THE BIG PICTURE

Benthic Imagery Analysis Workshop

Workshop report
April 2019

JNCC

19-21 March 2019, Birmingham, UK

Workshop report

Contents

1. This report

The Big Picture Benthic Imagery Analysis Workshop 2019 was organised by JNCC, with facilitation and support provided by 3KQ. This report – produced by 3KQ, in partnership with JNCC – summarises the presentations, discussions and recommendations produced by all participants during the course of the workshop.

The three-day workshop brought together a wide group of stakeholders from across the marine monitoring community in the UK and beyond (Figure 1), to begin the development of a collaborative action plan to deliver future benthic imagery standards and quality assurance work in the UK.

Figure 1. Participants of the Big Picture Benthic Imagery Analysis Workshop 2019.

Much of the workshop involved small group discussions and interactions, with key outputs recorded on flipcharts. As such, this report acts as both an *aide memoire* for workshop participants and a reflection of the main themes and topics covered for other readers. It does not capture every conversation or aim to interpret all outputs.

The key outputs from the workshop:

- Discussions suggesting options and ways forward;
- The Plan Development Group – a small group of workshop participants representing a mix of interests, tasked with using the workshop outputs contained within this report to inform the development of a Benthic Imagery Analysis Action Plan, and seeking ongoing communication and input from the wider group of participants as the plan emerges;
- Formation of an unofficial 'reference group' of benthic imagery analysis experts from different organisations across the UK and beyond.

If you require any further explanation or context, please contact:

Dr Henk van Rein

Marine Monitoring and Evidence Manager, JNCC

marinemonitoring@jncc.gov.uk

2. Project and workshop background

Benthic Imagery Analysis Investigation project

As a part of Defra's commitment to carry out world-class, cost-effective marine monitoring, JNCC has been tasked with leading a project to investigate benthic imagery analysis standards for the UK post EU Exit.

The scope of the work includes benthic biodiversity and environmental baseline assessment, and monitoring conducted to meet the requirements of national and international biodiversity monitoring drivers. Among these biodiversity monitoring drivers are the Convention for the Protection of the Marine Environment of the North-East Atlantic (the 'OSPAR Convention'[1]), Defra's 25 Year Environment Plan (25YEP[2]) and the United Kingdom's Marine Strategy (Marine Strategy Part 2[3]).

Of the range of biodiversity components, the project is focused on benthic communities and how imagery may be best used for the purposes of their baseline assessment and monitoring. The project aims to:

- Bring the UK community of interest together to develop a shared understanding of challenges and opportunities surrounding benthic imaging;
- Develop a collaborative action plan to coherently deliver future benthic imagery standards and quality assurance work in the UK (the '*Benthic Imagery Analysis Action Plan*' for the UK).

Context – why do we need a plan?

Video and still imagery are versatile media used to collect information about the marine biological communities that exist on the surface of seabed habitats. Use of imagery focused on the seabed, or 'benthic imagery', has become widespread across the UK, where it is currently used in almost every marine habitat and at different depths.

It has been incorporated into a wide variety of sampling platforms, including:

- SCUBA divers – holding underwater cameras;
- Research vessels – cameras deployed in drop frames or integrated into remotely operated vehicles (ROVs);
- Semi-autonomous platforms – cameras integrated into baited remote underwater video frames (BRUVs);
- Autonomous underwater vehicles (AUVs).

In some habitats, such as those of rocky or biogenic characteristics, the use of benthic imagery represents the only suitable means of collecting information.

Despite the widespread application and versatility of benthic imagery, a significant amount of research and development work is needed to ensure that the data generated from it are of high quality, comparable and, therefore, allow the detection of meaningful changes in seabed communities over time. There are also limitations to the use of this technology. Benthic imagery is not effective in areas of high turbidity or where rough surface conditions

[1] OSPAR: https://www.ospar.org/convention
[2] Defra 25 YEP: https://www.gov.uk/government/publications/25-year-environment-plan
[3] UK Marine Strategy Part 2: https://www.gov.uk/government/publications/marine-strategy-part-two-uk-marine-monitoring-programmes

and strong currents preclude stable images. It is also not effective at collecting information about cryptic species, such as those that live under canopies of seaweeds or behind boulders.

There are ongoing, cutting-edge developments in the processing and analysis of benthic imagery, which may improve the quality and application of benthic imagery. However, it is essential that these developments do not happen in isolation, or for singular benefit, if we are to address the common benthic imagery issues. A collaborative approach between organisations and across benthic imagery disciplines should help to ensure the most efficient and robust standards can be developed and utilised across the UK.

This approach will result in the creation of a Benthic Imagery Analysis Action Plan for the UK – a strategic framework to carry out necessary improvements to a wide range of imagery analysis standards in the UK.

Under the governance of HBDSEG (the Healthy and Biologically Diverse Seas Evidence Group), the Action Plan will lay out a 'road map' for the UK to follow that will encompass current needs for improvement as well as those expected to emerge with increasing use of new technologies. A key purpose of the Action Plan is to streamline the improvements into coherent work flows, so that national resources are used more efficiently. There will also be greater opportunities for knowledge exchange and collaborative working via the Action Plan.

It is hoped that the outcomes of this Plan will improve the science of marine benthic monitoring, enable more effective use of resources and, ultimately, enhance the effect of conservation efforts in the UK and beyond.

The Big Picture workshop

Several existing groups have worked on aspects of benthic imagery analysis for many years. The Big Picture workshop was designed to build on this work by bringing the community of interest together to begin scoping and developing the Action Plan. Around 50 stakeholders, representing a wide range of marine monitoring interests and expertise in the UK, and beyond, came together for three days in Birmingham over 19th, 20th and 21st March 2019.

Participants shared knowledge and experience during the three days, building a common picture of the issues, opportunities, ideas and possible actions to better harness the potential of benthic imagery and analysis for benthic biodiversity baseline assessment and monitoring.

The outputs from the workshop will be collated and synthesised by a small task and finish group (the 'Plan Development Group'; PDG). The PDG will draft an Action Plan, to be consulted on before being adopted. The collaborative approach taken throughout the workshop and in the development of the Plan aims to maximise the potential benefits and effectiveness of the Plan, to ensure the best possible results for the marine monitoring and conservation of benthic habitats in the UK.

Workshop objectives

- Review the project context and proposed governance structures;
- Review the approach to developing the Plan;
- Build a picture of current activity and gaps;
- Develop a shared understanding of inherent challenges and opportunities across organisations and work streams;
- Prioritise key challenges and opportunities, and develop a way forward for each of these to feed into the development of the Plan;

- Review resources, roles and responsibilities to contribute to the development and delivery of the Plan;
- Agree ongoing engagement and communication activities;
- Confirm project governance structures.

Range of interests involved

There were 31 organisations directly involved in the workshop (Table 1). See Appendix 2 for a full list of attendees.

Table 1. List of organisations that directly contributed to the workshop.

Organisation	Type of organisation
Joint Nature Conservation Committee (JNCC) Natural Resources Wales (NRW) Natural England (NE) Marine and Fisheries Division, Department of Agriculture, Environment and Rural Affairs (DAERA)	Statutory Nature Conservation Bodies for the UK (note DAERA is also a devolved government department in Northern Ireland)
Agri-Food Biosciences Institute (AFBI) Biofar (Faroe Islands) Centre for Environment, Fisheries and Aquaculture Science (CEFAS) National Oceanography Centre (NOC) Marine Scotland Science (MSS)	Government- and research council-funded research institutes
Eastern Inshore Fisheries and Conservation Authority (IFCA) Environment Agency (EA) Scottish Environmental Protection Agency (SEPA)	Government-funded environmental protection agencies and fisheries authorities
Marine Biological Association of the UK (MBA) Scottish Association for Marine Science (SAMS) Wageningen Marine Research (Netherlands)	Independent marine research institutes
University of Edinburgh University of Plymouth University of Ulster	Universities with academic researchers working in various fields of marine benthic imagery analysis
APEM Aquatic Survey and Monitoring Limited (ASML) Bureau Waardenburg BV (Netherlands) Cloudbase Productions Crangon Limited	Environmental consultancies (organisations and independents)

Envision Mapping Limited Fugro GB Marine Limited Gardline Marine EcoSol Ocean Ecology Limited Seastar Survey Limited	
Department of Environmental, Farming and Rural Affairs (DEFRA)	Government departments

Governance and links

There are plans to embed the project within the Healthy and Biologically Diverse Seas Evidence Group (HBDSEG; Table 2; Figure 2). HBDSEG will then take charge of the project and will require periodic progress updates from organisations implementing the Benthic Imagery Analysis Action Plan. Another key organisation is the North-East Atlantic Marine Biological Analytical Quality Control (NMBAQC), which also reports to HBDSEG (Table 2). All new standards and protocols that arise from the Action Plan will be published via the NMBAQC, who have a central role in maintaining standards and advice across the UK.

Table 2. List of working groups that work to deliver evidence, standards and advice for marine benthic monitoring.

Organisation	Type of organisation and role
Healthy and Biologically Diverse Seas Evidence Group (HBDSEG) Healthy and Biologically Diverse Seas Evidence Group – Benthic Sub-Group (HBDSEG BSG)	The HBDSEG exists within the UK Marine Monitoring and Assessment Strategy (UKMMAS) and reports to the Monitoring and Assessment Reporting Group (MARG). The BSG, within HBDSEG, coordinates just benthic biodiversity monitoring and assessment work. The project aims to have the Benthic Imagery Analysis Action Plan adopted by and implemented via the BSG.
North-East Atlantic Marine Biological Analytical Quality Control (NMBAQC)	The NMBAQC coordinate the development and implementation of analytical standards for UKMMAS, and periodically report to HBDSEG. The NMBAQC role is to support the Benthic Imagery Analysis Action Plan and to endorse the standards that arise from it. Any new standards and protocols will be published via the NMBAQC.
Marine Protected Areas Survey Coordination and Evidence Delivery Group (MPAG)	MPAG coordinate delivery of marine protected area monitoring and reporting in England (Secretary of State Waters). Their role is to advise and support the project through representatives common to both MPAG and the workshop.
Marine Monitoring Group (MMG)	The MMG coordinate the research and development of marine monitoring methods and tools for use by the Statutory Nature Conservation Bodies for the UK. Their role is to advise and support the project through representatives common to both MMG and the workshop.

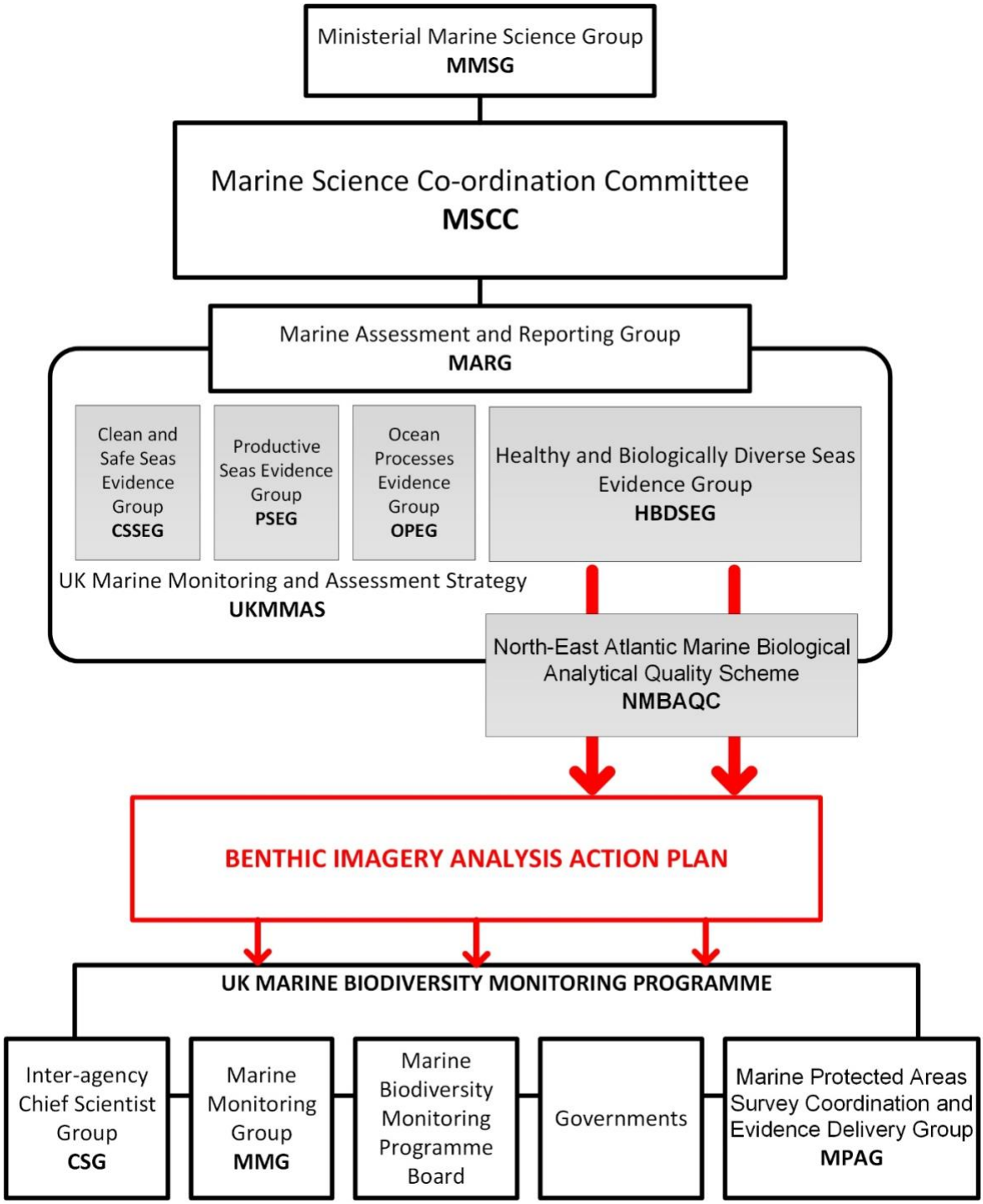


Figure 2. Proposed governance structure for the Benthic Imagery Analysis Action Plan (impacts indicated in red).

3. Workshop structure

Pre-workshop questionnaire

In advance of the workshop, participants and other interested stakeholders were sent a questionnaire, with the purpose of collating data on current purpose, acquisition methods, data extraction and quality control, as well as scoping the range of issues respondents wanted to discuss at the workshop. See Appendix 6 for a summary of the questionnaire findings.

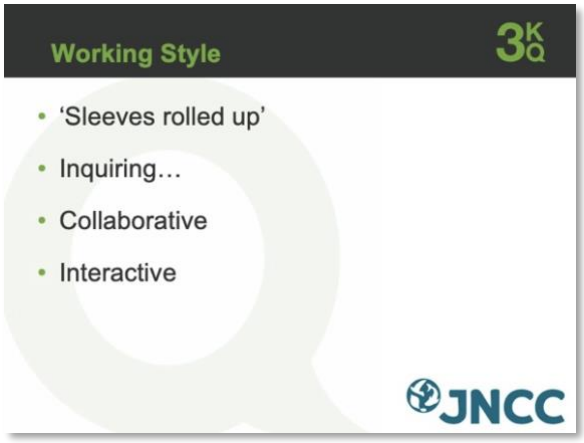
The questionnaire responses were used to directly shape the themes and topics for discussion at the workshop, resulting in a list of twelve themes:

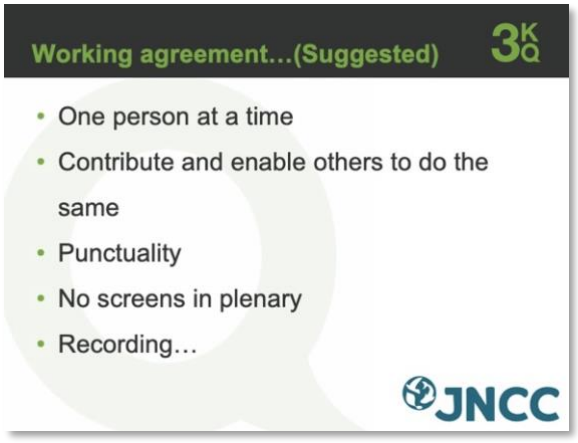
1. Need for improved levels of method standardisation and quality control;
2. Taxonomic identification;
3. Image annotation software, use of machine learning and improving cost and efficiency;
4. Development of standardised / specialised sampling approaches;
5. Image analysis contractual agreements and resources;
6. Biotope monitoring;
7. Sampling units;
8. Image reference collections;
9. Observer consistency;
10. Overall enumeration approaches;
11. Morphological classification systems;
12. Image analysis training.

These themes formed the basis for the bulk of the workshop activity, although participants were able to suggest other themes issues for discussion. Some of these, along with a 13th topic emerging from the questionnaire responses (data sharing and management), were discussed during the final working group sessions on day three – see Section 6 for details.

Style of working

The workshop was designed to be highly interactive and collaborative, with most of the work taking place in the form of small group discussions around the themes listed above. Participants were encouraged to enter into the workshop with a spirit of enquiry and a willingness to listen and learn, as well as to contribute their own knowledge and ideas.





Agenda and key workshop elements

The workshop was designed to enable participants to share knowledge and expand thinking around various themes. A range of themes were examined, focusing in on the key issues and connections between these. Participants were encouraged to explore options for the way forward with the aim of producing useful outputs to inform the work of the Plan Development Group.

The process was kept flexible and was reviewed and adapted throughout the workshop, to respond to participant feedback and emerging outputs.

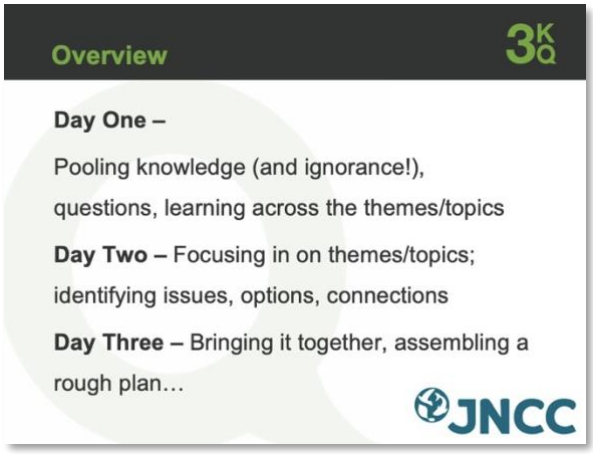


Table 3. Broad sessions and structure of the three-day workshop.

DAY 1: 1030-1630	DAY 2: 0900-1600	DAY 3: 0900-1400
Welcome, practicalities and introductions	Welcome and feedback from day 1	Welcome and feedback from day 2
Fairground: Scoping the themes – identifying questions	Identifying issues	Finalising in-depth work on themes, including reviewing and challenging emerging outputs
Presentations on key topics	In-depth work on themes: identifying options and connections	Picking up new and final themes for discussion
Marketplace: Sharing knowledge on the themes to answer questions	Plan Development Group meet up	Way forward and evaluation
Plan Development Group meet up	Shared dinner and networking	Plan Development Group meet up

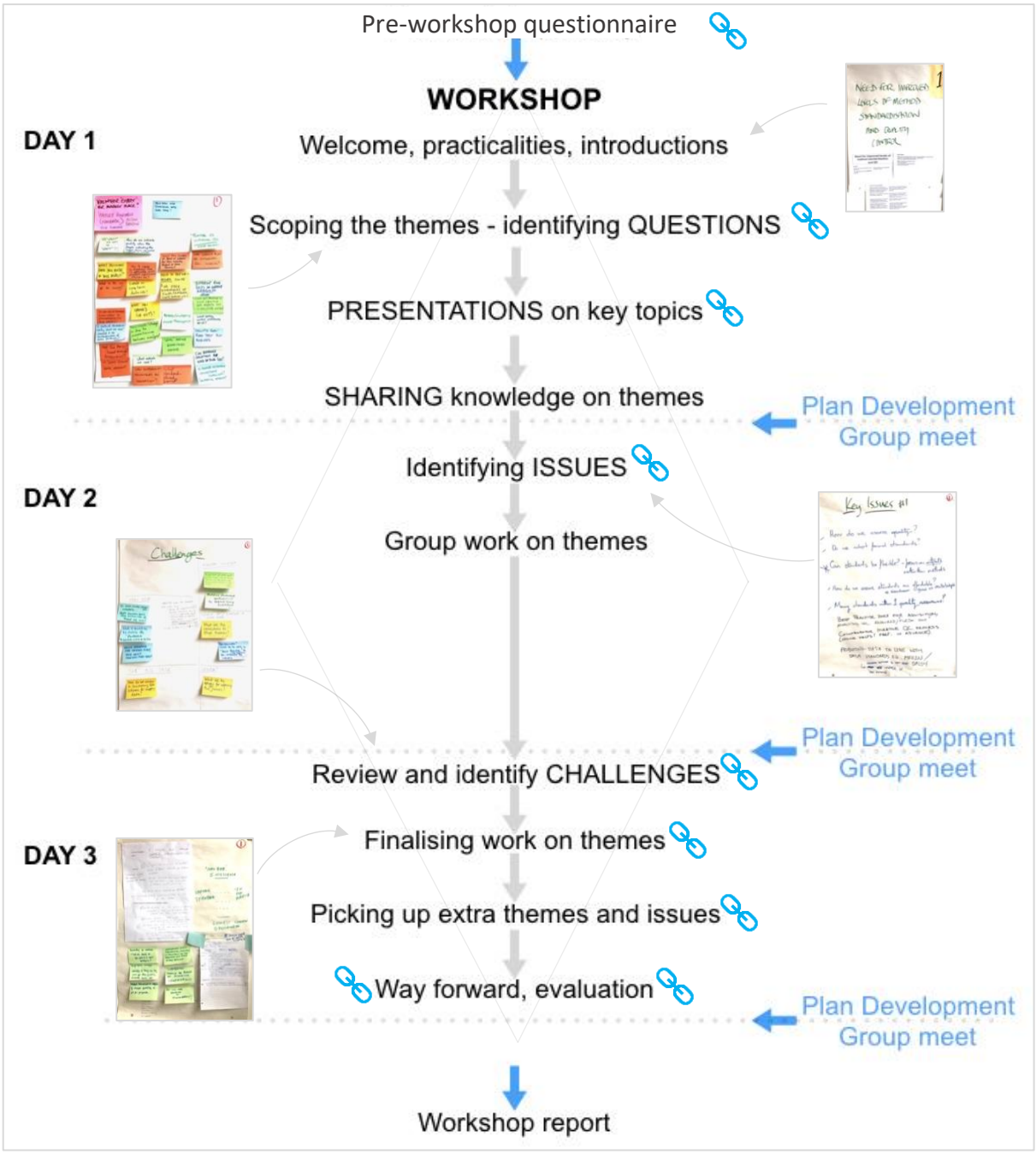


Figure 3. Workshop processes and work flows.

Evaluation

Participants provided formative feedback at the end of days 1 and 2, which helped to shape the following day's process in each case. At the end of the final day, participants completed an evaluation form – see Appendix 3 for the compiled results from these forms.

4. Presentations

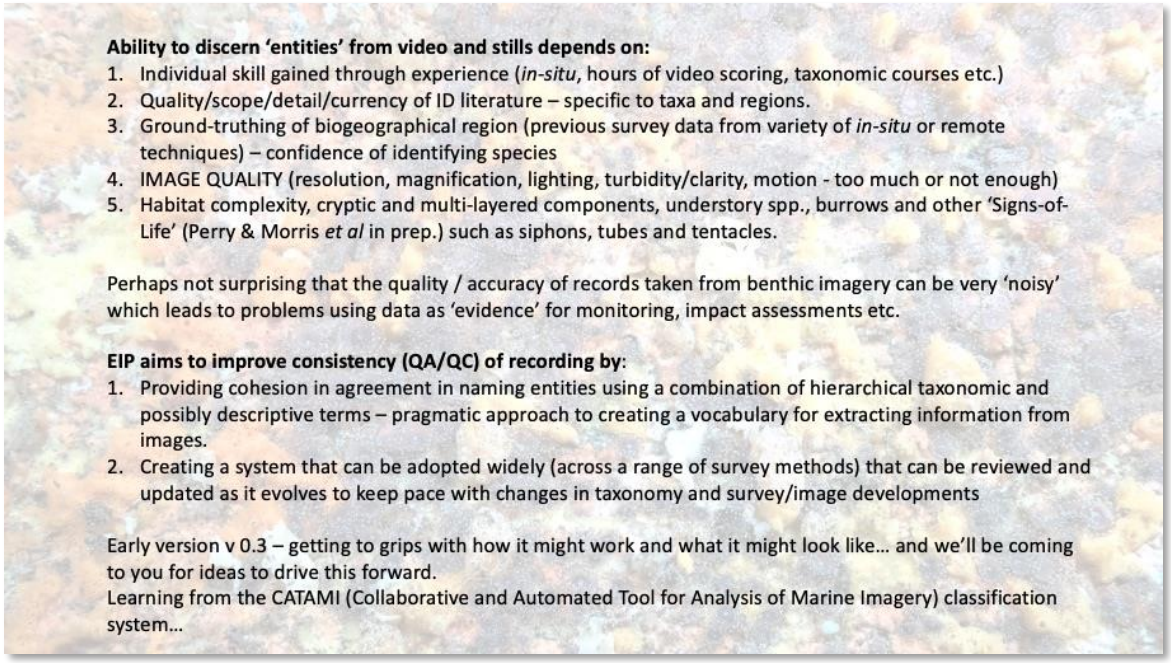
During day 1 of the workshop, four participants shared presentations with the group. These presentations were designed to describe recent or current work on specific aspects of benthic imagery analysis and highlight some of the key issues. This section shares relevant slides and provides a summary of the accompanying talks.

Presentation 1: Epifauna Identification Protocol (EIP) – Rohan Holt (Cloudbase Productions Ltd.)

"EIP is about standardisation of language, consistent use of terms, and what we can do with the visual image. At the moment, taxonomic language is used most commonly. There are various other ways, but language is not standardised in one place yet.

Even with a poor image there are a few things that might be identified.

Under improved resolution you can see what is discriminable – including species you can name to species level with about 99% confidence. Some things are less easy to identify, e.g. unidentified pink sponges. And you might have old literature that leads to species being called the wrong name. Standardising the species you have is important, otherwise you end up with pseudospecies and noisy data.

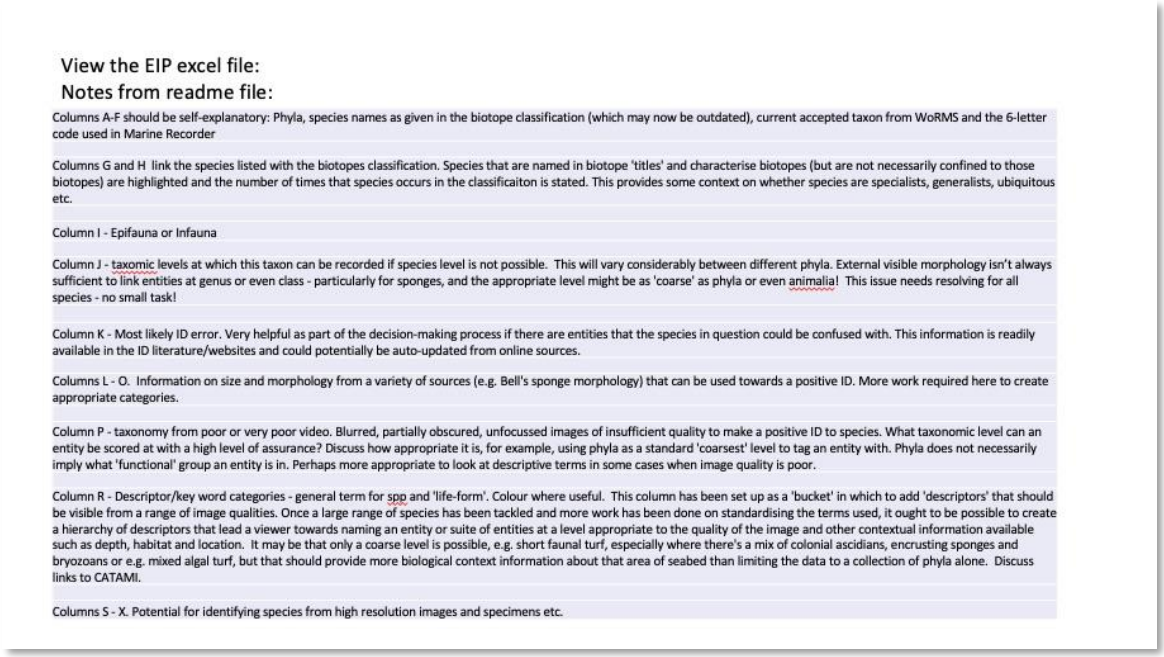


The ability to discern entities from video and stills depends on a number of things:

- Individual skill – varies depending on experience;
- Quality / scope / detail / currency of identification literature – the literature can be specific to taxa and regions. And as identification guides improve, there are surges in apparently 'new' things being found;
- Ground-truthing – this is very important to whether and what can be discriminated, and can enable identification even with blurry images, based on previous surveys;
- Image quality – a range of aspects such as lighting, magnification, turbidity and clarity of water all contribute. Motion can be useful sometimes, by resolving in three dimensions;
- Habitat complexity – this is a big factor when looking at the quality of an image. "Signs of life" is due to be published soon – looking at things like burrow shapes.

Data can be noisy. The key question is whether data is reliable enough for monitoring. EIP aims to improve the consistency of recording. Feedback on this will be really useful. EIP is currently up to version 0.3 at the moment. This also seems to be having a convergence with CATAMI identification.

The EIP spreadsheet includes a species list (around 800). It captures the scientific name (using WoRMS database for the latest names) and a suite of information such as the level of species names, where to stop, what you think you can derive from poor quality images, etc. Current work has added a lot of morphology descriptors and key terms that might be appropriate to use to try to discriminate in poor images."



View the EIP excel file:

Notes from readme file:

Columns A-F should be self-explanatory: Phyla, species names as given in the biotope classification (which may now be outdated), current accepted taxon from WoRMS and the 6-letter code used in Marine Recorder

Columns G and H link the species listed with the biotopes classification. Species that are named in biotope 'titles' and characterise biotopes (but are not necessarily confined to those biotopes) are highlighted and the number of times that species occurs in the classificaiton is stated. This provides some context on whether species are specialists, generalists, ubiquitous etc.

Column I - Epifauna or Infauna

Column J - taxomic levels at which this taxon can be recorded if species level is not possible. This will vary considerably between different phyla. External visible morphology isn't always sufficient to link entities at genus or even class - particularly for sponges, and the appropriate level might be as 'coarse' as phyla or even animalia! This issue needs resolving for all species - no small task!

Column K - Most likely ID error. Very helpful as part of the decision-making process if there are entities that the species in question could be confused with. This information is readily available in the ID literature/websites and could potentially be auto-updated from online sources.

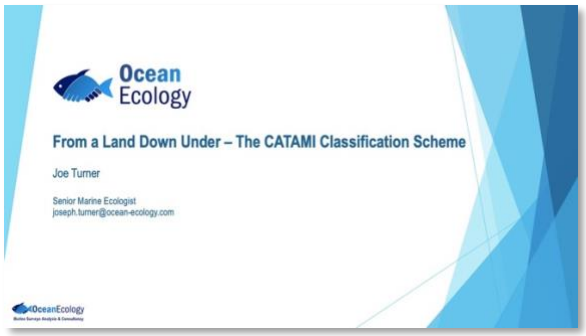
Columns L - O. Information on size and morphology from a variety of sources (e.g. Bell's sponge morphology) that can be used towards a positive ID. More work required here to create appropriate categories.

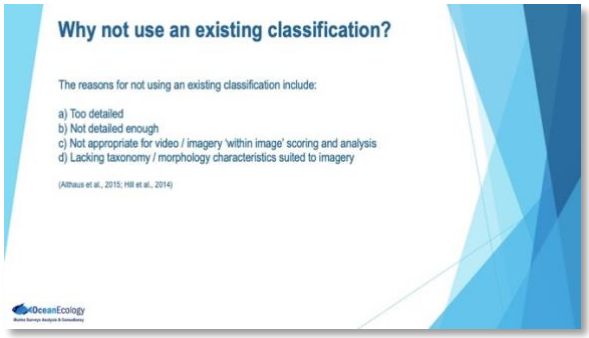
Column P - taxonomy from poor or very poor video. Blurred, partially obscured, unfocussed images of insufficient quality to make a positive ID to species. What taxonomic level can an entity be scored at with a high level of assurance? Discuss how appropriate it is, for example, using phyla as a standard 'coarsest' level to tag an entity with. Phyla does not necessarily imply what 'functional' group an entity is in. Perhaps more appropriate to look at descriptive terms in some cases when image quality is poor.

Column R - Descriptor/key word categories - general term for spp and 'life-form'. Colour where useful. This column has been set up as a 'bucket' in which to add 'descriptors' that should be visible from a range of image qualities. Once a large range of species has been tackled and more work has been done on standardising the terms used, it ought to be possible to create a hierarchy of descriptors that lead a viewer towards naming an entity or suite of entities at a level appropriate to the quality of the image and other contextual information available such as depth, habitat and location. It may be that only a coarse level is possible, e.g. short faunal turf, especially where there's a mix of colonial ascidians, encrusting sponges and bryozoans or e.g. mixed algal turf, but that should provide more biological context information about that area of seabed than limiting the data to a collection of phyla alone. Discuss links to CATAMI.

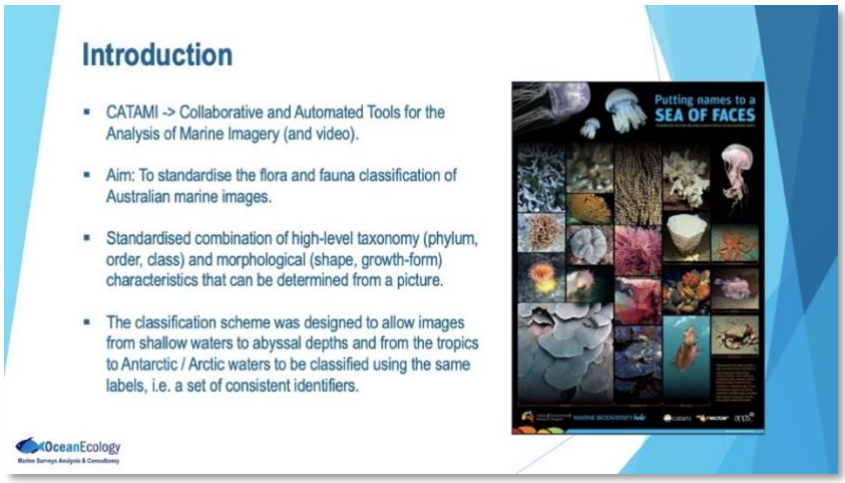
Columns S - X. Potential for identifying species from high resolution images and specimens etc.

Presentation 2: The CATAMI Classification Scheme – Joe Turner (Ocean Ecology Limited)





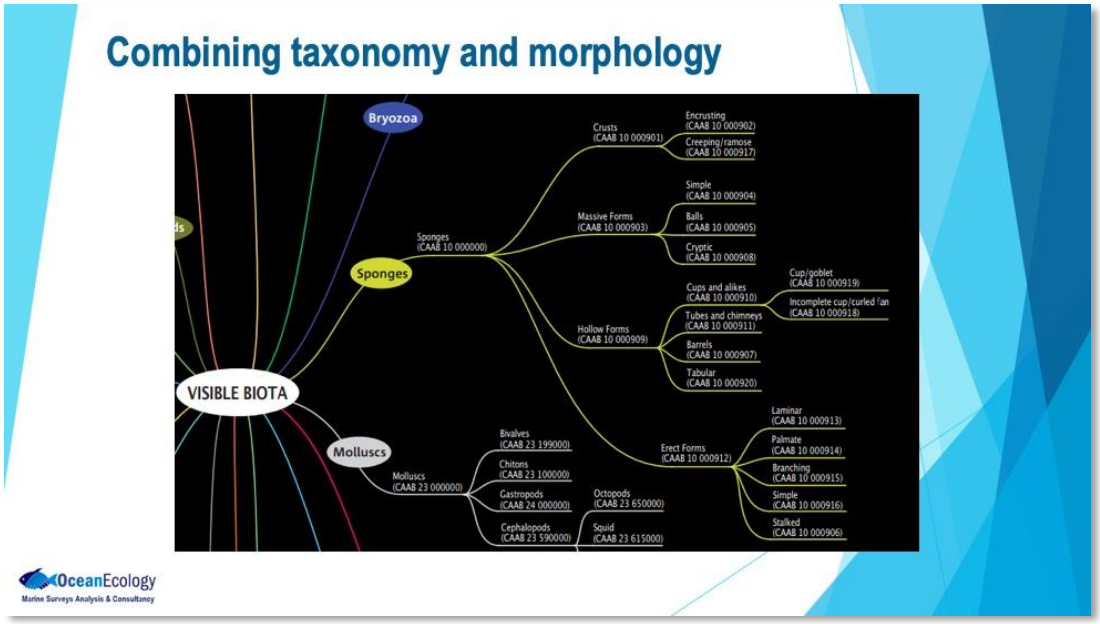
"A lot of people don't use classification schemes for a variety of reasons. They might find a particular scheme gives too much or too little detail, that it isn't appropriate for video or imagery analysis, or that it lacks the ability to record taxonomic or morphological characteristics.



CATAMI was developed in Australia to standardise analysis of all marine imagery across all environments.

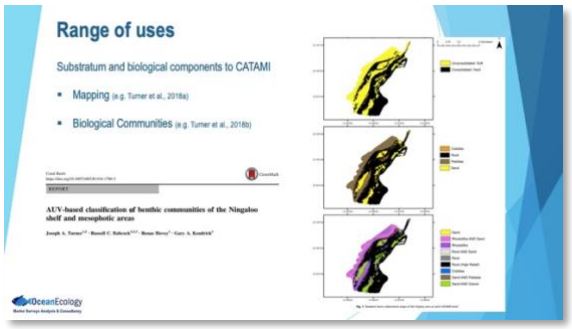
It has a relatively high level of taxonomic information, but combines this with morphology.

