

# **Deep-sea polymetallic nodules spatial modelling with machine learning algorithms and benthic plume monitoring**

## **Dissertation**

in fulfilment of the requirements for the degree of  
Doctorate of Natural Sciences (Dr. rer. nat.)  
of the Faculty of Mathematics and Natural Sciences  
at Christian-Albrecht University of Kiel

submitted by

**Iason Zois Gazis (MSc)**

Kiel (DE), 2025

First examiner: Prof. Dr. rer. nat. Jens Greinert

Second examiner: Prof. Dr.-Ing. Kevin Köser

Date of oral examination: 16.12.2024

## **Sworn Declaration**

I hereby confirm that apart from the supervisor's guidance, the content and design of the thesis is all the doctoral researcher's own work and only using the sources listed. The thesis has not been submitted either partially or wholly as part of a doctoral examination procedure to another examining body and it has not been published or submitted for publication. The thesis has been prepared subject to the Rules of Good Scientific Practice of the German Research Foundation, and I have never withdrawn an academic degree.

Kiel, 08.10.2024

Iason Zois Gazis (MSc)

## Acknowledgements

An African proverb says, "*If you want to go fast, go alone. If you want to go far, go together*". This phrase describes my PhD "trip", a life-changing experience that helped me (and forced me) to grow scientifically, mentally and emotionally. The DSM group is a vivid scientific "hive" that welcomed me from the very first moment. Since then, my colleagues and I spent countless hours together on scientific expeditions, labs, and offices, working as a team to accomplish our missions. Most times, it took us more than we expected, but the results always matched or exceeded our expectations: *DSM colleagues (former & current)*, thank you for our creative discussions, the valuable tools we learned together and your kind support inside and outside GEOMAR. Probably the biggest thanks goes to my Doctoral Father, *Prof. Dr. Jens Greinert*, who "took me under his wings" and provided the means to conduct my research seamlessly until today. Jens saw what I could achieve (even before I realised it) and pushed me gently towards it, providing me with the space, time and trust needed to grow as a scientist. Together, we formulate ideas and research papers that share our common passion for quality, innovation, applicability and interdisciplinarity. The PhD research sometimes challenges the work-life balance, and in these moments of stress, chatting and meeting with *good friends* (here and abroad) is probably the best recipe for recharging my batteries. A big thanks to all of you. A special thanks goes to my family, who always supports me in my steps abroad, although they miss me a lot. Finally, I am thankful to *Dr. Svetlana Kresova*, my beloved *wife*, who believed in me from the first moment we met here in Kiel. Sveta is there when I return happy from the office or expeditions having new research data and findings, but mainly when I return tired and frustrated after a long day without progress along the long but fruitful PhD journey.

## Preface

This dissertation has been submitted to the Faculty of Mathematics and Natural Sciences at Christian-Albrecht University of Kiel in compliance with the requirements for the *Doctorate of Natural Sciences (Dr. rer. nat.)* The research presented here was conducted entirely at the [DeepSea Monitoring Group](#) at GEOMAR Helmholtz Center for Ocean Research Kiel under the supervision of Prof. Dr. Jens Greinert.

The research topics are within the main objectives of the Program-Oriented Funding of Helmholtz Association of German Research Centers entitled '[Changing Earth – Sustaining our Future](#)' in the research field Earth and Environment 2021-2027. The research is part of Topic 6 (Marine and Polar Life: Sustaining Biodiversity, Biotic Interactions and Biogeochemical Functions) and in Topic 8 (Georesources for the Energy Transition and a High-Tech Society). The research was partly carried out in the framework of the European collaborative project [MarTERA-COMPASS](#)-Drimp, receiving national funding from the Federal Ministry for Economic Affairs and Climate Action (BMWi, grand no. 03SX466B) and the European collaborative project [MiningImpact 2](#) of the Joint Programming Initiative Healthy and Productive Seas and Oceans, receiving national funding from the German Ministry of Research (BMBF grant no. 03F0812A-H).

The dissertation is a collection of four published original peer-reviewed articles. All articles are open-access and licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). The research articles are based on fieldwork during offshore campaigns, acquisition and post-processing of hydroacoustic, optic and ground truth data, onshore spatial analysis, and visualization with the use of Geographic Information Systems (GIS), as well as computational analysis and predictive modelling using machine learning algorithms.

- Gazis, I.-Z.,** Schoening, T., Alevizos, E., & Greinert, J. (2018). Quantitative mapping and predictive modeling of Mn nodules' distribution from hydroacoustic and optical AUV data linked by random forests machine learning. *Biogeosciences*, *15*(23), 7347–7377. <https://doi.org/10.5194/bg-15-7347-2018>
- Gazis, I.-Z.,** & Greinert, J. (2021). Importance of spatial autocorrelation in machine learning modeling of polymetallic nodules, model uncertainty and transferability at local scale. *Minerals* (Basel, Switzerland), *11*(11), 1172. <https://doi.org/10.3390/min11111172>
- Gazis, I.-Z.,** Charlet, F., & Greinert, J. (2024). An interpretable multi-model machine learning approach for spatial mapping of deep-sea polymetallic nodule occurrences. *Natural Resources Research*. <https://doi.org/10.1007/s11053-024-10393-7>
- Gazis, I.-Z.,** Stigter H., Mohrmann J., Heger K., Diaz M., Gillard B., Matthias B., Veloso-Alarcón M., Purkiani K., Haeckel M., Vink A., Thomsen L., Greinert J. (2025). Monitoring benthic plumes, sediment redeposition and seafloor imprints caused by deep-sea polymetallic nodule mining. *Nature Communications* *16*, 1229 (2025). <https://doi.org/10.1038/s41467-025-56311-0>

# Contents

Abbreviations.....	7
Abstract.....	8
Zusammenfassung.....	9
Synopsis .....	10
Introduction .....	11
Section I ( <i>spatial modelling</i> ).....	14
Foreword.....	15
Nodules Spatial Distribution .....	16
Research Objectives .....	21
Future Work.....	58
Section II ( <i>plume monitoring</i> ).....	60
Foreword.....	61
Mining trial monitoring .....	61
Research objective.....	62
Future Work.....	68
Section III ( <i>scientific development</i> ).....	70
Co-authorship in Peer-Reviewed Scientific Publications.....	71
Data Publications .....	72
Scientific Expeditions.....	75
Conferences .....	77
Workshops & Summer Schools.....	80
Supervising students .....	81
Section IV ( <i>references</i> ) .....	82
References .....	83
Section V ( <i>research articles</i> ) .....	96
Published & Submitted Research Articles.....	96

## Abbreviations

AABW	Antarctic Bottom Water
ADCP	Acoustic Doppler Profiler
AOA	Area of Applicability
ARA	Angular Range Analysis
AUV	Autonomous Underwater Vehicle
BGR	Federal Institute for Geosciences and Natural Resources
CCD	Calcium Compensation Depth
CCZ	Clarion Clipperton Zone
CoMoNoD	Compact Morphology-based Nodule Delineation
CTD	Conductivity, Temperature, Depth sensor
CUBE	Combined Uncertainty and Bathymetry Estimator
CV	Cross-Validation
DEA	Disturbance and recolonization Experimental Area
DEM	Digital Elevation Model
DI	Dissimilarity Index
EEZ	Exclusive Economic Zone
GAM	Generalized Additive Models
GLM	Generalized Linear Models
GSR	Global Sea Mineral Resources
H <sup>2</sup> SOM	Hierarchically Growing Hyperbolic Self-Organizing Map
I	Moran Index of Spatial Autocorrelation
ISA	International Seabed Authority
LBL	Long Baseline (acoustic positioning system)
MAE	Mean Absolute Error
MBES	Multibeam Echosounder
NNM	Neural Networks Model
OBS	Optical Backscatter Sensor
p	Probability Value
PDP	Partial Dependence Plot
POC	Particulate Organic Carbon
QU	Quantile Uncertainty
R <sup>2</sup>	Coefficient of Determination
RFM	Random Forests Model
RMSE	Root Mean Square Error
ROV	Remotely Operated Vehicle
SBP	Sub-Bottom Profiler
SD	Standard Deviation
SE	Standard Error
SPM	Suspended Particulate Matter
SSS	Side-Scan Sonar
SVM	Support Vector Machines
TOC	Total Organic Content
USBL	Ultra-Short Baseline (acoustic positioning system)
VI	Variable Importance

## Abstract

This dissertation addresses two topics: a) deep-sea polymetallic nodules spatial modelling using machine learning and backscatter analysis, and b) mining-induced benthic plume monitoring. Deep-sea polymetallic nodules have a heterogeneous and spatially clustered distribution pattern within short distances (m to km). High-resolution data are needed to understand the reasons behind their patchy distribution, associate nodule coverage with benthic fauna occurrences, and enable accurate resource estimation and mining path planning. This study uses autonomous underwater vehicle (AUV), multibeam echosounder (MBES) data, seafloor images and photomosaics, providing a meter-to-millimeter spatial resolution and allowing detailed geomorphological analysis and accurate image-derived estimations of the polymetallic nodule seafloor coverage. Multi-model machine learning approaches (generalized linear models, generalized additive models, random forests, support vector machines, and neural networks) were used to link the two data types. They predict well ( $R^2 > 0.7$ ) for geographical areas where correlations between the response variable and predictors exist, the spatial autocorrelation is isolated from the model training, and an adequate feature space representation of the predicted space is guaranteed from the training data. Quasi-linear, monotonic and nonlinear relationships between the nodule coverage and spatial predictors were identified using interpretable machine learning tools such as multi-model variable importance and partial dependence plots (2D and 3D). The most important spatial predictors were the MBES backscatter, particularly from incident angles between  $25^\circ$  and  $55^\circ$ , bathymetry, slope, and slope orientation (eastness and northness), agreeing with the existing knowledge of nodule occurrence for larger areas.

Regarding benthic plume monitoring, the thesis shows that artificially generated sediment plumes initially spread as a high-density turbidity current perpendicular to the mining lanes, having little to no mixing with the ambient water. During the investigated mining experiment, the turbidity current was channelled through steeper seafloor sections ( $> 3^\circ$  slope), increasing the propagation distance and sediment mass transportation outside the mining site, a factor not considered so far in numerical models. After the weakening of the turbidity current dynamics (100 to 500 m, depending on the seafloor slope), a passive transfer of suspended particles by bottom currents continues to transport fine-grained particles. The suspended particulate matter concentration ( $\text{mg L}^{-1}$ ) decreases fast with distance and altitude, and values almost at background levels were recorded at the end of the monitored area at a 4.5 km distance and the highest monitoring altitude of 50 m. Particle flocculation was recorded in situ for the first time – another so far neglected factor, leading to fast (minutes to hours) sediment redeposition within short distances ( $< 500$  m) from the mining site. Within the mining site, mm-resolution seafloor photomosaic showed that the sediment redeposition was at least 3 cm, and the seafloor erosion was 5 cm. The study underlined the need to combine hydroacoustic and optic backscatter sensors on stable (landers) and moving (AUVs) platforms with seafloor images to achieve informative, high-resolution 4D plume monitoring.

## Zusammenfassung

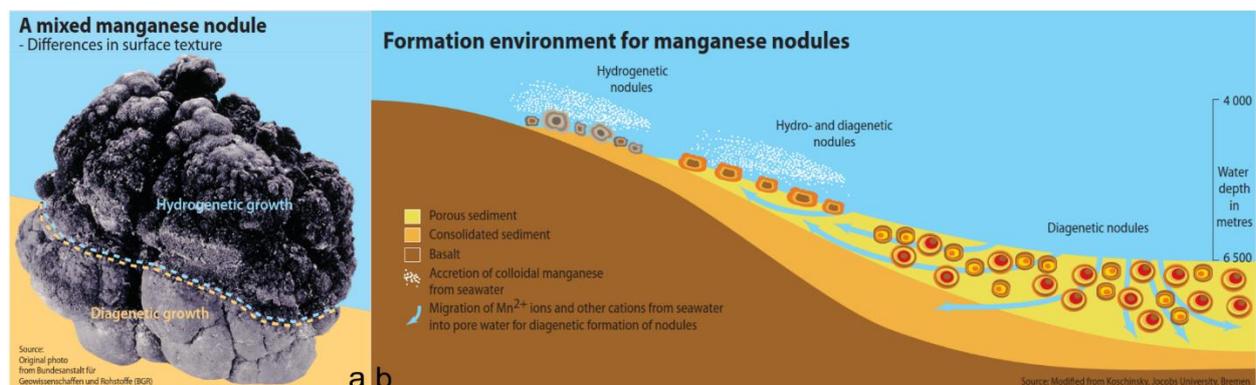
Diese Dissertation untersucht die räumliche Verbreitung von polymetallischen Knollen mittels Machine Learning Methoden, für deren Modellierung die Analyse der hydroakustischen Rückstreustärke wesentlich war. Weiter untersucht die Arbeit die Sedimentausbreitung im Rahmen eines Knollen-Kollektor Prototyp Tests, der 2021 als erster wirklich industrieller Test im zentralen Pazifik durchgeführt wurde. Polymetallische Knollen in der Tiefsee weisen auf kurzer Entfernungen (m bis km) eine heterogene und räumlich geclusterte Verteilung auf. Hochauflösende Daten sind erforderlich, um die Gründe für diese zum Teil lückenhafte Verteilung zu verstehen, die Knollenbedeckung mit der Verteilung benthischer Fauna zu korrelieren und eine genaue Ressourcenabschätzung und Abbaupfadplanung zu ermöglichen. In dieser Arbeit werden AUV MBES, Meeresbodenbilder und Fotomosaik verwendet, die eine räumliche Auflösung von einem Meter (Hydroakustik) bis zu einem Millimeter bieten (Optik). Hierdurch ist eine detaillierte geomorphologische Analyse sowie eine genaue, Quantifizierung der Knollenbedeckung am Meeresboden möglich. Die beiden Datensätze wurden dafür durch verschiedenen ML Ansätze (Generalized Linear Models, Generalized Additive Models, Random Forests Model, Support Vector Machines, and Neural Networks Model) miteinander kombiniert und deren Resultate verglichen. Alle Modelle machen gute Vorhersagen ( $R^2 > 0,7$ ) für geografische Gebiete, in denen Korrelationen zwischen abhängigen Variable und den Prädiktoren bestehen, die räumliche Autokorrelation vom Modeltraining isoliert ist und genügend Trainingsdaten für den gesamte Feature Space des vorherzusagenden Gebietes gegeben sind. Quasi-lineare, monotone und nicht-lineare Beziehungen zwischen der Knollenbedeckung und den räumlichen Prädiktoren wurden z.B. über die Variablen-Wichtigkeit aus mehreren Modellen und partiellen Abhängigkeitsdiagrammen (2D und 3D) ermittelt. Die wichtigsten räumlichen Prädiktoren waren die MBES-Rückstreuung (insbesondere für Einfallswinkeln zwischen  $25^\circ$  und  $55^\circ$ ), die Bathymetrie, die Steigung und die Steigungsausrichtung (Easting und Northing), was generell gut mit den vorhandenen Erkenntnissen für größere räumliche Skalen übereinstimmt.

Bezüglich der Ausbreitung der anthropogen induzierten benthischen Sedimentfahne belegt die Arbeit, dass sich initiale Trübestrome mit hoher Dichte senkrecht zu den Abbauspuren ausbreitet und sich nur wenig oder gar nicht mit dem Umgebungswasser vermischen. Diese Trübestrome werden durch steilere Abschnitte am Meeresbodens ( $> 3^\circ$  Steigung) kanalisiert, wodurch sich die Ausbreitungsdistanz und die Menge an Sediment außerhalb des eigentlichen Abbaubereiches erhöht, ein Faktor, der bisher in numerischen Modellen nicht berücksichtigt wurde. Nach der Abschwächung der Trübestromdynamik (in ca. 100 bis 500 m Entfernung, je nach Steigung des Meeresbodens) kommt es zum passiven Weitertransport von Schwebepartikeln durch Bodenströmungen. Die Partikel-Konzentration nimmt dabei mit der Entfernung und Höhe schnell ab. Dabei sinken Partikelkonzentrationen in 4.5km Entfernung und 50m Höhe über dem Meeresboden fast wieder auf das ursprüngliche Hintergrundniveau. Während der Untersuchungen wurden erstmals Partikelausflockung in situ festgestellt - ein weiterer bisher vernachlässigter Faktor, der zu einer schnellen (Minuten bis Stunden) Sedimentablagerung innerhalb kurzer Entfernungen ( $< 500$  m) vom Abbaubereich führt. Innerhalb des Abbaubereiches wurde durch hochauflösende Fotomosaiks und der daraus generierten Geländemodelle mit mm Auflösung neue Sedimentablagerung von bis zu 3 cm und Erosionen von bis zu 5 cm ermittelt. General zeigt das durchgeführte Monitoring der Sedimentausbreitung die Notwendigkeit, das hydroakustische und optische Rückstreuungssensoren auf stabilen und beweglichen Plattformen eingesetzt werden müssen um diese mit Bildern vom Meeresboden zu kombinieren, um eine aussagekräftige, vollumfängliche und hochauflösende 4D-Sedimentfahnenüberwachung zu ermöglichen.

# Synopsis

## Introduction

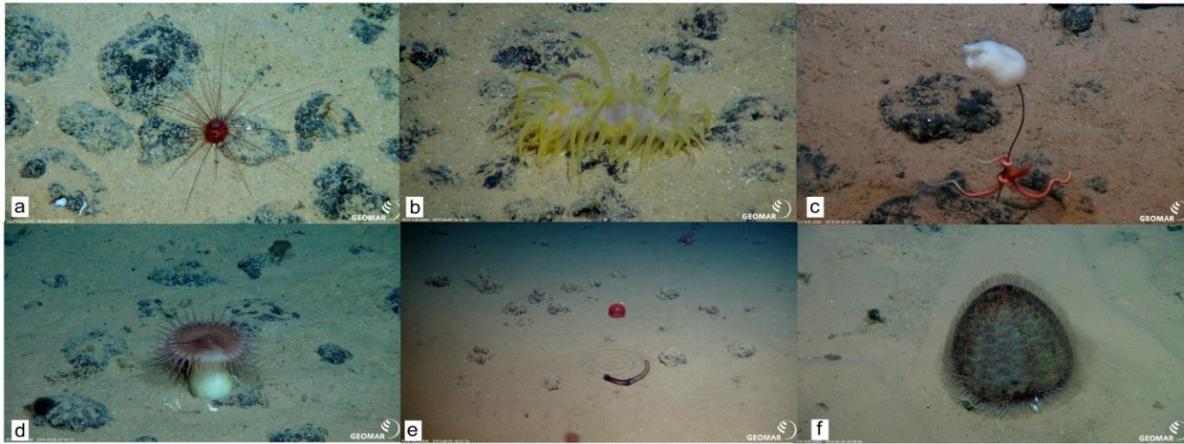
Deep-sea polymetallic nodules (hereafter nodules) are mineral concretions of mm to cm in diameter, embedded in the sediment surface of the ocean floor, mainly between 4,000 and 6,000 m water depth (Hein et al., 2020). They are formed by precipitation of Fe- and Mn-oxyhydroxides from seawater and sediment porewater (Figure 1), creating repeated concentric encrustations around a core (e.g., nodule fragment, rock debris, and shark's tooth) over millions of years (Kuhn et al., 2017; Hein et al., 2020). Nodules formed by hydrogenetic precipitation have an average growth rate of  $1\text{--}5\text{ mm Myr}^{-1}$ , while diagenetic precipitation results in a faster growth rate of up to  $250\text{ mm Myr}^{-1}$  (Kuhn et al., 2017 and references therein). Both precipitation types can co-occur in the same nodule (Figure 1), resulting in an asymmetrical growth and texture between the upper (exposed = hydrogenetic) and lower (embedded = diagenetic) parts (SPC, 2013).



**Figure 1.** a) Nodule with hydrogenetic and diagenetic growth. b) Simplified schematic illustration of nodule formation. Figure source: SPC, 2013.

The extremely slow growth rate is why scientists call nodules the "**black box**" of ocean history (ISA, 2009; Reykhard and Shulga, 2019). Indeed, the study of nodules has provided valuable information on the changes in physical and geochemical conditions of deep-sea basins. For example, a single nodule's magnetic and chemical scanning revealed the Antarctic Bottom Water (AABW) current weakening in the early Pliocene (Yi et al., 2023). Another example is the first in situ documentation of a fossil tooth of the megatooth shark *Otodus (Megaselachus) megalodon* from the deep-sea within a nodule field in the Pacific Ocean (Pollerspöck et al., 2023).

Apart from a priceless repository of ocean conditions, nodules are central to the deep-sea ecosystem as they are one of the few hard substrates on vast regions covered by otherwise soft siliceous clay (Figure 2). Several studies have linked nodule abundance ( $\text{kg m}^{-2}$ ), seafloor area coverage ( $\% \text{ m}^{-2}$ ), and nodule size (cm) to the abundance, spatial distribution, species richness, and community composition of benthic megafauna, macrofauna, meiofauna and microbial communities (Amon et al., 2016; Vanreusel et al., 2016; De Smet et al., 2017; Simon-Lledó et al., 2019; Molari et al., 2020; Pape et al., 2021). A recent study upgrades the ecosystemic role of nodules as it supports the idea of oxygen formation (the so-called "dark oxygen") through seawater electrolysis, with the necessary energy to come from the potential difference between metal ions within the nodule layers, particularly of larger nodules (Sweetman et al., 2024).



**Figure 2.** ROV images of benthic megafauna inside nodule fields in Clarion Clipperton Zone and Peru Basin. a) Echinoidea, b) Holothuroidea, c) Porifera and Asteroidea, d) Actiniaria, e) Holothuroidea and Scyphozoa, f) Echinoidea. The fauna annotation was based on the standardized taxonomic field guide by Simon-Lledó et al., 2023. Figure source: GEOMAR Helmholtz Centre for Ocean Research Kiel, ROV Team.

The importance of nodules for the benthic ecosystem is clearly depicted when nodules have been removed intentionally during disturbance experiments, e.g., ISOPE (2002) and Haalboom et al. (2022) or pilot mining trials, e.g., Burns (1980) and Vink et al. (2022) (Figure 3). The removal of nodules disturbs the benthic ecosystem; studies show alternations in sedimentological and geochemical parameters within the sediment, and a decrease in abundance, spatial density and structure of the benthic community even after 26-36 years (Miljutin et al., 2011; Jones et al., 2017; Volz et al., 2020; Vonnahme et al., 2020; Cuvelier et al., 2023; Lefaible et al., 2023). The sediment plume created during the nodules' removal has also been of high environmental concern as it increases the suspended particulate matter (SPM) concentration several times or magnitudes above the very low background levels of  $0.02 \text{ mg L}^{-1}$  that the deep-sea biota is adapted to. Recent experiments showed that SPM concentrations above  $17 \text{ mg L}^{-1}$  negatively affect deep-sea pelagic jellyfish species (Stenvers et al., 2023). The redeposited sediment following the sediment plume also negatively affects abundance and alternations in the benthic community structure (Guilhermic et al., 2023; Lefaible et al., 2023).



**Figure 3.** ROV images showing a) the disturbed seafloor 36 years after a pilot mining test in Clarion Clipperton Zone, b) the sediment plume during and mining imprints during a recent mining trial and c) the sediment redeposition 50 m from the mining site. Figure sources: a) GEOMAR Helmholtz Centre for Ocean Research Kiel, b and c) Federal Institute for Geosciences and Natural Resources and Gazis et al., 2025 <https://doi.org/10.1038/s41467-025-56311-0>

More importantly, benthic sediment plumes transport and redeposit the suspended sediment outside the mining sites, increasing the environmental footprint of mining activities (Ozturgut et al., 1981; Muñoz-Royo et al., 2022; Haalboom et al., 2023). Due to the lack of large-scale mining trials, the incomplete knowledge of deep-sea benthic ecosystems and the complexity of acquiring data on the resilience of deep-sea life to such disturbances, the spatiotemporal

evolution and impact of mining-induced plumes have been characterized as the "greatest unknown factor" in deep-sea mining activities (Amos and Roels, 1977; Amon et al., 2022).

Nodules could play a significant role in transitioning to green and decarbonized technologies, as they have a high average abundance (e.g., 10–15 kg m<sup>-2</sup> of wet weight) and are rich in metals as Mn (28.4 wt %), Fe (6.16 wt %), Ni (1.30 wt %), Co (0.215 wt %), Cu (1.07 wt %), Li (131 ppm) and rare earth elements such as Yttrium. These metals are currently in increasing demand as they are needed in technologies such as electric cars and wind turbines (Hein et al., 2020). The European Union alone will need 60 times more lithium and 15 times more cobalt by 2050 than today for this transition (EC, 2020). Thus, nodules are considered a valuable (now renewable) 2D resource (including the first five to ten cm of sediment column), and their existence has received attention. Considering the geological, ecosystemic and economic importance of nodules, the presented Thesis addresses the topics related to the scientific gaps in high-resolution spatial nodules mapping and the monitoring of mining-induced sediment plumes:

**Section I** addresses the main research topic of the thesis, which investigates spatial predictions of deep-sea polymetallic nodules by using hydroacoustic and optic data in machine learning algorithms (main research topic). Section I introduces the reasons that affect nodules' spatial distribution on global and regional scales and underlines the need for local scale studies using data obtained by autonomous underwater vehicles (AUVs). Afterwards, the research objectives are clarified, and the key findings from the three published research papers are summarized and related. Finally, an outlook for future work is given.

**Section II** deals with the secondary research topic of the Thesis, the mining-induced benthic plume monitoring. It summarizes the research findings of the provisionally accepted research paper related to in situ plume monitoring of the first modern trial of an industrial-scale pre-prototype seafloor nodule collector in the Pacific Ocean. The study underlines the need for a holistic scientific and technological approach to study such a complex topic and presents important outcomes regarding the technology needed to monitor sediment plumes in high spatiotemporal resolution in the future.

**Section III** describes the author's scientific development and achievements, including co-authorships in peer-reviewed original research articles and book chapters, scientific data publications, participation in international conferences, scientific expeditions, workshops and awards received during the PhD.

**Section IV** includes the references of sections I to III.

**Section V** contains the published and submitted research articles, their appendixes, and supplementary information.

# Section I

Deep-sea polymetallic nodules spatial modelling with machine learning algorithms and backscatter analysis.

*"Everything is related to everything else, but near things are more related than distant things."*

*The first law of Geography, (Waldo R. Tobler, 1970)*

## Foreword

Section I first introduces the research topic of deep-sea polymetallic nodules and summarizes the results and knowledge gained from addressing the various research objectives defined during the PhD time and presented after the research topic introduction. After that, each Research Objective is discussed independently but also in relation to each other. The derived conclusions are set in a broader context, as some findings are more generally relevant to the research fields of spatial predictive modelling and habitat mapping. The same section also mentions the research projects and further work initiated from these findings and the remaining scientific gaps, giving an outlook for future work. However, it does not describe the methods used, although it is a substantial part of this thesis. All methods used for hydroacoustic and image data analysis, together with the applied statistical and modelling techniques, are described in detail in each research article in Section IV. The research articles also describe in detail the geological settings of each study area and include more figures and tables. Within the Synopsis part, indicative examples of figures are presented to facilitate the purposes of the section. As the rationale for the synopsis is to outline the main findings, describing the methods and geomorphological settings would disorient the reader from the primary synopsis target.

It should be underlined that one of the most challenging aspects of Section I and II was the need to study and combine methodologies from various research fields:

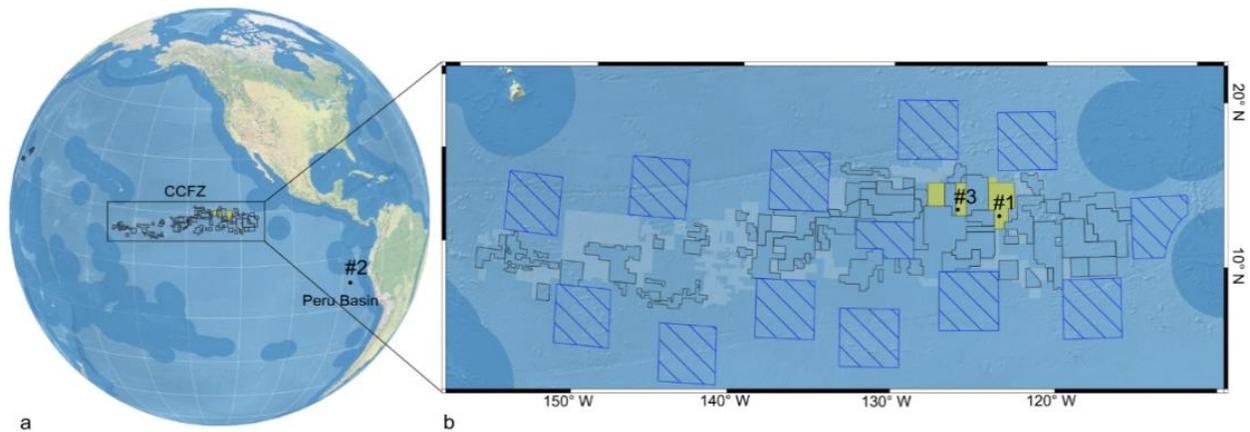
1. Underwater hydroacoustic (MBES and SSS data acquisition and post-processing)
2. Underwater imagery image analysis and orthophoto-mosaicking)
3. Statistics (e.g., univariate and multivariate analysis, spatial statistics and geo-statistics)
4. Machine learning (e.g., feature selection and modelling algorithms)
5. Programming (e.g., statistical, modelling and visualization packages in R)
6. Geographic Information Systems (e.g., spatial analysis and map production)
7. Marine Geology (e.g., formation and distribution of nodules)
8. Survey designing (e.g., AUV multi-purpose and multi-sensors survey planning)

The combination of these eight research fields and the skills acquired during the PhD time played a decisive role in the author's interdisciplinary scientific development, allowing him to understand the topic in-depth and independently conduct research from A (survey design) to Z (publishing results). Often, mishandlings at the data acquisition stage (e.g. suboptimal sensor settings) hinder a smooth data analysis and could even lead to erroneous conclusions. Similarly, poor data analysis (e.g., wrong data-post processing, statistical or coding errors) could lead to misleading results even for good-quality data. In addition, the blind use of machine learning approaches without domain knowledge results in sub-optimal conclusions, particularly in natural sciences. Marine geoscientists nowadays should have an interdisciplinary approach and learn various fast-evolving methods, tools, and software packages (commercial and open source) to address spatial problems with remote sensing methods. In the era of artificial intelligence, a good understanding of the data life cycle, the algorithms used and the causal natural mechanisms behind the investigated topic are essential more than ever. This PhD thesis has a holistic approach, creating a broad-minded marine geoscientist!

## Nodules Spatial Distribution

Since their discovery in 1873 during the *Challenger* expedition (Thomson, 1873; Belkin et al., 2021), nodules have been found in all oceans, the Pacific (e.g., Kuhn et al., 2017), Atlantic (e.g., Cronan, 1975), Indian (e.g., Sharma et al., 2013), as well as the Arctic (e.g., Hein et al., 2017) and Southern Ocean (e.g., Ostwald, and Frazer, 1973). However, they are unevenly distributed as several environmental factors affect nodule formation, growth, abundance, metal content, shape, size and texture on global, regional and local scales.

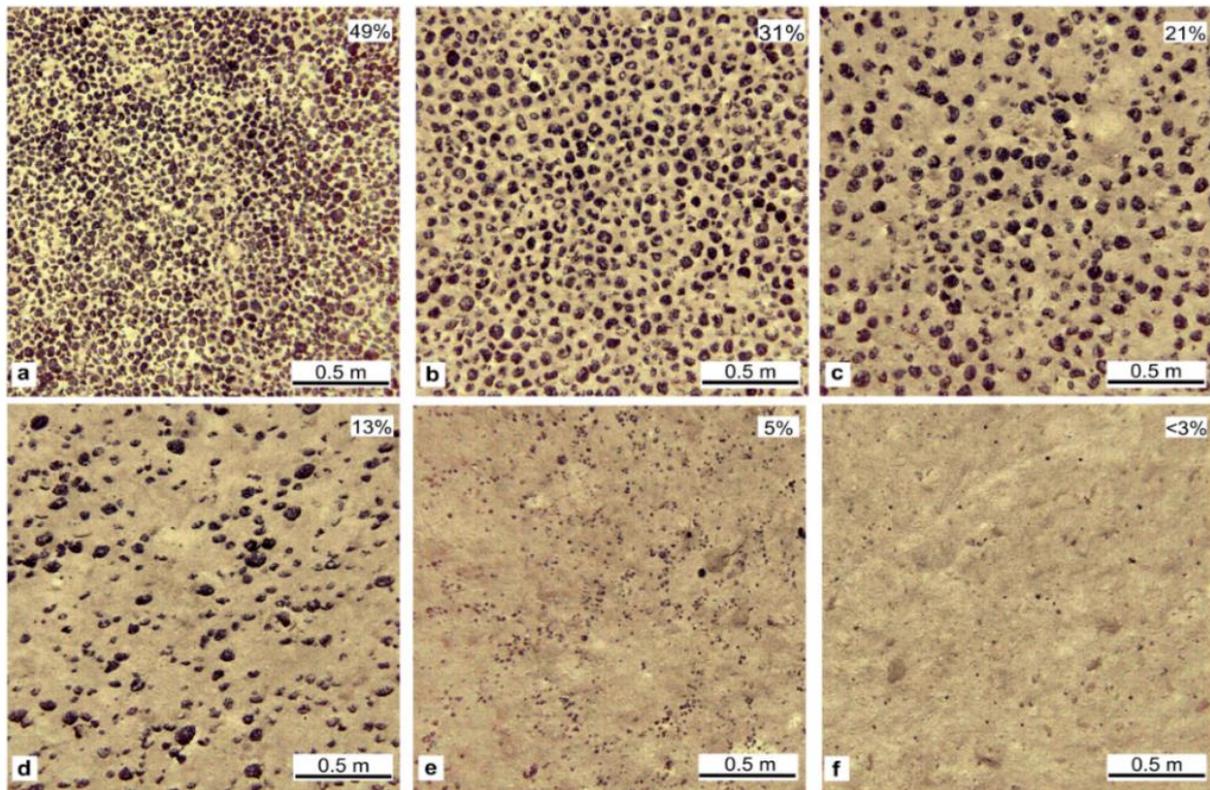
**On a global scale**, their spatial distribution is driven by the water depth and Calcium Compensation Depth (CCD), long-term low sedimentation rates, bottom water oxygen concentration, long-term bottom current velocity, seafloor lithology and sediment size, sea-surface productivity and total organic carbon content in seafloor sediments, seafloor biomass and plate tectonics (Dutkiewicz et al., 2020). Briefly, the seafloor 100-300 m below the CCD is most favourable for nodules as siliceous ooze on the seafloor allows metal exchange through the sediment pore water, enhancing diagenetic growth. Deeper than this, the diagenetic contribution is reduced due to lower particle flux rates of metal reaching the seafloor and the dominance of finer grain size (deep-sea clay), which has reduced sediment permeability and metal mobility in pore waters compared to siliceous ooze. Above the CCD, the dilution from precipitating carbonates and the higher content of calcareous ooze on the seafloor sediments hamper the formation and growth of nodules (Hein et al., 2013; Kuhn et al., 2017). Long-term low sedimentation rates ( $< 5\text{-}10\text{ cm kyr}^{-1}$ ) and low total organic carbon content ( $< 0.5\text{ wt }%$ ) in areas with medium to low surface productivity ( $< 300\text{ mg C m}^{-2}\text{ day}^{-1}$ ) are needed to support the slow mineral precipitation and retain well-oxygenated waters ( $150\text{-}210\text{ mmol m}^{-3}$ ) and oxic conditions on the surficial sediments and its pore water (ISA, 2010; Dutkiewicz et al., 2020). Long-term steady and well-oxygenated large-scale bottom currents, such as the AABW currents, inhibit sediment accumulation by eroding and removing the finest sediments and generating fragmental materials that serve as nodules' nuclei (ISA, 2010; Kuhn et al., 2017). Large plate movements, such as the northwest shift of the Pacific plate, also contribute to creating optimal (past) or suboptimal (present) conditions for nodule formation (ISA, 2010; Kuhn et al., 2017). Biotic processes also contribute to nodule growth and preservation. For example, free-living bacteria create a biofilm around nodule seeds, contributing to the precipitation of Mn- and Fe-oxyhydroxides (Wang and Müller, 2009; Zhang et al., 2023). The activity of benthic fauna around, under and on top of nodules (detritus feeding and foraging) is considered the primary mechanism sustaining nodules on the sediment surface, as they clean the fine sediment on top of nodules and deposit it underneath them (Gardner et al., 1984; Mullineaux, 1989; Dutkiewicz et al., 2020). The spatial variability of the abovementioned factors has resulted in regions with higher nodule abundance and metal content, gathering worldwide and national interest for exploration and future exploitation. Such regions include the Clarion Clipperton Zone (CCZ) in the NE equatorial Pacific, the Central Indian Ocean Basin, the Peru Basin, the Penrhyn Basin, and the Exclusive Economic Zone (EEZ) of the Cook Islands. Only within the CCZ, which has an area of ca. 4.5 million  $\text{km}^2$ , the International Seabed Authority (ISA) has issued 18 exploration licenses (ISA, 2023; Figure 4).



**Figure 4.** a) The globe map shows the CCZ between Central America and Hawaii, where most exploration contract areas for nodules have been issued, and the location of the Peru Basin. The EEZs are darker blue areas (FMI, 2019). b) A zoom-in map of the CCZ shows the exploration contract areas (black polygons), the Areas of Particular Environmental Interest (hashed blue squares), and the Reserved Areas (light blue blocks; ISA, 2023). The Global Sea Mineral Resources (GSR) exploration contract area is marked yellow, and the locations of the three study sites are provided. Each number corresponds to the number of each published research article chronologically ordered. Figure source: Gazis et al., 2025 <https://doi.org/10.1038/s41467-025-56311-0>

**Regional seafloor morphology** further influences the nodules' spatial distribution (Figure 5). Local bathymetry (Craig, 1979; Kodagali, 1988; von Stackelberg, 2000), terrain ruggedness (Kodagali, 1988; Skornyakova and Murdmaa, 1992), steep slopes, and outcropping bedrock (Dreisettl, 2016; Yoo et al., 2018; Alevizos et al., 2022), volcanism and the availability of nuclei (Glasby et al., 1973) have a direct effect on the nodules' spatial distribution. Synergistically, seafloor morphology alters the local hydrodynamics, e.g., when solitary knolls or seamount chains influence bottom current velocity and direction, causing erosive and depositional structures such as furrows and sediment drifts, respectively (Juan et al., 2018). Alternations in local sediment depositional patterns that influence sediment thickness (e.g., seafloor pits with increased sedimentation and conical hills with sediment accumulation at the lee side) contribute to the patchy nodule distribution (Wiedicke and Weber, 1996; von Stackelberg, 2000; Juan et al., 2018). The synergistic effect of seabed topography, bottom currents, and heterogeneous sedimentation patterns influences the geochemical sediment conditions that affect nodule formation, e.g., removal of clay-size particles by currents, variations in total organic carbon content and shifts of the Mn-redox depth in local depressions, highlighting the relationship between meter-scale seabed morphology and variation in geochemical conditions (Mewes et al., 2014; Volz et al., 2018; Paul et al., 2019).

**The thesis focuses on the geomorphological component of this interplay**, providing insights into the role of topography on nodule occurrences on local scales (meters to a few km), which are not studied as thoroughly as the regional (hundreds to thousands of km) and global scales mainly due to the lack of high-resolution data that are typically generated through AUV deployments.



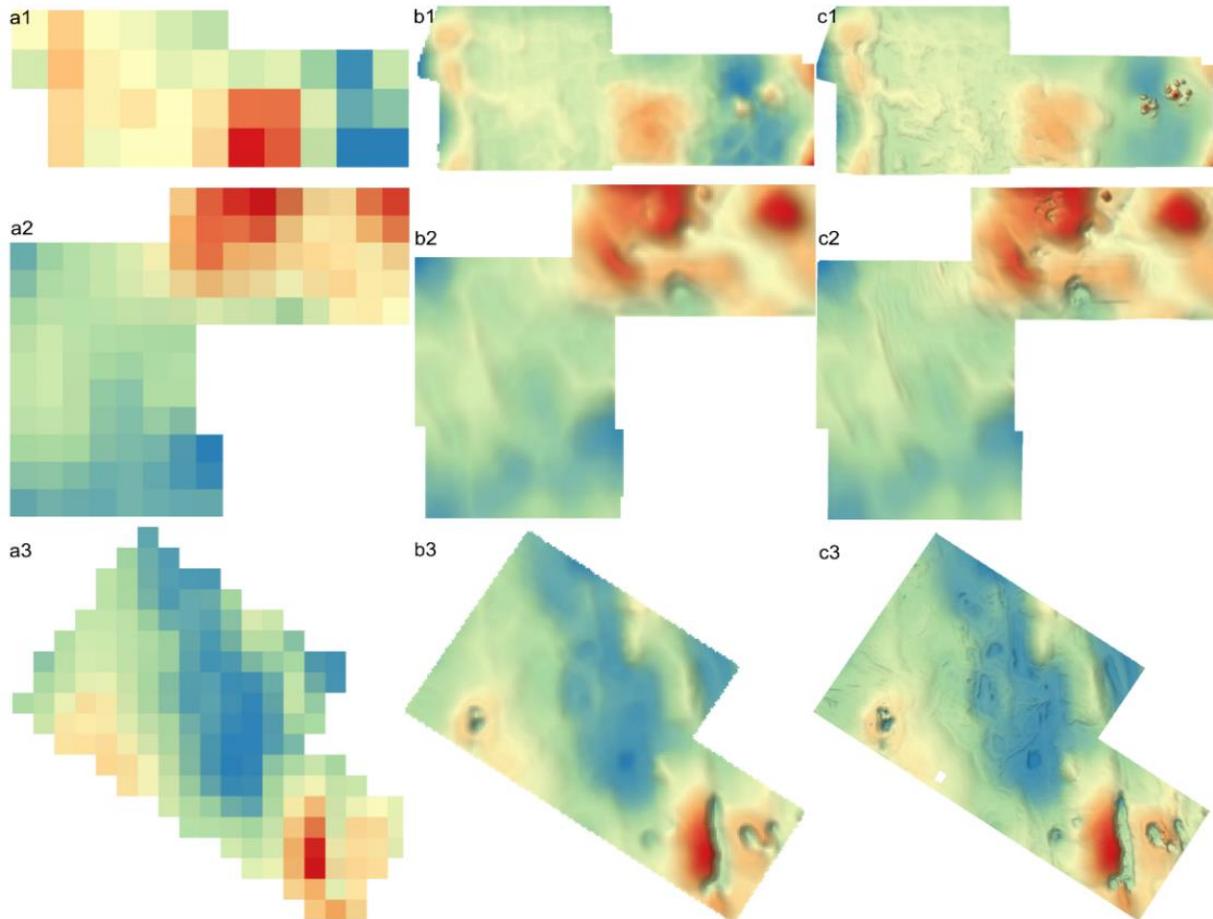
**Figure 5.** AUV colour-normalized (Köser et al., 2021) seafloor images showing the different nodule coverage within the same study site. The image-derived nodule coverage has been calculated using the Mangan Analyzer software (saltation GmbH & Co. KG). Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

**AUV Multibeam Echosounder (MBES) data:** This thesis significantly enriches the available high-resolution deep-sea data by **presenting the most extensive AUV-based MBES datasets on nodules as a deep-sea resource at the time of writing.** In total, 75.3 km<sup>2</sup> of seafloor were mapped in three study areas (Figure 6): 9.5 km<sup>2</sup> in the Global Sea Mineral Resources (GSR) B6 area of the CCZ (Research article I), 28.2 km<sup>2</sup> in the Disturbance and recolonization Experimental Area (DEA) area of the Peru Basin (Research article II) and another 37.6 km<sup>2</sup> in the GSR B4 area of the CCZ (Research article III). Notably, the GSR B4 area was ensonified by a 400 kHz MBES with a beam width of 0.7° × 0.7° from an acquisition altitude of 60 m above the seafloor; this resulted in a footprint\* of 0.7 m × 0.7 m. The smaller the beam width and the shorter the acquisition altitude, the greater the resolution as the size of the seafloor area in which a certain number of pings fall (footprint), is smaller (Hughes Clarke 2018). **This MBES footprint is the smallest in current literature regarding nodules and other deep-sea resources,** such as seafloor massive sulfides (Juliani et al., 2021; Haroon et al., 2023). AUV-based MBESs capture the nodules' backscatter reflectance with high-frequencies (< 100 – 700 kHz), a frequency spectrum in the past not used commonly.

---

\* The MBES footprint calculation is done for a flat seafloor according to the equations provided by the International Hydrographic Organization, IHO 2011.

Such high-resolution data are necessary to depict the meter-scale seafloor structures that influence the nodules' spatial distribution, which are not resolvable with ship-based MBES data (e.g. Peukert et al., 2018a and Figure 6). Low frequency ship-based MBESs and towed Side-Scan Sonar (SSS) systems ( $< 12 - 100$  kHz) map the nodule coverage on regional scales, i.e., hundreds to thousands of  $\text{km}^2$  (Scanlon and Masson, 1992; Lee and Kim, 2004; Chunhui et al., 2015; Machida et al., 2019; Kuhn and Rühlemann, 2021; Wang et al., 2021).



**Figure 6.** Bathymetric maps of the three study areas based on three different sources: a1-a3) GEBCO-based bathymetric grid with a cell size of 450 m (GEBCO, 2024). b1-b3) Ship-based MBES bathymetric grid with a cell size of  $50 \times 50$  m (Greinert, 2016; Gausepohl et al., 2019 a, b; Gazis, 2020). c1-c3) AUV-based MBES bathymetric grids with cell sizes of  $3 \times 3$  m (a3 and b3) and  $2 \times 2$  m (c3). Figure source: Gazis, 2025 (PhD dissertation).

**AUV optic data:** The recent advancements in underwater photography (Song et al., 2022; She et al., 2023; Song et al., 2024), nodule segmentation (Schoening et al., 2017; Song et al., 2019; Shao et al., 2023; Tomczak et al., 2024) and image clustering (Mbani et al., 2022) make seafloor images a valuable tool for quantitative nodule studies, providing thousands to millions of ground truth data in high-resolution (mm to cm scale). Based on thousands of images and image-derived nodule coverage estimations, robust spatial and statistical analyses have been made possible and clearly show that nodule distribution varies significantly even in short (m to km) distances (Schoening et al., 2017; Peukert et al., 2018; Alevizos et al., 2022).

**Machine learning:** The large amount of hydroacoustic and optic data collected during offshore surveys accelerated the application of machine learning at local (Wong et al., 2021),

regional (Hari et al., 2018; Kaikkonen et al., 2019; Kuhn & Rühlemann, 2021; Wasilewska-Błaszczuk & Mucha, 2021) and global scales (Dutkiewicz et al., 2020). Machine learning algorithms have also been used to a smaller extent for the spatial predictive modelling of ferromanganese crusts (Josso et al., 2023) and seafloor massive sulfides (Juliani et al., 2021; Haroon et al., 2023). Machine learning algorithms use MBES and SSS backscatter and bathymetric derivatives (quantitative descriptors of seafloor morphology) to predict the spatial distribution of nodules based on the associations built. These spatial predictors relate to nodule coverage or are proxies for the physical and geochemical processes that support or prevent nodule growth. For example, the slope orientation has been used as a proxy for the mean current direction in spatial models when continuous long-term information about bottom currents is unavailable (e.g. Uhlenkott et al., 2022). Variations in sediment thickness are reflected in the MBES backscatter intensity, Angular Range Analysis (ARA), and SSS backscatter mosaics (Yoo et al., 2018; Hillman et al., 2017; Gaida et al., 2019) and can be inferred from topographic indices such as the bathymetric position index (BPI; Alevizos et al., 2022).

In general, the thesis is not limited to getting an accurate prediction map. It conducts an in-depth machine learning methodological analysis, which includes comparing different machine learning algorithms and validation techniques and aims at identifying the limitations of model-based spatial predictive mapping. Particular attention was given to the interpretability of the machine learning models using interpretable machine learning tools such as variable importance ranking, partial dependence plots (2D and 3D) and domain knowledge. Such a combined analysis approach sheds light on the sometimes mystic character of machine learning models and provides valuable feedback regarding each predictor variable's statistical and natural contribution and highlights the role of the seafloor geomorphology on the spatial distribution of nodules. The thesis provides a detailed overview of the current literature on this topic.

## Research Objectives of Section I

Section I has five Research Objectives:

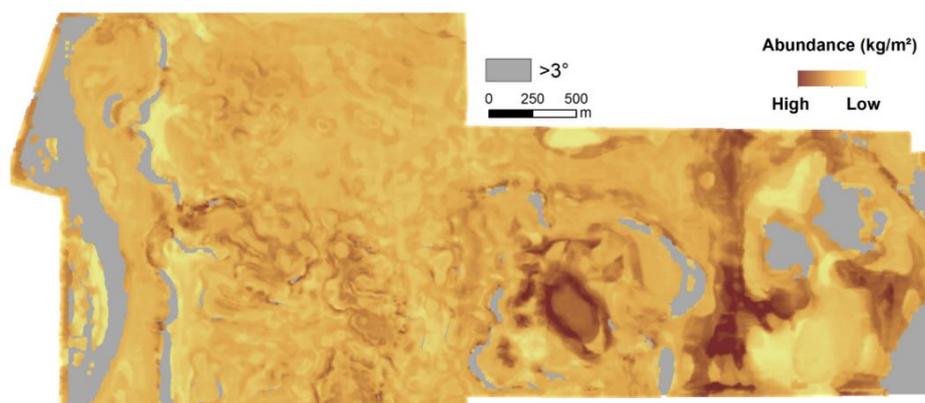
- I. **Investigate** the applicability of machine learning regression algorithms to predict the fine-scale spatial distribution of nodules using AUV-based hydroacoustic (MBES and SSS) and optic data (seafloor images), as well as a small number of ground truth data (i.e. box corers).
- II. **Study** the methodological limitations, pitfalls, and challenges of machine learning-based spatial modelling of nodules, related to spatial autocorrelation and feature space extrapolation.
- III. **Examine** the performance of five commonly used machine learning regression algorithms for the spatial prediction of nodules: generalized linear models, generalized additive models, random forests, support vector machines, and neural networks.
- IV. **Analyze** the MBES backscatter response of seafloor parts with different nodule coverage and study the contribution of different incident angles (angular range analysis) and backscatter components (impedance and volume) on the modelling performance.
- V. **Provide** insights into the geomorphological reasons behind the nodules' patchiness, by combining hydroacoustic and optic data, domain knowledge, and interpretable machine learning.

Three peer-reviewed original research articles addressed the abovementioned objectives (see Section V). No other peer-reviewed study with such objectives was published before these research articles.

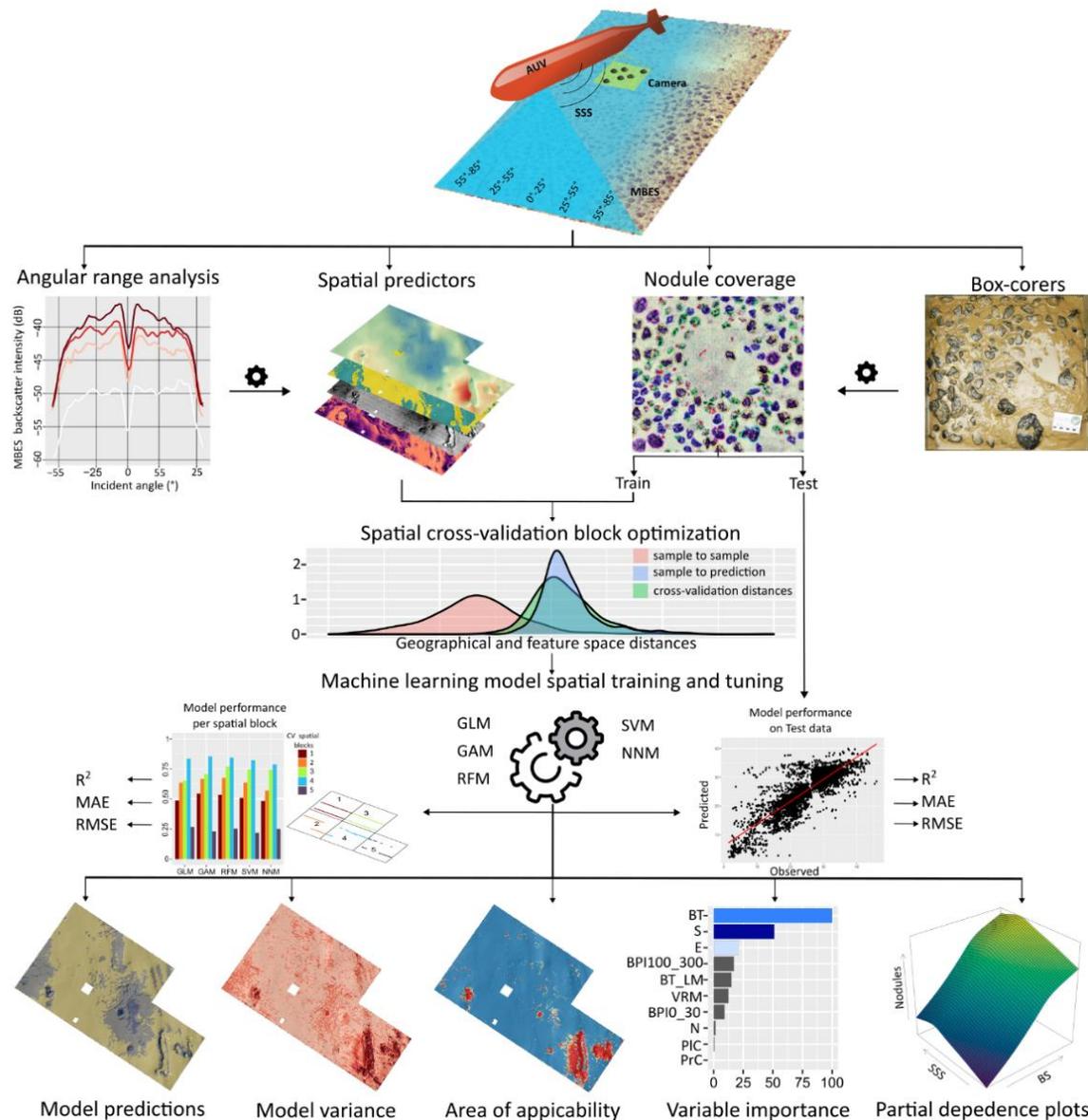
**Research Objective I:** *Investigate the applicability of machine learning regression algorithms to predict the fine-scale spatial distribution of nodules using AUV-based hydroacoustic (MBES and SSS) and optic data (seafloor images), as well as a small number of ground truth data (i.e. box corers).*

### Key findings

The three research articles show that machine learning algorithms extract meaningful information from thousands of seafloor images and hydroacoustic data, capturing linear, monotonic, nonlinear and non-monotonic relationships between nodules, backscatter signal and local geomorphology such as depth, slope, and rugosity (also Research Objectives III-V). Based on the relations built, one can develop detailed spatial maps on operational scales in a reasonable computational time, allowing near-real-time applications as soon as the image processing occurs fast enough (Figure 7). Research article I presents for the first time a workflow on how to combine AUV-based MBES data, image-derived values of nodules occurrences (number  $m^{-2}$ , % coverage  $m^{-2}$  and median size  $m^{-2}$ ) and ground truth data (box corers) using machine learning (e.g. random forests) to predict the fine-scale (meter to km scale) spatial distribution of nodules (Figure 8). Research article II focused on the methodological limitations of this approach and how they can be addressed (also Research Objective II), and research article III studied the differences between machine learning algorithms and performs an ARA of the MBES backscatter reflectance from different incident angles to derive the most sensitive incidence angles (also Research Objective III). MBES backscatter was the most important spatial predictor (also Research Objective IV) and, together with data generated from seafloor images, revealed a meter scale patchiness that cannot be resolved with box-corer data that are typically scarce and sparse (one sample per  $km^2$ ) and only represent a very small sampling area ( $0.25 m^2$ ). The spatial autocorrelation analysis with Anselin Local Moran's Index (Anselin, 1995), showed that nodules cluster into statistically significant ( $I > 0.5$ ,  $p < 0.05$ ) groups of higher and lower coverage following the fine-scale seafloor variations in all study areas, underlying the importance of a detailed terrain reconstruction for understanding and predicting nodule occurrences on a varying deep-sea seafloor (also Research Objective V).



**Figure 7.** Map showing the model predicted wet abundance ( $kg m^{-2}$ ) of nodules, accounting for the buried nodules (max 15 cm). The seafloor parts with slope  $\geq 3^\circ$  are masked, representing a future mining scenario. Figure source: Gazis et al., 2018. <https://doi.org/10.5194/bg-15-7347-2018>



**Figure 8.** Schematic flowchart of the main steps for the spatial prediction of nodules using AUV-based data and machine learning algorithms: From data acquisition to model outputs and interpretation. Each Research Objective discusses one or more steps of this workflow. Figure source: modified from Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

### Research initiatives

The key findings of Research Objective I initiated additional research topics that have been addressed in other projects or additional studies by the author:

a) Fine-scale distribution maps of nodules allow accurate resource estimation and optimized mining path planning on an operational scale and time. Such maps would be the base information for an autonomous and adaptive steering system for seabed nodule mining, which will determine the most operationally and environmentally efficient mining path on the seabed. Such operational calculation will integrate forecasted hydrodynamic and sediment transport data, optimal manoeuvrability of the vehicles and production efficiency while guaranteeing safe execution, e.g. by avoiding outcropping basement and increased slopes not resolved in ship-based MBES data (e.g. Peukert et al., 2018a). This idea initiated a 3.5-year research

program entitled COMPASS: Control of an Operational Mining Path Through an Auto-Adaptive Steering System. COMPASS strengthened the collaboration between GEOMAR Helmholtz Centre for Ocean Research Kiel and two industrial partners from Germany (saltation GmbH and Co. KG) and Belgium (Global Sea Mineral Resources NV). The MarTERA partners Flanders Innovation and Entrepreneurship (VLAIO), the German Federal Ministry of Economic Affairs and Energy (BMWi) and the European Union co-founded the COMPASS project (COMPASS, 2018). It should be highlighted that continuous surfaces with nodule coverage are among the best spatial predictors in deep-sea habitat species distribution modelling (Uhlenkott et al., 2020; Uhlenkott et al., 2023)

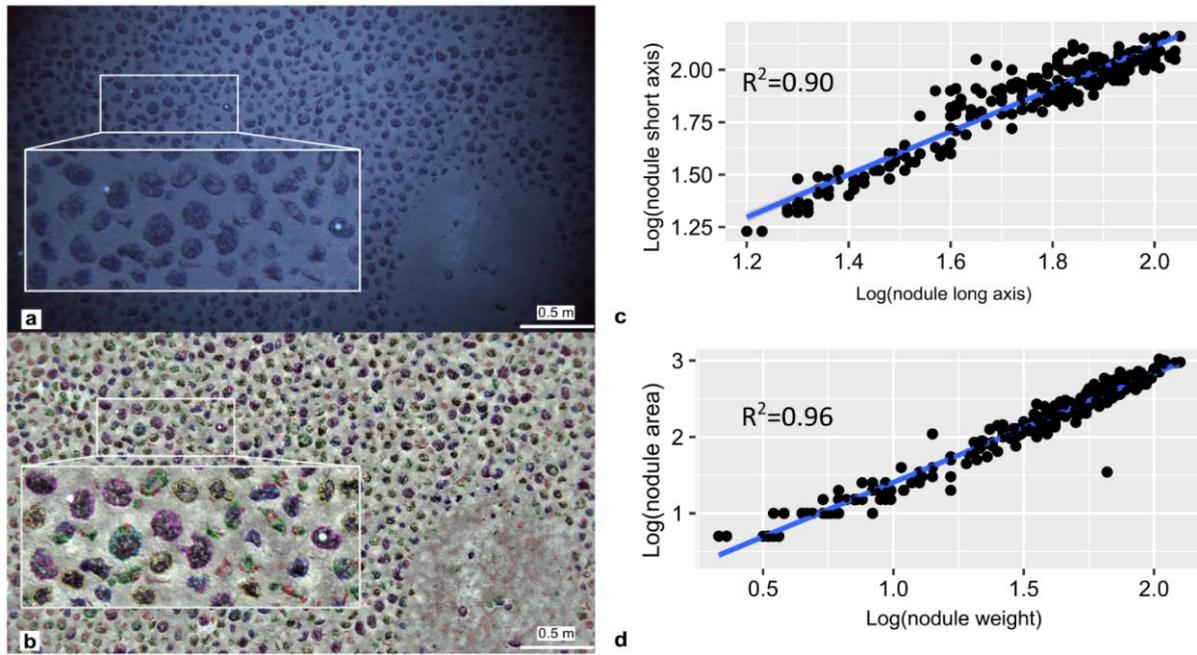
b) AUV-based images are crucial for Research Objective I, as they provide the % seafloor coverage of nodules (here after nodule coverage) constituting the training data for machine learning algorithms. This highlighted the need to empirically confirm that the ellipsoid model used by the Compact Morphology-based Nodule Delineation (CoMoNoD) and Hierarchically Growing Hyperbolic Self-Organizing Map (H<sup>2</sup>SOM) image processing algorithms yields realistic size and area estimations (Schoening et al., 2016; Schoening et al., 2017). The author, together with colleagues from the DeepSea monitoring group, measured the length, width, height, wet mass, volume, area, and buried depth of 1657 nodules from box corers collected in GSR and BGR (Federal Institute for Geosciences and Natural Resources) contract sites during the SO268 expedition (Haeckel and Linke, 2021). These measurements were published (see Data Publications) and were later used to prove that the nodule shape at these study sites agrees with the idealized ellipsoidal shape and that empirical equations (with a high degree of confidence;  $R^2 > 0.9$ ) can be established between the nodule weight and the major/minor nodule axis (Yu and Parianos, 2021; Figure 9). Such image-derived information regarding shape and size can be combined with the number, coverage and size of nodules derived from box-corer data, resulting in empirical formulas for the abundance, which are necessary as the images cannot provide information on the buried nodules and tend to underestimate the actual nodule coverage due to semi-buried nodules or nodules covered by faint sediment (Sharma 1993; 2017; Lipton et al., 2016; Ellefmo and Kuhn, T. 2021; Tsune 2021). The degree of underestimation varies between different nodule sizes, facies and substrates; underestimation is higher in sub-areas with thicker sediment layers and lower in sub-areas with rocky substrate and basement exposure (Sharma 1993, 2017; Parianos et al., 2021). In regions without surficial sediment on nodules and basement exposure-like the three study sites-results have shown a good agreement (e.g.  $R^2 > 50\%$ , locally up to 90%) between image-derived and box-corer-derived coverage (Lipton et al., 2016; Sharma 2017; GSR, 2018; Wasilewska-Błaszczyk and Mucha, J. 2021).

### **Technological challenges and solutions**

Investigations around Research Objective I highlight technological challenges related to hydroacoustic and optic data that hinder spatial modelling independently of the used machine learning algorithm.

a) Artifacts in MBES data products: AUV MBES surveys are typically conducted in low altitudes (e.g. 50-60 m above the seafloor) allowing the use of high frequency systems (e.g. 400 kHz) with small beam angles (e.g.  $0.7^\circ \times 0.7^\circ$ ) and high-ping rate (e.g. 50 Hz) achieving

high-resolution (1 to 3 m-scale). At the same time AUVs as platform avoid ship-based MBES acquisition problems like near-surface turbulence and bubbles close to transducers, ship noise, rough sea and significant sound velocity changes (Jakobson et al., 2016). However, AUVs are prone to navigation drifting over time, roll, pitch, yaw, latency artefacts, and pressure changes due to tides that result in artefacts, which are typically more prominent due to the very high resolution and are more complex to resolve than for ship-based MBES data.

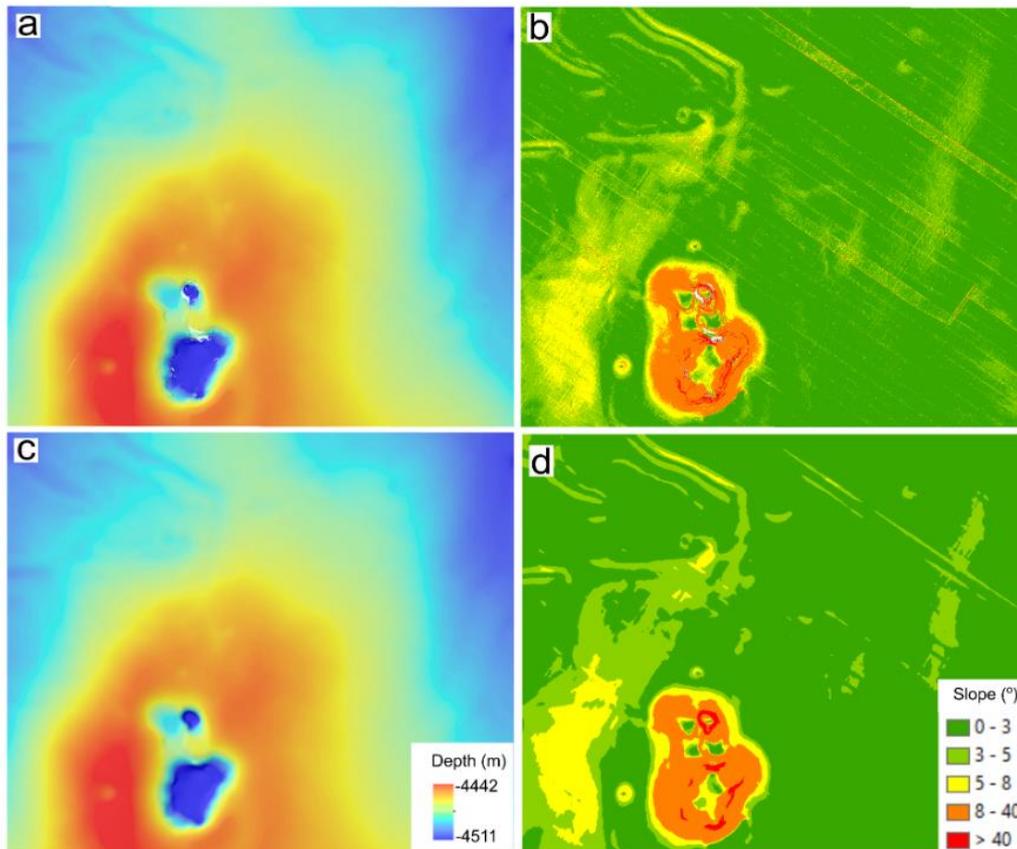


**Figure 9.** a) Raw seafloor image. The acquisition altitude is 5.33 m. The bright white spots are Epifauna indet. on top of two nodules. b) Nodule coverage segmentation and calculation using the Mangan analyzer software. The different colours correspond to different sizes. Scatterplots between the length of the two nodule axes (c) and between the nodule area (d) calculated based on the ellipsoid model and the nodule (wet) weight. Figure sources: a and b) modified from Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>, c and d) Gazis, 2025 (PhD dissertation).

Machine learning models are data-driven approaches that demand high-quality data. Unresolved acquisition artefacts affect the calculation of bathymetric derivatives (e.g., artificially increased slopes) and add unnecessary noise into the predictor variables which negatively affects spatial predictions (Lucieer et al., 2016; Lecours et al., 2017, Hughes Clarke, 2018). The performed in-depth MBES post-processing included roll bias correction, sound velocity ray tracing, tide correction. In addition, AUV turn removal, manual and automated filtering of erroneous soundings with the Combined Uncertainty and Bathymetry Estimator (CUBE) algorithm (Calder and Mayer, 2003), and low pass Gaussian filtering increased the quality of the final data product of high quality bathymetric and backscatter derivatives (Figure 10). In a last step, the final bathymetric surface is compared to the ship-based post-processed bathymetric surface, inspected for XY navigational offsets and adjusted if needed to gain an absolute georeferenced data layer (Mohrmann and Greinert, 2022).

**b1) AUV-related errors in absolute positioning:** Km-long AUV survey lines result in meter-scale positional drift over time, mainly when the AUVs use only their onboard inertial navigational systems without additional ultra-short baseline (USBL) or long baseline (LBL) acoustic positioning. Even if one or both systems are available, the km-long survey lines could

still result in positional drift. Fine-scale studies in heterogeneous seafloor demand a good positional match between the image location and the geomorphological parameter captured by the MBES- and SSS-derived predictors. Navigational post-processing (e.g. Kalman filtering) and predictor value extraction for each image location using bilinear interpolation of the four adjacent cells or from a given radius (depending on the dive-specific positional uncertainty) was applied to overcome navigational errors. This was done in accordance to findings of Naimi et al. (2011) and Stephens and Diesing (2014).

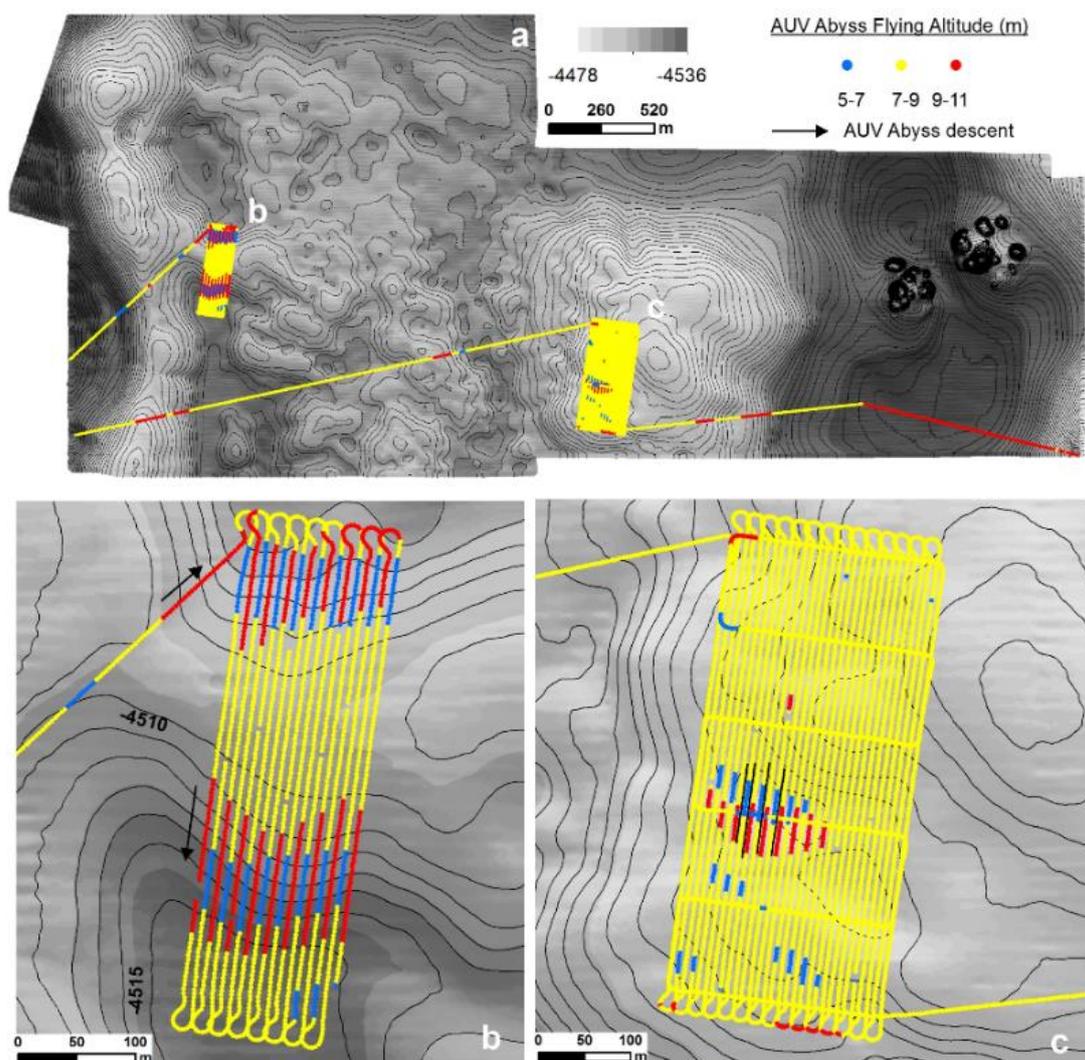


**Figure 10.** a-b) AUV-based bathymetric and slope map for a part of the study site used in Research article III, showing residual errors creating artificial slopes on a flat seafloor due to a non-thoroughly post-processing analysis. c-d) AUV-based bathymetric and slope map for the same part of the study site, showing the improvement after detailed post-processing analysis. Such residual artefacts are non-visible in bathymetric maps but become pronounced in bathymetric derivatives such as slope and rugosity, adding unnecessary noise to spatial predictors. Figure source: Gazis, 2025 (PhD dissertation).

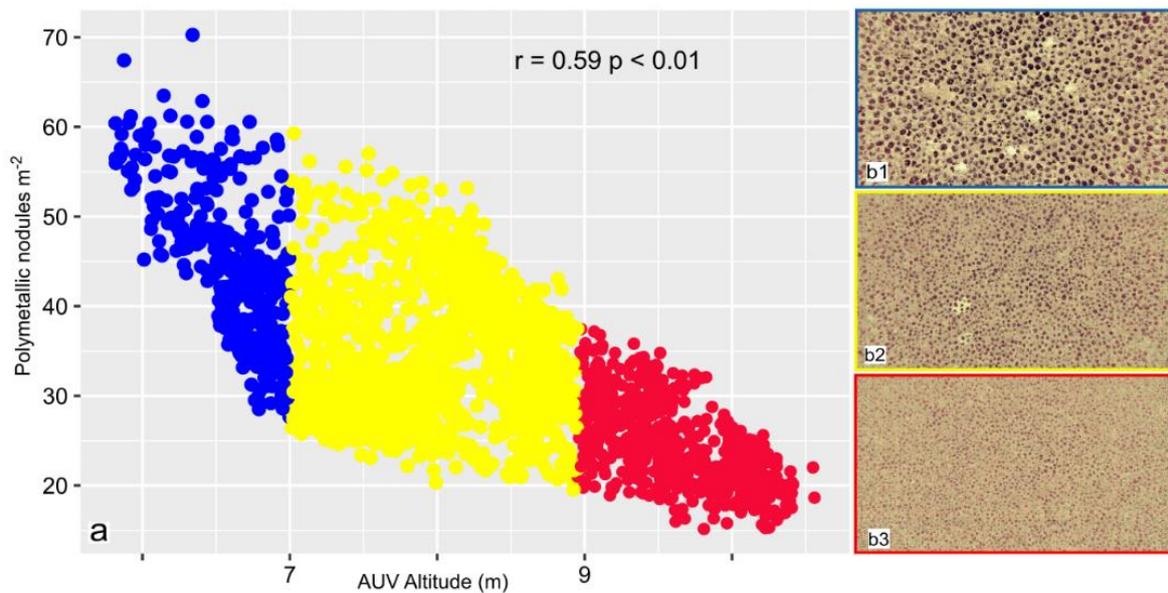
b2) AUV-related errors due to changing altitude: Flying perpendicular to slopes on a lawn mowing patterns alternates the acquisition altitude in uphill and downhill lines, resulting in seafloor images from two different altitudes within the same hillslope area which generates contradicting results in automated nodule coverage estimations of images (Figure 11 and Figure 12). Instead, flying parallel to the slope should be preferred if possible.

The lawn mowing pattern should also be avoided even in flat areas when the survey aims for orthophoto-mosaicking, as the different camera and flashlight orientations create uneven illumination and shadows that hinder the feature matching between photos from parallel survey lines. Instead, groups of survey lines with the same heading ensure uniform illumination for most mapped areas. AUV optic surveys are also affected negatively by changes (jumps) in

acquisition altitude due to erroneous altitude warnings. The image analysis articles by Schoening et al. (2016) and Schoening et al. (2017) showed that a constant AUV altitude of 4 to 5 m above the seafloor yields the best results, whilst images obtained from an altitude higher than 7 m should be excluded from nodules detection and orthophoto-mosaic reconstruction as they have reduced resolution, brightness and contrast hindering the feature matching during orthophoto-mosaicking and resulting in poor detection and underestimation of nodules (Figure 12). The author and his colleagues applied the knowledge from the abovementioned finding in the scientific expedition IP21-MANGAN (Vink et al., 2022), improving image acquisition, nodule detection and orthophoto-mosaic reconstruction.



**Figure 11.** a) AUV image survey altitude in GSR B6 contract area (CCZ). b) Slopes ( $> 3^\circ$ ) force the AUV to modify its altitude, flying closer to the seafloor in the ascending phase (blue lines) and farther from the seafloor in the descending phase. c) On a flat seafloor, the AUV flying altitude is constant (7–9 m) for almost the entire survey. Figure source: Gazis et al., 2018. <https://doi.org/10.5194/bg-15-7347-2018>



**Figure 12.** a) Scatterplot showing the statistically significant correlation between the AUV acquisition altitude and image-derived nodule number m<sup>-2</sup>. b1-b3) Color-normalized sea floor images (Köser et al., 2021) for each acquisition altitude. Figure source: modified from Gazis et al., 2018. <https://doi.org/10.5194/bg-15-7347-2018>

**Research Objective II:** *Study the methodological limitations, pitfalls, and challenges of machine learning-based spatial modelling of nodules, related to spatial autocorrelation and feature space extrapolation.*

### Key findings

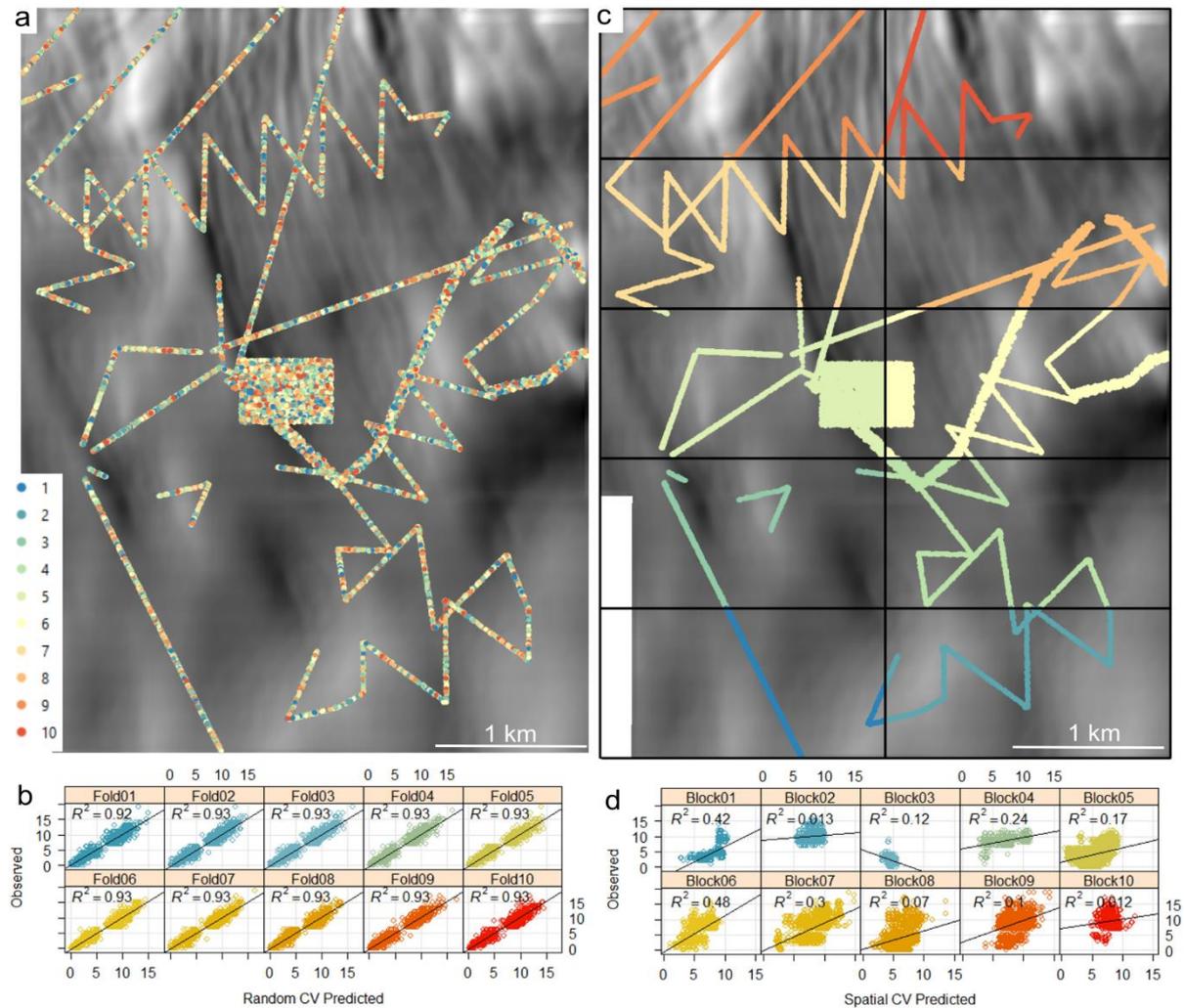
#### Spatial autocorrelation

Spatial autocorrelation describes the phenomenon where local similarity (neighbouring observations) is matched by value similarity (correlation between observations; Anselin, 1995). Independent of the machine learning algorithms used (see Research Objective III), all are prone to spatial overfitting when spatial autocorrelation exists (e.g. Roberts et al., 2017; Meyer et al., 2018; Misuk et al., 2019; Ploton, 2020; Kattenborn et al., 2022). When spatial autocorrelation is present, and a typical random k-fold cross-validation (CV) is used (i.e. the training points are assigned randomly to validation folds), the machine learning algorithm uses almost identical data for training and CV learning to reproduce neighbouring data with similar characteristics based on their spatial proximity and not the underlying relationships with the predictor variables (e.g. Dutkiewicz et al., 2017). Consequently, the fitted model predicts well in areas that fall within the range (i.e. influence) of spatial autocorrelation but poorly in areas beyond the spatial autocorrelation range. The model performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Coefficient of Determination ( $R^2$ ) are over-optimistic without reflecting the model's ability to learn from the available predictors. This phenomenon is called spatial overfitting, one of the biggest pitfalls in spatial predictive modelling with machine learning algorithms (Roberts et al., 2017; Meyer et al., 2018; Ploton, 2020; Karasiak et al., 2022). Addressing this issue is particularly important as all datasets used in the three research articles have significant ( $I > 0.5$ ,  $p < 0.05$ ) spatial autocorrelation in nodule occurrences (see Research Objective V).

Our analysis showed that the commonly applied random k-fold CV underestimates the model prediction error compared to the spatial k-block partitioning (Figure 13). The spatial k-block CV divides the study area into k-spatial blocks; the algorithm is trained k-times using the data from k-1 spatial blocks to train and validate the predictions in the remaining spatial block. The model is ready when all spatial blocks have been used for training and validation. The critical point in this method is the block size.

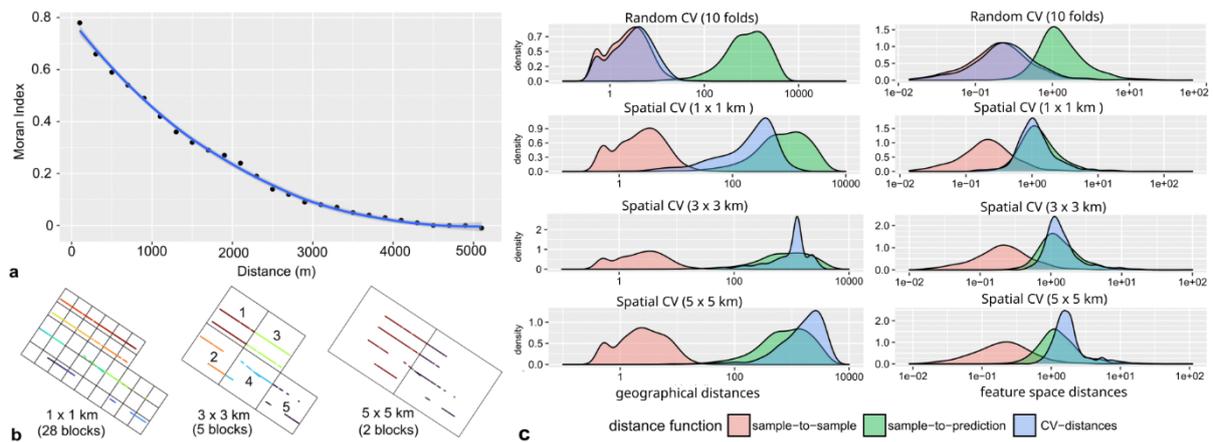
The block size (and consequently, the number of spatial blocks and training iterations) is a tradeoff between spatial autocorrelation range and model extrapolation (Roberts et al., 2017; Valavi et al., 2018). The optimum block size differs for each dataset, and should be optimized before the spatial modelling, following a three steps approach: First, the range, degree and significance of spatial autocorrelation are identified using the Incremental Moran Index analysis or/and variogram analysis (Figure 14); second, spatial blocks with dimensions close to the range of significant spatial autocorrelation are built, and the training data are assigned. Third, the density distributions of geographical and feature space distances between the training points, the CV blocks, are compared to the total geographical and future space area used for model predictions for each spatial block layout (Meyer and Pebesma, 2022). The spatial blocks with dimensions close to or beyond the range of significant spatial autocorrelation (e.g.  $I < 0.5$ ,

$p > 0.05$ ), that still adequately represent the geographical and feature space used in the spatial predictions, are selected.



**Figure 13.** a) Data splitting into ten random CV folds, b) The random forests CV performance overestimates the overall prediction performance (median  $R^2 = 0.93$ ), and it has a minimum variance among the random folds due to the almost identical samples. c) Data-splitting into ten spatial blocks. d) The model performance is poor (median  $R^2 = 0.13$ ) and varies between the spatial blocks. Figure source: modified from Gazis and Greinert, 2021 <https://doi.org/10.3390/min1111172>

For example, in research article III, the analysis showed that  $1 \times 1$  km spatial blocks yield overoptimistic results as the data still fall within the range of spatial autocorrelation. Spatial blocks that are too large (e.g.  $5 \times 5$  km) remove a large portion of training data in each training iteration, creating feature space gaps and causing unnecessary extrapolation that reduces modelling performance and results in over-pessimistic results (see feature space extrapolation). A  $3 \times 3$  km block size was the best choice for this dataset (Figure 14). The 'tradeoff' approach significantly reduces the effect of spatial autocorrelation but does not eliminate it, as the model residuals still exhibit low spatial autocorrelation values ( $I < 0.3$ ,  $p < 0.05$ ). These low values are not exclusively attributed to the limits of the spatial blocking approach but also to the missing predictor variables related to the spatial distribution of nodules (see Research Objective V).



**Figure 14.** a) Moran Index values at distances from 100 to 5100 m, with a step of 200 m. A locally estimated scatterplot smoothing (LOESS; Cleveland, 1981) with a 0.95 confidence interval has been plotted through the index values. b) Spatial block configuration for different block dimensions. The spatial blocks where data exist correspond to the spatial CV folds used during model training and evaluation. c) Geographical and normalized multivariate predictor space Euclidean distances among train dataset, CV folds, and predicted area for random-CV, spatial-CV (1 × 1 km), spatial-CV (3 × 3 km) and spatial-CV (5 × 5 km). Figure source: modified from Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

Research articles II and III showed that in addition to a better error estimation, a spatial k-block CV offers a better a) model interpretation and b) hyperparameter optimization:

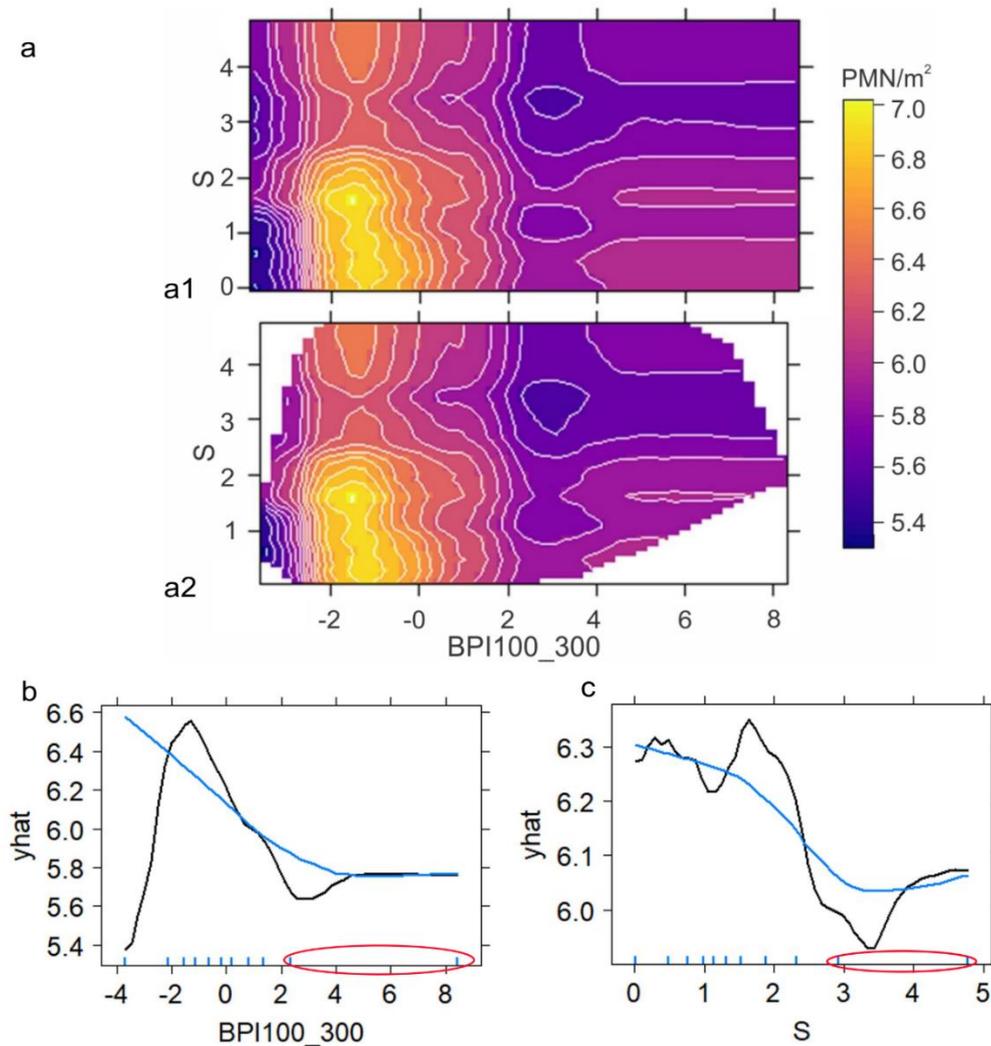
- a) While in random CV, the model performance is similar among the folds, in spatial k-block CV, the performance differs between two or more spatial blocks (Figure 13). Knowing that a spatial model is not predicting equally well within the study area is useful in the sense that it initiates further research to identify the reasons behind the poor performance in some sub-regions. Such information is beneficial in study sites with varying geomorphology and when different models are used and compared (i.e. identification of the spatial blocks where the predictions converge or diverge). When large differences exist between the spatial blocks, the prediction performance should not be averaged. Median values and box plot graphs should be preferred for the data representation instead.
- b) In random CV, where the hold-out samples are almost identical to the training samples, parsimonious predictors such as depth could be sufficient to fit the model. On the other hand, in spatial k-block cross-validation, more or all the available predictors are used in the model as the algorithm is based on the underlying relationships between predictors and response variables to gain predictive knowledge and reduce prediction error. The latter enhances the model interpretation, essential in researching nodule occurrences (see Research Objective V).

The spatial k-block CV has several advantages to spatial modelling but adds complexity to modelling procedure, error calculation and image data acquisition. Contrasting spatial blocks demands additional steps that increase the total time. An objective comparison of the local error from different spatial blocks demands an almost equal size of training and validation points for each spatial block. However, AUV surveys are typically/often not spatially balanced and dedicated to data acquisition for spatial predictive modelling, resulting in spatially imbalanced datasets. Further research should be conducted in this direction (e.g. Foster et al., 2014; Foster et al., 2020). The spatial k-block approach is considered the most conservative (pessimistic

approach) spatial CV method due to the feature space gaps that might be created during an arbitrary (and not optimized as here) spatial partitioning, particularly in sizeable heterogeneous study areas (Roberts et al., 2017; Hao et al., 2020; Wadoux et al., 2021). Nevertheless, this approach was preferred over other spatial methods such as spatial leave-one-location out (Meyer et al., 2018), Euclidean buffered distances (Hengl et al., 2018), nearest neighbour distance matching (Milà et al., 2022), spatial covariance weighting (Misiuk and Brown 2023) because of the dataset size and survey layout (i.e. 30,000 samples in linear AUV transects). The abovementioned methods have been designed and tested with fewer samples ( $\ll 1000$ ) distributed throughout the entire area and not in a linear AUV transect layout. In contrast, spatial blocking is appropriate for large datasets and regions, being computationally efficient (fewer training points per model iteration) and provides flexibility in partitioning the geographical space (Roberts et al., 2017; Valavi et al., 2019). The presence of geographical and feature space gaps represents better the error related to extrapolation and transferability over larger areas (Wegner and Olden, 2012; Roberts et al., 2017; Fourcade et al., 2018; Hao et al., 2020; Gazis and Greinert, 2021; Meyer and Pebesma, 2021). The local error analysis showed that the error magnitude in each spatial block is related to the feature space dissimilarity of this block compared to the rest of the training sample (see feature space extrapolation).

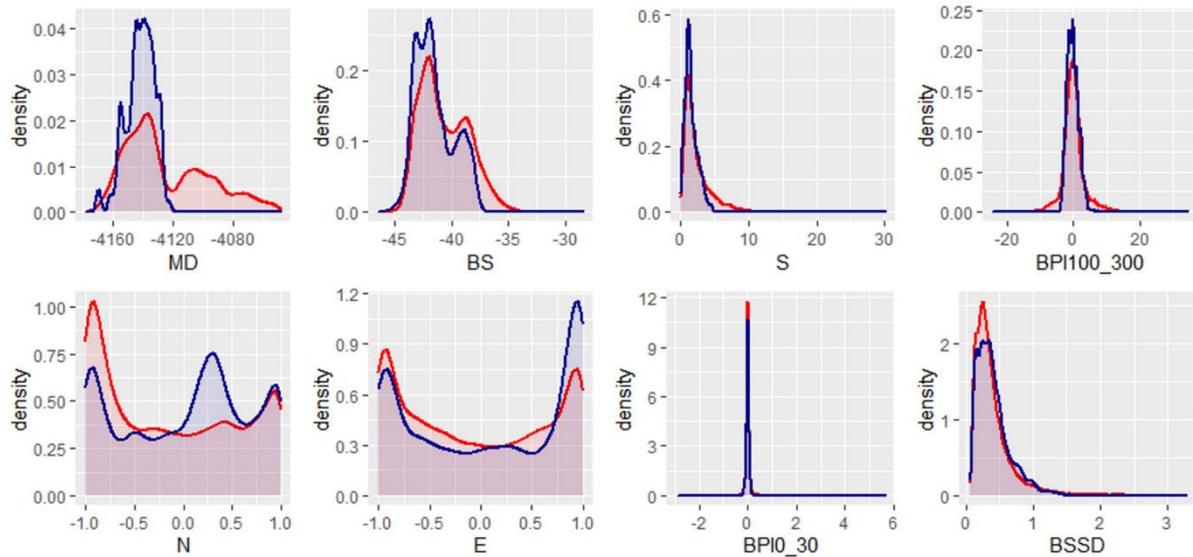
### **Feature space extrapolation**

Machine learning algorithms use data to capture the underlying relationships between response and predictor variables. The algorithm's training occurs within the convex hull, defined by the training data based on each predictor's univariate range and the predictors' inter-correlation structure (Yates et al., 2018). The developed model is applied to new geographical locations (i.e. spatial extrapolation), assuming that response-predictor relationships are stationary and remain beyond the calibration site. However, stationarity rarely holds in a heterogeneous seafloor like our study sites. Feature space extrapolation aims at predicting new data using predictor values outside the range, correlation structure, or value combinations of the existing training data (Elith et al., 2010; Zurell et al., 2012; Yates et al., 2018). Feature space extrapolation could also occur within the vacant regions of a convex hull (Figure 15), particularly in high-dimensional spaces (Kuhn and Johnson, 2013; Barbiero et al., 2020). Applying a machine learning model in a new geographical area where the feature space differs can result in poor and unreliable predictions (Mesgaran et al., 2014; Meyer and Pebesma, 2021). It is well-documented that the prediction error does not necessarily increase with the geographical distance from the training points but due to the environmental dissimilarity between reference (i.e. training sample) and predicted conditions (Meyer and Pebesma, 2021). Thus, there is a need for dissimilarity analysis between the training data and the new area where the predictions will be made (Ludwig et al., 2023).



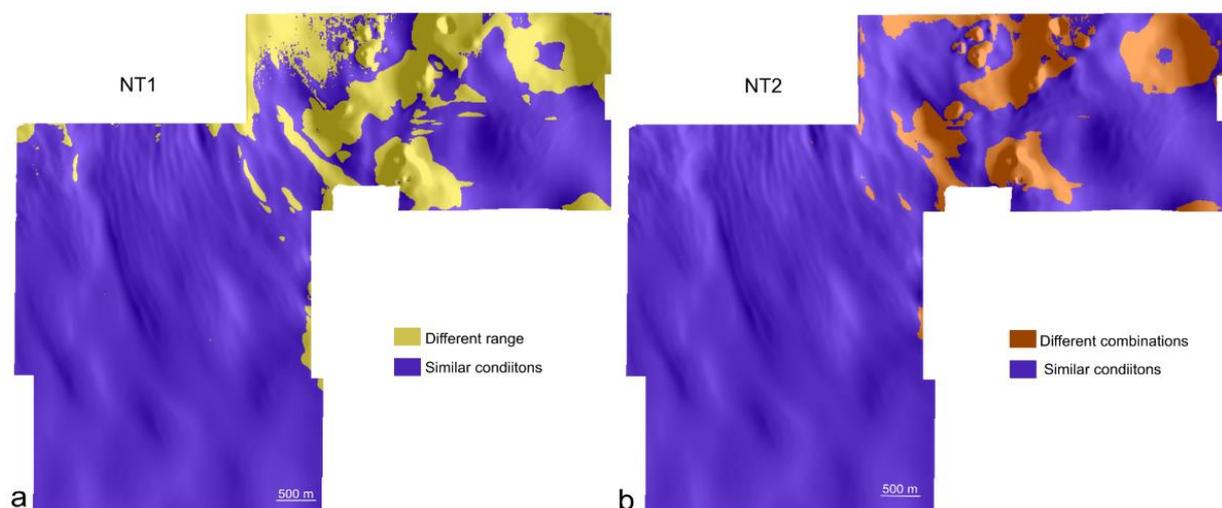
**Figure 15.** a) PDPs between slope (S), broad-scale (100-300 m) bathymetric position index (BPI100\_300), and image-derived number of nodules per square meter (PMN). a1) The PDP shows model predictions for the entire range of the new feature space. a2) The PDP shows model predictions only within the convex hull of the two variables as defined by the training data. b and c) PDP between BPI100\_300 and PMN showing the quantile regression forests (Meinshausen, 2006) model response (black line) and LOESS model response (blue line). These two response curves show the overall agreement and local differences between a random forests algorithm and a simpler non-parametric regression model. The blue ticks on the x-axis represent the data deciles. The red circles highlight the gaps in the feature space where no training samples exist, and the models extrapolate either with a constant (b) or monotonic (c) response. These gaps are shown in white in Figure a2. Figure source: modified from Gazis and Greinert, 2021 <https://doi.org/10.3390/min11111172>

Traditional non-spatial statistical approaches, such as visual inspection (e.g., boxplots, density plots, histograms), non-parametric tests (Kullback–Leibler divergence, two-samples Kolmogorov – Smirnov test) can show the differentiation between the two examined distributions. However, they cannot identify where these differences occur (Figure 16). When several predictors are used (typical scenario), the number of possible pair-wise comparisons (to examine the correlation structure) can quickly become large, making detailed analyses time-consuming and difficult for more than two variables.



**Figure 16.** Density plots show the difference in the univariate density distribution range of the spatial predictors between the training samples (blue) and the area where the model is applied (red). Figure source: Gazis & Greinert, 2021 <https://doi.org/10.3390/min11111172>

Spatial sample-based methods have been developed to address the abovementioned limitations. Multivariate Environmental Similarity Surfaces (Elith et al., 2010), Environmental Overlap (Zurell et al., 2012), Mobility-Oriented Parity (Owens et al., 2013) and Extrapolation Detection (Mesgaran et al., 2014) compare the univariate distribution range of each predictor and examine the presence of novel inter-correlations/combinations among the predictors using Euclidean or Mahalanobis distance metrics (Figure 17).



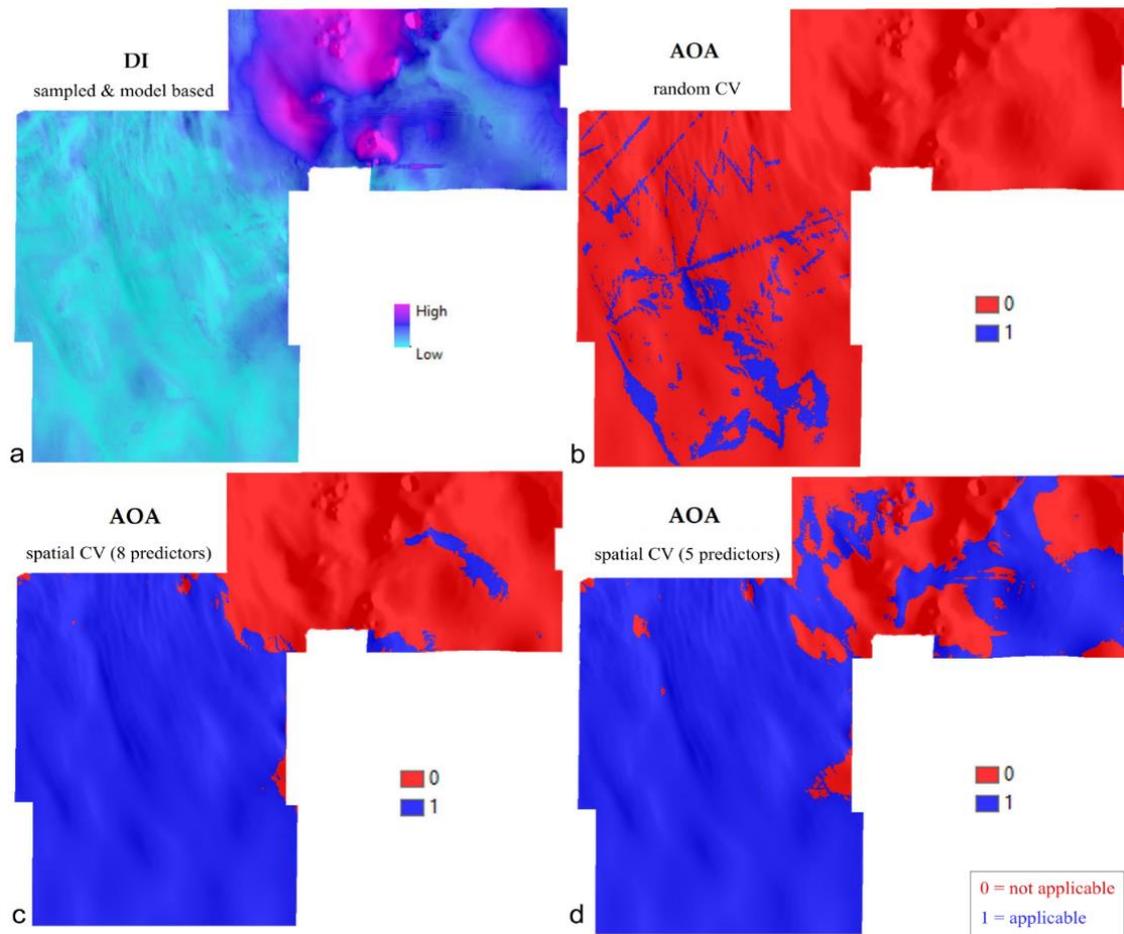
**Figure 17.** Sample-based extrapolation detection surfaces show a) the areas where at least one predictor variable is outside the univariate range of training data (aka NT1) and b) the areas where there are new combinations of predictor variables not existing in training data (aka NT2). The training locations are shown in Figure 13. Figure source: Gazis, 2025 (PhD dissertation).

While these methods retain the spatial component of the data, allowing for a combined geographical-feature space analysis, they do not account for the fact that deviations in the value range or relationship type in some predictors are more impactful to the model than others. In other words, they are not model-based, which is needed in a machine-learning application.

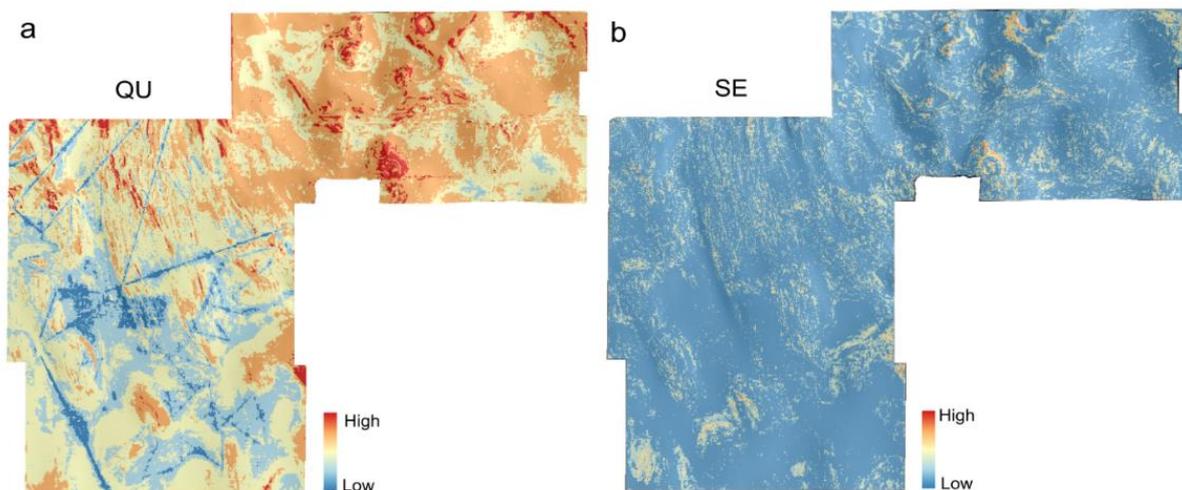
Meyer and Pebesma (2021) introduced the combined sampled- and model-based Area of Applicability (AOA) to address this issue. AOA assesses geographical areas with new feature space conditions where the prediction error of a given model exceeds the training CV error and thus being not representative and not trustworthy (Kuhn and Johnson, 2013). The AOA for each model is derived after calculating the Dissimilarity Index (DI) between the training dataset and the respective data of the new area (sample to pixel-based analysis). The DI uses the normalized Euclidean distance to the nearest training data point in the multivariate predictor space. Before the distance calculation, predictors are scaled and weighted according to their weights after the model is built. Thus, the distances of more important predictors account for the DI estimation more. A pixel in the new area falls outside the AOA when its feature space characteristics are beyond the 0.95<sup>th</sup> quantile (outlier removed) of the dissimilarity observed between the training data inside the CV folds.

Research articles II and III showed that:

- a) ***Feature space extrapolation occurs where sharp morphological transitions***, bedrock exposure, consolidated sediments, volcanic knolls and sediment furrows result in extreme derivative values and a new inter-correlation structure (Figure 18).
- b) ***The spatial block CV increases the AOA*** (Figure 18). Predictions derived from models trained with the typical random CV have the smallest AOA. The almost identical training points in CV folds and the over-optimistic error assessment limit the AOA to a small region around the training locations, where the prediction errors are comparable to the CV error. This finding agrees with the current literature (Ludwig et al., 2023; Meyer and Pebesma, 2021)
- c) ***Simpler models with fewer (but relevant) predictors increase the AOA*** (Figure 18). Removing predictors prone to overfitting increases the generalization performance, which other researchers have also mentioned (Mereow et al., 2014; Lauria et al., 2015; Moreno-Amat et al., 2015).
- d) The ***random forests prediction intervals cannot substitute DI and AOA***. The random forests prediction intervals can be calculated with different methods such as Monte Carlo simulations (Coulsto et al., 2016), infinitesimal Jackknife approach (Wager et al., 2014) and quantile regression forests (Meinshausen, 2006), with the latter to be the most commonly used (e.g. Diesing et al., 2020a; Fouedjio and Klump, 2019; Hengl et al., 2018; Kirkwood et al., 2017). They show the locations with increased prediction uncertainty but do not show the reasons behind this increased variation (e.g. univariate or multivariate feature space extrapolation, new inter-correlation structure). In addition, Meyer and Pebesma (2021) have shown that random forests prediction intervals are valid only for the data points enclosed within the convex hull of the training data, providing unrealistic and overoptimistic results outside the convex hull. Research article II also shows that there are sub-regions with high dissimilarity but no high analogue uncertainty (Figure 19).



**Figure 18.** a) DI between training samples and the study area accounting for the model weights. b) AOA for a model trained with random CV. The prediction error is comparable only around the training locations. c) AOA for a model trained with spatial CV. d) AOA for a model trained with spatial CV and fewer predictors. Figure source: Gazis & Greinert, 2021. <https://doi.org/10.3390/min11111172>



**Figure 19.** a) Random forest prediction uncertainty based on the quantile regression forests approach. The quantile uncertainty (QU) is minimized at and around the training locations and maximized in areas of higher DI being moderately correlated with the DI-derived surface ( $r = 0.65$ ,  $p < 0.05$ ). Nevertheless, areas of lower prediction uncertainty are within the non-applicable parts (see Figure 17). b) Random forests prediction standard error (SE) based on the infinitesimal Jackknife approach yields low standard errors for the entire study area, indicating over-optimistic results. Figure source: modified from Gazis & Greinert, 2021. <https://doi.org/10.3390/min11111172>

Based on the abovementioned findings, the conclusions are:

- 1) Feature space extrapolation occurs even in small but heterogeneous geographical areas. A thorough transferability assessment (e.g., sample- and model-based dissimilarity analysis) should precede the model application to a new area, notably if the new area lacks validation points. DI and AOA maps should accompany the spatial predictions and supplement the interpretation and decision-making process. In other words, the focus should not be on the model predictions but on how they were derived from the fitted model and available data.
- 2) AUV image surveys should aim at and achieve good geographical and feature space coverage and prioritize the latter when the data will be used for spatial predictive modelling. The acquired samples (i.e. images) should enclose the full range of environmental conditions as studies have shown that a poor representation reduces the predictive performance and model transferability in machine learning algorithms (e.g. Schmidt et al., 2014; Wadoux et al., 2019). DI and AOA could guide feature space sampling efforts and improve existing spatial models by prioritizing the most dissimilar sites in future AUV surveys. Progress in AUV hardware and software could allow for real-time generation of dissimilarity surfaces, improving the survey data collection through adaptive sampling. This conclusion initiated ongoing research for an optimized AUV path planning within the COMPASS project (see Research Objective I). In this direction, Shields et al. (2023) proposed an AUV-based future space sampling approach for a balanced geographical and feature space sample coverage.

**The work extends beyond the scientific community of deep-sea mining**, as machine learning is currently a widely used method in seafloor spatial modelling and habitat mapping (Misiuk and Brown, 2024). In the twenty years following the application of machine learning algorithms in marine habitat mapping, the focus has been on improving the model fit on a given dataset, mainly by using more complex models and predictors (Bowden et al., 2021; Misiuk and Brown, 2024). However, the spatial autocorrelation and feature space extrapolation should be considered in the modelling procedure and reported to avoid misleading spatial maps, independently of the machine algorithm used. By the time of writing research article II, random k-fold CV was the dominant approach in seafloor predictive modelling (e.g., Li et al., 2017; Diesing et al., 2017; Misiuk et al., 2018; Diesing et al., 2020a). **Research article II was the first to use spatial CV for random forests in seafloor predictive modelling with a continuous response variable (regression)**, only two other publications with seabed classification tasks considered this before (Misiuk et al., 2019; Dolan et al., 2021). In addition, **research article II was the first to use the DI and AOA in deep-sea environments and for nodule research**, only one application existed for shallower waters (Dolan et al., 2021). **Research article III was the first in seafloor modelling to extend the use of spatial block CV, DI and AOA in machine learning models other than random forests** (see Research Objective III). The latter was a call to the scientific community from the developers of the DI and AOA (Meyer and Pebesma, 2021).

**Research Objective III:** *Examine the performance of five commonly used machine learning regression algorithms for the spatial prediction of nodules: generalized linear models, generalized additive models, random forests, support vector machines, and neural networks.*

### **Key findings**

Research articles I and II used the random forests regression algorithm and quantile regression forests (see Research Objective II) as random forests algorithm and its variants are currently the "workhorse" in deep-sea predictive mapping (e.g., Lawson et al., 2017; Diesing 2020b; Uhlenkott et al., 2022; Josso et al., 2023) and marine habitat mapping in general (e.g., Wei et al., 2010; Diesing et al., 2017; Misiuk et al., 2018; Diesing et al., 2020a,b; Misiuk and Brown, 2024 and references therein). Research article III extended this analysis by adding four other commonly applied machine learning regression algorithms to build the following five models: generalized linear models (hereafter GLM), generalized additive models (hereafter GAM), random forests model (hereafter RFM), support vector machines (hereafter SVM), and neural networks model (hereafter NNM). To our knowledge, this is the first time that GAM, SVM and NNM were used for predictive nodule mapping and the first time that SVM and NNM were used for seafloor habitat mapping with a continuous response variable (regression), as literature research only provided applications for seafloor classification (Hasan et al., 2012; Diesing and Stephens, 2015; Alevizos and Greinert, 2018; Trzcinska et al., 2020; Janowski et al., 2021; Breyer et al., 2023; Haroon et al., 2023; Misiuk et al., 2024). The model evaluation included several aspects: predictive performance in test data, spatial and feature space extrapolation based on spatial block CV, hyperparameter optimization, complexity and interpretability.

### **Predictive performance in Test data**

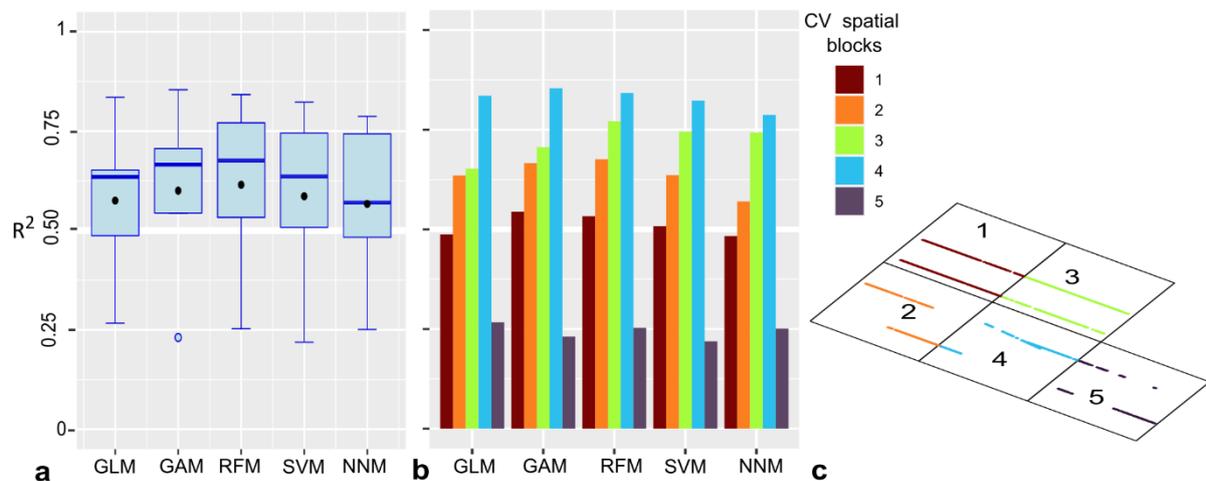
RFM had the best overall fit on the test dataset across all metrics used ( $R^2$ , adjusted  $R^2$ , MAE and RMSE), a finding that agrees with other studies showing the ability to outperform other algorithms in prediction error (Stephens and Diesing 2014; Diesing and Stephens 2015; Li et al., 2017). RFM reproduced the testing data distribution (e.g., mean, median, interquartile range) but not the minimum (min) and maximum (max) values (Table 1). The overestimation of the lower values and underestimation of the higher values is an inherent algorithm limitation due to averaging predicted values from each tree at the last calculation step (Breiman, 2001a). Similar to RFM, all other models did not manage to predict the maximum test value but were closer to the initial min-max range. Due to the linear extrapolation, GLM was the only model that predicted beyond the minimum value of the test data. NNM had the second-best predictive performance in the test dataset but the worst in spatial-block CV (see section spatial and feature space extrapolation). All models were run 30 times, having a stable predictive performance in all metrics. The residuals from all models were normally distributed at around zero, with low spatial autocorrelation ( $I < 0.3$ ,  $p < 0.05$ ), showing the absence of prediction bias in lower or higher values and implying that the errors are mainly due to the inherent variability of predictor and response variables as well as missing predictor variables (see Research Objective V).

**Table 1.** Model performance metrics on test dataset for the back-transformed predictions. Table source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

Model	R <sup>2</sup>	MAE	RMSE
GLM	0.72	2.80	3.92
GAM	0.75	2.60	3.68
RFM	0.84	2.04	2.94
SVM	0.75	2.44	3.63
NNM	0.76	2.54	3.56

### Spatial and feature space extrapolation

The analysis showed that less complex models such as GLM, GAM and RFM had better resampling performance in a spatial-block CV than SVM and NNM, indicating a better generalization ability when a big fraction of the training data is kept out (Figure 20). This fact is attributed to the smoother functions that are more transferable (and interpretable) over larger areas (Wengner and Olden, 2012; Lauria et al., 2015; Stienessen et al., 2021). RFM had the best average performance in CV folds, followed by GAM, which had the smallest interquartile of all models. The model performance varied among the spatial blocks in all models. All models had the lowest performance in the same spatial block. In the used data set, this spatial block was in the southeast of the study area (block 5 in Figure 20c), having complex geomorphology that differed from the rest. The presence of horst, cliff, trough, exposed indurated sediments, and different nodule facies resulted in extreme derivative values, model extrapolation and thus higher error.

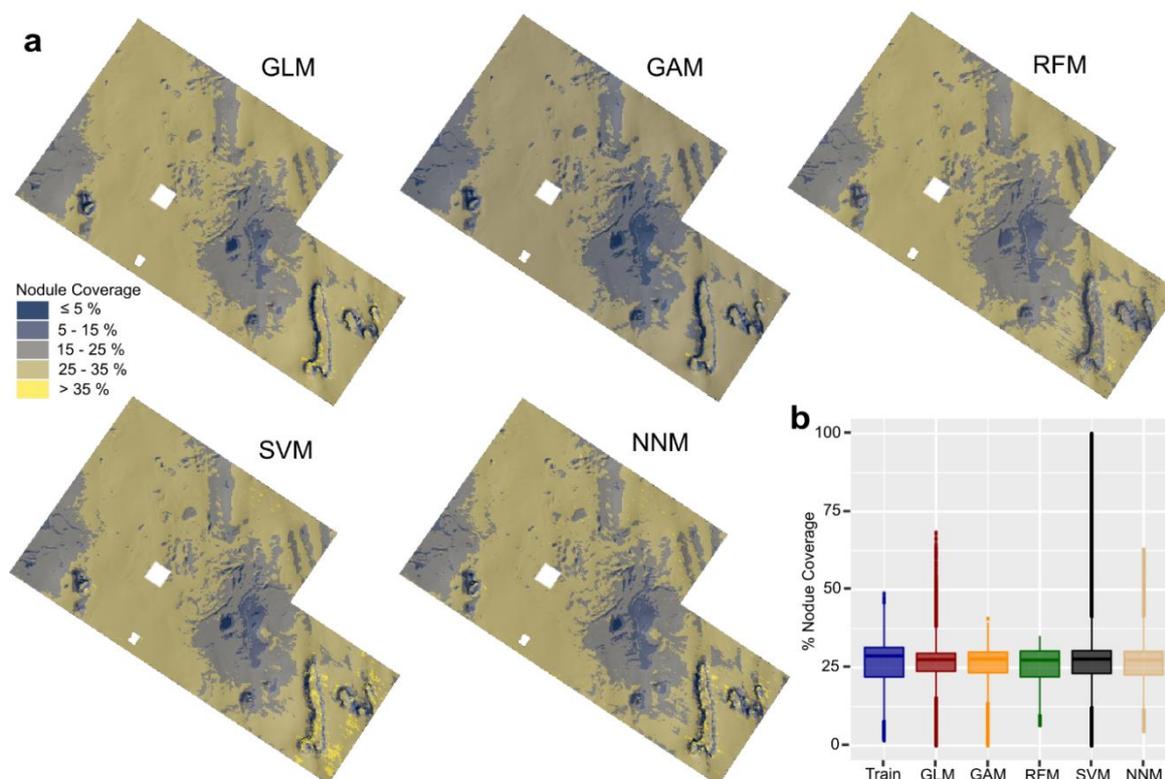


**Figure 20.** (a) CV performance for each model. The median (thick blue line) is higher than the mean (black dot) in all models apart from NNM, as the low performance in one block pulls down the mean. (b) In-fold performance for each spatial block and model. All models performed similarly in each CV spatial block. (c) Spatial orientation of the CV spatial blocks. Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

### Predicted surfaces of the entire study area

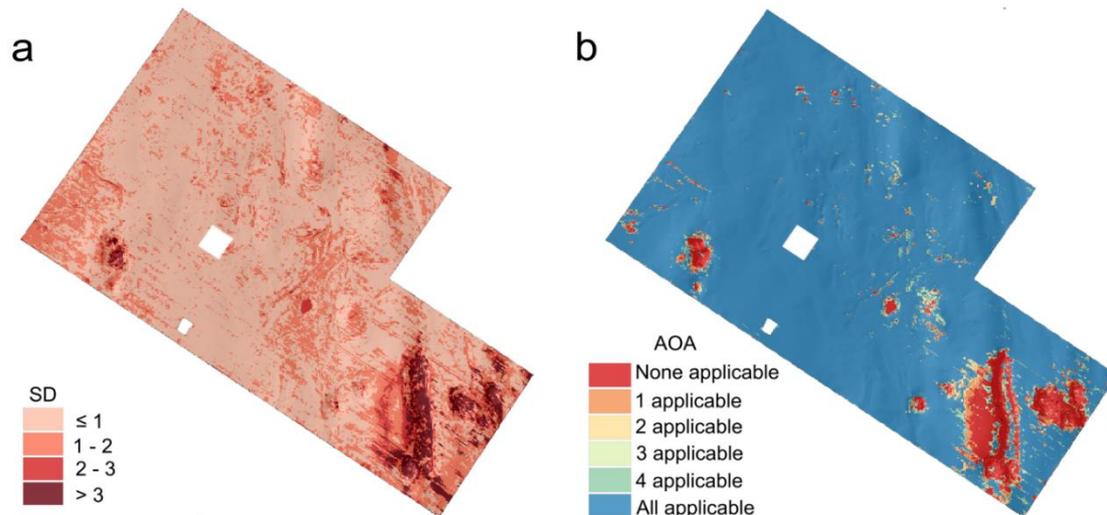
All models yielded similar spatial predictions on nodule coverage for most of the study site (Figure 21). The absolute decimal values differ marginally, but the spatial trends on predicted surfaces are similar. In less than 1 % of the total predicted surface, the non-tree-based core of GLM, GAM, SVM and NNM (i.e., linear, penalized B-splines, 3rd-degree polynomial and sigmoid activation function) allowed for a value range extrapolation beyond the value range of

the training data; with SVM reaching up to 100 % nodule coverage, although such high values were not present in training data. However, such high values were limited to < 1% of the study area.



**Figure 21.** a) Maps of the spatial predictions of nodule coverage and range of predicted values over the entire study area for each ML model. All models show that along a north-south axis at the west part of the study area, the seafloor is homogeneously covered with nodules (20–30% coverage). Higher coverage (> 35%) exists in the southeast but is somewhat localized. The central and deeper part has lower nodule coverage (15–25%) that progressively decreased, reaching almost zero values at the deepest part of the basin. Lower coverage also existed at the northwest corner (crater slope). b) Boxplots of nodule coverage predicted values for the entire study area. Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

A pixel-by-pixel comparison of model predictions showed the spatial agreement, expressed here as the standard deviation (SD) of the predicted values from all models for a given pixel (Figure 22). Overall, 68% of the study site has SD values < 1, indicating good agreement between the different models. Differences occur in geographically small sub-areas associated with sharp morphological transitions, bedrock exposure, consolidated sediments, and volcanic knolls, with the highest SD within the trough at the southeast part of the study area (Figure 22). The combination of deeper parts with steep slopes (causing lower image coverage) and high backscatter reflectance (partly attributed here to nodule coverage and partly to hard substrate), has led to contradictory predictions: RFM predicted low nodule coverage, while SVM yielded high nodule coverage. Such areas are outside of the AOA for all models. All models could be applied to 88% of the study area (pixel-by-pixel comparison regarding the pixels where all models applied simultaneously); for 6% of the area at least one model was applicable and for another 6% of the area no model was applicable (Figure 22).



