# A verified distribution model for the lesser sandeel *Ammodytes marinus*

**Rebecca Langton\*, Philip Boulcott, Peter J. Wright**

Marine Laboratory, Marine Scotland Science, 375 Victoria Road, Aberdeen, AB11 9DB, UK

ABSTRACT: The lesser sandeel *Ammodytes marinus* is a key component of the North East Atlantic ecosystem but little is known about its distribution outside of fished areas. In this study, species distribution models were developed to predict the occurrence and density of sandeels in parts of the North Sea and Celtic Seas regions. A hurdle model was found to be the best fitting model with the highest predictive performance; model evaluation with independent data demonstrated that it had significant discrimination ability across the study region. Percentage silt was the most important variable in predicting occurrence, and percentage sand had a strong influence on density, consistent with past local studies. Slope was also a significant explanatory variable, especially for predicting density, as buried sandeels avoided strongly sloping areas such as the edges of sand banks. A predicted preferred depth range of 30–50 m was consistent with many previous studies, although the depth response did appear partially biased by the depth range investigated in the training data set. Overall, the predicted distribution did not indicate that there were large areas of unexploited habitat. However, some small areas known to be important to sandeel predators were identified by the model. The distribution model helps refine past inferences about sandeel availability to predators and indicates to marine planners potential areas where anthropogenic impacts should be considered.

KEY WORDS: Species distribution model · *Ammodytes marinus* · Sandeels · Hurdle model · Fish distribution · Spatial management

## 1. INTRODUCTION

Sandeels (or sand lance) of the family Ammodytidae are small, slender fish that are an important component of food webs in the North Atlantic (Sherman et al. 1981, Harwood & Croxall 1988, Sparholt 1990, Wanless et al. 1998). In the North East Atlantic, the lesser sandeel *Ammodytes marinus* is the most abundant species, supporting the largest fishery in the North Sea and, previously, an inshore fishery off the Scottish west coast. In both regions they are a common prey of seabirds, seals and cetaceans (Halley et al. 1995, Santos et al. 2004, Wanless et al. 2018, Wilson & Hammond 2019) and are regarded as a species of conservation importance. Due to the magnitude of the fishery and the importance of sandeels to marine predators, there is considerable interest in their distribution

and population dynamics (Monaghan 1992, Carroll et al. 2017, Wilson & Hammond 2019, Hill et al. 2020). Studies of sandeel distribution have been useful in explaining the foraging locations of sandeel predators, such as piscivorous seabirds and cetaceans (Monaghan et al. 1996, Wright & Begg 1997, Wanless et al. 1998, Herr et al. 2009). However, information on sandeel distribution at a much larger scale is required in order to identify possible areas where foraging marine predators and human activity may be in conflict.

In common with other *Ammodytes* species, *A. marinus* has a close association with sandy substrates into which they burrow, following a planktonic larval phase (Reay 1970, Wright et al. 2000, Tien et al. 2017, Greene et al. 2020). Sediment preference of *A. marinus* is well established from both field-based and experimental studies, with low silt and high sand

fractions affecting presence and density (Wright et al. 2000, Holland et al. 2005, Tien et al. 2017). Buried sandeels are capable of sustaining their oxygen requirements by the advection of oxygen-rich water towards their mouth by gill ventilation (Behrens et al. 2007), which may be enhanced by hydrodynamic forcing related to the topographic relief of the sediment surface, where high densities often occur (Wright et al. 2000, Freeman et al. 2004). Despite the considerable evidence base on habitat preference in *A. marinus*, knowledge of their distribution outside fished areas is very limited and there is currently no distribution model for this species.

There are many sources of local information on the benthic distribution of *A. marinus* from grab and dredge samples, some of which have been used to support fishery assessments (Wright et al. 2000, Greenstreet et al. 2006, Engelhard et al. 2008, ICES 2010). Information on the distribution of fishing activity in the North Sea compiled from fishing fleets (Jensen et al. 2011) and vessel monitoring system activity (Engelhard et al. 2008, ICES 2010) also provides an indication of where large sandeel aggregations occur. However, fisheries targeting *A. marinus* operate where this species aggregates to feed, which is usually at the edge of the banks (Jensen 2001, Mackinson & van der Kooij 2006), and foraging schools may migrate several km from the areas where they bury each day (Engelhard et al. 2008, van der Kooij et al. 2008, Wright et al. 2019). Consequently, fishing distribution is only related to benthic distribution at a coarse scale and is also not relevant to areas where fishing is either not permitted or constrained by varying ground suitability. Many of the areas where seabirds, mammals and fish congregate to feed on sandeels have been found in regions where there is little or no data on sandeel presence or density (Monaghan et al. 1996, Temming et al. 2004). This has led several researchers to infer sandeel distribution from sediment characteristics alone (Macleod et al. 2004, Anderwald et al. 2012), which may fail to reflect important environmental characteristics of sandeel habitat.

Correlative species' distribution models (SDMs), in which field observations are coupled with corresponding environmental variables within a statistical modelling framework, are now being used increasingly for marine species (Melo-Merino et al. 2020). There are a wide variety of modelling algorithms that can be used for SDMs with varying degrees of complexity, ranging from 'simple' regression models to more complex machine learning techniques. One key application of SDMs is to make predictions be-

yond the geographic area originally sampled. However, model algorithms which fit complex, non-linear relationships, such as machine learning techniques or generalised additive models, can overfit the training data and therefore perform poorly when applied to new areas (Merow et al. 2014, Bell & Schlaepfer 2016, Gregr et al. 2019). This issue can be alleviated by methods that give the user more control over the form of response curves, leading to better predictions (Merow et al. 2014, Gregr et al. 2019). In addition to model complexity, predictive performance is also improved when transferring to areas with similar environmental conditions and for species with more specialised niches (Jarnevich et al. 2015, Yates et al. 2018, Qiao et al. 2019). Given the availability of appropriate environmental variables, SDMs can inform broader scale spatial management in species that are difficult and costly to sample, which is the case for many marine species (Reiss et al. 2015).

This study predicts the potential distribution of *A. marinus* in parts of the Greater North Sea and Celtic Seas OSPAR regions. SDMs were developed using data on lesser sandeel abundance and sediment from grab surveys in the east of Scotland, together with lower resolution environmental data. Grab sampling has been widely used to investigate sandeel distribution in several studies (Wright et al. 2000, Høines & Bergstad 2001, Holland et al. 2005, Baker et al. 2019). Zero counts in grab surveys may reflect unsuitable habitat or low densities of sandeel in suitable habitat due to the small area sampled by this gear (Holland et al. 2005). However, the sampling approach does allow the environmental relationships to be parameterised using data on sandeel abundance and associated sediment collected at a high spatial resolution. Model validation is important when extrapolating model outputs to non-sampled areas (Elith & Leathwick 2009), which could be a constraint in this study, so independent data from other benthic sampling in both the North Sea and Celtic Seas regions were used to validate model predictions.

## 2. MATERIALS AND METHODS

### 2.1. Determining the relationship between sandeels and environmental variables

#### 2.1.1. Data

Lesser sandeel *Ammodytes marinus* abundance data based on Day grab samples (sample area 0.096 m$^2$) from Holland et al. (2005, see their Table 1) were used

to parameterise the model as they provide point estimates of both sandeels and corresponding sediment composition. The surveys took place off the Firth of Forth, Scotland, between 1998 and 2003, inclusive. Samples were predominantly collected at times when sandeels were buried all day and during a period of low to average stock size (ICES 2020a) so that all but the most marginally suitable habitat may be expected to be colonised. Two core samples of sediment were taken for particle size analysis (PSA) from the Day grab before sorting for sandeels. The PSA of the sediment core samples were used to estimate percentage of silt (<63 µm), sand (63 µm to 2 mm) and gravel (>2 mm) (Holland et al. 2005). In total, there were 2885 data points. If pebbles prevent the grab closing, it is possible that using data from a grab survey may result in a bias towards softer sediments. However, generally the Day grab collected an adequate sample on the first deployment (Holland et al. 2005), and extensive RoxAnn surveys have shown that the study area has low roughness and hardness, indicating a predominance of softer sediments that are suited to the Day grab (Greenstreet et al. 2010).

Along with sediment characteristics, the model also included variables that were derived from bathymetry, as these have been found to be relevant for other benthic species and were available at an appropriate resolution for the extent of the study area. Bathymetry data were downloaded from the EMODnet Bathymetry Portal (https://portal.emodnet-bathymetry.eu/#) on 10 Jan 2019 at a resolution of 1/16 arc minutes (which is approximately 65 × 115 m in the Firth of Forth). A slope layer was generated with the 'terrain' function in the 'raster' (Hijmans 2019) package in R v.3.6.2 (R Core Team 2019). The depth and slope values for the grab locations were extracted from these original layers using the 'extract' function, also from the 'raster' package.

