EUSeaMap

Technical Appendix No. 2

Wave base data and thresholds

For further information see: http://jncc.defra.gov.uk/page-5020

Contents

1 Introduction

This technical appendix to the EUSeaMap project report describes the creation of a wave base layer and the use of habitat sample data to test the boundary between the circalittoral and deep circalittoral depth zones in the North and Celtic Seas.

1.1 Background

1.1.1 Biological zonation

The EUNIS habitat classification takes its structure and the majority of its habitat and biotope definitions from the Marine Habitat Classification for Britain and Ireland (Connor et al, 2004). The introductory text describes marine biological zones in terms of five main factors: immersion, thermal stability, light, salinity and wave action. The latter four factors are described in Table 1 for circalittoral and deep circalittoral/offshore circalittoral zones (both of which are fully immersed).

Table 1: Factors characterising circalittoral and deep circalittoral biological zones around Britain and Ireland in the Marine Habitat Classification for Britain and Ireland (modified from Connor et al, 2004).

Zone	Typical upper boundaries	Thermal stability	Light	Salinity	Wave action
Circalittoral	5 to 20 m	Moderately variable – mesothermal	Mesophotic (sparse algae, algal crusts)	Mesohaline/ Stenohaline	Moderately variable
Circalittoral offshore (or Deep circalittoral)	40 to 80 m	Stable - stenothermal	Aphotic	Stenohaline	Stable

The deep circalittoral zone is characterised by stable temperature and salinity, and a lack of light and wave action. Such conditions result in an animal-dominated environment where fauna that are intolerant of physical disturbance and rapid changes in temperature and salinity, such as fragile, large-bodied, erect, epifauna, dominate.

Mapping the boundary between the circalittoral and deep circalittoral zones requires full coverage layers of the relevant physical factors at a sufficient resolution. For the EUSeaMap modelling process, a temperature layer was not available for the project area and regional salinity data was only available for the Baltic region (see main report: Cameron and Askew, 2011). The position of the halocline was instead chosen to define the boundary in this region because of the highly variable in salinity and reduced wave action in the Baltic Sea. In the western Mediterranean, light alone was chosen to define this boundary because the spatial resolution of wave data was considered too coarse in contrast with the steeply sloping seabed of the Mediterranean's coastal areas, which lead to difficultly distinguishing zones based on data derived by wave models across the whole basin. For more information about how light availability at the seabed is modelled and how this boundary was determined for the western Mediterranean, see EUSeaMap Technical Appendix 1. In the less steeply sloping North and Celtic Seas, wave action alone was chosen to define this boundary. Therefore, this technical appendix describes the mapping of the boundary between the circalittoral and deep circalittoral for North and Celtic Seas only.

1.1.2 Modelling wave action at the seabed

In order to map the boundary between circalittoral and deep circalittoral zones, a threshold value of some variable can be used to define the seabed as either circalittoral or deep circalittoral. When this variable is wave action, as it is in this instance, the boundary is referred to as the wave base, either side of which can be referred to as "disturbed" or "undisturbed". The value chosen to define the position of the wave base at any location is the ratio of wave length of wind-driven waves to water depth, which is hereafter referred to as the wave base ratio. In the field, wave period, the time interval between subsequent crests or troughs of a wave, is measured rather than wave length.

The terms "wave length" and "wave period" refer to the properties of a regular wave with a constant frequency and velocity. However, the sea is made up of an infinite number of waves of different frequencies and velocities superimposed to produce an irregular surface (Michel, 1999). The range of wave frequencies and amplitudes that occur can be described in terms of a wave spectrum, where the energy under the curve is equivalent to the total energy in the system. If the spectrum is drawn in terms of the wave period rather than frequency, the peak wave period (T_p) can be defined as the period of the highest energy regular wave. This property, with depth (h), can be used to calculate the dominant wave length (L) at a specific point (Michel, 1999).

However, it is not T_p that is measured in the field (e.g. by wave buoys), but the zero up-crossing wave period (T_z), which is the time interval between subsequent upward crossings of the wave surface with the mean sea level. T_p can be expressed as a constant linear function of T_z with the value of the constant depending on the sea state. Following Wolf et al (2000), a T_p/T_z ratio of 0.777 was deemed appropriate for the particular spectrum used in this study (see West et al, 2010) and was used to transform the layer of T_z values to a layer of T_p values.

Soulsby (1997) describes a process of deriving L for variables h and T_p and a constant acceleration due to gravity (g), which is summarised here. L is larger for longer T, and smaller for decreasing h and can be described using a dispersion relationship (Equation 1), which is usually expressed in terms of the wave number k and radian frequency ω (Equations 2 and 3):

$$\omega^2 = gk\tanh(kh) \qquad (1)$$

where $$k = \frac{2\pi}{L} \qquad \omega = \frac{2\pi}{T} \qquad (2)\ \&\ (3)$$

It can be seen from the dispersion relationship, that it would be straightforward to extract ω and hence T from this equation if L is known. However, the recurrence of k and therefore L inside and outside of the hyperbolic tangent function means it is less straightforward to obtain L from T. To achieve this, the equation is first simplified by substituting in two non-dimensional intermediate terms, ξ and η:

$$\xi = \frac{\omega^2 h}{g} \qquad \eta = kh \qquad (4)\ \&\ (5)$$

So that Equation 3 becomes:

