
Jarvis, Susan G.; Rowe, Edwin C.; Henrys, Peter A.; Smart, Simon M.; Jones, Laurence; Garbutt, Angus. 2016. **Empirical realised niche models for British coastal plant species.** *Journal of Coastal Conservation*, 20 (2). 107-116. 10.1007/s11852-016-0422-3

Empirical realised niche models for British coastal plant species

Submitted to Journal of Coastal Conservation

Susan G. Jarvis[1]*, Edwin C. Rowe[2], Peter A. Henrys[1], Simon M. Smart[1], Laurence Jones[2], Angus Garbutt[2]

[1] Centre for Ecology and Hydrology Lancaster, Lancaster Environment Centre LA1 4AP, UK.

[2] Centre for Ecology and Hydrology Bangor, ECW, Deiniol Road, Bangor LL57 2UP, UK.

*Corresponding author; Email: susjar@ceh.ac.uk; Telephone: +44 (0)1524 595941

Abstract

Coastal environments host plant taxa adapted to a wide range of salinity conditions. Salinity, along with other abiotic variables, constrains the distribution of coastal plants in predictable ways, with relatively few taxa adapted to the most saline conditions. However, few attempts have been made to quantify these relationships to create niche models for coastal plants. Quantification of the effects of salinity, and other abiotic variables, on coastal plants is essential to predict the responses of coastal ecosystems to external drivers such as sea level rise. We constructed niche models for 132 coastal plant taxa in Great Britain based on eight abiotic variables. Paired measurements of vegetation composition and abiotic variables are rare in coastal habitats so four of the variables were defined using community mean values for Ellenberg indicators, i.e. scores assigned according to the typical alkalinity, fertility, moisture availability and salinity of sites where a species occurs. The remaining variables were the canopy height, annual precipitation, and maximum and minimum temperatures. Salinity and moisture indicator scores were significant terms in over 80% of models, suggesting the distributions of most coastal species are at least partly determined by these variables. When the models were used to predict species occurrence against an independent dataset 64% of models gave moderate to good predictions of species occurrence. This indicates that most models had successfully captured the key determinants of the niche. The models could potentially be applied to predict changes to habitats and species-dependent ecosystem services in response to rising sea levels.

Keywords

Climate; Ellenberg; Generalised linear model; Saltmarsh; Sand dune; Vegetation;

Introduction

A quantitative understanding of the habitat requirements of plant species is necessary if effects of environmental change on species occurrence and composition are to be predicted. Coastal habitats have high nature conservation value supporting many rare species, and coastal plant species provide important functions such as stabilising substrates, providing suitable habitat structure for bird feeding or nesting and sequestering carbon (Ranwell, 1972; Jones *et al.*, 2011a; Malpas *et al.*, 2013; Beaumont *et al.*, 2014). The development of realised niche models (Latour & Reiling, 1993; Smart *et al.*, 2010b) has provided capacity to predict change in species occurrence and species composition resulting from environmental change (de Vries *et al.*, 2010), but few such models exist for coastal plant species and none have been constructed for British taxa (Batriu *et al.*, 2011; Zhu *et al.*, 2013; Mendoza-González *et al.*, 2013). In the current study we describe the development of niche models that take into account the salinity requirements of species, among other influences, and that could be used to explore the effects of sea-level rise and other environmental changes on coastal habitats.

The distribution of coastal species is thought to be strongly constrained by the degree of exposure to seawater or salt spray and consequent substrate salinity (Ranwell, 1972), and therefore salinity would be expected to be important in determining the niches of coastal plants. Salt marsh vegetation is exposed to saline conditions through inundation while plants further inland are affected by salt input from coastal spray, which may penetrate up to a kilometre inland (Lowe *et al.*, 1996). Saline intrusion into groundwater and occasional inundation of low-lying land by sea water during storm surges can also increase soil salinity in coastal habitats. Variation in exposure to salinity between the intertidal and upper zones of the coastal ecosystem leads to clear shifts in the composition of vegetation (Emery *et al.*, 2001). However, the distribution of coastal plants may also be influenced by other variables such as the substrate water content, nutrient status and climate. Quantifying the relative impacts of drivers of coastal plant distributions, and any interactions between drivers, is important to predict the response of coastal plants to environmental change.

Quantifying the role of variables such as salinity in defining plant niches is best achieved by relating species occurrence to physico-chemical measurements of the environment. However, measurement of these variables is expensive and time-consuming, and in most floristic datasets the amount of environmental information associated with each stand or relevé is very limited. In particular, measures of salinity such as electrical conductivity are rarely taken in conjunction with floristic data. However, information is available directly from the plant community assemblage in the form of indicator scores used to describe the environmental requirements of each plant species (Ellenberg, 1991). Ellenberg scores were developed to assess variation in species optima along a number of ecological axes including pH, fertility and salinity (Ellenberg, 1991). For each axis species are given a score which reflects their preference, for example salinity is measured on a scale from 0 to 9 with species intolerant of salinity assigned a score of 0 and those occurring in highly saline conditions, e.g. *Salicornia europea* agg., assigned 9. The scores originally defined by Ellenberg were refined for the British flora by re-prediction using a two-way averaging method (Hill *et al.*, 2000, 2004). Although a single score on each environmental axis does not describe the niche breadth, the type of relationship with the variable (e.g. monotonal or unimodal) or how suitability is affected by interactions with other environmental factors, the mean indicator score for all species at a site has been demonstrated to give a robust indication of site conditions (Diekmann, 2003). Modelling species occurrence as a function of community mean scores for multiple Ellenberg indices can provide a useful description of the niche space occupied by a species (Smart *et al.*, 2010b).

The community mean Ellenberg score has also been shown to be strongly related to physico-chemical measurements by simple equations (Smart *et al.*, 2010b; Rowe *et al.*, 2011) and there may be several advantages to using mean indicator scores beyond the scarcity of information associated with floristic datasets. In some cases mean indicator scores may better represent biophysical properties of a site than physico-chemical measurements which are typically sampled at a single or few locations and may not capture heterogeneity within a site (Wagner *et al.*, 2007). Direct measurements are also subject to measurement error and variation in measurement techniques which are circumvented by the use of indicator scores. This is particularly true of highly dynamic properties such as soluble nutrient element concentrations, and properties for which a variety of measurement methods exists such as plant-available nitrogen. Finally, mean indicator scores do not suffer from the difficulty of ensuring that environmental measurements are from the same location as the floristic records, which is important when considering small-scale variation.

