Megafauna community assessment of polymetallic nodule fields with cameras: Platform and methodology comparison

Timm Schoening[1], Autun Purser[2], Daniel Langenkämper[3], Inken Suck[1], James Taylor[4],
Daphne Cuvelier[5,6], Lidia Lins[7], Erik Simon-Lledó[8], Yann Marcon[9], Daniel O.B. Jones[8],
Tim Nattkemper[3], Kevin Köser[1], Martin Zurowietz[3], Jose Gomes-Pereira[10,6], and Jens Greinert[1]

[1]GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany
[2]Alfred Wegener Helmholtz Institute for Polar and Marine Studies, Bremerhaven, Germany
[3]Bielefeld University, Bielefeld, Germany
[4]Senckenberg, Wilhelmshaven, Germany
[5]MARE - Marine and environmental sciences centre/IMAR - Instituto do Mar
[6]Centro OKEANOS - Universidade dos Açores, Horta, Portugal
[7]Ghent University, Ghent, Belgium
[8]National Oceanography Centre, University of Southampton Waterfront Campus, Southampton, UK
[9]MARUM, Bremen, Germany
[10]Naturalist, Lda. and Atlantic Naturalist Association, Horta, Portugal

Correspondence: Timm Schoening (tschoening@geomar.de)

Abstract. With the mining of polymetallic nodules from the deep sea seafloor again approaching commercial viability, decisions must be taken on how to most efficiently regulate and monitor physical and community disturbance in these remote ecosystems. Image based approaches allow non-destructive assessment of larger fauna abundances to be derived from survey data, with repeat surveys of areas possible to allow time series data collection. At time of writing key underwater imaging

5 platforms commonly used to map seafloor fauna abundances are Automated Underwater Vehicles (AUVs), Remotely Operated Vehicles (ROVs) and towed camera "Ocean Floor Observation Systems" (OFOSs). These systems are highly customisable, with mounted cameras, illumination systems and deployment protocols rapidly changing over time, and even within survey cruises. In this study 8 image datasets were collected from a discrete area of polymetallic nodule rich seafloor by an AUV and several OFOSs deployed at various altitudes above the seafloor. A fauna identification catalogue was used by 5 annotators to

10 estimate the abundances of 20 fauna categories from the different data sets. Results show that for many categories of megafauna differences in image resolution greatly influenced the estimations of fauna abundance determined by the annotators. This is an important finding for the development of future monitoring legislation for these areas. When and if commercial exploitation of these marine resources commences, to ensure best monitoring practice, unambiguous rules on how camera-based monitoring surveys should be conducted, and with what equipment, must be put in place.

1 Introduction

The increasing demand on high-tech metals for consumer and industrial high technology devices has again stirred interest in the potential use of the deep sea polymetallic nodule fields of the World Ocean as exploitable sources of these materials in the near future ((Peukert et al. (2018a); Volkmann and Lehnen (2018); Yamazaki and Brockett (2017))). This increasing interest, coupled
20 with the technological improvement of marine mining equipment and the granting of exploration licenses within the Clarion Clipperton Fracture Zone (CCFZ) ((Lodge et al. (2014))), has resulted in several recent European research projects (e. g. JPI Oceans MiningImpact 1&2 and MIDAS) being funded to focus on the study of these remote ecosystems, to derive a better understanding of the nodule distribution ((Peukert et al. (2018b)), community structure of macro-fauna ((De Smet et al. (2017)) and mega-fauna ((Simon-Lledó et al. (2019b)), ecosystem functioning and susceptibility to damage following anthropogenic
25 perturbation and / or resource removal ((Jones et al. (2017); Vanreusel et al. (2016)). Despite the occurrence of nodule fields in the Atlantic, Pacific and Indian Oceans, the majority of research efforts have been focused on the CCFZ, located in the North Central Pacific, as it has the highest known density of nodules ((Jones et al. (2017); Mullineaux (1987); Simon-Lledó et al. (2019b)) and the Peru Basin (South Central Pacific) ((Bluhm (2001); Purser et al. (2016); Simon-Lledó et al. (2019a)), both of which have been considered likely initial regions for early commercial exploitation. Focused scientific study commenced
30 in the 1980s, with simulated mining studies conducted in both areas, to allow the response of fauna to mining activities to be assessed ((Lam et al. (2006)). These studies are summarised in Jones et al. (2017), with the "DISturbance and COLonization" (DISCOL) long-term study in the Peru Basin being the most extensively perturbated region of seafloor studied to date ((Thiel (2001)). Prior to the 1980s, only occasional opportunistic fauna collection records had been published from these areas. Since the 1980s regular biological box core sampling has been conducted in the CCFZ, whereas the majority of fauna sampling
35 at the DISCOL area has been image based, augmenting some initial trawl sampling deployments. The DISCOL experiment was designed to show what effects physical disturbances, such as those caused by future commercial deep-sea mining, might have on the seafloor and its inhabitants. In 1989 a plough-harrow was used to create a large-scale disturbance on the seafloor. It destroyed megafauna within the plough tracks to a large extent and buried the manganese nodules in the area. As a result fauna that lived attached to the nodules were removed. The soft-bottom community, however, did show signs of recovery in
40 the seven years of the study. The repopulation of the disturbed areas by highly motile and scavenging animals started shortly after the area was ploughed ((Bluhm (2001)). Seven years later hemi-sessile animals had returned to the disturbed areas, but the total abundance of soft-bottom taxa was still low compared to the pre-impact study. Nearby reference areas not impacted by the experiment showed natural changes in animal abundances during the study ((Bluhm (2001)). The ploughing activities created a sediment plume that resettled in the surrounding areas. In these not directly impacted areas, animal densities declined
45 immediately after the ploughing event and although densities later (i.e. 3+ years) appeared to be greater than in the pre-impact study reference areas ((Bluhm (2001)), megafaunal community composition in these areas is still to date significantly different than that found within plough tracks and reference areas ((Simon-Lledó et al. (2019a)). As has been reported from many ecosystems, the methodologies used to quantify fauna abundances and species diversity can greatly influence assessments, rendering direct comparison between regions sampled differently problematic (Jaffe (2014); Lam et al. (2006); Murphy and

50 Jenkins (2010); Wilson et al. (2007)). Further, small variations in deployment techniques or sampling parameters (e.g variables such as mesh size or trawl speed for direct sampling, illumination, camera and lenses for remote sampling) can also influence the quality of the collected data (Purser (2015)). In this study a range of commonly used imaging platforms were deployed at varying altitudes above seafloor to survey megafauna across a defined region of the DISCOL Experimental Area (DEA) - a region of the Peru Basin with abundant seafloor nodule coverage, within which a plough harrow was deployed 78 times in 1989

55 with the aim of driving all polymetallic nodules from the sediment surface into the underlying soft sediments (figure 1) (Bluhm (2001)). These collected images were then placed into an online image annotation system (BIIGLE) (Langenkämper et al. (2017)) and fauna were identified in the different image sets by five annotators using a predetermined, species catalogue. The hypothesis tested was that imaging methodology impacts on conclusions made on both composition and abundance of fauna in polymetallic nodule fields. This study aims to provide useful information and guidance on how future optical monitoring

60 of these and other remote ecosystems should most efficiently be conducted, should commercial exploitation of these remote resource fields commence.

1.1 Polymetallic nodules and associated fauna

Polymetallic nodules, as well as representing a potential commercial resource (Burns and Burns (1977); Petersen et al. (2017); Watling (2015)), are a key hard substratum that, in combination with the background soft sediment, act to increase habitat

65 complexity promoting the occurrence of some of the most biologically diverse seafloor assemblages in the abyss (Vanreusel et al. (2016); Simon-Lledó et al. (2019c)). Nodule fields at the abyssal Pacific (the most commercially viable) can comprise of nodules of up to 25 cm diameter (Sharma (2017)), and at a range of abundance densities (e. g. 0-30 kg / m^2, (Mewes et al. (2014)). Methods of formation are uncertain, though each individual nodule tends to form around a small shell fragment, shark tooth or equivalent small hard foci. With growth, individual nodules become heavier and capable of supporting, as an anchor

70 or hard substrate, a range of larger filter feeding organisms (Simon-Lledó et al. (2019c); Tilot et al. (2018)), such as sponges (stalked (Kersken et al. (2018)) and encrusting (Lim et al. (2017))), stalked crinoids, soft and hard corals (Cairns (2016)), xenophyophores (Gooday et al. (2017)), sabellid worms etc. (Bluhm (2001)). Sessile organisms in turn support a diverse array of mobile and sessile epibenthic organisms, including further sponges, corals, crinoids and worms, as well as mobile and semi mobile fauna such as amphipods, isopods, anemones, brooding octopi (Purser et al. (2016); Beaulieu (2001)) and many others

75 (Vanreusel et al. (2016)). Although soft sediment stalked sponge fauna are found in nodule abundant regions, the nodule based epifauna supports increased local biodiversity and abundance in species. In addition to providing a hard substrate for living attachment, nodules also increase the range of hydrodynamic niches available to the local ecosystem fauna, as well as adding complexity to food fall transport pathways. Recent cruise observations from the DISCOL region showed rapid transport of dead salp, following a surface bloom, to the seafloor (Boetius (2015)). These dead salp were then hydro-dynamically trapped

80 by benthic currents alongside nodules, providing a local food supply to the nodule community which might otherwise have been transported from the region by the ambient benthic flow conditions. Though not addressed in the current study, polymetallic nodules also provide increased and varied habitat niches for infauna (across all infauna size classes) below and surrounding the nodules, with their presence influencing local biogeochemical activity and oxygen penetration pathways.

1.2 Potential impacts associated with nodule extraction

Nodule collection will locally remove the major source of hard substrate in nodule field areas, rendering the remaining habitat unsuitable for some fauna (i. e. suspension feeders), as observed in experimental mining studies in the CCFZ (Jones et al. (2017); Vanreusel et al. (2016)) and DISCOL areas (Simon-Lledó et al. (2019a)). Further, depending on the removal technique, the seafloor will likely be highly perturbed, with a range of depressions, ploughs or compaction tracks potentially formed (Jones et al. (2017)). These features will increase the complexity of biogeochemical activity in the region (Paul et al. (2018)), and influence local hydrodynamic conditions. Experimental tracks made with both epibenthic sled (Greinert (2015)) and plough harrow (Bluhm (2001)) have resulted in seafloor topography which greatly focused seafloor salp deposition abundances following a surface bloom event occurring during SO242-2 (Boetius (2015)). Such localised food abundance variability in the deep sea will likely result in a further modification of the fauna communities found in these exploited regions.

