JNCC Report
No. 641

Optimisation of Benthic Image Analysis Approaches

Moore, J., van Rein, H., Benson, A., Sotheran, I., Mercer, T. & Ferguson, M.

December 2019

ISSN 0963-8091

For further information please contact:

Marine Monitoring and Evidence (marinemonitoring@jncc.gov.uk)
Joint Nature Conservation Committee, Monkstone House
City Road, Peterborough, PE1 1JY
www.jncc.defra.gov.uk

Recommended citation:

Moore, J., van Rein, H., Benson, A., Sotheran, I., Mercer, T. & Ferguson, M. (2019). Optimisation of Benthic Image Analysis Approaches, *JNCC Report, No. 641*, JNCC, Peterborough, ISSN 0963-8091.

EQA:

This report is compliant with the JNCC Evidence Quality Assurance Policy http://jncc.Defra.gov.uk/default.aspx?page=6675.

Acknowledgements

The authors would like to thank the following imagery analysts for their annotation and comments on the extraction methods: Alison Benson, Envision, Hexham; Ian Sotheran, Envision, Hexham; Eloise Boblin, Envision, Hexham; Ellen Barrowclift, Envision, Hexham; Lucy Cross, Envision, Hexham; Tom Mercer, ASML, Frosterley, Co. Durham. Data analysis was led by Jon Moore, ASML, Cosheston, Pembrokeshire, and supported by Henk van Rein, JNCC, Peterborough. Finally, we thank Stefano Marra, JNCC, Peterborough, for his contribution in reviewing the report.

This work was funded by DEFRA Project FI009 Marine Methodologies.

Summary

This study explores the relative strengths and weaknesses of different data extraction methods for enumerating marine benthic taxa from still imagery collected by drop frame cameras. The imagery used in the study was collected from the Solan Bank Reef Special Area of Conservation in 2013. This study makes use of 100 high-resolution still images dominated by faunal crusts and turfs, encrusting sponges, hydroids, cup corals, serpulid worms and brittlestars to represent a high-density, high-diversity circalittoral bedrock reef community. To enumerate the bedrock reef community, the following six common data extraction methods were used on each still image: percentage cover, abundance count, the SACFOR scale, point intercept and two frequency of occurrence grids, one of 25 cells (5x5 grid) and the other of 100 cells per image (10x10 grid). The methods were applied to the 100 images by an experienced senior analyst. Several data metrics were calculated from the community data to aid comparison of the extraction methods. These include statistical precision, statistical power, efficiency of extraction, taxonomic richness and accumulation of data, and community impression.

To explore the consistency of data between different analysts, an additional five analysts (two more senior analysts and three junior analysts) also applied the six data extraction methods to a subset of 20 images. In addition to the consistency comparisons this also allowed comparisons of efficiency and taxonomic richness between the analysts. Note that the images and methods were analysed in a random order to minimise bias and learning effects. No guidance for taxonomic identification was provided to the analysts while analysing the imagery, other than to follow standard commercial practice for imagery analysis contracts in the UK. As a result, numerous taxonomic inconsistencies were observed in the community data. Although these issues were to some extent resolved by rigorous data filtering and manipulation practices (truncation), they reduced the robustness of the study. It is recommended that future studies of this nature make use of standard taxonomic lists to reduce taxonomic inconsistencies between analysts.

The numerous analyses in this study produced many interesting results. Overall, no one extraction method out-performed the other methods consistently but rather, different methods were better at different things. The traditional methods, percentage cover and abundance count, generated the most accurate community impressions. However, the data show that consistency, precision and power are relatively poor among these methods. When considering use of these methods, note they are usually used in tandem as neither is able to extract data from across the whole community (percentage cover for ground cover taxa and abundance count for solitary and motile taxa).

Comparisons of the SACFOR, frequency grid and point intercept data showed that methods with smaller data ranges (SACFOR: 1-6; 5x5 frequency grid: 0-25) tended to have high estimates of precision and power, as well as being the most efficient. Owing to the point intercept method's lack of spatial coverage, its data had relatively poor representation of taxa (poorest among solitary taxa). However, the point intercept data were the most consistent between the different analysts while the SACFOR data was the least consistent. SACFOR and the fine frequency grid (10x10) also generated the least realistic impressions of the community.

Comparison of the analysts' data revealed that the senior analysts worked through imagery twice as quick as the junior analysts and also generated larger taxa lists. Consistency between all analysts was strikingly poor in this study, however, the junior analysts were on average more consistent (between themselves) than the senior analysts.

Numerous recommendations arise from this study:

- although no single data extraction method performed the best at everything, the rank-based optimisation assessment in this study showed the Frequency of Occurrence (5x5 grid) recorded data that showed higher levels of precision, power and consistency than the majority of other methods in this study. This method recorded numerical data in a single unit that could represent the whole community sampled and the whole sample area (unlike percentage cover, abundance counts and point intercept). Further exploration of this method is recommended in future studies;

- image analysis approaches that reduce inconsistencies between analysts should be employed in future studies, including the use of image annotation software, fixed taxa lists, and methods that performed constantly in this study (point intercept and frequency grids). This will be particularly important for methods that produce data that is non-additive (frequency grids);

- future enumeration method comparison studies of this type should carry out an appropriate sampling unit analysis to reflect future monitoring needs;

- multi-metric method comparison studies should be carried out for different monitoring purposes, such as for target taxa (i.e. indicators) or different habitat types, in order to estimate the optimum approaches for those purposes.

Contents

1 Introduction

Benthic imagery, in the form of still images and video, is often used as an efficient, effective and non-destructive way to describe and illustrate seabed communities dominated by epibiota (particularly rocky reefs, biogenic reefs and seagrass beds). National and international legislation requires assessment and reporting on the condition of these communities and associated habitats (Hinchen 2014), so monitoring programmes are being developed, often using imagery methods. JNCC leads on a number of these programmes, particularly in offshore areas; working within the UK Marine Biodiversity Monitoring Programme (UK MBMP). UK MBMP aims to improve the quality of evidence on the state of marine biodiversity and to focus research into whether and how the biodiversity elements are changing in response to both natural and human induced pressures. However, it has proven difficult to standardise the collection and interpretation of benthic imagery for marine monitoring. A range of issues are involved, and many are widely recognised in national and international marine monitoring forums.

This report considers one specific issue: How to choose an appropriate method of enumerating taxa from still images of the benthos? Turner *et al.* (2016) discuss enumeration approaches but provide limited information on which is more suitable than another. Method comparison studies agree that monitoring methods must provide data that are accurate, precise and statistically robust to better detect changes over time, while also being efficient to collect and cost-effective to use in a monitoring programme (Drummond & Connell 2005; Beaumont *et al.* 2007; van Rein *et al.* 2011). Others add that inter-observer variability needs to be reduced so that patterns in the monitoring data better reflect community changes and not differences in the opinion of observers (Moore *et al.* 2015; Durden *et al.* 2016).

This study aims to further understand how the most commonly used data extraction methods, for still imagery, perform when compared against each other. More specifically:

To recommend the optimum non-automated, data extraction approach for benthic imagery acquired from marine benthic habitats by drop-down, towed camera systems.

The objectives were to:

1 *Extract biological community data from a set of benthic images using different data extraction methods and different observers.*

2 *Analyse biological community data and calculate metrics to compare data extraction methods and assess inter-observer effects.*

3 *Summarise the results and collate recommendations.*

2 Methods

2.1 Image files

The 100 sample images selected for this study were collected from within the Solan Bank Reef SAC (Special Area of Conservation), an area of bedrock and stony reef approximately 50km from the north coast of mainland Scotland. All sample images were classified as 'good' or 'excellent' quality using guidelines set by the North-East Atlantic Marine Biological Analytical Quality Control Scheme (NMBAQC; Turner *et al.* 2016). The SAC was designated for its Annex I reef habitat. Much of the reef is subjected to high levels of scour and therefore sparse in epifauna, but the selected images were from an area of circalittoral reef (approximately 50 to 60m depth) that was less scoured and characterized by higher biodiversity, including fragile sponges and anthozoan communities, as well as hydroids, bryozoans (erect and encrusting species), encrusting coralline algae, caryophyllid cup corals, ophiuroids and *Alcyonium digitatum*.

The images were captured during the JNCC/MSS Scotia 1714S survey, on vessel *MRV Scotia*, in October/November 2014. A 10 mega pixel camera with dedicated flash was deployed on a drop frame. Scaling was provided by two pairs of lasers, projecting a continuous centre square of 64mm onto the seabed, though the dots were usually bleached-out by the camera flash. Hanging below the drop-frame, was a weight of 64mm diameter, suspended 1.25m below the camera lens by a rope, which was generally visible within stills and provided a more reliable means of scaling objects. More details of the survey and its methodologies are given in Goudge *et al.* 2016.

Example images are shown in Figure 1.

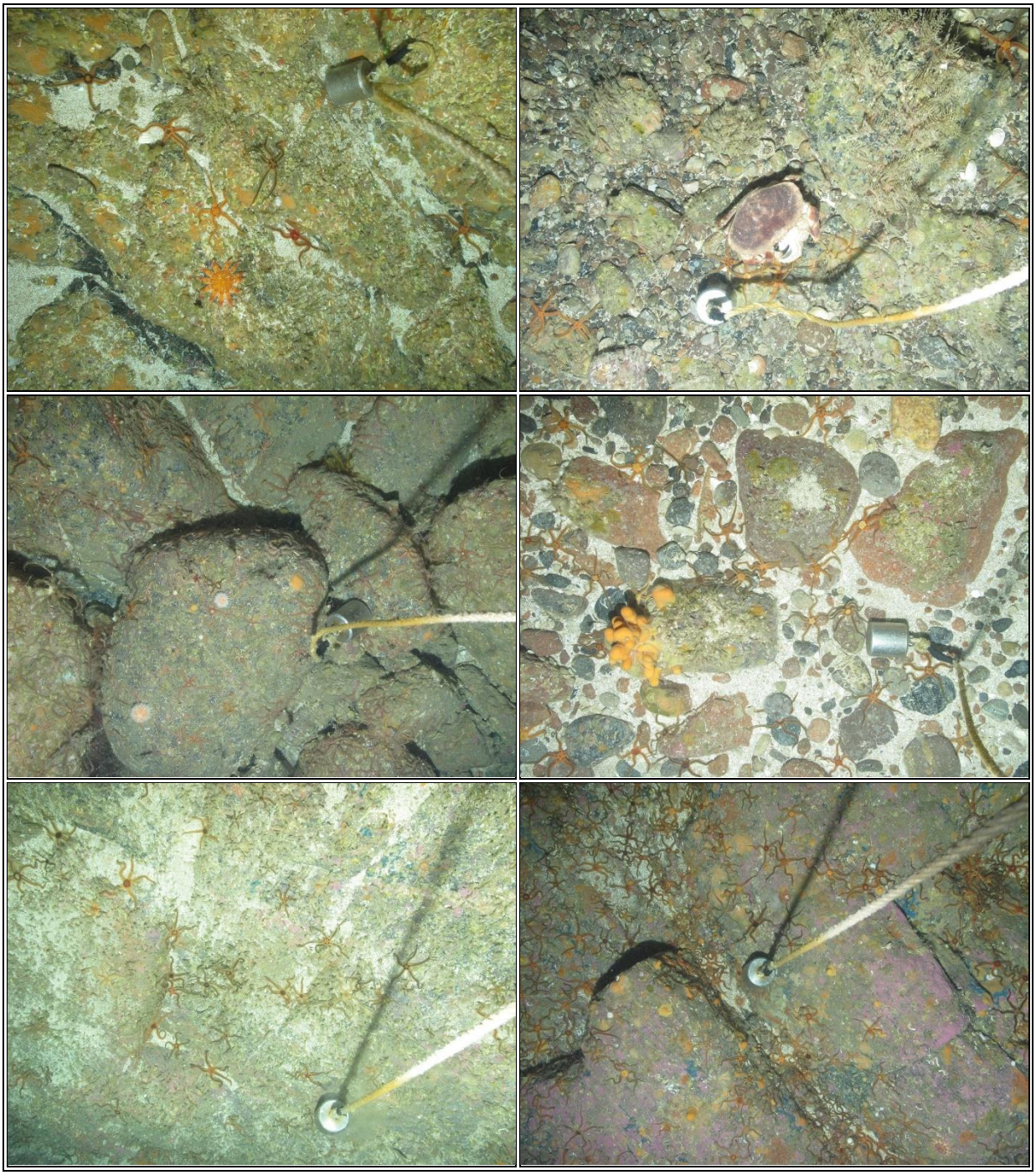


Figure 1. Example images selected to show a range of habitats and taxa.

2.2 Study design outline

Six methods for data extraction from images were chosen for this study:

- Percentage Cover – estimating the % cover of colonial/encrusting taxa
- Abundance Count – counting individuals of erect/solitary animals
- SACFOR – applying semi-quantitative abundances from the MNCR[1] abundance scales

[1] Marine Nature Conservation Review.

- Point Intercept – counting the frequency of each taxon directly under 100 gridded points
- 10 x 10 Frequency - counting the frequency of each taxon within 100 gridded cells
- 5 x 5 Frequency - counting the frequency of each taxon within 25 gridded cells

Primary dataset: An experienced analyst (Analyst A) applied each extraction method to the set of 100 images (i.e. every extraction method on every image, 6 x 100 = 600 samples), enumerating every identified taxon in each sample. The resulting matrix of 600 samples x multiple taxa forms the primary dataset for the comparisons between methods.

Secondary dataset: Further analyses were also carried out to study the consistency of recording by multiple analysts and the effect of analyst experience. These were carried out on a subset of 20 images[2] from the 100 images used above. Five other analysts (two experienced 'Senior' and three less experienced 'Junior') (Analysts B to F) then applied each extraction method to those 20 images (i.e. all five analysts applying every extraction method on every image, 5 x 6 x 20 = 600 samples). The same 20 images were analysed by all Analysts, to allow direct comparison of consistency. Added to the data for those same 20 images that were previously extracted by Analyst A, the secondary dataset provided a matrix of 720 samples (6 x 6 x 20) x multiple taxa for analysis.

The study design is summarised in Figure 2:

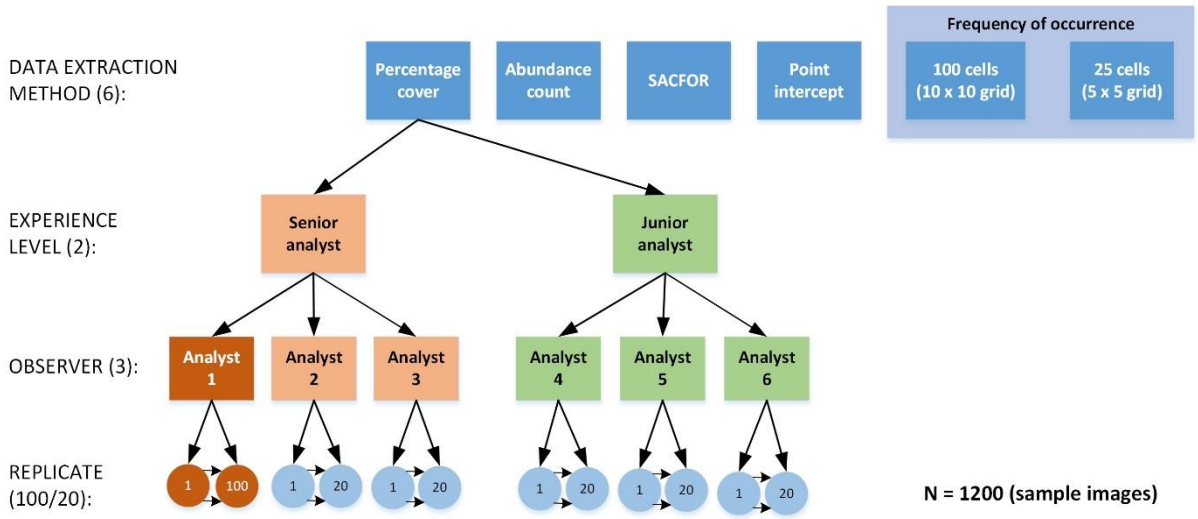


Figure 1. Experimental design for analysis of data from different extraction methods. Experimental factors are shown to the left-hand border and levels across the figure. Note all levels were repeated for each data extraction method (not shown here).

The six data extraction methods are described in Section 2.3.

The required experience of the analysts was defined as:

Senior analysts: having excellent taxonomic identification skills, especially in identifying epifauna from benthic imagery, and over 10 years' experience extracting data from benthic imagery;
Junior analysts: having a lower experience level.

[2] Only 20 images were analysed for this part of the study for resourcing and time constraint reasons.

2.3 Data extraction methods

2.3.1 Percentage Cover

Percentage cover was only estimated for ground-covering taxa (i.e. colonial/encrusting species/serpulid worms/barnacles/hydroids). Pre-gridded images were not required for this extraction method but were recommended. The image was overlain with the '10x10 grid' to enable a more accurate estimate of percentage cover to be made (see example in Figure d). Where the taxa were erect or solitary epifauna, the percentage cover not assessed, but the analyst still recorded its presence and entered N/A in the "Percentage Cover" sheet of the data entry spreadsheet.

2.3.2 Abundance Count

The total number of individuals of each taxonomic classification was counted for erect/solitary epifauna within the entire sample image field of view. Where the taxa were colonial or encrusting, the abundance was not counted, but was recorded as N/A. Pre-gridded images were not required for this extraction method, but the 5x5 gridded images were typically used to aid the counting process.

2.3.3 SACFOR

This semi-quantitative measure was used to record the relative occurrence of all organisms within the sample image, following the NMBAQC guidelines in Turner *et al.* (2016). Each taxon observed within the image was listed, then the abundance count for erect/solitary epifauna or percentage cover (for colonial/encrusting species/serpulid worms/barnacles/ hydroids) was assigned a SACFOR abundance according to the table of abundance scales shown in Table 1. Pre-gridded images were not required for this extraction method, but the 10x10 gridded images were typically used to aid estimation of ground cover taxa.

Table 1. Marine Nature Conservation Review SACFOR abundance scale.

	Growth form		Size of individuals/colonies				
% cover	**Crust/meadow**	**Massive/Turf**	**<1cm**	**1-3 cm**	**3-15 cm**	**>15 cm**	**Density**
>80%	S		S				>1/0.001 m^2 >10,000 / m^2 (1x1 cm)
40-79%	A	S	A	S			1-9/0.001 m^2 1000-9999 / m^2
20-39%	C	A	C	A	S		1-9 / 0.01 m^2 100-999 / m^2 (10 x 10 cm)
10-19%	F	C	F	C	A	S	1-9 / 0.1 m^2 10-99 / m^2
5-9%	O	F	O	F	C	A	1-9 / m^2
1-5% or density	R	O	R	O	F	C	1-9 / 10m^2 (3.16 x 3.16 m)
<1% or density		R		R	O	F	1-9 / 100 m^2 (10 x 10 m)
					R	O	1-9 / 1000 m^2 (31.6 x 31.6 m)
						R	<1/1000 m^2

The following protocols were applied when using the scale:

- whenever an attached species covers the substratum and percentage cover can be estimated, that scale should be used in preference to the density scale;
- use the massive/turf percentage cover scale for all species, excepting those given under crust/meadow;
- where two or more layers exist, for instance foliose algae overgrowing crustose algae, total percentage cover can be over 100% and abundance grade will reflect this;
- the species examples given in the guidance for the various growth forms and sizes (see "SACFOR scale and species.xlsx") take precedence over their actual size in deciding which scale to use;
- the scales have been amended from the original for this project, to improve consistency, by placing all Hydrozoa in the "massive/turf" category, and all Serpulid worms (e.g. *Spirobranchus*) and barnacles in the "crusts/meadows" category.

Note: the majority of images had a field of view of 1m^2 and therefore the appropriate rows of the SACFOR scales was (1-9 / m^2 or 10-99 / m^2 *etc*.). The area of view could be judged by the cross hairs visible on some of the images (interval 64mm x 64mm) and the metal weight (approx. 6cm in diameter). When the weight was in contact with the substrate and the rope was taut, the image field of view was approximately 1m^2. When the metal weight was resting on the bottom and the rope was loose, the field of view was less. An indicative field of view size for each of the images was provided in the spreadsheet of randomised order.

2.3.4 Point Intercept

Pre-gridded images were required for this extraction method. A regular grid of 100 points overlaid each image (see example in Figure b). Only the taxa that lay directly under the intersection of the cross hairs of the point were recorded, not the lines of the cross hair or the label number. All 100 points were recorded. If there was no recognisable taxon under a point it was assigned to 'No identifiable taxa'.

2.3.5 Frequency of occurrence: 25 cell grid and 100 cell grid

Pre-gridded images were required for this extraction method (see examples in Figure c and d). Each image was divided up into equal areas (both 100 cells with a 10x10 grid, and 25 cells with a 5x5 grid) and the presence of any unique taxa recorded within each of the grid cells (e.g. the number of cells that taxa occurs in were counted). The occurrence of each taxon in a cell represented 1% in the 100-cell grid and 4% in a 25-cell grid. A taxon lying under more than one cell (even if only a fraction) was counted for the total number of those cells.

2.4 Taxonomic assignment

All species were identified to the lowest taxonomic level possible. Where an analyst was uncertain at identifying individual species at a certain taxonomic level then a broader taxonomic level (e.g. from species to genera) was assigned. Notes could be added (in the "Notes" column of the data entry spreadsheet) detailing what the Analyst thought the finer level identification might be.

All species names should be 'accepted' within the World Register of Marine Species[3] (WoRMS), i.e. listed in the drop-down field in the "Taxa" column. Three additional entries in that list were: 'U. faunal turf', 'U. faunal crust' and 'No identifiable taxa'. When entering the

[3] WoRMS. URL: http://www.marinespecies.org/ (date last visited: 03/04/19).

"Taxa", if spelt incorrectly or not on the list, a notification informed the Analyst of a potential error. If a required taxon was not on the list, it was verified online on the WoRMS website and then copied into the field from another cell within the spreadsheet. Where detailed taxonomic identification was not possible a description of lifeform was given as a note or qualifier in the "Notes" column, e.g. "erect branching hydroid", with the higher taxonomic level recorded, e.g. "Hydrozoa" or "Porifera".

A variety of taxonomic guides and literature were used along with websites to confirm and assist with identifications. These included the Marine Life Information Network[4] (MarLIN) and the Encyclopaedia of Marine Life of Britain and Ireland[5] (Habitas).

2.5 Image analysis process

2.5.1 Preparation of the image files

Preparation of the images was required before application of the data extraction methods. Three of the methods require grids of lines (to create equally sized 'cells') or points to be overlaid on the images, and while grids are not required for the other three methods enumeration is often easier and recommended with a grid of cells.

Figure gives examples of an image with the three grid overlays: 100 evenly spaced points (required for Point Intercept method), 25 evenly spaced cells (for 5x5 Frequency method) and 100 evenly spaced cells (for 10x10 Frequency method). Methodology guidelines provided to the analysts recommended that the 10x10 grid images were also used for the Percentage Cover and SACFOR methods. Most analysts used the 5x5 grid images for the Abundance Count method. See Section 2.3 for further explanation.

[4] MarLIN, URL: https://www.webarchive.org.uk/wayback/archive/20130501175401/http://www.marlin.ac.uk/ (date last visited: 03/04/19).
[5] Habitas, URL: http://www.habitas.org.uk/marinelife/ (date last visited: 03/04/19).

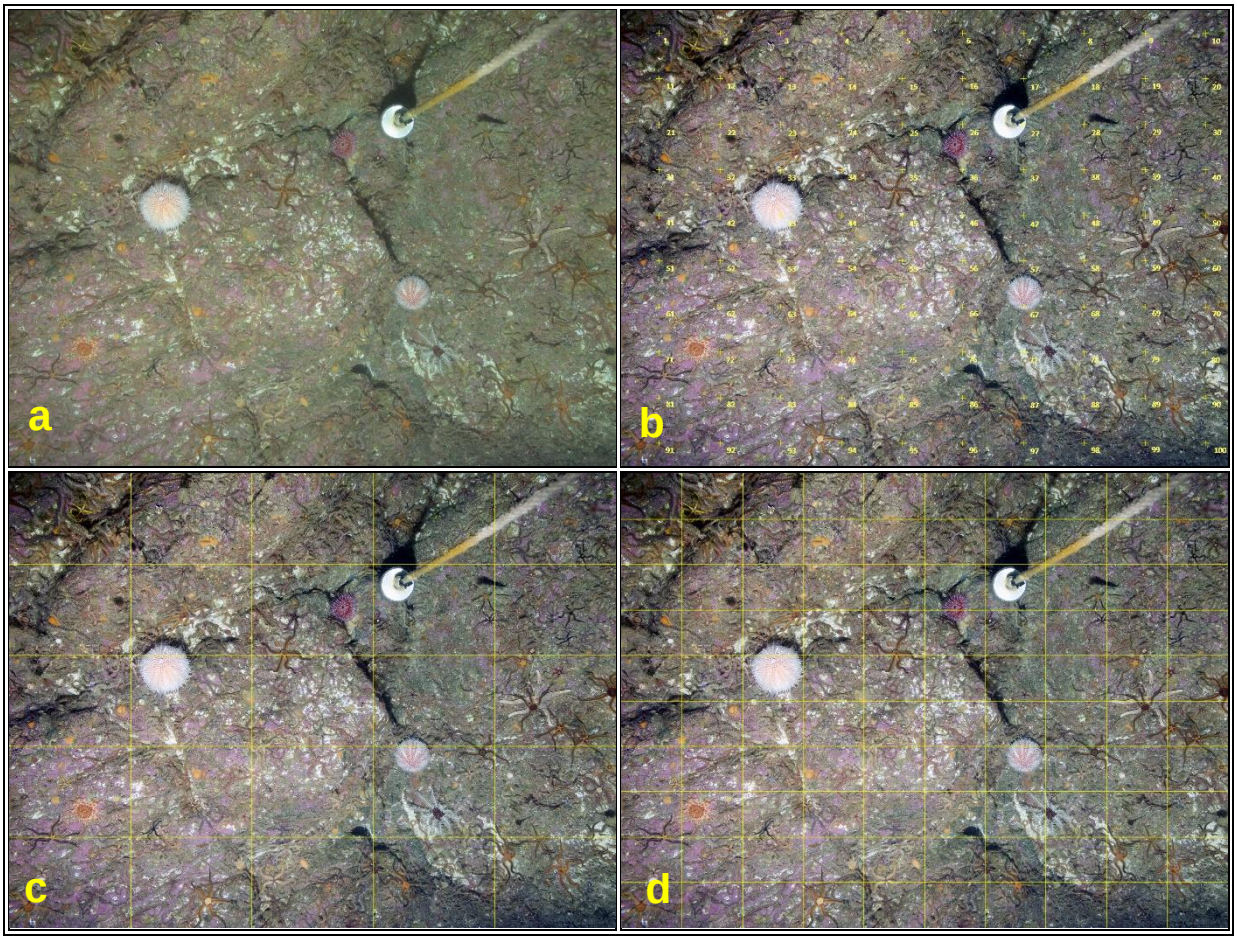


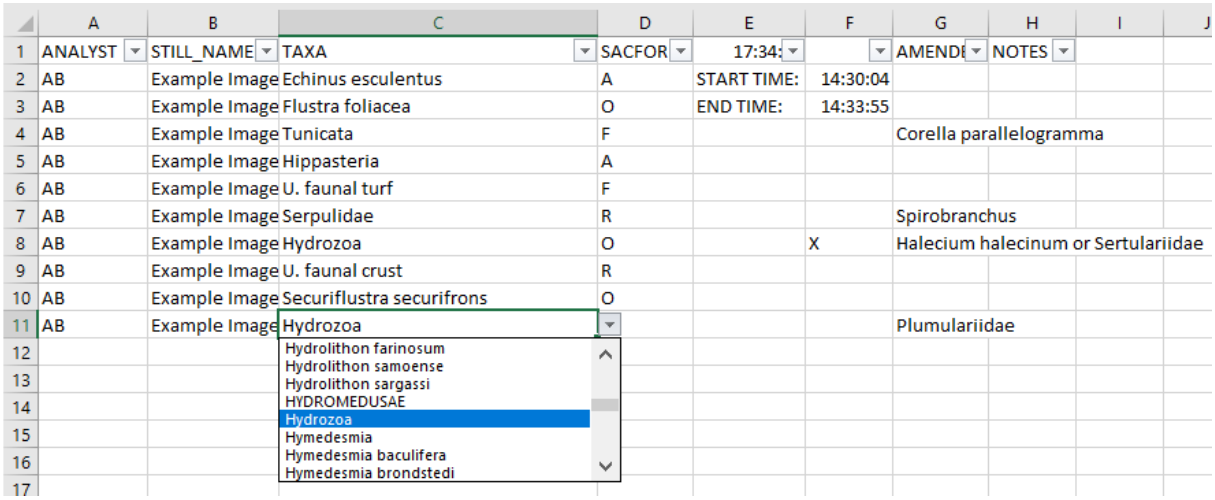
Figure 3. Example image with different grids. a) original image, b) point intercept, c) 5 x 5 grid, d) 10 x 10 grid.

