Prifysgol Abertawe
Swansea University

Swansea University

Habitat Suitability modelling for informing Zostera marina restoration in Wales

A report for WWF

Bertelli C.M., Bennett W., Bull J.C., Karunarathna, H., Reeve D.E., Unsworth R.K. F

WWF

seagrass ocean rescue

SEACAMS

Contents

List of Tables and Figures

1. Introduction

Seagrasses have been identified as significant contributors to climate change mitigation and adaptation (UNEP, 2020) as natural carbon sinks (Fourqurean et al., 2012), by stabilising coastal sediments and reducing coastal erosion (Potouroglou et al., 2017). Unfortunately, seagrasses are being lost at an alarming rate due to a number of pressures from direct physical destruction to water quality issues. In areas that were once covered by dense swathes of seagrass, it is so often the case that seagrass has declined or disappeared altogether. In recent decades many countries have made substantial changes to regulate and improve the health of coastal waters through water treatment and monitoring of water quality. Seagrasses are also recognised as being sentinels of healthy coastal ecosystems and have been widely implemented into conservation legislation. In some areas, changes have led to a recovery of seagrass meadows to some extent, but in areas where seagrass has significantly declined or been lost altogether, restoration is a viable route of action to aid recovery. Seagrass restoration has also been considered as a method of increasing carbon capture in the same way that planting trees is used for offsetting carbon emissions. Around the UK seagrass restoration is still in its relative infancy, but in other locations restoration has already been successfully achieved at a large scale, such as in Chesapeake Bay, USA (Shafer and Bergstrom, 2008). The benefits of previous research undertaken in Chesapeake Bay have been taken on board and adapted for smaller scale restorations projects within the UK, such as in Dale, west Wales (Unsworth et al., 2019). Site choice for restoration needs to be carefully considered to maximise success. Trying to restore areas where water quality is still an issue or where other pressures are still present will compromise plant growth and/or establishment.

Habitat suitability modelling (HSM) is becoming an increasing useful tool for determining species distributions and predicting distributions within spatial and temporal timeframes. These models can be used to predict species distribution at a spatial scale by identifying areas where environmental conditions are suitable. Also known as species distribution models (SDM) and ecological niche models (ENM), they all follow the same premise, and can be used for different aims. For example, if budgets for habitat surveys is limited, an SDM/HSM can be used to focus efforts to areas where the habitat or species is most likely to be found, using a non-subjective method. This method can also be used at a finer scale to determine the most suitable area for habitat restoration, so that chances of success can be maximised. This type of modelling uses environmental parameters which determine or affect the growth and presence of the species in question and then projects this to predict where that species should be able to exist if these parameters are met. To be able to do this successfully, the availability of good presence or distribution and environmental parameters is needed to be able to provide models with enough data for testing and training. Environmental parameters will be made up of the conditions that limit the growth and range of the species being studied. For *Zostera marina*, as a submerged aquatic vegetation, these parameters will include light availability, depth (which affects light availability through the water column), temperature, salinity, and physical factors such as exposure to wave energy and currents. *Z. marina* grows in shallow, sheltered areas around the coast of the UK. These locations are predominantly east or north-east face bays where they are sheltered from prevailing wind directions and therefore wave fetch, or within estuaries as they can cope with a range of salinities. A review of previous seagrass HSM studies found the most commonly used environmental parameters to be temperature, bathymetry, light availability, salinity, wave action, substrate and slope. However, the variables ultimately used will also be somewhat dependent upon what environmental data is readily available. Light availability, specifically Photosynthetically Active Radiation (PAR) which is the spectral range of light that is used for photosynthesis, is arguably one of the most important variables to consider as this will determine where plants can survive within coastal waters. Bathymetry is also important as it will influence light availability, as light is attenuated with

depth. Around Europe, *Z. marina* is usually found within a narrow depth range around the coast, typically up to 5-10 m deep depending on water clarity (Davison and Hughes, 1998; Jackson et al., 2013; Krause-Jensen et al., 2003; Nielsen et al., 2002). *Z. marina* is one of the most widely spread temperate species of seagrass across a wide temperature range from -1°C in Arctic regions to 30°C in the subtropics so it is well within its range around the coast of the UK. However, temperature will affect respiration rates within plants and will have significant influence on life stages, such as flowering and germination. *Z. marina* is also tolerant to a range of salinities and can be found within estuaries as well as fully oceanic conditions, from 18 psu to 40 psu (D'Avack et al., 2019).

One of the most significant physical factors determining seagrass presence is wave exposure, with *Z. marina* favouring bays sheltered from prevailing winds and waves where the substrate is sandy-muddy to accommodate its extensive root system (Beca-Carretero et al., 2019). Seagrasses will be able to grow more efficiently forming denser meadows with taller canopies where wave exposure is lower. Wave energy and wave height also has implications for seed burial and seedling development and is therefore one of the most important environmental variables to consider when choosing restoration sites. Successful seedling establishment has been found correspond with lowest maximum wave heights in Chesapeake Bay restoration experiments (Marion et al., 2020).

The aims of this study are to investigate the use of habitat suitability modelling for predicting the best locations for potential seagrass restoration around Wales and the benefits of utilising high resolution wave model data to give better predictions at fine-scale geographical areas. The use of the high-resolution environmental data is expected to improve HSM maps for predicting sites for seagrass restoration. The outcomes of this study will help WWF to make decisions on the best potential areas for seagrass restoration and those places where restoration is most likely to be successful based upon the environmental factors modelled. If wave modelled data improves results, there is a case for using them for other potential restoration locations around the UK.

