

Numbers from Images Updated 18th March 2014

In both marine and planetary science in the past, the images collected from the contrasting extreme environments were few in number.

Technological developments in computing power and imaging have changed this. With the vast numbers of images now routinely delivered from extreme ecosystems, statistical analysis of features has become much easier, and many new research hypotheses can be tested.

Within ROBEX, Jacobs University Bremen, together with DLR (and other partners if interested), is developing a 10 week course to introduce some freely available, easy to use and publishable techniques for both turning image data into numerical data, and for drawing scientific conclusions from this new data.

During this brief workshop, we would like to introduce to you some of the techniques from the course.

COURSE WEBSITE: <http://imageanalysis.weebly.com/>

There is a special page 'ROBEX workshop' on the 'Start the Course' tab, which summarises what is available for this meeting:

Basically, the techniques and data from weeks 2 – 4 of the course are available for playing with, and these introduce basic techniques for :

- Determining the numbers of features of interest in an image
- Determining the areas of features within an image
- Determining the lengths of features within an image
- Determining whether features in a transect CO-VARY in abundance

Also introduced are methods to compare the results from two (or more) images or sets of images to see if there are statistical differences in abundance.

The following questions are available for study:

MARINE

BASIC ANALYSIS AND COMPARISON STATS

1. Shrimp are more abundant in the vicinity of corals than away from corals
2. Shrimp cluster at different distances from each other depending on the substrate
3. *Geodia baretii* sponges are more abundant in areas with low live coral cover than areas with high coral cover
4. The role in assessment technique in ecosystem analysis

MARINE TRANSECT ANALYSIS

5. Fish abundance increases with proximity to a cold seep

SPACE

BASIC ANALYSIS AND COMPARISON STATS

6. Crater abundance is higher in the lunar highlands than in the lunar seas
7. Crater diameters are greater in the lunar highlands
8. Secondary craters are more abundant in the lunar lowlands
9. Ice ridges on icy moons are all of comparable length range
10. Ice ridges vary with latitude on Europa
11. The 'blueberries' of Mars are distinct in size class from terrestrial 'blueberries'

MARINE BASIC ANALYSIS AND COMPARISON STATS

1. Shrimp are more abundant in the vicinity of cold-water corals than away from corals

INTRODUCTION

Cold-water coral (CWC) reefs are dynamic, extreme environments. Commonly situated at depth far too great for divers to reach, the robotic exploration of these ecosystems has been on the increase during the last 10 years.

As with tropical coral reefs, CWC reefs are known for a high biodiversity and abundance of animal individuals. Sessile (fixed) animals like corals are relatively easy to count, whereas mobile animals can be more difficult, escaping from ROVs etc.

In this study, instrument platforms (Campods) were flown ~1m above the reef environment at ~1m s⁻¹. A camera faced vertically downward with a fixed lamp providing illumination. Two reefs on the Norwegian margin were visited – the Rost reef and the Sotbakken reef.

Shrimp are very abundant in areas of Norwegian waters, and several species are of commercial interest. In the last few years there has been anecdotal evidence published indicating that shrimp may be more abundant at CWC reefs than off reefs. This is of interest to fisheries as a loss of CWC environments may then have an impact on the commercial interests of fishermen. In this brief study we will investigate with a small dataset whether or not shrimp are indeed more abundant in the vicinity of corals than away from corals.

DATA

Two sets of 20 images are provided:

DATA_Rost_shrimp 20 images of the seafloor from the Rost reef
DATA_Sotbakken_shrimp 20 images of the seafloor from the Sotbakken reef

SOFTWARE AND WEB LINKS

(visit the resources page for links : <http://imageanalysis.weebly.com/resources.html>)

ImageJ will be required on your work computer

Access to an ANOVA calculator is needed, such as the **Turner** calculator.

TASKS

Several investigations can be made with this data. To investigate the hypothesis '**Shrimp are more abundant in the vicinity of cold-water corals than away from corals**' the following steps should be followed:

- 1) Images from the two datasets should be combined and then divided into image sets with no or few corals present, and those with lots of corals present.
- 2) Each of the images should be opened in **ImageJ** and the numbers of shrimp counted. The example below shows how the shrimp can be spotted, by the blurred bright reflections of their eyes.