1.3 Methodologies for fauna abundance assessment

Box coring or multi coring are commonplace survey methodologies in impact assessments and monitoring programmes, conducted to assess impacts on small fauna (e. g. less than 1 cm) following an anthropogenic impact event (Gage and Bett (2005)). For larger fauna, image-based surveys usually provide much more accurate estimations of benthic taxa richness and numerical density than traditional trawling techniques (Morris et al. (2014); Ayma et al. (2016)), with a minor impact. As we approach the third decade of the 21st century, and experience a continuous increase in public interest in maintaining as near to intact as possible even remote and inaccessible ecosystems, non-invasive monitoring is required (Bicknell et al. (2016)). When planning for future polymetallic nodule fauna abundance assessment following commercial exploitation of these remote resource fields, the associated human impacts of monitoring programmes should be as minor as possible. We therefore focus within this paper on the contrasting suitability of various image based approaches to assessing fauna abundance in polymetallic nodule abundant ecosystems. Furthermore, image data can be made publicly available to regulators, interested NGOs and other players easily via online platforms (Langenkämper et al. (2017)) allowing these stakeholders the opportunity to conduct their own studies or analysis with the same primary data. In the case of monitoring activities utilising directly collected fauna, from box core, multicore or ROV collection, much of the material will be processed once, by one lab, and then either be destroyed or degraded during the processing steps - preventing further studies. Image data also facilitates the straightforward archiving of collected data (Schoening et al. (2018)), for later comparison with subsequent images, which may potentially be collected decades after experimental or industrial disturbance, with the aim of gauging long-term recovery rates. Given the extremely long lifespans of many deep sea fauna (Norse et al. (2012); Roark et al. (2009)), this is an important consideration when developing monitoring strategies for efficient and useful impact assessment within these ecosystems.

1.4 Factors determining the quality of deep sea image data

When abundances of fauna are determined from box core, multicore or trawl samples, these abundances are based on the area of the sediment sampled (box core or multicore dimensions) or trawled (area covered by epibenthic sled or trawl). Although

the type of trawl or corer may influence the results obtained to some extent (i.e. net size and tow speed important for trawls, closing mechanism for box corers), there are possibly a greater number of factors which can influence the estimations of fauna abundance derived from image based data. The most significant of these factors are introduced below:

1.4.1 Camera optics

120 The area of seafloor which may be imaged by an optical platform is determined by the lens parameters used in the camera system, distance and orientation to the seafloor, sensitivity of the system to motion and illumination and a range of other factors (Jaffe (2014)). Larger areas of seafloor can be imaged with wide angle or 'fish eye' camera systems (Kwasnitschka et al. (2016)), though there is an associated vignetting effect rendering the details collected from the extremities of an image less detail rich than areas of seafloor more directly below the lens centre (Cauwerts et al. (2012); Purser et al. (2009)). The raw

125 images collected can appear quite distorted and manual labelling of fauna within these images is more difficult towards the edges of each image. Digital post-processing of these distorted images can be reasonably straightforward when the arrangement of optics for an imaging platform are known, and for larger fauna these processed images can be suitable for subsequent analysis (Schoening et al. (2016, 2017)). However, image processing cannot create 'new improved' data and therefore there will always be a loss of information at the image extremities following lens correction. Lenses of a more 'telephoto' or narrower angle will

130 allow collection of less distorted images, though these collected images will capture a significantly smaller area of seafloor than may be achieved with wider angle systems.

1.4.2 Illumination and power provision

The deep sea is a dark environment with no sunlight penetration. It is therefore essential that camera systems are supplemented by artificial illumination. To provide sufficient illumination for video and still camera systems, sufficient power reserves must

135 be mounted on the platform or delivered on demand via a cable from the support vessel. The amount of power which can be provided to a platform is determined by a range of design and operational parameters; Automatic Underwater Vehicles (AUVs) for example must remain reasonably light and must carry sufficient power to provide mobility and to take images at depth; towed camera systems in contrast are always attached to a cable, and this cable may be capable of supplying power, such as a coaxial fibre-optic cable, then considerable power may be provided to allow continuous seafloor illumination. Positioning of the

140 lights on an imaging platform can be difficult, and optimising the spread of light, i.e. maintaining an equal light balance across the imaged area, can be challenging. A vignetting of illumination in an image can be partially addressed prior to analysis by excluding the edges of collected images from analysis (Marcon and Purser (2015); Purser et al. (2009)). Given that AUVs must carry all required power (for mobility and imaging) with them can result in a less than optimal illumination of the seafloor (see 1.4.4). No doubt light emitting diode (LED) technology will continue to become more efficient, but at present these prevalent

145 lower light condition data sets constrains the seafloor resolution which may be achieved during imaging surveys. Additionally, when lights and camera are mounted close to each other (as compared to the altitude), a significant amount of light might be scattered by the water column into the camera, leading to a degraded "foggy" image, which is an issue for small platforms or high-altitude photography. Finally, also the colour spectrum of the light needs to be considered, as for instance the returned

yellow, orange and red components of the signal might be too weak to support taxonomy analyses, depending on the type of
150 light source. The illumination system is though in direct relation with the target altitude of survey above the seabed.

1.4.3 Data volume

Pioneer image-based studies in polymetallic nodule fields were conducted with analogue film based camera systems (although
live seafloor views were provided to towed systems via a basic TV camera setup) (Bluhm (2001)). This limitation constrained
deployments to the collection of a few 100s of images. At present, camera systems can deliver many images per second, even
155 under low light conditions. This potentially high flow of image data however requires adequate digital storage space on the
imaging platform (Kwasnitschka et al. (2016)), or the facility to be transferred directly to a shipboard storage system (Purser
et al. (2018)). This increased data flux allows for more complete spatial studies of the seafloor to be made with an imaging
platform, but to get this additional information from the data set, increased processing time is required.

1.4.4 Platform altitude

160 The distance to an object can greatly alter the quality of an image which can be taken with a particular range of optics
and illumination sources. Although this may sound a straightforward parameter it may play a hugely important role when
analysing fauna abundances in an area. Maintaining a uniform altitude throughout and between survey deployments is highly
desirable (i. e. to standardise the object/fauna detectability rates), but may be difficult. In regions of the World Ocean where
the seafloor is highly complex, such as at deep water coral reefs (Purser et al. (2009)) or within canyon systems (Orejas
165 et al. (2009)) it can be a struggle to maintain an equal distance from camera optics from towed, automated, remote and
submersible based imaging platforms to the seafloor. For polymetallic nodule fields however, the seafloor is generally fairly
uniform in depth, with very gentle slopes more the norm than occasional sudden slopes or cliff walls. Even so, towed platform
altitude stability can be greatly influenced by operator skill, experience, environmental conditions (i.e. wave conditions at
surface) or ship infrastructure (winch operational parameters / presence or absence of heave compensators). Automated AUV
170 imaging platforms are improving in stability and mission planning at a rapid rate (McPhail et al. (2010); Yu et al. (2018)), and
maintaining flight altitudes is now a standard surveying procedure. Operations with these expensive devices tend to err on the
side of caution; ground tracking often set with a conservative 5 - 10 m flight altitude. At these higher flight altitudes, more light
is required to illuminate the seafloor than when a comparable AUV is deployed close to the seafloor (see 2.1.6).

1.4.5 Dataset resolution

175 Within this study, image datasets are compared with regard to their resolution. This factor is a combination of the camera
optics and the deployment altitude and allows to compare image datasets numerically. The camera optics determine the pixel
resolution (usually in the tens of megapixels for state of the art camera systems). The field of view of the camera objective
lens and the deployment altitude determine the image footprint i.e. the area in square meters that is covered by a single image

acquisition. These two values can be combined to a measure of Megapixels per square meter (MPix / m^2) to analyse the
annotator performance and fauna density estimates consistently.

1.4.6 Time series studies

To determine the level of impact an event has had on a region of seafloor, repeat visits to a locale are required. Ideally a
number of surveys at differing times of the year would be conducted before an impacting event, to gauge the background fauna
community of a region, and to identify any seasonality in community patterns. These baseline studies would be subsequently
followed up with repeat surveys at different time points during and after the impacting event. These repeat visits should
allow identification of the duration and recovery of impacts. Planning such a study may sound straightforward, but given the
remoteness of many regions of deep sea seafloor, getting the same equipment and survey crews together for a full repeat visit
campaign may well be difficult. One such study, to gauge the impact of oil and gas exploration drilling on cold-water coral reefs
on the Norwegian Margin aimed to survey a number of reefs visually on 5 occasions (Purser (2015)). Despite these 5 survey
cruises taking place within a 3 year period to a relatively accessible area of Norwegian continental shelf, all used different ROV
systems to carry out the work. Analysis of collected data was further complicated by different camera systems, illumination
systems, flight altitudes and dive plans being used for each survey.

2 Methodology

For this comparative study of the effectiveness of various imaging platforms for gauging megafauna abundances in polymetallic
nodule ecosystems, 8 distinct image data sets DS_A - DS_H (see table 1) collected from a discrete area of seafloor were compared
(ca. 600 x 150 m^2). These 8 datasets were collected by three different towed camera platforms (one of which was deployed at
several altitudes above seafloor) and an AUV (deployed at two different altitudes above seafloor) during three research cruises.
One dataset (DS_C) was acquired during SO106, the other seven during SO242/1 (DS_A, DS_B, DS_D) and SO242/2 (DS_E-DS_H).
DS_H was created by producing a mosaic of the seafloor from overlapping AUV imagery, then dividing the mosaic into smaller
image tiles for fauna analysis. All image sets were analysed by five annotators a_1 - a_5, using a predesigned fauna catalogue
to label a selected group of 20 fauna categories ω_1 - $\omega_2 0$ fauna within each discrete image (see figure 6). From this labelling
effort, the densities of the various identified fauna categories in each data set were statistically compared.