Plant niche models based on Ellenberg indicator values and climatic axes have already been developed for 1,130 British plant species and have been shown to produce accurate predictions of species occurrence (Smart *et al.*, 2010b). In these models Ellenberg values corresponding to pH, fertility and moisture were used as predictor variables in generalised linear models along with canopy height and climatic variables. However, these authors excluded coastal species as they did not consider salinity which was expected to be a key driver of coastal plant niches. The current study extends this modelling approach to 132 coastal plant species by including mean Ellenberg salinity score as an additional explanatory variable to examine whether the occurrence of coastal plant species can be predicted on the basis of environmental gradients.

Methods

Data sources

Data used to construct the models were collated from several independent vegetation surveys covering the entire range of British habitats. The National Vegetation Classification, Countryside Survey, Broadleaved Woodland Survey and Key Habitat surveys are described in detail in (Smart *et al.*, 2010b). In addition, 138 quadrats from the Countryside Survey were added which were excluded from the previous modelling work due to their coastal location. Further coastal training quadrats were provided by a range of smaller surveys including the Threatened Plants Project, and ten national and local surveys of sand dunes and dune slacks (Jones *et al.*, 2004, 2011b; Plassmann *et al.*, 2009; Ford *et al.*, 2012; Curreli *et al.*, 2013; Rhymes *et al.*, 2014) to bring the total of quadrats used to train the model to 33,865. The vegetation composition of each quadrat was used to calculate the Ellenberg indices used as model variables. The Ellenberg indicators used were Ellenberg R (related to pH), Ellenberg N (related to fertility), Ellenberg F (related to moisture) and Ellenberg S (related to salinity). Scores recalculated for the British flora were used in place of the original values (Hill *et al.*, 2000, 2004). For each species, the mean community Ellenberg scores were calculated based on the associated species only to avoid circularity, i.e. removing the species in question from the calculation. In addition, an indicator of relative light availability in each plot was calculated by taking the mean height class (Grime & Hodgson, 1988), weighted by percentage cover of each species. Ellenberg indicators were not weighted by cover as unweighted values have been shown to be better correlated with environmental measures (Carpenter & Goodenough, 2014). Data on mean annual rainfall, average January temperature minimum and average July temperature maximum at 5 km resolution were obtained from the Met Office long-term average data (available at http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/). The vegetation surveys comprising the training dataset were conducted over several decades and therefore the period chosen to reflect average climate was variable between datasets. The bulk of the data was collected prior to 1990 so the 1961-1990 period was chosen, but for some of the more recent surveys the 1981-2010 period was used to cover the time of the survey. The eight explanatory variables included in the models are consistent with those chosen to model the niches of 1,130 species of higher and lower plants in Britain (Smart *et al.*, 2010b), with the addition of Ellenberg S to facilitate modelling of coastal taxa.

Model fitting and visualisation

Models were fitted for all plant species present in the training data which had an Ellenberg S score of one or more, taken to indicate the ability to occur in coastal environments. Models were fitted for 132 species in total (see Online Resource 2 for species list). Binary logistic generalised linear models (GLMs) were used to model occurrence of each species in relation to the eight explanatory variables (three climate variables, four Ellenberg indicators and cover-weighted canopy height) described above. Each explanatory variable was included as both linear and quadratic terms, and all two-way interactions of linear terms were also included, giving 44 potential model coefficients. Although simple, GLMs have been shown to outperform alternative niche modelling techniques such as classification trees and are easier to fit and interpret for large numbers of species than more complex models (Elith *et al.*, 2006; Meynard & Quinn, 2007; Smart *et al.*, 2010b). Use of GLMs as opposed to additive models also reduces the likelihood of overfitting models to the training data (Randin *et al.*, 2006; Smart *et al.*, 2010a). Full models containing all terms were fitted and two steps of model selection performed. Initially only significant terms with a critical threshold of $P < 0.05$ were selected from the full model, then stepwise backward selection based on the Akaike Information Criterion (AIC) was run on the shrunken model. The final model contained all terms that passed through both selection steps. As an additional test of the importance of the salinity term all GLM models were fitted without Ellenberg S, repeating the model fitting in Smart *et al.* (2010b) except for the addition of extra training data.

To provide a graphical representation of the modelled niche for a subset of four representative species, each model was used to predict the probability of species occurrence in relation to individual model terms, holding other parameters constant at median values. The probabilities of occurrence returned by model predictions are a function of both the explanatory variables and the prevalence of the species in the training data. Rare species will always have a low probability of occurrence, even where abiotic conditions are optimal. To facilitate the visualisation of differences in modelled niche space between species a scaling function was applied to the probability of occurrence (Real *et al.*, 2006). The scaling function adjusts the probability of occurrence using the prevalence of the species in the training data to produce an index of habitat suitability (HS) ranging from 0 to 1 as follows:

$$\mathrm{HS} = \frac{\frac{P}{1-P}}{\frac{n_1}{n_0} + \frac{P}{1-P}}$$

where P is the probability of occurrence from the GLM and n_1 and n_0 are the numbers of presences and absences in the dataset (Real *et al.*, 2006). The graphical representation of each niche was compared to species descriptions in Preston et al. (2002) and Stace (2010) as a qualitative evaluation of model performance.

Model evaluation

To quantitatively assess whether the models could accurately predict species occurrence they were tested against an independent set of 5,308 coastal vegetation plots collected by the Environment Agency between 2007 and 2012 as part of the Water Framework Directive marine angiosperm monitoring tool (Environment Agency 2013; available at http://www.geostore.com/environment-agency/WebStore?xml=environment-agency/xml/ogcDataDownload.xml). Community mean Ellenberg scores were calculated for each species in each plot, and climate data were collated from the 1981-2010 period which covers the time period of the surveys. The models were used to predict the presence or absence of taxa in the test dataset and the performance of the models was assessed by plotting receiver-operator characteristic plots (ROC plots) and assessing the area under the curve statistic (AUC) using the ROC-R package (Sing *et al.*, 2005). The ROC plot is constructed by plotting the number of true positives (where a species is predicted to be present and observed to be present) against the number of false positives (where a species is predicted to be present but is actually absent) for a range of probability thresholds. For a good model the number of true positives will increase faster than the number of false positives when the threshold used to convert the probability of occurrence to presence decreases and therefore the AUC will be large. For models with no ability to predict presences the true positive rate will increase at the same rate as the false positive rate (AUC will be 0.5).

