Analysis of deep-sea sulphide mounds using high-resolution MBES bathymetry and backscatter data.

A case study in the Indian Ocean including optimisation of the processing workflow and sulphide mound detection using deep learning.

DISSERTATION

to obtain the academic degree Doktor-Ingenieurin (Dr.-Ing.)

submitted to

HafenCity Universität Hamburg

in the field of

Geodesy and Geoinformatics

by

Tanja Dufek

Hamburg, 2025

Dissertation, HafenCity Universität Hamburg
Geodesy and Geoinformatics

1st Supervisor:	Prof. Dr.-Ing. Harald Sternberg	HafenCity Universität Hamburg
2nd Supervisor:	Prof. Dr. rer. nat. Poerbandono	Institut Teknologi Bandung
Additional Supervisor:	Prof. Dr.-Ing. Thomas Kersten	HafenCity Universität Hamburg

Submitted: January 27, 2025
Date of defense: July 2, 2025

IMPRINT
HafenCity Universität Hamburg
Henning-Voscherau-Platz 1
20457 Hamburg

Dufek, Tanja: *Analysis of deep-sea sulphide mounds using high-resolution MBES bathymetry and backscatter data. A case study in the Indian Ocean including optimisation of the processing workflow and sulphide mound detection using deep learning.*

DOI: 10.34712/142.80

Contents

List of Figures

List of Tables

Abstract

Deep-sea sulphide mounds and associated sulphide deposits form at hydrothermal venting sites, which are commonly known as black smokers. These deposits are of economic interest as potential resources because they contain a high concentration of metals. Current exploration methods focus mainly on active sites, as the plume of discharging hot hydrothermal fluid can be detected as a water column anomaly. However, inactive sites are of greater exploration interest, as they are likely to be older than active sites and therefore their deposits are thought to have accumulated more ore over time.

A workflow is developed in this study to detect and identify sulphide mounds not only at active but also at inactive hydrothermal sites using only bathymetry (2 m resolution) and backscatter (1 m resolution) data collected with a deep towed multibeam echo sounder (MBES) system. The MBES data has a coverage of about 782 km^2 and covers seven hydrothermal fields with 88 sites in the area of the Central and Southeast Indian Ocean Ridges.

The developed workflow consists firstly of improving the MBES bathymetry and backscatter processing workflow to reduce errors and artefacts. Secondly, a convolutional neural network (CNN) is used for segmentation to detect mound structures based on MBES bathymetry derivatives. Finally, the detected mounds are distinguished by their origin (i.e., volcanic or sulphide) based on the combination of MBES backscatter data, MBES bathymetry, and the prediction result obtained by the CNN using deep learning.

The evaluation of this approach shows that the CNN detects about 55% (6,441 mounds) of the manually labelled mounds. As these are too many for exploration, their number needs to be reduced by excluding mounds that are most likely volcanic domes. Depending on the method used, the remaining exploration targets can be reduced to about 43% or 20% of the originally identified possible targets. A comparison with known sulphide occurrences shows that the reduction increases the relative proportion of the sulphide mounds within these exploration targets by up to a factor of three to about 3%.

The results validate existing studies and show that an automated approach can be applied for sulphide mound detection. Furthermore, the deep learning model and the method for distinguishing mound types need to be further improved. In addition, the integration of data from other sensors (e.g., magnetics) is recommended for an increase in the success rate. Anyhow, it is a promising concept for reducing the amount of potential exploration targets to specific areas of greater exploration interest.

Zusammenfassung

Sulfidvorkommen in der Tiefsee sind oft als Sulfidhügel an der Meeresbodenoberfläche erkennbar. Sie entstehen an Austrittsstellen hydrothermaler Fluide und sind allgemein als "Schwarze Raucher" bekannt. Die Sulfidvorkommen sind als zukünftige Quelle metallischer Rohstoffe von Interesse.

Typische Methoden zur Erkundung hydrothermaler Vorkommen basieren auf der Detektion chemischer Anomalien im Wasser, die von austretenden hydrothermalen Fluiden herrühren. Im Allgemeinen sind erloschene Vorkommen für die Exploration von größerem Interesse, da sie in der Regel älter sind und daher mehr Zeit hatten, um Metallsulfide zu akkumulieren.

Im Rahmen dieser Arbeit wird ein Workflow entwickelt, mit dem Sulfidhügel nicht nur von aktiven, sondern auch von erloschenen Hydrothermalquellen detektiert werden können. Als Datenbasis dienen hochaufgelöste Fächerecholotdaten (Bathymetrie 2 m, Rückstreuintensitäten 1 m) eines tiefgeschleppten Systems. Das Untersuchungsgebiet liegt im Bereich des Zentral- sowie Südostindischen Rückens, hat eine Ausdehnung von 782 km^2, und umfasst sieben Hydrothermalfelder mit insgesamt 88 aktiven und erloschenen hydrothermalen Vorkommen.

Der Workflow beinhaltet zunächst eine optimierte Nachbearbeitung der Fächerecholotdaten. Anschließend erfolgt auf Basis der Bathymetrie (und ihrer Ableitungen) eine Segmentierung und damit eine "Hügeldetektion" mit Hilfe eines Convolutional Neural Networks (CNN). Abschließend werden die Ergebnisse des CNN, die Bathymetrie, und die Rückstreuintensitäten verwendet, um die Anzahl der detektierten Hügel zu reduzieren.

Die Untersuchung des entwickelten Ansatzes zeigt, dass das CNN 55% (6.441 Hügel) der zuvor manuell annotierten Hügel richtig segmentiert. Da nicht alle Hügel im Rahmen der Erkundung untersucht werden können, erfolgt eine Reduktion ihrer Anzahl durch Ausschluss der Hügel, die vermutlich vulkanischen Ursprungs sind. Je nach gewählter Methode kann so die Anzahl der Explorationsziele auf 43% bzw. 20% der ursprünglich detektierten Hügel reduziert werden. Der Vergleich mit den bekannten Hydrothermalfeldern zeigt, dass durch diese Reduktion der Anteil der Sulfidhügel an den verbleibenden, für die Exploration als interessant eingestuften Hügeln, erhöht werden kann. Ihr relativer Anteil kann um den Faktor drei auf 3% erhöht werden.

Die Arbeit bestätigt, dass ein automatisierter Prozess zur Erkennung von Sulfidhügeln eingesetzt werden kann. Zur Erhöhung der Erfolgsquote sollte das Deep-Learning-Modell und die Methode zur Unterscheidung der Hügelarten noch weiter optimiert werden. Des Weiteren kann die Klassifizierung durch Daten zusätzlicher Sensoren (zum Beispiel Magnetik) optimiert werden. Insgesamt kann gezeigt werden, dass es sich um einen vielversprechenden Ansatz zur Eingrenzung interessanter Gebiete für weitere Explorationsarbeiten handelt.

Chapter 1

Introduction

Sulphide mounds, or more precisely the associated sulphide deposits, have been identified as potential future resources for precious metals and are increasingly being targeted by deep-sea exploration (Beaulieu et al., 2015; Petersen et al., 2016). Deep-sea exploration relies primarily on hydro-acoustics such as multibeam echo sounders (MBES) as a mapping tool. The advantage of MBES systems lies in their large coverage and large data volumes, making them ideal for the application of artificial intelligence (AI) techniques such as deep learning.

This study investigates the general applicability of existing deep learning approaches for identifying potential locations for hydrothermal venting sites based on multibeam echo sounder bathymetry and backscatter data. Furthermore, this study improves the MBES data processing workflow for the purpose of deep learning assisted sulphide mounds detection, as the validity of deep learning approaches is strongly dependent on the quality of the underlying data.

The study is based on data collected in the exclusive exploration contract area of the BGR (Bundesanstalt für Geowissenschaften und Rohstoffe, Federal Institute for Geosciences and Natural Resources) in the central Indian Ocean.

This thesis is structured in eight chapters, of which the introductory one outlines the motivation for this study, the research gap, and the resulting study objectives. The second chapter presents the basic principles of sulphide mound formation and MBES based seafloor exploration, as well as some basic explanation of AI terminology. The third chapter introduces the case study data, including the instrumentation and exploration methods. Chapter 4 deals with the MBES bathymetry and backscatter processing – from a general outline of the workflow to its improvement for automated sulphide mound detection. The fifth chapter examines the characteristics of different seabed classes, and in particular of areas of hydrothermal activity in MBES backscatter data, for their

usability in seabed classification. Chapter six presents the results of an automated approach for mound detection (using the convolutional neural network (CNN) U-Net) and sulphide mound identification (using GLCM (grey-level co-occurrence matrix) and clustering). The seventh chapter discusses the results in respect to the classical deep-sea exploration approaches and the advent of deep learning assisted data analysis. The last chapter provides some concluding remarks and an outlook on the application of MBES and deep learning technologies in deep-sea exploration for hydrothermal activity.

1.1 Motivation

Sulphide mounds form at hydrothermal venting sites – commonly known as black smokers. These are discharge locations of hot hydrothermal fluid, which can reach temperatures of more than 400° C. Hydrothermal fluids are typically enriched in precious metals such as copper, zinc, and gold (Hannington et al., 2005). Where the hot hydrothermal fluids mix with cold seawater, the metals precipitate as sulphides (Petersen et al., 2016). These hydrothermally formed sulphides are either lost to the water column, in a so-called hydrothermal plume (from which they may settle as metalliferous sediments), or the sulphides accumulate in deposits at or below the seafloor. These deposits can grow to sizes of economic interest and are then referred to as seafloor massive sulphides (SMS).

Sulphide mounds form when hydrothermal venting is long lasting and the geological setting supports sustained SMS accumulation (German et al., 2016; Klischies, 2021). The types and sizes of SMS depend on several factors, including the geological setting, the activity of hydrothermal venting, and the preservation of the SMS over time. The first hydrothermal site was discovered less than 50 years ago (Corliss et al., 1979), as these sites typically occur in deep-sea areas, and the associated sulphide mounds rarely exceed 150 m in diameter. Due to their relatively small surface dimensions, their detection, localisation, and identification is challenging and requires a variety of specialised equipment.

Sulphide mounds associated with inactive hydrothermal activity are the preferred exploration target from an economic perspective (Jamieson et al., 2014; Koschinsky et al., 2018). Such inactive vent systems are typically assumed to be older than active ones and the potential for larger associated sulphide deposits is therefore higher (Murton et al., 2019). Inactive sites are also abandoned by the unique vent fauna and lack the typical dense communities of shrimps, snails, anemones, and other species found at active sites (SPC, 2013).

Exploration methods for inactive sulphide mounds have not yet been refined to the point where they can be used reliably and economically. In contrast, active hydrothermal vent sites are detected using classical exploration methods that search for the hydrothermal plume in the water column (Baker, 2017). This introduces a bias into the discovery of sulphide mounds and hinders a deeper understanding of the underlying processes and a comprehensive assessment of the resource potential of SMS (Hannington et al., 2015).

Exploration techniques are constantly evolving, with **MBES** playing a crucial role in locating new discharge location and mapping the surrounding geological setting (German et al., 2008). The first echo sounders were developed over a century ago (Grządziel and Wąż, 2018). Since then, the technology has been continuously improved, resulting in the availability of various hydro-acoustic sensors in the hydrographic field, including MBES or side scan sonars (Lurton, 2010). Ongoing developments in instrumentations have broadened the range of applications and have massively increased the quality and volume of hydro-acoustic data.

Today, MBES systems are the preferred tool for deep-sea exploration because they can remotely acquire seafloor bathymetry and backscatter information over large areas in a relatively short amount of time. They are standard equipment on research vessels and can be mounted on a variety of other platforms, including towed and autonomous underwater vehicles. However, hydro-acoustic data processing and analysis still remains a time-consuming and often manual task. Dependencies in data acquisition are complex, and obtaining reliable and accurate underwater positioning is a challenge.

With the advent of AI, machine and **deep learning** approaches have been successfully applied in the maritime domain and in hydro-acoustic data processing and analysis. Typical tasks in which they are applied are feature extraction, segmentation, and classification (Steiniger et al., 2022; Hinz et al., 2024). They are able to analyse large data sets in a short time and detect patterns or relationships within them that exceed human capabilities.

This study aims to improve exploration methods for inactive hydrothermal sites to reduce the bias in sulphide mound discoveries and to work towards a more realistic resource assessment of the deep seafloor. The work uses high-resolution MBES data, as it is common in seafloor exploration, to establish a data processing workflow for the useful application of deep learning and other AI approaches.

The data for this study has been collected over the last decade within the INDEX project (Indian Ocean Exploration) led by BGR. The data set comprises large volumes (covering 782 km^2) of high-resolution (2 m) MBES data from 88 active and inactive hydrothermal sites, allowing the application of AI methods

like deep learning. In the following sections, the research gap, the subsequent research questions, and the derived objectives of this thesis are outlined in greater detail.

1.2 Research Gap

The primary objective of this study is to contribute filling the gap in the field of inactive hydrothermal site exploration using high-resolution MBES data – bathymetry and backscatter – and techniques from the field of deep learning. In this section, the research gap within the context of the thesis is described in greater detail.

One exploration approach that has been successfully applied to detect inactive sites, involves the measurement of anomalies in the natural electric potential and thus the geoelectrical field. Such anomalies are caused be reactions like oxidation-reduction processes of sulphide deposits with their surroundings (Sato and Mooney, 1960; Sharma, 1997; Constable et al., 2016). These methods have the advantage of detecting not only inactive but also buried sites that are not visible on the seabed. The disadvantage of them is that these instruments utilised for electromagnetic mapping must be towed in very close proximity to the seabed (about 2 m to 50 m) at a constant altitude (Gehrmann et al., 2019; Müller et al., 2023). This can be challenging in highly morphologically structured environments like close to mid-ocean ridges as within the INDEX project. Additionally, the coverage of these techniques is limited and therefore small profile spacing is necessary. Consequently, such methods are typically employed in areas with known deposits to estimate their size or detect new ones within a specific area. Deployments over larger areas are not feasible.

MBES data has a comparable large coverage to other geophysical methods. Its data is therefore used for mapping and investigations of larger areas. Its backscatter data is often used for habitat mapping – typically in combination with bathymetry (Parnum et al., 2004; Che Hasan et al., 2014). In the context of hydrothermal venting, it can also aid geological mapping and interpretation (Graber et al., 2020). Data from ship-borne MBES with lower resolution is used for the interpretation of large-scale features of seafloor morphology (Eason et al., 2016; Klischies, 2021). The high-resolution data from submerged platforms is typically employed for general interpretation of structures' origins or for estimation of the extent of specific seabed covers such as sediment or pillow lavas (Ondréas et al., 2018).

High-resolution bathymetry and backscatter data is available from 13 hydrothermal fields within the INDEX project. Out of these, seven larger and well mapped

fields have been selected for this study. Their data is used for an investigation of the feasibility of using high-resolution MBES data for the site detection with focus on inactive sites. Such an analysis includes not only the visual backscatter interpretation, but also an automated approach that provides an estimate of possible hydrothermal site locations.

In recent years, machine and deep learning approaches have been integrated into the hydrographic field. Applications range from object detection such as boulders, ammunition, or ship wrecks (Feldens et al., 2019; McKay et al., 2017; Yu et al., 2021; Tang et al., 2023) to seabed classification and habitat mapping (Marsh and Brown, 2009; Alevizos et al., 2018; Qin et al., 2021). The integrating of deep learning into marine morphological feature detection and interpretation makes sense because a manual approach is often very time-consuming and the results can be subjective and lack consistency (Hinz et al., 2024).

Deep learning has also been applied in the field of hydrothermal venting, such as vent substrate classification (Vega et al., 2024) or plume detection in MBES water column data (Mimura et al., 2023) as well as in ROV video imagery for sampling (Wang et al., 2023). Regarding the above addressed shortcomings of inactive site detection methods, just recently, deep learning has also been successfully utilised for sulphide site prediction based on the geomorphological properties (Juliani and Juliani, 2021; Keohane and White, 2022; Haroon et al., 2023). The first two of these approaches are based solely on high-resolution bathymetric data, whereas the latest one listed also uses data from other instruments (e.g., magnetic) additionally. They use deep learning to identify mound or chimney structures in the data, and further classification distinguishes sulphide mounds from volcanic mounds or identifies hydrothermal areas.

These published approaches using deep learning for hydrothermal site detection have to deal with the general limitation of availability of high-resolution MBES and ground truth data. To develop robust automated solutions, a large amount of ground truthed training data is required (Feldens et al., 2021). Such data collection by submerged MBES and the ground truthing by video imagery in the deep sea is very time consuming and expensive. As a result, the available data is strongly limited. The number of confirmed mound/site locations used as training data in the named publications is therefore limited (ranging between two and 48 sulphide mounds / chimneys). The INDEX data can significantly contribute to the improvement of a deep learning model, having a coverage of 782 km^2 and including 88 sites with more than 200 confirmed hydrothermal mounds. In addition, the high-resolution MBES backscatter data will be utilised within the workflow to identify potential sulphide mounds, which has not been done yet in other publications. This might improve the overall results for sulphide mound prediction.

Research questions and objectives can be derived from the research gap concerning the shortcomings of exploration techniques for inactive vent sites. These questions and objectives are described in the following section, and it is outlined how they contribute to close the research gap.

1.3 Research Objectives and Research Questions

The goal to contribute to the exploration scheme for (inactive) hydrothermal sites using high-resolution MBES data and deep learning approaches, starts first and foremost with a closer inspection of the data quality. The MBES data constitutes the basis. It or its derivatives are used for quantitative and deep learning approaches. In particular, errors and residual artefacts in the data can negatively affect the results of automated detection methods. This leads to the first research objective (RO):

RO1: Improve the quality of the MBES data so that it can be used for automated detection and classification.

The data used in this study was collected with a submerged sensor platform (i.e., deep-towed sled) at an altitude of approximately 100 m above the seafloor in about 3,000 m of water depth. The derived raster resolution of the terrain model is about 2 m, which is better than the navigation solution of about 5 to 10 m in the ideal case and up to 50 m in the worst case. The navigation accuracy varies during a deployment, as it depends on the distance between vessel and sensor platform, which varies during a survey. Conventional commercial software is limited in its ability to correct for varying offsets. Since the focus in this study is the detection of small feature – hydrothermal venting sites and their associated sulphide mounds – offsets larger than the data resolution can distort the morphology and thus hinder the automated detection of specific structures. This yields the first research question (RQ):

RQ1.1: How can the high-resolution MBES bathymetry data be processed to minimise the effects of insufficient navigation accuracy?

The improved navigation solution benefits not only the bathymetry, but also the quality of the co-registered backscatter data. Typically, the acquisition settings for backscatter data are not optimised because the focus is usually on bathymetry. Therefore, gain variation, pulse form and shape, or other automated setting adjustments to improve the bottom detection and therefore the bathymetric data, can result in artefacts in the backscatter data. To date, backscatter data acquisition and processing have a rather black-box character, with certain corrections being applied automatically in the background during acquisition or by commercial processing software (Lurton and Lamarche, 2015). Even when using similar tools and settings, different software provides

different results (Schimel et al., 2018; Malik et al., 2019). Since the derived backscatter mosaics will be used for automated detection, the focus is on reducing visible data artefacts as this is likely to yield more reliable segmentation and classification results (Fakiris et al., 2012). This leads to the next research questions:

RQ1.2: How does different processing software and varying parameter settings affect the MBES backscatter data?

Typically, applications using backscatter data for habitat mapping are utilised in shallow water environments, such as coastal areas, where the bathymetry is flat and shows only smooth depth variations. Deep-sea applications typically use ship-borne MBES, which have a large coverage but low spatial resolution due to the great water depth. The data used in this study has been collected in a deep-sea environment, but with a high spatial resolution. The study area is located at a mid-ocean ridge and therefore at a spreading centre where new seafloor is being created and not much sediment cover is present. The tectonic and volcanic activity results in a morphologically harsh seafloor. These circumstances constitute a particular challenge for the backscatter processing, more precisely for modelling the MBES acquisition geometry and its correct consideration in the intensity adjustments (Malik et al., 2018). Certain default corrections and parameter settings in a processing software may not apply to the INDEX data. This leads to the next research question:

RQ1.3: Which processing workflow is best suited to process the high-resolution MBES backscatter data from the INDEX project area?

Before the optimised backscatter data results can be used in an automated approach for sulphide mound detection, it is necessary to understand how the environment of the study area is represented in the intensity mosaics. It is of particular interest, whether the various seabed classes such as sediment or pillow lava can be acoustically differentiated. Such analysis for seabed characteristics and classification using acoustic backscatter information has been performed for decades. Earlier on, mainly side scan sonars were used (Rusby and Somers, 1977; Morris et al., 1978). But as MBES technology improved over time, so did MBES backscatter data. It has the advantage of being recorded together with bathymetric information. Over time, the use of backscatter has been extended to various fields where information about the seabed properties are required as in geology or environmental studies. In the second objective, the MBES backscatter data is used to obtain information about the seabed characteristics:

RO2: Analyse the acoustic characteristics of the seabed classes in the study area.

There are various approaches for classifying the seabed using MBES backscatter data. Besides amplitude interpretation, the most common are texture analysis and angular dependence (Parnum et al., 2004). The seafloor in the study area is characterised by rugged morphology and limited sediment cover due to its close vicinity to the spreading centre. This poses a challenge in data processing when it comes to data corrections. The modelling of the local acquisition geometry is of great importance to apply particular adjustments correctly. Not all possible methods are suitable for the study area, the instrumentation used, and the data acquisition settings. This leads to the next research question:

RQ2.1: How can the seabed classes of the study area be differentiated in high-resolution MBES backscatter data?

With respect to hydrothermal environments, low-resolution backscatter data from ship-borne systems is often used for large-scale analysis of morphological features (Gràcia et al., 2000; Eason et al., 2016). High-resolution intensity data sets used in literature for site identification are mostly of different instrument types such as single-beam echo sounder (Durand et al., 2006) or side-scan sonar (Asada et al., 2015). Not as much literature is available on the use of high-resolution MBES backscatter data for site identification as Ondréas et al. (2018) in the Guaymas basin, Gulf of California. Typically, the data is mostly used in combination with other data sets as a further data layer with additional seabed information. As MBES backscatter data will be tested in this work for its usability in sulphide mound detection, it is of particular interest whether hydrothermal areas (active and inactive) have a unique and prominent backscatter signature:

RQ2.2: Are there certain backscatter characteristics that are unique to areas of hydrothermal activity?

In case such unique backscatter characteristics can be identified for hydrothermal sites, this information can be used for site detection and identification. In addition, hydrothermal sites are also associated with sulphide mounds, which exhibit a very specific morphological shape with steep slopes and conical tops (Jamieson et al., 2014). This distinguishes them from mounds with volcanic origin. In habitat mapping or landform classification, the combination of bathymetry and backscatter is often employed successfully as in Masetti et al. (2018) for example. The application of such combined approaches using distinctive bathymetric and backscatter characteristics of hydrothermal sites and deep learning is the third objective of this thesis:

RO3: Automate the process of predicting sulphide site locations based on high-resolution MBES data.

An automation of segmentation or classification is in general pursued as it reduces the time required, but also produces more objective and therefore reproducible results (Parnum et al., 2004). Such approaches based on the shape (morphology) and backscatter characteristics typically utilise tools that rely on data derivative from the terrain (i.e., slope, aspect, curvature, variability, bathymetric position index) (Lecours et al., 2016) or texture metrics derived from the grey-level co-occurrence matrix (GLCM).

By combining information from multiple types and sources (derived from bathymetry but also from backscatter data), conclusions on marine landform, the substrates, and marine habitats are possible as well as data segmentation and classification (Wilson et al., 2007; Janowski et al., 2018). It allows even further analysis of the local ecology as part of habitat mapping.

The approaches found in the literature vary strongly in terms of algorithms, type of attribution, and degree of automation using machine learning techniques (Diesing et al., 2014; Lecours et al., 2016). Especially in the last decade, with the improvement of computational power and the handling of large amounts of data, the utilisation of deep learning methods in marine applications has increased (Steiniger et al., 2022; Arosio et al., 2023). In the recent years, deep-learning approaches have been successfully applied as part of an automated prediction of sulphide deposits (Juliani and Juliani, 2021; Haroon et al., 2023). These approaches use the convolutional neural network U-Net to semantically segment mounds from background data. This leads to the research question:

RQ3.1: How successful is a mound prediction using U-Net for the data of the INDEX project?

The available U-Net model from Juliani and Juliani (2021) was trained with data from the Atlantic Mid-Ocean Ridge. It is of interest to investigate how well this model performs on the INDEX data and thus to what extent such models are universally valid for data of the same context. Since the availability of high-resolution deep-sea data is generally low, the training data set and especially the ground truth information is limited. Within the INDEX project, large amounts of data have been collected over the last years, which can be used for to train and evaluate a new model. Its results are compared to the ones of the pre-trained model. After obtaining predictions for mound locations, these are differentiated between sulphide and non-sulphide mounds, yielding the following research question:

RQ3.2: Can MBES backscatter data facilitate in distinguishing sulphide from volcanic mounds and thereby enable a reduction of possible exploration targets?

Verification of the origin of mounds is usually done using video imagery. As ship time is limited and expensive, a general reduction of exploration targets is necessary as the origin of all mapped mounds cannot be verified by video imagery. An automated assessment of the mounds and their likelihood of being associated with hydrothermal activity helps to reduce the number of possible targets and thereby guide the exploration. Areas with larger accumulations of AI-verified possible targets of interest can then be further explored by visual inspection or other geophysical techniques with limited coverage.

By answering the above listed research questions, a step towards an automated sulphide detection method that includes also inactive hydrothermal sites based solely on MBES data can be achieved. This will help to significantly reduce the research gap.

Chapter 2

Basic Principles

This chapter provides details on the general principles used in this study. This includes – first of all – some geological basics necessary for the understanding of hydrothermal venting and seafloor massive sulphides (SMS), as these are the investigation target of this thesis (Chapter 2.1). The chapter provides background on their formation, explains why they are of general interest, and gives an introduction to the environment in which they are found. Secondly, the fundamental techniques, geometry, and terminology of multibeam echo sounding are explained, as this instrument is the main data source for this work (Chapter 2.2). Finally, the various AI terms are differentiated and explained (Chapter 2.3). Since this work is aimed at readers with a geodetic/hydrographic background, the geological basics start very general, while certain terms and instrumentation within the field of seafloor mapping are assumed to be more familiar.

2.1 Geological Principles

Sulphide mounds form through geological processes that are controlled by Earth's structure and driven by plate tectonics, which are introduced in section 2.1.1. The metal content of sulphide mounds and therefore their economic value depend on the mineral composition (2.1.2). The mineral composition depends on the surrounding rocks (2.1.3), in which the hydrothermal circulation takes place (2.1.4).

2.1.1 Structure of the Earth and Plate Tectonics

The Earth's interior can be divided into different layers based on changes in density and chemical composition. In general, the density increases towards the Earth's centre as a result of the elevated pressure. In addition, abrupt density changes can occur due to variations in chemical composition, which

can create layer boundaries. The Earth's structure comprises three principle layers: the core (divided into an inner and outer section), the mantle, and the crust (Figure 2.1). The inner core has an approximate radius of 1,220 km and consists of solid material due to the high pressure. It is composed of an iron-nickel alloy and the estimated temperature varies from 3,500° C to 5,000° C in the centre. The outer core has a slightly lower density and consists not completely of iron and nickel, but to about 10% of oxygen and sulphur. Its structure is more fluid and it has a thickness of about 2,260 km (Grotzinger and Jordan, 2020).

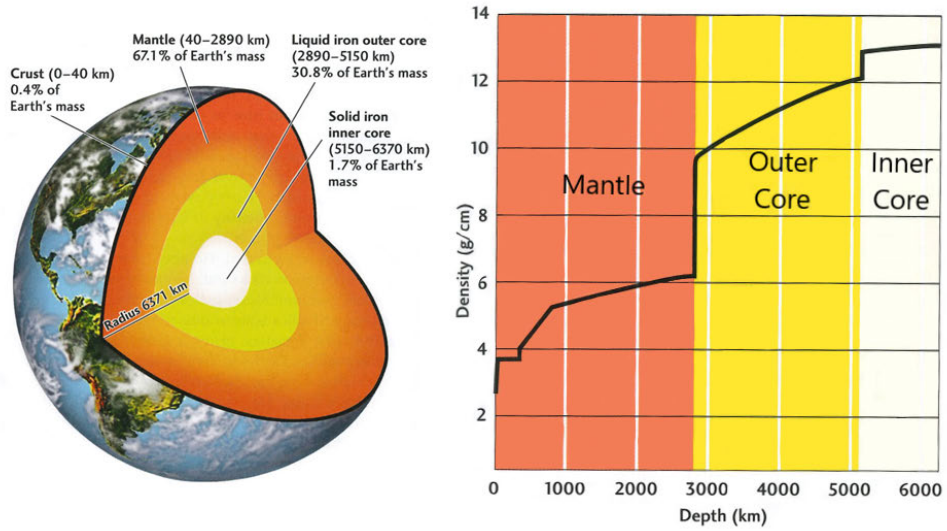


Figure 2.1: Left: Scheme of Earth's structure, right: density of the Earth's layers (Grotzinger and Jordan, 2020)

The solid mantle makes up the largest part of the Earth, accounting for more than two thirds of its mass. It has a lower density than the core and consists of silicate rock (mainly oxygen, magnesium, silicon but also calcium, aluminium, and iron). It is about 2,850 km thick and can be further differentiated based on density and chemical changes: the upper mantle (thickness about 410 km) with a lower density, the transition zone of about 250 km with continuously increasing density, and the lower mantle with higher density (Grotzinger and Jordan, 2020) (Figure 2.1, right). The crust consists of silicates (see also Chapter 2.1.2) of lower density and therefore "floats" on the denser mantle. It has a higher content of aluminium and potassium than the mantle, which in contrary has more magnesium and iron. The boundary between crust and mantle is called the Mohorovičić-Discontinuity, also abbreviated as Moho. There are two types of crust: continental and oceanic crust, which differ in general thickness and composition. The oceanic crust is in average about 7 km thick and is therefore thinner than the continental crust of approximately 40 km (Searle, 2014).

There is an alternative approach to divide the Earth into different layers not by its composition, but by solidity or mechanical properties. The outermost unit

is the lithosphere, which is more rigid due to cooling processes. Its thickness is in average 100 km, but varies strongly from only a few kilometres in volcanically active areas to over one hundred kilometres in older and cooler regions (Searle, 2014). It consists of the crust and parts of the upper mantle. Below the lithosphere is the more ductile asthenosphere. It is about 300 km thick and consists of the successive upper mantle. The lithosphere is made up of 13 larger and several smaller plates that move on top of the also moving asthenosphere at a rate of a few centimetres per year. This movement is caused by the heat within the inner part of the Earth: hot material from the mantle is ascending where plates move apart (ridge push). As they move away from each other, the lithosphere cools down and behaves rigidly and descends at borders where plates converge again due to gravity (slab pull) (Figure 2.2). This process of hot material ascending and cooler material descending is called convection (Grotzinger and Jordan, 2020).

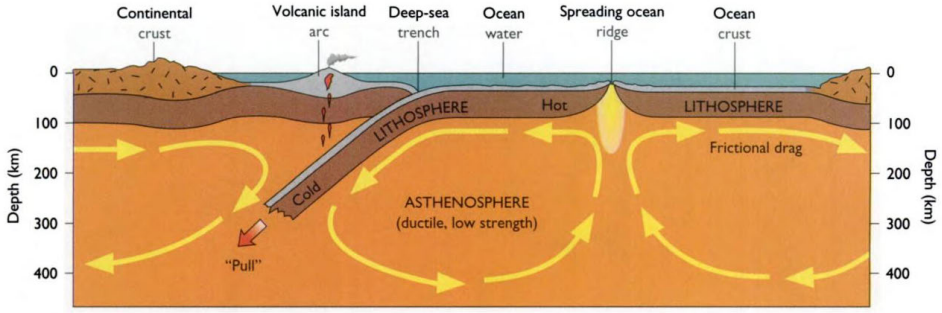


Figure 2.2: Overview of divergent and convergent plate boundary and the convection processes (Pinet, 2003).

Plate boundaries are differentiated by the direction of the plates' relative movement to each other. At **divergent plate boundaries** (constructive boundaries), the plates separate and move away from each other. This process is caused by convection as ascending hot mantle material increases pressure on the bottom of the lithosphere (Figure 2.2). The lithosphere is fractured and the two plates move away from each other, carried by the underlying convection current (Erickson, 2003). Hot melt ascends into magma chambers where it continues to rise via fissures. The melt creates new lithosphere as it solidifies in the ground (intrusive) or on the seafloor (extrusive) (Searle, 2014). Such constructive plate boundaries are also referred to as axis of seafloor spreading. They are visible in the bathymetry as mid-ocean ridges (MOR) and have a cumulative length of 65,000 km (Searle, 2014). The spreading rates of a MOR are differentiated into

- superfast (>13-15 cm/yr)
- fast (9-13 cm/yr)
- intermediate (5-9 cm/yr)

- slow (2-5 cm/yr)
- ultra-slow (< 2 cm/yr)

after Searle (2014). Fast-spreading ridges make up to about 25% of the total length of ocean ridges. They have a relatively thin crust and volcanic eruptions occur often. Slow-spreading ridges are more common (about 60%). There, the lithosphere is generally thicker and volcanic activity is lower (Hannington et al., 2005). The different properties of the spreading centres lead to a difference in appearance of their ridges within the bathymetry. Fast-spreading ridges show a axial high, whereas at slow-spreading ridges usually have a rift valley about 1-2 km deep and 20-30 km wide (Buck et al., 2005).

At **convergent plate boundaries** (destructive boundaries) plates move toward each other and collision or subduction occurs. Subduction usually takes place when two oceanic or oceanic and continental plates converge. The plate with the higher density is pushed below the other. The lithosphere of the descending plate immerses into the asthenosphere and melts. Such subduction zones can create trenches at the Earth's surface (Figure 2.2) (Grotzinger and Jordan, 2020).

The third type of plate boundaries are **transform faults** (conservative plate boundaries, transformation faults). At these faults, plates move past each other in opposite directions. Only horizontal movement occurs and no lithosphere is created or destroyed. Transform faults interrupt the linear MOR and divide it into segments - the spreading centres. Such segments can have lengths of up to several hundred kilometres, widths of about 60 km, and offsets between them of 30 km to a maximum of 1,000 km (Erickson, 2003; Searle, 2014). Each of them has their own volcanic source (Erickson, 2003). The extension of transform faults beyond the spreading axes is called fracture zones, where oceanic crust of different ages moves in the same direction (Grotzinger and Jordan, 2020) Smaller faults with lengths of less than 30 km that are not continuous are referred to as non-transform offsets or discontinuities (Searle, 2014). Figure 2.3 shows the INDEX project area along the Central and Southeast Indian Ridge with marked plate boundaries.

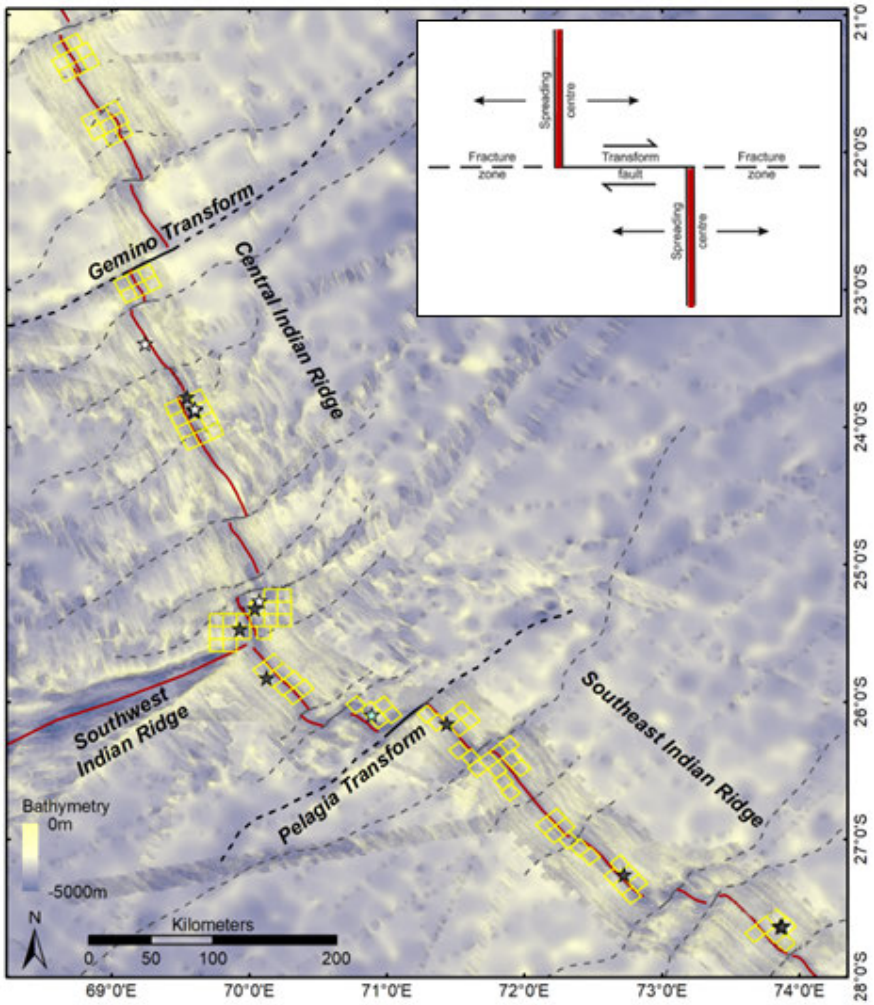


Figure 2.3: Bathymetric map of the Central and Southeast Indian Ridge showing the INDEX project area (contract blocks in yellow). The ridge axis /spreading centres are highlighted by a red line, the transform faults as continuous line, and their continuation as fracture zone by a dashed line. Locations of hydrothermal venting are marked by stars (Klischies, 2021). The schematic depiction in the top right corner shows the general moving direction of the tectonic plates at the boundaries (after Searle (2014)).

2.1.2 Minerals

Atoms consist of the nucleus, which is comprised of positively charged protons and neutral neutrons. The negatively charged electrons circle around the nucleus. Protons and neutrons have a mass number of 1, and since the mass of an electron is only a small fraction of that, most of the mass is in the nucleus. The number of protons is given by the atomic number of an element. Atoms have an equal number of protons and electrons and are therefore electrically neutral. But the number of neutrons, and therefore their atomic mass, can vary for atom types of the same element. They are called isotopes (Johnson et al., 2019).

Rock consists of minerals, which are compounds of atoms. Minerals occur naturally, are inorganic, solid, and have a defined chemical compositions and a regular crystal structure (Johnson et al., 2019). Most minerals are compounds

of more than one element through ionic or covalent bonding. For an ionic bond, an electron is passed from one atom to the other. Both atoms are no longer neutral, but developed into ions with opposite charges: a cation (positively charged) and an anion (negatively charged). The bond is formed based on their electrostatic attraction due to their opposite charges. Ionic bonding is the main bonding type for minerals. When atoms share electrons, it is called covalent bonding. In this bonding, no ions are created. Also a mixture of both types of bonding are possible (Johnson et al., 2019; Grotzinger and Jordan, 2020).

Due to chemical bonding, minerals have a crystal structure in which atoms and molecules are arranged homogeneously in a three-dimensional structure. Cations are smaller than anions and can be easily exchanged. Therefore, rock minerals are classified according to their anions. Even though there are thousands of known minerals, most are not common and there are only about 30 in general that are referred to as rock forming minerals. These form the Earth's crust, which consists mainly of only nine elements (see Chapter 2.1.1). The most important classes of rock forming minerals are listed in Table 2.1 (Grotzinger and Jordan, 2020).

Table 2.1: Some of the mineral classes of mineral differentiation based on their composition (Grotzinger and Jordan, 2020).

Class	Defining Anions	Example
Native elements	None: no charged ions	Copper metal (Cu)
Oxides	Oxygen ion (O^{2-})	Hematite (Fe_2O_3)
Halides	Chloride (Cl^-) Fluoride (F^-) Bromide (Br^-) Iodide (I^-)	Halite (NaCl)
Carbonates	Carbonate ions ($CO_3{}^{2-}$)	Calcite ($CaCO_3$)
Sulfates	Sulfate ion ($SO_4{}^{2-}$)	Anhydrite ($CaSO_4$)
Silicates	Silicate ion ($SiO_4{}^{4-}$)	Olivine $(Mg, Fe)_2SiO_4$
Sulphides	Suphfide ion (S^{2-})	Pyrite (FeS^2)

Ore minerals are minerals containing metals that can be extracted from them and are therefore an important industrial resource. Besides oxides and silicates, sulphides are the largest group of ore minerals. As mentioned above, the elements usually occur in rocks in a concentration similar to that found in the Earth's crust. Accumulations of ore minerals can be formed by enrichment processes. They become interesting for exploitation when reaching a certain size and metal concentration. Such ore deposits can occur, for example, in hydrothermal active areas (Grotzinger and Jordan, 2020).

Sulphide mounds consist of SMS, hence ore minerals, but mainly consit of

silicates (e-g-, pure silica), sulphates (e.g., gypsum), and oxides (e.g., iron hydroxide) (Hannington et al., 1995). The mineral composition of SMS varies, but is usually dominated by pyrite (FeS). The economic value of a SMS occurrences depends on its content in chalcopyrite (CuS), sphalerite (ZnS), galena (PbS), other metal sulphides, the trace metal content of the ore mineral (metal atoms captured within crystal structures) and inclusions of native elements (e.g., gold).

2.1.3 Igneous Rocks at Mid-Ocean Ridges

Rocks are classified into three main groups based on its formation: igneous (magmatic) rocks, sedimentary rocks, and metamorphic rocks. **Igneous rocks** are formed by the crystallisation/solidification of molten material, which can be magma (in the crust) or lava (on the surface). Over time, rocks on the Earth's surface undergo weathering, and small fragments break off as sediment and are transported away (eroded) by wind, water currents, gravity (e.g., landslides), or ice (Grotzinger and Jordan, 2020). Sediments accumulate when the transportation force decreases and are deposited in layers over time. Oceanic sediments are generally composed not only of mineral particles of terrigenous origin (clay), but also of organic sediments, consisting mostly of calcareous or siliceous ooze from phytoplankton and zooplankton shells. This biogenic material dissolves below the carbone-compensation depth (approx. 4,500 m) due to the high pressure and low temperature (Searle, 2014). Pelagic sediments therefore vary in composition depending on the water depth and the biological activity of the region. Sedimentation rates vary between a few centimetres per thousand years for oozes and even less for clays. Regardless of its composition, the further away from the MOR and the older the oceanic crust, the more sediment accumulates on top of it in layers. The older the sediment, the more it is buried and compacted by the increasing weight of newer overlying sediment. Cementation of mineral or organic material creates **sedimentary rocks**. When the texture or chemical composition of igneous or sedimentary rocks is altered by temperature, pressure, or fluids, they are called **metamorphic rocks** (Johnson et al., 2019). This can then again undergo melting and subsequent crystallization to form igneous rocks. The processes described above occur continuously and repeatedly and form the so-called rock cycle.

The solidification process of magma or lava occurs gradually and at different rates as it cools down. Different minerals crystallize at varying temperatures. This can result in a change of the chemical composition of the material. The speed of cooling plays a major role in the texture of arising igneous rocks. Lava usually cools down quickly on the surface, forming only small crystals that are barely recognizable to the eye. Such rocks with small crystals is called extrusive or volcanic rocks (e.g., basalt). It may have small holes due to gas

inclusions during the solidification process. If the cooling process happens very quickly before crystallisation can occur, volcanic glass is formed. The cooling process of magma in the ground is generally slower, allowing more time for crystal formation. The rocks are therefore coarser grained and is referred to as intrusive or plutonic igneous rocks (e.g., gabbro, granite) (Johnson et al., 2019). Basalt is the most common igneous rocks in the Earth's crust. It has the same composition as gabbro, but differs in texture. Figure 2.4 shows examples of basalt, gabbro, and granite.

Figure 2.4: Extrusive and intrusive igneous type of rocks. From left to right: basalt, gabbro, and granite (Grotzinger and Jordan, 2020).

Like solidification, melting occurs at different temperatures for different material, and is often gradual. At seafloor spreading centres, large magma chambers of several cubic kilometres form. At slow-spreading ridges, such magma lenses are generally deeper in the ground than at fast-spreading ridges, where they can be found even as shallow as 1 km to 3 km below the surface (Searle, 2014; Hannington et al., 2005). Magma chambers are formed by decompression melting: The separation of tectonic plates creates a low pressure zone and thus convection forces that facilitate the rising of mantle material. The melting process is further facilitated by the low pressure environment, as this leads to a reduction in melting temperature and increase of buoyancy.

The upper mantle consists mostly of peridotite, a dense and coarse-grained igneous rocks. Since not all of the material is completely melted, the more liquid rocks are separated from the remaining crystal mush. As shown in Figure 2.5, a layer of peridotite forms in the upper mantle below the Moho and at the bottom of the magma chamber. Overall, the magma influences the development of three different layers: [1] Some magma ascends to the surface where the plates diverge and forms basaltic pillow lava, [2] some magma solidifies during the rising process in the vertical cracks before reaching the surface and creates sheeted dikes of gabbro, and [3] some magma solidifies into a gabbro layer in vicinity to the magma chamber. These form the typical layers of the oceanic plate: pillow lava, sheeted dikes (approx. 1 km to 2 km), and massive gabbros (approx. 4 km to 6 km) (Grotzinger and Jordan, 2020; Seibold and Berger, 2017). They are typically covered by sediments when older and in closer vicinity to land. In the area of spreading centres with young crusts, the

sediment cover is typically very thin or even completely missing.

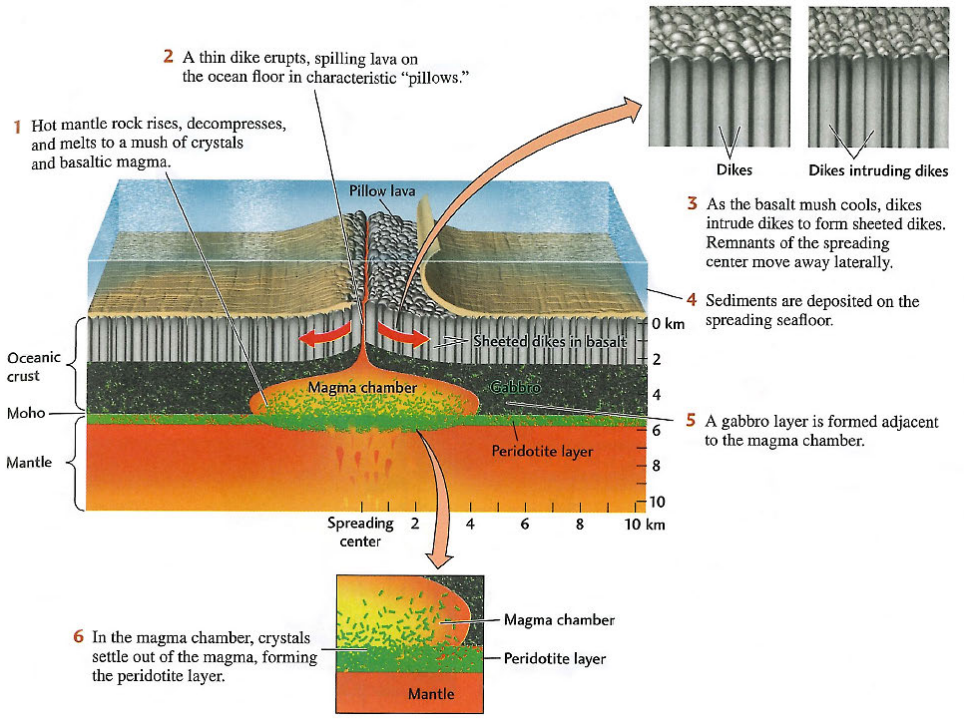


Figure 2.5: Decompression melting at seafloor spreading centres at MORs (Grotzinger and Jordan, 2020).

The most common seafloor cover in the area of slow-spreading centres are shown in Figure 2.6. Pillow lava obtained its name from its distinctive shape (Figure 2.6a). When lava erupts on the seafloor, it quickly solidifies on the outside when it comes in contact with cold water. The inside is still liquid and continues to flow. Glass can often be found on the surface of pillow lava. In general, they are not very stable and fragment easily (Figure 2.6b). Larger accumulations of broken rock material with an irregular and sharp shape accumulate at the bottom of steep slopes. They are called talus (Figure 2.6c).

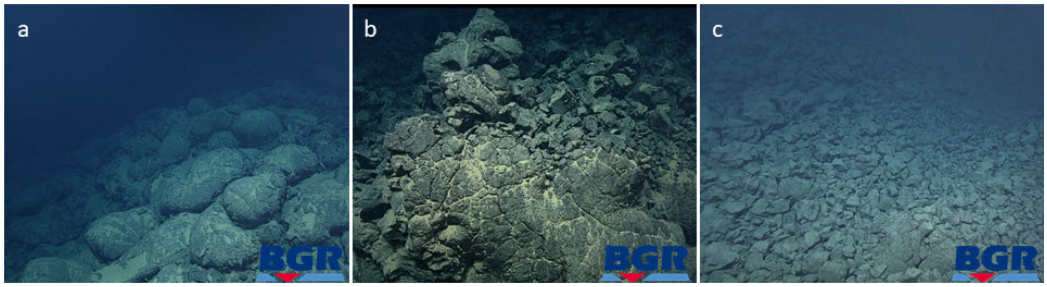


Figure 2.6: Different seafloor types typical for areas of slow-spreading ridges: a) pillow lava, b) fragmented pillow lava, c) talus.

2.1.4 Hydrothermal Fields at Mid-Ocean Ridges

Discharge sites of hydrothermal fluids are mostly found in tectonically active regions. About 65% of the known sites are located at MORs (Hannington et al., 2005). Due to the strong forces of seafloor spreading, these areas have a greater number of faults and fractures that allow seawater to penetrate up to several kilometres into the oceanic crust. At slow-spreading ridges the hydrothermal fluid descends deeper (5 km to 8 km) than at fast-spreading ridges (1 km to 2 km) because the depth of the magma chamber varies (Hannington et al., 2005). During the percolation, the water is heated by the underlying magma to over 400° C. In this way, seawater is transformed into a hot, reducing, and acidic fluid that dissolves metal from surrounding rocks, which enriches the hydrothermal fluid (Figure 2.7).

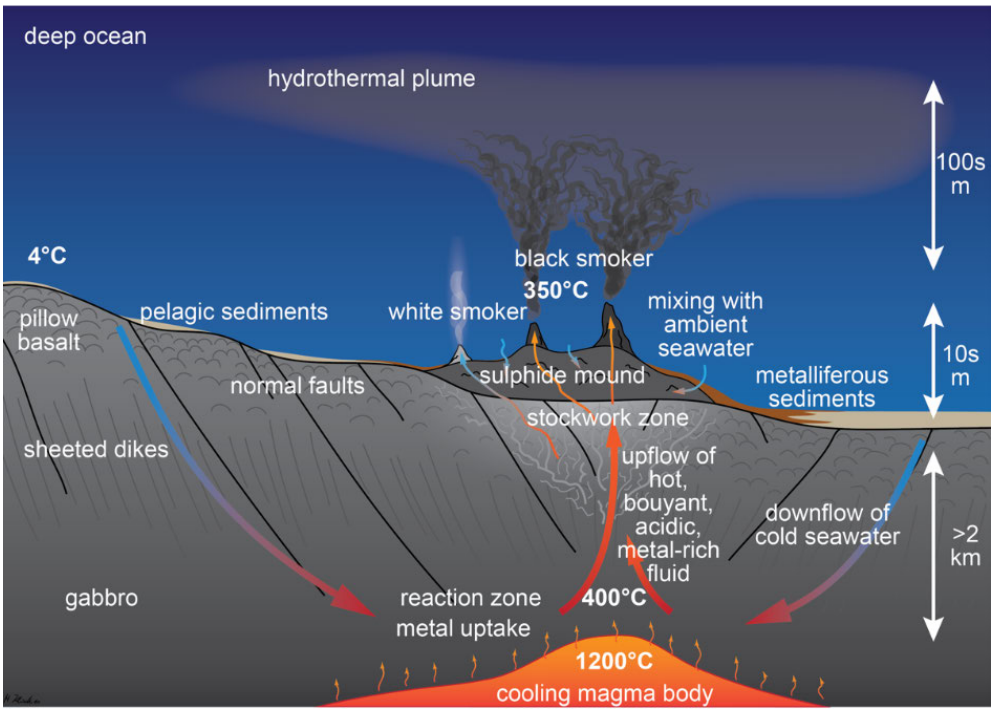


Figure 2.7: Scheme showing the principle of hydrothermal circulation and seafloor massive sulphide deposit creation (Klischies, 2021).

At a certain point, the fluid density has decreased and it begins ascending again until it is emitted from the seafloor as plumes of hot hydrothermal fluid. When the hot reducing fluid comes into contact with cold, oxic seawater, it suddenly cools down and precipitates the dissolved metals as sulphides. Some of this process occurs already within the stockwork zone. A portion of the sulphides will form small particles (1 µm to 3 µm) that are discharged into the water column as black plumes. Such plumes of hydrothermal vents are often black in colour and are therefore also called "black smokers". Depending on the chemical composition and temperature, the fluids may also be grey, white, or transparent, and the venting may be more diffuse. (SPC, 2013; Petersen et al., 2016)

After emerging at the surface, the hot hydrothermal fluid has a lower density than the surrounding seawater and therefore ascends through the water column. This rising part of the plume is referred to as buoyant plume. As it quickly mixes with the surrounding deep-sea water, it will only rise to between 100 m and 400 m above the seafloor, depending on factors such as the source vent fluids or the degree of stratification of the water column (Speer and Rona, 1989). Above that, the so called non-buoyant plume spreads horizontally with currents and can be detected chemically up to several or even tens to hundreds of kilometres from the discharge site (German et al., 2008).

Most of the dissolved metal in the hydrothermal fluid is carried away by the currents, but some precipitates near the venting location or forms chimney structures. Such chimneys can reach heights of several tens of metres and can also grow together to form beehive structures. The precipitation and collapse of the chimney structures results in the formation of sulphide mounds as part of the sulphide deposits at the site. Further processes involved in deposit formation are dissolution and replacement within and below the mounds (Petersen et al., 2018). If encouraged by the environment (accumulation/trapping, over a long period of time), seafloor massive sulphide (SMS) deposits can develop with high concentrations of economically interesting ore minerals of several hundreds of metres in diameter. Such larger SMS are assumed to occur at about half of the hydrothermal venting sites (Hannington et al., 2011).

Depending on the hydrothermal activity (e.g., temperature of fluid) and environment (e.g., tectonic setting, bulk chemical composition), the mineral composition of deposits varies. Strong regional variations are also possible. In general, the deposits are enriched in base metals (iron, zinc, copper, lead), special metals (cobalt, nickel, indium, germanium), or precious metals (gold and silver) (Boschen et al., 2013). High-temperature sites are of particular commercial interest because certain metals, like copper for example, dissolve only at very high temperatures (approx. 400° C), and therefore such high-intensity hydrothermal systems tend to concentrate mineral deposits (Boschen et al., 2013). Cooler vents, or also the outer part of a deposit, are often compromised by iron and zinc as their participation occurs at a lower temperature (Petersen et al., 2018). Deposits are therefore heterogeneous and to obtain information about the internal structure and composition of a deposit, methods of subsurface geophysics, sampling, and drilling are necessary (SPC, 2013).

Active venting sites terminate over time when the distance to the heat source increases due to seafloor spreading or the site is destructed by nearby eruptions. As mentioned before, magma chambers are shallower at fast-spreading then slow-spreading centres, and therefore hydrothermal sites are found in greater density at such plate boundaries (Hannington et al., 2011). However, higher spreading rates also increase the likelihood of eruptions and therefore

reduce the probability that a site will remain active for a long period of time and form a large deposit. Usually, smaller deposits of less than a few thousand tonnes have been observed at fast-spreading ridges (Hannington et al., 2005). Due to their lower spreading rate and less frequent eruptions, the hydrothermal systems are more durable and have greater potential to develop larger deposits (Hannington et al., 2005).

