AQUATIC CONSERVATION: MARINE AND FRESHWATER ECOSYSTEMS
Aquatic Conserv: Mar. Freshw. Ecosyst. **23**: 710–722 (2013)
DOI: 10.1002/aqc.2364

Applying distribution model projections for an uncertain future: the case of the Pacific oyster in UK waters

MIRANDA C. JONES[a,b,c,*], STEPHEN R. DYE[a,b], JOHN K. PINNEGAR[b], RACHEL WARREN[a,c]
and WILLIAM W.L. CHEUNG[d]
[a]*School of Environmental Sciences, University of East Anglia, Norwich, UK*
[b]*Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowestoft, UK*
[c]*Tyndall Centre for Climate Change Research, Norwich, UK*
[d]*Changing Ocean Research Unit, Fisheries Centre, University of British Columbia, Vancouver, BC, Canada*

ABSTRACT

1. The inherent complexity of the environment is such that attempts to model it must operate under simplifications and assumptions. Considering predictions from alternative models, with a range of assumptions and data requirements, therefore provides a more robust approach.

2. The intractability and uncertainty resulting from a suite of predictions may hinder the application of science in policy, where a single prediction with little ambiguity or uncertainty would be most desirable. Few studies modelling species' distributions attempt to present multi-model outputs in a format most useful to the non-modelling community, and none of these have done so for the marine environment.

3. The problem of uncertainty is particularly prevalent in predicting the distribution of invasive alien species under climate change. As invasive alien species are one of the main drivers of biodiversity loss and may incur significant economic costs, the benefit of applying predictions to highlight areas of possible establishment and inform policy and management may be large.

4. An ensemble prediction is used to assess the distribution of suitable habitat for the Pacific oyster, *Crassostrea gigas*, in UK waters both currently and in the future. The ensemble incorporates predictions from three species distribution models, using data from two global climate models. A method is developed highlighting the agreement of the ensemble, further applying threshold values to retain information from constituent predictions in the final map of agreement.

5. Ensemble predictions made here suggest that Pacific oyster will experience an opening of suitable habitat in northern UK waters, reaching the Faroe Islands and the eastern Norwegian Sea by 2050. Habitat suitability will increase with warming temperatures in the English Channel and Central North Sea for this species. The approaches applied here can be incorporated into risk assessment frameworks for invasive species, as stipulated in the Convention on Biological Diversity.

Received 23 August 2012; Revised 19 February 2013; Accepted 16 March 2013

KEY WORDS: ocean; distribution; habitat mapping; invertebrates; alien species; climate change

---

*Correspondence to: M.C. Jones, School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK.
Email: miranda.c.jones@gmail.com

INTRODUCTION

Science is often relied upon to help inform environmental policy-making and to provide answers in the face of political controversies. From the decision-maker's perspective, the most desirable form of advice would constitute a single prediction or projection with little ambiguity or uncertainty. Currently there is a tendency to expect that the introduction of better and more complete data will necessarily facilitate better and more effective policy-making. This is exemplified by calls for more research following politically sensitive or environmentally controversial decisions. However, while the need for environmental policy to be founded on sound evidence is not dismissed, it is stressed that there is sometimes a limit to what science can realistically offer and it is often unwise to delay decisions while awaiting better, or more accurate data. The complexity of the natural world is such that modelling must, by necessity, be reductionist. Unambiguous answers are rarely achieved and for pragmatic reasons modellers must make assumptions and simplifications, with even the most complex models being approximations of a real system (Collins *et al.*, 2012). Different modellers may also favour particular modelling frameworks/formulations with their inherent biases and peculiarities. Furthermore, rarely are a suite of equally plausible models tested and compared.

The wealth of methodologies for dealing with ecological complexity or uncertainty may not only lead to confusion, but also to criticism and scepticism among the non-modelling community. For example, model comparison studies have demonstrated the variation in outputs achievable using different Species Distribution Models (SDMs) (Araújo *et al.*, 2005; Pearson *et al.*, 2006; Araújo and New, 2007) and with inputs from alternative global climate projections (Jones *et al.*, 2012). A multi-model procedure (an 'ensemble' of models) is advocated in Jones *et al.* (2012) rather than assuming that any one model gives a 'true' picture of the ecosystem. For biological or ecosystem projections, the best possible policy outcome will often be achieved not by limiting outlook to a perceived 'best' model. Instead, a range of available projections from a variety of validated methodologies and sources should be taken into account.

If variability in predictions is great enough to cause confusion or misunderstanding, the utility of a multi-model approach in guiding policy will be limited and may even hinder the decision-making process. In such instances, the use of techniques to summarize the discrepancies and concordances within an 'ensemble' framework would seem desirable for presenting the maximum amount of information in a single figure. The statistical 'ensemble' approach was pioneered by J Willard Gibbs in 1878 as an idealization consisting of a large number of copies of a system, considered all at once, each of which represents a possible state that the real system might be in at some specified time. Ensemble methods aim to obtain better predictive performance than could be obtained from any of the constituent models runs, and the approach has become particularly prevalent among physical and meteorological modellers that participate in the Intergovernmental Panel on Climate Change (IPCC, 2007). It has been argued that significant improvements in the robustness of a forecast can be achieved if an ensemble approach is used and the results are analysed appropriately (Araújo and New, 2007). A non-ecological study has further shown that as long as individual forecasts contain some independent information, combined forecasts would yield lower mean error than any of the constituent forecasts (Bates and Granger, 2012). There are several methodologies for constructing ensemble forecasts (Araújo *et al.*, 2006; Araújo and New, 2007; Coetzee *et al.*, 2009; Diniz-Filho *et al.*, 2009; Marmion *et al.*, 2009) and these have been widely explored in species distribution modelling (Araújo *et al.*, 2006; Pearson *et al.*, 2006; Prasad *et al.*, 2006; Carvalho *et al.*, 2011). Several of these methods attempt to find the central tendency of forecasts through measures such as the mean or median (Araújo *et al.*, 2005; Marmion *et al.*, 2009). They attempt to distinguish the 'signal' of the combined predictions from the 'noise' surrounding it that might be associated with individual model error and uncertainty. Alternatively, an ensemble of predictions can be used to define the bounding box (Araújo and New, 2007). This method identifies the range in forecasts from the ensemble members and the maximum area of predicted distribution, without quantifying the probability distribution or conditional probabilities. No ensemble average or confidence limits around the average are calculated.

