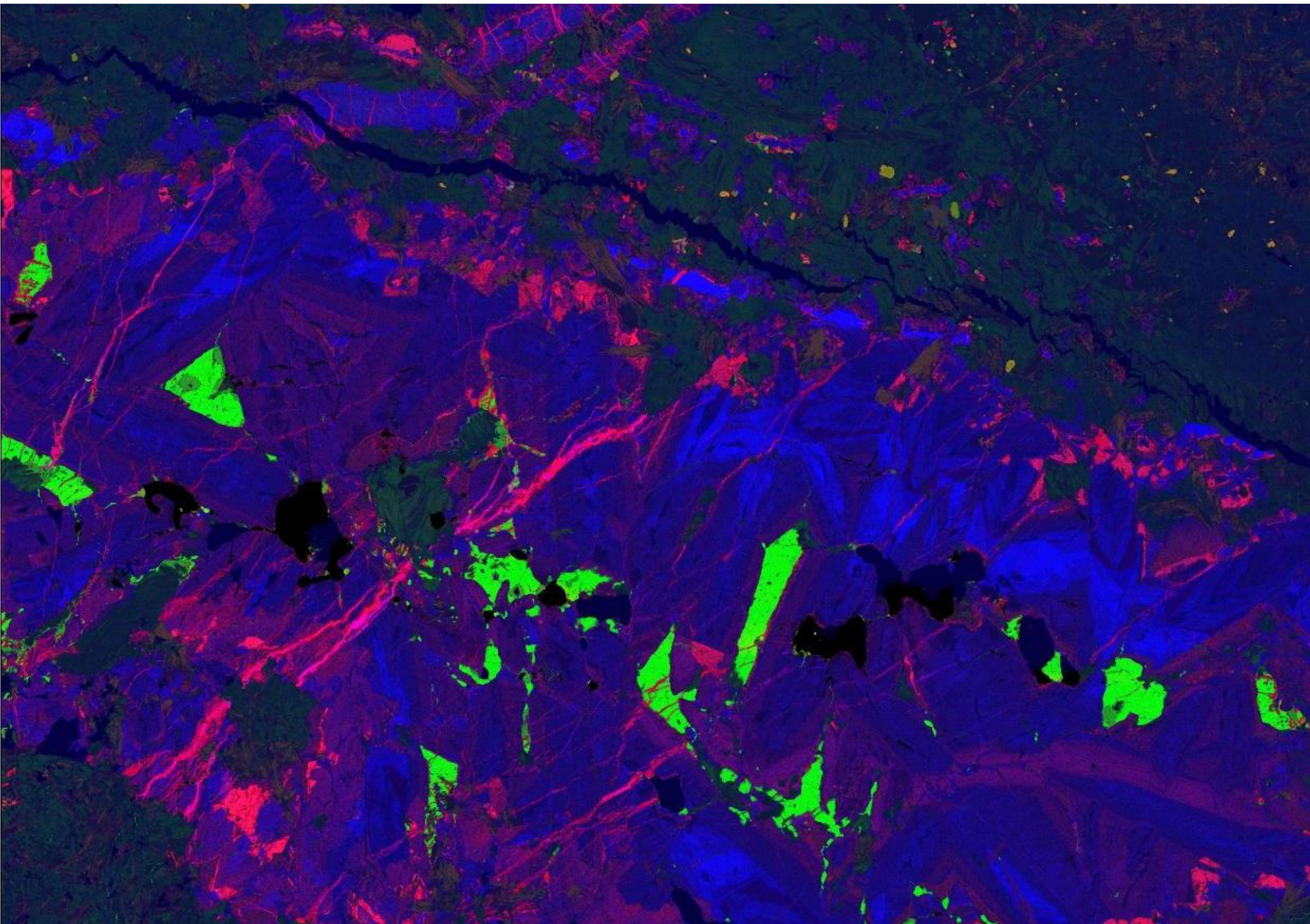
Spatially balanced designs for surveying and monitoring the impact of deep-sea mining on benthic biodiversity in NORI-D lease of the Clarion-Clipperton Zone

Project: Integrated Ecosystem Assessment and Ecosystem-Based Management Framework for Polymetallic Nodule Collection Activities

June 2025

El-Hachem M, Dunstan PK, Hosack GR, Hyman J, O'Hara TD, Parr JM, Schlacher TA, Woolley SNC, Dambacher JM



Prepared by the CSIRO Research Consortium for The Metals Company Australia Pty Ltd.

The 2021–2025 consortium, led by CSIRO, includes Museums Victoria, Griffith University, University of the Sunshine Coast and NIWA.

Citation

El-Hachem M, Dunstan PK, Hosack GR, Hyman J, O'Hara TD, Parr JM, Schlacher TA, Woolley SNC, Dambacher JM. 2025. *Spatially balanced designs for surveying and monitoring the impact of deep-sea mining on benthic biodiversity in NORI-D lease of the Clarion-Clipperton Zone*. CSIRO Marine Laboratories, Hobart.

Copyright

Important disclaimer

Acknowledgement of Country

CSIRO acknowledges the Traditional Owners of the lands, seas and waters, of the area that we live and work on across Australia and pays its respects to Elders past and present. CSIRO recognises that Aboriginal and Torres Strait Islander peoples have made, and will continue to make, extraordinary contributions to Australian life including in cultural, economic, and scientific domains.

Acknowledgements

The analysis and writing of this report was funded by The Metals Company Australia Pty Ltd through a contract to a research consortium that includes the Commonwealth Scientific and Industrial Research Organisation (CSIRO, Australia), National Institute of Water and Atmospheric Research (NIWA, New Zealand), Griffith University, University of the Sunshine Coast, and Museums Victoria (Australia).

Contents

Figures

Tables

1 Introduction

The assessment of the impact of deep-sea mining in the Clarion-Clipperton fracture zone (CCZ), specifically in NORI-D, requires an appropriate sampling methodology and a rigorous statistical analysis of the data before drawing conclusions that could inform the decision making (Hayes et al., 2019). The CCZ is a seafloor area where potential impacts of large-scale mining on faunal distribution and composition, and regional biodiversity are yet to be determined.

An appropriate statistical design helps ensure that the data collected and analysed answers the research questions under consideration, and represents the biological populations investigated. Further, an efficient design is rich in information so that uncertainty around inference is reduced as much as possible under budgetary and technical constraints of surveying (Przeslawski and Foster, 2024).

A survey is motivated by a research question that must be clearly defined. The primary research question here is to measure the impact of collection of polymetallic nodules on the fauna abundance, composition and distribution in NORI-D. The primary research question would require estimating the original abundance, composition and distribution of species in NORI-D, the immediate effect of nodules removal on the fauna, and the recovery rate and pattern distribution of the species after removal.

1.1 Motivation behind spatially balanced designs

A spatially balanced design could improve efficiency in natural resource sampling by maximising spatial independence between sample locations and by reducing variability in responses. By preventing samples from being placed too close together, spatially balanced designs reduce spatial autocorrelation, thereby minimising redundancy and maximising the unique information that each sample provides. Spatially balanced designs are more efficient than other random designs because they improve balance on environmental variables, ensuring that covariate means of the sample match the means of the population. Spatially balanced designs capture the spatial variability in the data more effectively than designs that lack spatial structure, thereby giving a better representation of the entire population and leading to more precise estimates of population parameters in statistical analysis.

We use the method of Balanced Acceptance Sampling (BAS) (Robertson et al., 2013) to ensure that randomly selected sample locations are spread out to maximise spatial independence and achieve even distribution across the area. The method, like all the methods that yield spatially balanced samples (Kermorvant et al., 2019), is advantageous if there is any spatial trend in the response. Examples of response in our study include the abundance and the composition of the species.

1.2 Concepts of spatially balanced survey designs

The method BAS is an extension of the Generalised Random-Tesselation Stratified (GRTS) design (Stevens and Olsen, 2004). GRTS maps two-dimensional space into a one-dimensional space. The method BAS improves spatial balance by allowing balancing over multiple dimensions or multiple factors (Brown et al., 2015). BAS selects spatial units following user-specified target inclusion probabilities. Unequal inclusion probabilities imply that each spatial unit has an unequal probability of selection.

Unequal inclusion probabilities could be used as a way of adding a third dimension to a two-dimensional representation. For example, we could intensify surveying depending on the depth associated to each spatial unit of the seafloor. Unequal inclusion probabilities could be used as a way of implementing strata without hard boundaries, where units can be sampled with different probabilities based on their characteristics rather than belonging to one stratum or one categorical division.

Using unequal inclusion probabilities does not make the design spatially unbalanced because the layer with higher inclusion probabilities only result in higher survey intensity (Brown et al., 2015). Altering inclusion probabilities is useful when covariates guiding inclusion probabilities are related to the sampling variable, but using too many covariates can reduce efficiency, and maintaining equal inclusion probabilities is often a more reliable and robust approach.

We use the implementation of the method BAS (Robertson et al., 2017) from the R package MBHdesign (Foster, 2020) for the examples presented in the document. The package allows also for the use of transects (Foster et al., 2019) in the design and for the inclusion of legacy sites (Foster et al., 2017).

1.3 Important considerations in the context of NORI-D

NORI-D is a rectangular lease in the CCZ, lying between 3200 and 4600 metres deep, with topographic features of seamounts, abyssal hills, troughs and ridges. Food is considered limited in the abyssal seafloor because the primary source of food is Particulate Organic Carbon (POC) which sinks from the upper ocean and decreases in concentration with water depth (Smith et al., 2008).

Topographic variations, POC flux, sediment characteristics and nodule distribution influence fauna densities and structure of communities (Washburn et al., 2021). Relevant topographic features derived from bathymetry are slope and bathymetric position index (BPI) (McQuaid et al., 2020). Gradients of environmental factors across CCZ, such as POC flux and topography features, may be considered for spatial trend in a survey design, but the habitats where some species may be found are not characteristics of the whole range of biodiversity in CCZ. For example, fauna could be buried at different layers of sediments, while sessile organisms may be attached to nodules (Gaikwad et al., 2024). Both hills and plains have varying abundance of species but sloping topography and enhanced water motion (Stefanoudis et al., 2016) could also influence abundance. It is known that the flux of PCO decreases from east to west and south to north in CCZ (McQuaid et al., 2020) but the change in flux may not be remarkable in NORI-D. Moreover, the flux is subject to oceanic currents (Smith et al., 2008) and may vary with complex weather patterns.

Important questions must be answered for a successful planning of balanced surveys specific to NORI-D. What is the metric used (count of species, mass, abundance, presence/absence)? What are the resources available (collection methods, gear, protocols) for surveying the area? Jamieson et al. (2013) describe methods and techniques available to sample the deep seafloor. Different collection methods, such as multicorers and optical underwater imaging, are going to accumulate different species and abundance data (Gaikwad et al., 2024) and should complement each other. Finally, time and budget constraints are limiting factors that impact the survey design.

TMC provided CSIRO with details of the proposed tracks of initial mining operation, called Project Zero. While these planned operations have not yet had regulatory approval, they nonetheless can be used for the identification of monitoring sites should approval be granted.

In this document, we explain how to design spatially balanced samplings in the context of measuring the impact of collection of polymetallic nodules on the fauna abundance, composition and distribution in NORI-D. We give examples of spatially balanced samplings depending on the type of deployments and the number of voyages.

2 Spatially balanced designs with unequal inclusion probabilities

2.1 Balanced Acceptance Sampling

Various methods and algorithms can be applied to achieve a spatially balanced design. In ecological or spatial survey, the stratification is combined with a spatially balanced sampling method. Recent methods developed to achieve a spatially balanced design include the Local Pivotal Method (Grafström et al., 2011) and **Balanced Acceptance Sampling (BAS)** (Brown et al., 2015), both designed to reduce variance in the sampling. Both methods allow balancing over multiple dimensions or multiple factors and the use of unequal inclusion probabilities. The method used in this document is BAS implemented in a software developed by CSIRO (Foster, 2020). Many software tools based on different methods are available as summarised in the review by Kermorvant (2019).

The BAS method relies on a Halton sequence to ensure that the selected locations are spatially dispersed. A Halton sequence is a deterministic sequence with low discrepancy that appears random. The "random" part is generated by skipping a random number of locations at the beginning of the sequence, in each dimension. A spatial cell or unit is selected following user-specified target inclusion probabilities or target inclusion density function.

