



Identifying marine invasion hotspots using stacked species distribution models

Devin A. Lyons · J. Ben Lowen · Thomas W. Therriault · David Brickman · Lanli Guo · Andrea M. Moore · M. Angelica Peña · Zeliang Wang · Claudio DiBacco

Received: 9 August 2019 / Accepted: 7 August 2020 / Published online: 29 August 2020
© Springer Nature Switzerland AG 2020

Abstract Early detection and management of aquatic invasive species requires identification of those areas most at risk of invasion (i.e., hotspots). Here we identify present-day and future hotspots of invasion risk for marine invertebrates and algae in nearshore habitats of the northwest Atlantic and northeast Pacific using more than 12 years of monitoring data in conjunction with other occurrence data and stacked species distribution models. The stacked species distribution models predicted the general patterns of observed invasive species richness in both study areas (Atlantic: $r^2 = 0.52$, Pacific: $r^2 = 0.42$). In the

northwest Atlantic, we identified hotspots through much of Massachusetts, New Hampshire and southern Maine, and in several bays in southwestern New Brunswick and Nova Scotia. In the northeast Pacific, much of the southern Salish Sea was identified as a hotspot, as were a few areas along the outer coast of Washington and Oregon. Projecting our species distribution modelling results to 2075 (climate scenario RCP 8.5), we found that existing hotspots are likely to expand slightly in the Atlantic, while in the Pacific existing hotspots are predicted to shift or expand, new hotspots are likely to appear, and areas with few invasive species attaining moderate invasive species richness. Our results suggest that climate change will have larger effects on the distributions of our focal invasive species on the Pacific coast compared to the Atlantic. Resultant hotspot maps provide an integrated perspective and guidance to managers tasked with prioritizing locations for monitoring and implementing policy related to marine invasive species, with projected hotspots making planning for future changes in invasion risk possible.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10530-020-02332-3>) contains supplementary material, which is available to authorized users.

D. A. Lyons (✉) · J. B. Lowen · D. Brickman · L. Guo · A. M. Moore · Z. Wang · C. DiBacco
Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, NS, Canada
e-mail: dalyons@dal.ca

D. A. Lyons
Department of Zoology, University of British Columbia, Vancouver, BC, Canada

T. W. Therriault
Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC, Canada

M. A. Peña
Institute of Ocean Sciences, Fisheries and Oceans Canada, Sidney, BC, Canada

Keywords Marine invasive species · Species distribution models · Climate change responses

Introduction

Invasive species are a major component of global environmental change, contributing to biotic homogenization and creating threats to biodiversity, ecosystem services, human health, and economic activities (Vitousek et al. 1997; Pejchar and Mooney 2009; Pyšek and Richardson 2010). However, knowledge gaps about the basic and applied ecology of species invasions and a research implementation gap between the knowledge generated by ecologists and that which is used by policy and management practitioners are persistent problems that impede the monitoring, management, and scientific understanding of invasive species (Esler et al. 2010; Bayliss et al. 2013). Tools that synthesize and transfer scientific knowledge effectively are needed in order to narrow these knowledge gaps, and to foster collaboration between the scientific, management, and policy development communities (Bayliss et al. 2013; Matzek et al. 2014; Giakoumi et al. 2016).

The spatial distributions of invasive species are a key piece of knowledge required for effective monitoring and management, but there is often considerable uncertainty about where invasive species do and do not occur. Unfortunately, few large-scale multi-species monitoring programs exist, and much of the information available on the distributions of invasive species comes from scattered, incidental occurrence records (e.g. Johnson et al. 2006; Gormley et al. 2011). Where dedicated surveys or monitoring programs have been implemented, they are often focused exclusively on areas of high activity by introduction and transport vector activity (e.g. ship traffic and aquaculture; Arenas et al. 2006; Carman et al. 2010). Surveys that sample beyond hubs of vector activity are almost inevitably limited to smaller spatial extents and/or coarser spatial resolutions than are needed for effective management action by logistical constraints. The deficiencies of occurrence information are exacerbated in marine ecosystems, where observing and monitoring invasive species is typically more difficult, and fewer potential observers are active, than on land.

Anthropogenically-induced environmental change is altering the biophysical properties of the ocean and has the potential to transform marine ecosystems and the human systems that rely on them (Cheung 2019). For example, the distributions of a wide variety of marine taxa, including important fisheries species, are

already tracking changes in local climate (Pinsky et al. 2013). There is strong scientific and practical interest in predicting how the distributions of invasive species are likely to change in response to future climate change, but present-day and historical occurrence records alone can tell us little about this (Jeschke and Strayer 2008; O'Donnell et al. 2012; Lowen et al. 2017). Consequently, reliable information about the current and future distributions of invasive species remains a major knowledge gap, and has contributed to the omission of invasive species from more than 95% of conservation plans, globally (Giakoumi et al. 2016; Mačić et al. 2018).

The limitations of existing occurrence data have increasingly caused scientists turn to species distribution models (SDMs), which predict the distribution of organisms based on the distribution of appropriate biophysical conditions (Elith and Leathwick 2009). Species distribution models allow us to use known, spatially incomplete occurrence locations to predict comprehensive distributions with spatially variable, quantitative, estimates of occurrence probability (Rondinini et al. 2006). When coupled to climate projections, they also allow us to obtain first-order predictions of how these distributions might shift in response to expected environmental change (Jeschke and Strayer 2008; O'Donnell et al. 2012; Lowen et al. 2017). Moreover, maps derived from individual species' distribution models are powerful communication tools that can be used to support policy development, guide management decision-making, inform systematic conservation planning, and optimize monitoring and detection programs (Ball et al. 2009; Venette et al. 2010; Baxter and Possingham 2011; Guisan et al. 2013). However, ecosystems are at risk of invasion by many non-indigenous species, each of which is likely to have its own distribution, environmental requirements, potential impacts and associations with vectors of introduction and spread, but understanding and communicating this complexity and detail is difficult. One way to deal with this complexity is to identify invasion hotspots, or areas subject to risk of invasion by multiple invaders (O'Donnell et al. 2012; Li et al. 2016). Just as biodiversity hotspots have been used extensively to focus conservation efforts, the concept of invasion hotspots can help to provide a focus for policy and management and support efforts to optimize our allocation of the limited resources available for

monitoring and management in ecosystems at risk of invasion by many species (O'Donnell et al. 2012).

The concept of invasion hotspots has been used frequently in invasion ecology and management, but its use has often been vague and instances where it has been formalized or applied in aquatic ecosystems are comparatively rare. Hotspots of marine invasions have been identified via counts of invasive species within individual embayments, governmental units, or ecoregions (e.g., Cohen and Carlton 1998; Ruiz et al. 2011; Marchini et al. 2015), or by using the intensity of vector activity as a proxy for invasion risk (e.g. Drake and Lodge 2004; Tidbury et al. 2016). Recent studies have combined information on global ship traffic patterns and the environmental distance between ports to model invasion risk and identify hotspots at the port and ecoregion level (Seebens et al. 2013, 2016). An alternate approach is to integrate (or “stack”) multiple high resolution species distribution models to identify areas of high invasion potential as indicators of invasion hotspots and risk, as has been done in some terrestrial systems (O'Donnell et al. 2012; Li et al. 2016). While the other approaches have proved informative and useful, they produce assessments that are focused on relatively small scales (e.g. single embayments), are focused on large scales but provide spatially homogenous estimates for large areas (e.g. entire ecoregions), or that provide spatially variable, but disjunct assessments over large scales (e.g. for large ports around the world). Stacked species distribution models are a promising tool because they can provide higher resolution, spatially comprehensive and variable assessments of present-day invasion risk over large spatial extents and predictions of how this risk might respond to environmental change. However, there are few studies using stacked species distribution models to assess the risk of multiple marine invasions or identify marine invasion hotspots (but see Lowen et al. 2017). In fact, even the number of studies modelling the ranges of individual marine invasive species and how they are likely to respond to climate change has been minimal compared to those on terrestrial or freshwater ecosystems (13, 314 and 98 studies, respectively) (Bellard et al. 2018).

In this paper, our goals were to use species distribution models to describe spatial patterns in predicted invasive species richness, to identify present-day and future hotspots as spatial indicators of risk of invasion, and to describe how maps of predicted

invasive species richness and invasion hotspots could be used as simple communication tools to guide monitoring and management decisions. We coupled occurrence data of invasive species on the east and west coasts of North America with high resolution seasonal climatologies for environmental variables via species distribution models to predict current and future distributions of 12 high-risk invasive species per coast. Integrating these species distribution models allows us to compare patterns of predicted invasive species richness to observed patterns, to identify hotspots and coldspots of invasion risk, and to examine how these patterns are predicted to change under climate change scenarios RCP 8.5 by 2075. Finally, we describe how hotspots of invasion risk can be used to inform managers and decision makers.

Methods

Study regions

We focused on the northwest Atlantic and the northeast Pacific for this study, concentrating on areas where more than a decade's worth of systematic marine invasive species monitoring data and high-resolution marine climate models from Fisheries and Oceans Canada could be leveraged. Within both regions, there is significant interest from government in using information about the present day and potential future distributions of marine invasive species in ecosystem management, environmental protection, and the design of monitoring programs. In the Atlantic, our model domain was from 32° to 53° N and 49° to 80° W. In the Pacific, our model domain was 24°–62° N and 111°–155° W. Although our present-day species distribution models covered these broader domains, we focused on smaller domains, where we were able to use 2075 climate projections (Atlantic: 41°–53°N, 52°–73° W; Pacific: 45.5°–56°N, 122.25°–135.75° W). A map of each study region and the places mentioned in the text can be found in Online Appendix 1.

Species occurrence data

To identify invasion hotspots, we used a subset of 12 invasive species on each coast. We focused on seaweeds and benthic/demersal invertebrates of

shallow-water coastal habitats as little information is available about potential pelagic and offshore invasive species in our study regions. However, we selected species covering a range of taxonomic and life history variation that have already invaded the broader region (northwest Atlantic or northeast Pacific) and that have been identified as moderate to high risk of invasion and impact in Canadian waters (Drolet et al. 2016). We selected well-established (> 20 years) species with occurrence records from at least 30 unique locations to ensure that our focal species would have had time to spread within the study region, and that we had sufficient data for species distribution modelling. For the Northwest Atlantic, our focal species included solitary (*Asciidiella aspersa*, *Ciona intestinalis*, *Styela clava*) and colonial (*Botrylloides violaceus*, *Botryllus schlosseri*, *Diplosoma listerianum*, *Didemnum vexillum*) tunicates, a skeleton shrimp (*Caprella mutica*), two shore crabs (*Carcinus maenas*, *Hemigrapsus sanguineus*), a colonial bryozoan (*Membranipora membranacea*) and a seaweed (*Codium fragile*). For the Northeast Pacific, the focal species included six of the same species used in the Atlantic (*B. violaceus*, *B. schlosseri*, *C. mutica*, *C. maenas*, *D. vexillum*, *S. clava*), four bivalves (*Mya arenaria*, *Crassostrea gigas*, *Venerupis philippinarum*, *Nuttallia obscurata*), a gastropod (*Ocenebrellus inornatus*), and a seaweed (*Sargassum muticum*). All these species are algae or ectotherms that have adapted differently across cold to warm temperate saltwater environments, and whose development rates and distributions are expected to be strongly delineated and constrained by temperature and salinity.