The application and utility of a non-statistical ensemble approach is demonstrated by considering the example of the Pacific oyster, *Crassostrea gigas*. The Pacific oyster is native to Japan and east Asia but has been introduced intentionally to countries such as Australia, France, the United States and the United Kingdom for aquaculture, and is now the most widely farmed and commercially important oyster in the world. Within the UK, 1400 tonnes of Pacific oyster were grown in aquaculture facilities in 2006 (GB Non-Native Species Secretariat (NNSS), 2012). Pacific oyster was deliberately introduced for commercial purposes after Ministry of Agriculture, Fisheries and Food (MAFF) trials in the 1960s and early 1970s had indicated that the species required temperatures clearly in excess of those in British waters for successful reproduction in the wild (Mann, 1979; Utting and Spencer, 1992). Accordingly, hatchery-produced seed were produced in large numbers for commercial rearing. However, natural recruitment has since occurred within the British Isles, although with regional differences. Some local recruitment occurred in estuaries of south-west England and north Wales following unusually warm summers in 1989 and 1990 (Spencer *et al.*, 1994) and there are now well established 'wild' adult populations. To date, there have been no substantiated records of spatfall in Scottish waters although maturation of the gonad and gamete release have been noted occasionally during atypically warm weather and in shallow, sheltered sea lochs (Maggs *et al.*, 2010). However, in 2005, temperatures in Strangford Lough, Northern Ireland, reached those sufficient to allow spat development and settlement, and were reflected by high frequencies of age classes recruited in 2005 (Guy and Roberts, 2010).

Despite the economic benefits, wild establishment of Pacific oyster may cause significant economic losses by outcompeting native species of bivalves, especially native oysters *Ostrea edulis*, as well as mussels and cockles. In the Wadden Sea, for example, Pacific oyster has been seen to displace mussel fisheries in some areas (Nehls *et al.*, 2006) and it has been suggested that their reefs may cause major shifts in the community of benthic filter feeders, with subsequent negative effects on bird populations (Smaal *et al.*, 2005). Furthermore, the presence of their sharp shells on the intertidal zone and mudflats may deter human leisure activities, thus negatively affecting tourism.

There are several terms used for describing species persisting outside their native range. Non-native refers to all species that have been deliberately or accidentally introduced to an area from their native range (Kolar and Lodge, 2001). While non-native species that establish themselves in a new range but do not cause negative impacts are referred to as naturalized or non-invasive (Kolar and Lodge, 2001), the term 'invasive alien species' is used to describe those that cause, or have the potential to cause, harm to the environment, economies, or human health (Global Invasive Species Programme, 1999). Invasive alien species are identified as one of the main drivers of biodiversity loss and ecosystem malfunction (Mcneely, 2001; Underwood *et al.*, 2003; Molnar *et al.*, 2008). With good evidence that climate change favours the spread of some non-native species (Stachowicz *et al.*, 2002; Sax *et al.*, 2007), the interaction between invasive alien species and climate change is thus becoming a pressing issue for conservation and fisheries economics. Although the extent to which the non-native Pacific oyster may affect UK native species and habitats is poorly understood, the increasingly favourable conditions caused by warming seawater temperatures are likely to benefit this species and promote its further spread. The UK is bound by international agreements such as the Convention on Biological Diversity, the United Nations Convention on the Law of the Sea, The Convention on the Conservation of Migratory Species of Wild Animals (Bonn, 1979), The Convention on the Conservation of European Wildlife and Natural Habitat (Bern, 1979), the EC Habitats and Species Directives, as well as the EU Water Framework Directive and Marine Strategy Framework Directive. All of these aim to protect biodiversity and most include provisions aimed at preventing the further introduction of, or control of, non-native species, especially those that pose a risk to native or protected species (JNCC, 2012).

Furthermore, the total cost of invasive alien and non-native species to the UK economy (both terrestrial and aquatic) is estimated at £1.7 billion per annum (Williams *et al*., 2010). Aquatic molluscs alone are thought to cause damages of $1 billion per year in the USA (Pimentel, 2005). Therefore early warning systems to highlight potentially invadable areas by species could be a useful first step in any proposed management, monitoring or prevention plan. This would seem especially useful in the case of Pacific oyster, for which no management plan exists and thus the most desirable route to preventing spread would be enhanced bio-security and vigilance against further deliberate and accidental introductions. The first UK Climate Change Risk Assessment (CCRA) published in January 2012 (a requirement under the UK Climate Change Act 2008) specifically argued that a more statistically rigorous and defensible study was needed of projection techniques for non-native aquatic species, as this report could only manage a very crude attempt at predicting future distribution. Species Distribution Models (SDMs) have been used to predict suitable habitat of other non-native marine species in Europe (e.g. Chinese mitten crab, *Eriocheir sinensis* (Herborg, 2007)) and have further been used to inform management efforts aimed at identifying areas at risk of zebra mussel (*Dreissena polymorpha*) invasion in the western USA (Drake and Bossenbroek, 2004). The application of an ensemble model approach and its utility in assessing the potential threat by an invasive species is explored as well as attempting to provide predictions of projected habitat suitability for the Pacific oyster in a format useable by policy-makers and the non-modelling community.

METHODS

Predictions of relative habitat suitability in UK for both current and future time periods were generated for the Pacific oyster using three different SDMs and outputs from two distinct Global Climate Models (GCMs). The SDMs were AquaMaps (Kaschner *et al*., 2006; Ready *et al*., 2010), Maxent (Phillips *et al*., 2006), and the Dynamic Bioclimate Envelope Model (DBEM) (Cheung *et al*., 2011). These models (described in detail in the above references and compared in Jones *et al*. (2012)) have been shown to produce plausible predictions of species' current distributions given occurrence data, which are used in model testing (Jones *et al*., 2012). Furthermore, comparisons between model hindcast and historical distribution changes of fishes and invertebrates from the 1970s to the 2000s in the Bering Sea and Northeast Atlantic suggest that DBEM has significant skill in predicting distribution shifts in these regions (Cheung *et al*., 2012).