A reliable way to survey a region where the spatial tendency of the response variable is unknown is to use **equal inclusion probabilities**, where each unit has the same probability to be selected. In the case of NORI-D, we could argue that features obtained from bathymetry, flux of PCO and presence of nodules are all indicative of abundance of different species and should be used to determine strata and subsequently to assign unequal inclusion probabilities to each unit. Altering inclusion probabilities is useful when covariates guiding inclusion probabilities are related to the sampling variable, which is the case here, but using too many covariates can reduce statistical efficiency (Foster et al., 2024). Equal inclusion probabilities is a reliable approach if we are looking at studying the abundance, distribution and diversity of species that could correlate positively and negatively with bathymetry features, PCO flux, and sediment type.

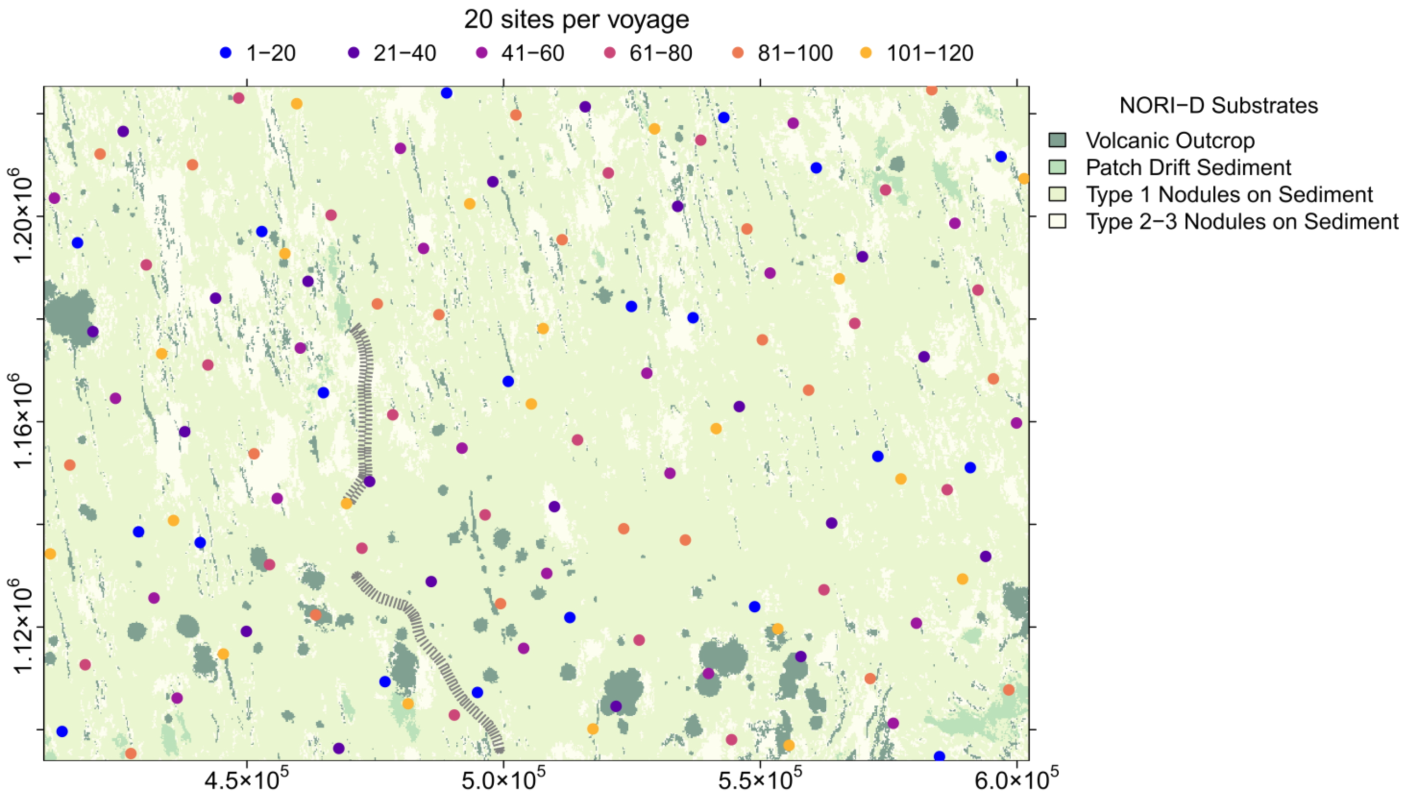


Figure 1 Master sample accommodating up to six voyages of 20 locations each obtained from equal inclusion probabilities

The spatially balanced design is obtained using the function quasiSamp.raster from the package MBHdesign with equal inclusion probabilities. The legend on the top links each voyage with 20 locations represented by discs of the same colour, where the centre of the disc is the location of the deployment. The map, supplied by The Metals Company, shows four types of substrates in four shades of green, as indicated by the legend on the right. The lines in grey represent the mining path in Project Zero, where polymetallic nodules of Type 1 may be collected. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611). The resolution of the map is of 1315 by 1919, where each cell measures 100 metres by 100 metres.

We show an example of a master sample spatially balanced in NORI-D in Figure 1, on a map of substrates provided by The Metals Company (2024). The polymetallic nodules on sediment are of types 1, 2 and 3 depending on their composition, their size and their shape. Linear volcanic features are growth faults that form scarps, and more massive volcanic features are volcanic knolls. Figure 1 shows the sampling locations obtained with BAS method, as implemented in the R package MBHdesign (see Appendix A1 for the R code), by using equal inclusion probabilities. The selection of locations in the sampling ignores the type of substrates (nodules sediment, volcanic outcrop, etc.) and the proposed design for the paths of the collector of polymetallic modules, or the mining paths. In the example shown, we assumed that six voyages are conducted through the project, and that each voyage consists of 20 deployments for benthic sampling (see Appendix A2 for the coordinates). Figure 1 warrants two specific comments. First, the design does not cover the blocks of the mining paths, illustrated in light grey in Figure 1. Second, the sampling locations were obtained using a map of 1315 by 1919 cells or spatial units, where each cell measures 100 metres by 100 metres, but we have obtained similar results, meaning similar positions and order of locations, where the resolution of the map is reduced as the size of each cell is increasing, and when the inclusion probability are equal for all map units.

2.2 How to draw smaller samples from a master sample

Six voyages of 20 deployments, as shown in Figure 1, may be aspirational and the access of all locations in each voyage may imply logistical hurdles, with some locations being inaccessible. A spatially balanced design obtained by BAS has the following property allowing scientists to create smaller subsets as needed: **a contiguous subset** of a spatially balanced design is also spatially balanced (van Dam-Bates et al., 2018). It is possible to generate a spatially balanced design by creating a large, spatially balanced **master sample** from which smaller samples can be drawn, depending on the number of deployments and the number of voyages feasible.

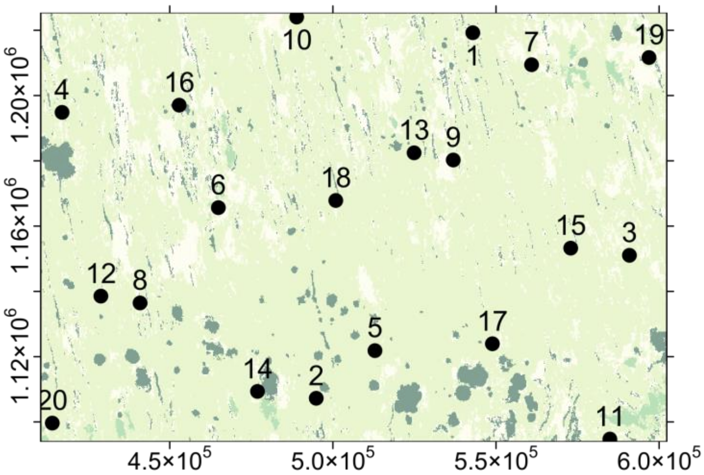


Figure 2 Contiguous sequence of selected locations in an example of spatially balanced design of 20 locations Spatially balanced design where the ordered list of locations 1-20 is obtained from the master sample in Figure 1. The order in the sequence is only pertinent when drawing a smaller sample from the sets of locations shown. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611).

The set of locations of a spatially balanced design is generated by MBHDesign as an ordered list, where the order is important when selecting a subsample. Figure 2 shows the first 20 locations from the master sample shown in Figure 1 where the number shown indicates the order in which the locations are generated or selected by the method BAS. The order number or rank in the sequence is useful to draw a contiguous number of locations that constitute a subset of a larger set of locations. For example, if we were planning three voyages with 20 deployments per voyage, three **contiguous sequences** of 20 locations that could be selected from the master sample are: 1-20, 21-40 and 41-60, or locations 2-21, 22-41 and 42-61, or locations 3-22, 23-42 and 43-62 or any contiguous sequence of 60 locations from the master sample.

2.3 Unequal inclusion probabilities

Unequal inclusion probabilities allow units to be selected with different probabilities based on one or many characteristics. We want to characterise each unit according to whether it has been mined. One approach is to define two strata, such as a “directly impacted” layer and an “indirectly impacted or untouched” layer. Limiting one stratum to the projected mining paths and immediate neighbouring cells may result in no locations being selected in the defined layer or may require inclusion probabilities in the “directly impacted” layer that are high compared to the inclusion probabilities in the rest of NORI-D.

Our approach is to calculate inclusion probabilities according to the distance to “directly impacted” layer, to smooth the difference between the inclusion probabilities in each layer. We calculate the distance to the directly impacted area, which is proximal to the mining tracks. **The tracks** correspond to two blocks of paths where the mining vehicle is planned to collect the nodules during Project Zero. The Metals Company has provided the geographical coordinates of the planned paths.

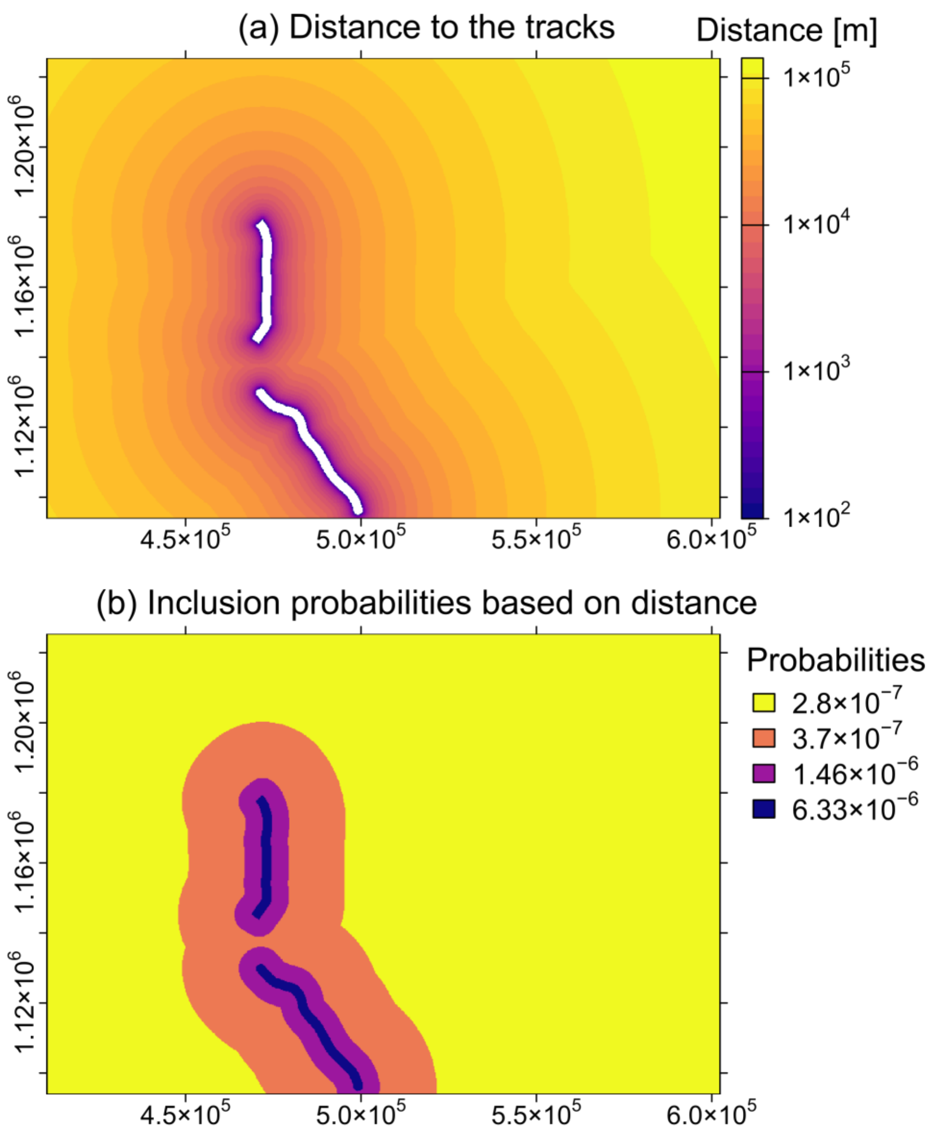


Figure 3 Inclusion probabilities obtained from distance to the tracks
(a) The distance to the tracks is calculated and shown on a logarithmic scale for the cells of NORI-D map. (b) The inclusion probabilities are calculated by dividing the studied region of NORI-D into categories, depending on the distance to the tracks. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611). Each cell measures 100 metres by 100 metres.

