Journal of Applied Ecology

British Ecological Society

Journal of Applied Ecology 2016 doi: 10.1111/1365-2664.12681

Consensus forecasting of intertidal seagrass habitat in the Wadden Sea

Eelke O. Folmer[1]*, Justus E.E. van Beusekom[2], Tobias Dolch[3], Ulf Gräwe[4], Marieke M. van Katwijk[5], Kerstin Kolbe[6] and Catharina J.M. Philippart[1,7]

[1]NIOZ Royal Netherlands Institute for Sea Research, Department of Coastal Systems and Utrecht University, PO Box 59, Den Burg Texel 1790 AB, the Netherlands; [2]Helmholtz-Zentrum Geesthacht, Institute for Coastal Research, Max-Planck-Strasse 1, Geesthacht 21502, Germany; [3]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Wadden Sea Station Sylt, Hafenstr. 43, List 25992, Germany; [4]Department for Physical Oceanography and Instrumentation, Leibniz Institute for Baltic Sea Research Warnemuende, Seestrasse 15D, Rostock-Warnemuende 18119, Germany; [5]Department of Environmental Science, Radboud University Nijmegen, Heyendaalseweg 135, Nijmegen 6525 AJ, the Netherlands; [6]Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency (NLWKN), Am Sportplatz 23, Norden D-26506, Germany; and [7]Department of Physical Geography, University of Utrecht, PO Box 80·115, Utrecht 3508 TC, the Netherlands

Summary

1. After the dramatic eutrophication-induced decline of intertidal seagrasses in the 1970s, the Wadden Sea has shown diverging developments. In the northern Wadden Sea, seagrass beds have expanded and become denser, while in the southern Wadden Sea, only small beds with low shoot densities are found. A lack of documentation of historical distributions hampers conservation management. Yet, the recovery in the northern Wadden Sea provides opportunity to construct robust habitat suitability models to support management.

2. We tuned habitat distribution models based on 17 years of seagrass surveys in the northern Wadden Sea and high-resolution hydrodynamics and geomorphology for the entire Wadden Sea using five machine learning approaches. To obtain geographically transferable models, hyperparameters were tuned on the basis of prediction accuracy assessed by non-random, spatial cross-validation. The spatial cross-validation methodology was combined with a consensus modelling approach.

3. The predicted suitability scores correlated amongst each other and with the hold-out observations in the training area indicating that the models converged and were transferable across space. Prediction accuracy was improved by averaging the predictions of the best models.

4. We graphically examined the relationship between the consensus suitability score and independent presence-only data from outside the training area using the area-adjusted seagrass frequency per suitability class (continuous Boyce index). The Boyce index was positively correlated with the suitability score indicating the adequacy of the prediction methodology.

5. We used the plot of the continuous Boyce index against habitat suitability score to demarcate three habitat classes – unsuitable, marginal and suitable – for the entire international Wadden Sea. This information is valuable for habitat conservation and restoration management.

6. Divergence between predicted suitability and actual distributions from the recent past indicates that unaccounted factors limit seagrass development in the southern Wadden Sea.

7. *Synthesis and applications*. Our methodology and data enabled us to produce a robust and validated consensus habitat suitability model. We identified highly suitable areas where intertidal seagrass meadows may establish and persist. Our work provides scientific underpinning for effective conservation planning in a dynamic landscape and sets monitoring priorities.

Key-words: Boyce index, ensemble forecasting, habitat distribution model, habitat management, machine learning, model transferability, seagrass, Wadden Sea, *Zostera marina*, *Zostera noltii*

*Correspondence author. E-mail: E.O.Folmer@gmail.com

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Introduction

Seagrass beds are highly valued for their ecosystem services (Costanza *et al.* 1997; Orth *et al.* 2006) related to nutrient cycling and provisioning resources and habitat for other species (Beck *et al.* 2001). On a global scale, seagrass distributions have drastically declined leaving behind unvegetated but potentially suitable habitat (Waycott *et al.* 2009). The effectiveness of seagrass conservation management such as safeguarding suitable habitats is often hampered by the lack of systematic documentation of historical seagrass distributions and by changes in environmental conditions in space or time. Another complication is that seagrass distributions in energetic environments like intertidal zones show local extinction–colonization dynamics which necessitates identification and protection of vegetated and suitable unvegetated habitat (Valle *et al.* 2013; Suykerbuyk *et al.* 2016). To identify habitat where populations are likely to establish and persist, habitat distribution models (HDMs) can be used (Elith & Leathwick 2009; Guisan *et al.* 2013).

To obtain predictions that are accurate beyond the training area, HDMs need to be generalizable or transferable between geographical areas (Wenger & Olden 2012). However, spatial and temporal heterogeneity in ecological relationships and environmental conditions may limit transferability leading to biased predictions (Randin *et al.* 2006; Bahn & McGill 2013). For instance, associations between variables may exist in a narrow geographical range but may not hold under a wider range or *vice versa*. Accuracy assessment and model selection on the basis of non-random cross-validation has been proposed to reduce the risk of overfitting and to improve transferability through space and time (Araújo *et al.* 2005; Wenger & Olden 2012). In the case of the spatial variant of this cross-validation method, data are split into geographical subsets that are more distinct than random subsets. The idea behind this method is that the validation sets differ from training sets similarly as the area for which predictions are required. Hence, this method yields more transferable models and thus more accurate predictions under new conditions than models that are tuned by random cross-validation (Wenger & Olden 2012).

Another complication may stem from prediction variability between modelling techniques (Elith & Leathwick 2009). This has led to the common practice of using multiple models followed by assessment and analysis of the agreement between predictions (Segurado & Araújo 2004; Thuiller 2004). Furthermore, various studies have shown that ensemble forecasting – forecasting by pooling of individual model predictions – may increase accuracy (e.g. Araújo & New 2007; Marmion *et al.* 2009). In a similar vein as Wenger *et al.* (2013), we propose combining both methods by pooling predictions of models that are tuned on the basis of transferability and to take into account the possibility that models may perform differently across geographical subsets of the data. This methodology thus combines the benefits of optimizing transferability of models by non-random cross-validation while improving accuracy by pooling predictions and the construction of a consensus forecast.

Before further detailing the modelling approach, we present a conceptual model describing how desiccation, hydrodynamics and geomorphology control the distributions of intertidal *Z. noltii* and *Z. marina* (Fig. 1) in the Wadden Sea (Den Hartog & Polderman 1975). In contrast to subtidal seagrass species, *Z. noltii* and *Z. marina* possess adaptations to reduce water tissue loss and maintain high rates of photosynthesis during emersion, which allows them to occur in intertidal zones. Nevertheless, upper intertidal distributions are limited by desiccation stress resulting from long-lasting aerial exposure (Leuschner, Landwehr & Mehlig 1998; van Katwijk *et al.* 2000; de Jong, Katwijk & Brinkman 2005; Shafer, Sherman & Wyllie-Echeverria 2007).

Hydrodynamic stress limits the distribution of seagrass by preventing establishment and damaging and uprooting of plants (Fonseca & Bell 1998; van Katwijk *et al.* 2000; Koch 2001). In intertidal systems, hydrodynamic stress increases along the depth gradient towards the subtidal area. We model seagrass distributions under the assumption that the hydrodynamic factors are exogenous, independent forcing factors. However, when discussing the modelling results, we keep in mind that seagrass meadows locally affect hydrodynamics and that feedbacks may exist (Fonseca & Bell 1998; Koch 2001; van der Heide *et al.* 2007; Folmer *et al.* 2012).

The geomorphology and sediment properties of soft sediment intertidal systems are controlled by hydrodynamics. Steep slopes occur at and near gullies where tidal currents are strong while shallow slopes occur in calmer areas. Therefore, the slope can be used as a proxy for past and present hydrodynamics and thus be used to predict seagrass distributions. Seagrass is expected to occur in areas with shallow slopes only; it is not assumed that the

Fig. 1. *Zostera noltii* and *Zostera marina* growing on the mudflats in the Schleswig–Holstein area in the German Wadden Sea.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

slope *per se* affects seagrass. The median grain size and mud content of the sediment can also be used as a proxy for hydrodynamic conditions. In addition, the sediment properties themselves may affect seagrass growth and occurrence and, *vice versa,* seagrass may affect sediment properties by promoting accretion of fine particles (Koch 2001; Bos *et al.* 2007; Folmer *et al.* 2012). However, in the Wadden Sea, the accumulated fine sediments are released during winter when the above-ground parts of intertidal seagrass meadows dissipate (Bos *et al.* 2007). Therefore, the sediment properties are modelled as exogenous factors.

Zostera plants tolerate a wide salinity range. Den Hartog (1970) reports a range of 5–42 psu and Nejrup & Pedersen (2008) report optimal growth of *Z. marina* between 10 and 35 psu. Even under narrower ranges, nearly the entire Wadden Sea is within the tolerable range. To avoid the risk of overfitting due to spurious correlations, salinity is not included as a predictor.