Here's an example of the sponge branch. It gives a good level of detail for describing assemblages.



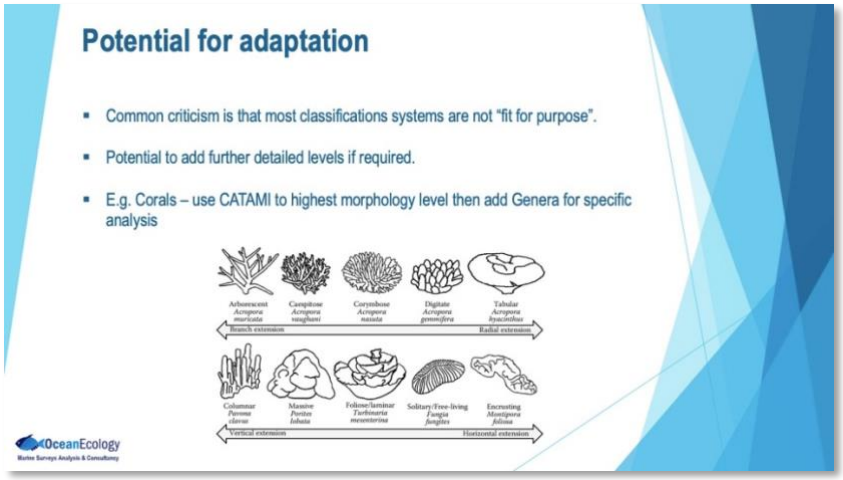
Current taxonomic classification approaches can make it hard to compare between studies. A more consistent approach can help, especially with monitoring change.

CATAMI classification tends to be relatively consistent between different observers. It can be used for mapping (e.g. PhD mapping project). It also enables more detailed analysis of biological communities, which can be correlated to environmental variables, using multivariable statistics.

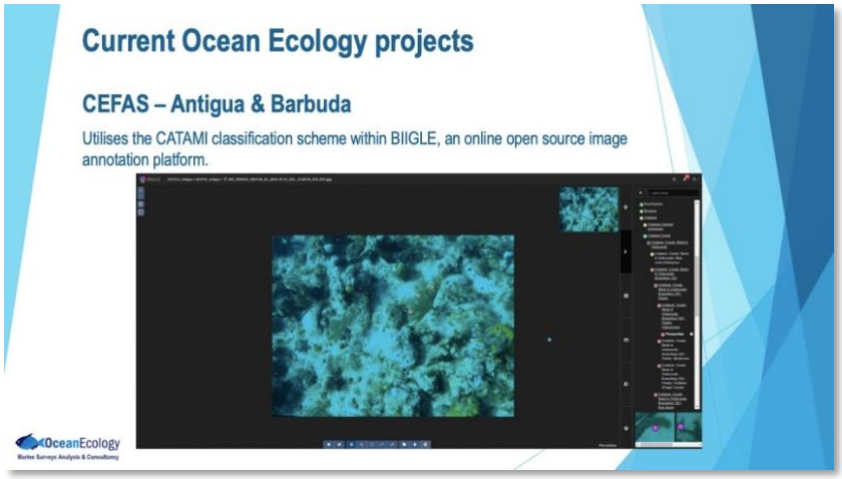


A full classification guide is available, with examples of substratum types and taxa and species information.

A good thing about CATAMI is that you can adapt it to include certain indicator species or add levels of detail (e.g. regarding coral assemblages) should the resolution of your imagery allow for certain and reliable identification. Extra categories, such as genera, species or morphospecies (e.g. Branching Sponge – Red), can be added for your particular study while still maintaining the underlying CATAMI structure.



Currently, Ocean Ecology have started working with Cefas on this, using CATAMI as a label tree within BIIGLE. This project has a tropical seabed focus but has been working very well and seems to have reduced the amount of time required to analyse the imagery when compared to the more traditional methods.

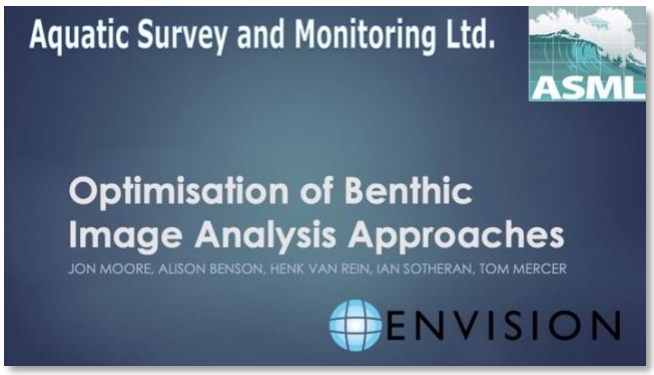


It can also be used in the UK. Temperate waters are already covered (there are some waters around Tasmania very similar to the UK). A UK version of the guidance document would need to be produced, highlighting any major classes that need adding to the current structure of the classification system. One approach would be to use EUNIS correlations for substratum, although it should be noted that EUNIS and CATAMI have very different purposes.

Rather than reinventing the wheel, it is possible to look at how this could fit into the UK without needing too much adaptation. It could be an excellent tool for improving consistency. See catami.github.io for more detail."

Presentation 3: Optimisation of Benthic Image Analysis Approaches – Jon Moore (Aquatic Survey and Monitoring Ltd.)

"This study is all about the enumeration of images – the methods of capturing the data we want from the images.



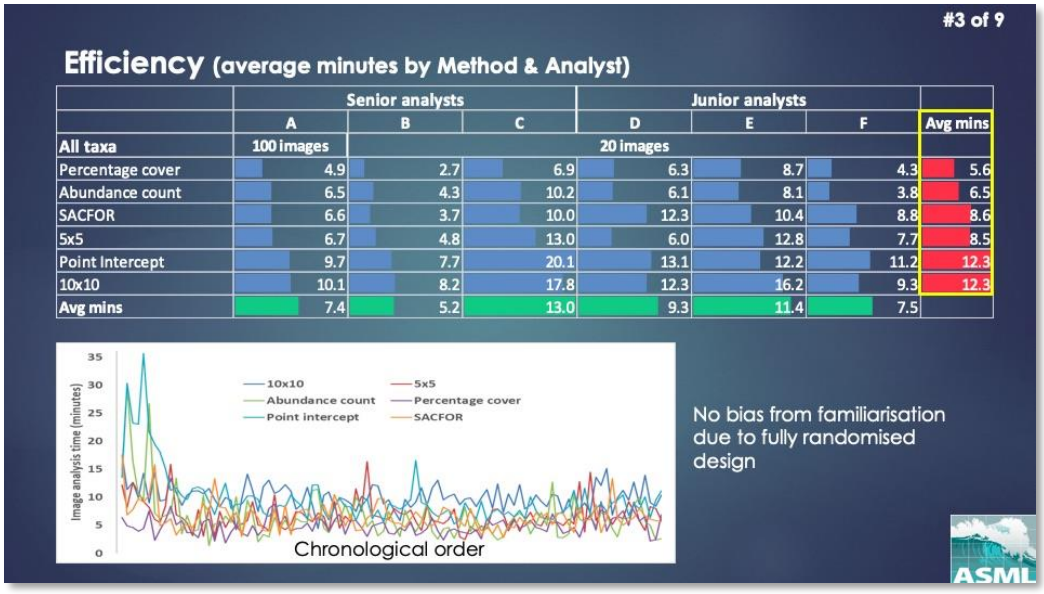
For the study, 100 images and six extraction methods were used:

- Percentage cover;
- Counts;
- SACFOR abundance scales;
- Point intercept (100 points per image);
- Cell frequencies (5x5);
- Cell frequencies (10x10).

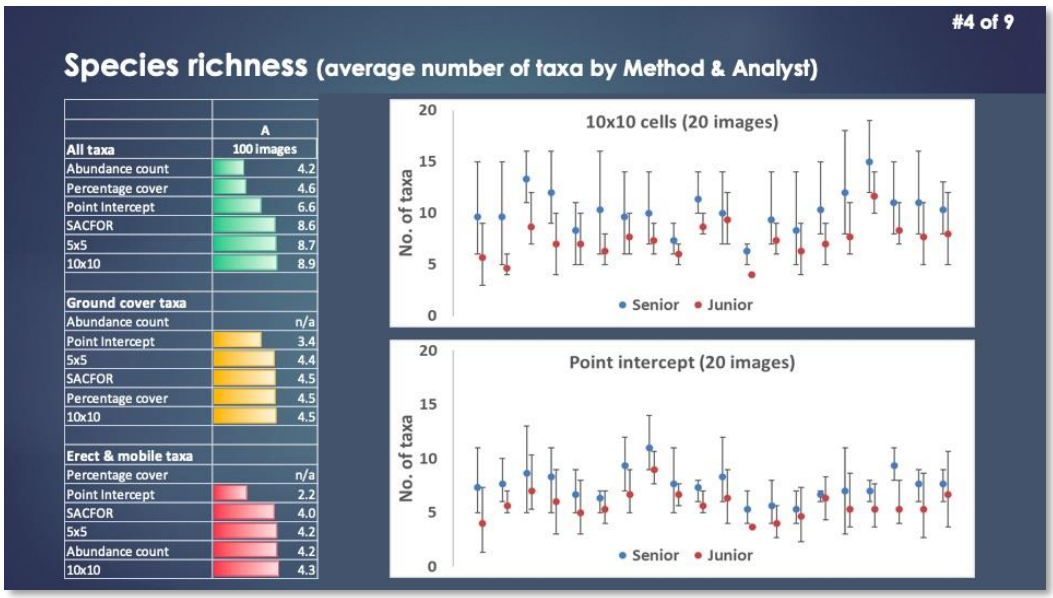
There were six analysts involved (three junior, three experienced). All 100 images were analysed by one experienced analyst, with 20 analysed by all of the others.

Overall, 417 different taxa were identified, because there was not a defined checklist. This reflected a realistic image analysis situation, with no preconceived ideas, essentially starting from scratch. That meant that initial consistency in taxonomic identification was poor, but this was improved considerably by truncation: the process of 'cleaning' the data by filtering and aggregating taxa, identifying where different descriptions were being used to identify the same taxon.

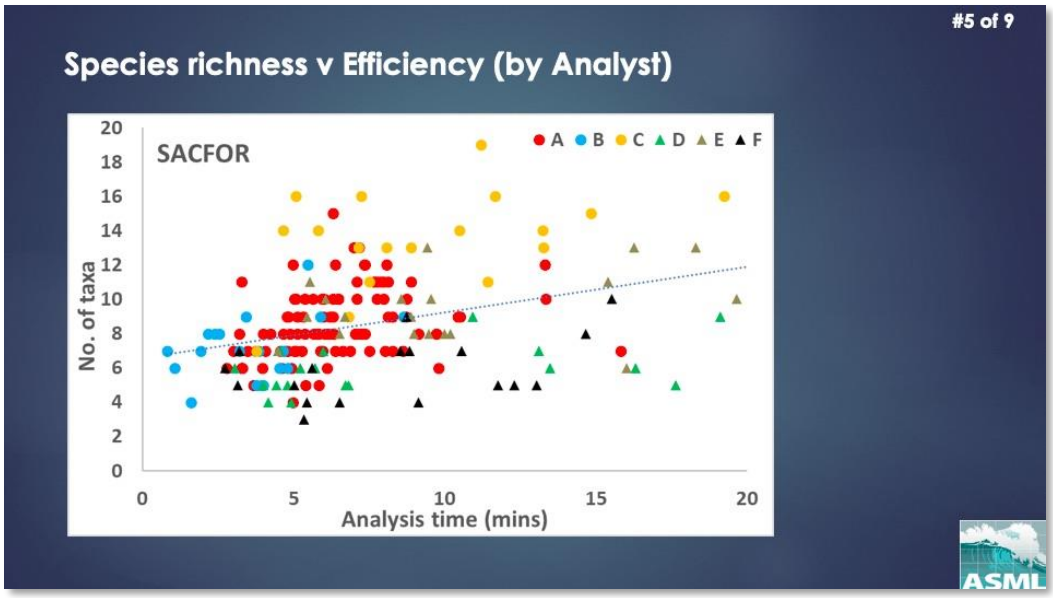
When the dataset was truncated, comparisons of data extracted by the different methods could be made. In the efficiency slide, the graph at the bottom is to prove that randomised design meant there was no familiarisation process (i.e. the number of minutes didn't reduce for any particular method). The average number of minutes by method and analyst is shown. Percentage cover and abundance count were the fastest – these can both be done in the time it takes for either point intercept or 10x10 cell frequency.



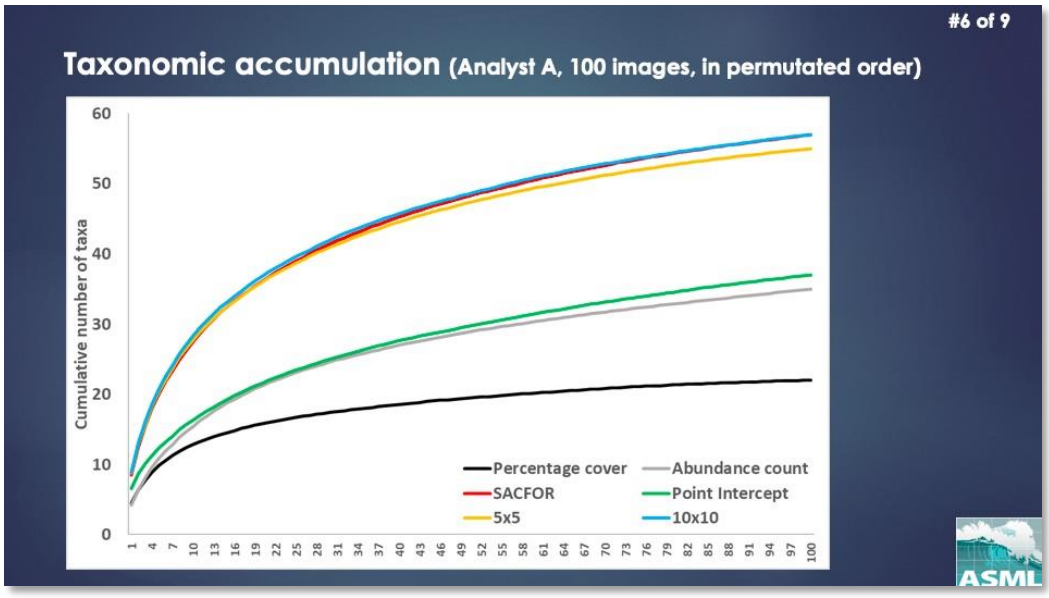
With species richness, all the taxa were investigated first, then only ground cover taxa and then only erect and mobile taxa. For ground cover taxa (analyst A only), the lowest number of species is by point intercept. The others have no significant difference. The same for erect and mobile taxa. Bringing in the other analysts, there is huge variability in the number of taxa recoded. Junior analysts were not recording as many as senior analysts, but there is still huge variation even among senior analysts. Point intercept is capturing fewer species per image and is maybe a little more consistent, but not a lot.



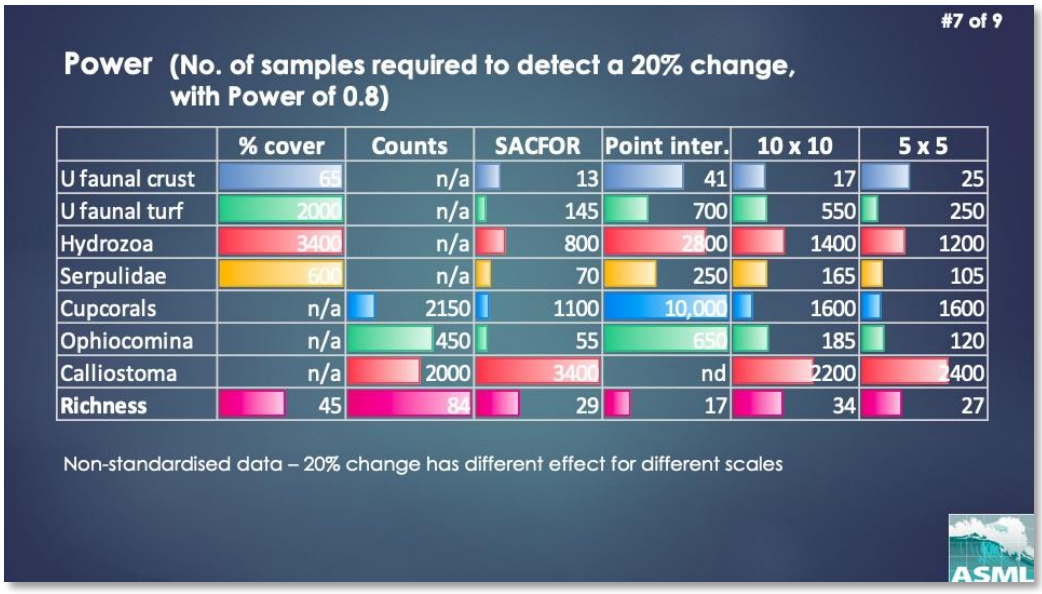
Putting richness and efficiency together, analysis time increases. For example, analyst C was generally capturing more species but taking more time. Analyst B was doing the opposite. They are both senior analysts but had different approaches in terms of the time they were willing to put into an image.



Taxonomic accumulation shows that the percentage cover line levels off much more quickly than for the other methods, so perhaps there is a potential advantage there. Although it might be to do with that specific location the samples were taken from.

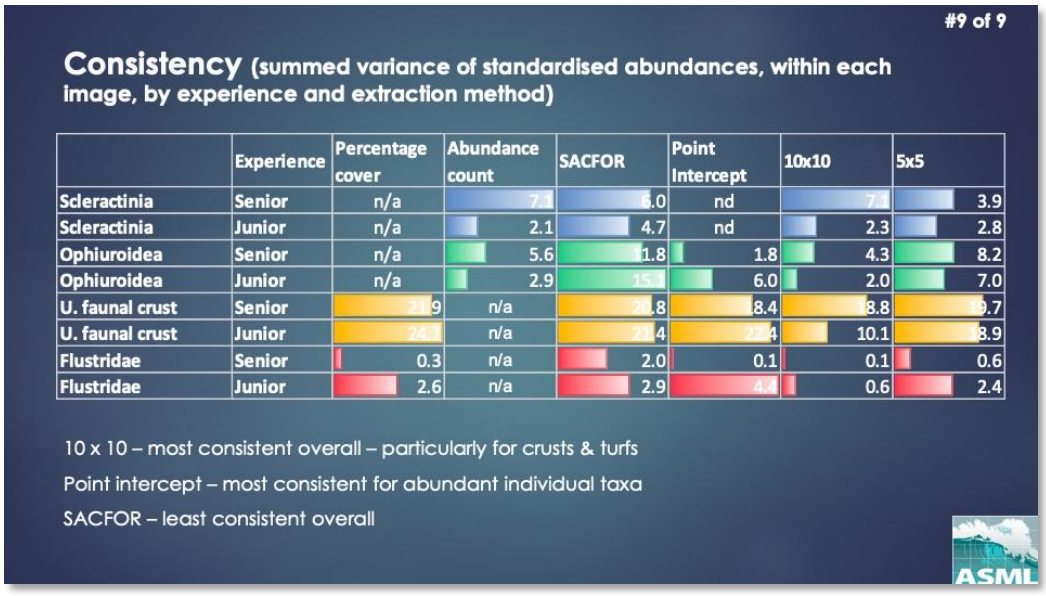


The robustness of the data was next analysed in terms of power (which was defined as the number of samples required to detect a 20% change, with significance of 5%, with a power of 0.8). There was no clear advantage to any method. A 20% change is different when applied to different methods as they are on different scales (e.g. it doesn't make much sense for SACFOR). For some species, a prohibitively large number of images is required to detect a 20% change in their abundance in the data.



	% cover	Counts	SACFOR	Point inter.	10 x 10	5 x 5
U faunal crust	65	n/a	13	41	17	25
U faunal turf	2000	n/a	145	700	550	250
Hydrozoa	3400	n/a	800	2800	1400	1200
Serpulidae	600	n/a	70	250	165	105
Cupcorals	n/a	2150	1100	10,000	1600	1600
Ophiocomina	n/a	450	55	650	185	120
Calliostoma	n/a	2000	3400	nd	2200	2400
Richness	45	84	29	17	34	27

There was a lot of inconsistency between analysts, partly because a standardised checklist was not used. The 10x10 frequency grid method is probably the most consistent overall, especially for crusts and turfs. SACFOR is the least consistent overall. Point intercept is the most consistent for abundant taxa but not for rarer ones. It seems every species is different in terms of which method is better for it.



#9 of 9

Consistency (summed variance of standardised abundances, within each image, by experience and extraction method)

	Experience	Percentage cover	Abundance count	SACFOR	Point Intercept	10x10	5x5
Scleractinia	Senior	n/a	7.1	6.0	nd	7.1	3.9
Scleractinia	Junior	n/a	2.1	4.7	nd	2.3	2.8
Ophiuroidea	Senior	n/a	5.6	11.8	1.8	4.3	8.2
Ophiuroidea	Junior	n/a	2.9	15.1	6.0	2.0	7.0
U. faunal crust	Senior	21.9	n/a	21.8	18.4	18.8	19.7
U. faunal crust	Junior	24.1	n/a	21.4	22.4	10.1	18.9
Flustridae	Senior	0.3	n/a	2.0	0.1	0.1	0.6
Flustridae	Junior	2.6	n/a	2.9	4.7	0.6	2.4

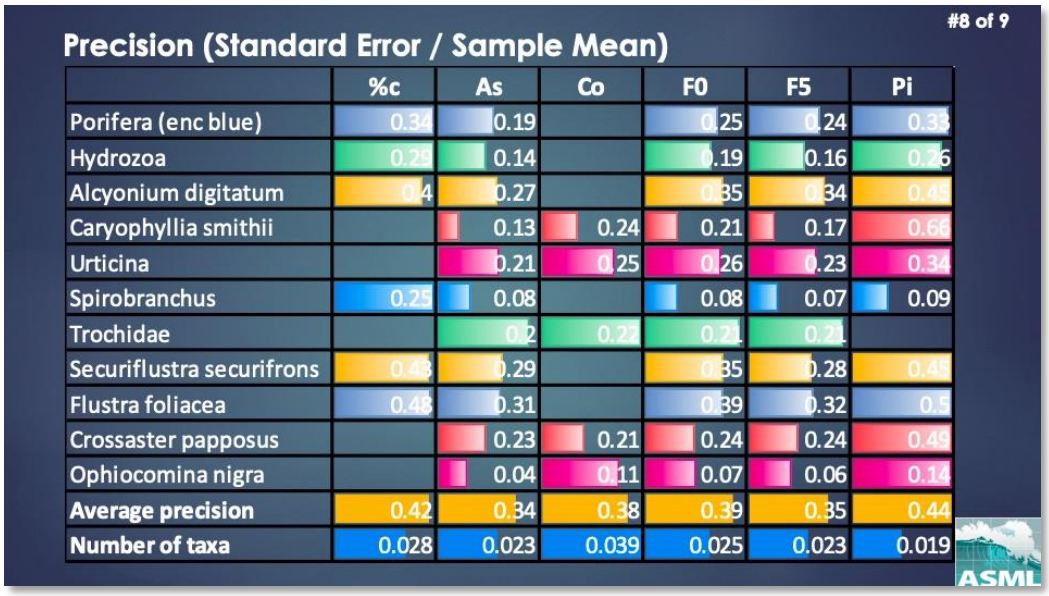
10 x 10 – most consistent overall – particularly for crusts & turfs

Point intercept – most consistent for abundant individual taxa

SACFOR – least consistent overall

ASML

Finally, precision relates to the variability within the data set for each species. Data with a high level of precision will vary less and be better suited to monitoring work as real changes over time may be more easily observed. However, there is not consistent pattern regarding one method over another. The least variability is perhaps with point intercept, but it is difficult to interpret."



#8 of 9

Precision (Standard Error / Sample Mean)

	%c	As	Co	F0	F5	Pi
Porifera (enc blue)	0.34	0.19		0.25	0.24	0.33
Hydrozoa	0.25	0.14		0.19	0.16	0.26
Alcyonium digitatum	0.4	0.27		0.35	0.34	0.45
Caryophyllia smithii		0.13	0.24	0.21	0.17	0.66
Urticina		0.21	0.25	0.26	0.23	0.34
Spirobranchus	0.25	0.08		0.08	0.07	0.09
Trochidae		0.2	0.23	0.21	0.21	
Securiflustra securifrons	0.43	0.29		0.35	0.28	0.45
Flustra foliacea	0.43	0.31		0.39	0.32	0.5
Crossaster papposus		0.23	0.21	0.24	0.24	0.49
Ophiocomina nigra		0.04	0.11	0.07	0.06	0.14
Average precision	0.42	0.34	0.38	0.39	0.35	0.44
Number of taxa	0.028	0.023	0.039	0.025	0.023	0.019

ASML

The results from this study will be published as a JNCC report in the next few months.

Questions

Q: Was one slide showing that counting species abundance was faster than point intercept?

Response: Average time was 12.3 minutes for point intercept and with that we were capturing a few more species. Average time was 6.5 minutes for abundance count.[4]

[4] Point of clarification added post-workshop: It should be noted that this was because 100 points were used. In reality 50 or less would be sufficient (as determined through precision analysis), meaning that average time for point intercept would likely be the same as abundance count.

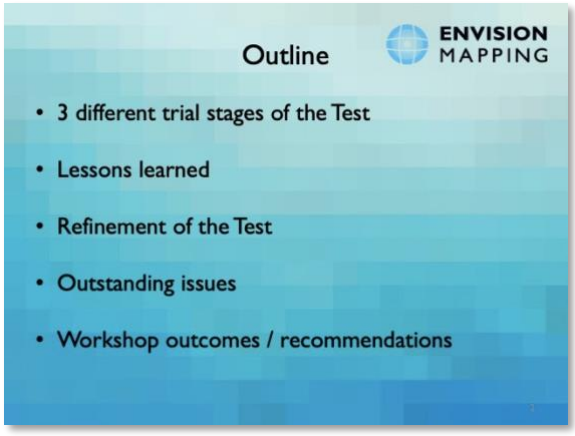
Q: Does a 20% change for power analysis come from the change agencies want to be able to detect? From a fisheries point of view we would say it's almost impossible to detect 20% change.

Response: That was our starting point – it's recommendation / guideline we've been trying from a planning perspective.

Q: Why 100 points?

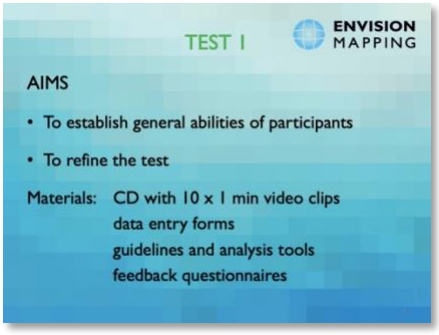
Response: You could increase or reduce the number of points. The more points, the more effort but the more resolution, while with fewer points there may be less resolution but more consistency.

Presentation 4: Development of the Video Ring Test Pilot: 2008-2009 – Alison Benson (Envision Mapping)

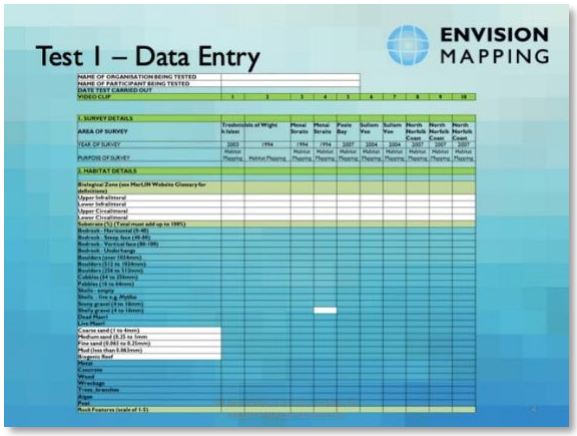


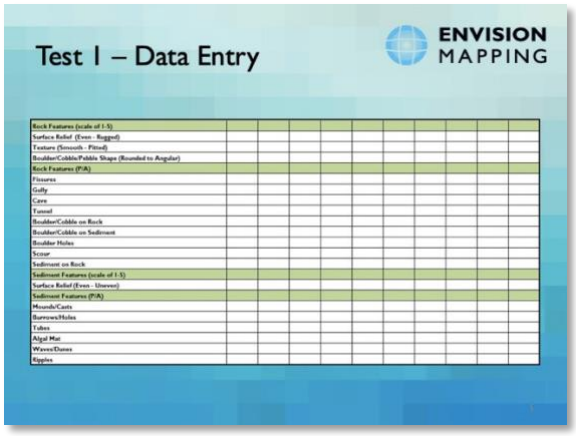
"This was a pilot project for a video ring test. The overall aim was to trial using a ring test with the same set of samples sent to all labs taking part, to give people feedback on their proficiency and to enhance quality.

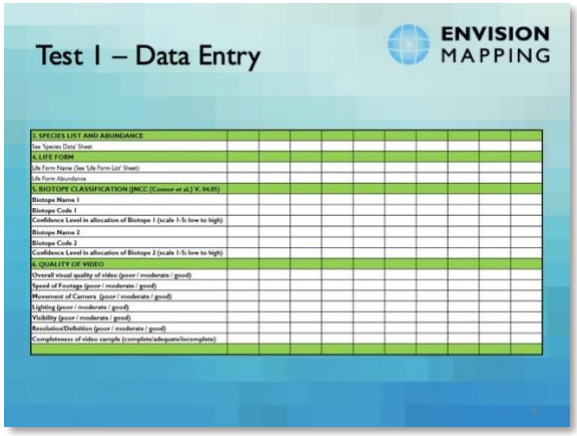
Three tests were trialled over the two years 2008 and 2009. Lessons were learned at each stage and to refine the test over time, to make it more workable. There were still issues at the end of the pilot and a workshop was held to try to resolve some of those issues.

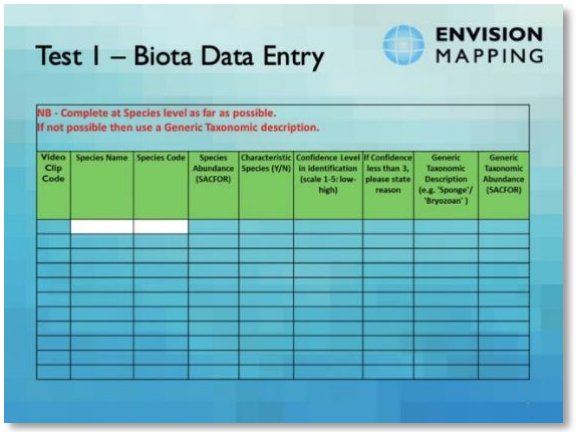


The aims of test 1 were to get a sense of the participants' ability and enable feedback to refine the test. An approach was adopted based on how video contracts were received at the time: participants were sent video clips and asked to analyse all of it. Hard copy data entry forms were provided, as well as all necessary guidance and feedback forms in hard copy.

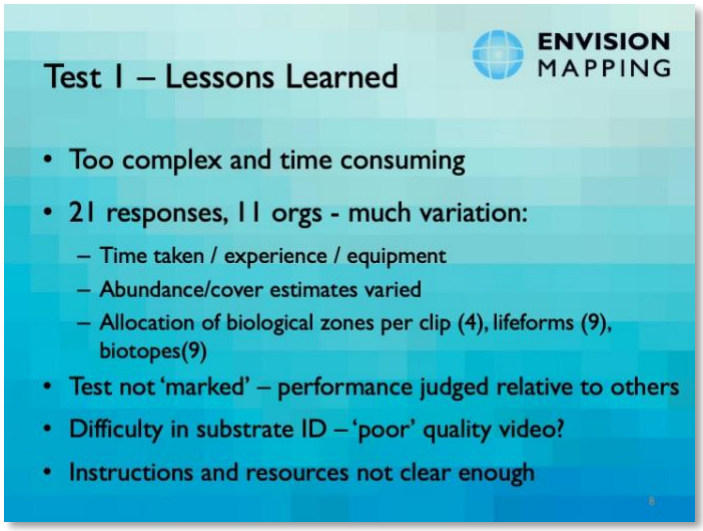








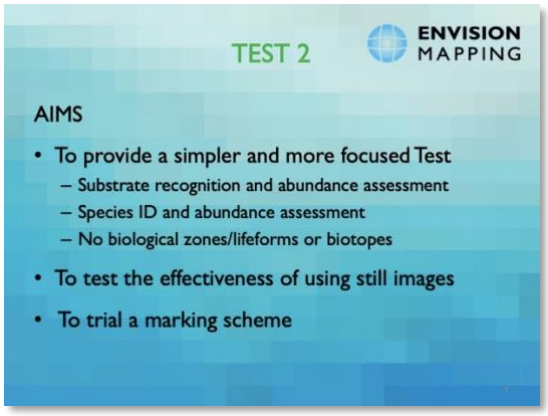
The data entry forms were based on the MNCR recording forms at the time. The only metadata participants were provided with was about the geospatial area where the video clips had come from. They were asked to fill in a lot of detail on habitat, rock and sediment features, lifeforms, biotopes and the quality of video. There was a lot of information to fill in. Species level identification was requested where possible, along with the level of confidence in that identification. SACFOR abundance data was requested too.



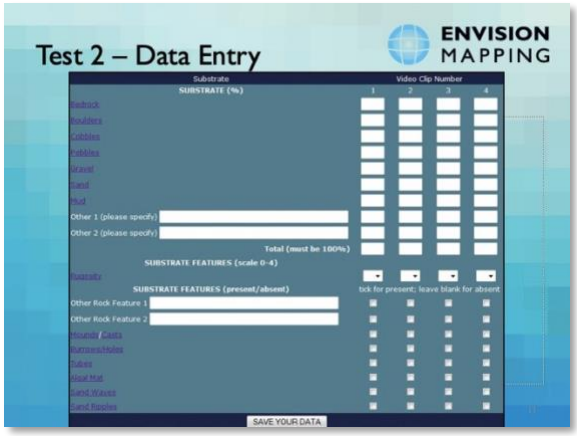
From test 1, it was apparent the process was too complex and time consuming. There were 21 participants from 11 organisations and the data varied a lot between them. Some individuals took a day to do the ring test while others took a week. The level of image analysis experience ranged from 0 to 12 years. The range of equipment used varied quite significantly too. Abundance and cover estimates were highly variable with very poor levels of consistency between different individuals.

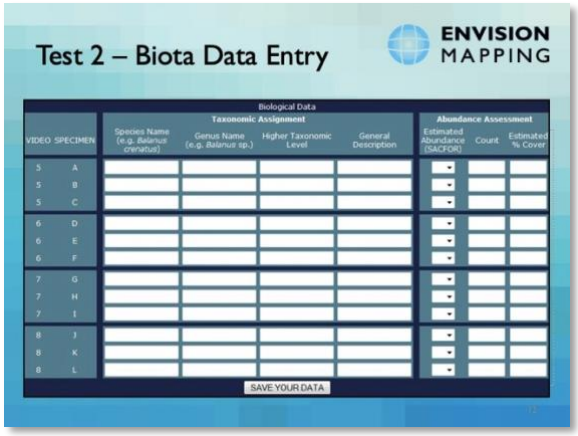
The tests weren't marked and there were no right or wrong answers. Instead a similarity test was applied to the results. There were lots of difficulties around analysing substrate. Some individuals said the video quality was poor, while others felt it was quite representative of video quality at the time. It was

realised that the instructions needed to be much clearer to reduce variability in their interpretation.

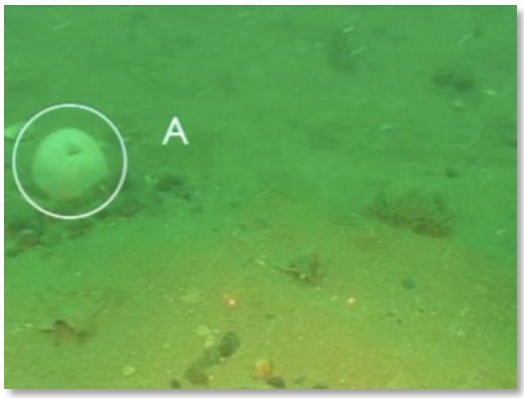


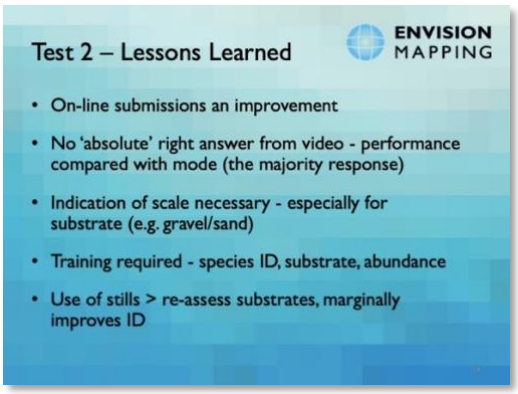
In response to feedback, test 2 was made a lot clearer and more focused. More metadata were provided with longer video clips this time. There were 10 clips: four for substrate, four for biota, two with still images. The organisms for counting were clearly annotated and a revised set tools of tools was provided to carry out the ring test. Included in this package were feedback questionnaires (for video quality too), updated reference lists and analysis tools, as well as a simplified SAFCOR scale from Seasearch.



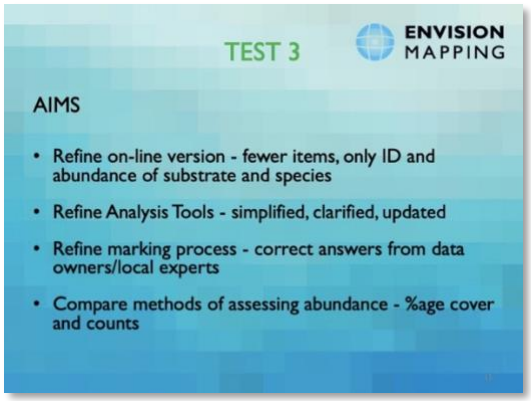


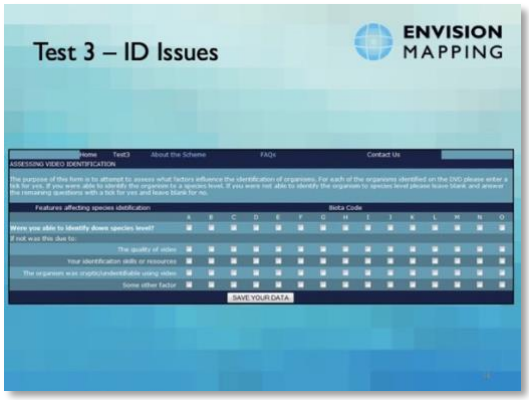
The main improvement was a purpose built online site for all documents, tools, data entry and feedback. The responses were a lot more targeted, with requests for basic substrate to be categorised, as well as presence / absence for other features. Three species were requested per video clip.