2.1 Imaging platforms, resolutions and deployment altitudes

2.1.1 DS_A (4.49 MPix/m^2) and DS_B (3.89 MPix/m^2): Low altitude imagery from AWI OFOS camera sled

Towed still image and video sleds are equipment often used for gleaning some information on seafloor physical and megafauna
community structure (examples can be found in figures 2 b, c, d). These devices consist of a solid frame which is connected to
a survey vessel by an umbilical cable, in many cases capable of supplying power and data transference between the ship and
the platform. To operate, an altitude above the seafloor is set by the users, as a function of seafloor topographical structure,

items of interest, vessel speed and weather conditions. A winch operator maintains the appropriate flight altitude above seafloor

210 as the survey vessel tows the device over the requested course. These systems can utilise reasonably simple cable systems to allow live TV signals from the seafloor to reach a towing support vessel, or modern fibre-optic cables through which high data loads can be transmitted in real-time. The simplicity and relatively cheap costs of these towed systems, coupled with their moderate personnel requirements have made them an attractive choice for use in scientific expeditions, and particularly for time series studies where the same equipment is required for each revisit to a location. For this current study, the AWI OFOS

215 system was used for collection of several of the data sets. Developed for time series analysis of the HAUSGARTEN marine time series station, the system has seen 15 years of regular use and numerous megafauna fauna papers have been published based on collected data (Bergmann et al. (2011); Pham et al. (2014); Purser et al. (2016); Taylor et al. (2016, 2017)). The AWI OFOS consists of a solid frame containing vertically downward facing still image and video cameras (figure 2). Additionally, the system mounts LED lights to supply light for the video camera, as well as powerful flash units to allow 26 megapixel still

220 images to be taken from an optimal altitude of 1.5 m above the seafloor. The AWI OFOS also incorporates 3 parallel lasers, to allow seafloor coverage (and fauna sizes) to be quantified in the images and video data collected. Figures 3 (a) and (b) show typical images collected from the DISCOL area from an operational altitude of 1.6 m and 1.7 m. For the current study, two deployments of the AWI OFOS system were made at altitudes of 1.6 m (DS_A) and 1.7 m (DS_B).

2.1.2 DS_C (1.05 MPix/m^2): High altitude, digitised analogue imagery from EXPLOS camera sled

225 Prior to the equipping of research vessels with fibre-optic cables allowing HD video to be transmitted directly to the support vessel during a dive, it was common practice to set up a low quality video link to the seafloor to allow the operators of a towed device to maintain an appropriate flight altitude above the seafloor during a deployment. The scientific data collected were still images triggered manually from the ship but recorded onto analogue photographic film using a PHOTOSEA 5000 camera mounted on the towed device. This requirement to mount actual film on the towed platforms tended to result in deployments

230 with <400 images, as this was the maximum number of photographs which could be taken with standard, extended 35 mm film magazines. In 1989, after the seafloor ploughing, such an analogue towed camera rig was used to image in the DISCOL area (figure 2 (a)). The photographs taken back then were recently digitised by the JPIO Mining Impact project and made available for this study. An example image is given in figure 3 (c).

2.1.3 DS_D (0.98 MPix/m^2): High altitude imagery from AWI OFOS camera sled

235 With increasing distance from seafloor, a particular optical system can image a greater area for a given set of optics, assuming correct focusing etc. can be achieved. With a doubling of distance however, effectiveness of illumination is reduced by 75%. For towed systems this may be compensated for by additional supply of power / a greater number of lights. For the current study however, the same AWI OFOS system introduced in section 2.1.1 was redeployed with the same standard lighting configuration at a flight altitude of 3.3 m. Figure 3 (d) shows a typical seafloor image taken from this altitude.

2.1.4 DS$_E$ (0.24 MPix/m^2): Low altitude imagery from AUV Abyss

During SO242-1, GEOMAR's AUV Abyss (Linke and Lackschewitz (2016)) has been deployed for several photographic mapping missions (see figure 2 (c)). The vehicle's original camera had been replaced by a Canon-6D DSLR camera and the Xenon strobe by an LED flash system (Kwasnitschka et al. (2016)), placed 2 m apart from one another. The low altitude vertical imagery of DS$_E$ was captured from a target altitude of 4.5 m, at a speed of 1.5 m/s and at a frame rate of 1 Hz. The system was equipped with a Canon 8-15mm fisheye lens (fixed to 15mm) centred in a dome port. Owing to weak illumination in the outer image regions, only the central 90° (across track) resp. 74° (along track) of the fisheye images were used and tri-linearly resampled to a picture that an ideal rectilinear 18mm lens would have taken. An example picture is shown in figure 4 (a).

2.1.5 DS$_F$ (0.16 MPix/m^2): Low altitude imagery from custom OFOS camera sled

During SO242-1 the area of interest was surveyed with a colour video camera (Oktopus GmbH) in conjunction with one Oktopus HID 50 light mounted vertically on a towed frame (See figure 2 (b)). The signal was transmitted to a deck unit (Oktopus GmbH VDT 3) and recorded using an external video converter (Hauppauge - HD PVR), which converted the signal to .mp4 files and was then recorded in a PC using ArcSoft Total Media Extream software. For this study frames were extracted from these video files at a rate of 0.1 Hz. The custom OFOS was put together in an 'ad-hoc' fashion, from a range of off the shelf components, to mimic "pioneer" image-based methodology, rather than as a fully designed and integrated device. An example image is given in figure 4 (b). Further details of the custom OFOS and its deployments can be found in (Greinert (2015)).

2.1.6 DS$_G$ (0.07 MPix/m^2): High altitude imagery from AUV Abyss

As a result of the fixed distance of roughly 2 m between camera and light source on AUV Abyss, images taken by the above system at higher altitudes were increasingly suffering from very strong backscatter, additional to the loss of colour resulting from the large distance from the light source to the seafloor and back into the camera. Although the AUV imaged at altitudes above 10 m, those images were deemed of a quality unsuited for fauna analysis. Consequently, besides the 4.2 m "low altitude" AUV imagery in DS$_G$, AUV imagery acquired at 7.5 m altitude represents the dataset of maximum altitude in this contribution. Apart from the different altitude, all capture parameters in DS$_G$ remained the same as in DS$_E$. An example image for this dataset is shown in figure 4 (c).

2.1.7 DS$_H$ (0.04 MPix/m^2): Low altitude imagery AUV Abyss and extracted from a photo mosaic

AUV images of station SO242-1_102 were collected at ca. 4.5m above seabed with 80% along track and 50% across track overlap in order to build one large photo mosaic out of the images. In order to alleviate water and illumination effects otherwise dominant in the final mosaic, a robust statistical estimate of the illumination component has been performed. For this, each image was robustly averaged with the seven images taken before and after, producing an image without nodules that represents the illumination effects. The raw image was then - pixel-wise - divided by the illumination image and multiplied by

the expected seafloor colour, which was obtained from box core photographs of the same cruise. For each track of a multi-track AUV mission the images were registered against each other, leading to relative AUV localisation information with sub-cm accuracy. Afterwards the photos were projected to the seafloor and rendered into a virtual ortho-photo with 5mm/pixel resolution (reflecting the best resolution in the fisheye images) of roughly seven hectares size. The photo mosaic was then subdivided into

275 ca. 11,000 tiles and uploaded to BIIGLE for megafaunal assessment (Simon-Lledó et al. (2019a)). An example tile is shown in figure 4 (d).

2.2 Image labelling methodology

Within the study, 1340 seafloor images (or mosaic tiles) were analysed for megafauna abundance and community structure estimation (see table 1). All images used in the study were imported into the BIIGLE online annotation system (Langenkämper

280 et al. (2017)). Once imported, five annotators inspected the images independently and annotated fauna by placing a circle around each fauna individual using the BIIGLE interface (see figure 5). To assist in this, a one page fauna identification guide with 20 categories was produced (see figure 6), from which the annotators could work.

2.3 Observer Agreement

Manual annotation was conducted independently. To compare results of the five annotators a_i (i = 1,..,5) inter-observer agree-

285 ment was computed (Schoening et al. (2012)). First, the individual annotations of each pair of two annotators were compared regarding to annotation location (i. e. the detection step) and annotation label (i.e. the classification step). Annotations of individual experts were then grouped to gold standard annotations to increase the robustness of the dataset comparison. Grouping was conducted by fusing overlapping annotations of similar size to one grouped annotation. The location and size of this grouped annotation was computed as the average of the annotation position and radius of the single expert annotations. The

290 support of one annotation quantifies how many experts found this individual and thus ranks between 1 and 5. The label of the grouped annotation was selected as the most frequent label within the grouped annotations. Annotations that were supported by only one annotator were discarded. Also, if no two annotators assigned the same label to an annotation it was discarded. As a further measure of observer agreement, Cohen's kappa was computed (McHugh (2012)).

2.4 Fauna-specific statistical analysis

295 The average abundance estimations of each individual fauna category computed for each of the 8 image sets was derived from the annotations made by each independent annotator. The five density estimates obtained for each fauna category, as generated from the labels made by the individual image annotators across the 8 imaging platform data sets were compared using nonparametric Kruskal-Wallis tests. These tests were conducted using the software package SPSS 17.0.

3 Results

3.1 Aggregated results for datasets

Figure 7 shows aggregated results for various characteristics of the eight datasets and annotations computed by averaging across all fauna categories (see also table 2). All figures except (g) further visualise the results of the grouped annotations. Most obvious is the increase of fauna density with imaging resolution (see 7 (a)). This trend is mirrored in the observation that the median size of the annotated fauna decreases with increasing resolution (see figure 7 (b)). Together it can be reasoned, that the increased resolution allows to annotate smaller objects, increasing the total amount of individuals annotated. Anyhow, it is also obvious that the increased resolution comes with an increase in observer disagreement. Figure 7 (c) shows that the standard deviation of fauna densities created by the five experts. Figure 7 (d) - (f) highlight the tradeoff between resolution and seafloor inspection effort. In (d) it is obvious that the increase in resolution comes with a decrease in acquisition efficiency in terms of the area per hour (m^2 / h) that can be imaged. Subfigure (f) shows that, although higher densities of fauna are detected for high resolution datasets, it still requires to manually inspect more megapixels per annotation compared to lower resolution datasets. The annotation effort for such high resolution data sets is thus over-proportionally large.