- 3) The shrimp numbers for the two types of images should be compared with an ANOVA test, following the instructions on the resources page.
- 4) Is there a significant difference between shrimp abundance and coral presence / absence? Further tests can also be run, comparing shrimp densities between the two reefs, for example.
- 5) Feel free to write a short summary of your findings. Please comment on : the suitability of the methodology in answering the research question – do you see any strengths or drawbacks with the experimental or analytical approaches?, on the statistical results given by the ANOVA and any other observations or opinions. Email your summary to a.purser@jacobs-university.de for checking and for combining with the other group data.

MARINE BASIC ANALYSIS AND COMPARISON STATS

2. Shrimp cluster at different distances from each other depending on the substrate

INTRODUCTION

Cold-water coral (CWC) reefs are dynamic, extreme environments. Commonly situated at depth far too great for divers to reach, the robotic exploration of these ecosystems has been on the increase during the last 10 years.

As with tropical coral reefs, CWC reefs are known for a high biodiversity and abundance of animal individuals. Sessile (fixed) animals like corals are relatively easy to count, whereas mobile animals can be more difficult, escaping from ROVs etc.

In this study, instrument platforms (Campods) were flown ~1m above the reef environment at ~1m s⁻¹. A camera faced vertically downward with a fixed lamp providing illumination. Two reefs on the Norwegian margin were visited – the Rost reef and the Sotbakken reef.

Many questions are still open as to how the various species at CWC reefs interact. With coral species being sessile, when they settle as juveniles, they cannot again reposition to maximise their environment for growth and development. Mobile organisms can do so however.

Shrimp have been anecdotally reported to feed off the material collected by CWC species tentacles'. If this is the case, then it would seem likely that the distribution of shrimp on different coral species may vary, with the gaps between shrimp perhaps reflecting differences in the coral polyp distribution. Possibly also the different coral species may collect different sorts of food, and thereby supply different concentrations and types of food to shrimp.

In this short study, the various minimal distances between shrimp on various substrates can be compared, to see if substrate (coral or seabed type) is related to shrimp spacing.

DATA

Two sets of 20 images are provided:

DATA_Rost_shrimp 20 images of the seafloor from the Rost reef
DATA_Sotbakken_shrimp 20 images of the seafloor from the Sotbakken reef

SOFTWARE AND WEB LINKS

(visit the resources page for links : <http://imageanalysis.weebly.com/resources.html>)

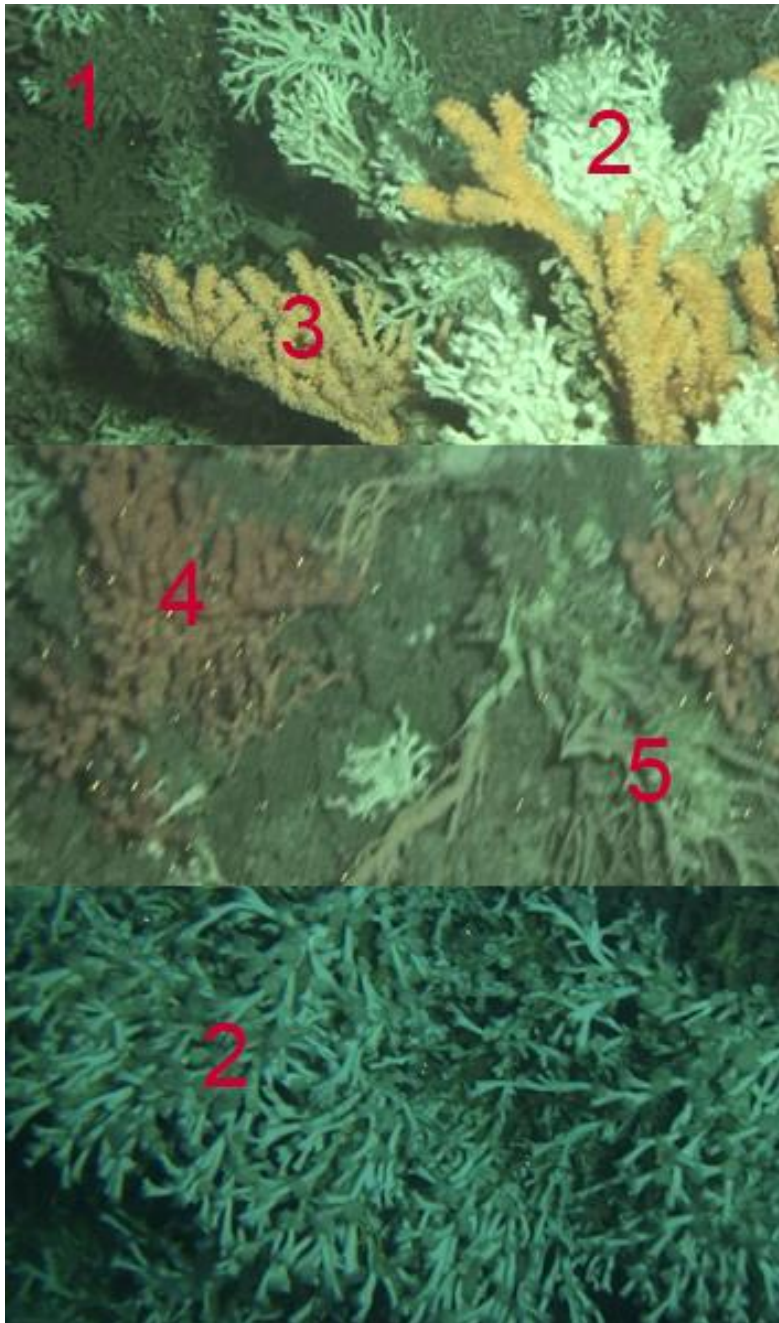
ImageJ will be required on your work computer