The distance between each cell of the map of NORI-D region and the tracks is calculated considering a resolution where each cell measures 100 metres by 100 metres. Figure 3 shows calculated distances to the tracks for all cells of NORI-D map. Yellow and orange colours indicate that the cells are at a distance of at least 10000 metres from the tracks. Purple colour identifies the cells closer to the tracks, at a distance between 100 metres to 1000 metres. Blank cells forming two thick paths are intersecting with the tracks. A distance of zero is attributed to each cell that intersects with the tracks.

The distribution of distances shown in Figure 3(a) is segmented into seven categories. The first three categories range from the distance zero to the first 20-percentile of the distances associated with the cells: $\{\{0, 100\}, \{100, 5000\}, \{5000, 21098.34\}\}$. The remaining categories correspond to the subsequent four 20th percentiles of cells, according to their distance to the tracks. The inclusion probabilities are computed to ensure an equal number of samples are selected from each category. The last four categories represent the 40th, 60th, 80th, and 100th percentiles of cells, each containing the same number of cells and thus resulting in the same inclusion probability being attributed to every cell within those categories, all represented in yellow in Figure 3(b).

The inclusion probabilities are inversely proportional to the number of cells in a category, meaning that a cell in a small area has more chance to be selected than a cell in a larger area. The high inclusion probabilities assigned to cells of small regions give them a high chance to be selected, (see the regions in orange and light purple in Figure 3(b) as well as the mining tracks shown in dark purple). The inclusion probabilities are then rescaled so that the sum of inclusions probabilities over all the cells is one.

In summary, a spatially balanced design is a sensible choice for surveying and monitoring the impact of mining on species diversity and abundance in NORI-D, because it allows spatial balance on a vast region with spatial heterogeneity. Fixing unequal inclusions probabilities following choices of covariates related to species distribution, such as bathymetry, PCO flux and nodules presence, would be pertinent if the response variable of the functional group of species studied is clearly related to the covariates chosen. In this work, a multitude of species and functional groups are under investigation, each answering differently to each covariate, therefore a spatially balanced design with equal inclusions probabilities is the way to maximise spatial coverage over all the covariates. Moreover, in the context of studying the mining impact, unequal inclusions probabilities based on distance to the tracks are a way of dividing the region into different layers of impact following mining.

3 Example of design for benthic sampling

Physical sampling techniques of ocean benthic floor deployed in NORI-D would be box corers or multicorers. A box corer (Kaiser et al., 2024) digs deep into the sediment layer while the multicorer collects multiple sediment samples simultaneously by minimising disturbance to the sediment layers (Gaikwad, 2024).

We present an example of spatially balanced design that could accommodate different scenarios of voyages with deployments of box corers or multicorers. We consider voyages with low, medium, and high number of deployments, corresponding to 8, 15 and 20 locations, respectively. The categories “low,” “medium” and “high” are not absolute, we use the categories as an example of constraints that could arise from budget and technical limitations. The grouping of locations presented in this section are obtained from the master sample presented below.

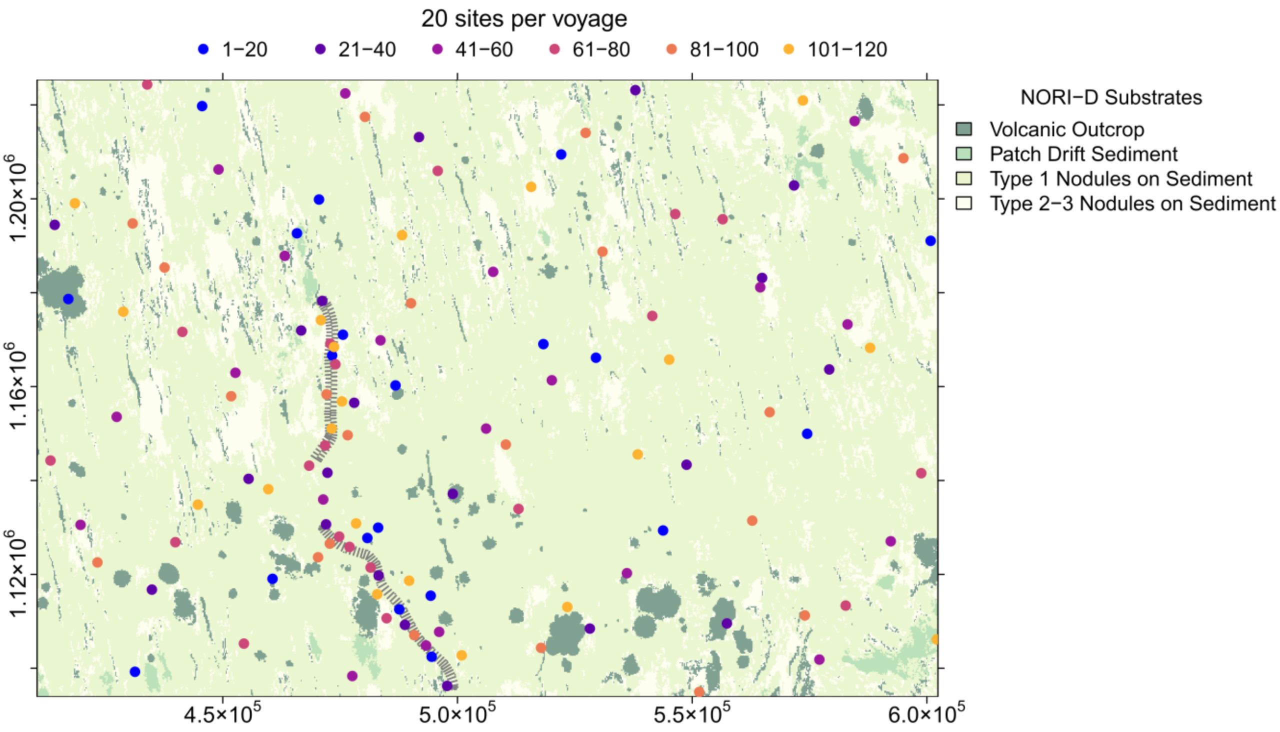


Figure 4 Master sample accommodating up to six voyages of 20 locations each obtained from unequal inclusions probabilities

The spatially balanced design is obtained using the function quasiSamp.raster from the package MBHdesign with the inclusion probabilities shown in Figure 3. The legend on the top links each voyage with 20 locations represented by discs of the same colour, where the centre of the disc is the location of the deployment. The map, supplied by The Metals Company, shows four types of substrates in four shades of green, as indicated by the legend on the right. The lines in grey represent the mining paths in Project Zero. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611). The resolution of the map is of 1315 by 1919, where each cell measures 100 metres by 100 metres.

We start by generating a master sample using unequal inclusion probabilities in the BAS method to define strata according to the distance to the tracks. The function quasiSamp.raster from the package MBHdesign gives an ordered list of 120 points, as requested, for the map of 1315 by 1919 spatial cells, where each cell measures 100 metres by 100 metres as shown in Figure 4 (see

Appendix A1 for R code). The polymetallic nodules on sediment are of types 1, 2 and 3 depending on their composition, their size and their shape. Linear volcanic features are growth faults that form scarps, and more massive volcanic features are volcanic knolls.

Contrarily to when we had equal inclusion probabilities, we do not obtain similar results when we increase the size of the spatial cell and diminish the resolution of the map, because the features of the map such as the shape of the mining tracks are degraded by pixelation. For example, a map resolution of 1315 by 1919 yields 16 locations within the tracks, illustrated in grey in Figure 4, while a resolution of 83 by 120 yields five locations that are really within mining tracks, due to the tracks being deformed by pixelation. The dependence of the design with unequal inclusion probabilities on the chosen resolution is important when the collected data in one location must be extrapolated to the whole cell where the collection is done. The size of the cell represented by the collected data varies with the resolution of the map used for the design with unequal inclusion probabilities.

The three scenarios presented in the next section are obtained by selecting a contiguous sequence of locations in the master sample from Figure 4. Each scenario is shown for six voyages, scheduled with respect to Project Zero, that consists into mining two blocks of paths illustrated in grey in Figure 4 and in the following figures. The first block or path to be mined is near the south border of NORI-D and is called “South path” in the figures. The second block or path is situated northernly to the first block and is named “North path.” The paths illustrated in grey are then coloured in red, once they are considered as mined.

We denominated the voyages according to how they should be distributed in time, as follows:

1. Baseline
2. Before mining South path (grey)
3. After mining South path (red)
4. Before mining North path (grey)
5. After mining North path (red)
6. After Project Zero

The baseline survey examines the natural state of the region before any mining disturbance. The survey before mining the South path would ideally measure the state of the region before the mining by specifically including at least a location in the South path. The voyage “After Project Zero” consisting into monitoring after Project Zero should occur regularly, to follow the evolution in time of the recovery after the mining impact. We show one voyage “After Project Zero” to unclutter the figures. It is obvious that more than one voyage “After Project Zero” is required to follow the evolution of how the mined region recovers.

The locations selected are obtained from contiguous sequences of 48, 90 and 120 locations from the master sample in Figure 4. The order is which the locations are positioned in the master sample is the order in which the locations are allocated to the voyages. The reader must not confuse the spatially balanced design given in example with a Before-After Control-Impact (BACI) design (Urban et al., 2022). If locations scheduled after impact of the collector intersect with the mining paths, it is only because surveying frequency have been intensified on the mining tracks with the help of unequal inclusion probabilities.

