

Estimating the distribution of harvested estuarine bivalves with natural-history-based habitat suitability models

Nathaniel S. Lewis^{a,*}, Eric W. Fox^b, Theodore H. DeWitt^c

^a ORISE Research Fellow, Pacific Coastal Ecology Branch, Western Ecology Division, U.S. Environmental Protection Agency, Newport, OR, 97365, USA

^b Department of Statistics and Biostatistics, California State University, East Bay, Hayward, CA, 94542, USA

^c Pacific Coastal Ecology Branch, Western Ecology Division, U.S. Environmental Protection Agency, Newport, OR, 97365, USA



ARTICLE INFO

Keywords:

Habitat suitability index (HSI)
Geographic information system (GIS)
Bay-clams
NE Pacific

ABSTRACT

Habitat suitability models are useful to estimate the potential distribution of a species of interest, particularly in the case of infaunal bivalves. Sampling for these bivalves is time- and cost-intensive, which is increasingly difficult for organizations or agencies that are limited by personnel and funds. Consequently, we developed a framework to identify suitable bivalve habitat in estuaries (FISBHE) – a habitat suitability index (HSI) modeling framework for NE Pacific estuaries that was parameterized with published natural-history information and existing habitat datasets, without requiring extensive field sampling of bivalves. Spatially explicit, rule-based habitat suitability models were constructed in a GIS for five species of bay-clams (*Clinocardium nuttallii*, *Mya arenaria*, *Tresus capax*, *Saxidomus gigantea*, and *Leukoma staminea*) that are popular targets for recreational and commercial harvest in estuaries of the U.S. Pacific Northwest. Habitat rasters were produced for Yaquina and Tillamook estuaries (Oregon, USA) using environmental data (bathymetric depth, sediment % silt-clay, wet-season salinity, and burrowing shrimp presence/absence) from multiple studies (1953–2015). These habitat rasters then served as inputs in the final model which produced HSI classes ranging from 0 to 4 (lowest to highest suitability), dependent upon the number of habitat variables that fell within the sensitivity limits for each species of bay-clam. The models were tested with validation analyses and bay-clam occurrence data (reported in benthic community studies, 1996–2012) within each HSI class; logistic regression and Kendall's correlation coefficient both showed correspondence between predicted HSI classes and bay-clam presence/absence. Results also showed that the greatest presence probabilities occurred within habitats of highest predicted suitability, with the exception of *M. arenaria* in Tillamook Bay. The advantage of FISBHE is that disparate, independent sets of existing data are sufficient to parameterize the models, as well as produce and validate maps of habitat suitability. This approach can be transferred to data-poor systems with modest investment, which can be useful for prioritizing estuarine land-use decisions and could be used to estimate the vulnerability of this valued ecosystem good to changes in habitat quality and distribution.

1. Introduction

Associations between environmental variables and the presence, absence, or abundance of flora and fauna are central to understanding the distribution and abundance of organisms in all ecosystems. Habitat suitability models (also called species distribution models or ecological niche models) provide a means to quantitatively explain or predict the occurrence of organisms within or across habitats (Austin, 2002; Boyce et al., 2002; Elith and Graham, 2009; Hijmans, 2012) based on these habitat-organism associations, and are limited only by the range and precision of input data (Valavanis et al., 2008; Robinson et al., 2011). Such ecological tools can help identify spatial patterns in valuable

ecosystem goods and services, as well as habitats susceptible to biological invasion or environmental change (Ysebaert et al., 2002; Guisan and Thuiller, 2005; Lenoir et al., 2011). These types of models are diverse and often not transferrable, even within similar habitats, due to differences in parameterization, scale (i.e., resolution and extent; Guisan and Thuiller, 2005), or among species (Randin et al., 2006).

Most habitat suitability models quantify specific relationships (e.g., response curves) between empirically derived organism occurrence data (presence/absence or presence only) and corresponding environmental variable data, which are then used to explain or predict the occurrence or distribution of organisms across a landscape (Guisan and Zimmermann, 2000; Ysebaert et al., 2002; Olivier and Wotherspoon,

* Corresponding author. 2111 SE Marine Science Dr, Newport, OR, 97365, USA.

E-mail address: lewis.nathaniel@epa.gov (N.S. Lewis).

Table 1

Each input habitat variable was classified into subcategories, within which, the suitability for each species was determined from pertinent literature (bottom row denotes the references used for each species). Within each species column, an 'X' denotes that the intersecting habitat subcategory falls within the sensitivity range of that species.