3.2 Observer agreement

Figure 7 (g) outlines the importance for any image-based study to rely on more than one annotator. It shows the generally poor observer agreement in this study when considering the single expert annotations (see also 2). It further highlights that the observer agreement drops with increasing image resolution reflecting the results in (c). When grouping the single observer annotation to gold standard annotations anyhow the observer agreement increases significantly (see 7 (h)). This increase is similarly reflected by the Cohen's kappa values that are all but one above 0.7 which is deemed as "substantial agreement" (0.6-0.8).

3.3 Fauna-specific statistical Analysis

The seafloor densities of the 20 categories of fauna and seafloor features, as quantified by the 5 independent annotators are given in figure 8 (mobile fauna) and figure 9 (sessile fauna). Kruskal-Wallis tests indicated that for all fauna categories (with the exception of 'molluscs') observed, individual densities differed by imaging platform at the 95% threshold ('Small Encrusting', 'Starfish') or <99% threshold (all other fauna categories). For sessile fauna, average individual densities observed were highest across fauna categories in DS$_A$. Generally the averaged densities for this dataset acquired at 1.6 m altitude were roughly double to triple those observed in DS$_B$ which was collected in the same year from a slightly higher median altitude of 1.7 m. Densities of sessile fauna derived from AUV data were generally lower than those derived from OFOS data. Sessile fauna densities derived from AUV data acquired at 4.2 m altitude (DS$_E$) were invariably higher than those derived from 7.5 m AUV data (DS$_G$). Sessile fauna densities determined from the mosaicked images were roughly equivalent or a little lower than the densities determined from the both the uncombined AUV data sets (see figure 9). For mobile fauna, trends in densities of

330   fauna categories were less dependent on observing platform. Although differences were indicated as significant for many fauna categories (see table 3), these differences were not clearly relatable to imaging platform deployment altitude or methodology and observers (see figure 8).

4   Discussion

4.1   Spatial and Temporal factors

335   The current study represents an attempt to estimate the impact of a range of devices imaging the same area of seafloor on experts' manual annotations. Given the inaccuracies of cs. 1% achievable with the POSIDONIA underwater positioning system used for the majority of imaging deployments (Peyronnet et al. (1998)) and the lack of distinct seafloor features in the DISCOL polymetallic nodule province ensuring that exactly the same areas of seafloor were imaged was not possible. Given the reasonably homogenous nature of the seafloor (at the meters to hundreds of meters scale) in the survey region, it seems likely

340   that reasonably comparable organisms were present across the area. Temporal differences in community structure, particularly between years, cannot be wholly discounted as explanatory factors of differences between data sets. Highly mobile fauna, such as fish and jellyfish, can vary in local abundances on temporal scales of minutes, and even the less mobile ophiuroids and holothurians can respond relatively swiftly to changes in seafloor conditions, such as in response to a food fall or hydrodynamic conditions. Even so, we assume here that temporal and spatial differences between the collected data are of minor significance

345   in explaining the differences in densities observed.

4.2   Deployment altitude and image resolution

Although it was not possible to deploy all platforms at different altitudes within the same cruise, it was possible to do so with the AUV (two altitudes) and the AWI OFOS (three altitudes). For virtually all fauna categories used, the highest density estimates were made from data collected at the lowest deployment altitude and highest pixel resolution. At these altitudes less water is

350   between the camera and the target, reducing distortion and light attenuation effects. The only exceptions to this trend were with the highly mobile, water column dwelling fauna, such as jellyfish and fish. Given the three dimensionality of the habitat utilised by these organisms, observation from a greater altitude is beneficial, and it is thus more likely to image such fauna. This is potentially coupled with avoidance mechanisms triggered in some species as they attempt to avoid the illumination mounted on the imaging platform or the sound of thrusters (in the case of the AUV deployments). The sensitivity of fauna

355   density estimations to deployment altitude does not appear to be linear, or comparable across fauna categories. Larger fauna, such as 'stalked sponges' (see figure 9 (d)) and 'starfish' (see figure 8 (j)) were spotted with equivalent ease across all data sets, whereas smaller fauna, such as 'sessile polychaetes' and 'sponges' (see figures 9 (b) and (i)) were logged with greater frequency in data collected from lower altitudes. These altitude-based trends in density estimation were observed in both AUV and OFOS data sets. Interestingly, an average deployment altitude difference of just 10 cm, from 1.7 to 1.6 m average altitude between

360   SO242-2 OFOS deployments corresponded to a much greater difference in fauna density estimations than the 1.6 m difference

in deployment altitudes between the 3.3 m and 1.7 m data sets. The attenuation of light in water, and the variable impact of this reduction on the wavelengths of reflected light, as well as the size of the fauna image received by the camera likely both play a role in determining the fauna abundance accuracy achievable from a data set. This extreme sensitivity to deployment altitude of derived density estimations is an important consideration when comparing results from different deployments.

4.3   Annotator skill / Observer effect

To label fauna to species level from imagery requires a certain amount of skill, and an awareness of fauna likely to occur in a particular survey region. Even with such knowledge, inter-observer differences in annotations can be significant (Schoening et al. (2012); Durden et al. (2016)). Here however, difference between platform altitude proved more significant than observer effect for all faunal categories. Given the sparsity of many deep sea fauna in nodule provinces (Simon-Lledó et al. (2019b)), key species are of more applicability when determining monitoring strategies for impact assessment, where statistically significant differences in abundances may allow differences in populations of pre-impacted or control areas and those within impacted areas to be determined. These key fauna would be different for different locations and ecosystems, and for deep sea manganese nodule provinces, the level of understanding of ecosystem functioning is probably insufficient to select species of major importance for the ecosystem. Certainly some easily annotated fauna play important roles as habitat engineer species, such as the stalked fauna which add the vertical axis to habitat niche availability (Purser et al. (2016); Vanreusel et al. (2016)). Biogeochemical processes within and at the sediment / seawater interface may well be influenced by mega, macro and meiofauna not visible in even the highest resolution image data, with some large fauna spending some or all their time below sediments, and with the smaller fauna not derivable within the image data. Though densities of these organism categories may be measured with a range of methodologies (Gollner et al. (2017)), the number of samples required coupled with the remoteness of resource sites renders these as probably inappropriate for cost effective monitoring. By providing annotators with a clear identification catalogue, ideally with a limited number of categories (as used in the current study, see figure 6) annotators with little or no experience can identify fauna within an image set with a degree of confidence. For complex studies of detailed community change trained scientific personnel would be required, which again would add a considerable financial cost to the monitoring program. In the future, it is probable that the ongoing developments of computer algorithms for resource quantification (Schoening et al. (2016, 2017)) and fauna identification (Aguzzi et al. (2009); Purser et al. (2009); Schoening et al. (2012); Siddiqui et al. (2017)) will allow a near real-time assessment of fauna abundances in a surveyed region, for a given platform and deployment strategy. At present however, as commercial nodule mining approaches viability, more traditional monitoring approaches are the only ones currently available for integration into regulatory frameworks and work plans.

5   Conclusions

The results from the current study highlight how tightly fauna abundance estimations in manganese nodule ecosystems may be related to investigative methodology. Small differences in imaging platform operational altitude, illumination and lens type all result in the collection of data from which a particular operator can derive quite different estimations of community struc-

ture. These results are not wholly novel, given similar studies conducted in shallow reef environments (Gardner and Struthers (2013)), though they are highly prescient given the commercial interest in these nodule resources and the current lack in back-

395   ground knowledge on ecosystem function. For the first time a quantification was given for the differences in observations based on different platform altitude and the resulting imagery resolution.

The authors of the current study do not propose to recommend a 'perfect' imaging platform for megafauna abundance monitoring in manganese nodule ecosystems, as more work is still needed to determine whether or not there are particular megafauna species particularly important for maintaining current community structures and biodiversity in these regions, and because

400   the commercial viability of the various platforms available for study will surely change during the forthcoming years. We would however give some general guidelines on how longterm monitoring studies in these regions should be planned, to allow collection of data of sufficient quality to allow time series analysis of larger fauna community composition:

1. For a given study location, a comparable survey deployment plan should be used at each time stage of analysis (same sensor payload, instrument platform altitude, deployment speed, seafloor area imaged, sample unit size).

405   2. A well documented camera system should be used (aperture, sensitivity, lens arrangement, mounting angle).

3. Illumination should be maintained across deployments (intensity, wavelength, mounting angle).

4. Annotations by several observers need to be collected and thoroughly merged to create robust data for interpretation

5. The lowest altitude above seabed that can be reached using a given platform will always provide more data and higher taxonomical resolution in the faunal identification

410   Although many of these points may seem obvious requirements for a study, the extended duration of deep sea surveys may lead to technological changes taking place between survey visits, or changes in personnel involved in conducting the work. Even during relatively short (3 year) studies conducted in medium depths offshore Norway most of these points were missed during a recent monitoring campaign (Purser (2015)). We highly recommend that in the developing industry of polymetallic nodule extraction, such guidelines be integrated into licensing agreements, with appropriate commitments made by companies

415   to ensure longterm adherence (commitments such as maintaining appropriate equipment for the duration of the monitoring campaign, providing accurate blueprints/design specification of platforms used at each monitoring stage etc). We also recommend an increase in vigour of studies focusing on the biogeochemical processes at work in these remote ecosystems, so a greater confidence can be made into the appropriateness and ecosystem relevance of any observations made of short or longterm fauna reductions associated with the exploitation of these resources.

420   *Author contributions.* Designed the study: JG, ESL, KK, TS, JNGP; Provided data: TS, AP, ESL, JNGP, KK, JG; Provided infrastructure: DL, TN, MZ, DOBJ, Annotated the images: AP, IS, JT, DC, LL; Analysed the data: TS, AP, DL, JT, DC, LL, ESL, JNGP, YM, MZ; Wrote the manuscript: TS, AP, DC, ESL, KK

Competing interests. The authors report no competing interests.

425   *Acknowledgements.* We thank the crew and scientific parties of cruises SO106, SO242/1 and SO242/2 for their indispensable support in making this study possible.

References

Aguzzi, J., Costa, C., Fujiwara, Y., Iwase, R., Ramirez-Llorda, E., and Menesatti, P.: A novel morphometry-based protocol of automated video-image analysis for species recognition and activity rhythms monitoring in deep-sea fauna, Sensors, 9, 8438–8455, 2009.

430   Ayma, A., Aguzzi, J., Canals, M., Lastras, G., Bahamon, N., Mechó, A., and Company, J.: Comparison between ROV video and Agassiz trawl methods for sampling deep water fauna of submarine canyons in the Northwestern Mediterranean Sea with observations on behavioural reactions of target species, Deep Sea Research Part I: Oceanographic Research Papers, 114, 149–159, 2016.