3.1 Low number of deployments

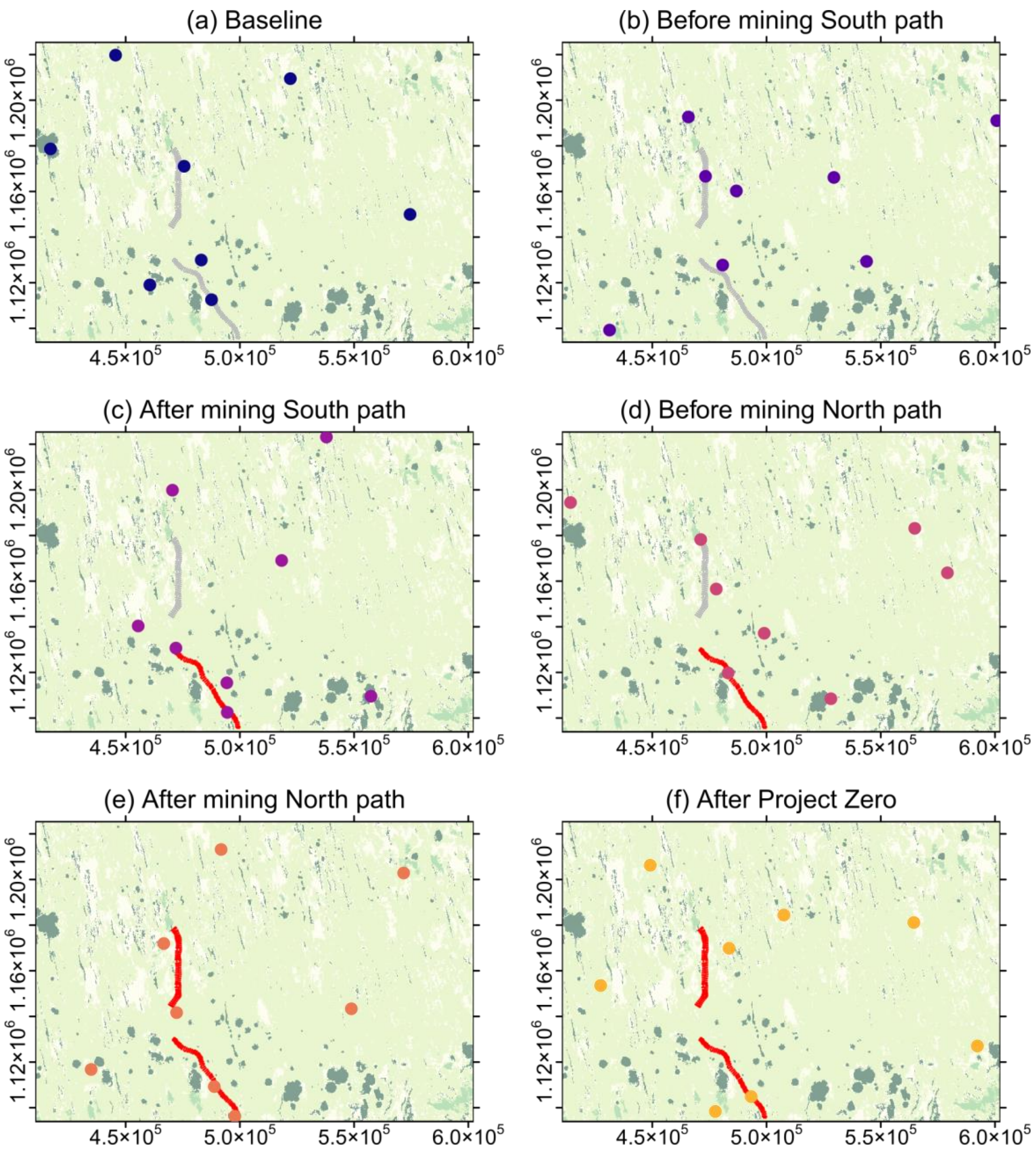


Figure 5 Spatially balanced sampling with eight deployments per voyage
The sites in voyages 1 to 6 are contiguous subsets from the master sample. Each voyage is represented by discs of the same colour, where the centre of the disc is the location of the deployment. Thick lines represent the mining paths: pre-mining (planned) paths in grey, post mining (actual) in red. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611).

Figure 5 shows a spatially balanced design of 48 locations dispatched over six voyages, equivalent to eight deployments per voyage. The low number of locations per voyage may look sufficient because it ensures that at least one deployment occurs in each of the following layer: the tracks, the immediate neighbourhood of the tracks and the far field of the tracks. However, the locations are too sparse to cover characteristics of NORI-D that vary in space: substrates type, PCO flux and bathymetry. Locations on volcanic outcrop appears only in two voyages, in Figure 5(a) and (c), no sampling locations are found on patch drift sediment.

3.2 Medium number of deployments

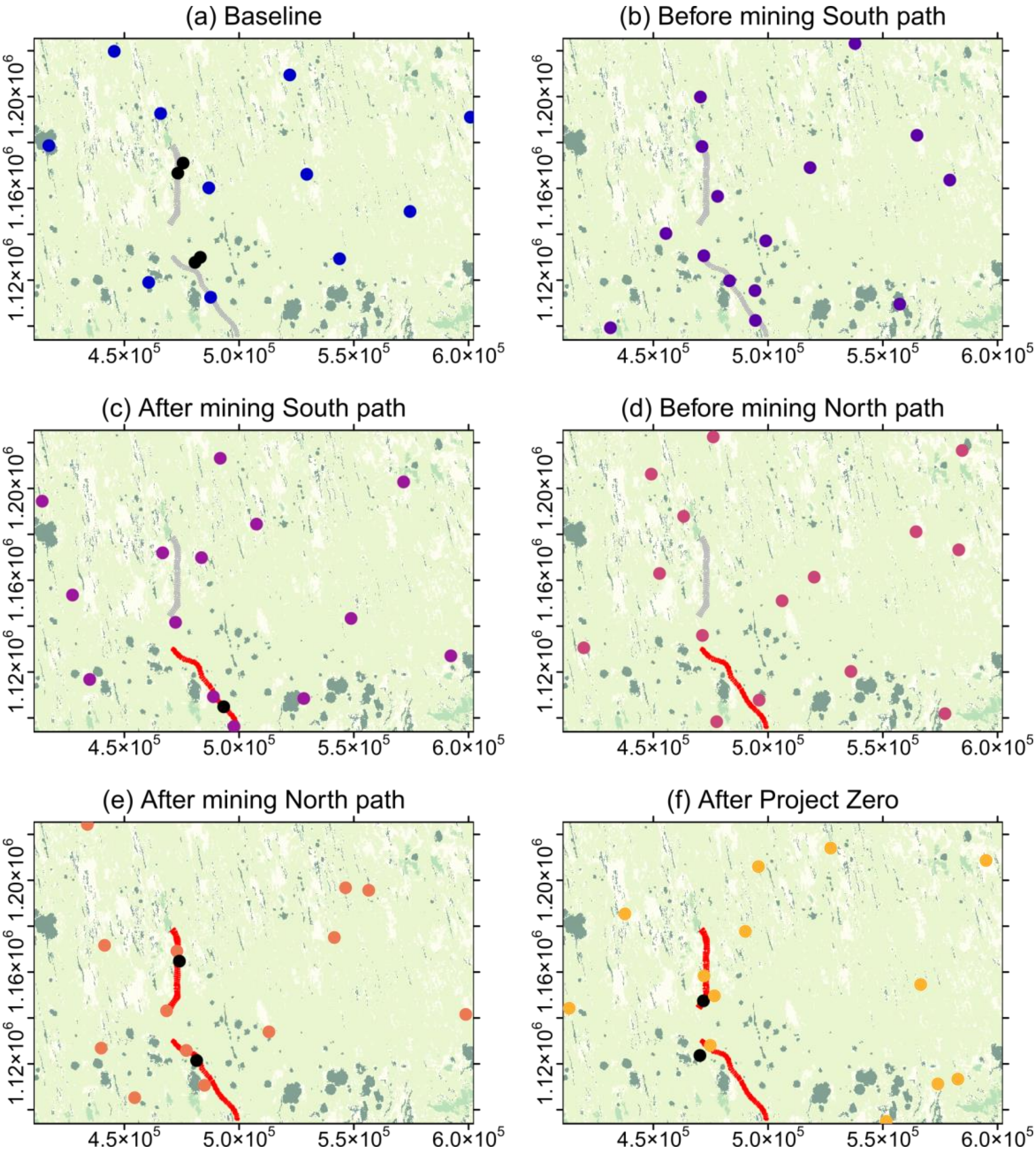


Figure 6 Spatially balanced sampling with 15 deployments per voyage

The sites in voyages 1 to 6 are contiguous subsets from the master sample. Each voyage is represented by discs of the same colour and discs coloured in black, where the centre of the disc is the location of the deployment. Thick lines represent the mining paths: pre-mining (planned) paths in grey, post mining (actual) in red. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611).

Figure 6 shows a spatially balanced design of 90 locations divided in six voyages, equivalent to 15 deployments per voyage. Deployments that are too close to each other such as deployments in the tracks and in the neighbourhood of the tracks, shown as black discs in Figure 6(a), (c), (e) and (f), could be switched between voyages or postponed to voyages where the paths have been mined, because the order in which the deployments are physically carried out does not impact the spatial balance of the design.

3.3 High number of deployments

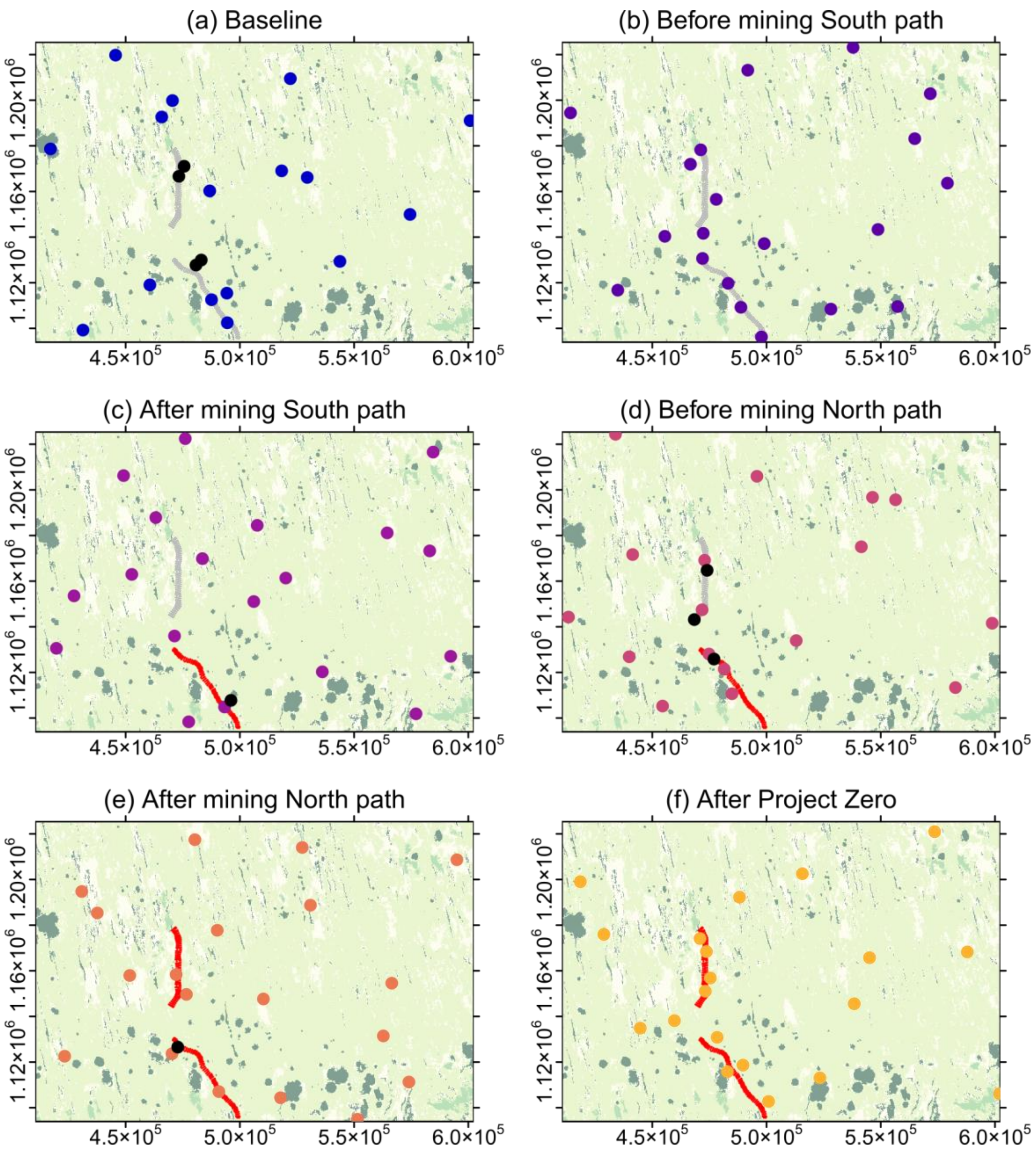


Figure 7 Spatially balanced sampling with 20 deployments per voyage
The sites in voyages 1 to 6 are contiguous subsets from the master sample. Each voyage is represented by discs of the same colour and discs coloured in black, where the centre of the disc is the location of the deployment. Thick lines represent the mining paths: pre-mining (planned) paths in grey, post mining (actual) in red. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611).

Figure 7 shows a spatially balanced design of 120 locations divided in six voyages, equivalent to 20 deployments per voyage. The high intensity of sampling associated with the tracks, where the cells have inclusion probabilities approximately 30 times higher than inclusion probabilities of the far field is the reason the tracks are mostly entirely covered by cumulated sampling locations through the six voyages. Deployments that are too close to each other, shown as black discs as in Figure 7(a), (c), (d) and (e), could be postponed in the timeline of voyages. For example, locations shown in black in Figure 7(a) could be postponed to voyages in Figure 7(c) or (e). Locations in black in Figure 7(c)--(e) could be rearranged between the three voyages, to be conducted after the paths have been mined.