**Figure 22.** a) Combined SD of the predicted values from all models. b) Combined AOA from all models.  
Figure source: modified from Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

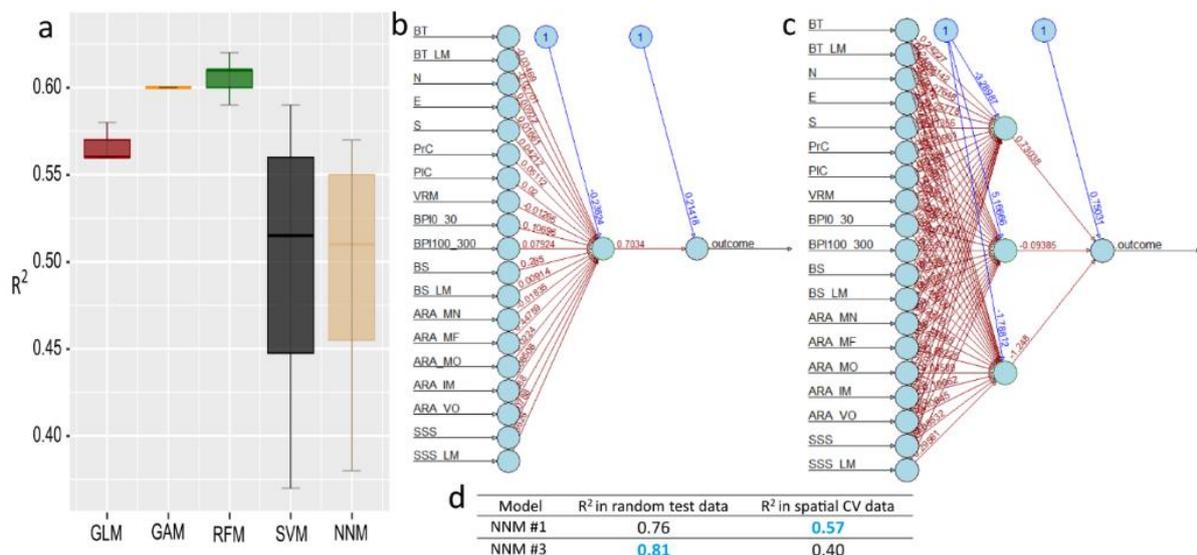
### Hyperparameter optimization and complexity

Machine learning algorithms have several hyperparameters that influence the model complexity, computational speed, fit on the training data, and predictive performance on new data. Several studies have applied machine learning algorithms across numerous benchmark datasets, examining different values for each hyperparameter and combinations of hyperparameters. Investigating different combinations often uses exhaustive techniques such as random search, grid search, simulated annealing, and genetic algorithms (Yu and Zhu, 2020; Bischl et al., 2021). Such techniques search the multidimensional space created by the hyperparameters (aka hyperparameter matrix) and get the best possible set of values, i.e., the best global solution that accounts for all joint effects. Six outcomes of these studies are: a) there is no ideal hyperparameter set-up for each machine learning algorithm that can be applied across different datasets, b) different hyperparameter search techniques could result in a different optimal hyperparameter set, c) not all algorithms are affected by the optimization process to the same degree, d) not all hyperparameters are of the same importance for each algorithm and e) in most cases, the default hyperparameter values in each software package provide an overall good solution and starting point (see also Probst et al., 2019a,b; Huang and Butros 2016), f) the time and computational power needed to perform an exhaustive search is a burden for most real case scenarios and datasets, with many researchers not having access to high-performance computing clusters that could handle this issue. An additional practical limitation is that not all software and statistical packages offer the option to tune all hyperparameters of each machine learning algorithm, with the majority focusing on the most influential for each algorithm (Probst et al., 2019b, Fernández-Delgado et al., 2019; Bischl et al., 2021). Moreover, some hyperparameters are dependent on "superordinate" hyperparameters and are tunable only if the "superordinate" hyperparameter is set to a specific value (Hofner et al., 2012; Probst et al., 2019a).

Considering all the above, a non-exhaustive but still detailed hyperparameter optimization was done for all of the five used models. In GLM, the number of boosting iterations was optimized. In GAM, the number of boosting iterations and the degrees of freedom in P-splines were

optimized. In RFM, the number of predictor variables randomly sampled at each node split, the split rule and the minimum node size were tuned. The kernel type, polynomial degree, scale, and cost hyperparameters were optimized in SVM. The number of hidden layers, neurons and activation functions were optimized in NNM. The optimization process occurred during the spatial CV, which is the preferred method when the data have spatial autocorrelation (Schratz et al., 2019).

The results showed that the default values are a good choice for GLM and GAM. No further improvement was made by increasing the number of boosting iterations and the degrees of freedom in P-splines. RFM default values regarding the number of predictor variables randomly sampled as candidates at each node split ( $p/3$ ) and split rule (variance) were the best for this dataset. However, the default value for the minimal node size to split was small for such a large training dataset (5 observations  $\approx 0.02\%$  of the training dataset), resulting in deep trees, which increased computational time and slightly reduced spatial transferability. A larger value ( $\approx 4.9\%$ ) was selected. SVM had the biggest variance and improvement during the hyperparameter optimization (Figure 23), a finding which agrees with previous comparative studies (Probst et al., 2019a). In NNM, more neurons and hidden layers resulted in better performance in random test data but an increased error in spatial blocks, indicating spatial overfitting (Figure 23). The model training and evaluation using spatial CV requires a good generalization ability, which simpler models typically offer. The simplest model was selected between two NNM with equal performance, considering the computational time and complexity following Occam's razor principle for model interpretability (Vandekerckhove et al., 2015).

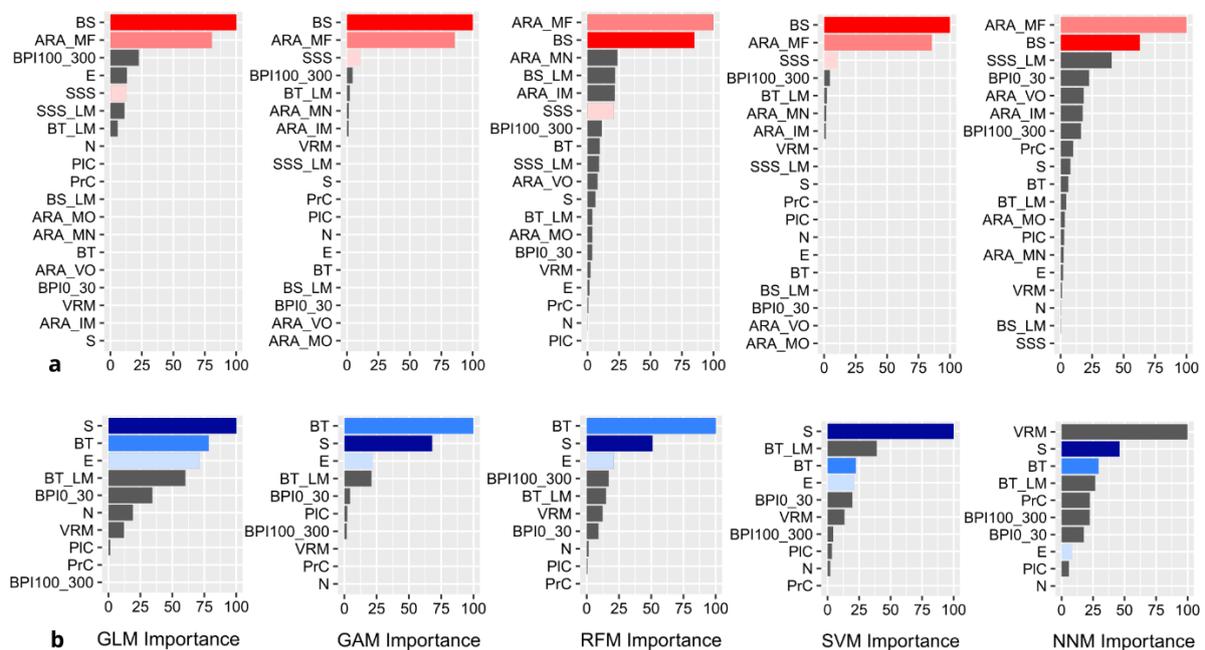


**Figure 23.** a) Boxplot of model performance on spatial CV data during the hyperparameter optimization. b and c) Visual representation of NNM with one hidden layer, one neuron (b), and three neurons (c). The model weights and biases are also shown. d) As the NNM becomes deeper, it predicts better the similar random test data, but worse the spatial block CV data indicating spatial overfitting. Figure source: Gazis, 2025 (PhD dissertation).

### Variable Importance (VI)

Each model estimated VI differently: In GLM, the absolute value of the coefficients from the final model was used (Kuhn, 2022). In GAM, the in-bag risk reductions per boosting iteration (the error difference between the current and the previous step for each base learner contained in the model) were calculated (Hothorn et al., 2021). In RFM, the average permutation VI from all trees was used (Breiman, 2001a). In SVM, there is no model-specific way to calculate the importance. Instead, a permutation approach was used, where the prediction error was measured before and after shuffling the values of the predictor variable. The larger the prediction error after shuffling, the higher the importance of this predictor variable. NNM assigns weights to each predictor to reduce the prediction error. These weights were used to derive VI based on the Garson algorithm (Garson, 1991; Goh, 1995). All VI outputs were scaled between 0 and 100 for better comparison across the different models (Kuhn, 2022).

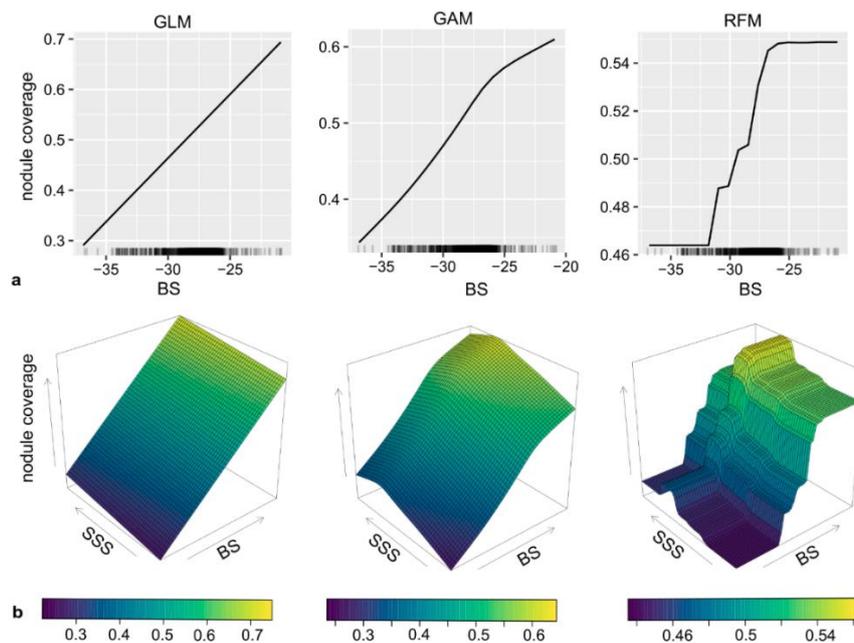
Despite the different ways to calculate VI, all models ranked the predictor variables similarly (Figure 24). MBES backscatter mosaic (BS), MBES backscatter from incident angle range of 25° to 55° (ARA\_MF), and SSS backscatter were the three most important predictors in all models (see Research Objective IV). All models, but particularly GLM, GAM and SVM, heavily relied on the hydroacoustic backscatter-related predictors and their derivatives for error minimization, shadowing the role of the seafloor geomorphology (Figure 24). The contribution of bathymetry (BT), slope (S), broad-scale BPI (BPI100\_300) and eastward slope orientation (E) became evident when the hydroacoustic backscatter-related predictors were not considered (see Research Objective V).



**Figure 24.** a) VI for each machine learning model trained with all available predictor variables. BS and ARA\_MF scored first in all models. b) VI for each model trained without the hydroacoustic backscatter-related predictors. For the abbreviations of predictor variables, please see the corresponding research article. Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

### Which algorithm to select?

The above described analysis showed that different machine learning algorithms yield different results because they capture the underlying relationships between predictors and response variable and the hidden relations among the predictors to a different extent and in different ways (Figure 25). The best fit and tradeoff between model performance and complexity are unknown a priori in real data (Kuhn and Johnson 2013). According to the *No Free Lunch theorem* in supervised statistical learning (Wolpert 1992; Schaffer 1993; Wolpert 1996; Wolpert 2002), there is not a universal best algorithm ('silver bullet algorithm') across all data-driven predictions on non-uniform data, and therefore, one needs to test several algorithms in each dataset (Shalev-Shwartz and Ben-David 2014; Sterkenburg and Grünwald 2021).

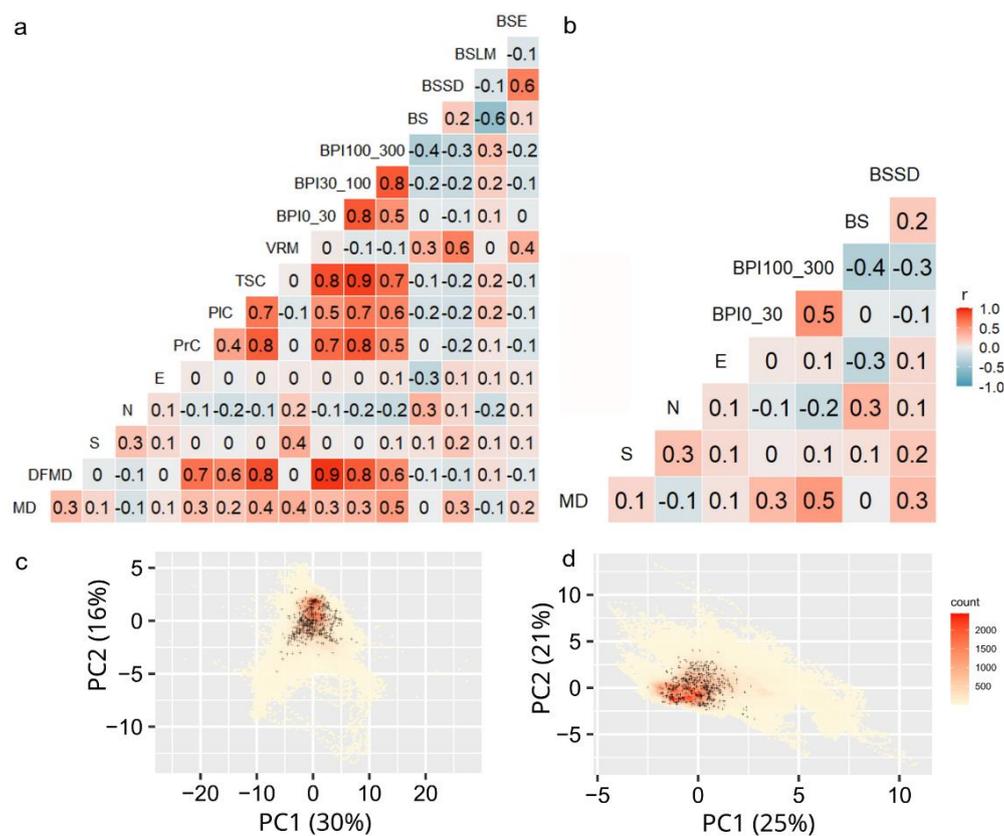


**Figure 25.** a) PDPs showing the relationship between nodule coverage (transformed values) and MBES BS for GLM, GAM and RFM. As the model complexity increases (left to right), the model captures more subtle relationships, resulting in less smooth regression lines. The black marks on the x-axis show the nodule coverage distribution concerning MBES BS values. b) PDPs showing the regression plane between nodule coverage, MBES BS and SSS. Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

In agreement with existing knowledge (e.g. Kuhn and Johnson, 2013), this study showed that the machine learning algorithms, independently of their core architecture, result in poor predictions when there is one or more of the following issues:

- (1) inadequate data pre-processing, such as acquisition artefacts (see Research Objective I), improper data georeferencing and scaling, or insufficient hyperparameter optimization, particularly in non-tree-based methods such as SVM and NNM (see Research Objective III)
- (2) spatial over-fitting when spatial autocorrelation is present and an inadequate model validation occurs (see Research Objective II),
- (3) extrapolation of the training feature space (see Research Objective II),
- (4) use of redundant, noisy or irrelevant predictors that unnecessarily increase the input dimensions and noise: Machine learning model complexity heavily depends on the number of input dimensions. Many dimensions need more training data, which could result in overfitting, loss of generalization ability, decreased interpretation (when multicollinearity occurs) and

increased computational time (Kuhn and Johnson, 2013; Guyon et al., 2003). The decrease in input dimensions can be achieved either with feature space extraction or feature selection methods (Alpaydin, 2020). The feature space extraction methods transform a high-dimensional space into a lower-dimensional space that still captures the most information (e.g., Principal Components Analysis) and can increase the model's predictive performance (e.g. Janecek et al., 2008). However, interpreting the transformed dimensions becomes complex as the principal components replace the original predictor variables. When performing feature selection on the other hand, the original variables are kept, and the number of input dimensions is reduced based on statistical, model-based approaches and domain knowledge (Li, 2019). The feature selection approach was selected to retain the original variables and enhance model interpretation: removal of zero variance predictors, Boruta ranking (Kursa and Rudnicki, 2010), and correlation filtering were used in the three research articles (e.g. Figure 26).



**Figure 26.** Correlation plots before (a) and after (b) the exclusion of correlated variables ( $r > 0.5$  and  $r < -0.5$ ,  $p < 0.05$ ). Both plots use the same scale. Between two correlated predictor variables, the variable with the highest Boruta score and domain relevance remained. For the abbreviations of predictor variables, please see the corresponding research article. The feature space and training locations before (c) and after (d) the exclusion of correlated variables are also provided. Figure sources: a and b) Gazis and Greinert, 2021 <https://doi.org/10.3390/min1111172>. c and d) Gazis, 2025 (PhD dissertation).

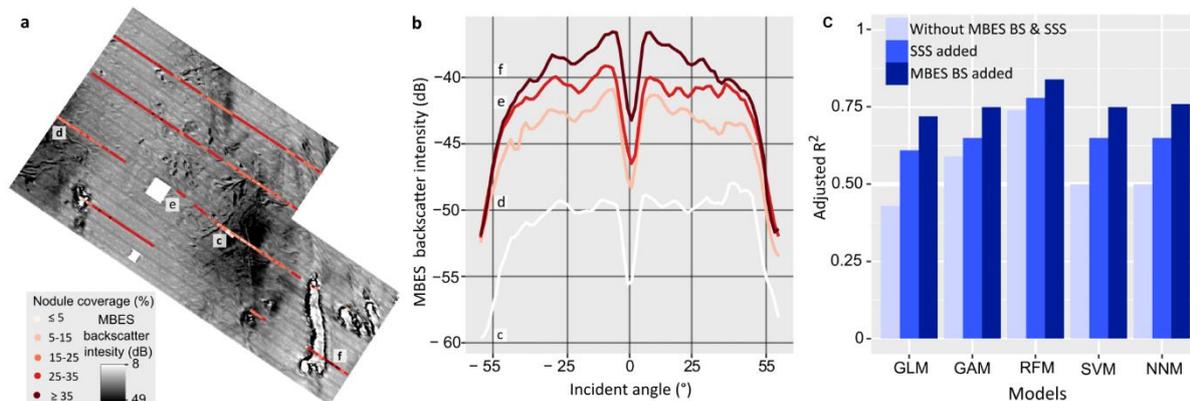
However, one should be aware that there are cases where irrelevant and artificial predictors increase predictive performance (e.g., spurious correlation), outperforming models even with domain-relevant predictors (Fourcade et al., 2018; Behrens and Viscarra Rossel, 2020). However, such models lack a causal interpretation, a vital aspect of marine (geo) sciences.

**On the road to the best model**, researchers should start by investigating the relationship between predictors and response variables (e.g. linear or nonlinear) using a thorough explanatory data analysis (e.g. distribution correlation tests, zero-variance predictors and graphs). It is vital to have a clear overview of each model's pros and cons and to know the exact target question and output result for the developed model. Does the model need to reproduce a specific test dataset and be applied to a similar area, or will it be used over a more extensive and dissimilar region (see also Research Objective II)? The ability of complex methods to create subtle regression lines results in a better fit in training data and (similar) test data but loses generalization ability if the new data differ significantly (spatial blocks), as noticed with the NNM that outperformed GAM. Moreover, researchers should be aware that while the primary interest of predictive modelling is to generate accurate predictions, it is equally important to understand how predictions occur. Ideally, the model with the highest performance metrics, the lowest complexity, and the best interpretability is chosen. However, the most accurate models are usually more complex and less interpretable. A tradeoff between interpretability and accuracy is thus needed, particularly in natural sciences, where the models are also used for causal explanation (Breiman, 2001b; Shmueli, 2010). Interpreting machine learning results is difficult, especially when domain knowledge or conceptual understanding is still developing and complex methods are used exclusively (e.g., neural networks; Merow et al., 2014).

In this direction, a multi-model approach offers the advantage of comparing and combining complex and simple models, finding the best model for the respective task. **The largest benefit of a multi-model approach is the confidence in areas (spatial and feature space) where the model predictions agree** (Figure 21). In addition, the contribution of each predictor can be evaluated across different models, providing better insights into their importance (Figure 24). **The model agreements on variable important ranking provides confidence in interpreting the role of each predictor** (see Research Objectives IV and V). To conclude, a multi-model approach should be prioritized. However, if there is no option for a multi-model approach (e.g., a need for real-time predictions on board a research vessel), then a good starting point would be to use an RFM. RFM demands the minimum data pre-processing (e.g. log-transformation, scaling), and the default hyperparameter values perform well. Furthermore, several habitat mapping studies use them already, giving additional confidence in their applicability (e.g. Li et al., 2016; Li et al., 2017; Diesing 2020a). RFM can capture linear, monotonic and nonlinear relationships and account for hidden interactions (Wright et al., 2016), enabling an adequate handling of real-life data. Finally, RFM provides variable importance without the need for an external function.

**Research Objective IV:** Analyze the MBES backscatter response of seafloor parts with different nodule coverage and study the contribution of different incident angles (angular range analysis) and backscatter components (impedance and volume) on the modelling performance.

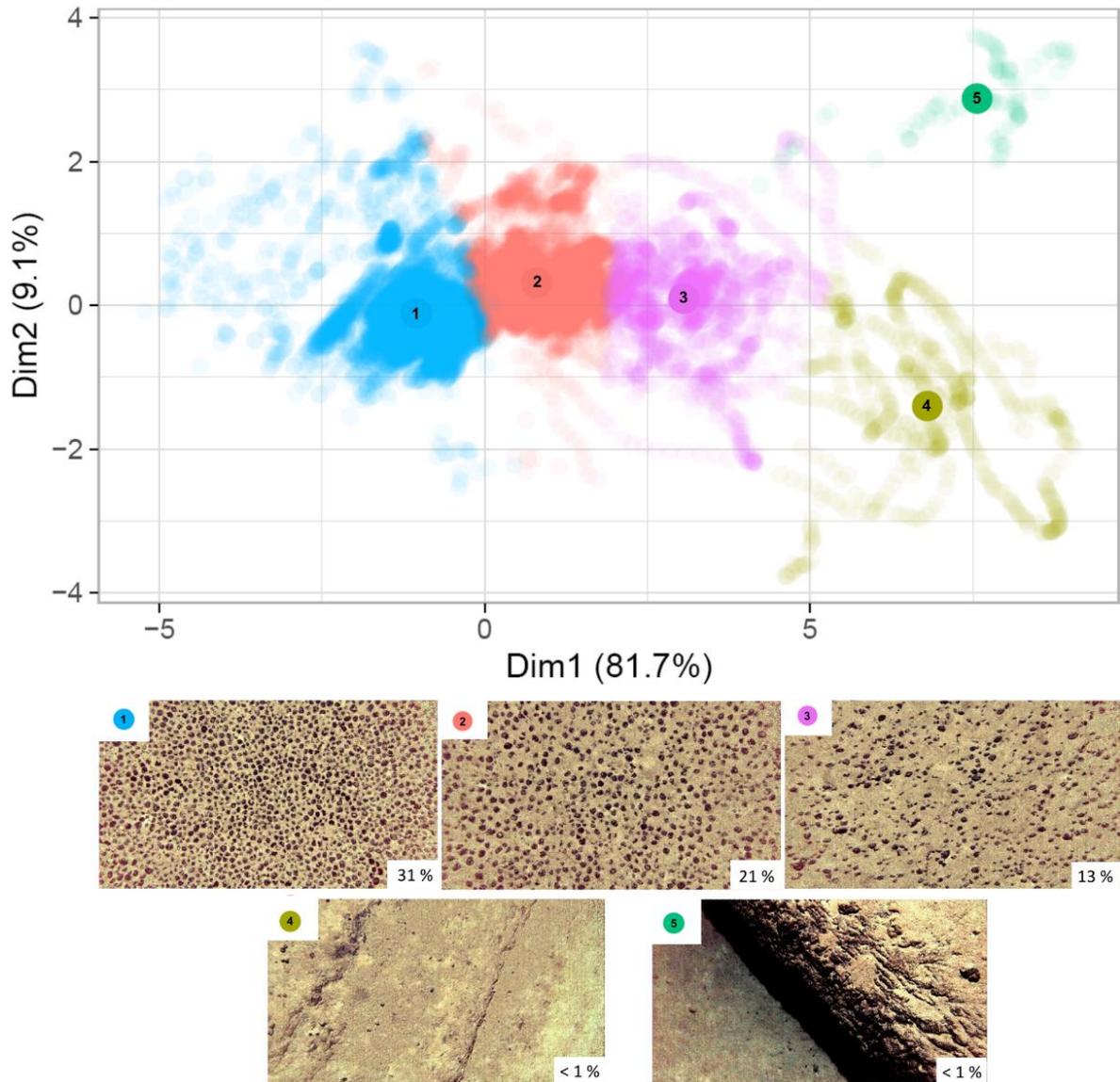
Research articles II and III showed that **MBES backscatter is the best spatial predictor for nodule coverage**. MBES backscatter scored first in all machine learning models (Figure 24). Independently of the model used, the prediction performance peaked only when the MBES backscatter and its derivatives were included as spatial predictors (Figure 27). The backscatter analysis showed a distinct difference ( $> 10$  dB) in backscatter intensity from seafloor parts with and without nodules. Moreover, the backscatter reflectance could discriminate seafloor parts with a 10 % variation in nodule coverage (Figure 27). The backscatter discriminative ability, even over transitional areas, is attributed to the combination of high frequency (400 kHz), low altitude survey (60 m) and small beam opening angles ( $0.7^\circ \times 0.7^\circ$ ), which resulted in a small footprint size ( $0.73 \text{ m} \times 0.73 \text{ m}$  in nadir on a flat seafloor) and short wavelength ( $\sim 4 \text{ mm}$ ), which was in the order or even smaller than the nodules' dimensions (mm to cm scale).



**Figure 27.** a) Post-processed MBES backscatter mosaic of the study area in research article III and image-derived nodule coverage. b) The MBES backscatter angular range analysis (ARA) from seafloor patches with different nodule coverage shows a clear difference in backscatter intensity for the entire range of incident angles. c) Barplots of the prediction performance of each model using only the MBES-derived geomorphological predictors (light blue), with added SSS backscatter intensity (medium blue) and with added MBES backscatter (dark blue). GLM had the largest improvement when the SSS and MBES backscatter data were added due to the monotonic/quasi-linear relationship between SSS and MBES backscatter with nodule coverage (see Figure 24). Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

Partial dependence plots (PDPs) of all models showed a monotonic relationship between backscatter and nodule coverage. The monotonic relationship has a different degree of smoothness for each model, but in all cases, it depicts higher backscatter reflectance with higher nodule coverage (Figure 25). It should be mentioned that high backscatter values could also be due to outcropping basement in small parts of the study area. This is a common issue in backscatter surveys within nodule fields, but it is limited in spatial extent and thus its contribution to the total backscatter (e.g. Yoo et al., 2018). The clustering of the feature space defined by the nodule coverage and the three most important backscatter-related predictors depicts this issue well in Figure 28. Spatial masking of such sub-areas/training data (here cluster 5) before modelling and the model's ability to learn from the geomorphological

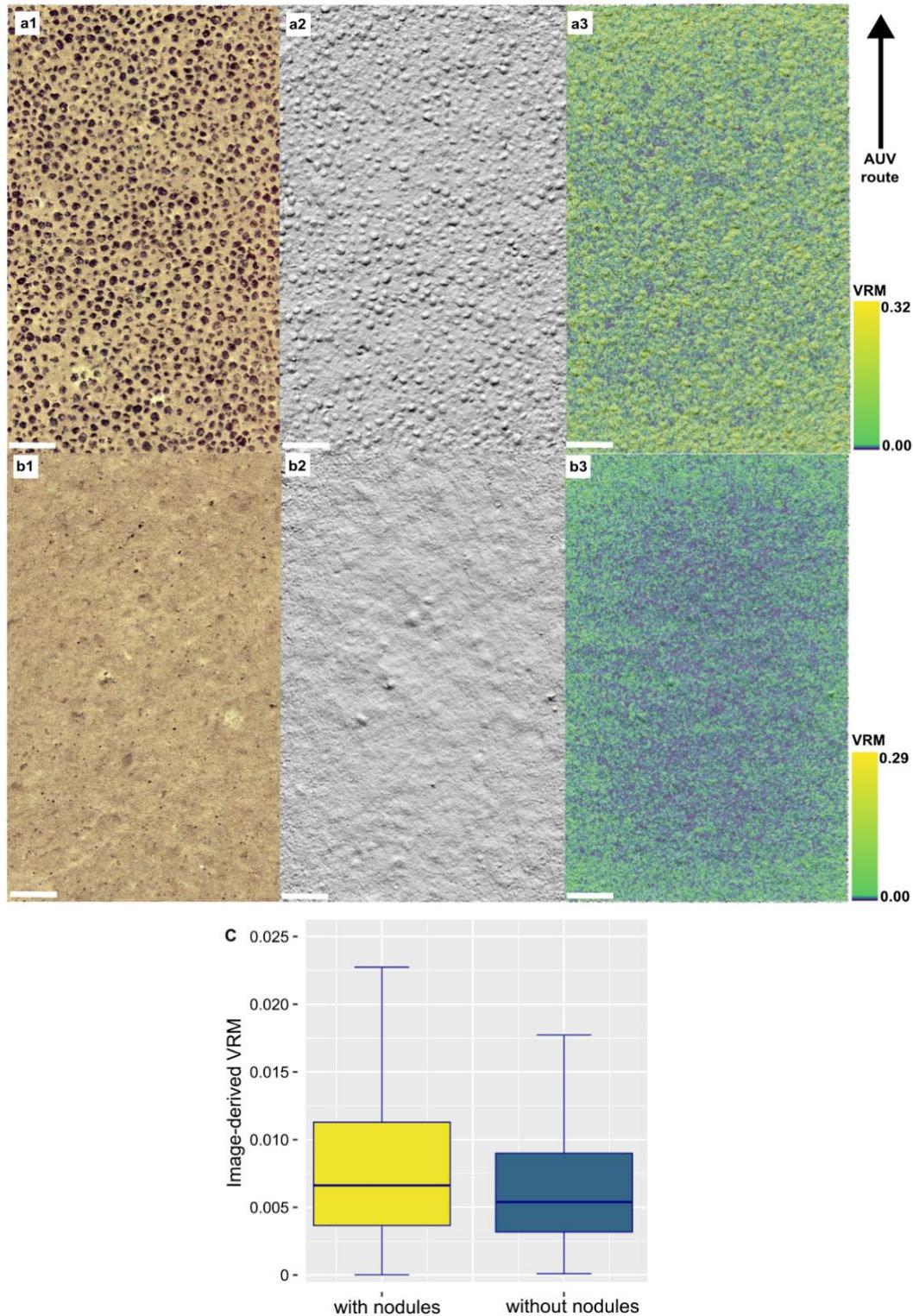
predictors (the geomorphology and its derivatives change at these locations) minimises this issue.



**Figure 28.** Feature space clustering of nodule coverage and the three most important backscatter-related predictors: MBES backscatter mosaic, MBES backscatter from incident angle range of  $25^{\circ}$  to  $55^{\circ}$ , and SSS backscatter mosaic. After centring and scaling the variables, the k-means++ algorithm (Arthur and Vassilvitskii, 2007) and the within-cluster sum of squares criterion yielded five clusters as the optimum feature space partitioning. The three primary (and cohesive) clusters (1-3) correspond to seabed images with varying nodule coverage (ca. 10% variation). Cluster 4 represents seabed images (locations) without nodule coverage on a flat seabed (low backscatter values). Cluster 5 (outlier cluster) shows seabed images without nodule coverage but on seabed parts with consolidated sediment and outcropping hard substrate (high backscatter values). The coloured circles are the centres of each cluster, and the images presented are the nearest to these centres. Figure source: Gazis, 2025 (PhD dissertation).

Older studies have described the effect of seabed roughness created by nodule occurrences on sound waves and the measured MBES backscatter intensity, particularly for low-frequency systems (Magnuson, 1983; Weydert, 1985; Moustier, 1985; Moustier, 1986; Weydert, 1990). The technological developments in underwater imagery now allow to create image-derived bathymetric surfaces and calculate bathymetric derivatives, such as roughness, similar to

MBES data but in mm resolution. This mm-scale seafloor roughness of nodule areas was quantitatively shown for the first time (Gazis et al. 2024) using image-derived roughness surfaces from areas with and without nodule coverage (Figure 29).



**Figure 29.** a1) Orthophoto-mosaic (2 mm pixel size) from 10 overlapping images of a seafloor part with nodule coverage. a2) Hill-shade of the image-derived Digital Elevation Model (DEM; 5 mm pixel size). a3) Image-derived seafloor roughness (5 mm pixel size) from the same seafloor part. b) Orthophoto-mosaic (2 mm pixel size) of 10 overlapping images from a seafloor part without nodule coverage. b2) Hill-shade of the image-derived DEM

(5 mm pixel size). b3) Image-derived seafloor roughness from the same seafloor part. The Vector Ruggedness Measure (VRM; Sappington et al., 2007) was used as a roughness index. The white scale bar represents 0.5 m on the seafloor. c) Boxplots of the two image-derived VRM surfaces (a3 and b3) without outliers. The Wilcoxon rank sum test showed a significant difference between the medians. Figure source: modified from Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

An additional step in backscatter analysis is the identification of the MBES incident angles that are most sensitive and contribute the most to backscatter reflectance and nodule detection. Generally, backscatter mosaic results from combining several angles from different swaths after filtering (angle varying gain, correction, nadir filtering, anti-aliasing, despeckling) and blending with inverse distance weighting, i.e. mosaicking, aiming for the best mosaic representation (Fonseca and Calder, 2005). However, the backscatter angular response for a specific frequency and incident angle is an inherent property of the seafloor, resulting in a different backscatter reflectance (Jackson and Briggs, 1992; Fonseca and Mayer, 2007). Despite this, nodule-related studies have used only the backscatter mosaic as a spatial predictor in machine learning so far (e.g., Kuhn and Rühlemann, 2021).

Research article II was the first to use MBES backscatter mosaic from 260 kHz, and research article III was the first to use MBES backscatter mosaic from 400 kHz and the products of a MBES backscatter angular range analysis (ARA) as predictors in machine learning for spatial mapping of nodules. The angular range analysis has a higher angular resolution (as only a few angles are combined) than the backscatter mosaic, but it has a decreased along-track resolution because it stacks 30 (or more) consecutive MBES pings for the half swath width on port and starboard side (Fonseca and Mayer, 2007; Fonseca et al., 2009). Despite the lower along-track resolution, ping stacking at different beam angles can identify the beam angles that contribute the most to the spatial predictions and reveal subtle differences between seafloor patches with different sediment properties. Beyond the extraction of the average angular response at certain incident angles, ARA attempts to characterize sediment properties by comparing the average angular response for each group of incident angles with mathematical models that have linked backscatter and sediment properties (Jackson et al., 1986; Fonseca and Mayer, 2007). Based on this model inversion, the backscatter impedance (hereafter impedance) and backscatter volume (hereafter volume) can be calculated. Impedance is the product of sediment bulk density and the sound velocity ratio between the sediment compressional wave speed and water sound speed, showing the acoustic contrast between the water and the sediment surface (Fonseca and Mayer, 2007). Sediments with lower water content and typically larger grain size provide higher impedance and a stronger backscatter signal (Jackson and Briggs, 1992). Volume scattering provides information on the sound attenuation within the sediment concerning sediment physical properties (e.g., density) and inhomogeneities within the sediment column (e.g., buried nodules or fragments of them, biogenic burrows, and gas bubbles). In fine-grain size sediments (e.g., deep-sea clay and silt) with low density and roughness, volume scattering could be the dominant contributor at lower frequencies, at intermediate oblique incident angles ( $> 15^\circ$  and  $< 60^\circ$ ) and when a bioturbated or harder/rougher layered substrate exists below the surficial thin, soft layer, e.g., through thickness variations in the semi-liquid bottom sediment layer (Jackson et al., 1986; Lurton and Lamarche, 2015).

The angular range analysis showed that the backscatter differences between seafloor parts with varying coverage on nodules were distinct for the entire range of incident angles but most prominent in the **mean-far angle range from 25° to 55°** (ARA\_MF), which scored first in two models (RFM and NNM) and second in the other three models (GLM, GAM and SVM). The mean-near range (< 25°; ARA\_MN) and mean-outer range (> 55°; ARA\_MO) ranked lower in all models (Figure 24).

Two practical outcomes from this finding are that 1) the angles ranging from 25 to 55° should also be included as an available spatial predictor and that 2) future AUV-based backscatter-dedicated surveys should prioritize the backscatter acquisition within the incident angles of 25–55°, a finding consistent with the existing guidelines for backscatter-optimized surveys (Lurton and Lamarche, 2015). It should be highlighted that deep-sea AUV surveys are typically optimized for acquiring bathymetric data. The overall increased operational AUV cost, the absence of available high-resolution bathymetry (with multiple users from different scientific fields), the lack of a well-established need for high angular resolution backscatter data, and the additional processing effort, typically discourage dedicated backscatter surveys. A backscatter-optimized survey would increase survey time, as the seafloor should be ensonified on specific angles, ideally obtained from the same direction (Lurton and Lamarche, 2015; Lurton et al., 2018). However, it would offer a backscatter surface of higher quality than a typical mosaic from ‘bathymetric surveys’, allowing for better detailed seafloor modelling and resource estimations.

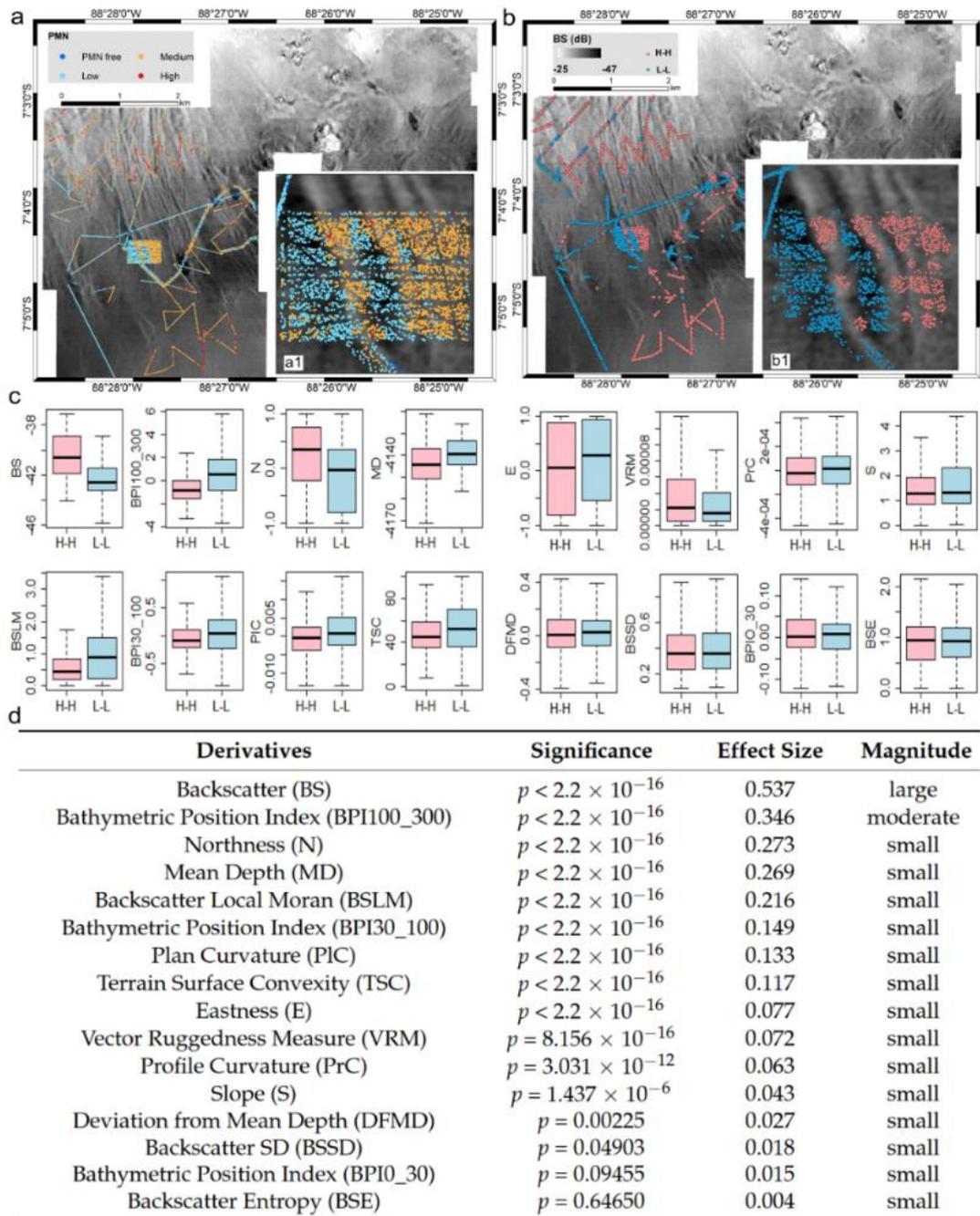
Regarding the ARA derivatives, impedance (primary) and volume (secondary) contributed to the model performance in all models except GLM (Figure 24). This finding agrees with other machine learning spatial studies, where ARA and its derivatives increased the machine learning prediction performance (Hasan et al., 2014; Alevizos and Greinert, 2018; Menandro et al., 2022; Misiuk and Brown, 2022; Porskamp et al., 2022). The lower contribution of volume backscattering is attributed to the limited penetration depth of the sound waves at 400 kHz, which is typically < 0.5 m for silt/clay sediments ( $\leq 62.5 \mu\text{m}$ ) and << 0.25 m for larger grain sizes (Huff, 2008; Gaida et al., 2018). The main component of volume scattering is expected to be the buried nodules located at a maximum depth of 5–10 cm.

In future studies, it would be interesting to use backscatter mosaics and ARA from multifrequency MBES because the use of higher and lower frequencies (e.g., 100–600 kHz) and also varying pulse lengths (e.g.,  $\mu\text{s}$  to ms range) could improve the machine learning predictive performance and interpretation (Menandro et al., 2022; Misiuk and Brown, 2022).

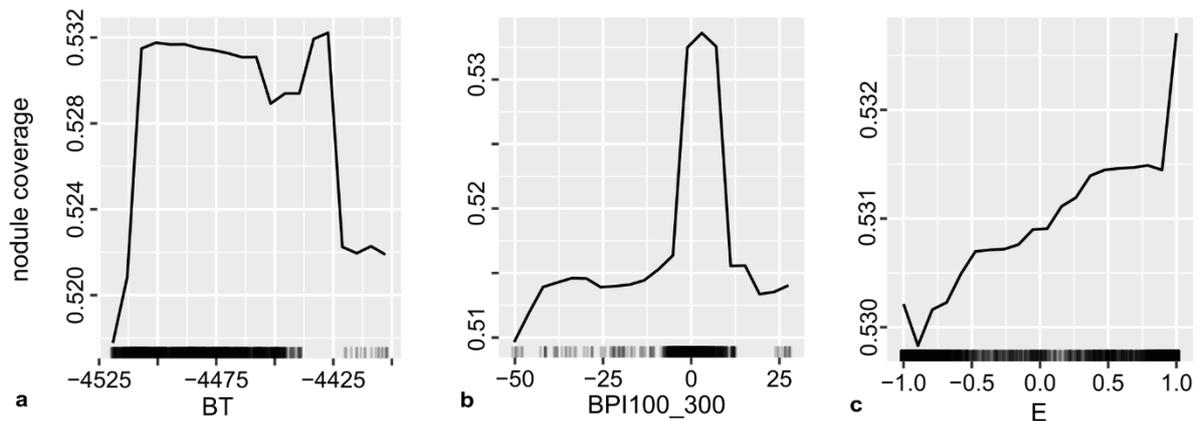
**Research Objective V:** *Provide insights into the geomorphological reasons behind the nodules' patchiness, by combining hydroacoustic and optic data, domain knowledge, and interpretable machine learning.*

The combined use and visual plotting of seafloor images, orthophoto-mosaics, backscatter data and ML-derived predicted surfaces showed that the **nodule coverage varied significantly even on a meter scale distance in all research areas** (Figure 7, Figure 21 and Figure 30). The nodules are clustered into distinctive higher or lower coverage patches, following the small-scale seafloor variations. These clusters have a statistically significant spatial autocorrelation, as proved by the Anselin Local Moran's Index, with very few (isolated) occurrences of spatial outliers, i.e. low values surrounded by high values or/and vice versa (Figure 30). The statistical analysis showed significant differences in the range of spatial predictors between the seafloor parts with higher and lower nodule coverage (Figure 30).

**Within the GSR B4 area of the CCZ** (Figure 21), higher nodule coverage and larger nodules occurred on relatively flat seafloor with gentle ( $< 3^\circ$ ), eastward-facing slopes. Both variable importance measures and partial dependency plots showed increased importance of eastward-facing slopes, which points toward a predominance of westward-moving bottom currents during the time of nodule formation (Figure 31). Long-term observations in the B4 domain of the GSR contract area show that the mean bottom current direction is (still) northwestward (Hayes, 1979; Juan et al., 2018; GSR, 2018). The predominant mean bottom flow towards NW is steady with a slow velocity (0–3 cm/s) only periodically interrupted by tidal currents towards NW and SE and sporadically short-time benthic storms (6–15 cm s<sup>-1</sup>), which are heading SE (Hayes, 1979; Juan et al., 2018; GSR, 2018). The steady and slow NW bottom currents contribute to favourable conditions for nodule growth (Skornyakova and Murdmaa, 1992; von Stackelberg, 1997; Mewes et al., 2014; Kuhn et al., 2017; Juan et al., 2018; Volz et al., 2018) by: (a) erosion and removal of fine sediments, leaving behind coarser particles that serve as nodule nuclei; and (b) continuous contact between nodule surface and well-oxygenated bottom water that allows the precipitation of hydroxides in nodules. Geochemical studies within the eastern CCZ have shown higher nodule coverage of larger sizes to sites with less clay-size particles and deeper oxygen penetration depths, particularly for sampling sites within the GSR contract area (Mewes et al., 2014; GSR, 2018; Volz et al., 2018). Older studies in the CCZ showed higher nodule coverage on flat seafloor parts with weak ( $\sim 3$  cm s<sup>-1</sup>) but steady near-bottom currents (Skornyakova and Murdmaa, 1992; Yamazaki and Sharma, 2000; Kuhn et al., 2017; Juan et al., 2018; Yoo et al., 2018; Alevizos et al., 2022). The importance of slope and slope orientation has also been shown in other machine-learning spatial studies within the CCZ (Kuhn and Rühlemann, 2021). The PDP of bathymetry showed a bathymetric range of  $\sim 100$  m, favouring the presence of nodule (Figure 31). Broader-scale studies in the CCZ and Peru Basin (Weber et al., 2000; von Stackelberg, 2000; Sharma, 2017) showed the maximum coverage and abundance occurring within a small bathymetric range (50–100 m) in water depths just below the local CCD, resulting in similar response curves as in the here presented studies.

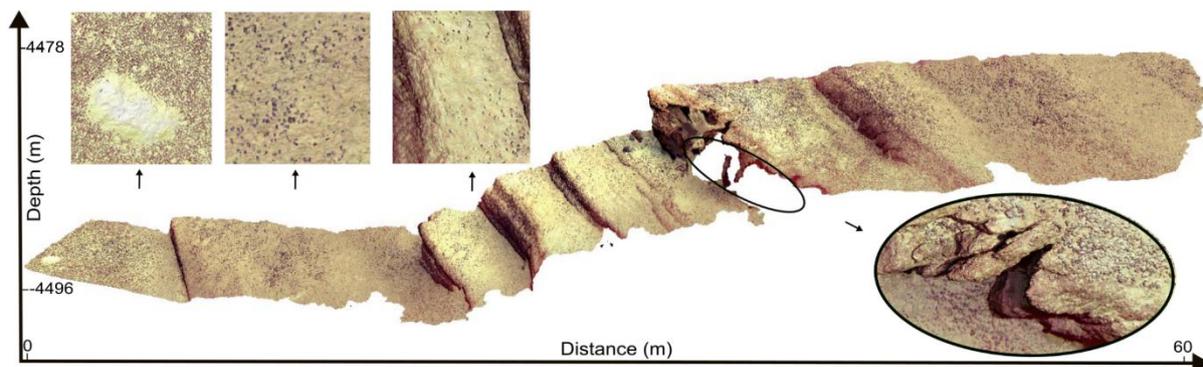


**Figure 30.** a-a1) Image-derived spatial distribution of nodules in DEA study area (Research Article II). b-b1) Spatial clustering of nodules into groups of high-high (H-H) values and low-low values (L-L) using the Anselin Local Moran's Index. The northwest and southeast parts have mainly or exclusively H-H values. Conversely, the eastern part has L-L values. b1) The clustering zonation follows the seafloor microrelief (abyssal furrows). c) Boxplots between spatial clusters and MBES-derived spatial predictors, showing the differences between H-H and L-L clusters. For example, higher BS values are linked with H-H clusters (see Research Objective IV). The bathymetry, broad-scale BPI (100–300 m), and slope orientation (Northness and Eastness) also have significant differences between the two groups (see following paragraphs). d) Wilcoxon–Mann–Whitney test (Wilcoxon, 1945; Mann and Whitney, 1947) results between statistically significant spatial clusters and MBES-derived spatial predictors. The statistical significance (p-value) is given together with the substantive significance (effect size) and magnitude classification for interpreting the results when the p-values are the same or similar. Figure source: modified from Gazis and Greinert, 2021 <https://doi.org/10.3390/min1111172>



**Figure 31.** PDPs between nodules and geomorphological derivatives for the GSR B4 area (Research Article III). (a) bathymetry (BT), (b) broad-scale bathymetric position index (BPI100\_300) and (c) Eastness (E). Black marks on the x-axis show the data distribution. Most nodules are on a gently sloping and flat seafloor (low BPI values around 0) facing eastwards ( $E=1$ ) within a limited bathymetric (ca. 100m). Depths outside that range were related to seafloor depressions or topographic highs with few to no nodules. Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

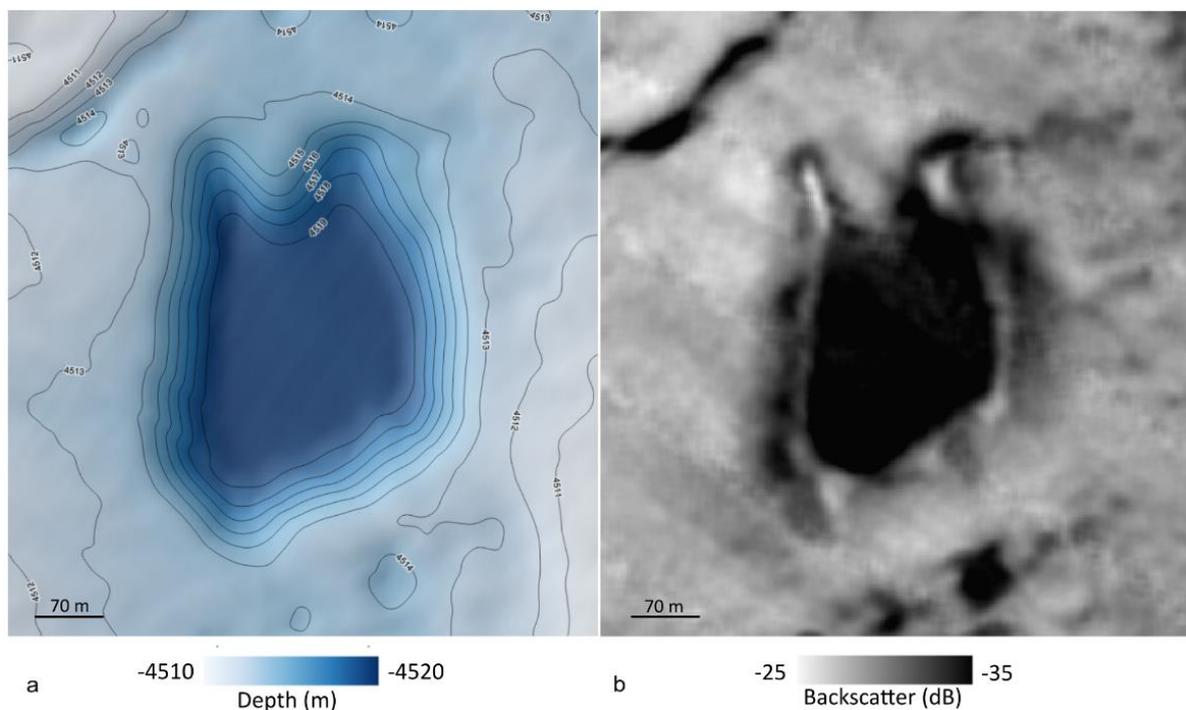
The highest nodule coverage of up to 50% was observed at geomorphologically prominent locations probably because of abundant nuclei material of fragments that might have been gravitationally transported downslope when steep slopes exist (Yamazaki and Sharma, 2000; Sharma, 2017; Li et al., 2021). However, the nodules in such high coverage areas were smaller than in the central and flatter part of the study area in the GSR are, a fact that can be attributed to several factors: a limited available horizontal space relative to a high number of nuclei, younger age, primarily slower hydrogenetic growth, and limited supply of dissolved metal through diagenetic precipitation due to a thinner or absent semi-liquid bottom sediment layer i.e., geochemically active layer (Cochonat et al., 1992; Lipton et al., 2016; Figure 32).



**Figure 32.** Colour-normalized ortho-photomosaic within a trough at the study area of Research Article III. Exposed carbonate-cemented sediment layers appear as successive steps along the eastern slope of the trough. Densely packed small nodules are observed at the deepest and flattest parts. On top of the compacted and layered indurated carbonate sediments, the loose sediment was winnowed away, and fewer to no small nodules lay. The grid cell size is 2 mm. Figure source: Gazis et al., 2024 <https://doi.org/10.1007/s11053-024-10393-7>

Lower nodule coverage was observed mainly at four sub-environments: (a) in seafloor parts with typically higher bottom current velocities such as gullies, furrows or the foot of slopes; (b) on topographic highs, with steep slopes such as horsts and volcanic knolls; (c) in deeper parts of basins around and inside sub-circular depressions and (d) in areas with basement exposure, rocky outcrops, and carbonate cemented sediment layers (Figure 32). In detail, the

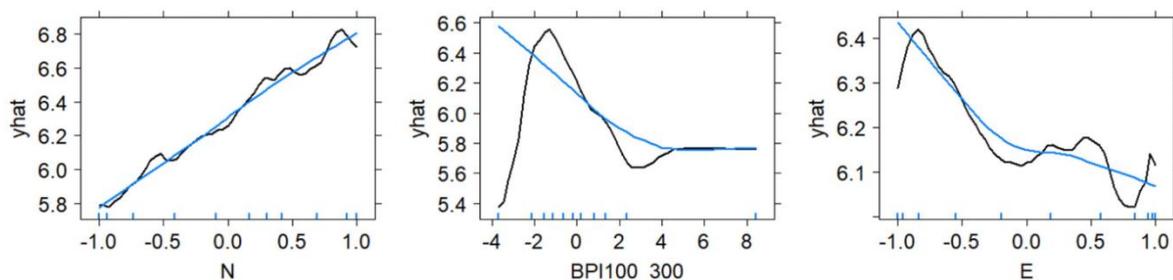
PDP of BPI100-300 (Figure 31) shows a small response to nodule coverage (i.e., few to no nodules) in local depressions with low backscatter reflectivity known as 'black patches' (Figure 33). Studies have shown that local depressions exhibit geochemical conditions different from their surroundings, having a higher amount of fine-grain sediments (clay) and organic carbon content. Fine sediments are trapped in the local depressions during sediment transportation (Wiedicke and Weber, 1996) and cannot act as nuclei for nodule formation due to their size and a typically small content of mobilized manganese, which hinders nodule formation and growth (Mewes et al., 2014). The increased organic carbon content shifts the Mn-redox closer to the seafloor, where the limited diagenetically mobilized manganese is released into the bottom water without supporting nodule formation and growth (Wiedicke and Weber, 1996; Marchig et al., 2001; Benites et al., 2018; Paul et al., 2019).



**Figure 33.** Sub-circular nodule-free depression in the study area of Research Article III. a) The depth difference from the surrounding seafloor is 10 m. Its base is remarkably flat and even. b) The backscatter difference from the surrounding seafloor is 10 dB, the minimum recorded in the study area. The even seafloor, without nodules and thicker layers of fine-grain sediment, reflects less acoustic energy to the MBES. Figure source: Gazis, 2025 (PhD dissertation).

Similar to local depressions, topographic highs also have limited nodule coverage related to increased slopes, reduced sediment thickness (even outcropping basement). Local variations in sediment geochemical characteristics such as particulate organic carbon (POC), total organic content (TOC), and Mn-redox depth occur as a result of the fine-sediment winnowing and downslope transportation from the locally intensified or deflected bottom currents as argued by Skornyakova and Murdmaa (1992), von Stackelberg (1997) and Volz et al. (2018). These studies mention that even small height differences of < 10 to 25 m are associated with such changes, which agrees with our findings.

**Within the DISCOL Experimental Area (DEA) in the Peru Basin,** the nodule coverage also follows the meter-scale seafloor variations created by the well-pronounced abyssal sediment furrows and the prevailing current regime (Figure 30). The observed abyssal furrows strike perpendicularly to the regional contours and exhibit a U-shape form. They are oriented parallel to the predominant NW bottom current flow (Klein, 1993; Klein, 1996). These long-term studies showed that the deep currents are not stable but have periods of quasi-unidirectional strong currents (>5 cm/s but sometimes up to 17 cm/s) and periods with slower omnidirectional currents (1–3 cm/s). The recorded bottom current velocities are within the velocity range that could preserve abyssal furrows (Flood, 1983). The abyssal furrows act as bottom current channels, occasionally eroded during short periods of higher current velocities, or they could be relics from an earlier period when the current regime was stronger (Flood and Hollister, 1980; Flood, 1983; Lonsdale and Spiess, 1997). Their preservation is supported by the low regional sedimentation rates of 0.4–2.0 cm/ka (Haeckel et al., 2001). Like in the CCZ, PDPs of Northness, BPI100\_300 and Eastness show that slope orientation and gently sloping seafloor are important predictors (Figure 34).

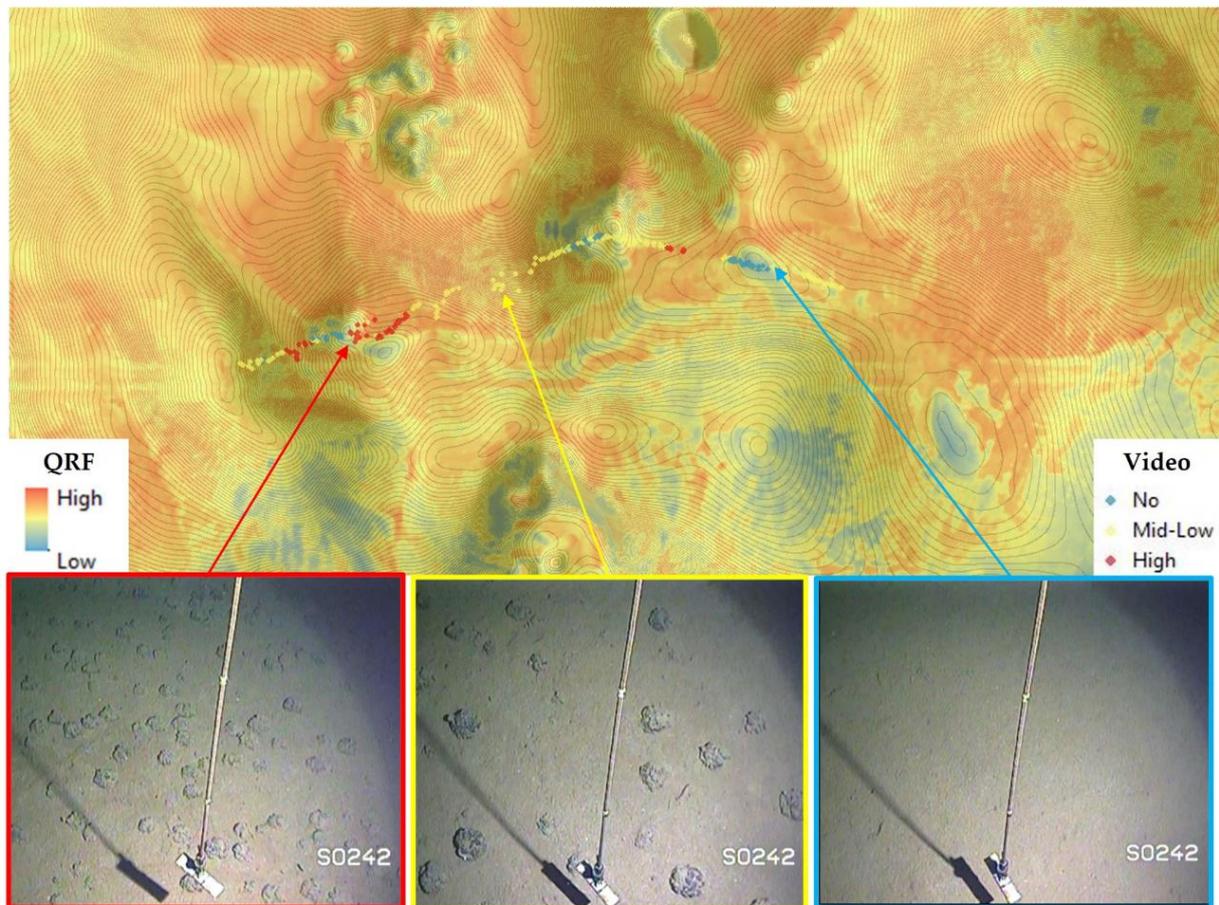


**Figure 34.** PDPs between nodules and geomorphological derivatives for the DEA area (Research Article II). The black line shows the random forests PDP, and the blue line shows the LOESS curve, which shows the overall agreement and local differences between a random forests algorithm and a simpler non-parametric regression model. Blue ticks on the x-axis represent the data deciles. For Northness (N) and Eastness (E), quasi-linear and monotonic relationships are observed. Most nodules are on a gently sloping and flat seafloor (low BPI values around 0) facing northwest (N=1 and E=-1) parallel to the dominant bottom current direction. Figure source: modified from Gazis and Greinert, 2021 <https://doi.org/10.3390/min11111172>

Nodule-free black patches are also observed in the DEA, which are easily recognizable due to their low backscatter reflectivity, similar to findings in the CCZ (Figure 30). They have an elliptical and elongated shape, with their long axes-oriented downslope. This downslope direction has been observed in other low-reflectivity nodule-free patches in the Peru Basin and has been associated with downslope sediment transport (Wiedicke and Weber, 1996). The depth of the black patches varies between 2–5 m with a remarkably flat central seafloor, again similar to the CCZ. A gravity core from inside such a DEA black patch revealed that its geochemical conditions differ significantly from the surroundings, with a strongly enriched organic carbon content (0.5 – 0.8 wt %; Paul et al., 2019) that could hinder nodule formation and growth in a similar way to the mechanism described for the CCZ above.

NE of the DEA, the seafloor is heterogeneous, with elongated and conical knolls and hills separated by local depressions. Visual inspection showed that slopes have a thin sediment cover, while local depressions act as sediment depocenters with increased sediment accumulation and talus debris (Greinert et al., 2015). The debris of basaltic fragments could

act as nuclei supply to form new nodules (von Stackelberg, 1997; Weber et al., 2000). Black patches are also present, and the nodule coverage varies, showing the need for high-resolution quantitative studies that use geomorphological and environmental predictors (Figure 35).



**Figure 35.** Quantile regression forests (QRF) predictions and video observations (SO242-1#135\_OFOS-6) NE of the DEA area. Predictions and observations have the same spatial trend, showing alternations between areas with higher, lower, and zero nodule coverage. The superimposed bathymetric contours have a 1 m interval. Figure source: Gazis and Greinert, 2021 <https://doi.org/10.3390/min11111172>

## Future Work

The model accuracy achieved in the above-mentioned studies ( $R^2 < 0.8$ ) shows that there is still space for improving the spatial models. The inclusion of more spatial predictors, the optimization of AUV image surveys and the use of different models are three main ways of achieving a low prediction error.

### Which other spatial predictors should be included?

Using AUV-based MBES and SSS derivatives as spatial predictors provides spatially large high-resolution maps (see introduction) and enables to derive the causal relationships (direct, synergistically or proxy) for modelling the spatial distribution of nodules. Additional predictors, unavailable in this thesis, could improve prediction performance and model interpretability in future studies towards holistic spatial modelling. These could be:

- a) Geophysical data obtained by Sub-Bottom Profilers (SBP), such as the thickness and composition of the semi-liquid bottom sediment layer. As discussed above, this is related to nodule coverage, particularly in areas with a high diagenetic contribution to the nodule growth (Cochonat et al., 1992; Yoo et al., 2018; Dreisetl, 2016; Lipton et al., 2016; Alevizos et al., 2022). An effort to collect SBP data was made for research article III; however, technical issues hindered the acquisition of meaningful data.
- b) Backscatter mosaics and ARA from a broad band multifrequency MBES with higher and lower frequencies (e.g., 100–600 kHz) and varying pulse lengths (e.g.,  $\mu\text{s}$  to ms range) has shown to improve machine learning predictive performance and interpretation (Menandro et al., 2022; Misiuk and Brown, 2022).
- c) More and well spatially distributed geochemical information from the sediment. A dense sampling network of sediment and pore water geochemical parameters (e.g., porosity, grain size, TOC,  $\text{O}_2$ ,  $\text{NO}_3^-$ ,  $\text{Fe}^{2+}$ ,  $\text{Mn}^{2+}$ ) would likely contribute the most to spatial modelling at this scale of analysis. Although the geochemical samples are local observations in space, they can be interpolated to continuous surfaces and used as input layers in spatial models despite the additional model uncertainty that interpolation brings. This approach is a common practice in land-based mineral potential assessments (e.g. Rodriguez-Galiano et al., 2015; Liu et al., 2022 and references therein) and marine studies in shallow areas, where hundreds of samples exist and can be interpolated using machine learning algorithms (e.g. Diesing et al., 2017; Diesing et al., 2020a, b; Spiegel et al., 2024). However, in deep-sea environments, such an application is limited to broad-scale studies (e.g. Maciąg and Harff 2020) due to data scarcity / sampling sites in relatively small areas. A typical ground truth sample in the deepsea (e.g., box corer, gravity corer or multi-corer) demands four hours on average to be recovered. The time doubles in case of deployment failure and acquisition of a replicate sample. ROVs are also used to acquire ground truth samples (e.g. short push-cores) for geochemical analysis. However, they exhibit long deployment times ( $\approx 12$  h), limited spatial coverage ( $\leq 500$  m from the research vessel) and high operation costs. Both approaches are susceptible to weather conditions, sea state and mechanical failures. Unless sampled with a multi-corer, the surficial sediments (0-10 cm) that are needed the most are disturbed and mixed by other sampling gear. Today, samples for geochemical and

associated biological studies with different sampling gears are typically recovered in small spatial areas not allowing comparisons between different geomorphological units. Studies have shown increased fine-scale spatial variability of geochemical properties even in the deep-sea, highlighting the need for intensifying the efforts on sediment sampling in the deep-sea (Mewes et al., 2014; Volz et al., 2018; Paul et al., 2019).

### **What else can be done to improve the modelling performance?**

The model performance could be further improved by optimizing the feature space coverage during image data acquisition (see also conclusions of Research Objective II). While scientific fields such as soil predictive modelling have realized the need for good feature space coverage in spatial models (e.g., Kidd et al., 2015; Brus, 2015; Wadoux et al., 2019; Bouasria et al., 2023), the deep-sea habitat mapping community lacks behind in this respect, potentially because the acquisition of samples demands more time and cost than during terrestrial surveys. This paradox is because satellite and airborne DEMs are typically available before field sampling in terrestrial studies. In deep-sea studies, the bathymetric and backscatter data are obtained shortly before, during (e.g. AUV) or even after the field sampling. Future expeditions should prioritize the MBES acquisition and near real-time processing of just acquired data. After that spatial predictors can be calculated, and informative sampling locations can be selected based on criteria/algorithms that maximize the geographical and feature space coverage, such as the conditional Latin Hypercube Sampling method (Minasny and McBratney, 2006). Such guidelines already exist in the official manuals for marine sampling to monitor Australian waters (Foster et al., 2020) and thus should be also considered in future activities elsewhere.

### **Notes on reproducibility**

Future work implies building and improving existing knowledge. In this direction, the reproducibility of the current work needs to be guaranteed. Regarding research article I, the machine learning modelling has been conducted using the [Marine Geospatial Ecology Tools](#) (MGET). MGET (Roberts et al., 2010) uses the random forests R package for classification and regression (Liaw and Wiener, 2002) inside the graphic user interface of [ArcMap™](#) 10.1, which is familiar and accessible to the majority of marine geoscientists. Research articles II and III used [R](#) ( R Core Team, 2022) and [R studio](#) (R Studio Team, 2022) to develop the machine learning modelling workflow combining different statistical libraries. The R script for research article II is uploaded as an R markdown .pdf document on the [GEOMAR cloud](#). The R script for research article III will also be uploaded as an R markdown .pdf document using the same link once finally generated.

---

MGET: <https://mgel.env.duke.edu/mget/download/>

ArcMap: <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>

R: <https://www.r-project.org/>

R studio: <https://posit.co/download/rstudio-desktop/>

GEOMAR cloud: <https://cloud.geomar.de/s/G2Jsi82ZTFLoH5A>

## **Section II**

### **Benthic plume monitoring**

*"...the disciplines are the place where we begin, but not where we end."*

*Introduction to Interdisciplinary Studies (Allen F. Repko, 2014)*

## Foreword

The protection of the deep-sea ecosystem and the related holistic scientific and technological approaches in deep-sea research have increased the questions' complexity, making it impossible to address them from only based on a single scientific field. Modern marine (geo)scientists thus need and should follow an interdisciplinary research approach to combine knowledge from different fields, effectively working within a team of experts. The *US National Academies of Sciences, Engineering and Medicine* defines that interdisciplinary research should have the following characteristics:

- It integrates information, data, techniques, tools, perspectives, concepts or theories from two or more disciplines or bodies of specialized knowledge.
- It can be done by teams or by individuals.
- It advances fundamental understanding or solves problems whose solutions are beyond the scope of a single discipline or area of research practice (NAS, 2005).

## Mining trial monitoring

All the above-mentioned characteristics for interdisciplinarity have been fulfilled during the scientific monitoring of the first modern deep-sea industrial mining test of a pre-prototype seafloor nodule collector vehicle (hereafter nodule collector), representing a critical step in the roll-out of this new and highly controversial industry. The author, with scientists from 29 European research institutes and universities of 8 European countries (MiningImpact, 2022; JPIO, 2022), joined forces to independently monitor the environmental impact of the industrial-scale mining of polymetallic nodules in the GSR contract area (CCZ), collecting biological, microbiological, biogeochemical, ecotoxicological, geological, and oceanographic data and testing needed monitoring approaches (Vink et al., 2022).

A cornerstone of this effort is monitoring **the sediment plume** (hereafter plume) generated behind the nodule collector from the discharged sediment-laden water and the collector's movement on the seafloor. The plume's SPM concentration, its spatial extent, and the amount and grain size of redeposited sediment are critical environmental properties. They define the impacted benthic ecosystem beyond the actual mining site and determine the time the ecosystem is under anthropogenic pressure. Such data are what marine biologists need to quantify the impacts on benthic life during and after the mining and at different levels of exposure. The presented data are essential input values to numerical models that extrapolate the plume evolution over larger areas and time spans beyond a limited trial. To collect these data, the author, together with his colleagues, deployed **the most spatially dense sensor network for deep-sea mining research** consisting of fifty acoustic [e.g., MBES, acoustic doppler profiler (ADCP), hydrophones] and optic backscatter sensors (e.g., JFE, WetLabs) and cameras (e.g., PartCam) mounted to twenty seafloor platforms, including tripods, benthic landers and moorings (Vink et al., 2022). Two remotely operated vehicles (ROVs) visually observed the plume and deployed/recovered the multi-sensor platforms. At the same time, an AUV acquired water-column and seafloor data MBES data, SSS backscatter data, seafloor images, and SPM concentration at different altitudes (5, 10, 30 and 50 m) and distances (500 m to 5 km) from the mining site. Sensors were intercalibrated before and after the cruise with

sediment from the same area to allow for a quantitative comparison of the sensor recordings in SPM concentration (Haalboom et al., 2022). All sensors measured ambient SPM concentration and background acoustic backscatter intensity before  $t$ , during and after the trial to allow a quantitative comparison.

## Research objective of Section II

Section II has only one research objective, which is:

- to understand the 4D spatiotemporal evolution of the mining-induced sediment plume and the technology needed for its monitoring.

Working on addressing this objective resulted in seven key findings that deal with the spreading of the mobilized sediment as turbidity current (1), gravity flow (2) and suspended sediment plume (3 and 4), the particle flocculation and fast sediment redeposition (5 and 6) and the direct seafloor disturbance (7).

### Key findings

1. The **plume spreads as a high-density gravity current** behind the nodule collector spread perpendicular to mining lanes independent of the bottom currents, driven by its turbulent flow dynamics. This finding agrees with those of Muñoz-Royo et al. (2022). Optical backscatter Sensor (OBS) data showed a maximum SPM concentration of  $264 \text{ mg L}^{-1}$  at 1 m altitude, 50 m from the mining lines but along the propagation axis of the turbidity current. This is 13,200 times higher than the background SPM concentration of  $0.02 \text{ mg L}^{-1}$  at 1 m altitude (Figure 36). During our measurements, the SPM concentration at the measuring site returned to background concentrations 14h after the mining ended.

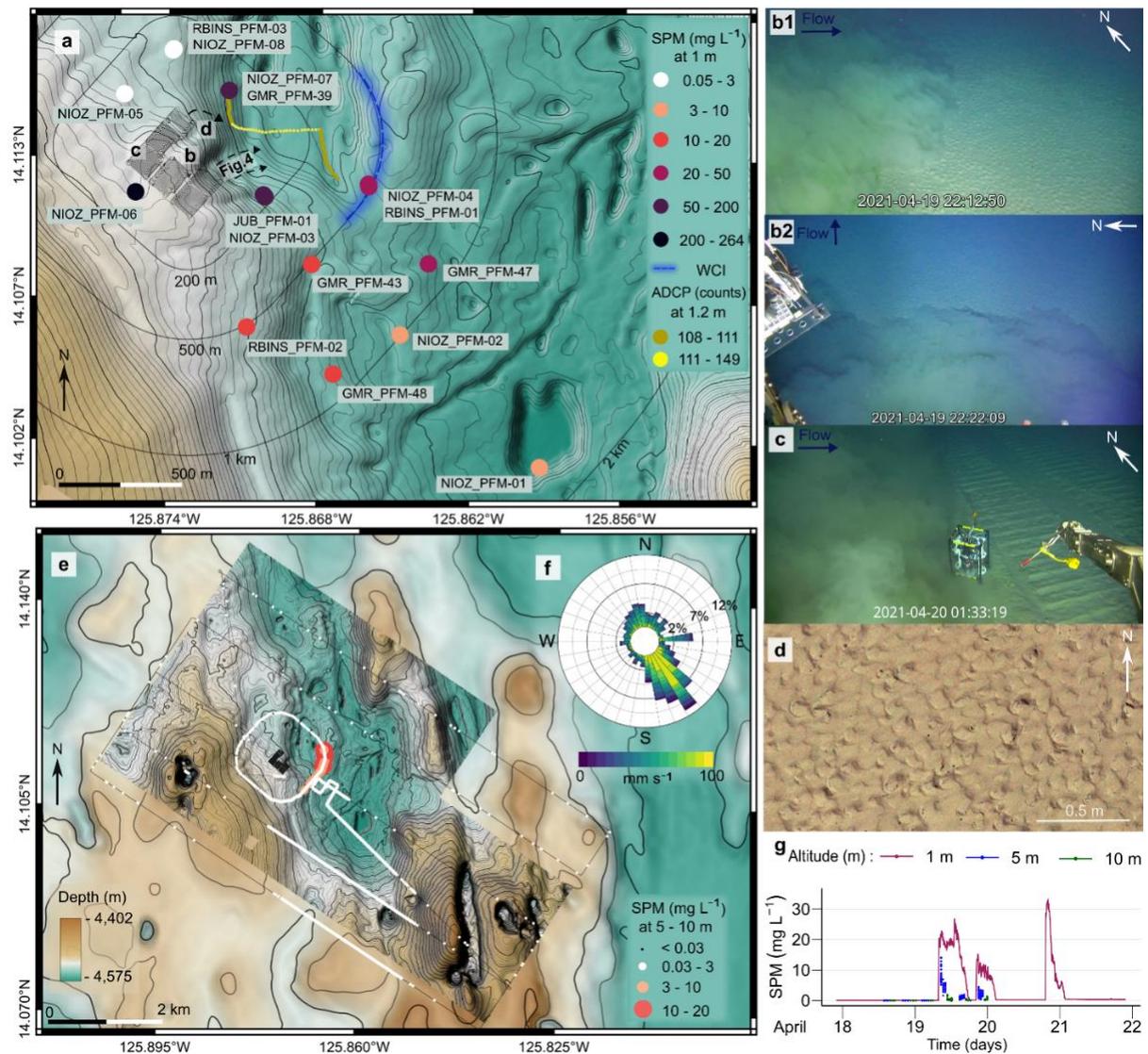
2. a) The gravity current is channelled through steeper seafloor sections ( $> 3^\circ$  slope) along its propagation path and moves independently of the bottom water current direction. The **gravitational forces at steeper seafloor sections increased the total propagation distance** five times (up to 500m) compared to flat seafloor stretches (up to 100 m). A topographic sill blocked the propagation at a certain distance, and the plume started spreading laterally (Figure 36). As a direct consequence, more resuspended sediment was actively transferred and redeposited outside of the mining site. **This is a new finding as no other study, including numerical modelling, had considered small-scale seafloor geomorphology as an essential factor for sediment plume propagation and sediment redeposition.** AUV-scale mapping of mining sites and a large enough surrounding should be an integral part of monitoring, numerical modelling, and mining site prioritization.

2. b) The **gravity current had a distinct backscatter signature** in ADCPs and OBS sensor data compared to the passive transfer (advection) of the sediment plume transported by bottom currents (Figure 37).

2. c) The **gravity current created sediment ripples along its downslope direction** at the steeper seafloor sections (Figure 36). The shape of ripple crests is sinuous to catenary, a typical bedform generated by a strong flow over cohesive sediment (Bass et al., 2016). The presence of ripples could imply erosion of surficial sediments (e.g. Koller et al. 2019), but erosion would

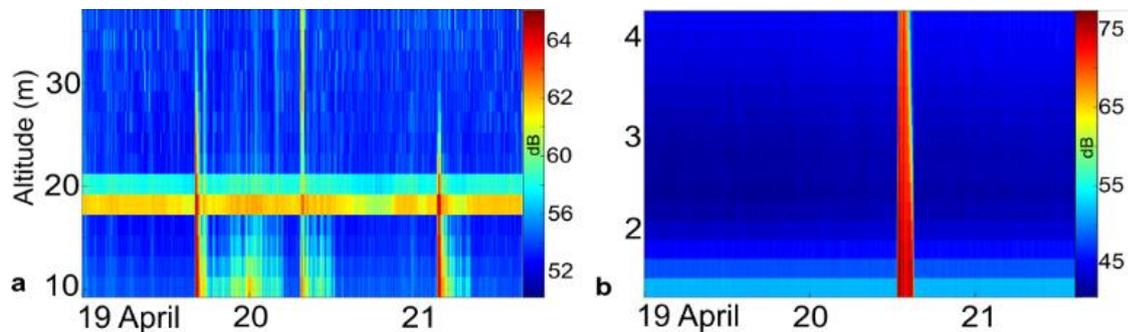
demand current velocities  $> 30 \text{ cm s}^{-1}$  (Thomsen and Gust, 2000; McCave and Hall, 2006), which were not measured in the field. The massive fallout of aggregated sediments within the gravity flow (see point 5) and the unidirectional currents in steeper seafloor parts could result in mud ripples generated by large non-cohesive plume aggregates, which behave hydraulically, equivalent to fine sands (Thomsen and McCave, 2000; Schieber et al., 2007).

3. The final plume-stage of sediment spreading is the **passive transport of fine suspended particles by the ambient bottom currents**. The AUV recorded the plume at distances up to 4.5 km from the mining site at the very end of the monitoring area (Figure 36).



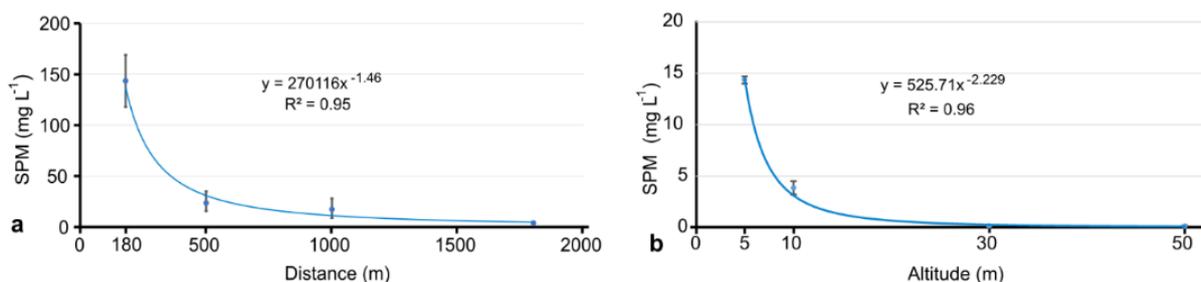
**Figure 36.** a) Maximum concentrations registered by OBS sensors at 1 m altitude on different platforms. The ADCP trajectory plot of NIOZ\_PFM\_07 showed an eastward displacement and increased backscatter intensity when the gravity current passed. The extent of MBES WCI data agrees with the OBS data from AUV (e). **b1-b2)** ROV video frame grabs inside the nodule impact site show the edge of the plume being distinct from the ambient seawater. Limited vertical mixing is observed as the ROV altitude is  $\approx 4.8 \text{ m}$ . **c)** ROV image showing the plume heading perpendicular to mining lanes. The sensor shown (Microprofiler) is  $\approx 80 \text{ cm}$  in height and  $60 \text{ cm}$  in width. **d)** Sediment ripples at the end of the third mining strip. **e)** The AUV recorded the plume's spatial extent during a ca. 54 h survey conducted parallel to the mining trial and continued after its termination. Scattered individual measurements above  $0.03 \text{ mg L}^{-1}$  do not correspond to the plume but to noise, e.g. from fish inside the sensor's view. The background map is the AUV MBES bathymetry (cell resolution  $2 \text{ m} \times 2 \text{ m}$ ) of the monitored area

superimposed on ship-based MBES bathymetry (Gazis, 2020). The contours have a 1 m and 25 m interval, respectively. **f)** The Rose diagram at NIOZ\_PFM-04 represents the current direction and speed during the nodule collector trial. The eastward spike is related to the gravity current. **g)** Concentration recorded at 1 m, 5 m, and 10 m altitude at NIOZ\_PFM-04. The measurements at 5 m and 10 m altitude had to last shorter to facilitate the far-field survey. Figure source: Gazis et al., 2025 <https://doi.org/10.1038/s41467-025-56311-0>



**Figure 37.** a) Backscatter time series recorded with an upward-looking ADCP at 300 kHz, 1800 km SE from the impact site (NIOZ\_PFM-01, see Figure 36). Although the background noise between 18-22 m above the seafloor (caused by a flotation body on the platform) dominates the signal, three distinct plume episodes remain visible. b) Time series of an upward-looking ADCP at 1.2 MHz (GMR\_PFM-39; Figure 36), showing the sharp increase in backscatter intensity when the gravity current passed. All backscatter values are range normalized and given in dB. Figure source: modified from Gazis et al., 2025 <https://doi.org/10.1038/s41467-025-56311-0>

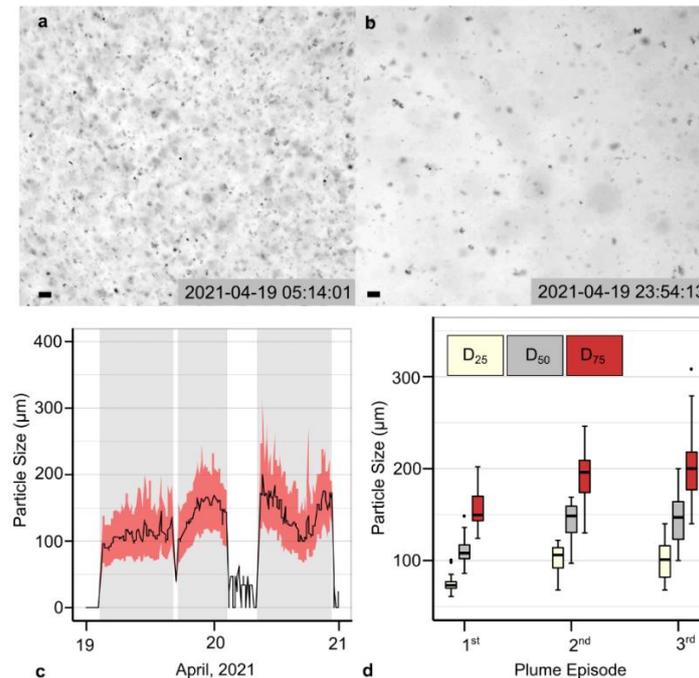
4. The **SPM concentration decreases fast with increasing distance and altitude**. A power-law decay of the concentration was detected at successive distances (180 to 1800 m, at 1 m altitude) and altitudes (5 to 500 m, at 500 m distance) from the mining site towards the ambient bottom current direction (Figure 38). Such empirical power-law shaped functions have been observed in diluted plumes such as in hydrothermal plumes (e.g., Bemis, 2002) and in plumes generated by towed fishing gear (e.g., Madron et al., 2005; O'Neill et al., 2013). This power-law relation shows the settling of the SPM (due to gravity), upward lifting and mixing (due to turbulence), and sediment transportation from bottom currents. The empirical power-law function cannot be applied at shorter distances (here < 200 m) and lower altitudes (here < 5 m). Here, the plume shows little mixing with a sharp front distinct from the ambient water, it is dominated by the near-field gravity current dynamics generated by the mining vehicle (Peacock and Ouillon, 2023) and seafloor geomorphology. The bulk of the plume stayed below 5 m, particularly within the first 500 m. The plume height increased with distance and time; SPM concentration similar to the background values was detected at the highest monitoring altitude of 50 m.



**Figure 38.** a) Maximum concentration recorded by OBS sensors at successive distances (from 180 m to 1800 m) and times (19-22.04.2021) in the SE direction from the impact site. The records from the following platforms were used: JUB\_PFM-01 and NIOZ\_PFM-03 at 180 m, NIOZ\_PFM-04, RBINS\_PFM-01, GMR\_PFM-43 and RBINS\_PFM-02 at 500 m, GMR\_PFM-47, NIOZ\_PFM-02 and GMR\_PFM-48 at 1 km, and NIOZ\_PFM-01 at

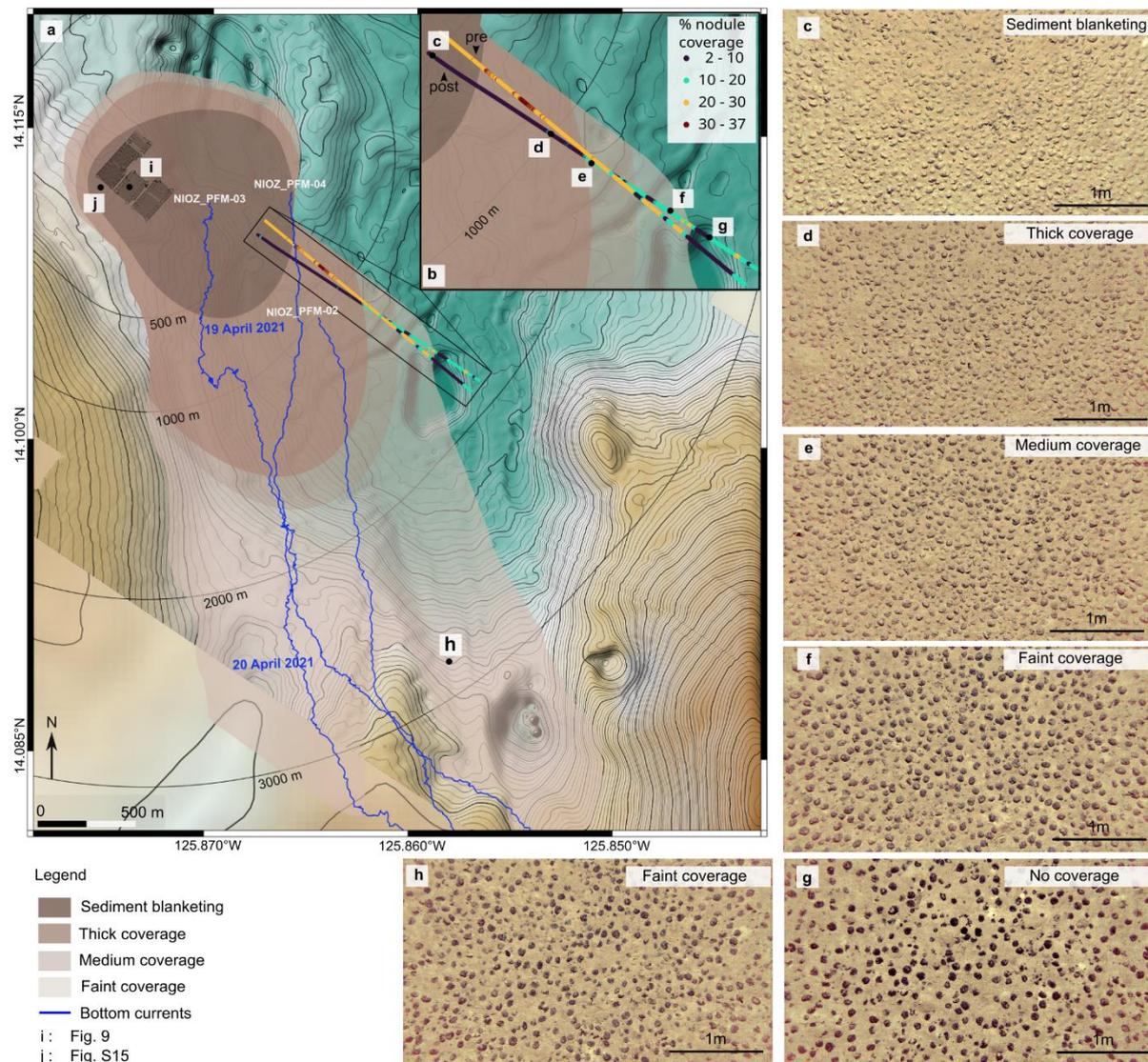
1.8 km (see Figure 36). When more than one OBS data set was available, the average maximum value was calculated and plotted. Vertical bars represent the highest and lowest maximum concentration values measured; this variability is due to the different locations across the longitudinal axis at the same distance from the impact site. b) The average maximum concentration recorded by the two OBS on the AUV at successive altitudes (from 5 m to 50 m) and times (19-20.04. 2021) while circling at a 500 m distance around the impact site. The bars represent the highest and lowest difference from the averaged maximum concentration between the two OBS sensors on the AUV. Figure source: modified from Gazis et al., 2025 <https://doi.org/10.1038/s41467-025-56311-0>

**5. Particle flocculation was recorded in situ**, confirming previous lab studies (Gillard et al., 2019) and mining studies at cobalt-rich crust sites (Spearman et al., 2020). The PartiCam system (Gillard et al., 2022) acquired water-column images, that were used to derive a image-based particle size distribution (Figure 39). While the median particle size ( $D_{50}$ ) of the top 5 cm of the seafloor sediment column is 0.012 mm (GSR, 2018), the plume  $D_{50}$  increased up to 0.147 mm. The  $D_{25}$ – $D_{75}$  of the plume ranged from 0.061 mm to 0.313 mm, with the largest particles to be detected when the plume was produced furthest away from the stationary camera location experiencing the longest aggregation time. The fact that the plume was recorded with ADCPs at 300 kHz with low sensitivity for the small primary particle size (median  $\approx$  0.02 mm; Haalboom et al., 2022) also indicates that at least parts of the mobilized fine-grained sediment must be present in an aggregated form that falls within the sensitivity window of the 300kHz frequency, confirming the larger grain sizes seen by the PartiCam system. This finding ends a long discussion in the scientific community on when sediment flocculation occurs within benthic plumes induced through nodule mining activities (e.g., Aleynik et al., 2017; Purkiani et al., 2021; Muñoz-Royo et al., 2022). It also underlines the need to include sediment flocculation in numerical plume dispersion and redeposition models.

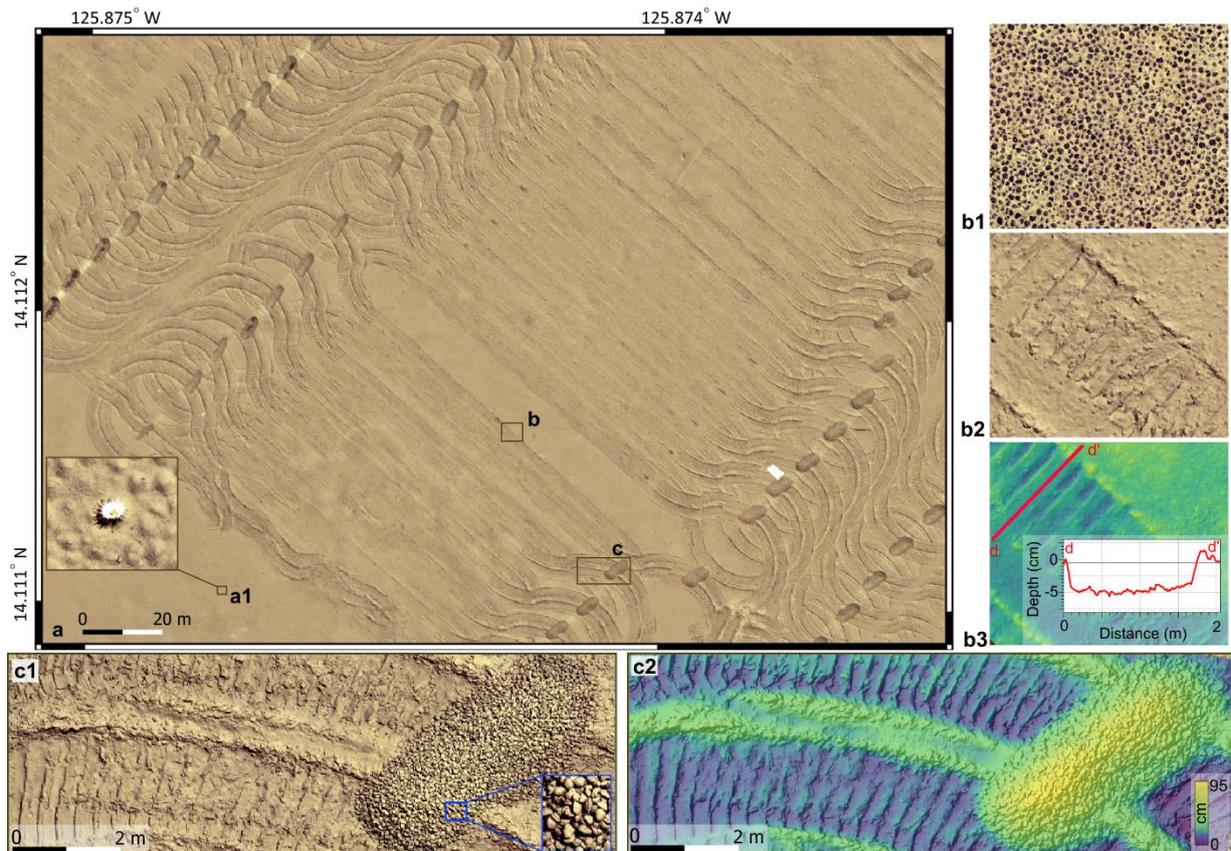


**Figure 39.** a and b) In-situ water-column photos obtained with the PartiCam system during the first day of the mining trial showed a decrease in particle number and an increase in particle size over time. The black scale bar corresponds to 0.27 mm. The minimum particle size that could be detected with that optical setup was 0.0153 mm with a pixel resolution of 0.0051 mm. c) Time series of particle size distribution during the three plume episodes (grey areas). The black line is the  $D_{50}$ , and the red area shows the upper  $D_{75}$  and lower  $D_{25}$  of the particle size distribution. d) Boxplots of the particle size for each distinct plume episode. Figure source: modified from Gazis et al., 2025 <https://doi.org/10.1038/s41467-025-56311-0>

6. **Sediment redeposition occurs fast** and within a short distance from the mining site; this is due to the friction of the developing gravity current and the seafloor and particle flocculation that increases the settling velocity (Figure 40). Seafloor images show the extent of the redeposited sediment, highlighting that most sediment was redeposited within the first 500 m; only a faint redeposition and nodule cover could be detected 3 km downstream (Figure 40).



7. Direct **seafloor disturbance** created by the movement of the nodule collector and suction of nodules **was mapped in high-resolution (cm to mm)** using AUV-mounted MBES, SSS and seafloor images (Figure 41). Image-derived digital elevation models (DEMs) before and after the mining trial were used to calculate the thickness of the redeposited sediment within the mining site, where a thick sediment blanketing occurred. The analysis showed that at least 3 cm of sediment was redeposited during the first 88-110 h after the mining trial ended. The depth of the average seafloor erosional inside distinct mining lanes was estimated at be around 5 cm (Figure 41).



**Figure 41.** a) Orthophoto-mosaic with 2 mm × 2 mm pixel resolution from part of the second mining strip showing the nodule collector imprints on the seafloor as mining lanes, the "light bulb" turns, and nodule piles. a1) A sea anemone (*Actiniaria*), capable of cleaning itself from sediment (Savoca et al., 2022), is shown as an example of the resolution achieved. b1) Orthophoto-mosaic of the seafloor before the mining trial. b2) The same part of the seafloor after the mining trial showed the caterpillar track imprints. b3) Image-derived DEM showing the depth of the caterpillar imprints on the seafloor. A depth profile across the caterpillar track (red line) shows the erosion depth caused by the nodule collector and the humps on each side of the track due to sideways sediment extrusion under the caterpillar tracks. The zero-depth refers to the unmined seafloor between the two caterpillar tracks along the same mining lane. The maximum depth shown has been formed after the sediment redeposition and consequently could be smaller than the actual erosion depth. c1) Orthophoto-mosaic of the track of the nodule collector while performing the light bulb turning manoeuvre at the end of a lane and the deposited nodule pile (≈ L 5 m × W 2 m × H 0.9 m). The sideways sediment extrusion under the caterpillar tracks is larger along the "light bulb" turn manoeuvres. c2) Image-derived DEM of the same area with a pixel resolution of 5 mm × 5 mm. Figure source: Gazis et al., 2025 <https://doi.org/10.1038/s41467-025-56311-0>

## Future Work

Four important outcomes regarding the monitoring technology of gravity flows and current driven sediment plumes in the deep-sea can be given, that also call for future investigations:

1. Need for a combined use of hydroacoustic and optical backscatter sensors. Both sensor types detected the plume signal and could differentiate between the active and passive phases of the mining activity. OBS systems are cheaper and lighter, can be attached to all platforms, have a high sampling rate, and can be calibrated with seafloor sediment in a small lab or onboard the research vessel. Their drawbacks include underestimating the SPM concentration when considerable particle flocculation occurs and the very small sampling volume of only several milliliters just in front of the sensor window. In contrast ADCPs and MBESs provide continuous measurements over long distances and altitudes (frequency dependent), allowing the integration over time and a large water volume. They benefit from the sediment aggregation over time and are less sensitive to very low SPM concentrations. ADCPs and MBESs technology is typically more expensive and challenging to deploy; they also have a disadvantage when being calibrated in the lab, where they demand deep tanks and complex sediment steering mechanisms. Ideally, each monitoring station on the seafloor, being part of a larger monitoring array, should have an ADCP system and an OBS, allowing for inter-calibration and data comparison between these two sensor types (e.g. Haalboom et al., 2022; Diaz et al., 2023).
2. Need for a combined use of sensors at fixed seafloor locations and on mobile platforms (i.e. AUVs). While sensors at fixed seafloor locations build a dense spatial network or array that allows measurement at successive distances (Eulerian approach), AUVs move together with the plume (Lagrangian approach) cover longer distances and carry several sensors. The technological advances in battery capacity (90+ h of survey) and plume tracking (von See et al., 2022; von See et al., 2025) pave the way towards automated plume monitoring. A fleet of AUVs that can exchange data (e.g., SPM concentration and location) with the mother ship and between each other can provide a real-time overview of the plume's spatiotemporal evolution. The AUVs are expected to be the cornerstone of monitoring the discharge plume at midwater depths or close to the seafloor. The current practice of water-column monitoring is based on ship-based sensors (e.g. CTD) and moorings deployment. Both approaches offer limited spatial coverage, have long deployment times, and depend on the sea state. Deep sea moorings cannot change locations without redeployment, and the ship-based CTDs have limited flexibility when needed to adapt the monitoring plan.
3. Need for real-time acoustic communication and data transfer to enable adaptive decision-making strategies, refined and adjusted based on the incoming data. Small-cabled sensor networks or repeated upload of AUV data at central data nodes can already be utilized today, allowing on-the-fly changes in the monitoring plan, similar to the work in the offshore dredging industry in shallower waters (Helmons et al., 2022).

Such an approach was tested during the second mining trial in the BGR contract area (Vink et al., 2022), where the real-time exchange of turbidity data provided valuable information at the time the ambient bottom currents changed direction.

4. Need for high-resolution underwater imagery of the water-column and the seafloor. This study shows that underwater image acquisition plays a vital role in plume and seafloor monitoring, providing, among others, the in-situ particle size, thickness and extent of sediment redeposition, excavation depth, and the direction of gravity current. Moreover, a data compilation of seafloor photographs and mosaics from different periods after the mining test shows the initial benthic ecosystem disturbance and its recovery. The author currently uses seafloor image-derived mosaics before (29-53 h), shortly after (88-110 h) and 1.5 years after the mining trial (SO295; Haeckel et al., 2023) to annotate the benthic megafauna and quantify the changes in abundance, biodiversity, spatial density and community composition within the mining site area.

The author participates in the third phase of the JPIO MiningImpact project, which will monitor the full-scale mining trial of a prototype nodule collector in CCZ. The knowledge acquired and research outcomes from the described mining trial will be applied (where feasible) and combined with simultaneous monitoring of the discharge plume.

## **Section III**

### Scientific development and achievements

*"The mind is not a vessel to be filled, but a fire to be kindled."*

*Plutarch, (45 – 120 BC)*

## Co-authorship in Peer-Reviewed Scientific Publications

Co-authorship is an essential part of scientific research aiming at synergy and skill transfer between research groups and institutes while improving the data analysis and enriching research outcomes, providing valuable insights from different scientific fields or sub-fields of the same discipline. During the PhD period, the author participated in writing five other published peer-reviewed research articles and one book chapter, grouped into three research fields.

### Underwater Imagery

1. Mbani, B., Schoening, T., **Gazis, I.-Z.**, Koch, R., & Greinert, J. (2022). Implementation of an automated workflow for image-based seafloor classification with examples from manganese-nodule covered seabed areas in the Central Pacific Ocean. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-19070-2>
2. Köser, K., She, M., Diller, N., Reissmann, S., Weiß, T., Heger, K., Song, Y., Schöntag, P., Nakath, D., Schoening, T., Mohrmann, J., **Gazis, I.-Z.**, Kampmeier, M., Rothenbeck, M., Wenzlaff, E., & Greinert, J. (2024). Autonomous visual 3D mapping of the ocean floor by underwater robots equipped with a single photo camera. In *Scanning Technologies for Autonomous Systems* (pp. 67–100). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-59531-8\\_3](https://doi.org/10.1007/978-3-031-59531-8_3)
3. Van Audenhaege, L., Mahamadaly, V., Price, D., Sneessens, A., Cawthra, H., Delamare, C., Danet, V., Delsol, S., Devillers, R., **Gazis, I.-Z.**, & Urbina-Barreto, I. (2023). Workshop on 3D mapping of habitats and biological communities with underwater photogrammetry. *Research Ideas and Outcomes*, 9. <https://doi.org/10.3897/rio.9.e115796>

### Deep-sea mining sediment plumes and oceanography-related monitoring

4. Haalboom, S., Schoening, T., Urban, P., **Gazis, I.-Z.**, de Stigter, H., Gillard, B., Baeye, M., Hollstein, M., Purkiani, K., Reichart, G.-J., Thomsen, L., Haeckel, M., Vink, A., & Greinert, J. (2022). Monitoring of anthropogenic sediment plumes in the Clarion-Clipperton Zone, NE equatorial Pacific Ocean. *Frontiers in Marine Science*, 9. <https://doi.org/10.3389/fmars.2022.882155>
5. Purkiani, K., Haeckel, M., Haalboom, S., Schmidt, K., Urban, P., **Gazis, I.-Z.**, de Stigter, H., Paul, A., Walter, M., & Vink, A. (2022). Impact of a long-lived anticyclonic mesoscale eddy on seawater anomalies in the northeastern tropical Pacific Ocean: a composite analysis from hydrographic measurements, sea level altimetry data, and reanalysis model products. *Ocean Science*, 18(4), 1163–1181. <https://doi.org/10.5194/os-18-1163-2022>

### Resources estimation and geostatistical analysis

6. Hasiotis, T., **Gazis, I.-Z.**, Anastasatou, M., Manoutsoglou, E., Velegrakis, A. F., Kapsimalis, V., Karditsa, A., & Stamatakis, M. (2021). Searching for potential marine sand resources to mitigate beach erosion in island settings. *Marine Georesources & Geotechnology*, 39(5), 527–542. <https://doi.org/10.1080/1064119x.2020.1721623>

## Data Publications

The publication of freely accessible research data (raw and post-processed ready-for-analysis data) is an integral part of scientific research as it enables other researchers or entities to reproduce, verify the results and conduct further research and meta-analysis. All datasets were published in the PANGAEA© data base (Data Publisher for Earth & Environmental Science) following the FAIR principles guidelines (Wilkinson et al., 2016). In total, 26 datasets were published during this PhD, with the author being the first author in 13 of these publications. It is worth mentioning that other researchers already used some of these datasets in high-ranked journals (e.g. Muñoz-Royo et al., 2021 in *Science Advances*). A chronologically sorted list of data publications is given below:

1. Stauffer, Julian B; Hoving, Henk-Jan T; **Gazis, Iason-Zois**; Mohrmann, Jochen; Kampmeier, Mareike; Veloso-Alarcón, Mario Enrique; Schumacher, Mia; Dierking, Jan (2024): Multibeam bathymetry raw data (Kongsberg EM712 entire dataset) of RV MARIA S. MERIAN during cruise MSM126 [dataset]. PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.968538>
2. Stauffer, Julian B; Hoving, Henk-Jan T; **Gazis, Iason-Zois**; Mohrmann, Jochen; Kampmeier, Mareike; Veloso-Alarcón, Mario Enrique; Schumacher, Mia; Dierking, Jan (2024): Multibeam bathymetry raw data (Kongsberg EM 122 entire dataset) of RV MARIA S. MERIAN during cruise MSM126 [dataset]. PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.968537>
3. **Gazis, I.-Z.**; Greinert, J (2023): Multibeam bathymetry raw data (Wärtsilä ELAC 3050 MK II entire dataset) of RV POSEIDON during cruise POS526, North Sea. PANGAEA, <https://doi.org/10.1594/PANGAEA.956775>
4. **Gazis, I.-Z.**; Thomadaki, M; Greinert, J (2023): Multibeam bathymetry processed data (Wärtsilä ELAC 3050 MK II entire dataset) of RV POSEIDON during cruise POS526, North Sea. PANGAEA, <https://doi.org/10.1594/PANGAEA.957183>
5. Haalboom, S; de Stigter, H; Urban, P; **Gazis, I.-Z.**; Gillard, B; Schoening, T (2022): Turbidity data measured during a dredge experiment during RV SONNE cruise SO268/2 in the German license area, CCZ, NE Pacific Ocean. PANGAEA, <https://doi.org/10.1594/PANGAEA.943396>
6. Baeye, M; Haalboom, S; Urban, P; **Gazis, I.-Z.**; Gillard, B; Schoening, T; de Stigter, H (2022): Turbidity data measured during a dredge experiment during RV SONNE cruise SO268/2 in the German license area, CCZ, NE Pacific Ocean. PANGAEA, <https://doi.org/10.1594/PANGAEA.942065>
7. Greinert, J; Haalboom, S; Urban, P; **Gazis, I.-Z.**; Gillard, B; Schoening, T; de Stigter, H (2022): Turbidity and current data measured during a dredge experiment during RV SONNE cruise SO268/2 in the German license area, CCZ, NE Pacific Ocean. PANGAEA, <https://doi.org/10.1594/PANGAEA.943331>
8. Haalboom, S; de Stigter, H; Urban, P; **Gazis, I.-Z.**; Gillard, B; Schoening, T (2022): Current data measured during a dredge experiment during RV SONNE cruise SO268/2 in the German license area, CCZ, NE Pacific Ocean. PANGAEA, <https://doi.org/10.1594/PANGAEA.943402>
9. **Gazis, I.-Z.**; Mohrmann, J; Schoening, T; Wöfl, A-C, 2021: Multibeam bathymetry processed data (Kongsberg EM 122 working area dataset) of RV MARIA S. MERIAN during cruise MSM96. PANGAEA, <https://doi.org/10.1594/PANGAEA.930063>

10. Mohrmann, J; **Gazis, I.-Z.**; Schoening, T; Wöfl, A-C, 2021: Multibeam bathymetry raw data (Kongsberg EM 122 entire dataset) of RV MARIA S. MERIAN during cruise MSM96. PANGAEA, <https://doi.org/10.1594/PANGAEA.929091>
11. Schoening, T; **Gazis, I.-Z.**; Mohrmann, J, 2021: Sediment echosounder raw data (Atlas Parasound P70 echosounder entire dataset) of RV MARIA S. MERIAN during cruise MSM96. PANGAEA, <https://doi.org/10.1594/PANGAEA.929548>
12. Schoening, T; Kopte, R; Mohrmann, J; **Gazis, I.-Z.**, 2020: ADCP current measurements (75 kHz) during Maria S. Merian cruise MSM96. PANGAEA, <https://doi.org/10.1594/PANGAEA.925938>
13. Schoening, T; Kopte, R; Mohrmann, J; **Gazis, I.-Z.**, 2020: ADCP current measurements (38 kHz) during Maria S. Merian cruise MSM96. PANGAEA, <https://doi.org/10.1594/PANGAEA.925936>
14. **Gazis, I.-Z.**; Urban, P.; Haeckel, M., 2020: Swath sonar multibeam EM122 bathymetry raw data during SONNE cruises SO268/1 and SO268/2, German License Area in Clarion Clipperton Zone, Pacific. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.org/10.1594/PANGAEA.919755>
15. **Gazis, I.-Z.**; Urban, P.; Haeckel, M., 2020: Swath sonar multibeam EM122 bathymetry raw data during SONNE cruises SO268/1 and SO268/2, Belgian License Area in Clarion Clipperton Zone, Pacific. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.org/10.1594/PANGAEA.919756>
16. **Gazis, I.-Z.**; Urban, P.; Haeckel, M.; Wöfl, A-C., 2020: Swath sonar multibeam EM122 bathymetry raw data during transits of SONNE cruises SO268/1 and SO268/2, Pacific. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.org/10.1594/PANGAEA.919757>
17. **Gazis, I.-Z.** 2020: Processed EM122 multibeam swath bathymetry collected during SONNE cruise SO268/1 inside the German Licensed Area in Clarion Clipperton Zone, Pacific. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA. <https://doi.org/10.1594/PANGAEA.915764>
18. **Gazis, I.-Z.** 2020: Processed AUV multibeam swath bathymetry collected during SONNE cruise SO268/1 inside the German Licensed Area in Clarion Clipperton Zone, Pacific. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA. <https://doi.org/10.1594/PANGAEA.915765>
19. **Gazis, I.-Z.** 2020: Processed EM122 multibeam swath bathymetry collected during SONNE cruise SO268/1 inside the Belgian Licensed Area in Clarion Clipperton Zone, Pacific. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA. <https://doi.org/10.1594/PANGAEA.915767>
20. **Gazis, I.-Z.**; Urban, P., 2020: Swath sonar multibeam EM122 water column data of R/V SONNE cruises SO268/1 and SO268/2 with links to wcd data files. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.914390>

21. **Gazis, I.-Z.**; Urban, P., 2020: Profile of sediment echo sounding during SONNE cruise SO268/1 with links to raw ParaSound data files. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.914607>
22. **Gazis, I.-Z.**; Urban, P., 2020: Profiles of vertical fish echo sounder Simrad EK60 during R/V SONNE cruises SO268/1 and SO268/2 with links to raw data files. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.914367>
23. **Gazis, I.-Z.**; Urban, P., 2020: Shipboard ADCP data during R/V SONNE cruises SO268/1 and SO268/2 with links to raw data files. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.913929>
24. Urban, P; **Gazis, I.-Z.**, 2020: Processed ROV multibeam swath bathymetry collected during SONNE cruise SO268/2 inside the Belgian Licensed Area in Clarion Clipperton Zone, Pacific. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.915766>
25. Schoening, T.; **Gazis, I.-Z.**, 2019: Sizes, weights, and volumes of polymetallic nodules from box cores taken during SONNE cruises SO268/1 and SO268/2. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.pangaea.de/10.1594/PANGAEA.904962>
26. Schoening, T; **Gazis, I.-Z.**, 2019: Summary of sizes, weights, counts and coverage of polymetallic nodules from box cores taken during SONNE cruises SO268/1 and SO268/2. GEOMAR-Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.org/10.1594/PANGAEA.904967>



**Figure 42.** Word cloud of the nine most frequently used words in the titles of the data publications. © Gazis, 2024 (unpublished)

## Scientific Expeditions

Offshore scientific expeditions are fundamental to marine research, particularly to deep-sea benthic and polymetallic nodule research. The author participated in six scientific expeditions and several field surveys to the Baltic Sea during his PhD years. The research expeditions and field surveys contributed significantly to the author's overall scientific development as they provided unprecedented field experience. The author was responsible for acquiring, storing, and post-processing hydroacoustic data, including MBES, SBES, ADCP, SSS, and SBP data. Furthermore, he was deeply involved in the design of AUV surveys (Research Objectives, time and dive setting optimization), image and oceanographic data acquisition, data post-processing and storage. The author was responsible for creating and updating the onboard GIS project (e.g., adding the planned and actual field stations, plotting acquired and existing data), supporting the onboard decision making and participating in writing the expedition reports. A chronologically sorted list of research expeditions is given below:

1. **IP21-MANGAN:** Independent scientific monitoring of two collector tests in the BGR and GSR contract areas for the exploration of polymetallic nodules in the equatorial NE Pacific. San Diego (USA) 04.04.2021 – San Diego (USA) 15.05.2021  
Cruise Report: Vink, et al., (2022). MANGAN 2021 Cruise Report: independent scientific monitoring of two collector tests in the BGR and GSR contract areas for the exploration of polymetallic nodules in the equatorial NE Pacific. BGR Report, 363 pp. <https://doi.org/10.25928/hw7d-fs42>
2. **MSM96 – METAL ML:** Heterogeneity and complexity of topography, habitats and geochemistry of the Atlantic Ocean seafloor. Emden (Germany) – Emden (Germany), 10.10.-10.11.2020  
Cruise Report: Schoening, Timm and Paul, Sophie (2021) Metal geochemistry meets machine learning: Assessing sedimentary and pore-water rare earth element variability from local to basin scale across the N.E. Atlantic Ocean, 10.10.2020 – 10.11.2020, Emden (Germany) – Emden (Germany) Metal-ML. Open Access. Maria S. Merian-Berichte, MSM96 . Gutachterpanel Forschungsschiffe, 77 pp. <https://oceanrep.geomar.de/id/eprint/52724/>
3. **SO268-1:** JPI Oceans-Environmental impacts and risks of deep-sea mining. Manzanillo (Mexico) – Manzanillo (Mexico), 18.02-27.03.2019  
Cruise Report: Haeckel, M., & Linke, P. (2021). RV SONNE Cruise Report SO268—Assessing the Impacts of Nodule Mining on the Deep-Sea Environment: Nodule Monitoring, Manzanillo (Mexico)—Vancouver (Canada), 17.02. – 27.05.2019 GEOMAR Report, N. Ser. 059. GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Kiel, Germany, 359 + Appendix (in all 802) pp. (2021). <https://oceanrep.geomar.de/id/eprint/54402/>
4. **MPII/112018:** Elucidating Patagonian Ice Sheets dynamics during the last deglaciation from the underwater geomorphology. Punta Arenas (Chile)- Punta Arenas (Chile), 2.11-18.11.2018
5. **POS530:** Environmental monitoring for the delaboration of munitions on the seabed (UDEMM project) inside the German EEZ (Baltic Sea). Kiel (Germany) – Kiel (Germany), 1.10 – 21.10. 2018  
Cruise Report: Kampmeier, Mareike, Greinert, Jens, Strehse, Jennifer, Beck, Aaron, Wichert, U., Diller, Nikolaj, Kurbjuhn, Torge, Wenzlaff, Emanuel and Schröder, Jens (2018) R.V. Poseidon POS530 Cruise Report "MineMoni 2018 "Mine Monitoring in the German Baltic Sea 2018; Environmental baseline study testing UDEMM Best Practices, 01rst – 21rst October 2018, Kiel (Germany) – Kiel (Germany). Open Access. GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, 45 pp. <https://oceanrep.geomar.de/id/eprint/47567/>

6. **SLOGARO M143:** Slope failures and active gas expulsion along the Romanian margin – investigating relations to gas hydrate distribution. Varna (Romania) – Heraklion (Greece), 12.12.-22.12.2017

Cruise Report: Riedel, Michael, Gausepohl, Florian, Gazis, Iason-Zois, Hähnel, Line, Kampmeier, Mareike, Urban, Peter and Bialas, Jörg, eds. (2018) RV METEOR Fahrtbericht / Cruise Report M143: SLOGARO: Slope failures and active gas expulsion along the Romanian margin – investigating relations to gas hydrate distribution, Varna (Bulgaria) – Heraklion (Greece), 12.12.-22.12.2017. Open Access. GEOMAR Report, N. Ser. 042 . GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany, 35 pp. DOI [https://doi.org/10.3289/GEOMAR\\_REP\\_NS\\_42\\_2018](https://doi.org/10.3289/GEOMAR_REP_NS_42_2018)



**Figure 43.** Photos during scientific expeditions: a) Sediment sampling in the Baltic Sea (POS530) together with Mareike Kampmeier (GEOMAR) and Dr. Jennifer Strehse (CAU). b) Smiles after successful data acquisition with Jochen Mohrmann (GEOMAR) and Karl Heger (GEOMAR) onboard MV Island Pride (IP21-MANGAN). c & d) MBES and SBP data acquisition and post-processing onboard RV Maria S.Merian (MSM969). e) Patagonian Ice Sheets Fjords mapping with Dr Mario Veloso-Alarcón onboard Marypaz II (GEOMAR; MPII/112018). f) Measuring the size of polymetallic nodules onboard the RV Sonne (SO268). g) Calibrating turbidity sensors together with Jochen Mohrmann (GEOMAR) onboard MV Island Pride (IP21-MANGAN). Figure source: Gazis, 2025 (PhD dissertation).

## Conferences

Participation in national and international scientific conferences has fostered the author's scientific exchange, interdisciplinary thinking, and collaboration, contributing to his research progress. Moreover, it has advanced the abilities of the author to present his research findings with confidence in the public. The author participated in 19 conferences as first author (12) and co-author (7). **The author received the first award for the best presentation among young researchers** in the Section “Energy and Digitalization in Sustainable Development” during the 12th Russian-German Raw Materials Conference (Saint Petersburg, Russia, 27-29 November 2019). A chronologically sorted list of conference participations is given below:

1. Stauffer J. B., Volpato M., Canning-Clode J., Dierking J., **Gazis I-Z.**, Greinert J., Mohrmann J., Hoving Henk-J.; 2025: Benthic habitat distribution and diversity on the slopes of Madeira. 17th Deep-Sea Biology Symposium, 12-17 January, Hongkong, China. (submitted)
2. **Gazis I.Z.**, Charlet F. Greinert J., 2024: Machine learning and backscatter for spatial mapping of deep-sea polymetallic nodules. 5-7 November, Rostock-Warnemünde, Germany. (accepted)
3. **Gazis I-Z.**, de Stigter H., Mohrmann J., Heger K., Diaz M., Gillard B., Baeye M., Veloso-Alarcón, M., Purkiani K., Haeckel M., Vink A., Thomsen L., Cuvelier D., Ramalho P., Greinert J., 2023: Deep-sea Mining Trial in the Clarion-Clipperton Zone: Advancements in resource estimation, plume monitoring, and environmental impact assessments. Underwater Minerals Conference, October 1-6, 2023, Rotterdam, Netherlands.
4. de Clippele L., Gafeira J., Arosio R., Huvenne V. A. I., Ragnarsson S., Kutti T., Larsson A. I., Sowers D., Diesing M., **Gazis I.-Z.**, Greinert J., Hebbeln D., Wienberg C., Titschack J., Wheeler A., Mohn C., Burmeister K., Schwarzkopf F., Orejas C., Roberts M., 2023: Using the novel CoMMA toolbox to map and characterize coral mounds across the Atlantic Ocean. 8th International Symposium on Deep-Sea Corals, Edinburgh, United Kingdom, 29 May-02 June 2023
5. **Gazis I.-Z.**, de Stigter H., Mohrmann J., Heger K., Diaz M., Gillard B., Baeye M., Veloso-Alarcón, M., Purkiani K., Thomsen L., Haeckel M., Vink A., Greinert J., 2023: Deep-sea Mining: Spatiotemporal seafloor and plume monitoring of in-situ trials in Clarion-Clipperton-Zone using hydroacoustic and optic methods, GEOHAB 2023 Marine Geological and Habitat Mapping Conference, La Réunion, France, 08-12 May 2023
6. **Gazis I.-Z.**, Charlet F., Greinert J., 2023: A multiple-model machine learning approach with increased interpretability for spatial prediction of polymetallic nodules using hydroacoustic and optic data (poster). GEOHAB 2023 Marine Geological and Habitat Mapping Conference, La Réunion, France, 08-12 May 2023
7. Diaz M., de Stigter H., Gillard B., **Gazis I.-Z.**, Mohrmann J., Heger K., Baeye M., Thomsen L., Greinert J., 2023: Monitoring flocculation during a deep-sea mining test in the Clarion-Clipperton Zone, eastern equatorial Pacific Ocean. EGU General Assembly 2023, Vienna, Austria & Online, 23-28 April 2023
8. **Gazis I.-Z.**, de Stigter H., Mohrmann J., Heger K., Diaz M., Gillard B., Baeye M., Purkiani K., Veloso-Alarcón, M., Hollstein M., Schoening T., Urban P., Haalboom S., Thomsen L., Haeckel M., Vink A., Charlet F., Greinert J., 2022: Deep-sea mining plume monitoring: Results from the first deep-sea trial of the Patania-II pre-prototype nodule collector and parallel plume monitoring with hydroacoustic and optic methods. Underwater Minerals Conference, October 2-7, 2022, The Vinoy Renaissance Resort, St. Petersburg, Florida, US

9. Purkiani K., **Gazis I.-Z.**, Mohrmann J., Heger K., Diaz M., Veloso-Alarcón, M., Gillard B., Baeye M., Paul A., de Stigter H., Haeckel M., Vink A., Greinert J., Thomsen L., Schulz M. 2022: Numerical simulation of deep-sea sediment dispersion induced by a pre-prototype nodule collector vehicle in the northeast tropical Pacific Ocean. Underwater Minerals Conference, October 2-7, 2022, The Vinoy Renaissance Resort, St. Petersburg, Florida, USA.
10. **Gazis I.Z.**, Greinert J., 2022: Machine learning-based modeling of deep-sea polymetallic nodules spatial distribution: spatial autocorrelation and model transferability at local scales. EGU General Assembly 2022, Vienna, Austria & Online, 23-27 May 2022.
11. Veloso-Alarcón, M., **Gazis I.-Z.**, Geiger, A., Bertrand Sebastien, and Cristián Rodrigo, 2022: Advances on submarine geomorphology at the Fjord system of Gran Campo Nevado (~52°S) EGU General Assembly 2022, Vienna, Austria & Online, 23-27 May 2022.
12. **Gazis I.Z.**, Greinert J., 2022: The effect of spatial autocorrelation in machine learning-based modeling of deep-sea polymetallic nodules spatial distribution. GEOHAB 2022 Marine Geological and Habitat Mapping Conference, Venice, Italy 16-20 May 2022.
13. **Gazis I.Z.**, Greinert J., Schoening T., 2019: Spatial quantitative mapping of polymetallic nodules based on machine learning. 12th Russian-German Raw Materials Conference, Saint Petersburg, Russia, 27-29 November 2019.
14. **Gazis I.Z.**, Greinert J., Schoening T., 2019: Manganese nodules spatial quantitative mapping based on AUV data and machine learning. Regional Conference on Geomorphology, Athens, Greece, 19-21 September 2019.
15. Geiger A.J., Smedley R.K., Garcia J.L., Rodriguez P.C., Herrera G., Ferri L., Veloso M., **Gazis I.Z.**, Bertrand S., 2019: A multi-method approach to constrain glacier geometry changes between the Gran Campo Nevado and Pacific Ocean during last glacial-deglacial phase. 20th Congress of the International Union for Quaternary Research (INQUA), Dublin, Ireland, 25-31 July 2019.
16. Alibakhshi A., Kumar A., **Gazis I.Z.**, Seelmann K., Juva K. 2018: Presenting the scientific research held at GEOMAR Helmholtz Centre for Ocean Research Kiel From the deep-sea to the Atmosphere. NexGen Conference, Helmholtz-Centre Potsdam, German Research Centre for Geosciences, 25-27 July 2018
17. **Gazis I.Z.**, Greinert J., Schoening T., 2018: Digital Data in Marine Resource Assessment: Challenges and Chances. GEOMAR DataScience Symposium, Kiel, Germany, 19-20 April 2018.
18. **Gazis I.Z.**, Greinert J., Schoening T., 2017: AUV-based small scale quantitative predictive mapping of Mn-nodules: Precondition for ideal mining path planning. Economical, Technological, and Environmental Aspects: Cooperative Solutions for Future Deep-sea Mining. 46th Underwater Mining Conference, Federation of German Industries (BDI), Berlin, Germany, 24-29 October 2017.
19. **Gazis I.Z.**, Hasiotis T., Velegrakis A.F., Anastasatou M., Karditsa A., Andreadis, O., Kapsimalis V., Stamatakis M., Poulos S., 2017: The potential of marine aggregate deposits off a highly eroded coastal area in Lesvos Isl. (Greece)-Implications for coastal management. 15th International Conference on Environmental Science and Technology, Rhodes, Greece, 31st August-2nd September 2017.