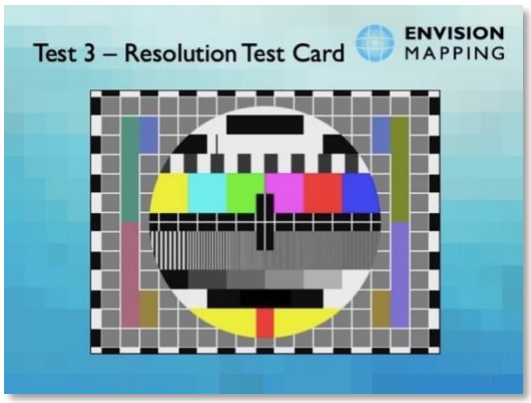


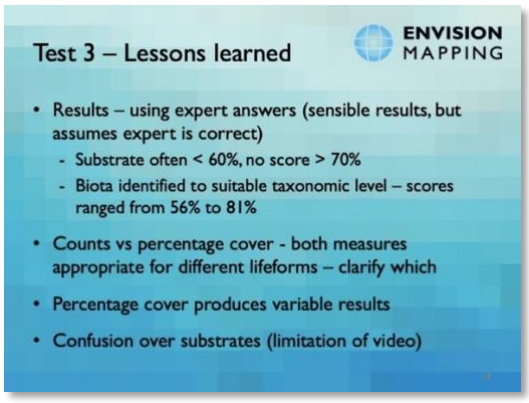
The online system and simpler approach were significant improvements to the ring test, but there was still a lot of variation in responses.



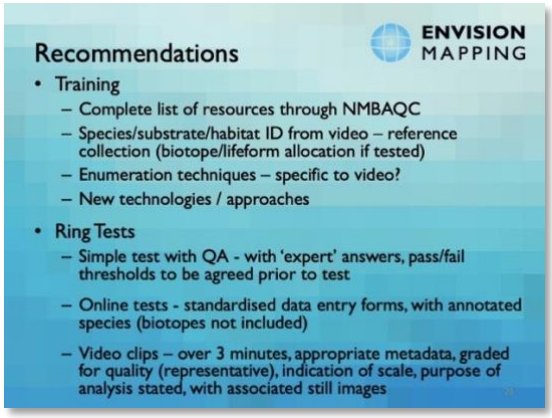


Using the feedback to improve test 3, individuals were requested to analyse even fewer items. Attempts were made to refine the tools and marking process too. The modal response was initially used for measuring quality. However, this had assumed the majority was always correct, which is not always the case. At this stage other aspects of the marking process were considered, such as weighting different elements of the test, or using ranges of correct answers (dependent on taxonomic level). There was still difficulty identifying substrate. And when the video clip was analysed with still images, there was not as greater improvement as expected. For test 3, the correct answers were provided from the data owners / experts to enable more feedback.





Providing expert answers produced more sensible results, but there was still a lack of consistency in terms of levels of agreement. This is important to note when thinking about past failure thresholds for tests. Envision have recently been asked to use external QA with over 80% agreement, but it is not currently realistic to use pass / fail thresholds for video imagery data. First the correct measures that are used in the tests need to be clarified. There is still confusion over substrates from video, which might be just one of the limitations of using video.

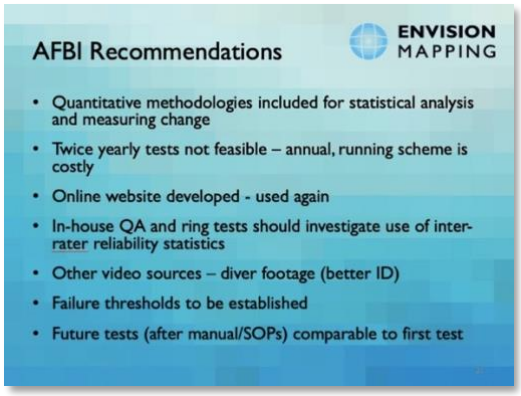


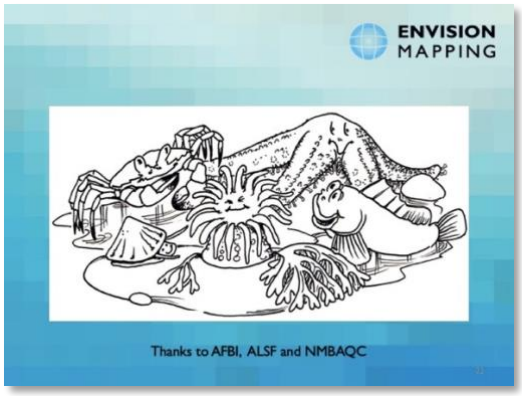
At the end of the pilot project, there were still issues remaining, which were discussed at the workshop – including level of detail for analysis, video quality, assessment and marking methods (and thresholds for passing), and training.

A lot of good points came up at the workshop. It was challenging to reach consensus, but some recommendations were made. Two have already been achieved (video analysis guidance and regular workshops).

Other recommendations focused on training and ring tests. For training, recommendations included a complete set of resources for this in a central repository, a reference collection for species, substrate, habitat and (if part of the tests) biotope and lifeform, and training in enumeration techniques and new technologies or approaches.

Ring tests recommendations included withdrawing use of using biotopes in ring tests, due to variation in interpretation, as well as biannual testing to account for personnel changes, timely feedback, consistent hardware (where possible), and analysts to receive training before tests (i.e. not take part in tests as part of their training).





AFBI also had recommendations as a result, including the possibility of resurrecting the website. Failure thresholds have still not been agreed across the board."

Post-workshop presentation: BIIGLE – Jon Hawes (Cefas)

At the end of day 3, after the close of the workshop, Jon Hawes gave a presentation on the use of BIIGLE, for those who were interested.

5. Connections and cross-cutting issues

Connections between themes

During their discussions, the working groups identified any clear connections or links between different themes or topics. These are shown in the Table 4. All of the themes are linked in some way, so this table shows only the strongest connections identified by the working groups.

Table 4. List of working groups that work to deliver evidence, standards and advice for marine benthic monitoring.

	1. Need for improved levels of method standardisation and quality control	2. Taxonomic identification	3. Image annotation software, use of machine learning and improving cost and efficiency	4. Development of standardised / specialised sampling approaches	5. Image analysis contractual agreements and resources	6. Biotope monitoring	7. Sampling units	8. Image reference collections	9. Observer consistency	10. Overall enumeration approaches	11. Morphological classification systems	12. Image analysis training
1				*								
2												
3												
4	*											
5												
6											**	
7												
8												
9												
10												
11						**						
12												

Blue: themes discussed together by a single working group – see clarifications below.

Light green: connection made in one direction (i.e. by one working group between their theme and another)

Darker green: connection made in both directions (i.e. by two working group between their two themes)

* Themes 1 and 4 were discussed together because the working group considered it important there be a joined up approach to data collection, analysis and interpretation, with very good information flow between those commissioning a survey (to define the objective, and therefore the "quality" of output needed), those conducing the survey (to be aware as to whether the quality of data being collected is likely to be "fit for purpose") and those processing and analysing survey results (to be aware of the limitations of the data provided and the purpose for which results are intended, which governs the level of detail required).

** Themes 6 and 11 were initially discussed together due to perceived linkages between the two, but were later separated again following further discussion, in order to develop clear and separate plans and recommendations.

Common questions, issues and challenges

Analysis of the range of questions, issues and challenges raised during the workshop (see Appendix 5) reveals a number of cross-cutting issues. These are summarised below.

OVERARCHING ISSUES

Standardisation and consistency

This was – perhaps not surprisingly – one of the most commonly mentioned issues. Key points included:

- The need to create a degree of standardisation, but recognising the challenges in doing so (see other issues below);
- Separation of standards for collection and analysis, including the following points:
    - Is standardisation of acquisition method possible or appropriate?
    - Focus on standardisation of data output / quality / analysis instead?
    - And/or possibly look at standardising specific aspects of acquisition, e.g. aspirational levels for kit, laser use, acquisition for analysis by annotation software, transect length, minimum number of images, etc?
- Enabling inter-comparability (of methods / data);
- Consistency of language, definitions, categories, guidance;
- Use of existing or emerging tools and guidelines as a starting point (e.g. current naming conventions, BIIGLE, CATAMI, EIP, MarLIN, MESH, WoRMS, previous biotope lifeform work, EIP), or applying learning between approaches (e.g. applying deep sea morphological approaches to shallower waters.);
- Identify which standards are needed, in what order, what their purpose is, and what they will cover.

Flexibility, relevance, context and appropriate use

Balanced with calls for standardisation were a large number of comments about the need to maintain flexibility and relevance, so that acquisition and analysis processes take account of:

- Overall purpose and aims;
- Survey conditions, location, timing and available equipment;
- Amount and type of data and metadata required and available;

- Difficult species, such as sponges and bryozoans.

There were several mentions of ‘appropriateness’ in relation to which methods are used for which purposes, particularly of biotope classification and SACFOR.

Related to the point about relevance were several comments regarding the need to define:

- The aims and scope of any survey – what questions are you trying to answer?
- The conditions of the survey;
- Key decisions and rationale for these;
- What level / scale of data is meaningful and relevant.

Dealing with and determining change

Change over time was raised in two contexts: dealing with change (how to future proof data, and deal with changing aims, classifications, equipment and techniques); and determining if change has happened (i.e. how much change is due to inconsistencies between data or analysts, what level of change is deemed significant or meaningful, how to detect changes to biotopes / communities).

Trade-offs and balance

Cutting across many of the conversations and written outputs of the workshop was the issue of trade-offs, and the need to balance the various parameters involved in image acquisition and analysis.

As well as the standardisation / flexibility balance touched on above, other aspects such as cost, time, effort, quality, consistency, level of detail, transparency, and confidence were frequently mentioned in various combinations.

TECHNOLOGY

Acquisition techniques and technology

Particularly in the early part of the workshop, some questions were raised about different acquisition methods and their uses.

Annotation technology and machine learning

As well as forming the focus of one of the workshop themes, annotation software and machine learning were mentioned in relation to other issues, including:

- Interoperability between organisations (regarding consistency and data sharing);
- Potential for automation to improve accuracy and consistency, but issue of improving technology creating inconsistencies;
- Possible loss of human expertise with respect to machine learning algorithms;
- Current imbalance in knowledge of annotation software / AI and in computing / data handling skills;
- Cost implications (time and resource)

- Potential for use as a supporting technology, e.g. role in QC, training, collaboration and development of reference collections; value in repetitive or large scale tasks.

DEFINING SAMPLES

The nature of sample units was commonly mentioned across several themes, with several comments about the need to determine size, quality, type (still / video, qualitative / quantitative, format) – sometimes with clarification that this should be appropriate to purpose / use.

Some comments asked whether there should be minimum / maximum parameters for some of these aspects (e.g. for resolution, size, video length, quality, replicability, etc) and guidance on aspects such as sample selection (e.g. randomness), metadata and storage (e.g. what format to save files in).

ANALYSIS

Level of confidence and resolution

Several comments across a number of themes ask questions about level of certainty and resolution required for different uses, and the need for clarity on this – with additional points about preserving raw data for potential future use, tackling uncertainty / differing views, and the use of reference collections.

QA methods and detail

QA and QC were issues cutting across most of the workshop themes. Points raised included:

- Standardising, e.g. internal / external, number of reviews, early or interim data quality checks or reviews. Potential for consistent QA / QC tools and software
- Minimum quality specifications and built in time / budget for QA (e.g. in contract);
- Potential use of annotation software;
- Ring tests (different tests for different purposes, use of reference collections, coordination, cost, governance);
- Setting grading and thresholds;
- How to moderate own samples;
- Resolving analysts and lab inconsistencies.
- Reference to Lin's Concordance Correlation Coefficient – a statistically robust method for continual intra- and inter-counter observer analysis.

COLLABORATION AND COMMUNICATION

Communications, collaboration, data sharing

Across many themes, the importance of collaboration and communication was apparent. At a high level this included reference to communications and partnerships between organisations, sometimes for specific purposes (e.g. sharing data or resources, training, influencing government funding), and more widely to communicate any new standards or resources.

More specifically, there were several queries about how data could better be stored or shared openly between organisations. Online resources (e.g. BIIGLE, SQUIDLE) were commonly mentioned, as well as crowdsourcing imagery or analysis, online training, social media and – most commonly – the potential for shared online reference collections (see discussion of theme 8 for further detail on this topic).

Governance and ownership

Alongside the many comments relating to collaboration and shared resources, a number of comments about governance, ownership, funding and management were raised, including for:

- Images, online collections;
- QA / QC processes;
- Training, standards and guidance;
- Key decisions (e.g. regarding taxonomy).

PEOPLE

Learning, training and sharing experience

As well as being the focus of one of the workshop themes, the issue of learning and training featured heavily throughout other topics, including the following points:

- Potential for workshops on specific topics (tools, taxonomy, enumeration, video identification, automated annotation, different biomes, SACFOR, NMBAQC);
- Can training improve consistency / at what point does additional training have limited impact on consistency?
- Potential for an online forum for analysts to share ideas, advice and experience;
- How would training be funded, and by whom (and would it be free or paid for?);
- Interaction between QC and training;
- Possibility of certification / accreditation scheme, and whether or not this is desirable (e.g. if it requires payment for certification)
- Refresher / retraining courses;
- Minimum training requirements for analysts;
- Online tools, e.g. species identification tests;
- Centre of excellence for taxonomy

Expertise, staffing and progression

A range of points were raised regarding staffing. These were primarily focused on the need to improve and retain expertise, as well as the need to address the risk of analyst fatigue. Contracts were also a common issue (see discussion of theme 5 for further detail on this topic).

6. Key outputs: working group theme discussions

The final outputs from the working sessions across the three days are presented in this section. This includes summary discussions of key issues under each of the twelve themes, followed by summary discussions of four additional themes.

This content forms the main guidance for the Plan Development Group as it begins to put together the draft Benthic Imagery Analysis Action Plan. Supporting information is provided in:

- Appendix 4: References.
- Appendix 5: Transcript of workshop discussions, by theme – including the full range of questions, issues and challenges feeding into the discussions below.

The twelve original themes:

1. Need for improved levels of method standardisation and quality control;
2. Taxonomic identification;
3. Image annotation software, use of machine learning and improving cost and efficiency;
4. Development of standardised / specialised sampling approaches;
5. Image analysis contractual agreements and resources;
6. Biotope monitoring;
7. Sampling units;
8. Image reference collections;
9. Observer consistency;
10. Overall enumeration approaches;
11. Morphological classification systems;
12. Image analysis training.

Additional themes:

- Data sharing and management (thirteenth theme from questionnaire outputs);
- Linking with marine industry, developers, Department for Business, Energy and Industrial Strategy and Marine Management Organisation;
- Culture around data;
- New tech pipeline (refers to scoping and validating new technology).

THEME 1. Need for improved levels of method standardisation and quality control & THEME 4. Development of standardised / specialised sampling approaches

Group response to peer review and challenge

"Our response has been to take on board comments (see Figure 4 below). They provided a new insight into aspects we hadn't looked at. We have incorporated some things into core suggestions, and some into practical actions. The timeline depends on human behaviour rather than actions."

ISSUE: Need for improved standardisation and quality control of outputs

Discussion

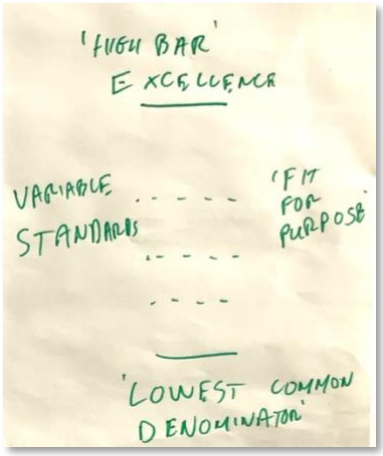
- Outputs need to be fit for purpose and so every monitoring purpose would require a different standard/ set of standards.
- Different standards can have very different cost implications, therefore, cost needs to be considered with the production of each standard.
- How do we ensure standards are affordable? Should this be the responsibility of the commissioning organisation (i.e. the client requiring the imagery analysis work) or should there be a centralised government fund (provided by Competent Monitoring Authorities?) to contribute to supporting the cost of quality control scheme?
- Over-prescription of methods (i.e. too detailed and specific an approach) can be counter-productive.
- Standardise outputs for acquisition, methods for processing.

Connections

- Flexible standards and those that focus on required outputs: these can be achieved via guidelines rather than rigid procedures.
- Affordable standards – needs greater clarity in contract specifications to allow realistic tenders e.g. potential for high percentage downtime.
- Quality control stage used while on survey – are survey conditions appropriate for data acquisition or go to downtime?
- Issues of time-constraint and conflict of interest on survey.
- Different outputs at acquisition and data interpretation.

Relevant notes / challenges

- Hierarchy of analysis – which level do you get to in your analysis?
- 'Repurpose images' – identify if they can be used for other purposes ⇒enable search later.
- Need standard to define if image quality is fit for purpose.
- Streamline data-processing, making it possible to 're-process' old data in the future.
- 'Standards' should be aimed at achieving compatibility.
- Do we need standards or recommendations?



Action / advice for PDG

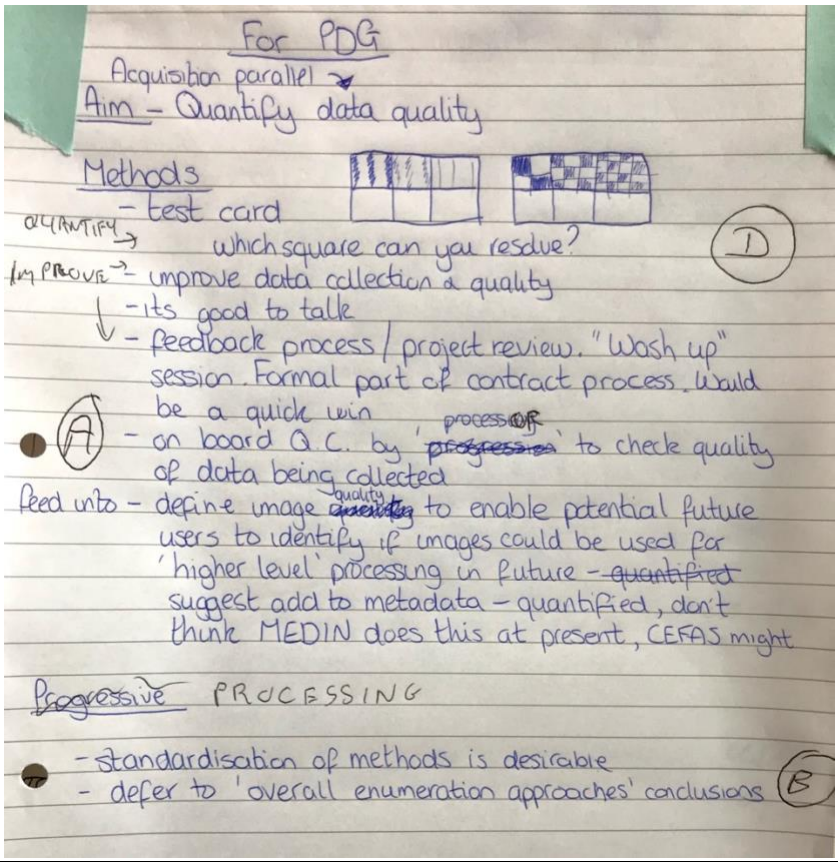


Figure 4. Recommended actions for the Plan Development Group to take forward for themes 1 and 4.

Note, letters indicate cross-references with specific comments or challenges. 'A' refers to having someone on the survey who has a vested interest in the results. 'B' is an example of the connections to other themes. 'D' links to a comment that standards should cover quality of image.

THEME 2. Taxonomic identification

Group response to peer review and challenge

"We only had comments about taxonomic identification (and one about reference collections, covered elsewhere). We have incorporated a couple of suggestions, and have annotated some of the challenges [see Appendix 5]. There are no timelines as such"

ISSUE A: How to increase the level of taxonomic expertise

Discussion

- Central, easily accessible reference collections and identification forums.
- More identification courses with more in-depth content, that are easily available. Webinars / online courses to reduce costs. Hosted by experts rather than citizen scientists.
- More sites like Habitas[5] but better!
- Funding for PhDs, linked to taxonomic analysis of previously collected imagery.
- Production and distribution of an Epifauna Identification Protocol → stop at level you are certain of (Taxonomic Discrimination Protocol style).
- Engagement with universities, museums, etc. – opportunities for funding?
- Subsidise training courses.

Connections

- 8. Reference collections.
- 9. Observer consistency.
- 12. Image analysis training.

ISSUE B: How to improve consistency of identification

Discussion

- Reference collections can provide a feedback system (HBDSEG hosted?).
- Can location tags be used to generate taxa lists relevant to that location?
- Identification courses.
- Consistency of enumeration techniques needs work.
- Consistency of nomenclature / morphological types needs work.
- Augment analysis with CATAMI[6] or similar hierarchical structures.

[5] Habitas: http://www.habitas.org.uk/marinelife/
[6] CATAMI: http://catami.github.io/

- Epifauna Identification Protocol (EIP) – when to stop! Distribution of EIP to all relevant parties to start using it asap.
- If in doubt, stop at a higher level and indicate suspected e.g. species in qualifier only.
- Consistency (of quality) in acquisition of data.
- Ring tests – problematic.
- Make use of Marine Biologist Forum – currently on Facebook.

Connections

- 8. Reference collections.
- 11. Morphological characterisation systems.
- 1. Method standardisation.
- 4. Standardised sampling procedures.
- 10. Enumeration techniques.
- 9. Observer consistency.
- 12. Image analysis training.

ISSUE C: How to improve expertise / consistency

Discussion

- Reference collection to be constructed (based upon existing resources – Marine Environmental Data and Information Network (MEDIN[7]), Data Archive for Marine Species and Habitats (DASSH[8]), Habitas[9], Marlin[10], National Oceanic and Atmospheric Administration (NOAA[11]), etc).
- Central website that is easy to find (could be JNCC biotope pages for example), that contains lots of resources, including books, websites etc as well as upcoming identification courses, forums, etc.
- Epifauna Identification Protocol to be developed and distributed – should be common knowledge – linking biological expertise with technological expertise.
- Training is key – need all courses to stress that over-confidence is a problem.
- Engagement with universities, museums, international taxonomists, consultancies, etc. is a priority. Experts must be involved in training – funding for this may be an issue.
- Ring tests – collaborative, could look at specific groups, indicator species.

[7] MEDIN: https://www.medin.org.uk/
[8] DASSH: https://www.dassh.ac.uk/
[9] Habitas: http://www.habitas.org.uk/marinelife/
[10] Marlin: https://www.webarchive.org.uk/wayback/archive/20130501175401/http://www.marlin.ac.uk/
[11] NOAA: https://www.noaa.gov/oceans-coasts

THEME 3. Image annotation software, use of machine learning… and improving cost and efficiency…

Group response to peer review and challenge

"Comments were more in the form of what is this and how can we use it, but there were some good comments about machine learning and the need for a training programme – there are likely to be a lot of calls for interest for introductions to image annotation and machine learning. Another comment was why do you need to get a huge data set together when have google available with hundreds of starfish (it gives a huge variation of context, but that could be useful). We've added a bit of timetabling for both elements. Requirement for workshops / training is the main thing."

ISSUE A: Machine learning

Discussion

- Recognise number of systems have been developed.
- Benthic monitoring underdeveloped: 'busy' images with poor contrast.
- Nascent stage.
- Bottleneck is having annotated data set for training (100k and images).
- Stress testing important.

Tasks

- Scope any models which are working in benthic – deep sea (Autonomous Underwater Vehicles; AUV).
- Conduit for creating an annotated training dataset (key priority).
- Periodic state of the art review.
- Develop introduction to machine learning for biologists.
- Define the tasks we as a community want machine learning to achieve.
- Scope out international expertise and work out how best to learn from them.
- Scope possible "x prize" competition for benthic image machine learning (after training data set built).

Reference collection collation

- Head start using Google? Good idea! Starting point algorithm.

Ensemble approach

- Multiple models can be used.
- Transfer learning from each model.
- Polygons better than points for training software.
- Consider positives of human-mediated Artificial Intelligence (AI).

Timeline

1. Strategy for data call (benthic imagery).

- Training dataset.

- Image types.
- Repositories.
- Participants.

2. Workshop to conceptualist the machine learning workflow for benthic biology (including both biology and machine learning specialists).

- High level: input criteria, output criteria.
- This feeds into image annotation timeline on other sheet.

Connections

- 10. Overall enumeration approaches.
- 11. Morphological classification systems
- 1. Standardisation of analysis methods.
- 7. Sampling units / knowing / recording relevant metadata.

ISSUE B: Image annotation in terms of cost efficiency

Discussion

- Universality of exports – CSV file with coordinate (XY) tags.
- Cloud – benefits or hindrances.
    - We don't need just one solution – cloud or local to suit.
- Interoperability. Pull and push species lists / stills.
- Needs to be usable in commercial setting.

Annotation tasks

- Identify a list of features which any platform must have.
- Preliminary task to identify applications across community (i.e. not just benthic imagery).
- Community building can start with creating an annotation user group.
- Define levels of image annotation (3 levels) to help machine learning.

Cost efficiency

- Initial cost outlay likely to be offset by efficiencies in future.

Timeline

1. Call for interest to be sent via workshop mailing list.

2. Identify a pool of experts to build supportive, collaborative community across stakeholders.

3. Define applications for annotated imagery.

4. Define features that the software platforms must have.

5. Build specification for experts (e.g. polygon vs point) data format.

6. Combined training course, to cover the following: awareness of why to use; key skills in using; key development issues and solutions.

Connections

- 1. Standardisation of analysis methods.
- 2. Taxonomic identification (creation of reference collections).
- 8. Image reference collections.
- 9. Observer consistency.
- 10. Overall enumeration approaches.
- 12. Image analysis training.

Overall point

- Introduction and training required – combined for annotation and machine learning. Link with international learning.
- Three-day course? With three entry levels: 1. Beginner; 2. Intermediate; 3. Expert.

THEME 5. Image analysis contractual agreements and resources

ISSUE: Clarity and purpose of contracts

Discussion

- Increase input of information at start to prevent over-working and provide a more in-depth scope.
- More information on purpose and analysis purpose (objectives of survey very useful). This needs to be clear with as much detail as possible to allow precise, efficient tendering.
- Wash up meetings needed – opportunity to provide feedback from analysis (quality or other issues) face to face. This information can also feedback to survey planning and implementation stages (added value).
- Communication throughout project so that it is two-way and adaptive.
- Remove poor quality images / video from data set or allow contractors more scope to not analyse.
- Appropriate levels of metadata to be standardised (e.g. images named and associated data checked) but don't expect contractor to do this or need to provide extra budget/time for this.
- Have an initial time period at start of contract to review data to decide what to remove and what detail is required for analysis. Can lead to adaptive contracts.

Connections

- 1. Method standardisation: Consistent methodologies efficiency in tender process.
- 1. Method standardisation: Quality Assurance (QA) consistency to stipulate precise requirements.

THEME 6. Biotope monitoring

Group response to peer review and challenge

"Comments were mostly positive, and there were some suggestions about how this might be used. We worked on refining the timescale. We came up with a mix of using morphological / CATAMI[12] classification to improve biotope monitoring and classification."

ISSUE: Identifying certain species from imagery is problematic where samples can't be collected

e.g. sponges, bryozoans, hydroids on circalittoral rock → can't classify the community to which an image 'belongs' with sufficient resolution.

Discussion

- Aim: use a morphological classification system to improve:
    - Ability to describe / resolve biotopes / biological communities (at level 5 of the EUNIS[13] habitat classification) – for mapping biological communities.
    - To be able to use imagery more effectively to describe and map biological communities, specifically those on sublittoral "rock" and other hard substrata. (Get better resolution imagery than we can currently get with biotope classification).
- Step 1: Develop a morphological classification system to describe biological communities of sublittoral hard substrate from imagery.
- Test 1: Test if the system can be applied to modify level 5 of EUNIS habitat classification.
- Test 2: Use it to test if morphology to detect change across a pressure gradient.
- Other tests to be suggested…

[12] CATAMI: http://catami.github.io/
[13] EUNIS: https://www.eea.europa.eu/data-and-maps/data/eunis-habitat-classification

THEME 7. Sampling units

Group response to peer review and challenge

"We had some good comments that made us feel justified in working on this. But we realised we need to provide upfront definition and action items for providing pictorial data to help people understand the context of sampling units."

ISSUE A: What is the appropriate sampling unit for the specific question of the study – size of sampling unit needs to be representative of the population

Discussion

- Step 1: Define your sampling population (template / pro-forma).
- Step 2: Define the size (m^2/numbers counted) and number of your sampling unit (template / pro-forma).
- Step 3: Record any limitations / reductions in your scope post-survey – if necessary go back to step 1 regarding sample unit changes in the field.
- Step 4: Keep documentation of steps 1-3 with data and results.

Action items for group

- Create a table for recording what the aim of the study is, what the sampling population is (spatial scale / size), prompts of what you should record.
- Flow chart for how to define the size of sampling unit based on existing knowledge, with prompts for what you need to consider and to record your rationale (potentially need a pilot study / simulation study).
- Replicates are sampling units – it depends on survey design – implicit in definition of sampling unit size and number.
- Collate information based on table [see first action item] to give examples of types of survey scope & decisions made.
- Need to make a clear terminology explanation on the relationship between sampling strategy / design and sampling unit – diagram to accompany table & flow chart.
- Create pictures of example sampling units to go with table and flow diagram.

Connections

- 4. Standardisation of acquisition methodology.
- 4. Specialised sampling approaches – best available technology.

Sample unit, definition

A sample unit is representative of a sample population. There is an equal chance of variation in every randomly located sample unit within a sample population. For example, a sample unit could be:

- A single image;
- A group of images;
- A video clip;
- A subset of video clips.

ISSUE B: Purpose specific survey requirements vs. trying to capture more information to "future proof"

Discussion

- You should not compromise the scope of your image collection – refer to your sample population.
- Additional information needs to be specified if it means changing sampling unit. Ask other people if aims can be extended to meet their needs.
- Augment your sampling by collecting more data (sampling units) than you need to ensure you have sufficient data to answer your question, when your data have been 'cleaned', e.g. buffer to account for poor image quality.

THEME 8. Image reference collections

Group response to peer review and challenge

"We incorporated most comments, although some were addressed by things we'd already looked at. We split this into five different elements. And the final thing is sorting out the expert groups and quality assurance (QA) data in reference collections, but it's about trying to do things at the same time where you can."

STEP 1, ISSUE A: Ownership

Can this be done?

Discussion

- Publicly funded – copyright, caveats, credited (open data).
- 'Opt in / out to' use data from commercial contracts.
- Subscription or freely available.
- [Incentives for industry to submit data, i.e. good publicity.]

Short term steps to take

1. Contact relevant institutes / companies – do they have an existing reference set?
2. Would they be prepared to release it? Tick box options for various uses.
3. Industry reward for participation (credited, free advertising, reduced subscription fees).

[Alternative sources of data, e.g. Seasearch, citizen science]

Connections

- 3. Image annotation software.
- 5. Image analysis contracts (commercially sensitive).

STEP 2, ISSUE B: What classification reference collections already exist?

[E.g. Annex 1 features, Priority Marine Features (PMFs), etc., Video?][14] Avoid duplication.

Discussion

- Centralised list of resources e.g. links to existing catalogues.
- Identify existing resources.
- Quality Control system set up.
- Scope / literature review of any previous processes in generating reference sets? E.g. ICES Working Group on *Nephrops norvegicus (Nephrops)* Surveys (WGNEPS).

Medium term

[14] Point of clarification added post-workshop: examples of usage include for species identification or abundance training.

- Developing a catalogue.
- Multiple images and range of conditions (image quality, resolution, etc.)

Connections

- 3. Image annotation software.
- 9. Observer consistency.

STEP 2, ISSUE C: What to include.

Discussion

- Identification of organisms (taxonomy vs. morphological / functional) → CATAMI[15]?
- Classification of habitats / biotope.
- Vulnerable marine ecosystems (VME) – cobble reef / *Sabellaria*.
- Abundance reference e.g. *Nephrops* burrows.
- For each of above:
    - Metadata;
    - Multiple images – different camera platform; different environmental conditions; image quality; different aspect of subject matter, e.g. oblique / vertical cameras – and angel subject is in relation to camera (square on, side on, etc.);
    - Hierarchical classification – taxonomy and morphology.

Steps

1. Identify what's required.
2. Identify priorities and trial dataset. Identify user groups. Test a beta version.

Metadata/imagery tags required:

Hierarchical classification – taxonomy & morphology.

Location, depth, temperature – refers to distribution ranges of species.

Connections

- 1. Improved method standardisation and quality control.
- 3. Image annotation software.
- 9. Observer consistency.
- 12. Image analysis training.

[15] CATAMI: http://catami.github.io/

STEP 3, ISSUE D: Validating reference sets.

Discussion

- Identify experts in that field.
- Include description and justification for name and include metadata.
- Maintain training for all users.
- Quality control (QC) / statistical confidence in identification.
- Have multiple examples.
- Quality / confidence score for identification, e.g. photo id only, or photo and sample id, in generating reference set.

Steps

- Identify experts to verify images.
- Put quality assurance and control (QA / QC) process in place.
- Identify confidence scores.
- Provide guidance on what level taxa can be identified to depending on purpose & quality.

Connections

- 1. Improved method standardisation and quality control.
- 3. Image annotation software.
- 9. Observer consistency.
- 12. Image analysis training.

STEP 3, ISSUE E: Setting up infrastructure.

Discussion

- Who will manage / host / pay for the processes?
- Work with others with similar interests, experience and / or knowledge.
- Identify knowledge gaps and address.
- Make available on and offline, and 'live' link to WoRMS.
- To include video and still options.
- Method to query available data for taxa, location, depth.

Steps

- Look and learn from existing systems.
- Find funding.
- Identify a host.
- Integrable with other platforms e.g. WoRMS (API interface = leverage tech to simplify exchange).

THEME 9. Observer consistency

Group response to peer review and challenge

"We divided this into two issues: improving consistency and dealing with instances we know we can't do anything about. We decided a staged approach would be a good one to go for. We had various suggestions about training, using annotation software, etc. And for the variation we know is there, we looked at a stage approach to taxonomy (e.g. define classifications based on detail able to get from video). We could use those different levels to provide confidence in the data.

Some of the key issues were looking at accepted error and also how the design of the analysis is undertaken – we suggested it should be randomised to avoid any auto correlation. Some found it a bit difficult to see how it would be practically implemented, but we thought about using databases to overcome that. Annotation software could be useful to revisit data. And we thought it would be useful to store metadata with data to enable analysis to further level of detail."

ISSUE A: How to improve observer consistency

Discussion

- In house suggested process:
- Stage 1: Review all photography (remove poor stills?). As a group of observers make a standard checklist (taxon and qualifier). Validate / calibrate observer ID.
    - At this stage, can consider reference collections, identification sources if accessible.
- Stage 2: Enumeration "training" (e.g. burrows video).
- Stage 3: Internal quality control (QC; e.g. 10% data checked).
- Stage 4: Evaluate analyst variability.

Other concurrent suggestions

- Introductory course to video identification (online? Skype? To reduce cost)
- Formal process and protocol to improve observer consistency. Consider observer fatigue, time restraints of analysis / annotation.
        - Carry out analysis as above (QC)
        - Evaluate variability and understand impact on analysis.
    - Also consider specifics of analysis (e.g. number of observers)
- Familiarise self with abundance assessment. This must be agreed with client/project manager.
- Ring test / own sample – feedback loop with regulatory body. Links to contracts – regulatory driven as client may not want "their data" in public domain.
- Equivalent of Seasearch competency test (region specific?). This may need annual reassessment but minimum competency may be appropriate
- Reference collections to improve consistency.
- Training to improve consistency needed.

- Image annotation software to improve consistency.
- Overall enumeration approaches with standards to improve consistency
- Morphological approaches standardised.
- To minimise intra-observer variability could randomise analysis.
- Actions: formalise process, establish procedure, analysis, QC, variability and importance of this per habitat / species. The create a database to incorporate stage approach to taxonomy. Provide guidance on process and protocol. Create an introductory course to video identification with examples of variability for specifics.

Connections

- 3. Image annotation software.
- 5. Links to contracts – regulatory driven as client may not want "their data" in public domain.
- 8. Image reference collections.
- 11. Morphological classification systems
- 10. Overall enumeration approaches.
- 12. Image analysis training.

ISSUE B: There will be inconsistencies. How do we deal with them?

Discussion

- Staged approach to taxonomy. Agree stage of classification with client (cost / quality / analyst dependent). Note that as stills are retained, dependent on image quality, may be able to ID further in future.
    - E.g. physical → morphological → CATAMI → Genus → Species.
    - AI would help here if revisiting identification. Also for QC.
- Excepted error dependent on question, contract, habitat, taxa, etc. → importance of variability in observer will be dependent on variability of habitat / taxa and relation of change that want to detect.
    - Excepted error at individual photo level versus final categorisation.
- Ability to aggregate data based on a defined structure (Epifauna Identification Protocol (EIP) could link) → develop database that could do this.
- Record specifics of analysis / annotation design → variability.