2. Methods

One of the most important factors for successful HSM is the availability of good quality species presence data. Inaccuracies in distribution data will lead to constrained models (Araújo et al., 2019). Presence data needs to consist of geographical locations of species and therefore point data (coordinates) is most useful and easiest to use. Data was obtained from sources in Table 1 and were uploaded into QGIS (ver. 3.16). This allows visualisation of data layers and identification of erroneous points were subjectively removed such as points on land and in deep ocean where occurrence of *Z. marina* is highly unlikely. Presence data combined from numerous sources enabled a good representation of species distribution around the UK but also included numerous duplicates of coordinates which were removed. Once data was refined, over 3000 presence points remained around the UK, Ireland, and Channel Islands. This data was used for testing suitable HSM methods that at a larger scale which would then be refined at a smaller spatial scale for potential restoration sites.

Table 1. Seagrass presence data sources and data type.

Data Source	Data Type
Natural Resources Wales (NRW) data on LLE Welsh government environmental data portal http://lle.gov.wales/home	Point and Polygon
EMODnet – Seabed habitats – EUNIS fine scale habitat maps (although layer includes all Zostera species) http://www.emodnet.eu/geoviewer/	Polygon
National Biodiversity Network (NBN) Atlas Wales https://wales-records.nbnatlas.org/occurrences/search?q=lsid:NHMSYS0021060374&fq=cl28:Wales#tab_mapView	Point
Botanical Society of Britain and Ireland - Included on NBN portal	Point
Seasearch Marine Conservation society - included on NBN portal	Point
Joint Nature Conservancy Council (JNCC) - habitat point data	Point
South East Wales Biodiversity Records Centre - included on NBN portal	Historical (Point)
Skomer Marine Reserve – reports	Polygon and points
United Nations Environment Programme (UNEP)	Point and polygon occurrence
Seagrass spotter – Project Seagrass	Point
Marine Environmental Data and Information Network (MEDIN)	Polygon

2.1. Environmental predictor variables for Zostera marina

For our initial HSM models, environmental variable layers and downloaded from open access sources and to the finest scale resolution available (Table 2). Available open-source environmental data comes in different resolutions, formats and formed using different methods. For example, photosynthetically available radiation (PAR) at the seabed (From EMODNet via EUSeaMap) is determined from field and satellite data for light in the water column and then calculating light attenuation from depth and proximity to coast (EUSeaMap, 2012). Kinetic energy at the seabed from currents and waves comes from modelled data provided from the National Oceanographic Centre (NOC) with forecast data available from Copernicus.eu for salinity and temperature. All environmental data layers were uploaded into QGIS and clipped to an area encompassing the coastal waters around the UK, Ireland and Channel Islands. Marine environmental datasets are often lacking coverage in shallow coastline and pixels covering both land and sea are often excluded (Yesson et al., 2015). To overcome this, a buffering tool can be used in QGIS to interpolate data from neighbouring pixels to cover coastal areas.

Table 2. Open-source environmental variables available for use in habitat suitability modelling.

Predictor variables	Source	Unit/file type	Spatial resolution	Temporal resolution
Light (PAR) availability at seabed	EMODnet - PAR https://www.emodnet-seabedhabitats.eu/access-data/launch-map-viewer/	Mol.phot.$m^{-2}.d^{-1}$ GeoTiff	~0.3km^2, 1.1 x 0.7 km (0.003°/10 arc seconds)	01 Jan2005 – 31 Dec 2009
Substrate	EMODnet Geology https://www.emodnet-geology.eu/	Folk 5 – 16 classification (.gdb)	(1:1,000,000)	N/A
Bathymetry	EMODnet Bathymetry and topography http://www.emodnet.eu/emodnet-maps-catalogue	Ascii raster (.asc) Xyz format	~0.115km (0.001°x 3.75 arc seconds) ~70m x 116m	N/A
Bathymetry	National Oceanographic centre, British Oceanographic Data Centre BODC (2020) https://www.bodc.ac.uk/data/ Data from single/multibeam, LIDAR, seismic surveys etc	NETcdf (.nc) GeoTiff ASCII	15 arc seconds 200x200m	N/A
Salinity	Copernicus Marine Environment Monitoring Service http://marine.copernicus.eu NorthWestShelf_ANALYSIS_FORCAST_PHY_004_013	Ppt NETcdf (.nc)	0.017° x 0.017°, ~1.7km x 1.7 km	2017-present hourly-instantaneous daily- mean
Salinity	Copernicus Marine Environment Monitoring Service http://marine.copernicus.eu NORTHWESTSHELF_ANALYSIS_FORECAST_PHYS_004_001_B	Ppt NETcdf (.nc)	0.11°x6.6°, ~11 km x 660 km	2018-present Daily-mean
Temperature and potential temperature at seabed	Copernicus Marine Environment Monitoring Service http://marine.copernicus.eu NorthWestShelf_ANALYSIS_FORCAST_PHY_004_013	°C NETcdf (.nc)	(0.017° x 0.017°) ~1.7km x 1.7 km	2017-present hourly-instantaneous daily- mean
Temperature potential temperature at seabed	Copernicus Marine Environment Monitoring Service http://marine.copernicus.eu NORTHWESTSHELF_ANALYSIS_FORECAST_PHYS_004_001_B	°C NETcdf (.nc)	0.11°x6.6°, ~11 km x 660 km	2018-present Daily-mean
Salinity and Temperature seabed and sea surface observation data	Marine Scotland climate Climatology of Surface and Near-bed Temperature and Salinity on the North-West European Continental Shelf https://data.marine.gov.scot/dataset/climatology-surface-and-near-bed-temperature-and-salinity-north-west-european-continental	°C/ppt (.csv)	1/6 longitude by 1/10 latitude	1971–2000 – monthly averages
Coefficient of light attenuation in water	EMODnet - KDPAR https://www.emodnet-seabedhabitats.eu/access-data/launch-map-viewer/	m^{-1} GeoTiff	3km	01 Jan2005 – 31 Dec 2009
Wave energy at seabed	EMODnet https://www.emodnet-seabedhabitats.eu/access-data/launch-map-viewer/	$N.m^2.s^{-1}$ TIFF	~0.3km (~342x195m)	29 Feb 2016-30 Mar 2018