Interpretation of AUC statistics without the ROC plots has been criticised (Jiménez-Valverde, 2012) and therefore all ROC plots were assessed visually (Online Resource 1). In addition, all test plots are from coastal areas to avoid artificially increasing AUC statistics by including areas unlikely to contain the modelled species (Lobo *et al.*, 2008). Only models for species occurring in both training and test datasets, a total of 45 taxa, could be tested. All model fitting and analysis was conducted in R v. 3.0.2 (R Foundation for Statistical Computing, 2013).

Results

In total, 132 coastal plant species were modelled using the training dataset. Of the fitted models for these species, 85% contained either a linear or quadratic Ellenberg S term, indicating that salinity was important in determining niche space for these species (Table 1, Online Resource 2). Ellenberg F terms, relating to moisture, were also included in over 80% of the models, whereas climate variables were only included in approximately half of all models. On average models contained 18 out of 44 possible terms, with a mean of eight two-way interactions included in each model.

Modelled responses to Ellenberg S for four representative taxa were investigated in detail by predicting habitat suitability for the full range of Ellenberg S values (0 to 9), holding all other indicator and climatic variables constant at the median values for that species (Figure 1). Relationships were shown to be variable between species in terms of the Ellenberg S optimum and the modelled niche breadth of the salinity axis. Large variation in modelled niche space in relation to Ellenberg S was observed between species with the same Ellenberg S score. This indicates that the range of suitable salinity conditions is variable between species. To visualise potential interactions between explanatory variables, three dimensional plots were produced by varying the two most common model terms (Ellenberg S and Ellenberg F; Figure 2). Three dimensional plots demonstrated that both salinity and moisture were important characteristics of niche space for the plant species shown, and that the modelled niches corresponded well to habitat descriptions in standard floras such as Stace (2010) and Preston et al. (2002).

Model evaluation was performed for the 45 niche models where the species was present in both training and test datasets (Figure 3, Table 2). The majority of models (64%) had AUC values above the threshold of 0.7 suggested to indicate useful performance (Swets, 1988; Manel *et al.*, 2001). Comparison with AUC values from GLM model fits without salinity terms (Table 2) showed that addition of Ellenberg S as an explanatory variable increased AUC by an average of 0.05. Removing the salinity term for *Trifolium fragiferum* decreased model predictive ability from good (0.80) to having no predictive power (0.39). Most models showed moderate declines in performance after removing the salinity term although models for three species (*Atriplex littoralis*, *Crithmum maritimum* and *Parapholis strigosa*) had substantially higher (over 20%) AUC values without the salinity term.

Discussion

The distributions of the coastal plant species examined here were shown to be primarily driven by variation in salinity and moisture, although fertility, pH, canopy height and climate were also found to be important for many species. Ellenberg indicator scores for salinity and moisture (Ellenberg S and F) were the most frequent explanatory variables, both occurring in over 80% of models. This result supports the finding that exposure to salinity and moisture availability are key controls on coastal plant niches (Batriu *et al.*, 2011; Yuan *et al.*, 2012). Models fitted with only the seven variables defined in Smart et al. (2010) and excluding Ellenberg S showed, on average, lower predictive power against an independent dataset than models fitted with the salinity term. Whilst the finding that salinity is an important component of coastal plant niches is not surprising, given that many coastal plants have specific adaptations to high salt concentrations, salinity has rarely been included in previous niche models of coastal plants due to the lack of measurements associated with vegetation surveys. Here, responses to salinity have been quantified for both taxa characteristic of intertidal saltmarsh e.g. *Aster tripolium* and for taxa that occur further inland and are influenced by factors such as salt spray e.g. *Trifolium fragiferum*.

The general paucity of available environmental data to inform species niche modelling efforts makes the use of proxies such as Ellenberg indicators as explanatory variables an attractive alternative. For environments such as coastal habitats, where there is little environmental data collection to complement floristic surveys, Ellenberg indicators allow the inclusion of relevant explanatory variables to niche modelling efforts. In the models presented here, the use of mean community Ellenberg scores allowed niche models to be constructed for a large floristic dataset where direct environmental measurements were not available. Comparison of modelled niche space to existing floristic descriptions for four representative taxa indicated a generally good correspondence. *Schoenoplectus tabernaemontani* is described as growing in “brackish water” (Preston *et al.*, 2002) and the modelled niche represented in Figure 2 showed the highest habitat suitability in the wettest conditions. The model for *Carex arenaria*, a “dominant of fixed dunes, dune-slacks, sandy flats and tracksides” (Preston *et al.*, 2002), showed high habitat suitability over a wide range of moisture conditions but low suitability in highly saline conditions, suggesting salinity may limit its distribution in coastal areas. The model for *Euphorbia paralias* showed less salt-tolerance than was implied by the relatively large Ellenberg Salinity score of 3, but species given a score of 3 clearly vary in their salt-tolerance, and the model shows that this score may be an overestimate for this species.

The use of Ellenberg indicators to model species occurrence has been criticised on the basis of the circularity inherent in predicting the occurrence of species based on values derived from the community of species present. In particular, relating Ellenberg scores to other variables derived from the species composition such as richness has been shown to exaggerate the importance of indicator scores as predictors compared to abiotic measurements (Zelený & Schaffers, 2012). However, by modelling only single species and removing the species in question from the dataset before calculation of the mean community Ellenberg score (Smart *et al.*, 2010b), the main potential cause of circularity has been avoided in this study. Some degree of circularity may remain given that Ellenberg values for plant taxa were initially defined partly based on co-occurring species, but the use of algorithmically refined scores (Hill *et al.*, 2000) reduces this circularity. Although accurate measurements of abiotic conditions at the site of floristic data collection will always be preferred in niche modelling exercises, the use of Ellenberg indicators enables useful models to be constructed where such data are not available. To enable models to be built on measured abiotic variables in place of Ellenberg indicators there is an urgent need for more collection of environmental data co-located with vegetation relevés. Measurements indicating salinity (e.g. electrical conductivity in a 10 g soil / 25 ml water slurry; or sodium concentration) would be particularly useful for establishing niches of coastal plants.