### 2.1.2. Model selection

Seventy percent of the grabs did not contain any sandeels. Zero-inflation, having more zeros than would be expected based on count distributions such as the Poisson or negative binomial, is common in fish abundance data. To address zero-inflation, absences can be separated into 2 broad categories. Firstly, there are the absences, where the species is not observed in a certain location because the habitat is unsuitable (true zeros or structural zeros). Then there are the absences when the habitat is suitable but the species is not recorded, possibly linked to low local abundance

or stochasiticity in the observation process resulting in an absence record, even when present in the area (false zeros or sampling zeros; Zuur et al. 2009).

Two groups of statistical models are available to analyse zero-inflated data: hurdle models and zero-inflated models. Hurdle models (a.k.a delta models) are a 2 step approach. During the first step, the data are modelled as presence/absence data. In the second step, the abundances observed in the presence-only data are modelled using an appropriate distribution that cannot produce zeros, such as a truncated Poisson (Zeileis et al. 2008). Zero-inflated models are a type of mixture model and also have 2 components. Again, the first component involves modelling the presence/absence data; however, in contrast to the hurdle model, the second component models the count data using a distribution that can produce zeros (which would represent true zeros) (Zeileis et al. 2008, Zuur et al. 2009). The difference between the 2 models is that the hurdle model does not distinguish between the different types of zero during model fitting, while the zero-inflated model separates the true and false zeros into the 2 components (Zuur et al. 2009).

Four regression models capable of dealing with zero-inflated count data were initially developed for the Firth of Forth grab data: hurdle models with a left-truncated Poisson or negative binomial distribution with a log link for the count component, and zero-inflated models assuming a Poisson or a negative binomial distribution with a log link. For all models, a binomial distribution with a logit link function was assumed for the zero component. A left-truncated discrete distribution was used for the hurdle model as this is more suited to count data than the lognormal distribution often used for fish abundance data (Lauretta et al. 2016). Using this regression framework gives a high level of control over the nature of the relationships that are fitted by the model, reducing the chance of overfitting to the survey data.

Collinearity between the candidate explanatory variables was assessed using scatterplots and Pearson correlation coefficients. Percentage of gravel and sand were highly correlated, therefore gravel was not included (see Fig. S1a in the Supplement at www.int-res.com/articles/suppl/667p145_supp.pdf). Variable inflation factors were calculated for the remaining variables and these were all under 2, so no further variables were removed. Data exploration also suggested that percentage silt and slope were right-skewed, and so these were square root transformed. The starting point for fitting the models was a linear additive relationship for √percentage silt, percentage sand, depth and √slope in both compo-

nents of the model. To determine if any of these variables could be dropped from the model without reducing the fit, a backwards stepwise selection process was applied by removing each term in turn to see if this resulted in a decrease in the Akaike information criterion (AIC). Once the AIC was at its minimum from backwards selection, a forwards selection process was undertaken for all dropped linear terms, second degree polynomials and 2 way interactions to allow non-linearity to be included. Forwards selection continued until adding another term either increased the AIC or decreased it by less than 2. Models were fitted with the 'countreg' package in R (Zeileis et al. 2008, Zeileis & Kleiber 2018).

Hanging rootograms indicated that the 2 models fitted using a Poisson distribution for the count component had a poorer fit compared to the ones fitted with a negative binomial distribution (Fig. S2) (Kleiber & Zeileis 2016). Further diagnostic plots for the hurdle and zero-inflated negative binomial models were similar, indicating no unexplained spatial or temporal patterns in sandeel abundance (Figs. S3–S6), and therefore model evaluation metrics were obtained for the 2 models using *k*-fold cross validation to determine which had the greatest predictive performance.

The Firth of Forth grab records were divided into 16 blocks of varying areas, which contained approximately the same number of data points. Each cross-validation fold included 4 random blocks. A block design was used to increase the independence between the training and test data compared to simple random partitioning of all data points. The 4-fold cross validation was repeated 125 times, each with different random groups of blocks, giving 500 evaluations. Each of the 500 random partitions were identical for the hurdle and zero-inflated model. Negative log likelihood, AIC and Bayesian information criterion (BIC) for the fitted model were recorded. The predictive performance of the model was assessed based on measures of discrimination, accuracy and calibration for the test data (Norberg et al. 2019). The ability of the model to discriminate between presence and absence was evaluated using the area under the receiver operating curve (AUC), and with the Spearman rank ($\rho$) and Pearson (r) correlation coefficients for abundance. Root mean squared error (RMSE) and average absolute error ($AVE_{error}$) were used as measures of accuracy (Norberg et al. 2019). A simple linear model was fitted between the observed and predicted abundances for the withheld data to evaluate model calibration. A perfectly calibrated model would have an intercept of 0 and a gradient of 1 (Potts & Elith 2006).

### 2.1.3. Deviance explained, variable importance and response curves

As deviances are not easily defined for hurdle or zero-inflated models, the McFadden pseudo-$R^2$ was calculated for the selected model. The McFadden pseudo-$R^2$ is calculated as 1 minus the ratio of log likelihood of the fitted model to the null model (McFadden 1977).

The importance of each variable on the predicted occurrence and count component was estimated using the same protocol implemented by Thuiller et al. (2009), where each variable is permuted randomly and the model fitted to the new data set. If a variable is important to the model prediction, there will be little correlation between the prediction from the original data set and the prediction when that variable has been permuted. The importance is therefore estimated as 1 − r. As this method involves random permutation of the data, there is stochasticity in the result, therefore the procedure was repeated 500 times for each variable.

Response plots for occurrence probability for each variable were created by changing the variable of interest across the range in the Firth of Forth data, and holding the remaining variables at the mean. The occurrence probabilities were calculated using the logistic model fitted with the 'glm' function (R Core Team 2019), to allow standard errors to be plotted. Partial responses were generated for each 2 way variable combination. Partial responses were calculated by predicting the density for all the data points used to fit the model, for each combination of the 2 variables of interest. The density of sandeels was predicted using the final fitted model and is, therefore, the product of the prediction from the 2 components of the model.

### 2.2. Geographic predictions of sandeel distribution

#### 2.2.1. Study regions

The model was used to predict sandeel distribution across UK areas of the Greater North Sea region from Shetland to south of the protected sandbanks off Norfolk, and the shelf sea areas in the north of the Celtic Seas OSPAR region (https://www.ospar.org/convention/the-north-east-atlantic) around the west of the Scotland, Northern Ireland and northern parts of Ireland (Fig. 1). Maps showing the depth, slope and sediment fractions for both study regions are provided in Fig. S7. Given the importance of sediment characteristics to sandeels, the extents of these

model domains were influenced by the availability of PSA records. Areas expected to be exposed rock, based on Downie et al. (2016), Brown et al. (2017) and Geological Survey of Ireland (2018), were masked out of the study region.

### 2.2.2. Environmental variables

Point records of sediment PSA data for the study regions were downloaded from the British Geological Survey (www.bgs.ac.uk/GeoIndex/offshore.htm [accessed January 2019]), Marine Recorder (JNCC 2019) and INFOMAR (Geological Survey of Ireland 2018). Layers of percentage silt, sand, and gravel were interpolated from the point records by inverse distance weighting using the 'gstat' package (Gräler et al. 2016). The layers were produced at a 200 m resolution in Lambert azimuthal equal area projection for Europe (https://spatialreference.org/ref/epsg/etrs89-etrs-laea/). The 3 component sediment layers were adjusted to ensure each cell summed to 100%.

The EMODnet Bathymetry data was again used for depth and slope. These were re-projected to match the resolution and coordinate reference system for the sediment layers. This was a lower resolution than the original source.

Extrapolating beyond the environmental conditions used in the initial fitting of the model can reduce predictive performance. For each of the 4 variables included in the model, the areas of the study regions that are outside the range used to fit the model were mapped to assess the spatial extent where extrapolation may occur. Collinearity shift, where the degree of correlation between explanatory variables in the new location differs from that in the data used to fit the model, can also reduce the predictive power of the model (Feng et al. 2019). Collinearity shift between the grab survey data and the 2 study regions was assessed visually using scatterplot matrices and Pearson correlation coefficients.

