



# Predicting sea pen (Pennatulacea) distribution on the UK continental shelf: evidence of range modification by benthic trawling

Anna-Leena Downie<sup>1,\*</sup>, Tamsyn Noble-James<sup>1</sup>, Ana Chaverra<sup>1,2</sup>, Kerry L. Howell<sup>2</sup>

<sup>1</sup>Centre for Environment, Fisheries and Aquaculture Science, Pakefield Rd, Lowestoft NR33 0HT, UK

<sup>2</sup>Marine Biology & Ecology Research Centre, Marine Institute, Plymouth University, Plymouth PL4 8AA, UK

**ABSTRACT:** Sea pen communities are United Nations General Assembly-designated Vulnerable Marine Ecosystems which occur worldwide in soft-bottom sediments where trawling often occurs. However, the ability of marine managers to assess, monitor and mitigate impacts to sea pens at national scales has been constrained by a limited understanding of their environmental requirements, geographical distribution and responses to trawling. In this study, we used random forest species distribution modelling (SDM) to predict the distribution of suitable habitat for 3 sea pen species (tall sea pen *Funiculina quadrangularis*, slender sea pen *Virgularia mirabilis* and phosphorescent sea pen *Pennatula phosphorea*) on the UK continental shelf, exploring the results relative to the distribution of fishing activity. Occurrence of all 3 species corresponded to areas of low current and wave velocity, where suspended matter in the water column was also low. However, for *F. quadrangularis*, the largest species, the models indicated substantially different drivers of distribution between the Greater North Sea and Celtic Seas ICES Ecoregions. This disparity appears to reflect modification to the range and realised niche of this species in the Greater North Sea, due to trawling impacts. *P. phosphorea* and *V. mirabilis* appear to be more resilient to trawling, with no clear negative relationships observed. Our findings illustrate the value of broadscale qualitative comparisons between SDMs and human activity data for insights on pressure–state relationships. When combined with robust distribution maps, this improved understanding of vulnerability will enable marine managers to make ecologically sound, defensible decisions and deliver tangible conservation outcomes for sea pen communities.

**KEY WORDS:** *Funiculina quadrangularis* · *Virgularia mirabilis* · *Pennatula phosphorea* · Species distribution modelling · Ecoregion · Pressure–state relationship · Sea pen · Habitat use

— Resale or republication not permitted without written consent of the publisher —

## 1. INTRODUCTION

Sea pens (Pennatulacea) are globally distributed colonial octocorals that are adapted to inhabit soft muddy or sandy sediments. Standing erect from the seabed, often in dense aggregations or ‘fields’, they provide structural complexity in otherwise featureless sediments, creating microhabitats, shelter and attachment substrata for sessile and motile fauna (Buhl-Mortensen & Buhl-Mortensen 2014, De Clippele et al. 2015). Such aggregations also function as

nursery grounds for commercially important fish (e.g. the redfish genus *Sebastes*; Baillon 2012) and serve as a food source for a range of invertebrates (García-Matucheski & Muniain 2011, Gale et al. 2013). Despite their wide distribution, high abundance and functional value, sea pens remain poorly studied in comparison to hard corals (Scleractinia).

As relatively slow-growing and long-lived organisms, sea pens are vulnerable to damage, displacement or removal by demersal fishing activities (Hixon & Tissot 2007, Malecha & Stone 2009, Lauria

\*Corresponding author: anna.downie@cefas.co.uk

et al. 2017). As such, their conservation status is acknowledged through various international legislative and policy instruments, with sea pen communities being recognised as Vulnerable Marine Ecosystems (VMEs) by the United Nations General Assembly (Rogers & Gianni 2010). Conservation measures for VMEs have advanced globally in recent decades (as summarised by Aguilar et al. 2017), with a variety of spatial management measures adopted at national and international scales, including the establishment of marine protected areas (MPAs). Although legislative provisions have improved, functional implementation of spatial measures has been slow, partly due to a poor understanding of environmental requirements and geographical distributions of VMEs (Aguilar et al. 2017). This is particularly relevant at the scale of individual nations, given the fundamental role of national frameworks in driving spatial conservation measures (Grip 2017). As sea pens are vulnerable to trawling impacts, a clear understanding of their environmental requirements and geographical ranges (both observed and theoretical) is essential to underpin spatial management of fishing activity within and beyond MPAs.

The UK continental shelf (UKCS) is a good example of a marine region that is extensively and chronically impacted by mobile demersal fishing, resulting in modification of biological assemblages and biogeochemical functions (Tillin et al. 2006, Thurstan et al. 2010, Kröger et al. 2018, Rijnsdorp et al. 2018). Sea pens in this region inhabit mud-rich, depositional habitats (Hill & Wilson 2000, Greathead et al. 2007, 2015), resulting in particular vulnerability to the commercially important *Nephrops norvegicus* fishery, which deploys otter trawls on muddy sediments, abrading the seabed surface and penetrating the top few centimetres (Ungfors et al. 2013, Eigaard et al. 2016). UK *Nephrops* landings have risen dramatically over recent decades in comparison to other European countries (Ungfors et al. 2013, Marine Management Organisation 2019), and the targeted muds are reported to be more significantly impacted by trawling, and to take longer to recover, than sand or gravel habitats (Hiddink et al. 2006, Rijnsdorp et al. 2016). Despite the potential threats to sea pen communities from demersal trawling, no studies have explored these impacts at the scale of the UKCS.

Sea pen communities on the UKCS comprise the tall sea pen *Funiculina quadrangularis*, slender sea pen *Virgularia mirabilis* and phosphorescent sea pen *Pennatula phosphorea*, the ecology and ranges of which have been explored in territorial Scottish

waters by Howson et al. (1994) and Greathead et al. (2007, 2015). Due to the relative paucity of research on the responses of these species to trawling abrasion, sensitivity assessments and evaluations of population viability in fished areas largely depend on physiological trait information and studies of analogous species (Hill & Wilson 2000, Ager 2003, Jones 2008), supported by a small number of local-scale trawling impact studies (Howson & Davies 1991, Tuck et al. 1998, Greathead et al. 2005, Murray et al. 2015). The available evidence suggests that the sensitivity of these species to trawling impacts may vary depending on physiology (i.e. flexibility and recovery potential) and retraction ability (Hill et al. 2020). *F. quadrangularis*, the largest and least often recorded of the 3 species (exceeding 200 cm in length) is considered the most sensitive to trawling impacts given its inability to retract its brittle axial rod beneath the sediment (Greathead et al. 2007) and the slow growth rate of analogous species (Wilson et al. 2002). The smaller and more often encountered *V. mirabilis* and *P. phosphorea* are able to retract into burrows and are generally considered less vulnerable than *F. quadrangularis* (Greathead et al. 2007, Ambroso et al. 2013). There is some evidence to support the lower sensitivity of these 2 species, although direct impact studies have thus far been confined to the north-western Scottish inshore (Tuck et al. 1998, Greathead et al. 2005).

Species distribution models (SDMs) are increasingly used in management of marine habitats, where determining actual species distribution can be logistically and financially challenging. SDMs provide geographically broad spatial predictions of environmental suitability for specified fauna, based on sample data and environmental data layers. The increasing accessibility of remotely sensed environmental data products now allows extrapolation from a limited amount of sample data to much larger areas of seabed (Tyberghein et al. 2012, He et al. 2015). A number of regional-scale studies have applied the SDM approach to predict sea pen distribution (Krigsman et al. 2012, Knudby et al. 2013, Kenchington et al. 2014, Beazley et al. 2016, Murillo et al. 2016, Lauria et al. 2017, Bastari et al. 2018, Kinlan et al. 2020), including a study of the Scottish West Coast, lochs and islands by Greathead et al. (2015). In addition to analysing species environmental requirements and extrapolating their distributions into unsampled areas, SDMs can also be used to shed light on external factors affecting natural distributions. The predicted distributions reflect the input data, and models using data from impacted populations will predict impacted distributions, allow-

ing comparison to natural distributions. These predicted distributions can then be evaluated relative to the known distribution of anthropogenic pressures, to support marine spatial planning and management (e.g. Marshall et al. 2014, Reiss et al. 2015). This linkage of species observations, environmental parameters and anthropogenic pressures enhances understanding of ‘pressure–state’ relationships, enabling defensible, evidence-based management decisions, for example exclusion of human activities from areas of high species vulnerability and establishment (and modification) of ecologically coherent, well-connected MPA networks (e.g. Smith et al. 2009, Sundblad et al. 2011). To this point, the majority of SDM studies on sea pens have focussed on spatial predictions of environmental suitability and have not linked predicted species distributions to anthropogenic impacts. A synergistic approach combining spatial predictions by SDMs and comparisons to anthropogenic activity data would improve the ability of marine managers to assess, mitigate and monitor impacts to sea pen communities. Thus far, to our knowledge, this connection has not yet been made at a national scale.

In this study, we present the first application of SDM to sea pens at the scale of the UKCS, using random forest classification models (Breiman 2001) to investigate environmental drivers of *F. quadrangularis*, *V. mirabilis* and *P. phosphorea* distribution and predict their occurrence across the UKCS. We use modelled environmental parameters to predict suitable habitat for the 3 species, training and validating the models using a large legacy trawl and imagery dataset. The predicted distributions are mapped and evaluated in the context of demersal fishing activity data. Finally, we use these results to explore whether anthropogenic pressures may have shaped the observed distribution of sea pens on the UKCS and discuss the implications for sea pen management and conservation.

## 2. METHODS

The study area (Fig. 1) covers most of the continental shelf inside the UK exclusive economic zone (EEZ). The spatial extent of the study was determined by the extent of the sediment composition layers (Mitchell et al. 2019a,b) used as predictor variables in the models. Consequently, the north-

ernmost tip of the shelf EEZ was excluded due to lack of environmental data coverage. The full study area is hereafter referred to as the UKCS. Based on preliminary results indicating differences in modelled sea pen distribution patterns and environmental responses between sea areas, the study area was further divided into 2 separate oceanographic ‘ecoregions’, the Celtic Seas (CELS) and the Greater North Sea (GRNS), as delineated by the International Council for the Exploration of the Sea (ICES 2020). All environmental data and species observations were aligned to a raster grid of the study area with a cell size of 0.002° (~200 m).

### 2.1. Species presence/absence data

Presence/absence sample data for *Funiculina quadrangularis*, *Virgularia mirabilis* and *Pennatula phosphorea* were collated from multiple legacy sources, comprising data collected using underwater imagery, dive surveys and benthic trawls and grabs over a period spanning 1961–2019. The majority of observations were derived from the publicly available UK Marine Recorder database (Public snapshot v20170825, available from <https://jncc.gov.uk/our-work/marine-recorder>, downloaded on 01/11/2017). These were supplemented by additional underwater imagery and grab samples collected as part of the UK Marine Protected Areas Programme, as well as Cefas benthic trawl surveys and grab sampling conducted under various research and monitoring programmes.