AquaMaps and Maxent are statistical models, which differ in complexity but which both generate predictions of a species' relative habitat suitability by associating presence-only data on a species' occurrence with a set of environmental variables. Species' occurrence data were obtained from two global online databases: the Ocean Biogeographic Information System (OBIS) (http://www.iobis.org) and the Global Biodiversity Information Facility (GBIF) (http://data.gbif.org), all last accessed in 2011. Occurrence records were rigorously filtered to minimize the recording error likely in using data compiled from many sources. Thus, additional information on species' environmental preferences and geographic limits (Fishbase, Froese and Pauly, 2011; FAO (2011) fact sheets: http://www.fao.org/fishery/species/3514/en; The Marine Life Information Network (MarLIN) www.marlin.ac.uk) was used to remove occurrence points located on land or outside expert-defined geographic ranges (obtained as latitudinal and longitudinal limits from FishBase) or FAO areas. Occurrence data were spatially aggregated at the level of 0.5° latitude x 0.5° longitude, giving a binary value of presence or absence for each cell.

Environmental datasets were obtained at 0.5° latitude x 0.5° longitude resolution from the Geophysical Fluid Dynamics Laboratory's Earth System Model (GFDL ESM2.1 (Dunne *et al*., 2010)). These included sea surface temperature, sea bottom temperature, salinity, primary productivity and depth. A further set of physical climate data were obtained from an ensemble of 12 different CMIP3 models that were assembled under the auspices of the fourth assessment of the Intergovernmental Panel on Climate Change (IPCC AR4). These were obtained from the World Climate Research Program (WCRP) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset (http://esg.llnl.gov:8080). Both climatic datasets were modelled under the 'high' emission SRES A2 scenario and are thus characterized by a heterogenous world with a continuously increasing global population and regionally orientated economic development (IPCC, 2000). Environmental envelopes generated by each model were used to predict the distribution of Pacific oyster, using a 30 year average of environmental data centred on 1985 (1970–2000), representing the current time period. This was then applied to a 30 year average centred on 2050 (2035–2065), representing the future, under climate change.

The DBEM, and associated *Sea Around Us Project* model contrast the above approaches by first defining a species' current distribution based on the following geographic and environmental tolerance limits: FAO area, latitudinal and depth ranges and habitat preferences (Close *et al.*, 2006; Jones *et al.*, 2012). This distribution is then used to define the species' bioclimatic envelope by its 'preference profile', formed by overlaying current 1970–2000 averaged environmental data over the maps of current relative suitability. In projecting the bioclimatic envelope under scenarios of climate change, the DBEM incorporates dispersal and the effects of oxygen level and acidification on species growth through incorporation of ecophysiological model components and a logistic population growth model (Cheung *et al.*, 2011). Predictions from each model and time period were standardized to give values lying between 0 and 1 and representing the relative habitat suitability of each cell of the study area for Pacific oyster.

The centroid of the distribution predicted for each time period (latitudinal centroids) was calculated for predictions made using each SDM-GCM combination and threshold using the following equation (Cheung *et al.*, 2009):

$$C = \frac{\sum_{i=1}^{n} Lat_i \cdot Abd_i}{\sum_{i=1}^{n} Abd_i}$$

where Lat_i is the latitude of the centre of the spatial cell (i), Abd is the predicted relative habitat suitability in the cell, and n is the total number of cells (Cheung *et al.*, 2012). The difference between latitudinal centroids in projected and reference years was then calculated in kilometres (km) (Cheung *et al.*, 2011):

$$Distance\ \ shift\,(km) = (Lat_m - Lat_n)\frac{\pi}{180} \times 6378.2$$

where Lat_m and Lat_n are the latitudinal centroids in 2050 and 1985, respectively, and 6478.2 km is the approximated equatorial radius of the Earth.

Further to predicting shifts in latitudinal centroid of a distribution, potential changes in range area between reference and projected years was calculated as the difference in the number of cells with habitat suitability > 0 (number in 2050 – number in 1985).

Model predictions were combined using a 'bounding box' method, retaining the information from each prediction using an index of model agreement. When applying species distribution models to an environmental problem there may be an element of perceived risk and cost. For example, the cost of acting on a forecast that gave a restricted estimate of invasion potential might be high if a non-native species has particularly damaging environmental effects. This idea was introduced into the ensemble forecasting process by applying thresholds. Thresholds may be used to transform the continuous predictions of relative suitability produced by SDMs into predictions of presence/absence. There are several methods for selecting threshold values, although there is currently no consensus on the most suitable method for applying thresholds to species' range projections (Liu *et al.*, 2005; Nenzén and Araújo, 2011). For example, if a low threshold is set, low values of habitat suitability will be converted into potential 'presence' areas. When investigating the spread of the Pacific oyster, such a technique could be beneficial in making a precautionary prediction of potentially invadable areas and assessing the species that might suffer the negative impact of invasion. However, if the main focus is to implement a strategy of prevention or mitigation, areas of incorrect prediction might incur considerable and unnecessary costs that might be better deployed elsewhere. Here, each of the six predictions in an ensemble was converted from a probability distribution to a binary prediction of presence or absence using one of three threshold cut-off values, providing a range of invasion outcomes.

The first threshold chosen was one that maximizes the accuracy of the model in predicting the observed occurrences/absences (maximum training sensitivity plus specificity (MaxSS)), as indicated in model testing. To find this value, occurrence data were split in two, with 75% being used to train and 25% to test the model. Model testing was implemented using the ROCR package in R (Sing *et al.*, 2007), a package designed for evaluating and visualizing classifier performance using R. Two fixed thresholds, of 0.5 and 0.7 were also chosen. These would produce more constrained predictions by retaining only cells with predicted relative habitat suitability values higher than the 50th and 70th percentile, respectively, of the relative habitat suitability distribution. Predicted habitat that is lower than the specified threshold value was thus allocated a value of 0 while that higher than the threshold value was allocated a value of 1. Having obtained a set of occurrence predictions across all model combinations for each threshold, projections within each set were summed to produce an index of agreement ranging between 0 and 6. This would result in a map that not only displays the maximum agreed areas of prediction but also the extent of agreement across 0.5° latitude × 0.5° longitude cells in the study area. Similarly, thresholds were used to retain information on the magnitude of change in presenting ensemble maps of the difference in relative habitat suitability (ΔRHS) values between projection and reference time periods (2050–1985, values ranging between -1 and 1). Having calculated ΔRHS for each SDM–GCM combination, cut offs of ±0.1, ±0.2 and ±0.4 were applied to create binary predictions as described above. These cut-off values reflected the range of ΔRHSs found while also allowing information on the degree of change to be portrayed in the map of agreement. These were summed across all SDM–GCM combinations to produce an index of agreement potentially ranging from -6 (maximum agreement of a decrease in habitat suitability across models) to 6 (maximum agreement of an increase in habitat suitability). As the index aimed to portray agreement of change, cells predicted to show no change remained at 0.