Connections

- 3. Image annotation software.
- 11. Morphological classification systems

ISSUE C: Actions to take forward (→ combine 9A and 9B)

Actions

- Questionnaire: what process / procedures does everyone use?
- Develop guidance:
    - Process / procedure (see 9A) → Agency led;
    - Specific QA (training) / QC (checks);
    - Variability.
- Provide examples of variability for specific habitats and species → assess acceptable error. MBA have data.
- Develop a database to incorporate staged approach to taxonomy where more info can be extracted → EIP / CATAMI[16] / users → funding required.
- Introductory course on video identification → online / Skype → evaluate Seasearch to see if can assign "competence". → Funding required.
- Ring test / own samples → consider "assessment" level and what is acceptable e.g. levels of identification and enumeration. Do macrofauna / macroalgae existing methods work for video imagery / stills? Needs investigating. → Funding required.

[16] CATAMI: http://catami.github.io/

THEME 10. Overall enumeration approaches

Group response to peer review and challenge

"We very quickly came to the realisation that different purposes will require different enumeration approaches. We thought a decision tree might be useful to move from purpose down to the best practice recommendation for an approach. One comment suggested not using the SACFOR system going forward. We had a bit of think about whether that might be sensible (given potential issues with comparison to historic data) and whether it might be worth coming back to the group as a whole about that. Essentially, we didn't get any wildly differing comments about short / medium / long term approaches."

ISSUES: Purposes of enumeration vary greatly – variety of approaches are suitable. Needs to be robust, efficient, cost-effective, valuable in the long term. The best approach is dependent on size of organisms, spatial scale, etc.

Discussion

Short term

- Review benthic habitat monitoring guidance.
- Carry out a questionnaire of SACFOR purposes by users to determine who is using it and what for.
- Review outcomes of comparison exercise & review existing research.
- Guidance on minimum size of organism that should be counted, e.g. 10mm in $1m^2$.
- Update / end SACFOR? Find out how people feel & think (key uses, improved guidance (Epifauna Identification Protocol (EIP)) about intercalibration. Identify appropriate uses of SACFOR and add to decision tree.

Medium term (0.5-1.0 years)

- Produce decision tree to support selection of enumeration approach for purpose & resources. Can we prescribe?
- Identify minimum requirements for each purpose?
- Option for developers to meet guidelines to feed into wider monitoring. If it's in the guidance, then it should be followed.

Long term

- Combining info on taxa with CATAMI[17] (could be short term if using BIIGLE) & enumeration approaches. Integrate data into CATAMI structure. Dependent on development of CATAMI & links to EIP scope. Should CATAMI level & presence/absence (P/A) be the minimum?

Overall point

- Products / terminology need to be accessible to a range of audiences. Both decision tree and EIP etc.

Connections

[17] CATAMI: http://catami.github.io/

- What is your purpose / question? Design, etc.
- 1. QC standards of data – what do the questions require.
- 2. Taxonomic identification - Taxa resolution, image quality.
- 3. Image annotation software / machine learning.
- 5. Contract specification requirement – cost efficiency.
- 7. Sampling unit / FoV minimum.
- 12. Training on enumeration techniques.

THEME 11. Morphological classification systems

Group response to peer review and challenge (as for theme 6, since these were discussed by the same working group)

"Comments were mostly positive, and there were some suggestions about how this might be used. We worked on refining the timescale. We came up with a mix of using morphological / CATAMI[18] classification to improve biotope monitoring and classification."

ISSUE: The ability to consistently and reliably describe biological communities from digital imagery with enough resolution

(by using a UK morphological classification system).

Discussion

- Two main stages to investigate use of morphological classification system and roll out for future implementation.
- Review classification systems e.g. CATAMI, Morphological Taxonomic Unit (MTU) catalogue.
- Pilot test on an image set (e.g. sponge and anthozoan indicator imagery or MPA monitoring survey) and adapt classification system to fit UK sublittoral (shallow – deep sea); would remove faults across years / surveys to assess performance and detect any changes in the communities. Incorporate Epifauna Identification Protocol to help define / resolve terms used.
- Method: Contract / ICES-style working groups.
- Can utilise platforms such as BIIGLE to enhance quality assurance and quality control (QA/QC) and produce reference collections.

Connections

- 1. Improved method standardisation.
- 2. Taxonomic identification.
- 8. Image reference collections.
- 9. Observer consistency.

[18] CATAMI: http://catami.github.io/

THEME 12. Image analysis training

ISSUE A: Consistency of training for image analysis

Discussion

- What entry level in needed?
- Personnel training.
    - In-house expert.
    - Outsourced company.
- Reference collection – workshops particularly on problem taxa, habitats, etc.
- Maintenance.
    - Geographically specific training.
    - Emerging taxonomic techniques (molecular).
- Video / image-based identification guides that are in alignment with the quality of true survey data – have these as standard, approved materials for training course.
- Online mandatory training module e.g. Marine Stewardship Council (MSC) audit training.
    - Workshops.
    - In developing training materials use expert knowledge for specific taxa.
    - Evaluate effectiveness of training (engaging, practical, fuller biological info – relevance, feedback on benefit of training).
    - Pass / fail.
    - Practice texts – refresher training.
    - Induction element.
    - Available multi-level (universities / government / private etc.).
    - Accreditation.
    - Achievable in short timeframe.
- Levels of ID training.
    - Online.
- Generic video / image analysis = global.
- Region specific taxa = local / UK.
    - Dedicated / additional modules or courses (e.g. Seasearch specific courses).
    - In-house or through 3rd party provider.

Tools

- Image reference collections.
- Contractual enforcement.

Timeline

2019

- Collate all current guidance & analysis tools – including current global seabed monitoring.
- Evaluation of above – research to prove worth / value of training.

2021

- Identify people's willingness to share & collect a reference set.
- Decide training materials – format / purpose and assign experts

2023

- Develop interface for online training.
- Identify facilitator / ownership – which assess pathway (link to issue 2).
- Decide on distribution model (outreach).
- 'B test' and evaluate to prove value of training in improving data.
- Specialist workshops based on reference collection available, machine learning available.

Dependent on timeline above and decisions on training materials and content, decide on common elements necessary for in-house training within organisations (in-house reference collection, sub-sample undertaken side by side then assessed decisions on quality and level of analysis).

Notes on organisations

- Field Studies Council (FSC) – experienced identification work – marine (Dale Fort).
- MSC – Marine Stewardship Council – fisheries certification audit.
- Part of CPD – continued professional development.
- Training should count toward CIEMS – endorse training.
- IES – Institute of Environmental Sciences (chartered institutes City & Guilds).

Connections

- 4. Sampling approaches
    - Quality of image data
- 5. Image analysis contractual agreements & resources
- 9. Observer consistency.
    - Training
        - Taxonomic ID
        - Enumeration techniques.

ISSUE B: Sharing and ownership of training – video analysis

Sub-issues

- Maintenance and validation.
- Costs of running – assessing performance.
- Updating with emerging techniques.
- Input from multiple sources (materials) and sectors.
- Direction from regulators / end users – consistency between them; updating.
- Expert input on validation.

Discussion

1. Completely public central repository

- Pros: open access, inclusive, all stakeholders; public interest – access to more funding.
- Cons: costs, funding.

(E.g. NMBAQC, Conservation Agency Research Institute (employ company – not in interests of competitor), Natural History Museum, MARLIN, Universities).

2. Restricted access – with good reason (no cost)

- Pros: more control over users and ring test results.
- Cons: exclude some users.

3. Buy licence for access

- Pros: it's self-funding.
- Cons: exclusive and opt out if possible.

Connections

- Reference collections.
- Quality Control and ring tests.
- Image annotation software and machine learning (BIIGLE).

ADDITIONAL THEME: Data sharing and management

Discussion

- Where is your data? Where should it go?
    - WoRMS;
    - OBIS;
    - Pangaea;
    - Marine Recorder → MEDIN → DASSH;
    - Capable of taking lots of images? (→ BoDe Pangaea);
    - Accessibility / download?
- JNCC diagram linking data infrastructure → Marine Monitoring Platform guidelines for ROV[19] and AUVs[20].
- Use the MEDIN data ingestion standards for images, annotations, metadata to develop a data submission standard for BIIGLE derived data (then BIIGLE data can be submitted to MEDIN).
- Where are the data / images currently – how managed by each agency / organisation?
- Define why we want / need an image data management system. Who will own / manage it?
- Improvements: what is the future of Marine Recorder?
- What data to archive? Metadata, raw, processed data?, annotations, summary.
- Is there ONE place to archive data / should there be?

[19] ROV guidelines: http://jncc.defra.gov.uk/page-7612
[20] AUV guidelines: http://jncc.defra.gov.uk/page-7618

ADDITIONAL THEME: Linking with marine industry, developers, Department for Business, Energy and Industrial Strategy and Marine Management Organisation

Discussion

- Define 'end user'.
- Define industry roles: contractor, operators, types (aquaculture, fisheries, mineral extraction).
- Release of data as a condition of regulated licenses (including for exploration / investigation) and after appropriate period (i.e. 2-5 years).
- Clarifying with industry that data and engagement isn't that sensitive and not 'likely' to be used against them at the regulatory level. Transparent requirements not guidelines support industry green credentials.
- 1st engagement and linkages with regulators, with stakeholders & define responsibilities / accountability.
- Engagement with regulators and industry with latest / upcoming methods.
- Simplified and robust policy briefings for open publication.
- Define levels of information required for hierarchy of purposes as anonymised data may be released sooner.
- Checklist for regulators on requirements in specifying and granting licenses.
- Example of license versus guidelines in marine mammal mitigation with wording in licence being minimum so tend to only be license that is complied with rather than intention of guidelines and licensing.
- Not regulated = not done.
- Developers / users link to the Marine Imaging Workshop.

ADDITIONAL THEME: New tech pipeline (refers to scoping and validating new technology)

Discussion

- New tech how to compare results with historic data – time series monitoring.
- New tech not available to smaller consultancies.

AI

- Opportunity for research to interact with developers in creating algorithm.
- Cost increase vs. saving.
- Important to know the limitations of 'R.' data manipulation software.
- *X-prize* / *Google* / *Microsoft* research programme.

3D photogrammetry (structure from motion)

- Data heavy (>Terabytes of data) – is it needed?
- Good for monitoring physical impacts?
- Good for public engagement.
- Needs well featured seabed.
- Opportunity to develop machine learning.

Video mosaicing

- Great for training.
- Oblique angle best.
- Doesn't work well on low feature sea beds.

Storage capacity

- Big data (opportunity and risk) and statistics.
- Cloud – security, transfer speeds.
- More use of UK Research Institute tools – Archerz, etc.
- Who is developing approaches to deal with new tech data?

New formats

- HD → 4K → 8K
- Extract stills from video.
- Comparative software.

Laser profiling

- E.g. Cathyx Ocean, stills from video, backscatter / forescatter reduction; R2 robotics.

Underwater human submersibles

- With millionaires – piggyback on their cruises?
- Cheap cameras with volunteer divers, improve coverage inshore areas, collect evidence of impacts / damage.

- Most recreational divers lack training in biosurveys – could this be addressed? Citizen science project?.

Low cost ROV (e.g. BlueROV2)

- Reduced kit cost vs. quality / ruggedness work in challenging conditions?
- Positioning micro-USB.

Virtual Reality

- Can it help with analysts?
- Can help with public communication.
- Can increase ROV accuracy.
- Resolution still not great.
- Motion sickness.

Acoustic imaging

- E.g. ARIS / DIDSON
- Limited coverage.
- Work in turbid water.
- Quantitative assessment relief.
- Ground-truthing tool.
- AUV swarms still at R&D stage.
- Forward compatibility of data storage & programmes – how would you now read a 7" floppy disc? A document in AMI Pro?

Action

- Produce a short state of the art vs future tech comparative and feasibility report / table.
- Contact equipment developers and see what they have planned.
- Potential to transfer knowledge and skills between developers and users.
- Which sectors benefit most from each type of tech.
- Maintain contacts with developers and researchers.

ADDITIONAL THEME: Culture around data

Discussion

- Link to training components of: image analyst training; image annotation & machine learning; data management.
- Action #1: Virtual spaces for asking / answering questions: Google groups; Slack; Listserve email list.
- Action #2: Find examples of where data sharing had a good impact / outcome.
- Action #3: Building a 'community of practice' around benthic image data – real-world / in-person workshops… hands on "how do I do…"
- How to encourage openness around data / sharing?
    - Ensure recognition / citation of data source if used.
    - 'Publish' data → recognition / citation.
    - Attempt to dispel fears around using other groups' data.
    - Different levels of open: commercial / in confidence → open academic / public funded.
- Story telling / narrative – why should I be open / contribute → 'bigger picture' contribution.
- Much research into data culture happening in U.S.

7. Overarching challenges and purpose of imagery

Two final topics for discussion and contribution were:

- The overarching challenges of improving benthic imagery – including developing an understanding of how the various issues all fit together.
- The purpose of imagery. This topic came up at various points during the workshop, with the point being made that understanding purpose is key to shaping the nature of acquisition and analysis.

Relevant outputs are shown below.

Overarching challenges

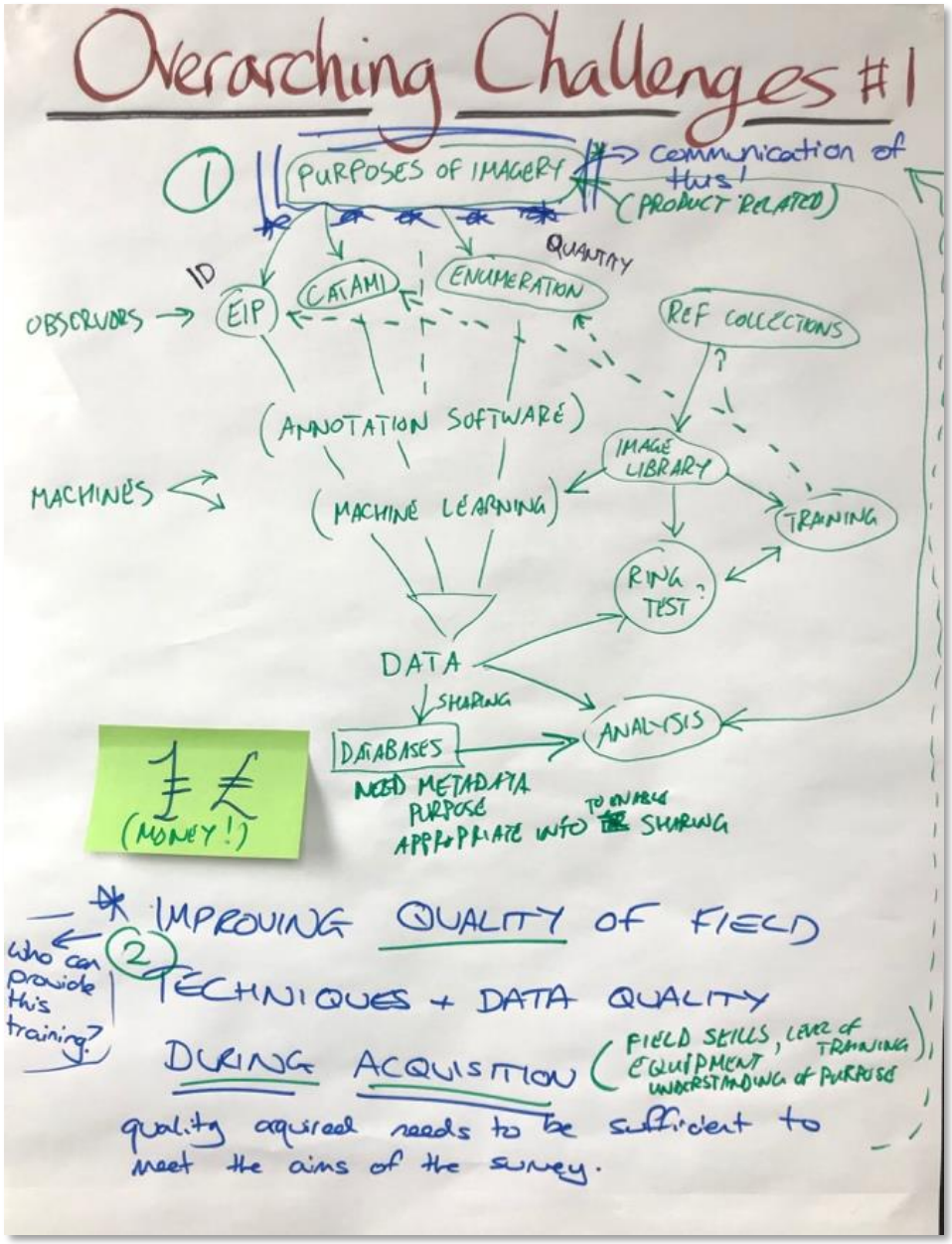


Figure 5. Overarching challenges – before additional overlays.

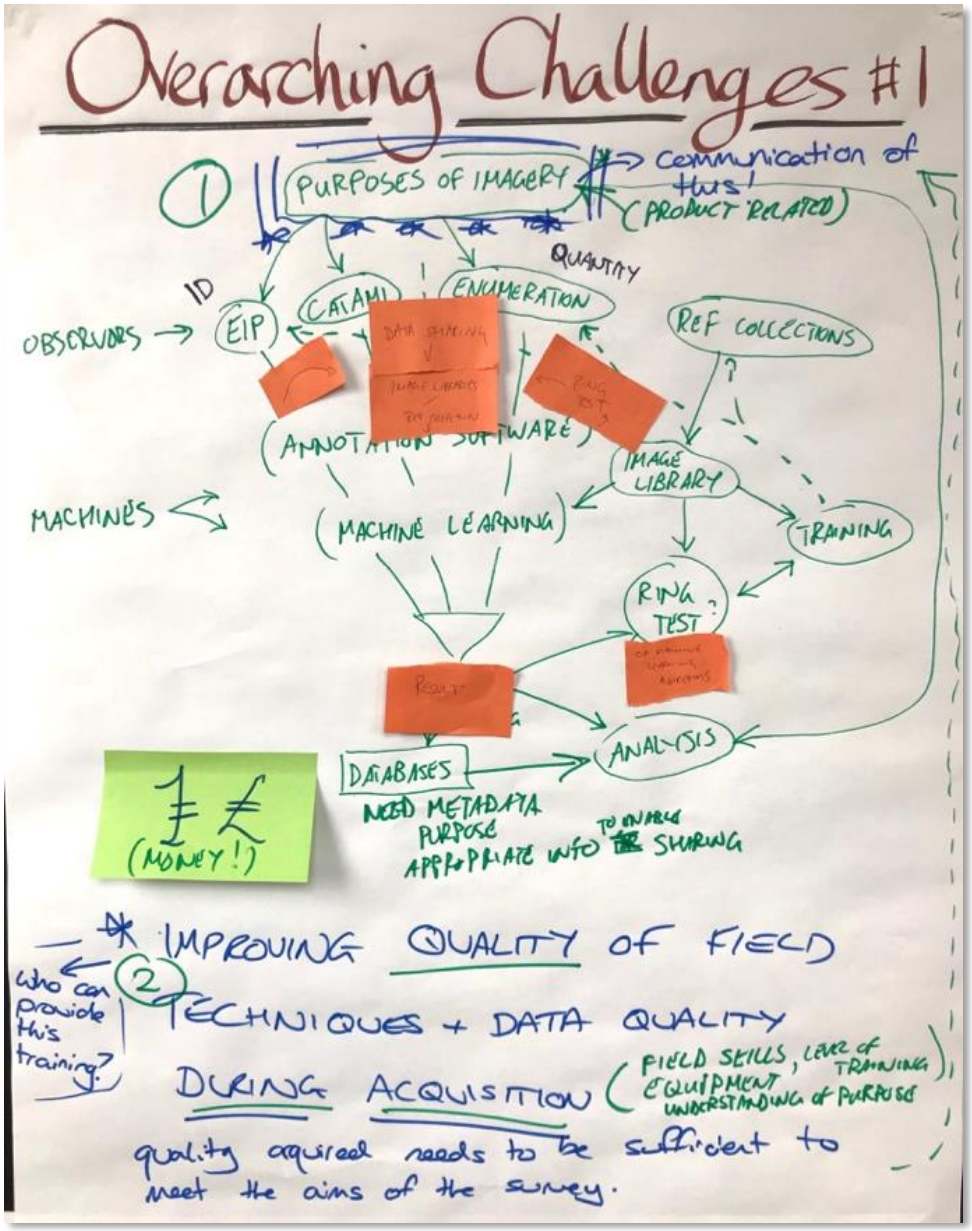


Figure 6. Overarching challenges – after additional overlays.[21]

Accompanying notes:

- 1. Why – what is the purpose of the survey, techniques and analysis type? What do you want to know? Know what analysis you want to perform before you step foot on a vessel! (/ go diving, etc.) But, horses for courses!
- 1a. If minimum requirements are to be introduced, these must be for specific purposes / job-by-job basis.

[21] Post-workshop clarification / suggestion: Annotation software is not a machine. Not everything needs to go through machine learning, and it should probably form its own separate loop from image / libraries / reference collections which then feeds back into enumeration (which should probably go where machine learning currently is).

Purposes of imagery

ACQUISITION	ANALYSIS
MPA monitoring	
Broadscale vs finescale purposes	
Research	
Testing novel technologies and techniques	
Publicising – public engagement, education ——————→ Funding	
Marine licensing	
Benthic monitoring (also wider sea)	
Ecological status	
Change (trends)	
Infrastructure inspection	
Search and recovery (MAIB investigations, kit recovery, etc.)	
Advice on operations	
	Impacts of activities
	Effectiveness of management
Impacts of unplanned incidents e.g. spills, illegal fishing events	
	Identify new species
Training	
Marine archaeology	
Heritage	
Technique / technology / method development	
Industry ←——————————————————————	Communicating analysis standards required by regulators and all steps for consistency
EIAs	
←————Stock assessment————→	
Invasive species	
Climate change effects	
Impacts and potential impacts of fisheries (inshore) – assessment of impacts from fishing – identification of 'habitat' types plus assessment of sensitivity of these habitats to various pressures from fishing	
Impacts from other activities, especially where there is the possibility for 'in combination'	
Production of 'outreach' material to wide audiences	

Table 5. Broad purposes of using imagery (from flipchart – see Figure 7 below)

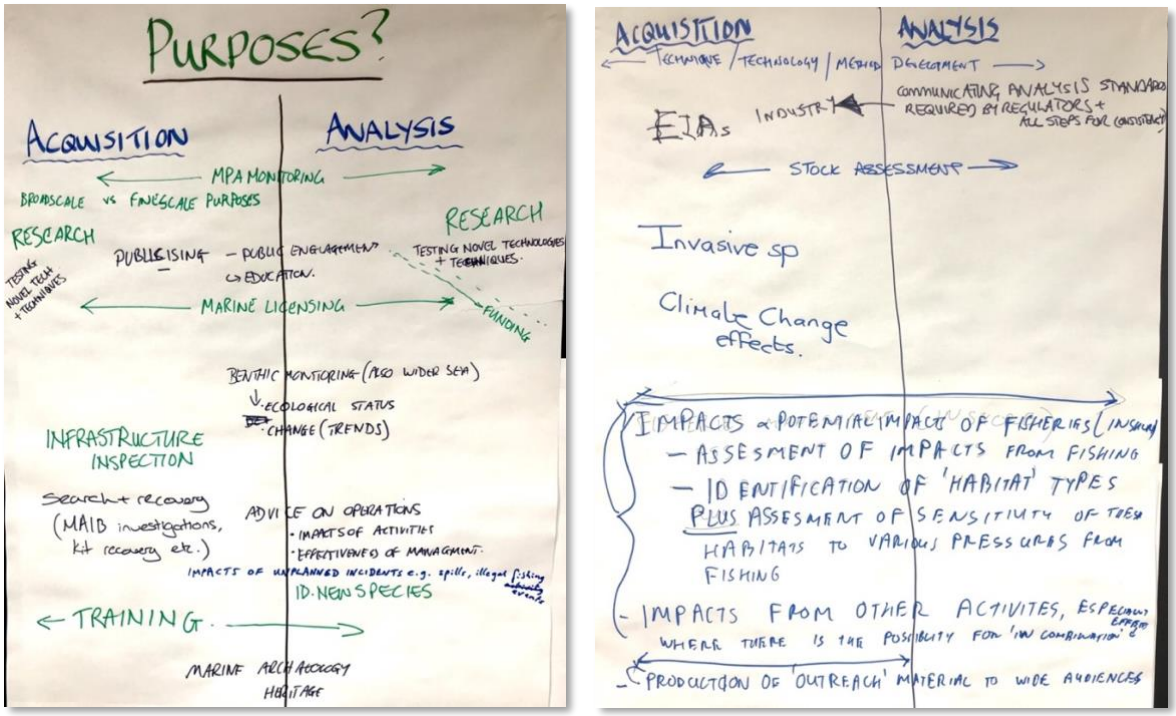


Figure 7. Broad purposes of using imagery (from flipchart).

8. Next steps

At the close of the workshop and within follow-up correspondence, next steps were confirmed as follows:

- Production of this workshop report, including opportunity for workshop participants to comment;
- Ongoing work by the Plan Development Group to begin synthesising workshop outputs into an emerging Benthic Imagery Analysis Action Plan, for further consultation and discussion;
- Ongoing communications with workshop participants and other interested parties from JNCC and the Plan Development Group, via a benthic imagery mailing list (following permission to share contact details);
- Investigation by JNCC of options for an online storage hub, to share workshop presentations, report and any other relevant future outputs;
- Development by JNCC of a short workshop webinar, following publication of this report, to share key discussions and findings with interested parties who could not attend the workshop;

Closing workshop words from Plan Development Group members

"I've been amazed by the information we have all collected here. It's allowed us to get a sense of the breadth of the issues, the overlapping topics and the connections coming out. The collaboration side of things is an important thing I am taking away. There are so many people with years and years of experience in this room, and we'd love if you carry on engaging with the process as we develop the plan – and perhaps reconvene in the future."

"I came here with a massive expectation and over the past few days I've got a sense of your enthusiasm, and of the overwhelming scale of what needs addressing! But we've come together, the conversation is progressing, and some possibilities are emerging. It's very much to be continued. Thank you everybody here for being so forthcoming and we look forward to continuing this work."

Other events of interest

3rd Marine Imaging Workshop, 24-28 June 2019, British Columbia, Canada

19-21 March 2019, Birmingham, UK

Appendices

Appendix 1: Abbreviations and acronyms

AFBI: Agri-Food and Biosciences Institute

AI: Artificial intelligence

API: Application programming interface

ARIS: Architecture of Integrated Information Systems

AUV: Autonomous underwater vehicle

AvTD: Average taxonomic distinctness

BEIS: Department for Business, Energy & Industrial Strategy

BIIGLE: Bio-Image Indexing and Graphical Labelling Environment

BODC: British Oceanographic Data Centre

BRUV: Baited remote underwater video

CATAMI: Collaborative and Automated Tools for Analysis of Marine Imagery

CCZ: Clarion Clipperton Zone

CEFAS: Centre for Environment, Fisheries and Aquaculture Science

CIESM: Commission Internationale pour l' Exploration Scientifique de la Méditerranée (Mediterranean Science Commission)

COCoast: Capturing Our Coast

DAERA: Department of Agriculture, Environment and Rural Affairs (Northern Ireland)

DASSH: Archive for marine species and habitats data

EA: Environment Agency

EIP: Epifauna Identification Protocol

EUNIS: European Nature Information System

FoV: Field of View

FSC: Field Studies Council

HBDSEG: Healthy and Biologically Diverse Seas Evidence Group

HBDSEG BSG: Benthic Sub-Group of the Healthy and Biologically Diverse Seas Evidence Group

HD / SD: High definition / standard definition

HOV: Human occupied vehicle

hROV: Hybrid remotely operated vehicle

ICES: International Council for the Exploration of the Sea

ICES CRR: ICES Cooperative Research Reports

ICES WGMHM: ICES Working Group on Marine Habitat Mapping

ICES WGNEPS: ICES Working Group on *Nephrops norvegicus (Nephrops)* Surveys

ICES WKNEPS: ICES Workshop on *Nephrops* burrow counting

JNCC: Joint Nature Conservation Committee

LIMS: Laboratory Information Management Systems

LISST: Optical Laser diffraction instruments

MarLIN: Marine Life Information Network

MarPAMM: Marine Protected Area Management and Monitoring

MBA: Marine Biological Association

MBARI: Monterey Bay Aquarium Research institute

MEDIN: Marine Environmental Data and Information Network

MESH: Mapping European Seabed Habitats

MMG: Marine Monitoring Group

MMO: Marine Management Organisation

MNCR: Marine Nature Conservation Review

MPA: Marine protected area

MPAG: Marine Protected Areas Group

MPATG: Marine Protected Areas Technical Group

MSC: Marine Stewardship Council

MSFD: Marine Strategy Framework Directive

MTU: Morphological Taxonomic Unit

NBN: National Biodiversity Network

NE: Natural England

NHM: Natural History Museum

North-East Atlantic Marine Biological Analytical Quality Control (NMBAQC)

NOAA: National Oceanic and Atmospheric Administration

NOC: National Oceanography Centre

NRW: Natural Resources Wales

OBIS: Ocean Biogeographic Information System

OSPAR: Convention for the Protection of the Marine Environment of the North-East Atlantic

PDG: Project delivery guidance

PT: Proficiency testing

QA: Quality assurance

QC: Quality control

RHIB: Rigid-hulled inflatable boat

ROG: Recommended operating guidelines (MESH)

ROV: Remotely operated vehicle

SACFOR: Abundance scale (Superabundant, Abundant, Common, Frequent, Occasional, Rare)

SEPA: Scottish Environmental Protection Agency

SISP: Series of ICES Survey Protocols

SNCB: Statutory nature conservation body

SNH: Scottish Natural Heritage

SOP: Standard operating procedure

SQUIDLE: Web-based tool for managing, exploring and annotating images, video and largescale mosaics

UKRI: UK Research and Innovation

UWTV: Underwater Television

VARS: Video Annotation and Reference System

VIAME: Video and Image Analytics for a Marine Environment

VME: Vulnerable marine ecosystems

VR: Virtual reality

WFD: Water Framework Directive

WoRMS: World Register of Marine Species

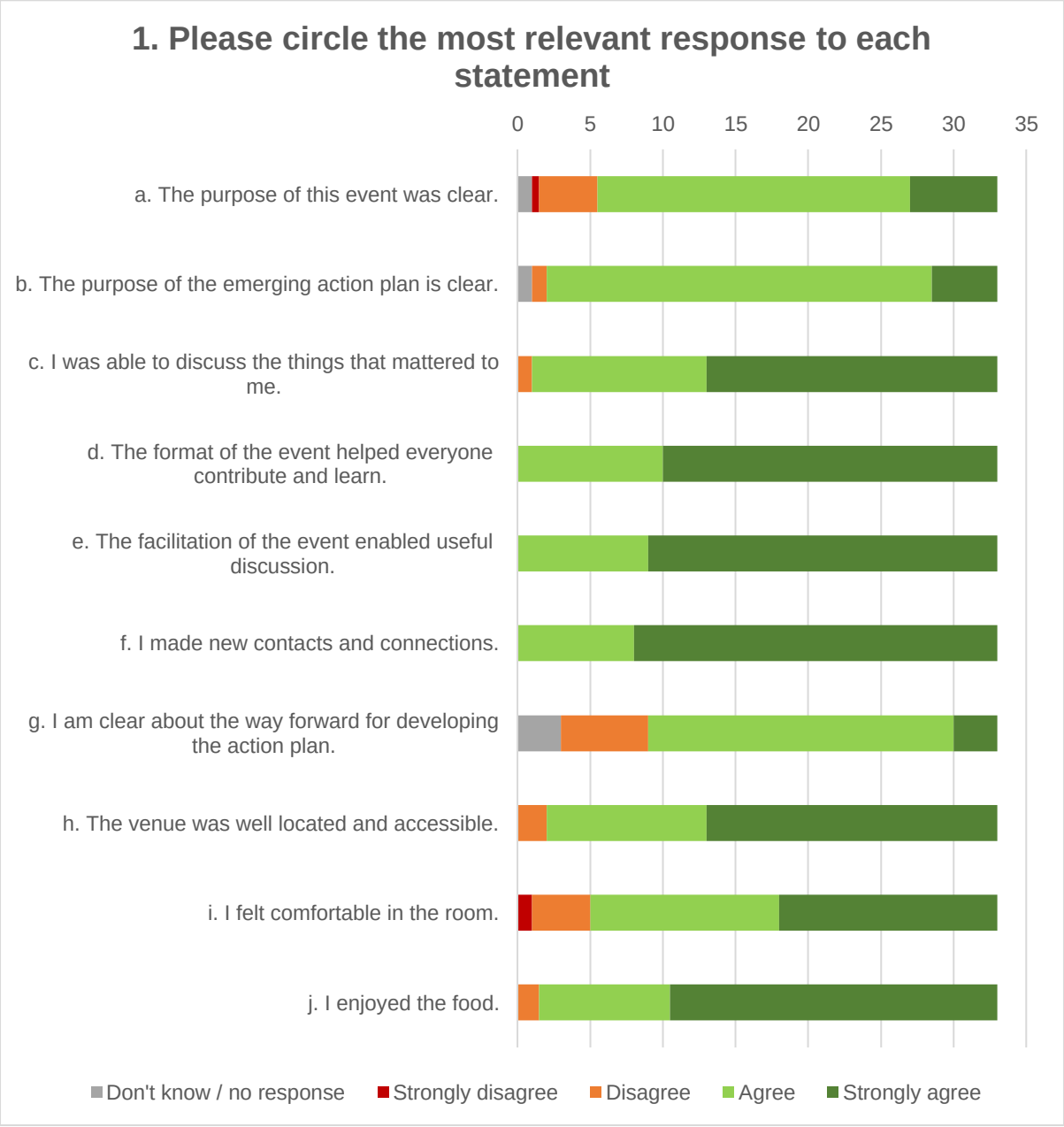
Appendix 2: Workshop attendees

Members of the Plan Development Group are marked **like this***

Name	Organisation
Ade Weetman	MSS
Alison Benson*	Envision Mapping Limited
Alison Tamkin	APEM
Ámundur Nolsø	Mar Bio Lab Faroes
Amy Cartwright	Plymouth University
Andrew Gates	NOC
Anna Downie	Cefas
Berta Ramiro-Sanchez	Edinburgh University
Charlie Lindenbaum*	NRW
Charlotte Johnston	Crangon Limited / SNH
Clara Alvarez	DAERA
Clive Fox	SAMS
Eloise Boblin	Envision Mapping Limited
Emma Gerrie	SEPA
Fionnuala McBreen	JNCC
George Graham	MBA
Hayley Hinchen*	JNCC
Henk van Rein*	JNCC
James Albrecht	JNCC
Jen Durden*	NOC
Jen O'Dell	Seastar Survey Limited
Joe Turner	Ocean Ecology Limited
Jon Hawes*	Cefas
Jon Moore	ASML
Karen Boswarva	SAMS
Karen Webb	JNCC
Katie Pryor	EA
Laura Bush	Fugro GB Marine Limited
Leigh-Ann Baker	Fugro GB Marine Limited
Matt Curtis	Cefas
Matt Green	NRW
Michael Thompson	Gardline
Mike Young	NE
Myles O'Reilly	SEPA
Dylan Todd	NE
Nicola Foster	Plymouth University
Oliver Bittner	Wageningen Marine Research
Phil Newman	NRW
Philip Boulcott*	MSS
Rachael Eyley-Roberts	Gardline
Rohan Holt	Cloudbase Productions
Ross Griffin	Ocean Ecology Limited
Samantha Hormbrey	Eastern IFCA
Stephen Thompson	Eastern IFCA
Steven Dewey	Seastar Survey Limited
Tabitha Pearman	NOC
Terri Souster	Ulster University
Tim Mackie	DAERA
Udo van Dongen	Bureau Waardenburg BV

Appendix 3: Workshop evaluation

Respondents: 33



2. Please use the space below to provide more detail about any of your responses

Purpose

- a. (disagree / agree) Not at the start.
- I felt that we could have outlined the issues and need for conference better – we with MPA know our issues, these may not be obvious to others.
- a. (agree) Purpose was clear, but not detailed – this may not be important. Good job – well done to all. A big job, but positive steps have been made.

- a. (disagree / agree) Re ‘agree’: After the initial “why are we here” talk. The purpose of the event became clearer after the first day – otherwise a little unclear as to the extent of the workshop. Access within Birmingham was great but parking proved a challenge to some.

Action plan (purpose and way forward)

- The only responses I have not “strongly” agreed with [b/g] is because much of this has to emerge after the workshop themes and issues are fully digested and synthesized. After the workshop summary report – I’m sure this will be clarified and understood this is the process.
- The next steps could have been a bit more clearly defined – having a clear platform for communicating progress would make it feel like we achieved more and keep the momentum going.
- I hope the resultant report will provide details of how we plan to take this forwards as there are a lot of unanswered questions with no one currently tasked to tackle and no time line.
- g. (disagree / agree) Some issues, not others.
- g. (agree) Not sure what happens after the action plan has been developed.

Scope

- The topic is very broad. It was not clear what aspects were to be covered during the course of the workshops. Perhaps attempting to cover too much?
- It was a great workshop but I felt it tried to achieve too much which made its focus a little unclear.

Format and process

- I felt the workshop overall was very well structured and valuable. I feel that real tangible work was done.
- Very well executed, ‘marketplace’ time was potentially too stretched out.
- At times progress did seem a bit slow – possibly could have completed work in 2 days.
- An agenda / indication of format and desired output provided well in advance would have allowed participants to be better prepared and understand the outcome.
- Maybe an opportunity for people to throw to the group one of their concerns and the groups can offer ideas, solutions, options. Great workshop given me a lot to think about. Thank you.
- Plenty of questions, not so many answers. Unsure about ability to influence some of the required outcomes.
- As someone very new to visual analysis I would have liked the chance to discuss some more basic aspects of visual analysis. It was all very overwhelming and the discussions while useful were for complicated matters I’m less familiar with. I feel like basic questions I had haven’t been answered but I now have made contacts I can communicate with to address my questions. Not beginner friendly.

Facilitation

- Having a facilitator who doesn’t have their own stake in the issue helped keep a very enthusiastic, yet diverse group on track without getting stuck on minutiae.
- It was well organized and facilitated.

- The venue and facilitators meant that the workshop was diverse and interesting and provided a suitable format for everyone to have their say.

