Automated seafloor massive sulfide detection through integrated image segmentation and geophysical data analysis: Revisiting the TAG hydrothermal field

Amir Haroon¹, Hendrik Paasche², Sebastian Graber³, Sven Petersen¹, Eric Attias⁴, Marion Jegen¹, Romina Gehrmann⁵, Sebastian Hölz¹, and Meike Klischies⁶

¹GEOMAR
²Unknown
³GEOMAR Helmholtz Centre of Ocean Research Kiel
⁴Institute for Geophysics
⁵Dalhousie University
⁶north.io GmbH

September 29, 2023

Abstract

Accessible seafloor minerals located near mid-ocean ridges are noticed to mitigate projected metal demands of the net-zero energy transition, promoting growing research interest in quantifying global distributions of seafloor massive sulfides (SMS). Mineral potentials are commonly estimated using geophysical and geological data that lastly rely on additional confirmation studies using sparsely available, locally limited, seafloor imagery, grab samples, and coring data. This raises the challenge of linking in-situ confirmation data to geophysical data acquired at disparate spatial scales to obtain quantitative mineral predictions. Although multivariate datasets for marine mineral research are incessantly acquired, robust, integrative data analysis requires cumbersome workflows and experienced interpreters. Here, we introduce an automated two-step machine learning approach that integrates automated mound detection with geophysical data to merge mineral predictors into distinct classes and reassess marine mineral potentials for distinct regions. The automated workflow employs a U-Net convolutional neural network to identify mound-like structures in bathymetry data and distinguishes different mound classes through classification of mound architectures and magnetic signatures. Finally, controlled source electromagnetic data is utilized to reassess predictions of potential SMS volumes. Our study focuses on the Trans-Atlantic Geotraverse (TAG) area, which is amid the most explored SMS area worldwide and includes 15 known SMS sites. The automated workflow classifies 14 of the 15 known mounds as exploration targets of either high- or medium-priority. This reduces the exploration area to less than 7% of the original survey area from 49 km² to 3.1 km².

Hosted file

Automated seafloor massive sulfide detection through integrated image segmentation and geophysical data analysis: Revisiting the TAG hydrothermal field

Amir Haroon\textsuperscript{1,2}, Hendrik Paasche\textsuperscript{3}, Sebastian Graber\textsuperscript{1,4}, Sven Petersen\textsuperscript{1}, Eric Attias\textsuperscript{5,6}, Marion Jegen\textsuperscript{1}, Romina Gehrmann\textsuperscript{7}, Sebastian Hölz\textsuperscript{1}, Meike Klischies\textsuperscript{1,8}

\textsuperscript{1}Helmholtz Centre for Ocean Research, GEOMAR Kiel, Wischhofstr. 1-3, 24148, Kiel Germany
\textsuperscript{2}Hawaii Institute of Geophysics and Planetology, School of Ocean and Earth Science and Technology, University of Hawai`i at Manoa, Honolulu, HI, USA
\textsuperscript{3}UFZ Helmholtz Centre for Environmental Research, Leipzig, Germany
\textsuperscript{4}Helmut-Schmidt Uni Hamburg
\textsuperscript{5}Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA
\textsuperscript{6}Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA
\textsuperscript{7}Dalhousie University, 6299 South St, Halifax, NS, B3H 4R2, Canada
\textsuperscript{8}North.io GmbH, Einsteinstraße 1, 24118 Kiel, Germany

ORCID IDs
Amir Haroon: 0000-0001-5138-6730; Hendrik Paasche: 0000-0003-1814-0645; Sebastian Graber: 0000-0003-1154-6968;
Eric Attias: 0000-0003-4308-1683; Romina Gehrmann: 0000-0002-3099-2771; Sebastian Hölz: 0000-0003-2839-9394;
Marion Jegen: 0000-0001-6307-8428; Sven Petersen: 0000-0002-5469-105X; Meike Klischies: 0000-0003-4319-4573

Key Points
- Developed two-step machine learning workflow to identify mound structures in bathymetry and classify their origins based on auxiliary data.
- Substantial increase of potential SMS edifices detected within the TAG hydrothermal field distributed within latitudinal bands.
- SMS mineral potential is likely lower than perviously assumed due to heterogeneously distributed mineralization within mounds.
Abstract

Accessible seafloor minerals located near mid-ocean ridges are noticed to mitigate projected metal demands of the net-zero energy transition, promoting growing research interest in quantifying global distributions of seafloor massive sulfides (SMS). Mineral potentials are commonly estimated using geophysical and geological data that lastly rely on additional confirmation studies using sparsely available, locally limited, seafloor imagery, grab samples, and coring data. This raises the challenge of linking in-situ confirmation data to geophysical data acquired at disparate spatial scales to obtain quantitative mineral predictions. Although multivariate datasets for marine mineral research are incessantly acquired, robust, integrative data analysis requires cumbersome workflows and experienced interpreters. Here, we introduce an automated two-step machine learning approach that integrates automated mound detection with geophysical data to merge mineral predictors into distinct classes and reassess marine mineral potentials for distinct regions. The automated workflow employs a U-Net convolutional neural network to identify mound-like structures in bathymetry data and distinguishes different mound classes through classification of mound architectures and magnetic signatures. Finally, controlled source electromagnetic data is utilized to reassess predictions of potential SMS volumes.

Our study focuses on the Trans-Atlantic Geotraverse (TAG) area, which is amid the most explored SMS area worldwide and includes 15 known SMS sites. The automated workflow classifies 14 of the 15 known mounds as exploration targets of either high- or medium-priority. This reduces the exploration area to less than 7% of the original survey area from 49 km$^2$ to 3.1 km$^2$.

Keywords: Convolution Neural Networks, Seafloor Massive Sulfides, Bathymetry, Magnetic Anomaly, CSEM

Data Availability Statement:

The bathymetry data used for training the U-Net model are available on open-access repositories as listed in Tab. A1. The AUV-bathymetry and magnetic anomaly maps are available from Peterson (2019) (https://doi.pangaea.de/10.1594/PANGAEA.899415) and CSEM data from Gehrmann (2019) (https://doi.pangaea.de/10.1594/PANGAEA.899073).
Introduction

Over 700 active, inactive and extinct hydrothermal venting sites (cf. definitions in Jamieson & Gartman, 2020) are known to exist along mid-ocean ridges, volcanic arcs, or back-arc spreading centers (Beaulieu et al. 2015; Beaulieu & Szafranski, 2020). Their existence is documented through hydrothermal plumes that are visually confirmed using a suite of underwater-vehicles (e.g., Murton et al., 2019), towed-camera systems (Beaulieu et al., 2013), or via in-situ probing such as gravity coring (Petersen et al., 2016) or seafloor drilling (e.g., Murton et al. 2019). Understanding the distributions of hydrothermal venting, often associated with the evolution of seafloor massive sulfides (SMS), remains a prevalent research topic motivated by the increased demand of strategic minerals needed to foster the net-zero energy transition. The economic and environmental challenges of modern society interfaces with various fields of marine research to predict where subsurface processes transport mineral-enriched, high-temperature fluids from the deep lithosphere towards the seafloor. In many cases these processes are associated with mineral accumulations that form distinct mound-like expressions (e.g., Fouquet et al. 2010) often referred to as SMS mounds. Current estimates suggest that the abundance of known SMS can contribute a fractional supply of strategic metals in the future (Hannington et al., 2011), albeit the known uncertainties with respect to size, distribution, volume, as well as the environmental impact that would succeed mining these potential deep-sea resources.

Tonnage estimates of marine minerals are interpreted from seafloor morphology (e.g. Fouquet et al. 2010; Graber et al. 2020), geophysical analyses (Haroon et al. 2018; Gehrmann et al. 2019; Murton et al. 2019; Galley et al. 2021), or global extrapolations from deposit occurrences along spreading centers (Hannington et al. 2011). Many studies have focused on previously known SMS sites and their immediate surroundings (e.g. Jamieson et al. 2014; Murton et al. 2019; Graber et al. 2020), likely leading to an underestimation of mapped SMS edifices within a given region. For example, Jamieson et al. (2014) discovered over 400 undocumented SMS edifices through manual analysis of high-resolution bathymetry data along the Endeavour Segment. These findings highlight the prevalent question of where to sample next and in which spatial resolution? A question often guided by experience, funding, technological constraints and ship-time availability. Moreover, marine scientists recognize the challenges of differentiating prospective SMS mounds from morphologically comparable volcanic constructions (Jamieson et al. 2014) based on bathymetry data alone. In such cases, multivariate databases that include geophysical, geochemical, and geological data acquired across disparate spatial scales can help to 1) identify regions of interest for more detailed in-situ sampling or visual confirmation studies, 2) optimize the use of ship-time through targeted, pre-informed surveying and 3)
improve volumetric estimates of SMS candidates via integrative geophysical analyses. However, to process, analyze and interpret this mass of information, integrative data workflows are pivotal for extracting valuable information optimally, improving decision-making tools for localized sampling, and providing more rigorous estimates of known SMS provinces. Here, machine learning (ML) offers avenues to develop coherent data workflows and processing chains sufficiently generic and thereby transferable across various geological domains and data layers. We demonstrate an ML application together with conventional geophysical analyses to automatically detect mound-like morphology on the seafloor and identify distinct high-interest areas based on auxiliary geological and geophysical information to characterize and quantify mineral-enriched SMS targets.

Geoscientific studies utilizing ML are progressively increasing in environmental and exploration research (e.g., Bouwer et al. 2022; Koedel et al. 2022). ML applications differ not only in the type and spatial and temporal resolution of the input data, but also in the applied techniques. In marine mineral research, ML applications have applied random forest classification (e.g. Gazis et al., 2018) or neural networks (e.g. Keohane & White, 2022; Juliani & Juliani, 2021), focusing mainly on one or two types of input layers (e.g. seafloor images and/or bathymetry data). However, as interdisciplinary SMS databases grow via contributions from geological, geophysical (e.g. Müller, Schwalenberg, Barckhausen, 2023), biological and geochemical applications, ML workflows are expected to also evolve into more generic implementations to facilitate this growing demand of interdisciplinary marine research.

Our study leverages existing ML applications previously conducted in the marine environment. We introduce a workflow that integrates concurrent data acquired at different spatial scales to better describe the mineral potential within the Trans-Atlantic Geo-traverse (TAG) hydrothermal field. First, a modified approach adapted from Juliani & Juliani (2021) is utilized to identify mound structures in bathymetry data using a U-Net convolutional neural network. Subsequently, the identified mounds are amalgamated with multivariate geophysical and geological databases to assess potential SMS edifices in greater detail. We test the developed workflow using data described in Petersen et al. (2016) and Murton et al. (2018) during two Blue Mining expeditions (https://bluemining.eu/) at the TAG hydrothermal field, and compare the results to manual classification studies of Graber et al. (2020) and Murton et al. (2019). This study extends previously published concepts of marine mineral research, which in most cases use a uni-variate database, by including more diversified, multivariate input layers.
such as reduced-to-the-pole (RTP) magnetics, controlled source and transient electromagnetics (CSEM and TEM), core and grab samples all acquired across various spatial scales.

2 Geological and Geophysical Data
The data used in this study were previously published in scientific literature, i.e. Petersen et al. (2016); Murton et al. (2018); Sztikar et al. (2019); Haroon et al. (2018); Gehrmann et al., (2019); Graber et al. (2020); Gehrmann et al. (2020); Galley et al. (2021). The following describes relevant aspects of the data, which is needed in the context of the ML implementation. Please refer to the above-mentioned literature for more details on data acquisition and geological/geophysical interpretations. It is important to note that from a data science perspective, the available data introduce \textit{a-priori} bias, as these were acquired with the specific purpose of imaging certain physical parameters that, from a geological perspective, are associated with the evolution of SMS. Thus, unknown correlations that extend beyond the current geological understanding of SMS evolution are likely neglected in the presented ML workflow.

2.1 Bathymetry Data
Seafloors host numerous focused fluid discharge sites that can appear as mounded manifestations in the seafloor topology (Olakunle et al., 2021). Linking these manifestations to either a volcanic or hydrothermal origin and deriving their potential for forming metalliferous accumulations requires auxiliary data acquired at each specific site. The high-resolution seafloor bathymetry data provides a spatial baseline of where to sample for potential SMS occurrences and also the structural framework for volumetric predictions (e.g., Jamieson et al., 2014; Graber et al., 2020).

The high-resolution bathymetry data interpreted and classified by the U-Net were collected during research cruise M127 (RV Meteor, 2016) using ship-based multibeam (Fig. 1a) and GEOMAR’s Autonomous-Underwater Vehicle (AUV) Abyss (Petersen et al., 2016). Data were acquired using a RESON Seabat 7125 multibeam echosounder, navigated at a speed of three knots using a frequency of 200 kHz. The line spacing between adjacent profiles was between 80 m and 100 m at an average altitude of 84 m relative to the seafloor, resulting in a 2 m grid resolution (Fig. 1b). The bathymetry data were processed using the software package MB Systems (https://www.mbari.org/technology/mb-system/) and georeferenced based on prominent seafloor features (Graber et al. (2020) and Sztikar et al. (2019)). This high-resolution bathymetry constructs the baseline of positioning morphological structures during automated segmentation.
To further optimize the U-Net model, we utilize available high-resolution AUV bathymetry data acquired at different SMS sites around the globe (e.g. Clague et al., 2015; Escartin & Petersen, 2017) and other openly-available bathymetry data. All of the applied bathymetry grids utilized to train, validate and test the U-Net are listed in Tab. A1.

