Satellite Derived Bathymetry for Times of Absent In Situ Data

Floortje Burgers

Deltares   **TU**Delft

Estimating satellite derived bathymetry for times of absent in situ data

by

Floortje Burgers

to obtain the degree of Master of Science

at the Delft University of Technology,

Student number: 4215087
Thesis committee: Prof. dr. ir. S.G.J. Aarninkhof, TU Delft
Dr. ir. R. C. Lindenbergh,   TU Delft
Dr. ir. A. P. Luijendijk,   TU Delft & Deltares

An electronic version of this thesis is available at *http://repository.tudelft.nl/.*

Cover image:
European Space Agency
Earth from Space: Changing faces of the Dutch Wadden Sea

TUDelft

Preface

This thesis marks the end of my master's degree in Geoscience and Remote Sensing at Delft University of Technology. The topic 'Satellite Derived Bathymetry' represents the program that I carefully put together to combine my interests for the ocean and remote sensing. Several graduation topics sparked my enthusiasm yet I am very grateful for the opportunity to research this particular topic at Deltares. Unfortunately the global pandemic completely changed my graduation year halfway through, from working in an inspiring environment to working from my tiny home. Alltogether, I am proud to present a year of hard work into this one document.

I would like to thank Arjen Luijendijk for introducing me to this exciting topic and for taking me on as a graduate student at Deltares. The opportunity to experience research outside of a university setting has been super valuable. Your practical point of view helped me to see the bigger picture when I got caught up in details. The idea for this thesis sprouted with you, I hope there is potential in my work to serve your projects at Deltares.

Thank you to Roderik Lindenbergh for supervising and guiding me. You included me in your group of research students and offered me somewhat of a handhold amidst COVID restrictions which really supported me. Your enthusiasm is contagious.

I offer my appreciation to Stefan Aarninkhof, who chaired my thesis committee. Although we only had a few meetings, your constructive feedback and supportive comments have helped me greatly. I will carry some of your wise words with me into my new chapter, whatever it may be.

Floortje Burgers
Delft, November 2020

Abstract

Knowledge of seafloor topography (bathymetry) is increasingly important as coastal environments are unprecedentedly stressed by climate change and anthropogenic pressure. The bathymetry of shallow nearshore waters is yet marginally monitored due to costly and time-intensive survey techniques. Methods to obtain satellite derived bathymetry (SDB) have become increasingly valuable. Mapping temporal change is however challenging, because the majority of these methods remain heavily dependent on situ observations. This thesis introduces an SDB approach to estimate temporal bathymetric changes, which omits the need for synchronous in situ data.

The approach is based on a reference image correction method that enables direct comparison of multitemporal imagery and temporal extrapolation of a conventionally-trained bathymetry estimation model (BEM). Research focused on pre-processing multispectral imagery, developing a bathymetry estimation model and estimating bathymetry for times of absent in situ data. The proposed method is demonstrated with a case study in the Dutch Wadden Sea; a site characterised by dynamic morphology, high turbidity and homogeneous bottom type. A log-linear estimation model is obtained by linear regression on in situ observations and the three visible bands of Sentinel-2 imagery. Scarcity of high-quality Sentinel-2 imagery is managed by combing multiple images into a six-month composite. The availability of two sets of *vaklodingen* in situ observations allowed for training and testing two bathymetry estimation models (BEM 2016 and BEM 2019) and for cross-validating the depth estimates after a three-year extrapolation of these models.

Bathymetry is estimated for times of absent in situ data by temporal extrapolation of the two estimation models. The extrapolation showed estimation of shallow bathymetric structures in up to four metre water depth with an RMSE of approximately one metre. Additionally, the migration direction of these bathymetric structures is successfully estimated. Within the tested three-year time frame, predictive power did not decrease. These results imply that estimation performance is governed by composite quality and predictive power of the bathymetry estimation model. The limited influence of temporal extrapolation on estimation performance suggests that the availability of high-quality satellite imagery and one set of non-synchronous in situ observations is sufficient to estimate bathymetry for times of absent in situ data. The proposed method potentially provides a tool for mapping temporal bathymetric changes of nearshore zones across the globe.

Contents

List of Figures

<div style="text-align: right">

1

</div>

Introduction

Coastal environments are unprecedentedly stressed by climate change and increasing anthropogenic pressure. Understanding and predicting future changes therefore become increasingly important. To assess potential impacts, detailed mapping and continuously monitoring of marine systems is required Accurate mapping of bathymetry is relevant to many areas of marine science and engineering, such as management of the marine environment, industrial development and coastal research studies.

Bathymetric surveys are traditionally ship-based observations from multi-beam sonars or aircraft-based measurements from lidar scanners. Ship-based surveys provide accurate depth observations at high spatial resolution, these time-intensive and costly surveys are however limited to one-dimensional transects in accessible, non-shallow waters. Airborne surveys alternatively offer swift access to remote areas but these campaigns are limited by weather conditions and in spatial coverage. In addition, lidar surveys are costly and constrained to optically clear waters.

Spaceborne remote sensing potentially provides a tool for global bathymetric mapping with frequently-recorded data, albeit at lower resolution than multi-beam sonar or airborne lidar surveys. Obtaining satellite derived bathymetry (SDB) from optical sensors has been extensively studied since the 1970s (Lyzenga, 1978). The spatiotemporal resolution of optical imagery has since increased significantly, for example with the launch of the Sentinel-2 satellites. Multispectral imagery from this mission is publicly available with a repeat-visit up to every five days and a spatial resolution of ten metre for its three visible bands (European Space Agency, 2015).

Estimation approaches that obtain bathymetry from spectral information vary from empirical to (semi-) analytical models. Analytical approaches simulate the propagation of light through the atmosphere and the water column by inversion of radiative transfer models (e.g. Hedley et al., 2009, Lee et al., 1999). Although analytical models require no in situ data for calibration, these approaches are strongly dependent on knowledge of the optical properties of the ocean and the characteristics of the seafloor (e.g. Gao, 2009, Hedley et al., 2009). Empirical approaches derive a relation between the intensity of a spectral image and in situ depth observations, for which methods vary from simplified theoretically-oriented models (e.g. Lyzenga, 1978) to extensive statistical machine learning techniques (e.g. Sagawa et al., 2019). The majority of bathymetry estimation approaches is heavily dependent on in situ observations, whether it be measurements of optical water properties as input for analytical models or depth observations to train empirical models. Obtaining in situ measurements is expensive and consequently not frequently repeated.

Bathymetry estimation is strongly limited by the scarce availability of in situ observations in both space and time. Mapping temporal bathymetric changes is especially challenging because it requires bathymetry estimation for times of absent in situ data. To overcome this problem, I propose an approach to estimate temporal bathymetric changes which omits the need for synchronous in situ data. The approach is based on temporal extrapolation of an empirically-derived bathymetry estimation model. To investigate the potential of the proposed method, the following research questions are defined:

<div style="text-align: center">

1

</div>

To what extent can bathymetry be derived from multispectral imagery, for times of absent in situ data?

In support of this main question, four sub-questions are defined.

- How to enable direct comparison of multitemporal satellite imagery in order to map temporal bathymetric changes??

- How to manage the expected scarcity of high-quality images due to challenging conditions associated with the study site?

- What approach is best-suited to relate water depth and multispectral imagery for the site-specific conditions?

- What is the performance of extrapolation in terms of detectable bathymetric changes, temporal limits and estimation quality?

This thesis is structured as follows: Chapter 2 provides some background information on bathymetric structures and remote sensing and presents insight into defining a bathymetry estimation model. Chapter 3 first introduces the case study and the data and then illustrates the proposed method to map temporal bathymetric changes. This consists of pre-processing the multispectral imagery to isolate the bathymetric signal, developing a bathymetry estimation model and extrapolation of this model. Chapter 4 discusses the results to assess to which extent the proposed method can be used to estimate bathymetry for times of absent in situ data. Chapter 5 summarises the main findings based on the research questions and provides future recommendations.

2

Background information

Estimating bathymetry from satellite imagery requires some understanding of the physical processes underlying both the dynamics of coastal zone and the propagation of light. Across this combined topic, spatial and temporal scales vary extremely from an electromagnetic wave with the speed of light to the development of a coastline over decades. Knowledge of the characteristics of the case study and the applications of remote sensing are essential to obtain a bathymetry estimation model appropriate for the local conditions. In support of the Wadden Sea case study (introduced in Section 3.1.1), this chapter first provides a brief introduction to the morphodynamics of a tide-dominated system (Section 2.1). Next, the chapter presents a background on electromagnetic radiation and its properties (Section 2.2) and an introduction on how to derive bathymetry from satellite images (Section 2.3).

2.1 Morphodynamic changes of a tide-dominated system

The coastal zone is controlled by a balance between constructive and destructive forcings, which form a complex system of coupled interactions between land and water. Sediment is continuously eroded, transported and deposited under the influence of physical and ecological processes in the water. A disparity between deposition and erosion governs the development of structures on the seabed. These structures shape the seafloor topography, which is also known as bathymetry.

The morphological structures on the seabed, known as bedforms, form as a result of the interaction between water and sediment. Their spatial and temporal scales are closely connected, large features develop over long periods of time. For example, large tidal channels and sand banks with surface dimensions from 5 to 20 km^2 change on a timescale from years to decades, while smaller structures with dimensions up to several km^2 evolve within years (Bosboom and Stive, 2012).

Tide-dominated shorelines are especially dynamic systems where bedforms and bathymetry are constantly changing. Tidal basins and barrier coasts were formed as a result of flooded land due to post-glacial sea level rise (De Swart and Zimmerman, 2009). The morphological development of these systems is controlled by the relative influence of hydrodynamics forcings by winds and tides.

Tidal basins are part of highly-dynamic tidal inlet systems, which consist of unique morphological features that each exhibit characteristic temporal evolution. Typical morphological elements of a tidal inlet systems are the ebb-tidal delta, the tidal channels, the tidal flats and the total channel volume. (van Veen et al., 2005(1950)) (Figure 2.1). The ebb-tidal delta is a fan-shaped accumulation of sediment seaward of the tidal inlet, that was deposited from the ebb current flowing out of the tidal basin. Alongshore sediment transport bypasses the inlet which may result in small depositional structures on top of the ebb-tidal delta. Over the period of a tidal cycle, large volumes of water are exchanged between the open ocean and the enclosed basin, which naturally keeps the inlets from closing. Flood currents enter the tidal basin through deep channels and across shallow intertidal flats, which fall dry during low water. The channels are characterised a deep main channel with tidal bars and a bifurcating pattern of side channels which gradually decrease in depth towards land (De Swart and Zimmerman, 2009, van Veen et al., 2005(1950)).
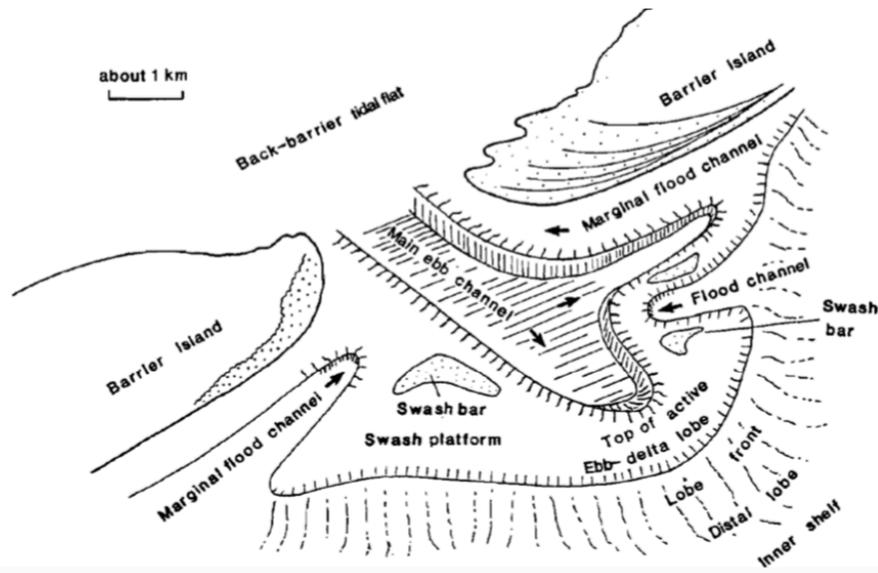
Figure 2.1: Illustration of the morphological elements of a tidal inlet system in the Dutch Wadden Sea. Source: (Sha, 1990)

The morphodynamic development of these tidal inlet systems occurs on a wide range of spatial and temporal scales. The smallest features form due to the interaction of water and sediment and develop within a day, while the elements of the tidal inlet system, such as the tidal basin, extend over kilometres and develop over centuries (Bosboom and Stive, 2012). The principal morphological features develop due to seasonal and interannual variations in weather and tidal forcings. An example are the tidal channels, which span hundreds of metres and evolve on a scale of years.

2.2 Electromagnetic radiation

Electromagnetic radiation emitted by the Sun can be used to obtain bathymetry. Due to interactions with the atmosphere and the ocean, the bathymetric signal only contributes marginally to the composed signal that is recorded by a satellite sensor.

2.2.1. An introduction to electromagnetic radiation

Electromagnetic radiation is radiant energy travelling through space by means of electromagnetic waves. This wave phenomenon is induced by periodically changing magnetic and electric fields (Ishimaru, 1991). Waves are characterised by their wavelength or frequency, which are related through the speed of the wave. In vacuum, electromagnetic waves travel with the speed of light. The energy carried by electromagnetic waves is proportional to their frequency, waves with high frequencies and short wavelengths hold more energy.

The electromagnetic spectrum describes the range electromagnetic waves based on their frequency and classifies them into separate spectral bands (Figure 2.2). The spectrum ranges from gamma waves at the high frequency end of the spectrum to radio rays at the low frequency end. Visible light covers only a narrow section of the electromagnetic spectrum, wavelengths range from approximately 400 nm to 700 nm. This section is commonly divided into even narrower bands of spectral colours, such as red, green and blue.
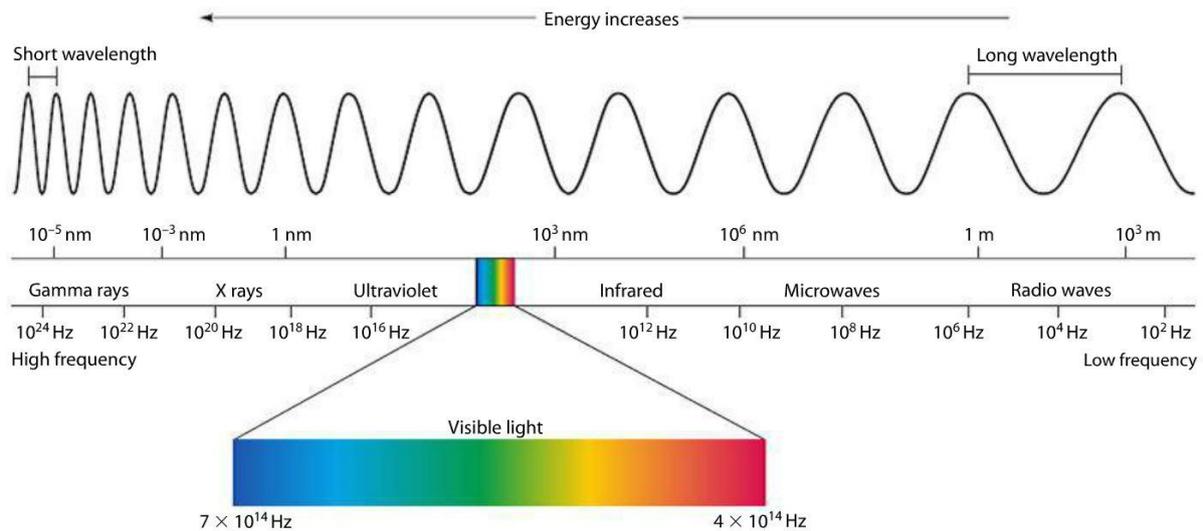
Figure 2.2: Visualisation of the electromagnetic spectrum, illustrates the range of electromagnetic waves based their frequency and wavelength and marks the separate bands. Source: *https://www.miniphysics.com/electromagnetic-spectrum_25.html*

2.2.2. Recording radiation with sensors

Sensors only record electromagnetic radiation from a specific part of the electromagnetic spectrum. For example, the human eye sees visible light while a radio is designed to detect radio waves. Active systems record electromagnetic radiation from their own artificial source, whereas passive sensors record naturally available radiation such as solar radiation. Optical sensors record radiation within the visible, near infrared and short wave infrared parts of the spectrum. The exact configuration of their recording bands however varies. Multispectral sensors record electromagnetic radiation in multiple spectral bands. The red, green and blue band of the visible spectrum are typically included. The instruments on board the Sentinel-2 satellites are an example of multispectral sensors currently in orbit around Earth. For this study, Sentinel-2 imagery is used. Details on the mission and its data are therefore provided in Section 3.1.2.

2.2.3. Attenuation of radiation through a medium

Electromagnetic radiation interacts with the medium through which it propagates; incident radiation may be reflected, scattered in a new direction or exchange energy with the medium. As a result of these interactions, the radiant intensity may decrease (attenuate). The intensity of propagating radiation is attenuated by scattering, which diffuses radiation, and by absorption, which converts radiation into another form of energy. The interaction of an electromagnetic wave and matter is controlled by the relation between the wavelength and the medium. This relative measure is described by the penetration depth, which is defined as the distance where the intensity has reduced to approximately one third. The spectrally dependent penetration depth can be used to express the attenuation of intensity. The law of Lambert-Beer defines the exponential decay of the intensity as function of the penetration depth and the path length of propagation into the medium:

$$I = I_0 \ \exp^{-\frac{\ell}{\alpha}} \tag{2.1}$$

Where I is the reduced intensity, I_0 the initial intensity, α the penetration depth and ℓ the path length.

2.2.4. Earth's energy balance

The Sun is the primal source of energy at Earth and its electromagnetic radiation is attenuated by interactions with the Earth and its atmosphere. Solar radiation is emitted across the shortwave end of the electromagnetic spectrum, ranging from ultraviolet to near infrared (250 to 2500 nm). Incoming solar radiation predominantly consists of visible light (approx. 46 %) and near infrared radiation (approx. 46 %) (Iqbal, 2012). Approximately one third of the incoming radiation is reflection back to outer space (Marshall and Plumb, 2007).

These incoming and outgoing radiation fluxes are part of the Earth's energy balance. Incoming shortwave radiation is absorbed by the surface of the Earth and subsequently emitted as heat and longwave radiation (Marshall and Plumb, 2007). Here, the focus is on incoming and outgoing shortwave radiation (Figure 2.3, adjusted from Ruddiman, 2000). Due to atmospheric interactions, only half of the incoming radiation reaches the surface of the Earth. From this flux, that actually reaches Earth's surface, the majority is absorbed and just a fraction is reflected back to space. The incoming radiation that never reaches Earth, is absorbed or redirected by scattering or reflection by the atmosphere. The outgoing signal predominantly consists of this radiation, that is scattered or reflected by the atmosphere (>85 %).
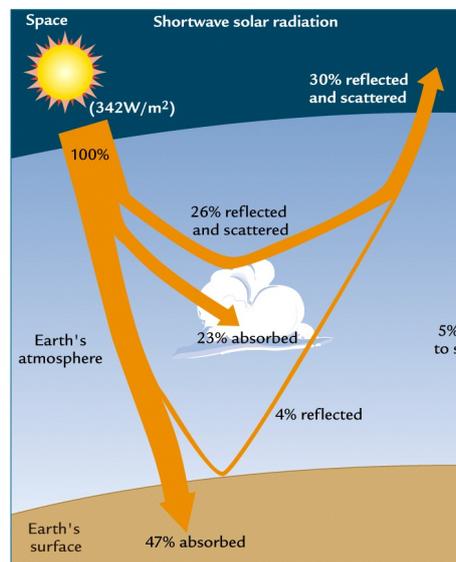
Figure 2.3: Incoming solar radiation interacts with the atmosphere and the Earth. The majority (70 %) of the incoming radiation is absorbed, the remainder (30 %) is reflected back to space. Reflections from the Earth's surface are a small fraction (approx. 15 %) of the outgoing radiation. This cartoon demonstrates the effect of the interactions between electromagnetic radiation and Earth's atmosphere. Source: Ruddiman (2000)

The Earth's atmosphere consists of a mix of gases, some of which absorb radiation from specific parts of the electromagnetic spectrum. The atmosphere can therefore only be penetrated by radiation of certain frequencies. In the visible spectrum, the atmosphere is almost completely transparent, while it is very opaque in the ultraviolet spectrum (Marshall and Plumb, 2007). This characteristic absorption spectrum of the atmosphere, which is due to its constituents, is one of the reasons for the decreasing intensity of electromagnetic radiation.

Reflections of electromagnetic radiation in the atmosphere are largely due to clouds. Scattering changes the direction of an electromagnetic wave, it can occur with and without the loss of energy (respectively inelastic and elastic scattering) (Platt et al., 2007). Also this type of interaction is determined by the relative relation between the wavelength of the electromagnetic wave and the size of the particle that it interacts with. Rayleigh scattering occurs upon interaction with molecules, while Mie scattering occur when electromagnetic radiation interacts with aerosols in the atmosphere (Platt et al., 2007). Radiation scattered and reflected by the atmosphere makes up the majority of the outgoing shortwave radiation.

2.2.5. The bathymetric signal

Approximately 15 % of the outgoing flux reflected on the surface of the Earth. The fraction of incoming solar radiation that is reflected is known as the albedo. Light surfaces, such as snow or desert, reflect more radiation and thus have a high albedo. Water mostly absorbs electromagnetic radiation, and therefore has an albedo of 2-10 % (Marshall and Plumb, 2007).