Energy at seabed due to currents	EMODnet https://www.emodnet-seabedhabitats.eu/access-data/launch-map-viewer/ Data from EU SeaMap (2016) Energy in the Celtic Sea and North Sea	$N.m^2.s^{-1}$ TIFF	~0.3km (~342x195m)	01 Jan 2001-01 Jan 2010
Slope	Calculated from bathymetry data with GIS software.	TIFF	Varies with input layer.	N/A

2.2. Broad-scale habitat suitability models

Initial habitat suitability models were developed using open-source environmental data layers and *Z. marina* presence data for the whole of the British Isles and Ireland. This was to test the effectiveness of model methods using the available environmental data. Formatted environmental data layers were uploaded into R as raster files along with *Z. marina* presence point coordinate data as spatial data frames. All environmental data layers were resampled, so they were all the same resolution, extent and format. A suite of methods were used for habitat suitability modelling using the 'sdm' package (Naimi and Araújo, 2016) in R (RStudio, version 4.0.2). This package allows the use of a wide range of the most common modelling algorithms covering parametric, non-parametric, regression and machine-learning methods to be used all at once. A range of algorithms were chosen to cover the different types of modelling approaches. These included Generalised Linear Models (GLM), Generalised Additive Models (GAM), Multivariate Adaptive Regression Splines (MARS), Random Forest (RF), Boosted Regression Trees (BRT) and Maximum Entropy (MaxEnt), to include some of the most popular methods in other HSM studies (Guisan et al., 2017; Valle et al., 2013). The models followed the formula of seagrass presence (coordinates) as a function of the predictor variables (a stack of the environmental variable layers) using the different approaches, repeated twice per approach. As we only had presence data, the modelling takes a range of 'background' points which are treated as pseudoabsences. The formula used takes a proportion of the presence point data to be used to test the model using bootstrapping method. Predictions of presence are then created for each approach (GLM etc.) and each run (repetition). Finally, an ensemble of all the methods was used to create prediction outputs in the form of a raster layer indicating probability of suitability for seagrass presence from 0-1. The ensemble method is useful as it allows various models to be combined to create an average of all the model outputs and removes the need to choose a single method.

2.3. Fine-scale habitat suitability models

Locations for potential restoration sites around Wales were identified as the areas in Figure 1. These locations were chosen as being potentially suitable due to existing or historical records of seagrass presence and comprising areas with appropriate substrate types such as sandy bays.

Figure 1. Sites for potential Zostera marina restoration around Wales used for high resolution habitat suitability modelling using high resolution wave model data. Areas start at Llyn Peninsula, north Wales, followed by west Anglesey coast, east Anglesey and finally south Wales, Pembrokeshire (area 4).

2.4. High resolution wave data

To provide high resolution wave data at each of the four study sites, the computational coastal modelling suite Delft3D was utilised (Lesser et al., 2004). The Delft3D-WAVE module utilises the spectral third generation SWAN model (Booij et al., 1999) to transform wave conditions in time and space, for a range of boundary conditions. A 1160 m x 1850 m numerical model domain encompassing the Irish Sea (Figure 2) was created to transform hourly European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 offshore wave conditions (Hersbach et al., 2013). High resolution near-shore domains were created for each of the four sites, with approximately 50 m x 50 m resolution covering the areas of interest, and coarser resolution offshore (Figure 3). Bathymetry for each of the high-resolution domains was created using EMODnet data, with the General Bathymetric Chart of the Oceans (GEBCO) dataset used to provide additional bathymetry for the larger domain. Each of the boundaries for the larger model were forced with spatially and temporally varying wave parameters (significant wave height and peak wave period). To provide atmospheric forcing, 0.5° resolution hourly ERA5 wind data was implemented across all of the model domain. Within the modelling framework, each of the four high resolution domains were nested within the larger domain, to enable transformation from coarser resolution to the desired 50 m x 50 m grid.

Validation of wave modelling was carried out through comparison with wave buoy observations at three locations within the larger domain (Figure 3). Centre for Environment, Fisheries and Aquaculture Science (CEFAS) W|VENET wave buoys at Liverpool Bay, West Pembrokeshire, and Scarweather provided hourly observations for differing contexts around the Welsh coast. The period 1st March – 31st March 2016 was utilised for model validation, due to consistent observations at each of the three buoys. Comparing the significant wave height at each location with the corresponding model output yielded R^2 values of 0.87 for the West Pembrokeshire and Scarweather buoys, and 0.93 for the Liverpool Bay buoy.