**Figure 44.** Photos from international conferences: **a)** First award in the Section Energy and Digitalization in Sustainable Development during the 12<sup>th</sup> Russian-German Raw Materials Conference, Saint Petersburg, Russia, 2019. Here, the Vice-Rector of the Saint Petersburg Mining University, Prof. Dr. Igor. B. Sergeyev presents the award to the author. For more details, please see <https://www.geomar.de/en/news/article/detection-of-raw-materials-research-by-machine-learning-honoured>. **b** and **c)** Oral and poster presentation at GEOHAB 2023 (Marine Geological and Habitat Mapping Conference) **d)** Oral presentation at GEOHAB 2022 (Marine Geological and Habitat Mapping Conference). Figure source: Gazis, 2025 (PhD dissertation).

## Workshops & Summer Schools

The GEOMAR Helmholtz Centre for Ocean Research Kiel, the Graduate Center at Kiel University and the Integrated School of Ocean Sciences (currently renamed to Foster Young Ocean Researcher Development/FYORD) offer various technical courses and workshops for PhD students. Participation in such workshops improves technical (e.g., programming, statistical analysis, scientific writing) and intellectual skills (e.g., analytical thinking and problem-solving, verbal communication and reasoning ability, knowledge transfer and teaching). The author participated in 16 workshops and summer schools, particularly in the first years of his PhD. The knowledge acquired was transferred to the PhD research. A list of workshops grouped based on their thematic areas is given below:

### Statistics, Programming and Machine Learning

- Statistics with R for marine scientists-Introduction to data exploration, regression, GLM and linear mixed-effects models, European University of the Seas, Kiel, 07-21 December 2021.
- Current Issues In Statistical Analysis, OpenGeoHub Summer School, Wageningen (virtual), 1-3 September 2021.
- Developing Machine Learning Algorithms for spatial and spatiotemporal data science problems, OpenGeoHub Summer School, Wageningen, 17-21 August 2020.
- First Steps in Python, Integrated School of Ocean Sciences, Kiel, 28 May 2019.
- Spatial Analysis in R, Integrated School of Ocean Sciences, Kiel, 9 March 2018.
- First Steps in R, Integrated School of Ocean Sciences, Kiel, 5 February 2018.

### MBES post-processing

- QPS Qimera training, GEOMAR/DSM Group Workshop, Kiel, 27-30 November 2017.

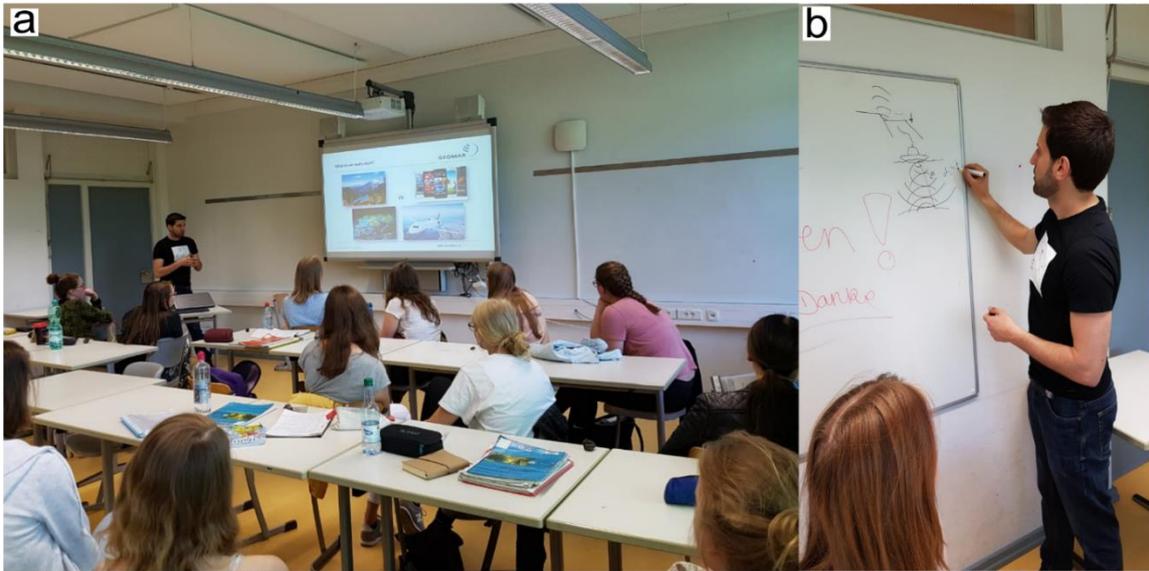
### Scientific writing and oral presentations

- Dealing with Questions – Communications Skills in Action, Graduate Center at Kiel University, Kiel, 1 February 2020.
- My publication – My Rights? GEOMAR Library Workshop, Kiel, 4 December 2017.
- Writing Scientific Publications, Graduate Center at Kiel University, Kiel, 22 September 2017.
- Speaking Clearly, Graduate Center at Kiel University, Kiel, 27-28 July 2017.

### Intellectual skills and teaching

- Understand, Formulate and Refute Arguments! Integrated School of Ocean Sciences, Kiel, June 7-8, 2022.
- Europa macht Schule 2018, DAAD and Department German as Foreign Language, Kiel, 30 May 2018
- With a School Class to Sylt, Integrated School of Ocean Sciences, Sylt, 6-7 February 2018.
- Management and Leadership in Academia, Integrated School of Ocean Sciences, Kiel, 7-8 September 2017.

- University Teaching, Integrated School of Ocean Sciences, Kiel, 13-14 July 2017.



**Figure 45.** Photos from workshops. A and b) Teaching students of the 11<sup>th</sup> class at the Regional Vocational Training Centre (aka RBZ1) in Königsweg, Kiel. The young pupils and teachers ranked this presentation as the best and most fascinating teaching session in a foreign language (English) of the Europa macht Schule 2018 project. Figure source: Gazis, 2025 (PhD dissertation).

### Section III – Supervision

#### Supervising students

Supervising master students doing an internship or being employed as student research assistants (aka HiWi) was a valuable experience for the author. He learned how to set up short-term projects that facilitate longer research aims, organize the research conducted, delegate weekly and daily tasks, and teach theoretical and technical skills simply and clearly. The author has **supervised four postgraduate** students and **five student research assistants** across five scientific fields: a) MBES and SSS post-processing, b) underwater orthophoto-mosaic reconstruction, c) deep-sea fauna annotation, d) spatial density estimation and e) machine learning. Their work has produced datasets and workflows that have been used and will be used in scientific and data publications.

## Section IV

### References

*" ...if I have seen further, it is by standing on the shoulders of giants."*

*Newton's letter to Robert Hooke, (Isaac Newton, 1675)*

## References

- Alevizos, E., & Greinert, J. (2018). The hyper-angular cube concept for improving the spatial and acoustic resolution of MBES backscatter angular response analysis. *Geosciences*, 8(12), 446. <https://doi.org/10.3390/geosciences8120446>
- Alevizos, E., Huvenne, V. A. I., Schoening, T., Simon-Lledó, E., Robert, K., & Jones, D. O. B. (2022). Linkages between sediment thickness, geomorphology and MN nodule occurrence: New evidence from AUV Geophysical mapping in the Clarion-clipperton zone. *deep-sea Research Part I Oceanographic Research Papers*, 179, 103645. <https://doi.org/10.1016/j.dsr.2021.103645>
- Aleynik, D., Inall, M. E., Dale, A., & Vink, A. (2017). Impact of remotely generated eddies on plume dispersion at abyssal mining sites in the Pacific. *Scientific Reports*, 7(1), 16959. <https://doi.org/10.1038/s41598-017-16912-2>
- Alpaydin, E. (2020). Introduction to machine learning. MIT press.
- Amon, D. J., Gollner, S., Morato, T., Smith, C. R., Chen, C., Christiansen, S., Currie, B., Drazen, J. C., Fukushima, T., Gianni, M., Gjerde, K. M., Gooday, A. J., Grillo, G. G., Haeckel, M., Joyini, T., Ju, S.-J., Levin, L. A., Metaxas, A., Mianowicz, K., ... Pickens, C. (2022). Assessment of scientific gaps related to the effective environmental management of deep-seabed mining. *Marine Policy*, 138(105006), 105006. <https://doi.org/10.1016/j.marpol.2022.105006>
- Amon, D. J., Ziegler, A. F., Dahlgren, T. G., Glover, A. G., Goineau, A., Gooday, A. J., Wiklund, H., & Smith, C. R. (2016). Insights into the abundance and diversity of abyssal megafauna in a polymetallic-nodule region in the eastern Clarion-Clipperton Zone. *Scientific Reports*, 6, 30492. <https://doi.org/10.1038/srep30492>
- Amos, A. F., & Roels, O. A. (1977). Environment aspects of manganese nodule mining. *Marine Policy*, 1(2), 156–163. [https://doi.org/10.1016/0308-597x\(77\)90050-1](https://doi.org/10.1016/0308-597x(77)90050-1)
- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. In Gabow, H. (Ed.). Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms. Society for Industrial and Applied Mathematics. pp. 1027–1035. Philadelphia, PA, USA.
- Baas, J. H., Best, J. L., & Peakall, J. (2016). Predicting bedforms and primary current stratification in cohesive mixtures of mud and sand. *Journal of the Geological Society*, 173(1), 12–45. <https://doi.org/10.1144/jgs2015-024>
- Barbiero, P., Squillero, G., & Tonda, A. (2020). *Modeling generalization in machine learning: A methodological and computational study*. arXiv:2006.15680. <https://doi.org/10.48550/ARXIV.2006.15680>
- Behrens, T., & Viscarra Rossel, R. A. (2020). On the interpretability of predictors in spatial data science: The information horizon. *Scientific Reports*, 10(1), 16737. <https://doi.org/10.1038/s41598-020-73773-y>
- Belkin, I. M., Andersson, P. S., & Langhof, J. (2021). On the discovery of ferromanganese nodules in the World Ocean. *Deep-Sea Research. Part I, Oceanographic Research Papers*, 175(103589), 103589. <https://doi.org/10.1016/j.dsr.2021.103589>
- Bemis, K. G. (2002). A comparison of black smoker hydrothermal plume behavior at Monolith Vent and at Clam 654 Acres Vent Field: Dependence on source configuration. *Marine Geophysical Researches*, 23, 81–96. <https://doi.org/10.1023/A:1022484931681>
- Benites, M., Millo, C., Hein, J., Nath, B., Murton, B., Galante, D., & Jovane, L. (2018). Integrated geochemical and morphological data provide insights into the genesis of ferromanganese nodules. *Minerals (Basel, Switzerland)*, 8(11), 488. <https://doi.org/10.3390/min8110488>
- Bischi, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, Algorithms, best practices, and open challenges. *WIREs Data Mining and Knowledge Discovery*, 13(2). <https://doi.org/10.1002/widm.1484>
- Bouasria, A., Bouslihim, Y., Gupta, S., Taghizadeh-Mehrjardi, R., & Hengl, T. (2023). Predictive performance of machine learning model with varying sampling designs, sample sizes, and spatial extents. *Ecological Informatics*, 78(102294), 102294. <https://doi.org/10.1016/j.ecoinf.2023.102294>
- Bowden, D. A., Anderson, O. F., Rowden, A. A., Stephenson, F., & Clark, M. R. (2021). Assessing habitat suitability models for the deep-sea: Is our ability to predict the distributions of seafloor fauna improving? *Frontiers in Marine Science*, 8. <https://doi.org/10.3389/fmars.2021.632389>
- Breiman, L. (2001a). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L. (2001b). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science A Review Journal of the Institute of Mathematical Statistics*, 16(3), 199–231.
- Breyer, G., Bartholomä, A., & Pesch, R. (2023). The suitability of machine-learning algorithms for the automatic acoustic seafloor classification of hard substrate habitats in the German bight. *Remote Sensing*, 15(16), 4113. <https://doi.org/10.3390/rs15164113>

- Brus, D. J. (2015). Balanced sampling: A versatile sampling approach for statistical soil surveys. *Geoderma*, 253–254, 111–121. <https://doi.org/10.1016/j.geoderma.2015.04.009>
- Burns, R. E. (1980). Assessment of environmental effects of deep ocean mining of manganese nodules. *Helgoländer Meeresuntersuchungen*, 33(1–4), 433–442. <https://doi.org/10.1007/bf02414768>
- Chunhui, T., Xiaobing, J., Aifei, B., Hongxing, L., Xianming, D., Jianping, Z., Chunhua, G., Tao, W., & Wilkens, R. (2015). Estimation of manganese nodule coverage using multi-beam amplitude data. *Marine Georesources & Geotechnology*, 33(4), 283–288. <https://doi.org/10.1080/1064119X.2013.806973>
- Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician*, 35(1), 54. <https://doi.org/10.2307/2683591>
- Cochonat, P., Le Suavé, R., Charles, C., Greger, B., Hoffert, M., Lenoble, J. P., Meunier, J., & Pautot, G. (1992). First in situ studies of nodule distribution and geotechnical measurements of associated deep-sea clay (Northeastern Pacific Ocean). *Marine Geology*, 103(1–3), 373–380. [https://doi.org/10.1016/0025-3227\(92\)90027-F](https://doi.org/10.1016/0025-3227(92)90027-F)
- COMPASS, 2018. Control of an Operational Mining Path Through an Auto-Adaptive Steering System Subproject: COMPASS-DRIMP-Producing DTMs and Mn-Nodule Reserve Data Layers for defining the Ideal-Mining-Path. <https://www.geomar.de/forschen/aktuelle-projekte/detailansicht/prj/355358?cHash=a4119113aebf77d79dad809f2f1ace56>
- Coulston, J. W., Blinn, C. E., Thomas, V. A., & Wynne, R. H. (2016). Approximating prediction uncertainty for random forest regression models. *Photogrammetric Engineering and Remote Sensing*, 82(3), 189–197. <https://doi.org/10.14358/pers.82.3.189>
- Craig, J. D. (1979). The relationship between bathymetry and ferromanganese deposits in the north equatorial Pacific. *Marine Geology*, 29(1–4), 165–186. [https://doi.org/10.1016/0025-3227\(79\)90107-5](https://doi.org/10.1016/0025-3227(79)90107-5)
- Cronan, D. S. (1975). Manganese nodules and other ferromanganese oxide deposits from the Atlantic Ocean. *Journal of Geophysical Research*, 80(27), 3831–3837. <https://doi.org/10.1029/jc080i027p03831>
- Cuvelier, D., Vigneron, M., Colaço, A., & Greinert, J. (2023). Delayed response of hermit crabs carrying anemones to a benthic impact experiment at the deep-sea nodule fields of the Peru Basin? *Marine Environmental Research*, 185(105899), 105899. <https://doi.org/10.1016/j.marenvres.2023.105899>
- De Smet, B., Pape, E., Riehl, T., Bonifácio, P., Colson, L., & Vanreusel, A. (2017). The community structure of deep-sea macrofauna associated with polymetallic nodules in the Eastern part of the Clarion-clipperton fracture zone. *Frontiers in Marine Science*, 4. <https://doi.org/10.3389/fmars.2017.00103>
- Diaz, M., Stigter, H., Gillard, B., Gazis, I.-Z., Mohrmann, J., Heger, K., Baeye, M., Thomsen, L., & Greinert, J. (2023). Monitoring flocculation during a deep-sea mining test in the Clarion-Clipperton Zone, eastern equatorial Pacific Ocean. <https://doi.org/10.5194/egusphere-egu23-17507>
- Diesing, M. (2020b). Deep-sea sediments of the global ocean. *Earth System Science Data*, 12(4), 3367–3381. <https://doi.org/10.5194/essd-12-3367-2020>
- Diesing, M., & Stephens, D. (2015). A multi-model ensemble approach to seabed mapping. *Journal of Sea Research*, 100, 62–69. <https://doi.org/10.1016/j.seares.2014.10.013>
- Diesing, M., Kröger, S., Parker, R., Jenkins, C., Mason, C., & Weston, K. (2017). Predicting the standing stock of organic carbon in surface sediments of the North–West European continental shelf. *Biogeochemistry*, 135(1–2), 183–200. <https://doi.org/10.1007/s10533-017-0310-4>
- Diesing, M., Thorsnes, T., & Bjarnadóttir, L. R. (2020a). Organic carbon in surface sediments of the North Sea and Skagerrak. <https://doi.org/10.5194/bg-2020-352>
- Dolan, M. F. J., Ross, R. E., Albretsen, J., Skarðhamar, J., Gonzalez-Mirelis, G., Bellec, V. K., Buhl-Mortensen, P., & Bjarnadóttir, L. R. (2021). Using spatial validity and uncertainty metrics to determine the relative suitability of alternative suites of oceanographic data for seabed biotope prediction. A case study from the Barents sea, Norway. *Geosciences*, 11(2), 48. <https://doi.org/10.3390/geosciences11020048>
- Dreisettl, I. (2016). deep-sea exploration for metal reserves—objectives, methods and look into the future. In T. Abramowski (Ed.), *deep-sea mining value chain: Organization, technology and development* (pp. 105–118). IOM.
- Durrieu de Madron, X., Ferré, B., Le Corre, G., Grenz, C., Conan, P., Pujó-Pay, M., Buscail, R., & Bodiou, O. (2005). Trawling-induced resuspension and dispersal of muddy sediments and dissolved elements in the Gulf of Lion (NW Mediterranean). *Continental Shelf Research*, 25(19–20), 2387–2409. <https://doi.org/10.1016/j.csr.2005.08.002>
- Dutkiewicz, A., Müller, R. D., Wang, X., O’Callaghan, S., Cannon, J., & Wright, N. M. (2017). Predicting sediment thickness on vanished ocean crust since 200 Ma. *Geochemistry, Geophysics, Geosystems: G(3)*, 18(12), 4586–4603. <https://doi.org/10.1002/2017gc007258>
- EC (2020) Critical Raw Materials Resilience: Charting a Path towards Greater Security and Sustainability. 2020. Communication from the commission to the European parliament, the council, the european economic and social committee and the committee of the regions. Critical Raw Materials Resilience: Charting a Path towards

- greater Security and Sustainability. COM/2020/474 final. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0474>
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species: The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), 330–342. <https://doi.org/10.1111/j.2041-210x.2010.00036.x>
- Ellefmo, S. L., & Kuhn, T. (2021). Application of soft data in nodule resource estimation. *Natural Resources Research*, 30(2), 1069–1091. <https://doi.org/10.1007/s11053-020-09777-2>
- Fernández-Delgado, M., Sirsat, M. S., Cernadas, E., Alawadi, S., Barro, S., & Febrero-Bande, M. (2019). An extensive experimental survey of regression methods. *Neural Networks*, 111, 11–34. <https://doi.org/10.1016/j.neunet.2018.12.010>
- Flood, R. D. (1983). Classification of sedimentary furrows and a model for furrow initiation and evolution. *Geological Society of America Bulletin*, 94(5), 630. [https://doi.org/10.1130/0016-7606\(1983\)94<630:cosfaa>2.0.co;2](https://doi.org/10.1130/0016-7606(1983)94<630:cosfaa>2.0.co;2)
- Flood, R. D., & Hollister, C. D. (1980). Submersible studies of deep-sea furrows and transverse ripples in cohesive sediments. *Marine Geology*, 36(1–2), M1–M9. [https://doi.org/10.1016/0025-3227\(80\)90035-3](https://doi.org/10.1016/0025-3227(80)90035-3)
- FMI-Flanders Marine Institute (2019). Maritime Boundaries Geodatabase, version 11. <https://www.marineregions.org/>
- Fonseca, L., & Calder, B. (2005). Geocoder: An efficient Backscatter map constructor. U.S. Hydrographic Conference <https://scholars.unh.edu/ccom/339/>
- Fonseca, L., & Mayer, L. (2007). Remote estimation of surficial seafloor properties through the application Angular Range Analysis to multibeam sonar data. *Marine Geophysical Research*, 28(2), 119–126. <https://doi.org/10.1007/s11001-007-9019-4>
- Fonseca, L., Brown, C., Calder, B., Mayer, L., & Rzhhanov, Y. (2009). Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and multibeam echosounder angular signatures. *Applied Acoustics*, 70(10), 1298–1304. <https://doi.org/10.1016/j.apacoust.2008.09.008>
- Foster SD, Monk J, Lawrence E, Hayes KR, Hosack GR, T. Langlois, Hooper G & Przeslawski R. (2020). Statistical considerations for monitoring and sampling. In *Field Manuals for Marine Sampling to Monitor Australian Waters, Version 2*. Przeslawski R, Foster S (Eds). National Environmental Science Program (NESP). <https://marine-sampling-field-manual.github.io/>
- Foster, S. D., Hosack, G. R., Hill, N. A., Barrett, N. S., & Lucieer, V. L. (2014). Choosing between strategies for designing surveys: autonomous underwater vehicles. *Methods in Ecology and Evolution*, 5(3), 287–297. <https://doi.org/10.1111/2041-210x.12156>
- Foster, S. D., Hosack, G. R., Monk, J., Lawrence, E., Barrett, N. S., Williams, A., & Przeslawski, R. (2020). Spatially balanced designs for transect-based surveys. *Methods in Ecology and Evolution*, 11(1), 95–105. <https://doi.org/10.1111/2041-210x.13321>
- Fouedjio, F., & Klump, J. (2019). Exploring prediction uncertainty of spatial data in geostatistical and machine learning approaches. *Environmental Earth Sciences*, 78(1). <https://doi.org/10.1007/s12665-018-8032-z>
- Fourcade, Y., Besnard, A. G., & Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Biogeography A Journal of Macroecology*, 27(2), 245–256. <https://doi.org/10.1111/geb.12684>
- Gaida, T. C., Mohammadloo, T. H., Snellen, M., & Simons, D. G. (2019). Mapping the seabed and shallow subsurface with multi-frequency multibeam echosounders. *Remote Sensing*, 12(1), 52. <https://doi.org/10.3390/rs12010052>
- Gaida, T., Tengku Ali, T., Snellen, M., Amiri-Simkooei, A., van Dijk, T., & Simons, D. (2018). A multispectral Bayesian classification method for increased acoustic discrimination of seabed sediments using multi-frequency multibeam backscatter data. *Geosciences*, 8(12), 455. <https://doi.org/10.3390/geosciences8120455>
- Gardner, W. D., Sullivan, L. G., & Thorndike, E. M. (1984). Long-term photographic, current, and nephelometer observations of manganese nodule environments in the Pacific. *Earth and Planetary Science Letters*, 70(1), 95–109. [https://doi.org/10.1016/0012-821x\(84\)90212-7](https://doi.org/10.1016/0012-821x(84)90212-7)
- Gausepohl, F., Hennke, A., Schoening, T., Köser, K., & Greinert, J. (2019a). *Bathymetric grid from the DISCOL working area of SONNE cruise SO242/1 in the Peru-basin* [Data set]. PANGAEA. <https://doi.org/10.1594/PANGAEA.905579>
- Gausepohl, F., Hennke, A., Schoening, T., Köser, K., & Greinert, J. (2019b). *Geo-referenced acoustic and optical data sets from the DISCOL area in the Peru-basin acquired during RV SONNE cruise SO242/1* [Data set]. PANGAEA. <https://doi.org/10.1594/PANGAEA.905616>
- Gazis, I.-Z. (2020). Processed EM122 multibeam swath bathymetry collected during SONNE cruise SO268/1 inside the Belgian License Area in Clarion Clipperton Zone, Pacific [Data set]. PANGAEA. <https://doi.org/10.1594/PANGAEA.915767>
- GEBCO (2024). Global ocean & land terrain models. [https://www.gebco.net/data\\_and\\_products/gridded\\_bathymetry\\_data/](https://www.gebco.net/data_and_products/gridded_bathymetry_data/)

- Gillard, B., Harbour, R. P., Nowald, N., Thomsen, L., & Iversen, M. H. (2022). Vertical distribution of particulate matter in the Clarion Clipperton Zone (German sector)—potential impacts from deep-sea mining discharge in the water column. *Frontiers in Marine Science*, 9. <https://doi.org/10.3389/fmars.2022.820947>
- Gillard, B., Purkiani, K., Chatzievangelou, D., Vink, A., Iversen, M. H., & Thomsen, L. (2019). Physical and hydrodynamic properties of deep-sea mining-generated, abyssal sediment plumes in the Clarion Clipperton Fracture Zone (eastern-central Pacific). *Elementa (Washington, D.C.)*, 7(1), 5. <https://doi.org/10.1525/elementa.343>
- Glasby, G. P. (1973). Distribution of manganese nodules and lebensspuren in underwater photographs from the Carlsberg Ridge, Indian Ocean. *New Zealand Journal of Geology and Geophysics*, 16(1), 1–17. <https://doi.org/10.1080/00288306.1973.10425383>
- Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143–151. [https://doi.org/10.1016/0954-1810\(94\)00011-s](https://doi.org/10.1016/0954-1810(94)00011-s)
- Greinert, J. (2016). Swath sonar multibeam EM122 bathymetry during SONNE cruise SO239 with links to raw data files [Data set]. PANGAEA. <https://doi.org/10.1594/PANGAEA.859456>
- Greinert, Jens, ed. (2015). RV SONNE Fahrtbericht / Cruise Report SO242-1 [SO242/1]: JPI OCEANS Ecological Aspects of Deep-Sea Mining, DISCOL Revisited, Guayaquil-Guayaquil (Equador), 28.07.-25.08.2015. GEOMAR Report, N. Ser. 026 . GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany, 290 pp. [https://doi.org/10.3289/GEOMAR\\_REP\\_NS\\_26\\_2015](https://doi.org/10.3289/GEOMAR_REP_NS_26_2015)
- GSR-Global Sea Mineral Resources NV (2018). Environmental Impact Statement. Small-Scale Testing of Nodule Collector Component on the Seafloor of the Clarion-Clipperton Fracture Zone and its Environmental Impact, p. 337. <https://www.isa.org.jm/minerals/environmental-impact-assessments>
- Guilhermic, C., Nardelli, M. P., Mouret, A., Le Moigne, D., & Howa, H. (2023). Short-term response of benthic foraminifera to fine-sediment depositional events simulated in microcosm. *Biogeosciences*, 20(15), 3329–3351. <https://doi.org/10.5194/bg-20-3329-2023>
- Guyon, I., Elisseeff, A., & Kaelbling, L. P. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3(7–8), 1157–1182.
- Haalboom, S., de Stigter, H. C., Mohn, C., Vandorpe, T., Smit, M., de Jonge, L., & Reichart, G.-J. (2023). Monitoring of a sediment plume produced by a deep-sea mining test in shallow water, Málaga Bight, Alboran Sea (southwestern Mediterranean Sea). *Marine Geology*, 456(106971), 106971. <https://doi.org/10.1016/j.margeo.2022.106971>
- Haalboom, S., de Stigter, H., Duineveld, G., van Haren, H., Reichart, G.-J., & Mienis, F. (2021). Suspended particulate matter in a submarine canyon (Whittard Canyon, Bay of Biscay, NE Atlantic Ocean): Assessment of commonly used instruments to record turbidity. *Marine Geology*, 434(106439), 106439. <https://doi.org/10.1016/j.margeo.2021.106439>
- Haalboom, S., Schoening, T., Urban, P., Gazis, I.-Z., de Stigter, H., Gillard, B., Baeye, M., Hollstein, M., Purkiani, K., Reichart, G.-J., Thomsen, L., Haeckel, M., Vink, A., & Greinert, J. (2022). Monitoring of anthropogenic sediment plumes in the Clarion-Clipperton Zone, NE equatorial Pacific ocean. *Frontiers in Marine Science*, 9. <https://doi.org/10.3389/fmars.2022.882155>
- Haeckel, M., & Linke, P. (2021). RV SONNE Cruise Report SO268—Assessing the Impacts of Nodule Mining on the Deep-Sea Environment: Nodule Monitoring, Manzanillo (Mexico)—Vancouver (Canada), 17.02. – 27.05.2019 GEOMAR Report, N. Ser. 059. GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Kiel, Germany, 359 + Appendix (in all 802) pp. (2021). <https://oceanrep.geomar.de/id/eprint/54402/>
- Haeckel, M., Janssen, F., and Martinez Arbizu, P. (2023) Assessing the impacts of polymetallic nodule mining on the deep-sea environment after an industrial collector test in the German and Belgian contract areas in the CCZ, Cruise No. SO295, 31 October 2022-23 December 2022, Port Hueneme (USA)-Port Hueneme (USA), NoduleMonitoring 2. Open Access . SONNE-Berichte, SO295 . Begutachterpanel Forschungsschiffe, Bonn, 598 pp. [https://doi.org/10.48433/cr\\_so295](https://doi.org/10.48433/cr_so295)
- Haeckel, M., König, I., Riech, V., Weber, M. E., & Suess, E. (2001). Pore water profiles and numerical modelling of biogeochemical processes in Peru Basin deep-sea sediments. *Deep-Sea Research. Part II, Topical Studies in Oceanography*, 48(17–18), 3713–3736. [https://doi.org/10.1016/s0967-0645\(01\)00064-9](https://doi.org/10.1016/s0967-0645(01)00064-9)
- Hao, T., Elith, J., Lahoz-Monfort, J. J., & Guillera-Aroita, G. (2020). Testing whether ensemble modelling is advantageous for maximizing predictive performance of species distribution models. *Ecography*, 43(4), 549–558. <https://doi.org/10.1111/ecog.04890>
- Hari, V. N., Kalyan, B., Chitre, M., & Ganesan, V. (2018). Spatial modeling of deep-sea ferromanganese nodules with limited data using neural networks. *IEEE Journal of Oceanic Engineering*, 43(4), 997–1014. <https://doi.org/10.1109/joe.2017.2752757>
- Haron, A., Paasche, H., Graber, S., Petersen, S., Attias, E., Jegen, M., Gehrmann, R., Hölz, S., & Klischies, M. (2023). Automated seafloor massive sulfide detection through integrated image segmentation and Geophysical Data Analysis: Revisiting the Tag Hydrothermal Field. *Geochemistry, Geophysics, Geosystems*, 24(12). <https://doi.org/10.1029/2023gc011250>

- Hasan, C., Ierodionou, R., Laurenson, D., & Schimel, L. (2014). Integrating Multibeam Backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. *PLoS ONE*, 9, e97339. <https://doi.org/10.1371/journal.pone.0097339>
- Hasan, R., Ierodionou, D., & Monk, J. (2012). Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar. *Remote Sensing*, 4(11), 3427–3443. <https://doi.org/10.3390/rs4113427>
- Hayes, S. P. (1979). Benthic current observations at DOMES sites A, B, and C in the tropical north Pacific Ocean. In J. L. Bischoff & D. Z. Piper (Eds.), *Marine geology and oceanography of the pacific manganese nodule province. Marine science.* (Vol. 9). Springer.
- Hein, J. R., Konstantinova, N., Mikesell, M., Mizell, K., Fitzsimmons, J. N., Lam, P. J., Jensen, L. T., Xiang, Y., Gartman, A., Cherkashov, G., Hutchinson, D. R., & Till, C. P. (2017). Arctic Deep Water ferromanganese-oxide deposits reflect the unique characteristics of the Arctic Ocean: AMERASIA ARCTIC OCEAN Fe-Mn DEPOSITS. *Geochemistry, Geophysics, Geosystems: G(3)*, 18(11), 3771–3800. <https://doi.org/10.1002/2017gc007186>
- Hein, J. R., Koschinsky, A., & Kuhn, T. (2020). Deep-ocean polymetallic nodules as a resource for critical materials. *Nature Reviews Earth & Environment*, 1(3), 158–169. <https://doi.org/10.1038/s43017-020-0027-0>
- Hein, J. R., Mizell, K., Koschinsky, A., & Conrad, T. A. (2013). Deep-ocean mineral deposits as a source of critical metals for high- and green-technology applications: Comparison with land-based resources. *Ore Geology Reviews*, 51, 1–14. <https://doi.org/10.1016/j.oregeorev.2012.12.001>
- Helmmons, R., de Wit, L., de Stigter, H., & Spearman, J. (2022). Dispersion of benthic plumes in deep-sea mining: What lessons can be learned from dredging? *Frontiers in Earth Science*, 10. <https://doi.org/10.3389/feart.2022.868701>
- Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B. M., & Gräler, B. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 6(e5518), e5518. <https://doi.org/10.7717/peerj.5518>
- Hillman, J. I. (2017). Validation of automated supervised segmentation of Multibeam Backscatter data from the Chatham Rise, New Zealand. *Marine Geophysical Research*, 39, 205–227. <https://doi.org/10.1007/s11001-016-9297-9>
- Huff, L. C. (2008). Acoustic remote sensing as a tool for habitat mapping in Alaska waters. *Marine habitat mapping technology for alaska* (pp. 29–46). Alaska Sea Grant, University of Alaska Fairbanks. <https://doi.org/10.4027/mhmta.2008.03>
- Hughes Clarke, J. (2018). The impact of acoustic imaging geometry on the fidelity of seabed bathymetric models. *Geosciences*, 8(4), 109. <https://doi.org/10.3390/geosciences8040109>
- IHO/International Hydrographic Organization (2011). Chapter 3 "Depth Determination" pp. 119-195 in *Manual on Hydrography*, Publication C-13. pp 548. [https://iho.int/uploads/user/pubs/cb/c-13/english/C-13\\_Chapter\\_1\\_and\\_contents.pdf](https://iho.int/uploads/user/pubs/cb/c-13/english/C-13_Chapter_1_and_contents.pdf)
- ISA/International Seabed Authority (2009). Establishment of a geological model of Polymetallic nodule deposits in the Clarion-clipperton fracture zone of the Equatorial north Pacific Ocean. Proceedings of the International Seabed Authority's Workshop held 13-20 May, 2003 in Nadi, Fiji. ISBN: 976-95217-3-5
- ISA/International Seabed Authority (2010). A geological model of polymetallic nodule deposits in the Clarion-Clipperton Fracture zone. *International Seabed Authority*, 211(6). <https://www.isa.org.jm/publications/technical-study-6-a-geological-model-of-polymetallic-nodule-deposits-in-the-clarion-clipperton-fracture-zone/>
- ISA/International Seabed Authority (2023). Maps: PMN Exploration Areas, PMN Reserved Areas, APEI Areas, <https://www.isa.org.jm/exploration-contracts/maps/>
- ISOPE/International Society of Offshore and Polar Engineers. (2002). *deep-seabed mining environment: Preliminary engineering and environmental assessment.*
- Jackson, D. R., & Briggs, K. B. (1992). High-frequency bottom backscattering: Roughness versus sediment volume scattering. *The Journal of the Acoustical Society of America*, 92(2), 962–977. <https://doi.org/10.1121/1.403966>
- Jackson, D. R., Winebrenner, D. P., & Ishimaru, A. (1986). Application of the composite roughness model to high-frequency bottom backscattering. *The Journal of the Acoustical Society of America*, 79(5), 1410–1422. <https://doi.org/10.1121/1.393669>
- Janecek, A., Gansterer, W., & Demel, M. (2008). *Gerhard Ecker Proceedings of the Workshop on New Challenges for Feature Selection in Data Mining and Knowledge Discovery at ECML/PKDD 2008.* 4, 90–105.
- Janowski, L., Wroblewski, R., Dworniczak, J., Kolakowski, M., Rogowska, K., Wojcik, M., & Gajewski, J. (2021). Offshore benthic habitat mapping based on object-based image analysis and geomorphometric approach. A case study from the Slupsk Bank, Southern Baltic Sea. *The Science of the Total Environment*, 801, 149712. <https://doi.org/10.1016/j.scitotenv.2021.149712>

- Jones, D. O. B., Kaiser, S., Sweetman, A. K., Smith, C. R., Menot, L., Vink, A., Trueblood, D., Greinert, J., Billett, D. S. M., Arbizu, P. M., Radziejewska, T., Singh, R., Ingole, B., Stratmann, T., Simon-Lledó, E., Durden, J. M., & Clark, M. R. (2017). Biological responses to disturbance from simulated deep-sea polymetallic nodule mining. *PLoS One*, 12(2), e0171750. <https://doi.org/10.1371/journal.pone.0171750>
- Josso, P., Hall, A., Williams, C., Le Bas, T., Lusty, P., & Murton, B. (2023). Application of random-forest machine learning algorithm for mineral predictive mapping of Fe-Mn crusts in the World Ocean. *Ore Geology Reviews*, 162(105671), 105671. <https://doi.org/10.1016/j.oregeorev.2023.105671>
- JPIO/Joint Programming Initiative Healthy and Productive Seas and Oceans: Ecological Aspects of Deep-Sea Mining, (2022). <https://jpi-oceans.eu/en/ecological-aspects-deep-sea-mining>
- Juan, C., Van Rooij, D., & De Bruycker, W. (2018). An assessment of bottom current controlled sedimentation in Pacific Ocean abyssal environments. *Marine Geology*, 403, 20–33. <https://doi.org/10.1016/j.margeo.2018.05.001>
- Juliani, C., & Juliani, E. (2021). Deep learning of terrain morphology and pattern discovery via network-based representational similarity analysis for deep-sea mineral exploration. *Ore Geology Reviews*, 129, 103936. <https://doi.org/10.1016/j.oregeorev.2020.103936>
- Kaikkonen, L., Virtanen, E. A., Kostamo, K., Lappalainen, J., & Kotilainen, A. T. (2019). Extensive coverage of marine mineral concretions revealed in shallow shelf sea areas. *Frontiers in Marine Science*, 6. <https://doi.org/10.3389/fmars.2019.00541>
- Karasiak, N., Dejoux, J.-F., Monteil, C., & Sheeren, D. (2022). Spatial dependence between training and test sets: Another pitfall of classification accuracy assessment in remote sensing. *Machine Learning*, 111(7), 2715–2740. <https://doi.org/10.1007/s10994-021-05972-1>
- Kattenborn, T., Schiefer, F., Frey, J., Feilhauer, H., Mahecha, M. D., & Dormann, C. F. (2022). Spatially autocorrelated training and validation samples inflate performance assessment of convolutional neural networks. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 5(100018), 100018. <https://doi.org/10.1016/j.ophoto.2022.100018>
- Kidd, D., Malone, B., McBratney, A., Minasny, B., & Webb, M. (2015). Operational sampling challenges to digital soil mapping in Tasmania, Australia. *Geoderma Regional*, 4, 1–10. <https://doi.org/10.1016/j.geodrs.2014.11.002>
- Kirkwood, C., Cave, M., Beamish, D., Grebby, S., & Ferreira, A. (2016). A machine learning approach to geochemical mapping. *Journal of Geochemical Exploration*, 167, 49–61. <https://doi.org/10.1016/j.jexplo.2016.05.003>
- Klein, H. (1993). Near-bottom currents in the deep Peru Basin, DISCOL experimental area. *Deutsche Hydrographische Zeitschrift*, 45(1), 31–42. <https://doi.org/10.1007/bf02226550>
- Klein, H. (1996). Near-bottom currents and bottom boundary layer variability over manganese nodule fields in the Peru Basin, se-pacific. *Deutsche Hydrographische Zeitschrift*, 48(2), 147–160. <https://doi.org/10.1007/bf02799384>
- Kodagali, V. (1988). Influence of regional and local topography on the distribution of polymetallic nodules in Central Indian Ocean Basin. *Geo-Marine Letters*, 8(3), 173–178. <https://doi.org/10.1007/bf02326094>
- Koller, D., Manica, R., Borges, A. de O., & Fedele, J. (2019). Experimental bedforms by saline density currents. *Brazilian Journal of Geology*, 49(2). <https://doi.org/10.1590/2317-4889201920180118>
- Köser, K., Song, Y., Petersen, L., Wenzlaff, E. & Woelk, F. (2021). Robustly removing deep-sea lighting effects for visual mapping of Abyssal Plains. arXiv.org (2021). <https://doi.org/10.48550/arXiv.2110.00480>
- Kuhn, M.; Johnson, K. (2013). *Applied Predictive Modeling*; Springer: New York, NY, USA; ISBN 978-1-4614-6848-6.
- Kuhn, T., & Röhlemann, C. (2021). Exploration of polymetallic nodules and resource assessment: A case study from the German contract area in the Clarion-Clipperton Zone of the tropical Northeast Pacific. *Minerals*, 11(6), 618. <https://doi.org/10.3390/min11060618>
- Kuhn, T., Wegorzewski, A., Röhlemann, C., & Vink, A. (2017). Composition, Formation, and Occurrence of Polymetallic Nodules. In R. Sharma (Ed.), *Deep-sea mining*. Springer. [https://doi.org/10.1007/978-3-319-52557-0\\_2](https://doi.org/10.1007/978-3-319-52557-0_2)
- Kursa, M. B., & Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. *J. Stat. Softw*, 36(11).
- Lauria, V., Power, A. M., Lordan, C., Weetman, A., & Johnson, M. P. (2015). Spatial transferability of habitat suitability models of *Nephrops norvegicus* among fished areas in the northeast Atlantic: Sufficiently stable for marine resource conservation? *PLoS ONE*, 10(2), e0117006. <https://doi.org/10.1371/journal.pone.0117006>
- Lawson, E., Smith, D., Sofge, D., Elmore, P., & Petry, F. (2017). Decision forests for machine learning classification of large, noisy seafloor feature sets. *Computers & Geosciences*, 99, 116–124. <https://doi.org/10.1016/j.cageo.2016.10.013>
- Lecours, V., Devillers, R., Edinger, E. N., Brown, C. J., & Lucieer, V. L. (2017). Influence of artefacts in marine digital terrain models on habitat maps and species distribution models: a multiscale assessment. *Remote Sensing in Ecology and Conservation*, 3(4), 232–246. <https://doi.org/10.1002/rse2.49>

- Lee, S. H., & Kim, K.-H. (2004). Side-scan sonar characteristics and manganese nodule abundance in the Clarion—Clipperton Fracture Zones, NE equatorial Pacific. *Marine Georesources & Geotechnology*, 22(1–2), 103–114. <https://doi.org/10.1080/10641190490473434>
- Lefaille, N., Macheriotou, L., Purkiani, K., Haeckel, M., Zeppilli, D., Pape, E., & Vanreusel, A. (2023). Digging deep: lessons learned from meiofaunal responses to a disturbance experiment in the Clarion-Clipperton Zone. *Marine Biodiversity*, 53(4). <https://doi.org/10.1007/s12526-023-01353-0>
- Li, J. (2019). A critical review of spatial predictive modeling process in environmental sciences with reproducible examples in R. *Applied Sciences (Basel, Switzerland)*, 9(10), 2048. <https://doi.org/10.3390/app9102048>
- Li, J., Alvarez, B., Siwabessy, J., Tran, M., Huang, Z., Przeslawski, R., Radke, L., Howard, F., & Nichol, S. (2017). Application of random forest, generalised linear model and their hybrid methods with geostatistical techniques to count data: Predicting sponge species richness. *Environmental Modelling & Software: With Environment Data News*, 97, 112–129. <https://doi.org/10.1016/j.envsoft.2017.07.016>
- Li, J., Tran, M., & Siwabessy, J. (2016). Selecting optimal random forest predictive models: A case study on predicting the spatial distribution of seabed hardness. *PLOS ONE*, 11(2). <https://doi.org/10.1371/journal.pone.0149089>
- Li, Z., Li, H., Hein, J. R., Dong, Y., Wang, M., Ren, X., Wu, Z., Li, X., & Chu, F. (2021). A possible link between seamount sector collapse and manganese nodule occurrence in the Abyssal Plains. *NW Pacific Ocean. Ore Geology Reviews*, 138, 104378. <https://doi.org/10.1016/j.oregeorev.2021.104378>
- Liaw, A. and Wiener, M. (2002) Classification and Regression by Randomforest. R News, 2, 18-22. [https://cran.r-project.org/doc/Rnews/Rnews\\_2002-3.pdf](https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf)
- Lipton, I., Nimmo, M., & Parianos, J. (2016). NI 43-101 Technical report. TOML Clarion Clipperton Zone Project, Pacific Ocean. <https://doi.org/10.13140/RG.2.2.23742.08000>
- Liu, Y., Carranza, E. J. M., & Xia, Q. (2022). Developments in quantitative assessment and modeling of mineral resource potential: An overview. *Natural Resources Research*, 31(4), 1825–1840. <https://doi.org/10.1007/s11053-022-10075-2>
- Lonsdale, P., & Spiess, F. N. (1977). Abyssal bedforms explored with a deeply towed instrument package. In *Developments in Sedimentology* (pp. 57–75). Elsevier.
- Lucieer, V., Huang, Z., & Siwabessy, J. (2016). Analyzing uncertainty in multibeam bathymetric data and the impact on derived seafloor attributes. *Marine Geodesy*, 39(1), 32–52. <https://doi.org/10.1080/01490419.2015.1121173>
- Ludwig, M., Moreno-Martinez, A., Hölzel, N., Pebesma, E., & Meyer, H. (2023). Assessing and improving the transferability of current global spatial prediction models. *Global Ecology and Biogeography: A Journal of Macroecology*, 32(3), 356–368. <https://doi.org/10.1111/geb.13635>
- Lurton, X. & Lamarche, G. (Eds) (2015). Backscatter measurements by seafloor-mapping sonars. Guidelines and Recommendations. 200p. <https://doi.org/10.5281/zenodo.10089261>
- Lurton, X., Eleftherakis, D., & Augustin, J.-M. (2018). Analysis of seafloor backscatter strength dependence on the survey azimuth using multibeam echosounder data. *Marine Geophysical Research*, 39(1–2), 183–203. <https://doi.org/10.1007/s11001-017-9318-3>
- Machida, S., Sato, T., Yasukawa, K., Nakamura, K., Iijima, K., Nozaki, T., & Kato, Y. (2019). Visualization method for the broad distribution of seafloor Ferromanganese deposits. *Marine Georesources & Geotechnology*, 39(3), 267–279. <https://doi.org/10.1080/1064119X.2019.1696432>
- Maciag, L., & Harff, J. (2020). Application of multivariate geostatistics for local-scale lithological mapping – case study of pelagic surface sediments from the Clarion-Clipperton Fracture Zone, north-eastern equatorial Pacific (Interoceanmetal claim area). *Computers & Geosciences*, 139(104474), 104474. <https://doi.org/10.1016/j.cageo.2020.104474>
- Magnuson, A. H. (1983). Manganese Nodule Abundance and Size from Bottom: Reflectivity Measurements. *Marine Mining*, 4(2–3), 265–296. <https://oceanrep.geomar.de/id/eprint/35811/>
- Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), 50–60. <https://doi.org/10.1214/aoms/1177730491>
- Marchig, V., von Stackelberg, U., Hufnagel, H., & Durn, G. (2001). Compositional changes of surface sediments and variability of manganese nodules in the Peru Basin. *deep-sea Research Part II Topical Studies in Oceanography*, 48(17–18), 3523–3547. [https://doi.org/10.1016/S0967-0645\(01\)00055-8](https://doi.org/10.1016/S0967-0645(01)00055-8)
- Mbani, B., Schoening, T., Gazis, I.-Z., Koch, R., & Greinert, J. (2022). Implementation of an automated workflow for image-based seafloor classification with examples from manganese-nodule covered seabed areas in the Central Pacific Ocean. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-19070-2>
- McCave, I. N., & Hall, I. R. (2006). Size sorting in marine muds: Processes, pitfalls, and prospects for paleoflow-speed proxies: SIZE SORTING IN MARINE MUDS. *Geochemistry, Geophysics, Geosystems: G(3)*, 7(10). <https://doi.org/10.1029/2006gc001284>
- Meinshausen, N. (2006). Quantile regression forests. *J. Mach. Learn. Res.*, 7, 983–999

- Menandro, P. S., Bastos, A. C., Misiuk, B., & Brown, C. J. (2022). Applying a multi-method framework to analyze the multispectral acoustic response of the seafloor. *Frontiers in Remote Sensing*, 3, 860282. <https://doi.org/10.3389/frsen.2022.860282>
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., Thuiller, W., Wüest, R. O., Zimmermann, N. E., & Eliith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37(12), 1267–1281. <https://doi.org/10.1111/ecog.00845>
- Mesgaran, M. B., Cousens, R. D., & Webber, B. L. (2014). Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Diversity & Distributions*, 20(10), 1147–1159. <https://doi.org/10.1111/ddi.12209>
- Mewes, K., Mogollón, J. M., Picard, A., Rühlemann, C., Kuhn, T., Nöthen, K., & Kasten, S. (2014). Impact of depositional and biogeochemical processes on small scale variations in nodule abundance in the Clarion-Clipperton Fracture Zone. *Deep-Sea Research Part I Oceanographic Research Papers*, 91, 125–141. <https://doi.org/10.1016/j.dsr.2014.06.001>
- Meyer, H., & Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, 12(9), 1620–1633. <https://doi.org/10.1111/2041-210x.13650>
- Meyer, H., & Pebesma, E. (2022). Machine learning-based global maps of ecological variables and the challenge of assessing them. *Nature Communications*, 13(1), 2208. <https://doi.org/10.1038/s41467-022-29838-9>
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software with Environment Data News*, 101, 1–9. <https://doi.org/10.1016/j.envsoft.2017.12.001>
- Milà, C., Mateu, J., Pebesma, E., & Meyer, H. (2022). Nearest neighbour distance matching Leave-One-Out Cross-Validation for map validation. *Methods in Ecology and Evolution*, 13(6), 1304–1316. <https://doi.org/10.1111/2041-210X.13851>
- Miljutin, D. M., Miljutina, M. A., Arbizu, P. M., & Galéron, J. (2011). Deep-sea nematode assemblage has not recovered 26 years after experimental mining of polymetallic nodules (Clarion-Clipperton Fracture Zone, Tropical Eastern Pacific). *Deep-Sea Research. Part I, Oceanographic Research Papers*, 58(8), 885–897. <https://doi.org/10.1016/j.dsr.2011.06.003>
- Minasny. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, 32, 1378–1388. <https://doi.org/10.1016/j.cageo.2005.12.009>
- MiningImpact (2022): Environmental impacts and risks of deep-sea mining. <https://miningimpact.geomar.de>
- Misiuk, B., & Brown, C. J. (2022). Multiple imputation of multibeam angular response data for high resolution full coverage seabed mapping. *Marine Geophysical Research*, 43(1), 7. <https://doi.org/10.1007/s11001-022-09471-3>
- Misiuk, B., & Brown, C. J. (2023). Improved environmental mapping and validation using bagging models with spatially clustered data. *Ecological Informatics*, 77, 102181. <https://doi.org/10.1016/j.ecoinf.2023.102181>
- Misiuk, B., & Brown, C. J. (2024). Benthic habitat mapping: A review of three decades of mapping biological patterns on the seafloor. *Estuarine, Coastal and Shelf Science*, 296(108599), 108599. <https://doi.org/10.1016/j.ecss.2023.108599>
- Misiuk, B., Diesing, M., Aitken, A., Brown, C. J., Edinger, E. N., & Bell, T. (2019). A spatially explicit comparison of quantitative and categorical modelling approaches for mapping seabed sediments using Random Forest. *Geosciences*, 9(6), 254. <https://doi.org/10.3390/geosciences9060254>
- Misiuk, B., Lecours, V., & Bell, T. (2018). A multiscale approach to mapping seabed sediments. *PloS One*, 13(2), e0193647. <https://doi.org/10.1371/journal.pone.0193647>
- Misiuk, B., Tan, Y. L., Li, M. Z., Trappenberg, T., Alleosfour, A., Church, I. W., Ferrini, V., & Brown, C. J. (2024). Multivariate mapping of seabed grain size parameters in the Bay of Fundy using convolutional neural networks. *Marine Geology*, 472(107299), 107299. <https://doi.org/10.1016/j.margeo.2024.107299>
- Mohrmann, J., & Greinert, J. (2022). AUV navigation correction based on automated multibeam tile matching. *Sensors (Basel, Switzerland)*, 22(3), 954. <https://doi.org/10.3390/s22030954>
- Molari, M., Janssen, F., Vonnahme, T. R., Wenzhöfer, F., & Boetius, A. (2020). The contribution of microbial communities in polymetallic nodules to the diversity of the deep-sea microbiome of the Peru Basin (4130–4198 m depth). *Biogeosciences*, 17(12), 3203–3222. <https://doi.org/10.5194/bg-17-3203-2020>
- Moreno-Amat, E., Mateo, R. G., Nieto-Lugilde, D., Morueta-Holme, N., Svenning, J.-C., & García-Amorena, I. (2015). Impact of model complexity on cross-temporal transferability in Maxent species distribution models: An assessment using paleobotanical data. *Ecological Modelling*, 312, 308–317. <https://doi.org/10.1016/j.ecolmodel.2015.05.035>
- Moustier, C. de (1986). Beyond bathymetry: Mapping acoustic backscattering from the deep-seafloor with Sea Beam. *The Journal of the Acoustical Society of America*, 79(2), 316–331. <https://doi.org/10.1121/1.393570>

- Moustier, C. de. (1985). Inference of manganese nodule coverage from Sea Beam acoustic backscattering data. *Geophysics*, 50(6), 989–1001. <https://doi.org/10.1190/1.1441976>
- Mullineaux, L. S. (1989). Vertical distributions of the epifauna on manganese nodules: Implications for settlement and feeding: Epifaunal distributions on nodules. *Limnology and Oceanography*, 34(7), 1247–1262. <https://doi.org/10.4319/lo.1989.34.7.1247>
- Muñoz-Royo, C., Ouillon, R., El Mousadik, S., Alford, M. H., & Peacock, T. (2022). An in situ study of abyssal turbidity-current sediment plumes generated by a deep-seabed polymetallic nodule mining preprototype collector vehicle. *Science Advances*, 8(38). <https://doi.org/10.1126/sciadv.abn1219>
- Naimi, B., Skidmore, A. K., Groen, T. A., & Hamm, N. A. S. (2011). Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling: Spatial autocorrelation and positional uncertainty. *Journal of Biogeography*, 38(8), 1497–1509. <https://doi.org/10.1111/j.1365-2699.2011.02523.x>
- NAS/National Academy of Sciences, National Academy of Engineering, and Institute of Medicine (2005). *Facilitating Interdisciplinary Research*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/11153>
- Newton, I., & Hooke, R. Isaac Newton letter to Robert Hooke, 1675. <https://digitallibrary.hsp.org/index.php/Detail/objects/9792>
- O'Neill, F. G., Simmons, S. M., Parsons, D. R., Best, J. L., Copland, P. J., Armstrong, F., Breen, M., & Summerbell, K. (2013). Monitoring the generation and evolution of the sediment plume behind towed fishing gears using a multibeam echosounder. *ICES Journal of Marine Science: Journal Du Conseil*, 70(4), 892–903. <https://doi.org/10.1093/icesjms/fst051>
- Ostwald, J., & Frazer, F. W. (1973). Chemical and mineralogical investigations on deep-sea manganese nodules from the Southern Ocean. *Mineralium Deposita*, 8(4). <https://doi.org/10.1007/bf00207513>
- Owens, H. L., Campbell, L. P., Dornak, L. L., Saupé, E. E., Barve, N., Soberón, J., Ingenloff, K., Lira-Noriega, A., Hensz, C. M., Myers, C. E., & Peterson, A. T. (2013). Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling*, 263, 10–18. <https://doi.org/10.1016/j.ecolmodel.2013.04.011>
- Ozturgut, E., Lavelle, J. W., & Burns, R. E. (1981). *Chapter 15 impacts of manganese nodule mining on the environment: 530 Results from pilot-scale mining tests in the North Equatorial Pacific*. Elsevier Oceanography Series.
- Pape, E., Bezerra, T. N., Gheerardyn, H., Buydens, M., Kieswetter, A., & Vanreusel, A. (2021). Potential impacts of polymetallic nodule removal on deep-sea meiofauna. *Scientific Reports*, 11(1), 19996. <https://doi.org/10.1038/s41598-021-99441-3>
- Parianos, J., Lipton, I., & Nimmo, M. (2021). Aspects of estimation and reporting of mineral resources of seabed polymetallic nodules: A contemporaneous case study. *Minerals*, 11(2), 200. <https://doi.org/10.3390/min11020200>
- Paul, S. A. L., Haeckel, M., Bau, M., Bajracharya, R., & Koschinsky, A. (2019). Small-scale heterogeneity of trace metals including rare earth elements and yttrium in deep-sea sediments and porewaters of the Peru Basin, southeastern equatorial Pacific. *Biogeosciences*, 16(24), 4829–4849. <https://doi.org/10.5194/bg-16-4829-2019>
- Peacock, T., & Ouillon, R. (2023). The Fluid Mechanics of deep-sea mining. *Annual Review of Fluid Mechanics*, 55(1). <https://doi.org/10.1146/annurev-fluid-031822-010257>
- Peukert, A., Petersen, S., Greinert, J., & Charlot, F. (2018a). Seabed Mining. In *Springer Geology* (pp. 481–502). Springer International Publishing. [https://doi.org/10.1007/978-3-319-57852-1\\_24](https://doi.org/10.1007/978-3-319-57852-1_24)
- Peukert, A., Schoening, T., Alevizos, E., Köser, K., Kwasnitschka, T., & Greinert, J. (2018b). Understanding Mn-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and optical data. *Biogeosciences*, 15(8), 2525–2549. <https://doi.org/10.5194/bg-15-2525-2018>
- Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., & Pélissier, R. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, 11(1), 4540. <https://doi.org/10.1038/s41467-020-18321-y>
- Pollerspöck, J., Cares, D., Ebert, D. A., Kelley, K. A., Pockalny, R., Robinson, R. S., Wagner, D., & Straube, N. (2023). First *in situ* documentation of a fossil tooth of the megatooth shark *Otodus (Megaselachus) megalodon* from the deep-sea in the Pacific Ocean. *Historical Biology*, 1–6. <https://doi.org/10.1080/08912963.2023.2291771>
- Porskamp, P., Schimel, A. C. G., Young, M., Rattray, A., Ladroit, Y., & Ierodiaconou, D. (2022). Integrating multibeam echosounder water-column data into benthic habitat mapping. *Limnology and Oceanography*, 67(8), 1701–1713. <https://doi.org/10.1002/lno.12160>
- Purkiani, K., Gillard, B., Paul, A., Haeckel, M., Haalboom, S., Greinert, J., de Stigter, H., Hollstein, M., Baeye, M., Vink, A., Thomsen, L., & Schulz, M. (2021). Numerical simulation of deep-sea sediment transport induced

- by a dredge experiment in the northeastern Pacific Ocean. *Frontiers in Marine Science*, 8. <https://doi.org/10.3389/fmars.2021.719463>
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- R Studio Team (2022). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>
- Repko, Allen F., Rick Szostak, and Michelle Phillips Buchberger. *Introduction to Interdisciplinary Studies*. 2014. Print.
- Reykhart, L. Y., & Shulga, N. A. (2019). Fe-Mn nodule morphotypes from the NE Clarion-Clipperton Fracture Zone, Pacific Ocean: Comparison of mineralogy, geochemistry and genesis. *Ore Geology Reviews*, 110(102933), 102933. <https://doi.org/10.1016/j.oregeorev.2019.102933>
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Roberts, J. J., Best, B. D., Dunn, D. C., Treml, E. A., & Halpin, P. N. (2010). Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. *Environmental Modelling & Software: With Environment Data News*, 25(10), 1197–1207. <https://doi.org/10.1016/j.envsoft.2010.03.029>
- Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, 71, 804–818. <https://doi.org/10.1016/j.oregeorev.2015.01.001>
- Sappington, J. M., Longshore, K. M., & Thompson, D. B. (2007). Quantifying landscape ruggedness for animal habitat analysis: A case study using bighorn sheep in the Mojave Desert. *The Journal of Wildlife Management*, 71(5), 1419–1426. <https://doi.org/10.2193/2005-723>
- Savoca, S., Di Fresco, D., Alesci, A., Capillo, G., & Spanò, N. (2022). Mucus secretions in cnidarian, an ecological, adaptive and Evolutive Tool. *Advances in Oceanography and Limnology*, 13. <https://doi.org/10.4081/aiol.2022.11054>
- Scanlon, K. M., & Masson, D. G. (1992). Fe–Mn nodule field indicated by GLORIA, north of the Puerto Rico Trench. *Geo-Marine Letters*, 12, 208–213. <https://doi.org/10.1007/BF02091840>
- Schaffer, C. (1993). Overfitting avoidance as bias. *Machine Learning*, 10(2), 153–178. <https://doi.org/10.1007/bf00993504>
- Schmidt, K., Behrens, T., Daumann, J., Ramirez-Lopez, L., Werban, U., Dietrich, P., & Scholten, T. (2014). A comparison of calibration sampling schemes at the field scale. *Geoderma*, 232–234, 243–256. <https://doi.org/10.1016/j.geoderma.2014.05.013>
- Schoening, T., Jones, D. O. B., & Greinert, J. (2017). Compact-morphology-based polymetallic nodule delineation. *Scientific Reports*, 7(1), 13338. <https://doi.org/10.1038/s41598-017-13335-x>
- Schoening, T., Kuhn, T., Jones, D. O. B., Simon-Lledo, E., & Nattkemper, T. W. (2016). Fully automated image segmentation for benthic resource assessment of poly-metallic nodules. *Methods in Oceanography*, 15–16, 78–89. <https://doi.org/10.1016/j.mio.2016.04.002>
- Schratz, P., Muenchow, J., Iturriza, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406, 109–120. <https://doi.org/10.1016/j.ecolmodel.2019.06.002>
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.
- Shao, M., Song, W., & Zhao, X. (2023). Polymetallic nodule resource assessment of seabed photography based on denoising diffusion probabilistic models. *Journal of Marine Science and Engineering*, 11(8), 1494. <https://doi.org/10.3390/jmse11081494>
- Sharma, R. (1993). Quantitative estimation of seafloor features from photographs and their application to nodule mining. *Marine Georesources & Geotechnology*, 11(4), 311–331. <https://doi.org/10.1080/10641199309379926>
- Sharma, R. (2017). Assessment of distribution characteristics of polymetallic nodules and their implications on deep-sea mining. In R. Sharma (Ed.), *Deep-Sea mining*. Springer. [https://doi.org/10.1007/978-3-319-52557-0\\_8](https://doi.org/10.1007/978-3-319-52557-0_8)
- She, M., Song, Y., Nakath, D., & Köser, K. (2023). Semihierarchical reconstruction and weak-area revisiting for robotic visual seafloor mapping. *arXiv [Cs.CV]*. <https://doi.org/10.48550/ARXIV.2308.06147>
- Shmueli, G. (2010). To explain or to predict? *Statistical Science A Review Journal of the Institute of Mathematical Statistics*, 25(3), 289–310.
- Simon-Lledó, E., Amon, D. J., Bribiesca-Contreras, G., Cuvelier, D., Durden, J. M., Ramalho, S. P., Uhlenkott, K., Arbizu, P. M., Benoist, N., Copley, J., Dahlgren, T. G., Glover, A. G., Fleming, B., Horton, T., Ju, S.-J.,

- Mejía-Saenz, A., McQuaid, K., Pape, E., Park, C., ... Jones, D. O. B. (2023). Carbonate compensation depth drives abyssal biogeography in the northeast Pacific. *Nature Ecology & Evolution*, 7(9), 1388–1397. <https://doi.org/10.1038/s41559-023-02122-9>
- Simon-Lledó, E., Bett, B. J., Huvenne, V. A. I., Schoening, T., Benoist, N. M. A., Jeffreys, R. M., Durden, J. M., & Jones, D. O. B. (2019). Megafaunal variation in the abyssal landscape of the Clarion Clipperton Zone. *Progress in Oceanography*, 170, 119–133. <https://doi.org/10.1016/j.pocean.2018.11.003>
- Skorniyakova, N. S., & Murdmaa, I. O. (1992). Local variations in distribution and composition of ferromanganese nodules in the Clarion-Clipperton Nodule Province. *Marine Geology*, 103(1–3), 381–405. [https://doi.org/10.1016/0025-3227\(92\)90028-G](https://doi.org/10.1016/0025-3227(92)90028-G)
- Song, W., Zheng, N., Liu, X., Qiu, L., & Zheng, R. (2019). An improved U-net convolutional networks for seabed mineral image segmentation. *IEEE Access: Practical Innovations, Open Solutions*, 7, 82744–82752. <https://doi.org/10.1109/access.2019.2923753>
- Song, Y., Nakath, D., She, M., & Köser, K. (2022). Optical imaging and image restoration techniques for deep ocean mapping: A comprehensive survey. *PGF – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90(3), 243–267. <https://doi.org/10.1007/s41064-022-00206-y>
- Song, Y., She, M., & Köser, K. (2024). Advanced underwater image restoration in complex illumination conditions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 209, 197–212. <https://doi.org/10.1016/j.isprsjprs.2024.02.004>
- SPC (2013). deep-sea Minerals: Manganese Nodules, a physical, biological, environmental, and technical review. Baker, E., and Beaudoin, Y. (Eds.) Vol. 1B, Secretariat of the Pacific Community. ISBN 978-82-7701-120-2
- Spearman, J., Taylor, J., Crossouard, N., Cooper, A., Turnbull, M., Manning, A., Lee, M., & Murton, B. (2020). Measurement and modelling of deep-sea sediment plumes and implications for deep-sea mining. *Scientific Reports*, 10(1), 5075. <https://doi.org/10.1038/s41598-020-61837-y>
- Spiegel, T., Diesing, M., Dale, A. W., Lenz, N., Schmidt, M., Sommer, S., Böttner, C., Fuhr, M., Kalapurakkal, H. T., Schulze, C.-S., & Wallmann, K. (2024). Modelling mass accumulation rates and 210PB rain rates in the Skagerrak: Lateral sediment transport dominates the sediment input. *Frontiers in Marine Science*, 11, 1331102. <https://doi.org/10.3389/fmars.2024.1331102>
- Stenvers, V. I., Hauss, H., Bayer, T., Havermans, C., Hentschel, U., Schmittmann, L., Sweetman, A. K., & Hoving, H.-J. T. (2023). Experimental mining plumes and ocean warming trigger stress in a deep pelagic jellyfish. *Nature Communications*, 14(1), 7352. <https://doi.org/10.1038/s41467-023-43023-6>
- Stephens, D., & Diesing, M. (2014). A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. *PloS One*, 9(4), e93950. <https://doi.org/10.1371/journal.pone.0093950>
- Sterkenburg, T. F., & Grünwald, P. D. (2021). The no-free-lunch theorems of supervised learning. *Synthese*, 199(3–4), 9979–10015. <https://doi.org/10.1007/s11229-021-03233-1>
- Stienessen, S. C., Rooper, C. N., Webe, T. C., Jones, D. T., Pirtle, J. L., & Wilson, C. D. (2021). Comparison of model types for prediction of seafloor trawlability in the Gulf of Alaska by using multibeam sonar data. *Fishery Bulletin*, 119(2–3), 184–196. <https://doi.org/10.7755/FB.119.2-3.7>
- Sweetman, A. K., Smith, A. J., de Jonge, D. S. W., Hahn, T., Schroedl, P., Silverstein, M., Andrade, C., Edwards, R. L., Lough, A. J. M., Wouds, C., Homoky, W. B., Koschinsky, A., Fuchs, S., Kuhn, T., Geiger, F., & Marlow, J. J. (2024). Evidence of dark oxygen production at the abyssal seafloor. *Nature Geoscience*, 17(8), 737–739. <https://doi.org/10.1038/s41561-024-01480-8>
- Thomsen, L., & Gust, G. (2000). Sediment erosion thresholds and characteristics of resuspended aggregates on the western European continental margin. *Deep-Sea Research. Part I, Oceanographic Research Papers*, 47(10), 1881–1897. [https://doi.org/10.1016/s0967-0637\(00\)00003-0](https://doi.org/10.1016/s0967-0637(00)00003-0)
- Thomson, W. Notes from the "Challenger" \*. *Nature* 8, 51–53 (1873).
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234. <https://doi.org/10.2307/143141>
- Tomczak, A., Kogut, T., Kabała, K., Abramowski, T., Ciałzela, J., & Giza, A. (2024). Automated estimation of offshore polymetallic nodule resources based on seafloor imagery using deep learning. <https://doi.org/10.2139/ssrn.4755951>
- Trzcinska, K., Janowski, L., Nowak, J., Rucinska-Zjadacz, M., Kruss, A., von Deimling, J. S., Pocwiardowski, P., & Tegowski, J. (2020). Spectral features of dual-frequency multibeam echosounder data for benthic habitat mapping. *Marine Geology*, 427, 106239. <https://doi.org/10.1016/j.margeo.2020.106239>
- Tsune, A. (2021). Quantitative expression of the burial phenomenon of deep-seafloor manganese nodules. *Minerals*, 11(2), 227. <https://doi.org/10.3390/min11020227>
- Uhlenkott, K., Simon-Lledó, E., Vink, A., & Martínez Arbizu, P. (2022). Investigating the benthic megafauna in the eastern Clarion Clipperton Fracture Zone (north-east Pacific) based on distribution models predicted with random forest. *Scientific Reports*, 12(1), 8229. <https://doi.org/10.1038/s41598-022-12323-0>

- Uhlenkott, K., Vink, A., Kuhn, T., & Martínez Arbizu, P. (2020). Predicting meiofauna abundance to define preservation and impact zones in a deep-sea mining context using random forest modelling. *The Journal of Applied Ecology*, 57(7), 1210–1221. <https://doi.org/10.1111/1365-2664.13621>
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). blockCV: An r package for generating spatially or environmentally separated folds for *k*-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10(2), 225–232. <https://doi.org/10.1111/2041-210X.13107>
- Vandekerckhove, J., Matzke, D., & Wagenmakers, E.-J. (2015). Model comparison and the principle of parsimony. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford handbook of computational and mathematical psychology* (pp. 300–319). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199957996.013.14>
- Vanreusel, A., Hilario, A., Ribeiro, P. A., Menot, L., & Arbizu, P. M. (2016). Threatened by mining, polymetallic nodules are required to preserve abyssal epifauna. *Scientific Reports*, 6, 26808. <https://doi.org/10.1038/srep26808>
- Vink, A. et al. (2022). MANGAN 2021 Cruise Report: independent scientific monitoring of two collector tests in the BGR and GSR contract areas for the exploration of polymetallic nodules in the equatorial NE Pacific. BGR Report, 363 pp. <https://doi.org/10.25928/hw7d-fs42>
- Volz, J. B., Haffert, L., Haeckel, M., Koschinsky, A., & Kasten, S. (2020). Impact of small-scale disturbances on geochemical conditions, biogeochemical processes and element fluxes in surface sediments of the eastern Clarion–Clipperton Zone, Pacific Ocean. *Biogeosciences*, 17(4), 1113–1131. <https://doi.org/10.5194/bg-17-1113-2020>
- Volz, J. B., Mogollón, J. M., Geibert, W., Arbizu, P. M., Koschinsky, A., & Kasten, S. (2018). Natural spatial variability of depositional conditions, biogeochemical processes and element fluxes in sediments of the eastern Clarion–Clipperton Zone, Pacific Ocean. *Deep-Sea Research Part I, Oceanographic Research Papers*, 140, 159–172. <https://doi.org/10.1016/j.dsr.2018.08.006>
- von See, T. B., Greinert, J., Meurer T., Multi-AUV sediment plume estimation using Bayesian optimization. *Front. Mar. Sci.* 11:1504099. <https://doi.org/10.3389/fmars.2024.1504099>
- von See, T. B., Meurer, T., & Greinert, J. (2022). Chemical plume tracking using an AUV with UKF based extremum seeking. *IFAC-PapersOnLine*, 55(31), 178–183. <https://doi.org/10.1016/j.ifacol.2022.10.428>
- von Stackelberg, U. (1997). Growth history of manganese nodules and crusts of the Peru Basin. *Geological Society Special Publication*, 119(1), 153–176. <https://doi.org/10.1144/gsl.sp.1997.119.01.11>
- von Stackelberg, U. (2000). Manganese nodules of the Peru Basin (Chapter 8). In Cronan, D.S. (Ed.), *Handbook of Marine Mineral Deposits*. CRC Marine Science Series, 17: pp. 197-238, CRC Press, Boca Raton, Florida, USA. ISBN 10: 084938429XISBN
- Vonnahme, T. R., Molari, M., Janssen, F., Wenzhöfer, F., Haeckel, M., Titschack, J., & Boetius, A. (2020). Effects of a deep-sea mining experiment on seafloor microbial communities and functions after 26 years. *Science Advances*, 6(18), eaaz5922. <https://doi.org/10.1126/sciadv.aaz5922>
- Wadoux, A. M. J.-C., Brus, D. J., & Heuvelink, G. B. M. (2019). Sampling design optimization for soil mapping with random forest. *Geoderma*, 355(113913), 113913. <https://doi.org/10.1016/j.geoderma.2019.113913>
- Wadoux, A.M.J.-C., Heuvelink, G. B. M., de Bruin, S., & Brus, D. J. (2021). Spatial cross-validation is not the right way to evaluate map accuracy. *Ecological Modelling*, 457, 109692. <https://doi.org/10.1016/j.ecolmodel.2021.109692>
- Wager, S., Hastie, T., & Efron, B. (2014). Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. *Journal of Machine Learning Research: JMLR*, 15(1), 1625–1651.
- Wang, M., Wu, Z., Best, J., Yang, F., Li, X., Zhao, D., & Zhou, J. (2021). Using multibeam backscatter strength to analyze the distribution of manganese nodules: A case study of seamounts in the Western Pacific Ocean. *Applied Acoustics*, 173, 107729. <https://doi.org/10.1016/j.apacoust.2020.107729>
- Wang, X., & Müller, W. E. G. (2009). Marine biominerals: perspectives and challenges for polymetallic nodules and crusts. *Trends in Biotechnology*, 27(6), 375–383. <https://doi.org/10.1016/j.tibtech.2009.03.004>
- Wasilewska-Błaszczczyk, M., & Mucha, J. (2021). Application of general linear models (GLM) to assess nodule abundance based on a photographic survey (Case Study from IOM area). *Pacific Ocean. Minerals*, 11. <https://doi.org/10.3390/min11040427>
- Weber, M. E., von Stackelberg, U., Marchig, V., Wiedicke, M., & Grupe, B. (2000). Variability of surface sediments in the Peru Basin: dependence on water depth, productivity, bottom water flow, and seafloor topography. *Marine Geology*, 163(1–4), 169–184. [https://doi.org/10.1016/s0025-3227\(99\)00103-6](https://doi.org/10.1016/s0025-3227(99)00103-6)
- Wei, C.-L., Rowe, G. T., Escobar-Briones, E., Boetius, A., Soltwedel, T., Caley, M. J., Soliman, Y., Huettmann, F., Qu, F., Yu, Z., Pitcher, C. R., Haedrich, R. L., Wicksten, M. K., Rex, M. A., Baguley, J. G., Sharma, J., Danovaro, R., MacDonald, I. R., Nunnally, C. C., ... Narayanaswamy, B. E. (2010). Global patterns and predictions of seafloor biomass using random forests. *PloS One*, 5(12), e15323. <https://doi.org/10.1371/journal.pone.0015323>

- Wenger, S. J., & Olden, J. D. (2012). Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, 3(2), 260–267. <https://doi.org/10.1111/j.2041-210X.2011.00170.x>
- Weydert, M. M. P. (1985). Measurements of the acoustic backscatter of manganese nodules. *The Journal of the Acoustical Society of America*, 78(6), 2115–2121. <https://doi.org/10.1121/1.392671>
- Weydert, M. M. P. (1990). Measurements of the acoustic backscatter of selected areas of the deep-seafloor and some implications for the assessment of manganese nodule resources. *The Journal of the Acoustical Society of America*, 88(1), 350–366. <https://doi.org/10.1121/1.399910>
- Wiedicke, M. H., & Weber, M. E. (1996). Small-scale variability of seafloor features in the northern Peru Basin: Results from acoustic survey methods. *Marine Geophysical Research*, 18(5), 507–526. <https://doi.org/10.1007/BF00310067>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80. <https://doi.org/10.2307/3001968>
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>
- Wolpert, D. H. (1992). On the connection between in-sample testing and generalization error. *Complex Systems*, 6, 47–94.
- Wolpert, D. H. (1996). The lack of A Priori distinctions between learning algorithms. *Neural Computation*, 8(7), 1341–1390. <https://doi.org/10.1162/neco.1996.8.7.1341>
- Wolpert, D.H. (2002). The Supervised Learning No-Free-Lunch Theorems. In: Roy, R., Köppen, M., Ovaska, S., Furuhashi, T., Hoffmann, F. (eds) *Soft Computing and Industry*. Springer, London. [https://doi.org/10.1007/978-1-4471-0123-9\\_3](https://doi.org/10.1007/978-1-4471-0123-9_3)
- Wong, L. J., Kalyan, B., Chitre, M., & Vishnu, H. (2021). Acoustic assessment of polymetallic nodule abundance using sidescan sonar and altimeter. *IEEE Journal of Oceanic Engineering*, 46(1), 132–142. <https://doi.org/10.1109/joe.2020.2967108>
- Wright, M. N., Ziegler, A., & König, I. R. (2016). Do little interactions get lost in dark random forests? *BMC Bioinformatics*, 17(1). <https://doi.org/10.1186/s12859-016-0995-8>
- Yamazaki, T., & Sharma, R. (2000). Morphological features of co-rich manganese deposits and their relation to seabed slopes. *Marine Georesources & Geotechnology*, 18(1), 43–76. <https://doi.org/10.1080/10641190009353782>
- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., ... Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology & Evolution*, 33(10), 790–802. <https://doi.org/10.1016/j.tree.2018.08.001>
- Yi, L., Medina-Elizalde, M., Tan, L., Kemp, D. B., Li, Y., Kletetschka, G., Xie, Q., Yao, H., He, H., Deng, C., & Ogg, J. G. (2023). Plio-Pleistocene deep-sea ventilation in the eastern Pacific and potential linkages with Northern Hemisphere glaciation. *Science Advances*, 9(8), eadd1467. <https://doi.org/10.1126/sciadv.add1467>
- Yoo, C. M., Joo, J., Lee, S. H., Ko, Y., Chi, S.-B., Kim, H. J., Seo, I., & Hyeong, K. (2018). Resource assessment of polymetallic nodules using acoustic backscatter intensity data from the Korean exploration area, northeastern equatorial pacific. *Ocean Science Journal OSJ*, 53(2), 381–394. <https://doi.org/10.1007/s12601-018-0028-9>
- Yu, G., & Parios, J. (2021). Empirical application of generalized Rayleigh distribution for mineral resource estimation of seabed polymetallic nodules. *Minerals*, 11(5), 449. <https://doi.org/10.3390/min11050449>
- Yu, T., and Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. *arXiv preprint arXiv:2003.05689*. <https://doi.org/10.48550/arXiv.2003.05689>
- Zurell, D., Elith, J., & Schröder, B. (2012). Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. *Diversity & Distributions*, 18(6), 628–634. <https://doi.org/10.1111/j.1472-4642.2012.00887.x>

## Section V

### Published & Submitted Research Articles

*"I do not mind if you think slowly. But I do object when you publish more quickly than you think."*

*Wolfgang Pauli (1900-1958)*



# Quantitative mapping and predictive modeling of Mn nodules' distribution from hydroacoustic and optical AUV data linked by random forests machine learning

Iason-Zois Gazis<sup>1</sup>, Timm Schoening<sup>1</sup>, Evangelos Alevizos<sup>1</sup>, and Jens Greinert<sup>1,2</sup>

<sup>1</sup>GEOMAR Helmholtz Centre for Ocean Research Kiel, Wischhofstrasse 1–3, 24148 Kiel, Germany

<sup>2</sup>Christian Albrechts University Kiel, Institute of Geosciences, Ludewig-Meyn-Str. 10–12, 24098 Kiel, Germany

**Correspondence:** Iason-Zois Gazis (igazis@geomar.de)

Received: 20 July 2018 – Discussion started: 3 August 2018

Revised: 26 October 2018 – Accepted: 5 November 2018 – Published: 13 December 2018

**Abstract.** In this study, high-resolution bathymetric multi-beam and optical image data, both obtained within the Belgian manganese (Mn) nodule mining license area by the autonomous underwater vehicle (AUV) *Abyss*, were combined in order to create a predictive random forests (RF) machine learning model. AUV bathymetry reveals small-scale terrain variations, allowing slope estimations and calculation of bathymetric derivatives such as slope, curvature, and ruggedness. Optical AUV imagery provides quantitative information regarding the distribution (number and median size) of Mn nodules. Within the area considered in this study, Mn nodules show a heterogeneous and spatially clustered pattern, and their number per square meter is negatively correlated with their median size. A prediction of the number of Mn nodules was achieved by combining information derived from the acoustic and optical data using a RF model. This model was tuned by examining the influence of the training set size, the number of growing trees (*n<sub>tree</sub>*), and the number of predictor variables to be randomly selected at each node (*m<sub>try</sub>*) on the RF prediction accuracy. The use of larger training data sets with higher *n<sub>tree</sub>* and *m<sub>try</sub>* values increases the accuracy. To estimate the Mn-nodule abundance, these predictions were linked to ground-truth data acquired by box coring. Linking optical and hydroacoustic data revealed a nonlinear relationship between the Mn-nodule distribution and topographic characteristics. This highlights the importance of a detailed terrain reconstruction for a predictive modeling of Mn-nodule abundance. In addition, this study underlines the necessity of a sufficient spatial distribution of the optical data to provide reliable modeling input for the RF.

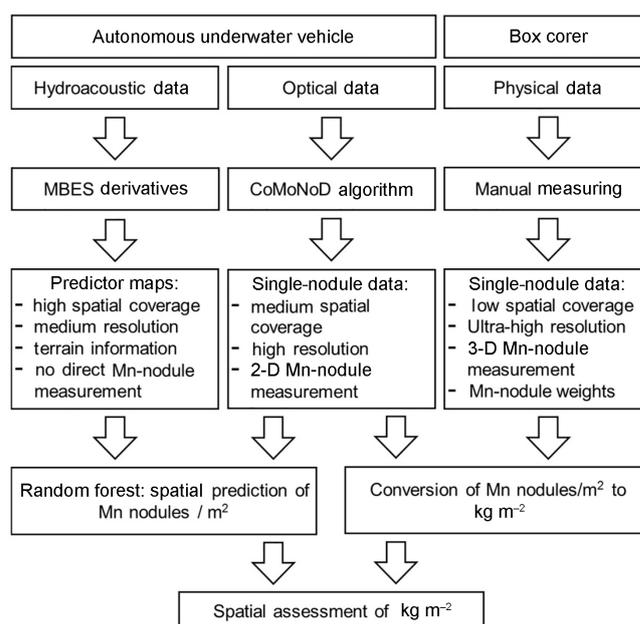
## 1 Introduction

High-resolution quantitative predictive mapping of the distribution and abundance of manganese nodules (Mn nodules) is of interest for both the deep-sea mining industry and scientific fields such as marine geology, geochemistry, and ecology. The distribution and abundance of Mn nodules are affected by several factors such as local bathymetry (Craig, 1979; Kodagali, 1988; Kodagali and Sudhakarand, 1993; Sharma and Kodagali, 1993), sedimentation rate (Glasby, 1976; Frazer and Fisk, 1981; von Stackelberg and Beiersdorf, 1991; Skornyakova and Murdmaa, 1992), availability of nucleus material (Glasby, 1973), and bottom current strength (Frazer and Fisk, 1981; Skornyakova and Murdmaa, 1992). As a consequence, the distribution and abundance of Mn nodules is heterogeneous (Craig, 1979; Frazer and Fisk, 1981; Kodagali, 1988; Kodagali and Sudhakar, 1993; Kodagali and Chakraborty, 1999; Kuhn et al., 2011), even on fine scales of 10 to 1000 m (Peukert et al., 2018a; Alevizos et al., 2018). This increases the difficulty of quantitative predictive mapping using remote-sensing methods. Vast areas of the seafloor can be mapped by ship-mounted, multibeam echosounder systems (MBESs). State-of-the-art MBESs feature a low frequency (12 kHz) and can map ca. 300 km<sup>2</sup> of seafloor in 4500 m water depth per hour. Hence, low-resolution regional maps can be created at a grid cell size of 50 to 100 m within which the main Mn-nodule occurrence can become apparent, based on the backscatter intensity (Kuhn et al., 2011; Rühlemann et al., 2011; Jung et al., 2001). A general separation in areas of high and low abundance (kg m<sup>-2</sup>) of Mn nodules seems possible, especially in flat areas where

sedimentological changes and physical influences on the footprint size and incidence angle of the transmitted acoustic pressure wave can be corrected accurately (De Moustier, 1986; Kodagali and Chakraborty, 1999; Chakraborty and Kodagali, 2004; Kuhn et al., 2010, 2011; Rühlemann et al., 2011, 2013). However, the patchy distribution of Mn nodules in size and number at meter-scale cannot be resolved with ship-mounted MBES data (Petersen, 2017). For an operational resource assessment, a higher resolution of a few meters grid cell size is needed to supply accurate depth, slope, and Mn-nodule distribution variability (Kuhn et al., 2011). Supplementary to the spatial mapping by acoustic sensors, point-based measurements from box corer samples are used as ground-truth data for training and validation of geostatistical techniques (e.g., kriging) in order to create quantitative maps of Mn-nodule abundance (Mucha et al., 2013; Rahn, 2017). However, the generally low number of ground-truth samples during surveys (usually below 10), their limited sampling area (typically  $0.25 \text{ m}^2$ ), and the relatively large distance between them ( $> 1 \text{ m}$ ) prevent an accurate correlation with the ship-based MBES data and thus a good prediction of the total Mn nodules' mass and distribution in large areas (Petersen, 2017). Importantly, the sparse sampling with box corers affects the performance of interpolation and geostatistical techniques, which are typically applied during data analysis (Li and Heap, 2011, 2014; Kuhn et al., 2016). In this article, we address this challenge by combining high-resolution hydroacoustic and optical data sets acquired with an autonomous underwater vehicle (AUV) and connecting those data with a machine learning (ML) algorithm (here random forests), in order to predict the spatial distribution of the number of Mn nodules per square meter. Unlike geostatistical methods, ML can be used to incorporate information from different bathymetric derivative layers and to detect complex relationships among predictor variables without making any prior assumptions about the type of their relationship or value distribution (Garzn et al., 2006; Lary et al., 2016). To this end, first predictions have already been achieved (Knobloch et al., 2017; Vishnu et al., 2017; Alevizos et al., 2018). Here, we present a complete data analysis workflow for potential mining exploration (Fig. 1).

### 1.1 AUV hydroacoustic mapping

AUVs have proven their usefulness for multibeam data acquisition in the deep-sea environment (Grasmueck et al., 2006; Deschamps et al., 2007; Haase et al., 2009; Wynn et al., 2014; Clague et al., 2014, 2018; Pierdomenico et al., 2015; Peukert et al., 2018a). They achieve higher spatial and vertical resolution compared to ship-mounted MBESs. This is due to their operation close to the seafloor, which results in a smaller footprint at a given beam angle and enables the use of higher frequencies (Henthorn et al., 2006; Mayer, 2006; Caress et al., 2008; Paduan et al., 2009). Additionally, AUVs avoid problems like near-surface turbulences, bubbles,



**Figure 1.** Schematic workflow of the data sets used in this study to enable the spatial assessment of Mn nodules inside the study area. The medium resolution of AUV MBES (meter scale) refers to the comparison of the optical and physical data (centimeter scale).

ship noise, and strong sound velocity changes (Kleinrock et al., 1992a, b; Jakobson et al., 2016; Paul et al., 2016). They work independently from the surface vessel and operate at a stable altitude. AUVs can efficiently conduct a dive pattern of dense survey lines and thus reduce survey effort and costs (Chance et al., 2000; Bellingham, 2001; Bingham et al., 2002; Danson, 2003; Roman and Mather, 2010). High-resolution bathymetry enables computing bathymetric derivatives like slope and rugosity with a similarly high resolution. These derivatives play an important role in predicting Mn nodules' distribution and abundance (Craig, 1979; Kodagali, 1988; Skornyakova and Murdmaa, 1992; Kodagali and Sudhakar, 1993; Sharma and Kodagali, 1993; Ko et al., 2006). However, a small number of recent studies have investigated this role on an AUV scale (Okazaki and Tsune, 2013; Peukert et al., 2018a; Alevizos et al., 2018).

### 1.2 Underwater optical data

Underwater optical data have generally played an important role in the qualitative analysis of the seafloor features and for the specific task of assessing Mn nodules' distribution explicitly (Glasby, 1973; Rogers, 1987; Skornyakova and Murdmaa, 1992; Sharma et al., 1993). The development of automated detection algorithms enabled quantitative optical image data analysis and subsequent statistical interpretation of Mn-nodule densities. The spatial coverage of optical imaging is much higher than for box core sampling. The data resolution remains high enough to reveal the high variance

in the spatial distribution of nodules at meter scale. Thus optical data can fill the investigation gap between ground-truth sampling and hydroacoustic remote sensing (Sharma et al., 2010, 2013; Schoening et al., 2012a, 2014, 2015, 2016, 2017a; Kuhn and Rathke, 2017). Moreover, mosaicking of optical data could reveal mining obstacles such as outcropping basements or volcanic pillow lava flows. In addition, seafloor photos are the source for evaluating benthic fauna occurrences and related habitats on a wider area (Schoening et al., 2012b; Durden et al., 2016).

### 1.3 Box corer sampling

Box coring is common to obtain physical samples of Mn nodules and sediments for resource assessments and biological studies. While optical data reveal only the exposed and semi-buried Mn nodules, box corers collect the top 30–50 cm of the seafloor with minimum disturbance, allowing an accurate measure of the Mn nodules' abundance ( $\text{kg m}^{-2}$ ). Box coring data are used for training and validation in geostatistical methods for quantitative and spatial predictions of Mn nodules (e.g., Mucha et al., 2013; Knobloch et al., 2017). The representativeness of box coring data is disputable as few deployments can be made due to time constraints (ca. 4 h per core) and as the spatial coverage of one sample is rather low (ca.  $0.25 \text{ m}^2$ ).

### 1.4 Random forests

Random forests (RF) is an ensemble machine learning (ML) method composed of multiple weaker learners, namely classification or regression trees (Breiman, 2001a). Within RF an ensemble of distinct tree models is trained using a random subsample of the training data for each tree until a maximum tree size is reached. In each tree, each node is split using the best among a subset of predictors randomly chosen at that node instead of using the best split among all variables (Liaw and Wiener, 2002). Thus, the process is double-randomized which further reduces the correlation between trees. About two-thirds of the training data are used to tune the RF while the remaining “out-of-bag” (OOB) samples are used for an internal validation. By aggregating the predictions of all trees (majority votes for classification, the average for regression), new values can be predicted. This aggregation keeps the bias low while it reduces the variance, resulting in a more powerful and accurate model. RFs have the ability to estimate the importance of each predictor variable, which enables data mining of the high-dimensional prediction data. Terrestrial studies use RFs in prospectivity mapping of mineral deposits (Carranza and Laborte, 2015a, b; 2016; Rodriguez-Galiano et al., 2014, 2015). In the marine environment, RFs have been used to combine MBES bathymetry, backscatter, their derivatives, sediment sampling, and optical data for various seabed classification and regression tasks (e.g., Li et al., 2010, 2011a; Che Hasan et al., 2014; Huang et al., 2014).

Further studies showed the robustness of RFs for selected data sets compared to other ML algorithms (Che Hasan et al., 2012; Stephens and Diesing, 2014; Diesing and Stephens, 2015; Herkul et al. 2017), as well as to geostatistical and deterministic interpolation methods (Li et al., 2010, 2011a, b; Diesing et al., 2014).

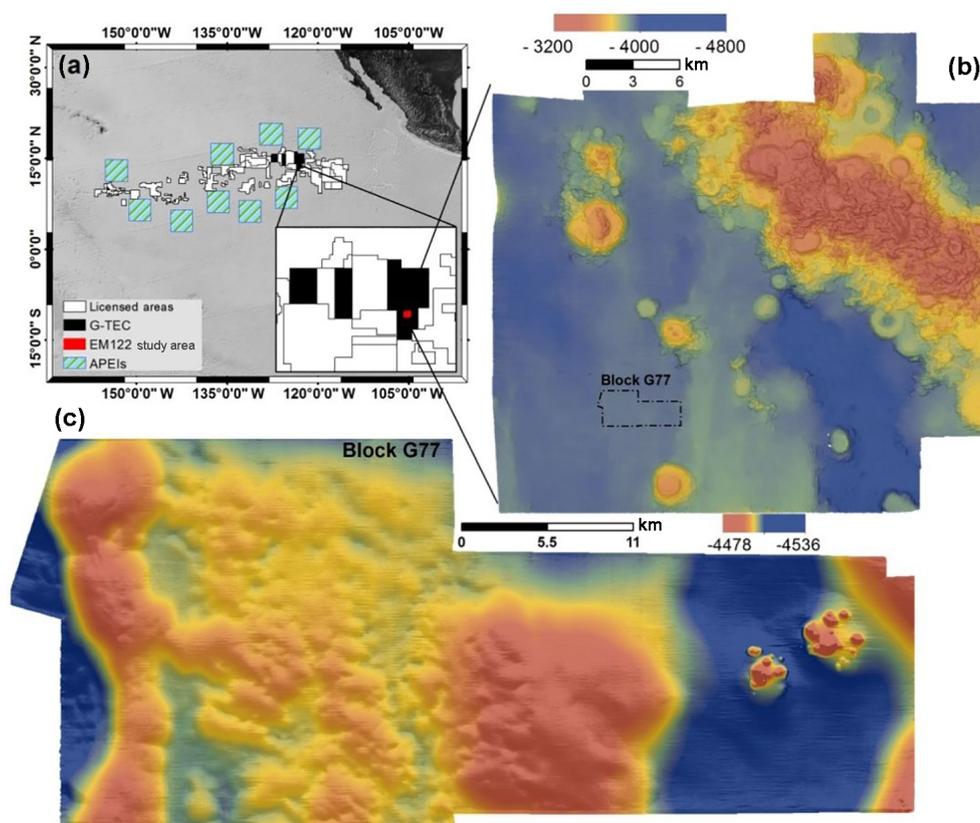
## 2 Study area

The study area lies in the Clarion–Clipperton Zone (CCZ; ca.  $4 \times 10^6 \text{ km}^2$ ) in the eastern central Pacific Ocean. The CCZ has triggered scientific and industrial interest for several decades due to its high resource potential in Mn-nodule deposits (Hein et al., 2013; Petersen et al., 2016) with an average nodule abundance of  $15 \text{ kg m}^{-2}$  (SPC, 2013). At the time of writing, the International Seabed Authority (ISA) has granted 17 exploration licences inside the CCZ (Fig. 2a). The study area described here is part of the Belgian GSR (Global Sea Mineral Resources) license area (Fig. 2b) and will be referred to as block G77 (Fig. 2c). Overall, this part of the Belgian license area has a high bathymetric range and complex morphology, due to the presence of submarine volcanoes, solitary seamounts, and seamount chains. Block G77 is characterized by a low bathymetric range (77 m) and mostly gentle slopes (95 % of the area below  $5^\circ$ ). An exception is located in the eastern part, where subrecent small-scale volcanic activity created 15 cone-shaped morphological features of up to 55 m height and 150 m width that are clustered in an area of ca.  $700 \text{ m} \times 380 \text{ m}$ . Despite the gentle slopes, block G77 is characterized by an uneven microrelief (according to Dikau, 1990) especially in the western part, where small (2–4 m) depressions coexist next to short (2–4 m) protrusions. In the central part, a 30 m high elevation acts as a natural barrier between the western part of the study area that features more relief and the eastern part that is deeper and has less relief (Fig. 2c).

## 3 Methodology

### 3.1 Hydroacoustic data acquisition and post-processing

The data (Greinert, 2016) were collected in March 2015 during cruise SO239 EcoResponse (Martínez Arbizu and Haeckel, 2015) with the German research vessel *Sonne*. Ship-based mapping was conducted with a hull-mounted Kongsberg EM 122 MBES (12 kHz,  $0.5^\circ$  along- and  $1^\circ$  across-track beam angle, 432 beams with  $120^\circ$  swath angle). High-resolution MBES data were acquired with AUV *Abyss* (GEOMAR, 2016) inside block G77 equipped with a Teledyne Reson Seabat 7125 MBES (200 kHz,  $2^\circ$  along- and  $1^\circ$  across-track beam angle, 256 beams with  $130^\circ$  swath angle). The data (60 km of survey lines) were acquired from 50 m altitude and with 100 % swath overlap resulting in an insonification of  $9.5 \text{ km}^2$ . Post-processing of the AUV data



**Figure 2.** (a) Areas of Particular Environmental Interest (APEIs), licensed areas (white), and the Belgium/GSR licenses area (black) within the CCZ. (b) Regional bathymetric map of the study area, created by the EM 122 MBES on R/V *Sonne* (cruise SO239). (c) Block G77, mapped by AUV *Abyss* with a Teledyne Reson Seabat 7125 MBES.

was conducted with the Teledyne PDS2000 software for data conversion of the raw data into s7k and GSF format. Further multibeam processing (sound velocity calibration, pitch/roll/yaw/latency artifacts correction) was performed using the Qimera™ software. The largest uncertainties during AUV operations result from inaccurate navigation and localization in the deep-sea environment (Paull et al., 2014). AUV *Abyss* has a combination of five different systems for navigation and positioning: Global Positioning System (GPS) when at the sea surface, Doppler velocity log (DVL) when 100 m or less from the ground, inertial navigation system (INS), long baseline acoustic navigation (LBL), and dead reckoning (GEOMAR, 2016). Each system has its own limitations that contribute to the total navigation error (Sibenac et al., 2004; Chen et al., 2013) that generally results in positioning drifts over time. Consequently, this affects the position accuracy of the MBES and optical data. Our AUV MBES data processing and an absolute geo-referencing of the resulting AUV bathymetry grid with the EM122 ship data, supplemented with the use of MBnavadjust in MB-Systems (Caress et al., 2017), resulted in a well-calibrated AUV bathymetric data set. The position of the AUV image data “only” relies on the abovementioned sensors with a “small” position error

that is not quantifiable. Backscatter data were excluded from the modeling procedure due to artifacts and a generally poor quality. The output grid cell size for the analyses was set to  $3\text{ m} \times 3\text{ m}$ . The depth raster was exported as ASCII format for further analysis in SAGA GIS v.6.3.0. SAGA includes numerous tools that focus on terrain analysis (Conrad, 2015). Eight bathymetric derivatives were computed (Table 1) with the SAGA algorithms (Appendix A).

### 3.2 Optical data acquisition and post-processing

High-resolution optical data (20.2 megapixels) were acquired by the DeepSurveyCamera system on board AUV *Abyss* (Kwasnitschka et al., 2016). During image acquisition, the altitude above ground was 5 to 11 m, resulting in an overlap between the images of ca. 60% in each direction. In total, 11 276 photos were acquired in block G77 (Greinert, 2017) and analyzed with the automated image analysis algorithm CoMoNoD (Schoening et al., 2017a, b, c). For each image this algorithm delineates each individual Mn nodule and provides quantitative information on each nodule (size in  $\text{cm}^2$ , alignment of main axis, geographical coordinate of the nodule). This information is further aggregated per image to provide the average number of Mn nodules per square me-

**Table 1.** The bathymetric derivatives computed in SAGA GIS and used as predictor variables.

Derivative	Description
Slope (S)	The first derivative of the bathymetry; describes the steepness of a surface
Plan curvature (Pl.C)	The second derivative of the bathymetry; perpendicular to the direction of the maximum slope (Zevenbergen and Thorne, 1987)
Profile curvature (Pr.C)	The second derivative of the bathymetry; parallel to the direction of the maximum slope (Zevenbergen and Thorne, 1987)
Topographic position index (TPI)	Compares the elevation of a single pixel to the average of multiple cells surrounding it at a defined distance (Weiss, 2001).
Broad-scale (TPI_B)	Distance: 150–400 m
Medium-scale (TPI_M)	Distance: 50–150 m
Fine-scale (TPI_F)	Distance: 0–50 m
Concavity (C)	In each cell its value is defined as the percentage of concave downward cells within a constant radius (Iwahashi and Pike, 2007). Here, a 10-cell radius was used.
Terrain ruggedness index (TRI)	A quantitative measure of surface heterogeneity; can be explained as the sum change in elevation between a central pixel and its neighborhood (Riley et al., 1999). Here, a 10-cell radius was used.

ter (Mn nodules  $\text{m}^{-2}$ ), the nodule coverage of the seafloor in percent, and the nodule size distribution in square-centimeter size quantiles. The algorithm has successfully been applied for quantitative assessment and predictive modeling of Mn nodules (Peukert et al., 2018a; Alevizos et al., 2018). Nevertheless, the derived number of Mn nodules  $\text{m}^{-2}$  is subject to uncertainties due to the limitations of the CoMoNoD algorithm and the nonconstant altitude of the AUV, especially in areas with slopes. The CoMoNoD algorithm cannot detect sediment-covered Mn nodules due to the low or non-existent contrast. It may count two or more adjacent small Mn nodules as one big nodule or misinterpret benthic fauna or rock fragments with similar visual features as Mn nodules. The CoMoNoD algorithm fits an ellipsoid around each detected Mn nodule, which limits the first two disadvantages as it splits huge Mn nodules and accounts for potentially buried parts (see discussions in Schoening et al., 2017a). In general, the first two disadvantages lead to underestimations while the third one results in an overestimation of the number of Mn nodules per square meter. These limitations are common, and the need for corrections (e.g., a factor that describes the ratio between the number of Mn nodules seen in the photo and the number of nodules counted in box corers, considering for the different spatial scales) has been acknowledged (Sharma and Kodagali, 1993; Sharma et al., 2010, 2013; Tsune and Okazaki, 2014; Kuhn and Rathke, 2017). Recent studies show that the difference between image estimates and the abundance in box corer data (due to sediment covered Mn nodules) can be 2–4 times higher (Kuhn and Rathke, 2017). In this study, none of the box corers was obtained exactly at a location for which optical data exist; thus, no direct com-

parison and verification exist. Taking box corer samples for verification requires ultrashort baseline (USBL) navigation and imaging of the seafloor prior to the physical sampling. The effects of the nonconstant flying altitude on the detection of Mn nodules per square meter are explained in detail below. For each photo location, the depth and the bathymetric derivative values were extracted from the hydroacoustic data. As no absolute geo-referencing could be performed for the AUV-based photo surveys, drifting sensor data will have an effect on the alignment between bathymetric and photo information, which was considered while interpreting the results.

### 3.3 Data exploration and spatial analysis

The data exploration, spatial plotting, and analysis was performed with ArcMap™ 10.1, PAST v3.19 (Hammer et al., 2001), and R (R Development Core Team, 2008). All data were projected as a UTM Zone 10N coordinate system (to enable spatial analysis). The existence of spatial autocorrelation in the distribution of Mn nodules  $\text{m}^{-2}$  was examined by the global Moran's index (GMI) and Anselin local Moran's index (LMI). Both GMI (Moran, 1948, 1950) and LMI (Anselin, 1995) are well-established for examining the overall (global) and local spatial autocorrelation, respectively (e.g., Goodchild, 1986; Fu et al., 2014). GMI attains values between  $-1$  and  $1$  with high positive values indicating strong spatial autocorrelation. High positive LMI index values indicate a local cluster. This cluster could be a group of observations with high–high (H–H) or low–low (L–L) values regarding the examined variable. A high negative index value im-

plies local outliers, like high–low (H–L) or low–high (L–H) clusters, in which an observation has a higher or lower value in comparison to its adjacent observations. Both Moran's index analyses were performed in ArcMap™ 10.1 (for parameter settings see Appendix A). One decimal was retained in the presentation of the results from statistical analysis and RF modeling.

### 3.4 Box corer data

A total of five box corers (0.5 m × 0.5 m surface area) were obtained close to the study area (coordinates not given due to confidentiality). However, one is located within block G77 (Fig. 3a); this is the result of independent sampling schemes and purposes during the cruise. Nevertheless, all box core samples (maximum distance < 1.5 km) were analyzed and used for further analyses. In the three box corers, the number, size, and weight of nodules were measured and the abundance (kg m<sup>-2</sup>) was estimated (mean value: 26.5 kg m<sup>-2</sup>). The total number of Mn nodules within each box corer was compared with the number of Mn nodules on the surface resulting in an average ratio of 1.32 (Table 2). This means that ≈ 25 % of the nodules are not seen on the surface but are completely buried within the sediment (down to a depth of about 15 cm).

### 3.5 RF predictive modeling

The RF modeling was performed with the Marine Geospatial Ecology Tools (MGET) toolbox in ArcMap™ 10.1. MGET (Roberts et al., 2010) uses the randomForests R package for classification and regression (Liaw and Wiener, 2002). Our target variable (number of Mn nodules m<sup>-2</sup>) is continuous, so regression was applied. We followed the three main steps to establish a good model by selecting predictor variables, and calibration/training of the model and finally validating the model results.

*Selection of predictor variables.* The depth (D) and its derivatives (Table 1) were used as predictor variables. Although RFs can handle a high number of predictor variables with similar information, the exclusion of highly correlated variables can improve the RF performance and decrease computation time (Che Hasan, 2014; Li et al., 2016). Thus, the correlation between derivatives was investigated using the Spearman's rank correlation coefficient. None of the variable pairs was perfectly correlated ( $\rho \geq 95$ ), and consequently, all of them were used for RF modeling (Appendix A).

*Calibration of the model.* During the calibration process, the RF parameters were adjusted as follows. The number of predictor variables to be randomly selected at each node (*mtry*), the minimum size of the terminal nodes (*nodesize*), and the number of trees to grow (*ntree*) were set to the default values, in order to investigate the optimum training size. For regression RF the default *mtry* value is 1/3 of the number of

predictor variables (rounded down), *nodesize* is 5 and *ntree* is 500 (Liaw and Wiener, 2002). RF has been demonstrated to be robust regarding these parameters, and the default values have given trustworthy results (e.g., Liaw and Wiener, 2002; Diaz-Uriarte and de Andres, 2006; Cutler et al., 2007; Okun and Priisalu, 2007; Li et al., 2016, 2017). With regards to the subsampling method (*replace*), the subsampling without replacement was selected. Although the initial implementations of the RF algorithm use subsampling with replacement (Breiman, 2001a), later studies showed that this process might cause a biased selection of predictor variables that vary in their scale and/or in their number of categories, resulting in a biased variable importance measurement (Strobl et al., 2007, 2009; Mitchell, 2011). Based on recent findings, the raw variable importance was preferred (*unscaled*) as the final parameter (Diaz-Uriarte and de Andres, 2006; Strobl et al., 2008, 2009; Strobl and Zeileis, 2008). Using these settings, the influence of the training sample size was examined (10 % to 90 % of the total sample in steps of 10 %) and compared based on the mean of squared residuals (MSR) using the respective equation provided in the randomForests R package (Liaw and Wiener, 2002). The different training groups need to be considered as representative of the total sample, in order to capture the heterogeneity of the Mn nodules' spatial distribution. The spatially random selection of subsamples by MGET ensured similar statistical characteristics in each group (Appendix A). For each case of different training sample size, the model was run 10 times and the results are presented as the average value of these 10 runs (Appendix B). Since the optimal training sample size was defined, the influence of the number of growing trees (*ntree*) and the influence of the number of predictor variables to be randomly selected at each node (*mtry*) was examined. Only for the already defined optimum training size were 10 different *ntree* values (100 to 1000 in steps of 100) and seven different *mtry* values (1 to 7 in steps of 1) tested and compared based on the MSR values. In each case of a different *ntree* and *mtry* parameter, the model was run 10 times, and the results are presented as the average value of these 10 runs (Appendix B).

*Selection and external validation of the optimal model.* Based on the abovementioned results and considering the sampling and computational cost, the optimal model was selected, run for 30 iterations, and applied to the entire study area. Its predicted values were validated with the observed values from the remaining data set that was not used. Several validation measures were used including the mean absolute error (MAE), the mean squared error (MSE), and the root mean squared error (RMSE). The combined use of MAE and RMSE is a well-established procedure as the MAE can evaluate better the overall performance of a model (all individual differences have equal weight), while the RMSE gives disproportionate weight to large errors showing an increased sensitivity to the presence of outliers. Due to this characteristic, RMSE is suitable for outlier detection analysis but

**Table 2.** The number of Mn nodules on the sediment surface, the total number of Mn nodules per box core, the ratio of those two values, and the distance of the box corer deployments from the study area in block G77.

Box corer station	Total number of Mn nodules	Number of Mn nodules at the surface	Ratio	Abundance (kg m <sup>-2</sup> )	Distance from G77 area (km)
BC20	40	27	1.5	–	0
BC21	67	58	1.1	27.1	1.4
BC22	29	21	1.4	27.1	0.6
BC23	32	20	1.6	25.2	0.1
BC24	17	16	1.0	–	1
Average	37	28	1.32	26.5	0.6

should not be used solely as an index for the model performance (Willmott and Matsuura, 2005). Both MAE and RMSE are measured in the same unit as the data. In addition, the  $R^2$ , Pearson ( $r$ ), and Spearman's rank correlation coefficients were used to identify the correlation between predicted and initial values. Finally, the descriptive statistics of predicted and initial values were compared and a residual analysis was performed.

### 3.6 Resource assessment

As the optimal RF model was applied to the entire block G77, an estimate of the abundance (kg m<sup>-2</sup>) was computed, based on the analogy between the corresponding abundance measured from the average number of Mn nodules in the box corer data and the number of Mn nodules m<sup>-2</sup> in each cell of the final result of the RF model. Considering that the collector can recover buried Mn nodules from a maximum depth of 10–15 cm (Sharma, 1993, 2010), the ratio of 1.32 was applied to account for Mn nodules not detected in the images, and areas with a slope of > 3° were excluded, assuming that a potential mining vehicle is limited to less steep slopes (UNOET, 1987).

## 4 Results

### 4.1 Data exploration

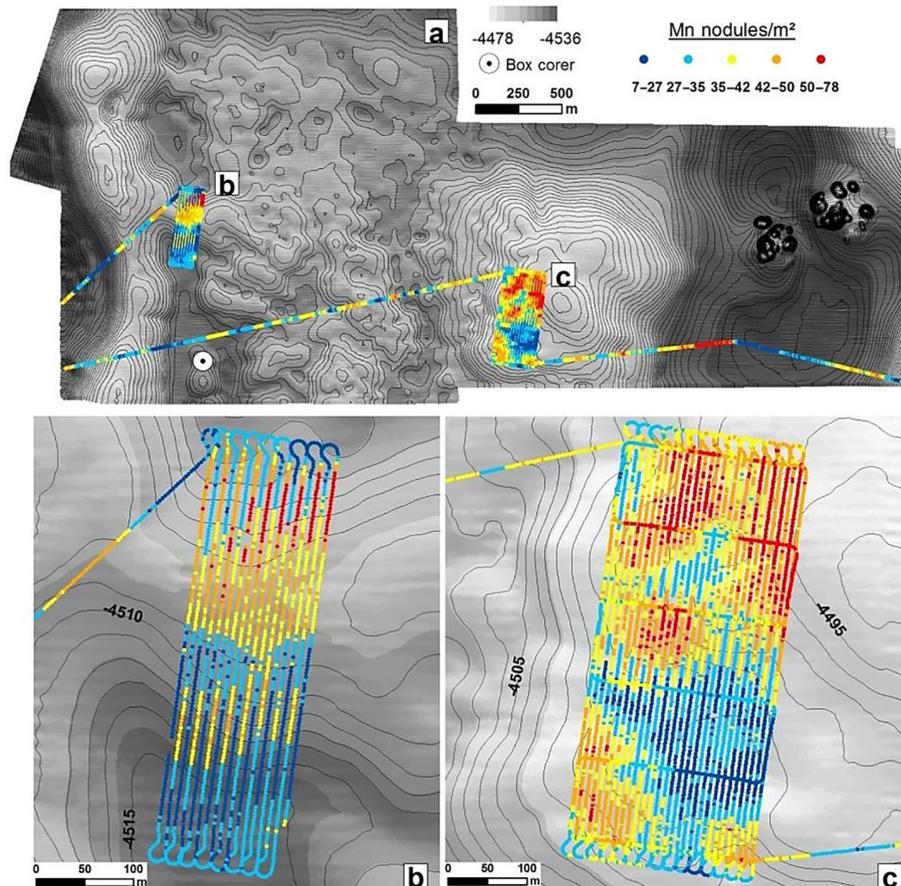
The analysis of AUV photos with the CoMoNoD algorithm (Schoening et al., 2017a) revealed a rather heterogeneous pattern of Mn nodules m<sup>-2</sup> in the study area, showing adjacent areas with high and low Mn-nodule number (Fig. 3a). The number of Mn nodules m<sup>-2</sup> changes within less than 100 m in the overall study area and in the two main subareas b and c (Fig. 3a–c). In half of the photos (48 %), the number of Mn nodules m<sup>-2</sup> varies from 30 to 43 with the mean value being 36.6 Mn nodules m<sup>-2</sup>. The very small change of 5 % trimmed mean value indicates the absence of extreme outliers, which is confirmed by box plot analysis (Appendix B). Further analysis of their descriptive and distribution characteristics was performed in order to assess the presence of normality in the data, with the result that the number of

Mn nodules m<sup>-2</sup> is approximately normally distributed (Appendix B). Although the presence of normality in data is not a prerequisite assumption in order to perform the RF (Breiman, 2001a), as it is with geostatistical interpolation techniques like kriging (e.g., Kuhn et al., 2016), this examination can give us a better understanding of the Mn nodules' distribution inside the study area, and it is an important step in order to examine potential extreme observations which may be derived from wrong measurements and could artificially change the training range during RF predictive modeling. Moreover, an absence of linear correlation was observed between Mn nodules m<sup>-2</sup> and the produced bathymetric derivatives, indicating the complexity of the phenomenon (Appendix B).

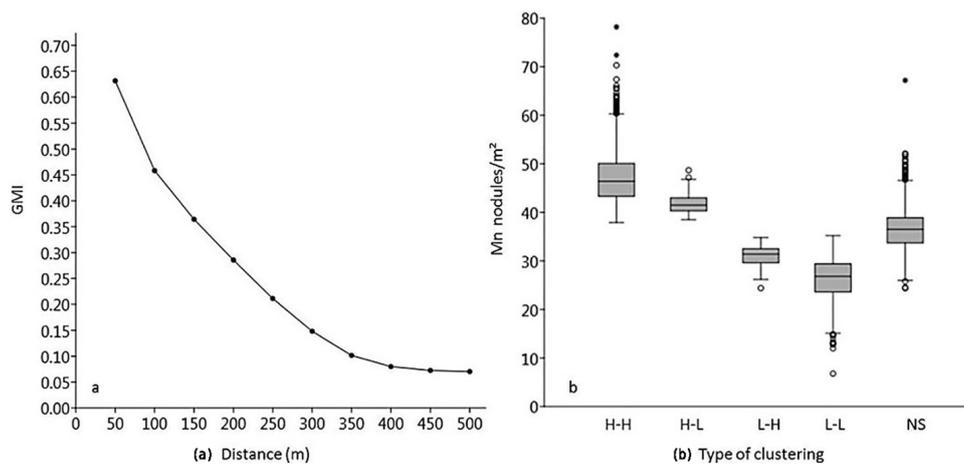
### 4.2 Spatial analyses

Spatial analyses revealed the presence of a spatial autocorrelation in the distribution of Mn nodules m<sup>-2</sup>. The GMI, with  $I = 0.6989$ ,  $p < 0.01$ , and a  $Z$  score > 2.58 indicates a positive spatial autocorrelation. According to the incremental analysis, the index takes its highest value in the first 50 m with a gradual decrease, approaching 0 values after a distance of 400 m (Fig. 4a). Similarly, the results from the LMI show that the main size of the spatial clusters does not exceed 400 m in either direction (Fig. 5a). The main types of these clusters are H–H and L–L groups (Fig. 4b and Table 3). A distinct “buffer/transitional zone” with Mn nodules was found between these two clusters, which does not show a significant autocorrelation (Fig. 5b, c). Approximately one-third of the data does not have a significant clustering (NS). In the subarea c, the few local H–L and L–H groups are located in the outer parts of these zones without significant spatial clustering. Both H–L and L–H (from the entire study area) only account for 2.1 % of the data (Table 3). The comparison of the number of Mn nodules m<sup>-2</sup> between the groups shows a clear discrimination between H–H and L–L clusters (Fig. 5b). The H–H clusters are in areas with 37.9–78.2 Mn nodules m<sup>-2</sup> whilst the L–L clusters are in areas with 6.8–35.2 Mn nodules m<sup>-2</sup>.

The application of the LMI reveals a bias that exists in the data due to the sampling procedure, especially in the subarea b (Fig. 5b). Here, the presence of the slope around 2.8° forced



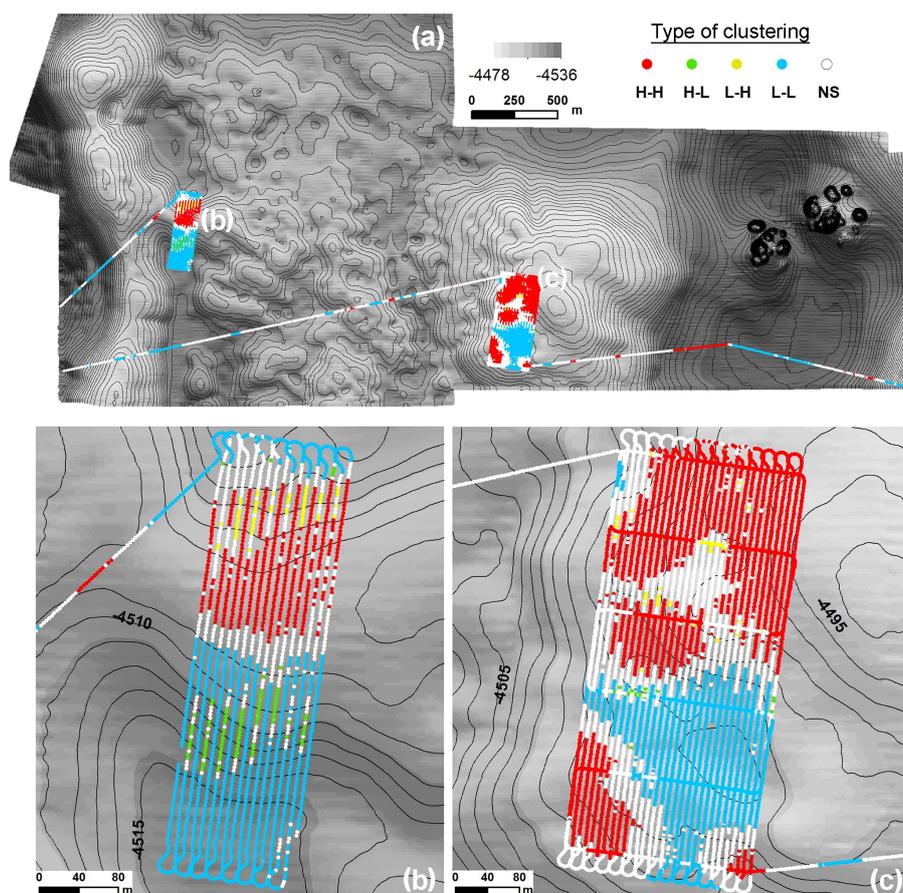
**Figure 3.** (a) The spatial distribution of Mn nodules  $\text{m}^{-2}$  inside block G77 and the box corer position. (b) The spatial distribution of Mn nodules  $\text{m}^{-2}$  inside the subarea b. (c) The spatial distribution of Mn nodules  $\text{m}^{-2}$  inside the subarea c.



**Figure 4.** (a) The GMI decrement due to increasing distance, after the first 50 m. (b) The range of Mn nodules  $\text{m}^{-2}$  in each clustered group.

the AUV to vary its altitude between the ascending and descending phase (Fig. 6b). This variation seems to affect the image quality, resulting in fewer nodules being counted for higher altitudes of the AUV (Figs. 7 and 8). This is also

confirmed by the distribution map of the Mn nodules  $\text{m}^{-2}$  (Fig. 3b). It is important to emphasize that this difference shows up clearly in the LMI results (Fig. 5b) and not in the distribution map (Fig. 3b); here the arbitrary choice of color



**Figure 5.** (a) The spatial distribution of the significant cluster types inside the block G77. (b) The spatial clusters inside the subarea b. (c) The spatial clusters inside the subarea (c).

**Table 3.** Number and percentage of samples in each type of spatial clustering.

Cluster type	H-H	H-L	L-H	L-L	NS
Counts ( <i>n</i> )	3472	121	113	3523	4047
Counts (%)	30.8	1.1	1.0	31.2	35.9

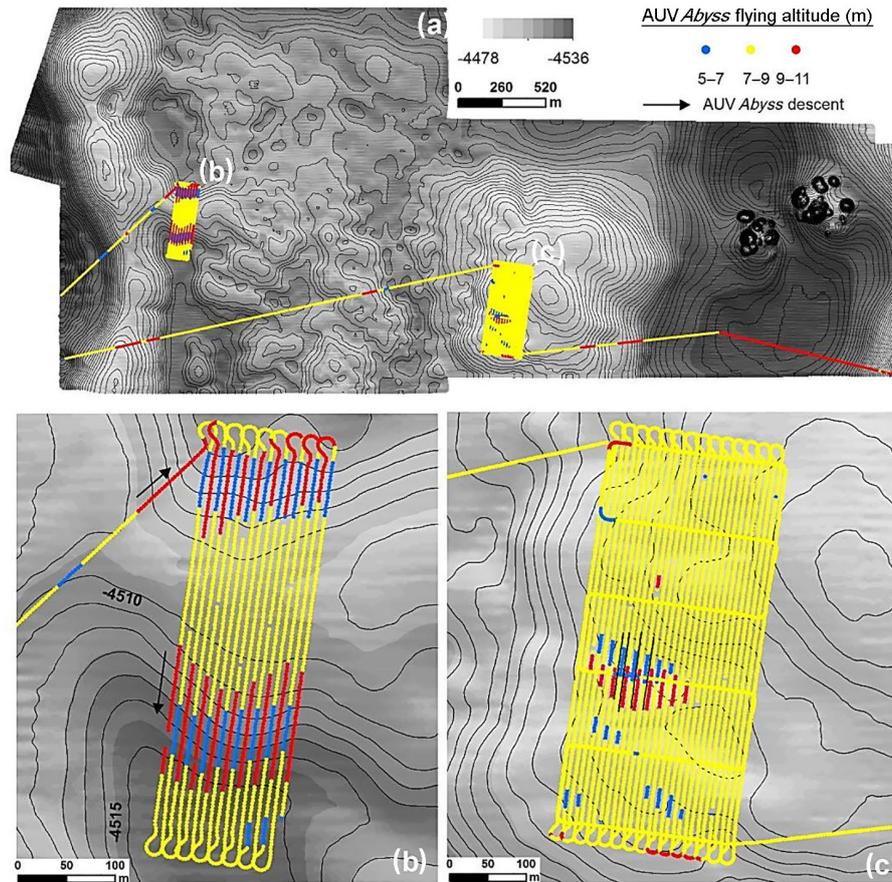
scale can hide this bias during plotting. The comparison of the detected Mn nodules  $\text{m}^{-2}$  in these adjacent lines, inside the small subarea b, gives a ratio  $\approx 1.4$  between photos that have been acquired at 7–9 m altitude and those at 9–11 m altitude. The ratio is higher ( $\approx 1.8$ ) between photos from 5–7 and 9–11 m altitude. In contrast, the ratio between photos from 5–7 m altitude and those at 7–9 m altitude is  $\approx 1.25$ , indicating that the problem mainly exists at upper and lower flying altitudes. Despite their different ratio, none of these groups contain extremely high or low values of Mn nodules  $\text{m}^{-2}$ . Moreover, in several parts of the block, the photos from higher altitude are the only source of information without the ability for further comparison, and consequently, they cannot be excluded from the modeling procedure.

*Spatial distribution of median size.* Plotting of the median size in square centimeters (Fig. 9) showed that the number of Mn nodules  $\text{m}^{-2}$  is anti-correlated to the median Mn-nodule size. The Spearman's rank correlation coefficient and  $R^2$  between these two variables are  $-0.50$  and  $0.25$ , respectively, supporting this observation (Fig. 10a); other studies found similar results (Okazaki and Tsune, 2013; Kuhn and Rathke, 2017; Peukert et al., 2018a). The box plot analysis of the median size values between the H–H and L–L clustered groups showed that although the L–L group contains the entire range of median size values ( $2.8$  to  $15.9 \text{ cm}^2$ ), the H–H group does not contain values above  $10 \text{ cm}^2$  ( $2.7$ – $10 \text{ cm}^2$ ). This means in consequence that in areas with significant clustering of higher numbers of Mn nodules  $\text{m}^2$ , the size of Mn nodules tends to be smaller (Fig. 10b).