### 2.2.3. Model validation

The predictive performance of the sandeel distribution model for the 2 study regions was tested with independent data that were not used in the development of the model. For the North Sea region, sandeel density data were available from sandeel assessment area 4 (SA4) dredge surveys conducted between 2008 and 2019 (Fig. 1a; ICES 2010). These surveys target sandeels; however, the dredge has both lower
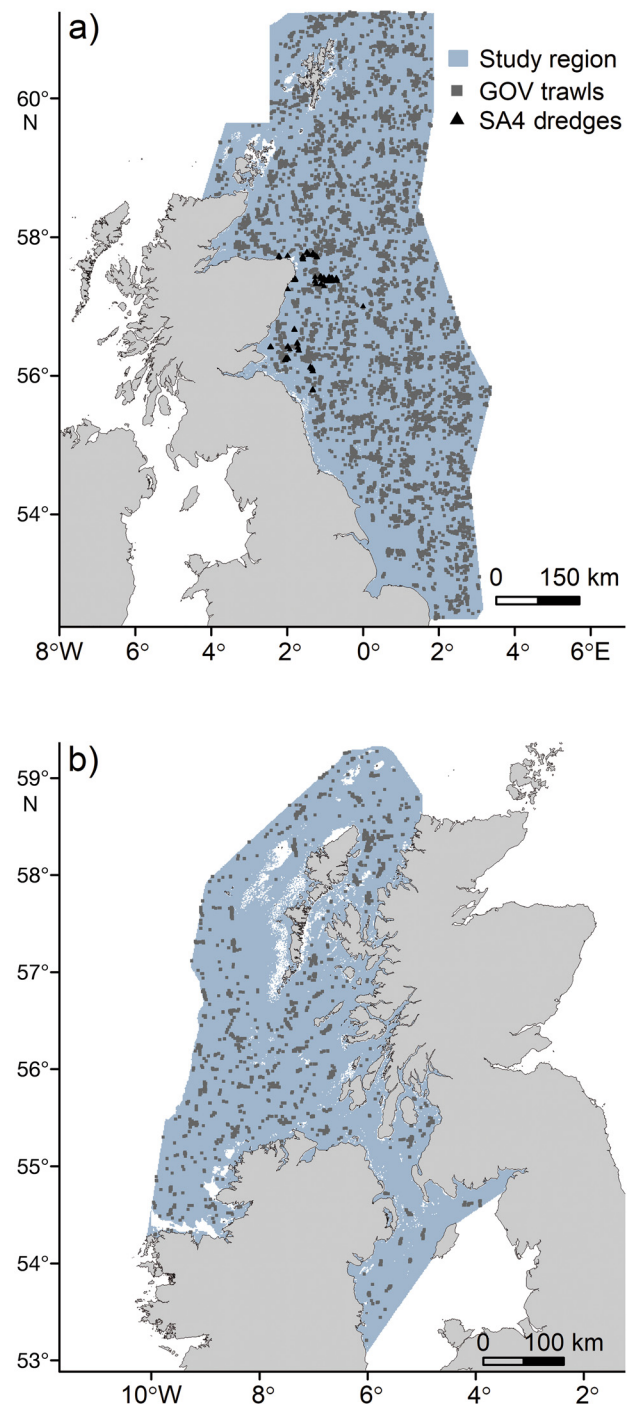


Fig. 1. Study region for parts of the (a) Greater North Sea region and (b) Celtic Seas region and the locations of data from benthic fishing surveys that were used to validate the model. In the North Sea region the Grande Ouverture Verticale (GOV) trawls are from the North Sea International Bottom Trawl Survey. The GOV trawls in the Celtic Seas region include the Scottish West Coast Bottom Trawl survey, the Scottish West Coast Ground-fish survey and the Irish Groundfish survey. The North Sea region also includes dredge surveys for sandeel assessment area 4 (SA4)

efficiency and higher size selectivity than grabs (Holland et al. 2005, Johnsen & Harbitz 2013) and so can only give an indication of relative abundance. The same model discrimination, accuracy and calibration metrics used during cross-validation were calculated for these data. The dredge data were randomly permuted 999 times to allow the statistical significance of AUC, RMSE and model calibration to be determined.

Bottom trawl survey data using Grande Ouverture Verticale (GOV) trawls for both regions were downloaded from the DATRAS database (ICES 2020b). The data included 1972–2020 and 1985–2018, for the North Sea and Celtic Seas regions, respectively (Fig. 1). Sandeels are generally too small to be fully selected in GOV trawl panels and so this gear cannot be used to reliably estimate abundance (Wright et al. 2019); therefore, only model discrimination metrics were calculated for these data. In addition, polygons of the locations of sandeel fishing grounds (Jensen et al. 2011) were mapped onto the predicted occurrence for the North Sea.

## 3. RESULTS

### 3.1. Relationship between sandeels and environmental variables

While the values of the evaluation metrics were similar between the 2 models, they were significantly different (Table 1). Considering negative log likelihood, AIC, RMSE, AUC, r and model calibration, the hurdle model appears to conform better than the zero-inflated model; the reverse is true for BIC and AVE$_{error}$. Given that the hurdle model generally had higher predictive performance, this model was chosen to be used in the subsequent analyses.

The binomial component of the hurdle model explained 30% of the deviance in lesser sandeel *Ammodytes marinus* occurrence. The final hurdle model had a McFadden pseudo-$R^2$ of 0.15. The McFadden pseudo-$R^2$ has values lower than $R^2$ derived from ordinary regression, and a value of 0.15 is estimated to be approximately equivalent to an $R^2$ of 0.35 (Domencich & McFadden 1974, McFadden 1977). In the zero component, the model included quadratic

Table 1. Evaluation metrics (mean ± SD) for the hurdle and zero-inflated model of sandeel abundance based on 125 rounds of 4-fold cross validation and statistical differences based on a paired *t*-test. The negative log-likelihood, Akaike information criterion (AIC), Bayesian information criterion (BIC), Spearman correlation (ρ), Pearson correlation coefficient (r), Area under the receiver operating curve (AUC), residual mean square error (RMSE), average absolute error (AVE$_{error}$) and the intercept and gradient of linear regression between observed and fitted sandeel abundance are shown. NA: not applicable

|                       | Hurdle            | Zero-inflated     | p       |
|-----------------------|-------------------|-------------------|---------|
| Negative log likelihood | 2631.70 ± 227.952 | 2646.98 ± 227.192 | <0.001  |
| AIC                   | 5309.40 ± 455.904 | 5323.97 ± 454.385 | <0.001  |
| BIC                   | 5440.03 ± 455.905 | 5414.84 ± 454.386 | <0.001  |
| Model discrimination  |                   |                   |         |
| AUC                   | 0.83 ± 0.033      | 0.83 ± 0.039      | 0.002   |
| ρ                     | 0.53 ± 0.078      | 0.53 ± 0.078      | 0.03    |
| r                     | 0.42 ± 0.071      | 0.40 ± 0.074      | <0.001  |
| Model accuracy        |                   |                   |         |
| RMSE                  | 4.92 ± 1.466      | 5.10 ± 1.590      | <0.001  |
| AVE$_{error}$         | −0.11 ± 0.578     | −0.02 ± 0.597     | <0.001  |
| Model calibration     |                   |                   |         |
| Intercept             | 0.14 ± 0.344      | 0.28 ± 0.478      | NA      |
| Gradient              | 0.96 ± 0.376      | 0.85 ± 0.378      | NA      |

terms for percentage silt, percentage sand, depth and slope, and the interactions between percentage sand and percentage silt, percentage sand and depth, and depth and slope (Table S1). The count component included percentage silt, percentage sand and quadratic terms for depth and slope (Table S1).

Silt content was the most important variable in determining occurrence (Fig. 2), with sandeels being absent when silt content is >~15% (Figs. 3 & 4). The interaction plots also highlight the strong influence of silt content, as density is predicted to be zero after the 15% threshold, regardless of the value of any other variable (Fig. S8). Sand content and depth had roughly equal influence on sandeel occurrence and abundance, both contributing around 30% to the count component (Fig. 2). For depth and sand, the presence and size of the peak sandeel density varies depending on the other environmental variables (Fig. S8). The response curves indicate that, under average conditions, occurrence peaks at a sand content of around 70% (Fig. 3), but that the greatest sandeel density occurs at a higher sand content as the silt content decreases (Fig. 4). The model predicts that sandeels are absent at depths >60 m and <20 m, with occurrence highest at 40 m (Fig. 3). As the final model includes interaction terms for depth and sand, and depth and slope, the shape of the depth response changes with these variables. The depth with the greatest density of sandeels is positively related to slope, but negatively related to sand content (Fig. 4). Slope is also important
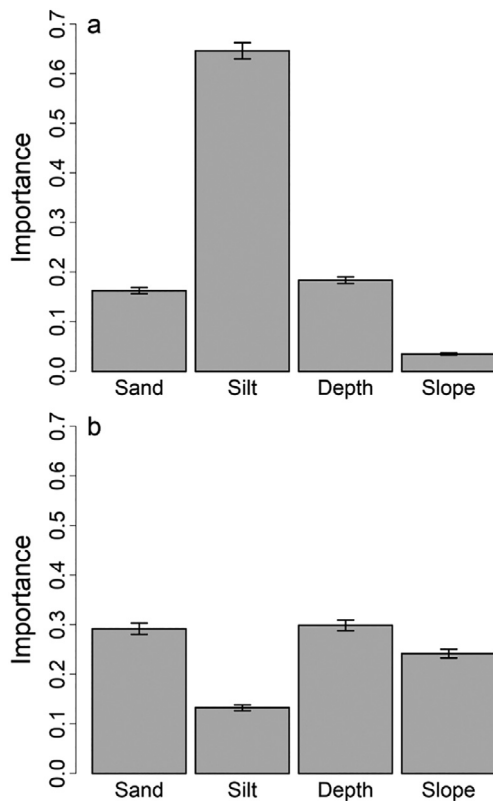
Fig. 2. Relative importance (mean ± SD, n = 500) of the explanatory variables for sandeel abundance from the (a) occurrence probability and (b) count component of the hurdle model

in determining sandeel abundance (Fig. 2), with sandeels preferring shallower slopes (Figs. 3 & 4).