Not all models performed well against test data, suggesting the main drivers for these taxa were not sufficiently represented in the models. For example, our models did not take into account substrate mobility which is important in determining niche space for some dune species (Maun & Perumal, 1999). However, dune mobility and salt exposure in sea spray are to a large extent co-correlated and dune plants living closest to the shoreline in the most mobile conditions have higher Ellenberg Salinity scores. For example the strandline and mobile dune species *Cakile maritima* and *Ammophila arenaria* both have scores of 3, compared with *Festuca rubra* with a score of 2, which although salt tolerant is typically found behind the leading dune. Salinity scores are therefore likely to capture much of this influence, but we recognise that they will not perform so well for species occupying blow-out conditions further inland. Further work focusing on dune species could test additional variables. Limiting the set of input variables allows all species to be modelled with the same predictors but increases the likelihood that the niches of some species are not captured by the input variables.

Methodological issues may also have affected performance for some models. For example, models for the genus *Atriplex* were particularly poor and might reflect inconsistency in identification of species between surveys. In addition, despite the large number of quadrats in the training dataset a relatively small proportion occurred in coastal habitats. The test dataset was restricted to coastal habitats and so some species had higher numbers of occurrences in the test dataset than the training dataset, a potential cause of poor performance if the entire species range was not captured in the training dataset. Therefore, despite the large dataset used for model building, the sampling effort in coastal regions was not high enough to create good models for all taxa. Increased survey effort in coastal regions will be required to address this, preferably including co-located environmental measurements.

Nonetheless, model evaluation showed that 64% of models tested were classed as good to excellent when compared against an independent test dataset (Swets, 1988; Manel *et al.*, 2001). Species occurrence can therefore be satisfactorily predicted for a majority of species with only eight explanatory variables. This indicates that the variables included capture the main drivers of coastal plant distributions. Although salinity and moisture were the dominant variables, multiple drivers were implicated in determining the distributions of the majority of species and most models contained multiple interaction terms, suggesting that species distributions are determined by a complex function of abiotic and biotic conditions. The results indicate that it will be necessary to consider the impacts of multiple drivers when considering how coastal plants may be affected by environmental change.

Coastal vegetation is threatened by rising sea levels which are likely to lead to increased submersion of lower saltmarshes (Boorman, 1992; Donnelly & Bertness, 2001). Vegetation further inland may also be affected by rising groundwater levels (Curreli *et al.*, 2013), and the potential habitat extent available will be reduced by coastal squeeze (Jones *et al.*, 2011a). The models developed here allow habitat suitability in relation to moisture and salinity to be described for a large proportion of British coastal plant species, and have potential applications in predicting likely habitat composition under future inundation regimes. Sufficient co-located floristic and environmental data are not currently available for deriving niche models directly with respect to abiotic conditions, but fewer data are necessary to derive relationships between mean indicator scores and the abiotic environment (e.g. Smart *et al.*, 2010b; Rowe *et al.*, 2011). If relationships can be established between mean Salinity score and aspects of the abiotic environment such as soil electrical conductivity or inundation depth and frequency, it will be possible to simulate the effects of sea-level rise on individual coastal species and species-assemblages. More immediately, current mean Ellenberg scores provide an indication of suitability for colonisation by other taxa, and therefore the likely success of vegetation restoration in managed re-alignment projects.

Conclusions

The models presented here demonstrate that the distributions of British coastal plants are driven by multiple interacting drivers, with salinity and moisture being the most important variables. By using relevant Ellenberg indicators as explanatory variables it was possible to describe responses to variables for which direct measurements are rarely associated with floristic datasets. There is now the possibility to use these models to predict impacts of environmental change on British coastal plant species.

Acknowledgements

We are grateful to Niall Phelan from the Environment Agency for providing the extensive saltmarsh test dataset and to Annette Burden for help collecting saltmarsh datasets in the field. This study was funded by Natural Environmental Research Council Centre for Ecology & Hydrology through the Ecological Processes and Resilience Science programme; Project Code NEC04654.

References

Batriu E, Pino J, Rovira P, Ninot JM. **2011**. Environmental control of plant species abundance in a microtidal Mediterranean saltmarsh. *Applied Vegetation Science* **14**: 358–366.

Beaumont NJ, Jones L, Garbutt A, Hansom JD, Toberman M. **2014**. The value of carbon sequestration and storage in coastal habitats. *Estuarine, Coastal and Shelf Science* **137**: 32–40.

Boorman LA. **1992**. The environmental consequences of climatic change on British salt marsh vegetation. *Wetlands Ecology and Management* **2**: 11–21.

Carpenter W, Goodenough AE. **2014**. How robust are community-based plant bioindicators? Empirical testing of the relationship between Ellenberg values and direct environmental measures in woodland communities. *Community Ecology* **15**: 1–11.

Curreli A, Wallace H, Freeman C, Hollingham M, Stratford C, Johnson H, Jones L. **2013**. Eco-hydrological requirements of dune slack vegetation and the implications of climate change. *Science of The Total Environment* **443**: 910–919.

Diekmann M. **2003**. Species indicator values as an important tool in applied plant ecology - a review. *Basic and Applied Ecology* **4**: 493–506.

Donnelly JP, Bertness MD. **2001**. Rapid shoreward encroachment of salt marsh cordgrass in response to accelerated sea-level rise. *Proceedings of the National Academy of Sciences* **98**: 14218–14223.

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, *et al.* **2006**. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**: 129–151.

Ellenberg H. **1991**. Zeigerwerte von Pflanzen in Mitteleuropa. *Scripta Geobotanica* **18**: 1–248.