Venue / room / food

- My only misgiving is that the venue did not cater well for dietary requirements (vegan / gluten free) even though they were made aware in advance. They did put together.
- h. (disagree/agree) Was dependent on travel method & ability to add more parking whilst in plenary. j. (agree) Didn't know there was vegan food available, just thought I had to make do with vegetarian.
- h. (agree – if by train; disagree – if not). Central Birmingham difficult for non-train commuters and timing for these further located necessitated staying overnight and no parking nearby.
- Room, bar today, was too hot and stuffy. Food was good but poor awareness of dietary requirements – distinct lack of gluten free snacks / fruit and vegetables.
- The acoustics made it hard for me to follow conversations in some occasions.
- Very few negative issues – 1 minor one: acoustics in the room variable and sometimes difficult to hear people speak.
- Meeting room too warm. Were not allowed to open doors to cool down a little.
- Room was WAY too hot.

Participants

- Wonder whether we should have invited developers? Certainly engage as soon as possible from now.

3a. What did you most like about the workshop?

Participants / networking

- The relevance of delegate experience and expertise.
- Contacts made and chance for networking.
- Chance to engage whole community.
- Networking with a bunch of enthusiastic and interesting people.
- Meeting others.
- Excellent networking op and well designed to all share opinions.
- Networking time to discuss.
- The high level of communication with a wide breadth of disciplines within the same field.
- The breadth of participation from different stakeholders i.e. industry, SNCB and academia.

Overall format and approach (including interactivity)

- The format of the workshop worked very well.
- Structured format. Provided balance between opportunity to contribute and need to keep moving forward.
- I thought the fairground / marketplace rotation style and small groups worked really well.

- Good mix of interaction and presentations, & small groups to discuss key issues worked well.
- Interaction (active rather than passive).
- Interaction between participants.
- Format was relaxed yet made people interact.
- Moving around, not sat listening. Interactive. Great discussions. People with mixed experience, some in acquisition, some in capture. Lots of enthusiasm.
- Very interactive and allowed me to express opinions.
- Open minded approach.
- Scope for plenty of discussion and sharing of ideas.
- Great opportunity for voicing thoughts and shaping areas of discussion. Raising profile of issues important to me.
- Open exchange of ideas.
- Ability to mix and meet different people and discuss / learn new information.
- It was good that we were all able to discuss / listen to [?] that were both presenting strengths and weaknesses.
- Not sitting all the time. Speaking to so many (new) people.
- The format of short bursts of concentration on issues and interactive format.

Detailed process

- 'Diamond' structure of idea expansion, tree-flow discussion, and then distillation.
- Good presentations.
- Breaking up of groups and having to achieve group outputs with different people.

Facilitation

- The fact that the organisers were free to partake due to Richard and Helen's attendance.
- Facilitation excellent!
- Richard of course!

Venue / refreshments / logistics

- Venue.

3b. What would you have changed?

Pre-workshop

- Pre knowledge of the event and a more detailed scope issued beforehand to enable pre-thoughts to form.
- Would prefer to have had a participant list distributed before meeting started. This would help us find and meet most relevant people more quickly.
- Scope – too big and unclear beforehand. Introduce everyone and their experience before starting (could be done online / using a short profile in advance / in person).

Participants

- Some intro about how all the participants are using imagery for benthic monitoring.

- More time to know others' experience and background, e.g. met someone in last 30 mins who would have been useful to have known on first day.
- Would have been good to have regulators involvement.

Content

- Less themes.
- More photos / video!
- More introduction / initial explanation may have proved useful.

Overall format

- 2 days.
- A little lacking in structure. Overuse of "the marketplace".
- Later start (post rush-hour!) / finish.
- A couple of introductory presentations to introduce the topic and what JNCC wanted would have helped for most.

Detailed process

- Maybe allow people with similar analysis to discuss issues together, rather than having to hunt for people to discuss issues with.
- Nominated leads for topics.
- Forced to mix groups up. Facilitated work area discussions and transcribing.
- Slightly shorter sessions during day & 2. Possibly focus in on the clear directions and advances that emerged during day 1. More progress could have been made.
- Perhaps drawing together some of the themes earlier (where there were strongest overlaps). More time to define theme headings at start – some ambiguity arising from different interpretation of theme headings.
- Potential time limits on each theme. More switching between themes may have been beneficial.
- In some cases the time to mull over things when circling round the room could have been just a little longer to be able to better process the issue. Not much longer though to avoid getting stuck.
- At the onset, identify where the "market stalls" were of overlapping topics and combine to produce fewer "stalls" = more time at each = easier to move around them all.
- Narrower remit – focus on fewer points / areas, complete this then move on to next area – reducing task / challenges for PDG.
- Did seem to go over and over same issues which got a bit decreasing benefit over time.

Venue / refreshments / logistics

- Acoustics in the room not great.
- Decaf coffee please!
- Nothing about the workshop, only the food (see comment above)!
- Parking was expensive.
- Venue w/ parking (or make it clear that you needed to pre-book).
- Seating at dinner.

Other

- Word count ‘germane’ ☺
- Zilch!

Appendix 4: References

The following is a list of references and contacts mentioned by workshop participants in relation to some of the themes (none were listed for themes 5, 6, 10 and 12). Several participants provided the names individuals and / or contact details. Any names and personal details have been removed from this report, but have been noted by the Plan Development Group.

Theme 1. Need for improved levels of method standardisation and quality control

- NMBAQC collection & interpretation guidelines – pilot ring test reports (http://www.nmbaqcs.org/scheme-components/epibiota/reports/).

Theme 2. Taxonomic identification

- WoRMS (http://www.marinespecies.org/).

Theme 3. Image annotation software, use of machine learning… and improving cost and efficiency…

- Kera Platform machine learning API keras.io
- Links TensorFlow capabilities. CNTK. Theano.
- Ocean.soton.ac.uk
- Gomes Pereira et al 2015 → Comparison of image annotation software (Progress in Oceanography).
- Perspectives in Visual Imaging for Marine Biology and Ecology by Durden et al (2016). Oceanography and Marine Biology an Annual Review Vol 54, pages 1-72.
- Article: Piechaud et al (in press), Automated identification of benthic epifauna with computer vision. Marine Ecology Progress Series. https://www.int-res.com/prepress/m12925.html

Theme 4. Development of standardised / specialised sampling approaches

- Epibiota Remote Monitoring from digital imagery: Operational Guidelines (http://www.nmbaqcs.org/scheme-components/epibiota/).
- Monitoring guidance for marine benthic habitats. Noble-James et al, 2017 (http://jncc.defra.gov.uk/page-7336).
- Underwater Photography Masterclass book (https://www.amustard.com/books/).
- AUV for use in marine benthic monitoring. JNCC, 2018 (http://jncc.defra.gov.uk/page-7618).

Theme 7. Sampling units

- Article: “Perspectives in visual imaging for marine biology and ecology” (2016) in: Oceanography and Marine Biology: An Annual Review, vol: 54, pages 1-72, by Durden et al. → example of stats needed.

- Article: “Megafaunal variation in the abyssal landscape of the clarion clipperton zone” (2019) in: Progress in Oceanography, vol: 70, p.119-133, by Simon-Lledo.

Theme 8. Image reference collections

- Driver – identification; classification of abundance?
- ICES – WGNEPS – SISP for *Nephrops* TV surveys; cooperative research report #340. http://prep.ices.dk/sites/pub/Publication%20Reports/Cooperative%20Research%20Report%20(CRR)/crr340/CRR340.pdf
- OBIS – online database. https://obis.org
- Deepsea ID, e.g. https://deepseacru.org/2016/12/16/deep-sea-species-image-catalogue/ and the Deep Sea ID App: http://www.nhm.ac.uk/our-science/our-work/biodiversity/deep-sea-systematics-ecology-group/world-register-of-deep-sea-species-app.html
- WoRMS (http://www.marinespecies.org/).
- MBARI Deep-sea Guide (UK?) – online. http://dsg.mbari.org/dsg/home
- International Seabed Authority Guide – online. https://www.isa.org.jm/documents-resources
- Serpent Project Guide – online. www.serpentproject.com/publications
- Fish Base. www.fishbase.org
- UYIC / DNC Guide.
- Amon et al article – in-situ photos and specimens (CCZ / UK-I claim).
- Rogacheva et al mid-atl ridge – in-situ photos, holothurians.
- JNCC image catalogue / (Cefas?).
- Shallow water – local diver knowledge; social media groups.

Theme 9. Observer consistency

- “Comparison of image annotation data generated by multiple experts for benthic ecology” by Durden et al (2016), Marine Ecology Progress series, vol 552, p61-70.

Theme 11. Morphological classification systems

- Catami.github.io → classification guide paper = Althaus et al, 2014.
- Deep sea ID catalogue = OTU database (University of Plymouth) – https://deepseacru.org/2016/12/16/deep-sea-species-image-catalogue/

Appendix 5: Transcript of workshop discussions, by theme

Theme 1. Need for improved levels of method standardisation and quality control

Questions

- How deal with changing aims over time?
- What has changed (or not)?
- How much 'change' is due to inconsistencies between analysts?
- To "waft" or not to "waft"?!
- How do we achieve quality when the people collecting the images have no vested interest in the results?
- Limit the number of points of interest but have multiple reviews of same footage?
- Need to define and record scope
- Or other differences of scope (platform, lights, season, etc).
- Added complexity lower throughout.
- 'Live' method guidelines online.
- Formal vs. informal QC. Cost drivers.
- Who would run an external QC scheme?
- And how would it work in regards ↑ time pressure deadlines.
- Different ring tests for different purposes / areas?
- Can we develop a cost effective QA module for video / stills analysis?
- What standards already exist.
- "Growth form" ring test also required.
- Can reference collections be used for ring test?
- Is shared reference collections sufficient? " Calibrating analyst."
- What decisions have you made in your project?
- How to create a standard that is applicable across different organisations, projects etc?
- What is the aim of the survey?
- Impacts on long term datasets?
- What methods are used?
- Can different techniques be 'standardised'?
- Do we have enough knowledge to standardise (in some areas)?
- Is method standardisation really what we need? Should it be standardisation of data products / assessment?
- Are Go Pro's good enough? Distortion!

- Is ‘good enough’ good enough?

Key issues

- How do we assess quality?
- Do we adopt formal standards?
- *Can standards be flexible? Focus on outputs rather than methods.
- How do we ensure standards are affordable? E.g. equipment – focus on results / outputs.
- Many standards within 1 quality assurance?
- Best practise docs for acquisition, analysis, QC required / “flesh out”
- Collaborative iterative QC process (online helps? Preferably in advance).
- Producing data in line with data standards e.g. MEDIN / DASSH. [Standard defined so that data usable in the future.]
- Who would run external QC / ring tests?
- Cost implications.
- Pass / fail too restrictive? [Specific to certain objectives, e.g. bronze / silver / gold standard.]
- Procedures and guidance rather than prescription and testing.
- Project / staff time / expertise for QA/QC.
- Standards provided in accordance with survey aim. Separation of ‘analysis’ and ‘collection standards.
- Set aspirational levels for survey ‘kit’ – needs tailored & with realism, i.e. lighting, camera image size. [Kit always evolving so not good idea to specify too tightly.]
- Link reference collection portal / database to ring tests, use ring test as peer reviewing of images to reduce costs.
- Online test for species ID.
- Lessons learned review of analysis procedures and outputs – metadata – to feedback to future methodologies and scoping of work.
- Not a focus on method standardisation but standardisation of data outputs.

Challenges

Already covered:

- A. A good start for getting good quality images will be having someone on the survey who has a vested interest in the results.
- C. Matching standardised methods to defined purpose / aim

Linked to / addressed in other areas:

- E. What are the connections to other themes? Many! (e.g. B.)
- “Guidelines” need to be sold as “best practise” by regulators to be embraced by industry. Theme #5.

Vague:

- What are the options for addressing the key issues?

Next steps

- Re: data / image / meta-data standards… adopt existing from other communities or develop our own?
- Task 1 should be to define the 'purposes' the standards need to be fit for.
- Which standards are needed first, and why? How will they help?
- Identify what the standards will cover, e.g.:
    - Quality of image (link to D.)
    - Reliable method of scaling.
    - Minimum area to be covered.
    - Minimum number of replicates.
    - Time stamp.
    - Survey window (DATES).
    - Positional precision needed.
- Linked to section 4, 7, 9 & 10 (E.)

'The big task'

- How do we achieve a functioning QA scheme for imagery data?

Addressing key issues

ISSUE: Need for improved ~~method~~ output standardisation and Q.C. output

Discussion

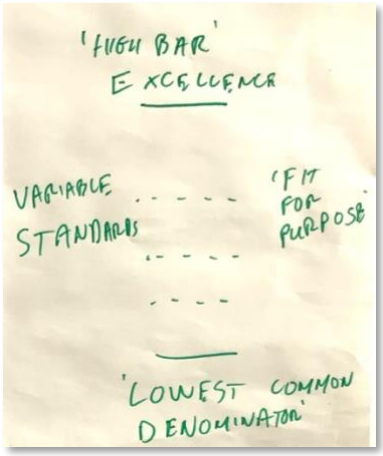
- Outputs need to be fit for purpose! Hence different standards for different purposes.
- Variable standards have hugely different cost implications, Over-prescription of method is counter-productive.
- How do we ensure standards are affordable? Commissioning organisation / CMA fund / contribute to cost of QC scheme.
- Standardise outputs for acquisition, methods for processing.

Connections

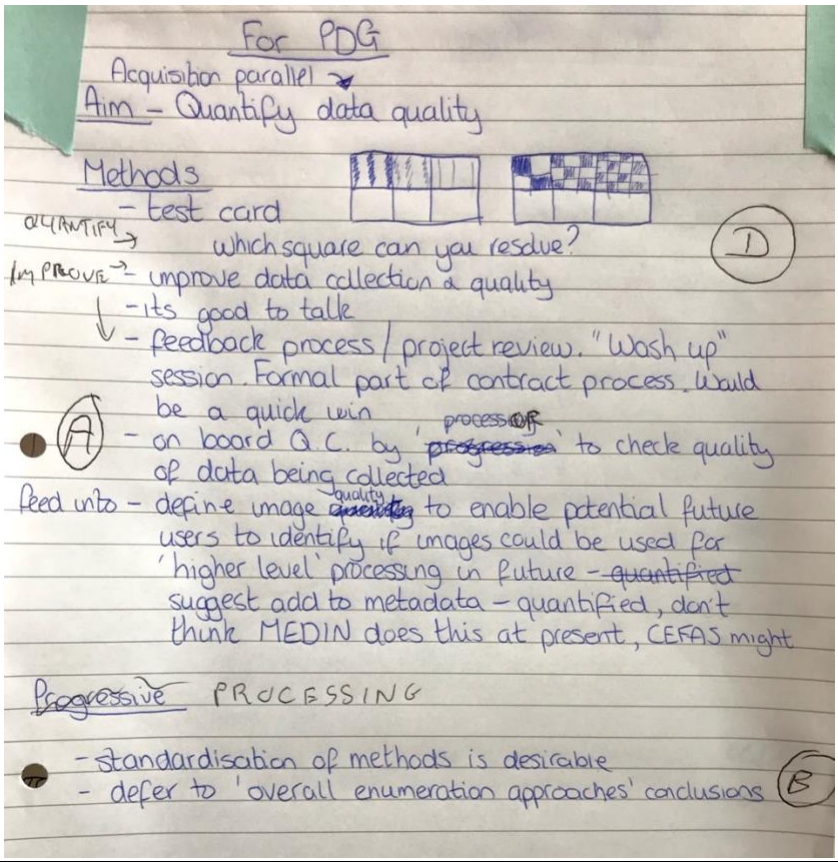
- Flexible standards – yes focus on required outputs – these can be achieved via guidelines rather than rigid procedures.
- Affordable standards – needs greater clarity in contract specifications to allow realistic tenders e.g. potential for high % downtime.
- On-board QC stage – are survey conditions appropriate for data acquisition or go to downtime?
- Issues of time-constraint / conflict of interest on survey.
- Different outputs at acquisition / and data interpretation.

Relevant notes / challenges

- Hierarchy of analysis – which level do you get to in your analysis?
- 'Repurpose images' – identify if they can be used for other purposes ⇒enable search later.
- Need standard to define if image quality is fit for purpose.
- Streamline data-processing, making it possible to 're-process' old data in the future.
- 'Standards' should be aimed at achieving compatibility.
- Do we need standards or recommendations?



Action / advice for PDG



Theme 2. Taxonomic identification

Questions

- Use of a shared image reference catalogue / database for regions?
- How do we make it more consistent?
- Fewer experts and new generation of taxonomists?
- Consider time series in taxonomic resolution.
- Published material? (e.g. guides / workshop outputs)
- ID workshops carried out?
- What level is meaningful? (e.g. sensitivity, survey objectives)
- Dependent on data quality and type (and survey strategy)
- Can we move towards indicator species so we don't have to ID everything?
- TDP possible? (Taxonomic Discrimination Protocol)
- Do we need to ID to species level? Project specific?
- How to deal with taxonomic changes?
- Can training of analysis improve consistency?
- How to standardise morphological / lifeforms categories across organisations?
- Lifeforms (e.g. tall epifauna) in parallel to taxonomy for lower quality images – how useful?
- Would traits analysis help with consistency? Is consistency more important than tax resolution?
- Confidence?
- What can you be certain of?
- Consistent meaning, e.g. "Open Nomenclature"
- How to resolve experts with different identities, i.e. lab inconsistency?
- Cost vs quality?
- Do we need a TDP for different imagery quality?
- What level of ID is needed for different uses? Doesn't always have to be to species?
- WoRMS codes?
- Standard lifeform approach → stratified by pressure?
- How do we transfer experience? Workshops, training, etc?

Key issues

- Different stages of analysis – quick rapid appraisal but then make video / images more widely available for detailed analyses.
- Don't be limited by T.I.
- Integrate all T.I. resources → one place (where; who owns; who maintains?)
- Publicise and utilise online reference image libraries.
- Record the scope of your study in terms of taxonomic I.D.

- Connection [geographic] to morphology / morphotypes [and lifeforms] → functionality → hierarchical approach.
- Defining the level of detail → answering hypotheses vs. producing universally applicable data.
- Bring in external experts (e.g. BIIGLE) → check ID.
- How will the data be analysed – potential for greater use of metrics such as AvTD.
- EiP means data intercomparible.
- Central reference collections / forums – should be readily and freely available.
- ID course lists published e.g. every year – to be easy to access.
- More and better ID courses – led by e.g. museum experts rather than just citizen scientists.
- Links with universities and taxonomic-based PhDs.
- Influence UKRI to invest / promote taxonomic training and links to research.
- Beware of trying to be too precise – sometimes family is more appropriate than species!
- Numbers of experts.
- Continuity of knowledge.
- Online training tool.
- Use morphospecies – more informative than going to higher taxonomic level.
- Multiple levels of ID – indicate what you think it is even if not sure.

Challenges

- Increased reliance on 'image annotation' and 'machine learning' will decrease focus on taxonomy training.
- Taxonomy training is key for enabling annotation and machine learning. Particularly for Novel!
- Balancing over-cautious v. over-confident ID.
- Some suggestions (e.g. confidence of ID) should already be in practice but we still have major issues.
- Many good options but what is the plan / steps to implement? (That are financially realistic.)
- Quality of video is better related to the resolving power (e.g. how fine / small you can see) rather than technical quality / disturbance / turbidity etc.
- EIP needs to have x-reference to biogeography. Can warn user that I.D. might be out of range.
- EIP – use links in EIP to I.D. websites & NBN etc.
- Use social media more. Link to Facebook ID groups or have EIP-dedicated F.B. page.
    - [Just a starting point. Should not be used for actual reference.]
    - [Create online forum / communication platform.]
    - [Only for people who actually know what they're talking about.]
    - […and what if you don't want to use F-B?!]

- Naming convention to indicate confidence in ID e.g. “Open Nomenclature”.
- NHM (and other collections) centre of excellence for taxonomy.
- Need to shift focus onto taxonomy so graduates have a basic level of taxonomy knowledge from university.
- Entry level courses e.g. Seasearch for citizen scientists. [Entry level insufficient!]
- Open University taxonomy module to raise expertise. Where to start? Basic taxonomy? Phyla?
- Is the taxonomic expertise available at UK universities to manage taxonomic-themed PhDs? [YES]
- Promote publication of taxonomic papers (academic). [Academic problem with paid for journal articles → Research Gate → contact authors.
- Roadshows. Plasticised marine life collection. Hosted by aquariums. (Sea creatures tour.) [Probably not good engagement → who would attend?]

Addressing key issues

ISSUE A: How to increase the level of taxonomic expertise

Discussion

- Central, easily accessible reference collections and identification forums.
- More and in-depth identification courses, easily available calendar. Webinars / online courses to reduce costs. Hosted by experts rather than citizen scientists.
- Habitas but better!
- Funding for PhDs, linked to taxonomic analysis of previously collected imagery.
- Production and distribution of EIP → stop at level you are certain of (TDP style).
- Engagement with universities, museums etc – funding?
- Subsidise training courses.

Connections

- 8. Reference collections.
- 9. Observer consistency.
- 12. Image analysis training.

ISSUE B: How to improve consistency of identification

Discussion

- Reference collections (→ feedback system; location tags; HBDSEG hosted?) identification courses etc.
- Consistency of enumeration techniques.
- Consistency of nomenclature / morphological types.
- Augment analysis with CATAMI or similar hierarchical structures.
- EIP – when to stop! Distribution of EIP to all relevant parties.

- If in doubt, stop at a higher level + indicate suspected e.g. species in qualifier only.
- Consistency (of quality) in acquisition of data.
- Ring tests – problematic.
- Marine Biologist Forum – currently on Facebook.

Connections

- 8. Ref collections.
- 11. Morphological characterisation systems.
- 1. Method standardisation.
- 4. Standardised sampling procedures.
- 10. Enumeration techniques.
- 9. Observer consistency.12. Image analysis training.

ISSUE C: How to improve expertise / consistency

Discussion

- Reference collection to be constructed (based upon existing resources – MEDIN, DASSH, Habitas, Marlin, NOAA, etc).
- Central website (easy to find! – could be JNCC biotope pages for example) containing lots of resources – books, websites etc as well as upcoming ID courses, forums etc.
- EIP to be developed and distributed – should be common knowledge – linking biological expertise with technological expertise.
- Training is key – need all courses to stress that over-confidence is a problem.
- Engagement with universities, museums, international taxonomists, consultancies etc. is a priority. Experts must be involved in training – funding for this may be an issue!
- Ring tests better than OS (own sample?)!! Collaborative, could look at specific groups, indicator species.

Theme 3. Image annotation software, use of machine learning… and improving cost and efficiency…

Questions

- Annotation. Time vs. transparency.
- Use in QA.
- Use of online resources? (BIIGLE, SQUIDLE)
- Automation of image analysis?
- Platform interoperability in-house vs cloud.
- Will this lead to a loss of (human) expertise?
- Quality of data extraction metrics?
- How can we ensure consistency across different types of software being used?
- Is this close to becoming a reality?
- Standardise acquisition for machine learning.
- How close are we to reliable automated species ID & enumeration?
- Data privacy.
- Machine learning as a QC procedure? (Remove rubbish images.)
- Is machine learning there yet for drop cam shallow habitats?
- Will improvements in tech create inconsistencies over time?
- Is this purely a cost cutting measure?
- Could machine learning be used to ID and eliminate poor quality videos / stills – leaving good quality for experts?
- Are there user friendly video annotation software available?
- Development of reference collections.
- AI is time consuming! Maybe manual analysis of less images.
- Importance in QC and training.
- Who is already leading on this?
- May be worth it for specific repetitive tasks – e.g. ID of *Nephrops* burrows.
- How to annotate analogue video? (Without degrading the quality.)
- Could we collaborate (using online resources and BIIGLE / SQUIDLE) to generate enough data to train AI system?

Key issues

- Standardised outputs.
- The ability to share information.
- Good training data for any machine learning model.
- Increase use and awareness of online resources such as BIIGLE and VARS and SQUIDLE and VIAME and BIIGLE.
- Settle on single online tool for all agencies / organisations.

- Code sharing. Open source – github.
- Estimate uncertainty in prediction / classification.
- Central processing → sharing capital assets; wider machine learning.
- Enough images annotated / labelled consistently / accurately.
- Remove rubbish / junk images.
- Should be a priority.
- Cloud based vs local.
- Image analysis already under-valued! [Clarify? Financially / expertise?]
- Appropriate use, not dependency.
- Understand limitations – expert validation necessary!
- Communication between developers.
- Image annotation software helps with issues around truncation in all frequency methods.
- Img.am software creates an audit trail and makes comparison and discussion easier.
- Could we have a workshop on the different tools and how to best utilise them / compare them. [Marine image workshop this year in Canada often has workshops & demos of software.]
- Value for screening large number of images.

Challenges

- ML a step too far for many? Need training in basics (software, storage, management).
- Many still not using annotation software at all
    - using any data base is an improvement (which one is less important) – improves consistency, facilitates randomisation, ease of QA / QC, etc.
    - need to deal with data culture issues that prevent buy-in (issues of ownership, privacy, sharing, etc.).
    - many annotation software are low cost, but implementation barriers exist (e.g. computing systems, lack of awareness of options, software knowledge, time, etc.)
- Imbalance in knowledge / implementation of annotation software / AI across groups.
- Can we have a workshop to teach people about automated annotation approaches?
- Fundamental computing / data handling skills = barrier to adoption.
- Need to sell why → not simply say “not using IA is clearly your problem” → be charismatic.
- Is the cost of development of a system that works justified?? Not convinced yet…!

Addressing key issues

ISSUE A: Machine learning

Discussion

- Recognise number of systems have been developed.
- Benthic monitoring underdeveloped: busy images with poor contrast.
- Nascent stage.

- Bottleneck is having annotated data set for training (100k + images).
- Stress testing.

Tasks

- Scope any models which are working in benthic – deep sear (AUV).
- Conduit for creating an annotated training dataset *priority.
- Periodic state of the art review.
- Develop introduction to machine learning for biologists.
- Define the tasks we as a community want machine learning to achieve.
- Scope out international expertise and work out how best to learn from them.
- Scope possible “x prize” competition for benthic image machine learning (after training data set built).

Reference collection collation

- Head start using Google? Good idea! Starting point algorithm.

Ensemble approach

- Multiple models.
- Transfer learning.
- Polygons better than points for training.
- Consider positives of human-mediated AI.

Timeline

3. Strategy for data call (benthic imagery).

- Training dataset.
- Image types.
- Repositories.
- Participants.
- Etc.

4. Workshop to conceptualist the machine learning workflow for benthic biology (including both biology and machine learning specialists).

- High level: input criteria, output criteria.
- This feeds into image annotation timeline on other sheet.

Connections

- Enumeration.
- Classification of community.
- Standardisation of analysis methods.
- Sampling units / knowing / recording relevant metadata.

ISSUE B: Image annotation → cost efficiency

Discussion

- Universality of exports – CSV with XY tags.
- Cloud – benefits or hindrances.
    - We don't need just one solution – cloud or local to suit.
- Interoperability. Pull and push species lists / stills.
- Needs to be usable in commercial setting.

Annotation tasks

- Identify a list of features which any platform must have.
- Preliminary task to ID applications across community (i.e. not just benthic imagery).
- Community building → create an annotation user group.
- Define levels of image annotation (3 levels) to help machine learning.

Cost efficiency

- Initial cost outlay likely to be offset by efficiencies in future.

Timeline

1. Call for interest – mail list.
2. Pool experts – to build supportive, collaborative community across stakeholders.
3. Define applications for annotated imagery.
4. Define features platforms must have.
5. Build specification for experts (e.g. polygon vs point) data format.
6. Combined training course, to cover: awareness of why to use; key skills in using; key development issues and solutions.

Connections

- Standardisation of analysis methods.
- Taxonomic ID (creation of reference collections).
- 8
- 9
- 10 Enumeration
- 12 Training.

Overall point

- Introduction and training required – combined for annotation and machine learning. Link with international learning. 3 day option? 1. Beginner; 2. Intermediate; 3. Expert.

Theme 4. Development of standardised / specialised sampling approaches

Questions

- What is the aim of your survey?
- What is the aim of the survey?
- Different approaches for systems / depths?
- Cross stakeholder regulatory decisions requirements?
- What is the max length a video clip should be?
- May have a specific Q. Need flexibility.
- Can we standardise broadscale mapping tools (e.g. WASSe) to determine EUNIS seabed types?
- Is standardisation of data output more relevant than sampling method?
- Sample area monitoring vs characterisation.
- How much does it matter?
- Field of view.
- Back compatibility 4K vs PAL!
- Standardise by habitat type?
- Recommendations on best equipment and gear to use: pros and cons etc.
- How would standardisation be enforced / followed in the commercial sector?
- How much is enough data for your question?
- External QA for analysis.
- How do we store data?
- What is your minimum quality?
- New technology within standards?
- What resolution is best?
- Why restrict survey capability?
- Ground truthing images.
- How to ensure a scale is provided with associated footage?
- Should we sub-sample the images (video and stills) to ensure comparative across institutes / consultants.
- How to standardise field of view across different media?
- What is “mesh”?
- Standardised Equipment: government procurement rules vs. standards.
- Photomosaicing (close ups covering whole quadrat in low visibility conditions)

Key issues

- Every job is different! [But this flexibility should balance with a minimum standard. Standard → metadata.

- Financial implications.
- How good does it really have to be for purpose?
- Confidence in conclusions (power to detect).
- Standard approaches to sampling for each high level aim – i.e. condition assessment, change, presence / absence etc.
- Separate 'collection standard' from 'analysis standards' → collect once use multiples.
- What are the existing metrics and which are critical to maximise use of data (would small changes to info gathered make a big difference).
- Quick post-survey quality classification.
- Commercial 'reps' on near shore surveys – data quality.
- Operational guidelines specific to habitat and / or task / data required. [Addition / updates for new / novel methods e.g. acoustic cameras, BRUVS etc.]
- What constitutes a sample? – task specific
- Mesh 'guidelines' were for mapping not monitoring.
- Guidelines cannot be too operationally specific – focus on outcomes.
- Guidance on sampling for monitoring change in biological communities should be possible.
- Level of compression (codes used and file format) in video acquisition – are we saving our data in the optimal format? Guidance?
- Research into optimal sampling strategy for image collection (scale / number of images).

Challenges

- 'Standardised approach' not standard methods, but standardised assessment of 'quality' & whether is fit for purpose.
- Very similar to #1 – q.v.
- Should consider perspective of 'end user' / bill payer & legislators / regulators.
- Can we set guidelines / recommendations rather than standards that must be achieved? Same standard not appropriate in all cases. [Link into decision tree for enumerator?]
- Cost vs benefit.
- Time and money and weather restraints.

Addressing key issues

See theme 1.

Theme 5. Image analysis contractual agreements and resources

Questions

- Allow experts to be promoted without moving them to admin.
- Is there a way to improve data sharing with public bodies?
- More flexibility for process review to meet aims?
- Lack of staff in agencies?
- Frameworks as per microbenthic etc?
- Where do we share data?
- How do we share more information / data?
- What is in place between agencies?
- Set standards for time-per-image (depending on complexity?) or per minute of video.
- Are agencies realistic when it comes to timings and prices?
- How do agencies determine analysis costs?
- Two stages – first to scope time needed then contract full job?
- Quantify quality and effort minimum?
- Do we consider client and regulator?
- Are frameworks appropriate?
- How do we ensure data are fit for the individual purpose?
- How do we deal with analysis when the purpose / indicators are unclear?
- Can we develop health & safety standards to ensure analysis is carried out as effectively as possible?
- Standard level of QA?
- Model agreements for generic approaches?
- How many images need to be analysed?

Key issues

- Being constantly asked for more and more detailed analysis for lower and lower cost. [That's life!]
- Streamlining the contractual process for imagery i.e. new framework.
- *Apparent devaluing of services.
- *Acquisition of data should also be available to commercial companies performing the analysis.
- *Problems with framework agreements, e.g. length, no. participants, etc.
- Clarity of objectives.
- Realistic timelines.*
- Specify time per unit video / image to be analysed.
- Procurement system more flexible → not just cost → unfair undercutting → poor quality.
- Bullet proof contract of what QC is expected.

- Keep data / image specification collection with same organisation who does the analysis.
- *Pre-check image quality before contracting. [As part of 2-stage contract.] [Supply thumbnail images.]
- Interim data check / QC with contractor could avoid issues late.
- Minimum quality (video) specified in contract?
- Keep contractor informed of the aim / use for the data – it will help them assess whether they are producing a fit for purpose product.
- In some departments people writing / issuing tenders do not understand the issues so can get no suitable bidders.
- Expectation management.
- Problem of low cost bids / undercutting true costs.
- Feedback from analysis – results.
- Invest some resources in appropriate scoping for contracts → efficient sampling / analysis (long term cost effectiveness).
- Ownership and access of footage / knowledge / data.

Challenges

- To acquire good data it helps if there is flexibility in timing of surveys – wait for good conditions This is very hard for larger organisations, but small organisations – with this flexibility – find it harder to win the contracts (expansion of certification / accreditation etc.)
- Implement 2-step contracts for video / stills analysis. 1. Quick QA check on image quality. 2. Issue full tender.
- Disconnect between acquisition and analysis needs to be reduced.
- Needs to be regulator driven.
- Needs input from contractors (scoping for most efficiency planning).

Addressing key issues

ISSUE: Clarity and purpose of contracts

Discussion

- Increase input of information at start to prevent over-working and provide a more in depth scope.
- *More information purpose and analysis purpose (objective of survey) – clear, as much detail as possible to allow precise tendering – efficient.*
- Wash up meeting – feedback from analysis (quality or other issues) face to face but feedback to survey planning and implementation.
- Communication throughout project → two way and adaptive.
- Remove poor quality images / video from data set – or allow contractors more scope to not analyse.
- Appropriate metadata – images named and associate data checked – don't expect contractor or allow budget.

- Initial time period of contract to review data – what to remove, detail to analyse → contractual changes.

Connections

- Consistent methodologies → efficiency in tender process.
- QA consistency – stipulate precise requirements.

Theme 6. Biotope monitoring

Questions

- Consider season [vs. cruise schedule]
- Ref collections for different EUNIS levels?
- Include confidence assessments for EUNIS categories.
- How can biotope monitoring feed into management?
- Can we develop a biotope image catalogue? [Shared video portal (QA'd) for good example video / image.
- What is considered the baseline?
- Is it possible to have multi-level standards based on reason for survey and available equipment.
- What about classification of community quality e.g. high, good, moderate, poor, etc.
- Can we grade the quality of biotope?
- Different interpretation of current → JNCC → handbook.
- What levels are really needed for different tasks?
- More quantifiable detail in biotope description i.e. offshore?
- What about developing an Epifaunal Quality Index?
- Are still images used for biotoping?
- Isn't the purpose of biotopes best for broad mapping?
- How draw biotope boundaries in gradual changes?
- Is it appropriate to assign biotope to images?
- Standard segment size for biotope video analysis.
- Can you monitor biotope change?
- Can we test for consistency in analysts determination of biotopes?
- Are biotopes too subjective to use in monitoring?
- Should we be monitoring change in biotopes? Are these the best indicators of change?
- What limitations of video for characterisation of soft sediment biotopes?
- Is substrate determination from video clear enough for biotope allocation (silt content).
- Is focus on biotopes the right thing? Indicator?
- Targeted monitoring: SDM models?
- How to define if there is a 'significant' change over time?
- How can we assess temporal change of communities?
- How can monitoring biotopes detect meaningful ecological change?

Key issues

- How to standardise something ('biotope') that can be difficult to achieve because of the image quality. [– indication of confidence in allocation of a 'biotope' name.]
- What change are you trying to detect?

- What scale are you measuring biotopes at?
- Currently it can be considered that biotopes [classifications] are too subjective to be an effective monitoring tool. Consider alternatives – either use community data or specific indicators more powerful?
- Is there any value to biotope monitoring? (It was designed for mapping.)
- Biotope classification doesn't deal well with (natural) fluctuations.
- Should biotopes be used?
- Biotope system based in part on infaunal communities – hard to apply epibenthic data/
- Detecting community change can be difficult due to inherent problems such as observer inconsistencies by using biotopes it adds an extra level of subjectivity.
- Weed out some of the 'artefact' problems in the classification.
- Classification updates.
- Gaps in offshore and NI inshore data used to define biotopes.
- Inappropriate biotope scale for some habitats.
- Higher resolution of CATAMI / morphological systems may make correlation with MNCR / EUNIS easier / improve biotope monitoring / classification.
- Great classification tool – uncertain of applicability in monitoring.

Challenges

- Should we disconnect the morphological approach from the issue of biotopes monitoring?
- Is there a need for sufficient 'background' data to identify the natural variation (spatial & temporal) against which to assess impacts? ~~£~~£~~£~~£~~£~~!!
- 2 purposes: 1. Improved [illegible] 2. Help in monitoring pressures using morphology.
- Biotopes not to be used for monitoring. [But how can we then compare with previous studies?]
- Outline clearly purpose and limitations of biotopes.
- If link CATAMI to biotopes, method needs to be readily available so users can convert historic for comparison.

Addressing key issues

ISSUE: Identifying certain species from imagery is problematic where samples can't be collected

e.g. sponges, bryozoans, hydroids on circalittoral rock → can't classify the community to which an image 'belongs' with sufficient resolution.

Discussion

- Aim: use a morphological classification system to improve:
    - Ability to describe / resolve biotopes / biological communities (at L5 of classification) – for mapping biological communities.

    - To be able to use imagery more effectively* to describe and map biological communities – specifically for sublittoral “rock” / hard stuff. *Get better resolution than we can currently get with biotope classification.