Both *Zostera* species can have annual and perennial life cycles in the Wadden Sea. However, the smaller *Z. noltii* often survives winter with rhizomes, while *Z. marina* largely recruits from seed each year. The predominantly subtidal, perennial morph of *Z. marina* went extinct during the 1930s and never returned in this area (Den Hartog 1987). Eutrophication was an important factor causing declines of intertidal *Z. noltii* and *Z. marina* populations between the 1970s and 1990s (Philippart & Dijkema 1995; van Katwijk *et al.* 2010; Dolch, Buschbaum & Reise 2013). It led to higher densities of epiphytes growing on seagrass leaves (Philippart & Dijkema 1995; Dolch, Buschbaum & Reise 2013) and increased growth of green algae that smothered seagrass beds (Reise, Herre & Sturm 1989). The nutrient loads may have also affected seagrass physiology because the metabolism of seagrasses is adapted to low nutrient conditions (Burkholder, Tomasko & Touchette 2007).

Data from disparate seagrass surveys in the Wadden Sea are available in grey and formally published literature. Dijkema, van Thienen & van Beek (1989) compiled data from various intertidal surveys between the 1950s and 1970s (see Appendix S1, Supporting information). Because declines probably already took place before this period, our knowledge of past distributions remains limited. It is clear, however, that seagrass extents in the southern Wadden Sea have remained low which has triggered substantial conservation concern and restoration efforts in the Dutch Wadden Sea. Up until now, restoration efforts have failed to establish new persisting seagrass meadows. In the northern Wadden Sea, however, seagrass extents have increased and the total surface area is currently larger than in the 1930s (Reise & Kohlus 2008; Dolch, Buschbaum & Reise 2013).

The conceptual model illustrates that nonlinear relationships and interactions between variables occur which may give rise to complex and spatiotemporally variable response surfaces. Skewed distributions and nonlinear relationships are more accurately described by machine learning approaches than by GLMs or climate envelopes for example (Barry & Elith 2006). Because our objective is to capture complex response surfaces to predict habitat suitability, we apply Random Forests, Gradient Boosting Machines, Multivariate Adaptive Regression Splines, Support Vector Machines and Generalized Additive Models (Hastie, Tibshirani & Friedman 2009; Kuhn & Johnson 2013). These approaches have performed well in other species and habitat distribution modelling studies (Barry & Elith 2006; Drake, Randin & Guisan 2006; Leathwick, Elith & Hastie 2006; Prasad, Iverson & Liaw 2006; Elith, Leathwick & Hastie 2008).

The choice of the modelling framework for developing HDMs depends on the aim of the study (Guisan & Thuiller 2005; Elith & Leathwick 2009; Dormann *et al.* 2012). When, for example, HDMs are developed to improve understanding of species' responses to environmental variables, statistical approaches such as GLM(M)s are often selected because relationships between responses and predictors are straightforward to analyse. However, failure to account for possible spatial autocorrelation in the response variable may lead to biased parameter estimates and influence statistical inference (Dormann *et al.* 2007; Folmer, Olff & Piersma 2012). When HDMs are developed to predict (potential) distributions of species' occurrences across geographical areas (as in our case), machine learning approaches are often preferred due to their predictive accuracy (Barry & Elith 2006; McCue, McGrath & Wiersma 2014). Furthermore, when models are tuned on the basis of transferability, they are not only guarded against overfitting, which may occur when there is heterogeneity in ecological relationships, but also against bias resulting from spatial autocorrelation (Wenger & Olden 2012). A disadvantage of many machine learning algorithms is that interpretation can be difficult because relationships between response and predictor variables are difficult to analyse which is why they are sometimes called 'black boxes'. However, for some machine learning algorithms, there are approaches which may help to gain insight into (marginal) dependencies between variables (Elith, Leathwick & Hastie 2008; Hastie, Tibshirani & Friedman 2009).

The main purpose of this paper was to develop, test and apply the 'transferability-consensus' approach to identify potential seagrass habitat in the Wadden Sea. Because the models will be used for geographical extrapolation, they are tuned on the basis of transferability. We use data from extensive seagrass surveys in the northern Wadden Sea where the extents are currently above the levels in the 1930s (Reise & Kohlus 2008; Dolch, Buschbaum & Reise 2013). We use predictor variables based on high-resolution hydrodynamic modelling outcomes, and geomorphology and sediment properties on the scale of the entire Wadden Sea. This combination of data provides ideal opportunities to develop robust HDMs to identify areas potentially suited for seagrass in the entire Wadden

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

Sea. After showing that the 'transferability-consensus' approach leads to improved predictions, we use the consensus predictions to map intertidal seagrass habitats for the Wadden Sea at large.

Materials and methods

STUDY AREA AND SPECIES

The Wadden Sea – located in the south-eastern coastal zone of the North Sea bordering Denmark, Germany and the Netherlands – is the largest coherent system of intertidal flats in the temperate zones of the world (Fig. 2). It spans a distance of nearly 500 km and the area of intertidal flats is about 5000 km^2. It was placed on UNESCO's World Heritage list because of its 'universally outstanding natural values'. It consists of intertidal flats, shallow subtidal flats, drainage gullies and deeper inlets and channels. Barrier islands are found in the entire Wadden Sea except in the central Wadden Sea. Tidal amplitudes range between 1·5 and 3·0 m in the north-eastern and south-western Wadden Sea and exceed 3·0 m in the central part. Tidal currents and exposure to waves strongly differ between regions due to differences in tidal range, geomorphology, fetch and the occurrence of barrier islands. The intertidal flats consist mostly of sand mixed with fine-grained muddy sediments; the fractions of fine-grained particles increase towards the shores.