Access to an ANOVA calculator is needed, such as the **Turner** calculator.

TASKS

Several investigations can be made with this data. To investigate the hypothesis '**Shrimp cluster at different distances from each other depending on the substrate**' the following steps should be followed:

- 1) There are a number of different substrates present in the presented images, as shown here:



1) Dead stony coral, 2) Live stony coral, 3) Primnoa gorgonian coral, 4) Paragorgia gorgonian coral, 5) Dead gorgonian coral.

Further categories of substrate may also be 6) no coral... feel free to explore the data and draw your own ideas.

- 2) The shrimp are identified by the bright eye reflections on the images (See **Task 1** for more info). For each of the images in the dataset, use **ImageJ** to estimate the distances for each shrimp on each substrate to its nearest neighbour on the same substrate. These should be logged into columns for each substrate in excel or some similar software application. For this work we will compare pixel distances and assume the camera is maintained at the same height throughout the study. If you want to take it further, feel free to scale the images within **ImageJ**, estimating an image diameter of 100 cm.

- 3) After determining the nearest neighbour distances, compare these with an ANOVA test, following the instructions on the resources page if using an online calculator.
- 4) Is there a significant difference between observed shrimp minimum distances on different substrates?
- 5) Feel free to write a short summary of your findings. Please comment on :
the suitability of the methodology in answering the research question – do you see any strengths or drawbacks with the experimental or analytical approaches?, on the statistical results given by the ANOVA and any other observations or opinions.
Email your summary to a.purser@jacobs-university.de for checking and for combining with the other group data.

MARINE BASIC ANALYSIS AND COMPARISON STATS

3. *Geodia baretii* sponges are more abundant in areas with low live coral cover than areas with high coral cover