RESULTS

Sea surface temperature in UK waters is predicted to rise by an average of 0.49°C (5%) using GFDL ESM 2.1 data and by 0.99°C (10%) using CMIP3 data from 1985 to 2050. All model combinations predict a northward range shift for Pacific oyster across its range between 1985 and 2050 (median = 467 km). The threshold applied to predictions makes little difference to the median prediction of range shift across SDM and climate datasets (Figure 1(a)), although variation in predictions is seen across threshold within SDM and climate datasets, in particular using AquaMaps (Table 1). The median latitudinal shift is also predicted to vary across SDMs, with Maxent predicting the most conservative median value and the smallest range (Figure 1(b)). However, the extent of difference between predictions of latitudinal shift using alternative SDMs is dependent on both the climate dataset and threshold used (Table 1). Thus, while results within SDM are relatively consistent using CMIP3-E data at different threshold levels, the greatest differences in latitudinal shift are caused by SDMs. Using GFDL data, the difference between SDM predictions increase with more restrictive (higher) threshold values.

Within UK waters (represented by the UK Exclusive Economic Zone (EEZ)) only AquaMaps and the DBEM, applied to GFDL climate data predicted an increase in the number of 0.5° latitude × 0.5° longitude cells containing suitable habitat (the range area) between 1985 and 2050, and thus range expansion across thresholds (Table 2). Using CMIP3-E data, a slight decrease in range area was predicted in some cases, while predictions from Maxent do not show change in range area within the UK EEZ.

An example of the change in habitat suitability that contributes to the predicted northward shifts in latitudinal centroid is shown for all SDMs with GFDL data and using the MaxSS threshold in Figure 2. Here, in particular using AquaMaps and Maxent, an increase in habitat suitability in seen in the central North Sea, and around the northern coast of Norway and Scotland, while a slight decrease in habitat suitability is seen south of the UK and towards the Mediterranean.

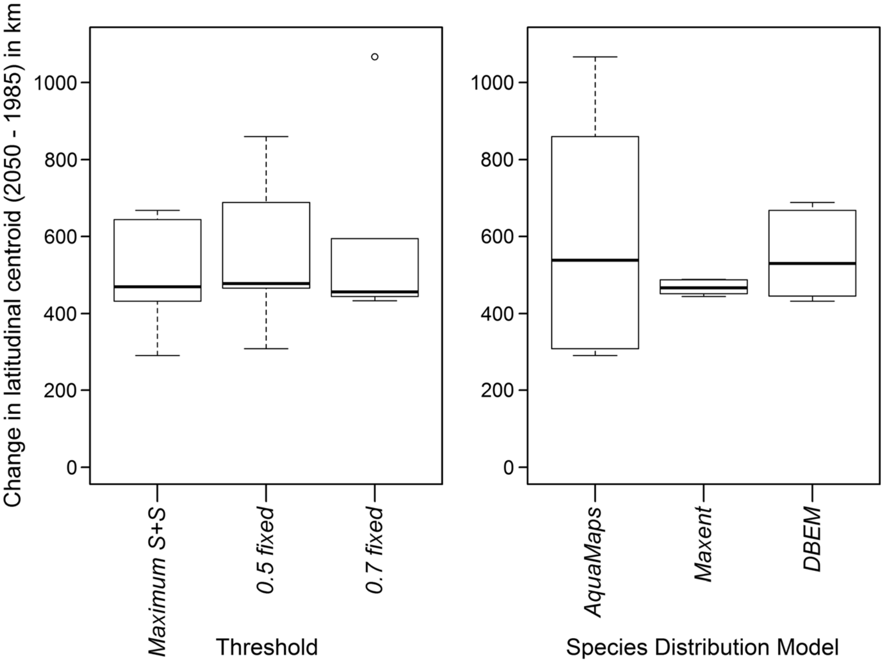

Figure 1. Shift in latitudinal centroids for Pacific oyster. Change in latitudinal centroid (2050–1985) in km (a) across different thresholds for all SDM–GCM combinations, (b) across different SDM models for both GCM datasets and all thresholds.

Table 1. Predicted latitudinal shifts in km using AquaMaps, Maxent and the DBEM with each threshold (maximum sensitivity + specificity (MaxSS), 0.5 and 0.7 fixed thresholds) and climate dataset (GFDL and CMIP3-E)

Global Climate Model and threshold applied	AquaMaps	Maxent	DBEM	Maximum Difference
GFDL, MaxSS	644.09	451.78	431.94	212.15
GFDL, 0.5	859.95	466.99	465.98	393.97
GFDL, 0.7	1067.07	444.39	445.71	622.68
Maximum difference (GFDL)	422.98	22.60	34.05	
CMIP3-E, MaxSS	290.50	487.55	667.81	487.23
CMIP3-E, 0.5	308.26	488.78	688.52	488.20
CMIP3-E, 0.7	433.42	466.31	594.11	466.44
Maximum difference (CMIP3-E)	142.92	22.47	94.40	

Table 2. Potential change in area, calculated as the number of 0.5° latitude x 0.5° longitude cells predicted as suitable in 1985 and 2050 in the UK EEZ using each threshold: Maximum Sensitivity + Specificity (MaxSS), 0.5 and 0.7 fixed thresholds

	MaxSS threshold		0.5 threshold		0.7 threshold	
	1985	2050	1985	2050	1985	2050
AquaMaps, GFDL	277	307	246	298	210	235
AquaMaps, CMIP3-E	307	307	303	298	289	280
Maxent, GFDL	320	320	320	320	320	320
Maxent, CMIP3-E	320	320	320	320	320	320
DBEM, GFDL	175	221	122	136	93	109
DBEM, CMIP3-E	230	230	201	186	148	150

Maps of agreement of Pacific oyster presence among the six combinations of SDM–GCMs in 2050 are shown in Figure 3, panels 1a, 2a, 3a. Here royal blue (0) denotes areas with no prediction of occurrence using any model, rather than no agreement between models. These projections also show the difference in maximum agreement when applying different thresholds. Thus, in applying a