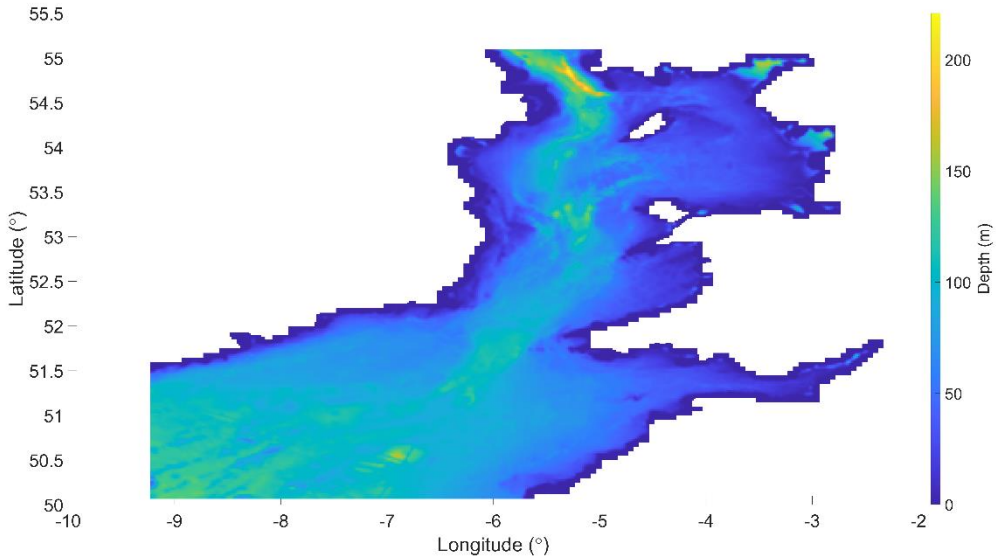

Figure 2. Delft3D Large Irish Sea domain and bathymetry, encompassing the Irish Sea and extending towards the Atlantic. Grid cells are 1160 m x1850 m resolution with bathymetry from GEBCO and EMODnet datasets.

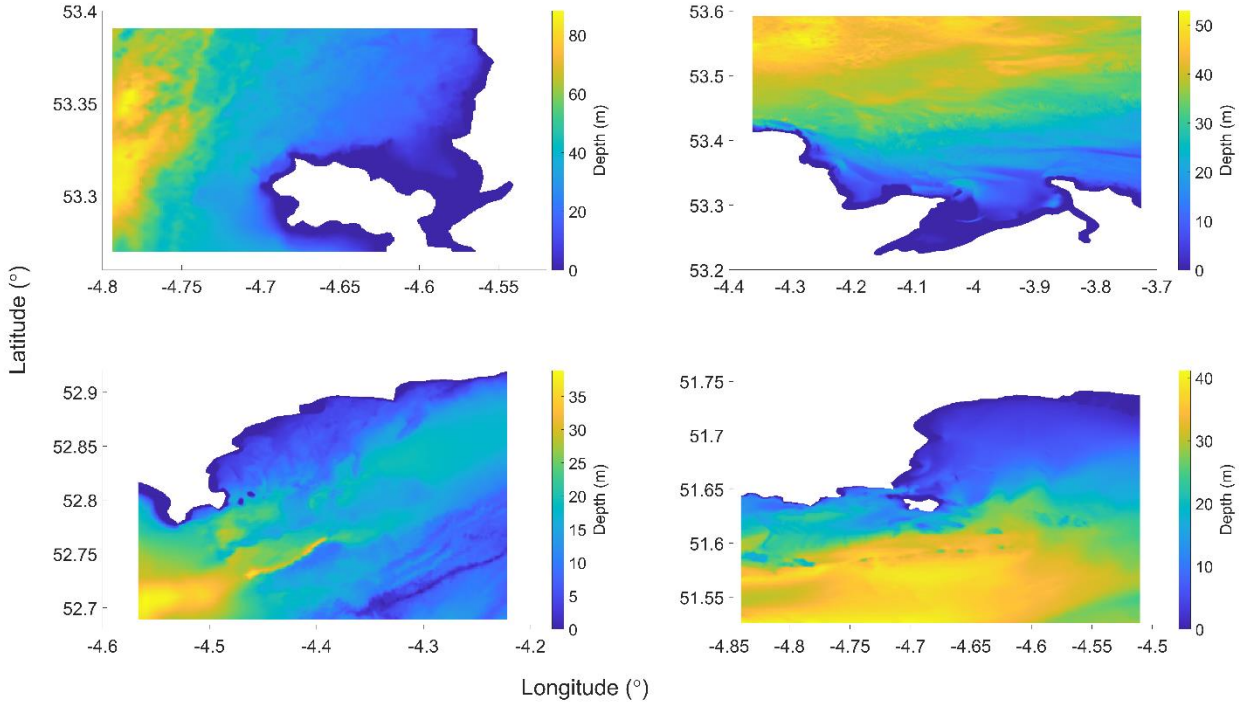

Figure 3. Delft3D near-shore domains. Top left – Anglesey, top right – North Coast, bottom left – Llyn Peninsula, bottom right – South Coast. Bathymetry provided by the EMODnet dataset.

2.5. HSM models using high resolution wave data

Wave data provided from the near-shore wave models including average significant wave height provides finer resolution wave data (~50 m x 50 m) than the open-source data available (~342 m x 195m). As the high-resolution wave data is calculated based on temporal data, four months were chosen to get averaged data. The months used for averages were chosen based upon seasons that are

significant for the life cycle of *Z. marina*. These months were: - January as a potential for highest wave action and wind speeds; March when seed germination and seedling emergence takes place; August for peak biomass and seed production and October for seed establishment and as such the suitable planting season for restoration work (Blok et al., 2018; Orth and Moore, 1986; Sand-Jensen, 1975). High resolution wave height data was converted from gridded data to a raster format in QGIS (Figure 4). Salinity and temperature data were obtained for each of these months as an average of the daily means. Other data was not temporal so remained as a single data layer (bathymetry, PAR, slope and currents). All environmental data layers were clipped to the areas of interest (HSM areas 1-4, Figure 1) along with the presence of *Z. marina* using QGIS. This data was then loaded into R and modelled based on the previously trialled models.

Figure 4. High resolution Delft3D point data (left) converted to raster layer (right) in QGIS v. 3.16, *Z. Marina* presence as orange points.

All the environmental predictor layers were tested for colinearity and colinear variables were removed. This left the model with the following environmental predictor variables:

- Average significant wave height January
- Average temperature March
- Average temperature August
- Average temperature October
- Average salinity August
- PAR at seabed
- Bathymetry
- Energy at seabed due to currents
- Slope of seabed

These layers were stacked and used as predictor variables for the HSM. GAM was found to no longer work, likely caused by the higher number of variables included and decrease in data available, so another algorithm was used – FDA (Flexible Discriminate Analysis) which also uses non-parametric functions. Each approach was run three times each, so the resulting ensemble included 18 prediction results using 6 methods. Variable importance changed significantly from the broadscale models, with wave height becoming much more influential in predicting seagrass presence (see Table 4).