$$\xi = \eta\tanh(\eta) \qquad (6)$$

Finally, the use of approximations allows Equation 6 to be solvable to within ±0.75 %:

$$\eta = \xi^{\frac{1}{2}}(1 + 0.2\xi) \qquad \text{for } \xi \leq 1 \qquad (7a)$$

$$\eta = \xi(1 + 0.2e^{2-2\xi}) \qquad \text{for } \xi > 1 \qquad (7b)$$

Following the application of either Equation 7a or 7b, the wave base ratio L/h can be extracted from the intermediate term η by substituting in the factors introduced in Equations 2 and 5:

$$\frac{L}{h} = \frac{\left(\frac{2\pi}{k}\right)}{h} = \frac{\left(\frac{2\pi}{\eta/h}\right)}{h} = \frac{2\pi}{\eta} \qquad (8)$$

For more detail about the derivations of wave length from wave period and depth data, refer to Soulsby (1997). When one has values to input into the wave base ratio all that remains to be derived is a threshold value that will define the zone boundary.

1.2 Aims and objectives

The aim of the work described in this report was to produce a map showing the lower limit of the circalittoral zone and the upper limit of the deep circalittoral zone in the North and Celtic Seas. In order to do this, two things were required:

1. Full coverage wave period and bathymetry data layers for the North Sea and Celtic Sea. Data needed to be continuous and comparable within the basin.

2. Values of the ratio between wave length and water depths that define the boundaries between the two depth zones within a “fuzzy membership” range (see main report section 4.3.1 for explanation; Cameron & Askew, 2011)). Threshold values needed to be determined based on statistical comparison with the presence or absence of reference species or habitats that biologically define these zones.

2 Data preparation

2.1 Bathymetry data

The EMODnet bathymetry data set was used as the primary bathymetry (depth) data source, with a resolution of 0.0020833°. General Bathymetric Chart of the Oceans (GEBCO) data was used for those areas outside of the coverage of the EMODnet data.

2.2 Wave data

The primary source of wave data is the National Oceanographic Centre’s ProWAM model (Monbaliu et al, 2000), which covers the area 48°07’N to 62°53’N and 11°50’W to 12°50’E. The model predicts wave height, period and direction at a resolution of 1/6° longitude by 1/9° latitude (~12 km) based on wind measurements, calibrated with *in situ* wave data. Beyond the western limit of the ProWAM model, the water depths are greater than 150 m and consequently, the effect of wave action on the seabed is assumed to be negligible (West et al, 2010). Within 6 km of the coast, the Danish Hydraulic Institute’s high resolution (100-300 m) Spectral Wave Model (West et al, 2010) is used in preference to ProWAM. Both wave height and zero up-crossing wave period (T_z) values were used to calculate kinetic energy at the seabed for the purpose of classifying the energy regimes of rock habitats (see

EUSeaMap Technical Appendix 3), while only T_z values were required in this method. Wind-driven waves were of primary interest rather than swell waves because the latter tend to be of lower amplitude and hence have less of an effect on the seabed. Swell waves were removed based on a wave steepness of 1/7.

Five versions of the T_z raster layers were available based on mean values over 6 years (2000-2005) and corresponding to different percentiles of the wave height data (50 %, 90 %, 98 %, 99 % and 100 %) i.e. the 100th percentile is the T_z value corresponding to the mean maximum wave height from 6 years, 50th percentile is the T_z value corresponding to the mean median wave height from 6 years, and so on. For production of the wave base ratio layer, the 50th percentile (median wave height) and 100th percentile (maximum wave heights) data were tested.

3 Methodology

3.1 Ground-truth data selection and preparation

In order to find an appropriate threshold for the wave base ratio in the North and Celtic Seas, values were tested against biotope presence and absence records from survey data around the UK and Ireland. The biological reference data was extracted from the benthic samples database, Marine Recorder, which is maintained at JNCC. In this database, samples are classified as a particular biotope or habitat according to the Marine Habitat Classification for Britain and Ireland (Connor et al, 2004). These samples were translated to the EUNIS habitat classification scheme and a subset of the data was extracted to provide presence/absence data for wave-disturbed habitats. Sediment habitat data was used without rock data because the rock section of the EUNIS habitat classification scheme only contains two poorly described biotopes referred to as "deep circalittoral" and sediment habitats are assumed to be affected mostly by waves due to their mobile nature. The thresholds analysis was repeated for two different datasets:

1. Dataset 1 included only those samples that had been classified to the equivalent of level 5 or 6 in the sediment section of the EUNIS habitats classification scheme, as these levels contain biological information and allow recognition of particular species that occur in more or less stable sediments. Of 145 samples included in this dataset, 57 (39 %) were classified as deep circalittoral (undisturbed by wave action) and the rest classed as infralittoral or circalittoral (disturbed by wave action). The sample size was relatively low as there are fewer data from diver surveys on sediment and at greater depths corresponding to the deep circalittoral zone. Figure 1 shows that the spatial distribution is strongly governed by where surveys have happened to take place (i.e. a sampling bias), which is likely to lead to spatial autocorrelation in the data.