The hydrothermal fluid discharge areas can be up to several kilometres away from the ridge axis (SPC, 2013). They often occur grouped within an area of 10 km in diameter (hydrothermal fields). Such sites may be active (currently emitting hydrothermal fluids), inactive (currently not emitting hydrothermal fluids but may become active again), or extinct (not expected to become active again) (Jamieson and Gartman, 2020). As the distinction between inactive and extinct is not easily possible without deeper knowledge of the area and the hydrothermal system, the term inactive will be used in this study commonly for sites without current fluid discharge. Deposits associated with inactive sites are generally assumed to be larger (German et al., 2016) as they tend to be older.

The first hydrothermal vent was discovered at the Galapagos Rift (Pacific Ocean) in 1977 (Corliss et al., 1979). Since then, many more have been identified. Currently, about 721 vent locations are known and listed in the InterRidge Global Database of Active Submarine Hydrothermal Vent Fields, version 3.4 of March 2020 (Figure 2.8) (Beaulieu and Szafranski, 2020). These venting sites have been identified by detection of anomalies in the water column (marked in yellow as inferred in Figure 2.8) or by direct observation (marked in red as confirmed in Figure 2.8). As can be noticed from the global distribution, hydrothermal vent sites are found predominantly in tectonically active areas such as MORs, back-arc basins, volcanic arcs, and also seldom at hot-spot volcanoes. About three quarters of known sites are active, as the exploration techniques are mostly based on the detection of water column anomalies in temperature or redox. Identifying inactive sites is more difficult as they do not cause water column anomalies and may be covered by sediment or lava (Petersen et al., 2018). Refer to Chapter 3.1 for further details on SMS exploration.

The detection and identification of hydrothermal sites is not trivial, as they are relatively small in size (starting from a few hundred square metres) for the great water depths at which they are mostly found. Although the number of known sites has been growing over the last years, many of the discovered sites have not yet been verified, and knowledge of the size and composition of the associated deposits is very limited as they are mainly lying underneath the seafloor. Drilling or at least sampling is necessary to gain a better understanding. As a result, estimates of the number and average spacing of hydrothermal sites as well as of the number and size of associated deposits vary

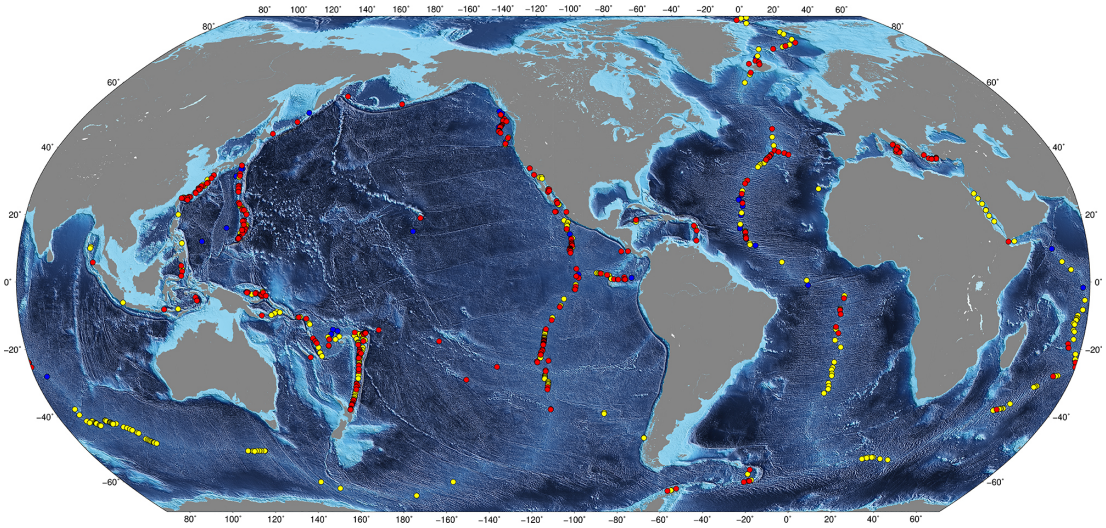


Figure 2.8: Vent locations after InterRidge Global Database of Active Submarine Hydrothermal Vent Fields, Version 3.4 from March 2020 coloured by status: confirmed (red), inferred (yellow), and new since last version 3.3 in 2015 (blue) and bathymetry from NOAA ETOPO1 (Beaulieu and Szafranski, 2020).

widely in literature. Such estimates are usually based on existing data like the InterRidge Global Database, literature reviews, measurements, and models. Beaulieu et al. (2015) predicts based on known occurrences that, in general, approximately 1,300 active hydrothermal sites can be expected. In Hannington et al. (2011), the total number of vent fields and mineral deposits, based on studies and models, is given as 500 to 5,000. Taking the *InterRidge Global Database* as a reference, a distance of about 12 to 220 km can be estimated between adjacent sites. Baker et al. (2016) concludes a spacing of about 3 to 20 km for fast- and intermediate-spreading ridges, taking into account also diffuse and particle-poor venting as well as grouped sites. This assumption is based on redox measurements from a towed platform along 1,470 km of fast- and intermediate- spreading ridge sections of the total 66,000 km length of MORs worldwide. In Hannington et al. (2011), distances of about 54 km at fast-spreading and 174 km at slow-spreading ridges are given for deposits based on published data. These large differences in the estimates shows how incomplete our current knowledge of existing hydrothermal sites and sulphide deposits is.

Exploration and research of hydrothermal venting sites is not only focused on their general resource potential, but also on their unique fauna. The development of biological ecosystems at vent sites is made possible by the chemosynthetic bacteria (SPC, 2013). The bacteria gains their energy not from photosynthesis but from chemical processes, allowing it to survive in hostile environments without sunlight, like the deep sea. About 600 vent-endemic species are known to exist only at hydrothermal sites. They survive in this environment that is toxic to most animals. Local densities can reach a few thousands of individuals per square metre. In the Indian Ocean, shrimps, anemones, and

snails are typically found at venting sites (SPC, 2013). Figure 2.9 shows pictures of typical vent fauna is depicted for active hydrothermal venting sites in the Central Indian Ocean. When active sites becomes inactive, the vent fauna vanishes and the sites become more accessible from an environmental perspective (Koschinsky et al., 2018).

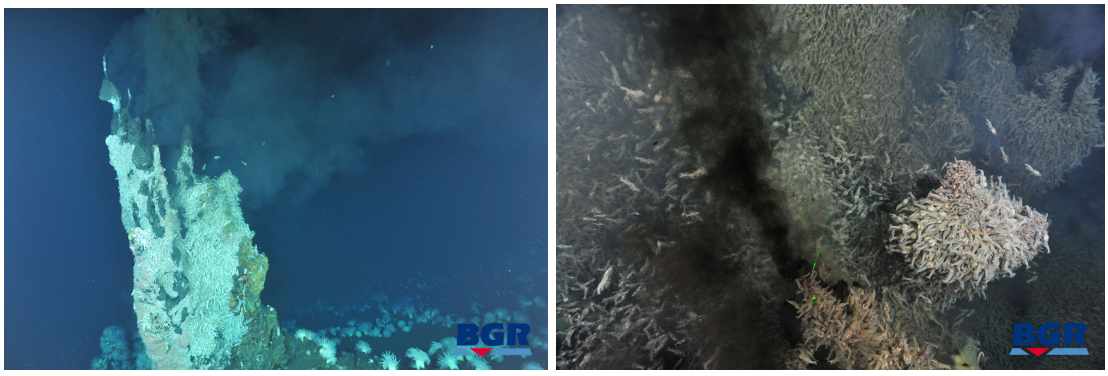


Figure 2.9: Hydrothermal venting sites at the Central Indian Ocean with typical vent fauna like shrimps in direct vicinity to the emitting venting fluids. (Pictures are courtesy of BGR.)

2.1.5 Section Summary

This section provides a brief summary of the most important geological principles in the context of hydrothermal venting. Within the scope of this study, no in-depth geological knowledge is expected or required. Therefore, the purpose of this chapter is to introduce hydrothermal venting and SMS, as well as certain terms and environmental characteristics that will be used at a later stage in this work. For further reading refer to German et al. (2004) or Klischies (2021).

The above comparison of different estimates found in literature regarding the number and density of hydrothermal sites, highlights the lack of knowledge and the challenge of detecting and identifying these relatively small features in a deep-sea environment. This illustrates the overall motivation of this study to contribute to the improvement of exploration methods for hydrothermal sites and SMS. The general BGR exploration workflow is outlined later in this work (Chapter 3.1). High-resolution multibeam echo sounder (MBES) data plays a crucial role in exploration and is also the main data source used in this study. The principles of MBES using a submerged platform are presented in the following chapter.

2.2 Principles of Seabed Mapping with Submerged Multibeam Echo Sounder Platforms

Echo sounders have been used to determine the water depth for more than 100 years. Prior to this, manual measurement techniques like lead lines were used. Independent of the technique, each depth measurement (sounding) requires its position in a global reference system to be known for georeferencing and subsequent mapping. Depth and position are usually determined by different instruments. For such a data acquisition by various sensors for a combined data usage, the relative location and orientation of the sensors to each other needs to be known and the individual data acquisition needs to be synchronised. Such an integrated system consisting of different sensors measuring the properties of their environment, is referred to as multi-sensor system.

In hydrography, moving multi-sensor systems or mobile mapping system (MMS) have been classically used since the development of echo sounders: a combined depth and position determination from a moving vessel. During the last decades, the applications of multi-sensor MMSs (e.g., geospatial data acquisition from mobile platforms) have increased dramatically and today a wide variety of different platforms and sensors are used in geodesy and related fields. For example, MMSs can be deployed in air, on roads, or in several thousand metres of water depth. They use sensors that collect point clouds, videos, or images of their surroundings. The configuration and complexity of an MMS depends on the individual application.

This section explains the principles necessary for the understanding of seabed mapping using multibeam echo sounder (MBES) on a submerged sensor platform, as this was the main instrument used to acquire the data analysed in this study. The principles include the basics of positioning 2.2.1, MBES bathymetry 2.2.2, and MBES backscatter data 2.2.3.

2.2.1 Positioning of Unmanned Underwater Multi-Sensor Platforms for MBES Data Acquisition

Each MBES measurement is initially a relative measurement, providing a 3D coordinate in relation to the transducer. Only by knowing the position and orientation of the echo sounder in a global reference system can the initial measurement be geo-referenced and spatially assessed. For this reason, the common MBES survey setup consists of several sensors: In addition to an MBES, positioning and attitude sensors are also utilised. Typically, an inertial navigation system (INS) is used for attitude determination and position estimates using dead-reckoning. Additional external sensor information are included to provide an integrated navigation solution. Such an integrated navigation solu-

tion increases the accuracy as well as reliability of the navigation by combining different positioning methods to reduce their individual disadvantages and exploit from their advantages. On the one hand, absolute positioning data from GNSS (Global Navigation Satellite System) or USBL (ultra-short base line) can often be noisy (low precision) or contain outliers. On the other hand, a dead-reckoning navigation solution drifts strongly if no auxiliary information is available (Groves, 2013). Depending on the setup and characteristics of the individual survey, the navigation data may be post-processed at the initial processing stage, to improve its quality. This typically involves the application of further corrections, the removal of erroneous GNSS/USBL measurements, or filtering of the integrated navigation solution with forward-backward-smoothing filter for position drift. This is usually done before any further MBES data processing.

Absolute positioning using GNSS is not possible for underwater applications. USBL is therefore used as alternative. An USBL system consists of a transceiver mounted under the vessel and a transponder installed at the underwater platform. The transceiver consists of an array of four transducers. The emitted acoustic signal from the transponder is received by the transceiver and based on the phase difference of the received signal from each transducers in the array is used to determine the direction to the transponder and also the slant range between transceiver and transponder (Lurton, 2010). Since the absolute position of the vessels is known (from GNSS positioning), as is its orientation – and that of the USBL, the absolute position of the transponder and the underwater vehicle can be determined and transformed into a global reference system.

The accuracy of USBL systems depend on the distance between transceiver and transponder, the validity of used sound speed profile, and the general setup (e.g., stability of the installation, accuracy of the level arm determination). Typically, the sole position accuracy and precision of an USBL is not sufficient for high-resolution MBES surveys in several thousand metres of water depth. In addition, this method does not provide the necessary data rate and attitude information of the underwater vehicle, which are essential for high-resolution MBES surveys. For this reason, INS are used not only on underwater vehicles but also on ships. An INS is a three-dimensional dead reckoning navigation system consisting of a set of inertial sensors, known as an inertial measurement unit (IMU), and a navigation processor, as shown in the scheme in Figure 2.10 (Groves, 2013).

Dead reckoning is the process of calculating one's current position from an initial position and then taking into account the determined speed and direction of movement (course). Such inertial sensors are accelerometers and gyroscopes measuring the force and angular rate along / around an axis. After integrating

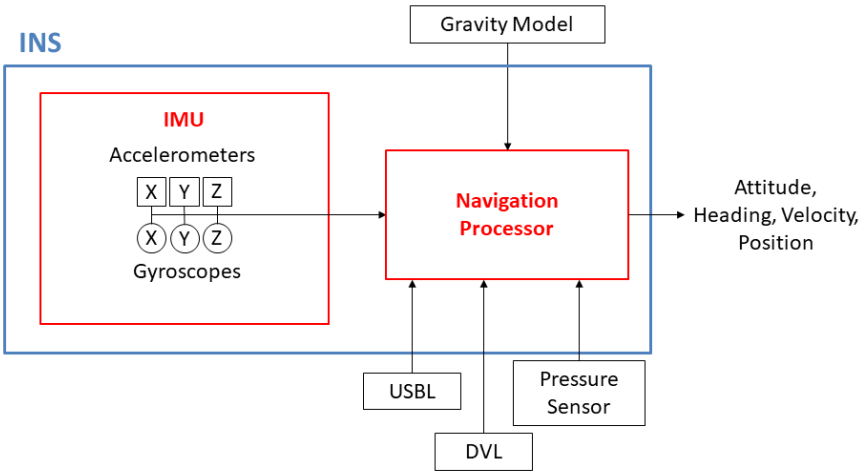


Figure 2.10: Schematic depiction of an submerged INS as used for MBES acquisition (modified after Groves (2013))

the accelerations twice and the angular velocity once, the translation along and rotation around the three axes have been determined. The navigation processor also takes into account the external information (e.g., gravity model, initial position) to output a position, velocity, and attitude solution. Such sole dead reckoning navigation method is subject to a strong drift as small errors accumulate over time (error propagation) if no further external data is used for aiding. To reduce the drift and improve the accuracy of the position estimate, additional sensor data is incorporated. For typical hydrographic applications using an underwater vehicle, these are besides USBL (necessary for absolute positioning), DVL (Doppler velocity log, providing speed over ground and altitude information), and CTDs (conductivity temperature depth probe, forwarding pressure and therefore vehicle depth information). (Groves, 2013)

The integration of data from various sources and with different characteristics and accuracies is done by using algorithms like Kalman filter for an integrated navigation solution. The disadvantages of the individual measurement techniques used are compensated for and the overall position estimate is improved and, in particular, the integrity and accuracy information can be provided.

Knowledge of the local and current water sound speed are crucial to the quality of any hydroacoustic measurement. Especially when echo sounders are operated from ships and their signal travels from the water surface down to the seabed, a sound speed profile (SSP) through the whole water column must be provided to the echo sounder system. This is important to avoid wrong runtime based calculations of the water depth and also to account for refraction effects, particularly from MBES outer beams (Hughes Clarke, 2012). In applications, where unmanned vehicles with MBES are deployed in greater water depths, the determination and updating of SSPs for the water column between the echo sounder transducer and the seafloor is not as critical as for measure-

ments using a MBES near to the water surface. At the water surface, the solar radiation, waves, and local currents can lead to strong temperature and salinity variations in the upper water column and thus to strong and also local changes sound speed (Lurton, 2010). Below the water depth of at least 2,000 m, temperature and salinity are compatibly stable and the sound speed depends only on the pressure (e.g., water depth). Is can therefore be estimated better than near the surface.

Although the SSP information is not regularly required for MBES measurement when using an underwater vehicle as MBES carrier platform in the deep sea, it is crucial for the USBL navigation. The length of the slant range between the two main acoustic components of the USBL and its direction. A current and valid SSP is important for accurate measurements, as the acoustic signal used for distance measurement, is also subject to refraction due to different water layers.

2.2.2 MBES Bathymetry Data

In general, an MBES uses a large number of narrow beams arranged in a fan (swath) to determine the depth within a stripe on the seafloor. The beams are generated by the transmitter (which insonifies a wide stripe of seabed in across track and a narrow in along track direction) and the receiver of the MBES (which is sensitive to multiple narrow stripes in across-track and a wide stripe in along-track direction). A depth is determined for the individual areas of insonification generated by the intersection of transmitter and receiver. They are called beams and the area of insonification by one beam is referred to as the footprint. Figure 2.11 shows the geometry of an MBES measurement. One of the footprints and the corresponding beam are highlighted in red. Furthermore, the important MBES parameters of swath width L, swath opening angle β, beam opening angle ϕ, beam tilting angle θ, footprint diameter a_{TX} / a_{RX} and distance between soundings d_{along} / d_{across} are shown schematically.

In general, the measurement geometry of an MBES is determined on the basis of various system, settings, or environment properties. Most parameters can be modified by the system operator during the survey within a certain range in order to adapt to given survey specifications or environmental changes.

Coverage

An important parameter for the coverage and therefore the efficiency of the MBES mapping is the swath width L, which depends on the swath opening angle β and the water depth D (IHO, 2011):

$$L = 2 \cdot D \cdot \tan(\frac{\beta}{2}). \tag{2.1}$$

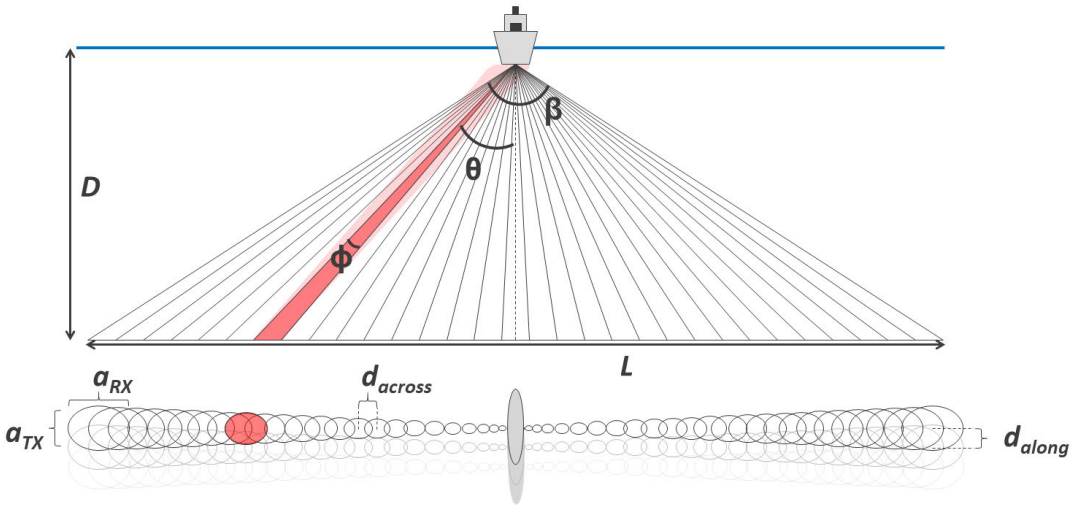


Figure 2.11: Scheme of MBES measurement geometry with water depth D, swath width L, swath opening angle β, beam opening angle ϕ, footprint diameter a_{TX} / a_{RX}, and distance between soundings d_{along} / d_{across} (own illustration).

The swath aperture β is manually adjusted during a survey depending on the data requirements, as a large swath width will increase the coverage but a decrease in across-track sounding density. Furthermore, a larger coverage also results in a decline in quality of the outer beam soundings due to the larger slant range, its greater susceptibility to refraction effects, and its weaker backscattered return signal.

Footprint Diameter / Horizontal Resolution

The footprint size is an important value in MBES data analysis, as it defines the horizontal resolution of a sounding. The backscattered signal from the seafloor is processed by the MBES using bottom detection methods to determine the travel time of the pulse and hence the water depth. The location of the target backscattering the acoustic pulse within the footprint is not known. Therefore, the depth measurement is usually assigned to the position of the beam centre. Larger beam opening angles result in a less detailed detection of small seafloor features within a given water depth.

The area of insonification by one beam depends on the beam opening angle ϕ (system dependent, typically 0.5° - 2°), the water depth D, and the beam tilting angle θ. As the transmitter and receiver beam opening angle can differ from each other, this also accounts for the footprint diameter in across- and along-track direction. The horizontal resolution is examined separately by distinguishing between the beam opening angle in along-track direction ϕ_{TX} and in across-track direction ϕ_{RX}. Therefore, the footprint diameter for a flat seafloor can be approximated according to IHO (2011) as

$$a_{TX} = \frac{2 \cdot D}{\cos(\theta)} \cdot \tan(\frac{\phi_{TX}}{2}), \tag{2.2}$$

$$a_{RX} = \frac{2 \cdot D}{\cos^2(\theta)} \cdot \tan(\frac{\phi_{RX}}{2}). \tag{2.3}$$

Vertical Resolution

The vertical resolution is the minimum vertical distance between two targets that can be separately separately by the echo sounder. In order to achieve this, the distance between the targets has to be greater than half the pulse length of the signal (IHO, 2011). For CW (continuous wave) signals, a shorter pulse length therefore gives a better vertical resolution. On the other hand, a longer pulse allows more energy to be transmitted into the water column, thus increasing the range. The pulse length of an echo sounder should be as short as possible to achieve a small vertical resolution, but must be long enough to reach the required range within the survey area (IHO, 2011). The vertical resolution can be derived from the sound speed c and the pulse duration τ according to de Jong (2006) as

$$\delta z = c \cdot \frac{\tau}{2} \tag{2.4}$$

Sounding Density

The number of beams varies depending on the echo sounder model. Modern MBES use up to several hundred beams per ping. Their arrangement within the swath can be adjusted by the operator. Typically, equidistant and equiangular modes are available, as well as sometimes also a high density mode with an increased number of derived soundings for a given swath opening angle. The equidistant mode is often used to achieve a homogeneous distribution of soundings within the swath. It increases the number of soundings in the inner swath and decreases their number in the outer swath. This mode is often used to increase the number of soundings with better quality, as the inner beams typically have. By setting the fan opening angle β and the mode (number of beams n), the across-track density d_{across} (distance between adjacent soundings) is defined in relation to a given water depth D. It can be approximated when using an equidistant mode by

$$d_{across} = \frac{L}{n} = \frac{2 \cdot D \cdot \tan(\theta)}{n} \tag{2.5}$$

The sounding density in along-track direction is a function of the ship speed v and the ping rate f.

$$d_{along} = v \cdot \frac{1}{f} \tag{2.6}$$

The ping rate depends on the two-way-travel time as an acoustic signal is often only sent after the previous is received. The ping rate is therefore related to

the water depth and the sound speed.

2.2.3 MBES Backscatter Data

In general, three different processes occur when an acoustic signal reaches the seafloor at an incident angle: A part of the incident signal is reflected in symmetrical direction away from the source (specular reflection), a part is scattered in all directions, and another part penetrates the seafloor. Within the sediments, the signal is deflected due to refraction and also volume scattering occurs. These processes are depicted schematically in Figure 2.12.

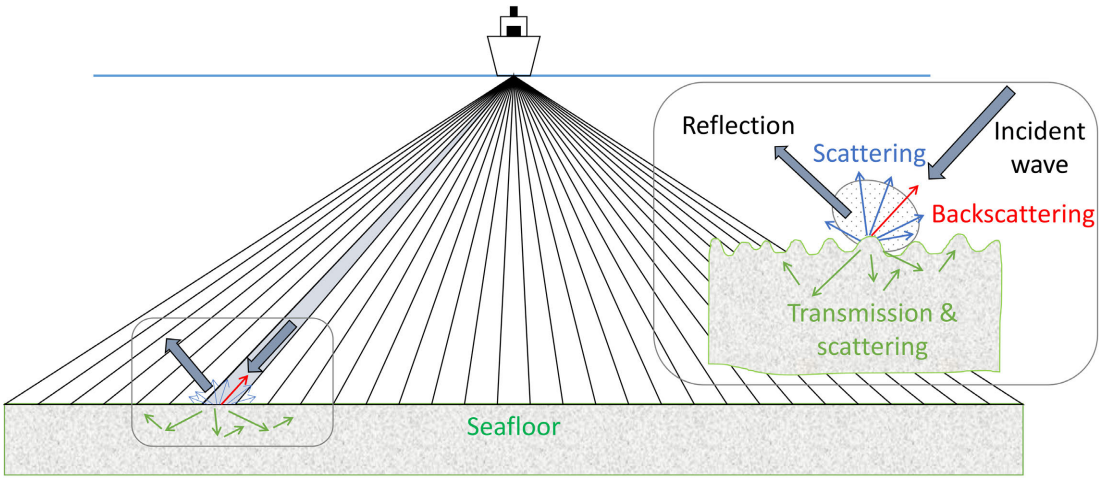


Figure 2.12: Schematic representation of the processes occurring when an incident acoustic signal reaches the seabed: reflection, transmission, and scattering (modified after Lurton (2010)).

The scattering of acoustic energy at the seafloor back towards the echo sounder is called backscattering (Lurton, 2010). The backscattered signal is the information which is received and processed by the MBES for water depth determination. How much of the incident energy is backscattered towards the echo sounder depends on several factors and parameters: echo sounder settings (e.g., transmission level), physical properties of the water (e.g., absorption), acquisition geometry, and seabed characteristics. Most MBES today not only determine the water depth, but also collect information on the received intensity signal. This can be used to derive information about the seabed characteristics after all other non seabed properties related effects have been eliminated. Changes in the backscattered intensities would therefore indicate changes in the seabed properties, allowing a prediction of the seabed type composition and distribution throughout the survey area.

The processes involved in an echo sounder measurement can be summarised in the sonar equation. For an active echo sounder, the sonar equation is

$$EL = SL - 2TL + TS. \tag{2.7}$$

The echo level *EL* is the echo received by the sonar. *SL* is the source level transmitted by the transducer in 1 m distance to it, *TL* is the transmission loss of the signal as it propagates through the water column, and *TS* is the target strength (intensity backscattered from the target to the echo sounder in proportion to how much was received at the target). If the echo level exceeds a certain threshold, the return is interpreted as a seafloor response and processed.

The components of the sonar equation are shown schematically in Figure 2.13. Often, the sonar equation includes an additional part that describes the noise (background noise, reverberation) detected by the echo sounder in addition to the target echo. Assuming that the echo level is higher than the noise, this part has been neglected at this point for simplification.

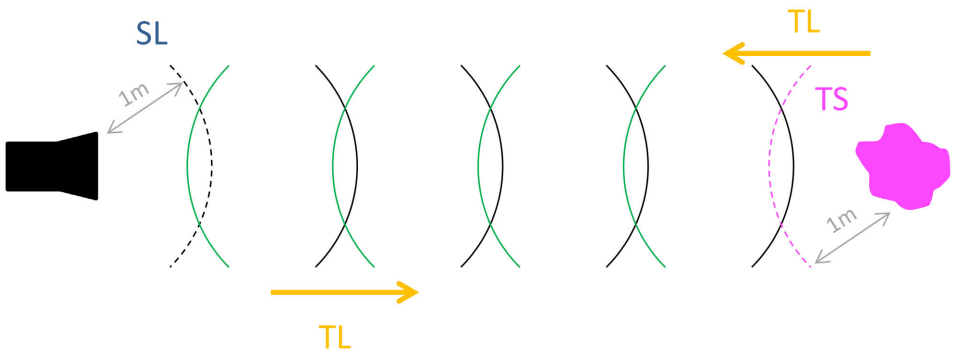


Figure 2.13: Scheme of the components summarised within the sonar equation (own illustration).

One important acoustic quantity is the acoustic intensity *I*, which describes the power per area unit and can be expressed for a plane wave after Lurton (2010) as

$$I = \frac{p^2}{\rho c} \left[\frac{W}{m^2}\right] \tag{2.8}$$

with the acoustic pressure *p*, density of the medium ρ, and sound velocity *c*. The product of density and sound velocity is also referred to as acoustic impedance and is a measure for the resistance of a medium to the formation of an acoustic wave. The quantities of the sonar equation are given as levels in decibel (dB, ten times the decadic logarithm) as this logarithmic unit allows a more functional handling of acoustic quantities. Expressing the acoustic intensity as intensity level IL in decibels (ten times the decadic logarithm of the ratio of the intensity *I* to a reference intensity I_0 yields:

$$IL = \log \frac{I}{I_0}[B] = 10 \log \frac{I}{I_0}[dB] \tag{2.9}$$

Going back to the sonar equation, the source level SL can therefore be expressed as

$$SL = 10 \log \frac{I}{I_0} || re 1 \mu Pa @ 1 m \tag{2.10}$$

This reference intensity is the intensity of an acoustic plane wave with a root mean square pressure of 1 μPa at a distance of 1 m (Lurton and Lamarche, 2015). As the signal propagates through the water, it experience energy losses due to geometrical spreading and absorption. This transmission loss *TL* is defined according to Lurton (2010) as

$$TL = 20 \log \frac{R}{R_{1m}} + \alpha R \tag{2.11}$$

where *R* is the range in metres and α is the absorption coefficient. The first summand accounts for the loss due to the spreading and the second summand for the loss due to absorption. Since the transmission loss occurs on the way away from the source as well as on its way back from the target, *TL* is included in the sonar equation with a factor of 2. The target strength *TS* is generally defined as the ratio of the intensity of the backscattered signal I_{bs} and the incident intensity I_i

$$TS = 10 \log \frac{I_{bs}}{I_i}. \tag{2.12}$$

Generally, a differentiation between point targets and extended targets is made when considering the target strength. Point targets like a fish or gas bubble are completely insonified by the echo sounder beam whereas extended targets are only partially insonified. Due to reasons of simplification, only the scattering of an extended target (like the seafloor) will be described here. For more information on point targets, please refer to further literature like (Lurton, 2010) or (Lurton and Lamarche, 2015).

The target strength of extended targets

$$TS = BS + 10 \log A \tag{2.13}$$

depends on the unit surface/volume backscattering strength BS (in dB, of 1 m^2 seafloor) and the area of insonification *A* (Lurton, 2010). The area of insonification *A* depends on the beam aperture, the signal duration, and the measurement configuration (e.g., range, angle). The backscattering strength is related to the seabed type and its properties, the incident angle, and the acous-

tic frequency. It is the parameter of interest in regard to backscatter analysis as it includes the seabed characteristics. Once all other influences have been removed from the measured echo level, and therefore the backscattering strength information has been retrieved, a backscatter analysis of the survey area can be performed. Such backscatter analyses are used in a variety of research areas such as habitat mapping, sediment grain sizes estimation, or – as with the INDEX project, to differentiate seabed characteristics and support geological mapping.

2.2.4 Section Summary

This chapter focuses on explaining the basic functionality of the MBES and sensor specific terms to give the reader some fundamental background information. Besides concentrating on the mapping sensor – MBES – some theory specific to the use of unmanned underwater vehicles as mobile mapping platforms is provided. The objective is to explain what other sensors are needed and their working principles. The explanation is kept more general and detailed specifications of the instruments used can be found in Chapter 3.2.1. The complete MBES processing workflow from validation, correction, and improvement to the generation of digital terrain models (DTM) and mosaics is outlined in Chapter 4. As the final data products (DTMs and mosaics) were used for an automatic detection of sulphide mounds using AI techniques, the following Chapter 2.3 provides some definitions of AI terms and an introductory insight into the different methods summarised under the general term of AI.

2.3 AI Technologies

In recent years, technologies using **Artificial Intelligence (AI)** have been increasingly introduced into a wide variety of branches and industries and thus into our everyday lives. It is a still growing field of technologies and tools that aim to build systems that carry out tasks that usually require human intelligence and typically involve some form of learning and problem solving (Berlyand and Jabin, 2023). Typical applications can be found in healthcare, media, finance, automation, or robotics (de Albuquerque et al., 2024). Such machines using AI have a certain autonomy and can, among other things, process large data sets, recognise objects, make decisions, and give recommendations.

Typical tasks solved by AI – and used in this study – include data segmentation, classification, and characterisation. These terms are closely related and can also be understood as different stages within one workflow. Segmentation originates from the field of image processing and describes the process of dividing an image into different regions – segments – based on the pixel values. In a classification, these groups, differentiated by their properties, are assigned

a predefined label or category. In a characterisation, the properties and characteristics of a class or segment are further described. (Blondel, 2003)

The term "AI" summarises a wide variety of technologies and tools. As within this study certain AI tools are used, this chapter provides some basic definitions and distinctions of terms used in the field of AI in order to provide some essential fundaments for further chapters.

Machine Learning (ML) is a subfield of AI and is the generic term used for algorithms that incorporate intelligence into machines by using some kind of learning. These approaches do not involve specific programming to define how the algorithm arrives at its results (Sze et al., 2017). Machine learning techniques are typically used for data modelling and pattern identification in the data. They can be divided into three types according to de Albuquerque et al. (2024):

- **Supervised Learning:** An ML model is trained on a data set with labelled input and output. The target output is known, and from this the model learns to predict the output for new data samples. Such supervised learning approaches are typically used in predictive analysis or classification.
- **Unsupervised Learning:** The training data is not labelled and the algorithm must create its own labels to identify patterns without a given explicit target output. Clustering or anomaly detection are typical application for unsupervised learning techniques.
- **Reinforced Learning:** These types of ML algorithms use feedback from the environment in the form of rewards and punishments. This allows these algorithms to learn faster and adapt to changing conditions in comparison to unsupervised learning approaches.

Deep Learning (DL) is a term for a subset of ML methods that mimic the structure of the human brain and consist of artificial neural networks. A neuron is a processing unit put together in multiple layers of neurons (Figure 2.14). The neurons are connected between the layers. Their outputs are called activations. Synapses are the weight between the output of one neuron and the input of the next one. They scale the signal by weight. Hence, they are referred to typically as weights. Each neuron has a threshold. The weighted sum of the input is only combined and further processed if the input exceeds the threshold value. As a network is trained, the weights are updated. The aim is to minimise the loss, which is the difference between the ideal result (labelled data) and the prediction made by the neural network based on the current weights (Sze et al., 2017).

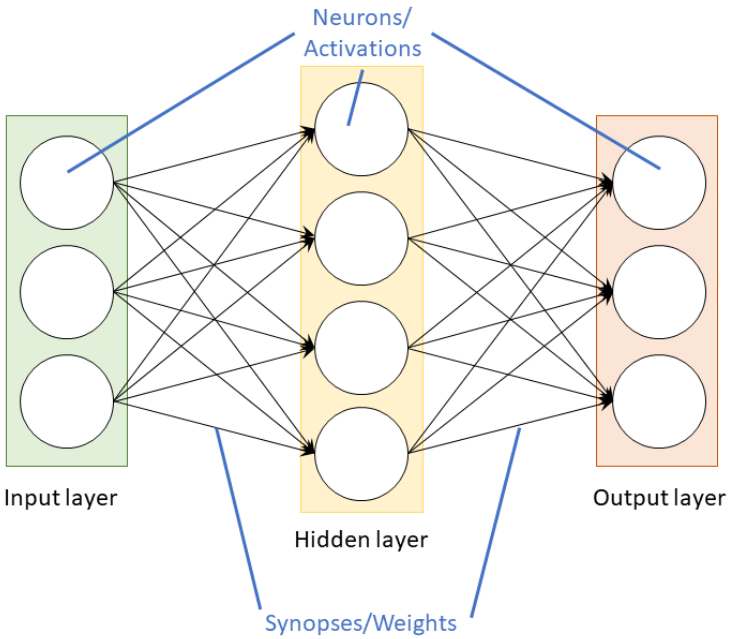


Figure 2.14: Simplified scheme of artificial neural network (Sze et al., 2017)

Neural networks with more than three layers (i.e., one input layer, one output layer, and more than one hidden layer) are considered to belong to DL (Sze et al., 2017). The more complex the structure of a neural network, the higher the level of abstraction and performance that can be achieved. Like ML algorithms, DL can use supervised, unsupervised, or reinforced learning. DL methods can process large amounts of data and are capable of automated decision making (de Albuquerque et al., 2024).

Convolutional neural networks (CNN) are neural networks that use convolutional layers to make the computation more efficient. Unlike layers in general deep neural networks, only a limited number of neurons of the input layer – called the local receptive field – are connected to neurons in the hidden layer. A neuron in the hidden layer therefore does not process the whole input, but only a part of it. The receptive field is defined by a filter. This approach follows the assumption that local pixel patterns are more relevant for classification than the overall data. Convolutional layers create abstractions of the input data – a so called feature map. In addition, a CNN reduces the number of weights by assigning the same weight to each of the neurons in the hidden layers. This means that all of them detect the same feature. (MathWorks, 2017)

Figure 2.15 shows a scheme of a convolutional layer for multi-dimensional image segmentation: In this example, the input activations for a convolutional layer are sets of 2D input feature maps (i.e., channels). These are convolved using filters composed of weights – a different one for each channel. The results from one location (over all channels) are summed to one channel of the output feature map. Additional filters may be used to generate a multi-

dimensional output. (Sze et al., 2017)

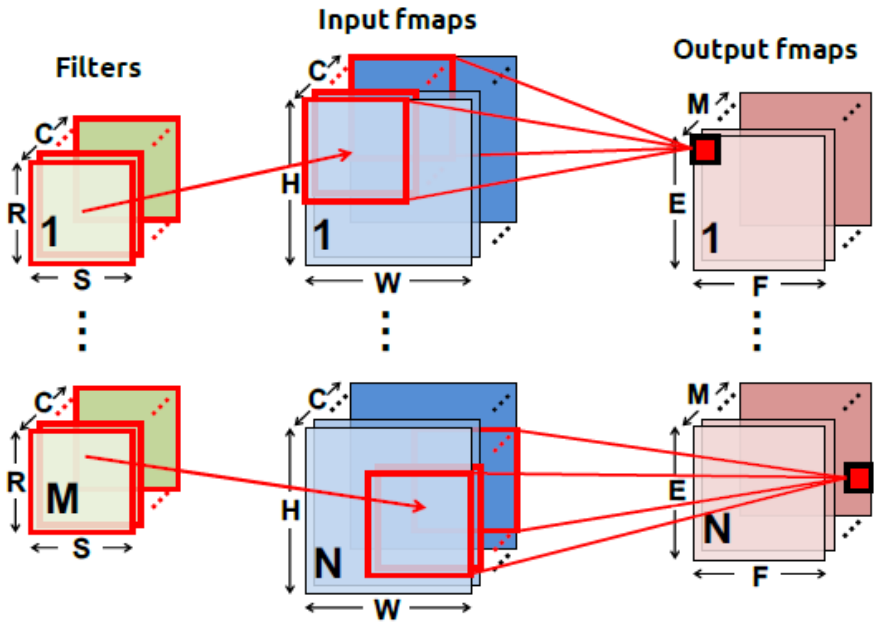


Figure 2.15: Scheme of multi-dimensional convolutions in CNN for image segmentation (Sze et al., 2017).

CNN do not exclusively employ convolutional layers. A convolutional layer is typically followed by a small number of fully-connected layers (all neuron of the input layer are connected to all neurons of the hidden layer) for classification. Another optional type of layers performs pooling. Pooling reduces the dimension of a feature map using operations such as maximum or average for example. Its utilisation makes the neural network more robust to small distortions (Sze et al., 2017). CNN can vary in their architecture, such as layout of layers and their connection, depending on their purpose. In this work, the CNN U-Net (Ronneberger et al., 2015) is used for mound detection. It is therefore explained in greater detail.

U-Net is a CNN originally developed for biomedical image segmentation (Ronneberger et al., 2015). It is designed to work with reduced training data, allowing for more accurate segmentation. Figure 2.16 shows the architecture of U-Net, which got its name from its distinct layer layout. It consists of two phases: the encoder (or contracting part), which extracts features and reduces the spatial dimension, and the decoder (expansive part), which increases the spatial dimension again to reconstruct the output. The encoder recursively consists of two 3x3 convolutional layers, each followed by a ReLU (rectified linear unit) for non-linearity, and a max pooling layer. The pooling layer reduces the spatial dimension of the feature maps by taking the maximum value of the 2x2 window (without overlap). The encoder reduces the spatial dimension of the maps, doubling the number of maps in each step. (Ronneberger et al.,

2015)

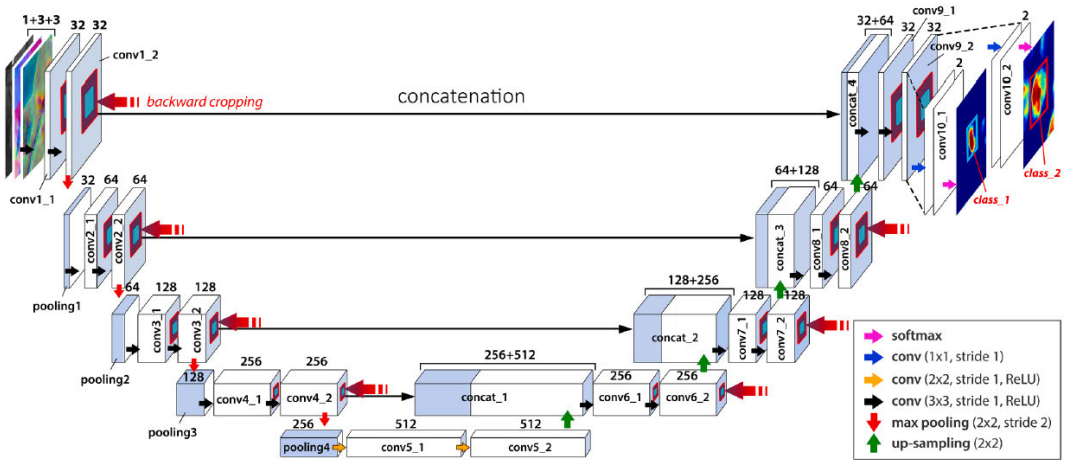


Figure 2.16: U-Net structure (Juliani and Juliani, 2021).

The decoder upsamples the data back to its original size, using at each step upsampling followed by a convolution to halve the number of feature map channels. This is followed by a concatenation with the corresponding encoder feature map and two 3x3 convolutions, each followed by a ReLU. The concatenation allows the layers to operate on and compare the encoder and decoder features. This additional information improved the algorithm in terms of computation time and training performance (Ronneberger et al., 2015). At the two final layers, 1x1 convolutions map the feature vectors to the desired classes. A softmax operation calculates the probabilities for each pixel to belong to one of the classes.

The U-Net code and pre-trained models used in this work were obtained from GitHub Juliani and Juliani (2021). It is scripted using Python 3.6, Tensorflow 1.14.0, and Scikit 0.24.2. The input data of 256 x 256 x 7 is augmented using geometric and colour-based augmentations to increase the amount of training data. Training is divided into epochs. An epoch is a run through the entire data set. After each epoch, the model parameters are adjusted and a validation is performed. A training consists of several epochs to allow the model to be improved iteratively.

Typically, the data for a CNN is split into three parts to be used for different purposes in the model building: training, validation, and testing. Training data is used to adjust model parameters to improve predictions. During training, the independent validation data is used to evaluate intermediate models. If the model performance does not improve on the validation data set, training can be stopped at that particular epoch to avoid overfitting. Overfitting occurs when the model learns not only to recognise certain underlying patterns, but also very specific data characteristics such as noise or artefacts. In such a case, the model will not perform well on new data that has not been used for

training (de Albuquerque et al., 2024). Once the model has been completed, the test data set is used to obtain an evaluation of the final model.

Chapter 3

Data Acquisition and Characteristics

This study investigates the potential of automated mound detection and identification based on high-resolution bathymetry and backscatter data in the deep sea. Seven hydrothermal fields in the western Indian Ocean were selected as study area. BGR (Federal Institute for Geosciences and Natural Resources) provided different types of data for this study, covering the the seven hydrothermal fields. The following data set types were provided by BGR – or were generated based on the data provided:

- MBES bathymetric DTM (2 m resolution)
- MBES backscatter mosaic (1 m resolution)
- Geological interpretation along ROV[1] tracks
- Geological maps
- Site locations of active and inactive hydrothermal venting sites
- ROV and VS (video sled) video material (of one field)

In this chapter, first, the study area and the general exploration concept are described before the MBES instrumentation, acquisition, and data sets are outlined. The other data sets used for ground truthing, such as the video footage and the resulting geological maps, are also presented. This chapter aims to provide all relevant information on the study area and the final data derivatives used for the work.

[1] An ROV (Remotely Operated Vehicle) is a submersible sensor platform that is connected to an operator station by a tether. This tether transmits data from the ROV sensors to the operator, who can monitor the ROV movement, its surrounding, and forward navigation or control commands. Depending on the application, ROVs can vary in size, payload, and capacities. They range from small observation class ROVs, mainly equipped with cameras (as well as lighting), to work class ROVs with multiple sensors and tools (Christ and Wernli, 2014).

3.1 Study Area and SMS Exploration

The United Nations Convention on the Law of the Sea (UNCLOS) was finalised 1982 and entered into force in 1994. It deals with the rights, obligations, and responsibilities of nations in regard of the seas. The high seas are beyond national jurisdiction and open to everyone, which includes the freedom of navigation, cable laying, fishing, or research (UN General Assembly, 1982). As part of UNCLOS, the International Seabed Authority (ISA) was established in Kingston, Jamaica. The aim of the ISA is the protection of the marine environment and it organises and controls mineral resources related activities in open waters, which cover about 54% of the ocean's surface (International Seabed Authority, 2021). Mineral resources of interest include polymetallic (manganese) nodules, cobalt-rich ferromanganese crusts, and seafloor massive sulphides of hydrothermal origin. The BGR has received such a contract for the exploration of marine mineral resources on behalf of the German government (Bundesanstalt für Geowissenschaften und Rohstoffe, 2021).

Since being granted the contract for the exploration of SMS deposits in 2015, BGR has been investigating active and inactive hydrothermal sites and possible seafloor massive sulphide deposits in the western Indian Ocean within the INDEX (INDian ocean EXploration) project. Annual cruises are carried out to the contract area to investigate known sites and identify new ones. The contract area consists of 12 clusters along the Central (CIR) and Southeastern Indian Ridge (SEIR). The spreading centres meet at the Rodriguez Triple Junction, where the three tectonic plates (Indo-Australian, Somalia/African, Antarctic) meet. Each cluster consists of 5 to 17 blocks of 10 x 10 km^2 (Figure 3.1).

All clusters are located close to the spreading centres of the slow-spreading CIR (spreading rate of 4.7 cm/yr) and the intermediate-spreading SEIR (spreading rate of 5.7 cm/yr) (Bartsch, 2014). As mentioned in Chapter 2.1.4, the majority of hydrothermal venting sites are found in the vicinity of MORs. Slow-spreading ridges in particular, are more likely to form large deposits because the sites move more slowly away from the heat source. In addition, also the probability of destruction by nearby eruptions is lower. The expected density of hydrothermal fields is lower than at fast-spreading ridges as the magma is located in greater depth (German et al., 2016).

The exploration of hydrothermal venting sites and possible associated SMS deposits involves their identification, delineation, and evaluation (SPC, 2013). Hydrothermal venting sites vary in size – ranging from just a few hundred square metres to several million square metres. Identifying such small seafloor features in great water depths of about 3,000 m, as in the INDEX contract area, is challenging because the size of individual sulphide mounds on the seafloor coincides roughly with the resolution of ship-borne MBES. As a result, they

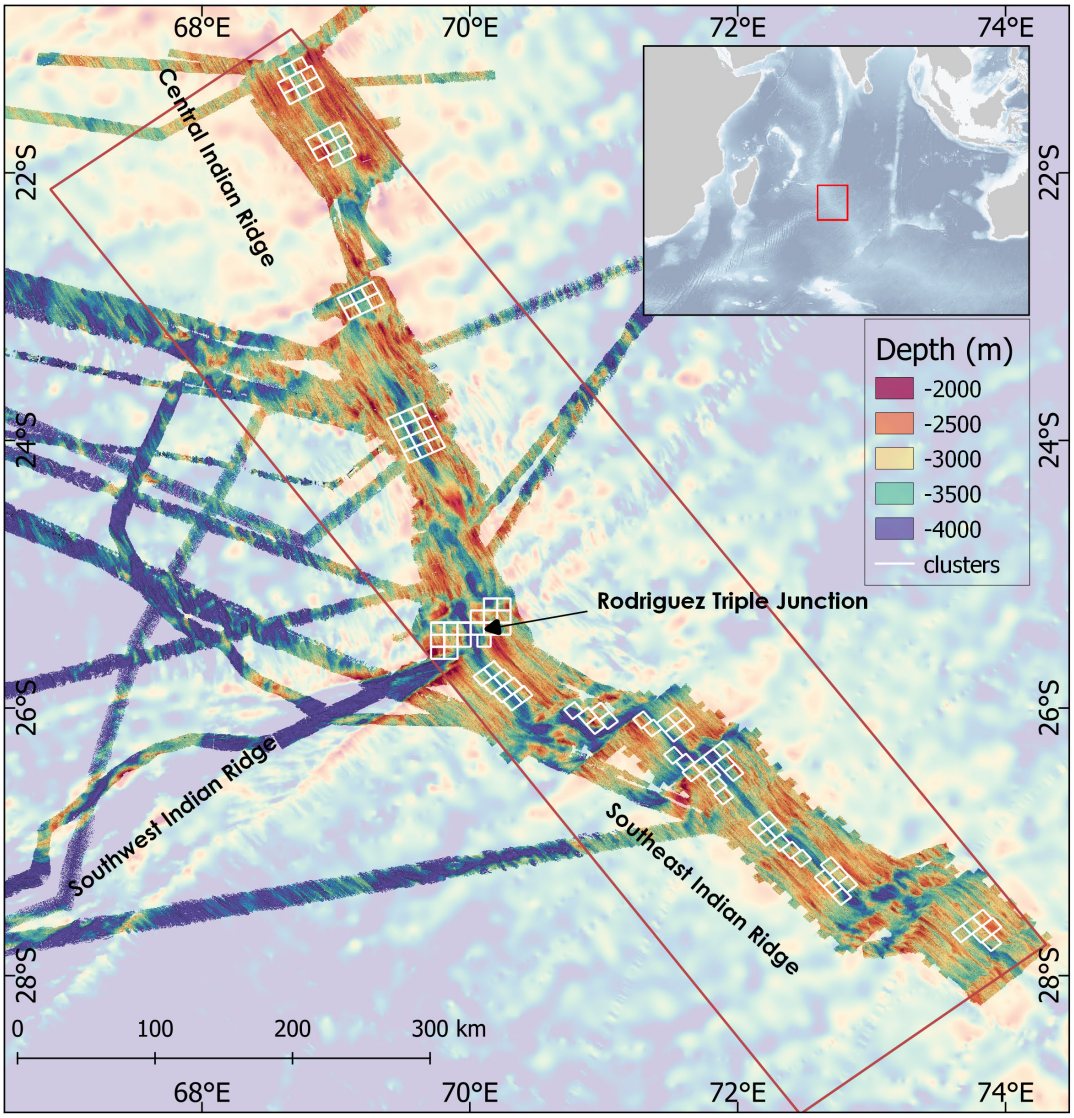


Figure 3.1: Overview of the contract area for exploration in the Indian Ocean (INDEX) consisting of 12 clusters and 100 blocks (white) with bathymetry in the background (GEBCO_2019 grid (GEBCO Compilation Group, 2019) and ship-borne MBES bathymetry (resolution 40 m)).

cannot be mapped using ship-borne methods. In the following, the exploration method to detect new hydrothermal sites is described based on the concept of BGR, which is complies with general approaches.

A bathymetric map and backscatter mosaic derived from ship-borne MBES data serve as basis to identify geological features and possible areas of interest with a higher probability of hydrothermal activity (Gràcia et al., 2000; Eason et al., 2016). In these areas, a sled equipped with chemical sensors is towed to detect water column anomalies that possibly originate from hydrothermal activity. Such "plume-hunting" sensors measure, for example, temperature, oxidation-reduction potential (ORP), and particle concentration (German et al., 2008; Tao et al., 2017; Baker et al., 2016; Peukert et al., 2018). BGR's "plume sled" SOPHI (Sled for Oceanographic Plume Hunting Investigation) is depicted

in Figure 3.2 (top left). It is also equipped with CTD sensors (conductivity, temperature, pressure, dissolved oxygen), rosettes for water sampling, and an altimeter. SOPHI is towed behind the vessel in a tow-yow concept, constantly altering the altitude over ground between 20 m and 400 m to ensure that water column anomalies caused by hydrothermal venting are detected. The data is monitored in real time and additional Miniature Autonomous Plume Recorders (MAPRs) with turbidity and ORP sensors are deployed on the towing cable to prevent missing plume caused anomalies at greater heights above the sled (Walker et al., 2007; Baker, 2017).

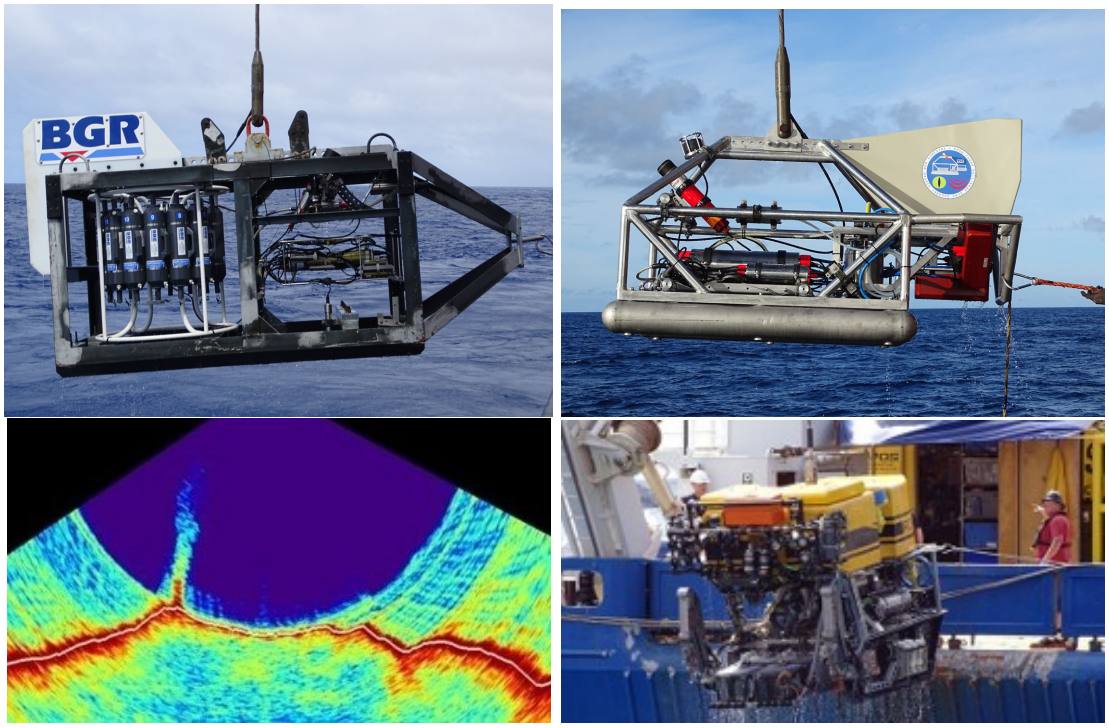


Figure 3.2: BGR exploration tools for hydrothermal venting sites. From top left to bottom right: SOPHI (Sled for Oceanographic Plume Hunting Investigation), HOMESIDE - deep-towed MBES sled, visualisation of MBES water column data showing a plume, ROV ROPOS. (Pictures are courtesy of BGR.)