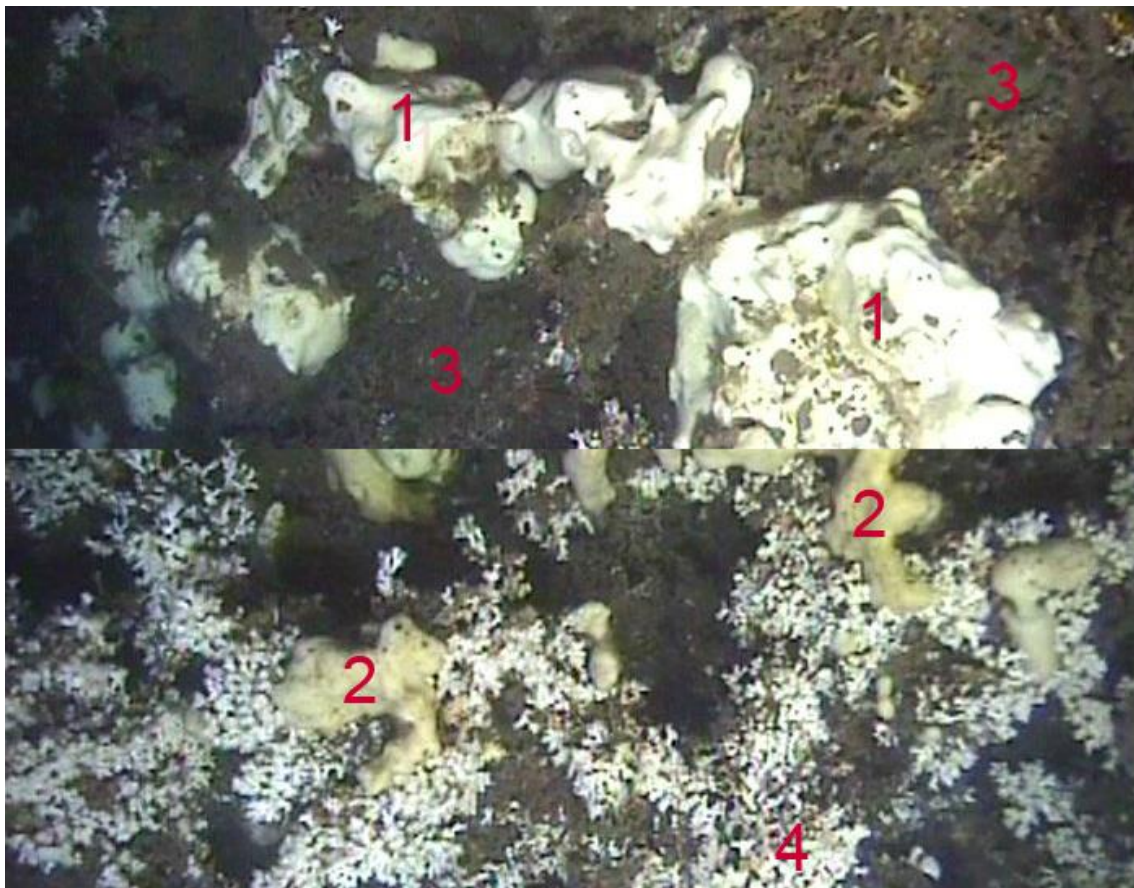
INTRODUCTION

Cold-water coral (CWC) reefs are dynamic, extreme environments. Commonly situated at depth far too great for divers to reach, the robotic exploration of these ecosystems has been on the increase during the last 10 years.

As with tropical coral reefs, CWC reefs are known for a high biodiversity and abundance of animal individuals. Sessile (fixed) animals like corals are relatively easy to count, whereas mobile animals can be more difficult, escaping from ROVs etc.

In this study an ROV was flown at a constant altitude over the Tisler CWC reef, Norway, filming the seafloor beneath it. The Tisler reef is made up of small coral reefs (~2m diameter) surrounded by areas of broken up coral rubble and the sandy seafloor. The reef is quite shallow, formed on a sill at ~90 – 130m depth.

Common components of CWC ecosystems are various sponge species. Sponges are sessile (fixed) animals which survive by filter feeding off material suspended in the water column. Species commonly associated with CWC reefs are *Geodia baretii* and *Mycale lingua*. In this small work we will investigate whether these sponge species are more abundant in areas with many live corals or in areas where there are few live corals.



- 1) *Geodia baretii* sponge, 2) *Mycale lingua* sponge (yellowish, many round holes), 3) Dead coral, 4) live stony coral

DATA

A set of 20 images is provided:

DATA_Tisler_100 grid 20 images of the seafloor from the Tisler reef

SOFTWARE AND WEB LINKS

(visit the resources page for links : <http://imageanalysis.weebly.com/resources.html>)

ImageJ will be required on your work computer

Access to an ANOVA calculator is needed, such as the **Turner** calculator.