We compiled occurrence data (latitude and longitude of each observation of each species) primarily from Fisheries and Oceans Canada's (DFO) Aquatic Invasive Species Monitoring, the Ocean Biogeographic Information System (OBIS 2018), as well as several other online databases and primary and grey literature publications (Online Appendix 1), using data only from the focal region. We chose to exclude data from other parts of the world to ensure our models would reflect the environmental responses of the genetic lineages that have invaded our study regions. For plots of observed invasive species richness, we rasterized the occurrence points at a resolution of 0.5°, counting the number of unique invasive species (from our focal species) in each grid cell. We did not include grid cells without observations, as it is often likely that

no sampling for our focal species was conducted in these areas. Prior to model fitting, we spatially rarified the occurrence points to a 10 km resolution using SDMtoolbox 2.0 for ArcGIS to reduce potential effects of spatial autocorrelation on our results (Brown et al. 2017).

Environmental predictors

We selected surface water temperature, salinity, and wave action (i.e., significant wave height, Atlantic only), which are known to influence the distributions of coastal marine invertebrates and algae (Burrows et al. 2008; Lowen et al. 2016), as predictors in our species distribution models. Our decision not to include other predictors reflects a lack of available data (particularly future projections) for potentially relevant variables when we began this work. We intentionally excluded variables such as depth, latitude, and distance from shore which are sometimes used as proxies for factors that control distributions, but whose relationships to those factors is likely to change in space and time. These decisions reflect recommendations to limit variables to functionally relevant predictors, and the dangers of using distal and proxy variables, which often make models more error-prone when they are extrapolated in space or time (Elith and Leathwick 2009). As our work involves extrapolation in time, including these proxies might improve model fit for the present day, but would risk significantly degrading future projections. Data for each predictor was averaged into seasonal climatology rasters, including winter (Atlantic: January to March, Pacific: December to February) spring (Atlantic: April to June, Pacific: March to May), summer (Atlantic: July to September, Pacific: June to August), and fall (Atlantic: October to December, Pacific: September to November). The difference in the definition of the seasons reflect how physical Oceanographers/ocean climate modellers typically define the seasons in these regions, and are related to the timing of events like stratification/destratification, warming, cooling, and the spring phytoplankton bloom.

In the Atlantic, sea surface temperature and salinity values were assembled at 0.01° resolution from AMSR-E Level 3 sea surface temperature satellite data (Advanced Very High Resolution Radiometer data, AVHRR Atlantic; 2002–2012; compiled by Fisheries and Oceans Canada) and global

oceanographic salinity composites (BioOracle; Tyberghein et al. 2012). Seasonal significant wave height values were derived from numerical models (Wavewatch III, Rascole and Ardhuin 2013, 0.167° resolution; Guo and Sheng 2017, 0.125° resolution). Future seasonal climatologies of salinity and temperature matching the resolution (0.01°) of “present day” values were derived from numerical model projections (BNAM RCP 8.5 2075 monthly anomalies; Brickman et al. 2016), as were future seasonal significant wave height climatologies (Wavewatch III driven by Canadian Regional Climate Model, CanRCM4, predictions for 2070–2071, 0.125° resolution, Guo and Sheng 2017). Lower resolution wave height values (present and future) were resampled, using the nearest neighbor value, to match the 0.01° temperature and salinity values. Our choice of scenario RCP 8.5 2075 reflects the fact that our oceans are already warming significantly faster than previously expected (Cheng et al. 2019). It also reflects the availability of ocean climate projections for this scenario on both the Atlantic and Pacific coasts produced by models developed by Fisheries and Oceans Canada (including several of the authors of this manuscript), and our familiarity with them.

In the Pacific, we compiled salinity and temperature values from hindcasts from the University of British Columbia’s Salish Sea Nucleus for European Modelling of the Ocean (NEMO) model (2014–2017 hindcast, 0.006° resolution, Soontiens et al. 2016; Soontiens and Allen 2017), a Regional Ocean Modelling System (ROMS) model of the British Columbia shelf (BC ROMS 1981–2010 hindcast, 0.04° resolution, Peña et al. 2019), as well as the MARSPEC database (0.00833° resolution, Sbrocco and Barber 2013). The higher resolution data from the NEMO model and MARSPEC were resampled using bilinear interpolation to match the 0.04° resolution of the ROMS model data. Future climatological scenarios for salinity and temperature were derived from projections of the BC ROMS model (RCP 8.5, 2041 to 2070, 0.04° resolution, Peña et al. 2018). Extrapolation detection analyses (Mesgaran et al. 2014) indicated neither novel values for individual predictors (i.e. values that fell outside the range of those in our present day values), nor novel combinations of covariates was present in our future climate scenario for either the Atlantic or Pacific study area (Fig. A1.3, Online Appendix 1), indicating that we did not need to

be concerned about our SDM projections involving extrapolation into non-analogous environmental conditions.

Although our focal species live primarily in the intertidal zone or very shallow subtidal areas close to shore, and the vast majority of our occurrence records were from depths shallower than 30 m, some focal species (e.g. *C. maenas*, *C. intestinalis*, *D. vexillum*) have been observed in deeper waters. We initially planned to limit our analysis to areas shallower than 100 m on both coasts. In the northwest Atlantic, this would have caused us to include data from large offshore banks, stretching up to several hundred km from shore in the background environmental data used to fit our SDMs. Our occurrence dataset comprised records almost entirely from coastal areas (most < 30 m), possibly because of increased sampling near shore. Such spatially biased sampling will generally lead to environmental differences between occurrence and background datasets that may result in inaccurate models (Phillips et al. 2009). Restricting the background environmental data used in SDMs is an effective way to reduce potential bias in model results (Kramer-Schadt et al. 2013; Brown et al. 2017). Thus, we clipped environmental data layers to areas shallower than 30-m depth in order to eliminate thousands of square km of offshore banks from 30 to 100 m. Of the 1314 occurrence records remaining after spatial rarefaction, 11 were from depths greater than 100 m, and an additional 31 were excluded as a result of this change. In the Pacific, the 100 m depth contour is typically much closer to shore than in the Atlantic, offshore banks do not exist to the same extent, and we had more records from areas 30 to 100 m deep. Consequently, we limited our analysis to areas shallower than 100 m, as originally intended.

Species distribution models

We modelled each species distribution using MaxEnt 3.4.1 (Phillips et al. 2017) with seasonal salinity, temperature, and significant wave height (Atlantic only) as predictors. MaxEnt is a presence-only method that estimates species distributions by identifying the distribution with maximum entropy, subject to constraints derived from the values of environmental covariates at presence locations. This is equivalent to minimizing the relative entropy between the probability distribution estimated for the covariates at

presence locations and that estimated for the covariates for the entire background landscape (Elith et al. 2011; Phillips et al. 2017). Although MaxEnt was originally developed from a machine learning perspective, it is equivalent to an inhomogeneous Poisson process model (Aarts et al. 2012; Renner and Warton 2013).

We used the default complementary log–log transform option within MaxEnt to produce estimates of occurrence probability, due to its better theoretical justification compared to the previously favored logistic transform (Phillips et al. 2017). To choose the feature class (i.e. potential response curve complexity) and regularization parameter (i.e. penalty for model complexity) settings for each species' model, we used the ENMeval package (Muscarella et al. 2014) for R version 3.5.2 (R Core Team 2018) to evaluate a broad range of combinations. We selected the model settings that resulted in the lowest corrected Akaike information criterion (AICc) for each species and fit a model using those settings in MaxEnt using the maximum possible number of background points (depending on which study area) and 30-fold random cross-validation to evaluate the models and estimate standard deviations for the model predictions. We also used this model, together with 2075 climate projection datasets to predict each species' future distribution.

Stacked species distribution models

We estimated grid cell-level invasive species richness by stacking our species distribution model predictions for present day and 2075 projections. We also calculated the centre of gravity (i.e. weighted centroid) of our present day and 2075 species richness estimates to examine the predicted latitudinal shift in the spatial distribution of invasive species richness within our study area. Our estimates of species richness are limited to the richness of our 12 focal species (per coast), so the potential range of our estimates (and observations) of species richness ranges from 0 to 12.

Species richness was estimated as the sum of the occurrence probability predictions from our individual models (Calabrese et al. 2014):

$$E(S_j) = \sum_{k=1}^K p_{j,k}$$

where $E(S_j)$ is the expected species richness (S) at site j , K is the number of species in the dataset (12 per coast), and $p_{j,k}$ is the occurrence probability prediction for species k at site j . This is equivalent to summing species suitability (based on occurrence probability) maps.

Assuming p_j are exact, known quantities, the variance of $E(S_j)$ should be estimated as (Calabrese et al. 2014):

$$\text{Var}(S_j) = \sum_{k=1}^K (1 - p_{j,k})p_{j,k}$$

However, in reality, p_j are estimated with uncertainty. To incorporate this uncertainty, we propagated the error estimated within our individual models by summing the estimated variances (Li and Wu 2006) and adding it to the expected variance, changing the equation to:

$$\text{Var}(S_j) = \sum_{k=1}^K (1 - p_{j,k})p_{j,k} + \sum_{k=1}^K sd_{j,k}^2$$

where $sd_{j,k}$ is the standard deviation of the model prediction for species k at site j . We used these estimates to construct Wald-type 95% confidence intervals for each site.

To check the performance of the stacked species distribution models we visually and statistically compared their predictions to observed species richness. For the statistical comparisons we aggregated our model predictions to the same 0.5° resolution used for the observed values, taking the mean of the finer resolution prediction grid cells. We then used linear regression to examine the relationship between the predictions and the observations. Ideally, these regression models would have an intercept near 0 and a slope near 1, and our predictions would statistically explain (predict) a significant proportion of the variation in observed species richness. However, we expected the proportion of explained variation to be modest and for model predictions to frequently exceed observed species richness because we believe some invaders may not have had the opportunity to disperse all potentially suitable areas, and that the observed number of the invaders in many areas is likely to be an underestimate of the true value because of insufficient sampling.