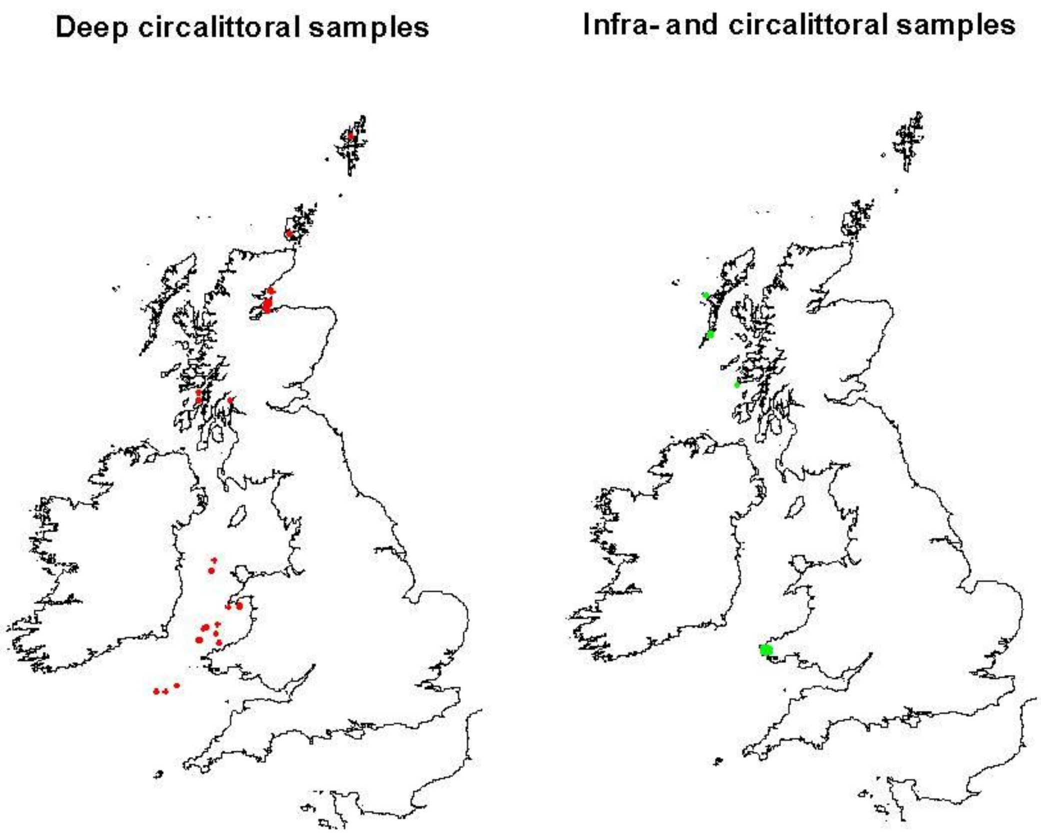

Figure 1: Spatial distribution of deep circalittoral zone samples (red) and infralittoral and circalittoral zone samples (green) used in dataset 1.

2. The analysis was repeated for a larger, but potentially less reliable, subset of the Marine Recorder database (Dataset 2) because Dataset 1 only had a small number of samples. In addition to the level 5 and 6 biotopes included in Dataset 1, Dataset 2 included samples that had been classified to the equivalent of level 4 in the sediment section of the EUNIS habitats classification scheme. At this level in the hierarchy, habitats are only described in terms of sediment type and biological zone. As no species are described, the classification of deep circalittoral or circalittoral by the surveyor may have only been made based on expert judgement of the wave exposure in an area or depth. A total of 2,324 samples were included in this dataset, of which only 355 (15 %) were classed as deep circalittoral (undisturbed by wave action).

Marine Recorder allows users to specify whether a record is "whole" or "partial" and whether the classifier is "certain" or "uncertain". For both datasets, only "whole" and "certain" records were used for the comparison. Values of h and T_p were extracted from the continuous data layers at locations corresponding with the biological samples and attached to the sample data. Finally, the wave base ratio L/h was calculated for each sample according to Equation 8 based on these parameters.

3.2 Modelling probability of presence in the deep circalittoral zone

A generalised additive model (GAM) was used to model the probability that a sample would be classed as deep circalittoral (rather than circalittoral or infralittoral) as a function of h/L. The ratio h/L

is the inverse of the wave base ratio and was used to prevent the occurrence of infinite values for zero values of depth.

Analysis was performed in the open-source statistical software R (R Development Core Team, 2011). A binomial error distribution was assumed because the response variable (presence in the deep circalittoral zone) is binary. A "logit" link function was applied, which transformed the response variable to allow values outside of the 0-1 range, and a single univariate smooth term was applied to the explanatory variable h/L, which enabled the probability to be described as a complex combination of polynomial functions of h/L.

A model of this type is flexible, requiring no *a priori* assumptions about the nature of the relationship between h/L and the probability of a sample being deep circalittoral. However, for this investigation prior assumptions were made, which can be used to restrict the shape of the probability curve. With the understanding that as depth, h, increases and as wave length, L, decreases waves are less likely to reach the seabed, it was assumed that:

1. The probability would always increase as h/L increases, i.e. the relationship is monotonically increasing.
2. Increases in h/L in shallower, more turbulent areas should have a proportionally greater impact than when it is more sheltered. Therefore, the rate of change of probability with h/L would decrease as h/L increased, levelling off as the probability approaches 1, i.e. the relationship is concave.

