Deep-Sea Research I 93 (2014) 72–82

Contents lists available at ScienceDirect

Deep-Sea Research I

journal homepage: www.elsevier.com/locate/dsri

Predicting the distribution of deep-sea vulnerable marine ecosystems using high-resolution data: Considerations and novel approaches

Anna M. Rengstorf [a,*], Christian Mohn [b], Colin Brown [c], Mary S. Wisz [d], Anthony J. Grehan [a]

[a] Earth and Ocean Sciences, School of Natural Science, NUI Galway, Galway, Ireland
[b] Department of Bioscience, Aarhus University, Roskilde, Denmark
[c] Ryan Institute for Environment, Marine and Energy Research, NUI Galway, Galway, Ireland
[d] Department of Ecology and Environment, DHI Water and Environment, Hørsholm, Denmark

ARTICLE INFO

Article history:
Received 18 March 2014
Received in revised form
7 July 2014
Accepted 16 July 2014
Available online 2 August 2014

Keywords:
Cold-water corals
Generalized linear model
Hydrodynamic modelling
Lophelia pertusa
Species distribution modelling

ABSTRACT

Little is known about species distribution patterns in deep-sea environments, primarily because sampling surveys in the high seas are expensive and time consuming. The increasing need to manage and protect vulnerable marine ecosystems, such as cold-water corals, has motivated the use of predictive modelling tools, which produce continuous maps of potential species or habitat distribution from limited point observations and full coverage environmental data. Rapid advances in acoustic remote sensing, oceanographic modelling and sampling technology now provide high quality datasets, facilitating model development with high spatial detail. This paper provides a short overview of existing methodologies for predicting deep-sea benthic species distribution, and illustrates emerging issues related to spatial and thematic data resolution, and the use of transect-derived species distribution data. In order to enhance the ecological relevance and reliability of deep-sea species distribution models, novel techniques are presented based on a case study predicting the distribution of the cold-water coral *Lophelia pertusa* in three carbonate mound provinces in Irish waters. Specifically, the study evaluates (1) the capacity of newly developed high-resolution (250 m grid cell size) hydrodynamic variables to explain local scale cold-water coral distribution patterns, (2) the potential value of species occurrence proportion data to maintain semi-quantitative information of coral prevalence (i.e. coverage) and sampling effort per grid cell within the response variable, and (3) mixed effect modelling to deal with spatially grouped transect data. The study shows that predictive models using vertical and horizontal flow parameters perform significantly better than models based on terrain parameters only. Semi-quantitative proportion data may decrease model uncertainty and increase model reliability, and provide a fruitful avenue of research for analysing large quantities of video data in a detailed yet time-efficient manner. The study concludes with an outlook of how species distribution models could improve our understanding of vulnerable marine ecosystem functioning and processes in the deep sea.

1. Introduction

Due to the great cost and time required for its exploration, the deep sea remains one of nature's great scientific frontiers. Merely 5% of the deep-sea floor has been explored remotely and less than 0.01% has been sampled (Ramirez-Llodra et al., 2010). Given the deep sea's value in providing important goods and services, such as commercially important fish, chemical compounds for pharmaceuticals, oil, gas and minerals, amongst others (Armstrong et al., 2012), and its vulnerability to increasing anthropogenic impacts (Davies et al., 2007), there is an urgent need to deepen our knowledge of deep-sea ecosystems, to map their spatial distribution, and to ensure their sustainable management and conservation. This has motivated the use of predictive modelling tools that relate species occurrence data with environmental predictor variables to estimate full-coverage species distribution in geographic space (Elith and Leathwick, 2009; Guisan and Zimmermann, 2000). Species distribution models (SDMs, also known as habitat suitability models, environmental niche models or resource selection functions) have been used for a variety of applications (Elith and Leathwick, 2009), including the prediction of the potential distribution of invasive species (Tyberghein et al., 2012), the estimation of the impact of a changing climate on future distributions (Tittensor et al., 2010), and conservation planning (Carroll, 2010). The effort to develop SDMs in the deep sea has mainly focused on cold-water corals (CWCs), driven by national and international obligations to develop conservation measures for these vulnerable marine ecosystems (VMEs, Rogers et al., 2008). Vierod et al. (2013) have recently reviewed the

---

[*] Corresponding author. Tel.: +44 759 777 5122.
E-mail address: amrengstorf@gmail.com (A.M. Rengstorf).

http://dx.doi.org/10.1016/j.dsr.2014.07.007
0967-0637/

application of SDMs to those deep-sea benthic species that typify VMEs. The authors draw attention to general modelling issues related to spatial scale, sampling bias (clustering of data in certain areas) and spatial pseudo-replication, amongst others, and particularly highlight the limitations of currently available deep-sea data. The present study addresses some of these data limitations by exploring the use of full coverage, high-resolution (~250 m) hydrodynamic variables derived from a newly developed oceanographic model, to explain and predict CWC distribution. It further proposes a novel modelling approach using high-quality, semi-quantitative occurrence proportion data derived from standardized video transects, ultimately aiming to enhance the ecological relevance and reliability of deep-sea SDMs.

Conventionally, SDMs are fit with information on both the presence and absence of a species (Guisan and Zimmermann, 2000). The difficulty to establish reliable absence data in deep-sea environments, however, has prompted most studies targeting deep-sea VMEs to make use of presence-only or presence-background modelling techniques (Vierod et al., 2013). Species presence data can originate from a variety of sources including fisheries bycatch, trawling and dredging surveys, boxcores and grab samples, video and photographic surveys. Such multi-source data vary significantly in quality, with potential issues including poor spatial precision, taxonomic errors and lack of metadata (e.g. information on the organism's life stage), amongst others. Quality control of readily available distribution data is therefore essential to ensure realistic representation of the target species (Ross et al., 2012; Rengstorf et al., 2013).