## 3.2. Predicted sandeel distribution and model validation

The model predicts that sandeels are absent from the majority of the area of the study regions. In the North Sea, the highest predicted sandeel occurrence and densities are over Dogger Bank and North Norfolk sandbanks in the southern North Sea (Fig. 5). The probability of sandeel presence was significantly higher within the sandeel grounds than outside (Mann-Whitney test = $1.5 \times 10^{12}$, p < 0.001), although there were fishing grounds in depths ≥60 m where the model predicts sandeel to be absent. In the Celtic Seas region, the model predicted high likelihood of sandeel presence and densities east of Dublin, the north east coast of Donegal, north and west of Islay and to the north of Lewis (Fig. 6).

Collinearity shift maybe occurring during the regional predictions, as both study regions have higher

correlation between sand and silt content than observed in the grab data (Fig. S1). There are areas in both study regions with a higher silt content than was observed in the Firth of Forth survey data used to fit the model (Fig. S9). Large parts of the 2 regions are outside the depth range of the grab survey data, particularly towards the north of the North Sea region and the north-west of the Celtic Seas region (Fig. S9).

AUC values ranged between 0 and 1, with values over 0.7 indicating adequate discrimination ability (Pearce & Ferrier 2000). The AUC values for occurrence predictions were >0.7 for all 3 validation data sets and close to 0.8 for the GOV trawl data (Table 2). The ρ and r were all positive and significant except r for the Celtic Seas GOV data (Table 2). The model is therefore able to discriminate areas where sandeels are likely to be present and reflects relative differences in density. The RMSE for the North Sea SA4 dredge data indicates that the model's accuracy at predicting density of sandeels is significantly better than random; however, the positive $AVE_{error}$ and gradient close to zero for model calibration shows that the model over-predicts density in the dredge (Table 2). This is as expected given the relative efficiency of grabs and dredges.

## 4. DISCUSSION

The model captured the general pattern and scale of lesser sandeel *Ammodytes marinus* distribution expected from the location of large fishing grounds reported in the central North Sea (Jensen et al. 2011) and the small grounds to the north (Wright 1996) and west of Scotland. Several of the predicted habitat patches are in areas where seabirds and marine mammals are thought to congregate to feed on sandeels (Wright & Begg 1997, Macleod et al. 2004, Anderwald et al. 2012). These included aggregations around Jura, Islay and Colonsay that are within the Inner Hebrides and the Minches Special Area of Conservation (SAC) and large areas in the Southern North Sea SAC, both of which are designated for harbour porpoise *Phocoena phocoena*. The model also predicts an area of high density of sandeels in Northeast Lewis Marine Protected Area, which was designated for sandeels in 2020 and previously fished.

A hurdle model was generally the best fitting model and had a higher predictive performance. Potts & Elith (2006) similarly found that the hurdle model out-performed alternatives. These predictions were also consistent with past evidence on the relationship between presence and percentage silt
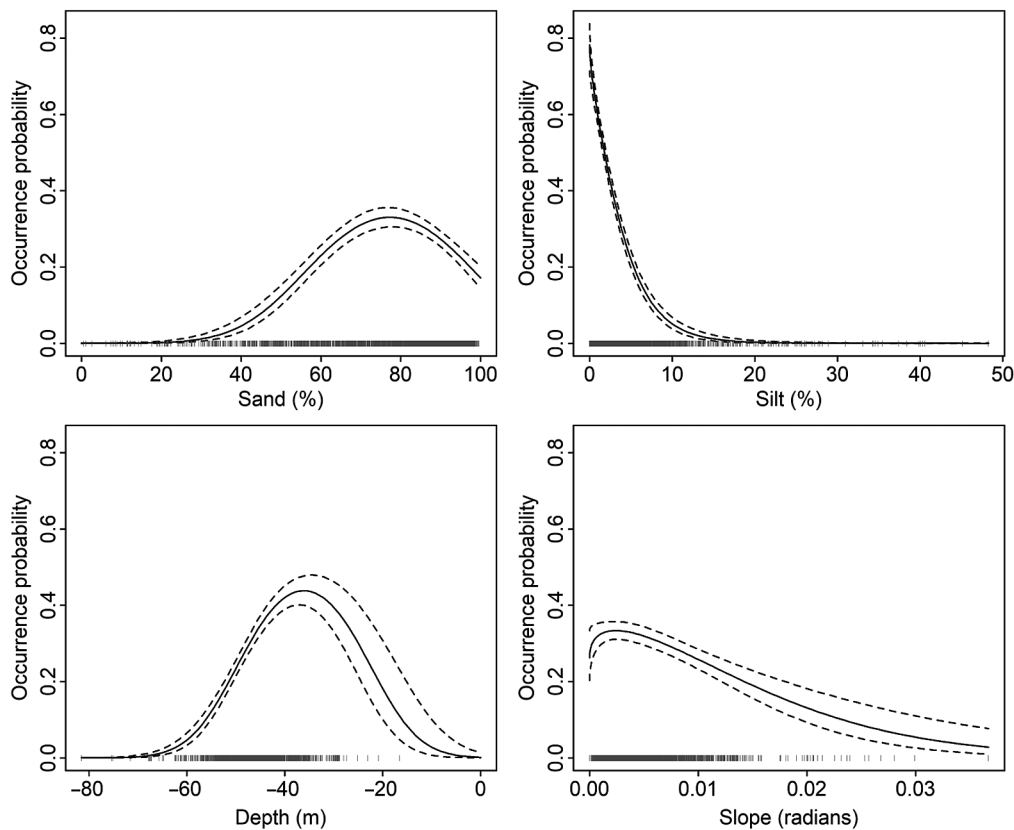
Fig. 3. Response curves for the occurrence probability of sandeel against the 4 explanatory variables (mean ± SE) from the hurdle model. All other variables were held constant at their mean in the grab data

(Wright et al. 2000, Holland et al. 2005, Tien et al. 2017). The high explained deviance in the occurrence model was probably related to differences in the spatial resolution of sample data corresponding to fish abundance and sediment particle size, as this would lead to less spatial error in the training data compared to most marine SDMs that compare catches from the mid-points of trawls with interpolated physical data of even coarser resolution (González-Irusta & Wright 2016). Although using the localised grab data to parameterise the relationships between sandeel abundance and environmental variables possibly led to high explanatory power, it resulted in extrapolation for silt and depth when making predictions at the regional scale (Fig. S9). Given the strength of the evidence that sandeels are absent when silt is greater than 15%, model predictions in areas with higher silt content should be reasonable. A significant proportion of the 2 regions are outside the depth range of the Firth of Forth survey data (Fig. S9), and predictions are less certain for these areas. Despite the presence of extrapolation and collinearity shift, both of which have been shown

to reduce the predictive performance of models (Feng et al. 2019), the evaluation of the model using independent data demonstrated that it had significant discrimination ability in both regions.