The R package *scam* (Pya, 2010; based on package *mgcv* by Wood, 2006) allows shape constraints to be placed on a model and so was chosen to run the GAM. Three models were constructed for the 100th percentile wave data, each with different model shape constraints set (Table 2).

Table 2: Summary of model constraints and numbering.

	Model number	
Model shape constraints	**Dataset 1 – level 5 & 6 biotopes**	**Dataset 2 – level 4, 5 & 6 biotopes**
No constraints	1.1	2.1
Monotonically increasing	1.2	2.2
Monotonically increasing and concave	1.3	2.3

The final choice of model was based on the quality of the model fit as well as the two assumptions described above. Before thresholds were determined, a comparison was made between the chosen h/L-based model and a model based purely upon depth, h. This allowed assessment of the relative contributions of depth to the predictive power of the h/L-based model and to ensure that a model based on h/L would produce a better fit than depth alone.

3.3 Defining thresholds

With a model selected, estimated probability values and corresponding h/L values were determined at increments of 5 % probability. The hard threshold for h/L which was to define the midpoint between the lower and upper parts of the fuzzy threshold was chosen as the value that occurs at a probability of 50 %, as above this value it is more likely that a sample will be classed as deep circalittoral and below this value it is more likely that a sample will be classed as infralittoral or circalittoral.

In the EUSeaMap GIS model, membership curves are composed of three straight edges, defined by an upper and a lower threshold that define the transition from one zone to another, with the midpoint between them being equal to the hard threshold described above (Figure 8 in main report; Cameron & Askew, 2011). The membership curve in this form is therefore a simplified version of a more sophisticated model based on probabilities. The upper and lower threshold values for the membership curve were chosen using visual analysis of the probability curve and an understanding of the transitionary nature of the zone boundaries as well as the limitations of the data used to construct the model.

Following this step, the chosen thresholds for the wave base ratio were propagated throughout the North Sea and Celtic Sea regions to delineate the lower boundary of the circalittoral zone.

4 Results

4.1 Modelling probability of presence in the deep circalittoral zone

4.1.1 Dataset 1: EUNIS level 5 and 6 biotopes only

Using Dataset 1 for the biological reference data, Models 1.1 and 1.2 produced non-significant models according to their p-values from χ^2 tests (Table 3), suggesting the best choice of model would be that which is constrained to be increasing and concave (Model 1.3 (maximum wave heights) and Model 1.4 (median wave heights)). However, the explanatory power of the model is decreased by 46 % with the addition of these two constraints. As a result of this and the small sample size (145), Dataset 2 was also used to produce models using the same constraints.

Table 3: Comparison of model fit with different shape constraints to the relationship between the probability a sample is classified as deep circalittoral and *h/L* for maximum wave heights (Models 1.1-3) and median wave heights (Model 1.4).

			For Smooth Term		
Model (variable, wave data used, constraints)	**UBRE Score**	**Deviance Explained**	**Estimated Degrees of Freedom**	χ^2	**P**
Null model	0.354	0.0 %	NA	NA	NA
1.1.*h/L*, *max* wave heights, no constraint	-0.269	53.7 %	7.0	6	0.60
1.2. *h/L*, *max* wave heights, monotonically increasing	-0.065	34.2 %	2.9	0.5	0.90
1.3. *h/L*, *max* wave heights, monotonically increasing & concave	-0.018	28.8 %	1.0	23	1.52×10^{-6}
1.4. *h/L*, *median* wave heights, monotonically increasing & concave	0.050	23.7 %	1.0	22	2.41×10^{-6}

4.1.2 Dataset 2: EUNIS levels 4, 5 and 6

Using Dataset 2 for the biological reference data, all sets of model constraints produced significant models according to their p-values from χ^2 tests and all help explain more than half of the deviance in the model (Table 4 and Figure 2).

Table 4: Comparison of fit of models with different shape constraints to the relationship between the probability a sample is classified as deep circalittoral and *h/L* (Models 2.1-3) and *h* only (Model 2.A).

			For Smooth Term		
Model (variable, wave data used, constraints)	**UBRE Score**	**Deviance Explained**	**Estimated Degrees of Freedom**	**χ^2**	**P**
Null model	-0.162	0.0 %	NA	NA	NA
2.1. *h/L*, *max* wave heights, no constraint	-0.672	61.7 %	8.5	334	6.25×10^{-67}
2.2. *h/L*, *max* wave heights, monotonically increasing	-0.669	61.1 %	5.4	350	3.66×10^{-73}
2.3. *h/L*, *max* wave heights, monotonically increasing & concave	-0.670	61.0 %	3.0	370	6.84×10^{-80}
2.4. *h/L*, *median* wave heights, monotonically increasing & concave	-0.655	51.2 %	3.0	294	1.55×10^{-63}
2.A. *h only*, monotonically increasing & concave	-0.702	64.8 %	3.0	401	1.60×10^{-86}