TASKS

Several investigations can be made with this data. To investigate the hypothesis '**Geodia baretii sponges are more abundant in areas with low live coral cover than areas with high coral cover**' the following steps should be followed:

- 1) 20 images of typical areas of Tisler reef seabed are provided in the dataset. These images cover an area of $\sim 1\text{m}^2$. There are 100 red circles on each of the images. Within the group identify in as many images as you desire the percentage cover of each type of substrate or fauna within each image. Decide on categories as you see fit, perhaps including categories as in the images above, but feel free to add others, such as 'starfish' or 'muddy seabed'. See the **point analysis** guide .pdf on the resources page if you are unsure how to do this.
- 2) After making the percentage estimates for a number of images, divide the results into categories which reflect coral cover percentage. You decide if $>50\%$ represents 'high coral cover' or if another percentage should mark this point. You can divide the images into a number of categories if you like, such as 'no coral', '<40% coral', '>40% coral')
- 3) After dividing the images, carry out an ANOVA analysis on the data to see if the 'Giodia baretii' coverage % are related to the coral coverage %.
- 4) Repeat the analysis for 'Mycale lingua' cover if you like and have time.
- 5) Feel free to write a short summary of your findings. Please comment on : the suitability of the methodology in answering the research question – do you see any strengths or drawbacks with the experimental or analytical approaches?, on the statistical results given by the ANOVA and any other observations or opinions. Email your summary to a.purser@jacobs-university.de for checking and for combining with the other group data.

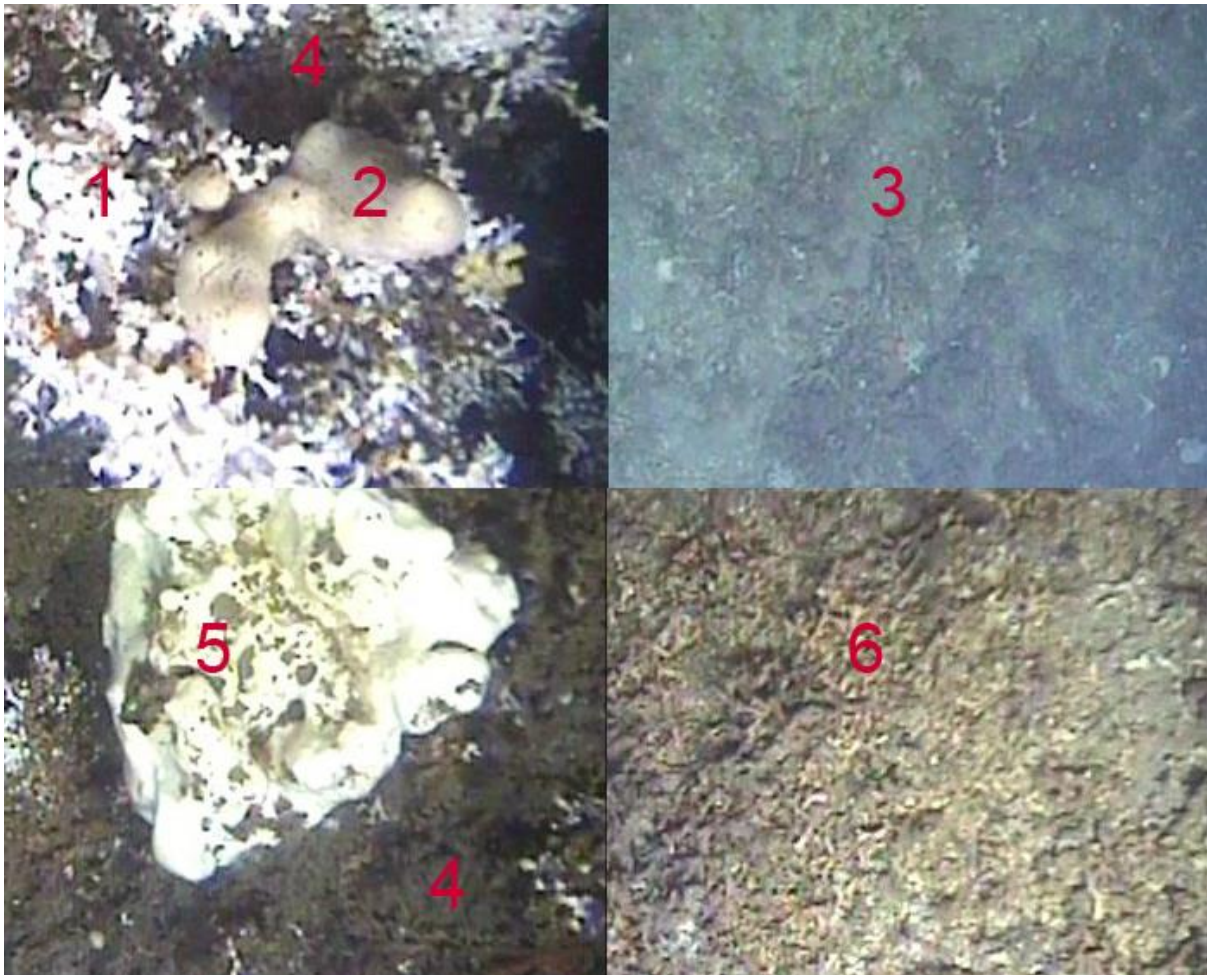
MARINE BASIC ANALYSIS AND COMPARISON STATS