Habitat Variable	Subcategories	Subcategory variable range	Cockle (<i>C. nuttallii</i>)	Softshell (<i>M. arenaria</i>)	Gaper (<i>T. capax</i>)	Butter (<i>S. gigantea</i>)	Littleneck (<i>L. staminea</i>)
Bathymetric depth (m; MLLW)	Subtidal	< -0.79	X		X	X	X
	Low-intertidal	(-0.78) - 0.37	X	X	X	X	X
	Mid-intertidal	0.38–1.56	X	X	X		X
	High-intertidal	> 1.56	X	X			
Sediment (% silt-clay)	Sand	< 7.51	X				
	Mud/Sand/Gravel	7.51–39.00	X	X	X	X	X
	Mud/Sand	39.01–60.00		X	X	X	X
	Mud	> 60.00		X	X		
Mean wet-season salinity (PSU)	Upper-estuary	< 16.00		X			
	Mid-estuary	16.01–23.00	X	X			
		23.01–27.00	X	X		X	X
	Lower-estuary	> 27.00	X	X	X	X	X
Burrowing shrimp (presence)	"Other" habitat	Absent	X	X	X	X	X
	Shrimp habitat	Present					
References		1,2,4,5,7,8, 9, 10,12,14, 15	2,5,7,9,10,13, 15	1,2,4,5,7,9, 10,11,15, 16,17	1,2,5,7,9,10, 15	1,2,3,5,6,7,9, 10,15	

References: ¹Bernard (1983), ²Coastal Biodiversity Risk Analysis Tool (CBRAT)**, ³Chew and Ma (1987), ⁴Coan et al., (2000), ⁵Dethier (2006), ⁶Dethier et al., (2012), ⁷Endreny and Sylvia, 2006, ⁸Gallucci and Gallucci (1982), ⁹Gaumer and McCrae (1990), ¹⁰Hancock et al., (1979), ¹¹Lauzier et al., (1998), ¹²Liu et al., (2010), ¹³Newell et al., (1986), ¹⁴Ratti (1977), ¹⁵Rudy et al., (2013), ¹⁶White and Orr (2011), ¹⁷Zhang and Campbell (2002).

** CBRAT (Lee and Reusser, 2014) species profiles were developed from independent references (available at <http://www.cbrat.org>).

2005; Hijmans, 2012). For environmental variable inputs, some models use data collected alongside species occurrence data (Schoeman and Richardson, 2002; Thrush et al., 2003), whereas others take advantage of existing data (Pereira and Itami, 1991; Lauver et al., 2002; Vincenzi et al., 2006). Spatially explicit models often use the environmental datasets to interpolate or estimate values over an area to predict the areal distribution of species or suitable habitat (Lenoir et al., 2011; Reiss et al., 2011). The data collection required for complex empirical models can be time-intensive and costly, which may limit their use by organizations or agencies that are limited by personnel and funds.

Habitat-based models can provide important insight into the distribution of organisms, particularly in marine and estuarine ecosystems, where organisms are often difficult to sample (Lenoir et al., 2011; Reiss et al., 2011). Estuaries offer a diverse range of habitats that are dynamic in space and time, and the organisms that inhabit these ecosystems are generally tolerant to a range of environmental conditions (Thrush et al., 2003). The suitability of habitat for estuarine organisms can then be considered scalar, ranging from unsuitable to most suitable, reflecting the likely distribution of species. Estuaries located in the U.S. Pacific Northwest (PNW) are inhabited by commercially and recreationally important shellfish, including the following bay-clam species: cockles (*Clinocardium nuttallii*), softshells (*Mya arenaria*), gapers (*Tresus capax*), butters (*Saxidomus gigantea*), and littlenecks (*Leukoma staminea*). These species are distributed among estuaries along much of the NE Pacific coastline, and their ranges often coincide in the U.S. PNW; however, wide larval dispersal and patchy settlement contribute to generally low estuary-wide population densities. The relative abundance and spatial distribution of bay-clam species varies within estuaries (Ainsworth et al., 2014). We posit that this is likely due to differences in habitat distribution among estuaries. Field distribution data are limited for bay-clam species in the PNW, likely because spatially extensive sampling requires expensive equipment (e.g., large boat, corer, dive gear, suction dredge), labor-intensive collection (e.g., raking, digging, coring, diving), and tedious sorting and identification of collected specimens.

Our goal was to develop a framework to identify suitable bivalve habitat in estuaries (FISBHE) – a spatially explicit habitat suitability index (HSI) modeling framework, using published natural-history

information (to establish bivalve species' physiological sensitivities to habitat variables) and existing habitat datasets, that could be used for multiple bivalve species and transferred across estuaries of the PNW. Bay-clam occurrence data were not necessary to build FISBHE, however independent bay-clam datasets were used for model validation. From a modeling standpoint, this could be considered a mechanistic model with empirical ties (Guisan and Zimmermann, 2000) that favors generality and reality over precision (Levins, 1966). Such a framework would be a time- and cost-efficient way to identify general areas of suitable habitat for bay-clam species.

2. Methods

2.1. Study area

Most estuaries in the PNW are small, drowned river basins (Emmett et al., 2000; Hickey and Banas, 2003) with extensive tideflats that support diverse infaunal communities (Ferraro and Cole, 2011, 2012), including populations of harvested bay-clams (Ainsworth et al., 2014). This study focused on two Oregon estuaries: Yaquina and Tillamook bays (Fig. A1.1 [see Appendix 1]). Yaquina Bay (15.8 km²) supports a predominately recreational bay-clam fishery, whereas Tillamook Bay (34 km²) supports subtidal and intertidal commercial bay-clam fisheries, in addition to one of the most productive recreational fisheries in the state (Ainsworth et al., 2014). Yaquina Bay has slightly more tidal variation (mean tidal range of 1.9 m) when compared to that of Tillamook Bay (mean tidal range 1.6 m) (Shirzad et al., 1988). We first developed FISBHE for use in Yaquina Bay, then applied the framework to Tillamook Bay to test for transferability across estuaries.

