



FEATURE ARTICLE

# Climate change winner in the deep sea? Predicting the impacts of climate change on the distribution of the glass sponge *Vazella pourtalesii*

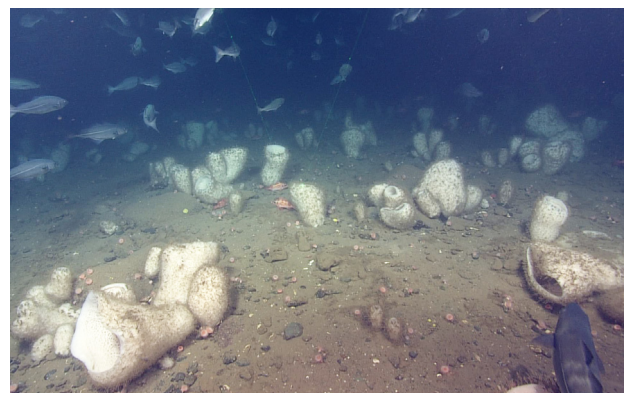
Lindsay Beazley<sup>1,\*</sup>, Ellen Kenchington<sup>1</sup>, Francisco Javier Murillo<sup>1</sup>, David Brickman<sup>1</sup>,  
Zeliang Wang<sup>1</sup>, Andrew J. Davies<sup>2</sup>, Emyr Martyn Roberts<sup>3</sup>, Hans Tore Rapp<sup>3,†</sup>

<sup>1</sup>Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia B2Y 4A2, Canada

<sup>2</sup>Department of Biological Sciences, University of Rhode Island, Kingston, Rhode Island 02881, USA

<sup>3</sup>Department of Biological Sciences and K. G. Jebsen Centre for Deep-Sea Research, University of Bergen, 5020 Bergen, Norway

**ABSTRACT:** Shallow-water sponges are often cited as being 'climate change winners' due to their resiliency against climate change effects compared to other benthic taxa. However, little is known of the impacts of climate change on deep-water sponges. The deep-water glass sponge *Vazella pourtalesii* is distributed off eastern North America, forming dense sponge grounds with enhanced biodiversity on the Scotian Shelf off Nova Scotia, Canada. While the strong natural environmental variability that characterizes these sponge grounds suggests this species is resilient to a changing environment, its physiological limitations remain unknown, and the impact of more persistent anthropogenic climate change on its distribution has never been assessed. We used Random Forest and generalized additive models to project the distribution of *V. pourtalesii* in the northwest Atlantic using environmental conditions simulated under moderate and worst-case CO<sub>2</sub> emission scenarios. Under future (2046–2065) climate change, the suitable habitat of *V. pourtalesii* will increase up to 4 times its present-day size and shift into deeper waters and higher latitudes, particularly in its northern range where ocean warming will serve to improve the habitat surrounding this originally sub-tropical species. However, not all areas projected as suitable habitat in the future will realistically be populated, and the reduced likelihood of occurrence in its core habitat on the Scotian Shelf suggests that the existing *Vazella* sponge grounds may be negatively impacted. An effective monitoring programme will require tracking changes in the density and distribution of *V. pourtalesii* at the margins between core habitat and where losses and gains were projected.



Species distribution models showed that the suitable habitat of the deep-water glass sponge *Vazella pourtalesii* will expand under climate change scenarios.

*Photo: Fisheries and Oceans Canada*

**KEY WORDS:** Glass sponge · *Vazella pourtalesii* · Northwest Atlantic · Deep sea · Climate change · Species distribution modelling · Random Forest · Generalized additive model

## 1. INTRODUCTION

Over the past century, upper-ocean temperatures have increased across much of the northwest Atlantic, consistent with the global trend of increasing sea surface temperature due to anthropogenic climate change (Bush & Lemmen 2019). The effects of ocean warming have already been observed in marine ecosystems

\*Corresponding author: lindsay.beazley@dfo-mpo.gc.ca

†Deceased

across the region, through northward shifts in the range of commercially harvested fish and their catch distribution (Nye et al. 2009, Pinsky & Fogarty 2012). Given the inevitable socio-economic impacts of a changing climate on the valuable fisheries that operate in the northeast USA and on the Scotian Shelf in Canadian waters, several studies have aimed to identify particular geographic areas of commercially harvested species considered vulnerable to present and future climate change (e.g. Hare et al. 2016, Rheuban et al. 2018). In the absence of information on physiological thresholds, species distribution models (SDMs) are often used to estimate potential climate change effects by predicting the degree of habitat loss or gain under various emission scenarios and time periods, based on the statistical association between occurrence records and spatial environmental data. Such correlative tools have been recently utilized to predict changes in the future distribution of commercially harvested groundfish (e.g. Kleisner et al. 2017, Stanley et al. 2018, McHenry et al. 2019) and motile invertebrate species (e.g. Tanaka et al. 2017, Stanley et al. 2018, Greenan et al. 2019) in the northwest Atlantic. However, much less is known about the impacts of climate change on the distribution of long-lived, sessile benthic invertebrates of conservation, but not commercial, importance in the region (although see Morato et al. 2020). Such communities may be more vulnerable than motile species, as latitudinal shifts in distribution are unlikely to occur over short time scales.

Sponges are key components of benthic marine ecosystems worldwide, from intertidal to abyssal depths (Maldonado et al. 2017). In the deep sea, sponge species can be widely distributed across vast areas as isolated individuals, but in certain locations where environmental conditions are favourable, they may form dense aggregations commonly known as sponge grounds. The importance of such aggregations in deep-sea environments has only recently emerged, through demonstrations of their role in biodiversity enhancement and habitat provision (Beazley et al. 2013, 2015, Hawkes et al. 2019, Murillo et al. 2020), and their importance in benthic–pelagic coupling and the cycling of nutrients (Kutti et al. 2013, Pham et al. 2019).

The Scotian Shelf off Nova Scotia, eastern Canada, is home to a globally unique aggregation of the glass sponge *Vazella pourtalesii*. This rosellid species has a wide but fragmented distribution along the continental margin of eastern North America between ~100 and 750 m depth (Beazley et al. 2018, NOAA 2019), occurring in low densities off Florida in the

southeastern USA, and forming the densest known monospecific aggregations of their kind in Emerald Basin on the Scotian Shelf. Through an analysis of *in situ* benthic imagery inside the 'Vazella sponge grounds' (sensu Beazley et al. 2018), Hawkes et al. (2019) found that on average, mean species density (i.e. richness) and abundance of other epibenthic megafauna was ~3–4 times higher in the presence of *V. pourtalesii*, demonstrating the importance of these sponge grounds in enhancing local biodiversity.

Beazley et al. (2018) used classification Random Forest modelling to predict the present-day distribution of *V. pourtalesii* across the Scotian Shelf. A suite of 35 predictor variables relating to ocean terrain, fishing effort, and biological and physical oceanographic characteristics considered to represent the current oceanographic climate were utilized. Additionally, historical trends in bottom temperature and salinity, 2 variables that were identified as important determinants of its present-day distribution, were examined through the hindcast Simple Ocean Data Assimilation model, revealing that the *Vazella* sponge grounds have historically been subjected to strong inter-annual and multi-decadal variability in water mass characteristics, with annually averaged bottom temperatures varying by up to 8°C (range 4–12°C). This variability was related to periods of larger than average transport of cold, fresh Labrador Slope Water by the Labrador Current, consistent with the influence of the Atlantic Multi-Decadal Oscillation (AMO), a mode of natural variability in the North Atlantic responsible for warming and cooling during positive and negative phases of the Atlantic Meridional Overturning Circulation, respectively. The persistence of *V. pourtalesii* in Emerald Basin since its discovery in the late 1800s (Honeyman 1889) suggests that this species may be inherently adapted to a highly dynamic physical environment, and able to withstand the particularly pronounced cold conditions that prevailed in Emerald Basin and across the Scotian Shelf in the 1960s (Beazley et al. 2018). While the impact of persistently warmer conditions that exceed the upper thermal range of the AMO (~12°C; Beazley et al. 2018) on *V. pourtalesii* has not been quantitatively assessed, the presence of this species in the subtropical waters off the mid- and southeastern USA suggests it is adapted to withstand a warmer climate. However, its ability to adapt to such changes over relatively short time scales and the physiological tolerance of individuals and populations in different parts of its range, remain unknown.

The response of sponges to ocean warming and acidification is well studied in shallow-water ecosys-

tems (see overviews in Carballo & Bell 2017 and Bell et al. 2018). In these environments, sponges have been referred to as potential ‘climate change winners’ (Bell et al. 2013, 2018), due mainly to the physiological tolerance of some species to high thermal anomalies that have caused bleaching events in corals. Observations from several experimental and field-based studies suggest that some shallow-water sponges have an upper thermal tolerance of  $\sim 2\text{--}3^\circ\text{C}$  above mean monthly values (Bennett et al. 2017, Ramsby et al. 2018). However, virtually nothing is known of the physiological effects of climate change on glass sponges, which are thought to be relatively sensitive to changes in temperature (Leys & Meech 2006), and how their physiological thresholds may drive community or ecological responses (e.g. distribution shifts) under a changing climate. In an atmospheric  $\text{CO}_2$  doubling experiment, upper ocean (0–300 m depth) temperatures on the northwest Atlantic shelf between  $35$  and  $45^\circ\text{N}$  were predicted to warm by  $\sim 3^\circ\text{C}$  (Saba et al. 2016), a rate 2–3 times faster than the global average. This change may have important implications for the distribution of the *Vazella* sponge grounds and the biodiversity that they support. Furthermore, understanding the effects of climate change on this vulnerable marine ecosystem indicator species (Fuller et al. 2008) will be necessary to determine whether the 2 bottom-fishery closures currently in place on the Scotian Shelf for the conservation of this species require more adaptive approaches to ensure their effectiveness into the future.

In the absence of information on the physiological tolerance of this long-lived, deep-water glass sponge, we here aim to deduce the potential impacts of future climate change on the distribution of *V. pourtalesii* using correlative SDM techniques. We employed the machine-learning method Random Forest, and generalized additive modelling (GAM) to project the distribution of this species in the northwest Atlantic using environmental variables simulated for the bi-decadal period 2046–2065 under moderate (representative concentration pathway [RCP] 4.5) and worst-case (RCP 8.5)  $\text{CO}_2$  emission scenarios. This period was chosen instead of the commonly modelled end-of-century (2100) to allow for consideration of the results in current conservation management strategies. We assessed whether this long-lived, deep-water sponge species is a potential ‘winner’, ‘loser’, or is neutral to the effects of climate change based on whether its suitable habitat expanded, contracted, or showed no change under future climatic conditions.

We assume that *V. pourtalesii* is locally adapted to the unique water mass characteristics that influence

the northern and southern portions of its range, and have applied models separately to the occurrences located on the Scotian Shelf in Canadian waters and those off the mid–southeast USA. Collectively, our approach will provide a better understanding of the potential upper thermal tolerance limit of this species and how it may affect its distribution in the northwest Atlantic into the future. To our knowledge, our study represents the first application of SDMs to evaluate climate-induced changes in the distribution of a deep-water, ground-forming sponge, and has important implications for the conservation of these relatively understudied ecosystems.

## 2. MATERIALS AND METHODS

### 2.1. Distribution of *Vazella pourtalesii* and study area

Occurrences of *V. pourtalesii* in the northwest Atlantic (Fig. 1) come from a variety of sources (Table 1). On the Scotian Shelf (see Breeze & Horsman 2005), significant effort has been made to map the distribution of the *Vazella* sponge grounds for conservation management purposes, and as a result, the occurrences (Fig. 1) are spatially biased towards this part of its range. There, records are collated from the Fisheries and Oceans Canada (DFO) multispecies research vessel trawl survey, DFO and Natural Resources of Canada (NRCAN) optical (camera/video) benthic surveys (comprising the majority of observations; Table 1), and commercial bycatch records from the Fisheries Observer Program. The DFO multispecies research vessel trawl survey is a depth-stratified random survey conducted primarily using Western IIA trawl gear. Tows are 30 min in duration, and the average linear distance covered is 3.17 km (Beazley et al. 2018). Commercial trawls in the region are often much longer in duration and length, collating catch data over distances of 10+ km. For both the research vessel and commercial trawl data, the start position of each tow was used to represent the presence of *V. pourtalesii* in the catch.