4 Example of design for ROV transects

Transects are used to represent samples when surveying is made through optical underwater imaging, where a Remotely Operated Vehicle (ROV) moves a camera along a designed path, collecting images and videos of the sea floor. Imagery allows for a non-invasive surveying and monitoring of megabenthic fauna and characteristics of undisturbed seafloor (Mbani et al., 2023).

We present an example of spatially balanced design that could accommodate different scenarios of number of voyages with deployments of ROV that capture underwater images or videos, following a path defined by transects. We consider voyages with low, medium and high number of deployments per voyage, corresponding respectively to 2, 4 and 6 deployments of ROV. As we mentioned in the previous section, the categories "low," "medium" and "high" are not absolute, we use the categories as an example of constraints that could arise from budget and technical limitations. The configurations of transects presented in this section are obtained from the master sample presented below.

4.1 Transects pattern

A transect consists of several points arranged in a certain pattern, a line for example. While some patterns are to be avoided, such as circular paths, there is no consensus on which transect path to prioritise (Foster et al., 2014).

In the following example, we implemented a zigzag pattern to demonstrate an alternative way to define a path of transects. In the documentation of the R package MBHdesign, a transect is not always a line, it can be any narrow pattern finite in length that covers more than one point, or cell, and that is defined by its centre location.

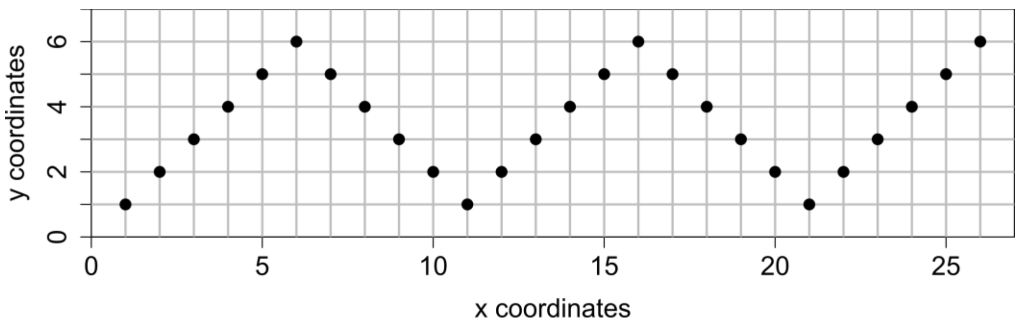


Figure 8 Pattern used for transects in example of survey design
Each planned deployment in the survey design of transects will follow a path in the shape of the zigzag illustrated, made of five segments. The grid is rescaled so that each unit in x-axis and each unit in the y-axis is equal to √80000 ≅282.84 metres, such that the length of a segment is 2000 metres.

The zigzag pattern is illustrated in Figure 8 and is inspired by the pattern used in the technical report of baseline sampling at the intention of TMC (Simon-Lledó et al., 2023). Each diagonal segment made of six points measures 2000 metres, such that each vertical or horizontal grid unit measures $\sqrt{80000} \cong 282.84$ metres. The pattern in Figure 8 can be rotated in any direction, and we chose to set the number of possible directions to 15, to limit the computational resources needed. Therefore, the pattern could be rotated by 24 degrees at minimum, at once. The number of points chosen for the designed pattern in Figure 8 is 26 points, because it is recommended (Foster et al., 2019) that the resolution of the path allows it to have at least one point in each cell, to accurately account for the inclusion probabilities.

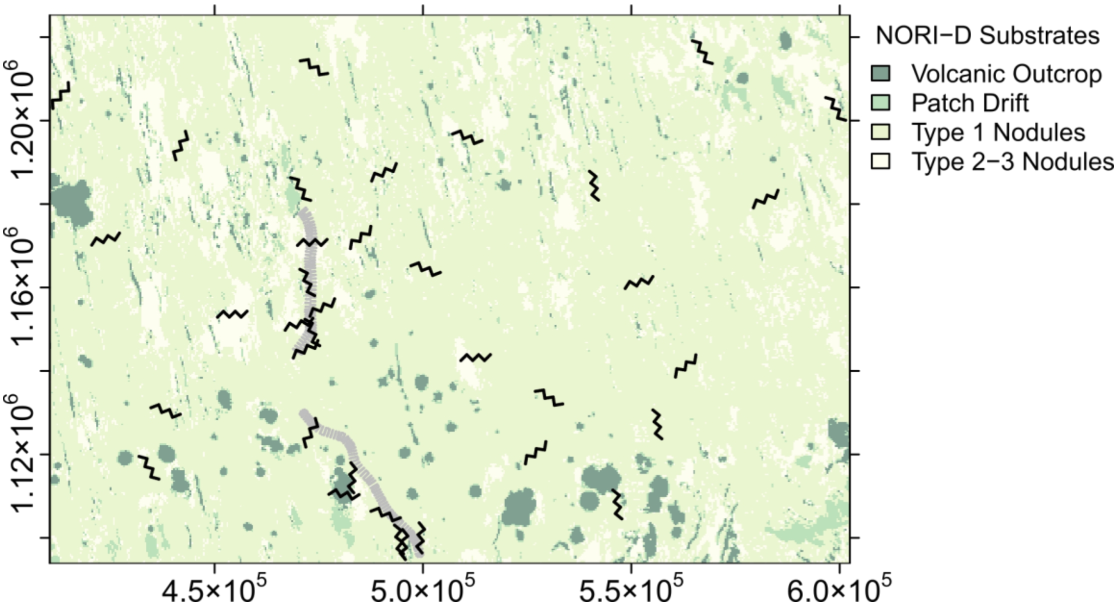


Figure 9 Master sample of transects obtained with unequal inclusion probabilities
The spatially balanced design is obtained using the function transectSamp from the package MBHdesign with inclusion probabilities $7.48 \times 10^{-5}, 2.52 \times 10^{-5}, 5.9 \times 10^{-6}$ and 4.5×10^{-6} corresponding to spatial units shown in respectively dark purple, light purple, orange and yellow, in Figure 3. The points constituting the zigzags are shown as black discs. The map, supplied by The Metals Company, shows four types of substrates in four shades of green. The lines in grey represent the mining paths in Project Zero. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611). The resolution of the map is of 329 by 480, where each cell measures 400 metres by 400 metres.

We start by generating a master sample using unequal inclusion probabilities in the BAS method to define strata according to the distance to the tracks, as we did in the previous section. The function transectSamp from the package MBHdesign gives an ordered list of 120 points for the map of 329 by 480 spatial cells, where each cell measures 400 metres by 400 metres as shown in Figure 9 (see Appendix A1 for R code). The resolution of the map has been reduced to obtain the master sample in a reasonable time. As mentioned precedingly, the resolution of the transect paths is such that at least one point of the path will be detectable in each spatial unit or cell of the map.

Three scenarios are obtained by selecting a contiguous sequence of transect paths in the master sample in Figure 9. Each scenario is shown for six voyages, scheduled with respect to Project Zero, that consists into mining two blocks of paths, illustrated in grey in Figure 9: South path and North path as defined in Section 3.

4.2 Low number of deployments

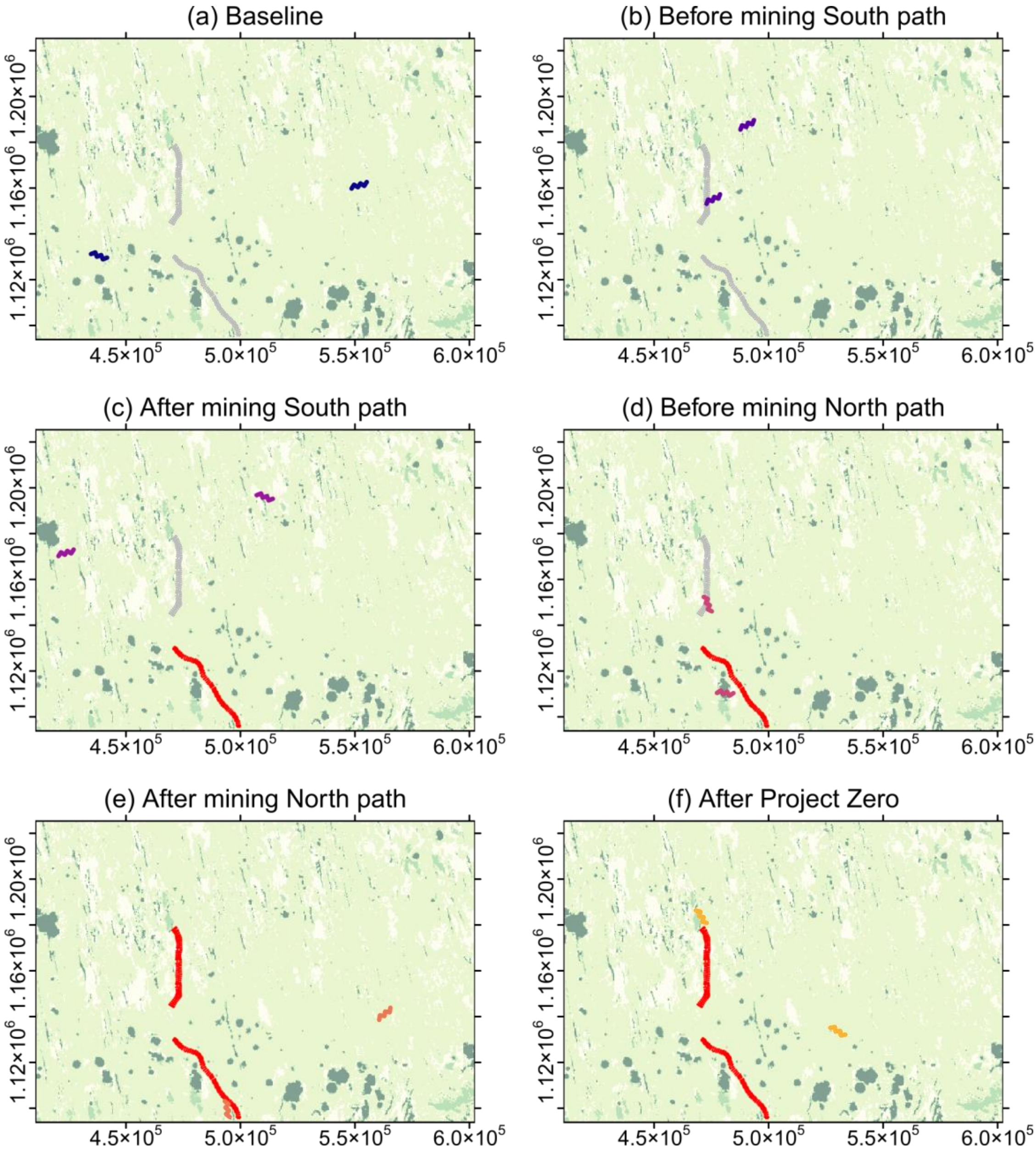


Figure 10 Spatially balanced design with two deployments per voyage
The transects in voyages 1 to 6 are contiguous subsets from the master sample. Each voyage is represented by dots of the same colour. Thick lines represent the mining paths: pre-mining (planned) paths in grey, post mining (actual) in red. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611).

Figure 10 shows a design of 12 deployments divided in six voyages, equivalent to two deployments per voyage, obtained from the master sample in Figure 9. The low number of deployments per voyage is not sufficient for the transects to span over spatial features present in NORI-D (sediments type, bathymetry, etc.), on the top of including voyages as in Figure 10 (a), (c) and (f) where the transects do not cross the mining paths.