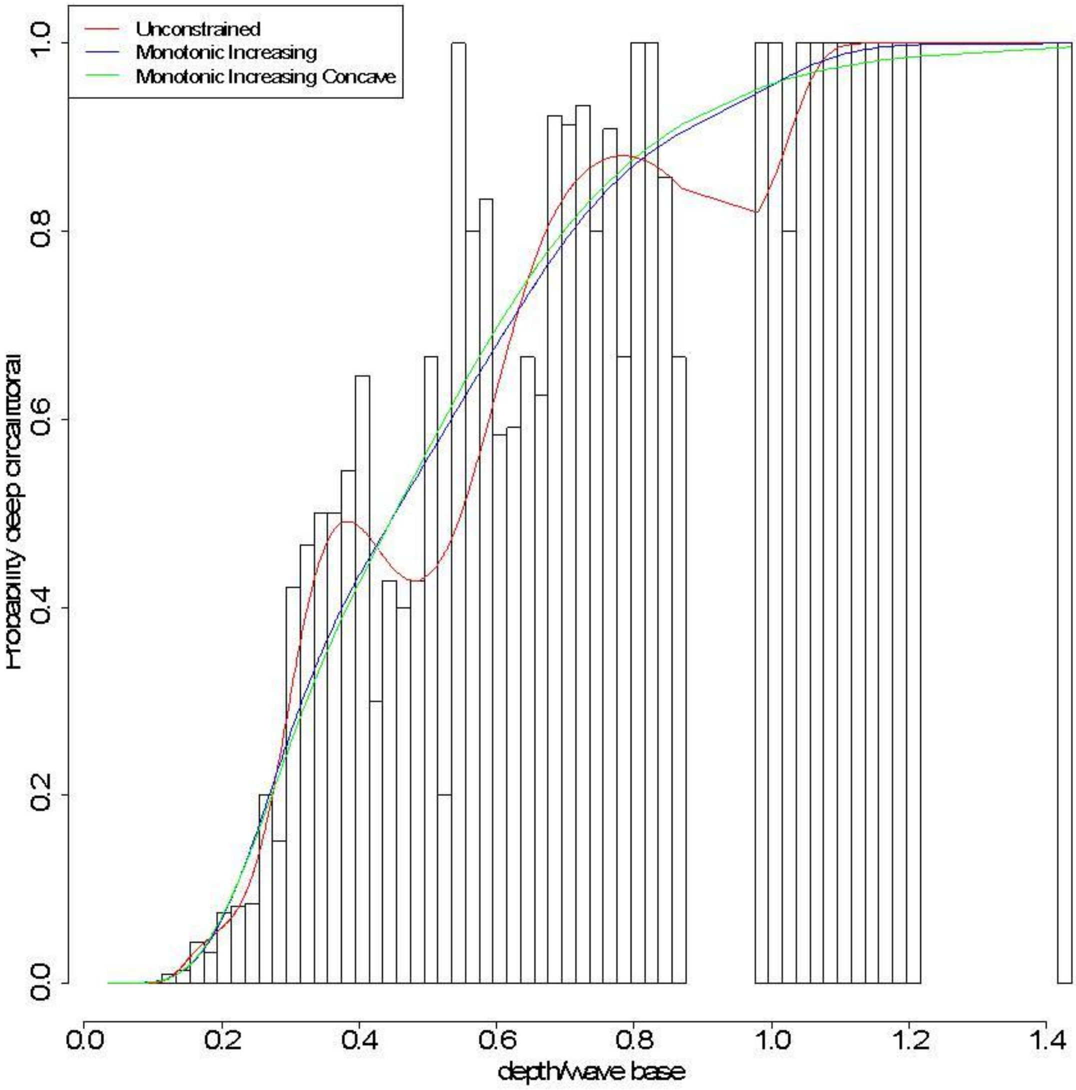

Figure 2: Relationship between probability a sample is classified as deep circalittoral and h/L, showing three alternative fitted models (2.1 to 2.3), with different constraints on the shape of the relationship. Bars show proportion of records that are deep circalittoral for each 0.02 wide interval of h/L.

The p-values in Table 4 suggest that Model 2.2 (monotonically increasing) has a marginally better fit than Model 2.1 (no constraints). Model 2.1 (red line in Figure 2) fits closely to the data causing the change in probability to decline in some places as *h/L* increases. As described in Section 3.2, the probability a sample is classified as deep circalittoral would be expected to increase with increasing *h/L*. Therefore, the fact that the data does not provide a steadily increasing proportion of deep circalittoral data as *h/L* increases suggests that this is almost certainly due to sampling error and gaps in the data (e.g. for $0.88 < h/L < 0.98$) and that Model 2.1 should be rejected.

According to the assumptions outlined in Section 3.2, a monotonically increasing and concave relationship might be expected to best represent the true relationship between *h/L* and the probability of presence in the deep circalittoral zone. Table 4 shows that the model fit and amount of deviance explained is very close for Model 2.2 and Model 2.3. Furthermore, the shape of the probability curves (Figure 2) are almost identical for the low values of *h/L* that are of most interest and define the transition from circalittoral to the deep circalittoral zone. As a result, Model 2.3 (blue line in Figure 2) would be preferred over Models 2.1 and 2.2 to define thresholds for *h/L* based on the *a priori* assumptions about the relevance to biology.

A comparison of Model 2.3 with the equivalent model using the median wave height data to produce wave length values (Model 2.4) suggests the former is a better fitting model with lower p-values and more deviance explained.

Finally, Model 2.3 was compared with a model using a monotonically increasing and concave function of water depth, *h*, alone (Model 2.A) to predict the probability a sample is deep circalittoral (Table 4 and Figure 3). Model 2.A explained 65 % of the deviance, compared to Model 2.3, which explained only 61 %. A χ^2 test also produces a lower p-value for Model 2.A than for Model 2.3; both of these factors suggest that a model based on depth alone fits better to the observations than a model that also takes wave length into account.