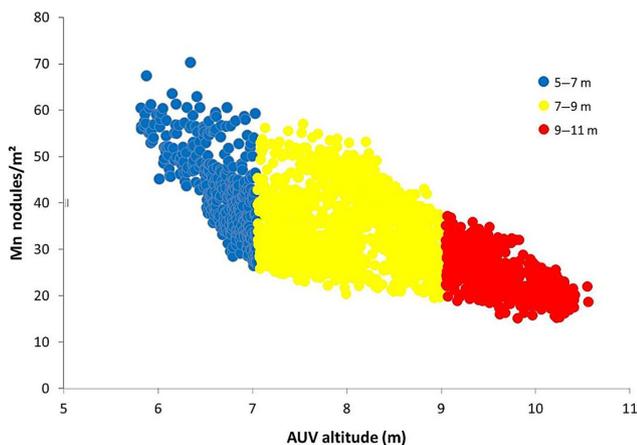
### 4.3 RF predictive modeling

#### 4.3.1 Effect of training sample size and *n*tree and *m*try parameters

The results of the modeling procedure demonstrate that the RF algorithm is influenced by the size of the training sample



**Figure 6.** (a) The altitude of AUV *Abyss* inside block G77. (b) The altitude inside the small subarea b, where the presence of the slope forces the AUV to modify its altitude, flying closer to the seafloor in the ascending phase (blue lines) and farther from the seafloor in the descending phase (red lines). (c) In the big subarea c, the AUV flying altitude is mainly constant between 7–9 m for the entire part.



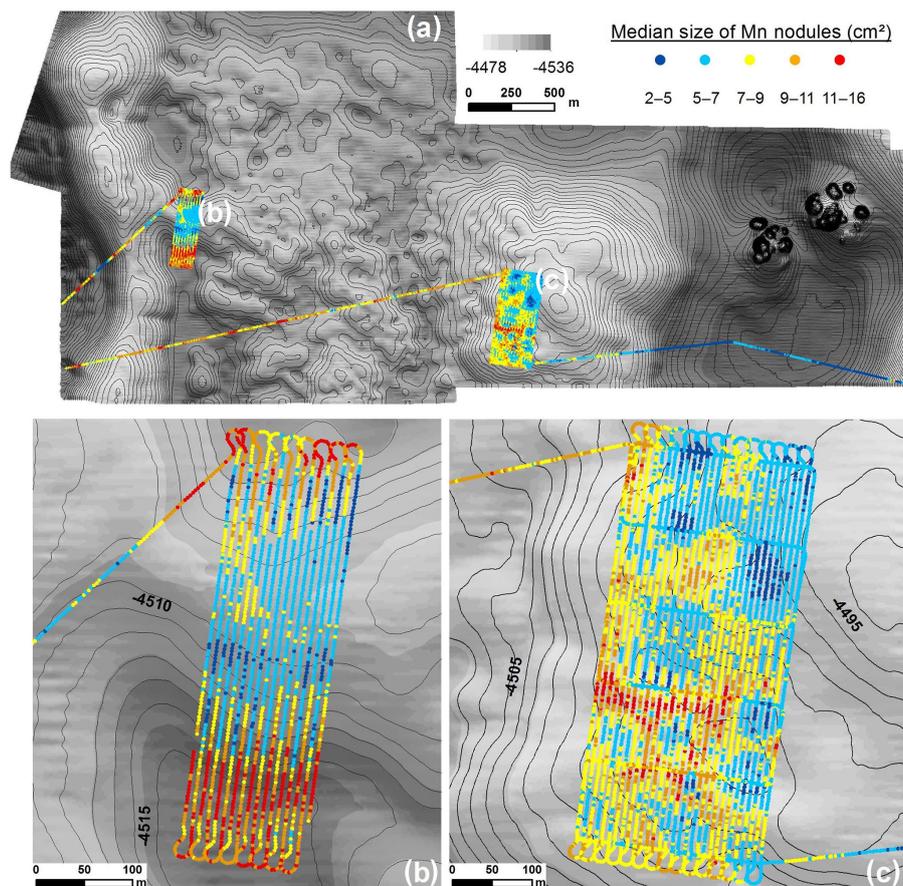
**Figure 7.** Scatterplot of the AUV altitude (m) and the estimated number of Mn nodules m<sup>-2</sup> inside subarea b. The colors correspond to the color scale in Fig. 7.

(Fig. 11a). This finding is in accordance with other studies, in which larger training samples tended to increase the per-

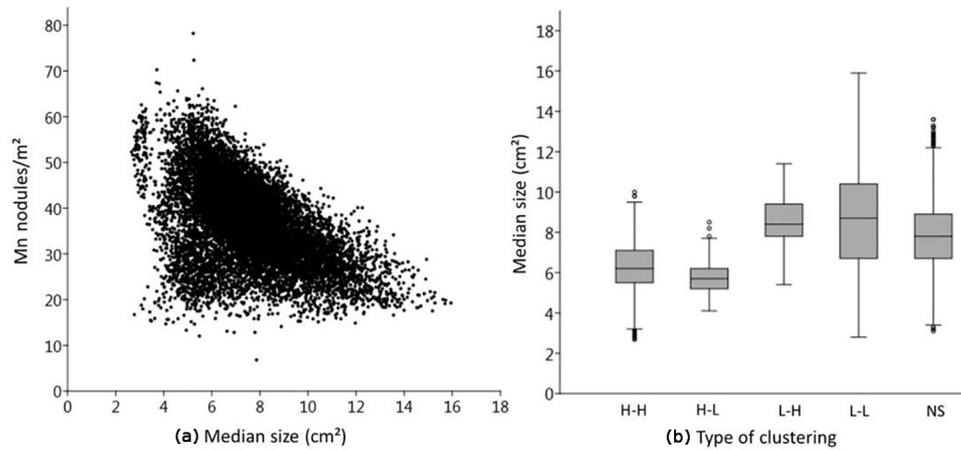
formance of RF (Li et al., 2010, 2011b; Millard and Richardson, 2015). The inclusion of a more representative range of the observed values, and consequently a larger spectrum of the causal underlying relationships, assists the RF to build a better model for the prediction of the value distribution inside the study area. For our data, the decrement becomes smaller when the size of the training sample increases further; it reaches a minimum value of 0.2 between 80 % and 90 %, showing that these additional 10 % do not notably benefit the RF model. However, the absence of stabilization of the error to a minimum value indicates that more optical data are needed from this block. The small decrement in error between 80 % and 90 % was the decisive factor to select 80 % of the data as training samples (also considering the larger number of remaining validation data and the reduced computational effort). Based on this data set, the examination of different numbers of trees showed that the RF error remains constant after 600 trees (Fig. 11b). Less trees result in a larger error; this becomes particularly evident with less than 300 trees. With more than 300 trees the range of the error is reduced (Appendix B). A higher number of trees en-



**Figure 8.** Adjacent AUV photos from consecutive dive tracks that were obtained inside subarea b from (a) lower (5–7 m) and (b) higher (9–11 m) altitudes. Note the decrement in the image brightness. (The area of the photos represents the central part of the photo, i.e., ca. 1/4 of the original photo size.)



**Figure 9.** (a) The spatial distribution of median Mn-nodule size (in  $\text{cm}^2$ ). (b) The estimation of median Mn-nodule size in subarea b and mainly in its southern part has been probably affected by the nonconstant altitude of the AUV. (c) The distribution of the median size inside subarea c shows also a clumped pattern.



**Figure 10.** (a) The plot of median size (cm<sup>2</sup>) and number of Mn nodules m<sup>-2</sup>. (b) The range of median size (cm<sup>2</sup>) in each type of cluster. Note the distinct difference in the range between the H–H and L–L cluster type.

ables higher *mtry* values as there are more opportunities for each variable to occur in several trees (Strobl et al., 2009). Similarly to the *ntree* parameter, a larger number of *mtry* values results in a reduced error (Fig. 11c). The error reaches a minimum and cannot be reduced further for *mtry* = 6; with values below 3, the error increases significantly. The different numbers of *ntree* reduced the error by only 0.6 in the MSR (from 18.8 to 18.2); in contrast, different *mtry* values reduced the error by 5.8 in the MSR (23.4 to 17.6), highlighting its importance for the prediction accuracy. In general a higher number of *mtry* values is suggested for RF studies with correlated variables to result in a less biased result regarding the importance of each variable; this is because the higher number increases the competition between highly correlated variables, giving more chances for different selections (Strobl et al., 2008). The finally selected *mtry* value of 6 coincides with the recommended approach for *mtry* (default, half of the default, and twice the default) suggested by Breiman (2001a). Despite the importance of this analysis, within the model with 80 % of the data as a training sample, the decrease in error by the use of RF tuned values instead of RF default values was only 0.7 in the MSR values, whilst the greatest reduction in error (16.5 in the MSR values) came from the increase in training data set size. This highlights the increased sensitivity of the method with respect to training data and that the recommended settings in the R randomForest package (Lia and Wiener, 2002) give trustworthy results, increasing its simplicity and operational character.

#### 4.3.2 Selection, application and external validation of the optimal model

Based on the abovementioned findings, the optimal RF regression model, which uses 80 % of training data, 600 trees, and 6 predictor variables to be randomly selected at each node, was selected and applied to the entire block G77. The

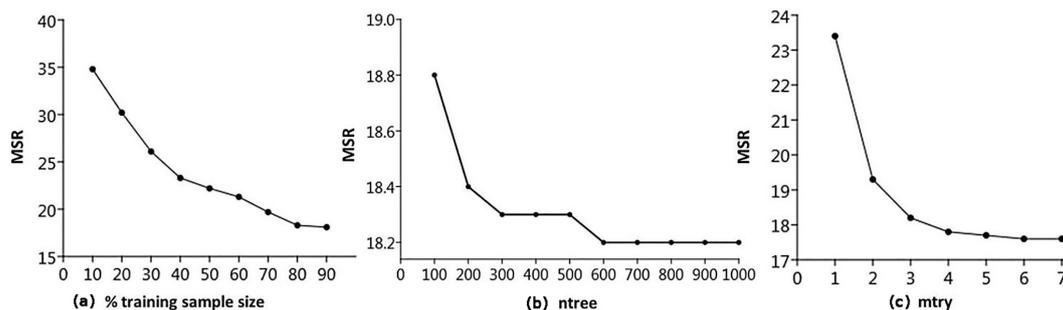
**Table 4.** The values of validation measures between predicted and observed data.

MAE	MSE	RMSE	R <sup>2</sup>	Pearson	Spearman
3.1	19.0	4.4	0.8	0.9	0.9

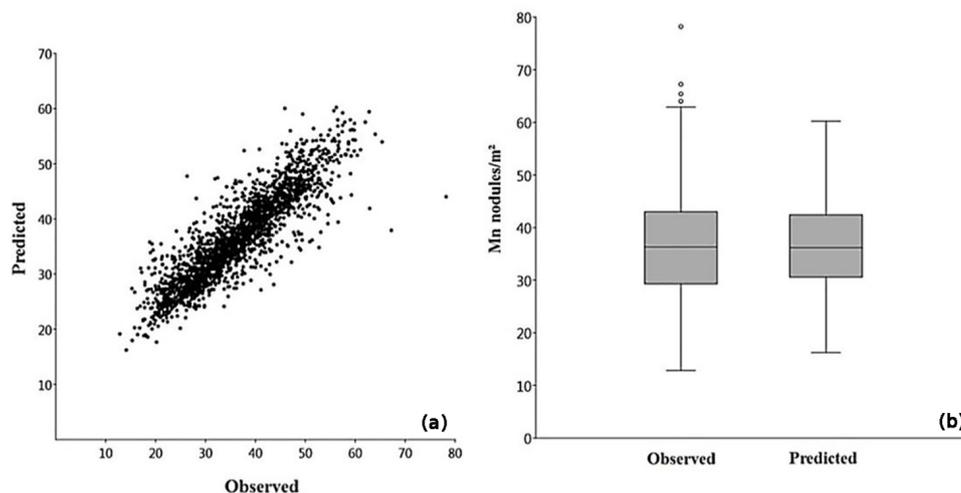
comparison of the predicted values with the observed values from the remaining 20 % (2255 observations) of validation data showed a good predictive performance (Table 4). Analytically, MAE and RMSE have very low values, R<sup>2</sup> has a high value, and both Pearson's and Spearman's correlation coefficients show a strong positive correlation between the predicted and observed values. The small deviation between MAE and RMSE and the same good correlation of the Pearson and Spearman factor point towards the absence of extremely high or low predicted values (outliers). Moreover, the performance is rather stable among all the iterations (Appendix B).

The scatterplot and box plot (Fig. 12a and b) illustrate this good match between predicted and observed values, as confirmed also by the descriptive statistics (Table 5). The residual analysis confirmed further the robustness of the model (Appendix B).

The statistical analysis also reveals the limitations of the RF model which cannot predict beyond the range of training values. It underestimates the maximum predicted values and overestimates the minimum values (Fig. 12b and Table 5), a limitation also mentioned by other authors (e.g., Horning, 2010). This happens because in regression RF, the result is the average value of all the predictions (Breiman, 2001a).



**Figure 11.** (a) The effect of training sample size in RF error (in MSR). (b) The effect of the *ntree* parameter in RF error (in MSR) for the 80 % training size. (c) The effect of the *mtry* parameter in RF error (in MSR) for the 80 % training size.



**Figure 12.** Comparison between observed and predicted values: scatterplot (a) and box plots (b).

### 4.3.3 RF-predicted distribution of Mn nodules $m^{-2}$

The final application of the RF model for the entire block G77 predicts that the majority of the area is covered by 30–45 Mn nodules  $m^{-2}$  (Fig. 13). In the central western part the distribution is quite uniform (at this scale) with few small areas of lower numbers. In the western part, there are two extended areas along the base of the hill with the lowest number of Mn nodules  $m^{-2}$ . Both of these areas have a linear shape in N–S direction and follow the seafloor topography with increased slope ( $>3^\circ$ ). The third main patch with minimum Mn nodules  $m^{-2}$  occurs in the eastern depression part. In contrast, areas with a higher number of Mn nodules  $m^{-2}$  are located mainly in the central upper part of the hill and eastward facing slope of eastern depression and south of the subrecent hydrothermally active area.

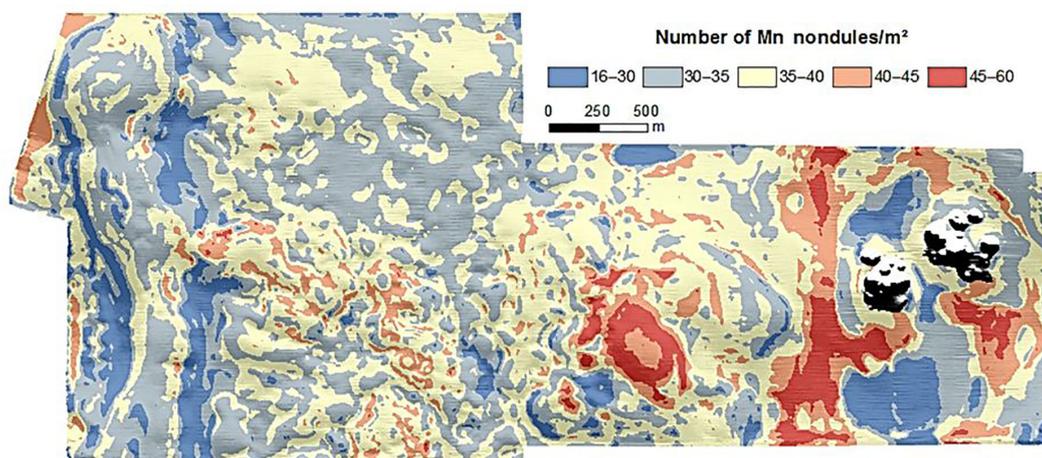
### 4.3.4 RF importance

The analysis of the RF variable importance showed that the best explanatory variable for the distribution of Mn nodules  $m^{-2}$  is depth (Fig. 14a). The partial dependence plot

of depth shows that there are specific depths, which promote higher numbers of Mn nodules  $m^{-2}$  aggregated in a nonlinear way (Fig. 14b). The two most important variables are the TPI\_B and TPI\_M. TRI, TPI\_F, C, and S follow in importance (Fig. 14a). All of them also contribute in a nonlinear way. (Appendix B). PI.C and Pr.C do not contribute significantly as explanatory variables in the performance of the RF model (Fig. 14a and Appendix B). Although the RF demonstrates good overall performance, the small study area and the arbitrary choice of the spatial scales for the TPI and other derivatives limit the potential of these variables as indicative explanatory variables on a broader scale. It is well established that surface derivatives are scale-dependent with different analysis scales to create alterations in results. Thus the combined use of different scales (here TPI) in the analysis and modeling procedure can produce models that do capture the natural variability and scale dependence (Wilson et al., 2007; Miller et al., 2014; Ismail et al., 2015; Leempoe et al., 2015). Due to the lack of relevant literature for AUV scale data sets, the C and TRI were created with the default scales of SAGA GIS v.6.3.0, while the three different TPI

**Table 5.** Descriptive statistics of observed and predicted values.

	Mean	SE	5 % trim. mean	Median	Mode	SD	Min	Max	C.L. (95 %)
Observed	36.5	0.2	36.3	36.3	40.8	9.4	12.8	78.2	0.4
Predicted	36.7	0.2	36.5	36.2	33.9	7.8	16.2	60.2	0.3

**Figure 13.** The RF-predicted distribution of Mn nodules  $\text{m}^{-2}$  inside block G77.

values were selected based on the minimum possible correlation among them.

#### 4.3.5 Estimation of abundance ( $\text{kg m}^{-2}$ ) of Mn nodules

The predicted Mn-nodule distribution was combined with the abundance from box corer data (and corrected with the ratio of buried to unburied Mn nodules, in order to include the top  $\sim 15$  cm of the sediment), resulting in the Mn nodules' abundance map shown in Fig. 15. According to this map, block G77 is a promising area for mining operations. The entire block is above the cutoff abundance of  $5 \text{ kg m}^{-2}$  (UNOET, 1987), with a mean value of  $33.8 \text{ kg m}^{-2}$ . We calculated that 84 % of block G77 has slopes below  $3^\circ$ ; steeper slopes are located mainly at the outer parts of the block, a fact that would ease establishing an ideal mining path. In this respect, the AUV scale mapping provides vital information for a potential mining path by decreasing the possibility of machine failure due to poorly mapped steep slopes not detected by, e.g., ship-based bathymetry (Peukert et al., 2018b). Mn-nodule distribution maps with this resolution increase the mining efficiency because local deposit variations can significantly affect the performance of the pickup rate, which is likely determined by technical parameters of the mining vehicle as well as the size, burial depth, and abundance of Mn nodules in the seafloor (Chung, 1996). The exclusion of areas with slopes  $> 3^\circ$  resulted in  $8 \text{ km}^2$  minable seafloor surface. Assuming a constant 80 % collection efficiency (Volkman and Lehnen, 2018) and a 30 % reduction in the Mn-nodule weight by the removal of water (Das and Anand, 2017), the dry mass of

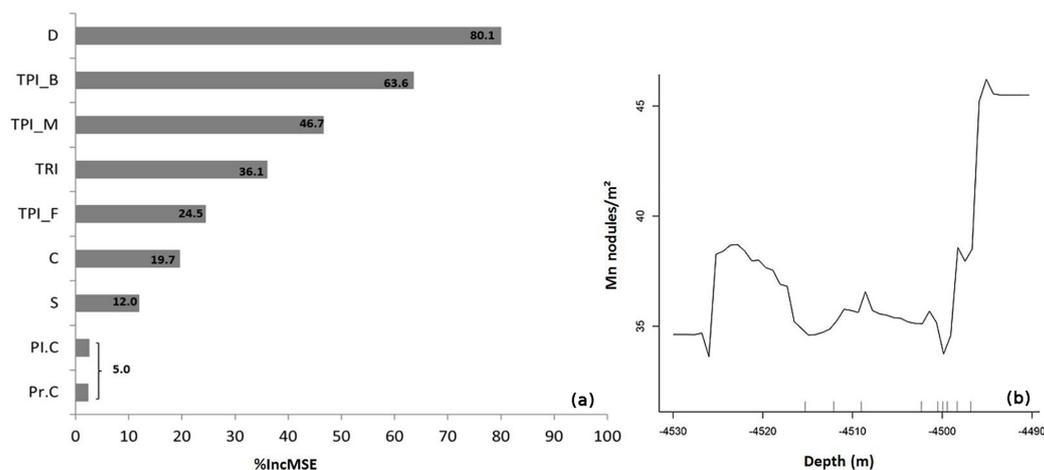
**Table 6.** The estimated amount of metal mass for five metals, based on the average values of metal content inside CCZ and a five-metal HCL-leach recovery method (Volkman, 2015).

Total wet mass (t):	270 400				
Total dry mass (t):	189,280				
Metal content	Mn	Ni	Cu	Co	Mo
Wt %	26.68	1.31	1.11	0.22	0.06
Equal to (t)	50 500	2480	2101	416	113
90 % metal recovery (t)	45 450	2232	1891	374	102

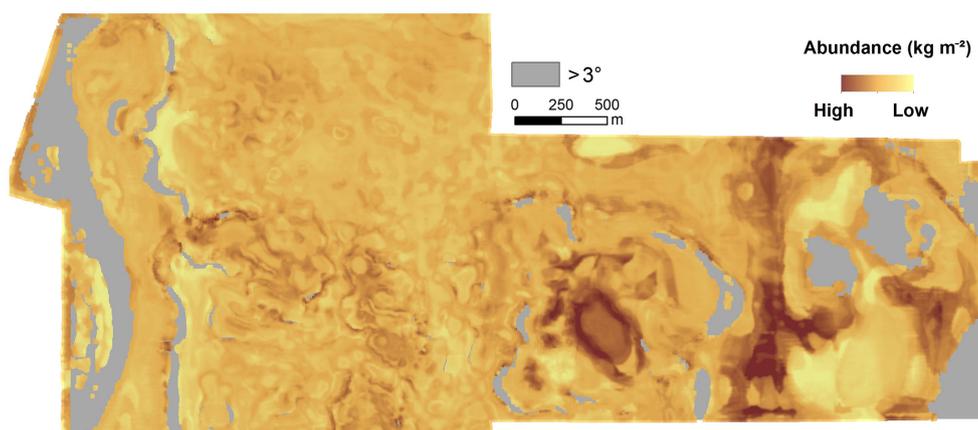
Mn nodules that can be extracted from the surface and the first 15 cm of the sediment column amounts to ca. 190 000 t. In a back-of-the-envelope calculation this quantity – assuming constant metal content inside the study area, equal to the average metal concentrations inside the CCZ (Table 6) (Volkman, 2015), and 90 % metal recovery efficiency – could result in an estimated resource haul of 45 450 t Mn, 2232 t Ni, 1891 t Cu, 374 t Co, and 102 t Mo (Table 6).

## 5 Discussion

We present a case study that highlights the applicability of the combination of AUV bathymetric and optical data for Mn-nodule resource modeling using RF machine learning. The use of AUVs for collecting hydroacoustic and optical



**Figure 14.** (a) The variable importance of each predictor in the RF model. (b) The partial dependence plot of depth. The ticks on the graphs indicate the deciles of the data.



**Figure 15.** The total abundance of Mn nodules from the surface and embedded in the sediment (max 15 cm), in areas with slope  $\leq 3^\circ$  inside block G77 (continuous values of abundance are not given due to confidentiality).

data in areas of scientific and commercial interest can provide more precise bathymetric and Mn-nodule distribution maps. Regarding the bathymetric maps, the accurate and detailed reconstruction of the seafloor bathymetry at meter-scale resolution enables to use bathymetry and its derivatives as source data layers within a high-resolution RF model. These data should have high-quality characteristics, as the presence of acquisition artefacts may affect the robustness of the modeling procedure (Preston, 2009; Herkül et al., 2017). The combined use of cameras as the DeepSurveyCamera (Kwasnitschka et al., 2016) for acquiring high-resolution photographs and an automated analysis with a state-of-the-art algorithm (Schoening et al., 2017a) provide essential quantitative information about the distribution of Mn nodules. Image analysis results are more robust for constant AUV altitudes (7–9 m) above flat areas ( $< 3^\circ$ ), while the alternation of the flying altitude and camera orientation during the ascending and descending phases limits the quality of the

obtained images and can affect the derived number of Mn nodules  $\text{m}^{-2}$ .

Inside block G77, the number of Mn nodules  $\text{m}^{-2}$  seems to follow a normal distribution without extreme outliers and without being linearly correlated with the predictor variables used. Spatially, a clumped autocorrelated pattern is demonstrated, mainly with clustered areas of H–H and L–L values. It is still unclear if this heterogeneity is caused by external processes (e.g., topographic characteristics, geochemical conditions, or the availability of nucleus material) or if it results from the interaction of neighboring Mn nodules. The areas with a higher number of Mn nodules could provide more fragments as potential nucleus material. However, the less available space in these areas may make individual Mn-nodule growth more difficult, resulting in smaller median sizes. Conversely, a recent study from Kuhn and Rathke (2017) showed that the blanketing of the Mn nodules by sediments is higher for larger Mn nodules and, as a re-

sult, fewer large nodules will be counted, resulting in biased results in areas where the Mn nodules are bigger. Probably all of these effects can happen at the same time (with different degrees of influence), promoting a given, scale-dependent spatial structure.

This study did not consider the geochemical properties of the sediments as input data in the modeling process, which might give additional clues as to why Mn nodules are distributed as they are. However, RF importance and partial dependence plots show that bathymetric and topographic factors tend to affect this distribution in a nonlinear way and with the bulk of data plotting in specific ranges of the bathymetric derivatives. Classic studies have shown that the bathymetry and the variation in the topographic characteristics of the seafloor affects the sediment deposition environment and bottom currents and thus also geochemical processes in the sediment. All these factors determine Mn-nodule growth and thus affect the distribution of Mn nodules on regional scales (e.g., Craig, 1979; Sharma and Kodagali, 1993). It is still unknown how these properties influence the Mn-nodule distribution on meter to tens of meter scales as seen in our AUV data. The nonlinear relationship between Mn nodules and bathymetry on such high-resolution scales only began to be investigated very recently (e.g., Peukert et al., 2018; Alevizos et al., 2018). To elaborate more on the hydrodynamic and geochemical reasons behind the observed distribution pattern, we would need more investigations at and in the sediment on the same scale.

It should be acknowledged that the aim of any ML predictive model is to derive accurate predictions based on an existing (large) number of measurements to capture a complex underlying relationship (e.g., nonlinear and multivariate) between different types of data, for which our theoretical knowledge or conceptual understanding is still under development (Schmueli, 2010; Lary et al., 2016). Especially due to the constantly increasing size of scientific multivariate data in marine sciences and the existence of such nonlinear relationships between predictor and response variables (e.g., Zhi et al., 2014; Li et al., 2017), ML and RF are considered important analytic tools that can objectively reveal patterns of a (unknown) phenomenon (Genuer et al., 2017; Kavenski et al., 2009; Lary et al., 2016). Such predictions may be used to derive causalities or may drive the creation of new hypotheses. In other words, for a predictive model, the “unguided” data analyses come first and the interpretation follows (Breiman, 2001b; Schmueli, 2010; Obermeyer and Emanuel, 2016). This “a priori” knowledge of the distribution of the Mn-nodule number and size on such a scale can contribute to the biological data survey planning, too. Recent studies showed that the abundance and species richness of nodule fauna inside the CCZ is affected by the abundance of Mn nodules (Amon et al., 2016; Vanreusel et al., 2016) as well as their size (Veillette et al., 2007). Thus, high-priority areas (e.g., those with the highest commercial interest) can be targeted for sampling based on the results of

optic data and RF modeling. The RF modeling takes advantage of the multilayer information (here: hydroacoustic and optical data), handling their complex relationships effectively while being resistant to overfitting (Breiman, 2001a). Moreover, the randomization of the input training points in each tree in each run results in a completely different training data set each time with mixed points from the entire study area. This random selection and mixing of points is appropriate for clustered data, as it ignores their spatial locations and consequently limits the influence of spatial autocorrelation (Appendix B). Along these lines, several authors have included the values of latitude/longitude and even the LMI values as predictor variables in order to increase the model performance (e.g., Li, 2013; Li et al., 2011b, 2013). RF has a high operational character due to its relatively simple calibration, which does not request extensive data preparation/transformation or the need for geostatistical assumptions (e.g., stationarity). The selection of the MGET toolbox (Roberts et al., 2010) further increased the simplicity of the workflow, as the RF modeling was performed entirely inside a graphic environment familiar to many geoscientists. As RF model runs can be implemented inside various software packages in future implementations of this workflow, it would be interesting to include the uncertainty for the associated predictions, e.g., with the use of the quantile regression forests (Meinshausen, 2006) from the `quantregForest` R package (Meinshausen, 2012). However, this will increase the computational time (Tung et al., 2014) and the simplicity of the procedure, especially if other recently proposed methodologies of estimating the uncertainty are used: the jackknife method (Wager et al., 2014), the Monte Carlo approach (Coulston et al., 2016), and the  $U$  statistics approach (Mentch and Hooker, 2016).

Similarly to other studies (e.g., Cutler et al., 2007; Millard and Richardson, 2015), RF showed increased stability in its performance, allowing a small number of iterations to compute sufficient results. The examination of the main two tuning parameters (*ntree* and *mtry*) showed that the model performance can be increased compared to default values. However, the largest improvement results from using more training data. In this respect, more photos would potentially improve the RF performance as no clear threshold was observed. Although the number of 11 276 photos seems to represent a large data set, the heterogeneity of the distribution and the occurrence of spatial clusters (patches) in different sizes and the inherent need of RF and ML in general for big training data sets (van der Ploeg et al., 2014; Obermeyer and Emanuel, 2016) stresses the need for collecting more and well-distributed data. The influence of the number of training data for model performance still remains a discussion point between studies showing an improvement by adding more data (e.g., Bishop, 2006) and other studies presenting stable performance of the model even if more data are added (e.g., Zhu et al., 2012). The availability of more data, especially if they are better distributed (i.e., data that will include the

entire range of the number of Mn nodules  $m^{-2}$  and come from all the different sub-terrains), would most likely reinforce the model to build better and wider relationships between the predictor and response variables, keeping also a larger number of validation data points.

Finally, the resource assessment showed that block G77 is a potential mining area with high average Mn-nodule density and gentle slopes. While the threshold of  $3^\circ$  (UNOET, 1987) was used here, newer plans for mining machines seem to enable operations on steeper slopes (Atmanand and Ramadass, 2017), increasing the total amount of collected Mn nodules within the area considered herein.

## 6 Conclusions

The results of this study show that the acquisition and analysis of optical seafloor data can provide quantitative information on the distribution of Mn nodules. This information can be combined with AUV-based MBES data using RF machine learning to compute predictions of Mn-nodule occurrence on small operational scales. Linking such spatial predictions with sampling-based physical Mn-nodule data provides an efficient and effective tool for mapping Mn-nodule abundance.

*Data availability.* The data used in this work are available at PAN-GAEA. This includes MBES ship-based data (Greinert, 2016), optical imagery (Greinert et al., 2017; Schoening, 2017c), and the source code of the CoMoNoD algorithm (Schoening, 2017b). The MBES AUV-based data are not publicly available due to the confidentiality of coordinates.

## Appendix A: Methodology

### A1 Hydroacoustic data acquisition and post-processing

The calculation of the bathymetric derivatives was performed with the SAGA GIS v6.3.0 Morphometry library ([http://www.saga-gis.org/saga\\_tool\\_doc/6.3.0/ta\\_morphometry.html](http://www.saga-gis.org/saga_tool_doc/6.3.0/ta_morphometry.html), last access: 6 December 2018).

### A2 Spatial statistics

Global Moran's I and local Moran's I were performed with the ArcMap™ 10.1 software, using the Spatial Statistics toolbox, according to the equations provided. As a null hypothesis, it is assumed that the examined attribute is randomly distributed among the features in the study area. For the optimal conceptualization of spatial relationships, the inverse Euclidian distance method was selected, as it is appropriate for modeling processes with continuous data in which the closer two samples are in space, the more likely they are to interact/influence each other or have been influenced for the same reasons. The distance threshold was set at 50 m, and the increment analysis was performed with a step of 50 m. Moreover, the spatial weights were standardized in order to minimize any bias that exists due to sampling design (uneven number of neighbors). Apart from the index value, the  $p$  value and  $Z$  score are also provided. The local Moran's I indicates statistically significant clusters and outliers for a 95 % confidence level. The high number of observations ( $\gg 30$ ) that was used ensures the robustness of the indexes.

### A3 RF predictive modeling (selection of predictor variables)

Correlation among the derivatives was checked by the Spearman's correlation coefficient ( $\rho$ ). This coefficient was preferred due to the skewed distribution of the values in the derivatives. The majority of the possible pairs is uncorrelated or weakly correlated. Only C vs. TPI\_F and TRI vs. S have a strong correlation. However, they should not be excluded as they express different topographic characteristics and they are not correlated with the remainder of derivatives. It should be mentioned that in similar studies even higher thresholds have been used during the selection of predictor variables (Che Hasan et al., 2014; Li et al., 2016, 2017).

The nine training samples with different sizes were created by the MGET tool "Randomly Split Table into Training and Testing Records". The spatial randomness of the procedure, combined with the many available data, resulted in training samples with similar descriptive statistics.

**Table A1.** Spearman's correlation coefficient for each pair of predictor variables.

	D	S	Pl.C	Pr.C	TPI_B	TPI_M	TPI_F	C	TRI
D									
S	-0.07								
Pl.C	0.06	-0.02							
Pr.C	0.08	-0.01	0.37						
TPI_B	0.76	-0.09	0.13	0.16					
TPI_M	0.36	-0.06	0.20	0.27	0.72				
TPI_F	0.23	-0.05	0.33	0.41	0.47	0.77			
C	-0.30	0.05	-0.25	-0.34	-0.54	-0.79	-0.90		
TRI	-0.10	0.91	-0.02	-0.03	-0.12	-0.06	0.04	0.05	

**Table A2.** Descriptive statistics of different training samples.

% training sample:	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %
Training set size	1127	2255	3383	4511	5638	6766	7894	9021	10148
Mean	36.5	36.3	36.6	36.6	36.6	36.7	36.6	36.7	36.6
SE	0.3	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1
SD	9.3	9.2	9.4	9.2	9.2	9.3	9.3	9.2	9.3
Minimum	7	13	12	13	12	14	7	7	7
Maximum	63	70	72	66	78	78	78	72	78

## Appendix B: Results

### B1 RF predictive modeling (calibration of the model)

The descriptive statistics of the performance of each model were used as decision factors for the number of iterations (Tables B1–B5). In all cases, the mean value with very low standard error, very low standard deviation, range, and the 95 % confidence interval indicate a rather stable performance, without the need for further iterations.

#### B1.1 Data exploration

The histogram of Mn nodules  $m^{-2}$  (Fig. B1) shows a good fit with the superimposed theoretical normal curve, with the shape of the distribution being rather symmetrical. This fact is supported by the equal 5 % trimmed mean and median and the slightly different mode (Table B6). Similarly, the visual inspection of the probability plot (Fig. B1) shows a good match as a linear pattern is observed for the greatest part, with slight deviation existing only in the outer parts of the curve. According to the box plot, there are only 21 mild outliers (according to Hoaglin et al., 1986; Dawson, 2011), which correspond to 0.18 % of the total observation. This percentage is smaller than the 0.8 % threshold that has been suggested for normal disturbed data (Dawson, 2011). The small values for skewness and kurtosis combined with the large sample size further support the normally distributed pattern of the data (Table B6). Especially for large data samples, measurements of skewness and kurtosis combined with the visual inspection of histogram and probability plot are recommended ways of examining the normality (D'Agostino et al., 1990; Yaziki and Yolacan, 2007; Field, 2009; Ghasemi and Zahediasl, 2012; Kim, 2013).

The potential linear correlation between depth, bathymetric derivatives, and the number of Mn nodules  $m^{-2}$  was investigated using the Spearman's rank correlation coefficient ( $\rho$ ) (Table B7) because of the skewed distribution and presence of extreme values in the depth and bathymetric derivative values (Mukaka, 2012).

#### B1.2 Selection, application, and external validation of the optimal model

Despite the fact that RF is a full nonparametric technique and there is no need for the residuals to follow specific assumptions (Breiman, 2001a), the examination of them can provide an in-depth look at RF performance characteristics. The scatterplot of residuals against predicted values shows a random pattern, which is also confirmed by the low values of Pearson, Spearman, and  $R^2$  coefficients between predicted values and residuals (Fig. B2 and Table B8). Moreover, the residuals tend to cluster towards the middle of the plot without being systematically high or low and having a zero mean value (Fig. B2 and Table B9). Their constant variance (homoscedasticity) implies that the distribution of error has the

same range for almost all fitted values. Indeed, 99.3 % of the residuals are inside the range  $\pm 15$  and 81.2 % are inside the range  $\pm 5$  (Table B10). The presence of outliers is very limited without affecting the main statistical characteristics of residuals (Table B9), indicating that the model adequately fits the overwhelming majority of the observations ( $> 2165$ ) and only random variation (that exists in any real, natural phenomenon) or noise can occur.

The spatial autocorrelation analysis of the residuals using the global Moran's index (same settings as Appendix A), showed low spatial autocorrelation ( $I = 0.112112$ ,  $p < 0.01$  and  $Z$  score  $> 2.58$ ). The index number of the residuals is relatively low compared with the high initial values of the original data ( $I = 0.69890$  and  $I = 0.697747$  for the entire data set and the 80 % training data set, respectively). Moreover, the spatial autocorrelation of the 5 % trimmed residuals is only 0.093832. According to similar studies (i.e., regression RF), the presence of spatial autocorrelation in the residuals of the model can result in underestimation of the true prediction error (Ruß und Kruse, 2010). The presence of low spatial autocorrelation values in the residuals of regression RF has been reported also by other authors (e.g., Mascaro et al., 2014; Xu et al., 2016), and it is a common problem in all the well-established machine learning methods (e.g., random forests, neural network, gradient boosting machines, and support vector machines) when dealing with regression predictions of spatial variables (Gilardi and Bengio, 2009; Ruß und Kruse, 2010; Santibanez et al., 2015a, b). The spatial plotting and visual examination of the residuals (Fig. B3) showed that this spatial clustering exists mainly in the small subarea b and especially in the areas which are associated with an increased slope ( $> 3^\circ$ ), where the AUV is forced to vary its altitude between the ascending and descending phase and consequently affects the image quality and the later modeling results.

#### B1.3 RF importance

The production of the RF partial dependence plots show the nonlinear character between the Mn nodules  $m^{-2}$  and the bathymetric derivatives.

**Table B1.** Descriptive statistics of MSR from different training set sizes, after 10 iterations with default settings.

% training sample:	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %
Mean	34.8	30.2	26.1	23.3	22.2	21.3	19.7	18.3	18.1
SE	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Median	34.8	30.3	26.1	23.2	22.2	21.3	19.7	18.3	18.1
Mode	34.7	30.3	26.1	23.2	22.2	21.3	19.7	18.3	18.1
SD	0.2	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.1
Minimum	34.5	30.1	25.9	23.2	22.1	21.2	19.6	18.2	18.1
Maximum	35.1	30.4	26.3	23.5	22.3	21.3	19.7	18.3	18.1
C.I. (95.0 %)	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0

**Table B2.** Descriptive statistics of MSR from a different number of the *ntree* parameter, after 10 iterations with 80% of the sample as training data and *mtry* = 3.

<i>ntree</i> :	100	200	300	400	500	600	700	800	900	1000
Mean	18.8	18.4	18.3	18.3	18.3	18.2	18.2	18.2	18.2	18.2
SE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Median	18.8	18.4	18.3	18.3	18.3	18.2	18.2	18.2	18.2	18.2
Mode	18.8	18.4	18.3	18.3	18.3	18.2	18.2	18.2	18.2	18.2
SD	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.0	0.0
Minimum	18.5	18.4	18.2	18.2	18.2	18.1	18.1	18.1	18.1	18.1
Maximum	18.9	18.5	18.5	18.4	18.3	18.3	18.3	18.3	18.2	18.2
C.I. (95.0 %)	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0

**Table B3.** Descriptive statistics of MSR from different number of the *mtry* parameter, after 10 iterations with 80% of the sample as training data and *ntree* = 600.

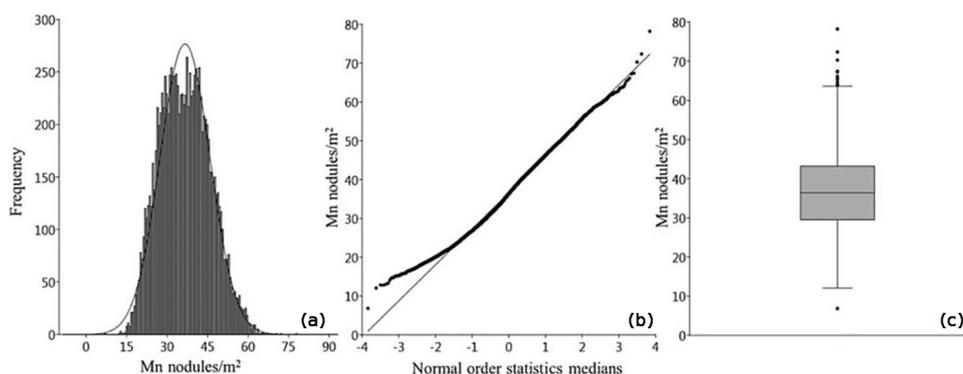
<i>mtry</i> :	1	2	3	4	5	6	7
Mean	23.4	19.3	18.2	17.9	17.7	17.6	17.6
SE	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Median	23.4	19.3	18.2	17.9	17.7	17.6	17.6
Mode	23.4	19.3	18.2	17.9	17.7	17.6	17.6
SD	0.0	0.1	0.1	0.1	0.0	0.0	0.0
Minimum	23.3	19.1	18.1	17.8	17.6	17.5	17.6
Maximum	23.5	19.4	18.3	17.9	17.7	17.7	17.7
C.I. (95.0 %)	0.0	0.1	0.0	0.0	0.0	0.0	0.0

**Table B4.** Descriptive statistics of MSR for the optimum selected RF model, after 30 iterations with 80% of the sample as training data, *ntree* = 600, and *mtry* = 6.

	Mean	SE	Median	Mode	SD	Minimum	Maximum	C.I. (95 %)
Optimum RF	17.6	0.0	17.6	17.6	0.0	17.5	17.7	0.0

**Table B5.** Descriptive statistics of RF importance for the optimum RF model, after 30 iterations with 80 % of the sample as training data,  $n_{tree} = 600$ , and  $mtry = 6$ .

RF importance:	Depth	TPI_B	TPI_M	TRI	TPI_F	C	S	Pl.C	Pr.C
Mean	80.1	63.6	46.7	36.1	24.5	19.7	12.0	2.6	2.4
SE	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Median	80.1	63.5	46.7	36.1	19.7	19.7	11.9	2.6	2.4
Mode	80.1	63.3	46.9	36.1	19.8	19.8	11.9	2.6	2.4
SD	0.4	0.6	0.6	0.2	0.2	0.2	0.2	0.0	0.0
Minimum	79.1	62.6	45.0	35.7	19.2	19.2	11.7	2.5	2.3
Maximum	80.8	64.9	47.7	36.4	20.1	20.1	12.4	2.6	2.5
C.I. (95.0 %)	0.1	0.2	0.2	0.1	0.1	0.1	0.1	0.0	0.0



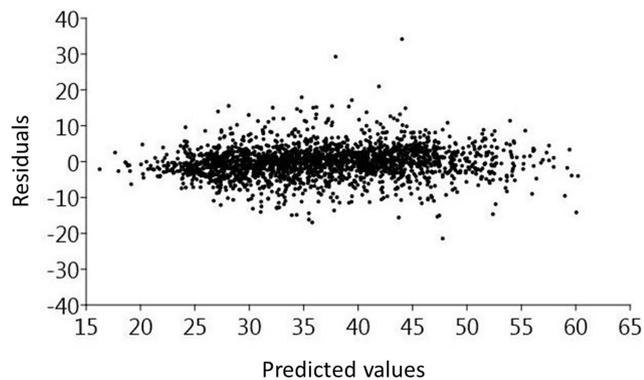
**Figure B1.** (a) Histogram of Mn nodules  $m^{-2}$  with the superimposed normal curve. (b) The normal probability plot of Mn nodules  $m^{-2}$ . (c) The box plot of Mn nodules  $m^{-2}$ .

**Table B6.** The descriptive statistics of the number of Mn nodules  $m^{-2}$ .

	Mean	5 % trim. mean	Median	Mode	SD	Min	Max	Skew.	Kurtosis
Mn nodules $m^{-2}$	36.6	36.4	36.4	39	9.2	6.8	78.2	0.1	-0.4

**Table B7.** The Spearman’s rank correlation coefficient between Mn nodules  $m^{-2}$ , depth, and bathymetric derivatives.

Depth	Slope	TRI	Pl.C	Pr.C	TPI_B	TPI_M	TPI_F	C.I. (95.0 %)
0.38	0.08	0.07	0.03	0.04	0.29	0.24	0.05	-0.14



**Figure B2.** Scatterplot between residuals and predicted values.

**Table B8.** Pearson, Spearman, and  $R^2$  correlation coefficients between residuals and predicted values.

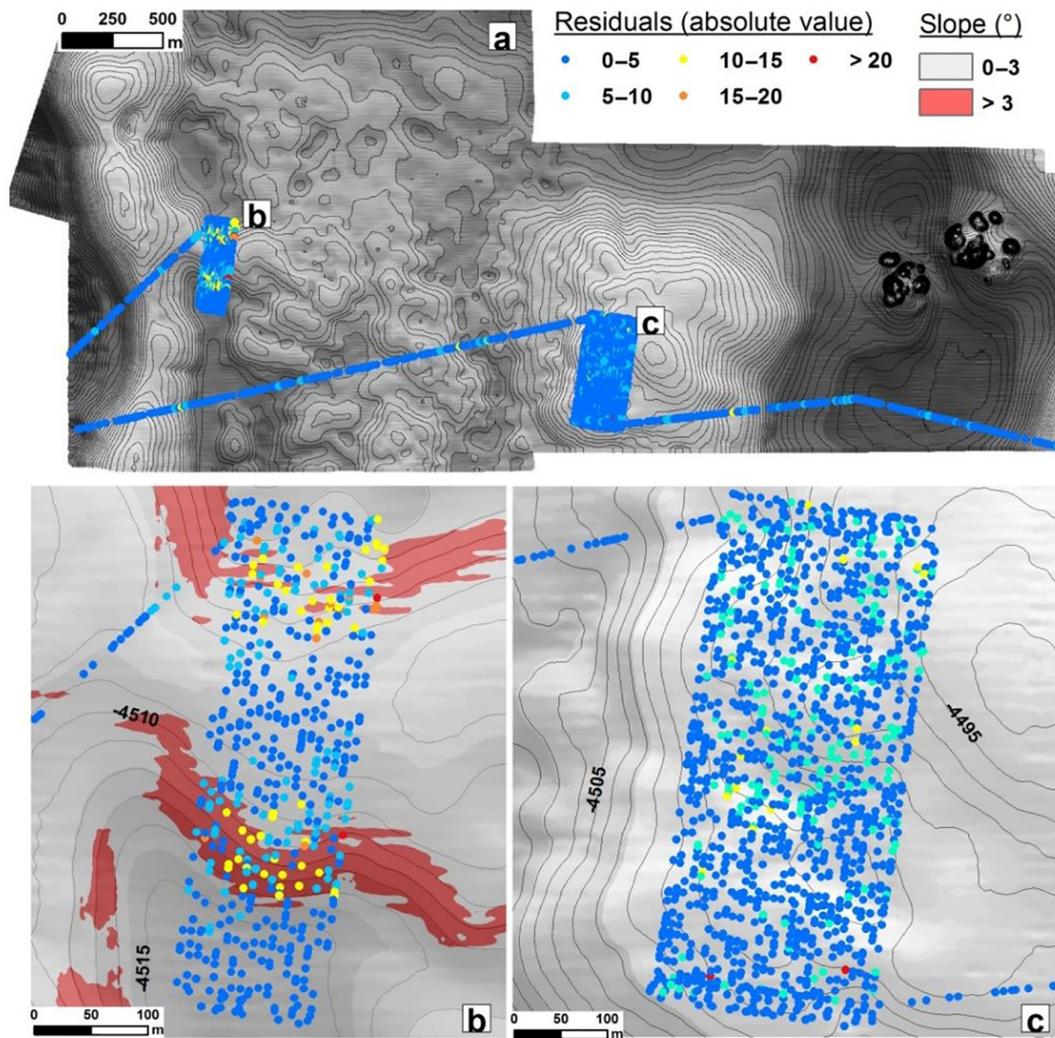
	Pearson	Spearman	$R^2$
Correlation of residuals and predicted values	0.1	0.2	0.0

**Table B9.** Main descriptive statistics of residuals and 5 % trimmed residuals.

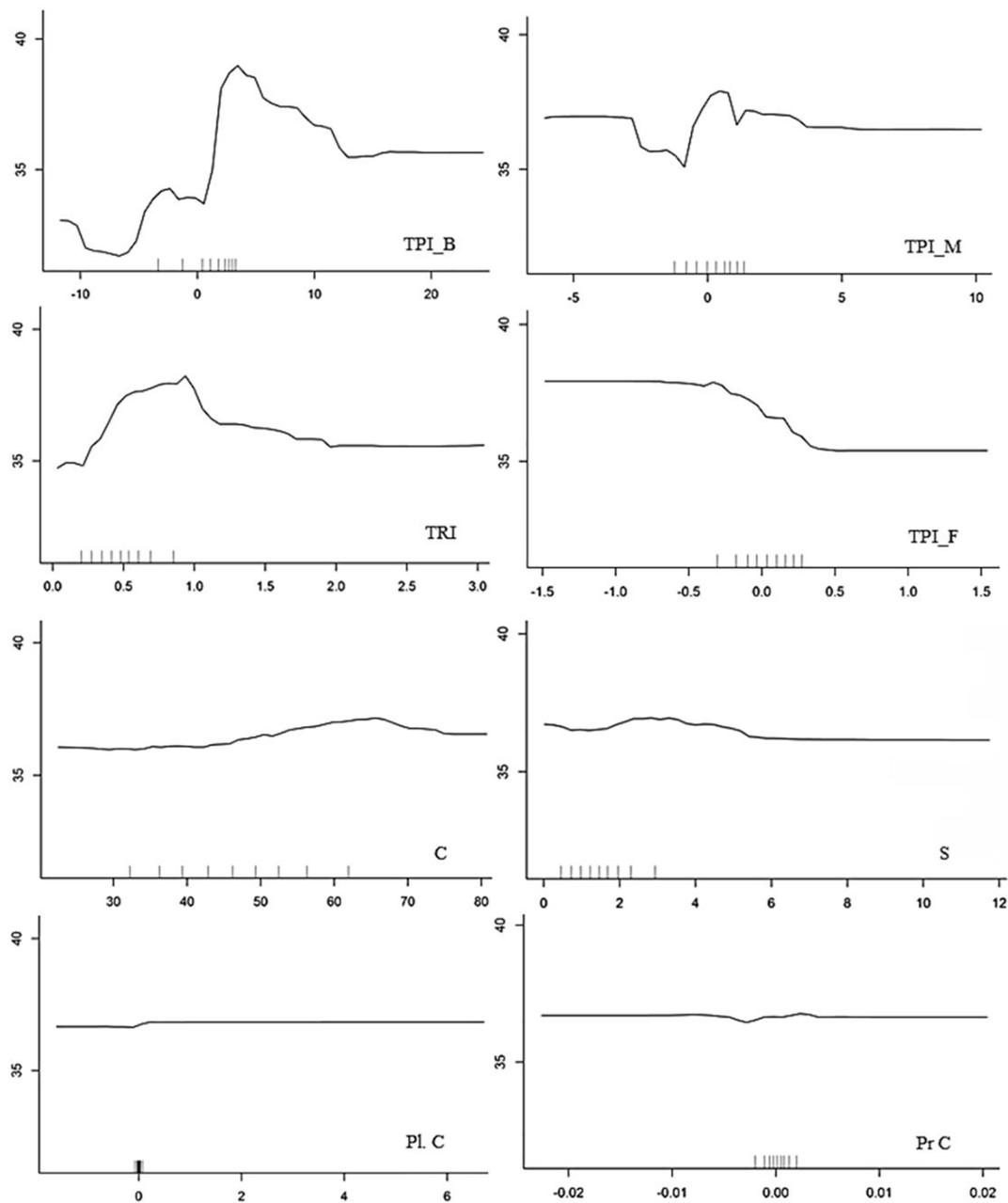
	Mean	SE	Median	Mode	SD
Residuals	-0.2	0.1	-0.2	0.6	4.4
5 % trimmed residuals	-0.2	0.1	-0.2	0.6	2.9

**Table B10.** Residuals range.

Residuals range	$\pm 20$	$\pm 15$	$\pm 10$	$\pm 5$
% of residuals	99.8	99.3	96.1	81.2



**Figure B3.** Spatial plotting of the RF residuals (absolute values). The intervals of their range are in accordance with the Table B10.



**Figure B4.** Partial dependence plots for each of the predictor variables. The y axis represents the number of Mn nodules  $m^{-2}$  and the x axis the values of each predictor variable (depth derivatives). The ticks on the graphs indicate the deciles of the data.

*Author contributions.* IZG processed the MBES and AUV data, performed the RF modeling, the statistical and GIS analysis, and wrote the paper. TS contributed to the survey design with respect to the optic data, developed the CoMoNoD algorithm, and processed the optic data. EA was involved in developing the idea of using RF for modeling and contributed to the GIS analysis. JG contributed to the survey by designing the MBES and the optic data survey planning, acquiring the MBES and the optic data, verifying the analytical methods, and supervising the project. All authors discussed the results, provided critical feedback, and contributed to the final paper.

*Competing interests.* The authors declare that they have no conflict of interest.

*Special issue statement.* This article is part of the special issue “Assessing environmental impacts of deep-sea mining – revisiting decade-old benthic disturbances in Pacific nodule areas”. It is not associated with a conference.

*Acknowledgements.* We thank the captain and crew of RV *Sonne* for their contribution to a successful cruise. We express our gratefulness to the GEOMAR AUV team for their support during the cruise. We thank Anja Steinführer for preprocessing of the AUV MBES data and Mareike Kampmeier for advice during the post-processing analysis of the MBES data. We thank Inken Preuss for proofreading the paper. Finally, we thank the GEOMAR Library team for its support in gathering the necessary bibliography. All data were acquired within the framework of the JPIO Project “Ecological Aspects of Deep-Sea Mining”, funded through BMBF grant 03F0707A. Funding for Iason-Zois Gazis was also made available through MarTERA grant COMPASS-Drimp from BMWi (03SX466B). This is publication 35 of the DeepSea Monitoring Group at GEOMAR Helmholtz Centre for Ocean Research Kiel.

Edited by: Daniel O. B. Jones

Reviewed by: two anonymous referees

## References

- Alevizos, E., Schoening, T., Koeser, K., Snellen, M., and Greinert, J.: Quantification of the fine-scale distribution of Mn-nodules: insights from AUV multi-beam and optical imagery data fusion, *Biogeosciences Discuss.*, <https://doi.org/10.5194/bg-2018-60>, in review, 2018.
- Anselin, L.: Local Indicators of Spatial Association – LISA, *Geogr. Anal.*, 27, 93–115, <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>, 1995.
- Atmanand, M. A. and Ramadass, G. A.: Concepts of Deep-Sea Mining Technologies, in: *Deep-Sea Mining*, edited by: Sharma, R., Resource Springer, Cham. Online ISBN 978-3-319-52557-0, [https://doi.org/10.1007/978-3-319-52557-0\\_6](https://doi.org/10.1007/978-3-319-52557-0_6), 2017.
- Bellingham, J.: Autonomous underwater vehicles (AUVs), in: *Encyclopedia of Ocean Sciences*, edited by: Steele, H., Academic Press, San Diego, 212–216, <https://doi.org/10.1006/rwos.2001.0303>, 2001.
- Bingham, D., Drake, T., Hill, A., and Lott, R.: The Application of Autonomous Underwater Vehicle (AUV) Technology in the Oil Industry – Vision and Experiences. FIG XXII International Congress Washington, DC USA, 19–26 April, 1–13, 2002.
- Breiman, L.: Random forests, *Machine Learning*, 45, 5–32, <https://doi.org/10.1023/A:101093340>, 2001a.
- Breiman, L.: Statistical Modeling: The Two Cultures, *Stat. Sci.*, 16, 199–215, <https://doi.org/10.1214/ss/1009213726>, 2001b.
- Caress, D. W. and Chayes, D. N.: MB-System: Mapping the Seafloor, available at: <https://www.mbari.org/products/research-software/mb-system> (last access: 6 December 2018), 2017.
- Caress, D. W., Thomas, H., Kirkwood, W. J., McEwen, R., Henthorn, R., Clague, D. A., Paull, C. K., and Paduan, J.: High-resolution multibeam, sidescan and subbottom surveys using the MBARI AUV D, in: *Marine Habitat Mapping Technology for Alaska*, edited by: Allan, B., Greene, H. G., and Reynolds, J. R., Alaska Sea Grant College Program, University of Alaska Fairbanks, 47–69, <https://doi.org/10.4027/mhmta.2008.04>, 2008.
- Carranza, E. J. M. and Laborte, A. G.: Random Forest Predictive Modelling of Mineral Prospectivity with Small Number of Prospects and Data with Missing Values in Abra (Philippines), *Comput. Geosci.*, 74, 60–70, <https://doi.org/10.1016/j.cageo.2014.10.004>, 2015a.
- Carranza, E. J. M. and Laborte, A. G.: Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm, *Ore Geol. Rev.*, 7, 777–787, <https://doi.org/10.1016/j.oregeorev.2014.08.010>, 2015b.
- Carranza, E. J. M. and Laborte, A. G.: Data-Driven Predictive Modeling of Mineral Prospectivity Using Random Forests: A Case Study in Catanduanes Island (Philippines), *Nat. Resour. Res.*, 25, 35–50, <https://doi.org/10.1007/s11053-015-9268-x>, 2016.
- Chakraborty, B. and Kodagali, V.: Characterizing Indian Ocean manganese nodule-bearing seafloor using multi-beam angular backscatter, *Geo-Mar. Lett.*, 24, 8–13, <https://doi.org/10.1007/s00367-003-0153-y>, 2004.
- Chance, T., Kleiner, A., and Northcutt, J.: The autonomous underwater vehicle (AUV): A cost-effective alternative to deep-towed technology, *Integrated Coastal Zone Management*, 2, 65–69, 2000.
- Che Hasan, R., Ierodiakonou, D., and Monk, J.: Evaluation of Four Supervised Learning Methods for BenthicHabitat Mapping Using Backscatter from Multi-Beam Sonar, *Remote Sens.*, 4, 3427–3443, <https://doi.org/10.3390/rs4113427>, 2012.
- Che Hasan, R., Ierodiakonou, D., Laurenson, L., and Schimel, A.: Integrating Multibeam Backscatter Angular Response, Mosaic and Bathymetry Data for Benthic Habitat Mapping, *PLoS ONE*, 9, e97339, <https://doi.org/10.1371/journal.pone.0097339>, 2014.
- Chen, L., Wang, S., McDonald-Maier K., and Hu, H.: Towards autonomous localization and mapping of AUVs: a survey, *Int. J. Intell. Syst.*, 1, 97–120, <https://doi.org/10.1108/20496421311330047>, 2013.
- Chung, J. S.: Deep-Ocean Mining: Technologies for Manganese Nodules and Crusts, *Int. J. Offshore Polar*, 6, 244–254, 1996.
- Clague, D. A., Dreyer, B. M., Paduan, J. B., Martin, J. F., Caress, D. W., Gill, J. B., Kelley D. S., Thomas, H., Portner, R. A., Delaney, J. R., Guilderson, T. P., and McGann,

- M. L.: Eruptive and tectonic history of the Endeavour Segment, Juan de Fuca Ridge, based on AUV mapping data and lava flow ages, *Geochem. Geophys. Geosys.*, 15, 3364–3391, <https://doi.org/10.1002/2014GC005415>, 2014.
- Clague, D. A., Caress, D. W., Dreyer, B. M., Lundsten, L., Paduan, J. B., Portner, R. A., Spelz-Madero, R., Bowles, J. A., Castillo, P. R., Guardado-France, R., Le Saout, M., Martin, J. F., Santa Rosadel Rio, M. A., and Zierenberg, R. A.: Geology of the Alarcon Rise, southern Gulf of California, *Geochem. Geophys. Geosys.*, 19, 807–837, <https://doi.org/10.1002/2017GC007348>, 2018.
- Clements, A. J., Strong, J. A., Flanagan, C., and Service, M.: Objective stratification and sampling-effort allocation of ground-truthing in benthic-mapping surveys, *ICES J. Mar. Sci.*, 67, 628–637, 2010.
- Cochran, W. G.: *Sampling Techniques*, 3rd edn. Wiley, New York, 1977.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J.: A system for Automated Geoscientific Analyses (SAGA) v. 2.1.4, *Geosci. Model Dev.*, 8, 1991–2007, <https://doi.org/10.5194/gmd-8-1991-2015>, 2015.
- Coulston, J. W., Blinn, C. E., Thomas, V. A., and Wynne, R. H.: Approximating prediction uncertainty for random forest regression models, *Photogramm. Eng. Rem. S.*, 807, 189–197, 2016.
- Craig, J. D.: The relationship between bathymetry and ferromanganese deposits in the north equatorial Pacific, *Mar. Geol.*, 29, 165–186, [https://doi.org/10.1016/0025-3227\(79\)90107-5](https://doi.org/10.1016/0025-3227(79)90107-5), 1979.
- Cutler, D. R., Edwards, T. C., Beard Karen, H., Cutler, A., Hess, K. T., Gibson, J. C., and Lawler, J. J.: Random forests for classification in ecology, *Ecology*, 88, 2783–2792, 2007.
- D’Agostino, R. B., Belanger, A., and D’Agostino Jr., R. B.: A Suggestion for Using Powerful and Informative Tests of Normality, *Am. Stat.*, 44, 316–321, <https://doi.org/10.2307/2684359>, 1990.
- Danson, E.: The Economies of Scale: Using Autonomous Underwater Vehicles (AUVs) for Wide-Area Hydrographic Survey and Ocean Data Acquisition, FIG XXII International Congress Washington, DC, USA, 19–26 April, 2002.
- Das, R. P. and Anand, S.: Metallurgical Processing of Polymetallic Ocean Nodules, in: *Deep-Sea Mining*, edited by: Sharma, R., Resource Springer, [https://doi.org/10.1007/978-3-319-52557-0\\_12](https://doi.org/10.1007/978-3-319-52557-0_12), 2007.
- Dawson, R.: How Significant is a Boxplot Outlier?, *J. Stat. Educat.*, 19, 1–13, <https://doi.org/10.1080/10691898.2011.11889610>, 2011.
- De Moustier, C.: Beyond bathymetry: Mapping acoustic backscattering from the deep seafloor with Sea Beam, *J. Acoust. Soc. Am.*, 79, 316–331, 1986.
- Deschamps, A., Maurice, T., Embley, R. W., and Chadwick, W. W.: Quantitative study of the deformation at Southern Explorer Ridge using high-resolution bathymetric data, *Earth Planet. Sc. Lett.*, 259, 1–17, <https://doi.org/10.1016/j.epsl.2007.04.007>, 2007.
- Diaz-Urriarte, R. and de Andres, A.: Gene selection and classification of microarray data using random forest, *BMC Bioinformatics*, 7, <https://doi.org/10.1186/1471-2105-7-3>, 2006.
- Diesing, M. and Stephens, D.: A multi-model ensemble approach to seabed mapping, *J. Sea Res.*, 100, 62–69, <https://doi.org/10.1016/j.seares.2014.10.013>, 2015.
- Diesing, M., Green, S. L., Stephens, D., Lark, R. M., Stewart, H. A., and Dove, D.: Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches, *Cont. Shelf Res.*, 84, 107–119, <https://doi.org/10.1016/j.csr.2014.05.004>, 2014.
- Dikau, R.: Geomorphic landform modelling based on hierarchy theory, in: *Proceedings of the 4th International Symposium on Spatial Data Handling*, edited by: Brassel, K. and Kishimoto, H., Department of Geography, University of Zürich, Zürich, Switzerland, 230–239, 1990.
- Durden, J. M., Schoening, T., Althaus, F., Friedman, A. Garcia, R., Glover, A. G., Greinert, J., Jacobsen, Stout, N., Jones, D. O. B., Jordt, A., Kaeli, J. W., Koser, K., Kuhn, L. A., Lindsay, D., Morris, K. J., Nattkemper, T. W., Osterloff, J., Ruhl, H. A., Singh, H., Tran, M., and Bett, B. J.: Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding, *Oceanogr. Mar. Biol.*, 54, 1–72, <https://doi.org/10.1201/9781315368597>, 2016.
- Field, A. P.: *Discovering statistics using SPSS: (and sex and drugs and rock “n” roll)*, (OKS Print.) Los Angeles (i.e. Thousand Oaks), California, SAGE Publications, 2009.
- Frazer, J. and Fisk, M. B.: Geological factors related to characteristics of sea-floor manganese nodule deposits, *Deep-Sea Res. Pt. A*, 28, 1533–1551, [https://doi.org/10.1016/0198-0149\(81\)90096-0](https://doi.org/10.1016/0198-0149(81)90096-0), 1981.
- Fu, W. J., Jiang, P. K., Zhou, G. M., and Zhao, K. L.: Using Moran’s I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China, *Biogeosciences*, 11, 2401–2409, <https://doi.org/10.5194/bg-11-2401-2014>, 2014.
- Garzón, M. B., Blazek, R., Neteler, M., Sánchez de Dios, R., Ollero, H. S., and Furlanello, C.: Predicting habitat suitability with machine learning models: The potential area of *Pinus sylvestris* L. in the Iberian Peninsula, *Ecol. Model.*, 197, 383–393, <https://doi.org/10.1016/j.ecolmodel.2006.03.015>, 2006.
- Genuer, R., Poggi J., Tuleau-Malot, C., and Villa-Vialaneix, N.: Random Forests for Big Data, *Big Data Research*, 9, 28–46, <https://doi.org/10.1016/j.bdr.2017.07.003>, 2017.
- GEOMAR: Helmholtz-Zentrum für Ozeanforschung, Autonomous Underwater Vehicle “ABYSS”, *Journal of large-scale research facilities*, 2, A79, 1–5, <https://doi.org/10.17815/jlsrf-2-149>, 2016.
- Ghasemi, A. and Zahediasl, S.: Normality Tests for Statistical Analysis: A Guide for Non-Statisticians, *Int. J. Endocrinol. Metab.*, 10, 486–489, <https://doi.org/10.5812/ijem.3505>, 2012.
- Gilardi, N. and Bengio, S.: Comparison of four machine learning algorithms for spatial data analysis, *Conf. Signals Syst. Comput.*, 17, 160–167, 2009.
- Glasby, G. P.: Distribution of manganese nodules and lebensspuren in underwater photographs from the Carlsberg Ridge, Indian Ocean, New Zealand, *J. Geol. Geophys.*, 16, 1–17, <https://doi.org/10.1080/00288306.1973.10425383>, 1973.
- Glasby, G. P.: Manganese nodules in the South Pacific: A review, New Zealand, *J. Geol. Geophys.*, 19, 707–736, <https://doi.org/10.1080/00288306.1976.10426315>, 1976.
- Goodchild, M. F.: Spatial autocorrelation. Concepts and Techniques in Modern Geography, 47, 1–56, 1986.
- Grasmueck, M., Eberli, G. P., Viggiano, D. A., Correa, T., Rathwell, G., and Luo, J.: Autonomous underwater vehicle (AUV) mapping reveals coral mound distribution, morphology, and oceanog-

- raphy in deep water of the Straits of Florida, *Geophys. Res. Lett.*, 33, L23616, <https://doi.org/10.1029/2006GL027734>, 2006.
- Greinert, J.: Swath sonar multibeam EM122 bathymetry during SONNE cruise SO239 with links to raw data files, PANGAEA, <https://doi.org/10.1594/PANGAEA.859456>, 2016.
- Greinert, J., Schoening, T., Köser, K., and Rothenbeck, M.: Seafloor images and raw context data along AUV tracks during SONNE cruises SO239 and SO242/1, GEOMAR – Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.org/10.1594/PANGAEA.882349>, 2017.
- Haase, K. M., Koschinsky, A., Petersen, S., Devey, C. W., German, C., Lackschewitz, K. S., Melchert, B., Seifert, R., Borowski, C., Giere, O., and Paulick, H.: M64/1, M68/1 and M78/2 Scientific Parties. Diking, young volcanism and diffuse hydrothermal activity on the southern Mid-Atlantic Ridge: the Lilliput field at 9°33′ S, *Mar. Geol.*, 266, 52–64, <https://doi.org/10.1016/j.margeo.2009.07.012>, 2009.
- Hammer, Ø., Harper, D. A. T., and Ryan, P. D.: PAST: Paleontological statistics software package for education and data analysis, *Palaeontol. Electron.*, 4, 9 pp., 2001.
- Hari, V. N., Kalyan, B., Chitre, M., and Ganesan, V.: Spatial Modeling of Deep-Sea Ferromanganese Nodules With Limited Data Using Neural Networks, *IEEE J. Ocean. Engin.*, 43, 1–18, <https://doi.org/10.1109/JOE.2017.2752757>, 2017.
- Herkül, K., Peterson, A., and Paekivi, S.: Applying multibeam sonar and mathematical modeling for mapping seabed substrate and biota of offshore shallows, *Estuarine, Coast. Shelf Sci.*, 192, 57–71, 2017.
- Henthorn, R., Caress, D. W., Thomas, H., McEwen, R., Kirkwood, W. J., Paull, C. K., and Keaten, R.: High-resolution multibeam and subbottom surveys of submarine canyons, deep-sea fan channels, and gas seeps using the MBARI mapping AUV, *Proceedings of the IEEE OCEANS Conference*, 1–6, <https://doi.org/10.1109/OCEANS.2006.307104>, 2006.
- Hoaglin, D. C., Iglewicz, B., and Tukey, J. W.: Performance of Some Resistant Rules for Outlier Labeling, *J. Am. Stat. Assoc.*, 81, 991–999, 1986.
- Horning, N.: RandomForests: An algorithm for image classification and generation of continuous fields data sets, *International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences (GIS-IDEAS)*, 9–11 December, Hanoi, Vietnam, 2010.
- Huang, Z., Justy, S., Scott, L. N., and Brendan, P. B.: Predictive mapping of seabed substrata using high-resolution multibeam sonar data: A case study from a shelf with complex geomorphology, *Mar. Geol.*, 357, 37–52, 2014.
- Ismail, K., Huvenne, V. A. L., and Masson, D. G.: Objective automated classification technique for marine landscape mapping in submarine canyons, *Mar. Geol.*, 362, 17–32, 2015.
- Iwahashi, J. and Pike, R. J.: Automated Classifications of Topography from DEMs by an Unsupervised Nested- Means Algorithm and a Three-Part Geometric Signature, *Geomorphology*, 86, 409–440, 2017.
- Jakobsson, M., Gyllencreutz, R., Mayer, L., Dowdeswell, J. A., Canals, M., Todd, B. J., Dowdeswell, E. K., Hogan, K. A., and Larter, R. D.: Mapping submarine glacial landforms using acoustic methods, *Geol. Soc., London, Memoirs*, 46, 17–40, <https://doi.org/10.1144/M46.182>, 2016.
- Jung, H. S., Ko, Y. T., and Moon, J. W.: Characteristics of Seafloor Morphology and Ferromanganese Nodule Occurrence in the Korea Deep-sea Environmental Study (KODES) Area, NE Equatorial Pacific, *Mar. Georesour. Geotec.*, 19, 167–180, <https://doi.org/10.1080/10641190109353811>, 2001.
- Kavenski, M., Pozdnukhov, A., and Timonin, V.: Machine learning for spatial environmental data. Theory, applications and software, EPFL Press, 1st Edn., <https://doi.org/10.1201/9781439808085>, 2009.
- Kim, H. Y.: Statistical notes for clinical researchers: assessing normal distribution using skewness and kurtosis, *Restorative Dentistry and Endodontics*, 38, 52–54, <https://doi.org/10.5395/rde.2013.38.1.52>, 2013.
- Kleinrock, M. C.: Capabilities of some systems used to Survey the Deep-Sea Floor, Chap. 2, in: *Handbook of geophysical exploration at sea*, edited by: Richard, G. A., 2nd Edn., Hard Minerals, CRC Press, p. 37, 1992a.
- Kleinrock, M. C., Hey, R. N., and Theberger Jr., A. E.: Practical geological comparison of some seafloor survey instruments, *Geophys. Res. Lett.*, 19, 1407–1410, <https://doi.org/10.1029/92GL01390>, 1992b.
- Knobloch, A., Kuhn, T., Rühlemann, C., Hertwig, T., Zeissler, K. O., and Noack, S.: Predictive Mapping of the Nodule Abundance and Mineral Resource Estimation in the Clarion-Clipperton Zone Using Artificial Neural Networks and Classical Geostatistical Methods, in: *Deep-Sea Mining*, edited by: Sharma R., Resource Springer, [https://doi.org/10.1007/978-3-319-52557-0\\_6](https://doi.org/10.1007/978-3-319-52557-0_6), 2017.
- Ko, Y., Lee, S., Kim, J., Kim, K. H., and Jung, M. S.: Relationship between Mn Nodule Abundance and Other Geological Factors in the Northeastern Pacific: Application of GIS and Probability Method, *Ocean Sci. J.*, 41, 149–161, <https://doi.org/10.1007/BF03022420>, 2006.
- Kodagali, V.: Influence of Regional and Local Topography on the Distribution of Polymetallic Nodules in Central Indian Ocean Basin, *Geo-Mar. Lett.*, 8, 173–178, 1988.
- Kodagali, V. and Chakraborty, B.: Multibeam Echosounder PseudoSidescan Images as a tool for Manganese Nodule Exploration, *Proceedings of the Third Ocean Mining Symposium Goa, India*, 8–10 November, 97–104, 1999.
- Kodagali, V. N. and Sudhakar, M.: Manganese nodule distribution in different topographic domains of the Central Indian Basin, *Mar. Georesour. Geotec.*, 11, 293–309, <https://doi.org/10.1080/10641199309379925>, 1993.
- Kuhn, T. and Rathke, M.: Report on visual data acquisition in the field and interpretation for SMnN. Deliverable D1.31 of the EU-Project Blue Mining, BGR Hannover, 34 pp., 2017.
- Kuhn, T., Wiedicke-Hombach, M., Barckhausen, U., Schwarz-Schampera, U., Rutkowski, J., and Lehmann, S.: New Insights of Mn Nodule Exploration from the German License Area in the Pacific Manganese Nodule Belt, *Toward the Sustainable Development of Marine Minerals: Geological, Technological, and Economic Aspects*, 39th Underwater Mining Institute, 4–9 October, Gelendzhik, Russia, 2010.
- Kuhn, T., Rühlemann, C., and Wiedicke-Hombach, M.: Development of Methods and Equipment for the Exploration of Manganese Nodules in the German License Area in the Central Equatorial Pacific, *Proceedings of the ninth ISOPE Ocean Mining Symposium, Maui, Hawaii, USA*, 19–24 June, 174–177, 2011.

- Kuhn, T., Rühlemann C., and Knobloch, A.: Classification of manganese nodule estimates: can we reach the “measured resource” level? Resource and Environmental Assessments for Seafloor Mining Development, 45th Underwater Mining Conference, Incheon, Korea, 9–13 October 2016.
- Kwasnitschka, T., Köser, K., Sticklus, J., Rothenbeck, M., Weiß, T., Wenzlaff, Em., Schoening, T., Triebe, L., Steinführer, An., Devey, C., and Greinert, J.: DeepSurveyCam – A Deep Ocean Optical Mapping System, *Sensors*, 16, 1–17, <https://doi.org/10.3390/s16020164>, 2016.
- Lary, D. J., Alavi, A. H., Gandomi, A. H., and Walker, A. L.: Machine learning in geosciences and remote sensing, *Geosci. Front.*, 7, 3–10, <https://doi.org/10.1016/j.gsf.2015.07.003>, 2016.
- Leempoel, K., Parisod, C., Geiser, C., Dapra, L., Vittoz, P., and Joost, S.: Very high-resolution digital elevation models: are multi-scale derived variables ecologically relevant?, *Methods Ecol. Evol.*, 6, 1373–1383, <https://doi.org/10.1111/2041-210X.12427>, 2015.
- Legendre, P. and Legendre, L.: *Numerical Ecology*, 2nd Edn., Elsevier, Amsterdam, p. 853, 1998.
- Li, J.: Predicting the spatial distribution of seabed gravel content using random forest, spatial interpolation methods and their hybrid methods. 20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December, 394–400, 2013.
- Li, J. and Heap, A. D.: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors, *Ecol. Inform.*, 6, 228–241, <https://doi.org/10.1016/j.ecoinf.2010.12.003>, 2011.
- Li, J. and Heap, A. D.: Spatial interpolation methods applied in the environmental sciences: A review, *Environ. Modell. Softw.*, 53, 173–189, <https://doi.org/10.1016/j.envsoft.2013.12.008>, 2014.
- Li, J., Potter, A., Huang, Z., Daniell, J. J., and Heap, A. D.: Predicting Seabed Mud Content across the Australian Margin: Comparison of Statistical and Mathematical Techniques Using a Simulation Experiment, *Geoscience Australia, Record 2010/11*, 146 pp., 2010.
- Li, J., Heap, A. D., Potter, A., and Daniell, J. J.: Application of machine learning methods to spatial interpolation of environmental variables, *Environ. Modell. Softw.*, 26, 1647–1659, <https://doi.org/10.1016/j.envsoft.2011.07.004>, 2011a.
- Li, J., Heap, A. D., Potter, A., Huang, Z., and Daniell, J. J.: Can we improve the spatial predictions of seabed sediments? A case study of spatial interpolation of mud content across the southwest Australian margin, *Cont. Shelf Res.*, 31, 365–1376, <https://doi.org/10.1016/j.csr.2011.05.015>, 2011b.
- Li, J., Siwabessy, P. J., Maggie, T., Zhi, H., and Andrew, D. H.: Predicting Seabed Hardness Using Random Forest in R, *Data Mining Applications with R*, Elsevier, edited by: Zhao, Y. and Cen, Y., 299–329, <https://doi.org/10.1016/B978-0-12-411511-8.00011-6>, 2013.
- Li, J., Tran, M., and Siwabessy, J.: Selecting Optimal Random Forest Predictive Models: A Case Study on Predicting the Spatial Distribution of Seabed Hardness, *PLoS ONE*, 11, e0149089, <https://doi.org/10.1371/journal.pone.0149089>, 2016.
- Li, J., Alvarez, B., Siwabessy, J., Tran, M., Huang, Z., Przeslawski, L., Radke, L., Howard, F., and Nichol, S.: Application of random forest, generalised linear model and their hybrid methods with geostatistical techniques to count data: Predicting sponge species richness, *Environ. Modell. Softw.*, 97, 112–129, 2017.
- Liaw, A. and Wiener, M.: Classification and regression by random Forest, *R News*, 2, 18–22, 2002.
- Martínez Arbizu, P. and Haeckel, M. (Eds.): *RV SONNE Fahrtbericht/Cruise Report SO239: EcoResponse Assessing the Ecology, Connectivity and Resilience of Polymetallic Nodule Field Systems*, Balboa (Panama) – Manzanillo (Mexico), 11.03.–30.04.2015. GEOMAR Report, N. Ser. 025. GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany, 204 pp., 2015.
- Mascaro, J., Asner, G. P., Knapp, D. E., Kennedy-Bowdoin, T., Martin, R. E., Anderson, C., Higgins, M., and Chadwick, D.: A Tale of Two “Forests”: Random Forest Machine Learning Aids Tropical Forest Carbon Mapping, *PLoS ONE*, 9, e85993, <https://doi.org/10.1371/journal.pone.0085993>, 2014.
- Mayer, L. A.: Frontiers in seafloor mapping and visualization, *Mar. Geophys. Res.*, 27, 7–17, <https://doi.org/10.1007/s11001-005-0267-x>, 2007.
- Meinshausen, N.: Quantile Regression Forests, *J. Mach. Learn. Res.*, 7, 983–999, 2006.
- Meinshausen, N.: *quantregForest 0.2–3*, R package, 2012.
- Mentch, L. and Hooker, G.: Quantifying uncertainty in random forests via confidence intervals and hypothesis tests, *J. Mach. Learn. Res.*, 17, 841–881, 2016.
- Millard, K. and Richardson, M.: On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping, *Remote Sens.*, 7, 8489–8515, <https://doi.org/10.3390/rs70708489>, 2015.
- Miller, B. A., Koszinski, S., Wehrhan, M., and Sommer, M.: Impact of multi-scale predictor selection for modeling soil properties, *Geoderma*, 239/240, 97–106, <https://doi.org/10.1016/j.geoderma.2014.09.018>, 2015.
- Mitchell, M. W.: Bias of the Random Forest Out-of-Bag (OOB) Error for Certain Input Parameters, *Open Journal of Statistics*, 1, 205–211, <https://doi.org/10.4236/ojs.2011.13024>, 2011.
- Moellering, H. and Tobler, W.: Geographical variances, *Geogr. Anal.*, 4, 34–50, <https://doi.org/10.1111/j.1538-4632.1972.tb00455.x>, 1972.
- Moran, P. A. P.: The interpretation of statistical maps, *J. Roy. Stat. Soc. B*, 10, 243–251, 1948.
- Moran, P. A. P.: Notes on Continuous Stochastic Phenomena, *Biometrika*, 37, 17–23, <https://doi.org/10.2307/2332142>, 1950.
- Mucha, J. and Wasilewska-Błaszczczyk, M.: Variability and Accuracy of Polymetallic Nodules Abundance Estimations in the IOM Area – Statistical and Geostatistical Approach, *Proceedings of the Tenth (2013) ISOPE Ocean Mining and Gas Hydrates Symposium, Szczecin, Poland, 22–26 September, 27–31*, 2013.
- Mukaka, M. M.: Statistics corner: A guide to appropriate use of correlation coefficient in medical research, *Malawi Med. J.*, 24, 69–71, 2012.
- Obermeyer, Z. and Emanuel, E. J.: Predicting the Future – Big Data, Machine Learning, and Clinical Medicine, *N. Engl. J. Med.*, 29, 1216–1219, <https://doi.org/10.1056/NEJMp1606181>, 2016.
- Okazaki, M. and Tsune, A.: Exploration of Polymetallic Nodule Using AUV in the Central Equatorial Pacific. *Proceedings of the Tenth ISOPE Ocean Mining and Gas Hydrates Symposium, Szczecin, Poland, 22–26 September*, 2013.
- Okun, O. and Priisalu, H.: Random Forest for Gene Expression Based Cancer Classification: Overlooked Issues, *Lect. Notes*

- Comput. Sc., 4478, 483–490, [https://doi.org/10.1007/978-3-540-72849-8\\_61](https://doi.org/10.1007/978-3-540-72849-8_61), 2007.
- Paduan, B. J., Caress, D. W., Clague D. A., Paull, C. K., and Thomas, H.: High-Resolution Mapping of Mass Wasting, Tectonic, and Volcanic Hazards Using the MBARI Mapping AUV, *Rend. online Soc. Geol. It.*, 7, 181–186, 2009.
- Paul, C. K., Anderson, K., Caress, D. W., Lundsten, E., and Gwiazda, R.: Fine Scale Morphology of Tubeworm Slump, Chap. 15, Monterey Canyon, in: Submarine mass movements and their consequences, 7th International Symposium, edited by: Lamarche, G., Mountjoy, J., Bull, S., Hubble, T., Krastel, S., Lane, E., Micallef, A., Moscardelli, L., Mueller, C., Pecher, I., and Woelz, S.: *Advances in Natural and Technological Hazards Research 41*, Springer, [https://doi.org/10.1007/978-3-319-20979-1\\_15](https://doi.org/10.1007/978-3-319-20979-1_15), 2016.
- Paull, L., Saeedi, S., Seto, M., and Li, H.: AUV Navigation and Localization: A Review, *IEEE J. Ocean. Engin.*, 39, 131–149, <https://doi.org/10.1109/JOE.2013.2278891>, 2014.
- Petersen S., Hannington, M., and Krättschell, A.: Technology developments in the exploration and evaluation of deep-sea mineral resources, *Responsabilité and Environment*, 85, 14–18, 2017.
- Peukert, A., Schoening, T., Alevizos, E., Köser, K., Kwasnitschka, T., and Greinert, J.: Understanding Mn-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and optical data, *Biogeosciences*, 15, 2525–2549, <https://doi.org/10.5194/bg-15-2525-2018>, 2018a.
- Peukert, A., Petersen, S., Greinert, J., and Charlot, F.: Seabed Mining, edited by: Micallef, A., Krastel, S., and Savini, A., *Submarine Geomorphology*, Springer Geology, Springer, Chamber, 481–502, <https://doi.org/10.1007/978-3-319-57852-1>, 2018b.
- Pierdomenico, M., Guida, V. G., Macelloni, L., Chiocci, F. L., Rona, P. A., Scranton, M. I., Asper, V., and Diercks, A.: Sedimentary facies, geomorphic features and habitat distribution at the Hudson Canyon head from AUV multibeam data, *Deep-Sea Res. Pt. II*, 121, 112–125, <https://doi.org/10.1016/j.dsr2.2015.04.016>, 2015.
- Preston, J.: Automated acoustic seabed classification of multibeam images of Stanton Banks, *Appl. Acoust.*, 70, 1277–1287, 2009.
- R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, 2008.
- Rahn, M.: Deposit models. Deliverable D3.11 of the EU-Project Blue Mining, BGR Hannover, 43–47, 2017.
- Riley, S. J., DeGloria, S. D., and Elliot, R.: A terrain ruggedness index that quantifies topographic heterogeneity, *Int. J. Sci.*, 5, 1–4, 1999.
- Roberts, J. J., Best, B. D., Dunn, D. C., Treml, E. A., and Halpin, P. N.: Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++, *Environ. Modell. Softw.*, 25, 1197–1207, <https://doi.org/10.1016/j.envsoft.2010.03.029>, 2015.
- Rodriguez-Galiano, V. F., Chica-Olmo, M., and Chica-Rivas, M.: Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, Southern Spain, *Int. J. Geogr. Informat. Sci.*, 28, 1336–1354, <https://doi.org/10.1080/13658816.2014.885527>, 2014.
- Rodriguez-Galiano, V. F., Sanchez-Castillo, M., Chica-Olmo, M., and Chica-Rivas, M.: Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines, *Ore Geol. Rev.*, 71, 804–818, <https://doi.org/10.1016/j.oregeorev.2015.01.001>, 2015.
- Rogers, J.: Seismic, Bathymetric, and Photographic Evidence of widespread erosion and a manganese-nodule pavement along the continental rise of the southeast Cape Basin, *Mar. Geol.*, 78, 57–76, [https://doi.org/10.1016/0025-3227\(87\)90068-5](https://doi.org/10.1016/0025-3227(87)90068-5), 1987.
- Roman, C. and Mather, R.: Autonomous underwater vehicles as tools for deep-submergence archaeology, *P. I. Mech. Eng. M-J. Eng.*, 224, 327–340, <https://doi.org/10.1243/14750902JEME202>, 2010.
- Rühlemann, C., Kuhn, T., and Wiedicke, M.: Current Status of Manganese Nodule Exploration in the German License Area Proceedings of the Ninth ISOPE Ocean Mining Symposium Maui, Hawaii, USA, 19–24 June, 168–173, 2011.
- Rühlemann, C., Kuhn, T., Vink, A., and Wiedicke, M.: Methods of Manganese Nodule Exploration in the German License Area. Recent Developments in Atlantic Seabed Minerals Exploration and Other Topics, 42nd Underwater Mining Institute, 21–29 October, Rio de Janeiro and Porto de Galinhas, Brazil, 2013.
- Ruß, G. and Kruse, R.: Regression Models for Spatial Data: An Example from Precision Agriculture. CDM 2010. Lecture Notes in Computer Science, Vol. 6171, Springer, Berlin, Heidelberg, [https://doi.org/10.1007/978-3-642-14400-4\\_35](https://doi.org/10.1007/978-3-642-14400-4_35), 2010.
- Santibanez, S., Lakes, T., and Kloft, M.: Performance Analysis of Some Machine Learning Algorithms for Regression Under Varying Spatial Autocorrelation, The 18th AGILE International Conference on Geographic Information Science, Lisboa, Portugal, 9–12 June, 2015a.
- Santibanez, S. F., Kloft, M., and Lakes T.: “Performance Analysis of Machine Learning Algorithms for Regression of Spatial Variables. A Case Study in the Real Estate Industry”, 13th International Conference of GeoComputation, Dallas, USA, 20–23 May, 2015b.
- Schmueli, G.: To Explain or to Predict?, *Stat. Sci.*, 25, 289–310, <https://doi.org/10.1214/10-STS330>, 2010.
- Schoening, T., Kuhn, T., and Nattkemper, T. W.: Estimation of poly-metallic nodule coverage in benthic images, in: Proceedings of the 41st Conference of the Underwater Mining Institute, UMI, 2012a.
- Schoening, T., Bergmann, M., Ontrup, J., Taylor, J., Dannheim, J., Gutt, J., Purser, A., and Nattkemper, T. W.: Semi-Automated Image Analysis for the Assessment of Megafaunal Densities at the Arctic Deep-Sea Observatory HAUSGARTEN, *PLoS ONE*, 7, e38179, <https://doi.org/10.1371/journal.pone.0038179>, 2012b.
- Schoening, T., Kuhn, T., and Nattkemper, T. W.: Seabed classification using a bag-of-prototypes feature representation, in: 2014 ICPR Workshop on Computer Vision for Analysis of Underwater Imagery (CVAUI), IEEE, 17–24, <https://doi.org/10.1109/CVAUI.2014.9>, 2014.
- Schoening, T., Thomas, K., Bergmann, M., and Nattkemper, T. W.: DELPHI – fast and adaptive computational laser point detection and visual footprint quantification for arbitrary underwater image collections, *Front. Mar. Sci.*, 2, 1–6, <https://doi.org/10.3389/fmars.2015.00020>, 2015.
- Schoening, T., Kuhn, T., Jones, D. O. B., Simon-Lledo, E., and Nattkemper, T. W.: Fully automated image segmentation for benthic resource assessment of poly-

- metallic nodules, *Methods Oceanogr.*, 15/16, 78–89, <https://doi.org/10.1016/j.mio.2016.04.002>, 2016.
- Schoening, T., Jones, D. O. B., and Greinert, J.: Compact-Morphology-based polymetallic Nodule Delineation, *Sci. Rep.*, 7, 13338, <https://doi.org/10.1038/s41598-017-13335-x>, 2017a.
- Schoening, T.: Source code for the Compact Morphology-based Nodule Delineation (CoMoNoD) algorithm, PANGAEA, <https://doi.org/10.1594/PANGAEA.875070>, Supplement to: Schoening T., Jones D. O. B., Greinert, J.: Compact-Morphology-based poly-metallic Nodule Delineation, *Sci. Rep.*, 7, 1–12, <https://doi.org/10.1038/s41598-017-13335-x>, 2017b.
- Schoening, T.: Results of nodule detection along AUV tracks during SONNE cruises SO239 and SO242/1, PANGAEA, <https://doi.org/10.1594/PANGAEA.883838>, 2017c.
- Sharma, R.: Quantitative estimation of seafloor features from photographs and their application to nodule, *Mar. Georesour. Geotec.*, 11, 311–331, <https://doi.org/10.1080/10641199309379926>, 1993.
- Sharma, R. and Kodagali, V.: Influence of seabed topography on the distribution of manganese nodules and associated features in the Central Indian Basin: A study based on photographic observations, *Mar. Geol.*, 110, 153–162, [https://doi.org/10.1016/0025-3227\(93\)90111-8](https://doi.org/10.1016/0025-3227(93)90111-8), 1993.
- Sharma, R. Sankar, S. J., Samanta, S., Sardar, A. A., and Gracious, D.: Image analysis of seafloor photographs for estimation of deep-sea minerals, *Geo-Mar. Lett.*, 30, 617–626, <https://doi.org/10.1007/s00367-010-0205-z>, 2010.
- Sharma, R., Khade, N. H., and Sankar, S. J.: Assessing the distribution and abundance of seabed minerals from seafloor photography data in the Central Indian Ocean Basin, *Int. J. Remote Sens.*, 34, 1691–1706, <https://doi.org/10.1080/01431161.2012.725485>, 2013.
- Sibenac, M., Podder, T., Kirkwood, W., and Thomas, H.: Autonomous Underwater Vehicles for Ocean Research: Current Needs and State of the Art Technologies, *Mar. Technol. Soc. J.*, 38, 63–72, <https://doi.org/10.4031/002533204787522848>, 2004.
- Skornyakova, N. and Murdmaa, I.: Local variations in distribution and composition of ferromanganese nodules in the Clarion-Clipperton Nodule Province, *Mar. Geol.*, 103, 381–405, [https://doi.org/10.1016/0025-3227\(92\)90028-G](https://doi.org/10.1016/0025-3227(92)90028-G), 1992.
- SPC: Deep Sea Minerals: Manganese Nodules, a physical, biological, environmental, and technical review, edited by: Baker, E. and Beaudoin, Y., Vol. 1B, Secretariat of the Pacific Community, 2013.
- Stephens, D. and Diesing, M.: A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data, *PLoS ONE*, 9, e93950, <https://doi.org/10.1371/journal.pone.0093950>, 2014.
- Strobl, C. and Zeileis, A.: Danger: High Power! – Exploring the Statistical Properties of a Test for Random Forest Variable Importance, *Proceedings of the 18th International Conference on Computational Statistics*, Porto, Portugal, 2008.
- Strobl, C. Boulesteix, A. L., Zeileis, A., and Hothorn, T.: Bias in random forest variable importance measures: Illustrations, sources and a solution, *BMC Bioinformatics*, 8, <https://doi.org/10.1186/1471-2105-8-25>, 2007.
- Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., and Zeileis, A.: Conditional variable importance for random forests, *BMC Bioinformatics*, 9, 307, <https://doi.org/10.1186/1471-2105-9-307>, 2008.
- Strobl, C., Malley, J., and Tutz, G.: An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests, *Psychol. Methods*, 14, 323–48, 2009.
- Tsune, A. and Okazaki, M.: Some Considerations about Image Analysis of Seafloor Photographs for Better Estimation of Parameters of Polymetallic Nodule Distribution, *Proceedings of the Twenty-fourth (2014) International Ocean and Polar Engineering Conference*, Busan, Korea, 15–20 June, 72–77, 2014.
- Tung, N. T., Huang, J. Z., Khan, I., Li, M. J., and Williams, G.: Extensions to Quantile Regression Forests for Very High-Dimensional Data, in: *Advances in Knowledge Discovery and Data Mining*, edited by: Tseng, V. S., Ho, T. B., Zhou, Z. H., Chen, A. L. P., and Kao, H. Y., *PAKDD 2014, Lecture Notes in Computer Science*, Vol. 8444, Springer, Cham [https://doi.org/10.1007/978-3-319-06605-9\\_21](https://doi.org/10.1007/978-3-319-06605-9_21), 2014.
- UNOET: Delineation of mine sites and potential in different sea areas, Vol. 9, *Seabed Minerals Series 4*, Graham and Trotman, London, 1987.
- van der Ploeg, T., Austin, P. C., and Steyerberg, E. W.: Modern modelling techniques are data hungry: a simulation study for predicting dichotomous end points, *Medical Research Methodology*, 14, 1–13, <https://doi.org/10.1186/1471-2288-14-137>, 2014.
- Volkman, S. E.: Concept for Sustainable Economic Evaluation. Deliverable D1.41 of the EU-Project Blue Mining, BGR Hannover, 43–47, 2017.
- Volkman, S. E. and Lehnen, F.: Production key figures for planning the mining of manganese nodules, *Mar. Georesour. Geotec.*, 36, 360–375, <https://doi.org/10.1080/1064119X.2017.1319448>, 2018.
- von Stackelberg, U. and Beiersdorf, H.: The formation of manganese nodules between the Clarion and Clipperton fracture zones southeast of Hawaii, *Mar. Geol.*, 98, 411–423, [https://doi.org/10.1016/0025-3227\(91\)90113-I](https://doi.org/10.1016/0025-3227(91)90113-I), 1991.
- Wager, S., Hastie, T., and Efron, B.: Confidence intervals for random forests: the jackknife and the infinitesimal jackknife, *J. Mach. Learn. Res.*, 15, 1625–1651, 2014.
- Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Clim. Res.*, 30, 79–82, <https://doi.org/10.3354/cr030079>, 2005.
- Wilson, M. F. J., O’Connell, B., Brown, C., Guinan J. C., and Grehan, A. J.: Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope, *Mar. Geodesy*, 30, 3–35, <https://doi.org/10.1080/01490410701295962>, 2007.
- Wynn, R. B., Huvenne, V. A. I., Le Bas, T. P., Murton, B. J., Connelly, D. P., Bett, B. J., Ruhl, H. A., Morris, K. J., Peakall, J., Parsons, D. R., Sunner, S. J., Darber S. E., Dorrell, R. M., and Hunt, J. E.: Autonomous Underwater Vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience, *Mar. Geol.*, 352, 451–468, <https://doi.org/10.1016/j.margeo.2014.03.012>, 2014.
- Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., and White, L.: Performance of nonparametric algorithms for spatial mapping of tropical forest structure, *Carbon Balance Manage*, 11, 1–14, <https://doi.org/10.1186/s13021-016-0062-9>, 2016.

Yazici, B. and Yolacan, S.: A comparison of various tests of normality, *J. Stat. Comput. Simul.*, 77, 175–183, <https://doi.org/10.1080/10629360600678310>, 2007.

Zevenbergen, L. W. and Thorne, C. R.: Quantitative analysis of land surface topography, *Earth Surf. Proc. Land.*, 12, 47–56, <https://doi.org/10.1002/esp.3290120107>, 1987.

Article

# Importance of Spatial Autocorrelation in Machine Learning Modeling of Polymetallic Nodules, Model Uncertainty and Transferability at Local Scale

Iason-Zois Gazis <sup>1,\*</sup>  and Jens Greinert <sup>1,2</sup> 

<sup>1</sup> GEOMAR Helmholtz Centre for Ocean Research Kiel, Wischhofstraße 1-3, 24148 Kiel, Germany; jgreinert@geomar.de

<sup>2</sup> Institute of Geosciences, Christian-Albrechts University Kiel, Ludewig-Meyn-Straße 10-12, 24098 Kiel, Germany

\* Correspondence: igazis@geomar.de

**Abstract:** Machine learning spatial modeling is used for mapping the distribution of deep-sea polymetallic nodules (PMN). However, the presence and influence of spatial autocorrelation (SAC) have not been extensively studied. SAC can provide information regarding the variable selection before modeling, and it results in erroneous validation performance when ignored. ML models are also problematic when applied in areas far away from the initial training locations, especially if the (new) area to be predicted covers another feature space. Here, we study the spatial distribution of PMN in a geomorphologically heterogeneous area of the Peru Basin, where SAC of PMN exists. The local Moran's I analysis showed that there are areas with a significantly higher or lower number of PMN, associated with different backscatter values, aspect orientation, and seafloor geomorphological characteristics. A quantile regression forests (QRF) model is used using three cross-validation (CV) techniques (random-, spatial-, and cluster-blocking). We used the recently proposed "Area of Applicability" method to quantify the geographical areas where feature space extrapolation occurs. The results show that QRF predicts well in morphologically similar areas, with spatial block cross-validation being the least unbiased method. Conversely, random-CV overestimates the prediction performance. Under new conditions, the model transferability is reduced even on local scales, highlighting the need for spatial model-based dissimilarity analysis and transferability assessment in new areas.



**Citation:** Gazis, I.-Z.; Greinert, J. Importance of Spatial Autocorrelation in Machine Learning Modeling of Polymetallic Nodules, Model Uncertainty and Transferability at Local Scale. *Minerals* **2021**, *11*, 1172. <https://doi.org/10.3390/min11111172>

Academic Editors: Pedro Madureira and Tomasz Abramowski

Received: 31 August 2021

Accepted: 12 October 2021

Published: 22 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** polymetallic nodules; spatial autocorrelation; cross-validation; model transferability

## 1. Introduction

The spatial distribution of deep-sea polymetallic nodules (PMN) is currently of high interest due to their high metal content of Mn, Fe, Ni, Co, Cu, or Li. These metals are needed for green and decarbonized technologies, such as electric cars and wind turbines [1,2]. The European Union alone will need 60 times more lithium and 15 times more cobalt by 2050 than today [3] for this transition. Deep-sea resources such as PMN could support this transition, with predictive spatial mapping having a key role in mining-block prioritization.

PMN spatial mapping advanced by using autonomous underwater vehicles (AUVs) to acquire large volumes of hydroacoustic and image data allows for high-resolution seafloor reconstruction at meter and even down to centimeter scales. The quantitative analysis of images enlightens the PMN distribution, narrowing the spatial gap that arises from the sparse ground-truth box-corer samples (usually > 1.8 km) with very limited sampling area (0.25 m<sup>2</sup>) [4–9].

Multibeam echosounder (MBES) data (bathymetry and backscatter), seafloor lithology, environmental information (e.g., organic carbon content), and ground-truth information (photos, information from box-corer sampling) have been analyzed with machine learning

(ML) methods, providing the spatial distribution of PMN [7,10–15]. However, the influence of spatial autocorrelation (SAC) on ML modeling has not been considered.

Spatial autocorrelation describes the phenomenon where local similarity (neighboring observations) is matched by value similarity (correlation between observations) [16]. When the examined variable is spatially autocorrelated, the assumption of independence among the observations in the cross-validation (CV) data is violated (i.e., the fitted model uses almost identical data for training and testing).

Recent studies in terrestrial and marine spatial ML modeling showed that if the commonly used, non-spatial, random k-fold CV is used, the prediction performance is over-optimistic when SAC exists in the data. The magnitude of the spatial overfitting varies based on the degree of SAC among the training points, the environmental similarity among the regions, and the ML method used [17–25]. In order to address the influence of SAC, different CV schemes have been proposed; the most common are buffer distances among training locations [18,26,27], successive distances [28], leaving one training location out [19,29], and spatial block training. Spatial blocks can be based on geographical coordinates clustering [30], latitudinal-blocking [23], systematic assignment, and environmental similarity clustering [17,23,31]. SAC also influences the ML feature selection methods and hyperparameter optimization, resulting in suboptimal variable selection and model parameters [19,25,29,32].

Applying ML models can also result in poor and unreliable predictions when they extrapolate in a new geographical area where the feature space varies [17,33–35]. Thus, there is a need for dissimilarity analysis between the training data and the new area the prediction should be performed for. Traditional non-spatial approaches, such as density plots and boxplots, can show the overall degree of differentiation but cannot identify where it occurs. Moreover, examining and interpreting such plots in multidimensional space is difficult and could lead to erroneous conclusions [36]. Spatial sample-based methods have been developed, relying on the univariate distribution range of each predictor and the new correlations among the predictors within their univariate distribution range [37–40]. Model-based methods have also been introduced, using the model weights and prediction error in the dissimilarity calculation [17,35,41].

This paper addresses the presence of SAC in the PMN distribution of a specific site and how this could influence the results in the various modelling steps of the ML workflow by studying:

(a) The use of SAC as source information for the feature selection before modeling. The proposed workflow uses the spatial clusters resulting from the local indicators of spatial association (LISA) and specifically from the local Moran's I [16] and investigates their relations with the seafloor predictors using boxplots and the non-parametric Wilcoxon–Mann–Whitney test [42–44]. The Boruta algorithm [45] is used as an alternative automated ML method. Boruta is an all-relevant feature selection approach that has shown good performance in high and low-dimensional datasets [46–48]. It has been widely used in ML seafloor mapping studies, providing increased interpretability and prediction performance [48–53]. However, its performance under the presence of SAC has not been extensively investigated.

(b) The influence of SAC on the cross-validation workflow steps. Three techniques are studied: random k-fold CV, systematic spatial block, and feature space clustering. The latter two techniques were selected as they can highlight the biases due to SAC and environmental dissimilarity [17]. They are appropriate for big spatial data, as they require less computation time compared with other methods such as buffer distances [18].

(c) The random forests (RF) [54] spatial predictive uncertainty. The prediction uncertainty is an integral part of spatial modeling, as it provides an in-depth analysis of the prediction validity. It can also prioritize areas for future sampling; this is of interest particularly in deep-sea research, where available ground-truth data are typically scarce. The RF prediction uncertainty can be estimated with different methods [18]. Here, the quantile regression forests (QRF) [55] was selected, as studies showed that it could outperform

other methods while it produces informative maps without the need for extensive data preprocessing and model assumptions [18,56–59].

(d) The recently proposed method of dissimilarity index and area of applicability (AOA) [35]. AOA can identify geographical areas with novel feature space conditions that could hinder a reliable model transfer.

To our current knowledge, this is the first time that these topics are investigated for a regression random forests model applied for predicting PMN spatial distribution. Moreover, the literature research yielded only two studies in the marine environment (habitat mapping) that consider and include spatial-CV or/and the AOA, highlighting the need for further investigation. Similarly, QRF has been applied only rarely for seafloor spatial predictive modeling, despite the popularity of the RF algorithm in general [60].

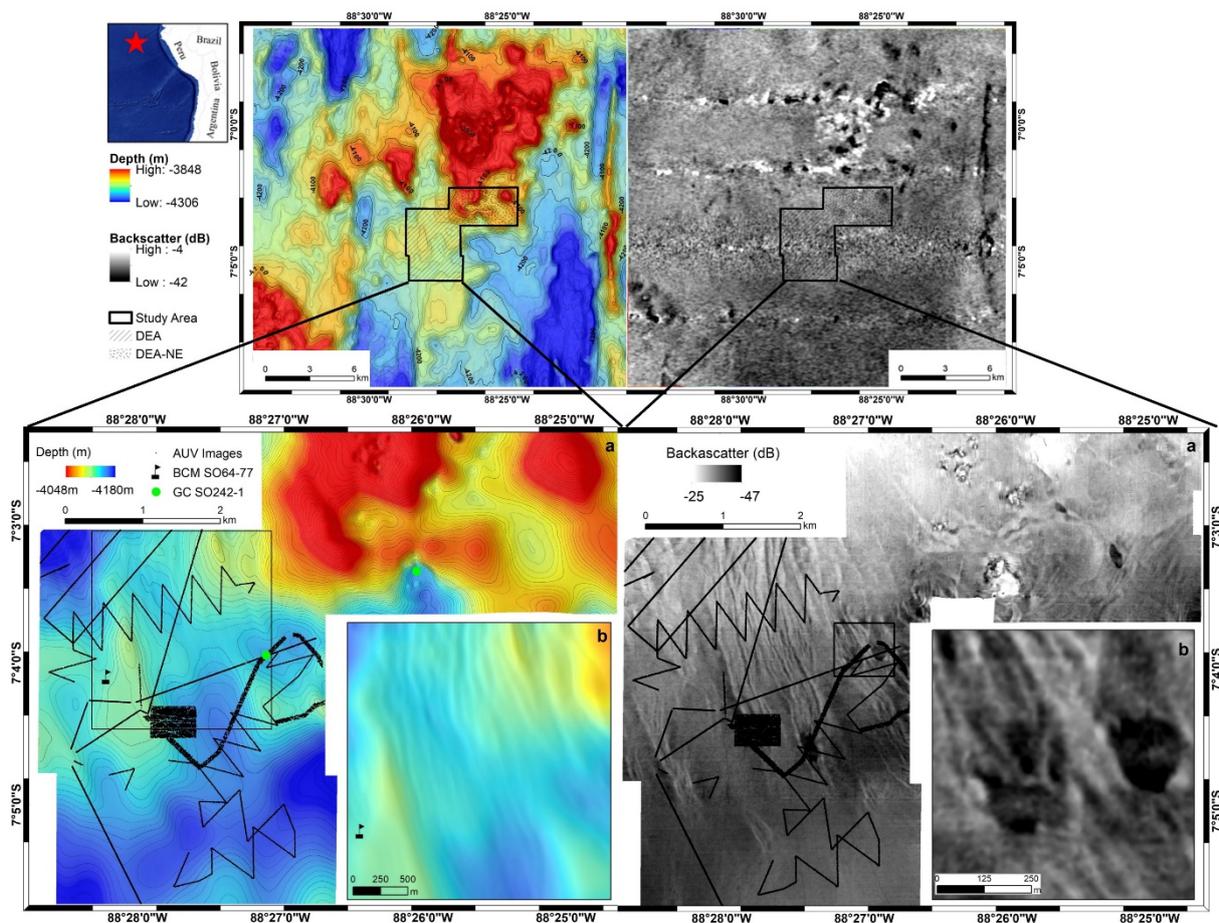
## 2. Materials and Methods

### 2.1. Study Area

The Peru Basin is one of the largest PMN fields globally, with an average abundance of 10 kg/m<sup>2</sup> [61]. Although the abundance and metal grade are of economic value [1], only a few studies have examined the PMN spatial distribution for economic reasons, focusing on the northern part of the investigated area that exhibits a substantial spatial variance [62–65]. In this northern part, lies the “DISturbance and reCOLonization” (DISCOL) experimental area (DEA). Inside the DEA (Figure 1), disturbance experiments were conducted to assess the environmental impact produced by a plough harrow [66,67]. The DEA is spatially heterogeneous regarding the seafloor morphology, geochemical properties, and PMN distribution, with many authors highlighting the need for further research [62–64,68,69]. The mapped area extends north and south of the DEA (in the N–S direction) and towards the northeast (hereafter DEA-NE). The entire region lies between –4047 m and –4179 m water depth, which is slightly above the regional carbonate compensation depth (CCD) at –4250 m [63]. The DEA itself has low relief, gentle slopes (<3°), longitudinal abyssal furrows, and areas with low reflectivity backscatter values and no PMN, hereafter called black patches (Figure 1). The sediments are layered clayey silts and silty clays, with foraminiferous residues and shell fragments [67,70].

The observed abyssal furrows strike perpendicularly to the regional contours and have a U-shape form. They are oriented parallel to the predominant NW bottom current flow [71,72]. Long-term studies showed that this deep flow is not stable but has periods of quasi-unidirectional strong currents (>5 cm/s but sometimes up to 17 cm/s) and periods with slower omnidirectional currents (1–3 cm/s). The recorded bottom current velocities are inside the velocity range that could preserve abyssal furrows [73]. The abyssal furrows act as bottom current channels, occasionally eroded during short periods of higher current velocities or even being relics from an earlier basin period, when the current regime was stronger [73–75]. Their preservation is supported by the low regional sedimentation rates of 0.4–2.0 cm/ka [76]. Past erosional bottom currents have also been proposed to be responsible for the formation of the black patches; these show Plio-Miocene carbonate-rich sediments as infilling [63,68,72].

The PMN-free black patches are easily recognizable due to their low backscatter reflectivity (Figure 1). They have an ellipsoid shape, with their long axes oriented downslope. Their depth varies between 2–5 m with remarkable flat and horizontal seafloors [67]. This downslope direction has been observed in other low reflectivity PMN-free patches in Peru Basin, connected with downslope sediment transport [68].



**Figure 1.** Top left: bathymetric map of the surveyed area in the northern Peru Basin (Southeast Pacific) during research cruise SO242-1 [77]. The area is characterized by N–S orthogonal striking graben and horst structures, sea mountains, knolls and hills with steep slopes, local depressions, and pit structures [67]. Top right: ship-based MBES backscatter mosaic map of the same area. The AUV mapped DEA and DEA-NE areas are indicated as a shaded black polygon. Bottom left: (a) DEA and DEA-NE AUV-based bathymetry with photo locations crisscrossing the area (black points). The contours have a 2 m depth interval. GC: gravity corer [69], BCM: bottom current measurements [71]. Map (b) is an enlarged map showing the abyssal furrows. The furrow height is up to 2 m. Bottom right: (a) DEA and DEA-NE AUV-based backscatter map. The northern part has a higher backscatter intensity than the central and southern parts because of the higher number of PMN. Backscatter alternations are also created by the presence of abyssal furrows; (b) two of the three low reflectivity black patches (PMN-free) that can be identified in the DEA.

Larger-scale PMN-free areas with low backscatter reflectivity have also been observed for the Clarion-Clipperton Zone [15]. A recent gravity core inside a black patch in DEA (Figure 1) revealed that its geochemical conditions differ significantly from the surroundings, having increased organic carbon content (0.5 wt%–0.8 wt%) [69]. The increased organic carbon content might have shifted the Mn-redox closer to the seafloor, where the diagenetically mobilized Mn is released into the bottom water and thus not supporting the PMN formation [68,69,78,79]. Photos inside the black patch confirmed the absence of surficial PMN (Figure A1).

In the DEA-NE area, the seafloor is heterogeneous, with elongated and conical knolls and hills separated by local depressions. Visual inspection showed that the slopes have an extremely thin sediment coverage, while the local depressions act as sediment depocenters with increased sediment accumulation and talus debris [77]. The debris of basaltic fragments could act as nuclei supply to form new PMN [63,64]. PMN surficial coverage in this area varies, with black patches also present (Figure 1). The high-reflectivity sub-areas are mainly due to an exposed basement, such as the two 10 m high outcropping volcanic cones

of pillow basalt inside a crate-shaped structure (Figures 1 and A2). The visual inspection revealed the absence of sediments in their current-exposed, steep sides. PMN were found in the base of the crater, mixed with talus debris [67,77]. Geochemical analyses of the GC sediments revealed depth-profile variations for several chemical compounds, which exceed the area variability of other stations, highlighting the need for further research [69].

## 2.2. Hydroacoustic Data

High-resolution MBES data were acquired with the AUV Abyss from GEOMAR [80,81]. Bathymetric data processing was performed with the QPS Qimera 1.7 and MB-System 5.7 software (Monterey Bay Aquarium Research Institute (MBARI) University of New Hampshire and MARUM, Handelsweg, The Netherlands) [82]. The backscatter processing was carried out with QPS FMGT 7. The finally generated GeoTIFF grids have a 3 m × 3 m cell size, projected in Universal Transverse Mercator (UTM) zone 16S. Parts of the MBES data have been presented already by [67,83], but here they have been reprocessed and merged into one unified dataset.

## Seafloor Geomorphological Analysis

Sixteen derivatives of the bathymetric data were computed (Table 1), following the recommendations of the current literature in the field of quantitatively geomorphometric analysis of seafloor data [50,84]. Derivatives were calculated based on a 10-cell (30 m) neighborhood, which is the minimum defined size for some of the derivatives (e.g., vector ruggedness measure) [85]. However, there are bathymetric derivatives that are calculated from a 3 × 3 cell neighborhood (e.g., slope, aspect). For those derivatives, the mean bathymetry was calculated first, and afterwards the derivative was determined. This approach shows better results than grid resampling or derivative averaging [86]. The 30 m neighborhood relates to the AUV positioning uncertainty, ensuring that the correct seafloor derivatives values will be extracted for each photo location. In this respect, a spatial autocorrelation of the predictors can further reduce the impact of positional uncertainty on the ML accuracy [87,88]. Moreover, the spatial autocorrelation could eliminate existing MBES artifacts that could affect the predictive modeling [52,89–91]. A multiscale approach was applied for the bathymetric position index, with finer and broader scales to better depict seafloor heterogeneity (0–30 m, 30–100 m, and 100–300 m). The aspect was transformed to northness and eastness [92–94]. Abbreviations and references to the algorithms of the calculated 14 derivatives are given in Table 1.

**Table 1.** MBES derivatives, their abbreviations, and references to their calculation algorithms.

MBES Derivatives	Abbreviation	Algorithm
Mean Depth	MD	Focal statistics <sup>*1</sup>
Deviation from Mean Depth	DFMD	Focal statistics <sup>*1</sup>
Slope	S	Zevenbergen and Thorne, 1987 <sup>*1</sup> [95]
Northness	N	Olaya, 2009 <sup>*2</sup> [96]
Eastness	E	Olaya, 2009 <sup>*2</sup>
Profile Curvature	PrC	Zevenbergen and Thorne, 1987 <sup>*1</sup>
Plan Curvature	PIC	Zevenbergen and Thorne, 1987 <sup>*1</sup>
Terrain Surface Convexity	TSC	Iwahashi and Pike, 2006 <sup>*1</sup> [85]
Vector Ruggedness Measure	VRM	Sappington et al., 2007 <sup>*1</sup> [97]
Bathymetric Position Index	BPI	Weiss, 2000 [98], Wilson et al., 2007 <sup>*1</sup> [99]
Backscatter	BS	Focal statistics <sup>*1</sup>
Backscatter SD	BSSD	Focal statistics <sup>*1</sup>
Backscatter Local Moran	BSLM	Anselin, 1995 <sup>*3</sup> [16]
Backscatter Entropy	BSE	Haralick et al. 1973 <sup>*4</sup> [100]

<sup>\*1</sup> SAGA GIS [101], <sup>\*2</sup> Benthic Terrain Modeler [102], <sup>\*3</sup> raster package [103], <sup>\*4</sup> glcm package [104].

### 2.3. Optic Data

High-resolution photos were acquired by the deep survey camera system onboard the AUV Abyss [105]. The compact morphology-based nodule delineation (CoMoNoD) algorithm [9] was used for automated image analysis, providing the number of PMNs/m<sup>2</sup>. CoMoNoD has been used for quantitative assessment and predictive modeling already [7,8], in which the advantages and limitations of the algorithm have been discussed.

We need to highlight that the primary goal of SO242-1 was to re-map in high-resolution the seafloor (acoustically and visually) inside and outside the DEA, which had been ploughed in 1989 [66,67], to have data that can provide insight into the current state of the environmental status and change when compared to previous data acquired between 1989 and 1996 [67,77]. The crisscrossing AUV photo surveys were baseline exploration surveys about the general PMN occurrence and faunal distribution and were not meant to be a proper resource estimation survey. Generally, the optic data underestimate the number and abundance of PMN, and local correction factors (based on box-corer data from the photo locations) must be applied for a more realistic resource assessment [14,15,106–109].

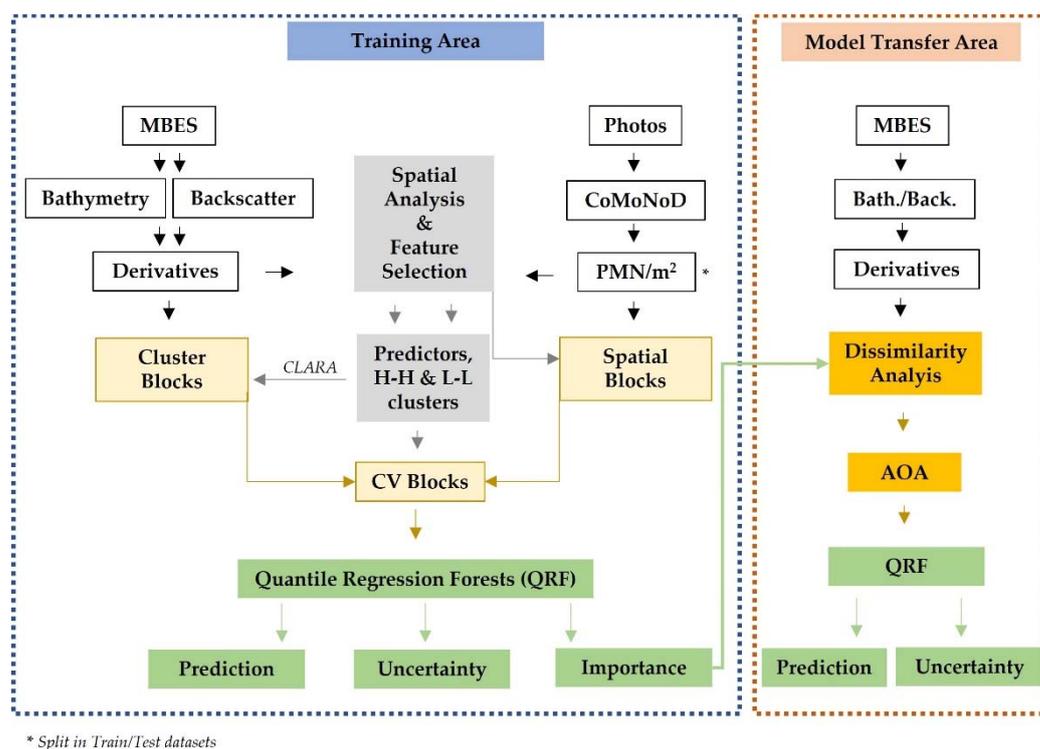
For compiling the ground-truth dataset, all photos inside and next to the plough disturbance tracks were excluded, as they do not represent the original seafloor state [67]. The degree of blanketing around the tracks varies [67]. Nevertheless, the PMN abundance estimation in the area of the full coverage photomosaic (Figure 1) showed that the PMN could be effectively quantified, while it revealed the first signs of a correlation between PMN occurrence and seafloor morphology [9].

The optic data sampling in general has good geographical coverage (Figure 1), which is vital to efficiently depict the PMN spatial distribution trend, especially in local-scale studies such as ours [7,8,110]. The correlation points towards an underlying relationship between PMN and seafloor morphology that, synergistically with other environmental factors, influenced the PMN genesis and current spatial distribution. This is true, although the observed PMN number only represents the minimum number of nodules, as another part might occur within the sediment or due to a sediment cover that was not detected by the CoMoNoD algorithm.

In total, analyses of 30,000 photos were considered reliable and exported as ESRI shapefile in UTM 16S for further spatial analysis steps. The exported dataset was split into two independent datasets, creating the train and test dataset with 20,000 and 10,000 photos, respectively. The random data split without replacement was performed using the Subset Features tool (Geostatistical Analyst) in ArcMap 10.6 (© Environmental Systems Research Institute Inc. (ESRI), West Redlands, CA, USA); this tool ensures the same geographical coverage and statistical characteristics of the two generated datasets (Figure A3). The training dataset was used for the spatial autocorrelation analysis, feature selection, and spatial modeling, and the test dataset was used only for the final model evaluation. An overview of the different processing steps of the presented modelling workflow is given in Figure 2.

### 2.4. Spatial Data Analysis

The presence of spatial autocorrelation was investigated using the local indicators of spatial association (LISA) and particularly the local Moran's I [16]. The local Moran's I identifies clusters with significant spatial aggregation of similar high (H-H) or low (L-L) values (i.e., many or few PMN). The index was calculated using the Cluster and Outlier Analysis (Anselin Local Moran's I) Tool (Spatial Statistics) in the ArcMap 10.6, according to the equations and null hypothesis provided as default from the software (i.e., the examined attribute is randomly distributed). The spatial relationship was based on the inverse Euclidian distance, and spatial weights were standardized to eliminate any bias that could be induced due to the uneven number of spatial neighbors. The analysis was based on 999 permutations. A false discovery rate correction was applied as the recommended approach to deal with the multiple testing and spatial dependency biases in large datasets; it provides significant clusters and outliers for a 95 per cent confidence level [111,112].



**Figure 2.** The general analyses and modeling workflow with its different processing steps. For description, see text.

A visual boxplot analysis was performed between the two spatial clusters and their underlying MBES derivative values. In addition, the non-parametric Wilcoxon–Mann–Whitney [42–44] was used to identify significant differences in the MBES derivatives between the two spatial clusters. Due to the large sample size (inherent sampling variability), the substantive significance (effect size) was additionally used together with the statistical significance ( $p$ -value) for the interpretation of the results, when the  $p$ -values are similar or the same [113]. The base [114] and rstatix R packages were used here [115].

### 2.5. Feature Selection

The results from the spatial autocorrelation analysis were used jointly with the Boruta analysis for the feature selection. The Boruta algorithm creates a shuffled copy of each predictor variable and calculates the average variable importance using the ranger-RF algorithm [45]. The variables with higher importance than their shuffled copies are considered relevant to the target variable (PMN). Since all the relevant variables have been identified, Spearman’s rank correlation coefficient [116] was used to exclude correlated features with coefficient values  $> 0.5$ . Between two correlating features, the one with the higher relevance was kept. The Spearman correlation was preferred instead of, e.g., the Pearson correlation, as it is a non-parametric measure, which assesses the presence of monotonic relationships while being robust to outliers and deviations from normality [117,118]. For executing the features selection, the Boruta [45] and GGally [119] R packages were used.

### 2.6. Quantile Regression Forests

Random forests is an ensemble machine learning method designed of multiple classification or regression trees [54]. Each tree uses a random subset of the training data through bootstrapping. The remaining data (out-of-bag data) are used for internal error validation. Each tree node is split using the best subset of predictors randomly chosen, minimizing the correlation among trees. A tree is developed until the maximum depth is reached; the final predicted value of the regression results after averaging all trees predictions is finished. The QRF extends the random forests approach by keeping all the predicted values for each

observation. This information assesses the conditional distribution and, thus, quantiles can be estimated. The range between the maximum and minimum quantile for a single prediction expresses the model uncertainty for this prediction. Here, the 0.05th and 0.95th quantiles were used for the lower and upper prediction uncertainty value.

RF performs well with the recommended default hyperparameter values (e.g., minimal node size, maximal tree depth) [7,48,120]. Nevertheless, the default caret [121] ranger-RF optimization process was applied, focusing on the number of variables available for splitting (mtry) and the splitting criterion. The permutation variable importance and the partial dependence plots (PDP) were also calculated using the pdp [122] R package. In a subpart of this study area, previous RF modelling showed an overall good prediction performance based on the internal OOB data [83].

## 2.7. Cross-Validation Techniques

Three different cross-validation techniques have been tested.