4.The role in assessment technique in ecosystem analysis

The last 20 years have seen a progressively greater flux of image data returned from extreme ecosystems each passing year. How best to analyse this data is a paramount research question.

Many new methods for turning image data into numbers (be these numbers species occupation of the seafloor, substrate type etc) are focusing on machine learning approaches and semi-automated analysis.

To understand how images can be converted into numerical data we will in this brief study take a more basic approach. We will use various image analysis techniques to manually assess seabed coverage of various coral and fauna types in a Cold-Water Coral (CWC) reef environment, and draw some conclusions as to the relative merits of each technique.



1) Stony coral, 2) Mycale lingua sponge, 3) Seafloor, 4) Dead coral, 5) Giodia baretti sponge, 6) coral rubble

In this diagram 6 key categories of image coverage commonly observed at the Tisler coral reef (See **Task 3** for more info) can be seen. In this study we will look at 4 techniques for estimating the coverage of the seafloor in a set of images by each of these categories.

DATA

A set of 20 images is provided:

DATA_Tisler_Raw	20 images of the seafloor from the Tisler reef with point grid.
DATA_Tisler_10 grid	20 images of the seafloor from the Tisler reef with point grid.
DATA_Tisler_100 grid	20 images of the seafloor from the Tisler reef with point grid.

SOFTWARE AND WEB LINKS

(visit the resources page for links : <http://imageanalysis.weebly.com/resources.html>)

ImageJ will be required on your work computer

Access to an ANOVA calculator is needed, such as the **Turner** calculator.

The '**Quantifying area coverage**' .pdf from the resource page is also needed

TASKS

Several investigations can be made with this data. To investigate '**The role in assessment technique in ecosystem analysis**' the following steps should be followed:

- 1) Read the '**Quantifying area coverage**' .pdf from the resources page.
- 2) Firstly, set up a spreadsheet with the six coverage categories columns (as described above - 1) Stony coral, 2) Mycale lingua sponge, 3) Seafloor, 4) Dead coral, 5) Giodia baretta sponge, 6) coral rubble)
- 3) Go through the 20 images in the **DATA_Tisler_Raw** dataset and guess at the percentage coverages of each category in each image.
- 4) Go through the 20 images in the **DATA_Tisler_10grid** dataset and make percentage estimates for coverage based on the 10 marked points.
- 5) Go through as many of the 20 images in the **DATA_Tisler_100grid** dataset as you can face, and make percentage estimates for coverage based on the 100 marked points.
- 6) Go through some of the 20 images in the **DATA_Tisler_Raw** dataset , and map the coverage of each of the six categories using ImageJ. Refer to the ImageJ instructions if required.
- 7) Choose a category, such as 'stony coral', and compare the results from each technique using an ANOVA test.
- 8) If you have sufficient time, compare some of the other categories with ANOVA tests.
- 9) Feel free to write a short summary of your findings. Please comment on : the suitability of the methodology in answering the research question – do you see any strengths or drawbacks with the experimental or analytical approaches?, on the statistical results given by the ANOVA and any other observations or opinions. Email your summary to a.purser@jacobs-university.de for checking and for combining with the other group data.

