SCUFN33-08.2A

Information Paper for Consideration by SCUFN

Undersea Feature Detection Project (UFDP)

Submitted by:	Canadian Hydrographic Service of Fisheries and Oceans Canada
Executive Summary:	Show the recent progress of the Undersea Feature Discovery Project
Related Documents:	B-6
Related Projects:	Development of an S-100 compliant standard for Undersea Features Projects from the SCUFN Undersea Feature Names Project Team; projects from the SCUFN Generic Terms Working Group.

Introduction / Background

Fisheries and Oceans Canada (DFO), is a member of the Geographical Names Board of Canada (GNBC). As Canada's national naming authority, the GNBC has a sub-committee dedicated to undersea feature naming - the Advisory Committee for Undersea Feature Names (ACUFN). The Canadian Hydrographic Service (CHS) of DFO, leads this Committee and maintains the Canadian Gazetteer of Undersea Feature Names.

With the availability of new and finer quality bathymetric data, ACUFN wanted to discover NEW unnamed undersea features in Canadian waters, in order to generate NEW name submissions. Thus began the Undersea Feature Discovery Project (UFDP). There is no technological investment cost attached to this project, other than the human resource. The analysis is based on open bathymetric data and GIS. To be clear, the purpose was only to discover NEW unnamed undersea features, and not to validate or check existing named features.

Up until SCUFN 32 (2019), we had tested several methods to detect undersea features based on the definitions in B-6. We had applied these methods to bathymetric datasets with the objective of standardizing the settings and thresholds used to discover different features (different generic terms). with a desire to build consistency in the result of the analysis. We were hoping that these settings could be standardized.

Analysis/Discussion

This year we compiled a list of attributes for each generic term, for which a setting or threshold would be needed (Annex 1). This information was shared with the Chair of the Generic Terms WG (GTWG), who asked us to set a value for those attributes. The values would be assess by the GTWG, in terms of geological/geophysical contexts. The UFNPT extracted existing values from the GEBCO database (Annex 2) that were not transferred into the list on Annex 1.

By then, our team of University intern students, were studying the contour patterns that are associated with certain undersea features. We then changed our focus onto how to train a machine to detect undersea features, by using the existing records in GEBCO as training data. If this were possible, then maybe we wouldn't need to determine all attributes, settings and thresholds because the machine would learn at least some of them on its own.

The GTWG was updated on the progress made in testing contour pattern recognition methods used for Seamounts (Annex 3)

Conclusions

In our efforts to discover new undersea features in Canadian waters in order to generate new name submissions, we have tested several methods, encountered challenges while trying to automate these methods, and also when assigning quantifiable values to qualitative data. Machine learning, might provide us with the option to bypass the assignation of values, if a machine could do this on its own by using existing GEBCO records as training data. This has the potential to achieve the same desired goal to consistently interpret each generic term, and the successful automation of the discovery and identification of more undersea features.

Recommendations

Continue the investigation of machine learning methods

Justification and Impacts

Standardizing the detection of undersea features to name them, fits into the objective of SCUFN and benefits our understanding about the seafloor.

There will be no impact on existing named features.

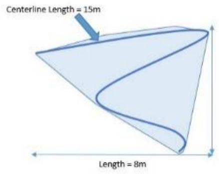
Action Required of SCUFN33

Encourage support and feedback from experts from the following fields of study: GIS, topography, oceanography, geology, earth science, computer science, especially those who are investigating machine learning methods.

Annex 1

Undersea feature
encoding attributes.xls

Attributes: Centerline Horizontal Length/ Centerline Horizontal Width, were used mainly for linear features. They allow us to differentiate between the length of the feature and the length of the area that feature occupies.



Annex 2

Metrics for UFs that have already been adopted into the GEBCO database. (these were compiled by Marine Regions)

SCUFN Generic terms metrics.xlsx

Annex 3

Undersea Feature Detection Project: Update

Loucas Diamant-Boustead | July 21st, 2020

Overview

The automation and specification of parameters for the detection of undersea features was proposed by the Canadian Hydrographic Service (CHS) at SCUFN31, where the initial research findings for Seamounts and Basins was shared. This led to the creation of the Undersea Feature Detection Project (UFDP). The goals of this team were not only to create a python script and ArcGIS toolbox to easily automate the detection undersea features, but also standardize the parameters of these features. Creating a standardized set of parameters for each feature will allow for a better, much clearer understanding of what is required to be considered a specific feature. For example, 37 of the 43 generic terms laid out in the International Hydrography Organizations (IHO) B-6 manual do not contain any quantifiable parameters. The B-6 is the foremost resource when attempting to determine what classification to give to a feature.

CHS's team of students developed parameters and created a detection method for 7 undersea features, including: basins, seamounts, guyots, abyssal plains, shelves, sea channels, and ridges. Of these, basins, seamounts, and guyots had been automated in time for SCUFN32, the others had not yet been automated. Since SCUFN32, automation has been completed for all but sea channels and ridges. The work was completed by Shenghao Shi, Samir Sellars, Oliver Farwell, and Erin Tunrbull. Automation and scripting was completed by Erin Turnbull.

It is important to note that the methods created by these students was just the beginning of this project, and has since been updated and improved upon by Erin Turnbull. Further information about this new method will be detailed further in this report.