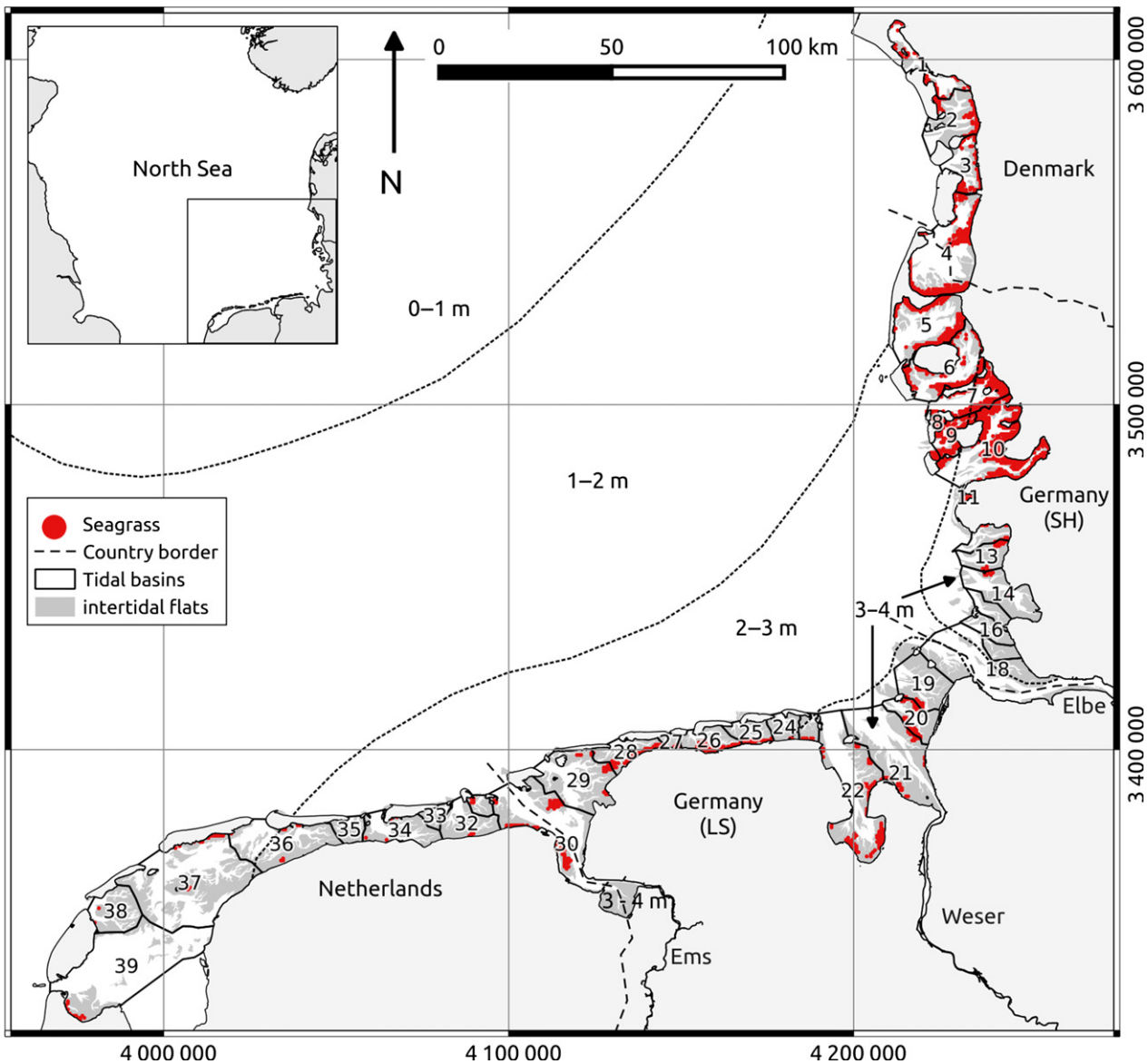

Fig. 2. The Wadden Sea with all known intertidal seagrass occurrences. There are regional differences in the monitoring intensity and methodology underlying the presented seagrass data. The seagrass beds (in red) are drawn larger than their actual sizes to increase visibility of small seagrass beds in the Netherlands and Lower Saxony (LS). The number ranges (i.e. 0–1 m, 1–2 m, 2–3, 3–4 m) separated by dashed lines denote the tidal ranges at spring tide (redrawn from Wiersma *et al.* 2009). The geographical extent ranges between 52°57′ – 55°37′ north and 4°44′ – 8°12′ east.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

DATA

Seagrass areal extents

We used intertidal seagrass distribution data from Schleswig–Holstein for the years 1995–2011 (Reise & Kohlus 2008; Dolch, Buschbaum & Reise 2013). Each year, between June and September when seagrass attains its seasonal maximum cover, three surveyors in a four-seated Cessna aircraft flying during low tide at heights between 300 and 500 m surveyed intertidal flats for seagrass. Each surveyor mapped seagrass beds on paper with preprinted satellite images showing the contours of land, sea and intertidal flats. Then, a synthesis of all three maps was created and the seagrass mapping results were digitized as vector files in a GIS. Since 2007, yearly ground-truth field surveys were performed in July and August; they showed that a minimum cover density of about 20% is required to detect seagrass from the air with satisfying reliability (Dolch, Buschbaum & Reise 2013).

The response variable seagrass prevalence (p) was constructed by summarizing the yearly contours in a single raster grid with cell size of 200 × 200 m covering the intertidal area of Schleswig–Holstein (Fig. 3, n = 39 645 grid cells). The origin and resolution of the raster is identical to the raster with the predictor variables described below. Prevalence is the fraction of the time span of 17 years that the represented mudflat patch was covered with seagrass (0: never covered, 1: always covered) (Fig. 3a). The models were tuned on log-transformed prevalence. The log transformation was used to reduce the relative difference between high prevalence values while maintaining relatively large differences between the lower prevalence values. To avoid taking the log of zero, we added half of the smallest non-zero value to all values ($0{\cdot}5 \times (1/17) \approx 0{\cdot}03$).

We furthermore compiled a spatial seagrass data set (i.e. a vector file) of all known intertidal seagrass occurrences in Denmark, Lower Saxony and the Netherlands between the 1970s and 2015 to investigate the quality of the model predictions outside Schleswig–Holstein. Due to differences in methodology and timing of the surveys, these data sets are distinct and they were not used for model tuning; they were used as presence-only data for validation (Boyce *et al.* 2002) (details of these data sets are described in Appendix S1 and included in Fig. 2).

Predictor variables

Predictor variables summarizing the hydrodynamics were obtained from the General Estuarine Transport Model (GETM, Burchard & Bolding 2002), which is designed for coastal ocean simulations with drying and flooding of intertidal flats. GETM is a three-dimensional baroclinic open source model with hydrostatic and Boussinesq assumptions. In GETM, the three-dimensional hydrostatic momentum equations are solved on a staggered C-grid, which for the present application is horizontally Cartesian. To calculate vertical turbulent fluxes, GETM uses turbulence closure models from the General Ocean Turbulence Model (GOTM, Umlauf & Burchard 2003). In the present study, a k-epsilon model is used with an algebraic second moment closure. To vertically discretize the model equations, the water column is resolved with 25 adaptive terrain-following depth layers. A bathymetry with resolution 200 m was constructed on the basis of data made available by Rijkswaterstaat for the Dutch Wadden Sea (resolution 50 m), the project AufMod (resolution 50–200 m) for the German Wadden Sea and the Danish Maritime Safety Administration (resolution 200 m) for Danish waters. The numerical set-up that is based on this 200 × 200 m topography is the end member of a hierarchy of four nested models covering the North Atlantic, the North Sea, and the southern North Sea (Gräwe *et al.* 2015). The bottom layer in the model has a maximum thickness of 0·3 m to resolve properly the bottom boundary layer. The forward integration of the model is done with a time step of 40 s. Thus, every 40 s, the current velocity, temperature, salinity and turbulent quantities are prognostically computed.

Atmospheric forcing is taken from the operational model of the German Weather Service (DWD) with a spatial resolution of 7 km. Atmospheric fields, which include precipitation, are provided every 3 h. In between, GETM uses linear interpolation. The model is further forced by freshwater run-off from rivers. For all rivers, observations at daily or finer resolutions are available. The Dutch run-off data were provided by the Royal Netherlands Institute for Sea Research and Rijkswaterstaat (temporal resolution 15 min), the data for the East Frisian Wadden Sea by the German Federal Institute of Hydrology (temporal resolution 24 h) and those for the North Frisian and the Danish Wadden Sea by Schleswig-Holstein's Government-Owned Company for Coastal Protection, National Parks and Ocean Protection (temporal resolution 24 h). The run-off data ($m^3\ s^{-1}$) with a temporal resolution of 24 h were linearly interpolated to obtain run-off data at the 15 min resolution. For each 40 s time step, the run-off ($m^3\ s^{-1}$ at the 15 min interval) is multiplied by 40 s to obtain run-off volumes (m^3) on the resolution of 40 s. The model was run for the period 2008–2011. The year 2008 was not included in the computation of the summary statistics since it was considered as spin up for the hydrodynamics. For a detailed description of the model set-up and performance, we refer to Duran-Matute *et al.* (2014). The model output was used to obtain estimates of mean exposure time (i.e. the mean fraction of time that the seabed is exposed to the air, *exptime*) and mean bottom shear stress due to currents (τ, *taub*) over the period 2009–2011.

The bathymetry was used to calculate the *slope* of each grid cell with the R package 'raster' (Hijmans 2013); the algorithms used are appropriate for smooth surfaces. We used the values of four neighbouring cells to calculate the *slope*.

High-resolution sediment data (*mud*: fraction sediment <63 μm; and *mgs*: median grain size) covering the Dutch and German Wadden Sea, developed within the AufMod project, were provided by the German Federal Maritime and Hydrographic Agency. Missing data for Denmark were imputed with the K-nearest neighbour model. This approach computes the new sample by taking the average value of the samples in the complete set that are closest to it; K was set to 5. Prior to imputation, the predictor variables were transformed by means of the Box–Cox family of power transformations (Venables & Ripley 2002). To impute *mgs* and *mud* for the Danish part of the Wadden Sea, the variables *slope*, *taub*, *taub.Q75* (75th percentile of bottom shear stress), *exptime*, *dshore* (shortest distance to shoreline) were used.

HABITAT DISTRIBUTION MODELLING

Training

Habitat distribution models (HDMs) were developed by relating log-transformed seagrass prevalence to the five Box–Cox

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

transformed environmental predictor variables, that is $log(p+\varepsilon) = f(slope, exptime, taub, mud, mgs)$. Because there was no reason to assume that one particular modelling approach would perform best, we used five approaches: Random Forests (RF), Gradient Boosting Machines (GBM), Multivariate Adaptive Regression Splines (MARS), Support Vector Machines (SVM) and Generalized Additive Models (GAM). The modelling principles are described in Appendix S2.

We tuned and compared the predictive capacity of the HDMs by means of non-random, spatial cross-validation. In a similar vein to Araújo *et al.* (2005) and Wenger & Olden (2012), we split the data into seven regionally distinct subsets which are more independent than random subsets. We used the borders of the tidal basins of the Wadden Sea (Folmer *et al.* 2014) to delineate the seven regions. Observations from small adjacent basins were combined so that the subsets contained roughly equal numbers of data points while ensuring that all subsets contained non-zero observations (Fig. 3b and Table 1). To assess possible bias caused by ignoring spatial heterogeneity, we also tuned the models by conventional random cross-validation where the data set was randomly split into 10 equally sized, non-overlapping, stratified subsets. Stratification across bins of the response variable ensured that the probability distributions of the response variable between subsets were similar (Kuhn & Johnson 2013).