Emery NC, Ewanchuk PJ, Bertness MD. **2001**. Competition and salt-marsh plant zonation: stress tolerators may be dominant competitors. *Ecology* **82**: 2471–2485.

Ford H, Garbutt A, Jones DL, Jones L. **2012**. Impacts of grazing abandonment on ecosystem service provision: Coastal grassland as a model system. *Agriculture, Ecosystems & Environment* **162**: 108–115.

Grime JP, Hodgson JG. **1988**. *Comparative plant ecology: a functional approach to common British species*. London: Unwin Hyman.

Hill MO, Preston CD, Roy D. **2004**. *PLANTATT - attributes of British and Irish plants: status, size, life history, geography and habitats*. Abbotts Ripton: Centre for Ecology & Hydrology.

Hill MO, Roy DB, Mountford JO, Bunce RG. **2000**. Extending Ellenberg's indicator values to a new area: an algorithmic approach. *Journal of Applied Ecology* **37**: 3–15.

Jiménez-Valverde A. **2012**. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography* **21**: 498–507.

Jones L, Angus S, Cooper A, Doody P, Everad M, Garbutt A, Gilchrist P, Hansom J, Nicholls R, Pye K, *et al.* **2011a**. Coastal Margins. UK National Ecosystem Assessment. Understanding nature's value to society. Technical Report. Cambridge, UK: UK National Ecosystem Assessment, UNEP-WCMC, 411–457.

Jones L, Wallace H, M Prosser. **2011b**. *Baseline botanical monitoring and establishing permanent quadrats in embryo primary slack, Whiteford Burrows, 7th and 8th July 2011.* Report to the Countryside Council for Wales.

Jones MLM, Wallace HL, Norris D, Brittain SA, Haria S, Jones RE, Rhind PM, Reynolds BR, Emmett BA. **2004**. Changes in vegetation and soil characteristics in coastal sand dunes along a gradient of atmospheric nitrogen deposition. *Plant Biology* **6**: 598–605.

Latour JB, Reiling R. **1993**. A multiple stress model for vegetation ('move'): a tool for scenario studies and standard-setting. *Science of the Total Environment* **134**: 1513–1526.

Lobo JM, Jiménez-Valverde A, Real R. **2008**. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* **17**: 145–151.

Lowe JAH, Reynolds B, Fowler D, Brittain SA, Hughes S. **1996**. *Orographic enhancement of acidic deposition in Snowdonia. Final report to Welsh Office*. Bangor: Institute of Terrestrial Ecology.

Malpas LR, Smart J, Drewitt A, Sharps E, Garbutt A. **2013**. Continued declines of Redshank *Tringa totanus* breeding on saltmarsh in Great Britain: is there a solution to this conservation problem? *Bird Study* **60**: 370–383.

Manel S, Williams HC, Ormerod SJ. **2001**. Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**: 921–931.

Maun MA, Perumal J. **1999**. Zonation of vegetation on lacustrine coastal dunes: effects of burial by sand. *Ecology Letters* **2**: 14–18.

Mendoza-González G, Martínez ML, Rojas-Soto OR, Vázquez G, Gallego-Fernández JB. **2013**. Ecological niche modeling of coastal dune plants and future potential distribution in response to climate change and sea level rise. *Global Change Biology* **19**: 2524–2535.

Meynard CN, Quinn JF. **2007**. Predicting species distributions: a critical comparison of the most common statistical models using artificial species: Comparison of species-distribution models. *Journal of Biogeography* **34**: 1455–1469.

Plassmann K, Edwards-Jones G, Jones MLM. **2009**. The effects of low levels of nitrogen deposition and grazing on dune grassland. *Science of The Total Environment* **407**: 1391–1404.

Preston CD, Pearman DA, Dines TD. **2002**. *New Atlas of the British and Irish Flora: An atlas of the vascular plants of Britain, Ireland, The Isle of Man and the Channel Islands*. UK: Oxford University Press.

Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A. **2006**. Are niche-based species distribution models transferable in space? *Journal of Biogeography* **33**: 1689–1703.

Ranwell DS. **1972**. *Ecology of Salt Marshes and Sand Dunes*. London: Chapman and Hall.

Real R, Barbosa AM, Vargas JM. **2006**. Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics* **13**: 237–245.

R Foundation for Statistical Computing. **2013**. *R: A language and environment for statistical computing.* Vienna, Austria.

Rhymes J, Wallace H, Fenner N, Jones L. **2014**. Evidence for sensitivity of dune wetlands to groundwater nutrients. *Science of the Total Environment* **490**: 106–113.

Rowe EC, Emmett BA, Smart SM, Frogbrook ZL. **2011**. A new net mineralizable nitrogen assay improves predictions of floristic composition. *Journal of Vegetation Science* **22**: 251–261.

Sing T, Sander O, Beerenwinkel N, Lengauer T. **2005**. ROCR: visualizing classifier performance in R. *Bioinformatics* **21**: 3940–3941.

Smart SM, Henrys PA, Scott WA, Hall J. R, Evans CD, Crowe A, Rowe EC, Dragosits U, Page T, Whyatt JD, *et al.* 2010a. Impacts of pollution and climate change on ombrotrophic *Sphagnum* species in the UK: analysis of uncertainties in two empirical niche models. *Climate Research* **45**: 163–177.

Smart SM, Scott AW, Whitaker J, Hill MO, Roy DB, Critchley NC, Marini L, Evans C, Emmett BA, Rowe EC, *et al.* 2010b. Empirical realised niche models for British higher and lower plants – development and preliminary testing. *Journal of Vegetation Science* **21**: 643–656.

Stace C. **2010**. *New Flora of the British Isles*. UK: Cambridge University Press.

Swets JA. **1988**. Measuring the accuracy of diagnostic systems. *Science* **240**: 1285–1293.

de Vries W, Wamelink GWW, Dobben H van, Kros J, Reinds GJ, Mol-Dijkstra JP, Smart SM, Evans CD, Rowe EC, Belyazid S, *et al.* 2010. Use of dynamic soil–vegetation models to assess impacts of nitrogen deposition on plant species composition: an overview. *Ecological Applications* **20**: 60–79.