Records located off the mid–southeast USA (Table 1) were extracted from a combination of sources including the primary literature and online repositories, and through direct video annotations collected for the purpose of this study. Three specimens collected in 1868 off Florida that were re-examined by Reiswig (1996) are included here, as well as remotely operated vehicle (ROV) and submersible records extracted from the NOAA Deep-Sea Coral Data Portal

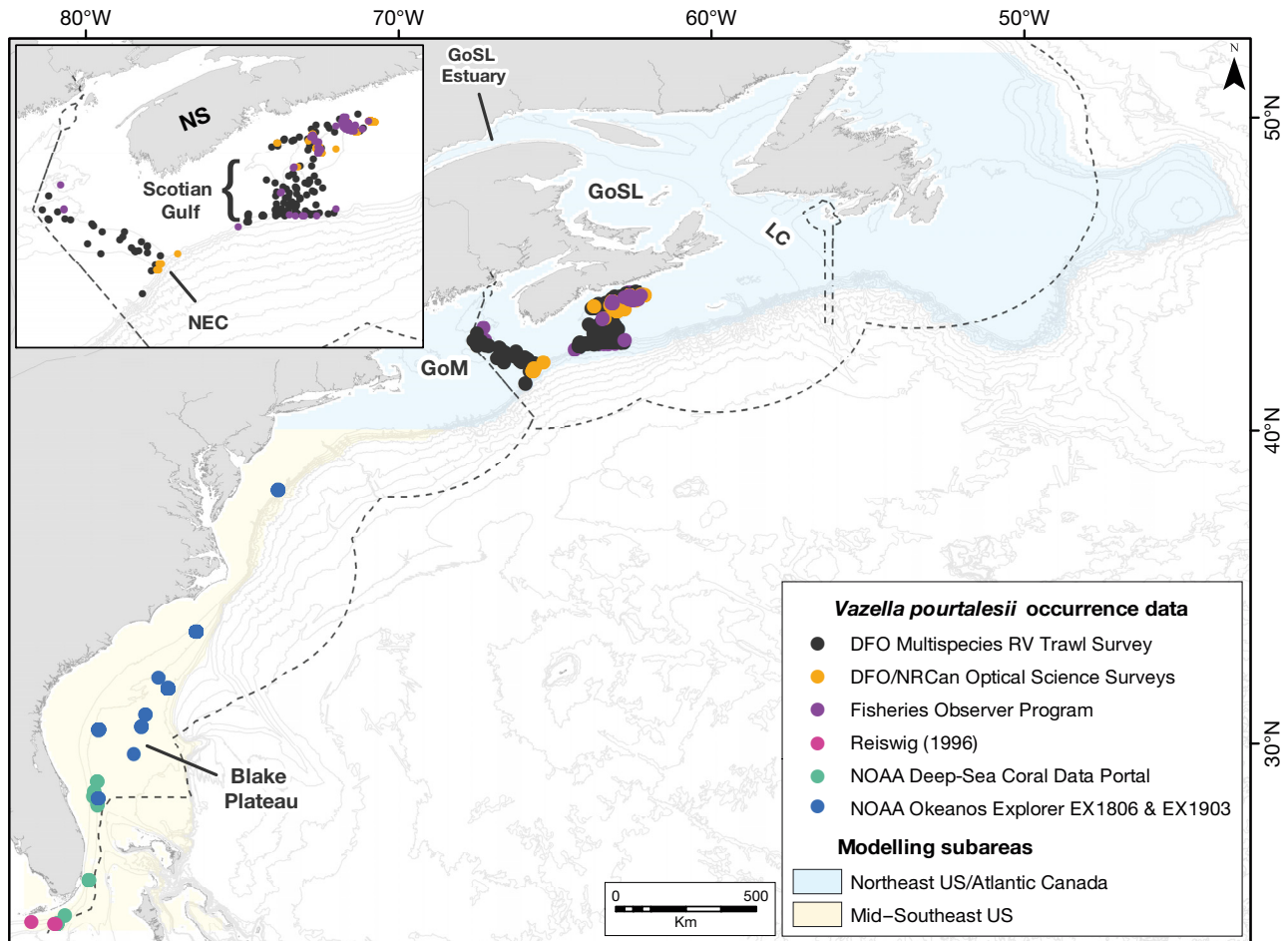


Fig. 1. Occurrences of *Vazella pourtalesii*. Presences are colour-coded according to their source as outlined in Table 1. The black dashed lines show the exclusive economic zones of countries within the study region. Inset shows a close-up of the distribution of records on the Scotian Shelf off Nova Scotia (NS). NEC: Northeast Channel; GoM: Gulf of Maine; LC: Laurentian Channel; GoSL: Gulf of St. Lawrence. Also shown are the modelling subareas in which the species distribution models were trained

Table 1. Sources of *Vazella pourtalesii* occurrence data, including general location, record/gear type, date range of data source, and original number of presences prior to filtering the data for modelling purposes

Source	Location	Record type	Date collected	Number of presences
DFO Multispecies RV Trawl Survey	Scotian Shelf	Research vessel trawl catch	2007–2019	131
DFO/NRCan Optical Science Surveys	Scotian Shelf	Still images, video observations	1989–2018	4216
Fisheries Observer Program	Scotian Shelf	Commercial trawl bycatch	1997–2007, 2010–2015	104
Reiswig (1996)	Southeast USA	Collected specimens	1868	3
NOAA Deep-Sea Coral Data Portal	Southeast USA	Still images, video observations	2009–2011	38
NOAA Okeanos Explorer EX1806 & EX1903	Mid- and southeast USA	Video observations	2018–2019	1051

(<https://deepseacoraldata.noaa.gov/>). In both 2018 and 2019, NOAA conducted 2 ROV surveys off the mid- and southeastern USA ('Windows to the Deep'

Okeanos Explorer EX1806 and EX1903 missions in 2018 and 2019, respectively) and reported the presence of *V. pourtalesii*. In order to increase the number

of presences in this relatively under-sampled (compared to the Scotian Shelf) portion of its range, all dives within its expected depth range (<1000 m) from these 2 missions were examined and presences recorded using the Ocean Networks Canada SeaTube V2 annotation software (<https://data.oceannetworks.ca/SeaTubeV2>). A total of 1051 *V. pourtalesii* individuals were observed from 9 different stations (see Table S1 in the Supplement at [www.int-res.com/articles/suppl/m657p001\\_supp.pdf](http://www.int-res.com/articles/suppl/m657p001_supp.pdf), for details). Most observations were from the Blake Plateau where *V. pourtalesii* was found in localized, dense aggregations in association with the reef-building coral *Lophelia pertusa* (and its rubble). These observations extend the maximum depth of *V. pourtalesii* in the northwest Atlantic to 935 m (from 754 m previously reported from observations on Cape Canaveral North off Florida; NOAA 2019). Finally, a search for *V. pourtalesii* records not encompassed by the sources described above was made in the Ocean Biogeographic Information System (<https://obis.org/>) to ensure that all records were considered. No new records resulted from this search.

Given the wide latitudinal gradient and presumed environmental heterogeneity over which *V. pourtalesii* is distributed, key environmental variables (depth, bottom temperature, and salinity) shown to be important for the distribution of this species (Beazley et al. 2018) were evaluated at the *V. pourtalesii* presence localities in order to assess the potential generality and future transferability of a model trained on its entire range (see Wenger & Olden 2012). Temperature–salinity plots and bivariate plots of the relationships between depth–temperature and depth–salinity at the location of the occurrences were examined (see Fig. S1). At comparable depths, records of *V. pourtalesii* on the Scotian Shelf were located in a colder, less saline environment than those located off the mid–southeast USA. We therefore chose to model the *V. pourtalesii* occurrences in the northern portion its range separately from its southern distribution. Two modelling subareas were created (see Fig. 1) and split at 40°N, which roughly corresponds to the boundary of the 2 regimes of the Ekman portion of the Atlantic Meridional Overturning Circulation (see Fig. 9 of Wang et al. 2019), reflecting the anticyclonic and cyclonic atmosphere circulations over the sub-

tropical and subpolar gyres. The subarea north of 40°N is referred to as the ‘Northeast US/Atlantic Canada’ subarea, which includes the northeast USA and Atlantic Canada regions, but also the Flemish Cap situated in international waters off Newfoundland, Canada (Fig. 1). The spatial extent south of 40°N is referred to as the ‘Mid–Southeast US’ subarea. Both subareas extend seaward to 2000 m depth (based on the resampled General Bathymetric Chart of the Oceans [GEBCO] 2019 bathymetry layer), which was the upper depth limit of model extrapolation based on previous modelling exercises of this species (see Beazley et al. 2016, 2018). A 5 km land buffer was applied to all coastline within each subarea, and any environmental grid cell (see Section 2.2, below) falling within its boundaries was removed to avoid the inclusion of land points.

## 2.2. Environmental predictor layers

A total of 10 environmental predictor layers (Table 2), including 2 static (stable in time, i.e. depth, slope) and 8 dynamic (varying in time, e.g. temperature, salinity) layers were extracted for consideration in the models. Although previous modelling works (Beazley et al. 2018) highlighted the importance of both hydrographic conditions (e.g. bottom and surface temperature, bottom salinity) and ocean productivity (e.g. sea surface primary production) in predicting the distribution of this species, only a limited

Table 2. Environmental predictor variables considered for modelling the distribution of *Vazella pourtalesii*. Dynamic variables are from the Bedford Institute of Oceanography North Atlantic Model and are available for the RCP 4.5 and 8.5 scenarios. All variables were gridded to the final 0.088° cell size. Note that salinity as incorporated here is considered unitless. Correlated variables were removed prior to distribution modelling

Variable	Unit	Native resolution	Used in models
<b>Dynamic</b>			
Mean Bottom Temperature	°C	1/12°	Yes
Mean Bottom Salinity	–	1/12°	Yes
Mean Bottom Current Velocity	m s <sup>-1</sup>	1/12°	Yes
Mean Bottom Shear	Pa	1/12°	No
Mean Surface Temperature	°C	1/12°	Yes
Mean Surface Salinity	–	1/12°	Yes
Mean Surface Current Velocity	m s <sup>-1</sup>	1/12°	Yes
Mean Maximum Mixed Layer Depth	m	1/12°	Yes
<b>Static</b>			
Depth	m	15-arc seconds (0.00417°)	No
Slope	°	15-arc seconds (0.00417°)	Yes

number of physical hydrographical variables (bottom and surface temperature, salinity, and current velocity; bottom shear stress and mixed layer depth [MLD]) were available for both present and future climatic conditions over the full modelling domain. The static variables depth and slope were chosen based on their continuous coverage across the modelling domain and importance in predicting the distribution of deep-sea sponges elsewhere (Howell et al. 2016). Mean bottom and surface temperature, salinity, and current velocity, as well as mean bottom shear and mean maximum MLD were extracted from the Bedford Institute of Oceanography North Atlantic Model (BNAM; Wang et al. 2018) to represent mean present-day climatic conditions. The BNAM is based on the Nucleus for European Modelling of the Ocean (NEMO) 2.3 re-analysis model, which has a horizontal resolution of  $1/12^\circ$  and consists of 50 vertical layers of varying thickness, ranging from 1 m at the surface to up to 450 m in the bottom layer (Wang et al. 2018). Monthly data for the time period 1990–2015 were extracted from BNAM and averaged across months and years to represent present-day mean climatic conditions.

A version of BNAM was developed to simulate future climatologies for the bi-decadal period 2046–2065 (referred to hereafter as the ‘2055 climatology’) under 2 different IPCC (IPCC 2013) emission scenarios: RCP 4.5 and RCP 8.5 (Brickman et al. 2016). RCP 4.5 is considered an emission-stabilizing scenario in which radiative forcing increases to 4.5 watts  $m^{-2}$  and stabilizes around 2100, when trajectories of atmospheric  $CO_2$  concentration and the median global temperature anomaly reach  $\sim 650$  ppm and  $2.4^\circ C$ , respectively (Moss et al. 2010). RCP 8.5 is a high-emission scenario in which radiative forcing increases to 8.5 watts  $m^{-2}$  by 2100 ( $\sim 1370$  ppm  $CO_2$  and a median temperature anomaly of  $4.9^\circ C$ ; Moss et al. 2010), but does not stabilize until 2300 when it reaches 12 watts  $m^{-2}$  (van Vuuren et al. 2011). Averaged annual anomalies representing the difference between present day and future (2055) conditions were extracted from BNAM simulations (see Brickman et al. 2016) and applied (added) to the averaged present-day climatology layers modelled in this study to create predictor layers representing future forecasted conditions.

Depth was extracted from the 2019 GEBCO grid ([https://www.gebco.net/data\\_and\\_products/gridded\\_bathymetry\\_data/gebco\\_2019/gebco\\_2019\\_info.html](https://www.gebco.net/data_and_products/gridded_bathymetry_data/gebco_2019/gebco_2019_info.html)). Slope was derived from this bathymetry using the ‘Slope’ tool in ArcGIS Desktop 10.7 (ESRI 2019). Prior to the generation of slope, depth was projected in an Albers equal area conic projection where units are represented in metres.

Continuous raster surfaces of the BNAM point data were created in ArcGIS using the ‘Point to Raster’ tool, with point data displayed using geographic coordinates and a WGS 1984 datum. The final raster cell size was  $0.088^\circ$ , which is approximately equivalent to 7.5 km horizontal resolution at the centre of the study area ( $\sim 40^\circ N$ ). Bilinear interpolation was used to resample both depth and slope to match the resolution of the climatic variables ( $0.088^\circ$ ).

### 2.3. Pseudo-absence generation

The most widely used SDM techniques can be broadly categorized into those that utilize only presences, and those that require information on both presences and absences (Barbet-Massin et al. 2012). In cases where real absence data are not available, absence points (‘background data’ or ‘pseudo-absences’) can be simulated and extracted from the model domain or ‘background’. Despite popular belief, MaxEnt, which is often falsely referred to as a ‘presence-only’ method, requires information on where a species is considered absent (i.e. pseudo-absences) in order to contrast the conditions at those locations against where it is present (VanDerWal et al. 2009, Barbet-Massin et al. 2012, Guillera-Arroita et al. 2015). Such models are more accurately considered ‘presence–background’ models (Guillera-Arroita et al. 2015), and should closely resemble the results of a presence–absence model in cases where a species is well-sampled but rare, and background points closely emulate the distribution of true absences (Ward et al. 2009).

As the majority of the *V. pourtalesii* presence data were extracted from video surveys and discrete sampling locations, real absence data were not available over the majority of the subarea extents and at a scale relevant to the environmental variables modelled in our study. However, null catches from the DFO multispecies trawl survey operating on the Scotian Shelf in the Northeast US/Atlantic Canada subarea could be used to approximate absence locations, an approach which was recently used by Beazley et al. (2018) in a presence–absence Random Forest classification modelling exercise. We evaluated the results of Random Forest and GAM models generated using absences from null catches collected between 2007 and 2017 from the DFO multispecies trawl survey on the Scotian Shelf, augmented with randomly generated pseudo-absences across the remainder of the Northeast US/Atlantic Canada subarea (see Fig. S2). Absences were augmented with

pseudo-absences to avoid potential extrapolation of models built on data collected only from the Scotian Shelf to other, environmentally dissimilar locations in the Northeast US/Atlantic Canada subarea.

The Random Forest and GAM models built on absences augmented with pseudo-absences had lower accuracy measures (see Table S2) compared to Beazley et al. (2018) and to the final models presented in this study (see Table 3), performing particularly poorly in their ability to correctly predict the absence data (lower 'specificity'). This was possibly due to the lower number of environmental variables used here, and close proximity of trawl survey absences to the presence data in its core habitat in Emerald Basin and the inability of the models to consistently classify or fit a model to the presences and absences that co-occur in the same water mass. The resulting predicted/projected distribution surfaces (see Fig. S3) appeared to over-extend the suitable habitat of this species into more coastal and shelf depths, where absences from the trawl survey are distributed. Therefore, we chose to run our models using randomly generated pseudo-absences across the entire Northeast US/Atlantic Canada subarea. A description of the methods used for pseudo-absence generation can be found in Text S1.

