

# Spatial predictive distribution modelling of the kelp species *Laminaria hyperborea*

Trine Bekkby, Eli Rinde, Lars Erikstad, and Vegar Bakkestuen

Bekkby, T., Rinde, E., Erikstad, L., and Bakkestuen, V. 2009. Spatial predictive distribution modelling of the kelp species *Laminaria hyperborea*. – ICES Journal of Marine Science, 66: 2106–2115.

The kelp species *Laminaria hyperborea* constitutes highly productive kelp forest systems hosting a broad diversity of species and providing the basis for commercial kelp harvesting and, through its productivity, the fishing industry. Spatial planning and management of this important habitat and resource needs to be based on distribution maps and detailed knowledge of the main factors influencing the distribution. However, in countries with a long and complex coastline, such as Norway, detailed mapping is practically and economically difficult. Consequently, alternative methods are required. Based on modelled and field-measured geophysical variables and presence/absence data of *L. hyperborea*, a spatial predictive probability model for kelp distribution is developed. The influence of depth, slope, terrain curvature, light exposure, wave exposure, and current speed on the distribution of *L. hyperborea* are modelled using a generalized additive model. Using the Akaike Information Criterion, we found that the most important geophysical factors explaining the distribution of kelp were depth, terrain curvature, and wave and light exposure. The resulting predictive model was very reliable, showing good ability to predict the presence and absence of kelp.

**Keywords:** geographical distribution, GIS, habitat mapping, kelp, *Laminaria hyperborea*, Norway, Norwegian Sea, predictive modelling.

Received 24 February 2009; accepted 9 June 2009; advance access publication 4 July 2009.

T. Bekkby and E. Rinde: Norwegian Institute for Water Research, Gaustadalléen 21, N-0349 Oslo, Norway. L. Erikstad and V. Bakkestuen: Norwegian Institute for Nature Research, Gaustadalléen 21, N-0349 Oslo, Norway. V. Bakkestuen: Department of Botany, NHM, University of Oslo, PO Box 1172, Blindern, N-0318 Oslo, Norway. Correspondence to T. Bekkby: tel: +47 22 185100; fax: +47 22 185200; e-mail: trine.bekkby@niva.no.

## Introduction

The Norwegian coast accommodates a high density of seaweeds (Jensen, 1998), the kelp *Laminaria hyperborea* being the dominant species (Sjötun *et al.*, 1995). The species is found in the Northeast Atlantic, from Portugal (Kain, 1971) to Russia (Schoschina, 1997), with optimal conditions at latitudes of mid-Norway (63–65°N; Kain, 1967; Rinde and Sjötun, 2005), on subtidal, shallow (<30 m), and rocky substratum in exposed and moderately exposed areas (Kain, 1971).

The kelp forest is a productive system (Kain, 1971; Sjötun *et al.*, 1995; Abdullah and Fredriksen, 2004), with a diverse association of flora and fauna (Moore, 1972; Norton *et al.*, 1977; Schultze *et al.*, 1990; Christie *et al.*, 1998, 2003), providing habitat for several fish species of commercial interest (Norderhaug *et al.*, 2005). In southwestern Norway, *L. hyperborea* has been harvested to produce alginate since the 1970s (Jensen, 1998). In northern Norway, it has been grazed heavily by the green sea urchin, *Strongylocentrotus droebachiensis* for the past 30 years (Skadsheim *et al.*, 1995; Sivertsen, 1997), resulting in large areas of kelp forest being replaced by persisting high densities of sea urchins (Sivertsen, 1997).

Management of the kelp forests, both as a commercially exploited resource and as an important habitat (cf. Rinde *et al.* 2006), is difficult without a map of their distribution and knowledge of the main factors influencing their occurrence. In countries with a long and complex coastline, such as Norway, detailed field-mapping of all areas is practically and economically difficult.

Moreover, simply mapping the kelp distribution does not capture the dynamic nature of the species, identify the reference (i.e. natural) conditions in degraded areas (e.g. through sea urchin grazing), or explain the reasons for changes or deviations. Hence, it does not give managers the knowledge they need for spatial planning and management of the extensive forest. Managers require the output from a cost-effective sampling programme that yields information on the influence of environmental factors on kelp. This is important for understanding regional and local variation, as well as forming a baseline for predicting temporal change. Information and knowledge that can be transferred from few samples to a wider area is of great value. Hence, establishing area coverage GIS layers of environmental factors (based on modelling or field measures), mapping kelp distribution through a well-planned field design, and transferring knowledge of the statistical relationships between distribution patterns and environmental factors into spatial predictive models (i.e. distribution maps) would be a cost-effective and valuable tool for mapping, spatial planning, and managing kelp forests.

Spatial predictive modelling has been applied in several studies (Lehmann, 1998; Guisan and Zimmermann, 2000; Kelly *et al.*, 2001; Lehmann *et al.*, 2003; Elith *et al.*, 2006; Wilson *et al.*, 2007; Bekkby *et al.*, 2008b, c), and is increasingly used for mapping and nature management, for instance, as part of the Norwegian mapping programme on marine biodiversity (Rinde *et al.*, 2006) and as base maps of Gullmarsfjord, Sweden (Bekkby and Rosenberg, 2006).

Bekkby *et al.* (2002) made the first attempt to develop a predictive map of the distribution of *L. hyperborea* forests, based on literature on the environmental requirements of the species. The aims of the present study are to develop a new model by analysing empirical data of *L. hyperborea* distribution along geophysical gradients (depth, slope, terrain curvature, light exposure, wave exposure, and current speed), and transforming the statistical relationships into a predictive probability map of *L. hyperborea* distribution. Previous studies analysed *L. hyperborea* recruitment, age, growth, size and density by depth, wave exposure, and light gradients (Kain, 1962, 1963, 1971; Lobban and Harrison, 1994; Sjøtun and Fredriksen, 1995; Sjøtun *et al.*, 1995; Rinde and Sjøtun, 2005). However, wave exposure has usually only been subjectively or semi-quantitatively defined. Hence, the influence of wave exposure and the interaction with other factors have to a limited extent been analysed up to now, and the results have generally been difficult to compare.

## Methods

### Study site and field sample characteristics

Field data were collected at 778 stations in Sandøy municipality, Møre and Romsdal, Norway (62°N, Figure 1), from 19 September to 1 October 2003. The study area is typical of the outer central west coast of Norway, with small islands, underwater shallows and rocks, and high tidal amplitude (~1.80 m). Stations were selected to cover the slope, wave exposure, and current speed gradients within the study area at seabed depths down to 45 m (see Table 1 for detail). The minimum distance between stations was 15 m. At each station, kelp coverage was recorded using a water glass (in shallow areas, down to ~5 m) or an underwater camera (UWC, in deeper water). The UWC was supplied with light and connected to a monitor on the boat through a cable 100-m long (UWC and cable purchased from www.tronitech.no). The position of each station was recorded on a GPS (Garmin GPSmap 76CSx, accuracy ± 5 m).



**Figure 1.** Large-scale map of the study area (encircled), Sandøy municipality, Møre and Romsdal, Norway (62°N). Grey areas are land.