2.2. Model parameterization

FISBHE centers on established sensitivities of each species to several environmental variables that potentially contribute to their habitat 'preferences' or associations. We conducted an extensive review of literature (predominantly from NE Pacific estuaries, including the U.S. PNW) on the natural-history, distribution, and physiology of the five

bay-clam species (*Clinocardium nuttallii*, *Mya arenaria*, *Tresus capax*, *Saxidomus gigantea*, and *Leukoma staminea*) to obtain information on species-specific sensitivities (i.e., survival, mortality, presence-absence) to several environmental variables (see Appendix 2). Although many environmental variables can affect the spatial distribution of bivalves within an estuary, particularly during short-term events (e.g., green macroalgal bloom effects on cockles; Lewis and DeWitt, 2017), four variables that were consistently cited as influential on survival and distribution were: bathymetric depth, sediment silt-clay %, low salinity, and the presence of burrowing shrimp. Henceforth, these are the “habitat variables” used in FISBHE. We established species-specific sensitivity ranges for each habitat variable as the average range (derived from the collective literature) within which the bay-clam species could consistently survive (Table 1, Tables A2.1 – A2.4 [see Appendix 2]); for a more detailed description, see Text A2.1 in Appendix 2. The literature review included data from throughout the species’ geographic ranges, and thus we made no distinction for possible physiologically distinct subpopulations when establishing sensitivity limits. These sensitivity ranges essentially served as simple suitability index (or response) curves. Interactions among habitat variables, or with other environmental variables were not evaluated nor considered within FISBHE.

Mature bay-clam species have differential habitat ‘preferences’ or associations based on physiological differences, such as body size, siphon length, metabolic rate, and ventilation rate (Bernard, 1983). The range of bathymetric depths at which bivalve species reside can be indicative of the ability to tolerate temperature gradients (Bernard, 1983). Lower bathymetric zones (i.e., subtidal and low intertidal) experience more consistent temperature regimes than higher bathymetric zones (i.e., mid-high intertidal), so these strata are important drivers of species distributions for bivalves with different temperature requirements (Bernard, 1983). Physical sediment characteristics are important to bay-clams because these species are infaunal, meaning most of their lifetime is spent within the sediment. For instance, *T. capax* prefers stable, consolidated sediment (characteristic of fine mud or clay) because it is a sedentary deeper-lived species (Lauzier et al., 1998; Dethier, 2006), whereas *C. nuttallii* is a more mobile species (Lewis and DeWitt, 2017) that often inhabits unconsolidated surface sediments (sandy or mixed sediments) (Gaumer and McCrae, 1990; Dethier, 2006) better suited for active movement. Extended periods of low salinities are limiting to these bivalve species (Gaumer and McCrae, 1990; Lauzier et al., 1998; Endreny and Sylvia, 2006; Dethier, 2006; Rudy et al., 2013). For modeling purposes, low salinity was defined as the mean wet-season salinity (November–April; period of maximum rainfall for the coastal PNW; Emmett et al., 2000). We used the mean salinity value (regardless of location in the water column and tide stage) at established stations sampled during the wet season due to the high spatiotemporal variation inherent with salinity in a mixed semi-diurnal tidal cycle. Burrowing shrimp are ecosystem engineering species that significantly affect soft-sediment habitat through bioturbation and bioirrigation, leading to sediment destabilization and increased turbidity, which can smother burrowing bivalves and interfere with suspension-feeding bivalves (Murphy, 1985; Feldman et al., 2000).

Existing spatially explicit data for the four habitat variables in Yaquina and Tillamook bays were obtained from state (Oregon Dept. of Fish and Wildlife, and Oregon Dept. of Environmental Quality) and federal (U.S. Environmental Protection Agency, U.S. Dept. of Agriculture, U.S. Army Corps of Engineers, and National Ocean Service) natural resource agencies (Table A1.1). For each habitat variable, data were transformed into the same units of measurement and spatial reference (NAD 1983 UTM Zone 10N), then combined into a geospatial dataset. The temporal extent of the abiotic datasets reached up to 62 years (1953–2015), whereas the burrowing shrimp datasets never surpassed 16 years (1996–2012). We included the full time-series of data for each variable to generally yield more robust and representative datasets, assuming that available datasets are often disparate and collected infrequently. Temporal variation in physical habitat variables

(bathymetric elevation, mean salinity, and sediment characteristics) would be expected to be naturally minimal; however, changes in the biological variable (burrowing shrimp presence) could occur on a shorter time scale, so the shorter temporal extent for those data is appropriate. After importing each dataset into ArcGIS (V 10.2.2), the data were audited to remove spatially inaccurate points that fell outside estuary boundaries or points that had inaccurate values outside variable ranges observed in PNW estuaries.