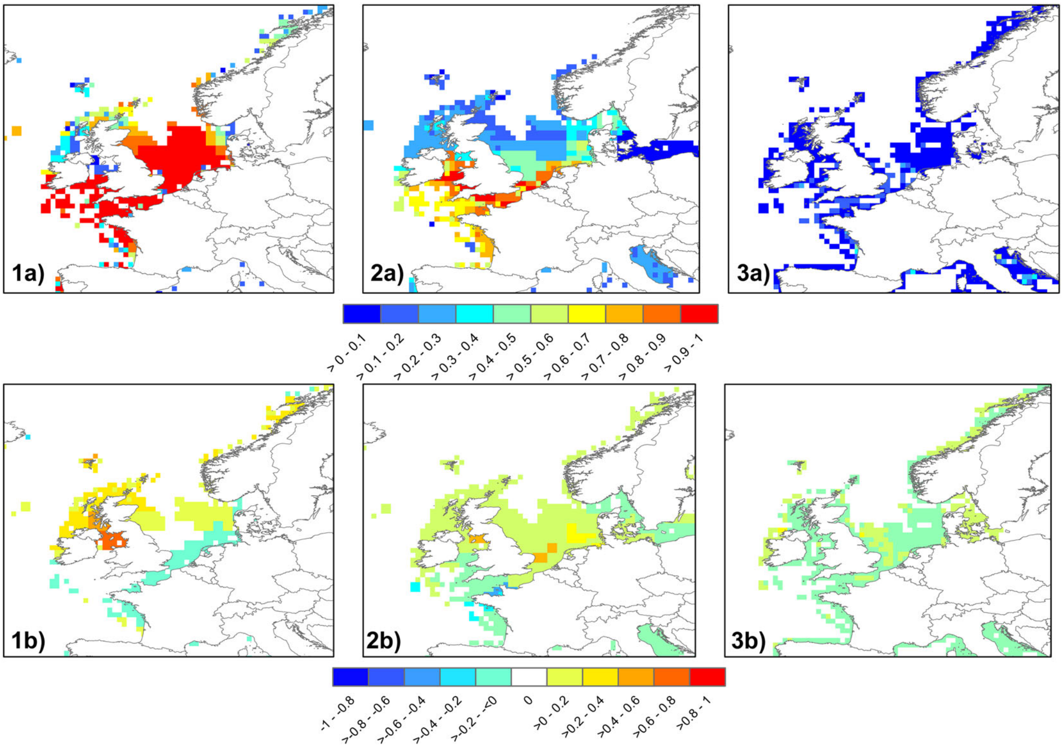

Figure 2. Predicted distributions of relative habitat suitability for Pacific oyster. Predictions made for the Pacific oyster using GFDL Topaz ESM2.1 data and species distribution models (1) AquaMaps (2) Maxent (3) DBEM with maximum sensitivity and specificity threshold and depicting (a) relative habitat suitability (0–1) in 1985 and (b) difference in relative habitat suitability (2050–1985 values).

relatively low threshold, maximum sensitivity plus specificity, the outputs show strong agreement in predicted habitat suitability throughout the southern North Sea and English Channel as well as coastal waters around the UK and the west coast of France. This area of maximum agreement decreases as the thresholds become more restrictive (i.e. panels 2a and 3a), with the strongest agreement under the 0.7 threshold being predominantly restricted to the south-west, south and south-east coasts of England and Wales. Larger areas of agreement were obtained for outputs in 2050 than 1985, indicating the increased suitability of the North Sea, Norwegian coast and waters around Scotland and Northern Ireland to Pacific oyster under climatic change. This is highlighted in Figure 3, panels 1b, 2b, 3b, which show the agreement of change in habitat suitability (2050–1985) across threshold cut-off values of 0.1, 0.2, and 0.4, respectively. There is a prevalence of positive differences in RHS, with decreases in RHS values predominantly being restricted to the English Channel and southern European waters. Although maximum agreement (-6 or 6) in ΔRHS was not achieved in UK waters, there is relatively high model agreement in >0.1 ΔRHS (RHS values lying between 0 and 1), in particular three or four models predicting positive change around the west and north coasts of Scotland and west coast of Norway. Three models also predict increasing habitat suitability change of >0.4 ΔRHS in the north-east Irish Sea. There was relatively poor agreement between model predictions in the Skagerrak, Kattegat, Baltic and coastal waters of the Mediterranean seas.

DISCUSSION

Overall, these findings suggest that UK waters are projected to become more suitable for Pacific oyster populations, potentially allowing further expansion of oyster distribution. Although an increase in suitable habitat area for the Pacific oyster is predicted within the UK Exclusive Economic Zone for some model combinations,