Figure 1 here!

2.2 AUV-Based Magnetic Data

The magnetic properties of seafloor basalts are dictated by two alteration processes, namely deuteritic oxidation during the initial cooling phase and the superimposed regional hydrothermal alteration that occurs at younger ages (Ade-Hall et al. 1971). During the latter process, high-temperature fluids can cause permanent demagnetization of basalt due to the alteration of titanomagnetite (Sjitkar et al. 2020). Thus, RTP magnetic lows constitute an exploration criterion for the recognition of high-intensity hydrothermal discharge zones and potential SMS deposits in the TAG region (Rona, 1978; Rona, 1980). They constitute a meaningful geophysical indicator to differentiate between mounds of hydrothermal and volcanic origin within the TAG hydrothermal field.

During M127, AUV Abyss was augmented by an Applied Physics System (APS) 1540 Digital three-axis miniature Fluxgate-magnetometer recording at 10 Hz. At the time of the cruise, the Earth's inducing field vector had an inclination of 42°, declination of −15°, and field strength of about 38290 nT (Galley et al., 2021). Induced and permanent magnetization effects caused by the AUV itself were removed from the magnetic data by conducting figure-eight calibration dives to solve for the AUV's magnetic properties following Honsho et al. (2013). The magnetic data illustrated in Fig. 2a have been interpreted regionally by Sjitkar et al. (2019) and locally around the TAG mound by Galley et al. (2021), and are openly available as a 10 m raster (Petersen, 2019).

AUV drift relative to the bathymetry leads to indeterminate errors that may propagate into the workflow. Using an inertial system, the AUV's lateral position is tracked from the initially calibrated position. However, water column currents may induce gradual shifts away from the inferred position. In comparison to the vertical position of the AUV that is determined through altimeter and depth readings, a lateral shift between the magnetic anomalies and bathymetric features can either be geology driven (cf. Sjitkar et al., 2019) or, alternatively, result from positioning errors; both are relevant
constraints for the data integration process of the described ML workflow and are addressed in section 3.3.

2.3 Electrical Conductivity

Accumulations of SMS exhibit a distinct contrast in the electrical resistivity compared to the surrounding basalt (Morgan, 2012; Spagnoli et al., 2016). Sulfide mounds are generally more porous compared to the background basalt (Murton et al., 2019), host high-temperature fluids when active, and contain metalliferous minerals and clays, all attributes that contribute to a decrease of the electrical resistivity. Several electromagnetic applications have been proposed to detect and characterize volumes of minerals for example at TAG (Haroon et al., 2018; Gehrmann et al., 2019) or in the Okinawa Trough (Constable et al., 2018; Ishizu et al., 2019; MacGregor et al., 2021).

2.3.1 Controlled Source Electromagnetic Measurements

Our controlled source electromagnetic (CSEM) data were acquired using two, fixed-offset Vulcan receivers (Constable et al. 2016) towed at distances of 350 m and 505 m behind a 50-m horizontal electric dipole (HED) source (Sinha et al., 1990). The resulting 2D resistivity models computed with MARE2DEM (Key, 2016) are interpreted and discussed by Haroon et al. (2018) and Gehrmann et al. (2019, 2020). In summary, the CSEM conductivity models highlight distinct regions of known SMS through increased electrical conductivity (cf. Fig. 3a). Here, we use the acquired CSEM data to reassess the electrical resistivity distributions at identified high-priority sites. Electrical conductance illustrated along CSEM profiles qualitatively define spatial extents of conductive, possibly mineral-enriched seafloor, and predict if morphological expressions are associated with hydrothermal (conductive) or volcanogenic (resistive) activity. Notably, Gehrmann et al. (2020) demonstrate that navigational uncertainty of the CSEM system is not trivial in such complex bathymetry, which can result in inversion artifacts. The authors mitigate these artifacts by estimating the data quality on navigational uncertainties such as instrument position with respect to the bathymetry. Here, we further reduce potential inversion artifacts by confining CSEM inversion models to distinct regions associated with mounds, mitigating potential misinterpretations caused by over-fitting the data at irrelevant locations.

2.3.2 Transient Electromagnetic Measurements

Marine transient electromagnetic (TEM) data were acquired at two specific sites within the TAG hydrothermal field using GEOMAR’s MARTEMIS system (Fig. 3b and 3c). This coincident loop
system consists of one transmitter and one receiver loop, which are housed in a 6.3 × 6.3 m\(^2\) frame. In regions where seafloor conductivity exceeds the water column conductivity, TEM data will exhibit an increased induced voltage allowing a localized inference of SMS distributions. The system is towed at <1 kn and 5 – 15 m above the seafloor and records the electromagnetic response of a 50% duty-cycle transmitter signal at 10 kHz. The acquired time series are processed considering the distorting effects described by Reeck et al. (2020) and transformed into full-space apparent conductivity curves following Eq. 1 of Haroon et al. (2018). Regions of increased apparent conductivity are generally associated with areas where the seafloor is more conductive than the water column resistivity (\(\rho < 0.3\ \Omega m\)), which in our geological setting is indicative of SMS occurrences (Swidinsky et al. 2012). MARTEMIS positioning was computed through an ultra-short baseline (USBL) transponder attached to the tow cable and merged with the processed apparent conductivity data. The spacing between adjacent stations is approximately 10 m (Fig. 3b and 3c).

2.4 In-situ Data

Overall, 33 gravity cores of max. 3-m length were acquired within the TAG hydrothermal region during the M127 cruise (Petersen et al., 2016). Locations of possible coring sites were selected with the help of the high-resolution AUV bathymetry data. Twenty-three cores contained abundant sediment, eight contained only fragments of gravel, basalt and traces of sediments in the core catcher, and two were empty (Fig. 3, white markers). Among the 23 sediment cores, 10 had visible hydrothermally-influenced indications (Fig. 3, green triangles); the other cores had the visual appearance of background sediments (carbonate ooze) or showed layers of volcanic origin (Fig. 3, red and blue triangles, respectively). Note that the presence of background sediment or empty cores does not rule out hydrothermal activity at greater depth, as penetration of this coring technique was limited to a maximum of three meters.

In addition to gravity cores, rock drill samples have been drilled to a maximum depth of 12.5 m below seafloor (Murton et al., 2019). The obtained samples from the Southern, Rona and MIR mounds show high concentrations of minerals, confirming the hydrothermal origin of these three mounds. Here, we link core sites indicative of hydrothermal alteration with collocated EM resistivity data and models to distinguish spatial extents of mineral zones on these mounds and reassess tonnage estimates.
3 Methods

The workflow is split into four steps as illustrated in Fig. 4: 1) selecting and preparing suitable mid-ocean ridge (MOR) bathymetry data, e.g. from accessible open-source data repositories (Tab. A1), 2) training, validating and testing the U-Net model, 3) post-processing of the model output to derive mound architectures and integrate with concurrent RTP magnetic data, and 4) classification and geophysical analysis of identified mounds. The workflow is scripted in Python (Ver. 3.8.12) and uses the Tensorflow (Ver. 2.4.1) library for machine learning tools.

3.1. Data Preparation: Bathymetry Data

Bathymetry data used for training are identified as suitable, if a large spatial coverage is acquired at either MORs or at specific hydrothermal fields. Bathymetry rasters are subdivided into overlapping patches of 256 x 256 pixels with a step length of 128 pixels. These pixels are manually annotated using a binary representation, where pixels associated with mounds are labeled as True and all other pixels are labeled as False. In total, 1899 mounds were manually annotated using the bathymetry data listed in Tab. A1.

To appear in a common standard that highlights rounded convex and concave morphology through distinct representations, we use a multi-directional slope analysis by mapping the normalized aspect, slope, and the $\partial y$ derivative onto Red, Green and Blue channels of a standardized RGB image, respectively (Fig. 5). The chosen approach converting certain derivatives of bathymetry data into single RGB images facilitates a generalized visual interpretation and aids the model performance. All resulting images show the north flank of mounds in yellow to green moving west to east. Southern mound flanks appear white to blue. Concave features such as pits appear in a reversed manner.

Note that directional dependencies of background features in the preprocessed images remain unaltered through pre-processing (Figure 5). To increase the amount of training data and mitigate learning of directional dependencies in various settings, input bathymetry was augmented by means of a 90° rotation. As mound structures are near-circular structures, they remain rotational invariant although background strike differs (cf. Fig. 5a and b). In total, 2280 RGB images were produced to train, validate and test the U-Net model, each consisting of 65,536 pixels.
3.2 U-Net Implementation, Training and Evaluation

The model architecture yields an end-to-end trainable neural network including segmentation of input images into partitioned pixel sets of corresponding classes. This type of network was first introduced for biomedical image segmentation by Ronneberger et al. (2015) and resembles a symmetric “U” (Fig. 6). In our specific case, the U-Net model distinguishes mound from background features and provides values of probabilities (Hu et al., 2015) as outputs.

For training and testing, only images with at least two percent of the pixels annotated as mounds were considered. Of these, 75 percent were used for training, 20 percent for validation and 5 percent for testing. We use the binary cross entropy loss function defined as:

\[
H_p(q) = \frac{1}{N} \sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))
\]

where \( y_i \) refers to the corresponding binary label of each pixel and \( p(y_i) \) to the predicted value between 0 and 1 within each Epoch \( i \) of training. The accuracy, true positive and false negative metrics were computed to determine a point of early stopping, i.e. a model with sufficient accuracy and minimal over-fitting (Fig. 7). The model outputs are contoured at values of >0.5 to outline lateral mound dimensions of the mound pedestal.

After training, the pre-processed AUV bathymetry image from the TAG area (Petersen, 2016) is presented to the U-Net model as overlapping patches of 256 x 256 pixels. The image reconstruction process is illustrated in Fig. 8. To prevent inaccurate predictions along the image edges, an image smoothing considers only the central 128 x 128 pixels. The outer edges of each predicted patch are neglected. Using an overlapping process, each pixel of the bathymetry grid is included four times within the output prediction. The maximum probability from the four predictions is used as the final pixel probability.

3.3 Post-processing of mound structures

The output mound contours define the location and lateral footprints of each mound, which is calculated (in m²). A minimum threshold of ~1040 m² mound footprint (290 pixels) is introduced to remove most of the falsely detected mound structures caused by geological noise and to focus the analysis on potentially significant SMS volumes (cf. Murton et al. (2019) for discussion). From each of the mound contours, we compute the lateral footprint, maximum height, and median slope using only
pixels located within each mound contour. These parameters are used to describe the general mound architecture and are used as inputs for classification.

For integrating the magnetic anomaly data with the detected mounds from the UNet, an image overlay of gray-scaled hill shade and a diverging red-to-blue magnetic anomaly map (Fig. 2) is cropped and centered around each mound contour, including also peripheral areas. Using the specific red-blue color representation, RTP anomalies appear either red (if positive) or blue (if negative), both being primary color channels within an RGB color spectrum. Color histograms and correlations of the three RGB channels depict positive, negative or a mixture of RTP anomalies into three single values, depending on whether the image is blue, red, or a blue/red blend. The channel correlations serve as inputs for the subsequent classification of the mounds.

3.4. Classification & Evaluation

We applied spectral clustering using the Sci-Kit Learn Python Library on the derived parameters for each mound contour. Where available, we added peripheral SMS indicators derived from gravity cores, electromagnetic data and known SMS edifices to determine mound evolution and assess the mineral potential at confirmed high-priority sites. This ensemble of derived morphological expressions and geological/geophysical characteristics was integrated into a new model for the formation of SMS mounds in the TAG hydrothermal field. Further, it provides the basis to re-discuss the resource potential of the field.

4 Results

U-Net analysis

The U-Net metrics indicate an optimal point of early stopping at around Epoch 158 (Fig. 7). There, the network reached a prediction accuracy of greater than 98.6 and 97.8 percent in training and validation, respectively. The training loss reached 0.032 and the validation loss 0.075 (green and black curves, respectively, in Fig. 7a) using a learning rate of $10^{-4}$. A learning trend is observable in the ensuing epochs especially within the training data. Yet, the trend is less pronounced within the validation data, indicating that predictions will not improve for unseen images. To analyze the efficiency of the model, accuracy, true positives and false negatives are also utilized to understand the general characteristics of U-Net predictions (Fig. 7 b-d and Tab. 1).
Manually annotated mound structures make up, on average, less than 8.2 and 9.1 percent of the total pixels in each of the training and validation images, respectively. This significant imbalance compared to background leads to a high starting accuracy of approximately 91 percent, assuming that all pixels are predicted as background, i.e. False. Other metrics, such as true positives and false negatives are more significant for such imbalanced problems. At the point of early stopping, the trained network can retrieve 87 percent of the true positive pixels in the validation data, meaning that manually classified mound pixels are also classified as mound affiliated pixels by the U-Net. Similarly, false negatives are minimized to less than 1 percent of the total pixels per image during training and validation, indicating only few background formations are being falsely classified as mounds.