### 2.7.1. Random k-fold Cross-Validation

Here, the model is repeatedly trained through random 10-fold CV, and its prediction performance is evaluated on the left-out fold data (k-1). Ten folds are recommended for large datasets, as they provide a good bias-variance trade-off [123].

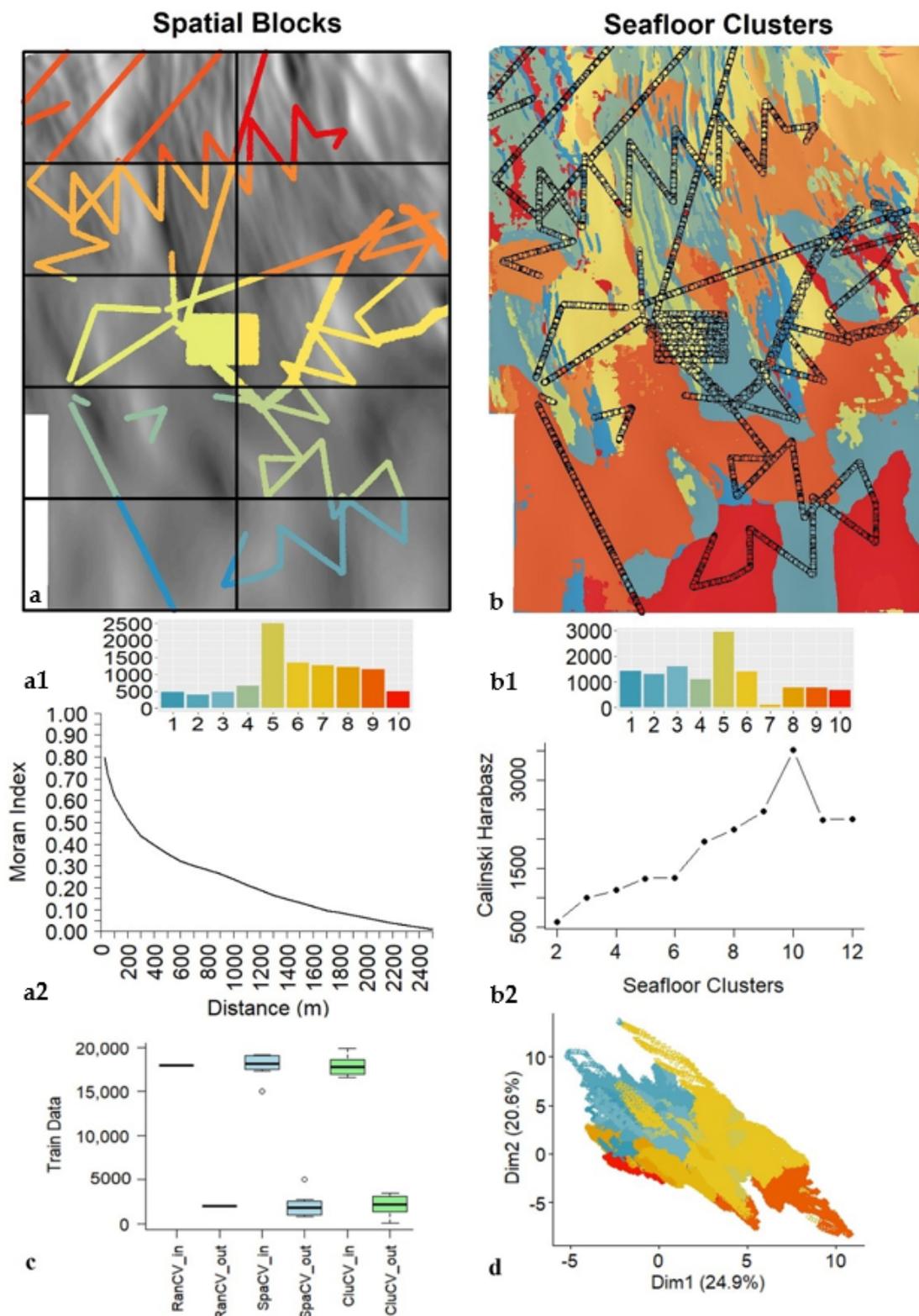
### 2.7.2. Systematic k-fold Spatial-Blocking Cross-Validation

Ten spatial non-overlapping folds (sub-areas) with equal geographical coverage (2 km × 1 km) were created using ArcMap 10.6 [Create Fishnet Tool (Data Management)] (Figure 3). The size of the block is a trade-off between the spatial autocorrelation range and the need for extrapolation [17,22,23,31]; it should minimize the influence of spatial autocorrelation in the training locations without creating extensive feature space differentiation among the blocks [17]. The Moran's I incremental analysis [Incremental Spatial Autocorrelation Tool (Spatial Statistics)] showed that the spatial autocorrelation quickly drops after the 1st km (<0.25) and approaches almost zero (<0.05) from 2 km onward (Figure 3).

### 2.7.3. Feature Space k-fold Clustering Cross-Validation

As third method, the clustering large applications (CLARA) algorithm was used. CLARA [124] is a fast implementation of the partitioning around medoids [125] algorithm designed for large datasets. It uses an actual data point (medoid) as the center of each class, in which the sum of pairwise dissimilarities in this cluster is minimal; this method is robust to outliers. Only the predictors that resulted from the feature selection were clustered (Figure 3). The optimal number of clusters was based on the Calinski-Harabasz index [126]. The clustering was performed with the R packages cluster [127], clusterCrit [128], and RStoolbox [129].

It needs to be mentioned that both the random and the spatial k-fold CV can only be applied inside the DEA, where AUV footage exists and an objective comparison between the three CV techniques is possible. Nevertheless, it provides the opportunity for transferring the model to a different neighboring area, the DEA-NE. The spatial/cluster-blocking integration within the model training was performed with the CAST R package [130].



**Figure 3.** (a) Systematic spatial block CV data assignment, consisting of 10 spatial blocks with equally sized sub-areas ( $2\text{ km} \times 1\text{ km}$ ). The background map is the bathymetric hillshade. (b) Seafloor clustering CV consisting of 10 clusters with varying areas between  $0.80\text{--}3.81\text{ km}^2$ . (a1,b1) The sampling effort (here expressed as training points per  $\text{km}^2$  inside each spatial/cluster block) is unevenly distributed among the blocks/clusters and within each block/cluster. Noteworthy, the AUV data acquisition did not aim at predictive modeling and was not optimized to achieve a balanced spatial sampling. More samples have been acquired in the central part of the area due to dedicated photomosaic survey [4,9]. (a2) PMN Moran’s I value at varying distances; the Moran’s I still does not reach zero beyond 2 km. (b2) Calinski–Harabasz index

score for different numbers of clusters. Ten (10) is the optimum number in CLARA clustering, which is relatively high for the small area ( $\approx 17 \text{ km}^2$ ), indicating considerable spatial heterogeneity. The central and northern parts have higher variability. (c) Boxplot of training (hold-in) and validation (hold-out) data that are used in every training iteration for each CV scheme (RanCV = random-CV, SpaCV = spatial-CV, CluCV = cluster-CV). We noticed that in each training iteration a similar number of training data is used, reducing the unequal number of sampling points in each block or cluster. (d) Two-dimensional representation of the clustered feature space with both independent and overlapping clusters. Due to the area heterogeneity, the first two principal components retained variation accounts only for  $\approx 50\%$  of the total variation.

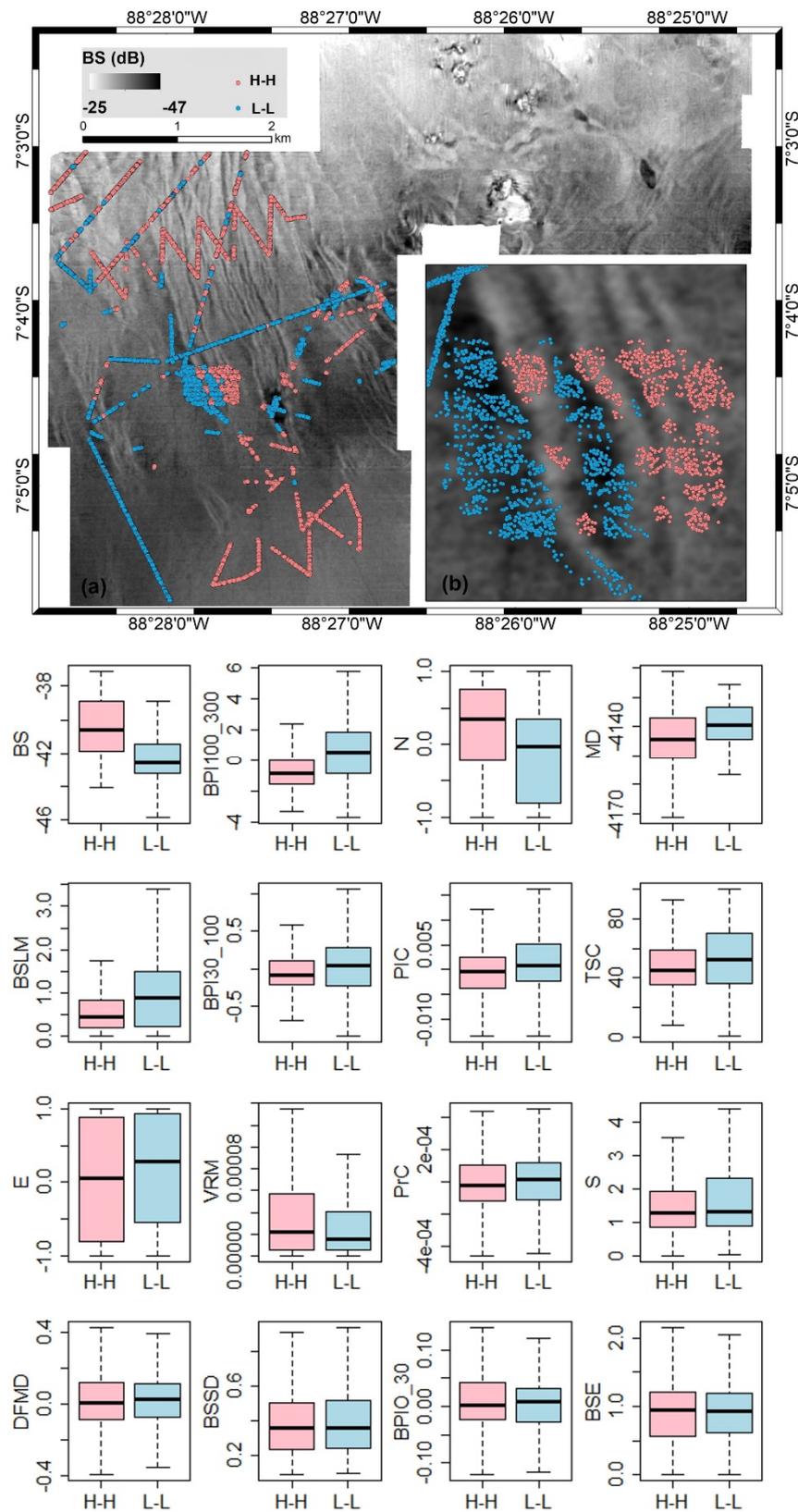
### 2.8. Dissimilarity Index and Area of Applicability

The recently proposed approach of using the dissimilarity index (DI) and area of applicability (AOA) [35] is a model-based method that assesses geographical areas, which have new feature space conditions and where the prediction error of a given pre-trained model exceeds the training CV error. The DI is based on the Euclidean feature space distances between the training dataset and the respective data of the new area. Before the distance calculation, predictors are scaled and weighted according to their variable importance of the model training. Thus, the distances of more important predictors account more to the DI estimation. The 0.95th quantile (outlier removed) of the DI is used as the AOA threshold. DI values between 0 and 1 indicate similar conditions, while values  $>1$  indicate dissimilarity. Conversely, AOA values closer to 0 indicate unknown conditions for the extrapolation. The CAST R package was used.

## 3. Results

### 3.1. PMN Spatial Distribution and Spatial Autocorrelation

Plotting of the CoMoNoD results in a spatial context revealed a local scale heterogeneous PMN distribution, with higher numbers in the northeast and southwest parts of the studied area. Of particular interest is the central area within the photomosaic survey, where patches of high PMN numbers coexist next to areas with lower numbers. The PMN distribution follows the small-scale seafloor variations created by the abyssal furrows and the prevailing current regime (Figure A4). The local Moran's I reveals that the majority of PMN (62%) are spatially aggregated into areas of H-H (27.5%) and L-L (34.5%) distributions (Figure 4). The boxplot analysis and the Wilcoxon–Mann–Whitney test show differences in the derivative distribution between the H-H and L-L clusters (Figure 4 and Table 2). In detail, higher BS values are linked with H-H clusters having a large effect in the test. The broad-scale BPI (100–300 m) also has significant differences between the two groups, with an observable moderate substantive significance (effect size). In contrast, the small-scale BPI (0–30 m) has no statistically significant variation between the two clusters. Similar to the backscatter entropy (BSE), the aspect of the seafloor surface plays a role in H-H clustering, as areas with higher numbers of PMN tend to be north-faced oriented (i.e., northness values closer to 1).



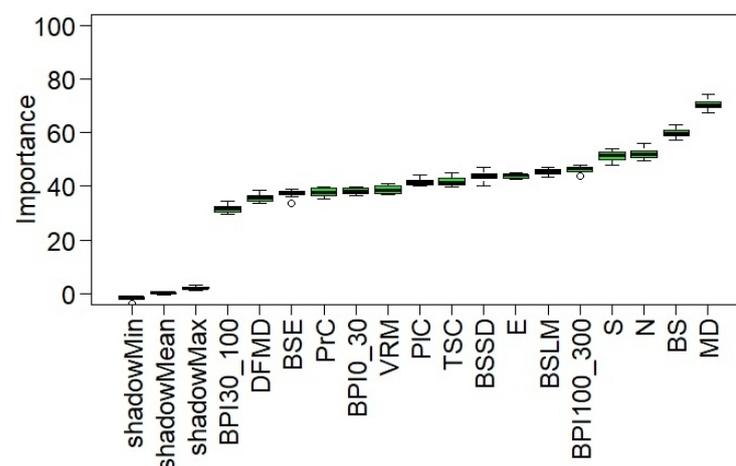
**Figure 4.** Top: PMN spatial H-H and L-L cluster distribution based on the local Moran’s I analysis. (a) The northwest and southeast parts have mainly or exclusively H-H values. Conversely, the eastern part has L-L values. (b) The clustering zonation follows the seafloor microrelief (abyssal furrows). Bottom: boxplots between statistically significant spatial clusters and MBES derivatives, showing the differences between H-H and L-L clusters.

**Table 2.** Wilcoxon–Mann–Whitney test results between statistically significant spatial clusters and MBES derivatives. The statistical significance ( $p$ -value) is given together with the substantive significance (effect size) for the interpretation of the results when the  $p$ -values are the same or similar. Its calculation and magnitude classification were based on [115].

Derivatives	Significance	Effect Size	Magnitude
Backscatter (BS)	$p < 2.2 \times 10^{-16}$	0.537	large
Bathymetric Position Index (BPI100_300)	$p < 2.2 \times 10^{-16}$	0.346	moderate
Northness (N)	$p < 2.2 \times 10^{-16}$	0.273	small
Mean Depth (MD)	$p < 2.2 \times 10^{-16}$	0.269	small
Backscatter Local Moran (BSLM)	$p < 2.2 \times 10^{-16}$	0.216	small
Bathymetric Position Index (BPI30_100)	$p < 2.2 \times 10^{-16}$	0.149	small
Plan Curvature (PIC)	$p < 2.2 \times 10^{-16}$	0.133	small
Terrain Surface Convexity (TSC)	$p < 2.2 \times 10^{-16}$	0.117	small
Eastness (E)	$p < 2.2 \times 10^{-16}$	0.077	small
Vector Ruggedness Measure (VRM)	$p = 8.156 \times 10^{-16}$	0.072	small
Profile Curvature (PrC)	$p = 3.031 \times 10^{-12}$	0.063	small
Slope (S)	$p = 1.437 \times 10^{-6}$	0.043	small
Deviation from Mean Depth (DFMD)	$p = 0.00225$	0.027	small
Backscatter SD (BSSD)	$p = 0.04903$	0.018	small
Bathymetric Position Index (BPI0_30)	$p = 0.09455$	0.015	small
Backscatter Entropy (BSE)	$p = 0.64650$	0.004	small

### 3.2. Boruta Analysis and Feature Selection

Similar to the spatial analysis, the Boruta algorithm shows that BS, BPI100\_300, and northness are important and relevant predictors. However, the MD and not the BS is the most relevant predictor (Figure 5). Moreover, S and BSSD are ranked high, although both have not a significant difference between H-H and L-L groups (Table 2). Opposite to the Wilcoxon–Mann–Whitney test, Boruta results indicate that all derivatives are relevant to predict the PMN distribution. Using all variables leads to a highly correlated dataset (Figure A5) but excluding highly-correlated variables ( $r > \pm 0.5$ ; in the Boruta importance score) results in the following predictors: MD, S, BS, N, BPI100\_300, E, BSSD, and BPI0\_30.



**Figure 5.** Boruta importance. See explanation in the text.

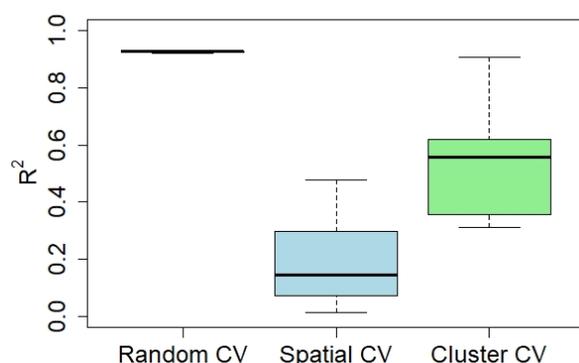
### 3.3. Model Training and CV Results

Based on the random-CV assessment, the QRF prediction performance was  $R^2 = 0.93$ . The same performance also results from the RF internal OOB error. However, when the spatial- and cluster-CV schemes were used, the prediction performance was reduced to  $R^2 = 0.19$  and  $R^2 = 0.53$ , respectively (Table 3). The performance is minimized when the spatial-blocking CV is applied. This shows that the model cannot efficiently transfer the

predictions towards spatial blocks, where the SAC is low/absent and feature space extrapolation occurs to different degrees (Figure A6). The effect of feature space extrapolation is also depicted in the clustering CV error assessment. Here, in every training repetition, the model is trying to predict a new feature space cluster. This feature space can be completely new, or it has overlap with other clusters to a varying degree (Figure 3). This fact combined with the varying degree of spatial distance among the training points (and consequently varying SAC) results in a higher training error variance than spatial- and random-CV (Figure 6). The random-CV has a minimum training error variance due to the almost identical spatial and feature space characteristics of the randomly resampling folders. The performance analysis of each spatial/cluster block could provide in-depth information regarding the specific areas and seafloor clusters for which the model performs worst, potentially guiding future sampling (Figure A7).

**Table 3.** QRF performance ( $R^2$ ) for the three different CV schemes. The internal RF OOB error assessment is also provided.

Training Data	OOB	Random-CV	Spatial-CV	Cluster-CV
H-H and L-L data (12,327)	0.93	0.93	0.19	0.53
All training data (19,952)	0.87	0.87	0.14	0.46



**Figure 6.** CV performance for each method.

The RF hyperparameter optimization differs between the spatial/cluster and random CV. In spatial- and cluster-CV, the optimum  $mtry$  value is eight (8), whereas it is five (5) in random-CV. This difference might result from the increased difficulty the model faces in spatial- and cluster-CV to extrapolate the predictions; it uses all the available information to gain predictive knowledge. In random-CV, the validation hold-out samples are identical to the training samples. Consequently, the model can predict well using fewer predictors.

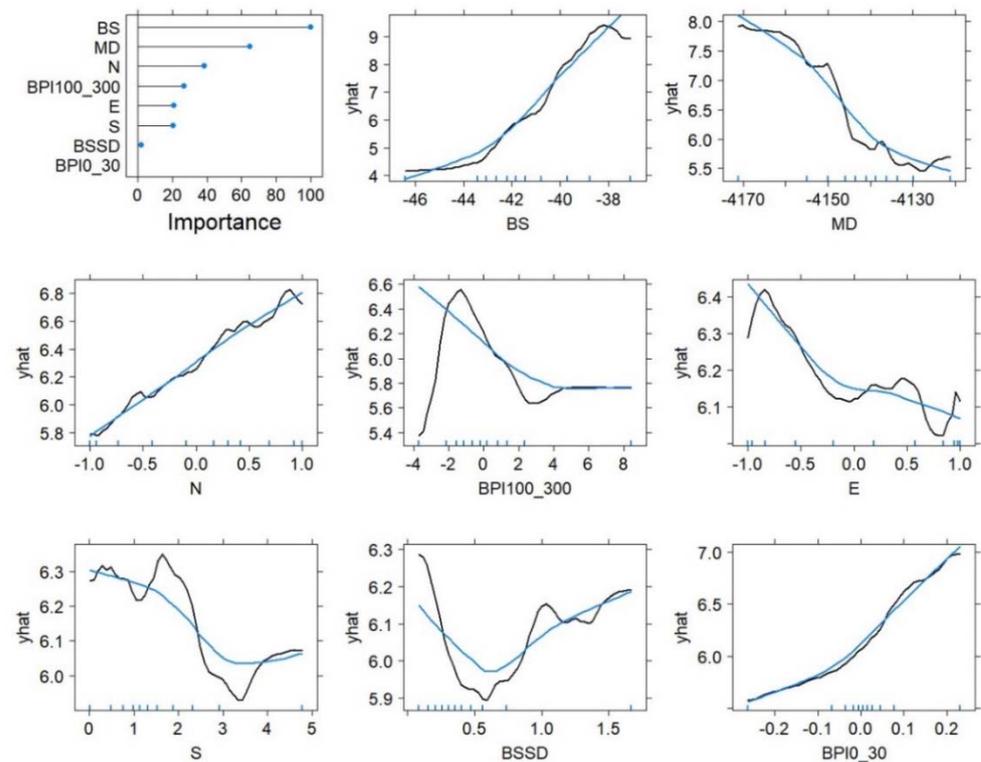
### 3.3.1. Model Training and Sample Size

No significant correlation was found between the number of training samples used from the hold-in data and the prediction performance in the remaining ( $k-1$ ) hold-out data (random-CV:  $-0.03$   $p = 0.93$ ; spatial-CV:  $-0.11$   $p = 0.75$ ; cluster-CV:  $0.23$   $p = 0.51$ ). The inclusion of all the available training points did not further improve the modeling performance, but resulted in a decrease, particularly in spatial- and cluster-CV (Table 3). This decrement is attributed to the noise added to the model by using the additional 38% of data that are not spatially aggregated. This ‘noise’ could be the result of the inherent natural variability of PMN distribution and/or wrong analyses of the CoMoNoD algorithm, e.g., through noise or lower resolution image data (greater distance to the seafloor).

### 3.3.2. Model Performance in Test Data

The performance for the test data is the same for the three models ( $R^2 = 0.76$ ), independently of the CV scheme followed. The higher predictive performance during the

random-CV implies that data overfitting occurred. This is not the case for the spatial- and cluster-CV. At the end of the training cycle, the model has seen the same training data from the random folds, spatial blocks, and clusters and similar weights and relationships between predictors and response variables are established. The relationship type (e.g., monotonic, complex) and the marginal effect on the response variable are depicted as PDP in Figures 7 and A8–A10. The low-correlated dataset ensures that the assumption of independence (uncorrelated predictors) holds, and the PDP calculation is not violated [131,132].



**Figure 7.** Top left: The QRF variable importance and PDP between PMN and seafloor derivatives. The QRF response curve is represented by a black line; the LOESS curve [133] (blue line) shows the difference between a complex ML algorithm and a simpler non-parametric regression model [122]. Blue ticks on the x-axes represent the data deciles. All RF relationships have a non-linear character, with BS, MD, N, E, and BSSD having a monotonic response. For BS, a clear trend between higher backscatter values and an increasing number of PMN can be observed. A similar trend is noticed for the N and E features, with the highest PNM numbers observed towards NW dipping slopes, which are parallel the dominant bottom current direction. The BPI100\_300 and slope S have complex relationships, with data aggregating in a specific range of the variable. This information could, in general, be a valuable indication for favorable geomorphological characteristics of PMN occurrence (e.g., slopes  $< 2^\circ$ ), but clearer indications are not given due to the sampling distribution in relation to the values of the derivatives. In most cases, the training samples are well distributed across the feature range, providing confidence regarding the established response curves. For BPI100\_300 and BPI0\_30, the data are aggregated in a small range, creating none-sampled regions inside the training feature space that consequently force model extrapolation. The model extrapolation is visualized better in the two- and three-way interaction PDP, where the data convex hulls between two predictor variables are presented (Figures A8–A10). All models have the same variable importance ranking and PDP and, thus, only this from spatial-CV is presented.

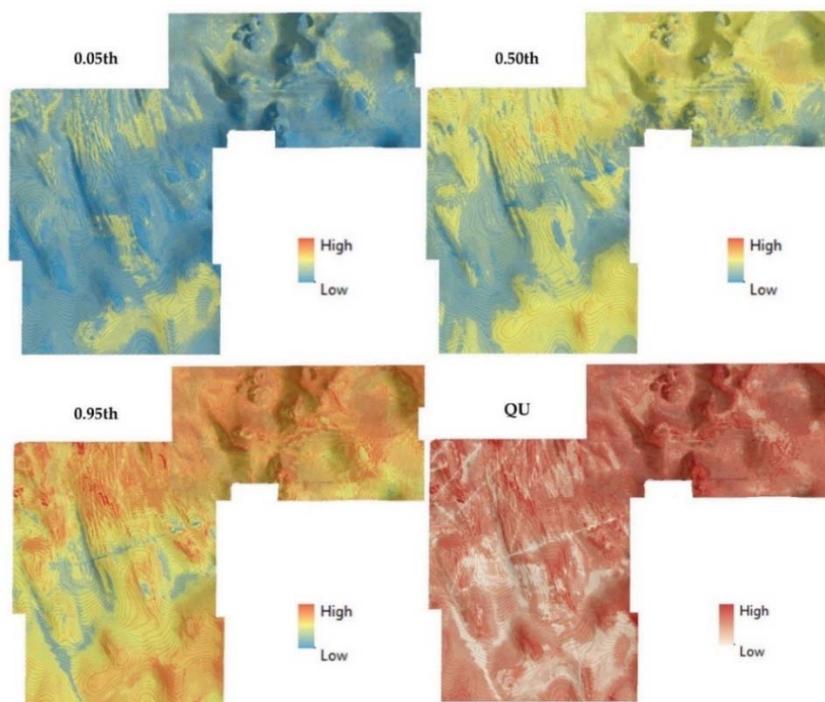
### 3.3.3. QRF Variable Importance

The backscatter (BS) has the highest variable importance, followed by mean depth (MD), northness (N), and the coarse bathymetric position index (BPI100\_300). Slope (S) and eastness (E) contribute less, while the backscatter standard deviation (BSSD) and BPI0\_30

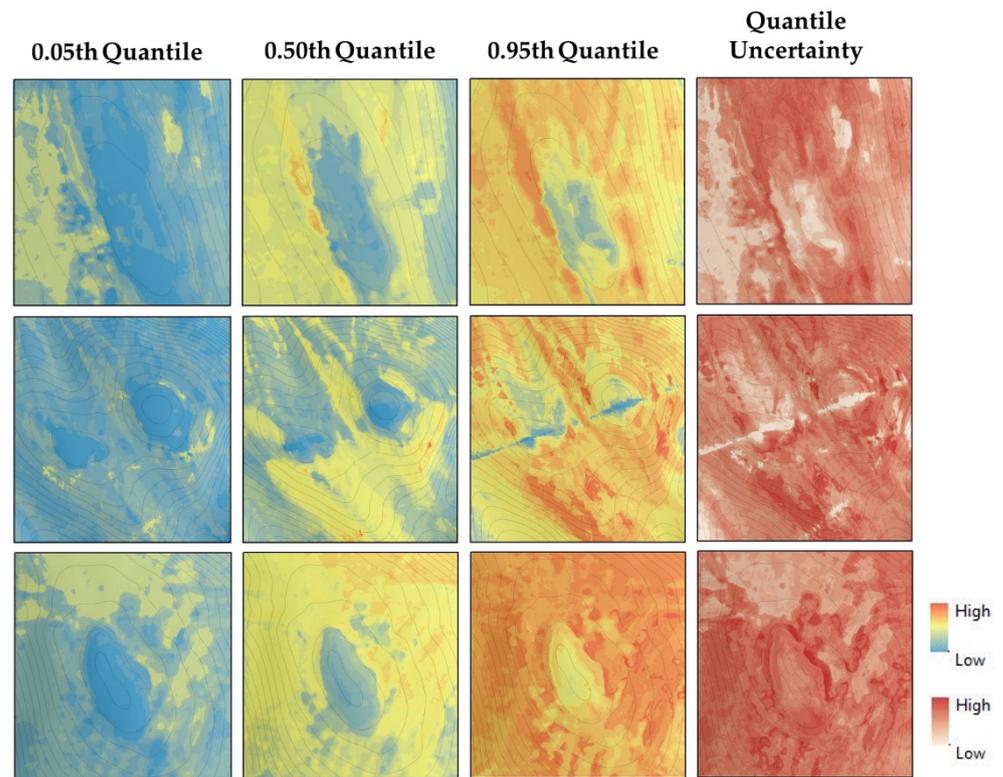
have marginal or no contribution at all (Figure 7). This ranking is closer to the spatial analysis results than from the Boruta analyses, where the BSSD, S, and MD have scored higher. The overall higher agreement between local Moran's I analysis and a spatially trained QRF shows that the Boruta may result in sub-optimum importance ranking, due to the overfitting that occurs when non-spatial training is performed.

### 3.4. QRF Spatial Predictions and Uncertainty

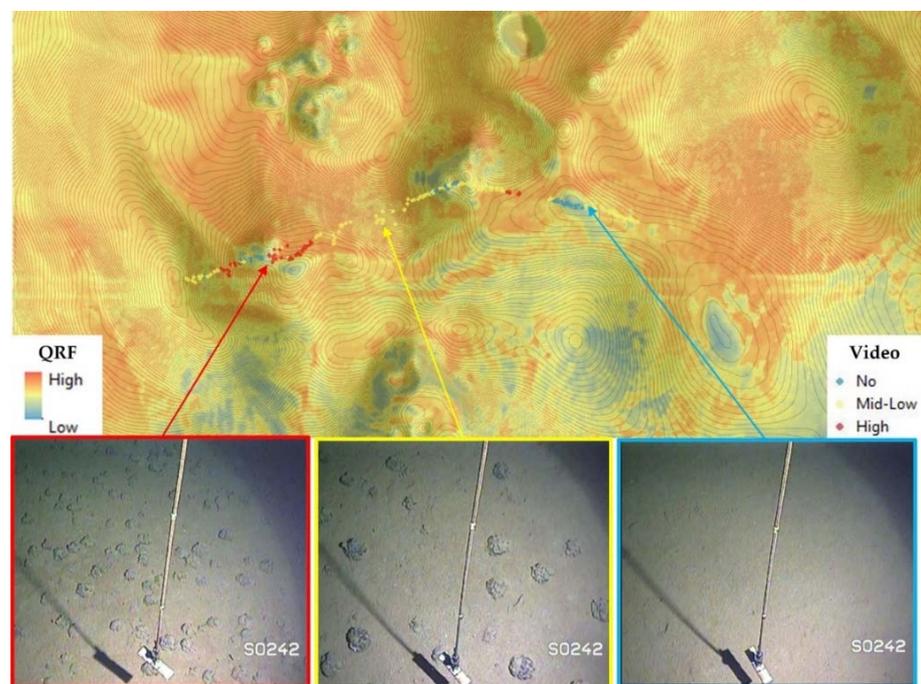
The model prediction shows a heterogeneous PMN distribution, with a higher number of PMN aggregated in the northern and southern areas of the DEA, that follow the overall seafloor topography and the bottom current regime (Figure 8). The lower and upper quantile predictions differ, with the 0.05th quantile being less affected by sampling artifacts, which are prominent in the 0.95th quantile. Moreover, the 0.05th and 0.50th quantile predict the spatial extent of PMN-free patches better (Figure 9). Inside the DEA, the central and southern parts have the lowest quantile difference, due to the increased sampling effort. The DEA-NE area also has parts with alternating high and low numbers of PMN (Figure 8). In Figure 10, the RF model results were plotted together with the visual observations from the ocean floor observation system, showing an overall spatial agreement. However, the predictions inside the DEA-NE area have high uncertainty, as the model has seen no training data from this area and the seafloor is morphologically more complex than the DEA (Figure 8).



**Figure 8.** QRF spatial predictions for the entire AUV-mapped area. PMN are spatially aggregated in the northern and southern parts of the DEA, with the eastern part being the least covered. Despite the overall spatial agreement, the 0.05th and 0.50th quantiles predict the extent of the PMN free black patches better. In addition, the 0.95th quantile has linear artifacts; these are caused due to the difference between the model trend and sample values. The quantile uncertainty (QU) is minimized in the areas that have samples (here images) and have a similar seafloor geomorphology. The superimposed bathymetric contours have a 1 m depth interval.



**Figure 9.** Enlarged maps of the PMN abundance prediction from three areas with black patches (no nodules). From left to right are the 0.05th, 0.50th, and 0.95th quantiles shown, followed by the quantile uncertainty. The superimposed bathymetric contours have a 1 m depth interval.



**Figure 10.** Semi-quantitative validation between the RF predictions and video observations (S0242-1#135\_OFOS-6). Prediction and observation have the same spatial trend, showing alternations between areas with higher, lower, and no PMN. The superimposed bathymetric contours have a 1 m interval.

### 3.5. Dissimilarity Analysis and Area of Applicability

The dissimilarity analysis showed that the DEA-NE is dissimilar to the training samples, especially in the areas with rugged seafloor, increased slopes, and shallower bathymetric depth range. These circumstances reduce the area of applicability of the model, showing that additional samples must be added for better predictions. Data derived from the random-CV model have the smallest AOA. The overoptimistic error assessment that occurs limits the model AOA to a small region around the training locations, where the CV error is still small and comparable. Conversely, spatial- and cluster-CV have a larger prediction error. Hence, this error applies towards a wider area (Figure 11). The AOA also increased when excluding the MD, BSSD, and BPI0\_30 features. The latter two were excluded due to their negligible contribution in the final model. Although, they are well sampled (Figure 11). MD was excluded, as the training samples do not cover the entire depth range of the study area (Figure 11). We have prior knowledge from video observations (e.g., Figure 10 and previous studies [62,63,69]) that the DEA-NE is inside the bathymetric depth range that favors the aggregation of PMN. The exclusion of the mean depth could provide a simpler and better transferable model (on local scales), putting a greater importance on relative bathymetric variations, such as local elevations and depressions expressed by the BPI. The exclusion of the three variables MD, BSSD, and BPI0\_30 resulted in a decreased dissimilarity and a larger AOA (Figure 11), but the predictive performance for the test data was reduced ( $R^2 = 0.73$ ). In general, several parts of the DEA-NE remain unsuitable for predictions using the developed model, due to the complex geomorphology and lack of samples that cover this complex terrain.

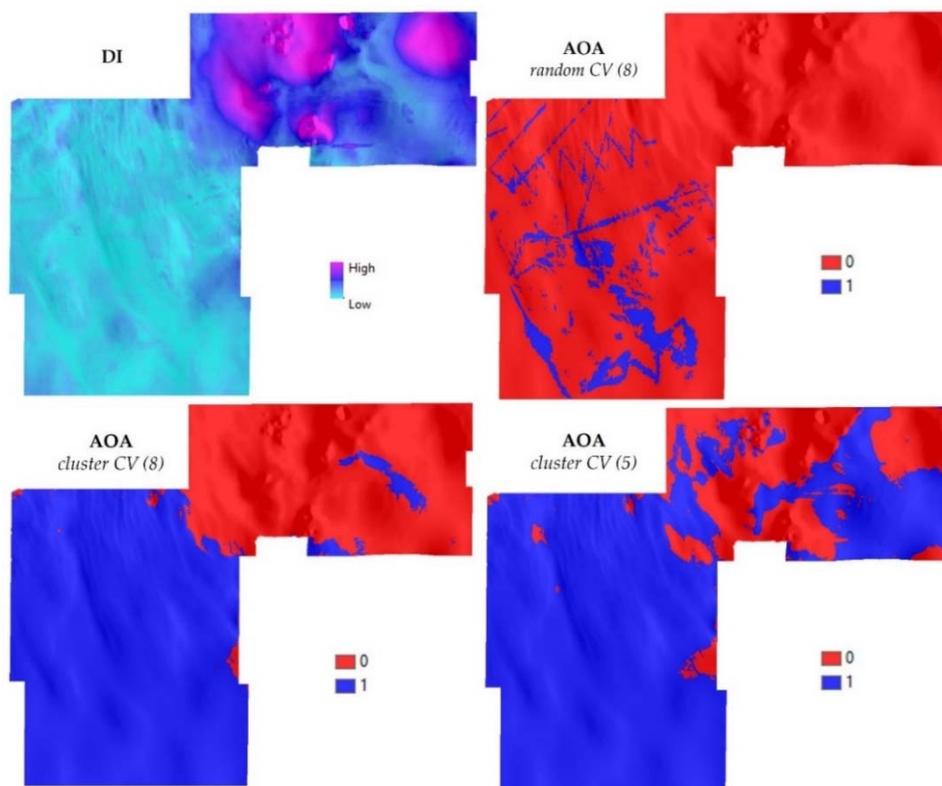
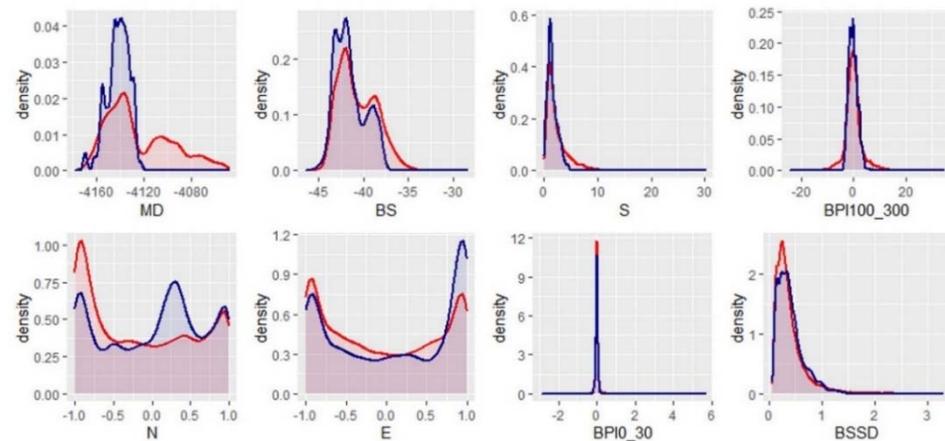


Figure 11. Cont.



**Figure 11.** Top left: the dissimilarity index (DI) between training samples and the study area resulted from the QRF model weights. Top right: the random-CV has the most limited AOA, which is restricted to the area directly around the training samples. For block- and cluster-CV (middle left), the prediction error is comparable for the entire DEA and a small part of the DEA-NE (0 = Not applicable, 1 = Applicable). Without the mean depth (MD) as predictor variable, the AOA increases (middle right). The spatial- and cluster-CV have very similar AOA statistic values ( $r = 0.96$ ,  $p < 2.2 \times 10^{-16}$ ). Thus, only the cluster-CV AOA is presented in the density plots (bottom). The density plots show the difference between the training samples (blue) and the entire study area (red). Mean depth (MD) has the most significant difference, followed by backscatter (BS), northness (N), and eastness (E).

#### 4. Discussion

This article presents a PMN prediction case study based on AUV hydroacoustic and image data from the Peru Basin in the Pacific. The study highlights the presence of spatial autocorrelation within the polymetallic nodule (PMN) distribution, with the PMN being spatially aggregated due to local geomorphological settings and the prevailing bottom current regime. The northwest oriented bottom current is channeled through abyssal furrows and erodes fine sediments. This leaves abundant fragmental materials (fish teeth, basalt debris, tiny pieces of broken shell, or nodules) as nuclei for the formation of PMN [134]. The higher number of PMN in the northern part of the DEA and DEA-NE is supported by the vicinity to seamounts (Figure 1), which provides additional nucleus material, initiating the nodule formation presence [135]. It is still unclear to what degree such influence has in comparison or synergistically to the geochemical properties of the sediment, which play an essential role in the PMN spatial distribution, their abundance, and size [134,136]. Using spatial geochemical information as input data in the modeling process would have been advantageous, as omitting such drivers limits the predictive performance. Unfortunately, the existing sampling methods (gravity cores, push cores, multi-cores, and box-corers) have limited spatial coverage, are typically clustered in small sub-regions due to the need for multiple replicates in biological studies, and generally do not occur in great numbers due to the time-consuming sampling (4–5 h per sample) and subsequent geochemical analyses.

Although our knowledge of the real reasons and their interrelationship with PMN formation is incomplete, analyzing the spatial autocorrelation has a strong potential to explore complex geomorphological relations, as it is an inherent part of the natural PMN distribution and variability across different spatial scales. In this respect, calculating the local Moran's I can contribute in three ways:

(a) Local Moran's I reveals significant PMN spatial clusters. Further investigation of such areas revealed the spatial dependency of PMN by the seafloor morphology (exogenous autocorrelation), as the MBES derivatives values differ significantly between the H-H and L-L groups.

(b) Local Moran's I can effectively identify predictors for spatial modeling. In contrast, the non-spatial Boruta algorithm showed increased relevance in predictors such as the BSSD and BPI0\_30 that had a minor to zero contribution in the final RF model.

(c) Local Moran's I reduces the number of training samples needed. Only the data that contribute the most are kept, and the model becomes computationally lighter while its performance increases. Studies have highlighted the need for better data (i.e., representative with low sampling uncertainty and noise, having enough variation to capture critical patterns in the data, and well distributed across the entire feature space) in data-driven approaches such as geostatistical and ML predictive modeling [18,137,138]. In this respect, the local Moran's I analysis could be an efficient tool.

Spatial autocorrelation (SAC) can help to address the issues mentioned above, at the same time not considering spatial autocorrelation may result in over-optimistic CV predictions. Similar to recent studies [17,19–21,23,24,139,140], we showed that the conventional and commonly applied random 10-fold CV could not assess the model performance when spatial autocorrelation is present. In such cases, spatial- and cluster-blocking perform better. Spatial-blocking is the least biased and results in lower variance than cluster-blocking; the higher variance of cluster-blocking has also been mentioned by others [17].

The spatial- and cluster-CV also yielded different mtry hyperparameter values. Five predictors were adequate to predict the hold-out samples inside the random (and identical) folders. However, for the spatial- and cluster-CV, all available predictors (eight) were needed. The final models have the same performance in the unseen (but similar) test data. By applying different CV methods, we changed our assessment method, not the model itself. This fact shows that a spatially similar test dataset can show how well the model can reproduce the data, but it cannot inform about how the model performs when it is transferred to another area. In this case, the spatial-CV is recommended.

The random-CV could still be preferable in case of a) small datasets with geographical separation within the training points that exceed the area of spatial autocorrelation (influence zone) or b) when the predictions are restricted to nearby locations of the training samples [20,33,141]. Random-CV is the most straightforward implementation. Contrasting spatial-blocking demands a prior correlogram calculation to identify the block size, and cluster-blocking demands the use of clustering indexes (e.g., Calinski–Harabasz) to find the optimum number of clusters along with data preprocessing (e.g., scaling of input MBES derivatives). These procedures usually demand the use of more than one software, increasing the time and complexity of the processing.

However, the biggest challenge lies in realizing the need for training data that are simultaneously well spread across the geographical and feature space of the covariates used for the analysis. Despite the significant advances in terrestrial spatial sampling [142–146], deep-sea studies have only lately started developing methods to optimize AUV photo sampling in a way that maximizes the environmental information [147–150]. In this respect, the QRF and AOA could guide sampling efforts in previously sampled areas or even during sampling surveys (adaptive sampling or/and active learning).

The poor representation of the feature space, especially the range of the most important predictor variables, causes a reduced performance and transferability of regression random forests modelling [151,152]. In our case, the dissimilarity analysis (DI and density curves) showed differences between the training samples and the targeted feature space. This becomes visible in the quantile uncertainty (QU), which is maximized in the DEA-NE. The QU is a helpful tool to explore the model variation inside the convex hull of the training points (Figures A8 and A9). QU is correlated with DI ( $r = 0.65$ ,  $p < 2.2 \times 10^{-16}$ ). However, there are subareas inside the DEA-NE with high dissimilarity but no high analogue uncertainty. This is a misleading extrapolation effect that has been described by the AOA authors [35]. In other words, QRF can provide locations that have increased model variation, but it cannot identify if this is caused due to the inherent uncertainty of the hydroacoustic and image data or due to extrapolation.

Any model extrapolation beyond the training domain should be accompanied by a thorough transferability assessment, ideally with an external evaluation using data from the new area [22,33,153]. This was not possible here, due to a lack of data. Instead, we used the recently proposed AOA method. Within the DEA-NE area, the expected prediction error is higher than in the trained area of the DEA, reducing the model applicability. The exclusion of predictors that are prone to overfitting, such as elevation/depth, increases the generalization performance, which has also been mentioned by [22]. The advantage of RF to build complex non-linear response curves with the training data, outperforming other regression methods [154], could result in a disadvantage when these associations occur only locally [155]. A successful transfer thus relies on the assumption that the relationship between response variable and predictor exists in both areas. The generalization performance is maximized when the data information and feature selection are combined with the domain knowledge, even if the domain knowledge is basic or imperfect [156,157]. Other studies also showed that machine learning models with only a few predictor variables are more transferable in marine habitat mapping than complex ones [158].

Similarly, less complex models (e.g., partial least squares regression) could generalize better when spatial-CV schemes are applied, highlighting the trade-off decision to be made between accuracy and generalization performance [21,25,27]. Similarly, ensemble ML models that average predictions of different single models could also yield better predictions across different areas, as presented by [23,141]. The comparison of different models was beyond the scope of this paper. Here, we focused on the widely used random forests algorithm and its quantile uncertainty (i.e., QRF). We should underline that transferability is not the primary goal when applying machine learning-based spatial modelling. ML is designed to derive accurate predictions based on an existing measurement that captures the underlying relationship, for which our knowledge or conceptual understanding is still developing. Towards this direction, the RF importance score is a valuable measure to test known hypotheses, but also to generate new ones [155,159,160].

Future research in other parts of the Peru Basin and other known fields with polymetallic nodules (e.g., the Clarion-Clipperton Zone) could further enlighten the drivers of spatial autocorrelation. In addition, the spatial analysis in various scales would also provide insights for the underlying mechanisms that influence the spatial distribution of polymetallic nodules.

## 5. Conclusions

Our case study shows that spatial predictions of polymetallic nodules with ML methods need to be spatial-cross-validated when the spatial autocorrelation is present, and that the seafloor morphology varies. Similarly, model transfer to areas with scarce or no data should be evaluated by regarding the new area similarity with the training domain. Ideally, each ML predictive spatial map should be accompanied with its cross-validation strategy, uncertainty prediction, and area of application analyses, for supporting the model interpretation and decision making. In other words, the focus should not only lie in generating the final prediction map, but also on how this map has been derived from the fitted model and data.

**Author Contributions:** The authors of this paper contributed as follows: conceptualization, I.-Z.G. and J.G.; methodology, I.-Z.G. and J.G.; software, I.-Z.G.; validation, I.-Z.G.; formal analysis, I.-Z.G.; investigation, I.-Z.G.; resources, J.G.; data curation, I.-Z.G. and J.G.; writing—original draft preparation, I.-Z.G.; writing—review and editing, I.-Z.G. and J.G.; visualization, I.-Z.G.; supervision, J.G.; project administration, J.G.; funding acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

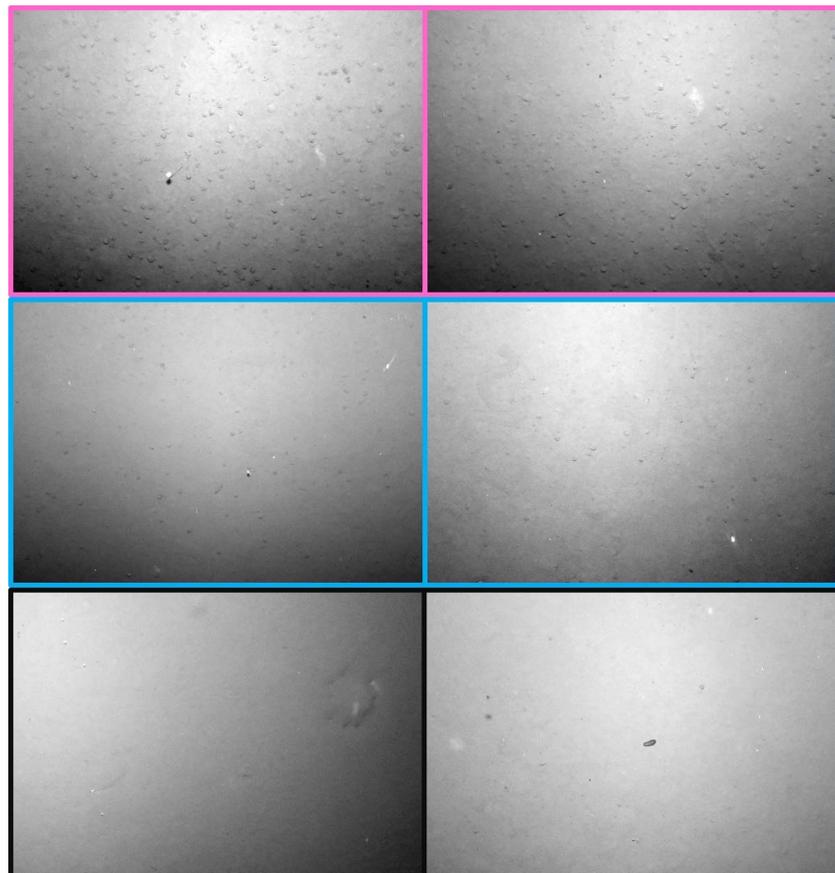
**Funding:** All data were acquired within the JPIO Project “Ecological Aspects of Deep-Sea Mining” framework, funded through BMBF grant 03F0707A. Funding was made available through MarTERA grant COMPASS-Drimp from BMWi grant 03SX466B.

**Data Availability Statement:** The following data are available on PANGAEA® Data Publisher: SO242-1 ship-based MBES raw data at <https://doi.org/10.1594/PANGAEA.859528>; SO242-1 AUV raw images at <https://doi.org/10.1594/PANGAEA.882349>; SO242-1 AUV processed images at <https://doi.org/10.1594/PANGAEA.883838>; Source code of the CoMoNoD algorithm at <https://doi.pangaea.de/10.1594/PANGAEA.875070>. The datasets and scripts generated for this study are available upon reasonable request to the corresponding author. These include: the processed ship-based MBES bathymetry, the processed and merged AUV-based MBES bathymetry, backscatter, their derivatives, the final selected subset of images for spatial analysis and modeling, the ArcMap spatial autocorrelation analysis outputs, and the R scripts. Moreover, the SO242-1 OFOS video footage and annotations.

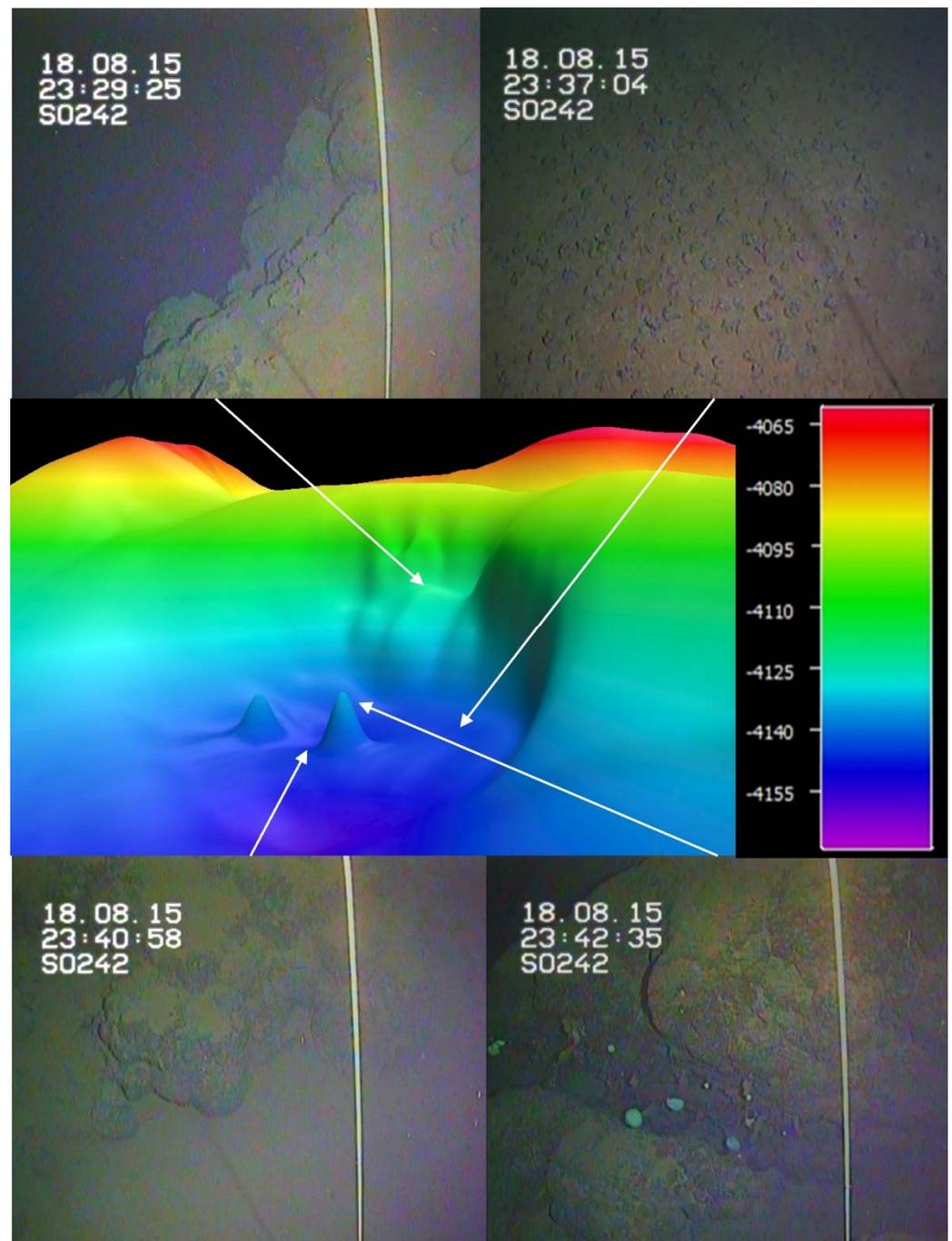
**Acknowledgments:** We thank the captain and crew of RV Sonne for their contribution to a successful cruise. We express our gratefulness to the GEOMAR AUV team for their support during the cruise and the preprocessing of the MBES data. We thank Timm Schoenning and Peter Urban for fruitful discussions during the conceptualization stage. We value Astrid Ulbrich for the proofreading of the manuscript. Finally, we thank the GEOMAR Library team for its support in gathering the necessary bibliography. This is publication 49 of the DeepSea Monitoring Group at GEOMAR Helmholtz Centre for Ocean Research Kiel.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript or in the decision to publish the results.

## Appendix A. Study Area Geomorphology

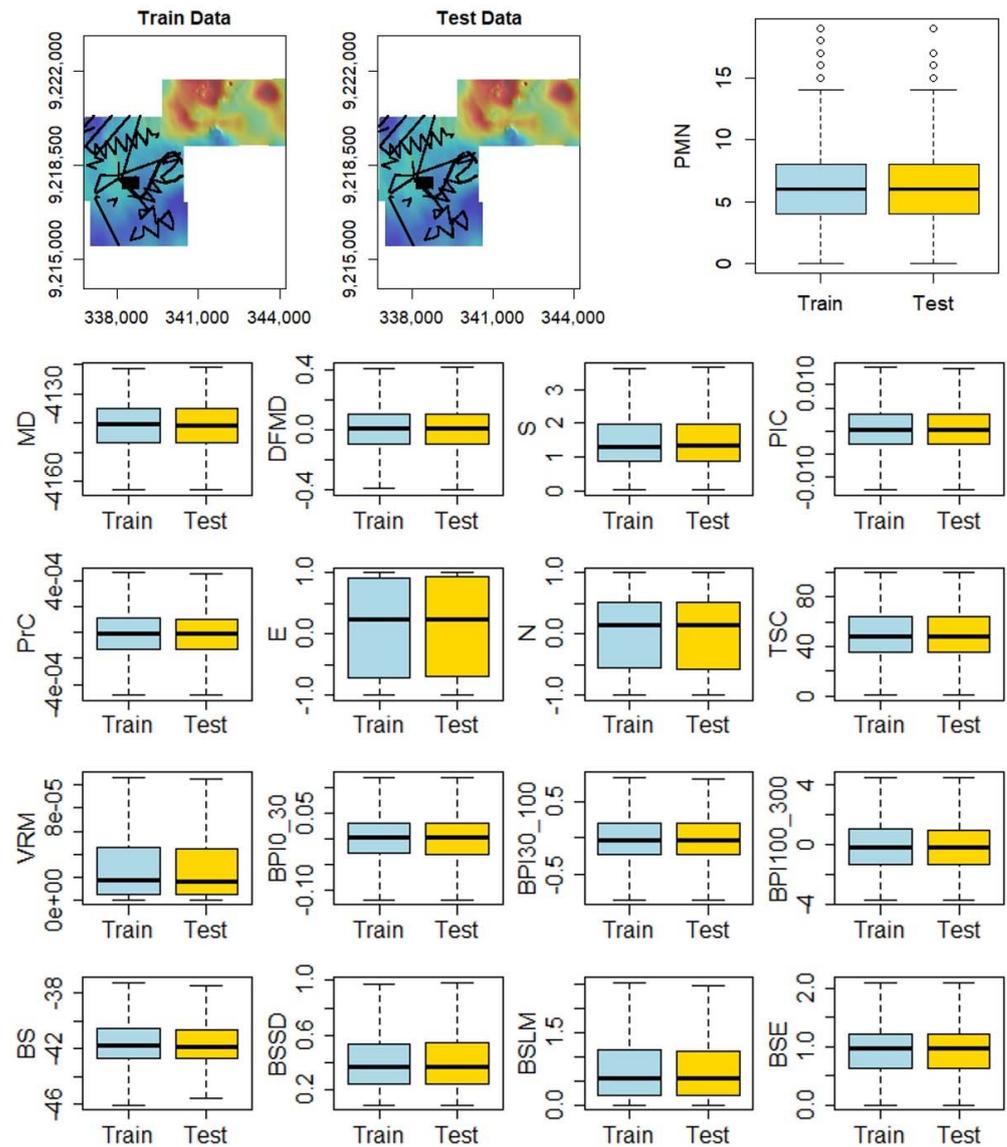


**Figure A1.** AUV grayscale images from within the DEA area. Top: high number of PMN; middle: lower number of PMN; bottom: no PMN. The colored frames correspond to the spatial clusters (Section 3.1).

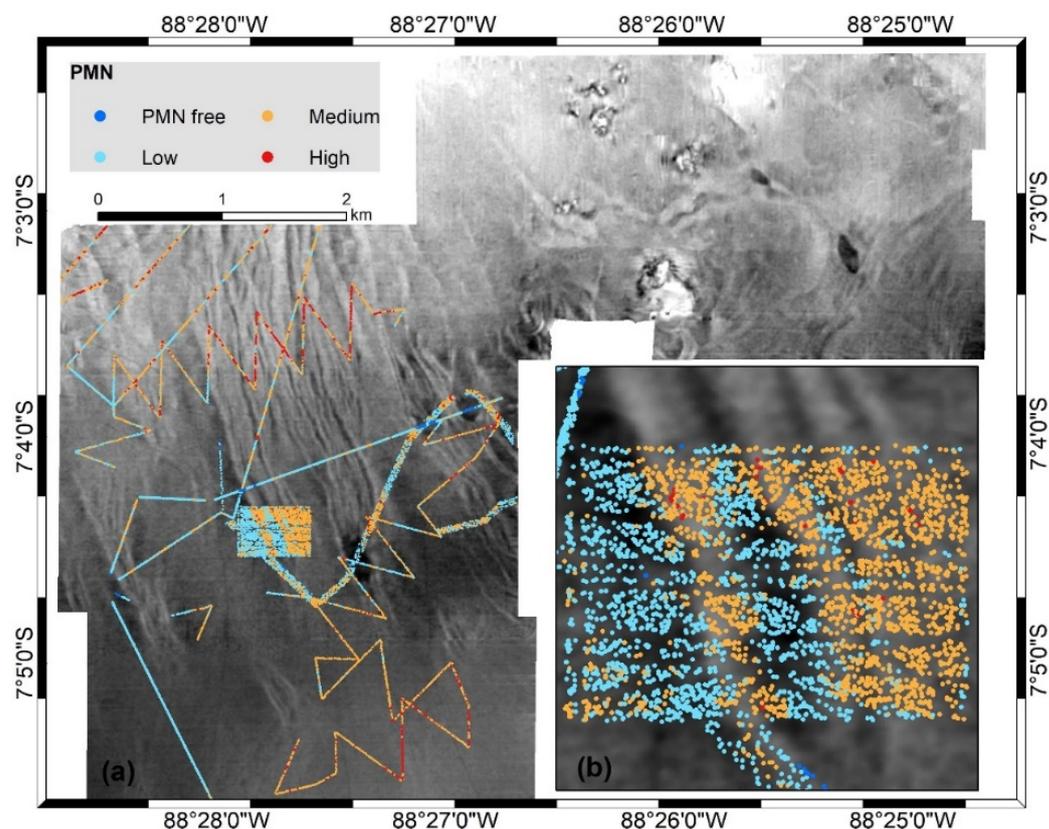


**Figure A2.** Examples of OFOS photos (SO242-1\_135\_OFOS06) inside the crate-shaped structure with two outcropping volcanic cones of pillow basalts (DEA-NE). Top left: the short hillcrest with increased slopes; top right: PMN mixed with talus debris in the foot slope and crater floor; middle: 3D bathymetric representation (exaggerated  $\times 5$ ); bottom: the base of the volcanic cone (left) and the summit (right).

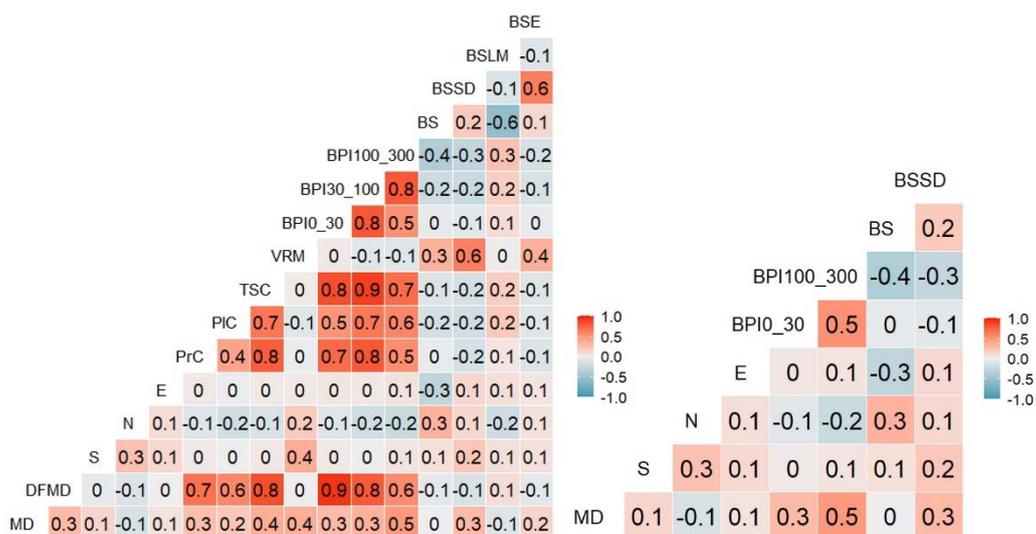
Appendix B. Methodology



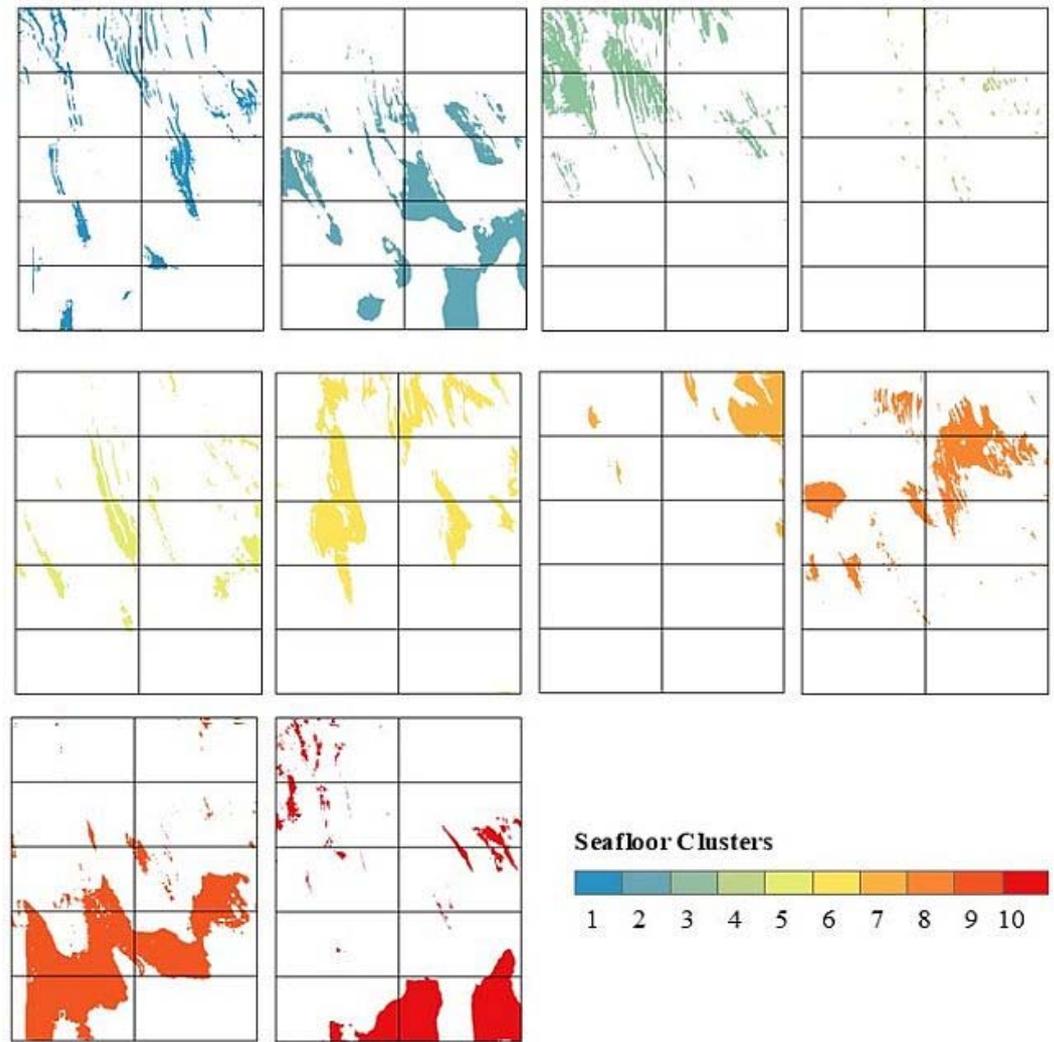
**Figure A3.** Top left: train and test dataset distribution have the same geographical coverage and distribution characteristics. The MBES derivative values were extracted at each photo location using the Extract Multi Values to Points tool (Spatial Analyst) in ArcMap 10.6.



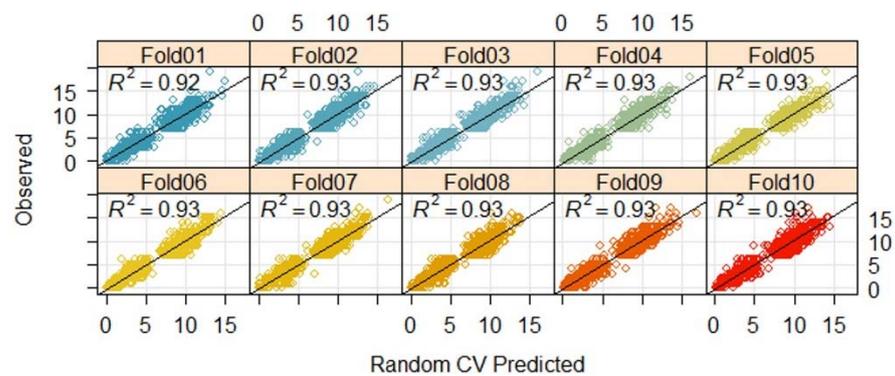
**Figure A4.** (a) PMN distribution within the study area (DEA). (b) The central part of the area is characterized by successive alternations of higher and lower numbers of PMN. These alternations seem to follow the abyssal furrows microrelief. The background map is the backscatter intensity, with brighter areas to imply higher reflectivity. A higher-resolution map of PMN distribution inside the photomosaic area is provided by [9].



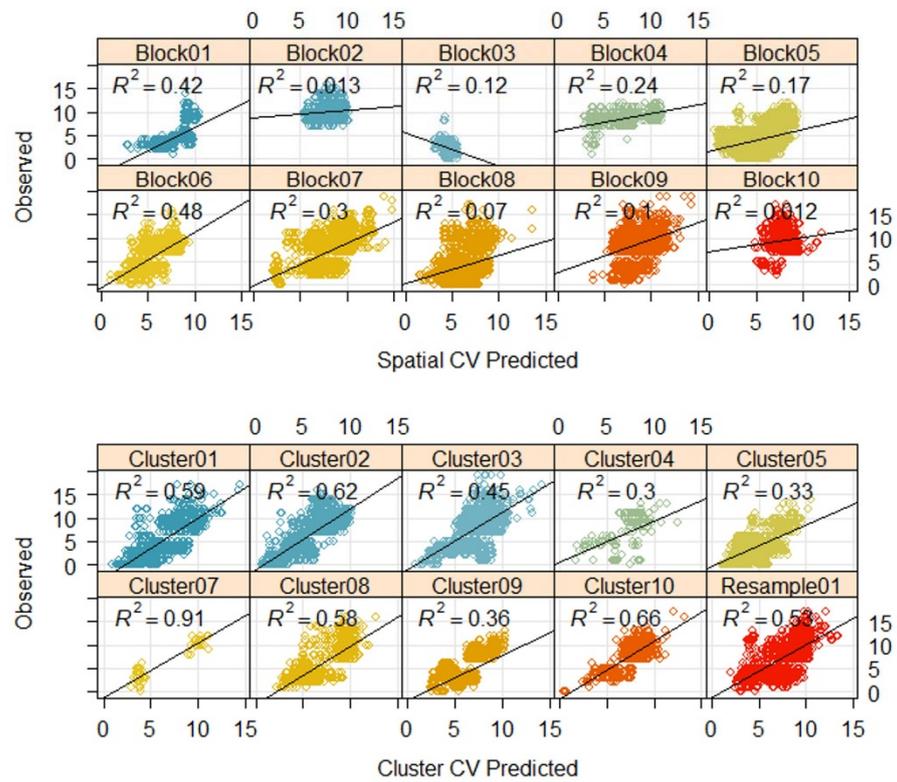
**Figure A5.** MBES derivatives correlation plot (left) and low-correlated predictors according to the Boruta ranking (right).



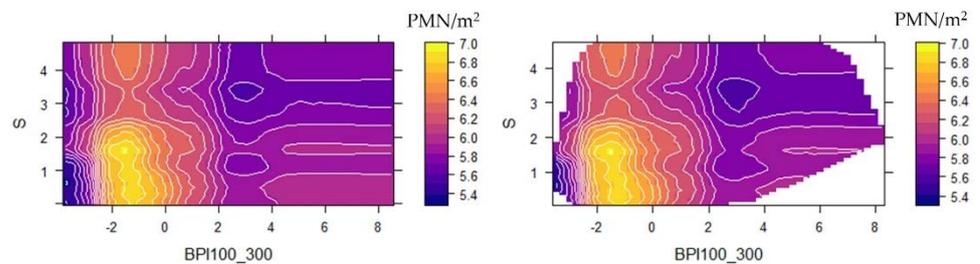
**Figure A6.** Seafloor clusters within each spatial block; at least one cluster is absent in each of the spatial blocks. This causes spatial-CV extrapolation, which creates results such as the cluster-CV approach.



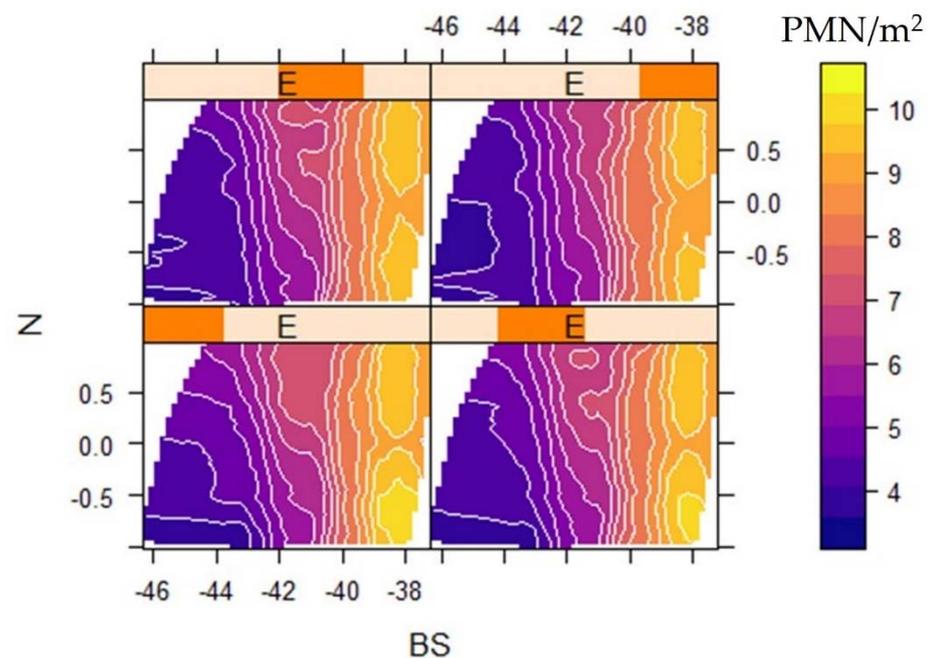
**Figure A7.** Cont.



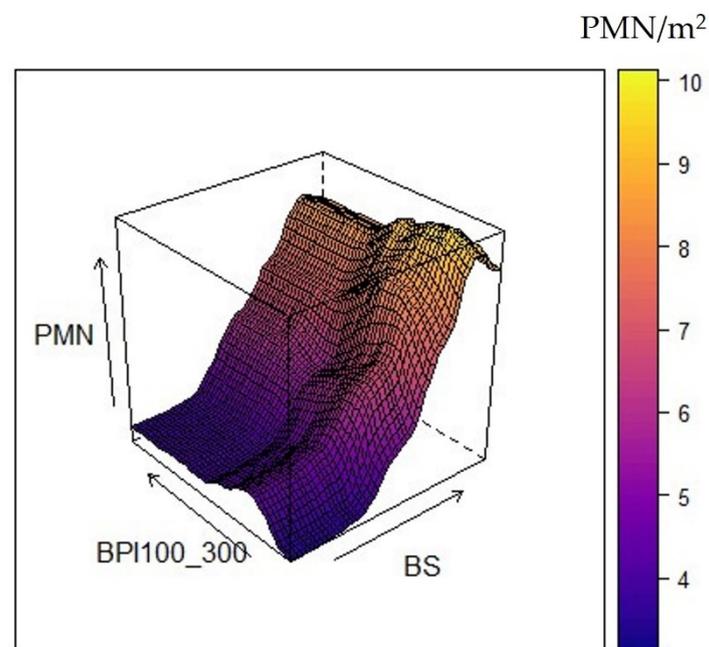
**Figure A7.** Random fold (top), spatial block (middle), and cluster block (bottom) resampling prediction performance during CV. The model uses the information included in the nine folds/blocks/clusters to predict the remaining fold/block/cluster.



**Figure A8.** Shown are 2D PDPs between slope, BPI100\_300, and PMN. Left: extrapolated feature space, Right: convex hull of the two variables and the non-described empty feature space. The interpretation outside of the convex hull is inadvisable.



**Figure A9.** Three-way interaction PDP show the PMN response (colored) related to the BS, N, and E. The convex hull of the three variables is presented. The number of PMN increases with higher backscatter intensity and to the NW direction. In order to display the 3rd continuous variable (N), the PDP package converts it into shingles (a data structure that consists of a numeric vector along with four intervals that define the classes of the shingle. The intervals overlap is 10%).



**Figure A10.** A 3D two-way interaction PDP shows the PMN response (colored surface) related to the BS and BPI100\_300.

## References

1. Hein, J.R.; Koschinsky, A.; Kuhn, T. Deep-ocean polymetallic nodules as a resource for critical materials. *Nat. Rev. Earth Environ.* **2020**, *1*, 158–169. [[CrossRef](#)]
2. Hein, J.R.; Mizell, K.; Koschinsky, A.; Conrad, T.A. Deep-ocean mineral deposits as a source of critical metals for high- and green-technology applications: Comparison with land-based resources. *Ore Geol. Rev.* **2013**, *51*, 1–14. [[CrossRef](#)]

3. EC Communication COM, 474, F. Critical Raw Materials Resilience: Charting a Path towards Greater Security and Sustainability. 2020. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0474&from=EN> (accessed on 30 August 2021).
4. Schoening, T.; Purser, A.; Langenkämper, D.; Suck, I.; Taylor, J.; Cuvelier, D.; Lins, L.; Simon-Lledó, E.; Marcon, Y.; Jones, D.O.B.; et al. Megafauna community assessment of polymetallic-nodule fields with cameras: Platform and methodology comparison. *Biogeosciences* **2020**, *17*, 3115–3133. [[CrossRef](#)]
5. Schoening, T.; Köser, K.; Greinert, J. An acquisition, curation and management workflow for sustainable, terabyte-scale marine image analysis. *Sci. Data* **2018**, *5*, 180181. [[CrossRef](#)] [[PubMed](#)]
6. Simon-Lledó, E.; Bett, B.J.; Huvenne, V.A.I.; Köser, K.; Schoening, T.; Greinert, J.; Jones, D.O.B. Biological effects 26 years after simulated deep-sea mining. *Sci. Rep.* **2019**, *9*, 8040. [[CrossRef](#)] [[PubMed](#)]
7. Gazis, I.Z.; Schoening, T.; Alevizos, E.; Greinert, J. Quantitative mapping and predictive modeling of Mn nodules' distribution from hydroacoustic and optical AUV data linked by random forests machine learning. *Biogeosciences* **2018**, *15*, 7347–7377. [[CrossRef](#)]
8. Peukert, A.; Schoening, T.; Alevizos, E.; Köser, K.; Kwasnitschka, T.; Greinert, J. Understanding Mn-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and optical data. *Biogeosciences* **2018**, *15*, 2525–2549. [[CrossRef](#)]
9. Schoening, T.; Jones, D.O.B.; Greinert, J. Compact-Morphology-based poly-metallic Nodule Delineation. *Sci. Rep.* **2017**, *7*, 13338. [[CrossRef](#)]
10. Hari, V.N.; Kalyan, B.; Chitre, M.; Ganesan, V. Spatial Modeling of Deep-Sea Ferromanganese Nodules with Limited Data Using Neural Networks. *IEEE J. Ocean. Eng.* **2018**, *43*, 997–1014. [[CrossRef](#)]
11. Kaikkonen, L.; Virtanen, E.A.; Kostamo, K.; Lappalainen, J.; Kotilainen, A.T. Extensive Coverage of Marine Mineral Concretions Revealed in Shallow Shelf Sea Areas. *Front. Mar. Sci.* **2019**, *6*, 541. [[CrossRef](#)]
12. Wong, L.J.; Kalyan, B.; Chitre, M.; Vishnu, H. Acoustic Assessment of Polymetallic Nodule Abundance Using Sidescan Sonar and Altimeter. *IEEE J. Ocean. Eng.* **2021**, *46*, 132–142. [[CrossRef](#)]
13. Dutkiewicz, A.; Judge, A.; Müller, R.D. Environmental predictors of deep-sea polymetallic nodule occurrence in the global ocean. *Geology* **2020**, *48*, 293–297. [[CrossRef](#)]
14. Wasilewska-Błaszczuk, M.; Mucha, J. Application of General Linear Models (GLM) to assess nodule abundance based on a photographic survey (case study from IOM Area, Pacific Ocean). *Minerals* **2021**, *11*, 427. [[CrossRef](#)]
15. Kuhn, T.; Rühlemann, C. Exploration of polymetallic nodules and resource assessment: A case study from the German contract area in the clarion-clipperton zone of the tropical northeast pacific. *Minerals* **2021**, *11*, 618. [[CrossRef](#)]
16. Anselin, L. Local Indicators of Spatial Association-LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [[CrossRef](#)]
17. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [[CrossRef](#)]
18. Hengl, T.; Nussbaum, M.; Wright, M.N.; Heuvelink, G.B.M.; Gräler, B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ* **2018**, *6*, e5518. [[CrossRef](#)]
19. Meyer, H.; Reudenbach, C.; Hengl, T.; Katurji, M.; Nauss, T. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environ. Model. Softw.* **2018**, *101*, 1–9. [[CrossRef](#)]
20. Misiuk, B.; Diesing, M.; Aitken, A.; Brown, C.J.; Edinger, E.N.; Bell, T. A Spatially Explicit Comparison of Quantitative and Categorical Modelling Approaches for Mapping Seabed Sediments Using Random Forest. *Geosciences* **2019**, *9*, 254. [[CrossRef](#)]
21. Ploton, P.; Mortier, F.; Réjou-Méchain, M.; Barbier, N.; Picard, N.; Rossi, V.; Dormann, C.; Cornu, G.; Viennois, G.; Bayol, N.; et al. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nat. Commun.* **2020**, *11*, 4540. [[CrossRef](#)]
22. Wenger, S.J.; Olden, J.D. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods Ecol. Evol.* **2012**, *3*, 260–267. [[CrossRef](#)]
23. Hao, T.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography* **2020**, *43*, 549–558. [[CrossRef](#)]
24. Dolan, M.F.J.; Ross, R.E.; Albretsen, J.; Skarðhamar, J.; Gonzalez-Mirelis, G.; Bellec, V.K.; Buhl-Mortensen, P.; Bjarnadóttir, L.R. Using Spatial Validity and Uncertainty Metrics to Determine the Relative Suitability of Alternative Suites of Oceanographic Data for Seabed Biotope Prediction. A Case Study from the Barents Sea, Norway. *Geosciences* **2021**, *11*, 48. [[CrossRef](#)]
25. Schratz, P.; Muenchow, J.; Iturrutxa, E.; Richter, J.; Brenning, A. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecol. Modell.* **2019**, *406*, 109–120. [[CrossRef](#)]
26. Pohjankukka, J.; Pahikkala, T.; Nevalainen, P.; Heikkonen, J. Estimating the prediction performance of spatial models via spatial k-fold cross validation. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 2001–2019. [[CrossRef](#)]
27. Parmentier, I.; Harrigan, R.J.; Buermann, W.; Mitchard, E.T.A.; Saatchi, S.; Malhi, Y.; Bongers, F.; Hawthorne, W.D.; Leal, M.E.; Lewis, S.L.; et al. Predicting alpha diversity of African rain forests: Models based on climate and satellite-derived data do not perform better than a purely spatial model. *J. Biogeogr.* **2011**, *38*, 1164–1176. [[CrossRef](#)]
28. Trachsel, M.; Telford, R.J. Technical note: Estimating unbiased transfer-function performances in spatially structured environments. *Clim. Past* **2016**, *12*, 1215–1223. [[CrossRef](#)]

29. Le Rest, K.; Pinaud, D.; Monestiez, P.; Chadoeuf, J.; Bretagnolle, V. Spatial leave-one-out cross-validation for variable selection in the presence of spatial autocorrelation. *Glob. Ecol. Biogeogr.* **2014**, *23*, 811–820. [[CrossRef](#)]
30. Ruß, G.; Brenning, A. Spatial Variable Importance Assessment for Yield Prediction in Precision Agriculture. In *Advances in Intelligent Data Analysis IX*; Lecture Notes in Computer Science; Cohen, P.R., Adams, N.M., Berthold, M.R., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6065. [[CrossRef](#)]
31. Valavi, R.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. blockCV: An r package for generating spatially or environmentally separated folds for *k*-fold cross-validation of species distribution models. *Methods Ecol. Evol.* **2019**, *10*, 225–232. [[CrossRef](#)]
32. Meyer, H.; Reudenbach, C.; Wöllauer, S.; Nauss, T. Importance of spatial predictor variable selection in machine learning applications—Moving from data reproduction to spatial prediction. *Ecol. Modell.* **2019**, *411*, 108815. [[CrossRef](#)]
33. Randin, C.F.; Dirnböck, T.; Dullinger, S.; Zimmermann, N.E.; Zappa, M.; Guisan, A. Are niche-based species distribution models transferable in space? *J. Biogeogr.* **2006**, *33*, 1689–1703. [[CrossRef](#)]
34. Yates, K.L.; Bouchet, P.J.; Caley, M.J.; Mengersen, K.; Randin, C.F.; Parnell, S.; Fielding, A.H.; Bamford, A.J.; Ban, S.; Barbosa, A.M.; et al. Outstanding Challenges in the Transferability of Ecological Models. *Trends Ecol. Evol.* **2018**, *33*, 790–802. [[CrossRef](#)]
35. Meyer, H.; Pebesma, E. Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods Ecol. Evol.* **2021**, *12*, 2041–2210. [[CrossRef](#)]
36. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013; ISBN 978-1-4614-6848-6.
37. Elith, J.; Kearney, M.; Phillips, S. The art of modelling range-shifting species. *Methods Ecol. Evol.* **2010**, *1*, 330–342. [[CrossRef](#)]
38. Zurell, D.; Elith, J.; Schröder, B. Predicting to new environments: Tools for visualizing model behaviour and impacts on mapped distributions. *Divers. Distrib.* **2012**, *18*, 628–634. [[CrossRef](#)]
39. Owens, H.L.; Campbell, L.P.; Dornak, L.L.; Saupe, E.E.; Barve, N.; Soberón, J.; Ingenloff, K.; Lira-Noriega, A.; Hensz, C.M.; Myers, C.E.; et al. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecol. Modell.* **2013**, *263*, 10–18. [[CrossRef](#)]
40. Mesgaran, M.B.; Cousens, R.D.; Webber, B.L. Here be dragons: A tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Divers. Distrib.* **2014**, *20*, 1147–1159. [[CrossRef](#)]
41. Rödder, D.; Engler, J.O. Disentangling Interpolation and Extrapolation Uncertainties in Species Distribution Models: A Novel Visualization Technique for the Spatial Variation of Predictor Variable Colinearity. *Biodivers. Inform.* **2012**, *8*, 4326. [[CrossRef](#)]
42. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biom. Bull.* **1945**, *1*, 80. [[CrossRef](#)]
43. Mann, H.B.; Whitney, D.R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [[CrossRef](#)]
44. Kruskal, W.H. Historical Notes on the Wilcoxon Unpaired Two-Sample Test. *J. Am. Stat. Assoc.* **1957**, *52*, 356. [[CrossRef](#)]
45. Kursá, M.B.; Rudnicki, W.R. Feature Selection with the Boruta Package. *J. Stat. Softw.* **2010**, *36*, 11. [[CrossRef](#)]
46. Kursá, M.B. Robustness of Random Forest-based gene selection methods. *BMC Bioinform.* **2014**, *15*, 8. [[CrossRef](#)] [[PubMed](#)]
47. Degenhardt, F.; Seifert, S.; Szymczak, S. Evaluation of variable selection methods for random forests and omics data sets. *Brief. Bioinform.* **2019**, *20*, 492–503. [[CrossRef](#)] [[PubMed](#)]
48. Li, J.; Tran, M.; Siwabessy, J. Selecting Optimal Random Forest Predictive Models: A Case Study on Predicting the Spatial Distribution of Seabed Hardness. *PLoS ONE* **2016**, *11*, e0149089. [[CrossRef](#)]
49. Li, J.; Alvarez, B.; Siwabessy, J.; Tran, M.; Huang, Z.; Przeslawski, R.; Radke, L.; Howard, F.; Nichol, S. Application of random forest, generalised linear model and their hybrid methods with geostatistical techniques to count data: Predicting sponge species richness. *Environ. Model. Softw.* **2017**, *97*, 112–129. [[CrossRef](#)]
50. Li, J. A Critical Review of Spatial Predictive Modeling Process in Environmental Sciences with Reproducible Examples in R. *Appl. Sci.* **2019**, *9*, 2048. [[CrossRef](#)]
51. Diesing, M.; Thorsnes, T. Mapping of Cold-Water Coral Carbonate Mounds Based on Geomorphometric Features: An Object-Based Approach. *Geosciences* **2018**, *8*, 34. [[CrossRef](#)]
52. Diesing, M.; Mitchell, P.J.; O’Keeffe, E.; Gavazzi, G.O.A.M.; Bas, T. Le Limitations of Predicting Substrate Classes on a Sedimentary Complex but Morphologically Simple Seabed. *Remote Sens.* **2020**, *12*, 3398. [[CrossRef](#)]
53. Diesing, M. Deep-sea sediments of the global ocean. *Earth Syst. Sci. Data* **2020**, *12*, 3367–3381. [[CrossRef](#)]
54. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
55. Meinshausen, N. Quantile regression forests. *J. Mach. Learn. Res.* **2006**, *7*, 983–999.
56. Kirkwood, C.; Cave, M.; Beamish, D.; Grebby, S.; Ferreira, A. A machine learning approach to geochemical mapping. *J. Geochem. Explor.* **2016**, *167*, 49–61. [[CrossRef](#)]
57. Vaysse, K.; Lagacherie, P. Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma* **2017**, *291*, 55–64. [[CrossRef](#)]
58. Fouedjio, F.; Klump, J. Exploring prediction uncertainty of spatial data in geostatistical and machine learning approaches. *Environ. Earth Sci.* **2019**, *78*, 38. [[CrossRef](#)]
59. Szatmári, G.; Pásztor, L. Comparison of various uncertainty modelling approaches based on geostatistics and machine learning algorithms. *Geoderma* **2019**, *337*, 1329–1340. [[CrossRef](#)]
60. Diesing, M.; Kröger, S.; Parker, R.; Jenkins, C.; Mason, C.; Weston, K. Predicting the standing stock of organic carbon in surface sediments of the North–West European continental shelf. *Biogeochemistry* **2017**, *135*, 183–200. [[CrossRef](#)]

61. Baker, E.; Beaudoin, Y. *Deep Sea Minerals: A Physical, Biological, Environmental, and Technical Review*; Secretariat of the Pacific Community: Suva, Fiji, 2013; ISBN 978-827-701-12-02.
62. Marchig, V.; Reyss, J.L. Diagenetic mobilization of manganese in Peru Basin sediments. *Geochim. Cosmochim. Acta* **1984**, *48*, 1349–1352. [[CrossRef](#)]
63. Von Stackelberg, U. Growth history of manganese nodules and crusts of the Peru Basin. *Geol. Soc. Lond. Spec. Publ.* **1997**, *119*, 153–176. [[CrossRef](#)]
64. Weber, M.; von Stackelberg, U.; Marchig, V.; Wiedicke, M.; Grupe, B. Variability of surface sediments in the Peru basin: Dependence on water depth, productivity, bottom water flow, and seafloor topography. *Mar. Geol.* **2000**, *163*, 169–184. [[CrossRef](#)]
65. Toro, N.; Jeldres, R.I.; Órdenes, J.A.; Robles, P.; Navarra, A. Manganese Nodules in Chile, an Alternative for the Production of Co and Mn in the Future—A Review. *Minerals* **2020**, *10*, 674. [[CrossRef](#)]
66. Thiel, H.; Schriever, G.; Ahnert, A.; Bluhm, H.; Borowski, C.; Vopel, K. The large-scale environmental impact experiment DISCOL—reflection and foresight. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2001**, *48*, 3869–3882. [[CrossRef](#)]
67. Gausepohl, F.; Hennke, A.; Schoening, T.; Köser, K.; Greinert, J. Scars in the abyss: Reconstructing sequence, location and temporal change of the 78 plough tracks of the 1989 DISCOL deep-sea disturbance experiment in the Peru Basin. *Biogeosciences* **2020**, *17*, 1463–1493. [[CrossRef](#)]
68. Wiedicke, M.H.; Weber, M.E. Small-scale variability of seafloor features in the northern Peru Basin: Results from acoustic survey methods. *Mar. Geophys. Res.* **1996**, *18*, 507–526. [[CrossRef](#)]
69. Paul, S.A.L.; Haeckel, M.; Bau, M.; Bajracharya, R.; Koschinsky, A. Small-scale heterogeneity of trace metals including rare earth elements and yttrium in deep-sea sediments and porewaters of the Peru Basin, southeastern equatorial Pacific. *Biogeosciences* **2019**, *16*, 4829–4849. [[CrossRef](#)]
70. Grupe, B.; Becker, H.J.; Oebius, H.U. Geotechnical and sedimentological investigations of deep-sea sediments from a manganese nodule field of the Peru Basin. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2001**, *48*, 3593–3608. [[CrossRef](#)]
71. Klein, H. Near-bottom currents in the deep Peru Basin, DISCOL experimental area. *Dtsch. Hydrogr. Z.* **1993**, *45*, 31–42. [[CrossRef](#)]
72. Klein, H. Near-bottom currents and bottom boundary layer variability over manganese nodule fields in the peru basin, se-pacific. *Dtsch. Hydrogr. Z.* **1996**, *48*, 147–160. [[CrossRef](#)]
73. Flood, R.D. Classification of sedimentary furrows and a model for furrow initiation and evolution. *Geol. Soc. Am. Bull.* **1983**, *94*, 630. [[CrossRef](#)]
74. Lonsdale, P.; Spiess, F.N. Abyssal Bedforms Explored with a Deeply Towed Instrument Package. *Dev. Sedimentol.* **1977**, *23*, 57–75.
75. Flood, R.D.; Hollister, C.D. Submersible studies of deep-sea furrows and transverse ripples in cohesive sediments. *Mar. Geol.* **1980**, *36*, M1–M9. [[CrossRef](#)]
76. Haeckel, M.; König, I.; Riech, V.; Weber, M.E.; Suess, E. Pore water profiles and numerical modelling of biogeochemical processes in Peru Basin deep-sea sediments. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2001**, *48*, 3713–3736. [[CrossRef](#)]
77. Greinert, J. *RV Sonne Fahrtbericht/Cruise Report SO242-1 [SO242/1], JPI Oceans Ecological Aspects of Deep-Sea Mining, DISCOL Revisited, Guayaquil-Guayaquil, 28 July–25 August 2015*; GEOMAR Report, N. Ser. 026; GEOMAR Helmholtz-Zentrum für Ozeanforschung; Kiel, Germany, 2015; Volume 7.
78. Benites, M.; Millo, C.; Hein, J.; Nath, B.; Murton, B.; Galante, D.; Jovane, L. Integrated Geochemical and Morphological Data Provide Insights into the Genesis of Ferromanganese Nodules. *Minerals* **2018**, *8*, 488. [[CrossRef](#)]
79. Burdige, D.J. The biogeochemistry of manganese and iron reduction in marine sediments. *Earth-Sci. Rev.* **1993**, *35*, 249–284. [[CrossRef](#)]
80. Linke, P.; Lackschewitz, K. Autonomous Underwater Vehicle “ABYSS”. *J. Large-Scale Res. Facil.* **2016**, *2*, A79. [[CrossRef](#)]
81. Klischies, M.; Rothenbeck, M.; Steinfuhrer, A.; Yeo, I.A.; dos Santos Ferreira, C.; Mohrmann, J.; Faber, C.; Schirmick, C. AUV Abyss workflow: Autonomous deep sea exploration for ocean research. In Proceedings of the 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV), Porto, Portugal, 6–9 November 2018; pp. 1–6.
82. Caress, D.W.; Chayes, D.N. MB-System: Mapping the Seafloor. 2017. Available online: <http://www.mbari.org/products/research-software/mb-system/> (accessed on 18 October 2021).
83. Alevizos, E.; Schoening, T.; Koeser, K.; Snellen, M.; Greinert, J. Quantification of the fine-scale distribution of Mn-nodules: Insights from AUV multi-beam and optical imagery data fusion. *Biogeosciences* **2018**, 1–29. [[CrossRef](#)]
84. Lecours, V.; Dolan, M.F.J.; Micallef, A.; Lucieer, V.L. A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 3207–3244. [[CrossRef](#)]
85. Iwahashi, J.; Pike, R.J. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology* **2007**, *86*, 409–440. [[CrossRef](#)]
86. Dolan, M.F.J.; Lucieer, V.L. Variation and Uncertainty in Bathymetric Slope Calculations Using Geographic Information Systems. *Mar. Geod.* **2014**, *37*, 187–219. [[CrossRef](#)]
87. Naimi, B.; Skidmore, A.K.; Groen, T.A.; Hamm, N.A.S. Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *J. Biogeogr.* **2011**, *38*, 1497–1509. [[CrossRef](#)]
88. Stephens, D.; Diesing, M. A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data. *PLoS ONE* **2014**, *9*, e93950. [[CrossRef](#)]
89. Lucieer, V.; Huang, Z.; Siwabessy, J. Analyzing Uncertainty in Multibeam Bathymetric Data and the Impact on Derived Seafloor Attributes. *Mar. Geod.* **2016**, *39*, 32–52. [[CrossRef](#)]

90. Lecours, V.; Devillers, R.; Edinger, E.N.; Brown, C.J.; Lucieer, V.L. Influence of artefacts in marine digital terrain models on habitat maps and species distribution models: A multiscale assessment. *Remote Sens. Ecol. Conserv.* **2017**, *3*, 232–246. [CrossRef]
91. Hughes Clarke, J. The Impact of Acoustic Imaging Geometry on the Fidelity of Seabed Bathymetric Models. *Geosciences* **2018**, *8*, 109. [CrossRef]
92. Florinsky, I.V. An illustrated introduction to general geomorphometry. *Prog. Phys. Geogr.* **2017**, *41*, 723–752. [CrossRef]
93. Misiuk, B.; Lecours, V.; Bell, T. A multiscale approach to mapping seabed sediments. *PLoS ONE* **2018**, *13*, e0193647. [CrossRef]
94. Cremers, J.; Klugkist, I. One Direction? A Tutorial for Circular Data Analysis Using R With Examples in Cognitive Psychology. *Front. Psychol.* **2018**, *9*. [CrossRef] [PubMed]
95. Zevenbergen, L.W.; Thorne, C.R. Quantitative analysis of land surface topography. *Earth Surf. Process. Landf.* **1987**, *12*, 47–56. [CrossRef]
96. Olaya, V. Chapter 6 Basic Land-Surface Parameters. *Dev. Soil Sci.* **2009**, *33*, 141–169.
97. Sappington, J.M.; Longshore, K.M.; Thompson, D.B. Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert. *J. Wildl. Manage.* **2007**, *71*, 1419–1426. [CrossRef]
98. Weiss, A. Topographic position and landforms analysis. *Poster Present. ESRI User Conf.* **2001**, *64*, 227–245.
99. Wilson, M.F.J.; O’Connell, B.; Brown, C.; Guinan, J.C.; Grehan, A.J. Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. *Mar. Geod.* **2007**, *30*, 3–35. [CrossRef]
100. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man. Cybern.* **1973**, *SMC-3*, 610–621. [CrossRef]
101. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* **2015**, *8*, 1991–2007. [CrossRef]
102. Walbridge, S.; Slocum, N.; Pobuda, M.; Wright, D. Unified Geomorphological Analysis Workflows with Benthic Terrain Modeler. *Geosciences* **2018**, *8*, 94. [CrossRef]
103. Hijmans, R.J. Raster: Geographic Data Analysis and Modeling. R Package Version 3.4-13. 2021. Available online: <https://CRAN.R-project.org/package=raster> (accessed on 19 October 2021).
104. Zvoleff, A. glm: Calculate Textures from Grey-Level Co-Occurrence Matrices (GLCMs). R Package Version 1.6.5. 2020. Available online: <https://CRAN.R-project.org/package=glm> (accessed on 19 October 2021).
105. Kwasnitschka, T.; Köser, K.; Sticklus, J.; Rothenbeck, M.; Weiß, T.; Wenzlaff, E.; Schoening, T.; Triebe, L.; Steinführer, A.; Devey, C.; et al. DeepSurveyCam—A Deep Ocean Optical Mapping System. *Sensors* **2016**, *16*, 164. [CrossRef]
106. Ellefmo, S.L.; Kuhn, T. Application of Soft Data in Nodule Resource Estimation. *Nat. Resour. Res.* **2021**, *30*, 1069–1091. [CrossRef]
107. Wasilewska-Błaszczuk, M.; Mucha, J. Possibilities and Limitations of the Use of Seafloor Photographs for Estimating Polymetallic Nodule Resources—Case Study from IOM Area, Pacific Ocean. *Minerals* **2020**, *10*, 1123. [CrossRef]
108. Yu, G.; Parianos, J. Empirical Application of Generalized Rayleigh Distribution for Mineral Resource Estimation of Seabed Polymetallic Nodules. *Minerals* **2021**, *11*, 449. [CrossRef]
109. Tsune, A. Quantitative Expression of the Burial Phenomenon of Deep Seafloor Manganese Nodules. *Minerals* **2021**, *11*, 227. [CrossRef]
110. Simon-Lledó, E.; Bett, B.J.; Huvenne, V.A.I.; Schoening, T.; Benoist, N.M.A.; Jones, D.O.B. Ecology of a polymetallic nodule occurrence gradient: Implications for deep-sea mining. *Limnol. Oceanogr.* **2019**, *64*, 1883–1894. [CrossRef]
111. Caldas de Castro, M.; Singer, B.H. Controlling the False Discovery Rate: A New Application to Account for Multiple and Dependent Tests in Local Statistics of Spatial Association. *Geogr. Anal.* **2006**, *38*, 180–208. [CrossRef]
112. Benjamini FDR\_Benjamin\_1995. *Ital. J. Food Sci.* **2009**, *21*, 89–95.
113. Sullivan, G.M.; Feinn, R. Using Effect Size—or Why the *p* Value Is Not Enough. *J. Grad. Med. Educ.* **2012**, *4*, 279–282. [CrossRef]
114. R, Core, T. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2021. Available online: <https://www.R-project.org/> (accessed on 19 October 2021).
115. Kassambara, A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R Package Version 0.7.0. 2021. Available online: <https://CRAN.R-project.org/package=rstatix> (accessed on 19 October 2021).
116. Spearman, C. The proof and measurement of association between two things. *Int. J. Epidemiol.* **2010**, *39*, 1137–1150. [CrossRef]
117. Makowski, D.; Ben-Shachar, M.; Patil, I.; Lüdecke, D. Methods and Algorithms for Correlation Analysis in R. *J. Open Source Softw.* **2020**, *5*, 2306. [CrossRef]
118. Mukaka, M.M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* **2012**, *24*, 69–71.
119. Schloerke, B.; Cook, D.; Larmarange, J.; Briatte, F.; Marbach, M.; Thoen, E.; Elberg, A.; Toomet, O.; Crowley, J.; Hofman, H.; et al. GGally: Extension to “ggplot2”. R Package Version 2.1.2. 2021. Available online: <https://CRAN.R-project.org/package=GGally> (accessed on 19 October 2021).
120. Probst, P.; Wright, M.N.; Boulesteix, A. Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2019**, *9*, 1301. [CrossRef]
121. Kuhn, M. caret: Classification and Regression Training. R package version 6.0-88. 2021. Available online: <https://CRAN.R-project.org/package=caret> (accessed on 19 October 2021).
122. Greenwell, B.M. pdp: An R Package for Constructing Partial Dependence Plots. *R J.* **2017**, *9*, 421–436. Available online: <https://journal.r-project.org/archive/2017/RJ-2017-016/index.html> (accessed on 19 October 2021). [CrossRef]

123. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An introduction to Statistical Learning. *Curr. Med. Chem.* **2000**, *7*, 995–1039. [[CrossRef](#)]
124. Kaufman, L.; Rousseeuw, P.J. Clustering Large Applications (Program CLARA). In *Finding Groups in Data: An Introduction to Cluster Analysis*; Wiley: Hoboken, NJ, USA, 1990; pp. 126–163.
125. Kaufman, L.; Rousseeuw, P.J. Partitioning Around Medoids (Program PAM). In *Finding Groups in Data: An Introduction to Cluster Analysis*; Wiley: Hoboken, NJ, USA, 1990; pp. 68–125.
126. Calinski, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.—Theory Methods* **1974**, *3*, 1–27. [[CrossRef](#)]
127. Maechler, M.; Rousseeuw, P.; Struyf, A.; Hubert, M.; Hornik, K. Cluster: Cluster Analysis Basics and Extensions. R Package Version 2.1.2. 2021. Available online: <https://CRAN.R-project.org/package=cluster> (accessed on 19 October 2021).
128. Desgraupes, B. clusterCrit: Clustering Indices. R Package Version 1.2.8. 2018. Available online: <https://CRAN.R-project.org/package=clusterCrit> (accessed on 19 October 2021).
129. Leutner, B.; Horning, N.; Schwalb-Willmann, J.; Hijmans, R.J. RStoolbox: Tools for Remote Sensing Data Analysis. R Package Version 0.2.6. 2019. Available online: <https://CRAN.R-project.org/package=RStoolbox> (accessed on 19 October 2021).
130. Meyer, H.; Reudenbach, C.; Ludwig, M.; Nauss, T.; Pebesma, E. CAST: “caret” Applications for Spatial-Temporal Models. R Package Version 0.5.1. 2021. Available online: <https://CRAN.R-project.org/package=CAST> (accessed on 19 October 2021).
131. Friedman Jerome, H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232.
132. Molnar, C. Interpretable Machine Learning. A Guide for Making Black Box Models Explainable. 2019. Available online: <https://christophm.github.io/interpretable-ml-book/> (accessed on 19 October 2021).
133. Cleveland, W.S. LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression. *Am. Stat.* **1981**, *35*, 54. [[CrossRef](#)]
134. Verlaan, P.A.; Cronan, D.S. Origin and variability of resource-grade marine ferromanganese nodules and crusts in the Pacific Ocean: A review of biogeochemical and physical controls. *Geochemistry* **2021**, 125741. [[CrossRef](#)]
135. Kuhn, T.; Wegorzewski, A.; Rühlemann, C.; Vink, A. Composition, Formation, and Occurrence of Polymetallic Nodules BT—Deep-Sea Mining: Resource Potential, Technical and Environmental Considerations. In *Deep-Sea Mining*; Sharma, R., Ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 23–63. ISBN 978-3-319-52557-0.
136. Skowronek, A.; Maciag, Ł.; Zawadzki, D.; Strzelecka, A.; Baláž, P.; Mianowicz, K.; Abramowski, T.; Konečný, P.; Krawcewicz, A. Chemostratigraphic and Textural Indicators of Nucleation and Growth of Polymetallic Nodules from the Clarion-Clipperton Fracture Zone (IOM Claim Area). *Minerals* **2021**, *11*, 868. [[CrossRef](#)]
137. Hengl, T.; Heuvelink, G.B.M.; Rossiter, D.G. About regression-kriging: From equations to case studies. *Comput. Geosci.* **2007**, *33*, 1301–1315. [[CrossRef](#)]
138. Lobo, J.M. More complex distribution models or more representative data? *Biodivers. Inform.* **2008**, *5*, 40. [[CrossRef](#)]
139. Mets, K.D.; Armenteras, D.; Dávalos, L.M. Spatial autocorrelation reduces model precision and predictive power in deforestation analyses. *Ecosphere* **2017**, *8*, e01824. [[CrossRef](#)]
140. Hengl, T.; Walsh, M.G.; Sanderman, J.; Wheeler, I.; Harrison, S.P.; Prentice, I.C. Global mapping of potential natural vegetation: An assessment of machine learning algorithms for estimating land potential. *PeerJ* **2018**, *6*, e5457. [[CrossRef](#)] [[PubMed](#)]
141. Robert, K.; Jones, D.O.B.; Roberts, J.M.; Huvenne, V.A.I. Improving predictive mapping of deep-water habitats: Considering multiple model outputs and ensemble techniques. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **2016**, *113*, 80–89. [[CrossRef](#)]
142. Wang, J.-F.; Stein, A.; Gao, B.-B.; Ge, Y. A review of spatial sampling. *Spat. Stat.* **2012**, *2*, 1–14. [[CrossRef](#)]
143. Li, J.; Heap, A.D. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecol. Inform.* **2011**, *6*, 228–241. [[CrossRef](#)]
144. Hengl, T.; Rossiter, D.G.; Stein, A. Soil sampling strategies for spatial prediction by correlation with auxiliary maps. *Soil Res.* **2003**, *41*, 1403. [[CrossRef](#)]
145. Brus, D.J. Sampling for digital soil mapping: A tutorial supported by R scripts. *Geoderma* **2019**, *338*, 464–480. [[CrossRef](#)]
146. Malone, B.P.; Minansy, B.; Brungard, C. Some methods to improve the utility of conditioned Latin hypercube sampling. *PeerJ* **2019**, *7*, e6451. [[CrossRef](#)]
147. Foster, S.D.; Hosack, G.R.; Hill, N.A.; Barrett, N.S.; Lucieer, V.L. Choosing between strategies for designing surveys: Autonomous underwater vehicles. *Methods Ecol. Evol.* **2014**, *5*, 287–297. [[CrossRef](#)]
148. Yilmaz, N.K.; Evangelinos, C.; Lermusiaux, P.; Patrikalakis, N.M. Path Planning of Autonomous Underwater Vehicles for Adaptive Sampling Using Mixed Integer Linear Programming. *IEEE J. Ocean. Eng.* **2008**, *33*, 522–537. [[CrossRef](#)]
149. Foster, S.D.; Hosack, G.R.; Monk, J.; Lawrence, E.; Barrett, N.S.; Williams, A.; Przeslawski, R. Spatially balanced designs for transect-based surveys. *Methods Ecol. Evol.* **2020**, *11*, 95–105. [[CrossRef](#)]
150. Hughes, R.N.; Hughes, D.J.; Smith, I.P.; Dale, A.C. (Eds.) *Oceanography and Marine Biology*; CRC Press: Boca Raton, FL, USA, 2016; ISBN 978-131-536-8-597.
151. Schmidt, K.; Behrens, T.; Daumann, J.; Ramirez-Lopez, L.; Werban, U.; Dietrich, P.; Scholten, T. A comparison of calibration sampling schemes at the field scale. *Geoderma* **2014**, 232–234, 243–256. [[CrossRef](#)]
152. Wadoux, A.M.-C.; Brus, D.J.; Heuvelink, G.B.M. Sampling design optimization for soil mapping with random forest. *Geoderma* **2019**, *355*, 113913. [[CrossRef](#)]
153. Bowden, D.A.; Anderson, O.F.; Rowden, A.A.; Stephenson, F.; Clark, M.R. Assessing Habitat Suitability Models for the Deep Sea: Is Our Ability to Predict the Distributions of Seafloor Fauna Improving? *Front. Mar. Sci.* **2021**, *8*, 632389. [[CrossRef](#)]

154. Fernández-Delgado, M.; Sirsat, M.S.; Cernadas, E.; Alawadi, S.; Barro, S.; Febrero-Bande, M. An extensive experimental survey of regression methods. *Neural Netw.* **2019**, *111*, 11–34. [[CrossRef](#)] [[PubMed](#)]
155. Merow, C.; Smith, M.J.; Edwards, T.C.; Guisan, A.; McMahon, S.M.; Normand, S.; Thuiller, W.; Wüest, R.O.; Zimmermann, N.E.; Elith, J. What do we gain from simplicity versus complexity in species distribution models? *Ecography* **2014**, *37*, 1267–1281. [[CrossRef](#)]
156. Bochare, A.; Gangopadhyay, A.; Yesha, Y.; Joshi, A.; Yesha, Y.; Brady, M.; Grasso, M.A.; Rische, N. Integrating domain knowledge in supervised machine learning to assess the risk of breast cancer. *Int. J. Med. Eng. Inform.* **2014**, *6*, 87. [[CrossRef](#)]
157. Guan, X.; Runger, G.; Liu, L. Dynamic incorporation of prior knowledge from multiple domains in biomarker discovery. *BMC Bioinform.* **2020**, *21*, 77. [[CrossRef](#)]
158. Lauria, V.; Power, A.M.; Lordan, C.; Weetman, A.; Johnson, M.P. Spatial Transferability of Habitat Suitability Models of *Nephrops norvegicus* among Fished Areas in the Northeast Atlantic: Sufficiently Stable for Marine Resource Conservation? *PLoS ONE* **2015**, *10*, e0117006. [[CrossRef](#)] [[PubMed](#)]
159. Shmueli, G. To Explain or to Predict? *Stat. Sci.* **2010**, *25*, 330. [[CrossRef](#)]
160. Breiman, L. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Stat. Sci.* **2001**, *16*, 726. [[CrossRef](#)]



Original Paper

# An Interpretable Multi-Model Machine Learning Approach for Spatial Mapping of Deep-Sea Polymetallic Nodule Occurrences

Iason-Zois Gazis<sup>1,4</sup>, Francois Charlet,<sup>2</sup> and Jens Greinert<sup>1,3</sup>

Received 27 March 2024; accepted 19 July 2024

High-resolution mapping of deep-sea polymetallic nodules is needed (a) to understand the reasons behind their patchy distribution, (b) to associate nodule coverage with benthic fauna occurrences, and (c) to enable an accurate resource estimation and mining path planning. This study used an autonomous underwater vehicle to map 37 km<sup>2</sup> of a geomorphologically complex site in the Eastern Clarion–Clipperton Fracture Zone. A multibeam echosounder system (MBES) at 400 kHz and a side scan sonar at 230 kHz were used to investigate the nodule backscatter response. More than 30,000 seafloor images were analyzed to obtain the nodule coverage and train five machine learning (ML) algorithms: generalized linear models, generalized additive models, support vector machines, random forests (RFs) and neural networks (NNs). All models ML yielded similar maps of nodule coverage with differences occurring in the range of predicted values, particularly at parts with irregular topography. RFs had the best fit and NNs had the worst spatial transferability. Attention was given to the interpretability of model outputs using variable importance ranking across all models, partial dependence plots and domain knowledge. The nodule coverage is higher on relatively flat seafloor (< 3°) with eastward-facing slopes. The most important predictor was the MBES backscatter, particularly from incident angles between 25 and 55°. Bathymetry, slope, and slope orientation were important geomorphological predictors. For the first time, at a water depth of 4500 m, orthophoto-mosaics and image-derived digital elevation models with 2-mm and 5-mm spatial resolutions supported the geomorphological analysis, interpretation of polymetallic nodules occurrences, and backscatter response.

**KEY WORDS:** Backscatter, Nodule coverage, Machine learning, Orthophoto-mosaics.

## INTRODUCTION

Deep-sea polymetallic nodules (hereafter nodules) are mineral concretions of mm to cm in diameter, embedded in the sediment surface of the ocean floor, mainly between 4000 and 6000 m water depth. They are formed by precipitation of Fe- and Mn-oxyhydroxides from seawater and sediment porewater, creating repeated concentric encrustations around a core (e.g., nodule fragment, rock debris, and shark's tooth) over millions of years

<sup>1</sup>GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany.

<sup>2</sup>DEME Group, GSR Global Sea Minerals Resources NV, Antwerp, Belgium.

<sup>3</sup>Christian-Albrechts University Kiel, Kiel, Germany.

<sup>4</sup>To whom correspondence should be addressed; e-mail: igazis@geomar.de

(Kuhn et al., 2017; Hein et al., 2020). Within the Clarion–Clipperton fracture zone (CCFZ) in the eastern equatorial Pacific Ocean, nodule abundances are typically 10–15 kg m<sup>-2</sup> (wet weight). Nodules are rich in Mn (28.4 wt%) and Fe (6.16 wt%) while they also have high contents in Ni, Co, and Cu, which are usually > 2.5 wt% for these three metals combined (Kuhn et al., 2017; Hein et al., 2020). Due to their abundance, metal content, and the currently increasing demand for those metals (e.g., batteries for electric vehicles), nodules are of commercial interest, and 18 exploration licenses in the CCFZ have been issued and supervised by the International Seabed Authority (ISA, 2023). Apart from a potential metal resource, nodules are central to the deep-sea ecosystem as they are one of the few hard substrates on vast regions covered by otherwise soft siliceous clay. Several studies have linked nodule abundance (kg m<sup>-2</sup>), seafloor area coverage (% m<sup>-2</sup>), and nodule size (cm) to the abundance, spatial distribution, species richness, and community composition of benthic fauna (Amon et al., 2016; Vanreusel et al., 2016; Simon-Lledó et al., 2019a, b, 2020; Pape et al., 2021; Mbani et al., 2023; Uhlenkott et al., 2023). Because nodules are a valuable metal resource and deep-sea habitat, the study of their spatial distribution has been advanced.

In regional scales (hundreds to thousands km), ship-based multibeam echosounder systems (MBESs) and deep-towed side scan sonar (SSS) systems have been used to discriminate areas of nodule presence and absence using the backscatter response, typically at 12 kHz (Scanlon & Masson, 1992; Lee & Kim, 2004; Chunhui et al., 2015; Machida et al., 2019; Kuhn & Rühlemann, 2021; Wang et al., 2021). These data are linked with sparse (usually > 1.8 km apart) ground truth samples, e.g., box corers (sampling area of 0.25 m<sup>2</sup>) and geostatistical techniques (e.g., kriging) to provide nodule distribution and abundance (Kuhn & Rühlemann, 2021). Within contract areas, high-resolution seafloor mapping with autonomous underwater vehicles (AUVs) has revealed the fine-scale (meters to few km) spatial distribution of nodules (Schoening et al., 2017; Gazis et al., 2018; Peukert et al., 2018; Gazis & Greinert, 2021; Alevizos et al., 2022) and associated fauna (Simon-Lledó et al., 2019a,b; De Smet et al., 2021). The large amounts of hydroacoustic and optic data collected during such surveys have driven the application of machine learning (ML) at local (Gazis et al., 2018; Gazis & Greinert, 2021; Wong et al., 2021; Mbani et al., 2022), regional (Hari et al., 2018;

Kaikkonen et al., 2019; Kuhn & Rühlemann, 2021; Wasilewska-Błaszczuk & Mucha, 2021) and global scales (Dutkiewicz et al., 2020). ML provides reliable predictions for geographical areas where good correlations between the response variable and covariates/predictors exist, and an adequate feature space representation for the training data is guaranteed (Gazis & Greinert, 2021).

Regional and local seafloor parameters influence directly the nodule spatial distribution, e.g., bathymetry (Craig, 1979; Kodagali, 1988; von Stackelberg, 2000), terrain ruggedness (Kodagali, 1988; Skornyakova & Murdmaa, 1992), outcropping bedrock (Dreisettl, 2016; Yoo et al., 2018; Alevizos et al., 2022), and volcanism and availability of nuclei (Glasby et al., 1973). Synergistically, the seafloor morphology alters the local hydrodynamics, e.g., when solitary knolls or seamount chains influence the bottom current magnitude and direction, causing erosive and depositional structures such as furrows and sediment drifts, respectively (Juan et al., 2018). The slope orientation (i.e., aspect) has been used as a proxy for the mean current direction in spatial models when continuous long-term information about bottom currents is unavailable (Gazis & Greinert, 2021; Uhlenkott et al., 2022). Alternations in local sediment depositional patterns that influence sediment thickness (e.g., seafloor pits with increased sedimentation and conical hills with sediment accumulation at the lee side) contribute to the patchy nodule distribution (Wiedicke & Weber, 1996; von Stackelberg, 2000; Juan et al., 2018; Gazis & Greinert, 2021). Such variations in sediment thickness are reflected in the MBES backscatter intensity, angular range analysis (ARA), and SSS backscatter mosaic (Yoo et al., 2018; Hillman et al., 2017; Gaida et al., 2019), while they can be inferred from topographic indices such as the bathymetric position index (BPI) (Gazis & Greinert, 2021; Alevizos et al., 2022) and can be measured in geophysical data obtained by sub-bottom profiler (SBP), e.g., thickness of semi-liquid bottom sediment layer (Cochonat et al., 1992; Yoo et al., 2018; Alevizos et al., 2022). The synergistic effect of seabed topography, bottom currents, and heterogeneous sedimentation patterns influences the geochemical parameters that affect nodule formation, e.g., removal of clay-size particles by currents, variations in total organic carbon content and shifts on Mn-redox depth in local depressions, highlighting the relationship between meter-scale seabed morphology and variation in geo-

## An Interpretable Multi-Model Machine Learning Approach

chemical parameters (Mewes et al., 2014; Volz et al., 2018; Paul et al., 2019).

Despite the ability to derive accurate predictions based on data that capture the underlying relationships, different ML algorithms result in divergent prediction performance even when using the same dataset (Stephens & Diesing, 2014; Diesing & Stephens, 2015; Robert et al., 2016; Li et al., 2017). Multi-model approaches could help decrease the prediction error, identify models that better generalize over broad geographical areas or a more complex feature space, and increase confidence in locations with good spatial agreement (Wenger & Olden, 2012; Diesing & Stephens, 2015; Robert et al., 2016). To our knowledge, the multi-model approach has not been used until now for fine-scale spatial prediction of nodules. Here, we use five well-established supervised ML algorithms: generalized linear models (GLMs), generalized additive models (GAMs), support vector machines (SVMs), random forests (RFM), and neural networks model (NNM). These algorithms represent five of the most common modeling architectures with different degrees of model complexity; they are available in all standard data analysis programming languages (e.g., R, Python) and some GIS software (e.g., ArcGIS, QGIS).

While prediction accuracy on an independent dataset is the standard way to assess model performance, it is equally important to understand how model predictions occurred. Studies have shown that accurate model predictions can be achieved even when irrelevant and artificial predictors are used, outperforming models with domain-relevant predictors (Fourcade et al., 2018; Behrens & Viscarra Rossel, 2020). A tradeoff between interpretability and accuracy is thus needed, particularly in natural sciences, where the models are used for causal explanation (Breiman, 2001b; Shmueli, 2010). However, interpreting ML results is difficult, especially when our knowledge or conceptual understanding is still developing and complex methods (e.g., neural networks) are used exclusively (Merow et al., 2014). In a multi-model approach, the contribution of each predictor can be evaluated across different models, providing better insights into their importance (Li et al., 2017). Besides variable importance measures are partial dependence plots (PDPs), which depict the relationship type (linear, monotonic, or complex) built between a response and the predictor variables during model training (Friedman, 2001). PDPs are one of the most widely used model-agnostic methods in regression prob-

lems with large datasets due to their simple interpretation and computation efficiency (Molnar, 2018).

In addition to the multi-model approach followed, an ARA was done to better interpret the contribution of MBES backscatter as a spatial predictor. The backscatter angular response for a specific frequency and incident angle is an inherent property of the seafloor (Jackson & Briggs, 1992; Fonseca & Mayer, 2007). The ARA has a higher angular resolution than the processed backscatter mosaic (where several angles from different swaths have been blended and AVG filtering has been applied), but it has a decreased along-track resolution because it is a stack of 30 consecutive MBES pings for the half swath width on port and starboard side (Fonseca & Mayer, 2007; Fonseca et al., 2009). Despite the lower along-track resolution, ping stacking at different beam angles could help to identify the beam angles that contribute the most to the spatial predictions and reveal subtle differences between seafloor patches with different sediment properties. Beyond the extraction of average angular response at certain incident angles, ARA attempts to characterize sediment properties by comparing the average angular response for each group of incident angles with mathematical models that have linked backscatter and sediment properties (Jackson et al., 1986; Fonseca & Mayer, 2007). Based on this model inversion, the backscatter impedance (hereafter impedance) and backscatter volume (hereafter volume) were calculated. Impedance is the product of sediment bulk density and sound velocity ratio between sediment compressional wave speed and water sound speed, showing the acoustic contrast between water and sediment surface (Fonseca & Mayer, 2007). Sediments with lower water content and typically larger grain size provide higher impedance and stronger backscatter signal (Jackson & Briggs, 1992). Volume scattering provides information on the sound attenuation within the sediment concerning sediment physical properties (e.g., density) and inhomogeneities within the sediment column (e.g., buried nodules or fragments of them, burrows, and gas bubbles). In fine-grain size sediments (e.g., deep-sea clay and silt) with low density and roughness, volume scattering could be the dominant contributor at lower frequencies, intermediate oblique incident angles ( $> 15^\circ$  and  $< 60^\circ$ ) and when a bioturbated or harder/rougher layered substrate exists below the surficial thin, soft layer, e.g., thickness variations in the semi-liquid bottom

sediment layer (Jackson et al., 1986; Lurton & Lamarche, 2015).

To enhance the geomorphological analysis and better understand the spatial distribution of nodule coverage, we constructed high-resolution orthophoto-mosaics of the seafloor at a spatial resolution of 2 mm pixel, which has not been acquired elsewhere in such detail and water depth until now. In addition, image-derived digital elevation models (DEMs) from seafloor parts with and without nodule coverage supplemented the analysis and interpretation of MBES and SSS backscatter.

## STUDY AREA

The study area was in the Eastern CCFZ, within the B4 domain of the global sea mineral resources (GSR) contract area (Fig. 1). This stretch of the deep-sea, with an area of 37.34 km<sup>2</sup> and – 4520 to – 4392 m water depth, is one of the study sites of the in situ mining trials of the Patania II pre-prototype seafloor nodule collector vehicle (Vink et al., 2022). It is part of a larger geomorphological system of abyssal hills and valleys (horsts and grabens, respectively) aligned to the N–S/NW–SE axis, parallel to each other (Supplementary Fig. S1), a typical structure in the eastern CCFZ (Macdonald et al., 1996; Parianos et al., 2022).