Models were fitted to data remaining after holding out one subset; the procedure was repeated for each combination of holdout set and remaining data. A grid search was used to tune across model-specific hyperparameters (described in Appendix S2) by minimizing root-mean-squared error (RMSE) between predictions and observations of the holdout sets.

Consensus habitat suitability within and outside the training area

Within training area. The non-random cross-validation procedure using five models m resulted in five sets of predictions per holdout set which were used to compute the consensus predictions. For each grid cell i in the holdout sets, the consensus suitability score was computed by taking the average of the five predicted values, S_{mi}, weighted by the inverse of the individual variances,

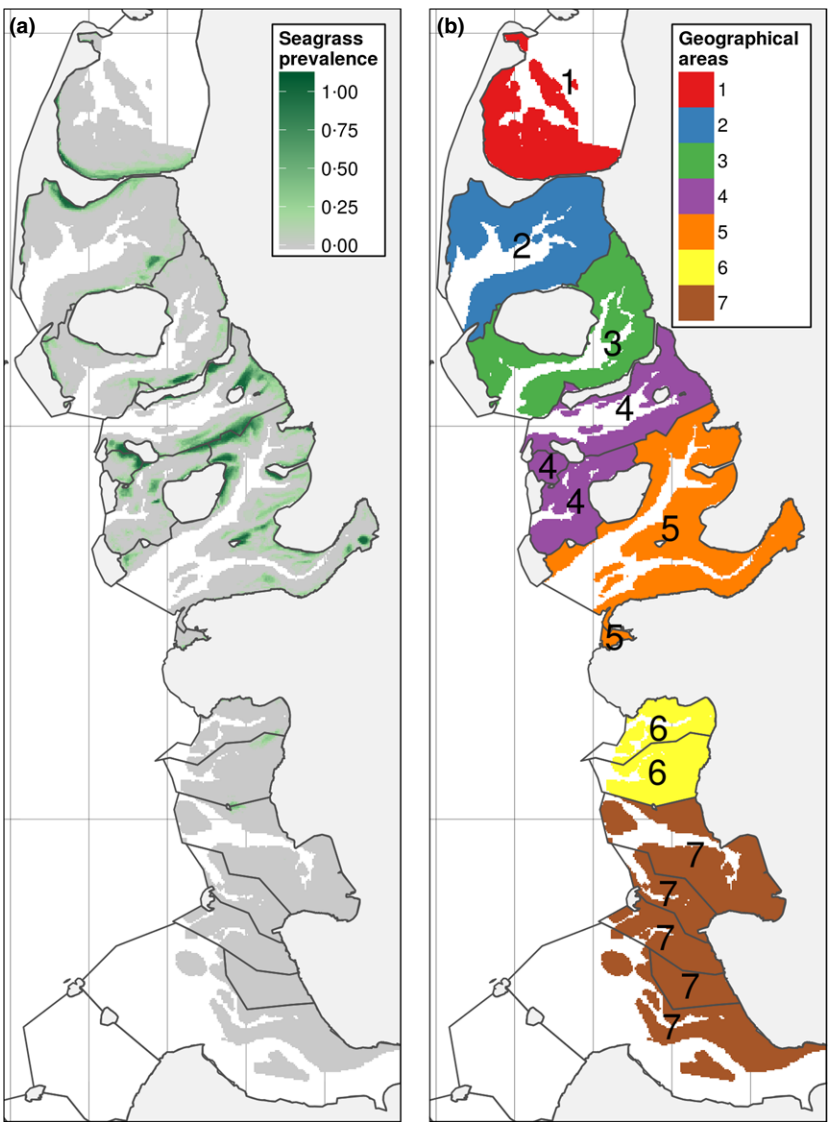

Fig. 3. The Schleswig–Holstein training area used for developing habitat distribution models. (a) Seagrass prevalence in the period 1995–2011. (b) The seven tidal basins or sets of tidal basins used to split the data for spatial cross-validation are indicated by different colours.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

that is $\bar{S}_i = \frac{\sum_{m=1}^{m=5} w_{mi} S_{mi}}{\sum_{m=1}^{m=5} w_{mi}}$ and $w_{mi} = \sigma_{mi}^{-2}$. The transferability of the five models and that of the consensus predictions to each of the holdout sets were compared by the RMSE.

Outside training area. The consensus habitat suitability for the entire Wadden Sea was computed on the basis of five models fitted to seven holdout sets in the training area. This resulted in 35 predicted suitability scores per grid cell. The consensus suitability score was computed similarly as above except that in this case there were 35 predicted scores to average. We also computed the ensemble spread; it is similar to the standard variance of the predictions except that it is based on the weighted average,

$$\bar{\sigma}_i = \frac{\sum_{m=1}^{m=35} w_m (S_{im} - \bar{S}_i)^2}{\sum_{m=1}^{m=35} w_{im}}$$

Presence-only validation and classification

The presence-only data from the Netherlands, Lower Saxony and Denmark were used to examine the consensus predictions outside the training area of Schleswig–Holstein. Boyce *et al*. (2002) present an index to validate modelled suitabilities (in their case 'resource selection functions') with out-of-sample, presence-only data. The index is the frequency of presence per suitability class divided by the area (or number of grid cells) of that class. We used an improved version of the index – labelled 'continuous Boyce index' – where rolling means of the habitat suitability score (HSS) are used instead of fixed classes (Hirzel *et al*. 2006). We constructed a plot of the index against the rolling means of HSS and used it to set boundaries between habitat classes. If the index was less than one, the habitat was classified 'unsuitable' because there were fewer occurrences than expected if the seagrass observations were randomly distributed in space. The second class is the 'marginal' class where the index is greater than one but substantially lower than the third 'suitable' class (Hirzel *et al*. 2006). The value of the HSS that separates the marginal and suitable class was chosen on the basis of the shape of the curve.

Data management, preparation and analysis were done with R (www.r-project.org). The R package *caret* (Kuhn 2014) was used for transformations, imputation and tuning.

Results

TRANSFERABILITY WITHIN THE TRAINING AREA

None of the five models predicted the observations best in each of the seven subsets. Furthermore, the differences in transferability between the models are relatively small (Table 1, see Appendix S3 for parameters). SVM scored best (lowest RMSE, highest transferability) three times, MARS three times (of which two ties), GBM tied twice, and RF and GAM tied once. The average RMSE of MARS was lowest followed by GBM, SVM, GAM and RF. The average transferability of the consensus model equalled the transferability of the MARS model. Furthermore, the RMSEs of the consensus predictions were close to the best predicting models for holdout sets 1–6; for holdout set 7, it was 39% higher.

The average RMSEs based on conventional random cross-validation were substantially lower than the average RMSEs based on non-random cross-validation. RF especially showed poor transferability, while its average RMSE was lowest in the case of random cross-validation.

HABITAT SUITABILITY WITHIN THE TRAINING AREA

The five sets of predictions correspond well with each other and with the observed seagrass prevalences (Fig. 4). The consensus model correctly identifies all locations where seagrass occurred (no false negatives) and it rarely predicts high suitability where no seagrass has been observed. The mapped residuals (Fig. 4: bottom row, third panel) visualize the differences between consensus predictions and observations. The figure shows that the consensus suitability scores tend to be lower than the observed values at locations where the prevalence is high (i.e. underprediction of high values, red areas in the last panel of Fig. 4). Overall, the consensus predictions fit the observations well, and in the following section, an ensemble of 35 models based on the entire training area will be used to predict habitat suitability outside Schleswig–Holstein.

Table 1. Transferability assessment of the best tuned models (one for each modelling framework) on the basis of non-random cross-validation for the seven spatial subsets within the training area of Schleswig–Holstein. RMSEs are given for RF, GBM, MARS, SVM, GAM predictions and for the consensus prediction (CONS). The row RMSE (random CV) gives the average RMSE of the models that were tuned by conventional random cross-validation

Holdout	ncells	RF	GBM	MARS	SVM	GAM	CONS
1	3609	0·93	0·83	0·72	0·85	0·92	0·80
2	5328	0·77	0·73	0·73	0·78	0·76	0·73
3	4516	0·81	0·70	0·68	0·65	0·73	0·68
4	5220	1·16	1·16	1·18	1·34	1·20	1·19
5	7094	1·13	1·09	1·09	1·25	1·09	1·11
6	3236	0·55	0·52	0·53	0·46	0·56	0·49
7	10 642	0·59	0·55	0·53	0·36	0·62	0·50
RMSE (non-random CV)		0·84	0·79	0·78	0·79	0·83	0·78
RMSE (random CV)		0·59	0·64	0·75	0·74	0·76	