In addition to these training metrics, 114 images were used as a test dataset to further assess the U-Net’s efficiency and prediction characteristics for unseen data. A summary of the test data metrics is listed in Tab. 1. A prediction accuracy of >97.6 percent is achieved for the test data. Model efficiency in predicting true positives and avoiding false negatives is comparable to the validation data. Note that there is some bias to consider in the evaluation of these pixel-based metrics. Discrepancies in mound dimensions between manual annotation and automated segmentation will reduce model performance, although a mound is essentially detected by the network. In the majority of studied cases within the test data, mounds were detected and metric deficiencies arise from discrepancies in lateral mound extensions between manual and automated annotation (see Fig. S1 of supplementary materials).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Training</th>
<th>Validation</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benchmark: 0.91603</td>
<td>Benchmark: 0.9103</td>
<td>Benchmark: 0.89777</td>
</tr>
<tr>
<td></td>
<td>Prediction: 0.98696</td>
<td>Prediction: 0.97790</td>
<td>Prediction: 0.976976</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True Positives</td>
<td>Benchmark: 0.08397</td>
<td>Benchmark: 0.08970</td>
<td>Benchmark: 0.10223</td>
</tr>
<tr>
<td></td>
<td>Prediction: 0.07659</td>
<td>Prediction: 0.07816</td>
<td>Prediction: 0.088996</td>
</tr>
<tr>
<td></td>
<td>91.2%</td>
<td>87.1%</td>
<td>87.1%</td>
</tr>
<tr>
<td>False Negatives</td>
<td>Prediction: 0.00738</td>
<td>Prediction: 0.01154</td>
<td>Prediction: 0.01323</td>
</tr>
</tbody>
</table>

The trained U-Net detects a total of 323 mounds within the mapped 49 km$^2$ of the AUV bathymetry data (Fig. 9), each with a lateral footprint greater than 1040 m$^2$ (= 290 pixel). The predictions include all previously identified SMS mounds (cf. Fig.1 and Fig. 9). The lateral mound dimensions match, in most cases, the manual annotation. However, the U-Net model underestimates the spatial footprint for Southern and Double Mound. Both mounds show a tectonized surface texture, deviating from an idealized mound shape, which may explain the reduced model performance. The total number of
known mounds accumulates to 16 (compared to 15 of Gehrmann et al. 2019) because Double Mound is segmented as two individual peaks by the U-Net classification.

The output mounds can be classified into three distinct clusters (Fig. 9) using the elbow method applied during spectral clustering (K-means). Table 2 lists some statistics for each cluster, including the minimum, maximum, and mean mound dimensions. The number of associated known SMS sites defines the priority of each cluster to host SMS. Cluster 1 is assigned a ‘high’ priority, as 10 out of 98 mounds are known to host SMS. A ‘low’ SMS priority is assigned to cluster 3 (1 known SMS site out of 114 mounds), and ‘medium’ to cluster 2 (4 known SMS sites out of 111 mounds).

Table 2: Clustering statistics of output mounds, including the total area of all mounds and associated mound dimensions. The number of known SMS sites associated with each cluster defines the SMS priority of the cluster.

<table>
<thead>
<tr>
<th>Cluster # (SMS priority)</th>
<th>Number of Mounds</th>
<th>Number of known SMS sites per Cluster</th>
<th>Total Area of all Mounds</th>
<th>Footprint</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (high)</td>
<td>98</td>
<td>10</td>
<td>982314 m²</td>
<td>Max: 141310 m² Min: 1052 m² Mean: 10024 m²</td>
<td>Max: 58.62 m Min: 1.35 m Mean: 14.46 m</td>
</tr>
<tr>
<td>2 (medium)</td>
<td>111</td>
<td>4</td>
<td>1188213 m²</td>
<td>Max: 78268 m² Min: 1205 m² Mean: 10704 m²</td>
<td>Max: 52.53 m Min: 0.35 m Mean: 14.98 m</td>
</tr>
<tr>
<td>3 (low)</td>
<td>114</td>
<td>1</td>
<td>942423 m²</td>
<td>Max: 75269 m² Min: 1094 m² Mean: 8266 m²</td>
<td>Max: 56.34 m Min: 1.12 m Mean: 13.62 m</td>
</tr>
<tr>
<td>Total</td>
<td>323</td>
<td>15</td>
<td>3112950 m²</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The clustering is primarily driven by the magnetic anomaly data (cf. Fig. S2 and S3); therefore, 10 of the 15 known SMS sites fall in Cluster 1, showing a distinct negative magnetic anomaly. At these 10 mounds, Szitkar et al. (2019) interprets the magnetic anomalies to represent a vertical hydrothermal conduit centered above their corresponding source. Known SMS sites of Cluster 2 (4) show both positive and negative RTP magnetic anomalies indicating either geological alteration or poor AUV navigation. Cluster 3 contains only one previously known SMS site (Mound #24 from Graber et al. 2020) that is associated with a positive magnetic anomaly. Images of all clustered mounds are displayed in Fig. S2 of the supplementary materials.
Despite their magnetic signature, mound morphology is similar in all clusters with respect to their mean heights and mound footprints (cf. Tab. 2 and Fig. S3 in the supplementary materials). Other morphological features indicative of hydrothermal activity, such as jagged contours and number of peaks (Jamieson et al. 2014), could not be identified by tour workflow as important parameters for differentiating mound evolution in the TAG area. This leaves the magnetic anomaly as the strongest spatial indicator in the available data set.

Spatial Distribution of Morphological Features and Magnetic Footprint

The high-priority sites (Cluster 1) occur spatially confined in three bands (labeled Southern, Central and Northern Band in Fig. 9). All bands strike NW-SE, roughly perpendicular to the axis of the Mid-Atlantic Ridge, and coincide well with interfaces between different structural domains identified by Graber et al. (2020). The southern band lies within a region of oblique faults and fissures (Fig. 8 of Graber et al. 2020), that have been suggested to promote upward migration of hot fluids at TAG and other regions (Anderson et al. 2015). The central band of mounds is located within the NW-SE extension of the, so-called, Three-mound area (cf. Graber et al., 2020), runs parallel to mapped corrugations, and connects the Three-Mound regions to the MIR zone (Fig. 9). The northern band lies north of a zone with chaotic seafloor morphology and positive magnetic signature, separating the smooth bathymetry and negative anomalies of the central and northern bands.

The alignment of Cluster 1 mounds is interrupted in several locations. In the southern band, at around 26.138°N and 44.837°W, Cluster 1 mounds are not associated with oblique fissures mapped by (Graber et al. 2020). Instead, negative magnetic anomalies correlate with ‘fresh’ pillow mounds. Such potentially younger magmatic features may mask the oblique fissures typical for the southern band.

The northern band contains only one known hydrothermal site named Shimmering, but multiple gravity cores indicate an abundance of hydrothermally-altered sediments in the area (cf. Fig. 9). However, mounds located at the western section of the northern band are structurally interpreted as pillow mounds (Graber et al. 2020).

Analysis using Electromagnetic Data

SMS potentials have often assumed homogeneously distributed metal grades across a mounds morphological footprint and its corresponding stockwork zone. Although tonnage estimates are often based on in-situ measurements (i.e. core-log data/seafloor drilling), derived mineral potentials are
likely too optimistic because in-situ data are 1) available at only few representative mounds, 2) generally obtained at the points of highest interest, i.e. where SMS is apparent in seafloor imagery and 3) penetrate only few meters into the subsurface. In the majority of cases, the structural heterogeneity of individual mounds is either neglected or considered too simplistic, assuming that SMS with high metal grades are distributed across the entire lateral extent of each mound. To constrain this rather generalized assumption, resistivity models derived from CSEM or TEM data appear more suitable to understand the degree of mineralization away from the point-scale core-log data.

Here, we focus on the electromagnetic data acquired only along high-priority mounds (Cluster 1) or verified mounds that have been documented in preceding literature. CSEM resistivity models that were recomputed for Cluster 1 mounds are illustrated in Fig. 10 using data acquired by Gerhmann et al. (2019). Note that only those mounds are considered that are intersected by CSEM transects. Regions of low resistivity are illustrated by red to orange coloring, background resistivity through green and blue coloring.

The CSEM resistivity models illustrate that the majority of investigated Cluster 1 mounds are associated with a distinct low-resistivity anomaly of variable magnitude. This, together with core-log data and grab samples (Peterson et al. (2016); Murton et al. (2019)) confirms a certain degree of mineralization at each of the prospective sites. It needs to be noted that in a few instances, i.e. Fig. 10d and 10g, CSEM data has only limited resolution due to the large vertical offset between measurement system (denoted by black markers) and seafloor, thus, limiting a quantitative analysis of mineral grades using CSEM data. Laterally, resistivity distributions and amplitudes do indicate a high degree of certainty and illustrate characteristics of mound composition. Shinkai and Southern mounds, residing in the central band (Fig. 10c through 10e) exhibit a relatively homogeneous conductive structure. These two mounds appear as low-resistive anomalies across their entire lateral footprint (cf. apparent conductivity data in Fig. 3b) indicating that previous mineral estimates could be accurate. In contrast, Double (Fig. 10e) and Rona (Fig. 3b) mound exhibit a low resistive anomaly only in the vicinity of their peaks with no notable contrast to the background resistivity at their pedestal. Although, Rockdrill cores of up to 12 m acquired at the peak of Rona confirm high metal concentrations within a sulfide layer (Murton et al., 2019), TEM data indicates that the majority of Rona’s volume is of lower economic value due to a missing apparent conductivity anomaly (Fig. 3b).
MIR mound has the largest spatial footprint of Cluster 1 mounds in the study area (Fig. 10f). Previous predictions based on mound volumes and extrapolated metal grades derived from gravity cores and rock drill data suggest MIR to be the most economical site within the TAG hydrothermal field (Graber et al. 2020). CSEM inversion of MIR illustrated in Fig. 10f shows an irregular distribution of low-resistive zones across the mound transect indicating the presence of mineralized sediments, but not in the quantity suggested by previous studies. The higher resolution TEM data acquired along multiple transects across the MIR mound support this hypothesis (Fig. 3c) contradicting the notion of MIRs high mineral potential. The point-scale gravity core and Rockdrill data (Peterson et al. (2016); Murton et al. (2019)) were acquired in the northwestern region of the MIR contour, where high apparent conductivities exist. Most of the other regions of the mound structure are not associated with a distinct resistivity anomaly, thus, indicating that mineralization is irregularly distributed across MIR and that tonnage predictions for MIR may be significantly overestimated.

TAG is arguably the most prominent mound in the study area. Multiple geophysical and geological surveys have focused on the internal mound structure, including the Ocean Drilling Program (ODP) Leg 158 experiment (e.g. Humphris et al., 1995). As such, the internal structure of TAG is well constrained and serves as a blueprint for estimating mineral potentials for other mounds, where less knowledge about internal structure and data are available. The CSEM resistivity models of TAG (Fig. 10g and 10h) show an E-W and N-S transect crossing the mound, respectively. Unfortunately, navigation during the E-W transect was chosen too conservative with towing altitudes exceeding 100 m above seafloor, which resulted in inadequate resolution for the TAG’s resistivity structure. However, the N-S profile (Fig. 10h) is intriguing, as it supports the asymmetric distribution of mineralization presented in previous studies (e.g Galley et al. 2020). Moreover, the CSEM resistivity model indicates that a significant resistivity contrast compared to the background basalt may only exists for the massive pyrite, pyrite-anhydrite, pyrite-silica and possibly the pyrite silica units (see Knott et al. 1998). Hence, the applied CSEM configuration is likely unsuitable for detecting the stockwork structure and requires higher resolution CSEM data acquired in a 3D survey as demonstrated by MacGregor et al. (2021).

5 Discussion

5.1 Automated SMS Mapping using Machine Learning

The presented workflow can be used as a blueprint for prioritizing SMS exploration targets at mid-ocean ridges and understanding distributions of mineral potentials. The workflow reduces the total area of interest from the surveyed 49 km² to 3.1 km², which in turn can be further reduced to either 1.92 km²
(Clusters 1 and 2) or 0.98 km² (only Cluster 1) using additional magnetic constraints (i.e. mounds with magnetic lows). Moreover, the latitudinal bands of hydrothermal activity identified through this integrated analysis reveal prospective areas where to search for SMS and may also exist at other MORs. The workflow is fully automated, allowing us to identify regions of interest in quasi real-time (if a pre-trained model exists) when new data is acquired, thus, reducing exploration costs considerably and permitting more focused surveying. Additionally, the workflow is adaptable to future developments in marine mineral exploration and research and its application in other survey areas with similar or additional data layers appears feasible.

The data preparation part of our workflow seeks to unify the bathymetry data, acquired at different spatial resolution and at various sites across the globe, into a common representation independent of the actual depth, slope, and curvature within a given area. Juliani & Juliani (2021) propose a principal component analysis (PCA) consisting of both a change in slope and a multi-directional shading of elevation data in order to reduce the bathymetric inputs. However, our analysis shows that PCA may not generalize well for bathymetry data acquired at different regions with variable roughness and geological strike. A key advantage of our proposed processing scheme is notably that bathymetry data from different regions will unify onto a single coherent visualization.

Moreover, the workflow also allows to test and use other types of ML segmentation tools, and to include additional data layers (e.g. high-resolution backscatter, self-potential, etc). As many of these additional data layers are currently not available in open-access repositories, they can be integrated best within the post-processing step. As backscatter and self-potential data become more readily available, it is also feasible to train the U-Net directly for different types of mound characteristics. The integration of such additional data will likely increase ambiguity, but presumably achieve higher certainty in identifying SMS sites.