Hotspot/coldspot identification

In many other fields, hotspots and coldspots are defined using measures of local spatial association, such as Getis-Ord G_i^* , which allow one to identify clusters of values in weighted point data that are particularly high, or low, relative to surrounding areas (Getis and Ord 1992). In order to evaluate invasion risk and impact, the magnitudes of the predicted values are more important than their spatial association, so we defined hotspots as the areas within the highest invasive species richness and coldspots as the areas with the lowest species richness, without considering spatial clustering. Areas of intermediate species richness can be considered “warmspots”. Several approaches could be used to distinguish hotspots and coldspots from other areas. We used two to demonstrate how and why they might be used. The first evaluated whether the invasive species richness estimate in each pixel was significantly above or below specific thresholds (above 8 for hotspots, below 4 for coldspots), according to their 95% confidence interval. This approach takes uncertainty in model estimates into account and is effective in communicating how the magnitude and spatial distribution of invasion risk might change through time. We used this approach to identify hotspots and coldspots across our entire study regions. The second approach delineated hotspots and coldspots according to whether pixels fell within the top and bottom deciles (i.e. top and bottom 10%) of the estimates. This approach enables one to focus attention on a fixed proportion of a focal area, regardless of the actual values, and whether they change through time or between different focal areas. It will be useful in situations where logistical constraints impose a limit on the area that can be studied, sampled, monitored, or protected, or where one wants to produce similar results across changing contexts. We used this approach to identify hotspots and coldspots in different ecoregions within our broader study areas that have different numbers of invasive species. The species richness and quantile thresholds we set in our examples were arbitrary but, as for choosing a method to define or delineate hot- and cold-spots, there may be legislative, policy-based, scientific, or logistical reasons for choosing thresholds in other instances.

Results

Species distribution models

Our individual species distribution models exhibited good performance, with the area under the receiver operating characteristic curve (AUC) ranging from 0.760 to 0.950 and Boyce Index ranging from 0.7 to 0.98 in the Atlantic (Table 1). In the Pacific, AUC ranged from 0.815 to 0.948 and Boyce Index ranged from 0.69 to 0.97. Seasonal temperature variables were generally the best predictors of the distributions of our focal species in the Atlantic, with salinity and wave height variables each playing secondary or tertiary roles for different species (Table 1a). In the Pacific, seasonal temperature variables also were more important than salinity for all the species, except for 2 molluscs; *Nuttallia obscurata*, and *Ocenebrellus inornatus* (Table 1b).

Observed invasive species richness and stacked species distribution model predictions

In the Atlantic, most coastal grid cells contained observations of at least one of our focal species, except in much of Quebec and parts of Newfoundland (Fig. 1a). Where data were available, the number of our focal invasive species observed in the 0.5° grid cells ranged from 1 to 12, with particularly high numbers of invaders observed in parts of southwestern Nova Scotia, southern Maine, New Hampshire, Massachusetts, and Rhode Island. Few invaders were observed in the inner Bay of Fundy, northern New Brunswick, Quebec, and most of Newfoundland (see Figure A1.1 in Online Appendix for map of locations). Our present-day stacked species distribution model estimates for invasive species richness ranged from 0.2 to 11.1 of the possible 12 species, (4.0 ± 2.6 ; mean \pm Std Dev) (Fig. 1b), and exhibited a broadly similar spatial pattern to observed richness of our focal invasive species (Fig. 1a). The stacked species distribution model did a good job of predicting observed species richness, with the intercept and slope close to their ideal values and more than half the variation in observed species richness explained by predicted richness (intercept: -0.20 , slope: 0.91 , $r^2 = 0.52$, Online Appendix 2, Fig. A2.1). As expected (see Methods), over-predictions of observed species richness (62% of cells) were more common than under-

Table 1 Evaluations of (a) Atlantic and (b) Pacific model performance and variable importance for each species

Species	N occurrence	Mean AUC	AUC SD	CV folds	Continuous Boyce index	\sum Seasonal temperature importance	\sum Seasonal salinity importance	\sum Seasonal sig. wave ht. importance
(a) Atlantic								
<i>Ascidella aspersa</i>	89 (22)	0.885	0.21	19	0.92	62.8	10.7	26.7
<i>Botrylloides violaceus</i>	470 (126)	0.856	0.08	30	0.95	67.9	21.7	10.3
<i>Botryllus schlosseri</i>	528 (168)	0.826	0.075	30	0.97	61.5	20.4	18
<i>Caprella mutica</i>	292 (121)	0.834	0.077	30	0.98	61.5	29.2	9.2
<i>Carcinus maenas</i>	1357 (284)	0.805	0.051	30	0.97	68.6	17	14.4
<i>Ciona intestinalis</i>	424 (107)	0.898	0.038	30	0.94	69.5	21.1	9.4
<i>Codium fragile</i>	610 (122)	0.836	0.081	30	0.97	47.3	19.1	33.6
<i>Didemnum vexillum</i>	100 (32)	0.95	0.055	28	0.89	63.4	26.4	10.1
<i>Diplosoma listerianum</i>	57 (28)	0.76	0.3	26	0.70	94.9	4.2	1
<i>Hemigrapsus sanguineus</i>	270 (76)	0.93	0.039	30	0.97	79.2	0.5	20.3
<i>Membranipora membranacea</i>	418 (175)	0.837	0.062	30	0.89	80.7	10.2	9
<i>Styela clava</i>	281 (52)	0.894	0.076	30	0.93	70.8	13.9	15.3
(b) Pacific								
<i>Botrylloides violaceus</i>	386 (103)	0.865	0.12	30	0.90	76	23.9	–
<i>Botryllus schlosseri</i>	287 (70)	0.891	0.103	30	0.95	82.4	17.7	–
<i>Caprella mutica</i>	145 (49)	0.815	0.148	30	0.92	91.3	8.8	–
<i>Carcinus maenas</i>	3060 (57)	0.919	0.044	30	0.91	60.8	39.1	–
<i>Crassostrea gigas</i>	541 (129)	0.896	0.057	30	0.95	55.5	44.5	–
<i>Didemnum vexillum</i>	113 (33)	0.883	0.174	30	0.88	100	0	–
<i>Mya arenaria</i>	856 (203)	0.843	0.053	30	0.95	65.2	34.8	–
<i>Nuttallia obscurata</i>	466 (112)	0.928	0.032	30	0.95	40.8	59.3	–
<i>Ocinebrellus inornatus</i>	37 (22)	0.856	0.131	19	0.69	40.1	59.8	–
<i>Sargassum muticum</i>	466 (146)	0.845	0.085	30	0.97	81.9	18	–
<i>Styela clava</i>	210 (41)	0.948	0.057	30	0.91	94.8	5.1	–
<i>Venerupis philippinarum</i>	813 (159)	0.890	0.046	30	0.97	69.8	30.4	–

Mean area under the curve (AUC) and standard deviation (SD) were calculated for the test data from 30-fold cross validation (CV), except where sample size limited cross validation to fewer folds. The number of occurrences (N) presented here is the number of unique occurrence locations prior to spatial rarefaction, with the number remaining after rarefaction in parenthesis. For clarity and brevity we present the sum of the permutation importance scores for the seasonal climatologies of each environmental parameter. The values of most important environmental parameter are in bold

predictions (38% of cells). Our stacked species distribution model for 2075 predicts the average number of invasive species present per cell to increase slightly, from 4.0 to 4.5 (± 2.6 ; range 0.3–11.0; Fig. 1c). These relatively small changes in the mean and range of predicted invasive species richness reflect pixel-level changes that were mostly small and positive (range -1.8 to 2.5) as highlighted in the anomaly plot (Fig. 1d). The centre of gravity of our SSDM estimates shifted 33.6 km to the northeast, from (45.30 N, 64.60 W) to (45.47 N, 64.25 W).

In the Pacific, occurrence data were available from almost all the grid cells in our study area and the number of observed invaders in these grid cells ranged from 1 to 11 (Fig. 2a). High invasive species richness was observed primarily in the Salish Sea and some areas of West Coast Vancouver Island (Fig. 2a), low richness was observed along the central and northern coasts of British Columbia (BC), and low to moderate invasive species richness was observed along the outer coast of Washington and Oregon (see Figure A1.2 in Online Appendix for map of locations). Our present-

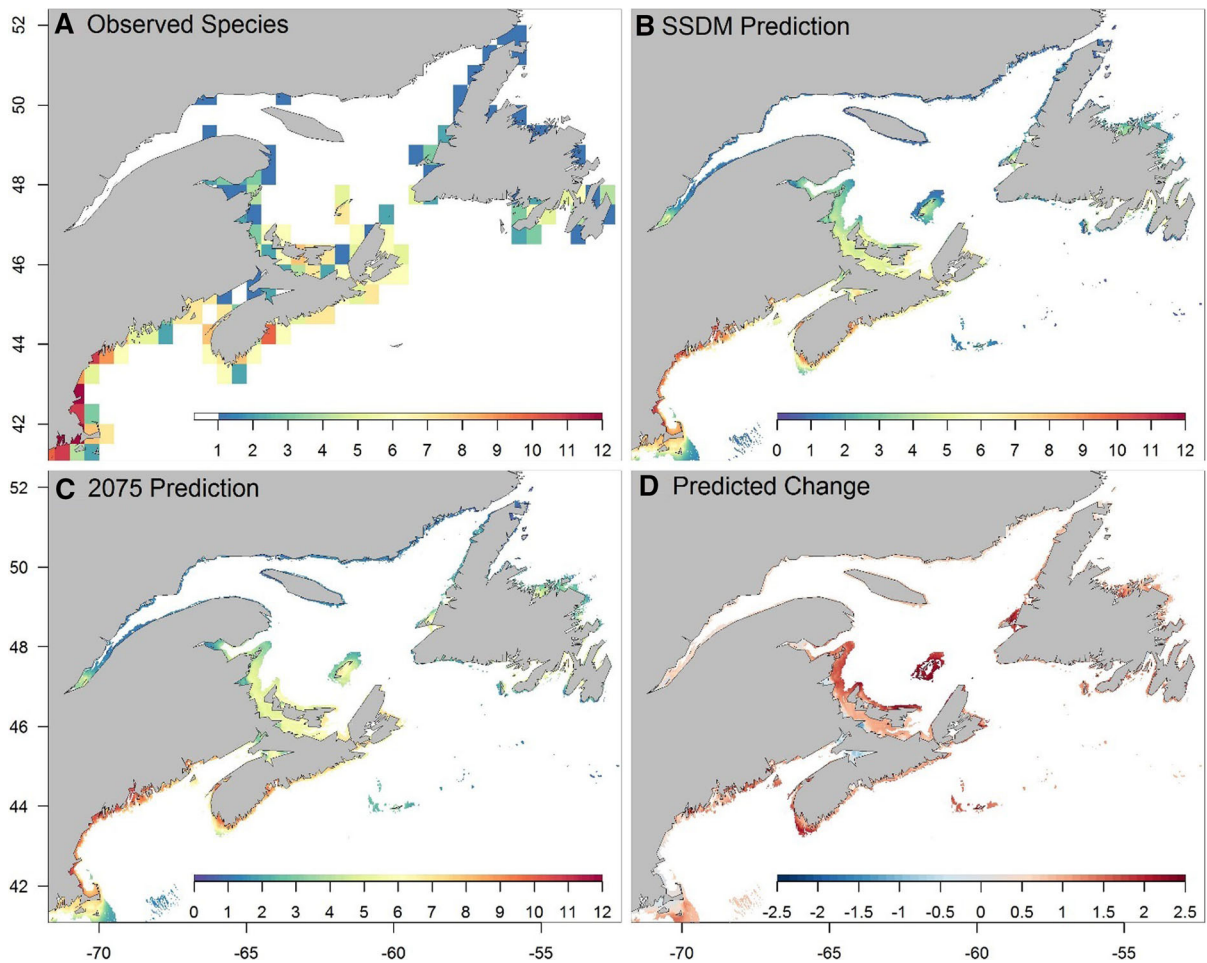


Fig. 1 Number of focal invasive species in the northwest Atlantic according to observations (a), present day stacked species distribution model (SSDM) (b), and 2075 SSDM

day stacked species distribution model estimates for invasive species richness ranged from 0.4 to 11.6 species, (4.6 ± 2.4). The stacked species distribution model again exhibited a broadly similar spatial pattern to observed species richness, though it predicted that more invasive species would be present along much of the outer Washington and Oregon coastline (Fig. 2b). The stacked species distribution model did a similar job of predicting observed species richness to the one for the Atlantic (intercept: 0.01, slope: 0.81 $r^2 = 0.42$, Online Appendix 2, Fig. A2.2), and overpredictions of observed species richness (69% of cells) were more common than underpredictions (31% of cells). The stacked species distribution model for 2075 predicted the average number of invasive species present per cell to increase from 4.6 to 6.4 (± 1.5 ; range 2.8–10. 6;

projection (c), and the predicted change in number of invaders according to SSDMs (d)

Fig. 2c). The change in predicted richness (1.9 ± 1.4 ; range -5.8 to 6.2, Fig. 2d) was positive in 90% of the grid cells in our study area, but there were a few small areas in the southern Strait of Juan de Fuca, Padilla Bay, Skagit Bay, Samish Bay, and Bellingham Bay where the expected number of these invaders was predicted to fall. The centre of gravity of our SSDM estimates shifted 95 km to the northwest, from (50.56 N, -126.61 W) to (51.14 N, -127.60 W).