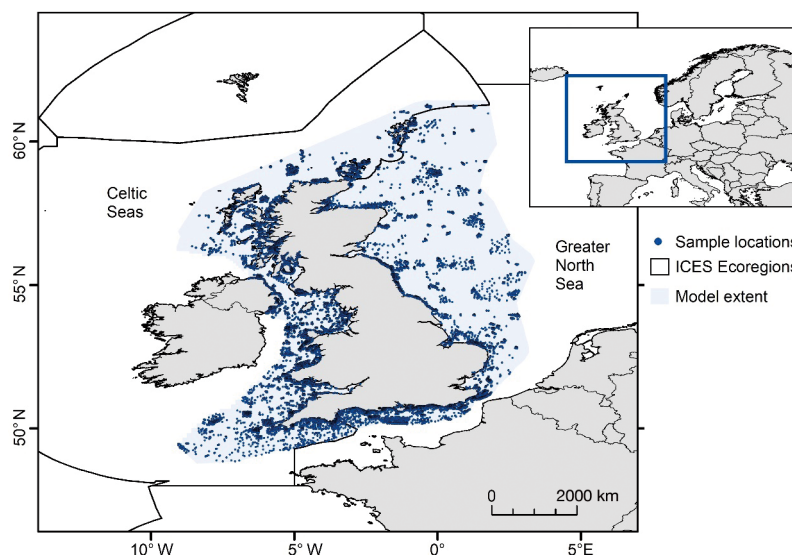


Fig. 1. Study area, extent of distribution models, boundaries of the ICES Ecoregions used to delineate the 2 target regions and the distribution of sea pen presence/absence point observations

Presence/absence observations for each of the 3 sea pen species were reduced to 1 observation per 0.002° (~200 m) raster cell. For trawls and video tows, where coordinates were recorded at the beginning and end, the positional midpoint was used as the observation location. The precision of positioning for the observations differed based on both their vintage and sampling method, but this was mitigated by the size of the raster cells and consequently the spatial scale of prediction. Each raster cell with at least 1 intersecting presence observation was classified as a presence. Absence observations from grab samples were excluded, as they are considered to under-sample large, sparsely distributed epifauna such as sea pens. To limit the effects of spatial autocorrelation and redundancy, the data in raster cells were further spatially subsampled to exclude neighbouring points closer than 250 m. Datasets with very low prevalence lack the ability to adequately represent the nature of species dependence on environmental conditions (Santika 2011). Hence, a final subsampling step was conducted individually for each species and region to randomly down-sample absence records, achieving a 10% prevalence for each dataset used in models. Table 1 shows the total number of raster cells with a presence/absence record and the number of cells included for modelling in each region.

## 2.2. Environmental predictor layers

The environmental predictor layers included in model development, their units, sources and spatial resolution of the source data are listed in Table 2. All layers were resampled to the native resolution of the bathymetry layer (0.002°, or ~200 m). Where necessary, values were interpolated using the empirical Bayesian kriging function in ArcGIS10.5 Geostatistical Analyst (with default settings).

SAGA GIS tools for QGIS (v. 3.2; Conrad et al. 2015) were used to calculate a set of regional terrain variables that, whilst terrestrial in origin, can be used to represent aspects of seabed topography. The regional terrain variables calculated include channel network base level, channel network distance, valley depth, closed depressions, relative slope position and standardised height (Shaw et al. 2004). The concept of the channel network base level is used to distinguish topographic highs and lows, by using digital elevation modelling (DEM) to create a channel network. The channel network base level is an interpolated elevation surface connecting the channel elevations. The channel network distance is calculated as the vertical distance between the DEM elevation and the channel network base level elevation. Valley depth is calculated as the vertical distance to the lowest elevation of source flow. Closed depressions represent sinks in the topography, surrounded on all sides by higher ground, with values increasing with the height difference to the surrounding higher ground. The relative slope position ranges from 0 to 1 as an index of location along the entire length of a slope. Standardised height is the product of normalised height (value ranging from 0 to 1 from the lowest to the highest position within a respective reference) multiplied by absolute height, relating a location to its wider surrounding topography (Shaw et al. 2004).

Layers used for current and wave velocity at the seafloor, winter suspended particulate matter, sand, mud and gravel fractions, as well as sand to gravel log ratio and mud to gravel log ratio are those produced by Mitchell et al. (2019c), available for download through the Cefas Data Hub (Mitchell et al. 2019a,b; Table 2). Mean, maximum and minimum bottom temperatures, along with the mean and maximum annual temperature ranges for 2017–2019, were calculated from daily mean seabed temperatures (°C), obtained from a high-resolution North-

Table 1. Summary of *Funiculina quadrangularis*, *Pennatula phosphorea* and *Virgularia mirabilis* presence/absence records used in the species distribution models (SDMs). Each record corresponds to a raster cell with at least 1 point observation. The total number of records in the full dataset is given for the whole study extent (UK continental shelf: UKCS). Number of records in the final datasets subsampled to 10% prevalence are given for each model region (GRNS: Greater North Sea; CELS: Celtic Seas)

Species	Total number of records (presences/absences)			
	All records UKCS	UKCS	Subsampled records GRNS	CELS
<i>Funiculina quadrangularis</i>	30103 (379/29724)	3861 (351/3510)	627 (57/570)	3234 (294/2940)
<i>Pennatula phosphorea</i>	30541 (1677/28864)	11132 (1012/10120)	7225 (656/6569)	3267 (297/2970)
<i>Virgularia mirabilis</i>	30301 (1383/28918)	9790 (890/8900)	4884 (444/4440)	4906 (446/4460)

Table 2. Environmental layers included in model development. Variables selected for the final models are indicated by: Y; na: not applicable

Variable	Unit	Selected	Source resolution	Source
Bathymetry	m		0.002°	EMODnet Digital Bathymetry (EMODnet 2016)
Valley depth	m	Y		
Relative slope position	0–1			
Distance from channel network	m	Y	0.002°	Calculated from Bathymetry with SAGA for QGIS – Basic terrain analysis tools
Standardised height	m	Y		
Channel network baseline	m			
Closed depressions	Unitless	Y		
Current velocity	m s <sup>-1</sup>	Y	0.002°	Cefas Data Hub, Mitchell et al. (2019a) <a href="https://doi.org/10.14466/CefasDataHub.62">https://doi.org/10.14466/CefasDataHub.62</a>
Wave velocity	m s <sup>-1</sup>	Y		
Winter suspended particulate matter	g m <sup>-3</sup>	Y		
Sand fraction	%	Y	0.002°	Cefas Data Hub, Mitchell et al. (2019b) <a href="https://doi.org/10.14466/CefasDataHub.63">https://doi.org/10.14466/CefasDataHub.63</a>
Mud fraction	%	Y		
Gravel fraction	%			
Sand to gravel log ratio	Ratio			
Mud to gravel log ratio	Ratio	Y		
ICES Statistical Areas	na		Vector	ICES Data Portal (ICES 2005)
ICES Ecoregions	na	Y	Vector	ICES Data Portal (ICES 2015)
Mean spring chl <i>a</i> concentration (2017–2019)	mg m <sup>-3</sup>	Y	1 km	OCEANCOLOUR_ATL_CHL_L3_NRT_OBSERVATIONS_009_036
Mean summer chl <i>a</i> concentration (2017–2019)	mg m <sup>-3</sup>		1 km	from <a href="http://marine.copernicus.eu/">http://marine.copernicus.eu/</a>
Mean autumn chl <i>a</i> concentration (2017–2019)	mg m <sup>-3</sup>		1 km	
Mean spring primary production (2017–2019)	C mg m <sup>-3</sup> d <sup>-1</sup>		7 km	NORTHWESTSHELF_REANALYSIS_IS_BIO_004_011 from
Mean summer primary production (2017–2019)	C mg m <sup>-3</sup> d <sup>-1</sup>	Y	7 km	<a href="http://marine.copernicus.eu/">http://marine.copernicus.eu/</a>
Mean autumn primary production (2017–2019)	C mg m <sup>-3</sup> d <sup>-1</sup>		7 km	
Annual range in bottom temperature (2017–2019)	°C		1.5 km	NORTHWESTSHELF_ANALYSIS_FORECAST_PHY_004_013 from
Maximum bottom temperature (2017–2019)	°C		1.5 km	<a href="http://marine.copernicus.eu/">http://marine.copernicus.eu/</a>
Mean bottom temperature (2017–2019)	°C	Y	1.5 km	
Minimum bottom temperature (2017–2019)	°C	Y	1.5 km	

West European Shelf forecasting ocean assimilation model, downloaded from <http://marine.copernicus.eu/> (Table 2). Seasonal averages for spring (March, April, May), summer (June, July, August) and autumn (September, October, November) chlorophyll *a* concentrations (mg m<sup>-3</sup>) and net primary productivity of carbon (mg m<sup>-3</sup> d<sup>-1</sup>) between 2017 and 2019 were computed from monthly composites of ESA Ocean Colour CCI Remote Sensing Reflectance data and the UK Met Office Operational Suite Atlantic Margin Model FOAM output downloaded from <http://marine.copernicus.eu/> (Table 2). Vector layers of ICES Ecoregions (ICES 2015) and ICES Statistical Areas (ICES 2005) were rasterised and included as factor variables acting as proxies for biogeographical attributes and oceanographic conditions in the various sea areas surrounding the UK.

### 2.3. Model training and validation

Originating from the field of machine learning, the random forest method is increasingly used by ecologists

due to its high classification accuracy and ability to characterize complex interactions between variables (Prasad et al. 2006, Cutler et al. 2007). Three random forest distribution models (Breiman 2001) were built for each sea pen species: (1) a full model trained using data from the entire study area (hereafter referred to as the UKCS model), (2) a model trained using only data from the GRNS (hereafter referred to as the GRNS model) and (3) a model trained using only data from the CELS (the CELS model). Each model was used to predict the distribution over the entire study area (UKCS). Transferability of models between the 2 ecoregions was tested by calculating model performance statistics on test data separately for each. In the case of the UKCS model, accuracy was tested for the full UKCS test dataset and for each ecoregion.

All analyses were run using the R statistical computing software (v. 3.5.1, R Core Team 2018). Pre-selection of predictor variables was conducted using the 'Boruta' algorithm in the 'Boruta' R package (Kursa & Rudnicki 2010), to reduce the number of variables included in the model for ease of interpretation.



tation and to allow comparison between species. The algorithm compares the importance of a variable calculated by a random forest model to the importance of a random permutation of the same variable over several iterations. Variables with higher than random importance for at least 1 species were considered for inclusion in the final predictor variable set. Of the correlated predictor variables (correlation coefficient  $>0.6$ ), the variable with the highest average importance across all 3 species (calculated by 'Boruta') was selected to be included in the final set of variables. The UKCS model also included ecoregion as a factor variable. The final set of predictor variables included in models is indicated in Table 2.