In liquid state, water absorbs radiation across a wide range of the electromagnetic spectrum. Absorption across the visible spectrum is however weak (e.g. Buiteveld et al., 1994, Morel, 1974) (Figure 2.4). Absorption across this narrow band of visible light is variable, wavelengths with long wavelengths and low energy are absorbed first. Blue light is therefore absorbed the least, because it has the shortest wavelength in the visible spectrum. As a result, blue light penetrates deeper into the water column than red light. The law of Lambert-Beer can be used to quantify this attenuation of light in water (Section 2.2.3).
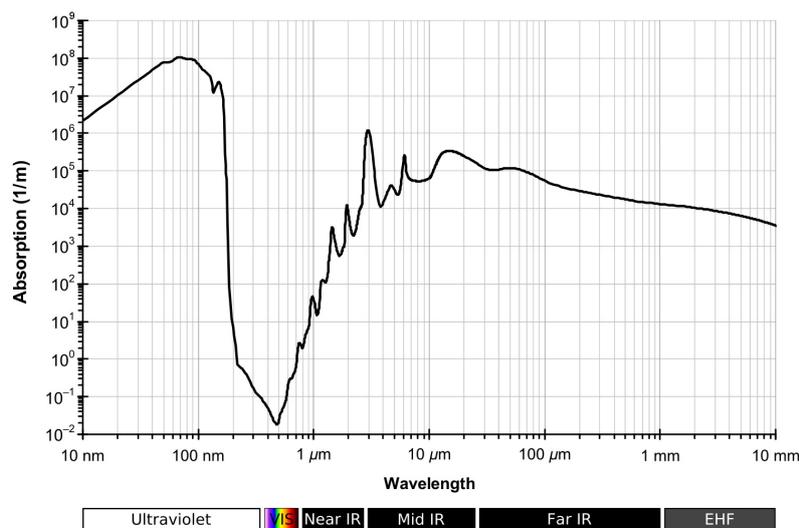
Figure 2.4: Absorption spectrum of liquid water. The majority of electromagnetic radiation is strongly absorbed by water. Visible light is however absorbed weakly; only electromagnetic radiation at these wavelengths propagates into the water column. Source: *https://commons.wikimedia.org/wiki/File:Absorption_ spectrum_ of_ liquid_ water.png*

In nearshore waters, incoming solar radiation may reflect on the seafloor if the intensity of the incoming radiation is sufficiently large and the depth sufficiently shallow. In this case, the reflected signal is a product of the water depth and the reflective properties of the bottom type. This bathymetric signal, which reflected on the seafloor, propagates towards space. All outgoing reflections result from interactions between electromagnetic radiation and the medium through which it propagated. The total outgoing radiation flux is thus the summation of all the different processes, of which the bathymetric signal is only a very small fraction.

2.3 Satellite derived bathymetry

The intensity of the reflected signal can be used as a proxy for water depth, due to the relation between the intensity of radiation and the properties of the water, as defined by the law of Lambert-Beer (Section 2.2.3, Equation 2.1). As a result of the absorption spectrum of water (Section 2.2.5), visible light can be used to obtain information of the water column and thus optical remote sensing is used to derive bathymetry.

2.3.1. Empirical models vs. analytical models

Algorithms to derive bathymetry from optical satellite imagery are predominantly based on the physics of light attenuation in water. These bathymetry estimation models are commonly classified into analytical and empirical approaches. This section briefly discusses analytical and empirical models in order to provide context for the empirical estimation model that is used in this study.

Analytical approaches employ the inversion of radiative transfer models which are priory tuned. Forward modelling is used to calculate a range of reflected intensities from the properties of the water column and the bottom type. For each pixel, the observed intensity is compared to the modelled intensities and the optimal model is inverted to estimate depth (Hedley et al., 2009, Kerr and Purkis, 2018, Lee et al., 1999). Although analytical models require no in situ data for calibration, expert knowledge of a range of water properties and seafloor characteristics is essential to apply a radiative transfer model (Gao, 2009). By simultaneously deriving water depth and reflected intensities, analytical approaches are robust and not sensitive to variations in bottom type (Bierwirth et al., 1993).

Empirical algorithms derive light attenuation properties and water depth by performing regression analysis on in situ depth observations. These empirically calibrated approaches are much simpler and require less input parameters than the complex and computationally-expensive analytical models (Gao, 2009). The performance of empirical models is limited by water turbidity and the influence of the seafloor surface, they are however widely used (e.g. Lyzenga, 1978, Pacheco et al., 2015, Stumpf et al., 2003). Although, analytical models are extremely sensitive to atmospheric corrections (Casal et al., 2020, Hedley et al., 2018), accurate corrections are also important in empirical approaches. Consistent corrections are in particular crucial for temporal analysis of SDB, because imagery must be directly comparable (Caballero and Stumpf, 2020a, Casal et al., 2019)

2.3.2. Atmospheric corrections

Atmospheric effects have a significant influence on the signal received by a satellite-based sensor (Section 2.2.4). The contribution of atmospheric effects to the recorded signal is in the visible spectrum up to 90 % (Gordon and Morel, 1983). Correcting for these atmospheric influences is therefore crucial in order to isolate the bathymetric signal and obtain accurate depth estimates. Currently, a wide variety of correction schemes is available. Atmospheric corrections can be classified into image-based approaches and methods that rely on atmospheric radiative transfer models, the latter requiring in situ observations (Hadjimitsis et al., 2004).

The darkest object subtraction (DOS) method, first introduced by Chavez Jr (1988), is a widely-used atmospheric correction for bathymetric purposes (e.g. Geyman and Maloof, 2019, Lyzenga, 1978, Pacheco et al., 2015). The DOS method is an image-based correction which derives the atmospheric effects from an image itself, it thus requires no external data and can therefore be applied to any multispectral image. The reflectance from a significantly dark object is assumed to predominantly consist of signal reflected or scattered by the atmosphere. The atmosphere is assumed to be constant across the image such that each pixel is affected equally by the atmospheric effects and its constant offsetting effect can be removed by subtracting the intensity of the darkest pixel.

Sen2Cor is a physics-based algorithm to apply atmospheric corrections on multispectral imagery that was developed by the European Space Agency (ESA) to correct imagery from their Sentinel-2 satellites (European Space Agency, 2015). The Sen2Cor correction is based on the ATCOR model (Richter, 1990) which employs the inversion of radiative transfer model by using lookup tables (Main-Knorn et al., 2017). ESA distributes Sen2Cor freely and offers a selection of readily processed Sentinel-2 imagery. The Sen2Cor corrections have been used in recent studies on the derivation of bathymetry from multispectral images (e.g. Casal et al., 2019, Evagorou et al., 2019, Poursanidis et al., 2019).

2.3.3. Defining a bathymetry estimation model

The reflected signal recorded by a multispectral sensor can be used to estimate water depth according to the law of Lambert-Beer (Section 2.2.3). Lyzenga (1978) was the first to define the log-linear relation between observed water depths and the intensity of the recorded signal within a given spectral band:

$$z = a \ \log I_k + b \tag{2.2}$$

Where z is the water depth, I the intensity of the recorded radiance, k denotes the spectral band and a and b are the coefficients describing the linear relationship.

The relation between depth and intensity is better constrained by combining multiple spectral bands (e.g. Lyzenga, 1978, Stumpf et al., 2003). A linear combination of multiple spectral bands is given by:

$$z = b + \sum_k a_k \log I_k \tag{2.3}$$

Where z the water depth is now a function of each spectral band k, with I_k the intensity of the recorded radiance and a_k and b the coefficients defining the linear relationship.

The log-linear model (Equation 2.3) assumes a uniform bottom type, which makes the algorithm very sensitive to bottom type heterogeneity. Stumpf et al. (2003) provide an alternative independent of bottom type, by using the ratio of different spectral bands. The simple algorithm from Lyzenga (1978) is however widely applicable and has since been used in many studies that estimate bathymetry from optical satellite imagery (e.g. Pacheco et al., 2015, Poursanidis et al., 2019, Traganos et al., 2018).

2.3.4. Using in situ observations to obtain the estimation model

In situ observations can be used to obtain the coefficients defining the relationship between water depth and intensity (Equation 2.3). Regression analysis establishes the relationship between independent and dependent variables in order to predict quantitative outputs (Hastie et al., 2009). Empirical bathymetry models apply regression analysis on in situ depth observations and recorded radiances to estimate water depth.

For a data set of N_e depth observations and the three visible bands of an image (red, green and blue), Equation 2.3 can be written in matrix form as a set of N_e equations and 4 unknown coefficients:

$$\mathbf{z} = \mathbf{A}\mathbf{x} = \begin{bmatrix} \log I_{r_1} & \log I_{g_1} & \log I_{b_1} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \log I_{r_{n_e}} & \log I_{g_{n_e}} & \log I_{b_{n_e}} & 1 \end{bmatrix} \begin{bmatrix} a_r \\ a_g \\ a_b \\ b \end{bmatrix} \tag{2.4}$$

Where $\mathbf{z}$ is the vector of depth observations (size $N_e \times 1$), $\mathbf{x}$ the vector with the unknown coefficients (size 4×1) and $\mathbf{A}$ (size $4 \times N_e$) the design matrix relating these two vectors through the log-transformed recorded radiances I_k.

The log-linear relation between water depth and intensity allows for linear regression by applying least squares estimation. Least-squares estimation is a common approach in regression analysis to manage overdetermined systems, which have more observations available than necessarily required to estimate the unknown parameters of interest. Often there is no exact solution due to measurements errors, and thus are overdetermined systems generally inconsistent (Verhagen and Teunissen, 2017). By including the errors in the linear system and subsequently minimising the sum of the squared errors, the least-squares solution is a unique fit to the data.

The linear system including measurement errors is then defined as:

$$\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{e} \tag{2.5}$$

Where $\mathbf{e}$ is the vector of errors (size $N_e \times 1$).

The unknown coefficients in $\mathbf{x}$ are obtained by minimising the squared sum of errors. The estimated solution to the unknown parameters is the least squares estimator $\hat{\mathbf{x}}$, which is given by:

$$\hat{\mathbf{x}} = \arg\min(\mathbf{e}^\top\mathbf{e}) = \arg\min(\mathbf{z} - \mathbf{A}\mathbf{x})^\top(\mathbf{z} - \mathbf{A}\mathbf{x}) = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{z} \qquad (2.6)$$

Where $\hat{\mathbf{x}}$ is the least squares estimator which is the solution to the unknown coefficients in the vector $\mathbf{x}$ that is obtained by using the observations vector $\mathbf{z}$, design matrix $\mathbf{A}$ and error vector $\mathbf{e}$.

3

Methodology

This chapter presents a method to estimate temporal changes from satellite derived bathymetry. The proposed method is based on tailored processing of multispectral imagery and temporal extrapolation of a conventionally-trained bathymetry estimation model. The chapter introduces the Wadden Sea case study and data that are used in this study (Section 3.1) and the workflow which is structured into three components (Figure 3.1). First, the processing of multispectral imagery to isolate the bathymetric signal (Section 3.2). Second, the development of a bathymetry estimation model (Section 3.3). Third, the temporal extrapolation of the trained estimation model to estimate bathymetry for times of absent in situ data (Section 3.4).

Figure 3.1: Visual summary of the presented workflow that enables estimating bathymetry from multispectral imagery for times of absent in situ data. The workflow is divided into three sections; processing multispectral imagery to isolate the bathymetric signal, deriving a bathymetry estimation model and estimating temporal bathymetric changes.

To verify the capabilities of temporal analysis using high-quality multispectral images, research has focused on pre-processing satellite images which includes removal of clouds and a novel reference image correction method for the correction of large-scale non-bathymetric signals. In addition, an optimised approach for the creation of composite images is studied (Section 3.2).

To describe the relation between depth and the recorded intensity of a multispectral image, a depth estimation model is employed which is based on the propagation of visible light in water (Section 3.3). Study site characteristics such as turbidity govern the attenuation of light whereas bottom habitat controls the reflection of incoming radiation. Therefore, a bathymetry estimation model should be adapted to the case study. The bathymetry estimation model is obtained by using in situ depth observations to train a linear regression model. The availability of two sets of in situ observations allows for training and testing two bathymetry estimation models (BEM 2016 and BEM 2019).

By temporal extrapolation of the readily-trained bathymetry estimation model, bathymetry is estimated for times of absent in situ data (Section 3.4). Extrapolated estimates are compared with in situ observation to verify the estimation performance. Additionally, predictive power of the extrapolated estimation model is compared with a benchmark performance.

3.1 Case study

Testing and validating the presented method requires a case study with a dynamic environment for which repeated in situ observations are available. The Wadden Sea is a tidal system where bathymetry continuously changes, its Dutch waters are surveyed by the national agency Rijkswaterstaat approximately every three years. Multispectral imagery is available from a wide range of sources. Imagery of ESA's Sentinel-2 mission is especially suited given its high revisit frequency of approximately every five days and its open access. This section presents the characteristics of the Wadden Sea case study (Section 3.1.1) and the details of the available data sets that are used for this study (Section 3.1.2).

3.1.1. The Wadden Sea tidal system

The Wadden Sea is world's largest tidal flat system which stretches along the coasts of The Netherlands, Germany and Denmark. It is recognised for its global importance and has been listed as a UNESCO world heritage site since 2009. The coastal wetlands consist of a variety of habitats which include tidal areas, estuaries and salt marches. The tidal flats with its complex pattern of channels are the most important habitat, covering more than 10,000 km^2. A series of barrier islands shelters the tidal flats from the North Sea, the back-barrier Wadden Sea is only connected through the tidal inlets.

With its barrier islands and tidal inlets it forms the highly dynamic Wadden system, where bathymetry is continuously changing (Wang et al., 2012). Characteristic features such as a large ebb-tidal channel and deep tidal channels (Section 2.1), indicate that its major inlets are predominantly shaped by tidal forcings (Elias et al., 2012). Sediment is transported alongshore by waves, bypassing the inlets and forming bedforms on top of the ebb-tidal delta (Sha, 1989). Its dynamic environment makes the Wadden Sea a suitable area to investigate the potential of mapping these morphodynamic changes by using multispectral imagery.

The Wadden Sea consist of a large variety of habitats such as its tidal flats and seagrass beds. In an ecological context, habitat is often used to describe the physical environment of a living organism. In remote sensing however, it just refers to surface features, which can have for instance geological, geomorphological or biological origin (Green et al., 2000). Sand is the principle constituent of the sediment in both tidal channels and on the tidal flats (van Straaten, 1954). Therefore, the tidal flats and tidal channels form a homogeneous habitat with a uniform sandy bottom type in view of remote sensing purposes.

High concentrations of suspended sediment are the primary reason for high turbidity in many coastal waters, including the Wadden Sea (Postma, 1961). Turbidity affects the propagation of electromagnetic radiation through the water column and attenuates incoming solar radiation or the reflected bathymetric signal. Deriving bathymetry from multispectral imagery for the Wadden Sea is therefore limited by high turbidity levels in the water column.

A study site was selected which is representative of the dynamics Wadden Sea system; bathymetric surveys demonstrate the presence of deep tidal channel and shallow tidal flat structures within this site. The study site has a homogeneous bottom type due to absence of seafloor covering habitats such as seagrass beds. The site of approximately 10 by 4 km is located north of the Dutch island Rottumeroog and west of the German island Borkum (Figure 3.2).
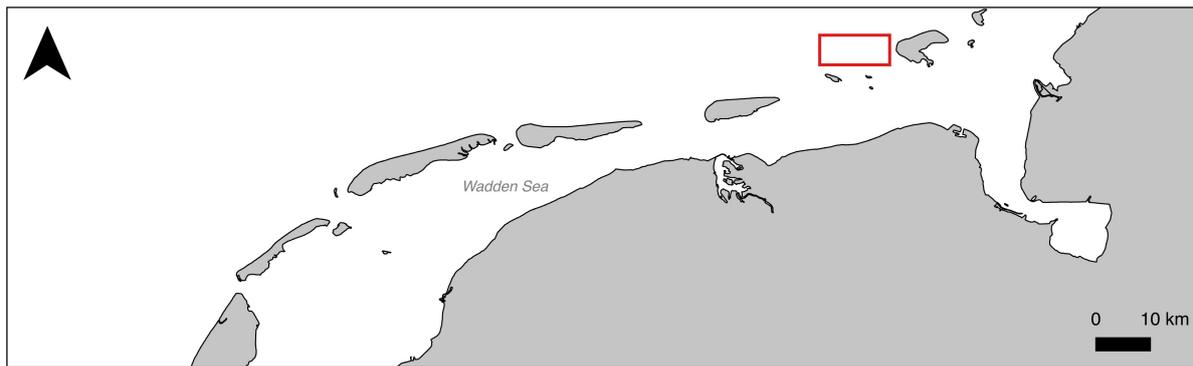
Figure 3.2: Map of the Wadden Sea across the north of The Netherlands. The red square marks the selected study site which is representative of the dynamic environment of the Wadden system due to the presence of a deep tidal channel structure and shallower tidal flats. The channel and flat structures are the only surface features present within the study site, which therefore has a uniform sandy bottom type.

3.1.2. Available data

In situ depth observations in the Dutch Wadden Sea are routinely recorded by Rijkswaterstaat. Approximately every years the bathymetry of delineated zones is acquired. Depths up to approximately -20 m NAP are measured by ship-based multibeam echosounders, emerged areas are additionally surveyed by using laser altimetry. For this study, part of tile KB1110 × 134 was used. Data is interpolated onto a 20 × 20 m grid and available from the OpenEarth database via Deltares (Deltares, 2018). This bathymetry dataset is commonly known as *vaklodingen* and therefore hereafter referred to as such. Figure 3.3 shows an example of the vaklodingen observations for the Dutch Wadden Sea. For this study, the vaklodingen data is which was acquired from July to September 2016 and July to September 2019. Figure 3.4 shows the two vaklodingen data set (2016 and 2019) for the study site.

Figure 3.3: Visualisation of the bathymetry of the Dutch Wadden Sea from a vaklodingen data set. Deep ares such as the tidal channels are depicted in dark blue, whereas shallow waters such as the tidal flats are shown in light colours.
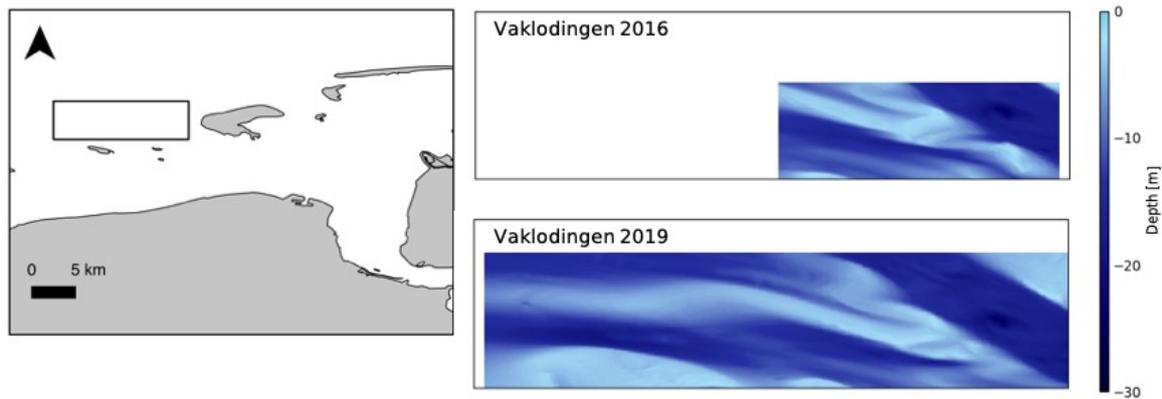
Figure 3.4: Visualisation of the 2016 and 2019 vaklodingen data for the study site marked on the zoomed in map of the Wadden Sea. Deep depths are shown in dark blue, shallow areas are depicted in light blue.

Multispectral imagery from Sentinel-2 is freely available through ESA's Copernicus programme. The Sentinel-2 mission consists of two twin polar-orbiting satellites; Sentinel-2A was launched on 23 June 2015 and Sentinel-2B on 7 March 2017 (European Space Agency, 2015). The Sentinel-2 satellites are equipped with a multispectral sensor that records electromagnetic radiation in thirteen spectral bands (443 to 2190 nm). The three visible bands have a spatial resolution of 10 m, these bands were used for this study. Data is available as tiles of a fixed size of 100×100 km^2.

Sentinel-2 imagery is provided to users at different levels of processing; products are available at Level-1C (L1C) and Level-2A (L2A). L1C data are Top Of Atmosphere (TOA) reflectances that have been corrected radiometrically and geometrically, which includes orthorectification and geo-referencing to the WGS84 datum (European Space Agency, 2015). L2A data are Bottom Of Atmosphere (BOA) reflectances derived from the L1C product by using the Sen2Cor processor (Section 2.3.2). ESA offers users access to the Sen2Cor processor but also has the corrected L2A product readily available since March 2018, for the area of northwestern Europe. Products are embedded with a cloud-type specifying bitmask, which is based on combining information from the blue spectral band and the three short-wavelength infrared (SWIR) bands. Pixels are specified as cloud-free, dense cloud or cirrus cloud (European Space Agency, 2015).

This study employs Sentinel-2 imagery of tile 31U at L1C (TOA) and L2A (BOA) level. Data was accessed through the online JavaScript interface of Google Earth Engine (GEE). GEE is an online cloud platform that provides access to a collection of free geospatial datasets and optical satellite imagery, and offers a high- performance computing system (Gorelick et al., 2017). The original tile was clipped to the study site of approximately 10×4 km^2 (Section 3.1.1). For this research, the three visible bands (Section 2.2.5) and the embedded cloud mask were used.