#### 2.4. Species distribution modelling, evaluation, and prediction

Prior to model fitting, correlation between environmental variable pairs was examined using R statistical software version 3.6.1 (R Core Team 2019), and those variables (Depth and Mean Bottom Shear) considered highly correlated (Spearman's  $\rho > 0.7$ ) with others were not considered further. Two different SDM approaches were employed to model the distribution of *V. pourtalesii* in each subarea: Random Forest classification (Breiman 2001a) and GAMs (Hastie & Tibshirani 1986). Both techniques have been recently used in applications dealing with presences and pseudo-absences (González-Irusta et al. 2015, Morato et al. 2020), and were employed here to maximize both model prediction and inference, and to increase reliability of model predictions in areas outside the environmental domain of the training data. Black box approaches such as Random Forest tend to optimize model prediction (i.e. the ability to predict a response variable based on the set of independent variables; Breiman 2001b) without particular focus on identifying the relationship between the predictors and response, so long as accurate predictions of

the response are yielded (James et al. 2013). In contrast, regression methods such as GAMs are better at identifying the type (linear or non-linear), strength, and direction of the relationship between the response and predictors (James et al. 2013), and are therefore more useful in applications requiring biological inference (i.e. understanding how changes in the predicts variables affect the response). Furthermore, when extrapolating outside the domain of the training data, where different physical conditions from those used to train the model may exist, Random Forest predicts the same value as it would for the closest value in the tree for which it had training data (Breiman et al. 1984). Regression techniques are better at predicting trends in the data (Hengl et al. 2018), which is important when extrapolating species–environment relationships into the future.

Random Forest models were fitted using R package 'randomForest' (Liaw & Wiener 2002) with default parameters and 500 trees. GAMs were fitted with a binomial error distribution and cloglog link (see Zuur et al. 2009) using the 'gam' function in R package 'mgcv' (Wood 2011), after a series of generalized linear models indicated non-linear patterns between the response and some predictor variables. The final GAMs in each subarea were selected using a forward stepwise approach, which involved fitting single-variable models, identifying the model with the lowest Akaike's information criterion (AIC) and highest significance, and iteratively adding the other variables and selecting the best combination that resulted in the lowest AIC, and repeating the process until there was no further improvement in AIC. The final GAM built in the Northeast US/Atlantic Canada subarea included 7 variables (bottom and surface temperature and salinity, MLD, bottom current velocity, and slope), while 4 variables (bottom and surface temperature, surface current velocity, and slope) were included in the GAM built in the Mid–Southeast US subarea. Smoothers were initially applied to all terms, and the effective degrees of freedom (edf) were evaluated. Following Wood (2001), smooth terms were replaced with parametric linear terms for variables with edf close to 1. Smoothers were applied to bottom and surface temperature and salinity, and MLD in the Northeast US/Atlantic Canada subarea GAM, and to bottom temperature in the Mid–Southeast US subarea GAM to account for their non-linear relationship with the response, with knots constrained to a maximum of 4 to avoid overfitting.

Model accuracy metrics and threshold probabilities were derived using 5-fold spatial block cross-validation (CV), which reduces the effects of spatial

autocorrelation on model prediction error (Roberts et al. 2017), and has been applied in recent deep-sea modelling efforts (Guinotte & Davies 2014). Using the 'spatialBlock' and 'spatialAutoRange' functions in the R package 'blockCV' (Valavi et al. 2019), each sub-area was partitioned into regularly-shaped 'blocks', and the presence/pseudo-absence data falling within each block were randomly assigned to either the training or testing dataset of each of the 5 folds. To ensure similar prevalence rates between folds, latitudinal and longitudinal offsets were applied to shift the blocks on their longitudinal and latitudinal axes in the Northeast US/Atlantic Canada subarea, and the 'numLimit' parameter was set to 0 to ensure that the most evenly dispersed number of records was chosen (with 500 iterations). The number of presences and pseudo-absences assigned to the training and testing datasets of each fold is shown in Table S3.

For each subarea, Random Forest models and GAMs were fitted to the training datasets from 4 folds and validated on the testing dataset from the fifth fold. The process was repeated 5 times, resulting in 5 CV runs from which accuracy metrics were derived. The threshold-independent area under the receiver operating characteristic curve (AUC) was calculated for each CV run, from which the mean and standard deviation were derived. The 'optimal.thresholds' function in the R package 'PresenceAbsence' (Freeman & Moisen 2008) was used to calculate several common threshold values (e.g. maximum Kappa, sensitivity = specificity, observed prevalence; Liu et al. 2005) above which a given relative likelihood of occurrence is considered a presence. We selected the threshold which maximizes the sum of sensitivity and specificity (MSS), where sensitivity and specificity represent the proportion of accurately predicted presences and absences, respectively. This threshold was used to convert the likelihood of occurrence outcomes from each CV run into predicted outcomes (0 or 1) that are then summarized into a  $2 \times 2$  confusion matrix. Sensitivity, specificity, and the true skill statistic (TSS; Allouche et al. 2006) were derived and used to assess model performance along with AUC.

From presence–background data, it is not possible to determine whether a species is well surveyed but rare, or common but under-surveyed, a function which requires presence–absence or occupancy detection data (Guillera-Arroita et al. 2015). Consequently, predictions from models that use pseudo-absences over real absence data are more indicative of a species' relative likelihood of occurrence than probability of presence (Guillera-Arroita et al. 2015).

Therefore, relative likelihood of occurrence was adopted herein to refer to the Random Forest and GAM outputs. The final predictions/projections of the relative likelihood of *V. pourtalesii* occurrence under present-day and future climatic conditions were generated from models trained on the full presence/pseudo-absence datasets in each subarea: 118 presences and 3373 pseudo-absence in the Northeast US/Atlantic Canada subarea for both Random Forest and GAM (see Text S1) and 18 presences and 1164 pseudo-absences for GAM, and 18 presences and 237 pseudo-absences for Random Forest in the Mid–Southeast US subarea. As all the data are used (important for small datasets), this approach was chosen over averaging the predictions generated from models built in each CV run. Its disadvantage is that the error estimates from CV do not apply perfectly to the final models, as slightly different datasets were used to train them. However, the error rates from CV are likely conservative, with better performance expected from models built on the full dataset (Roberts et al. 2017).

Model uncertainty was evaluated by calculating the standard deviation of the predictions generated from the 5 CV models. This estimate of standard deviation highlights areas of higher or lower variability between models built on different, spatially partitioned subsets of data, and areas of higher uncertainty may be reflective of local non-stationarity (Nephin et al. 2020). Similar to the accuracy metrics described above, this measure of uncertainty is considered conservative, as standard deviation was derived from the 5 CV models and not the final models used for prediction. To aid in identifying where model outputs may be less certain, especially those of Random Forest, areas of model extrapolation were identified by highlighting each grid cell across both subareas where at least 1 environmental variable had values above or below the present-day variables used to train the final models.

Predictions of relative likelihood of occurrence were thresholded into a binary depiction of suitable vs. unsuitable habitat using the MSS values used to threshold the confusion matrices of each model built in each subarea. Suitable habitat defined here is based solely on the relationship of the *V. pourtalesii* presences and pseudo-absences with the 8 present-day physical oceanographic and ocean terrain environmental predictor variables used in the models of this study. The relative likelihood of occurrence projected to the 2 future climatic scenarios was also thresholded using MSS, and areas that experienced a gain, loss, or no change in suitable habitat from present-day conditions were evaluated.



## 2.5. Variable importance and functional response curves

The importance of the predictor variables in Random Forest models was assessed using the mean decrease in Gini index extracted using the ‘importance’ function in the package ‘randomForest’. For GAMs, the ‘varImp’ function in the package ‘caret’ (Kuhn 2020) was used to extract the importance of the predictor variables based on their p-value. Following Lopes et al. (2019), ‘functional response curves’ were generated to evaluate changes in the relative likelihood of occurrence across environmental variable gradients. This method allowed for a direct comparison of the relative likelihood of occurrence–environment relationships for the present-day, RCP 4.5, and RCP 8.5 scenarios. Relative likelihood of occurrence and the value of environmental predictors for each raster cell across each study area were extracted and smoothed using the ‘loess’ method with span = 0.5 applied to each curve. Curves were displayed with 95 % standard error confidence intervals.

## 3. RESULTS

### 3.1. Changes in environmental conditions between present-day and future climatologies

The overall change in environmental conditions between present-day and the 2 future climatic scenarios was more prominent in the northern portion of *Vazella pourtalesii*'s range than across its southern distribution (see Table S4, Figs. S4 & S5). Bottom and surface temperature showed the largest change, with average Mean Bottom Temperature increasing by 0.80 and 0.38°C between present-day and RCP 8.5 in the Northeast US/Atlantic Canada and Mid–Southeast US subareas, respectively, and Mean Surface Temperature by 0.90 and 0.58°C (Table S4). This was followed by maximum MLD, which became shallower across both subareas in the future (change in average MLD: –0.59 and –0.35 m in the Northeast US/Atlantic Canada and Mid–Southeast US subareas, respectively). Changes in bottom salinity, bottom and surface current, and bottom shear between the present-day and future forecasted conditions were negligible in both subareas, while sur-

face salinity decreased in both subareas but more drastically in the northern subarea (change in Mean Surface Salinity –0.25 and –0.09 in the Northeast US/Atlantic Canada and Mid–Southeast US subareas, respectively; Table S4).

### 3.2. Model performance

Accuracy metrics (AUC, sensitivity, specificity, and TSS) of the Random Forest models and GAMs built in each subarea are shown in Table 3, while additional summaries of the parametric and smoothed terms of the GAMs are shown in Table S5. In the Northeast US/Atlantic Canada subarea, both Random Forest and GAM had an excellent performance rating based on AUC (>0.90), with GAM marginally outperforming Random Forest based on all 4 metrics. Models generated in the Mid–Southeast US subarea had a comparatively poor performance, with AUC values in the low 0.80 range, and low TSS values (0.52 and 0.49 for Random Forest and GAM, respectively). Models in both subareas performed similarly in terms of their ability to accurately predict the presences (i.e. sensitivity) and pseudo-absences (i.e. specificity) correctly, with specificity lower than sensitivity for all models (Table 3). The environmental variables considered most important for the present-day distribution of *V. pourtalesii* (see Fig. S6) were Mean Bottom Temperature and Mean Surface Salinity for Random Forest and GAM, respectively in the Northeast US/Atlantic Canada subarea. Mean Surface Temperature and Slope were the most important variables for these models in the Mid–Southeast US subarea (Fig. S6).

Table 3. Accuracy measures for the Random Forest models and generalized additive models (GAMs) trained and tested on the *Vazella pourtalesii* presence/pseudo-absence data in both the Northeast US/Atlantic Canada and Mid–Southeast US subareas. Cross-validation was done via 5-fold spatial blocking with random assignment of blocks into folds. Sensitivity, specificity, and the true skill statistic (TSS) were generated from a confusion matrix of tabulated outcomes that was thresholded using the maximum of sensitivity + specificity (MSS) identified for each model and subarea. AUC: area under the receiver operating characteristic curve

Model type	Mean AUC ± SD	Sensitivity	Specificity	TSS	MSS threshold
<b>Northeast US/Atlantic Canada subarea</b>					
Random Forest	0.99 ± 0.01	0.95	0.94	0.89	0.10
GAM	0.99 ± 0.01	0.95	0.94	0.89	0.03
<b>Mid–Southeast US subarea</b>					
Random Forest	0.82 ± 0.08	0.83	0.69	0.52	0.07
GAM	0.82 ± 0.07	0.89	0.60	0.49	0.01

### 3.3. Present-day and future distribution — Northeast US/Atlantic Canada

The areas predicted with a higher relative likelihood of occurrence of *V. pourtalesii* under present-day conditions were relatively consistent between Random Forest and GAM in the Northeast US/Atlantic Canada subarea (Fig. 2). In accordance with the spatial distribution of the occurrence data (Fig. 1), the Scotian Gulf, an inlet which opens up into the LaHave and Emerald Basins on the inner shelf, and the Northeast Channel, a deep channel at the mouth of the Gulf of Maine (see Beazley et al. 2018 for further details), were predicted with high relative likelihood of occurrence of *V. pourtalesii* by both models. However, under future environmental conditions, the spatial distribution of the projections diverged between models. Random Forest projected a higher relative likelihood of occurrence of *V. pourtalesii* in the Gulf of Maine than GAM, while GAM projected a higher relative likelihood of occurrence in the Laurentian Channel/Gulf of St. Lawrence than Random Forest. The spread into the Laurentian Channel/Gulf of St. Lawrence intensified under the RCP 8.5 scenario for both models, although GAM projected a higher relative likelihood of occurrence further north, into the mouth of the St. Lawrence estuary.