As our dataset was compiled from multiple sources, consideration was given to retaining 1 or more of the datasets for independent validation. However, combinations of all of the component datasets were required to achieve full geographical coverage of the study area for both training and testing the models. Consequently, a multi-run cross-validation on the full dataset was determined to be the more appropriate validation strategy, affording the most comprehensive training and test datasets, together with an approximation of model stability over multiple subsets of data. The 'biomod2' package (v. 3.4.6, Thuiller et al. 2009, 2020) was used to train and test 10 cross-validation (CV) runs of each model to control for any artefacts resulting from a single split of the data into training and test data sets by random selection. The 10 training and test data splits were created by assigning two 5-fold cross-CV sets with equal numbers of presences included in each for balance, resulting in an 80/20 split into training/testing data.

Random forest models were built using 500 trees and 5 variables randomly sampled as candidates at each split. Variable importance statistics and partial response curves were extracted from each CV run. Predictions from each CV run were dichotomized into presence/absence using a threshold that optimizes the true skill statistic (TSS; Allouche et al. 2006). TSS was selected based on its insensitivity to prevalence and its equal weighting of sensitivity and specificity, avoiding pitfalls of both the Kappa statistic and the area under the receiver operating characteristic curve highlighted in numerous studies (including Manel et al. 2001, Lobo et al. 2008, Jiménez-Valverde 2012). The final predicted distribution layer from each model combined the dichotomized outputs from all 10 CV runs, classified into 3 categories: (1) absence, (2) low likelihood of presence (1–5 CV runs predict presence) and (3)

high likelihood of presence ( $>5$  CV runs predict presence).

The accuracy of predicted presences/absences was examined using the 'PresenceAbsence' package in R (Freeman & Moisen 2008). Specificity, sensitivity (Fielding & Bell 1997), the Kappa statistic and TSS were calculated for the corresponding test data in each CV run. For the UKCS model, separate statistics were calculated by filtering the test data by each ecoregion in turn. For testing the transferability of the model trained on data from one ecoregion to the other, the test data from the corresponding CV run from the other ecoregion were used.

#### 2.4. Overlap with demersal fisheries

The potential contribution of bottom contact fishing impacts to the observed difference in environmental preferences of *F. quadrangularis* between the 2 ecoregions was investigated further. Spatial fishing intensity data were overlain with presence as predicted by the CELS model, and the ratio of observed absences in impacted to unimpacted sample locations was compared in areas where presence was predicted.

The spatial distribution of potential impacts from fishing activity was determined by combining information on bottom contact fishing from 2 sources. The spatial distribution of offshore physical disturbance by mobile bottom contacting fishing gears was derived from gridded Vessel Monitoring Systems (VMS) data (ICES 2019), acquired from vessels  $\geq 12$  m in length. The gridded VMS layer was created by calculating the mean value of the average swept area ratios (SARs) for 2009–2017 from annual 5 km resolution 'OSPAR Bottom Fishing Intensity – Surface' data layers, downloaded from the OSPAR Data and Information System ([https://odims.ospar.org/odims\\_data\\_files/](https://odims.ospar.org/odims_data_files/)). For inshore waters, an additional layer was obtained, representing the number of small ( $<15$  m) fishing vessels using bottom contacting gears operating in coastal areas (Kafas et al. 2014). This layer was produced by adding the number of vessels for scallop dredgers, *Nephrops* trawlers and other bottom trawlers, as reported by the ScotMap Inshore Fisheries Mapping Project (2007–2011; downloaded from the Marine Scotland Data portal, <https://dx.doi.org/10.7489/1554-1>).

The fishing intensity layers ('mean SAR'/no. vessels') were intersected with sample points for *F. quadrangularis*. Histograms of presence observations across increasing fishing intensity (both layers)

were used to determine cut-off values for fishing intensity where *F. quadrangularis* was less numerous, inferring a negative impact on this species. The cut-off values were applied to each fishing intensity layer, and both were combined into 1 layer delineating the areas where fishing is likely to impact *F. quadrangularis*.

The number of absence observations from areas with a predicted high likelihood of presence (>5 CV runs predict presence), low likelihood of presence (1–5 CV runs predict a presence) and absence were examined to compare the fraction of absences in 'impacted' vs. 'unimpacted' areas in both ecoregions.

### 3. RESULTS

#### 3.1. Model performance and transferability

The UKCS models for all 3 sea pen species showed good performance (Table 3). *Funiculina quadrangularis* was modelled with the highest overall accuracy (TSS = 0.90), followed by *Pennatula phosphorea*

(TSS = 0.85), then *Virgularia mirabilis* (TSS = 0.77). Inclusion of ecoregion as a predictor variable in the UKCS models resulted in accurate predictions in both the CELS and GRNS areas. For *P. phosphorea* and *V. mirabilis*, the UKCS model showed fractionally higher sensitivity and overall performance in the GRNS than the CELS (Table 3). *F. quadrangularis* models trained on data from one ecoregion did not transfer well to the other. The transferred models severely underpredicted presence, with very low sensitivity (0.04 and 0.12, GRNS to CELS and CELS to GRNS, respectively) and consequently poor overall performance in the opposite region (GRNS to CELS, TSS = 0.02; CELS to GRNS, TSS = 0.04). For *P. phosphorea* and *V. mirabilis*, transferability was much better from the CELS model to the GRNS region than from the GRNS model to the CELS, with an equal or better ability to predict presences and only a marginally poorer specificity (Table 3). The poorer sensitivity and specificity in the *P. phosphorea* and *V. mirabilis* GRNS models transferred to the CELS indicated that presence was underpredicted in some areas of this ecoregion and overpredicted in

Table 3. Model validation results ( $\pm$ SD). Results for full models are shown for validation datasets covering the entire UK continental shelf (UKCS) study area, as well as separately for the Greater North Sea (GRNS) and Celtic Seas (CELS) ecoregions. Results for the regional models are shown for both ecoregions. N: number of samples included in each training data set; P: number of presences in each training data set; TSS: true skill statistic

Model	N	P	Test data	Sensitivity	Specificity	Kappa	TSS
<i><b>Funiculina quadrangularis</b></i>							
UKCS	3089	281	UKCS	0.98 (±0.02)	0.92 (±0.03)	0.68 (±0.09)	0.9 (±0.03)
			GRNS	0.91 (±0.09)	0.93 (±0.04)	0.53 (±0.14)	0.84 (±0.08)
			CELS	0.98 (±0.02)	0.92 (±0.03)	0.72 (±0.08)	0.9 (±0.03)
GRNS	502	46	GRNS	0.98 (±0.04)	0.93 (±0.03)	0.7 (±0.09)	0.91 (±0.05)
			CELS	0.04 (±0.08)	0.98 (±0.02)	0.03 (±0.07)	0.02 (±0.06)
CELS	2587	235	GRNS	0.12 (±0.14)	0.92 (±0.04)	0.03 (±0.12)	0.04 (±0.12)
			CELS	0.97 (±0.01)	0.93 (±0.02)	0.7 (±0.07)	0.9 (±0.02)
<i><b>Pennatula phosphorea</b></i>							
UKCS	8906	810	UKCS	0.96 (±0.02)	0.9 (±0.01)	0.6 (±0.03)	0.85 (±0.01)
			GRNS	0.98 (±0.01)	0.88 (±0.01)	0.7 (±0.03)	0.87 (±0.02)
			CELS	0.88 (±0.05)	0.9 (±0.02)	0.42 (±0.03)	0.78 (±0.04)
GRNS	5780	525	GRNS	0.96 (±0.02)	0.92 (±0.02)	0.67 (±0.07)	0.88 (±0.01)
			CELS	0.59 (±0.14)	0.85 (±0.07)	0.3 (±0.07)	0.44 (±0.1)
CELS	2614	238	GRNS	0.96 (±0.03)	0.83 (±0.03)	0.46 (±0.05)	0.79 (±0.03)
			CELS	0.93 (±0.03)	0.87 (±0.03)	0.52 (±0.06)	0.8 (±0.04)
<i><b>Virgularia mirabilis</b></i>							
UKCS	7832	712	UKCS	0.92 (±0.03)	0.85 (±0.03)	0.48 (±0.05)	0.77 (±0.03)
			GRNS	0.95 (±0.03)	0.89 (±0.03)	0.63 (±0.08)	0.84 (±0.04)
			CELS	0.89 (±0.04)	0.83 (±0.03)	0.38 (±0.04)	0.71 (±0.04)
GRNS	3907	355	GRNS	0.91 (±0.02)	0.95 (±0.02)	0.73 (±0.06)	0.86 (±0.03)
			CELS	0.44 (±0.12)	0.85 (±0.07)	0.22 (±0.05)	0.3 (±0.07)
CELS	3925	357	GRNS	0.94 (±0.04)	0.71 (±0.07)	0.3 (±0.07)	0.66 (±0.06)
			CELS	0.9 (±0.04)	0.82 (±0.04)	0.41 (±0.05)	0.72 (±0.02)

others, suggesting poor specification of the species niche by the models (Table 3).

### 3.2. Factors driving sea pen distributions

*F. quadrangularis*, *V. mirabilis* and *P. phosphorea* all display broadly similar environmental requirements when modelled and tested using the full dataset (including both ICES ecoregions). The most important predictor variables for all 3 species include mean bottom temperature, wave and current velocities, and the concentration of winter suspended particulate matter in the water column. Primary productivity is also a notable contributor to the *P. phosphorea* and *V. mirabilis* models. The topographic and sediment variables contribute to a lesser degree, but all increase model performance when included. However, the order of importance differs between the ecoregions. The GRNS models are all strongly driven by temperature, whilst the CELS models are more influenced by suspended matter and current

and/or wave velocities. Sand and mud content are also more influential in the CELS models than the GRNS models (Fig. 2).