Following Szitkar et al. (2019) and Rona (1978; 1980), hydrothermal mounds within the TAG hydrothermal field are associated with a distinct negative RTP magnetic anomaly, whereas volcanic edifices typically display positive values. This characteristic proves suitable for clustering depicted mound contours into groups to identify their potential origins. However, three aspects must be considered to integrate the magnetic footprint of a mound or a group of adjacent mounds with the corresponding bathymetry attributes.
1. Magnetic anomaly data are acquired at a resolution of 10 m grid spacing compared to the 2 m resolution of the bathymetry data. A pixel-wise comparison between magnetic anomalies and morphological features requires an up-sampling of the magnetic anomaly data, which may lead to interpolation artifacts.

2. An RTP magnetic anomaly shows magnetization and geomagnetic field vectors that render vertical anomalies above the causative body. However, Szitkar et al. (2019) discuss that tectonic events can tilt the crustal block causing altered shapes of the magnetic anomalies, leading to incoherent magnetic anomalies associated with morphological expression. A similar effect is also observable if tectonic forces act on a previously deposited mound.

3. AUV magnetic data are susceptible to errors that arise from inaccuracies in AUV positioning relative to its calibrated coordinates. This may lead to a lateral shift between the morphological expression and the corresponding RTP magnetic anomaly.

All three undetermined circumstances lead to uncertainties in a pixel-wise integration of the magnetic and corresponding bathymetry pixels. Hence, a relaxation of spatial similarities between mound structures and resulting magnetic anomalies is required and was implemented in our analysis.

The approach would benefit from a greater number of annotated bathymetry data available in online repositories to improve model training. The chosen study area belongs to the most studied hydrothermal sites globally, and, provides a solid first training set. The developed workflow focuses on the analysis of high-resolution bathymetry data, which resemble the most common collected data in seafloor exploration. Given the steady increase of sea-going SMS research, manual assessments of each individual data layer becomes increasingly difficult and automatization of workflows will be inevitable. Therefore, application of the workflow in other hydrothermal areas either at mid-ocean ridges or other geological environments with complex, rough terrain, is a crucial future task.

5.2 Implications for hydrothermal activity in the TAG area

Notably, marine mineral exploration is a complex endeavor unlikely solved by a silver bullet approach. Thus, various ML strategies and conventional geoscientific research will attest a feasibility for detecting SMS and estimating mineral potentials. The proposed ML strategy does not contradict this notion, but instead, offers a means of integrating multivariate data into a common interpretation scheme that is easily audited. Note that the delineation of convex structures in bathymetry data underlies some variability resulting from terrain analysis and, even if conducted manually, remains ambiguous due to
interfering geomorphic processes that mask or distort typical mound morphologies. Hence, although
mapping the correct mound dimensions is significant for addressing mineral potentials, discrepancies
between manual and automated segmentation are expected.

Despite these explainable deviations, the spatial alignment of Cluster 1 bands is clearly visible and the
correlation to structural domains defined by Graber et al. (2020) is apparent. This may support the
hypotheses of a structural heterogeneity within the hydrothermal field dictating the distribution of SMS
edifices, as proposed by Graber et al. (2020). Both the spatial extent of the alignment and the
occurrence of deviating areas indicate a structural constraint in the deep subsurface. This dominant
structural constraint supports an interpretation where a strongly distorted subsurface structuring (e.g.
bend detachment fault) leads to focused fluid flow in the deep subsurface that results in linear, off-axis,
distribution of hydrothermal edifices. Moreover, the various upflow zones are likely to span a region
larger than the investigated area of study. Further sites may be located north of Shimmering and also
south of TAG, as well as in the westward extension of the three bands. Conclusively, the presented
workflow has demonstrated a successful amalgamation of spatially acquired bathymetry and magnetic
data, which could be used to inform future AUV bathymetry and magnetic surveying.

Although the number of potential SMS sites drastically increased through this automated analysis,
mineral potential of the TAG hydrothermal field is likely lower than originally presumed.
Electromagnetic data illustrates that mineralized zones for the largest proposed SMS sites are generally
heterogeneously distributed across the mound contour, thus, contradicting the proposed high tonnage
estimates using homogeneously distributed mineral grades derived from point-scale measurements. We
propose that future analysis of SMS tonnages should incorporate multiple seafloor drillings conducted
across the mound contours with additional data layers (e.g. backscatter, seismics, and 3D CSEM
inversion models) to achieve a high degree of certainty in the tonnage estimation. If such high-
resolution survey strategy for SMS sites is economical, remains beyond the scope of this study.

The current SMS priority depends mainly on AUV magnetic data and on the number of known SMS
sites within a given cluster. The former is not typically considered a conventional data layer in SMS
exploration and should be added to the necessary SMS exploration criteria. Furthermore, it cannot be
ruled out that future endeavors may potentially change the presented prioritization of SMS mounds
through more SMS discoveries or through the acquisition of additional data layers (e.g. high-resolution
backscatter, resistivity or self-potential data). It is expected that not all mounds within the high-priority
cluster are associated with SMS edifices and that additional data layers will improve the certainty of the automated SMS prediction. Overall, a more diversified dataset measured at numerous SMS sites across the global MORs will only help improve our understanding of SMS predictors and improve future developments of ML workflows.

Marine CSEM and TEM data have demonstrated additional value when conducting tonnage estimates for SMS sites. However, more development is required to improve the significance of resistivity models to help quantifying volumes of mineralized zones within a mound edifice. Most notably, high-resolution 3D surveys using AUV technology are likely required to accurately derive spatial extensions of conductive material in these remote settings. MacGregor et al. (2021) have already presented a first 3D application and inversion of CSEM data and others are likely to follow suit, given the increased value of electrical resistivity models to constrain volumetric predictions.

6 Conclusion

A workflow to conduct automated SMS site detection using multivariate geoscientific data is presented that employs a U-Net neural network to identify prominent mound-like morphologies in bathymetry data. The predicted contours are subsequently integrated with other spatial data layers (e.g. AUV magnetic data) to identify high-priority sites for SMS prospecting. Within the 49 km$^2$ grid of high-resolution bathymetry data, 323 mounds were detected. 98 of these were classified as high priority due to their architecture and magnetic signature. Moreover, from the automated analysis 14 (10 high, 4 medium) of the 15 known SMS sites in the TAG area were identified as either high or medium priority. Only one known site was classified within the low-priority group. The high-priority sites were spatially distributed into latitudinal bands, supporting the hypotheses that focused fluid-flow at depth leads to linear distributions of off-axis SMS edifices in the area.

The presented workflow cannot only be used to improve analysis and interpretation of previously surveyed areas, but also serve as a blueprint to optimize SMS exploration at sea. The trained model can be applied on newly acquired bathymetry data in quasi real-time to determine prospective zones for more detailed confirmation/visualization studies. Thus, optimizing the use of ship time and reducing exploration costs. The workflow is very adaptable to include additional data layer such as backscatter, self-potential, turbidity and other water column data maps if available.
Electrical resistivity models demonstrate that mineralization of SMS mounds are less homogeneous than often considered. Thus, indicating that high-grade mineral contents in SMS are not equally distributed across the entire mound and that tonnage estimates may be significantly overestimated. Consequently, although the workflow detects many more potential SMS edifices than previously known, the overall resource potential of the TAG hydrothermal field is likely lower than previously assumed.

7 Acknowledgments
AH was in part funded by the Digital Earth and SMART Projects at GEOMAR Helmholtz Centre of Ocean Research Kiel.

8 References


Villinger, Heinrich; Strack, Anne; Gaide, Stefanie; Thal, Janis (2018). Gridded bathymetry of North Pond (MAR) from multibeam echosounder EM120 and EM122 data of cruises MSM20/5 (2012) and MSM37 (2014). Department of Geosciences, Bremen University, PANGAEA, https://doi.org/10.1594/PANGAEA.889439.
### Table A1: List of open-access bathymetry data used for training the U-Net.

<table>
<thead>
<tr>
<th>Index</th>
<th>Longitude Min.</th>
<th>Longitude Max.</th>
<th>Latitude Min.</th>
<th>Latitude Max</th>
<th>Resolution</th>
<th>Source / Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-44.99625</td>
<td>-44.69042</td>
<td>25.98625</td>
<td>26.28625</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic <a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>2</td>
<td>-43.23542</td>
<td>-41.31541</td>
<td>29.69375</td>
<td>30.78042</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic <a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>3</td>
<td>-34.45958</td>
<td>-33.12792</td>
<td>35.45792</td>
<td>37.50708</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic <a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>4</td>
<td>-27.14292</td>
<td>-23.76958</td>
<td>61.70625</td>
<td>63.44625</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic <a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>6</td>
<td>-44.85162</td>
<td>-44.75758</td>
<td>26.10058</td>
<td>26.18684</td>
<td>~ 2 m</td>
<td>Petersen (2019) <a href="https://doi.org/10.1594/PANGAEA.899415">https://doi.org/10.1594/PANGAEA.899415</a></td>
</tr>
<tr>
<td>7</td>
<td>-44.97116</td>
<td>-44.86794</td>
<td>13.27914</td>
<td>13.34156</td>
<td>~ 2 m</td>
<td>Escartin &amp; Petersen (2017) and Escartin et al. (2017)</td>
</tr>
<tr>
<td>8</td>
<td>-45.00966</td>
<td>-44.88124</td>
<td>13.48664</td>
<td>13.52226</td>
<td>~ 2 m</td>
<td>Escartin &amp; Petersen (2017) and Escartin et al. (2017)</td>
</tr>
<tr>
<td>10</td>
<td>-34.99333</td>
<td>-31.87833</td>
<td>56.69666</td>
<td>58.00000</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic <a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>11</td>
<td>-42.40119</td>
<td>-41.72925</td>
<td>29.82935</td>
<td>30.36156</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic <a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>12</td>
<td>-129.1490</td>
<td>-129.0070</td>
<td>47.88199</td>
<td>48.08900</td>
<td>~ 1 m</td>
<td>Clague et al. (2015) <a href="http://get.iedadata.org/doi/321990">http://get.iedadata.org/doi/321990</a></td>
</tr>
<tr>
<td>14</td>
<td>-29.93012</td>
<td>-22.78952</td>
<td>35.99997</td>
<td>39.50023</td>
<td>~ 45 m</td>
<td>Hübscher &amp; Beier (2022) <a href="https://doi.pangaea.de/10.1594/PANGAEA.945528">https://doi.pangaea.de/10.1594/PANGAEA.945528</a></td>
</tr>
</tbody>
</table>
Figure 1: Bathymetry maps of the study area. (a) GEBCO bathymetry (shaded map) overlain by the ship-based bathymetry data acquired during RV Meteor cruise M127 with a spatial resolution of 30 m. Gray outlines denote visible mound structures whereas the black outlined region denotes the high-resolution bathymetry survey illustrated in (b). (b) AUV bathymetry data acquired with a spatial resolution of 2 m using the same color scale as in (a). Known SMS mounds are outlined and labeled as depicted by Graber et al. (2020).
Figure 2: Overlay map of the hillshade bathymetry (2 m resolution) and the magnetic anomaly map with 10 m spatial resolution from Pedersen (2016). Outlined in black are the known SMS mounds depicted by Graber et al. (2020).
Figure 3: (a) Hillshade map of the bathymetry data with a spatial resolution of 2 m. Computed electrical conductance values derived from 2D CSEM resistivity models of Gehrmann et al. 2019 are displayed with color-coded markers. Light colors denote a low and hot colors a high conductance. (b) Zoom-in of the Three-Mound region overlain by the transformed apparent conductivity values obtained by TEM measurements. (c) Same as (b) but for the MIR zone. Outlines of the mounds denote the manually-annotated lateral mound dimensions from Graber et al. (2020). Triangular markers in (a) through (c) illustrate the locations of the 3 m gravity cores and the lithology observed within the core samples (Petersen, 2016).
Figure 4: Schematic of the applied workflow including all relevant steps applied, i.e. data preparation, U-Net implementation, post-processing and classification.
Figure 5: Image unification example of relevant bathymetry features into a common visual representation that is generically applicable to coherently address bathymetry data acquired at different regions across the globe. The aspect, slope, and $\partial y$ derivative of the bathymetry are mapped onto the red, green and blue channels of a standardized 0-255 RGB image (left column of each panel). In this representation, mounds appear directionally invariant with coherent color representation. (a) The original input data and (b) the original input data rotated by 90°. Contrarily, the background bathymetry differs based on the predominant strike direction of the seafloor morphology, whereas prominent mound features remain rotational invariant.
Figure 6: Schematic of the U-Net architecture used for semantic segmentation of the bathymetry data (modified after Ronneberger et al., 2015).
Figure 7: (a) Binary cross-entropy loss function used during training (green) and validation (black). Additional metrics, i.e. (b) accuracy, (c) true positives and (d) false negatives are also used to assess the model performance. Note that true positives and false negatives are normalized to represent percentages of pixels per image. The vertical dotted line denotes the point of early stopping whereas the horizontally dashed lines in (c) represent the average number of pixels affiliated with mound structure within training and validation data.
Figure 8: Workflow of mound prediction for the AUV bathymetry raster. The pre-processed AUV bathymetry map is cropped into overlapping 256 x 256 patches, which are presented to the U-Net for prediction. The output is smoothed using displayed window where pixels within the black region are neglected due to edge effects that deteriorate predictions (as observed within the testing phase). The right panel shows the final prediction map that is utilized for further processing.
Figure 9: Bathymetry map containing the 323 mounds with a lateral of footprint greater than 290 pixels (1040 m$^2$) as predicted by the U-Net. The mounds are illustrated by white contours and are classified as either Cluster 1 (high priority), Cluster 2 (medium priority) or Cluster 3 (low priority). Dotted gray lines illustrate the interpreted boundaries of the three latitudinal bands containing hydrothermal edifices.
Figure 10: (a) Hillshade bathymetry data with classified mounds as illustrated in Fig. 9. Panels (b) through (i) show electrical resistivity models for each of the high priority mounds intersecting a CSEM profile. The lateral extent of each profile is illustrated by red lines in (a). In (b) through (i), red color shading indicates low resistivity (mineralization), whereas green/blue are more resistive background basalts. Black markers denote transmitter positions along profile.
Automated seafloor massive sulfide detection through integrated image segmentation and geophysical data analysis: Revisiting the TAG hydrothermal field