The central part is a wide N–S oriented U-shape valley limited to the south-west by a volcanic knoll with a crater on top. To the south-east, it is limited by a horst, and to the south, there is a sill with slopes of up to 5° and two cone-shaped morphological features (Fig. 1, Supplementary Fig. S1). Several sub-circular and irregularly shaped depressions exist inside the central basin, creating an uneven seafloor. The largest sub-circular depression has a diameter of ~ 400 m and a water depth of – 4520 m, the deepest part of the study area (Fig. 1, Supplementary Fig. S2). Based on the box corer samples retrieved from the central part (Haeckel & Linke, 2021), the seafloor is covered by unconsolidated, very soft, yellowish to medium grayish-brown clayey silt with traces of fine sand and biogenic siliceous components, being homogeneous at the first 5 cm (Supplementary Fig. S2). The sediment had an increased water content (55.63–72.28%) and bulk density of 1.35 t m<sup>-3</sup> (GSR, 2018).

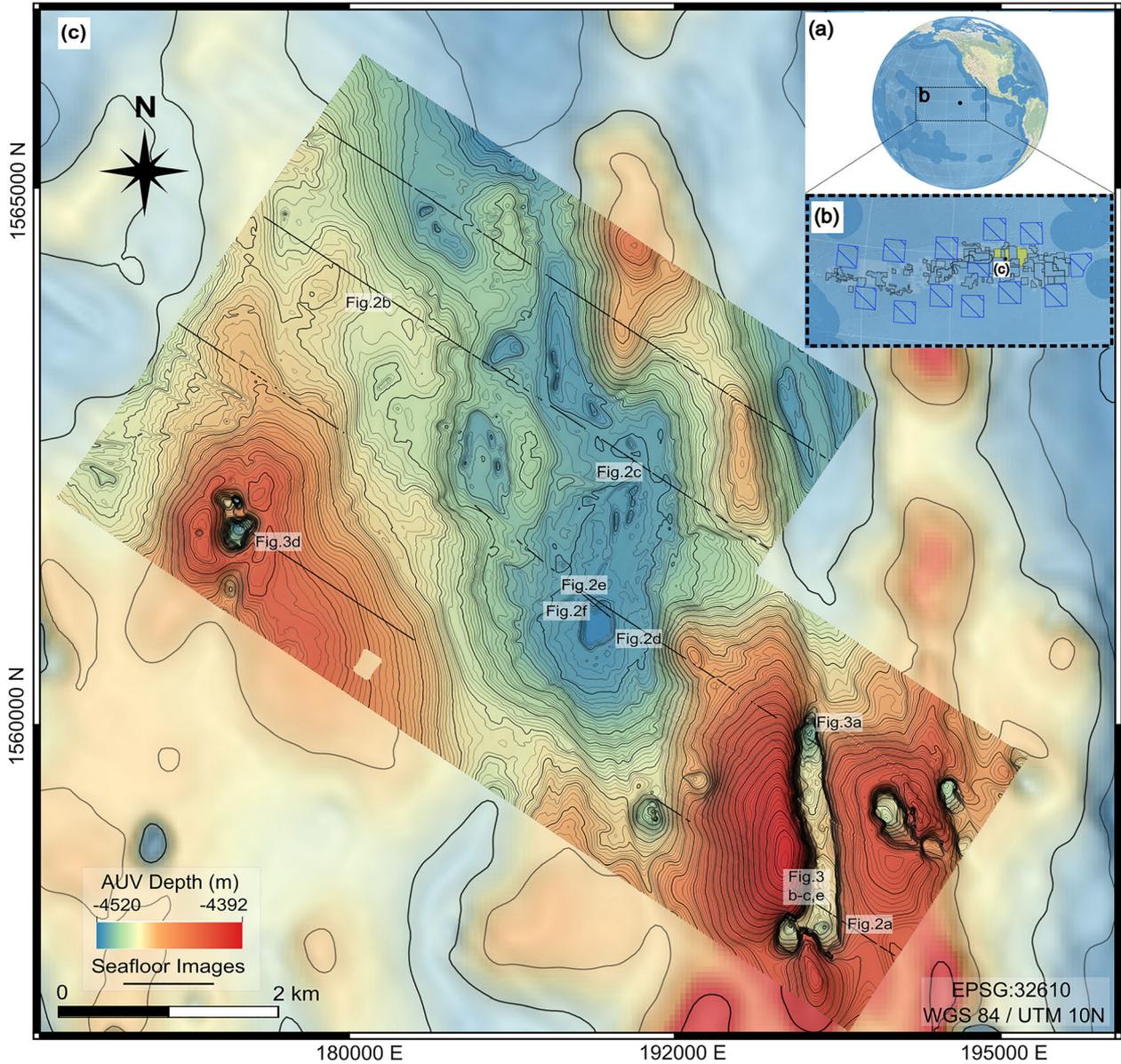
The nodules were mainly of diagenetic origin with Mn/Fe ratio > 5 (Halbach et al., 1981; GSR, 2018). Their shape was ellipsoidal, with the major

and minor axes having a linear relationship (Schoening & Gazis, 2019; Yu & Parianos, 2021). Their texture was smooth or botryoidal on top, separated by a visible equatorial belt from the rough buried bottom part (Yu & Parianos, 2021). The nodule density was 1.99 g cm<sup>-3</sup> (GSR, 2018). As regards the central part of the study area, the median abundance, including surficial and buried nodules, was 22.4 kg m<sup>-2</sup> of wet weight (Schoening & Gazis, 2019). The percentage of buried nodules (first 10 cm in the sediment column) accounted for 14.3% of the total number of nodules within the box corer samples (Schoening & Gazis, 2019). The % seafloor nodule coverage (nodule coverage) based on the few box corer samples ranged from 4 to 43%, with a median of 38% (Schoening & Gazis, 2019). This nodule coverage range agrees with the nodule coverage derived from the seafloor images within the central part (Fig. 2). The nodules were large individuals, not interconnected to other nodules (Fig. 2). Lower nodule coverage and smaller nodules existed around local depressions. Within the local depressions, the nodule coverage was < 3% (Fig. 2). Mounds of sediment accumulation without nodules on top caused by bioturbation (i.e., tumuli) co-existed between the nodules throughout the study area (Supplementary Fig. S3). Traces of mobile surface fauna and bioturbated sediments were found in seafloor images, box corer, and gravity corer samples (e.g., sinuous burrows).

The volcanic knoll's northern and eastern/southeastern parts have gentle slopes (3°) and relatively smooth terrain. The northwestern part had steeper slopes (5–7°) and downslope channels, indicating sediment movement downslope (Fig. 1, Supplementary Fig. S2). Lower nodule coverage and smaller nodules were observed here (Supplementary Fig. S4). The main crater was 50–60 m deep (Supplementary Fig. S2). Based on seafloor images, there were nodule-free areas, consolidated sediments, and rocky outcrops around and inside the crater (Fig. 3).

The horst at the south-east part of the study area had a length of 2.3 km and maximum width of ~ 1 km. Its western side was nodule-covered and had a maximum slope of 10°, which forms a topographic high at 4392 m water depth, the shallowest part of the study area (Fig. 1, Supplementary Fig. S2). The opposite east side of the horst structure was bounded by inward-dipping normal faults with a sediment-free vertical cliff (60–80°). The base of the cliff was an enclosed U-shape collapsed trough (>80 m altitude difference), related to a combina-

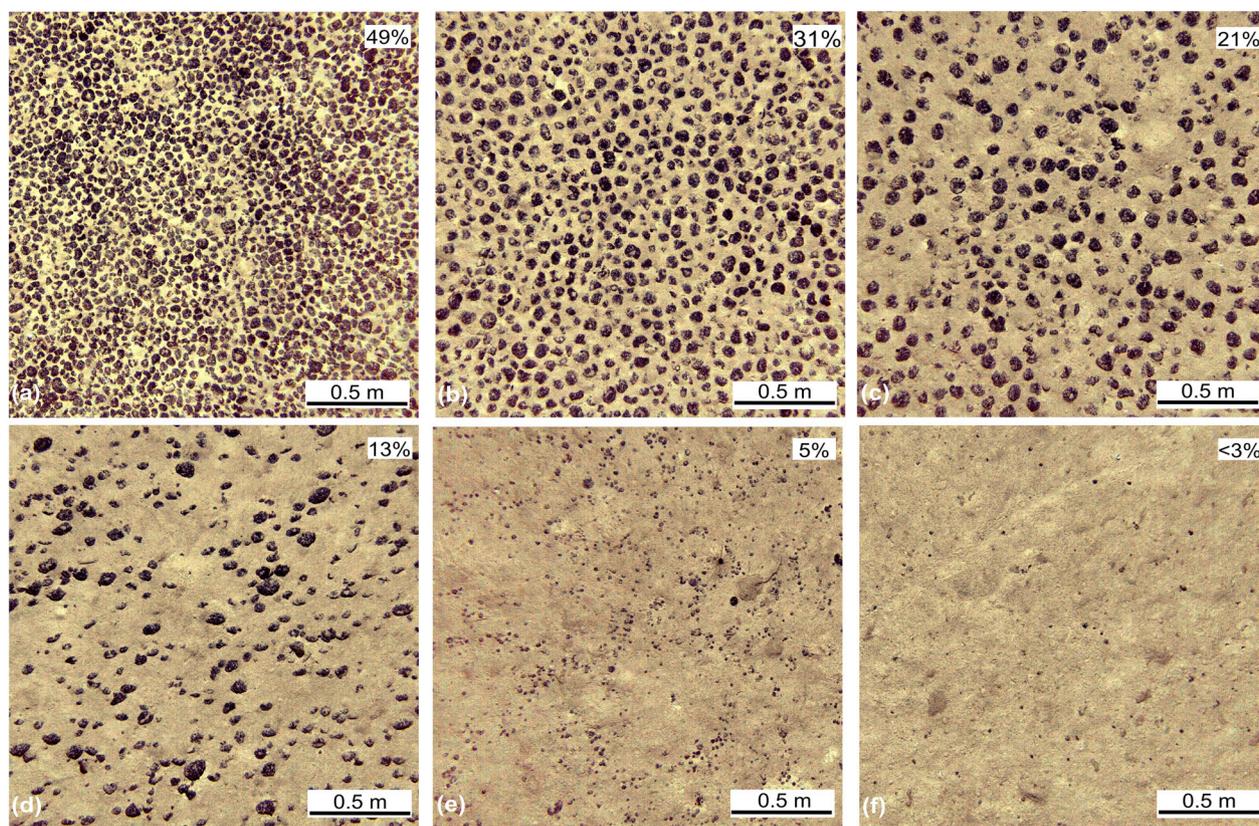
## An Interpretable Multi-Model Machine Learning Approach



**Figure 1.** (a) Map showing the CCF between North America and Hawaii. The darker blue regions are the exclusive economic zones (FMI, 2019). (b) Map showing the nodule exploration contract areas (black polygons), reserved areas (light blue polygons) around the exploration contract areas) and the areas of particular environmental interest (hashed blue squares) in the CCFZ (ISA, 2023). The GSR exploration contract area is divided into three blocks (marked in yellow). The black dot indicates the location of the study area. (c) AUV-based bathymetry of the study area (grid cell is  $2\text{ m} \times 2\text{ m}$ ; contour interval 1 m). The background bathymetry is from ship-based MBES data, with grid cell size of  $50\text{ m} \times 50\text{ m}$  and contour interval of 25 m (Gaziz, 2020). The black lines are the AUV seafloor images used for spatial modeling. Seafloor images exist along the entire AUV transect, but only the images obtained at an altitude  $\leq 7\text{ m}$  were used for spatial modeling and shown here. The locations of inset images in Figures 2 and 3 are indicated.

tion of faulting and carbonate dissolution (Mayer, 1981; Fig. 1). Its width varied from  $\sim 100\text{ m}$  in the northern part to  $\sim 700\text{ m}$  in the southern part (Fig. 1). Different nodule facies existed here: In the northern part, the seafloor had smaller nodules with

high spatial density. This part was characterized by exposed compacted and layered indurated sediments at successive steps. On top of the exposed parts, the loose sediment was winnowed away, and small nodules lie (Fig. 3). Such structures within



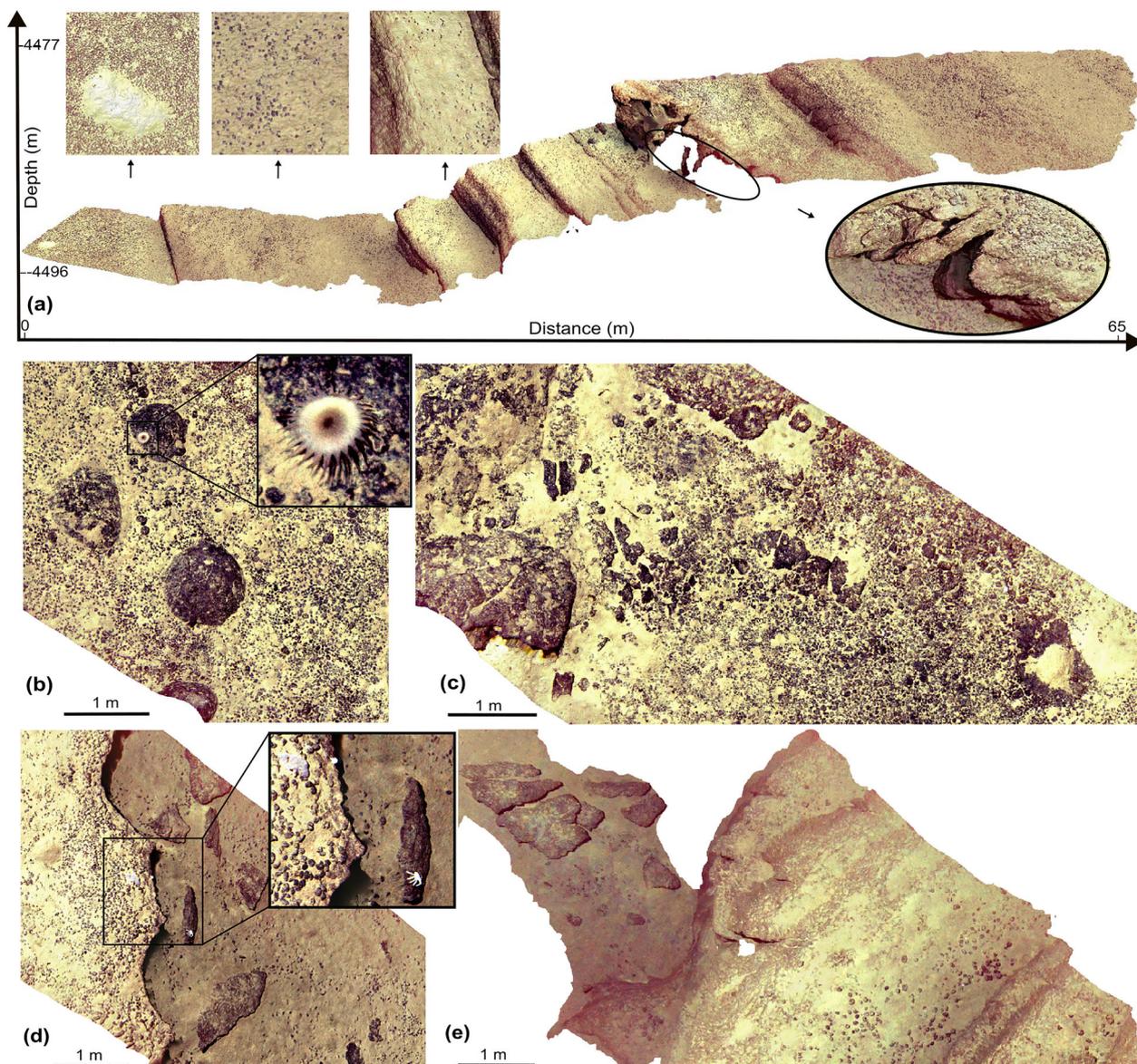
**Figure 2.** Color-normalized AUV-generated seafloor images from different locations within the study area. Each image shows a different nodule coverage. The location of each image is given in Fig. 1.

deep-sea troughs have been mapped in the past in CCFZ (Mayer, 1981). In the southern part of the trough, the seafloor was covered by densely packed small nodules, debris, pillow basalts, basaltic fragments, and rocky outcrops (Fig. 3).

The highest nodule coverage (up to 50%) was observed at the southeastern edge of the study area (Fig. 2). Geomorphologically, prominent sites have been linked to high nodule coverage because of abundant nuclei material that existed as fragments or was gravitationally transferred downslope when steep slopes exist (Yamazaki & Sharma, 2000; Sharma, 2017; Li et al., 2021). The nodules here were smaller than in the central and flatter part of the study area (Fig. 2), a fact that could be attributed to one or several factors: the limited available horizontal space relative to a high number of nuclei, younger age, primary hydrogenetic growth, and limited supply of dissolved metal through diagenetic precipitation due to thinner or absent semi-liquid

bottom sediment layer (i.e., geochemically active layer) due to frequent basement exposure (Cochonat et al., 1992; Lipton et al., 2016). Further to the southeast, three pits (or potholes) of varying depths (10–50 m) were presented (Fig. 1, Supplementary Fig. S1). They had sharp walls (locally steeper than 60°) with layered consolidated sediments (Supplementary Fig. S5). Pits existed along all the main horsts within the broader study area, having varying depths and sizes (Supplementary Fig. S1). Their presence has been linked to the dissolution of outcropping carbonate and Miocene limestone because the study area is at the limit of the calcite compensation depth (CCD; Berger et al., 1976) augmented by turbulent erosion and hydrothermal discharge through conduits created by faulting and differential compaction between the asymmetrical flanks (Mayer, 1981; Bekins et al., 2007; Moore et al., 2007).

## An Interpretable Multi-Model Machine Learning Approach



**Figure 3.** Color-normalized ortho-photomosaics from different seafloor locations within the study area (the location of each section is given in Fig. 1). The grid cell size is  $2\text{ mm} \times 2\text{ mm}$ . (a) Exposed carbonate cemented sediment layers appear as successive steps along the eastern slope of the trough. Densely packed small nodules are observed at the deepest and flattest part, while the carbonate steps exhibit fewer nodules. (b) Four solitary pillow basalt blocks in the southern part of the trough with associated fauna (*Actiniaria*). (c) and (d) Basement exposure, basaltic debris, small nodules, and nodule fragments generate a hard substrate inside the trough. The sharp-edged steps with detached, undissolved rocky fragments indicate a cobalt-rich ferromanganese crust. Sessile (*Porifera*, top left) and motile fauna (*Arthropoda*, bottom right) are shown within the black inset in (d). The fauna annotation was based on the standardized taxonomic field guide presented by Simon-Lledó et al. (2023). (e) The outer crater shows low-to-no nodule coverage and angular rocks as seafloor pavement.

## METHODOLOGY

### MBES Data Acquisition and Post-Processing

The used HUGIN 6000 AUV was equipped with a Kongsberg EM2040 MKII operating at 400 kHz, with beam opening angles of  $0.7^\circ \times 0.7^\circ$ . The swath opening angle was  $120^\circ$ , and the flying altitude was kept constant at 60 m above the seafloor, resulting in uniform swath width and  $\sim 50\%$  overlap between neighboring lines. A detailed description of the acquisition system settings (e.g., beam spacing pattern, pulse mode, and ping rate) is provided in the expedition report (Vink et al., 2022). The raw data were stored in the Kongsberg .all format. The navigation data were post-processed onboard using Kongsberg's proprietary NavLab software. The post-processing was done with the QPS Qimera v1.7 software and included sound velocity profile ray tracing correction, roll bias correction, and tide correction based on pressure changes of two fixed deployed SBE16 CTDs (Vink et al., 2022), AUV turn removal, automated CUBE filtering (Calder & Mayer, 2003), and manual removal of erroneous soundings. The generated bathymetric grid was exported as GeoTIFF with  $2\text{ m} \times 2\text{ m}$  cell size in Universal Transverse Mercator coordinates zone 10N (hereafter UTM 10N). The absence of navigational drift was confirmed using the available ship-based MBES data (Gazis, 2020) as a reference layer.

The processed lines were exported to MBES generic sensor format (.gsf) and imported to the QPS Fledermaus Geocoder Toolbox (hereafter FMGT) v7.8 for the backscatter post-processing. The FMGT built-in workflow was followed (i.e., radiometric correction, slope correction based on the exported bathymetric grid, angle varying gain (AVG) correction, nadir filtering, anti-aliasing, despeckling, and mosaicking) to create the backscatter mosaic (Fonseca & Calder, 2005). The backscatter mosaic was exported as a GeoTIFF grid with  $2\text{ m} \times 2\text{ m}$  cell size in UTM 10N. In addition to a backscatter mosaic, the ARA in the near range (incident angles within  $0\text{--}25^\circ$ ), far range (incident angles within  $25\text{--}55^\circ$ ), and outer range (incident angles within  $55\text{--}85^\circ$ ) were exported as GeoTIFF grids in UTM10N. The ARA beam details (i.e., ping number, beam number, easting, northing, depth, backscatter value, corrected backscatter value, and true angle) were exported in .txt format for specific MBES lines for comparing patches with variations in

backscatter intensity and nodule coverage. Based on the ARA model inversion, the acoustic impedance (hereafter impedance) and volume scattering (hereafter volume) were calculated and exported as GeoTIFF grids in UTM10N.

### SSS Data Acquisition and Post-Processing

An AUV-mounted EdgeTech 2205 SSS was used to acquire seafloor backscatter data at 230 kHz, parallel to MBES mapping (Vink et al., 2022). No interference was observed between the two systems. The raw data were stored in EdgeTech .jsf format (in which the post-processed navigation was integrated) and processed with the SonarWiz v7.09 software. The post-processing included bottom tracking based on the provided threshold detection method, slat range and bottom slope correction, gain correction using the empirical gain normalization method, nadir correction, and turn removal. The high overlap ( $\sim 60\%$ ) between the neighboring lines resulted in the absence of nadir gaps in the final mosaic. The final mosaicking was done using the average blending method. The SSS backscatter mosaic extent was clipped to the MBES grid extent and exported as a GeoTIFF grid in  $2\text{ m} \times 2\text{ m}$  cell size in UTM10N.

### MBES and SSS Derivatives

Following results from previous studies (e.g., Gazis & Greinert, 2021), 19 derivatives were produced (Table 1). Before calculating the derivatives, the mining site area and adjacent periphery ( $\sim 450 \times 430\text{ m}$ ) were clipped out (in QGIS v3.24) as the anthropogenic nodule removal and seafloor disturbance alternated the natural bathymetric and backscatter values. The three source grids (MBES bathymetry, MBES backscatter, SSS backscatter) were filtered using a Gaussian filter (in SAGA GIS v8) with standard deviation (SD) of 1 and a kernel radius of 2 pixels to eliminate any remaining single-cell outliers that could affect negatively the derivative calculation and modeling performance (Lucieer et al., 2016; Lecours et al., 2017). Afterward, the derivatives were produced. The aspect was transformed into eastness and northness (Florinsky, 2017). None of the derivatives had zero variance (Supplementary Table S1) or was perfectly correlated ( $r > 0.85$ ,  $p < 0.05$ ) with another derivative

## An Interpretable Multi-Model Machine Learning Approach

**Table 1.** MBES and SSS derivatives used for the spatial modeling

MBES and SSS derivatives	Abbreviation	Algorithm	Software
<i>MBES bathymetry</i>			
Bathymetry	BT	CUBE and manual filtering	QPS Qimera
Bathymetry Local Moran	BT_LM	Anselin (1995)	R raster package, Hijmans (2022)
Slope	S	Zevenbergen and Thorne (1987)	ArcGIS BTM, Walbridge et al. (2018)
Northness	N	cos (Aspect in radians)	ArcGIS (BTM)
Eastness	E	sin (Aspect in radians)	ArcGIS (BTM)
Profile Curvature	PrC	Zevenbergen and Thorne (1987)	ArcGIS (BTM)
Plan Curvature	PIC	Zevenbergen and Thorne (1987)	ArcGIS (BTM)
Vector Ruggedness Measure	VRM	Sappington et al. (2007)	ArcGIS (BTM)
Bathymetric Position Index	BPI	Wilson et al. (2007)	ArcGIS (BTM)
Fine scale	BPI0_30	Inner radius: 0 m Outer radius: 30 m	
Broad scale	BPI100_300	Inner radius: 100 m Outer radius: 300 m	
<i>MBES backscatter</i>			
Backscatter	BS	Fonseca and Calder (2005)	QPS (FMGT)
Backscatter Local Moran	BS_LM	Anselin (1995)	R (raster package)
ARA Mean Near	ARA_MN	Fonseca and Mayer (2007)	QPS (FMGT)
ARA Mean Far	ARA_MF	Fonseca and Mayer (2007)	QPS (FMGT)
ARA Mean Outer	ARA_MO	Fonseca and Mayer (2007)	QPS (FMGT)
ARA Impedance	ARA_IM	Fonseca and Mayer (2007)	QPS (FMGT)
ARA volume	ARA_VO	Fonseca and Mayer (2007)	QPS (FMGT)
<i>SSS backscatter</i>			
SSS backscatter	SSS	SonarWiz	SonarWiz
SSS backscatter Local Moran	SSS_LM	Anselin (1995)	R (raster package)

(Supplementary Fig. S6). The maximum correlation was among the ARA outputs. Because the ARA contribution to modeling performance was one of the study's interests, all ARA outputs were kept. The caret (Kuhn, 2022) and GGally (Schloerke et al., 2023) R packages were used for the zero variance and correlation analyses.

### Image Data Acquisition and Post-Processing

High-resolution seafloor images ( $4096 \times 3072$  pixels) were acquired by the AUV HD CathX Colour Still Camera (M12 A1000) and stored in JPEG format (Vink et al., 2022). Like the hydroacoustic data, the AUV navigation from the photo-surveys was post-processed using Kongsberg's proprietary NavLab software to define each image location, camera altitude and the photographed seafloor image area. The nodule coverage was calculated using the saltation GmbH & Co. KG proprietary H<sup>2</sup>SOM software (Supplementary Fig. S7), which automatically detects and segments polymetallic nodules in underwater images using a hierarchically growing hyperbolic self-organizing map approach (Schoening et al., 2016). Other researchers also have used

the H<sup>2</sup>SOM software to quantify the nodule coverage in seafloor images (Ellefmo & Kuhn, 2021; Kuhn & Rühlemann, 2021). The fine-tuning of the software parameters for this dataset was done by saltation GmbH & Co. KG. During the image acquisition, the AUV flying altitude was set at 5 m above the seafloor, which could not always be kept constant due to altimeter bottom detection warnings, which forced the AUV to fly higher in some parts of the survey lines (up to 15 m in some cases). Image evaluations showed that the photos obtained above 7 m could not offer meaningful information and thus were excluded. Images along the AUV turns and those outside the MBES grid extent were also excluded. In total, 31,409 images were used to extract the nodule coverage for model training and testing.

Although not a direct topic of this study, the conversion of nodule coverage from seafloor images (%) to nodule abundance ( $\text{kg m}^{-2}$ ) was not possible due to the absence of needed information to establish local equations between the seafloor images and the box corer data. Apart from the available image-derived nodule coverage, the image-derived size distribution and box corer-derived nodule abundance, coverage, and size distribution were needed at the seafloor image locations. In addition, the

seafloor images tended to underestimate the actual nodule coverage due to semi-buried nodules or nodules covered by faint sediment (Sharma, 1993; Ellefmo & Kuhn, 2021; Tsune, 2021). The degree of underestimation varied among different nodule sizes, facies and substrates, being higher in sub-areas with thicker sediment cover and lower in sub-areas with rocky substrate and basement exposure (Sharma, 1993, 2017; Parianos et al., 2021). Finally, seafloor images cannot provide information on the buried nodules (Sharma et al., 2013; Lipton et al., 2016; Sharma, 2017), and correction factors based on box corer data must be applied. Nevertheless, studies have related higher nodule coverage with higher nodule abundance in areas without surficial sediment on nodules—like our study site (Lipton et al., 2016; GSR, 2018; Wasilewska-Błaszczuk & Mucha, 2021). Thus, image-derived nodule coverage and model predictions (see **Results** section) could guide future box corer sampling at seafloor image locations and sub-areas with different nodule coverage.

### Orthophoto-Mosaics and Image-Derived Digital Elevation Models

The seafloor images were taken at 5 Hz, resulting in a high overlap (~75%) in the along-track direction, which allowed the computation of high-resolution orthophoto-mosaics and image-derived DEM. Before orthophoto-mosaicking, the original seafloor color was restored in underwater images using the in-house software Image Normalization (Köser et al., 2021), which removes lighting artifacts and illumination cones (Supplementary Fig. S7). Orthophoto-mosaics and the respective DEMs were calculated using the Agisoft Metashape Pro v8.3 software, following the built-in workflow (Camera calibration, Image alignment, Dense Point Cloud creation, Dense Point Cloud Confidence filtering, DEM calculation, and orthophoto mosaic reconstruction). The orthophoto-mosaics and DEMs were exported as GeoTIFF grids with 2 mm × 2 mm and 5 mm × 5 mm pixel/cell sizes, respectively.

### Spatial Modeling

#### *Data Pre-Processing*

The predictor values of the MBES and SSS data sets were extracted from each image location using a

bilinear interpolation of the four adjacent cells to reduce the effect of AUV positioning uncertainty. The data were split into training (80%) and testing (20%) datasets using stratified sampling without replacement (i.e., no duplicate samples between training and testing datasets), ensuring good representativeness of the univariate density distribution and spatial coverage between the two datasets. The splitTools (Mayer, 2022) R package was used for this purpose. The testing dataset was kept from further steps such as correlation analysis, data transformation, cross-validation (CV), and model training to avoid data leakage.

#### *Data Transformation*

As the nodule coverage (response variable) was given in % relative to the observed seafloor, the arcsine transformation was used to avoid extrapolation below 0% or above 100%. In addition, it enhanced the model comparison and interpretability at the extreme ends because it stretched out the data in these two areas. The transformation was done in R using the formula:  $y = \arcsin(\sqrt{x})$ , where  $x$  is a real number from 0 to 1, representing the nodule coverage in percentage. The model predictions were back-transformed using the formula:  $x = \sin(y)^2$ , and multiplied by 100 to return the predicted percentages of nodule coverage. The performance metrics on the testing dataset were calculated using the back-transformed data.

#### *ML Algorithms*

The following regression algorithms were applied using the caret package in R (R Core Team, 2022):

GLM is an extension of the linear regression model that allows for incorporating non-normally distributed data (e.g., Poisson, binomial) using link functions that relate the predicted response variable to predictors (McCullagh & Nelder, 1989). Here, we used a GLM augmented with component-wise gradient boosting (Bühlmann & Hothorn, 2007) as implemented in the mboost (Hothorn et al., 2021) R package.

GAM is a data-driven approach that divides data into sections (knots), fitting independent smooth functions (e.g., splines) for each section,

## An Interpretable Multi-Model Machine Learning Approach

assuming that the built functions are additive. It can handle non-linear and non-monotonic relationships between the response and predictors without assumptions regarding the distribution of the response variable and predictors (Hastie & Tibshirani, 1986). Like the GLM approach, a component-wise gradient-boosted GAM that uses penalized B-splines (a.k.a. P-splines; Eilers & Marx, 1996) as a base learner was applied using the `mboost` R package.

RFM is an ensemble of regression trees that grow independently using a random subset of the training data and a random subset of predictor variables (Breiman, 2001a). Thus, correlations among trees are minimized. In each tree, the 'parent' nodes are split using the best subset of predictors that minimizes the variance of the 'child' nodes until the maximum defined depth is reached. The final prediction is the average prediction of all trees (in regression problems). The `ranger` R (Wright & Ziegler, 2017) package was used.

SVM is a non-parametric technique that uses kernel functions to divide the hyperparameter data space into subspaces (hyperplane), aiming for good separation of the data (Cortes & Vapnik, 1995). The kernel functions can be linear, polynomial, sigmoidal or radial basis and can project data from a low-dimensional space to a higher dimensional space, the so-called 'kernel trick' (Karatzoglou et al., 2004). The prior choice of kernel function is arbitrary, and several kernel functions should be tested as the performance may vary depending on the data size and relationships among data. The `kernelab` (Karatzoglou et al., 2023) R package was used.

NNM is a data-driven approach to extract patterns (i.e., relationships) from data using algorithms (e.g., forward propagation, backpropagation) and activation functions (e.g., logistic, tanh). It creates weighted associations between predictors and response variables at the hidden layers, which receive information from all nodes of the previous layers and transfer it to the activation function and next layer(s). Non-linear activation functions allow the network to deal with non-linearities between predictors and response variable (Lippmann, 1987). We used the `neuralnet` (Günther & Fritsch, 2010) R package to develop an NNM that uses resilient backpropagation with weight backtracking (a.k.a., `rprop +`; Riedmiller & Braun, 2003; Riedmiller, 1994).

Hyperparameter optimization (aka tuning) was done for all models. GLM tuning included the

number of boosting iterations (Supplementary Table S2). GAM tuning included the number of boosting iterations and the degrees of freedom in P-splines (Supplementary Table S3). RFM tuning included the number of predictor variables randomly sampled at each node split, the split rule and the minimum node size (Supplementary Tables S4–S5). SVM tuning optimized the kernel type, polynomial degree, scale and cost hyperparameters (Supplementary Tables S6–S7). The number of hidden layers, neurons and activation functions were optimized in NNM (Supplementary Tables S8–S9). An exhaustive hyperparameter tuning was out of the scope of this paper. The optimization process occurred during the spatial CV, which is the preferred method when data are spatially autocorrelated (Schratz et al., 2019; Gazis & Greinert, 2021).

The variable importance was extracted for all models. Each model estimated variable importance differently: In GLM, the absolute value of the coefficients from the final model was used (Kuhn, 2022). In GAM, the in-bag risk reductions per boosting iteration (error difference between the current and the previous step for each base learner contained in the model) were calculated (Hothorn et al., 2021). In RFM, the average permutation variable importance from all trees was used (Breiman, 2001a). In SVM, there is no model-specific way to calculate the importance. Instead, we used a permutation approach, where the prediction error was measured before and after shuffling the values of the predictor variable. The larger the prediction error after shuffling, the higher the importance of this predictor variable; the `iml` (Molnar, 2018) R package was used. In NNM, we used the model weights assigned to each predictor to measure variable importance; the weights were calculated based on the Garson algorithm (Garson, 1991; see also Goh, 1995), as it is implemented in `Neural Net Tools` (Beck, 2018) R package. All variable importance outputs were scaled between 0 and 100 for better comparison across the different models (Kuhn, 2022). PDPs were calculated using the `iml` and `pdp` (Greenwell, 2017) R packages. All needed graphs were produced using the `ggplot2` (Wickham, 2016) R package.

### *CV Strategy and Extrapolation Assessment*

ML models can yield over-optimistic results when a non-spatial, random k-fold CV is used for

spatially autocorrelated data (Roberts et al., 2017; Meyer et al., 2018; Ploton et al., 2020; Karasiak et al., 2022). Here, a spatial k-folds CV strategy was applied. The folds' dimensions (and consequently, number) are a tradeoff among spatial autocorrelation range, data representativeness, and model extrapolation (Roberts et al., 2017; Valavi et al., 2019). Small folds could not eliminate the influence of spatial autocorrelation in the training locations (over-optimistic predictions), while too large folds can result in the removal of a big portion of training data, which creates feature space gaps and causes unnecessary extrapolation that reduces modeling performance (over-pessimistic results). We used an incremental Moran's Index analysis to identify the maximum distance of spatial autocorrelation. This analysis was done in ArcMap v10.6 according to the provided equations, the null hypothesis (i.e., the examined attribute is randomly distributed) and the significance level ( $p < 0.05$ ). The conceptualization of spatial relationship was based on the inverse Euclidian distance, and the spatial weights were standardized, eliminating any bias that could be induced due to the different number of spatial neighbors (Mitchell & Griffin, 2021). Three different numbers of non-overlapping spatial folds were created and compared:  $1 \times 1$  km (28 blocks),  $3 \times 3$  km (5 blocks) and  $5 \times 5$  km (2 blocks; Supplementary Fig. S8).

For each of the three spatial block configurations, the density distributions of geographical and feature space distances were compared among the training dataset, the CV folds and the total geographical and future space area used for model predictions aiming to identify the best size of the spatial blocks (tradeoff between the minimization of spatial autocorrelation and need for extrapolation). This comparison was made using the CAST (Meyer et al., 2024) R package, which implements the abovementioned method proposed by Meyer and Pebesma (2022). The degree of how well a model extrapolates in space was assessed using the area of applicability (AOA) method (Meyer & Pebesma, 2021). This method outputs the geographical areas where the multivariate feature space conditions are similar to those captured by the model, and the CV predictive performance still holds. The geographical parts with novel feature space conditions were returned as non-applicable areas (Meyer & Pebesma, 2021). The feature space dissimilarity analysis is based on the normalized and weighted (according to model weights, a.k.a. variable importance) Eu-

clidean distances between training data and the total feature space of the entire study area. The AOA was calculated using the CAST R package.

The schematic flowchart that illustrates the main methodological and modeling steps is presented in Supplementary Figure S9.

## RESULTS

### CV Block Size

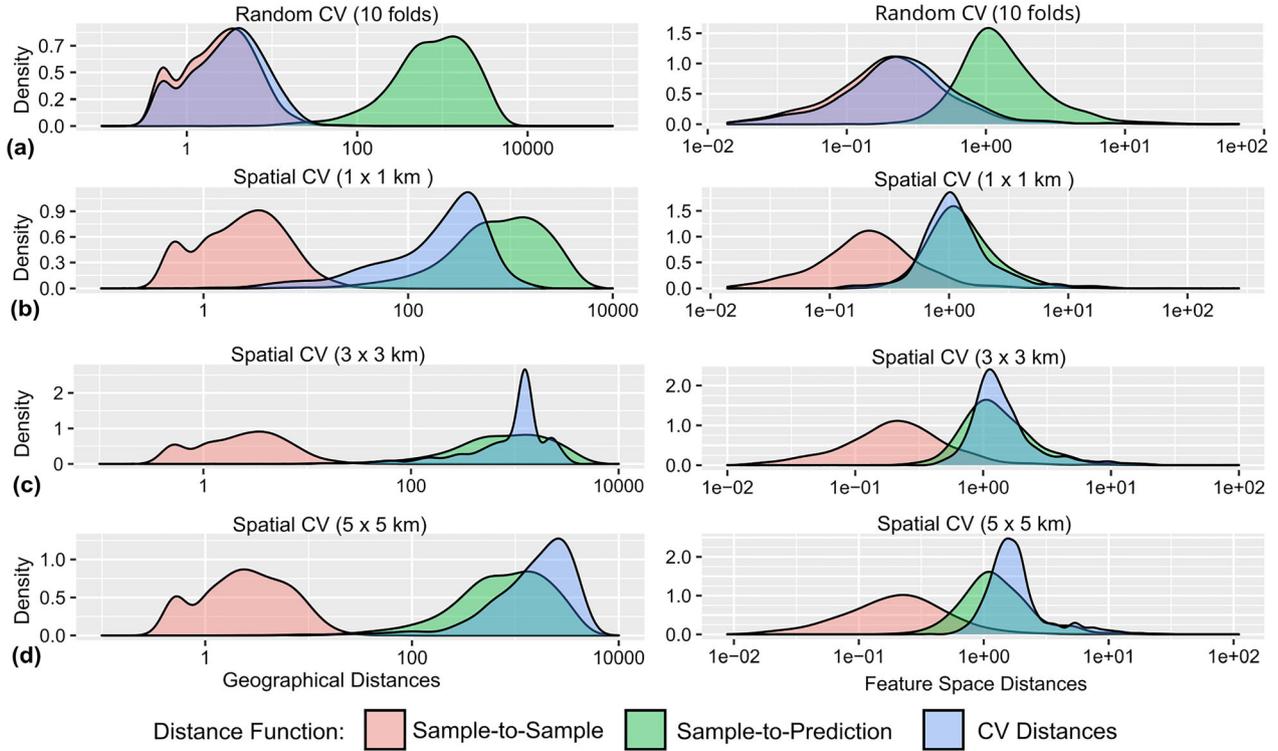
The incremental Moran Index analysis showed that nodule coverage was spatially autocorrelated, particularly in distances  $< 1$  km (Moran I statistic  $> 0.5$ ,  $p < 0.05$ ). At 3 km distance, the Moran I statistic was  $< 0.1$  ( $p < 0.05$ ) with decreasing trend. It reached zero at distances over 4.5 km (Supplementary Fig. S8). The analysis showed that, at 1 km, the geographical distances were still smaller than those needed for representative predictions, and the spatial autocorrelation was still considerable, potentially leading to over-optimistic results (Fig. 4). At 5 km, the geographical and feature space distances inside CV blocks were larger than those encountered during prediction, resulting in unnecessary extrapolation (Fig. 4). The  $3 \times 3$  km CV approach showed the best representativeness between the training data and the model space, and it was selected for model training, tuning and evaluation (Fig. 4). All block configurations outperformed the random-CV strategy (Fig. 4).

### Predictive Performance, Nodule Coverage and AOA

RFM had the best predictive performance on the testing data, followed by NNM (Table 2). SVM and GAM had the same coefficient of determination (hereafter  $R^2$ ) between model predictions and observations on the testing data. However, SVM had a lower mean absolute error (hereafter MAE) and root mean squared error (hereafter RMSE) than GAM. GLM had the worst fit on the testing data ( $R^2 = 0.72$ ; Table 2).

All models reproduced the testing data distribution (e.g., mean, median, interquartile range) well but not the minimum (min) and maximum (max) values of the response variable (Table 3). Despite the best overall performance, RFM yielded the lowest value range, overestimating the lower values

## An Interpretable Multi-Model Machine Learning Approach



**Figure 4.** Geographical and normalized multivariate predictor space Euclidean distances among train dataset, CV folds, and predicted area for (a) random-CV, (b) spatial-CV (1 × 1 km), (c) spatial-CV (3 × 3 km) and (d) spatial-CV (5 × 5 km).

**Table 2.** Model performance metrics on test dataset for the back-transformed predictions

Model	R <sup>2</sup>	MAE	RMSE
GLM	0.72	2.80	3.92
GAM	0.75	2.60	3.68
RFM	0.84	2.04	2.94
SVM	0.75	2.44	3.63
NNM	0.76	2.54	3.56

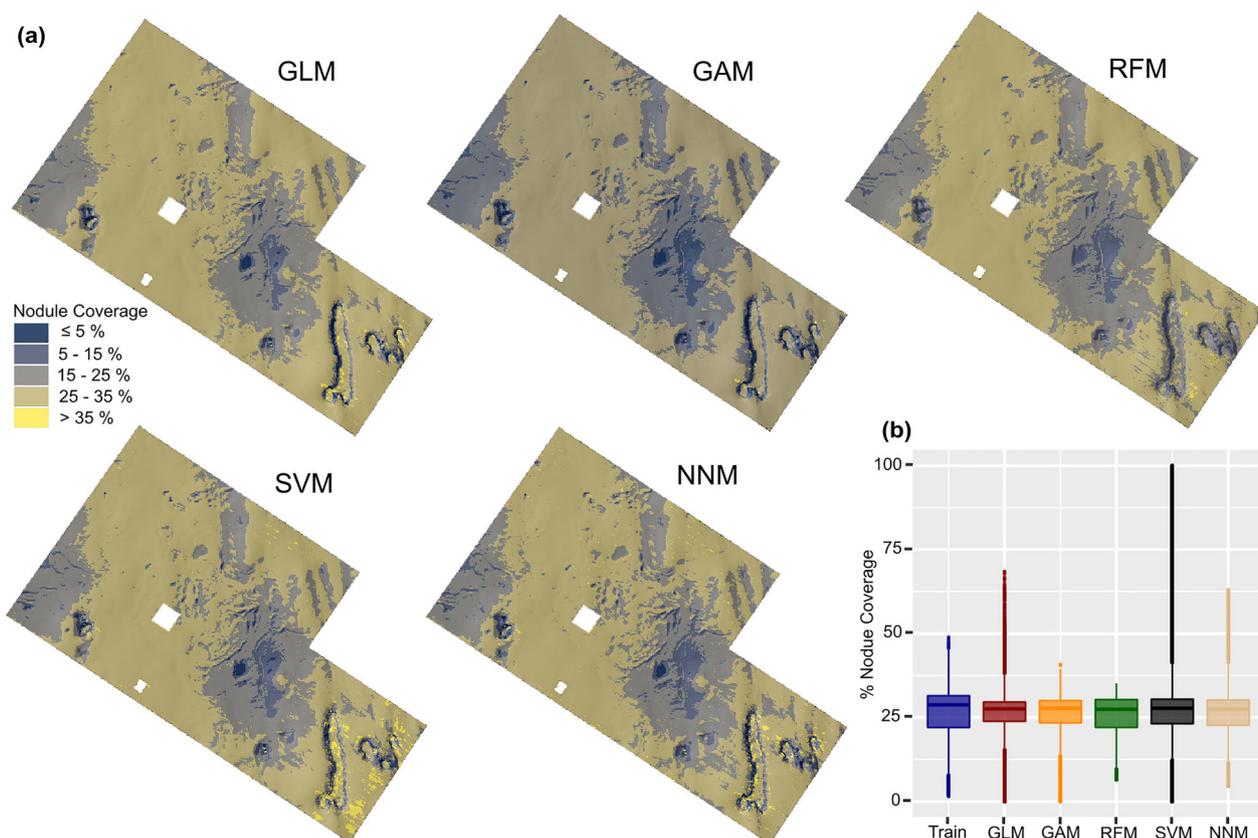
**Table 3.** Summary statistics of test dataset and model predictions on the test dataset

	Min.	1st Qu.	Median	Mean	3 <sup>rd</sup> Qu.	Max.
Test data	1	22	29	26	32	46
GLM	0	24	28	26	30	40
GAM	2	23	29	26	31	34
RFM	7	22	29	26	31	35
SVM	2	23	29	26	31	40
NNM	5	23	29	26	31	41

and underestimating the higher values (Table 3). NNM, SVM and GAM predictions were closer to the initial min–max range. GLM was the only model that predicted beyond the min value of the testing data, but it also underestimated the upper range (Table 3).

All models yielded similar nodule coverage maps and interquartile ranges (Fig. 5, Supplementary Table S10). Differences in the predicted min–max range were observed as GAM extrapolated only in the lower range of nodule coverage, NNM only in the upper range, and GLM and SVM extrapolated in both directions (Fig. 5). SVM predicted 100% nodule coverage in isolated pixels, which were not observed in the seafloor images or predicted by any other model (Fig. 5). The model predictions beyond the min–max range of training data had limited spatial extent ( $\leq 0.23\%$  of the total area; Supplementary Fig. S10).

The weighted-average prediction surface from all models showed excellent spatial agreement with the seafloor images from the testing data and box



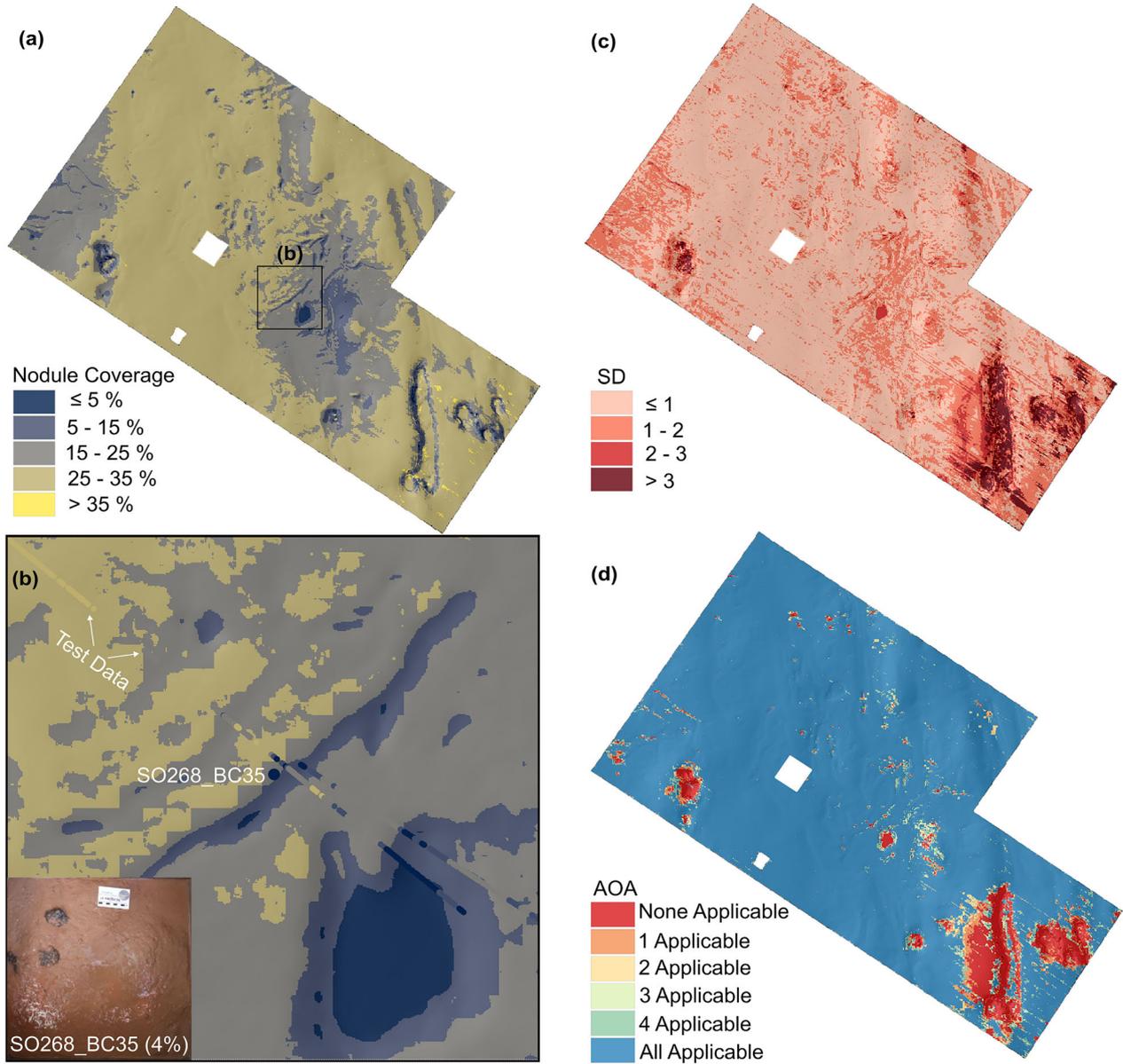
**Figure 5.** (a) Maps of the spatial predictions of nodule coverage and range of predicted values (over the entire study area for each ML model). Along a north–south axis at the west part of the study area, the seafloor is quasi-homogeneously covered with nodules (20–30% coverage). Higher coverage exists in the southeast but is rather localized ( $> 35\%$ ). The central and deeper part has lower nodule coverage ( $\sim 15\text{--}25\%$ ) that progressively decreased, reaching almost zero values at the deepest part of the basin. Lower coverage also existed at the northwest corner (crater slope). (b) Boxplots of nodule coverage predicted values for the entire study area.

corer samples (Fig. 6). Comparing the models' predictions pixel-by-pixel, we obtained the spatial differences (here expressed as SD of the predicted values). The largest differences were in geographically small sub-areas associated with sharp morphological transitions, bedrock exposure, consolidated sediments, and volcanic knolls. The highest SD was within the trough at the southeast part of the study area (Fig. 6). The combination of deeper parts with steep slopes (linked with lower coverage) and high backscatter reflectance (linked with higher nodule coverage), which is partly attributed here to nodule coverage and partly to hard substrate, led to contradictory predictions: RFM predicted low nodule coverage, while SVM yielded high nodule coverage (Fig. 6).

Regarding model transferability, the spatial-CV approach showed that all models performed well when a large part of the geographical and feature

space was kept out, a fact partially attributed to the representativeness of the univariate density distribution of each predictor by the training sample (Supplementary Fig. S11). The CV residuals from all models had weak spatial autocorrelation and were normally distributed (Supplementary Fig. S12). RFM had the best average performance in CV folds, followed by GAM, which had the smallest interquartile of all models (Fig. 7). However, the model performance varied among the spatial blocks, with median being higher than the mean in all models (Fig. 7). This variation was due to the poor performance of all models in the same spatial block (Fig. 7). This spatial block was in the southeast of the study area (Fig. 7), having a complex geomorphology which differed from the rest of the study area. The presence of horst, cliff, trough, exposed indurated sediments and different nodule facies (Figs. 1, 2, 3) resulted in extreme derivative values

## An Interpretable Multi-Model Machine Learning Approach

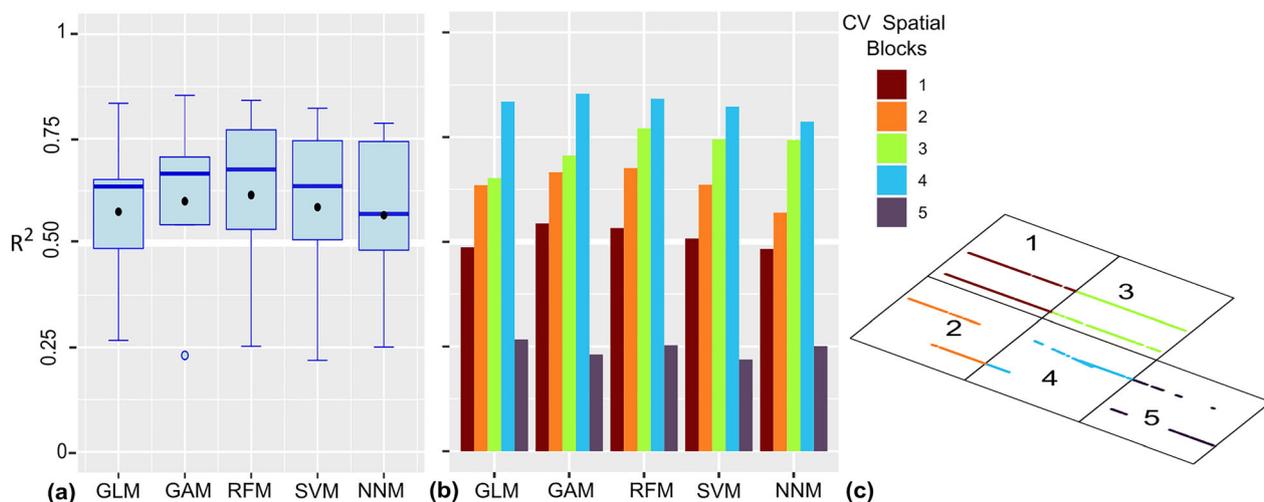


**Figure 6.** (a) Nodule coverage map of the weighted-mean prediction from all models. (b) Map inset from the central part of the study area, which shows the spatial fit between the predicted surface, test data (derived from seafloor images) and box corer data. (c) Variance of the models expressed as SD of the predicted values from each model at each cell. (d) Map showing the areas where models extrapolate based on the AOA method.

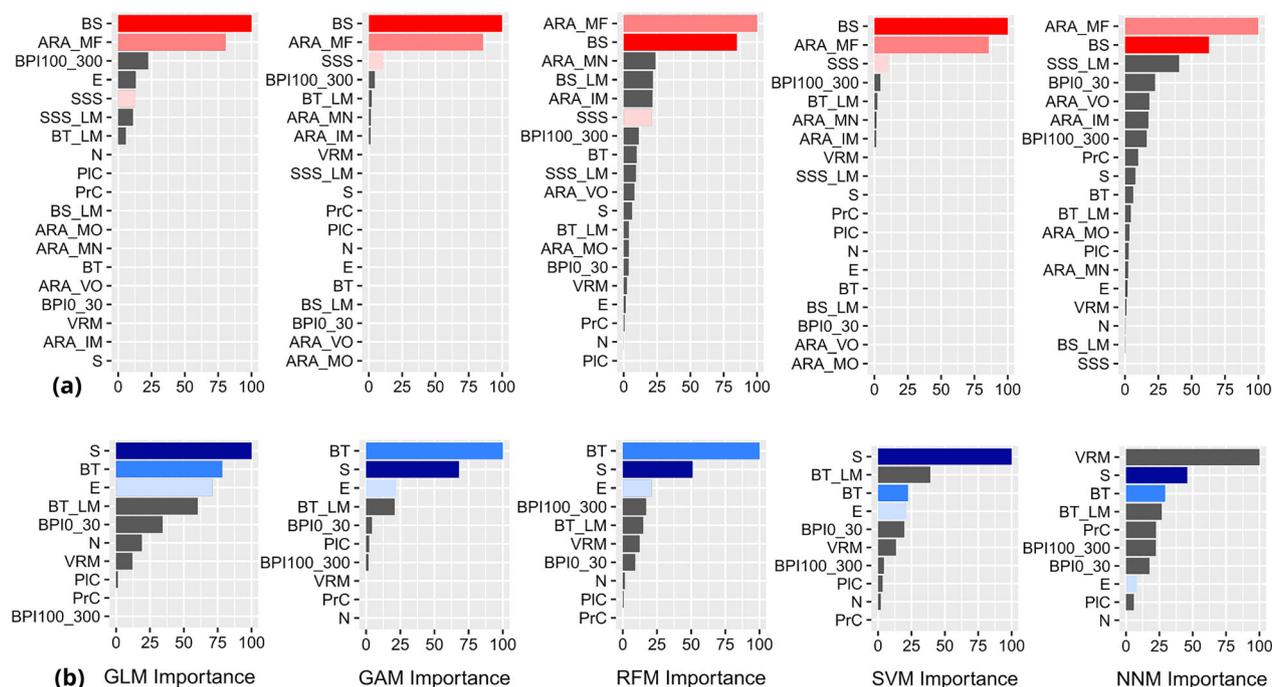
that forced the models to extrapolate beyond the training feature space, as shown by the feature space dissimilarity analysis and AOA (Fig. 6).

The sub-areas where severe model extrapolation occurred were limited because each model was applicable in more than 91% of the total area (Supplementary Fig. S13). All models had similar

AOA, with marginal differences to occur locally (Supplementary Fig. S13). All models applied to 88% of the study area (pixel-by-pixel comparison regarding the pixels where all models applied simultaneously); there was 6% for which only some models were applicable (at least one) and another 6% for which no model was applicable (Fig. 6).



**Figure 7.** (a) CV performance for each ML model. The median (thick blue line) is higher than the mean (black dot) in GLM, GAM, SVM and RFM. The low performance in one block pulls down the mean  $R^2$ . In NNM, the mean and median are the same. Detailed statistics are provided in Tables S11–S13. (b) In-fold performance for each spatial block and model. All models performed similarly in each CV spatial block. (c) Spatial orientation of the CV spatial blocks. Block 5 has the smallest AOA (see Fig. 6).



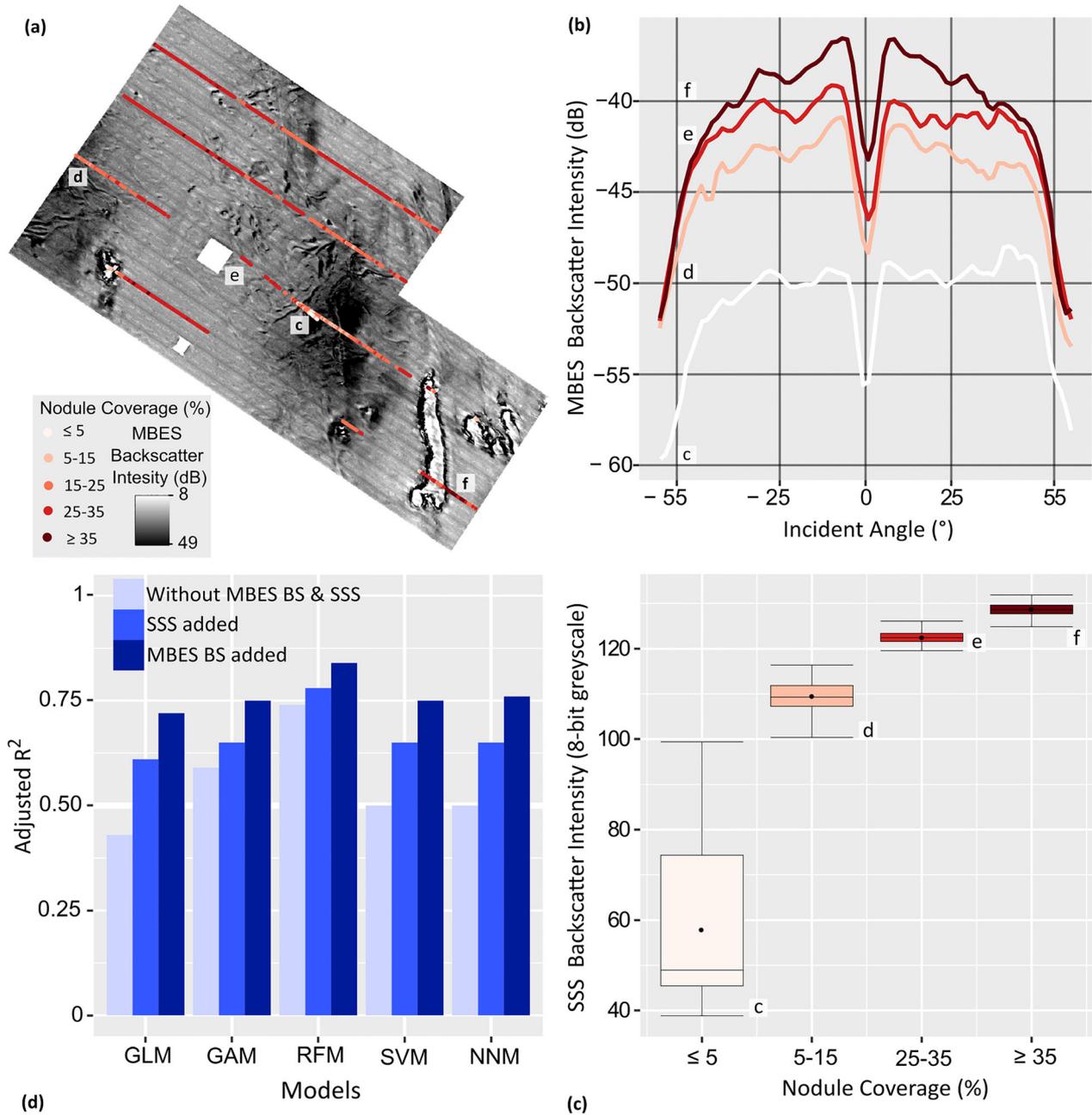
**Figure 8.** (a) Variable importance for each model trained with all available predictor variables. BS and ARA\_MF scored first in all models. (b) Variable importance for each model trained without the hydroacoustic backscatter-related predictors.

### Variable Importance and ML Interpretation

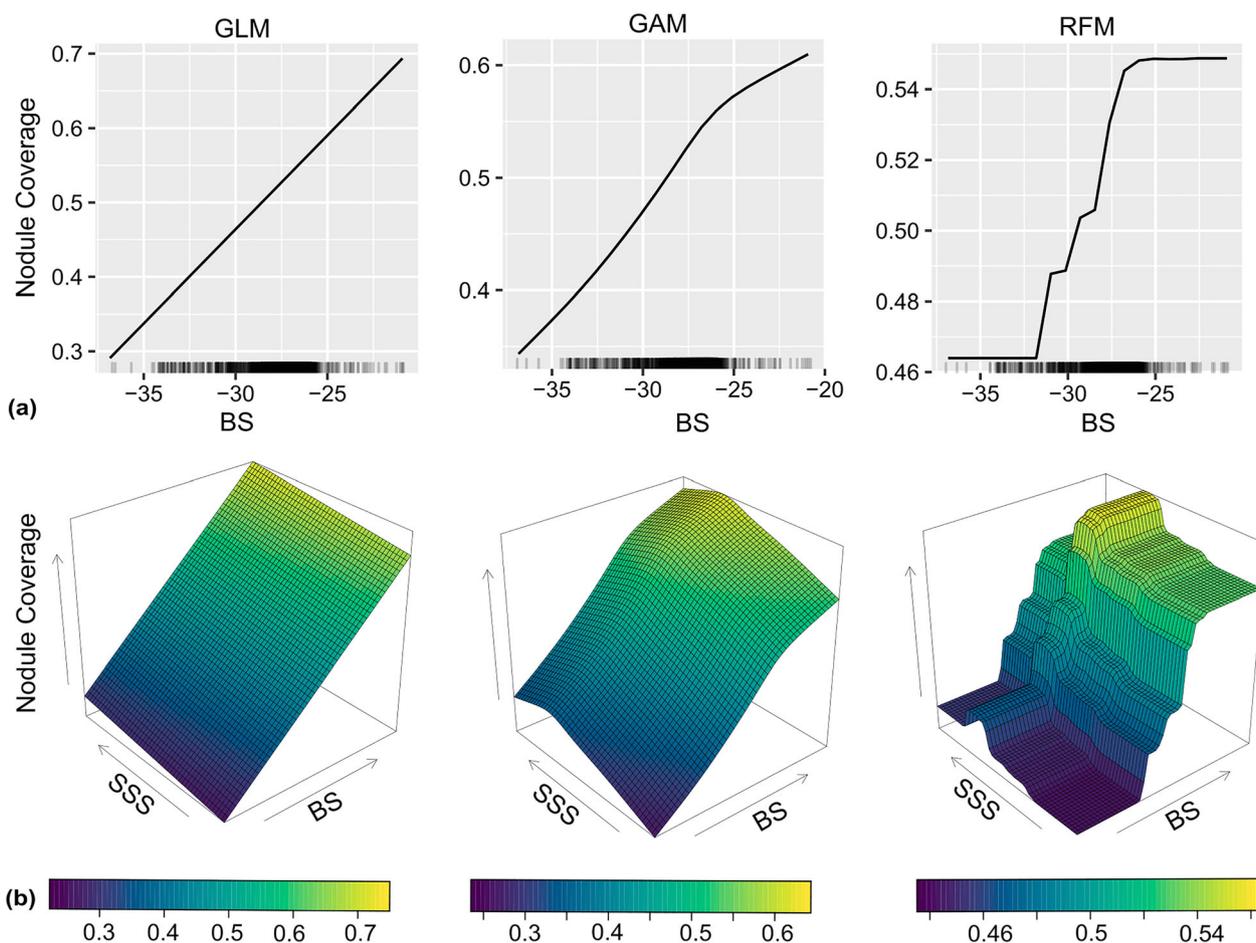
The variable importance calculation showed that BS, ARA\_FM and SSS were the three most

important predictors for the spatial distribution of nodule coverage (Fig. 8). BS and ARA\_FM ranked in the two first positions in all models. The swapping between first and second played among the models

## An Interpretable Multi-Model Machine Learning Approach



**Figure 9.** (a) Post-processed MBES backscatter mosaic of the study area and image-derived nodule coverage. (b) ARA curves from seafloor patches with different nodule coverage. MBES soundings from locations with different nodule coverage show a clear difference in backscatter intensity for the entire range of incident angles. (c) Boxplots of SSS backscatter intensity from seafloor patches with different nodule coverage. The black horizontal line corresponds to the median, and the black dot corresponds to the mean value. (d) Bar plots of the prediction performance of each model using only the geomorphological predictors derived from MBES data (light blue), geomorphological predictors derived from MBES data and SSS backscatter intensity (medium blue), all available predictors including MBES backscatter data and ARA (dark blue). GLM had the biggest predictive improvement when the SSS and MBES backscatter data were added (Supplementary Table S6) due to the monotonic/quasi-linear relationship between SSS and MBES backscatter with nodule coverage (Fig. 10).



**Figure 10.** (a) PDPs showing the relationship between nodule coverage (transformed values) and MBES BS for GLM, GAM and RFM. As the model complexity increases (left to right), the model captures more subtle relationships, resulting in less smooth regression lines. The black marks in the x-axis show the nodule coverage distribution to MBES BS values. (b) PDPs showing the regression plane between nodule coverage, MBES BS and SSS.

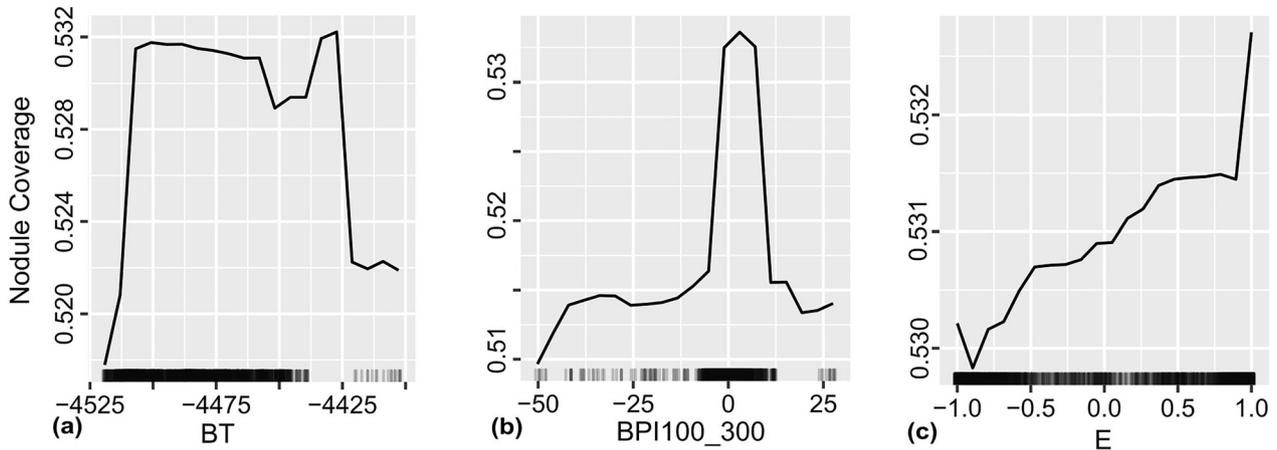
was probably due to the correlation ( $r = 0.8, p < 0.05$ ; Supplementary Fig. S6). SSS was the third most important predictor for two models (GAM, SVM) and scored high in GLM and RFM (Fig. 8). In NNM, the replacement of SSS from SSS\_LM, which ranked as the third most important predictor, was observed. SSS and SSS\_LM were also correlated predictors ( $r = 0.7, p < 0.05$ ; Supplementary Fig. S6).

The MBES backscatter mosaic and ARA analyses depicted the differences in nodule coverage (Fig. 9). A distinct difference ( $> 5$  dB) in backscatter intensity from seafloor parts without and with nodules existed (Fig. 9). Moreover, nodule coverage variations in the order of 10% were detectable, highlighting the sensitivity of the used 400 kHz frequency (Fig. 9). All models captured the monotonic relationship between MBES backscatter and nodule

coverage similarly but with different degrees of complexity (Fig. 10). Although the backscatter differences were distinct for the entire range of incident angles, they were more prominent in the mean-far angle range from 25 to 55°, i.e., ARA\_MF (Fig. 9). The inner and outer incident angles (ARA\_MN and ARA\_MO, respectively) ranked lower than the ARA\_MF. ARA derivatives (ARA\_IM and ARA\_VO) also contributed to the model performance. The SSS backscatter mosaic also discriminated well between seafloor parts with different nodule coverage (variations in the order of 10%) and increased the model performance (Figs. 9, 10, Supplementary Table S15).

All models, particularly GLM, GAM and SVM, heavily relied on the hydroacoustic backscatter-related predictors (BS, ARA, SSS and their derivatives)

## An Interpretable Multi-Model Machine Learning Approach



**Figure 11.** RFM PDPs of nodule coverage concerning (a) bathymetry (BT), (b) broad-scale bathymetric position index (BPI100\_300) and (c) Easting (E).

for error minimization and predictions. The broad-scale seafloor geomorphology, such as bathymetry (BT), slope (S), broad-scale BPI (BPI100\_300) and the eastward slope orientation (E) also related to the spatial distribution of nodule coverage, which became more evident when the hydroacoustic backscatter-related predictors are not used (Fig. 8).

The PDP of BT showed a bathymetric range between  $-4425$  and  $-4500$  m, in which the nodule coverage was higher in this study area. Depths beyond that range were related to seafloor depressions or topographic highs where no nodule existed (Fig. 11). Similarly, the PDP of BPI100\_300 showed that the nodule coverage was higher where BPI values were around 0, i.e., flat or gently sloping seafloor (Fig. 11). E also had a monotonic relationship with nodule coverage. Higher E values were associated with higher nodule coverage (Fig. 11).

At a rather detailed 2-m pixel size, PrC, PIC, and VRM seemed to make only a small contribution to the spatial distribution of nodule coverage (Fig. 8). However, the image-derived DEM from seafloor parts with and without nodule coverage showed distinct differences in seafloor roughness along seafloor parts, although the sediment characteristics stayed similar (Fig. 12). These roughness differences were well captured and integrated into the MBES and SSS backscatter at 400 and 260 kHz, respectively.

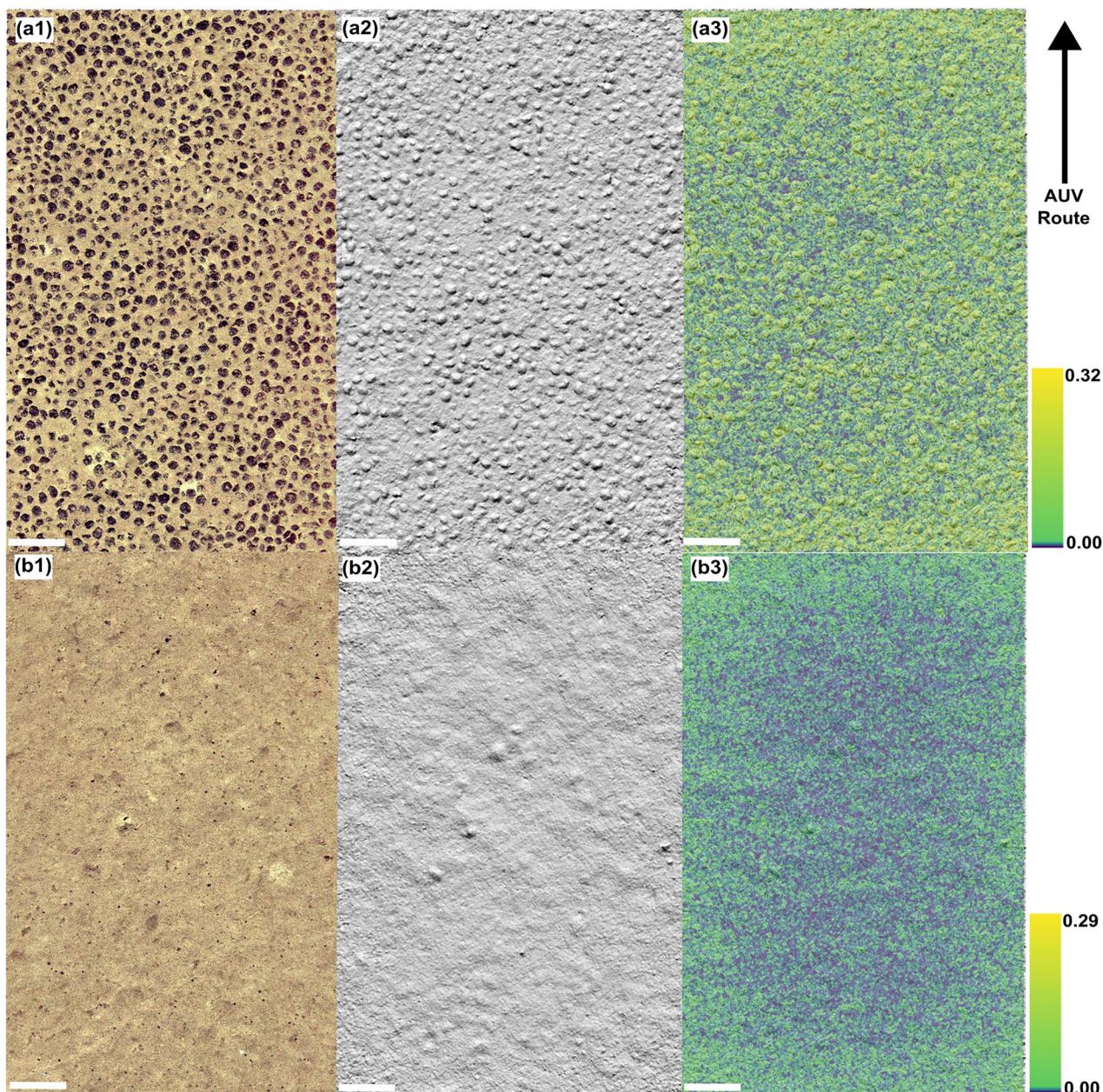
## DISCUSSION

The acquisition of high-resolution data with aerial coverage of  $37.34 \text{ km}^2$ , the largest AUV-based

dataset presented until now in deep-sea nodule research, provided the unique opportunity to study the spatial distribution of nodules in a geomorphologically complex terrain using a spatial resolution ranging between 2 mm (orthophoto-mosaics) and 2 m pixel size (MBES grids).

The combined use of seafloor images, MBES, SSS data, and different ML algorithms showed that the nodule coverage varied significantly (locally up to 100%) even on a meter scale (Figs. 5, 6). Such variations have been found in other parts of eastern CCFZ and Peru Basin (Wiedicke & Weber, 1996; Gazis et al., 2018; Peukert et al., 2018; Gazis & Greinert, 2021; Alevizos et al., 2022) and are related to the interaction of the seafloor topography, sediment thickness, bottom currents, and the spatial variability of biogeochemical processes (Cochonat et al., 1992; Mewes et al., 2014; Sharma, 2017; Volz et al., 2018; Paul et al., 2019; Hein et al., 2020). This study examined only the geomorphological component of this interplay, providing insights into the role of local topography on nodule occurrences.

Higher nodule coverage and larger nodules were on relatively flat seafloor with gentle ( $< 3^\circ$ ), eastward-facing slopes (Figs. 5, 6, 11). Both variable importance measures (Fig. 8) and PDPs (Fig. 11) showed an increased importance of eastward-facing slopes, which points toward a predominance of westward-moving bottom currents. Long-term observations in the B4 domain of the GSR contract area show that the mean bottom current direction is northwestward (Hayes, 1979; Juan et al., 2018; GSR, 2018). The predominant mean bottom flow toward NW is steady with slow velocity (0–3 cm/s) and only



**Figure 12.** (a1) Orthophoto-mosaic (2 mm pixel size) from 10 overlapping images of a seafloor part with nodule coverage, (a2) hill-shade of the image-derived DEM (5 mm pixel size) and (a3) image-derived seafloor roughness (5 mm pixel size) from the same seafloor part. (b) Orthophoto-mosaic (2 mm pixel size) of 10 overlapping images from a seafloor part without nodule coverage), (b2) hill-shade of the image-derived DEM (5 mm pixel size) and (b3) image-derived seafloor roughness from the same seafloor part. The two image-derived roughness surfaces have different univariate distributions and value ranges (Supplementary Fig. S14). The VRM was used as a roughness index. The white scale bar represents 0.5 m on the seafloor.

periodically interrupted by tidal currents (NW and SE) and sporadic by short-in-time benthic storms (6–15 cm/s), which are heading SE (Hayes, 1979; Juan et al., 2018; GSR, 2018). The NW steady and slow bottom currents contribute to the creation of

favorable conditions for the polymetallic nodule growth (Skornyakova & Murdmaa, 1992; von Stackelberg, 1997; Mewes et al., 2014; Kuhn et al., 2017; Juan et al., 2018; Volz et al., 2018): (a) erosion and removal of fine sediments, leaving coarser par-

## An Interpretable Multi-Model Machine Learning Approach

ticles that serve as nodule nuclei behind; and (b) continuous contact between nodule surface and well-oxygenated bottom water that allows the precipitation of hydroxides in polymetallic nodules. Geochemical studies within the eastern CCFZ have shown higher nodule coverage of larger sizes to sites with less clay-size particles and deeper oxygen penetration depths, particularly for sampling sites within the GSR contract area (Mewes et al., 2014; GSR, 2018; Volz et al., 2018). Older studies in the CCFZ showed higher nodule coverage on flat seafloor parts with weak ( $\sim 3 \text{ cm s}^{-1}$ ) but steady near-bottom currents (Skorniyakova & Murdmaa, 1992; Yamazaki & Sharma, 2000; Kuhn et al., 2017; Juan et al., 2018; Yoo et al., 2018; Alevizos et al., 2022). The importance of slope and slope orientation has also been shown in other ML spatial studies within the CCFZ (Kuhn & Rühlemann, 2021) and the Peru Basin at a nodule field with well-pronounced sediment furrows (Gazis & Greinert, 2021).

Smaller nodule coverage was observed mainly at four sub-environments (Figs. 5, 6): (a) in seafloor parts with typically higher bottom current velocities such as gullies, furrows or the foot of slopes; (b) in topographic highs, with steep slopes such as horsts and volcanic knolls; (c) in the deeper parts of basins around and inside sub-circular depressions; and (d) in areas with basement exposure, rocky outcrops, and carbonate cemented sediment layers.

The PDP of bathymetry showed a bathymetric range of  $\sim 100 \text{ m}$  that favored the nodule presence (Fig. 11). Broader-scale studies in the CCFZ and Peru Basin showed the maximum coverage and abundance occurring within a small bathymetric range (50–100 m) in water depths nearly below the local CCD, resulting in similar to our study response curves (Weber et al. 2000; von Stackelberg, 2000; Sharma, 2017). The PDP of BPI100-300 (Fig. 11) showed a small response to nodule coverage (i.e., few to no polymetallic nodules) in local depressions. Studies have shown that local depressions exhibit geochemical conditions different from their surroundings, having: (a) higher amount of fine-grain sediments (clay), which is trapped in the local depressions during sediment transportation (Wiedicke & Weber, 1996); fine sediments cannot act as nuclei for polymetallic nodule formation due to their size and typically small content of mobilized manganese, which hinders polymetallic nodule formation and growth (Mewes et al., 2014); and (b) increased organic carbon content that shifts the Mn-redox closer to the seafloor, where the limited (see

point a) diagenetically mobilized manganese is released into the bottom water without supporting the polymetallic nodule formation and growth (Wiedicke & Weber, 1996; Marchig et al., 2001; Paul et al., 2019).

Similar to local depressions, topographic highs also have limited nodule coverage related to increased slopes, reduced sediment thickness (e.g., shallower or outcropping basement), and local variations in sediment geochemical characteristics (e.g., POC fluxes, TOC contents, and Mn-redox depth) as a result of the fine-sediment winnowing and downslope transportation from the locally intensified or deflected bottom currents (Skorniyakova & Murdmaa, 1992; von Stackelberg, 1997; Volz et al., 2018). The studies mentioned above showed that even small height differences of  $< 10\text{--}25 \text{ m}$  are associated with such changes, which agrees with our findings.

The construction of seafloor orthophoto-mosaics further enhanced the geomorphological analysis and interpretation of nodule occurrences (Figs. 3, Supplementary Fig. S5). Deep sea orthophoto-mosaicking is an emerging field with several challenges, particularly regarding AUV-based camera systems (Song et al., 2022; She et al., 2023; Song et al., 2024), and only a few studies have used AUV-based orthophoto-mosaics from such depth (but in lower resolution) within nodules fields (Simon-Lledó et al., 2019c; Gausepohl et al., 2020). This study's results showed a great potential for deep-sea orthophoto-mosaicking even when it is based on long AUV transects usually done for fauna or nodule coverage mapping.

Although the geomorphological predictors describe the nodule coverage spatial distribution in an interpretable way linked to natural processes, the hydroacoustic backscatter contributed the most in all ML algorithms. MBES BS and ARA\_MF were the most important predictors in all ML models (Fig. 8). Older studies have described the effect of seafloor roughness created by nodules on sound waves and the measured backscatter intensity in MBES and SSS (Weydert, 1991; Scanlon & Masson, 1992; Chunhui et al., 2015; Machida et al., 2019). Here, we showed for the first time this mm-scale seafloor roughness from parts with and without nodule coverage in a quantitative way using the image-derived DEM generated during the orthophoto mosaic processing (Fig. 12). MBES and SSS backscatter captured the nodule coverage variations, being able to discriminate not only sub-environ-

ments with and without nodule coverage but also sub-environments with varying degrees of nodule coverage (Fig. 9). MBES backscatter and SSS can be used complementarily, as excluding one results in poorer predictive performance (Fig. 9, Supplementary Table S15). This finding agrees with a previous study in which the combined use of both systems increased interpretation and modeling performance (Janowski et al., 2021).

MBES ARA analyses showed that MBES backscatter acquisition is most meaningful for incident angles within 25–55°, a finding that is consistent with the existing guidelines for backscatter-optimized surveys (Lurton & Lamarche, 2015). The MBES backscatter discriminative ability, even over transitional areas, is attributed to the combination of high frequency (400 kHz), low altitude survey (60 m) and small beam opening angles ( $0.7^\circ \times 0.7^\circ$ ), which resulted in a small footprint size ( $0.73 \text{ m} \times 0.73 \text{ m}$  in nadir on a flat seafloor) and short wavelength ( $\sim 4 \text{ mm}$ ), which was in the order or even smaller than the nodules' dimensions (mm to cm scale). MBES ARA derivatives impedance (primary) and volume (secondary) contributed to the model performance in all models except GLM (Fig. 8). This finding agrees with other ML spatial studies, where ARA and its derivatives increased the ML prediction performance (Hasan et al., 2014; Alevizos & Greinert, 2018; Menandro et al., 2022; Misiuk & Brown, 2022; Porskamp et al., 2022). The lower contribution of volume backscattering is attributed to the limited penetration depth of the sound waves at 400 kHz, which is typically  $< 0.5 \text{ m}$  for silt/clay sediments ( $\phi > 4.5$ ) and  $< 0.25 \text{ m}$  for larger grain sizes (Huff, 2008; Gaida et al., 2018). The main component of volume scattering is expected to be the buried nodules located at a maximum depth of 5–10 cm.

In future studies, it would be interesting to use backscatter mosaics and ARA from multifrequency MBES because the use of higher and lower frequencies (e.g., 100–600 kHz) and pulse lengths (e.g.,  $\mu\text{s}$  to ms range) has shown to improve ML predictive performance and interpretation (Menandro et al., 2022; Misiuk & Brown, 2022). A holistic modeling approach should also include geophysical-derived data obtained by SBP, such as the thickness and composition of the semi-liquid bottom sediment layer, which is related to nodule coverage (Cochonat et al., 1992; Yoo et al., 2018; Dreisetl, 2016; Lipton et al., 2016; Alevizos et al., 2022). Similarly, the use of geochemical information would contribute to

spatial modeling. This approach is used in spatial studies in shallower depths, where hundreds of well-distributed samples exist and can be interpolated using ML algorithms (e.g., Diesing et al., 2017; 2020a; Spiegel et al., 2024). However, in deep-sea environments, geochemical data rely on sparse local observations of cm diameter (e.g., gravity corers), which are typically clustered around areas of particular interest, hindering their use in spatial models where continuous surfaces are needed as input data to generate spatially continuous output data. Studies have shown increased fine-scale spatial variability of geochemical properties related to the spatial distribution and size of deep-sea polymetallic nodules, highlighting the need for intensifying efforts on deep-sea seafloor sampling (Mewes et al., 2014; Volz et al., 2018; Paul et al., 2019).

Regarding the modeling performance, RFM had the best fit, which agrees with other studies showing its ability to outperform other ML algorithms in prediction error (Stephens & Diesing, 2014; Diesing & Stephens, 2015; Herkül et al., 2017; Li et al., 2017; Fernández-Delgado et al., 2019). RFM is currently the “workhorse” in deep-sea predictive mapping (Lawson et al., 2017; Gazis et al., 2018; Diesing, 2020b; Gazis & Greinert, 2021; Uhlenkott et al., 2022; Josso et al., 2023). However, here, it yielded the smallest range in predicted values due to the averaging step between the predicted values from each tree at the last step of the RFM algorithm (Breiman, 2001a). The non-tree-based core of the other four models (i.e., linear, penalized B-splines, 3<sup>rd</sup>-degree polynomial and sigmoid activation function) allowed for a more substantial value range extrapolation (Fig. 5). NNM had the second-best predictive performance on the testing data but the worst in spatial block CV. Less complex models such as GLM and GAM had better resampling performance in spatial block CV than SVM and NNM, indicating a better generalization ability when a big fraction of the training data is kept out (Fig. 7). This fact is attributed to the smoother functions that are more transferable (and interpretable) over larger areas (Wenger & Olden, 2012; Lauria et al., 2015; Stienessen et al., 2021). In addition, NNM is more sensitive to hyperparameter optimization than GLM, GAM, and RFM (Supplementary Tables S2–S9). The same applies to SVM, where the kernel type and other parameters (e.g., polynomial degree and cost function) largely influenced the model performance (Supplementary Tables S6–S7), a finding which agrees with previous

## An Interpretable Multi-Model Machine Learning Approach

comparative studies (e.g., Probst et al., 2019). To our knowledge, this is the first time that NNM and SVM were used for a regression task in seafloor habitat mapping, as literature research only provided applications for seafloor classification (Hasan et al., 2012; Diesing & Stephens, 2015; Alevizos & Greinert, 2018; Trzcinska et al., 2020; Janowski et al., 2021; Breyer et al., 2023).

Independent of the model used, the spatial block CV approach could result in a large CV error (pessimistic approach) due to the gaps created in feature space (Roberts et al., 2017; Hao et al., 2020; Wadoux et al., 2021). The block optimization minimized the geographical and feature space distances between training data and the prediction area to a degree but not entirely (Fig. 4). Nevertheless, this approach was preferred over other spatial CV methods such as spatial leave-one-location out (Meyer et al., 2018), Euclidean buffered distances (Hengl et al., 2018), nearest neighbor distance matching (Milà et al., 2022), and spatial covariance weighting (Misiuk & Brown, 2023) because of dataset size and survey layout (Figs. 1, 7). The abovementioned methods have been designed and tested with fewer samples ( $< < 1000$ ) distributed throughout an entire area but not in linear transects like here. In contrast, spatial blocking is appropriate for large datasets and large regions, being computationally efficient (fewer training points per model iteration) and providing flexibility in partitioning the geographical space (Roberts et al., 2017; Valavi et al., 2019). The model evaluation based on a spatial block CV was done using held-out data from each spatial block separately (Fig. 7). An alternative approach would be to use all held-out data from spatial blocks at once (global  $R^2$ ), which better represents a model's ability to predict over larger areas (Meyer & Pebesma, 2022). However, we preferred the local  $R^2$  from each spatial block to get (a) the strictest evaluation for each model (Supplementary Table S14) and (b) the comparison between different spatial blocks, identifying the spatial blocks that model predictions converged or diverged (Fig. 7). The partitioning of the geographical and feature space represents better the error related to extrapolation and transferability although it increases the AOA (Wenger & Olden, 2012; Roberts et al., 2017; Fourcade et al., 2018; Hao et al., 2020; Gazis & Greinert, 2021; Meyer & Pebesma, 2021).

Here, the AOA was large and similar for all models ( $> 91\%$  of the study area) and enlightened the reasons behind the poor performance of all

models in a specific resampling spatial block (Figs. 6, 7). The AOA and SD among model predictions provide valuable information about the models' trustworthiness and should be offered next to their predictions. These two approaches could guide future sampling within sub-areas with high SD and low model applicability. Despite the importance of spatial CV and extrapolation assessment when using ML algorithms, the literature research yielded only few studies that have offered such information aside from ML-based seafloor spatial predictions (Misiuk et al., 2019; Diesing, 2020a; Dolan et al., 2021; Gazis & Greinert, 2021; Spiegel et al., 2024).

It is worth mentioning that a detailed comparison of performance metrics was not the primary goal of this study. Each spatial problem is unique, and each algorithm has several issues to address, such as model accuracy, extrapolation, interpretability, and computational speed. Here, no ML method could outperform the others in all aspects. The multi-model approach provides confidence in the derived nodule coverage distribution because all models predicted similar distributions and variable ranking. The true benefit of a multi-model approach is the increased confidence in areas with good spatial agreement (i.e., low SD) and understanding of the contribution of specific predictors. Complex and simple models perform similarly well in the presence of meaningful predictors and training data that capture the feature space and variability well (Rudin, 2019). As such, we see ML models as derivatives for predictions based on data that have captured the underlying relationships and interactions between the response variable and predictors, but for which our conceptual understanding is still developing (Breiman, 2001b; Shmueli, 2010). Thus, although computational algorithms are used for the prediction, the initial phase of such studies is and needs to be very data-centric, emphasizing acquisition of a large amount of high-quality data at the right locations. A priori knowledge of the right (and technically feasible) sampling locations considering each study area's unique characteristics is difficult, particularly in the absence of legacy data (a typical problem in deep-sea research). Recently, proposed AUV-based future sampling approaches (Shields et al., 2023) and domain knowledge could help achieve a balanced geographical and feature space sample coverage.

## CONCLUSIONS

Deep-sea mining research is characterized by data scarcity, knowledge gaps, and lack of high-resolution studies (Amon et al., 2022). This work provided clear insights into the spatial distribution of polymetallic nodules within a potential mining field in the Clarion–Clipperton fracture zone, providing the basis for resource estimation, mining path planning, and benthic habitat mapping therein. The multi-model and interpretable analysis combined with angular range analysis and orthophoto-mosaicking can shed light on machine learning models' 'black box' character, giving us valuable feedback regarding each predictor variable's statistical and natural contribution and advancing our knowledge of deep-sea polymetallic nodule occurrences.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the captains and crews of RV SONNE and MV Island Pride for their essential assistance during the SO268 and MANGAN 2021 expeditions. We thank the Ocean Infinity AUV team for completing the extensive AUV dives and acquiring the presented data. We thank GSR (DEME group) for chartering the AUV used in this study. We thank saltation GmbH & Co. KG for fine-tuning the H<sup>2</sup>SOM software for this dataset. We are grateful to Karl Heger (Data Science Unit, GEOMAR) for his valuable help in image color normalization, Anastasios Poulos-Sidiropoulos for his assistance with the MBES and SSS data post-processing and Mario Veloso-Alarcon for his contribution to the drawing of the AUV in Supplementary Figure S9. Finally, we thank the GEOMAR Library team for its support in gathering the needed bibliography. This research was initiated by the European collaborative project MiningImpact 2 and received national funding through the German Ministry of Education and Research (grant no. 860 03F0812A). This is publication #65 of the Deep-Sea Monitoring group at GEOMAR.

## FUNDING

Open Access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY

All data associated with the results are presented in the paper and as supplementary material.

## DECLARATIONS

**Conflict of Interest** The authors declare no competing interests.

## OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## SUPPLEMENTARY INFORMATION

The online version contains supplementary material available at <https://doi.org/10.1007/s11053-024-10393-7>.

## REFERENCES

- Alevizos, E., & Greinert, J. (2018). The hyper-angular cube concept for improving the spatial and acoustic resolution of MBES backscatter angular response analysis. *Geosciences*, 8(12), 446.
- Alevizos, E., Huvenne, V. A. I., Schoening, T., Simon-Lledó, E., Robert, K., & Jones, D. O. B. (2022). Linkages between sediment thickness, geomorphology and MN nodule occurrence: New evidence from AUV Geophysical mapping in the clarion-clipperton zone. *Deep Sea Research Part I Oceanographic Research Papers*, 179, 103645.
- Amon, D. J., Gollner, S., Morato, T., Smith, C. R., Chen, C., Christiansen, S., Currie, B., Drazen, J. C., Fukushima, T., Gianni, M., Gjerde, K. M., Gooday, A. J., Grillo, G. G., Haeckel, M., Joyini, T., Ju, S.-J., Levin, L. A., Metaxas, A., Mianowicz, K., & Molodtsova, T. N. (2022). Assessment of

## An Interpretable Multi-Model Machine Learning Approach

- scientific gaps related to the effective environmental management of deep-seabed mining. *Marine Policy*, 138, 105006.
- Amon, D. J., Ziegler, A. F., Dahlgren, T. G., Glover, A. G., Goineau, A., Gooday, A. J., Wiklund, H., & Smith, C. R. (2016). Insights into the abundance and diversity of abyssal megafauna in a polymetallic-nodule region in the Eastern Clarion-Clipperton Zone. *Scientific Reports*, 6(1), 30492.
- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2), 93–115.
- Beck, M. W. (2018). NeuralNetTools: Visualization and analysis tools for neural networks. *Journal of Statistical Software*, 85(11), 1–20.
- Behrens, T., & Viscarra Rossel, R. A. (2020). On the interpretability of predictors in spatial data science: The information horizon. *Scientific Reports*, 10(1), 16737.
- Bekins, B. A., Spivack, A. J., Davis, E. E., & Mayer, L. A. (2007). Dissolution of biogenic ooze over basement edifices in the equatorial Pacific with implications for hydrothermal ventilation of the oceanic crust. *Geology*, 35(8), 679.
- Berger, W. H., Adelseck, C. G., Jr., & Mayer, L. A. (1976). Distribution of carbonate in surface sediments of the Pacific Ocean. *Journal of Geophysical Research*, 81(15), 2617–2627.
- Breiman, L. (2001a). Random forests. *Machine Learning*, 45(1), 5–32.
- Breiman, L. (2001b). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science A Review Journal of the Institute of Mathematical Statistics*, 16(3), 199–231.
- Breyer, G., Bartholomä, A., & Pesch, R. (2023). The suitability of machine-learning algorithms for the automatic acoustic seafloor classification of hard substrate habitats in the German bight. *Remote Sensing*, 15(16), 4113.
- Bühlmann, P., & Hothorn, T. (2007). Boosting algorithms: Regularization, prediction and model fitting. *Statistical Science A Review Journal of the Institute of Mathematical Statistics*, 22(4), 477–505.
- Calder, B. R., & Mayer, L. A. (2003). Automatic processing of high-rate, high-density multibeam echosounder data. *Geochemistry Geophysics Geosystems*. <https://doi.org/10.1029/2002gc000486>.
- Chunhui, T., Xiaobing, J., Aifei, B., Hongxing, L., Xianming, D., Jianping, Z., Chunhua, G., Tao, W., & Wilkens, R. (2015). Estimation of manganese nodule coverage using multi-beam amplitude data. *Marine Georesources & Geotechnology*, 33(4), 283–288.
- Cochonat, P., Le Suavé, R., Charles, C., Greger, B., Hoffert, M., Lenoble, J. P., Meunier, J., & Pautot, G. (1992). First in situ studies of nodule distribution and geotechnical measurements of associated deep-sea clay (Northeastern Pacific Ocean). *Marine Geology*, 103(1–3), 373–380.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297.
- Craig, J. D. (1979). The relationship between bathymetry and ferromanganese deposits in the north equatorial Pacific. *Marine Geology*, 29(1–4), 165–186.
- De Smet, B., Simon-Lledó, E., Mevenkamp, L., Pape, E., Pasotti, F., Jones, D. O. B., & Vanreusel, A. (2021). The megafauna community from an abyssal area of interest for mining of polymetallic nodules. *Deep Sea Research Part I Oceanographic Research Papers*, 172, 103530.
- Diesing, M. (2020b). Deep-sea sediments of the global ocean. *Earth System Science Data*, 12(4), 3367–3381.
- Diesing, M., Kröger, S., Parker, R., Jenkins, C., Mason, C., & Weston, K. (2017). Predicting the standing stock of organic carbon in surface sediments of the North-West European continental shelf. *Biogeochemistry*, 135(1–2), 183–200.
- Diesing, M., & Stephens, D. (2015). A multi-model ensemble approach to seabed mapping. *Journal of Sea Research*, 100, 62–69.
- Diesing, M., Thorsnes, T., & Bjarnadóttir, L. R. (2020a). Organic carbon in surface sediments of the North Sea and Skagerrak. *Biogeosciences Discussions*, 2020, 1–30.
- Dolan, M. F. J., Ross, R. E., Albretsen, J., Skarðhamar, J., Gonzalez-Mirelis, G., Bellec, V. K., Buhl-Mortensen, P., & Bjarnadóttir, L. R. (2021). Using spatial validity and uncertainty metrics to determine the relative suitability of alternative suites of oceanographic data for seabed biotope prediction. A case study from the Barents sea, Norway. *Geosciences*, 11(2), 48.
- Dreisetl, I. (2016). Deep sea exploration for metal reserves—objectives, methods and look into the future. In T. Abramowski (Ed.), *Deep sea mining value chain: Organization, technology and development* (pp. 105–118). IOM.
- Dutkiewicz, A., Judge, A., & Müller, R. D. (2020). Environmental predictors of deep-sea polymetallic nodule occurrence in the global ocean. *Geology*, 48(3), 293–297.
- Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science A Review Journal of the Institute of Mathematical Statistics*, 11(2), 89–121.
- Ellefmø, S. L., & Kuhn, T. (2021). Application of soft data in nodule resource estimation. *Natural Resources Research*, 30(2), 1069–1091.
- Fernández-Delgado, M., Sirsat, M. S., Cernadas, E., Alawadi, S., Barro, S., & Febrero-Bande, M. (2019). An extensive experimental survey of regression methods. *Neural Networks*, 111, 11–34.
- Florinsky, I. V. (2017). An illustrated introduction to general geomorphometry. *Progress in Physical Geography*, 41(6), 723–752.
- FMI/Flanders Marine Institute. (2019). Maritime Boundaries Geodatabase, version 11, Retrieved April 24, 2023, from <https://www.marineregions.org/>.
- Fonseca, L., & Calder, B. (2005). Geocoder: An efficient Backscatter map constructor. U.S. Hydrographic Conference Retrieved April 24, 2023, from <https://scholars.unh.edu/ccom/339/>.
- Fonseca, L., Brown, C., Calder, B., Mayer, L., & Rzhano, Y. (2009). Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and multibeam echosounder angular signatures. *Applied Acoustics*, 70(10), 1298–1304.
- Fonseca, L., & Mayer, L. (2007). Remote estimation of surficial seafloor properties through the application Angular Range Analysis to multibeam sonar data. *Marine Geophysical Research*, 28(2), 119–126.
- Fourcade, Y., Besnard, A. G., & Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Biogeography A Journal of Macroecology*, 27(2), 245–256.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>.
- Gaida, T. C., Mohammadloo, T. H., Snellen, M., & Simons, D. G. (2019). Mapping the seabed and shallow subsurface with multi-frequency multibeam echosounders. *Remote Sensing*, 12(1), 52.
- Gaida, T., Tengku Ali, T., Snellen, M., Amiri-Simkooei, A., van Dijk, T., & Simons, D. (2018). A multispectral Bayesian classification method for increased acoustic discrimination of seabed sediments using multi-frequency multibeam backscatter data. *Geosciences*, 8(12), 455.
- Garson, D. G. (1991). Interpreting neural-network connection weights. *Artificial Intelligence Expert*, 6(4), 46–51.
- Gausepohl, F., Hennke, A., Schoening, T., Köser, K., & Greinert, J. (2020). Scars in the abyss: Reconstructing sequence, location and temporal change of the 78 plough tracks of the 1989

- DISCOL deep-sea disturbance experiment in the Peru Basin. *Biogeosciences*, 17(6), 1463–1493.
- Gazis, I.-Z. (2020). Processed EM122 multibeam swath bathymetry collected during SONNE cruise SO268/1 inside the Belgian License Area in Clarion Clipperton Zone, Pacific. GEOMAR—Helmholtz Centre for Ocean Research Kiel, PANGAEA. (2020). Retrieved April 24, 2023, from <https://doi.org/10.1594/PANGAEA.915767>.
- Gazis, I.-Z., & Greinert, J. (2021). Importance of spatial autocorrelation in machine learning modeling of polymetallic nodules, model uncertainty and transferability at local scale. *Minerals*, 11(11), 1172.
- Gazis, I.-Z., Schoening, T., Alevizos, E., & Greinert, J. (2018). Quantitative mapping and predictive modeling of MN Nodules' distribution from hydroacoustic and optical AUV data linked by random forests machine learning. *Biogeosciences*, 15(23), 7347–7377.
- Glasby, G. P. (1973). Distribution of manganese nodules and lebensspuren in underwater photographs from the Carlsberg Ridge, Indian Ocean. *New Zealand Journal of Geology and Geophysics*, 16(1), 1–17.
- Global Sea Mineral Resources NV (GSR). (2018). Environmental Impact Statement. Small-Scale Testing of Nodule Collector Component on the Seafloor of the Clarion-Clipperton Fracture Zone and its Environmental Impact, p. 337. Retrieved April 24, 2023, from <https://www.isa.org/jm/minerals/environmental-impact-assessments>.
- Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143–151.
- Greenwell, B. (2017). pdp: An R package for constructing partial dependence plots. *The R Journal*, 9(1), 421.
- Günther, F., & Fritsch, S. (2010). Neuralnet: Training of neural networks. *The R Journal*, 2(1), 30.
- Haeckel, M., & Linke, P. (2021). RV SONNE Cruise Report SO268—Assessing the Impacts of Nodule Mining on the Deep-Sea Environment: Nodule Monitoring, Manzanillo (Mexico)—Vancouver (Canada), 17.02. – 27.05.2019 GEOMAR Report, N. Ser. 059. GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Kiel, Germany, 359 + Appendix (in all 802) pp. (2021). Retrieved April 24, 2023, from <https://oceanrep.geomar.de/id/eprint/54402/>.
- Halbach, P., Scherhag, C., Hebisch, U., & Marchig, V. (1981). Geochemical and mineralogical control of different genetic types of deep-sea nodules from the Pacific Ocean. *Mineralium Deposita*, 16(1), 59–84.
- Hao, T., Elith, J., Lahoz-Monfort, J. J., & Guillera-Aroita, G. (2020). Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography*, 43(4), 549–558.
- Hari, V. N., Kalyan, B., Chitre, M., & Ganesan, V. (2018). Spatial modeling of deep-sea ferromanganese nodules with limited data using neural networks. *IEEE Journal of Oceanic Engineering*, 43(4), 997–1014.
- Hasan, C., Ierodiaconou, R., Laurenson, D., & Schimel, L. (2014). Integrating Multibeam Backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. *PLoS ONE*, 9, e97339.
- Hasan, R., Ierodiaconou, D., & Monk, J. (2012). Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar. *Remote Sensing*, 4(11), 3427–3443.
- Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science A Review Journal of the Institute of Mathematical Statistics*, 1(3), 297–310.
- Hayes, S. P. (1979). Benthic current observations at DOMES sites A, B, and C in the tropical north pacific ocean. In J. L. Bischoff & D. Z. Piper (Eds.), *Marine geology and oceanography of the pacific manganese nodule province*. *Marine science*. (Vol. 9). Springer.
- Hein, J. R., Koschinsky, A., & Kuhn, T. (2020). Deep-ocean polymetallic nodules as a resource for critical materials. *Nature Reviews Earth & Environment*, 1(3), 158–169.
- Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B. M., & Gräler, B. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 6(e5518), e5518.
- Herkül, K., Peterson, A., & Paekivi, S. (2017). Applying multi-beam sonar and mathematical modeling for mapping seabed substrate and biota of offshore shallows. *Estuarine Coastal and Shelf Science*, 192, 57–71.
- Hijmans, R. (2022). raster: Geographic data analysis and modeling, R package version 3.5-21. Retrieved June 24, 2023, from <https://CRAN.R-project.org/package=raster>.
- Hillman, J. I. (2017). Validation of automated supervised segmentation of Multibeam Backscatter data from the Chatham Rise, New Zealand. *Marine Geophysical Research*, 39, 205–227.
- Hothorn, T., Buehlmann, P., Kneib, T., Schmid, M., & Hofner, B. (2021). mboost: Model-based boosting, R package version 2.9-5. Retrieved April 24, 2023, from <https://CRAN.R-project.org/package=mboost>.
- Huff, L. C. (2008). Acoustic remote sensing as a tool for habitat mapping in Alaska waters. *Marine habitat mapping technology for alaska* (pp. 29–46). Alaska Sea Grant, University of Alaska Fairbanks. <https://doi.org/10.4027/mhmta.2008.03>.
- International Seabed Authority (ISA). (2023). PMN Exploration Areas, PMN Reserved Areas, APEI Areas. Retrieved April 24, 2023, from <https://www.isa.org/jm/exploration-contracts/maps/>.
- Jackson, D. R., & Briggs, K. B. (1992). High-frequency bottom backscattering: Roughness versus sediment volume scattering. *The Journal of the Acoustical Society of America*, 92(2), 962–977.
- Jackson, D. R., Winebrenner, D. P., & Ishimaru, A. (1986). Application of the composite roughness model to high-frequency bottom backscattering. *The Journal of the Acoustical Society of America*, 79(5), 1410–1422.
- Janowski, L., Wroblewski, R., Dworniczak, J., Kolakowski, M., Rogowska, K., Wojcik, M., & Gajewski, J. (2021). Offshore benthic habitat mapping based on object-based image analysis and geomorphometric approach. A case study from the Slupsk Bank, Southern Baltic Sea. *The Science of the Total Environment*, 801, 149712.
- Josso, P., Hall, A., Williams, C., Le Bas, T., Lusty, P., & Murton, B. (2023). Application of random-forest machine learning algorithm for mineral predictive mapping of Fe-Mn crusts in the World Ocean. *Ore Geology Reviews*, 162(105671), 105671.
- Juan, C., Van Rooij, D., & De Bruycker, W. (2018). An assessment of bottom current controlled sedimentation in Pacific Ocean abyssal environments. *Marine Geology*, 403, 20–33.
- Kaikkonen, L., Virtanen, E. A., Kostamo, K., Lappalainen, J., & Kotilainen, A. T. (2019). Extensive coverage of marine mineral concretions revealed in shallow shelf sea areas. *Frontiers in Marine Science*, 6, 541.
- Karasiak, N., Dejoux, J.-F., Monteil, C., & Sheeren, D. (2022). Spatial dependence between training and test sets: Another pitfall of classification accuracy assessment in remote sensing. *Machine Learning*, 111(7), 2715–2740.
- Karatzoglou, A., Smola, A., & Hornik, K. (2023). kernlab: Kernel-Based Machine Learning Lab. R package version 0.9-32, 2023. Retrieved April 24, 2023, from <https://CRAN.R-project.org/package=kernlab>.
- Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). kernlab—An S4 package for Kernel Methods in R. *Journal of Statistical Software*, 11(9), 1–20.

## An Interpretable Multi-Model Machine Learning Approach

- Kodagali, V. (1988). Influence of regional and local topography on the distribution of polymetallic nodules in Central Indian Ocean Basin. *Geo-Marine Letters*, 8(3), 173–178. <https://doi.org/10.1007/bf02326094>.
- Köser, K., Song, Y., Petersen, L., Wenzlaff, E. & Woelk, F. (2021). Robustly removing deep sea lighting effects for visual mapping of Abyssal Plains. arXiv.org (2021). Retrieved June 24, 2023, from <https://doi.org/10.48550/arXiv.2110.00480>.
- Kuhn, M. (2022). caret: Classification and regression training. R package version 6.0-92. Retrieved April 24, 2023, from <http://CRAN.R-project.org/package=caret>.
- Kuhn, T., & Rühlmann, C. (2021). Exploration of polymetallic nodules and resource assessment: A case study from the German contract area in the Clarion-Clipperton Zone of the tropical Northeast Pacific. *Minerals*, 11(6), 618.
- Kuhn, T., Wegorzewski, A., Rühlmann, C., & Vink, A. (2017). Composition, Formation, and Occurrence of Polymetallic Nodules. In R. Sharma (Ed.), *Deep-sea mining*. Springer. [https://doi.org/10.1007/978-3-319-52557-0\\_2](https://doi.org/10.1007/978-3-319-52557-0_2).
- Lauria, V., Power, A. M., Lordan, C., Weetman, A., & Johnson, M. P. (2015). Spatial transferability of habitat suitability models of Nephrops norvegicus among fished areas in the northeast Atlantic: Sufficiently stable for marine resource conservation? *PLoS ONE*, 10(2), e0117006.
- Lawson, E., Smith, D., Sofge, D., Elmore, P., & Petry, F. (2017). Decision forests for machine learning classification of large, noisy seafloor feature sets. *Computers & Geosciences*, 99, 116–124.
- Lecours, V., Devillers, R., Edinger, E. N., Brown, C. J., & Lucieer, V. L. (2017). Influence of artefacts in marine digital terrain models on habitat maps and species distribution models: A multiscale assessment. *Remote Sensing in Ecology and Conservation*, 3(4), 232–246.
- Lee, S. H., & Kim, K.-H. (2004). Side-scan sonar characteristics and manganese nodule abundance in the Clarion—Clipperton Fracture Zones, NE equatorial Pacific. *Marine Georesources & Geotechnology*, 22(1–2), 103–114.
- Li, J., Alvarez, B., Siwabessy, J., Tran, M., Huang, Z., Przeslawski, R., Radke, L., Howard, F., & Nichol, S. (2017). Application of random forest, generalised linear model and their hybrid methods with geostatistical techniques to count data: Predicting sponge species richness. *Environmental Modelling & Software*, 97, 112–129.
- Li, Z., Li, H., Hein, J. R., Dong, Y., Wang, M., Ren, X., Wu, Z., Li, X., & Chu, F. (2021). A possible link between seamount sector collapse and manganese nodule occurrence in the Abyssal Plains. *NW Pacific Ocean. Ore Geology Reviews*, 138, 104378.
- Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4(2), 4–22.
- Lipton, I., Nimmo, M., & Parianos, J. (2016). NI 43-101 Technical report. TOML Clarion Clipperton Zone Project, Pacific Ocean. Retrieved April 24, 2023, from <https://doi.org/10.13140/RG.2.2.23742.08000>.
- Lucieer, V., Huang, Z., & Siwabessy, J. (2016). Analyzing uncertainty in multibeam bathymetric data and the impact on derived seafloor attributes. *Marine Geodesy*, 39(1), 32–52.
- Lurton, X. & Lamarche, G. (Eds) (2015). Backscatter measurements by seafloor-mapping sonars. Guidelines and Recommendations. 200p. Retrieved January 24 1, 2024, from <https://doi.org/10.5281/zenodo.10089261>.
- Macdonald, K. C., Fox, P. J., Alexander, R. T., Pockalny, R., & Gente, P. (1996). Volcanic growth faults and the origin of Pacific abyssal hills. *Nature*, 380(6570), 125–129.
- Machida, S., Sato, T., Yasukawa, K., Nakamura, K., Iijima, K., Nozaki, T., & Kato, Y. (2019). Visualisation method for the broad distribution of seafloor Ferromanganese deposits. *Marine Georesources & Geotechnology*, 39(3), 267–279.
- Marchig, V., von Stackelberg, U., Hufnagel, H., & Durn, G. (2001). Compositional changes of surface sediments and variability of manganese nodules in the Peru Basin. *Deep Sea Research Part II Topical Studies in Oceanography*, 48(17–18), 3523–3547.
- Mayer, M. (2022). splitTools: Tools for data splitting. R package version 0.3.2). Retrieved June 24, 2023, from <https://CRAN.R-project.org/package=splitTools>.
- Mayer, L. (1981). Erosional troughs in deep-sea carbonates and their relationship to basement structure. *Marine Geology*, 39(1–2), 59–80.
- Mbani, B., Buck, V., & Greinert, J. (2023). An automated image-based workflow for detecting megabenthic fauna in optical images with examples from the Clarion-Clipperton Zone. *Scientific Reports*, 13(1), 8350.
- Mbani, B., Schoening, T., Gazis, I.-Z., Koch, R., & Greinert, J. (2022). Implementation of an automated workflow for image-based seafloor classification with examples from manganese-nodule covered seabed areas in the Central Pacific Ocean. *Scientific Reports*, 12(1), 15338.
- McCullagh, P. J., & Nelder, J. A. (1989). *Generalized linear models*. Chapman & Hall.
- Menandro, P. S., Bastos, A. C., Misiuk, B., & Brown, C. J. (2022). Applying a multi-method framework to analyze the multi-spectral acoustic response of the seafloor. *Frontiers in Remote Sensing*, 3, 860282.
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., Thuiller, W., Wüest, R. O., Zimmermann, N. E., & Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37(12), 1267–1281.
- Mewes, K., Mogollón, J. M., Picard, A., Rühlmann, C., Kuhn, T., Nöthen, K., & Kasten, S. (2014). Impact of depositional and biogeochemical processes on small scale variations in nodule abundance in the Clarion-Clipperton Fracture Zone. *Deep-Sea Research Part I Oceanographic Research Papers*, 91, 125–141.
- Meyer, H., Milà, C., Ludwig, M., & Linnenbrink, J. (2024). CAST: 'caret' Applications for Spatial-Temporal Models. R package version 0.9.0. Retrieved January 24, 2024, from <https://hanna-meyer.github.io/CAST/>.
- Meyer, H., & Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, 12(9), 1620–1633.
- Meyer, H., & Pebesma, E. (2022). Machine learning-based global maps of ecological variables and the challenge of assessing them. *Nature Communications*, 13(1), 2208. <https://doi.org/10.1038/s41467-022-29838-9>.
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software with Environment Data News*, 101, 1–9.
- Milà, C., Mateu, J., Pebesma, E., & Meyer, H. (2022). Nearest neighbour distance matching Leave-One-Out Cross-Validation for map validation. *Methods in Ecology and Evolution*, 13(6), 1304–1316.
- Misiuk, B., & Brown, C. J. (2022). Multiple imputation of multi-beam angular response data for high resolution full coverage seabed mapping. *Marine Geophysical Research*, 43(1), 7.
- Misiuk, B., & Brown, C. J. (2023). Improved environmental mapping and validation using bagging models with spatially clustered data. *Ecological Informatics*, 77, 102181.
- Misiuk, B., Diesing, M., Aitken, A., Brown, C. J., Edinger, E. N., & Bell, T. (2019). A spatially explicit comparison of quantitative and categorical modelling approaches for mapping seabed sediments using Random Forest. *Geosciences*, 9(6), 254.

- Mitchell, A., & Griffin, L. S. (2021). *Spatial measurements and statistics*. Esri Press.
- Molnar, C. (2018). iml: An R package for interpretable machine learning. *Journal of Open Source Software*, 3(26), 786.
- Moore, T. C., Jr., Mitchell, N. C., Lyle, M., Backman, J., & Pálíke, H. (2007). Hydrothermal pits in the biogenic sediments of the equatorial Pacific Ocean. *Geochemistry Geophysics Geosystems* G(3). <https://doi.org/10.1029/2006gc001501>.
- Pape, E., Bezerra, T. N., Gheerardyn, H., Buydens, M., Kieswetter, A., & Vanreusel, A. (2021). Potential impacts of polymetallic nodule removal on deep-sea meiofauna. *Scientific Reports*, 11(1), 19996.
- Parianos, J., Lipton, I., & Nimmo, M. (2021). Aspects of estimation and reporting of mineral resources of seabed polymetallic nodules: A contemporaneous case study. *Minerals*, 11(2), 200.
- Parianos, J., O'Sullivan, A., & Madureira, P. (2022). Geology of parts of the central and eastern Clarion Clipperton Zone. *Journal of Maps*, 18(2), 232–245.
- Paul, S. A. L., Haeckel, M., Bau, M., Bajracharya, R., & Koschinsky, A. (2019). Small-scale heterogeneity of trace metals including rare earth elements and yttrium in deep-sea sediments and porewaters of the Peru Basin, southeastern equatorial Pacific. *Biogeosciences*, 16(24), 4829–4849.
- Peukert, A., Schoening, T., Alevizos, E., Köser, K., Kwasnitschka, T., & Greinert, J. (2018). Understanding MN-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and Optical Data. *Biogeosciences*, 15(8), 2525–2549.
- Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., & Pélissier, R. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, 11(1), 4540.
- Porskamp, P., Schimel, A. C. G., Young, M., Rattray, A., Ladroit, Y., & Ierodiaconou, D. (2022). Integrating multibeam echosounder water-column data into benthic habitat mapping. *Limnology and Oceanography*, 67(8), 1701–1713.
- Probst, P., Bischl, B., & Boulesteix, A.-L. (2019). Tunability: Importance of hyperparameters of machine learning algorithms. *Journal of Machine Learning Research*, 20, 1–32.
- R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Riedmiller, M., & Braun, H. (2003). A direct adaptive method for faster backpropagation learning: the RPROP algorithm. In *IEEE International Conference on Neural Networks*.
- Riedmiller, M. (1994). Advanced supervised learning in multilayer perceptrons—From backpropagation to adaptive learning algorithms. *Computer Standards & Interfaces*, 16(3), 265–278.
- Robert, K., Jones, D. O. B., Roberts, J. M., & Huvenne, V. A. I. (2016). Improving predictive mapping of deep-water habitats: Considering multiple model outputs and ensemble techniques. *Deep Sea Research Part I Oceanographic Research Papers*, 113, 80–89.
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guisler-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.
- Sappington, J. M., Longshore, K. M., & Thompson, D. B. (2007). Quantifying landscape ruggedness for animal habitat analysis: A case study using bighorn sheep in the Mojave Desert. *The Journal of Wildlife Management*, 71(5), 1419–1426.
- Scanlon, K. M., & Masson, D. G. (1992). Fe–Mn nodule field indicated by GLORIA, north of the Puerto Rico Trench. *Geo-Marine Letters*, 12, 208–213. <https://doi.org/10.1007/BF02091840>.
- Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A., & Crowley, J. (2023). GGally: Extension to 'ggplot2'. R package version 2.2.0. Retrieved April 24, 2023, from <https://github.com/ggobi/ggally>.
- Schoening, T., Gazis, I.-Z. (2019). Summary of sizes, weights, counts and coverage of poly-metallic nodules from box cores taken during SONNE cruises SO268/1 and SO268/2. GEOMAR - Helmholtz Centre for Ocean Research Kiel, PANGAEA. Retrieved April 24, 2023, from <https://doi.org/10.1594/PANGAEA.904967>.
- Schoening, T., Jones, D. O., & Greinert, J. (2017). Compact-morphology-based poly-metallic nodule delineation. *Scientific Reports*, 7(1), 13338.
- Schoening, T., Kuhn, T., Jones, D. O. B., Simon-Lledo, E., & Nattkemper, T. W. (2016). Fully automated image segmentation for benthic resource assessment of poly-metallic nodules. *Methods in Oceanography*, 15–16, 78–89.
- Schratz, P., Muenchow, J., Iturriza, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using Spatial Data. *Ecological Modelling*, 406, 109–120. <https://doi.org/10.1016/j.ecolmodel.2019.06.002>.
- Sharma, R. (1993). Quantitative estimation of seafloor features from photographs and their application to nodule mining. *Marine Georesources & Geotechnology*, 11(4), 311–331.
- Sharma, R. (2017). Assessment of distribution characteristics of polymetallic nodules and their implications on deep-sea mining. In R. Sharma (Ed.), *Deep-Sea mining*. Springer. [https://doi.org/10.1007/978-3-319-52557-0\\_8](https://doi.org/10.1007/978-3-319-52557-0_8).
- Sharma, R., Khadge, N. H., & Jai Sankar, S. (2013). Assessing the distribution and abundance of seabed minerals from seafloor photographic data in the Central Indian Ocean Basin. *International Journal of Remote Sensing*, 34(5), 1691–1706.
- She, M., Song, Y., Nakath, D., Köser, K., (2023). Efficient Large-scale AUV-based Visual Seafloor Mapping. arXiv preprint arXiv:2308.0614, <https://doi.org/10.48550/arXiv.2308.06147>.
- Shields, J., Pizarro, O., & Williams, S. (2023). Feature space exploration for planning initial benthic AUV surveys. *Field Robotics*, 3(1), 652–686.
- Shmueli, G. (2010). To explain or to predict? *Statistical Science A Review Journal of the Institute of Mathematical Statistics*, 25(3), 289–310.
- Simon-Lledó, E., Amon, D. J., Bribiesca-Contreras, G., Cuvelier, D., Durden, J. M., Ramalho, S. P., Uhlenkott, K., Arbizu, P. M., Benoist, N., Copley, J., Dahlgren, T. G., Glover, A. G., Fleming, B., Horton, T., Ju, S.-J., Mejía-Saenz, A., McQuaid, K., Pape, E., Park, C., & Jones, D. O. B. (2023). Carbonate compensation depth drives abyssal biogeography in the northeast Pacific. *Nature Ecology & Evolution*, 7(9), 1388–1397.
- Simon-Lledó, E., Bett, B. J., Huvenne, V. A. I., Köser, K., Schoening, T., Greinert, J., & Jones, D. O. B. (2019a). Biological effects 26 years after simulated deep-sea mining. *Scientific Reports*, 9(1), 8040.
- Simon-Lledó, E., Bett, B. J., Huvenne, V. A. I., Schoening, T., Benoist, N. M. A., Jeffreys, R. M., Durden, J. M., & Jones, D. O. B. (2019b). Megafaunal variation in the abyssal landscape of the Clarion Clipperton Zone. *Progress in Oceanography*, 170, 119–133.
- Simon-Lledó, E., Bett, B. J., Huvenne, V. A., Schoening, T., Benoist, N. M., & Jones, D. O. (2019c). Ecology of a polymetallic nodule occurrence gradient: Implications for deep-sea mining. *Limnology and Oceanography*, 64(5), 1883–1894.