Once an anomaly has been detected in the water column, possible areas of origin are identified based on the water depth of the anomaly and the local bathymetry. In these areas of interest, high-resolution mapping is performed using an unmanned sensor platform equipped with MBES deployed at a low altitude (between 30 m and 100 m) above the seafloor (German et al., 2008; Ondréas et al., 2018). Within the INDEX project, AUVs (Autonomous Underwater Vehicles) and ROVs (Remotely Operated Vehicles) are employed occasionally, but mainly the HOMESIDE sled developed by BGR is utilised (Figure 3.2 top right). More detailed information on HOMESIDE and the installed sensors can be found in Chapter 3.2. The high-resolution bathymetry allows a more detailed geological interpretation of the local seafloor morphology and even the visual

distinction of possible sulphide mounds associated with hydrothermal venting (Jamieson et al., 2014). More importantly, the additional chemical sensors and the real-time visualisation of the MBES water column data provide information on the presence and location of active venting sites (Figure 3.2 bottom left). Once new discharge sites of hydrothermal fluids have been located, a video sled or an ROV with cameras is used to collect video footage of the site and its surroundings (Figure 3.2 bottom right). Geological, biological, or fluid samples can be taken with special gear attached to the ROV.

The exploration concept outlined above is limited to active sites with current discharge of hydrothermal fluid. Furthermore, all these methods focus on surface investigations, which provide only limited information about the size and composition of the underlying deposit. Therefore, marine geophysical active (electromagnetic, induced polarisation) and passive (self-potential, magnetics) methods are used to identify buried deposits (e.g., inactive sites) and to obtain some estimates of the dimension of a deposit (Müller et al., 2023). To gain information about the composition of a deposit, only sampling by drilling provides conclusive information.

Hydrothermal venting does not occur isolated. Often a number of active and inactive vents can be found in close proximity (approx. 30 m) to each other. They are then typically located on a mound or several connected mounds. These are referred to as sites. An accumulation of sites in an area is called a hydrothermal field. As of the end of 2023, 14 hydrothermal fields are known within the INDEX contract area. 13 of these have been discovered since obtaining the exploration contract in 2015. Most of the hydrothermal venting sites are located to the east of the ridge axis at a distance of up to 14 km from the spreading centre.

The work of this study focuses on seven hydrothermal fields in the BGR contract area. They are listed in Table 3.1 with their maximum extent as well as the number of active and overall sites. The number of known sites and the extent of the fields are regularly updated as new data is collected each year and thereby further knowledge is gained about the fields.

Table 3.1: Overview of hydrothermal fields and their sites (status 01/2024).

Field	Extent	Total/Active Sites
Field 1	2.2 km	14 / 2
Field 2	2 km	7 / 1
Field 3	1.7 km	12 / 5
Field 4	3.6 km	15 / 1
Field 5	3.7 km	14 / 6
Field 6	2.6 km	16 / 9
Field 7	2.4 km	10 / 0

These individual hydrothermal fields are shown and described below. As their data is used for the CNN training and testing for mound detection in this work, the local geomorphological differences between the fields may influence the result. The mound shapes are also addressed, with emphasis on the differences between sulphide and volcanic mounds.

Field 1 consists of 14 sites and a total of seven active venting locations in two of these sites. It was discovered in 2015 following the detection of water column anomalies. It covers an area of 1.8 km x 1 km with water depths ranging from 3,200 m to 3,000 m (Figure 3.3). *Field 1* is divided into two parts by a large north-east tending fault. This fault has an orientation of about 120° and creates a graben in the middle of *Field 1*. The individual sites are located along the graben structure in a linear configuration and with a similar orientation. The northern sites have a slightly different orientation to the linear alignment of approx. 130°.

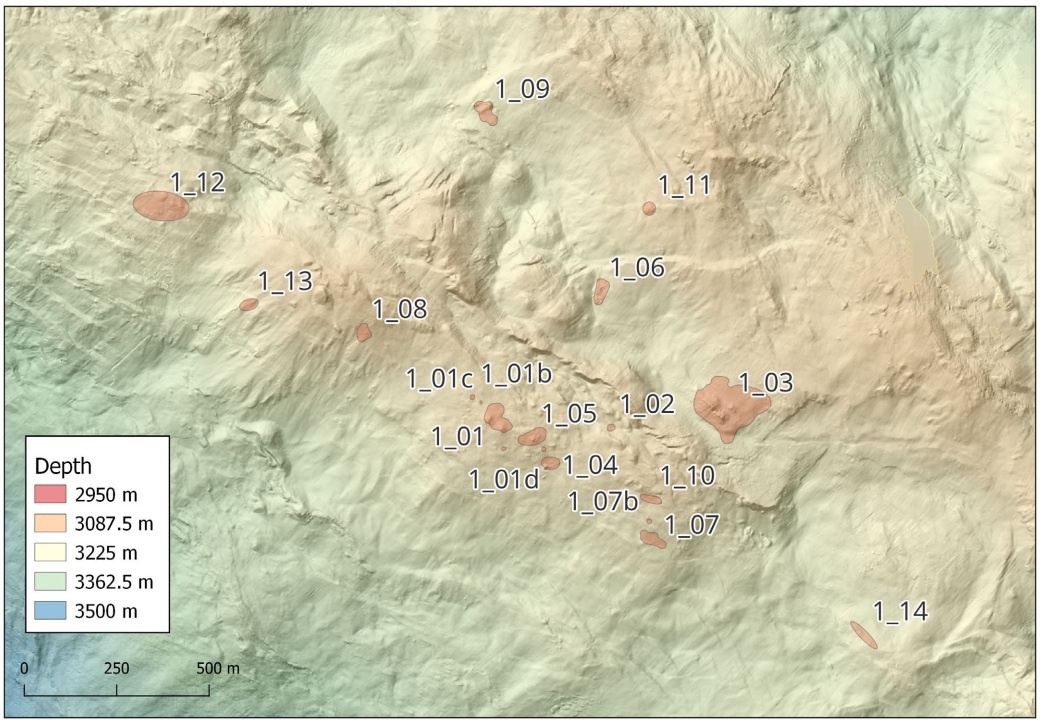


Figure 3.3: Overview of *Field 1* with sites locations.

Two of the sites within *Field 1* are active. They are depicted in Figure 3.4 with bathymetry cross sections and images of the vents. The peaks of the two mounds in Figure 3.4a are about 45 m apart. There are one or taw active vents on each of them. The mounds have a size of about 35 m x 35 m each and a height of approx. 10 m above the surrounding area. There is also evidence of past activity in the form of weathered and collapsed chimney remnants in their vicinity. The site in Figure 3.4b and Figure 3.4c has several vents, including a large beehive chimney structure with various active and inactive chimneys.

There are several mounds: One with a diameter of about 80 m (Figure 3.4b) and two mounds in close proximity with a combined size of approx. 200 m x 100 m (Figure 3.4c). The height of all these mounds is about 10 m to 15 m.

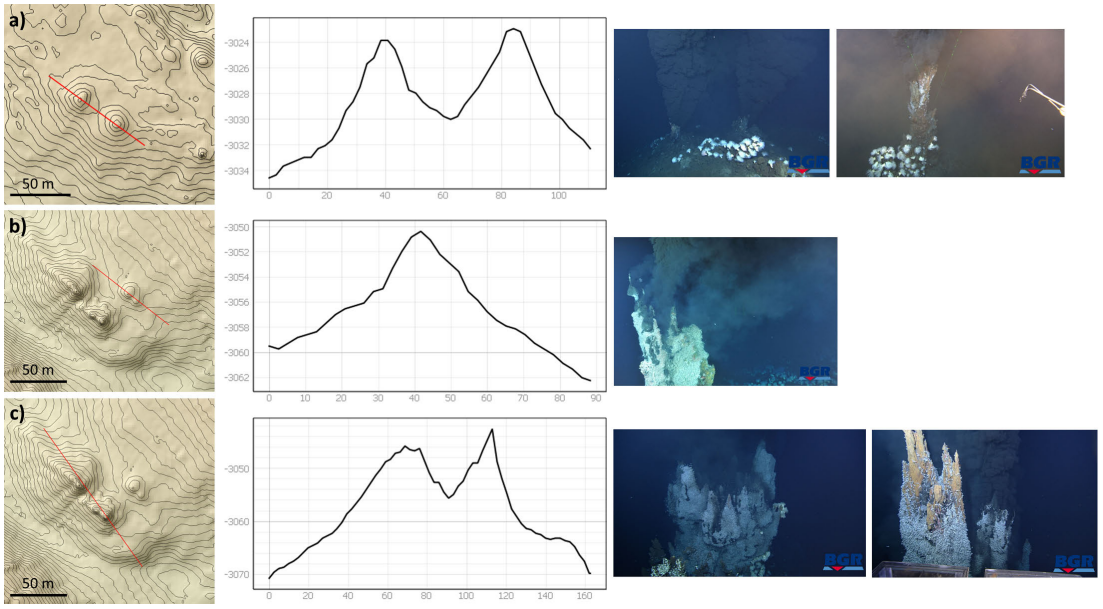


Figure 3.4: Active vent locations in *Field 1* with bathymetry (2 m contour interval), cross section, and pictures. (Pictures are courtesy of BGR.)

From the cross sections above, certain shape characteristics of sulphide mounds can be noted: They have steep slopes, the top is conical, and their sides are often irregularly shaped (Jamieson et al., 2014; Murton et al., 2019). Figure 3.5 shows several inactive sites of *Field 1* with bathymetric cross sections and images. The typical sulphide mound shape is still visible in the ones where only limited erosion or weathering has occurred. Figure 3.5a is an example of such a site that has recently become extinct. It is a large site where some of the tall inactive chimneys are still standing upright. Older inactive sites are more likely to be weathered and be covered by sediment. The typical sulphide mound morphology is no longer as distinct and the slope and conical top appear smoother (Figure 3.5b) (Murton et al., 2019). Weathering and cover can progress to a stage, where the mound structure is no longer morphologically apparent and only stained sediments or very small chimney remnants indicate past hydrothermal activity, as at the site shown in Figure 3.5c.

Lava domes of volcanic origin also form mounds of similar size to sulphide mounds. These volcanic mounds are formed by solidified lava and consist of pillow lavas. Their shape is usually slightly different from that of sulphide mounds, as the top is less conical evolved and has a more smooth concave shape, as shown in Figure 3.6.

Mound analysis based on bathymetry alone is not trivial, and ground-truth data such as video imagery is important for a reliable distinction of its type (i.e., sulphide or volcanic). There are two reasons why interpretation is difficult:

firstly, the geometric outline of mounds is not easy to define, as the transition from mound to surrounding terrain is gradual. Secondly, the shape of mounds, especially inactive sulphide mounds, changes due to erosion and weathering. In addition, the (partial) coverage of a mound with talus or other material may further alter its morphological appearance.

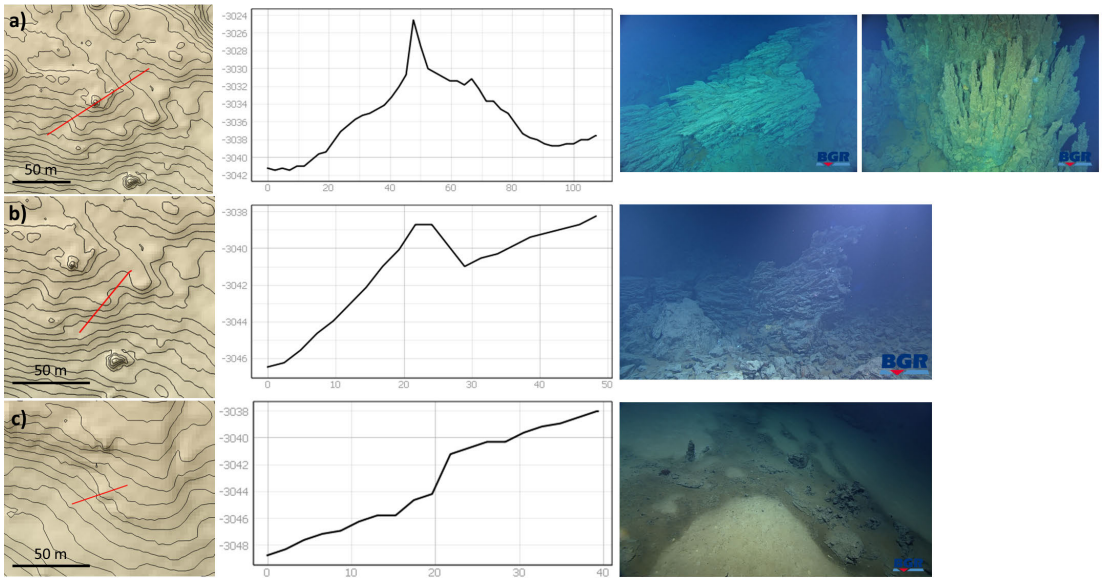


Figure 3.5: Inactive vent locations in *Field 1* with bathymetry (2 m contour interval), cross section, and pictures. (Pictures are courtesy of BGR.)

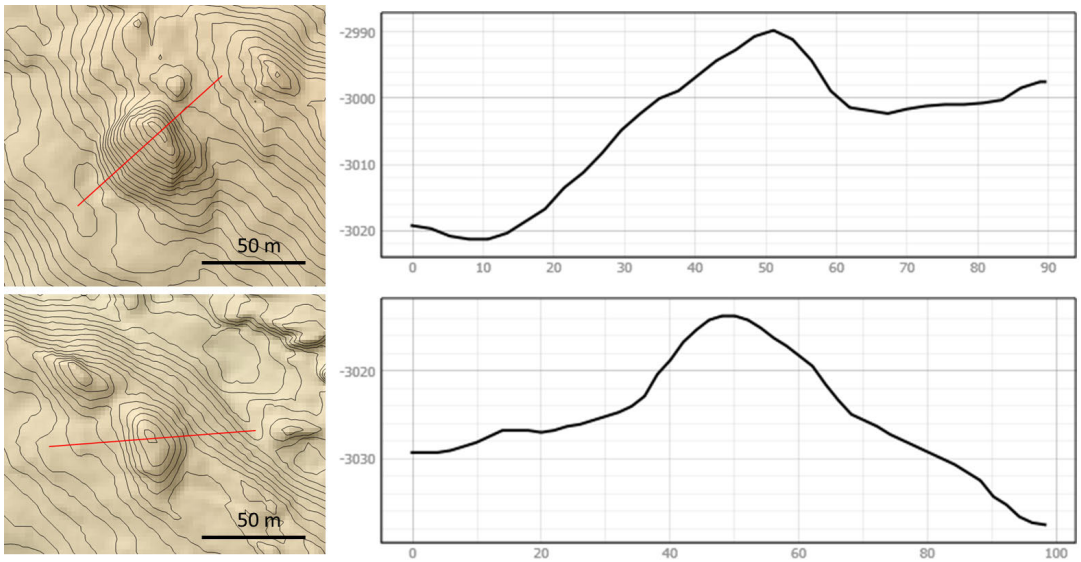
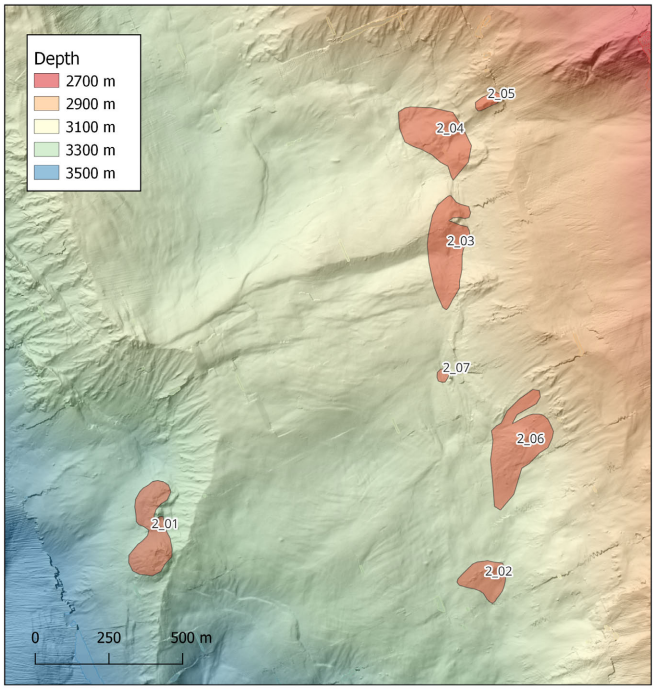
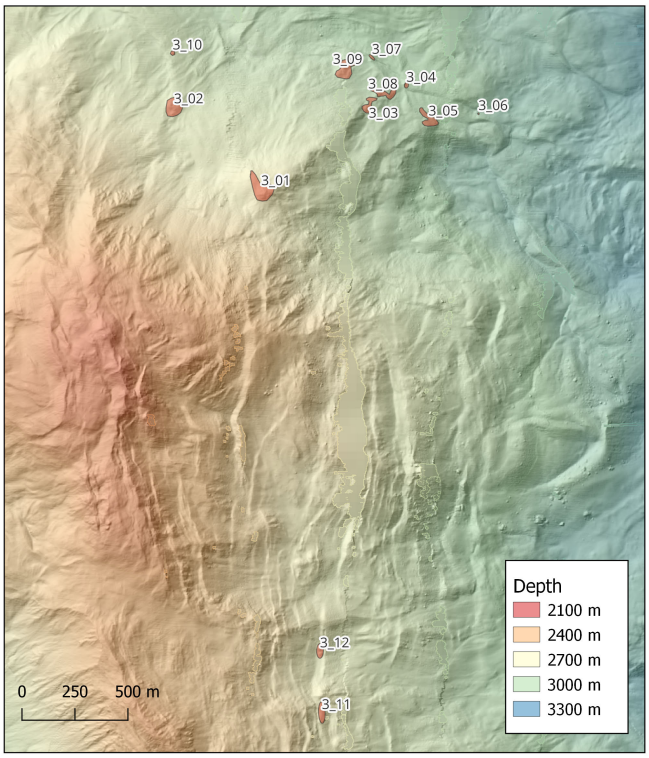


Figure 3.6: Volcanic domes in *Field 1* with bathymetry (2 m contour interval) and cross section.



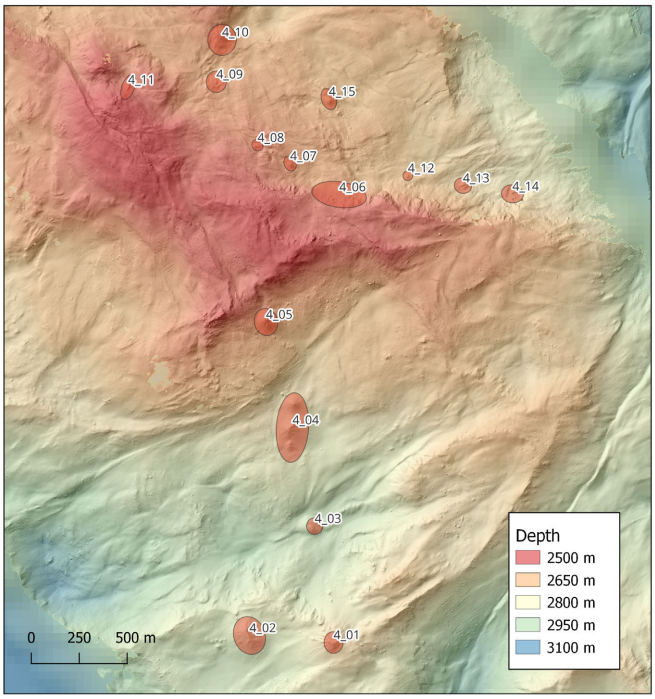
Field 2 consists of seven sites, of which only site 1 is active. This site consists of several discharge sites on two overlapping mounds (about 150 m x 150 m each). It is located approx. 1.3 km west of the other sites, which extend over a distance of 1 km to 8 km in north-south direction. The average water depth of the sites is about 3,200 m to 3,400 m. They are located on the western flank of a larger topographic ridge. Mound structures are not visible at any of the inactive sites and are mostly weathered.

Figure 3.7: Overview of *Field 2* and its sites.



Field 3 extents over 3.3 km in north-south and 1.5 km east-west. Of its twelve sites, five are active (sites 1, 4-6, 8). Sites 1 and 4 have small mounds of 50 m to 100 m in diameter. Similarly, at sites 5 and 8 only a small mound of less than 30 m in diameter is visible. It is mostly covered by inactive and weathered chimneys and only diffuse venting was observed. Similarly, site 6 consists of an isolated patch of diffuse venting. Smooth mounds can be observed at sites 7 and 9. No such bathymetric structures can be identified at sites 3, 7, 10, and 11. At site 12 two small overlapping mounds with a combined diameter of max. 80 m can be noted.

Figure 3.8: Overview of *Field 3* and its sites.



Within *Field 4* there are 14 currently known sites, of which only one is active. The sites extend over 2 km in an east-west direction and about 3.2 km in north-south direction, in water depths between 2,600 m and 2,900 m. At most of the sites, mounds with typical dimensions of 100 m to 150 m in diameter and 20 m to 50 m in height can be observed. The sites are mostly located along the northern slope of a north-west to east trending ridge. The sites south of the ridge are more heavily weathered than those to the north.

Figure 3.9: Overview of *Field 4* and its sites.

Field 5 to *7* are in close proximity to each other. The sites of *Field 5* and *Field 6* are arranged linearly on the south-western slopes of two NW-SE trending flanks. Both are parallel to each other at a distance of 1.2 km. *Field 5* and *Field 6* have 14 and 16 sites respectively, of which 6 and 9 are active. A plateau extends further east where *Field 7* is located in about 1 km from *Field 6*. It was discovered during INDEX 2023 and ten inactive sites were confirmed by ROV dives during this cruise. However, there are a large number of possible sulphide mounds visible in the bathymetry that have not been verified yet. So there may be more sites, but this needs to be confirmed by future ROV dives. The water depths are very similar for all fields ranging between 2,300 m and 2,650 m. Overviews of the fields are shown in Figure 3.10 to 3.12.

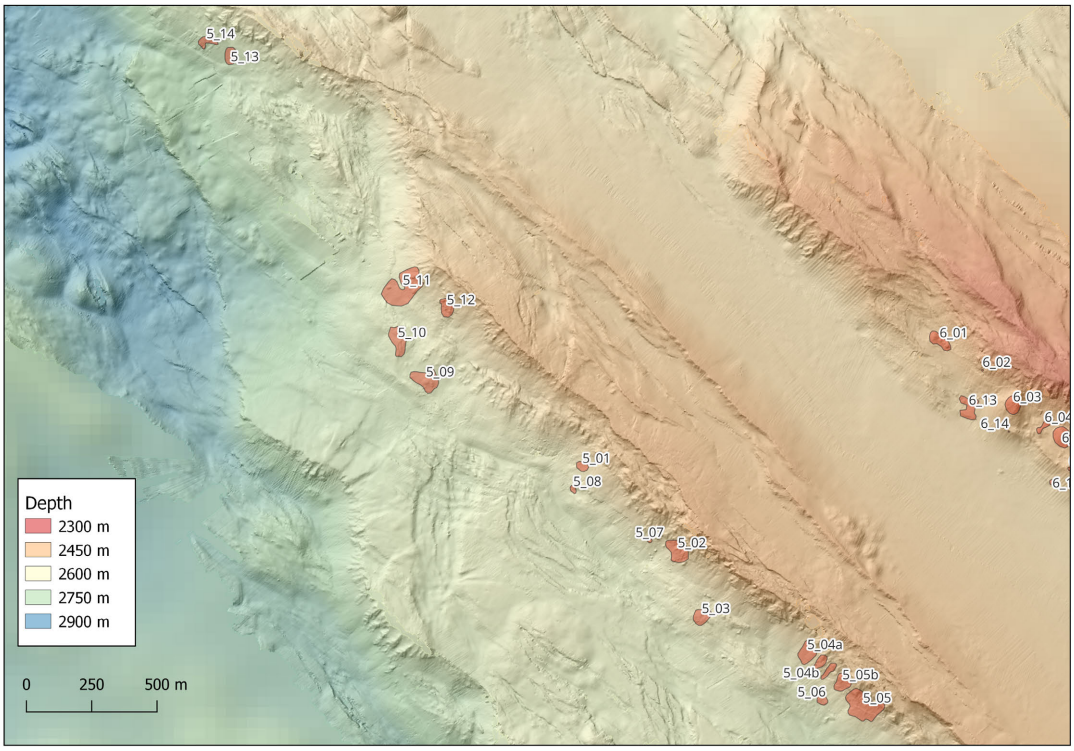


Figure 3.10: Overview of *Field 5* with sites.

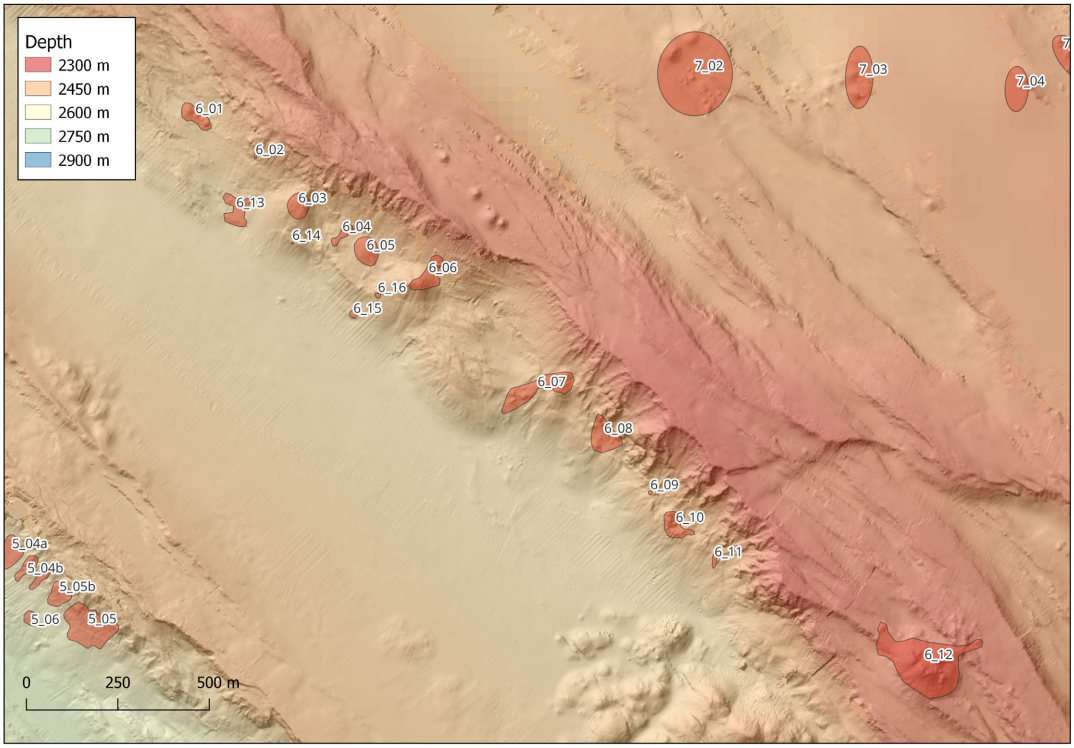


Figure 3.11: Overview of *Field 6* with sites.

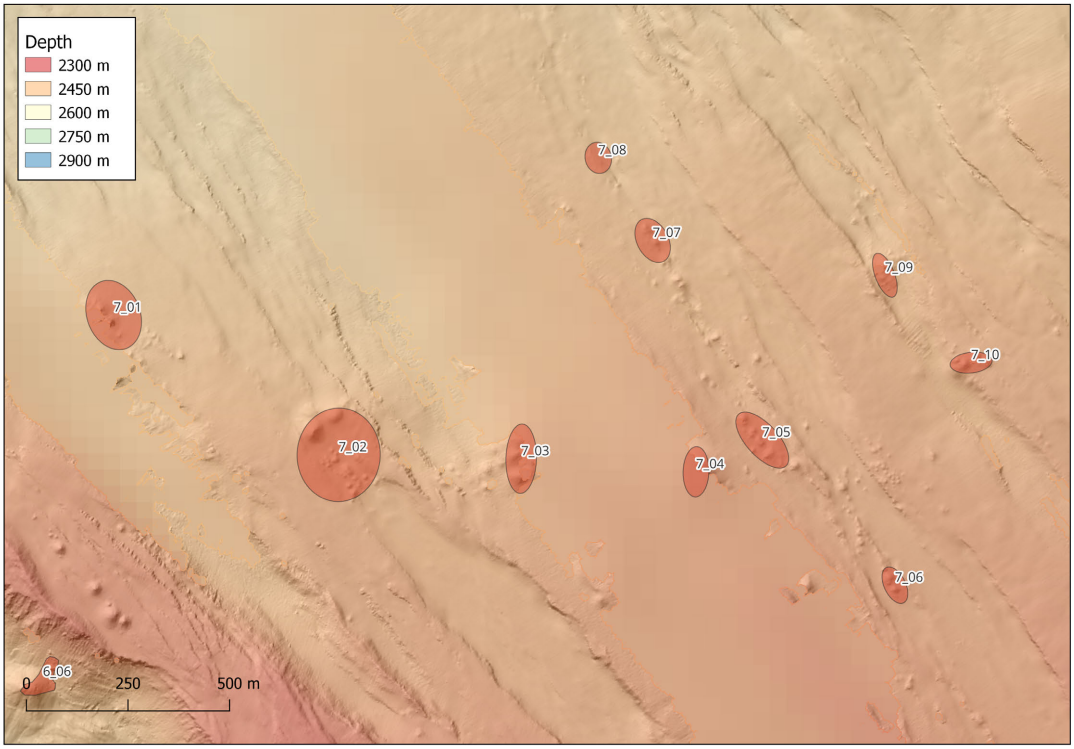


Figure 3.12: Overview of *Field 7* with sites.

3.2 High-Resolution MBES Data of Deep-towed Sled HOMESIDE

3.2.1 Instrumentation

HOMESIDE ("Hochauflösendes Multibeam-Echolot, Sidescan-Depressor"; english: "high-resolution MBES, sidescan depressor") is a deep-towed sled developed by BGR (Figure 3.13). It has been used since INDEX 2015 to obtain high-resolution bathymetric and backscatter maps. Its main acquisition sensor is a multi-beam echo sounder. HOMESIDE is towed behind the vessel at a speed of about 1 kn and an altitude of approx. 100 m above the seafloor. In addition to the MBES, various navigation and data acquisition sensors are installed for hydrothermal site detection. As the instrument is constantly being improved, some of the sensors have been used since the first deployments, while others have been added at a later stage. In this chapter, the individual sensors and the functionality of HOMESIDE are explained in more detail.

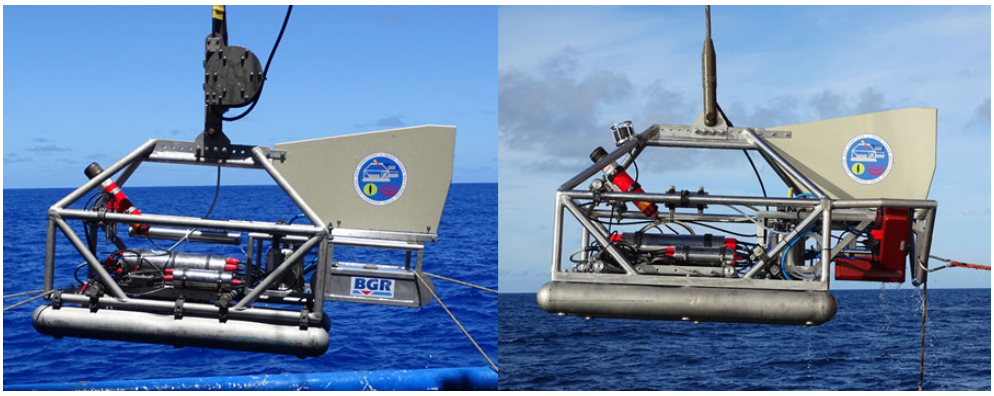


Figure 3.13: BGR HOMESIDE with single-RX setup during cruise INDEX 2018 (left) and with dual-RX setup during INDEX 2019 (right) (Bundesanstalt für Geowissenschaften und Rohstoffe, 2018, 2019)

The HOMESIDE MBES is a Kongsberg EM 2040. This MBES was used with a single receiver array (RX) from 2015 to 2018 and was upgraded to a dual-RX version before INDEX 2019. The EM 2040 allows the user to select a nominal measurement frequency of 200, 300, or 400 kHz. The angular across-track coverage is specified by the manufacturer as 140° for the single-receiver configuration. In the dual-receiver configuration and under normal conditions, an swath opening angle of 160°can be used. The beam opening angle varies between 0.7° x 0.7° (400 kHz) and 1.5° x 1.5°(200 kHz) and it has 256 preformed beam (Kongsberg, 2017). The number of soundings of a ping can be increased to 400 in “high-density mode” and the signal can be set to continuous wave or frequency modulated (for 200 kHz and 300 kHz only).

Since 2019, the typical HOMESIDE setup includes a string attached to the aft that detects self-potential anomalies – that is, anomalies in the general natural

electric potential field caused by electrochemical processes (Sharma, 1997). A buried sulphide deposit reacts with its surroundings. This includes oxidation-reduction processes which affect the potential and thereby the geoelectrical field, detectable as a negative anomaly – (i.e., self-potential anomaly)(Sato and Mooney, 1960). The string attached to HOMESIDE consists of electrodes whose voltage difference is measured. It is partially visible on the far right in Figure 3.13 right. This method has proved to be very valuable in detecting deposits of inactive hydrothermal sites that cannot be detected by the classical chemical sensors or the water column backscatter data (Kawada and Kasaya, 2017). Its disadvantage is its sensitivity to changes in altitude and its limited coverage. The towing altitude of HOMESIDE is typically about 100 to 120 m during regular deployments (e.g., with attached self-potential string and focus on increasing coverage and vent detection).

During some of the geophysical cruises within the INDEX project, HOMESIDE was used as a towing vehicle for deep-towed three-axial electric field receivers (Vulcans) (Constable et al., 2016; Müller et al., 2023). They work similar as the self-potential string, but as the name suggests, they allow three dimensional measurements of the electrical field vector. In these surveys, the altitude is reduced to about 50 m and the profile spacing to about 100 m to 150 m. Such deployments focus on further investigation of known sites and usually no additional bathymetric coverage is achieved. However, these data sets have a higher resolution. Overall, the resulting swath coverage of HOMESIDE depends on the acoustic frequency, the altitude, MBES settings, and local terrain. It varies from about 500 m swath width for regular high-resolution surveys for vent detection to approx. 300 m during ultra-high-resolution deployments with attached Vulcans.

As mentioned in Chapter 2.2.1, the positioning accuracy is crucial when collecting high-resolution MBES data. To meet the high accuracy requirements, an INS Exail PHINS 6000 is installed on HOMESIDE (Table 3.2). Its IMU (inertial measurement unit) also includes three accelerometers and three fibre-optic gyroscopes (FOGs) (Exail, 2016). The use of FOGs does not require any additional heading information. The additional navigation aiding sensors installed are the DVL Teledyne RDI Workhorse Navigator 600, which provides additional speed over ground and altitude information, and the CTD Valeport Midas SVX, which provides additional sensor depth information.

As the INDEX cruises always take place on different research vessels, HOMESIDE deployments have been carried out with different USBL systems. Table 3.3 lists the main USBL systems used for HOMESIDE data collection in recent years. Typically, each research vessel has its own USBL system installed. In 2018, BGR purchased the USBL Sonardyne Ranger 2, which was used during INDEX 2018 and from that onwards. Sometimes it was used as the only ab-

Table 3.2: Specifications of the INS Exail PHINS 6000 (Exail, 2018).

Position accuracy with USBL	Three times better than USBL accuracy
Position accuracy with DVL	0.1% of travel distance
Position accuracy with no aiding for 1 min / 2 min	0.8 m / 3.2 m
Heading accuracy with DVL/USBL	0.02° secant latitude
Roll and pitch accuracy	0.01°
Heave accuracy	5 cm or 5%

solute position source, sometimes only the ship system, and sometimes both systems together.

Table 3.3: Vessel and USBL systems used during HOMESIDE deployments.

INDEX Cruise	Vessel	USBL
INDEX 2015	R/V Pelagia	HiPAP 100
INDEX 2016	R/V Maria S. Merian	Sonardyne Ranger 2
INDEX 2017	R/V Sonne	Posidonia
INDEX 2018	R/V Pelagia	HiPAP 100 and Sonardyne Ranger 2
INDEX 2019	R/V Sonne	Posidonia and Sonardyne Ranger 2
INDEX 2021	R/V Pelagia	Posidonia and Sonardyne Ranger 2
INDEX 2022	R/V Pelagia	HiPAP 100 and Sonardyne Ranger 2
INDEX 2023	R/V Sonne	Sonardyne Ranger 2

All of the USBLs are deep-sea systems, operating with low acoustic frequencies and similar positioning accuracies. The navigation accuracy of HOMESIDE depends on the distance between the ship-mounted transceiver and the transponder in HOMESIDE, as can be noted from Table 3.4, which lists the characteristics of the main used USBL systems.

Table 3.4: Overview of USBL properties used in HOMESIDE deployments (Kongsberg, 2016; Exail, 2023; Sonardyne, 2014)

USBL	Acoustic Frequency	Range	Position Accuracy
Kongsberg HiPAP	TX bandwidth 10 – 12.5 kHz RX bandwidth 13 – 15.5 kHz	> 10,000 m	0.2% slant range
Exail Posidonia	TX bandwidth 8 – 14 kHz RX bandwidth 14 – 18 kHz	> 10,000 m	0.12% slant range (with OCTANS) 0.06% slant range (with PHINS)
Sonardyne Ranger 2	19 – 34 kHz	> 6,000 m	0.2% slant range (“typical”) 0.1% slant range (“optimised”)

The distance between the transceiver and the transponder varies depending on whether the data is recorded in a turn or at the beginning of a station (slant range of approx. 3,000 m) or after a longer period of profile survey (slant range of approx. 5,000 m). The positioning accuracy of the USBL can theoretically be expected to be between 3 m (0.1% slant range, 3,000 m) and 10 m (0.2% slant range, 5,000 m) under ideal conditions – provided the USBL is calibrated and the correct sound velocity is used. As HOMESIDE is equipped with an INS, the final navigation accuracy could theoretically be in the range of 1 m

to 4 m under optimum conditions. In practice, an accuracy of 10 m between overlapping profiles (i.e., 5 m in each direction) can be achieved. Mismatches of 20 m or more between overlapping profiles are corrected in post-processing.

3.2.2 HOMESIDE MBES Data Sets

This study focuses on the high-resolution bathymetry and backscatter data acquired with HOMESIDE of the seven hydrothermal fields and their surroundings presented above. Overall, four DTMs with a bathymetry of 2 m resolution and four backscatter mosaics with 1 m resolution have been used in this thesis. They cover all seven hydrothermal fields. They are compilations of HOMESIDE data from several years with possibly slightly different MBES settings and varying swath overlap. Table 3.5 provides an overview of the data compilations used, including the number of individual data sets and their year of acquisition.

Table 3.5: Overview of the data sets used, listing the hydrothermal fields, year of data collection, number of stations, and area covered.

Data Set	Field	Cruise	Number of MBES Deployments (Stations)	Area
A	Field 1-2	2015, 2017-2019, 2021-2023	30	272 km^2
B	Field 3	2018, 2019, 2021, 2023	31	186 km^2
C	Field 4	2022 & 2023	16	130 km^2
D	Field 5-7	2017-2019, 2021-2023	30	194 km^2
		SUM:	**107**	**782 km^2**

Combining multiple data sets, especially with different accuracy, is a crucial step in deriving comprehensive data compilations that can be used for an automated feature detection. Part of this study includes therefore the development of a processing workflow that ensures high quality data products. In Chapter 4 the processing workflow is described in more detail for the bathymetry collected in *Field 1* and *Field 2* and for the backscatter for *Field 1*. Some of the data was processed by other members of the working group. However, the development of the processing workflow was part of this study.

Within this study, *Field 1* was chosen as a sample data set for a closer analysis of the backscatter data in Chapter 4, because several HOMESIDE data sets of different years were collected there with a large overlap. This allows a detailed analysis of the MBES data acquired with different settings and in various directions. Therefore, the data of *Field 1* will be described in more detail below than the data of the other hydrothermal fields used in this work. However, the MBES settings and acquisition characteristics are also representative of the data collected throughout the INDEX area.

Figure 3.14 shows an overview of the MBES data sets with HOMESIDE track lines collected in the *Field 1*. Each deployment is divided into individual profiles – each consisting of a straight data line and is numbered chronologically. The main profile direction in *Field 1* is 120°, along the bathymetry. The station of 2015 was one of the first HOMESIDE deployments and therefore the track lines deviate from the optimum straight line configuration. In 2017, the MBES data set was extended southwards. A systematic mapping of the *Field 1* took place during two stations in 2019. Parallel and crossing profiles were placed according to the *Field 1* site locations.

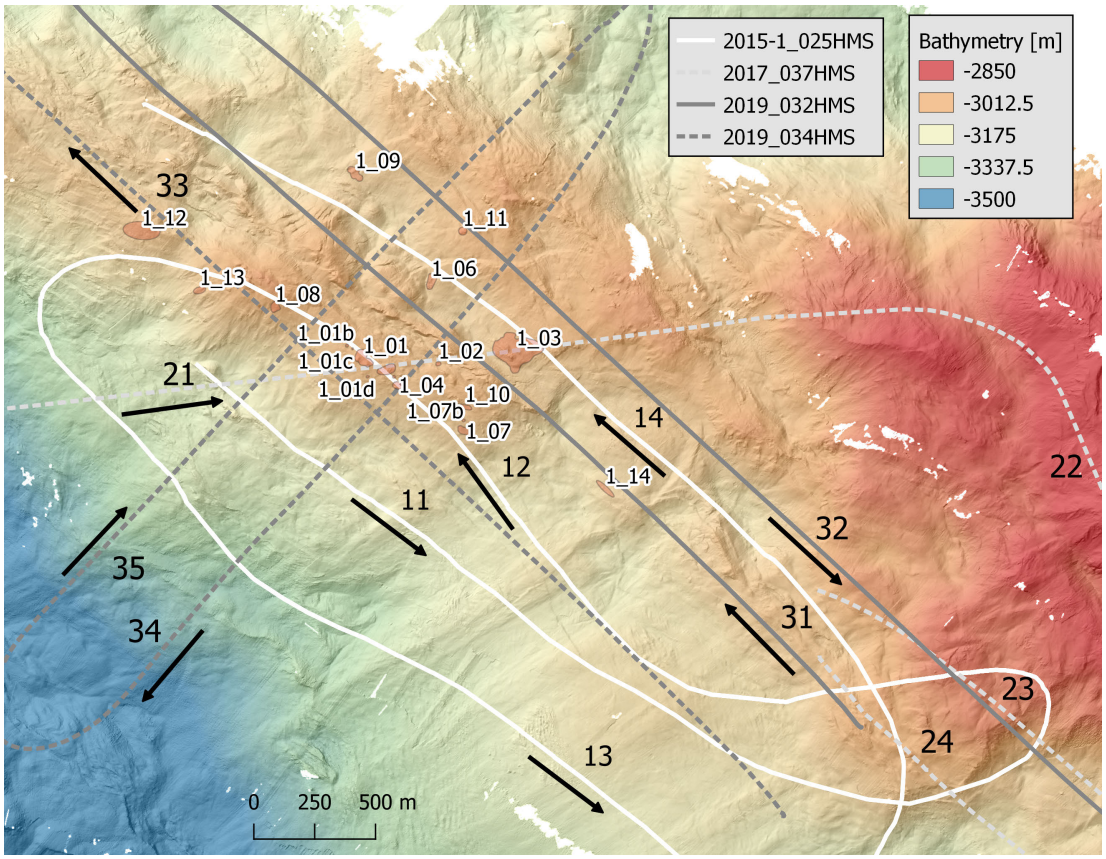


Figure 3.14: Overview of HOMESIDE tracks (except INDEX 2022 data) in *Field 1* with number and acquisition direction of the main data profiles used for detailed analysis.

Table 3.6 gives an overview of all data profiles with the used acoustic frequency, pulse width, and swath opening. Data profiles not used in the detailed backscatter analysis (Chapter 4.2.2) are coloured in grey. The MBES parameters, such as acoustic frequency or pulse type, can be changed during data collection to adjust to the current acquisition geometry (e.g., reducing the acoustic frequency to increase the range with larger altitudes). The characteristics of MBES data can therefore vary along a profile, but are typically attempted to keep fixed. In particular, at the beginning and in the end of a profile, the altitude is usually higher, as the distance from the HOMESIDE to the seabed is increased in turns to avoid the risk of bottom contact when the ship changes sailing direction and abrupt reductions in drag are expected.

Table 3.6 shows that in 2015 an acoustic frequency of 300 kHz and an FM pulse type are used during regular HOMESIDE deployments for hydrothermal site exploration and extension of bathymetric coverage. Profile 21 was recorded with the MBES "additional soundings" functionality enabled, which is only available for a CW pulse type. In 2019, the dual receiver, and therefore a larger swath opening, was available when carrying out the specific survey of *Field 1*. The preferred setting for this survey would have been 300 kHz and FM, but it was noted early in the data acquisition that HOMESIDE was slightly off the planned profile line. The settings were changed accordingly to increase the coverage and ensure that the venting sites were covered by the data. In 2022, the altitude was 50 m instead of approx. 100 m as in during regular surveys because the Vulcans were attached. As the distance to the seafloor was lower, the acoustic frequency could be increased to 400 kHz. There is no FM pulse available at 400 kHz.

Table 3.6: HOMESIDE data profiles recorded in the area of *Field 1* with the corresponding MBES settings. Data sets used for detailed backscatter analysis are coloured in black.

Station	Profile	Files	Dual / Single RX	Frequency	Pulse Type	Swath Opening	TX Pulse Length
2015-1_025HMS	11	0-18	single RX	300 kHz	FM	70° / 350 m	675 µm
2015-1_025HMS	12	36-50	single RX	300 kHz	FM	70° / 350 m	675 µm
2015-1_025HMS	13	58-83	single RX	300 kHz	FM	70° / 350 m	675 µm
2015-1_025HMS	14	91-111	single RX	300 kHz	FM	70° / 350 m	675 µm
2017_037HMS	21	2-18	single RX	300 kHz	CW	70° / 500 m	602 µm
2017_037HMS	22	21-43	single RX	200 kHz	FM	70° / 500 m	619 µm
2017_037HMS	23	46-68	single RX	200 kHz	FM	70° / 500 m	619 µm
2017_037HMS	24	73-92	single RX	200 kHz	FM	70° / 500 m	619 µm
2019_032HMS	31	1-14	dual RX	200 kHz	FM	80° / 500 m	517 µm
2019_032HMS	32	18-32	dual RX	200 kHz	FM	80° / 500 m	517 µm
2019_034HMS	33	1-15	dual RX	200 kHz	FM	80° / 500 m	517 µm
2019_034HMS	34	27-38	dual RX	200 kHz	FM	80° / 500 m	517 µm
2019_034HMS	35	43-56	dual RX	200 kHz	FM	80° / 500 m	517 µm
2022_084HMS	all	1-211	dual RX	400 kHz	CW	80° / 500 m	107 µm

3.3 Seabed Classification Based on Video Imagery

The video imagery collected by ROVs for visual inspection can be transferred to a geological map. This is usually done by geologists on board. As part of this study, an additional and finer scaled seafloor classification has been done for *Field 1*, focusing on the acoustic characteristics. The map provides the ground truth information for backscatter analysis. BGR provided the geological maps of all investigated seven fields as well as the initial geological classification of the ROV track, which is the initial step towards a geological map.

The imagery data of *Field 1* processed in this study was collected during INDEX 2015, INDEX 2022, and INDEX 2023 with the ROV ROPOS (Remotely

Operated Platform for Ocean Science) from CSSF (Canadian Scientific Submersible Facility). ROPOS is equipped with several cameras. In addition to a wide-angle pilot camera for navigation purposes, the primary camera is an *Insite Pacific Zeus-Plus* HD camera (10 x zoom 5.2 to 52 mm) and a secondary *Insite Pacific Mini-Zeus* HD camera. Both can be tilted by the operator. In addition, a 36.3 megapixel *Nikon D820* digital still camera with 14-24 mm AF-S f2.8 lens is installed to capture high resolution images of specific biological or geological interesting spots (CSSF, 2021).

In addition, visual data was collected during INDEX 2015 using the video sled from NIOZ (Royal Netherlands Institute for Sea Research). Unlike an ROV, a sled does not have its own propulsion system, but is towed behind the vessel. The NIOZ video sled is equipped with two cameras: a Sony full HD 1920x1080 pixels block camera (downward looking) and one Sony HD block camera (forward looking) (Bundesanstalt für Geowissenschaften und Rohstoffe, 2015). Both video platforms are equipped with two lasers 10 cm apart (ROPOS) for scaling and both systems are positioned by using USBL. ROPOS is additionally equipped with an *Exail Octans Subsea 3000 Gen4* for attitude determination.

A total of eleven ROV and three video sled stations were conducted within *Field 1* (Table 3.7). The video tracks are shown in Figure 3.15. Altogether, approx. 82 hours of video material were collected and used as ground truthing for backscatter data. This was done by visually inspecting the videos and classifying it according to the seabed type.

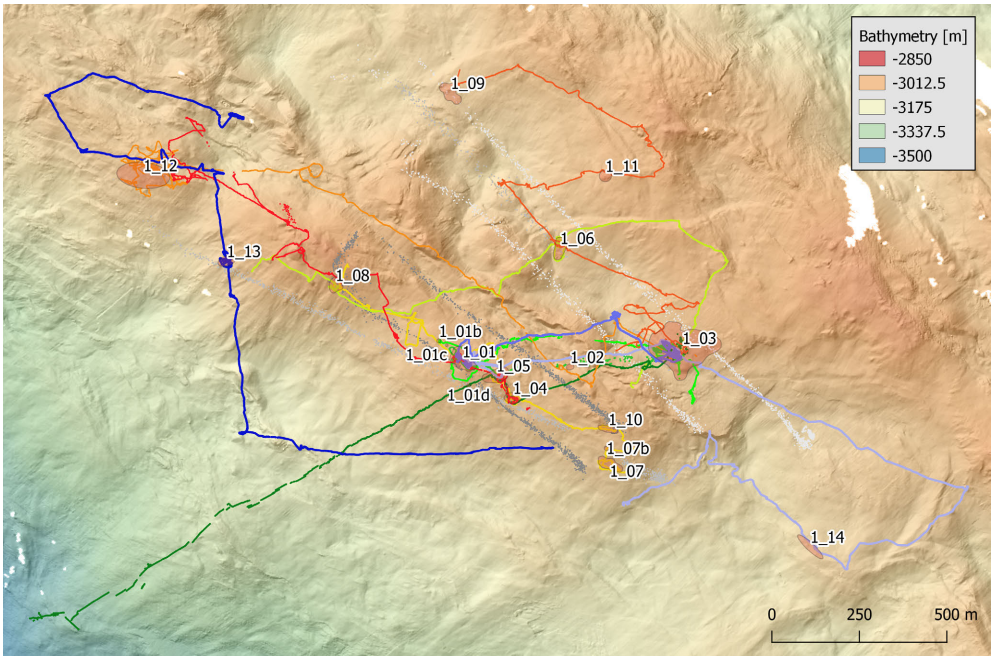


Figure 3.15: Overview of ROV ROPOS and video sled tracks within *Field 1*.

To create a geological map, a shapefile of line features was created for each dive. Along the video track, this shapefile was filled with lines coloured accord-

Table 3.7: ROV ROPOS and NIOZ video sled deployments in *Field 1* with start and end time (UTC) of video recording.

Station	Start Time (UTC)	End Time (UTC)
2015_037_ROV02	27.11.15 06:44	27.11.15 11:20
2015_041_VS-1	28.11.15 10:39	28.11.15 13:01
2015_041_VS-2	28.11.15 15:00	28.11.15 17:09
2015_043_ROV03	29.11.15 04:38	29.11.15 11:19
2015_045_ROV04	30.11.15 04:46	30.11.15 11:29
2015_046_VS-1	30.11.15 16:50	30.11.15 19:02
2015_046_VS-2	30.11.15 21:35	30.11.15 23:35
2015_047_ROV05	01.12.15 04:38	01.12.15 11:10
2015_048_VS-1	01.12.15 15:24	01.12.15 17:53
2015_048_VS-2	01.12.15 20:20	01.12.15 22:35
2015_049_ROV06	02.12.15 06:44	02.12.15 11:25
2022_069_ROV	29.11.22 05:09	29.11.22 12:39
2022_071_ROV	30.11.22 05:20	30.11.22 12:37
2022_073_ROV	01.12.22 05:17	01.12.22 11:48
2023_034_ROV	08.10.23 05:52	08.12.23 12:02
2023_037_ROV	09.10.23 06:43	09.10.23 12:29
2023_049_ROV	13.10.23 05:48	13.10.23 12:00

ing to the seabed type visible in the videos and images. The 2015 ROPOS videos were not time-stamped and the html protocol (containing time-stamped pictures and descriptions) was used to correlate the visual information with time and hence the location. The bathymetric map from the HOMESIDE data was used as a basis and it also helped to locate the imagery data, as navigation accuracy of the 2015 data in particular was low. As mentioned above, the classification results were used for an MBES backscatter analysis. The focus regarding the seabed classes was therefore lying on the acoustically distinguishable seabed types. Figure 3.16 shows examples of the selected classes are.

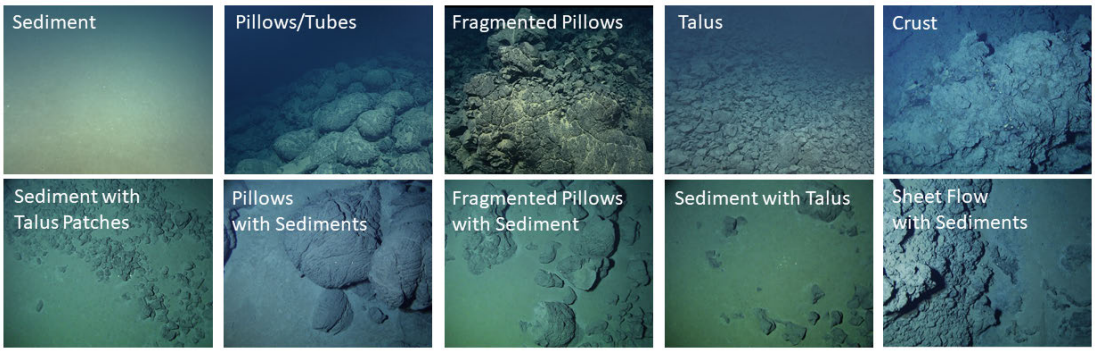


Figure 3.16: Distinguished seabed types of the video based classification of *Field 1*.

The main seabed classes are *sediment*, *talus*, *pillow*, and *crust*. In addition, mixed types, such as *fragmented pillows* – a combination of *pillow* and *talus* – have been classified. *Pillows*, sometimes referred to as tubes depending on their shape, are formed by the cooling of lava flows. These distinct pillow and tube shapes are formed as the surface material cools down more rapidly while the lava below continuous to flow. They are of basaltic origin and can reach sizes of about one metre and more (at least in *Field 1*). Due to their fine structure, they are generally porous and easily fractured. The fragments of these basaltic flows are called basaltic talus, an can range in size from a few centimetres to metres. Whereas “basaltic” provides information about the material and origin, the term “talus” provides information about the structure and refers to the fractured material in general. Unless otherwise stated in this study, talus typically means basaltic talus.

Sediment of varying thickness covers the entire study area. Accumulations of sediment typically occur in depressions or between pillows and are caused by currents. Areas where the sediment completely covers the underlying structures are classified as *sediment*. *Crusts* are formed by the chemical accumulation of metals dissolved in seawater. They can occur anywhere in the ocean. However, the local chemical composition of the seawater and hydrothermal fluids at hydrothermal active areas encourages the formation of hydrothermal crusts. These have a distinct yellowish to reddish colour.