Amir Haroon\textsuperscript{1,2}, Hendrik Paasche\textsuperscript{3}, Sebastian Graber\textsuperscript{1,4}, Sven Petersen\textsuperscript{1}, Eric Attias\textsuperscript{5,6}, Marion Jegen\textsuperscript{1}, Romina Gehrmann\textsuperscript{7}, Sebastian Hölz\textsuperscript{1}, Meike Klischies\textsuperscript{1,8}

\textsuperscript{1}Helmholtz Centre for Ocean Research, GEOMAR Kiel, Wischhofstr. 1-3, 24148, Kiel Germany
\textsuperscript{2}Hawaii Institute of Geophysics and Planetology, School of Ocean and Earth Science and Technology, University of Hawai‘i at Manoa, Honolulu, HI, USA
\textsuperscript{3}UFZ Helmholtz Centre for Environmental Research, Leipzig, Germany
\textsuperscript{4}Helmut-Schmidt Uni Hamburg
\textsuperscript{5}Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA
\textsuperscript{6}Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA
\textsuperscript{7}Dalhousie University, 6299 South St, Halifax, NS, B3H 4R2, Canada
\textsuperscript{8}North.io GmbH, Einsteinstraße 1, 24118 Kiel, Germany

ORCID IDs
Amir Haroon: 0000-0001-5138-6730; Hendrik Paasche: 0000-0003-1814-0645; Sebastian Graber: 0000-0003-1154-6968; Eric Attias: 0000-0003-4308-1683; Romina Gehrmann: 0000-0002-3099-2771; Sebastian Hölz: 0000-0003-2839-9394; Marion Jegen: 0000-0001-6307-8428; Sven Petersen: 0000-0002-5469-105X; Meike Klischies: 0000-0003-4319-4573

Key Points
- Developed two-step machine learning workflow to identify mound structures in bathymetry and classify their origins based on auxiliary data.
- Substantial increase of potential SMS edifices detected within the TAG hydrothermal field distributed within latitudinal bands.
- SMS mineral potential is likely lower than perviously assumed due to heterogeneously distributed mineralization within mounds.
Abstract

Accessible seafloor minerals located near mid-ocean ridges are noticed to mitigate projected metal demands of the net-zero energy transition, promoting growing research interest in quantifying global distributions of seafloor massive sulfides (SMS). Mineral potentials are commonly estimated using geophysical and geological data that lastly rely on additional confirmation studies using sparsely available, locally limited, seafloor imagery, grab samples, and coring data. This raises the challenge of linking in-situ confirmation data to geophysical data acquired at disparate spatial scales to obtain quantitative mineral predictions. Although multivariate datasets for marine mineral research are incessantly acquired, robust, integrative data analysis requires cumbersome workflows and experienced interpreters. Here, we introduce an automated two-step machine learning approach that integrates automated mound detection with geophysical data to merge mineral predictors into distinct classes and reassess marine mineral potentials for distinct regions. The automated workflow employs a U-Net convolutional neural network to identify mound-like structures in bathymetry data and distinguishes different mound classes through classification of mound architectures and magnetic signatures. Finally, controlled source electromagnetic data is utilized to reassess predictions of potential SMS volumes. Our study focuses on the Trans-Atlantic Geotraverse (TAG) area, which is amid the most explored SMS area worldwide and includes 15 known SMS sites. The automated workflow classifies 14 of the 15 known mounds as exploration targets of either high- or medium-priority. This reduces the exploration area to less than 7% of the original survey area from 49 km$^2$ to 3.1 km$^2$.

Keywords: Convolution Neural Networks, Seafloor Massive Sulfides, Bathymetry, Magnetic Anomaly, CSEM

Data Availability Statement:
The bathymetry data used for training the U-Net model are available on open-access repositories as listed in Tab. A1. The AUV-bathymetry and magnetic anomaly maps are available from Peterson (2019) [https://doi.pangaea.de/10.1594/PANGAEA.899415] and CSEM data from Gehrmann (2019) [https://doi.pangaea.de/10.1594/PANGAEA.899073].
1 Introduction

Over 700 active, inactive and extinct hydrothermal venting sites (cf. definitions in Jamieson & Gartman, 2020) are known to exist along mid-ocean ridges, volcanic arcs, or back-arc spreading centers (Beaulieu et al. 2015; Beaulieu & Szafranski, 2020). Their existence is documented through hydrothermal plumes that are visually confirmed using a suite of underwater-vehicles (e.g., Murton et al., 2019), towed-camera systems (Beaulieu et al., 2013), or via in-situ probing such as gravity coring (Petersen et al., 2016) or seafloor drilling (e.g., Murton et al. 2019). Understanding the distributions of hydrothermal venting, often associated with the evolution of seafloor massive sulfides (SMS), remains a prevalent research topic motivated by the increased demand of strategic minerals needed to foster the net-zero energy transition. The economic and environmental challenges of modern society interfaces with various fields of marine research to predict where subsurface processes transport mineral-enriched, high-temperature fluids from the deep lithosphere towards the seafloor. In many cases these processes are associated with mineral accumulations that form distinct mound-like expressions (e.g., Fouquet et al. 2010) often referred to as SMS mounds. Current estimates suggest that the abundance of known SMS can contribute a fractional supply of strategic metals in the future (Hannington et al., 2011), albeit the known uncertainties with respect to size, distribution, volume, as well as the environmental impact that would succeed mining these potential deep-sea resources.

Tonnage estimates of marine minerals are interpreted from seafloor morphology (e.g. Fouquet et al. 2010; Graber et al. 2020), geophysical analyses (Haroon et al. 2018; Gehrmann et al. 2019; Murton et al. 2019; Galley et al. 2021), or global extrapolations from deposit occurrences along spreading centers (Hannington et al. 2011). Many studies have focused on previously known SMS sites and their immediate surroundings (e.g. Jamieson et al. 2014; Murton et al. 2019; Graber et al. 2020), likely leading to an underestimation of mapped SMS edifices within a given region. For example, Jamieson et al. (2014) discovered over 400 undocumented SMS edifices through manual analysis of high-resolution bathymetry data along the Endeavour Segment. These findings highlight the prevalent question of where to sample next and in which spatial resolution? A question often guided by experience, funding, technological constraints and ship-time availability. Moreover, marine scientists recognize the challenges of differentiating prospective SMS mounds from morphologically comparable volcanic constructions (Jamieson et al. 2014) based on bathymetry data alone. In such cases, multivariate databases that include geophysical, geochemical, and geological data acquired across disparate spatial scales can help to 1) identify regions of interest for more detailed in-situ sampling or visual confirmation studies, 2) optimize the use of ship-time through targeted, pre-informed surveying and 3)
improve volumetric estimates of SMS candidates via integrative geophysical analyses. However, to process, analyze and interpret this mass of information, integrative data workflows are pivotal for extracting valuable information optimally, improving decision-making tools for localized sampling, and providing more rigorous estimates of known SMS provinces. Here, machine learning (ML) offers avenues to develop coherent data workflows and processing chains sufficiently generic and thereby transferable across various geological domains and data layers. We demonstrate an ML application together with conventional geophysical analyses to automatically detect mound-like morphology on the seafloor and identify distinct high-interest areas based on auxiliary geological and geophysical information to characterize and quantify mineral-enriched SMS targets.

Geoscientific studies utilizing ML are progressively increasing in environmental and exploration research (e.g., Bouwer et al. 2022; Koedel et al. 2022). ML applications differ not only in the type and spatial and temporal resolution of the input data, but also in the applied techniques. In marine mineral research, ML applications have applied random forest classification (e.g. Gazis et al., 2018) or neural networks (e.g. Keohane & White, 2022; Juliani & Juliani, 2021), focusing mainly on one or two types of input layers (e.g. seafloor images and/or bathymetry data). However, as interdisciplinary SMS databases grow via contributions from geological, geophysical (e.g. Müller, Schwalenberg, Barckhausen, 2023), biological and geochemical applications, ML workflows are expected to also evolve into more generic implementations to facilitate this growing demand of interdisciplinary marine research.

Our study leverages existing ML applications previously conducted in the marine environment. We introduce a workflow that integrates concurrent data acquired at different spatial scales to better describe the mineral potential within the Trans-Atlantic Geo-traverse (TAG) hydrothermal field. First, a modified approach adapted from Juliani & Juliani (2021) is utilized to identify mound structures in bathymetry data using a U-Net convolutional neural network. Subsequently, the identified mounds are amalgamated with multivariate geophysical and geological databases to assess potential SMS edifices in greater detail. We test the developed workflow using data described in Petersen et al. (2016) and Murton et al. (2018) during two Blue Mining expeditions (https://bluemining.eu/) at the TAG hydrothermal field, and compare the results to manual classification studies of Graber et al. (2020) and Murton et al. (2019). This study extends previously published concepts of marine mineral research, which in most cases use a uni-variate database, by including more diversified, multivariate input layers
such as reduced-to-the-pole (RTP) magnetics, controlled source and transient electromagnetics (CSEM and TEM), core and grab samples all acquired across various spatial scales.

2 Geological and Geophysical Data

The data used in this study were previously published in scientific literature, i.e. Petersen et al. (2016); Murton et al. (2018); Szitkar et al. (2019); Haroon et al. (2018); Gehrmann et al., (2019); Graber et al. (2020); Gehrmann et al. (2020); Galley et al. (2021). The following describes relevant aspects of the data, which is needed in the context of the ML implementation. Please refer to the above-mentioned literature for more details on data acquisition and geological/geophysical interpretations. It is important to note that from a data science perspective, the available data introduce \textit{a-priori} bias, as these were acquired with the specific purpose of imaging certain physical parameters that, from a geological perspective, are associated with the evolution of SMS. Thus, unknown correlations that extend beyond the current geological understanding of SMS evolution are likely neglected in the presented ML workflow.

2.1 Bathymetry Data

Seafloors host numerous focused fluid discharge sites that can appear as mounded manifestations in the seafloor topology (Olakunle et al., 2021). Linking these manifestations to either a volcanic or hydrothermal origin and deriving their potential for forming metalliferous accumulations requires auxiliary data acquired at each specific site. The high-resolution seafloor bathymetry data provides a spatial baseline of where to sample for potential SMS occurrences and also the structural framework for volumetric predictions (e.g., Jamieson et al., 2014; Graber et al., 2020).

The high-resolution bathymetry data interpreted and classified by the U-Net were collected during research cruise M127 (RV Meteor, 2016) using ship-based multibeam (Fig. 1a) and GEOMAR’s Autonomous-Underwater Vehicle (AUV) Abyss (Petersen et al., 2016). Data were acquired using a RESON Seabat 7125 multibeam echosounder, navigated at a speed of three knots using a frequency of 200 kHz. The line spacing between adjacent profiles was between 80 m and 100 m at an average altitude of 84 m relative to the seafloor, resulting in a 2 m grid resolution (Fig. 1b). The bathymetry data were processed using the software package MB Systems (https://www.mbari.org/technology/mb-system/) and georeferenced based on prominent seafloor features (Graber et al. (2020) and Szitkar et al. (2019)). This high-resolution bathymetry constructs the baseline of positioning morphological structures during automated segmentation.
To further optimize the U-Net model, we utilize available high-resolution AUV bathymetry data acquired at different SMS sites around the globe (e.g. Clague et al., 2015; Escartin & Petersen, 2017) and other openly-available bathymetry data. All of the applied bathymetry grids utilized to train, validate and test the U-Net are listed in Tab. A1.

2.2 AUV-Based Magnetic Data

The magnetic properties of seafloor basalts are dictated by two alteration processes, namely deuteric oxidation during the initial cooling phase and the superimposed regional hydrothermal alteration that occurs at younger ages (Ade-Hall et al. 1971). During the latter process, high-temperature fluids can cause permanent demagnetization of basalt due to the alteration of titanomagnetite (Sztitkar et al. 2020). Thus, RTP magnetic lows constitute an exploration criterion for the recognition of high-intensity hydrothermal discharge zones and potential SMS deposits in the TAG region (Rona, 1978; Rona, 1980). They constitute a meaningful geophysical indicator to differentiate between mounds of hydrothermal and volcanic origin within the TAG hydrothermal field.

During M127, AUV Abyss was augmented by an Applied Physics System (APS) 1540 Digital three-axis miniature Fluxgate-magnetometer recording at 10 Hz. At the time of the cruise, the Earth's inducing field vector had an inclination of 42°, declination of −15°, and field strength of about 38290 nT (Galley et al., 2021). Induced and permanent magnetization effects caused by the AUV itself were removed from the magnetic data by conducting figure-eight calibration dives to solve for the AUV's magnetic properties following Honsho et al. (2013). The magnetic data illustrated in Fig. 2a have been interpreted regionally by Szitkar et al. (2019) and locally around the TAG mound by Galley et al. (2021), and are openly available as a 10 m raster (Petersen, 2019).