All species in both ecoregions are more likely to occur in low current and wave velocities, where concentration of suspended matter in the water column is low. *F. quadrangularis* and *P. phosphorea* show a preference for current and wave velocities up to 0.3 and 0.5 m s<sup>-1</sup>, respectively, and suspended matter at less than 2 g m<sup>-3</sup>, with 95 % of all presences occurring in these conditions. *V. mirabilis* shows a somewhat higher tolerance of wave conditions and turbidity, up to 1.1 m s<sup>-1</sup> for wave velocity and 5.5 g m<sup>-3</sup> for suspended matter. The relationship with primary productivity is more complicated. *P. phosphorea* is more likely to occur in areas with low primary productivity, whereas *V. mirabilis* shows a positive association in the CELS model but a negative one in the GRNS model. Responses to bottom temperature and sediment, and in the case of *F. quadrangularis* to bottom topography, differ between the ecoregions. Temperature ranges of all species are similar (7.8–12.9°C)

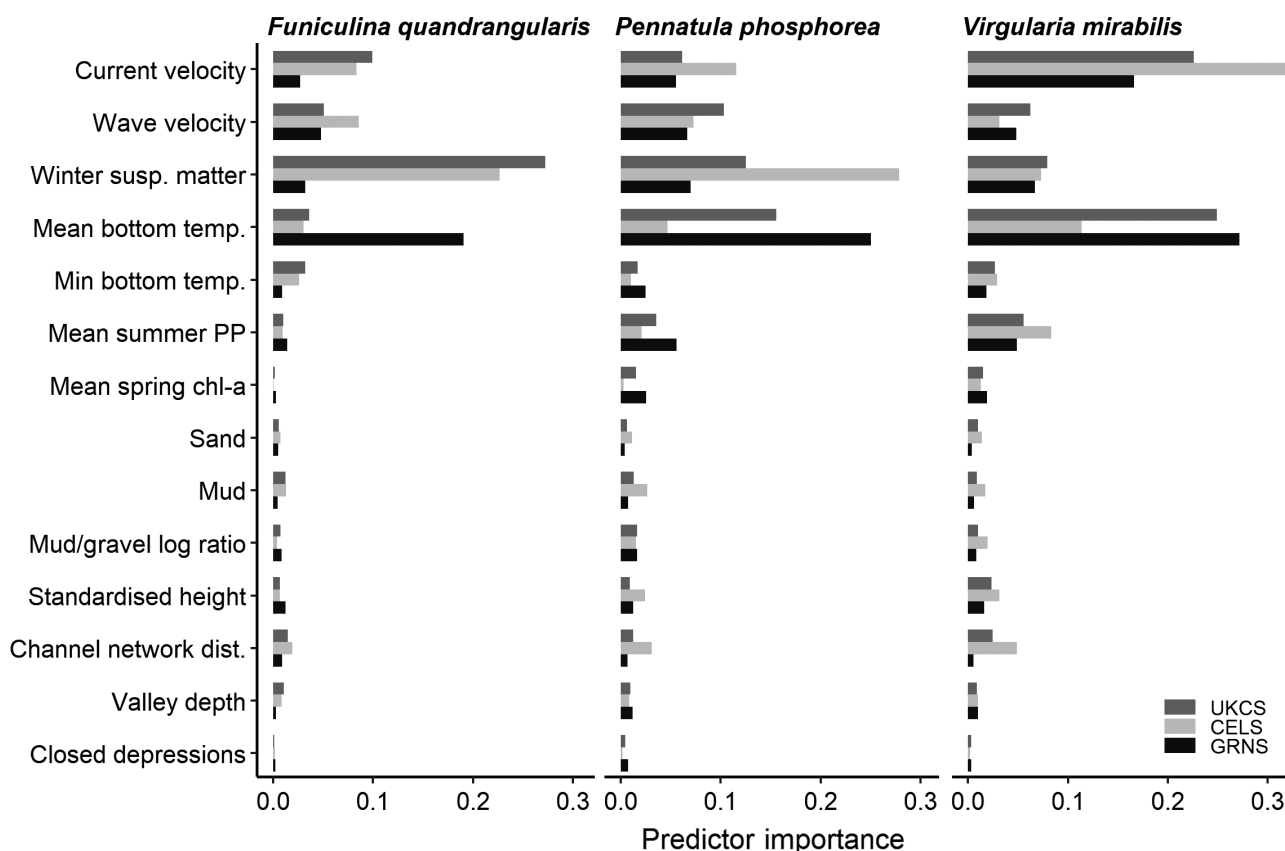


Fig. 2. Predictor variable contributions to models by species and regional model. PP: Primary productivity; UKCS: UK continental shelf; CELS: Celtic Seas; GRNS: Greater North Sea. Variable contribution is inferred from a reduction in model accuracy when each predictor variable in turn is randomised. Predictor importance = 1 minus the correlation coefficient between model accuracies for the full model and the model with the randomised variable



but extend into much warmer mean temperatures in the CELS than in the GRNS. The clearest difference is seen in *F. quadrangularis*, which in the North Sea is most likely to occur between 8 and 9°C, whilst in the CELS, this extends to ~10.5°C (Fig. 3).

*P. phosphorea* and *V. mirabilis* consistently occur in areas with sheltered concave topography with increased 'enclosedness', represented by higher values of valley depth and closed depressions along with low standardised height. *F. quadrangularis* follows these trends in the CELS, but in the GRNS, it occurs almost exclusively offshore, and on flat or slightly elevated ground (Fig. 3).

Although the effect of mud and sand content is much lower in the GRNS models, the presence probability for all species, and in both ecoregions, is increased by higher mud content. The CELS models show higher probability of presence for all species with lower sand content, whilst in the GRNS models, a slight increase is seen with high sand content. The mud to gravel ratio also shows an opposite trend of influence between the ecoregions. Whilst presences are more likely with a high mud to gravel ratio in the GRNS, the opposite is true of CELS. The differences are most pronounced for *F. quadrangularis* (Fig. 3).

The predicted distributions of *F. quadrangularis* across the UKCS by the CELS and GRNS models are

almost entirely different (Fig. 4). The GRNS model does not predict the presences observed in the warmer inshore bottom temperatures of the CELS, instead predicting presences further offshore in cooler areas with higher sand content. The CELS model, on the other hand, predicts into muddy basins in a wide temperature range, but fails to predict most of the presences in the North Sea that occur in patches of sandier sediments.

### 3.3. Predicted distributions of sea pens on the UKCS

The UKCS models indicate a largely overlapping distribution, with the main habitat for each species occurring in North Atlantic waters off the coast of Scotland and in the northern half of the North Sea. In terms of spatial extent, the area of predicted distribution is greatest for *P. phosphorea*, followed by *V. mirabilis* and *F. quadrangularis*. *F. quadrangularis* has the most restricted distribution, being confined to the northern waters and, in the south, the shelf edge where canyon heads cut into the shelf break. A small suitable patch is also predicted in the western English Channel. *P. phosphorea* is also predicted to occur in pockets of suitable habitat in the western English Channel and southern Celtic Sea towards the shelf edge. *V. mirabilis* habitat has the widest geographical distribution, with predicted patches in all sea areas excluding the eastern English Channel and southern North Sea (Fig. 5).

### 3.4. Role of bottom contact fishing impacts in the observed distribution of *F. quadrangularis*

Comparison of the CELS and GRNS modelled presence distributions and the fishing impact layers revealed spatial trends which related to the ecoregion model disparities observed for *F. quadrangularis*. *F. quadrangularis* is notably absent from the area predicted as suitable habitat by the CELS model (Fig. 6a). The CELS model, therefore, fails to accurately describe the observed distribution of *F. quadrangularis* in the GRNS, despite encompassing the range of environmental

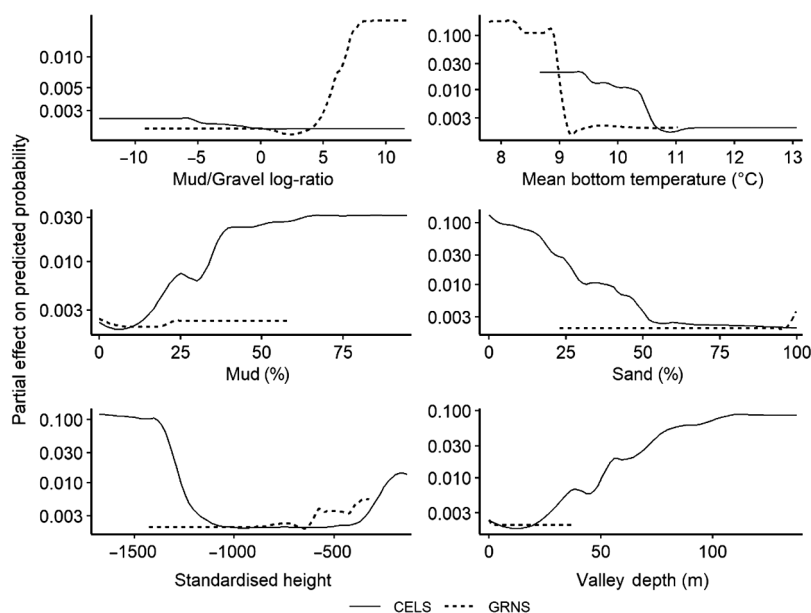


Fig. 3. Random forest model partial response curves illustrating the different responses of *Funiculina quadrangularis* presence/absence to substrate and topography variables observed between the Celtic Seas (CELS) and Greater North Sea (GRNS) ecoregions. The plots show the relative logit contribution of each variable to the probability of presence. The y-axis scale has been log-transformed to allow plotting of curves with a large range of values from both models onto the same axis

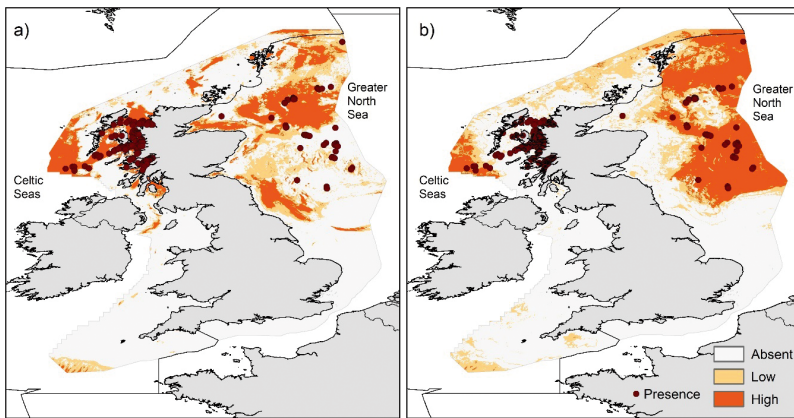


Fig. 4. Distribution of *Funiculina quadrangularis* predicted by the (a) Celtic Seas (CELS) and (b) Greater North Sea (GRNS) models, with observed presences. Confidence in the predicted distribution is represented by 2 categories Low (high) likelihood of presence is shown where  $<5$  ( $\geq 5$ ) of 10 cross-validation model iterations predicted presence

conditions existing within it. A large proportion (71 %) of the muddier seabed in the GRNS, predicted as suitable *F. quadrangularis* habitat by the CELS model, is impacted by benthic trawling by vessels  $\geq 12$  m in length, in comparison with only 39 % in the CELS (Table 4). An additional 5 % of the GRNS and 15 % of the CELS ecoregions are covered by grid cells where inshore fisheries vessels ( $<15$  m in length) are known to operate. Where presences are observed in areas predicted as suitable by the CELS model, they occur outside of the mapped extent of fishing impact (Fig. 6a(i)). *F. quadrangularis* are also observed on sandier ground adjacent to the edges of the mud basins predicted as suitable habitat by the CELS model. This effect is exemplified in Fig. 6a(ii), which shows the presence of *F. quadrangularis* in patches of sandier and mixed sediments in one of the large mud

basins in the Fladen Grounds. In contrast, *P. phosphorea*, which has a very similar predicted distribution, does occur widely across the areas predicted as suitable for *F. quadrangularis*, where *F. quadrangularis* is absent (Fig. 6b). An example of *F. quadrangularis* on the sandy mixed ground patches in Fig. 6a(ii) is shown in Fig. 7.

