Predictive Ensemble Maps for cold-water coral distributions in the Cap de Creus Canyon (NW Mediterranean). Advantages and limitations

Claudio Lo Iacono[1,2], Katleen Robert[1], Rita Gonzalez-Villanueva[3], Andrea Gori[2], Covadonga Orejas[4] and Josep Maria Gili[2]

[1]Marine Geoscience, National Oceanography Centre, Southampton, United Kingdom, cllo@noc.ac.uk

[2]Instituto de Ciencias del Mar, Consejo Superior de Investigaciones Científicas, Barcelona, Spain

[3]Departamento de Geociencias Marinas y O.T. Universidad de Vigo, Spain

[4]Instituto Español de Oceanografía, Centro Oceanográfico de Baleares, Palma de Mallorca, Spain

Predictive habitat mapping has shown great promise to improve the understanding of the spatial distribution of benthic habitats and is a valuable means to highlight species-environment relationships where field data may be limited. However, how much of what spatial distribution models tell us is true? Although they surely represent an important step forward in process-based ecosystem management, their predictive efficiency is not always tested by independent groundtruthing data. This is particularly true for the deep-sea environment, where sample data are always limited compared to the large extent of the areas to be mapped. Furthermore, the usual paucity of high resolution maps in deep-sea settings makes reliable predictive models even harder to obtain. The aim of this study is to apply and test different spatial models to statistically predict the distribution of three Cold-Water Coral (CWC) species (*Madrepora oculata*, *Lophelia pertusa* and *Dendrophyllia cornigera*) in the Cap de Creus Canyon (NW Mediterranean), based on high-resolution swath-bathymetry data (pixel resolution: 5m) and video observations from the submersible JAGO (IFM-GEOMAR). Presence/absence of CWCs was estimated in each 5m resolution pixel from video imagery. CWCs correspond to a habitat forming deep-sea community, which provide habitat and shelter for a wide range of species, including commercially viable fish. Moreover, submarine canyons act as specific hosting areas for these habitats, owing to their favourable environmental conditions.

Maximum Entropy (MaxEnt), General Additive Model (GAM) and decision tree model (Random Forest) were independently applied to represent non-linear species-environment relationships using terrain variables derived from multibeam bathymetry (slope, geomorphologic category, rugosity, aspect, backscatter). Relevant differences between the three models were observed. MaxEnt gave an outstanding performance for the three species, with the area under the curve (AUC) from the sensitivity-specificity approach between 0.98 and 0.99, and with slope, aspect and rugosity as the most relevant variables responsible for CWC distribution. On the contrary, the presence/absence models (GAM and Random Forest) gave a lower performance, with the considered physical elements having a minor ability to explain CWC distribution. These findings could be partly attributable to the effect of autocorrelation between available samples and also suggest that the addition of oceanographic variables (e.g. current speed & direction) would potentially improve the model performance. Nonetheless, the predicted areas where CWCs should be found with higher probabilities coincided for the three methods when a lower 50 m spatial scale was considered. According to the models, CWCs are most likely to be found on the medium to steeply sloping, rough walls of the southern flank of the canyon, aligning with the known CWC ecology acquired from previous studies in the area.

As a final step, a probabilistic predictive ensemble has been produced merging the outcomes of the three models considered, providing a more robust prediction for the three species. The main insight is that important discrepancies can arise in using different species distribution models, especially when high spatial resolutions are considered. This could in part be the result of the different statistical philosophies behind each of the models. Single models may not always be the most appropriate and definitive option, particularly when a limited number of observations is available. We suggest that a more reliable prediction could be obtained by merging models into spatial ensembles, able to reduce differences and associated uncertainties, showing hence a strong potential as an objective approach in the planning and management of natural resources.