AUV drift relative to the bathymetry leads to indeterminate errors that may propagate into the workflow. Using an inertial system, the AUV's lateral position is tracked from the initially calibrated position. However, water column currents may induce gradual shifts away from the inferred position. In comparison to the vertical position of the AUV that is determined through altimeter and depth readings, a lateral shift between the magnetic anomalies and bathymetric features can either be geology driven (cf. Szitkar et al., 2019) or, alternatively, result from positioning errors; both are relevant.
constraints for the data integration process of the described ML workflow and are addressed in section 3.3.

2.3 Electrical Conductivity

Accumulations of SMS exhibit a distinct contrast in the electrical resistivity compared to the surrounding basalt (Morgan, 2012; Spagnoli et al., 2016). Sulfide mounds are generally more porous compared to the background basalt (Murton et al., 2019), host high-temperature fluids when active, and contain metalliferous minerals and clays, all attributes that contribute to a decrease of the electrical resistivity. Several electromagnetic applications have been proposed to detect and characterize volumes of minerals for example at TAG (Haroon et al., 2018; Gehrmann et al., 2019) or in the Okinawa Trough (Constable et al., 2018; Ishizu et al., 2019; MacGregor et al., 2021).

2.3.1 Controlled Source Electromagnetic Measurements

Our controlled source electromagnetic (CSEM) data were acquired using two, fixed-offset Vulcan receivers (Constable et al. 2016) towed at distances of 350 m and 505 m behind a 50-m horizontal electric dipole (HED) source (Sinha et al., 1990). The resulting 2D resistivity models computed with MARE2DEM (Key, 2016) are interpreted and discussed by Haroon et al. (2018) and Gehrmann et al. (2019, 2020). In summary, the CSEM conductivity models highlight distinct regions of known SMS through increased electrical conductivity (cf. Fig. 3a). Here, we use the acquired CSEM data to reassess the electrical resistivity distributions at identified high-priority sites. Electrical conductance illustrated along CSEM profiles qualitatively define spatial extents of conductive, possibly mineral-enriched seafloor, and predict if morphological expressions are associated with hydrothermal (conductive) or volcanogenic (resistive) activity. Notably, Gehrmann et al. (2020) demonstrate that navigational uncertainty of the CSEM system is not trivial in such complex bathymetry, which can result in inversion artifacts. The authors mitigate these artifacts by estimating the data quality on navigational uncertainties such as instrument position with respect to the bathymetry. Here, we further reduce potential inversion artifacts by confining CSEM inversion models to distinct regions associated with mounds, mitigating potential misinterpretations caused by over-fitting the data at irrelevant locations.

2.3.2 Transient Electromagnetic Measurements

Marine transient electromagnetic (TEM) data were acquired at two specific sites within the TAG hydrothermal field using GEOMAR’s MARTEMIS system (Fig. 3b and 3c). This coincident loop
system consists of one transmitter and one receiver loop, which are housed in a \( 6.3 \times 6.3 \, \text{m}^2 \) frame. In regions where seafloor conductivity exceeds the water column conductivity, TEM data will exhibit an increased induced voltage allowing a localized inference of SMS distributions. The system is towed at \(<1 \, \text{kn}\) and 5 – 15 m above the seafloor and records the electromagnetic response of a 50% duty-cycle transmitter signal at 10 kHz. The acquired time series are processed considering the distorting effects described by Reeck et al. (2020) and transformed into full-space apparent conductivity curves following Eq. 1 of Haroon et al. (2018). Regions of increased apparent conductivity are generally associated with areas where the seafloor is more conductive than the water column resistivity (\( \rho < 0.3 \, \Omega \text{m} \)), which in our geological setting is indicative of SMS occurrences (Swidinsky et al. 2012). MARTEMIS positioning was computed through an ultra-short baseline (USBL) transponder attached to the tow cable and merged with the processed apparent conductivity data. The spacing between adjacent stations is approximately 10 m (Fig. 3b and 3c).

**2.4 In-situ Data**

Overall, 33 gravity cores of max. 3-m length were acquired within the TAG hydrothermal region during the M127 cruise (Petersen et al., 2016). Locations of possible coring sites were selected with the help of the high-resolution AUV bathymetry data. Twenty-three cores contained abundant sediment, eight contained only fragments of gravel, basalt and traces of sediments in the core catcher, and two were empty (Fig. 3, white markers). Among the 23 sediment cores, 10 had visible hydrothermally-influenced indications (Fig. 3, green triangles); the other cores had the visual appearance of background sediments (carbonate ooze) or showed layers of volcanic origin (Fig. 3, red and blue triangles, respectively). Note that the presence of background sediment or empty cores does not rule out hydrothermal activity at greater depth, as penetration of this coring technique was limited to a maximum of three meters.

In addition to gravity cores, rock drill samples have been drilled to a maximum depth of 12.5 m below seafloor (Murton et al., 2019). The obtained samples from the Southern, Rona and MIR mounds show high concentrations of minerals, confirming the hydrothermal origin of these three mounds. Here, we link core sites indicative of hydrothermal alteration with collocated EM resistivity data and models to distinguish spatial extents of mineral zones on these mounds and reassess tonnage estimates.
3 Methods

The workflow is split into four steps as illustrated in Fig. 4: 1) selecting and preparing suitable mid-ocean ridge (MOR) bathymetry data, e.g. from accessible open-source data repositories (Tab. A1), 2) training, validating and testing the U-Net model, 3) post-processing of the model output to derive mound architectures and integrate with concurrent RTP magnetic data, and 4) classification and geophysical analysis of identified mounds. The workflow is scripted in Python (Ver. 3.8.12) and uses the Tensorflow (Ver. 2.4.1) library for machine learning tools.

Figure 4 here!

3.1. Data Preparation: Bathymetry Data

Bathymetry data used for training are identified as suitable, if a large spatial coverage is acquired at either MORs or at specific hydrothermal fields. Bathymetry rasters are subdivided into overlapping patches of 256 x 256 pixels with a step length of 128 pixels. These pixels are manually annotated using a binary representation, where pixels associated with mounds are labeled as True and all other pixels are labeled as False. In total, 1899 mounds were manually annotated using the bathymetry data listed in Tab. A1.

To appear in a common standard that highlights rounded convex and concave morphology through distinct representations, we use a multi-directional slope analysis by mapping the normalized aspect, slope, and the ∂y derivative onto Red, Green and Blue channels of a standardized RGB image, respectively (Fig. 5). The chosen approach converting certain derivatives of bathymetry data into single RGB images facilitates a generalized visual interpretation and aids the model performance. All resulting images show the north flank of mounds in yellow to green moving west to east. Southern mound flanks appear white to blue. Concave features such as pits appear in a reversed manner.

Note that directional dependencies of background features in the preprocessed images remain unaltered through pre-processing (Figure 5). To increase the amount of training data and mitigate learning of directional dependencies in various settings, input bathymetry was augmented by means of a 90° rotation. As mound structures are near-circular structures, they remain rotational invariant although background strike differs (cf. Fig. 5a and b). In total, 2280 RGB images were produced to train, validate and test the U-Net model, each consisting of 65,536 pixels.

Figure 5 here!
3.2 U-Net Implementation, Training and Evaluation

The model architecture yields an end-to-end trainable neural network including segmentation of input images into partitioned pixel sets of corresponding classes. This type of network was first introduced for biomedical image segmentation by Ronneberger et al. (2015) and resembles a symmetric “U” (Fig. 6). In our specific case, the U-Net model distinguishes mound from background features and provides values of probabilities (Hu et al., 2015) as outputs.

For training and testing, only images with at least two percent of the pixels annotated as mounds were considered. Of these, 75 percent were used for training, 20 percent for validation and 5 percent for testing. We use the binary cross entropy loss function defined as:

\[ H_p(q) = \frac{1}{N} \sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i)) \]

where \( y_i \) refers to the corresponding binary label of each pixel and \( p(y_i) \) to the predicted value between 0 and 1 within each Epoch \( i \) of training. The accuracy, true positive and false negative metrics were computed to determine a point of early stopping, i.e. a model with sufficient accuracy and minimal over-fitting (Fig. 7). The model outputs are contoured at values of >0.5 to outline lateral mound dimensions of the mound pedestal.

After training, the pre-processed AUV bathymetry image from the TAG area (Petersen, 2016) is presented to the U-Net model as overlapping patches of 256 x 256 pixels. The image reconstruction process is illustrated in Fig. 8. To prevent inaccurate predictions along the image edges, an image smoothing considers only the central 128 x 128 pixels. The outer edges of each predicted patch are neglected. Using an overlapping process, each pixel of the bathymetry grid is included four times within the output prediction. The maximum probability from the four predictions is used as the final pixel probability.

3.3. Post-processing of mound structures

The output mound contours define the location and lateral footprints of each mound, which is calculated (in m\(^2\)). A minimum threshold of ~1040 m\(^2\) mound footprint (290 pixels) is introduced to remove most of the falsely detected mound structures caused by geological noise and to focus the analysis on potentially significant SMS volumes (cf. Murton et al. (2019) for discussion). From each of the mound contours, we compute the lateral footprint, maximum height, and median slope using only
pixels located within each mound contour. These parameters are used to describe the general mound architecture and are used as inputs for classification.

For integrating the magnetic anomaly data with the detected mounds from the UNet, an image overlay of gray-scaled hill shade and a diverging red-to-blue magnetic anomaly map (Fig. 2) is cropped and centered around each mound contour, including also peripheral areas. Using the specific red-blue color representation, RTP anomalies appear either red (if positive) or blue (if negative), both being primary color channels within an RGB color spectrum. Color histograms and correlations of the three RGB channels depict positive, negative or a mixture of RTP anomalies into three single values, depending on whether the image is blue, red, or a blue/red blend. The channel correlations serve as inputs for the subsequent classification of the mounds.

3.4. Classification & Evaluation

We applied spectral clustering using the Sci-Kit Learn Python Library on the derived parameters for each mound contour. Where available, we added peripheral SMS indicators derived from gravity cores, electromagnetic data and known SMS edifices to determine mound evolution and assess the mineral potential at confirmed high-priority sites. This ensemble of derived morphological expressions and geological/geophysical characteristics was integrated into a new model for the formation of SMS mounds in the TAG hydrothermal field. Further, it provides the basis to re-discuss the resource potential of the field.

4 Results

U-Net analysis

The U-Net metrics indicate an optimal point of early stopping at around Epoch 158 (Fig. 7). There, the network reached a prediction accuracy of greater than 98.6 and 97.8 percent in training and validation, respectively. The training loss reached 0.032 and the validation loss 0.075 (green and black curves, respectively, in Fig. 7a) using a learning rate of $10^{-4}$. A learning trend is observable in the ensuing epochs especially within the training data. Yet, the trend is less pronounced within the validation data, indicating that predictions will not improve for unseen images. To analyze the efficiency of the model, accuracy, true positives and false negatives are also utilized to understand the general characteristics of U-Net predictions (Fig. 7 b-d and Tab. 1).
Manually annotated mound structures make up, on average, less than 8.2 and 9.1 percent of the total pixels in each of the training and validation images, respectively. This significant imbalance compared to background leads to a high starting accuracy of approximately 91 percent, assuming that all pixels are predicted as background, i.e. False. Other metrics, such as true positives and false negatives are more significant for such imbalanced problems. At the point of early stopping, the trained network can retrieve 87 percent of the true positive pixels in the validation data, meaning that manually classified mound pixels are also classified as mound affiliated pixels by the U-Net. Similarly, false negatives are minimized to less than 1 percent of the total pixels per image during training and validation, indicating only few background formations are being falsely classified as mounds.

In addition to these training metrics, 114 images were used as a test dataset to further assess the U-Net’s efficiency and prediction characteristics for unseen data. A summary of the test data metrics is listed in Tab. 1. A prediction accuracy of >97.6 percent is achieved for the test data. Model efficiency in predicting true positives and avoiding false negatives is comparable to the validation data. Note that there is some bias to consider in the evaluation of these pixel-based metrics. Discrepancies in mound dimensions between manual annotation and automated segmentation will reduce model performance, although a mound is essentially detected by the network. In the majority of studied cases within the test data, mounds were detected and metric deficiencies arise from discrepancies in lateral mound extensions between manual and automated annotation (see Fig. S1 of supplementary materials).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Training</th>
<th>Validation</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>Benchmark: 0.91603 Prediction: 0.98696</td>
<td>Benchmark: 0.9103 Prediction: 0.97790</td>
<td>Benchmark: 0.89777 Prediction: 0.976976</td>
</tr>
<tr>
<td>True Positives</td>
<td>Benchmark: 0.08397 Prediction: 0.07659</td>
<td>91.2% Benchmark: 0.08970 Prediction: 0.07816</td>
<td>87.1% Benchmark: 0.10223 Prediction: 0.088996 87.1%</td>
</tr>
<tr>
<td>False Negatives</td>
<td>Prediction: 0.00738</td>
<td>Prediction: 0.01154</td>
<td>Prediction: 0.01323</td>
</tr>
</tbody>
</table>

The trained U-Net detects a total of 323 mounds within the mapped 49 km$^2$ of the AUV bathymetry data (Fig. 9), each with a lateral footprint greater than 1040 m$^2$ (= 290 pixel). The predictions include all previously identified SMS mounds (cf. Fig.1 and Fig. 9). The lateral mound dimensions match, in most cases, the manual annotation. However, the U-Net model underestimates the spatial footprint for Southern and Double Mound. Both mounds show a tectonized surface texture, deviating from an idealized mound shape, which may explain the reduced model performance. The total number of
known mounds accumulates to 16 (compared to 15 of Gehrmann et al. 2019) because Double Mound is segmented as two individual peaks by the U-Net classification.