2.5.2 Extraction of biological community data (image analysis)

The images were reviewed, processed and analysed in accordance with national guidelines (including standards for analysis in Visual Seabed Surveys (BS EN 16260:2012), Mapping European Seabed Habitats (MESH; Coggan *et al.* 2007) and NMBAQC (Turner *et al.* 2016)).

Images were viewed on high resolution computer screen with a variety of viewing software, using zoom tools to enlarge the image to assist with recognition. Analyst A analysed all 100 images. Separately, a subset of 20 images (from the 100) were randomly selected (using the random number generator in Excel) and these were also analysed by Analysts 2 and 3 (also Senior), and Analysts D, E and F (Junior). The same 20 images were analysed by all Analysts, but Analyst A did not know which 20 images had been selected. The images were analysed by each Analyst in a random order (using the random number generator) of both the sample image number and of the data extraction method, to reduce any bias in the data collected from each method or from each image and to reduce 'learning effects' among the Analysts. This random order was different for each Analyst. This meant that an analyst might analyse Image 49 with the SACFOR method, then Image 22 with the 5x5 frequency method, then Image 91 with the Percentage Cover method, and so on, in a completely random order. This inevitably took longer for the analysis than normal, as the usual familiarisation process was impeded.

Each Analyst was provided with the prepared images, a spreadsheet listing the random order of images / extraction methods, and a data entry spreadsheet with six tabs: one for each extraction method. The layout of the latter is shown in Figure 2.



	A	B	C	D	E	F	G	H	I	J
1	ANALYST	STILL_NAME	TAXA	SACFOR	17:34:		AMENDE	NOTES		
2	AB	Example Image	Echinus esculentus	A	START TIME:	14:30:04				
3	AB	Example Image	Flustra foliacea	O	END TIME:	14:33:55				
4	AB	Example Image	Tunicata	F			Corella parallelogramma			
5	AB	Example Image	Hippasteria	A						
6	AB	Example Image	U. faunal turf	F						
7	AB	Example Image	Serpulidae	R			Spirobranchus			
8	AB	Example Image	Hydrozoa	O		X	Halecium halecinum or Sertulariidae			
9	AB	Example Image	U. faunal crust	R						
10	AB	Example Image	Securiflustra securifrons	O						
11	AB	Example Image	Hydrozoa				Plumulariidae			
12			Hydrolithon farinosum							
13			Hydrolithon samoense							
14			Hydrolithon sargassi							
15			HYDROMEDUSAE							
16			Hydrozoa							
17			Hymedesmia							
			Hymedesmia baculifera							
			Hymedesmia brondstedi							

Figure 2. Data entry spreadsheet layout, an example for the SACFOR method. Sheets for the other five methods had a different column D.

The data entry spreadsheet also included Lookup tables of taxa (the full UK list from the World Register of Marine Species; WoRMS), Analysts and Abundance categories. These were available as drop-down lists in the relevant columns (see example in Figure 2). The Amended column was used to indicate that an amendment had been made by an Analyst to the records from a previous method/image (e.g. to change a previous identification after seeing a better image of the species). This allowed an Analyst to correct his/her dataset, as one would normally do in a realistic scenario, while indicating (crudely) that some extra time had been taken. The start and finish times to complete data extraction of each image were recorded to measure efficiency.

The image analysis aimed to replicate a realistic survey scenario as much as possible, i.e. where the analysts have no prior knowledge of the site and where no checklist of species was provided. The Analysts were not supposed to copy previous lists from method to method, but to use the drop-down list and assign whatever they considered appropriate to each entity they observed in the images. The lack of any standard checklist resulted in a considerable variety of different taxonomic assignments.

2.6 Reference collection

A reference collection was built up during the analysis process to aid quality assurance procedures, with good quality images noted and collated for each taxon or species identified by each Analyst. Each image was then reviewed and the taxon/species was highlighted (see example in Figure 3) with a box or circle. The file was then saved with the species taxon name and the site identification forming the filename structure.

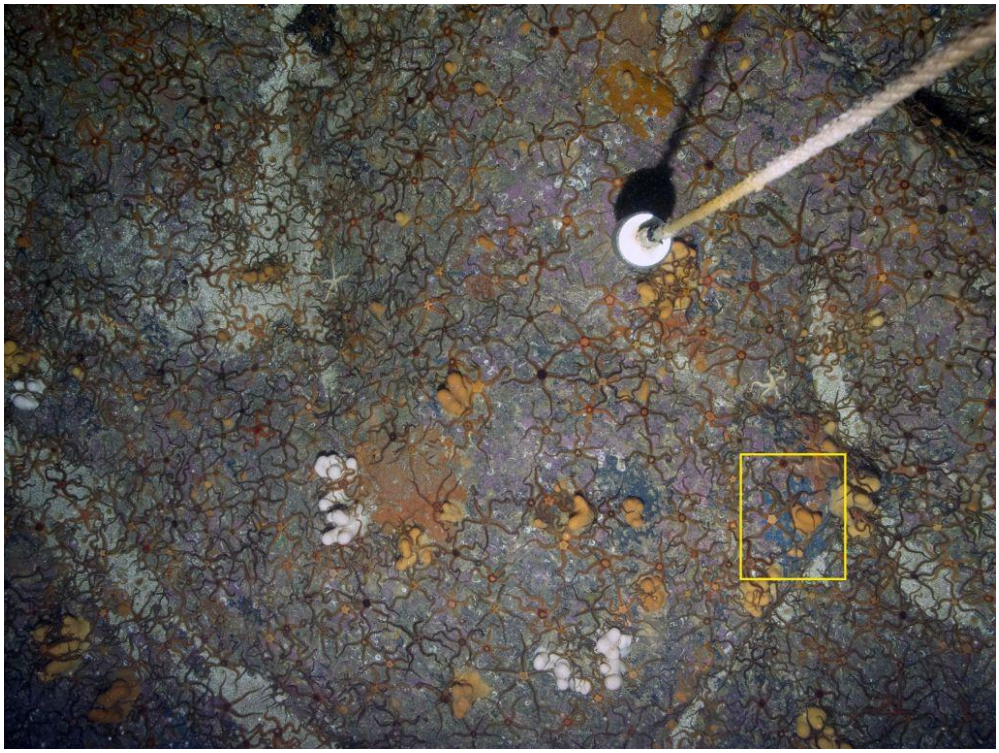


Figure 3. Example reference image: Porifera_blue encrusting_Still_JNCC0093.jpg

2.7 Data preparation and truncation

2.7.1 Data preparation

Data from the individual analysts were initially collated in Excel and multiple quality control checks were carried out, including:

- nomenclature and taxonomy – checking spelling and correct nomenclature using the Taxon match tool on the Marine Species of the British Isles and Adjacent Seas (MSBIAS) website. Species Codes, based on those from the Species Directory (Howson & Picton 1997) were added to the data, to facilitate taxonomic sorting for easier inspection.
- cleaning and tidying – checking to ensure data all in correct format and appropriate for the extraction method used. SACFOR scale data were converted to the numerical equivalent, R=1, O=2, *etc*). Checks that all taxa and their qualifiers were unique and sensible (e.g. no duplicates with different spelling).

The data when then imported into Microsoft Access and restructured into a series of tables (Records, Samples, Images, Methods, Analysts, Taxa) for more secure and robust management.

A Sample can be considered the unique combination of *Extraction Method*, *Analyst* and *Image*. A Record within a Sample consisted of a taxonomic *Entity*[6] and its recorded *Abundance*. These sample metadata fields are tagged as 'Factors' in the Plymouth Routines in Multivariate Ecological Research statistical software package[7] (PRIMER).

[6] **Taxa and Entities** – Entity / Entities are generic terms used here to refer to unique species, taxa, or life forms, sometimes including qualifiers (e.g. enc, erect, tube). The main dataset used in the analyses for this study includes 97 unique entities. However, the more commonly used terms Taxon / Taxa are used in this report to mean Entity / Entities.

[7] PRIMER, URL: https://www.primer-e.com/ (date last visited: 03/04/19).

Additional factors included the time taken to analyse each image (by method and analyst). Various metadata fields (including classification data derived from WoRMS) were added to the Taxa table. In addition, two fields were added to tag the ground cover taxa and the erect/solitary taxa. These taxon metadata fields are tagged as 'Indicators' in PRIMER.

2.7.2 Truncation

The lack of a standardised checklist of taxa (for the analysts to work from) resulted in a considerable variety of different taxonomic assignments and associated notes. Some analysts were very cautious, assigning to a higher-level taxon (e.g. Phylum, Class or Order), with notes that often suggested a Family, Genus or even Species; other analysts were more confident, more often assigning to Family, Genus or Species. After the initial data collation and cleaning it was clear that the extreme variability in taxonomic assignments would complicate comparison of the data extraction methods. This was the case whether analyses were carried out using the Taxa field alone or with a concatenation of the Taxa and Notes field. However, further inspection of the data then showed that a large number of the different assignments (taxon and associated notes) could be re-assigned, to provide a simpler list of taxa and qualifiers while minimising the loss of information (i.e. retaining the essential taxonomic richness). Appendix 2 provides a complete listing of the 417 assigned taxa and associated notes, with the re-assignments to just 97 taxa and associated qualifiers.

While retaining taxonomic richness was an important objective of the re-assignment process, a certain amount of taxonomic aggregation was required. There was no risk of error to data extracted with the Abundance Count, Percentage Cover or Point Intercept methods. However, aggregation of data extracted with the SACFOR, 5x5 or 10x10 methods needed to be treated differently, as either summing or taking the maximum value would likely result in an error. Luckily the number of records requiring aggregation for those methods was not large and the abundances were aggregated manually to minimise likely error.

The final dataset for analysis was still characterised by an extreme variability in taxonomic assignments (see examples in Tables 23 and 24 in Appendix 3).

2.8 Data analysis

The Access database was used routinely for production of data extracts, tabulations and some statistics, using queries and functions. Data extracts were typically exported back to Excel for production of graphs, formatted tables (often using the conditional formatting tools) and statistical analyses. Analysis of variance, including post-hoc Tukey HSD tests, and power analyses were carried out in the statistical package *R*.

Full matrices (raw and standardised) were exported from Access, with multiple factors and indicators, for import into PRIMER. PRIMER 7 and the add-on PERMANOVA were used for multivariate analyses, including production of resemblance matrices (Bray Curtis similarity), nMDS plots (non-metric multi-dimensional scaling), similarity groupings from CLUSTER analyses (Group Average clustering), Species accumulation data, and significance testing with PERMANOVA. All default settings in PRIMER were accepted unless otherwise specified.

2.9 Rank-based method for comparison of data extraction methods

Owing to fundamental differences between the data extraction methods, direct comparison of the results in this study is difficult. Perhaps the two key issues are that the scales of measurement used between the methods are different and that the Percentage Cover and

Abundance Count methods can only record either ground-cover taxa only or erect/solitary taxa only, respectively. These two issues alone complicate any comparison between the methods and can render such efforts meaningless.

To tackle the first issue, it was decided to adopt a rank-based approach as it does not rely on scales of measurement. A rank-based approach, rather, can assess the relative performance of methods against each other and could be used across all data metrics in this study. It is useful to rank results against a baseline result, usually the lowest or highest score in any test. Such an approach could reveal which methods had results that ranked the highest or lowest relative to the others. However, owing to the second issue of taxonomic recording for the Percentage Cover and Abundance Count methods, such an approach would not be fair. Rank-based results could either bias these methods as performing better or worse than the others, without a true comparison across all taxa in the imagery being made.

So, to tackle the second issue it was decided to set a baseline for method comparison by combining the results of the Percentage Cover and Abundance Count methods. As already mentioned, these two methods are frequently carried out together to measure different types of taxa in the community, i.e. those that occupy space and are typically colonial (ground cover taxa) and those that grow in a solitary fashion (erect/solitary taxa). Thus, a ranked comparison of the results from all the methods in the study, across the data metrics, against the combined results from the Percentage Cover and Abundance Count methods would show how the performance of alternative methods compares against the methods most commonly used for imagery analysis work currently in the UK.

Next a broad selection of results from the data metrics were examined to enable robust comparison of the methods across all data metrics investigated in this study. In total, 33 datasets were assessed, including nine Precision datasets, eight Power datasets, two Efficiency datasets, two Taxonomic Richness datasets, the one Taxonomic Accumulation dataset, two Community Impression datasets and eleven Consistency datasets (See Appendix 3 (Table 29) for full details). Where multiple datasets were used for a data metric an average was calculated for that metric. In every instance, the results of every method were compared to the results of the combined Percentage Cover and Abundance Count methods. The rank score for a method increased negatively or positively according to how it compared to Percentage Cover and Abundance Count methods results as well as the other methods results. Where the results were the same as the Percentage Cover and Abundance Count methods then the rank score was ‘0’. Where the results were the same as another method the rank score was tied. Finally, an overall average rank score was calculated from all rank scores in the study to enable relative comparison of the methods performance across the entire study.

3 Results and Discussion

3.1 Summary description of data

The final dataset prepared for analysis comprises 8269 records from 1200 samples (each sample defined by the combination of image, analyst and extraction method). A total of 97 taxa were recorded, although the highest number of taxa recorded by any individual analyst was 64. Table 2 lists the most frequently recorded taxa by Analyst A and highlights the dominance of five taxa: Unidentified faunal crust, Unidentified faunal turf, *Spirobranchus*, Ophiuroidea and *Ophiocomina nigra*.

The range of recorded abundance values varied between methods: Percentage Cover: 0 to 98%, Abundance Counts 0 to 344 individuals, SACFOR: 0 to 6, 5x5 Frequency 0 to 25 cells, 10x10 Frequency 0 to 100 cells, Point Intercept 0 to 90 points.

The area of seabed within a single image was not measured precisely but estimates ranged from approximately 0.3m^2 to 1m^2. Using these estimates, the total area of all 100 images combined was 79.8m^2. The total area of designated reef habitat within which the images were collected was estimated to be approximately 387km^2 (JNCC 2012), so the proportion of the reef captured by the images was approximately 0.000002%.

Table 2. List of taxa (in taxonomic order) recorded by Analyst A from >5 images (max =100 images).

Taxa	**Images**
U. faunal crust	100
U. faunal turf	95
U. faunal turf (Hydrozoa or Crinoidea)	18
U. faunal turf (Porifera or Bryozoa)	10
Porifera (enc blue)	33
Hydrozoa	47
Nemertesia antennina	12
Alcyonium digitatum	19
Actiniaria	20
Urticina	24
Scleractinia	13
Caryophyllia smithii	58
Serpulidae	69
Spirobranchus	100
Decapoda (crab)	6
Caridea	12
Munida	12
Trochidae	32
Bryozoa (staghorn)	6
Flustridae	28
Flustra foliacea	15
Securiflustra securifrons	19
Antedon	6
Asteroidea	14
Luidia ciliaris	7

Crossaster papposus	19
Ophiuroidea	93
Ophiothrix fragilis	40
Ophiocomina nigra	91
Echinus esculentus	7
Tunicata	9

3.2 Frequency distributions of data

An appreciation of how extraction method affects the distribution of data is a useful initial comparison. Table 3 gives examples of frequency distributions of data, for selected (characterising and most frequently recorded) taxa, showing notable effects of the methodology. Thus, in the Unidentified faunal crust ('U.faunal crust') data, the Point Intercept data has an approximately normal distribution and shows a wide range of values, while data from the other methods are skewed. The Percentage Cover data is least skewed and still describes a wide range of values, while the 5x5 Frequency data is highly skewed with 91 images having the maximum values (i.e. 25 cells). The SACFOR data is also very skewed but distinguishes a greater range of values. For this taxon, which was generally common in the form of patches of variable size scattered across each image (i.e. not clumped), the point intercept and percentage cover methods provide a good representation of ground cover, and thereby of the physical quantity of crust material.

The 5x5 and 10x10 Frequency methods do not provide a good representation of ground cover, but they do represent the frequency by which the 'U. faunal crust' occurs. Is this more important for assessing condition of that taxon group or population? It would be very possible to have a notable change in percentage cover with very little change in frequency of occurrence. The opposite is also possible, but less likely for that taxon. In some circumstances it would also be possible to have an increase in one and a decrease in the other (e.g. fewer larger patches after a period of steady growth but poor recruitment, or more smaller patches after a period of decline followed by recent high recruitment).

The Unidentified faunal turf ('U. faunal turf') data has a very different pattern of data distribution, with the 5x5 Frequency data showing the widest spread of values while the Point Intercept and Percentage Cover data are strongly skewed to the low abundances. This is an example of a taxon present in most images, as variable sized patches, but fairly clumped.

The frequency distributions for data representing cup corals and the brittle star, *Ophiocomina,* provide examples of erect/solitary taxa, each showing notably different effects of the extraction methods. The cup corals were generally well scattered within each image but they were often present in low abundance, so most of the distributions are strongly skewed to that end. The SACFOR method, however, is much less skewed, showing the effect of the semi-quantitative scale for small individuals. The Point Intercept method often missed the cup corals present in the images, producing a very strongly skewed distribution. The *Ophiocomina* had a clumped spatial distribution, with large densities observed in a small number of images, but otherwise fairly low abundances. The frequency distributions for counts and point intercepts are, therefore, strongly skewed to the lower percentiles. As for the SACFOR method, however, the frequency distribution was skewed to the upper percentiles. The frequency of occurrence methods generated *Ophiocomina* data that showed a broad spread across the percentiles.

Table 3. Comparison of frequency distributions of the data extracted by the six extraction methods, for selected taxa. Numbers within each cell are the number of images for a particular percentile value. Data recorded by Analyst A from 100 images (i.e. each column adds up to 100). Each row represents a range of percentiles in multiples of 10, i.e. represents 1/10th of the range of data recorded for that taxon.

Unidentified faunal crust frequency distribution of data

percentile	%c	As	Co	F0	F5	Pi
0.1	2	1		3	6	1
0.2	3					3
0.3	4	1		1		10
0.4	6					12
0.5	11	1				15
0.6	14					20
0.7	10	14				16
0.8	21			6		12
0.9	19	53		4	3	6
1	10	30		86	91	5

Unidentified faunal turf frequency distribution of data

percentile	%c	As	Co	F0	F5	Pi
0.1	11	14		17	16	39
0.2	80			29	9	19
0.3	6	47		19	17	22
0.4	2			14	14	9
0.5		35		11	9	5
0.6				4	6	2
0.7		3		1	11	
0.8				1	4	2
0.9					4	1
1	1	1		4	10	1

Cup coral frequency distribution of data

percentile	%c	As	Co	F0	F5	Pi
0.1		58	55	50	53	97
0.2			36	29	18	
0.3			3	9	14	
0.4			3	7	1	2
0.5			2	1	5	
0.6				2	3	
0.7		24			1	
0.8					2	
0.9		15			2	
1		3	1	2	1	1

Ophiocomina frequency distribution of data

percentile	%c	As	Co	F0	F5	Pi
0.1		13	10	11	10	25
0.2			38	11		48
0.3			23	20	10	16
0.4			15	16	10	9
0.5			5	13	15	
0.6			4	12	2	1
0.7			3	9	10	
0.8		16		5	13	
0.9		61		2	16	
1		10	2	1	14	1

Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

3.3 Precision

Precision (Standard Error / Sample Mean) measures the overall variability in the data across all images, relative to the sample mean, and it improves (value reduces) with increasing numbers of samples. It is often used with pilot survey data to aid decisions on the number and size of sampling units for ongoing survey design, but it is here used to compare the variability in data collected with the six extraction methods.

Precision was calculated from abundance data for taxa recorded by Analyst A from 100 images. Taxa that were recorded from fewer than five images by any extraction method were excluded. The full results table is given in Appendix 3 and shows that there is, as expected, a strong correlation between precision and the number of records with abundances >0). It was not always possible to compare all data extraction methods for all individual taxa due to the paucity/absence of positive records associated with some methods. In particular, Point Intercept contained few positive records for many taxa.

The data for individual taxa (Table , Appendix 3) show complex patterns but the differences between any of the extraction methods appear to be small. However, an attempt was then made to look for patterns in the precision data with respect to the life form of the taxa. Life forms were assigned from knowledge of the taxa and inspection of the images. Initially, an attempt was made to test the effect of life form with a two-way ANOVA, but the results showed nothing meaningful, due to the strong influence of the number of records. The precision values were then averaged by life form and the results are given in Table 4. They suggest that out of 9 growth forms, SACFOR was always one of the most precise and that in 6 of the 9 growth forms, the 5x5 Frequency method was also relatively precise. It is likely that this is due to the reduced discrimination of both of those methods, i.e. as they both have the smallest data ranges of all methods used in this study.

Table 4. Precision (Standard Error / Sample Mean), averaged within life-form categories, for all taxa recorded frequently by Analyst A from 100 images.

Form	Records	%c	As	Co	F0	F5	Pi
Crust/Bed	489	0.04	0.02		0.02	0.03	0.03
Crust_dis	112	0.34	0.19		0.25	0.24	0.33
Indiv_lrg	212		0.30	0.31	0.35	0.33	0.48
Indiv_med	139		0.45	0.46	0.48	0.50	0.64
Indiv_ram	1139		0.27	0.30	0.31	0.29	0.31
Indiv_sml	227		0.31	0.44	0.38	0.28	0.66
Massive	70	0.40	0.27		0.35	0.34	0.45
Turf	902	0.40	0.25		0.34	0.29	0.45
Turf_indiv	513	0.48	0.39		0.41	0.36	0.49

Form is a categorisation of the life form of each taxon: Crust/bed= Encrusting/bed forming; Crust_dis= Encrusting and distinct; Indiv_lrg= large individuals; Indiv_med= medium sized individuals; Indiv_sml= small individuals; Indiv_ram= ramose individuals; Massive= massive form; Turf= turf forming; Turf_individ= ground covering individuals. Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

3.4 Sampling Power

Power analyses were carried out using the statistical package R, with functions (in the *emon* package) written by Barry and Maxwell (2017) and scripts and guidance developed within JNCC (Griffiths 2018). The *emon* functions were designed for benthic (infaunal) data, with abundances as whole numbers, so the percentage cover data (with some values <1), were

first multiplied by 10. With large numbers of zeros in the data for some taxa it was expected that a transformation appropriate to a negative binomial distribution should be used. Where sufficient suitable data existed, however, calculations were carried out using a normal distribution (indicated in Table 5).

The power analyses were carried out with the data for Analyst A only (100 images). Taxa selected for analysis were ones for which there was sufficient data for the purposes of making comparisons.

The results in Table 5 are presented for comparison between the extraction methods and show that the minimum number of samples required (as a proxy for statistical power) varies greatly for different taxa. However, before direct comparison is made it must be appreciated that the abundance data from the six extraction methods are on different scales. Thus, most obviously, a 20% change in abundance of the different methods will result in different abundances. For example, a 20% change in 10x10 Frequency data (range 0 to 100), from 60/100 to 80/100, would be equivalent to a change of 15/25 to 20/25 for 5x5 Frequency data (range 0 to 25). Most notably, the SACFOR method is on a categorical scale (range 0 to 6) and each increment on that scale is supposed to follow a logarithmic scale, so a 20% change in mean SACFOR abundance is illogical. However, for the purposes of broad comparison in this study, these issues are being overlooked. In a real-world application of the Power data metric, via power analysis, care would be used to ensure data are fit for purpose and that the change to be estimated is meaningful.

A general comparison of the data extraction methods shows SACFOR to have the lowest sample estimate in 5 out of 8 data-sets, with low estimates in all other data-sets (Table 5). This makes it the method with the most power overall in this study. The Frequency grid methods have the next lowest sample estimates, with the 5x5 being slightly lower overall compared to the 10x10 grid (Table 5). When considering that both SACFOR and Frequency 5x5 grid methods have the smallest data ranges in this study, it is likely this contributes to their higher power estimates. It is likely that any method with a small data range has less potential to vary within that range relative to a method with a larger data range. This will enable a higher sampling power overall.

By removing the SACFOR and 5x5 Frequency method data from this analysis, we are better able to consider methods of more similar data ranges, such as the Percentage Cover, Abundance Count, Point Intercept and 10x10 Frequency methods. Among these four methods, the data generated by the 10x10 frequency method has the lowest sampling estimates in 6 out of the 7 taxa compared (Table 5). Interestingly, when taxonomic richness is compared between the four methods the lowest sample numbers (17), and the highest power, are estimated for the point intercept method, with the 10x10 Frequency method having the second lowest estimate (34), which is double that of the Point Intercept method. This is likely because the Point Intercept method intercept method generated data that contained fewer records with values greater than 0 (zero).

Table 5. Number of samples required to detect a significant 20% change in mean values of selected taxa, with a power of 0.8, by extraction method, using data from only Analyst A (100 images). Calculated using emon function power.groupsByFactors in R, with distribution set as listed in table and with 1000 simulations (p=<0.05).

Taxa	%c	As	Co	F0	F5	Pi	Distribution
U faunal crust	65	13		17	25	41	Normal
U faunal turf	2000	145		550	250	700	Normal
Hydrozoa	3400	800		1400	1200	2800	Normal
Serpulidae	1300	90		70	50	300	Normal
Cupcorals		1100	2150	1600	1600	10,000	Negative binomial
Ophiocomina		55	450	185	120	650	Normal
Calliostoma		3400	2000	2200	2400		Negative binomial
Richness	45	29	84	34	27	17	Lognormal

Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

3.5 Efficiency

Time taken to analyse the images will always be an important consideration as it will have large effects on budget.

Randomization of the image analysis, as described in Section 2.5, resulted in a much longer analysis time, but was necessary to minimise the effect of the inevitable familiarity as analysts worked through the images. Most importantly it minimised the potential bias in analysis time between extraction methods. Assessment of the effectiveness of the randomization process was carried out before a comparison of the methods. The results of that assessment are given in Appendix 3 and clearly demonstrated a lack of bias to any method.

The results in Figure 4 and Table 7 clearly show large differences in analysis time, with the Point Intercept and 10x10 Frequency methods taking much longer than other methods and percentage cover taking the least time.

[Note: the protocol for the percentage cover method also included recording erect/solitary taxa with abundance recorded as N/A, so the analysis time would have been even less if they had not done that].

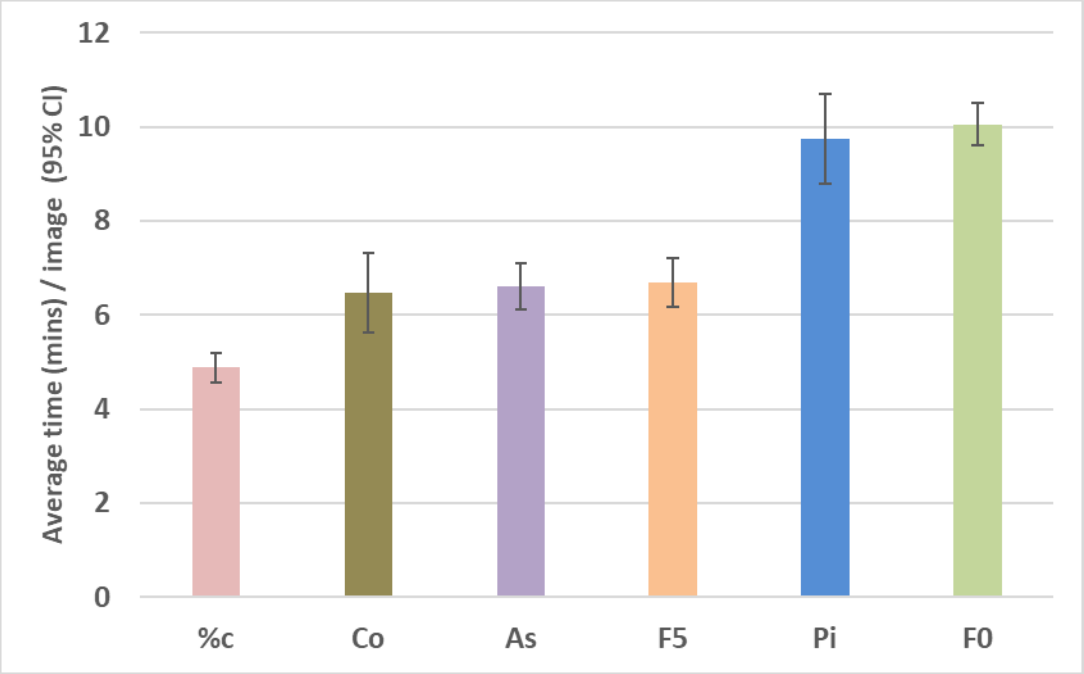


Figure 4. Average (with 95% confidence interval) of minutes taken to analyse each image (N = 100) by Analyst A, by extraction method. Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

A one-way analysis of variance (on untransformed data, following checks for normal distribution) showed that there was a highly significant difference between the methods [F = 41.2, p <0.0001]. A post hoc test (Tukey HSD) showed significant differences between many of the methods (Table 6). Of more interest is a lack of difference between the times taken to extract data using the SACFOR, Abundance Count and 5x5 Frequency methods, as well as between the Point Intercept and 10x10 Frequency method data (Table 6). This shows the SACFOR, Abundance Count and 5x5 Frequency methods are as efficient as each other, and the Point Intercept and 10x10 Frequency methods are also as efficient as each other at extracting community data from the sample images.