Fig. 8 shows that presences of *F. quadrangularis* across the whole study area are highly skewed towards areas of low fishing impact. Presences are overwhelmingly observed where no fishing occurs according to SAR from VMS data (Fig. 8a). Similarly, in the inshore region, the majority of presences are observed in areas where fewer than 4 inshore vessels ( $<15$  m) are known to regularly fish using bottom contacting gear (Fig. 8b). The majority of absences of *F. quadrangularis* observed where the CELS model indicates a high likelihood of presence are from areas impacted by bottom contact fisheries (70 and 60 % in the North Sea and CELS, although these values are likely to be inflated to some degree due to a greater area of predicted suitable habitat being trawled in the North Sea than the CELS). The comparable values where the model predicts absence are only 24 and 20 %, respectively (Fig. 8c).

#### 4. DISCUSSION

The random forest models were highly effective in predicting occurrences of *Funiculina quadrangularis*, *Virgularia mirabilis* and *Pennatula phosphorea* on the

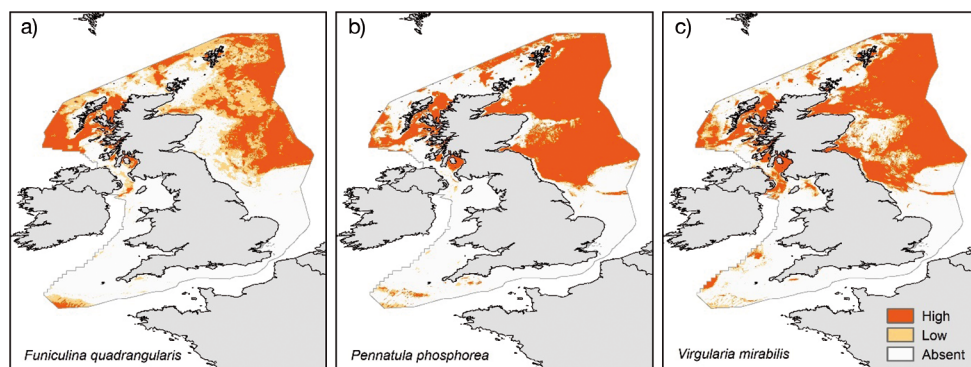


Fig. 5. Distributions of (a) *Funiculina quadrangularis*, (b) *Pennatula phosphorea* and (c) *Virgularia mirabilis* predicted by the full UK continental shelf (UKCS) model. Confidence in the predicted distribution is represented by 2 categories. Low (high) likelihood of presence is shown where  $<5$  ( $\geq 5$ ) of 10 cross-validation model iterations predicted presence

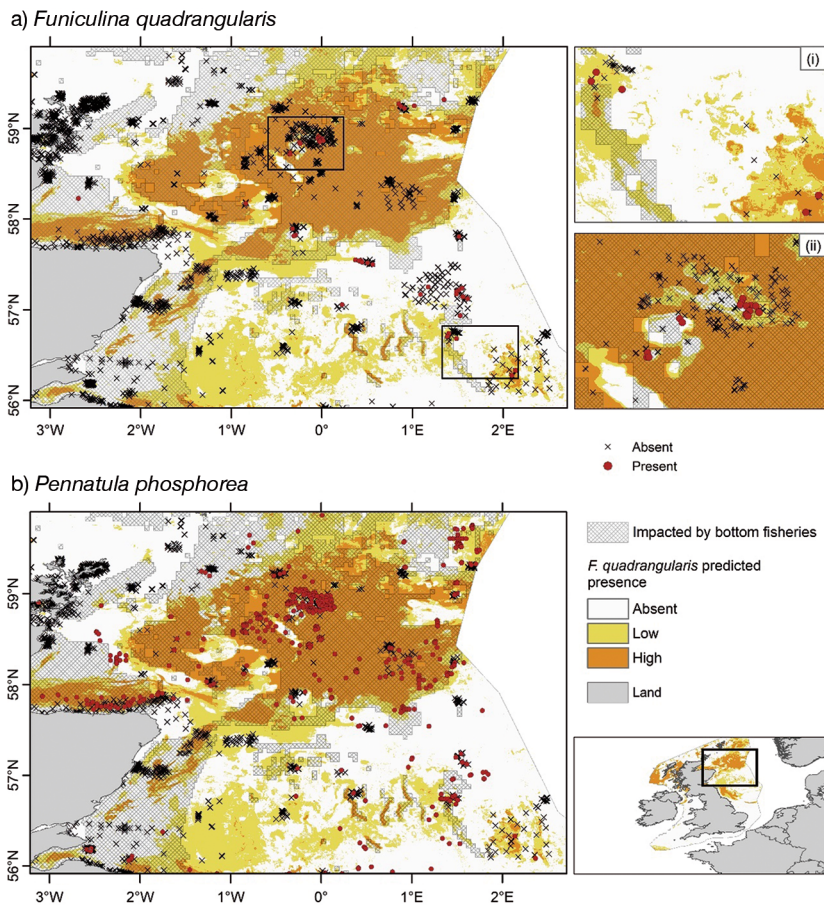


Fig. 6. Distribution of *Funiculina quadrangularis* predicted by the Celtic Seas (CELS) model, overlain with the extent of bottom contact fisheries (swept area ratio >0 or >4 vessels [see Table 4] reported fishing using bottom contacting gears) and the observed presence/absence of (a) *F. quadrangularis* and (b) *Pennatula phosphorea*. Insets in panel (a) show close-up details of *F. quadrangularis* presence in (i) areas predicted by the CELS model in the absence of fishing pressure and (ii) sandier patches outside the predicted habitat

UKCS. *F. quadrangularis* was predicted with the highest accuracy, followed by *P. phosphorea* and *V. mirabilis*. Preliminary investigations, however, indicated disparities in the responses of sea pens to environmental parameters between the CELS and GRNS ecoregions, particularly for *F. quadrangularis* (the species most sensitive to physical disturbance). This was corroborated by models trained on data from only 1 ecoregion, which showed different partial responses to environmental variables in each ecoregion. A formal assessment testing the accuracy of a model trained on data from one ecoregion using the inde-

Table 4. Fraction of the area predicted as suitable habitat by the Celtic Seas (CELS) model (presence in >5 cross-validation iterations) overlapping fished areas. Area fished by offshore vessels comprises the area covered by OSPAR gridded Vessel Monitoring System data where swept area ratio (SAR) > 0 (OSPAR Data and Information System, [https://odims.ospar.org/odims\\_data\\_files/](https://odims.ospar.org/odims_data_files/)) and by inshore vessels the area where >4 dredgers and trawlers were reported fishing in the ScotMap Inshore Fisheries Mapping Project (Kafas et al. 2014)

Ecoregion	Fishing impact	Area (km <sup>2</sup> )	Percentage of area
<b>Celtic Seas</b>			
	Offshore vessels	10193.192	39
	Inshore vessels	4021.080	15
	Total fished	14214.272	53
	Not fished	12310.761	46
	Total	26525.033	100
<b>Greater North Sea</b>			
	Offshore vessels	30388.753	71
	Inshore vessels	2030.666	5
	Total fished	32419.419	76
	Not fished	10281.401	24
	Total	42700.820	100

pendent data from the other ecoregion confirmed the poor transferability of models between ecoregions. Further qualitative examination of the predicted and observed distributions, in conjunction with fishing data, revealed that these differences likely reflect the modification of sea pen communities by demersal trawling. *F. quadrangularis* appeared to be largely absent from its expected range in the GRNS, whilst *P. phosphorea* and *V. mirabilis* appeared less vulnerable to trawling impacts.

The broadscale geographical distributions predicted by the full UKCS model (using data from both ecoregions) are largely consistent with sea pen observation data held by the Ocean Biogeographic Information System (OBIS 2020) and the UK National Biodiversity Network Atlas (NBN 2020). This is unsurprising, given the inclusion of NBN data in the model; however, there were some inconsistencies in predicted and observed distributions. All 3 sea pens (particularly *F. quadrangularis*) are predicted to occur at the shelf break at the extreme south-western tip of the UK EEZ, but are not represented by direct observations in this particular dataset. Fields of *F. quadrangularis* were observed on the shelf break in



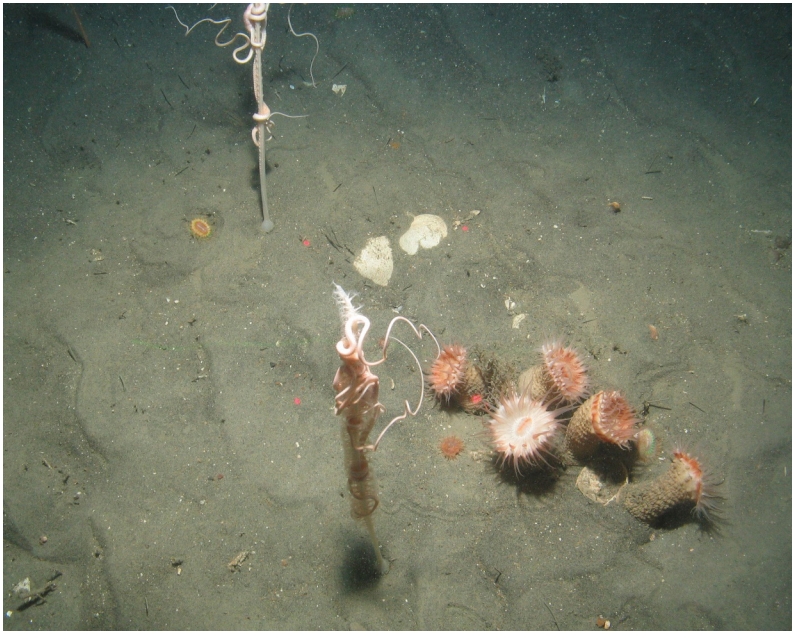


Fig. 7. *Funiculina quadrangularis* (and commensal brittlestar *Asteronyx lovenii*) on sandy mixed sediment in the North Sea. Photograph: Joint Nature Conservation Committee (JNCC) and Cefas 2013

this region during the Flanders Research Foundation funded MINIMOUNDS project (2013–2016; K. L. Howell unpubl. data). This provides an example of the efficacy of the model, even where direct observations are absent. The model also predicts all 3 sea pen species at limited patches in the Western Channel. *P. phosphorea* and *V. mirabilis* are both recorded in this area from NBN and Cefas data, but *F. quadrangularis* is not (also being absent from the OBIS data). It remains unclear whether this absence is due to overprediction by the model (which in this area has been assigned low confidence), a sparse distribution leading to a low probability of encounter, or absence due to human impacts. The different time periods covered by the sea pen observation data (1961–2019) and environmental data, such as temperature and primary productivity (2017–2019), add a potential source of error where the conditions may have been different at the time the samples were collected.