Wagner M, Kahmen A, Schlumprecht H, Audorff V, Perner J, Buchmann N, Weisser WW. **2007**. Prediction of herbage yield in grassland: How well do Ellenberg N-values perform? *Applied Vegetation Science* **10**: 15–24.

Yuan X, Ma K, Wang D. **2012**. Partitioning the effects of environmental and spatial heterogeneity on distribution of plant diversity in the Yellow River Estuary. *Science China Life Sciences* **55**: 542–550.

Zelený D, Schaffers AP. **2012**. Too good to be true: pitfalls of using mean Ellenberg indicator values in vegetation analyses. *Journal of Vegetation Science* **23**: 419–431.

Zhu G, Gao Y, Zhu L. **2013**. Delimiting the coastal geographic background to predict potential distribution of *Spartina alterniflora*. *Hydrobiologia* **717**: 177–187.

Figure captions

Fig. 1 Modelled habitat suitability in relation to Ellenberg S score for four representative coastal plant taxa; thrift (*Armeria maritima*), sea spurge (*Euphorbia paralias*), grey club-rush (*Schoenoplectus tabernaemontani*) and sand sedge (*Carex arenaria*) (Hill *et al.*, 2004). Vertical dotted lines indicate the Ellenberg S score. Figure created in R version 3.0.2

Fig. 2 Modelled habitat suitability in relation to Ellenberg S and F scores for four coastal taxa: *Armeria maritima*, found predominantly in saltmarshes and on sea cliffs; *Euphorbia paralias*, often found on mobile or semi-stable sand-dunes; *Schoenoplectus tabernaemontani*, frequent in wet areas near the sea e.g. tidal channels and dune-slacks; and *Carex arenaria*, a dominant plant of fixed dunes and grassy maritime areas (Preston *et al.*, 2002; Stace, 2010). Figure created in R version 3.0.2

Fig. 3 Histogram of AUC values obtained from model testing of 45 coastal plant species niche models against an independent dataset. Fill represents the AUC thresholds of <0.7 being indicative of poor model predictive ability (white), 0.7-0.9 suggesting moderate performance (grey) and over 0.9 suggesting good model performance (black) (Swets, 1988; Manel *et al.*, 2001). Figure created in R version 3.0.2

Table 1. Number of coastal plant species niche models containing the eight explanatory variables included in the analysis (out of 132 models in total). Variables were entered as both linear and quadratic terms and all two way linear interactions were also included (see Online Resource 2 for model coefficients).

Model term		**Number of models containing term**
Ellenberg S	Linear	56
	Quadratic	110
Ellenberg F	Linear	83
	Quadratic	108
Ellenberg R	Linear	69
	Quadratic	89
Ellenberg N	Linear	47
	Quadratic	97
Canopy height	Linear	52
	Quadratic	97
Max. July temperature	Linear	74
	Quadratic	82
Min. January temperature	Linear	57
	Quadratic	77
Annual precipitation	Linear	55
	Quadratic	58

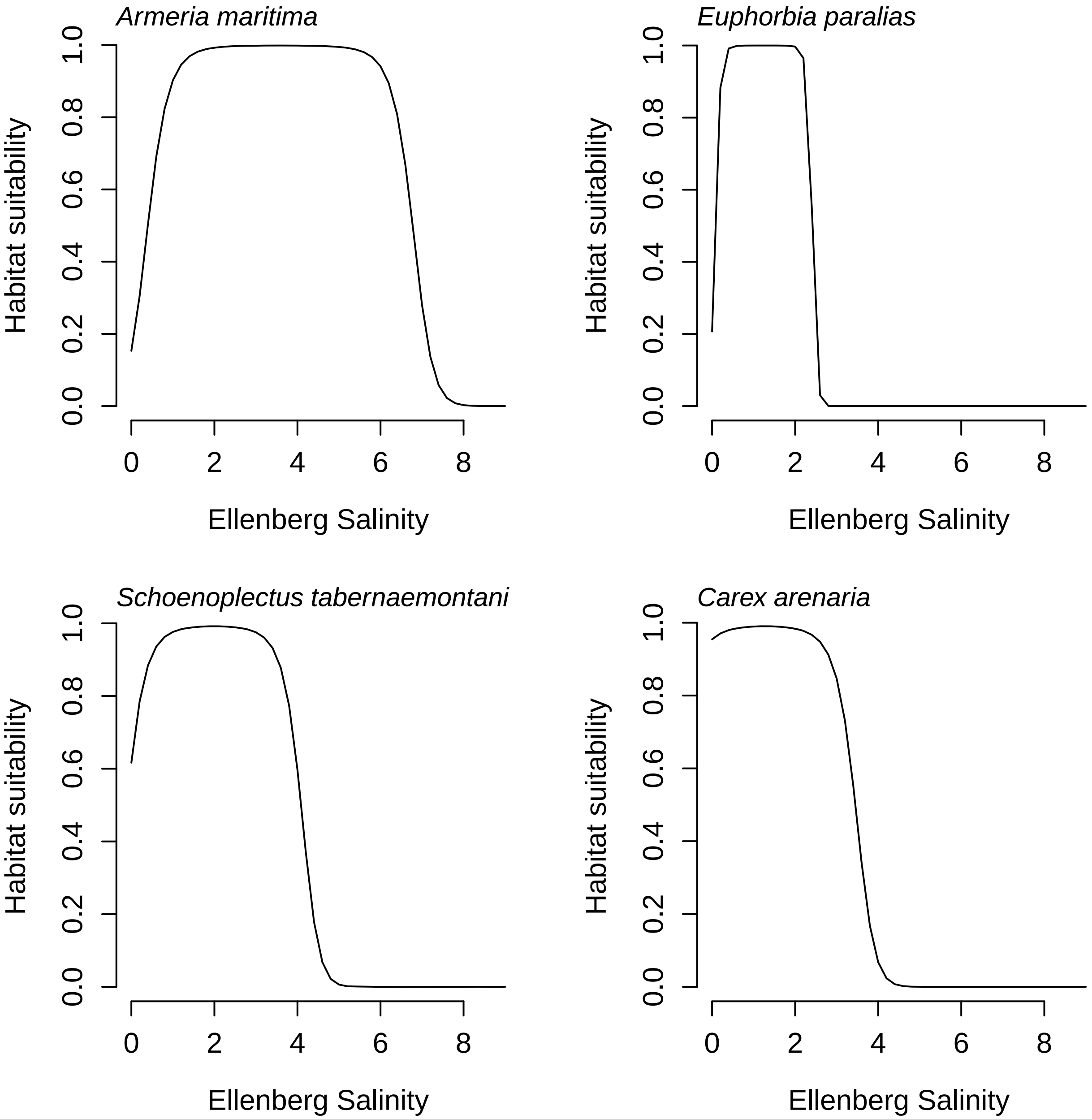
Table 2. Table of AUC values obtained from testing 45 coastal plant species niche models against an independent dataset.