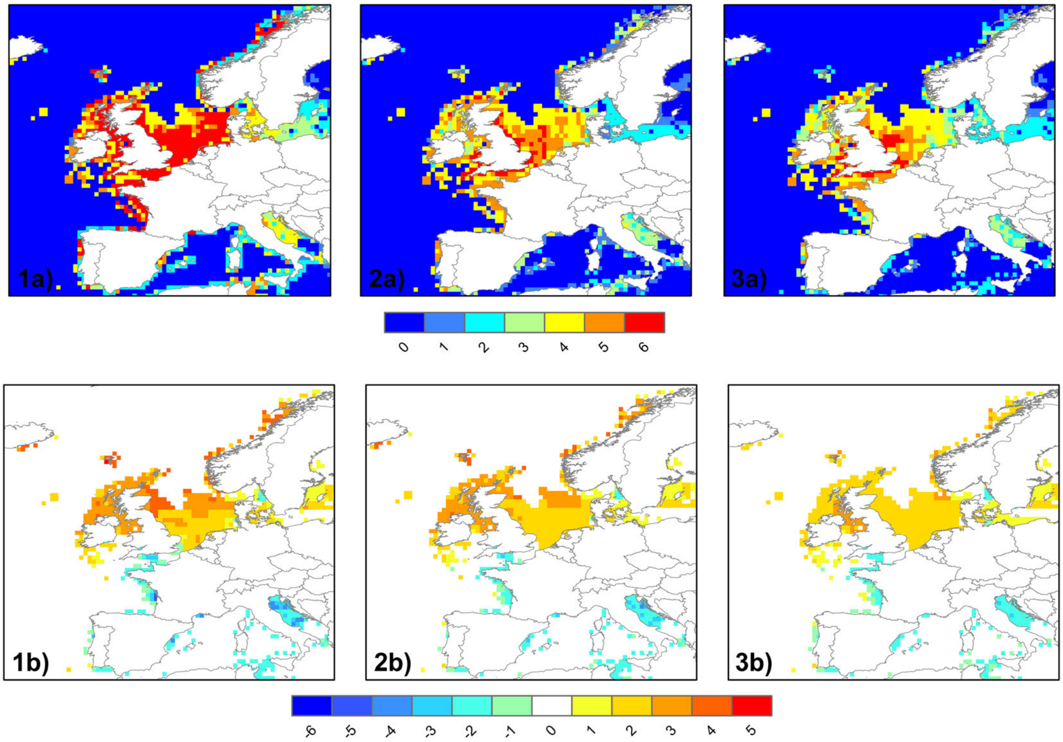

Figure 3. Ensemble predictions for Pacific oyster using the SDMs AquaMaps, Maxent and DBEM and environmental data from GFDL ESM 2.1 and CMIP3-E ensemble GCMs. (a) Agreement in relative habitat suitability predicted for 2050 and threshold cut-off values set at (1) maximum sensitivity and specificity, (2) 0.5 (3) 0.7. Cells with no prediction using any models remain at 0 (blue). (b) Agreement in change in relative habitat suitability values between projection (2050) and reference (1985) time periods, with threshold cut-off values set at (1) ±0.1, (2) ±0.2, (3) ±0.4. Within each threshold, cells predicted to show no change across all models remain at 0 (white).

range expansion may also occur through colonization of currently suitable but unoccupied habitats. Also consistent with hypotheses of potential range expansion for Pacific oyster are the predictions of increased habitat suitability within the range, in particular reflected in the calculation of latitudinal centroids.

Before applying the ensemble approach, results presented here highlight the variability in projections obtainable using different SDM and GCM combinations. Although predictions contrast previous studies that found the DBEM to predict the greatest distributional shifts (Jones *et al*., 2013), results are consistent with the inclusion of dispersal, the influence of which depends on species-specific parameters of dispersal. As the Pacific oyster is sedentary, no adult movement is included. Variation between SDMs agrees with previous studies suggesting differences to be predominantly influenced by the specific model characteristics and techniques applied (Jones *et al*., 2012). However, projections from AquaMaps and the DBEM show a larger amount of variation caused by the climate dataset used (GFDL or CMIP3-E). This may result from the varying ability of the GCMs to model the shelf sea and coastal regions where the Pacific oyster is predominantly found. In this case, Maxent is more robust to variations in the climate data. However, as many uncertainties exist in predicting environmental and oceanographic change in shelf seas, considering impacts from different predictions is important. This work may thus benefit from the inclusion of datasets from additional climate models, discussed below.

Applying predictions from SDMs to assess invasive alien species or develop management plans would benefit from systematic ground-truthing and continued monitoring to assess the rate of spread of a species into predicted suitable habitat. Although using occurrence data following previous range expansion into UK waters more accurately portrays locations that Pacific oyster might currently inhabit, some of these areas may not be inhabited at present due to dispersal limitation in the time period this habitat has become suitable. For example, no records of occurrence were obtained for the east coast of Scotland, although the environmental envelopes generated here predicted suitable habitat for both the current and future time period. For this reason also, absence data would not contribute valid information on the habitat preferences of an invasive alien species. Although absence data might contribute useful information for species occupying their entire environmental niche space, these data would be misleading if absence at a location were caused by factors not used in constructing an environmental envelope, such as dispersal. These maps therefore provide information on sites of likely invasion both currently and under climate change. This information may be subsequently informed and refined using spatially explicit information on suitable substrate for the Pacific oyster. For example, although Pacific oyster can be found on mud and sand–mud substrate, settlement and invasion may be more likely in areas of hard or rocky substrates, on which they preferentially attach (FAO, 2012). Although the benefits of applying SDMs to invasive alien species and the foresight obtainable are highlighted here, this application also demonstrates one of the difficulties in modelling species for which no systematic sampling has been undertaken, or for which locations of occurrence may be rapidly changing.

Given the variability in both model algorithms and output characteristics (Jones *et al.*, 2012), a consensus approach to ensemble forecasting was not considered appropriate in this case, and unlikely to match the truth (Thuiller *et al.*, 2004; Araújo and New, 2007). A method of compiling predictions more analogous to the bounding box method was therefore applied. The ensemble maps compiled provide a useful method of conveying the uncertainty and variation in species distributions resulting from a multi-model approach.

However, it is important to note that the set of SDMs and GCM projections applied here does not provide comprehensive coverage of model characteristics and possible sources of variability. Variation and uncertainty may result from SDMs according to their complexity, data requirements and algorithms. Similarly, uncertainty is introduced into GCMs due to the current climate, downscaling method, greenhouse gas emission scenario, and the climate model itself. Thus while an ensemble allows variability according to differences in component models to be captured, certain model similarities cause other variability to be excluded. For example, although the ensemble compiled here takes into account differences in species environmental envelopes, the data and mechanisms used to construct them, none of the SDMs consider inter-specific interactions or evolutionary adaptation, all relying on the assumption that a species is in pseudo-equilibrium with its environment. This assumption may not be upheld if models fail to take into account biotic interactions that prevent species occupying otherwise suitable habitat, or in the case of modelling a species for which there are few, or outdated, data, as discussed above. Furthermore, both climate models from which data were obtained are global and may be criticized for poor resolution of the topography and dynamics of the coastal shelf sea. The utility of spatial projections that do not incorporate the full ranges of uncertainty has been cautioned against (Planque *et al.*, 2011). However, this ensemble technique presents information as a range of possibilities, thereby providing a useful method to project climate-shifted distributions that can be updated and refined as alternative techniques and data become available. Further exploration of this example in the UK shelf sea may thus benefit from incorporation of a down-scaled regional climate model, ground-truthing of input data and incorporation of further predictors such as substrate, as suggested above. Although climate data may also incorporate systematic bias (Stock *et al.*, 2011), no bias correction was undertaken for this study due to insufficient observation data for all environmental variables incorporated in this study. However, both building and projecting species distribution models using modelled climate data ensured a consistency that minimized the effect of this source of bias on conclusions drawn.

With the exception of Maxent, which was found to be insensitive to changes in predicted range size using this set of thresholds, the use of thresholds enables some of the information contained in the prediction (the probability distribution) to be maintained while being converted into a simpler, more easily communicated, policy-relevant format. An application may be informed by a threshold selected according to the costs and risks involved in addressing a particular environmental problem. For an invasive alien species, this would involve weighing up the costs involved with preventing establishment compared with economic losses following establishment and costs of management or eradication. Eradication programmes can be very expensive. For example, the cost of eradication of the current (very small) UK population of carpet sea squirt (*Didemnum vexillum*) from marinas was estimated at £2.4 million in a recent study for the UK Department for Environment Food and Rural Affairs (Pinnegar *et al.*, 2012). If the carpet sea squirt were to spread over the whole of the UK, then the overall cost of eradication could rise to £72 million (Williams *et al.*, 2010). Eradication of Pacific oyster would be even more costly. On the other hand, range expansion of Pacific oyster may favour the development of highly profitable oyster farming. The risk of ecological impacts from invasion of Pacific oyster and the economic benefits from oyster farming need to be properly assessed; approaches employed here would be useful for such an assessment.