Hotspots and coldspots

Our northwest Atlantic analysis using species richness thresholds identified hotspots through much of the area between Cape Cod and southern Maine, in Passamaquoddy Bay, and in several bays in southwestern

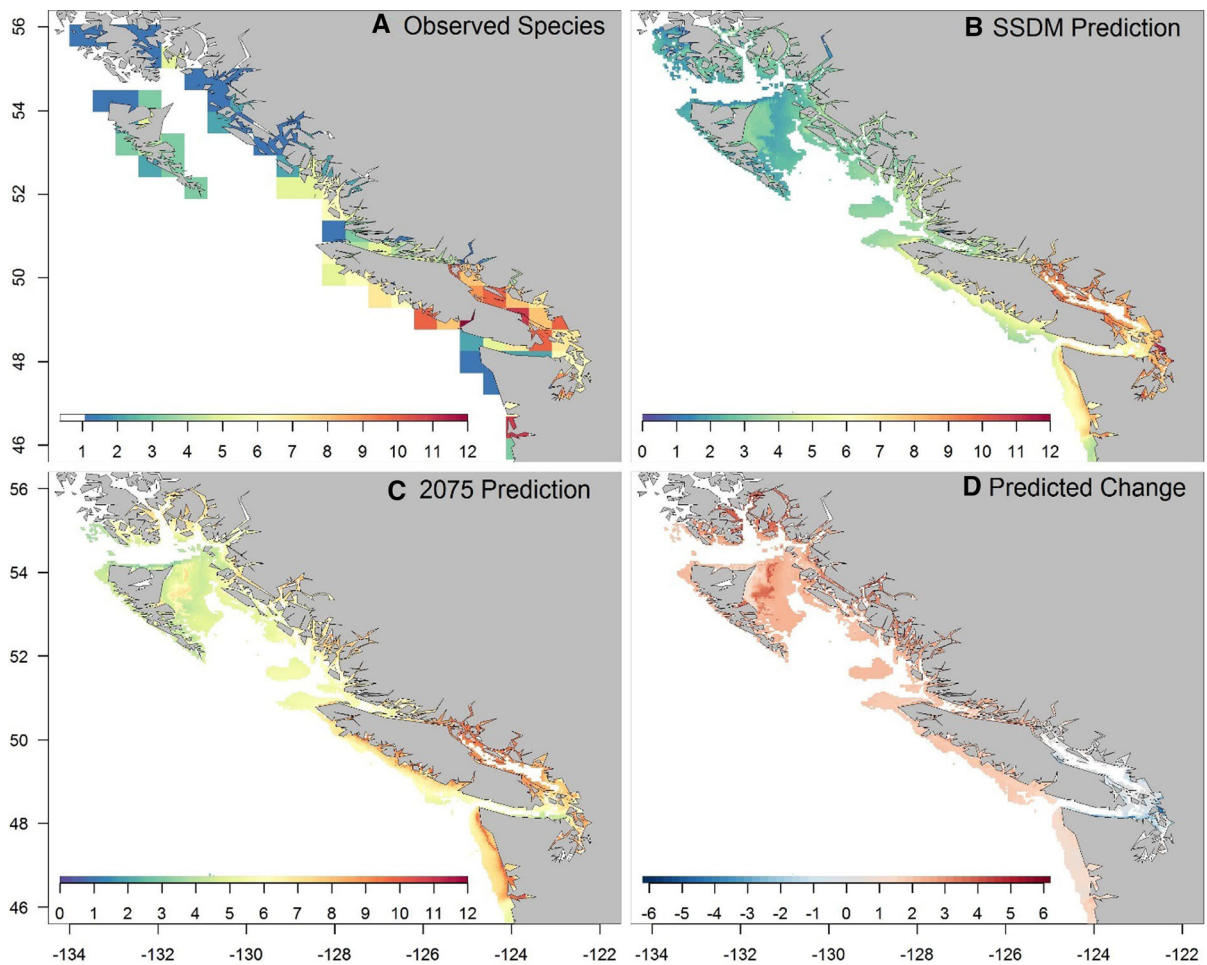


Fig. 2 Number of focal invasive species in the northeast Pacific according to observations (a), present day stacked species distribution model (SSDM) (b), and 2075 SSDM projection (c), and the predicted change in number of invaders according to

SSDMs (d). Note that there are slight differences in the coverage of present day and 2075 environmental datasets and, thus, the coverage of associated model predictions. Predicted change is calculated only for areas with data for both time periods

Nova Scotia for the present day (Fig. 3a). Coldspots were confined to the northern Gulf of St. Lawrence, around Newfoundland, southeast of Cape Cod, and small shallow areas far offshore. For 2075, there was little change in the overall pattern of hotspots and coldspots, apart from slight expansion of existing hotspots and contraction of existing coldspots (Fig. 3b). In the Pacific, much of the southern Salish Sea was identified as present-day hotspot, as were a few areas along the outer coast of Vancouver Island, Washington and Oregon (Fig. 3c). Large areas around Haida Gwaii, northern BC and southern Alaska were identified as coldspots. Our analysis predicts that coldspots will be almost completely eliminated by 2075, replaced by large areas of intermediate invasive

species richness (warmspots), and hotspots will expand to cover many areas on the western coast of Vancouver Island, the central coast of British Columbia, and most of the outer coast of Washington (Fig. 3d). Conversely, hotspots areas are predicted to disappear from parts of the southern Salish Sea, replaced by warmspots.

Using deciles to delineate hotspots at the ecoregion level, we found similar present-day hotspots in the same areas of Scotian Shelf as we identified in the broader-scale species richness-based analysis (Fig. 4a). Coldspots were identified in offshore areas around Sable Island, south of Yarmouth, and in the inner Bay of Fundy. These patterns changed only slightly for the 2075 projections, with some hotspots

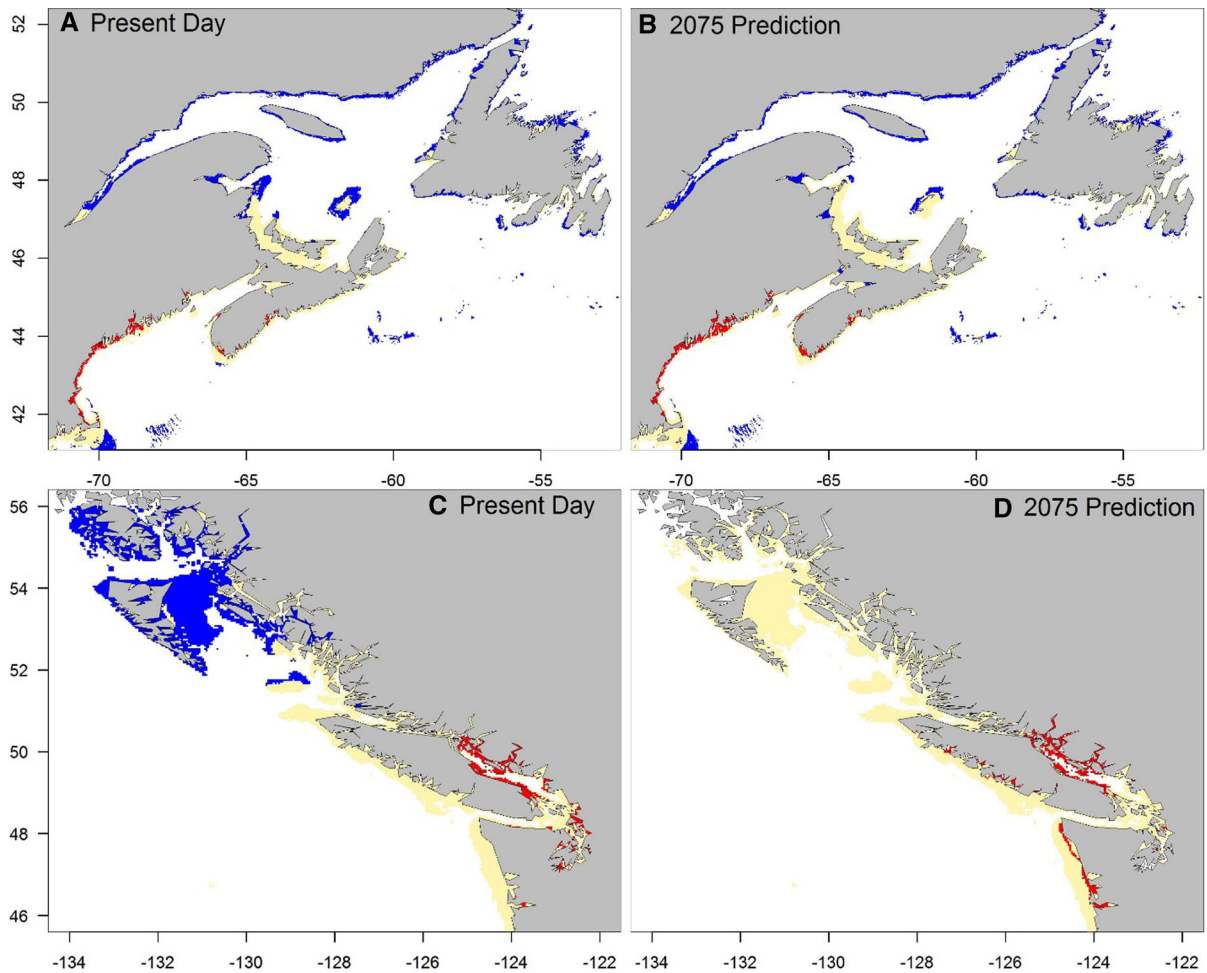


Fig. 3 Location of present and future hot and cold spots in the northwest Atlantic (a, b) and northeast Pacific (c, d). Hot (red) and cold (blue) spots identified as areas with predicted species