### *Laminaria hyperborea* presence and absence

At each of the 778 stations, kelp coverage was defined semi-quantitatively in one of the four classes: 0 (no kelp,  $n = 522$ ), 1 (single plants, i.e. 1–2 plants  $\text{m}^{-2}$ ,  $n = 76$ ), 2 (medium dense, i.e. the seabed is visible through the canopy, usually 2–8 plants  $\text{m}^{-2}$ ,  $n = 54$ ), and 3 (dominating/dense, i.e. the seabed is not visible through the canopy, usually 10 plants  $\text{m}^{-2}$ ,  $n = 126$ ). For the spatial predictive modelling, the dataset was divided into two classes, kelp and no kelp. Any coverage of kelp was defined as presence. The dataset had many more absences ( $n = 522$ ) than presences ( $n = 256$ ). To obtain a more balanced dataset for the statistical analyses, 256 of the absences were selected at random to represent areas without kelp. Of the resulting 512 stations, 384 were used to develop the predictive model (Figure 2, left), and 128 stations were randomly selected for independent model validation (Figure 2, right). In all, 178 of the 384 stations used for modelling had kelp present, and kelp was absent from the other 206 stations. To validate the model, 78 of the 128 stations had kelp (i.e. presence) and 50 stations had no kelp (i.e. absence).

### Predictor variables

The variables depth, slope, terrain curvature, light and wave exposure, and current speed were available as GIS layers (as described in Table 1). The correlations between these variables were  $\leq 0.41$ . All data and models were integrated in ArcGIS 9.2. For the statistical analyses used to develop the algorithm for the predictive modelling, we used field-measured depth rather than GIS-modelled depth (see Table 1 for the method). The current speed models had a spatial resolution of 100 m, and the other models had a spatial resolution of 10 m. To preserve most of the information in the layers and to gain an output predictive model at 10-m spatial resolution, the current speed model was resampled to 10 m before the predictive modelling.

We analysed the effect of two terrain curvature models, one with a 500-m calculation window and the other with a 1-km calculation window. We further analysed the effect of surface and depth-attenuated wave exposure models. For current speed, depth-averaged models (averaged over the water column) of mean, median, and 90th percentile were tested. Table 1 gives further detail of the models.

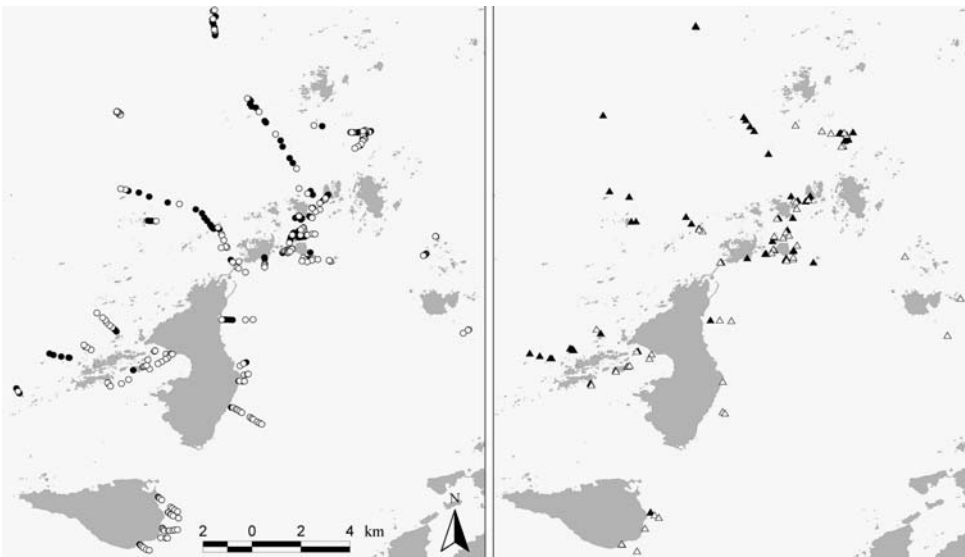
### Generalized additive model and Akaike Information Criterion model selection

We analysed the statistical influence of depth, slope, terrain curvature, light and wave exposure, and current speed using generalized additive models (GAMs; Hastie and Tibshirani, 1990; occurrence as a binomial dataset, 2 d.f. for the smoothing spline function) in S-PLUS 2000. A GAM permits the response probability distribution to be any member of the exponential family of distributions, and the method is commonly used when there are many independent variables in the model.

As a tool for model selection, we used the Akaike Information Criterion (AIC; Burnham and Anderson, 2001) in GRASP (an extension to S-PLUS 2000; Lehmann *et al.*, 2003, 2004). The candidate models were tested and ranked relative to each other. The best model according to AIC would be the model receiving most support from the data, and which at the same time used a small number of explanatory factors (the principle of parsimony, i.e. the trade-off between squared bias and variance against the number of parameters in the model). We used the AICc

**Table 1.** Description of the predictors used in the statistical analyses and spatial predictive modelling.

|                   |  |
|-------------------|--|
| Depth             | For the statistical analyses used to develop the algorithm for the predictive modelling, we used depth values measured in the field, recorded using a rigidly mounted echosounder (with a dual frequency 200/83 kHz sonar system) linked to a GPS (Garmin GPSmap 76CSx, accuracy: $\pm 5$ m). The depth values were adjusted for temporal tidal differences (using tables provided by the Norwegian Mapping Authority, NMA). For the spatial predictive modelling, a digital terrain model (DTM, 10 m spatial resolution) was developed using the Kriging method (Krige, 1967), interpolating bathymetric and topographic data purchased from NMA (14–91 map readings per 100 m <sup>2</sup> )   |
| Slope             | Calculated as maximum change from each DTM cell to its eight closest neighbours (in degrees) using the ArcGIS 9.2 “Slope” function. The slope range within the study area was 0–50°  |
| Terrain curvature | Calculated in ArcGIS 9.2 as the difference between the depth in each DTM grid point and the mean depth in a moving neighbourhood, defined by the size of the calculation window (similar to the bathymetric position index, described by Wilson <i>et al.</i> , 2007). Based on ongoing work on finding the optimal calculation window size for capturing landform characteristics, we used 500-m and 1-km calculation windows (as in Bekkby <i>et al.</i> , 2008b, c). A negative grid cell value indicates that the grid cell is within a basin, a positive value indicates a shoal. The more negative the value, the deeper the basin; the more positive the value, the greater the rise in the shoal   |
| Light exposure    | Calculates the deviance from optimal influx of light based on estimates of vertical slope (using the ArcGIS 9.2 “Slope” function) and orientation (using the ArcGIS 9.2 “Aspect” function) calculated from the DTM. This index was developed for terrestrial vegetation (Parker, 1988; discussed and developed further by Økland, 1990, 1996) and developed as an ArcView 3.3 script for this study. At the optimal slope ( $\sim 45^\circ$ ) and aspect ( $202.5^\circ$ ; Økland 1990, 1996), the light exposure is optimal, with an index value equal to 1. The index is positive at aspects $202.5 \pm 90^\circ$ (regardless of slope) and negative at $(202.5 - 180^\circ) \pm 90^\circ$   |
| Wave exposure     | Modelled from the DTM using data on fetch (distance to nearest shore, island, or coast), wind strength, and direction averaged over a 5-year period, with a spatial resolution of 10 m (more detail in Isæus, 2004). This model has been validated in the Stockholm archipelago (Isæus, 2004) and applied in studies on the effect of boating activities on aquatic vegetation and fish recruitment in the Baltic (Eriksson <i>et al.</i> , 2004; Sandström <i>et al.</i> , 2005). A depth-attenuated version of the model (Bekkby <i>et al.</i> , 2008a) was also tested. The wave exposure values of the study area ranged from sheltered to very exposed (classes developed by Rinde <i>et al.</i> , 2006, similar to the EUNIS system of Davies and Moss, 2003). The dominating winds (cf. data from the Norwegian Meteorological Institute) came from the SSW (i.e. $195 - 225^\circ$ ) |
| Current speed     | Estimated with the three-dimensional numerical ocean model ROMS (Shchepetkin and McWilliams, 2005) in a two-level nesting procedure for a 9-d period. Level 1: ocean currents, atmospheric forcing from forecasts by the Norwegian Meteorological Institute and climatological river flow rates were used to drive an ocean model at a 500-m spatial resolution. Level 2: the fields from the 500-m model were used to drive an inner model at a resolution of 100 m, resampled to 10 m. We selected the depth-averaged component of the model (averaged over the water column), and estimated the mean, median (50th percentile), and 90th percentile. The range of current speeds within the study area was $0 - 0.9 \text{ m s}^{-1}$ .   |