Important environmental drivers affecting the spatial distribution of deep-sea benthic species include primary productivity, near-seabed water chemistry and hydrodynamics, terrain morphology and seabed substrate (e.g. Davies and Guinotte, 2011; Howell et al., 2011; Rengstorf et al., 2013; Yesson et al., 2012). Full-coverage environmental data can be derived from ship- or satellite borne remote sensing (such as sea surface temperature and primary productivity), point-interpolations from in-situ measurements, ocean circulation models, and hybrids of these. A wide range of global scale data layers are readily available and have facilitated spatial predictions of cold-water sclaractinians and octocorals on a global scale (Tittensor et al., 2009; Davies and Guinotte, 2011; Yesson et al., 2012). While the resulting maps provide valuable biogeographic information and guidance for future survey efforts, their reliability and application to marine management is limited by the relative coarseness of the underlying environmental data (Vierod et al., 2013).

An important consideration when predicting the distribution of sessile species, such as CWCs, is that all habitat requirements must coincide at exactly the same spatial location (Guisan and Thuiller, 2005). Thus, high-resolution data are required to minimize spatial mismatch between distribution records and corresponding environmental conditions, and between the different environmental conditions. Davies et al. (2008), for example, showed that a 1° by 1° temperature grid was too coarse to resolve changes in water temperature, leading to a mismatch between CWC occurrences and temperature values beyond the species' thermal tolerance limit. Further spatial scale issues may arise when the environmental variables used in the predictions do not resolve, and thus do not reflect, the environmental factors or processes actually shaping the species distribution. Rengstorf et al. (2012) found that bathymetric data of 1 km resolution are too coarse to resolve the often small carbonate mounds that support CWC growth in Irish waters, and stressed the need for high-resolution bathymetry data to avoid over-estimation of the predicted coral habitat. Further, several authors (e.g. Frederiksen et al., 1992; White et al., 2005) have demonstrated, using in-situ current measurements, the preference of benthic suspension feeders for enhanced current flows. However, broad-scale SDMs predicting CWC distribution have failed to identify this relationship, a fact attributed to the coarse resolution of the hydrographic data (SODA, Carton et al., 2005), resulting in a lack of information on topographically influenced and locally varying factors (Davies and Guinotte 2011; Yesson et al., 2012). Ocean circulation models of high spatial resolution can simulate ecologically relevant processes, such as internal wave dynamics and resonant tidal amplification (Mohn et al., 2014). Integration of such hydrodynamic data is highly desirable, as it could add to the SDMs ecological relevance and provide valuable insight in critical processes such as larval dispersal, sediment settling rates, and – most importantly – food supply. The present study investigates the value of such high resolution hydrodynamic variables for predicting deep-sea benthic species distributions.

The spatial scale issues discussed above, coupled with the need for more detailed maps relevant for marine spatial planning, have motivated an increasing number of SDMs at finer spatial scales (Vierod et al., 2013). In contrast to global, explorative SDMs, these regional or local scale models are usually targeted towards specific areas of interest, where VMEs are expected to occur. Due to the lack of high-resolution, near-seabed oceanographic data, most 'fine scale' models to date are exclusively based on bathymetric data, which can be derived at high resolution from multi-beam echosounders (Brown et al., 2011). Bathymetry-derived parameters such as seabed slope and bathymetric position index (Weiss, 2001) act as proxies for seabed substratum (Dunn and Halpin, 2009) and current flow (Genin et al., 1986), and have been used to predict distribution patterns of a variety of benthic biota (e.g. Kostylev et al., 2001; Holmes et al., 2008), including CWCs (e.g. Dolan et al., 2008; Guinan et al., 2009a; Woodby et al., 2009). Video and photographic surveys have evolved to be the standard method for such regional habitat mapping and modelling efforts (Brown et al., 2011), as they are non-destructive, provide information on seabed substrata and biological assemblages, and allow for precise spatial matching with high-resolution environmental data (e.g. Dolan et al., 2008).

In a few fine-scale studies, standardized video transects could even provide (relatively reliable) coral absence data, enabling the use of conventional presence–absence modelling techniques, such as generalized linear models (GLM, Guisan and Zimmermann, 2000), to predict CWC distribution (Woodby et al., 2009; Marshall, 2012). Grid cells containing at least one occurrence observation are coded as "present" (1), whereas grid cells where the species was not observed are coded as "absent" (0), assuming that the limited area surveyed by the camera's footprint is representative of the entire area encompassed by the grid cell. Fig. 1 illustrates two main shortcomings of this simplification; it eliminates information on the species' prevalence (i.e. abundance or coverage) within the grid cell, and it eliminates information on how much of the grid cell has been actually surveyed (i.e. sampling effort). In this paper we explore the

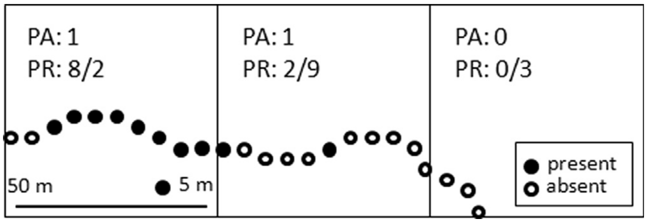

Fig. 1. Scheme showing observations of species presence (filled circle) and species absence (open circle) along a video transect in context with the environmental data grid cells (large squares). The corresponding response codes specified for a presence–absence model (PA) and a proportion model (PR) are given for each cell. The use of proportion data retains information on both prevalence (*n presences*/*n absences*) and sampling effort (*n presences*+*n absences*) within each grid cell.