Areas with evidence of sulphides or active/inactive hydrothermal activity are distinguished from other classes and further subdivided into *sulphide block*, *hydrothermal crust*, and *hydrothermal sediment*. Sulphide blocks typically originate from fallen chimneys and are often covered by hydrothermal crust. Areas of reddish and yellowish colouration are classified as hydrothermal sediments, indicating the possible presence of hydrothermal and iron oxhydroxides components. The final seabed classification is shown in Figure 3.17.

BGR provided geological maps of seven hydrothermal fields for this study. Compared to the map above, these maps have a higher level of generalisation in term of seabed classes. Maps for geological interpretation must allow for a broader inspection and therefore have to include larger structures and a more general overview of the area. Therefore, these maps include geological landforms and structures that are essential for geological analysis. The geological map created specifically for the backscatter analysis includes mixed classes, allowing for a more detailed examination of the visualisation of individual seabed types in the backscatter data (Chapter 5).

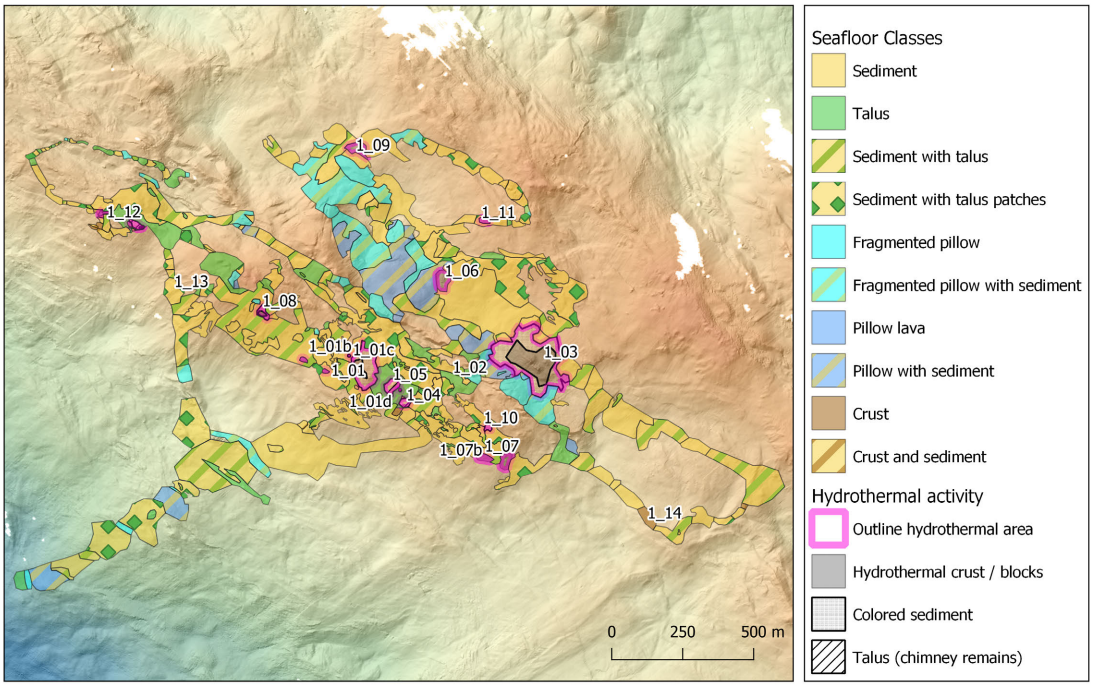


Figure 3.17: Results of visual seabed classification based on video and imagery footage of ROV ROPOS and NIOZ video sled in the area of *Field 1*.

Chapter 4

MBES Data Processing

Post-processing is the process of checking, correcting, and improving the acquired data to ensure high quality derivatives are created in the final stage. In this chapter, the processing workflow of MBES bathymetry and backscatter data, its improvement, and the generation of final products is described in more detail using the example of the data set collected in the hydrothermal venting *Field 1*.

4.1 Bathymetry Processing

Parts of this chapter have been presented a the "FIG Working Week 2024" and published as conference paper in the "International Hydrographic Review".

Dufek, T.; Freitag, R.; Kuhn, T.; Sternberg, H. (2024): Exploration of hydrothermal venting sites using deep-towed multibeam echo sounder data. *International Hydrographic Review*, 30(1), pp. 128-134. DOI: 10.58440/ihr-30-1-c01.

During an MBES survey, a number of sensor data is collected simultaneously from a moving platform. These include at least MBES, positioning, attitude, and SSP data. The individual sensor data is merged in a first step and then checked to ensure that each data set is valid and conclusive. If necessary, each sensor data can be improved individually and then updated for the overall data and products, auch as DTMs, produced in the final step. The following subchapters describe the general workflow as well as the more detailed one for HOMESIDE and ship-borne bathymetric MBES data.

4.1.1 General Bathymetry Processing Workflow and DTM Generation

In general, the MBES raw data is stored in a manufacturer specific data format. Depending on the software used for the processing, the data is converted

into a software specific format during import. Typically, the individual sensor data is not collected at the same moment in time, as each has its own measurement frequency. Therefore, the first step is to interpolate the data from the auxiliary sensors so that each MBES measurement has position and attitude information describing the orientation and location of the platform at the time of measurement. All the data information is then merged, taking into account the lever arms and SSP, to derive a 3D point coordinate in a global coordinate reference system. Corrections within the multi-sensor system setup of the (e.g., changing of sensor coordinates within the sensor system) can be applied at this stage if necessary, or also updated data (e.g., corrected navigation data) can be imported. SSP measurements can be re-applied to the data using different methods to those used during acquisition.

Typical MBES software allows each sensor data to be visualised individually so that it can be checked for outliers. The combined result is also inspected for errors by visualising the sounding data as terrain model in a regular raster. This facilitates the identification of gross or systematic errors, especially in overlapping areas of the swath. Different editors allow the sounding data to be viewed from various directions and erroneous depth values to be removed by flagging them as “disabled/rejected”. They will then be ignored in the subsequent product generation. This flagging can be done manually or by applying a predefined filter. The settings of such automatic filters must be adjusted to the individual data sets and their properties such as sounding density or water depth.

Once the data has been corrected, final products can be derived by gridding the soundings, which are spatially inhomogeneously distributed data points, into a regular spaced raster/grid. These resulting digital terrain models (DTMs) can then further be used in various visualisation or GIS (Geographic Information System) software for presentation and further analysis. A DTM should represent the seabed as accurately and realistically as possible. Various algorithms are available for the calculation of such a DTM. Typical approaches are based on calculating the DTM cell value by taking into account the soundings within a certain geometric radius - typically in the unit of the number of neighbouring raster cells. The calculation method can be basic, such as the mean, median, shallowest, or deepest value of all soundings considered. Or weighted approaches such as *Weighted Moving Average (WMA)* can be used. In this the weight of the influence of each sounding considered depends on its proximity to the location of the raster cell centre for which the depth value is being determined. An algorithm with a different approach is the *CUBE (Combined uncertainty and bathymetric estimator)* algorithm. It was developed as a research project of the Center for Coastal and Ocean Mapping and NOAA/UNH Joint Hydrographic Center (CCOM/JHC) at the University of New Hampshire. It is an error-based gridding method that takes into account the total propa-

gated uncertainty (TPU) of a sounding. The TPU is the accumulated a priori error taking into account all error sources of the sensors and their performance during data acquisition (i.e., survey system, auxiliary sensors, survey platform configuration). Following the law of error propagation, the total vertical and horizontal uncertainties (TVU and THU) are determined for each sounding. Further details can be found in Calder and Wells (2007).

Each of these calculation algorithms are accompanied with a number of user-defined parameter settings. Irrespective of the DTM calculation method chosen, the user always has to define the raster cell size and therefore the resolution of the DTM. On the one hand, the DTM resolution should be chosen as small as possible to allow small features to be still visible within in the DTM. On the other hand, the cell size should not be chosen too small, as this could result in gaps, which should be avoided. The DTM resolution should be chosen according to the density and beam footprint size of the data set. The footprint size defines the resolution of the data set and a cell size smaller than the data resolution will not improve the resolution of small seafloor features. The other parameters depend on the chosen gridding algorithm and should be set taking into account the data and environmental characteristics.

4.1.2 HOMESIDE MBES Bathymetry Processing

A major focus of the HOMESIDE data processing is the navigation correction, as underwater positioning is more prone to errors than GNSS-based ship navigation on the water surface. Regular USBL reception losses are caused by unfavourable acquisition geometry (e.g., slant range reaching maximum range, multipathing, non-optimal orientation of transceiver and transponder). Due to the high resolution of the MBES data, even small offsets in the navigation can lead to visually unpleasant and disadvantageous effects in the resulting DTM. As the focus is on the identification and interpretation of sulphide mounds, which have only small dimensions of 150 m or less in diameter, the reduction of artefacts that can overlay such small bathymetric features is crucial.

The INS navigation of HOMESIDE is post-processed using the processing software *Exail DELPH INS*. It allows to view the individual sensor data, to discard certain time-based data in case of failure, to edit the navigation data, or include additional sensors that were not integrated during data acquisition. If the “optimised” processing option is selected, a three pass algorithm (forward-backward-smoothing) is performed, where the data is processed in the forward direction in the first step (i.e., the data is integrated chronologically and filtered by Kalman filter). Afterwards, this process is then repeated in reverse chronological order, starting with the last data received. In the third step, both results are merged (Exail, 2015).

Bathymetric data processing was performed using the software *QPS Qimera*. After importing the raw MBES files in *.all format, the navigation processed with *Exail DELPH INS* is added as a *.txt file and merged with the raw sonar data by time stamp. The INS post-processing software was only available in the INDEX project from the middle of the cruise in 2017. As a result, the data sets collected in 2015, 2016, and partly in 2017 showed larger navigation errors and had to be first manually corrected in *QPS Qimera*. These corrections included removing jumps (rejecting certain navigation values and interpolating in between the remaining ones) or, in the case of offsets greater than 50 m, the entering of new navigation points using other high-resolution or ship-borne bathymetry as a reference. Figure 4.1 shows the navigation of one of the first HOMESIDE dives in 2015 with the resulting bathymetry before (left) and after (right) manual navigation correction.

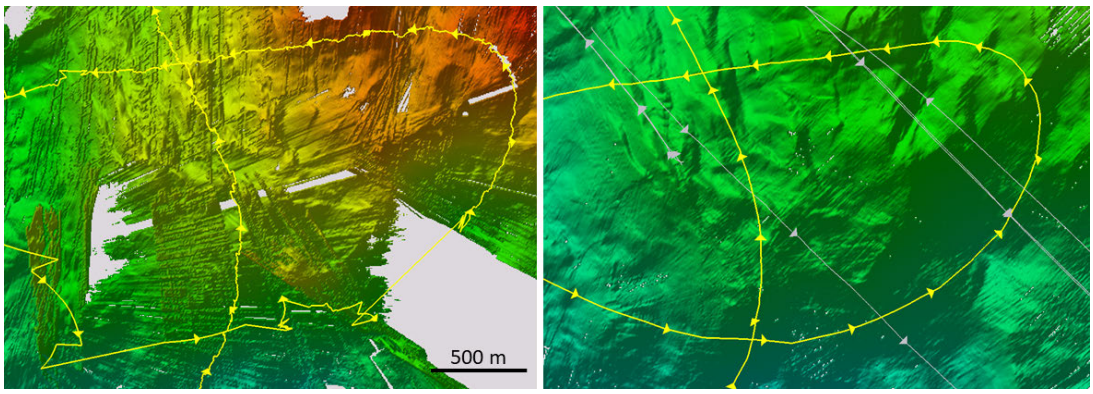


Figure 4.1: Visualisation of one of the very first HOMESIDE dives in 2015 showing the navigation by arrows before (left) and after (right) manual navigation improvement.

Systematic USBL offsets caused, for example, by a wrong SSP or incorrect calibration values in the USBL system cannot be automatically detected or corrected by the INS processing software. It assumes that the absolute USBL values are correct as there is no comparison with true values. The overall navigation solution is smoothed and from small jumps and drifts are removed. But an overall systematic offsets of up to 50 m may still be present. Typically, such offsets caused by inappropriate USBL settings are not static. Their magnitude varies according to the slant range between the ship (transceiver) and platform (transponder), and their direction depends in the direction of acquisition.

To detect such systematic offsets that exceed the required positioning accuracy, it is necessary to have overlapping HOMESIDE data acquired in different directions. If larger offsets are detected, they are first corrected manually in *QPS Qimera*. Since even smaller offsets in the range of 20 m can cause artefacts in the final DTM, the manual adjustment using adjacent profiles or ship-borne data has shown to be insufficient. Therefore, the open-source software *MB-System* is used when there is a certain overlap between adjacent data sets and only small offsets are present. *MB-System* is a software to

process and visualise MBES and side scan data and was developed by the Lamount-Doherty Earth Observatory of Columbia University and is now maintained by the Monterey Bay Aquarium Research Institute (MBARI) University of New Hampshire and MARUM (MBARI, 2017). *MB-System* is included in the Poseidon distribution based on Ubuntu Linux. *MB-System* is a package of several programmes that can be executed in a terminal window. The main component od *MB-System* used for this work was the graphical programme *mbnavadjust*. This tool allows the manual and/or automated relative navigation adjustment of data files based on features present in overlapping areas. In this work, *Poseidon 9* with *MB-System 5.5.2315* was used within Oracle VM VirtualBox 7.0 on Windows 10. The processing workflow including *MB-System* for the navigation correction is shown in Figure 4.2.

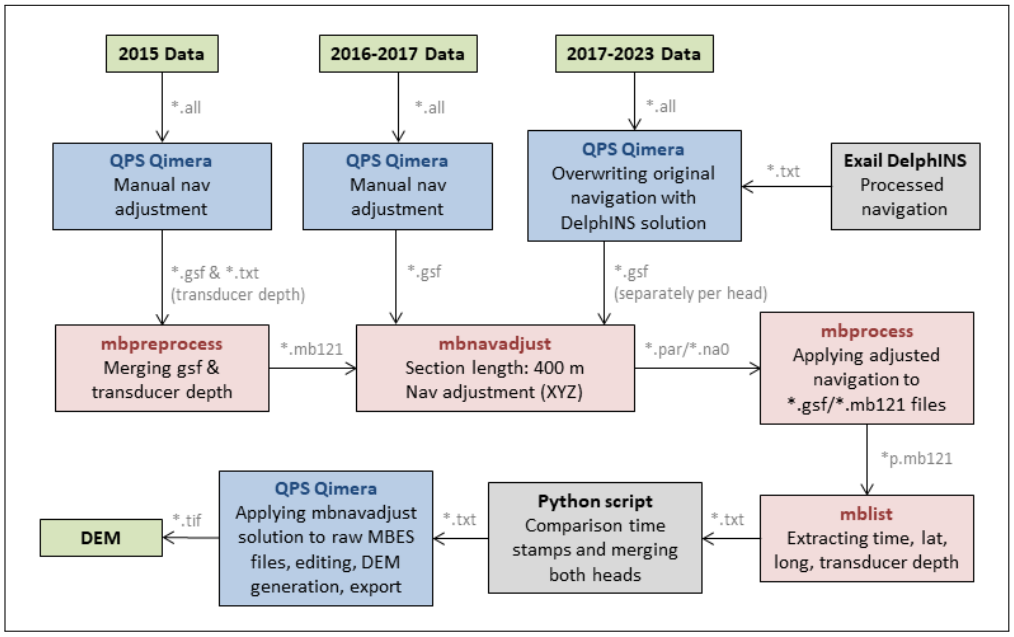


Figure 4.2: Processing workflow for HOMESIDE bathymetric data using *QPS Qimera* and *MB-System* for navigation adjustment (own illustration.)

To prepare the data for *MB-System*, the MBES raw *.all files and the adjusted *DELPH INS* navigation (only available from 2017 onwards) are imported into *QPS Qimera*. If necessary, the navigation is adjusted manually (correction for small jumps or drifts when no *DELPH INS* processing possible, or larger systematic offsets). The data is then exported in *.gsf format (generic sensor format). For the HOMESIDE EM2040 dual-receiver configuration, each head is exported separately as otherwise *MB-System* cannot interpret the data correctly. For the 2015 data, the transducer depth is additionally exported in text format and applied to the MBES data using the programme *mbpreprocess*, as erroneous sensor depth information is automatically exported in the original *.gsf files due to an initial incorrect sensor assignment. The MBES data with the updated sensor depth was saved in MB121 format, which is an *MB-System* format for *.gsf files.

When importing the data import into *mbnavadjust*, the ten minute data acquisition files, which are about 350 m long, are split into sections of a user defined length. In order to maintain the approximate file length, the section length is set to 40 m. Ten navigation control points are automatically defined at regular intervals for each section. This results in one control point every 40 m of navigation track for this data set, corresponding to intervals of about one minute of data acquisition.

The sections are analysed for the percentage of overlap (crossings). Where the track lines of two sections directly cross each other, theses are referred to as true crossings. The crossings are listed and the user can select a crossing and open the visualisation of the corresponding sections, the track line, and the bathymetry within the NavError display (Figure 4.3). The depth information is displayed as contours for each section. They are visualised with the current navigation. Offsets are therefore visible by mismatching contours. The NavError display also shows the RMS (root mean square) of the horizontal and vertical misalignment ("misfit") between the two sections (in the bottom left corner). These two RMS displays only apply to the current extent of the data and are therefore updated when zooming in or out.

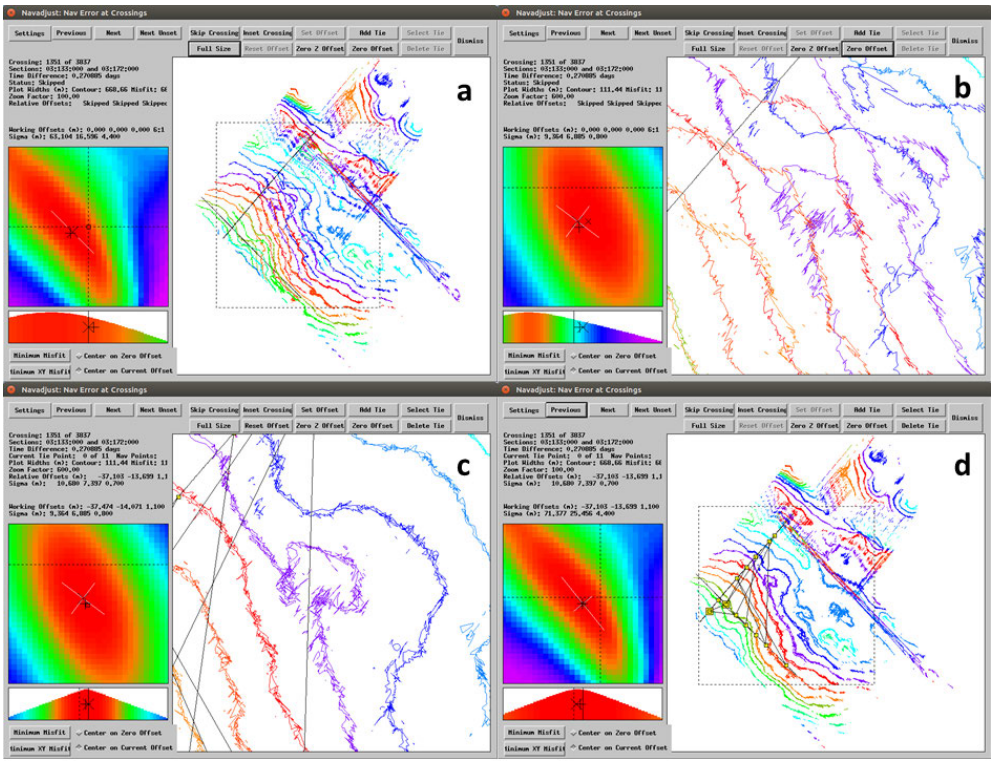


Figure 4.3: Different stages of navigation adjustment for the overlap of two sections in *MB-System*: a) overview with zero offset, b) detailed view showing a distinct bathymetric feature with zero offset, c) after automatic 3D adjustment (minimum misfit offsets: -37.5 m / -14.1 m / 1.1 m), d) overview after adjustment showing the tie points as lines connecting corresponding navigation control points (yellow).

To adjust the navigation, prominent bathymetric features in the overlap of two

sections are selected. That part of the data is zoomed in and the automatic 3D offset correction (minimum misfit) determines the navigation offsets between the navigation control points based on the lowest RMS. The track navigation and thereby the contours are automatically adjusted to these values, allowing visual verification by the user. These offsets can be further adjusted manually by moving the sections horizontally until the contours are aligned. A tie point is set to assign the final offset to the navigation control points. The offsets for each tie point and the corresponding navigation control points are then saved. Between one and eleven tie points are set for each crossing, depending on the amount of overlap and visibility of bathymetric features.

Figure 4.3 shows the different stages of the navigation adjustment in the NavError display are depicted: Figure 4.3a shows two sections with contours and track without any offsets applied. The contour lines are not aligned. Figure 4.3b and c depict a bathymetric feature before and after navigation adjustment. After the adjustment of -37.5 m in X, -14.1 m in Y, and 1.1 m in Z, the contours of the two sections of the 2019 data (INDEX2019_32HMS, head 2) coincide. Figure 4.3d shows the crossing after navigation adjustment. The tie points are depicted as lines connecting the navigation control points (yellow squares). For each crossing, as many tie points as possible are defined, homogeneously distributed over the overlap.

The MBES data set of *Field 1* was chosen as main focus and workflow development for this work, as several data sets collected in different years overlap. In addition, an ultra-high resolution survey was carried out in 2022 (see Chapter 3.2.2). The INDEX high-resolution processing workflow is organised so that data collected in a certain area (i.e., the same cluster and the same side of graben flank) is processed together. Therefore, the data over *Field 1* was not processed alone, but together with the data sets of *Field 2*, as the data sets are not only in close proximity, but also overlap. In total, the data set consists of over 3,271 data files and 23,398 crossings. All crossings with more than 10% overlap were individually checked (12,393) and tie points defined where possible (24,293). Figure 4.4 shows an overview of the bathymetric data with indicated tie points and navigation control points for crossings of more than 50%.

An inversion is then performed by *mbnavadjust* to determine the optimal navigation solution based on the defined tie points and adjustments of the navigation control points. As some of the defined offsets may be in conflict, a least squares adjustment (Paige and Saunders, 1982) is performed (MBARI, 2017). The navigation quality of individual surveys (i.e., consecutive data files without temporal interruption) can be defined by the user as poor, good (default), or fixed. Data defined as good is used regularly in the inversion. Fixed data will not be assigned an offset, and poor data will be adjusted to the previously

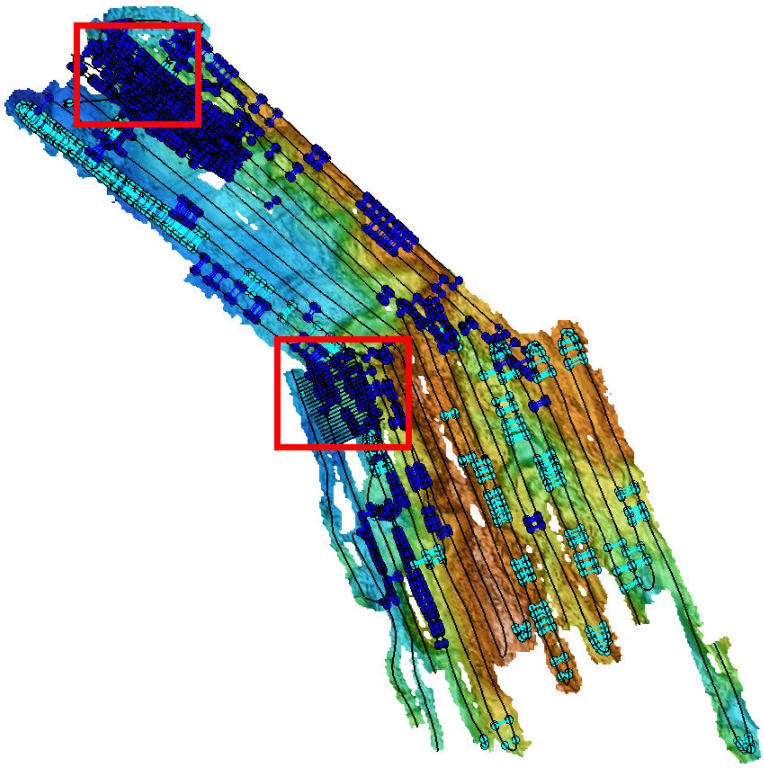


Figure 4.4: Overview of the MBES data set with navigation control points (circles) and tie points (lines connecting corresponding navigation control points) for crossings with more than 50% overlap. Cyan coloured circles and lines represent ties within a survey (a block of data with no time gap) and dark circles and lines represent ties between two different surveys. The locations of the hydrothermal fields are indicated by red boxes (*Field 1* in the north, *Field 2* in the south).

already corrected data (i.e., *good* data) at a later stage, without explicitly influencing the overall adjustment. All individual data sets are compared with existing ROV dive trajectories in the *Field 1* during INDEX 2022. The ultra-high-resolution HOMESIDE data collected in 2022 shows the smallest offsets and one of the profiles is selected as fixed. Data sets with large offsets were set as *poor*. Table 4.1 shows the assigned quality status of the navigation status for *MB-System*.

Table 4.1: List of HOMESIDE data sets collected in the area of *Field 1* and *Field 2* with their status for the navigation adjustment in *mbnavadjust* in *MB-System*.

Data set	MB Navigation Status
2015	poor
2016	good
2017	good
2018	poor
2019	poor
2021	good
2022 (except file 99–128)	good
2022 (file 99–128)	fixed

After the inversion, the statistical results can be examined. If necessary, individual ties can be changed before the inversion is repeated. Figure 4.5 shows

the *mbnavadjust* navigation adjustment model for longitude, latitude, and depth (from top to bottom) for the navigation points by time series. Vertical green lines distinguish the surveys, which are defined in the order of import. The maximum values of the lateral offsets reach almost 90 m. This occurs mainly for the very first HOMESIDE stations of 2015, which had a very poor navigation quality and which were manually adjusted before *MB-System* based on the ship-borne bathymetry. The 2015 data also shows a systematic deviation of about 10 m in depth from the other data sets. As can be also seen in the model plot that the 2018 data set has large east-west offsets that were due to erroneous USBL calibration values. The fixed surveys (i.e., one for each head) are marked with red circles in the model plot. They have no offsets.

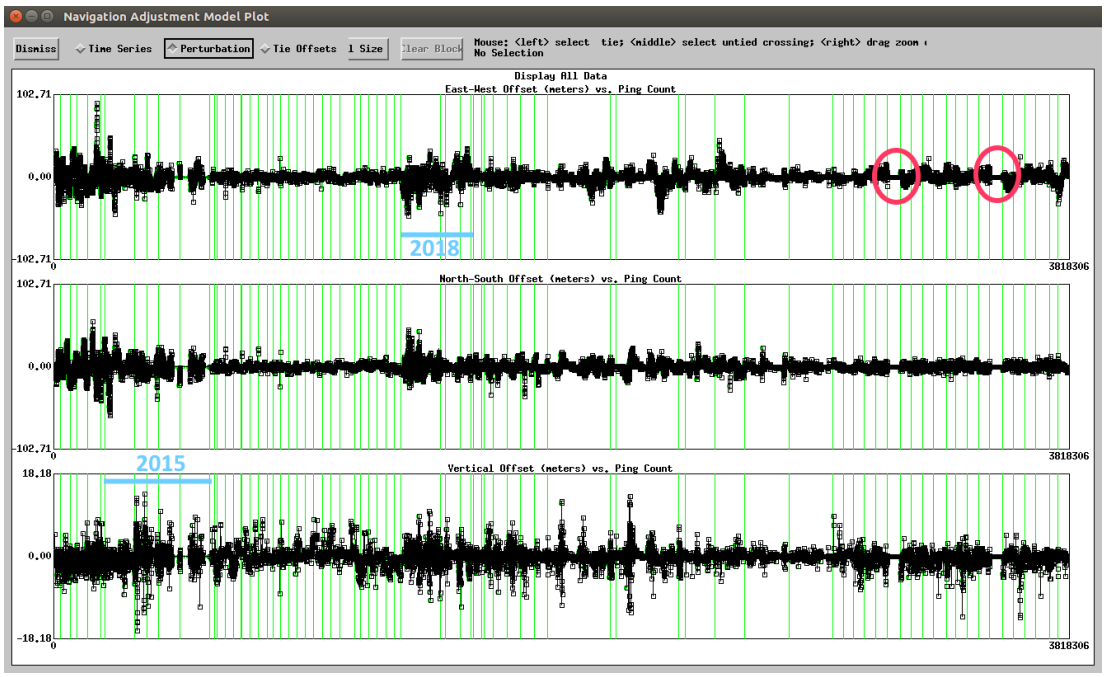


Figure 4.5: *MB-System mbnavadjust* model plot after inversion showing the offsets in longitude (top), latitude (middle), as well as depth (bottom) for each ping in a time series.

The final inversion results are saved by *mbnavadjust* within an adjusted navigation file (*.na0) and an *MB-System* parameter file (*.par) for each MBES file. The parameter file stores all changes made by any of the *MB-System* processing tools. The *mbprocess* programme merges the adjusted navigation and parameter file with the MBES file, resulting in a processed file in *MB-System* format (*p.mb121). From these processed files the navigation information (time, lat, long, sensor depth) is extracted by using *mblist*.

The results for both heads are combined by first checking adjusting the time stamps for missing entries and then averaging the values for xyz. The derived navigation data is then applied to the raw MBES data in *QPS Qimera* for manual inspection and removal of remaining misalignments in overlaps and general editing. Figure 4.6 shows the high-resolution bathymetric data collected in the

area of *Field 1* and *Field 2* at different stages of processing in an overview (top) and in more detail for *Field 1* (bottom). The data is coloured by uncertainty with a 95% confidence interval (2-sigma) to visualise the improvement from navigation adjustment. Figure 4.6 left shows the data with the original navigation retrieved in real-time or by post-processing with *DELPH INS* – if available. The middle depiction shows the result after applying the *mbnavadjust* result. Figure 4.6 right shows the final result after manual removal of the remaining mismatching overlaps. With the original navigation, large discrepancies of 20 m and more are visible in large areas of the data set. These have been significantly removed in the right plot, with uncertainty values barely exceeding 10 m.

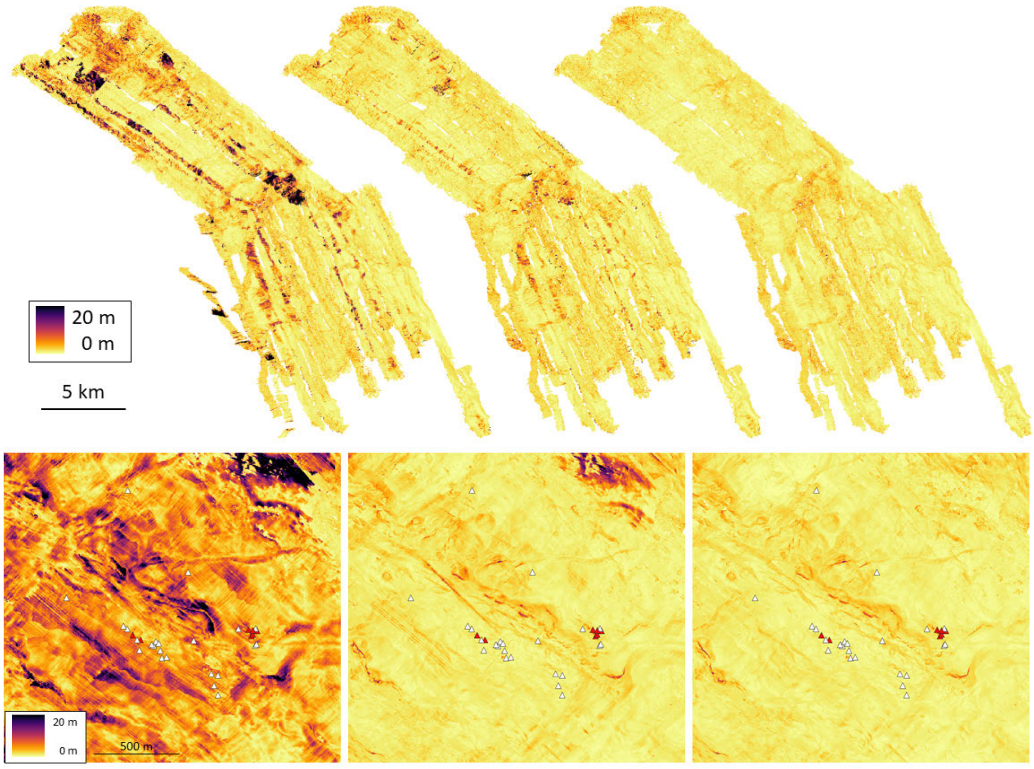


Figure 4.6: Uncertainty of the HOMESIDE MBES bathymetry compilation (2 m resolution, at 95% confidence level) at different stages of navigation processing: partially edited with no additional navigation improvement (left), with final *MB-System* navigation (middle), and finally edited (right) for the whole data set (top) and in a detailed view of *Field 1* with indicated site locations (bottom).

The importance of improving the navigational data becomes even more apparent when viewing the high-resolution bathymetry of *Field 1* with different navigation solutions in 3D, as shown in Figures 4.7. If the data overlaps and does not match within a few metres, artefacts are apparent in the elevation model. These artefacts can overlay small bathymetric features that are difficult to identify as a result. This is particular important as the resulting DTMs are used to analyse and identify possible sulphide mounds within the size of such small bathymetric features, which are adversely affected by the artefacts

caused by mismatching navigation.

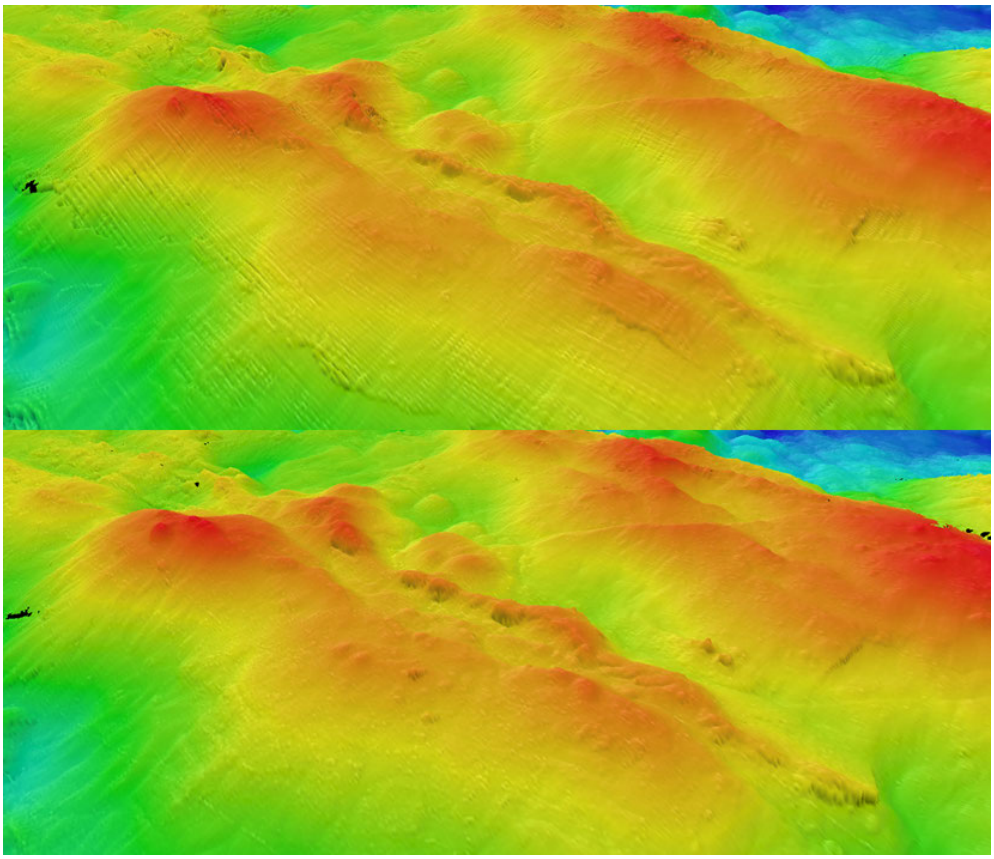


Figure 4.7: 3D view of the DTM (2 m resolution) for *Field 1* before (top) and after (bottom) navigation adjustment.

The final step in the processing is to produce digital terrain models. A DTM of the HOMESIDE MBES data within the INDEX area should highlight small bathymetric features as they may represent sulphide mounds and thus indicate current or past hydrothermal activity. On the other hand, data artefacts (such as those caused by mismatches due to navigation offsets) should not be highlighted and the number of gaps must be as small as possible. This should be done without interpolating and thus smoothing the data too much. These requirements are partially contradictory and therefore the choice of grid cell size and gridding method is not trivial.

As mentioned in Chapter 3.2.1, the beam opening angle of the Kongsberg EM2040 on HOMESIDE varies between 0.7° x 0.7° (400 kHz) and 1.5° x 1.5° (200 kHz) and therefore the general footprint size of the combined data ranges between 1.5 m (0.7°, nadir) and 12 m (1.5°, @60°) for an average altitude of 120 m. The theoretical distance between two soundings is about 0.5 m in along-track direction and about 1 m in across-track direction (without overlap). Taking all this into account, a DTM resolution of 2 m is chosen as this resembles the lower range of the footprint size and is still larger than the average sounding density, ensuring that there is at least one sounding in most of the grid cells. By using a small resolution with respect to the footprint diameter,

small-scale bathymetric features are still represented in the data.

Various gridding algorithms have been tested for the HOMESIDE data and, as in theory, the data shows that each of them has its advantages and disadvantages. Therefore, the HOMESIDE data is generally gridded using four different algorithms: *mean-interpolated* and *shallow-interpolated* (taking into account only the soundings within the corresponding cell), *WMA 3x3* (3x3 referring to the 8 by 8 cell matrix surrounding the according cell, to which all soundings contribute according to the weight assigned to them, which depends on their distance), and *CUBE*. Figure 4.8 shows a 3D visualisation of a 200 m diameter sulphide mound in *Field 6* created using different gridding algorithms.

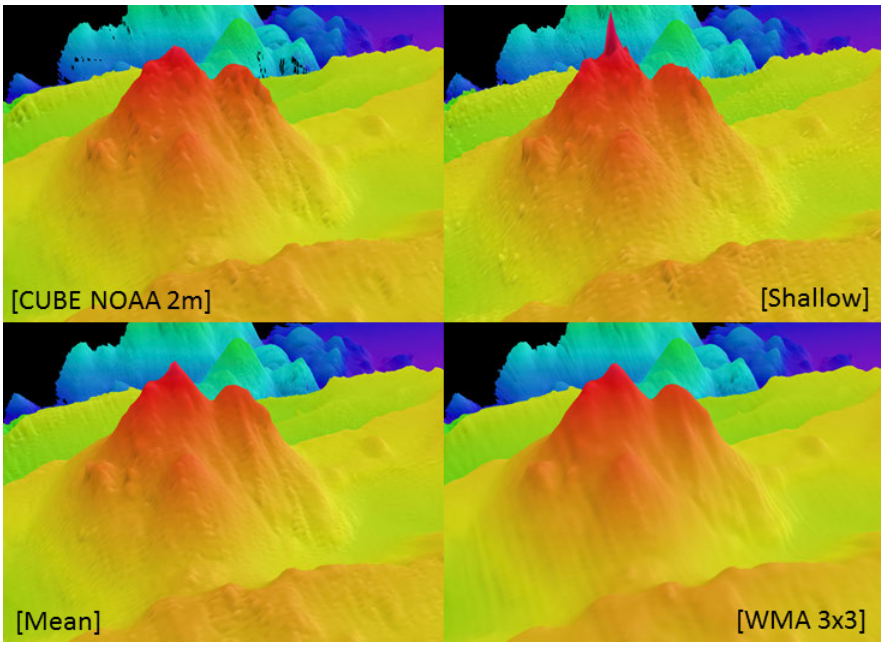


Figure 4.8: 3D view DTMs showing a sulphide mound (200 m in diameter) created using different gridding algorithms.

The advantage of the *mean* and *shallow* algorithms is that they only consider soundings within the corresponding raster cell, so there is less smoothing. In particular, the *shallow* method highlights local bathymetric highs that may represent chimney structures. On the other hand, these approaches show some gaps, as the chosen cell size is within the scale of the sounding density. Therefore, these DTMs had to be interpolated to obtain a continuous terrain representation. *WMA 3x3* produces a closed surface with less noise and artefacts, but the highlighting of small bathymetric features is significantly reduced. This representation, like *mean-interpolated*, is good for a more general overview of the survey area, whereas the *shallow-interpolated* is more useful for close inspection of chimney and mound structures where artefacts are not posing a problem. *CUBE* seems less suitable for HOMESIDE data as the noise is high and systematic mismatches occur often in overlaps. Such erroneous data is clearly visible. Small chimney features are highlighted, but there are also peaks

within the mismatching overlap that can easily be misinterpreted as chimneys. This is because *CUBE* works with hypotheses. While *mean* averages the data, *CUBE* can also obtain two or more possible raster values (hypotheses), one of which is selected depending on the sending. If the decision for the hypothesis changes between neighbouring cells, visible peaks are created. The advantage of *CUBE* is that the number of gaps is small, although there is no strong smoothing as with *WMA*.

4.2 HOMESIDE Backscatter Processing

Important seabed information can be obtained from MBES backscatter data. It provides similar information to side-scan sonar data, but has the advantage of being acquired together with the MBES bathymetry data. Therefore, no additional survey is required to have depth information available. About 20 years ago, the *Geocoder* as a set of algorithms to process MBES backscatter data, has been developed by the University of New Hampshire (Fonseca and Calder, 2005). Shortly after, this tool was implemented in the major commercial software and MBES backscatter processing became more common. The processing is usually highly automated and the influence of the user is very limited. Since the processing is not standardised and certain information is not disclosed, backscatter processing is generally of a “back-box” character. This chapter describes the general MBES backscatter processing workflow before discussing the in the derived mosaics from different software to obtain the best processing result of the HOMESIDE MBES backscatter data.

4.2.1 General MBES Backscatter Processing Workflow

As mentioned in Chapter 2.2.3, the general idea of backscatter data analysis is to derive seabed characteristics from the received acoustic signal after removing all influences that are not dependent on seabed characteristics. MBES backscatter processing involves a number of steps to generate the final product for further analysis and interpretation is retrieved. Typically, two different products are used in backscatter analysis: angular response curves and mosaics. Angular response curves show the variation in backscattering strength for different incident angles for one or numerous consecutive pings. Such angular responses vary for different seabed types, and Angular Range Analysis (ARA) can be used to infer the local sediment grain size. Mosaics are georeferenced seafloor images and show the backscattering strength for the area of interest as if the whole data was collected with an acoustic signal of the same angular incidence.

The process of obtaining angular response curves and mosaics, which involves the removal and correction of unwanted influences on the backscatter data, is

not standardised in detail and varies between the MBES manufacturers, MBES models, MBES modes, and software. The general workflow consists of the following steps: data decoding, georeferencing, radiometric correction, angle dependence removal, pre-mosaicking corrections, and mosaicking (Figure 4.9). In typical commercial processing software, all the necessary corrections are applied automatically by the software, and the user has only little opportunity to intervene. Usually only certain parameter settings can be adjusted.

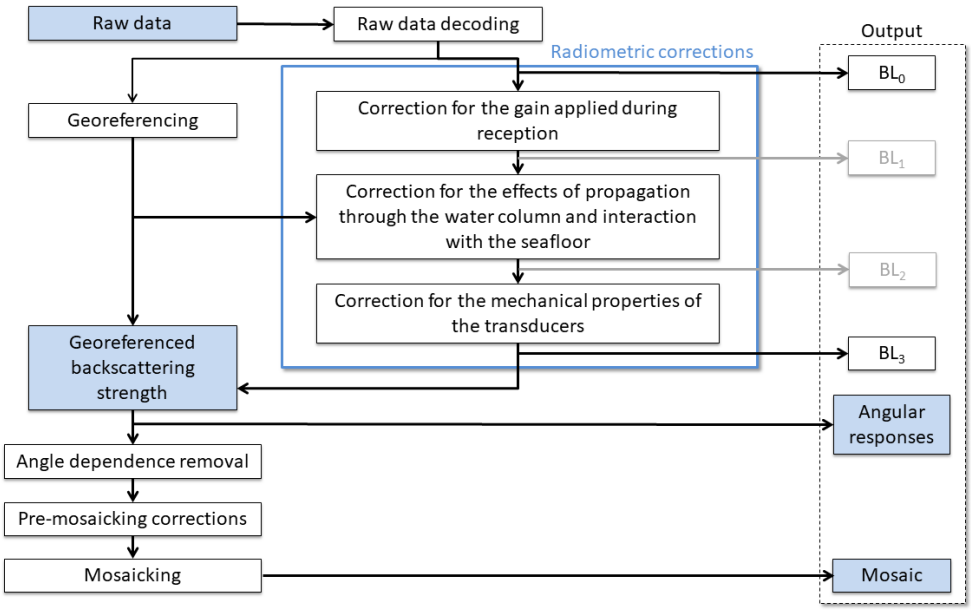


Figure 4.9: MBES backscatter processing workflow showing the individual corrections from raw data to angular response curves and mosaics (after Malik et al. (2019); Schimel et al. (2018))

Ten years ago, the GeoHab group (Marine Geological and Biological Habitat Mapping Group) was formed. The Backscatter Working Group (BSWG) is part of it and was formed to focus on the standardisation in the acquisition and processing of MBES backscatter data. The work of this and other groups has led to a more open discussion and exchange of information on backscatter processing. A good example of this is the *Backscatter Processing Software Inter-Comparison Project*, which has resulted in the ability to extract certain intermediate backscatter processing products now from commercial software (named BL_0 to BL_3 in Figure 4.9) Malik et al. (2019). This allows the user a better understanding of the adjustments done by the software. A more detailed description of the theoretical steps in backscatter processing is given below to show that it is not trivial and to highlight that many corrections are performed by the software without user intervention.

The first step in the backscatter processing workflow is **raw data decoding**, where the backscatter data is retrieved from the raw manufacturer files. Backscatter information can be stored in a variety of formats. For Kongsberg MBES – like on HOMESIDE – two formats are available: the *beam average* format

stores the intensity value for each beam and the *time series* format consists of a series of intensity samples for each beam. After decoding, the backscatter level BL_0 is retrieved, which is the raw imported backscatter data before any correction has been applied.

As with bathymetric data, the raw data is only relatively referenced to the transducer. So in the next step – **georeferencing** – the derived backscatter values are georeferenced in a global coordinate reference system, taking into account GNSS/USBL/INS positioning, attitude, sensor alignment, sound speed profile, and MBES geometry. This process is generally quite straightforward for the beam average format, as the position of the beam on the seafloor is already known from the bathymetry processing. For other formats, such as the beam time series format, additional processing steps are required to georeference the individual samples. One approach uses the location of the beam from bathymetric processing (as for the beam average format) and then interpolates the bathymetry between adjacent beams for georeferencing. Another approach converts the beam time series to a half-swath time series and then applies a slant range correction, as in side-scan sonar data (Schimel et al., 2018).

The **radiometric corrections** involve the removal of unwanted dependencies from the received backscatter level to retrieve data that depends only on the incidence angle, the acoustic frequency, and the seabed properties. According to Schimel et al. (2018), three types of radiometric corrections can be distinguished. The first is the gain correction applied during the reception, which accounts for all modifications to the signal between reception and recording. This is primarily a time-varying gain (TVG), which accounts for the attenuation of the signal over time. Such gains can be dynamically adapted to system settings (e.g., pulse length) or environmental characteristics (e.g., depth). Further gains are applied during the signal processing. Gain designs vary between manufacturers, MBES models, and MBES modes. They are often based on assumptions or estimates and are not always well documented. After the correction for the gains applied during reception, the received level BL_1 is retrieved (Schimel et al., 2018).

The second group of radiometric corrections are those that address the transmission loss of the signal as it propagates through the water and interacts with the seabed. The transmission loss depends on the range and the absorption coefficient (see Equation 2.11). The latter depends not only on the frequency, but also on the current local properties such as temperature, salinity, and pH, which vary with depth. The absorption coefficient can be modelled for different depths based on CTD measurements, but generally an average value is used to correct for transmission loss (Lurton and Lamarche, 2015). The area of insonification can be estimated using various models of different complexity.

According to Schimel et al. (2018), they range from assuming a flat seabed to models that take into account refraction, pitch, transmit steering angles, and the 3D normal vector to the seafloor are taken into account (Beaudoin et al., 2004). After applying these corrections, the resulting level BL_2 is obtained, which depends on the backscattering strength and the mechanical characteristics of the transducer.

The mechanical characteristics of the transducers to be corrected, are the source level and the transmit as well as receive beam pattern (directivity modulations). It is the third group of radiometric corrections. A generic value for the source level is often provided by the manufacturer, unlike beam pattern information, which is only rarely available. Such generic values are obtained from tank tests and therefore the actual values for an MBES are likely to differ due to variations from the prototype and abrasion over time. In order to apply a thorough and accurate correction for the mechanical characteristics of the transducer, a calibration should be performed. Due to the complexity of such a calibration, it is not usually conducted, but most software allows for a *beam pattern correction* that accounts for the source level, beam pattern, and sensitivity of the receiver array (Schimel et al., 2018). To do this, data collected over a flat and homogeneous surface is compared to a model and the deviations from the model are taken into account for the entire data set. After removing the mechanical characteristics of the transducer, BL_3 is obtained or, if georeferenced, the angular response curves are derived and can be used for analysis.

To retrieve a mosaic, the **removal of angular dependence** must be carried out. Otherwise, artefacts along the swath would be present in the mosaic, showing a generally lower intensity in the outer beams than for the inner ones at nadir. This normalisation in across-track direction is called AVG (angle-varying gain) and takes into account the incident angle of the signal on the seafloor. This angle has already been determined in the correction for insonified area. However, as the angular dependence is also influenced by the type of seabed, certain adjustments have to be made to the physical models. Since the seafloor type is usually unknown, the correction is typically determined empirically based on the data itself. This is done by selecting a data subset of consecutive pings. This subset is then divided into angular bins and the average signal level for each bin is calculated. A reference level is determined at a reference angle. The differences between the average angular response and the level at the reference angle are then used for the correction (Lurton and Lamarche, 2015).

The **pre-mosaicking corrections** improve the visual quality of the mosaic to be generated. This usually involves downsampling, as the original data has a higher resolution than the derived mosaic, and to reduce aliasing. Another important tool is the low-pass filter to minimize noise in the image (de-speckle).

Finally, within the **mosaicking**, the angel independent and filtered backscattering strength is compiled into a georeferenced image of the seabed. This includes the gridding process, where the backscatter samples are combined into an equally spaced grid. The resolution of the grid cells and the algorithm used to deal with overlapping samples are the most important aspects to consider in this step. These can often be influenced, at least to a certain extend, by the operator. Other image enhancement algorithms can also be applied to the grid at this stage, as in pre-mosaicking. Afterwards, the decision has to be made how to store the values (8 or 16-bit scale) and how to adapt the colour scale is adjusted to the mosaic values (Schimel et al., 2018).

This sequence of corrections described above can vary between different commercial backscatter processing software as the processing is not standardised. In addition, different models, algorithms, and assumptions have to be made and applied throughout the process, and these can also vary between software. As a result, the mosaics from the same data set are likely to vary when produced by different software. To assess the magnitude of these differences for the INDEX data collected with HOMESIDE, a comparison is made between different MBES backscatter processing software.

4.2.2 Software Comparison for HOMESIDE MBES Backscatter Data Processing

Backscatter mosaics are one basis for the seabed analysis and the classification of sulphide mounds in this study. The results are therefore highly dependent on the quality of the derived mosaics. Especially when using automatic approaches such as machine learning, it is important to derive of homogeneous seafloor images with minimal artefacts are important. In order to investigate how large the differences between software are when using the same processing steps and how much they influence the final mosaics, a software comparison was performed. To investigate the differences within the steps of the workflow using different software, the intermediate processing stages of BL_0 (beam average after decoding) and BL_3 (after radiometric correction) were extracted as in Malik et al. (2019). A first analysis compares the results for these backscatter levels. A second one focuses on the comparison of the mosaics, which are the main results of the backscatter processing of the HOMESIDE data. The mosaics are compared in terms of remaining artefacts and representation of different seabed classes derived from the video footage.

The commercial software *QPS FMGeocoder Toolbox 7.9.5 (FMGT)*, *Teledyne CARIS HIPS and SIPS 11.3.17 (CARIS)*, and *IFREMER SonarScope (SSc) R2019b_2020-09-29 / R2019b_2021-02-20* (Augustin, 2023) were used in this comparison. *FMGT* has been used for years to process the INDEX data. *Teledyne CARIS* is also one of the main hydrographic processing software. Like

the other two, *SonarScope* is also a commercial software, but is available for free under an academic non-commercial user licence. It has also been widely used in publications of the GEOHAB group and offers some backscatter processing tools that exceed the capabilities of other software. This comparison was carried out in late 2020 / early 2021.

4.2.2.1 Comparison of Intermediate Backscatter Level

For this comparison, four data files were used that overlap and cover areas of hydrothermal venting sites (Table 4.2). These files were also acquired with different acoustic frequencies and signal types, and with different MBES setups (single and dual receiver). They thus represent the variety of characteristics of the HOMESIDE data sets collected over the years.

Table 4.2: HOMESIDE data files used for the comparison of BL_0 and BL_3.

Station	File	Receiver Type	Frequency	Pulse Type
INDEX2017_37HMS Line 1	11 & 12	single	300 kHz	CW
INDEX2019_34HMS Line 4	32 & 33	dual	200 kHz	FM

As the export of intermediate backscatter levels is not a typical data product, there are certain limitations in retrieving this output. *CARIS* offers several backscatter processing engines (*SIPS Backscatter, SIPS Backscatter With Area Based AVG, SIPS Side Scan,* and *Geocoder*) but the export of intermediate backscatter level is only available for the *SIPS backscatter engine*. Therefore, this one was used in this comparison. It was observed that some of the functions within *SonarScope* were only possible to execute with version 2020, while others were only possible with the latest version 2021. The software version had to be changed during processing. Unfortunately, neither version was able to calculate the incident angle for the 2019 data. A possible reason for this is the dual-receiver setup. Therefore, only the 2017 data could be used for this comparison.

Radiometric corrections are based on information stored in the data or are provided by the manufacturer. Some details of the corrections are only disclosed between the MBES manufacturer and the software producer. Therefore, the user's ability to intervene in this stage of processing is very limited. The operator can influence the transmission loss corrections by applying a DTM or by setting the absorption coefficient. Table 4.3 shows the correction settings with the corresponding parameters. These settings can be selected and changed by the user. In all software, the corrected navigation and transducer depths were imported as well as the final bathymetry. The time series format was used as data type.

Figure 4.10 shows the derived backscatter levels before (BL_0, top) and after

Table 4.3: Backscatter radiometric corrections settings for the comparison of BL_0 and BL_3 in *FMGeocoder Toolbox*, *CARIS (SIPS backscatter engine)*, and *SonarScope*.

Software	Data Type	Absorption	DTM
FMGT	Time Series	[applies values stored in raw data]	v
CARIS	Time Series	Default: Temp: 8 Salinity: 35	v
SSc	Time Series	[via own algorithm according to depth]	v

(BL_3, bottom) the radiometric corrections for the two test data sets extracted using different software. For the files collected during INDEX 2017 (11-12), gaps can be observed in the *CARIS* output. These gaps of 500 pings occur systematically in each file and are therefore assumed to be caused by a software error in the export, as they so not occur in the final computation of the mosaic. The backscatter levels before radiometric correction BL_0 look similar in all software. The levels obtained for BL_0 by *CARIS* seem to be slightly higher than those of *FMGT* and *SSc*. At this stage only the decoding has been done and even smaller differences might not be expected at this stage. Larger discrepancies can be observed after the radiometric corrections have been applied in BL_3. Also after the corrections, the obtained backscatter level seems to be higher in *CARIS* than in *FMGT*. The values derived by *SSc* are even lower.

