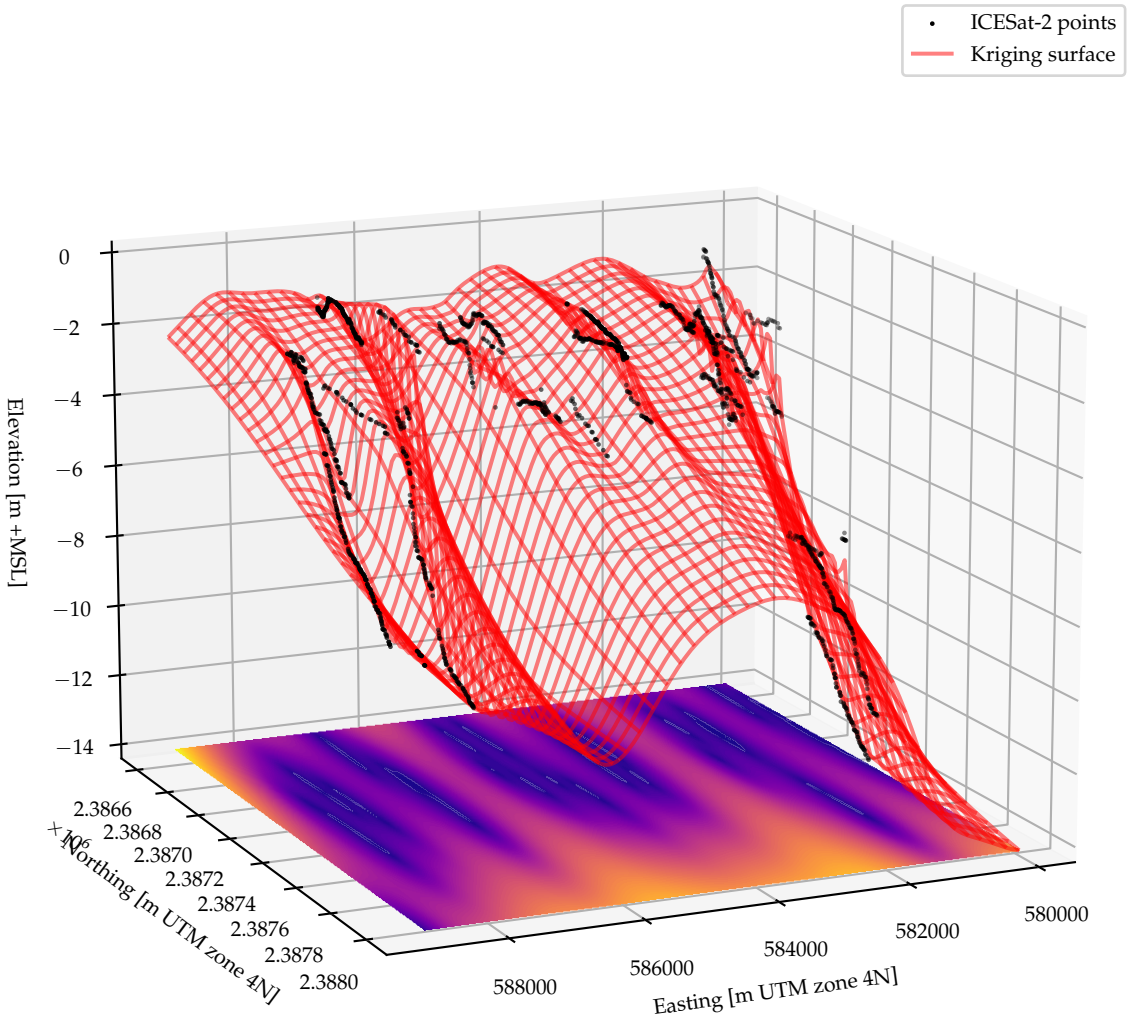


Figure C.49: Oahu Subsite 7: bathymetry points from KDE signal finding.
Basemap data:

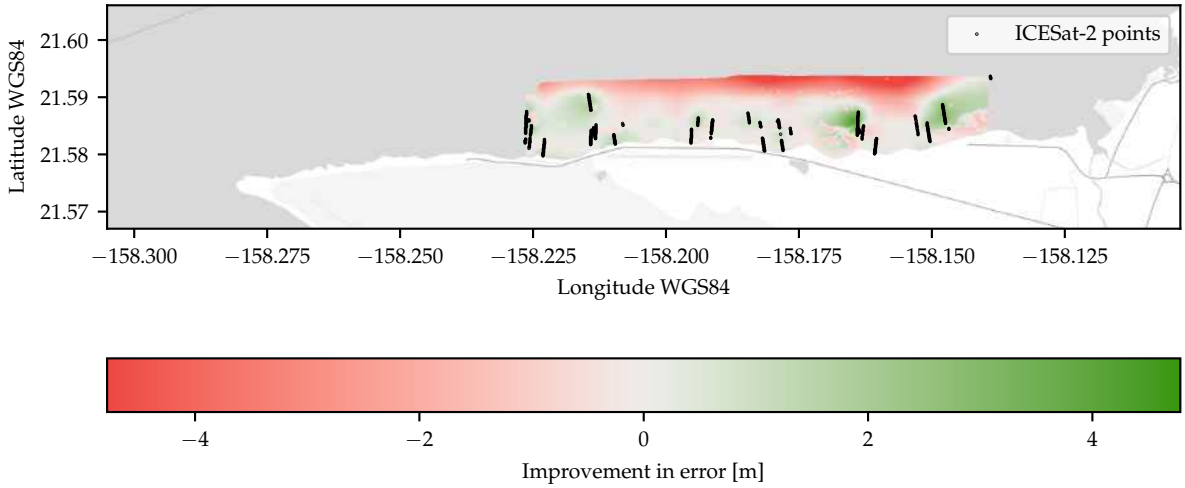


Figure C.50: Error Improvement in Oahu Subsite 7.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

C.11. Oahu Subsite 8

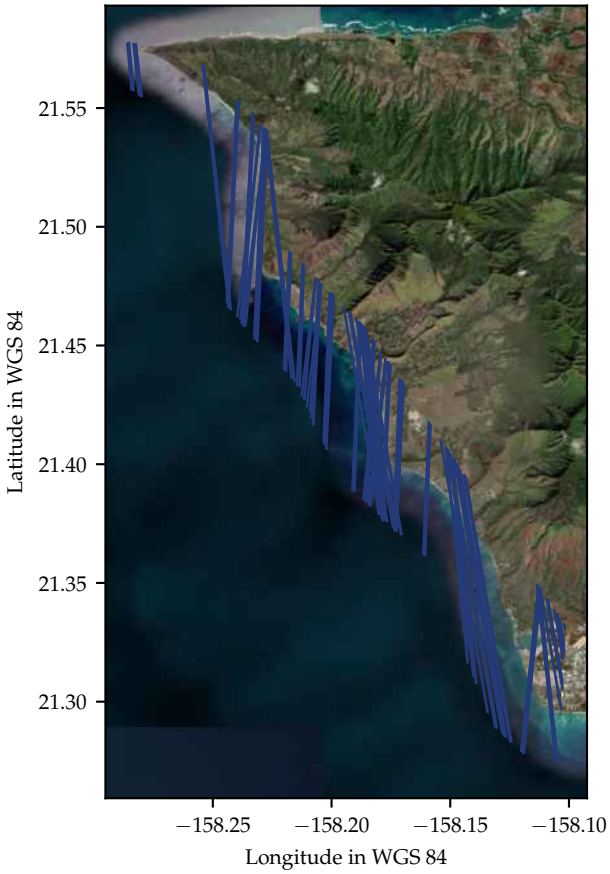


Figure C.51: Oahu Subsite 8: ICESat-2 tracklines over site.
Basemap data ©Esri – Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