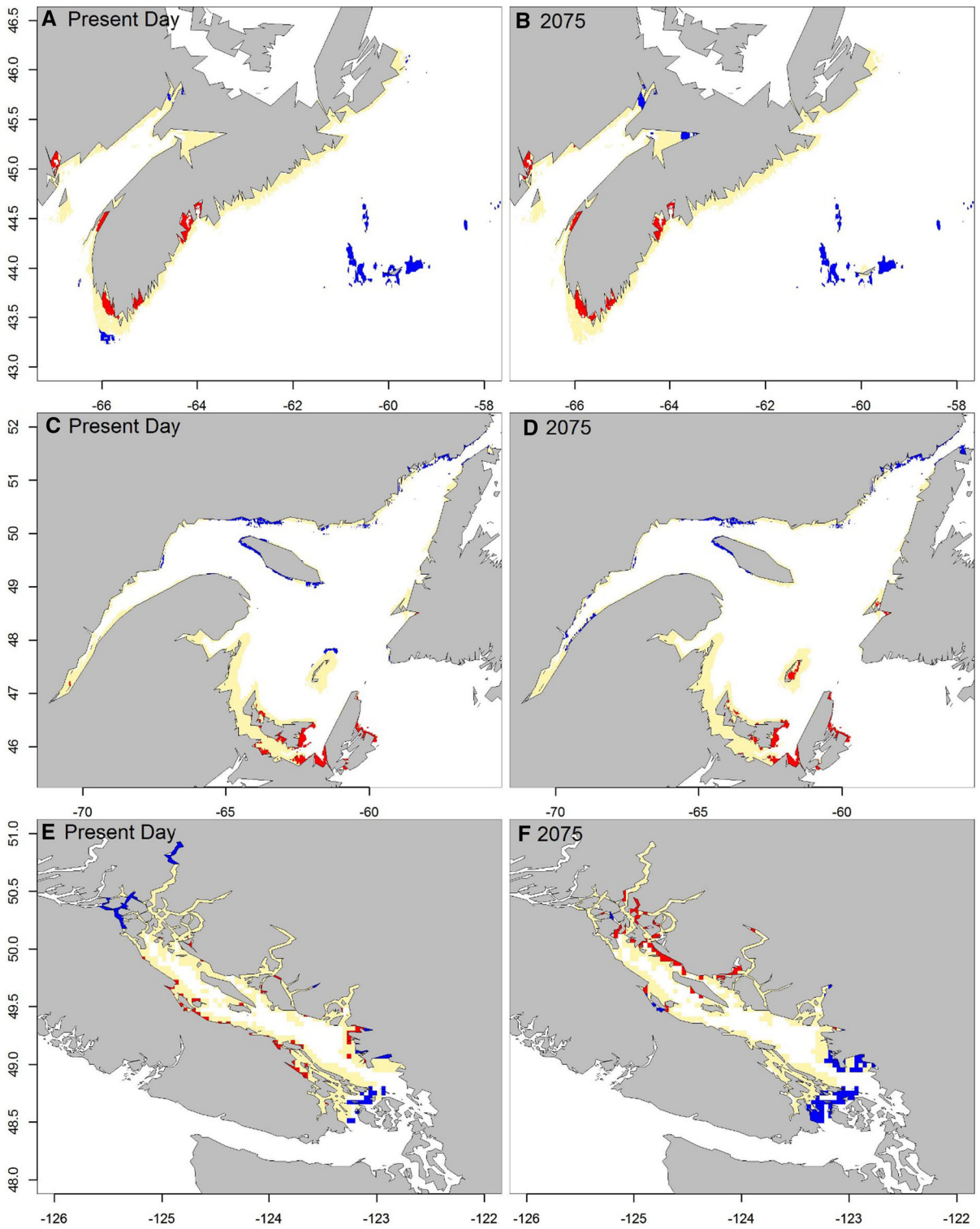
richness significantly higher than 8 and lower than 4, respectively. Other areas (i.e. “warm spots”) are depicted in yellow

and coldspots increasing or decreasing in size (Fig. 4b). However, new coldspots were predicted to develop in Minas Basin. In the Gulf of St. Lawrence ecoregion, hotspots were concentrated around the eastern Northumberland Strait and northern Cape Breton Island, with coldspots spread around the northern Gulf (Fig. 4c). For 2075, new hotspots are predicted to occur in Port au Port Bay, Newfoundland and a formerly very small hotspot in St. Georges Bay, Newfoundland is predicted to expand considerably (Fig. 4d). A coldspot north of the Magdalen Islands is predicted to disappear, while new coldspot areas are predicted to occur in the Saint Lawrence Estuary. In the Strait of Georgia, hotspots are concentrated in the middle part of the strait on the Vancouver Island and

mainland coastlines, with coldspots to the southeast and northwest (Fig. 4e). For 2075, hotspot areas are largely predicted to move to the northwest while coldspots are predicted to shrink in the northwest and expand in the southeast (Fig. 4e).

Discussion

Hotspots of invasion risk based on the history of invasions and the predicted distributions of invasive species within an area can help to identify areas where invasive species might pose the greatest ecological and socioeconomic threats, and where limited resources available for prevention, early detection,



◀ **Fig. 4** Location of present and future hot and cold spots in the Scotian Shelf (a, b), Gulf of St. Lawrence (c, d), and Strait of Georgia (e, f) ecoregions. Hot (red) and cold (blue) spots identified as areas with top and bottom 10% of predicted species richness within the ecoregion, respectively. Other areas that are neither hotspot nor coldspot (i.e. “warmspots”) are depicted in yellow

and management should be focused (Kulhanek et al. 2011; O’Donnell et al. 2012; Bellard et al. 2014; Li et al. 2016). There is a need to assess where invasive species are most likely to establish since it is critical for marine ecosystem management, but a lack of this information prevents invasive species from being effectively integrated into biosecurity monitoring programs and conservation plans (Giakoumi et al. 2016). Previous research has used species distribution models to examine the current and potential future distributions of individual marine invasive species (e.g. de Rivera et al. 2011; Lowen et al. 2016), or closely related guilds of invasive species that exhibit differences in their temperature and salinity tolerances that translate to differences in their potential to spread across cold to warm temperate and sub-arctic environments (Lowen et al. 2017). We advanced this approach by using species distribution models to identify areas that are at a high risk of invasion from a larger, more diverse suite of moderate to high risk invasive species with well known invasion histories under present-day and future climatological conditions. Results based on this methodology provide a broader, more integrative assessment of the spatial patterns of the threat posed by marine invasive species in the coastal waters of the northwest Atlantic and northeast Pacific.

Spatial patterns in invasion risk

Previous work focusing exclusively on invasive tunicates predicted moderate to high numbers of invasive species to occur in most coastal areas of SW New Brunswick, the Atlantic coast of Nova Scotia, and in the southern Gulf of St. Lawrence, and few species to occur in most areas around Newfoundland and the northern Gulf of St. Lawrence (Lowen et al. 2017). Although we focused on a broader suite of invasive species, and used slightly different environmental and occurrence different datasets, our results

were broadly similar. Focusing on the full extent of our study areas, we found that the predicted number of invasive species generally decreased with latitude, such that present day invasion hotspots were concentrated within the southern portion of our focal area and coldspots were generally concentrated in the north. This pattern is related to our choice of species, which are thought to have been initially introduced to the south of our focal areas, or towards their southern end, and subsequently spread northward (Fofonoff et al. 2018). Moreover, conditions in our focal area also tend to be towards the cooler end of the thermal tolerances of these species, except *Membranipora membranacea* (Fofonoff et al. 2018). However, we are not aware of marine invasive species thought to have been initially introduced to North America in the northern waters. Several lines of evidence support the existence of a latitudinal gradient in marine invasive species richness, which mirrors the general latitudinal diversity gradient. The existence of a global latitudinal gradient in coastal marine invertebrate and plant biodiversity (Worm and Tittensor 2018) mean that the pool of potential invasive species declines with latitude. Although invasive species richness is often negatively correlated with native biodiversity in experiments, and at very fine spatial scales, the opposite is typically true at coarser spatial scales, and latitudinal gradients of invasive species richness have been observed in other taxa (Fridley et al. 2007; Sax 2001). Moreover, the U.S. has a higher concentration of ports, and thus, greater exposure to shipping as a vector of introduction and spread. Thus, the southern end of our study areas have historically experienced a higher likelihood and number of introductions of marine invasive species. Finally, southern portions of the study areas also have a better climate match to European and Asian source regions than those in the north.

Widely cited global analyses have used the intensity of commercial vessel traffic in major ports (e.g. Drake and Lodge 2004), or vessel traffic and environmental conditions within major ports (e.g. Seebens et al. 2013, 2016) to assess regional-scale invasion risk. Neither patterns of observed marine invasive species richness, nor predicted hotspots, closely matched the locations of major ports within our study area (Figs. 1, 2, 3). Our results suggest that conditions in large ports do not necessarily reflect regional-scale invasion risk or experience the highest risk of invasion, despite high levels of vessel activity. Although

the ports of Boston, Seattle, and Vancouver overlap with hotspots, they have lower observed and/or predicted invasive species richness (for our focal species) than nearby non-port areas, as do other large ports that fall outside hotspots [e.g. Halifax, Saint John, St. John's, Prince Rupert, (see Figure A1.1 and A1.2 in Online Appendix for maps of port locations)]. This underscores the importance of considering where species can spread via natural dispersal following initial introduction, as well as the influence of other vectors of initial introduction and spread, such as aquaculture (Naylor et al. 2001), recreational vessels (Clarke Murray et al. 2011), oil platforms (Pajuelo et al. 2016), and the trade of live organisms for food, bait, education, research, and public and private aquaria (Weigle et al. 2005) when assessing local and regional-scale invasion risk.

Our results suggest that projected change in invasive species richness by 2075 will be far more dramatic in the northeast Pacific than the northwest Atlantic, with many species currently confined to warmer southern waters expected to move north, thus reducing or eliminating present day coldspots concentrated on the northeast Pacific Coast (Figs. 2, 3). Whereas most Pacific species expanded their range to the north in our future projections, there is less change in the Atlantic (Online Appendix 3). This may be due to differences between the models used to project the future climate scenario, including the magnitude of the projected changes in temperature and salinity, which are larger in the Pacific than the Atlantic (Brickman et al. 2016; Peña et al. 2018), or to a larger number of species expected cross thermal thresholds under future scenarios in the Pacific.

Applications of invasion hotspot maps

Different scientific, management, and communication objectives require different information, and should guide how one chooses to define invasion hotspots. Maps of predicted invasive species richness provide the best depiction of our results, but discretized hotspot maps help to simplify communication and focus attention. In some cases, there may be thresholds in the number of invaders that we are, or are not, concerned about. In other cases, resource limitations might constrain the number of sites or total area we can monitor, manage, or mitigate, so prioritizing hotspots (i.e., according to quantiles) make sense. We

focus on applications of hotspots and coldspots because of the frequent need to prioritize (or deprioritize) specific areas for attention. However, this does not mean that areas of intermediate invasion risk should be ignored. Moreover, the more detailed species richness maps and/or the full coldspot-warmspot-hotspot rankings can be used when decision makers have greater capacity to deal with complexity or to allocate resources more broadly.