3. Results

3.1. Results of broad-scale habitat suitability models

Of all the predictor variables tested, the coefficient of light attenuation and PAR at the seabed were highly correlated so only PAR was included in the models. Bathymetry was the most influential variable (63 % based on Correlation matrix, 32.3% based on AUC) followed PAR at the seabed (Table 3). Results are shown in figure 5.

Table 3. Results for the UK from the models run in 'sdm' package showing averaged AUC (Area Under Curve), COR (Correlation coefficient), TSS (True Summary Statistic) and the deviance (how much variation remains after the model is fitted) on the left and the summary of the variable importance in % on the right.

Methods	AUC	COR	TSS	Deviance	*Summary of relative variable importance*	*Based on Correlation metric (%)*	*Based on AUC metric (%)*
GLM	0.9	0.76	0.73	0.53	*Currents*	1	0.7
GAM	0.95	0.86	0.86	0.41	*Waves*	0.8	0.9
RF	0.99	0.93	0.93	0.16	*Temperature*	4.7	3.6
BRT	0.95	0.84	0.81	0.49	*Salinity*	4.1	2
MARS	0.94	0.84	0.81	0.38	*Bathymetry*	63	32.3
MaxEnt	0.93	0.79	0.78	1.22	*PAR at seabed*	19.3	12.5

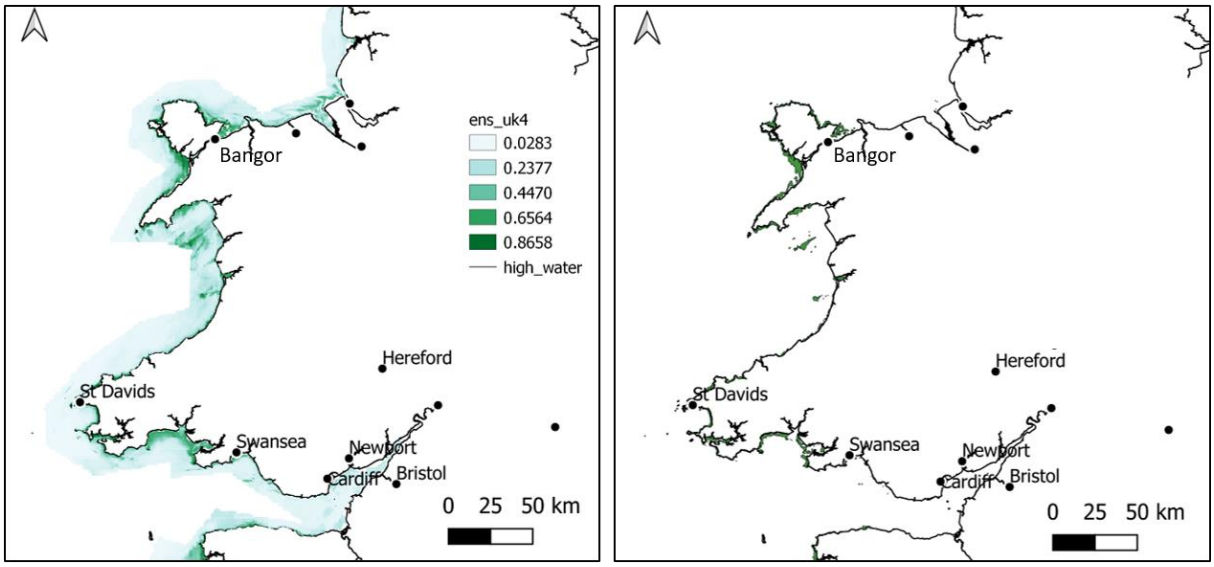

Figure 5. Results from habitat suitability ensemble model results for UK. The legend shows colour scale for the probability of habitat suitability on the left, and map on the right is showing probability of suitability over 0.6 in green.

Results from the UK wide model runs were considerably influenced by bathymetry (Table 3), and outputs from the ensemble model appear to over predict suitable habitat for seagrass around much of the shallow areas around the coast of north Wales especially. The model also predicts areas that are unsuitable for Z. marina growth such as around the tip of the Llyn Peninsula and areas offshore

known to be considerably affected by wave action. This highlighted the need for more accurate predictor variables, particularly wave energy which was found to have little influence on the models as a predictor of presence.

3.2. Results of fine-scale habitat suitability models

The resulting predictions are shown in Figures 6 and 7 for each of the restoration site areas around Wales. Model results are shown in Table 4. All model algorithms resulted in AUC scores of equal to or greater than 0.9 which can be interpreted as good or even excellent predictions based on scales defined in other studies (Araujo et al., 2005). The TSS (True Summary Statistic) is also a useful model validation measure which is independent of prevalence or size of dataset, which is limited for the smaller restoration areas where presence data is low (Allouche et al., 2006). TSS values closer to 1 show higher prediction accuracy, and all models scored above 0.8 (Table 4).

Table 4. Results from the models run in 'sdm' package showing averaged AUC (Area Under Curve), COR (Correlation coefficient), TSS (True Summary Statistic) and the deviance (how much variation remains after the model is fitted) on the left and the summary of the variable importance in % on the right.