MARINE

MARINE TRANSECT ANALYSIS

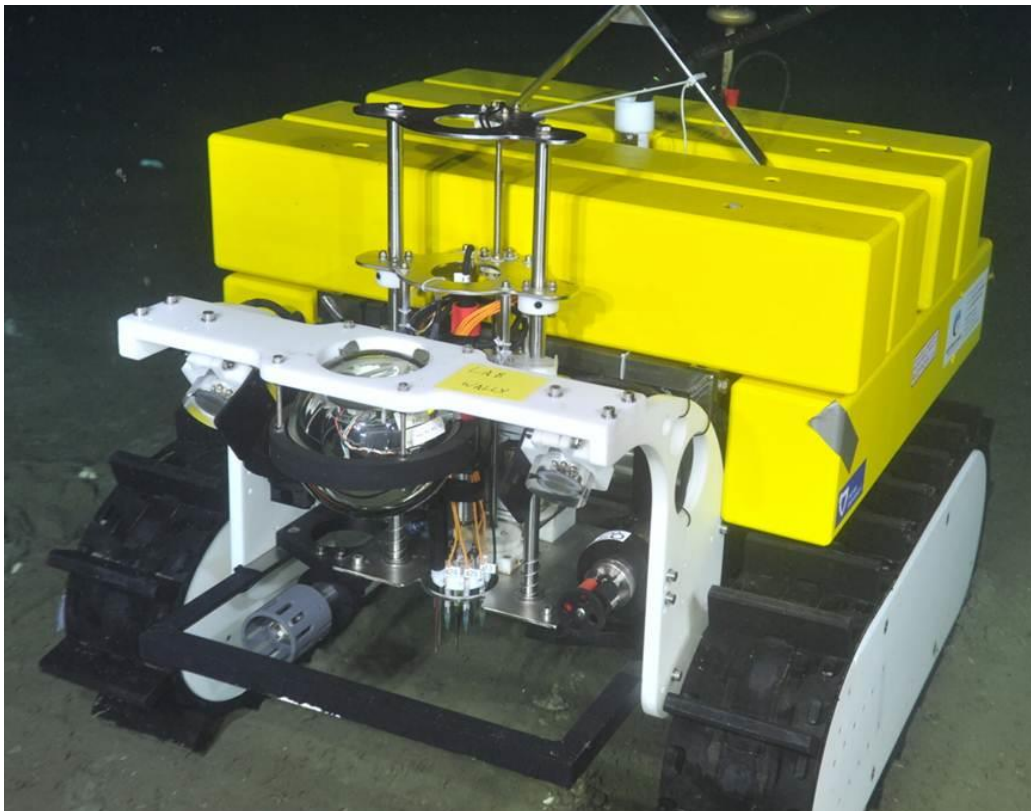
5. Fish abundance increases with proximity to a cold seep or varies with date/time

Robotic exploration of extreme environments in the deep sea has mostly been based on either short duration Remote Operated Vehicle (ROV) investigations or longer duration studies using fixed observatory landers.

The short duration ROV studies allow for spatial analysis of the environment, whereas the fixed observatories allow for temporal analysis.

The long term deployment of mobile instrument platforms, supplied with power and internet connectivity by either cable or via surface communications buoy, allows both spatial and temporal analysis to be carried out simultaneously.

At present, the most significant long-term mobile deep sea extreme ecosystem research vehicle is the Jacobs University Bremen Deep Sea Crawler, Wally.



The Jacobs University Bremen deep sea research crawler, 'Wally'.

For the last three years, Wally has been collecting data from the vicinity of a cold seep pockmark, situated at 900m depth in the Barkley Canyon, Pacific Canada. This vehicle is equipped with standard oceanographic sensors such as Conductivity, Temperature, Depth sensors (CTDs), flow meters, flurometers etc. and several imaging systems of various resolution and use.

One group of collaborators, from CSIC, Barcelona, have been using Wally to film video transects from the cold seep pockmark to a point ~30 m away. They have been logging the abundance of all macro and megafauna over each transect, and repeating these at numerous times within a seven day period each month.

They are interested in two key animal behaviours:

- 1) Do animals keep away from or prefer the proximity of a methane pockmark
- 2) Are the behavioural patterns (abundances) of animals varying over time – either over short diurnal cycles or over longer periods of time.

For the ROBEX workshop, a number of these transects have been made available and Pedro can show you where they are.