The sea pen distributions observed for all 3 species are strongly influenced by lower seabed temperatures, low wave and current velocities, and low levels of suspended particulate matter, confirming a requirement for deep, depositional environments. Somewhat surprisingly, sediment composition was determined to be less important. This is likely because sediment composition is equally driven by the more influential variables (e.g. in a negative cor-

relation between wave and current velocities and fine sediments). All 3 species did, however, display a strong affinity for muds and muddy sands. Primary productivity, which was influential in the models for *P. phosphorea* and *V. mirabilis*, is likely to act as a proxy for coastal and enclosed water bodies. The varying responses to primary productivity in the 2 ecoregions may result from the inshore/offshore balance of sample locations between the regions. A high number of samples in otherwise suitable habitat in the CELS ecoregion are from inlets and lochs with corresponding high primary productivity. On the other hand, the deeper mud basins are associated with lower primary productivity in the GRNS, being located offshore in non-photoc waters.

Despite the similarities in habitat requirements, the 3 species displayed some divergence in niche breadth, with *V. mirabilis* and *P. phosphorea* tolerating a wider range of environmental conditions than *F. quadrangularis*, reflected in superior model accuracy for the latter species. *V. mirabilis* and *P. phosphorea* were more tolerant of increased sand fractions, and slightly higher current velocities. *V. mirabilis* showed the greatest tolerance to increased wave velocities and concentration of particulate matter in the water column. This explains its distribution, which extends further inshore in the North Sea than that of *P. phosphorea*, and its presence in the Celtic and Irish Seas, which have higher suspended sediment loads. These findings are supported by those of Greathead et al. (2015), who noted that *V. mirabilis* showed the widest tolerance of a broader range of environmental conditions in Scottish inshore waters, and that *F. quadrangularis* occupied the narrowest niche. However, it must be noted that we are only discussing the distribution of these species on the continental shelf. *F. quadrangularis* is a cosmopolitan species that, although most commonly reported on the continental shelf and at the top of continental slopes, has been recorded as deep as 2000 m (De Clippele et al. 2015, Lauria et al. 2017, Bastari et al. 2018). Consequently, the models do not capture the full distribution, which extends to deeper, colder water than the UKCS.

Model transferability between ecoregions was exceptionally poor for *F. quadrangularis*, with almost

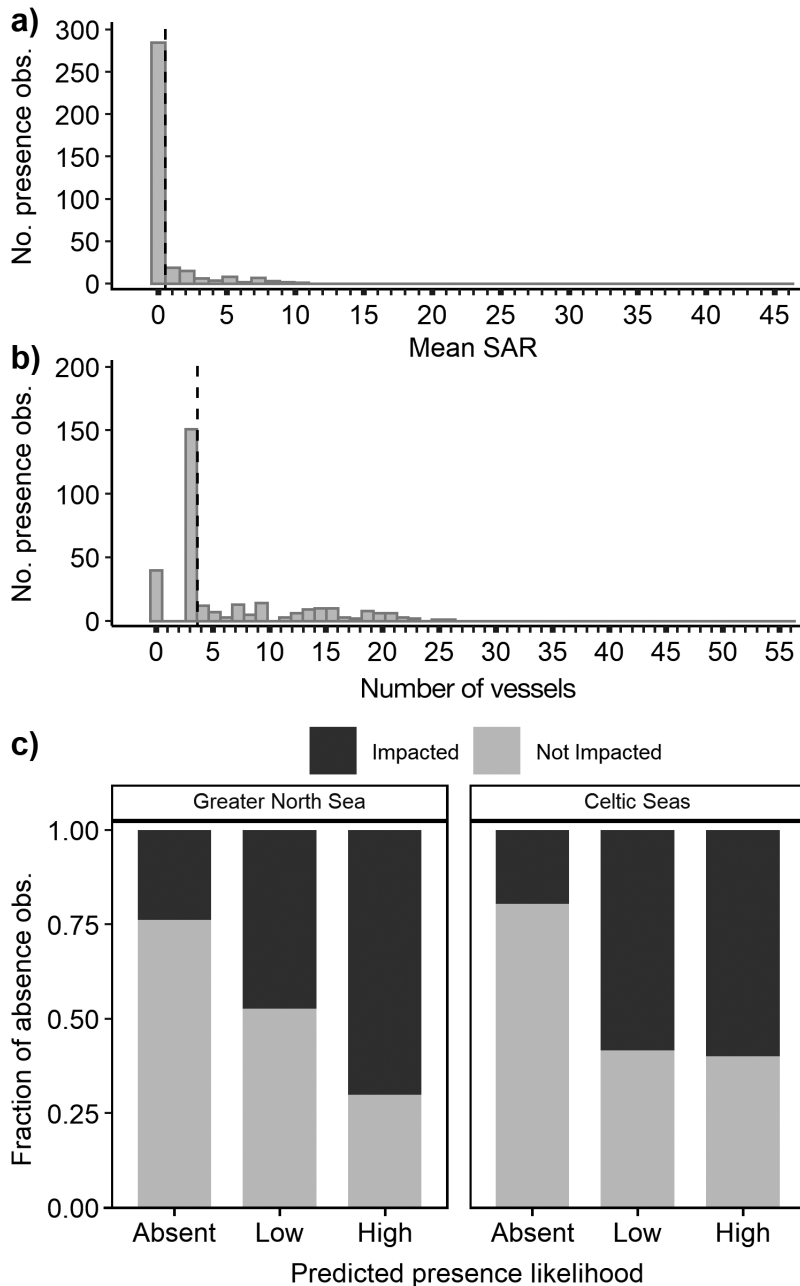


Fig. 8. Frequency of *Funiculina quadrangularis* presences across (a) the range of observed swept area ratio (SAR) and (b) number of vessels across the full UK continental shelf (UKCS) study area. Dashed lines indicate the threshold values that were used to split each fishing intensity layer into impacted and non-impacted categories. These were used to plot (c) the fraction of *F. quadrangularis* observed absences in impacted and non-impacted areas, where presence was either not predicted, predicted with low likelihood or predicted with high likelihood

no overlap in the observed conditions in which this species was predicted. The GRNS model predicted *F. quadrangularis* in topographically elevated areas of sediment with higher sand content. This result entirely contradicts the CELS model and the findings

of previous Scottish inshore studies (Greathead et al. 2007, 2015), which predicted *F. quadrangularis* in topographically concave muddy basins such as inshore lochs and straits. The apparent 'preference' of *F. quadrangularis* for sandy elevated sediments in the GRNS is therefore thought to reflect artefactual variation in the sea pen dataset. Such a striking disparity in agreement points towards the depletion of *F. quadrangularis* within its natural range in the GRNS (and possibly in the western English Channel, where it has been predicted but not observed), caused by a strongly influential factor which is not accounted for by the environmental layers, and which varies in intensity and spatial distribution between the 2 ecoregions.

Demersal trawling is widely acknowledged as one of the main pressures on marine benthic ecosystems (Halpern et al. 2008, Hiddink et al. 2017, Rijnsdorp et al. 2018) and is well known to modify faunal assemblages and affect species distributions (e.g. Tillin et al. 2006, Hinz et al. 2009). Where prevalent, this can result in sensitive species not occupying their full predicted range or being restricted to suboptimal 'edge of niche' habitats (Braunisch et al. 2008). Evidence from the North Sea shows that otter trawlers targeting *Nephrops* and demersal flatfish focus effort in muddy depressions with low bed shear stress (van der Reijden et al. 2018), habitat which is predicted as optimal for sea pens by the CELS model. The Fladen Grounds is one such area, which has been reported to support the largest *Nephrops* fishery in the world (Ungfors et al. 2013). *F. quadrangularis* was not observed in the muddy depressions of the Fladen Grounds, where gridded VMS data indicate benthic abrasion by

trawling (and where the CELS model predicted its occurrence). This species did, however, occur in patches of interspersed sandier mixed sediment (as confirmed from imagery data; see Fig. 7) that do not support *Nephrops* and are therefore unlikely to be targeted



by fishers. We therefore hypothesise that *F. quadrangularis*, the most sensitive of the 3 species (Ager 2003, Greathead et al. 2007), is subject to range modification in the GRNS, with the observed occurrences representing a remnant population occupying the elevated sandy areas with lower disturbance. This finding reflects a significant shift in the niche of *F. quadrangularis* in the GRNS.

Estimating the accurate spatial range and intensity of demersal fishing activities is inherently challenging, often limiting the extent to which pressure–state relationships can be explored at finer spatial scales (Lee et al. 2010, Lambert et al. 2012). In this study, we cannot categorically demonstrate cause and effect, given the coarse resolution of the OSPAR VMS data units in relation to the sea pen occurrence data. However, the overwhelming majority of *F. quadrangularis* was observed in grid cells where the swept area ratio was zero, providing anecdotal support for our hypothesis. This hypothesis is further supported by the findings of Greathead et al. (2005) that the density of *F. quadrangularis* in the Scottish Cuillin Sound (an area of high *Nephrops* trawling intensity), was lower in contrast to the environmentally comparable (but low trawling intensity) Hebridean Sound. Further afield, Pierdomenico et al. (2018) and Fabri et al. (2014) both observed low abundances of *F. quadrangularis* in chronically trawled Mediterranean canyons, whilst Malecha & Stone (2009) found that experimental breakage of the morphologically similar *Halipteris willemoesi* resulted in extremely high mortality rates. Although the available evidence suggests that *F. quadrangularis* is vulnerable to trawling impacts, there is no such indication that creel fishing for *Nephrops* has a negative impact on its distribution. Eno et al. (2001) observed that *F. quadrangularis* was able to re-insert itself into the sediment after uprooting by creel pots. The theory that *F. quadrangularis* is more resilient to creeling impacts is supported by our study, as *F. quadrangularis* was shown to proliferate in Scottish sea lochs where smaller fishing vessels operate, in particular Loch Linnhe, where the *Nephrops* fishery is predominantly creel-based (Berx et al. 2015).

Whilst the high sensitivity of *F. quadrangularis* to trawling appears to severely diminish the transferability of the CELS and GRNS models, this effect was not as pronounced for *V. mirabilis* and *P. phosphorea*. Actual presences in the North Sea were captured effectively by the CELS model, although this model tended to overpredict. Conversely, the GRNS model underpredicted presences in the CELS. This trend suggests that the distributions of these species are

more limited in the GRNS and perhaps do not cover the entire natural range of environmental variability covered by observations in the CELS. *V. mirabilis* and *P. phosphorea* both display far broader geographical ranges than *F. quadrangularis*, being frequent in areas of demersal fishing activity, both in the CELS and GRNS. There is therefore no clear evidence that trawling limits their distribution on the UKCS, corroborating the findings of previous studies that *V. mirabilis* and *P. phosphorea* are less vulnerable to fishing impacts than *F. quadrangularis* (Howson & Davies 1991, Tuck et al. 1998, Greathead et al. 2005). This is likely due to a combination of wider niche breadth, higher abundance, physiological traits (e.g. flexibility) and the ability to withdraw into a burrow (Ambroso et al. 2013, Chimienti et al. 2018).