The binary depiction of suitable versus unsuitable habitat and comparison of the percent change in areas projected as suitable habitat between the different climatic scenarios (Figs. 3 & 4) also exemplify the overall range expansion of *V. pourtalesii* under future climate change. However, certain areas of suitable habitat in the Laurentian Channel and Gulf of St. Lawrence projected by GAM were also asso-

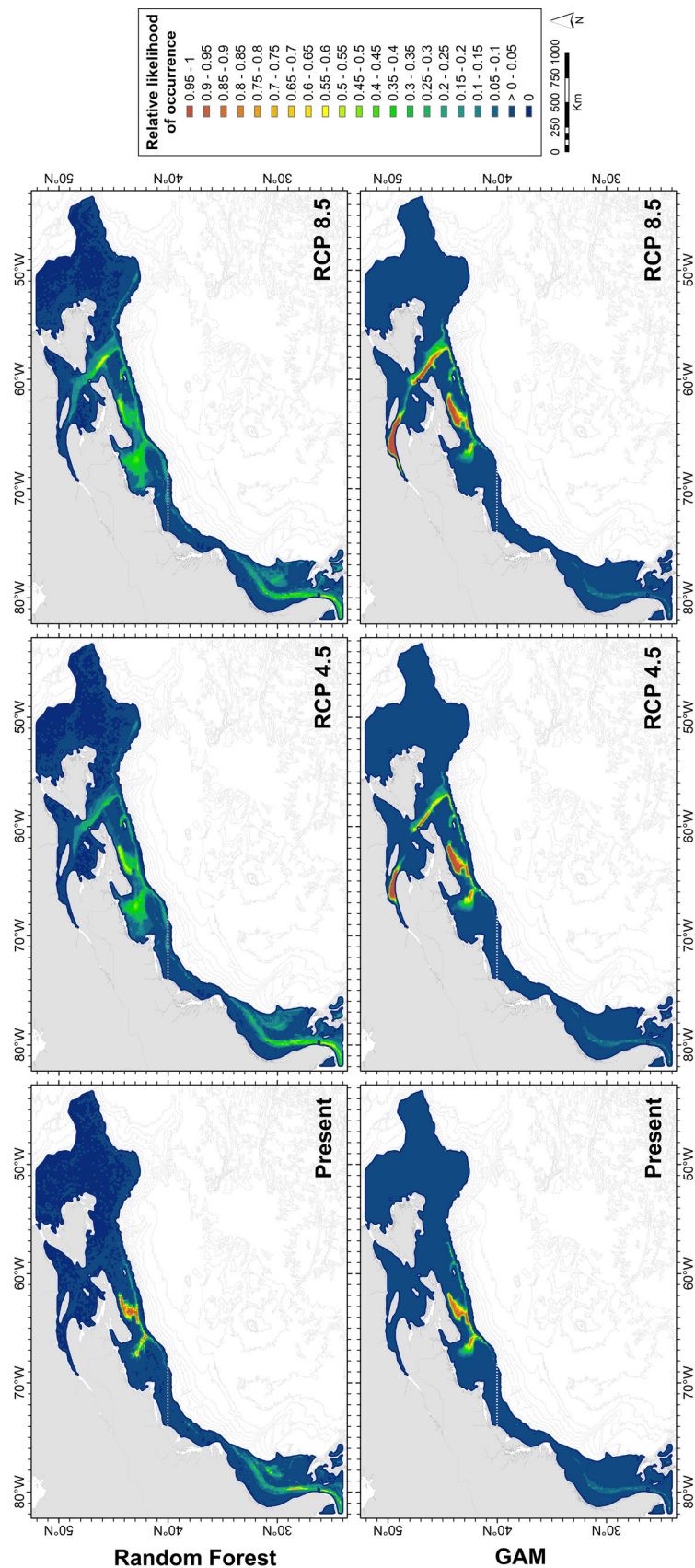


Fig. 2. Relative likelihood of occurrence of *Vazella pourtalesii* from Random Forest (top row) and generalized additive models (GAMs, bottom row) under present day, and representative concentration pathway (RCP) 4.5 and RCP 8.5 future climatic conditions. Models were built and applied within their respective subareas, which are separated at 40° N (indicated by the white dashed line)

ciated with high model uncertainty (Fig. 5). Evaluation of the prediction/projection outputs from each of the 5 CV models (not shown) showed that projections of relative likelihood of occurrence from the model based on the data partitioned into one particular fold (Fold 3) were erroneously high in the lower Laurentian Channel, mouth of the Gulf of St. Lawrence estuary, and northeast Scotian Gulf, congruent with the areas of higher uncertainty in Fig. 5. The training data of this fold had the lowest number of presences (70 compared to 89–114) and the highest number of pseudo-absences (2804 compared to 2645–2716) of all folds (see Table S3), and in contrast to the other folds, the presence data selected by spatial blocking for Fold 3 were located in the lower Scotian Gulf and Northeast Channel, south of the main sponge grounds in Emerald Basin.

While the CV models for Random Forest were built on the same subsets of data as the GAMs, their model projections were not as greatly affected by the varying numbers of presences and pseudo-absences in each fold, nor by their spatial distribution. This was reflected by the relatively low standard deviation across model predictions (Fig. 5). Random Forest is more prone to poor model results when predicting/projecting outside the environmental envelope of the training data. However, those areas where model uncertainty was high for both Random Forest and GAM were not considered extrapolated (see Fig. S7) with the exception of the area where GAM projected suitable habitat in the mouth of the Gulf of St. Lawrence estuary under the RCP 4.5 and 8.5 scenarios.

Another notable difference between model projections in this subarea was the decrease in rela-

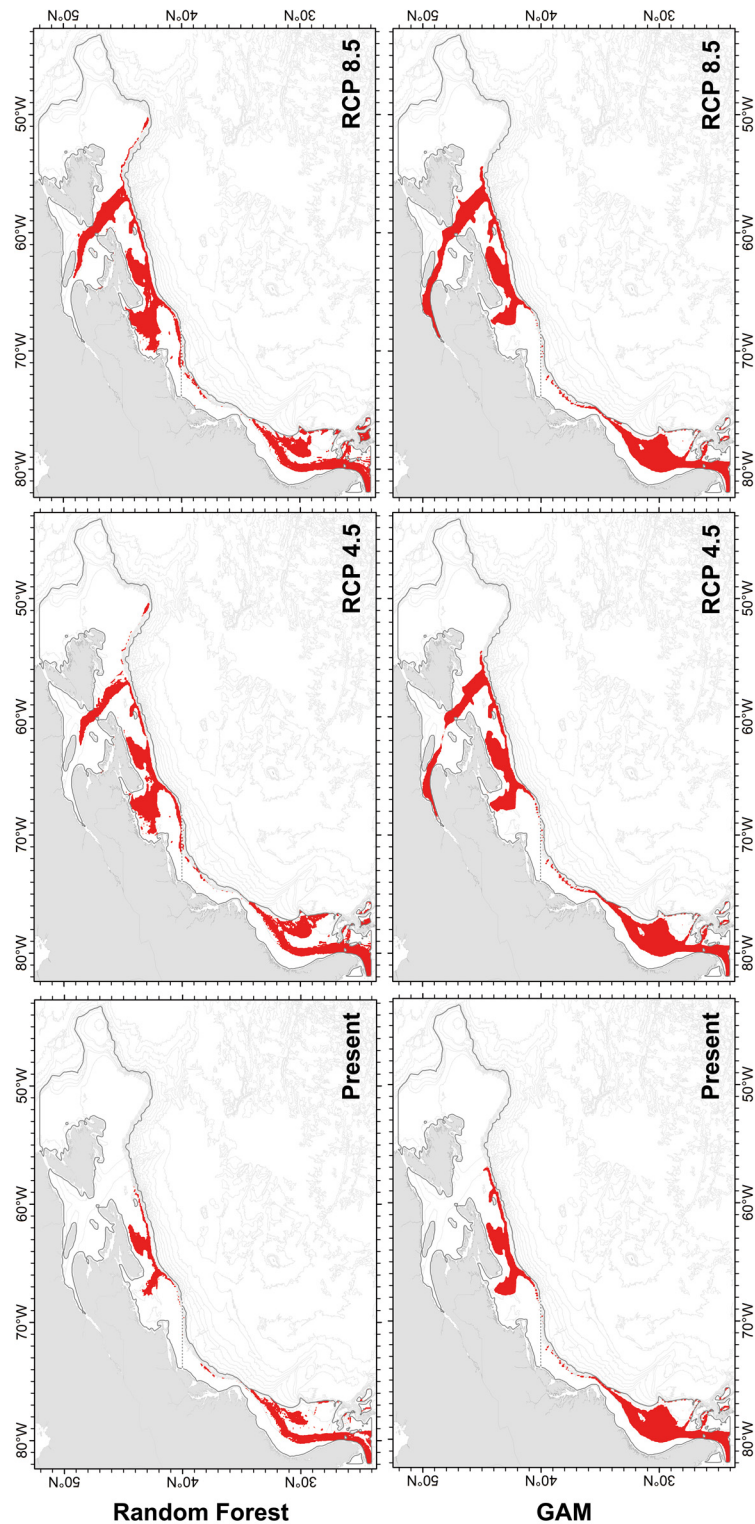


Fig. 3. *Vazella pourtalesii* suitable habitat (red) predicted/projected from Random Forest models (top row) and GAMs (bottom row) under present day, and representative concentration pathways (RCP) 4.5 and RCP 8.5 future climatic conditions. Relative likelihood of occurrence surfaces from Fig. 2 were thresholded using the maximum of sensitivity + specificity in Table 3. 40° N is indicated by the grey dashed line

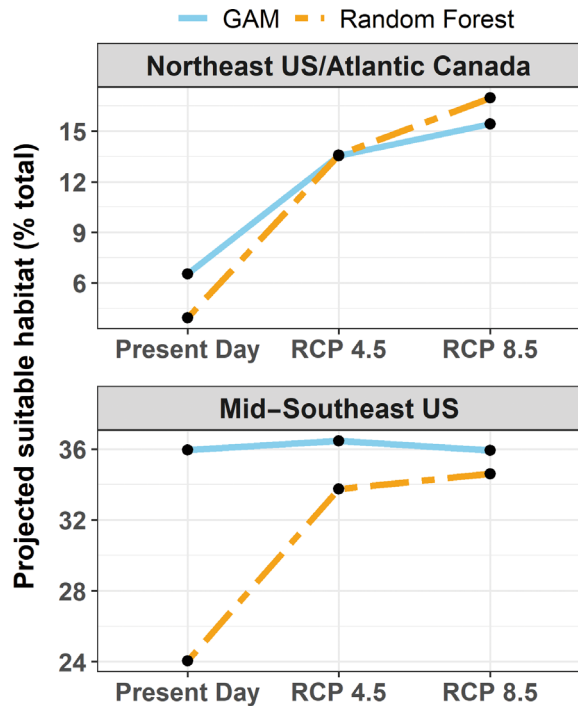


Fig. 4. Percentage of total area in the Northeast US/Atlantic Canada and Mid-Southeast US subareas predicted/projected as habitat suitable for *Vazella pourtalesii* by Random Forest models and GAMs under present-day environmental conditions and conditions projected under the representative concentration pathway (RCP) 4.5 and 8.5 emission scenarios for 2055

tive likelihood of occurrence of *V. pourtalesii* in its core habitat in the Scotian Gulf and Northeast Channel projected by Random Forest for both RCP scenarios (Fig. 2). In contrast, GAM projections of relative likelihood of occurrence intensified in the future. Although these areas were generally associated with higher model uncertainty (Fig. 5), this high uncertainty was due to the projections resulting from the GAM fit to the training data in Fold 3. With Fold 3 excluded (not shown), the standard deviation across the GAM predicted/projected surfaces was relatively low, and highest in the Gulf of St. Lawrence estuary. The slight northward shift and contraction in the areas of highest relative likelihood of occurrence in the Scotian Gulf and Northeast Channel, respectively, as projected by GAM under RCP 8.5 suggests that these model results collectively indicate that some impacts to the future distribution of *V. pourtalesii* may be incurred with further climate change.

Fig. 6 shows where Random Forest and GAM projected a gain, loss, and no change (i.e. the core suitable habitat) in suitable habitat between present-day and future forecasted conditions. For both models,

the core of this species' habitat in the Scotian Gulf and Northeast Channel in the Northeast US/Atlantic Canada subarea remained relatively stable between the present-day and future forecasted conditions, suggesting these areas may serve as refugia against climate change in the future. Areas projected to gain suitable habitat under the RCP 4.5 and 8.5 scenarios were, on average, deeper than the present-day habitat of *V. pourtalesii* (see Table S6), indicating an extension of this species into deeper waters as a result of climate change. For Random Forest, these areas were associated with warmer mean bottom temperatures ( $7.39 \pm 1.63^\circ\text{C}$  SD, up to  $12.48^\circ\text{C}$ ; RCP 8.5) compared to the present-day habitat in the Scotian Gulf and Northeast Channel/Gulf of Maine (Fig. 6), where temperatures were on average  $6.97 \pm 0.78^\circ\text{C}$  and reached a maximum of  $9.20^\circ\text{C}$ . In contrast, areas where GAM projected a gain in suitable habitat were, on average, colder than the areas projected by Random Forest ( $6.04 \pm 0.72^\circ\text{C}$  for RCP 8.5, maximum:  $9.33^\circ\text{C}$ ), and also slightly colder than average bottom temperature in its present-day habitat ( $6.77 \pm 0.92^\circ\text{C}$  in present day; although note the aforementioned higher uncertainty in GAM projections into the Gulf of St. Lawrence estuary). This difference is likely due to Random Forest projecting a larger gain in habitat in the Gulf of Maine, where simulated bottom temperatures are warmer than in the Laurentian Channel where GAM projected a larger gain than Random Forest (Fig. S4, Table S7). Table S6 shows that while the average depth of present-day suitable habitat predicted by both Random Forest and GAM is similar ( $203 \pm 125$  and  $219 \pm 177$  m, respectively), the average depth of suitable habitat projected by GAM under RCP 8.5 is over 50 m deeper ( $335 \pm 164$  m) than that projected by Random Forest ( $285 \pm 178$  m), likely explaining why average bottom temperatures are slightly cooler in its future suitable habitat.

The slight peak in relative likelihood of occurrence at  $6^\circ\text{C}$  and larger mode peaking at  $\sim 9^\circ\text{C}$  in the Random Forest functional curve for Mean Bottom Temperature (Fig. 7a) also reflects this pattern, with the smaller mode corresponding to temperatures in the Laurentian Channel/Gulf of St. Lawrence and the larger mode corresponding to the Scotian Gulf and Northeast Channel/Gulf of Maine where this model projected a larger gain in suitable habitat. For GAM, relative likelihood of occurrence under RCP 4.5 and 8.5 peaked at cooler temperatures ( $\sim 6^\circ\text{C}$ ; Fig. 7a), corresponding to the Laurentian Channel and Gulf of St. Lawrence where it projected a larger gain than Random Forest. Bottom temperatures in

both the Gulf of Maine and Laurentian Channel/Gulf of St. Lawrence warmed from their present-day conditions (Fig. S4, Table S7), suggesting that the habitat of *V. pourtalesii* may respond positively to warming bottom waters.