2.3. Model creation

Using the discrete point datasets generated from existing habitat data, we created an estuary-wide estimation surface (i.e., raster) for each habitat variable. These estimation surfaces were interpolated using ordinary geostatistical kriging (ArcGIS Geostatistical Analyst), which also produced an uncertainty (\pm SE) prediction surface for each variable estimation surface. This method allowed us to adjust the finer aspects of the interpolation (e.g., variogram and search neighborhood parameters) that were important considerations for the fit of these distinct habitat variables created from varied quantities and distributions of data points. Variograms (or stable semi-variograms, as used in our framework) are geostatistical spatial autocovariance structures that differentiate data variation with distance (Burrough, 2001); in other words, a semi-variogram guides geostatistical kriging to estimate nearby features more predictably and with less variability than features far apart. As another means of controlling the disparity between habitat datasets, parameters of the search neighborhoods (i.e., number of points and semiaxes) were scaled to reflect the total number of points in a dataset, the distance covered by a dataset, as well as the physical properties (e.g., spatiotemporal variation) of the variable being interpolated. For each habitat variable, we determined standardized proportional ranges of total stations in a dataset to use as the number of points in a search neighborhood and developed a standardized equation to determine the distance of the neighborhood semiaxes.

The product of each habitat-variable interpolation was a raster, or estimation surface, of continuous values approximately bounded by the range of values within the observed dataset, providing a surface of spatially and temporally averaged conditions. Habitat-variable values at each cell on the estimation surface were then assigned a binary suitability score for each species (Table 1); values within each species’ sensitivity range received a suitability score of ‘1’, while values outside of the range received a suitability score of ‘0’. Each cell of the habitat-variable rasters contained seven value fields: the interpolated estimate of the variable, estimated SE, and the binary suitability score (1 or 0) for each of five bay-clam species. Suitability scores were summed across habitat variables for each bay-clam species in each raster cell to produce an overall habitat-suitability-index (HSI) raster with classes ranging between 0 (not suitable; i.e., no habitat-variable values within the suitable sensitivity range) and 4 (most suitable). This method differs from many models that calculate HSI on a continuous scale between 0.0 and 1.0, but was necessary due to the use of binary suitability scores and equal weights among habitat variables. Additionally, the scaling of HSI into categorical classes minimized uncertainty and complexity. As a qualitative measure of confidence in the HSI predictions, we developed a confidence index (0.0–1.0) by subtracting the cumulative habitat variability from 1.0 (or 100% confidence); cumulative habitat variability was calculated by normalizing each habitat SE raster as a proportion of the observed range, multiplying each by 0.25, and summing the four resulting rasters. This confidence index is descriptive in nature and was not used to make formal statistical inferences.

2.4. Model validation

Model validation increases user confidence and the resulting utility of a model (Power, 1993). To produce a validation dataset, we followed the same procedure used to create the geospatial habitat datasets; we

obtained field-collected bivalve data from multiple natural resource agencies (dating from 1996 to 2012 for Yaquina Bay, and 1999 to 2011 for Tillamook Bay), transformed bay-clam species data into presence/absence, projected those data into the same geospatial reference (NAD 1983 UTM Zone 10N) within ArcGIS (V 10.2.2), and audited the layers to remove spatially inaccurate points that fell outside estuary boundaries. The resulting dataset for Yaquina Bay contained 780 unique stations with presence-absence data for each bay-clam species, whereas the dataset for Tillamook Bay contained 635 unique stations (Fig. A1.2 [see Appendix 1]). These validation datasets were normalized using bivalve presence/absence to minimize potential effects of heteroscedasticity among multiple studies and the patchy settlement of bivalves. Logistic regression and Kendall's correlation coefficient (tau-b) analyses were then used to validate the HSI models.

We created a binary logistic regression model for each species using the 'glm' function in the statistical software R (R Development Core Team, 2016); species presence/absence was the response variable and the HSI class was the explanatory variable:

$$p(x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}$$

where $p(x)$ was the species presence probability, x was the HSI class (treated as a continuous variable), and β_0 and β_1 were slope parameters. The estimated logistic regression model can be used to predict the probability of a species' presence for a given HSI class. For example, setting $x = 2$ in the estimate of the regression model gives the predicted presence probability when the HSI class is 2. Logistic regression models were used to determine whether there was a correspondence between the HSI classes (from 0 to 4) and the predicted presence probabilities for each bivalve species. Confidence intervals were calculated (Hosmer and Lemeshow, 2000, p. 20) to quantify the uncertainty in the estimation of the presence probabilities for each HSI class, and thus provided a way to determine whether there was significant separation between the presence probabilities at different HSI classes.

Kendall's correlation coefficient (tau-b) was used as a nonparametric alternative validation method that made no statistical modeling assumptions. We implemented this analysis by using the 'cor.test' function in R (R Development Core Team, 2016), which measured the ordinal association between two variables, and provided strength and assessment measures. This analysis measured the model's ability to predict the presence/absence of each bay clam species. In contrast to logistic regression, which treated HSI as a continuous variable, Kendall's tau-b accounted for discrete data.