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

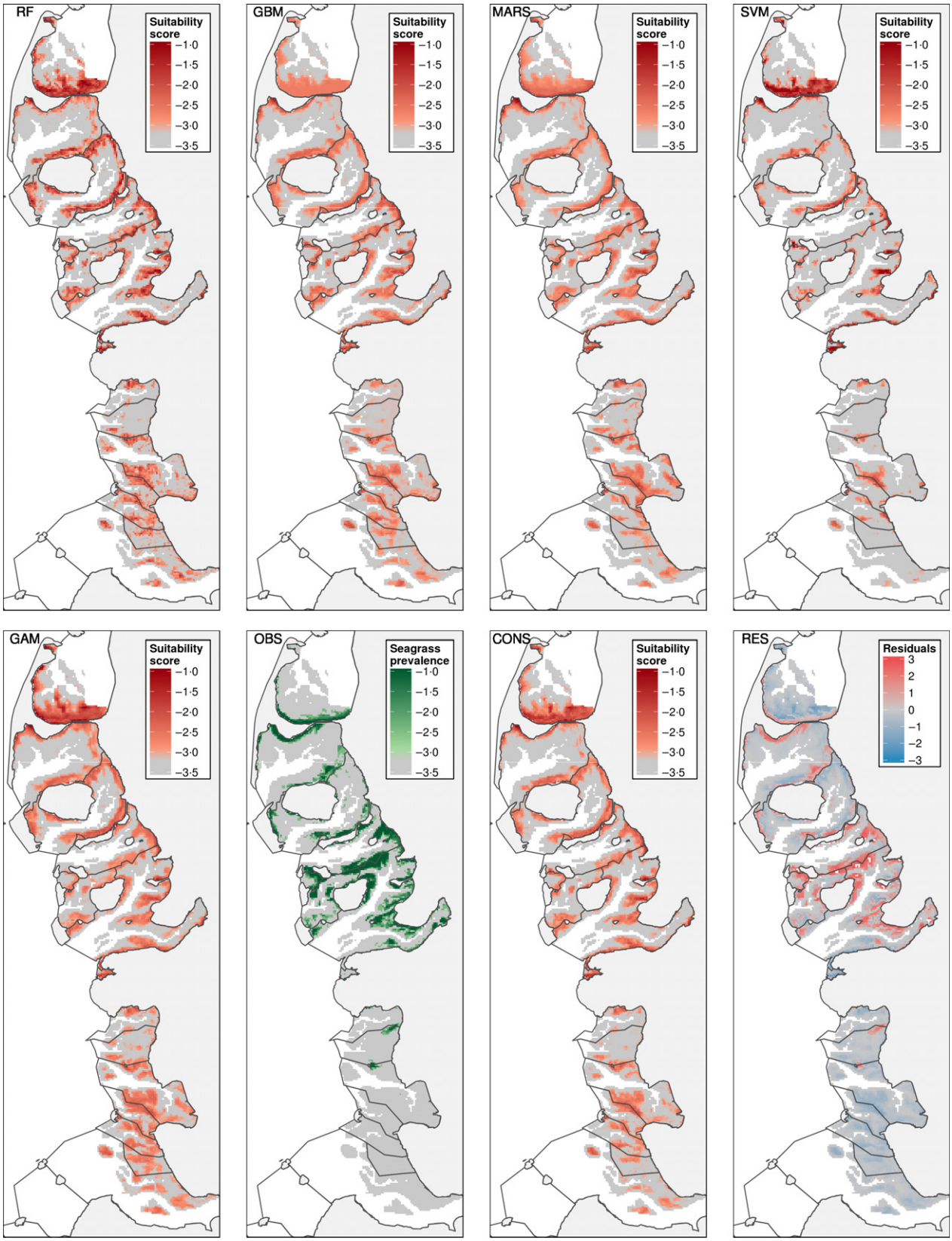

Fig. 4. Habitat suitability scores within the training area of Schleswig–Holstein by RF, GBM, MARS, SVM, GAM. The predictions for each of the spatial subsets are based on a model-tuning procedure where the data of the area for which predictions were made were held out. OBS is the log-transformed prevalence and CONS the prediction by the consensus model. RES are the consensus model residuals (RES = OBS – CONS).

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

HABITAT SUITABILITY AND CLASSIFICATION OF THE ENTIRE WADDEN SEA

The area-adjusted frequency of the presence of seagrass (continuous Boyce index on the basis of the presence-only data outside the training area) increases with HSS (Fig. 5). This means that the probability of observing seagrass clearly increases when predicted habitat suitability increases. Habitat is unsuitable when the continuous Boyce index is smaller than 1 which corresponds to a HSS of −3·16; marginal habitat is defined as habitat with a HSS between −3·16 and −2·42, while suitable habitat has a HSS larger than −2·42 (Fig. 5).

The ensemble spread is small compared to the consensus predictions (Figs 6 and 7) which indicates that the predictions of the 35 models converge. Particularly, the largest ensemble spread is never >6% of the consensus prediction.

The locations of the presence-only seagrass observations correspond well with the predicted suitability and most locations where seagrass has occurred since the 1970s are classified as suitable (Figs 6 and 7). Some exceptions occur the eastern part of Lower Saxony where seagrass presence is observed on marginal and unsuitable flats (Fig. 6). There are also large extents of suitable flats where seagrass was not observed during the period considered. The largest suitable but unoccupied extents occur in the southern Wadden Sea (Fig. 6). The divergence in this area is most striking south of the barrier islands.

The percentage area of suitable tidal flats ranges between 8% in Schleswig–Holstein and 13% in the Netherlands (Table 2). The percentages of unsuitable and marginal habitat are of similar magnitude between the regions.

Discussion

To allow for the recovery of seagrass populations, there is urgent need for identification of suitable habitat. We developed a robust consensus habitat suitability model for intertidal seagrass by combining five HDMs tuned on the basis of transferability performance. We identified areas of high suitability in the Wadden Sea where seagrass meadows may establish and persist. We discuss methodology and possible causes of the absence of seagrass meadows in the southern Wadden Sea.

TRANSFERABILITY ASSESSMENT AND POOLING

Our study shows substantial differences between the prediction accuracy of models tuned by conventional random cross-validation and of models tuned by non-random, spatial cross-validation. The former insufficiently took account of spatial heterogeneity leading to overfitted and poorly transferable models. In particular, Random Forest showed high accuracy under random cross-validation but poor transferability. Note that Wenger & Olden (2012) also found that Random Forest had high in-sample accuracy but suffered from poor transferability. Although our study is limited to analysis of spatial heterogeneity and transferability, there is no reason to believe that it would be any different for temporal transferability. Our results are in line with Araújo *et al.* (2005) and Wenger & Olden (2012).

We furthermore found that none of the five modelling approaches consistently predicted the holdout observations best and that the ensemble approach improved accuracy and robustness of predictions. The ensemble approach also provided measures of spread which were low due to the fact that model predictions converged. From a management point of view, the ensemble spread is not of concern because even the largest spreads are below 6% of the consensus HSS.

HDM DEVELOPMENT WITHIN TRAINING AREA

The consensus predictions were favourable in that all locations within the training area where the prevalence was high were correctly identified as suitable habitat. The observation that most of the suitable habitat in Schleswig–Holstein is occupied is in line with the observation that the seagrass extent in this area has exceeded the levels of the 1930s and that the increase has levelled off over the last years (Dolch, Buschbaum & Reise 2013; T. Dolch, non-published data).

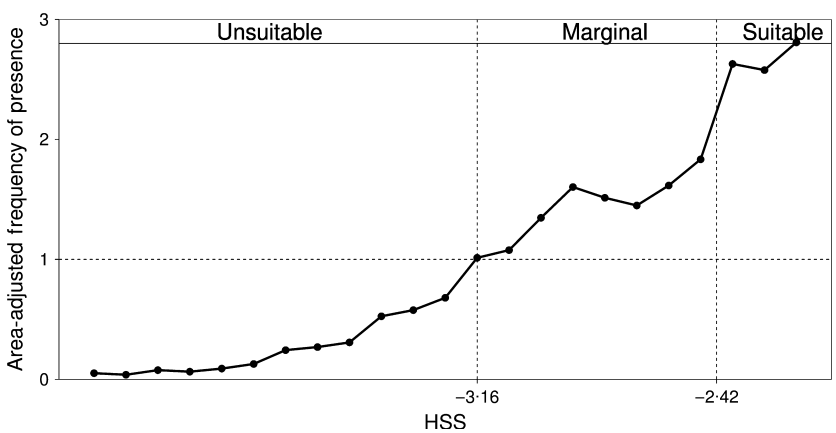

Fig. 5. Area-adjusted frequency of the presence of seagrass 'continuous Boyce index' vs. HSS (habitat suitability score) on the basis of a rolling mean. The vertical dashed lines show the values of the HSSs used to classify unsuitable, marginal and suitable habitat.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

The consensus HDM underpredicted suitability at locations with high prevalence. There are two main reasons for this. First, regression models capture the variables that correlate with the phenomenon under study; they do not capture extremities resulting from randomness. Particularly, randomness underlies 'regression towards the mean' (e.g. Kelly *et al.* 2005). Secondly, static regression models do not fully capture dynamic ecological processes such as reproduction and dispersal which underlie spatial and temporal autocorrelation in the data (Guisan & Thuiller 2005). Particularly, our modelling approach ignores that the development of a seagrass bed is influenced by local production and retention of seed and rhizomes and by possible feedbacks with hydrodynamics (Fonseca & Bell 1998). This means that the impact of the presence of seagrass in 1 year on the presence of seagrass in the following year is not captured. Hence, a relatively frequently occupied location is likely, but not necessarily, more suitable than an infrequently occupied location. For instance, the former might by chance have been recolonized earlier in the monitoring period. Models that are tuned on the basis of transferability are guarded against overfitting in the case of autocorrelation which increases their applicability (Wenger & Olden 2012).