**Figure 2.** A detailed map of the study area (Sandøy municipality, Møre and Romsdal, Norway, 62°N). Grey areas are land. Circles (left panel) identify the stations used for the statistical analyses and predictive modelling (384 stations), triangles (right panel) stations used for independent model validation (128 stations). Closed symbols are stations with kelp; open symbols are stations without kelp. Note that some stations may be hidden behind others.

calculations, as recommended by Burnham and Anderson (2004), which is the AIC adjusted to fit small sample sizes (AIC and AICc being equal at large sample sizes).

**Spatial probability prediction and validation**

Based on the response curves from the GAM analysis, GRASP develops a matrix of prediction. In ArcView 3.3, the matrix for

**Table 2.** Environmental parameter statistics for the four different coverage classes 0 (no kelp,  $n = 256$ ), 1 (single plants,  $n = 76$ ), 2 (medium dense,  $n = 54$ ), and 3 (dominating/dense,  $n = 126$ ).

| Parameter         | Coverage | Mean       | Standard deviation | Median (50th percentile) | Maximum      | Minimum   |
|-------------------|----------|------------|--------------------|--------------------------|--------------|-----------|
| Depth             | 0        | -11.24     | 9.17               | -8.84                    | 0.32         | -43.10    |
|                   | 1        | -10.63     | 8.89               | -7.50                    | -0.47        | -31.94    |
|                   | 2        | -6.96      | 6.35               | -4.28                    | -0.34        | -21.15    |
|                   | 3        | -4.36      | 4.09               | -3.21                    | -0.03        | -16.43    |
| Slope             | 0        | 4.55       | 6.42               | 2.32                     | 46.59        | 0.26      |
|                   | 1        | 3.14       | 3.63               | 2.13                     | 20.21        | 0.12      |
|                   | 2        | 4.46       | 3.82               | 3.21                     | 15.70        | 0.47      |
|                   | 3        | 4.68       | 3.79               | 3.74                     | 18.59        | 0.14      |
| Terrain curvature | 0        | 3.62       | 10.89              | 1.13                     | 54.63        | -12.91    |
|                   | 1        | 0.35       | 3.91               | -0.19                    | 15.00        | -11.82    |
|                   | 2        | 3.72       | 4.05               | 3.12                     | 14.36        | -5.00     |
|                   | 3        | 6.01       | 4.49               | 6.27                     | 16.29        | -4.31     |
| Light exposure    | 0        | 0.02       | 0.10               | 0.00                     | 0.61         | -0.39     |
|                   | 1        | 0.00       | 0.06               | 0.00                     | 0.20         | -0.28     |
|                   | 2        | 0.03       | 0.09               | 0.02                     | 0.27         | -0.25     |
|                   | 3        | 0.03       | 0.07               | 0.02                     | 0.33         | -0.11     |
| Wave exposure     | 0        | 416 718.33 | 521 000.02         | 135 497.50               | 1 880 347.00 | 8 716.00  |
|                   | 1        | 742 233.54 | 672 622.39         | 586 923.50               | 1 856 411.00 | 23 704.00 |
|                   | 2        | 583 793.98 | 639 351.44         | 209 753.00               | 1 865 978.00 | 23 842.00 |
|                   | 3        | 674 426.12 | 538 084.91         | 693 411.50               | 1 869 373.00 | 13 982.00 |
| Current speed     | 0        | 0.05       | 0.04               | 0.04                     | 0.23         | 0.00      |
|                   | 1        | 0.06       | 0.05               | 0.04                     | 0.21         | 0.00      |
|                   | 2        | 0.06       | 0.05               | 0.05                     | 0.20         | 0.00      |
|                   | 3        | 0.05       | 0.04               | 0.04                     | 0.22         | 0.00      |

The statistics include the data used for both modelling and validation. Terrain curvature is measured with a 1-km calculation window. A negative value indicates a basin and a positive value a shoal. The more negative the value, the deeper the basin; the more positive the value, the greater the rise in the shoal. The light exposure index calculates the deviance from optimal influx of light based on information on slope and orientation ("aspect"). At the optimal slope ( $45^\circ$ ) and orientation ( $202.5^\circ$ ), the light exposure index is 1. Wave exposure is calculated as a 5-year mean, and current speed is measured as the median (50th percentile) over a 9-d period. More details on the predictors are given in Table 1.

the selected statistical model was used to produce an output spatial probability model (at a spatial resolution of 10 m) from GIS layers of the variables.

The predicted probability distribution was validated using both a cross-validation test and an independent dataset. In both cases, we used the area under the receiver operating characteristic (ROC) curve, known as the AUC value (Swets, 1988). The cross-validation ROC test (Fielding and Bell, 1997) was made with five subsets (folds) of the same dataset used for modelling (fivefold cv-ROC), and is an "internal" test of the predictive model against the dataset used for modelling. The independent dataset tests the model against a dataset not included in the statistical model building or prediction (in all 128 stations, 78 presences and 50 absences).

### Predicted probability against coverage

Based on statistical analyses of the presence/absence data ( $n = 384$ ), the output spatial predictive model provides estimates of the probability of finding kelp at each of the stations. We compared these probabilities (arcsin-transformed) for the four kelp coverage classes observed in the field (0, no kelp,  $n = 206$ ; 1, single plants,  $n = 53$ ; 2, medium dense,  $n = 39$ ; 3, dominating/dense,  $n = 86$ ) using ANOVA in StatGraphics Plus 5.1. Multiple range tests provide information on which means that are significantly different from others.

### Results

In our area, the maximum depth of *L. hyperborea* (regardless of coverage) was 32 m. Kelp dominated as a dense forest (i.e.

coverage class 3) down to 16 m and as medium dense (i.e. coverage class 2) down to 21 m. Environmental parameter statistics for the different coverage classes are listed in Table 2. The analyses show that the distribution of kelp is best determined by the combined effect of depth, terrain curvature (with a 1-km calculation window), wave and light exposure, and slope (Model 1 in Table 3; fivefold cv-ROC = 0.78, independent data ROC = 0.79). The result is not improved by using terrain curvature with a 500-m calculation window or the depth-attenuated wave exposure ( $\Delta_i$  was always larger). Model 1 is only slightly better ( $\Delta_i = 1.17$ ; Table 3) than the model not including slope, i.e. based on depth, terrain curvature (with a 1-km calculation window), and wave and light exposure (models with a  $\Delta_i \leq 2$  have substantial support; Burnham and Anderson, 2001). Consequently, to preserve the principle of parsimony, this simpler model (Model 2 in Table 3; fivefold cv-ROC = 0.78, independent data ROC = 0.79) was selected for the prediction.