Marine invasion hotspots and coldspots are most directly applicable in invasive species monitoring and management. Guisan et al. (2013) discuss how species distribution models can be used in all stages of a structured decision-making process for dealing with biological invasions, and the principles they discuss apply equally to our stacked species distribution models and hotspot maps. First, the models can help to identify potential problems, such as increasing invasion risk in an area due to climate change, and help to provide a frame of reference to define objectives and possible actions. By indicating where species are most likely to become established now and spread in the future, predictive models facilitate early detection through optimization of effective, cost-efficient monitoring programs (Honrado et al. 2016), and can allow one to anticipate and plan regulatory or management interventions to prevent, mitigate or eradicate invasions. Then, the models can help to assess how effective different actions might be and trade-offs between the costs and benefits of different options prior to a final decision (e.g. Baxter and Possingham 2011). Finally, error associated with model estimates contribute to uncertainty assessments throughout the decision-making process.

Marine invasive species intersect with other conservation and marine management objectives through their vectors of introduction and dispersal, and through their ecological, economic, and social/cultural effects in affected areas. Vessel traffic (ballast water, hull fouling) and aquaculture are considered the primary vectors responsible for the introduction and spread of marine invasive species (Williams et al. 2013), including all the species in our study. In recent years, the threat posed by introduced and invasive species has been reduced though regulatory measures and voluntary codes of practice, including ballast water management (Scriven et al. 2015), and control of introductions and transfers of aquaculture species, and movement of aquaculture gear (ICES 2005). However,

a degree of threat remains. Rates of compliance with existing regulations are less than 100%, and many vessels (e.g. recreational vessels, vessels moving within or between regions within national boundaries) and aquaculture activities are not yet subject to regulations (Castro et al. 2018). Moreover, hull fouling is a major component of the risk posed by vessel traffic but, unlike ballast water, it is largely unregulated (Sylvester et al. 2011). Hotspot maps depicting areas that are at high and low risk of invasion by multiple species could be used to focus new and existing regulatory measures in ways that minimize the risk of introduction and spread of invasive species by these vectors and the regulatory burden placed on both industries.

Examining how invasion hotspots intersect with areas of conservation concern (e.g. habitats of threatened species, marine protected areas, biodiversity hotspots) or areas providing important economic and ecosystem service benefits (e.g. ports, marinas, beaches) is another way to focus biosecurity measures and guide decisions about conservation and marine spatial planning. For instance, in addition to being a vector of marine invasive species, the aquaculture industry suffers some of their most significant economic impacts, through bio-fouling of aquaculture organisms and gear, as well as losses due to competition and predation by marine invasive species (Beal and Kraus 2002; Carver et al. 2003). Maps of invasion hotspots could be used to assess the current and future threat to existing aquaculture operations and potentially influence where new aquaculture leases are located. Aquaculture operations in the southern Gulf of St. Lawrence have been strongly affected by marine invasions (Locke et al. 2009), but this area has only moderate invasive species richness relative to the entire northeast Atlantic study region, and no hotspots were identified for this area in our broadest scale analyses (Figs. 2, 3). However, defining hotspots according to regionally specific thresholds reveals that hotspots within the Gulf of St. Lawrence ecoregion are concentrated in the southern Gulf, and that a high proportion of aquaculture leases are in current or projected invasion hotspots (Figs. 4, 5). This demonstrates the value of setting regionally and/or objective-specific criteria to define and delineate hotspots and illustrates how potentially serious and significant impacts on ecosystem goods and services from invasive species exist outside the broadest-scale hotspots.

Our analyses provide information about invasion risk, in terms of the identity and number of invasive species that are likely to become established in an area. However, in many management applications, risk of impacts may be more important than risk of invasion. Estimating cumulative invasion impacts requires information about the abundance of each invasive species, as well as knowledge of the ecological and human systems present in an area, and how particular invasive species affect them. If such information cannot be obtained, it may sometimes be possible to use cumulative invasion risk as a proxy for cumulative risk of impact, but only with great caution.

Limitations

Our analysis of invasive species richness and invasion risk hotspots is subject to the same limitations as other analyses using species distribution models. All estimates provided by species distribution models are subject to errors because of data deficiencies and uncertainty introduced through imperfect model specification (Barry and Elith 2006). Predictions will be imperfect because occurrence data used to fit the models is incomplete and may be biased towards areas where observations are more likely; long term climatologies will not capture intra-annual variability's effects on species distributions; even the best high resolution climate data will not perfectly match the environmental conditions organisms actually experience; and species distribution models never capture all the factors that control a species' distribution (Rondinini et al. 2006).

The differences between observed species richness and model-based estimates can be informative. These differences can be caused by shortcomings of either the SDMs or the observation data (e.g. underestimates due to insufficient sampling). Where model predictions are higher than observed species richness three scenarios are worth investigating. First, some invasion(s) may have gone undetected due to insufficient sampling. Second, although our focal species are likely to have had many opportunities to spread to representative (in terms of environmental suitability) habitats in the 20-plus years since each invasion began, they may not have reached or established in all suitable areas. In suitable areas where species are not yet established there may be an opportunity for preventative management. Third, there may be a

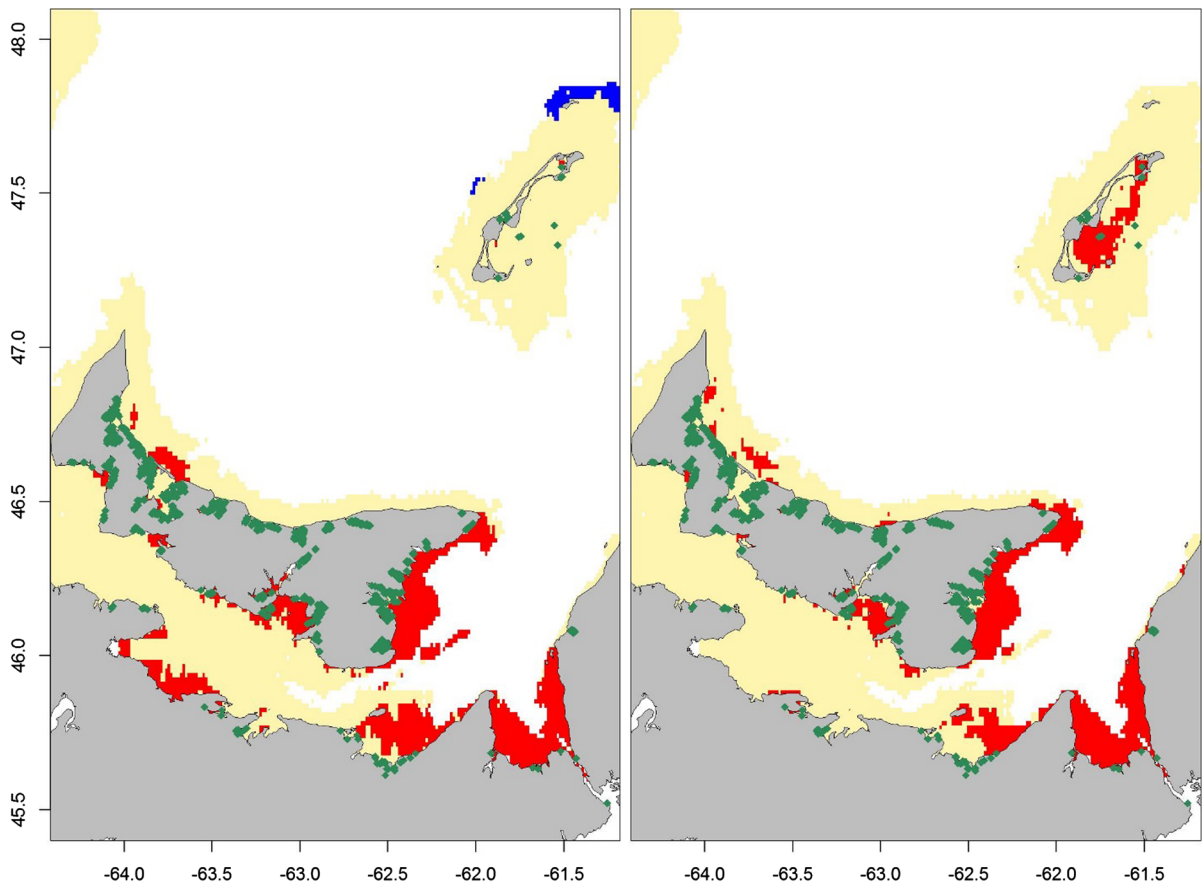


Fig. 5 Existing aquaculture sites (green diamonds) in relation to present day and future invasion risk hotspots and coldspots in the southern Gulf of St. Lawrence, demonstrating a projected increase in the severity of invasions near aquaculture sites in the Magdalen Islands. Hotspots (red) and coldspots (blue) are

delineated as top and bottom 10% of predicted invasive species richness in the Gulf of St. Lawrence ecoregion (as in Fig. 4). Other areas that are neither hotspot nor coldspot (i.e. “warm-spots”) are depicted in yellow

limiting environmental factor or biotic interactions in that area that has not been captured in the model. Underpredictions of observed species richness were half as common as overpredictions in our results. Underpredictions could be indicative of source-sink dynamics, where repeated introductions or on-going dispersal (via natural or anthropogenic vectors) help to support populations in marginal habitats (Lockwood et al. 2005; Franklin 2010). Future research combining predictions of invasive species richness and information about currents and vector activity could help to identify areas where high dispersal maintains higher than predicted species richness, as well as environmentally suitable areas that go uncolonized due to dispersal limitation.

Including additional environmental parameters (e.g. nutrients, pH, etc.) might have further improved our models. We did not include additional predictor variables, either because we did not have access to environmental projections that would allow us to project our models, or because no appropriate present-day data exist. For example, most of the species we investigated rely on hard substrates (Fofonoff et al. 2018). To our knowledge, high resolution substrate data throughout our study areas do not exist. Without it, our models are likely to over-estimate invasive species richness in sandy or muddy areas, unless sufficient rocky outcrops or anthropogenic structures are present within each of the grids to which a probability is assigned. Thankfully, the accessibility of high resolution, historical and present-day data for

an increasingly wide array of environmental parameters that can be used to model the distributions of marine species is steadily increasing (e.g. BioORACLE v2.0, Assis et al. 2018). Acquiring projections for environmental parameters under future climate scenarios is more difficult, and remains an impediment to further improving predictions of the future distributions of invasive and native species.