- Step 1: Develop a mixed model classification system to describe biological communities of sublittoral hard substrate from imagery.
- Test 1: Test if the system can be applied to modify L5 of classification.
- Test 2: Use it to test if morphology to detect change across a pressure gradient.
- Other tests to be suggested…

Theme 7. Sampling units

Questions

- Qualitative vs. quantitative.
- Quantitative or qualitative?
- Sampling units stratified to feature?
- Mesh guidelines – update?
- Quality of sampling unit? Rough weather & heave… "SNOW" speed…
- What is a transect v. drop-down camera in a moving vessel?
- Lasers → standardise their use.
- What is the scope? And have you recorded it?
- What question are we asking?
- How do we sample patchy distributions?
- Is SACFOR used 'properly' (consistently)
- At what scale?
- Should this be based on area of seabed or number of individuals?
- How do you factor temporal variability into sampling units based on spatial variability?
- Matching sediment video and grabs → units?
- Should we set a standard for minimum transect length / minimum number of images for different substrates?
- How to select samples? (Randomisation)
- Can species curves be used to inform transect length?
- How do you set standards for transect lengths, frames analyses etc across different habitats / sites etc?
- How many still images do you need to analyse?
- What should be standard for video vs. stills?
- Do we alter sampling units on basis of visibility?
- Frequency of still images v. image bias.

Key issues

- Kit dependent. [Must be in the scope / plan upfront.] [and fit for purpose and consistent planning guidance (survey design).] [and spatial planning exercises like power analysis.]
- What is the question that needs answering?
- Standards are needed to guide project planning / sampling design to ensure correct sampling units are used / data are robust. But these need to be flexible to accommodate different projects.
- The sampling unit needs to reflect the spatial scale of your question. Important factors: patchiness, organism size spectrum, habitat variability, spatial scale of study.
- Biotopes are sampling units too! [No experimental stratifications.]

- Purpose specific survey requirements vs. trying to capture more info / which may never get analysed.
- Catalogue of survey & video etc which could be used for student projects etc. Generally no funds to analyse everything.
- Replication? Within station variability.
- Purpose – monitoring vs characterisation [vs verification (presence / absence).
- Sampling units should be able to produce quantitative data even if qualitative used e.g. making sure footage has a scale (lasers, altimeter, etc.). [YES – but is this realistic for small inshore operators?]
- Use of SDMs to guide survey design in new areas.
- Meet statistical assumptions e.g. randomness. [Spatial autocorrelation.]
- Stills = quantitative.
- Video = qualitative? [Quantitative too (objective specific).]
- Issue of cost benefit of collecting more data (sampling units) than you need to answer your project → collect more data that could be analysed in future & subsample this for your project?

Challenges

- Missing link to standardisation / metadata? How is sampling unit recorded in metadata?
- How will awareness of this 'decision tree' be spread around the industry? Needs to be common knowledge.
- → PDG → final work package → shared resource → eventually find platform
- Good points made, but where would standardising help, and how? Justify… [→ Provide an introduction to the issue with examples. Not a restrictive standard. Providing a process with an audit trail to document planning.]
- Cost will win. [→ Defining sampling unit will help to improve survey plan, which may make survey more efficient, so cost may increase, decrease or stay same. If we want standardisation, and to monitor appropriately, it may cost more for everyone.]
- Sometimes sampling unit can't be defined until in the field. [→ OK! Repeat process and document. (Added to step 3.)]
- Defined sampling unit but do we need to define sample populations. [→ Yes. Developing a template for this is an action item.]

Addressing key issues

ISSUE A: What is the appropriate sampling unit for the specific question of the study – size of sampling unit needs to be representative of the population

Discussion

- Step 1: Define your sampling population (template / pro-forma).
- Step 2: Define the size (m^2/numbers counted) and number of your sampling unit (template / pro-forma).
- Step 3: Record any limitations / reductions in your scope post-survey – if necessary go back to step 1 re sample unit changes in the field.

- Step 4: Keep documentation of steps 1-3 with data and results.

Action items for group

- Create a table for recording what the aim of the study is, what the sampling population is (spatial scale / size), prompts of what you should record.
- Flow chart for how to define the size of sampling unit based on existing knowledge – with prompts for what you need to consider and to record your rationale (potentially need a pilot study / simulation study).
- Replicates are sampling units – it depends on survey design – implicit in definition of sampling unit size and number.
- Collate information based on table [see first action item] to give examples of types of survey scope & decisions made.
- Need to make a clear terminology explanation on the relationship between sampling strategy / design and sampling unit – diagram to accompany table & flow chart.
- Create pictures of example sampling units to go with table & flow diagram.

Connections

- Standardisation of acquisition methodology
- Available technology

Sample unit, definition

A sample unit is representative of a sample population. There is an equal chance of variation in every randomly located sample unit within a sample population. For example, a sample unit could be:

- A single image
- A group of images
- A video clip
- A subset of video clips

ISSUE B: Purpose specific survey requirements vs. trying to capture more information to "future proof"

Discussion

- You should not compromise the scope of your image collection – refer to your sample population.
- Additional information needs to be specified if it means changing sampling unit. → Ask other people if aims can be extended to meet their needs.
- Augment your sampling by collecting more data (sampling units) than you need to ensure you have sufficient data to answer your question, when your data have been 'cleaned', e.g. buffer to account for poor image quality.

Theme 8. Image reference collections

Questions

- How to address biogeography in image collections?
- Reference collection for different conditions.
- Keyword multi-project connections “pool similar images”
- Who pays to do the collation and maintenance.
- Using crowd-sourced imagery with suitable QA/QC.
- Generate using annotation software? BIIGLE?
- Publicising such facility exists & encourage uptake.
- Need validated images c.f. “type” specimens.
- Confidence in image ID.
- Do we have the infrastructure for a UA (universally accessible?) reference collection.
- Ability to contribute to taxonomic ID online?
- What quality of images required? Broader taxa = poorer quality?
- What should the format of the images be?
- Could we have multilevel reference collection that identifies images based on taxonomy and life form.
- Ownership for centralised collection online (everyone contribute)?
- Minimum standard for photos, i.e. what photos bin.
- Who owns the images?
- RESTful / SOAP API interface for cross-platform access?
- How to make widely available – link to existing web-based facilities?
- Can there be a searchable image database with attributes?
- Can we identify general spatial location to help with potential taxa from area of interest.
- Would there need to be a DAC for this?
- Linked to current naming conventions.
- Who QCs the ref collection?
- Can we add on to a pre-existing database / info system? E.g. WoRMS.
- Is it free access?
- How to get past copyright?
- Once published can images be open access? (Public)
- A reference collection that could be accessed offline? Be useful for offshore survey.
- Who would moderate – be the expert on taxonomic decisions?
- Collate reference catalogues.
- Who manages / updates?
- Who manages a system to ensure consistency across reference images?

- How do we coordinate ring tests?
- Who would QC photos before they go “live”?

Key issues

- Guaranteed funding.
- Public accessible.
- Maintained & updated.
- Range of image qualities included.
- With metadata.
- Species / lifeforms / habitats.
- Link to other systems (e.g. CATAMI, WoRMS).
- Wide range of validated & checked sources.
- Validation of data (DNA analysis? Taxonomic expert?)
- Keyword tagging.
- Two different reference collections: 1. Species identification; 2. Abundance training.
- Online but also downloadable for offshore use.
- Cross-platform link – e.g. WoRMS.
- Biotope reference collection, confirmed by others.
- Reference collection portal / database online that people can upload to and peer review.
- Inter-lab reference collection. Share resources.
- Reference library to include local/regional biotope and species variants.
- Single image library → app developed to take out on survey.
- Image library to be nested, so include best, adequate and shite from each major approach (go pro → ROV etc) (inshore → deep sea) (turbid → clear).
- Priority solved (in part) by annotation software.
- Linked with EIP and ring test.
- Owned and managed centrally, but community / expert updated – i.e. WoRMS.
- User feedback a priority.
- Who would host and moderate?
- Include regional variation.
- Who will fund it / own it / manage it?
- Lon-term management.

Challenges

- Anonymising data can get more participation from industry. [See “tick box” options (1st step sheet).]
- Mandate / demand publicly funded data is accessible. [It is!]
- [We feel these have been tackled in our plans.]

    - Marlin provides images of biotopes… don't create something new…upgrade / maintain.
    - Medin / DASSH image archive. Add to / increase / develop existing.
    - Easier to build on existing resources (Habitas, Marlin etc. etc.) that start something from scratch (?)
    - Avoid duplication by using existing infrastructure / data capability.
    - Similar to current deep sea efforts – contact Kerry Howell.
    - Can this be part of something more ambitious? I.e. a UK Imagery Data archive, e.g. Australia.
- A deep sea image ref collection & database is being developed – structure & approach to be published soon and this could form a useful base.
- Consider substrate, region, habitat etc. [→ step 2]
- Do we have right tools / software for consistent QA/QC across groups. Should everyone use the same tools / software? [QA to same standard "industry/public body"? Links with theme 12.]

Addressing key issues

STEP 1, ISSUE A: Ownership

Can this be done?

Discussion

- Publicly funded – copyright, caveats, credited (open data).
- Opt in / out to use data from commercial contracts.
- Subscription or freely available.
- [Incentives for industry to submit data, i.e. good publicity.]

Short term steps to take

1. Contact relevant institutes / companies – do they have an existing reference set?

2. Would they be prepared to release it? Tick box options for various uses.

3. Industry reward for participation (credited, free advertising, reduced subscription fees).

[Alternative sources of data, e.g. Seasearch, citizen science]

Connections

- 5 Image analysis contracts (commercially sensitive)
- 3 Annotation

STEP 2, ISSUE B: What classification reference collections already exist?

[E.g. Annex 1 features, PMFs etc. Video?] Avoid duplication.

Discussion

- Centralised list of resources e.g. links to existing catalogues.

- Identify existing resources.
- QC.
- Scope / literature review of any previous processes in generating reference sets? E.g. WGNEPS.

Medium term

- Developing a catalogue.
- [Multiple images + range of conditions (image quality, resolution, etc.)]

Connections

- Observer consistency
- Annotation

STEP 2, ISSUE C: What to include.

Discussion

- Identification of organisms (taxonomy vs. morphological / functional) → CATAMI?
- Classification of habitats / biotope.
- VME – cobble reef / *Sabellaria*, e.g. borderline.
- Abundance reference e.g. *Nephrops* burrows.
- For each of above:
    - Metadata
    - Multiple images – different camera platform; different environmental conditions; image quality?; different aspect of subject matter, e.g. oblique / vertical cameras – and angel subject is in relation to camera (square on, side on, etc.)
    - Hierarchical classification – taxonomy and morphology.

Steps

1. Identify what's required.

2. Identify priorities and trial dataset. Identify user groups. Test a beta version.

[Hierarchical classification – taxonomy and morphology.]

[Location, depth, temperature – refers to distribution ranges of species.]

Connections

- Training
- Annotation
- Observer consistency
- 12 Image analysis training
- Standards

STEP 3, ISSUE D: Validating reference sets.

Discussion

- Identify experts in that field.
- Include description and justification for name and include metadata.
- Maintain training for all users.
- QC / statistical confidence in identification.
- Have multiple examples.
- [Quality / confidence score for id, e.g. photo id only, or photo and sample identification, in generating reference set.]

Steps

- Identify experts to verify images.
- Put QA / QC process in place.
- Identify confidence scores.
- Provide guidance on what level taxa can be identified to depending on purpose & quality.

Connections

- 9 observer consistency
- 2 taxa identification
- 12 image analysis training
- 1

STEP 3, ISSUE E: Setting up infrastructure.

Discussion

- Who will manage / host / pay for the processes.
- Work with others with similar interests, experience and / or knowledge.
- Identify knowledge caps and address.
- Make available on and offline, and 'live' – link to WoRMS.
- To include video + still options.
- Method to query available data – taxa, location, depth.

Steps

- Look & learn from existing systems.
- Find funding.
- Identify a host.
- [Integrable with other platforms e.g. WoRMS.]
- [API interface = leverage tech to simplify exchange (e.g. WoRMS).]

Theme 9. Observer consistency

Questions

- Use of common reference libraries?
- Is consistency more important than taxonomic resolution?
- Can traits help us improve consistency rather than species ID?
- Consistency over time – how?
- Verification with metabarcoding, e-DNA techniques.
- Cost of ring tests? QA.
- Level of experience. How to share expertise? Junior and senior.
- Need control organisation or “club” to join – is this NMBAQC?
- A forum for analysts to seek advice, share ideas etc. How to run/who would organise?
- Time & staff cost of several observations.
- How to improve expertise?
- Analysis by contracts is competitive rather than collaborative.
- Training workshops & costs?
- Consistent training.
- Use of machine learning?
- Techniques to overcome? E.g. sponge morphology single species mon.
- Level of training between observers…
- How to annotate footage (video and stills) for training? Analogue and HD.
- How to compare imagery acquired using different technology? (e.g. HD, not HD.)
- Would use of a shared image reference catalogue help with consistency?
- How to train new analysts.
- Consistency in annotation of what? Species? Substrates?
- Use of online annotation software (Biigle?) may allow collaboration / QA between organisations.
- How to overcome analysis fatigue.
- Is analyst fatigue taken seriously?
- Double analyse? How much.
- Randomise analysis order.
- Use AI to prioritise effort.
- Data quality?
- Annotation software.

Answers?

- Workshops:
    - Multiple, split by phyla

    - Willing to pay (small amounts)
- Online
    - Ref collections
    - Image guide (i.e. with EIP)
    - Commercial colleagues – massive image database
- Annotation software platforms
    - Collaborative working
    - Real time QC
    - Easy / repeatable QC
- AI event marking
- Address analyst fatigue

Key issues

- Online resources and forums.
- Won't get rid of all inconsistencies when using human analysis. [What should the excepted error between annotators be?]
- How does this really impact what we can do with the data?
- At which point does extra training stop delivering substantial improvements in consistency?
- Accepting less taxonomic resolution (→CATAMI?) to improve consistency (encourage) and confidence in data.
- Ensuring observers have access and use (standardised) training and guidance (contractual and workshops).
- Set a threshold when training observers. Lin's CCC. [& relationship between how much is asked and amount error. Ask simple thing = less error.
- Use of a shared reference collection / morphological identification system would help with consistency.
- Staff time and resulting costs.
- Training courses should be readily accessible and very reasonably priced.
- Inter-lab PT scheme?
- Own sample?
- Fatigue.
- Could we develop online training exercises?
- Study reference collection from similar areas / habitats before beginning analysis to have a more consistent "search image" from the start (usually this develops as view more images).
- →BIAS
- Document how you tested observer consistency and what the results were.
- Must have robust QC and methodology first. →Training.
- Accurate survey planning, reporting and image analysis methodology and metadata.

- NMBAQC (or similar) survey method audit for first time video survey (commercial and government). [Covered elsewhere.]
- Pilot ring test for deep sea images reliant on training, image library, id skills etc etc! [Better quality images?]
- Supportive technology e.g. annotation software.

Challenges

- Define "consistency" and "excepted error".
- Contract issues – time consuming and costly within contracts at low rates and short timescales). [Needs to be regulator driven "best practise" → sell to client as "best product" → refer to guidance doc →]
- How to grade inconsistencies and set thresholds – use statistical analysis to provide unbiased reflection of the work (e.g. Lin's concordance correlation coefficient, Lin 1989) [Will need to further consider evaluation process.] [Assess advantages gained → refer to available medical literature and connect to other working groups in other fields.]
- How will OS (own sample) module be moderated – any real chance for meaningful discussion? [Consider regional "experts" to check OS (e.g. UK v. global)] [Re global: Is this applicable?]
    - Definitely, this approach has been very successful with other elements where audit feeds into workshops / ID books.
- Is funding required for an intro course on video ID? If a locally generated ref set and accompanying S.O.P. is generated this can be circulated vastly reducing costs. [Don't disagree BUT can we not have both for extra trainee buy in? (IF FUNDING AVAILABLE.)
- Consideration needs to be paid to how 'ring tests' etc can be organised – i.e. through BIIGLE – reduce time needed. [Agree.]
- How to avoid bias? [What bias? Suggested include randomisation step to avoid intraobserver bias and interobserver bias. And a QC step. → What else is suggested here?
- What common tools / software should everyone use? [Not a priority can recommend but as long as follow the process improvements will be made.] [→Recommend high resolution screens!]
- Should everyone be using the same tool / software.
- Who will provide management and funding for database and introductory course? [Excellent question!] [Consider tie-in to Habitas / Marlin / Algaebase / Fishbase etc.]

Addressing key issues

ISSUE A: How to improve observer consistency

Discussion

- In house suggested process:
- Stage 1: Review all photography – remove poor stills? As a group of observers make a standard checklist (taxon + qualifier). Validate / calibrate observer identification.
    - At this stage, can consider reference collections, identification sources (accessible?).
- Stage 2: Enumeration "training" (e.g. burrows video).

- Stage 3: Internal QC (e.g. 10%).
- Stage 4: Evaluate analyst variability.

Other concurrent suggestions

- Introductory course to video identification (online? Skype? To reduce cost)
- Formal process & protocol. Consider observer fatigue, time restraints of analysis / annotation?
    - Do analysis as above → QC
    - Evaluate variability (see*)
    - Accept inconsistent
    - → Also consider specifics of analysis (e.g. # observers)
- Familiarise self with abundance assessment → agree with client!
- Ring test / own sample – feedback loop with regulatory body. Links to contracts – regulatory driven as client may not want "their data" in public domain.
- Equivalent of Seasearch competency test (region specific?) → May need annual reassessment but minimum competency may be appropriate
- Reference collections
- Training
- Image annotation software
- Overall enumeration approaches
- Morphological approaches
- [Cefas have similar process in place.]
- [Could be contract issues? Cost unknown!]
- [*Also, evaluate analyst variability and understand impact on analysis.
- [In addition to fatigue → intraobserver variability. Therefor randomise analysis?]
- [Examples of variability in specific "habitats" "species" etc.]
- [Actions: formalise process, establish procedure, analysis, QC, variability and importance of this per habitat / species. Database to incorporate stage approach to taxonomy. Guidance on process and protocol. Introductory course to video ID. Provide examples of variability for specifics – MBA suggest have data? Ring test / own samples. FUNDING.]

Connections

- Reference collections.
- Training.
- Image annotation software.
- Overall enumeration approaches.
- Morphological approaches.
- Links to contracts – regulatory driven as client may not want "their data" in public domain.

ISSUE B: There will be inconsistencies. How do we deal with them?

Discussion

- Staged approach to taxonomy. Agree stage of classification with client (cost / quality / analyst dependent). Note that as stills are retained, dependent on image quality, may be able to ID further in future.
    - E.g. physical → morphological → CATAMI → Genus → Species.
    - IA would help here if revisiting ID. Also for QC.
- Excepted error dependent on question, contract, habitat, taxa etc. → importance of variability in observer will be dependent on variability of habitat / taxa + relation of change that want to detect.
    - Excepted error at individual photo level v. final categorisation.
- Ability to aggregate – defined structure (EIP could link) → develop database that could do this.
- [Record specifics of analysis / annotation design → variability.]

Connections

- Image annotation software?
- Morphological approaches.

ISSUE C: Actions to take forward (→ combine 9A + 9B)

Actions

- Questionnaire: what process / procedures does everyone use?
- Develop guidance
    - Process / procedure (see 9A) → Agency led.
    - Specific QA (training) / QC (checks)
    - Variability
- Provide examples of variability for specific habitats and species → assess acceptable error. MBA have data.
- Develop a database to incorporate staged approach to taxonomy where more info can be extracted → EIP / CATAMI / users → funding required.
- Introductory course on video ID → online / Skype → evaluate Seasearch to see if can assign “competence”. → Funding required.
- Ring test / own samples → consider “assessment” level e.g. ID / enumeration etc – what’s ok? Do macrofauna / macroalgae existing methods work for video / still? Can be done in advance of “funding” (MSC?) Needs investigating. → Funding required.

Theme 10. Overall enumeration approaches

Questions

- Is presence / absence suitable for some taxa?
- Are there some case studies of where SACFOR has been successfully?
- Should we use more quantitative approaches to assess change?
- Why do we use SACFOR? (It was developed for intertidal areas.)
- Numerous assessment methods available. Still vs. video.
- How do we assess reef height and rugosity?
- Which technique is most consistent? Which is most useful for purpose?
- Most appropriate approach?
- Is a consistent approach more valuable than an accurate approach?
- Objective specific: ground-truthing, biotopes, analysis.
- Should we abandon SACFOR?? What is a better alternative?
- Can a new, more appropriate key / scale be developed?
- Are there criteria for a reliable absence record?
- How do you assess absence? I.e. decide what to look for?
- SACFOR – issues. How do you analyse.
- What are we enumerating for? Aim specific data, keep the raw files.
- How can enumeration adapt without ever improving quality of imagery?
- We should set a minimum size for enumeration but how do we do this?
- How much sampling needed to include rare species?
- How to select most appropriate key / scale.
- Do we need a list of species per biotope (to reduce list for recording absence).
- Is SACFOR used, do workshops / training exist?
- How do we account for what species can be reliably enumerated using different equipment e.g. ROV vs AUV?
- Re-purpose existing image data e.g. pipeline survey / ROV inspection.

Key issues

- Standardised area (mainly relates to video).
- Use of SACFOR – not always feasible or applicable to all indices.
- Ability to truncate data (e.g. cell frequency issues). [possible with software to annotate grid cells?]
- Repeatability!
- Some level of discrimination in lifeform descriptors, e.g. sponge-branching etc.
- What is use of data, e.g. high level data may be difficult to use for formal biodiversity analyses – need to think how to use indices like presence / absence for monitoring.

- Agree on most consistent enumeration technique between observers, even if results less detailed / takes more time (dependent on lifeform?)
- Quality of data.
- Not investigating more recent / more appropriate methods.
- Defining method suitability for pressure / indices monitored.
- Presence / absence can be fit for some circumstances.
- Consider enumeration in annotation in relation to data analysis → scope. [I don't understand.]
- Build in time for pre-survey enumeration standardisation / QC → all surveys. [Visual familiarisation exercise?]
- Don't do too much – avoid fatigue.
- Morphology assessments (e.g. reefiness).
- Health and safety standards / protocol to avoid observer fatigue / error.
- Requirement for actually enumerating presence absence may be enough for assessing functionality via AvTD etc. Allowing for combining epibiota data with mobile species i.e. the whole community.
- Methods to subsample imagery.
- ID purpose – don't always need to enumerate everything.
- Which method is most cost-effective for the purpose.

Challenges

- Missing-connection to training on enumeration techniques.
- Missing link to machine learning → automation.
- Evaluate research (or commission further) to develop guidance.
- Combining with CATAMI can be very short term if used with BIIGLE.
- SACFOR has a strong following. Quick / easy therefore useful for broadscale especially intertidal.
- Issues with SACFOR usually due to inappropriate use? Don't bin it…read guidance!
- Review / comparison exercise SACFOR vs abundance – are communities different?

Addressing key issues

ISSUES: Purposes of enumeration vary greatly – variety of approaches are suitable. Needs to be robust, efficient, cost-effective, valuable in the long term. The best approach is dependent on size of organisms, spatial scale etc.

Discussion

Short term

- Review benthic habitat monitoring guidance.
- Questionnaire of SACFOR purposes and enumeration techniques.
- Review outcomes of comparison exercise & review existing research.

- Guidance on minimum size of organism that should be counted, e.g. 10mm in research.
- Update / end SACFOR? Find out how people feel & think (key uses, improved guidance (EIP)) about intercalibration. ID appropriate uses of SACFOR & add to decision tree.

Medium term (0.5-1.0 years)

- Produce decision tree to support selection of enumeration approach for purpose & resources. Can we prescribe?
- ID minimum requirement for each purpose?
- Option for developers to meet guidelines to feed into wider monitoring. If in guidance it will be followed.

Long term

- Combining info on taxa with CATAMI (could be short term if using BIIGLE) & enumeration approaches. Integrate data into CATAMI structure. Dependent on development of CATAMI & links to EIP scope. Should CATAMI level & P/A be the minimum?

Overall point

- Products / terminology need to be accessible to a range of audiences. Both decision tree and EIP etc.

Connections

- What is your purpose / question? Design etc.
- Image annotation software / machine learning.
- Contract specs – cost efficiency.
- Tax resolution, image quality.
- Sampling unit / FoV minimum.
- QC standards of data – what do the questions require.
- Training on enumeration techniques.

Theme 11. Morphological classification systems

Questions

- Will CATAMI replace or augment current monitoring?
- How confident do we need to be before committing to time series?
- Role of species database vs CATAMI.
- Research question? Be clear.
- No experience of CATAMI system.
- At what level of detail? Taxonomic / functional.
- Standardisation
- Common UK morphotype catalogues?
- Could we apply deep sea morphological approaches to shallower waters?
- Species?
- How do we deal with things we can't pin to a phylum? (E.g. is it a sponge or a bryozoan.)
- Sponge issue. Not all species fit well.
- Is its use more appropriate for "habitat" level classification? (E.g. stony reef, *Sabellaria*.)
- How would we decide what to add to CATAMI to make it more suitable for the areas we work in?
- Can sensitivity be related to morphology?
- Comparability over time? Consistency of approach.
- Where do we draw the line on detail?
- How does it fit with biotope classification?
- How do we decide the taxonomic level to be counted?
- How do we tackle uncertainty / different views about morphology?
- If people add CATAMI categories (project specific) who will collate? Centralise?
- We need it for certain groups – sponges, algal turfs, bryozoan turfs, *Sabellaria*.
- Do we risk losing our high level taxonomic expertise?
- Defined operations taxonomic units for broad groups?

Key issues

- How easily can CATAMI be adapted to use in the UK? E.g. guidance documents with UK examples.
- Can we apply CATAMI to existing datasets?
- How can we embed the taxonomy into CATAMI structure? [Is in there and can add additional category for specific surveys.]
- Is CATAMI best system out there?
- Is danger we may lose species data which we need to link changes to specific (non physical) pressures?
- Seasonal morphological changes (sponges?) may affect monitoring.

- How linked to biotope classification system (previous biotope lifeform work explored?). [→ Test CATAMI on sublittoral rock section of biotope classification (that's where problems with image analysis and id to species / genus – e.g. sponges "turfs".
- Should be used in tandem with Taxon. ID.
- Ignores indicator species. [Does it?]
- Fund UK pilots.
- Calibrate across key pressure gradients.
- Link higher level biotopes (i.e. 3 & 4 complexes) to morphological classification.
- Work to build in a function element to a UK / global morpho scheme.
- Link with EIP (CATAMI still subjective).
- Define method of deriving species acc curve for morphological system.

Addressing key issues

ISSUE: The ability to consistently + reliably describe biological communities from digital imagery with enough resolution

(by using a UK morphological classification system).

Discussion

- Two main stages to investigate use of morphological classification system and roll out for future implementation.
- Review classification systems e.g. CATAMI, Morphological Taxonomic Unit (MTU) catalogue.
- Pilot test on an image set (e.g. sponge and anthozoan indicator imagery or MPA monitoring survey) and adapt classification system to fit UK sublittoral (shallow – deep sea); can compare faults across years / surveys to assess performance. Incorporate EIP to help define / resolve terms used.
- Method: Contract / ICES-style working groups.
- Can utilise platforms such as BIIGLE to enhance QA/QC and produce reference collections.

Connections

- 1 QA /QC standards.
- 2 Taxonomic ID.
- 9 Observer consistency.
- 8 Reference library.

Theme 12. Image analysis training

Questions

- NMBAQC workshops?
- How to get hands-on experience with organisms – especially deep-sea, perhaps V-R?
- Crowd-sourced analysis?
- Are there any formal image analysis training courses available or does everyone learn in-house?
- How do you know your training is working?
- Consider that QC and training can feed into each other.
- Would there be certification?
- Who would give the training?
- Who need the training?
- How to reduce / control costs of training?
- What is the scope of the "analysis" and the training?
- Would industry (commercial) experts be considered?
- Is this just about ID or is it also about process?
- How could we better learn from each other's experience?
- How to deal with geographic differences in fauna / flora – especially in ring tests.
- What internal training do companies provide currently?
- How often does re-training occur?
- How keep people refreshed / current?
- How to stop untrained staff being used?
- Can a S.O.P be developed?
- How to advertise in-house training workshops?
- Availability of data to train from?
- Online course over personal training during 'probation' period of new job.

Key issues

- Lack of communication between government and private sector (funding must be written in contract).
- [IP issues] How do we share best practice? How do we share knowledge? Local knowledge; site related; species related.
- How do we make training costs more affordable?
- Who will provide.
- Assessment for competency / accreditation?
- Ring test costs.
- Funded in-house training for commercial companies (subsidised by government agencies).

- Who will run the ring test?
- SNCB agreement on what is needed.
- Useful resources → online training? E.g. BIIGLE.
- Online portal to signpost all available training
    - Updated with new stuff
    - Link to minimum expected training standards for project delivery.
    - Biogeographic flavour to training critical.
- Targeted issue expert workshops (results of QC, developments, learn from each other).
- Why are universities / colleges not teaching these industry relevant skills?
- Larger profile, attracting people to develop the expertise needed.
- Build consistent training process into contracts – preliminary analysis of subset required and assessed.
- Lack of S.O.P. referencing training resources and procedures.

Challenges

- Graduates have less basic (hands on) taxonomy than previous.
    - Promote taxonomy training as career options.
    - Can be saved to some extent by the training options suggested.
    - Generic long-term problem recognised by Royal Society.
- Funding for commercial companies by government agencies sounds great but is it realistic?
    - To support training material (to keep information & materials consistent).
- Training element as line item in budget of contract.
    - Commercial decision – company specific.
- Cost of developing numerous courses – deep sea, inshore, tropical, temperate, etc.
    - There will always be a cost challenge.
    - Reusable (inline) materials cost-effective in the long-term.
    - Cost of bringing experts in each biome.
    - Retired community assistance? (FSC)
    - Open source (option 1) would potentially bring in additional £ - funding sources.
- All in the plan:
    - Training must be seen to be worthwhile / efficient, not just ‘box ticking’.
    - Need to identify the specific problem training would help to resolve – then set it up…
    - UK specific or worldwide?
    - Accreditation for training course i.e. not ring test but individual certificate based system? [Cheap!]
    - Learning will always be better and quicker when we are engaged and interested. I think any training plan should build that in from the start. Otherwise it’s just

telling people how to distinguish blobs on a screen. So, incorporate the biology and the value of the results into the training.

Addressing key issues

ISSUE A: Consistency of training for image analysis

Discussion

- Entry level?
- Personnel training.
    - In-house expert.
    - Outsourced company.
- Reference collection – workshops particularly on problem taxa, habitats etc.
- Maintenance.
    - Geographically specific training.
    - Emerging taxonomic techniques (molecular).
- Video / image based identification guides (image from above – representative of true survey data) – approved materials.
- Online mandatory training module e.g. MSC audit training.
    - Workshops.
    - In developing training materials use expert knowledge for specific taxa.
    - Evaluate effectiveness of training (engaging, practical, fuller biological info – relevance, feedback on benefit of training).
    - Pass / fail.
    - Practice texts – refresher training.
    - Induction element.
    - Available multi-level (universities / government / private etc).
    - Accreditation.
    - Achievable in short timeframe.
- Levels of ID training
    - Online
- Generic video / image analysis = global
- Region specific taxa = local / UK
    - Dedicated / additional modules or courses (e.g. Seasearch specific courses)
    - In-house or through 3rd party provider.

Tools

- Image reference collections.
- Contractual enforcement.

Timeline

2019

- Collate all current guidance & analysis tools – including current global seabed monitoring.
- Evaluation of above – research to prove worth / value of training.

2021

- Identify people's willingness to share & collect a reference set.
- Decide training materials – format / purpose + assign experts

2023

- Develop interface for online training.
- Identify facilitator / ownership – which assess pathway (link to issue 2).
- Decide on distribution model (outreach).
- B test + evaluate to prove value of training in improving data.
- Specialist workshops → reference collection available, machine learning available.

Dependent on timeline above + decisions on training materials + content, decide on common elements necessary for in-house training within organisations (in-house reference collection, sub-sample undertaken side by side → then assessed decisions on quality + level of analysis).

Notes on organisations

- Field Studies Council (FSC) – experienced ID – marine (Dale Fort).
- MSC – Marine Stewardship Council – fisheries certification audit.
- Part of CPD – continued professional development.
- Training should count toward CIEMS – endorse training.
- IES – Institute of Environmental Sciences (chartered institutes City & Guilds).

Connections

- Observer consistency.
    - Training
        - Taxonomic ID
        - Enumeration techniques.
- Sampling approaches
    - Quality of image data
- Image analysis contractual agreements & resources

ISSUE B: Sharing and ownership of training – video analysis

Sub-issues

- Maintenance and validation.
- Costs of running – assessing performance.
- Updating with emerging techniques.

- Input from multiple sources (materials) and sectors.
- Direction from regulators / end users – consistency between them; updating.
- Expert input on validation.

Discussion

1. Completely public central repository

- Pros: open access, inclusive, all stakeholders; public interest – access to more funding.
- Cons: costs, funding.

E.g. NMBAQC, Conservation Agency Research Institute (employ company – not in interests of competitor), Natural History Museum, MARLIN, Universities.

2. Restricted access – with good reason (no cost)

- Pros: more control over users + ring test results.
- Cons: exclude some users.

3. Buy licence for access

- Pros: it's self-funding.
- Cons: exclusive and opt out if possible.

Connections

- Reference collections.
- Quality Control and ring tests.
- Image annotation software and machine learning (BIIGLE)

ADDITIONAL THEME: Data sharing and management

Discussion

- Where is your data? Where should it go?
    - WoRMS
    - OBIS
    - Pangaea
    - Marine Recorder → MEDIN → DASSH
    - Capable of taking lots of images? (→ BoDe Pangaea)
    - Accessibility / download?
- JNCC diagram linking data infrastructure → ROV guidelines pub.
- MEDIN call for data ingestion – images, annotations, metadata. → Standard to BIIGLE for MEDIN standard expert.
- Where is the data / images currently – how managed by each agency / organisation?
- Define why we want / need an image data management system. Who will own / manage it?
- Improvements: Marine Recorder – future?
- What data to archive? Metadata, raw, processed?, annotations, summary.
- Is there ONE place to archive data / should there be?

ADDITIONAL THEME: Linking with marine industry, developers, Department for Business, Energy and Industrial Strategy and Marine Management Organisation

Discussion

- Define 'end user'.
- Define industry roles: contractor, operators, types (aquaculture, fishers, mineral extraction).
- Release of data as a condition of regulated licenses (including for exploration / investigation) and after appropriate period! I.e. 5 years.
- Clarifying with industry that data and engagement isn't that sensitive and won't 'likely' to be used against them @ regulatory level. Transparent requirements not guidelines support industry green credentials.
- 1st engagement and linkages with regulators, with stakeholders & define responsibilities / accountability.
- Engagement with regulators and industry with latest / upcoming methods.
- Simplified and robust policy briefings for open publication.
- Define levels of information required for hierarchy of purposes as anonymised data may be released sooner.
- Checklist for regulators on requirements in specifying and granting licenses.
- Example of license versus guidelines in marine mammal mitigation with wording in licence being minimum so tend to only be license that is compiled with rather than intention of guidelines and licensing.
- Not regulated = not done.
- Developers / users link at Marine Imaging Workshop.

ADDITIONAL THEME: New tech pipeline (opportunities / risks)

Discussion

- New tech how to compare results with historic data – time series monitoring.
- New tech not available to smaller consultancies.

AI

- It's going to take my job! [No!]
- Opportunity for research to interact with developers in creating algorithm.
- Cost increase vs. saving.
- R. overselling itself: know its limitations.
- X-prize / Google / Microsoft research programme.

3D photography (SFM)

- Data heavy (>Tbs) – is it needed?
- Good for monitoring physical impacts?
- Good for public engagement.
- Needs well featured seabed.
- Opportunity to develop machine learning.

Video mosaicking

- Great for training.
- Oblique angle best.
- Doesn't work well on low feature sea beds.

Storage capacity

- Big data (opportunity & risk).
- Cloud – security, transfer speeds.
- More use of UKRI stuff – Archerz etc.
- [Arrow from storage capacity] Statistics
- Who is developing approaches to deal with new tech data?

New formats

- HD → 4K → 8K
- Extract stills from video.
- Comparative software & PCs.

Laser profiling

- E.g. Cathyx Ocean, stills from video, backscatter / forescatter reduction; R2 robotics.

Underwater human submersibles

- With millionaires – piggyback on their cruises.
- Cheap cameras with volunteer divers, improve coverage inshore areas, collect evidence of impacts / damage.

- Most rec divers lack training in biosurveys.

Low cost ROV (e.g. BlueROV2)

- Reduced kit cost vs. quality / ruggedness work in challenging conditions?
- Positioning micro-USB.

V.R.

- Can it help with analysts?
- Can help with public communication.
- Can increase ROV accuracy.
- Resolution still not great.
- Motion sickness.

Acoustic imaging

- E.g. ARIS / DIDSON
- Limited coverage.
- Work in turbid water.
- Quantita assessment relief.
- Ground-truthing tool.
- AUV swarms still at R&D stage.
- Forward compatibility of data storage & programmes – how would you now read a 7" floppy disc? A document in AMI Pro?

Action

- Produce a short state of the art vs future tech comparative report / table. [and feasibility]
- Contact equipment developers and see what they have planned.
- Potential to transfer knowledge & skills between developers and users.
- Which sectors benefit most from each type of tech.
- Maintain contacts with developers and researchers.