Previous publications have proposed that sea pen presence and/or abundance could serve as indicators of condition in mud habitats (Greathead et al. 2007, Murray et al. 2015). Our study suggests that the presence of *F. quadrangularis* could potentially act as an indicator of healthy, undisturbed conditions, although the absence of this species in suitable habitat should not automatically be assumed to indicate poor habitat condition, as sea pens are known to display patchy or aggregated distributions (Langton et al. 1990, Greathead et al. 2007). Although the predicted distributions of *V. mirabilis* and *P. phosphorea* did not appear to be adversely affected by demersal fishing, it should be noted that these models only reflect presence or absence of sea pens, offering no insight on density or condition. In their 2015 study of the Fladen Grounds, Murray et al. (2015) found that the density of *P. phosphorea* decreased along a gradient of increasing trawling pressure, suggesting that density may have potential as an indicator at higher levels of fishing intensity. We suggest that further studies be conducted to explore the relationships between trawling activity and density of *P. phosphorea* and *V. mirabilis*. A greater understanding of these relationships would allow evaluation of sea pen management measures, in turn enabling ecologically sound management decisions.

A number of UK MPAs have been designated to protect sea pens; however, management measures within MPAs are spatially variable and trawling is not necessarily under active management in areas of suitable habitat. The vast majority of *F. quadrangularis* observations and areas of suitable habitat identified in this study are not currently covered by MPAs designated to protect sea pen communities (particularly in the GRNS, although they are found within those managed for different habitat types and could benefit from management measures in these areas).

The predicted distribution models presented in this study (alongside further pressure-state studies on *P. phosphorea* and *V. mirabilis*) could be used to refine management of specific MPA zones and gear types, thus optimizing the balance between access to commercially important fishing grounds and protection of sea pen communities.

This study has demonstrated the effectiveness of the random forest method for modelling vulnerable species distributions and highlighted the utility of national (and international) data platforms for addressing conservation questions at large scales. Whilst challenges still remain in reconciling the spatial and temporal disparities of sample data, environmental parameters and human pressures, our findings illustrate the value of broadscale qualitative comparisons between SDMs and human activity data for understanding fundamental pressure-state relationships. Benthic trawling is generally accepted as the greatest global threat to marine benthic fauna, yet the increasing accessibility of marine data creates opportunities for informed and targeted mitigation. We anticipate that the approach used in this study could be widely applied to investigate trawling impacts on sensitive species at large scales, enabling marine spatial planners to make better management decisions and deliver tangible conservation outcomes.

**Acknowledgements.** We thank Peter Mitchell (Cefas) for advice on accessing and interpreting the environmental data layers obtained through the Cefas Data Hub; David Clare and Michaela Schratzberger (Cefas), Karema Randall and Will Hutchinson (Defra) for providing valuable feedback on drafts of this article; and the Joint Nature Conservation Committee (JNCC) for supplying sea pen data and network information for Scottish MPAs.

#### LITERATURE CITED

- ✦ Ager OED (2003) The tall sea pen (*Funiculina quadrangularis*). In: Tyler-Walters H, Hiscock K (eds) Marine Life Information Network: biology and sensitivity key information reviews. Marine Biological Association of the United Kingdom, Plymouth. [www.marlin.ac.uk/species/detail/1154](http://www.marlin.ac.uk/species/detail/1154) (accessed 21 Oct 2020)
- Aguilar R, Perry AL, López J (2017) Conservation and management of vulnerable marine benthic ecosystems. In: Rossi S, Bramanti L, Gori A, Orejas C (eds) Marine animal forests. Springer, Cham, p 1–28
- ✦ Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43: 1223–1232
- ✦ Ambroso S, Dominguez-Carrió C, Grinyó J, López-González PJ and others (2013) *In situ* observations on withdrawal behaviour of the sea pen *Virgularia mirabilis*. *Mar Biodivers* 43:257–258
- ✦ Baillon S (2012) Deep cold-water corals as nurseries for fish larvae. *Front Ecol Environ* 10:351–356
- ✦ Bastari A, Pica D, Ferretti F, Micheli F, Cerrano C (2018) Sea pens in the Mediterranean Sea: habitat suitability and opportunities for ecosystem recovery. *ICES J Mar Sci* 75: 1722–1732
- ✦ Beazley L, Kenchington E, Murillo FJ, Lirette C, Guijarro J, McMillan A, Knudby A (2016) Species distribution modelling of corals and sponges in the Maritimes region for use in the identification of significant benthic areas. *Can Tech Rep Fish Aquat Sci* 3172, doi:10.13140/RG.2.1.1393.3044
- ✦ Berx B, Gallego A, Heath M and the MASTS Community (2015) Loch Linnhe and Firth of Lorn MASTS Case Study Workshop Report. *Scott Mar Freshw Sci* 6. Scottish Government, Edinburgh, doi:10.7489/1539-1
- ✦ Braunisch V, Bollmann K, Graf RF, Hirzel AH (2008) Living on the edge—modelling habitat suitability for species at the edge of their fundamental niche. *Ecol Model* 214: 153–167
- ✦ Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- ✦ Buhl-Mortensen P, Buhl-Mortensen L (2014) Diverse and vulnerable deep-water biotopes in the Hardangerfjord. *Mar Biol Res* 10:253–267
- ✦ Chimienti G, Angeletti L, Mastrototaro F (2018) Withdrawal behaviour of the red sea pen *Pennatulula rubra* (Cnidaria: Pennatulacea). *Eur Zool J* 85:64–70
- ✦ Conrad O, Bechtel B, Bock M, Dietrich H and others (2015) System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci Model Dev* 8:1991–2007
- ✦ Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. *Ecology* 88:2783–2792
- ✦ De Clippele LH, Buhl-Mortensen P, Buhl-Mortensen L (2015) Fauna associated with cold water gorgonians and sea pens. *Cont Shelf Res* 105:67–78
- ✦ Eigaard OR, Bastardie F, Breen M, Dinesen GE and others (2016) Estimating seabed pressure from demersal trawls, seines, and dredges based on gear design and dimensions. *ICES J Mar Sci* 73:i27–i43
- ✦ EMODnet Bathymetry Consortium (2016) EMODnet Digital Bathymetry (DTM 2016). Available online: <http://portal.emodnet-bathymetry.eu/> (accessed on 30 March 2018)
- ✦ Eno NC, MacDonald DS, Kinnear JAM, Amos SC and others (2001) Effects of crustacean traps on benthic fauna. *ICES J Mar Sci* 58:11–20
- ✦ Fabri MC, Pedel L, Beuck L, Galgani F, Hebbeln D, Freiwald A (2014) Megafauna of vulnerable marine ecosystems in French Mediterranean submarine canyons: spatial distribution and anthropogenic impacts. *Deep Sea Res II* 104: 184–207
- ✦ Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- ✦ Freeman EA, Moisen G (2008) PresenceAbsence: an R package for presence absence analysis. *J Stat Softw* 23:1–31
- ✦ Gale KSP, Hamel JF, Mercier A (2013) Trophic ecology of deep-sea Asteroidea (Echinodermata) from eastern Canada. *Deep Sea Res I* 80:25–36
- ✦ García-Matucheski S, Muniain C (2011) Predation by the nudibranch *Tritonia odhneri* (Opisthobranchia: Tritoniidae) on octocorals from the South Atlantic Ocean. *Mar Biodivers* 41:287–297
- Greathead CF, Donnan DW, Mair JM (2005) Impact of *Nephrops* trawling on the distribution of the sea pens