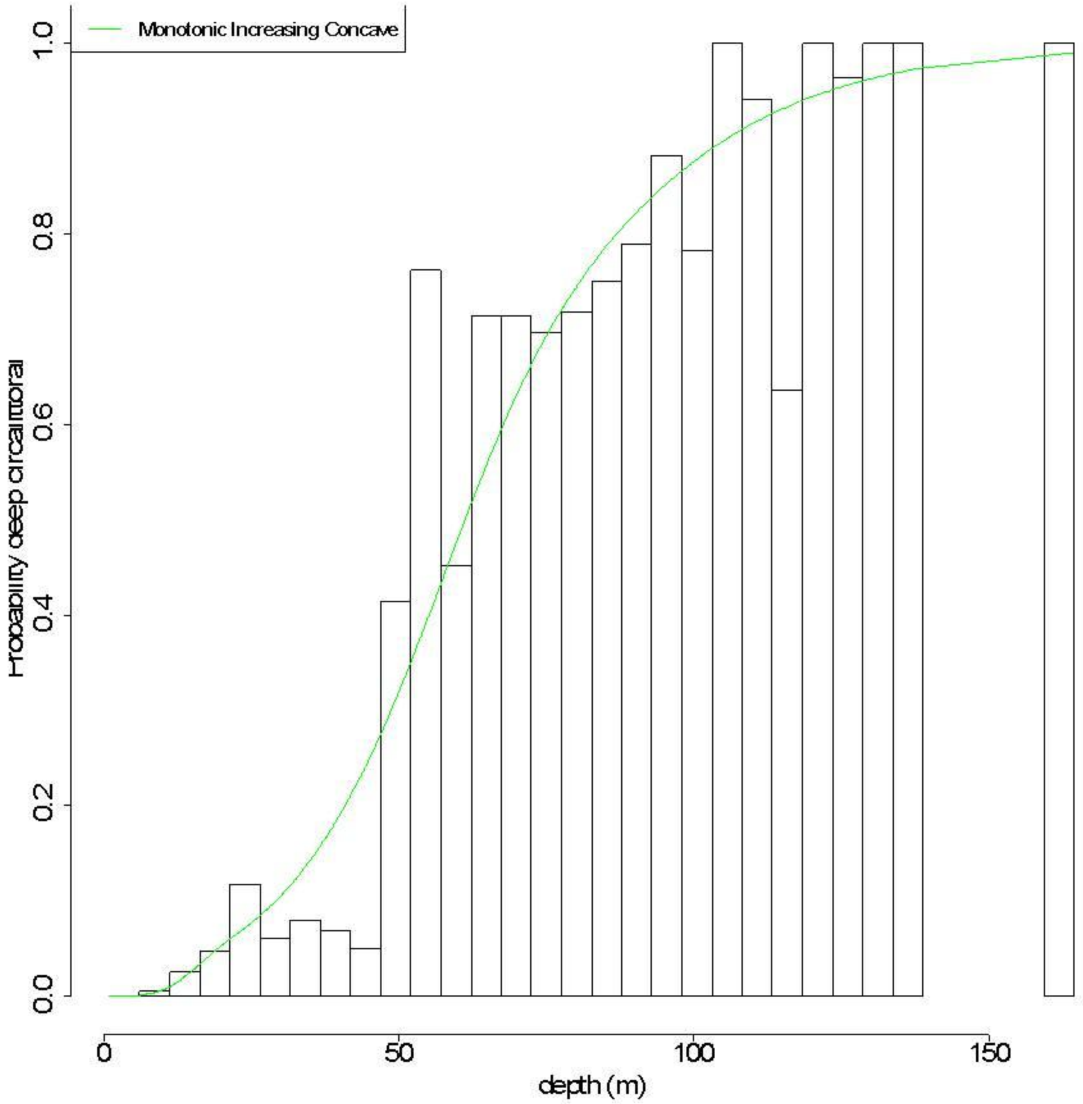

Figure 3: Relationship between probability a sample is classified as deep circalittoral and depth, *h*, for a model constrained to be monotonically increasing and concave (Model 2.A). Bars show proportion of records that are deep circalittoral for each 5 m wide interval of *h*.

In summary, the explanatory power of the *h/L* model is stronger using maximum wave height data compared with median wave height data. However, it is stronger still using only depth, not taking wave length into account at all.

4.2 Defining thresholds

Thresholds were calculated for several models and then compared because of the similarity in fit of the various models. For each model, a set of *h/L* values and equivalent wave base ratios (inverse of *h/L*) were characterised for the range of probability values produced in the model, evenly spaced at 5 % intervals (e.g. Table 5).

Table 5: Characterisation of different threshold values for the wave base ratio *L/h* based upon Model 2.3.