A set of thresholds may also be treated as a range of scenarios on extent of spread of a particular species. It is here emphasized that decision-making may be aided and policy enhanced through consideration of a range of the available science and the extent of agreement between alternative model formulations, rather than the use of a perceived 'best' model. Ensemble models may provide a useful solution for policy-making for the future when there is uncertainty concerning the reliability and accuracy of data and model outputs or disparity in the assumptions of particular models. However, the ensemble approach should not be viewed as an alternative to improving and developing models and collecting better data. In modelling a changing environment, data may be continually updated and models refined. While this is occurring, ensemble forecasts can go some way to providing confidence and thereby prevent inaction due to uncertainty and barriers in understanding between modellers and environmental policy-makers

Ensemble predictions made here suggest that Pacific oyster will experience suitable habitat as far north as the Faroe Islands and the eastern Norwegian Sea by 2050. In the worst case scenario, there is substantial agreement in this prediction between SDMs and alternative climatic datasets. A potential northward movement will also coincide with a decrease in habitat suitability in the English Channel and an increase in the southern and central North Sea. However, whether the species will be able to fully exploit this potential environmental niche will depend on its ability to disperse and settle. A more optimistic possibility is suggested by limiting areas of potential habitat suitability to only the most suitable habitat. In this case, although areas around the north of Scotland, Shetland and Faroe Islands are still predicted as suitable both currently and in the future, there is less agreement between models. Predictions such as these may thus be combined with analyses of the perceived costs, and risk of establishment and eradication, to target management plans most efficiently. They would further facilitate application of the precautionary approach to non-native species that is emphasized in the Convention on Biological Diversity (CBD) when there is lack of firm scientific evidence. This technique may be usefully incorporated into frameworks such as the 'GB risk analysis mechanism', developed to promote risk assessment and the precautionary approach, as stipulated in the CBD.

ACKNOWLEDGEMENTS

This work was supported by the UK Department for Environment, Food and Rural Affairs (DEFRA)/ Centre for Environment, Fisheries and Aquaculture Science (CEFAS) under project ME5102- ACME. We would like to thank the *Sea Around Us Project*, Fisheries Centre, University of British Columbia, the International Council for the Exploration of the Sea (ICES), the World Climate Research Program (WCRP) Coupled Model Intercomparison Project phase 3 (CMIP3) (http://esg.llnl.gov:8080) and the Geophysical Fluid Dynamics Laboratory (GFDL) (http://nomads.gfdl.noaa.gov) for data. William Cheung acknowledges funding support from the

National Geographic Society and Natural Sciences and Engineering Research Council of Canada. John Pinnegar acknowledges funding support from the EU FP7 project 'VECTORS'.

REFERENCES

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

Araújo MB, Whittaker RJ, Ladle RJ, Erhard M. 2005. Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography* **14**: 529–538.

Araújo MB, Thuiller W, Pearson RG. 2006. Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography* **33**: 1712–1728.

Bates JM, Granger CWJ. 2012. The combination of forecasts. *Operational Research Quarterly* **20**: 451–468.

Carvalho SB, Brito JC, Crespo EG, Watts ME, Possingham HP. 2011. Conservation planning under climate change: toward accounting for uncertainty in predicted species distributions to increase confidence in conservation investments in space and time. *Biological Conservation* **144**: 2020–2030.

Cheung WWL, Lam VWY, Sarmiento JL, Kearney K, Watson R, Pauly D. 2009. Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries* **10**: 235–251.

Cheung WWL, Dunne J, Sarmiento JL, Pauly D. 2011. Integrating ecophysiology and plankton dynamics into projected maximum fisheries catch potential under climate change in the Northeast Atlantic. *ICES Journal of Marine Science* **68**: 1008–1018.

Cheung WWL, Sarmiento JL, Dunne J, Frolicher TL, Lam VWY, Deng Palomares ML, Watson R, Pauly D. 2012. Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. *Nature Climate Change* **3**: 254–258.

Close C, Cheung W, Hodgson S, Lam V, Watson R, Pauly D. 2006. Distribution ranges of commercial fishes and invertebrates. *Fisheries Centre Research Report* **14**: 17–37.

Coetzee BWT, Robertson MP, Erasmus BFN, van Rensburg BJ, Thuiller W. 2009. Ensemble models predict Important Bird Areas in southern Africa will become less effective for conserving endemic birds under climate change. *Global Ecology and Biogeography* **18**: 701–710.

Collins M, Chandler R, Cox PM, Huthnance JM, Rougier J, Stephenson DB. 2012. Quantifying future climate change. *Nature Climate Change* **2**: 403–409.

Diniz-Filho JAF, Mauricio Bini L, Fernando Rangel T, Loyola RD, Hof C, Nogués-Bravo D, Araújo MB. 2009. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography* **32**: 897–906.

Drake JM, Bossenbroek JM. 2004. The potential distribution of zebra mussels in the United States. *BioScience* **54**: 931–941.

Dunne JP, Gnanadesikan A, Sarmiento JL, Slater RD. 2010. Technical description of the prototype version (v0) of tracers of phytoplankton with allometric zooplankton (TOPAZ) ocean biogeochemical model as used in the Princeton IFMIP model. *Biogeosciences* **7** (Suppl.): 3593. http://www.biogeosciences.net/7/3593/2010/bg-7-3593-2010-supplement.pdf

FAO. 2011. http://www.fao.org/fishery/species/search/en [March 2011].

FAO. 2012. *Crassostrea gigas*, FAO factsheet. Fisheries and Aquaculture Department, FAO. http://www.fao.org/fishery/culturedspecies/Crassostrea_gigas/en [June 2012].

Froese R, Pauly D (eds). FishBase. World Wide Web electronic publication 2011. www.fishbase.org, version (02/2011) [December 2011].

Global Invasive Species Programme. 1999. http://www.gisp.org/about/IAS.asp [July, 2012].