- Virgularia mirabilis*, *Pennatula phosphorea* and *Funiculina quadrangularis* in Scottish waters. Fisheries Research Services Internal Report No 02/05. Scottish Government, Aberdeen. <https://www.webarchive.org.uk/wayback/archive/3000/https://www.gov.scot/Uploads/Documents/IR0205.pdf> (accessed 21 Oct 2020)
- Greathead CF, Donnan DW, Mair JM, Saunders GR (2007) The sea pens *Virgularia mirabilis*, *Pennatula phosphorea* and *Funiculina quadrangularis*: distribution and conservation issues in Scottish waters. *J Mar Biol Assoc UK* 87: 1095–1103
- Greathead C, González-Irusta JM, Clarke J, Boulcott P, Blackadder L, Weetman A, Wright PJ (2015) Environmental requirements for three sea pen species: relevance to distribution and conservation. *ICES J Mar Sci* 72:576–586
- Grip K (2017) International marine environmental governance: a review. *Ambio* 46:413–427
- Halpern BS, Walbridge S, Selkoe KA, Kappel CV and others (2008) A global map of human impact on marine ecosystems. *Science* 319:948–952
- He KS, Bradley BA, Cord AF, Rocchini D and others (2015) Will remote sensing shape the next generation of species distribution models? *Remote Sens Ecol Conserv* 1:4–18
- Hiddink JG, Jennings S, Kaiser MJ (2006) Indicators of the ecological impact of bottom-trawl disturbance on seabed communities. *Ecosystems* 9:1190–1199
- Hiddink JG, Jennings S, Sciberras M, Szostek CL and others (2017) Global analysis of depletion and recovery of seabed biota after bottom trawling disturbance. *Proc Natl Acad Sci USA* 114:8301–8306
- Hill JM, Wilson E (2000) Slender sea pen (*Virgularia mirabilis*). In: Tyler-Walters H, Hiscock K (eds) Marine Life Information Network: biology and sensitivity key information reviews. Marine Biological Association of the United Kingdom, Plymouth. [www.marlin.ac.uk/species/detail/1396](http://www.marlin.ac.uk/species/detail/1396) (accessed 21 Oct 2020)
- Hill JM, Tyler-Walters H, Garrard S (2020) Seapens and burrowing megafauna in circalittoral fine mud. In: Tyler-Walters H, Hiscock K (eds) Marine Life Information Network: biology and sensitivity key information reviews. Marine Biological Association of the United Kingdom, Plymouth. [www.marlin.ac.uk/habitats/detail/131](http://www.marlin.ac.uk/habitats/detail/131) (accessed 21 Oct 2020)
- Hinz H, Prieto V, Kaiser M (2009) Trawl disturbance on benthic communities: chronic effects and experimental predictions. *Ecol Appl* 19:761–773
- Hixon MA, Tissot BN (2007) Comparison of trawled vs untrawled mud seafloor assemblages of fishes and macroinvertebrates at Coquille Bank, Oregon. *J Exp Mar Biol Ecol* 344:23–34
- Howson CM, Davies LM (1991) Marine Nature Conservation Review, surveys of Scottish sealochs. A towed video survey of Loch Fyne. Vol 1. Report to the Nature Conservancy Council from the University Marine Biological Station, Millport
- Howson CM, Connor DW, Holt RHF (1994) The Scottish Sea lochs—an account of surveys undertaken for the MNCR. Joint Nature Conservation Committee Rep 164 (Marine Nature Conservation Review Report No. MNCR/SR/27). Joint Nature Conservation Committee, Peterborough
- ICES (2005) ICES Statistical Areas. [https://gis.ices.dk/gis/rest/services/ICES\\_reference\\_layers/ICES\\_Areas/MapServer](https://gis.ices.dk/gis/rest/services/ICES_reference_layers/ICES_Areas/MapServer) (accessed 21 Oct 2020)
- ICES (2015) ICES Ecoregions. <https://gis.ices.dk/geonet-work/srv/api/records/4745e824-a612-4a1f-bc56-b540772166eb> (accessed 21 Oct 2020)
- ICES (2019) Spatial distribution of fishing effort and physical disturbance of benthic habitats by mobile bottom trawl fishing gear using VMS. Technical Guidelines. Report of the ICES Advisory Committee. ICES Advice 2019. <https://doi.org/10.17895/ices.advice.4683>
- ICES (2020) Definition and rationale for ICES ecoregions. Report of the ICES Advisory Committee, 2020. ICES Advice 2020. <https://doi.org/10.17895/ices.advice.6014>
- Jiménez-Valverde A (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Glob Ecol Biogeogr* 21:498–507
- Jones H (2008) Phosphorescent sea pen (*Pennatula phosphorea*). In: Tyler-Walters H, Hiscock K (eds) Marine Life Information Network: biology and sensitivity key information reviews. Marine Biological Association of the United Kingdom, Plymouth, UK. [www.marlin.ac.uk/species/detail/1817](http://www.marlin.ac.uk/species/detail/1817) (accessed 21 Oct 2020)
- Kafas A, McLay A, Chimienti M, Gubbins M (2014) ScotMap inshore fisheries mapping in Scotland: recording fishermen's use of the sea. *Scott Mar Freshw Sci* 5. Scottish Government, Edinburgh, doi:10.4789/1554-1
- Kenchington E, Murillo FJ, Lirette C, Sacau M and others (2014) Kernel density surface modelling as a means to identify significant concentrations of Vulnerable Marine Ecosystem indicators. *PLOS ONE* 9:e109365
- Kinlan BP, Poti M, Drohan AF, Packer DB, Dorfman DS, Nizinski MS (2020) Predictive modeling of suitable habitat for deep-sea corals offshore the Northeast United States. *Deep Sea Res I* 158:103229
- Knudby A, Lirette C, Kenchington E, Murillo FJ (2013) Species distribution models of black corals, large gorgonian corals and sea pens in the NAFO Regulatory Area. NAFO SCR Doc. 13/078. Serial No. N6276. Northwest Atlantic Fisheries Organization, SC Ecosystem Science and Assessment Working Group (WGESA). <https://www.nafo.int/Portals/0/PDFs/sc/2013/scr13-078.pdf>
- Krigsman LM, Yoklavich MM, Dick EJ, Cochrane GR (2012) Models and maps: predicting the distribution of corals and other benthic macro-invertebrates in shelf habitats. *Ecosphere* 3(1):art3
- Kröger S, Parker R, Cripps G, Williamson P (2018) Shelf seas: the engine of productivity. Policy Report on NERC-Defra Shelf Sea Biogeochemistry Programme. Cefas, Lowestoft. [https://www.uk-ssb.org/shelf\\_seas\\_report.html](https://www.uk-ssb.org/shelf_seas_report.html) (accessed 21 Oct 2020)
- Kursa MB, Rudnicki WR (2010) Feature selection with the boruta package. *J Stat Softw* 36:1–13
- Lambert GI, Jennings S, Hiddink JG, Hintzen NT, Hinz H, Kaiser MJ, Murray LG (2012) Implications of using alternative methods of vessel monitoring system (VMS) data analysis to describe fishing activities and impacts. *ICES J Mar Sci* 69:682–693
- Langton RW, Langton EW, Theroux RB, Uzman JR (1990) Distribution, behavior and abundance of sea pens, *Pennatula aculeata*, in the Gulf of Maine. *Mar Biol* 107:463–469
- Lauria V, Garofalo G, Fiorentino F, Massi D and others (2017) Species distribution models of two critically endangered deep-sea octocorals reveal fishing impacts on vulnerable marine ecosystems in central Mediterranean Sea. *Sci Rep* 7:8049
- Lee J, South AB, Jennings S (2010) Developing reliable, repeatable, and accessible methods to provide high-res-

- olution estimates of fishing-effort distributions from vessel monitoring system (VMS) data. *ICES J Mar Sci* 67: 1260–1271
- ✦ Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Glob Ecol Biogeogr* 17:145–151
- ✦ Malecha PW, Stone RP (2009) Response of the sea whip *Halipteris willemoesi* to simulated trawl disturbance and its vulnerability to subsequent predation. *Mar Ecol Prog Ser* 388:197–206
- ✦ Manel S, Ceri Williams H, Ormerod SJ (2001) Evaluating presence/absence models in ecology: the need to account for prevalence. *J Appl Ecol* 38:921–931
- ✦ Marine Management Organisation (2019) UK sea fisheries statistics 2018. <https://www.gov.uk/government/statistics/uk-sea-fisheries-annual-statistics-report-2018> (accessed 21 Oct 2020)
- ✦ Marshall CE, Glegg GA, Howell KL (2014) Species distribution modelling to support marine conservation planning: the next steps. *Mar Policy* 45:330–332
- Mitchell P, Aldridge J, Diesing M (2019a) Predictor variables and groundtruth samples for north-west European continental shelf quantitative sediment analysis. Cefas, Lowestoft. doi:10.14466/CefasDataHub.62
- Mitchell P, Aldridge J, Diesing M (2019b) Quantitative sediment composition predictions for the north-west European continental shelf. Cefas, Lowestoft. doi:10.14466/CefasDataHub.63
- ✦ Mitchell PJ, Aldridge J, Diesing M (2019c) Legacy data: how decades of seabed sampling can produce robust predictions and versatile products. *Geosciences* 9:182
- Murillo FJ, Kenchington E, Beazley L, Lirette C and others (2016) Distribution modelling of sea pens, sponges, stalked tunicates and soft corals from research vessel survey data in the Gulf of St. Lawrence for use in the identification of significant benthic areas. *Can Tech Rep Fish Aquat Sci* 3170, doi:10.13140/RG.2.1.2966.1689
- ✦ Murray J, Jenkins C, Eggleton J, Whomersley P, Robson L, Flavell B, Hinchin H (2015) The development of monitoring options for UK MPAs: Fladen Grounds R&D case study. Joint Nature Conservation Committee/Cefas Partnership Report Series No. 9. JNCC, Peterborough. <https://hub.jncc.gov.uk/assets/2d594d86-06f1-419d-8cb6-db2f61b5be9c>
- ✦ NBN (2020) National Biodiversity Network Atlas. <https://www.nbnatlas.org> (accessed 21 Oct 2020)
- ✦ OBIS (2020) Global map of species distribution using gridded data. Ocean Biogeographic Information System. [www.obis.org](http://www.obis.org) (accessed 21 Oct 2020)
- ✦ Pierdomenico M, Russo T, Ambroso S, Gori A and others (2018) Effects of trawling activity on the bamboo-coral *Isidella elongata* and the sea pen *Funiculina quadrangularis* along the Gioia Canyon (Western Mediterranean, southern Tyrrhenian Sea). *Prog Oceanogr* 169:214–226
- ✦ Prasad AM, Iverson LR, Liaw A (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9:181–199
- R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- ✦ Reiss H, Birchenough S, Borja A, Buhl-Mortensen L and others (2015) Benthos distribution modelling and its relevance for marine ecosystem management. *ICES J Mar Sci* 72:297–315
- ✦ Rijnsdorp AD, Bastardie F, Bolam SG, Buhl-Mortensen L and others (2016) Towards a framework for the quantitative assessment of trawling impact on the seabed and benthic ecosystem. *ICES J Mar Sci* 73(Suppl 1): i127–i138
- ✦ Rijnsdorp AD, Bolam SG, Garcia C, Hiddink JG, Hintzen NT, van Denderen PD, van Kooten T (2018) Estimating sensitivity of seabed habitats to disturbance by bottom trawling based on the longevity of benthic fauna. *Ecol Appl* 28:1302–1312
- ✦ Rogers AD, Gianni M (2010) The implementation of the UNGA Resolutions 61/105 and 64/72 in the management of deep-sea fisheries on the high seas. Report prepared for the Deep-Sea Conservation Coalition, International Programme on the State of the Ocean, London. [www.savethehighseas.org/publicdocs/61105-Implementation-finalreport.pdf](http://www.savethehighseas.org/publicdocs/61105-Implementation-finalreport.pdf) (accessed 21 Oct 2020).
- ✦ Santika T (2011) Assessing the effect of prevalence on the predictive performance of species distribution models using simulated data. *Glob Ecol Biogeogr* 20:181–192
- ✦ Shaw RL, Booth A, Sutton AJ, Miller T and others (2004) Finding qualitative research: an evaluation of search strategies. *BMC Med Res Methodol* 4:5
- ✦ Smith RJ, Eastwood PD, Ota Y, Rogers SI (2009) Developing best practice for using Marxan to locate Marine Protected Areas in European waters. *ICES J Mar Sci* 66: 188–194
- ✦ Sundblad G, Bergström U, Sandström A (2011) Ecological coherence of marine protected area networks: a spatial assessment using species distribution models. *J Appl Ecol* 48:112–120
- ✦ Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIO-MOD—a platform for ensemble forecasting of species distributions. *Ecography* 32:369–373
- ✦ Thuiller W, Georges D, Engler R, Breiner F (2020) Biomod2: ensemble platform for species distribution modeling. R package version 3.4.6. <https://cran.r-project.org/web/packages/biomod2/biomod2.pdf> (accessed 21 Oct 2020)
- ✦ Thurstan RH, Brockington S, Roberts CM (2010) The effects of 118 years of industrial fishing on UK bottom trawl fisheries. *Nat Commun* 1:15
- ✦ Tillin HM, Hiddink JG, Jennings S, Kaiser MJ (2006) Chronic bottom trawling alters the functional composition of benthic invertebrate communities on a sea-basin scale. *Mar Ecol Prog Ser* 318:31–45
- ✦ Tuck ID, Hall SJ, Robertson MR, Armstrong E, Basford DJ (1998) Effects of physical trawling disturbance in a previously unfished sheltered Scottish sea loch. *Mar Ecol Prog Ser* 162:227–242
- ✦ Tyberghein L, Verbruggen H, Pauly K, Troupin C, Mineur F, De Clerck O (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Glob Ecol Biogeogr* 21:272–281
- ✦ Ungfors A, Bell E, Johnson ML, Cowing D, Dobson NC, Bublit R, Sandell J (2013) *Nephrops* fisheries in European waters. *Adv Mar Biol* 64:247–314
- ✦ van der Reijden KJ, Hintzen NT, Govers LL, Rijnsdorp AD, Olff H (2018) North Sea demersal fisheries prefer specific benthic habitats. *PLOS ONE* 13:e0208338
- ✦ Wilson MT, Andrews AH, Brown AL, Cordes EE (2002) Axial rod growth and age estimation of the sea pen, *Halipteris willemoesi* Kölliker. *Hydrobiologia* 471: 133–142