Patterns in the functional response curves presented here were similar to the presence–absence Random Forest partial dependence plots of Beazley et al. (2018) for similar environmental variables. For instance, both Random Forest and GAM models showed an increase in relative likelihood of occurrence at  $\sim 6^{\circ}\text{C}$  and  $\sim 34$  along the gradients in Mean Bottom Temperature and Salinity, respectively, similar to the peaks in Minimum Bottom Temperature ( $5\text{--}6^{\circ}\text{C}$ ) and Salinity (33.5) as shown by Beazley et al. (2018), suggesting that our use of pseudo-absences closely emulates the predicted distribution of *V. pourtalesii* based on real absence data.

### 3.4. Present-day and future distribution — Mid–Southeast US subarea

Patterns in the present-day and future predicted/projected distribution of *V. pourtalesii* in the Mid–Southeast US subarea were not as clearly defined as in the northern portion of its range. Under present-day environmental conditions, both Random Forest and GAM predicted low to moderate *V. pourtalesii* relative likelihood of occurrence in a band following the Florida peninsula (Fig. 2). This band is consistent with the location of *V. pourtalesii* occurrences around the peninsula and on Blake Plateau (Fig. 1). While uncertainty associated with both models was relatively low (Fig. 5), predictions/projections were low at the location of presence localities from both models (5 % relative like-

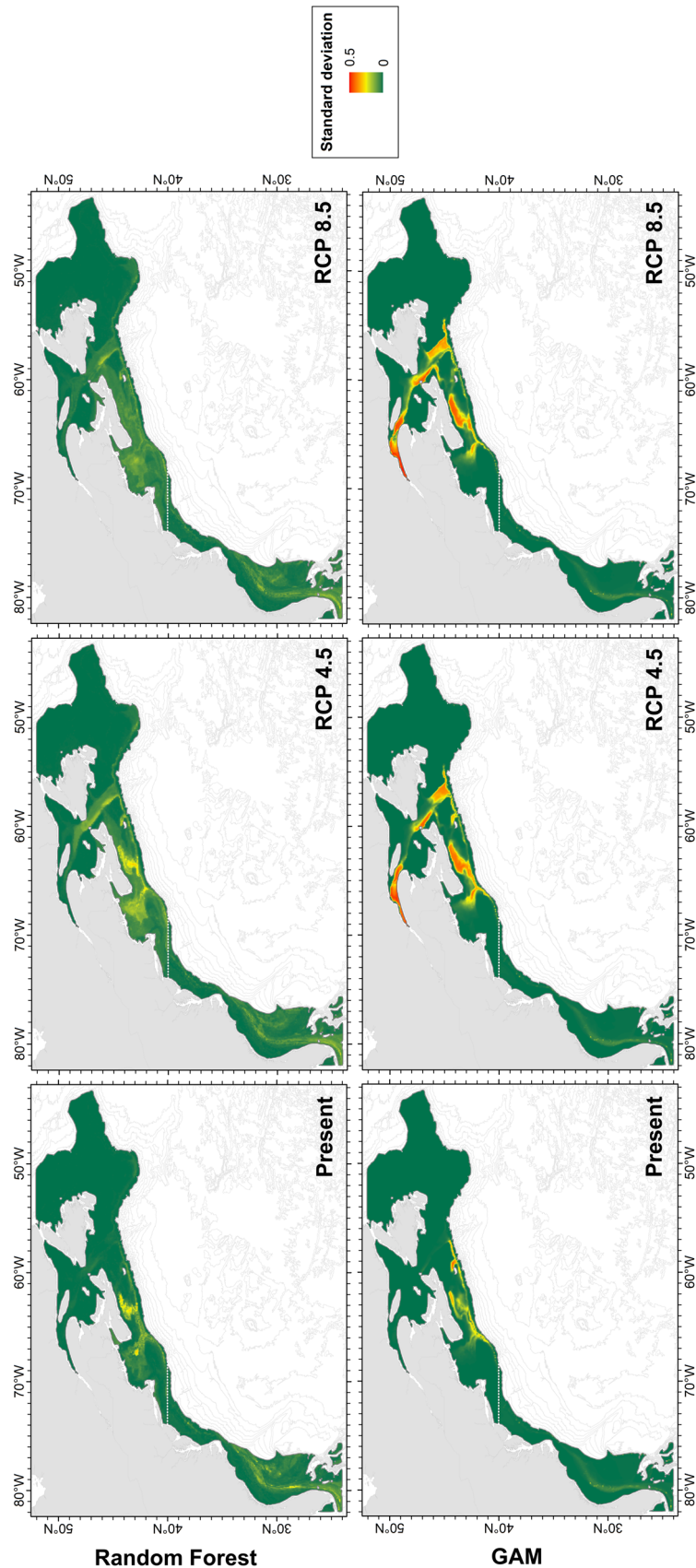


Fig. 5. Model uncertainty measured as the SD of present-day predictions and future projections of *Vazella pourtalesii* relative likelihood of occurrence from Random Forest models (top row) and GAMs (bottom row) generated in 5-fold spatial block cross-validation

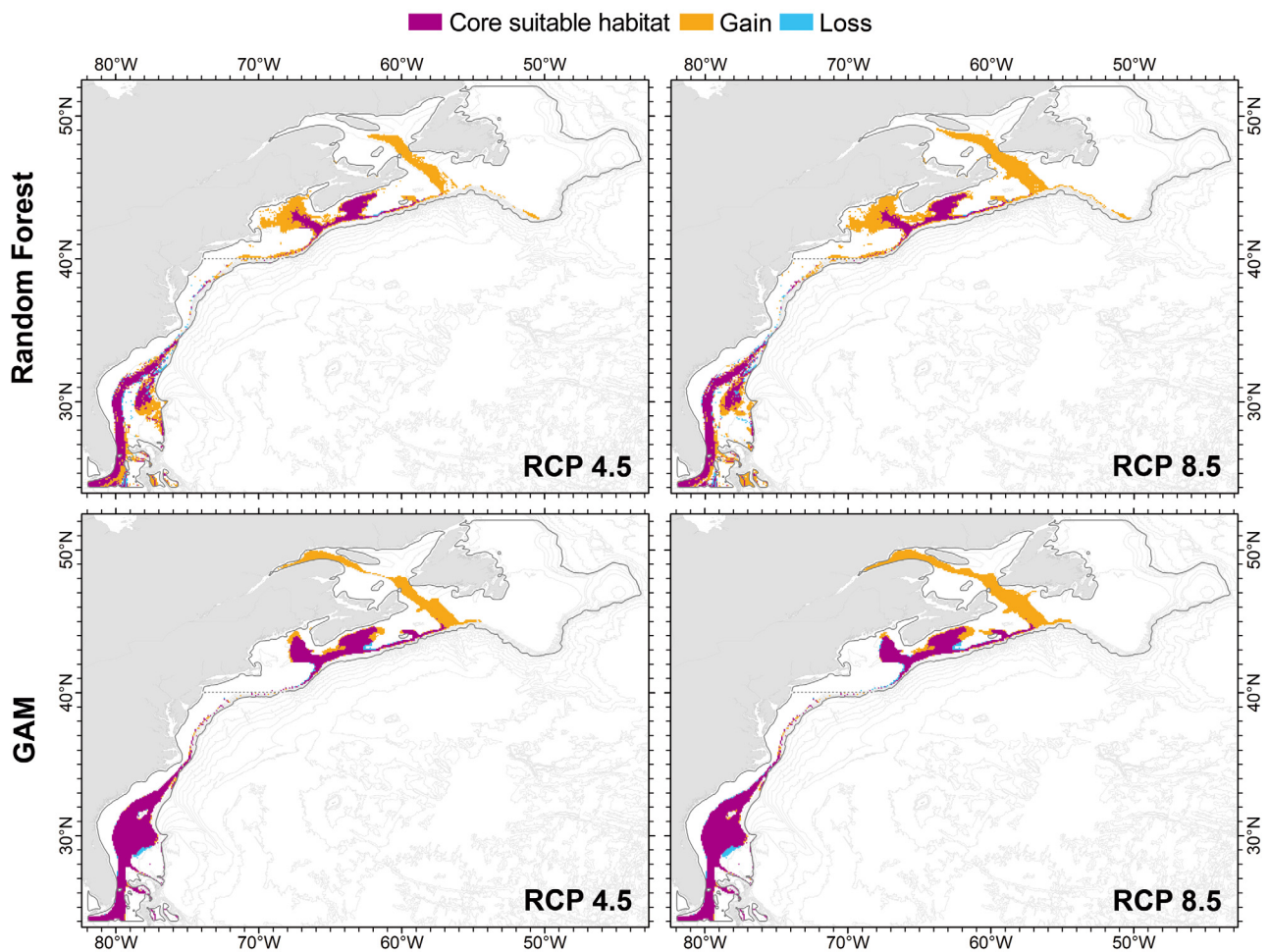


Fig. 6. Areas projected by Random Forest models and GAMs under relative concentration pathway (RCP) 4.5 and 8.5 emission scenarios to gain, lose, and show no change (i.e. core suitable habitat) in suitable habitat from present-day conditions. 40° N is indicated by the grey dashed line

likelihood of occurrence by GAM and up to 50% by Random Forest), reiterating their poor performance in this subarea (Table 3; Table S5). Changes in the intensity and distribution of the relative likelihood of occurrence under the 2 future emission scenarios were relatively small for Random Forest and virtually unchanging for GAM (Figs. 2 & 3).

Random Forest projected a slight gain in suitable habitat in the deeper waters of the Blake Plateau under RCP 4.5 conditions (Fig. 3). Similar to the pattern observed in the northern portion of its range, these areas corresponded to slightly colder temperatures than the present-day habitat of this species (see Table S8). However, the total area projected as suitable habitat by Random Forest showed only a negligible increase from RCP 4.5 to 8.5 (Fig. 4), suggesting that the warmer portion of this species' range will not benefit from additional ocean warming. GAM also projected a slight increase in suitable habitat under

RCP 4.5, followed by a decrease back to its present-day size under RCP 8.5 (Fig. 4). Fig. 6 shows a loss of suitable habitat in the deeper waters of the Blake Plateau as well as along the flanks of the refugia areas, which was slightly more pronounced for RCP 8.5 than RCP 4.5.

Relative likelihood of occurrence predicted/projected along the present-day and future gradients in Mean Bottom Temperature were highest at ~10–11°C for both Random Forest and GAM in the Mid-Southeast US subarea (Fig. 7b). However, Mean Bottom Temperature was not the top environmental predictor variable in either model (Fig. S6), suggesting that any inferences based on temperature should be made with caution. The selection of Mean Surface Temperature by Random Forest as its top predictor is likely due to the spatial congruence between the location of *V. pourtalesii* occurrences (Fig. 1) and the Florida Current, a swift, northward-flowing segment

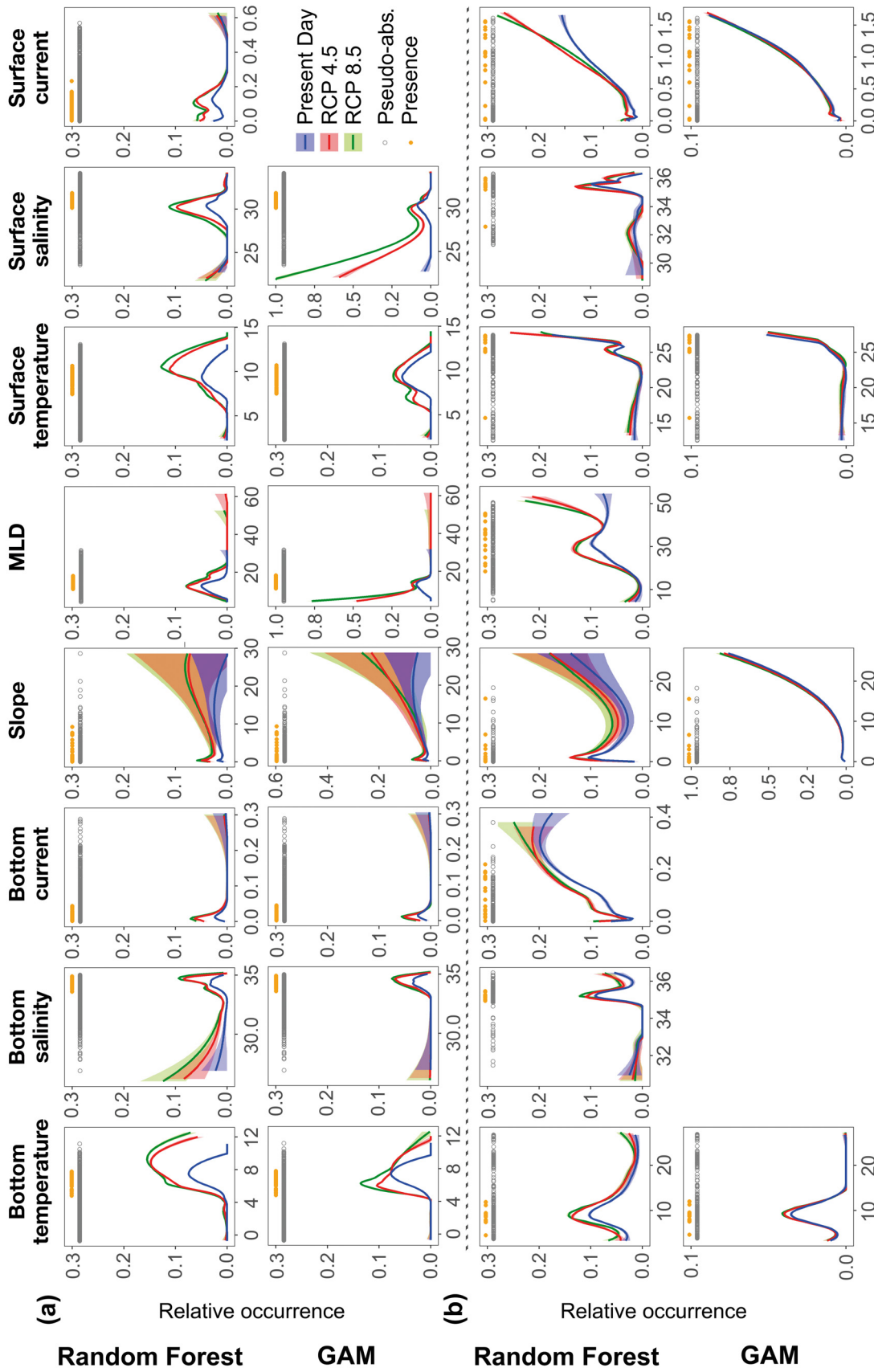


Fig. 7. Functional response curves for Random Forest models and GAMs built in the (a) Northeast US/Atlantic Canada and (b) Mid-Southeast US subareas showing the relationship between smoothed predictions/projections of the relative likelihood of occurrence of *Vazella pourtalesii* along the gradients of each environmental predictor variable from the present day, RCP 4.5, and RCP 8.5 climatologies. The absence of a variable means that variable was not included in the model. At the top of each plot are the *V. pourtalesii* presence (orange) and pseudo-absence data points (grey) across the range of each environmental predictor. Shaded regions represent approximate 95% confidence intervals. Curves and associated confidence intervals were bounded by 0 on the y-axis. Note the different y-axis scales for some variables. MLD: mixed layer depth. The units of each variable are as follows: Bottom and Surface Temperature ( $^{\circ}\text{C}$ ), Bottom and Surface Current ( $\text{m s}^{-1}$ ), Slope ( $^{\circ}$ ), MLD (m); salinity is unitless