Predicted probability	Threshold value of	Equivalent wave base	Total number of	Number of deep circa	% of all samples	% of deep circa	% of samples above

sample is deep circalittoral	h/L	ratio (L/h)	samples above threshold	samples above threshold	above threshold	samples above threshold	threshold that are deep circalittoral
0.05	0.18	5.4	753	327	32 %	95 %	43 %
0.1	0.22	4.6	581	317	25 %	92 %	55 %
0.15	0.25	4.0	492	309	21 %	90 %	63 %
0.2	0.27	3.7	458	303	20 %	88 %	66 %
0.25	0.30	3.4	424	298	18 %	87 %	70 %
0.3	0.32	3.1	395	284	17 %	83 %	72 %
0.35	0.35	2.9	381	278	16 %	81 %	73 %
0.4	0.38	2.6	363	269	16 %	78 %	74 %
0.45	0.42	2.4	337	253	14 %	74 %	75 %
0.5	**0.45**	**2.2**	**297**	**241**	**13 %**	**70 %**	**81 %**
0.55	0.49	2.1	274	230	12 %	67 %	84 %
0.6	0.52	1.9	267	227	11 %	66 %	85 %
0.65	0.56	1.8	255	219	11 %	64 %	86 %
0.7	0.60	1.7	238	208	10 %	61 %	87 %
0.75	0.65	1.6	209	188	9 %	55 %	90 %
0.8	0.70	1.4	172	161	7 %	47 %	94 %
0.85	0.76	1.3	129	122	6 %	36 %	95 %
0.9	0.84	1.2	109	105	5 %	31 %	96 %
0.95	0.97	1.0	99	97	4 %	28 %	98 %

For Dataset 2, 15 % of all samples were classified as deep circalittoral. This means that, if the model were 100 % accurate in its predictions, so that 100 % of all deep circalittoral samples were above the threshold value of *h/L* and 100 % of samples above the threshold were deep circalittoral, then 15 % of all samples would be included above the threshold value. Conversely, if we predicted at random which samples were infralittoral then we would still expect to be correct 15 % of the time.

As explained in Section 3.3, a probability of 50 % has been chosen to define the most suitable threshold value of *h/L* (row highlighted bold in Table 5), above which samples are more likely to be deep circalittoral than circalittoral according to the model. Values corresponding to a probability of 50 % are compared in Table 6 for various increasing and concave models described in the previous section.

Table 6: Comparison of thresholds for different explanatory variables corresponding to a deep circalittoral probability of 50 %.

Model #	Explanatory variable (wave height percentile)	Threshold wave base ratio (2 sig. figures)	Total number of samples above threshold	Number of deep circa samples above threshold	% of all samples above threshold	% of deep circa samples above threshold	% of samples above threshold that are deep circalittoral
1.4	*L/h* (50)	*L/h* = 2.2	33	24	23 %	42 %	73 %
1.3	*L/h* (100)	*L/h* = 4.4	38	29	26 %	51 %	76 %
2.4	*L/h* (50)	*L/h* = 1.1	245	158	11 %	64 %	64 %
2.3	*L/h* (100)	*L/h* = 2.2	297	241	13 %	70 %	81 %

2.A	h	h = 61 m	314	265	14 %	77 %	84 %

The best fitting model, Model 2.3, gives a hard threshold wave base ratio of L/h = 2.2. This threshold includes 70 % of deep circalittoral samples, and 81 % of samples satisfying this threshold are deep circalittoral. The depth-only model, Model 2.A, gives a hard threshold value of h = 61 m; this threshold includes 77 % of deep circalittoral samples, and 84 % of samples satisfying this threshold are deep circalittoral.

5 Discussion

Several problems and challenges were encountered in the process of defining the boundary between the circalittoral and deep circalittoral zones in the North and Celtic seas:

1. **Quantity and quality of biological reference data**: ideally, a statistical analysis would only include biotopes that are described in terms of their species assemblages (levels 5 and 6 in EUNIS). This is to ensure that any analysis would be between the environmental variables and the biology, rather than potentially being between environmental variables and perceived environmental variables in the field, as may be the case for samples classed to level 3 or 4 of the EUNIS habitat classification scheme. However, as the sample size for the biology-based dataset (dataset 1) was only 165, the sample size was increased to include level 4 biotopes to the possible detriment of the quality of the reference data. Furthermore, the larger dataset (dataset 2; 2,324 samples) only contained 15 % deep circalittoral samples and suffered from data gaps, e.g. for $0.88 < h/L < 0.98$ (Figure 2).

2. **Definition of 'circalittoral' and 'deep circalittoral'**: the four factors (temperature, light, salinity and wave action) that vary between the two zones according to the classification scheme that defined them (Connor et al, 2004) are given in Table 1. The interactions of these four factors can result in a range of different environmental conditions in between [mesothermic, mesophotic, meso/stenohaline, moderately variable wave action] and [stenothermal, aphotic, stenohaline, stable wave action] resulting in a potentially wide transition zone within which species that are sensitive to some variables but not others occur. This transition zone is not reflected in the classification system and as such all must be classified as occurring in one zone only. Furthermore, there may also be biotopes that are tolerant to variations in all four of the variables, but because of the structure of the classification, must be classified as occurring in one zone only. In addition, many assemblages in nature transition from one to another along an environmental gradient with the loss or addition of a few characterising species and do not shift with thresholds. These factors inevitably result in a boundary that is not clear-cut and biological reference data that may not be directly linked to the variables under investigation.

3. **Availability of physical data**: the modelling process may have been improved by the inclusion of temperature, salinity and light data to reflect the definition of the zones (described above in 2).