Species name	Number of occurrences in training data	Number of occurrences in test data	AUC from full model	AUC without Ellenberg S score	Difference (AUC from full model - AUC without Ellenberg S score)
Agrostis stolonifera	5585	1028	0.86	0.76	0.10
Alopecurus geniculatus	389	3	0.96	0.96	0.00
Apium graveolens	24	30	0.54	0.54	0.00
Armeria maritima	1621	484	0.93	0.90	0.03
Aster tripolium	903	2820	0.69	0.68	0.01
Atriplex littoralis	35	146	0.38	0.61	-0.23
Atriplex patula	60	9	0.43	0.49	-0.06
Atriplex portulacoides	764	1500	0.68	0.67	0.01
Beta vulgaris	59	144	0.71	0.56	0.14
Bolboschoenus maritimus	86	423	0.65	0.46	0.19
Carex arenaria	1481	5	0.92	0.92	0.00
Carex distans	141	15	0.98	0.84	0.14
Carex extensa	116	39	0.81	0.70	0.11
Carex otrubae	110	10	0.97	0.65	0.32
Cochlearia anglica	185	575	0.61	0.57	0.04
Crithmum maritimum	108	4	0.69	0.84	-0.16
Elytrigia atherica	498	1058	0.65	0.51	0.14
Elytrigia repens	951	90	0.71	0.81	-0.10
Festuca rubra agg.	8033	1503	0.81	0.80	0.01
Glaux maritima	902	1241	0.85	0.80	0.05
Honckenya peploides	128	7	0.76	0.86	-0.10
Hordeum marinum	30	8	0.62	0.62	0.00
Inula crithmoides	66	23	0.9	0.79	0.11
Juncus gerardii	617	545	0.82	0.70	0.12
Juncus maritimus	274	234	0.49	0.50	-0.01
Limonium humile	31	166	0.78	0.76	0.02
Limonium vulgare	411	378	0.83	0.81	0.02
Oenanthe lachenalii	244	90	0.83	0.48	0.36
Parapholis strigosa	67	102	0.5	0.71	-0.20
Phragmites australis	888	218	0.68	0.50	0.18
Plantago coronopus	737	52	0.92	0.86	0.06
Plantago maritima	2068	1513	0.88	0.84	0.04
Puccinellia maritima	1138	2446	0.76	0.72	0.03
Salicornia europaea agg.	472	388	0.55	0.62	-0.06
Sarcocornia perennis	251	98	0.92	0.91	0.01
Schoenoplectus tabernaemontani	19	17	0.66	0.39	0.27
Scorzoneroides autumnalis	1576	58	0.93	0.88	0.05
Sedum anglicum	324	3	1	0.95	0.04
Sonchus arvensis	325	29	0.93	0.79	0.13
Spergularia marina	166	232	0.77	0.77	0.00
Spergularia media	385	794	0.81	0.75	0.06
Suaeda maritima	649	1608	0.81	0.75	0.06
Suaeda vera	214	42	0.68	0.70	-0.02
Trifolium fragiferum	50	9	0.8	0.39	0.41
Triglochin maritima	776	1582	0.78	0.79	-0.01