PREDICTIONS FOR THE WADDEN SEA, RECOMMENDATIONS FOR MANAGEMENT AND FUTURE RESEARCH

Because the predictions for the spatial holdout data sets within the training area were accurate, the ensemble of models provided a solid foundation for predicting habitat suitability outside the training area. The relatively low prediction variances yielded further confidence in the reliability of the consensus suitability scores. Finally, the monotonic relationship between the area-adjusted frequency of presence and HSS implies that the predictions are accurate beyond the training area.

We identified extensive areas of suitable habitat in the southern Wadden Sea where sustaining meadows had not been observed in the recent past. It is encouraging that in the last few years, seagrass plants have established and persisted on suitable locations, where they had not been observed before (Appendix S1). Particularly, plants of *Z. noltii* and *Z. marina* were found on tidal flats south of

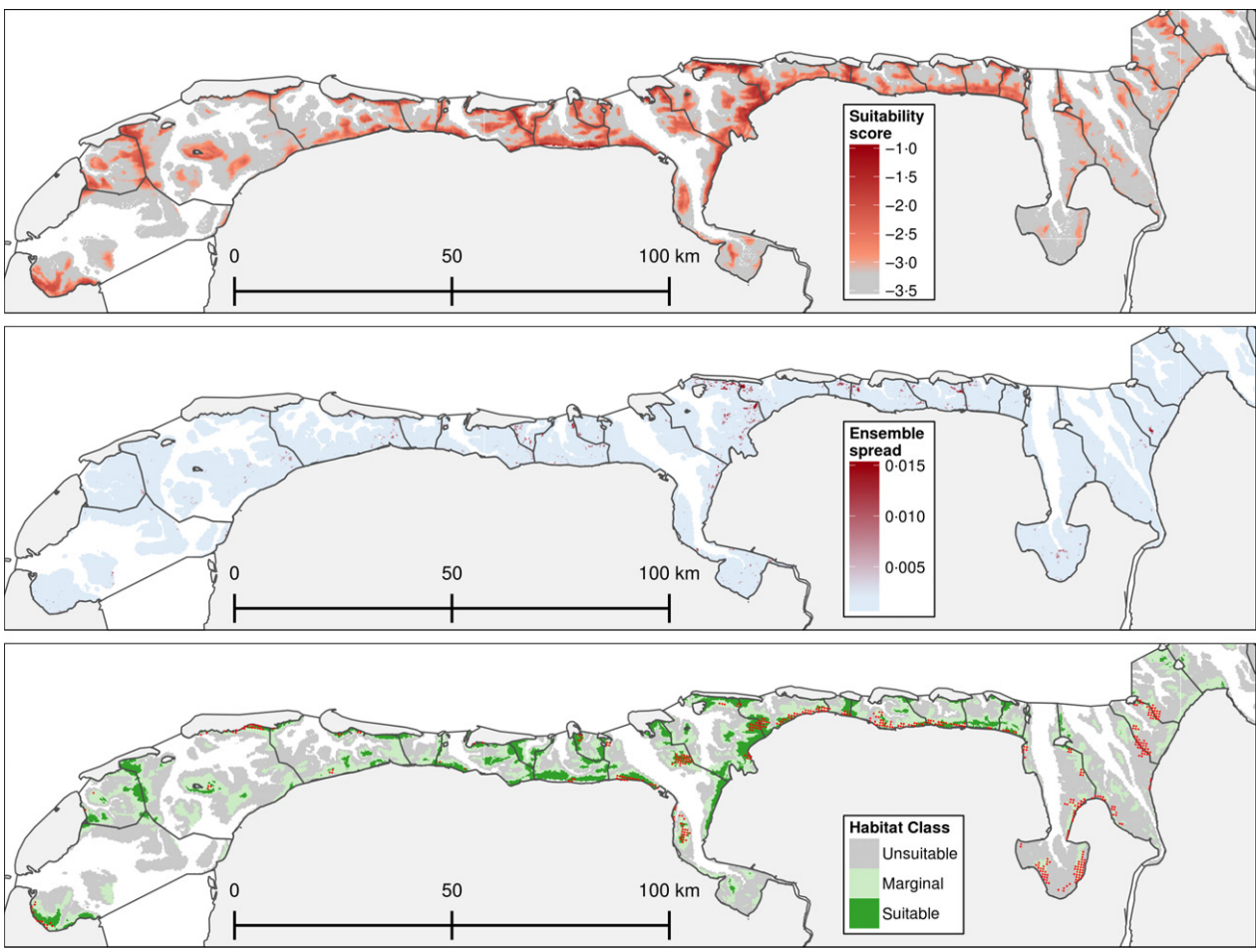

Fig. 6. Consensus predicted suitability scores, ensemble spread, habitat classes and observed intertidal seagrass in the southern Wadden Sea. The maps are rotated 12° clockwise.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

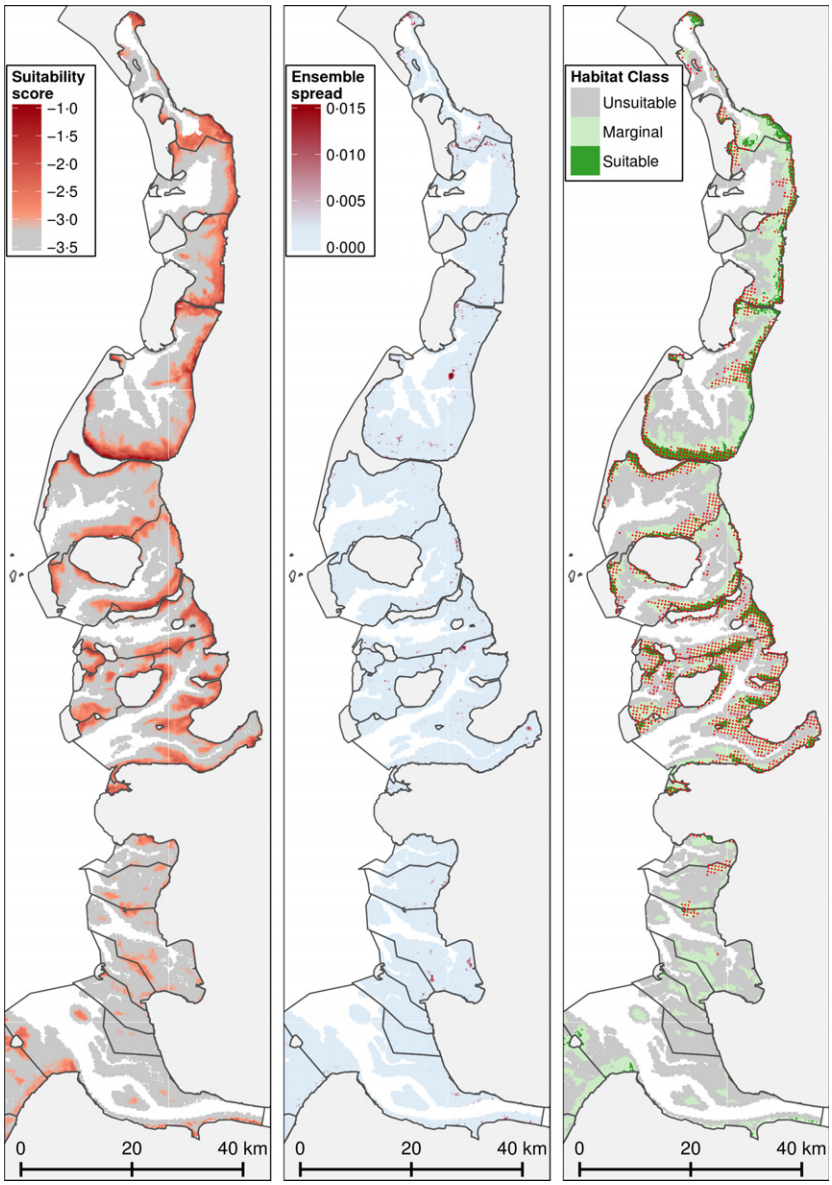

Fig. 7. Consensus predicted suitability scores, ensemble spread, habitat classes and observed intertidal seagrass in the northern Wadden Sea. The maps are rotated 12° clockwise.