Figure 3.5: Visualisation of a Sentinel-2 BOA image (22-09-2020) of the Dutch Wadden Sea (processed with Sen2Cor)

3.2 Isolating the bathymetric signal in multispectral imagery

Obtaining clean, high-quality multispectral images suited to temporal analysis, requires specifically tailored image processing. To assess temporal changes of the bathymetric signal, it is important that multispectral imagery is directly comparable. This section presents an approach to isolate the bathymetric signal from non-bathymetric constituents of the recorded signal (Section 2.2.4) developed especially for temporal analysis. First, a reference image is selected (Section 3.2.1) and next, the three components of the presented method are introduced (Figure 3.6); the detection and removal of clouds (Section 3.2.2), a correction for large-scale non-bathymetric variations with the use of a reference image (Section 3.2.3) and a correction of local and high-frequent non-bathymetric signals by creating composite images (Section 3.2.4).
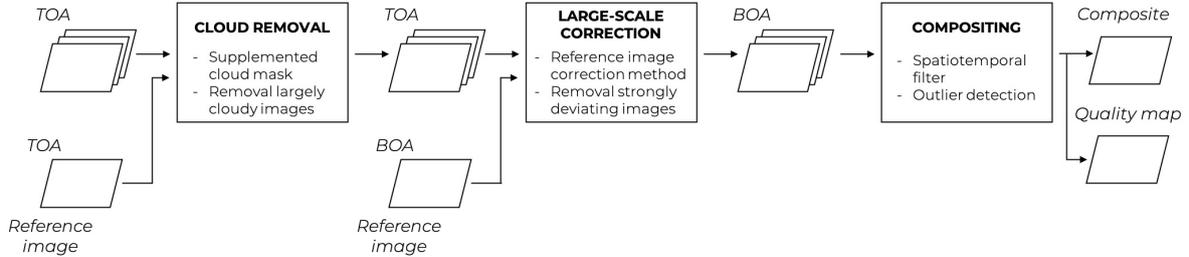
Figure 3.6: Visualisation of the presented workflow to isolate the bathymetric signal from non-bathymetric constituents in multispectral imagery, tailored to temporal analysis. The workflow is structured into three components; i) the removal of clouds, ii) correcting large-scale non-bathymetric variations with the use of a reference image and, iii) correcting local and high-frequent non-bathymetric signals by creating composite images. The correction level of an image is indicated as Top Of Atmosphere (TOA) which is the uncorrected imagery, or Bottom Of Atmosphere (BOA) which is the level after correction.

Clouds obstruct multispectral imagery and pixels affected by cloud coverage should be discarded. The effect of clouds is expressed locally by a binary classifier for each pixel known as a cloud mask. This study employs an ESA cloud mask embedded in the Sentinel-2 product (European Space Agency, 2015) (Section 3.1.2), supplemented with a cloud mask based on the intensity recorded in the red band, for the local detection of clouds (Section 3.2.2). Images largely affected by clouds are completely removed from the data collection.

To correct for large-scale non-bathymetric signals, a novel reference image correction method is designed and applied to the multispectral imagery (Section 3.2.3). Atmospheric and aquatic variations alter multispectral imagery on a large scale, which allows isolation of the local bathymetric signal by using a reference image. Images obstructed by large-scale non-bathymetric effects that cannot be corrected are removed from the data set.

To correct for local and high-frequent non-bathymetric signals, composite images are created. Combining multiple images into one composite image also manages the limited availability of high-quality multispectral imagery. To create clean composites, this study employs a spatiotemporal filter which includes outlier detection to correct for local or high-frequent non-bathymetric variations (Section 3.2.4).

3.2.1. Reference image

One single image is used as reference to create a correction method for large-scale non-bathymetric signals. To achieve this, the bathymetric signal is this image should be clear and unobstructed by for example clouds or high water levels. To minimise atmospheric effects, an readily corrected BOA image is selection from the Sen2Cor-processed collection of Sentinel-2 products. The image is manually selected to ensure high quality and a clear bathymetric signal. Multiple images satisfied the criteria of which one image is selected. The reference image used in this study is obtained by the Sentinel-2 mission at August 26, 2019 (Figure 3.7). This section briefly introduces the reference image and its applications in this study.

3.7).

Figure 3.7: Manually-selected clean and cloud-free reference image (26-08-2019) processed to BOA level by the Sen2Cor processor.

The reference image is central to the correction method for multispectral imagery which will be introduced in Section 3.2.3). This reference image correction method employed this BOA image as reference to correct all uncorrected TOA imagery for large-scale non-bathymetric variations. Furthermore, the reference image is used to identify images obstructed by large-scale effects which cannot be removed through corrections (Section 3.2.3).

Additionally, the cloud-free reference image is used to determine a relative threshold for the supplementary red band cloud mask (Section 3.2.2). For this purpose, to enable comparison with the uncorrected TOA imagery, the reference image is used at L1C TOA level. The reference image is only used at TOA level to determine this relative red band threshold. Throughout this thesis, the term *reference image* therefore refers to the image at L2A BOA level.

3.2.2. Detecting local clouds

Multispectral imagery is often obstructed by clouds; incoming radiation with wavelengths in the visible spectrum cannot penetrate and is reflected back to space (Section 2.2.4). The intensity of the reflected signal is relatively large, because clouds strongly reflect light across all three visible spectral bands. In the combined multispectral image, this results in white pixels which contain no information of the Earth's surface. Image-wide bathymetry estimation is hindered by these cloudy pixels and thus should these pixels be excluded from the depth estimation workflow.

Methods of cloud detection vary strongly depending on their application. Sentinel-2 imagery is embedded with a cloud mask by ESA (Section 3.1.2), for which the algorithm was developed for global cloud detection. On a local scale, not every cloudy pixel is identified (Figure 3.8).
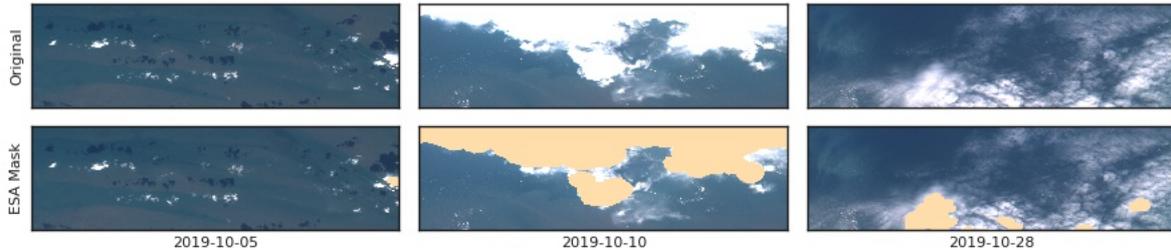
Figure 3.8: Three (partly) cloudy images of different dates, overlaid with the ESA cloud mask (yellow) which did not identify each individual cloudy pixel.

In this study, the absence of any other surface reflectors such as land and ice, allows for supplementary identification of clouds by using the characteristically high reflectances of clouds. Minimal light at the red wavelength is reflected from the ocean and thus high intensities for this spectral band are expected to be indicative of clouds (Section 2.2.5). Therefore, to supplement ESA's cloud-detection, a threshold is set on the intensity of the red band. The threshold is determined relative to a cloud-free image. For this, a L1C TOA version of the reference image is used to enable comparison of the reference image and the uncorrected TOA imagery. The threshold is defined as twice the 99th percentile of the red band of the L1C reference image. By setting this threshold, a secondary binary cloud mask is created which is suited for identifying clouds under the study-specific conditions. Images with a total cloud coverage exceeding 20 % are removed from the data set.

3.2.3. Correcting large-scale variations with a reference image

The recorded multispectral signal predominantly consists of atmospheric components and thus are atmospheric corrections crucial in the derivation of bathymetry, for which many tailored algorithms exist (Section 2.3.2). Each individual image is corrected to its present atmospheric and aquatic conditions, because the correction methods apply an image-based correction or require synchronous in situ observations. The error introduced by the insufficiency of these corrections, therefore differs at any epoch. As a result, images from different epochs display strong variations in intensity.

To illustrate this, images from three different dates are presented side-by-side (Figure 3.9). To visualise the effect of different atmospheric corrections, original uncorrected TOA imagery (Section 3.1.2) are shown (top row) in comparison to images corrected by the DOS method (middle row) and the Sen2Cor processor (bottom row) (Section 2.3.2). A column-wise comparison demonstrates the strong variation between images of different dates, which is a consequence of the time-varying errors that arise from the discrepancy between physical processes and their modelled corrections.
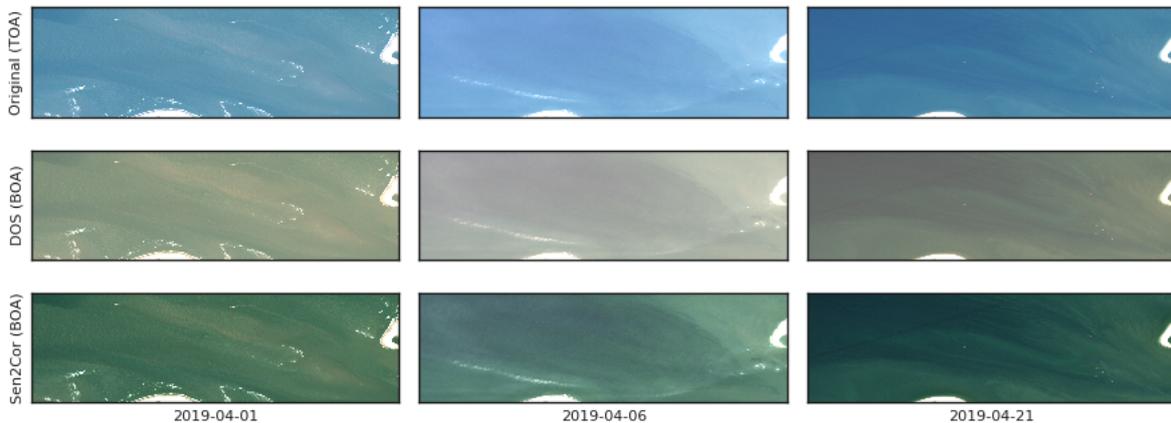
Figure 3.9: Visualisation of strong variability of multispectral imagery from different epochs. The top row displays 'original' TOA images which are not corrected for atmospheric effects. Variation between these three images demonstrates the heterogeneity in space and time of atmosphere and ocean. DOS-corrected imagery in the middle row and Sen2Cor-corrected imagery in the bottom row also display strong temporal variation despite correcting for atmospheric and aquatic effects. Strong variations between images from different epochs illustrate the incomparibility of multitemporal imagery for a time-independent relation between intensity and depth.

The temporal variation of the errors introduced by the insufficiency of the correction methods limits the comparability of multispectral imagery. Multitemporal studies on satellite derived bathymetry therefore often focus on minimising these errors by improving the (atmospheric) correction methods (e.g. Caballero and Stumpf, 2020a). This study presents an alternative method that instead aims to minimise the temporal variation of the errors from the insufficiency of the correction method.

To obtain time-independent correction errors, I developed an approach that corrects for non-bathymetric variations by utilising a single reference image. The idea behind this approach is that bathymetric changes are local, which causes the majority of the image to be stable over time. The local bathymetric signal can be isolated from the non-bathymetric constituents that cause large-scale variations because of their different spatial scales. Subtracting the non-bathymetric constituents from an image established a correction for large-scale variations. The approach of subtracting the non-bathymetric sources is similar to the DOS method (Section 2.3.2), which models these sources as a constant offset. In this study, the heterogeneity of the atmosphere and the ocean is taken into account by defining a spatially dependent function for the non-bathymetric sources.

The correction for large-scale non-bathymetric variations is implemented using the manually selected high-quality reference image (Section 3.2.1). The bathymetric signal is isolated by minimising the difference between an image and the reference image under the constrain of a model that estimates the non-bathymetric sources.

The correction is a function of the reference image and the non-bathymetric constituents. Since the atmospheric and aquatic conditions are heterogeneous in space, the modelled non-bathymetric sources

C_k are spatially dependent:

$$I_k^{cor} = I_k - C_k(x, y, I_k) \tag{3.1}$$

Where I_k^{cor} and I_k are the intensity of the recorded radiance in spectral band k for the corrected image and the image to be corrected respectively, and C_k is a spatially dependent model that estimates the non-bathymetric constituents.

The non-bathymetric perturbations are modelled as a linear approximation:

$$C_k = \alpha_k(x, y) \ I_k + \beta_k(x, y) \tag{3.2}$$

Where C_k is the model which corrects for non-bathymetric constituents in spectral band $k = 1, 2, 3$ for red, green and blue respectively, I_k is the intensity of the recorded radiance in spectral band k and α_k and β_k are the spatially dependent (x, y) coefficients defining the linear relationship of the non-bathymetric sources estimation model.

On a small scale, large-scale variations can be approached by a linearisation. Therefore, the spatial dependence of the coefficients α_k and β_k is given by a linear approximation:

$$\alpha_k(x, y) = \alpha_k^1 x + \alpha_k^2 y + \alpha_k^3 \quad \text{and} \quad \beta_k(x, y) = \beta_k^1 x + \beta_k^2 y + \beta_k^3 \tag{3.3}$$

Where α_k^1, α_k^2 and α_k^3 are the coefficients defining the linear spatial dependence of correction coefficient α_k of location (x, y). Idem for β_k.

The use of the reference image allows for empirically solving the coefficients α_k and β_k that define this non-bathymetric sources estimation model C_k. Least squares estimation is used to fit the model C_k to the difference between the reference image and the image that is to be corrected. This overdetermined linear system is given by:

$$I_k - I_k^{ref} = \alpha_k^1 x \ I_k + \alpha_k^2 y \ I_k + \alpha_k^3 \ I_k + \beta_k^1 x + \beta_k^2 y + \beta_k^3 \tag{3.4}$$

By using a reference, the temporal variation of the errors resulting from the insufficiency of the reference image correction method is minimised. As a result of this approach, multitemporal imagery is now directly comparable (Figure 3.10).
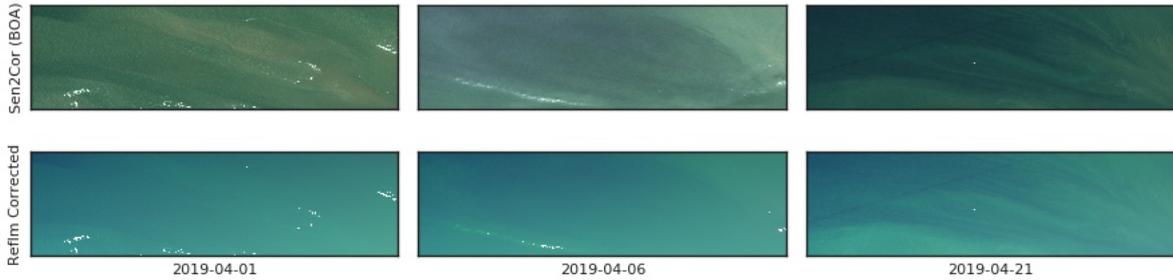
Figure 3.10: Visualisation of the reference image correction method (bottom row) for three example images of different dates in comparison to the Sen2Cor correction (top row). The reference image correction method reduced changes between images significantly, the imagery is now directly comparable.

Images obstructed by large-scale effects can be removed based on the prior assumption that bathymetric changes only occur locally. After the corrections, all imagery should resemble the reference image. However, large-scale effects such as haze or turbidity can alter images significantly. These strongly deviating images are identified by comparing the corrected imagery to the reference image. To quantify the resemblance of an image, the average deviation per pixel relative to the reference image is defined. Distinction of strongly deviating images is found to be best for comparison of the green band only. To remove these images with large-scale variations, a threshold is set on the relative deviation in the green spectral band at 8.5%. This threshold is determined by manually tuning the parameter and visually assessing its identification performance.

3.2.4. Correcting local and high-frequent variations with a composite

Cloud coverage affects multispectral imagery such that, despite the high revisit frequency of Sentinel-2, the majority of data is not usable for SDB (or any purpose). To deal with this problem, imagery is often manually handled to find a suitable image. Combining a stack of multitemporal images into one cloud-free composite is an alternative to manhandling data that also manages the problem of data sparsity due to cloud coverage.

In addition to handling cloudy images, compositing acts as a low-pass filter for other high-frequent variations. Single images are highly susceptible to high-frequent variations such as tidal water levels or the degree of turbidity. Generally, knowledge of these parameters is required to compare multitemporal images. High-frequent variations are averaged by combining multiple images into one composite. As a result, additional information of local conditions is no longer required to compare multitemporal imagery.

To create composites of consistent quality that contain minimal local non-bathymetric sources, a spatiotemporal filter is developed that detects and removes outliers. Small-scale non-representative pixels, such as cloud shadows, coast lines or abrupt changes, are removed.

A composite is created from a stack of N_i atmospherically-corrected images from epoch t_0 to t_{N_i} (Figure 3.11). Each pixel (i,j) of the composite is generated by evaluating the pixels in a spatiotemporal box, which has the size of the spatially adjacent pixels and the number of images in the stack (size $3 \times 3 \times N_i$ indexed by (i_f, j_f, t_f)). For each pixel (i,j) all pixels of the spatiotemporal filter are assessed; non-representative pixels (depicted in red) are removed. Identification of these non-representative pixels is done with the supplemented cloud mask (Section 3.2.2) and by iterative outlier detection.
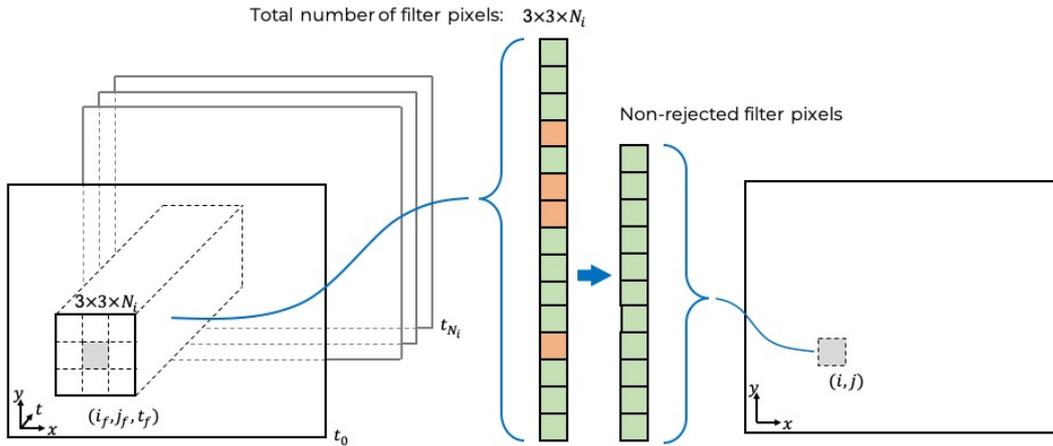
Figure 3.11: Graphic illustration of the spatiotemporal filter. Each pixel (i,j) in a composite image is generated based on the pixels from a spatiotemporal box with the size of the spatially adjacent pixels and the number of images in a stack (size $3 \times 3 \times N_i$, indexed by i_f, j_f, t_f). Cloudy and non-representative pixels (depicted in red) are removed from the selection of filter pixels. The remaining filter pixels (depicted in green) are used to obtain the value of the composite pixel.

The iterative outlier detection employed the generalised extreme studentized deviate (ESD) test from Rosner (1983). Outliers are iteratively removed from the collection of spatiotemporal filter pixels based on their relative deviation from the mean (Figure 3.12). For each composite pixel (i,j), a test statistic $\overline{R}$ is defined to score the combined relative deviation to the mean of the red, green and blue spectral bands. The largest outlier is identified as the pixel (i_f, j_f, t_f) that corresponds to the maximum test statistics. The outlier is removed if its test statistics exceeds the maximum relative deviation while maintaining a minimum number of filter pixels.
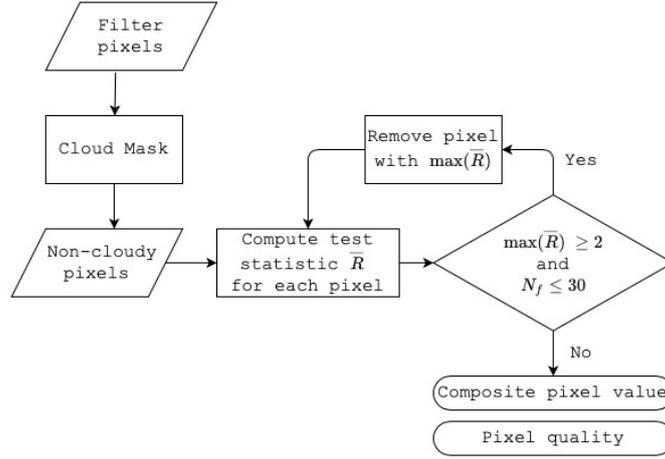
Figure 3.12: Graphic illustration of the iterative outlier detection. For each filter pixel (i_f, j_f, t_f) a test statistic $\overline{R}$ is defined to express its deviation from the mean. Outliers are removed if maximum deviation is exceeded while maintaining a minimum number of filter pixels.

The number of pixels in the filter N_f is defined as the contributing pixels which are not rejected by the cloud mask or the outlier detection. It is obtained by applying a binary operator Π, to all pixels in the spatiotemporal box and summing over its three dimensions. Operator Π represents the combined outcome of the rejection criteria for a pixel (i_f, j_f, t_f), the pixel is included when $\Pi = 1$ and excluded when $\Pi = 0$.

$$N_f = \Sigma\Sigma\Sigma \; \Pi(i_f, j_f, t_f) \tag{3.5}$$

For each pixel (i, j) the mean of each spectral band $\langle I_k \rangle$ is determined over the spatiotemporal filter. Furthermore, for each pixel (i, j) the variance-covariance matrix of the spectral bands $cov(I_k, I_\zeta)$ is defined. From the variance-covariance matrix the standard deviations for each spectral band σ_{I_k} follow.

$$\langle I_k \rangle \big|_{i,j} = \sum_{i_f=i-1}^{i+1} \sum_{j_f=j-1}^{j+1} \sum_{t_f=1}^{N_i} \frac{I_k(i_f, j_f, t_f)}{N_f} \; \Pi(i_f, j_f, t_f) \tag{3.6}$$

$$cov(I_k, I_\zeta) \big|_{i,j} = \frac{\sum_{i_f=i-1}^{i+1} \sum_{j_f=j-1}^{j+1} \sum_{t_f=1}^{N_i} \left(I_k(i_f, j_f, t_f) - \langle I_k \rangle \big|_{i,j} \right) \left(I_\zeta(i_f, j_f, t_f) - \langle I_\zeta \rangle \big|_{i,j} \right)}{N_f} \tag{3.7}$$

$$\sigma_{I_k} \big|_{i,j} = \sqrt{\frac{\sum_{i_f=i-1}^{i+1} \sum_{j_f=j-1}^{j+1} \sum_{t_f=1}^{N_i} \left(I_k(i_f, j_f, t_f) - \langle I_k \rangle \big|_{i,j} \right)^2}{N_f}} \tag{3.8}$$

Where $\langle I_k \rangle$, $cov(I_k, I_\zeta)$ and σ_{I_k} are the mean, covariance and standard deviation of the intensity for spectral bands k, ζ evaluated for composite pixel (i, j).