Table 6. Results of a post hoc test (Tukey HSD) showing the significance of differences in time taken by Analyst A to analyse each image using six different extraction methods. Significance levels: ns = not significant, * 0.05, ** 0.01, *** 0.001.

	%c	As	Co	F5	Pi
As	**				
Co	**	ns			
F5	***	ns	ns		
Pi	***	***	***	***	
F0	***	***	***	***	ns

Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

Table 7 provides data from all of the analysts and shows that the results from Analysts B to F follow much the same pattern to those from Analyst A, but there are some differences. Most notably the data show some large differences between the length of time spent by different analysts. Thus, Analyst B (a senior analyst) took less than half the time spent by Analyst C (another senior analyst), and the difference was statistically highly significant (p<0.0001). Analysts C also took longer than any other analyst, but overall the Junior analysts took approximately 20% longer, on average, than the Senior analysts. That difference was also statistically significant [F = 15.3, p <0.0001]. An interesting result is that

in general the senior analysts seemed to be proportionately quicker at estimating SACFOR data. This could indicate that sufficient experience is required to use this method efficiently.

The results in Table 7 suggest that if Percentage Cover or Abundance Count, were used together their total averaged time would be very similar to the times to use the 10x10 Frequency and Point Intercept methods. In practice, the Percentage Cover or Abundance Count methods are rarely used in isolation; they are usually used together to estimate different aspects of the benthic community. Thus, comparing methods (or combinations of methods) that describe the whole community, finds the SACFOR and 5x5 Frequency methods to be the most efficient overall.

Table 7. Average number of minutes to analyse an image, by extraction method and analyst. Data from 100 images for Analyst A, but only 20 images for the other analysts.

Data extraction method	A	B	C	D	E	F	Avg mins
Percentage Cover	4.9	2.7	6.9	6.3	8.7	4.3	5.6
Abundance Count	6.5	4.3	10.2	6.1	8.1	3.8	6.5
5x5 Frequency	6.7	4.8	13.0	6.0	12.8	7.7	8.5
SACFOR	6.6	3.7	10.0	12.3	10.4	8.8	8.6
10x10 Frequency	10.1	8.2	17.8	12.3	16.2	9.3	12.3
Point Intercept	9.7	7.7	20.1	13.1	12.2	11.2	12.3
Avg mins	7.4	5.2	13.0	9.3	11.4	7.5	

3.6 Taxonomic richness

There were notable and statistically significant differences in the number of taxa recorded by some extraction methods and by each analyst ($F = 159.5$, $p < 0.0001$, Table 8). Figure 5 and Table 9 shows that the Percentage Cover method and Abundance Count methods captured fewer taxa than any of the other methods. This was expected, given the extraction protocols described in Sections 2.3.1 and 2.3.2. Furthermore, Table 9 shows that when the data are confined to ground covering taxa or erect/solitary taxa, then the differences between the methods mostly disappear. The expected exception is the Point Intercept method, which consistently captured fewer taxa, due to the much smaller proportion of each image analysed.

Table 8. Results of a post hoc test (Tukey HSD) showing the significance of differences in number of taxa recorded using six different extraction methods. Significance levels: ns = not significant, * 0.05, ** 0.01, *** 0.001.

	%c	As	Co	F0	F5
As	***				
Co	ns	***			
F0	***	ns	***		
F5	***	ns	***	ns	
Pi	***	***	***	***	***

Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

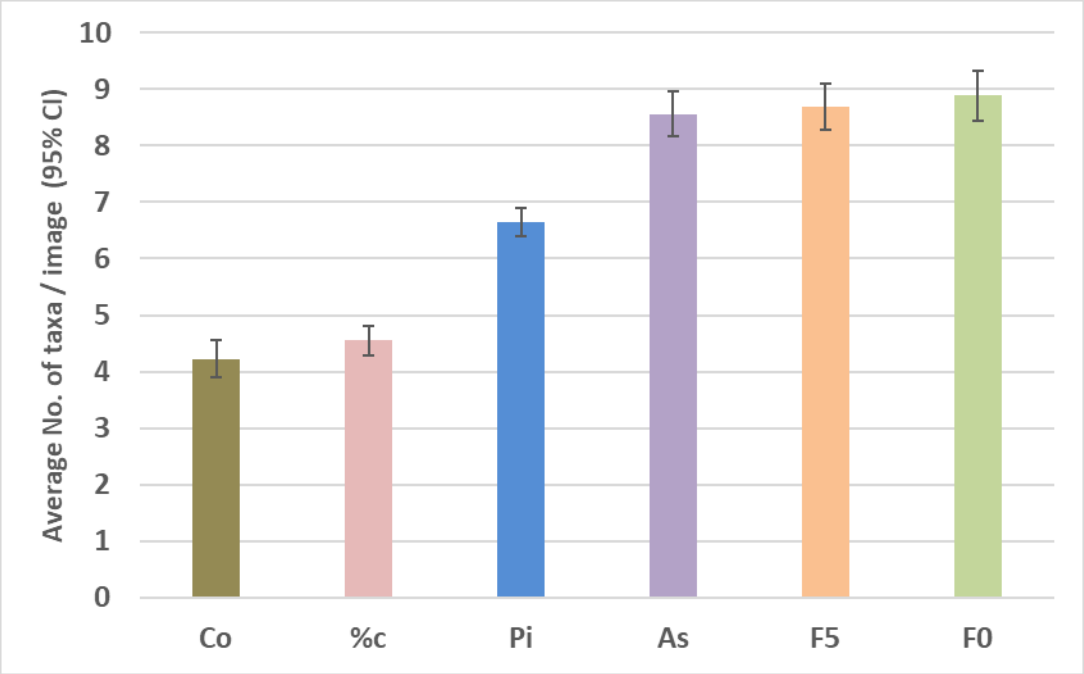


Figure 5. Mean number of taxa (with 95% confidence interval) recorded per image by Analyst A, by extraction method. Data from 100 images. Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

The data in Table 9 shows, as expected, that Senior analysts recorded more species than Junior analysts (on average and in total). However, it also shows that there were large differences between individual Senior analysts; Analyst C recording considerably more taxa than Analyst B (on average and in total).

Further inspection of the data shows that the average number of taxa recorded by the Percentage cover method was slightly higher than that by the Abundance count method; however, the total number of taxa recorded by the latter was much larger than the former. See Section 3.7 for an explanation.

The likely reason why the SACFOR, 10x10 and 5x5 Frequency methods extracted statistically similar numbers of taxa per image (Figure 5; Table 8) is that they enabled Analysts to sample 100% of the image area and all the taxa in it. However, it is notable then that they each have slightly different mean numbers of taxa per image (Figure 5) and slightly different total numbers of taxa per image (Table 9). This slight, but not significant, difference between these results shows that the Analysts recorded slightly different numbers of taxa each time they viewed the same sample image. This issue highlights a potential source of error when requesting image analysts to extract full taxa lists from imagery.

Table 9. Average number of taxa recorded per image and total number of taxa recorded, by each analyst, from 100 images (Analyst A) or 20 images (other analysts), by extraction method.

	Senior analysts			Junior analysts		
	A	B	C	D	E	F
All taxa	100 images			20 images		
Percentage Cover	4.55	3.9	7.1	3.4	5.2	3.2
Abundance Count	4.22	3.75	6.9	2.95	4.85	2.7
SACFOR	8.56	7.05	13.6	6.1	9.45	5.9
Point Intercept	6.64	6	9.65	5.1	6.7	5.3
10x10 Frequency	8.89	7.6	14.25	6.45	9.55	5.95
5x5 Frequency	8.69	7.4	14.3	6.2	9.25	5.95
Ground cover taxa						
Percentage Cover	4.5	3.9	7	3.4	5	3.2
Abundance Count	n/a	n/a	n/a	n/a	n/a	n/a
SACFOR	4.46	3.95	7	3.3	4.9	3.2
Point Intercept	3.42	2.9	5.9	2.7	3.75	2.8
10x10 Frequency	4.53	3.95	7.4	3.3	4.9	3.25
5x5 Frequency	4.43	3.9	7.45	3.3	4.9	3.25
Erect & solitary taxa						
Percentage Cover	n/a	n/a	n/a	n/a	n/a	n/a
Abundance Count	4.2	3	6.65	2.9	4.6	2.65
SACFOR	4.03	2.95	6.55	2.8	4.3	2.65
Point Intercept	2.2	2.1	2.65	1.4	1.95	1.5
10x10 Frequency	4.31	3.35	6.75	3.15	4.4	2.7
5x5 Frequency	4.19	3.3	6.75	2.9	4.2	2.65
Total number of taxa						
Percentage Cover	22	12	22	10	13	7
Abundance Count	35	32	39	17	27	18
SACFOR	57	33	58	26	35	25
Point Intercept	37	22	36	22	20	17
10x10 Frequency	57	43	58	26	37	24
5x5 Frequency	55	39	58	28	36	25

3.7 Taxonomic accumulation

Analysis of taxonomic accumulation was carried out in PRIMER, using only the data collected by Analyst A from 100 images. Accumulation plots were made using different subsets of taxa (all taxa, ground cover taxa only and erect/solitary taxa only), first using the original order that the images were analysed, secondly using multiple permutations (max 999) of the order to average the cumulative count and smooth the graph. Figure 6 show the results for all taxa using the permutated order. Other plots are given in Appendix 3.

Figure 6 shows how the rate of taxonomic accumulation reduces with increasing number of images. The line for Percentage Cover levels off with relatively few sample images, appearing to reach the asymptote at fewer images (approx. 50 images) compared to lines for the other methods. The methods that have the greater representation of taxa (SACFOR, 10x10 and 5x5 Frequency) have steeper curves with fewer sample images but they require many more samples to reach the asymptote (approx. 80-100 samples). Indeed, these latter methods may not reach the asymptote until well after 100 images.

Overall, the rate of taxonomic accumulation for all methods reduces considerably after 20 images. It would be interesting to compare that for different habitats, with different groups of taxa. Figure 7 plots taxonomic accumulation, from only 20 images, for each analyst using the extraction method that gave most taxa (10x10 Frequency). The lines for the junior analysts, particularly D & F, level off with fewer sample images than those for the senior analysts.

Lastly, comparison of Figure 6 (100 images) with Figure 7 (20 images) shows that more taxa were recorded by Analyst A from the selected subset of 20 images (43 taxa) than from a permutated subset of 20 from the full set of images (approx. 37 taxa). This means that the randomised selection of 20 images still resulted in a collection of images with more conspicuous taxa than the average.

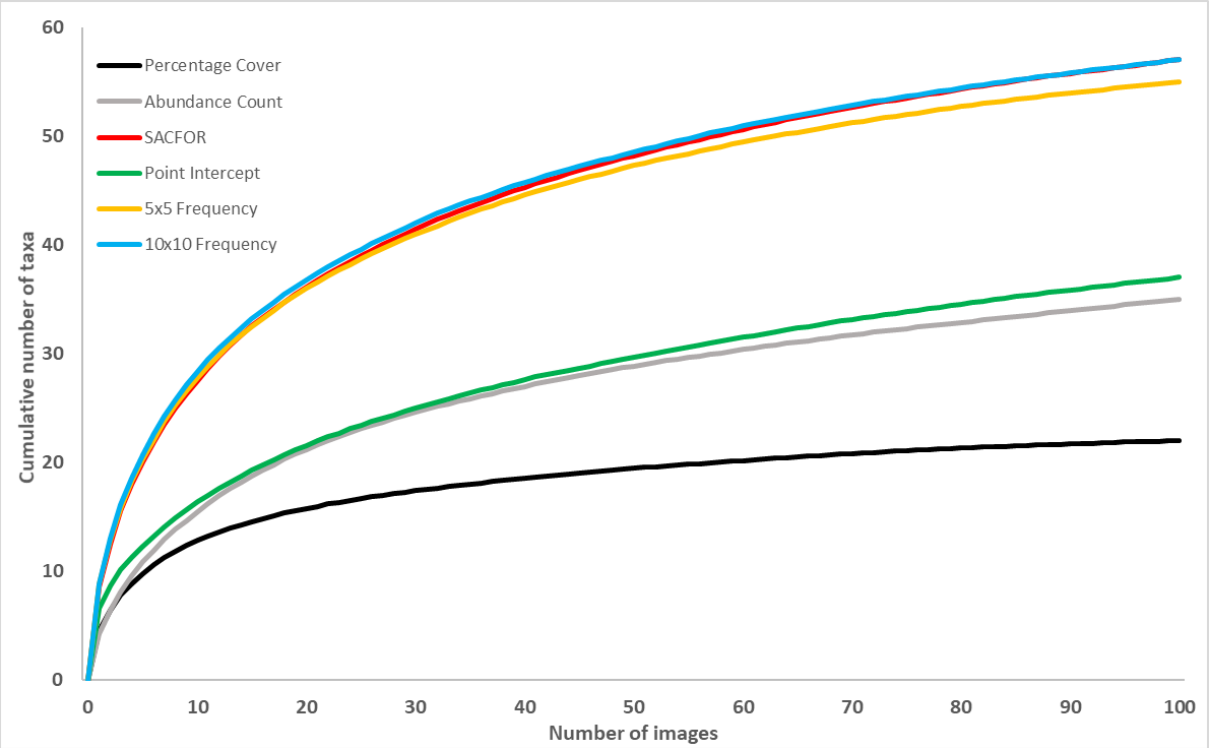


Figure 6. Cumulative number of taxa recorded from 100 images (in permutated order) by Analyst A, by extraction method.

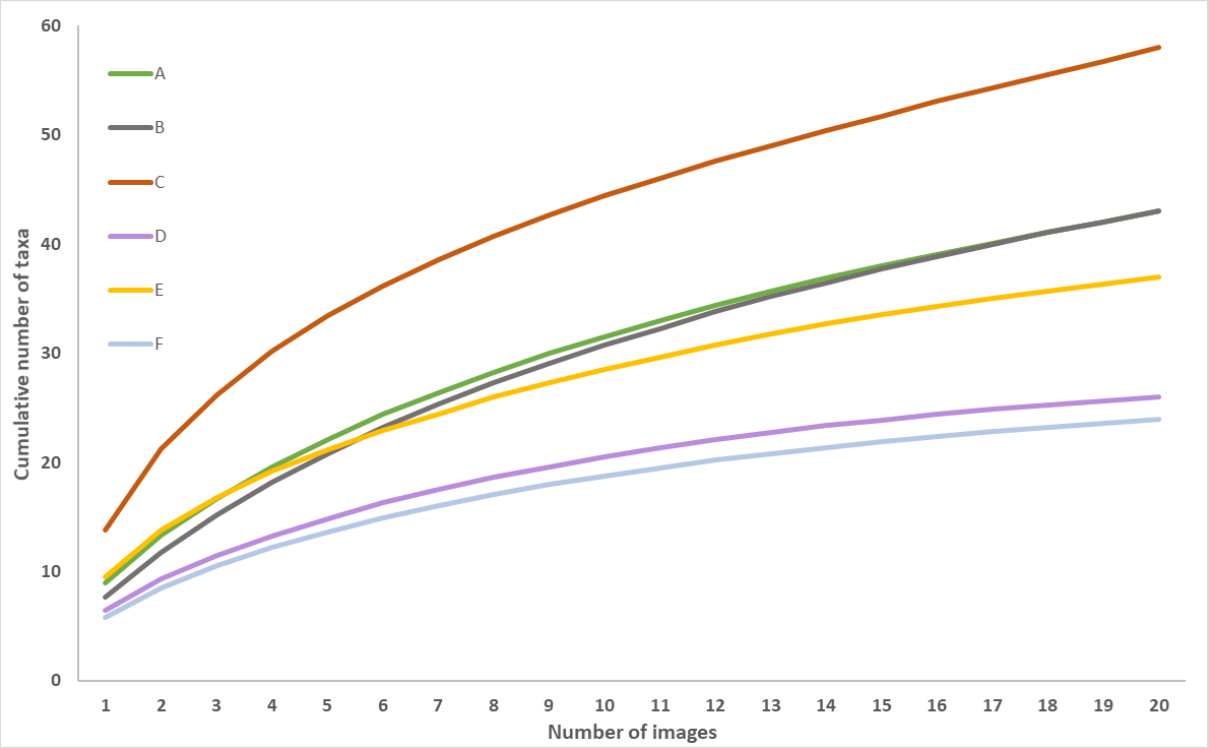


Figure 7. Cumulative number of taxa recorded from 20 images (in permutated order) using 10x10 Frequency method, by Analyst. Analysts are represented by letters A-F.

3.8 Community impression

The plots and discussion in Section 3.1 has already shown that different taxa, within the broad groups of ground cover taxa and individual erect/solitary taxa, have different distributions that are captured in different ways by the extraction methods. This section considers this in more detail and looks for methods that provide similar impressions of the community. Univariate and multivariate analysis approaches were used to explore the ground cover taxa and erect/solitary taxa separately.

3.8.1 Univariate analyses

Ground cover taxa

A notable part of the benthic community at the Solan Bank Reef SAC is comprised of colonial 'ground covering taxa', particularly sponges, soft corals, hydroids, bryozoan, and a few sedentary non-colonial taxa, e.g. anemones and serpulid worms. These latter taxa are also considered here as ground covering taxa. Some of those taxa were difficult to distinguish in the photographs, so the most frequently recorded ground covering entities were *unidentified faunal crust* (U. Faunal crust) and *unidentified faunal turf* (U. Faunal turf).

Notionally (i.e. without considering accuracy), percentage cover can be considered to provide the most representative enumeration measure for ground cover taxa in this study (biomass, or a measure of ecological function, might be better, but is not feasible from still images). Indeed, as none of the other methods are as closely related to the actual spatial area covered by each ground cover taxon, Percentage Cover estimates were therefore used as a means of relative comparison between the data acquired by the other methods.

Three other extraction methods provide data that can be directly compared with percentage cover (i.e. 10x10 Frequency, 5x5 Frequency and Point intercept), although the 5x5

Frequency data are on a shorter scale (range 0 to 25). Data from the SACFOR, on a much shorter (range 0 to 6) categorical scale (with different scales for crusts and turfs) is less easily compared so was removed from this comparison.

Analyst A recorded 22 ground covering taxa from the 100 images, 19 of which were recorded more than once. Table 10 compares the average values recorded with each method (except for Abundance Counts), for each of those taxa. It shows that the mean values from the two Frequency methods were almost always much higher than the Percentage Cover values. The ratios and coloured bars in the last four columns of Table 10 highlight the magnitude of that increase. Taxa tending to be present as a small number of small sized patches (e.g. the hydroids *Thuiaria thuja* and *Nemertesia antennina* and the cup form sponges) were grossly over estimated by the Frequency methods (relative to Percentage Cover). However, taxa tending to be present as larger sized patches or densely aggregated individuals (e.g. unidentified faunal crusts, Flustrid bryozoa and Hexacorallia) had mean values that were more similar to the Percentage Cover values.

Of the taxa shown in Table 10, data from the Point Intercept method appeared to differ the least overall from that made using the Percentage Cover. In every instance provided by the taxa, the Point Intercept differed less than the two Frequency methods. These results suggest that the Point Intercept method is the most accurate of the three methods when compared to data from the Percentage Cover method, assuming the latter is the most accurate overall. This result is interesting because the Point Intercept method has the known limitation of sampling only a discrete area of each still image and yet, has seemingly enumerated various taxa with accuracy nonetheless.

Table 10. Comparison of ground cover metrics for relevant taxa, recorded by Analyst A (from 100 images).

	Average values (from 100 images)				Ratio to percentage cover (%c)		
Entity	**%c**	**F0**	**F5**	**Pi**	**F0**	**F5**	**Pi**
Porifera (cup)	0.0	0.1	0.0	0.0	25.00	20.00	0.00
Polymastia	0.0	0.0	0.0	0.0	20.00	15.00	0.00
Alcyonidium	0.0	0.0	0.0	0.0	20.00	10.00	0.00
Thuiaria thuja	0.0	0.1	0.1	0.0	23.33	26.67	0.00
Bryozoa (staghorn)	0.0	0.1	0.1	0.0	14.00	12.00	2.00
Nemertesia antennina	0.0	0.2	0.1	0.0	30.00	21.67	3.33
U. faunal turf (Porifera or Bryozoa)	0.0	0.6	0.4	0.0	20.71	15.36	0.71
Alcyonium digitatum	0.2	1.7	0.9	0.1	7.34	3.93	0.39
U. faunal turf (Hydrozoa or Crinoidea)	0.2	1.4	0.7	0.3	5.65	2.97	1.17
Flustridae	0.3	0.8	0.4	0.2	3.10	1.55	0.89
Porifera (enc blue)	0.3	4.4	2.4	0.4	14.00	7.56	1.21
Hydrozoa	0.3	2.6	1.4	0.5	7.78	4.10	1.49
Securiflustra securifrons	0.4	3.5	1.6	0.6	9.46	4.31	1.70
Hexacorallia	0.5	0.8	0.4	0.3	1.79	0.77	0.71
Serpulidae	0.6	17.1	4.9	0.1	27.25	7.80	0.19
Flustra foliacea	0.7	1.4	0.6	0.6	2.10	0.94	0.84
U. faunal turf	1.2	18.7	9.1	2.2	16.15	7.85	1.88
Spirobranchus	1.4	49.9	15.6	2.3	34.84	10.87	1.58
U. faunal crust	58.7	92.4	23.3	51.9	1.57	0.40	0.88
				Average ratio	**14.95**	**9.14**	**1.00**

Terms: Percentage Cover (%c); Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

Erect/solitary taxa

The Solan Bank Reef community also includes many mobile (e.g. crustacea, snails, starfish and fish), erect (e.g. some tall thin hydroids) and solitary (e.g. anemones, cup corals and tube worms) taxa that are more easily counted than estimated for their ground cover.

While Abundance Counts may be easiest and provide the greatest range of values (range 0 to 344 in this dataset), the range of body size of individuals can be large. As such it does not necessarily provide a good representation of ecological function of the community. However, none of the other methods are as closely related to the actual occurrence of erect or solitary taxa in an image. Much like Percentage cover estimates for ground cover taxa, Abundance Count values were used as a means of relative comparison between the data acquired by the other methods (except Percentage Cover which was not measured for these taxa).

Analyst A recorded Abundance Counts for 35 taxa from the 100 images, 25 of which were recorded more than once. Table 11 compares the average abundance recorded with each method (except Percentage Cover), for each of those taxa.

The ratio between 10x10 Frequency and Abundance Counts data vary, with frequencies mostly higher than counts. The biggest discrepancies are for large bodied animals (e.g. large crabs and starfish) which cover large numbers of cells but are only counted as a single

individual. The ratios for most of the smaller bodied animals (e.g. *Ophiothrix*) were much closer to one (i.e. frequencies represented counts well), but some small closely aggregated taxa (e.g. Scleractinia) were under-represented by frequencies. 5x5 Frequencies had similar characteristics to the 10x10 Frequencies, but more taxa were under-represented, compared to counts. The relationship between Abundance Counts and SACFOR abundance records is also complicated and the ratio between them varies considerably.

Table 11. Comparison of abundances for counted taxa, recorded by Analyst A (from 100 images).

	Average values (from 100 images)					Ratio to Abundance Counts (Co)			
Entity	**Co**	**As**	**F0**	**F5**	**Pi**	**As**	**F0**	**F5**	**Pi**
Actiniaria	0.2	0.3	0.2	0.1	0.0	1.87	1.53	0.93	0.07
Urticina	0.2	0.8	0.6	0.3	0.1	4.81	3.50	1.75	0.69
Scleractinia	1.7	0.1	0.2	0.0	0.0	0.07	0.11	0.01	0.00
Caryophyllia smithii	3.0	1.3	3.5	2.4	0.1	0.45	1.18	0.79	0.02
Sabella (tube)	0.0	0.1	0.0	0.0	0.0	4.00	2.00	0.50	0.50
Decapoda (crab)	0.1	0.2	0.1	0.1	0.0	3.20	1.80	1.40	0.00
Caridea	0.1	0.4	0.1	0.1	0.0	2.57	0.86	0.64	0.07
Paguridae	0.0	0.1	0.1	0.0	0.0	3.00	1.25	1.00	0.00
Munida	0.1	0.2	0.3	0.1	0.0	1.85	1.92	0.69	0.15
Cancer pagurus	0.0	0.1	0.1	0.1	0.0	5.00	5.50	2.50	1.50
Trochidae	0.4	0.6	0.5	0.4	0.0	1.69	1.50	1.11	0.03
Pecten maximus	0.0	0.2	0.0	0.0	0.0	5.33	0.67	0.67	0.33
Antedon	0.1	0.2	0.1	0.1	0.0	2.67	1.50	1.33	0.17
Asteroidea	0.1	0.4	0.2	0.2	0.0	2.85	1.69	1.31	0.08
Luidia ciliaris	0.1	0.3	0.3	0.2	0.1	4.86	4.57	2.43	1.00
Porania pulvillus	0.1	0.1	0.1	0.1	0.0	2.40	1.60	1.20	0.00
Crossaster papposus	0.2	0.6	0.4	0.3	0.0	3.20	1.95	1.50	0.20
Asterias rubens	0.0	0.1	0.1	0.0	0.0	2.67	2.00	1.00	0.00
Ophiuroidea	69.5	4.8	49.3	17.1	8.9	0.07	0.71	0.25	0.13
Ophiothrix fragilis	7.6	1.7	8.0	4.5	0.7	0.22	1.05	0.60	0.10
Ophiocomina nigra	22.7	4.3	31.6	13.5	2.9	0.19	1.39	0.60	0.13
Echinoidea	0.1	0.2	0.1	0.1	0.0	2.86	1.29	0.86	0.29
Echinus esculentus	0.1	0.3	0.3	0.2	0.1	3.75	3.38	2.38	0.75
Tunicata	0.1	0.3	0.1	0.1	0.0	6.40	2.60	1.60	0.60
Gadidae	0.0	0.2	0.1	0.0	0.0	4.25	2.00	1.00	0.00
					Average ratio	**2.81**	**1.90**	**1.12**	**0.27**

Terms: Abundance Count (Co) SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

3.8.2 Multivariate analysis

The data for these analyses (100 images, Analyst A) were first standardised (before importing into PRIMER), using formula:

$$\frac{x}{maxM} x\ 100$$

Where: x = recorded abundance
maxM = the maximum value in the whole dataset for that extraction method

This was found to be better than using the standardise transformation in PRIMER, which would have standardised each sample individually rather than by each method.

Data was then imported into PRIMER and fourth root transformed (all data). Other transformations were considered and inspected with shade plots, and it is likely that each extraction method would have benefited from a different transformation, but fourth root was considered the best overall.

Three subsets of the data were created to explore the community impressions of the a) ground cover taxa (with samples from all extraction methods except Abundance Counts), b) erect/solitary taxa (with samples from all extraction methods except Percentage Cover) and c) all taxa from only four of the methods (SACFOR, Frequency 10x10 and 5x5 grids and Point Intercept methods, without any Abundance Count and Percentage Cover method data), created by the different data extraction methods. Initial tests with erect/solitary taxa dataset resulted in errors ('undefined resemblance between sample:') caused by lack of sufficient data in some samples, including many Point Intercept samples and also samples from particular images (52, 60 & 61). These samples were therefore removed from the analyses.

PERMANOVA testing with a single factor design, using the fixed factor 'Method', showed that the impressions of the community created by each method significantly differed in each of the three subsets of data (Table 12). Pairwise PERMANOVA tests of the fixed factor 'Method' show that these differences in community impression occur across the methods and between data sets, apart from one instance: where the community impressions of the Frequency of Occurrence methods are similar when only ground cover taxa are concerned ($t = 1.6$, $P = 0.051$, permutations = 9949; full results in Appendix 3).

The components of variation in Table 12 indicate the amount of variability of the data within the factor 'Method' and between replicate sample images. Each result may be interpreted as the amount of variability (on a relative scale) that is accounted for within experimental factors and, thus, may serve as a proxy for multivariate variability in general (Anderson *et al.* 2008). In this instance, the highest amount of variability between replicate still image data is in the erect/solitary taxa data set. The highest amount of variability between the different impressions of the community, determined by the different methods, was in the all taxa data set, perhaps owing to the full taxa list being used in this data set and not in the others.

Table 12. PERMANOVA test results for a one-factor PERMANOVA model, using the fixed factor 'Method', with 9999 unrestricted permutations. Tests were conducted on standardised, fourth root-transformed data of (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, frequency of occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover method data). Bold results indicate a significant effect ($P \leq 0.05$). The components of variation (Comp. var.) indicate the multivariate variability of data between the fixed factors, as indicated, and between replicate sample images (Residual). perms = permutations.