of the Gulf Stream whose influence extends to the seafloor both off Florida in water depths of 700 m and further north on the Blake Plateau where depths reach 1000 m (Richardson 2019). While temperatures associated with this current are projected to increase in the future (Fig. S5), the associated projected occurrences remained relatively stable (Fig. 2), suggesting that other variables or interactions between variables may have a greater influence in this subarea. The functional response curves for both models (Fig. 7b) showed little change in the relative likelihood of occurrence predicted/projected for the present-day and future forecasted conditions, especially for GAM. Of the 4 environmental predictors included in GAM (Mean Bottom Temperature, Mean Surface Current, Slope, and Mean Surface Temperature), the static variable Slope was considered the most important (Fig. S6), possibly explaining why spatial patterns in predicted/projected relative likelihood of occurrence and suitable habitat (Figs. 2 & 3) remained relatively unchanged between present-day and future forecasted conditions.

Overall, the performance of both Random Forest and GAM models in this subarea was relatively poor (Table 3; Table S5), with poor congruence between the spatial predictions of relative likelihood of occurrence and the location of presence points under present-day environmental conditions (especially for GAM; see Figs. 1 & 2), suggesting that any inferences derived from these models should be taken with caution.

#### 4. DISCUSSION

The results of our study show that under future climatic conditions, the potential suitable habitat of the glass sponge *Vazella pourtalesii* will increase in the northwest Atlantic. Using Random Forest and GAM techniques, we projected a gain in habitat of up to ~4 times its present-day size using environmental conditions simulated under moderate (RCP 4.5) and worst-case (RCP 8.5) CO<sub>2</sub> emission scenarios for 2046–2065. In the northern portion of this species' range, we projected a shift into deeper waters and higher latitudes, a pattern consistent with previous studies of the effects of climate change on marine species in the region (Nye et al. 2009, Greenan et al. 2019, Morato et al. 2020). This shift in distribution was likely due to the warming of the colder waters surrounding its current habitat, and the availability of new, unoccupied niche for this deep-water, originally subtropical species (Schmidt 1870).

##### 4.1. Effects of climate change on sponges: a deep-water example

Fitness and adaptability of individuals within the same population often differ between individuals located at the core and leading/trailing edges of its range, resulting in observed contractions and expansions at the trailing (southern) and leading (northern/colder) edges of its distribution, respectively, in response to climate change (Rilov et al. 2019). Our projected gain in *V. pourtalesii* suitable habitat in the future, particularly in the northern portion of its range where a colder climate prevails, is not entirely unexpected. Its current distribution on the Scotian Shelf is highly associated with areas that experience regular incursion of Warm Slope Water (Beazley et al. 2018), a warmer, saltier water mass originating from the Gulf Stream. Under a CO<sub>2</sub> doubling scenario, Saba et al. (2016) predicted that the enhanced warming on the northwest Atlantic shelf was the result of increased incursion of Warm Slope Water into the region due to a northerly shift in the Gulf Stream and the retreat of the cold and fresh Labrador Current. The Laurentian Channel/Gulf of St. Lawrence, where significant gains in suitable habitat are projected, is projected to warm at least 1°C (see Fig. S4 and Table S7), shifting the climate closer to the temperature preference of *V. pourtalesii* as defined by its present-day distribution (~6°C; Fig. 7a). The southern range of *V. pourtalesii* did not contract under further warming as expected, and was instead projected (by Random Forest) to nearly double in size under RCP 4.5 conditions. However, no further expansion was projected by this model under RCP 8.5. The area predicted/projected as suitable habitat by GAM was similar between the present-day and future climatologies, but showed a slight decrease in size from RCP 4.5 to 8.5, suggesting that this species does not benefit from further warming in the warmer part of its range.

In contrast to the Northeast US/Atlantic Canada subarea, the mechanism for the projected shift of *V. pourtalesii* into deeper waters in the southern portion of its range may be due to the warming of waters beyond its upper thermal tolerance limit, causing it to retreat into deeper, cooler waters. Although both Random Forest and GAM models showed a loss of suitable habitat along the flanks of its core distribution, a clear extension of the leading (cooler) edge and decline of the trailing (warmer) edge of its core habitat was not apparent in this subarea (Fig. 6). Both models predicted poorly in that area, possibly due to the low number of presence points (18 presences).



Wisz et al. (2008) evaluated the predictions of 12 different modelling algorithms applied to 46 species occurring at 3 different frequencies (10, 30, and 100 presence records), and found that no model predicted consistently well when sample size was less than 30. This suggests that model predictions/projections in this subarea should be considered for exploratory purposes only until additional information on the distribution of *V. pourtalesii* is collected.

Species affiliated with narrow environmental envelopes are often easier to model than those distributed over wide environmental gradients or mosaics (McPherson & Jetz 2007, Tsoar et al. 2007). High environmental variability associated with only a few *V. pourtalesii* presences (18) and/or local heterogeneity that is difficult to capture in SDMs could possibly explain the poor model results in the Mid–Southeast US subarea. For instance, the densest aggregations of *V. pourtalesii* off Florida co-occurred with vast regions of dead *Lophelia pertusa* reef. Niche overlap between *L. pertusa* reefs and sponges has been previously noted (van Oevelen et al. 2018), and the use of *L. pertusa* as settlement substrate by *V. pourtalesii* may be to take advantage of the current-induced flow and enhanced food supply due to interactions between reef topography and local hydrodynamic conditions (Davies et al. 2009, Mohn et al. 2014). Such small-scale associations would not be accurately captured by our coarser-scale data (Austin & Van Niel 2011).

Studies of shallow-water sponges have shown that while many appear to be relatively resilient to the combined effects of ocean warming and acidification compared to other benthic taxa, their tolerance to various climate stressors is highly species-specific (Bell et al. 2018). Bell et al. (2018) conducted a review of studies on the singular and combined effects of ocean warming and acidification on sponges and found that of the 44 studies focused solely on ocean warming, 30 demonstrated negative effects to host physiology, including pumping rate, filtration efficiency, choanocyte chamber size, and density; to gene expression, feeding ecology, and reproductive output; and to microbial community composition and function. Ramsby et al. (2018) examined the effects of incrementally increasing ambient seawater temperature on the bioeroding sponge *Cliona orientalis* and its symbiotic dinoflagellates (*Symbiodinium*), and found little effect of raising temperatures 2°C above monthly mean values. However, at 3°C above average (consistent with temperatures predicted to occur in year 2100 under RCP 8.5), *C. orientalis* bleached and reduced its energy reserves, consistent with the response of sympatric corals. Bennett et al. (2017) demonstrated via *ex situ* experimen-

tation a tolerance of several Great Barrier Reef sponges to ocean temperatures forecasted under moderate (RCP 6.0; 3°C above ambient), but not high (RCP 8.5; 4.5°C above ambient) CO<sub>2</sub> emission scenarios.

We modelled changes in the distribution of the deep-water sponge *V. pourtalesii* under moderately strong climate policy (RCP 4.5) and worst-case (RCP 8.5) CO<sub>2</sub> emission scenarios, where forecasted anomalies in Mean Bottom Temperature in the Northeast US/Atlantic Canada subarea in 2046–2065 were, on average,  $+0.57 \pm 0.37^\circ\text{C}$  (SD) and  $+0.80 \pm 0.56^\circ\text{C}$  under RCP 4.5 and 8.5, respectively. In the core of this species' habitat in the Scotian Gulf and Northeast Channel, such anomalies would raise average bottom temperatures by a maximum of  $\sim 2^\circ\text{C}$  (to  $\sim 10^\circ\text{C}$  under RCP 8.5) from their present-day average maximum ( $\sim 8^\circ\text{C}$ ), which is below the thermal change often shown to cause deleterious effects in shallow-water sponges ( $\sim 3^\circ\text{C}$  above ambient). Emerald Basin, where *V. pourtalesii* forms the densest known aggregations of its kind, is subjected to strong inter-annual and multi-decadal variability in water mass characteristics, where empirical maximum bottom temperatures of  $12^\circ\text{C}$  were recorded from CTD and Argo float data collected between 1950 and 2015 (Beazley et al. 2018). Exposure to short-term fluctuations in environmental conditions may have important consequences for species' ranges (Wetthey et al. 2011), and have shown to cause mass mortality events in deep-water sponge grounds elsewhere (Guihen et al. 2012). Our use of average conditions over monthly or seasonal data likely do not capture the overall temperature range experienced by these sponge grounds historically, nor the extremes that will likely be incurred in the future. Nonetheless, our correlative SDMs using mean climatic conditions show that while the population may persist into the future, further warming of *V. pourtalesii*'s core habitat in Emerald Basin may result in a reduction in its relative likelihood of occurrence there compared to the present day. While it has long been assumed that the deep sea is a highly stable environment relatively buffered from the effects of short-term changes in atmospheric or surface conditions, recent studies suggest that deep-sea ecosystems may be relatively sensitive to even seasonal shifts in upper ocean conditions (Glover et al. 2010). As the bottom temperatures forecasted under RCP 4.5 and 8.5 in the core habitat of *V. pourtalesii* are not outside the range experienced historically by these sponge grounds, or warmer than the conditions associated with *V. pourtalesii* presence in the southern portion of its range, the reduction in relative likelihood of occurrence suggests that

even marginal increases in temperature in the deep sea may have deleterious impacts on animal physiology and habitat quality, and/or that variables other than bottom temperature, or possibly synergies between variables, may be responsible for this decline.

The occurrence of *V. pourtalesii* in certain localities off the mid–southeast USA, where empirical maximum bottom temperatures of 11.2°C were recorded (Table S1), points to an ability of this species to acclimate to temperatures higher than the average conditions experienced on the Scotian Shelf, suggesting some hope for the persistence of the dense sponge grounds that reside there. However, the densest aggregations of *V. pourtalesii* in the Mid–Southeast US subarea (up to 663 individuals on a single transect) were typically associated with bottom temperatures no higher than 8.6°C, suggesting that individuals found at ~12°C may be more representative of marginal than central populations, on the brink of their thermal tolerance limit (Bennett et al. 2019). The formation of sponge grounds at these higher temperatures is likely less probable. Evaluation of the Random Forest and GAM response curves for the Mid–Southeast US subarea (Fig. 7b) also support a preference for bottom temperatures <12°C (maximum relative likelihood of occurrence at ~10°C). Interpretation of our model results would greatly benefit from laboratory-based experiments where the physiological effects of incrementally increasing temperatures beyond 10°C are measured, and the upper thermal tolerance of *V. pourtalesii* definitively identified.

#### 4.2. Limitations and importance of environmental interpretation

While we attempted to ensure a similar prevalence (i.e. proportion of observed presences) between data partitions (folds) used in 5-fold spatial-block cross-validation, the lower number of presences and higher number of pseudo-absences, along with their spatial distribution, had an impact on the model outputs for GAM (but not for Random Forest). For irregularly sampled data, spatially clustered data, or those with a highly unbalanced prevalence such as ours, Roberts et al. (2017) suggested non-gridded and/or irregularly-shaped blocks (i.e. pie slices) to ensure even sampling of presence and absence data, but also cautioned that non-regular blocks may not address the issue of autocorrelation consistently. Caution must also be taken when using spatial blocking, as blocking structures that follow environmental gradients can lead to entire portions of envi-

ronmental predictor space (i.e. ranges and/or combinations of predictor variables) being held out in the testing dataset, resulting in extrapolation between folds (Roberts et al. 2017). The CV GAM model trained on the data partitioned in Fold 3 led to erroneously high projections of relative likelihood of occurrence in the lower Laurentian Channel, mouth of the Gulf of St. Lawrence estuary, and northeast Scotian Gulf, consistent with the locations of high model uncertainty. Similar applications of Random Forest modelling in the Gulf of St. Lawrence using groundfish trawl survey presences and absences (see Murillo et al. 2016) indicated an absence of *V. pourtalesii* in those areas. While we did not evaluate whether the CV model was extrapolating beyond the data used to train the model, unlike the data partitioned in the other folds, the dataset of Fold 3 excluded presences from the core sponge grounds in Emerald Basin. We recommend that alternate approaches to rectangular spatial blocking be investigated in the future for highly clustered species datasets such as those modelled here, to ensure uniform data selection between folds across the environmental domain of the species.