Methods	AUC	COR	TSS	Deviance	*Summary of relative variable importance*	*Based on Correlation metric (%)*	*Based on AUC metric (%)*
GLM	0.95	0.67	0.92	0.63	*Ave.significant wave height.jan*	30.1	20.3
BRT	0.96	0.71	0.91	0.59	*Ave.temp.mar.2019*	5.6	2.3
FDA	0.9	0.54	0.82	0.94	*Ave.temp.aug.2019*	17.3	11.3
RF	0.98	0.83	0.95	0.32	*Ave.temp.oct.2019*	8.1	5.1
MARS	0.94	0.76	0.9	4.71	*Ave.sal.aug.2019*	3.8	2
MaxEnt	0.96	0.74	0.91	0.43	*PAR at seabed*	27.2	16.2
					Bathymetry	10.1	7
					Currents	2.2	1.2
					Slope	15.7	14.2

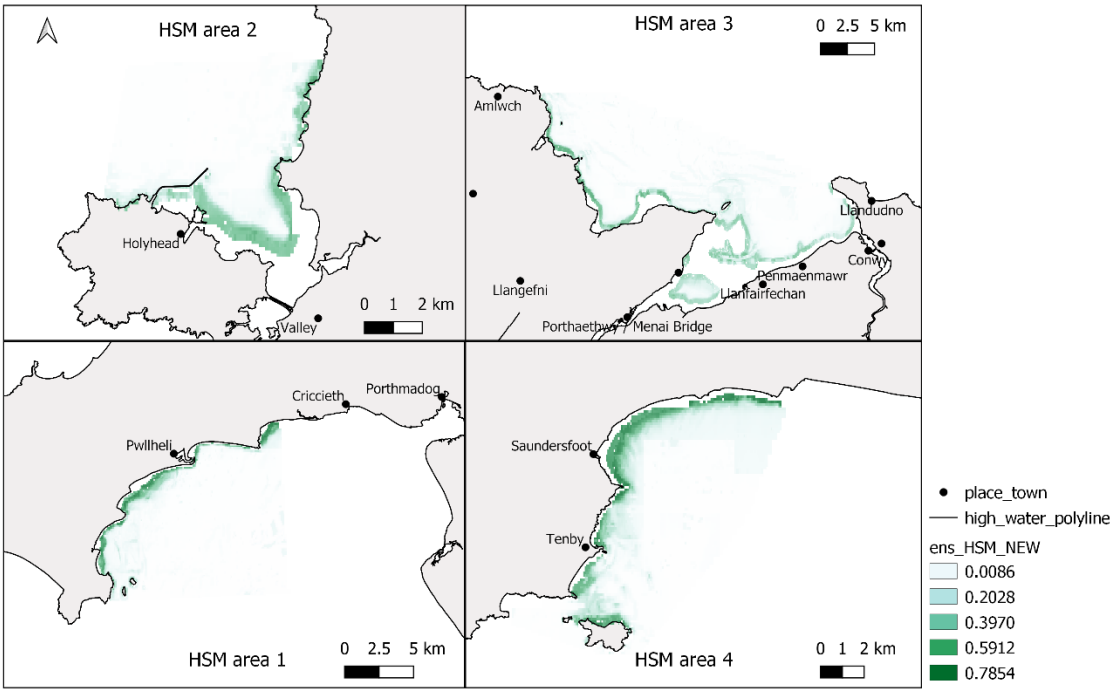

Figure 6. Areas of suitable seagrass habitat in green for the four chosen areas around Wales. Maps show results from ensemble model of 6 methods run 3 times each, using only presence data and all the non-colinear variables.

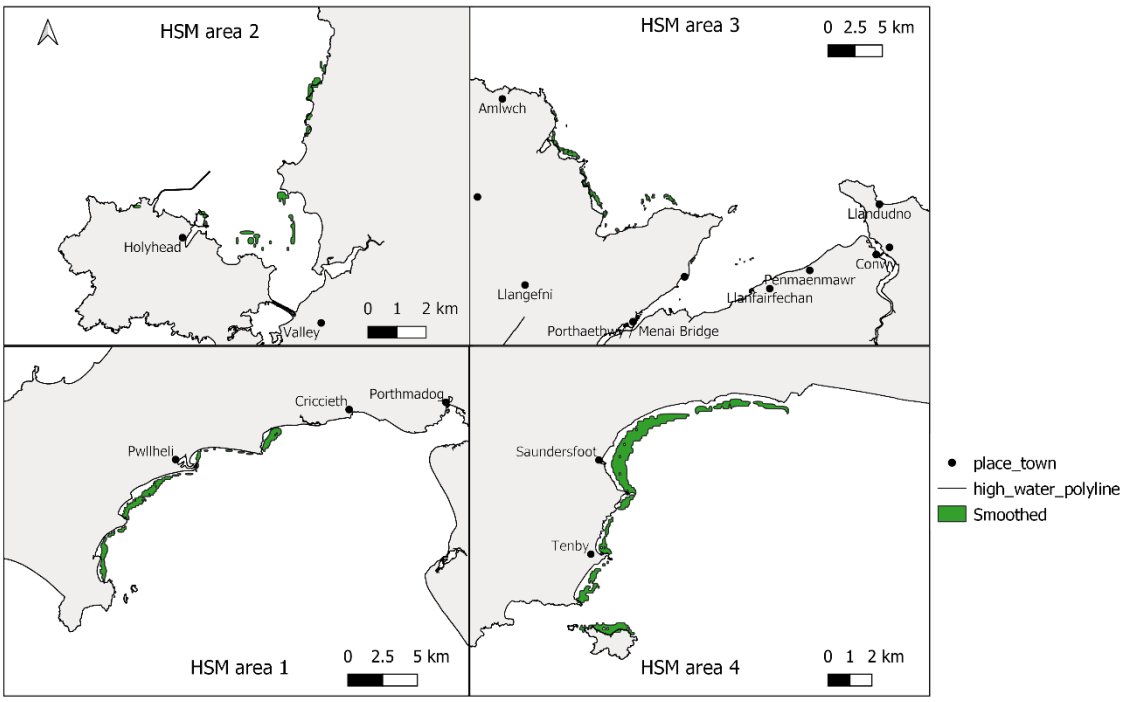

Figure 7. Same results as Fig. 6 with the layer clipped to a threshold of probabilities ≥0.6.