Silt fraction was the most important variable in predicting occurrence, which is consistent with the avoidance of this sediment type by sandeels in choice experiments (Pinto et al. 1984, Wright et al. 2000). The modelled presence response to silt percentage was similar to past studies where occurrence declined from 0–10% silt (Wright et al. 2000, Tien et al. 2017). Permeability and the characteristics of the boundary layer affect the rate of water percolation through sediments (Huettel et al. 1996). As permeability of sediments is a function of grain size and porosity (Chilingar 1964), silt-rich sediments tend to have small interstitial water volume and low rates of water exchange. Given that sandeels need to be able to draw water from the surface through pores in the sediment (Behrens et al. 2007), which could be clogged by even small amounts of silt (reducing permeability), there is a clear physiological explanation for the avoidance of silt-rich sediment.
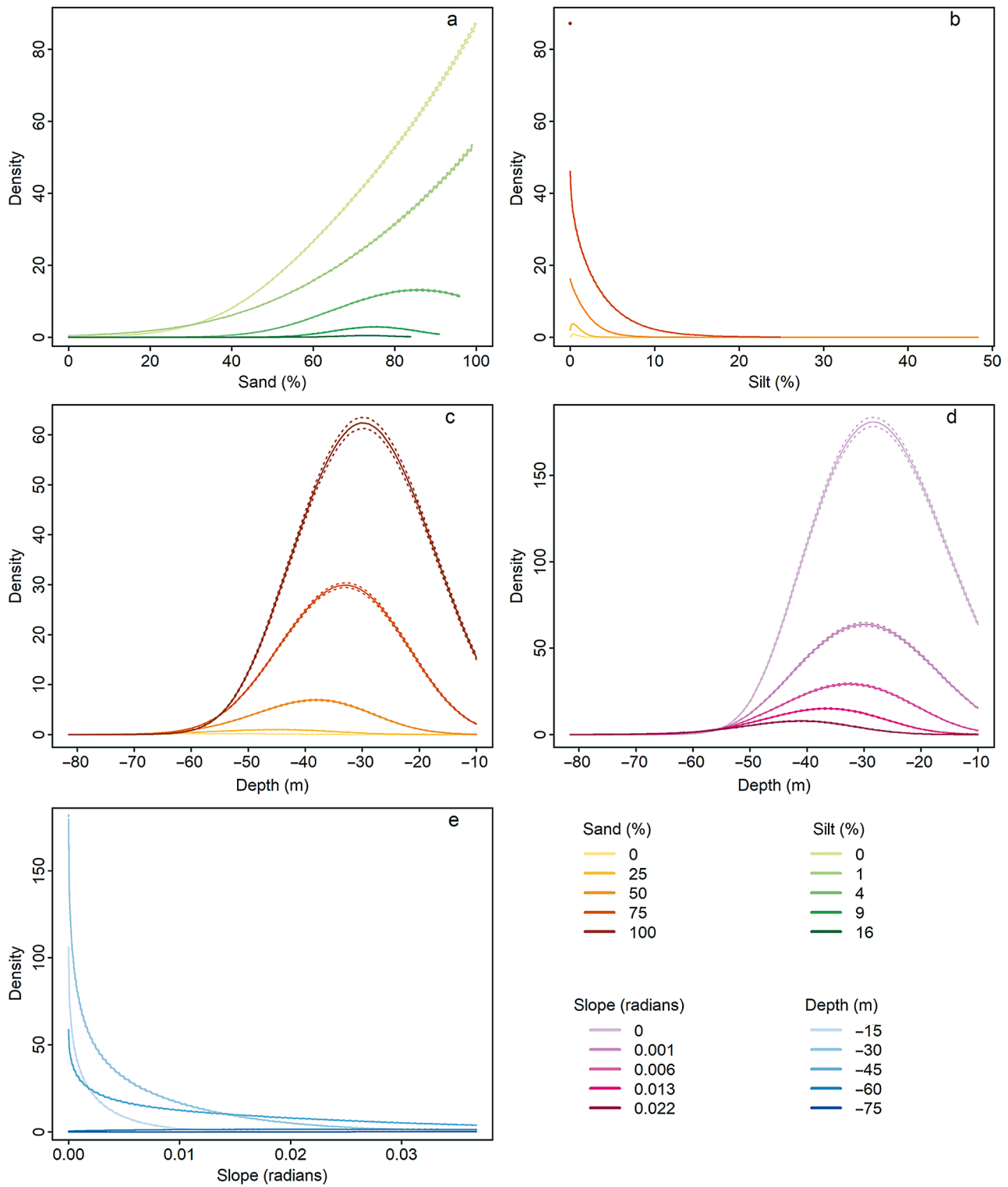
Fig. 4. Interaction plots showing the partial response curves for the predicted density of sandeels (no. m$^{-2}$) for each explanatory variable and key interactions included in the final model; (a) the response to sand content at different levels of silt, (b) the response to silt content for different levels of sand, (c) the response to depth at different levels of sand, (d) the response to depth at different slopes and (e) the response to slope at different depths. For each combination of variables being tested, predictions were made for all Firth of Forth survey records; mean ± SE are displayed
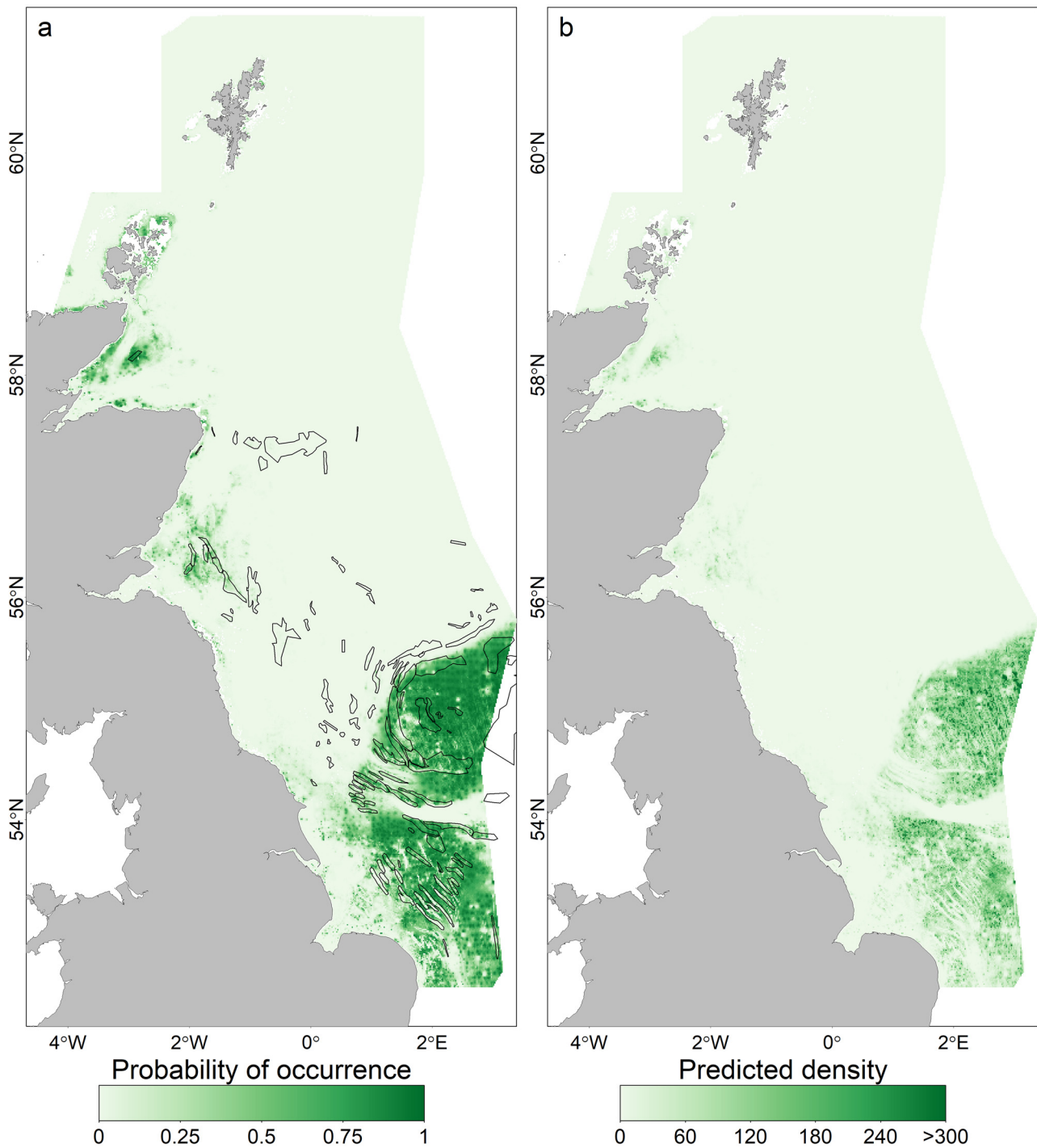
Fig. 5. (a) Predicted probability of occurrence of sandeel with known sandeel fishing grounds (Jensen et al. 2011), and (b) predicted density of sandeels (no. m[−2]) for the North Sea region. The upper limit of the density scale is greater than 99.9[th] percentile

The positive effect of sand fraction on density is consistent with early reports that sandeels tend to inhabit areas of 'clean sand' (Macer 1966, Reay 1970, Meyer et al. 1979) and the seeming importance of sand to sandeel presence in the Firth of Forth grounds