Figure 3 depicts histograms of the distribution of the field observations along the slope, light exposure, terrain curvature, wave exposure, current speed, and depth gradients. The probability of finding kelp is high in shallow and wave-exposed areas with good light conditions and intermediate terrain curvature (Figure 4). The probability of finding kelp is close to zero at 30 m depth and low in the very sheltered areas. It increases with increasing light exposure and as the terrain moves from being a basin towards being a shoal (i.e. as the terrain curvature index turns positive; Figure 4). For terrain curvature, the probability increases to a certain point, then decreases as the shoals become steeper (i.e. as the terrain curvature index peaks).



The statistical analyses show that as a single factor, depth was most important, followed by (in decreasing importance) terrain curvature, wave exposure, light exposure, and slope (Table 2,

**Table 3.** Results from the AICc model selection (sorted with ascending AICc values; the lower the value, the better the model).

| Model | Predictors  | AICc   | $\Delta_i$ | $W_i$ |
|-------|---|--------|------------|-------|
| 1     | Depth, terrain curvature, wave exposure, light exposure, slope                | 423.14 | 0          | 0.56  |
| 2     | Depth, terrain curvature, wave exposure, light exposure                       | 424.31 | 1.17       | 0.31  |
| 3     | Depth, terrain curvature, wave exposure, light exposure, slope, current speed | 426.71 | 3.57       | 0.09  |
| 4     | Depth, terrain curvature, wave exposure                                       | 428.46 | 5.32       | 0.04  |
| 5     | Depth, wave exposure  | 440.06 | 16.92      | 0.00  |
| 6     | Terrain curvature, wave exposure, light exposure                              | 485.14 | 62.00      | 0.00  |
| 7     | Terrain curvature, wave exposure  | 486.11 | 62.97      | 0.00  |
| 8     | Depth, terrain curvature, light exposure                                      | 490.26 | 67.12      | 0.00  |
| 9     | Depth, terrain curvature  | 490.65 | 67.51      | 0.00  |
| 10    | Depth   | 498.88 | 75.74      | 0.00  |
| 11    | Terrain curvature   | 509.97 | 86.83      | 0.00  |
| 12    | Wave exposure   | 516.77 | 93.63      | 0.00  |
| 13    | Light exposure  | 529.09 | 105.95     | 0.03  |
| 14    | Slope   | 529.58 | 106.44     | 0.00  |
| 15    | Current speed   | 531.18 | 108.04     | 0.00  |

The response variable was kelp *L. hyperborea* occurrence (presence/absence). The predictor variables used in the analyses were depth, terrain curvature (with 1-km calculation window), wave exposure, light exposure, slope, and current speed.  $\Delta_i$  is the discrepancy from the best model. Models with  $\Delta_i < 2$  have substantial support, those with  $4 < \Delta_i < 7$  have considerable less support, and those with  $\Delta_i > 10$  have no support.  $W_i$  is the probability (range 0–1) that the model is in fact the most appropriate of those tested (the Akaike weight).

Figure 5). Figure 6 shows the resulting map of the probability of finding kelp using the response curves for depth, terrain curvature, and light and wave exposure (Figure 4).

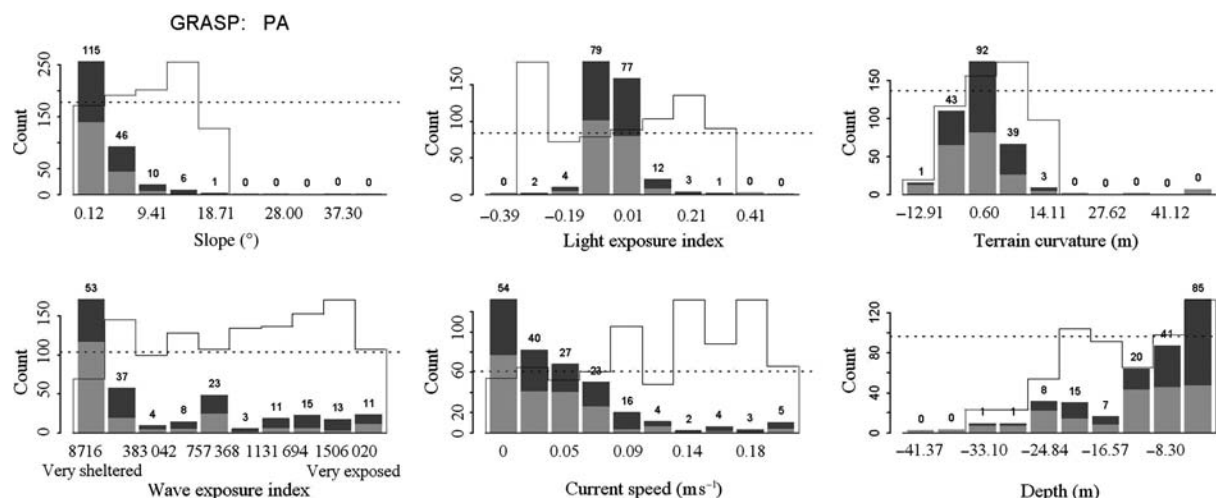
The ANOVA analysis comparing the predicted probabilities (arcsin-transformed) for the four different coverage classes observed in the field (0, no kelp; 1, single plants; 2, medium dense; 3, dominating/dense) shows that there is an overall significant difference in mean probability from one level of coverage to another ( $F = 51.98$ ,  $p < 0.001$ ; Table 4). The probability of finding kelp increases with observed coverage (Figure 7). The multiple range tests (Table 3) show that the probability values of coverage class 1 (single plants) is not significantly different from class 2 (medium dense). All other classes were significantly different from each other.

## Discussion and conclusions

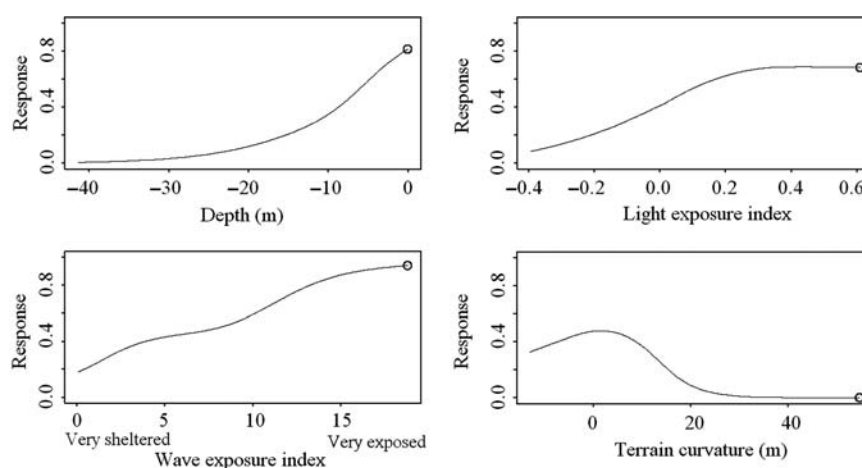
The spatial predictive model including depth, terrain curvature, and wave and light exposure was able to discriminate between sites where the kelp *L. hyperborea* is present from those where it is absent ( $AUC = 0.79$ ). In our study area, *L. hyperborea* mainly lives in shallow and wave-exposed areas with good light (calculated from slope and orientation) and intermediate terrain curvature (the latter indicating rocky, but not too steep, seabed).