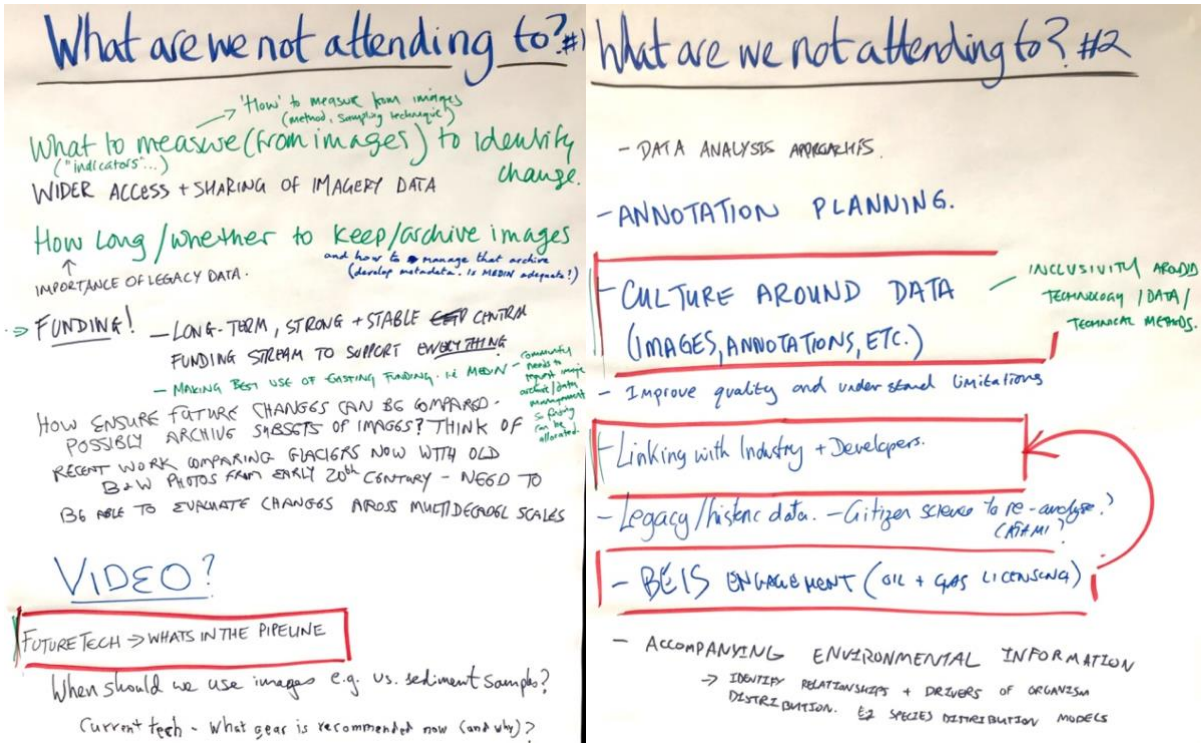
ADDITIONAL THEME: Culture around data

Discussion

- Link to training components of: image analyst training; image annotation and machine learning; data management.
- Action #1: Virtual spaces for asking / answering questions: Google groups; Slack; Listserve email list.
- Action #2: Find examples of where data sharing had a good impact / outcome.
- Action #3: Building a 'community of practice' around benthic image data – real-world / in-person workshops… hands on "how do I do…". [Dispel fears around using other groups' data.] [Regular series of events. Themes?]
- How to encourage openness around data / sharing?
    - Ensure recognition / citation of data source if used.
    - 'Publish' data → DOI (digital object identifier?); recognition / citation.
    - Different levels of open: commercial / in confidence → open academic / public funded.
- Story telling / narrative – why should I be open / contribute → 'bigger picture' contribution.
- Much research into data culture happening in U.S.

What are we not attending to?

During the working group discussion of themes, participants contributed thoughts on topics or questions that might be slipping between the cracks or being overlooked. These are shown below – some were picked up during working group discussions (those boxed in red), while others might be worth further consideration by the Plan Development Group.

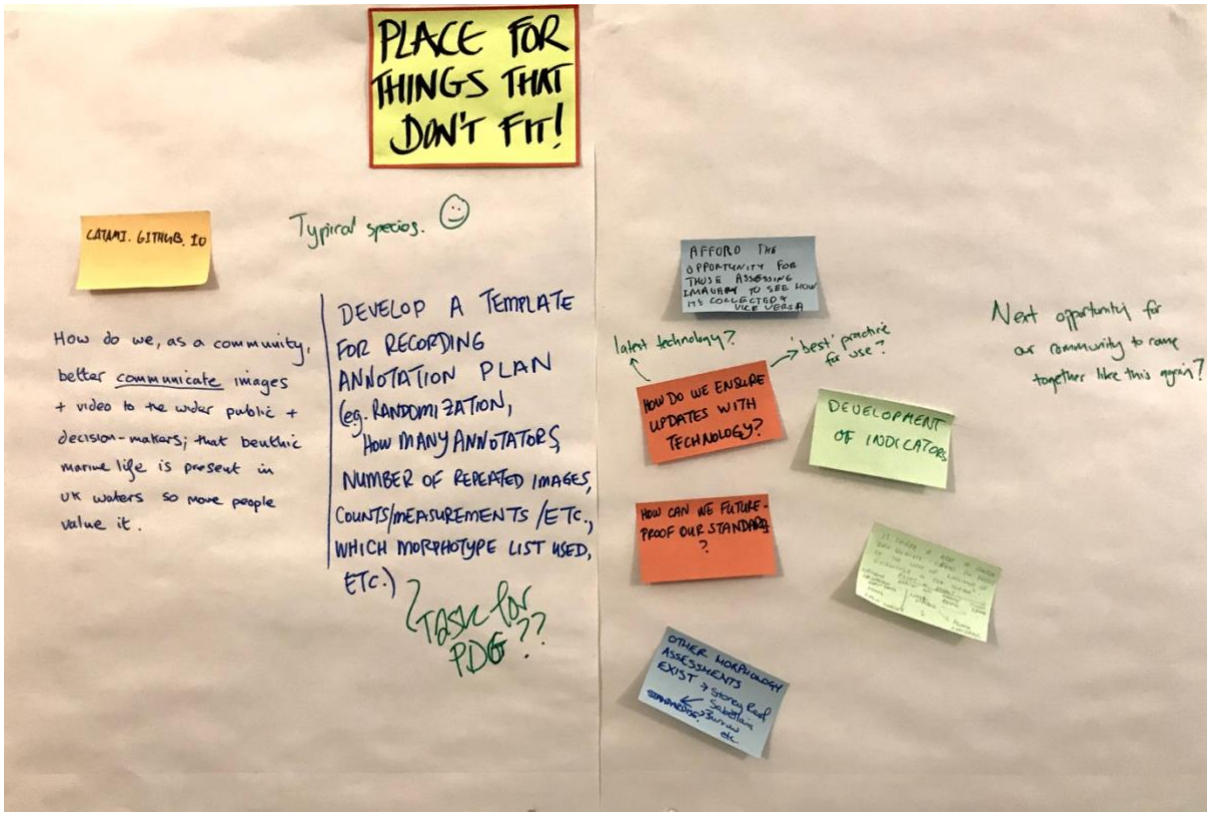


- What to measure (from images) to identify change ("indicators").
    - 'How' to measure from images (method, sampling technique).
- Wider access and sharing of imagery data.
- How long / whether to keep / archive images.
    - And how to manage that archive (develop metadata, is MEDIN adequate?)
    - Importance of legacy data.
- Funding!
    - Long-term, strong and stable central funding stream to support everything.
    - Making best use of existing funding, i.e. MEDIN. [Community needs to request image archive / data management so funding can be allocated.]
- How ensure future changes can be compared – possibly archive subsets of images? Think of recent work comparing glaciers now with old black and white photos from early 20th century – need to be able to evaluate changes across multidecadal scales.
- Video?
- Future tech → what's in the pipeline.
- When should we use images e.g. vs. sediment samples?
- Current tech – what gear is recommended now (and why)?
- Data analysis approaches.

- Annotation planning.
- Culture around data (images, annotation, etc.)
    - Inclusivity around technology / date / technical methods.
- Improve quality and understand limitations.
- Linking with industry and develops.
- BEIS engagement (oil and gas licensing).
- Legacy / historic data – citizen science to re-analyse? CATAMI?
- Accompanying environmental information → identify relationships and drivers of organism distribution. E.g. species distribution models.

Things that don't fit

Throughout the workshop, participants contributed thoughts on additional issues or things of note that did not fit into any of the discussions. These are shown below – some were picked up during working group discussion, while others might be worth further consideration by the Plan Development Group.



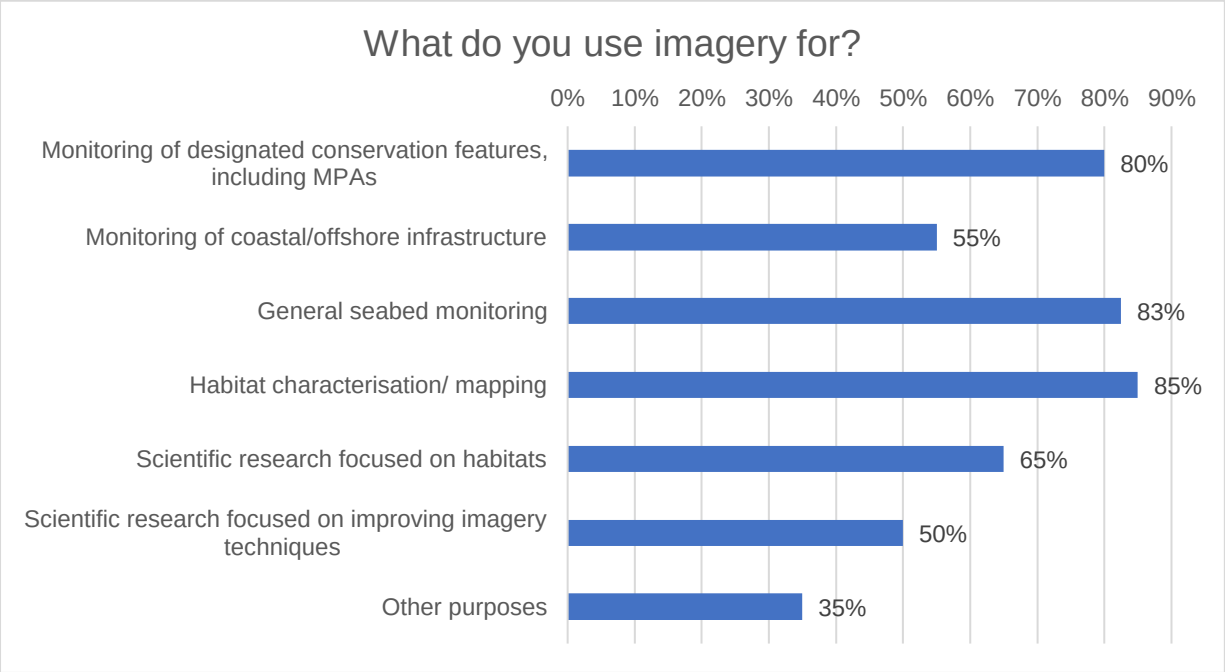
- Typical species ☺
- catami.github.io
- How do we, as a community, better communicate images and video to the wider public and decision-makers; that benthic marine life is present in UK waters so more people value it.

- Develop a template for recording annotation plan (e.g. randomisation, how many annotators, number of repeated images, counts / measurements / etc., which morphotype list used, etc.) [Task for PDG?]
- Afford the opportunity for those assessing imagery to see how it's collected & vice versa.
- How do we ensure updates with technology? [Latest technology?] ['Best' practice for use?]
- Development of indicators.
- How can we future-proof our standards?
- Next opportunity for our community to come together like this again?
- Other morphology assessments exist → stony reef, *Sabellaria* burrows, etc. → standardise?
- Is there a need to consider 'data quality' – fitness for purpose in the light of likelihood of challenge to the output?
    - Ability to detect
    - Habitat type Presence / absence Current species
    - Lowest standards → higher standards

Appendix 6: Pre-workshop questionnaire summary

Number of respondents: 40

1. What do you use imagery for?



Other purposes include…

Other scientific research / assessment

- Species reference material
- Development of indicators to assess state of marine environment.
- EIA / reef assessment.
- Habitat and species cataloguing during diving surveys to support subsequent species sampling and identification.
- Other scientific research (natural history, feeding behaviour, bioturbation, activity rates etc. etc.)

Fisheries and stock assessment

- Fisheries.
- Fisheries (*Nephrops* and scallop and some trawling imagery data are acquired).
- Stock assessment.
- *Nephrops* stock assessment, Queen scallop stock assessment.

Education

- Educational, documentary filming.
- Outreach and interpretation
- Production of imagery for outreach / interpretation.

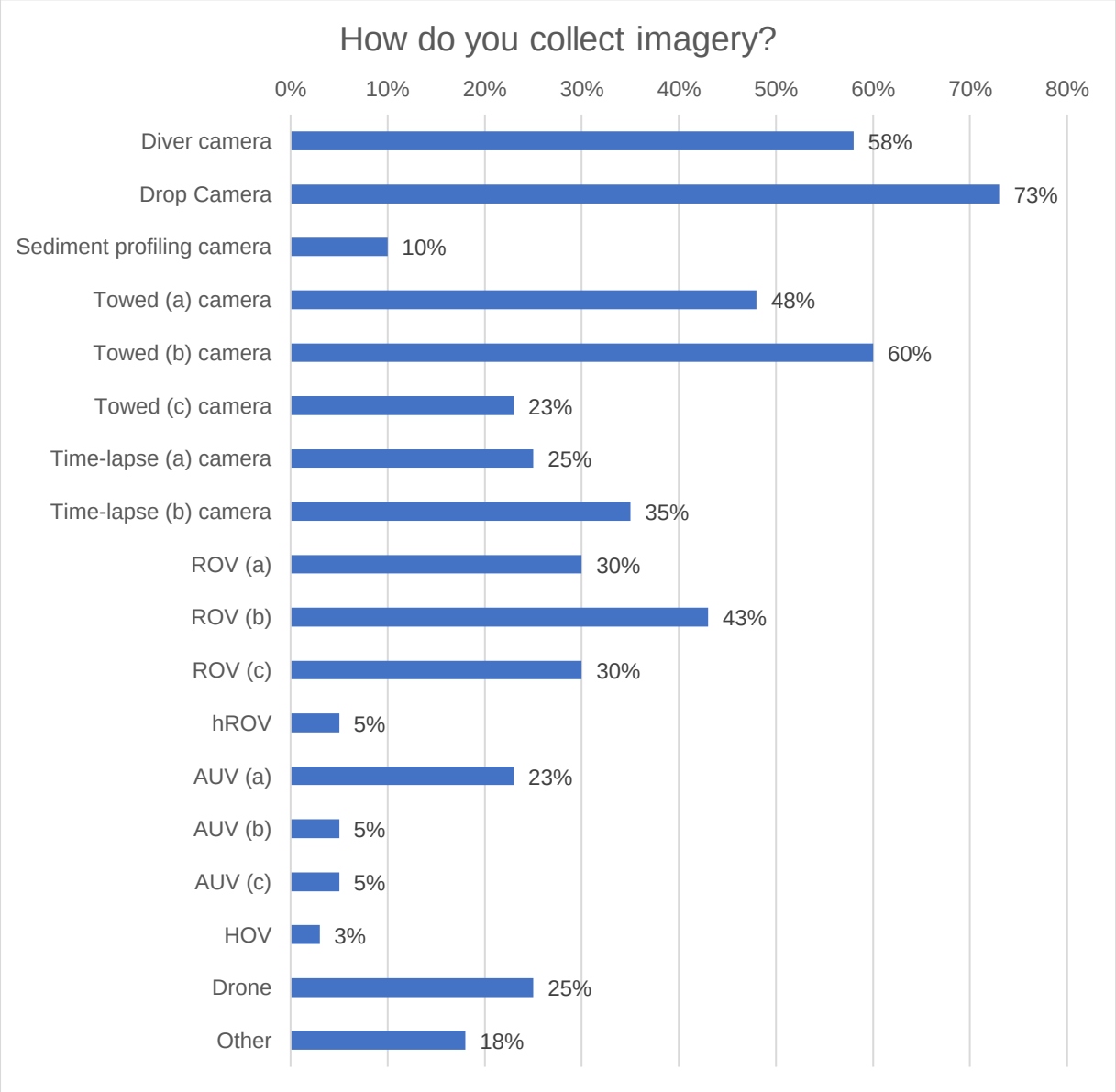
Regulation and designations

- MPA designations.
- Aquaculture regulation.

Other

- Marine accident investigation.
- Groundtruthing of acoustic data prior to offshore industry activity, e.g. oil and gas installations.

2. How do you collect imagery?



Other includes…

- Use of "third party" remote data, and integration to our work - e.g. satellite imagery,

Google Earth.

- Use of sonar techniques to generate imagery - side scan sonar, ARIS sonar camera - and integration to other survey results. Not visual imagery, but is useful.
- LISST HoloCam (microscopic plankton and particles); other pelagic cameras
- Acoustic camera.
- Laser Scanning Camera.
- Freshwater lens drop camera.
- Kits, helium balloon surveys of intertidal.
- Polecam.

Deploying platforms

Across all acquisition methods, average usage of different deploying platforms among respondents is as follows (note, several respondents use more than one):

1. Rigid-Hulled Inflatable Boat (RHIB): **22%**
2. Inshore survey vessel (day use < 20 m length): **45%**
3. Inshore survey vessel (night use < 50 m): **22%**
4. Offshore and deep-sea survey vessel (> 50 m)): **46%**

Features targeted

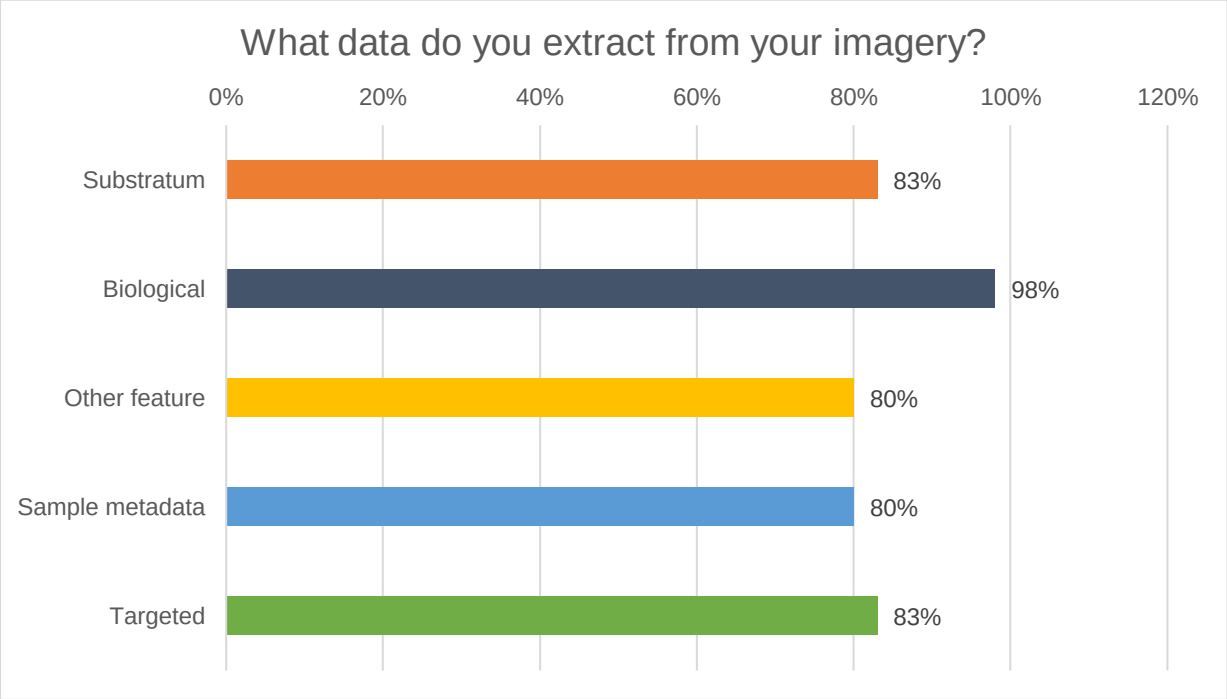
Across all acquisition methods, the most commonly targeted features for respondents are:

- Habitats (including reefs): average **53%**
- Species: average **35%**
- Sediment / substrate: average **24%**
- Communities: average **12%**
- Designated or potential conservation features: average **12%**

Stills and videos

Across all acquisition methods, stills are collected by an average of 80% and videos by an average of 64% of respondents using each acquisition method.

3. What data do you extract from your imagery and how?



Nature of methods used to extract data, across all data types

	Bespoke	Standardised	Not specified
Substratum	45%	21%	33%
Biological	48%	15%	38%
Other feature	38%	3%	59%
Sample metadata	44%	13%	44%
Targeted	42%	15%	42%
AVERAGE	**42%**	**15%**	**42%**

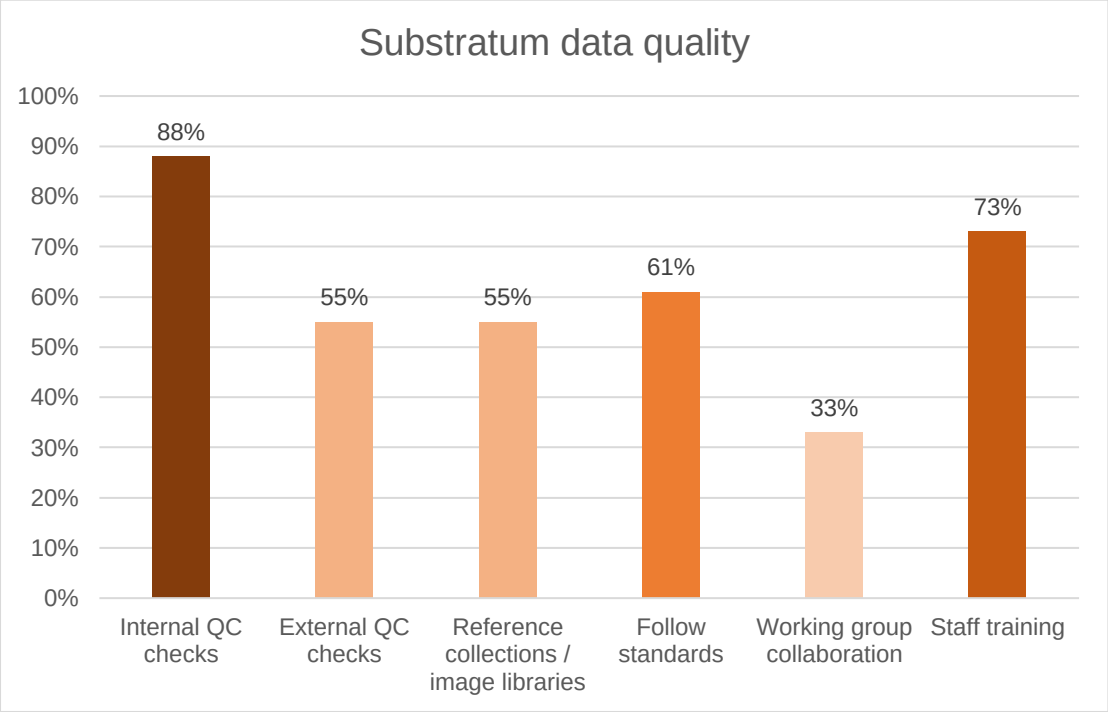
Submission to databases

Across all data types, an average of 31% submit to Marine Recorder and 29% to MEDIN, with an average of 12% confirming they submit, but not specifying where to, and an average of 6% referencing a bespoke or in-house database. Other databases or submission points include:

- BODC.
- CEDAR.
- DASSH.
- Dutch National Waterboard (RWS).
- Gems.

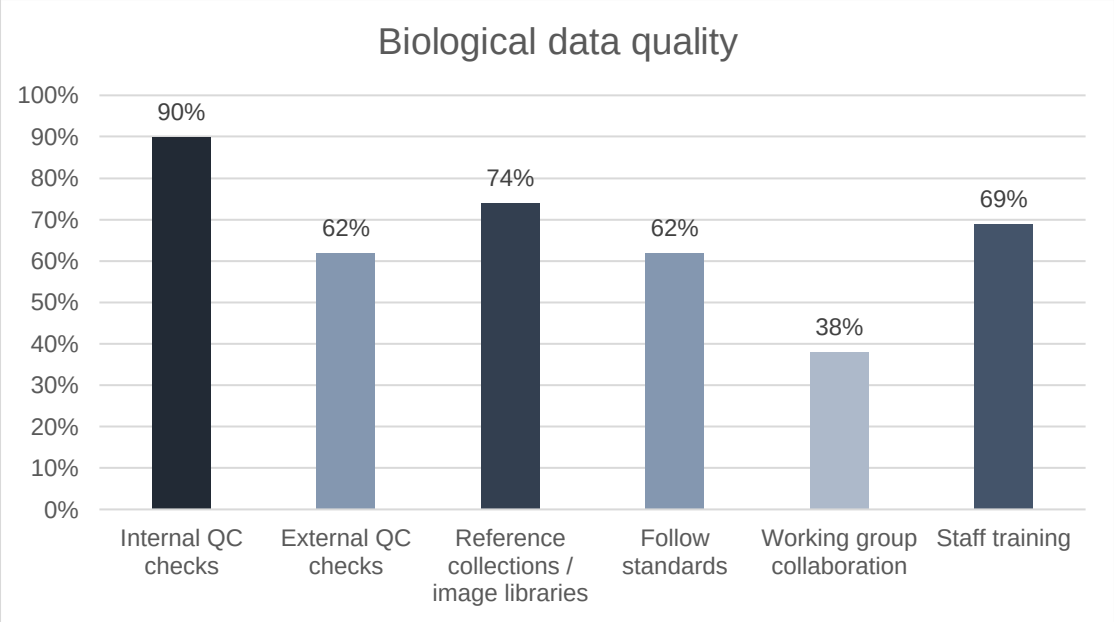
- Grey literature.
- ICES.
- OBIS
- SNH / Marine Scotland.
- To client.
- WWF.

4. How do you ensure / improve the quality of your data?



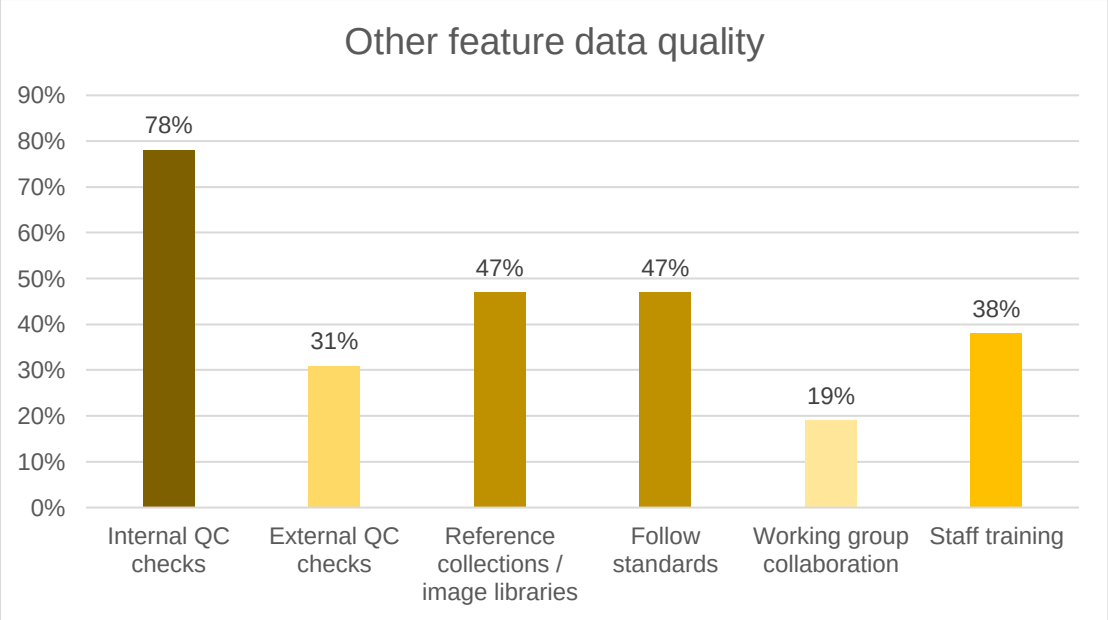
Specific experiences / specialities:

- Comparison with PSA data to be trialled in 2019.
- MarLIN training, NMBAQC, external biotope training, CoCOast, internal training on species ID and video analysis, partnership working with AFBI etc.
- Substratum imagery ground truthed by PSA sampling.
- Deriving data to support habitat classification and MPA monitoring.
- Personally undertake additional diving to familiarise myself with reality of the appearance and composition of substrates as opposed to video imagery which I feel can be misleading. Additionally grab samples are often taken in conjunction with video data, which can add to overall substrate appreciation.
- Involved in previous NMBAQC ring tests.
- Each project has a side-by-side 'training period', with regular of communication throughout jobs with designated QA time per week, and a rolling reference collection available to all analysts.
- In-house training.



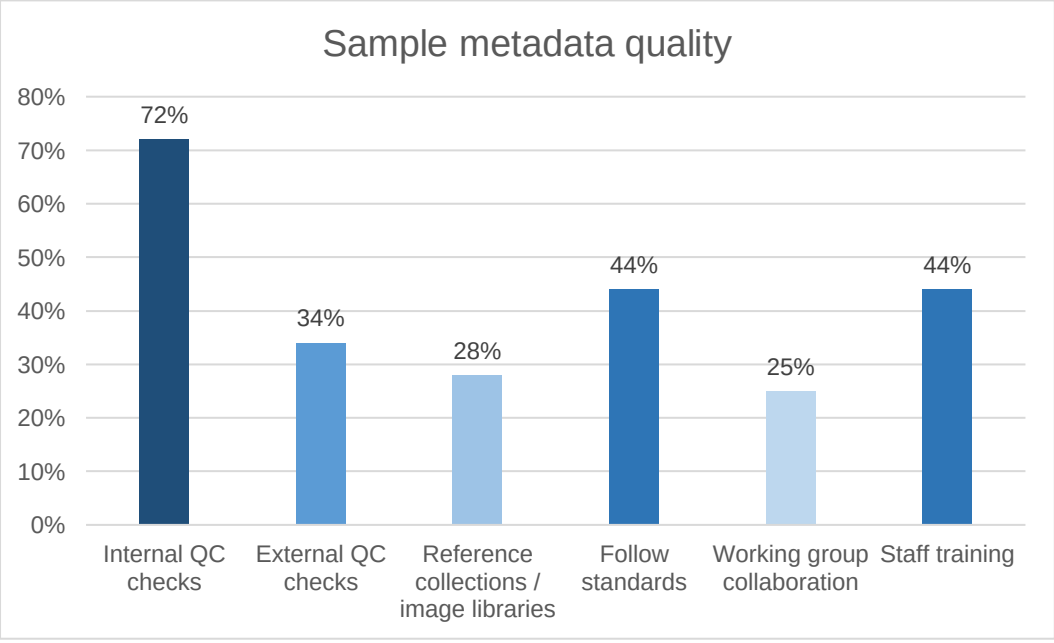
Specific experiences / specialities:

- MPA Baci work.
- Taxonomic QC.
- Training new staff members and volunteers - taxa familiarisation and system use.
- Deep sea North Atlantic.
- In house working group on imagery acquisition and analysis. Literature review underway on current state of the field.
- MarLIN training, NMBAQC, external biotope training, CoCOast, internal training on species ID and video analysis, partnership working with AFBI etc.
- Deriving data to support habitat classification and MPA monitoring.
- Staff have attended in house training with external experts in important local taxa.
- Development of towed video for shellfish surveys is new so no formal QC available yet.
- Have published on the subject (10.3354/meps11775).
- Historically undertook years of lab training, which has been maintained for epifauna in a QC capacity. Knowledge consolidated by extensive work based and recreational diving, including sample collection and species identification.
- Involved in previous NMBAQC ring tests.
- Each project has a side-by-side 'training period', with regular of communication throughout jobs with designated QA time per week, and a rolling reference collection available to all analysts.
- In house training with external experts in important local taxa.



Specific experiences / specialities:

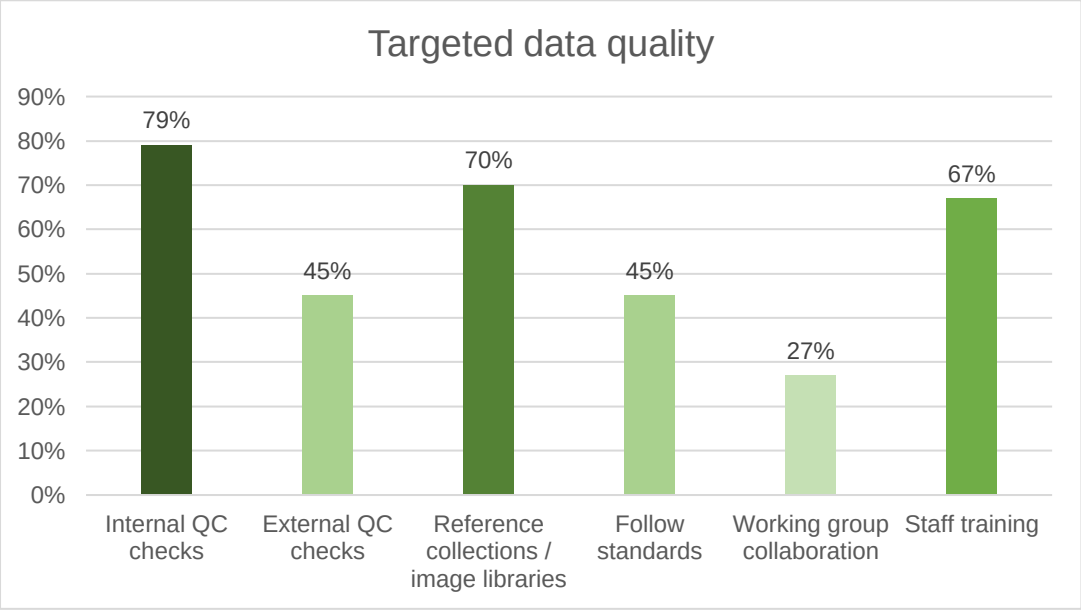
- Presence / absence.
- Have not felt the need for 'standards' etc re debris/litter/trawl marks? Maybe I've not understood the question?
- Each project has a side-by-side 'training period' if seeking a specific feature / impact. See above.



Specific experiences / specialities:

- Depends on acquisition system.
- MarLIN training, NMBAQC, external biotope training, CoCOast, internal training on species ID and video analysis, partnership working with AFBI etc.
- WFD monitoring and Aquaculture regulation.

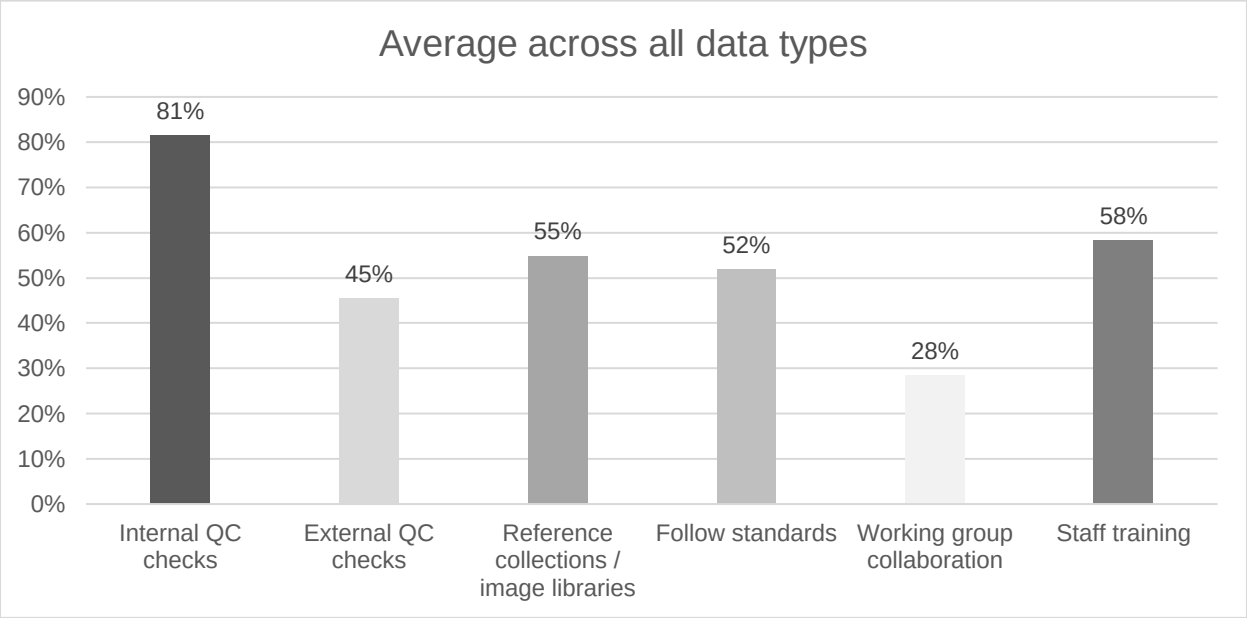
- Each project has a side-by-side 'training period' if seeking a specific feature / impact. See above.