4.3 Medium number of deployments

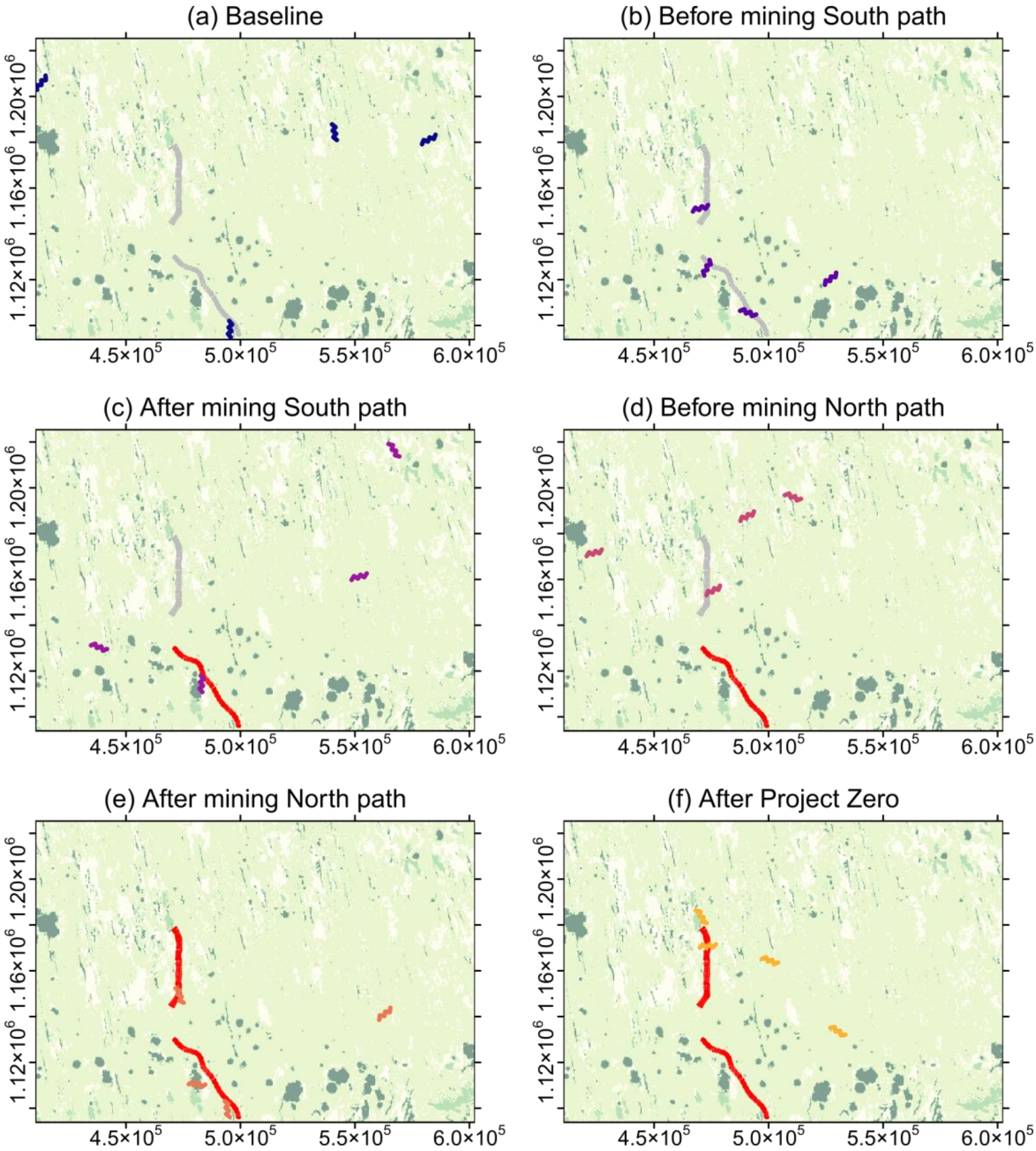


Figure 11 Spatially balanced sampling with four deployments per voyage
The transects in voyages 1 to 6 are contiguous subsets from the master sample. Each voyage is represented by dots of the same colour. Thick lines represent the mining paths: pre-mining (planned) paths in grey, post mining (actual) in red. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611).

Figure 11 shows a spatially balanced design of 24 deployments divided in six voyages, equivalent to four deployments per voyage, obtained from the master sample in Figure 9. The number of deployments per voyage does not always ensure that at least one deployment cross the mining paths. A medium number of deployments per voyage could be a viable option for surveying if deployments crossing the mining paths, like in Figure 11(a)-(b), are scheduled after the mining South and North paths.

4.4 High number of deployments

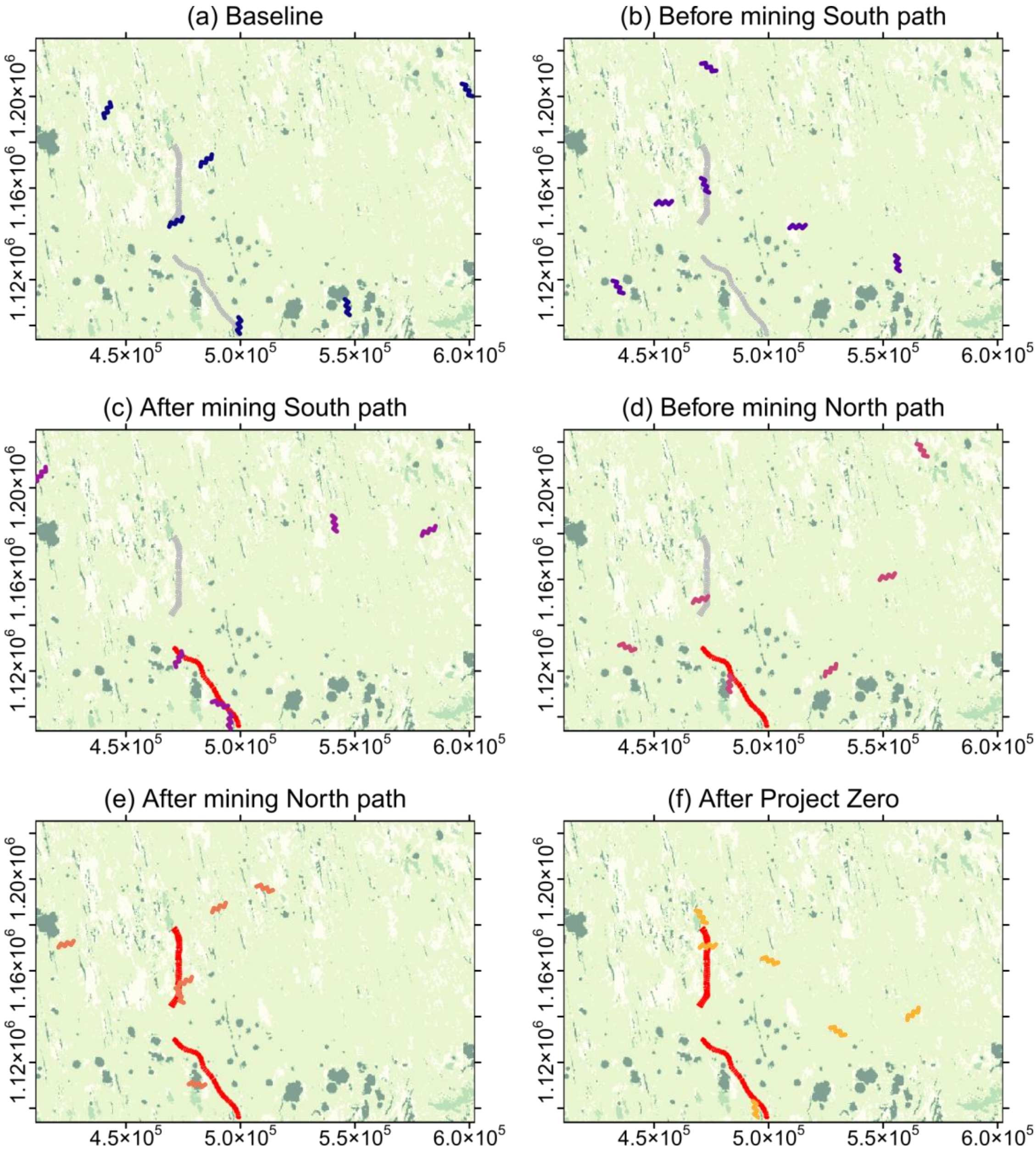


Figure 12 Spatially balanced sampling with six deployments per voyage
The transects in voyages 1 to 6 are contiguous subsets from the master sample. Each voyage is represented by dots of the same colour. Thick lines represent the mining paths: pre-mining (planned) paths in grey, post mining (actual) in red. The coordinates are shown in metres, in the coordinate reference system WGS 84 / UTM Zone 11N (EPSG:32611).

Figure 12 shows a spatially balanced design of 36 deployments divided in six voyages, equivalent to six deployments per voyage, obtained from the master sample in Figure 9. The high number of deployments cover the tracks, their surrounding and the rest of NORI-D. The number of deployments is acceptable if the 36 deployments in Figure 12(a)-(f) are taken altogether as they span over the main spatial features of NORI-D. Transects locations can be revisited often as the survey method by ROV imaging is not destructive.

5 Discussion and conclusion

5.1 Main applications of spatially balanced designs in NORI-D

Spatially balanced sampling offers many features that are practical for surveying the environmental impact of deep-sea mining in NORI-D. First, the design yields a geographical coverage of NORI-D where the samples are spatially balanced over different spatial scales: sample locations are distributed over a regional scale, and they are also conducted on a local scale in each location, to collect tiny specimens such as microorganisms.

Second, the design permits intensification of surveying in a particular zone through unequal inclusion probabilities. In this way, we can divide NORI-D into different layers depending on the distance to the mining tracks and adjust the inclusion probabilities to intensify surveying near and on the mining tracks.

Third, selecting a subset from a master sample and oversampling allow to distribute the samples over different voyages, and furthermore to implement rotating panel designs (see van Dam-Bates P et al. (2018) for how to use BAS for rotating panel design). In the context of NORI-D, benthic sampling in one location may disturb the seafloor in a way that a subsequent benthic sampling close in time at the same location may lead to poor results. Therefore, selecting different subsamples from a master sample would be practical to create longitudinal studies on biodiversity in NORI-D, in the untouched layers as well as in the mining tracks. Revisiting frequently the same locations to create longitudinal data would be suitable for non-destructive sampling, such as ROV imaging.

Finally, the BAS method is not computationally expensive, making possible the calculations with portable computational resources (except for transects designs with large number of rotation directions) and making the generation and the testing of different designs an easy task.

5.2 Standardisation of sampling methodologies

Care must be given to determine the standard sampling protocols and to make sure that the same standards are followed through the survey. Indeed, negative correlations have been found between indicators of food availability and macrofaunal density, a result that the authors of the study (Kaiser et al., 2024) explain by different sampling methodologies, beyond the factor of temporal variation in PCO flux.

Different sampling methods target different environments and yield different results. For example, multicorer (Montagna et al., 2016) and giant box corer (Washburn et al., 2021) are suitable for quantitative assessments, with multicorer preserving the vertical distribution of the sediment and the specimens. Standard box corers and trawls (Chimienti et al., 2018) are used for qualitative assessments. Non-destructive method of surveying by underwater optic imagery allows also for quantitative assessment of abundance, however physical collection of specimens is necessary to deduce biometric correlations to obtain estimates of size and biomass from imaging (Chimienti et al., 2018).

Each sampling method would require a different type of analysis of the data collected; therefore, it is not necessary to ensure that samples obtained with a collection method (for example, multicorer) are spatially balanced with respect to samples obtained with another collection method (for example, ROV imaging). The key takeaway from these differences in sampling that arise from different collection methods is that the selection of a sampling gear should be paired with a standard protocol of collection that should be documented for future use.

5.3 Estimation of population response

Horvitz-Thompson (HT) estimator (Horvitz and Thompson, 1952) can be used to get an unbiased estimate of the characteristic measured from a population when samples are obtained with unequal inclusions probabilities. The population total for a response, for example the population total of a species depending on the efficacy of the collection method, is given by $\hat{X} = \sum_{i=1}^{N} \frac{x_i}{n\pi_i} I(i)$, where $I(i) = 1$ if the spatial unit i is in the sample and $I(i) = 0$ if the spatial unit i is not in the sample, x_i is the response at cell i, π_i is the inclusion probability of cell i such that $\sum \pi_i = 1$, the number of samples in the design is n, and the total number of spatial cell is N. Most importantly, the authors Horvitz and Thompson (1952) give the formula for the variance estimator that the design is supposed to minimise.

Although many variance estimators are available in the literature, the **local neighbourhood variance estimator** (Stevens and Olsen, 2003) is known to perform well with the BAS method (Robertson et al., 2013). Spatially based designs regularly outperform random-based designs in comparative studies where authors use variance estimators to compare designs, with real or simulated data (Robertson et al., 2013; Foster et al., 2017; Dumelle et al., 2020).

Model-based estimators, such as generalised additive models, improve accuracy in estimating the population compared to expansion estimators such as HT estimator (Hogland and Affleck, 2021). A good strategy to improve estimation of the population is combining data collection through a spatially balanced design with population estimation through a model-based estimator that considers inclusion probabilities from the design (Foster et al., 2021).