## An Interpretable Multi-Model Machine Learning Approach

- Simon-Lledó, E., Pomee, C., Ahokava, A., Drazen, J. C., Leitner, A. B., Flynn, A., Parianos, J., & Jones, D. O. B. (2020). Multi-scale variations in invertebrate and fish megafauna in the mid-eastern Clarion Clipperton Zone. *Progress in Oceanography*, 187, 102405.
- Skornyakova, N. S., & Murdmaa, I. O. (1992). Local variations in distribution and composition of ferromanganese nodules in the Clarion-Clipperton Nodule Province. *Marine Geology*, 103(1–3), 381–405.
- Song, Y., Nakath, D., She, M., & Köser, K. (2022). Optical imaging and image restoration techniques for deep ocean mapping: A comprehensive survey. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90(3), 243–267. <https://doi.org/10.1007/s41064-022-00206-y>.
- Song, Y., She, M., & Köser, K. (2024). Advanced underwater image restoration in complex illumination conditions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 209, 197–212. <https://doi.org/10.1016/j.isprsjprs.2024.02.004>.
- Spiegel, T., Diesing, M., Dale, A. W., Lenz, N., Schmidt, M., Sommer, S., Böttner, C., Fuhr, M., Kalapurakkal, H. T., Schulze, C.-S., & Wallmann, K. (2024). Modelling mass accumulation rates and 210PB rain rates in the Skagerrak: Lateral sediment transport dominates the sediment input. *Frontiers in Marine Science*, 11, 1331102.
- von Stackelberg, U. (2000). Manganese nodules of the Peru Basin (Chapter 8). In Cronan, D.S. (Ed.), *Handbook of Marine Mineral Deposits*. CRC Marine Science Series, 17: pp. 197–238, CRC Press, Boca Raton, Florida, USA. ISBN 10: 084938429X ISBN.
- Stephens, D., & Diesing, M. (2014). A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. *PLoS ONE*, 9(4), e93950.
- Stienessen, S. C., Rooper, C. N., Webe, T. C., Jones, D. T., Pirtle, J. L., & Wilson, C. D. (2021). Comparison of model types for prediction of seafloor trawlability in the Gulf of Alaska by using multibeam sonar data. *Fishery Bulletin*, 119(2–3), 184–196.
- Trzcinska, K., Janowski, L., Nowak, J., Rucinska-Zjadacz, M., Kruss, A., von Deimling, J. S., Pocwiardowski, P., & Tegowski, J. (2020). Spectral features of dual-frequency multibeam echosounder data for benthic habitat mapping. *Marine Geology*, 427, 106239.
- Tsune, A. (2021). Quantitative expression of the burial phenomenon of deep seafloor manganese nodules. *Minerals*, 11(2), 227.
- Uhlenkott, K., Meyn, K., Vink, A., & Martínez Arbizu, P. (2023). A review of megafauna diversity and abundance in an exploration area for polymetallic nodules in the eastern part of the Clarion Clipperton Fracture Zone (North East Pacific), and implications for potential future deep-sea mining in this area. *Marine Biodiversity*, 53(2), 22.
- Uhlenkott, K., Simon-Lledó, E., Vink, A., & Martínez Arbizu, P. (2022). Investigating the benthic megafauna in the eastern Clarion Clipperton Fracture Zone (north-east Pacific) based on distribution models predicted with random forest. *Scientific Reports*, 12(1), 8229.
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10(2), 225–232.
- Vanreusel, A., Hilario, A., Ribeiro, P. A., Menot, L., & Arbizu, P. M. (2016). Threatened by mining, polymetallic nodules are required to preserve abyssal epifauna. *Scientific Reports*, 6(1), 26808.
- Vink, A. et al. (2022). MANGAN 2021 Cruise Report: Independent scientific monitoring of two collector tests in the BGR and GSR contract areas for the exploration of polymetallic nodules in the equatorial NE Pacific. BGR Report, 363 pp. Retrieved April 24, 2023, from <https://doi.org/10.25928/hw7d-fs42>.
- Volz, J. B., Mogollón, J. M., Geibert, W., Arbizu, P. M., Koschinsky, A., & Kasten, S. (2018). Natural spatial variability of depositional conditions, biogeochemical processes and element fluxes in sediments of the eastern Clarion-Clipperton Zone, Pacific Ocean. *Deep-Sea Research Part I, Oceanographic Research Papers*, 140, 159–172.
- von Stackelberg, U. (1997). Growth history of manganese nodules and crusts of the Peru Basin. *Geological Society, London Special Publications*, 119(1), 153–176.
- Wadoux, A.M.J.-C., Heuvelink, G. B. M., de Bruin, S., & Brus, D. J. (2021). Spatial cross-validation is not the right way to evaluate map accuracy. *Ecological Modelling*, 457, 109692.
- Walbridge, S., Slocum, N., Pobuda, M., & Wright, D. (2018). Unified geomorphological analysis workflows with benthic Terrain Modeler. *Geosciences*, 8(3), 94.
- Wang, M., Wu, Z., Best, J., Yang, F., Li, X., Zhao, D., & Zhou, J. (2021). Using multibeam backscatter strength to analyze the distribution of manganese nodules: A case study of seamounts in the Western Pacific Ocean. *Applied Acoustics*, 173, 107729.
- Wasilewska-Błaszczuk, M., & Mucha, J. (2021). Application of general linear models (GLM) to assess nodule abundance based on a photographic survey (Case Study from IOM area), Pacific Ocean. *Minerals*, 11, 427.
- Weber, M. E., von Stackelberg, U., Marchig, V., Wiedicke, M., & Grube, B. (2000). Variability of surface sediments in the Peru Basin: Dependence on water depth, productivity, bottom water flow, and seafloor topography. *Marine Geology*, 163(1–4), 169–184. [https://doi.org/10.1016/s0025-3227\(99\)00103-6](https://doi.org/10.1016/s0025-3227(99)00103-6).
- Wenger, S. J., & Olden, J. D. (2012). Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, 3(2), 260–267.
- Weydert, M. (1991). Design of a system to assess manganese nodule resources acoustically. *Ultrasonics*, 29(2), 150–158.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer International Publishing.
- Wiedicke, M. H., & Weber, M. E. (1996). Small-scale variability of seafloor features in the northern Peru Basin: Results from acoustic survey methods. *Marine Geophysical Research*, 18(5), 507–526.
- Wilson, M. F. J., O’Connell, B., Brown, C., Guinan, J. C., & Grehan, A. J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30(1–2), 3–35.
- Wong, L. J., Kalyan, B., Chitre, M., & Vishnu, H. (2021). Acoustic assessment of polymetallic nodule abundance using sidescan sonar and altimeter. *IEEE Journal of Oceanic Engineering*, 46(1), 132–142.
- Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77(1), 1–17.
- Yamazaki, T., & Sharma, R. (2000). Morphological features of corrich manganese deposits and their relation to seabed slopes. *Marine Georesources & Geotechnology*, 18(1), 43–76.
- Yoo, C. M., Joo, J., Lee, S. H., Ko, Y., Chi, S.-B., Kim, H. J., Seo, I., & Hyeong, K. (2018). Resource assessment of polymetallic nodules using acoustic backscatter intensity data from the Korean exploration area, northeastern equatorial Pacific. *Ocean Science Journal OSJ*, 53(2), 381–394.
- Yu, G., & Parianos, J. (2021). Empirical application of generalized Rayleigh distribution for mineral resource estimation of seabed polymetallic nodules. *Minerals*, 11(5), 449.
- Zevenbergen, L. W., & Thorne, C. R. (1987). Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms*, 12(1), 47–56.

## Supplementary Material

### An interpretable multi-model machine learning approach for spatial mapping of deep-sea polymetallic nodule occurrences.

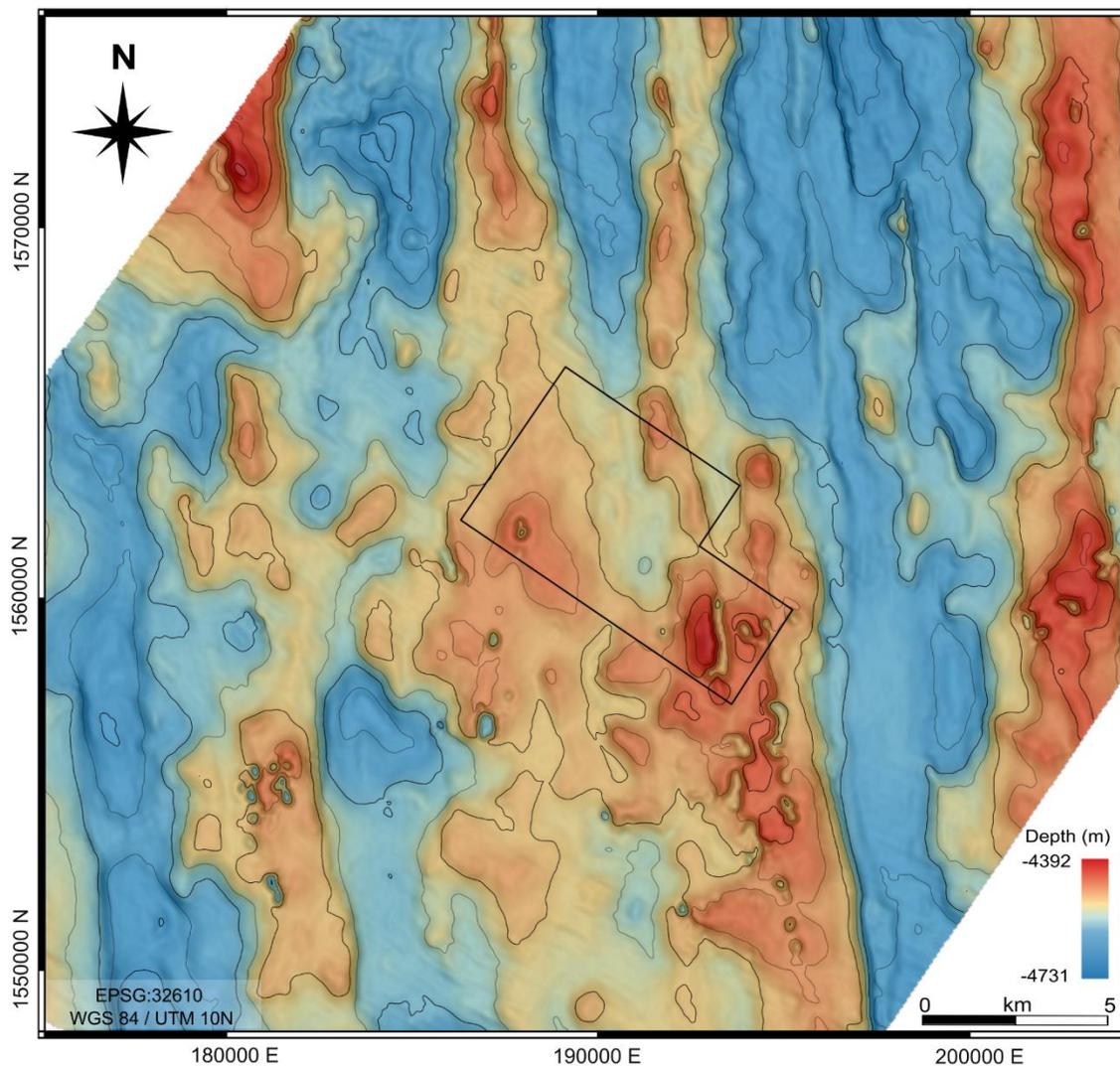
Iason-Zois Gazis<sup>1\*</sup>, Francois Charlet<sup>2</sup>, Jens Greinert<sup>1,3</sup>

<sup>1</sup> GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany

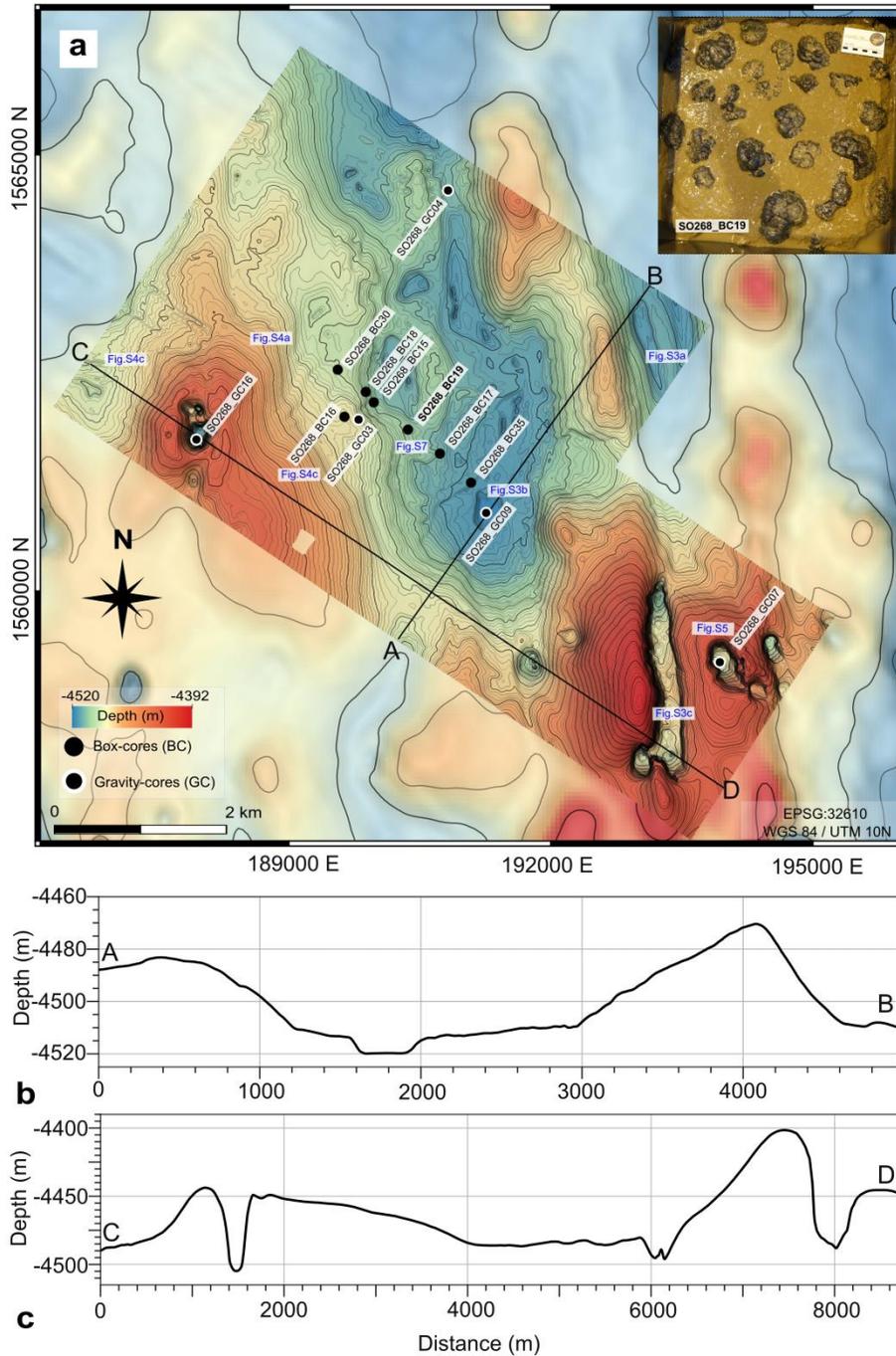
<sup>2</sup> GSR Global Sea Minerals Resources NV, DEME Group, Belgium

<sup>3</sup> Christian-Albrechts University Kiel, Germany

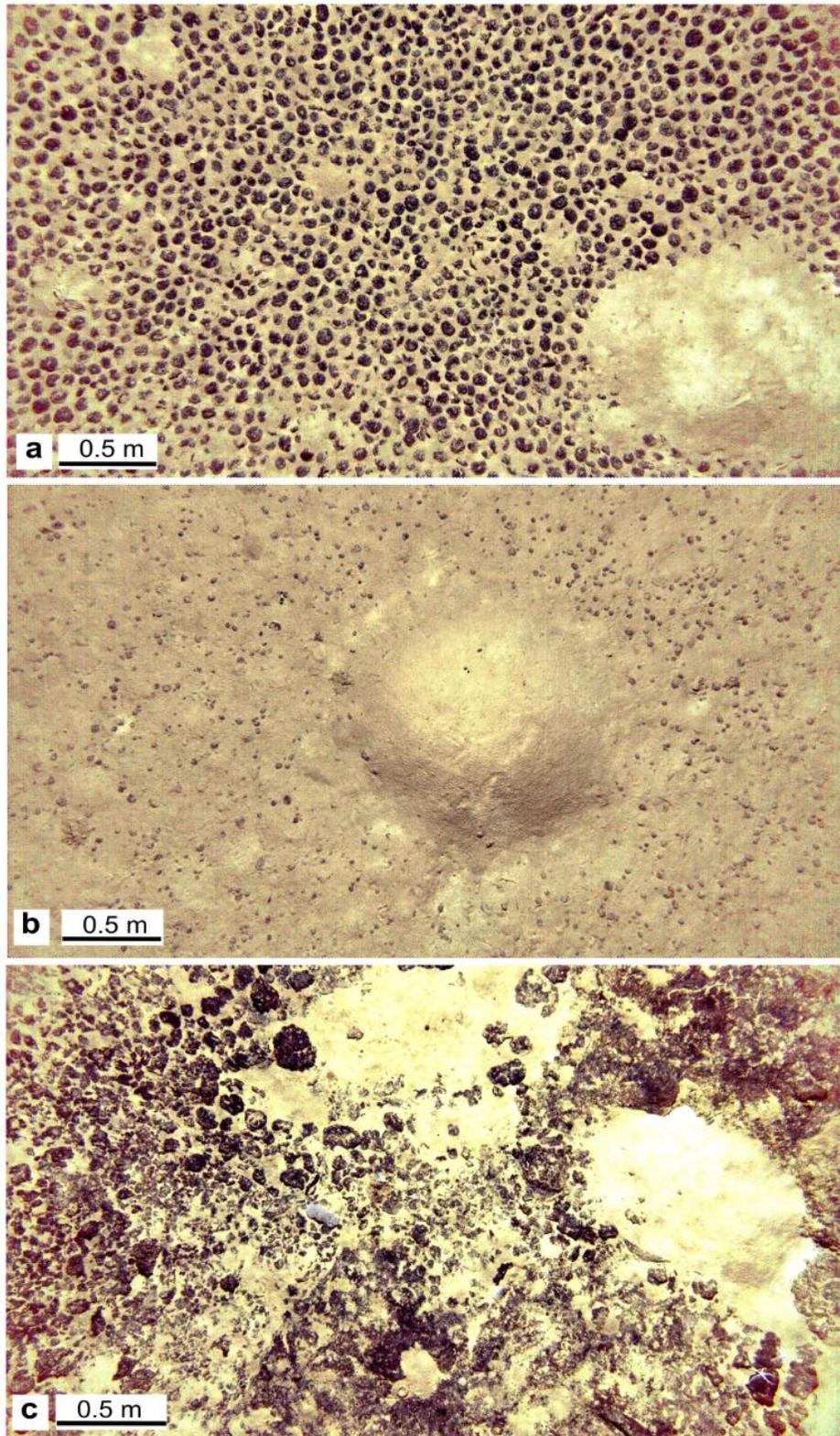
\* Corresponding Author.



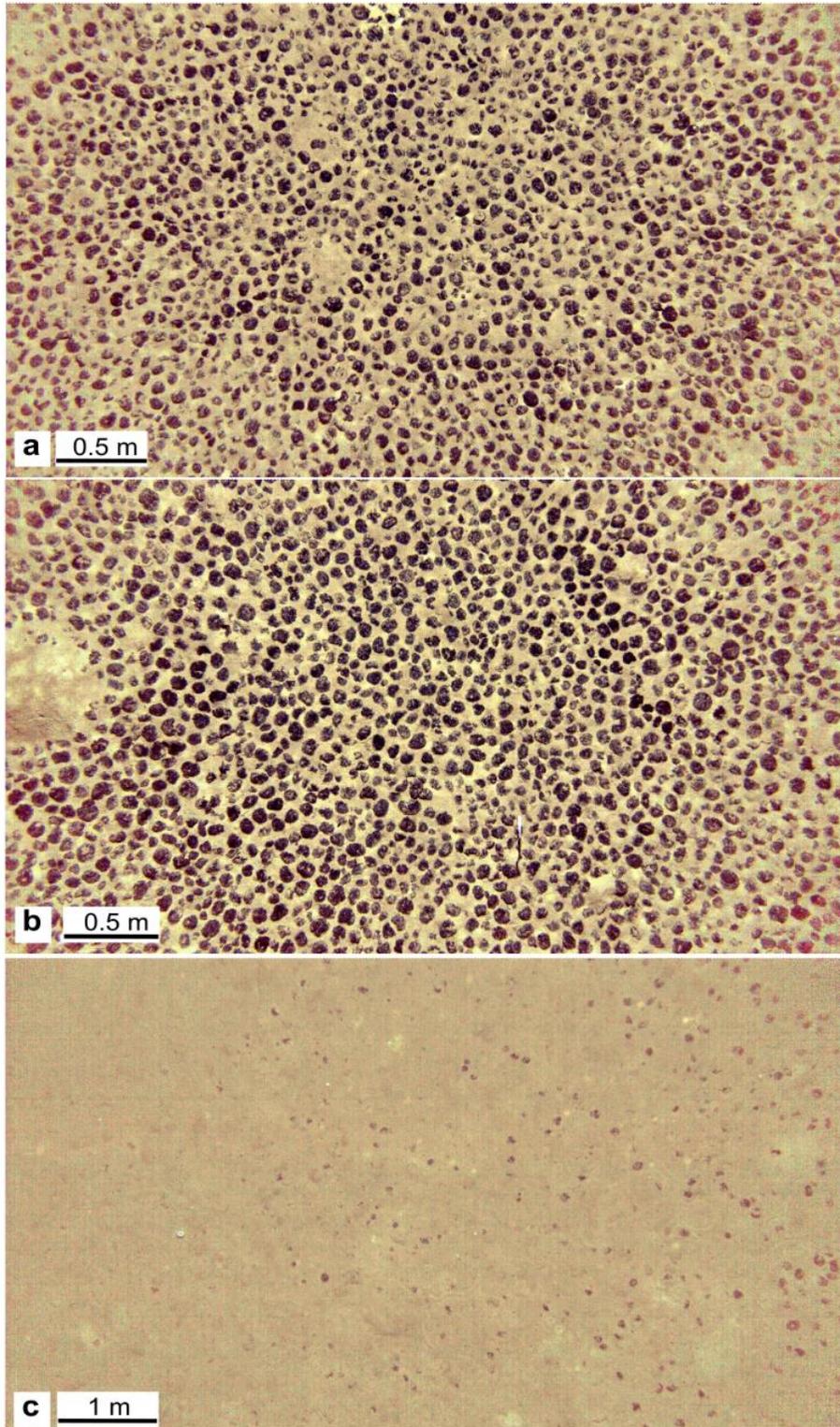
**Figure S1.** MBES bathymetry of the broader study area at 50 m x 50 m grid cell size (Gazis 2020). Parallel N-S oriented horst and graben structures dominate the seafloor. Several pits of varying depths and sizes exist along the horst structures. The contours have a 25 m interval. The black polygon indicates the study area with available AUV-based bathymetry.



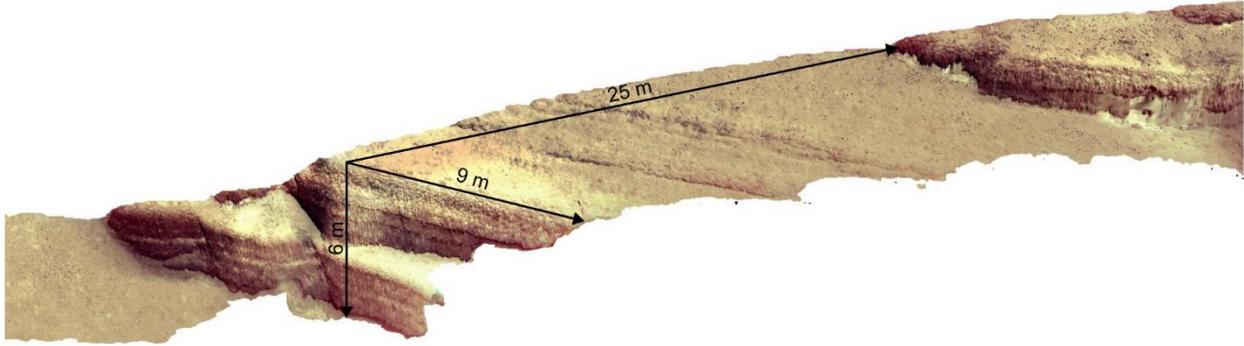
**Figure S2.** a) AUV MBES bathymetry of the study area with indicated box corer and gravity corer sampling locations. The box-corer photo from the central part shows yellowish-to-brownish clay on the seafloor surface and medium-to-large size nodules. A detailed description of the box and gravity corers exists in Haeckel and Linke (2021). The grid cell size is 2 m x 2 m; contours have 1 m intervals. The background bathymetry is from ship-based MBES data at a grid cell size of 50 m x 50 m (Gazis 2020). b) The bathymetric profile along the NE-SW axis shows the wide U-shaped valley at the centre of the study area, which is the deepest part of the area. c) Bathymetric profile along the NW-SE axis shows the depth of the crater and its asymmetrical sides, with the southeast side having gentler slopes that progressively smooth out. The shallowest part of the study area is on the summit of the horst structure.



**Figure S3.** Mounds of sediment accumulation without nodules on top (i.e., tumuli) caused by bioturbation. **a)** with high nodule coverage of large nodules, **b)** with low nodule coverage of small nodules and **c)** within hard substrate. The location of each image is given in Figure S2.



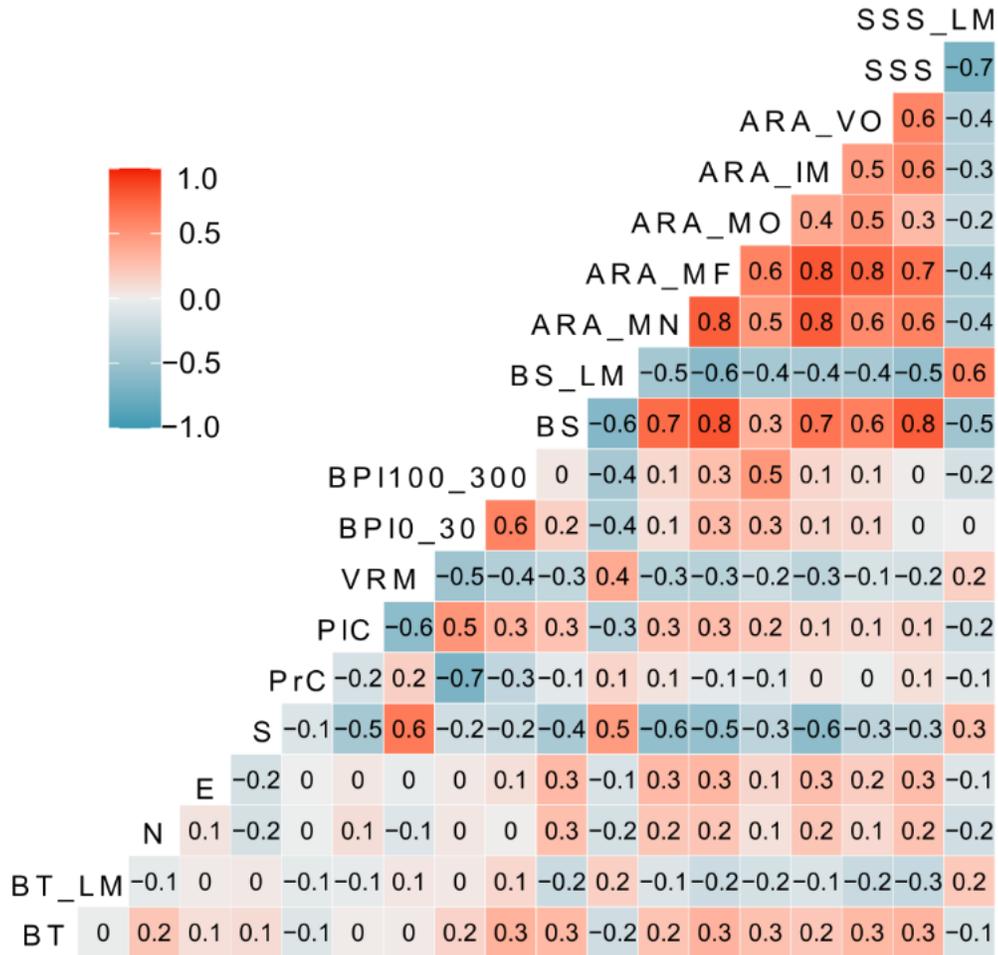
**Figure S4.** Seafloor images from the **a)** northeast, **b)** southeast, and **c)** northwest base of the slopes around the volcanic knoll (Figure S2). The northwest side has steeper slopes and incised channels, indicating sediment transport downslope. This dynamic environment is related to a lower nodule coverage. The location of each image is given in Figure S2. The AUV altitude is higher in image c. Nevertheless, the lower nodule coverage is still distinguishable.



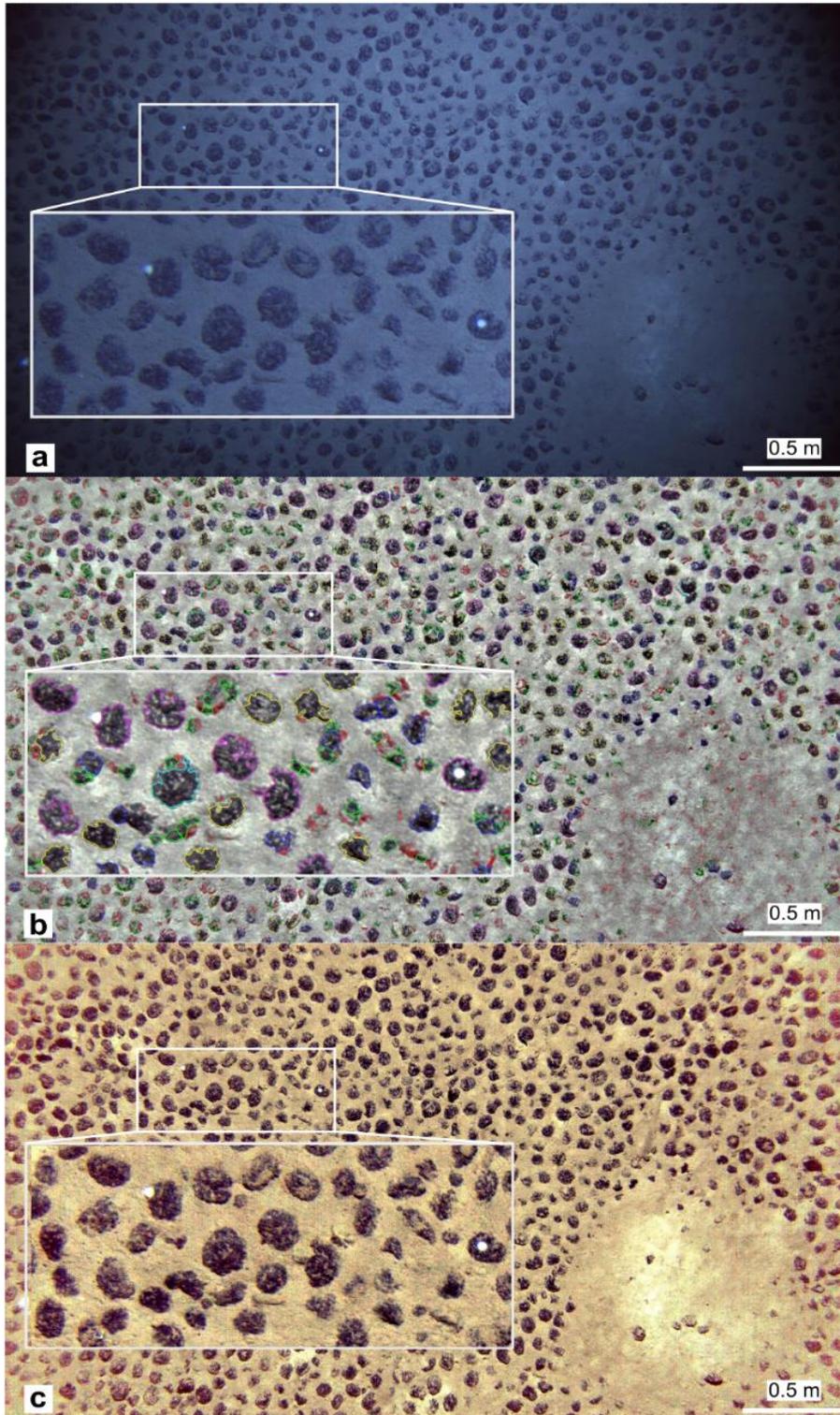
**Figure S5.** Color-normalized ortho-photomosaic of the seafloor at the outer edge of a pit at the southeast part of the study area. Steep walls and layered consolidated sediments are present. The location is given in Figure S2.

**Table S1.** Zero variance analysis of MBES and SSS derivatives using the R caret package (Kuhn 2022). The frequency ratio shows the frequency of the most prevalent value over the second most frequent value. The percent of unique values shows the number of unique values divided by the total number of samples (times 100). The recommended criteria for frequency ratio (95/5) and percent of unique values (<10) were used for the analysis (Kuhn 2022). Both criteria must apply to flag a derivative as a zero-variance predictor.

MBES & SSS derivatives	Abbreviation	Frequency ratio	Percent of unique values	zero-variance predictor
<b>MBES bathymetry</b>				
Bathymetry	BT	1.03	28.60	FALSE
Bathymetry Local Moran	BT_LM	1.25	97.96	FALSE
Slope	N	1.00	99.24	FALSE
Northness	E	1.13	95.07	FALSE
Eastness	S	1.00	97.71	FALSE
Profile Curvature	PrC	1.00	99.48	FALSE
Plan Curvature	PIC	1.50	99.43	FALSE
Vector Ruggedness Measure	VRM	1.00	98.48	FALSE
Bathymetric Position Index				
Fine Scale	BPI0_30	1.00	99.32	FALSE
Broad Scale	BPI100_300	1.00	99.24	FALSE
<b>MBES backscatter</b>				
Backscatter	BS	1.00	68.67	FALSE
Backscatter Local Moran	BS_LM	1.50	99.17	FALSE
ARA Mean Near	ARA_MN	1.36	4.95	FALSE
ARA Mean Far	ARA_MF	1.27	11.46	FALSE
ARA Mean Outer	ARA_MO	2.27	3.61	FALSE
ARA Impedance	ARA_IM	1.12	7.64	FALSE
ARA Volume	ARA_VO	22.61	6.34	FALSE
<b>SSS backscatter</b>				
SSS backscatter	SSS	1.10	58.48	FALSE
SSS backscatter Local Moran	SSS_LM	1.00	99.25	FALSE



**Figure S6.** Pearson's product-moment correlation among the used derivatives. The derivative acronyms are provided in Table S1.



**Figure S7. a)** Raw seafloor image. The acquisition altitude is 5.33 m. Epifauna indet. is shown on top of two nodules. **b)** H<sup>2</sup>SOM nodule coverage segmentation and calculation. The different colours correspond to different sizes classes. **c)** Color-normalized image using the in-house software Image Normalization (Köser et al. 2021). The location is given in Figure S2.

**Table S2.** GLM hyperparameter optimization: The *mstop* restricts the number of boosting iterations to avoid overfitting (Mayr and Hofner 2018). It is the main hyperparameter in model-boosting algorithms with a generalized linear model as the base learner.

Model	<i>mstop</i>	R <sup>2</sup> (spatial CV)
GLM	<b>50</b>	<b>0.58</b>
	100	0.56
	150	0.56

**Table S3.** GAM hyperparameter joint optimization: Three *mstop* and *dfbase* values were tuned. The *dfbase* defines the degrees of freedom and the smoothness of the additive P-splines used as a base learner in GAM (Eilers and Marx, 1996; Perperoglou et al. 2019). Larger *dfbase* values result in a more complex model with a better local fit but prone to overfitting. In model-based boosting with additive learners like here, large *dfbase* values could also lead to a high learning rate and faster convergence, resulting in the inability to converge at the global minimum of the cost function and divergence or oscillation around the optimal point, aka overshooting (Hofner et al. 2014). The model did not improve with increasing boosting iterations and degrees of freedom. We selected the simplest model considering the model complexity and computational time following Occam's razor principle (Vandekerckhove et al. 2015). The final model used two boundary knots anchored at the extremes of the data range and 20 internal equidistant knots connected with additive P-splines of 4 degrees of freedom. This model scheme is also recommended as a good overall solution (Hofner et al. 2014).

Model	<i>mstop</i>	<i>dfbase</i>	R <sup>2</sup> (spatial CV)
GAM	<b>50</b>	<b>4</b>	<b>0.60</b>
	100	4	0.60
	150	4	0.60
	50	5	0.60
	100	5	0.60
	150	5	0.60
	50	6	0.60
	100	6	0.60
	150	6	0.60

**Table S4.** RFM hyperparameter optimization: The *mtry* defines the number of predictor variables randomly sampled as candidates at each node split (Liaw and Wiener, 2002). The random sampling was done without replacement to reduce the bias in variable important ranking and correlation among trees (Strobl et al. 2007). The *mtry* hyperparameter is one of the most critical factors regarding predictive power, computational time, model bias, and variable importance (Probst et al. 2019a). The higher the *mtry* value, the higher the correlation between trees (bias) and computational time. Large *mtry* values could mask the weaker predictors as strong and weak predictors will fall more times into the same sample, with the stronger preferred for the split (Probst et al. 2019b). The optimum value here is equal to the default value for regression [rounded down (predictor variables/3)]; Liaw and Wiener 2002).

Model	<i>mtry</i>	R <sup>2</sup> (spatial CV)
RFM	1	0.59
	3	0.61
	<b>6</b>	<b>0.62</b>
	9	0.61
	12	0.61
	15	0.60

**Table S5.** RFM hyperparameter optimization: The *min.node.size* defines the minimal node size to split, controlling the tree depth. Smaller values result in deeper trees (i.e. more splits) with a better local fit but are prone to overfitting (Segal 2003; Probst et al. 2019b). Computational time increases exponentially with smaller values (Probst et al. 2019b). Five *min.node.size* values were tested. The value in brackets corresponds to the percentage of observations kept from the total training sample. In all runs, the number of trees was kept constant at 500 (default) as several studies

have shown good results and relatively low computational time (e.g. Cutler et al. 2007; Li et al. 2016; Li et al. 2017; Diesing 2020). The *variance reduction* (Breiman 2001) and *extra trees* (Guerts et al. 2006) methods were compared as a split criterion. The *variance reduction* (default for regression) was preferred as it had a better predictive performance ( $R_{\text{variance}}^2=0.62$  and  $R_{\text{extratrees}}^2=0.60$ ).

Model	<i>mtry</i>	<i>min.node.size</i>	R <sup>2</sup> (spatial CV)
RFM	6	5 (0.02%)	0.60
	6	50 (0.2%)	0.61
	6	250 (1.0%)	0.61
	<b>6</b>	<b>1250 (4.9%)</b>	<b>0.62</b>
	6	2500 (9.95%)	0.61

**Table S6.** SVM hyperparameter joint optimization. Three kernel types and C values were tested. The kernel type defines the hyperplane shape in the higher-dimensional space. The C value controls the width of the margins (distance between the hyperplane and support vectors). Larger C values penalize the error more, resulting in lower margins but prone to overfitting (Karatzoglou et al. 2004).

Model	<i>kernel</i>	<i>C</i>	R <sup>2</sup> (spatial CV)
SVM	linear	0.25	0.51
	<b>polynomial</b>	<b>0.25</b>	<b>0.59</b>
	radial basis function	0.25	0.50
	linear	0.50	0.51
	polynomial	0.50	0.58
	radial basis function	0.50	0.50
	linear	1.00	0.51
	polynomial	1.00	0.56
	radial basis function	1.00	0.50

**Table S7.** SVM hyperparameter joint optimization of a polynomial kernel. The three different values for *degree*, *scale* and *C* were jointly tuned. The *degree* defines the polynomial degree, and the *scale* defines the polynomial scaling factor (Karatzoglou et al. 2004). A 3-degree polynomial gave the best fit. SVM had the biggest improvement during the hyperparameter optimization ( $R_{\text{min}}^2=0.16$  and  $R_{\text{max}}^2=0.59$ ) compared to all other models, a finding which agrees with previous comparative studies (Probst et al. 2019a).

Model	<i>degree</i>	<i>scale</i>	<i>C</i>	R <sup>2</sup> (spatial CV)
SVM	1	0.001	0.25	0.56
	1	0.001	0.50	0.55
	1	0.001	1.00	0.54
	1	0.010	0.25	0.53
	1	0.010	0.50	0.52
	1	0.010	1.00	0.52
	1	0.100	0.25	0.51
	1	0.100	0.50	0.51
	1	0.100	1.00	0.51
	2	0.001	0.25	0.57
	2	0.001	0.50	0.58
	2	0.001	1.00	0.58
	2	0.010	0.25	0.50
	2	0.010	0.50	0.47
	2	0.010	1.00	0.45
	2	0.100	0.25	0.40
	2	0.100	0.50	0.39
	2	0.100	1.00	0.37
	<b>3</b>	<b>0.001</b>	<b>0.25</b>	<b>0.59</b>

	3	0.001	0.50	0.58
	3	0.001	1.00	0.56
	3	0.010	0.25	0.44
	3	0.010	0.50	0.42
	3	0.010	1.00	0.43
	3	0.100	0.25	0.16
	3	0.100	0.50	0.16
	3	0.100	1.00	0.16

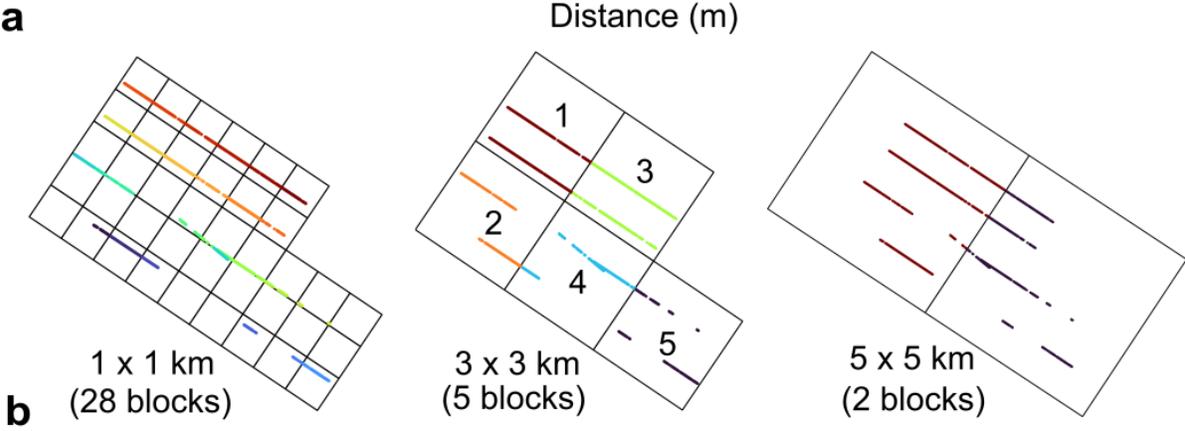
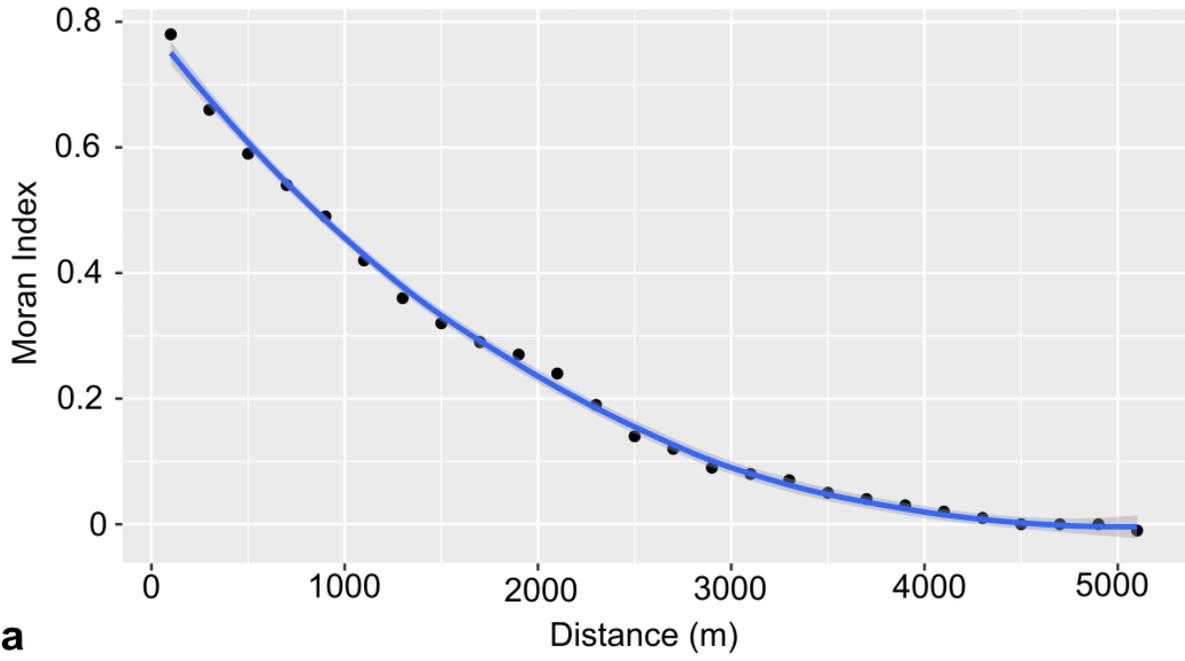
**Table S8.** NNM hyperparameter joint optimization of the *number of neurons* and type of *activation function* at the first hidden layer. The neuron (or neurons) receives the sum of weighted values from the input layer and a bias (a constant value that shifts the activation function across the hyperplane for optimum fit and guarantees activation even if nodes with zero inputs exist) and passes it to the activation function, which is differentiable (e.g., logistic, tanh, relu) and introduces non-linearity into the model, allowing it to learn complex relationships.

Model	hidden layer 1 <i>number of neurons</i>	<i>Activation Function</i>	R <sup>2</sup> (spatial CV)
NNM	<b>1</b>	<b>logisitic</b>	<b>0.57</b>
	3	logisitic	0.49
	5	logisitic	0.44
	1	tanh	0.54
	3	tanh	0.48
	5	tanh	0.49
	1	relu	0.54
	3	relu	0.38
	5	relu	0.48

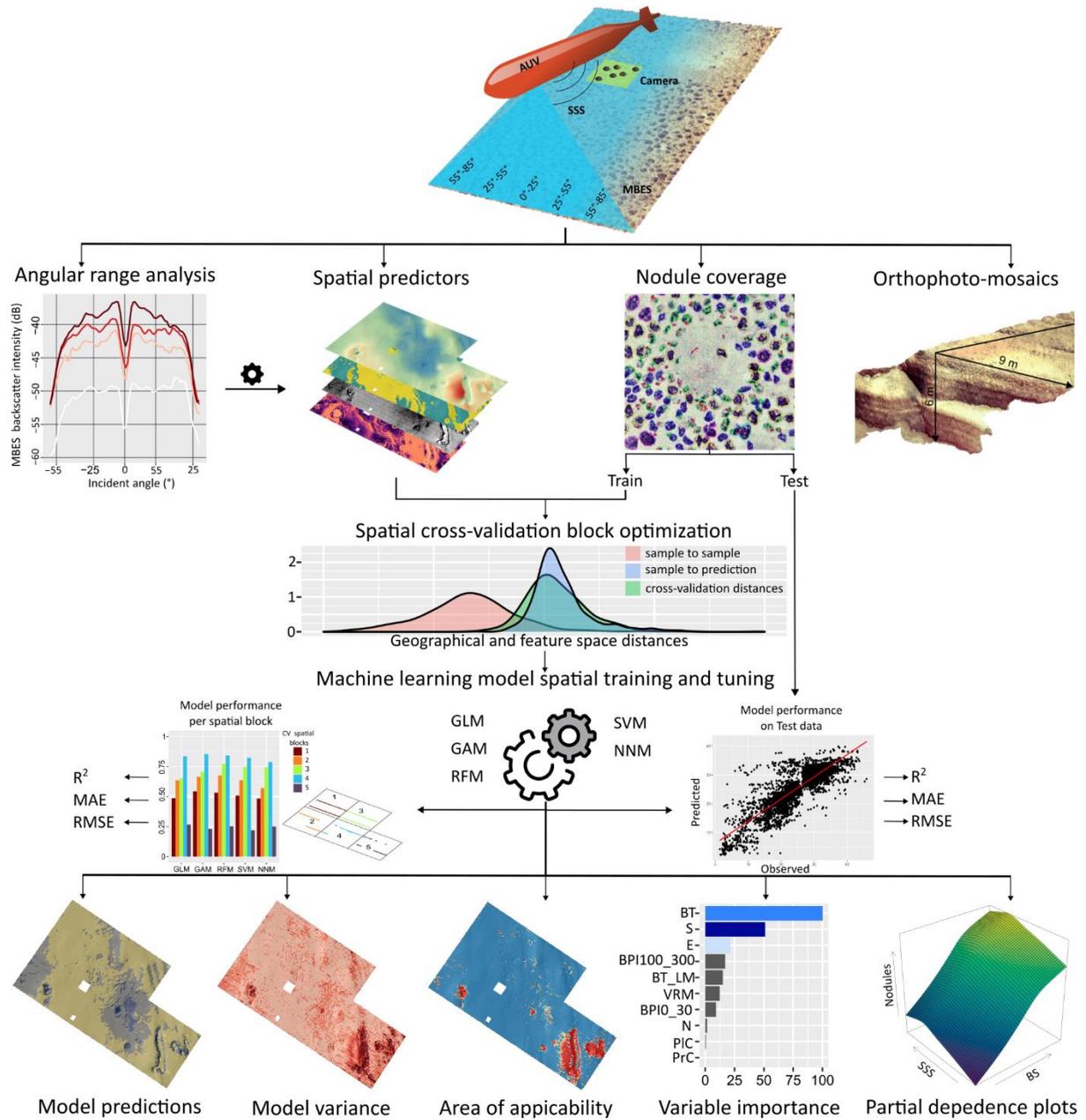
**Table S9.** NNM hyperparameter joint optimization of the number of *neurons* at the *first, second and third hidden layer*. The higher the number of hidden layers and neurons per hidden layer, the more complex the neural network, but prone to overfitting. The hyperparameters *stepmax* (maximum steps to stop the training process), *threshold* (the threshold for the partial derivatives of the error function as stopping criterion) and *activation function* (see Table S8) were kept constant at  $1e^{+05}$ , 0.5, and logistic, respectively. In this study, the increased number of neurons and hidden layers did not further improve the model performance. Model training and evaluation using spatial CV requires a good generalization ability, typically offered by simpler models. Between two NNM with equal performance, we selected the simplest model considering the model complexity and computational time following Occam's razor principle for model interpretability (Vandekerckhove et al. 2015). The NNM with optimum hyperparameter values were run 30 times, showing stable performance.

Model	hidden layer 1 <i>number of neurons</i>	hidden layer 2 <i>number of neurons</i>	hidden layer 3 <i>number of neurons</i>	R <sup>2</sup> (spatial CV)
	<b>1</b>	<b>0</b>	<b>0</b>	<b>0.57</b>
	3	0	0	0.49
	5	0	0	0.44
	1	1	0	0.57
	1	3	0	0.57
	1	5	0	0.55
	3	1	0	0.57
	3	3	0	0.50
	3	5	0	0.45
	5	1	0	0.44
	5	3	0	0.40
	5	5	0	0.44
	1	1	1	0.57

NNM	1	1	3	0.55
	1	1	5	0.57
	1	3	1	0.55
	1	3	3	0.54
	1	3	5	0.54
	1	5	1	0.54
	1	5	3	0.56
	1	5	5	0.56
	3	1	1	0.46
	3	1	3	0.55
	3	1	5	0.50
	3	3	1	0.52
	3	3	3	0.51
	3	3	5	0.50
	3	5	1	0.55
	3	5	3	0.53
	3	5	5	0.44
	5	1	1	0.49
	5	1	3	0.56
	5	1	5	0.47
	5	3	1	0.49
	5	3	3	0.43
	5	3	5	0.43
	5	5	1	0.38
	5	5	3	0.48
	5	5	5	0.44



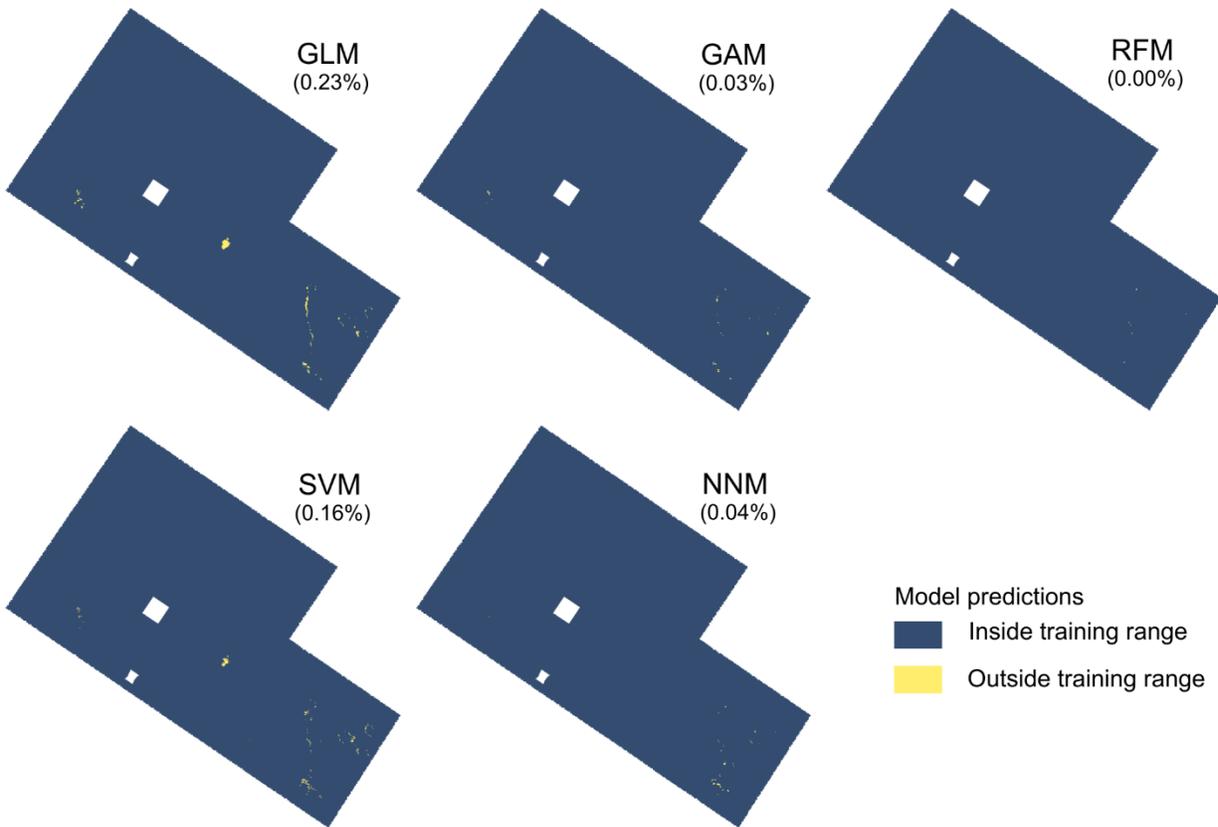
**Figure S8.** **a)** Moran Index values at distances from 100 to 5100 m, with a step of 200m. A LOESS smoother with a 0.95 confidence interval has been plotted through the index values using the R ggplot2 package (Wickham 2016). **b)** Spatial block configuration for different block dimensions. The spatial blocks where data exist correspond to the spatial CV folds used during model training and evaluation.



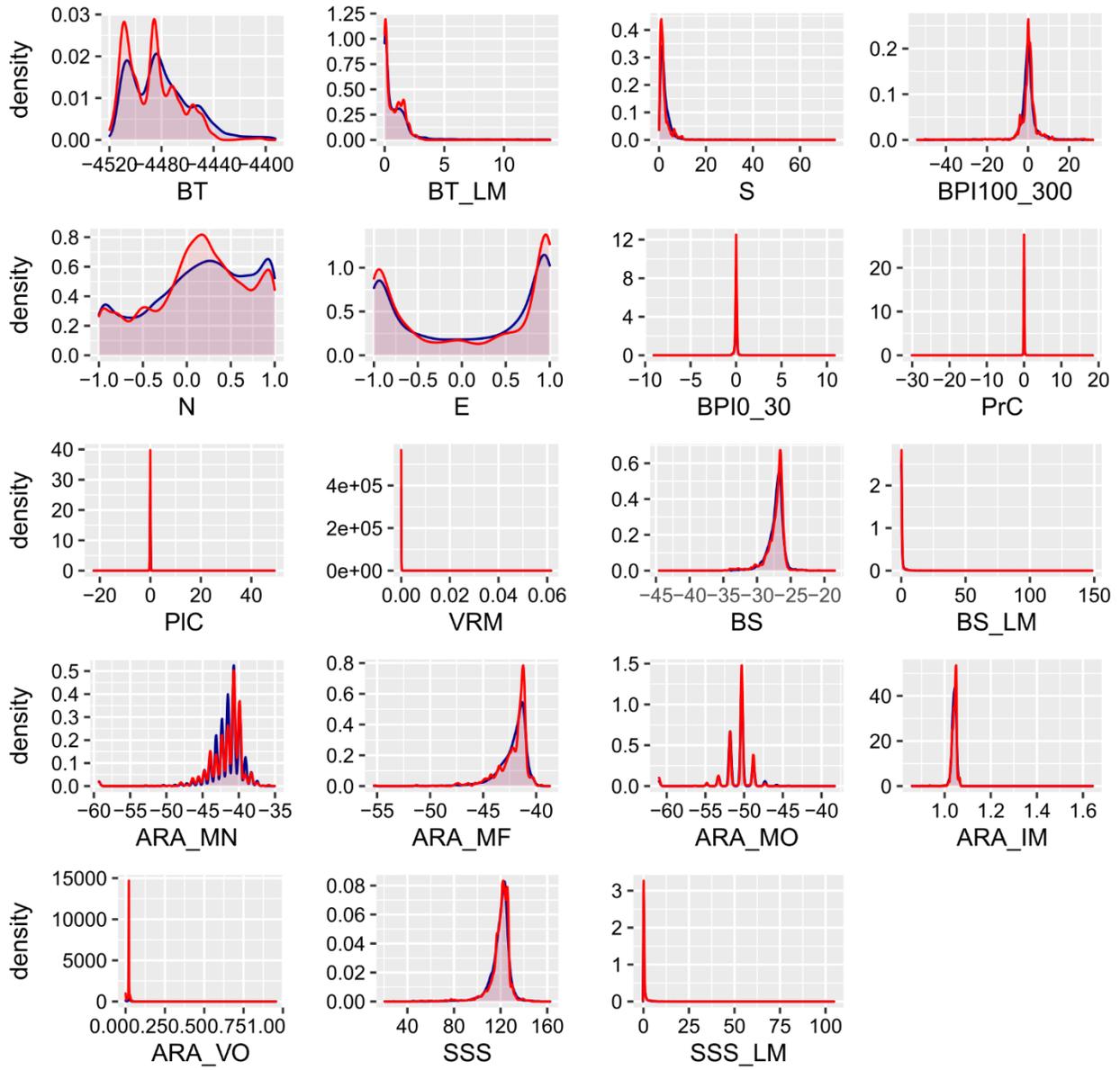
**Figure S8.** Schematic flowchart of this study's main methodological and modelling steps: From data acquisition to model outputs and interpretation.

**Table S10.** Summary statistics of train dataset and model predictions within the entire study area.

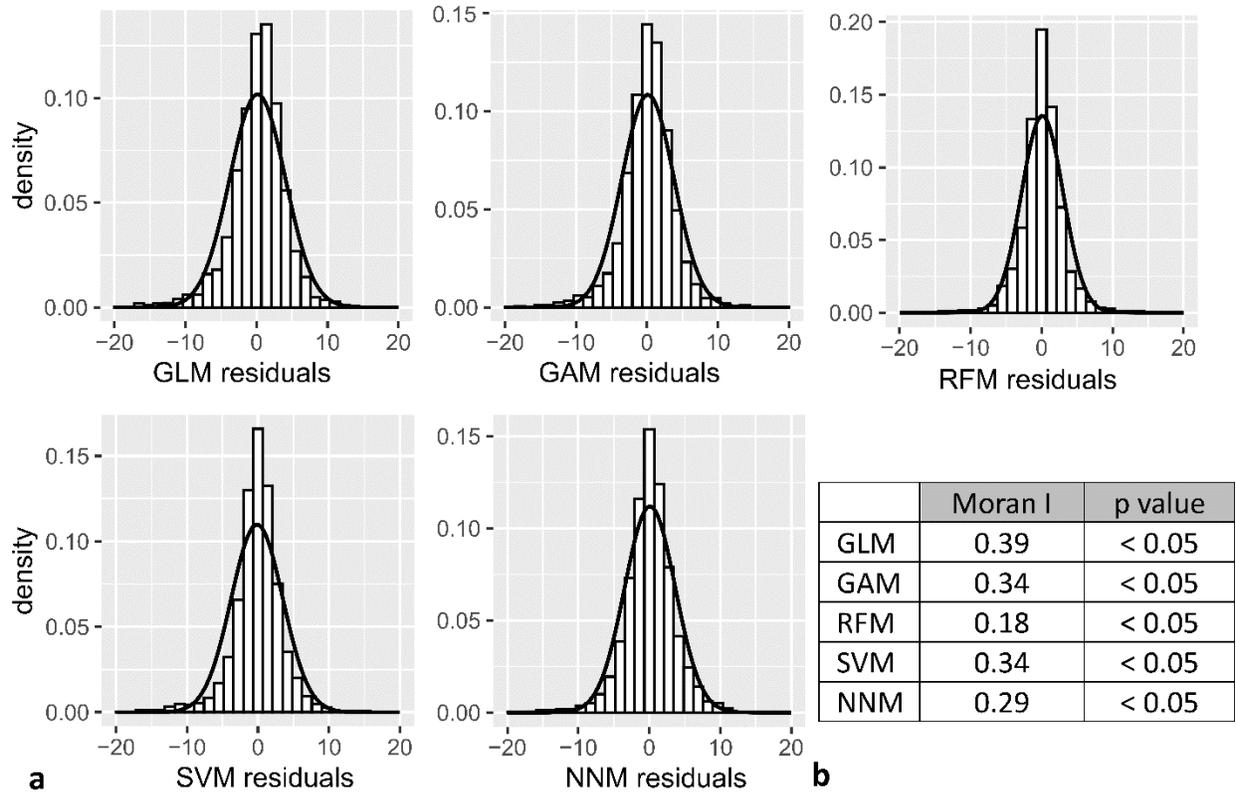
	Min.	1st Qu.	Median	Mean	3 <sup>rd</sup> Qu.	Max.
<b>Train data</b>	<b>1.5</b>	<b>22</b>	<b>29</b>	<b>26</b>	<b>31</b>	<b>49</b>
GLM	0	24	28	26	30	69
GAM	0	23	28	26	30	42
RFM	7	22	27	26	30	35
SVM	0	23	28	26	30	100
NNM	5	23	27	26	30	63



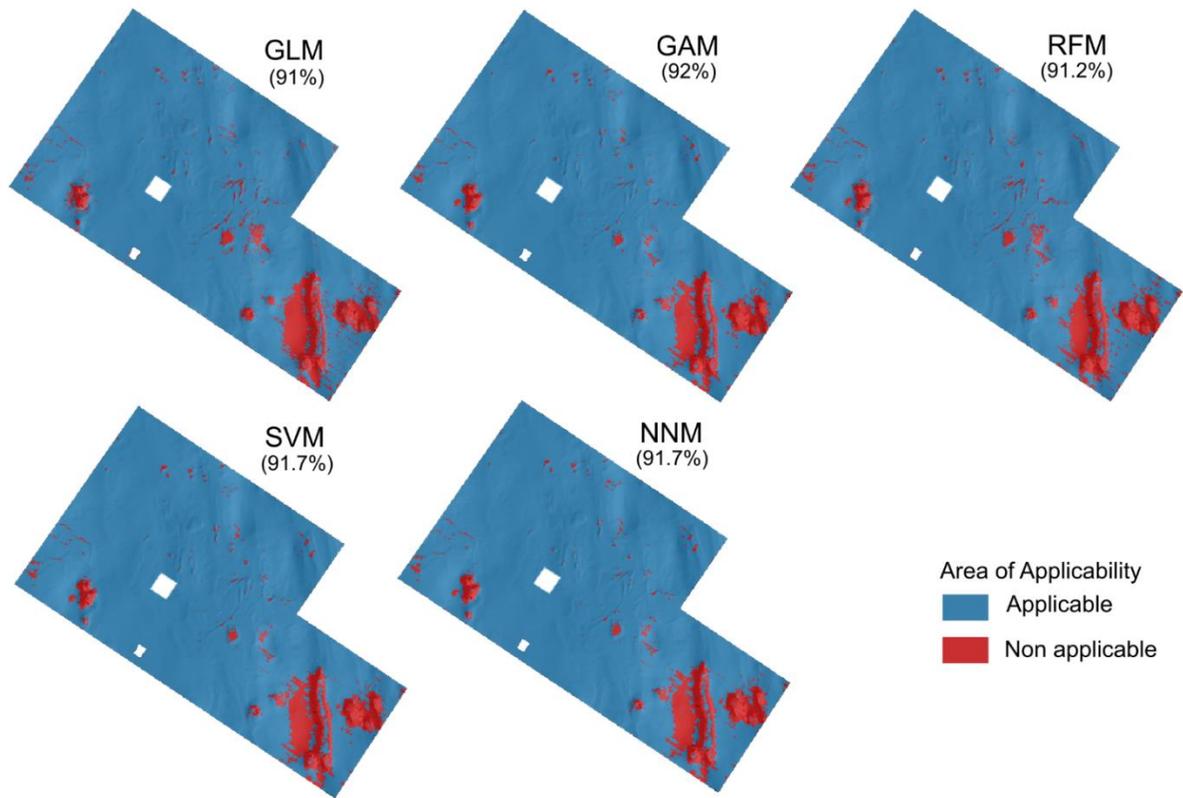
**Figure S10.** Maps of the predicted nodule coverage value range concerning the (min-max) range of training data.



**Figure S11.** Density distribution plots of predictors from the study area (blue) and the training data set (red).



**Figure S12. a)** Histograms and superimposed normal density distribution for the residuals of each model. **b)** Moran Index (Moran I) values for the residuals of each model.



**Figure S13.** Maps of the AOA spatial distribution for each model. The AOA is provided in brackets as a percentage of the total area.

**Table S11.** R<sup>2</sup> for each CV spatial block of each model.

CV spatial block	GLM	GAM	RFM	SVM	NNM
1	0.49	0.54	0.53	0.51	0.48
2	0.64	0.67	0.68	0.64	0.57
3	0.65	0.71	0.77	0.75	0.74
4	0.84	0.85	0.84	0.82	0.79
5	0.27	0.23	0.25	0.22	0.25
mean	0.58	0.60	0.62	0.59	0.57
median	0.64	0.67	0.68	0.64	0.57

**Table S12.** Number of samples for each spatial block that are kept out for model validation, and the remaining number of samples from the other spatial blocks used for training.

CV spatial block	# of samples in each validation spatial block	# of samples in (remaining) training spatial blocks
1	7664	17456
2	3930	21190
3	6495	18625
4	4378	20742
5	2653	22467

**Table S13.** Pearson's product-moment correlation between the number of training samples used in each model iteration and R<sup>2</sup> in each test spatial block. The unequal number of samples kept out in each iteration did not influenced the model performance.

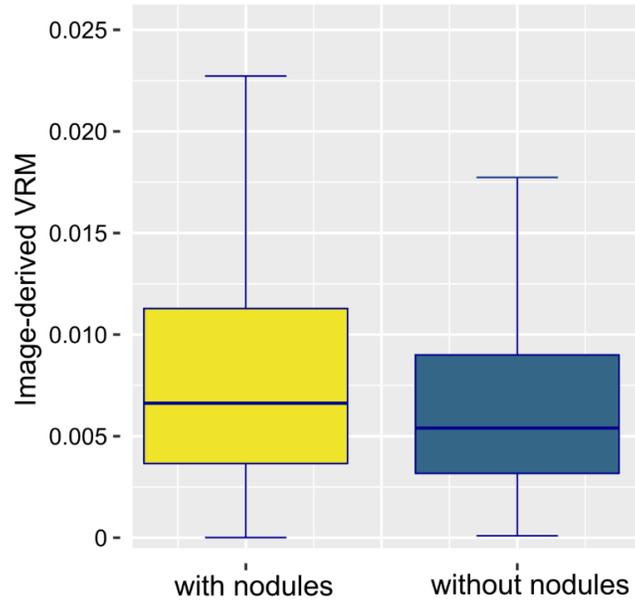
	GLM	GAM	RFM	SVM	NNM
correlation (r)	-0.22	-0.34	-0.35	-0.36	-0.37
p-value	0.72	0.57	0.57	0.55	0.54

**Table S14.** Comparison between local and global R<sup>2</sup> in the CV held-out data.

Model	Local spatial R <sup>2</sup>	Global spatial R <sup>2</sup>
GLM	0.58	0.63
GAM	0.60	0.67
RFM	0.62	0.69
SVM	0.59	0.66
NNM	0.57	0.63

**Table S15.** Adjusted R<sup>2</sup> for each model based on the different subsets of predictors

	Geomorphological predictors only (A)	Geomorphological Predictors + SSS (B)	Geomorphological Predictors, SSS + MBES backscatter (C)	Difference (C-A)
GLM	0.43	0.61	0.72	0.29
GAM	0.59	0.65	0.75	0.16
RFM	0.74	0.78	0.84	0.10
SVM	0.50	0.65	0.75	0.25
NNM	0.50	0.65	0.76	0.26



**Figure S14.** Boxplots of image-derived VRM surfaces without outliers. The Wilcoxon rank sum test showed that there is a significant difference between the two datasets ( $p < 0.05$ ).

# Monitoring benthic plumes, sediment redeposition and seafloor imprints caused by deep-sea polymetallic nodule mining

Received: 28 July 2023

Accepted: 15 January 2025

Published online: 31 January 2025

 Check for updates

Iason-Zois Gazis<sup>1</sup>✉, Henko de Stigter<sup>2</sup>, Jochen Mohrmann<sup>1</sup>, Karl Heger<sup>1</sup>, Melanie Diaz<sup>2</sup>, Benjamin Gillard<sup>3</sup>, Matthias Baeye<sup>4</sup>, Mario E. Veloso-Alarcón<sup>1</sup>, Kaveh Purkiani<sup>1,5,6</sup>, Matthias Haeckel<sup>1</sup>, Annemiek Vink<sup>7</sup>, Laurenz Thomsen<sup>3,8</sup> & Jens Greinert<sup>1,9</sup>

A deep-sea (4500 m) trial of a pre-prototype polymetallic nodule collector with independent scientific monitoring revealed that a gravity current formed behind the collector channeled through steeper seafloor sections and traveled 500 m downslope. The prevailing bottom currents dominated sediment dispersion up to the end of the monitoring area at 4.5 km distance. The maximum suspended particle concentration recorded 50 m from mining lanes was up to four orders of magnitude higher than ambient values but decreased rapidly with increasing time, distance, and altitude. Most of the plume remained close to the seafloor, with the highest concentrations at 1 m monitoring altitude and reaching background concentrations at 50 m. Rapid particle flocculation was followed by fast and substantial sediment redeposition. A mm-scale photogrammetric seafloor reconstruction allowed quantitative estimates of the thickness of redeposited sediment next to mining lanes of  $\approx 3$  cm and a minimum erosional depth of 5 cm.

Almost 150 years after the first deep-sea polymetallic nodules (hereafter nodules) were recovered during the HMS Challenger expedition (1872–1876)<sup>1</sup>, the industrial deep-sea trial of a tracked pre-prototype seafloor nodule collector vehicle (hereafter nodule collector) in the Clarion-Clipperton Zone (CCZ; eastern equatorial Pacific)<sup>2</sup> offered the opportunity to quantify the mobilisation, dispersal and redeposition of the benthic sediment plume (hereafter plume). Nodules are mineral concretions of millimetre to decimetre size formed over millions of years by precipitation of iron and manganese oxyhydroxides from seawater and sediment porewater around a nucleus<sup>3</sup>. They are embedded on the surface of ocean floor siliceous clays, typically between 3500 m and 6500 m water depth<sup>3,4</sup>. Due to their abundance

and appreciable contents of copper, nickel, cobalt, and a range of rare earth elements, nodules have received commercial interest since the 1960s<sup>4</sup>. The surge of green, digital, and decarbonising technologies has increased this interest in the last two decades<sup>4</sup>. Mandated through the United Nations Convention on the Law of the Sea<sup>5</sup>, the International Seabed Authority (ISA) is currently draughting the parts of the Mining Code<sup>6</sup> related to exploiting these resources. However, our knowledge of the monitoring technology needed, plume dispersion, sediment redeposition and likely impacts of deep-sea mining activities is still incomplete due to the recent birth of this industry and the complexity of acquiring data on the resilience of deep-sea life to such disturbances<sup>7</sup>. Valuable insights have been obtained from mining trials,

<sup>1</sup>GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany. <sup>2</sup>NIOZ Royal Netherlands Institute for Sea Research, Texel, The Netherlands. <sup>3</sup>Constructor University, Bremen, Germany. <sup>4</sup>Royal Belgian Institute of Natural Sciences, Brussels, Belgium. <sup>5</sup>MARUM Center for Marine Environmental Sciences and Faculty of Geosciences, University of Bremen, Bremen, Germany. <sup>6</sup>Now at Federal Maritime and Hydrographic Agency, Hamburg, Germany. <sup>7</sup>Federal Institute for Geosciences and Natural Resources, Hannover, Germany. <sup>8</sup>Now at the University of Gothenburg, Gothenburg, Sweden. <sup>9</sup>Institute of Geosciences, Christian-Albrecht University of Kiel, Kiel, Germany. ✉e-mail: [igazis@geomar.de](mailto:igazis@geomar.de)

and benthic impact experiments carried out since the 1970s, although those disturbances were of a different nature, spatial scale and intensity than those expected from industrial mining technologies<sup>8–16</sup>.

Herein, the most detailed, up-to-date in situ view of deep-sea plume monitoring using hydroacoustic and optic sensors is presented. The Patania II nodule collector of the Belgium-sponsored ISA contractor Global Sea Mineral Resources NV (GSR) is smaller in size and weight than an industrial-scale prototype collector and was tested without a riser system<sup>2,17</sup>. However, it does integrate caterpillar tracks for locomotion and hydraulic suction for nodule collection, a technological scheme widely considered feasible for future mining activities<sup>18</sup>. The 4 m wide hydraulic nodule collector head uses water jets to lift the nodules (Coandă Effect)<sup>17</sup>, thereby also mobilising the topmost bioactive layer of the seafloor with its epi- and endofauna as well as its microbial communities<sup>19</sup>. The resulting mixture of nodules, sediment and water passes through a separator, which diverts the nodules into the collector bucket, while the sediment-laden water passes through the diffuser located at the rear side and is discharged as a plume at 2.5–3.2 m above the seafloor<sup>17,20</sup>. Together with sediment stirred up by the locomotion on the seafloor, the initial input to the plume during the trial was estimated to be  $12 \pm 3 \text{ kg s}^{-1}$  of sediment<sup>20</sup>.

Independent scientific monitoring of the plume dispersion and benthic ecosystem disturbance was conducted in the framework of the third-party-funded MiningImpact 2 project<sup>21</sup> of the European Joint Programming Initiative entitled Healthy and Productive Seas and Oceans<sup>22</sup> using the MV Island Pride<sup>2</sup>. More than ten years of oceanographic observations, detailed analyses of the seabed sediment characteristics and spatial distribution of nodules, insights into the aggregation potential of suspended sediment from laboratory and small-scale deep-sea disturbance experiments, as well as numerical models of plume behaviour, provided the basis for the seafloor and plume monitoring strategy applied during the mining trial<sup>12,13,16,19,23–29</sup>. Many of the monitoring platforms and sensor settings had already been tested during a dredge disturbance experiment in the CCZ<sup>12</sup>. In total, 44 oceanographic sensors were mounted on 23 stationary seafloor platforms<sup>2</sup> distributed at various distances (50–1800 m) from the collector impact site (hereafter impact site), including a PartiCam<sup>29</sup> sensor for in situ particle size monitoring. Two Remotely Operated Vehicles (ROVs) were used to visually observe the plume and deploy/recover the multi-sensor platforms<sup>2</sup>. An Autonomous Underwater Vehicle (AUV) acquired multibeam echosounder (MBES), side-scan sonar (SSS), and seafloor image data, as well as suspended particulate matter (SPM) concentration (hereafter concentration) at different altitudes (5, 10, 30 and 50 m) and distances (up to 5 km) from the impact site<sup>2</sup>. The optical backscatter sensors (OBS) for measuring concentration ( $\text{mg L}^{-1}$ ) were calibrated to reference suspensions prepared from CCZ surface sediments (see Methods). All OBS and acoustic doppler current profilers (ADCPs) measured ambient concentration and background acoustic backscatter intensity before the trial.

## Results

### Mining imprints on the seafloor

On 19 April 2021, the nodule collector was lowered to the seafloor and was remotely operated for 41.33 h, covering a seafloor distance of 21.37 km<sup>2</sup>. It picked up nodules along 171 lanes, each 50 m long and 4 m wide, resulting in a total mined area of 0.034 km<sup>2</sup>. The displaced nodules ( $\approx 660 \text{ t}$ ) were deposited on the seafloor as piles at the end of each driven lane (“light bulb” turn manoeuvre; Figs. 1–2 and Supplementary Fig. S1). The lanes were distributed over three strips (first strip: 55 lanes, second strip: 31 lanes, third strip: 85 lanes; Fig. 1). The first and third strips were driven from SW to NE, whilst the second strip was driven from NE to SW. The two northern strips were located on a gentle NE-sloping seafloor ( $0^\circ$ – $3^\circ$ ), whereas the southern strip ended towards the NE in a steeper ( $3^\circ$ – $6.5^\circ$ ) slump scar-like structure sloping

down into a wider N-S oriented enclosed basin (Fig. 3). The SSS mosaic shows that  $\approx 0.08 \text{ km}^2$  of the seafloor was directly disturbed by the nodule collector, including the turns (Fig. 1). The mined lanes are cleared from nodules and eroded by the water jets. Based on the Digital Elevation Model (DEM) derived from seafloor images acquired 88–110 h after the trial within the impact site, an erosional depth of  $\approx 5 \text{ cm}$  was measured. However, this depth is measured after substantial sediment redeposition occurred (see Sediment redeposition). The caterpillar tracks created an additional small-scale topography due to sideways sediment extrusion (Fig. 2 and Supplementary Fig. S2).

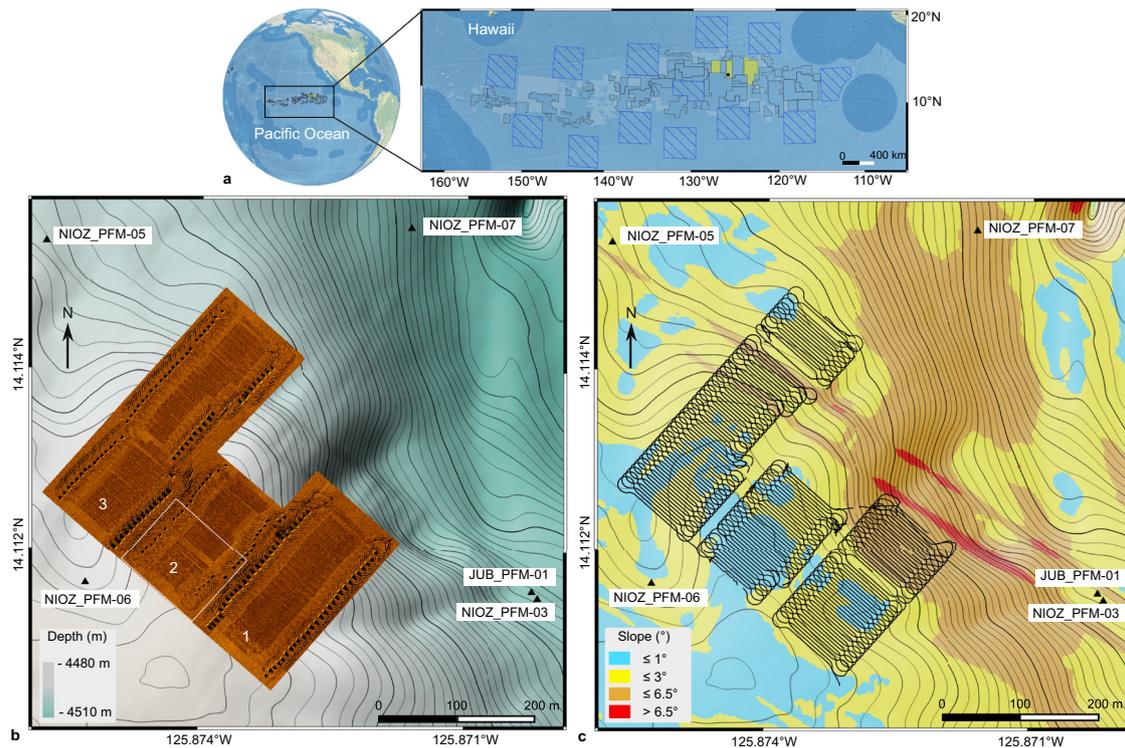
### Spatial distribution of the plume

At 50 m distance from the edge of the impact site, the maximum concentration recorded was  $264 \text{ mg L}^{-1}$  at 1 m altitude (NIOZ\_PFM-06; Fig. 3). This was four orders of magnitude above the measured background level of  $0.02$ – $0.03 \text{ mg L}^{-1}$ , but returned to ambient values after 14 h (Supplementary Figs. S2, S3). At the flattest parts within the impact site (Fig. 1), the ROV recorded a low-lying plume with a sharp front distinct from the ambient water and headed perpendicular to mining lanes. Although a gentle oscillation was observed, the net movement was downslope in the form of a gravity current (Fig. 3).

NE of the impact site, at a distance of  $\approx 200 \text{ m}$  downslope, a sharp increase in the concentration at 1 m altitude was recorded at NIOZ\_PFM-07, which lasted for 2.3 h before returning to background values (Supplementary Fig. S2). The plume arrival was marked by a distinct increase in acoustic backscatter intensity, current speed and current direction change from S to E based on ADCPs at NIOZ\_PFM-07 and GMR\_PFM-39 (Fig. 3 and Supplementary Figs. S4, S5). A progressive vector plot at 1.2 m altitude revealed a 380 m eastward displacement during the passage of the gravity current (Fig. 3). The gravity current created sediment ripples with an NE–E current direction (Figs. 3, 4). The presence of sediment ripples could imply the mobilisation of additional surface sediments due to erosion<sup>30</sup>. This would, however, demand current velocities  $> 30 \text{ cm s}^{-1}$  as determined during experiments on sediment resuspension in a large seawater flume using original sediments from the CCZ or as determined for deep-sea sediments of similar particle composition<sup>29,31,32</sup>.

We consider a more likely explanation is that the massive fallout of aggregated sediments within the gravity flow<sup>33</sup> (see Particle flocculation) combined with the unidirectional currents in steeper seafloor parts resulted in mud ripples generated by large non-cohesive plume aggregates. They behave hydraulically equivalent to fine sands<sup>34,35</sup>. Results from laboratory studies of aggregated plume sediments from the study site<sup>23</sup> using the approach of Thomsen & Gust<sup>31</sup> revealed that bedload transport generated ripples began to form at a critical shear stress of  $0.04 \text{ N m}^{-2}$  and resuspension of recently settled plume sediments starts at  $0.1 \text{ N m}^{-2}$  (see Online Video Content). Multiple turbidity currents originating from multiple lanes could enhance the shape and dimensions of ripples. Larger sediment ripples were mapped at the NE end of the first mining strip (downslope direction), where the gravity current focused into the steeper NE running valley (Fig. 4). Here, the nearest OBS (JUB\_PFM-01) recorded the second-highest concentration ( $169 \text{ mg L}^{-1}$ ; Fig. 3), while the ADCP data at NIOZ\_PFM-03 shows a change in current direction from S to E-NE at the same time (Supplementary Fig. S6). No sediment ripples were found upslope of the mining strips, but nodules were entirely covered by redeposited sediment (Fig. 2).

SE of the impact site, at 500 m distance downslope, the 2 MHz ADCP at NIOZ\_PFM-04 recorded a change in current direction from SE to E, but not in magnitude (Supplementary Fig. S7). This implies that the gravity current had weakened at that distance. However, the comparison with the ADCP at NIOZ\_PFM-07 might be flawed by the relatively low sampling interval of  $5 \text{ min}^2$  and the absence of video data during the time of ADCP measurements that could ensure that the ADCPs measured the same part (e.g., head) of the gravity current.



**Fig. 1 | The three mining lanes and the seafloor morphology of the nodule impact site. a** Global map<sup>65</sup> and the Clarion-Clipperton Zone (CCZ), between Central America and Hawaii, where most of the current polymetallic nodule exploration contract areas are located (black polygons)<sup>66</sup>. The Global Sea Mineral Resources NV exploration contract area, which is divided into three blocks, is marked in yellow. The red dot shows the location of the mining trial described in this study. The blue hashed squares represent Areas of Particular Environmental Interest (i.e., no mining areas), and the light blue blocks are Reserved Areas (i.e., areas reserved for developing nations)<sup>66</sup>. Exclusive Economic Zones<sup>67</sup> are shown in

darker blue. **b** Side-scan sonar map showing the nodule collector mining lanes and strips superimposed on the bathymetry (0.5 m isobath contours) and the locations of nearby sensor platforms. The mining lanes across the three parallel strips are visible due to their lower backscatter intensity related to nodule removal. The white box inside the second strip indicates part of the orthophoto-mosaic area presented in Fig. 2. **c** Slope distribution map of the impact site. Multibeam echosounder residual artifacts cause the NW-SE linear trending seafloor parts with slopes > 6.5° (red coloration).

Nevertheless, this was the most distant location where such a change in the current direction was observed, implying the end of the active phase of sediment transportation as a gravity current. A 6 m high sill with slopes of up to 3° along the plume's path blocked the further eastward propagation, and the plume spread laterally (up to 690 m) around NIOZ\_PFM-04. Here, the AUV mapped the lateral spreading of the plume's lower part (< 5 m altitude) using MBES Water-Column Imaging (WCI) data (Figs. 3, 5).

Beyond this distance, the plume drifted passively with the ambient bottom currents in the SE to S direction (Supplementary Figs. S8–S10). The OBS mounted on the most remote platform at ≈ 1800 m from the impact site (NIOZ\_PFM-01; Fig. 3) recorded a maximum concentration of 3.9 mg L<sup>-1</sup> at 1 m altitude (Fig. 3 and Supplementary Fig. S2). This concentration is three orders of magnitude less than that recorded 50 m from the impact site but is still 100-fold higher than the background concentrations. The plot of the maximum concentrations from OBS located SE of the impact site showed a power-law decay in plume concentration with increasing distance and time at 1 m altitude (Fig. 6). This was visually confirmed during an ROV dive, where the top of the plume was followed. This approach worked well inside and near the impact site where the plume was still quite dense, but towards the SE at ≈ 1700 m distance, it became increasingly difficult to distinguish the top of the plume from overlying clear water.

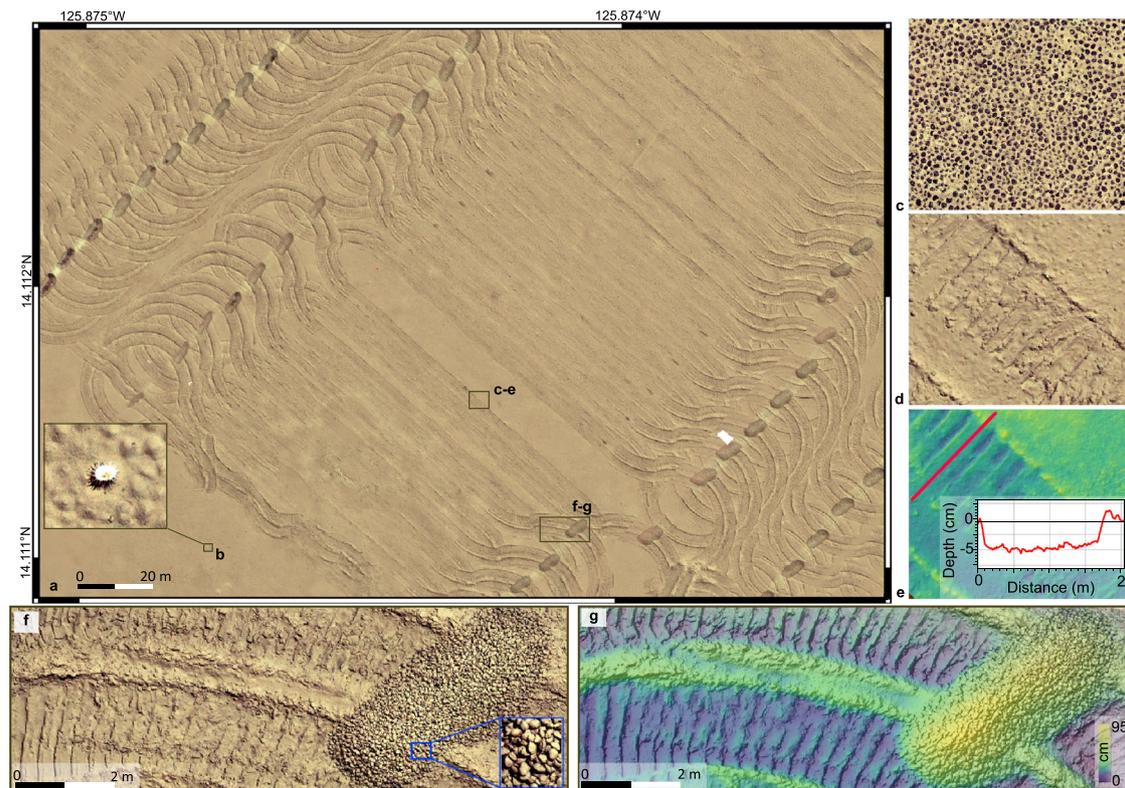
With the AUV flying at 5 m and 10 m altitudes, it was possible to track the plume up to 4.5 km S of the impact site 35 h after the start of the trial. At that distance, the concentration was 0.1 mg L<sup>-1</sup>. The two AUV tracks north of the impact site only recorded background-level values, confirming the SE-S passive advection of the plume by bottom

currents (Fig. 3 and Supplementary Fig. S11). The plume was detected over an area of ≈ 9 km<sup>2</sup>.

N and NW of the impact site, at a distance of ≈ 200 m upslope and upstream, the lowest maximum concentrations (0.05–0.14 mg L<sup>-1</sup>) and backscatter values (34–49 dB) were recorded (Fig. 3 and Supplementary Figs. S2, S12, S13).

With regards to the vertical spreading of the plume, the multi-altitude (5, 10, 30, and 50 m) passes of the AUV at 500 m distance SE from the impact site also showed a power-law decrease in maximum concentrations for the survey period (Fig. 6). The upward-looking ADCPs showed a rapid decrease in backscatter intensity with increasing altitude (Fig. 7 and Supplementary Figs. 1,3–6,10). Moreover, they revealed that while the lower (< 5 m altitude) and denser (up to 35 mg L<sup>-1</sup> at 1 m altitude) part of the plume moved towards the E, the upper part (with a maximum concentration of 3.2 mg L<sup>-1</sup> at 10 m altitude) moved towards SE like the prevailing bottom currents (Supplementary Fig. S7). The 300 kHz ADCPs recorded the plume at altitudes > 30 m at 1800 m SE from the impact site. Together with the AUV OBS records of ≈ 0.07 and 0.06 mg L<sup>-1</sup> at 30 and 50 m, respectively (500 m from the impact site; Fig. 3 and Supplementary Fig. S11), these data show that suspended particles can reach greater heights with distance and time but in low concentration.

The plume evolution has been compiled in Digital Earth Viewer<sup>36</sup>, a web-based 4D visualisation software that allows easy navigation through time and space. Moreover, a time-lapse video from a still camera (GMR\_CAM-37) next to GMR\_PFM-43<sup>2</sup> shows the succession of the first two plume episodes at < 1 m altitude (see Online Video Content).



**Fig. 2 | Orthophoto-mosaics obtained at a depth of 4500 m show the mining imprints on the seafloor. a** Orthophoto-mosaic with 2 mm × 2 mm pixel resolution from part of the second mining strip (Fig. 1b), showing the nodule collector imprints on the seafloor as mining lanes, the “light bulb” turns, and nodule piles. **b** A sea anemone (*Actiniaria*)<sup>88</sup>, a benthic animal capable of cleaning itself from sediment<sup>89</sup>, is shown as an example of the resolution achieved. **c, d** Orthophoto-mosaic of the same seafloor part before (**c**) and after (**d**) the mining trial. **e** Image-derived digital elevation model (DEM) of the seafloor part shown in (**d**). A depth profile across the caterpillar track (red line) shows the erosion depth caused by the nodule collector and the humps on each side of the track due to sideways sediment

extrusion under the caterpillar tracks. The zero-depth refers to the unmined seafloor between the two caterpillar tracks along the same mining lane. The maximum depth shown has been formed after the sediment redeposition and consequently could be smaller than the true erosion depth. **f** Orthophoto-mosaic of the track of the nodule collector while performing the light bulb turning manoeuvre at the end of a lane and the deposited nodule pile ( $\approx L 5 \text{ m} \times W 2 \text{ m} \times H 0.9 \text{ m}$ ). The sideways sediment extrusion under the caterpillar tracks is larger along the “light bulb” turn manoeuvres. **g** Image-derived DEM of the same area with a pixel resolution of 5 mm × 5 mm. A three-dimensional orthophoto-mosaic of this area is also shown in Supplementary Fig. S1.

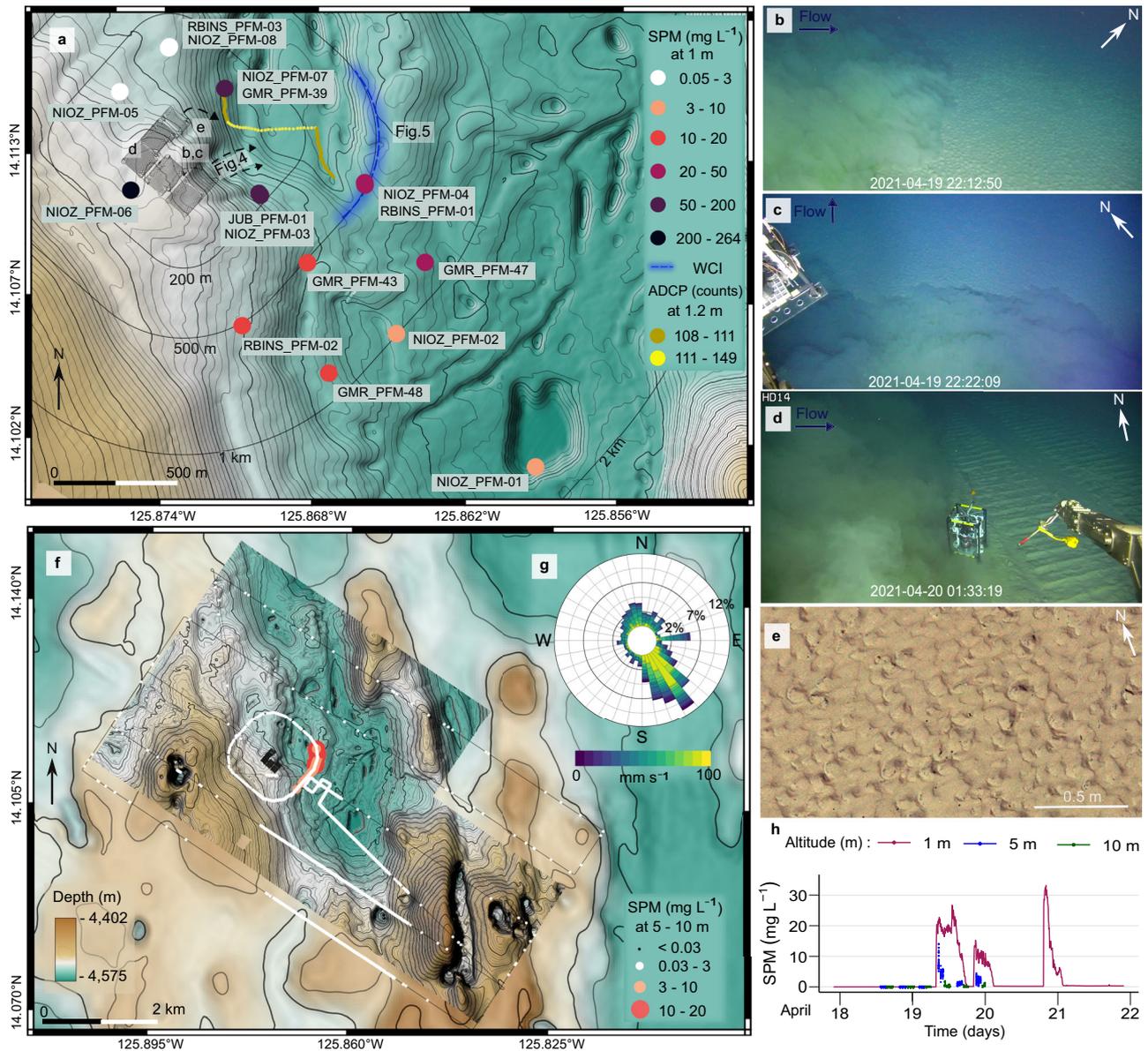
### Particle flocculation

The fact that the plume can be observed with ADCPs at 300 kHz with low sensitivity for primary particle size (median  $\approx 0.02 \text{ mm}$ )<sup>37</sup> indicates that at least part of the mobilised fine-grained sediment must be present in an aggregated form that falls within the detection window of these devices. The PartiCam (JUB\_PFM-01) showed that whereas the median particle size ( $D_{50}$ ) of the top 5 cm of the seafloor sediment is 0.012 mm<sup>17</sup>, the plume  $D_{50}$  increased up to 0.147 mm during the second and third plume episodes (Fig. 8). Throughout the trial, which lasted longer than the minimum time needed to form flocs with CCZ sediment in the lab<sup>23,38</sup>, the first ( $D_{25}$ ) and third quantile ( $D_{75}$ ) ranged from 0.061 mm to 0.313 mm (Fig. 8, Table S1 and Supplementary Fig. S14). The largest particles were detected during the third plume episode when the plume was produced furthest away from the camera and exhibited the longest aggregation time (Fig. 8, Supplementary Table S1 and Supplementary Fig. S14). Although the settling velocity was not measured in situ, laboratory experiments with similar-sized aggregated plume particles from the study area revealed settling rates of  $\approx 100 \text{ m per day}$ <sup>23,39</sup>. Such rates resulted in short residence times within a 2–3 m high plume of 30 – 45 min, with more than 80 % of all aggregated particles settled by then<sup>39</sup>. If these particles settle within the first 500 m away from the nodule collector, the corresponding flow velocity of the gravity current would be  $\approx 0.15 \text{ m s}^{-1}$  for the  $D_{50}$  and  $\approx 0.2 \text{ m s}^{-1}$  for larger particles  $D_{50} - D_{75}$ . These values are close to the

discharge buoyancy velocity of  $\approx 0.21 \text{ m s}^{-1}$  mentioned in the near-field study<sup>20</sup>. They also agree with the in situ observations from the OBS sensor at NIOZ\_PFM-06, which registered the plume for the first time on 19.04.2021 at 19:25:00. Assuming an average and constant speed of  $\approx 0.2 \text{ m s}^{-1}$  for a period of 600 s, the nodule collector should have been  $\approx 120 \text{ m}$  away from the platform at the time of sediment discharge. Plotting the nodule collector track positions shows that on 19.04.2021 at 19:15:00, the collector was 124 m away from NIOZ\_PFM-06 on a flat seafloor.

### Sediment transport and redeposition

The seafloor orthophoto-mosaic and ROV footage that followed the trial showed that the unmined seafloor within the impact site and in the close vicinity ( $< 100 \text{ m}$ ) was completely draped by redeposited sediment so that nodules did not appear as dark objects anymore, i.e., sediment blanketing occurred (Fig. 9 and Supplementary Fig. S15). As the height of the nodules above the sediment surface at this site is 1–3 cm<sup>17,40</sup>, the sediment redeposition was at least 3 cm in these locations (Fig. 9 and Supplementary Fig. S15). The redeposited sediment smoothed out the convex shape of nodules and filled the gaps between neighbouring nodules, resulting in decreased seafloor ruggedness (Fig. 9). Throughout the impact site’s first 500 m SE downslope (along the gravity current propagation path), AUV and ROV seafloor images also showed substantial sediment blanketing



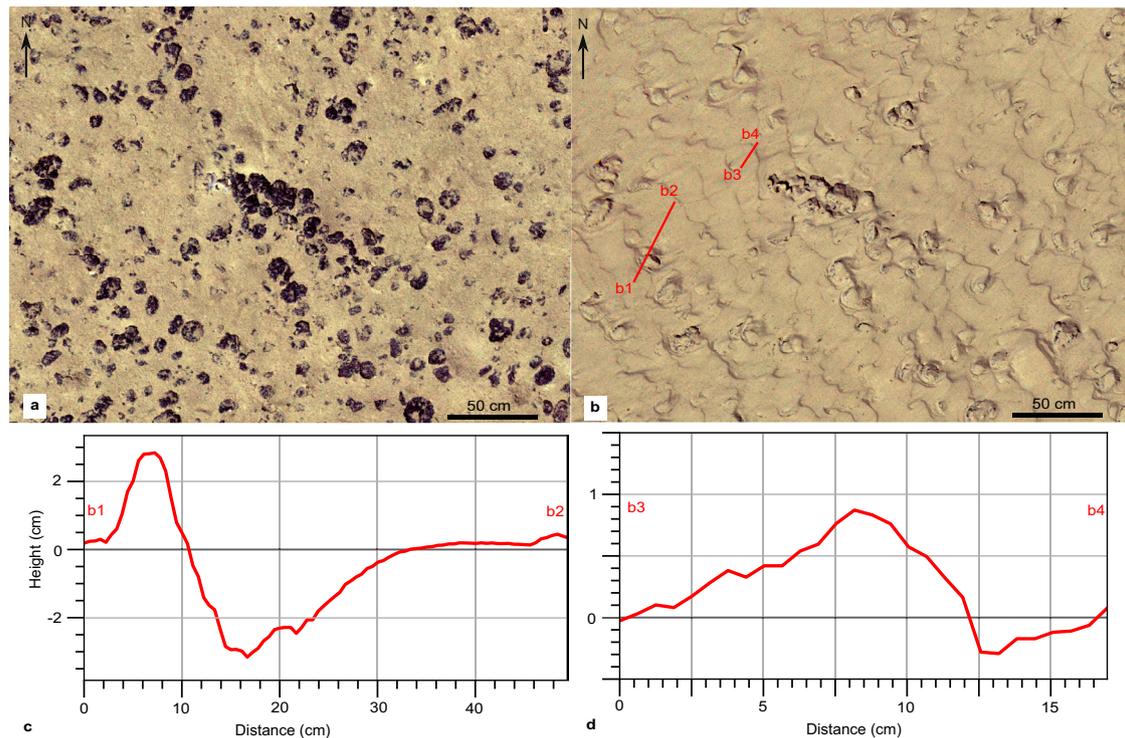
**Fig. 3 | Spatiotemporal distribution of the benthic sediment plume.** **a** Maximum concentration registered by optical backscatter sensors (OBS) at 1 m altitude. Time-series graphs of concentration for all OBS are provided in Supplementary Fig. S2. The acoustic doppler current profiler (ADCP) trajectory plot of the NIOZ\_PFM\_07 platform showed an eastward displacement and increased backscatter intensity when the gravity current passed. The extent of the multibeam echosounder water-column imaging (MBES WCI) data agrees with the OBS data from the autonomous underwater vehicle (AUV) (**e**). **b, c** Video frame grabs inside the nodule impact site show the edge of the benthic sediment plume (plume) being distinct from the ambient seawater. Limited vertical mixing is also observed as the altitude of the remotely operated vehicle (ROV) is  $\approx 4.8$  m. Parts of the ROV footage are available online (see Online Video Content). **d** ROV image<sup>2</sup> showing the plume heading

perpendicular to mining lanes. The sensor shown (Microprofiler)<sup>2</sup> is  $\approx 80$  cm in height and 60 cm in width. **e** Sediment ripples at the end of the third mining strip. **f** The AUV recorded the spatial extent of the plume during a ca. 54 h survey conducted parallel to the mining trial and continued after its termination. Scattered individual measurements above  $0.03 \text{ mg L}^{-1}$  correspond to noise, e.g., from fish inside the sensor's view. The background map is the AUV-based MBES bathymetry of the monitored area superimposed on the ship-based MBES bathymetry<sup>70</sup>. **g** The Rose diagram at NIOZ\_PFM-04 represents the current direction and speed during the nodule collector trial. The eastward spike is related to the gravity current. **h** Concentration recorded at 1 m, 5 m, and 10 m altitude at NIOZ\_PFM-04. The measurements at 5 m and 10 m altitude had to last shorter to facilitate the far-field survey. The locations of Figs. 4 and 5 are also noted.

(Fig. 10 and Supplementary Figs. S15–S17). A simple estimation of the sediment mass transported at 500 m SE from the impact site was attempted using the MBES WCI data and the OBS data from NIOZ\_PFM-04 and RBIS\_PFM-01 (see “Methods”). During a two-hour mining activity in the first strip, 11.2–12.3 % of discharged sediment mass passed through the lower 2 m and an additional 0.81–1.02 % of sediment mass was estimated to pass at 5–6 m altitude (Supplementary Fig. S18). This estimation sets a lower limit, with the actual

percentage expected to be higher as the locations of platforms are on the southern edge along the recorded MBES WCI data, where lower concentrations were observed (Fig. 3). Moreover, the OBS at those platforms measured concentration at 1 m altitude. However, a portion of the plume had already ascended to higher altitudes (see Discussion).

At greater distances from the source, computer-based quantitative analysis of seafloor images (see “Methods”) showed that the



**Fig. 4 | Formation of ripples downstream and downslope from the first mining strip.** **a** Orthophoto-mosaic of seafloor part downslope from the end of the first mining strip before the mining trial. **b** The same seafloor part after the mining trial. The formation of ripples heading towards the East (downslope direction) is depicted. **c, d** Topographic profiles based on the image-derived digital elevation

model (DEM) after the mining trial. The depth of sediment erosion behind the nodule (c) and the ripple's lee side are shown (d). The ripple geometry (stoss side, crest, and lee side) is outlined (d). The ripple crests have a shape sinuous to catenary, which is typical for strong flow over cohesive sediments<sup>90</sup>.

detectable % seafloor nodule coverage ranged between 20 % and 30 % before the mining trial. After the passage of the plume, it dropped to 2–10 % (Fig. 10 and Supplementary Fig. S17). We observed three types of sediment redeposition: thick, medium, and faint coverage (see Methods for classification scheme). The sediment redeposition progressively faded towards the SE direction, where coverage could no longer be visually distinguished on the seafloor images at 1800 m. The faint blanketing is more extended in a southward direction (Fig. 10), whilst no sediment blanketing was observed along the two north-eastern AUV lines, agreeing with the AUV OBS data (Fig. 3).

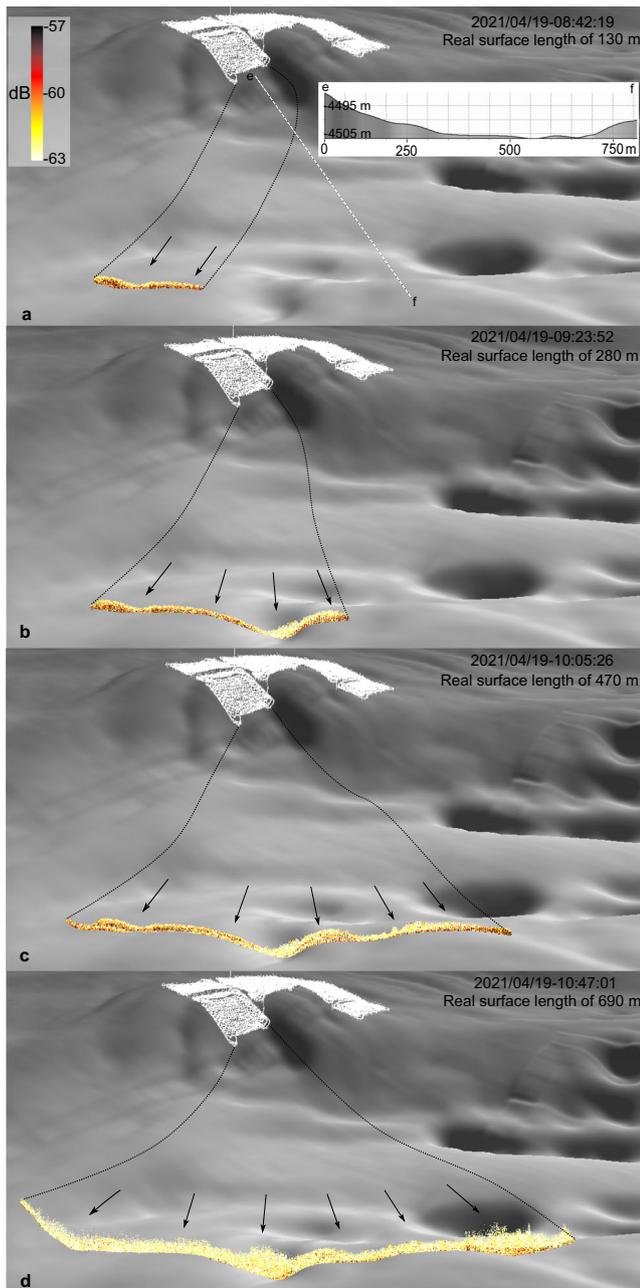
Within the limits of space and time during our monitoring survey, sediment redeposition was detected over an area of  $\approx 6 \text{ km}^2$  (Fig. 10). Estimating conservatively the total volume of mobilised sediment to be  $1700 \text{ m}^3$  (8500 m of collector lanes, 4 m collector width, 0.05 m average erosion depth, without considering sediment suspension around the collector heads and caterpillar tracks), and assuming that all mobilised sediment would settle out over  $\approx 6 \text{ km}^2$ , an average coverage thickness of  $\approx 0.2 \text{ mm}$  is hypothesised (assuming sediment density equal to that of the undisturbed top 0–5 cm layer)<sup>17</sup>. Knowing that a small fraction of the mobilised sediment likely travelled out of the monitored area and that deposition within the monitored area was focused close to the impact site (< 500 m), this 0.2 mm of coverage represents the upper limit in distal parts, which we expect to be even smaller.

## Discussion

Our results demonstrate that the sediments mobilised by the nodule collector generated a gravity current that propagated perpendicular to the mining lanes to distances of up to 500 m downslope from the impact site. Up to this distance, the seafloor slope determined the E-ward direction of the gravity current. The ambient SE-directed

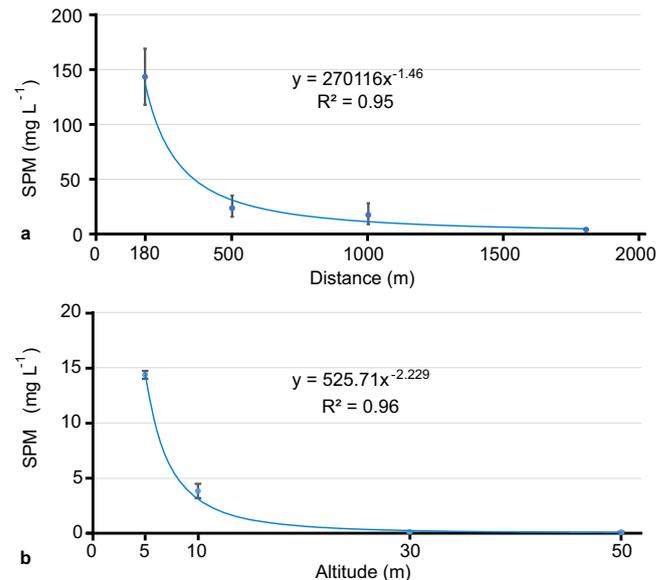
crossflow did not deflect the gravity current from its downslope trajectory within the first 500 m. At the base of the slope, a stretch of flat seafloor (< 1°) bounded towards the East by a sill forming an enclosed basin. The sill blocked the propagation of the gravity current further to the East, and the plume started spreading laterally (Fig. 5). Such flow refraction/diversion of gravity currents has also been shown in other studies<sup>41,42</sup>. From there onward, ambient currents determined the plume's direction towards SE, with finer particles still in suspension dominating the plume composition. The gravity current travelled downslope for longer distances than reported from initial tests of the same nodule collector on flat terrain (< 1°)<sup>20</sup>. As our results indicate that seafloor morphology plays an essential role in gravity current propagation, increasing the area of thick sediment redeposition, we suggest that the AUV-scale mapping would be necessary for future monitoring and plume modelling, while the mining could be prioritised to flat areas limiting the spatial extent of thick sediment redeposition.

The fast sediment redeposition within a limited spatial extent was supported by particle flocculation, observed in situ, corroborating previous lab<sup>23,33,38</sup> and field studies with nodule collectors<sup>8,43,44</sup>, deep-sea epi-benthic sledges<sup>10–13,16</sup>, ROV-induced plumes<sup>45</sup>, trawling gear in muddy sediments<sup>46,47</sup> and redeposition models derived from those tests<sup>11,28,48</sup>. Similarly, empirical power-law shape functions have been reported to describe the sediment concentration decay in low-concentration plumes under currents<sup>49,50</sup>, deep-sea hydrothermal plumes<sup>51</sup> and plumes generated by towed fishing gear<sup>46,47</sup>. This power trend, which expresses the balance between the settling of the SPM (due to gravity) and upward lifting (due to turbulence), can be described by the generalised advection-diffusion equation<sup>52</sup>, if mixing and bottom current advection dominates the plume (which is what occurred here in larger distances and altitudes). At shorter distances



**Fig. 5 | Multibeam echosounder water-column backscatter data 500 m SE of the impact site. a–d** Time-series plot of the multibeam echosounder (MBES) water-column backscatter data 500 m SE from the impact site, showing the evolution of the benthic sediment plume's lower part (< 5 m), corresponding to the first mining strip. The seafloor topography blocked further eastward propagation, and the plume spread laterally. The time (in UTC) corresponds to the start of each MBES file. The seafloor vertical exaggeration is  $\times 9$ .

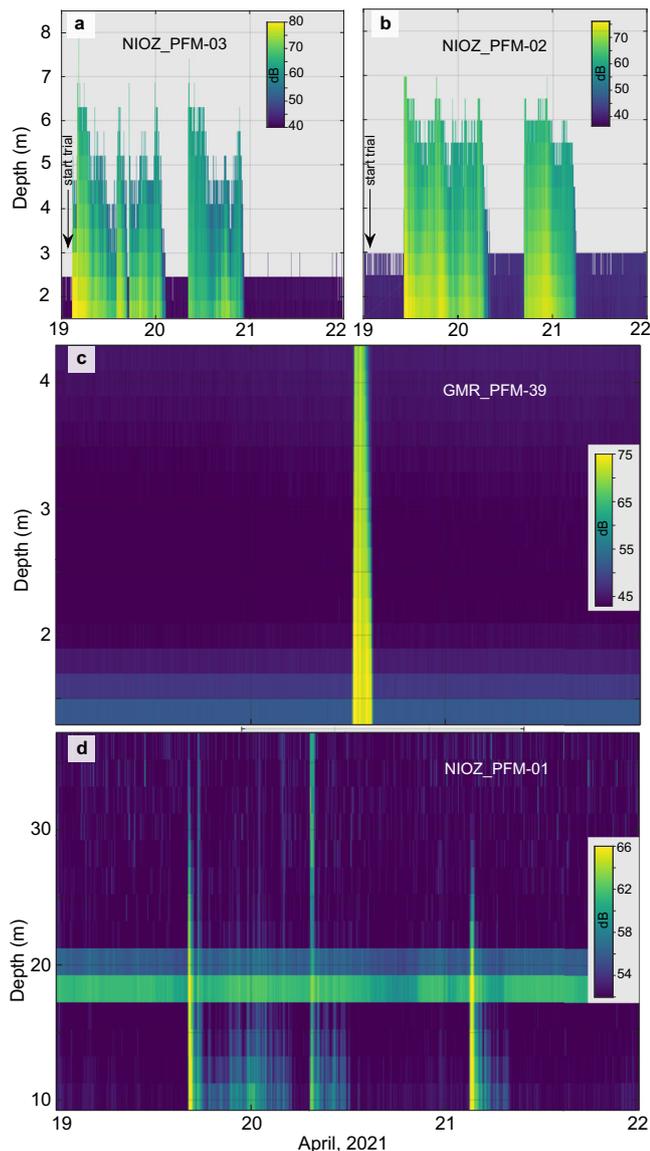
(here < 200 m) and lower altitudes (here < 5 m), the plume shows little mixing, and it is dominated by the near-field gravity current dynamics generated by the mining vehicle<sup>20,53,54</sup>. In future surveys, a denser OBS and ADCP grid closer to the source and along the gravity current propagation path is favourable to depict the concentration decay at short distances and capture the ignition, auto suspension and dissipation phase of the gravity flow<sup>55</sup>. The required time for the mean floc size to form and break up again and its effect on settling velocity within the gravity current and at greater distances should also be



**Fig. 6 | Power-law trend line as functions of distance and altitude for maximum concentrations of suspended particulate matter observed. a** Maximum suspended particulate matter (SPM) concentration recorded by optical backscatter sensors (OBS) at successive distances (from 180 m to 1800 m) and times (19 to 22 April 2021) in the SE direction from the impact site. The maximum values from the OBS on the following platforms were used: JUB\_PFM-01 and NIOZ\_PFM-03 at 180 m ( $n = 3$ ), NIOZ\_PFM-04, RBINS\_PFM-01, GMR\_PFM-43 and RBINS\_PFM-02 at 500 m ( $n = 4$ ), GMR\_PFM-47, NIOZ\_PFM-02 and GMR\_PFM-48 at 1 km ( $n = 3$ ), and NIOZ\_PFM-01 at 1.8 km ( $n = 1$ ). The sensors' locations are given in Fig. 3, and their time series data in Supplementary Fig. S2. When more than one OBS were available, the averaged maximum value was calculated and plotted (blue circle). The vertical bars represent the highest and lowest maximum values measured; this variability is due to the different locations across the longitudinal axis at the same distance from the impact site. **b** The averaged maximum SPM concentration recorded by the two ( $n = 2$ ) OBS mounted on the autonomous underwater vehicle (AUV) at successive altitudes (from 5 m to 50 m) and times (19 to 20 April 2021) while circling at a 500 m distance around the impact site. As two OBS were available at all altitudes ( $n = 2$ ), the average maximum value was calculated and plotted (blue circle). The vertical bars represent the highest and lowest maximum values measured; this variability is due to a) different mounted locations on AUV and looking directions (forward and upward). For a detailed comparison, see Supplementary Fig. S11.

investigated. Here, it was not possible to establish a clear relationship between the concentration gradient and the total percentage of sediment mass passing by in the lower 5 m above the seafloor, as only one measurement at 1 m altitude exists. Our empirical data cannot be extrapolated and applied generally over larger scales or for other trials with different initial conditions (e.g., discharge rate) or other nodule collectors, different sediment size distribution, seafloor morphology and bottom currents without proper understanding of the physical processes contributing to the observed dispersion pattern. However, it does provide a detailed representation of the investigated trial.

The undertaken AUV-mounted MBES survey proved that the lower plume cross-section can be observed if the AUV flies across the plume axis during a low altitude survey and the plume is still highly concentrated, which shows the potential of MBES WCI data to estimate concentrations and grain size, at least for larger aggregates<sup>56–59</sup>. The MBES frequency and settings used<sup>2</sup> could not detect the plume during the far-field AUV survey, where the concentration was  $\approx 0.1 \text{ mg L}^{-1}$  at 5 m altitude and detectable by the OBS on the AUV (Fig. 3). The combined use of two (or more) AUVs that ensniff the plume along and across the spreading axes using higher MBES frequencies (700–1400 kHz) and seafloor platforms with forward-looking<sup>59</sup> or 360° rotating MBES<sup>60</sup>, would close the gap towards a 3D mapping of the



**Fig. 7 | Acoustic doppler current profilers recorded the benthic sediment plume.** **a, b** Water-column backscatter data from two upward-looking acoustic doppler profilers (ADCPs) at 2 MHz located at different distances SE from the impact site. The three benthic sediment plume (plume) episodes are visible due to the sharp increase in backscatter intensity. A  $\approx 10$  dB decrease in maximum backscatter intensity was observed between the two ADCPs, indicating fewer particles at greater distances. **c** An upward-looking ADCP at 1.2 MHz shows a sharp increase in backscatter intensity when the gravity current passes. **d** An upward-looking ADCP at 300 kHz, almost 2 km SE from the impact site, shows the three plume episodes having a larger height. A flotation body on the platform causes the background noise between 18–22 m above the seafloor. All backscatter echo values are normalised. Each ADCP could capture only part of the plume's vertical extent as different frequencies and settings were used<sup>2</sup>.

plume volume at least for the near-field monitoring. The fate of the plume, as it spreads over longer distances and time spans, and the thickness of the redeposited sediment can currently only be derived from 4D numerical models. Improved by in situ data such as those presented here and accounting for particle flocculation<sup>23</sup> and seafloor morphology, such models can provide reliable forecasts under different initial conditions and take account of sporadic environmental events, such as mesoscale eddies reaching the seafloor<sup>61–63</sup>. In this regard, the deployment of long-term seafloor observatories at successive distances from the impact site could enrich our knowledge of

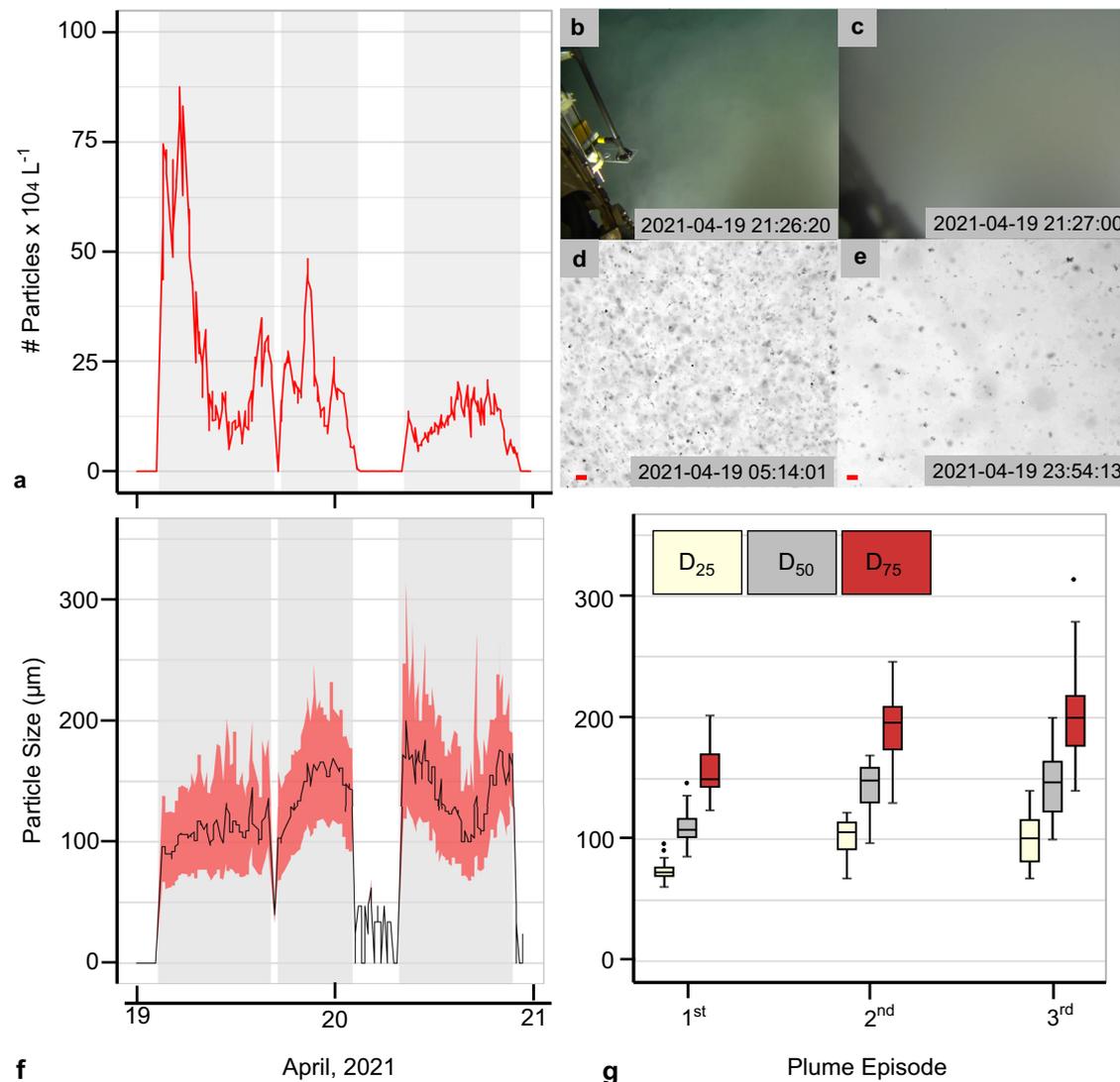
far-field plume dispersion, as an older mining trial with a towed nodule collector and different monitoring technology detected the plume in distances up to 10 km downstream<sup>8</sup>.

Whereas observations made at short range from the nodule collector showed that the plume stayed mostly below 2 m altitude<sup>20</sup>, our results demonstrate that it expanded in height, albeit at a much lower concentration, while spreading from the source. Similar observations have been reported during deep-sea mining pre-prototype collector tests in shallow water<sup>43</sup> and benthic disturbance experiments<sup>45</sup>. In addition, the maximum concentration recorded here is higher than reported from initial testing of the same nodule collector<sup>20</sup>. These differences show that more in situ trials must be closely monitored, particularly with larger nodule collectors or nodule collectors that use different technology. It is important to note that the larger size and longer operation time of a full-scale nodule collector, increasing from a few days to multiple years, will - in comparison with this trial - result in a much larger seafloor mining area ( $\approx 200$  km<sup>2</sup> per year) and volume of resuspended and redeposited sediment over the lifetime of the mine<sup>18</sup>. The method employed for mining, whether along extensive routes or focused on smaller areas, as well as the seafloor slope, would also affect plume dispersion.