For each pixel (i_f, j_f, t_f) of the N_f filter pixels, a relative deviation is expressed in each spectral band k by a test statistic R_k. The test statistic is defined as the distance to the mean $\langle I_k \rangle$ normalised by the standard deviation σ_{I_k}:

$$R_k(i_f, j_f, t_f) = \frac{\left| I_k(i_f, j_f, t_f) - \langle I_k \rangle \big|_{i,j} \right|}{\sigma_{I_k} \big|_{i,j}} \tag{3.9}$$

Subsequently, each filter pixel is scored as the average of the test statistics of the three spectral bands:

$$\overline{R} = \frac{1}{3} \sum_{k=1}^{3} R_k \tag{3.10}$$

Where $\overline{R}$ is the average test statistic for filter pixel (i_f, j_f, t_f), which represents the relative deviation combined for the spectral bands red, green and blue ($k = 1, 2, 3$).

A relative threshold is set for the combined deviation $\overline{R}$ of filter pixel (i_f, j_f, t_f), which is set at two times the standard deviation $2\sigma_{I_k}$. Because the test statistic is normalised by the standard deviation, the threshold is correspondingly normalised, which sets the threshold at 2.

The average test statistic $\overline{R}$ is calculated for each pixel (i_f, j_f, t_f) of the N_f filter pixels. If the maximum value exceeded the threshold, the corresponding pixel is removed. The process is iterated until no maximum test statistic score exceeded the threshold or until the arbitrary minimum number of pixels ($N_{min} = 30$) is reached:

$$\text{if } \max\left(\overline{R}(i_f, j_f, t_f)\right) > 2 : \text{remove filter pixel } (i_f, j_f, t_f) \text{ and recompute } \overline{R} \tag{3.11}$$

A composite image is created by applying the spatiotemporal filter to each pixel. For each pixel (i, j), the mean of the intensity I_k in spectral band k of the remaining N_f pixels in the filter, is assigned as the pixel value of the intensity of the composite I_k^C in that spectral band:

$$I_k^C(i, j) = \langle I_k \rangle\big|_{i,j} \tag{3.12}$$

A quality map is constructed to support the composite image. The standard deviation of the intensity I_k in spectral band k of the remaining N_f pixels in the filter is used for each pixel (i, j) as a quality measure of the pixel after filtering:

$$\sigma_{I_k}^C(i, j) = \sigma_{I_k}\big|_{i,j} \tag{3.13}$$

Where $\sigma_{I_k}^C$ is the standard deviation of the composite value I_k^C of the intensity I_k in spectral band k

Composite images manage limited data availability and high-frequent non-bathymetric variations, albeit at the cost of smoothening the signal of bathymetric changes. The spatial and temporal scales of the bathymetric structures of interest should be taken into consideration when selecting the duration of a composite image. For this study, a composite duration of six months is selected to ensure observable dynamics within a yearly timescale (Section 2.1) while handling the scarcity of Sentinel-2 imagery. To model temporal bathymetry changes, composites are generated between 1 January 2016 and 31 December 2019, by using a moving-average filter. Partially overlapping composite time-frames are created with their start dates one month apart, which resulted in forty-three time-windows.

3.3 Developing a bathymetry estimation model

A bathymetry estimation model describes the relation between water depth and the recorded radiance of multispectral imagery. In situ depth observations are required to train the estimation model. In this study, these depth observations were additionally used to constrain the depth estimation model to the penetration depth of visible light into the water column (Section 2.2.3). This section presents a description of the depth estimation model (Section 3.3.1) and an explanation of how this model was trained (Section 3.3.2).

3.3.1. Model description

To define the relation between the corrected and composited imagery and the in situ observations from the vaklodingen, a log-linear relationship was used which is based on the model from Lyzenga (1978) (Section 2.3.1). This log-linear approach is well-suited to the homogeneous habitat of the study site in the Wadden Sea (Section 3.1.1). Due to the absorption spectrum of water, only visible light can be used for the estimation of bathymetry (Section 2.2.5). For this study, water depth is therefore defined as a function of the intensity of the composite image for the three visible spectral bands:

$$z = b + \sum_{k=1}^{3} a_k \log I_k^C \tag{3.14}$$

Where z is the water depth, I_k^C the intensity of composite image for the $k = 1, 2, 3$ spectral bands, and a_k and b the coefficients describing the linear relationship.

The log-linear relation is constrained to the penetration depth of visible light into the water column (Section 2.2.3). The turbid waters of the Wadden Sea strongly limit the penetration of light and subsequently the ability to derive bathymetry for deep depths from an optical image. The local penetration depth was assessed by comparing recorded radiances of the reference image to observed water depths from the corresponding 2019 vaklodingen. Depths observed deeper than this penetration depth, cannot be estimated from multispectral imagery with the log-linear model. To implement this constrain, a threshold was set on the in situ observations that were used to derive the estimation model.

3.3.2. Model training

The coefficients a_k and b which define the bathymetry estimation model (Equation 3.14) are obtained empirically, by using in situ observations and the corresponding composite image to train the model. The vaklodingen data is interpolated onto the 10 m grid of the Sentinel-2 multispectral imagery. Both data sets are randomly partitioned in subsets for training and testing (80/20). The observed penetration depth (Section 3.3.1) is set as a threshold on the vaklodingen data to constrain the depth estimation model (Equation 3.14). Regression analysis is performed on the constrained training subsets of the composite image and the interpolated vaklodingen by using least squares estimation (Section 2.3.4).

Synchronous in situ observations and multispectral imagery are required to train the depth estimation model. Two sets of vaklodingen data are obtained since the launch of the Sentinel-2 mission, respectively the 2016 and 2019 surveys (Section 3.1.2). Two synchronous six-month composite images (Section 3.2.4) are therefore suited for training and validating the depth estimation model (Equation 3.14); the composites from July 2016 to January 2017 and from July 2019 to January 2020. Hereafter, the two resulting estimation models are referred to as the 2016 and 2019 bathymetry estimation models respectively (BEM 2016 and BEM 2019). To assess the affect of averaging images in a composite, an additional bathymetry estimation model is trained on the reference image.

3.4 Estimating bathymetry

The trained bathymetry estimation models (BEM 2016 and BEM 2019) are used to estimate depths for times of absent in situ data. As a result of the reference image correction method, multispectral imagery of different epochs is directly comparable, which allows for using a bathymetry estimation model that is constant over time. Temporal extrapolation is performed by applying the readily-trained estimation model to non-synchronous imagery. This section presents how composite images are used to estimate bathymetry and map temporal changes (Section 3.4.1), provides a quality assessment of the estimated depths (Section 3.4.2) and shows how depth estimates are validated (Section 3.4.3).

3.4.1. Estimating bathymetry and mapping temporal changes

Bathymetry is estimated with the two bathymetry estimation models which are trained with the vaklodingen (Section 3.3.2). First, the performance of the estimation models for the Wadden Sea case study is assessed. The bathymetry estimation models are applied to the remaining testing subset of the composite that was used to define the models. Applying the estimation models to their training composite sets a benchmark for estimation performance which can be used to evaluate the influence of temporal extrapolation. The estimated depths are visualised on two-dimensional maps and along one-dimensional transects.

The effect of temporal extrapolation on estimation performance is cross-validated with non-synchronous vaklodingen data. An extrapolation of three-years is performed by applying BEM 2016 to the 2019 composite and BEM 2019 to the 2016 composite. For each estimation model, the performance is compared to its the benchmark performance. The extrapolated depth estimates are correspondingly visualised on two-dimensional bathymetry maps and along one-dimensional transects.

Temporal bathymetric changes over a period of four years are estimated by applying the two bathymetry estimation models to consecutive partially-overlapping composites. For six locations the temporal development is visualised with timeseries. To provide context to the temporal changes, the development of a bathymetric sandbank structure is estimated. Contourlines are used to define the outline of this structure. The development of the structure is also estimated over a four-year period with the use of four composite from consecutive years.

3.4.2. Quality of estimated depths

To obtain an estimate of the depth estimation quality, the covariance matrix for each pixel (i, j) of the composite is propagated (Equation 3.7). The bathymetry estimation model is non-linear, therefore a linearisation is used for the error propagation (Verhagen and Teunissen, 2017). The resulting depth variance σ_z^2 for pixel (i, j) is given by:

$$\sigma_z^2 \Big|_{i,j} \approx \mathbf{J} \mathbf{Q}_{II} \mathbf{J}^\top \tag{3.15}$$

Where $\mathbf{J}$ is the Jacobian matrix of the bathymetry estimation model:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial z}{\partial I_r} & \frac{\partial z}{\partial I_g} & \frac{\partial z}{\partial I_b} \end{bmatrix} = \begin{bmatrix} \frac{a_r}{I_r} & \frac{a_g}{I_g} & \frac{a_b}{I_b} \end{bmatrix} \tag{3.16}$$

And where $\mathbf{Q}_{II}$ is the variance-covariance matrix:

$$\mathbf{Q}_{II} \Big|_{i,j} = \begin{bmatrix} \sigma_{I_r}^2 & \mathrm{cov}(I_g, I_r) & \mathrm{cov}(I_b, I_r) \\ \mathrm{cov}(I_r, I_g) & \sigma_{I_g}^2 & \mathrm{cov}(I_b, I_g) \\ \mathrm{cov}(I_r, I_b) & \mathrm{cov}(I_g, I_b) & \sigma_{I_b}^2 \end{bmatrix}_{i,j} \tag{3.17}$$

3.4.3. Validation of estimation

Bathymetry estimates are compared to in situ observations to assess estimation performance. To set a benchmark for the performance of the estimation models for the Wadden Sea case study, the bathymetry estimation models are applied to the remaining testing subset of the composite used for training. Next, the influence of temporal extrapolation is cross-validated by applying the estimation models to a non-synchronous composite. By combining the two estimation models and the two six-month training composites, four scenarios are defined (Table 3.1).

Benchmark (Synchronous)	Extrapolation (Non-synchronous)
BEM 2016 + Composite 2016	BEM 2016 + Composite 2019
BEM 2019 + Composite 2019	BEM 2019 + Composite 2016

Table 3.1: Setting a benchmark of estimation performance by applying the bathymetry estimation models to the remainder of the training subset.

Estimated depths are qualitatively validated by visual comparison with the in situ observations. The widely-used root mean square error (RMSE) is used to quantify the differences between observed and estimated values. The RMSE metric is sensitive to outliers, it should therefore be evaluated when used to quantify estimation performance.

4

Results & Discussion

As discussed in Chapter 3, a three-strep workflow to map temporal bathymetric changes is tested on a Wadden Sea study site. Here, I present the results of processing the multispectral imagery to isolate the bathymetric signal (Section 4.1), validating the bathymetry estimation model (Section 4.2) and estimating bathymetry for times of absent in situ data (Section 4.3).

4.1 Isolating the bathymetric signal in multispectral imagery

To obtain clean, high-quality multispectral imagery suited to temporal analysis, an approach to process multispectral imagery is presented that is specifically tailored to mapping temporal bathymetric changes. (Section 3.2). This includes the removal of clouds, a novel reference image correction method to correct for large-scale non-bathymetric signals and a modified approach for the creation of composite images to correct for local and high-frequent non-bathymetric variations.

4.1.1. Detecting and removing clouds

Pixels obstructed by clouds in Sentinel-2 imagery are classified by the ESA and identified with a binary cloud mask (European Space Agency, 2015) developed for global applications. To locally supplement cloud detection for the Wadden Sea case study, an additional cloud mask is implemented which is based on the reflectance recorded in the red band. Cloudy pixels are identified by the characteristically high reflectance in the red band, for which the threshold is set as twice the 99th percentile of the red band reflectance of the TOA reference image. In this section, the performance of the supplementary red band mask is assessed.

Three Sentinel-2 images with varying cloud coverage are selected to illustrate differences between the embedded ESA cloud mask and the supplementary red bank cloud mask (Figure 4.1, columns). To evaluate the performance of the cloud masks, unmasked original TOA imagery is shown (top row) overlaid with the ESA cloud mask (middle row) and the supplementary red band mask (bottom row). For each epoch, the cloud coverage is computed using the ESA mask and using the supplemented ESA and red band mask. By supplementing the ESA cloud mask, the percentage of cloudy pixels for the first image increased from zero to 3 %, for the second image from 39 % to 59 % and for the third image from 5 % to 41 %.
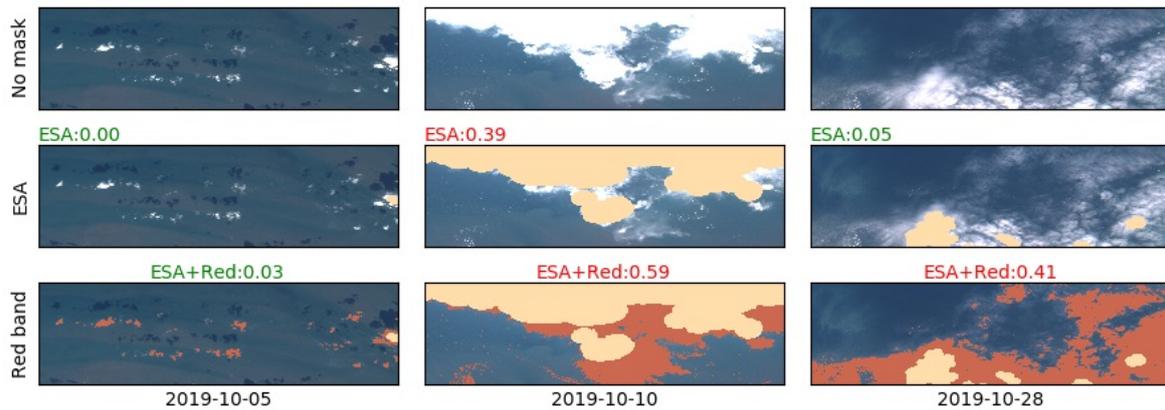
Figure 4.1: Example of multispectral imagery with varying degrees of cloud coverage from three different epochs (columns) to illustrate cloud mask performance. Unmasked TOA imagery (top row) is overlaid with the ESA cloud mask in yellow (middle row) and the supplementary red band cloud mask in orange (bottom row). A maximum cloud coverage of 20 % was accepted. Supplementing the ESA cloud mask with the red band cloud mask appropriately removes the third image in addition to the second image (displayed in red).

The site-specific red band cloud mask successfully supplemented cloud detection by the ESA cloud mask. For all three epochs, the red band mask correctly identified additional cloudy pixels. Cloud detection is complex because of the heterogeneous effects of clouds on multispectral imagery and is especially difficult on a global scale. Cloud coverage can be very local such as on the first day, which leaves the majority of pixels in the image unaffected. Often, the effects of clouds are image-wide and few pixels contain bathymetric information. The heterogeneity of clouds and their image-wide effects is visible for the images of the second en third day; dense clouds are more easily detected than thinner layers of clouds. Upon inspection, the remaining unmasked parts of these images show no structure at all or display a pattern of clouds. The unmasked pixels of cloudy images with a significant percentage of cloud coverage appear to contain no reflections from the Earth's surface, which demonstrates the difficulty of cloud detection. To avoid including these undetected non-representative pixels, images with a cloud coverage computed to be more than 20 % are removed from the four-year collection of Sentinel-2 data. Supplementing the ESA cloud mask with the red band cloud mask appropriately removed the third image in addition to the second image.

4.1.2. Correcting large-scale variations

The reference image correction method is applied to the multispectral imagery to correct for large-scale non-bathymetric signals (Section 3.2.3). The local bathymetric bathymetric signal is isolated from large-scale atmospheric and aquatic variations by using a reference image. Images obstructed by large-scale non-bathymetric effects are identified by their relative deviation from the reference image, which is quantified as the average deviation per pixel in the green spectral band. This section evaluates imagery corrected with the reference image correction method and illustrates how images obstructed by large-scale non-bathymetric effects are removed.

Three images with varying degrees of clarity are selected to demonstrate the effects of the reference image correction method and to illustrate large-scale obstructions (Figure 4.2). The intensity of the bathymetric signal increases from the first to the third image. The relative deviation in the green band with respect to the reference image is displayed above each image.

Figure 4.2: Visualisation of directly comparable multispectral imagery which was corrected for large-scale variations with the reference image correction method. The bathymetric signal is increasingly visible for these three epochs. The first image is strongly obstructed by some large-scale effects. To remove obstructed images, large-scale obstructions were identified by quantifying the resemblance to the clean reference image as the average relative deviation in the green band. The threshold for these residuals in the green band (RG) was manually set to 8.5 % which appropriately removes the first image (displayed in red) but includes the second and third image (displayed in green).

Multispectral imagery was directly comparable after applying the reference image correction method. The corrected images display uniformity in colour, clarity and brightness, in contrast to the uncorrected images, the DOS-corrected images and the Sen2Cor corrected images (Figure 3.9). The clear bathymetric signal in the third image demonstrates that large-scale non-bathymetric variations are successfully removed, while the local bathymetry is conserved. These preserved local bathymetric structures confirm the assumption that bathymetric changes are local and that they can be distinguished from large-scale variations (section 3.2.3).

Despite the corrections and the preservation of local bathymetric structures, the three images display varying degrees of image quality. These large-scale distortions can be due to atmospheric or aquatic conditions such as haze or turbidity. Image quality is resembled by the average deviation per pixel in the green band relative to the reference image. A larger deviation from the reference image was indicative of images strongly obstructed by large-scale non-bathymetric effects. The first image displays little bathymetric information and deviates the most from the reference image (8.9 %). The bathymetric signal is best visible in the third image, which correspondingly deviates the least from the reference image (7.3 %). The threshold on the relative deviation is set to 8.5 % manually, which appropriately included the second and third image but excluded the first image from the data set. Setting this threshold however potentially removes images that contain bathymetric information. Selecting images by setting this threshold was therefore a trade-off between including low-quality images and excluding high-quality images.

4.1.3. Compositing

Multiple images are combined into a composite to correct for local and high-frequent non-bathymetric effects. Compositing also handled data scarcity of high-quality images. This section demonstrates the need for composite images.

From 01 July 2016 to 01 January 2017 fourteen images are initially selected that had a cloud coverage lower than fifty percent (Figure 4.3). The images are corrected with the reference image correction method and overlaid with the ESA cloud mask (depicted in yellow) and the supplementary red band cloud mask (depicted in orange). Cloud coverage is calculated from the ESA mask and from the supplemented cloud mask, both values are displayed above each image. The red band mask extended the ESA mask, with cloud detection predominantly centred around pixels already identified as cloudy. Furthermore, images obstructed by large-scale non-bathymetric effects are identified by expressing the deviation relative to the reference image. This relative deviation in the green band is also displayed above the individual images. Thresholds are set to a maximum cloud coverage of 20 % and maximum average relative deviation of 8.5 %. Values exceeding the threshold are depicted in red, accepted values in green. To create a high-quality composite image, individual images should comply with both thresholds.
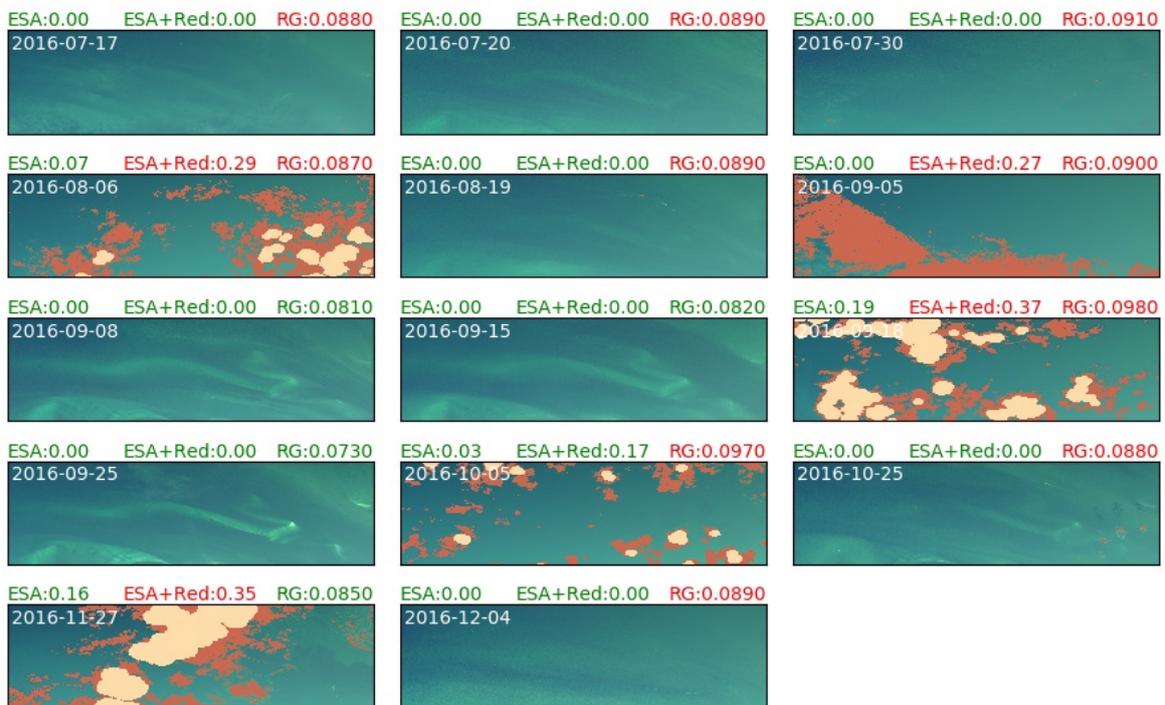
Figure 4.3: Visualisation of all corrected images from 01 July 2016 to 01 January 2017 with a total cloud coverage less than fifty percent. Images are overlaid with ESA cloud mask (yellow) and Red Band cloud mask (orange) and display three percentages: (ESA) cloud coverage based on ESA cloud mask, (ESA+Red) cloud coverage based on ESA and Red band cloud masks and (RG) deviation relative to reference image expressed as Residual of the Green band. Values are displayed in red if they exceed corresponding threshold (ESA: 0.2, ESA+Red: 0.2, RG: 0.085). Only images with all percentages shown in green were considered satisfactory for compositing.