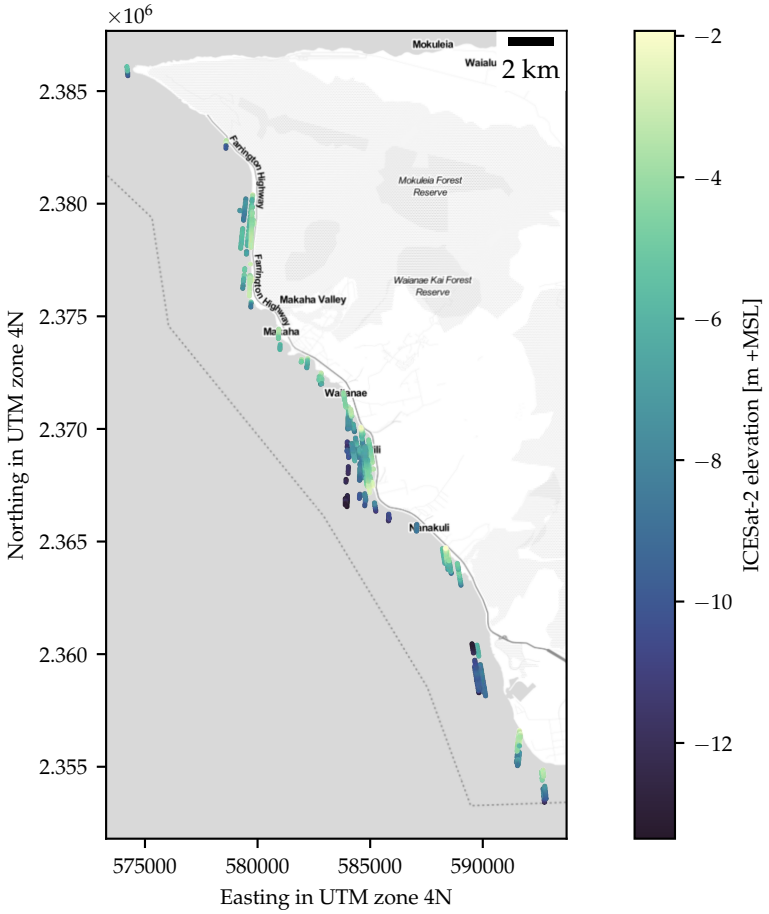


Figure C.52: Oahu Subsite 8: bathymetry points from KDE signal finding.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

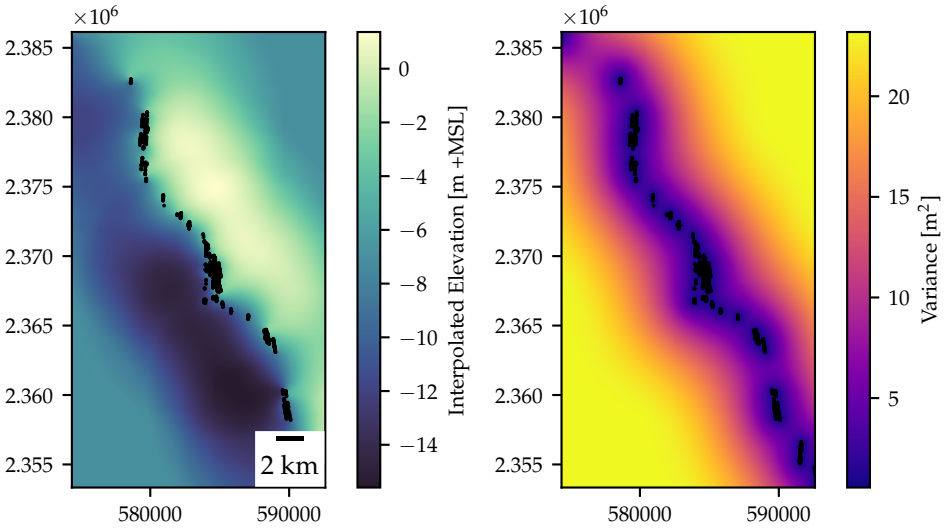


Figure C.53: Oahu Subsite 8: kriging results.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors

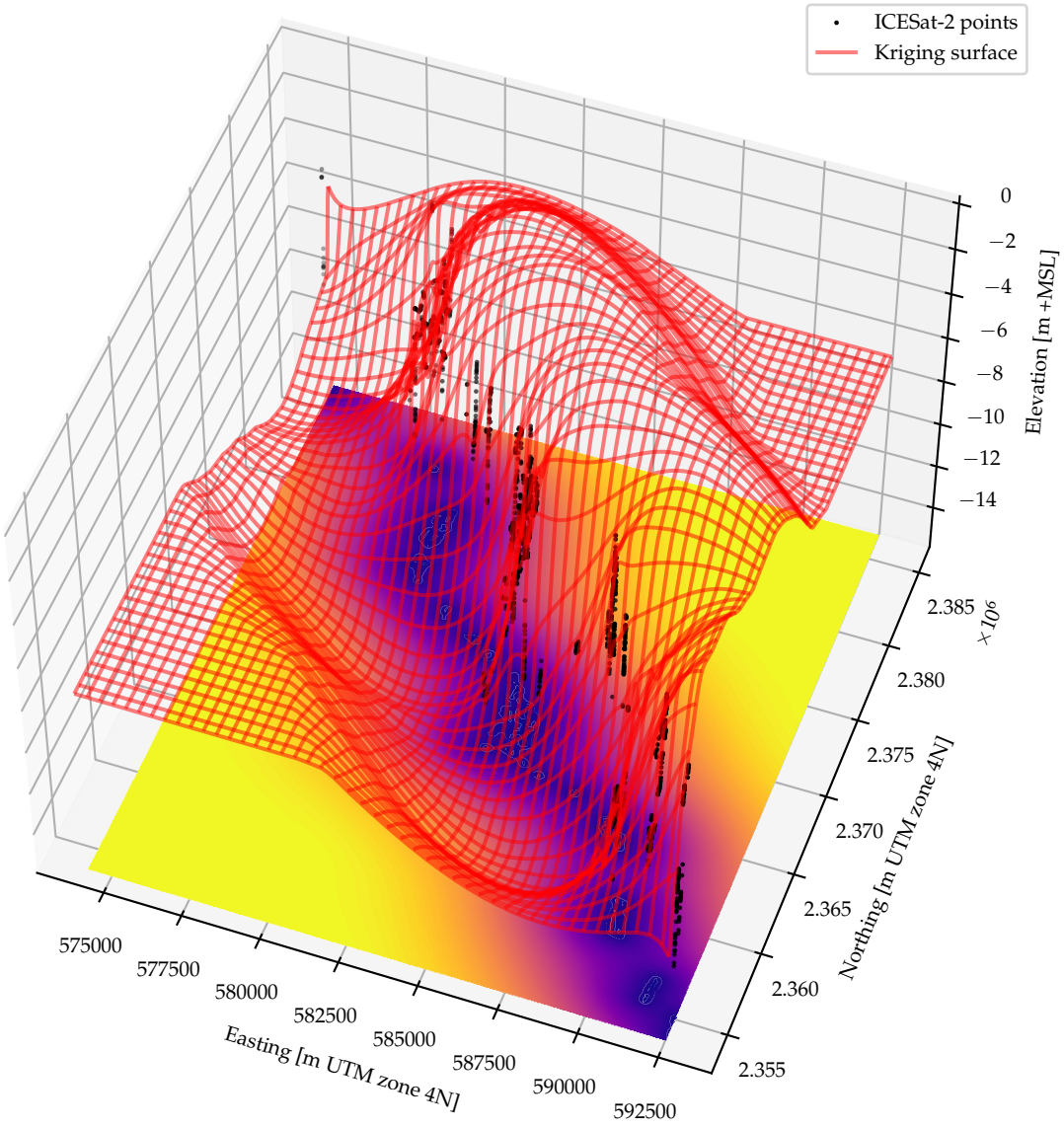


Figure C.54: Oahu Subsite 8: bathymetry points from KDE signal finding.
Basemap data:

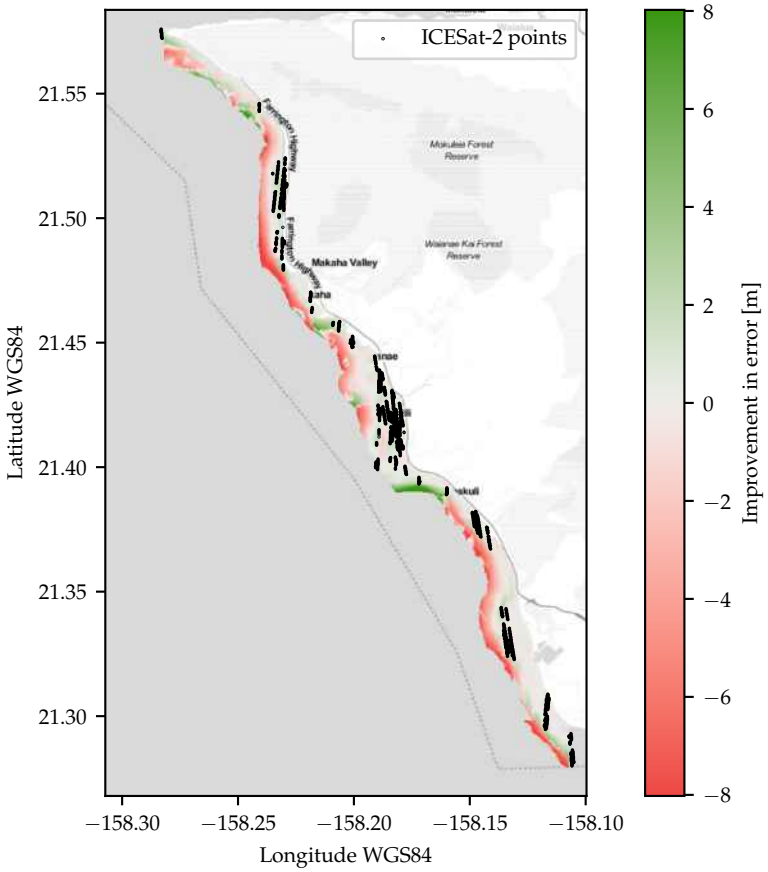


Figure C.55: Error Improvement in Oahu Subsite 8.
Basemap data: Map tiles by Stamen Design, CC BY 3.0 – Map data ©OpenStreetMap contributors