While our model results are an important first step in understanding the impacts of climate change on deep-water sponges, the caveats and uncertainties inherent to this particular type of habitat suitability modelling should be considered prior to the development of strategies and policies for the management of this species in the future. Climate trajectories inherently include a broad range of assumptions on future greenhouse gas emissions, population and economic growth, and technological change (van Vuuren et al. 2011), the uncertainties of which are carried forward into ecological models (Payne et al. 2016). RCP 8.5 trajectories are associated with the highest level of uncertainty, and changes in projected distribution under this RCP should be considered the worst-case scenario. Many factors can influence the response of a species to changes in the environment, including recruitment dynamics, physiological tolerances, food availability, competition, and community composition of existing or receiving communities (Poloczanska et al. 2016), which are difficult to capture in single-species models such as those presented here.

Correlative SDMs such as those applied here are based on the statistical association between occurrence records and spatial environmental data, and inherently capture complex interactions between a species and its current environment. When projecting future distributions, these interactions are assumed to be preserved in the new environmental

space, an assumption that may not be valid under future climate change where novel environmental scenarios and non-equilibrium species distributions are likely (Kearney et al. 2010). Furthermore, we defined the habitat of *V. pourtalesii* in relation to only a few dynamic environmental variables (bottom and surface temperature, salinity, current velocity, and MLD). Other abiotic factors not captured here may be equally as important for the distribution of this species, such as nutrient availability and primary production (Beazley et al. 2018, Kazanidis et al. 2019), both of which are expected to decrease on the Scotian Shelf in the future (Pepin et al. 2013, Lavoie et al. 2019).

Further interpretation of the results of these models with other information important for species' distributions is necessary in order to evaluate their uncertainty. For instance, while both Random Forest and GAM were consistent in projecting gains in suitable habitat in the deeper waters of the Laurentian Channel and Gulf of St. Lawrence, the direction of flow of oceanographic currents in the future, which was not captured by our models, does not support the transport of larvae beyond the edge of the Scotian Shelf. Although the east-to-west flow of the Labrador Current along the edge of the Scotian Shelf is predicted to weaken (Saba et al. 2016), no significant transport from the Scotian Gulf to the Laurentian Channel is anticipated under future climate change (see Fig. 6 of Saba et al. 2016). In contrast, the proportion of warm and salty Slope Water entering the Gulf of Maine through the Northeast Channel is predicted to increase (Saba et al. 2016), suggesting that the gain in suitable habitat in the Gulf of Maine projected by Random Forest (Fig. 6) is a much more likely event. Furthermore, the presence of hard substrate is crucial for the settlement of *V. pourtalesii*, which attaches to substrate ranging in size from pebbles to boulders (Hawkes et al. 2019). The Laurentian Channel is dominated by mud and clay substrate (DFO 2011), further reiterating that the settlement and proliferation of this species there is unlikely. This highlights the need for further interpretation with other factors not included in these models that may be important for the spatial distribution of a species.

#### 4.3. Deep-water sponges as 'climate change winners'?

The predicted impacts of climate change on the home ranges of other marine species residing on the continental shelves of the northeast USA and Atlantic Canada are, to date, highly species-specific, with

emergence of clear 'winners' and 'losers' (Greenan et al. 2019, McHenry et al. 2019, Morato et al. 2020). For the distribution of 125 benthic invertebrates, and pelagic and demersal fish on the northeast USA shelf, McHenry et al. (2019) predicted more losers than winners, with those species highly associated with the seabed (benthic invertebrates and demersal fish) experiencing the most severe habitat loss and/or fragmentation. In contrast, highly mobile species such as pelagic fish, cephalopods, and elasmobranchs showed a greater ability to shift their home ranges north to help mitigate climate-induced effects. For long-lived, habitat-forming cold-water corals in the North Atlantic (including our study area), Morato et al. (2020) predicted severe habitat losses ranging between 28 and 100%. These results suggest that the impacts on benthic or sedentary species may be more severe than on highly mobile species, varying by taxa likely as a consequence of longevity, reproductive periodicity, and dispersal capacity. Nonetheless, the projections of McHenry et al. (2019) and Morato et al. (2020) were made using a different suite of environmental variables and for a later time period than modelled here, when the severity of climate-induced effects on marine habitats will likely be higher.

If based solely on the projected expansion in suitable habitat in the future, our results may lend further weight to the concept that sponges, even in deep-sea environments, are potential 'winners' against climate change compared to other benthic groups. Whether the distribution of boreal or arctic sponge grounds in the North Atlantic will benefit from ocean warming in the same way remains unknown. Boreal assemblages may have some ability to shift their distribution north to help mitigate the effects of ocean warming. However, a laboratory-based study of the effects of increasing temperatures on the transcriptomic response of the extremophile Antarctic demosponge *Isodictya* sp. (González-Aravena et al. 2019) suggests that cold-adapted sponges may have an even more limited ability to tolerate increased temperatures compared to warm-adapted species. This suggests that Arctic assemblages located at their maximum latitudinal limit may therefore experience an overall loss of habitat due to contraction of their southern range as temperatures warm. Nonetheless, the local formation of sponge grounds has been strongly linked to factors other than temperature, such as the occurrence of internal tidal waves at the seabed (Klitgaard et al. 1997, Klitgaard & Tendal 2004) and the presence of water mass fronts (Klitgaard & Tendal 2004, Roberts et al. 2018, Meyer et al. 2019), which should be accounted for when making inferences of

climate-induced distribution changes. Such characteristics are often poorly resolved in climate forecasting models.

#### 4.4. Implications for conservation management

The size of the *V. pourtalesii* core habitat in the Scotian Gulf and Northeast Channel remained relatively unchanged between our present-day predictions and future projections (Fig. 6), suggesting that these areas may serve as important refugia into the future. However, the projected relative likelihood of occurrence of *V. pourtalesii* under both RCP scenarios was lower in these areas compared to present-day conditions (see Random Forest outputs; Fig. 2), including at the location of the 2 bottom-fishery closures implemented by DFO in 2013 for the protection of this species from bottom-fishing activities. While relative likelihood of occurrence or presence probability is not a direct correlate of abundance or biomass, it could be considered a proxy for habitat quality, where a higher-quality environment would presumably support higher species' abundances. In 2013, DFO closed 2 areas equating to nearly 260 km<sup>2</sup> in Emerald Basin to all bottom-tending fishing gears to protect 2 of the 5 densest concentrations of *V. pourtalesii* in the Scotian Gulf. As the purpose of the 2 sponge conservation areas is to protect the most significant concentrations (based on biomass/density) of this species, range shifts and potential abundance reductions under future climate change present a significant challenge for maintaining the effectiveness of these fishery closures. Effective conservation management of *V. pourtalesii* will require an iterative monitoring programme designed to track changes in its density and distribution within and outside these closed areas, with emphasis placed on the margins between core areas and those where a loss or gain of habitat are projected.

## 5. CONCLUSIONS

The results of our correlative SDMs show that the suitable habitat of the deep-water glass sponge *V. pourtalesii* will expand in the northwest Atlantic under future climate change, particularly in the northern portion of its range (Atlantic Canada), where ocean warming will serve to improve the conditions surrounding the current habitat of this species and increase the availability of currently unoccupied niche. While not all of these areas are likely to be populated

in the future, any expansion of *V. pourtalesii*'s niche will likely have positive implications for other benthic species, which were shown to be more abundant and diverse in the presence of this habitat-forming species (Hawkes et al. 2019), a function they have in common with other sponge grounds across the North Atlantic (Klitgaard 1995, Bo et al. 2012, Beazley et al. 2013, 2015). Although our projected expansion of suitable *V. pourtalesii* habitat suggests that this species may be a 'winner' against future climate change, the reduction in relative likelihood of occurrence in its core habitat on the Scotian Shelf suggests that the *Vazella* sponge grounds that reside there may experience losses as the northwest Atlantic continues to warm. We recommend that future studies aimed at elucidating the effects of climate change on the distribution of *V. pourtalesii* combine mechanistic approaches based on functional traits and physiological constraints with correlative SDM techniques (see Kearney et al. 2010) such as those presented here, in order to strengthen their results. Nonetheless, our results are an important first step in evaluating the impacts of climate change on a ground-forming deep-water sponge, serve as a basis for hypothesis testing in future laboratory-based physiological studies, and provide a starting point for its effective conservation management in light of climate change.

While the impacts of climate change on the distribution, abundance, and composition of other sponge grounds in the North Atlantic remain unknown, given their importance to various ecological functions, any positive or negative impacts of anthropogenically induced change are likely to have cascading effects throughout these benthic ecosystems.

*Acknowledgements.* This research was funded through the Fisheries and Oceans Canada (DFO) Strategic Program for Ecosystem-Based Research and Advice (SPERA) project 'Evaluation of the Effectiveness of Two Sponge Conservation Areas in the Maritimes Region: Identifying Patterns of Dispersal, Connectivity, and Recovery Potential of the Russian Hat Sponge *Vazella pourtalesii*' led by L.B. and E.K. and the H2020 EU Framework Programme for Research and Innovation Project SponGES (Deep-sea Sponge Grounds Ecosystems of the North Atlantic: an integrated approach towards their preservation and sustainable exploitation) (Grant Agreement no. 679849). This document reflects only the authors' views, and the Executive Agency for Small and Medium-sized Enterprises (EASME) is not responsible for any use that may be made of the information it contains. Data/imagery from the NOAA Deep-Sea Coral Data Portal and the NOAA Okeanos Explorer EX1806 and EX1903 missions courtesy of NOAA's Office of Ocean Exploration and Research. This study is in memory of our colleague and co-author Prof. H. T. Rapp, the coordinator of project SponGES, who passed away in March 2020.

## LITERATURE CITED

- ✦ Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43: 1223–1232
- ✦ Austin MP, Van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. *J Biogeogr* 38:1–8
- ✦ Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol Evol* 3: 327–338
- ✦ Beazley LI, Kenchington EL, Murillo FJ, Sacau MM (2013) Deep-sea sponge grounds enhance diversity and abundance of epibenthic megafauna in the Northwest Atlantic. *ICES J Mar Sci* 70:1471–1490
- ✦ Beazley L, Kenchington E, Yashayaev I, Murillo, FJ (2015) Drivers of epibenthic megafaunal composition in the sponge grounds of the Sackville Spur, northwest Atlantic. *Deep Sea Res I* 98:102–114
- Beazley L, Kenchington, E, Murillo FJ, Lirette C and others (2016) Species distribution modelling of corals and sponges in the Maritimes region for use in the identification of significant benthic areas. *Can Tech Rep Fish Aquat Sci* 3172:vi + 189p
- ✦ Beazley L, Wang, Z, Kenchington E, Yashayaev I and others (2018) Predicted distribution of the glass sponge *Vazella pourtalesi* on the Scotian Shelf and its persistence in the face of climatic variability. *PLOS ONE* 13:e0205505
- ✦ Bell JJ, Davy SK, Jones T, Taylor MW, Webster NS (2013) Could some coral reefs become sponge reefs as our climate changes? *Glob Change Biol* 19:2613–2624
- ✦ Bell JJ, Bennett HM, Rovellini A, Webster NS (2018) Sponges to be winners under near-future climate scenarios. *BioScience* 68:955–968
- ✦ Bennett HM, Altenrath C, Woods L, Davy SK, Webster NS, Bell JJ (2017) Interactive effects of temperature and pCO<sub>2</sub> on sponges: from the cradle to the grave. *Glob Change Biol* 23:2031–2046
- ✦ Bennett S, Duarte CM, Marbà N, Wernberg T (2019) Integrating within-species variation in thermal physiology into climate change ecology. *Philos Trans R Soc B* 374: 20180550
- ✦ Bo M, Bertolino M, Bavestrello G, Canese S and others (2012) Role of deep sponge grounds in the Mediterranean Sea: a case study in northern Italy. *Hydrobiologia* 687:167–177
- Breeze H, Horsman T (eds) (2005) *The Scotian Shelf: an atlas of human activities*. Oceans and Coastal Management Division, Fisheries and Oceans Canada, Dartmouth, NS. <https://waves-vagues.dfo-mpo.gc.ca/Library/321387.pdf> (accessed 9 July 2020)
- ✦ Breiman L (2001a) Random forests. *Mach Learn* 45:5–32
- ✦ Breiman L (2001b) Statistical modeling: the two cultures. *Stat Sci* 16:199–231
- Breiman L, Friedman JH, Olshen R, Stone CJ (1984) *Classification and regression trees*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA
- 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:xiv + 143 pp
- Bush E, Lemmen DS (eds) (2019) *Canada's changing climate report*. Government of Canada, Ottawa. [https://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/energy/Climate-change/pdf/CCCR\\_FULLREPORT-EN-FINAL.pdf](https://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/energy/Climate-change/pdf/CCCR_FULLREPORT-EN-FINAL.pdf)
- Carballo JL, Bell JJ (eds) (2017) *Climate change, ocean acidification and sponges: impacts across multiple levels of organization*. Springer International Publishing, Cham
- Davies AJ, Duineveld GCA, Lavaleye MSS, Bergman, MJN, van Haren H, Roberts JM (2009) Downwelling and deep-water bottom currents as food supply mechanisms to the cold-water coral *Lophelia pertusa* (Scleractinia) at the Mingulay Reef Complex. *Limnol Oceanogr* 54:620–629
- DFO (Fisheries and Oceans Canada) (2011) *Biophysical overview of the Laurentian Channel Area of Interest (AOI)*. DFO Can Sci Advis Secret Sci Advis Rep 2010/076. <https://waves-vagues.dfo-mpo.gc.ca/Library/343558.pdf>
- ESRI (2019) *ArcGIS Desktop: Release 10.7*. Environmental Systems Research Institute, Redlands, CA
- ✦ Freeman EA, Moisen G (2008) PresenceAbsence: an R package for presence absence analysis. *J Stat Softw* 23:1–31
- Fuller SD, Murillo Perez FJ, Wareham V, Kenchington E (2008) Vulnerable marine ecosystems dominated by deep-water corals and sponges in the NAFO Convention Area. *NAFO SCR Doc 08/22 Ser No N5524*. <https://www.nafo.int/Portals/0/PDFs/sc/2008/scr08-022.pdf>
- ✦ Glover AG, Gooday AJ, Bailey DM, Billett DSM and others (2010) Temporal change in deep-sea benthic ecosystems: a review of the evidence from recent time-series studies. *Adv Mar Biol* 58:1–95
- ✦ González-Aravena M, Kenny NJ, Osorio M, Font A, Riesgo A, Cárdenas CA (2019) Warm temperatures, cool sponges: the effect of increased temperatures on the Antarctic sponge *Isodictya* sp. *PeerJ* 7:e8088
- ✦ González-Irusta JM, González-Porto M, Sarralde R, Arrese B, Almón B, Martín-Sosa P (2015) Comparing species distribution models: a case study of four deep sea urchin species. *Hydrobiologia* 745:43–57
- ✦ Greenan BJW, Shackell NL, Ferguson K, Greyson P and others (2019) Climate change vulnerability of American lobster fishing communities in Atlantic Canada. *Front Mar Sci* 6:579
- ✦ Guihen D, White M, Lundälv T (2012) Temperature shocks and ecological implications at a cold-water coral reef. *Mar Biodivers Rec* 5:e68
- ✦ Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, Gordon A and others (2015) Is my species distribution model fit for purpose? Matching data and models to applications. *Glob Ecol Biogeogr* 24:276–292
- ✦ Guinotte JM, Davies AJ (2014) Predicted deep-sea coral habitat suitability for the US West Coast. *PLOS ONE* 9: e93918
- ✦ Hare JA, Morrison WE, Nelson MW, Stachura MM and others (2016) A vulnerability assessment of fish and invertebrates to climate change on the northeast US Continental Shelf. *PLOS ONE* 11:e0146756
- ✦ Hastie TJ, Tibshirani RJ (1986) Generalized additive models. *Stat Sci* 1:297–310
- ✦ Hawkes N, Korabik M, Beazley L, Rapp HT, Xavier JR, Kenchington E (2019) Glass sponge grounds on the Scotian Shelf and their associated biodiversity. *Mar Ecol Prog Ser* 614:91–109
- ✦ Hengl T, Nussbaum M, Wright MN, Heuvelink GBM, Gräler B (2018) Random forest as a generic framework for predicting modeling of spatial and spatio-temporal variables. *PeerJ* 6:e5518
- Honeyman D (1889) Glacial boulders of our fisheries and invertebrates, attached and detached. *Proc Trans N S*