4. Discussion

Habitat Suitability models are a useful tool for identifying areas where seagrass should be able to exist based upon the environmental data that is available. Although no model is perfect, using ensemble modelling enables the inclusion of a range of model algorithms and a combination of approaches. This removes the need for selecting a single 'best' model which limits model selection bias and reduces the need for removing predictor variables which is often the case for step-wise model selection. This method is useful for predictions in new areas where conditions may be different or where data is lacking (Guisan et al., 2017). Therefore, this model method should be more transferrable to other sites around the UK, where a source of fine-scale wave height data such as Delft3D modelled data used in this study.

The use of fine resolution Delft3D wave data has enabled habitat suitability models to be created to a fine resolution within the four small areas identified for potential seagrass restoration around Wales. These results appear to be more appropriate for seagrass growth than the broadscale outputs for the UK and north Wales. The initial trials including lower resolution data resulted in overprediction of suitability in areas that are not appropriate for seagrass growth, such as very large swathes of exposed coastlines (around Anglesey Island) and deep (>15 m) regions such as offshore from the mid Wales coast.

The results of the ensemble model using Delft3D wave data appear to show realistic predictions for seagrass growth, with average wave height for January obtained from the wave model found to be the most important variable for predicting seagrass suitability. January was predicted to have the highest wave heights and so will determine where exposure to wave energy is highest. The PAR at the seabed was the second most important variable for predicting seagrass presence in the four areas around Wales. The open-source wave energy data used in the same ensemble model was found to be less important as a predictor variable, with bathymetry being the most important factor by far.

The interpretation of HSM outputs can be tricky and relatively subjective based upon prior knowledge of sites and the species in question. The model outputs (AUC, COR and TSS) allow model comparisons and will indicate those which best fit, however when the outputs are very similar, visual interpretation is needed to assess if the results are realistic. The results from the HSM, shown in figures 6 and 7 predict suitability in areas that appear to be feasible for restoration. The ensemble model shows higher levels of probability in site areas 1 and 4 (Llŷn Peninsula and southwest Wales coast) than areas 2 and 3 (east and west of Anglesey). This could be due to areas 2 and 3 having higher areas of intertidal shoreline, where seagrass has been found. Due to large resource requirements, the wave model within Delft3D was not coupled to the tidal processes, and as such wave modelling was carried out for the mean water level. The wave modelling therefore does not provide data in the upper intertidal, limiting prediction of wave heights in shallow extents, especially where land is also obstructing potential wave energy, and therefore this restricts the model areas to subtidal zones. The wave height data was extrapolated to get some extra overlap in the shallows, but only by 3 pixels to maintain data integrity. Ground truthing the highlighted areas will help to improve site selection for restoration. If restoration is deemed suitable for the intertidal areas in these sites, then the area could be extended somewhat shorewards. Outputs from the HSM will aid in focussing field survey efforts within these sites for choosing restoration sites. Other factors that would be useful for site choice which could also be fed into the HSM would be a quantitative survey of substrate looking at particle size. Only categorical values could be obtained for this study which could not be integrated into the HSM. However, all of the areas that have been predicted to be suitable for seagrass growth are in areas of soft sediment overall which is often where wave exposure is low and due to existing seagrass presence data available for the model. Ground-truthing before restoration takes place will allow any unsuitable

substrates to be avoided. Alongside, other environmental variables would need to be measured at these locations importantly factors affecting water quality and seagrass growth such as nutrients, turbidity, dissolved oxygen, and temperature. Other factors may also help to refine areas for restoration including overlap with other stakeholders including harbour areas.

5. Conclusions and recommendations

The use of high-resolution wave data has enabled habitat suitability modelling to be refined for highlighting areas suitable for seagrass growth at a fine-scale. The model outputs from the ensemble models suggest that there are large potentially suitable areas within the spatial extents around Wales chosen for possible restoration. All HSM areas have areas over 0.6 probability of being suitable for *Z. marina* to grow, especially off the south Llyn peninsula and south Pembrokeshire.

It is recommended that these refined areas are surveyed to see if any seagrass is already present and if other conditions appear conducive to seagrass growth. There are some factors not explored within the HSM models, such as substrate and nutrients. This was because the data was either unavailable or not available in a format that needed for the modelling methods used. Nutrient loading is one of the biggest issues for existing seagrass meadows around Wales and the UK (Bertelli et al., 2021). Unless these issues are addressed, restoration will have a lower chance of being successful.

5. References

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Araujo, M., Pearson, R., Thuiller, W., Erhard, M., 2005. Validation of species-climate impact models under climate change. Glob. Chang. Biol. 11, 1504–1513. https://doi.org/10.1111/j.1365-2486.2005.001000.x

Araújo, M.B., Anderson, R.P., Barbosa, A.M., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R.B., Zimmermann, N.E., Rahbek, C., 2019. Standards for distribution models in biodiversity assessments. Sci. Adv. 5, 1–12. https://doi.org/10.1126/sciadv.aat4858

Beca-Carretero, P., Stanschewski, C.S., Julia-Miralles, M., Sanchez-Gallego, A., Stengel, D.B., 2019. Temporal and depth-associated changes in the structure, morphometry and production of near-pristine Zostera marina meadows in western Ireland. Aquat. Bot. 155, 5–17. https://doi.org/10.1016/j.aquabot.2019.02.003

Bertelli, C.M., Bull, J.C., Cullen-Unsworth, L.C., Unsworth, R.K.F., 2021. Unravelling the Spatial and Temporal Plasticity of Eelgrass Meadows. Front. Plant Sci. 12, 1–12. https://doi.org/10.3389/fpls.2021.664523