3. Results

3.1. Habitat variables

Habitat interpolations yielded estimation surfaces that reasonably represented the distribution of habitat-variable values relative to data available for Yaquina and Tillamook bays (Fig. 1, Fig. A1.3 [see Appendix 1]). We also calculated the uncertainty (\pm SE) associated with each estimation surface (Figs. A1.4, A1.5 [see Appendix 1]) to assess the accuracy of the interpolations. Bathymetric interpolation (m; Mean Lower Low Water datum [MLLW]) of the estuaries accurately showed the main channels, tidal flats, and tidal channels with little uncertainty due to the quantity of input points used for the interpolation ($> 49,000$). Sediments contained a higher predicted percentage of silt-clay (fines) near input creeks and streams; in the main stem of the estuaries, silt-clay percentage generally decreased (i.e., sediment became coarser) downstream toward the ocean, culminating in predominately sandy sediment near the estuaries' mouths, as previously reported (Peterson et al., 1984). Most of data used to interpolate sediment silt-clay were derived from the intertidal zone, so interpolated estimates for tideflats showed the least variability. Mean wet-season

salinity decreased from the estuaries' mouths (> 27 PSU in Yaquina, and > 23 PSU in Tillamook) to < 5 PSU at the upper reaches of the estuaries and showed the lowest variability near the main channels. Burrowing shrimp occurrence was greatest within the central portions of both estuaries, particularly on intertidal flats.

3.2. Yaquina Bay HSI models

HSI model output for each species showed distinct patterns in the distribution of suitable habitats (Fig. 2). The confidence index raster, which descriptively assessed cumulative variability in habitat interpolations, showed high confidence (0.81–0.97, 0.93 average) in the FISBHE predictions throughout Yaquina Bay (Fig. A1.6 [see Appendix 1]).

Predicted presence probabilities from logistic regression analyses showed strong correspondence with the HSI classes for Yaquina Bay (Fig. 3). Thus, as the HSI increased from 1 to 4 (no areas were predicted to have HSI 0 habitat), so did the presence probability for each bay-clam species, as predicted by logistic regression. The 90% confidence intervals for the presence probabilities also showed reasonable separation between the different HSI classes (1–4). For *C. nuttallii*, *M. arenaria*, and *T. capax* there was some overlap between the 90% confidence intervals at consecutive HSI classes (e.g., between HSI 1 and 2); however, the confidence intervals did not overlap for larger differences (e.g., between HSI 1 and 3, or HSI 2 and 4). For *L. staminea* and *S. gigantea* species, only a small proportion ($\leq 3.6\%$) of sites had observed presences; hence the overlap between the confidence intervals for these species was in part due to the large amount of uncertainty in estimating the presence probabilities at each HSI class. Finally, the estimates of the slope parameters (β_1 , expressed in terms of the logit) for the Yaquina Bay logistic regression models were highly significant ($p < 0.01$) for the three most-common species. Thus, the correspondences between the logistic probabilities and HSI classes were genuine, and not likely due to random variation.

Kendall's correlation coefficient (tau-b [τ_b]) results were in agreement with the logistic regression results; the presence of *C. nuttallii*, *M. arenaria*, and *T. capax* were significantly ($p < 0.01$) correlated with the predicted HSI classes for Yaquina Bay. Statistically nonsignificant results from the validation analyses of *S. gigantea* and *L. staminea* do not necessarily mean that the modeling framework was unsuccessful for these species. Due to the naturally low population densities of these species in Oregon, model validation for the latter two species was potentially limited by an insufficient quantity or unequal distribution of validation points.

3.3. Tillamook Bay HSI models

Similar to Yaquina Bay, HSI model output for Tillamook Bay showed distinct patterns in the distribution of suitable habitats for each species (Fig. 4), and the descriptive confidence index raster displayed high confidence (0.75–0.97, 0.94 average) in FISBHE's predictions (Fig. A1.6 [see Appendix 1]). No areas were predicted to have an HSI of 4 for these species in Tillamook Bay, likely due to the wet-season salinity layer, which may have been too homogenous throughout the bay to be a good discriminator.

Logistic regression validation results for Tillamook Bay were similar to those of Yaquina Bay, indicating that the framework was successfully transferred between estuaries. In general, the predicted presence probabilities from the Tillamook logistic regression models corresponded with the HSI classes; the only exception was *M. arenaria*, which showed no significant association between the predicted probabilities and HSI (Fig. 5). That species was the only one in which the 90% confidence intervals overlapped across several HSI classes. Logistic presence probabilities and confidence intervals could only be extrapolated for *T. capax*, *L. staminea*, and *S. gigantea* at HSI 4 since habitats in Tillamook Bay were not predicted to have an HSI of 4 for those