(Greenstreet et al. 2010). Similarly, Tien et al. (2017) found a significant positive relationship with medium-coarse sand and *A. marinus* density. This might be related to its suitability for burrowing. In support, although experiments indicate that sandeels are
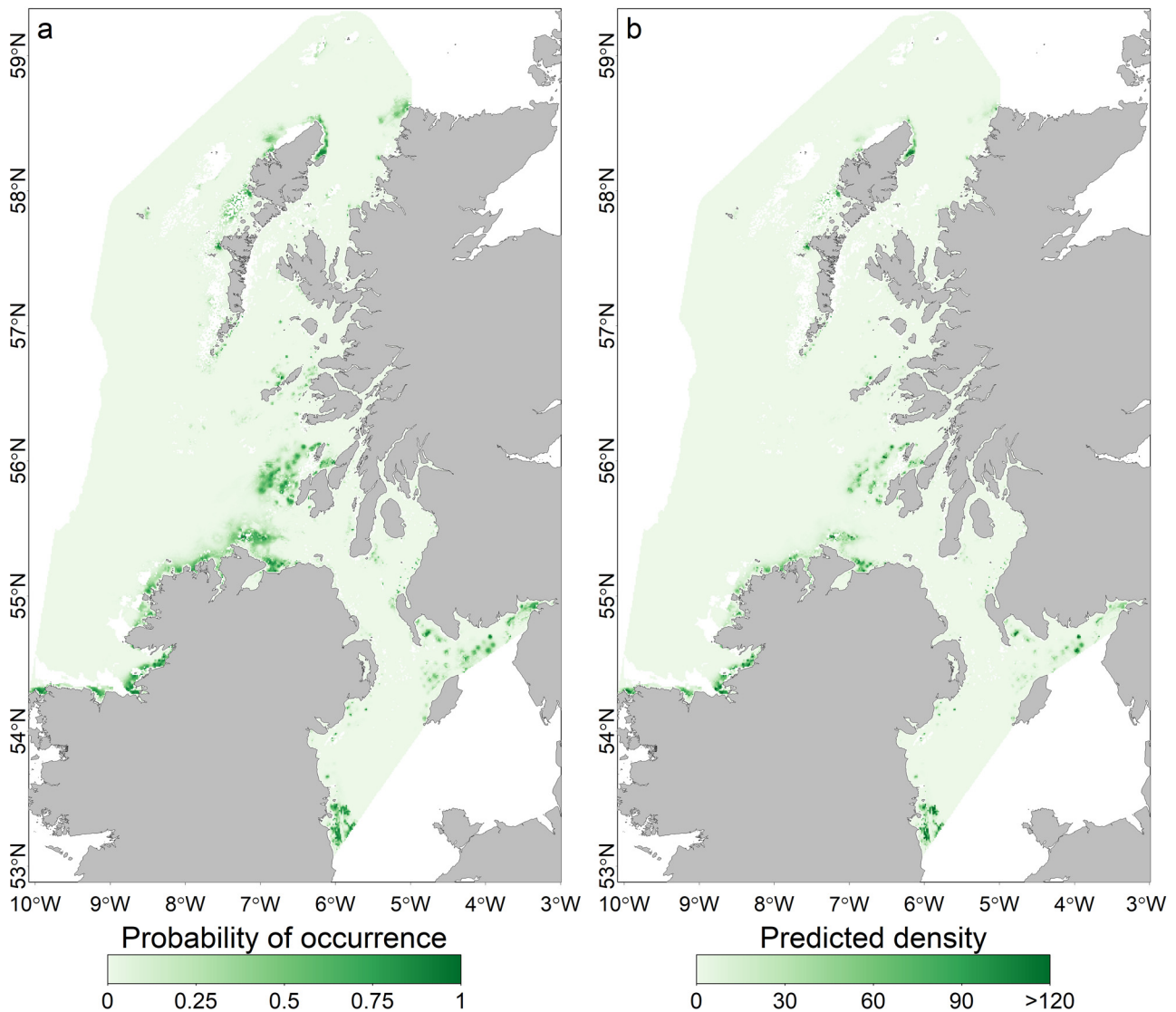
Fig. 6. (a) Predicted probability of occurrence of sandeel and (b) the predicted density of sandeel (no. m$^{-2}$) for the Celtic Seas region. The upper limit of the density scale is greater than 99.9$^{th}$ percentile

capable of penetrating even large gravel ≥16 mm (Pinto et al. 1984), they prefer sediments of low shear strength (Endo et al. 2019), and numbers decline when offered sediments with a high percentage of gravel (Wright et al. 2000). Estimates of maximum median grain size occupied by *A. japonicus*, which is linked to shear stress (Endo et al. 2019), were also similar to that reported for *A. marinus* (Wright et al. 2000). Observations on the time taken to enter the gravel-rich sediment suggests that this preference could also be related to the slow speed of penetration (J. Verspoor & P. J. Wright unpubl. data). Hence, while sandeels do occasionally occupy high gravel content sediment (Holland et al. 2005), the effort and potential risk from predators during entry may ex-

plain why sand is preferred. Consequently, if coarser sediment types were under represented in the grab data, it is unlikely to reduce the reliability of the fitted model.

Within the limitations of preferred sediment, slope was a significant explanatory variable, especially for predicting density. The avoidance of strongly sloping habitat is consistent with reports of buried sandeels on the tops of sand banks (Jensen 2001, Engelhard et al. 2008). The model prediction indicates that high sloping sites are avoided, and this is supported by acoustic observations that indicate sandeels emerge and then move off the tops of the banks to the side of banks to feed on plankton (Jensen 2001, van der Kooij et al. 2008).

Table 2. Validation of the hurdle model of sandeel distribution and data from North Sea sandeel assessment area 4 (SA4) dredge surveys and Grande Ouverture Verticale (GOV) surveys. Spearman correlation (ρ), Pearson correlation coefficient (r), Area under the receiver operating curve (AUC), residual mean square error (RMSE), average absolute error (AVE$_{error}$) and the intercept and gradient of linear regression between observed and fitted sandeel density are shown. The statistical significance of AUC, RMSE and model calibration results were determined using random permutations of the data (n = 999). NA: not applicable. Model accuracy and model calibration metrics were not calculated for the GOV surveys as this gear cannot be used to reliably estimate sandeel abundance

|  | North Sea SA4 dredge | | GOV | | | |
|  | | | North Sea | | Celtic Seas | |
|  | Estimate | p | Estimate | p | Estimate | p |
| --- | --- | --- | --- | --- | --- | --- |
| Model discrimination | | | | | | |
| AUC | 0.71 | <0.001 | 0.82 | <0.001 | 0.79 | <0.001 |
| ρ | 0.41 | <0.001 | 0.38 | 0 | 0.32 | <0.001 |
| r | 0.44 | <0.001 | 0.03 | <0.001 | 0.05 | 0.056 |
| Model accuracy | | | | | | |
| RMSE | 32.5 | <0.001 | | | | |
| AVE$_{error}$ | 17.0 | NA | | | | |
| Model calibration | | | | | | |
| Intercept | 0.07 | <0.001 | | | | |
| Gradient | 0.003 | <0.001 | | | | |

The preferred depth range of 30–50 m for *A. marinus* is consistent with the recorded depth distribution of most grounds in the North Sea where sandeel fisheries operate (Macer 1966, Jensen et al. 2011) and falls within the range reported for Shetland (Wright et al. 2000). However, *A. marinus* can occur between depths of 12 and 120 m (Wright et al. 1998, Tien et al. 2017) and some important grounds in the northern North Sea are at depths of around 70 m, such as the Turbot Bank off north-east Scotland (Jensen et al. 2011). That similar areas of depth were not in the training data set explains why this and other deep grounds were not predicted by the model — demonstrating a weakness in parameterising the model from one area, as the derived depth relationship may simply reflect local conditions that are a proxy for some other physical influence. For example, the general trend towards finer sediments and declining water velocity with depth may be relevant, as most aggregations are found in areas of sand ripples where the net residual flow is likely to be >0.5 m s⁻¹ (Wright et al. 1998, 2000). Clearly, further work is needed to consider local variation in sandeel abundance in relation to residual flow rate, and how this interacts with depth.

When comparing the distribution of sandeel habitat predicted by the model with current and historic fishing grounds (Wright 1996, Jensen et al. 2011), there appears to be few unexploited areas. In gen- eral, large expanses of suitable sediment occur in tidally formed sand banks and ridges, which are largely accessible to fishing. However, there are some areas where patches of sand are mixed in among rocky areas, leading to some very small sandeel fishing grounds around the Shetland Isles (Wright 1996), and even smaller patches only accessible to foraging predators (Monaghan et al. 1996). As habitat for burying is likely to be a physical constraint on carrying capacity for sandeels, the extent of predicted habitat from this study might explain regional differences in maximum fishery landings. For example, the Scottish west coast only comprised small patches of habitat, with landings from the fishing grounds in this region only peaking at 24.4 kt. Similarly, the peak landings from grounds around the Shetland Isles peaked at 52 kt. In contrast, peak landings in SA4 and SA1 were 147 and 605 kt, respectively (ICES 2020a). Habitat patchiness inferred from the model, when combined with information on larval transport, can also help explain the high levels of connectivity in *A. marinus* across the central North Sea, but not with the Firth of Forth Banks (Wright et al. 2019) given the low density of habitat patches between them. The small aggregations around Shetland and the large interpatch distances to grounds further south around Orkney may also help explain the limited exchange between these areas (Wright 1996, Gibb et al. 2017). Due to high rates of larval transport, larvae from the small grounds north of the Outer Hebrides tend to mix with those from Orkney (Proctor et al. 1998), and occasional transport from Orkney to Shetland appears important to recruitment in that area (Wright 1996, Gibb et al. 2017). The relative hydrographic isolation of banks off the Firth of Forth and those further south results in differences in early environmental exposure that are reflected in different recruitment trends (ICES 2020a) and the annual breeding success of kittiwakes *Rissa tridatyla* at colonies utilising these regions (Olin et al. 2020).