Old Method – Raster Analysis

The initial method that was created analyzed a raster of the bathymetric data. Each cell within the raster represents a specific area, and is given one depth value. This means a raster that covers an area of 5x5 km could potentially contain tens of thousands of cells, depending on the data source. This method individually analyses each cell within the specified raster, and in some cases compares it to the surrounding cells. The majority of the methods used incorporate already created tools within ArcGIS, and some, such as basins and seamounts, require other statistical methods. Parameters were tested and fine-tuned based off analysis of both the east and west coast of Canada. The list of new descriptions, which are based off and include the new parameters, is as follows*:

1. **Basins:** A circular area where the edges are steeper and 300 m shallower than the center, with a minimum area of 4 cells. The circular shape is defined by having an ideal circle ratio $[I = \frac{\pi r_{major}}{P}]$ of 1.00 ± 0.15, where r_{major} is the long axis if the rectangular bounding box by width around the feature.

2. **Seamounts:** A rise that surpasses 1000 m in height with a slope that is no less than 2°, while maintaining a circular shape. The circular shape is defined by having a Polsby-Popper ratio

greater than 0.5, and having the distance between the peak and feature centroid be no greater than 20% of the ideal circle radius ($(\sqrt{A})\pi$).

3. **Guyots:** Same as a seamount but is defined by having more than 30% of the top 400 m of the feature sloping is less than 2°.

4. **Abyssal Plain:** An extensive area larger than 50,000 km^2 that overlies a flat or gently sloping area with a standard deviation of less than 50 m and is found at a depth between 4000 and 6000 meters.

5. **Shelf Break:** A line defined by the shoreward boundary of an area which deepens at a rate of at least 400 m per 30,000 m away from shore or has a localized slope of at least 1°. It may include adjacent areas which deepen at a rate of 350 m per 30,000 m or have a localized slope of at least 0.5°. It must span depths from 250 m to 1,000 m.

6. **Shelf:** The flat or gently sloping region adjacent to a continent or around an island that extends from the low water line to the shelf-break.

7. **Ridge:** A large feature that must have a topographical position index of greater than 1.5, indicating a substantial rise, and consist of an area greater than 30 km^2.

8. **Sea Channel:** A meandering linear depression with sinuosity typically greater than 1.3, at least no less than 1.15. Feature needs to be no shorter than 150 km, usually occurring on a gently sloping plain or FAN where local slope standard deviation is typically less than 0.5°, at least no more than 0.65°.

*These are not official in any means, and are purely based off the detection methods created.

Drawbacks of This Method

Once automation of all features (except ridge and sea channel) was complete, the detection algorithm was run on a world bathymetric dataset. When comparing the polygons discovered by the features listed in the GEBCO database, the results were as follows:

Feature	*Accuracy*
Abyssal Pain	23.2%
Basin	0%
Guyot	6.3%
Seamount	42.3%
Shelf	70.6%

In this case, accuracy is referring to the percentage of GEBCO features that fell within a generated polygon.

It is clear that this method needs refinement. One issue is that the parameters were created based off testing solely on the east and west coast of Canada. Undersea features vary greatly throughout the world, and therefore features off the Canadian coasts likely do not adequately represent this variety. Another issue is that the sheer processing power and time required to deal with the analyses of a raster are substantial. It requires a large amount of memory to run a single tool on a small sized raster, let alone a bundle of tools and processes on a large raster. With this, much of the tools that are being used, create specific values and require certain parameters that are not present outside of the GIS domain. This makes creating a good definition that clearly outlines the requirements for these features exceedingly difficult, as much of the parameters rely on specific statistical values that are likely unknown by the average observer. The drawbacks accompanied with this method are what led to the current development of a new method, which primarily focuses on finding and associating patterns in contours to specific features.

New Method – Contour Pattern Recognition

As mentioned, this new method uses the contours, rather than the raster cell values. This method cuts down on processing times; however there is still significant processing time, and allows for more leniency with regards to variation, as it is based on patterns rather than individual values. This method is still in progress, and the concept is largely theoretical, but so far the results, when tested on seamounts, are promising. The method attempts to group the features into four categories:

1. **Closed Loop Features**: Features defined by closed contour rings. In hydrology terms, these represent areas that are higher or lower than the surrounding terrain. There may be narrow gaps in the contour line.
2. **Trend Variation Features**: Features defined by localized variations in the trend of a contour line. These represent areas that likely are the result of erosion or sedimentation altering the typical contour line.
3. **Slope Variation Features:** Features defined by variation in the spacing between approximately parallel contour lines. These represent areas where the slope is significantly different from the surrounding area.
4. **Negative Features:** Features defined by narrow gaps between other features. These represent features such as passages that are breaks in other features.

So far, the only group that has had any substantial progress has been the Closed Loop Features, the accuracies of this group are as follows:

Feature	*Accuracy*
Seamount	96%
Knoll	64%
Guyots	98%
Banks	80%
Hills	64%

It is clear that this method is much better than the original, raster analysis method, when looking at the accuracies. The reason the smaller features (knolls and hills) had such a low accuracy is simply because the resolution of the data used was not sufficient enough to clearly identify these features, this can be remedied by using better resolution data. Progress is being made with regards to the other three groups, but it is becoming increasingly complex and difficult.

Conclusion

This update reviewed the methods created and implemented by the CHS, including the first method created, which analyzes a raster, and the new method currently being developed which uses contours. Ultimately, the work completed by the UFDP will assist the CHS and IHO in making better descriptions, and further increasing the understanding and accessibility to detecting and naming undersea features.

I would recommend CHS move forward with the new method being developed, and not consider the old method as official. The validation of the old method did not offer enough confidence to deem it accurate, and the methods used are not streamlined and can be refined. This is especially true when compared to the results seen from the new method. Furthermore, the parameters determined with the old method cannot effectively be scaled past the Canadian coasts.