How to compare the animal abundances spatially and temporally from the video transects:

- 1) Copy the NEPTUNE Canada guide to species from the laptop (file: marine_life_guide_neptune_canada.pdf)
- 2) Copy the 'ROBEX_BarcelonaFishGroup.xls' spreadsheet, or download it from the website.
- 3) Explore these two documents. You can see the animals logged during two of the transects in the .xls document. The WALLY filecodes for the videos give the timing information, except for the last digit – a '1' means the film was recorded driving away from the hydrate, a '2' means the film was recorded driving toward the hydrate.
- 4) Watch some videos of interest and log the occurrences of animals by time within a video.
- 5) When you have analysed several videos from several months, then these counts can be compared with an ANOVA... do particular fish vary in abundance by month?
- 6) Do the abundances vary with proximity to the seep? For a simple analysis of this the data can be split, either in the middle of the transect, or into more blocks. The numbers of animals in each block can then be compared. The counts from several transects should be combined for this analysis.

SPACE

BASIC ANALYSIS AND COMPARISON STATS

6. Crater abundance is higher in the lunar highlands than in the lunar seas

As for the deep sea, the numbers of images from planetary bodies (extreme ecosystems) has been greatly increasing as a result of imaging, computing, power and data transfer technological improvements.

As in the deep sea, it is desirable to generate numerical data from these image sets for the use in statistically assessing certain research hypotheses.



A glance at the moon shows it to be heavily cratered. To the naked eye it would appear that the craters are more abundant in the highlands (the whiter areas) than the lowlands (the darker areas).

Is this actually the case?

DATA

Download a number of images via the NASA web portals (as demonstrated by the DLR participants in the workshop!) from the lunar highlands and the lunar seas. These images should have reasonably similar spatial coverage areas.

ImageJ will be required on your work computer

Access to an ANOVA calculator is needed, such as the calculator on the webpage.

TASKS

To address this research question, to determine whether or not craters are more abundant in the highlands than the lunar seas, the following steps can be followed:

- 1) Read the '**iass_guidetoimageJ.docx**' from the resources page.
- 2) Download a number of images via the NASA web portals (as demonstrated by the DLR participants in the workshop!) from the lunar highlands and the lunar seas. These images should have reasonably similar spatial coverage areas, and download the same number of images from each category of region.

- 3) Use ImageJ to count the numbers of craters in the images from each area.
- 4) Use an ANOVA test to compare abundances between the two regions.
- 5) If you desire to go further with the analysis, also look at boulders in the two areas.
- 6) Feel free to write a short summary of your findings. Please comment on :
the suitability of the methodology in answering the research question – do you see any strengths or drawbacks with the experimental or analytical approaches?, on the statistical results given by the ANOVA and any other observations or opinions.
Email your summary to a.purser@jacobs-university.de for checking and for combining with the other group data.

SPACE

BASIC ANALYSIS AND COMPARISON STATS

7. Crater diameters are greater in the lunar highlands

As for the deep sea, the numbers of images from planetary bodies (extreme ecosystems) has been greatly increasing as a result of imaging, computing, power and data transfer technological improvements.

As in the deep sea, it is desirable to generate numerical data from these image sets for the use in statistically assessing certain research hypotheses.

The moon offers many opportunities for spatial analysis. One simple question which can readily be investigated is whether or not the diameters of the lunar craters is greater in the lunar sea areas than the lunar highlands.

DATA

Download a number of images via the NASA web portals (as demonstrated by the DLR participants in the workshop!) from the lunar highlands and the lunar seas. These images should have reasonably similar spatial coverage areas.

ImageJ will be required on your work computer

Access to an ANOVA calculator is needed, such as the calculator on the webpage.

TASKS

To address this research question, to determine whether or not craters are more abundant in the highlands than the lunar seas, the following steps can be followed:

- 7) Read the '**iass_guidetoimageJ.docx**' from the resources page. Much of this document is laid out with measuring lunar crater diameters as an example.
- 8) Download a number of images via the NASA web portals (as demonstrated by the DLR participants in the workshop!) from the lunar highlands and the lunar seas. These images should have reasonably similar spatial coverage areas, and download the same number of images from each category of region.
- 9) Use ImageJ to measure the diameter of a number of craters in each image. These should be selected at random and give a good cross-section of those present. The best approach is to measure all the craters in an image, to avoid observer bias.
- 10) Use an ANOVA test to compare diameters between the two regions.
- 11) Feel free to write a short summary of your findings. Please comment on :
the suitability of the methodology in answering the research question – do you see any strengths or drawbacks with the experimental or analytical approaches?, on the statistical results given by the ANOVA and any other observations or opinions.
Email your summary to a.purser@jacobs-university.de for checking and for combining with the other group data.