Beaulieu, S.: Life on glass houses: sponge stalk communities in the deep sea, Marine Biology, 138, 803–817, 2001.

Bergmann, M., Soltwedel, T., and Klages, M.: The interannual variability of megafaunal assemblages in the Arctic deep sea: Preliminary
435   results from the HAUSGARTEN observatory (79 N), Deep Sea Research Part I: Oceanographic Research Papers, 58, 711–723, 2011.

Bicknell, A. W., Godley, B. J., Sheehan, E. V., Votier, S. C., and Witt, M. J.: Camera technology for monitoring marine biodiversity and human impact, Frontiers in Ecology and the Environment, 14, 424–432, 2016.

Bluhm, H.: Re-establishment of an abyssal megabenthic community after experimental physical disturbance of the seafloor, Deep Sea Research Part II: Topical Studies in Oceanography, 48, 3841–3868, 2001.

440   Boetius, A.: RV Sonne Fahrtbericht/Cruise Report SO242-2: JPI OCEANS Ecological Aspects of Deep-Sea Mining, DISCOL Revisited, Guayaquil - Guayaquil (Equador), 28.08.-01.10. 2015, Kiel: Helmholtz-Zentrum fü r Ozeanforschung, 2015.

Burns, R. and Burns, V. M.: The mineralogy and crystal chemistry of deep-sea manganese nodules, a polymetallic resource of the twenty-first century, Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 286, 283–301, 1977.

Cairns, S. D.: New abyssal Primnoidae (Anthozoa: Octocorallia) from the Clarion-Clipperton Fracture Zone, equatorial northeastern Pacific,
445   Marine Biodiversity, 46, 141–150, 2016.

Cauwerts, C., PhD, M. B., and Deneyer, A.: Comparison of the vignetting effects of two identical fisheye lenses, Leukos, 8, 181–203, 2012.

De Smet, B., Pape, E., Riehl, T., Bonifácio, P., Colson, L., and Vanreusel, A.: 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, 103, 2017.

Durden, J. M., Schoening, T., Althaus, F., Friedman, A., Garcia, R., Glover, A. G., Greinert, J., Stout, N. J., Jones, D. O., Jordt, A., et al.:
450   Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding, in: Oceanography and Marine Biology, pp. 9–80, CRC Press, 2016.

Gage, J. and Bett, B.: Deep-sea benthic sampling, Methods for the Study of the Marine Benthos. Oxford, UK: Blackwell, pp. 273–325, 2005.

Gardner, J. and Struthers, C.: Comparisons among survey methodologies to test for abundance and size of a highly targeted fish species, Journal of Fish Biology, 82, 242–262, 2013.

455   Gollner, S., Kaiser, S., Menzel, L., Jones, D. O., Brown, A., Mestre, N. C., Van Oevelen, D., Menot, L., Colaço, A., Canals, M., et al.: Resilience of benthic deep-sea fauna to mining activities, Marine Environmental Research, 129, 76–101, 2017.

Gooday, A. J., Holzmann, M., Caulle, C., Goineau, A., Kamenskaya, O., Weber, A. A.-T., and Pawlowski, J.: Giant protists (xenophyophores, Foraminifera) are exceptionally diverse in parts of the abyssal eastern Pacific licensed for polymetallic nodule exploration, Biological conservation, 207, 106–116, 2017.

460   Greinert, J.: 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, 2015.

Jaffe, J. S.: Underwater optical imaging: the past, the present, and the prospects, IEEE Journal of Oceanic Engineering, 40, 683–700, 2014.

Jones, D. O., Kaiser, S., Sweetman, A. K., Smith, C. R., Menot, L., Vink, A., Trueblood, D., Greinert, J., Billett, D. S., Arbizu, P. M., et al.: Biological responses to disturbance from simulated deep-sea polymetallic nodule mining, PLoS One, 12, e0171 750, 2017.

465 Kersken, D., Kocot, K., Janussen, D., Schell, T., Pfenninger, M., and Arbizu, P. M.: First insights into the phylogeny of deep-sea glass sponges (Hexactinellida) from polymetallic nodule fields in the Clarion-Clipperton Fracture Zone (CCFZ), northeastern Pacific, Hydrobiologia, 811, 283–293, 2018.

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, 16, 164, 2016.

470 Lam, K., Shin, P. K., Bradbeer, R., Randall, D., Ku, K. K., Hodgson, P., and Cheung, S. G.: A comparison of video and point intercept transect methods for monitoring subtropical coral communities, Journal of Experimental Marine Biology and Ecology, 333, 115–128, 2006.

Langenkämper, D., Zurowietz, M., Schoening, T., and Nattkemper, T. W.: Biigle 2.0-browsing and annotating large marine image collections, Frontiers in Marine Science, 4, 83, 2017.

Lim, S.-C., Wiklund, H., Glover, A. G., Dahlgren, T. G., and Tan, K.-S.: A new genus and species of abyssal sponge commonly encrusting 475 polymetallic nodules in the Clarion-Clipperton Zone, East Pacific Ocean, Systematics and biodiversity, 15, 507–519, 2017.

Linke, P. and Lackschewitz, K.: Autonomous Underwater Vehicle ABYSS, Journal of large-scale research facilities JLSRF, 2, 79, 2016.

Lodge, M., Johnson, D., Le Gurun, G., Wengler, M., Weaver, P., and Gunn, V.: Seabed mining: International Seabed Authority environmental management plan for the Clarion–Clipperton Zone. A partnership approach, Marine Policy, 49, 66–72, 2014.

Marcon, Y. and Purser, A.: PAPARA(ZZ)I 1.6: an open-source software interface for annotating deep-sea imagery data, 480 https://doi.org/10.1594/PANGAEA.855568, https://doi.org/10.1594/PANGAEA.855568, 2015.

McHugh, M. L.: Interrater reliability: the kappa statistic, Biochemia medica: Biochemia medica, 22, 276–282, 2012.

McPhail, S., Furlong, M., and Pebody, M.: Low-altitude terrain following and collision avoidance in a flight-class autonomous underwater vehicle, Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment, 224, 279–292, 2010.

485 Mewes, K., Mogollón, J. M., Picard, A., Rühlemann, C., Kuhn, T., Nöthen, K., and Kasten, S.: 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, 2014.

Morris, K. J., Bett, B. J., Durden, J. M., Huvenne, V. A., Milligan, R., Jones, D. O., McPhail, S., Robert, K., Bailey, D. M., and Ruhl, H. A.: A new method for ecological surveying of the abyss using autonomous underwater vehicle photography, Limnology and Oceanography: 490 Methods, 12, 795–809, 2014.

Mullineaux, L. S.: Organisms living on manganese nodules and crusts: distribution and abundance at three North Pacific sites, Deep Sea Research Part A. Oceanographic Research Papers, 34, 165–184, 1987.

Murphy, H. M. and Jenkins, G. P.: Observational methods used in marine spatial monitoring of fishes and associated habitats: a review, Marine and Freshwater Research, 61, 236–252, 2010.

495 Norse, E. A., Brooke, S., Cheung, W. W., Clark, M. R., Ekeland, I., Froese, R., Gjerde, K. M., Haedrich, R. L., Heppell, S. S., Morato, T., et al.: Sustainability of deep-sea fisheries, Marine policy, 36, 307–320, 2012.

Orejas, C., Gori, A., Iacono, C. L., Puig, P., Gili, J.-M., and Dale, M. R.: Cold-water corals in the Cap de Creus canyon, northwestern Mediterranean: spatial distribution, density and anthropogenic impact, Marine Ecology Progress Series, 397, 37–51, 2009.

Paul, S. A., Gaye, B., Haeckel, M., Kasten, S., and Koschinsky, A.: Biogeochemical regeneration of a nodule mining disturbance site: trace 500 metals, DOC and amino acids in deep-sea sediments and pore waters, Frontiers in Marine Science, 5, 117, 2018.

Petersen, S., Hannington, M., and Krätschell, A.: Technology developments in the exploration and evaluation of deep-sea mineral resources, in: Annales des Mines-Responsabilite et environnement, 1, pp. 14–18, FFE, 2017.

Peukert, A., Petersen, S., Greinert, J., and Charlot, F.: Seabed mining, in: Submarine Geomorphology, pp. 481–502, Springer, 2018a.

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 (BG), 15, 2525–2549, 2018b.

Peyronnet, J.-P., Person, R., and Rybicki, F.: Posidonia 6000: a new long range highly accurate ultra short base line positioning system, in: IEEE Oceanic Engineering Society. OCEANS'98. Conference Proceedings (Cat. No. 98CH36259), vol. 3, pp. 1721–1727, IEEE, 1998.

Pham, C. K., Ramirez-Llodra, E., Alt, C. H., Amaro, T., Bergmann, M., Canals, M., Davies, J., Duineveld, G., Galgani, F., Howell, K. L., et al.: Marine litter distribution and density in European seas, from the shelves to deep basins, PloS one, 9, e95 839, 2014.

Purser, A.: A time series study of Lophelia pertusa and reef megafauna responses to drill cuttings exposure on the Norwegian margin, PloS one, 10, e0134 076, 2015.

Purser, A., Bergmann, M., Lundälv, T., Ontrup, J., and Nattkemper, T. W.: Use of machine-learning algorithms for the automated detection of cold-water coral habitats: a pilot study, Marine Ecology Progress Series, 397, 241–251, 2009.

Purser, A., Marcon, Y., Hoving, H.-J. T., Vecchione, M., Piatkowski, U., Eason, D., Bluhm, H., and Boetius, A.: Association of deep-sea incirrate octopods with manganese crusts and nodule fields in the Pacific Ocean, Current Biology, 26, R1268–R1269, 2016.

Purser, A., Marcon, Y., Dreutter, S., Hoge, U., Sablotny, B., Hehemann, L., Lemburg, J., Dorschel, B., Biebow, H., and Boetius, A.: Ocean Floor Observation and Bathymetry System (OFOBS): a new towed camera/sonar system for deep-sea habitat surveys, IEEE Journal of Oceanic Engineering, 44, 87–99, 2018.