5.4 Conclusion

In this document, we presented the concepts and applications of spatially balanced designs for surveying the impact of deep-sea mining in NORI-D. We showed how to use equal and unequal inclusions probabilities in the design. We suggested that unequal inclusions probabilities could be based on the distance to the mining tracks, to measure different levels of impact. We also showed how to plan a survey with varying number of voyages and deployments by using a master sample.

Survey designs keep evolving with increasing need to gather data and with rapid advances in statistical analysis. Spatially balanced designs are an appropriate method to survey NORI-D where the region to study is large, benthic locations are hard to access, and many species are rare. In the future, where extensive data could be available and some models of population distribution and recovery could be validated, scientists can consider how extending spatially balanced designs could expand knowledge on benthic biodiversity and deepen insights into the impact of deep-sea mining.

Appendix A R Code

A.1 R code to generate samples and Figures 1, 4 and 9

```
library(MBHdesign) #spatially balanced designs
library(RColorBrewer) #nice colours
library(viridis) #nice colours
library(terra) #spatial data analysis
library(sf) #st_read

# Please set your working directory
working.dir = ""
setwd(working.dir)

# The number of samples to generate
nSamp <- 120

# Read substrates map
map.sh <- st_read("./geo/FP_NORI_D_substrates_v2_211016.shp")
y = rast(ext(map.sh), 1315, 1919)
substrates.rast = rasterize(as(map.sh, "SpatVector"), y, field = "label")
crs(substrates.rast) <- "EPSG:32611"

# Read mining tracks
north.mining <- read.csv("./Northern Block Timestamp_RevB.csv")
south.mining <- read.csv("./Southern Block Timestamp_RevC.csv")

### FUNCTION that calculates unequal inclusion probabilities
### depending on distance to tracks
### Parameters: substrates map, North tracks, South tracks
### Output: Data frame of the distance map (same extent as substrates map)
### with inclusions probabilities for each cell
create.distance.inclusion.prob <- function(substrates.rast,
                                           north.mining,
                                           south.mining)
{
  # Process the mining tracks to add a buffer of 11 metres
  north.points <- vect(
    data.frame(x = north.mining$XPT, y = north.mining$YPT),
    geom = c("x", "y"),
    crs = "EPSG:32611"
  )
  south.points <- vect(
    data.frame(x = south.mining$XPT, y = south.mining$YPT),
    geom = c("x", "y"),
    crs = "EPSG:32611"
  )
  north.lines <- as.lines(north.points)
  south.lines <- as.lines(south.points)
  north.points.buffered <- buffer(north.lines, 11)
  south.points.buffered <- buffer(south.lines, 11)

  # Transform North and South tracks into two polygons
  north.polygons <- as.polygons(north.points.buffered)
  south.polygons <- as.polygons(south.points.buffered)

  # Create a map of distance to the mining tracks
  ground0.rast <- substrates.rast
  values(ground0.rast) <- 1
  ground0.mask <- mask(ground0.rast, north.polygons, inverse = TRUE)
  ground0.mask <- mask(ground0.mask, south.polygons, inverse = TRUE)
  ground0.mask <- is.na(ground0.mask) * 1
  ground0.mask[ground0.mask == 0] <- NA
  distance.rast <- distance(ground0.mask)
```

```
  names(distance.rast) <- 'dist'
  distance.df <- as.data.frame(distance.rast, xy = TRUE)

  # Create seven categories of distance to the mining tracks
  distance.Bins <- c(
    min(distance.df$dist),
    min(distance.df$dist[distance.df$dist > 0]),
    5000,
    quantile(distance.df$dist, probs = seq(20, 100, 20) / 100)
  )
  distance.df$category <- cut(distance.df[, 3], distance.Bins, include.lowest =
                                TRUE)
  ## Define inclusion probabilities with equal expected
  ## sample size for each category.
  incProb.tmp <- 1 / table(distance.df$category) / 7
  distance.df$inclusion.prob <- NA
  for (ii in 1:length(incProb.tmp))
    distance.df[distance.df$category == names(incProb.tmp)[ii], "inclusion.prob"] <-
incProb.tmp[ii]
  return(distance.df)
} ################### END OF FUNCTION create.distance.inclusion.prob


################### POINTS SAMPLING ##################
# Calculate inclusion probabilities
distance.df <- create.distance.inclusion.prob(
  substrates.rast, north.mining, south.mining
  )
# Put in raster to use quasiSamp.raster because it is faster
incProb <- rast(
  distance.df[, c("x", "y", "inclusion.prob")],
  type = "xyz", crs = crs("EPSG:32611")
  )
plot(incProb)

# Reproducible code
set.seed(2125)
# Generate the master sample with unequal inclusions probabilities
unevenSample.dist <- quasiSamp.raster(
  n = nSamp,
  inclusion.probs = incProb,
  nSampsToConsider = 1000 * nSamp,
  nStartsToConsider = 100 * nSamp
)
unevenSample.dist$no <- seq.int(nrow(unevenSample.dist))
# Save the results, just in case
saveRDS(unevenSample.dist, "results_uneven.rds")


# Generate the master sample with equal inclusions probabilities
values(incProb) <- 1 / (1315 * 1919)
evenSample.dist <- quasiSamp.raster(
  n = nSamp,
  inclusion.probs = incProb,
  nSampsToConsider = 1000 * nSamp,
  nStartsToConsider = 100 * nSamp
)
evenSample.dist$no <- seq.int(nrow(evenSample.dist))
# Save the results, just in case
saveRDS(evenSample.dist, "results_even.rds")


################### TRANSECTS SAMPLING ##################
# The number of zigzags transects to generate
nSamp <- 36

#Create the zigzag pattern of 5 segments
x.pts <- seq(1, 26, 1)
y.pts <- c(seq(1, 6), seq(5, 1, -1), seq(2, 6), seq(5, 1, -1), seq(2, 6))
zigzag.pattern <- matrix(
  c(x.pts, y.pts),
```

```
  nrow = length(x.pts),
  ncol = 2,
  byrow = F
)
nb.pts.transect = nrow(zigzag.pattern)
segment.length = 2000 # length of each segment must be 2000 m
# (5x)^2+(5x)^2=2000^2
# x = sqrt(2000^2/(50))
unit.length = sqrt(segment.length * segment.length / 50)
zigzag.pattern <- zigzag.pattern * unit.length

# Parameters describing the transect pattern and the directions of rotation
my.control <- list(
  #the type of transect
  transect.pattern = zigzag.pattern,
  #the number of putative directions that a transect can take
  nRotate = 15,
  mc.cores = 4
)
# Compress map to reduce computation time
substrates.rast <- aggregate(substrates.rast,
                             fact = 4,
                             fun = "mean",
                             na.rm = TRUE)


# Reproducible code
set.seed(2125) # change seed to NULL if you don't want same results

# Calculate inclusion probabilities
distance.df <- create.distance.inclusion.prob(
  substrates.rast, north.mining, south.mining
  )
# potential.sites needs to be ordered so that
# The first dimension changes more rapidly than the second
distance.df <- distance.df[order(distance.df$y), ]


# Generate the master sample of transects with equal inclusions probabilities
transect.samp.dist <- transectSamp(
  n = nSamp,
  potential.sites = distance.df[, c("x", "y")],
  inclusion.probs = distance.df[, "inclusion.prob"],
  control = my.control
)

# It takes a long time to generate the transects, let's not lose them
saveRDS(transect.samp.dist, "results_transects.rds")

#### PLOT FIGURES #########################
levels(substrates.rast) <- data.frame(
  id = 0:4,
  cover = c(
    "Volcanic Outcrop",
    "Volcanic Outcrop",
    "NORI D Patch Drift Sediment",
    "NORI D Type 1 Nodules on Sediment",
    "NORI D Type 2-3 Nodules on Sediment"
  )
)
my.legend = c("1-20", "21-40", "41-60", "61-80", "81-100", "101-120")
my.title = "20 sites per voyage"
pax <- list(
  xat = c(4.5e5, 5e5, 5.5e5, 6e5),
  yat = c(1.10e6, 1.12e6, 1.14e6, 1.16e6, 1.18e6, 1.2e6, 1.22e6),
  ylabs = c(
    "",
    expression("1.12\u00D710"^"6"),
    "",
    expression("1.16\u00D710"^"6"),
    "",
```

```
    expression("1.20\u00D710"^"6"),
    ""
  ),
  xlabs = c(
    expression("4.5\u00D710"^"5"),
    expression("5.0\u00D710"^"5"),
    expression("5.5\u00D710"^"5"),
    expression("6.0\u00D710"^"5")
  ),
  tick = 1:4,
  cex.axis = 1.5
)
legend.substrates = rev(
  c(
    "Type 2-3 Nodules on Sediment",
    "Type 1 Nodules on Sediment",
    "Patch Drift Sediment",
    "Volcanic Outcrop"
  )
)

#### Plot Figure 1: Point samples with equal incl. prob. ####
pdf(paste0("Figure1.pdf"), width = 12, height = 6)
par(
  fig = c(0, 1, 0, 1),
  oma = c(0, 0, 1, 4),
  mar = c(5, 4, 4, 2) + 0.1
)
plot(
  substrates.rast,
  col = rev(brewer.pal(n = 9, name = "YlGn"))[c(1, 5, 7, 9)],
  alpha = 0.5,
  pax = pax,
  legend = "top",
  plg = list(
    cex = 1.25,
    title = "NORI-D Substrates",
    legend = legend.substrates
  )
)
# Where are the mining paths
points(north.mining$XPT[seq(1, 16351, 10)],
       north.mining$YPT[seq(1, 16351, 10)],
       cex = 0.05,
       col = "grey")
points(south.mining$XPT[seq(1, 16351, 10)],
       south.mining$YPT[seq(1, 16351, 10)],
       cex = 0.05,
       col = "grey")
# Where are the samples
# 6 voyages of six colours
for (i in 1:6)
  points(
    evenSample.dist[(1:20) + (i - 1) * 20, c("x", "y")],
    cex = 1.25,
    pch = 20,
    col = viridis_pal(option = "plasma")(7)[i]
  )
legend(
  "topleft",
  inset = c(0.25, -0.1),
  xpd = TRUE,
  cex = 1,
  bty = "n",
  horiz = TRUE,
  bg = "white",
  legend = my.legend,
  pch = 19,
  col = viridis_pal(option = "plasma")(7)[1:6]
)
title(
```

```
  my.title,
  cex.main = 1.2,
  font.main = 1,
  adj = 0.5,
  outer = "TRUE"
)
dev.off()

#### Plot Figure 4: Point samples with unequal incl. prob. ####
pdf(paste0("Figure4.pdf"), width = 12, height = 6)
par(
  fig = c(0, 1, 0, 1),
  oma = c(0, 0, 1, 4),
  mar = c(5, 4, 4, 2) + 0.1
)
plot(
  substrates.rast,
  col = rev(brewer.pal(n = 9, name = "YlGn"))[c(1, 5, 7, 9)],
  alpha = 0.5,
  pax = pax,
  legend = "top",
  plg = list(
    cex = 1.25,
    title = "NORI-D Substrates",
    legend = legend.substrates
  )
)
# Where are the mining paths
points(north.mining$XPT[seq(1, 16351, 10)],
       north.mining$YPT[seq(1, 16351, 10)],
       cex = 0.05,
       col = "grey")
points(south.mining$XPT[seq(1, 16351, 10)],
       south.mining$YPT[seq(1, 16351, 10)],
       cex = 0.05,
       col = "grey")
# Where are the samples
# 6 voyages of six colours
for (i in 1:6)
  points(
    unevenSample.dist[(1:20) + (i - 1) * 20, c("x", "y")],
    cex = 1.25,
    pch = 20,
    col = viridis_pal(option = "plasma")(7)[i]
  )
legend(
  "topleft",
  inset = c(0.25, -0.1),
  xpd = TRUE,
  cex = 1,
  bty = "n",
  horiz = TRUE,
  bg = "white",
  legend = my.legend,
  pch = 19,
  col = viridis_pal(option = "plasma")(7)[1:6]
)
title(
  my.title,
  cex.main = 1.2,
  font.main = 1,
  adj = 0.5,
  outer = "TRUE"
)
dev.off()

#### Write coordinates (Fig. 4) as long/lat in csv file ####
from_crs <- "EPSG:32611"
to_crs <- "EPSG:4326"  # WGS84 (lat/lon)
new_coords <- project(as.matrix(unevenSample.dist[, c("x", "y")], ncol =
                                2),
```

```
                               from = from_crs,
                               to = to_crs)
write.csv(new_coords, "./coodinates_benthic_sampling.csv")

#### Plot Figure 9: transects with unequal incl. prob. ####
pdf("Figure9.pdf", width = 10, height = 4)
par(
  fig = c(0, 1, 0, 1),
  oma = c(0, 0, 1, 0),
  mar = c(5, 4, 4, 2) + 0.1
)
plot(
  substrates.rast,
  col = rev(brewer.pal(n = 9, name = "YlGn"))[c(1, 5, 7, 9)],
  alpha = 0.5,
  pax = pax,
  legend = "top",
  plg = list(
    cex = 1,
    title = "NORI-D Substrates",
    legend = legend.substrates
  )
)
# Where are the mining paths
points(north.mining$XPT[seq(1, 16351, 10)],
       north.mining$YPT[seq(1, 16351, 10)],
       cex = 0.05,
       col = "grey")
points(south.mining$XPT[seq(1, 16351, 10)],
       south.mining$YPT[seq(1, 16351, 10)],
       cex = 0.05,
       col = "grey")
# Where are the samples
points(
  transect.samp.dist$points[, c("x", "y")],
  cex = 0.1,
  pch = 1,
  col = "black"
)
dev.off()
```