Depth is the single most important factor, followed by (in decreasing importance) terrain curvature, wave exposure, and light exposure. Depth does not have a direct impact on kelp distribution, but it is a proxy for light attenuation. Others have found a decrease in recruitment, growth, and density of *L. hyperborea* with increasing depth (Kain, 1971; Sjøtun and Fredriksen, 1995; Sjøtun *et al.*, 1995).

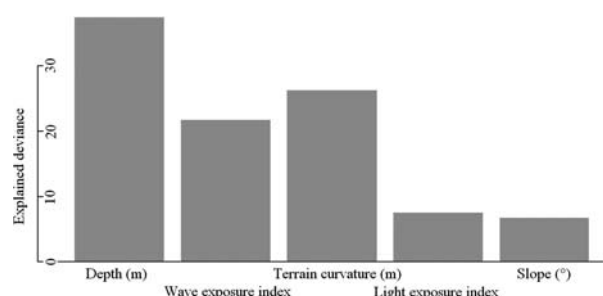
The main effect of terrain curvature (the second most important factor, similar to the bathymetric position index described by Wilson *et al.*, 2007) is most likely its indication of substratum. Our results show that the probability of finding kelp increases as the terrain changes from basin to shoal (e.g. where the soft sediments associated with basins are replaced by the rocky seabed found on shoals). Terrain curvature is a good indicator



**Figure 3.** Histogram of the field observations used for statistical analyses along slope, light exposure, terrain curvature, wave exposure, median current speed, and depth gradients. The height of the bars represents the total number of observations (in all, 178 presences and 206 absences). The light grey parts of the bars represent the number of kelp absences, and the dark grey parts of the bars and the number on the top of each bar represent the number of presences. Dashed horizontal lines represent the overall proportion of presences compared with the total number of observations; solid outline shows the proportion of presences for each bar. The light exposure index calculates the deviance from optimal influx of light based on information on slope and orientation ("aspect"). At the optimal slope ( $45^\circ$ ) and aspect ( $202.5^\circ$ ), the light exposure index is 1. Terrain curvature is measured with a 1-km calculation window. A negative value indicates a basin and a positive value a shoal. The more negative the value, the deeper the basin; the more positive the value, the greater the rise in the shoal.



**Figure 4.** Response curves for kelp *L. hyperborea* occurrence (presence/absence) against depth, light exposure, wave exposure, and terrain curvature. The y-axis is scaled according to the response, and is a number between 0 and 1. The light exposure index calculates the deviance from optimal influx of light based on information on slope and orientation ("aspect"). At the optimal slope ( $45^\circ$ ) and aspect ( $202.5^\circ$ ), the light exposure index is 1. The wave exposure index values (Isæus, 2004) are in  $\times 10^5$ . Terrain curvature is measured with a 1-km calculation window. A negative value indicates a basin and a positive value a shoal. The more negative the value, the deeper the basin; the more positive the value, the greater the rise in the shoal.



**Figure 5.** The deviance explained by each of the predictors in the best model (defined by the AICc analyses; see Table 3). The light exposure index calculates the deviance from optimal influx of light based on information on slope and orientation ("aspect"). At the optimal slope ( $45^\circ$ ) and aspect ( $202.5^\circ$ ), the light exposure index is 1. Terrain curvature is measured with a 1-km calculation window. A negative value indicates a basin and a positive value a shoal. The more negative the value, the deeper the basin; the more positive the value, the greater the rise in the shoal. More details on the predictors are given in Table 1.

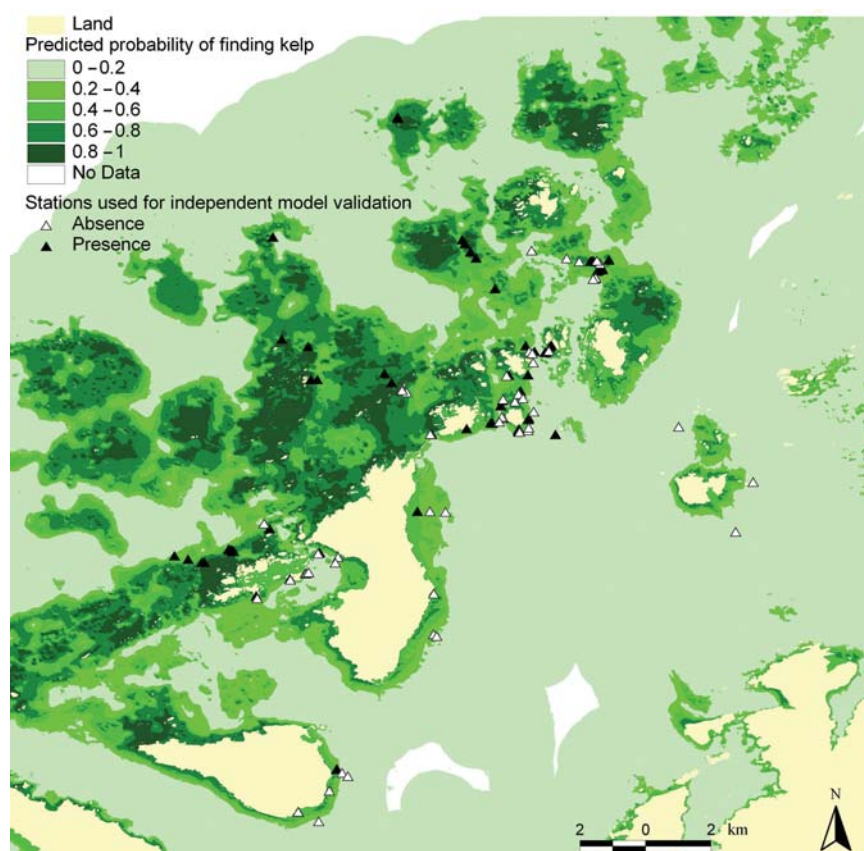
of seabed substratum, and has been used to identify soft sediment basins (Bekkby *et al.*, 2008b). As *L. hyperborea* needs rocky substratum to attach, this relationship is reasonable. The probability of finding kelp decreases as the slope of the shoals steepens, indicating that extremely steep slopes in general, including around shoals, are unsuitable for kelp growth.

The terrain curvature estimate has a weakness. We calculated the difference between the depth at each grid point and mean depth in a moving neighbourhood. A negative grid cell value indicates that the grid cell is within a basin, and a positive value indicates a shoal. However, grid cells in the middle of a slope will be assigned terrain curvature values close to zero, because the depth value is close to the neighbourhood average. Hence, mid-slopes may not be separated from flat terrain. Figure 4 may consequently be used to separate the effect of basins from shoals, but not that of flat terrain and mid-slope.