Data set	Factors/ Source	df	SS	MS	Pseudo -F	P	Unique perms	Comp. Var.
a. **Ground cover taxa**	**Method**	4	121830.0	30458.0	35.1	**0.0001**	9907	17.2
	Residual	495	429430.0	867.5				29.5
	Total	499	551260.0					
b. **Erect/ solitary taxa**	**Method**	4	158590.0	39649.0	28.0	**0.0001**	9890	19.8
	Residual	482	683750.0	1418.6				37.7
	Total	486	842350.0					
c. **All taxa** (but no count or cover data)	**Method**	3	158570.0	52857.0	56.1	**0.0001**	9901	22.8
	Residual	396	372840.0	941.5				30.7
	Total	399	531410.0					

Non-parametric Multi-Dimensional Scaling (nMDS) plots of the data sets show the similarities and dissimilarities in the community impression between each method, in each data set (Figure 10). The data from the SACFOR and Frequency of Occurrence methods appear to be the most similar is all data sets. It is no surprise that the Frequency of Occurrence grids (10x10 and 5x5) appear to generate data that are the most similar in each data set, as these methods function in a very similar way, differing only by their data resolution. It is interesting that the data made using the Point Intercept method appears similar to that from the methods assumed to be the most accurate, the Percentage Cover and Abundance Count methods. These results support the results from the univariate analyses of community impression (Section 3.8.1), which highlight that although the Point Intercept samples a discrete area of each still image it seems to generate an accurate impression of the community nonetheless.

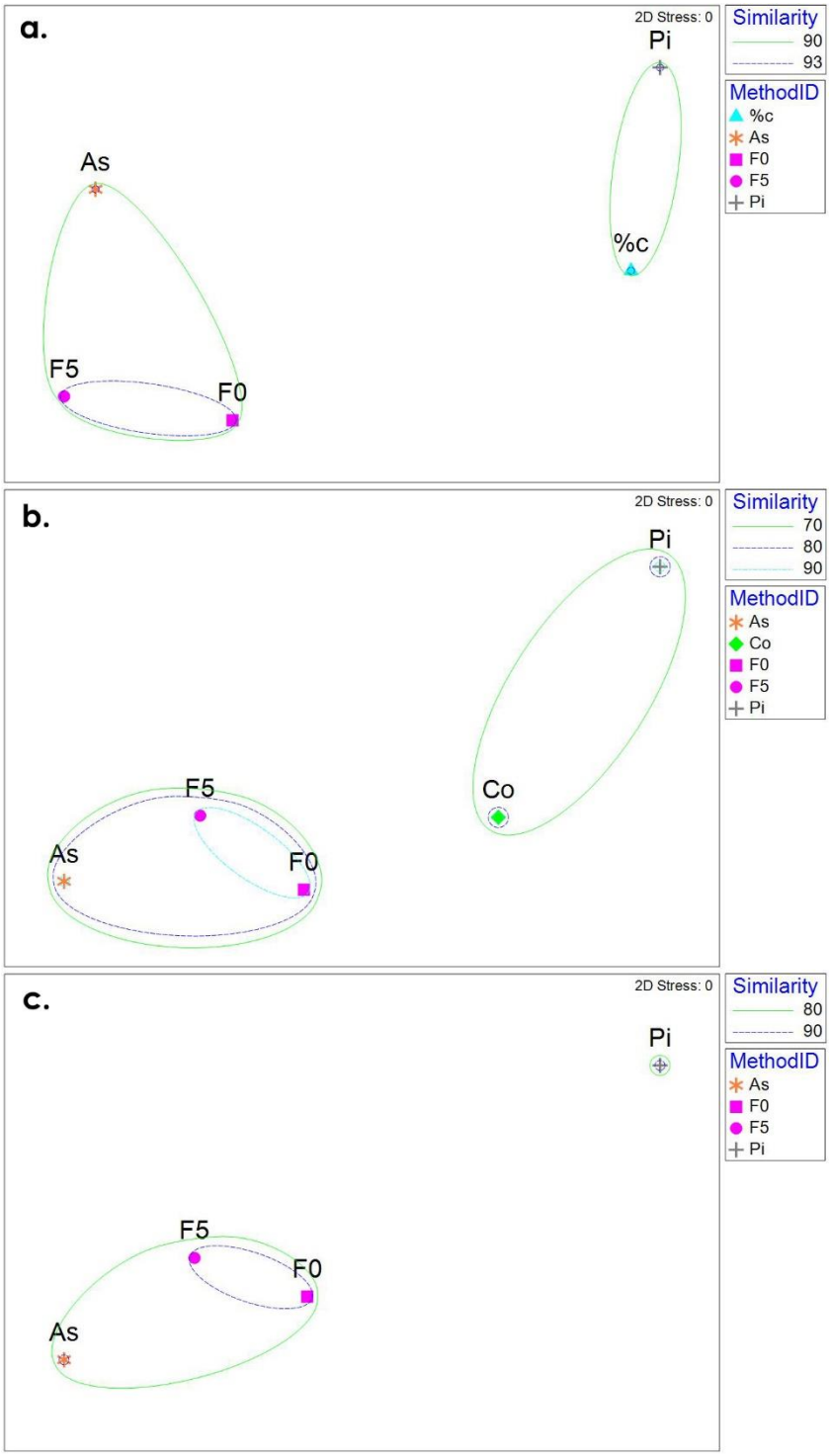


Figure 10. Non-parametric multi-dimensional scaling (nMDS) plots of sample data extracted by analyst A using six data extraction methods (MethodID) for three data sets: (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, Frequency of Occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover data). All data were averaged by extraction method and analyst (excluding outliers), and fourth-root transformed before calculating a Bray-Curtis similarity matrix. Overlaid similarity groupings from Group Average CLUSTER analysis. Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

3.9 Consistency

Inconsistency of recording is a notable issue for many monitoring programmes, confounding our ability to reliably detect changes or accurately describe natural fluctuations. The considerable inconsistencies present in the data from this image analysis exercise are obvious, but this section compares the data extracted by the six methods to assess whether any methods were more consistent (less inconsistent) than others.

3.9.1 Consistency of species richness

There is considerable variability in the number of taxa recorded between Senior and Junior Analysists within each method, as well as between the methods themselves (Table 13; variability of individual images between the Analysts, across the methods can be seen in Appendix 3). The variance seems to be related to the number of species recorded (Figure 5, Table 9, Table 13), so it is lower for those methods that inherently captured fewer species (Percentage Cover, Abundance Counts and Point Intercept).

Table 13. Summed variance of the number of taxa recorded from 20 images, by extraction method and experience.

Experience	Percentage Cover	SACFOR	Abundance Count	10x10 Frequency	5x5 Frequency	Point Intercept
All	65	204	75	220	228	82
Senior	77	278	101	290	313	114
Junior	41	111	43	106	97	31

Figure 11 provides one example to illustrate the effect of analyst experience. While the senior analysts usually recorded more taxa than the junior analysts, their variability was also usually greater. This was evident for all methods (see Appendix 3 for results of all methods).

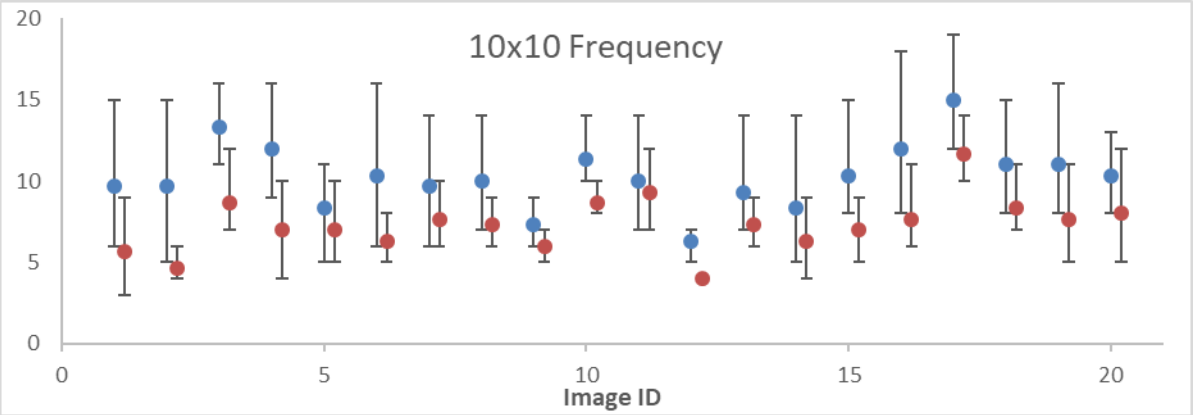


Figure 11. Variability (mean and range) in species richness, by analyst experience, for 20 images using the 10x10 Frequency extraction method. ● Senior analysts, ● Junior analysts. The latter have been offset slightly to the right for readability.

3.9.2 Consistency of selected species records

Variability of abundance records between analysts was calculated for selected taxa and then compared between extraction methods. All enumeration data were first standardised (z = (x – mean of x) / standard deviation, where x is the recorded value and z is the standardised value) to compensate for the different scales of the extraction methods. Then, for each selected taxon, the variance of abundance records was calculated for each image and summed across all images.

Selection of taxa for this analysis was complicated by the extreme inconsistency of identification, even after extensive truncation. Identification of encrusting and turf fauna were particularly inconsistent, with sponges, bryozoa and hydrozoa combined or described in a large variety of ways that could not be reliably reassigned to a standardised list of taxa. The taxa in Table 14 are ones that were frequently recorded by all or most of the analysts with relatively little identity confusion.

Comparing the extraction methods, there are few clear patterns, but some methods appear to be better than others for certain taxa. Thus, the Point Intercept method provided relatively consistent values for the conspicuous echinoderms, while the 10x10 Frequency method was most consistent for crusts and turfs. Overall, the results from the SACFOR method were relatively inconsistent.

Analyst experience also had mixed results; with senior analysts apparently being more consistent for Flustridae and other faunal turf but junior analysts were more consistent in their identification of Scleractinia. It is likely that most of those apparent effects are due to variable identifications between the different analysts.

Table 14. Summed variance of standardised abundances of selected taxa, from 20 images, by extraction method and experience.

	Experience	Percentage Cover	SACFOR	Abundance Count	10x10 Frequency	5x5 Frequency	Point Intercept
Scleractinia	**All**	n/a	5	4	4	3	nd
	Senior	n/a	6	7	7	4	nd
	Junior	n/a	5	2	2	3	nd
Echinoidea	**All**	n/a	2.0	2.7	1.9	2.2	0.9
	Senior	n/a	0.6	0.7	0.5	1.9	0.0
	Junior	n/a	2.4	3.7	3.1	1.6	1.2
Ophiuroidea	**All**	n/a	12	5	3	7	3
	Senior	n/a	12	6	4	8	2
	Junior	n/a	15	3	2	7	6
Ophiocomina nigra	**Senior**	n/a	17	5	9	16	7
Ophiothrix fragilis	**Senior**	n/a	18	20	13	12	14
U. faunal crust	**All**	20	20	n/a	16	19	17
	Senior	22	21	n/a	19	20	18
	Junior	24	21	n/a	10	19	22
U. faunal turf	**All**	18	16	n/a	10	13	13
	Senior	14	12	n/a	7	9	5
	Junior	15	16	n/a	11	18	16
Flustridae	**All**	2.8	3.0	n/a	0.4	2.9	1.9
	Senior	0.3	2.0	n/a	0.1	0.6	0.1
	Junior	2.6	2.9	n/a	0.6	2.4	4.4

3.9.3 Multivariate analysis of consistency – comparisons of entire community between methods, experience level and individual analysts

Dividing the data into three sets allowed testing of ground cover taxa only (without Abundance Count data), erect/solitary taxa only (without Percentage Cover data) and of all the taxa identified in the imagery, but only by the methods that could record all the taxa (SACFOR, Frequency 10x10 and 5x5 grids and Point Intercept methods, without any Abundance Count and Percentage Cover method data). This made the comparisons more appropriate to the taxa identified by each method.

PERMANOVA test results show that differences exist between the data acquired by different methods, between analysts and the level of experience of those analysts (Table 15).

Importantly, these results suggest there is little consistency of community impression between the observers seen when using the same method on the same set of 20 images. Further comparisons of the data are made using the components of variation from each set of tests, which as mentioned before, indicate the amount of variability of the data within each factor and between replicate sample images. Each result may be interpreted as the amount of variability (on a relative scale) that is accounted for within that factor and, thus, may serve as a proxy for multivariate variability in general (Anderson *et al.* 2008).

Table 15. PERMANOVA test results for a three-factor PERMANOVA model, using the fixed factors 'Method', 'Experience' and 'Analyst' (nested within Experience) with 9999 unrestricted permutations. Tests were conducted on fourth root-transformed data of (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, frequency of occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover method data). Bold results indicate a significant effect (P ≤0.05). The components of variation (Comp. var.) indicate the multivariate variability of data between the fixed factors, as indicated, and between replicate sample images (Residual). perms = permutations.

Data set	Factors/Source	df	SS	MS	Pseudo -F	P	Unique perms	Comp. var.
a. **Ground cover taxa**	Method (Me)	4	108790.0	27197.0	34.55	**0.0001**	9903	14.8
	Experience (Ex)	1	109250.0	109250.0	138.78	**0.0001**	9947	19.0
	Analyst (An/Ex)	4	297890.0	74471.0	94.60	**0.0001**	9908	27.1
	Me x Ex	4	19006.0	4751.5	6.04	**0.0001**	9900	8.1
	Me x An/Ex	16	38350.0	2396.9	3.04	**0.0001**	9812	9.0
	Residual	570	448720.0	787.2				28.1
	Total	599	1022000.0					
b. **Erect/ solitary taxa**	Method (Me)	4	145650.0	36412.0	17.11	**0.0001**	9890	17.2
	Experience (Ex)	1	154550.0	154550.0	72.63	**0.0001**	9930	22.9
	Analyst (An/Ex)	4	171040.0	42761.0	20.10	**0.0001**	9885	20.5
	Me x Ex	4	24803.0	6200.7	2.91	**0.0001**	9887	8.4
	Me x An/Ex	16	25990.0	1624.4	0.76	0.9835	9760	-5.1
	Residual	552	1174600.0	2127.8				46.1
	Total	581	1696800.0					
c. All taxa (but no count or cover data)	Method (Me)	3	190600.0	63533.0	60.17	**0.0001**	9911	22.8
	Experience (Ex)	1	103310.0	103310.0	97.85	**0.0001**	9938	20.6
	Analyst (An/Ex)	4	212200.0	53051.0	50.24	**0.0001**	9878	25.5
	Me x Ex	3	8851.1	2950.4	2.79	**0.0001**	9907	5.6
	Me x An/Ex	12	26725.0	2227.1	2.11	**0.0001**	9791	7.7
	Residual	456	481480.0	1055.9				32.5
	Total	479	1023200.0					

The components of variation show that for each data set the majority of variability is accounted for among the replicates (see residual values in Table 15). This effect is most pronounced among the replicates of erect/solitary taxa (Residual = 46.1), indicating that this data set has the highest amount of residual variability, i.e. least consistency. This effect is least pronounced among the ground cover taxa (Residual = 28.1). The components of variation also provide a useful means to understand the relative variability within each experimental factor in each data set. Within the ground cover taxa data set, the majority of variability within the factors is among the analysts (Analyst = 27.1). This result is clearly visible in Figure 12a, in which differences in the ground cover taxa data set seem largely influenced by the analyst more than any other factor, including method. As there is also no clear pattern in relation to analyst experience, these results indicate that estimates of ground

cover taxa are more influenced by the analyst than their experience level or the method they use to extract the data from the imagery.

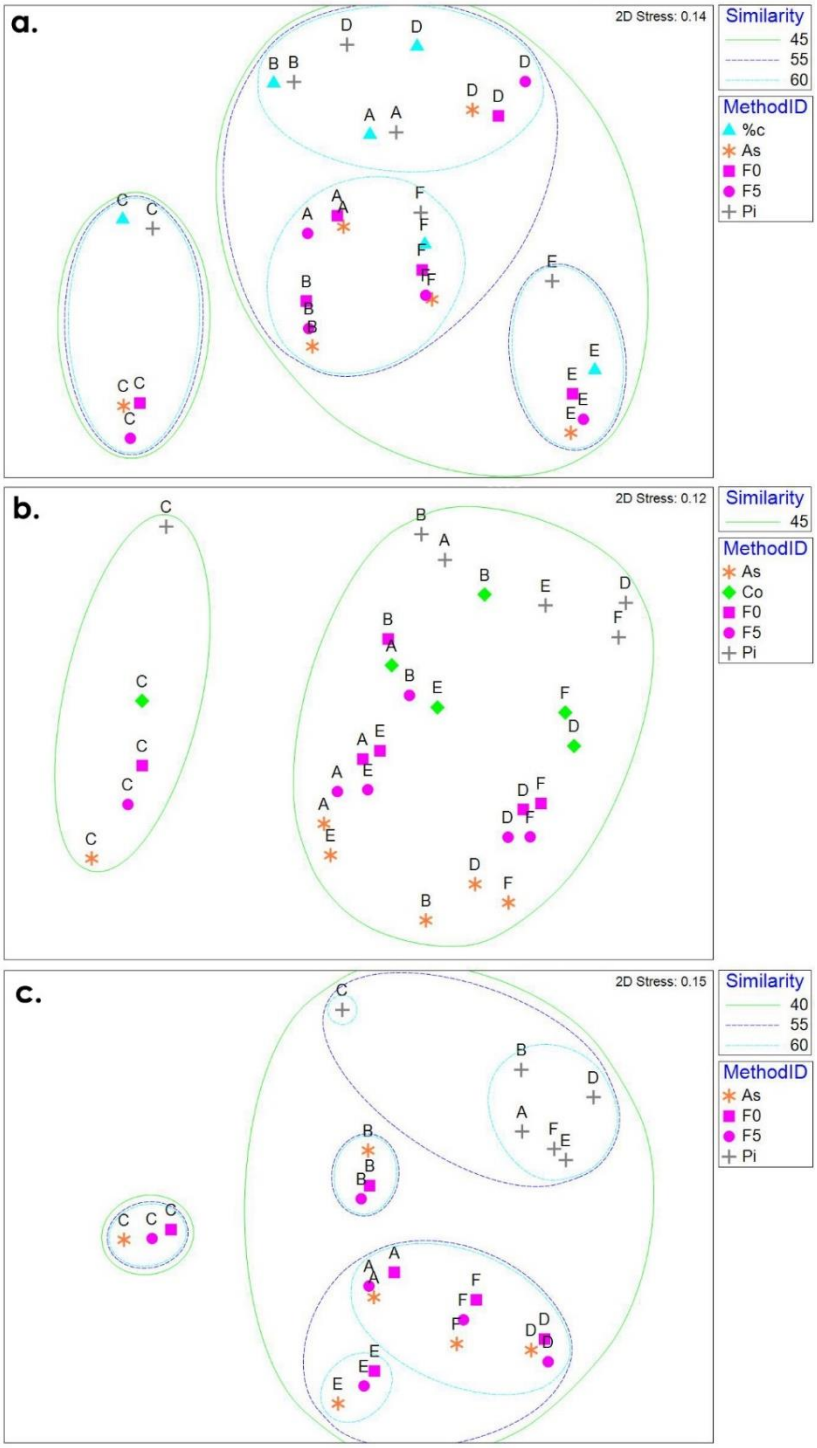


Figure 12. Non-parametric multi-dimensional scaling (nMDS) plots of sample data extracted by six analysts (A-F as shown) using six data extraction methods (MethodID) for three data sets: (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, Frequency of Occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover data). All data were averaged by extraction method and analyst (excluding outliers), and fourth-root transformed before calculating a Bray-Curtis similarity matrix. Overlaid similarity groupings from Group Average CLUSTER analysis. Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

The components of variation for the erect/solitary taxa data set show, unlike the ground cover data set, the experience level of observers accounts for slightly more variability (Experience = 22.9) than that between analysts (Analyst = 20.5) or between methods (Method = 17.2). The nMDS plot shows that data extracted by analyst C (a senior analyst) seems a little different to all the others, while amongst the others (analysts A, B, D, E and F) data differences seem to be more influenced by the data extraction method than in the ground cover data set (Figure 12b).

The results from the final data set is the most interesting of the three, in which the components of variation for the methods are the highest of all three data sets (Method = 22.8). The nMDS plot of this data set show that all community impressions made by the Point Intercept method are clustering together, even for Analyst C (Figure 12c). These results would indicate a degree of consistency is being achieved by the analysts using the Point Intercept method when the entire taxa list is considered. This is to be expected for a method that focuses the analysts' attention on particular points and represents a relatively small proportion of the images and of the diversity captured by the other methods. Looking at data from the other methods, however, all impressions cluster around the analyst again, similar to the other data sets but more clearly. In every instance, the impressions of the community made by each analyst seem similar regardless of whether they used a Frequency of Occurrence method or SACFOR (Figure 12c).

Multiple post-hoc pairwise PERMANOVA tests investigated the factorial interactions in the main PERMANOVA tests (in Table 15; (key results presented in Appendix 3). They also enabled further exploration of some of the similarities and patterns seen in the data sets so far. By tabulating the average similarities of data annotated by each analyst using the methods (generated by pairwise testing of factors in PRIMER), a multivariate means of consistency may be estimated for the data sets. Using this approach, the highest levels of consistency between all analysts in all three data sets occurs using the Point Intercept method (Table 16). This result applies to both groups of senior and junior analysts alike. It is likely these results are related to the fact that the Point Intercept method records fewer taxa than the other methods (when all taxa are concerned; See section 3.6) and, thus, has less capacity to vary. With lower variability comes the possibility of being able to achieve greater levels of consistency when utilised between different analysts.

The methods with the next highest level of consistency among all three data sets are the Frequency of Occurrence methods, with levels slightly higher overall for the data made using the 10x10 grid (Table 16). Unlike the Point Intercept method, these methods do sample the whole still image and so have the potential to generate more variable data. They represent, therefore, the most consistent methods of those used in this study when considering the whole community within the entire area of an image. Conversely, the least consistent method that considered the whole community in the entire area of a still image was the SACFOR method (Table 16). This observation applies to all data sets and to both groups of senior and junior analysts alike. It is not understood whether this result is related to the limited scale range of the SACFOR method, varying only from 0-6, or due to inconsistencies between the records made by the analysts. Either or, this result indicates this method should be used with extreme caution in benthic monitoring studies using stills imagery, in which the consistency of data recorded by different analysts over time is a key concern.

Table 16. Multivariate consistency of recording (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, frequency of occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover method data) using six different data extraction methods. Average similarities among data within each method taken as a proxy for consistency, calculated by pairwise PERMANOVA tests for the factor 'Method' in a three-factor PERMANOVA model, using the fixed factors 'Method', 'Experience' and 'Analyst' (nested within Experience) with 9999 unrestricted permutations. Tests were conducted on Bray-Curtis similarity matrices using fourth root-transformed data. Results are grouped into 'All analysts' (A-F), 'Senior' analysts (A-C) and 'Junior' analysts (D-F). Bold values indicate methods with highest level of similarity in data set (considering whole numbers).

Data set	Method	Average similarity between groups		
		All analysts	Senior	Junior
a. **Ground cover taxa**	%c	46.42	48.01	52.29
	As	45.17	47.77	51.19
	F0	48.21	**50.67**	55.11
	F5	45.62	**50.79**	51.00
	Pi	**51.45**	**50.73**	**56.84**
b. **Erect/ solitary taxa**	As	27.52	24.90	34.93
	Co	28.33	26.23	37.72
	F0	29.18	28.86	38.17
	F5	30.04	28.41	38.64
	Pi	**40.45**	**40.02**	**58.33**
c. **All taxa** (but no count or cover data)	As	36.45	36.12	43.57
	F0	40.99	42.29	49.00
	F5	39.50	41.99	46.41
	Pi	**55.59**	**54.22**	**62.83**

Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As); 10x10 Frequency (F0); 5x5 Frequency (F5); Point intercept (Pi).

Another interesting result in Table 16 is that the community impressions generated by the junior analysts appear to be more consistent than that generated by the senior analysts for every data extraction method and in every data set. Perhaps this relates to the lower numbers of taxa recorded by the junior analysts overall (see section 3.6). With fewer taxa records the potential for data variability is reduced and the potential for higher levels of consistency is increased. This reason is similar to that proposed to explain the high levels of consistency of the Point Intercept data.

One final result of interest is that overall, the consistency of data was higher in the ground cover data set, i.e. among taxa that tend to be sessile and grow colonially (Table 16). Overall, the lowest levels of consistency among the data sets was in the erect/solitary taxa (Table 16). It is not clear whether this is attributable to the difficulties in recording these taxa (i.e. analyst error) or the spatial distributions and behaviour of such taxa.

3.10 Rank-based comparison of extraction methods across data metrics

The rank-based method of comparison allowed comparison of the relative performance of all methods in this study against that of the combined results of the Percentage Cover and Abundance Count methods. The ranked results show both positive and negative differences between the data extraction methods across all data metrics in this study (Figure 13; full results in Appendix 3, Table 29). The overall difference in performance of the two frequency grid methods (Frequency 5x5 and Frequency 10x10) and the SACFOR method was higher

than the performance of the benchmark method (combined Percentage Cover and Abundance Count methods). This suggests these methods generally performed better than the Percentage Cover and Abundance Count methods across all the data metrics. The difference in performance of the Point Intercept method, however, is lower than the benchmark method, suggesting it is the poorest performing method in this study across the data metrics.

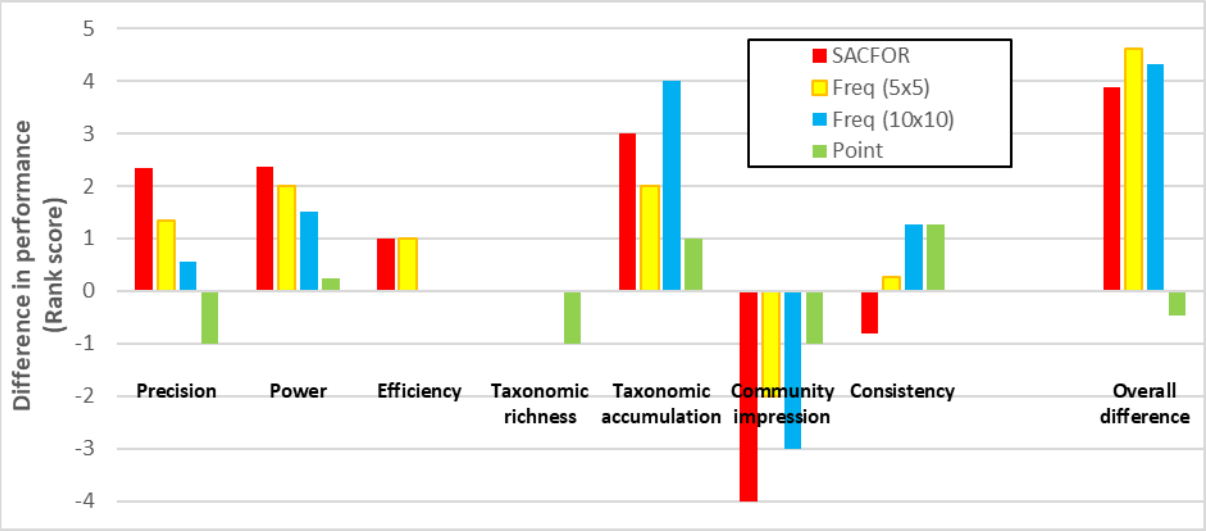


Figure 13. Relative performance of data extraction methods, as indicated by rank scores relative to combined scores of Abundance Count (erect/solitary taxa) and Percentage Cover (ground cover taxa) for all key data sets of every data metric in this study.

The Frequency 5x5 grid had the highest overall rank performance scores of the study, followed closely by the Frequency 10x10 grid method (Figure 13). These results suggest that across all the data metrics in this study these methods represent the optimum approaches for data extraction across all data metrics, from the set of images used. Closer examination of the rank scores in each data metric shows that the Frequency 5x5 method has higher levels of Precision, Power and Efficiency than the Frequency 10x10 method. In turn, the Frequency 10x10 method has quicker accumulation of taxa and higher levels of Consistency than the Frequency 5x5 method. Estimates of Community Impression (a proxy for accuracy) are poor for both methods but less poor for the Frequency 5x5 method.

When trying to understand the relative strengths and weaknesses of the frequency methods, the key difference between them relates to data resolution, with one method using up to 25 and the other up to 100 data points per taxon. In this instance, it seems as though an increase in data resolution (by using more frequency cells per image) increases the levels of Consistency between different analysts, as well as the speed at which taxa accumulate in the data. However, this is at the cost of decreasing Precision, Power, Efficiency and accuracy (inferred from community impression results) of the data. If frequency methods were to be used for future data extraction work from imagery, the data needs of the study must be considered carefully before deciding which level of data resolution to use.

Like the frequency methods, the SACFOR method also produced high rank scores across the data metrics, especially for Precision, Power, Efficiency and Taxonomic Accumulation (Figure 13). As discussed in earlier sections (sections 3.3 and 3.4), it is likely that the relatively high rank scores of the SACFOR and Frequency 5x5 grid methods may be due to, in part, the small data ranges over which they operate (1-6 and 0-25 respectively). These results support the notion that data extraction methods that operate over smaller data ranges will perform well in terms of Precision, Power and Efficiency. These data metrics are all important to the success of benthic monitoring programmes (Drummond & Connell 2005;

Leujak & Ormond 2007; van Rein *et al.* 2011). However, such methods will suffer in terms of data discrimination and resolution owing to their smaller data ranges.