A.2 Geographical coordinates of benthic sampling locations

Table 1 Geographical coordinates of sampling locations in Figure 4

ORDER IN MASTER SAMPLE	LONGITUDE	LATITUDE	ORDER IN MASTER SAMPLE	LONGITUDE	LATITUDE
1	-117.103	10.5683	**61**	-116.415	10.54065
2	-116.337	10.08378	**62**	-117.619	9.938256
3	-117.432	10.4803	**63**	-116.524	10.8637
4	-117.79	11.00774	**64**	-116.964	10.20293
5	-116.912	9.995798	**65**	-117.36	10.37909
6	-116.692	10.52476	**66**	-116.265	10.24642
7	-117.24	9.922463	**67**	-117.47	10.81939
8	-117.514	10.3626	**68**	-116.867	10.68768
9	-117.076	10.09891	**69**	-117.086	10.02926
10	-116.747	10.23087	**70**	-117.196	10.55741
11	-117.186	10.62705	**71**	-117.771	10.46881

12	-116.966	10.80347	**72**	-117.332	10.20547
13	-116.487	10.67101	**73**	-116.676	9.984799
14	-116.595	10.99403	**74**	-117.113	10.38199
15	-117.035	9.981337	**75**	-117.223	10.91011
16	-117.309	10.50938	**76**	-117.004	9.911695
17	-116.214	10.02473	**77**	-116.128	10.30667
18	-117.199	10.15765	**78**	-116.73	10.74818
19	-117.419	10.9497	**79**	-117.168	10.08802
20	-117.337	10.69205	**80**	-116.95	10.61625
21	-117.117	10.16402	**81**	-116.491	10.13199
22	-116.35	10.73586	**82**	-117.698	11.0559
23	-117.72	10.60367	**83**	-117.312	9.970339
24	-116.624	10.11951	**84**	-116.215	10.89472
25	-116.076	10.44127	**85**	-117.422	10.41067
26	-117.588	9.958209	**86**	-116.901	10.27889
27	-116.823	10.88391	**87**	-117.34	10.67486
28	-117.261	10.22272	**88**	-117.121	10.69492
29	-117.042	10.39919	**89**	-117.23	10.16577
30	-117.645	10.72141	**90**	-116.353	10.56223
31	-117.207	10.45791	**91**	-117.723	10.07732
32	-116.548	10.5897	**92**	-116.627	11.00312
33	-117.343	10.10507	**93**	-117.066	9.990377
34	-116.246	11.02954	**94**	-117.176	10.51853
35	-116.85	9.904426	**95**	-117.574	10.84092
36	-117.289	10.30139	**96**	-117.134	10.57734
37	-117.07	10.82969	**97**	-116.805	10.35663
38	-116.302	10.34417	**98**	-117.244	10.75363
39	-117.399	10.74171	**99**	-117.025	10.22464
40	-117.794	10.21191	**100**	-116.421	10.7975
41	-117.138	10.12512	**101**	-116.532	10.06331
42	-117.248	10.65324	**102**	-117.327	10.9887
43	-116.152	10.0508	**103**	-117.107	10.10795
44	-117.522	10.97577	**104**	-116.23	10.50317
45	-116.754	10.49225	**105**	-117.217	10.63608
46	-116.974	10.00757	**106**	-117.435	10.01904
47	-116.452	10.93231	**107**	-117.271	10.28331
48	-117.658	10.32979	**108**	-117.052	10.30332
49	-116.563	10.19811	**109**	-116.285	10.17049
50	-117.275	10.77079	**110**	-117.381	10.56719

51	-117.384	10.15386	**111**	-117.818	10.7427
52	-116.936	11.07931	**112**	-116.94	10.08263
53	-117.154	10.43805	**113**	-116.72	10.61161
54	-116.389	9.952742	**114**	-117.268	10.12685
55	-116.662	10.39359	**115**	-116.172	11.05105
56	-117.209	10.26165	**116**	-117.543	10.44938
57	-116.113	10.71629	**117**	-116.776	9.965888
58	-117.566	10.58409	**118**	-117.214	10.36204
59	-117.237	10.49678	**119**	-116.995	10.89029
60	-117.017	11.02505	**120**	-117.679	10.27367

References

Brown JA, Robertson BL, McDonald T. (2015). Spatially Balanced Sampling: Application to Environmental Surveys. Procedia Environmental Sciences, 27, 6-9. doi:10.1016/j.proenv.2015.07.108

Chimienti G, Angeletti L, Rizzo L, et al. (2018). ROV trawling approaches in the study of benthic communities: the case of *Pennatula rubra* (Cnidaria: Pennatulacea). Journal of the Marine Biological Association of the United Kingdom, 98, 1859-1869. doi:10.1017/s0025315418000851

Dumelle M, Higham M, Ver Hoef JM, et al. (2022). A comparison of design-based and model-based approaches for finite population spatial sampling and inference. Methods in Ecology and Evolution, 13, 20182029. doi:10.1111/2041-210X.13919

Foster SD. (2020). MBHdesign: An R-package for efficient spatial survey designs. Methods in Ecology and Evolution, 12, 415-420. doi:10.1111/2041-210x.13535

Foster SD, Hosack GR, Hill NA, et al. (2014). Choosing between strategies for designing surveys: autonomous underwater vehicles. Methods in Ecology and Evolution, 5, 287-297. doi:10.1111/2041-210X.12156

Foster SD, Hosack GR, Lawrence E, et al. (2017). Spatially balanced designs that incorporate legacy sites. Methods in Ecology and Evolution, 8, 1433-1442. doi:10.1111/2041-210x.12782

Foster SD, Hosack GR, Monk J, et al. (2019). Spatially balanced designs for transect-based surveys. Methods in Ecology and Evolution, 11, 95-105. doi:10.1111/2041-210x.13321

Foster SD, Monk J, Lawrence E, et al. (2024). Statistical considerations for monitoring and sampling. In Field Manuals for Marine Sampling to Monitor Australian Waters, Version 3. Przeslawski R, Foster S (Eds). National Environmental Science Program (NESP).

Foster SD, Vanhatalo J, Trenkel VM, et al. (2021). Effects of ignoring survey design information for data reuse. Ecological Applications, 31, e02360. doi:10.1002/eap.2360

Gaikwad S, Sautya S, Damare S, et al. (2024). Macrobenthic communities in the polymetallic nodule field, Indian Ocean, based on multicore and box core analysis. Frontiers in Marine Science, 11. doi:10.3389/fmars.2024.1395892

Hayes KR, Hosack GR, Lawrence E, et al. (2019). Designing Monitoring Programs for Marine Protected Areas Within an Evidence Based Decision Making Paradigm. Frontiers in Marine Science, 6. doi:10.3389/fmars.2019.00746

Hogland J, Affleck DLR. (2021). Improving Estimates of Natural Resources Using Model-Based Estimators: Impacts of Sample Design. Estimation Technique, and Strengths of Association. Remote Sensing, 13, 3893. doi:10.3390/rs13193893

Horvitz DG, Thompson DJ. (1952). A Generalization of Sampling Without Replacement From a Finite Universe. Journal of the American Statistical Association, 47, 663-685.

Jamieson AJ, Boorman B, Jones, DOB. (2013). Deep-Sea Benthic Sampling. In Methods for the Study of Marine Benthos, A. Eleftheriou (Ed.). doi:10.1002/9781118542392.ch7

Kaiser S, Bonifácio P, Kihara TC, et al. (2024). Effects of environmental and climatic drivers on abyssal macrobenthic infaunal communities from the NE Pacific nodule province. Marine Biodiversity, 54. doi:10.1007/s12526-024-01427-7

Kermorvant C, D'Amico F, Bru N, et al. (2019). Spatially balanced sampling designs for environmental surveys. Environmental Monitoring and Assessment, 191. doi:10.1007/s10661-019-7666-y

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. doi:10.1038/s41598-023-35518-5

McQuaid KA, Attrill MJ, Clark MR, et al. (2020). Using Habitat Classification to Assess Representativity of a Protected Area Network in a Large, Data-Poor Area Targeted for Deep-Sea Mining. Frontiers in Marine Science, 7. doi:10.3389/fmars.2020.558860

Montagna PA, Baguley JG, Hsiang CY, et al. (2016). Comparison of sampling methods for deep-sea infauna. Limnology and Oceanography: Methods, 15, 166-183. doi:10.1002/lom3.10150

Przeslawski R and Foster S. (2024). Field Manuals for Marine Sampling in Australian Waters, Version 3. National Environmental Science Program (NESP). Marine and Coastal Hub. doi:10.25607/OBP-918.2

Robertson BL, Brown JA, McDonald T, et al. (2013). BAS: Balanced Acceptance Sampling of Natural Resources. Biometrics, 69, 776-784. doi:10.1111/biom.12059

Robertson BL, McDonald T, Price CJ et al. (2017). A modification of balanced acceptance sampling. Statistics & Probability Letters, 129, 107-112. doi:10.1016/j.spl.2017.05.004

Simon-Lledó E, Fleming B, Van Audenhaege L et al. (2023) Quantitative benthic megafauna baseline ecological characterisation at NORI-D. Technical report to TMC. National Oceanography Centre, Southampton.

Smith C, Deleo F, Bernardino A, et al. (2008). Abyssal food limitation, ecosystem structure and climate change. Trends in Ecology & Evolution, 23, 518-528. doi:10.1016/j.tree.2008.05.002

Stefanoudis PV, Bett BJ, Gooday AJ. (2016). Abyssal hills: Influence of topography on benthic foraminiferal assemblages. Progress in Oceanography, 148, 44-55. doi:10.1016/j.pocean.2016.09.005

Stevens DL and Olsen AR. (2003). Variance estimation for spatially balanced samples of environmental resources. Environmetrics, 14, 593-610. doi:10.1002/env.606

Stevens DL and Olsen AR. (2004). Spatially Balanced Sampling of Natural Resources. Journal of the American Statistical Association, 99, 262-278. doi:10.1198/016214504000000250

The Metals Company (2024). Substrates Map of NORI-D in the CCZ. [Shapefile]. Provided by The Metals Company.

Urban E, Seger KD, Schmitter-Soto JJ et al. eds. (2022). Before-After Control-Impact (BACI) Studies in the Ocean. Frontiers Media SA.

van Dam-Bates P, Gansell O, Robertson B (2018). Using balanced acceptance sampling as a master sample for environmental surveys. Methods in Ecology and Evolution, 9, 1718-1726. doi:10.1111/2041-210x.13003

Washburn TW, Menot L, Bonifácio P, et al. (2021). Patterns of Macrofaunal Biodiversity Across the Clarion-Clipperton Zone: An Area Targeted for Seabed Mining. Frontiers in Marine Science, 8. doi:10.3389/fmars.2021.626571

As Australia's national science agency, CSIRO is solving the greatest challenges through innovative science and technology.

Contact us

1300 363 400
+61 3 9545 2176
csiro.au/contact
csiro.au

For further information

CSIRO Environment
Dr Piers Dunstan
+61 3 6232 5382
piers.dunstan@csiro.au
research.csiro.au/dsm