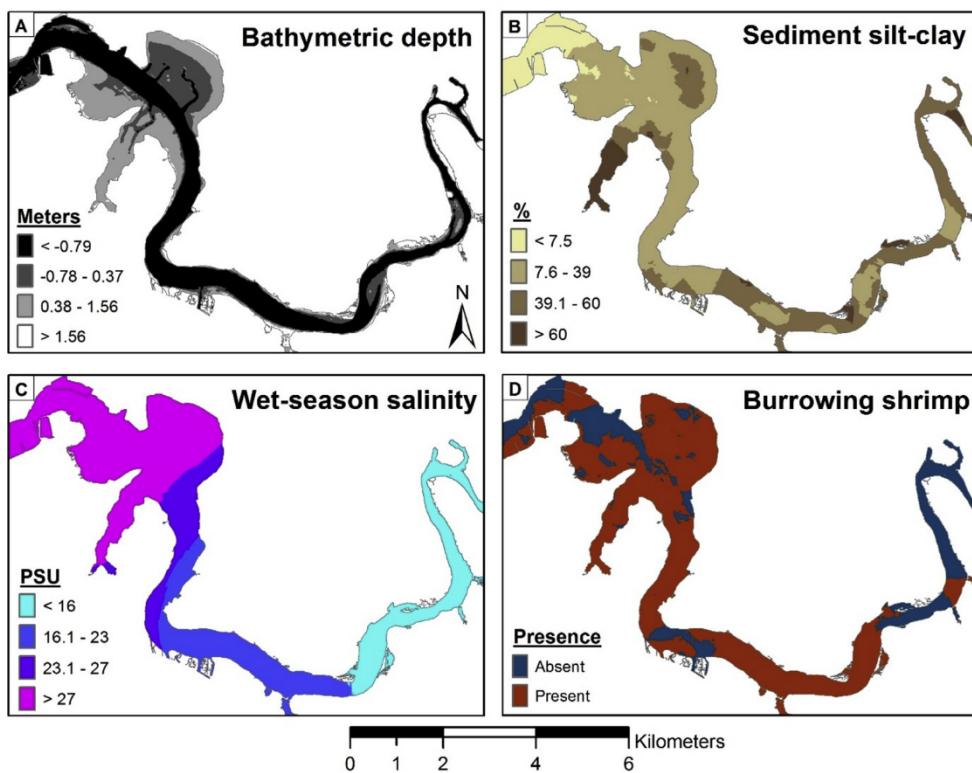


Fig. 1. Estuary-wide estimations (for Yaquina Bay, OR) of the input habitat variables used in FISBHE: [A] bathymetric depth (m; MLLW); [B] sediment silt-clay (%); [C] wet-season salinity (PSU); [D] burrowing shrimp (presence). Estimation surfaces were interpolated using geostatistical kriging.

species. Finally, the estimated slope parameters for the logistic regression models were highly significant ($p < 0.001$) for all species other than *M. arenaria*.

Results from Kendall's correlation coefficient (τ_b) analyses reflected the findings of logistic regression; the presence of *C. nuttallii*, *T. capax*, *S. gigantea*, and *L. staminea* were significantly ($p < 0.001$) correlated with the predicted HSI classes in Tillamook Bay. Thus, the HSI was generally a statistically significant predictor of bay-clam presence/absence in both bays.

4. Discussion

Using existing, disparate environmental datasets and natural-history information, we developed, parameterized, and validated an approach for modeling habitat suitability for harvested bivalves in the PNW that was successfully applied to predict areas of suitable habitat for five bay-clam species within two estuaries. This utilitarian approach offers a time- and cost-efficient means to identify where ecologically and economically important clam stocks most-likely occur, forecast how environmental changes can alter stock distributions, and inform estuarine land-use decisions on issues such as environmental management, fisheries harvest locations, shellfish preserves, habitat restoration, and aquaculture designations. Implementation costs are minimal and relate to GIS software access and the time of at least one user with technical skills in estuarine ecology, GIS, and statistics. Although the use of this modeling approach has tradeoffs (like any model), we believe the advantages far outweigh the detriments when applied under the right circumstances.

We developed FISBHE to be transferable across bay-clam species and estuaries in the PNW, so it is predicated on generality, which is a driver of the associated advantages and detriments of the framework. The ideal input dataset for spatial interpolation contains a high density of sample points equally dispersed throughout the entire bay, but such is rarely the case when working with existing datasets. Small-scale

variation in habitat variables may not be accurately captured by our models because each input layer is interpolated from a different set of data points, thus the accuracy of variable estimates at a given point will vary. Temporal resolution also varies within each dataset, which limits us to examining patterns generalized across years, and requires us to assume that the habitat values were static across the time range of the datasets. However, these general estimations allowed habitat suitability to be modeled throughout entire estuaries at low labor costs.

Traditional HSI models (built from concurrently collected biophysical and bivalve data) are more laborious to parameterize than FISBHE, but are likely to fit bivalve distribution data better and produce more spatially-specific estimates for the estuary in which the model was developed. However, such models differ from our fundamental approach in that they require bivalve data to characterize bivalve-habitat relationships (e.g., via variable weighting). Variable weighting is often used within traditional HSI models to improve the fit of a model, but the resulting model will typically be specific to individual species within individual estuaries. Each input variable is weighted equally in FISBHE and all habitat variables are included, rather than a subset. Since FISBHE was parameterized with bivalve-habitat relationships estimated from collective literature, weighting habitat variables would have increased the level of uncertainty in our results because variables would need to be ranked in order of importance and to the degree of importance between the rankings – information that was not possible to glean from regional literature. Similarly, interactions between habitat variables were not considered due to the additional increase in uncertainty that would have resulted from estimating the magnitude and direction of complex interactions between multiple variables in the absence of data.