Distribution models based on species' environmental responses typically do not account for biotic interactions, or how they might change in response to environmental change (Davis et al. 1998), and simple stacking of distribution models involves an implicit assumption that species occurrence probabilities are independent after accounting for the effects of the environmental predictors (Calabrese et al. 2014). Nor do species distribution models typically capture the influence of genetic variability, phenotypic plasticity, or evolutionary changes (Elith and Leathwick 2009). Accounting for genetic population structure, such that all local populations are not assumed to have the same physiological responses and environmental tolerances, significantly improves model predictions (Lowen et al. 2019). Thus, a lack of information about the plasticity, genetic structure, and evolutionary responses of populations, along with the uncertainty associated with climate projections, will constrain the predictability of species' future distributions.

Predicting species distributions in the context of biological invasions and climate change can cause additional difficulties because the models assume that species are at equilibrium with their environment and that the relevant environmental gradients have been well sampled, yet such contexts can involve predictions for novel and non-equilibrium conditions (Elith and Leathwick 2009). We have tried to reduce the challenges posed by invasions by selecting species that were well established and have already spread over broad spatial and environmental ranges, but it is not possible to know if they have all reached equilibrium.

Despite these limitations, correlative species distribution models remain one of the few practical approaches to predicting past, current, and future species distributions, and their ability to use incomplete data to produce spatially comprehensive estimates for the present day and future distributions is a major advantage (Elith and Leathwick 2009; Rondinini et al. 2006). With today's high-resolution

oceanographic data and models, these estimates can be made for a broad range of spatial extents and resolutions that are useful in the context of management applications. These estimates are likely to improve as data for more factors that control species distributions' become available and existing oceanographic models are refined further. Finally, like any analysis of cumulative invasion risk or impact, our findings are also limited by the set of species included in the analysis. Although we chose to select invasive species covering a broad range of taxonomic and life history diversity in hopes of making our results more representative, they may not reflect the patterns of the species we could not include due to lack of data or logistical constraints.

Local adaptation can cause distribution models for different populations and genetic lineages to differ markedly from one another, and for combined models to differ markedly from models based on the occurrence data for the entire species (Hällfors et al. 2016; Lowen et al. 2019). Our study uses occurrence data from within each of our focal regions to model the species distributions within each of those regions. We chose not to include occurrence data from other parts of the species' global distributions so that our models would reflect the potential distributions and environmental responses of the genetic lineages that have established populations in each our focal regions. However, this means that our models may not capture the entire range each species could potentially achieve if secondary introductions bring additional genetic lineages with different responses to temperature, salinity, and wave action to our shores.

Conclusions

Understanding, managing, and monitoring species invasions are complex, difficult tasks. There are many established and potential invaders to consider, each associated with environmental requirements, dispersal vectors and pathways, potential impacts and mitigation measures. Additionally, environmental change is altering the susceptibility of native ecosystems to invasion (Stachowicz et al. 2002; Bellard et al. 2018). Tools that synthesize across this complexity are needed to clarify our thinking, to identify where the cumulative risks (and thus potential impacts) of invasive species are greatest, and to shape

management responses. By stacking species distribution models for 12 invasive species in the northwest Atlantic and northeast Pacific to identify invasion hotspots, this study demonstrates that there is significant spatial heterogeneity in cumulative invasion risk in the coastal northwest Atlantic and northeast Pacific, and that these patterns are likely to change over the next 55 years in response to climate change, particularly in the Pacific. Stacked species distribution models and the hotspot maps they generate provide an integrated perspective that can help focus attention and provide guidance to those tasked with generating new scientific knowledge, prioritizing locations for monitoring and implementing policy around marine invasive species, typically with limited resources.

Acknowledgements We thank the editor, two anonymous reviewers, P. L. Thompson and A. M. Gehman for comments on a previous version of this manuscript. We also thank M. Sizer for help with occurrence data, and M. E. Auger-Méthé and R. R. E. Stanley for coding tips. This research was supported by funding from Fisheries and Oceans Canada's Strategic Program for Ecosystem-Based Research and Advice (SPERA).

References

- Aarts G, Fieberg J, Matthiopoulos J (2012) Comparative interpretation of count, presence-absence and point methods for species distribution models. *Methods Ecol Evol* 3:177–187. <https://doi.org/10.1111/j.2041-210X.2011.00141.x>
- Arenas F, Bishop JDD, Carlton JT et al (2006) Alien species and other notable records from a rapid assessment survey of marinas on the south coast of England. *J Mar Biol Assoc UK* 86:1329–1337. <https://doi.org/10.1017/S0025315406014354>
- Assis J, Tyberghein L, Bosch S et al (2018) Bio-ORACLE v2.0: extending marine data layers for bioclimatic modelling. *Glob Ecol Biogeogr* 27:227–284. <https://doi.org/10.1111/geb.12693>
- Ball IR, Possingham HP, Watts ME (2009) Marxan and relatives: software for spatial conservation prioritization. In: Moilanen A, Wilson KA, Possingham HP (eds) *Spatial conservation prioritisation: quantitative methods and computational tools*. Oxford University Press, Oxford, pp 185–195
- Barry S, Elith J (2006) Error and uncertainty in habitat models. *J Appl Ecol* 43:413–423. <https://doi.org/10.1111/j.1365-2664.2006.01136.x>
- Baxter PWJ, Possingham HP (2011) Optimizing search strategies for invasive pests: learn before you leap. *J Appl Ecol* 48:86–95. <https://doi.org/10.1111/j.1365-2664.2010.01893.x>
- Bayliss H, Stewart G, Wilcox A, Randall N (2013) A perceived gap between invasive species research and stakeholder priorities. *NeoBiota* 19:67–82. <https://doi.org/10.3897/neobiota.19.4897>
- Beal BF, Kraus MG (2002) Interactive effects of initial size, stocking density, and type of predator deterrent netting on survival and growth of cultured juveniles of the soft-shell clam, *Mya arenaria* L., in eastern Maine. *Aquaculture* 208:81–111. [https://doi.org/10.1016/S0044-8486\(01\)00900-0](https://doi.org/10.1016/S0044-8486(01)00900-0)
- Bellard C, Leclerc C, Leroy B et al (2014) Vulnerability of biodiversity hotspots to global change. *Glob Ecol Biogeogr* 23:1376–1386. <https://doi.org/10.1111/geb.12228>
- Bellard C, Jeschke JM, Leroy B, Mace GM (2018) Insights from modeling studies on how climate change affects invasive alien species geography. *Ecol Evol* 8:5688–5700. <https://doi.org/10.1002/ece3.4098>
- Brickman D, Wang Z, DeTracey B (2016) High resolution future climate ocean model simulations for the northwest Atlantic shelf region. *Can Tech Rep Hydrogr Ocean Sci* 315:1–143
- Brown JL, Bennett JR, French C (2017) SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ*. <https://doi.org/10.7717/peerj.4095>
- Burrows MT, Harvey R, Robb L (2008) Wave exposure indices from digital coastlines and the prediction of rocky shore community structure. *Mar Ecol Prog Ser* 353:1–12. <https://doi.org/10.3354/meps07284>
- Calabrese JM, Certain G, Kraan C, Dormann CF (2014) Stacking species distribution models and adjusting bias by linking them to macroecological models. *Glob Ecol Biogeogr* 23:99–112. <https://doi.org/10.1111/geb.12102>
- Carman MR, Morris JA, Karney RC, Grunden DW (2010) An initial assessment of native and invasive tunicates in shellfish aquaculture of the North American east coast. *J Appl Ichthyol* 26:8–11. <https://doi.org/10.1111/j.1439-0426.2010.01495.x>
- Carver CE, Chisholm A, Mallet AL (2003) Strategies to mitigate the impact of *Ciona intestinalis* (L.) biofouling on shellfish production. *J Shellfish Res* 22:621–631
- Castro MCT, Hall-Spencer JM, Poggian CF, Fileman TW (2018) Ten years of Brazilian ballast water management. *J Sea Res* 133:36–42. <https://doi.org/10.1016/j.seares.2017.02.003>
- Cheng L, Abraham J, Hausfather Z, Trenberth KE (2019) How fast are the oceans warming? *Science* 363:128–129. <https://doi.org/10.1126/science.aav7619>
- Cheung WWL (2019) Predicting the future ocean: pathways to global ocean sustainability. In: Cisneros-Montemayor AM, Cheung WWL, Ota Y (eds) *Predicting future oceans: sustainability of ocean and human systems amidst global environmental change*. Elsevier, Amsterdam, pp 3–15
- Clarke Murray C, Pakhomov EA, Theriault TW (2011) Recreational boating: a large unregulated vector transporting marine invasive species. *Divers Distrib* 17:1161–1172. <https://doi.org/10.1111/j.1472-4642.2011.00798.x>
- Cohen AN, Carlton JT (1998) Accelerating invasion rate in a highly invaded estuary. *Science* (80-) 279:555–558. <https://doi.org/10.1126/science.279.5350.555>