Blok, S.E., Olesen, B., Krause-Jensen, D., 2018. Life history events of eelgrass Zostera marina L. populations across gradients of latitude and temperature. Mar. Ecol. Prog. Ser. 590, 79–93. https://doi.org/10.3354/meps12479

Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions 1. Model description and validation. J. Geophys. Res. Ocean. 104, 7649–7666. https://doi.org/10.1029/98JC02622

D'Avack, E.A.S., Tyler-Walters, H., Wilding, C., Garrard, S.M., 2019. Zostera (Zostera) marina beds on lower shore or infralittoral clean or muddy sand In Tyler-Walters H. and Hiscock K. (eds) Marine Life Information Network: Biology and Sensitivity Key Information Reviews [WWW Document]. URL https://www.marlin.ac.uk/habitats/detail/257

Davison, D.M., Hughes, D.J., 1998. Vol.I. Zostera biotopes 1. An overview of dynamic and sensitivity characteristics for conservation management of marine SACs.

EUSeaMap, 2012. EUSeaMap Technical Appendix No. 1 Light data and thresholds.

Fourqurean, J.W., Duarte, C.M., Kennedy, H., Marba, N., Holmer, M., Mateo, M.A., Apostolaki, E.T., Kendrick, G.A., Krause-Jensen, D., McGlathery, K.J., Serrano, O., 2012. Seagrass ecosystems as a globally significant carbon stock. Nat. Geosci. 5, 505–509. https://doi.org/http://www.nature.com/ngeo/journal/v5/n7/abs/ngeo1477.html#supplementary-information

Guisan, A., Thuiller, W., Zimmermann, N.E., 2017. Habitat Suitability and Distribution Models With Applications in R. Cambridge University Press. https://doi.org/10.1017/9781139028271

Hersbach, H., Peubey, C., Simmons, A., Poli, P., 2013. 16 ERA-20CM: a twentieth century atmospheric model ensemble. Report 46.

Jackson, E.L., Griffiths, C.A., Durkin, O., 2013. A guide to assessing and managing anthropogenic impact on marine angiosperm habitat - Part 1: Literature review. Natural England Commissioned Reports, Number 111.

Krause-Jensen, D., Pedersen, M.F., Jensen, C., 2003. Regulation of eelgrass (Zostera marina) cover along depth gradients in Danish coastal waters. Estuaries 26, 866–877. https://doi.org/10.1007/BF02803345

Lesser, G.R., Roelvink, J.A., van Kester, J.A.T.M., Stelling, G.S., 2004. Development and validation of a three-dimensional morphological model. Coast. Eng. 51, 883–915. https://doi.org/10.1016/j.coastaleng.2004.07.014

Marion, S.R., Orth, R.J., Fonseca, M., Malhotra, A., 2020. Seed Burial Alleviates Wave Energy Constraints on Zostera marina (Eelgrass) Seedling Establishment at Restoration-Relevant Scales. Estuaries and Coasts.

Naimi, B., Araújo, M.B., 2016. Sdm: A reproducible and extensible R platform for species distribution modelling. Ecography (Cop.). 39, 368–375. https://doi.org/10.1111/ecog.01881

Nielsen, S., Sand-Jensen, K., Borum, J., Geertz-Hansen, O., 2002. Depth colonization of eelgrass (Zostera marina) and macroalgae as determined by water transparency in Danish coastal waters. Estuaries and Coasts 25, 1025–1032. https://doi.org/10.1007/bf02691349

Orth, R.J., Moore, K.A., 1986. Seasonal and year-to-year variations in the growth of Zostera marina L. (eelgrass) in the lower Chesapeake Bay. Aquat. Bot. 24, 335–341. https://doi.org/10.1017/CBO9781107415324.004

Potouroglou, M., Bull, J.C., Krauss, K.W., Kennedy, H.A., Fusi, M., Daffonchio, D., Mangora, M.M., Githaiga, M.N., Diele, K., Huxham, M., 2017. Measuring the role of seagrasses in regulating sediment surface elevation. Sci. Rep. 1–11. https://doi.org/10.1038/s41598-017-12354-y

Sand-Jensen, K., 1975. Biomass, net production and growth dynamics in an eelgrass (Zostera marina L.) population in Vellerup Vig, Denmark. Ophelia 14, 185–201. https://doi.org/10.1080/00785236.1975.10422501

Shafer, D., Bergstrom, P., 2008. An Introduction to a Special Issue on Large-Scale Submerged Aquatic Vegetation Restoration Research in the Chesapeake Bay: 2003-2008. Restor. Ecol. 18, 481–489. https://doi.org/10.1111/j.1526-100X.2010.00689.x

Unsworth, R.K.F.R.K.F., Bertelli, C.M.C.M., Cullen-Unsworth, L.C.L.C., Esteban, N., Jones, B.L.B.L., Lilley, R., Lowe, C., Nuuttila, H.K.H.K., Rees, S.C.S.C., 2019. Sowing the seeds of seagrass recovery using hessian bags. Front. Ecol. Evol. 7, 1–7. https://doi.org/10.3389/fevo.2019.00311

Valle, M., van Katwijk, M.M., de Jong, D.J., Bouma, T.J., Schipper, A.M., Chust, G., Benito, B.M., Garmendia, J.M., Borja, Á., 2013. Comparing the performance of species distribution models of Zostera marina: Implications for conservation. J. Sea Res. 83, 56–64. https://doi.org/10.1016/j.seares.2013.03.002

Yesson, C., Bush, L.E., Davies, A.J., Maggs, C.A., Brodie, J., 2015. The distribution and environmental requirements of large brown seaweeds in the British Isles. J. Mar. Biol. Assoc. United Kingdom 95, 669–680. https://doi.org/10.1017/S0025315414001453