Table 2. The area (km^2) of unsuitable, marginal and suitable seagrass habitat and the total area of intertidal flats in the Danish (DK), Schleswig–Holstein (SH), Lower Saxon (LS) and Dutch Wadden Sea (NL). The percentage of the total intertidal flat area is given in parentheses

Region	Unsuitable (%)	Marginal (%)	Suitable (%)	Total
DK	223 (49)	180 (39)	56 (12)	460
SH	999 (63)	460 (29)	126 (8)	1585
LS	976 (58)	537 (32)	179 (11)	1692
NL	897 (56)	501 (31)	211 (13)	1609

Rottumeroog and Rottumerplaat between 2012 and 2015 (Fig. 2, coordinates 4089000, 3385000; B. Ebbinge and M. Bunskoek pers. comm.) and plants of *Z. noltii* were found east of Griend in summer 2015 (coordinates 4007000, 3360000, S. Holthuijsen pers. comm). Given these developments, it is likely that plants and meadows will continue to emerge and expand at these and in other, presently unvegetated, locations. Detailed surveillance monitoring in combination with targeted research and systematic conservation planning may help the protection of seagrass habitat and the development of seagrass populations. For instance, it would be useful to know how natural and anthropogenic conditions at (newly) occupied locations compare to other suitable locations where seagrass meadows are not (yet) developing. For example, cockle fisheries by means of hand dredging is an activity that removes standing seagrass plants and possibly prevent establishment. Although cockle dredging is illegal inside existing seagrass beds, it is currently allowed inside a major part of the suitable habitat in the Dutch Wadden Sea (Appendix S4).

Our HDMs performed well, but it is important to note that they – like any model – incorporate uncertainties. For instance, possible data errors in the bathymetry will

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

cause errors in the modelled hydrodynamics and the sediment data may contain interpolation errors. These data errors affect the HDMs. In this context, it is important to distinguish random error from systematic bias in all observations. The former case would reduce prediction accuracy, while the latter case would not change predictions because correlations remain intact. Nevertheless, hydrodynamic models are currently the only source of consistent and synoptic data at the relevant spatial and temporal scales. GETM was set up for the entire Wadden Sea and simulated the hydrodynamics for a period of 3 years. The model was combined and validated with observations of different variables and performed adequately (Duran-Matute *et al.* 2014). Hence, the summary statistics used in this paper are representative for the true conditions. The sediment data are also reliable because the data set was constructed and validated with an extensive set of observations covering large spatial and temporal scales. Nevertheless, advances in hydrodynamic simulation and monitoring methodology and their integration can help to improve the quality of predictor variables and hence enable more accurate predictions.

Another limitation of our HDMs is that they are based on hydrodynamical and geomorphological predictors only. There are other (dynamic) factors of biotic, chemical and geological nature that we have not accounted for, that may influence settlement, growth and survival of seagrass meadows. For instance, eutrophication may negatively affect seagrass growth via direct negative impact of nutrients (Burkholder, Tomasko & Touchette 2007). Eutrophication may also indirectly affect seagrass by stimulating growth of epiphytes on leafs and growth of green macroalgae which may cover and smother seagrass stands (Reise, Herre & Sturm 1989; Philippart 1995; Burkholder, Tomasko & Touchette 2007). Bioturbation caused by, for example, lugworms *Arenicola marina* may also hamper settlement and growth of seagrass (Philippart 1994). On the other hand, below-surface clay layers from inundated saltmarshes and polders may provide stable underground for seagrass roots enabling seagrass to occur at relatively exposed sites (Reise & Kohlus 2008) while preventing bioturbation by lugworms (Philippart 1994).

A comprehensive framework for modelling seagrass populations requires not only insight into habitat suitability but also understanding and quantification of demographic processes including actual and potential dispersal between source and sink locations. An interesting challenge is to further develop the modelling framework and data collection so that the relationships between habitat and demographic variables can be taken into account. Although integration of slow and fast and of exogenous and endogenous variables is not easy, we expect that this type of refinement not only improves predictions of habitat suitability, but will also allow for more realistic metapopulation models which may help to provide insight into the reasons why seagrass meadows have not established on the extensive areas of seemingly suitable habitat in the southern Wadden Sea. We recommend that the production and dispersal of seagrass propagules are included in monitoring programmes so that source–sink dynamics can be modelled by integrating HDMs, transport models and distribution data.

In conclusion, our consensus HDM identified areas where seagrass meadows may potentially establish and persist in the international Wadden Sea. This information provides managers and ecologists with scientific underpinning to focus on particular areas for conservation, restoration and further research. Habitat distribution modelling is a common practice in conservation science and the number of applications is increasing as a result of the wide availability of GIS data, models and software. Although we think that this is a positive development, our work illustrates that, on a general note, caution should be taken because inaccurate articulation of the purpose of modelling and the failure to account for spatial heterogeneity and spatial autocorrelation may lead to undesired outcomes and biased predictions. Therefore, it is crucial that scientists and managers collaborate to avoid misunderstanding and to develop insights and models that are apt for the problems at hand.

Acknowledgements

This study was carried out within the WaLTER project. The study was financially supported by the Waddenfonds, provinces of Fryslân and Noord Holland and the Program Towards a Rich Wadden Sea. The study to investigate the hydrodynamics of the entire Wadden Sea (project PACE) was funded by the German Federal Ministry of Research and Education (BMBF) (grant FKZ 03F0634A) and the Netherlands Organisation for Scientific Research (NWO) (ZKO project 839·11·003). We thank the following people and organizations for providing data: Karsten Reise (AWI, Wadden Sea Station Sylt), National Park Administration Wadden Sea Lower Saxony and the Dutch Ministry of Transport, Water Management and Public Works. We are grateful to Folkert de Jong for stimulating trilateral research between Denmark, Germany and the Netherlands. We thank Dick de Jong for providing digitized mapping of historical seagrass beds in the Netherlands, Jennifer Valerius for sediment data and Els van der Zee for vector files describing the designation of the Dutch Wadden Sea regarding cockle dredging. Finally, we thank the anonymous reviewer and the non-anonymous reviewer Dr. Seth Wenger for helpful comments on earlier versions of this manuscript.

Data accessibility

The seagrass distribution data, hydrodynamic data and R code used in this study have been deposited in the Dryad Digital Repository doi: 10.5061/dryad.q9c54 (Folmer *et al.* 2016). Sediment data from the AufMod project were provided by the German Federal Maritime and Hydrographic Agency (BSH). They can be requested from Jennifer Valerius Jennifer.Valerius@bsh.de.

References

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42–47.

Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.

Bahn, V. & McGill, B.J. (2013) Testing the predictive performance of distribution models. *Oikos*, **122**, 321–331.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

Barry, S. & Elith, J. (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology*, **43**, 413–423.

Beck, M.W., Heck, K.L., Able, K.W., Childers, D.L., Eggleston, D.B., Gillanders, B.M. *et al.* (2001) The identification, conservation, and management of estuarine and marine nurseries for fish and invertebrates. *BioScience*, **51**, 633–641.

Bos, A.R., Bouma, T.J., de Kort, G.L.J. & van Katwijk, M.M. (2007) Ecosystem engineering by annual intertidal seagrass beds: sediment accretion and modification. *Estuarine, Coastal and Shelf Science*, **74**, 344–348.

Boyce, M.S., Vernier, P.R., Nielsen, S.E. & Schmiegelow, F.K.A. (2002) Evaluating resource selection functions. *Ecological Modelling*, **157**, 281–300.

Burchard, H. & Bolding, K. (2002) GETM - a general estuarine transport model. Scientific documentation. Technical Report EUR 20253 EN.

Burkholder, J.M., Tomasko, D.A. & Touchette, B.W. (2007) Seagrasses and eutrophication. *Journal of Experimental Marine Biology and Ecology*, **350**, 46–72.

Costanza, R., dArge, R., deGroot, R., Farber, S., Grasso, M., Hannon, B. *et al.* (1997) The value of the world's ecosystem services and natural capital. *Nature*, **387**, 253–260.

Den Hartog, C. (1970) *The Sea-grasses of the World*. North Holland Publ. Co., Amsterdam.

Den Hartog, C. (1987) "Wasting disease" and other dynamic phenomena in *Zostera* beds. *Aquatic Botany*, **27**, 3–14.

Den Hartog, C. & Polderman, P.J.G. (1975) Changes in the seagrass populations of the Dutch Waddenzee. *Aquatic Botany*, **1**, 141–147.