- Davis AJ, Jenkinson LS, Lawton JH et al (1998) Making mistakes when predicting shifts in species range in response to global warming. *Nature* 391:783–786. <https://doi.org/10.1038/35842>
- De Rivera CE, Steves BP, Fofonoff PW et al (2011) Potential for high-latitude marine invasions along western North America. *Divers Distrib* 17:1198–1209. <https://doi.org/10.1111/j.1472-4642.2011.00790.x>
- Drake JM, Lodge DM (2004) Global hot spots of biological invasions: evaluating options for ballast-water management. *Proc R Soc B Biol Sci* 271:575–580. <https://doi.org/10.1098/rspb.2003.2629>
- Drolet D, DiBacco C, Locke A et al (2016) Evaluation of a new screening-level risk assessment tool applied to non-indigenous marine invertebrates in Canadian coastal waters. *Biol Invasions* 18:279–294. <https://doi.org/10.1007/s10530-015-1008-y>
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst* 40:677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Elith J, Phillips SJ, Hastie T et al (2011) A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17:43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Esler KJ, Prozesky H, Sharma GP, McGeoch M (2010) How wide is the “knowing-doing” gap in invasion biology? *Biol Invasions* 12:4065–4075. <https://doi.org/10.1007/s10530-010-9812-x>
- Fofonoff PW, Ruiz GM, Steves B, Simkanin C, Carlton J (2018) National exotic marine and estuarine species information system. <https://invasions.si.edu/nemesis/>. Accessed 14 Mar 2018
- Franklin J (2010) Mapping species distributions: spatial inference and prediction ecology, biodiversity and conservation ecology, biodiversity and conservation. Cambridge University Press, Cambridge
- Fridley JD, Stachowicz JJ, Naeem S et al (2007) The invasion paradox: reconciling pattern and process in species invasions. *Ecology* 88:3–17. [https://doi.org/10.1890/0012-9658\(2007\)88\[3:TIPRPA\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2007)88[3:TIPRPA]2.0.CO;2)
- Getis A, Ord JK (1992) The analysis of spatial association by use of distance statistics. *Geogr Anal* 24:189–206. <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>
- Giakoumi S, Guilhaumon F, Kark S et al (2016) Space invaders; biological invasions in marine conservation planning. *Divers Distrib* 22:1220–1231. <https://doi.org/10.1111/ddi.12491>
- Gormley AM, Forsyth DM, Griffioen P et al (2011) Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species. *J Appl Ecol* 48:25–34. <https://doi.org/10.1111/j.1365-2664.2010.01911.x>
- Guisan A, Tingley R, Baumgartner JB et al (2013) Predicting species distributions for conservation decisions. *Ecol Lett* 16:1424–1435. <https://doi.org/10.1111/ele.12189>
- Guo L, Sheng J (2017) Impacts of climate changes on ocean surface gravity waves over the eastern Canadian shelf. *Ocean Dyn* 67:621–637. <https://doi.org/10.1007/s10236-017-1046-3>
- Hällfors MH, Liao J, Dzurisin J et al (2016) Addressing potential local adaptation in species distribution models: implications for conservation under climate change. *Ecol Appl* 26:1154–1169. <https://doi.org/10.1890/15-0926>
- Honrado JP, Pereira HM, Guisan A (2016) Fostering integration between biodiversity monitoring and modelling. *J Appl Ecol* 53:1299–1304. <https://doi.org/10.1111/1365-2664.12777>
- International Council for the Exploration of the Sea (2005) ICES code of practice on the introductions and transfers of marine organisms 2005
- Jeschke JM, Strayer DL (2008) Usefulness of bioclimatic models for studying climate change and invasive species. *Ann NY Acad Sci* 1134:1–24
- Johnson LE, Bossenbroek JM, Kraft CE (2006) Patterns and pathways in the post-establishment spread of non-indigenous aquatic species: the slowing invasion of North American inland lakes by the zebra mussel. *Biol Invasions* 8:475–489. <https://doi.org/10.1007/s10530-005-6412-2>
- Kramer-Schadt S, Niedballa J, Pilgrim JD et al (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. *Divers Distrib* 19:1366–1379. <https://doi.org/10.1111/ddi.12096>
- Kulhanek SA, Leung B, Ricciardi A (2011) Using ecological niche models to predict the abundance and impact of invasive species: application to the common carp. *Ecol Appl* 21:203–213. <https://doi.org/10.1890/09-1639.1>
- Li H, Wu J (2006) Uncertainty analysis in ecological studies: an overview. In: Wu J, Jones KB, Li H, Loucks OL (eds) *Scaling and uncertainty analysis in ecology: methods and applications*. Springer, Dordrecht, pp 45–66
- Li X, Liu X, Kraus F et al (2016) Risk of biological invasions is concentrated in biodiversity hotspots. *Front Ecol Environ* 14:411–417. <https://doi.org/10.1002/fee.1321>
- Locke A, Doe KG, Fairchild WL et al (2009) Preliminary evaluation of effects of invasive tunicate management with acetic acid and calcium hydroxide on non-target marine organisms in Prince Edward Island, Canada. *Aquat Invasions* 4:221–236. <https://doi.org/10.3391/ai.2009.4.1.23>
- Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. *Trends Ecol Evol* 20:223–228. <https://doi.org/10.1016/j.tree.2005.02.004>
- Lowen JB, McKindsey CW, Therriault TW, DiBacco C (2016) Effects of spatial resolution on predicting the distribution of aquatic invasive species in nearshore marine environments. *Mar Ecol Prog Ser* 556:17–30. <https://doi.org/10.3354/meps11765>
- Lowen JB, DiBacco C, Lowen JB et al (2017) Distributional changes in a guild of non-indigenous tunicates in the NW Atlantic under high-resolution climate projections. *Mar Ecol Prog Ser* 570:173–186. <https://doi.org/10.3354/meps12077>
- Lowen JB, Hart DR, Stanley RRE et al (2019) Assessing effects of genetic, environmental, and biotic gradients in species distribution modelling. *ICES J Mar Sci*. <https://doi.org/10.1093/icesjms/fsz049>
- Mačić V, Albano PG, Alpanidou V et al (2018) Biological invasions in conservation planning: a global systematic review. *Front Mar Sci*. <https://doi.org/10.3389/fmars.2018.00178>
- Marchini A, Ferrario J, Sfriso A, Occhipinti-Ambrogi A (2015) Current status and trends of biological invasions in the

- Lagoon of Venice, a hotspot of marine NIS introductions in the Mediterranean Sea. *Biol Invasions* 17:2943–2962. <https://doi.org/10.1007/s10530-015-0922-3>
- Matzek V, Covino J, Funk JL, Saunders M (2014) Closing the knowing-doing gap in invasive plant management: accessibility and interdisciplinarity of scientific research. *Conserv Lett* 7:208–215. <https://doi.org/10.1111/conl.12042>
- Mesgaran MB, Cousens RD, Webber BL (2014) Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Divers Distrib* 20:1147–1159. <https://doi.org/10.1111/ddi.12209>
- Muscarella R, Galante PJ, Soley-Guardia M et al (2014) ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods Ecol Evol* 5:1198–1205. <https://doi.org/10.1111/2041-210x.12261>
- Naylor RL, Williams SL, Strong DR (2001) Aquaculture—a gateway for exotic species. *Science* 294:1655–1656. <https://doi.org/10.1126/science.1064875>
- O'Donnell J, Gallagher RV, Wilson PD et al (2012) Invasion hotspots for non-native plants in Australia under current and future climates. *Glob Chang Biol* 18:617–629. <https://doi.org/10.1111/j.1365-2486.2011.02537.x>
- OBIS (2018) Ocean biogeographic information system. Intergovernmental Oceanographic Commission of UNESCO. www.iobis.org. Accessed 17 Mar 2017
- Pajuelo JG, González JA, Triay-Portella R et al (2016) Introduction of non-native marine fish species to the Canary Islands waters through oil platforms as vectors. *J Mar Syst* 163:23–30. <https://doi.org/10.1016/j.jmarsys.2016.06.008>
- Pejchar L, Mooney HA (2009) Invasive species, ecosystem services and human well-being. *Trends Ecol Evol* 24:497–504
- Peña A, Fine I, Masson D (2018) Towards climate change projections of biogeochemical conditions along the British Columbia coast. In: Jang CJ, Curchitser E (eds) Report of working group 29 on regional climate modeling. PICES scientific report No. 54, pp 114–124
- Peña MA, Fine I, Callendar W (2019) Interannual variability in primary production and shelf-offshore transport of nutrients along the northeast Pacific Ocean margin. *Deep Sea Res Part II Top Stud Oceanogr*. <https://doi.org/10.1016/j.dsr2.2019.104637>
- Phillips SJ, Dudík M, Elith J et al (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol Appl* 19:181–197. <https://doi.org/10.1890/07-2153.1>
- Phillips SJ, Anderson RP, Dudík M et al (2017) Opening the black box: an open-source release of Maxent. *Ecography (Cop)* 40:887–893. <https://doi.org/10.1111/ecog.03049>
- Pinsky ML, Worm B, Fogarty MJ et al (2013) Marine taxa track local climate velocities. *Science* (80-) 341:1239LP–1242LP. <https://doi.org/10.1126/science.1239352>
- Pyšek P, Richardson DM (2010) Invasive species, environmental change and management, and health. *Annu Rev Environ Resour* 35:25–55. <https://doi.org/10.1146/annurev-environ-033009-095548>
- R Development Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Rascle N, Arduin F (2013) A global wave parameter database for geophysical applications. Part 2: model validation with improved source term parameterization. *Ocean Model* 70:174–188. <https://doi.org/10.1016/j.ocemod.2012.12.001>
- Renner IW, Warton DI (2013) Equivalence of MAXENT and poisson point process models for species distribution modeling in ecology. *Biometrics* 69:274–281. <https://doi.org/10.1111/j.1541-0420.2012.01824.x>
- Rondinini C, Wilson KA, Boitani L et al (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecol Lett* 9:1136–1145. <https://doi.org/10.1111/j.1461-0248.2006.00970.x>
- Ruiz GM, Fofonoff PW, Steves B et al (2011) Marine invasion history and vector analysis of California: a hotspot for western North America. *Divers Distrib* 17:362–373. <https://doi.org/10.1111/j.1472-4642.2011.00742.x>
- Sax DF (2001) Latitudinal gradients and geographic ranges of exotic species: implications for biogeography. *J Biogeogr* 28:139–150. <https://doi.org/10.1046/j.1365-2699.2001.00536.x>
- Sbrocco EJ, Barber PH (2013) MARSPEC: ocean climate layers for marine spatial ecology. *Ecology* 94:979. <https://doi.org/10.1890/12-1358.1>
- Scriven DR, DiBacco C, Locke A, Therriault TW (2015) Ballast water management in Canada: a historical perspective and implications for the future. *Mar Policy* 59:121–133. <https://doi.org/10.1016/j.marpol.2015.05.014>
- Seebens H, Gastner MT, Blasius B (2013) The risk of marine bioinvasion caused by global shipping – Supplementary information. *Ecol Lett* 16:782–790. <https://doi.org/10.1111/ele.12111>
- Seebens H, Schwartz N, Schupp PJ, Blasius B (2016) Predicting the spread of marine species introduced by global shipping. *Proc Natl Acad Sci* 113:5646–5651. <https://doi.org/10.1073/pnas.1524427113>
- Soontiens N, Allen SE (2017) Modelling sensitivities to mixing and advection in a sill-basin estuarine system. *Ocean Model* 112:17–32. <https://doi.org/10.1016/j.ocemod.2017.02.008>
- Soontiens N, Allen SE, Latonell D et al (2016) Storm surges in the Strait of Georgia simulated with a regional model. *Atmos Ocean* 54:1–21. <https://doi.org/10.1080/07055900.2015.1108899>
- Stachowicz JJ, Terwin JR, Whitlatch RB, Osman RW (2002) Linking climate change and biological invasions: ocean warming facilitates nonindigenous species invasions. *Proc Natl Acad Sci USA* 99:15497–15500. <https://doi.org/10.1073/pnas.242437499>
- Sylvester F, Kalaci O, Leung B et al (2011) Hull fouling as an invasion vector: can simple models explain a complex problem? *J Appl Ecol* 48:415–423. <https://doi.org/10.1111/j.1365-2664.2011.01957.x>
- Tidbury HJ, Taylor NGH, Copp GH et al (2016) Predicting and mapping the risk of introduction of marine non-indigenous species into Great Britain and Ireland. *Biol Invasions*. <https://doi.org/10.1007/s10530-016-1219-x>

- Tyberghein L, Verbruggen H, Pauly K et al (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Glob Ecol Biogeogr* 21:272–281. <https://doi.org/10.1111/j.1466-8238.2011.00656.x>
- Venette RC, Kriticos DJ, Magarey RD et al (2010) Pest risk maps for invasive alien species: a roadmap for improvement. *Bioscience* 60:349–362. <https://doi.org/10.1525/bio.2010.60.5.5>
- Vitousek PM, D'Antonio CM, Loope LL et al (1997) Introduced species: a significant component of human-caused global change. *N Z J Ecol* 21:1–16
- Weigle SM, Smith LD, Carlton JT, Pederson J (2005) Assessing the risk of introducing exotic species via the live marine species trade. *Conserv Biol* 19:213–223. <https://doi.org/10.1111/j.1523-1739.2005.00412.x>
- Worm B, Tittensor DP (2018) *A theory of global diversity*. Princeton University Press, Princeton

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.