Guy C, Roberts D. 2010. Can the spread of non-native oysters (*Crassostrea gigas*) at the early stages of population expansion be managed? *Marine Pollution Bulletin* **60**: 1059–1064.

Herborg LM. 2007. Predicting the range of Chinese Mitten Crabs in Europe. *Conservation Biology* **21**: 1316–1323.

IPCC. 2000. *Emission Scenarios, Summary for Policymakers*, Nakicenovic N, Swart R (eds). Cambridge University Press: Cambridge.

IPCC. 2007. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt K, Tignor M, Miller H (eds). Cambridge University Press: Cambridge, UK and New York, USA.

JNCC Marine Advice: Non-Native Species. 2012. http://jncc.defra.gov.uk/page-1528 [June 2012].

Jones MC, Dye SR, Pinnegar JK, Warren R, Cheung WWL. 2012. Modelling commercial fish distributions: prediction and assessment using different approaches. *Ecological Modelling* **225**: 133–145.

Jones MC, Dye SR, Fernandes JA, Frölicher TL, Pinnegar JK, Warren R, Cheung WWL. 2013. Predicting the impact of climate change on threatened species in UK waters. *PloS One* **8**: e54216.

Kaschner K, Watson R, Trites A, Pauly D. 2006. Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. *Marine Ecology Progress Series* **316**: 285–310.

Kolar CS, Lodge DM. 2001. Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution* **16**: 199–204.

Liu C, Berry PM, Dawson TP, Pearson RG. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**: 385–393.

Marine Life Information Network (MarLIN). www.marlin.ac.uk. [March 2011].

Maggs C, Mineur F, Bishop J, McCollin T. 2010. Non-natives. In *MCCIP Annual Report Card 2010–11, MCCIP Science Review*, 1–11.

Mann R. 1979. Some biochemical and physiological aspects of growth and gametogenesis in *Crassostrea gigas* and *Ostrea edulis* grown at sustained elevated temperatures. *Journal of the Marine Biological Association UK* **59**: 95–110.

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W. 2009. Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* **15**: 59–69.

Mcneely J. 2001. Invasive species: a costly catastrophe for native biodiversity. *Land Use and Water Resources Research* **1**: 1–10.

Molnar JL, Gamboa RL, Revenga C, Spalding MD. 2008. Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment* **6**: 485–492.

NNSS. 2012. Pacific oyster, *Crassostrea gigas*. https://secure.fera.defra.gov.uk/nonnativespecies/factsheet/factsheet.cfm?speciesId=1013 [June 2012].

Nehls G, Diederich S, Thieltges D, Stasser M. 2006. Waden Sea mussel beds invaded by oysters and slipper limpets: competition or climate control? *Helgoland Marine Research* **60**: 135–143.

Nenzén HK, Araújo MB. 2011. Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling* **222**: 3346–3354.

Pearson RG, Thuiller W, Araújo MB, Martinez-Meyer E, Brotons L, McClean C, Miles L, Segurado P, Dawson TP, Lees DC. 2006. Model-based uncertainty in species range prediction. *Journal of Biogeography* **33**: 1704–1711.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modelling of species geographic distributions. *Ecological Modelling* **190**: 231–259.

Pimentel D. 2005. Aquatic nuisance species in the New York state canal and Hudson River systems and the Great Lakes basin: an economic and environmental assessment. *Environmental Management* **35**: 692–702.

Pinnegar J, Watt T, Kennedy K. 2012. Climate Change Risk Assessment for the Marine and Fisheries Sector, UK Climate Change Risk Assessment 2012. www.defra.gov.uk/environment/climate/government/risk-assessment/#sectors [January 2013].

Planque B, Bellier E, Loots C. 2011. Uncertainties in projecting spatial distributions of marine populations. *ICES Journal of Marine Science* **68**: 1045–1050.

Prasad AM, Iverson LR, Liaw A. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* **9**: 181–199.

Ready J, Kaschner K, South AB, Eastwood PD, Rees T, Rius J, Agbayani E, Kullander S, Froese R. 2010. Predicting the distributions of marine organisms at the global scale. *Ecological Modelling* **221**: 467–478.

Sax D, Stachowicz J, Brown J, Bruno J, Dawson M, Gaines S, Grosberg R, Hastings A, Holt R, Mayfield M, *et al*. 2007. Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution* **22**: 465–471.

Sing T, Sander O, Beerenwinkel N, Lengauer T. 2007. The ROCR Package. http://rocr.bioinf.mpi-sb.mpg.de/

Smaal A, van Stralen M, Craeymeersch J. 2005. Does the introduction of the Pacific Oyster *Crassostrea gigas* lead to species shifts in the Wadden Sea? In *The Comparitive Roles of Suspension Feeders in Ecosystems. NATO Searies IV - Earth and Environmental Sciences*, Dame R, Olenin S (eds). Springer: Dordrecht, Netherlands.

Spencer BE, Edwards DB, Kaiser MJ, Richardson CA. 1994. Spatfalls of the non-native Pacific oyster, *Crassostrea gigas*, in British waters. *Aquatic Conservation: Marine and Freshwater Ecosystems* **4**: 203–217.

Stachowicz JJ, Terwin JR, Whitlatch RB, Osman RW. 2002. Linking climate change and biological invasions: ocean warming facilitates nonindigenous species invasions. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 15497–15500.

Stock CA, Alexander MA, Bond NA, Brander KM, Cheung WWL, Curchitser EN, Delworth TL, Dunne JP, Griffies SM, Haltuch MA, *et al*. 2011. On the use of IPCC-class models to assess the impact of climate on Living Marine Resources. *Progress in Oceanography* **88**: 1–27.

Thuiller W, Araújo MB, Pearson RG, Whittaker RJ, Brotons L, Lavorel S. 2004. Biodiversity conservation: uncertainty in predictions of extinction risk. *Nature* **430**: 145–148.

Underwood E, Ustin S, DiPietro D. 2003. Mapping non-native plants using hyperspectral imagery. *Remote Sensing of Environment* **86**: 150–161.

Utting S, Spencer B. 1992. Introductions of bivalve molluscs into the United Kingdom for commercial culture - case histories. *ICES Marine Science Symposium* **194**: 84–91.

Williams F, Eschen R, Harris A, Djeddour D, Pratt C, Shaw R, Varia S, Lamontagne-Godwin J, Thomas S, Murphy S. 2010. *The Economic Cost of Invasive Non-Native Species on Great Britain*. CABI Europe: UK.