A number of fish and mammalian predators can catch sandeels in the sediment (Hobson 1986), and certain diving seabirds have been found to aggregate over such areas of preferred sandeel habitat (Monaghan et al. 1996). Sandeel distribution can be

an important influence on the distribution of such predators at sea, as in the case with common guillemots *Uria aalge* during the breeding season (Wright & Begg 1997), and foraging for sandeels may explain why some predators aggregate over patches of sand and avoid mud (Aarts et al. 2008). As a consequence, studies have used evidence of sand as a predictor of sandeel availability to predators (Wanless et al. 1998, Macleod et al. 2004, Anderwald et al. 2012). However, it is likely that such studies will have overestimated the extent of sandeel habitat. For example, while the present model did predict habitat to the west of Mull as suggested by Macleod et al. (2004), the extent of these patches were smaller. Nevertheless, the habitat patches predicted by the present study were consistent with the June sightings of Minke whale *Balaenoptera acutorostrata* in that study, supporting the view that this species may congregate on sandeel grounds at certain times of year. Due to technological advances in GPS tagging, it is now possible to analyse habitat use by marine top predators at a relatively fine spatial scale (Aarts et al. 2008, Wakefield et al. 2017) and identify where foraging behaviour occurs (Bennison et al. 2018). Future comparisons between the sandeel distribution model and predator foraging sites could help identify the key sandeel areas used by predators. This capability illustrates that the distribution model from this study could help refine inferences of sandeel presence and density for spatial predator–prey studies.

The present study demonstrates the effective development of a predictive distribution model for an ecologically and commercially important fish species with a dependence on a specific habitat. This model will alert marine planners to potential areas where possible anthropogenic impacts on sandeels should be considered, such as sand extraction, marine renewable developments and fishing with gears that cause significant benthic disturbance, such as dredges (Eleftheriou & Robertson 1992). It will also improve the basis for spatial comparisons between predators and sandeels and identify areas where further direct observations are needed. The predicted distribution map for sandeels could then be updated to incorporate new data on sandeel occurrence or environmental variables, and this would ensure that the best available information is being used for the management of the species at a regional scale.

LITERATURE CITED

Aarts G, MacKenzie M, McConnell B, Fedak M, Matthiopoulos J (2008) Estimating space-use and habitat preference from wildlife telemetry data. Ecography 31: 140−160

Anderwald P, Evans PGH, Dyer R, Dale A, Wright PJ, Hoelzel AR (2012) Spatial scale and environmental determinants in minke whale habitat use and foraging. Mar Ecol Prog Ser 450:259−274

Baker MR, Matta ME, Beaulieu M, Paris N and others (2019) Intra-seasonal and inter-annual patterns in the demographics of sand lance and response to environmental drivers in the North Pacific. Mar Ecol Prog Ser 617-618: 221−244

Behrens JW, Stahl HJ, Steffensen JF, Glud RN (2007) Oxygen dynamics around buried lesser sandeels *Ammodytes tobianus* (Linnaeus 1785): mode of ventilation and oxygen requirements. J Exp Biol 210:1006−1014

Bell DM, Schlaepfer DR (2016) On the dangers of model complexity without ecological justification in species distribution modelling. Ecol Modell 330:50−59

Bennison A, Bearhop S, Bodey TW, Votier SC and others (2018) Search and foraging behaviors from movement data: a comparison of methods. Ecol Evol 8:13−24

Brown LS, Green SL, Stewart HA, Diesing M, Downie AL, Cooper R, Lillis H (2017) Semi-automated mapping of rock in the Irish Sea, Minches, western Scotland and Scottish continental shelf. JNCC Report No. 609. Joint Nature Conservation Committee, Peterborough

Carroll MJ, Bolton M, Owen E, Anderson GQA, Mackley EK, Dunn EK, Furness RW (2017) Kittiwake breeding success in the southern North Sea correlates with prior sandeel fishing mortality. Aquat Conserv 27:1164−1175

Chilingar GV (1964) Relationship between porosity, permeability and grain-size distribution of sands and sandstones. Dev Sedimentol 1:71−75

Domencich TA, McFadden D (1974) Urban travel demand: a behavioural analysis. A Charles river Associates research study. North Holland Publishing Company, Amsterdam

Downie AL, Dove D, Westhead RK, Diesing M, Cooper R (2016) Semi-automated mapping of rock in the North Sea. JNCC Report No. 592. Joint Nature Conservation Committee, Peterborough

Eleftheriou A, Robertson MR (1992) The effects of experimental scallop dredging on the fauna and physical environment of a shallow sandy community. Neth J Sea Res 30:289−299

Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst 40:677−697

Endo A, Iwasaki N, Shibata JY, Tomiyama T, Sakai Y (2019) The burrowing sand lance *Ammodytes japonicus* (Teleostei, Ammodytidae) prefers benthic sediments of low shear strength. J Ethol 37:213−219

Engelhard GH, van der Kooij J, Bell ED, Pinnegar JK, Blanchard JL, Mackinson S, Righton DA (2008) Fishing mortality versus natural predation on diurnally migrating

sandeels *Ammodytes marinus.* Mar Ecol Prog Ser 369: 213–227

Feng X, Park DS, Liang Y, Pandey R, Papeş M (2019) Collinearity in ecological niche modeling: confusions and challenges. Ecol Evol 9:10365–10376

Freeman S, Mackinson S, Flatt R (2004) Diel patterns in the habitat utilisation of sandeels revealed using integrated acoustic surveys. J Exp Mar Biol Ecol 305:141–154

Geological Survey of Ireland (2018) Integrated mapping for the sustainable development of Ireland's marine resource (INFOMAR): shapefiles. https://jetstream.gsi.ie/iwdds/delivery/INFOMAR_VIEWER/index.html (accessed 9 Oct 2019)

Gibb FM, Régnier T, Donald K, Wright PJ (2017) Connectivity in the early life history of sandeel inferred from otolith microchemistry. J Sea Res 119:8–16

González-Irusta JM, Wright PJ (2016) Spawning grounds of haddock (*Melanogrammus aeglefinus*) in the North Sea and West of Scotland. Fish Res 183:180–191

Gräler B, Pebesma EJ, Heuvelink G (2016) Spatio-temporal interpolation using gstat. R J 8:204–218

Greene HG, Baker M, Aschoff J (2020) A dynamic bedforms habitat for the forage fish Pacific sand lance, San Juan Islands, WA, United States. In: Harris PT, Baker E (eds) Seafloor geomorphology as benthic habitat. Elsevier, Amsterdam, p 267–279

Greenstreet SPR, Armstrong E, Mosegaard H, Jensen H and others (2006) Variation in the abundance of sandeels *Ammodytes marinus* off southeast Scotland: an evaluation of area-closure fisheries management and stock abundance assessment methods. ICES J Mar Sci 63:1530–1550

Greenstreet SPR, Holland GJ, Guirey EJ, Armstrong E, Fraser HM, Gibb IM (2010) Combining hydroacoustic seabed survey and grab sampling techniques to assess 'local' sandeel population abundance. ICES J Mar Sci 67: 971–984

Gregr EJ, Palacios DM, Thompson A, Chan KMA (2019) Why less complexity produces better forecasts: an independent data evaluation of kelp habitat models. Ecography 42:428–443

Halley DJ, Harrison N, Webb A, Thompson DR (1995) Seasonal and geographical variations in the diet of common guillemots *Uria aalge* off western Scotland. Seabird 17: 12–20

Harwood J, Croxall JP (1988) The assessment of competition between seals and commercial fisheries in the North Sea and the Antarctic. Mar Mamm Sci 4:13–33

Herr H, Fock HO, Siebert U (2009) Spatio-temporal associations between harbour porpoise *Phocoena phocoena* and specific fisheries in the German Bight. Biol Conserv 142: 2962–2972

Hijmans RJ (2019) raster: geographic data analysis and modeling. R package version 2.8-19. https://cran.r-project.org/web/packages/raster/index.html

Hill SL, Hinke J, Bertrand S, Fritz L and others (2020) Reference points for predators will progress ecosystem-based management of fisheries. Fish Fish 21:368–378

Hobson ES (1986) Predation on the Pacific sand lance, *Ammodytes hexapterus* (Pisces: Ammodytidae), during the transition between day and night in southeastern Alaska. Copeia 1986:223–226

Høines ÅS, Bergstad OA (2001) Density of wintering sand eel in the sand recorded by grab catches. Fish Res 49: 295–301

Holland GJ, Greenstreet SPR, Gibb IM, Fraser HM, Robert-

son MR (2005) Identifying sandeel *Ammodytes marinus* sediment habitat preferences in the marine environment. Mar Ecol Prog Ser 303:269–282