Wave exposure is the third most important factor in the model, and the probability of finding kelp is low in very sheltered areas and increases with enhanced wave exposure. Others have found that growth and density of *L. hyperborea* are influenced by wave exposure (Kain, 1971; Sjøtun and Fredriksen, 1995; Sjøtun *et al.*, 1995), and that wave action has a positive effect through moving algal fronds maximizing the area available to trap light, as well as maintaining a high nutrient flux (Lobban and Harrison, 1994).

We expected the depth-attenuated wave exposure model (developed by Bekkby *et al.*, 2008a) to improve the prediction of kelp, but this was not the case. Although the depth-attenuated model includes the decline in wave exposure with depth, it does not take into account the effect of ocean swell on waves, resulting in an underestimate that might influence the depth-attenuated model more than the non-attenuated version.

Light exposure is the fourth most important factor in the model. This is consistent with the results of other studies showing profound impacts of light on growth in *L. hyperborea* (Lobban and Harrison, 1994). The probability of finding kelp increases as the light exposure index increases (Figure 4). The index represents the exposure of light as a function of slope and orientation, and the maximum value (an index of 1) represents the optimal angle of both and consequently the highest level of light influx. We believe that this is the first time that such a pattern, indicating that vegetation occurrence peaks at optimal sunlight exposure angles, has been proposed for the marine environment. Such patterns have been observed on land in studies in Norway (Økland, 1996). It is worth mentioning that none of our stations had optimal light conditions, because most of them had a slope of  $<20^\circ$ . Hence, the effect of light exposure was tested across the whole orientation range, but not across the whole slope range. Also, the optimal orientation for maximum light influx is SSW ( $202.5^\circ$ ), i.e. within the range of the dominating winds in the area ( $195\text{--}225^\circ$ ). As wind-based wave exposure has an effect on the probability of finding kelp, the effect of light exposure is difficult to separate from



**Figure 6.** Predicted probability of finding kelp (*L. hyperborea*) within the study area (Sandøy municipality, Møre and Romsdal, Norway, 62°N) based on the response curves for depth, terrain curvature, and wave and light exposure (cf. Figure 4). The darker the green, the higher the probability of kelp presence. The spatial resolution of the model is 10 m. White areas are not covered by the model. Yellow areas are land. Triangles indicate stations used for independent model validation (128 stations). Closed symbols are stations with kelp and open symbols stations without kelp. Note that some stations may be hidden behind others.

**Table 4.** Results from the Duncan's multiple range tests, providing information on which kelp (*L. hyperborea*) coverage class had predicted probability means that were significantly different from the other coverage classes (probability values arcsin-transformed).

| Coverage                                | 0<br>(no kelp,<br><i>n</i> = 206) | 1<br>(single<br>plants,<br><i>n</i> = 53) | 2<br>(medium<br>dense,<br><i>n</i> = 39) | 3<br>(dominating/<br>dense,<br><i>n</i> = 86) |
|---|-----------------------------------|---|--|---|
| 0 (no kelp,<br><i>n</i> = 206)          |                                   | +   | +  | +   |
| 1 (single plants,<br><i>n</i> = 53)     | +                                 |   | –  | +   |
| 2 (medium dense,<br><i>n</i> = 39)      | +                                 | –   |  | +   |
| 3 (dominating/<br>dense, <i>n</i> = 86) | +                                 | +   | +  |   |

The minus sign indicates no significant difference, the plus sign a significant difference ( $p \leq 0.05$ ). The different levels of coverage were 0 (no kelp,  $n = 206$ ), 1 (single plants,  $n = 53$ ), 2 (medium dense,  $n = 39$ ), and 3 (dominating/dense,  $n = 86$ ).

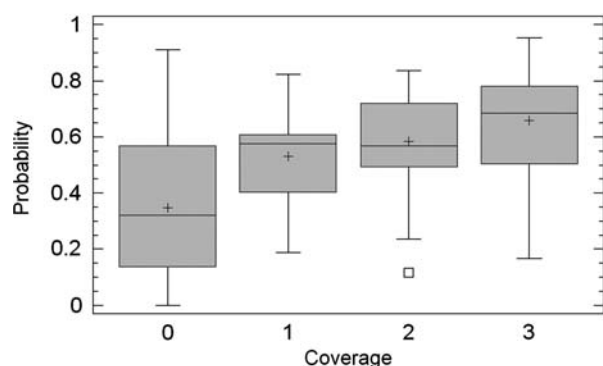
wave-exposure effects. However, the correlation between the wave exposure and the light exposure is small ( $r^2 = 0.15$ ), indicating an independent light exposure influence.

The model including slope in addition to depth, terrain curvature, wave exposure, and light exposure is only marginally better

than the model without slope (slope made the least contribution to the model). As steep slopes indicate rocky seabed (Bekkby *et al.*, in press), we expected slope to have an impact. A possible explanation is that the terrain curvature captures the landscape features and seabed substratum types in a better way than slope. Terrain curvature is notably better at identifying the soft sediment basins (Bekkby *et al.*, 2008b), which are unsuitable habitats for kelp. In flat terrain in exposed areas, glacial debris, e.g. rocks and boulders, provides the opportunity for kelp growth in a terrain otherwise assumed to have soft sediments. This might additionally explain why slope is not improving the model; kelp was in fact present in both flat and steep terrain in areas where wave exposure limits the sedimentation.

Current speed does not improve the predictive model. This result contrasted our expectations that current speed would have similar effect to wave exposure, i.e. through drag (Thomsen *et al.*, 2004), moving algal fronds maximizing the area to trap light and maintaining a high nutrient flux (Lobban and Harrison, 1994). One possible explanation for this lack of response may be that the influence of wave exposure is more important than currents within the area. The range of wave exposure is broad, ranging from sheltered to very exposed. However, the range of current speeds is not as broad. We believe that an effect of current speed might have been detected if the study included areas of ultra-sheltered locations with very high current speeds (*L. hyperborea* has been observed





**Figure 7.** Box-and-whisker plot showing the variability in the predictive probability values within each coverage class 0 (no kelp,  $n = 206$ ), 1 (single plants,  $n = 53$ ), 2 (medium dense,  $n = 39$ ), and 3 (dominating/dense,  $n = 86$ ). The Duncan's multiple range tests showed that the probability values (arcsin-transformed) of coverage class 1 were not significantly different from class 2, but that all other classes were significantly different from each other. The data in the plot are divided into four areas of equal frequency (quartiles). The grey box in the plot encloses the middle 50%, where the median is drawn as a vertical line inside the box and the mean as a +. The lower whisker is drawn from the lower quartile to the smallest point within 1.5 interquartile ranges from the lower quartile. The upper whisker is drawn from the upper quartile to the largest point within 1.5 interquartile ranges from the upper quartile. The open square represents a value that falls beyond the whiskers (an outlier), but which is within three interquartile ranges.

under such conditions in narrow sounds and fjords; H. Christie, NIVA, pers. comm.). A second explanation might be that the current speed model resolution (100 m) is too coarse to detect current effects at the scale of our study.

The probability of finding kelp, as predicted from the model, rises with an increase in observed kelp coverage. This is logical, because the predicted probability represents the chance of finding kelp at a given combination of geophysical factors. As shown in Figure 7, areas in which no kelp was recorded in the field had a median probability, defined by the predictive model as 0.32. For areas of single plants and medium dense kelp forest, the median probability values were in both cases  $\sim 0.57$ . For dominating/dense kelp forest, the value was 0.68. Hence, the probability model may be transferred into a model of coverage, and can be used to identify areas with dense kelp forest. This is of value for nature and resource management of kelp, because dense kelp forests are regarded as having a high ecological value (cf. the Norwegian mapping programme on marine biodiversity; Rinde *et al.*, 2006) and the highest yield potential with respect to harvesting.