- Inst Nat Sci 7:205–213
- ✦ Howell KL, Piechaud N, Downie AL, Kenny A (2016) The distribution of deep-sea sponge aggregations in the North Atlantic and implications for their effective spatial management. *Deep Sea Res I* 115:309–320
- IPCC (2013) Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
- James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning. Springer, New York
- ✦ Kazanidis G, Vad J, Henry LA, Neat F, Berx B, Georgoulas K, Roberts JM (2019) Distribution of deep-sea sponge aggregations in an area of multisectoral activities and changing oceanic conditions. *Front Mar Sci* 6:163
- ✦ Kearney MR, Wintle BA, Porter WP (2010) Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. *Conserv Lett* 3: 203–213
- ✦ Kleisner KM, Fogarty MJ, McGee S, Hare JA, Moret S, Perretti CT, Saba VS (2017) Marine species distribution shifts on the US Northeast Continental shelf under continued ocean warming. *Prog Oceanogr* 153:24–36
- ✦ Klitgaard AB (1995) The fauna associated with outer shelf and upper slope sponges (Porifera, Demospongiae) at the Faroe Islands, northeastern Atlantic. *Sarsia* 80: 1–22
- ✦ Klitgaard AB, Tendal OS (2004) Distribution and species composition of mass occurrences of large-sized sponges in the northeast Atlantic. *Prog Oceanogr* 61:57–98
- Klitgaard AB, Tendal OS, Westerberg H (1997) Mass occurrences of large sponges (Porifera) in Faroe Island (NE Atlantic) shelf and slope areas: characteristics, distribution and possible causes. In: Hawings LE, Hutchinson S (eds) The responses of marine organisms to their environments. Proc 30<sup>th</sup> Eur Mar Biol Symp, University of Southampton, UK, p 129–142
- Kuhn M (2020) caret: classification and regression training. R package version 6.0-86. <https://CRAN.R-project.org/package=caret>
- ✦ Kutti T, Bannister RJ, Fosså JH (2013) Community structure and ecological function of deep-water sponge grounds in the Traenadypet MPA—northern Norwegian continental shelf. *Cont Shelf Res* 69:21–30
- ✦ Lavoie D, Lambert N, Gilbert D (2019) Projections of future trends in biogeochemical conditions in the northwest Atlantic using CMIP5 Earth Systems Models. *Atmos Ocean* 57:18–40
- ✦ Leys SP, Meech RW (2006) Physiology of coordination in sponges. *Can J Zool* 84:288–306
- Liaw A, Wiener M (2002) Classification and regression by randomForest. *R News* 2–3:18–22
- ✦ Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28:385–393
- ✦ Lopes PFM, Verba JT, Begossi A, Pennino MG (2019) Predicting species distribution from fishers' local ecological knowledge: a new alternative for data-poor management. *Can J Fish Aquat Sci* 76:1423–1431
- Maldonado M, Aguilar R, Bannister RJ, Bell JJ and others (2017) Sponge grounds as key marine habitats: a synthetic review of types, structure, functional roles, and conservation concerns. In: Rossi S, Bramanti L, Gori A, Orejas Saco del Valle C (eds) Marine animal forests: the ecology of benthic biodiversity hotspots. Springer International Publishing, Cham, p 145–183
- ✦ McHenry J, Welch H, Lester SE, Saba V (2019) Projecting marine species range shifts from only temperature can mask climate vulnerability. *Glob Change Biol* 25:4208–4221
- ✦ McPherson JM, Jetz W (2007) Effects of species' ecology on the accuracy of species distribution models. *Ecography* 30:135–151
- ✦ Meyer HK, Roberts EM, Rapp HT, Davies AJ (2019) Spatial patterns of arctic sponge ground fauna and demersal fish are detectable in autonomous underwater vehicle (AUV) imagery. *Deep Sea Res I* 153:103137
- ✦ Mohn C, Rengstorf A, White M, Duineveld G, Mienis F, Soetaert K, Grehan A (2014) Linking benthic hydrodynamics and cold-water coral occurrences: a high-resolution model study at three cold-water coral provinces in the NE Atlantic. *Prog Oceanogr* 122:92–104
- ✦ Morato T, González-Irusta JM, Dominguez-Carrió C, Wei CL and others (2020) Climate-induced changes in the habitat suitability of cold-water corals and commercially important deep-sea fishes in the North Atlantic. *Glob Change Biol* 26:2181–2202
- ✦ Moss RH, Edmonds JA, Hibbard KA, Manning MR and others (2010) The next generation of scenarios for climate change research and assessment. *Nature* 463:747–756
- Murillo FJ, Kenchington E, Beazley L, Lirette C and others (2016) Distribution modelling of sea pens, sponges, stalked tunicates and soft corals from research vessel survey data in the Gulf of St. Lawrence for use in the identification of significant benthic areas. *Can Tech Rep Fish Aquat Sci* 3170:vi + 132p
- ✦ Murillo FJ, Kenchington E, Koen-Alonso M, Guijarro J and others (2020) Mapping ecological diversity and interactions with bottom-contact fishing on the Flemish Cap (northwest Atlantic). *Ecol Indic* 112:106135
- Nephtin J, Gregr EJ, St. Germain C, Fields C, Finney JL (2020) Development of a species distribution modelling framework and its application to twelve species on Canada's Pacific coast. DFO Can Sci Advis Sec Res Doc. 2020/004: xii + 107 p. <https://waves.vagues.dfo-mpo.gc.ca/Library/40875325.pdf>
- NOAA (2019) National Database for Deep-Sea Corals and Sponges (version 20190117-0). NOAA Deep Sea Coral Research & Technology Program. <https://deepseacoral-data.noaa.gov/> (accessed 18 Oct 2019)
- ✦ Nye JA, Link JS, Hare JA, Overholtz WJ (2009) Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. *Mar Ecol Prog Ser* 393:111–129
- ✦ Payne MR, Barange M, Cheung WWL, MacKenzie BR and others (2016) Uncertainties in projecting climate-change impacts in marine ecosystems. *ICES J Mar Sci* 73: 1272–1282
- Pepin P, Maillet GL, Lavoie D, Johnson C (2013) Temporal trends in nutrient concentrations in the northwest Atlantic basin. In: Loder JW, Han G, Galbraith PS, Chassé J, van der Baaren A (eds) Aspects of climate change in the Northwest Atlantic off Canada. *Can Tech Rep Fish Aquat Sci* 3045, p 127–150
- ✦ Pham CK, Murillo FJ, Lirette C, Maldonado M, Colaço A, Ottaviani D, Kenchington E (2019) Removal of deep-sea sponges by bottom trawling in the Flemish Cap area: conservation, ecology and economic assessment. *Sci Rep* 9:15843
- ✦ Pinsky ML, Fogarty M (2012) Lagged socio-ecological responses to climate and range shifts in fisheries. *Clim*

- Change 115:883–891
- ✦ Poloczanska ES, Burrows MT, Brown CJ, García Molinos J and others (2016) Responses of marine organisms to climate change across oceans. *Front Mar Sci* 3:62
- R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- ✦ Ramsby BD, Hoogenboom MO, Smith HA, Whalan S, Webster NS (2018) The bioeroding sponge *Cliona orientalis* will not tolerate future projected ocean warming. *Sci Rep* 8:8302
- Reiswig HM (1996) Redescription and placement of the rossellid genus *Vazella* Gray (Hexactinellida: Lyssacinosa). *Bull Inst R Sci Nat Belg* 66(Suppl):135–141
- ✦ Rheuban JE, Doney SC, Cooley SR, Hart DR (2018) Predicted impacts of future climate change, ocean acidification, and management on the US Atlantic sea scallop (*Placopecten magellanicus*) fishery. *PLOS ONE* 13: e0203536
- Richardson PL (2019) Florida Current, Gulf Stream, and Labrador Current. In: Cochran JK, Bokuniewicz HJ, Yager PL (eds) *Encyclopedia of ocean sciences*, Vol 1, 3rd edn. Marine biogeochemistry. Elsevier, London, p 351–361
- ✦ Rilov G, Mazaris AD, Stelzenmüller V, Helmuth B and others (2019) Adaptive marine conservation planning in the face of climate change: What can we learn from physiological, ecological and genetic studies? *Glob Ecol Conserv* 17:e00566
- ✦ Roberts DR, Bahn V, Ciuti S, Boyce MS and others (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40:913–929
- ✦ Roberts EM, Mienis F, Rapp HT, Hanz U, Meyer HK, Davies AJ (2018) Oceanographic setting and short-timescale environmental variability at an Arctic seamount sponge ground. *Deep Sea Res I* 138:98–113
- ✦ Saba VS, Griffies SM, Anderson WG, Winton M and others (2016) Enhanced warming of the Northwest Atlantic Ocean under climate change. *J Geophys Res Oceans* 121:118–132
- Schmidt EO (1870) *Grundzüge einer Spongien-Fauna des atlantischen Gebietes*. Wilhelm Engelmann, Leipzig
- ✦ Stanley RRE, DiBacco C, Lowen B, Beiko RG and others (2018) A climate-associated multispecies cryptic cline in the northwest Atlantic. *Sci Adv* 43:eaaq0929
- ✦ Tanaka KR, Belknap SL, Homola JJ, Chen Y (2017) A statistical model for monitoring shell disease in inshore lobster fisheries: a case study in Long Island Sound. *PLOS ONE* 12:e0172123
- ✦ Tsoar A, Allouche O, Steinitz O, Rotem D, Kadmon R (2007) A comparative evaluation of presence-only methods for modelling species distribution. *Divers Distrib* 13:397–405
- ✦ Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2019) BlockCV: an R package for generating spatially or environmentally separated folds for *k*-fold cross-validation of species distribution models. *Methods Ecol Evol* 10:225–232
- ✦ VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecol Model* 220:589–594
- ✦ van Oevelen D, Mueller CE, Lundälv T, van Duyl FC, de Goeij JM, Middelburg JJ (2018) Niche overlap between a cold-water coral and an associated sponge for isotopically-enriched particulate food sources. *PLOS ONE* 13: e0194659
- ✦ van Vuuren DP, Edmonds J, Kainuma M, Riahi K and others (2011) The representative concentration pathways: an overview. *Clim Change* 109:5
- Wang Z, Lu Y, Greenan B, Brickman D (2018) BNAM: An eddy-resolving North Atlantic Ocean model to support ocean monitoring. *Can Tech Rep Hydrogr Ocean Sci* 327: vii + 18p
- ✦ Wang Z, Brickman D, Greenan BJW (2019) Characteristic evolution of the Atlantic Meridional Overturning Circulation from 1990 to 2015: an eddy-resolving ocean model study. *Deep Sea Res I* 149:103056
- ✦ Ward G, Hastie T, Barry S, Elith J, Leathwick JR (2009) Presence-only data and the EM algorithm. *Biometrics* 65: 554–563
- ✦ Wenger SJ, Olden JD (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods Ecol Evol* 3:260–267
- ✦ Wethey DS, Woodin SA, Hilbish TJ, Jones SJ, Lima FP, Brannock PM (2011) Response of intertidal populations to climate change: effects of extreme events versus long term change. *J Exp Mar Biol Ecol* 400:132–144
- ✦ Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group (2008) Effects of sample size on the performance of species distribution models. *Divers Distrib* 14:763–773
- Wood SN (2001) mgcv: GAMs and generalized ridge regression for R. *R News* 1:20–25
- ✦ Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc B* 73:3–36
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) *Mixed effects models and extensions in Ecology with R*. Springer Science+Business Media, Berlin

*Editorial responsibility: Myron Peck,  
Hamburg, Germany  
Reviewed by: 3 anonymous referees*

*Submitted: March 30, 2020  
Accepted: November 4, 2020  
Proofs received from author(s): December 18, 2020*