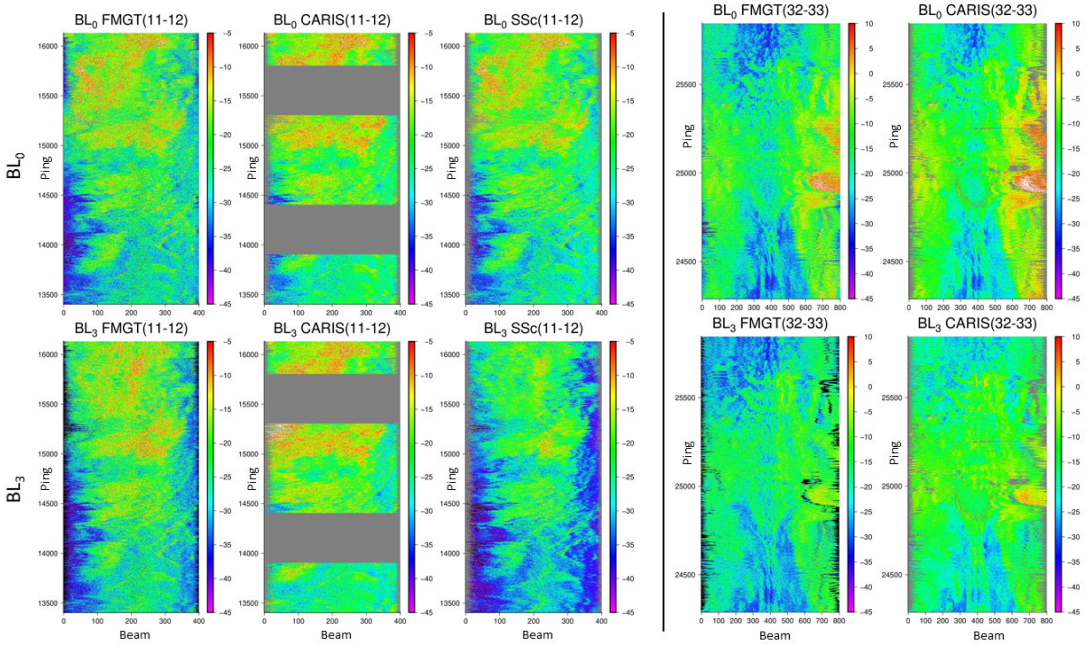


Figure 4.10: Intermediate backscatter levels BL_0 and BL_3 for file 11 and 12 collected 2017 (left) and file 32 and 33 from 2019 (right) obtained from different software.

For the derived backscatter levels, the mean value for each beam (top) and

ping (bottom) is shown in Figure 4.11. The previously made observation is also clearly visible here: *CARIS* (in blue) provides higher backscatter levels than *FMGT* (in green) and especially *SSc* (red). Comparing BL_0 (lighter colours) to BL_3 (darker colours) for each software, *FMGT* and *CARIS* show a similar pattern: The values in the nadir are slightly increased, while the values for the outer beams are decreased during the radiometric correction. This is particular evident for the 2019 data, as the the dual-receiver configuration enhances the coverage and therefore the certain effects. An exception to this regularity are the beam average values of INDEX 2017 data in *CARIS*. On the port side, the backscatter values are increased by the radiometric correction.

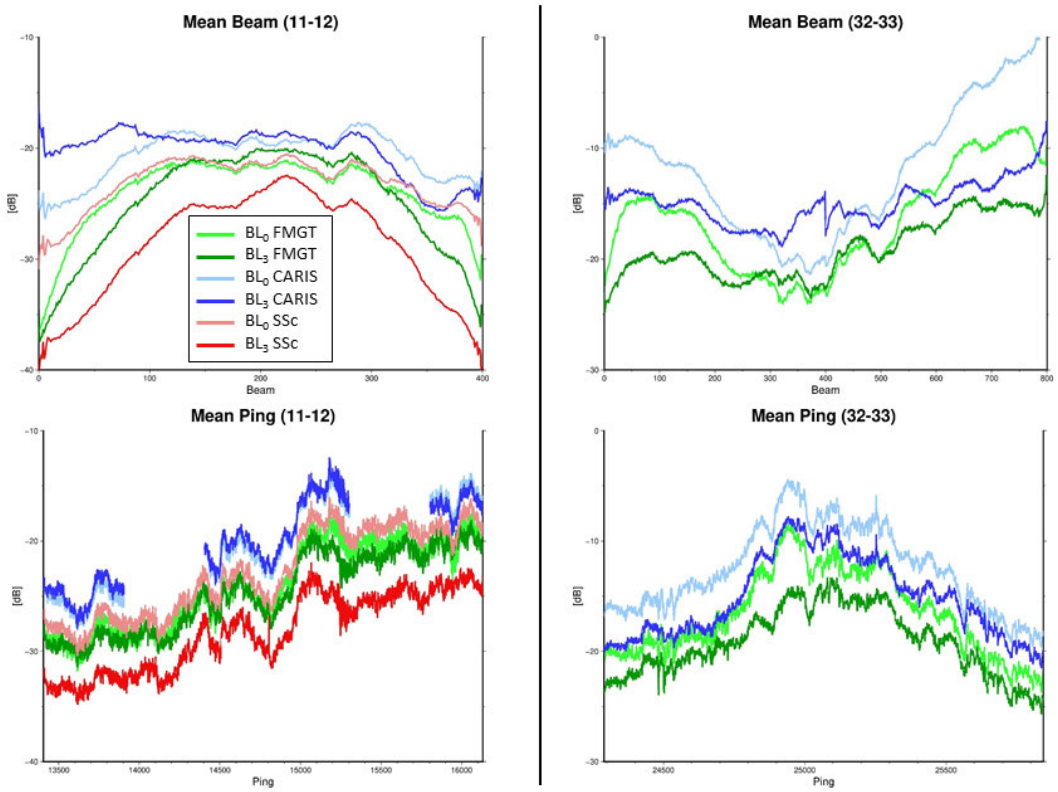


Figure 4.11: Mean backscatter level BL_0 and BL_3 calculated for the beams (top) and pings (bottom) for files 11 and 12 (INDEX 2017, left) and files 32 and 33 (INDEX 2019, right).

The mean values of the backscatter levels over the pings (Figure 4.11) allow an overall interpretation of the backscatter level. For the 2017 data, the backscatter levels BL_0 and BL_3 are relatively similar in *FMGT* and *CARIS*. This indicates that the increase and decrease of values within a swath are balanced. The previously noticed stronger decrease in mean values for the outer beams of the 2019 data from BL_0 to BL_3, is also noticed in the ping plot: The backscatter levels after the radiometric corrections are significantly lower than before the correction (bottom right). The change in backscatter levels in *SonarScope* is different from the other two software (i.e., decrease in outer beams and increase in inner beams). The decrease in values is relatively homogeneous

over the swath and higher by about 5 to 10 dB.

Figure 4.12 shows the differences between BL_3 and BL_0, and thus the changes in backscatter level due to the radiometric corrections for each software and data set. The colour scales are the same for all images. The bathymetry is also depicted. The visualised backscatter levels show a similar pattern (as noted already in the graph of the mean beam values): There is an increase in backscatter level at the nadir and a decrease at the outer beams. However, the spatial pattern is very similar for *FMGT* and *SonarScope*, but very different for *CARIS*. *CARIS* shows a stripe-like pattern in along-track direction, as if a certain correction is being applied consistently to a particular beam interval. *FMGT* and *SonarScope* show a less homogeneous distribution but a more turbulent pattern, but the correction values of adjacent beams can vary greatly. There is also a certain correlation with the bathymetry. It seems that the bathymetry and especially small bathymetric changes have a greater influence on the correction. Unfortunately it is not possible to investigate this further as the intermediate levels (BL_1 and BL_2) are not available. Most likely, the larger differences in *CARIS* are due to the specific algorithms used by the processing engine, as in this comparison no Geocoder approach was used in *CARIS*.

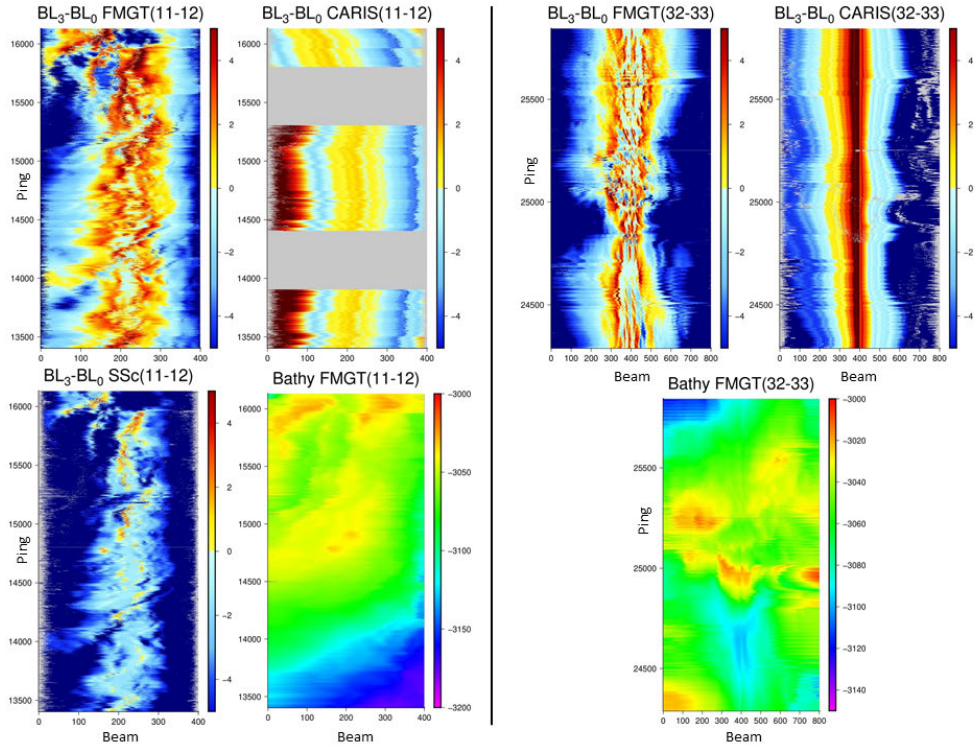


Figure 4.12: Differences between BL_3 and BL_0 for all data sets (2017 left, 2019 right) and software.

In summary, differences between BL_3 (after radiometric corrections) and even already between BL_0 levels (before radiometric corrections) are noticeable. Even if different software uses the same corrections, the individual algorithms

and their integration into the software can differ to such an extent that different results are obtained. This observation correlates with the study of intermediate processing steps by Malik et al. (2019), where this analysis was carried out in more detail. Furthermore, Malik et al. (2019) concludes that the differences in the BL_3 level for Kongsberg data are mainly caused by the difference in data reading in the initial step, rather than by the radiometric corrections itself.

4.2.2.2 Visual Comparison of Mosaics

After radiometric corrections, the angular dependence removal and image enhancement filtering are conducted to obtain the final mosaic. The possibility of user interference is greater at this stage – compared to the previous radiometric corrections – as they are more related to the specific study area, survey layout, and data characteristics. There are two sources that influence the **angular dependency** of the backscatter levels: echo sounder induced imperfections in the beam pattern and the acquisition geometry.

Beam pattern errors are typically visible as a constant intensity offset of a particular group of beams, resulting in along track artefacts. They are usually corrected using a beam pattern file. This can often be created by the operator and should be generated using a small subset of data from a flat and homogeneous area of consistent and preferably known seabed type. Deviations from the expected angular response are then considered to be echo sounder induced beam pattern artefacts. These are then removed from the data. If the operator does not create such a file, the correction can often be done automatically by the software by creating a beam pattern file on its own (*CARIS*) or by modelling the correction within the AVG correction (*FMGT*). Such a combined approach is also followed by *SonarScope*. As the high-resolution MBES data within the INDEX project is acquired in the area of a mid-ocean spreading ridge with distinct bathymetric features and frequent changes in seabed type, there are no locations available for optimal beam pattern file generation. Therefore, in the following comparison, no manual beam pattern file was created.

The second correction for angular dependency removal – the AVG correction – is also generally determined using a subset of the data. This data subset is binned by angle (e.g., in 1° intervals) and the values for the bins are averaged. In addition, a reference level is calculate at a reference angle to which the data should be normalised. This can be an narrow angular range like 43°-47° as suggested by (Lamarche et al., 2011) and implemented in *SonarScope*. Or also a wider range can be used as in FMGT (20°- 60°) or CARIS (30°- 60°) (Schimel et al., 2018). Subtracting the reference level from the averaged level of the angular bin provides the correction, which can then be applied to data samples taking into account the local angle. As a result, the retrieved backscatter level appears to be derived from a constant insonification angle (Lurton and

Lamarche, 2015). The selection of the subset for determining the correction values is crucial, as it has a strong influence on the visual result. *FMGT* and *CARIS* follow an approach of a sliding window to determine a subset. For each data sample to be adjusted, a local subset of a certain number of previous and subsequent pings is created (Lurton and Lamarche, 2015). The number of pings can be specified by the operator. The approaches and algorithms implemented for mosaicking are different for the software tested. Therefore, each processing workflow and possible parameter settings to produce an optimal mosaic were first investigated separately within each software before being compared.

As the name suggests, the MBES backscatter processing in **QPS Fledermaus Geocoder Toolbox** is based on the Geocoder tool developed by the University of New Hampshire (Fonseca and Calder, 2005). The user has limited options to influence the mosaic result. As mentioned above, the user can create a beam pattern file based on a user defined subset of data, or this correction will be taken care of by the AVG correction. As the test for this test data does not include a flat and homogeneous seabed area, the beam pattern file cannot be created. The user can also specify the window size of the AVG correction sliding window (default 300). This window size expresses the number of consecutive pings of a data sample to be considered for AVG correction. When processing HOMESIDE MBES backscatter data, this has been increased from the default value of 300 to prevent the removal of larger intensity patterns. A data file of HOMESIDE data has a recording length of 10 min and includes about 1,000 to 1,500 pings. A window size of 900 has been chosen as this is typically about 60% to 90% of the file length. Figure 4.13 shows a mosaic created with an AVG window of 300 (top) and 900 (bottom). The mosaic with the larger sliding window has two visual advantages. First, larger scale backscatter patterns are more apparent (see red mark in Figure 4.13). Secondly, certain abrupt changes in backscatter level in along track direction are reduced (see blue mark in Figure 4.13).

The user can choose between the AVG algorithms *flat, trend,* and *adaptive* (default). These algorithms define how the reference level is calculated from the selected data subset (“window”). The *flat* algorithm aims to reduce noise in the data by smoothing out small variations. The *trend* algorithm additionally checks for a trend in the data and maintains it. This is particularly useful for data collected in areas with strong slopes in the bathymetry, such as the INDEX data. The *adaptive* algorithm is a combination of the other two (QPS, 2022). When comparing mosaics created with all three methods, the result using the *adaptive* method was hardly distinguishable from the one of the *trend* method. Overall, the resulting mosaic created with the settings listed in Table 4.4 shows the least artefacts and has the most homogeneous appearance. A section

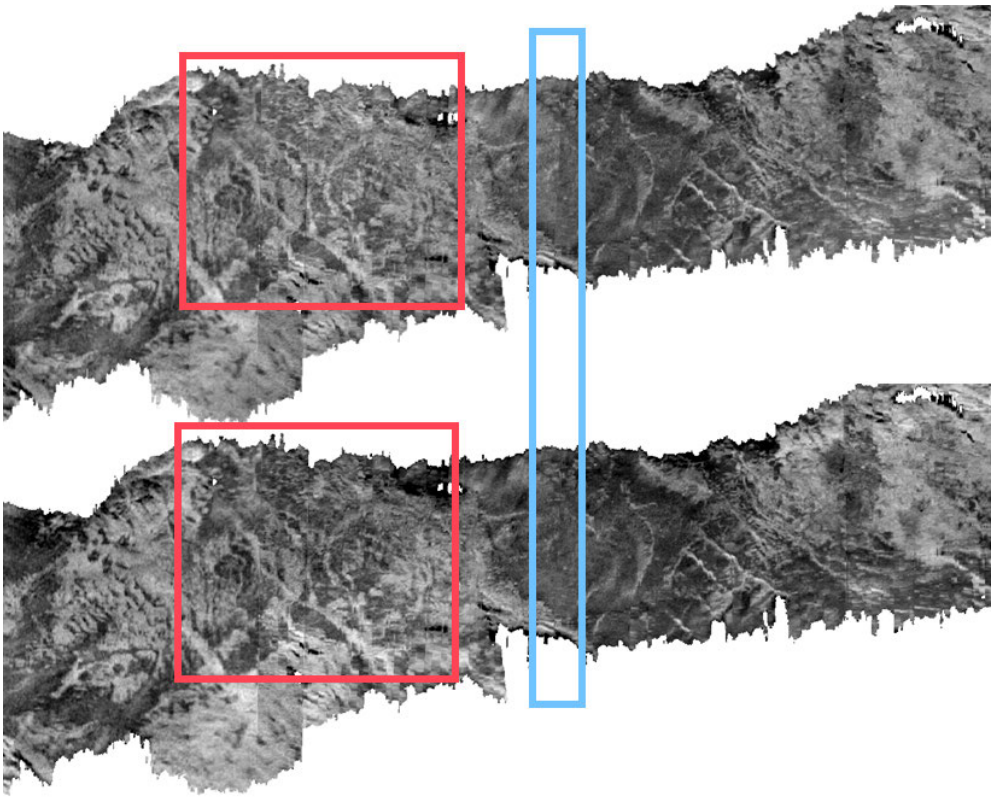


Figure 4.13: Mosaic sections (1 m resolution) of backscatter data collected during INDEX 2017. The upper mosaic was created with an AVG sliding window of 300 and the lower one with 900 in *QPS Fledermaus Geocoder Toolbox*.

of it is shown in Figure 4.14. Some blurring effects (marked blue) and some across-track artefacts (marked red) are still visible. A possible cause of the latter artefacts, if they occur at file boundaries, can be the limitation of the AVG correction to the file length. Such artefacts are also caused by mode changes. They cannot be resolved during post-processing. They have to be avoided during data acquisition by keeping the mode constant and not changing the frequency or pulse type. Such changes in acquisition settings occur regularly when collecting HOMESIDE data, as the main focus is on bathymetric data acquisition. Adjustments to the settings need to be made to maintain high quality bathymetry during altitude variations or environmental changes.

Table 4.4: Recommended backscatter processing settings in *FMGT* for HOMESIDE MBES data.

Correction	Setting
Beam Pattern Correction	Off
AVG Method	Adaptive / Trend
AVG Window Size	900

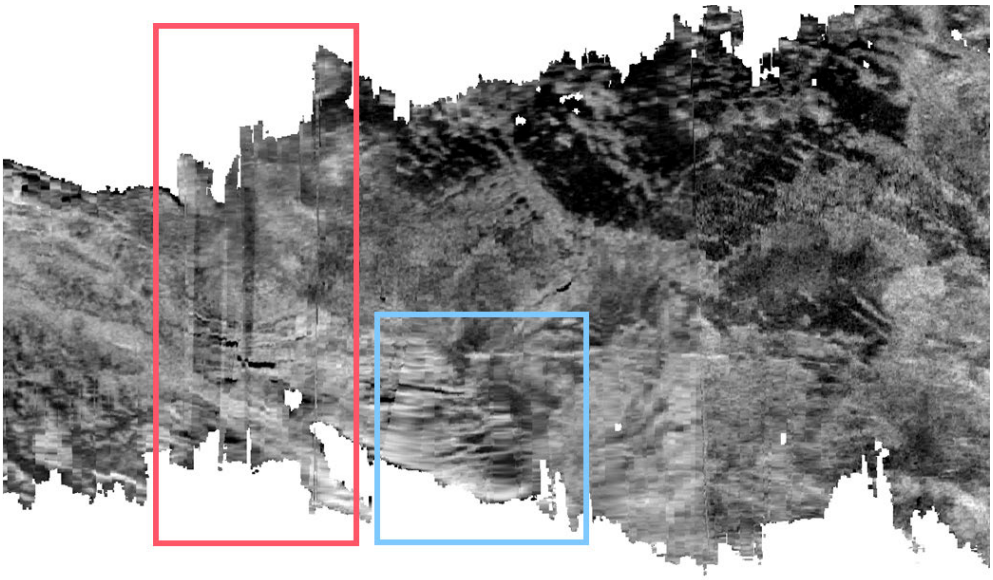


Figure 4.14: Example of the final mosaic (1 m resolution) created using *FMGT*. Some across-track artefacts (red) and blurring effects (blue) are visible.

Three different mosaicking engines are available in **CARIS**: *SIPS Backscatter*, *SIPS Backscatter with area based AVG*, and *Geocoder*. Although the *Geocoder* engine is based on the same algorithm as also incorporated in *FMGT*, due to differences in implementation, different results can be observed (Dufek, 2012). Unfortunately, this engine crashes when using 2019 data. It also shows artefacts caused by strong changes in the overall backscatter level for the 2017 data (Figure 4.15 top, location highlighted in red). These are most likely related to mode changes already observed in *FMGT*. However, using this *CARIS* engine, the artefacts become so large that smaller backscatter patterns become barely visible due to the large change in intensity level. The *SIPS Backscatter* and *SIPS Backscatter (WMA with area based AVG)* engines give generally good results.

The *SIPS Backscatter* also uses a sliding window AVG correction. Therefore, the window size is also kept large with a value of 900. The value of 900 is the maximum possible for the data sets tested, as a value of 1,000 results in large gaps in the mosaic. The beam pattern correction is switched on automatic, as this reduces an intensity difference between the two heads. The option *Correct for Acquisition Modes* seems to produce across-track artefacts in certain areas. This correction differentiates beam pattern files for the different acquisition mode (pulse length, wave form). Due to the frequent mode change in the INDEX data, this approach does not improve the data quality and is therefore not used.

The *SIPS Backscatter (WMA with area based AVG)* engine also takes into account data overlaps for AVG correction. It also uses a sliding window, but as understood from the sparse information provided by the user manual, the

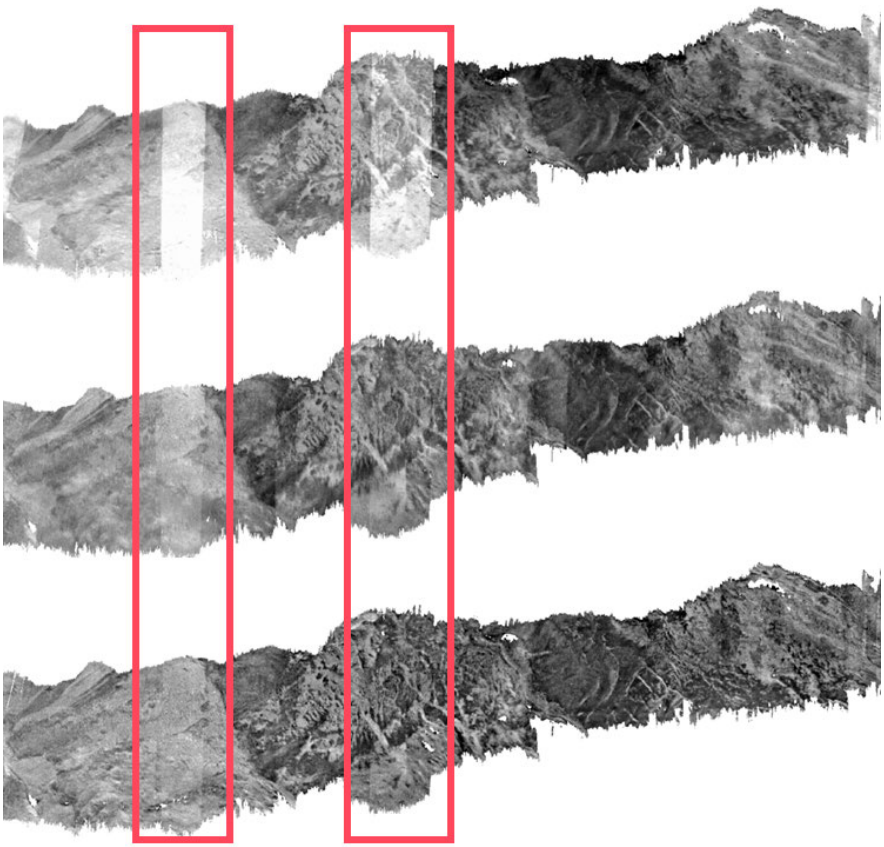


Figure 4.15: Mosaics (1 m resolution) generated with *CARIS* using different backscatter engines: *Geocoder* (top), *SIPS Backscatter* (middle), *SIPS Backscatter with WMA and area based AVG* (bottom). The abrupt shift in backscatter level caused by mode changes are visible as artefacts (highlighted in red) in all mosaics to varying degrees.

data subsets used for AVG correction are not line or swath based as. Normally HOMESIDE MBES swath data has hardly any overlap between adjacent lines. However, INDEX data typically shows small-scale intensity variations due to the rough seabed and its abrupt changes in seabed type. For this case a non-swath based approach may be advantageous and has been tested.

The mosaic generated with *CARIS SIPS Backscatter (WMA with area based AVG)* engine (Figure 4.15) shows less artefacts. Also some artefacts in the nadir area likely caused by the dual-head setup can be further removed. The window size for samples being grouped can be set using the *updating size* parameter. This setting is crucial to avoid artefacts. If no value is given, *CARIS* chooses one automatically. As this results in square shaped artefacts probably due to some gridding issues, the value was set to 5 m to avoid such problems. Although, the mosaics created with this engine look more homogeneous than those generated using the *SIPS engine*, it has the disadvantage of a much longer computation time. The mosaic generation took 34 min instead of less than 2 min for the 2017 test data. The mosaic created of the 12 files collected 2019 even took 110 min. The engine is therefore not suited for large data sets. A list of recommended processing settings is given in Table 4.5.

Table 4.5: Recommended backscatter processing settings in *CARIS* for HOMESIDE MBES data.

SIPS Backscatter		**SIPS Backscatter WMA with area based AVG**	
Beam Pattern	Automatic	Search Radius From Footprint	Off
Correct for Acquisition Mode	Off	Beam Pattern	Automatic
AVG	Adaptive	Correct for Acquisition Mode	Off
AVG Window Size	900	AVG	On
		Updating Size	5 m
		Chunk Size Multiplier	Not Set

The three different approaches for mosaic creation implemented in **SonarScope** (*no corrections*, *statistic compensation*, and *sonar calibration*) are tested using HOMESIDE data of 2017. As mentioned previously, the current software version is not able to process the data from the HOMESIDE dual-receiver setup, so the data collected from 2019 onwards cannot be tested. The *no corrections* method plots the georeferenced data stored in the *.all files. The result shows a more homogeneous mosaic than may be expected, as Kongsberg applies various real-time corrections during data acquisition (Figure 4.16a). As these corrections are based on assumptions and simplified models (Lurton and Lamarche, 2015), they are usually initially removed by the backscatter processing software before more complex corrections are applied.

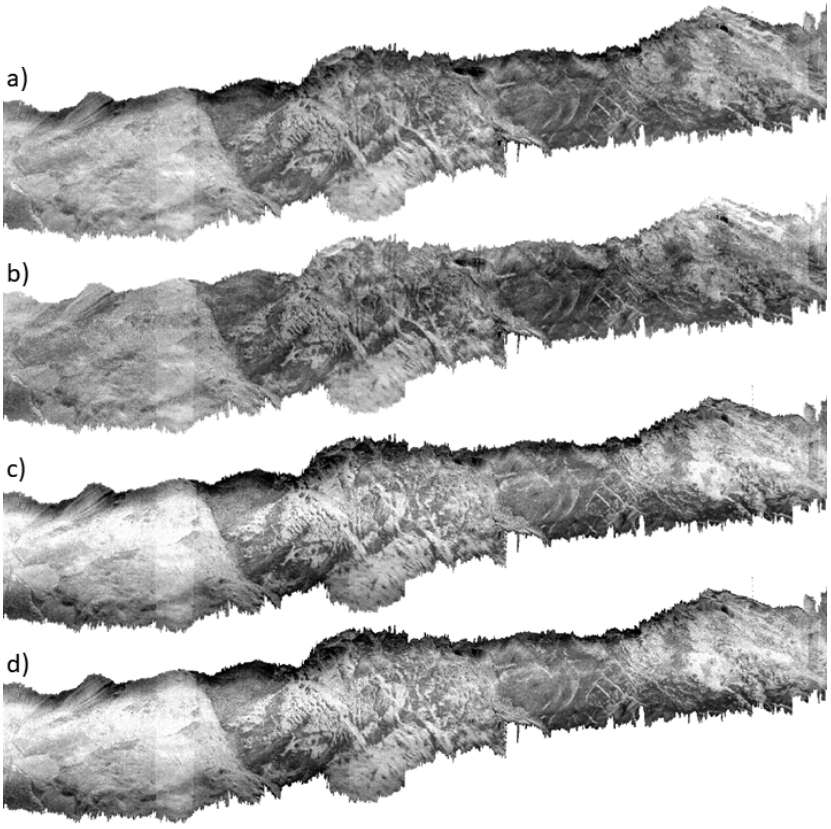


Figure 4.16: Mosaics (1 m resolution) created using *SonarScope* with the settings a) no corrections, b) static compensation, c) BS calibration step 1, and d) BS calibration step 3.

The *statistic compensation* in *SonarScope* provides data based enhancement of the backscatter data. An area with the least amount of backscatter variation is selected and a compensation curve is calculated and applied to the data set (Figure 4.16b). Compared to the mosaic without correction, this backscatter image shows a significant improvement in the systematic lower backscatter level of the outer beams on the port side (to the north). However, this improvement does not apply to the whole mosaic, because further east, where the original data (Figure 4.16a) shows no systematic pattern in the outer beams, the compensation artificially increases the backscatter level of the outer beams on the northern side. When using such a statistical approach to correct of backscatter data, the data needs to be carefully examined. The subset of data used to construct the compensation curve and the selection of data to which the correction is applied must be carefully chosen to obtain results with reduced systematic along track artefacts.

Another available option is to perform a backscatter calibration to generate a mosaic of absolute backscatter values. However, as typically not all of the necessary parameters involved are known, *SonarScope* provides some models for their estimation. The processing therefore requires a relatively high degree of user intervention. The individual steps are applied file-wise in a non-georeferenced domain of beam/ping geometry. Towards the end of the processing chain using HOMESIDE data, the software accounted some errors which prevented the whole process from being completed. *SonarScope* divides the processing parts into stages. The last stage up to which HOMESIDE data can be processed is called step 3. Its status as well as the correction status of step 1 are listed in Table 4.6. Figure 4.16 shows the corresponding mosaics of step 1 (c) and step 3 (d).

Table 4.6: Correction stages of the created mosaics in *SonarScope*.

Correction	**Mosaic Step 1**	**Mosaic Step 3**
Transmission Loss	Real-time compensated	Compensated by SSc
TxDiag	Real-time compensated	Not compensated
Incidence Angle	Real-time compensated	Compensated by SSc
Backscatter	Not compensated	Compensated by model

Step 1 is an early stage where most of the manufacturers real-time corrections are still implemented and only the backscatter adjustment is removed. However, the mosaic generated after processing step 3 is visually nearly indistinguishable from the one produced after step 1. Both show strong angular dependent artefacts as not all necessary compensations are applied.

4.2.2.3 Conclusion of the Section

The software comparison is conducted to determine the optimal processing workflow (e.g., software and parameter settings) for the HOMESIDE MBES backscatter data. The generated mosaic should be homogeneous and not highlight data artefacts for further analysis using AI algorithms. The mosaic should also emphasise the visualization of individual seabed classes. Chapter 5.2.2 focuses on the later part and analyses the influence of acquisition parameters and processing settings on backscatter data.

Commercial backscatter processing software is typically "black-box" in nature, with very limited user interaction. The option to compare intermediate backscatter levels provides some insight into the individual corrections applied by the software. This comparison shows the differences in backscatter level adjustment between various software already at an early stage in the processing workflow.

In the second comparison, the final results – georeferenced mosaics – are compared visually. This comparison is done to decide on the software for HOMESIDE backscatter data. But it also showed how important it is to adapt the default values of certain corrections to the specific properties of the data (e.g., AVG window of 900 pings). In particular, the INDEX data has very "atypical" characteristics, as backscatter analysis typically uses shallow water data collected in proximity to land over a flat and sedimented terrain.

When evaluating the individual software as well as the final mosaics, *SonarScope* allows the most user intervention and options – even allowing a full calibration. However, the HOMESIDE data could not be final processed using that software, partly due to the lack of information needed for adequate calibration. *FMGT* and *CARIS* successfully generate homogeneous mosaics. The best overall result is obtained using *CARIS* and the *SIPS Backscatter with WMA and area based AVG* engine (most homogeneous mosaic). However, this approach is not suitable for processing large data sets, as the computation time is much longer then using other approaches (by a factor of about 15). There are also other problems with the 2019 data.

The mosaics generated using *FMGT* show some blurring effects, but the software has the advantage of being relatively low in computation time and being able to handle large amounts of data comparatively well. It is therefore chosen as the overall backscatter processing software, also because its results significantly improved since version 7.10.3 (available since May 2023). Since then, a new AVG option has been available which allows the AVG calculation to include the previous and subsequent data file. This approach has been shown to significantly reduce artefacts in the mosaics at file breaks. As already noted in the comparison when *CARIS SIPS Backscatter with WMA and area based*

AVG gave the best result, it is beneficial for the final mosaic using HOMESIDE data if the AVG correction is not limited to the file length.

4.2.3 Final HOMESIDE MBES Backscatter Data Processing Workflow

Based on the conclusions of the above comparisons, a general processing workflow for the HOMESIDE data was developed. It consists of six steps (Figure 4.17). First, the raw MBES files in Kongsberg format (*.all), the corrected *DELPH INS* or *mbnavadjust* navigation (*.txt), and the DTM are imported. The backscatter correction settings (processing parameters) as well as the mosaicking parameters need to be set by the operator. For typical high-resolution HOMESIDE data, a resolution of 1 m is usually chosen. Once the mosaic has been computed, manual adjustments can be made using the tools provided within the software. These include, for example, the manual backscatter level improvement of a file or the complete removal of a data section. As the *time series* format is used, the data amount of an larger area of overlapping HOMESIDE data can reach a size where the the software/hardware computation limit is reached. Therefore, each year's data is processed separately and then the mosaics are merged in QGIS. The merging process doe not average overlapping data, but the user selects which mosaic should overlay the other. This reduces addition artefacts created by averaging mismatching mosaic values.

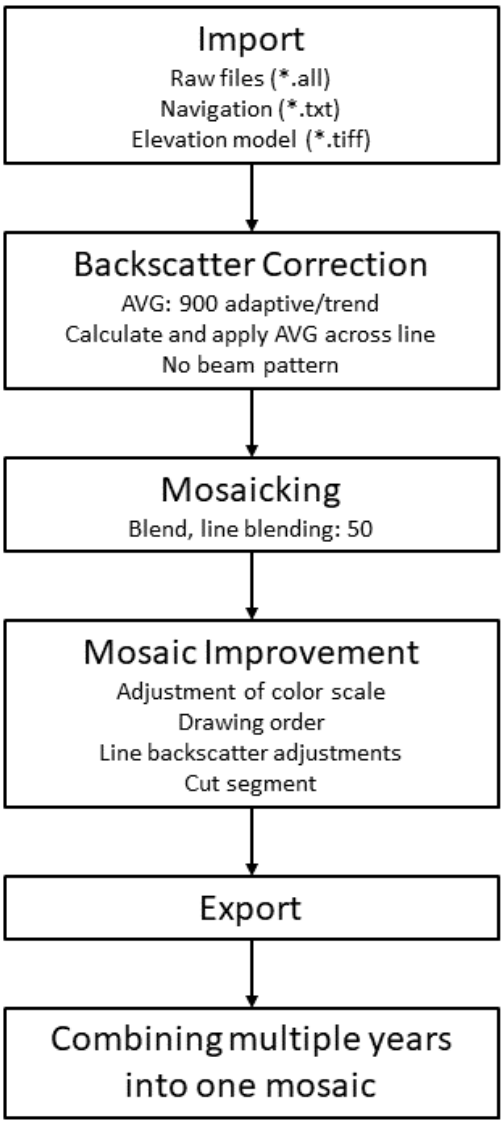


Figure 4.17: Scheme of HOMESIDE backscatter processing workflow using *QPS FMGT* (own illustration).

Chapter 5

Acoustic Characterisation of Seafloor Classes Including Hydrothermal Areas

The initial motivation for this study is to investigate the feasibility of using HOMESIDE MBES backscatter data for the detection of (inactive) hydrothermal sites. Active hydrothermal discharge locations are typically detected by water column anomalies caused by the hydrothermal fluid. However, from an exploration point of view, inactive fields are of greater interest, as they may contain larger sulphide deposits and no vent fauna is present. Prior to the addition of the self-potential string to HOMESIDE as part of the typical setup, inactive hydrothermal sites were only discovered by video inspection of the surroundings of active areas. This chapter of the thesis focuses on the acoustic characterisation of the different seafloor types within the study area and in particular of active and inactive hydrothermal sites. The aim is to investigate whether specific backscatter signatures can be established for each of the seafloor classes.

Ideally the backscatter data should be calibrated to allow a quantitative analysis. Due to the nature and remoteness of the study area, a calibration of the backscatter data is too complex and time consuming to be feasible. In addition, different system setups and sensors have been used over the years. An absolute quantitative analysis of the backscatter data is therefore not possible. Due to the harsh morphology and the MBES instrument and settings used, approaches such as using the angular response are not feasible for the INDEX data as well. However, the INDEX data has the advantage that ground truth data is available in the form of ROV or video sled imagery over large areas. In this section, the backscatter mosaics are compared with the ground truth seabed classes to examine their visualisation – first in general and then explicitly for areas of hydrothermal activity. Afterwards, a textural analysis is also carried out.

5.1 Acoustic Characterisation of General Seafloor Classes

BGR provided geological maps of seven hydrothermal fields for comparison of seafloor classes and MBES backscatter data. These were generated by geologists from the Laurentian University (Canada). In addition, an own seabed classification of *Field 1* was performed based on eleven ROV and three video sled deployments (Chapter 3.3). For each geological map, a histogram of the backscatter values of each of the major classes (e.g., sediment, pillow, talus, sulphide blocks) and the area covered by the map was generated (Figure 5.1). The total mapped area is approximately 11 km^2.

As mentioned above, *Field 1* has been mapped twice: once by the geologists and again within this study. The additional classification of *Field 1* was done to produce a map with a more detailed classification – with more classes and therefore a stronger correlation with the MBES backscatter data. The geological maps by BGR used in this study do not use mixed classes, but only list the most dominant class. They have a higher degree of generalisation. In Figure 5.1, only the main classes of the BGR geological maps have been used – also for the own classification. Only the class "sediment with talus" has been added, as it is one of the main data classes and increases the amount of data shown.

Seafloor coverage varies significantly between the different classes and fields shown in Figure 5.1. In particular, "sulphide blocks" would not be visible in the histogram plot if its bin size had not been increased. The purpose of the histograms is to visualise the backscatter values of the corresponding MBES mosaic for each seafloor class and their relationship to each other. The backscatter values shown are those taken from the individual mosaics. As the backscatter data is not calibrated, no direct comparison between seafloor class and absolute backscatter value can be made.

In all histograms, sediment and talus are the most dominant classes overall. Areas covered by pillow or gabbro (both hard rock classes) are limited. The data is mostly normally distributed, except for *Field 2*, where the backscatter values representing the areas covered by sediment are not symmetrical. Or for the talus class in *Field 3*, which shows a lower number of occurrences around the mean values. Some specific regularities can be noted from the histograms: The curve representing sediment is typically shifted to the lower backscatter values than talus or pillow / gabbro (mostly). Areas classified as "sulphide block" have a similar backscatter response range to talus: the histogram ranges are comparable in three out of the six fields. In the other three, sulphides have slightly higher backscatter values than talus. This correlates with other studies (Searle et al., 2010; Asada et al., 2015).

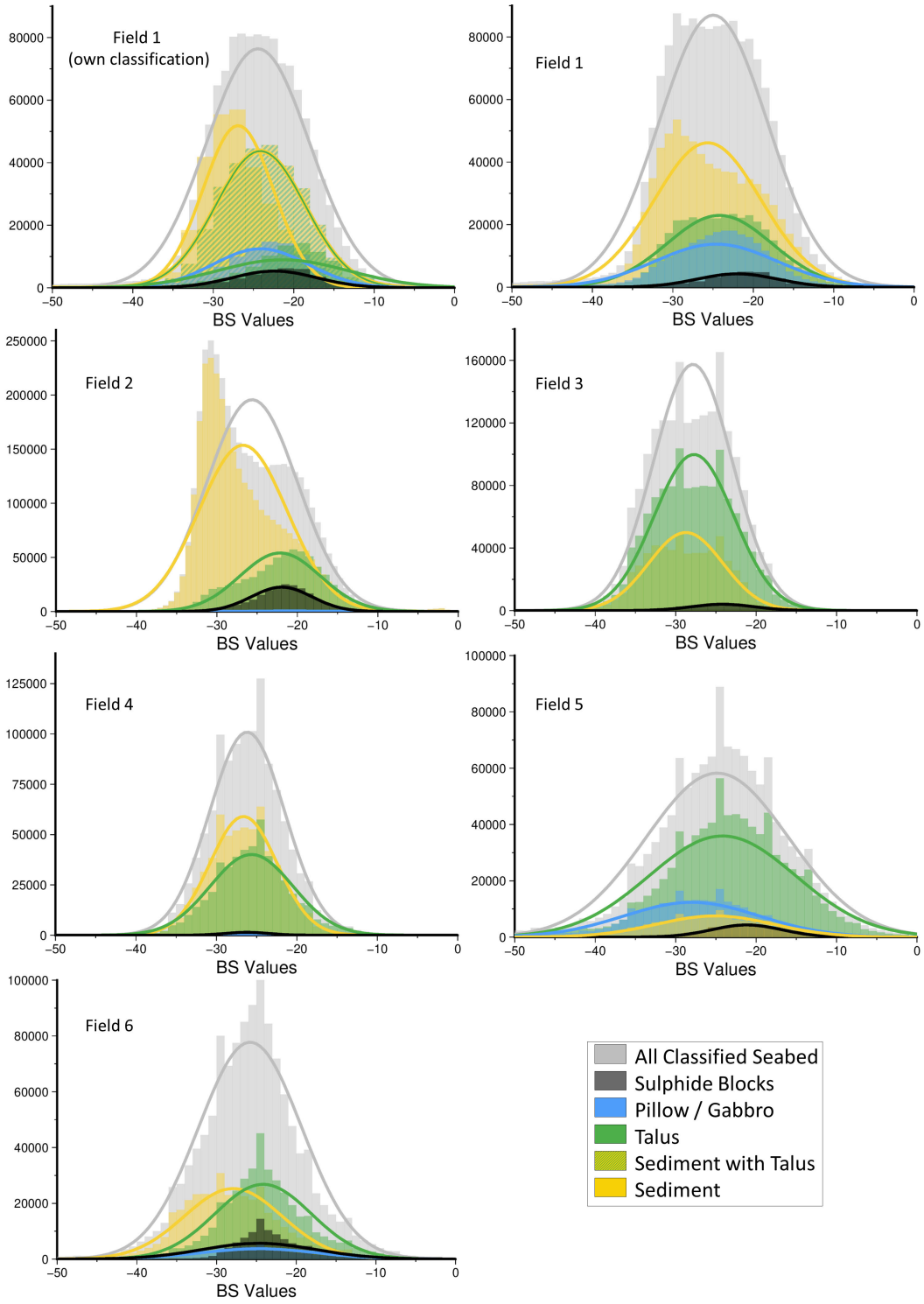


Figure 5.1: Histograms of backscatter values of seabed classes identified in video imagery.

Figure 5.2 shows another plot of the same data as above. The backscatter values have been normalised to the 3-sigma range for the particular data set. The box plots with whiskers represent the distribution of the data (2-sigma minimum, 25% quantile, median, 75% quantile, 2-sigma maximum). The width of the boxes represents the relative area coverage of each class within the

individual plot.

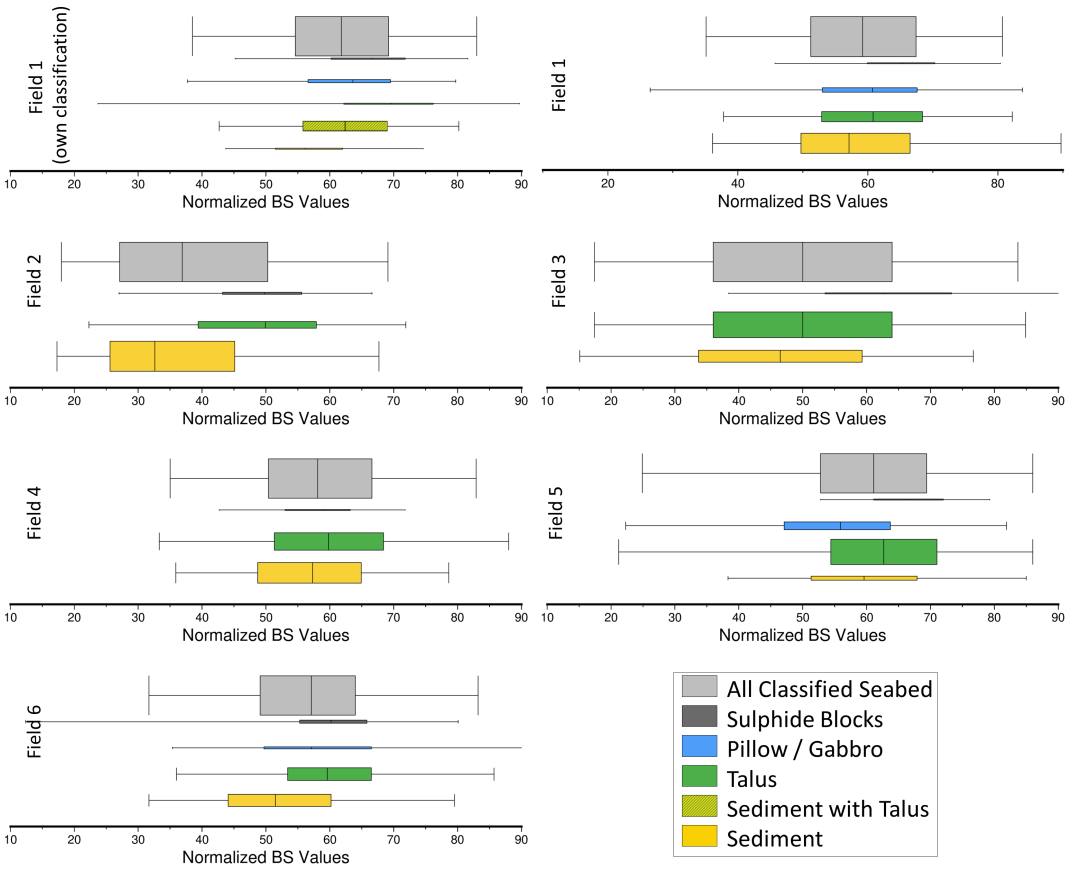


Figure 5.2: Box plots of normalised backscatter values of seabed classes identified from on video imagery. Whiskers indicate the 2-sigma minimum/maximum data range in additional to the 25%-quantile, median, and 75%-quantile represented by the boxes. The box width represents the relative area coverage of each seafloor class within the individual plot.

Although the data distribution varies slightly between classes and fields, the same particular regularities can be seen as in the previous figure. In addition, the 2-sigma minimum and maximum values for all classes are roughly the same across the data sets: the whiskers fall within a similar range across the data sets. The extreme values of the classes are therefore similar and do not provide any additional information. This could be due to remaining artefacts causing extreme values in the data (such as shadowing effects or direct acoustic reflection). An examination of the majority of the data around the median (e.g., between the 25% and 75% quantiles) provides more meaningful information about the characteristics of a class in the backscatter data and can be summarised as follows:

Sediment: The backscatter values and therefore the median of the mapped class "sediment" are generally lower than any of the other classes. Such a lower backscatter response from sediment is expected as it allows more of the acoustic energy to be transmitted to the upper layers, resulting in

less scattering on this typically smooth surface.

Talus: The irregularly shaped and hard talus has a higher overall backscattering strength than sediment. It is a generally rough and irregularly shaped material that does not allow any energy to be transmitted into the ground and facilitates scattering.

Pillow: The pillow or tube shaped cooled down lava flows are typically one metre in size. Pillow lava, like also talus, is a hard material that does not allow the acoustic signal to penetrate into the ground. Its average backscatter value is therefore higher than that of sediment, but its smoother surface than talus facilitates direct reflection and therefore the scattered energy is lower.

Sulphide Block: Areas of predominantly sulphide blocks have a similar backscatter signature to talus. This is due to the very similar characteristics of both seabed classes: They have a gravelly / irregularly shaped structure and consist of hard material that facilitates the scattering of the acoustic energy.

In summary, the different seabed classes show similar characteristics in the geological maps of all six hydrothermal fields. However, even when normalised backscatter values are used, as in Figure 5.2, the values for each class vary over the data range and cannot be assigned to a particular data range interval. There are several reasons for this: First, the MBES backscatter data has not been calibrated, and different setups and parameter settings have been used over the years, resulting in mismatching backscatter ranges between adjacent data sets. Although attempts have been made to manually adjust backscatter levels from different data sources, it is not possible to compare absolute values between different deployments. Inadequate data correction during processing can also cause such range differences. Although all influences of local acquisition geometry and gain settings should have been removed, the study area is located in a challenging area for backscatter data acquisition. During data collection, the focus was not on optimal settings for backscatter data in terms of profile overlap or consistent signal settings.

Second, the definition of seafloor classes and the level of detail play a crucial role in the geological mapping. The imagery data was collected over eight years and the geological interpretation was carried out by different people over time. The decision on the locally dominant class as well as the level of detail (and spatial extrapolation) can vary between maps as they depend on the individual's decision. In addition, ROV navigation can be erroneous during some deployments, leading to spatial discrepancies between bathymetry and geological interpretation. The attempt within this study to create a more detailed seabed classification, with particular emphasis on differentiation in the

backscatter data, resulted in a map with a low level of generalisation and therefore many classes. As the area coverage per class is therefore low, no general conclusions can be drawn as the influence of remaining artefacts (extremely high and low backscatter values) is too high. This again shows the importance of a properly post-processed and corrected backscatter mosaic for numerical analyses.

In addition to the above limitations, it should be noted that the hydrothermal fields under investigation extend over a distance of approximately 620 km, illustrating the vastness of the whole INDEX area. The area covered by geological maps of 11 km^2 seems tiny in comparison. However, it is to important to considered how time-consuming the geological map creation process is. It takes at least as long as all the video deployments, and even longer considering further editing afterwards. The video playing time is about 82 hours for *Field 1* alone. Even with a relatively small spatial coverage and a sub-optimal backscatter data quality due to the acoustically challenging environment and acquisition parameters, it has been shown above that some basic conclusions about the characteristics of the individual seabed classes can be drawn from the MBES backscatter data. Mosaics can be used to extrapolate seabed classes in geological mapping.

However, to return to the original question of identifying a unique backscatter response from areas showing evidence of sulphides, these areas are examined in more detail in the following section. Such areas with evidence of hydrothermal activity – from hydrothermal alteration to massive sulphide blocks and chimneys – represent only a small part of the total mapped area. However, their MBES backscatter response, or more precisely the identification of a potentially unique response, is of great interest for exploration.

5.2 Acoustic Characterisation of Seabed Areas Showing Evidence of Hydrothermal Activity

The acoustic signature of hydrothermal areas is of particular interest because the identification of unique parameters could help to detect both active and inactive hydrothermal sites, and hence potential associated sulphide deposits. However, as discussed in the previous section, the influence of acquisition settings on the resulting backscatter mosaics and therefore the visual appearance of a site cannot be neglected. In the following two subsections, firstly, backscatter data from different sites is compared visually to determine whether hydrothermal sites have the same backscatter characteristics (Chapter 5.2.1). Secondly, the sensitivity of the acquisition settings on the site visualisation is analysed (Chapter 5.2.2) by visually comparing sites acquired using different acquisition geometries and MBES settings. This analysis is also done visually,

as it has been shown above that a numerical based analysis is not advisable.

5.2.1 Visual Inspection of Backscatter Intensities at Hydrothermal Sites

In the geological maps provided, the following four seabed classes have been distinguished in areas of hydrothermal evidence: hydrothermal sediment, hydrothermal crust, sulphide talus, and sulphide block. Figure 5.3 shows images of these hydrothermal classes. They all exhibit the typical iron-manganese colouration, ranging from yellow to red. However, their visual characteristics differ.

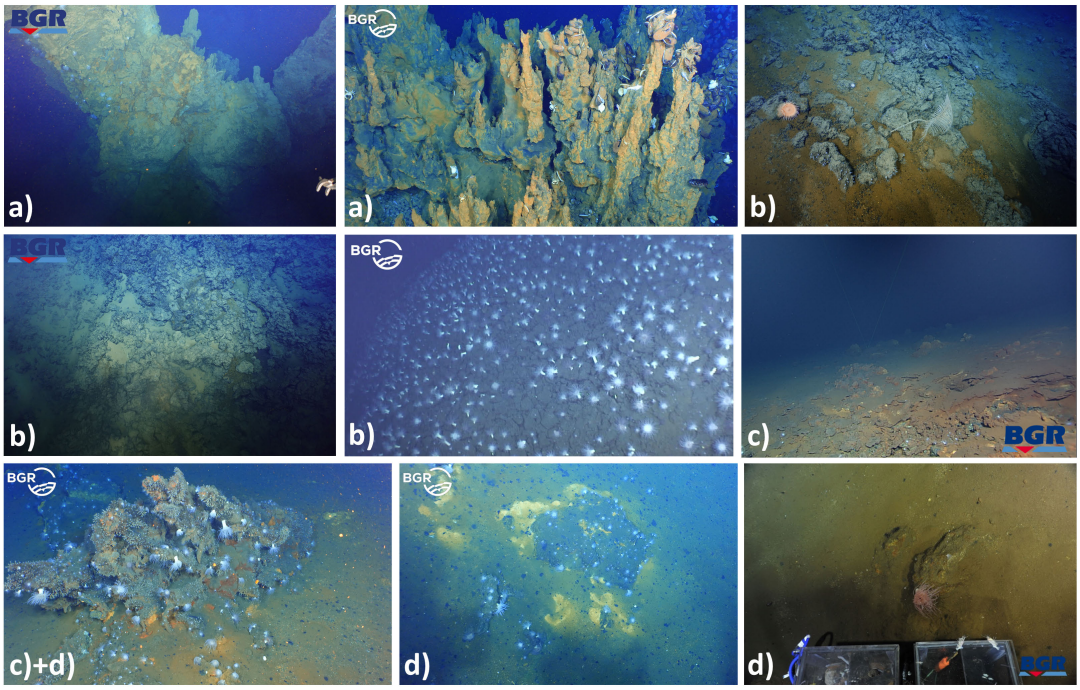


Figure 5.3: Images of different seafloor classes associated with the presence of sulphides: a) sulphide block b) sulphide talus c) hydrothermal crust d) hydrothermal sediment.