Dijkema, K.S., van Thienen, G. & van Beek, J.G. (1989) Habitats of the Netherlands, German and Danish Wadden Sea. Tech. rep., Research Institute for Nature Management, Veth Foundation, Leiden, Texel.

Dolch, T., Buschbaum, C. & Reise, K. (2013) Persisting intertidal seagrass beds in the northern Wadden Sea since the 1930s. *Journal of Sea Research*, **82**, 134–141.

Dormann, C.F., McPherson, J.M., Araujo, M.B., Bivand, R., Bolliger, J., Carl, G. *et al.* (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., Graham, C., Hartig, F. *et al.* (2012) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, **39**, 2119–2131.

Drake, J.M., Randin, C. & Guisan, A. (2006) Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, **43**, 424–432.

Duran-Matute, M., Gerkema, T., de Boer, G.J., Nauw, J.J. & Gräwe, U. (2014) Residual circulation and freshwater transport in the Dutch Wadden Sea: a numerical modelling study. *Ocean Science*, **10**, 611–632.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.

Elith, J., Leathwick, J.R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.

Folmer, E.O., Olff, H. & Piersma, T. (2012) The spatial distribution of flocking foragers: disentangling the effects of food availability, interference and conspecific attraction by means of spatial autoregressive modeling. *Oikos*, **121**, 551–561.

Folmer, E.O., Geest, M.v.d., Jansen, E., Olff, H., Anderson, T.M., Piersma, T. & van Gils, J.A. (2012) Seagrass–sediment feedback: an exploration using a non-recursive structural equation model. *Ecosystems*, **15**, 1380–1393.

Folmer, E.O., Drent, J., Troost, K., Büttger, H., Dankers, N., Jansen, J. *et al.* (2014) Large-scale spatial dynamics of intertidal mussel (*Mytilus edulis* L.) bed coverage in the German and Dutch Wadden Sea. *Ecosystems*, **17**, 550–566.

Folmer, E.O., van Beusekom, J.E.E., Dolch, T., Gräwe, U., van Katwijk, M., Kolbe, K. & Philippart, C.J.M. (2016) Data from: Consensus forecasting of intertidal seagrass habitat in the Wadden Sea. Dryad Digital Repository, http://dx.doi.org/10.5061/dryad.q9c54.

Fonseca, M.S. & Bell, S.S. (1998) Influence of physical setting on seagrass landscapes near Beaufort, North Carolina, USA. *Marine Ecology-Progress Series*, **171**, 109–121.

Gräwe, U., Holtermann, P., Klingbeil, K. & Burchard, H. (2015) Advantages of vertically adaptive coordinates in numerical models of stratified shelf seas. *Ocean Modelling*, **92**, 56–68.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I. *et al.* (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424–1435.

Hastie, T., Tibshirani, R. & Friedman, J.H. (2009) *The elements of statistical learning. Data mining, inference, and prediction*. Springer, New York.

van der Heide, T., van Nes, E., Geerling, G., Smolders, A., Bouma, T.J. & van Katwijk, M. (2007) Positive feedbacks in seagrass ecosystems: implications for success in conservation and restoration. *Ecosystems*, **10**, 1311–1322.

Hijmans, R.J. (2013) raster: Geographic data analysis and modeling.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, **199**, 142–152.

de Jong, D.J., Katwijk, M. & Brinkman, A. (2005) Kansenkaart Zeegras Waddenzee. Tech. rep., RIKZ, Middelburg.

van Katwijk, M.M., Hermus, D.C., de Jong, D.J., Asmus, R.M. & de Jonge, V.N. (2000) Habitat suitability of the Wadden Sea for restoration of *Zostera marina* beds. *Helgoland Marine Research*, **54**, 117–128.

van Katwijk, M.M., Bos, A.R., Kennis, P. & de Vries, R. (2010) Vulnerability to eutrophication of a semi-annual life history: a lesson learnt from an extinct eelgrass (*Zostera marina*) population. *Biological Conservation*, **143**, 248–254.

Kelly, C., Price, T.D., West, A.E.S.A. & Whitlock, E.M.C. (2005) Correcting for regression to the mean in behavior and ecology. *The American Naturalist*, **166**, 700–707.

Koch, E. (2001) Beyond light: physical, geological, and geochemical parameters as possible submersed aquatic vegetation habitat requirements. *Estuaries*, **24**, 1–17.

Kuhn, M. (2014) caret: Classification and regression training. R package version 6.0-37.

Kuhn, M. & Johnson, K. (2013) *Applied predictive modeling*. Springer, New York, NY.

Leathwick, J.R., Elith, J. & Hastie, T. (2006) Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological Modelling*, **199**, 188–196.

Leuschner, C., Landwehr, S. & Mehlig, U. (1998) Limitation of carbon assimilation of intertidal *Zostera noltii* and *Z. marina* by desiccation at low tide. *Aquatic Botany*, **62**, 171–176.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, **15**, 59–69.

McCue, A.J., McGrath, M.J. & Wiersma, Y.F. (2014) Benefits and drawbacks of two modelling approaches for a generalist carnivore: can models predict where Wile E. Coyote will turn up next? *International Journal of Geographical Information Science*, **28**, 1590–1609.

Nejrup, L.B. & Pedersen, M.F. (2008) Effects of salinity and water temperature on the ecological performance of *Zostera marina*. *Aquatic Botany*, **88**, 239–246.

Orth, R.J., Carruthers, T.J.B., Dennison, W.C., Duarte, C.M., Fourqurean, J.W., Heck, J. *et al.* (2006) A global crisis for seagrass ecosystems. *BioScience*, **56**, 987–996.

Philippart, C. (1994) Interactions between *Arenicola marina* and *Zostera noltii* on a tidal flat in the Wadden Sea. *Marine Ecology Progress Series*, **111**, 251–257.

Philippart, C.J.M. (1995) Seasonal variation in growth and biomass of an intertidal *Zostera noltii* stand in the Dutch wadden sea. *Netherlands Journal of Sea Research*, **33**, 205–218.

Philippart, C.J.M. & Dijkema, K.S. (1995) Wax and wane of *Zostera noltii* Hornem. in the Dutch Wadden Sea. *Aquatic Botany*, **49**, 255–268.

Prasad, A.M., Iverson, L.R. & Liaw, A. (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, **9**, 181–199.

Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*, **33**, 1689–1703.

Reise, K., Herre, E. & Sturm, M. (1989) Historical changes in the benthos of the Wadden Sea around the island of Sylt in the North Sea. *Helgoländer Meeresuntersuchungen*, **43**, 417–433.

Reise, K. & Kohlus, J. (2008) Seagrass recovery in the northern wadden sea? *Helgoland Marine Research*, **62**, 77–84.

Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,

Shafer, D.J., Sherman, T.D. & Wyllie-Echeverria, S. (2007) Do desiccation tolerances control the vertical distribution of intertidal seagrasses? *Aquatic Botany*, **87**, 161–166.

Suykerbuyk, W., Govers, L.L., Bouma, T.J., Giesen, W.B., de Jong, D.J., van de Voort, R., Giesen, K., Giesen, P.T. & van Katwijk, M.M. (2016) Unpredictability in seagrass restoration: analysing the role of positive feedback and environmental stress on *Zostera noltii* transplants. *Journal of Applied Ecology*, doi:10.1111/1365-2664.12614.

Thuiller, W. (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.

Umlauf, L. & Burchard, H. (2003) A generic length-scale equation for geophysical turbulence models. *Journal of Marine Research*, **61**, 235–265.

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. *Journal of Sea Research*, **83**, 56–64.

Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S*. Springer, New York.

Waycott, M., Duarte, C.M., Carruthers, T.J.B., Orth, R.J., Dennison, W.C., Olyarnik, S. *et al*. (2009) Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences*, **106**, 12377–12381.

Wenger, S.J. & Olden, J.D. (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, **3**, 260–267.

Wenger, S.J., Som, N.A., Dauwalter, D.C., Isaak, D.J., Neville, H.M., Luce, C.H. *et al*. (2013) Probabilistic accounting of uncertainty in forecasts of species distributions under climate change. *Global Change Biology*, **19**, 3343–3354.

Wiersma, A., Oost, A.P., van der Berg, M., Vos, P., Marges, V. & de Vries, S. (2009) Geomorphology. thematic report no. 9.(eds H. Marencic & de Vlas J.). Thematic Report 9, Common Wadden Sea Secretariat, Trilateral Monitoring and Assessment Group, Wilhelmshaven, Germany.

Received 24 November 2015; accepted 15 April 2016
Handling Editor: Yolanda Wiersma

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Inventory and trends of littoral seagrass in the Wadden Sea.

Appendix S2. Machine learning approaches.

Appendix S3. Tuning parameters.

Appendix S4. Cockle dredging in the Dutch Wadden Sea.

© 2016 The Authors. *Journal of Applied Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society,