The quality of the output predictive model is never better than that of the input models used to develop the statistical algorithm. In our study, the grid resolution of most of the input models was quite high, higher than for most areas along the Norwegian coast. However, a spatial resolution of 10 m means that we have one depth value representing an area of 100 m<sup>2</sup>. This bias might have a considerable effect in areas of high terrain variability, and will be transferred to all derived parameters (e.g. slope and terrain curvature). Knowing the influence of this bias and how it changes with scale is important in making the right decisions for data sampling and modelling, and for transferring prediction algorithms from the areas in which they are developed to areas

of different resolution of predictor models. Ongoing analyses (Lars Erikstad, NINA, pers. comm.) test the effect of input model resolution on the quality of the derived terrain indices and the usability of the output geomorphology and habitat models. The results from those studies will provide knowledge of whether or not (and hopefully how) algorithms developed in areas with high-resolution models may be applied to other areas.

We used the AUC for model validation. Because it is independent of prevalence, the method is regarded as good for assessing the discriminatory power of spatial predictive models (Fielding, 2002; Pearce and Ferrier, 2000; Schröder, 2006), although this has recently been discussed (Lobo *et al.*, 2008).

This work has illustrated how GIS-based modelling can be used to integrate information on geophysical conditions to analyse and predict the distribution of the kelp *L. hyperborea*. We regard the methodology and the resulting map as an improvement to the approach described in Bekkby *et al.* (2002) in several ways. First, the variables used for prediction are easily available for many areas along the Norwegian coast. Second, the resulting model identifies areas of different *L. hyperborea* occurrence probabilities. Probability maps provide users with the choice of applying the precautionary principle or a more focused approach. Evaluating all areas (also those of low probabilities) is a relevant starting point for applying the precautionary principle. However, if priority has to be given to areas with respect to establishing marine protected areas, for instance, it will be useful to identify areas with the greatest probabilities only. Finally, our study links the predicted probability and coverage of kelp observed in the field, which makes the results easy to understand and disseminate to various stakeholders.

As the algorithm and approach may be applied to other areas, the potential of the method as a tool for mapping and nature management of kelp forests is good. The algorithm we have developed has recently been applied to other areas, as a basis for sampling design, and the approach has been applied too to studies of the distribution of barren grounds caused by green sea urchins grazing in northern Norway (unpublished). Further, the modelled habitat maps provide the information needed to report on the natural status of the marine environment, and may be used to develop hypotheses for reasons for deviations from the natural state.

## Acknowledgements

The project was funded by the Research Council of Norway, the Norwegian Institute for Nature Research (NINA), and the Norwegian Institute for Water Research (NIVA). We thank Martin Isæus (AquaBiota Water Research, Sweden) for the wave-exposure modelling, Pål Erik Isachsen (Norwegian Meteorological Institute) for the current-speed modelling, Anna Engdahl (AquaBiota Water Research, Sweden) for help with AUC measures, and Hartvig Christie (NIVA) and the reviewers for valuable comments on the manuscript. We also acknowledge Oddvar Longva (Geological Survey of Norway), Ole Christensen (Electromagnetic Geoservice, Norway), and Ellen Soldal (Norwegian University of Life Sciences) for the fun and intense fieldwork, and the staff at Finnøy Sjøhus for taking care of us during the fieldwork.

## References

- Abdullah, M. I., and Fredriksen, S. 2004. Production, respiration and exudation of dissolved organic matter by the kelp *Laminaria*