Sulphide blocks typically include mound and chimney structures found near discharge sites of hydrothermal fluid. When a site becomes inactive, the sulphide blocks weather over time, becoming more porous and reducing the size of the mounds and chimneys. Hydrothermal crust generally has a smoother surface than the sulphide blocks or talus. It is a hydrothermal alteration of the seafloor and typically indicates hydrothermal activity in the vicinity. *Sulphide talus* refers to sulphide that is broken into smaller pieces than sulphide blocks. This can include chimney remnants, for example, but also additional evidence of hydrothermal activity such as coloured sediment or vent fauna like anemones. Hydrothermal sediment is visually distinguishable from non-hydrothermal sediment by its colour. The typical yellow to red colouration is caused by iron precipitating from the plume.

Each of these classes can also occur at different stages of erosion and weath-

ering, which also affects the backscatter response. The classes sulphide talus, hydrothermal crust, and hydrothermal sediment are expected to be acoustically indistinguishable from the "non-hydrothermal" equivalent or general classes – and therefore also numerically indistinguishable in the mosaic. Therefore, only the *sulphide block* class was used for the comparison in the previous chapter.

As mentioned in the previous chapter, the *sulphide block* class has a backscatter response similar to or higher than that of talus. Figure 5.4 depicts some active and inactive sites from different hydrothermal fields, with geological maps overlaying the bathymetry (left) and backscatter mosaic (right). Areas showing (former) hydrothermal activity are outlined in pink in both types of images.

Comparing the two different representations for each site in Figure 5.4, there is no clear distinction between seafloor areas with and without evidence of (former) hydrothermal activity in the backscatter mosaic. Similarly, the four different classes representing sulphide presence and hydrothermal activity show no visually discernible demarcation from each other in the backscatter data. In some examples, hydrothermal sites are represented by generally higher backscatter values (light grey), while in others they appear with lower backscatter values (dark grey). Visually, therefore, there is no specific signature of hydrothermal (in)activity. The appearance of a site within the backscatter data depends on the surrounding seabed and its backscatter characteristics.

One of the conclusions of the analysis in the previous chapter was that talus (coloured green in the geological maps) has a higher backscatter response than sediment (depicted in yellow in the geological maps). This is due to its hard and rough surface, which prevents acoustic transmission and allows strong scattering in all directions. One would expect a clear visual difference between these two classes in the backscatter data, but this cannot be confirmed from Figure 5.4. The main reason for this is probably the degree of generalisation in geological mapping. Mixed seabed classes such as "sediment covered talus" are more common than plain ones. The proportion of each class in a mixed area may also vary. In mapping, only the dominant one is represented, resulting in clear boundaries between different classes in geological maps, whereas in reality – as well as in the high-resolution backscatter data – the transition between classes is typically smoother.

As mentioned above, the maps in Figure 5.4 were provided by BGR. They are used for geological interpretation of the area on a larger scale and are therefore slightly generalised. Figure 5.5 shows a section of the geological map of *Field 1* provided by BGR (left) alongside the seabed classification made as part of this study (right). There are some similarities in the general classification between the maps: a diagonal stripe from northwest to southeast in the middle of the selected section shows predominantly (fractured) pillow lava. To-

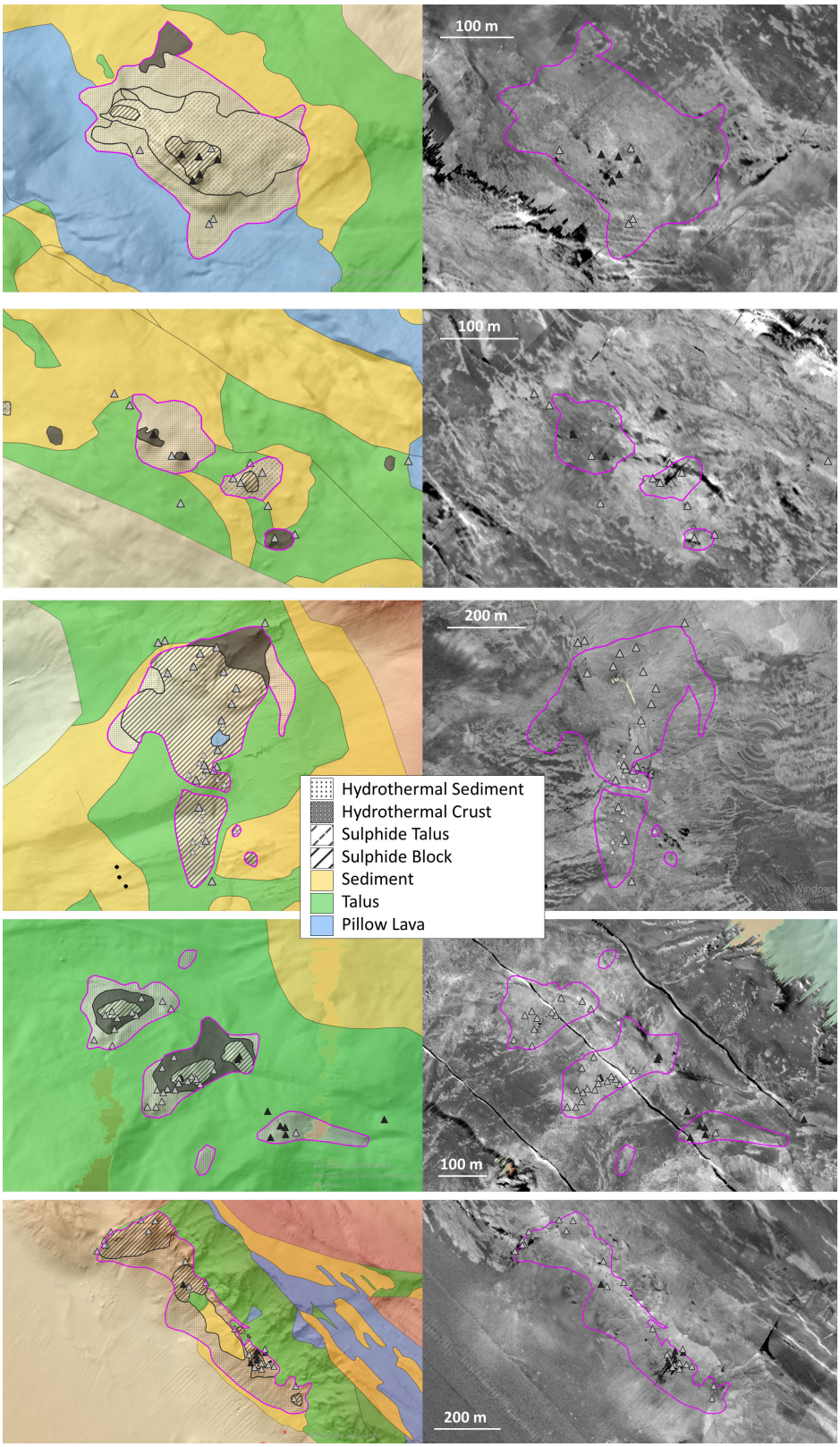


Figure 5.4: Various active and inactive hydrothermal sites shown as geological map (left) and backscatter mosaic (right). Areas with indication of hydrothermal activity are outlined (pink). Active and inactive vent locations are marked by black and grey triangles.

wards the northeast and southwest, sediment and talus are more dominant. The locations of seabed classes with sulphide occurrences are also similar.

However, the shape of these areas and the assignment of individual classes varies between the maps. This is also evident when comparing the outlines of the hydrothermal areas, illustrating also the problem of manually drawing class boundaries when the transition between classes is smoother in reality. In addition, the own classification includes more – and especially mixed – seabed classes, resulting in a stronger visual correlation between the class boundaries and the backscatter values.

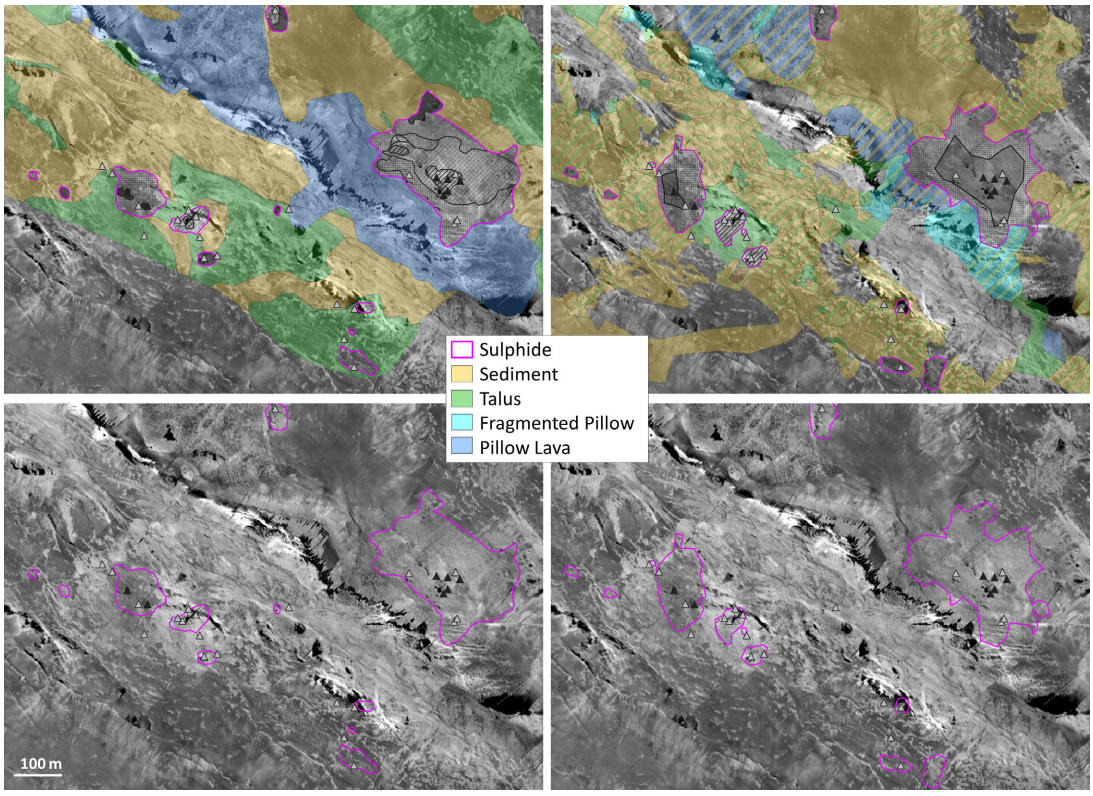


Figure 5.5: Section of a geological map provided by BGR (left) and made as part of this study (right) of *Field 1*.

In summary, there is no clear visual distinction between classes with and without signs of hydrothermal activity in the backscatter mosaic. In the following section, the visual analysis of hydrothermal sites is extended to individual data files collected at different times, using various settings, and processed with different software, in order to examine the influence of these parameters on the visualisation of the sites.

5.2.2 Influence of Acquisition Parameters and Processing Settings on Backscatter Data of Hydrothermal Sites

HOMESIDE data is typically acquired with the same MBES settings (300 kHz, FM pulse). However, these acquisition parameters (e.g., acoustic frequency, the pulse form), as well as the acquisition geometry, and processing settings, affect the backscatter data. The effect of different processing software with various settings for the backscatter mosaic is discussed in more detail in Chapter

4.2.2. The data used in the previous investigation was collected with different acquisition settings and geometry in hydrothermal areas in *Field 1*. In this section, the same data will be used to draw conclusions about the sensitivity of the site representation in the final backscatter mosaic. Again, different software is used for this visual analysis.

Two areas within *Field 1*, covered by HOMESIDE data in 2017 and 2019, were selected for this analysis. The first area (Figure 5.6) includes an active site with two active discharge sites – each located on a mound. There are a number of inactive sites in the vicinity. These include several chimneys remnants, ranging from very large standing or fallen structures to very small altered and broken chimney remnants.

Figure 5.6 shows the self-made seabed classification (with a lower level of generalisation) and the mosaics generated from different data sets and different processing settings. The seafloor areas with evidence of hydrothermal activity are outlined in pink. In all mosaics, the three larger sites in the middle of the data section – especially the active one – have lower backscatter values than their surroundings. In their immediate vicinity, mostly talus and sediment with talus patches have been mapped. The active site is dominated by hydrothermal sediment. This is consistent with the general expectation and previous observations that hydrothermal sediment has a lower backscatter response than talus.

The two inactive sites in the centre of Figure 5.6 have been classified as sulphide block. The backscatter intensities of these two sites are only slightly lower than those of the surrounding area and therefore the boundary is less obvious. In the numerical analysis of the different seafloor classes in Chapter 5.1, it was concluded that sulphide blocks typically have a similar or higher backscatter response than talus. This general statement contradicts the visual observation of the site made using Figure 5.6. This contradiction is caused by the generalisation of mapping. Both classes have a very similar backscatter response. In the maps used for the numerical comparison, talus and talus mixed with sediment may have been combined into one class in certain areas. In addition, sulphide blocks may show different degrees of weathering and erosion. These processes alter their appearance and formally rough surfaces may be smoothed, resulting in a lower backscatter response.

In summary, the numerical analysis of seabed classes and their corresponding backscatter values generally indicates that talus has a lower or similar backscatter response to sulphide blocks. However, in some cases this may be different. The purity of a class and processes such as erosion can influence the backscattering strength.

Comparing the different mosaics generated from data with various acquisition

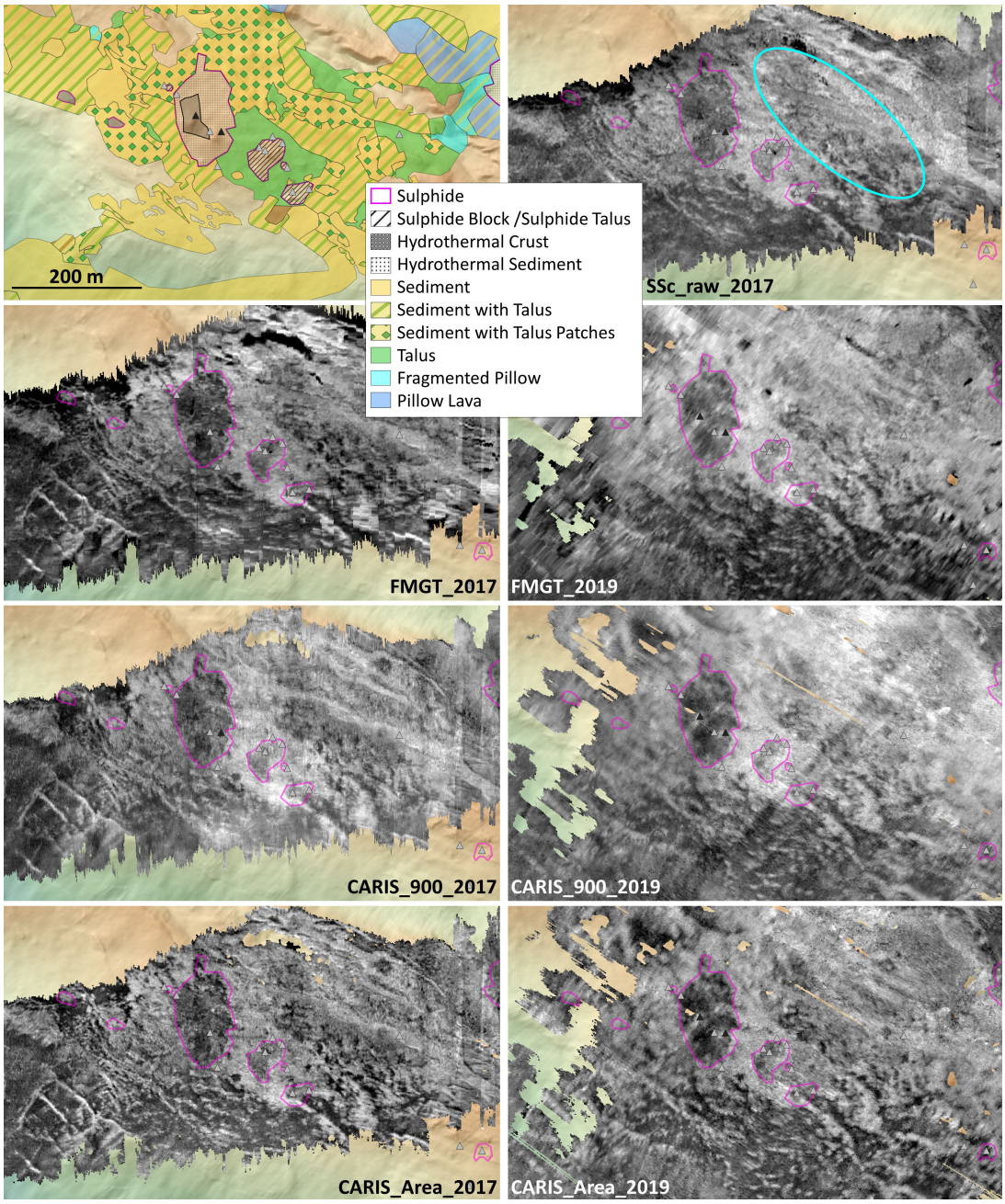


Figure 5.6: Example 1: Seabed classification and different backscatter mosaics (1 m resolution) created from data from different deployments in 2017 (300 kHz) and 2019 (200 kHz) and with various software and software settings (SSc_raw = *IFREMER SonarScope* without correction, FMGT = *QPS FM Geocoder Toolbox*, CARIS_900 = *CARIS HIPS and SIPS - SIPS Backscatter engine* with AVG 900, CARIS_Area = *CARIS HIPS and SIPS - SIPS Backscatter With Area Based AVG engine*). An area with greater differences in backscatter representation is highlighted in cyan.

and processing settings in Figure 5.6, the representation of the hydrothermal sites is similar in all of the mosaics. The greatest difference in representation can be seen for an area to the east of the inactive sites (highlighted in cyan in Figure 5.6 top right). This area appears darker and more heterogeneous in the 300 kHz data (2017) than in the 200 kHz data (2019). It consists of a mixture of sediment and talus, which explains the overall decrease in backscatter level. Only in the bottom mosaics does this area look similar. Therefore, the most

likely cause of the difference in representation is a combination of its location within the swath and the corrections used by the software.

Overall, there is a difference in resolution between the 2019 (200 kHz) and 2017 (300 kHz) data, which is to be expected as higher acoustic frequencies typically provide better resolution. The difference is further enhanced by the fact that the active site was recorded in the outer part of the swath in 2019. In that year, HOMESIDE had already been upgraded to a dual-receiver configuration. This section was therefore insonified with a larger incidence angle, which further reduces the quality.

Figure 5.7 shows the second study area. It covers a large active site consisting of several discharge sites of hydrothermal fluids with multiply chimneys grown together into a beehive structure. About 400 m to the north-west is an inactive site with only small chimney remnants. Even more so than in the previous example, the outlines of the sites with indications on the seabed for hydrothermal activity cannot be clearly distinguished from the surrounding area in either mosaic. The mosaics at both sites show intermediate and homogeneous backscatter values. At both sites sulphide blocks and sulphide talus is predominate. Hydrothermal sediment has been mapped around the large active site.

Both sites are mostly surrounded by sediment and partly by the harder material (fragmented) pillow lava (with sediment). The hydrothermal sediment from the active site is indistinguishable from the non-hydrothermal sediment, as expected. The inner part of the site, where more sulphide blocks and sulphide talus have been mapped, is indistinguishable in the mosaic. However, towards the south and towards the area of predominantly (fragmented) pillow lava with some sediment, a distinction can be seen in the backscatter, especially in the 2019 data.

The inactive site has a slightly brighter backscatter level than the active site. It shows some difference from the adjacent sediment to the east and the pillow lava to the southwest. The higher backscatter values of the pillow covered area (highlighted in cyan in Figure 5.7 bottom right) are particularly visible in the 2019 data. As noticed in the previous example, the visual difference is less obvious in the mosaics at the very bottom where the areal AVG correction (CARIS_Area) has been used. This confirms the above conclusion: The difference in visual appearance is most likely due to differences in acquisition geometry. This particular area, which is part of a trench and coincides with the pillow lava cover, was acquired with the MBES in 2019 down- and upslope (i.e., with the ship's direction orthogonal to the slope). Typically, the AVG correction is applied to a larger number of consecutive pings. If the MBES swath covers the bottom and top of a larger bathymetric structure that also has differ-

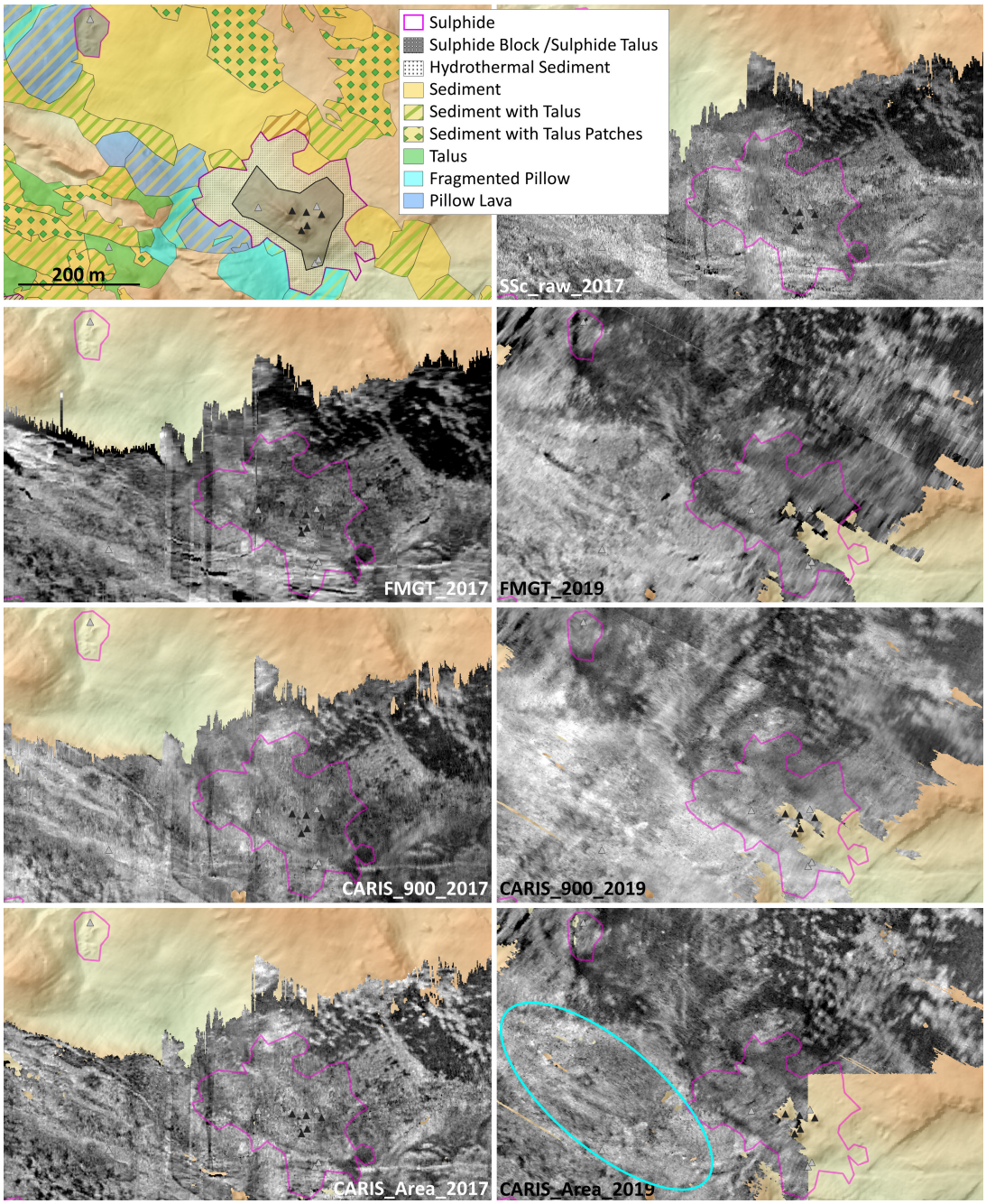


Figure 5.7: Example 2: Seabed classification and different backscatter mosaics (1 m resolution) created from data from different deployments in 2017 (300 kHz) and 2019 (200 kHz) and with various software and software settings (SSc_raw = *IFREMER SonarScope* without correction, FMGT = *QPS FM Geocoder Toolbox*, CARIS_900 = *CARIS HIPS and SIPS - SIPS Backscatter engine* with AVG 900, CARIS_Area = *CARIS HIPS and SIPS - SIPS Backscatter With Area Based AVG engine*). An area with greater differences in backscatter representation is highlighted in cyan.

ent seabed characteristics, the corrections in post-processing may reduce the difference in backscatter caused by the different characteristics.

In summary, it can be concluded from this example, as previously observed, that the discriminability of different seabed classes (especially of those showing evidence of hydrothermal activity) depends on how much neighbouring classes differ in their acoustic properties (e.g., soft/hard, smooth/rough). This

can only be partially influenced by the acquisition geometry or processing settings. For optimal backscatter acquisition, the acoustic signal (e.g., frequency, pulse form, pulse length, gain) should be kept stable (to avoid artefacts), the highest frequency should be used for high resolution, and areas of interest should be covered at an incidence angle of about 20° to 40°. As the focus of HOMESIDE deployments is on collecting bathymetry and water column data, the settings for the backscatter data are not optimised and therefore the quality of the final mosaics is compromised.

Each of the backscatter processing software suites tested had advantages and disadvantages. *FMGT* is prone to smearing effects, probably caused by some filtering or blending algorithm. These effects slightly reduce the resolution of the mosaic. The same AVG approach in *CARIS* shows no such effects, but struggles to compensate for the dual-receiver setup in the nadir area. The *CARIS area-based* AVG algorithms gave very good outcomes, but can result in small features being more prominent than larger backscatter patterns in the data. This can be influenced by adjusting the parameter updating size during the processing, but due to the long processing time and the likelihood of stripe shaped artefact, this needs to be done with careful observation of the data and the results. It also showed a larger and artificial cut out, which is most likely caused by a bug in the software's gridding (Figure 5.7 bottom right).

5.3 Textural Analysis

The previous chapters have shown that a numerical backscatter analysis of the high-resolution MBES backscatter data of the INDEX data cannot be used directly to identify hydrothermal sites. Another approach focuses not only on the backscatter values within a certain area, but also on their relative position to each other. The variation between neighbouring pixel values – their spatial organisation – provides information about the textural properties, which can be rough, smooth, homogeneous, or random for example (Blondel et al., 1998).

Several tools have been developed for texture analysis. One of the most widely used methods is GLCM (Grey-Level Co-occurrence Matrix) after Haralick et al. (1973); Haralick (1979). It has been successfully applied in the field of remote sensing for terrestrial and marine applications (Wang et al., 2016; Blondel et al., 1998). The GLCM lists how often different combinations of two pixel values occur in a given distance and direction from each other within a small section (window) of an image. Different measures can be derived from GLCM, taking into account the spatial direction (second order) or not (first order). Typical first (^{1}) and second order (^{2}) texture measures according to Haralick et al. (1973), Hall-Beyer (2017a), and GRASS Development Team (2024) are:

- **Contrast**[2] provides information about the difference in intensity between neighbouring pixels. A high contrast indicates a large grey value variation between neighbouring cells and therefore a coarse texture or sharp transitions/edges, whereas a low contrast suggests a smooth texture.
- **Homogeneity** / **Inverse Difference Moment (IDM)**[2] measures how similar neighbouring pixels are. It is inversely related to contrast. High values indicate smooth and homogeneous textures.
- **Angular Second Moment (ASM)** / **Energy**[2] gives information about the uniformity or orderliness of a texture and is the opposite of entropy. It is the sum of the squared elements of the GLCM. A high ASM value indicates that the pixel values are similar.
- **Entropy**[1] expresses the degree of randomness or disorder, predicting the unpredictability and complexity of a texture. It returns a high value when the pixel values have a random texture. The number of occurrences of pixel pairs is therefore relatively equal throughout the GLCM and will therefore have lower values. It is typically inversely correlated with ASM.
- **Variance**[1] gives the dispersion of the values around the mean.
- **Correlation**[2] measures the linear dependency of neighbouring grey levels. A high correlation indicates that the pixel relationship is predictable.

Haralick et al. (1973) defines 14 different texture measures and there are generally many more presented in further literature. In practice, however, only some of them are used because they are correlated. The choice of measures depends on the individual data set and the intended classification. Clausi (2002) and Hall-Beyer (2017b) for example recommend the combined use of contrast, correlation, and entropy for terrestrial applications. Similarly, entropy and homogeneity have been successfully applied for texture analysis of side scan sonar data and MBES backscatter data (Blondel et al., 1998; Huvenne et al., 2007). In order to derive different texture classes from the texture measures, clustering can be used for example (Blondel and Gómez Sichi, 2009).

GLCM and texture indices were calculated in QGIS 3.36.3 using the GRASS GIS implemented algorithms (QGIS Development Team, 2024; GRASS Development Team et al., 2024). In a first step, the backscatter data was normalised and grey values were reduced to 11 (for entropy) and 40 values (for all other GLCM textures), to facilitate the GLCM computation. Different texture measures with varying window sizes (3, 5, 25) and distances (1, 2, 10) were generated for a texture analysis of the INDEX data. This was done to investigate whether, and if so which, texture measures can be used to distinguish areas of hydrothermal activity from non-hydrothermal active areas. Figure 5.8 shows some examples of texture measures as well as the ruggedness derived from

the backscatter data in areas of hydrothermal sites. Ruggedness is calculated as the square root of the sum of the squared difference from each neighbouring cell and is an indicator of raster complexity/variability (Riley et al., 1999).

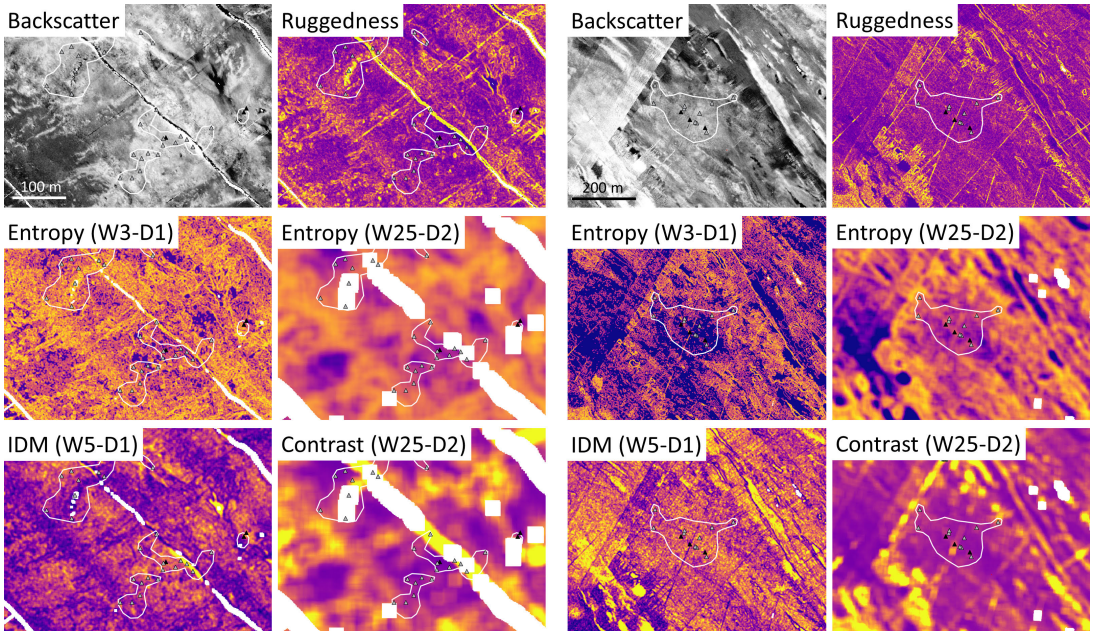


Figure 5.8: Backscatter, backscatter ruggedness, and various texture measures with different GLCM window (W) and distance (D) sizes of hydrothermal areas (site extent marked in white).

The selected examples in Figure 5.8 illustrate two problems with backscatter data when generating derivatives such as the texture measures: Data gaps (left example) and data artefacts (right example). Gaps typically occur in the INDEX data where the swath overlap is insufficient or in areas where a slope faces geometrically away from the echo sounder and it cannot be ensonified (i.e., shadowing). The derived texture values at data edges may be unrepresentative and biased due to the lack of data continuity. This can affect numerical calculations and derived statistics. When working with large window sizes for the GLCM computation, the gaps become even larger, resulting in a loss of information. The second data problem – artefacts – occurs when intensity levels change abruptly. This typically occurs when the system automatically adjusts gain and signal settings to improve the bottom detection. This can happen, for example, when the altitude changes or when individual pings fail to detect the bottom. Such sudden changes in intensity will affects the texture measures and therefore further numerical analysis.

Overall, a visual inspection of the INDEX hydrothermal sites and the backscatter derived measures reveal some similarities between certain sites. At most sites, the entropy (W3 D1) value is high in the vicinity of active chimneys, indicating a random texture pattern. This can be due to the presence of chimneys and the resulting effects of shadowing and direct reflection of the acoustic signal. Areas around the sites identified as sulphide block in the geological maps,

are often represented by low entropy values. This area is typically covered in sediment-like material from plume precipitation, making the texture more homogeneous. At some sites, the absence of locally high entropy values can be noted, as in the examples in Figure 5.8. This indicates a smaller or less pronounced chimney or chimney remains/blocks directly at the site location, as the texture is more homogeneous.

The values of IDM (also called homogeneity) are mostly inversely related to entropy, with low values and therefore low homogeneity directly at the active vent sites with larger chimneys. As contrast is inversely related to IDM, it does not provide any further insight into the data. Ruggedness generally shows lower values at sites. It does not appear to be very sensitive to small variations, but tends to show high values in areas of very rough seabed class changes (not taking into account data gaps and artefacts).

The seafloor in the INDEX area is generally rough and inhomogeneous due to the high tectonic activity and the thin sediment cover. Hydrothermal areas cannot be distinguished from non-hydrothermal areas by general numerical values or visual patterns using texture measures. There is no clear visual distinction between sulphide cover and regular seabed. In particular, data gaps and artefacts distort an areal analysis as the gaps increase. Small window sizes should be used to reduce the gaps in the GLCM derivatives. Although areal texture analysis is not feasible for the INDEX data, the use of backscatter texture properties after mound classification using deep learning may be beneficial. In this case, the texture properties are only used from a small local area – covering the pixels of a detected mound – to distinguish between sulphide and volcanic mounds.

5.4 Chapter Summary

Seafloor areas showing indications of hydrothermal activity do not have a unique backscatter or backscatter texture signature. Their representation in the backscatter data is highly dependent on the characteristics of the surrounding and how much they differ. Another factor in distinguishing sites from the general seafloor is their nature: typically, seabed types have a smooth transition between each other rather than a sharp boundary. They also tend to be mixed. The acoustic response of hydrothermal areas can be very similar to the other seabed types.

Whether a site is active or has been extinct for a long time is important, as weathering and erosion alter its characteristics. In addition, debris can (partially) bury sites. It has also been noticed that large chimneys can cause shadows in the backscatter data. However, their visibility depends on the acquisition

geometry and large chimneys do not occur at every site. Therefore, this particular feature cannot be generalised to all sites and incorporated into a possible method in backscatter mosaics.

The overall acquisition geometry, MBES settings, and post-processing corrections do affect the final mosaic. However, these factors do not significantly alter the representation of the individual seabed classes. It is more important to ensure that artefacts are removed as good as possible in post-processing to reduce sudden changes in backscatter levels that are not caused by variations in the seabed.

This chapter shows that backscatter data allows certain conclusions to be drawn about seabed characteristics (e.g., hardness), but cannot be used alone to identify sites of (former) hydrothermal activity. Such an approach might work for certain sites that are (or were until recently) active over a long period of time and where the seabed characteristics differ significantly from the surrounding seafloor.

However, MBES backscatter data is acquired together with bathymetry. At hydrothermal sites, mound structures are typically formed by precipitation of the dissolved metal particles from the plume and falling chimneys (Chapter 3.1). An approach using both – bathymetry and backscatter data – is likely to facilitate the detection of hydrothermal sites in high-resolution MBES data. Such a bathymetry-based mound detection for the identification of hydrothermal sites using deep learning methods has been investigated and tested by several research groups. In the following chapter, such a deep learning approach is tested on the INDEX data.

Chapter 6

Automated Detection and Classification of Sulphide Mounds

Most exploration methods for hydrothermal venting sites and associated sulphide deposits are based on the detection of chemical anomalies caused by active venting sites. In this chapter, an approach is tested that firstly uses deep learning to detect mounds based on bathymetry. Secondly, these derived possible exploration targets are differentiated by type using bathymetry and backscatter data, to extract possible sulphide mounds. This approach aims to reduce the required time if all mounds had to be evaluated manually. In addition, such an automated process also promises to increase the objectivity and consistency of such classification.

In the first section of this chapter (6.1), some similar approaches found in the literature are presented. They also utilised CNN to identify possible sulphide deposits. The second section (6.2) presents the data preparation for the CNN used, which includes the creation of data derivatives as input layers, but also the labelling of the training data. The third section (6.3) presents the results. First, a pre-trained and publicly available U-Net model is used before the model is trained with the study data. Section 6.4 deals with the subsequent differentiation of volcanic and sulphide mounds before summarizing the whole chapter.

6.1 Literature Review

The task of first detecting the mounds and then identifying the sulphide ones among them is not trivial. However, such approaches have already been tested and applied. Three studies are presented in greater detail in Table 6.1.

All of them include a CNN, mostly U-Net, for detection using a segmentation. U-Net was originally developed for biomedical image segmentation (Ronneberger et al., 2015). Since then, it (or further developed versions) has been

used in different fields of application such as landform classification studies to distinguish between geomorphological types (Li et al., 2020; Torres et al., 2020) or also in marine applications for target segmentation (Tang et al., 2023). Also, in a comparison by of different approaches (heuristic and deep learning) for peak detection on land, U-Net showed the best results (Torres et al., 2020).

All three approaches in Table 6.1 are based on high resolution bathymetry. Mounds are complex features and their detection is not easy using only elevation data (Juliani and Juliani, 2021). Therefore, in all three approaches, multiple derivatives (also called secondary features) from the bathymetry are used in all three approaches as input layers for the CNN.

Distinguishing between sulphide and volcanic mounds from bathymetry alone is even more difficult. As illustrated in Chapter 3.1 with examples of sites in *Field 1*, hydrothermal mounds can most commonly be distinguished from volcanic domes in bathymetric data by the slope of the peak: Sulphide mounds have steep and conical tops with irregular sides (Jamieson et al., 2014). Volcanic mounds have a dome-shaped top.

The paper presented by Juliani and Juliani (2021) includes an evaluation and several experiments on the method applied for detection. The U-Net hyperparameters are discussed and tested in more detail to optimise the approach. The study uses high-resolution bathymetry data of a 20 km^2 area at the Mohns Ridge, covering the Loki's Castle venting site. The elevation data is transformed into a seven-dimensional input data set for the semantic segmentation using U-Net. In a semantic segmentation, each pixel value is assigned to a class, here: "mound" and "no mound". In a second step, a similarity analysis is performed and the segmented mounds are clustered using the convolutional signals. A final comparison with the ground truth data identifies the group containing the most known sulphide mounds.

The approach of Haroon et al. (2023) focuses more on the estimation of sulphide deposits. The elevation data is transformed into a three-dimensional data representation for the training, evaluation, and testing. As in Juliani and Juliani (2021), the U-Net is used to identify the mounds, but in addition also magnetic, controlled source electromagnetic, and in-situ data is used for the subsequent clustering and classification. The study area is the hydrothermal field TAG, located at the MOR of the Northern Atlantic.

Both of the above approaches use the same data resolution of 2 m and a patch size of 256 x 256 pixels, in which the input layers are clipped. The approach of Keohane and White (2022) focuses more on chimney detection. The original data resolution is slightly better at 1 m and the patch size for the CNN is only 15 x 15 pixels, as this better represents the dimension of the object to be detected. In addition to the chimneys, other landforms such as flat terrain, fault

scarp edges, or round features are also classified. No additional classification is performed after the detection.

Table 6.1: Overview of different approaches using machine learning methods for (sulphide) mound detection.

	Method 1 Juliani et al. (2021)	**Method 2 Haroon et al. (2023)**	**Method 3 Keohane et al. (2022)**
Method	U-Net	U-Net [modified after Juliani et al. (2021)]	CNN "Chimney Identification Tool"
Aim of CNN	Mound detection	Mound detection	Detection of chimneys (flat, faults scarp edge, round feature)
Input raster layers	**7 bands:** 1) RGB $\rightarrow$ 3 PCs[1]: Slope, roughness, TPI [2] 2) RGB $\rightarrow$ 3 PCs: Shaded relief 3) Bathymetry	**3 bands:** R$\rightarrow$ Norm. aspect G $\rightarrow$ Slope B $\rightarrow \partial y$	**3 bands:** Bathymetry Slope TPI
Grid resolution	2 m	2 m	1 m
Tile size [pixel]	256 x 256	256 x 256	15 x 15
Data augmentation	Geometric: Scale, flip, transl., rotation Colour: Brightness, contrast, blur, noise	Rotation and flip	Rotation and flip
After CNN	Morphological similarity analysis and clustering of segmented features using convolutional signals (CV & Data Proc)	1) Mound classification based on mound architecture and magn. signature 2) Re-asses SMS volume via CSEM & in-situ data	–
Further data source	–	Magnetic, CSEM[3]	–
Location	Loki's Castle hydrothermal field & Axial volcanic ridge section, Mohn's Ridge, MOR, Arctic	TAG hydrothermal field, MOR, Atlantic (& other free data)	Endeavor Ridge & Galapagos Spreading Centre, Pacific
Data extent	Approx. 30 km^2 369 tiles	Multiple free data sets, test area: 49 km^2, overall 2,280 tiles	–
Annotated instances	1,115 mounds 899 peaks 7.6% (2.2 km^2) of area	–	528 hydrothermal chimneys
No of sites	2 sulphide mounds	Test area: 15 sulphide mounds	48 out of 528 annotated confirmed
Model results	Accuracy mound detection: 84% pixel-wise 80% object-wise	323 mounds detected 14 of 15 known sulphide mounds were classified as exploration targets	96% agreement with annotation, 39 of 48 chimneys correctly detected

[1] Three principle components resulting from a principal component analysis
[2] Topographic Position Index
[3] Controlled Source Electro-Magnetic

A large amount of labelled data should be available to train a CNN. Since the acquisition of high-resolution bathymetry in deep-sea areas is very time consuming and expensive, large area coverage is typically not available. Also in all three examples in Table 6.1 the area covered is limited. Juliani and Juliani (2021) uses data with a coverage of 29 km^2 for training, validation and testing, which includes only two known sulphide mounds. Juliani and Juliani (2021) further distinguishes between two classes: the peak (class 1) and the base of the mound (class 2). A total of 1,115 mounds and 899 peaks were annotated in the data set.

In Haroon et al. (2023) different 14 freely available data sets are used for training. However, these data sets differ in spatial resolution from 2 m to about 90 m. The testing data set used to evaluate the CNN performance covers an area of 49 km^2, including 15 SMS sites. In the "chimney identification tool", 528 chimneys were labelled for training, validation, and testing. 48 of these locations are confirmed chimney locations.

U-net has the advantage of being able to produce good results with relatively sparse training data (Torres et al., 2020). Nevertheless, all three approaches use data augmentation to increase the amount of training data. Typically, at least rotations and flips have been used to increase amount of data by a factor of eight (Table 6.1).

A detection focusing on chimneys, as in Keohane and White (2022), was not considered for the data within this study, as the focus is not on chimneys but on mounds. Not every hydrothermal site – especially the inactive ones – has chimneys of sufficient size to be visible in the bathymetry. The aim is to detect not only chimneys but all the mounds in general. Therefore the detection method published by Juliani and Juliani (2021) is used in this study. Haroon et al. (2023) shows that additional data sets improve the subsequent differentiation of the sulphide mounds from all detected ones. Therefore, in this study, the use of MBES backscatter data is used to identify sulphide mounds. the data preparation is presented in the following section.

6.2 Data Preparation

The input data for U-Net is prepared as described in Juliani and Juliani (2021): As a first step, the following geomophometric terrain attributes are created using high-resolution bathymetry with SAGA GIS 9.3.2 (Conrad et al., 2015):

- **Slope**: Slope represents the rate of elevation change in terrain, expressed as the angle between a horizontal plane and a tangential plane at the location of interest (Olaya, 2009).

- **Roughness** / **Terrain Ruggedness Index (TRI)**: This value describes the heterogeneity of a grid by expressing the change from one grid cell to its eight neighbouring cells. According to Riley et al. (1999), it is calculated for each cell by the square root of the sum of the squared difference to each neighbour.
- **Topographic Position Index (TPI)** / **Slope Position**: This value expresses the local morphology by comparing the focal point with the heights of the neighbouring cells. A positive value indicates that the location is higher than its surroundings and a negative value indicates that the location is lower than its surroundings. Values around zero indicate a flat surface or areas of constant slope (Weiss, 2000) (Evans, 2018).
- **Shaded Relief**: Shaded reliefs are commonly used to enhance the visualisation of topographic features by providing a 3D representation of the surface as if being illuminated by the sun from a given azimuth (e.g., direction) and elevation above the horizon. The pixel values for this secondary feature are determined by applying an insolation model based on slope and aspect. The more a location is illuminated, the higher its pixel value. (Olaya, 2009)

The results are then simplified using two Principal Component Analyses (PCAs) (Murtagh and Heck, 1987) in SAGA GIS 9.3.2. A PCA is an image processing tool that reduces the dimension and complexity of a raster input data set using four steps according to Eleftherakis (2013): The input data sets are standardised by subtracting the mean from each value and then dividing by the standard deviation. The covariance matrix is then determined. The diagonal elements have a value of 1 due to the previous standardisation process and the other elements represent the correlation coefficient between the individual features. In the third step, the eigenvalues are calculated before the data matrix is multiplied by the eigenvector matrix to derive the principal component (PC) matrix. A subset of PCs is then determined, consisting of the smallest number of PCs that contain most of the original data but are not correlated. They are the output of the PCA and are typically ordered by importance (Bahrenberg et al., 2008).

In this work, the rasters representing slope, roughness, and TPI (10x10, 20x20, 30x30) are used as input for PCA1, and the shaded reliefs with different illumination angles (North, 45°, 135°, 225°, 315°) are used for PCA2 (see also Figure 6.1). Each of the PCAs produces three PCs, which are then combined into two RGB composites. The individual PC results have been inverted where necessary to give the same visual representation for each area. Together with the bathymetry, they represent the seven input data set layers for U-Net.

For the **mound annotation**, polygons along the mound bases (class 2) and

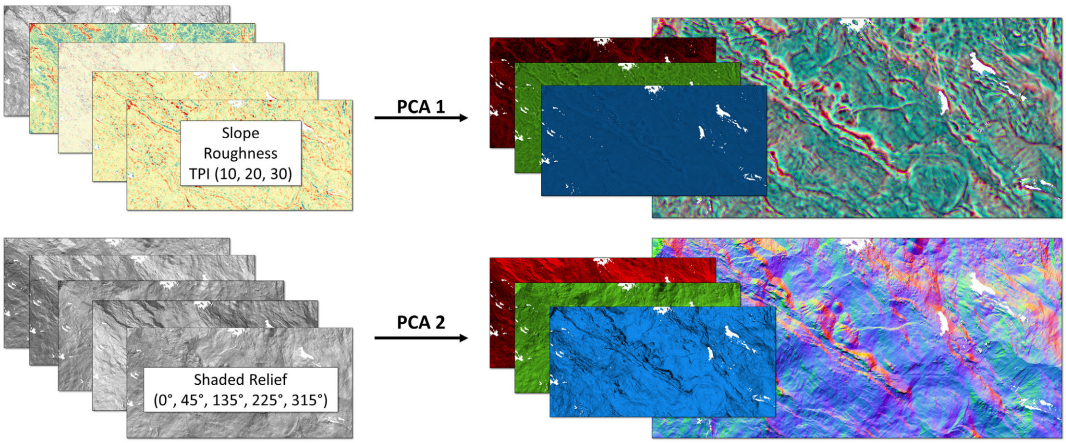


Figure 6.1: Depiction of the creation of two RGB data representations using PCA from different bathymetry derivatives (own illustration).

peaks (class 1) were manually generated in QGIS 3.36.3. Several layers were used in combination to identify the mounds and their bases and peaks: both RGB images, shaded relief, and contour lines at 1 m intervals. Typically, individual mounds with an elliptical base shape were annotated. While Juliani and Juliani (2021) limits the mound annotation to features with a diameter of 100 m, in this study mounds up to 250 m were annotated, as this was the maximum size of the known sulphide mounds observed in the data. Instances smaller than 20 m were not classified by base and peak, but only as peak. These small mounds were not equally present in all four data sets used. They are more abundant in data sets C and D. These morphological features can origin from boulder or more isolated chimneys or small mounds (Asada et al., 2015).

Volcanic domes can overlap to some extent – appearing to have "grown together". Such overlapping features have not been included because the bases of the individual structures cannot be identified. Also, only complete features that are not cut by fissures or faults have been considered. Mounds (especially sulphide mounds) may have multiple peaks. However, mounds without a clear peak are present in the data. Figure 6.2 shows some examples of labelled mounds.

The vector format of the mound annotation layer has been converted to raster. To prepare the larger data sets for PCA1, PCA2, elevation, and annotated mounds for U-Net, they need to be normalised and cut into small tiles of 256 x 256 pixels. This size has been shown to give the best results (Juliani and Juliani, 2021). Adjacent tiles overlap by 128 pixels to ensure that each feature is covered by at least one tile. Splitting this data into tiles and converting it to the format required by U-Net (8-bit and 16-bit integer) is done in QGIS. Tiles not containing any annotations are removed to improve the object-background relationship towards a more balanced data set. Table 6.2 gives an overview of the number of tiles, known sulphide mounds, and the labelled mounds.

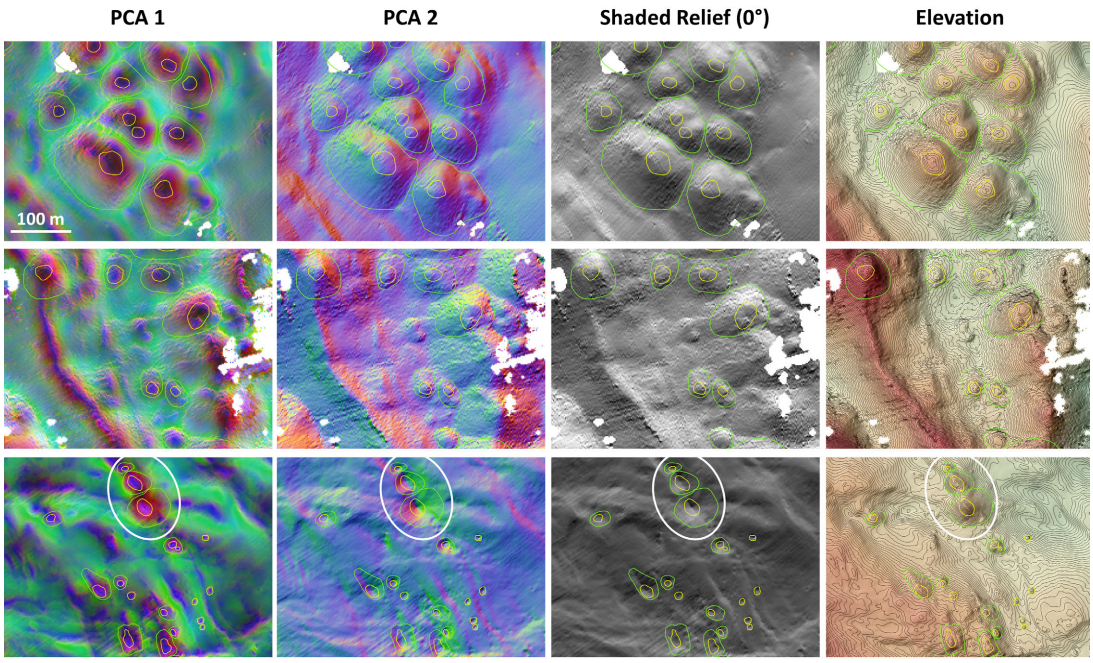


Figure 6.2: Examples of labelled mounds (class 1 = yellow / class 2 = green) with the different CNN input data as background. Sulphide mounds are highlighted with a white outline.

Table 6.2: Number of tiles (for each data layer), known hydrothermal sites/mounds, and annotated mounds for each data set.

Data Set	Field	Area	Tiles	Number of hydrothermal sites	Number of sulphide mounds (C1/C2)	Number of labelled mounds (C1/C2)
A	1 & 2	272 km^2	1,818	21	44 / 26	1,650 / 1,447
B	3	186 km^2	1,877	12	22 / 15	2,563 / 2,071
C	4	130 km^2	1,019	15	24 / 20	1,576 / 721
D	5 - 7	194 km^2	1,998	40	148 / 67	3,195 / 2,205
Sum:		**782 km^2**	**6,712**	**88**	**238 / 128**	**8,984 / 6,444**

A total of 6,712 tiles are available for the training of the U-Net model. The data covers 88 hydrothermal sites with 238 annotated peaks and 128 bases. In total, 8,984 mounds peaks and 6,444 bases were manually annotated. The following section presents the results of the CNN.

6.3 Mound Prediction

All four data set are segmented twice: first, using the pre-trained model from Juliani and Juliani (2021). Second, by using the INDEX data for training and validation of the model. Their results are presented separately in the next two sections.

6.3.1 Applying Pre-Trained Model to INDEX Data (No Training)

The pre-trained U-Net model from Juliani and Juliani (2021) is used for segmentation and thus prediction of mound locations in the INDEX data. The results allow an evaluation of the general universality of the model. The CNN provides prediction maps, which are raster representations of the study area with possible mounds coloured in red and background data in blue. The predicted mound locations are extracted and compared to the manually annotated mounds (ground truth). A labelled mound is considered correctly classified if the prediction and the annotation intersect. Additionally, the metrics *recall*, *precision*, and *dice coefficient* (DC) are also used to evaluate the segmentation results when comparing the annotated ground truth data with the predicted output (Figure 6.3) (Sokolova and Lapalme, 2009; Sørensen, 1948).

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$DC = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Ground Truth \ Prediction	Positive	Negative
Positive	**True Positive (TP)**	**False Negative (FN)**
Negative	**False Positive (FP)**	**True Negative (TN)**

Figure 6.3: Metrics to evaluate segmentation results.

Recall is also referred to as sensitivity and indicates the ability to detect all relevant positives: a high recall suggests that most annotated instances have been successfully detected (i.e., a low number of false negatives). **Precision** expresses the accuracy of positive predictions by showing how many of the output predicted instances are correct: a low precision indicates that many unannotated areas are wrongly classified (i.e., a high number of false positives). Ideally, both values should be high (i.e., close to one) (Sokolova and Lapalme, 2009). The **DC** is the harmonic mean of recall and precision. It is also called *F1 score* and is often used in image segmentation as this task typically involves imbalanced data sets. In imbalanced data sets, the instances to be detected are significantly smaller in number or area than those that should not be detected (i.e., the background). Since DC includes false negatives as well as false positives, it provides a more balanced view of a model's performance. As for recall and precision, a value of close to one is favourable as it represents a good similarity between prediction and ground truth.

Table 6.3 lists the metrics of the segmentation results with the pre-trained model. For each data set and class, the number of annotated mounds ("mounds ground truth") is listed next to the number of correctly predicted mounds, as well as the recall, precision, and DC values. These values are obtained by

comparing the resulting prediction map and the ground truth data (by intersection of the instances). The DC values are also provided based on a pixel by pixel evaluation of the model. The numbers for the detected sulphide mounds are obtained visually from the data.

Table 6.3: Segmentation results (class 1 = peak and class 2 = base) using the pre-trained model without additional training. The majority of the values listed are derived from an instance-wise comparison (intersection = correct identification). Only the DC value is also provided by a pixel-wise evaluation.