Unlike the Frequency 5x5 grid data, which scored consistently well across all data metrics, the SACFOR method produced the lowest rank scores for Community Impression (accuracy) and Consistency of all data extraction methods trialled in this study (Figure 13). These results indicate that SACFOR data are inaccurate and vary highly between different analysts. Although the method performs well in many other metrics, poor scores in these data metrics must be noted and carefully considered when applying this method in benthic monitoring scenarios.

Finally, the poorest rank scores overall are seen in the Point Intercept data, which had the lowest rank scores for Precision, Power, Efficiency, Taxonomic Richness and Taxonomic Accumulation (Figure 13). When compared with all the other methods in this study, this method may be considered the least optimal method. It is likely that this method's poor data coverage per image, with data created from only 100 dots per image, has caused this result. Interestingly, it is likely that this more restricted means of sampling an image resulted in Point Intercept having the joint highest levels of consistency among the methods (tied with the Frequency 10x10 method). The method also had high levels of accuracy indicating that the impressions of community observed using Point Intercept were similar to those expected in real life.

4 Conclusions

As mentioned in the introduction, this report considers one specific issue: How to choose an appropriate method of enumerating taxa from still images of the benthos? The numerous analyses in this study have outlined a robust approach that considered seven data metrics useful for any monitoring programme: Precision, Power, Efficiency, Taxonomic Richness, Taxonomic Accumulation, Community Impression and Consistency of data between analysts. The relative 'performance' of each data extraction method was ranked within those data metrics and summarised to select the optimum method overall, which in this study was the Frequency of Occurrence 5x5 method. However, issues with taxonomic classification, different data ranges and operational approaches of the methods confound this conclusion.

Further understanding of the exact monitoring objectives and priorities are needed before method recommendations should be made. Questions such as 'what taxa are to be monitored?' and 'which data metrics are the most important for the taxa to be monitored?' and 'what level of data resolution do I need for my data?' should be asked first. The results of this study suggest that differences in answers to such questions would result in different method recommendations being made. Some of the methods performed well in certain data-sets within certain different metrics and yet not in others. Although the 'generalist', well-performing Frequency 5x5 grid is considered optimal across all metrics in this study, it may not be the most appropriate method for certain taxa, communities, objectives and associated metrics.

4.1 Advantages and disadvantages for each extraction method

4.1.1 Percentage Cover (%c) and Abundance Count (Co)

The Percentage Cover and Abundance Count methods are usually used in tandem to enumerate all taxa in benthic monitoring imagery. As one method records ground cover taxa and the other erect/solitary taxa they complement each other well and provide data that is intuitively understandable to all users. Although they were quick to carry out individually in this study, relative to the other methods, they each only recorded roughly half the taxa in an image. When they are used in tandem to enumerate all taxa in the image, as for most everyday applications of the methods, however, this time-saving advantage would be lost. Both of their data types also allow for aggregation of taxa, by adding enumeration values together through truncation procedures.

The data generated by these two methods was considered the most accurate in this study and so it was used as the measure to which the values of the other methods were compared to. Although they were accurate, these methods did not generate data that could be considered precise overall, nor with high power or very consistent between different analysts. It is likely that this is because both methods provide the greatest discrimination of any method, but when they provide inconsistent data the greater discrimination will be misleading. These results should be considered carefully when choosing these methods for use in marine benthic monitoring programmes, in which data consistency, precision and power are key for detecting changes in benthic communities.

4.1.2 SACFOR (As)

The Marine Nature Conservation Review abundance scales (i.e. SACFOR method) were designed for rapid mapping and assessment surveys of larger areas of seabed than are shown in still images. Even though there is a mismatch of sample size and scale, the SACFOR method is routinely used for still image analysis in marine benthic monitoring programmes. This is at least partly due to the rapidity at which it can be applied, as clearly

demonstrated by the results of this study by being one of the most efficient methods under investigation.

The SACFOR method owes its efficiency to the broad categorisation of abundances in each scale, which generate semi-quantitative ordinal data. This type of data made it difficult to compare with the other methods in this study. It also makes it difficult to use in marine benthic monitoring programmes. In this study, as in many others, the SACFOR scale values were converted to a numerical scale (1-6) to enable statistical analysis of the data. However, considering the logarithmic nature of the SACFOR scale and that there are different scales for different growth forms and sizes of individuals, perhaps this was an inappropriate approach.

It is interesting that the SACFOR method generated data, post-conversion, that were the most precise of all the methods in this study. However, this may have been an indirect result of converting the semi-quantitative ordinal data to just a 1-6 scale, whereas most other methods used 0-100 scale that has more potential to show variability in the data. If having a smaller range of data led to the SACFOR data being considered 'precise', it is contradictory that the same data were also among the least consistent of all the methods under investigation. Coupled with this inconsistency, the data are not-additive, so aggregation of taxa is not straightforward for truncation procedures. Although use of this method in marine benthic monitoring programmes is widespread, the key issues of sample size, spatial scale of operation, inconsistent interpretation of abundance scales and difficulties of use in a statistical analysis should be carefully considered every time it is employed.

4.1.3 Frequency of occurrence: 10x10 (F0) & 5x5 (F5)

Both Frequency of Occurrence methods, 10x10 and 5x5 grids, sampled the entire sample image and all discernible taxa, much like the SACFOR method. As such the Taxonomic Richness values were among the highest of the study for both methods. Providing abundances for all taxa (i.e. both ground covering and solitary) on the same scale is particularly useful for multivariate analyses. Having the two grids, 10x10 and 5x5 cells, enabled an assessment of the effect that increasing the resolution of the data would have on the performance of the data extraction approach. In this study, it seemed to reduce precision, power, efficiency and relative accuracy of the data while increasing the accumulation of taxa and consistency of data between analysts. While the 10x10 Frequency method took significantly more time than any method (except Point Intercept), the 5x5 Frequency was one of the most efficient of the study.

One characteristic of these frequency methods, particularly 5x5 Frequency, is that it tends to down-weight the most abundant ground covering taxa and up-weight the small bodied widely scattered taxa. This is similar to the pre-treatment transformations typically recommended for multivariate analyses of community data in PRIMER. Frequency of Occurrence data may be considered partially pre-transformed, but with less discrimination compared to Abundance Counts. It is likely that this inherent 'transformation' effect is why both Frequency-based methods generated data with higher precision, power and consistency overall, relative to the other methods. These results indicate that the Frequency methods could be very reliable for use in marine benthic monitoring programmes, in which methods that record data with low variability and high consistency between analysts are favourable. Compared with SACFOR the Frequency data had similar or higher precision, power and consistency, while also providing greater discrimination.

Before the Frequency methods are used for marine benthic monitoring programmes, however, careful consideration needs to be made of how to aggregate taxa for truncation procedures. Currently, this is not easily possible with this data type without reanalysing

sample images, similar to the SACFOR method. It is possible that increased use of image annotation software may tackle this issue to some extent.

4.1.4 Point Intercept (Pi)

The Point Intercept method was the least efficient of all the methods in this study. This is in contradiction to the findings of other method comparison studies (van Rein *et al.* 2012), probably due to differences in the method protocols between the two studies. This method also captured fewer taxa than any of the other methods that could record all taxa (i.e. fewer than SACFOR or either of the Frequency of Occurrence methods). This is clearly a result of the method only sampling a very limited area of the sample image. It is likely that this is also why the Point Intercept method did not record data as precisely and with as much power as the Frequency of Occurrence methods. Even so, the Point Intercept method recorded data that seemed to represent the actual abundances (physical amount) of those taxa, unlike the Frequency of Occurrence methods. It also recorded data with high levels of consistency between Analysts. This was particularly the case for ground cover taxa and large bodied mobile animals. The Point Intercept data seemed relatively unreliable, however, for small bodied taxa, unless they are present in more than moderate abundance.

The high levels of consistency and ease of use for Point Intercepts have led to these methods being widely used in marine benthic monitoring programmes. They seem best suited to monitoring ground cover taxa where they may incur fewer errors (e.g. missing records) than other methods. The inherently systematic protocol allows high levels of quality control to be implemented, i.e. each point may be easily verified by other analysts. Important for truncation procedures, Point Intercept data may also be easily aggregated if needs be.

4.1.5 Non-additive abundances – a warning!

As described above, three of the extraction methods (SACFOR, Frequency 10x10 and 5x5) are non-additive. Aggregation of taxa is therefore likely to introduce errors, particularly for SACFOR scores and 5x5 Frequencies, if the scores or frequencies are estimated or counted manually (without the aid of software). Aggregation of abundance data to higher taxa may be required prior to analysis of data that were collected using different protocols for species identification or when samples have been analysed by multiple analysts with differing identification skills. This could create a major data quality issue over the duration of a monitoring time series. Every additional survey can result in more errors, unless a standardised list of taxa is applied, with detailed identification protocols and analysts receive the necessary training. Direct comparison of data from a recent survey with one from the beginning of the time series is likely to be much more difficult if the data are non-additive. Many monitoring programmes have had such problems with non-additive abundance data, often requiring data managers to spend large amounts of time manually re-configuring data sets to minimise potential errors in analyses. The magnitude of the problem will depend on the type of analyses and the questions being asked of the data.

As indicated earlier, a potential solution to this issue can be explored for Frequency of Occurrence methods but not SACFOR methods. By using image annotation software, which marks the location on the image where each taxon is found, accurate recalculation of cell frequencies of taxa may be possible if aggregation is required.

4.2 Recommendations for future studies

There are many interesting results in this study that enable the following recommendations:

- It seems that no one data extraction method performs the best at everything: generating data that are precise, with more power, taxa rich and consistent between different image analysts. The target / focus of the monitoring study is vital in selecting the optimum approach to extract data from samples collected for it. Of the approaches explored in this study, the Frequency of Occurrence methods recorded data that were more precise, with more power and with higher levels of consistency than the majority of other methods in this study. They recorded numerical data in a single unit that could represent the whole community sampled and the whole sample area (unlike Percentage Cover, Abundance Counts and Point Intercept). Further exploration of these approaches is recommended but only with the following points considered first:
    - an approach to enable aggregation of Frequency of Occurrence data must be developed so that it may be truncated in future studies (suggested via the use of image annotation software);
    - efforts should be made to consider how to enable the aggregation of Frequency of Occurrence data from multiple images into larger image, representing an appropriate sample size. Perhaps this could be achieved by standardising Frequency of Occurrence cell sizes so that they may be used in images of all sizes. For example, grids made up of cells measuring 10 x 10 cm could be arranged in a grid across a still image using image annotation software.

- The methods applied in this study were all carried out manually, without the aid of software. Image annotation software is rapidly developing and appears to offer many benefits, some of which will affect the application of the methods studied. A similar study design, to compare these methods, but using image annotation software (i.e. extending the study described by Durden *et al.* 2016), should be considered.

- The sample sizes in this study were determined by the sizes of the images collected during the JNCC/MSS *Scotia* 1714S survey, which typically ranged from sizes of 0.3 to 1m^2 per image. There is increasing interest in aggregating data from such images after data extraction to determine the appropriate sampling unit size for the taxa under investigation. This may result in artificial images of sizes of 3-5m^2 in area. This type of analysis will require enumeration methods to be robust and flexible, so that they their data may be aggregated to allow such analyses. It is recommended that any future enumeration method comparison studies of this type carry out an appropriate sampling unit analysis be applied to reflect future monitoring needs.

- Extremely high levels of inconsistency were present in the dataset used by this study, due in large part to the lack of a standard checklist of taxa and qualifiers. This was a realistic scenario for many surveys, particularly one-off surveys in a previously un-surveyed areas, but unrealistic (one would hope!) for repeat monitoring surveys. Comparison of the extraction methods was still achievable, but the comparisons of consistency and precision between the methods would benefit from further study using data extracted using a standard checklist of taxa (representing a range of forms) known from the location under investigation. This additional study, again with multiple images, analysts and methods, would allow a better comparison of consistency and precision of the enumeration values. If the images were chosen to represent two or three similar but slightly different habitats (stratified random sampling) it might be possible to test the ability of each method to distinguish the habitats;

- It is recommended that, where possible, multi-metric method comparison studies should be carried out to estimate optimum approaches for different monitoring purposes (e.g. taxa, habitats). Such studies may incur additional expense at first but should enable the selection of the optimum data extraction (annotation) methods for the future purposes of that monitoring work. Using methods that represent the optimum balance of data metrics, such as precision, power, efficiency, accuracy and consistency, will improve the chances of detecting meaningful change in well-structured monitoring datasets over time.

5 References

Anderson, M.J., Gorley, R.N. & Clarke, K.R. 2008. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. PRIMER-E: Plymouth.

Barry, J. & Maxwell, D. 2017. emon: tools for environmental and ecological survey design and analysis. https://CRAN.R-project.org/package=emon

Beaumont, J.C., Brown, C.J. & Sayer, M.D.J. 2007. Evaluation of techniques used in the assessment of subtidal epibiotic assemblage structure. Biofouling 23 (5): 343-356.

BS EN 16260. 2012. Water quality. Visual seabed surveys using remotely operated and/or towed observation gear for collection of environmental data.

Coggan, R., Mitchell, A., White, J. & Golding, N. 2007. Recommended operating guidelines (ROG) for underwater video and photographic imaging techniques. MESH Project.

Drummond, S.P. & Connell, S.D. 2005. Quantifying percentage cover of subtidal organisms on rocky coasts: a comparison of the costs and benefits of standard methods. Marine and Freshwater Research 56: 865-876.

Durden, J.M., Bett., B.J., Schoening, T., Morris, K.J., Nattkemper, T.W. & Ruhl., H.A. 2016. Comparison of image annotation data generated by multiple investigators for benthic ecology. Marine Ecology Progress Series 552: 61-70.

Goudge, H., Morris-Webb, E., Stamp, T., Perry, F., Deamer-John, A. & O'Connor, J. 2016. *Analysis of seabed video and stills data collected by drop down camera on the Solan Bank Reef SCI (1714S) (2014)*. JNCC Report No. 582. JNCC, Peterborough. [http://jncc.defra.gov.uk/pdf/report_582_final_web.pdf].

Griffiths, Y. (2018, in draft) Guidance for conducting a power analysis of benthic community data using 'emon' package in 'R'. JNCC internal document.

Hinchen, H., 2014. Review of marine biodiversity assessment obligations in the UK. Part I: A summary of the marine biodiversity assessment obligations stipulated within national and international legislative and policy instruments, JNCC Report No. 497, JNCC, Peterborough. http://jncc.defra.gov.uk/page-6673 (Accessed 24 July 2017).

JNCC. 2012. Solan Bank Reef SAC: Inshore and Offshore Special Area of Conservation: Solan Bank Reef: SAC Selection Assessment Document. A report for JNCC and SNH, Peterborough. URL: http://jncc.defra.gov.uk/pdf/SolanBank_SAC_SAD_v_5_0.pdf (Date last visited 20/05/2019).

Leujak, W. & Ormond, R.F.G. 2007. Comparative accuracy and efficiency of six coral community survey methods. Journal of Experimental Marine Biology and Ecology 351: 168–187.

Moore, J.J., Bunker, F., van Rein, H. & Jones, J. 2015. Methodological trials: Recording subtidal epibiota in situ and in photographs, Portrush, August 2013 and Sound of Mull, August 2014. JNCC Report No. 561, JNCC, Peterborough.

Turner, J.A., Hitchin, R., Verling, E. & van Rein, H. 2016. Epibiota remote monitoring from digital imagery: Interpretation guidelines.

Van Rein, H., Schoeman, D.S., Brown, C. J., Quinn, R. & Breen, J. 2011. Development of benthic monitoring methods using photoquadrats and SCUBA on heterogeneous hard substrata: A boulder-slope community case study. Aquatic Conservation 21 (7): 676-689.

Van Rein, H.B., Schoeman, D., Brown, C.B., Quinn, R. & Breen, J. 2012. Development of low-cost image mosaics of hard-bottom sessile communities using SCUBA: comparisons of optical media and proxy measures of community structure. Journal of the Marine Biological Association of the United Kingdom 92 (1): 49-62

Appendix 1: Image analysis protocol notes

The following notes by five of the analysts explain some of the slightly different ways that they applied the six extraction methods

Percentage Cover

Analyst A	I used the 10 x 10 grid overlay image and made a list of all the taxa I could see in the taxa spreadsheet column, with any notes accompanying, initially at full screen, then zoomed in to approx. 1/3rd – ½ of the screen (added to the list if I noticed other taxa later during analysis of image). Filled in N/A for the measurement column for all erect or solitary epifauna taxa. For the cover estimates: For taxa where percentage cover was high (e.g. U. faunal crust) I estimated the number of grid cells that were not covered and deducted from 100. If cover was less but still had a wide spread throughout the image, I made an estimation from looking at the whole image (by trying to estimate the number of rows and cells the cover of that taxon would squeeze into/fill up if all together), and where appropriate backed this up with the method of only counting cells if more than 50% of that cell was covered, and this tended to corroborate well. For cover of less than 20%, but where the coverage was clumped, I counted the cells where the taxa were present. Where coverage was small and spread throughout the image, I zoomed in and moved through the cells adding up by 1% at a time, as and when I felt one full grid cell would have been filled by the cover of that taxon.
Analyst E	I used the 10x10 grid in order to view and list all taxa. I did this square by square and zoomed in for certainty in the identification. N/A was filled in for erect/solitary epifauna taxa. Percentage cover was estimated by estimating the percentage cover of taxa in each square and then overall, same as Analyst A's approach.
Analyst F	Viewed and listed taxa square by square, as Analyst E did. Filled in N/A for erect/solitary epifauna. Percentage cover estimated same as Alison's approach.
Analyst D	Same approach
Analyst B	I've spent my life looking at strung quadrats, so I think my approach was pretty much identical to Analyst A's. Though I obviously have identified the taxa whilst zoomed in, I generally then assessed the % cover in full view – mentally "stuffing cells" to obtain a total, particularly for the less abundant taxa. For the abundant taxa, then often it is a Giz thing – knowing what I think 80%, 60% *etc*. *etc*. look like.

Abundance Count

Analyst A	I used the 5x5 grid overlay for this, and made a list of all the taxa I could see in the taxa spreadsheet column, with any notes accompanying, initially at full screen, then zoomed in to approx. 1/3rd – ½ of the screen (added to the list if I noticed other taxa later during analysis of image). Filled in N/A for the measurement column for all colonial or encrusting taxa. I then went through each taxa one by one, counting all the individuals in 2 of the 5x5 rows of cells at a time (zoomed in) moving from one side of image to another, and finally the remaining row at the bottom. For the brittle stars, I counted all the brittle stars I could see on the first count, then filled that number in as the measurement with 'Ophiuroidea' in the taxa column, and 'uncertain' in the notes column. I then counted all the brittle stars I thought I could potentially identify as *Ophiocomina nigra* and filled in that number for the measurement with 'Ophiuroidea' in the taxa column, and 'Ophiocomina nigra' in the notes column. I then did the same for the potential *Ophiothrix fragilis*. All other individuals I considered 'uncertain' – so I subtracted the number of potential *Ophiocomina nigra* and *Ophiothrix fragilis* from the initial total number. In this way, if the Ophiuroidea had to be pooled up to just 'Ophiuroidea' and the abundance count added for the 3 potential types, it would still represent the total number of brittle stars I had counted in the image, and there wouldn't be duplication in counting the same individual twice by mistake.
Analyst E	Using 5x5 grid, looked through each individual square (zooming when required) and noted all taxa. N/A was used for colonial/encrusting taxa. I then repeated this technique in order to count taxa individuals. For brittle stars, a tally was made for Ophiuroidea (when further classification is not possible) and for species that were clearly identifiable, this was then noted in the count. It felt tricky to count all brittle stars accurately, especially when large numbers and overlapping was present.
Analyst F	Used 5x5 grid to identify and note all taxa. N/A used for colonial/encrusting taxa. Went through each taxa and counted them in 2 squares at a time. For high numbers of Ophiuroidea, I swapped to the 10x10 grid and counted 2 squares at a time.
Analyst D	Same approach - although more conservative and mostly classified as 'Ophiuroidea'. Agree with Analyst E's point about counting very large, overlapping brittle star numbers.
Analyst B	I've dived over countless brittlestar beds as well as id'd hundreds of them in the lab. So I was pretty confident about the differences. So I just counted all taxa I recognised. Where I've recorded ophiuroidea separately I realise now I was referring to the genus Ophiura (usually seen half buried in the silt…..) ooops! Sorry. For most taxa I went through each cell, counting all the individuals in 2 of the 5x5 rows of cells at a time (zoomed in) moving from one side of image to another, and finally the remaining row at the bottom.

SACFOR

Analyst A	I used the 10 x 10 grid overlay image and made a list of all the taxa I could see in the taxa spreadsheet column, with any notes accompanying, initially at full screen, then zoomed in to approx. 1/3rd – ½ of the screen (added to the list if I noticed other taxa later during analysis of image). I then went through the same process for estimating the percentage cover of encrusting/colonial taxa (U. faunal crust, sponges, serpulid worms) as the percent cover data extraction technique, but with less precision when it was obvious that the percentage cover would definitely fit into one or other of the coverage categories (e.g. >80%, 40-79%, 20-39%, 10-19%, 5-9%, 1-5%, <1%) and the SACFOR unit/letter in the first column of the SACFOR scale spreadsheet ('crust/meadows'). The same process applied for faunal turfs (e.g. hydrozoa, alcyonium, erect bryozoa) using the second column of the SACFOR scale ('massive/turf'). For erect/solitary species, I chose the SACFOR scale column corresponding to the size of the taxa I was observing (and the instructions on which taxa should be in which size category when it was named specifically in the rows below). If the image was $1m^2$ and there were 1-9 individuals in the image, I would choose the SACFOR category in the 5th row down of the SACFOR scale categories (i.e. <1cm = O, 1-3 cm = F, 3-15 cm = C, >15 cm = A. If there were 10-99 individuals of that taxa present, I would select the categories from the row above. If there were 100-999 individuals of that taxa present, I would select the categories from the row above that. If the image was $0.5m^2$, I would count the individuals, then multiply by 4 ($0.5m^2$ is ¼ of the area covered by $1m^2$) [8]and then follow the rules in the paragraph above with the multiplied-up numbers according the size class of the individual/taxa. If the image was $0.3m^2$, I would count the individuals, then multiply by 9 ($0.3m^2$ is approximately 1/9th of the area covered by $1m^2$). Where there was overlap between coverage categories e.g. 1-5% and 5-9%, I usually took 1-5% to mean less than 5% coverage, and the 5-9% category as over 5%. I estimated percentage cover for all sponges, though not specified.
Analyst E	Same method as for percentage cover, investigating each square one by one and noting species present. The rest is the same as Analyst A.
Analyst F	Same method as used in percentage cover for encrusting/colonial species. Then the same as Analyst A – I did not estimate percentage cover of sponges.
Analyst D	Same approach
Analyst B	Apart from my comment pretty similar to all above. Though I was frustrated by the abundance scales – Why are Pododesmus, Porella and Alcyonidium counted as if they are mobile /solitary….. saddle oysters on several photos should definitely have been assesses by % cover?

[8] TM: $0.5m^2$ in my head is half of $1m^2$ so I only multiplied by 2 NOT 4? Sorry …. Though I think it will only affect a handful of records. In future we must write FOV = 0.5x0.5m = $0.25m^2$.

Point Intercept

Analyst A	Using the 10x10 points overlay, I zoomed in the image until I could see 5 points across the width of the image on the screen, and filled in the taxa and notes column on the spreadsheet for each taxa that was under each point, moving through the point numbers sequentially. Where the point fell on either shadow or the weight/rope, I entered 'No identifiable taxa', and I also used this for when whatever that was under the point was just not clear enough to be identified as anything for certain. When the cross hairs of the point fell just on the edge of a brittle star arm or serpulid worm or any type of fauna, I would always try to enter what was next to the cross hairs of the point in the bottom left quadrant, for consistency.
Analyst E	I zoomed in so that at least 3 points of one line of points were visible, taxa were noted for each point and then the image was dragged across to reveal the next set out points. If a point fell on a shadow/weight/rope/unidentifiable individuals, "No identifiable taxa" was entered.
Analyst F	Same approach as Analyst E.
Analyst D	Zoomed in to approx. 3 points, recorded the taxa at each and scrolled across to record the next points. Only recorded taxa if the centre point of the cross hairs were on it. If this couldn't be distinguished, "No identifiable taxa" was entered.
Analyst B	Identical to Analyst A's but I tried to use the intersect itself and extrapolation – I can see that Analyst A's method is very sensible as you can then actually see the place you're recording from and not the cross hairs.

Frequency of occurrence (10x10 & 5x5)

Analyst A	I used the 5x5 or 10x10 grid overlay respectively, and made a list of all the taxa I could see in the taxa spreadsheet column, with any notes accompanying, initially at full screen, then zoomed in to approx. 1/3rd – ½ of the screen (added to the list if I noticed other taxa later during analysis of image). I then went through each taxa one by one, counting every cell with the presence of that taxa, using 2 of the 5x5 rows of cells at a time (zoomed in) moving from one side of the image to the other, and finally the remaining row at the bottom. For the 10x10 grid I counted all cells containing a taxa, using 4 of the 10x10 rows of cells at a time (zoomed in) moving from one side of the image to the other, and finally the two remaining rows at the bottom. For the brittle stars, I initially counted all the cells with the individuals I could potentially identify as *Ophiocomina nigra* or *Ophiothrix fragilis* (recorded in the notes column), separately. I then counted all cells containing the other individuals I considered 'uncertain' – and tried to take care not to count cells with only those I potentially identified as *Ophiocomina nigra* or *Ophiothrix fragilis* (recorded in the notes column). For the Ophiuroidea: uncertain counts, I tried to concentrate on those individuals where only arms could be seen, or those that could not be seen clearly, but this was a confusing process as you had to count cells where even just the tip of an arm was present in a cell, and therefore tracing all the arms back to the bodies did not feel very accurate, especially in the 10x10 counts.
Analyst E	Either the 5x5 or 10x10 grid was used, depending on the selected method. Again, I looked through each individual square close up and noted the taxa present (and zoomed in further when required). I then returned to the first square and counted the number of individuals of each taxa before moving on to the next square. For brittle stars, I made a tally with Ophiuroidea (when species level identification was not possible) and species of Ophiuroidea such as *Ophiocomina nigra*. I tallied all the cells that counted each species or classification.
Analyst F	Same approach as Analyst E. If taxa were very abundant for example brittle stars or faunal crust, I counted the cells that did not contain these and then took this away from 25 or 100 depending on the size of grid being used.
Analyst D	Same approach - again more conservative with 'Ophiuroidea' identification. Counted cells even if one the tip of a brittle star arm was present.
Analyst B	Very similar to all above, though I was probably pretty comfortable with the brittlestars. Same comment as before with my ophiuroidea …. I meant Ophiura!

General comments / admissions

Analyst A	When I changed my mind about certainty of taxa name, or the content of the text in the notes column, I would search for that image number in all data extraction techniques, find the taxa entry and make the necessary amendments, then enter a cross in the 'amended' column. Occasionally, when I was getting tired or towards the day end and needed to finish more analysis than there was time, I would copy over the taxa list from one data extraction technique to another, along with the 'notes' column, but not the counts. Where counts or percentage covers were around the borderline between SACFOR categories, I would occasionally check what I had put for the abundance or percent cover if that data extraction technique had already been completed, and vice versa. Occasionally I missed putting a start or end time in the Random Image Order spreadsheet. If this happened, I would try to fill in a sensible time, but obviously these were estimations. There was also instances where the time for analysis was affected by a phone call/other distraction, and this could not be avoided completely. [9]
Analyst E	When I changed my mind about certainty of taxa name, or the content of the text in the notes column, I would search for that image number in all data extraction techniques, find the taxa entry and make the necessary amendments, then enter a cross in the 'amended' column. At the start of the process I copied taxa lists from abundance counts/percentage cover/SACFOR for use for another data extraction technique, however after 1 or 2 days I stopped doing this. Where counts or percentage covers were around the borderline between SACFOR categories, I would occasionally check what I had put for the abundance or percent cover if that data extraction technique had already been completed, and vice versa. The abundance counts made for brittle stars can be difficult at times and it is possible that individuals were missed due to overlapping or poor visibility.
Analyst F	Same as Analysts A and E, if I changed my mind about a taxa name. When there was a large number of taxa present in an image, I would copy taxa lists from other extraction techniques. For SACFOR I would use data from percentage cover/abundance count technique, if I had already completed these. Otherwise I would put in the notes column of the SACFOR technique the percentage/abundance I had used to calculate SACFOR, then when I came to percentage cover/abundance I already had the data recorded.
Analyst D	Same approach regarding updates to taxa names or changing my mind (I would edit for all method types but would put a cross in the 'amended' column) - this was particularly the case when first processing the images as I became more familiar with taxa ID. As part of this, I would cross check the taxa listed for images when I identified a new taxon not previously identified. Agree with inaccurate brittle star counts where numbers were large (>50).
Analyst B	All techniques feel (an probably are) less accurate in all the outer cells (top row, bottom row and sides) where the photograph is often poorly lit and the focus is at its worst and therefore the taxa are often blurred and unrecognisable. Perhaps the photographs should be edited so that the grids *etc* are only placed over "in focus / well lit" sea-bed.