Although the spatial and temporal generalizability of FISBHE allows for application to other PNW estuaries, or modification for species in other locales, the framework could possibly mischaracterize habitat conditions at locations where data are sparse. Similarly, bivalve datasets lacking in size or distribution could potentially lead to validation

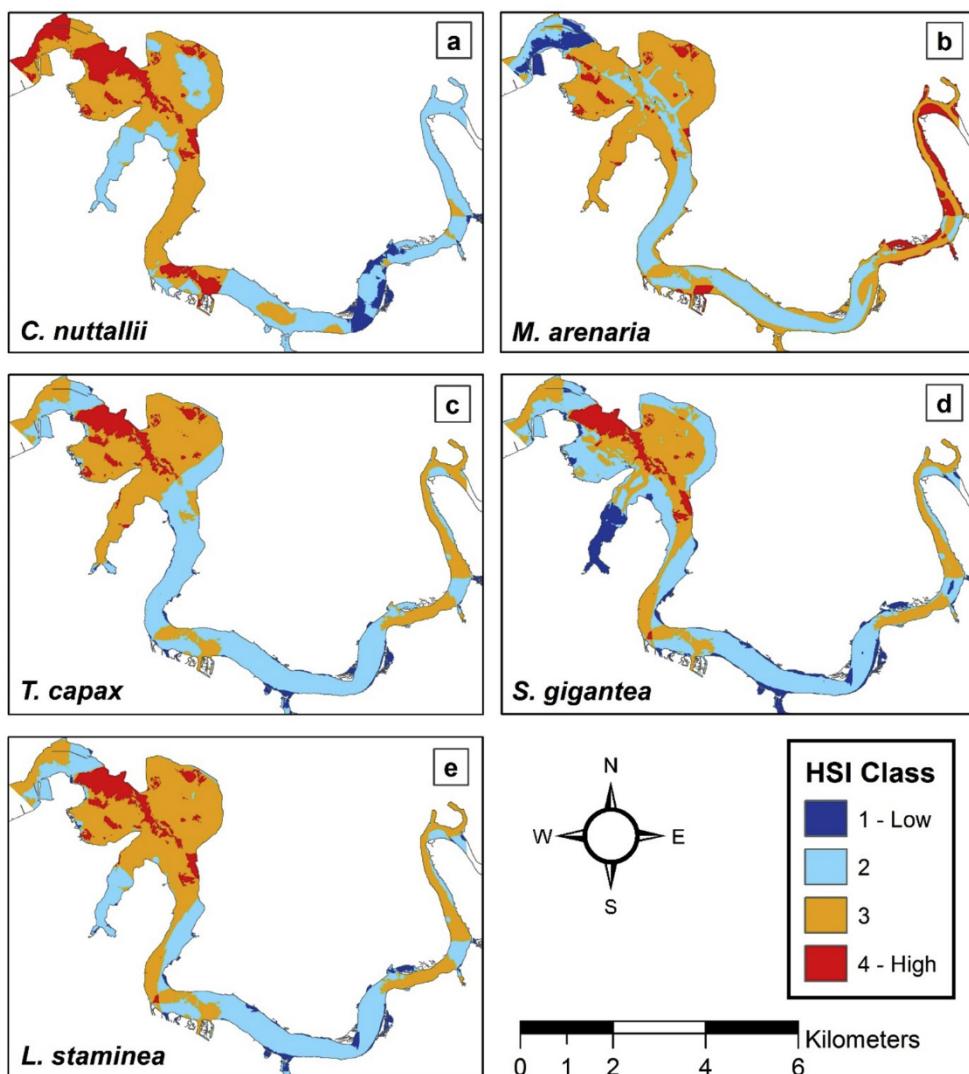


Fig. 2. Results from FISBHE in Yaquina Bay, OR; estuarine-wide HSI predictions for each species considered (a – *Clinocardium nuttallii*, b – *Mya arenaria*, c – *Tresus capax*, d – *Saxidomus gigantea*, e – *Leukoma staminea*). Predicted habitat suitability increases from HSI 1–4 (blue to red).