The selected images demonstrate the performance of the supplemented cloud mask and the reference image correction method, but also illustrate the limited number of high-quality images to combine into a composite image. The supplementary red band cloud mask sufficiently detected additional cloudy pixels. The maximum cloud coverage based on this combined cloud mask aptly identified three more images obstructed such that they contain no bathymetric information. Out of the fourteen images, five images exceeded the accepted cloud coverage. Setting the threshold for relative deviation in the green band at 8.5 %, appropriately identified another six images obstructed by large-scale effects. From the three remaining high-quality images (2016-09-08, 2016-09-15, 2016-09-25) a composite image is created.

To demonstrate the limited availability of high-quality data, Figure 4.4 provides an overview of the high-quality images among all Sentinel-2 imagery available from 1 January 2016 to 1 January 2020. Data availability increased strongly with some delay after the launch of the Sentinel-2B satellite in 2017. Seasonal trends are somewhat visible, between April and October more high-quality images are available than from October to April, which corresponds to Dutch summer and winter respectively.
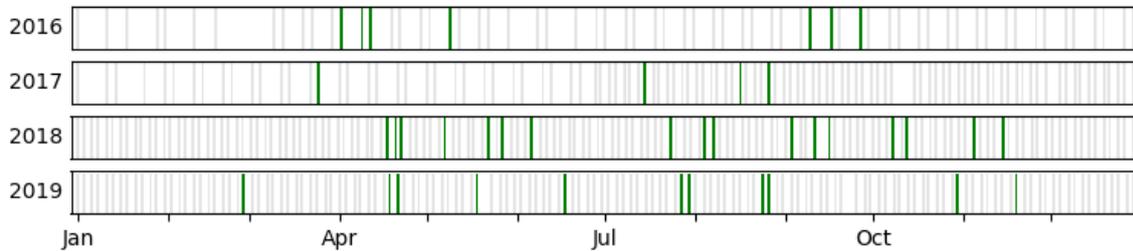
Figure 4.4: Visualisation of Sentinel-2 imagery from 2016 to 2019 which is usable for the derivation of bathymetry (indicated in green) based on a maximum total cloud coverage of 20 % and a maximum deviation in the green band of 8.5 %.

Figure 4.4 accentuates the limited availability of high-quality images suited for the estimation of bathymetry. Despite the increase in data due to the launch of Sentinel-2B, the number of high-quality images is limited. Although partially-overlapping composite images with a duration of six-month provided a solution to the data sparsity, for some time periods no high-quality imagery is available (for example from 1 October 2017 to 1 April 2018). These data show that data scarcity can be managed by increasing the duration of the composite. From a perspective that aims to maximise the number of images in a composite, the suitability of six-month composites is argued by the limited availability of high-quality data. Composite duration is however a trade-off between data scarcity and the observability of morphological structures (Section 2.1).

The six-month composite from 1 July 2016 to 1 January 2017 is created from three individual images (Figure 4.3). The composite is created by the spatiotemporal filter which included pixel-based outlier detection (section 3.2.4) (Figure 4.5 (top)). Bathymetric structures from the individual images are preserved while creating a smooth image. The composite image from 1 July 2019 to 1 January 2020 is of lower quality; the bathymetric structures are just faintly visible (Figure 4.5 (bottom)). The RGB values of both composites are scaled from 0 to 700 to enable fair visual comparison.
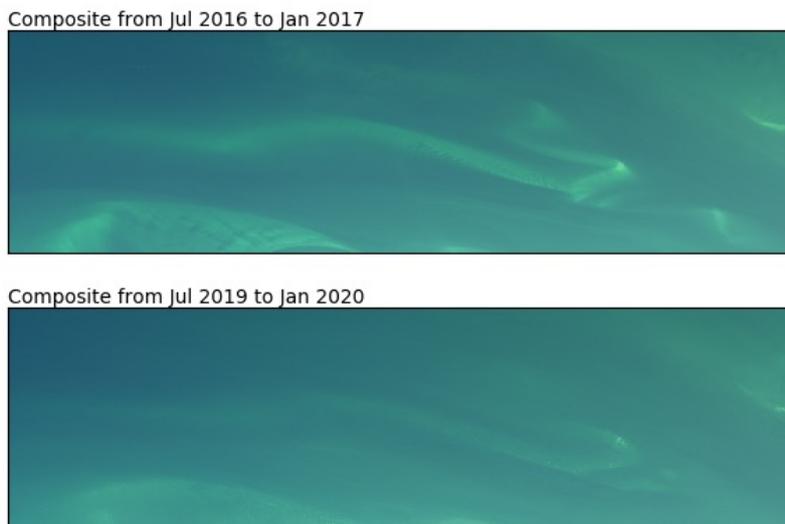
Figure 4.5: Two six-month composite images (2016 and 2019) created with the spatiotemporal filter. The bathymetric signal is significantly stronger in the 2016 composite (top image); the quality of the 2019 composite (bottom image) is lower.

4.2 Bathymetry estimation model

The relationship between water depth and intensity of the multispectral image is described by the bathymetry estimation model. In situ depth observations are used to obtain this relation and to constrain the estimation model to the local penetration depth of visible light into the water column. First, this section discusses the applicability of the log-linear model for the case study (Section 4.2.1). Next, it presents the estimation benchmark results (Section 4.3) and highlights limitations on those depth estimates (Section 4.2.3). This section concludes with bathymetry estimates of the benchmarks scenarios along two-dimensional transects (Section 4.2.4).

4.2.1. Applicability of the log-linear model

To validate the log-linear bathymetry model (Section 3.3.1), the reference image is compared to the corresponding 2019 vaklodingen. The log-transformed intensity of each spectral band is evaluated against the in situ depths (Figure 4.6). The recorded intensity in the red band is somewhat lower than the intensities in the green and blue bands. The three spectral bands display similar trends nonetheless; each band shows a linear relationship for depths up to approximately -5 m. Below this penetration depth, the relation between depth and intensity is ambiguous and a near-vertical trend is visible.
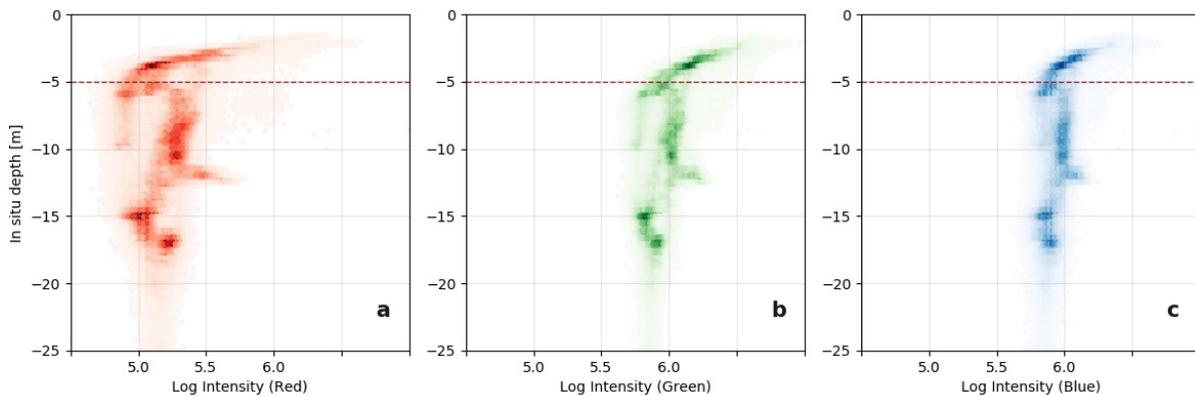
Figure 4.6: Comparison of log-transformed intensities of the three visible bands from the reference image (x-axis) and observed water depths from the 2019 vaklodingen (y-axis). A clear linear relationship is visible for depths up to approximately 5 m, below this depth that relationship is ambiguous.

The observed penetration depth of -5 m is determined by local aquatic conditions which control the attenuation of light in the water column (Section 2.2.5). Due to high turbidity in the Wadden Sea waters, little light actually reflects on the seafloor. Little reflected radiance is therefore recorded for all depths deeper than the penetration depth, and all deep water pixels are approximately equally dark. When comparing these low intensities of recorded radiance to observed water depths, a near-vertical trend develops, where all deeper depth correspond to approximately the same intensity. The log-linear relationship only holds for depths above the penetration depth, no clear relation can be defined for depths under that threshold and the intensities of the multispectral image. The log-linear model is therefore only suited to the first five meters of the water column; for this study area deeper depths cannot be estimated. Consequently, only pixels corresponding to observed depths shallower than the penetration depth are used to train the regression model.

4.2.2. From composites to bathymetry

To set a benchmark for the performance the bathymetry estimation model, the BEM 2016 and BEM 2019 models (Section 3.3.2) are applied to the remaining testing subset of their training composite. BEM 2016 is applied to the 2016 composite (Figure 4.5) and the resulting depth estimates are compared to the 2016 vaklodingen (Figure 4.7. Similarly, BEM 2019 is applied to the 2019 composite (Figure 4.5) and resulting estimates are compared to the depth observations from the 2019 vaklodingen (Figure 4.8. The estimation models are constrained to the penetration depth of 5 m (Section 4.2.1) by setting a threshold on the observations used for training (Section 3.3.2).

Estimated depths (Figures 4.7a and 4.8a) display shallow bathymetric structures similar to the observed

vaklodingen (Figures 4.7c and 4.8c). Deep areas are estimated approximately equally shallow, around 5 m. The shallow bathymetric structures are clear and well-defined in the 2016 estimates (Figure 4.7a), the 2019 estimates (Figure 4.8a) display in comparison less depth contrast and therefore less clear and defined structures. A quality map is created from the standard deviation the depth estimate of each pixel (Section 3.4.2) (Figures 4.7b and 4.8b). For the 2016 scenario (Figure 4.7b) the average standard deviation appears to be larger than for 2019 (Figure 4.8b). However, in 2016 deep area estimates come with the largest uncertainties, whereas the 2019 case shows largest uncertainties in the estimation of the bathymetric structures of interest. To provide insight on the order of estimation errors, estimated depths are compared to the observed depths (Figures 4.7d and 4.8d). Both estimation models incorrectly estimate the deep areas to be greatly shallower. By applying BEM 2016 to the 2016 composite, shallow bathymetric are estimated slightly too deep with locally larger errors (Figure 4.7d). Based on the 2019 composite, BEM 2019 estimated the shallow structures generally somewhat deeper (Figure 4.8d).
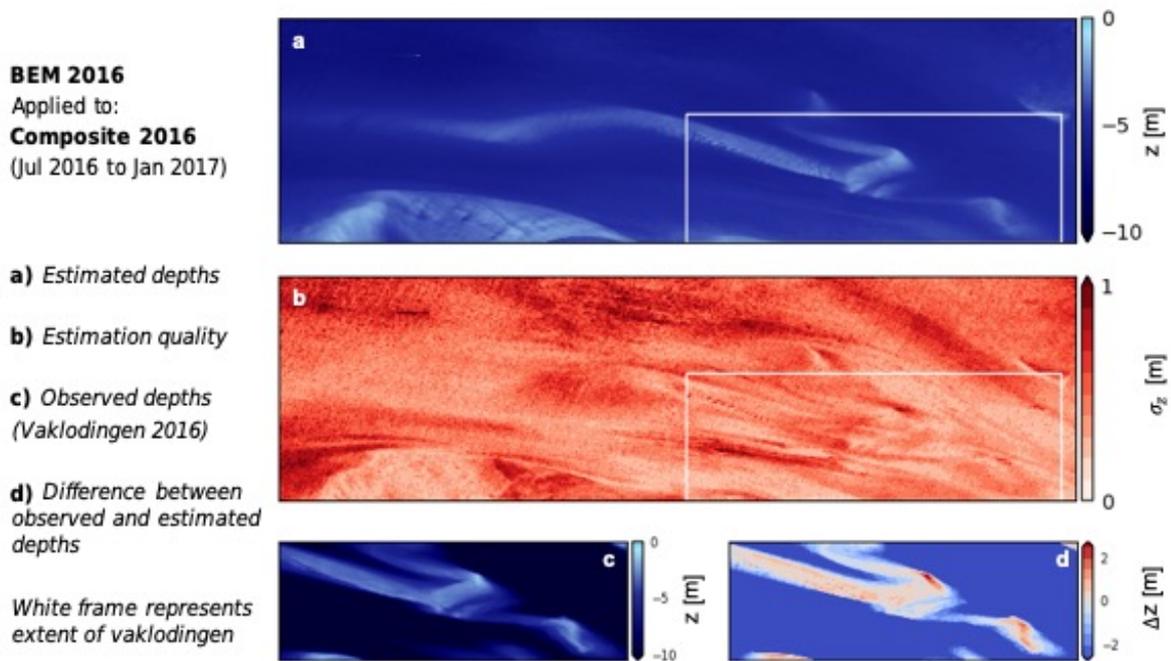
Figure 4.7: Bathymetry estimates and estimation quality from BEM 2016 with 2016 composite, compared with 2016 vaklodingen.
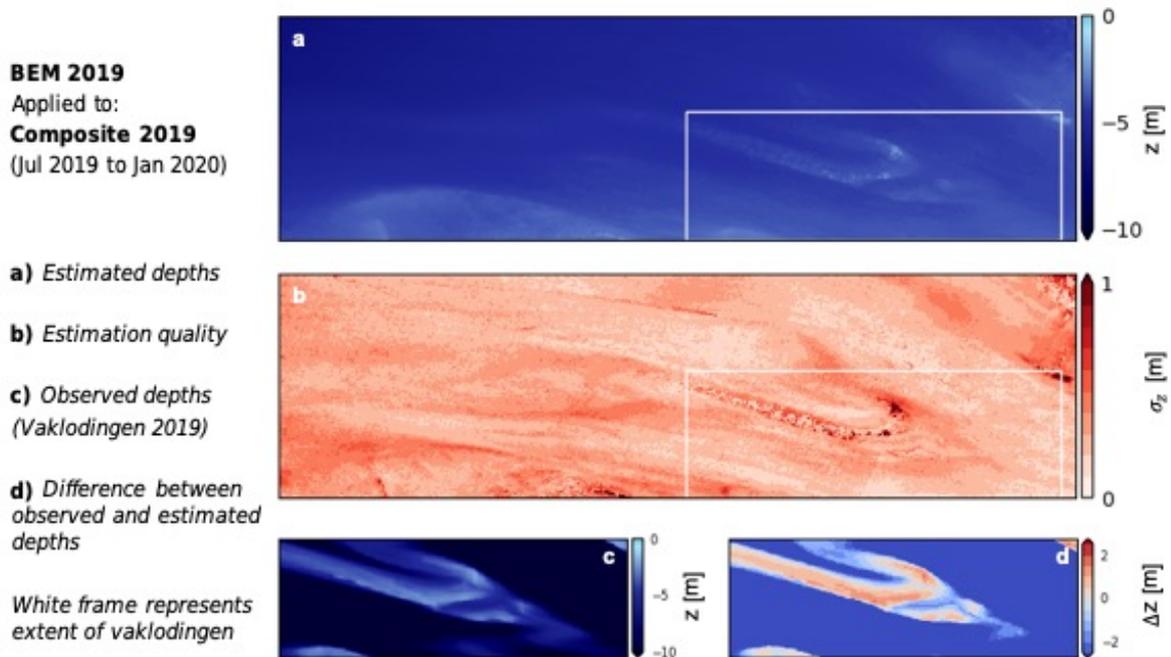
Figure 4.8: Bathymetry estimates and estimation quality from BEM 2019 with 2019 composite, compared with 2019 vaklodingen.

The bathymetry estimation models successfully estimated shallow bathymetric structures up to approx-imately 5 m. Deep depths are however estimated significantly too shallow, which results in large estimation errors of over 20 m. Estimation uncertainties are generally largest for deep areas. These results correspond to the findings that depths deeper than the penetration depth cannot be estimated by the log-linear model (Section 4.2.1). Although uncertainties appear larger for the BEM 2016 in combination with the 2016 composite, the average standard deviation is approximately equal to the 2019 scenario when only taking into account depths observed shallower than 5 m (≈ 0.25 m). By combining these results of bathymetries estimated for times with synchronous in situ data available, a benchmark is set for the performance of the bathymetry estimation models.

4.2.3. Limitations on the depth estimates

To improve the performance of the bathymetry estimation model, the limiting effect of the penetration depth on predictable depths is assessed. To illustrate this, the reference image is used to estimated depths which are compared to observed depths (Figure 4.6). The bathymetry estimation model is trained on the training subsets of the reference image and the 2019 vaklodingen for which (based on results in Figure 4.6) only pixels corresponding to to depths up to 5 m are selected for training. For shallow depths, the data shows a linear trend corresponding to the 1:1 line, which represents correct estimation (See inner figure). In other words, depths up to five metres are estimated correctly by using the log-linear relation between depth and intensity and setting the penetration threshold at 5 m. However, deeper depths are estimated incorrectly, Figure 4.9 displays estimation errors as large as 20 m.
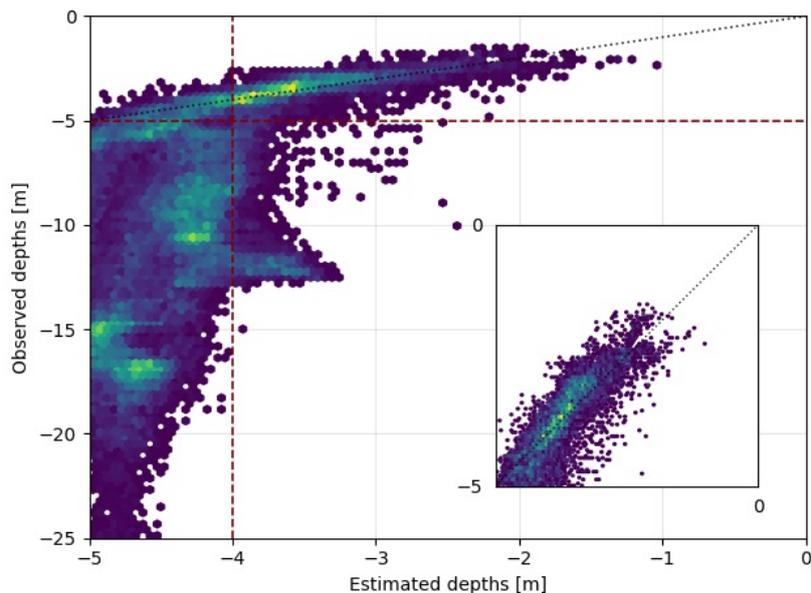
Figure 4.9: Comparison of estimated depths (x-axis) and observed depths (y-axis) from the reference image (26-08-2019) and the 2019 vaklodingen. The dark, dotted line represents the 1:1 line.

Figure 4.9 demonstrates the inability to estimate deep depths. The estimation model is constrained to the penetration depth because the log-linear relation does not hold for deeper depths. Depths below the penetration threshold reflect equally low radiation and are incorrectly estimated to be shallower by the constrained model. This bias towards shallow depths strongly affects the RMSE which is used to express the estimation performance.

If in situ observations are available, the penetration depth can be used to eliminate all depths observed deeper from the data set in order to calculate the RMSE. By using this threshold on observed depths, this RMSE provided an estimate of the minimal estimation error under the local conditions. For depths that are observed shallower than 5 m, the RMSE computed on the reference image was 0.40 m, which corresponds to an estimation error of approximately 8 %.

In the absence of in situ data, it is not possible to eliminate those deep depths that give rise to large estimation errors. The RMSE cannot be determined by setting a threshold on the observations since knowledge is only available on estimated depths, rather than observed depths. To illustrate this principle, the RMSE is calculated with the penetration depth of 5 m as a threshold on the estimated depths. For depths estimated shallower than 5 m, the RMSE computed on the reference image was 7.22 m; an estimation error even larger than the range of estimated depths.

A closer look at Figure 4.9 illustrates the problem. No clear relationship is observable between observed depths and depths that are estimated shallower than 4 m (indicated by the vertical dashed red line). Deep depths are projected by the bathymetry estimation model as depths between 4 and 5 m. This suggests that predictable depths are not limited to the penetration depth but that reliable estimates are even more restricted.

To evaluate to what threshold estimated depths are reliable, a comparison is made of the RMSE based on different cutoffs on the estimated depths (Figure 4.10). The two estimation models, BEM 2016 and BEM 2019, are used to estimate depths on the remaining testing subset of their training composite (Section 3.3.2). By setting the penetration depth (5m) as threshold on the estimated depths, large estimation errors are obtained similar to the the RMSE prior calculated for the reference image. For depths estimated shallower than 5 m, the RMSE computed with BEM 2016 was 5.14 m and the RMSE computed with BEM 2019 was 7.78 m. Although reducing the threshold on the estimated depths strongly decreased the RMSE for both scenarios, estimation especially improved for the BEM 2016 scenario. By setting the cutoff on estimated depths at 4 m (as suggested in Figure 4.9), the RMSE for BEM 2016 is reduced to 0.94 m.
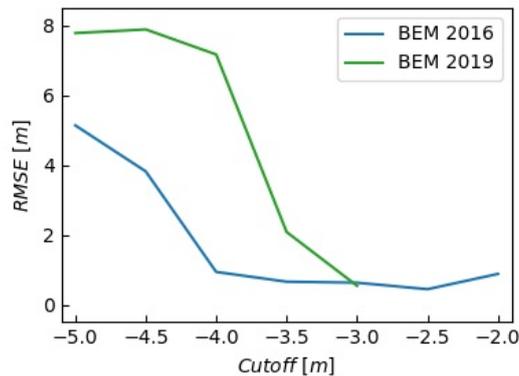
Figure 4.10: Visualisation of RMSE calculated with a threshold on estimated depths to evaluate the cutoff to which estimated depths are reliable. Two scenarios are illustrated; the bathymetry estimation model trained on a six-month composite and vaklodingen in 2016 (BEM 2016) applied to the remainder of its training composite, and the bathymetry estimation model trained on a six-month composite and vaklodingen from 2019 (BEM 2019) applied to the remaining testing subset of its training composite. RMSE errors decrease significantly by defining a shallower cutoff, which implies that estimation is more accurate for shallower depths.