Data Set	Class	Mounds					Sulphide Mounds	
		Ground Truth	Predicted Correctly	Recall/ Precision	DC	DC (pixel)	Ground Truth	Predicted Correctly
A	1	1,650	1,315 (80%)	0.80 / 0.09	0.16	0.06	44	31 (70%)
	2	1,447	1,393 (96%)	0.96 / 0.09	0.16	0.13	27	24 (89%)
B	1	2,563	1,896 (74%)	0.74 / 0.12	0.21	0.07	23	16 (70%)
	2	2,071	1,850 (89%)	0.89 / 0.13	0.22	0.16	15	14 (93%)
C	1	1,576	789 (50%)	0.50 / 0.07	0.12	0.09	24	21 (88%)
	2	721	672 (93%)	0.93 / 0.06	0.12	0.21	20	18 (90%)
D	1	3,195	2,123 (66%)	0.66 / 0.20	0.31	0.06	148	59 (40%)
	2	2,205	1,973 (89%)	0.89 / 0.19	0.31	0.14	67	41 (61%)

It is possible to draw some general conclusions from the results:

- **All data sets have a high recall and a low precision.** Ideally, both should be high and close to one. The recall values indicate that most of the annotated mounds are detected (high sensitivity and low number of false negatives). The success rate is corresponds to that of the study by Juliani and Juliani (2021), where about 80% of the instances are correctly predicted. The precision values suggest that most of the instances predicted by the model do not overlap with any ground truth annotation and therefore result in a high number of false positives. Taken together, this means that the model detects many more mounds than are annotated for the ground truth, which facilitates the value of true positives. The recall values of class 2 (mound base) are higher than those of class 1 (mound peak) and are therefore more successfully identified by the model.
- **The DC values are low – both for the instance and for the pixel-wise detection evaluation.** This fact indicates a low similarity between predicted and annotated instances. As DC takes into account recall and precision, and thus false negatives and false positives, it provides a more balanced view of the model performance.
- **Overall sulphide mound detection is within the range of general mound detection.**

- **All four data sets have similar metric values, although they may differ in terms of the geological setting and hence local morphology.** Such local variations in morphology influence the appearance of individual (sulphide) mounds and their surroundings. The sulphide mound detection in data set D is significantly lower than in the other data sets, suggesting a more "atypical" morphology for this area.

These observations derived from the metrics can also be noticed when visually inspecting the data derivatives and especially the prediction maps with the overlayed ground truth data (Figure 6.4). The most obvious observations are that [1] the number of predicted instances is significantly higher than the number of annotated instances – for class 2 by a factor of 5 to 15 and [2] the predicted mound outlines are larger and rounder than those of the ground truth. This can have two reasons: first, the differences in the characteristics of the original training data to the one of this study, and second, the differences in operators and, consequently, in the way the mounds are annotated. It can be assumed that both causes together contribute to the noted differences between prediction and ground truth.

One of the differences in characteristics between the original training data used for the model and the INDEX data may be the number of isolated class 1 objects (peaks) observed in the latter one that are not surrounded by a class 2 object. Also in Table 6.3, a difference in the number of class 1 and 2 instances between the data sets can be seen: For all data, the number of class 1 objects is higher than the number of class 2 objects. But where the class 2 instances for data sets A and B are about 88% and 81% of the class 1 instances respectively, the ratio is about 45% and 69% for data sets C and D. Both of these data sets have a larger number of isolated peaks with a diameter less than 20 m.

These isolated class 1 objects seem to be more difficult for the model to detect, as the correct prediction rate for them is lower in Table 6.3 and as it can also be seen in Figure 6.4 (middle column). It can therefore be concluded that the training data did not contain such small isolated bathymetric features. Whether they did not exist or were not labelled by the operator cannot be concluded. However, they are included in the annotation of the INDEX data because some of them have been identified at hydrothermal sites in data set D. Probably caused by chimneys (or their remnants). Although these isolated peaks are also present in data set C, they are not as abundant in the hydrothermal areas (Table 6.3). In areas of non-hydrothermal activity, they most likely originate from large boulders.

The influence of the different operators and labelling approaches can be seen by comparing the shape of the predicted instances. The predicted peaks are generally larger and have a more circular shape than the labelled ones (Figure

6.4 middle column).

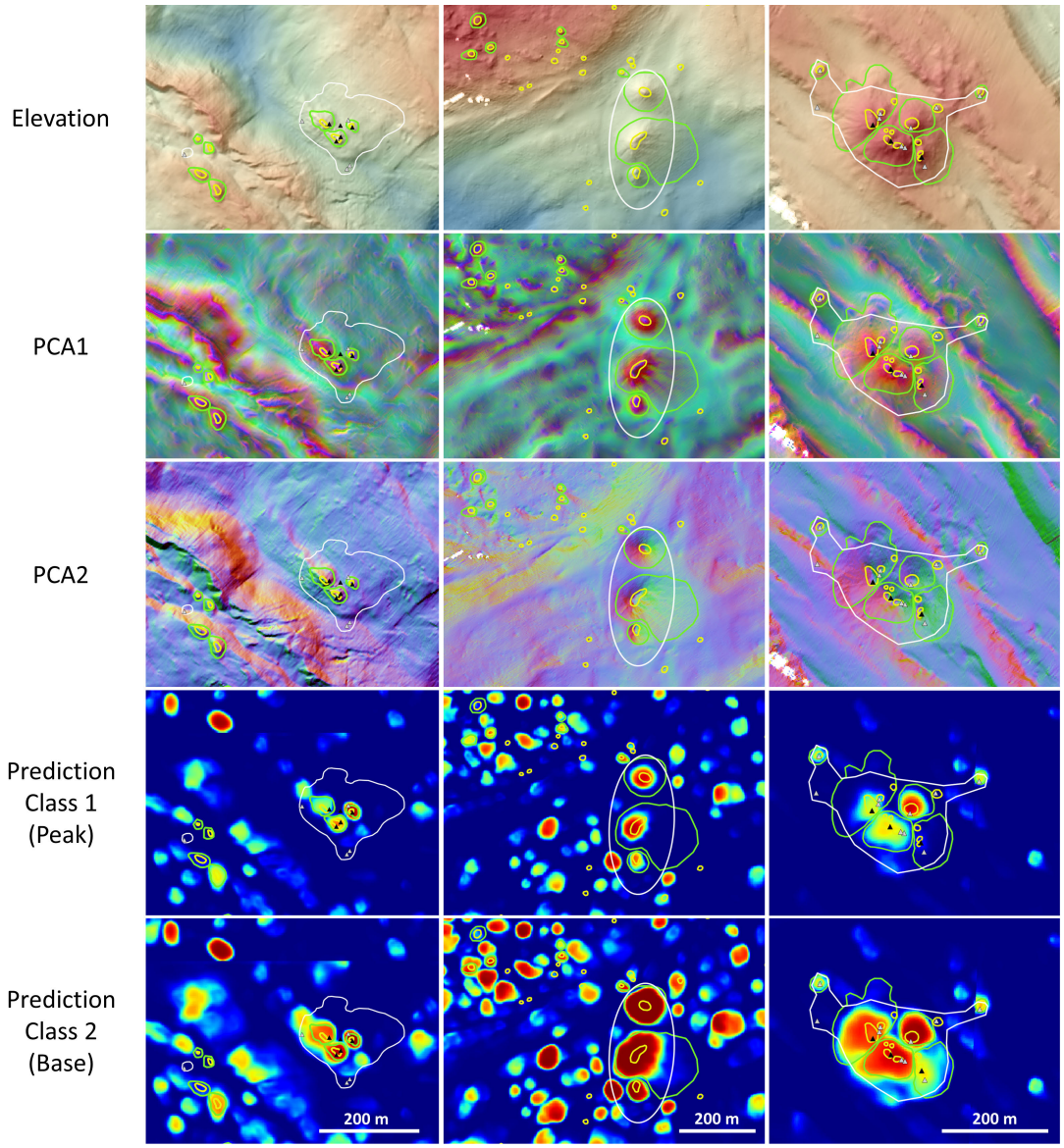


Figure 6.4: Examples of annotated mounds (class 1 = yellow, class 2 = green) in areas of hydrothermal sites (white outline).

In summary, the previously trained U-Net model achieves overall good segmentation results for the mound prediction. It works particularly well for isolated and distinct mounds with sizes in the range of base diameters from 40 m to 150 m. However, the overall sensitivity of the model is too high, as mounds are also predicted on very weakly developed seabed structures. Overall, a more reliable model is required to reduce the number of false positives. Therefore, the following section presents the results of a U-Net model trained on the INDEX data.

6.3.2 Training U-Net For Mound Detection

In order to improve the model and therefore the mound prediction, U-Net is trained with INDEX data. Of the total 6,712 tiles, approximately 15% were ran-

domly selected for validation and 15% for testing. The tiles were all normalised and automatically augmented geometrically and colour-wise using the script from Juliani and Juliani (2021).

Within each training epoch, all training tiles are passed through and the weights are updated to improve the model. Validation is performed after each epoch to obtain information about the current performance of the model. The model can then be tuned according to the validation results. This ensures, for example, that the model is not over- or underfitting – meaning that it is too well adapted to the training data or not trained enough to recognise underlying patterns. The test data is used at the very end, after the training, for an independent evaluation of the final model using data that was not included in the training phase.

To improve the training and thus increase the performance of the final model, the hyperparameter settings are crucial (Passos and Mishra, 2022). These are parameters describing for example training parameters, network architecture, or training data (Barnefske and Sternberg, 2022). Their settings control the training process. While the network architecture (e.g., layer types, layer combination) was not changed, different training parameters were tested. For the final training, the batch size was set to 4, the dropout rate to zero, and the number of epochs to 500.

During training, it was noticed that the imbalanced distribution of areas covered with and without mounds (object vs. background) poses a problem. The resulting model tends to struggle to classify the mounds correctly, but provides a more pessimistic predication, as predicting only the background area ("no mound") provides better model accuracy metrics. To improve the model results, empty tiles without mounds were removed. This process reduces the imbalance and therefore improves the overall results. Other possible solutions to this problem of imbalanced classes include (a) data augmentation of the training data so that the under-represented class is artificially increase, (b) algorithm-based methods (e.g., by modifying the network architecture), or (c) hybrid methods (combination of method- and algorithm-based) (Barnefske and Sternberg, 2022).

There are several possibilities and metrics to evaluate the training of a deep learning neural network model. One way is to look at the learning experience plot, which shows, for example, the average training and validation loss or accuracy over epochs, as shown in Figure 6.5.

The loss is calculated by comparing the prediction with the annotated data. A small value close to zero indicates a good learning result (Passos and Mishra, 2022). This applies for the training and validation loss shown in Figure 6.5 left. Both losses are low and slightly decreasing, suggesting that the model is not

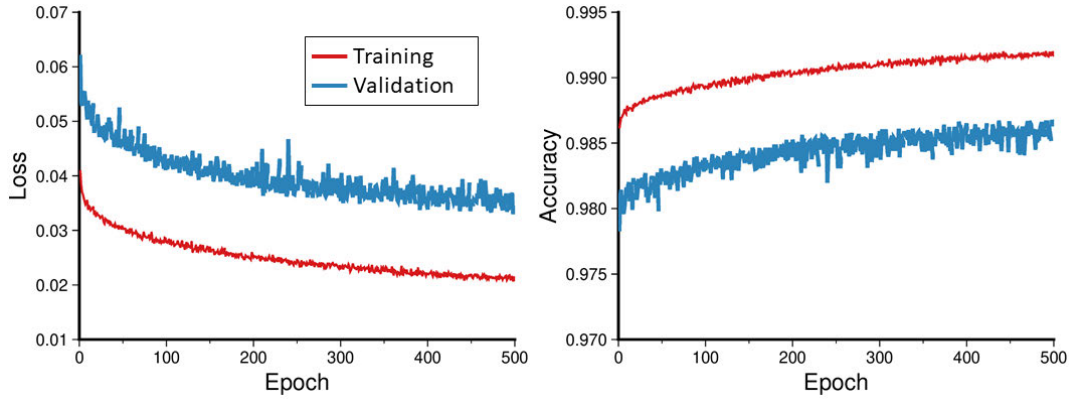


Figure 6.5: Loss (left) and accuracy (right) graphs for training and validation.

underfitting (suggested by constantly high loss values or strong decreasing until the last epochs) or overfitting (decreasing training and increasing validation loss). The training loss is constantly lower than the validation loss, which is typical for training deep learning models. It is also referred to as the generalisation gap (Brownlee, 2019). It indicates how well the model trained on the training data can perform on the "new" validation data.

Similar behaviour can be seen in the training and validation accuracy graph. Accuracy is a measure of how well a pixel is correctly classified (object or background) and should ideally be high as shown in Figure 6.5 right. During the learning process, the accuracy increases, with the training data showing a slightly higher accuracy than the validation data. This is consistent with the observations from the loss curves. The other learning metrics (recall, precision, DC) also show the same trend: improvement over the learning process with training showing higher (better) values than the validation data set (Figure 6.6).

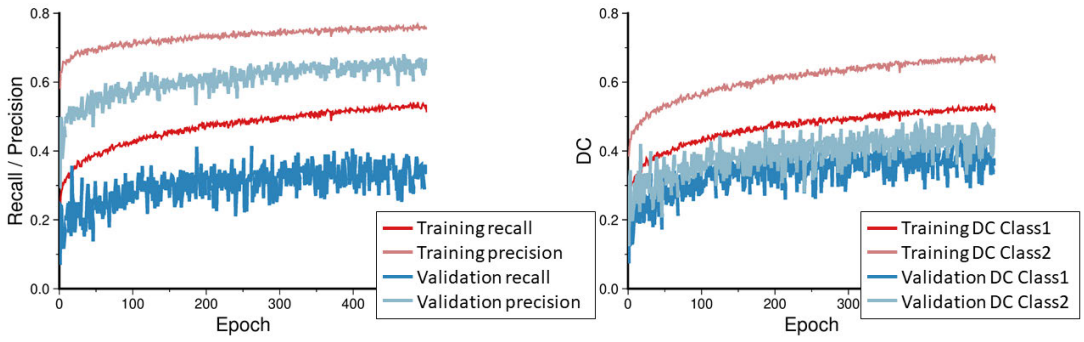


Figure 6.6: Left: Graph showing the recall and precision for training and validation. Right: Graphs showing the DCs for training and validation of class 1 and 2.

Strictly speaking, none of the displayed curves reach a point of stability. This is typically expected in the learning plots. It could be due to the small range of loss/accuracy values in the graph. A decrease of 0.004 in loss and an increase of about 0.0025 in accuracy over 250 epochs can be considered as stable. The

number of epochs and therefore the learning time could have been increased to investigate the behaviour of loss and accuracy. However, U-Net, and particular the data augmentation configuration used, is optimised for a small amount of training data. The number of epochs is already high for the large amount of data. So no significant improvement can be expected and the risk of model overfitting increases. In fact, given low values already achieved at the beginning of the training, the learning phase can be reduced to about 200 epochs to decrease computation time.

The pre-trained model is used as the initial model for the training with the INDEX data (transfer learning). This fact may also contribute to the low initial loss and high accuracy values. However, also a new model was also trained. The predictions did not deviate significantly and the loss values were only slightly higher in the first 15 epochs before reaching similar values as with the pre-trained model.

The training is carried out four times. Each time a different of the four available data sets is left out of the training and used only for testing. Except for data set D. Due to its size and high number of sulphide mounds, only the southern half was used for testing. The testing area includes most of hydrothermal *Field 5*, all of *Field 6*, and two sites from *Field 7*. All results (also again for the pre-trained model segmentation) are listed in Table 6.4. The same metrics are used as in the evaluation of mound prediction using the pre-trained models.

The statistics in the table above show that the newly trained model has a lower success rate in detecting mounds compared to the labelled ones (recall). Only about 55% of the annotated mounds are detected when training the model, compared to 81% when using the pre-trained model (averaged also over the two classes). However, the precision is three to four times higher with own training, suggesting that the number of false positives is much lower than without training. This is also indicated by the increase in DC. The new model provides a prediction that is overall closer to the ground truth. To summarise the metrics, the pre-trained model predicted a large number of instances, resulting in a high value of true positives (high recall) but also of false positives (low precision). The newly trained model is more conservative and does not predict mounds as easily. This more conservative approach results in fewer true positives, but also in fewer false positives, and an overall better fit between prediction and ground truth (high DC).

Unfortunately, this comes with a subsequent reduction in the number of correctly predicted sulphide mounds. The newly trained model correctly predicts only about a quarter of the peaks and less than half of the bases of the annotated sulphide mounds. Figure 6.7 shows the segmentation result, with likely mound locations are coloured red and the background blue. The ground truth

Table 6.4: Segmentation results (class 1 = peak and class 2 = base) using a model trained with the INDEX data ("With Training") and another (pre-trained) model which was not trained with INDEX data ("Without Training"). Most entry values originate from a instance wise comparison (intersection = correctly identified). The pixel-wise DC was provided as statistical output from testing and refers to a pixel-wise evaluation.

			Mounds					Sulphide Mounds		
	Training	Class	Ground Truth	Predicted Correctly	Recall/ Precision	DC	DC (pixel)	Ground Truth	Pred. Corr.	Pred. Add.
A	With	1	1,650	944 (57%)	0.57 / 0.41	0.48	0.49	44	22 (50%)	1
		2	1,447	1,305 (90%)	0.76 / 0.46	0.58	0.55	27	17 (63%)	0
	Without	1	1,650	1,315 (80%)	0.80 / 0.09	0.16	0.06	44	31 (70%)	12
		2	1,447	1,393 (96%)	0.96 / 0.09	0.16	0.13	27	24 (89%)	8
B	With	1	2,563	1,083 (42%)	0.42 / 0.74	0.54	0.45	23	6 (26%)	0
		2	2,071	1,141 (55%)	0.55 / 0.63	0.59	0.56	15	7 (47%)	0
	Without	1	2,563	1,896 (74%)	0.74 / 0.12	0.21	0.07	23	16 (70%)	4
		2	2,071	1,850 (89%)	0.89 / 0.13	0.22	0.16	15	14 (93%)	4
C	With	1	1,576	541 (34%)	0.34 / 0.64	0.45	0.45	24	8 (33%)	1
		2	721	429 (60%)	0.60 / 0.43	0.50	0.53	20	9 (45%)	0
	Without	1	1,576	789 (50%)	0.50 / 0.07	0.12	0.09	24	21 (88%)	8
		2	721	672 (93%)	0.93 / 0.06	0.23	0.21	20	18 (90%)	13
D	With	1	2,009	909 (45%)	0.45 / 0.78	0.57	0.36	104	27 (26%)	1
		2	1,566	903 (58%)	0.58 / 0.67	0.62	0.41	43	19 (44%)	13
	Without	1	2,009	1,508 (75%)	0.75 / 0.21	0.33	0.19	104	42 (40%)	4
		2	1,566	1,432 (91%)	0.91 / 0.21	0.34	0.37	43	28 (65%)	17

data (own annotations) is also shown. This visual representation underlines the conclusions drawn from the statistics when comparing the segmentation from the pre-trained mode (top) and the own trained model (bottom): The self-trained model is more conservative.

Another observation can be made from these visualized segmentation results: The predicted mounds are larger when the pre-trained model is used. The segmented outlines fit well to the annotated boundaries of the predicted mounds by the self-trained model. This must be taken into account when comparing the two models, as the success rate is measured by the intersection of the prediction and the ground truth. An overly large predicted mound might intersect with several annotations, all of which are then considered to be detected correctly. This may falsify the statistics in favour of the pre-trained model. An example of such a situation is depicted in Figure 6.7 for class 1 in a larger hydrothermal site: the pre-trained model predicts one large peak that intersects with three labelled ones. Therefore, all three are statistically registered as correctly recognized.

In conclusion, three statements can be summarized as outcome:

- **The model trained with INDEX data has an overall better performance (higher DC, similar value for recall and precision), but the success**

rate for true positives (recall) is low with less than 50%. The difference in model performance may be due to different labelling of the mounds for training. The INDEX data appears to be more cautiously labelled, so the resulting model is more conservative and the number of predictions is lower. However, a difference in morphology may also be considered.

- **Mound peaks (class 1) are more often annotated in ground truthing than bases (class 2) in all data sets.** However, they also have a lower prediction success rate for prediction in all data sets. This may indicate that the different layers used as training input data are not optimal for the model to recognise and learn their distinct features.
- **The relatively low success rate in predicting sulphide mounds using the self-trained model suggests that the model has not been successful in learning to recognise specific morphological structures.** The number of volcanic domes in the data is much higher than the number of sulphide mounds, and the latter often deviate from the stringent shape. In particular, sulphide mounds with more complex shapes that deviate from the typical cone shape are not recognised by the model. Figure 6.8 shows an example of a site with more complex mound structures, where only two out of eight peaks and three bases are recognised by the model.

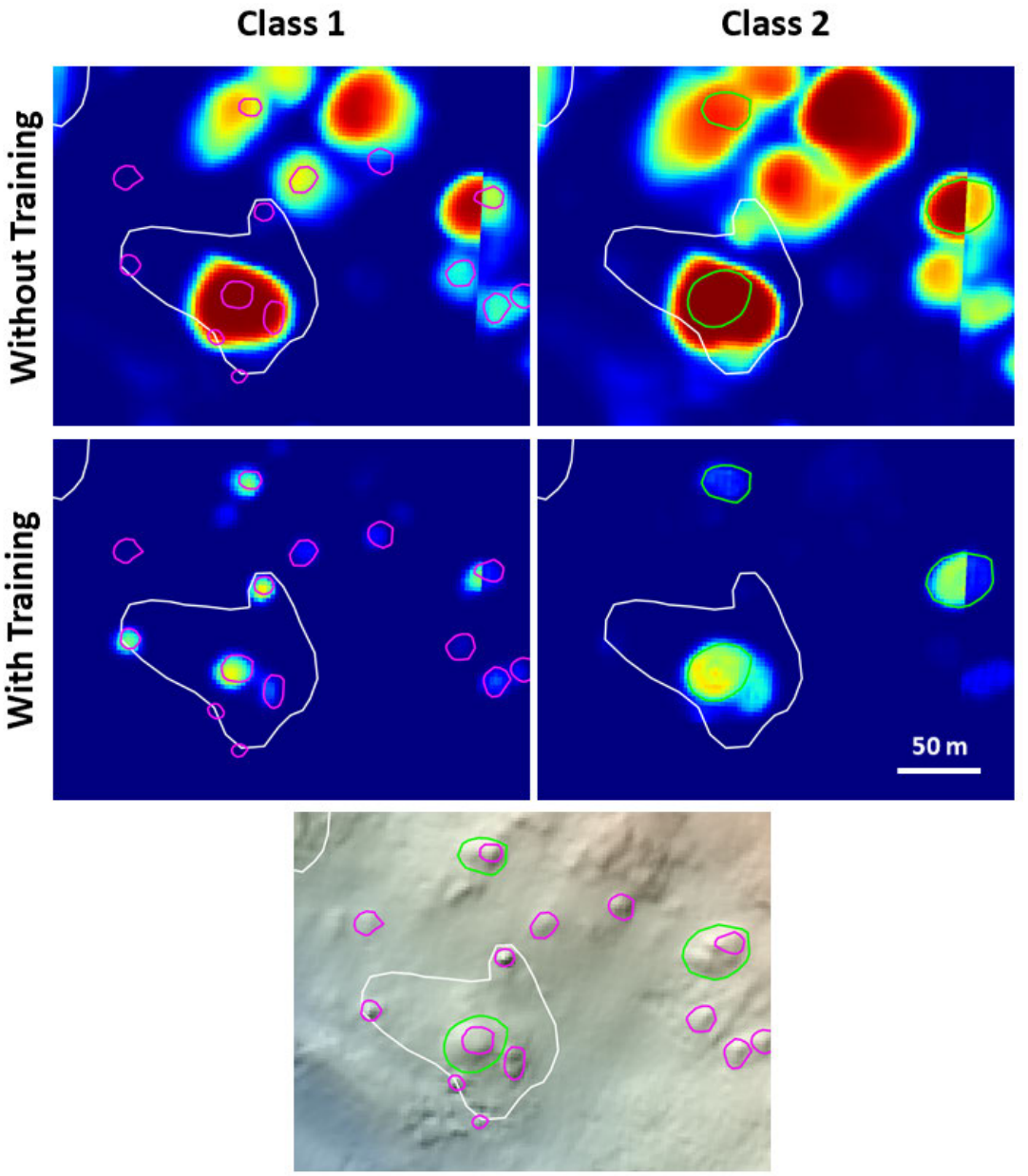


Figure 6.7: Prediction of a site (white outline) with (bottom) and without (top) model training for class 1 (left, annotation in pink) and class 2 (right, annotation in bright green).

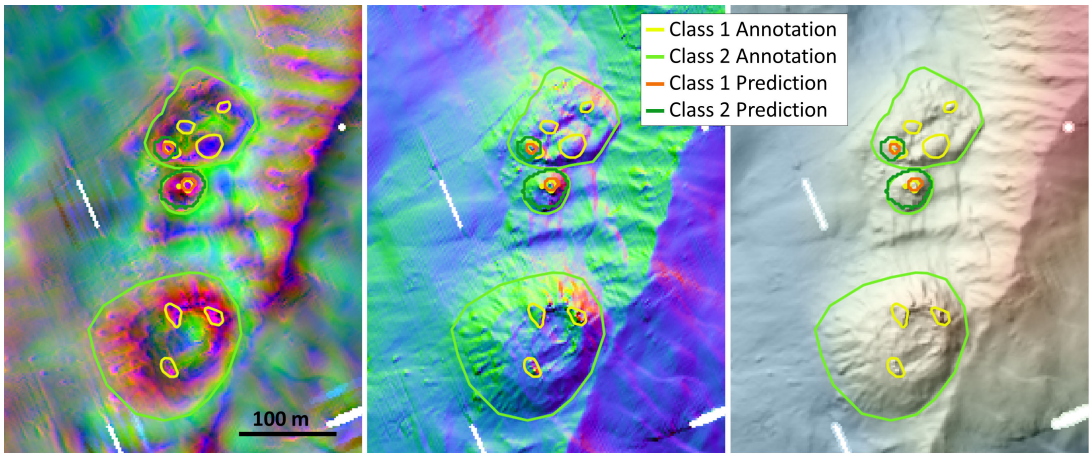


Figure 6.8: Annotated (bright green and yellow marks) and predicted (dark green and orange marks) mounds of an hydrothermal site with more complex structure.

6.4 Differentiation of Sulphide and Volcanic Mounds

Once the mounds have been successfully detected in the bathymetric data, the next step is to extract of interesting exploration targets. To do this, sulphide mounds must be distinguished from volcanic domes. Three possible data sets (and their derivatives) are available for this task: MBES bathymetry, MBES backscatter, and CNN mound prediction results.

This chapter presents the results of two different approaches to such differentiation. In section 6.4.1, parameters of backscatter texture and bathymetry derivatives are analysed for their feasibility in differentiating mound origins. In addition, a clustering approach based on the visual representation of the CNN results is derived (6.4.2). Finally, both results are applied to the mound predictions derived from the self-trained CNN model to evaluate the success of exploration target reduction (6.4.3).

6.4.1 Mound Differentiation Based on Backscatter Texture and Bathymetric Parameters

Previously it was shown that backscatter intensities or its texture derivatives do not provide a visual or numerical unambiguous identification of areas of recent or current hydrothermal activity (Chapter 5). Data gaps and especially areas of low data quality (e.g., artefacts, no calibration) in the mosaics affect the quantitative analysis when investigating seabed characteristics of larger continuous areas. U-Net provides a prediction of mound locations within the study area. This allows to restrict an data analyses to be restricted those small features of limited spatial extent and high probability of representing a mound.

The first step is to identify suitable data types for differentiating between mounds. Furthermore, the data values and thresholds that identify sulphide mounds need to be determined. Chapter 5.3 describes that sulphide mounds generally have higher entropy (randomness) and lower IDM (homogeneity) values. Figure 6.9 shows the values of these two parameters for all manually annotated mound peaks (class 1). The known sulphide mounds are marked with black stars.

The data distribution in Figure 6.9 shows a correlation between the two parameters, as the data points are arranged in a straight line: high entropy values coincide with low IDM and low entropy with high IDM. Most of sulphide mound peaks in data sets A, B, and C have an entropy value of at least 1 and an IDM value not exceeding 0.5.

These observed data characteristics are different for data set D. The general correlation and therefore linear trend between entropy and IDM is not as evident in the graph. The values for the sulphide mounds are also more generally

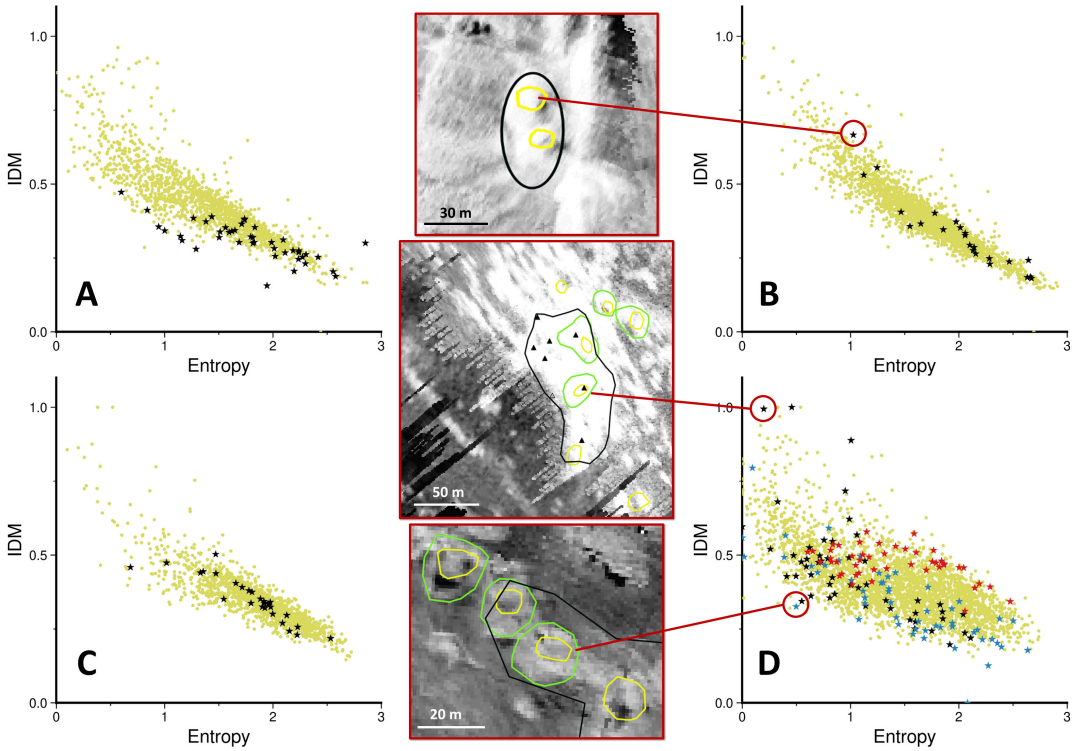


Figure 6.9: Entropy and IDM for all annotated mound peaks (class 1) for each data set. Known sulphide mounds are marked with black stars. The three different hydrothermal fields in data set D are coloured differently [black: Field 5, blue: Field 6, red: Field 7].

distributed. The sulphide mound peaks are additionally coloured by field in Figure 6.9 to visualise that all three fields differ from the previous observation. This discrepancy is most likely due to the location of the fields on a fault scarp, which provides an acoustically challenging environment with strong slopes, thus reducing the quality of the MBES data and hence the numerical values.

Figure 6.9 centre also shows some examples of backscatter data from sulphide mound peaks outside the general values. The inverse relation of low entropy and high IDM seems to occur when an area is acoustically over- or under-exposed. Such insufficient backscatter data can be caused by changes in signal settings or general unfavourable local acquisition geometry. As a result, neighbouring pixel values are too similar and the texture properties cannot be correctly derived. The top two data examples in Figure 6.9 show such a case where the backscattering is very high, resulting in lower entropy and higher IDM as expected.

The Lower data example in Figure 6.9 shows a sulphide mound peak which has a low IDM – as expected – but an overall lower entropy than most of the peaks. The texture has low homogeneity and low randomness. In this case the statistical deviation occurs because the annotated peak is less random in texture. This example also illustrates the influence of annotation on the numerical results. If the peak boundary are moved slightly, the texture values would be different. As the peak delimitation is not as explicit as the base, it is

important to use large amounts of data to obtain statistically useful conclusions.

To examine other data characteristics common to sulphide mounds, the area covered by the peaks is set in relation to the area covered by the bases of the annotated mounds. This is plotted together with the mean slope (in degrees) for each of the four data sets separately in Figure 6.10. From the graphs it can be seen that most of the peaks do not cover more than 30% of the base area. Data set B differs from the other three in this respect, as its mean is closer to 30%, whereas the others are closer to 10% - 20%. The value of the known sulphide occurrences is within the average and does not show any specific characteristics.

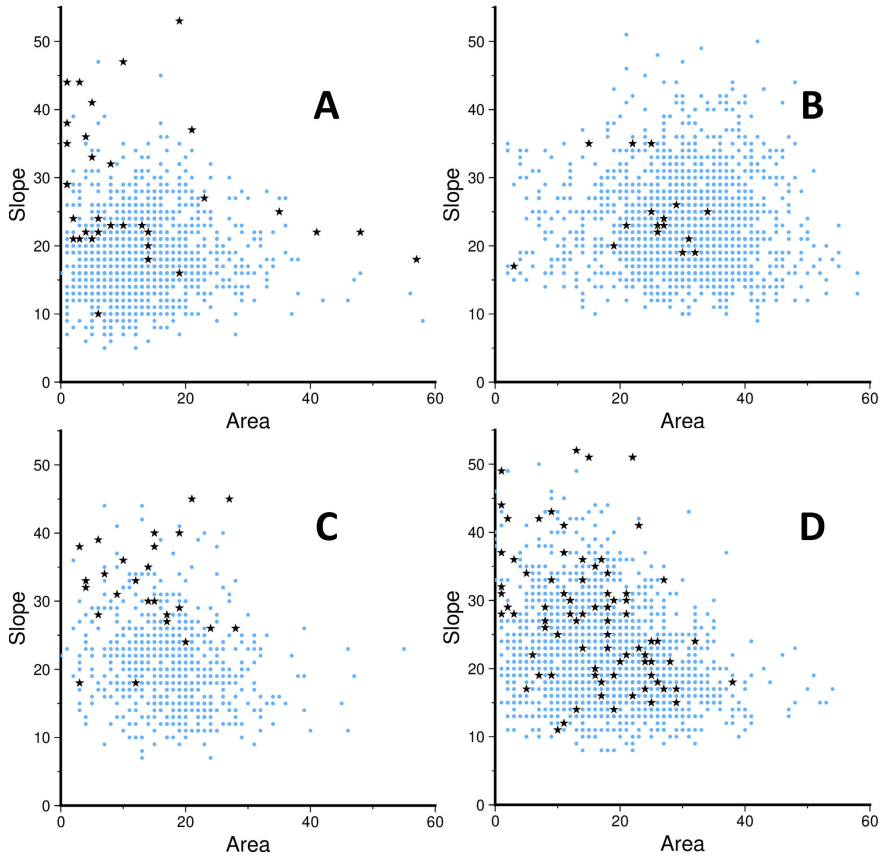


Figure 6.10: Peak coverage to base (in percent) and mean slope (in degrees) of peak covered area derived from annotated mounds. Known sulphide mounds are marked bys black stars.

With regard to the mean peak slope value, it can be seen from Figure 6.10 that the mean value for sulphide sites is generally 20°or higher and therefore in the upper part of the general data. Data set D is again an exemption, showing different data characteristics to the other data sets.

The mean slope value for sulphide mounds is not very different from other mounds. However, by using it for general differentiation, some of the mounds can be further excluded as exploration targets. On the other hand, the areal ratio of peak to base is not found to be a useful criterion and is not included.

To summarize the findings of this section, it can be concluded that sulphide mound peaks typically have the following characteristics based on the backscatter texture and bathymetric characteristics:

- High entropy (randomness) of at least 1.
- Low IDM (homogeneity) of 0.5 or less.
- High mean slope value of at least 20°.

These conclusions can be applied to the CNN predicted mounds to reduce the number of exploration targets to some extent. To further improve this estimation of sulphide mounds, a second approach based on clustering of the visual representation of the CNN prediction results is included. This approach is presented in the following section.

6.4.2 Mound Differentiation Based on Clustering of CNN Predictions

The U-Net mound prediction is shown as a raster, with possible mound locations coloured in red and background data in blue. In the previous investigations of this study, areas of high mound probability were converted to vector format for comparison to the ground truth data. The approach described in this section, uses the raster data directly.

In a first step, the circular features representing the mound predictions are extracted. Then, for each of the mound prediction, the GLCM is computed and the texture metrics such as contrast, dissimilarity, IDM, and energy are extracted (see 5.3 for a description). Furthermore, the circularity of the extracted features is determined using edge detection. The features are then clustered into four groups by using the k-means algorithm (Lloyd, 1982). The latter is a commonly used algorithm in machine learning to group data into a defined number of clusters. The data in each cluster has certain similarities. It is an unsupervised process and is iterative. The whole process was done in Python using the scikit-image library for image processing. The Python script was created with the support of ChatGPT (OpenAI, 2024).

Figure 6.11 shows some examples of extracted mound predictions for each of the cluster. It can be seen that the shape and colour gradient of the individual mound prediction influence the cluster in which it is grouped. For example, cluster 3 for class 1 seems to cluster red and more circular features compared to the other clusters. Or cluster 1 for class 2 contains mound predictions with a less circular shape, including double mounds.

The cluster into which the sulphide mounds are grouped are of particular interest, as this would allow conclusions to be drawn about their textural and circular properties. The likelihood that other sulphide mounds are grouped in

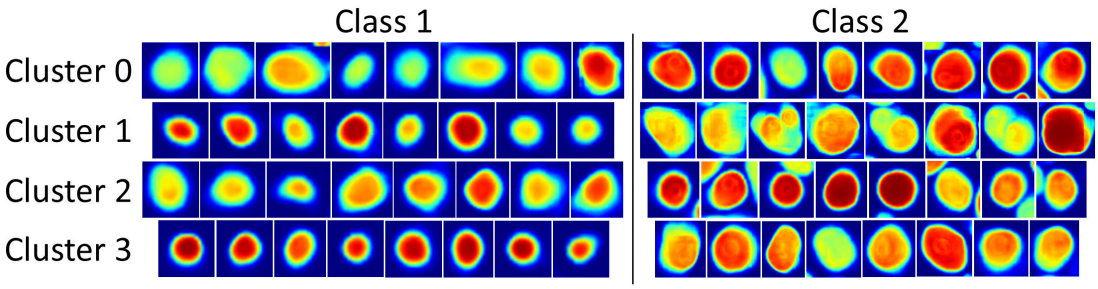


Figure 6.11: Examples of predicted mounds sorted by class and cluster.

the same clusters is therefore higher. Table 6.5 lists the number of predicted mounds per cluster as well as the known sulphide mounds. As the shape of the peak is the main morphological feature distinguishing sulphide from volcanic mounds, one would expect the clustering to group sulphide mound peaks more unambiguous than the base of the mound. However, it appears to be the other way round.

A clear definition of a preferred cluster for sulphide mound peaks (class 1) cannot be made. Most of them, 66%, are assigned to clusters 1 and 2. These two group mound predictions that are intermediate circular and moderate coloured. The majority of known sulphide mound bases (class 2), 49%, are assigned to cluster 3. This group combines intermediate circular features with medium and irregular colouring.

Table 6.5: Clustering results of predicted mounds and cluster assignment of sulphide mound peaks (class 1) and bases (class 2).

		Class 1				**Class 2**			
Cluster		**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Total Mounds		1,664	985	1,200	495	1,076	2,081	336	1,944
Sulphide Mounds	A	2	6	5	3	3	2	4	6
	B	3	1	1	0	1	2	0	1
	C	2	4	0	1	1	2	1	3
	D	6	8	10	1	4	0	3	12
Total Sulphide Mounds:		13	**19**	**16**	5	9	6	8	**22**

Figure 6.12 shows the class 1 and class 2 predictions of some of the known sulphide mounds with assigned clusters. It shows the diversity of the prediction representation. In this regard, it is investigated if certain cluster combinations for class 1 and 2 for sulphide mounds are more likely to occur. The most common cluster combinations are cluster 1 and 3, and 2 and 3 (for class 1 and class 2), which corresponds to the above class-wise assignment of sulphide mounds to the clusters.

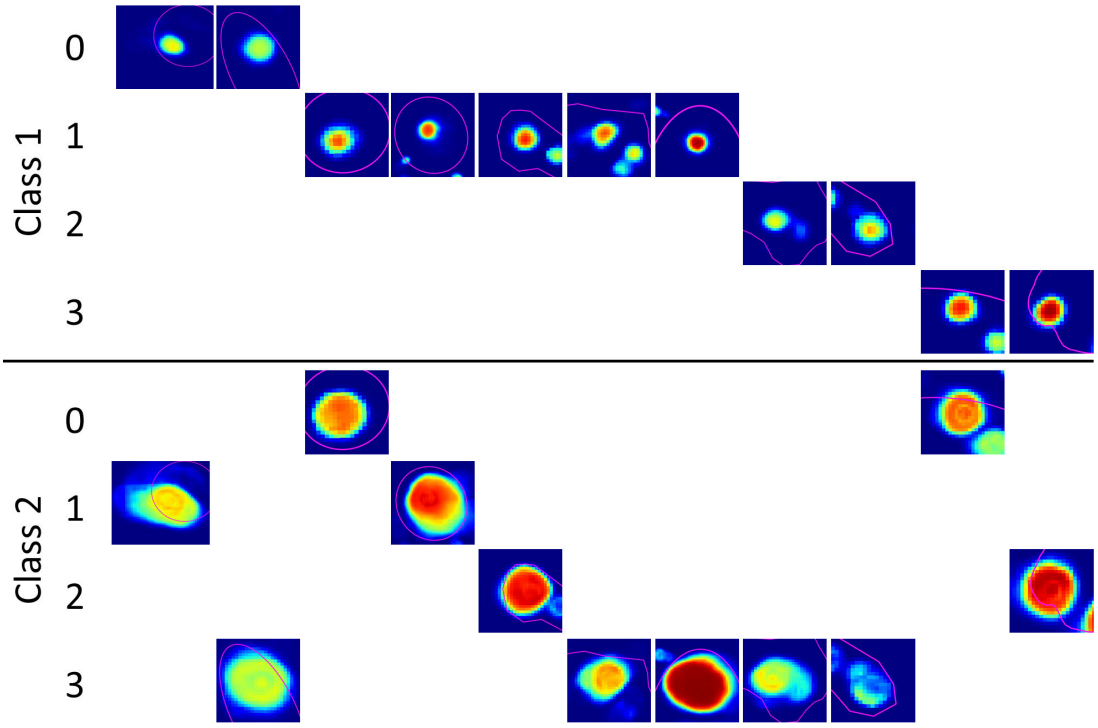


Figure 6.12: Examples of predicted sulphide mounds sorted by class and cluster.

6.4.3 Combination of Both Approaches to Reduce Exploration Targets

The previous two sections have established various parameters that can contribute to reduce the exploration targets by excluding mounds with a high probability of being of volcanic origin. The remaining mounds are more likely to be associated with sulphide deposits. Areas with accumulations of mounds that are more likely to be of hydrothermal origin are of particular interest and can be specifically considered for exploration planning. Such areas provide a suitable initial point for further data evaluation and analysis by geologists and locations for the deployment of further geophysical instruments with limited coverage.

Table 6.6 shows the number of predicted mound peaks (class 1) using different reduction methods: [1] without any reduction, [2] with reduction using the threshold values of mean slope ($\geq 20^{\circ}$), entropy (≥ 1), and IDM (≤ 0.5), [3] the same as in [2] plus consideration of only those assigned to cluster 1 and 2 (class 1), and [4] with the additional requirement that the corresponding base must be assigned to class 2.

From Table 6.6 it can be seen that in three of the four data sets the number of remaining targets is reduced to less than half of the predicted mounds using only the bathymetric and backscatter based reduction criteria. Some of the known sulphide mounds are also removed as possible targets. However, their relative proportion increases as their reduction is less pronounced.

Table 6.6: Reduction of exploration targets by applying different data criteria to the predicted mounds.

Data set		No Target Reduction (Class 1)	Mean Slope & Texture	+ Clustering Class 1 (1 or 2)	+ Clustering Class 2 (3)
A	Mounds:	2,305	927 (40%)	361 (16%)	148 (6%)
	Sulph. Md.:	22	21 (95%)	16 (72%)	9 (41%)
B	Mounds:	1,459	596 (41%)	280 (19%)	95 (7%)
	Sulph. Md.:	6	4 (67%)	1 (17%)	1 (17%)
C	Mounds:	845	494 (58%)	212 (25%)	70 (8%)
	Sulph. Md.:	9	8 (89%)	4 (44%)	2 (22%)
D	Mounds:	1,832	753 (41%)	451 (25%)	258 (14%)
	Sulph. Md.:	45	25 (56%)	19 (42%)	8 (18%)
Sum	Mounds:	6,441	2,770 (43%)	1,204 (20%)	571 (9%)
	Sulph. Md.:	82	58 (71%)	40 (49%)	20 (24%)

When the clustering requirement for the peak is added, the remaining mounds are reduced to about 21% of the originally predicted mounds. The known sulphide mounds are reduced to only 43% and now account for about 2% of the total. Adding the clustering criteria for class 2 reduces the total number of mounds to 10% of which about 3% are known sulphide mounds. Such a large reduction is not advisable as drastically limits the exploration targets. This last approach will therefore not be pursued further.

All the reduction criteria and thresholds developed are not strict parameters. As shown in the previous two sections, it cannot be fully expected that all sulphide mounds will fall with in the defined parameter range. Volcanic and sulphide mounds are natural features that vary in shape, size, and type and thus vary greatly in data representation. Therefore, data reduction may not strictly distinguish between mounds of different origin, but may indicate areas of increased exploration interest. Consideration should be given to using the various reduction criteria in combination for exploration planning, depending on the general area of interest and its characteristics.

Figure 6.13 shows a section of *Field 3* (top) and *Field 1* (bottom) with known hydrothermal sites (black outline) and possible exploration targets based on the criteria slope and texture (left, in orange) and additionally with the criteria of being assigned to peak cluster 1 or 2 (right, in red). The larger number of possible exploration targets in the left images is clearly noticeable. It can also be seen that most of the visually recognisable sulphide peaks are maintained and proposed as exploration targets.

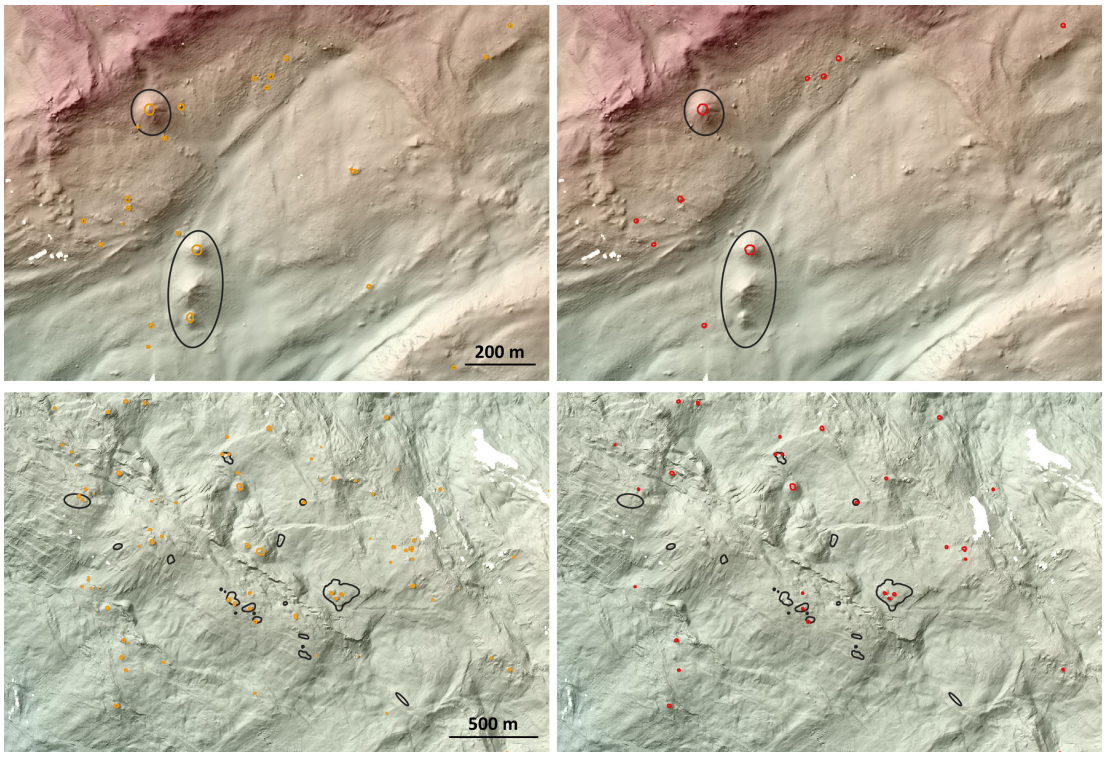


Figure 6.13: Results of the exploration target reduction in *Field 3* (top) and *Field 1* (bottom) using different criteria. Left: Mean slope and texture (entropy and IDM). Right: Mean slope, texture (entropy and IDM) and class 1 being assigned to cluster 1 or 2.

6.5 Chapter Summary - Deep Learning Assisted Sulphide Mound Detection

The study shows that deep learning (in this case CNN U-Net) can be successfully applied to deep-sea mound detection using high-resolution MBES data. The training and especially the annotation of the training data is crucial as it directly influences the trained model. The trained model accomplishes to correctly identify about 55% of the annotated mounds. Mound bases have a higher success rate (65%) than mound peaks (45%). The rate of correctly detected sulphide mounds is about 42% on average and therefore lower overall.

The subsequent reduction of the AI-detected mounds aims to exclude volcanic mounds so that the overall proportion of sulphide mounds in the remaining possible exploration targets increases. An evaluation of possible parameters identifies mean slope, entropy and IDM of backscatter texture, and certain clustered groups as the most suitable for target reduction. The parameters considered may be adjusted depending on the individual area and data. They result in a reduction of exploration targets to about 45% (without considering clustering) and 21% (with considering class 1 clustering) of the original AI-based detected mounds.

Chapter 7

Discussion

Figure 7.1 shows the workflow developed in this study for an automated prediction of sulphide mound locations. The overall process consists of four main parts: MBES data processing, data preparation and implementation into the CNN for mound detection, the processing and clustering of the CNN results, and, finally, the selection of exploration targets.

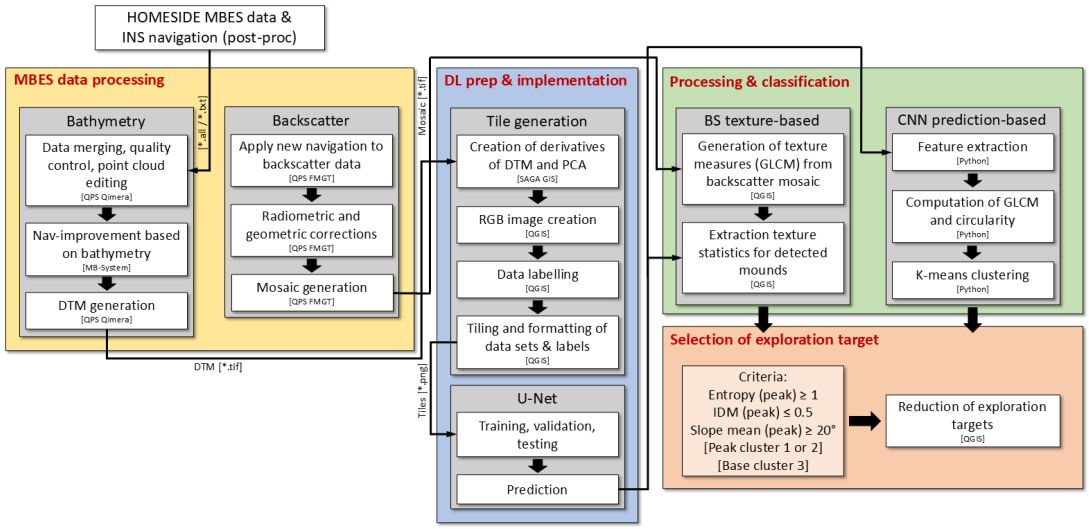


Figure 7.1: Workflow for an automated prediction of sulphide mounds as exploration targets.

The innovation of this approach lies in its exclusive use of high-resolution MBES bathymetry and backscatter data and that it is not only limited to active but also includes inactive sites. It is discussed in detail in the following sections, in the order of the three research objectives and associated research questions outlined in Chapter 1.3.

7.1 RO1: Improve the quality of the MBES data so that it can be used for automated detection and classification.

RQ1.1: How can the high-resolution MBES bathymetry data be processed to minimise the effects of insufficient navigation accuracy?

MB-System and *mbnavadjust* prove to be powerful tools for adjusting remaining **navigation offsets** in MBES data based on bathymetric features in overlapping data. The implemented ability of varying offsets within a line using a least squares adjustment, the possibility to assign offsets in xyz simultaneously, and the partially automated processes for determining the local offset are the major advantages of the software. Other studies using similar data (high-resolution MBES of underwater systems) confirm the importance of such bathymetry-based approach (Ferrini et al., 2007; Klischies, 2021).

The disadvantage of this approach is the time-consuming assignment of the individual offsets – as there where more than 24,000 tie points manually set for data set A (Chapter 4.1.2). This issue needs to be addressed further to make the approach more generally applicable. An automation of tie point assignment should be pursued in the future.

At the time of the workflow improvement, no other software with similar capabilities could be found and utilised. Recently, also commercial software has started to implement more flexible tools for navigation adjustment based on bathymetry data. However, these are often still limited in terms of a combined adjustment in xyz or the consideration of different offsets. However, with the increased use of unmanned and underwater sensor platforms, the demand for such tools with a high degree of automation is increasing, and so is the likelihood of progress in this area.

The introduction of *MB-System* into the HOMESIDE MBES processing workflow allows the mismatches in the overlaps of adjacent MBES swathes to be significantly reduced. Combined with manual editing of the overlaps, the resulting DTMs and their derivatives have been improved for automated detection and classification of small mounds, as artefacts are reduced and therefore small bathymetric structures are less superimposed.

As the final underwater navigation estimate is highly dependent on the absolute positioning of the USBL, care must be taken with its calibration and the applied sound speed information used. Due to time constraints during cruises, the SSP and calibration values may not be updated as often as ideally preferable. In such cases, an adequate overlap between adjacent swaths or even cross lines is even more important to detect systematic offsets. This should already be taken into account when planning the acquisition profiles. Due to

the nature of towed instruments and the unpredictability of bottom currents, HOMESIDE cannot always be towed exactly along the planned line. Offsets of approximately 50 m and more are possible. Adjustments in the field are slow. Planning with larger swath overlaps by reducing the profile spacing by at least 50 m will allow better data error analysis and subsequent data adjustment.

RQ1.2: How does different processing software and varying parameter settings affect the MBES backscatter?

For the **backscatter processing**, the evaluation and comparisons show that the possibilities for user interference in this overall complex process are limited. The selected instrument settings, and in particular their continuous maintenance during the acquisition, are crucial for obtaining good results, as the processing possibilities are limited. Ideally, the backscatter data should be calibrated to obtain homogeneous and absolute backscatter values. However, in projects such as INDEX, the complexity of such a calibration is very high, as the MBES is located on a submerged platform in a morphologically rough deep-sea environment.

The analysis shows that an AVG correction is the most important setting in backscatter processing where one can influence and improve the results – when working with INDEX data. An AVG correction, which is not limited by the data file boundaries, significantly improves the quality of the backscatter data in such harsh seabed environments as MORs.

RQ1.3: Which processing workflow is best suited to process the high-resolution MBES backscatter data from the INDEX project area?