In addition, the sediment in suspension around the collector heads and caterpillar tracks should be considered. This resuspended sediment is not entrained into the nodule collector, having different characteristics (e.g., sediment size distribution) and behaviour than the discharge plume. It contributes to the total plume concentration within the impact site, but it is unclear if it is integrated into the gravity current and far-field plume transportation. Moreover, the merging of multiple gravity currents generated by the mining vehicle(s) is a field that should be investigated in detail<sup>64</sup>. Both could be of interest if one or more nodule collectors operate continuously for weeks or months over the seafloor. The erosion depth under the caterpillar tracks is expected to be of similar scale as long as the mass per surface contact area remains similar to the pre-prototype used here and the sediment characteristics (e.g., shear strength) are the same. Using image-derived DEMs to calculate the erosion depth provides good spatial coverage, but it only detects the minimum depth as it cannot account for redeposited sediment that has already settled before the image data acquisition. The slurry produced by onboard processing of collected nodules and discharged via a riser system near the seafloor will increase the SPM in the benthic environment with potentially cumulative effects. However, release close to the seabed may help limit the spreading of the discharge plume and reduce the environmental footprint, thus being preferred over midwater release.

This data-driven study tested existing sensors and technologies for deep-sea plume monitoring as an analogue to the equipment and approaches that we think will most likely be used in future industrial mining scenarios. The results show that the available sensor technology and the combination of fixed and moving platforms can quantify the evolution of plume concentration in space and time, while the image-derived 3D reconstruction of the impact site yielded a resolution down to mm scale that was not available until now. Despite these advancements, two main issues should be addressed in future studies.

Firstly, the post-processing, analysis, collation, visualisation, and interpretation of terabytes of data from different sensors is presently a task that demands computational power and time during onshore work. The current practice does not allow effective decision-making informed by real-time data ingest during future mining activities. Seagoing high-performance computers for image, MBES, SSS analysis, and adaptive spatiotemporal plume modelling (similar to weather forecasts) should become part of the monitoring equipment on mining vessels, an approach already introduced on research vessels<sup>65</sup>. Using such data and numerical models linked to 4D visualisation capabilities that enable extracting, exploring, and visualising quantitative and semantic information in near-real time (digital twin) would allow for



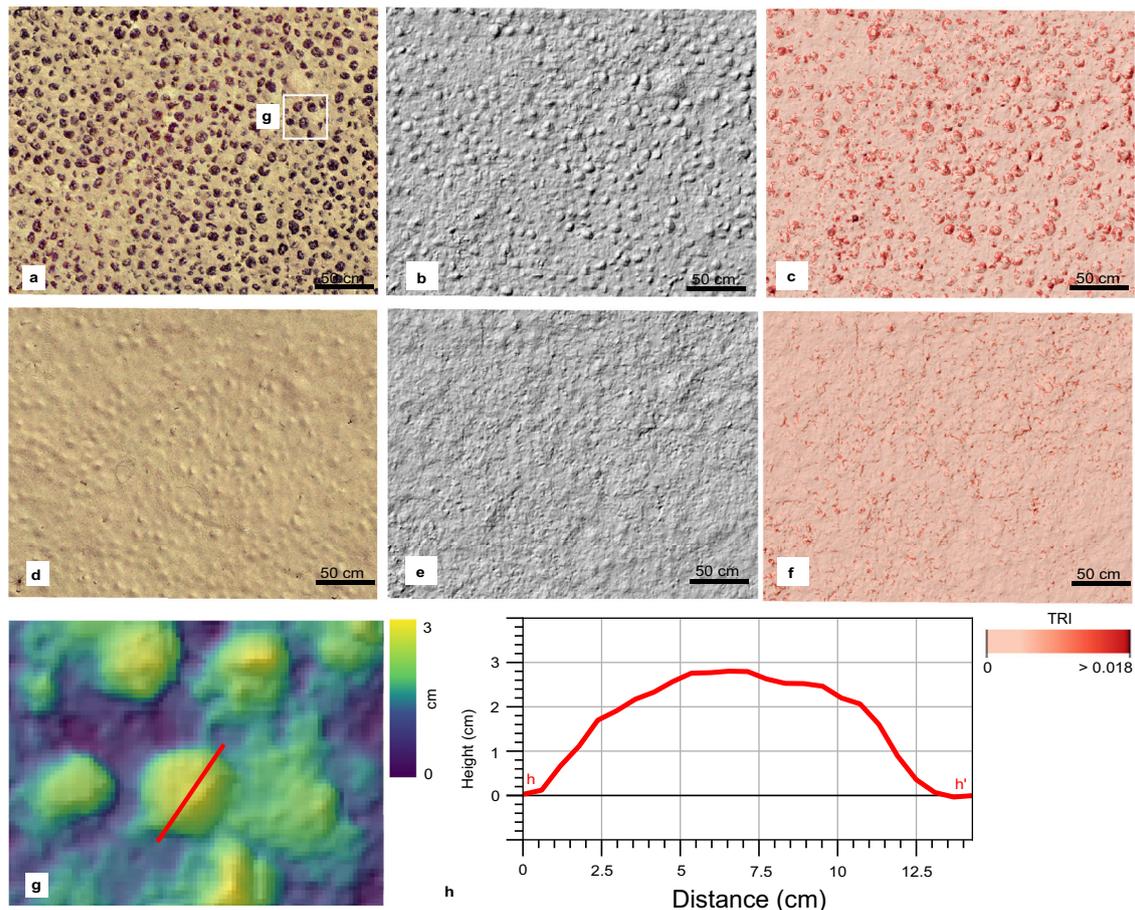
**Fig. 8 | Number of suspended particles and particle size during the mining trial.** **a** Particle number during the mining trial using the PartiCam data (JUB\_PFM-01 platform). Three successive benthic sediment plume (plume) episodes (grey areas) corresponding to the three mining strips were recorded. **b** Video frame grab acquired at 4.7 m altitude above the bulk of the plume. **c** Video frame grab acquired at 4.3 m altitude but now within the plume at the same location (JUB\_PFM-01 platform) 40 s later. The reduced visibility becomes noticeable, with the sensor (Microprofiler)<sup>2</sup> hanging from the arm of the remotely operated vehicle (ROV) at about 1 m from the camera not being visible in (**c**). **d**, **e** In situ PartiCam images during the first day of the mining trial show a decrease in the number of particles

and an increase in particle size with time. The red scale bar is 0.27 mm. The minimum particle size that could be detected with that optical setup was 0.0153 mm with a pixel resolution of 0.0051 mm. **f** Particle size time-series distribution during the mining trial. The black line corresponds to the median particle size ( $D_{50}$ ), and the pink area corresponds to the first ( $D_{25}$ ) and third ( $D_{75}$ ) particle size quantile range. **g** Boxplots of the particle sizes  $D_{25}$ ,  $D_{50}$  and  $D_{75}$  for the 1<sup>st</sup> ( $n = 92$ ), 2<sup>nd</sup> ( $n = 66$ ) and 3<sup>rd</sup> ( $n = 100$ ) plume episodes ( $n = 258$  observations). The black centre line is the median value (50<sup>th</sup> percentile), and the boxes contain each dataset's 25<sup>th</sup> to 75<sup>th</sup> percentiles. The black whiskers mark the 5<sup>th</sup> and 95<sup>th</sup> percentiles, and values beyond these bounds are outliers (black dots).

meaningful and tailored adaptive environmental monitoring during future mining activities. Long-ranging AUVs (endurance of 90 h or more) with integrated multi-sensor instrumentation, such as the one used in this study, support limited near-real-time data communication, e.g., turbidity and position data, allowing an adaptive change of the AUV mission path based on the interpretation of the incoming data. Acoustic communication, small-cabled sensor networks, or repeated upload of AUV data at central data nodes can already be utilised today. Flexible decision-making strategies that are refined based on the incoming environmental data are already applied in the offshore dredging industry, and similar schemes have been proposed for deep-sea mining<sup>66</sup>. Towards automated plume monitoring, existing studies on AUV deep-sea hydrothermal plume tracking<sup>67</sup> could be adjusted for deep-sea nodule mining.

Secondly, the most difficult challenge is to define quantitative tolerance threshold levels of representative deep-sea biota to plume concentrations and thickness of sediment redeposition as a measure to protect the wide range of benthic life inhabiting the seafloor areas targeted by mining. What tolerance limits do deep-sea fauna have for suspended and redeposited sediment of certain particle size distribution and chemical composition, and for what time of exposure before distress or mortality occurs<sup>68,69</sup>? Relating impacts on benthic life observed after the trial to different levels of exposure to suspended and redeposited sediment, as derived from our study, is a first step in that direction but requires time.

This in situ study is part of a worldwide effort to increase our understanding of deep-sea mining and the requirements that future monitoring technologies should meet at a time that the ISA is



**Fig. 9 | Sediment redeposition within the impact site.** **a** Seafloor orthophoto-mosaic before the mining trial within the impact site. The exact location is shown in Fig. 10. **b** Hill shade of the same seafloor part based on the image-derived digital elevation model (DEM). **c** Terrain Ruggedness Index (TRI) of the image-derived DEM (see Methods). **d–f** Orthophoto-mosaic (**d**), hill shade (**e**), and TRI (**f**) of the same seafloor part after the mining trial. The sediment redeposition has covered

the nodules entirely (**d**) and smoothed out the seafloor cm-scale relief (**e–f**). **g** Image-derived DEM of the seafloor before the mining trial showing large nodules and a height profile over a large nodule (red line). **h** The height profile shows that the exposed part of the nodule reaches almost 3 cm above the seafloor before the sediment redeposition and is covered with redeposited sediment after the mining trial (**d**).

draughting regulations for the exploitation of deep-sea minerals<sup>6</sup>, and several nodule collectors are under development<sup>18</sup>.

#### Online video content

A 4D visualisation of SPM concentration ( $\text{mg L}^{-1}$ ) evolution is accessible here: <https://cloud.geomar.de/s/6cZC4Mg8ZPFQIFG>

A timelapse video showing the first two plume episodes is accessible here: <https://cloud.geomar.de/s/BjkHskFmG47onei>

Part of the ROV video footage during the mining trial is accessible here: [https://www.youtube.com/watch?v=wVGmmfqltA&ab\\_channel=GeoChannelBGRLBEG](https://www.youtube.com/watch?v=wVGmmfqltA&ab_channel=GeoChannelBGRLBEG)

[https://www.youtube.com/watch?v=EQ0J5hd2EPY&ab\\_channel=GeoChannelBGRLBEG](https://www.youtube.com/watch?v=EQ0J5hd2EPY&ab_channel=GeoChannelBGRLBEG)

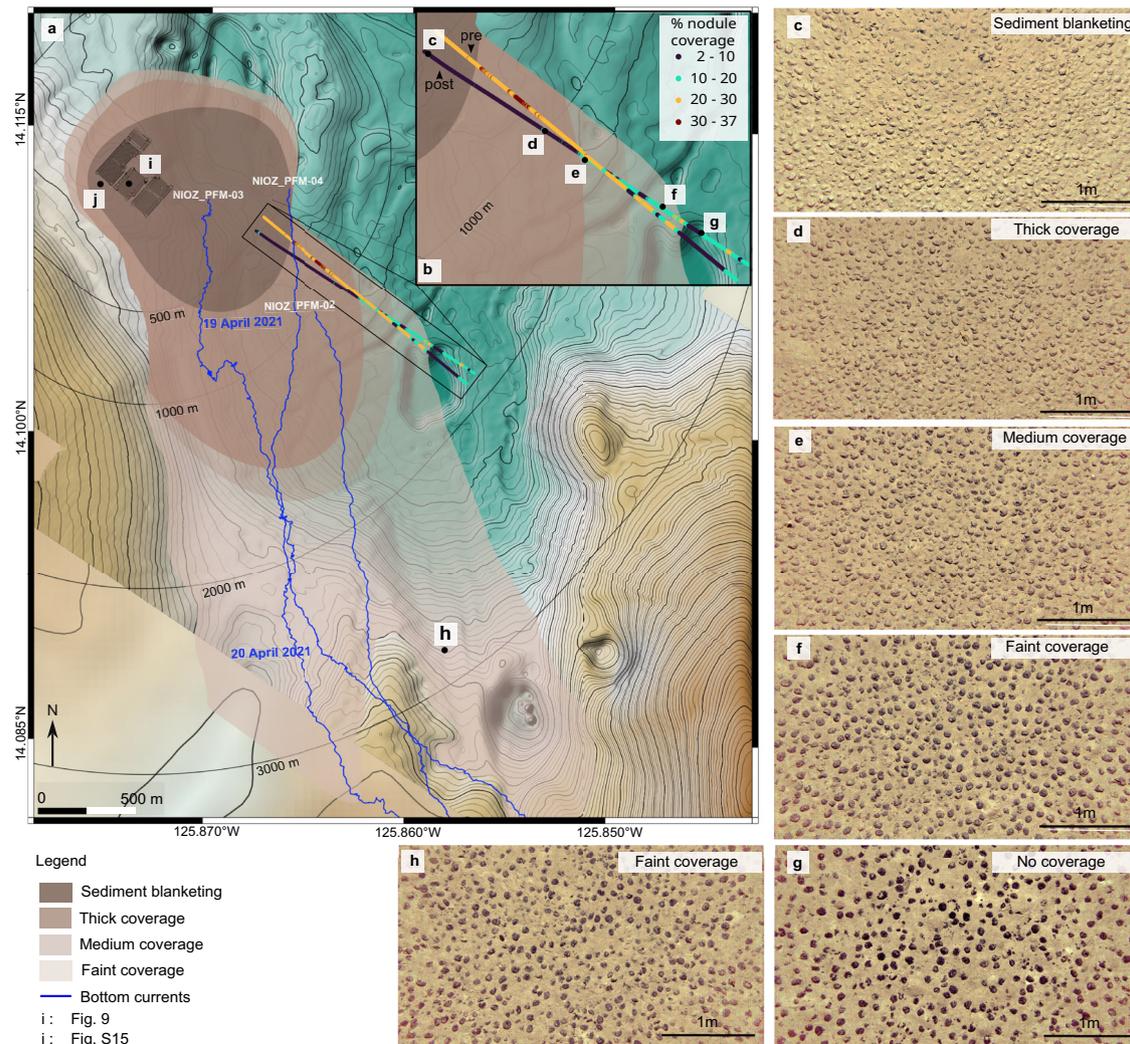
A video on resuspension and bedload transport of recently settled plume sediments at critical shear stress of  $0.04 \text{ N m}^{-2}$  is accessible here: <https://cloud.geomar.de/s/6qR8Df95n25fB6d>

#### Methods

##### AUV MBES & SSS data

A HUGIN 6000 AUV (© Kongsberg SA) was used to acquire MBES, SSS, OBS data and seafloor images. Detailed mission objectives, survey routes and statistics are provided in the expedition report<sup>2</sup>. Multibeam bathymetric and backscatter data from the seafloor and the water column were acquired using the Kongsberg EM2040 at 400 kHz. The

description of the acquisition system settings (e.g., opening angle, beam spacing pattern, pulse mode, AUV altitude) is provided in the expedition report<sup>2</sup>. The raw data post-processing was done with QPS Qimera v1.7 and included full-depth Sound Velocity Profile (SVP) ray tracing correction, roll bias correction, turn removal, filtering of erroneous soundings, and tide correction based on the pressure sensors (Digiquartz<sup>®</sup>) of two SBE16 CTDs at fixed seafloor locations<sup>2</sup>. The navigation data was post-processed onboard using the Kongsberg SA proprietary NavLab software. The absence of an absolute navigational shift was confirmed using the processed ship-based MBES data<sup>70</sup> as a reference layer. The final bathymetric grid was projected on the Universal Transverse Mercator (UTM) zone 10 N and exported as GeoTIFF with a  $2 \text{ m} \times 2 \text{ m}$  cell size. The MBES WCI data were processed with FMMidwater v7 (built-in workflow) and exported as an ASCII dataset. In parallel to MBES mapping, the EdgeTech 2205 SSS running at 230 kHz acquired SSS backscatter data<sup>2</sup>. Data post-processing was done with the CT SonarWiz v7.09 software and included bottom tracking (built-in threshold detection method), gain correction (built-in Empirical Gain Normalisation method), nadir correction, and turn removal. The high overlap between the neighbouring lines resulted in the absence of nadir gaps, while the final mosaicking was done using the average method for blending. A 22 m navigation drift to the west was measured between the SSS and MBES grids, as they have been recorded on different dives. The SSS grid was corrected accordingly using the ESRI



**Fig. 10 | Sediment redeposition map for the entire study area.** **a** The extent of the redeposited sediment based on AUV and ROV footage. A classification scheme of sediment blanketing followed by thick, medium, and faint coverage was used (see “Methods”). The bottom current trajectory plots from three acoustic doppler profilers (ADCPs) are shown in blue. The timestamp is also given for the trajectory plot of NIOZ\_PFM-03. **b** The degree of sediment redeposition expressed as % of

seafloor nodule coverage (see “Methods”) is shown before (pre) and after (post) the passage of the benthic sediment plume. **c–g** Autonomous underwater vehicle (AUV) images showing reductions of sediment coverage at successive distances in the SE direction. An extensive version of (c–g) is given in Supplementary Fig. S17. **h** AUV image of a faint sediment coverage, 2.5 km S-SE from the impact site. **i** Location of Fig. 9. **j** Location of Supplementary Fig. S15.

ArcMap Georeferencing Toolbox. The final backscatter grid was projected in the Universal Transverse Mercator zone 10 N (UTM 10 N) and exported as GeoTIFF files with 0.20 m × 0.20 m cell size.

#### AUV Image data

An AUV HD CathX Colour Still Camera (M12 A1000) with a resolution of 12.5MP (4096 × 3072) was used to acquire seafloor images. The acquired images were colour corrected/normalised using the DeepSea Monitoring group (GEOMAR Helmholtz Centre for Ocean Research Kiel) in-house software Image Normalisation<sup>71</sup>, which removes lighting artefacts and illumination cones created in underwater images while aiming to restore the original seafloor colour. The colour-corrected images from the impact site were used to reconstruct the seafloor orthophoto-mosaics and DEM. The Agisoft Metashape Pro v8.3 software was used following the built-in workflow (Image alignment, Dense Point Cloud creation, Dense Point Cloud Confidence filtering, DEM, and orthophoto mosaic reconstruction). As the AUV camera was provided without underwater calibration parameters, the Agisoft Metashape Pro built-in camera calibration workflow was used

to estimate the camera orientation parameters (e.g., the principal point coordinates). Due to the large number of images (≈ 500, 000), the photomosaic area was divided into chunks of 20,000 images to make computation feasible. The chunk alignment was carried out based on shared features inside the overlapping parts. The absolute seafloor position was corrected where necessary based on the seafloor and mining features derived from the MBES and SSS data using the ESRI ArcMap Georeferencing Toolbox. The final orthophoto-mosaics and DEMs were exported in 2 mm × 2 mm and 5 mm × 5 mm cell sizes, respectively. The TRI was calculated using the Riley formula<sup>72</sup> as applied in the free and open-source software QGIS v3.24, which was also used to create all maps presented here.

The percentage of nodule seafloor coverage was quantified using the saltation GmbH & Co. KG proprietary software Mangan Analyser, fine-tuned for this explicit dataset. The software is based on a Hierarchically Growing Hyperbolic Self-Organising Map (H<sup>2</sup>SOM) approach to segment and detects the % of seafloor nodule coverage per square metre<sup>73</sup>. The algorithm has been used in the past to quantify nodules’ spatial distribution and support resource estimation studies<sup>74–76</sup>. The

same methodology has been applied to quantify the spatial extent of sediment redeposition during an epibenthic sledge disturbance experiment in the CCZ<sup>13</sup>. The difference in detectable % seafloor nodule coverage between the seafloor images before and after the plume passed and (partially or totally) covered up the nodules shows the spatial extent and degree of sediment redeposition. Such an approach works well when the redeposited sediment completely covers the polymetallic nodules (i.e., 0% detection) but not when there is a faint sediment cover (powder) on top of the nodules before the additional sediment deposition<sup>66,75</sup>. The latter is not the case in our study, as shown in seafloor images acquired before the trial (e.g., Supplementary Fig. S16). The images were obtained along two AUV transects in the SE direction  $\approx 29$  h before and  $\approx 72$  h after the mining trial. The two transects have a similar length of  $\approx 1400$  m and a maximum distance of  $\approx 85$  m between them. Only the images acquired in flying altitudes 5–7 m were analysed as they have better quality characteristics (e.g., brightness, contrast). Another limitation is that taking images from the exact same area of seafloor with an AUV is difficult due to slight (cm- to m-scale) navigation uncertainties. In our case, we use images from survey lines that cross each other at a slight angle due to operational restrictions during the AUV dive: two ROVs were used at the same time to recover equipment on the seafloor, and the AUV needed to keep a safety distance, prohibiting flying the same survey lines as before the plume settled. In total, 2678 and 2710 images were analysed before and after the mining trial. The results were exported as an ESRI point shapefile and projected in UTM 10 N. Parallel to detectable % seafloor nodule coverage, qualitative criteria were used to distinguish the degree of sediment redeposition. **Blanketing**: Nodules are entirely covered by redeposited sediments. The open spaces between neighbouring nodules are filled by redeposited sediment. The nodule shape is not distinguishable, and only the nodules' top part, protruding slightly from the seafloor, is still visible. The AUV seafloor photos are yellowish, and the nodules' dark (black) colour is not visible anymore. **Thick**: Nodules are entirely covered by redeposited sediments, but their shape is distinguishable. The nodules' dark colour is still not visible. **Medium**: Nodules are partly covered by redeposited sediments. The nodules' dark colour is still discernable. **Thin**: Nodules are partly covered by redeposited sediments. The brighter sediment slightly shades their dark colour.

### OBS data

Four different types of OBS were mounted on the stationary seafloor platforms<sup>2</sup>: (i) stand-alone JFE Advantech Infinity Series ATUD-USB (JFE), (ii) stand-alone AQUALogger® 310TY (Aqualogger), (iii) WetLabs ECO FLNTU (FLNTU) coupled on SBE16 and SBI9 + CTDs and (iv) Seapoint STM-S (Seapoint) coupled on SBI9 + CTDs. The JFE works at 800 nm, with a 0–1000 FTU range, precision of 0.03 FTU and accuracy of  $\pm 0.3$  FTU ( $\pm 2\%$ ). The AQL works at 880 nm, with a 0–10,000 FTU range, precision  $<1$  FTU and accuracy of  $<2\%$ . The FLNTU works at 700 nm wavelength, with a 0–25 NTU measurement range, sensitivity of 0.013 NTU and precision of  $\approx 0.0065$  NTU. The Seapoint works at 880 nm, with a 0–25 NTU measurement range, sensitivity of 200 mV/FTU (at 100x gain) and linearity of  $\pm 2\%$ . The acquisition settings for all sensors are provided in the expedition report<sup>2</sup>. The AUV carried two OBS: (i) a fixed-mounted FLNTU coupled on a SAIV A/S SD208 6 K CTD and a stand-alone JFE attached on the AUV dorsal fin<sup>2</sup>. The JFE raw data had an offset of one FTU in clear bottom water compared to the calibration dive and the lab calibration. The JFE sensor window was checked and cleaned before each dive, ensuring no sediment, dirt, or damage affected the measurement. The sensor recorded 0.04 FTU in clear CCZ calibration water during the lab calibration, showing no sensor drift over time. A reasonable explanation for the offset is that something (part of the AUV fin, security rope, or a loose end of the tape) was in the sensor's field of view during the AUV dive. Based on comparisons with the FLNTU data, the offset occurred exclusively for

lower concentrations. A hand-fitted offset of 1 (0.98–1.04) FTU was corrected, resulting in a good match with the FLNTU data ( $R^2 = 0.94$ ).

### OBS calibration

All OBS were calibrated to reference suspensions prepared from CCZ surface sediment (top 5 cm of the multicore IP21-070MUC)<sup>2</sup> dispersed in clear CCZ bottom water (IP21-072CTD)<sup>2</sup> at a temperature of 1.5 °C. The principle of calibrating sensors by reference to artificially prepared sediment suspensions has been described earlier by Guillen et al. (2000)<sup>77</sup> and for CCZ sediment by Haalboom et al.<sup>12</sup>. Briefly, the suspensions were contained in a cylindrical calibration vessel made of PVC, sprayed matt black inside to reduce light reflections. A stirring rotor and submerged aquarium wave pump kept the suspension homogeneous, while baffles prevented the formation of a vortex that might create air bubbles. The lid on top had a porthole through which the sensors were inserted, shielded from incident light. For the first calibration step, the vessel was filled with 16 L of CCZ bottom water with an SPM concentration of 0.14 mg L<sup>-1</sup>. Sensors were immersed one after the other in the water for 1 min while measuring at a rate of 1 Hz. For the next calibration step, half the volume of water in the vessel was replaced with an equal volume of prepared sediment suspension, producing a suspension with a concentration of 726 mg L<sup>-1</sup>. In each of the ten subsequent steps, half the volume of suspension in the vessel was replaced by an equal volume of CCZ bottom water, thus producing a sequence of logarithmically decreasing concentrations. The measurements were averaged per calibration step for each sensor, plotted against the concentration of the reference suspensions and fitted with a linear function of the form  $OB = c \times SPM$ , where OB is optical backscatter recorded as FTU for the JFE/Aqualogger and as output voltage for the FLNTU/Seapoint sensors; SPM is the concentration in (mg L<sup>-1</sup>) and c the regression constant. For all JFE, the  $R^2$  of the regression was 1.0000. For all other sensors,  $R^2$  was  $>0.9980$ . Turbidity recorded in the field was converted to SPM concentration (mg L<sup>-1</sup>) by application of the respective regression functions determined for each sensor. Since the sediment plume investigated in our study was artificially produced by the mining machine stirring up fine-grained seafloor sediment into the bottom water, suspensions produced in the lab from seabed sediment and bottom water collected from the same region provide a good analogue for calibrating sensors in the lab. The data quality was assessed by a feasibility check to identify spurious values or instrumental drift. Two FLNTU sensors on GMR\_PFM-47 and GMR\_PFM-48 had an offset of +0.17 NTU and +0.175 NTU, respectively, which was subtracted. For the JFE, internal quality control was used to discard erroneous data. Data recorded during the platforms' deployment, relocation, and recovery were dismissed. The AUV FLNTU sensor, belonging to MV Island Pride, was intercalibrated as it remained onboard and could not be calibrated with in situ sediment at land. Using the relationship between SPM concentration (mg L<sup>-1</sup>) and optical backscatter (FTU/NTU) established during lab calibration of the JFE sensor and the relationship between AUV JFE and AUV FLNTU data, the AUV FLNTU turbidity data could be converted to SPM concentration (mg L<sup>-1</sup>). The total least squares – orthogonal regression (TLS) was used to calculate the relationship between the corrected and calibrated AUV JFE and AUV FLNTU turbidity values. TLS was preferred over the commonly applied Ordinary Least Squares (OLS) as TLS accounts for errors in independent and dependent variables. The AUV FLNTU turbidity in SPM (mg L<sup>-1</sup>) was found to be  $SPM_{FLNTU} = \frac{OBS_{FLNTU} - 0.02278291}{1.0415194 \cdot c}$ , with c being the regression constant.

### PartiCam data

The PartiCam is composed of two separate pressure housings, one containing the camera and one containing the flash. It was equipped with a Canon EOS 760D SLR with a resolution of 24.2 megapixels and a Canon EF-S 60 mm f 2.8 macro lens. Using a macro lens allows the images to be taken closer to the lens, providing a higher optical

resolution and a more even illumination of the sample volume. A collimated light source (Yongnuo YN-468 II Speedlite strobe) was lined up and facing the camera with a  $\approx 2$  cm slit between the flash and camera window through which water could flow. A short flash duration of 1/200 s allowed the acquisition of images containing particles that were in focus without motion blur effects. This optical setup's minimum detectable particle size is 0.0153 mm, with a pixel resolution of 0.0051 mm. The MATLAB Image Processing toolbox<sup>78</sup> was used following the methodology of Iversen et al.<sup>79</sup>, and the ParChar architecture code<sup>80</sup> was run to remove the image background and extract the size of the particles in each image. A rolling median ( $k=3$ , with  $k$  being the integer width of the rolling window) was applied to remove a few isolated outliers from the raw data and allow better visualisation of the trend of increasing  $D_{25}$ ,  $D_{50}$ , and  $D_{75}$  during the mining trial. The R zoo package was used<sup>81</sup>. The raw data before and after applying the rolling median are shown in Supplementary Fig. S14.

### ADCP data

In total, 5 Teledyne RDI Workhorse Sentinel ADCPs at 300 kHz, one at 1.2 MHz and 7 Nortek ADCPs at 2 MHz were deployed on fixed sea-floor platforms. A detailed overview of the sensor specifications and acquisition settings is provided in the expedition report<sup>2</sup>. Data from 10 devices are presented here (see Supplementary Material). Firstly, the data recorded during platform deployment, recovery, or any relocation by the ROV were removed. Secondly, the data in bins where the acoustic backscatter was lower than 25 counts for high-frequency (2 MHz) ADCPs and 40 counts for other lower-frequency ADCP sensors were discarded to ensure the reliability of measured current speed and direction. The manufacturers provide these values as the noise floor, i.e., the lower backscatter threshold, to ensure good data. In addition to the set ensemble averaging performed internally by each device during the measurement, the data were averaged over the user-defined burst duration. The measured raw acoustic backscattered signal in counts was converted to echo strength (dB) and corrected for sound attenuation away from the source by both acoustic spreading and absorption by water<sup>82</sup>. Due to the naturally low concentration of suspended particles in the deep ocean water column<sup>12,17,83</sup>, even during the plume episodes (mostly below  $200 \text{ mg L}^{-1}$ ), the sound attenuation by the suspended particles was assumed to be negligible<sup>84</sup>.

### Sediment transport estimation

The AUV MBES WCI data captured the plume's lower part ( $< 5$  m) during a total time of 3.5 h (19 April 2021; from 08:40 to 12:05) on the first day of the mining trial (1st strip, Fig. 3 and Fig. 5). The recorded data showed a variable spatial extent while the plume was spreading over a maximum distance of  $\approx 700$  m, which was recorded continuously for the last two hours (Fig. 5). The 2 MHz ADCP at NIOZ\_PFM-04 showed that the average current direction and magnitude at 2.25 m altitude when the plume was passing by, was  $117^\circ$  and  $0.062 \text{ m s}^{-1}$  respectively. According to the OBS data at NIOZ\_PFM-04 and RBNIS\_PFM-01, the average concentration was  $18.37 \text{ mg L}^{-1}$  at 1 m altitude. The cross-sectional area defined by the MBES WCI maximum spatial extent along the AUV flight path was nearly normal to the plume axis (defined by the average current direction) with a  $22^\circ$  deviation from this. It has a length of 640 m, a height of 5 m and a  $95^\circ$  angle between its normal vector and the N (Supplementary Fig. S18). We assumed a constant concentration, current direction and magnitude (equal to the above values) for the entire cross-sectional area of the first two metres of altitude during the two hours with the maximum recorded plume spread. The altitude of 2 m is based on previous work<sup>2</sup>, which showed that 92 to 98% of the sediment mass remains below 2 m. The nodule collector average flow rate was also assumed constant for these two hours ( $12 \text{ kg s}^{-1}$ )<sup>20</sup>. The equations used for estimating the

sediment transport are as follows:

$$\Phi = SPM_c U_{ADCP} \cos \alpha \quad (1)$$

$$F_r = \Phi A_{cs} \quad (2)$$

$$M = F_r \Delta t \quad (3)$$

$\Phi$  : Particles flux [ $\text{kg m}^{-2} \text{ s}^{-1}$ ]

$SPM_c$  : Suspended particulate matter [ $\text{kg m}^{-3}$ ]

$U_{ADCP}$  : Current magnitude given by ADCP measurements [ $\text{ms}^{-1}$ ]

$\alpha$  : Angle in degrees between the gravity current direction (ADCP measurement) and the unit vector normal to cross – sectional area

$F_r$  : Particles flow rate across cross – sectional area [ $\text{kg s}^{-1}$ ]

$A_{cs}$  : Cross – sectional area [ $\text{m}^2$ ]

$M$  : Sediment mass [ $\text{Kg}$ ]

$\Delta t$  : Time window [ $\text{sec}$ ]

The same equations were used to calculate the sediment mass at 5 m altitude along the same AUV flight path 500 m from the impact site. The two AUV OBS data sets revealed a similar lateral extent as the MBES WCI (slightly extended to SE direction), corroborating the hydroacoustic results (Fig. 3). The concentration spatial distribution along the AUV flight path is similar to the MBES WCI data, with higher concentration in the eastern part (Fig. 3). The average concentration along this pathway is  $3.15 \text{ mg L}^{-1}$ . When the plume was passing by, the average current direction and magnitude at 5.25 m altitude were  $133^\circ$  SE and  $0.06 \text{ m s}^{-1}$ , respectively. As with the sediment mass calculation above, we assumed constant concentration, current direction and magnitude for the entire cross-sectional area between 5 and 6 m altitude during the two hours of plume recording.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All data associated with the results of this study are presented in the paper, Supplementary Material and online video content. The data will be published on PANGAEA® Data Publisher for Earth & Environmental Science (<https://www.pangaea.de/>) and can be obtained from the corresponding author upon personal request.

### References

- Thomson, W. Notes from the “Challenger” \*. *Nature* **8**, 51–53 (1873).
- Vink, A. et al. MANGAN 2021 Cruise Report: Independent scientific monitoring of two collector tests in the BGR and GSR contract areas for the exploration of polymetallic nodules in the equatorial NE Pacific. <https://doi.org/10.25928/hw7d-fs42> (2022).
- Halbach, P. et al. Geochemical and mineralogical control of different genetic types of deep-sea nodules from the Pacific Ocean. *Miner. Deposita* **16**, 59–84 (1981).
- Hein, J. R., Koschinsky, A. & Kuhn, T. Deep-ocean polymetallic nodules as a resource for critical materials. *Nat. Rev. Earth Environ.* **1**, 158–169 (2020).
- UN General Assembly, Convention on the Law of the Sea, 10 December 1982. [https://www.un.org/Depts/los/convention\\_agreements/convention\\_overview\\_convention.htm](https://www.un.org/Depts/los/convention_agreements/convention_overview_convention.htm). (2023).
- International Seabed Authority, The Mining Code, Draft regulations for exploitation of mineral resources in the Area. <https://www.isa.org/jm/the-mining-code/draft-exploitation-regulations/> (2023).
- Amon, D. J. et al. Assessment of scientific gaps related to the effective environmental management of deep-seated mining. *Mar. Policy* **138**, 105006 (2022).

8. Ozturgut, E., Lavelle, J. W. & Burns, R. E. Chapter 15 impacts of manganese nodule mining on the environment: Results from pilot-scale mining tests in the North Equatorial Pacific. *Elsevier Oceanogr. Ser.* **27**, 437–474 (1981).
9. Burns, R. E. Assessment of environmental effects of deep ocean mining of manganese nodules. *Helgoländer Meeresunters.* **33**, 433–442 (1980).
10. Gausepohl, F., Hennke, A., Schoening, T., Köser, K. & Greinert, J. Scars in the abyss: Reconstructing sequence, location and temporal change of the 78 plough tracks of the 1989 Discol Deep-sea disturbance experiment in the Peru Basin. *Biogeosciences* **17**, 1463–1493 (2020).
11. Baeye, M. et al. Tidally driven dispersion of a deep-sea sediment plume originating from seafloor disturbance in the discol area (Se-Pacific Ocean). *Geosciences* **12**, 8 (2021).
12. Haalboom, S. et al. Monitoring of anthropogenic sediment plumes in the Clarion-Clipperton Zone, NE equatorial Pacific Ocean. *Front. Mar. Sci.* **9**, <https://doi.org/10.3389/fmars.2022.882155> (2022).
13. Peukert, A. et al. Understanding MN-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and Optical Data. *Biogeosciences* **15**, 2525–2549 (2018).
14. Jones, D. O. et al. Biological responses to disturbance from simulated deep-sea polymetallic nodule mining. *PLOS ONE* **12**, e0171750 (2017).
15. Simon-Lledó, E. et al. Biological effects 26 years after simulated deep-sea mining. *Sci. Rep.* **9**, 8040 (2019).
16. Lefaille, N. et al. Digging deep: Lessons learned from meiofaunal responses to a disturbance experiment in the Clarion-Clipperton Zone. *Mar. Biodivers.* **53**, 48 (2023).
17. Global Sea Mineral Resources NV (GSR) Environmental Impact Statement. Small-scale testing of nodule collector component on the seafloor of the clarion-clipperton fracture zone and its environmental impact, <https://www.isa.org/jm/minerals/environmental-impact-assessments> (2018).
18. Weaver, P. P. E. et al. Assessing plume impacts caused by polymetallic nodule mining vehicles. *Mar. Policy* **139**, 105011 (2022).
19. Volz, J. B. et al. Natural spatial variability of depositional conditions, biogeochemical processes and element fluxes in sediments of the Eastern Clarion-Clipperton Zone, Pacific Ocean. *Deep Sea Res. Part 1 Oceanogr. Res. Pap.* **140**, 159–172 (2018).
20. Muñoz-Royo, C., Ouillon, R., El Mousadik, S., Alford, M. H. & Peacock, T. An in situ study of abyssal turbidity-current sediment plumes generated by a deep seabed polymetallic nodule mining preprototype collector vehicle. *Sci. Adv.* **8**, eabn1219 (2022).
21. MiningImpact: Environmental impacts and risks of deep-sea mining. <https://miningimpact.geomar.de> (2023).
22. Joint Programming Initiative Healthy and Productive Seas and Oceans: Ecological Aspects of Deep-Sea Mining. <https://jpi-oceans.eu/en/ecological-aspects-deep-sea-mining> (2022).
23. Gillard, B. et al. Physical and hydrodynamic properties of deep sea mining-generated, abyssal sediment plumes in the Clarion Clipperton Fracture Zone (Eastern-central Pacific). <https://doi.org/10.1525/elementa.194> (2017).
24. Martinez Arbizu, P. & Haeckel M. RV SONNE Fahrtbericht / Cruise Report SO239: EcoResponse Assessing the Ecology, Connectivity and Resilience of Polymetallic Nodule Field Systems, Balboa (Panama) – Manzanillo (Mexico), 11.03.-30.04.2015 GEOMAR Report, N. Ser. 025. GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany, 204 pp. (2015).
25. Haeckel, M. & Linke, P. RV SONNE Cruise Report SO268 –Assessing the Impacts of Nodule Mining on the Deep-Sea Environment: Nodule Monitoring, Manzanillo (Mexico) – Vancouver (Canada), 17.02. – 27.05.2019 GEOMAR Report, N. Ser. 059. GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Kiel, Germany, 359 + Appendix (in all 802) pp. [https://doi.org/10.3289/GEOMAR\\_REP\\_NS\\_59\\_20](https://doi.org/10.3289/GEOMAR_REP_NS_59_20) (2021).
26. Purkiani, K. et al. Evidence of eddy-related deep-ocean current variability in the northeast tropical Pacific Ocean induced by Remote Gap Winds. *Biogeosciences* **17**, 6527–6544 (2020).
27. Purkiani, K. et al. Impact of a long-lived anticyclonic mesoscale eddy on seawater anomalies in the northeastern tropical Pacific Ocean: A composite analysis from hydrographic measurements, sea level altimetry data, and reanalysis model products. *Ocean Sci.* **18**, 1163–1181 (2022).
28. Purkiani, K. et al. Numerical simulation of deep-sea sediment transport induced by a dredge experiment in the northeastern Pacific Ocean. *Front. Mar. Sci.* **8**, <https://doi.org/10.3389/fmars.2021.719463> (2021).
29. Gillard, B., Harbour, R. P., Nowald, N., Thomsen, L. & Iversen, M. H. Vertical Distribution of particulate matter in the Clarion Clipperton Zone (German sector)—potential impacts from deep-sea mining discharge in the water column. *Front. Mar. Sci.* **9**, <https://doi.org/10.3389/fmars.2022.820947> (2022).
30. Koller, D., Manica, R., Borges, A. de & Fedele, J. Experimental bed-forms by Saline Density Currents. *Braz. J. Geol.* **49**, <https://doi.org/10.1590/2317-4889201920180118> (2019).
31. Thomsen, L. & Gust, G. Sediment erosion thresholds and characteristics of resuspended aggregates on the Western European continental margin. *Deep Sea Res. Part 1 Oceanogr. Res. Pap.* **47**, 1881–1897 (2000).
32. McCave, I. N. & Hall, I. R. Size sorting in marine muds: Processes, pitfalls, and prospects for paleoflow-speed proxies. *Geochem. Geophys. Geosyst.* **7**, <https://doi.org/10.1029/2006GC001284> (2006).
33. Ali, W., Enthoven, D., Kirichek, A., Chassagne, C. & Helmons, R. Effect of flocculation on turbidity currents. *Front. Mar. Sci.* **10**, <https://doi.org/10.3389/fmars.2022.1014170> (2022).
34. Thomsen, L. A. & McCave, I. N. Aggregation processes in the benthic boundary layer at the Celtic Sea Continental margin. *Deep Sea Res. Part 1 Oceanogr. Res. Pap.* **47**, 1389–1404 (2000).
35. Schieber, J., Southard, J. & Thaisen, K. Accretion of mudstone beds from migrating floccule ripples. *Science* **318**, 1760–1763 (2007).
36. Buck, V., Stäbler, F., Mohrmann, J., González, E. & Greinert, J. Visualising geospatial time series datasets in realtime with the digital earth viewer. *Comput. Graph.* **103**, 121–128 (2022).
37. Haalboom, S. et al. Suspended particulate matter in a submarine canyon (Whittard Canyon, Bay of Biscay, Ne Atlantic Ocean): Assessment of commonly used instruments to record turbidity. *Mar. Geol.* **434**, 106439 (2021).
38. Becker, H. J., Grupe, B., Oebius, H. U. & Liu, F. The behaviour of deep-sea sediments under the impact of nodule mining processes. *Deep Sea Res. Part 2 Top. Stud. Oceanogr.* **48**, 3609–3627 (2001).
39. Gillard & Thomsen (2020). Characterization of sediment plumes behind mining vehicles in the NORI area Collector test study. Environmental Impact Statement, Appendix 6. Report to the ISA, March, 2022. <https://metals.co/nori/> (2024).
40. Schoening, T. & Gazis, I.-Z. Sizes, weights and volumes of polymetallic nodules from box cores taken during SONNE cruises SO268/1 and SO268/2. GEOMAR - Helmholtz Centre for Ocean Research Kiel, PANGAEA, (2019).
41. Howlett, D. M. et al. Response of unconfined turbidity current to deep-water fold and thrust belt topography: Orthogonal incidence on solitary and segmented folds. *Sedimentology* **66**, 2425–2454 (2019).
42. Patacci, M., Haughton, P. D. W. & McCaffrey, W. D. Flow behavior of ponded turbidity currents. *J. Sediment. Res.* **85**, 885–902 (2015).
43. Haalboom, S. et al. Monitoring of a sediment plume produced by a deep-sea mining test in shallow water, Málaga Bight, Alboran Sea (Southwestern Mediterranean Sea). *Mar. Geol.* **456**, 106971 (2023).

44. Mousadik, S. E. et al. In situ optical measurement of particles in sediment plumes generated by a pre-prototype polymetallic nodule collector. *Sci. Rep.* **14**, 23894 (2024).
45. Spearman, J. et al. Measurement and modelling of deep sea sediment plumes and implications for deep sea mining. *Sci. Rep.* **10**, 5075 (2020).
46. Durrieu de Madron, X. et al. Trawling-induced resuspension and dispersal of muddy sediments and dissolved elements in the Gulf of Lion (NW mediterranean). *Cont. Shelf Res.* **25**, 2387–2409 (2005).
47. O'Neill, F. G. et al. Monitoring the generation and evolution of the sediment plume behind towed fishing gears using a multibeam echosounder. *ICES J. Mar. Sci.* **70**, 892–903 (2013).
48. Lavelle, J., Ozturgut, E., Swift, S. & Erickson, B. Dispersal and re-sedimentation of the benthic plume from deep-sea mining operations: a model with calibration. *Mar. Mining* **3**, 59–93 (1981).
49. Bolaños, R., Thorne, P. D. & Wolf, J. Comparison of measurements and models of bed stress, bedforms and suspended sediments under combined currents and waves. *Coast. Eng.* **62**, 19–30 (2012).
50. Kumar, S., Singh, H., Balaji, S., Hanmaiahgari, P. & Pu, J. Inclusive hyper- to dilute-concentrated suspended sediment transport study using modified rouse model: Parametrized power-linear coupled approach using machine learning. *Fluids* **7**, 261 (2022).
51. Bemis, K. G. et al. A comparison of black smoker hydrothermal plume behavior at Monolith Vent and at Clam Acres Vent Field: Dependence on source configuration. *Mar. Geophys. Res.* **23**, 81–96 (2002).
52. Cantero-Chinchilla, F. N., Castro-Orgaz, O. & Dey, S. Distribution of suspended sediment concentration in wide sediment-laden streams: A novel power-law theory. *Sedimentology* **63**, 1620–1633 (2016).
53. Peacock, T. & Ouillon, R. The fluid mechanics of deep-sea mining. *Annu. Rev. Fluid Mech.* **55**, 403–430 (2023).
54. Ouillon, R., Kakoutas, C., Meiburg, E. & Peacock, T. Gravity currents from moving sources. *J. Fluid Mech.* **924**, <https://doi.org/10.1017/jfm.2021.654> (2021).
55. Heerema, C. J. et al. What determines the downstream evolution of turbidity currents? *Earth Planet. Sci. Lett.* **532**, 116023 (2020).
56. Fromant, G. et al. Suspended sediment concentration field quantified from a calibrated multibeam EchoSounder. *Appl. Acoust.* **180**, 108107 (2021).
57. Simmons, S. M. et al. An evaluation of the use of a multibeam echosounder for observations of suspended sediment. *Appl. Acoust.* **126**, 81–90 (2017).
58. Praet, N. et al. The potential of multibeam sonars as 3D turbidity and SPM monitoring tool in the North Sea. *Remote Sens.* **15**, 4918 (2023).
59. Hughes Clarke, J. E. First wide-angle view of channelized turbidity currents links migrating cyclic steps to flow characteristics. *Nat. Commun.* **7**, 11896 (2016).
60. Marcon, Y., Kelley, D., Thornton, B., Manalang, D. & Bohrmann, G. Variability of natural methane bubble release at Southern Hydrate Ridge. *Geochem. Geophys. Geosyst.* **22**, <https://doi.org/10.1029/2021GC009894> (2021).
61. Aleynik, D., Inall, M. E., Dale, A. & Vink, A. Impact of remotely generated eddies on plume dispersion at abyssal mining sites in the Pacific. *Sci. Rep.* **7**, 16959 (2017).
62. Kontar, E. A. & Sokov, A. V. A benthic storm in the northeastern tropical Pacific over the fields of manganese nodules. *Deep Sea Res. Part 1 Oceanogr. Res. Pap.* **41**, 1069–1089 (1994).
63. Hollister, C. & McCave, I. Sedimentation under deep-sea storms. *Nature* **309**, 220–225 (1984).
64. Alhaddad, S. & Elerian, M. Mitigating suspended-sediment environmental pressure in subsea engineering through colliding turbidity currents. *Res. Eng.* **21**, 101916 (2024).
65. Schoening, T. SHIPCC—a sea-going high-performance compute cluster for image analysis. *Front. Mar. Sci.* **6**, 6 (2019).
66. Helmons, R., de Wit, L., de Stigter, H. & Spearman, J. Dispersion of benthic plumes in deep-sea mining: What lessons can be learned from dredging? *Front. Earth Sci.* **10**, <https://doi.org/10.3389/feart.2022.868701> (2022).
67. von See, T. B., Meurer, T. & Greinert, J. Chemical Plume tracking using an AUV with UKF based extremum seeking. *IFAC-PapersOnLine* **55**, 178–183 (2022).
68. Stenvers, V. I. et al. Experimental mining plumes and ocean warming trigger stress in a deep pelagic jellyfish. *Nat. Commun.* **14**, 7352 (2023).
69. Lefaible, N. et al. Industrial mining trial for polymetallic nodules in the Clarion-Clipperton Zone indicates complex and variable disturbances of meiofaunal communities. *Front. Mar. Sci.* **11**, <https://doi.org/10.3389/fmars.2024.1380530> (2024).
70. Gazis, I.-Z. Processed EM122 multibeam swath bathymetry collected during SONNE cruise SO268/1 inside the Belgian License Area in Clarion Clipperton Zone, Pacific. GEOMAR - Helmholtz Centre for Ocean Research Kiel, PANGAEA, <https://doi.org/10.1594/PANGAEA.915764> (2020).
71. Heger, K. & Petersen, L. Tomato (Software). (2024).
72. Riley, S. J., DeGloria, S. D. & Elliot, R. A terrain ruggedness index that quantifies topographic heterogeneity. *Int. J. Sci.* **5**, 1–4 (1999).
73. Schoening, T., Kuhn, T., Jones, D. O. B., Simon-Lledo, E. & Nattkemper, T. W. Fully automated image segmentation for benthic resource assessment of polymetallic nodules. *Methods Oceanogr.* **15–16**, 78–89 (2016).
74. Kuhn, T. & Rühlemann, C. Exploration of polymetallic nodules and resource assessment: A case study from the German contract area in the Clarion-Clipperton Zone of the Tropical Northeast Pacific. *Minerals* **11**, 618 (2021).
75. Ellefmo, S. L. & Kuhn, T. Application of soft data in nodule resource estimation. *Nat. Resour. Res.* **30**, 1069–1091 (2020).
76. Gazis, I.-Z., Charlet, F. & Greinert, J. An interpretable multi-model machine learning approach for spatial mapping of deep-sea polymetallic nodule occurrences. *Nat. Resour. Res.* **33**, 2473–2501 (2024).
77. Guillén, J., Palanques, A., Puig, P., Durrieu de Madron, X. & Nyffeler, F. Field calibration of optical sensors for measuring suspended sediment concentration in the western Mediterranean. *Sci. Mar.* **64**, 427–435 (2000).
78. The MathWorks Inc. Image Processing toolbox (R2016a), Natick, Massachusetts: The MathWorks Inc. (2016) <https://de.mathworks.com/help/doc-archives.html> (2023).
79. Iversen, M. H., Nowald, N., Ploug, H., Jackson, G. A. & Fischer, G. High resolution profiles of vertical particulate organic matter export off Cape Blanc, Mauritania: Degradation processes and ballasting effects. *Deep Sea Res. Part 1 Oceanogr. Res. Pap.* **57**, 771–784 (2010).
80. Markussen, T. N. Parchar – characterization of suspended particles through image processing in matlab. *J. Open Res. Softw.* **4**, 1–7 (2016).
81. Zeileis, A. & Grothendieck, G. zoo: S3Infrastructure for Regular and Irregular Time Series. *J. Stat. Softw.* **14**, 1–27 (2005).
82. Manik, H. M., Gultom, D. A., Firdaus & Elson, L. Evaluation of ADCP backscatter computation for quantifying suspended sediment concentration. *IOP Conf. Ser. Earth Environ. Sci.* **429**, 012035 (2020).
83. Gardner, W. D., Sullivan, L. G. & Thorndike, E. M. Long-term photographic, current, and nephelometer observations of manganese nodule environments in the Pacific. *Earth Planet. Sci. Lett.* **70**, 95–109 (1984).
84. Ha, H. K., Maa, J. P.-Y., Park, K. & Kim, Y. H. Estimation of high-resolution sediment concentration profiles in bottom boundary layer using pulse-coherent acoustic Doppler current profilers. *Mar. Geol.* **279**, 199–209 (2011).

85. Made with Natural Earth. Free vector and raster map data @ [naturalearthdata.com](http://naturalearthdata.com)
86. International Seabed Authority, Maps: PMN Exploration Areas, PMN Reserved Areas, APEI Areas. <https://www.isa.org/jm/exploration-contracts/maps/> (2023).
87. Flanders Marine Institute. Maritime Boundaries Geodatabase, version 11, <https://www.marineregions.org/> (2019).
88. Simon-Lledó, E. et al. Carbonate compensation depth drives abyssal biogeography in the northeast Pacific. *Nat. Ecol. Evol.* **7**, 1388–1397 (2023).
89. Savoca, S., Di Fresco, D., Alesci, A., Capillo, G. & Spanò, N. Mucus secretions in cnidarian, an ecological, adaptive and Evolutionary Tool. *Adv. Oceanogr. Limnol.* **13**, <https://doi.org/10.4081/aiol.2022.11054> (2022).
90. Baas, J. H., Best, J. L. & Peakall, J. Predicting bedforms and primary current stratification in cohesive mixtures of mud and sand. *J. Geol. Soc. London* **173**, 12–45 (2015).

## Acknowledgements

We gratefully acknowledge the captain and crew of MV Island Pride for their assistance during the MANGAN 2021 expedition. We thank the Ocean Infinity ROV and AUV teams for extensive dives. We are grateful to the German Federal Institute for Geosciences and Natural Resources (Bundesanstalt für Geowissenschaften und Rohstoffe - BGR) for inviting MiningImpact 2 scientists to join their research cruise on MV Island Pride and for chartering both ROVs. We are thankful to the Global Sea Mineral Resources NV (GSR) / DEME Group, who chartered the AUV onboard MV *Island Pride*. We express our gratitude to Francois Charlet (GSR/ DEME Group) for his support during the MANGAN 2021 expedition, particularly for the communication between the GSR team onboard MV Normand Energy and our team onboard MV Island Pride. Special thanks go to the Oceanic Machine Vision Group at GEOMAR for their valuable recommendations on the AUV photomosaic surveys. We thank the GEOMAR Library team for its support in gathering the bibliography and Anastasios Poulos-Sidiropoulos for his assistance with the MBES and SSS data post-processing. This research was carried out in the European collaborative project MiningImpact 2 and received national funding through the Joint Programming Initiative Healthy and Productive Seas and Oceans (JPI Oceans): German Ministry of Research grant no. 03F0812A-H; Dutch Research Council grant no. 856.18.002. Additional funds for representing the data within the 4D Digital Earth Viewer came through the Helmholtz Project “Digital Earth” grant ZT-0025. M.D. acknowledges funding through the Plumefloc project; Dutch Research Council grant no. TWM.BL.019.004. This is publication # 68 of the DeepSea Monitoring Group at GEOMAR Helmholtz Centre for Ocean Research Kiel.

## Author contributions

All authors collaboratively contributed to the conception and design of the study. H.S., M.H., A.V., M.B., B.G., L.T. and J.G. offered scientific equipment. I.Z.G., H.S., J.M., K.H., M.H. and A.V. acquired the data. I.Z.G., H.S., J.M., K.H., M.D., B.G., L.T., M.B. and M.V.A. contributed to the data

post-processing. I.Z.G., M.D. and M.V.A. collaboratively contributed to the data visualisation. J.M., K.H., L.T. and A.V. collaboratively contributed to the creation of video data. I.Z.G., H.S., M.V.A. and K.P. collaboratively contributed to the sediment transport estimation. M.H. acted as EU MiningImpact 2 project leader. L.T. and J.G. acted as EU work-package leaders for this monitoring mission, and L.T. was significantly involved in finalising the manuscript. I.Z.G. compiled the results and wrote the manuscript. I.Z.G., L.T., H.S. and J.G. revised the manuscript. All co-authors commented on the manuscript and provided input to its final version.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-56311-0>.

**Correspondence** and requests for materials should be addressed to Iason-Zois Gazis.

**Peer review information** *Nature Communications* thanks Zhihan Fan, Rudy Helmons and Isobel Yeo for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

## Supplementary Information

### Monitoring benthic plumes, sediment redeposition and seafloor imprints caused by deep-sea polymetallic nodule mining

Iason-Zois Gazis<sup>1\*</sup>, Henko de Stigter<sup>2</sup>, Jochen Mohrmann<sup>1</sup>, Karl Heger<sup>1</sup>, Melanie Diaz<sup>2</sup>, Benjamin Gillard<sup>3</sup>, Matthias Baeye<sup>4</sup>, Mario E. Veloso-Alarcón<sup>1</sup>, Kaveh Purkiani<sup>1,5,6</sup>, Matthias Haeckel<sup>1</sup>, Annemiek Vink<sup>7</sup>, Laurenz Thomsen<sup>3,8</sup>, Jens Greinert<sup>1,9</sup>

<sup>1</sup> GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany

<sup>2</sup> NIOZ Royal Netherlands Institute for Sea Research, Texel, The Netherlands

<sup>3</sup> Constructor University, Bremen, Germany

<sup>4</sup> Royal Belgian Institute of Natural Sciences, Brussels, Belgium

<sup>5</sup> MARUM Center for Marine Environmental Sciences and Faculty of Geosciences, University of Bremen, Bremen, Germany

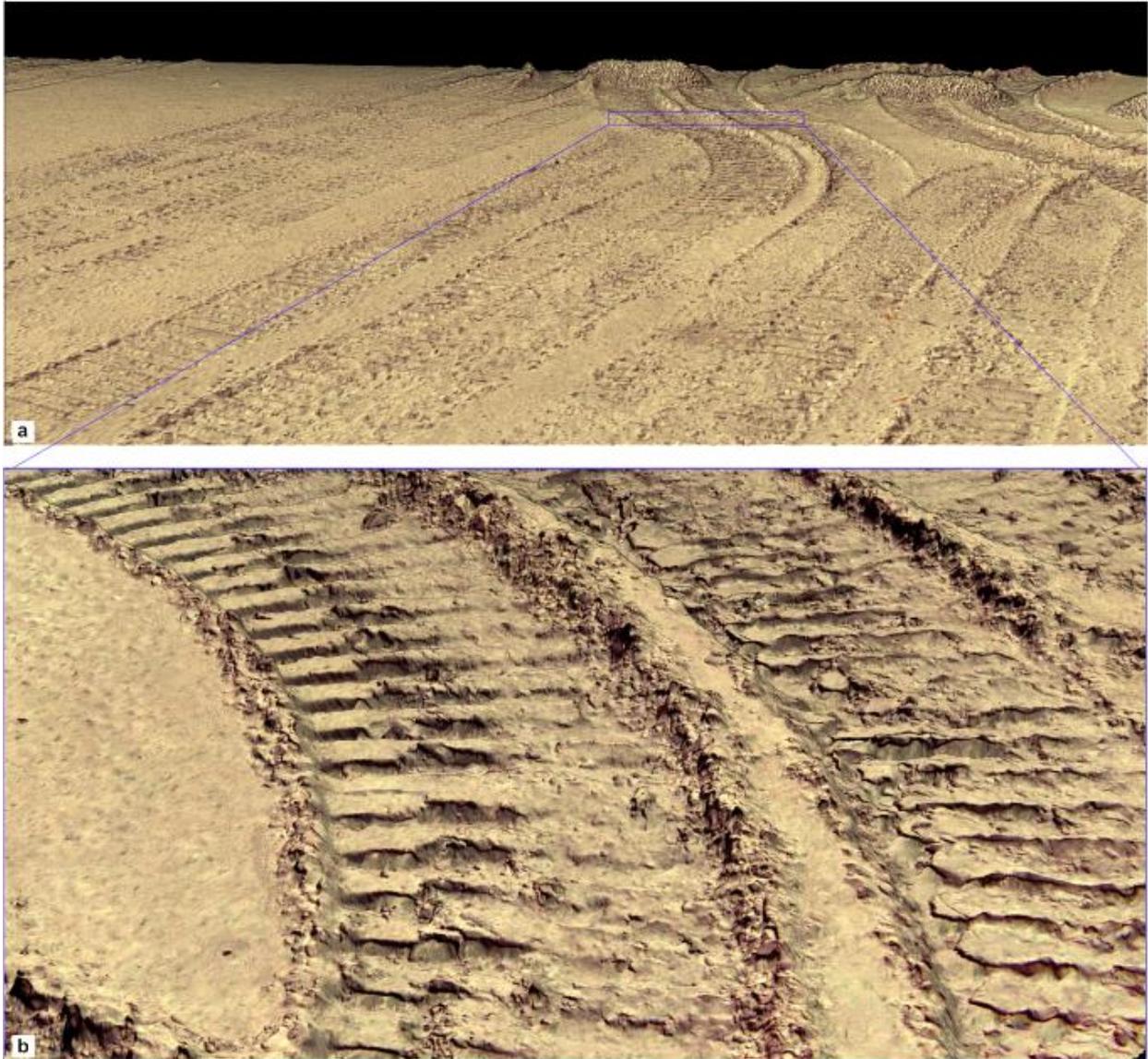
<sup>6</sup> Now at Federal Maritime and Hydrographic Agency, Hamburg, Germany

<sup>7</sup> Federal Institute for Geosciences and Natural Resources, Hannover, Germany

<sup>8</sup> Now at the University of Gothenburg, Sweden

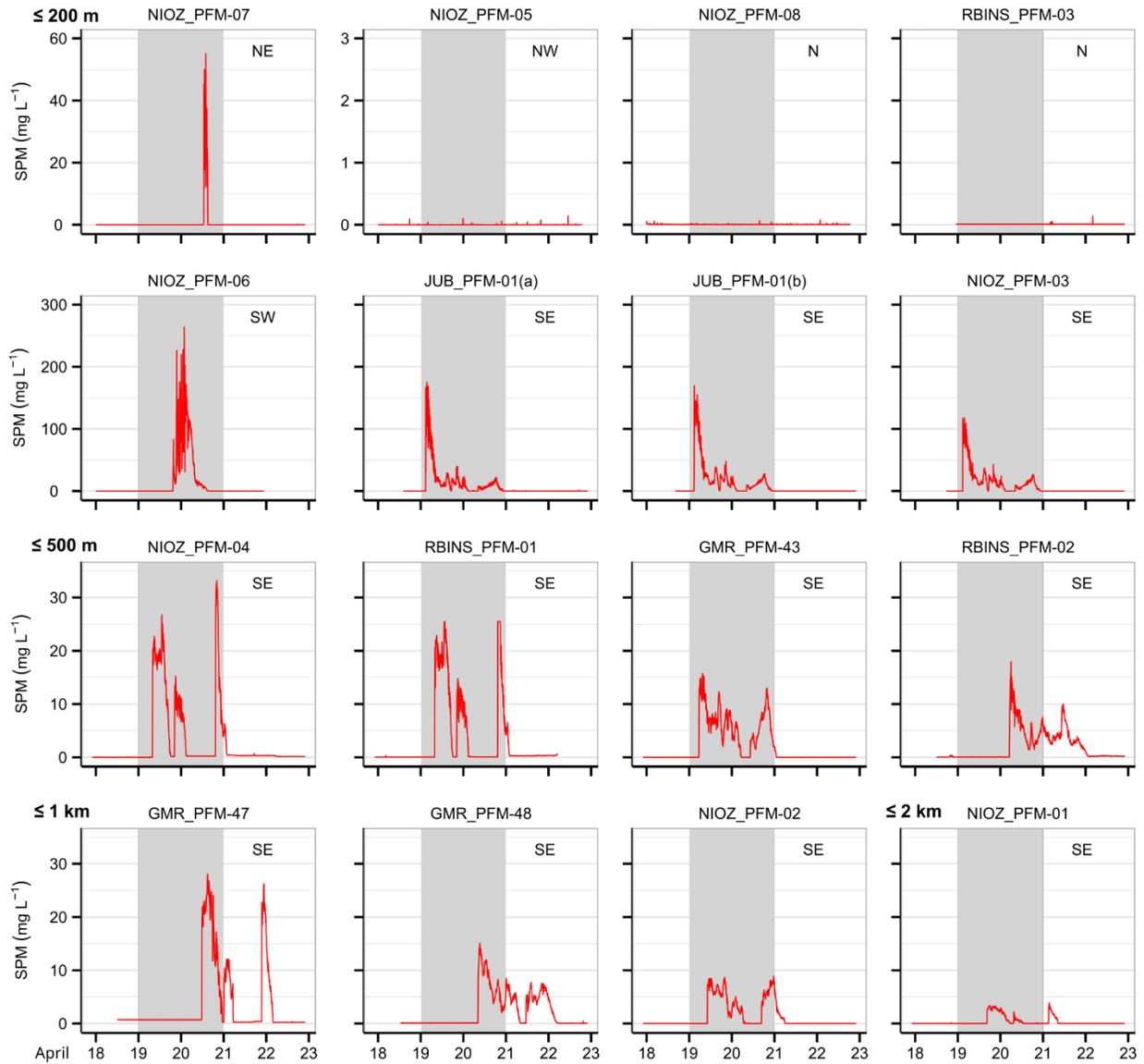
<sup>9</sup> Institute of Geosciences, Christian-Albrecht University of Kiel, Germany

\* Corresponding author: [igazis@geomar.de](mailto:igazis@geomar.de)



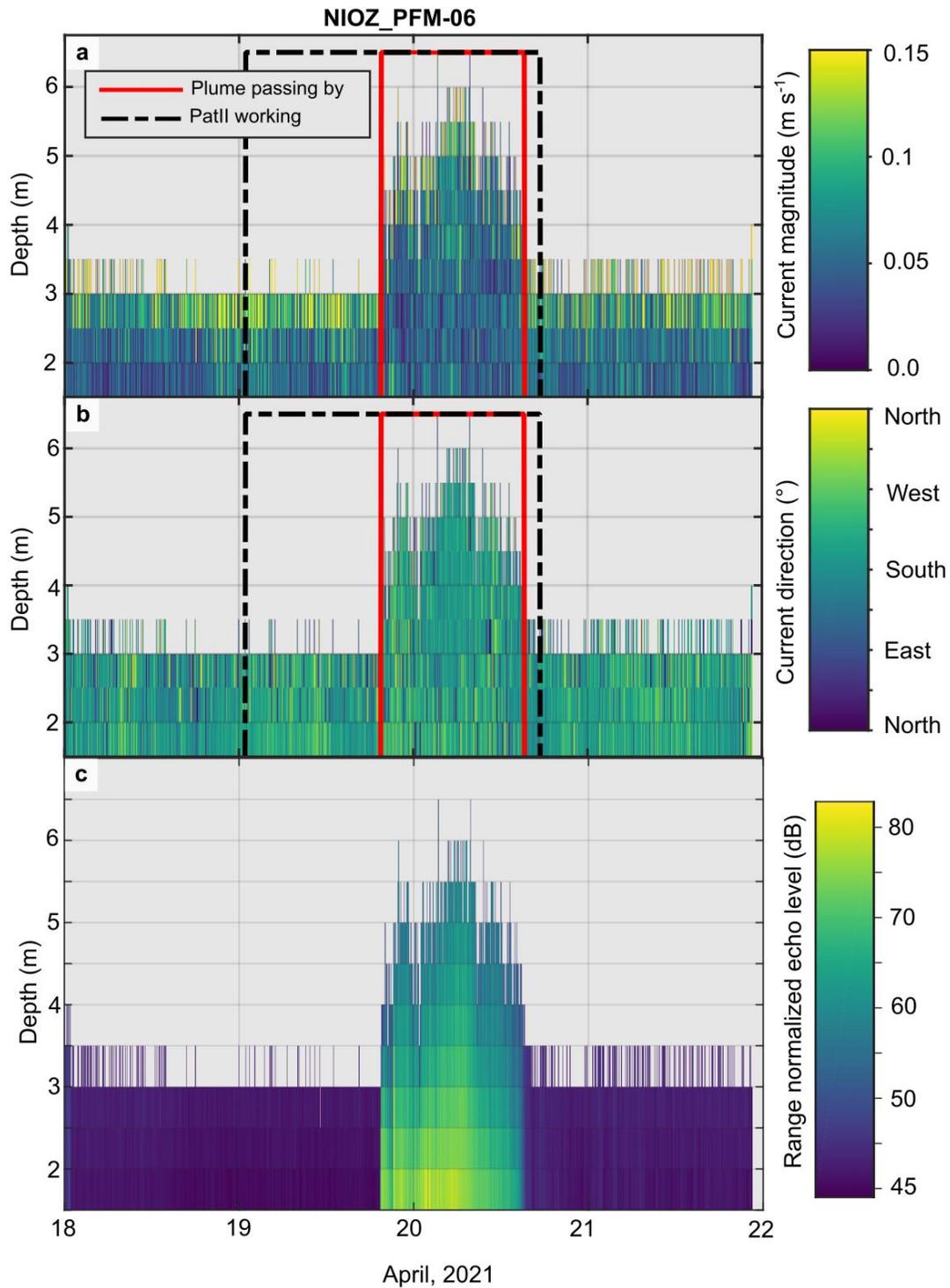
**Fig: S1 | 3D orthophoto-mosaic of the nodule collector trial site.**

**a)** 3D perspective view from an oblique angle of the caterpillar track imprints along the mining lanes (part of the second strip, as presented in Fig. 2). The redeposited sediment covers the unmined seafloor between the mining lanes and only the humps of the unmined nodules are partly visible. **b)** Close-up of the small-scale seafloor morphology (cm scale) created by the nodule collector caterpillar tracks along the "light bulb" turning maneuver shortly before the deposition of the nodule pile at the line end. These imprints are locally deeper (up to 8 cm) than those made along the mining lanes due to the additional nodule payload (~ 3 t). The two caterpillar imprints have a total length of 4 m and average depth of 5 cm.



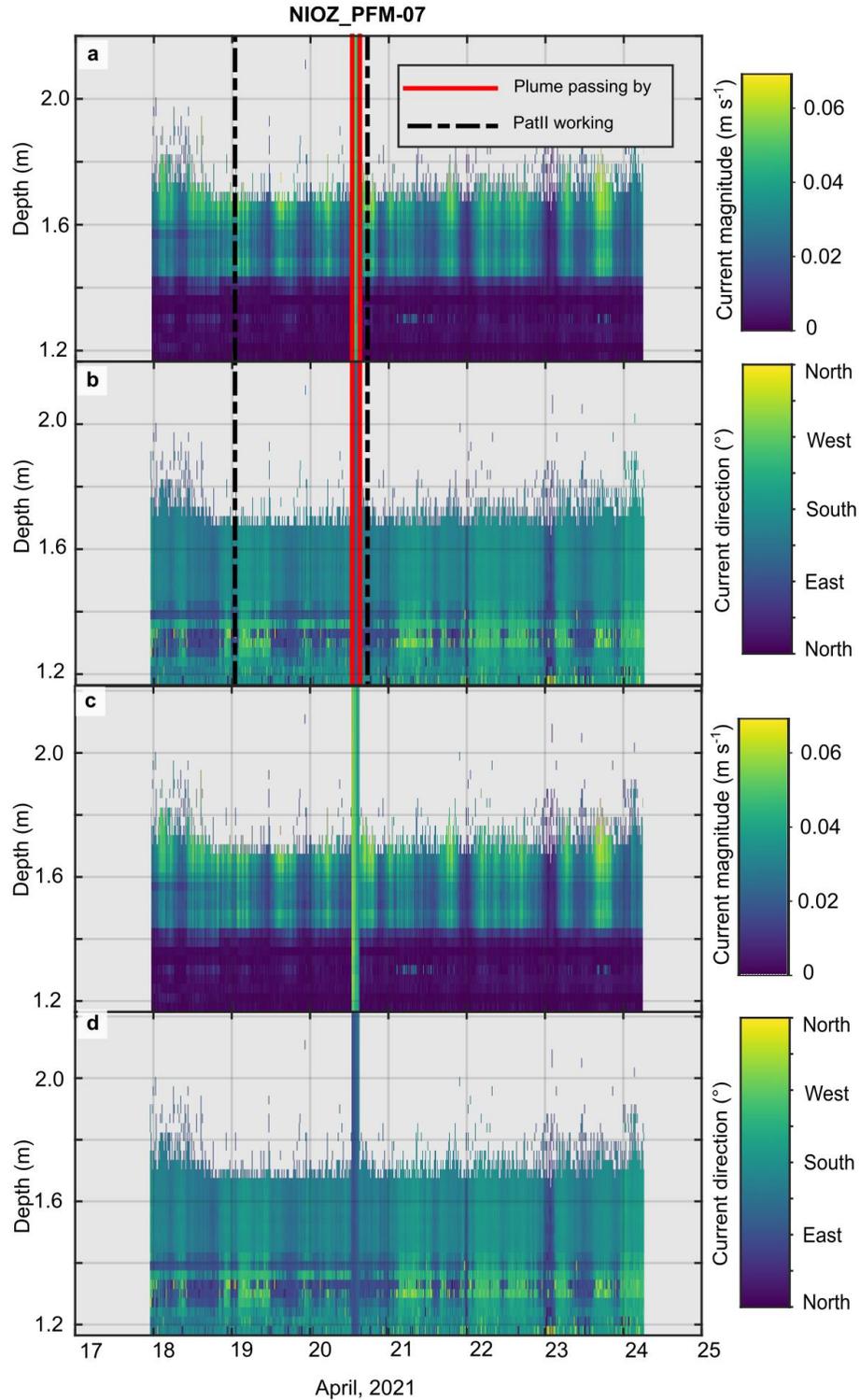
**Fig. S2 | Time-series of suspended particulate matter concentration**

The optical backscatter sensors (OBS) recorded the suspended particulate matter (SPM) concentrations 1 m above the seafloor. The seafloor platforms have been grouped by distance and orientation to the impact site. The grey sections indicate the nodule collector operation time (from 19 April 2021 00:52 UTC to 20 April 2021 17:25 UTC).



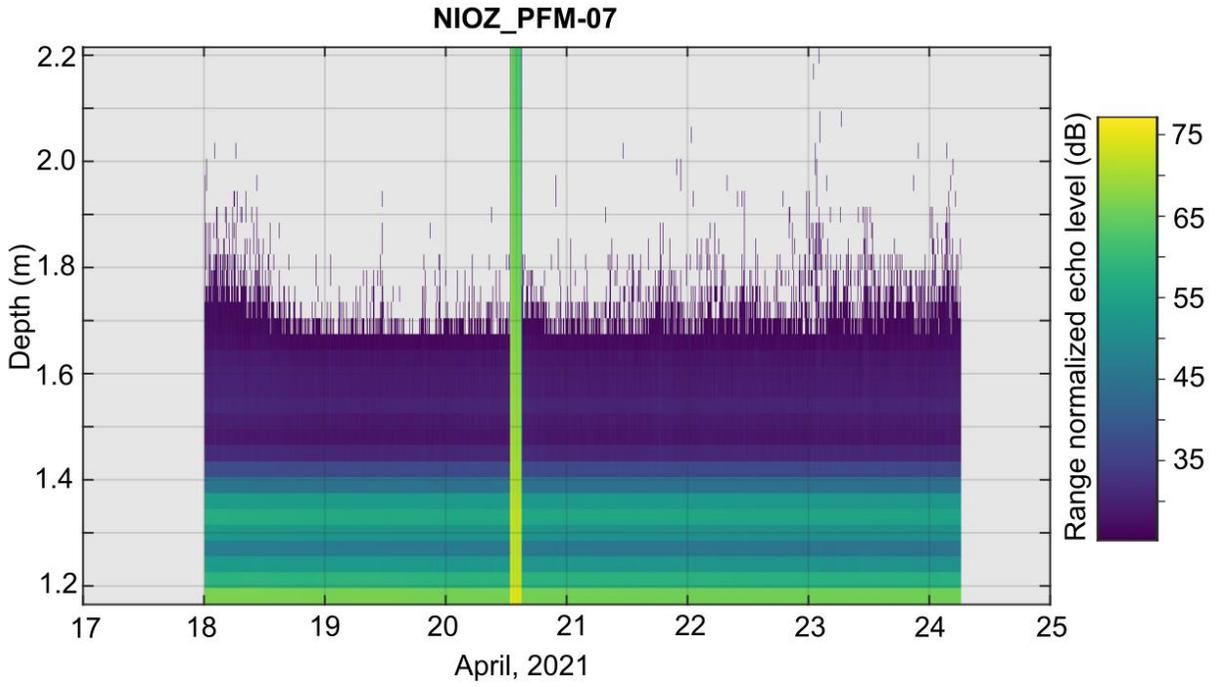
**Fig. S3 | Acoustic doppler current profiler (ADCP) data at the NIOZ\_PFM-06 platform.**

**a)** Current magnitude, **b)** current direction and **c)** range normalized backscatter intensity at the NIOZ\_PFM-06 platform between 18 and 22 April 2021. The nodule collector (PatII) operation time and the passage of the benthic sediment plume are indicated in Figures a and b.



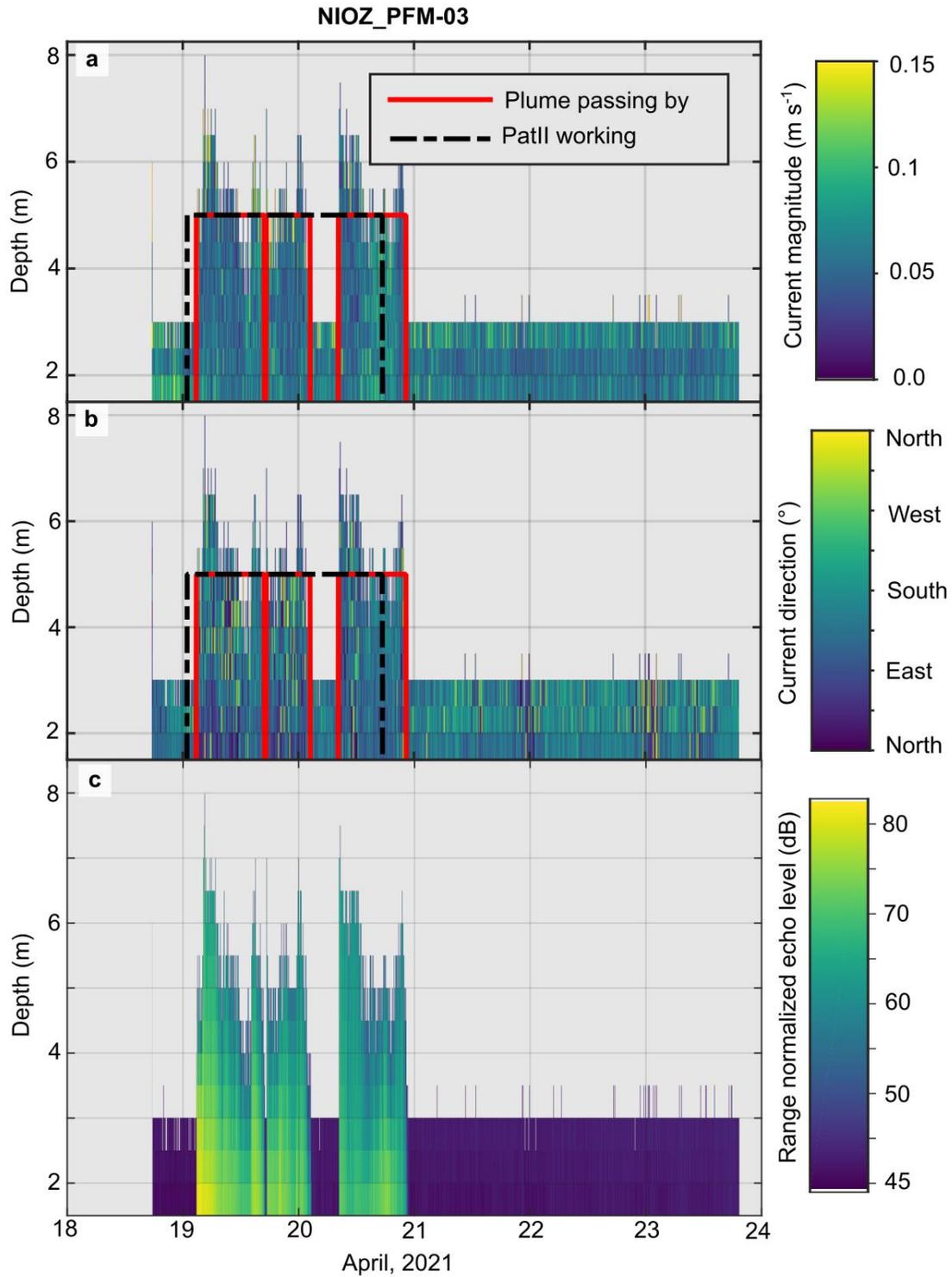
**Fig. S4 | Acoustic doppler current profiler (ADCP) data at the NIOZ\_PFM-07 platform.**

Current magnitude, **b**) current direction and **c**) range normalized backscatter intensity at the NIOZ\_PFM-07 platform between 18 and 25 April 2021, with (a and b) and without (c and d) the nodule collector (PatII) activity and benthic sediment plume passage time. The gravity current created a distinct change in current speed and direction.



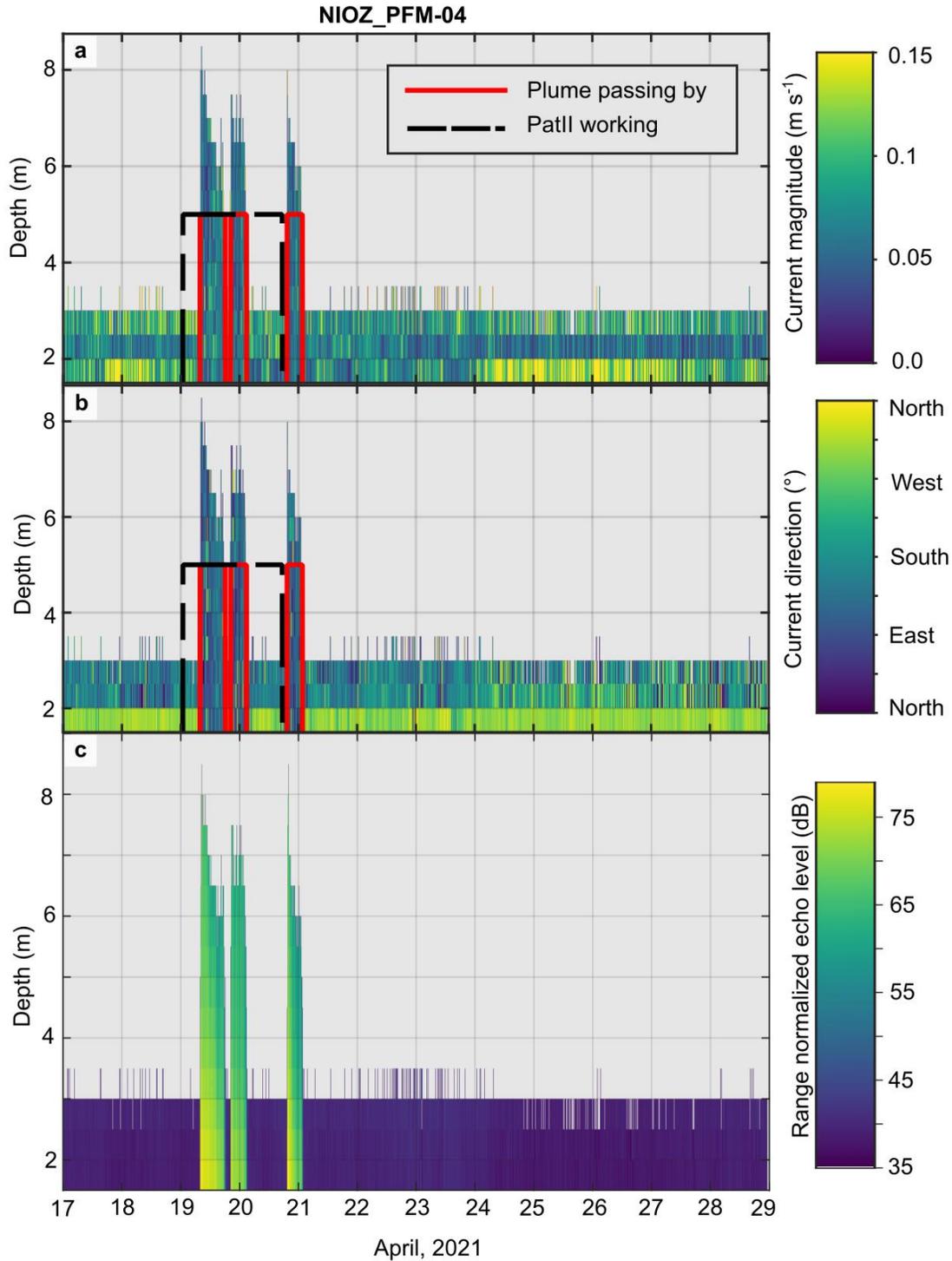
**Fig. S5 | Range normalized water-column backscatter at the NIOZ\_PFM-07 platform.**

Range normalized water-column backscatter intensity from the upward-looking ADV (2MHz) at the NIOZ\_PFM-07 platform between 18 and 25 April 2021, showing the gravity current's passage.



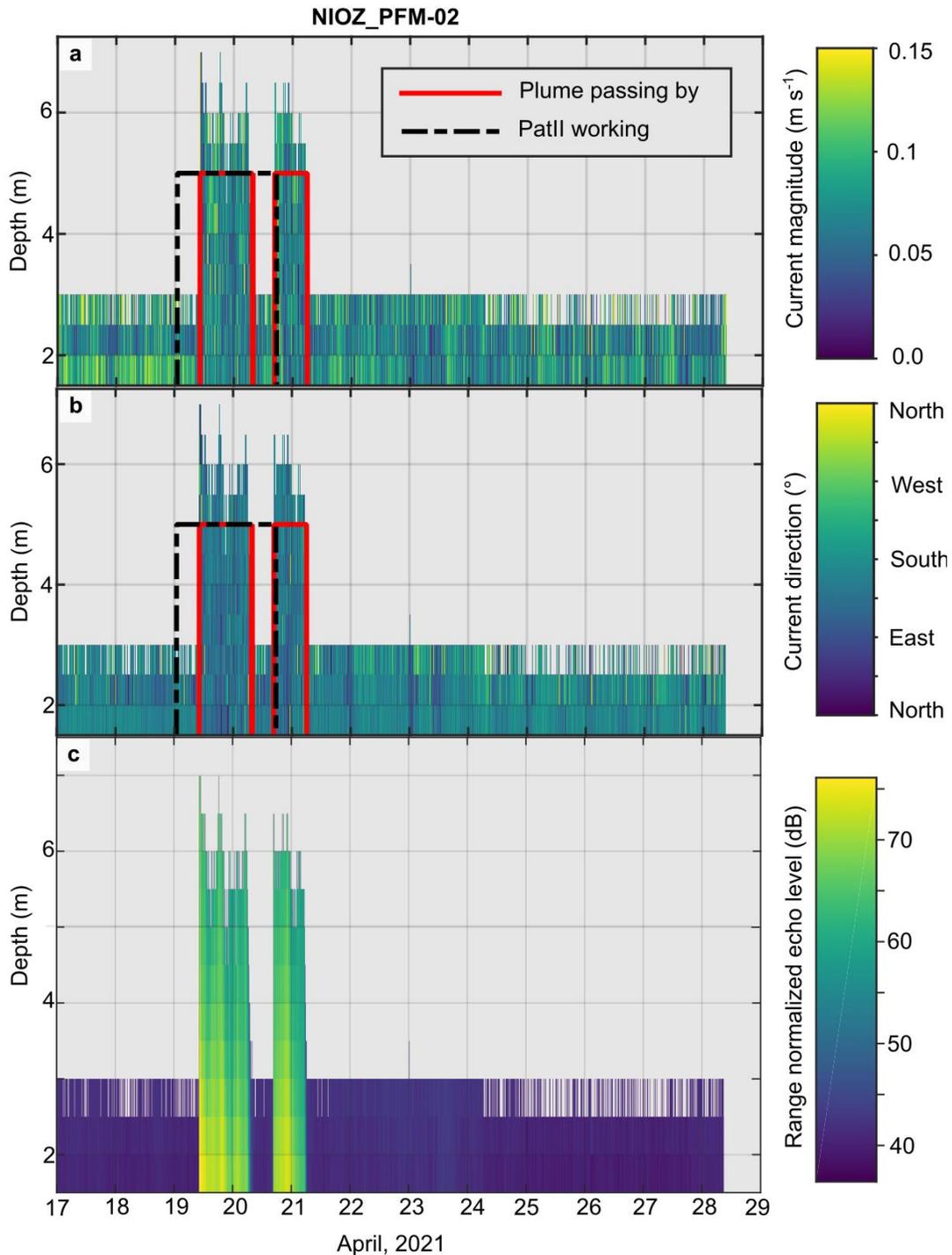
**Fig. S6 | Acoustic doppler current profiler (ADCP) data at the NIOZ\_PFM-03 platform.**

**a)** Current magnitude, **b)** current direction, and **c)** range normalized backscatter intensity from the ADCP (2 MHz) at the NIOZ\_PFM-03 platform between 18 April and 24 April 2021. The nodule collector (PatII) operation time and the passage of the benthic sediment plume are indicated in Figures a and b.



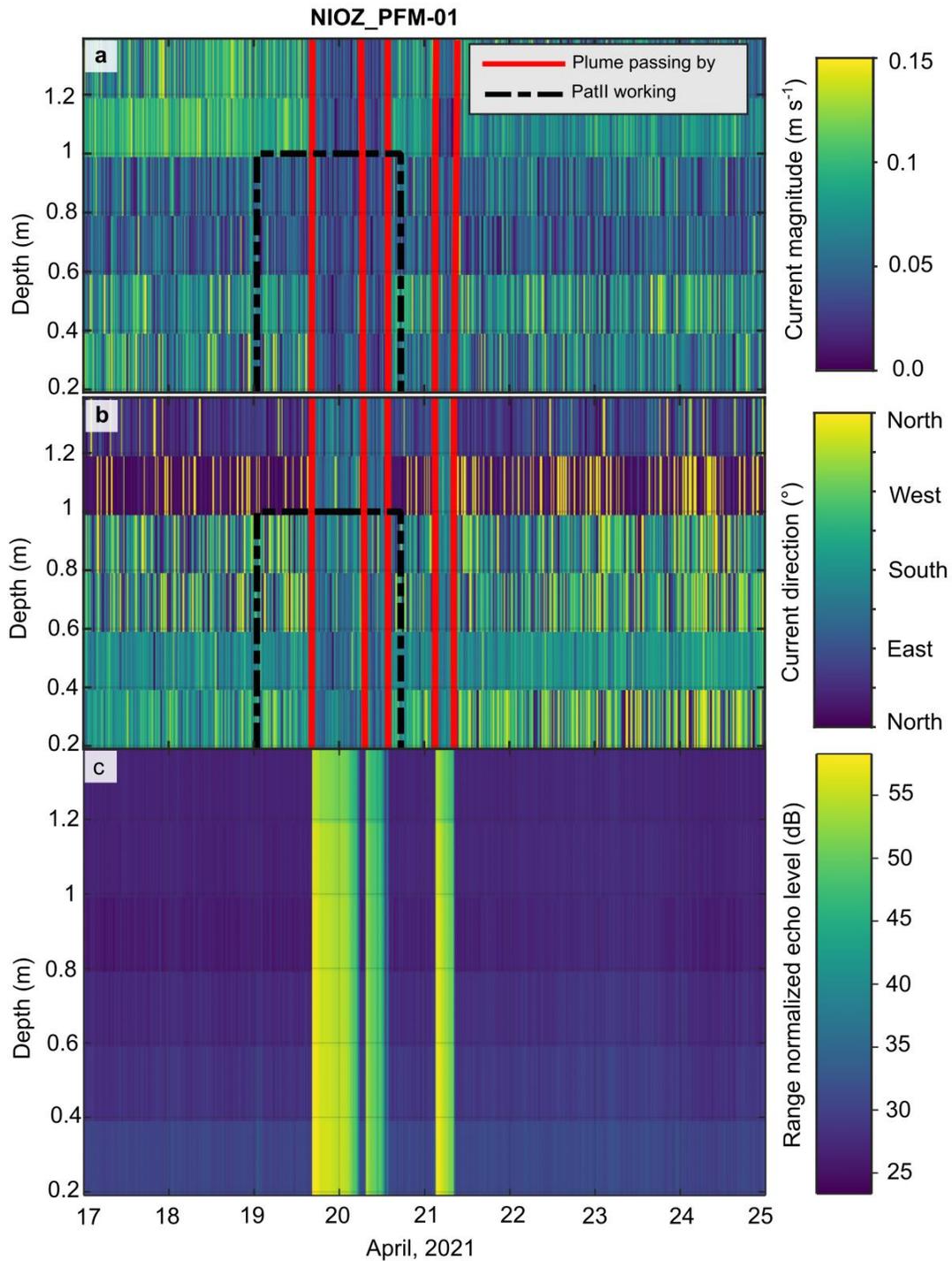
**Fig. S7 | Acoustic doppler current profiler (ADCP) data at the NIOZ\_PFM-04 platform.**

**a)** Current magnitude, **b)** current direction and **c)** range normalized backscatter intensity from the upward-looking ADCP (2 MHz) at the NIOZ\_PFM-04 platform between 17 and 29 April 2021. The ADCP at platform NIOZ\_PFM-04 recorded the most distant change in the current direction when the benthic sediment plume passed. The nodule collector (PatII) operation time and the passage of the benthic sediment plume are indicated in Figures a and b.



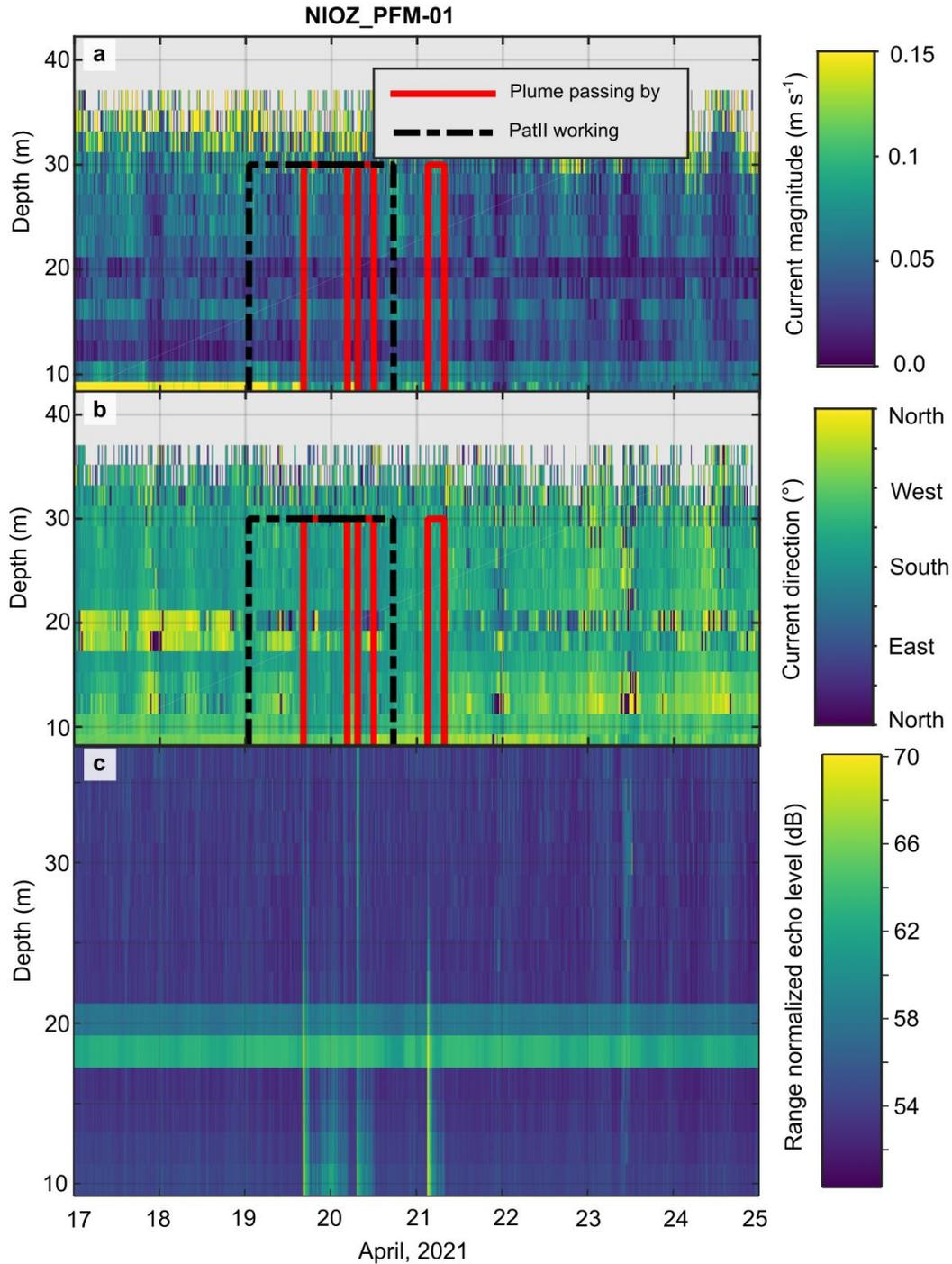
**Fig. S8 | Acoustic doppler current profiler (ADCP) data at the NIOZ\_PFM-02 platform.**

**a)** Current magnitude, **b)** current direction and **c)** range normalized backscatter intensity from the upward-looking ADCP (2 MHz) at the NIOZ\_PFM-02 platform between 17 and 29 April 2021. At 1 km SE from the impact site, the benthic sediment plume has the same direction as the ambient bottom currents. The nodule collector (PatII) operation time and the passage of the benthic sediment plume are indicated in Figures a and b.



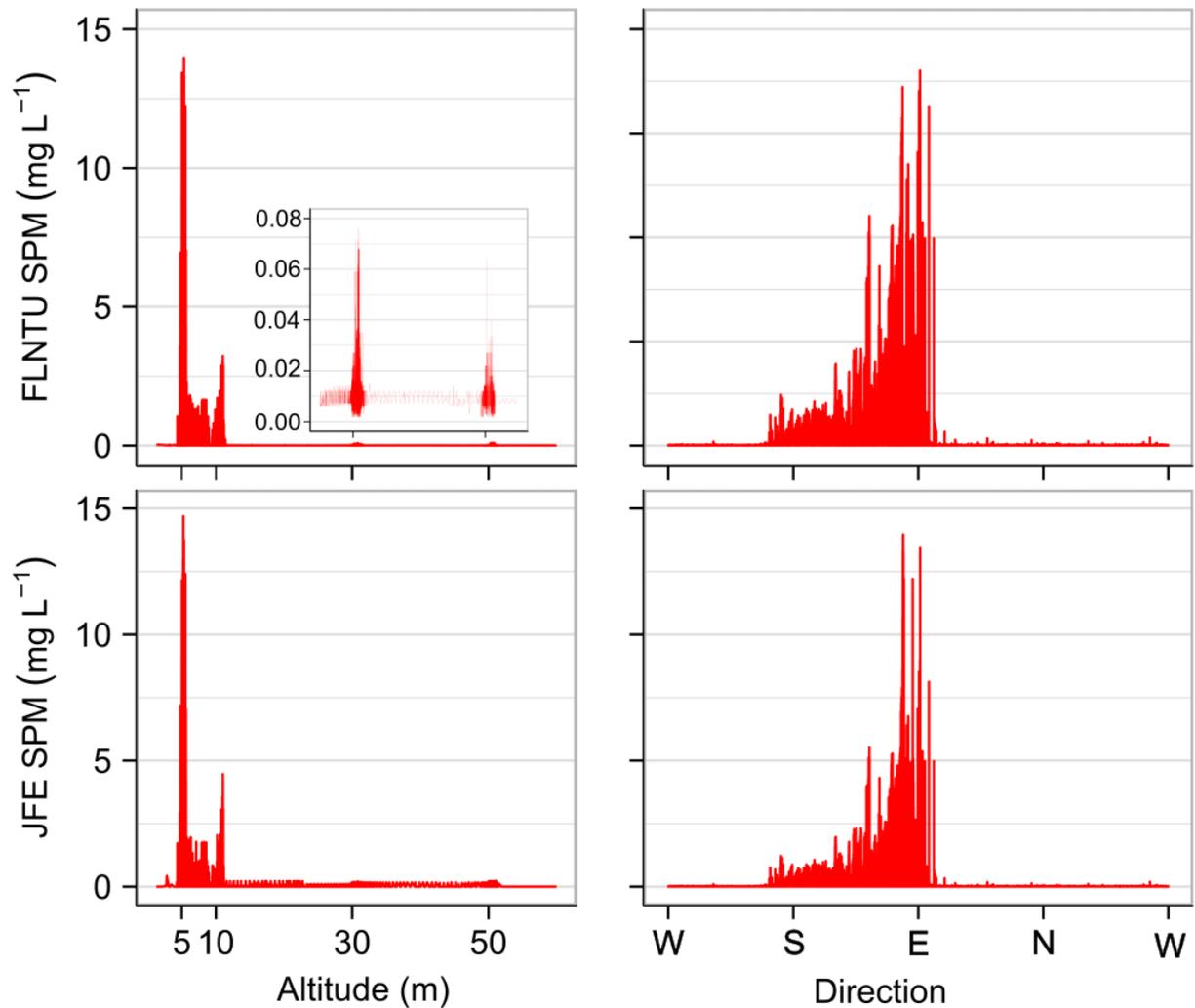
**Fig. S9 | Acoustic doppler current profiler (ADCP) downward-looking data at the NIOZ\_PFM-01 platform.**

**a)** Current magnitude, **b)** current direction and **c)** range normalized backscatter intensity from the downward-looking ADCP (1.2 MHz) at the NIOZ\_PFM-01 platform between 17 April and 25 April 2021. The most remote downward-looking ADCP showed that the benthic sediment plume moves in the SE direction. The nodule collector (PatII) operation time and the passage of the benthic sediment plume are indicated in Figures a and b.



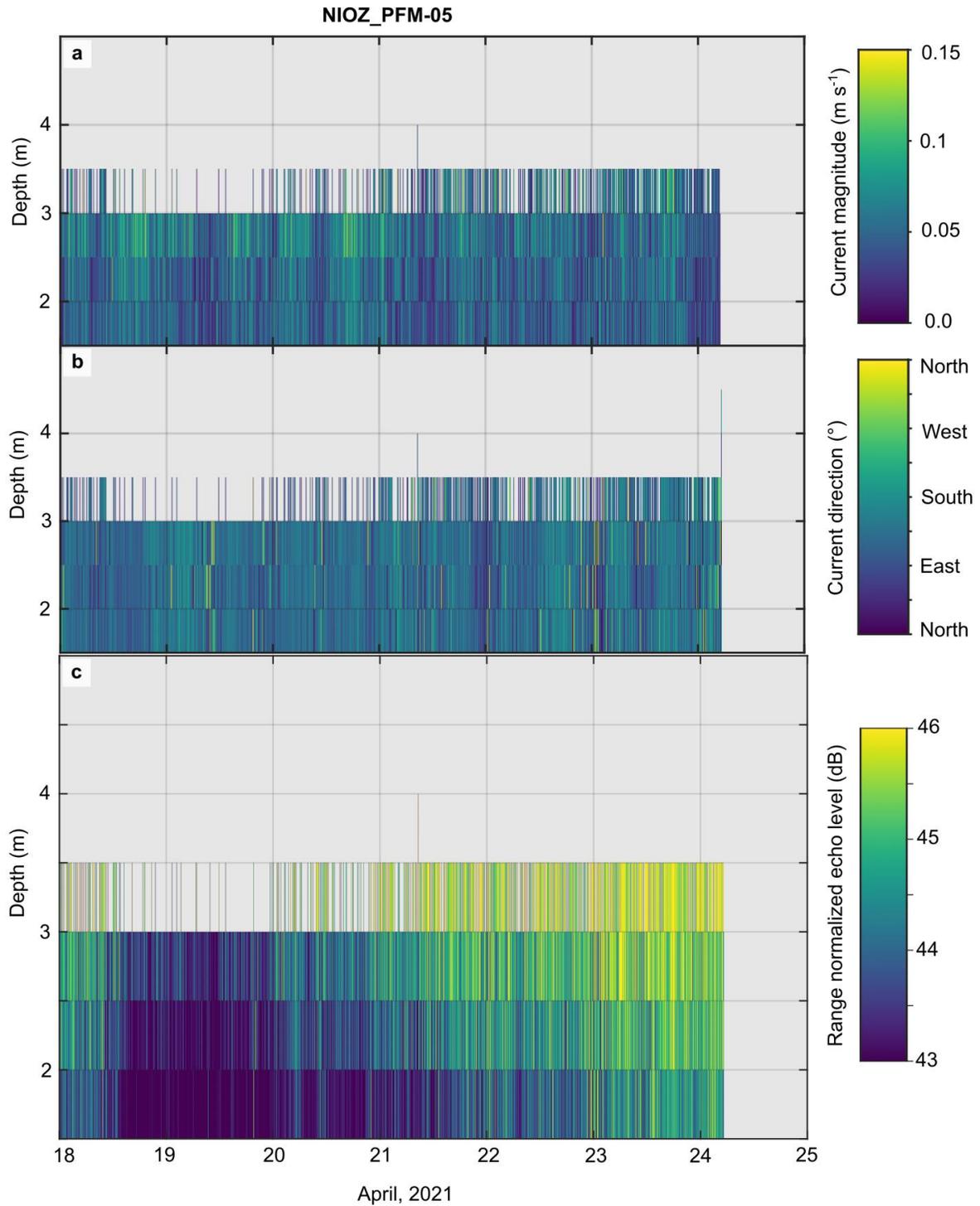
**Fig. S10 | Acoustic doppler current profiler (ADCP) upward-looking data at the NIOZ\_PFM-01 platform.**

a) Current magnitude, b) current direction and c) range normalized backscatter intensity from the upward-looking ADCP (300 kHz) at the NIOZ\_PFM-01 platform between 17 April and 25<sup>th</sup> April 2021. The most remote upward-looking ADCP at 1800m SE from the impact site recorded the suspended particles at 30+ m altitude. The nodule collector (PatII) operation time and the passage of the benthic sediment plume are indicated in Figures a and b.



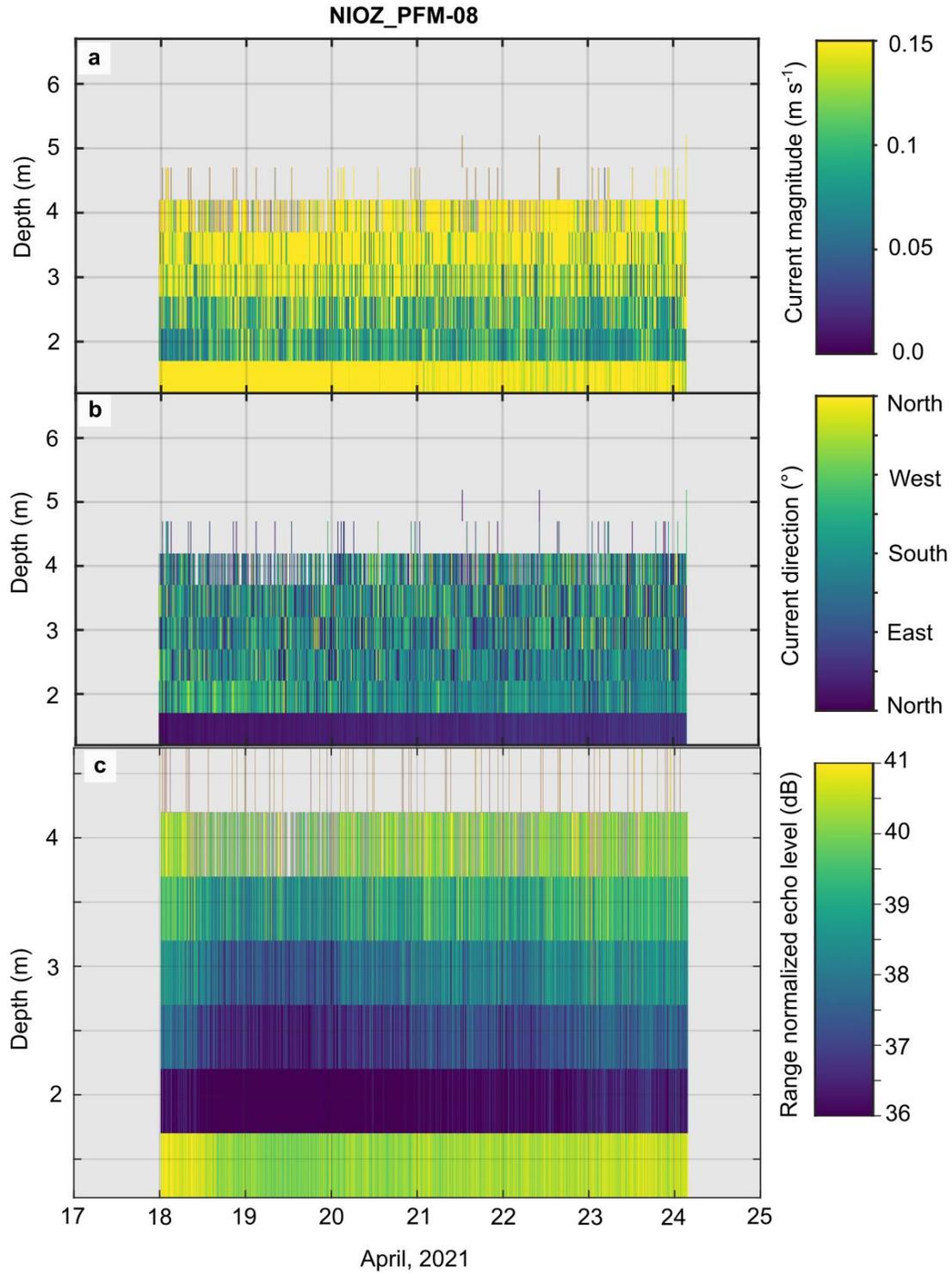
**Fig. S11 | Suspended particulate matter (SPM) concentration recorded by the autonomous underwater vehicle (AUV)**

The AUV conducted a ca. 54 h survey parallel to the mining trial and immediately after recording the SPM concentration with altitude and direction with two different optical backscatter sensors (OBS), namely FLNTU and JFE (see Methods). **a-b)** FLNTU SPM concentration for different monitoring altitudes and directions, **c-d)** JFE SPM concentration for different monitoring altitudes and directions. The data clearly show that the benthic sediment plume spread in the S to E direction and that the highest concentrations were registered at 5 m, which was the lowest monitoring altitude by the AUV. Even the highest peak in SPM concentration is less than half of the SPM concentration recorded at 1 m altitude at the same distance and direction from the impact site (Fig. S29).



**Fig. S12 | Acoustic doppler current profiler (ADCP) data at the NIOZ\_PFM-05 platform.**

**a)** Current magnitude, **b)** current direction, and **c)** range normalized backscatter intensity from the upward-looking ADV (2 MHz) at the NIOZ\_PFM-05 platform between 18 April and 25 April 2021. The benthic sediment plume was not detected NW of the impact site.

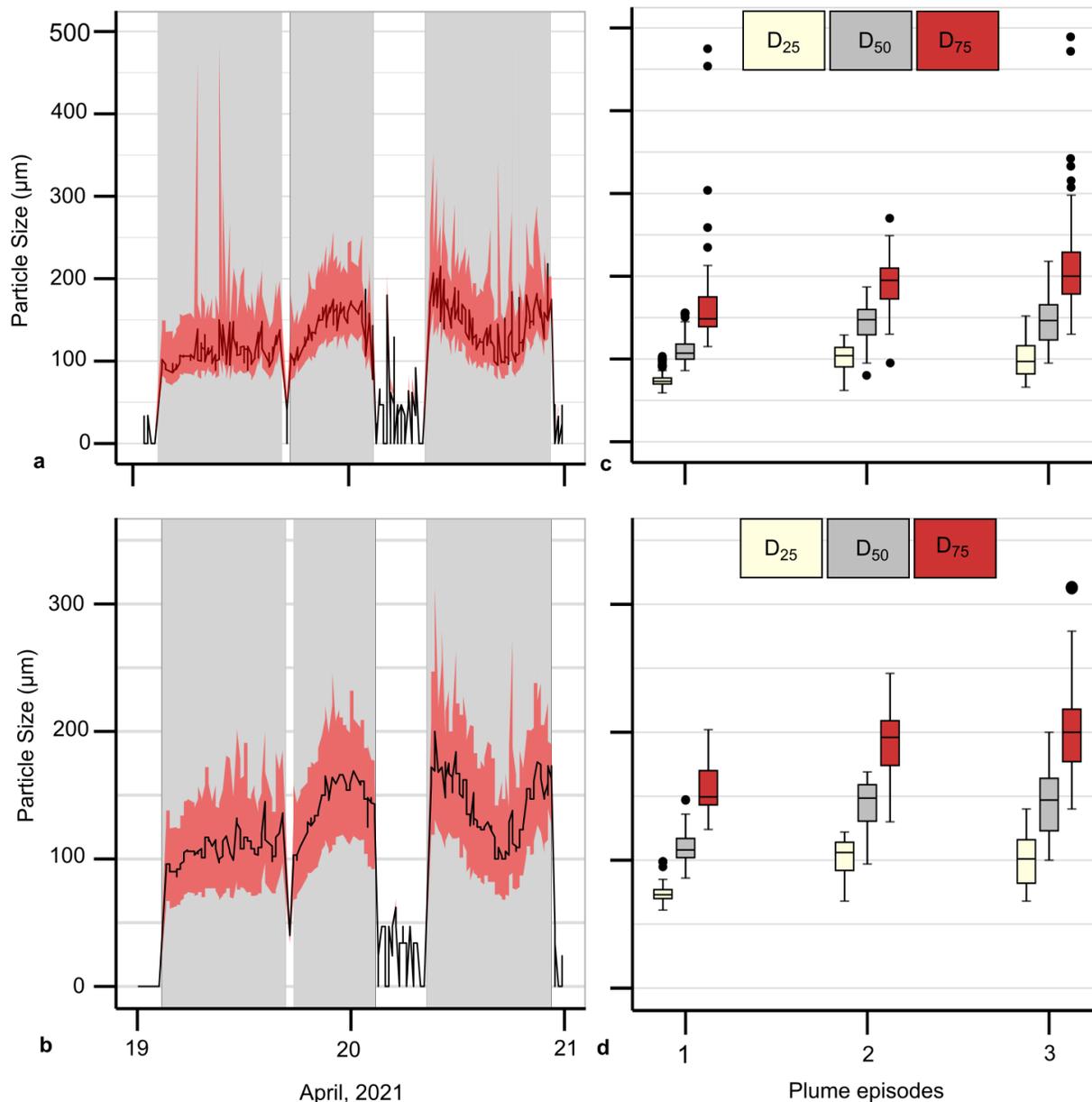


**Fig. S13 | Acoustic doppler current profiler (ADCP) data at the NIOZ\_PFM-08 platform.**

**a)** Current magnitude, **b)** current direction, and **c)** range normalized backscatter intensity from the upward-looking ADV (2 MHz) at the NIOZ\_PFM-08 platform between 18 April and 24 April 2021. The benthic sediment plume was not detected at the N of the impact site.

**Table S1** | Summary statistics of the particle size ( $D_{25}$ ,  $D_{50}$ ,  $D_{75}$ ) for each benthic sediment plume episode (plume).

1 <sup>st</sup> Plume Episode	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
$D_{25}$ (raw)	59.0	69.7	73.0	74.7	77.0	101.0
$D_{25}$ (after roll median)	61.0	70.0	73.0	74.2	77.0	97.0
$D_{50}$ (raw)	86.0	99.7	107.0	110.5	118.0	150.00
$D_{50}$ (after roll median)	86.0	102.0	108.0	109.3	117.0	145.00
$D_{75}$ (raw)	115.0	139.0	148.5	164.9	175.0	473.0
$D_{75}$ (after roll median)	124.00	143.2	149.5	155.7	170.0	202.0
2 <sup>nd</sup> Plume Episode	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
$D_{25}$ (raw)	62.0	90.5	104.0	100.6	114.0	129.0
$D_{25}$ (after roll median)	68.0	92.5	106.0	101.0	114.0	122.0
$D_{50}$ (raw)	78.0	129.2	147.5	142.9	159.8	187.0
$D_{50}$ (after roll median)	97.0	130.5	148.5	143.5	159.0	169.0
$D_{75}$ (raw)	93.0	172.5	195.0	190.6	209.8	267.0
$D_{75}$ (after roll median)	130.0	174.0	195.0	189.8	209.0	246.0
3 <sup>rd</sup> Plume Episode	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
$D_{25}$ (raw)	66.0	82.0	97.0	98.6	116.0	152.0
$D_{25}$ (after roll median)	68.0	82.0	101.0	98.5	116.0	140.0
$D_{50}$ (raw)	95.0	123.0	146.5	144.6	165.5	218.0
$D_{50}$ (after roll median)	100.0	123.0	147.0	142.0	164.0	200.0
$D_{75}$ (raw)	130	178.5	200.0	210.0	229.0	487.0
$D_{75}$ (after roll median)	140	177.0	200.0	199.6	218.0	313.0



**Fig. S14 | Particle number and size before and after applying a rolling median**

**a)** Time series of particle size (raw data) during the three benthic sediment plume episodes, which are shown in grey. The black line is the  $D_{50}$ , and the red area shows the upper  $D_{75}$  and lower  $D_{25}$  of the particle size distribution. Few isolated outliers are observed in raw data. **b)** Same as a, after applying a rolling median ( $k=3$ , with  $k$  being the integer width of the rolling window). **c)** Boxplots of the particle sizes  $D_{25}$ ,  $D_{50}$  and  $D_{75}$  for the first (1), second (2) and third (3) plume episodes using the raw data ( $n=258$  observations). **d)** Same as c, after applying a rolling median ( $k=3$ ). Both in c and d the black centre line is the median value (50<sup>th</sup> percentile), and the boxes contain each dataset's 25<sup>th</sup> to 75<sup>th</sup> percentiles. The black whiskers mark the 5<sup>th</sup> and 95<sup>th</sup> percentiles, and values beyond these bounds are outliers (black dots).



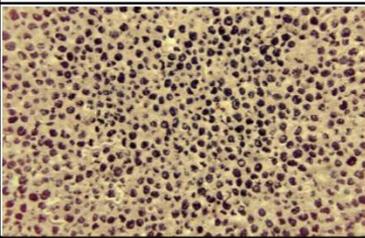
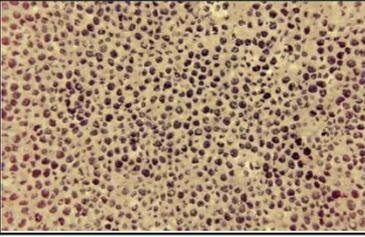
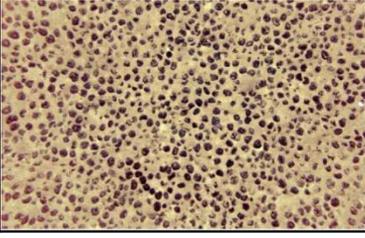
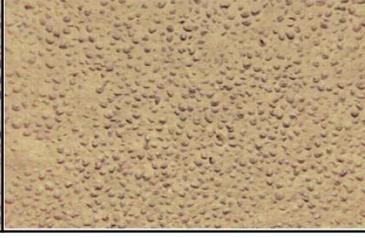
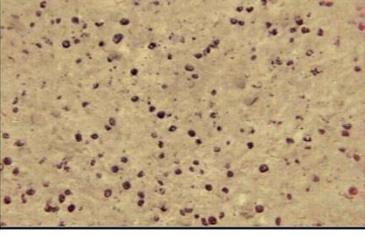
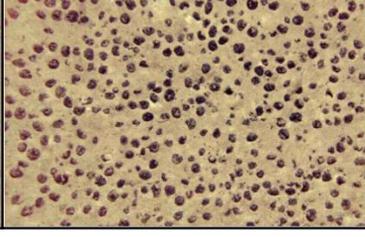
**Fig. S15 | Seafloor image inside the mining site**

A seafloor image obtained by the remotely operated vehicle 5 m from a mining lane shows that redeposited sediment completely covers nodules (sediment blanketing). A sea anemone (*Actiniaria*) was recorded on the redeposited sediment. The image location is shown in Fig. 10.



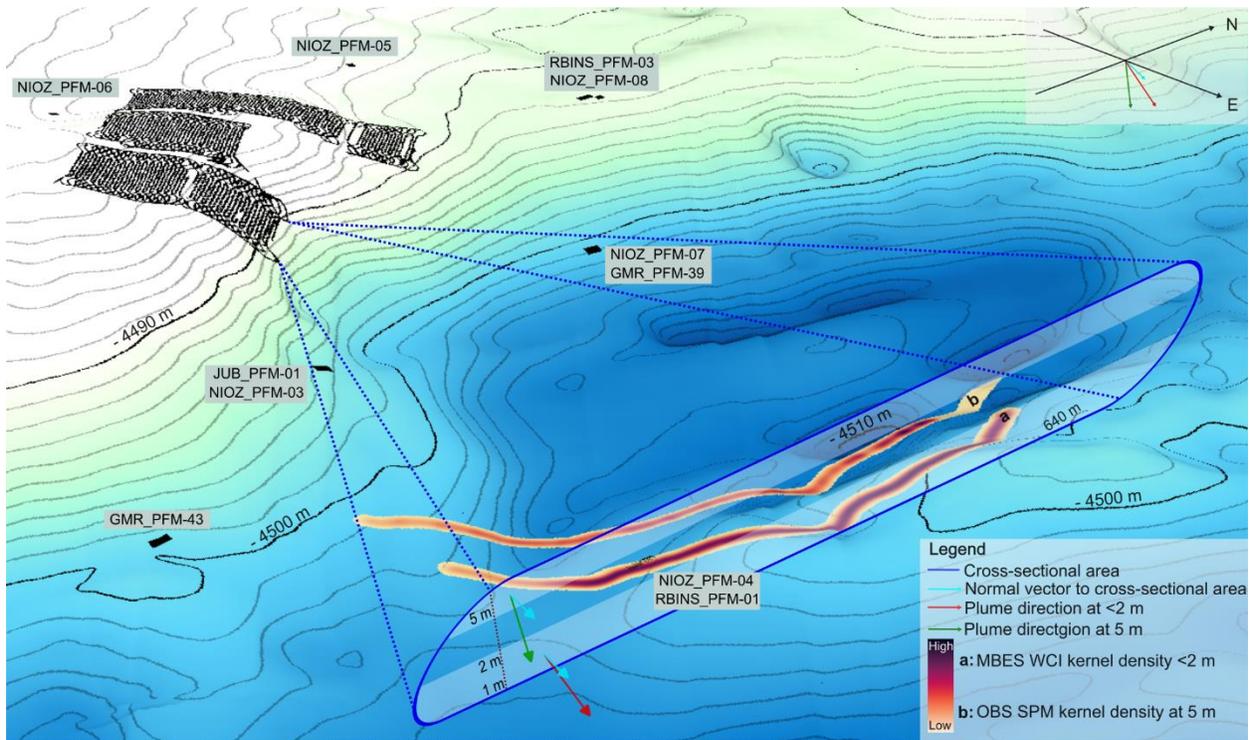
**Fig. S16 | Sediment redeposition 500 m E-SE from the impact site.**

Video frame grab shows the seafloor, GMR\_PFM-40, and GMR\_PFM-43 platforms at 500 SE from the impact site. The distance between A and B is 0.75 m. **b)** Video frame grab that shows the seafloor and NIOZ\_PFM-04 and RBINS\_PFM-01 platforms at 500 m E from the impact site. In both images, the redeposited sediment has covered the nodules. Please see Fig. 3 for the location of the platforms.

<b>a</b>		Pre-impact	Post-impact
Distance from mining lanes (m)	50		
Distance from nearest OBS (m)	0.15		
Maximum SPM (mg L <sup>-1</sup> )	264		
Pre-detectable coverage (%)	44		
Post-detectable coverage (%)	1.7		
<b>b</b>			
Distance from mining lanes (m)	50		
Distance from nearest OBS (m)	200		
Maximum SPM (mg L <sup>-1</sup> )	35		
Pre-detectable coverage (%)	27.5		
Post-detectable coverage (%)	10.5		
Distance from mining lanes (m)	1000		
Distance from nearest OBS (m)	160		
Maximum SPM (mg L <sup>-1</sup> )	28		
Pre-detectable coverage (%)	27.1		
Post-detectable coverage (%)	9.5		
Distance from mining lanes (m)	1800		
Distance from nearest OBS (m)	300		
Maximum SPM (mg L <sup>-1</sup> )	3.9		
Pre-detectable coverage (%)	6.5		
Post-detectable coverage (%)	20.1		

**Fig. S17 | Seafloor images before and after the sediment redeposition**

**a)** 50 m W from the impact site (next to the NIOZ\_PFM-06 platform). **b-d)** 500–1800 m SE from the impact site along the AUV survey lines shown in Figure 10. **d)** The images are at the end of the survey lines outside the zone of sediment redeposition (see Fig.10). The lower nodule coverage in the pre-impact image is due to the seafloor spatial variability in nodule coverage; the pre-impact image was taken within a depression that has lower nodule coverage<sup>1</sup>. The lateral distance from the nearest optical backscatter sensor (OBS), the maximum suspended particulate matter (SPM) concentration and the % detectable seafloor nodule coverage per square meter (see Methods) are provided.



**Fig. S18 | Cross-sectional area of the benthic sediment plume for the first 5 m altitude**

Schematic representation of the cross-sectional area of the benthic sediment plume, generated during the first mining strip. The multibeam echosounder water-column imaging (WCI) and the suspended particulate matter (SPM) concentrations recorded by the optical backscatter sensors (OBS) mounted on the autonomous underwater vehicle have the same lateral extent. The cross-section is located 500 m E-SE from the impact site, it has a height of 5 m and a length of 640 m. The seafloor vertical exaggeration is  $\times 8$ . The location of seafloor platforms is also given.

## References

<sup>1</sup>Gazis, I.-Z., Charlet, F., & Greinert, J. An interpretable multi-model machine learning approach for spatial mapping of deep-sea polymetallic nodule occurrences. *Natural Resources Research*. (2024).