Defining an appropriate threshold on the estimated depths is important for the interpretation of the results. RMSE errors decreased significantly when estimated depths are limited by lowering this threshold. This reduction in RMSE implies that estimation accuracy improves for shallower depths. Lowering the threshold also reduced the range of predictable depths. Defining this threshold is therefore a trade-off between improving estimation accuracy and maintaining a range of predictable depths. Based on these criteria, the threshold on estimated depths is set to 4 m.

Estimation errors were significantly smaller for the BEM 2016 scenario than for BEM 2019. These lower RMSE values may suggest that estimation is more accurate with BEM 2016 instead of BEM 2019. The RMSE is however not only an indicator of the performance of the estimation model, it is especially an indication of the quality of the composite that was used. From these RMSE values it is difficult to distinguish whether it was training the estimation model or applying this model to estimate depths, that is limited by composite quality.

4.2.4. Bathymetry along transects

To understand the potential of the bathymetry estimation model to map bathymetric structures in space, estimated depths (Section 4.2.2) are visualised along three transects (Figures 4.11 and 4.12). Evaluating the capability of estimating small structures is more straightforward from these two-dimensional transects instead of the three-dimensional bathymetry maps (Figures 4.7d and 4.8d). Displayed are three transects across a local sandbank, its morphological development is observed by comparing observed depths (depicted in yellow) from the 2016 vaklodingen to the 2019 vaklodingen. Locally, bathymetric changes of Estimated depths are supported by the (propagated) standard deviation for depth estimates (Figures 4.7b and 4.8b) as a confidence interval.

The capability of the bathymetry estimation model to estimate shallow bathymetric structures and the inability to estimate deeper depths is once again demonstrated. These transects clearly show that deep depths are projected as 4 to 5 m depths. The general shape of shallow bathymetric structures is estimated successfully for both the 2016 and 2019 scenarios. However, estimates from BEM 2016 and the 2016 composite are more accurate than the estimates for 2019. Especially the relatively large structure along transect B is underestimated in 2019. Uncertainties were generally larger for estimates of deeper depths. Small-scale structures observed on top of the larger structures are visible along the estimated depth profiles, although not equally along each transect.
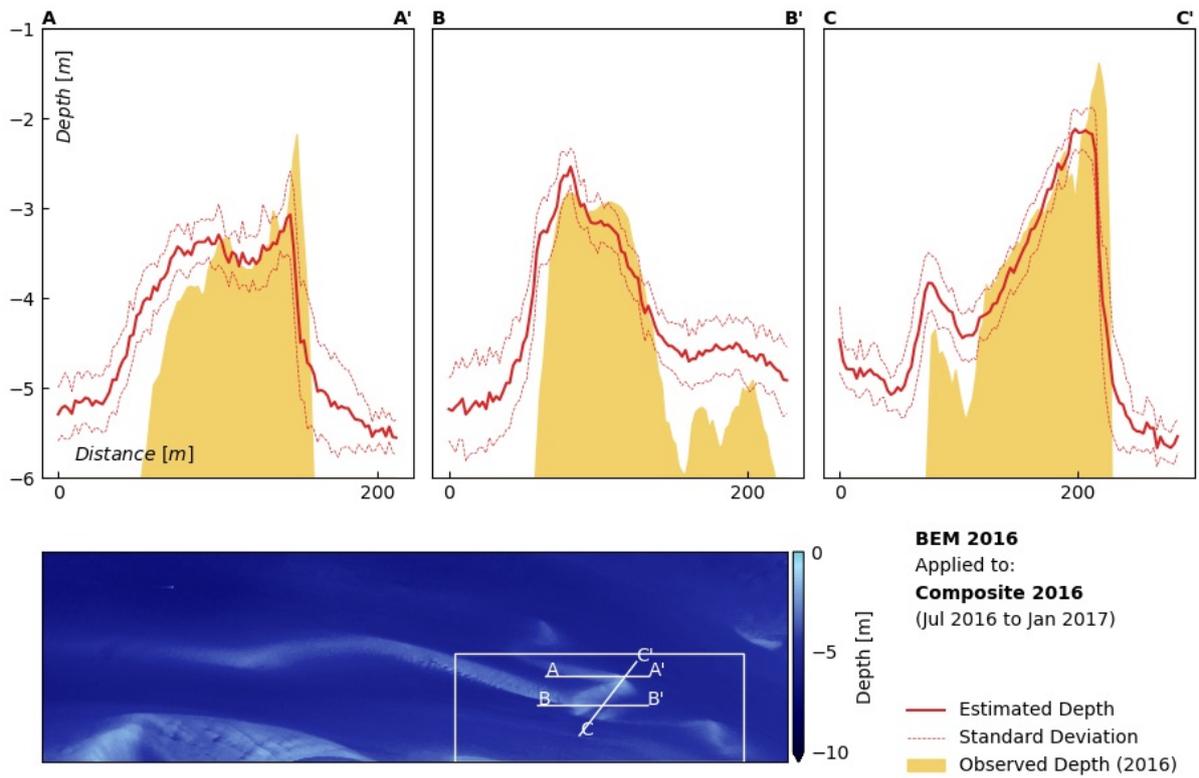
Figure 4.11: Estimated depths along transects from BEM 2016 applied to the remaining testing subset of its 2016 training composite. Shallow bathymetric structures are accurately estimated with relatively small confidence intervals. Deeper depths are estimated incorrectly and projected around 5 m. Shallow structures of varying spatial scales are observable along the estimated depths.
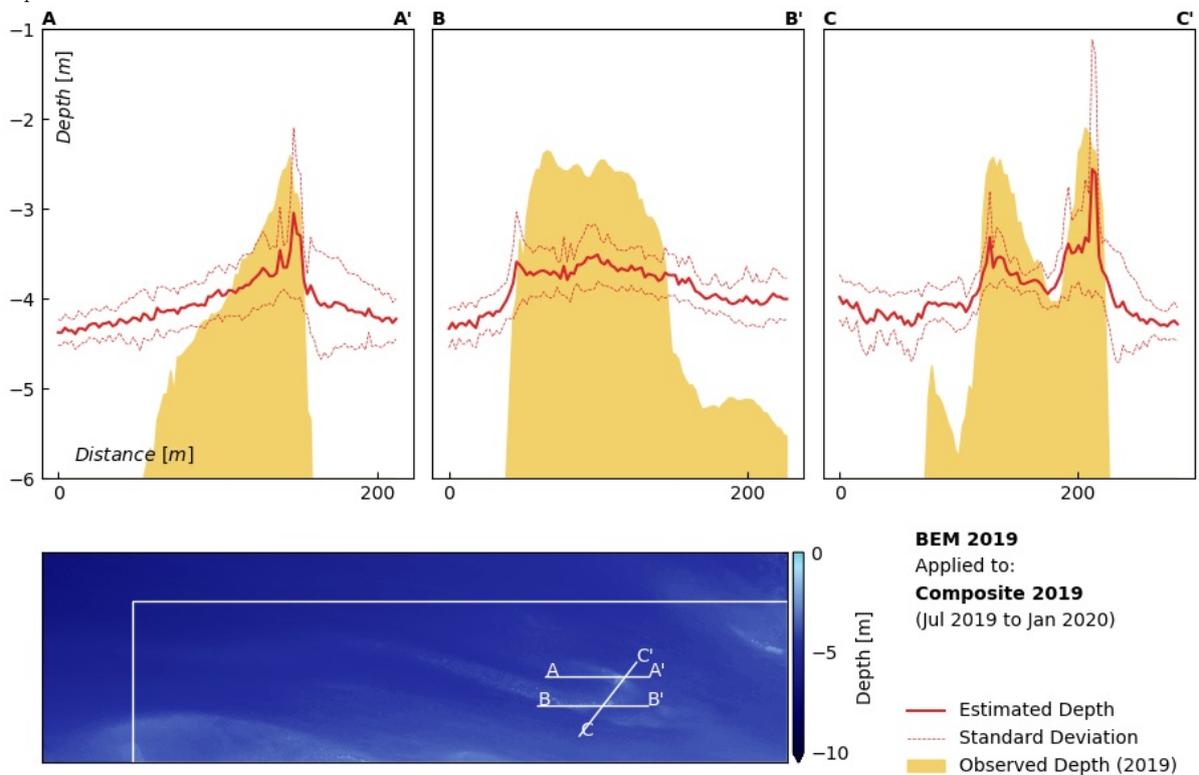
Figure 4.12: Estimated depths along transects from BEM 2019 applied to the remaining testing subset of its 2019 training composite. Shallow bathymetric structures are estimated somewhat deeper. Along transects A and C the general bathymetric profile is visible, the large scale shallow structure along transect B is however hardly observable along the estimated profile. Deep depths are incorrectly estimated to be approximately 4 m.

4.3 Bathymetry for absent in situ data

Now that a benchmark for the performance of the log-linear bathymetry estimation model is defined, the estimation model can be extrapolated in time to estimate bathymetry for times of absent in situ data. The performance of this temporal extrapolation approach is first verified by using the non-synchronous vaklodingen to validate extrapolated results.

4.3.1. Validating temporal extrapolation with bathymetry maps

The influence on estimation performance from temporal extrapolation of the bathymetry estimation model is assessed. The availability of two sets of vaklodingen allowed for cross-validating the two bathymetry estimation models. Depth estimates are obtained from BEM 2016 applied to the 2019 composite and compared to the 2019 vaklodingen (Figure 4.13). Similarly, depth estimates from BEM 2019 and the 2016 composite are compared to the 2016 vaklodingen (Figure 4.14). Prior to estimation, the models were constrained to the penetration depth of 5 m (Section 4.2.1) by setting a threshold on the observations used for training (Section 3.3.2).

Estimated depths (Figures 4.13a and 4.14a) display similar shallow bathymetric structures to depth estimates as the result from the estimation model that were trained on synchronous in situ data (Figures 4.7a and 4.8a). Shallow structures are observable and deep depths are as expected incorrectly projected to be equally shallow at 4 to 5 m. The bathymetric structures are better defined and more accurately estimated by BEM 2019 in combination with the 2016 composite (Figure 4.14d) than by the BEM 2016 applied to the 2019 composite (Figure 4.13d). Sharp edges between shallow and deeper areas are observable, even in the speckly estimates from BEM 2016 with the 2019 composite (Figure 4.8a). Uncertainties for the depth estimates are correspondingly largest for the shallow structures of interest estimated by BEM 2016 with the 2019 composite (Figure 4.13b). By applying BEM 2016 to the 2019 composite, shallow areas are estimated significantly too deep (Figure 4.13d). Based on the 2016 composite, the BEM 2019 estimates the shallow structures only slightly too deep (Figure 4.14d).

BEM 2016
Applied to:
Composite 2019
(Jul 2019 to Jan 2020)

a) Estimated depths

b) Estimation quality

c) Observed depths
(Vaklodingen 2019)

d) Difference between
observed and estimated
depths

White frame represents
extent of vaklodingen
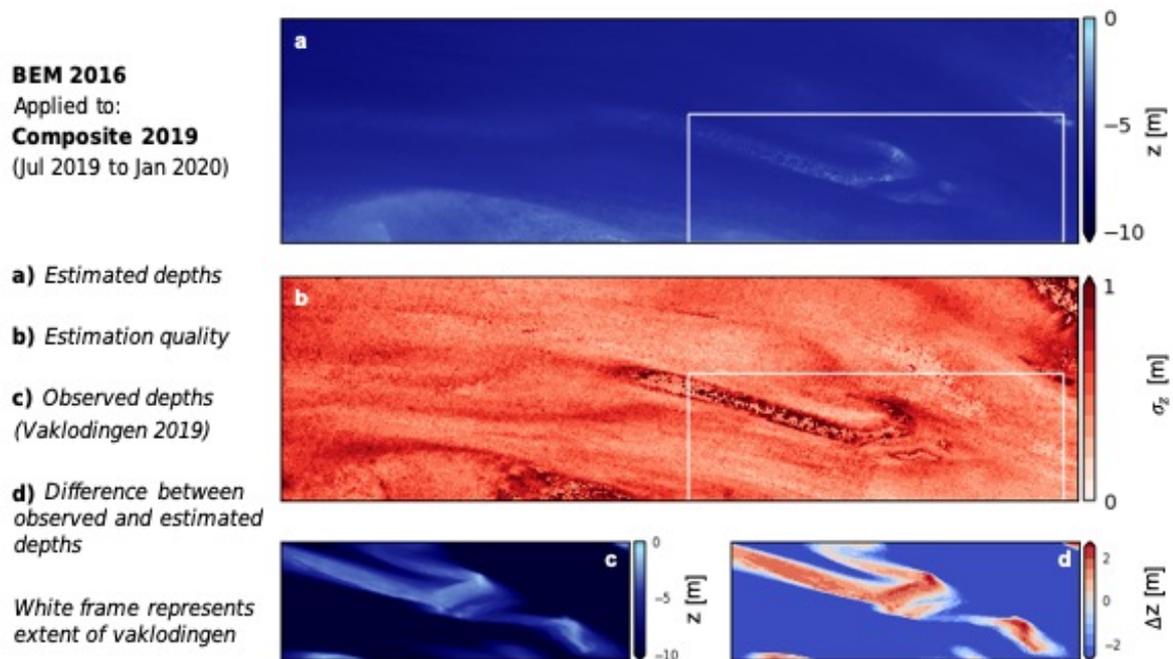
Figure 4.13: Bathymetry estimates and estimation quality from BEM 2016 with 2019 composite, compared with 2019 vaklodingen.

BEM 2019
Applied to:
Composite 2016
(Jul 2016 to Jan 2017)

a) Estimated depths

b) Estimation quality

c) Observed depths
(Vaklodingen 2016)

d) Difference between
observed and estimated
depths

White frame represents
extent of vaklodingen
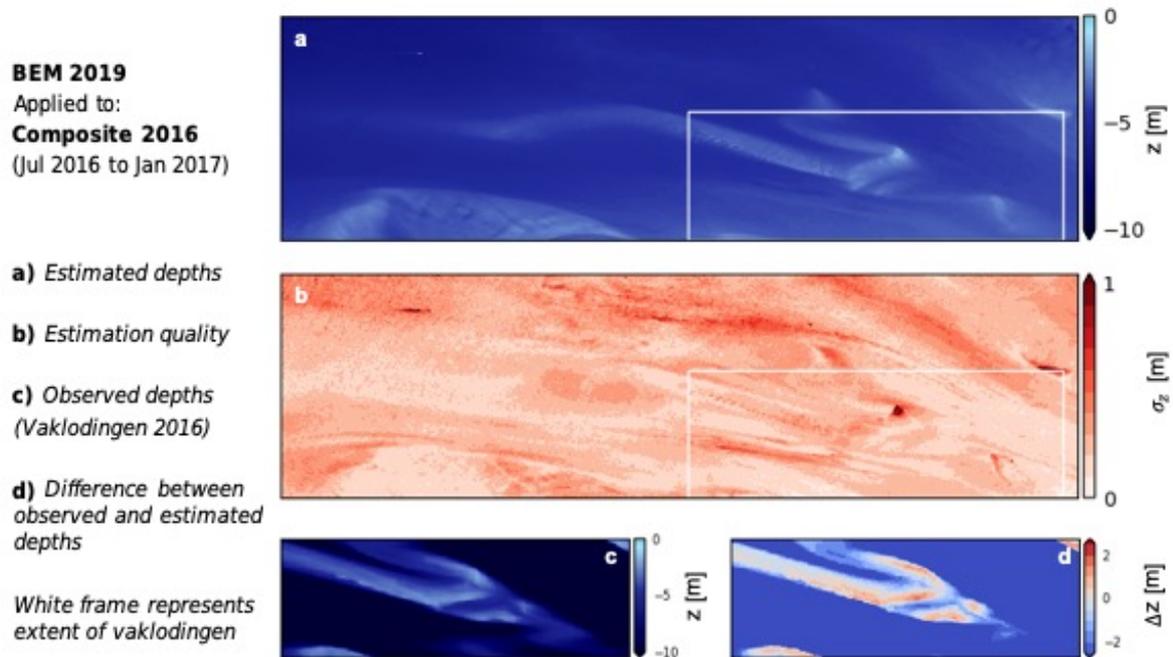
Figure 4.14: Bathymetry estimates and estimation quality from BEM 2019 with 2016 composite, compared with 2016 vaklodingen.

Comparing these results of the bathymetry estimation models applied to non-synchronous composites (Figures 4.13 and 4.14) to the estimates of the estimation models and their synchronous training composite (Figures 4.7 and 4.8), provides insight on the performance of the bathymetry estimation models. When applied to the synchronous composites, BEM 2016 outperformed BEM 2019. However, when applied to the non-synchronous composites, BEM 2019 outperformed BEM 2016. This seeming contradiction implies that composite quality is more important to estimating depths than to training the estimation model. Both BEM 2016 and BEM 2019 perform significantly lower when applied to the low-quality composite from 2019 (Figure 4.5).

4.3.2. Validating temporal extrapolation with bathymetry along transects
To illustrate the potential of estimating bathymetry without the use of synchronous in situ data, depth estimates from Section 4.3.1 are visualised along three transects (Figures 4.15 and 4.16). Estimated depths are plotted along depth observations from the vaklodingen and propagated uncertainties are depicted as confidence intervals.

Although the outline of the shallow bathymetric structures is somewhat observable in the estimates, uncertainties are irregular and large for both deep and shallow estimates. Estimates from BEM 2019 and the 2016 composite are smooth and successfully estimate the shallow bathymetric structures (Figure 4.16). Small-scale structures observed on top of the large structure are visible along the estimated profiles, although not equally along each transect. As expected, deep depths are not estimated correctly and are estimated at approximately 4.5 m. The small peaking structure at approximately 5 m depth in Transect C was estimated, albeit too shallow.
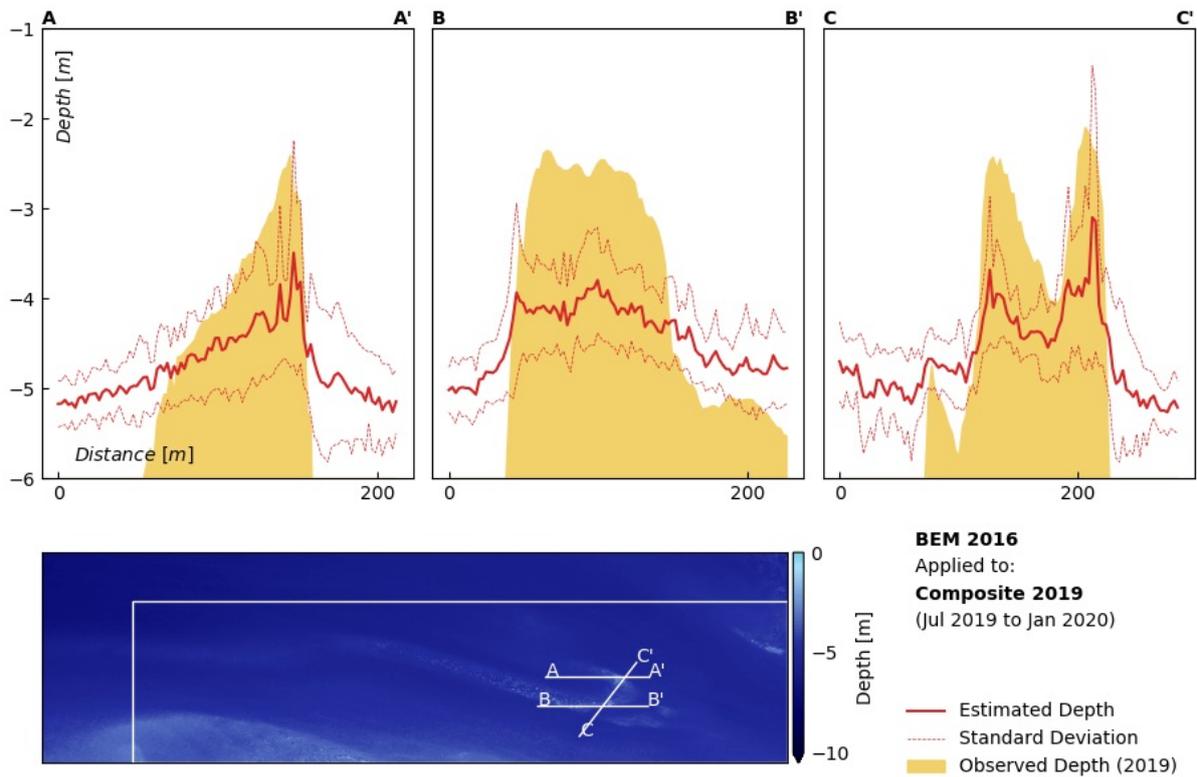
Figure 4.15: Estimated depths along transects from BEM 2016 applied to the 2019 composite. Large-scale structures are somewhat observable from the irregular estimated profile. However, uncertainties of both shallow and deep estimates are large.
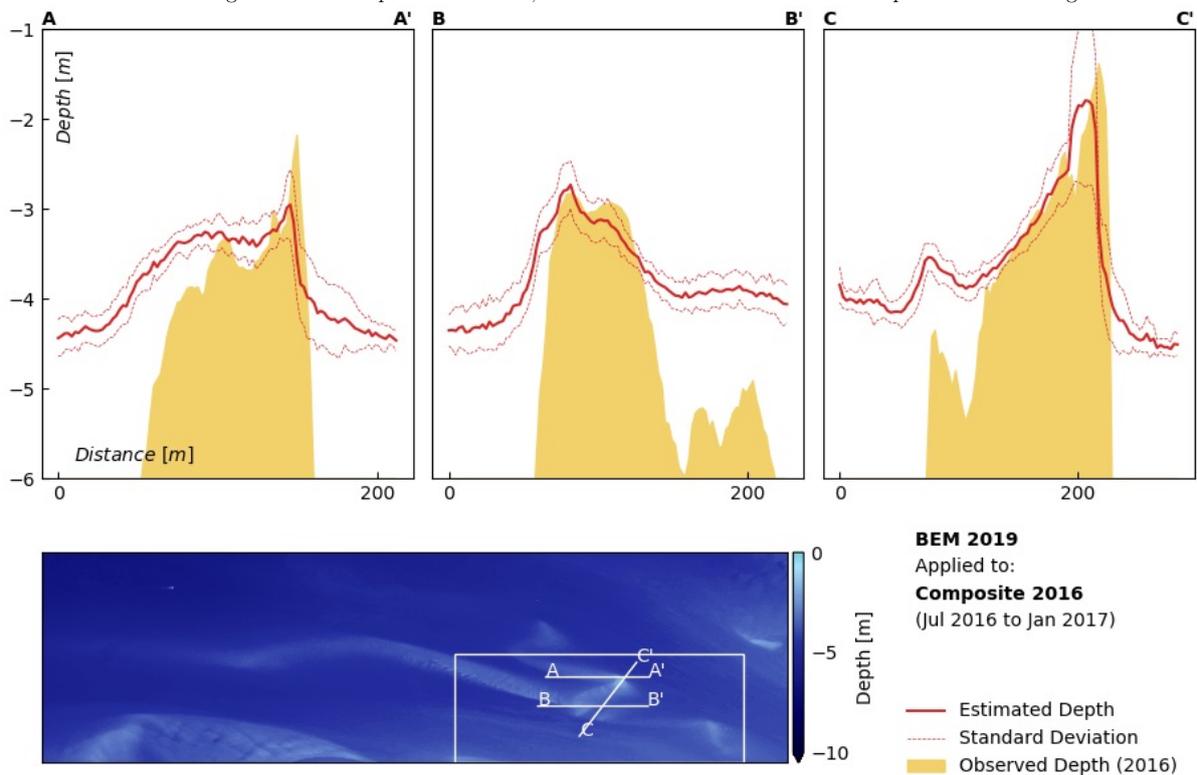
Figure 4.16: Estimated depths along transects from BEM 2019 applied to the 2016 composite. Both large-scale and small-scale shallow structures are correctly estimated. Deep depths are incorrectly projected at approximately 4.5 m.