- hyperborea* along the west coast of Norway. Journal of the Marine Biological Association of the UK, 84: 887–894.
- Bekkby, T., Erikstad, L., Bakkestuen, V., and Bjørge, A. 2002. A landscape ecological approach to coastal zone applications. *Sarsia*, 87: 396–408.
- Bekkby, T., Isachsen, P. E., Isæus, M., and Bakkestuen, V. 2008a. GIS modelling of wave exposure at the seabed—a depth-attenuated wave exposure model. *Marine Geodesy*, 31: 117–127.
- Bekkby, T., Moy, F., Kroglund, T., Gitmark, J., Walday, M., Rinde, E., and Norderhaug, K. M. Identifying rocky seabed using GIS modelled predictor variables. *Marine Geodesy*, in press.
- Bekkby, T., Nilsson, H. C., Rygg, B., Isachsen, P. E., Olsgard, F., and Isæus, M. 2008b. Identifying soft sediments at sea using GIS-modelled predictor variables and Sediment Profile Image (SPI) measured response variables. *Estuarine, Coastal and Shelf Science*, 79: 631–636.
- Bekkby, T., Rinde, E., Erikstad, L., Bakkestuen, V., Longva, O., Christensen, O., Isæus, M., et al. 2008c. Spatial probability modelling of eelgrass *Zostera marina* L. distribution on the west coast of Norway. *ICES Journal of Marine Science*, 65: 1093–1101.
- Bekkby, T., and Rosenberg, R. 2006. Marine habitats utbredelse—terrengmodellering i Gullmarsfjorden. Report to the Västra Götaland administrative county, Vattenvårdsenheten, Report 2006:07. ISSN 1403-168X. 33 pp. (in Norwegian/Swedish with English abstract). [http://www.o.lst.se/o/Publikationer/Rapporter/2006/2006\\_07.htm](http://www.o.lst.se/o/Publikationer/Rapporter/2006/2006_07.htm).
- Burnham, K. P., and Anderson, D. R. 2001. Kullback–Leibler information as a basis for strong inference in ecological studies. *Wildlife Research*, 28: 111–119.
- Burnham, K. P., and Anderson, D. R. 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33: 261–304.
- Christie, H., Fredriksen, S., and Rinde, E. 1998. Regrowth of kelp and colonization of epiphyte and fauna community after kelp trawling at the coast of Norway. *Hydrobiologia*, 375/376: 49–58.
- Christie, H., Jørgensen, N. M., Norderhaug, K. M., and Waage-Nielsen, E. 2003. Species distribution and habitat exploitation of fauna associated with kelp (*Laminaria hyperborea*) along the Norwegian coast. *Journal of the Marine Biological Association of the UK*, 83: 687–699.
- Davies, C. E., and Moss, D. 2003. EUNIS Habitat Classification. European Topic Centre on Nature Protection and Biodiversity, Paris. <http://eunis.eea.eu.int/habitats.jsp>.
- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29: 129–151.
- Eriksson, B. K., Sandström, A., Isæus, M., Schreiber, H., and Karås, P. 2004. Effects of boating activities on aquatic vegetation in the Stockholm archipelago, Baltic Sea. *Estuarine, Coastal and Shelf Science*, 61: 339–349.
- Fielding, A. H. 2002. What are the appropriate characteristics of an accuracy measure? In *Predicting Species Occurrences. Issues of Accuracy and Scale*, pp. 271–280. Ed. by J. M. Scott, P. J. Heglund, M. Morrison, J. B. Haufler, M. G. Raphael, W. B. Wall, and F. Samson. Island Press, Washington.
- Fielding, A. H., and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24: 38–46.
- Guisan, A., and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135: 147–186.
- Hastie, T. J., and Tibshirani, R. J. 1990. *Generalized Additive Models*. Chapman and Hall, London.
- Isæus, M. 2004. Factors structuring *Fucus* communities at open and complex coastlines in the Baltic Sea. Doctoral thesis, Department of Botany, Stockholm University, Sweden. 165 pp. [http://www.aquabiota.se/publications/pdf/Avhandling\\_Isaesus.pdf](http://www.aquabiota.se/publications/pdf/Avhandling_Isaesus.pdf).
- Jensen, A. 1998. The seaweed resources of Norway. In *Seaweed Resources of the World*, pp. 200–209. Ed. by A. T. Critchley, and M. Ohno. Japan International Cooperation Agency, Yokosuka.
- Kain, J. M. 1962. Aspects of the biology of *Laminaria hyperborea*. 1. Vertical distribution. *Journal of the Marine Biological Association of the UK*, 42: 377–385.
- Kain, J. M. 1963. Aspects of the biology of *Laminaria hyperborea*. 2. Age, weight and length. *Journal of the Marine Biological Association of the UK*, 43: 129–151.
- Kain, J. M. 1967. Populations of *Laminaria hyperborea* at various latitudes. *Helgoländer Wissenschaftliche Meeresuntersuchungen*, 15: 489–499.
- Kain, J. M. 1971. The biology of *Laminaria hyperborea* 6. Some Norwegian populations. *Journal of the Marine Biological Association of the UK*, 51: 387–408.
- Kelly, N. M., Fonseca, M., and Whitfield, P. 2001. Predictive mapping for management and conservation of seagrass beds in North Carolina. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 11: 437–451.
- Krige, D. G. 1967. Two-dimensional moving-average trend surface for ore evaluation. *Journal of the South African Institution of Mining and Metallurgy*, 67: 21–29.
- Lehmann, A. 1998. GIS modeling of submerged macrophyte distribution using generalized additive models. *Plant Ecology*, 139: 113–124.
- Lehmann, A., Leathwick, J. R., and Overton, J. M. 2004. GRASP v.3.1. User's Manual. Swiss Centre for Faunal Cartography, Switzerland.
- Lehmann, A., Overton, J. M., and Leathwick, J. R. 2003. GRASP: generalized regression analysis and spatial prediction. *Ecological Modelling*, 160: 165–183.
- Lobban, C. S., and Harrison, P. J. 1994. *Seaweed Ecology and Physiology*. 1. Cambridge University Press, Cambridge, UK.
- Lobo, J. M., Jiménez-Valverde, A., and Real, R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17: 145–151.
- Moore, P. G. 1972. The kelp fauna of North East Britain. 1. Introduction and the physical environment. *Journal of Experimental Marine Biology and Ecology*, 13: 97–125.
- Norderhaug, K. M., Christie, H., Fosså, J. H., and Fredriksen, S. 2005. Fish–macrofauna interactions in a kelp (*Laminaria hyperborea*) forest. *Journal of the Marine Biological Association of the UK*, 85: 1279–1286.
- Norton, T. A., Hiscock, K., and Kitching, J. A. 1977. The ecology of Lough Ine. 20. The *Laminaria* forest at Carrigathorna. *Journal of Ecology*, 65: 919–941.
- Økland, T. 1990. Vegetational and ecological monitoring of boreal forests in Norway. 1. Rausjømarka in Akershus county, SE Norway. *Sommerfeltia*, 10: 1–52.
- Økland, T. 1996. Vegetation–environment relationships of boreal spruce forests in ten monitoring reference areas in Norway. *Sommerfeltia*, 22: 1–349.
- Parker, K. C. 1988. Environmental relationships and vegetation associates of columnar cacti in the northern Sonoran desert. *Vegetation*, 78: 125–140.
- Pearce, J., and Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133: 225–245.
- Rinde, E., Rygg, B., Bekkby, T., Isæus, M., Erikstad, L., Sloreid, S-E., and Longva, O. 2006. Documentation of marine nature type models included in Directorate of Nature Management's database Naturbase. First generation models for the municipalities mapping of marine biodiversity 2007. NIVA Report LNR 5321-2006 (in Norwegian with English abstract).
- Rinde, E., and Sjøtun, K. 2005. Demographic variation in the kelp *Laminaria hyperborea* along a latitudinal gradient. *Marine Biology*, 146: 1051–1062.

- Sandström, A., Eriksson, B. K., Karås, P., Isæus, M., and Schreiber, H. 2005. Boating activities influence the recruitment of near-shore fishes in a Baltic Sea archipelago area. *Ambio*, 34: 125–130.
- Schoschina, E. V. 1997. *Laminaria hyperborea* (Laminariales, Phaeophyceae) on the Murman coast of the Barents Sea. *Sarsia*, 82: 371–373.
- Schröder, B. 2006. ROC and AUC-calculation—evaluating the predictive performance of habitat models. Computer program, version 1.3, [www.unipotsdam.de/users/schroeder/download.html](http://www.unipotsdam.de/users/schroeder/download.html).
- Schultze, K., Janke, K., Krüß, A., and Weidemann, W. 1990. The macrofauna and macroflora associated with *Laminaria digitata* and *L. hyperborea* at the island of Helgoland (German Bight, North Sea). *Helgoländer Wissenschaftliche Meeresuntersuchungen*, 44: 39–51.
- Shchepetkin, A. F., and McWilliams, J. C. 2005. Regional ocean model system: a split-explicit ocean model with a free surface and topography-following vertical coordinate. *Ocean Modelling*, 9: 347–404.
- Sivertsen, K. 1997. Geographic and environmental factors affecting the distribution of kelp beds and barren grounds and changes in biota associated with kelp reduction at sites along the Norwegian coast. *Canadian Journal of Fisheries and Aquatic Sciences*, 54: 2872–2887.
- Sjötun, K., and Fredriksen, S. 1995. Growth allocation in *Laminaria hyperborea* (Laminariales, Phaeophyceae) in relation to age and wave exposure. *Marine Ecology Progress Series*, 126: 213–222.
- Sjötun, K., Fredriksen, S., Rueness, J., and Lein, T. E. 1995. Ecological studies of the kelp *Laminaria hyperborea* (Gunnerus) Foslie in Norway. In *Ecology of Fjords and Coastal Waters*, pp. 525–536. Ed. by H. R. Skjoldal, C. Hopkins, K. E. Erikstad, and H. P. Leinaas. Elsevier Science, Amsterdam.
- Skadsheim, A., Christie, H., and Leinaas, H. P. 1995. Population reductions of *Strongylocentrotus droebachiensis* (Echinodermata) in Norway and the distribution of its endoparasite *Echinomermella matsi* (Nematoda). *Marine Ecology Progress Series*, 119: 199–209.
- Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. *Science*, 240: 1285–1293.
- Thomsen, M. S., Wernberg, T., and Kendrick, G. A. 2004. The effect of thallus size, life stage, aggregation, wave exposure and substratum conditions on the forces required to break or dislodge the small kelp *Ecklonia radiata*. *Botanica Marina*, 47: 454–460.
- Wilson, M. F. J., O'Connell, B., Brown, C., Guinan, C. J., and Grehan, A. J. 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30: 3–35.

doi:10.1093/icesjms/fsp195