Huettel M, Ziebis W, Forster S (1996) Flow-induced uptake of particulate matter in permeable sediments. Limnol Oceanogr 41:309–322

ICES (2010) Report of the benchmark workshop on sandeel (WKSAN). ICES, Copenhagen

ICES (2020a) Sandeel (*Ammodytes* spp.) in Divisions 4.a and 4.b, Sandeel Area 4 (northern and central North Sea). In: Report of the ICES Advisory Committee, 2020. ICES Advice 2020, san.sa.4, https://doi.org/10.17895/ices.advice.5763

ICES (2020b) ICES database of trawl surveys (DATRAS). https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx (accessed 29 September 2020)

Jarnevich CS, Stohlgren TJ, Kumar S, Morisette JT, Holcombe TR (2015) Caveats for correlative species distribution modeling. Ecol Inform 29:6–15

Jensen H (2001) Settlement dynamics in the lesser sandeel *Ammodytes marinus* in the North Sea. PhD dissertation, University of Aberdeen

Jensen H, Rindorf A, Wright PJ, Mosegaard H (2011) Inferring the location and scale of mixing between habitat areas of lesser sandeel through information from the fishery. ICES J Mar Sci 68:43–51

JNCC (Joint Nature Conservation Committee) (2019) Marine Recorder Public UK snapshot—v20190212. https://hub.jncc.gov.uk/assets/4df6ab95-8d06-44cf-b4f8-e1c3db68174a (accessed 29 Apr 2019)

Johnsen E, Harbitz A (2013) Small-scale spatial structuring of burrowed sandeels and the catching properties of the dredge. ICES J Mar Sci 70:379–386

Kleiber C, Zeileis A (2016) Visualizing count data regressions using rootograms. Am Stat 70:296–303

Lauretta MV, Walter JF, Christman MC (2016) Some considerations for CPUE standardization; variance estimation and distributional considerations. Col Vol Sci Pap ICCAT 72:2304–2312

Macer CT (1966) Sand eels (Ammodytidae) in the southwestern North Sea; their biology and fishery. Fishery Investigations Series 2, Vol 24. Great Britain Ministry of Agriculture, Fisheries and Food, London, p 1–55

Mackinson S, van der Kooij J (2006) Perceptions of fish distribution, abundance and behaviour: observations revealed by alternative survey strategies made by scientific and fishing vessels. Fish Res 81:306–315

Macleod K, Fairbairns R, Gill A, Fairbairns B, Gordon J, Blair-Myers C, Parsons ECM (2004) Seasonal distribution of minke whales *Balaenoptera acutorostrata* in relation to physiography and prey off the Isle of Mull, Scotland. Mar Ecol Prog Ser 277:263–274

McFadden D (1977) Quantitative methods for analysing travel behaviour of individuals: some recent developments. Cowles Foundation Discussion Papers, No. 474. Cowles Foundation for Research in Economics, Yale University, New Haven, CT

Melo-Merino SM, Reyes-Bonilla H, Lira-Noriega A (2020) Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. Ecol Modell 415:108837

Merow C, Smith MJ, Edwards TC Jr, Guisan A and others (2014) What do we gain from simplicity versus complexity in species distribution models? Ecography 37: 1267–1281

Meyer CT, Cooper RA, Langstone RW (1979) Relative abundance, behaviour and food habits of the American sand lance, *Ammodytes americanus*, from Gulf of Maine. Fish Bull 77:243–254

Monaghan P (1992) Seabirds and sandeels: the conflict between exploitation and conservation in the northern North Sea. Biodivers Conserv 1:98–111

Monaghan P, Wright PJ, Bailey MC, Uttley JD, Walton P (1996) The influence of changes in food abundance on diving and surface feeding seabirds. In: Montevecchi WD (ed) Studies of high-latitude seabirds, Vol 4: trophic relationships and energetics of endotherms in cold ocean systems. Canadian Wildlife Service Occasional Paper No. 91. Ocean Sciences Center, Memorial University of Newfoundland, St. John's, p 10–19

Norberg A, Abrego N, Blanchet FG, Adler FR and others (2019) A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecol Monogr 89:e01370

Olin AB, Banas NS, Wright PJ, Heath MR, Nager RG (2020) Spatial synchrony of breeding success in the blacklegged kittiwake *Rissa tridactyla* reflects the spatial dynamics of its sandeel prey. Mar Ecol Prog Ser 638:177–190

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecol Modell 133:225–245

Pinto JM, Pearson WH, Anderson JW (1984) Sediment preferences and oil contamination in the Pacific sand lance *Ammodytes hexapterus.* Mar Biol 83:193–204

Potts JM, Elith J (2006) Comparing species abundance models. Ecol Modell 199:153–163

Proctor R, Wright PJ, Everitt A (1998) Modelling the transport of larval sandeels on the north west European shelf. Fish Oceanogr 7:347–354

Qiao H, Feng X, Escobar LE, Peterson AT, Soberón J, Zhu G, Papeş M (2019) An evaluation of transferability of ecological niche models. Ecography 42:521–534

R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Reay P (1970) Synopsis of biological data on North Atlantic sandeels of the genus *Ammodytes* (*A. tobianus*, *A. dubius*, *A. americanus* and *A. marinus*). FAO, Rome

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

Santos MB, Pierce GJ, Learmonth JA, Reid RJ and others (2004) Variability in the diet of harbour porpoise (*Phocoena phocoena*) in Scottish Waters 1992–2003. Mar Mamm Sci 20:1–27

Sherman K, Jones C, Sullivan L, Smith W, Berrien P, Ejsymont L (1981) Congruent shifts in sand eel abundance in western and eastern North Atlantic ecosystems. Nature 291:486–489

Sparholt H (1990) An estimate of the total biomass of fish in the North Sea. ICES J Mar Sci 46:200–210

Temming A, Götz S, Mergardt N, Ehrich S (2004) Predation of whiting and haddock on sandeel: aggregative

response, competition and diel periodicity. J Fish Biol 64: 1351–1372

Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD — a platform for ensemble forecasting of species distributions. Ecography 32:369–373

Tien NSH, Craeymeersch J, van Damme C, Couperus AS, Adema J, Tulp I (2017) Burrow distribution of three sandeel species relates to beam trawl fishing, sediment composition and water velocity, in Dutch coastal waters. J Sea Res 127:194–202

van der Kooij J, Scott BE, Mackinson S (2008) The effects of environmental factors on daytime sandeel distribution and abundance on the Dogger Bank. J Sea Res 60:201–209

Wakefield ED, Owen E, Baer J, Carroll MJ and others (2017) Breeding density, fine-scale tracking, and large-scale modeling reveal the regional distribution of four seabird species. Ecol Appl 27:2074–2091

Wanless S, Harris MP, Greenstreet SPR (1998) Summer sandeel consumption by seabirds breeding in the Firth of Forth, south-east Scotland. ICES J Mar Sci 55:1141–1151

Wanless S, Harris MP, Newell MA, Speakman JR, Daunt F (2018) Community-wide decline in the occurrence of lesser sandeels *Ammodytes marinus* in seabird chick diets at a North Sea colony. Mar Ecol Prog Ser 600:193–206

Wilson LJ, Hammond PS (2019) The diet of harbour and grey seals around Britain: examining the role of prey as a potential cause of harbour seal declines. Aquat Conserv 29:71–85

Wright PJ (1996) Is there a conflict between sandeel fisheries and seabirds? A case study at Shetland. In: Greenstree SPR, Tasker ML (eds) Aquatic predators and their prey. Fishing News Books, Blackwell Science, Oxford, p 154–165

Wright PJ, Begg GS (1997) A spatial comparison of common guillemots and sandeels in Scottish waters. ICES J Mar Sci 54:578–592

Wright PJ, Pedersen SA, Donald L, Anderson C, Lewy P, Proctor R (1998) The influence of physical factors on distribution of lesser sandeel, *Ammodytes marinus*, and its relevance to fishing pressure in the North Sea. ICES CM 1998/AA:3:1–9

Wright PJ, Jensen H, Tuck ID (2000) The influence of sediment type on the distribution of the lesser sandeel, *Ammodytes marinus.* J Sea Res 44:243–256

Wright PJ, Christensen A, Régnier T, Rindorf A, van Deurs M (2019) Integrating the scale of population processes into fisheries management, as illustrated in the sandeel, *Ammodytes marinus.* ICES J Mar Sci 76:1453–1463

Yates KL, Bouchet PJ, Caley MJ, Mengersen K and others (2018) Outstanding challenges in the transferability of ecological models. Trends Ecol Evol 33:790–802

Zeileis A, Kleiber C (2018) countreg: count data regression. R package version 0.2-1. https://r-forge.r-project.org/R/?group_id=522

Zeileis A, Kleiber C, Jackman S (2008) Regression models for count data in R. J Stat Softw 27:1–25

Zuur A, Ieno EN, Walker N, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer Science + Business Media, New York, NY