Solely comparing the transect depth estimates from BEM 2016 with the synchronous 2016 composite (Figure 4.11) to the transects from the non-synchronous 2019 composite (Figure 4.15 could suggest extrapolation in time reduces estimation performance. Estimates along these transects from BEM 2019 however visibly improved with the non-synchronous 2016 composite (Figures 4.12 and 4.16). The visualisation of estimated depths along the transects accentuates the limiting effect of applying a bathymetry estimation model to the low-quality 2019 composite.

The accurate estimation of shallow structures without the use of synchronous in situ data demonstrate the potential of mapping temporal variations (Figure 4.16). Both large-scale and small-scale structures are visible along the estimated profile. In comparison to BEM 2016 (Figure 4.11), BEM 2019 shallow structure estimates for 2016 are slightly more limited to approximately 4.5 m instead of 5.5 m depth. This maximum predictable depth is likely an effect of the low-quality 2019 composite which was used to train BEM 2019.

For transects A and B the eastern edge of the profile is estimated with a small horizontal bias. This offset could be due to the incidence angle of the sun which results in differences in illuminated sides of the bathymetric structure. This hypothesis is supported by the absence of a horizontal bias in transect C, which is oriented somewhat perpendicular to transects A and B.

Confidence intervals are approximately constant along the estimated profiles, although depths estimated around the prediction threshold of 4 to 4.5 m are slightly more uncertain. Large uncertainty is especially visible for the steep transition of shallow to deep bathymetry. These large uncertainties could be to the influence of neighbouring pixels on the bathymetric signal by for example refraction. As a result, small horizontal errors occur around steep transitions from shallow to deep depth. On steep slopes, small horizontal variations coincide with large vertical differences, which are visible along the estimated profiles.

Bathymetry from the bathymetry estimation models and non-synchronous composites were cross-validated by using the non-synchrnous vaklodingen data. RMSE erros for multiple thresholds on estimated depths from were compared for the 2016 (triangles) and 2019 (circles) composites and three bathymetry estimation models; BEM 2016 (blue), BEM 2019 (green) and an estimation trained on the reference image (red) (Figure 4.10).
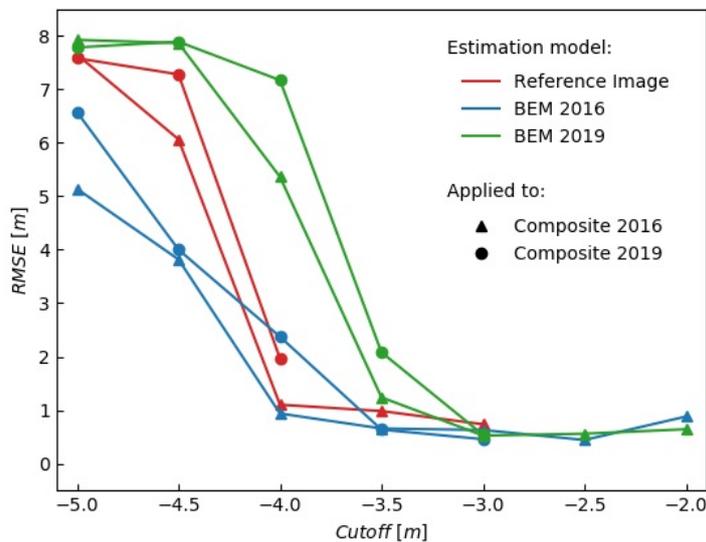
Figure 4.17: Visualisation of RMSE calculated with a threshold on estimated depths. Depths and RMSEs were calculated for three different bathymetry estimation models (BEM 2016, BEM 2019 and trained on the reference image) and applied to the 2016 and 2019 composites.

For all three bathymetry estimation models, the RMSE for the synchronous and non-synchronous scenarios are similar. This indicates that temporal extrapolation, by applying a model to an image

three years apart from the training image, does not strongly affect the error of the estimation. These findings verify the assumption that a general model relating depth to intensity exists after correcting the multispectral imagery for all non-bathymetric variations. The estimation errors for the synchronous scenario are expected to be less than those of the non-synchronous case. However, the RMSE for the synchronous composite is not consistently lower. For all three models, the estimation errors are larger when applied to the 2019 composite. Even the 2019 model applied to its synchronous composite, results in larger errors than when applied to the 2016 composite. This suggests that the quality of the composite image has a larger effect on the estimation errors than temporal extrapolation. This also implies that reducing the estimation error by implementing a different model will benefit the extrapolated results equally. Figure 4.17 highlights the potential of the extrapolating a readily-trained bathymetry estimation model. As the RMSE does not significantly increases by extrapolating these models in time, the approach successfully prepares multispectral imagery for extrapolation. Implementing another bathymetry estimation model may therefore improve extrapolated results.

Figure 4.17 demonstrates the effect of averaging. The model trained on the reference image, cannot estimate depths shallower than 3 [m] when applied to the 2016 and 2019 composites. By creating a composite, peak signals of individual images are averaged. This discrepancy between a single image and composite results in a bias when applying a single image model to a composite image, and vice versa. The model trained on the reference image has a negative bias when it is applied to a composite image, hence the most shallow depths completely absent in the estimation. To optimise bathymetry estimation on composites, the model should therefore be trained on a composite image too.

4.3.3. Temporal bathymetric changes in timeseries

Temporal bathymetric changes are estimated for six locations by using successive partially-overlapping composites (Figure 4.18). Timeseries are created by applying the estimation models BEM 2016 (blue) and BEM 2019 (green) to composite Sentinel-2 images spanning from 2016 to 2020 (Section 3.4.1). Estimated depths are plotted against the vaklodingen observations (star symbol).

Depth increased between the vaklodingen observations for five locations, a decrease in depth is observed only for location three (P3). Estimates display comparable trends from 2016 to 2019. Estimates from both bathymetry estimation models display display similar profiles with fluctuations of varying scales. The majority of the depth estimates from BEM 2016 is approximately 0.2 m to 1 m deeper than estimates from BEM 2019. BEM 2016 estimates depths up to 5 m, whereas estimates from BEM 2019 are limited to 4 m.
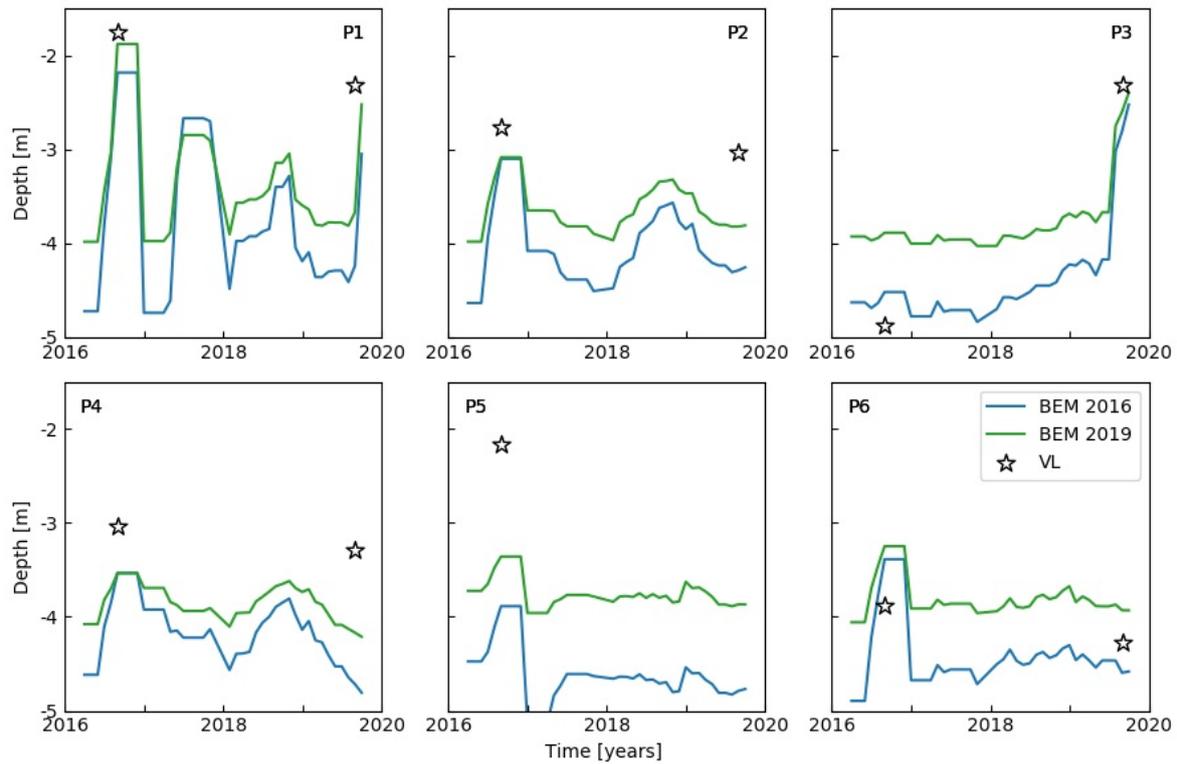
Figure 4.18: Temporal bathymetric changes for six point locations estimated by applying the BEM 2016 (blue) and BEM 2019 (green) estimation models to partially-overlapping six-month composites in comparison to vaklodingen observations (depicted by the star symbol).
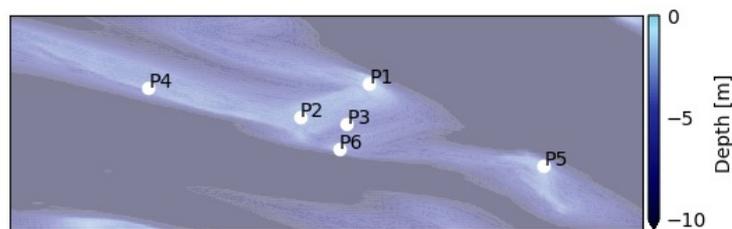
Figure 4.19: Point locations from Figure 4.18 visualised on top of the 2016 vaklodingen.

As expected, deep depths (eg. P3 2016 or P6 2019) are projected to the maximum predictable depth; approximately 4 m for BEM 2019 and approximately 5 m for BEM 2016. The shallower profile of BEM 2019 therefore appears to better approach the selected shallow observations. BEM 2019 however closely estimated the large difference between observations in P3. Based on these timeseries, both bathymetry estimation models perform approximately equal.

The steep fluctuations which are especially visible for location one (P1), appear to be somewhat unrealistic in comparison to the small difference that is observed between vaklodingen. These large temporal fluctuations could be caused by location one's position on a steep slope. On steep slopes, small horizontal variations correspond with large vertical differences. Reflections from neighbouring pixels due to for example refraction of light influence the bathymetric signal. As a result, small horizontal errors occur around steep bathymetric edges. This effect is also visible in depth estimates along transects (eg. Figure 4.16). This method of visualising temporal changes of point locations, is expected to be very susceptible to large depth fluctuations as a result of small horizontal errors.

4.3.4. Temporal bathymetric changes of two-dimensional structures

To provide context to the temporal changes, estimates are visualised as part of a bathymetric structure. Here, temporal changes of (part of) a sandbank structure are estimated by using four composites of consecutive years (Figure 4.20). The outline of the bathymetric structure is define as the 4 m contourline. The observed outlines of the sandbank are depicted in grey; 2016 (dotted) and 2019 (dashed) (Figure 4.20a). Between these observations, the outline of the structure retreated towards the right and the narrow tip changed direction. To estimate this temporal trend, intermediate annual changes are estimated from BEM 2019 in combination with a six-month composite for each year (depicted in pink) (Figure 4.20b). The lightest colour displays the sandbank estimated for 2016, the darkest colour represents the structure in 2019. From light to dark the gradual retreat of the bathymetric structure is visible. To visually improve the image, both estimated depth and vaklodingen lines were smoothed by a Gaussian low-pass filter. Because vaklodingen data were used to anchor the bathymetry estimation models to depth, estimates are referred to the vaklodingen chart datum, which is Normaal Amsterdams Peil.
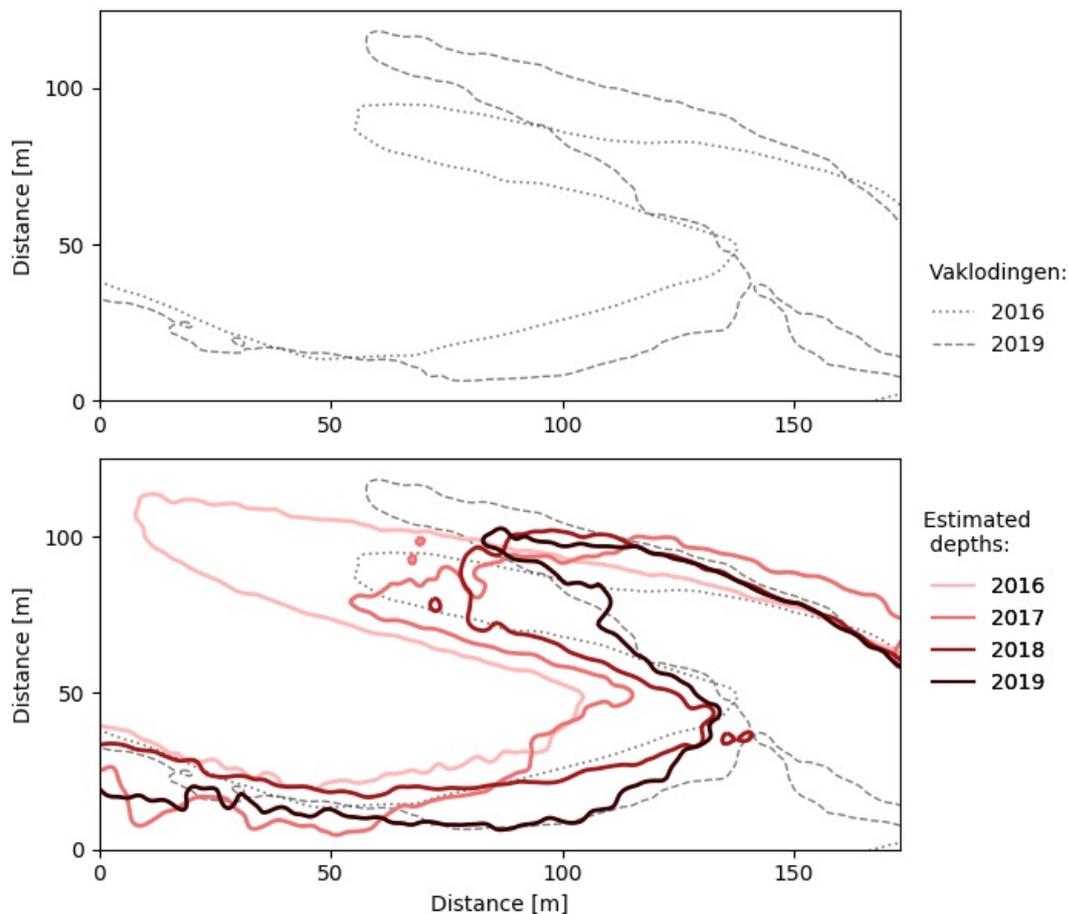
Figure 4.20: Outline of a bathymetric structure defined by the -4 m contour line. a) Temporal development of this sandbank structure from 2016 to 2019 from vaklodingen observations. b) Intermediate annual development estimated from BEM 2019 and six-month composites from four consecutive years.

The estimated depths show a temporal trend similar to the vaklodingen. The 2019 vaklodingen show the start of the breakthrough of the sandbank in the right lower corner of the figure. Although the estimates did not match this morphodynamic change completely, the retreating pattern is however clearly visible. Composite quality strongly affects the performance of depth estimation, the contourlines are therefore not solely representative of the performance of extrapolation. Increasing the duration of the composite would increase the number of images and consequently improve its quality. It can be argued, for example, that the breakthrough of the bathymetric structure would be better estimated by improving the 2019 composite image. However, by doing so, the balance between observable changes and and

quality images needs to be maintained. By showing only a single contour line, the extent of the line is very sensitive to small deviations in depth and should thus be interpreted with care. The orientation of the structures depicted by the contour lines is less sensitive to small errors and is therefore a good representation of the structures.

4.4 Additional points of discussion

To finalise this chapter, I would like to address some additional points of discussion. This section aims to elaborate on the consequences of selecting a different reference image (Section 4.4.1), the impact of tidal water level variations and data scarcity on the quality of a composite (Section 4.4.2), the influence of several modifications to the bathymetry estimation model (Section 4.4.3) and effects of temporal extrapolation (Section 4.4.4).

4.4.1. Reference image

The reference image correction method enabled direct comparison of multitemporal images and allowed for the use of a time-independent bathymetry estimation model. The quality and characteristics of the reference image therefore have an extensive influence, since all multispectral imagery is corrected to this image. Here, I would like to elaborate on the influence of (i) a different high-quality reference image and (ii) a low-quality reference image.

A high-quality reference image is required for the reference image correction method because it is used to correct for large-scale non-bathymetric variations in the Sentinel-2 imagery (Section 2.3.2). To achieve this, the bathymetric signal in the reference image should be unobstructed; a high-quality cloud-free image is required where propagation of the shallow bathymetry signal is not limited by turbidity or high water levels. An image at low water is preferred; at low water an increased level of details is displayed because the bathymetric signal is significantly stronger. To obtain an image that satisfied these criteria, the reference image was manually selected from a collection atmospherically processed images. The reference image correction method is a relative correction which brings all imagery to the level of the reference image. Although selecting another satisfactory image would change the reference level, that change is compensated by the bathymetry estimation model which anchors the imagery by using the vaklodingen. Therefore, I expect that another clean and high-quality reference image will have little effect on the estimated depths. To verify this line of reasoning, another high-quality reference image should be tested and results should be compared.

The impact of a low-quality reference image would be more extensive. Any adverse features or errors in the reference image are widely propagated because all multispectral imagery is corrected to this image. Correcting the imagery to an obstructed reference image would therefore complicate isolating the bathymetric signal from these images. Furthermore, the reference image is used to identify images obstructed by large-scale non-bathymetric effects that cannot be removed by an atmospheric correction (Section 2.3.2). Selecting a low-quality reference image would disable identification of images affected by those large-scale effects. As a result, the quality of composite images would decrease by including these low-quality images into a composite.

4.4.2. Composites

Composite images are used to train a bathymetry estimation model and subsequently to estimate depths by serving as input for the trained model. Results are therefore strongly influenced by the quality of the composite images. Composite quality is controlled by the images that are included into the composite and their relative differences, both a product of the composite's timespan. Here, I would like to discuss (i) the influence of differences between multitemporal images due to tidal water level variations. Furthermore, I would like to emphasise (ii) the impact of the scarcity of high-quality data on composite quality, because creating composites is a trade-off between image quality and image quantity.

Tidal water level variations may induce large-scale variations between multitemporal imagery. Repeat visits of the Sentinel-2 satellites of the Wadden Sea area are at an approximately equal time of day. Tides are the periodic motion of the oceans on Earth due to changes in the gravitational forces acting upon the rotating Earth. The gravitational pull of the Moon, which orbits the Earth in one lunar day (24 hours and 50 minutes), creates approximately two tidal cycles in one solar day (24 hours).