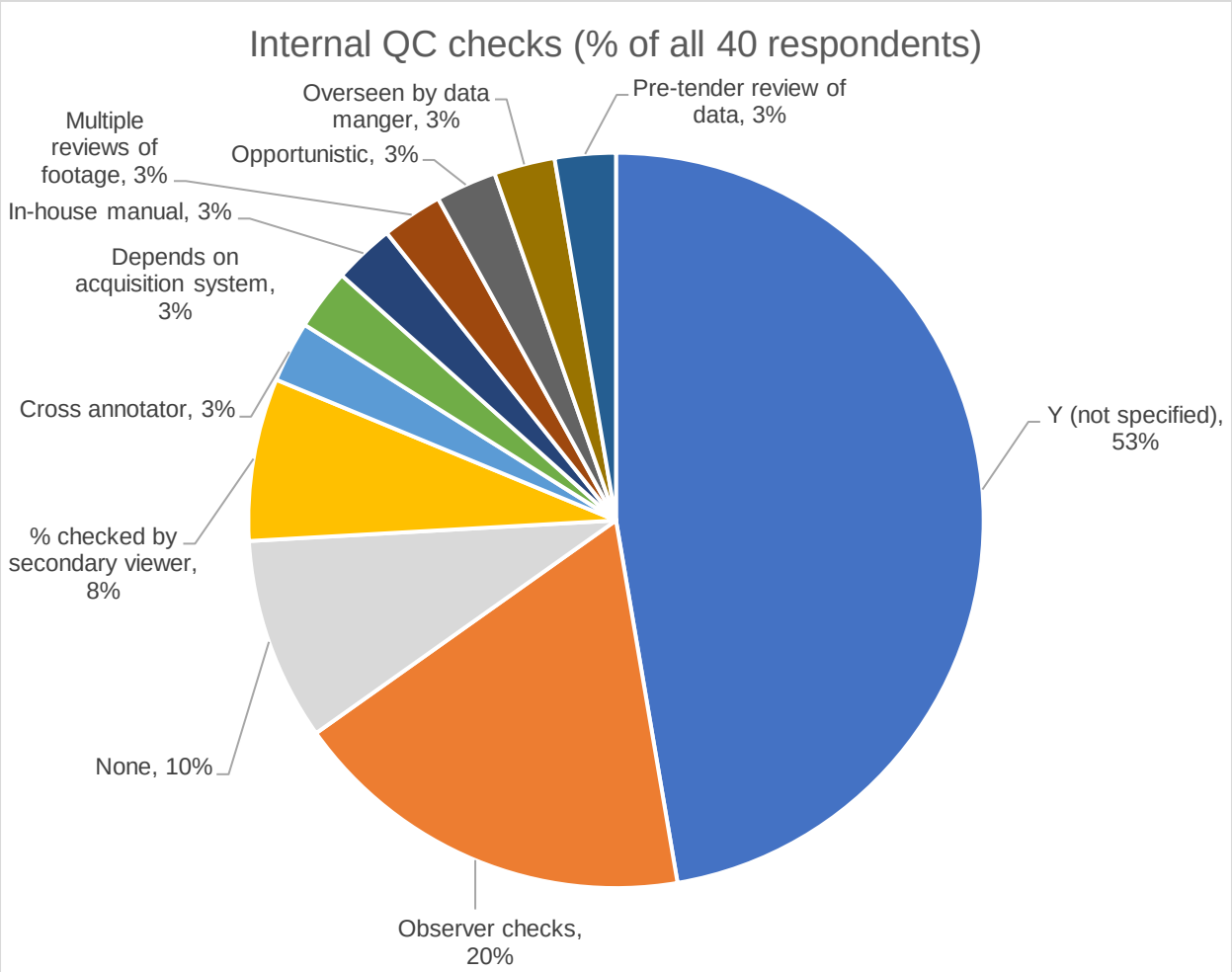
Specific experiences / specialities:

- Species distribution models.
- In house working group on imagery acquisition and analysis includes discussions with fisheries teams.
- MarLIN training, NMBAQC, external biotope training, CoCOast, internal training on spp ID and video analysis, partnership working with AFBI etc.
- WFD monitoring and Aquaculture regulation.
- Burrow identification and use of quality control for counts.
- Much of this work is with *Sabellaria* - staff have attended numerous workshops etc. on the approaches t monitoring *Sabellaria*.
- Each project has a side-by-side 'training period' if seeking a specific feature / impact. See above.
- In-house training.

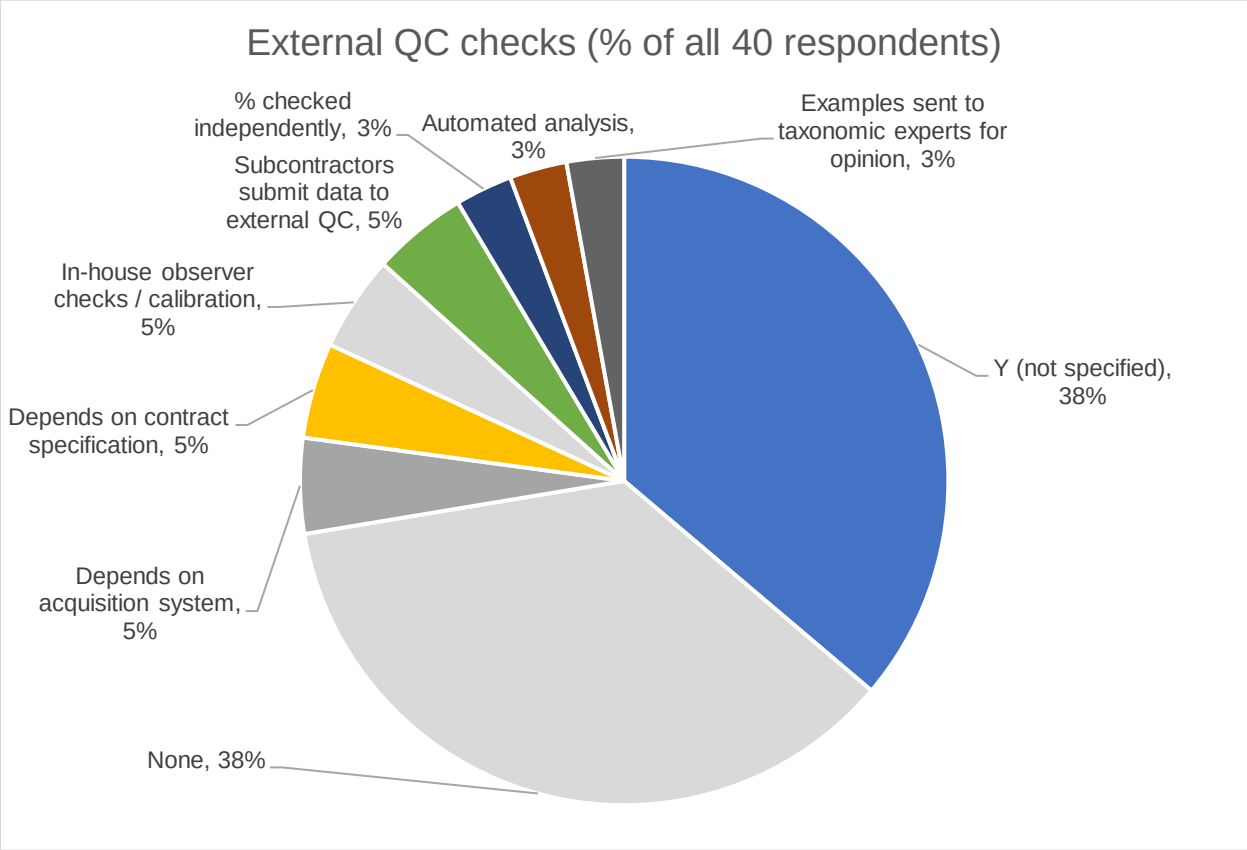
Average figures



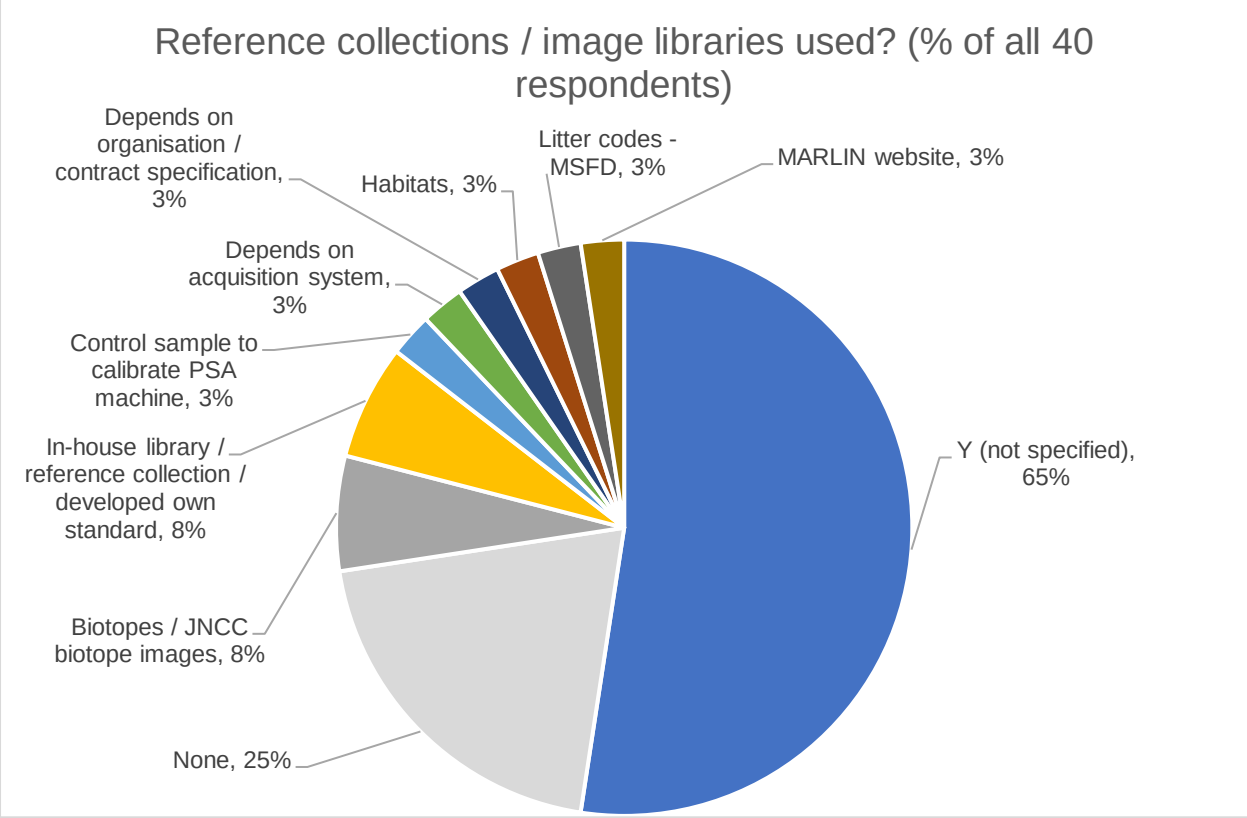
Internal QC checks



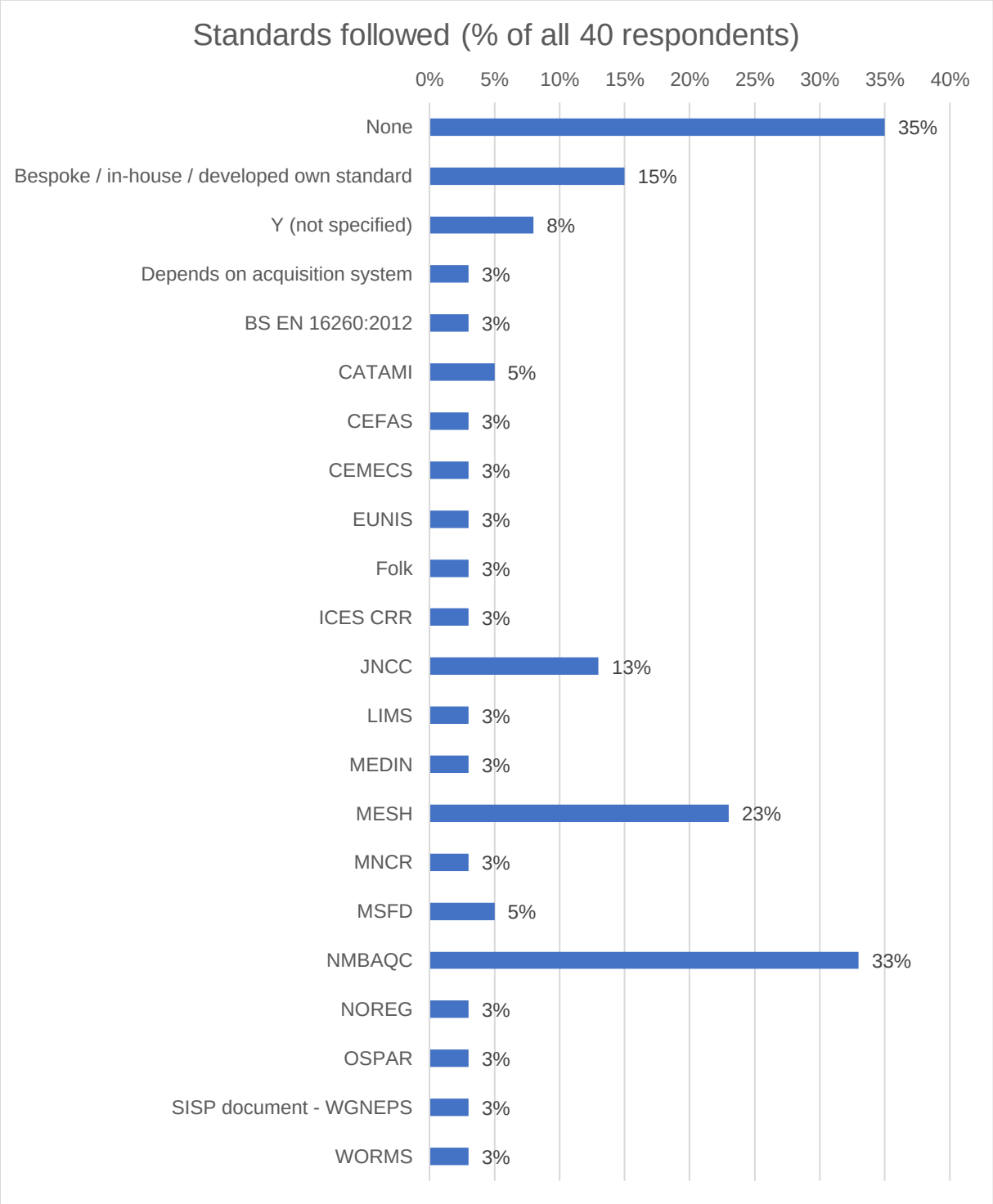
External QC checks



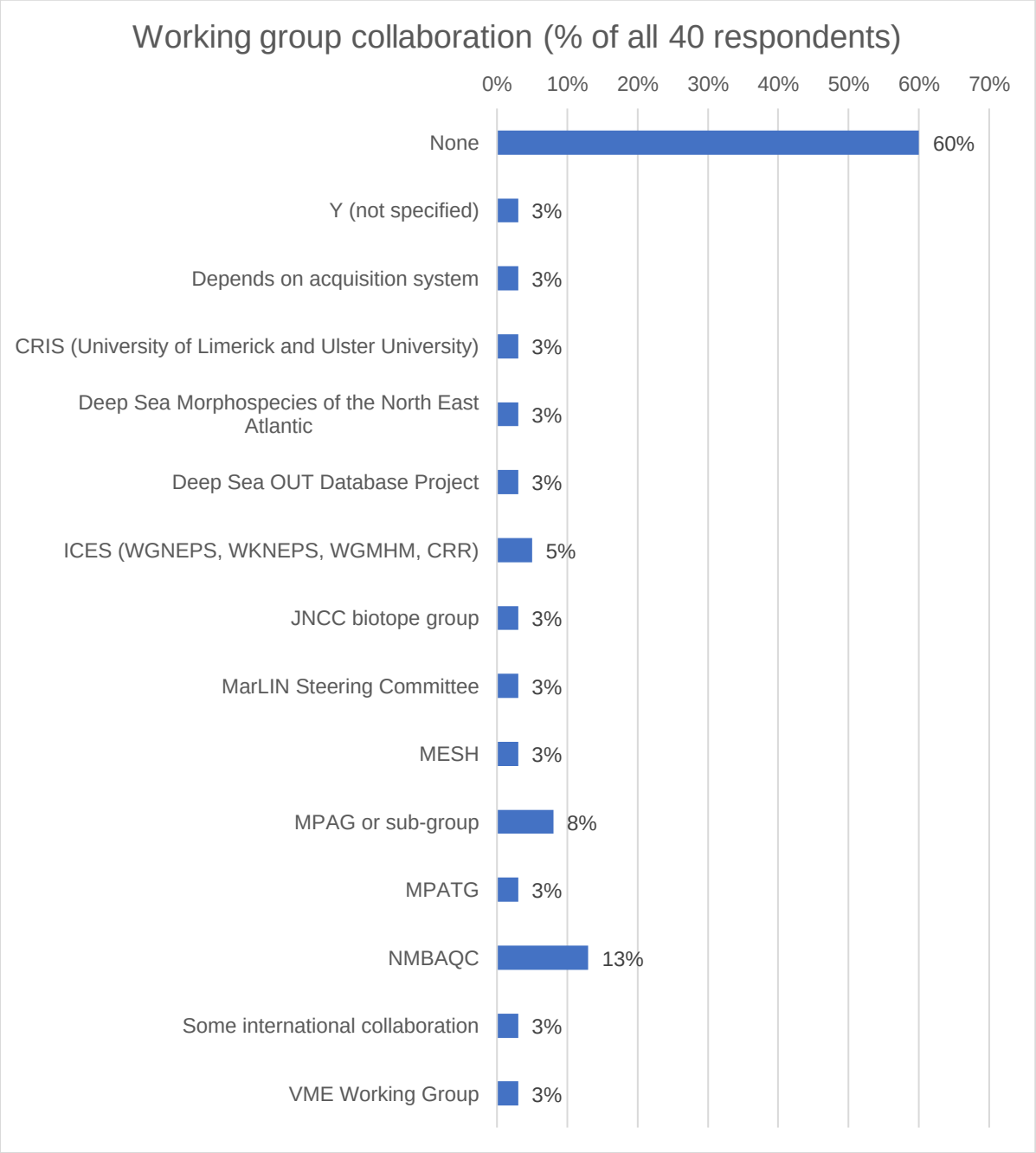
Reference collections / image libraries



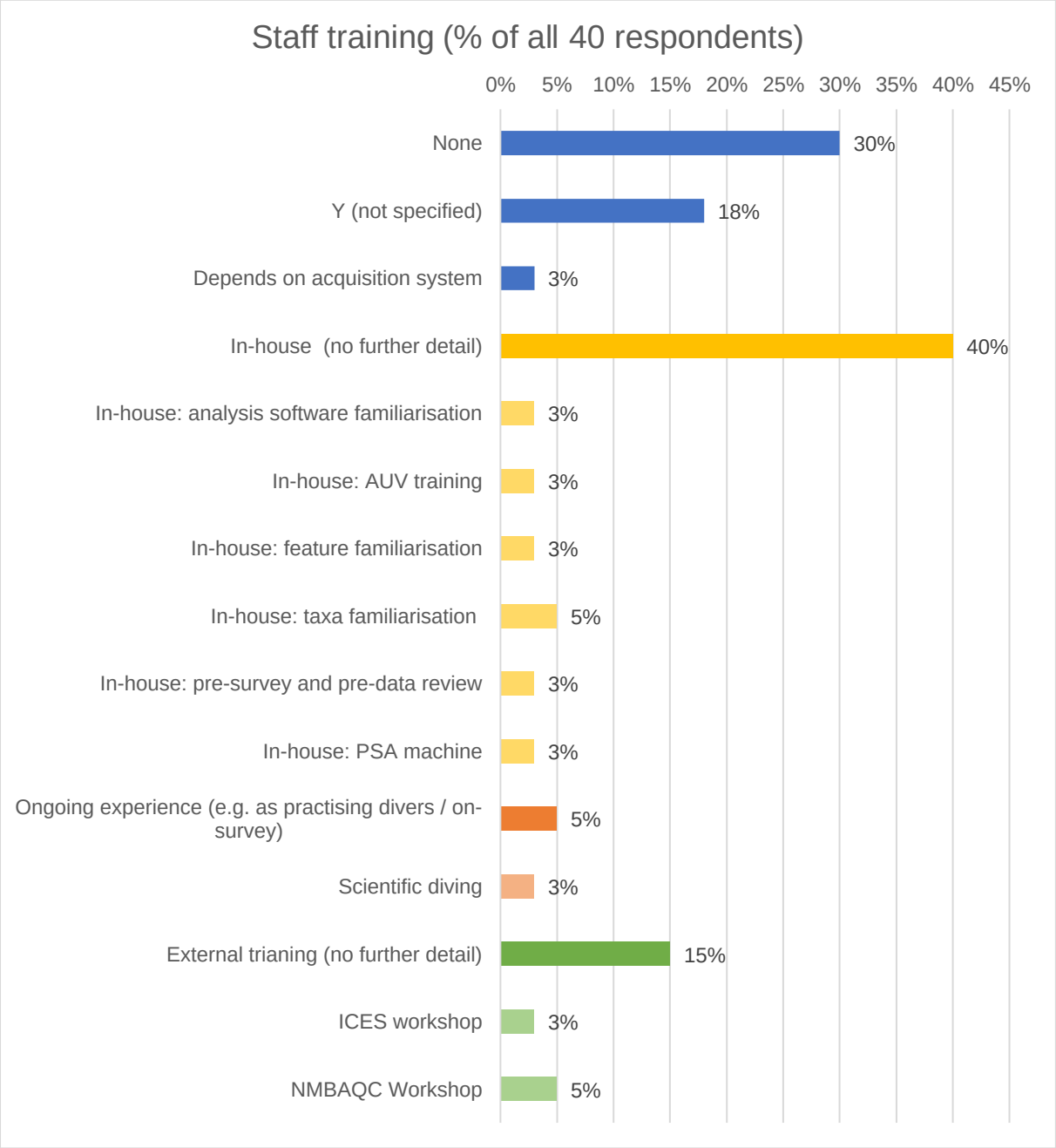
Standards


Standards followed (% of all 40 respondents)
0%
5%
10%
15%
20%
25%
30%
35%
40%
None
35%
Bespoke / in-house / developed own standard
15%
Y (not specified)
8%
Depends on acquisition system
3%
BS EN 16260:2012
3%
CATAMI
5%
CEFAS
3%
CEMECS
3%
EUNIS
3%
Folk
3%
ICES CRR
3%
JNCC
13%
LIMS
3%
MEDIN
3%
MESH
23%
MNCR
3%
MSFD
5%
NMBAQC
33%
NOREG
3%
OSPAR
3%
SISP document - WGNEPS
3%
WORMS
3%

Working group collaborations



Training


Staff training (% of all 40 respondents)
0%
5%
10%
15%
20%
25%
30%
35%
40%
45%
None
30%
Y (not specified)
18%
Depends on acquisition system
3%
In-house (no further detail)
40%
In-house: analysis software familiarisation
3%
In-house: AUV training
3%
In-house: feature familiarisation
3%
In-house: taxa familiarisation
5%
In-house: pre-survey and pre-data review
3%
In-house: PSA machine
3%
Ongoing experience (e.g. as practising divers / on-survey)
5%
Scientific diving
3%
External trianing (no further detail)
15%
ICES workshop
3%
NMBAQC Workshop
5%

5. Text responses to questions 5 and 7

Q5. Please describe any particular issues in your work area that you wish to be discussed in the workshop.

Q7. Please share any general monitoring questions or comments for the workshop participants.

General monitoring and management standardisation for MPAs

- Standard methodologies for MPA monitoring and management.
- What habitats / species work well and give us information about feature condition (marine indicators).
- We can collect all sorts of data but what is it actually telling us - what do we need to be looking for to give us meaningful assessments of feature condition.
- How sure can you be that a change has occurred? (Issues with power, survey design, random image selection, numbers of points in an image analysed, etc).
- Make sure that any monitoring objectives are reasonable, quantifiable and have a defined threshold or classification of concern.

B**iotope monitoring (Video-based?)**

- Most appropriate use of video imagery for biotope monitoring.
- Differences in data captured between video and stills. It is hard to capture patterns in distribution from stills alone for taxa with low density and sparse/patchy distributions in space.
- Video imagery assessment is used by SEPA for regulation purposes. Improved standardised methodology and QC would be beneficial where video data might be used for regulatory enforcement.
- Issues with current guidance (which is a vast improvement on past), specifically biotope assignment of mixed grounds / gravels.
- Interpretation issues with SACFOR from video.

Acquisition of imagery

Improving quality of images acquired

- Use of quality scales to ensure consistency yet quality often poor

Site weather effects

- Working in inshore environment without dynamic positioning - survey vessels are often subject to tides/swell/poor visibility - how can these be mitigated or data quality improved?

Specialised equipment (or not!)

- Use of freshwater lens systems in poor vis environments.
- Options for high res dropcam systems, currently we are still using a SD resolution

camera system.

- Experiences in (baited) camera systems, fish identification, ecosystem development monitoring.
- Our work involves the application of knowledge of the location and condition of features and habitats in connection with the management of commercial fisheries. We are obliged to implement management measures based on best available evidence - and the key word for us is often "available"! In many instances, it is better to have something which is "fit for purpose" (i.e. "good enough") now rather than waiting for perfection which never arrives. Digital imagery is a field which is making extremely rapid strides, and what was very special a few years ago is now commonplace. That being the case - do we always need to go to the latest "all singing all dancing" equipment - which is usually expensive, large, heavy and power hungry? If a small light cheap bit of kit available now can do the same job that the "professional" level kit did a few years ago - when that professional kit was considered what was needed and appropriate - do we need to go to the level of today's professional level kit? After all - "good enough is good enough", and if image quality from five years ago was "good enough" then, why is the same image quality not good enough now?
- Application of eyeball class mini-ROV user experiences.

Development of standardised sampling approaches

- Using towed system for density & condition of species e.g. Pink sea-fan.
- Getting the most out of the images your equipment can produce. Developing a standardised method for imagery using this particular AUV as all AUVs have different capabilities. What are the most appropriate benthic imagery working practices for the imagery this particular AUV can produce?
- We are one of the few companies in the Netherlands with a team of commercial divers who are specialised in biological monitoring and reef restoration. Personally I have many years' experience in underwater photography with DSLR systems. In the near future we plan to buy an eyeball class mini-ROV.
- Suitable methods of ground truthing side scan sonar / multibeam surveys, minimum specification of camera systems.
- Drop Camera standardisation and quantification of imagery data for monitoring purposes / definition of what imagery data can be used for.

Sampling units

- How much data? And when to subsample when an area of up to 1km square can be imaged in its entirety – lines, box surveys or 100% coverage?
- Definition of appropriate sampling units.
- Required area to be sampled for video monitoring.
- Very variable ways of using still images for seabed characterisation from drop down video/stills sleds. Some organisations stipulate recording each still as an individual sample (can be thousands - pointless repetition within the same habitat - and seabed area too small to estimate abundance properly) whereas others stipulate recording overall abundances from all stills taken for example from within one habitat (much quicker and more accurate representation of abundance).
- How and when to subsample footage for habitat mapping applications.
- Are people experiencing how different agencies use stills and videos; some entering each image as a sample, others (sensibly) lumping their images to assess species and

substrates and entering the video habitat splits as a sample (augmented by rapid analysis of 50images for instance), rather than individual analysis of every image. Although this is needed for some monitoring trials?

- How are people using SACFOR scales, particularly with respect to large species in small sample areas? Do others diligently calculate the area of images for each tow / habitat? With a 25m2 minimum survey area to assign biotopes and, I think I read in early MNCR / SACFOR guidance, that 100m area is what SACFOR was developed for. I am concerned for inconsistencies and misinterpretation here. And I know different agencies are using them differently.
- Difficulties in achieving MESH standards due to lack of precision (± 10m rather than cm's) when recording sample station co-ordinates.

Data extraction (annotation)

Need for improved levels of method standardisation and QC

- Existing regulatory guidance and requirements are very broad scale which leads to multiple interpretation resulting in dilution of data quality. This is particularly true regarding definition of OSPAR habitats.
- Issues with standardising enumeration and identification of taxa from stills imagery (specifically ensuring 'area of survey' is large enough to use SACFOR scales as they were designed), and how to minimise inter-surveyor variability.
- Currently there are difference in the standard requested from different agencies (e.g. CEFAS, SNH, JNCC, etc.) that create redundancy in the process. Ideally a single standard should be implemented across the whole industry and it should be recognised by all the regulators. A dedicated and approved software would be great to (especially for data entry that takes the majority of time).
- Repeatability of sampling method and repeatability of annotation / QC.
- Updating best practice in deriving substratum and biological data (MESH ROG is now quite old!), and in particular how to address quality control of these data e.g. ring tests, use of independent counters etc. Data management/archiving best practice (HD & stills).
- Quality of data acquisition and standardised approaches for analysis of imagery. For example as provided by the NMBAQC scheme but to a standard implemented for macrofaunal benthic samples under NMBAQC scheme would be good. With expert workshops and training.
- Given the increasing reliance on imagery for habitat mapping and MPA monitoring, how is quality control (and assurance) best designed and implemented for such studies? Can we learn from repeatability type statistics as used for *Nephrops* UWTV surveys?
- If possible, please chose to specify requirements rather than recommendations to enforce wider industry adoption.

Image annotation software

- Annotation platforms and data extraction.
- A dedicated and approved software would be great to (especially for data entry that takes the majority of time).

Use of machine-learning for image annotation

- Automated recognition of species.

Image analysis training

- Training opportunities for methods.
- Quality of data acquisition and standardised approaches for analysis of imagery. For example as provided by the NMBAQC scheme but to a standard implemented for macrofaunal benthic samples under NMBAQC scheme would be good. With expert workshops and training.
- I am keen to learn and acquire hands-on skills in data acquisition, processing and analysis of benthic imagery acquired from a range of techniques. This will form a better part of my PhD which is focussed on monitoring biological features in special areas of conservation using standardized and refined methods.

Improving cost and efficiency of image analysis

- We are currently working on the analysis of a large video survey we performed during summer 2018 on the Cleaver Bank (North Sea). Footage was acquired with a large ROV (Saab Seaeye Panther Panther XT Plus) at open sea with a dedicated 4K video camera system (SubC). Analysis is very time consuming.

Image analysis contractual agreements and resources

- Much of the guidance on analysis and QA does not seem to link with available budget allocation or reasonable timeframes, so I worry about the robustness of some analysis / evidence that is inevitably rushed.
- Duplication of records in Marine Recorder (when entering enumerated and SACFOR data from the same surveys). For example, some imagery can easily be viewed and analysed to 'standards' within 3x the video time, but on more complex habitats, or unusual methods or some focussed trials, analysis can take 10x (or even more) of the video time to be robust. Many agencies who let contracts do not appreciate this, and neither do many contractors until they've done a lot of imagery work. Just two weeks ago somebody contacted me regarding an agency contract with an estimated analysis time of 15 days for one person. I strongly suspect, that same contract could take 3 people at least 60 days full time, before external QA, MR data entry or reporting. I've heard of some companies (not our company) actually asking their employees to work 12 hr shifts (night and day), 7 days a week to get the analysis of some mis-resourced jobs done within the tight time frames. This is not only bad practice, but also a health and safety concern as imagery is not nice on the eyes, and will drive good companies away from taking on work in future (if it doesn't bankrupt them first). The rush / surveyor fatigue also risks driving down the quality and, with more surveyors required to meet deadlines, increases inter-surveyor variability. I really think this is something that needs discussing to raise awareness between agencies and contractors, especially as imagery gets better quality resulting in some old standards on resource allocation becoming increasingly out of date.
- Obtaining feedback on analysis carried out.
- Are other people finding many contracts unfeasible? How have they dealt with increasing size jobs to deliver good quality evidence within time scales?
- Are what are becoming 'standard' QA procedures feasible with resource limitations? Or is it driving the more experienced analysts out of the competition?

Taxonomic identification

- Identification of taxa from stills imagery.
- Resolution of taxonomic ID - how low should you go?
- Using a consistent system of species coding which is standard across organisations -

there seem to be many different ones in use but which is best?

- My own experience is not with Imaging or Image analyses, but with taxonomic identification QC and nomenclature, and databases.
- Identification of benthic fauna from stills and video material, what level of ID is sensible/achievable.
- Limited standardisation of identification between surveys means data requires a lot of re-processing to make comparable, which often isn't possible in permitting timelines.
- Potential for standardised adherence to TDP.
- Shared reference collection for analysts to improve consistency.
- Very difficult to standardise how 'accurate' or fine level of resolution you can identify species to using various methods. I know that Becky Hitchin is working on this - can provide input if required. In the past NMBAQC was working towards rather rigid protocols for identification, even though in practice it's often down to individual's experience at similar locations (e.g. through diving) that enables or increases capacity to identify species from variable quality video and stills.
- How do you decide at what linnean rank to identify fauna to, with confidence. How is this reported in a standardised manner?

Morphological classification systems

- Morphological classification systems and their use CATAMI.

Image reference collections

- Shared reference collection for analysts to improve consistency.
- Results aren't clearly shared from publicly funded work so often difficult to know details/results or find reliable sources of imagery ID to match datasets.

Enumeration approaches

- Comparison of various methods for data extraction.
- Consensus on which species should be enumerated using stills or video.
- Semi-quantitative nature of imagery analysis versus need for quantitative data to inform assessment of change to biological communities over time.
- I'd like to discuss the role that absence data has in modelling and our need for it.

SACFOR

- SACFOR is not particularly useful, Image quality (including distance from the seabed) and numbers of points that are analysed, % cover from video is also not useful / reliable
- Providing abundance over SACFOR is important. Presence data for species is useful, but complimentary absence data very rarely recorded. This restricts modelling efforts.
- How are people using SACFOR scales, particularly with respect to large species in small sample areas? Do others diligently calculate the area of images for each tow / habitat? With a 25m2 minimum survey area to assign biotopes and, I think I read in early MNCR / SACFOR guidance, that 100m area is what SACFOR was developed for. I am concerned for inconsistencies and misinterpretation here. And I know different agencies are using them differently.

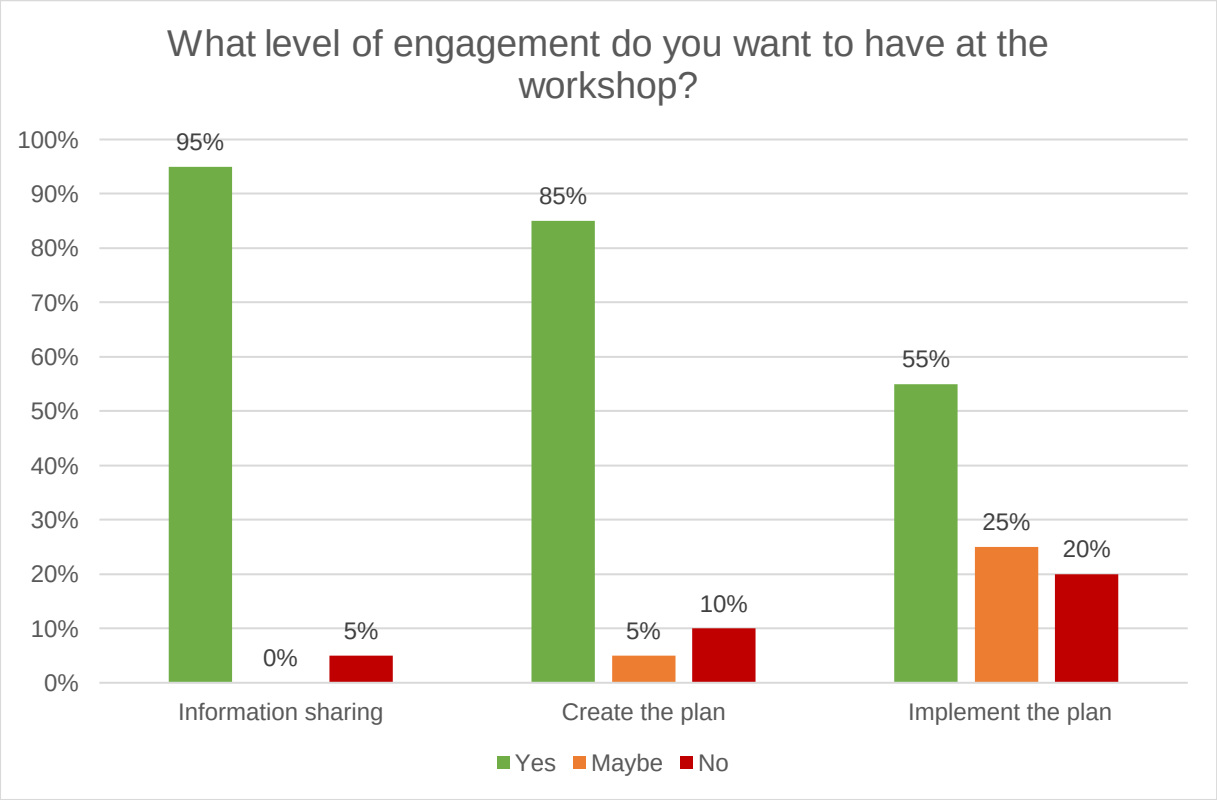
Observer consistency

- Issues around consistency of data i.e. how to ensure a consistent approach has been used to analyse data across different surveys and thus allowing for less noise when collating and using data from multiple surveys. Concerns over surveyor error/different levels of expertise.
- Comparability of imagery datasets over time (e.g. inter analyst variability).

Data sharing and management

- Open nomenclature and appropriate data banking.
- Results aren't clearly shared from publicly funded work so often difficult to know details/results or find reliable sources of imagery ID to match datasets.
- We collect quite a lot of underwater imagery through various projects - these are often PhD and students and main purpose is research, but we are also involved in projects like MarPAMM. I am particularly interested in how we need to be archiving imagery to make it useable to wider community including government stakeholders.
- Some agencies enter species records into Marine Recorder with two species lists (one enumerated and one SACFOR). I am concerned that this will be misinterpreted in future by evidence / managers looking for frequency of records in datasets.
- Please make sure any ID guidance/results are shared openly including allowance for use for commercial purposes.

6. What level of engagement do you want to have at the workshop?



Number of respondents

	Information sharing	Create the plan	Implement the plan
Yes	38	34	22
Maybe	0	2	10
No	2	4	8

Comments

Information sharing

- Yes – anything which will enable our work to be more effective, or efficient (or preferably - both!) will be of value. I am particularly interested in learning about "water lens" cameras and similar.
- Yes - interested to hear about what other organisations are doing and if we can make our work more effective and efficient.

Create the plan

- Y (if my time can be paid for!)
- Yes - BUT it must be borne in mind that protocols and standardise methods must be "fit for purpose" - i.e. GOOD ENOUGH, without necessarily being "gold plated".

- Will never have the time.
- Maybe, to some extent.
- Yes - keen to ensure our methods our standardised and to work towards achieving standards and improving the quality of our data.

Implement the plan

- Y (if my time can be paid for!).
- Limited.
- Maybe?
- Depending on the subject.
- In principle - yes, I'd like to stay in touch with the process - BUT any commitment will depend on the level of resources needed to fulfil it, and the benefits to my organisation.
- Y (with limitation on time)
- Y, if time allows.
- Y if time allows.
- Y (but unable to until 2020 as not currently working full time, but have also deliberately avoided large imagery jobs after 10 years of running them on large scales, as I think the timescales and resource allocation for many are becoming unfeasible on larger jobs. Since last year, we are now only taking selected jobs for JNCC, NRW and private clients.
- No, unless contracted.
- Yes - keen to involved with projects that could benefit our work (subject to having the resources needed).