Roark, E. B., Guilderson, T. P., Dunbar, R. B., Fallon, S. J., and Mucciarone, D. A.: Extreme longevity in proteinaceous deep-sea corals, Proceedings of the National Academy of Sciences, 106, 5204–5208, 2009.

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, e38 179, 2012.

Schoening, T., Kuhn, T., Jones, D. O., Simon-Lledo, E., and Nattkemper, T. W.: Fully automated image segmentation for benthic resource assessment of poly-metallic nodules, Methods in Oceanography, 15, 78–89, 2016.

Schoening, T., Jones, D. O., and Greinert, J.: Compact-morphology-based poly-metallic nodule delineation, Scientific reports, 7, 13 338, 2017.

Schoening, T., Köser, K., and Greinert, J.: An acquisition, curation and management workflow for sustainable, terabyte-scale marine image analysis, Scientific data, 5, 180 181, 2018.

Sharma, R.: Deep-Sea Mining, Springer, 2017.

Siddiqui, S. A., Salman, A., Malik, M. I., Shafait, F., Mian, A., Shortis, M. R., and Harvey, E. S.: Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data, ICES Journal of Marine Science, 75, 374–389, 2017.

Simon-Lledó, E., Bett, B. J., Huvenne, V. A., Köser, K., Schoening, T., Greinert, J., and Jones, D. O.: Biological effects 26 years after simulated deep-sea mining, Scientific reports, 9, 8040, 2019a.

Simon-Lledó, E., Bett, B. J., Huvenne, V. A., Schoening, T., Benoist, N. M., Jeffreys, R. M., Durden, J. M., and Jones, D. O.: Megafaunal variation in the abyssal landscape of the Clarion Clipperton Zone, Progress in oceanography, 170, 119–133, 2019b.

Simon-Lledó, E., Bett, B. J., Huvenne, V. A., Schoening, T., Benoist, N. M., and Jones, D. O.: Ecology of a polymetallic nodule occurrence gradient: Implications for deep-sea mining, Limnology and Oceanography, 2019c.

Taylor, J., Krumpen, T., Soltwedel, T., Gutt, J., and Bergmann, M.: Regional-and local-scale variations in benthic megafaunal composition
540   at the Arctic deep-sea observatory HAUSGARTEN, Deep Sea Research Part I: Oceanographic Research Papers, 108, 58–72, 2016.

Taylor, J., Krumpen, T., Soltwedel, T., Gutt, J., and Bergmann, M.: Dynamic benthic megafaunal communities: Assessing temporal variations
in structure, composition and diversity at the Arctic deep-sea observatory HAUSGARTEN between 2004 and 2015, Deep Sea Research
Part I: Oceanographic Research Papers, 122, 81–94, 2017.

Thiel, H.: Evaluation of the environmental consequences of polymetallic nodule mining based on the results of the TUSCH Research Asso-
545   ciation, Deep Sea Research Part II: Topical Studies in Oceanography, 48, 3433–3452, 2001.

Tilot, V., Ormond, R., Moreno Navas, J., and Catalá, T. S.: The benthic megafaunal assemblages of the CCZ (Eastern Pacific) and an approach
to their management in the face of threatened anthropogenic impacts, Frontiers in Marine Science, 5, 7, 2018.

Vanreusel, A., Hilario, A., Ribeiro, P. A., Menot, L., and Arbizu, P. M.: Threatened by mining, polymetallic nodules are required to preserve
abyssal epifauna, Scientific reports, 6, 26 808, 2016.

550   Volkmann, S. E. and Lehnen, F.: Production key figures for planning the mining of manganese nodules, Marine Georesources & Geotech-
nology, 36, 360–375, 2018.

Watling, H.: Review of biohydrometallurgical metals extraction from polymetallic mineral resources, Minerals, 5, 1–60, 2015.

Wilson, S., Graham, N., and Polunin, N.: Appraisal of visual assessments of habitat complexity and benthic composition on coral reefs,
Marine Biology, 151, 1069–1076, 2007.

555   Yamazaki, T. and Brockett, F. H.: History of Deep-Ocean Mining, Encyclopedia of Maritime and Offshore Engineering, pp. 1–9, 2017.

Yu, C., Xiang, X., Wilson, P., and Zhang, Q.: Guidance-error-based robust fuzzy adaptive control for bottom following of a flight-style AUV
with delayed and saturated control surfaces, IEEE Transactions on Cybernetics, 2018.

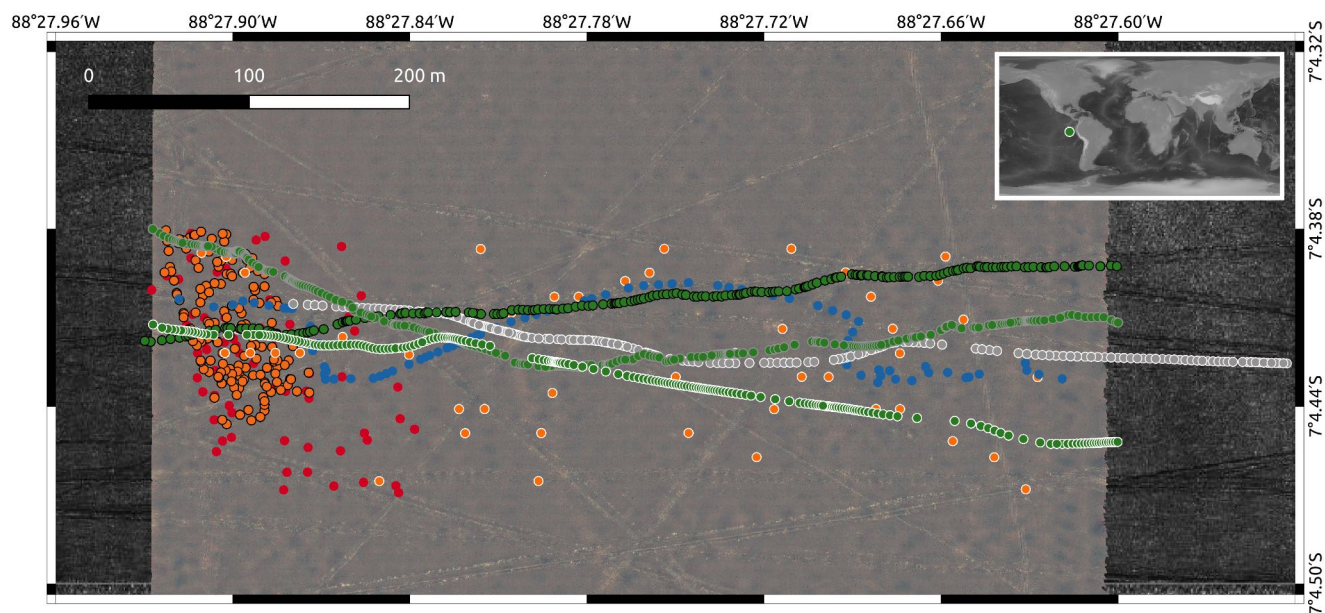



Figure 1. Overview map of imaging locations of the eight different datasets. DS_A (green dots, grey border), DS_B (green dots, black border), DS_C (blue dots), DS_D (green dots, white border), DS_E (orange dots, black border), DS_F (grey dots), DS_G (orange dots, white border), DS_H (red dots). The world map in the top right corner shows the geographical location of the DISCOL area in the Eastern South Pacific (green dot). The study area covers ca. 600x150 m^2. The background map shows another photo mosaic, created from the full image set of which DS_G is a subset. Criss-crossing lines are plough tracks by the mining simulation in 1989.

Table 1. Summary of image data collected for each dataset considered in this study. Columns marked by (*) represent median values across the dataset.

Dataset	Station	Date	Platform	Resolution* [MPix / m^2]	Altitude* [m]	Footprint* [m^2 / image]	Number of images
DS_A	SO242-2_171	25/09/2015	AWI OFOS	4.49	1.6	4.9	311
DS_B	SO242-2_155	25/09/2015	AWI OFOS	3.89	1.7	5.7	206
DS_C	SO106_OFOS35	1997	EXPLOS OFOS	1.05	3.4	12.5	80
DS_D	SO242-2_233	25/09/2015	AWI OFOS	0.98	3.2	22.5	209
DS_E	SO242-1_107	17/08/2015	AUV Abyss	0.24	4.2	52.9	154
DS_F	SO242-1_111	18/08/2015	Custom OFOS	0.16	2.0	2.6	272
DS_G	SO242-1_083	13/08/2015	AUV Abyss	0.07	7.5	169.1	46
DS_H	SO242-1_102 (Mosaic)	16/08/2015	AUV Abyss	0.04	4.5	32.8	62