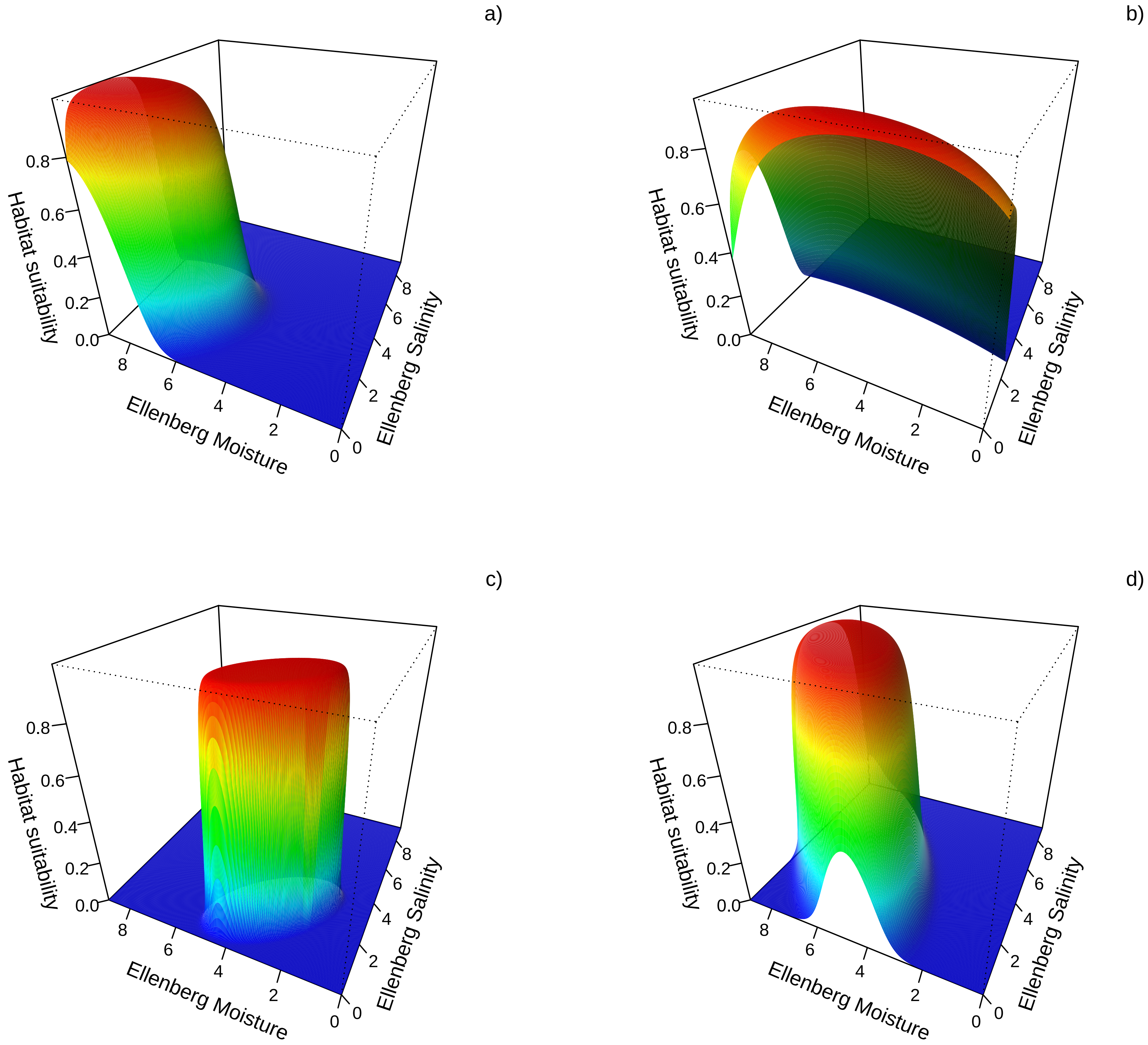
Supplementary Material

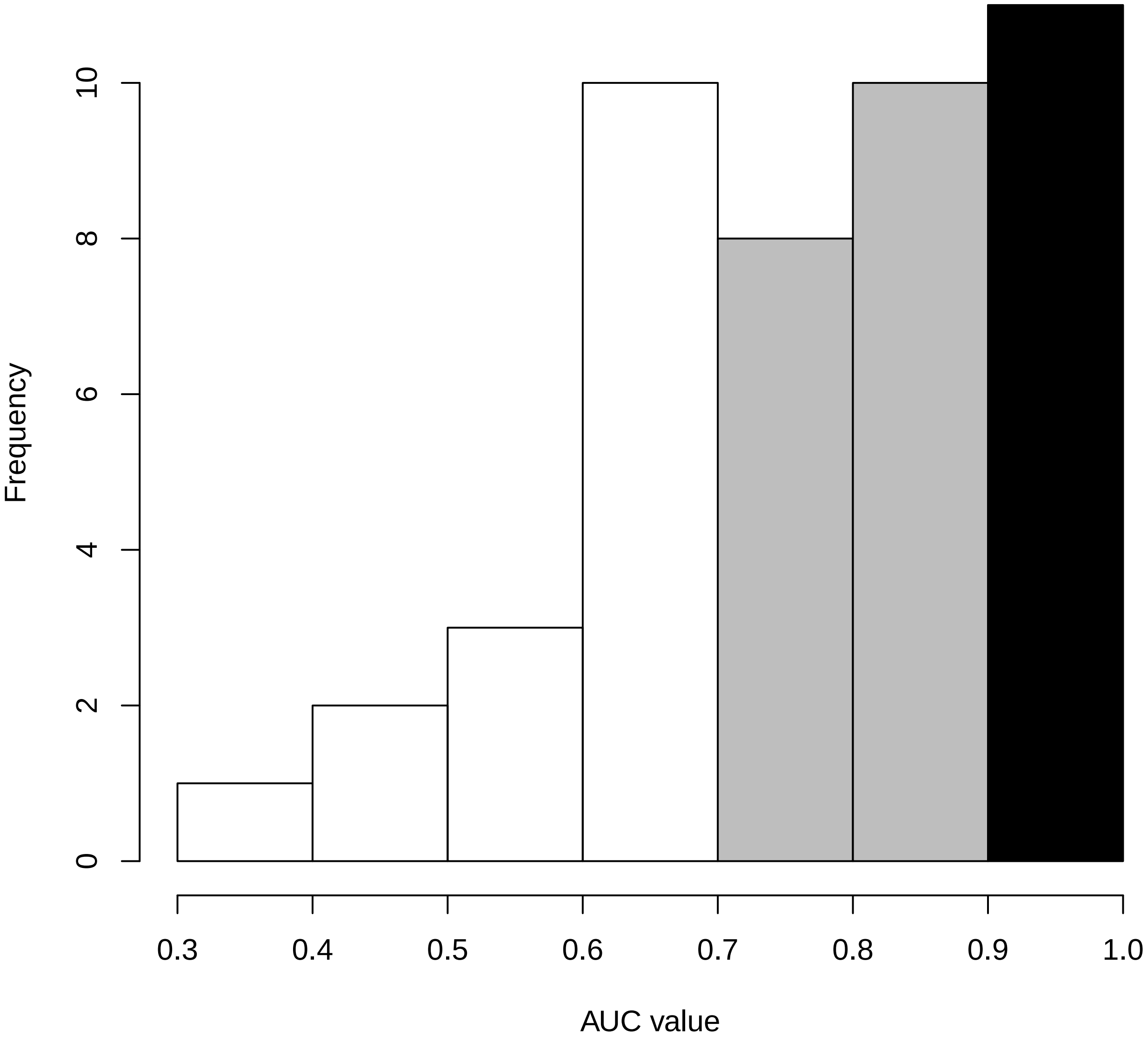
Online Resource 1. Receiver operator characteristic plots for all coastal plant niche models a) with the full model b) without the salinity term.

Online Resource 2. Table of model coefficients for 132 niche models of coastal plant species



Armeria maritima
Euphorbia paralias
Schoenoplectus tabernaemontani
Carex arenaria
Habitat suitability
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
Ellenberg Salinity


a)
b)
c)
d)
Habitat suitability
0.0
0.2
0.4
0.6
0.8
Ellenberg Moisture
0
2
4
6
8
Ellenberg Salinity
0
2
4
6
8



0
2
4
6
8
10
Frequency
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
AUC value