[9] TM: That seemed to happen a lot – glad it wasn't just me.

Appendix 2: Data truncation protocol and resulting assignments

See Section 2.7.2 for explanation

Truncation protocol

Tcode	Entity being assessed	Related entities in dataset	Change made
1	Wrong spelling		Spelling corrected
2	Different versions of essentially the same entity (qualifier)		Standardise entity (qualifier)
3	Uncertainty (and sometimes reasons) expressed in qualifier		Remove uncertainty (and reasons), but take account of it when considering entity assignment
4	Genus only assigned, but there is only one likely species in that genus		Re-assigned to that species
5	Record(s) assigned to a higher taxon	Multiple other records assigned to a different higher taxon that still covers all likely taxa and seems more appropriate	Standardise to one higher taxon
6	Record(s) assigned to a higher taxon with notes that suggest analyst thought it was a particular species or genus	Multiple other records assigned to that species or genus	Re-assigned to that species or genus
7	Record(s) assigned to a higher taxon with notes that suggest analyst thought it was a particular species or genus	No other records (or only 1 or 2 records) assigned to that species or genus	Entity remains at higher taxon
8	Qualifier contains unnecessary notes		Notes removed from qualifier
9	Genus only assigned, and while there is more than one species possible it isn't likely habitat and it is distinctly different from expected	No mention in the dataset of the other possible species and the most experienced analyst assigned all records to the expected Species	Re-assigned to Species
10	Species assigned, and there is another possible species	Majority of other records for that Genus are assigned to Genus only	Re-assigned to Genus
11	Species assigned, and there is another possible species	Multiple other records for that Genus are assigned to Species. No records for other possible species.	Entity remains at Species
12	Genus only assigned, there is more than one species possible and they are not necessarily easy to distinguish	[the most experienced analyst assigned all records to the Genus]	Entity remains at Genus

Tcode	Entity being assessed	Related entities in dataset	Change made
13	Genus only assigned, there is more than one species possible and they are not necessarily easy to distinguish	[the most experienced analyst assigned all records to the Species]	Re-assigned to Species
14	Record(s) assigned to a higher taxon with notes describing a variety of forms & colours		Used notes to define a number of likely standard forms/colours
15	Single record assigned to a taxon by one analyst.	Inspection of other records by that analyst for the same image, but different methods, show that he/she usually assigned it to a different taxon	Re-assigned to the most frequently assigned taxon

Truncation results

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
5	A	Fauna	sp A	Fauna (sp A)	2	A (Species a (r2c2))
5	A	Fauna	sp A	Fauna (sp A)	2	A
5	A	Fauna	sp A	Fauna (sp A)	13	Species A
5	A	Fauna	sp C	Fauna (sp C)	2	C
5	A	Fauna	sp C	Fauna (sp C)	2	C (Species C)
5	A	Fauna	sp C	Fauna (sp C)	5	Species B
8	A	U. faunal crust		U. faunal crust	1	U. faunal crust (Hydrozoa)
	A	U. faunal crust		U. faunal crust	1148	U. faunal crust
3	A	U. faunal crust		U. faunal crust	1	U. faunal crust (? Densest bit bottom right out of focus)
3	A	U. faunal crust		U. faunal crust	2	U. faunal crust (?? Poor focus)
3	A	U. faunal crust		U. faunal crust	3	U. faunal crust (???)
8	A	U. faunal crust		U. faunal crust	5	U. faunal crust (Bryozoans and or sponges - poor clarity in Zoom mode)
3	A	U. faunal crust		U. faunal crust	5	U. faunal crust (Educated guess - focus poor)
3	A	U. faunal crust		U. faunal crust	1	U. faunal crust (Guess - silt cloud from weight)
8	A	U. faunal crust		U. faunal crust	1	U. faunal crust (Rock)
8	A	U. faunal crust		U. faunal crust	1	U. faunal crust (Sandy)
8	A	U. faunal crust		U. faunal crust	4	U. faunal crust (Sponges and bryozoans)
8	A	U. faunal crust		U. faunal crust	1	U. faunal crust (Caryophylliidae)
	A	U. faunal turf		U. faunal turf	938	U. faunal turf
3	A	U. faunal turf		U. faunal turf	5	U. faunal turf (Educated guess - focus poor)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
5	A	U. faunal turf	Hydrozoa or Crinoidea	U. faunal turf (Hydrozoa or Crinoidea)	7	U. faunal turf (Hydrozoa or Crinoidea)
5	A	U. faunal turf	Hydrozoa or Crinoidea	U. faunal turf (Hydrozoa or Crinoidea)	5	U. faunal turf (Crinoidea or hydrozoa)
5	A	U. faunal turf	Hydrozoa or Crinoidea	U. faunal turf (Hydrozoa or Crinoidea)	2	Hydrozoa (Hydrozoa (Diphasia alata?) or Crinoidea (Antedon?))
5	A	U. faunal turf	Hydrozoa or Crinoidea	U. faunal turf (Hydrozoa or Crinoidea)	85	U. faunal turf (Hydrozoa (Diphasia alata?) or Crinoidea (Antedon?))
5	A	U. faunal turf	Porifera or Bryozoa	U. faunal turf (Porifera or Bryozoa)	1	U. faunal turf (Possible flustra truf)
5	A	U. faunal turf	Porifera or Bryozoa	U. faunal turf (Porifera or Bryozoa)	39	U. faunal turf (Pale branching sponge-bryozoan)
5	A	U. faunal turf	Porifera or Bryozoa	U. faunal turf (Porifera or Bryozoa)	1	U. faunal turf (Pale branching sponge-bryozoan)
5	A	U. faunal turf	Porifera or Bryozoa	U. faunal turf (Porifera or Bryozoa)	2	U. faunal turf (Pale branching sponge/bryozoan)
5	A	U. faunal turf	Porifera or Bryozoa	U. faunal turf (Porifera or Bryozoa)	1	U. faunal turf (pale branching sponge - bryozoan)
14	C000001	Porifera		Porifera	83	Porifera
14	C000001	Porifera		Porifera	9	Porifera indet.
14	C000001	Porifera	cup	Porifera (cup)	1	Porifera (Cup)
14	C000001	Porifera	cup	Porifera (cup)	1	Porifera (Cup, uncertain)
6	C000001	Porifera	cup	Porifera (cup)	4	U. faunal turf (cup-like sponge)
14	C000001	Porifera	cup	Porifera (cup)	1	Porifera (Cup r1c8)
14	C000001	Porifera	cup	Porifera (cup)	7	Porifera (cup-like Axinellidae)
14	C000001	Porifera	cup	Porifera (cup)	3	Porifera (cuplike)
14	C000001	Porifera	cup	Porifera (cup)	3	Porifera (Cup like)
14	C000001	Porifera	cup	Porifera (cup)	1	Porifera (solitary cuplike)
14	C000001	Porifera	cup	Porifera (cup)	2	Porifera (cup-like sponge - Axinellidae)
14	C000001	Porifera	enc	Porifera (enc)	5	Porifera (orange encrusting)
14	C000001	Porifera	enc	Porifera (enc)	76	Porifera indet crusts
6	C000001	Porifera	enc	Porifera (enc)	2	u. faunal crust (Porifera)
14	C000001	Porifera	enc	Porifera (enc)	5	Porifera (encrusting)
14	C000001	Porifera	enc	Porifera (enc)	1	Porifera (Crust)
6	C000001	Porifera	enc	Porifera (enc)	2	U. faunal crust (Porifera orange)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
14	C000001	Porifera	enc	Porifera (enc)	1	Porifera (Encrusting)
14	C000001	Porifera	enc	Porifera (enc)	2	Porifera indet. (Crust yellow)
14	C000001	Porifera	enc	Porifera (enc)	1	Porifera indet. (Grey crustose sponge)
14	C000001	Porifera	enc	Porifera (enc)	5	Porifera indet. (Grey crust)
14	C000001	Porifera	enc	Porifera (enc)	2	Porifera indet crusts (poor focus questionable)
14	C000001	Porifera	enc	Porifera (enc)	5	Porifera indet crusts (Buff)
14	C000001	Porifera	enc	Porifera (enc)	2	Porifera indet. (Crust orange)
14	C000001	Porifera	enc	Porifera (enc)	1	Porifera indet. (crust - buff)
14	C000001	Porifera	enc	Porifera (enc)	2	Porifera indet crusts (<1%)
14	C000001	Porifera	enc	Porifera (enc)	4	Porifera indet crusts (Orange)
14	C000001	Porifera	enc blue	Porifera (enc blue)	9	Porifera indet crusts (Blue - Hymedesmia paupertas)
14	C000001	Porifera	enc blue	Porifera (enc blue)	3	Porifera indet crusts (Blue)
14	C000001	Porifera	enc blue	Porifera (enc blue)	4	Porifera (Blue crust)
14	C000001	Porifera	enc blue	Porifera (enc blue)	32	Porifera indet. (Blue - Hymedesmia paupertas)
14	C000001	Porifera	enc blue	Porifera (enc blue)	1	Porifera (blue encrusting)
14	C000001	Porifera	enc blue	Porifera (enc blue)	137	Porifera (Blue encrusting)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	3	Porifera indet. (Flabellate)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	3	Porifera (?Massive yellow)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	5	Porifera indet. (Flabellate -buff)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	3	Porifera indet. (Erect (buff))
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	6	Porifera indet. (Papillate yellow)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	1	Porifera indet. (Erect)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	3	Porifera indet. (Erect (yellow))
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	4	Porifera indet. (Flabellate yellow)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	2	Porifera indet (Erect massive yellow)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	1	Porifera indet. (Orange massive)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	2	Porifera indet. (small massive yellow)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	4	Porifera indet. (Buff erect)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	1	Porifera indet crusts (Orange massive)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	2	Porifera (Massive yellow)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	4	Porifera indet. (Buff flabellate)
14	C000001	Porifera	non-encrusting	Porifera (non-encrusting)	1	Porifera indet. (Grey flabellate sponge)
10	C002560	Polymastia		Polymastia	5	Polymastia boletiformis
8	C002560	Polymastia		Polymastia	1	Polymastia (Polymastia)
6	C002560	Polymastia		Polymastia	17	Porifera (Polymastia)
	C002560	Polymastia		Polymastia	6	Polymastia
8	D001050	Hydrozoa		Hydrozoa	5	Hydrozoa (a)
8	D001050	Hydrozoa		Hydrozoa	5	Hydrozoa (b)
8	D001050	Hydrozoa		Hydrozoa	1	Hydrozoa (Branching hydroid)
8	D001050	Hydrozoa		Hydrozoa	1	Hydrozoa (Erect hydroid)
3	D001050	Hydrozoa		Hydrozoa	1	Hydrozoa (uncertain)
8	D001050	Hydrozoa		Hydrozoa	103	Hydrozoa (Halecium halecinum or Sertulariidae)
8	D001050	Hydrozoa		Hydrozoa	2	Hydrozoa (Hydroid)
	D001050	Hydrozoa		Hydrozoa	281	Hydrozoa
8	D001050	Hydrozoa		Hydrozoa	2	Hydrozoa (Hydroid turf)
8	D001050	Hydrozoa		Hydrozoa	4	Hydrozoa (Erect)
6	D001050	Hydrozoa		Hydrozoa	12	U. faunal turf (Hydrozoa)
6	D001050	Hydrozoa		Hydrozoa	5	U. faunal turf (Hydroids)
	D005230	Halecium		Halecium	5	Halecium
15	D005970	Nemertesia antennina		Nemertesia antennina	1	Hydrozoa (Plumulariidae)
	D005970	Nemertesia antennina		Nemertesia antennina	18	Nemertesia antennina
6	D005970	Nemertesia antennina		Nemertesia antennina	26	Hydrozoa (Nemertesia antennina)
5	D006240	Sertulariidae		Sertulariidae	5	Hydrozoa (sertulariidae)
6	D006250	Abietinaria		Abietinaria	17	Hydrozoa (Abietinaria)
	D006250	Abietinaria		Abietinaria	8	Abietinaria
6	D006250	Abietinaria		Abietinaria	1	Hydrozoa (Abietinaria?)
	D006260	Abietinaria abietina		Abietinaria abietina	2	Abietinaria abietina
6	D006260	Abietinaria abietina		Abietinaria abietina	2	Sertulariidae (Abietinaria abietina)
	D006600	Thuiaria thuja		Thuiaria thuja	43	Thuiaria thuja
6	D006600	Thuiaria thuja		Thuiaria thuja	1	Hydrozoa (Thiuaria thuja)
6	D006600	Thuiaria thuja		Thuiaria thuja	3	Hydrozoa (Thuiaria thuja)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
5	D006940	Campanulariidae		Campanulariidae	2	Hydrozoa (Very small - Obelia?)
5	D006940	Campanulariidae		Campanulariidae	1	Hydrozoa (Obelia type)
8	D010060	Anthozoa		Anthozoa	1	Anthozoa (R9C7)
	D010060	Anthozoa		Anthozoa	42	Anthozoa
3	D010060	Anthozoa		Anthozoa	1	Anthozoa (??)
3	D010060	Anthozoa		Anthozoa	2	Anthozoa (?? Can't really tell this is a complete guess)
3	D010060	Anthozoa		Anthozoa	1	Anthozoa (?? Poor focus)
8	D010060	Anthozoa		Anthozoa	1	Anthozoa (Small)
	D010060	Anthozoa	sp B	Anthozoa (sp B)	1	Species B (Anthozoa)
	D010060	Anthozoa	sp B	Anthozoa (sp B)	4	Species B (Anthozoa?)
10	D010240	Alcyonium digitatum		Alcyonium digitatum	11	Alcyonium digitatum
9	D010240	Alcyonium digitatum		Alcyonium digitatum	1	Alcyonium (Possible)
10	D010240	Alcyonium digitatum		Alcyonium digitatum	1	Alcyonium digitatum (one small ind.)
1	D010240	Alcyonium digitatum		Alcyonium digitatum	1	Alyconium (Alcyonium digitatum)
9	D010240	Alcyonium digitatum		Alcyonium digitatum	13	Alcyonium
9	D010240	Alcyonium digitatum		Alcyonium digitatum	84	Alcyonium (Alcyonium digitatum)
6	D010710	Hexacorallia		Hexacorallia	2	U. faunal turf (Anthozoa, zoantharia)
	D010710	Hexacorallia		Hexacorallia	6	Hexacorallia
7	D010710	Hexacorallia		Hexacorallia	6	Hexacorallia (Corynactis viridis)
6	D010710	Hexacorallia		Hexacorallia	2	U. faunal turf (Anthozoa - zoantharia)
5	D010710	Hexacorallia		Hexacorallia	5	Anthozoa (Zoantharia)
6	D010710	Hexacorallia		Hexacorallia	1	U. faunal turf (white anthozoa - zoantharia)
	D011310	Actiniaria		Actiniaria	69	Actiniaria
5	D011310	Actiniaria		Actiniaria	2	Anthozoa (Actiniaria?)
5	D011310	Actiniaria		Actiniaria	1	Anthozoa (Actiniaria uncertain)
7	D011310	Actiniaria		Actiniaria	5	Actiniaria (Sagartiidae)
5	D011310	Actiniaria		Actiniaria	1	Anthozoa (Sagartia/Cereus?)
5	D011310	Actiniaria		Actiniaria	1	Anthozoa (Anemone r4 c1)
5	D011310	Actiniaria		Actiniaria	2	Anthozoa (Ceriantharia)
5	D011310	Actiniaria		Actiniaria	2	Anthozoa (Sagartia?)
5	D011310	Actiniaria		Actiniaria	1	Anthozoa (Sagartia)
5	D011310	Actiniaria		Actiniaria	1	Anthozoa (Actiniaria)
3	D011310	Actiniaria		Actiniaria	12	Actiniaria (uncertain)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
6	D011670	Urticina		Urticina	14	Actiniaria (Urticina closed)
10	D011670	Urticina		Urticina	39	Urticina felina
3	D011670	Urticina		Urticina	1	Urticina (?)
6	D011670	Urticina		Urticina	11	Actiniaria (Urticina felina)
6	D011670	Urticina		Urticina	79	Actiniaria (Urticina)
6	D011670	Urticina		Urticina	1	Anthozoa (Urtcina)
6	D011670	Urticina		Urticina	13	Anthozoa (Urticina)
12	D011670	Urticina		Urticina	25	Urticina
	D012370	Cereus pedunculatus		Cereus pedunculatus	9	Cereus pedunculatus
3	D012370	Cereus pedunculatus		Cereus pedunculatus	1	Cereus pedunculatus (? Poor focus)
	D013610	Scleractinia		Scleractinia	58	Scleractinia
6	D013700	Caryophyllia smithii		Caryophyllia smithii	1	Scleractinia (Caryphyllidae)
13	D013700	Caryophyllia smithii		Caryophyllia smithii	35	Caryophyllia
6	D013700	Caryophyllia smithii		Caryophyllia smithii	1	Scleractinia (Caryophyllia)
6	D013700	Caryophyllia smithii		Caryophyllia smithii	51	Scleractinia (Caryophyllia)
6	D013700	Caryophyllia smithii		Caryophyllia smithii	51	Scleractinia (Cup coral)
6	D013700	Caryophyllia smithii		Caryophyllia smithii	210	Scleractinia (Caryophylliidae)
	D013700	Caryophyllia smithii		Caryophyllia smithii	59	Caryophyllia smithii
6	D013700	Caryophyllia smithii		Caryophyllia smithii	1	Scleractinia (Caryophillia)
6	D013700	Caryophyllia smithii		Caryophyllia smithii	3	Scleractinia (Caryolphyllidae)
6	D013700	Caryophyllia smithii		Caryophyllia smithii	8	Scleractinia (Caryophillia)
6	D013700	Caryophyllia smithii		Caryophyllia smithii	8	Scleractinia (Caryophyillia)
2	P000010	Polychaeta	tube	Polychaeta (tube)	1	Polychaeta (Tube)
1	P020310	Lanice conchilega		Lanice conchilega	3	Lanice concheliga
6	P020310	Lanice conchilega		Lanice conchilega	1	Terebellidae
6	P020310	Lanice conchilega		Lanice conchilega	5	Terebellidae (Lanice)
6	P020310	Lanice conchilega		Lanice conchilega	4	Terebellidae (Lanice?)
4	P020310	Lanice conchilega		Lanice conchilega	4	Lanice

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
4	P020310	Lanice conchilega		Lanice conchilega	1	Lanice (Lanice)
	P022590	Sabella	tube	Sabella (tube)	12	Sabella (tube)
3	P022720	Serpulidae		Serpulidae	1	Serpulidae (Caryophylliidae)
3	P022720	Serpulidae		Serpulidae	2	Serpulidae (?? Can't really tell this is a complete guess)
	P022720	Serpulidae		Serpulidae	379	Serpulidae
3	P022720	Serpulidae		Serpulidae	1	Serpulidae (??)
6	P023020	Spirobranchus		Spirobranchus	1	U. faunal crust (spirobranchus)
6	P023020	Spirobranchus		Spirobranchus	514	Serpulidae (Spirobranchus)
6	P023020	Spirobranchus		Spirobranchus	5	Serpulidae (Spirobranchus *et al.*)
3	P023460	Protula tubularia		Protula tubularia	4	Protula tubularia (?)
3	P023460	Protula tubularia		Protula tubularia	1	Protula tubularia (Empty tube)
6	P023550	Spirorbidae		Spirorbidae	1	u. faunal crust (Spirobida)
	P023550	Spirorbidae		Spirorbidae	100	Spirorbidae
7	R000210	Cirripedia		Cirripedia	1	Cirripedia (Semibalanus)
	R000210	Cirripedia		Cirripedia	20	Cirripedia
3	R000210	Cirripedia		Cirripedia	2	Cirripedia (??)
3	R000210	Cirripedia		Cirripedia	2	Cirripedia (?? Can't really tell this is a complete guess)
3	R000210	Cirripedia		Cirripedia	1	Cirripedia (?? Poor focus)
3	R000210	Cirripedia		Cirripedia	2	Cirripedia (poor focus questionable)
3	R000210	Cirripedia		Cirripedia	1	Cirripedia (Guesstimate in lower half of photo)
5	S021440	Decapoda		Decapoda	11	Crustacea (Legs)
5	S021440	Decapoda		Decapoda	9	Crustacea
5	S021440	Decapoda		Decapoda	1	Crustacea (Eye r1 c1)
5	S021440	Decapoda		Decapoda	1	Crustacea (Leg)
5	S021440	Decapoda		Decapoda	1	Crustacea (R2C1)
5	S021440	Decapoda		Decapoda	1	Crustacea (Small r2c1)
5	S021440	Decapoda		Decapoda	27	Decapoda
5	S021440	Decapoda	crab	Decapoda (crab)	8	Brachyura (uncertain)
5	S021440	Decapoda	crab	Decapoda (crab)	1	Brachyura
5	S021440	Decapoda	crab	Decapoda (crab)	2	Crustacea (Crab)
5	S021440	Decapoda	crab	Decapoda (crab)	1	Decapoda (?Crab)
5	S021440	Decapoda	crab	Decapoda (crab)	1	Decapoda (Inachus ?)
5	S021440	Decapoda	crab	Decapoda (crab)	1	Decapoda (Crab carapace, claw and leg)
5	S021440	Decapoda	crab	Decapoda (crab)	3	Decapoda (Crab carapace, claw and legs)
5	S021440	Decapoda	crab	Decapoda (crab)	5	Decapoda (Crab)
5	S021440	Decapoda	crab	Decapoda (crab)	1	Crustacea (Small crab r2 c5)
5	S021440	Decapoda	crab	Decapoda (crab)	13	Decapoda (Brachyura)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
5	S021440	Decapoda	crab	Decapoda (crab)	4	Decapoda (Inachus?)
8	S021690	Caridea		Caridea	1	Caridae
8	S021690	Caridea		Caridea	1	Caridae (antenna)
	S021690	Caridea		Caridea	75	Caridea
3	S021690	Caridea		Caridea	1	Caridea (?? Poor focus)
8	S021690	Caridea		Caridea	4	Caridea (Prawn)
8	S024440	Paguridae		Paguridae	1	Paguridae (Two)
3	S024440	Paguridae		Paguridae	1	Paguridae (2?)
3	S024440	Paguridae		Paguridae	1	Paguridae (uncertain)
7	S024440	Paguridae		Paguridae	5	Paguridae (Pagurus prideaux)
7	S024440	Paguridae		Paguridae	2	Paguridae (Pagurus bernhardus)
3	S024440	Paguridae		Paguridae	4	Paguridae (?)
	S024440	Paguridae		Paguridae	60	Paguridae
5	S024440	Paguridae		Paguridae	5	Paguroidea
7	S024440	Paguridae		Paguridae	4	Decapoda (Pagarus)
3	S024440	Paguridae		Paguridae	1	Paguridae (? Poor focus)
	S024820	Galatheoidea		Galatheoidea	17	Galatheoidea
3	S024820	Galatheoidea		Galatheoidea	1	Galatheoidea (claws only)
6	S024940	Munida		Munida	34	Galatheoidea (Munida)
6	S024940	Munida		Munida	1	Crustacea (Mudida r3 c5)
	S024940	Munida		Munida	5	Munida
6	S024940	Munida		Munida	1	Galatheoidea (Munida - claws only)
6	S024940	Munida		Munida	1	Galatheoidea (Mundia)
6	S024940	Munida		Munida	1	Decapoda (Munida)
6	S024940	Munida		Munida	3	Galatheoidea (Munida claws only)
5	S025000	Porcellanidae		Porcellanidae	2	Crustacea (porcelain carbs)
5	S025000	Porcellanidae		Porcellanidae	1	Crustacea (Crab (Small porcelin))
8	S025430	Ebalia tuberosa		Ebalia tuberosa	1	Ebalia tuberosa (Possibly 2)
3	S025430	Ebalia tuberosa		Ebalia tuberosa	1	Ebalia tuberosa (??2)
3	S025430	Ebalia tuberosa		Ebalia tuberosa	3	Ebalia tuberosa (2?)
	S026460	Cancer pagurus		Cancer pagurus	41	Cancer pagurus
6	S026720	Necora puber		Necora puber	4	Brachyura (Necora? Legs only)
6	S026720	Necora puber		Necora puber	1	Brachyura (Necora uncertain)
	S026720	Necora puber		Necora puber	5	Necora puber
	W000500	Polyplacophora		Polyplacophora	9	Polyplacophora
8	W000500	Polyplacophora		Polyplacophora	1	Polyplacophora (Chiton)
	W000920	Gastropoda		Gastropoda	5	Gastropoda
8	W001235	Acmaeidae		Acmaeidae	2	Acmaeidae (In WORMS!)
8	W001235	Acmaeidae		Acmaeidae	1	Acmaeidae (Not in list but in WORMS)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
8	W001235	Acmaeidae		Acmaeidae	2	Acmaeidae (IN WORMS?)
5	W001570	Trochidae		Trochidae	1	Gastropoda (Gibbula)
5	W001570	Trochidae		Trochidae	2	Gastropoda (?Gibbula)
5	W001570	Trochidae		Trochidae	5	Mollusca (Top shell)
5	W001570	Trochidae		Trochidae	120	Gastropoda (Trochidae)
	W002000	Calliostoma zizyphinum		Calliostoma zizyphinum	11	Calliostoma zizyphinum
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	5	Calliostoma zizyphinum (Could be pagurids)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	1	Calliostoma zizyphinum (??)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	2	Calliostoma zizyphinum (Possibly pagurid)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	5	Calliostoma zizyphinum (Paguridae?)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	3	Calliostoma zizyphinum (? Paguridae)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	2	Calliostoma zizyphinum (?Pagurid)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	5	Calliostoma zizyphinum (Could be paguridae)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	5	Calliostoma zizyphinum (?paguridae)
3	W002000	Calliostoma zizyphinum	possibly Paguridae	Calliostoma zizyphinum (possibly Paguridae)	1	Calliostoma zizyphinum (? Paguridae)
3	W016120	Bivalvia		Bivalvia	1	Bivalvia (Uncertain)
	W017740	Pectinidae		Pectinidae	26	Pectinidae
6	W018050	Aequipecten opercularis		Aequipecten opercularis	1	Pectinidae (Queenie r4 c5)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
	W018050	Aequipecten opercularis		Aequipecten opercularis	15	Aequipecten opercularis
6	W018050	Aequipecten opercularis		Aequipecten opercularis	1	Pectinidae (Chlamys opercularis)
6	W018050	Aequipecten opercularis		Aequipecten opercularis	5	Pectinidae (Aequipecten opercularis)
6	W018090	Pecten maximus		Pecten maximus	15	Pectinidae (Pecten maximus)
	W018090	Pecten maximus		Pecten maximus	6	Pecten maximus
3	W018130	Anomiidae		Anomiidae	2	Anomiidae (Guesstimate)
3	W018130	Anomiidae		Anomiidae	1	Anomiidae (Guess poor focus)
15	W018130	Anomiidae		Anomiidae	1	Pododesmus
3	W018130	Anomiidae		Anomiidae	2	Anomiidae (Probably more)
3	W018130	Anomiidae		Anomiidae	11	Anomiidae (Pododesmus?)
7	W018130	Anomiidae		Anomiidae	9	Anomiidae (Pododesmus)
3	W018130	Anomiidae		Anomiidae	4	Anomiidae (Could be miles out - covered in brittlestars)
	W018130	Anomiidae		Anomiidae	29	Anomiidae
7	W018130	Anomiidae		Anomiidae	1	Anomiidae (Pododesmus - estimate)
3	W018130	Anomiidae		Anomiidae	1	Anomiidae (?? Poor focus)
3	W018130	Anomiidae		Anomiidae	1	Anomiidae (Guess beneath brittlestars)
7	W018130	Anomiidae		Anomiidae	1	Anomiidae (educated guess - Pododesmus)
3	W018130	Anomiidae		Anomiidae	1	Anomiidae (Guess)
3	W018130	Anomiidae		Anomiidae	1	Anomiidae (Guess - poorly lit)
3	W018130	Anomiidae		Anomiidae	1	Anomiidae (Estimate)
1	W018250	Veneroidea		Veneroidea	7	Veneroida
	W019650	Cardiidae		Cardiidae	3	Cardiidae
1	Y000001	Bryozoa		Bryozoa	1	Bryozoa (Bryozaoan)
	Y000001	Bryozoa		Bryozoa	26	Bryozoa
2	Y000001	Bryozoa	enc	Bryozoa (enc)	1	Bryozoa indet crusts (?Gibbula)
2	Y000001	Bryozoa	enc	Bryozoa (enc)	4	Bryozoa indet crusts (Covered in brittlestars)
2	Y000001	Bryozoa	enc	Bryozoa (enc)	5	Bryozoa indet crusts (Educated guess - focus poor)
2	Y000001	Bryozoa	enc	Bryozoa (enc)	1	Bryozoa indet crusts (Orange)
2	Y000001	Bryozoa	enc	Bryozoa (enc)	1	Bryozoa indet crusts (Orange - or sponge poor focus)
6	Y000001	Bryozoa	enc	Bryozoa (enc)	1	U. faunal crust (Bryozoa)
2	Y000001	Bryozoa	enc	Bryozoa (enc)	88	Bryozoa indet crusts
2	Y000001	Bryozoa	enc	Bryozoa (enc)	1	Bryozoa indet crusts (??)
6	Y000001	Bryozoa	enc	Bryozoa (enc)	1	U. faunal crust (Bryozoan crust)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
2	Y000001	Bryozoa	enc	Bryozoa (enc)	1	Bryozoa indet crusts (Sand)
2	Y000001	Bryozoa	erect	Bryozoa (erect)	5	Bryozoa (Erect)
	Y000001	Bryozoa	erect	Bryozoa (erect)	1	Bryozoa (Erect - flabellate)
2	Y000001	Bryozoa	erect	Bryozoa (erect)	29	Bryozoa (Erect)
2	Y000001	Bryozoa	staghorn	Bryozoa (staghorn)	2	Bryozoa (White branching)
6	Y000001	Bryozoa	staghorn	Bryozoa (staghorn)	1	U. faunal turf (small branching staghorn)
2	Y000001	Bryozoa	staghorn	Bryozoa (staghorn)	8	Bryozoa (Branching)
2	Y000001	Bryozoa	staghorn	Bryozoa (staghorn)	5	Bryozoa (Reteporella)
2	Y000001	Bryozoa	staghorn	Bryozoa (staghorn)	18	Bryozoa (small branching staghorn)
	Y001340	Alcyonidium		Alcyonidium	9	Alcyonidium
3	Y006920	Flustridae		Flustridae	2	Flustridae (Flustra Securiflustra mix)
3	Y006920	Flustridae		Flustridae	1	Flustridae (Flustra & Securiflustra mix)
6	Y006920	Flustridae		Flustridae	3	U. faunal turf (Flustridae)
3	Y006920	Flustridae		Flustridae	1	Flustridae (Securiflustra and Flustra mix)
	Y006920	Flustridae		Flustridae	172	Flustridae
6	Y006940	Flustra foliacea		Flustra foliacea	1	Flustridae (f.foliacea)
	Y006940	Flustra foliacea		Flustra foliacea	106	Flustra foliacea
4	Y006940	Flustra foliacea		Flustra	55	Flustra
6	Y006940	Flustra foliacea		Flustra foliacea	16	Flustridae (Flustra foliacea)
	Y007100	Securiflustra securifrons		Securiflustra securifrons	121	Securiflustra securifrons
6	Y007100	Securiflustra securifrons		Securiflustra securifrons	39	Flustridae (Securiflustra securifrons)
3	Y007100	Securiflustra securifrons		Securiflustra securifrons	1	Securiflustra securifrons (??)
6	Y007100	Securiflustra securifrons		Securiflustra securifrons	2	Flustridae (Securiflustra)
3	Y007100	Securiflustra securifrons		Securiflustra securifrons	1	Securiflustra securifrons (? Sponge)
	ZB00000	Echinodermata		Echinodermata	1	Echinodermata
3	ZB00010	Crinoidea		Crinoidea	1	Crinoidea (uncertain)
	ZB00010	Crinoidea		Crinoidea	101	Crinoidea
6	ZB00100	Antedon		Antedon	23	Crinoidea (Antedon)
	ZB00100	Antedon		Antedon	5	Antedon
11	ZB00110	Antedon bifida		Antedon bifida	27	Antedon bifida
11	ZB00220	Leptometra celtica		Leptometra celtica	5	Leptometra celtica
	ZB00310	Asteroidea		Asteroidea	45	Asteroidea