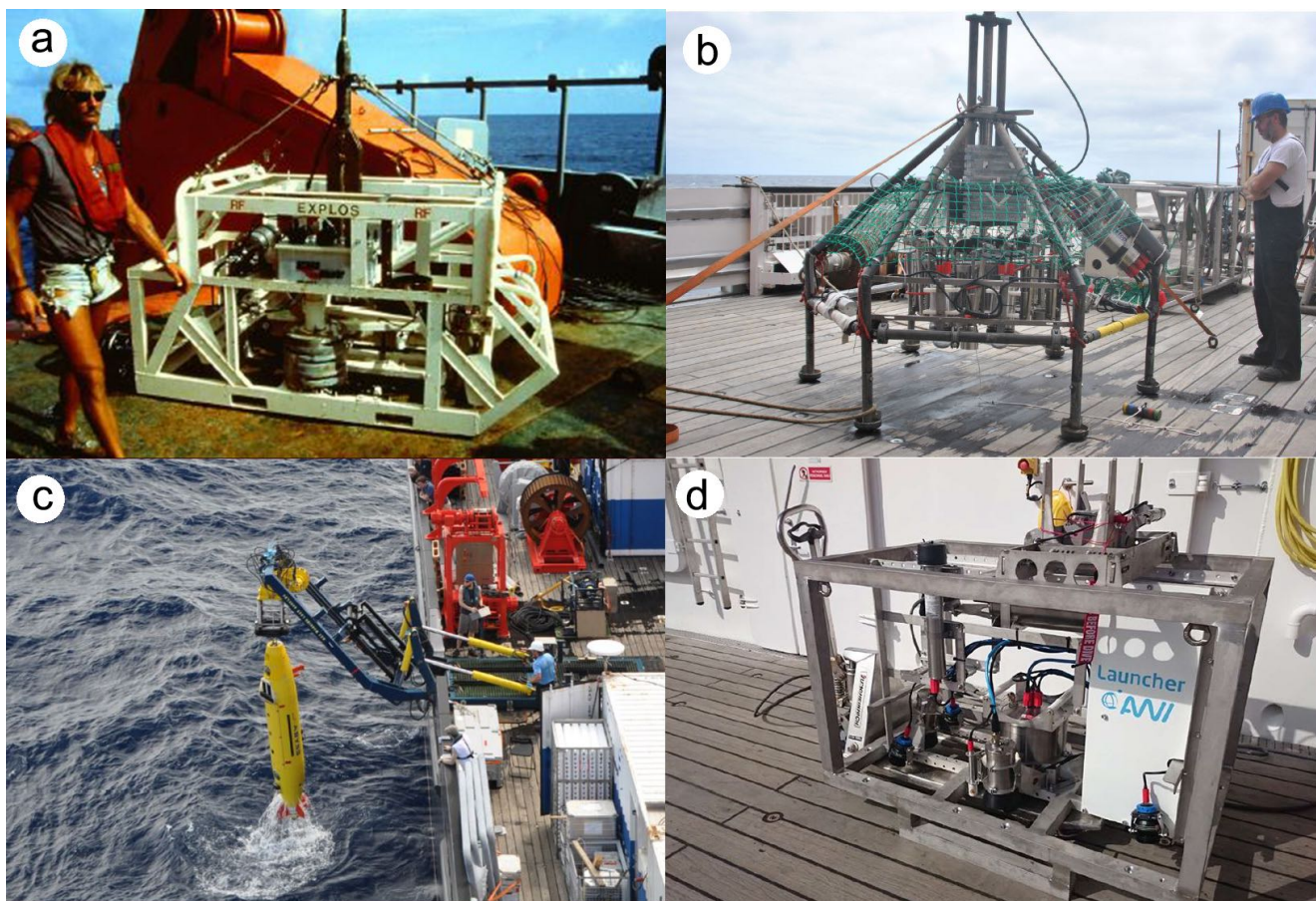
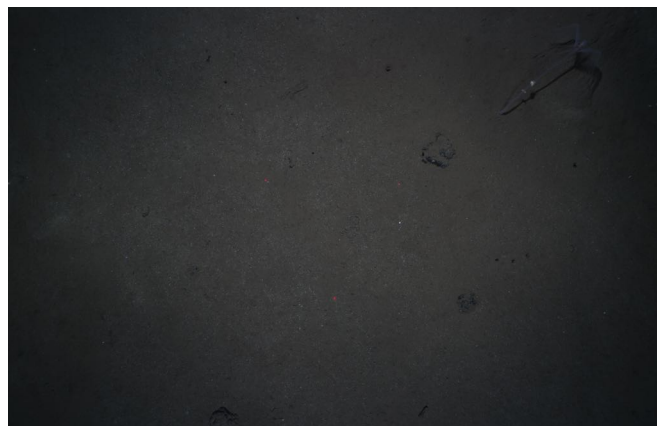



Figure 2. Imaging platforms used in the current study. a) The EXPLOS OFOS analog camera sled from 1997 b) A custom OFOS used during SO241/1 c) GEOMAR AUV Abyss d) AWI OFOS

Table 2. Annotation results for the eight different datasets considered in this study.

Dataset	No. annotations (grouped)	No. Categories found	Observer Agreement (single annotators)	Observer Agreement (grouped)	Cohens' Kappa (grouped)	Fauna Density [ind. / m^2]
DS$_A$	741	22	0.06	0.65	0.75	0.0194
DS$_B$	264	22	0.11	0.66	0.76	0.0092
DS$_C$	78	18	0.12	0.71	0.82	0.0085
DS$_D$	1077	22	0.14	0.66	0.81	0.0065
DS$_E$	231	22	0.20	0.69	0.82	0.0009
DS$_F$	70	15	0.24	0.40	0.49	0.0029
DS$_G$	61	13	0.16	0.65	0.74	0.0007
DS$_H$	202	22	0.23	0.66	0.77	0.0030

(a) DS$_A$

(b) DS$_B$

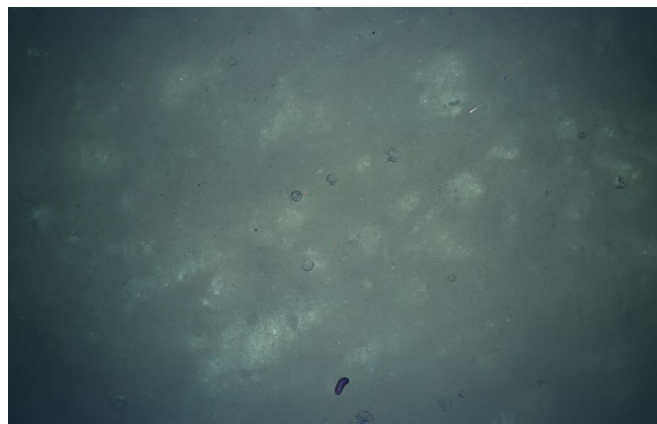
(c) DS$_C$

(d) DS$_D$

Figure 3. Example images of datasets DS$_A$ - DS$_D$.

(a) DS$_E$

(b) DS$_F$

(c) DS$_G$

(d) DS$_H$

Figure 4. Example images of datasets DS$_E$ - DS$_H$.

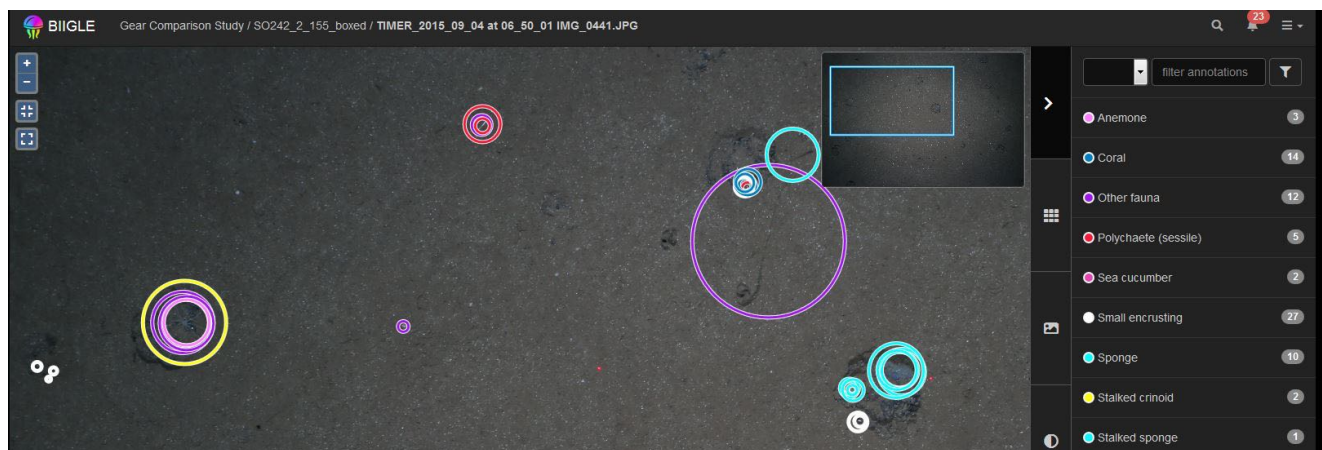



Figure 5. Circular fauna identifications made by an operator using the BIIGLE software application

Table 3. Kruskal-Wallis test assessment of whether differences in fauna abundance derived from the DISCOL seafloor data are significant for each fauna category used in the current study. P values of less than 0.05 indicate significance at the 95% percentile.

Fauna	H	N	df	p
Anemone	34.09	5	7	<0.001
Coral	34.63	5	7	<0.001
Crustacean	24.20	5	7	<0.001
Epifauna	33.61	5	7	<0.001
Ipnops fish	36.92	5	7	<0.001
Jellyfish	32.86	5	7	<0.001
Litter	25.68	5	7	<0.001
Mollusc	13.65	5	7	0.46
Other fish	29.09	5	7	<0.001
Polychaete mobile	27.14	5	7	<0.001
Polychaete sessile	35.16	5	7	<0.001
Sea Cucumber	23.73	5	7	<0.001
Sea Urchin	25.22	5	7	<0.001
Small encrusting	16.56	5	7	0.013
Spiral worm	25.37	5	7	<0.001
Sponge	32.011	5	7	<0.001
Stalked crinoid	35.54	5	7	<0.001
Stalked sponge	23.99	5	7	<0.001
Stalk no head	25.82	5	7	<0.001
Starfish	16.93	5	7	0.011

Figure 6. Fauna categories used in the current study for the DISCOL area.