The output mounds can be classified into three distinct clusters (Fig. 9) using the elbow method applied during spectral clustering (K-means). Table 2 lists some statistics for each cluster, including the minimum, maximum, and mean mound dimensions. The number of associated known SMS sites defines the priority of each cluster to host SMS. Cluster 1 is assigned a ‘high’ priority, as 10 out of 98 mounds are known to host SMS. A ‘low’ SMS priority is assigned to cluster 3 (1 known SMS site out of 114 mounds), and ‘medium’ to cluster 2 (4 known SMS sites out of 111 mounds).

Table 2: Clustering statistics of output mounds, including the total area of all mounds and associated mound dimensions. The number of known SMS sites associated with each cluster defines the SMS priority of the cluster.

<table>
<thead>
<tr>
<th>Cluster # (SMS priority)</th>
<th>Number of Mounds</th>
<th>Number of known SMS sites per Cluster</th>
<th>Total Area of all Mounds</th>
<th>Footprint</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (high)</td>
<td>98</td>
<td>10</td>
<td>982314 m²</td>
<td>Max: 141310 m² Min: 1052 m² Mean: 10024 m²</td>
<td>Max: 58.62 m Min: 1.35 m Mean: 14.46 m</td>
</tr>
<tr>
<td>2 (medium)</td>
<td>111</td>
<td>4</td>
<td>1188213 m²</td>
<td>Max: 78268 m² Min: 1205 m² Mean: 10704 m²</td>
<td>Max: 52.53 m Min: 0.35 m Mean: 14.98 m</td>
</tr>
<tr>
<td>3 (low)</td>
<td>114</td>
<td>1</td>
<td>942423 m²</td>
<td>Max: 75269 m² Min: 1094 m² Mean: 8266 m²</td>
<td>Max: 56.34 m Min: 1.12 m Mean: 13.62 m</td>
</tr>
<tr>
<td>Total</td>
<td>323</td>
<td>15</td>
<td>3112950 m²</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The clustering is primarily driven by the magnetic anomaly data (cf. Fig. S2 and S3); therefore, 10 of the 15 known SMS sites fall in Cluster 1, showing a distinct negative magnetic anomaly. At these 10 mounds, Sztikar et al. (2019) interprets the magnetic anomalies to represent a vertical hydrothermal conduit centered above their corresponding source. Known SMS sites of Cluster 2 (4) show both positive and negative RTP magnetic anomalies indicating either geological alteration or poor AUV navigation. Cluster 3 contains only one previously known SMS site (Mound #24 from Graber et al. 2020) that is associated with a positive magnetic anomaly. Images of all clustered mounds are displayed in Fig. S2 of the supplementary materials.
Despite their magnetic signature, mound morphology is similar in all clusters with respect to their mean heights and mound footprints (cf. Tab. 2 and Fig. S3 in the supplementary materials). Other morphological features indicative of hydrothermal activity, such as jagged contours and number of peaks (Jamieson et al. 2014), could not be identified by our workflow as important parameters for differentiating mound evolution in the TAG area. This leaves the magnetic anomaly as the strongest spatial indicator in the available data set.

**Spatial Distribution of Morphological Features and Magnetic Footprint**

The high-priority sites (Cluster 1) occur spatially confined in three bands (labeled Southern, Central and Northern Band in Fig. 9). All bands strike NW-SE, roughly perpendicular to the axis of the Mid-Atlantic Ridge, and coincide well with interfaces between different structural domains identified by Graber et al. (2020). The southern band lies within a region of oblique faults and fissures (Fig. 8 of Graber et al. 2020), that have been suggested to promote upward migration of hot fluids at TAG and other regions (Anderson et al. 2015). The central band of mounds is located within the NW-SE extension of the, so-called, Three-mound area (cf. Graber et al., 2020), runs parallel to mapped corrugations, and connects the Three-Mound regions to the MIR zone (Fig. 9). The northern band lies north of a zone with chaotic seafloor morphology and positive magnetic signature, separating the smooth bathymetry and negative anomalies of the central and northern bands.

The alignment of Cluster 1 mounds is interrupted in several locations. In the southern band, at around 26.138°N and 44.837°W, Cluster 1 mounds are not associated with oblique fissures mapped by (Graber et al. 2020). Instead, negative magnetic anomalies correlate with ‘fresh’ pillow mounds. Such potentially younger magmatic features may mask the oblique fissures typical for the southern band. The northern band contains only one known hydrothermal site named Shimmering, but multiple gravity cores indicate an abundance of hydrothermally-altered sediments in the area (cf. Fig. 9). However, mounds located at the western section of the northern band are structurally interpreted as pillow mounds (Graber et al. 2020).

**Analysis using Electromagnetic Data**

SMS potentials have often assumed homogeneously distributed metal grades across a mound morphological footprint and its corresponding stockwork zone. Although tonnage estimates are often based on in-situ measurements (i.e. core-log data/seafloor drilling), derived mineral potentials are
likely too optimistic because in-situ data are 1) available at only few representative mounds, 2) generally obtained at the points of highest interest, i.e. where SMS is apparent in seafloor imagery and 3) penetrate only few meters into the subsurface. In the majority of cases, the structural heterogeneity of individual mounds is either neglected or considered too simplistic, assuming that SMS with high metal grades are distributed across the entire lateral extent of each mound. To constrain this rather generalized assumption, resistivity models derived from CSEM or TEM data appear more suitable to understand the degree of mineralization away from the point-scale core-log data.

Here, we focus on the electromagnetic data acquired only along high-priority mounds (Cluster 1) or verified mounds that have been documented in preceding literature. CSEM resistivity models that were recomputed for Cluster 1 mounds are illustrated in Fig. 10 using data acquired by Gerhmann et al. (2019). Note that only those mounds are considered that are intersected by CSEM transects. Regions of low resistivity are illustrated by red to orange coloring, background resistivity through green and blue coloring.

The CSEM resistivity models illustrate that the majority of investigated Cluster 1 mounds are associated with a distinct low-resistivity anomaly of variable magnitude. This, together with core-log data and grab samples (Peterson et al. (2016); Murton et al. (2019)) confirms a certain degree of mineralization at each of the prospective sites. It needs to be noted that in a few instances, i.e. Fig. 10d and 10g, CSEM data has only limited resolution due to the large vertical offset between measurement system (denoted by black markers) and seafloor, thus, limiting a quantitative analysis of mineral grades using CSEM data. Laterally, resistivity distributions and amplitudes do indicate a high degree of certainty and illustrate characteristics of mound composition. Shinkai and Southern mounds, residing in the central band (Fig. 10c through 10e) exhibit a relatively homogeneous conductive structure. These two mounds appear as low-resistive anomalies across their entire lateral footprint (cf. apparent conductivity data in Fig. 3b) indicating that previous mineral estimates could be accurate. In contrast, Double (Fig. 10e) and Rona (Fig. 3b) mound exhibit a low resistive anomaly only in the vicinity of their peaks with no notable contrast to the background resistivity at their pedestal. Although, Rockdrill cores of up to 12 m acquired at the peak of Rona confirm high metal concentrations within a sulfide layer (Murton et al., 2019), TEM data indicates that the majority of Rona’s volume is of lower economic value due to a missing apparent conductivity anomaly (Fig. 3b).
MIR mound has the largest spatial footprint of Cluster 1 mounds in the study area (Fig. 10f). Previous predictions based on mound volumes and extrapolated metal grades derived from gravity cores and rock drill data suggest MIR to be the most economical site within the TAG hydrothermal field (Graber et al. 2020). CSEM inversion of MIR illustrated in Fig. 10f shows an irregular distribution of low-resistive zones across the mound transect indicating the presence of mineralized sediments, but not in the quantity suggested by previous studies. The higher resolution TEM data acquired along multiple transects across the MIR mound support this hypothesis (Fig. 3c) contradicting the notion of MIRs high mineral potential. The point-scale gravity core and Rockdrill data (Peterson et al. (2016); Murton et al. (2019)) were acquired in the northwestern region of the MIR contour, where high apparent conductivities exist. Most of the other regions of the mound structure are not associated with a distinct resistivity anomaly, thus, indicating that mineralization is irregularly distributed across MIR and that tonnage predictions for MIR may be significantly overestimated.

TAG is arguably the most prominent mound in the study area. Multiple geophysical and geological surveys have focused on the internal mound structure, including the Ocean Drilling Program (ODP) Leg 158 experiment (e.g. Humphris et al., 1995). As such, the internal structure of TAG is well constrained and serves as a blueprint for estimating mineral potentials for other mounds, where less knowledge about internal structure and data are available. The CSEM resistivity models of TAG (Fig. 10g and 10h) show an E-W and N-S transect crossing the mound, respectively. Unfortunately, navigation during the E-W transect was chosen too conservative with towing altitudes exceeding 100 m above seafloor, which resulted in inadequate resolution for the TAG’s resistivity structure. However, the N-S profile (Fig. 10h) is intriguing, as it supports the asymmetric distribution of mineralization presented in previous studies (e.g Galley et al. 2020). Moreover, the CSEM resistivity model indicates that a significant resistivity contrast compared to the background basalt may only exists for the massive pyrite, pyrite-anhydrite, pyrite-silica and possibly the pyrite silica units (see Knott et al. 1998). Hence, the applied CSEM configuration is likely unsuitable for detecting the stockwork structure and requires higher resolution CSEM data acquired in a 3D survey as demonstrated by MacGregor et al. (2021).

5 Discussion

5.1 Automated SMS Mapping using Machine Learning

The presented workflow can be used as a blueprint for prioritizing SMS exploration targets at mid-ocean ridges and understanding distributions of mineral potentials. The workflow reduces the total area of interest from the surveyed 49 km^2 to 3.1 km^2, which in turn can be further reduced to either 1.92 km^2
(Clusters 1 and 2) or 0.98 km² (only Cluster 1) using additional magnetic constraints (i.e. mounds with magnetic lows). Moreover, the latitudinal bands of hydrothermal activity identified through this integrated analysis reveal prospective areas where to search for SMS and may also exist at other MORs. The workflow is fully automated, allowing us to identify regions of interest in quasi real-time (if a pre-trained model exists) when new data is acquired, thus, reducing exploration costs considerably and permitting more focused surveying. Additionally, the workflow is adaptable to future developments in marine mineral exploration and research and its application in other survey areas with similar or additional data layers appears feasible.

The data preparation part of our workflow seeks to unify the bathymetry data, acquired at different spatial resolution and at various sites across the globe, into a common representation independent of the actual depth, slope, and curvature within a given area. Juliani & Juliani (2021) propose a principal component analysis (PCA) consisting of both a change in slope and a multi-directional shading of elevation data in order to reduce the bathymetric inputs. However, our analysis shows that PCA may not generalize well for bathymetry data acquired at different regions with variable roughness and geological strike. A key advantage of our proposed processing scheme is notably that bathymetry data from different regions will unify onto a single coherent visualization.

Moreover, the workflow also allows to test and use other types of ML segmentation tools, and to include additional data layers (e.g. high-resolution backscatter, self-potential, etc). As many of these additional data layers are currently not available in open-access repositories, they can be integrated best within the post-processing step. As backscatter and self-potential data become more readily available, it is also feasible to train the U-Net directly for different types of mound characteristics. The integration of such additional data will likely increase ambiguity, but presumably achieve higher certainty in identifying SMS sites.

Following Szitkar et al. (2019) and Rona (1978; 1980), hydrothermal mounds within the TAG hydrothermal field are associated with a distinct negative RTP magnetic anomaly, whereas volcanic edifices typically display positive values. This characteristic proves suitable for clustering depicted mound contours into groups to identify their potential origins. However, three aspects must be considered to integrate the magnetic footprint of a mound or a group of adjacent mounds with the corresponding bathymetry attributes.
1. Magnetic anomaly data are acquired at a resolution of 10 m grid spacing compared to the 2 m resolution of the bathymetry data. A pixel-wise comparison between magnetic anomalies and morphological features requires an up-sampling of the magnetic anomaly data, which may lead to interpolation artifacts.

2. An RTP magnetic anomaly shows magnetization and geomagnetic field vectors that render vertical anomalies above the causative body. However, Szitkar et al. (2019) discuss that tectonic events can tilt the crustal block causing altered shapes of the magnetic anomalies, leading to incoherent magnetic anomalies associated with morphological expression. A similar effect is also observable if tectonic forces act on a previously deposited mound.

3. AUV magnetic data are susceptible to errors that arise from inaccuracies in AUV positioning relative to its calibrated coordinates. This may lead to a lateral shift between the morphological expression and the corresponding RTP magnetic anomaly.

All three undetermined circumstances lead to uncertainties in a pixel-wise integration of the magnetic and corresponding bathymetry pixels. Hence, a relaxation of spatial similarities between mound structures and resulting magnetic anomalies is required and was implemented in our analysis.

The approach would benefit from a greater number of annotated bathymetry data available in online repositories to improve model training. The chosen study area belongs to the most studied hydrothermal sites globally, and, provides a solid first training set. The developed workflow focuses on the analysis of high-resolution bathymetry data, which resemble the most common collected data in seafloor exploration. Given the steady increase of sea-going SMS research, manual assessments of each individual data layer becomes increasingly difficult and automatization of workflows will be inevitable. Therefore, application of the workflow in other hydrothermal areas either at mid-ocean ridges or other geological environments with complex, rough terrain, is a crucial future task.