The tidal cycle shifts every solar day due to the discrepancy between a solar and lunar day. The gravitational pull of the Sun introduces an additional cycle of tidal water level variation that occurs within a month (spring/neap). As a result, satellite images covering a timespan of several months record water levels across both high/low and spring/neap variations. These tidal water level variations may induce large-scale differences between multitemporal images. The reference image correction method is developed to minimise such large-scale variations between images, therefore it also acts as a correction for tidal water level variations. Corrected imagery and temporal differences are additionally averaged by creating a composite image. Therefore, the influence of tidal water level variations on composites and depth estimates appears to be minimal. To verify this hypothesis, a composite of images with large tidal water level variations should be created and used to train a bathymetry estimation model. Subsequently, its depth estimates should be tested against those of a composite with minimal variation in tidal water levels.

Local conditions such as frequent cloud coverage and high turbidity levels result in a scarcity of high-quality Sentinel-2 images. In this study, high-quality images are identified by thresholds on cloud coverage and relative deviation in the green band. These threshold parameters are tuned to maximise the number of high-quality images identified, while simultaneously minimising the number of low-quality image incorrectly marked as high-quality. The identification of high-quality images is especially sensitive to the threshold on relative deviation in the green band. A balance between high-quality imagery and a sufficient number of images was maintained by setting this threshold at 8.5 %. In this study, increasing the threshold with 0.5 % approximately doubled the number of images in the total four-year collection of Sentinel-2 images, whereas reducing it with 0.5 % halved that number. Determining which images to include is a trade-off between quality and quantity; including low-quality images solves the problem of scarce high-quality data yet it reduces composite quality. Because of the sensitivity to this threshold, the threshold on the relative deviation in the green band should be determined with care by manually inspecting (some of) the corrected imagery. The identification of high-quality images was less sensitive to the threshold on cloud coverage. Lowering the threshold from 20 % to 10 % removed one image from the thirty-nine images in the four-year collection. Increasing the threshold to allow up to 40 % cloud coverage, increased the number of images in the four-year collection with seven. A sensitivity study could quantify the influence of both the cloud coverage threshold and the threshold on the relative deviation in the green band.

4.4.3. Bathymetry estimation model

The bathymetry estimation model in this study is obtained by linear regression on a composite image and in situ depth observations. The three-year extrapolation that is performed on this estimation model did not decrease its predictive power; the root mean square errors were approximately equal at one metre. The absence of error growth suggests that temporal extrapolation does not introduce significant estimation errors. This appears to imply that improving composite quality or the bathymetry estimation model will improve extrapolated estimation results accordingly. Therefore, I would like to discuss (i) the influence of training the estimation model on a composite image and (ii) the expandability of the presented work by evaluating four alternatives for the estimation model.

Estimation results differed for a bathymetry estimation model trained on a composite or on a single image. The effect of averaging multiple images into a composite is visible by comparing the estimation performance of three different estimation models (Section 4.3.2). The bathymetry estimation model that is trained on a single image displayed a bias in the depth estimates obtained from a composite. Peak signals are flattened in a composite image, this effect is particularly observable for very shallow water that appeared to be somewhat deeper. The peak signals, which describe minimum depths, are much higher in an individual image. A bathymetry estimation model that is trained on the averaged signal of a composite is unable to predict the most shallow depths represented by the high peak signals in an individual image. To avoid this estimation bias, the image type on which the bathymetry estimation model is trained should be equal to that of the image used for estimation. Hence, in this study, estimating depths from composite images required training the estimation model on a composite.

The application of the log-linear bathymetry estimation model is physically limited to the penetration depth. Due to the projection of deeper depths onto the penetration depth, reliable estimates were further constrained (Section 4.2.3). Both the penetration depth and a threshold on the reliability of

estimated depths can be obtained by using in situ depth observations. In clear water conditions the penetration depth is larger, which enables estimating deeper depths. Additionally, the projection of deep depths onto the penetration depth is minimal if the water is sufficiently clear and bathymetry sufficiently shallow. Application of the presented method to an alternative study site with clear water conditions is therefore favourable, as depth estimation under these conditions is not dominated by limitations of the penetration depth. Despite an increased range of predictable depths of the log-linear estimation model, bathymetry derivation in optically clear waters is often complicated by other factors such as heterogeneous bottom habitats.

Bathymetry estimation models based on ratios of spectral bands are demonstrated to be robust to spatial variations in bottom type (e.g. Stumpf et al., 2003). The homogeneous bottom type of the Wadden Sea study site allowed for use of the log-linear estimation model. Implementing a band ratio model will expand the applicability of the presented method and enable the estimation of bathymetry across spatially-varying habitats. Bathymetry estimation may be additionally complicated by spatial variations across the water column such as turbidity (Caballero and Stumpf, 2020b). Substituting the log-linear estimation model may also improve depth estimation for the turbid Wadden Sea study site, as band ratio models are also demonstrated to be relatively robust to those spatial variations (Stumpf et al., 2003). The presented extrapolation method is potentially applicable to estimate nearshore shallow bathymetry at any global location if there is an estimation model available that can account for the site's spatial characteristics.

The log-linear estimation model may be improved by a different combination of the spectral bands. Here, a linear combination of the log-transformed red, green and blue bands of Sentinel-2 imagery is used. Removing red light from this equation for its limited capacity to penetrate the water column may improve depth estimation (e.g. Traganos et al., 2018). Alternatively, additional Sentinel-2 bands may be exploited to develop an estimation model that employs an increased number of spectral bands. Furthermore, machine-learning techniques are demonstrated to improve estimation (e.g. Misra et al., 2018, Sagawa et al., 2019). To verify the hypothesis that an improved estimation model directly improves extrapolated estimation results, a comparative study could quantify predictive power for multiple estimation models.

Although this study omits the need for synchronous in situ data, it is still dependent on at least one set of in situ depth observations. In a recent study, Thomas et al. (2020) demonstrated a method to substitute in situ depth observations with spaceborne lidar data from the ICESat-2 satellite. Incorporating ICESat-2 data to train the bathymetry estimation would greatly expand the applicability of the work presented in this thesis. ICESat-2's continuous and global recordings are publicly available; its data may provide new observations of locations yet to be surveyed by traditional campaigns or offer a high-frequent alternative for areas that are regularly surveyed such as the Wadden Sea. Because the proposed method enables temporal extrapolation of a bathymetry estimation model, it may be used to utilise recent ICESat-2 observations to map historical bathymetric changes.

4.4.4. Temporal extrapolation

Within in the tested three-year time frame, temporal extrapolation did not decrease the predictive power of bathymetry estimation in the Wadden Sea. These results imply that estimation performance is predominantly governed by composite quality and the predictive power of the bathymetry estimation model. This suggests that improving composite quality or the estimation model will directly improve extrapolated results. Quantifying the effect of temporal extrapolation is however difficult, because bathymetry estimation for this case study is strongly limited by the relatively shallow penetration depth due to highly turbid waters. The estimation errors due to high-turbidity are significantly larger than those introduced by the temporal extrapolation. To isolate the estimation errors from temporal extrapolation, the presented approach should be tested for a less-constrained case study. Applying temporal extrapolation of a bathymetry estimation model in optically clear waters will provide more insight into its effect on predictive power and resulting temporal limits.

5

Conclusions

5.1 Conclusions

Mapping temporal bathymetric changes from satellite derived bathymetry is challenging because the majority of bathymetry estimation methods is heavily dependent on in situ observations. As a result of this dependency, the availability of synchronous in situ data governs the derivation of bathymetry from multispectral imagery. For this thesis, I developed an approach to estimate temporal changes from satellite derived bathymetry by estimating bathymetry for times of absent in situ data. This approach included a reference image correction method for large-scale non-bathymetric variations, composite images created with a spatiotemporal filter and temporal extrapolation of a readily-trained bathymetry estimation model. To investigate the potential of the approach, four sub-questions are defined in support of one main research question. Here, I summarise my findings for the Wadden Sea case study based on these five questions.

1. *How to enable direct comparison of multitemporal satellite imagery in order to map temporal bathymetric changes?*

   Although tailored processing of multispectral imagery is in general important to obtain satellite derived bathymetry, temporal analysis specifically requires imagery to be directly comparable. A reference image correction method is developed to correct individual images for large-scale non-bathymetric variations. This use of a reference image to correct multispectral imagery is found to deliver results in which temporal signal variations are dominated by bathymetric changes. The isolation of the bathymetric signal allowed for temporal extrapolation of a time-independent bathymetry estimation model. The extrapolation enabled estimating bathymetry for times of absent in situ data and subsequently mapping temporal changes.

2. *How to manage the expected scarcity of high-quality images due to challenging conditions associated with the study site?*

   High-quality images are effectively selected using two selection criteria; a threshold of 20 % on the maximum cloud coverage and a threshold of 8.5 % on the average relative pixel deviation in the green band between an image and the reference image. To manage the limited availability of high-quality images, composites with a duration of six months are constructed. A spatiotemporal filter that included an outlier detection algorithm successfully created clean and smooth composites.

3. *What approach is best-suited to relate water depth and multispectral imagery for the site-specific conditions?*

   The homogeneous bottom type of the case study allowed for a log-linear bathymetry estimation model. The estimation model is applicable to the upper part of the water column into which visible light penetrates. Linear correlation between depth and log-transformed intensities of the multispectral image is observed for the first five metres of the water column. The bathymetry estimation model is therefore trained only on depths that were observed shallower than five metre. The estimation model projected deep depths onto this five metre zone due to the ambiguous

relation between depth and intensity below the penetration depth. As a result, only depth estimates up to four metre were reliable. Structures shallower than this four metre threshold are estimated with root mean square errors of approximately one metre. Optimal estimation is obtained by training the bathymetry estimation model on a composite, instead of training on a single image. The application of an estimation model trained on a single image to a composite image resulted in a negative bias that is due to the averaging of spectral peaks in individual images. Submetre structures and large-scale formations are correctly estimated by using a composite to train the estimation model.

4. *What is the performance of extrapolation in terms of detectable bathymetric changes, temporal limits and estimation quality?*

Temporal changes are derived by extrapolating the readily-trained bathymetry estimation model to consecutive composites with the use of a moving-average filter. With this method, the position and the direction of the lateral migration of bathymetric structures are successfully estimated. No difference in estimation quality was found between the application of a bathymetry estimation model to a synchronous composite and an extrapolation of three years; the root mean square errors were approximately equal at one metre. The absence of error growth suggests that predictive power does not decrease with time. Within the tested three-year time frame, there is no temporal limit of the extrapolation.

To what extent can bathymetry be derived from multispectral imagery, for times of absent in situ data?

The presented method has shown to estimate shallow bathymetric structures with approximately one metre accuracy for times of absent in situ data. Furthermore, the migration direction of these structures is correctly estimated. Composite images with a duration of six months are used to successfully estimate depths up to four metre. Predictive power did not decrease with the three-year extrapolation performed on the bathymetry estimation model, therefore there seems no temporal limit to the extent of the extrapolation within that time frame. The limited influence of extrapolation on the predictive power suggests that the availability of high-quality satellite imagery and one set of non-synchronous in situ observations is sufficient to estimate bathymetry for times of absent in situ data.

5.2 Recommendations

Based on the findings and conclusions of this thesis, the following recommendations are suggested.

Deriving bathymetry from multispectral imagery is generally challenging for waters with high turbidity such as the Wadden Sea. Assessing the influence of temporal extrapolation on a bathymetry estimation model is therefore increasingly difficult because the effects of turbidity dominated the estimation. To quantify the effect of temporal extrapolation on estimation and to further investigate its temporal limits, I recommend to test the presented approach under less constraining conditions such as optically clear waters.

Estimation quality is strongly dependent on the performance of the bathymetry estimation model. Implementing a more-advanced or improved estimation model will directly improve extrapolated depth estimates. Selecting a robust estimation model that is well-suited to the characteristics of a study site is therefore crucial. For the Wadden Sea case study I suggest improving the log-linear model by changing the number of bands or substituting it for a band ratio model.

Estimation quality is furthermore governed by the quality of the composite image. The composite is created by balancing quantity and quality of individual images on a timescale that represents the dynamics of the morphological structures of interest. In that context, I recommend to explore different composite configurations to establish an optimum. This specifically includes adjusting the duration of the composite and testing the sensitivity to the two quality thresholds: the maximum cloud coverage and the relative deviation in the green band. To quantify composite quality, I suggest to establish a metric that describes the sensitivity to these tuning parameters. One way to this, would be expressing the relative deviation of the composite with respect to the reference image, in a way similar to the green band deviation that is used for individual images.

Bibliography

Bierwirth, P., Lee, T. and Burne, R. (1993), 'Shallow sea-floor reflectance and water depth derived by unmixing multispectral imagery', *Photogrammetric Engineering and Remote Sensing;(United States)* **59**(3).

Bosboom, J. and Stive, M. J. (2012), *Coastal dynamics I: lectures notes CIE4305*.

Buiteveld, H., Hakvoort, J. and Donze, M. (1994), Optical properties of pure water, *in* 'Ocean Optics XII', Vol. 2258, International Society for Optics and Photonics, pp. 174–183.

Caballero, I. and Stumpf, R. P. (2020*a*), 'Atmospheric correction for satellite-derived bathymetry in the Caribbean waters: from a single image to multi-temporal approaches using Sentinel-2A/B', *Optics Express* **28**(8), 11742–11766.

Caballero, I. and Stumpf, R. P. (2020*b*), 'Towards routine mapping of shallow bathymetry in environments with variable turbidity: contribution of Sentinel-2A/B satellites mission', *Remote Sensing* **12**(3), 451.

Casal, G., Hedley, J. D., Monteys, X., Harris, P., Cahalane, C. and McCarthy, T. (2020), 'Satellite-derived bathymetry in optically complex waters using a model inversion approach and Sentinel-2 data', *Estuarine, Coastal and Shelf Science* p. 106814.

Casal, G., Monteys, X., Hedley, J., Harris, P., Cahalane, C. and McCarthy, T. (2019), 'Assessment of empirical algorithms for bathymetry extraction using Sentinel-2 data', *International Journal of Remote Sensing* **40**(8), 2855–2879.

Chavez Jr, P. S. (1988), 'An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data', *Remote sensing of environment* **24**(3), 459–479.

De Swart, H. and Zimmerman, J. (2009), 'Morphodynamics of tidal inlet systems', *Annual review of fluid mechanics* **41**, 203–229.

Deltares (2018), 'OpenEarth Database'.
URL: *https://opendap.deltares.nl/thredds/catalog/opendap/rijkswaterstaat/vaklodingen/catalog.html*

Elias, E., Van der Spek, A., Wang, Z. B. and De Ronde, J. (2012), 'Morphodynamic development and sediment budget of the Dutch Wadden Sea over the last century', *Netherlands Journal of Geosciences* **91**(3), 293–310.

European Space Agency (2015), 'User handbook', *ESA Standard Document* **64**.

Evagorou, E. G., Mettas, C., Agapiou, A., Themistocleous, K. and Hadjimitsis, D. G. (2019), 'Bathymetric maps from multi-temporal analysis of Sentinel-2 data: the case study of Limassol, Cyprus', *Advances in Geosciences* .

Gao, J. (2009), 'Bathymetric mapping by means of remote sensing: methods, accuracy and limitations', *Progress in Physical Geography* **33**(1), 103–116.

Geyman, E. C. and Maloof, A. C. (2019), 'A simple method for extracting water depth from multispectral satellite imagery in regions of variable bottom type', *Earth and Space Science* **6**(3), 527–537.

Gordon, H. R. and Morel, A. Y. (1983), *Remote assessment of ocean color for interpretation of satellite visible imagery: a review*, Vol. 1, Springer Science & Business Media.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. (2017), 'Google Earth Engine: Planetary-scale geospatial analysis for everyone', *Remote Sensing of Environment* **202**, 18–27.

Green, E., Mumby, P., Edwards, A. and Clark, C. (2000), *Remote sensing: handbook for tropical coastal management*, United Nations Educational, Scientific and Cultural Organization (UNESCO).

Hadjimitsis, D. G., Clayton, C. and Hope, V. (2004), 'An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs', *International Journal of Remote Sensing* **25**(18), 3651–3674.

Hastie, T., Tibshirani, R. and Friedman, J. (2009), *The elements of statistical learning: data mining, inference, and prediction*, Springer Science & Business Media.

Hedley, J. D., Roelfsema, C., Brando, V., Giardino, C., Kutser, T., Phinn, S., Mumby, P. J., Barrilero, O., Laporte, J. and Koetz, B. (2018), 'Coral reef applications of Sentinel-2: Coverage, characteristics, bathymetry and benthic mapping with comparison to Landsat 8', *Remote Sensing of Environment* **216**, 598–614.

Hedley, J., Roelfsema, C. and Phinn, S. R. (2009), 'Efficient radiative transfer model inversion for remote sensing applications', *Remote Sensing of Environment* **113**(11), 2527–2532.

Iqbal, M. (2012), *An introduction to solar radiation*, Elsevier.

Ishimaru, A. (1991), *Electromagnetic wave propagation, radiation, and scattering: from fundamentals to applications*, John Wiley & Sons.

Kerr, J. M. and Purkis, S. (2018), 'An algorithm for optically-deriving water depth from multispectral imagery in coral reef landscapes in the absence of ground-truth data', *Remote sensing of environment* **210**, 307–324.

Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G. and Patch, J. S. (1999), 'Hyperspectral remote sensing for shallow waters: 2. deriving bottom depths and water properties by optimization', *Applied optics* **38**(18), 3831–3843.

Lyzenga, D. R. (1978), 'Passive remote sensing techniques for mapping water depth and bottom features', *Applied optics* **17**(3), 379–383.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U. and Gascon, F. (2017), Sen2Cor for Sentinel-2, *in* 'Image and Signal Processing for Remote Sensing XXIII', Vol. 10427, International Society for Optics and Photonics, p. 1042704.

Marshall, J. and Plumb, R. (2007), *Atmosphere, Ocean and Climate Dynamics: An Introductory Text*, International Geophysics, Elsevier Science.

Misra, A., Vojinovic, Z., Ramakrishnan, B., Luijendijk, A. and Ranasinghe, R. (2018), 'Shallow water bathymetry mapping using Support Vector Machine (SVM) technique and multispectral imagery', *International journal of remote sensing* **39**(13), 4431–4450.

Morel, A. (1974), Optical properties of pure water and pure sea water., *in* 'Optical aspects of oceanography', Vol. 12, Academic Press London.

Pacheco, A., Horta, J., Loureiro, C. and Ferreira, Ó. (2015), 'Retrieval of nearshore bathymetry from Landsat 8 images: A tool for coastal monitoring in shallow waters', *Remote Sensing of Environment* **159**, 102–116.

Platt, U., Pfeilsticker, K. and Vollmer, M. (2007), *Radiation and Optics in the Atmosphere*, Springer New York, pp. 1165–1203.

Postma, H. (1961), 'Transport and accumulation of suspended matter in the dutch wadden sea', *Netherlands Journal of Sea Research* **1**(1-2), 148–190.

Poursanidis, D., Traganos, D., Reinartz, P. and Chrysoulakis, N. (2019), 'On the use of Sentinel-2 for coastal habitat mapping and satellite-derived bathymetry estimation using downscaled coastal aerosol band', *International Journal of Applied Earth Observation and Geoinformation* **80**, 58–70.

Richter, R. (1990), 'A fast atmospheric correction algorithm applied to Landsat TM images', *REMOTE SENSING* **11**(1), 159–166.

Rosner, B. (1983), 'Percentage points for a generalized ESD many-outlier procedure', *Technometrics* **25**(2), 165–172.

Ruddiman, W. (2000), *Earth's Climate: Past and Future*, W. H. Freeman.

Sagawa, T., Yamashita, Y., Okumura, T. and Yamanokuchi, T. (2019), 'Satellite Derived Bathymetry Using Machine Learning and Multi-Temporal Satellite Images', *Remote Sensing* **11**(10), 1155.

Sha, L. (1989), 'Sand transport patterns in the ebb-tidal delta off Texel Inlet, Wadden Sea, The Netherlands', *Marine Geology* **86**(2-3), 137–154.

Sha, L. (1990), 'Surface sediments and sequence models in the ebb-tidal delta of Texel Inlet, Wadden Sea, The Netherlands', *Sedimentary Geology* **68**(1-2), 125–141.

Stumpf, R. P., Holderied, K. and Sinclair, M. (2003), 'Determination of water depth with high-resolution satellite imagery over variable bottom types', *Limnology and Oceanography* **48**, 547–556.

Thomas, N. M., Pertiwi, A. P., Traganos, D., Lagomasino, D., Poursanidis, D., Moreno, S. G. and Fatoyinbo, T. (2020), 'Space-borne cloud-native satellite-derived bathymetry (SDB) models using ICESat-2 and Sentinel-2', *Earth and Space Science Open Archive ESSOAr* .

Traganos, D., Poursanidis, D., Aggarwal, B., Chrysoulakis, N. and Reinartz, P. (2018), 'Estimating satellite-derived bathymetry (SDB) with the Google Earth Engine and Sentinel-2', *Remote Sensing* **10**(6), 859.

van Straaten, L. M. J. U. (1954), 'Composition and structure of recent marine sediments in the netherlands', *Leidse Geologische Mededelingen* **19**(1), 1–108.

van Veen, J., Van Der SPEK, A. J., Stive, M. J. and Zitman, T. (2005(1950)), 'Ebb and flood channel systems in the Netherlands tidal waters', *Journal of Coastal Research* **21**(6 (216)), 1107–1120.

Verhagen, S. and Teunissen, P. J. (2017), Least-squares estimation and Kalman filtering, *in* 'Springer Handbook of Global Navigation Satellite Systems', Springer, pp. 639–660.

Wang, Z., Hoekstra, P., Burchard, H., Ridderinkhof, H., De Swart, H. and Stive, M. (2012), 'Morphodynamics of the Wadden Sea and its barrier island system', *Ocean & coastal management* **68**, 39–57.