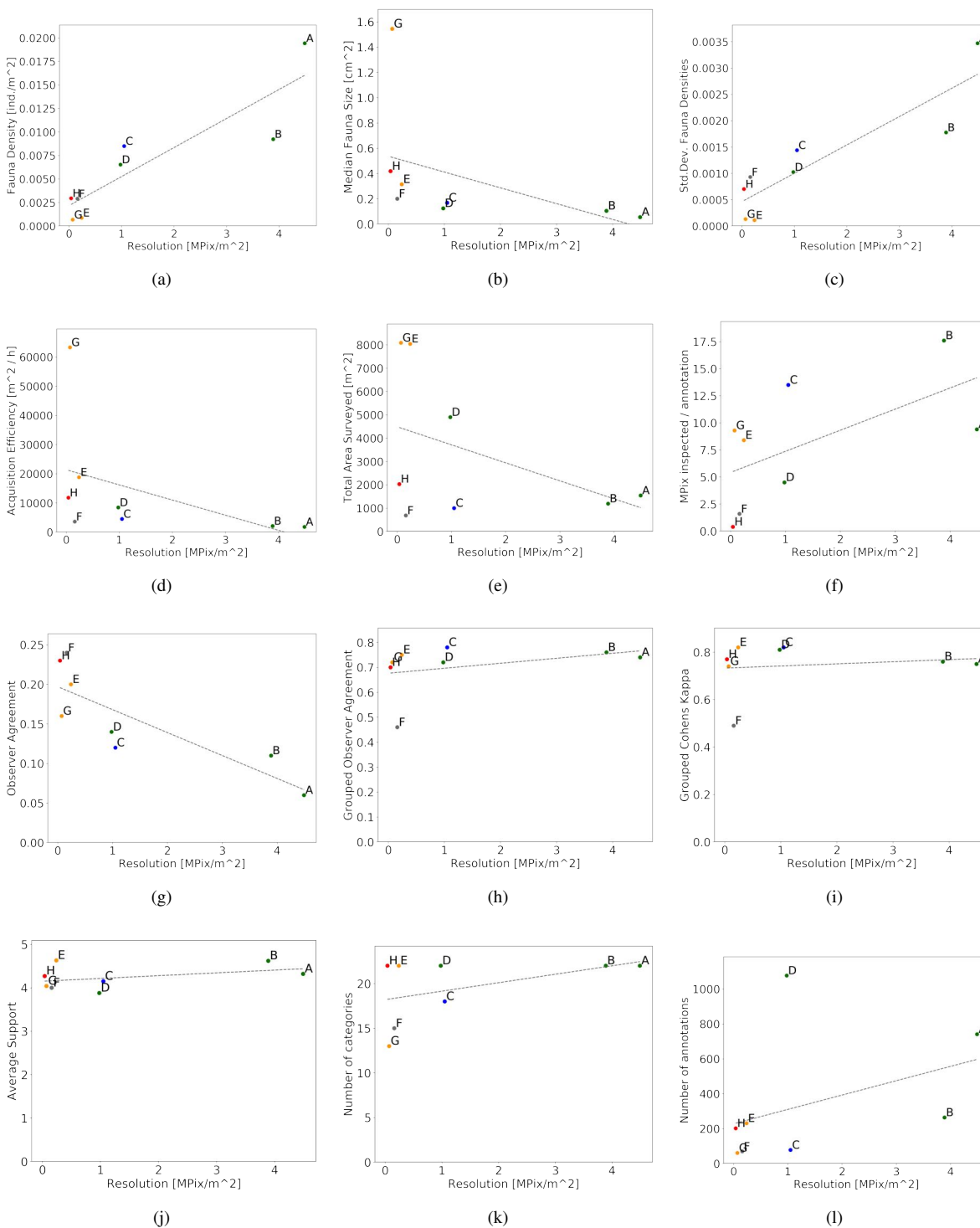



Figure 7. Aggregated results of fauna annotations for the eight datasets (dots A-H, green: AWI OFOS, blue: EXPLOS OFOS, grey: custom OFOS, orange: AUV Abyss, red: AUV Abyss mosaic). Dashed lines show linear regressions.

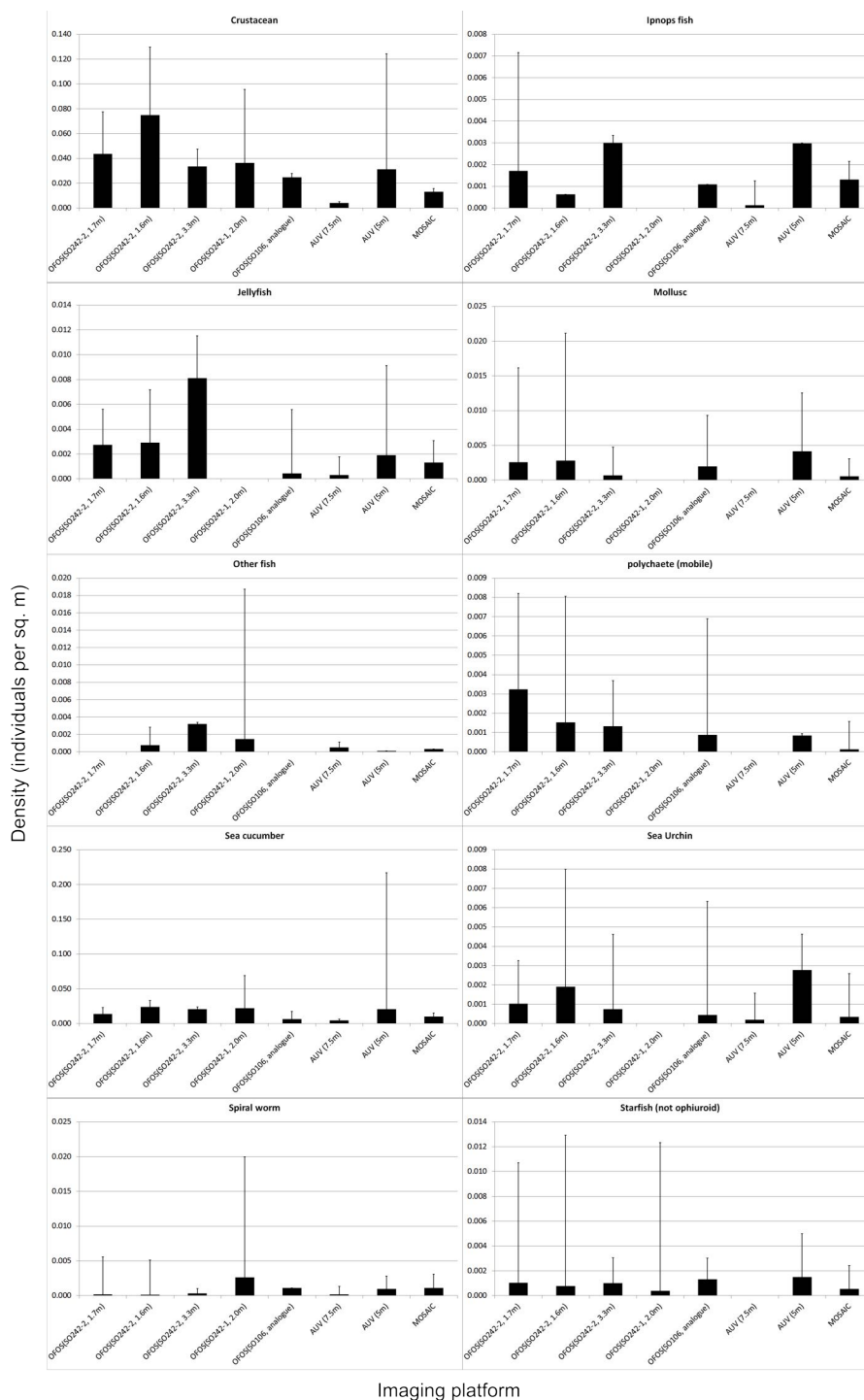



Figure 8. Mobile fauna abundances determined by 5 annotators independently annotating image data collected during the eight survey deployments.

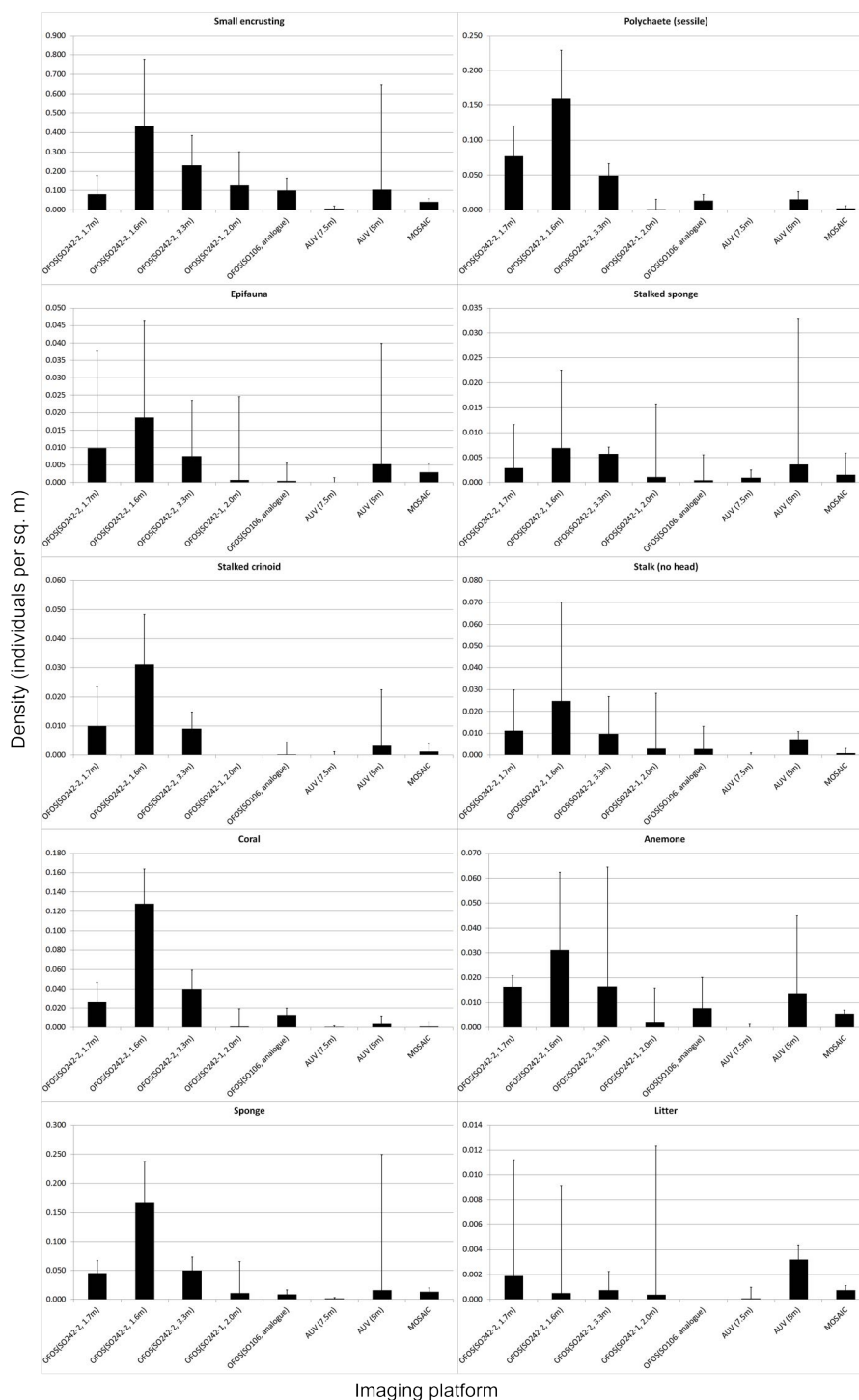



Figure 9. Sessile fauna abundances determined by 5 annotators independently annotating image data collected during the eight survey deployments.