4. **Quality of physical data**: the relatively low resolution of the depth and wave data compared with the precise positions of the biotopes inevitably means that incorrect h and L values will have been attributed to some samples. This is especially true for the wave data, with a resolution of only ~12 km in some places, with a maximum resolution of 100 m near the coast.

5. **Definition of the most representative wave statistics**: wave data was available in several percentiles to describe the mean conditions over a six-year period (see Section 2.2). In this investigation a stronger model was achieved with the wave period values associated with the maximum heights per year than with the median heights per year. This suggests that the most extreme annual conditions drive the community composition and distribution, reinforcing previous hypotheses on disturbance theory (for example, Rees et al, 2007).

Considering the outputs of the analysis in Section 4 and the discussion above, it was decided that the modelling process in this case was not robust enough to alter the convention (Coltman et al, 2008; McBreen et al, 2011) of using a wave base ratio of *L/h* = 2 to define the hard threshold for boundary between the circalittoral and deep circalittoral zones. This value is, however, close to the threshold produced using Model 2.3 (*L/h* = 2.2; Table 6), adding confidence to this decision. However, an indication of uncertainty in the chosen hard threshold is also required for the EUSeaMap model (upper and lower fuzzy thresholds). This was chosen to be ±25 %, which was considered appropriate to indicate the relatively large uncertainty in this threshold. Therefore, the fuzzy threshold chosen to define the boundary between the circalittoral and deep circalittoral biological zones is $1.5 < L/h \leq 2.5$.

6 Conclusion

Hard and fuzzy thresholds have been produced to describe the lower boundary of the circalittoral zone in the North and Celtic Seas in terms of wave disturbance at the seabed. A threshold analysis was carried out using the best data that was available. Difficulties arose due to a combination of: a lack of complete biological reference data at an appropriate resolution; relatively low resolution for depth and, in particular, wave data; and the complicated definition of the circalittoral and deep circalittoral zones in the habitat classification scheme. As a result, it was decided to remain using the conventional threshold of a wave base ratio of 2, with a fuzzy range of 1.5 to 2.5.

This analysis may be improved in future by:

1. Comparing and determining the most appropriate wave statistics to represent the biological response
2. Further investigation of the effect of light, temperature and salinity
3. A more consistent approach in all EUSeaMap regions in terms of variable selection and modelling procedure.

7 References

Cameron, A. and Askew, N. (eds.). 2011. EUSeaMap - Preparatory Action for development and assessment of a European broad-scale seabed habitat map final report. Available online at http://jncc.defra.gov.uk/euseamap [Accessed 8th August 2012]

Coltman, N., Golding, N. & Verling, E. 2008. Developing a broadscale predictive EUNIS habitat map for the MESH study area. Available online at http://www.searchmesh.net/pdf/MESH%20EUNIS%20model.pdf [Accessed 8th August 2012]

Connor, D. W., Allen, J. H, Golding, N., Howell, K.L., Lieberknecht, L. M., Northen, K. O. and Reker, J. B. 2004. The Marine Habitat Classification for Britain and Ireland Introduction Version 04.05 JNCC, Peterborough ISBN 1 861 07561 8 Available online at http://jncc.gov.uk/marinehabitatclassification [Accessed 8th August 2012]

McBreen, F., Askew, N., Cameron, A., Connor, D., Ellwood, H. & Carter, A. 2011. UKSeaMap 2010: Predictive mapping of seabed habitats in UK waters. *JNCC Report No. 446*. Available online at http://jncc.defra.gov.uk/ukseamap [Accessed 8th August 2012]

Michel, W. H. (1999). Sea Spectra Revisited. Marine Technology, vol. 36, no. 4, Winter 1999. Pp. 211-227.

Monbaliu, J., Padilla-Hernández, R., Hargreaves, J., Carretero Albiach, J. C., Luo, W., Sclavo, M., & Günther, H., (2000) The spectral wave model, WAM, adapted for applications with high spatial resolution, Coastal Engineering, 41, 41-62.

Pya, N. 2010. *Additive models with shape constraints*. PhD thesis. University of Bath. Department of Mathematical Sciences. Available online at http://opus.bath.ac.uk/27546 [Accessed 9th September].

R Development Core Team. 2011. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rees, H. L., Eggleton, J. D., Rachor, E., and Vanden Berghe, E. (Eds). 2007. Structure and dynamics of the North Sea benthos. ICES Cooperative Research Report No. 288. 258 pp. Available online at http://www.ices.dk/pubs/crr/crr288/CRR%20288.pdf [Access 10th September 2012]

Soulsby, R. L. (1997). Dynamics of marine sands. A manual for practical applications. Thomas Telford, London.

West, N., Swift, R., & Bell, C., 2010. Accessing and developing the required biophysical datasets and data layers for Marine Protected Areas network planning and wider marine spatial planning purposes: Report No 10: Task 2E. Models of Fetch and Wave Exposure. Available online at http://randd.defra.gov.uk/Document.aspx?Document=MB0102_9939_TRP.pdf [Accessed 10th September 2012]

Wolf, J., Hargreaves, J.C. & Flather, R.A. (2000). Application of the SWAN shallow water wave model to some U.K. coastal sites, POL Report no. 57. Available online at http://nora.nerc.ac.uk/3917/1/ir57.pdf [Accessed 10th September 2012]

Wood S.N. 2006. *Generalized additive models, an introduction with R.* Chapman & Hall/CRC, London