5.2 Implications for hydrothermal activity in the TAG area

Notably, marine mineral exploration is a complex endeavor unlikely solved by a silver bullet approach. Thus, various ML strategies and conventional geoscientific research will attest a feasibility for detecting SMS and estimating mineral potentials. The proposed ML strategy does not contradict this notion, but instead, offers a means of integrating multivariate data into a common interpretation scheme that is easily audited. Note that the delineation of convex structures in bathymetry data underlies some variability resulting from terrain analysis and, even if conducted manually, remains ambiguous due to
interfering geomorphic processes that mask or distort typical mound morphologies. Hence, although
mapping the correct mound dimensions is significant for addressing mineral potentials, discrepancies
between manual and automated segmentation are expected.

Despite these explainable deviations, the spatial alignment of Cluster 1 bands is clearly visible and the
correlation to structural domains defined by Graber et al. (2020) is apparent. This may support the
hypotheses of a structural heterogeneity within the hydrothermal field dictating the distribution of SMS
edifices, as proposed by Graber et al. (2020). Both the spatial extent of the alignment and the
occurrence of deviating areas indicate a structural constraint in the deep subsurface. This dominant
structural constraint supports an interpretation where a strongly distorted subsurface structuring (e.g.
bend detachment fault) leads to focused fluid flow in the deep subsurface that results in linear, off-axis,
distribution of hydrothermal edifices. Moreover, the various upflow zones are likely to span a region
larger than the investigated area of study. Further sites may be located north of Shimmering and also
south of TAG, as well as in the westward extension of the three bands. Conclusively, the presented
workflow has demonstrated a successful amalgamation of spatially acquired bathymetry and magnetic
data, which could be used to inform future AUV bathymetry and magnetic surveying.

Although the number of potential SMS sites drastically increased through this automated analysis,
mineral potential of the TAG hydrothermal field is likely lower than originally presumed. Electromagnetic data illustrates that mineralized zones for the largest proposed SMS sites are generally
heterogeneously distributed across the mound contour, thus, contradicting the proposed high tonnage
estimates using homogeneously distributed mineral grades derived from point-scale measurements. We
propose that future analysis of SMS tonnages should incorporate multiple seafloor drillings conducted
across the mound contours with additional data layers (e.g. backscatter, seismics, and 3D CSEM
inversion models) to achieve a high degree of certainty in the tonnage estimation. If such high-
resolution survey strategy for SMS sites is economical, remains beyond the scope of this study.

The current SMS priority depends mainly on AUV magnetic data and on the number of known SMS
sites within a given cluster. The former is not typically considered a conventional data layer in SMS
exploration and should be added to the necessary SMS exploration criteria. Furthermore, it cannot be
ruled out that future endeavors may potentially change the presented prioritization of SMS mounds
through more SMS discoveries or through the acquisition of additional data layers (e.g. high-resolution
backscatter, resistivity or self-potential data). It is expected that not all mounds within the high-priority
cluster are associated with SMS edifices and that additional data layers will improve the certainty of the automated SMS prediction. Overall, a more diversified dataset measured at numerous SMS sites across the global MORs will only help improve our understanding of SMS predictors and improve future developments of ML workflows.

Marine CSEM and TEM data have demonstrated additional value when conducting tonnage estimates for SMS sites. However, more development is required to improve the significance of resistivity models to help quantifying volumes of mineralized zones within a mound edifice. Most notably, high-resolution 3D surveys using AUV technology are likely required to accurately derive spatial extensions of conductive material in these remote settings. MacGregor et al. (2021) have already presented a first 3D application and inversion of CSEM data and others are likely to follow suit, given the increased value of electrical resistivity models to constrain volumetric predictions.

6 Conclusion

A workflow to conduct automated SMS site detection using multivariate geoscientific data is presented that employs a U-Net neural network to identify prominent mound-like morphologies in bathymetry data. The predicted contours are subsequently integrated with other spatial data layers (e.g. AUV magnetic data) to identify high-priority sites for SMS prospecting. Within the 49 km² grid of high-resolution bathymetry data, 323 mounds were detected. 98 of these were classified as high priority due to their architecture and magnetic signature. Moreover, from the automated analysis 14 (10 high, 4 medium) of the 15 known SMS sites in the TAG area were identified as either high or medium priority. Only one known site was classified within the low-priority group. The high-priority sites were spatially distributed into latitudinal bands, supporting the hypotheses that focused fluid-flow at depth leads to linear distributions of off-axis SMS edifices in the area.

The presented workflow cannot only be used to improve analysis and interpretation of previously surveyed areas, but also serve as a blueprint to optimize SMS exploration at sea. The trained model can be applied on newly acquired bathymetry data in quasi real-time to determine prospective zones for more detailed confirmation/visualization studies. Thus, optimizing the use of ship time and reducing exploration costs. The workflow is very adaptable to include additional data layer such as backscatter, self-potential, turbidity and other water column data maps if available.
Electrical resistivity models demonstrate that mineralization of SMS mounds are less homogeneous than often considered. Thus, indicating that high-grade mineral contents in SMS are not equally distributed across the entire mound and that tonnage estimates may be significantly overestimated. Consequently, although the workflow detects many more potential SMS edifices than previously known, the overall resource potential of the TAG hydrothermal field is likely lower than previously assumed.

7 Acknowledgments

AH was in part funded by the Digital Earth and SMART Projects at GEOMAR Helmholtz Centre of Ocean Research Kiel.

8 References


Source Electromagnetics at the TAG Hydrothermal Field, 26°N Mid-Atlantic Ridge.


Villinger, Heinrich; Strack, Anne; Gaide, Stefanie; Thal, Janis (2018). Gridded bathymetry of North Pond (MAR) from multibeam echosounder EM120 and EM122 data of cruises MSM20/5 (2012) and MSM37 (2014). Department of Geosciences, Bremen University, PANGAEA, https://doi.org/10.1594/PANGAEA.889439.
**Table A1: List of open-access bathymetry data used for training the U-Net.**

<table>
<thead>
<tr>
<th>Index</th>
<th>Longitude Min.</th>
<th>Longitude Max.</th>
<th>Latitude Min.</th>
<th>Latitude Max</th>
<th>Resolution</th>
<th>Source / Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-44.99625</td>
<td>-44.69042</td>
<td>25.98625</td>
<td>26.28625</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>2</td>
<td>-43.23542</td>
<td>-41.31541</td>
<td>29.69375</td>
<td>30.78042</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>3</td>
<td>-34.45958</td>
<td>-33.12792</td>
<td>35.45792</td>
<td>37.50708</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>4</td>
<td>-27.14292</td>
<td>-23.76958</td>
<td>61.70625</td>
<td>63.44625</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>5</td>
<td>-45.21799</td>
<td>-44.40485</td>
<td>25.77404</td>
<td>26.54754</td>
<td>~ 30 m</td>
<td>Petersen (2019)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://doi.org/10.1594/PANGAEA.899415">https://doi.org/10.1594/PANGAEA.899415</a></td>
</tr>
<tr>
<td>6</td>
<td>-44.85162</td>
<td>-44.75758</td>
<td>26.10058</td>
<td>26.18684</td>
<td>~ 2 m</td>
<td>Petersen (2019)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://doi.org/10.1594/PANGAEA.899415">https://doi.org/10.1594/PANGAEA.899415</a></td>
</tr>
<tr>
<td>7</td>
<td>-44.97116</td>
<td>-44.86794</td>
<td>13.27914</td>
<td>13.34156</td>
<td>~ 2 m</td>
<td>Escartin &amp; Petersen (2017) and Escartin et al. (2017)</td>
</tr>
<tr>
<td>8</td>
<td>-45.00966</td>
<td>-44.88124</td>
<td>13.48664</td>
<td>13.52226</td>
<td>~ 2 m</td>
<td>Escartin &amp; Petersen (2017) and Escartin et al. (2017)</td>
</tr>
<tr>
<td>9</td>
<td>-46.45929</td>
<td>-45.81298</td>
<td>22.36806</td>
<td>23.05779</td>
<td>~ 10 m</td>
<td>Villinger et al. (2018)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://doi.org/10.1594/PANGAEA.A.889439">https://doi.org/10.1594/PANGAEA.A.889439</a></td>
</tr>
<tr>
<td>10</td>
<td>-34.99333</td>
<td>-31.87833</td>
<td>56.69666</td>
<td>58.00000</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>11</td>
<td>-42.40119</td>
<td>-41.72925</td>
<td>29.82935</td>
<td>30.36156</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>12</td>
<td>-129.1490</td>
<td>-129.0070</td>
<td>47.88199</td>
<td>48.08900</td>
<td>~ 1 m</td>
<td>Clague et al. (2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="http://get.iedadata.org/doi/321990">http://get.iedadata.org/doi/321990</a></td>
</tr>
<tr>
<td>13</td>
<td>-177.1383</td>
<td>-176.3533</td>
<td>-23.0983</td>
<td>-21.5792</td>
<td>~ 90 m</td>
<td>Multibeam Mosaic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://www.ncei.noaa.gov/maps/bathymetry/">https://www.ncei.noaa.gov/maps/bathymetry/</a></td>
</tr>
<tr>
<td>14</td>
<td>-29.93012</td>
<td>-22.78952</td>
<td>35.99997</td>
<td>39.50023</td>
<td>~ 45 m</td>
<td>Hübscher &amp; Beier (2022)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.945528">https://doi.pangaea.de/10.1594/PANGAEA.945528</a></td>
</tr>
</tbody>
</table>
Figure 1: Bathymetry maps of the study area. (a) GEBCO bathymetry (shaded map) overlain by the ship-based bathymetry data acquired during RV Meteor cruise M127 with a spatial resolution of 30 m. Gray outlines denote visible mound structures whereas the black outlined region denotes the high-resolution bathymetry survey illustrated in (b). (b) AUV bathymetry data acquired with a spatial resolution of 2 m using the same color scale as in (a). Known SMS mounds are outlined and labeled as depicted by Graber et al. (2020).
Figure 2: Overlay map of the hillshade bathymetry (2 m resolution) and the magnetic anomaly map with 10 m spatial resolution from Pedersen (2016). Outlined in black are the known SMS mounds depicted by Graber et al. (2020).
Figure 3: (a) Hillshade map of the bathymetry data with a spatial resolution of 2 m. Computed electrical conductance values derived from 2D CSEM resistivity models of Gehrmann et al. 2019 are displayed with color-coded markers. Light colors denote a low and hot colors a high conductance. (b) Zoom-in of the Three-Mound region overlain by the transformed apparent conductivity values obtained by TEM measurements. (c) Same as (b) but for the MIR zone. Outlines of the mounds denote the manually-annotated lateral mound dimensions from Graber et al. (2020). Triangular markers in (a) through (c) illustrate the locations of the 3 m gravity cores and the lithology observed within the core samples (Petersen, 2016).
Figure 4: Schematic of the applied workflow including all relevant steps applied, i.e. data preparation, U-Net implementation, post-processing and classification.
Figure 5: Image unification example of relevant bathymetry features into a common visual representation that is generically applicable to coherently address bathymetry data acquired at different regions across the globe. The aspect, slope, and $\partial y$ derivative of the bathymetry are mapped onto the red, green and blue channels of a standardized 0-255 RGB image (left column of each panel). In this representation, mounds appear directionally invariant with coherent color representation. (a) The original input data and (b) the original input data rotated by 90°. Contrarily, the background bathymetry differs based on the predominant strike direction of the seafloor morphology, whereas prominent mound features remain rotational invariant.
Figure 6: Schematic of the U-Net architecture used for semantic segmentation of the bathymetry data (modified after Ronneberger et al., 2015).
Figure 7: (a) Binary cross-entropy loss function used during training (green) and validation (black). Additional metrics, i.e. (b) accuracy, (c) true positives and (d) false negatives are also used to assess the model performance. Note that true positives and false negatives are normalized to represent percentages of pixels per image. The vertical dotted line denotes the point of early stopping whereas the horizontally dashed lines in (c) represent the average number of pixels affiliated with mound structure within training and validation data.
Figure 8: Workflow of mound prediction for the AUV bathymetry raster. The pre-processed AUV bathymetry map is cropped into overlapping 256 x 256 patches, which are presented to the U-Net for prediction. The output is smoothed using displayed window where pixels within the black region are neglected due to edge effects that deteriorate predictions (as observed within the testing phase). The right panel shows the final prediction map that is utilized for further processing.
Figure 9: Bathymetry map containing the 323 mounds with a lateral of footprint greater than 290 pixels ($1040 \text{ m}^2$) as predicted by the U-Net. The mounds are illustrated by white contours and are classified as either Cluster 1 (high priority), Cluster 2 (medium priority) or Cluster 3 (low priority). Dotted gray lines illustrate the interpreted boundaries of the three latitudinal bands containing hydrothermal edifices.
Figure 10: (a) Hillshade bathymetry data with classified mounds as illustrated in Fig. 9. Panels (b) through (i) show electrical resistivity models for each of the high priority mounds intersecting a CSEM profile. The lateral extent of each profile is illustrated by red lines in (a). In (b) through (i), red color shading indicates low resistivity (mineralization), whereas green/blue are more resistive background basalts. Black markers denote transmitter positions along profile.