The best workflow for the INDEX data was developed as part of this study. However, as with navigation adjustment, backscatter processing software and the tools implemented in it are constantly changing and improving. The developed workflows are therefore only valid for a limited period of time. Available updates need to be evaluated for project relevance and should be incorporated into the workflow if beneficial. For example, since the initial comparison for this work, current QPS FMGT versions now include an AVG correction that takes into account previous and subsequent data files. Recent efforts to address the large gaps in backscatter standardisation and quality control by the GeoHab group and others are also showing very promising ideas and initiatives, such as the open source software *Iskaffee* (Schimel et al., 2024).

In addition, new data, possibly collected by different instruments, is constantly being added to the existing data compilations. The characteristics and processing requirements of the new data need to be considered and implemented. However, the study shows that a close examination and careful selection of the processing tools, adjustments, and corrections, as well as their settings, can

improve the processing results and thus the quality of the final MBES data product.

7.2 RO2: Analyse the acoustic characteristics of the seabed classes in the study area.

RQ2.1: How can the seabed classes of the study area be differentiated in high-resolution MBES backscatter data?

The backscatter analysis uses geological maps as ground truth information. These have been created using video imagery from ROVs and video sleds. The resulting maps have the limitation that their coverage does not coincide with the one of MBES, but is much smaller. They are also subject to generalisation. Seabed types can occur mixed and often also alternate within a small area. In order to understand the influence of the generalisation in geological mapping on the backscatter analysis, a separate map was created with a more detailed differentiation of seabed types, including also mixed classes. When this map is used in the backscatter analysis, the area coverage of each classes is too small to be statistically significant. Therefore, a certain degree of generalisation is necessary in geological maps. However, this aspect and its influence must be considered when numerically analysing backscatter data used as ground truth data.

The study shows that an areal numerical backscatter analysis of the INDEX data suffers from two major problems: first, the backscatter intensities are not normalized or calibrated and therefore only relative ranges can be considered. Secondly, although the quality of the mosaic can be significantly improved by post-processing, the remaining data artefacts and shadowing effects on steep morphological structures make statistical and quantitative analyses difficult. This strongly affects the ability to differentiate seabed classes and in particular to identify hydrothermal mounds in the backscatter data.

Despite these problems, the numerical analysis of relative backscatter values shows that the major seabed classes, such as sediment and talus/pillow, which have very different acoustic backscatter responses, can be differentiated from each other to some extent. A more detailed differentiation of talus and pillow lavas or hydrothermal influenced areas is not possible. This is especially true for the use of the backscatter mosaic and the immediate backscatter intensity values for areal analyses.

Texture analysis using GLCM texture metrics is also shown to be more negatively affected by artefacts, shadowing effects, and data gaps than when using backscatter intensities directly. In particular, data gaps due to lack of overlap between profiles or shadowing are exacerbated by the use of larger window

sizes. It cannot be used for areal analysis, confirming other studies such as Sperle et al. (2015).

In summary, a comprehensive and complete numerical seabed classification for large areas is not possible with the current INDEX data. This applies not only for an analysis based on the relative backscatter values, but also for textural derivatives – especially if no additional data sources or ground truthing are used.

RQ2.2: Are there certain backscatter characteristics that are unique to areas of hydrothermal activity?

First of all, visible hydrothermal activity on the seabed cannot be summarized in ONE seabed class, but can appear as sulphide blocks, sulphide talus, hydrothermal sediment, or hydrothermal crust. All these types have different acoustic scattering characteristics, which are often expected to be similar or even equal to non-hydrothermal classes (e.g., hydrothermal sediment and sediment).

Due to their similarity in acoustic backscatter to non-hydrothermal classes, no unique signature can be identified. In certain parts, hydrothermal areas can be relatively distinguished from the non-hydrothermal environment when both types have a very different backscatter response.

The study shows that certain textural properties (high entropy, low IDM) are often observed at the known hydrothermal mounds. These originate from shadowing effects caused by the steep slope of the sulphide mound peak. This characteristic cannot be utilised over the entire data set as it is not a unique characteristic. Due to the rough terrain, shadowing can be observed regularly. However, when combined with other information such as possible mound locations, it can become useful in distinguishing between sulphide mounds and volcanic mounds. This is discussed further in the second research question of the following section.

7.3 RO3: Automate the process of predicting sulphide side locations based on high-resolution MBES data.

RQ3.1: How successful is a mound prediction using U-Net for the data of the INDEX project?

Before this can be answered, another question needs to be addressed: **How can the performance of the U-Net model be evaluated and when is an instance correctly segmented?** The CNN result is a raster representation of the test data with a numerical indication of possible mound locations. In addition, some statistical values from training, validation, and testing are provided,

which indicate how well the model results match with the labelled data.

The overall success of training a CNN model is provided by the parameters loss, accuracy, recall, and precision parameters given for each epoch. In this study, these parameters are all within the desired range. Even the initial values indicate already good model performance. The improvement of the model over the epochs is relatively small taking into account the computational effort (about 50 min of computation time for each epoch). The U-Net architecture is designed to train effectively with a smaller amount of training data. In addition, data augmentation is performed. Considering the performance improvement of the model over the epochs, it might be possible to reduce the training time by reducing the number of epochs if needed.

The automatically generated evaluation metrics are based on a pixel-by-pixel comparison of the model segmentation and the annotated data. However, this approach may not be suited for all studies – as in this case for the mound detection. In this study, it is not important whether the boundaries of a mound are correctly segmented, but whether the mound is detected within a certain radius. This radius is in the order of tens of metres, as this is the distance at which an ROV deployed for site confirmation can locate a mound without too much effort. A pixel-by-pixel evaluation may distort the results in this study case.

The raster with the mound prediction is therefore converted into polygons in vector format, each representing an area with a high probability of being covered by a mound. These polygons are checked for their intersection with instances of the ground truth data (annotations). This conversion is done on the basis of colour, which raises the question of where the boundary should be defined in the obtained continuously coloured segmentation results that do not provide a clear boundary. In this study, the areas coloured red and yellow are extracted as predicted mound locations, as this is consistent with human intuitive visual interpretation. The colour parameters were carefully chosen during the development of the workflow, as they have a strong influence on the statistical results, since correctness is achieved through overlap and is therefore to some extent size dependent.

When comparing the segmentation success, the statistics of the bases are better than those of the peaks. This indicates that they are better recognised by the model. This is probably due to the fact that the boundary of the base can be more clearly distinguished from the surrounding seafloor as the elevation changes abruptly. The peaks, on the other hand, are more difficult to distinguish, as the change in elevation is not as obvious.

The instance-wise comparison of the automated segmentation of the model trained with INDEX data shows that only about 55% of the instances are cor-

rectly detected. The success rate for detecting sulphide mounds is even lower at about 42%. This means that about half of the known sulphide mounds are excluded from further investigations. This is not satisfying as the aim is to get closer to the ground truth data with higher segmentation success rates. This leads to the next intermediate question: **Why is the success rate of mound detection relatively low?**

The generally low success rate of mound detection may be due to the imbalance between mounds and background areas. As the area covered by mounds is significantly smaller than the area not covered by mounds, the model may tend to make pessimistic predictions. An attempt has already been made to counteract this by removing tiles from the training and validation that have no or very little mound coverage. However, the imbalance could not be completely eliminated. Further data augmentation to artificially increase the mound ratio has not been tested but may help to improve this issue.

An additional reason for the low success rate could be the inadequacy of the input data. It is possible that further or alternative data derivatives such as contour lines may improve the results. Also in the approach used by Haroon et al. (2023), which is also based on the method of Juliani and Juliani (2021), other data input layers are used, as it gives better results. An evaluation of the individual importance of the data layers – also with the help of algorithms like *Boruta* – may be considered (Janowski et al., 2018).

As the number of correctly detected sulphide mounds is generally lower than the general number, this indicates that the shapes are likely to be more different from the volcanic domes, which are much more common in the data. As sulphide mounds are overall under-represented, the CNN may have difficulty learning to recognise these features.

Comparing the results of using the self-trained model and the pre-trained one leads to the next question: **Why does the pre-trained model recognise more mounds than when using the self-trained model?**

The visual comparison of the predictions resulting from the two models leads to the conclusion that the mound labelling for the ground truthing of the mound locations is done differently. On the one hand, the predicted mound sizes are much larger for the pre-trained model. This erroneously improves the success rate, as larger mound predictions are more likely to overlap with the ground truth data, thereby increasing the number of correctly identified instances. On the other hand, the predictions of the pre-trained model are more optimistic and more areas are wrongly considered as mound locations (e.g., high value of false positives, low precision). These two circumstances improve the overall mound detection rate, but the overall performance of the model is not necessarily better as it is equally important to reduce the number of incorrectly

detected mounds. Otherwise the number of possible exploration targets will be falsely increased. The higher DC value of self-trained model indicates that the overall performance, taking into account recall and precision, is better. This, in turn, supports the point made above about the importance of examining the metrics for CNN evaluation with greater care.

This presumed bias caused by different ways of annotating data by various operators should be avoided. Where multiple operators are involved in data labelling, standards, guidelines, and training should be in place. The consistency and agreement of annotations from different sources should be measured regularly and adjusted if necessary.

The difference between the two models may also be due to the different characteristics of the study area. This leads to the next intermediate question: **Can the mound detection models be applied to other study areas?** The study area of Juliani and Juliani (2021), the Mohns Ridge, is an ultraslow-spreading ridge with about 15 to 16 mm per year (Mosar et al., 2002). The data in this study is located at the slow-spreading Central Indian Ridge (data set A) and the intermediate-spreading Southeast Indian Ridge (data sets B, C, D) (Klischies, 2021).

This difference in spreading rate may affect the applicability of the pre-trained model for this study, as it can be assumed that the general environment may be slightly different. However, The differences are not expected to be significant compared to fast- or superfast-spreading ridges, where the environment is less dominated by pillow lavas and general large-scale tectonic features (as in the study area), but where sheet and ponded lava flows are present (Fornari and Embley, 2010). Seafloor massive sulphides are less common in such regions.

Overall, the success rate of all the hydrothermal fields studied varies considerably – for both models tested. The fields are spread over a distance of approximately 500 km, and their local geological settings differ from each other. The success rate of the mound detection also varies: *Field 1* shows the best mound detection rate (57%) and *Field 5* the worst (35%).

The differences between the fields can be seen in the ratio of the number of peaks to the number of bases. Particularly in the more southerly fields along the intermediate spreading axis, the number of isolated peaks (diameter less than 30 m) is higher. This becomes also evident in the ratio of annotated peaks to bases, which varies greatly. These isolated peaks have also been annotated as in certain fields they occur at sites of hydrothermal activity.

To summarise the above listed discussion and answer the research question: U-Net can be successfully applied to the INDEX data for mound detection, but certain aspects need to be improved. These include, first and foremost,

increasing the success rate of mound prediction and in particular sulphide mound prediction. Secondly, the operator bias should be reduced. Thirdly, the integration of data from other areas and different geological settings can improve the general usability of the CNN model.

RQ3.2: Can MBES backscatter data facilitate in distinguishing sulphide from volcanic mounds and thereby enable a reduction of possible exploration targets?

The issue of defining mound boundaries in CNN prediction has already been mentioned in the context of the metrics for CNN model evaluation. This aspects is also important for the step of excluding volcanic mounds in order to ideally extract mounds that are more likely to be of hydrothermal origin. In this part, the numerical values for these mound boundaries are extracted, based on which the mounds are differentiated. It is necessary to ensure that the polygons representing mounds cover the desired area and thus the statistical values representing certain desired characteristics.

Another aspect to consider is the use of clustering to differentiate mound types. Ideally, when adding a smaller piece of newly collected data – such as by the annual cruises – it should be easy to apply the workflow directly to the new data. The integration of clustering, as done in the developed approach, expects that all data to be clustered to identify the cluster containing the most sulphide mounds. This is not practical and an approach that takes into account the parameters influencing the clustering (e.g., GLCM texture and circularity) is generally sought.

The study shows that backscatter data and its textural derivatives are not sufficient to exclude a significant proportion of the detected mounds as exploration targets. This confirms previous investigations in this study that backscatter data can only be used to a very limited extent to distinguish hydrothermal from non-hydrothermal areas. The shape of the mounds – and therefore the bathymetry – is in most cases a better discriminator.

Therefore, both bathymetric data and direct CNN predictions are used as mound differentiation criteria to improve the target reduction. The overall result of combining all these different approaches tends in the right direction, but is not yet fully satisfactory.

Chapter 8

Conclusion

The study demonstrates that AI tools can be successfully used to automatically retrieve exploration targets that indicate locations with a high probability being of (former) hydrothermal activity and thus potentially associated with sulphide deposits. It is a proof of the concept that such a reduction in exploration targets is possible based on MBES data alone.

Using this developed approach on the study data, the exploration targets (i.e., all segmented mounds) can be reduced to 43% (using backscatter texture and slope) or 20% (by additionally considering clustering). This reduction increases the proportion of known sulphide mounds from an original 1.2% to 2.1% (using backscatter texture and slope) and 3.1% (by additionally considering clustering) respectively. This is an increase by a factor of up to three and is estimated to by also applicable to unknown sulphide mounds. The overall reduction in potential areas of interest not only saves time and thereby money, but more importantly increases the exploration success. However, the workflow has its limitations and certain aspects of it need to be improved in order to use it successfully and reliably in practice.

First, the approach does not distinguish between active and inactive **hydrothermal sites** per se, but is strictly limited to sites with distinct mound structures that are not buried by talus or sediment. The typical sulphide mound shape is expected to diminish over time as a site becomes inactive due to weathering processes. Some hydrothermal sites may not develop a mound at all (Jamieson et al., 2014). The approach works best when a mound with a steep and conical shaped peak or a chimney is located on a flat and even surrounding seafloor. This limitation is caused by the decision to use only MBES data, and is of course expected, but needs to be taken into account when applying it, or when considering exploration methods for hydrothermal sulphide deposits in a broader context.

Secondly, the success rate of mound detection using **deep learning** needs

to be improved as about 45% of the manually labelled mounds are not recognized. Possible starting points are the way of annotation or the consideration of other CNN such as BHP-U-Net, which is said to perform better in fuzzy boundary recognition (Tang et al., 2023). Furthermore, additional or other input data derivatives may be considered as also suggested in the study of Haroon et al. (2023) (e.g., aspect).

Thirdly, the whole process is based on **MBES data** only. This single data source, combined with the fact of low quality backscatter data (i.e., frequent artefacts and gaps), means that it alone in this state is not sufficient to be used to differentiate between volcanic and sulphide mounds. Improvement of the backscatter quality by optimised acquisition settings (i.e., avoidance of automatic system adjustments concerning the signal or gain), processing corrections, or calibration. In particular, more effort should be made during acquisition to avoid data gaps, as texture or other algorithms based on neighbourhood analysis or a moving window cannot be applied as successfully when gaps are present. In addition, direct integration of the backscatter data into the CNN may be considered, as deep learning may be able to detect certain patterns that are not visible to the human eye. Such a change in the method could improve the usability of backscatter data.

It is also strongly recommended that other data sources are added in the mound differentiation process. An integration of further data sources has shown to maximise the machine learning results (Brandes et al., 2021). Suitable data sources may include magnetic data for example as it has also been used by Haroon et al. (2023) for the same purpose (sulphide mound detection). In addition, geological maps of structural features that influence the formation of hydrothermal systems (e.g., faults) may be added.

Coming back to the main question: Is the research gap closed? Can the developed approach fill the gap in the field of hydrothermal site exploration (active and inactive) using MBES data and techniques from the field of deep learning techniques?

Yes, the study shows that such an approach can reduce the number of exploration targets for sulphide mounds. But in order to guide the exploration reliably and sufficiently, the workflow needs to be further improved. The study also shows that backscatter data can be used to some extent for reducing the detected mounds, but that the backscatter data quality is crucial. It also it illustrates the importance of data (and model) sharing within the research community to achieve more universal solutions, especially when working in such highly specialised research fields with limited data availability.

The approach developed provides exploration targets. Although further refinement is required, clusters of probable target locations identify areas for fur-

ther manual geological inspection. Such regions may also be considered for the use of other geophysical instruments such as receivers for the three-axial electric field. The workflow can therefore help to reduce the overall exploration area, but at this stage cannot completely replace the human interaction in this context.

Bibliography

Alevizos, E., Schoening, T., Koeser, K., Snellen, M., and Greinert, J. (2018): Quantification of the fine-scale distribution of Mn-nodules: Insights from AUV multibeam and optical imagery data fusion. *Biogeosciences Discussions*, pp. 1–29, DOI: 10.5194/bg-2018-60.

Arosio, R., Hobley, B., Wheeler, A. J., Sacchetti, F., Conti, L. A., Furey, T., and Lim, A. (2023): Fully convolutional neural networks applied to large-scale marine morphology mapping. *Frontiers in Marine Science*, 10, DOI: 10.3389/fmars.2023.1228867.

Asada, M., Yoshikawa, S., Mochizuki, N., Nogi, Y., and Okino, K. (2015): Brief report of side-scan sonar imagery observations of the archaean, pika, and urashima hydrothermal sites. In: Maiani, L., Abousahl, S., and Plastino, W., (Eds), *International Cooperation for Enhancing Nuclear Safety, Security, Safeguards and Non-proliferation-60 Years of IAEA and EURATOM: Proceedings of the XX Edoardo Amaldi Conference, Accademia Nazionale dei Lincei, Rome, Italy, October 9-10, 2017*, pp. 479–487. Springer, Tokyo, ISBN: 978-4-431-54864-5, DOI: 10.1007/978-4-431-54865-2_37.

Augustin, J.-M. (2023): *SonarScope software*. SEANOE, DOI: 10.17882/87777, https://www.seanoe.org/data/00766/87777/.

Bahrenberg, G., Giese, E., Mevenkamp, N., and Nipper, J. (2008): *Statistische Methoden in der Geographie, Band 2: Multivariate Statistik*. Gebrüder Borntraeger Verlagsbuchhandlung, 3rd Edition, ISBN: 978-3-443-07144-8.

Baker, E. T. (2017): Exploring the ocean for hydrothermal venting: New techniques, new discoveries, new insights. *Ore Geology Reviews*, 86:55–69, ISSN: 01691368, DOI: 10.1016/j.oregeorev.2017.02.006.

Baker, E. T., Resing, J. A., Haymon, R. M., Tunnicliffe, V., Lavelle, J. W., Martinez, F., Ferrini, V., Walker, S. L., and Nakamura, K. (2016): How many vent fields? New estimates of vent field populations on ocean ridges from precise mapping of hydrothermal discharge locations. *Earth and Planetary Science Letters*, 449:186–196, DOI: 10.1016/j.epsl.2016.05.031.

Barnefske, E. and Sternberg, H. (2022): Evaluation of class distribution and class combinations on semantic segmentation of 3D point clouds with PointNet. *IEEE Access*, 11:3826–3845, DOI: 10.1109/ACCESS.2022.3233411.

Bartsch, C. (2014): *Structural and magnetic investigations of two spreading systems around the Rodriguez Triple Junction with respect to hydrothermal activities*. PhD thesis, Gottfried Wilhelm Leibniz Universität Hannover, DOI: 10.15488/8221.

Beaudoin, J. D., Hughes Clarke, J. E., and Bartlett, J. E. (2004): Application of surface sound speed measurements in post-processing for multisector multibeam echosounders. *The International Hydrographic Review*, ISSN: 00206946, https://journals.lib.unb.ca/index.php/ihr/article/view/20675. Accessed 03/2021.

Beaulieu, S. E., Baker, E. T., and German, C. R. (2015): Where are the undiscovered hydrothermal vents on oceanic spreading ridges? *Deep Sea Research Part II: Topical Studies in Oceanography*, 121:202–212, DOI: 10.1016/j.dsr2.2015.05.001.

Beaulieu, S. E. and Szafranski, K. M. (2020): *InterRidge Global Database of active submarine hydrothermal vent fields version 3.4*. PANGAEA - Data Publisher for Earth & Environmental Science, DOI: 10.1594/PANGAEA.917894. Accessed 06/2021.

Berlyand, L. and Jabin, P.-E. (2023): *Mathematics of deep learning: An introduction*. Walter de Gruyter GmbH, Berlin and Boston, DOI: 10.1515/9783111025551.

Blondel, P. (2003): Seabed classification at ocean margins. In: Wefer, G., Billett, D., Hebbeln, D., Jørgensen, B. B., Schlüter, M., and van Weering, T. C. E., (Eds), *Ocean Margin Systems*, pp. 125–141. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-07872-9, DOI: 10.1007/978-3-662-05127-6_8.

Blondel, P. and Gómez Sichi, O. (2009): Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf. *Applied Acoustics*, 70(10):1288–1297, ISSN: 0003682X, DOI: 10.1016/j.apacoust.2008.07.015.

Blondel, P., Parson, L. M., and Robigou, V. (1998): TexAn: Textural analysis of sidescan sonar imagery and generic seafloor characterisation. In: *IEEE Oceanic Engineering Society. OCEANS'98. Conference Proceedings (Cat. No.98CH36259)*, pp. 419–423. IEEE, DOI: 10.1109/OCEANS.1998.725780.

Boschen, R. E., Rowden, A. A., Clark, M. R., and Gardner, J. (2013): Mining of deep-sea seafloor massive sulfides: A review of the deposits,

their benthic communities, impacts from mining, regulatory frameworks and management strategies. *Ocean & Coastal Management*, 84:54–67, DOI: 10.1016/j.ocecoaman.2013.07.005.

Brandes, T. S., Ballard, B., Ramakrishnan, S., Lockhart, E., Marchand, B., and Rabenold, P. (2021): Environmentally adaptive automated recognition of underwater mines with synthetic aperture sonar imagery. *The Journal of the Acoustical Society of America*, 150(2):851, DOI: 10.1121/10.0005811.

Brownlee, J. (2019): How to use learning curves to diagnose machine learning model performance. *Machine Learning Mastery*, https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/. Accessed 11/2024.

Buck, W. R., Lavier, L. L., and Poliakov, A. N. B. (2005): Modes of faulting at mid-ocean ridges. *Nature*, 434(7034):719–723, DOI: 10.1038/nature03358.

Bundesanstalt für Geowissenschaften und Rohstoffe (07/2021): BGR Homepage. https://www.bgr.bund.de. Accessed 07/2021.

Bundesanstalt für Geowissenschaften und Rohstoffe (2015): *INDEX2015 Cruise Report*.

Bundesanstalt für Geowissenschaften und Rohstoffe (2018): *INDEX2018 Cruise Report*.

Bundesanstalt für Geowissenschaften und Rohstoffe (2019): *INDEX2019 Cruise Report*.

Calder, B. R. and Wells, D. E. (2007): CUBE user's manual. *Centre for Coastal and Ocean Mapping 1217*, https://scholars.unh.edu/ccom/1217/?utm_source=scholars.unh.edu%2Fccom%2F1217&utm_medium=PDF&utm_campaign=PDFCoverPages. Accessed 11/2023.

Che Hasan, R., Ierodiaconou, D., Laurenson, L., and Schimel, A. (2014): Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. *PloS ONE*, 9(5):e97339, DOI: 10.1371/journal.pone.0097339.

Christ, R. D. and Wernli, R. L. (2014): *The ROV manual: A user guide for remotely operated vehicles*. Butterworth-Heinemann, 2nd Edition, ISBN: 978-0-7506-8148-3, DOI: 10.1016/C2011-0-07796-7.

Clausi, D. A. (2002): An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing*, 28(1):45–62, ISSN: 0703-8992, DOI: 10.5589/m02-004.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J. (2015): System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, 8(7):1991–2007, DOI: 10.5194/gmd-8-1991-2015.

Constable, S., Kannberg, P. K., and Weitemeyer, K. (2016): Vulcan: A deep–towed CSEM receiver. *Geochemistry, Geophysics, Geosystems*, 17(3):1042–1064, ISSN: 15252027, DOI: 10.1002/2015GC006174.

Corliss, J. B., Dymond, J., Gordon, L. I., Edmond, J. M., von Herzen, R. P., Ballard, R. D., Green, K., Williams, D., Bainbridge, A., Crane, K., and van Andel, T. H. (1979): Submarine thermal springs on the Galapagos Rift. *Science (New York)*, 203(4385):1073–1083, DOI: 10.1126/science.203.4385.1073.

CSSF (2021): *ROPOS*. https://www.ropos.com. Accessed 04/2021.

de Albuquerque, V. H. C., Raj, P., and Yadav, S. P. (2024): *Toward artificial general intelligence: Deep learning, neural networks, generative AI*. De Gruyter, Berlin, DOI: 10.1515/9783111323749.

de Jong, C. D. (2006): *Hydrography*. Series on mathematical geodesy and positioning. DUP Blue Print, Delft, 1st Edition.

Diesing, M., Green, S. L., Stephens, D., Lark, R. M., Stewart, H. A., and Dove, D. (2014): Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches. *Continental Shelf Research*, 84:107–119, ISSN: 02784343, DOI: 10.1016/j.csr.2014.05.004.

Dufek, T. (2012): *Backscatter analysis of multibeam sonar data in the area of the Valdivia Fracture Zone using Geocoder in CARIS HIPS&SIPS and IVS3D Fledermaus*. Master thesis, HafenCity University Hamburg, https://epic.awi.de/id/eprint/31617/. Accessed 10/2024.

Durand, S., Legendre, P., and Juniper, S. K. (2006): Sonar backscatter differentiation of dominant macrohabitat types in a hydrothermal vent field. *Ecological Applications*, 16(4):1421–1435, ISSN: 1939-5582, https://esajournals.onlinelibrary.wiley.com/doi/full/10.1890/1051-0761%282006%29016%5B1421%3ASBDODM%5D2.0.CO%3B2. Accessed 04/2024.

Eason, D. E., Dunn, R. A., Pablo Canales, J., and Sohn, R. A. (2016): Segment-scale variations in seafloor volcanic and tectonic processes from multibeam sonar imaging, Mid-Atlantic Ridge Rainbow region (35°45'-36°35'N). *Geochemistry, Geophysics, Geosystems*, 17(9):3560–3579, DOI: 10.1002/2016GC006433.

Eleftherakis, D. (2013): *Classifying sediments on Dutch riverbeds using multi-beam echo-sounder systems*. PhD thesis, Delft University of Technology, DOI: 10.4233/uuid:f12389ee-9ae9-4c7e-bb99-a27da0449f6f.

Erickson, J. (2003): *Marine Geology: Exploring the new frontiers of the ocean*. The living Earth. Facts On File, New York, rev. Edition.

Evans, J. (2018): *Geomorphometry & gradient metrics*. https://github.com/jeffreyevans/GradientMetrics/releases/tag/2.0.0. Accessed 04/2024.

Exail (2015): *DELPH INS V.3.0 Subsea Edition - User Guide*.

Exail (2016): *Phins III - User manual - Inertial products - Principle & conventions*.

Exail (2018): *PHINS 6000 datasheet*. https://rts.as/wp-content/uploads/2018/09/iXBlue-PHINS-6000-Subsea.pdf. Accessed 10/2023.

Exail (2023): *Posidonia datasheet*. https://www.ixblue.com/wp-content/uploads/2022/01/Datasheet_Posidonia.pdf. Accessed 10/2023.

Fakiris, E., Rzhanov, Y., and Zoura, D. (2012): On importance of acoustic backscatter corrections for texture-based seafloor characterization. *Proceedings of the 11th European Conference on Underwater Acoustics*, (819).

Feldens, P., Darr, A., Feldens, A., and Tauber, F. (2019): Detection of boulders in side scan sonar mosaics by a neural network. *Geosciences*, 9(4):159, DOI: 10.3390/geosciences9040159.

Feldens, P., Westfeld, P., Valerius, J., Feldens, A., and Papenmeier, S. (2021): Automatic detection of boulders by neural networks - a comparison of multibeam echo sounder and side-scan sonar performance. *Hydrographische Nachrichten*, (119):6–17, DOI: 10.23784/HN119-01.

Ferrini, V. L., Fornari, D. J., Shank, T. M., Kinsey, J. C., Tivey, M. A., Soule, S. A., Carbotte, S. M., Whitcomb, L. L., Yoerger, D., and Howland, J. (2007): Submeter bathymetric mapping of volcanic and hydrothermal features on the east pacific rise crest at 9°50'N. *Geochemistry, Geophysics, Geosystems*, 8(1), ISSN: 15252027, DOI: 10.1029/2006GC001333.

Fonseca, L. E. and Calder, B. R. (2005): Geocoder: An efficient backscatter map constructor. https://scholars.unh.edu/cgi/viewcontent.cgi?article=1338&context=ccom. Accessed 05/2023.

Fornari, D. J. and Embley, R. W. (2010): Tectonic and volcanic controls on hydrothermal processes at the mid-ocean ridge: an overview based on near-bottom and submersible studies. In: Humphris, S. E., Zierenberg, R. A., Mullineaux, L. S., Thomson, R. E., and Humphris, S., (Eds),

Seafloor hydrothermal systems, Geophysical monograph 0065-8448, pp. 1–46. American Geophysical Union, Washington, DC, ISBN: 9781118663998, DOI: 10.1029/GM091p0001.

GEBCO Compilation Group (2019): GEBCO 2019 Grid. DOI: 10.5285/836f016a-33be-6ddc-e053-6c86abc0788e.

Gehrmann, R. A. S., Haroon, A., Morton, M., Djanni, A. T., and Minshull, T. A. (2019): Seafloor massive sulphide exploration using deep-towed controlled source electromagnetics: Navigational uncertainties. *Geophysical Journal International*, 220(2):1215–1227, ISSN: 0956-540X, DOI: 10.1093/gji/ggz513.

German, C., Lin, J., and Parson, L., (Eds) (2004): *Mid-Ocean Ridges*. Geophysical Monograph Series. American Geophysical Union, Washington, D. C., ISBN: 9781118665879.

German, C. R., Petersen, S., and Hannington, M. D. (2016): Hydrothermal exploration of mid-ocean ridges: Where might the largest sulfide deposits be forming? *Chemical Geology*, 420:114–126, DOI: 10.1016/j.chemgeo.2015.11.006.

German, C. R., Yoerger, D. R., Jakuba, M., Shank, T. M., Langmuir, C. H., and Nakamura, K. (2008): Hydrothermal exploration with the autonomous benthic explorer. *Deep Sea Research Part I: Oceanographic Research Papers*, 55(2):203–219, ISSN: 09670637, DOI: 10.1016/j.dsr.2007.11.004.

Graber, S., Petersen, S., Yeo, I., Szitkar, F., Klischies, M., Jamieson, J., Hannington, M., Rothenbeck, M., Wenzlaff, E., Augustin, N., and Stobbs, I. (2020): Structural control, evolution, and accumulation rates of massive sulfides in the TAG hydrothermal field. *Geochemistry, Geophysics, Geosystems*, 21(9), ISSN: 15252027, DOI: 10.1029/2020GC009185.

Gràcia, E., Charlou, J. L., Radford-Knoery, J., and Parson, L. M. (2000): Non-transform offsets along the Mid-Atlantic Ridge south of the Azores (38°N–34°N): Ultramafic exposures and hosting of hydrothermal vents. *Earth and Planetary Science Letters*, 177(1):89–103, ISSN: 0012821X, DOI: 10.1016/S0012-821X(00)00034-0.

GRASS Development Team (2024): *Geographic Resources Analysis Support System (GRASS GIS) Reference Manual, Version 8.4.1dev*. Open Source Geospatial Foundation, USA, https://grass.osgeo.org/grass84/manuals/index.html. Accessed 09/2024.

GRASS Development Team, Landa, M., Neteler, M., Metz, M., Petrášová, A., Petráš, V., Clements, G., Zigo, T., Larsson, N., Kladivová, L., Haedrich, C., Blumentrath, S., Andreo, V., Cho, H., Gebbert, S., Nartišs, M., Kudrnovsky,

H., Delucchi, L., Zambelli, P., Lennert, M., Mitášová, H., Chemin, Y., Pešek, O., Barton, M., Tawalika, C., Ovsienko, D., and Bowman, H. (2024): GRASS GIS. DOI: 10.5281/ZENODO.5176030.

Grotzinger, J. and Jordan, T. (2020): *Understanding Earth*. Macmillan International Higher Education. Macmillan International, New York, 8th Edition.

Groves, P. D. (2013): *Principles of GNSS, inertial, and multisensor integrated navigation systems*. Artech House, Incorporated, Norwood, 2nd Edition, DOI: 10.1017/S0373463313000672.

Grządziel, A. and Wąż, M. (2018): The invention and developing of multibeam echosounder technology. *Polish Hyperbaric Research*, 62(1):33–41, DOI: 10.2478/phr-2018-0002.

Hall-Beyer, M. (2017a): GLCM texture tutorial. http://www.fp.ucalgary.ca/mhallbey/tutorial.htm. Accessed 10/2024.

Hall-Beyer, M. (2017b): Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *International Journal of Remote Sensing*, 38(5):1312–1338, ISSN: 0143-1161, DOI: 10.1080/01431161.2016.1278314.

Hannington, M. D., de Ronde, C. E. J., and Petersen, S. (2005): Sea-floor tectonics and submarine hydrothermal systems. In: Hedenquist, J. W., Thompson, J. F. H., Goldfarb, R. J., and Richards, J. P., (Eds), *Economic Geology*. Society of Economic Geologists, Littleton, Colorado, DOI: 10.5382/AV100.06.

Hannington, M. D., Jamieson, J., Monecke, T., Petersen, S., and Beaulieu, S. (2011): The abundance of seafloor massive sulfide deposits. *Geology*, 39(12):1155–1158, DOI: 10.1130/G32468.1.

Hannington, M. D., Jamieson, J., and Petersen, S. (2015): Seafloor massive sulfide deposits: Continuing efforts toward a global estimate of seafloor massive sulfides. In: *Oceans 2015 - Genova*, pp. 1–3, Piscataway, NJ. IEEE, ISBN: 978-1-4799-8736-8, DOI: 10.1109/OCEANS-Genova.2015.7271526.

Hannington, M. D., Jonasson, I. R., Herzig, P. M., and Petersen, S. (1995): Physical and chemical processes of seafloor mineralization at mid-ocean ridges. In: *Seafloor Hydrothermal Systems: Physical, Chemical, Biological, and Geological Interactions*, pp. 115–157. American Geophysical Union (AGU), ISBN: 9781118663998, DOI: 10.1029/GM091p0115.

Haralick, R. M. (1979): Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804, ISSN: 0018-9219, DOI: 10.1109/PROC.1979.11328.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973): Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, ISSN: 0018-9472, DOI: 10.1109/TSMC.1973.4309314.

Haroon, A., Paasche, H., Graber, S., Petersen, S., Attias, E., Jegen, M., Gehrmann, R., Hölz, S., and Klischies, M. (2023): Automated seafloor massive sulfide detection through integrated image segmentation and geophysical data analysis: Revisiting the TAG hydrothermal field. *Geochemistry, Geophysics, Geosystems*, 24(12), DOI: 10.1029/2023GC011250.

Hinz, M., Westfeld, P., Feldens, P., Feldens, A., Themann, S., and Papenmeier, S. (2024): AI-based boulder detection in sonar data – bridging the gap from experimentation to application. *The International Hydrographic Review*, 30(1):78–98, ISSN: 00206946, DOI: 10.58440/ihr-30-1-a08.

Hughes Clarke, J. E. (2012): Optimal use of multibeam technology in the study of shelf morphodynamics. In: Li, M. Z., Sherwood, C. R., and Hill, P. R., (Eds), *Sediments, Morphology and Sedimentary Processes*, pp. 1–28. DOI: 10.1002/9781118311172.ch1.

Huvenne, V. A. I., Hühnerbach, V., Blondel, P., Gomez-Sichi, O., and Le Bas, T. P. (2007): Detailed mapping of shallow-water environments using image texture analysis on sidescan sonar and multibeam backscatter imagery. *Proceedings of the 2nd underwater acoustic measurements conference*, https://people.bath.ac.uk/pyspb/research/selected_publications/Huvenne_etal_UAM2007.pdf. Accessed 05/2024.

IHO (2011): *Manual on Hydrography: Publication M-13*. Monaco, 1st Edition, http://iho.int/iho_pubs/IHO_Download.htm. Accessed 06/2019.

International Seabed Authority (07/2021): ISA Homepage. https://www.isa.org/. Accessed 07/2021.

Jamieson, J. W., Clague, D. A., and Hannington, M. D. (2014): Hydrothermal sulfide accumulation along the Endeavour Segment, Juan de Fuca Ridge. *Earth and Planetary Science Letters*, 395:136–148, DOI: 10.1016/j.epsl.2014.03.035.

Jamieson, J. W. and Gartman, A. (2020): Defining active, inactive, and extinct seafloor massive sulfide deposits. *Marine Policy*, 117:103926, DOI: 10.1016/j.marpol.2020.103926.

Janowski, L., Trzcinska, K., Tegowski, J., Kruss, A., Rucinska-Zjadacz, M., and Pocwiardowski, P. (2018): Nearshore benthic habitat mapping based on multi-frequency, multibeam echosounder data using a combined

object-based approach: A case study from the Rowy Site in the southern Baltic Sea. *Remote Sensing*, 10(12):1983, ISSN: 2072-4292, DOI: 10.3390/rs10121983.

Johnson, C., Affolter, M. D., Inkenbrandt, P., and Mosher, C. (2019): *An Introduction to Geology*. https://geo.libretexts.org/@go/page/6838. Accessed 08/2021.

Juliani, C. and Juliani, E. (2021): Deep learning of terrain morphology and pattern discovery via network-based representational similarity analysis for deep-sea mineral exploration. *Ore Geology Reviews*, 129:103936, DOI: 10.1016/j.oregeorev.2020.103936.

Kawada, Y. and Kasaya, T. (2017): Marine self-potential survey for exploring seafloor hydrothermal ore deposits. *Scientific Reports*, 7(1), DOI: 10.1038/s41598-017-13920-0.

Keohane, I. and White, S. (2022): Chimney identification tool for automated detection of hydrothermal chimneys from high-resolution bathymetry using machine learning. *Geosciences*, 12(4):176, DOI: 10.3390/geosciences12040176.

Klischies, M. (2021): *Structural controls on hydrothermal systems on slow- and intermediate-spreading mid-ocean ridges*. https://macau.uni-kiel.de/receive/macau_mods_00002683. Accessed 08/2023.

Kongsberg (2016): *Datasheet HiPAP*. https://www.kongsberg.com/globalassets/kongsberg-discovery/navigation--positioning/hipap/documents/400578f-hipap-pd.pdf. Accessed 10/2023.

Kongsberg (2017): *EM 2040 product specification*. https://www.km.kongsberg.com/ks/web/nokbg0397.nsf/AllWeb/248996D7F1021D46C12575E500285652/$file/332644_em2040_product_specification.pdf. Accessed 02/2019.

Koschinsky, A., Heinrich, L., Boehnke, K., Cohrs, J. C., Markus, T., Shani, M., Singh, P., Smith Stegen, K., and Werner, W. (2018): Deep-sea mining: Interdisciplinary research on potential environmental, legal, economic, and societal implications. *Integrated environmental assessment and management*, 14(6):672–691, DOI: 10.1002/ieam.4071.

Lamarche, G., Lurton, X., Verdier, A.-L., and Augustin, J.-M. (2011): Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter – application to Cook Strait, New Zealand. *Continental Shelf Research*, 31(2):pp. 93–109, ISSN: 02784343, DOI: 10.1016/j.csr.2010.06.001.

Lecours, V., Dolan, M. F. J., Micallef, A., and Lucieer, V. L. (2016): A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Sciences*, 20(8):3207–3244, ISSN: 1027-5606, DOI: 10.5194/hess-20-3207-2016.

Li, S., Xiong, L., Tang, G., and Strobl, J. (2020): Deep learning-based approach for landform classification from integrated data sources of digital elevation model and imagery. *Geomorphology*, 354:107045, ISSN: 0169-555X, DOI: 10.1016/j.geomorph.2020.107045.

Lloyd, S. (1982): Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, ISSN: 0018-9448, DOI: 10.1109/TIT.1982.1056489.

Lurton, X. (2010): *An Introduction to Underwater Acoustics - Principles and Applications*. Springer and Praxis Publishing, 2nd Edition.

Lurton, X. and Lamarche, G., (Eds) (2015): *Backscatter measurements by seafloor-mapping sonars: Guidelines and Recommendations*. DOI: 10.5281/zenodo.10089261.

Malik, M., Lurton, X., and Mayer, L. (2018): A framework to quantify uncertainties of seafloor backscatter from swath mapping echosounders. *Marine Geophysical Research*, 39(1):151–168, ISSN: 0025-3235, DOI: 10.1007/s11001-018-9346-7.

Malik, M., Schimel, A. C. G., Masetti, G., Roche, M., Le Deunf, J., Dolan, M. F., Beaudoin, J., Augustin, J.-M., Hamilton, T., and Parnum, I. (2019): Results from the first phase of the seafloor backscatter processing software inter-comparison project. *Geosciences*, 9(12):516, DOI: 10.3390/geosciences9120516.

Marsh, I. and Brown, C. (2009): Neural network classification of multibeam backscatter and bathymetry data from Stanton Bank (Area IV). *Applied Acoustics*, 70(10):1269–1276, ISSN: 0003682X, DOI: 10.1016/j.apacoust.2008.07.012.

Masetti, G., Mayer, L., and Ward, L. (2018): A bathymetry- and reflectivity-based approach for seafloor segmentation. *Geosciences*, 8(1):14, DOI: 10.3390/geosciences8010014, https://www.mdpi.com/2076-3263/8/1/14.

MathWorks (2017): What are convolutional neural networks? – introduction to deep learning. https://de.mathworks.com/videos. Accessed 07/2024.

MBARI (2017): Homepage of Monterey Bay Aquarium Research Institute. http://iho.int/iho_pubs/IHO_Download.htm. Accessed 08/2020.

McKay, J., Gerg, I., Monga, V., and Raj, R. G. (2017): What's mine is yours: Pretrained CNNs for limited training sonar ATR. *OCEANS 2017 MTS/IEEE Anchorage*, DOI: 10.48550/arXiv.1706.09858.

Mimura, K., Nakamura, K., Takao, K., Yasukawa, K., and Kato, Y. (2023): Automated detection of hydrothermal emission signatures from multibeam echo sounder images using deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16:2703–2710, ISSN: 1939-1404, DOI: 10.1109/JSTARS.2023.3247467.

Morris, H., E.L., H., Bucker, H., and Backmann, R. (1978): *Interaction of sound with the ocean bottom: a three-year summary*. Naval Ocean Systems Center, San Diego.

Mosar, J., Lewis, G., and Torsvik, T. H. (2002): North Atlantic sea-floor spreading rates: Implications for the Tertiary development of inversion structures of the Norwegian-Greenland Sea. *Journal of the Geological Society*, 159(5):503–515, ISSN: 0016-7649.

Müller, H., Schwalenberg, K., and Barckhausen, U. (2023): Magnetic and electromagnetic exploration of SMS deposits; perspectives and frontiers. In: *84th EAGE Annual Conference & Exhibition*, pp. 1–5. European Association of Geoscientists & Engineers, DOI: 10.3997/2214-4609.2023101269.

Murtagh, F. and Heck, A. (1987): *Multivariate data analysis*, volume 131. Springer, DOI: 10.1007/978-94-009-3789-5.

Murton, B. J., Lehrmann, B., Dutrieux, A. M., Martins, S., de La Iglesia, A. G., Stobbs, I. J., Barriga, F. J., Bialas, J., Dannowski, A., Vardy, M. E., North, L. J., Yeo, I. A., Lusty, P. A., and Petersen, S. (2019): Geological fate of seafloor massive sulphides at the TAG hydrothermal field (Mid-Atlantic Ridge). *Ore Geology Reviews*, 107:903–925, ISSN: 01691368, DOI: 10.1016/j.oregeorev.2019.03.005.

Olaya, V. (2009): Chapter 6 – Basic land-surface parameters. In: Hengl, T. and Reuter, H. I., (Eds), *Developments in Soil Science : Geomorphometry*, volume 33, pp. 141–169. Elsevier, DOI: 10.1016/S0166-2481(08)00006-8.

Ondréas, H., Scalabrin, C., Fouquet, Y., and Godfroy, A. (2018): Recent high-resolution mapping of Guaymas hydrothermal fields (Southern Trough). *Bulletin de la Société Géologique de France*, 189(1), ISSN: 0037-9409, DOI: 10.1051/bsgf/2018005.

OpenAI (2024): ChatGPT (Version October 2023). https://chat.openai.com/. Accessed 12/2024.

Paige, C. C. and Saunders, M. A. (1982): LSQR: Sparse equations and least-squares problems. *ACM Transactions on Mathematic Software*, (Vol. 8, No. 2):195–209.

Parnum, I., Siwabessy, P., and Gavrilov, A. (2004): Identification of seafloor habitats in coastal shelf waters using a multibeam echosounder. In: *Processings of ACOUSTICS 2004*, pp. 181–186. https://cmst.curtin.edu.au/wp-content/uploads/sites/4/2016/05/2004-41.pdf. Accessed 08/2024.

Passos, D. and Mishra, P. (2022): A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks. *Chemometrics and Intelligent Laboratory Systems*, 223:104520, ISSN: 0169-7439, DOI: 10.1016/j.chemolab.2022.104520, https://www.sciencedirect.com/science/article/pii/S0169743922000314. Accessed 09/2024.

Petersen, S., Krätschell, A., Augustin, N., Jamieson, J., Hein, J. R., and Hannington, M. D. (2016): News from the seabed – geological characteristics and resource potential of deep-sea mineral resources. *Marine Policy*, 70:175–187, ISSN: 0308597X, DOI: 10.1016/j.marpol.2016.03.012.

Petersen, S., Lehrmann, B., and Murton, B. J. (2018): Modern seafloor hydrothermal systems: New perspectives on ancient ore-forming processes. *Elements*, 14(5):307–312, DOI: 10.2138/gselements.14.5.307.

Peukert, A., Petersen, S., Greinert, J., and Charlot, F. (2018): Seabed mining. In: Micallef, A., Krastel, S., and Savini, A., (Eds), *Submarine Geomorphology*, Springer Geology, pp. 481–502. Springer International Publishing, Cham, ISBN: 978-3-319-57851-4, DOI: 10.1007/978-3-319-57852-1_24.

Pinet, P. R. (2003): *Invitation to Oceanography*. Oceanography Series. Jones and Bertlett Publishers, 3rd Edition, ISBN: 9780763721367.

QGIS Development Team (2024): *QGIS Geographic Information System*. QGIS Association, https://www.qgis.org. Accessed 07/2024.

Qin, X., Luo, X., Wu, Z., and Shang, J. (2021): Optimizing the sediment classification of small side-scan sonar images based on deep learning. *IEEE Access*, 9:29416–29428, DOI: 10.1109/ACCESS.2021.3052206.

QPS (2022): *Fledermaus 7.8 Documentation, FMGT 7.10 / FMMidwater 7.9*, https://publicdownload.qps.nl/Fledermaus/Fledermaus-7.8-Documentation_v3.pdf. Accessed 03/2023.

Riley, S. J., de Gloria, S. D., and Elliot, R. (1999): A terrain ruggedness that quantifies topographic hetergeneity. *Intermountain Journal of Science*, 1-4(5):23–27.

Ronneberger, O., Fischer, P., and Brox, T. (2015): U-Net: Convolutional networks for biomedical image segmentation. http://arxiv.org/pdf/1505.04597.pdf. Accessed 03/2024.

Rusby, J. and Somers, M. (1977): The development of the GLORIA sonar system from 1970 to 1975. *A Voyage of Discovery*, pp. 611–625.

Sato, M. and Mooney, H. M. (1960): The electrochemical mechanism of sulfide self-potentials. *Geophysics*, 25(1):226–249, DOI: 10.1190/1.1438689.

Schimel, A. C. G., Beaudoin, J., Parnum, I. M., Le Bas, T., Schmidt, V., Keith, G., and Ierodiaconou, D. (2018): Multibeam sonar backscatter data processing. *Marine Geophysical Research*, 39(1-2):121–137, DOI: 10.1007/s11001-018-9341-z.

Schimel, A. C. G., Dolan, M., Chand, S., Thorsnes, T., and Bjarnadottir, L. R. (2024): Iskaffee - multibeam backscatter quality control software. https://github.com/alexschimel/Iskaffe?tab=readme-ov-file. Accessed 12/2024.

Searle, R. C. (2014): *Mid-Ocean Ridges*. Cambridge University Press, Cambridge, DOI: 10.1017/CBO9781139084260.

Searle, R. C., Murton, B. J., Achenbach, K., LeBas, T., Tivey, M., Yeo, I., Cormier, M. H., Carlut, J., Ferreira, P., Mallows, C., Morris, K., Schroth, N., van Calsteren, P., and Waters, C. (2010): Structure and development of an axial volcanic ridge: Mid-Atlantic Ridge, 45°N. *Earth and Planetary Science Letters*, 299(1-2):228–241, ISSN: 0012821X, DOI: 10.1016/j.epsl.2010.09.003.

Seibold, E. and Berger, W. H. (2017): *The sea floor: An introduction to marine geology*. Springer Textbooks in Earth Sciences, Geography and Environment. Springer, 4th Edition.

Sharma, P. V. (1997): *Environmental and engineering geophysics*. Cambridge University Press, Cambridge, U.K. and New York, NY, USA, ISBN: 9780521572408, DOI: 10.1017/CBO9781139171168.

Sokolova, M. and Lapalme, G. (2009): A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427–437, ISSN: 0306-4573, DOI: 10.1016/j.ipm.2009.03.002.

Sonardyne (2014): *Datasheet Sonardyne Ranger 2 USBL*. https://www.uniquegroup.com/wp-content/uploads/2022/10/Sonardyne-Ranger-2-USBL.pdf. Accessed 10/2023.

Sørensen, T. (1948): A method of establishing groups of equal amplitude in plant sociobiology based on similarity of species content and its application to analyses of the vegetation on danish commons. In: *Biologiske Skrifter/Kongelige Danske Videnskabernes Selskab*, volume 5, pp. 1–34.

SPC (2013): *Deep Sea Minerals: Sea-floor massive sulphides, a physical, biological, environmental, and technical review*. Backer, E. and Beaudoin, Y. (Eds.) Vol 1A, Secretariat of the Pacific Community.

Speer, K. G. and Rona, P. A. (1989): A model of an Atlantic and Pacific hydrothermal plume. *Journal of Geophysical Research: Solid Earth*, 94(C5):6213–6220, ISSN: 01480227, DOI: 10.1029/JC094iC05p06213.

Sperle, M., Negri, E., and Ternes, C. (2015): Automatic classification of sidescan sonar images for mapping marine mineral resources. In: *2015 IEEE/OES Acoustics in Underwater Geosciences Symposium*, pp. 1–5, Piscataway, NJ. IEEE, ISBN: 978-1-4673-7019-6, DOI: 10.1109/RIOAcoustics.2015.7473620.

Steiniger, Y., Kraus, D., and Meisen, T. (2022): Survey on deep learning based computer vision for sonar imagery. *Engineering Applications of Artificial Intelligence*, 114:105157, ISSN: 0952-1976, DOI: 10.1016/j.engappai.2022.105157.

Sze, V., Chen, Y.-H., Yang, T.-J., and Emer, J. (2017): Efficient processing of deep neural networks: A tutorial and survey. http://arxiv.org/pdf/1703.09039. Accessed 04/2024.

Tang, Y., Wang, L., Jin, S., Zhao, J., Huang, C., and Yu, Y. (2023): AUV-based side-scan sonar real-time method for underwater-target detection. *Journal of Marine Science and Engineering*, 11(4):690, DOI: 10.3390/jmse11040690.

Tao, C., Chen, S., Baker, E. T., Li, H., Liang, J., Liao, S., Chen, Y. J., Deng, X., Zhang, G., Gu, C., and Wu, J. (2017): Hydrothermal plume mapping as a prospecting tool for seafloor sulfide deposits: a case study at the Zouyu-1 and Zouyu-2 hydrothermal fields in the southern Mid-Atlantic Ridge. *Marine Geophysical Research*, 38(1-2):3–16, ISSN: 0025-3235, DOI: 10.1007/s11001-016-9275-2.

Torres, R. N., Fraternali, P., Milani, F., and Frajberg, D. (2020): Mountain summit detection with deep learning: evaluation and comparison with heuristic methods. *Applied Geomatics*, 12(2):225–246, ISSN: 1866-9298, DOI: 10.1007/s12518-019-00295-2.

UN General Assembly (1982): *Convention on the Law of the Seas*.

Vega, P. J. S., Papadakis, P., Matabos, M., van Audenhaege, L., Ramiere, A., Sarrazin, J., and Da Costa, G. A. O. P. (2024): Convolutional neural networks for hydrothermal vents substratum classification: An introspective study. *Ecological Informatics*, 80:102535, ISSN: 1574-9541, DOI: 10.1016/j.ecoinf.2024.102535.

Walker, S. L., Baker, E. T., Resing, J. A., Nakamura, K., and McLain, P. D. (2007): A new tool for detecting hydrothermal plumes: an ORP sensor for the PMEL MAPR. *AGU Fall Meeting Abstracts*, 88.

Wang, L., Shi, C., Diao, C., Ji, W., and Yin, D. (2016): A survey of methods incorporating spatial information in image classification and spectral unmixing. *International Journal of Remote Sensing*, 37(16):3870–3910, ISSN: 0143-1161, DOI: 10.1080/01431161.2016.1204032.

Wang, X., Cao, Y., Wu, S., and Yang, C. (2023): Real-time detection of deep-sea hydrothermal plume based on machine vision and deep learning. *Frontiers in Marine Science*, 10, DOI: 10.3389/fmars.2023.1124185.

Weiss, A. D. (2000): Topographic position and landforms analysis. https://www.jennessent.com/downloads/TPI-poster-TNC_18x22.pdf. Accessed 09/2024.

Wilson, M. F. J., O'Connell, B., Brown, C., Guinan, J. C., and Grehan, A. J. (2007): Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30(1-2):3–35, DOI: 10.1080/01490410701295962.

Yu, Y., Zhao, J., Gong, Q., Huang, C., Zheng, G., and Ma, J. (2021): Real-time underwater maritime object detection in side-scan sonar images based on Transformer-YOLOv5. *Remote Sensing*, 13(18):3555, ISSN: 2072-4292, DOI: 10.3390/rs13183555.

Acknowledgements

First of all, I would like to thank my first supervisor Prof. Dr.-Ing. Harald Sternberg for his support, trust, and guidance not only during this thesis, but also during my time at HCU.

Furthermore, I would like to thank my second supervisor Prof. Dr. rer. nat. Poerbandono and Prof. Dr.-Ing. Thomas Kersten for their support of this thesis.

I would like to thank Dr. Ralf Freitag, Dr. Thomas Kuhn, and Dr. Ulrich Schwarz-Schampera for giving me the opportunity to be part of the INDEX project, for being open to my questions, for their continuous support, and for their very helpful feedback and advice. I would also like to thank the entire INDEX team and the crew members of all the cruises for making the participation in the project such a great experience.

I further like to thank Dr.-Ing. Eike Barnefske and Güren Dinga for sharing their AI expertise and providing me with very helpful feedback and ideas on how to overcome certain AI hurdles!

Thanks to all my HCU colleagues and friends who have crossed my path over the last ten years and who made it such a great time. Special thanks to Felix Tschirschwitz & Franziska Dietrich, Hannes Kröger, Mona Lütjens, Dilip Adhikari, Friedrich Keller, and Thomas Willemsen for a great working atmosphere, helpful discussions, and of course unforgettable Einstein evenings. ;o)

Thanks to Franziska Jurisch for the help with all the practical questions I constantly had during my start at HCU ... and still have ...

I would also like to thank Dr.-Ing. Annette Scheider and Dr. Meike Klischies for their constant and very helpful discussions and their feedback on this thesis. Thank you for your companionship and support!

Thanks to thank my parents, Angela and Pavel, and my sister Judith for their constant encouragement and support.

Last but not least, a big dankjewel to Arjen for his patience, the uplifting conversations, and especially his understanding throughout the last year. Without that, the thesis would not have been possible. ♡