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
3	ZB00310	Asteroidea		Asteroidea	9	Asteroidea (uncertain)
8	ZB00310	Asteroidea		Asteroidea	2	Asteroidea (R2C1)
7	ZB00310	Asteroidea		Asteroidea	1	Asteroidea (Henricia?)
7	ZB00310	Asteroidea		Asteroidea	2	Asteroidea (Henricia)
7	ZB00310	Asteroidea		Asteroidea	1	Asteroidea (Cushion star)
7	ZB00310	Asteroidea		Asteroidea	1	Asteroidea (Cushion star)
8	ZB00310	Asteroidea		Asteroidea	25	Asteroidea (small)
9	ZB00670	Luidia ciliaris		Luidia ciliaris	2	Luidia
	ZB00670	Luidia ciliaris		Luidia ciliaris	25	Luidia ciliaris
6	ZB00670	Luidia ciliaris		Luidia ciliaris	10	Asteroidea (Luidia)
	ZB00940	Hippasteria		Hippasteria	15	Hippasteria
	ZB01010	Porania pulvillus		Porania pulvillus	6	Porania pulvillus
6	ZB01010	Porania pulvillus		Porania pulvillus	15	Asteroidea (Porania)
	ZB01490	Crossaster papposus		Crossaster papposus	141	Crossaster papposus
4	ZB01490	Crossaster papposus		Crossaster papposus	1	Crossaster (Crossaster)
4	ZB01490	Crossaster papposus		Crossaster papposus	7	Crossaster
6	ZB01490	Crossaster papposus		Crossaster papposus	1	Asteroidea (Crossaster)
6	ZB01490	Crossaster papposus		Crossaster papposus	1	Asteroidea (Crossaster uncertain)
6	ZB01490	Crossaster papposus		Crossaster papposus	12	Asteroidea (Crossaster papposus)
8	ZB01900	Asterias rubens		Asterias rubens	1	Asterias rubens (1xsmall 1xlarge)
	ZB01900	Asterias rubens		Asterias rubens	42	Asterias rubens
4	ZB01900	Asterias rubens		Asterias rubens	6	Asterias
6	ZB01900	Asterias rubens		Asterias rubens	3	Asteroidea (Asterias)
6	ZB01900	Asterias rubens		Asterias rubens	7	Asteroidea (Asterias rubens)
4	ZB01900	Asterias rubens		Asterias rubens	3	Asterias (Juv.)
8	ZB01950	Leptasterias muelleri		Leptasterias muelleri	4	Leptasterias muelleri (Buried)
	ZB01950	Leptasterias muelleri		Leptasterias muelleri	1	Leptasterias muelleri
5	ZB02040	Ophiuroidea		Ophiuroidea	5	Ophiuridae (Ophiocomina nigra & Ophiothrix mix)
5	ZB02040	Ophiuroidea		Ophiuroidea	1	Ophiuridae (Too jumbled to count accurately in places but well over 100 which are superabundant)
5	ZB02040	Ophiuroidea		Ophiuroidea	1	Ophiuridae (Ophiocomina nigra & Ophiothrix mix, superabundant 100+)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
5	ZB02040	Ophiuroidea		Ophiuroidea	1	Ophiuridae (Ophiura albida over turned)
5	ZB02040	Ophiuroidea		Ophiuroidea	1	Ophiuridae (x)
	ZB02040	Ophiuroidea		Ophiuroidea	443	Ophiuroidea
5	ZB02040	Ophiuroidea		Ophiuroidea	60	Ophiuridae
x	ZB02040	Ophiuroidea		Ophiuroidea	3	Amphiura
3	ZB02040	Ophiuroidea		Ophiuroidea	7	Ophiuroidea (uncertain)
3	ZB02040	Ophiuroidea		Ophiuroidea	424	Ophiuroidea (uncertain)
8	ZB02040	Ophiuroidea		Ophiuroidea	1	Ophiuroidea (Caryophylliidae)
	ZB02780	Ophiopholis aculeata		Ophiopholis aculeata	16	Ophiopholis aculeata
8	ZB03100	Ophiocomina nigra		Ophiocomina nigra	1	Ophiocomina nigra (Sand)
8	ZB03100	Ophiocomina nigra		Ophiocomina nigra	1	Ophiocomina nigra (Estimate)
	ZB03100	Ophiocomina nigra		Ophiocomina nigra	173	Ophiocomina nigra
6	ZB03100	Ophiocomina nigra		Ophiocomina nigra	1	Ophiuroidea (Ophiocomina nigra)
6	ZB03100	Ophiocomina nigra		Ophiocomina nigra	5	Ophiuridae (Ophiocomina nigra)
6	ZB03100	Ophiocomina nigra		Ophiocomina nigra	36	Ophiuridae (Ophiocomina nigra)
6	ZB03100	Ophiocomina nigra		Ophiocomina nigra	512	Ophiuroidea (Ophiocomina nigra)
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	1	Ophiuridae (Ophithrix)
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	2	Ophiuroidea (Ophiothrix fragilis)
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	2	Ophiuridae (Ophiothrix)
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	2	Ophiuridae (Ophiothirx)
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	1	Ophiuridae (Mostly Ophithrix)
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	1	Ophiuridae (Ophiothrix)
	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	123	Ophiothrix fragilis
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	17	Ophiuridae (Ophiothrix fragilis)
8	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	2	Ophiothrix fragilis (Estimate)
6	ZB03100	Ophiothrix fragilis		Ophiothrix fragilis	183	Ophiuroidea (Ophiothrix fragilis)
5	ZB03380	Echinoidea		Echinoidea	1	Echinodermata-Echinidae
5	ZB03380	Echinoidea		Echinoidea	4	Echinodermata (Echinidae)

Tcode	TaxCode	TaxaNew	QualifierNew	EntityNew	Records	Original Taxon & Qualifier
8	ZB03380	Echinoidea		Echinoidea	1	Echinoidea (small)
	ZB03380	Echinoidea		Echinoidea	48	Echinoidea
8	ZB03620	Echinus esculentus		Echinus esculentus	1	Echinus esculentus (Possibly a third)
6	ZB03620	Echinus esculentus		Echinus esculentus	1	Echinoidea (Echinus?)
6	ZB03620	Echinus esculentus		Echinus esculentus	2	Echinoidea (Echinus esculentus)
	ZB03620	Echinus esculentus		Echinus esculentus	111	Echinus esculentus
	ZB03870	Echinocyamus		Echinocyamus	5	Echinocyamus
	ZB04790	Aslia lefevrei		Aslia lefevrei	5	Aslia lefevrei
4	ZB04790	Aslia lefevrei		Aslia lefevrei	1	Aslia
7	ZD00001	Tunicata		Tunicata	34	Tunicata
3	ZD00001	Tunicata		Tunicata	1	Tunicata (V Dubious R10 C6)
3	ZD00001	Tunicata		Tunicata	27	Tunicata (uncertain)
7	ZD00001	Tunicata		Tunicata	5	Tunicata (Corella parallelogramma)
3	ZD01480	Ascidia		Ascidia	1	Ascidia (??)
3	ZD01480	Ascidia		Ascidia	1	Ascidia (? Not clear)
3	ZD01480	Ascidia		Ascidia	1	Ascidia (?? Poor focus)
6	ZD01480	Ascidia		Ascidia	2	Tunicata (Ascidia sp.)
11	ZD01500	Ascidia mentula		Ascidia mentula	5	Ascidia mentula
3	ZD01500	Ascidia mentula		Ascidia mentula	1	Ascidia mentula (Ascidia virginia?)
3	ZE	Actinopterygii		Actinopterygii	1	Actinopterygii (uncertain)
5	ZE	Actinopterygii		Actinopterygii	10	Pisces
	ZE	Actinopterygii		Actinopterygii	23	Actinopterygii
5	ZE	Actinopterygii		Actinopterygii	9	Teleostei
7	ZG01500	Gadidae		Gadidae	19	Actinopterygii (trisopterus)
7	ZG01500	Gadidae		Gadidae	1	Actinopterygii (Pouting? Trisopterus)
3	ZG01500	Gadidae		Gadidae	3	Gadidae (Bib?)
	ZG01500	Gadidae		Gadidae	7	Gadidae
3	ZX	Alga	enc	Alga (enc)	1	U. algal crust (Guess - poorly lit)
7	ZX	Alga	enc	Alga (enc)	7	U. algal crust (L. sonderi)
7	ZX	Alga	enc	Alga (enc)	36	U. algal crust (Lithothamnion sonderi)
	ZX	Alga	enc	Alga (enc)	35	U. algal crust
	ZZ	No identifiable taxa		No identifiable taxa	201	No identifiable taxa

Appendix 3: Additional results

Precision

Table 18 can be used to compare the relative precision for individual taxa. Thus, for the cup coral *Caryophyllia smithii*, the extraction method with the highest precision (lowest value: 0.13) was SACFOR, while the Point Intercept method had the lowest precision (highest value: 0.66) [though, as noted earlier, this is likely due to the low number of positive records].

At the bottom of Table the average precision across all listed taxa suggests that SACFOR had the overall highest precision and Point Intercept the lowest. Again, however, the influence of the number of records may be significant, and in any case the differences between any of the extraction methods appear to be small.

The last row in the Table gives precision values for the number of taxa recorded. As all of these values are based on 100 images, they are directly comparable. They suggest that Point Intercept had the highest precision and Abundance Counts the lowest when the number of taxa recorded is considered.

Table 18. Precision (Standard Error / Sample Mean) for all taxa recorded frequently by Analyst A, ordered by the number of records with values>0 (Recs). N = 100 images.

Form	Entity	Recs	%c	As	Co	F0	F5	Pi
Crust/Bed	U. faunal crust	489	0.04	0.02		0.02	0.03	0.03
Indiv_ram	Ophiuroidea	447		0.03	0.12	0.07	0.05	0.11
Indiv_ram	*Ophiocomina nigra*	431		0.04	0.11	0.07	0.06	0.14
Turf	U. faunal turf	403	0.24	0.06		0.11	0.09	0.14
Turf_indiv	*Spirobranchus*	361	0.25	0.08		0.08	0.07	0.09
Indiv_sml	*Caryophyllia smithii*	171		0.13	0.24	0.21	0.17	0.66
Turf	Hydrozoa	163	0.29	0.14		0.19	0.16	0.26
Indiv_ram	*Ophiothrix fragilis*	155		0.15	0.21	0.19	0.18	0.29
Turf_indiv	Serpulidae	119	0.23	0.17		0.18	0.19	0.42
Crust_dis	Porifera (enc blue)	112	0.34	0.19		0.25	0.24	0.33
Indiv_lrg	*Urticina*	78		0.21	0.25	0.26	0.23	0.34
Indiv_lrg	*Crossaster papposus*	75		0.23	0.21	0.24	0.24	0.49
Turf	Flustridae	75	0.42	0.23		0.48	0.29	0.41
Turf	U. faunal turf (Hydrozoa or Crinoidea)	74	0.40	0.24		0.36	0.34	0.45
Massive	*Alcyonium digitatum*	70	0.40	0.27		0.35	0.34	0.45
Turf	*Securiflustra securifrons*	66	0.43	0.29		0.35	0.28	0.45
Turf	*Flustra foliacea*	57	0.48	0.31		0.39	0.32	0.50
Indiv_med	Actiniaria	43		0.37	0.27	0.31	0.34	
Indiv_sml	Asteroidea	40		0.31	0.39	0.36	0.39	
Indiv_ram	Caridea	38		0.32	0.34	0.32	0.36	
Turf	U. faunal turf (Porifera or Bryozoa)	35	0.51	0.36		0.43	0.44	0.70
Indiv_ram	*Munida*	33		0.40	0.34	0.36	0.36	0.70
Indiv_med	Tunicata	33		0.34	0.44	0.39	0.34	0.57
Indiv_lrg	*Luidia ciliaris*	30		0.37	0.37	0.46	0.47	0.55
Turf	*Nemertesia antennina*	29	0.40	0.37		0.41	0.40	0.70
Indiv_lrg	*Echinus esculentus*	29		0.40	0.42	0.42	0.38	0.52
Indiv_med	Echinoidea	21		0.45	0.47	0.48	0.52	0.70
Turf_indiv	Bryozoa (staghorn)	20	0.44	0.57		0.47	0.46	
Indiv_ram	Decapoda (crab)	19		0.49	0.52	0.42	0.47	
Indiv_ram	*Antedon*	16		0.49	0.46	0.74	0.52	
Indiv_sml	Scleractinia	16		0.50	0.69	0.58		
Indiv_med	Gadidae	15		0.50	0.49	0.49	0.61	
Indiv_med	Paguridae	15		0.57	0.49	0.52	0.49	
Turf_indiv	Hexacorallia	13	1.00	0.74		0.92	0.70	0.97
	Average precision		0.42	0.34	0.38	0.39	0.35	0.44
	Precision of No. of taxa		0.028	0.023	0.039	0.025	0.023	0.019

Form is a categorisation of the life form of each taxon: Crust/bed= Encrusting/bed forming; Crust_dis= Encrusting and distinct; Individ_lrg= large individuals; Individ_med= medium sized individuals; Individ_sml= small individuals; Individ_ram= ramose individuals; Massive= massive form; Turf= turf forming; Turf_individ= ground covering individuals. Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

Sampling Power

Table 19. Mean and (variance) for the taxa listed in Table 5.

	%c	Co	As	Pi	F0	F5
U faunal crust	58.7 (594.2)		5.1 (0.7)	51.9 (264.3)	92.4 (388.1)	23.3 (35.6)
U faunal turf	1.2 (7.7)		1.3 (0.7)	2.2 (9.4)	18.7 (429.7)	9.1 (61.4)
Hydrozoa	0.33 (0.89)		0.56 (0.61)	0.49 (1.67)	2.56 (23.24)	1.35 (4.96)
Serpulidae	2.1 (13.3)		1.1 (0.3)	2.4 (4.2)	67 (828.4)	20.5 (55.6)
Cup corals		4.7 (183.1)	1.5 (3.2)	0.1 (0.1)	3.7 (55.5)	2.4 (16.2)
Ophiocomina		22.7 (581.6)	4.3 (3)	2.9 (15.6)	31.6 (514.1)	13.5 (60.6)
Calliostoma		0.36 (0.62)	0.61 (1.51)	0.01 (0.01)	0.54 (1.3)	0.4 (0.73)
Richness	4.6 (1.6)	4.2 (2.7)	8.6 (4)	6.6 (1.5)	8.9 (4.9)	8.7 (4.2)

Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

Efficiency

Prior to comparisons between methods (Section 3.5) assessment of the effectiveness of the randomization process was achieved by comparing the mid-point in time for completion of the extraction (by method and analyst) and comparing this with the whole analysis period (by analyst). The results are shown in Table 20. It shows that the greatest potential bias was for Analyst B, who completed most of the Abundance Count samples many hours before completing the Percentage Cover samples, potentially having greater familiarity, and therefore reduced analysis time, for the Percentage Cover samples. However, comparing that with the length of the whole analysis period shows that the bias was likely small.

Table 20. Effectiveness of randomization process, by extraction method and analyst (as indicated A-F). Period is the time between starting the first image and the completing the last image (not the time spent analysing). Hours difference from average start time should be compared with that Period.

Analyst	A	B	C	D	E	F	Average
Period (hours)	1054	431	236	76	246	193	373
Hours difference from average start time:							
Abundance count	-31.6	-48.8	6.8	2.8	14.7	-8.1	-10.7
Point Intercept	-14.3	8.1	-0.8	3.0	-10.3	-0.2	-2.4
5x5	11.2	-0.6	-2.4	-1.1	-11.4	-2.3	-1.1
SACFOR	8.9	-1.3	-13.8	-6.3	18.6	-3.7	0.4
10x10	8.4	-4.0	8.5	4.4	-3.3	5.6	3.3
Percentage cover	17.4	46.5	1.6	-2.8	-8.3	8.7	10.5

The data in Figure 14, for Analyst A, also suggests a lack of any notable bias in analysis time for any extraction method. Interestingly, the only apparent improvement in analysis time was between the first and the tenth image analysed. Average analysis time between the tenth and the 100th image reduced only very slightly. Previous experience of benthic image analyses suggests that there would have been further improvements if only one extraction method had been used.

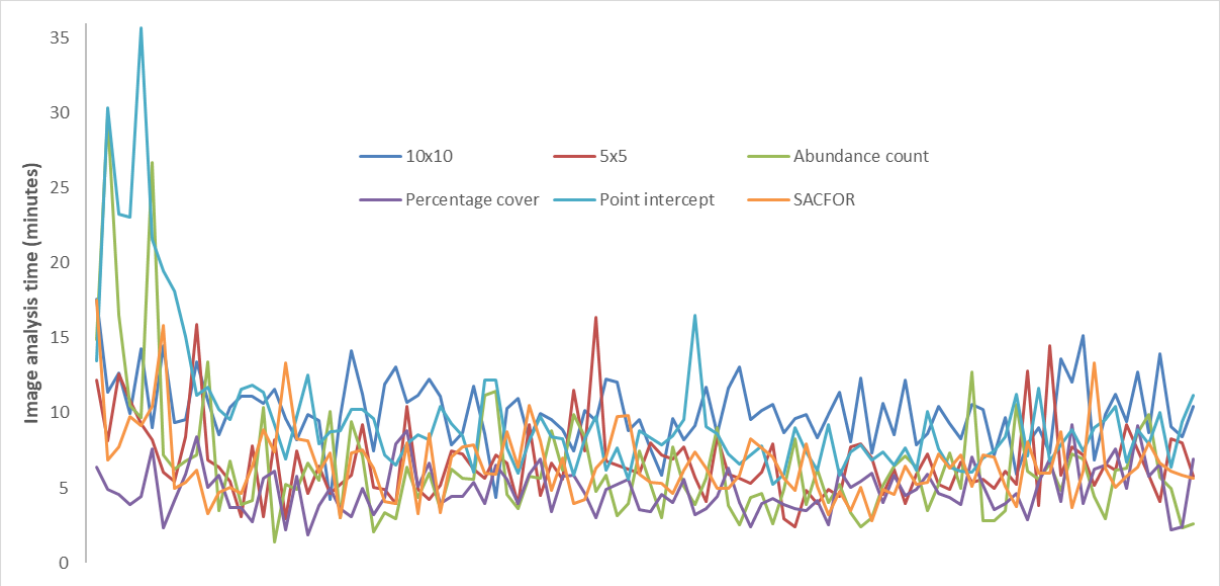


Figure 14. Number of minutes taken to analyse each image by Analyst A, in chronological order, by extraction method.

Table 21. Minimum and maximum number of minutes to analyse an image, by extraction method and analyst (as indicated A-F). Note: Data from 100 images for Analyst A, but only 20 images for the other analysts.

Minimum mins by Method & Analyst

Data extraction method	A	B	C	D	E	F	Avg
Percentage cover	1.9	1.0	2.0	3.7	2.3	0.4	1.9
Abundance count	1.4	2.3	4.9	1.7	1.6	0.3	2.0
SACFOR	2.8	0.8	3.8	3.1	4.5	2.7	2.9
5x5	2.4	2.2	7.2	2.7	3.6	1.1	3.2
10x10	4.2	2.5	4.9	4.5	7.5	2.4	4.3
Point Intercept	5.3	5.0	8.4	6.4	4.6	5.7	5.9
Avg	3.0	2.3	5.2	3.7	4.0	2.1	

Maximum mins by Method & Analyst

Data extraction method	A	B	C	D	E	F	Avg
Abundance count	29.8	6.3	18.9	15.5	17.0	13.4	16.8
Percentage cover	9.2	8.4	18.7	15.3	33.6	26.9	18.7
SACFOR	17.5	8.7	20.6	89.9	19.7	19.0	29.2
5x5	16.4	8.1	24.1	13.8	59.9	59.9	30.4
10x10	17.6	15.8	27.0	33.0	44.8	49.2	31.2
Point Intercept	35.7	9.8	37.3	28.0	62.0	24.7	32.9
Avg	21.0	9.5	24.4	32.6	39.5	32.2	

Methods: 10x10 Frequency (10x10); 5x5 Frequency (5x5).

Taxonomic accumulation

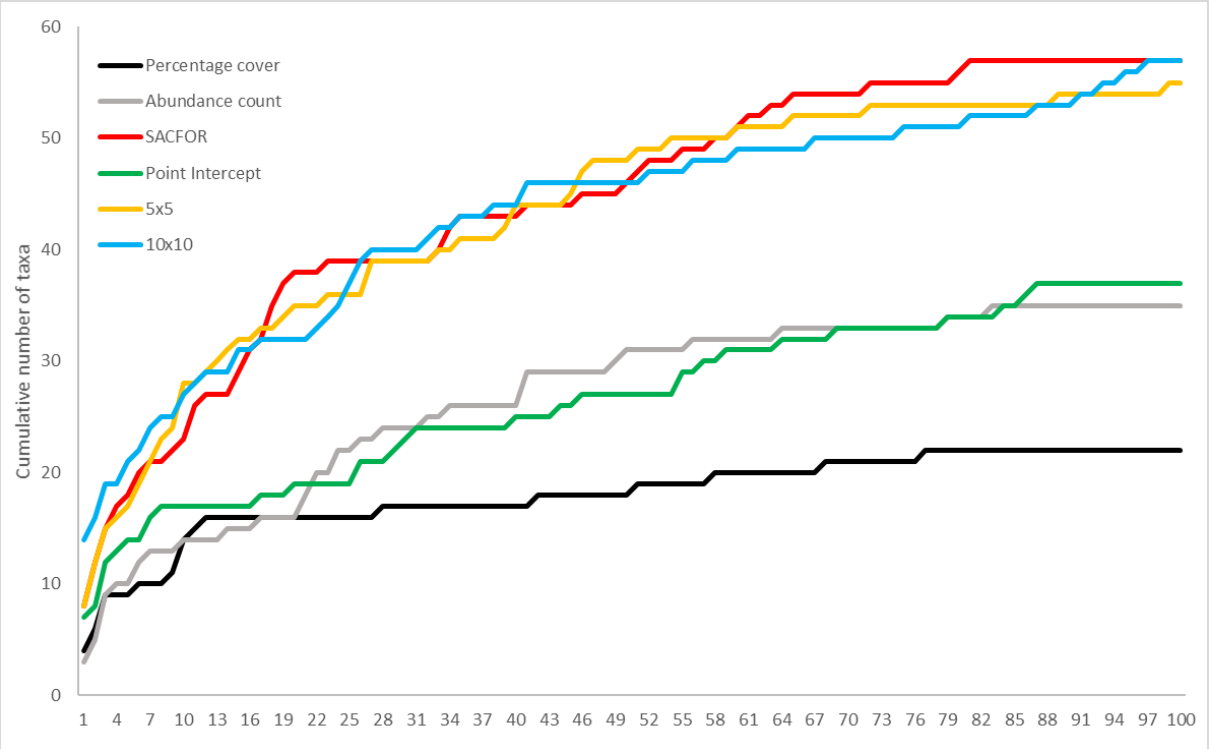


Figure 15. Cumulative number of taxa recorded from 100 images (in original order) by Analyst A, by extraction method. Methods: 10x10 Frequency (10x10); 5x5 Frequency (5x5).

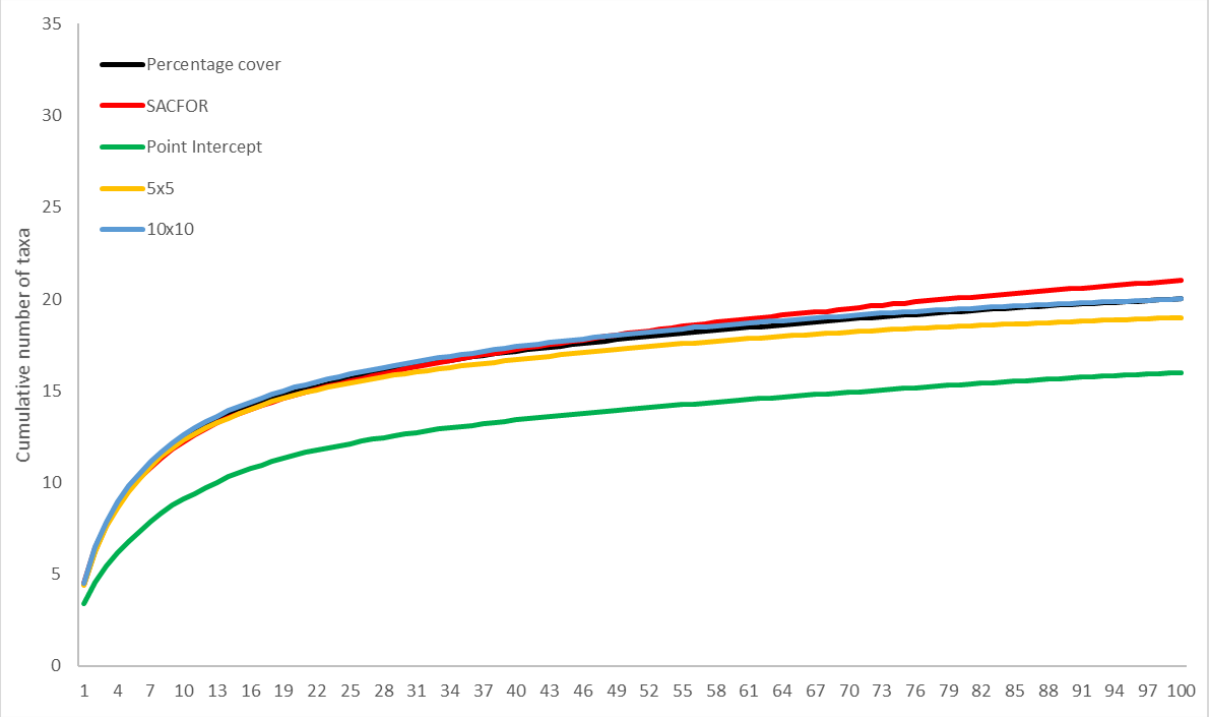


Figure 16. Cumulative number of ground cover taxa recorded from 100 images (in permutated order) by Analyst A, by extraction method. Methods: 10x10 Frequency (10x10); 5x5 Frequency (5x5).

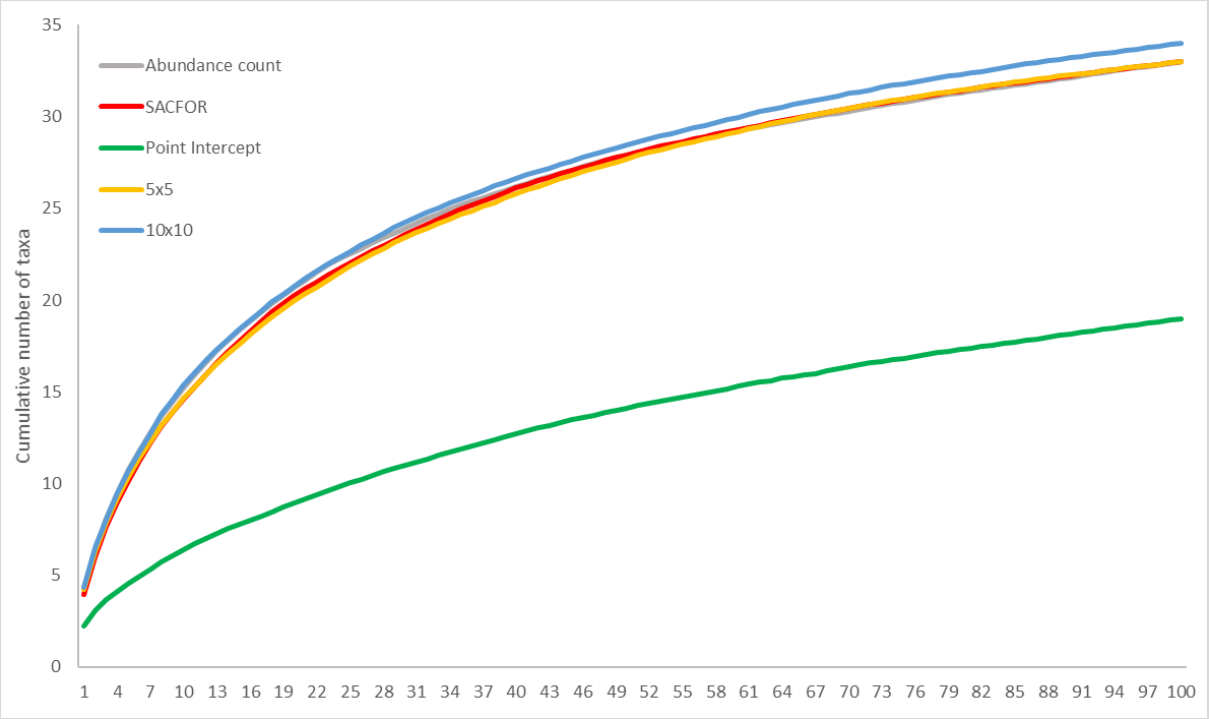


Figure 17. Cumulative number of erect/solitary taxa recorded from 100 images (in permutated order) by Analyst A, by extraction method.

Community impression

The ratio of Point Intercept to Percentage Cover mean values in Table 10 were much smaller (up or down), but also varied greatly. Taxa present as large patches (e.g. unidentified faunal crusts and Flustrid bryozoa) were relatively well represented by the average Point Intercept values, but any other taxa could be underestimated or overestimated, purely by chance. Representativeness increases with abundance, although Figure 18 shows that discrepancies between %c and Pi were generally between ±10-20% at all abundances >5%.

Results in Table 22 show that the impressions generated by each method were mostly dissimilar when data from the whole community are compared. However, when only the ground cover taxa are considered the impressions of the community made using the two frequency methods are just similar (t= 1.6, P = 0.0506, perms = 9949). This implies that either method could be used to generate a similar impression of ground cover taxa.

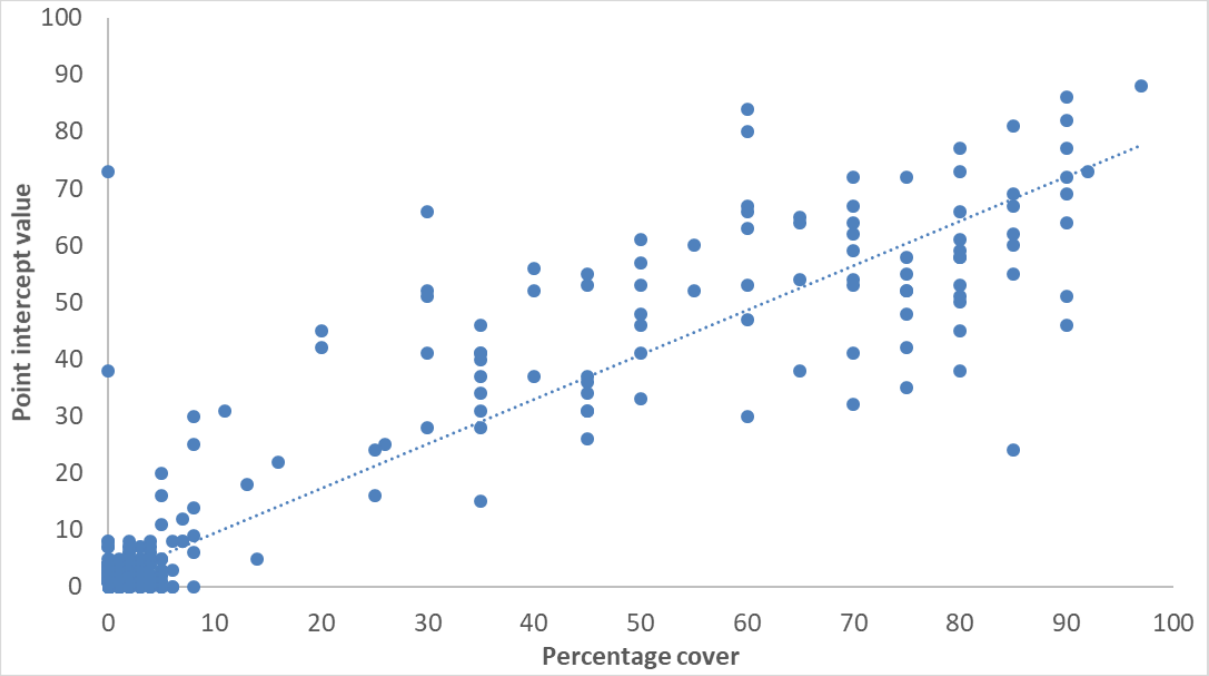


Figure 18. Relationship between Percentage Cover and Point Intercept values. Records by Analyst A from 100 images (496 points from 22 ground cover taxa). Note: the two points lying high on the point intercept axis (0 % cover) were for unidentified faunal crust, but no equivalent crust taxa were recorded as % cover in those images (i.e. missing records).

Table 22. Pairwise PERMANOVA test results for a one-factor PERMANOVA model, using the fixed factor using the fixed factors 'Method', 'Experience' and 'Analyst' (nested within Experience) 'Method', with 9999 unrestricted permutations. Tests were conducted on standardised, fourth root-transformed data of (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, frequency of occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover method data). Bold results indicate a significant effect (P ≤0.05).

Data set	Methods	t	P	Unique perms
a. **Ground cover taxa**	%c, As	6.7	**0.0001**	9945
	%c, F0	6.6	**0.0001**	9949
	%c, F5	7.6	**0.0001**	9954
	%c, Pi	3.3	**0.0001**	9953
	As, F0	3.0	**0.0001**	9949
	As, F5	3.2	**0.0001**	9949
	As, Pi	7.6	**0.0001**	9947
	F0, F5	1.6	0.0506	9949
	F0, Pi	7.5	**0.0001**	9951
	F5, Pi	8.4	**0.0001**	9935
b. **Erect/ solitary taxa**	As, Co	6.0	**0.0001**	9934
	As, F0	3.6	**0.0001**	9929
	As, F5	2.2	**0.0001**	9942
	As, Pi	8.8	**0.0001**	9941
	Co, F0	3.5	**0.0001**	9927
	Co, F5	4.9	**0.0001**	9941
	Co, Pi	3.8	**0.0001**	9931
	F0, F5	1.8	**0.0030**	9933

	F0, Pi	6.8	**0.0001**	9945
	F5, Pi	8.1	**0.0001**	9941
c. **All taxa** (but no count or cover data)	As, F0	3.6	**0.0001**	9927
	As, F5	3.0	**0.0001**	9928
	As, Pi	11.2	**0.0001**	9947
	F0, F5	1.9	**0.0027**	9937
	F0, Pi	10.4	**0.0001**	9943
	F5, Pi	11.4	**0.0001**	9937

Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As); 10x10 Frequency (F0); 5x5 Frequency (F5); Point Intercept (Pi).

Consistency

Tables 23 and 24 show range of inconsistencies of taxonomic identification among analysts working in one image. These tables also show the differences of identification made using the different methods by each analysts, something which was expected to be consistent. These results indicate that analysts may make a different classification of a taxon each time they look at an image.

Figure 22 shows how estimates for taxonomic richness vary in each image analysed by the six analysts using the six methods. The estimates for taxonomic richness made using the Percentage Cover, Abundance Count and Point Intercept seem to vary the least.

Tables 25 and 26 show the results of pairwise PERMANOVA tests. Both set of tests show that when data from the whole community are considered, the impressions of the community made by the senior and junior analysts differ no matter which of the methods is used (Table 26). They also show that, in the majority of cases, the community impressions made by the methods also differed (Table 25). Only in a few instances, were the impressions of community made by the frequency methods similar, suggesting the methods may substitute for one another.

Table 14. Examples of inconsistency of identification (and abundance) of Ophiuroids between analysts (ID) from a single image. (see Section 2.7.2)

ID	Method	Ophiuroidea	*Ophiothrix fragilis*	*Ophiocomina nigra*
A	SACFOR	6	5	5
A	Abundance Counts	236	19	4
A	10x10	92	29	9
A	5x5	25	17	5
A	Point intercept	15	1	
B	SACFOR	6		
B	Abundance Counts	150		
B	10x10	95		
B	5x5		25	
B	Point intercept	8	19	1
C	SACFOR		6	5
C	Abundance Counts		220	11
C	10x10		97	17
C	5x5		25	5
C	Point intercept		29	3
D	SACFOR	6		5
D	Abundance Counts	142		
D	10x10	96		
D	5x5	25		
D	Point intercept	17		
E	SACFOR	6	6	3
E	Abundance Counts	80	55	3
E	10x10	96	27	
E	5x5	25	10	
E	Point intercept	29		
F	SACFOR	6		
F	Abundance Counts		104	
F	10x10	97		
F	5x5	25		
F	Point intercept	30		

Table 24. Examples of inconsistency of identification (and abundance) of Bryozoa, Porifera and Faunal turf between analysts (ID) from a single image. (see Section 2.7.2)

ID	Method	Bryozoa	Bryozoa (erect)	Porifera	U_ faunal turf	U_ faunal turf (Porifera or Bryozoa)
A	Percentage cover				2	<1
A	SACFOR				2	1
A	10x10				53	3
A	5x5				25	
B	Percentage cover			<1	1	
B	SACFOR	4		3	2	
B	10x10				75	
B	5x5	2		2	21	
C	Percentage cover		<1		2	
C	SACFOR		4		2	
C	10x10		6		34	
C	5x5		6		13	
D	Percentage cover			6	1	
D	SACFOR			1	2	
D	10x10			59		
D	5x5			25	13	
E	Percentage cover			2	6	
E	SACFOR				4	
E	10x10				13	
E	5x5				5	
F	Percentage cover				15	
F	SACFOR				4	
F	10x10				35	
F	5x5				20	

Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As); 10x10 Frequency (F0); 5x5 Frequency (F5); Point Intercept (Pi).

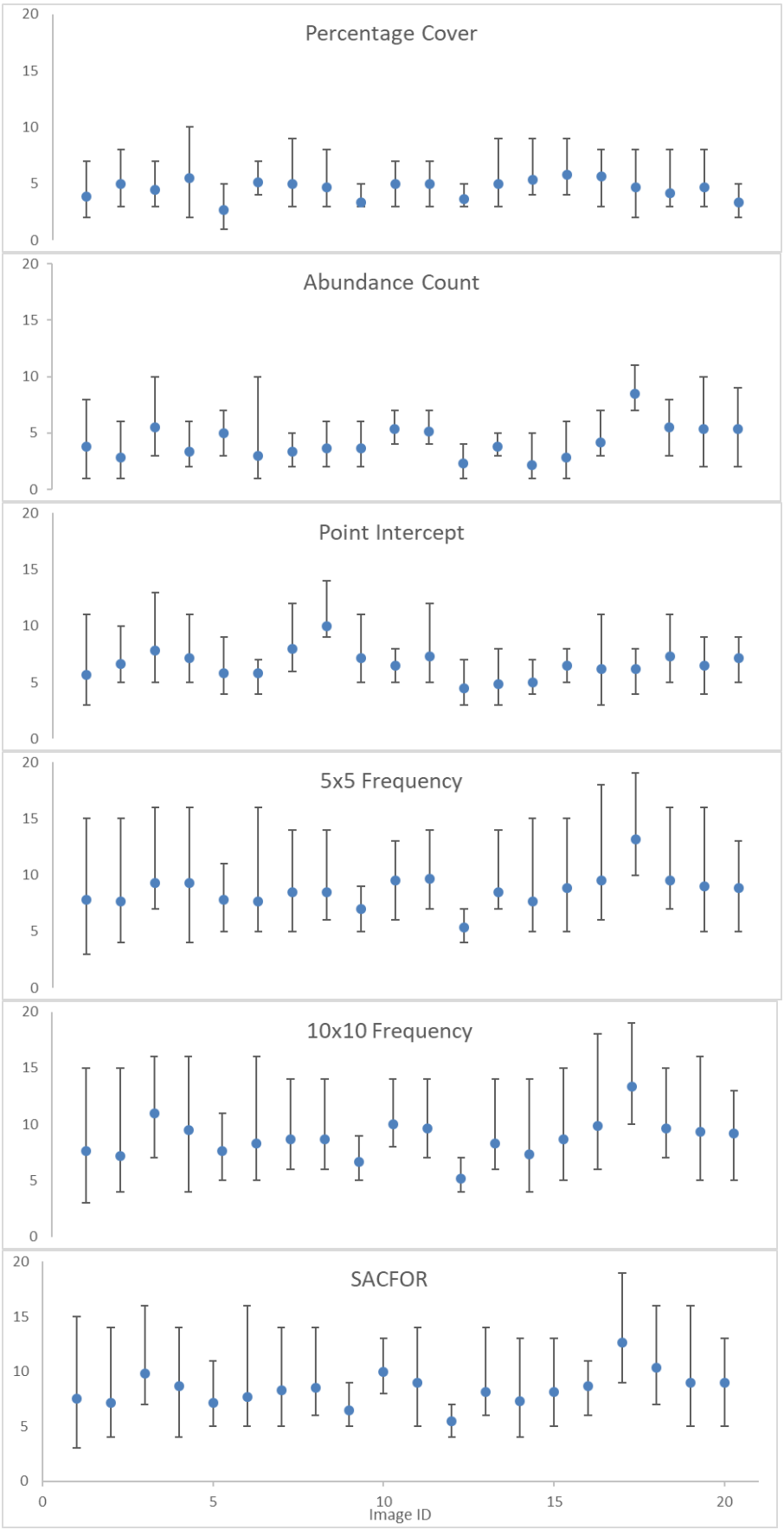


Figure 8. Variability (mean and range) in species richness between the Analysts for all 20 images and 6 extraction methods.

Table 25. Pairwise PERMANOVA test results for factor 'Experience' from test for 'Method' x 'Experience' in a three-factor PERMANOVA model, using the fixed using the fixed factors 'Method', 'Experience' and 'Analyst' (nested within Experience) with 9999 unrestricted permutations. Tests were conducted on standardised, fourth root-transformed data of (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, frequency of occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover method data). Bold results indicate a significant effect (P ≤0.05).

Data set	Experience	Methods	t	P	Unique perms
a. **Ground cover taxa** (no count data)	Senior	%c, As	5.2016	**0.0001**	9939
		%c, F0	5.5622	**0.0001**	9949
		%c, F5	6.4074	**0.0001**	9950
		%c, Pi	1.6849	**0.0115**	9954
		As, F0	2.4051	**0.0001**	9950
		As, F5	2.8738	**0.0001**	9961
		As, Pi	5.2563	**0.0001**	9934
		F0, F5	1.3449	0.1237	9940
		F0, Pi	5.6419	**0.0001**	9940
		F5, Pi	6.5903	**0.0001**	9953
	Junior	%c, As	4.1735	**0.0001**	9959
		%c, F0	3.9226	**0.0001**	9942
		%c, F5	4.6497	**0.0001**	9958
		%c, Pi	2.7882	**0.0001**	9960
		As, F0	2.4375	**0.0002**	9953
		As, F5	1.9916	**0.0052**	9963
		As, Pi	5.6839	**0.0001**	9946
		F0, F5	1.5404	0.0577	9962
		F0, Pi	5.4719	**0.0001**	9949
		F5, Pi	6.0521	**0.0001**	9941
b. **Erect/ solitary taxa** (no cover data)	Senior	As, Co	3.3152	**0.0001**	9924
		As, F0	2.2863	**0.0001**	9933
		As, F5	1.5604	**0.0144**	9936
		As, Pi	4.8619	**0.0001**	9947
		Co, F0	1.8634	**0.0016**	9933
		Co, F5	2.5327	**0.0001**	9917
		Co, Pi	2.8336	**0.0001**	9944
		F0, F5	1.1311	0.248	9935
		F0, Pi	3.8917	**0.0001**	9939
		F5, Pi	4.4897	**0.0001**	9939
	Junior	As, Co	3.6360	**0.0001**	9930
		As, F0	2.0756	**0.0002**	9931
		As, F5	1.3982	0.0577	9942
		As, Pi	5.6670	**0.0001**	9931
		Co, F0	2.2003	**0.0001**	9935
		Co, F5	2.9507	**0.0001**	9941
		Co, Pi	3.0275	**0.0001**	9934
		F0, F5	1.0620	0.3393	9944
		F0, Pi	4.6128	**0.0001**	9934
		F5, Pi	5.2610	**0.0001**	9951
	Senior	As, F0	2.6164	**0.0001**	9942

c. **All taxa** (but no count or cover data)		As, F5	2.4038	**0.0001**	9947
		As, Pi	7.3953	**0.0001**	9944
		F0, F5	1.3644	0.0552	9930
		F0, Pi	7.4786	**0.0001**	9933
		F5, Pi	8.1366	**0.0001**	9935
	Junior	As, F0	2.6760	**0.0001**	9956
		As, F5	1.8811	**0.0009**	9948
		As, Pi	8.5052	**0.0001**	9951
		F0, F5	1.5126	**0.0317**	9947
		F0, Pi	8.4631	**0.0001**	9954
		F5, Pi	8.8809	**0.0001**	9946

Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As); 10x10 Frequency (F0); 5x5 Frequency (F5); Point Intercept (Pi).

Table 26. Pairwise PERMANOVA test results for factor 'Method' from test for 'Method' x 'Experience' in a three-factor PERMANOVA model, using the fixed using the fixed factors 'Method', 'Experience' and 'Analyst' (nested within Experience) with 9999 unrestricted permutations. Tests were conducted on standardised, fourth root-transformed data of (a) ground cover taxa (no erect/solitary taxa), (b) erect/solitary taxa (no ground cover taxa) and (c) all taxa from the SACFOR, frequency of occurrence 10x10 and 5x5 grids and point intercept methods (without any abundance count and percentage cover method data). Bold results indicate a significant effect (P ≤0.05).

Data set	Method	Experience	t	P	Unique perms
a**. Ground cover taxa** (no count data)	%c	Senior, Junior	5.40	**0.0001**	9949
	As	Senior, Junior	5.84	**0.0001**	9950
	F0	Senior, Junior	6.36	**0.0001**	9949
	F5	Senior, Junior	6.62	**0.0001**	9953
	Pi	Senior, Junior	4.65	**0.0001**	9956
b. **Erect/ solitary taxa** (no cover data)	As	Senior, Junior	2.95	**0.0001**	9936
	Co	Senior, Junior	3.69	**0.0001**	9936
	F0	Senior, Junior	4.01	**0.0001**	9936
	F5	Senior, Junior	3.69	**0.0001**	9937
	Pi	Senior, Junior	6.56	**0.0001**	9952
c. **All taxa** (but no count or cover data)	As	Senior, Junior	4.42	**0.0001**	9934
	F0	Senior, Junior	5.60	**0.0001**	9937
	F5	Senior, Junior	5.64	**0.0001**	9939
	Pi	Senior, Junior	4.95	**0.0001**	9935

Extraction methods: Percentage Cover (%c); Abundance Count (Co); SACFOR (As); 10x10 Frequency (F0); 5x5 Frequency (F5); Point Intercept (Pi).

An alternative method to calculate multivariate analysis of consistency by making comparisons within images was also developed and is described here. The multivariate analyses in Section 3.9.3 are based on calculations of similarity between all samples, but the issue of consistency is concerned primarily with the similarity of samples from the same image (between the six analysts). In Section 3.9.3 the similarities between samples from different images include the inherent real differences between those images. PERMANOVA analyses can separate the variance between images and within images (and between extraction methods and analysts), but it could be simpler to remove the between image variance by not including the between image similarities in the analyses.

This does not appear to be achievable within PRIMER, so the Bray-Curtis resemblance matrices were exported into Excel and certain image similarities were deleted. Between the six analysts this gave 15 similarity scores within each image (for each extraction method), 300 across the 20 images. As the analyses in Section 3.9.3 had also shown the difficulties of direct comparison of extraction methods with standardised data, these analyses were carried out on unstandardised data, but with the following transformations to reduce the effect of high abundances:

Percentage Cover, Point Intercept, Frequently 5x5 and 10x10 data:	Square root transformation
Abundance counts:	Log(x+1) transformation
SACFOR:	No transformation

Resemblance matrices were then made separately for each extraction method, before exporting to Excel. The analyses were also made on different subsets of taxa: all taxa, ground cover taxa only and erect/solitary taxa only.

Consistency was assessed by calculating the average Bray Curtis similarity for each extraction method. As the Bray Curtis similarities are all on a scale from 0 (no similarity) to 100 (completely the same), they can be compared directly. Table 27 gives the results.

Table 27. Average Bray Curtis similarity between samples, by method, experience and subset of taxa.

	All analysts			Senior Analysts			Junior Analysts		
	All taxa	Ground cover taxa	Mobile & erect taxa	All taxa	Ground cover taxa	Erect & solitary taxa	All taxa	Ground cover taxa	Erect & solitary taxa
Percentage Cover	46.1	46.2	n/a	45.8	45.9	n/a	46.3	46.6	n/a
Abundance Counts	46.0	n/a	48.5	36.9	n/a	40.3	67.2	n/a	68.0
SACFOR	46.6	45.2	49.0	39.2	43.9	36.7	57.4	50.5	65.1
Point intercept	63.8	58.0	56.4	61.4	55.0	54.9	68.2	59.4	77.3
5 x 5 Frequency	49.1	48.9	49.5	50.3	55.0	40.3	54.5	48.1	67.8
10 x 10 Freqency	51.6	52.2	50.1	52.4	55.2	44.8	58.7	52.7	70.9

The greatest similarity, i.e. most consistent, is clearly shown in the Point Intercept data, followed by the 10x10 Frequency data. The lowest values are shown in the Percentage Cover, Abundance Counts and SACFOR data, but these rank differently in different datasets. One-way analysis of variance (all taxa, on untransformed data, following checks for normal distribution (see Appendix 3)) showed that there was a highly significant difference between the methods [$F = 39.7$, $p < 0.0001$]. A post hoc test (Tukey HSD) showed that the Point Intercept similarities were different to those of the other methods (Table 28). The 10x10 Frequency Bray-Curtis similarities were also significantly different to the other methods, all apart from the similarities among the 5x5 Frequency method. All other methods had similar levels of Bray-Curtis similarity.

Table 28. Results of a post hoc test (Tukey HSD) showing the significance of differences in Bray Curtis similarities from six different extraction methods (all taxa, all analysts). Significance levels: ns = not significant, * 0.05, ** 0.01, *** 0.001. Terms: Percentage Cover (%c); Abundance Count (Co); SACFOR (As), Point Intercept (Pi); 10x10 Frequency (F0); 5x5 Frequency (F5).

	%cover	SACFOR	Counts	10x10	5x5
SACFOR	ns				
Counts	ns	ns			
10x10	*	*	*		
5x5	ns	ns	ns	ns	
Point Int	***	***	***	***	***

Significance tests on datasets with other subsets of taxa had similar results, except that for the erect/solitary taxa the 10x10 Frequency similarities were not significantly different from those of the SACFOR, Abundance Counts and 5x5 Frequency methods.

Table 27 also shows that Junior analysts had higher average similarities than Senior analysts for most extraction methods. Analysis of variance showed that this difference was statistically highly significant for the erect/solitary taxa subset but that there was no difference with experience for the ground cover taxa subset. Inspection of the data behind these results suggests that the greater similarity between the junior analysts for the erect/solitary taxa is because they distinguished fewer of those taxa than the senior analysts (i.e. greater consistency in assigning to higher level taxa), while there was greater consistency between all analysts in records of the fewer, but relatively more abundant, ground cover taxa.

Ranked-based method comparisons

Table 29 shows the full range of rank scores for each method relative to the combined results from the Abundance Count and Percentage Cover methods, for each of the datasets tested under each data metric.

Table 29. Comparison of four extraction methods with combined results from Abundance Count and Percentage Cover, ranked for each of the analysed metrics and datasets tested. Ranks derived by authors judgement from the results in Sections 3.3 to 3.9 on a scale moving incrementally away from the combined results from Abundance Count and Percentage Cover (set at zero): if a method performed better than its rank scores are positive and if worse they are negative.

Data metric (data used)	**Dataset**	**Annotation method and data range**			
		SACFOR (1-6)	**Freq (5x5) (1-25)**	**Freq (10x10) (1-100)**	**Point (1-100)**
Precision (Analyst A)	Crust/Bed	3	1	3	1
	Crust_dis	4	3	2	1
	Indiv_lrg	1	-1	-2	-3
	Indiv_med	1	-2	-1	-3
	Indiv_ram	2	1	-1	-1
	Indiv_sml	2	3	1	-1
	Massive	3	2	1	-1
	Turf	3	2	1	-1
	Turf_indiv	2	3	1	-1
	Mean	**2.3**	**1.3**	**0.6**	**-1.0**
Power (Analyst A)	U. faunal crust	4	2	3	1
	U. faunal turf	4	3	2	1
	Hydrozoa	4	3	2	1

	Serpulidae	2	4	3	1
	Cupcorals	3	1	1	-1
	Ophiocomina sp.	3	2	1	-1
	Calliostoma sp.	-3	-2	-1	-4
	Richness	2	3	1	4
	Mean	**2.4**	**2.0**	**1.5**	**0.3**
Efficiency	Analyst A	1	1	0	0
	All analysts	1	1	0	0
	Mean	**1.0**	**1.0**	**0.0**	**0.0**
Taxonomic richness	Analyst A	0	0	0	-1
	All analysts	0	0	0	-1
	Mean	**0.0**	**0.0**	**0.0**	**-1.0**
Taxonomic accumulation	Analyst A	3	2	4	1
Community impression	Ground cover taxa	-4	-2	-3	-1
(Analyst A)	Erect/solitary taxa	-4	-2	-3	-1
	Mean	**-4**	**-2**	**-3**	**-1**
Consistency	Taxa richness	-2	-4	-3	-1
(All analysts)	Scleractinia	-1	1	0	-2
	Echinoidea	2	1	3	4
	Ophiuroidea	-2	-1	1	1
	Ophiocomina nigra	-4	-3	-2	-1
	Ophiothrix fragilis	1	4	3	2
	U. Faunal crust	0	1	3	2
	U. Faunal turf	1	2	4	2
	Flustridae	-1	1	3	2
	Ground cover taxa*	-2	-1	1	2
	Erect/solitary taxa*	-1	2	1	3
	Mean	**-0.8**	**0.3**	**1.3**	**1.3**
	Total of means	**3.9**	**4.6**	**4.3**	**-0.5**

Form is a categorisation of the life form of each taxon: Crust/bed= Encrusting/bed forming; Crust_dis= Encrusting and distinct; Individ_lrg= large individuals; Individ_med= medium sized individuals; Individ_sml= small individuals; Individ_ram= ramose individuals; Massive= massive form; Turf= turf forming; Turf_individ= ground covering individuals. Multivariate analyses indicated by *.