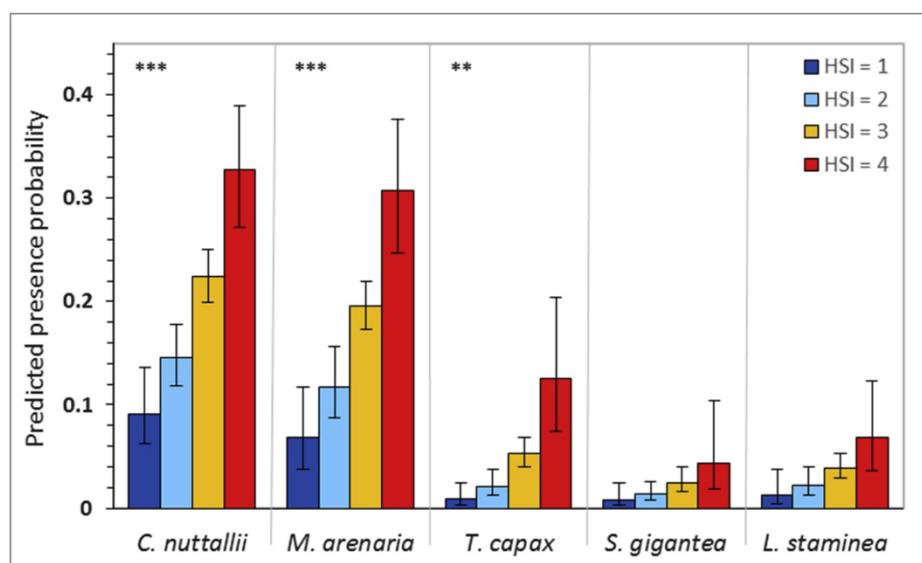


Fig. 3. Logistic regression results (predicted presence probabilities with 90% confidence intervals) used to validate FISBHE for bay-clam species in Yaquina Bay, OR. Significance levels (** $p \leq 0.001$, ** $p \leq 0.01$) indicate a correspondence between HSI classes (from 1 to 4) and predicted presence probabilities for a given species.

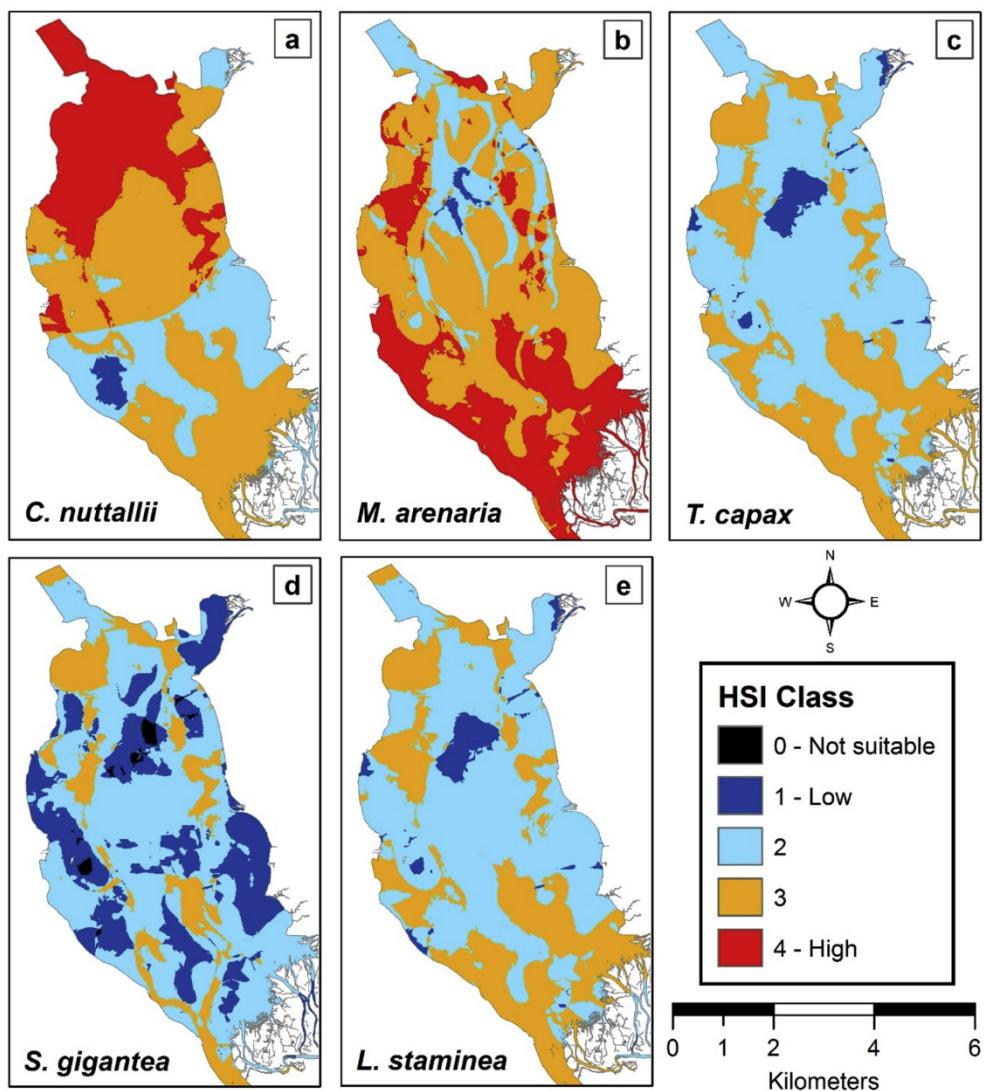


Fig. 4. Results from FISBHE in Tillamook Bay, OR; estuary-wide HSI predictions for each species considered (a – *Clinocardium nuttallii*, b – *Mya arenaria*, c – *Tresus capax*, d – *Saxidomus gigantea*, e – *Leukoma staminea*). Predicted habitat suitability increases from HSI 0–4 (black to red).

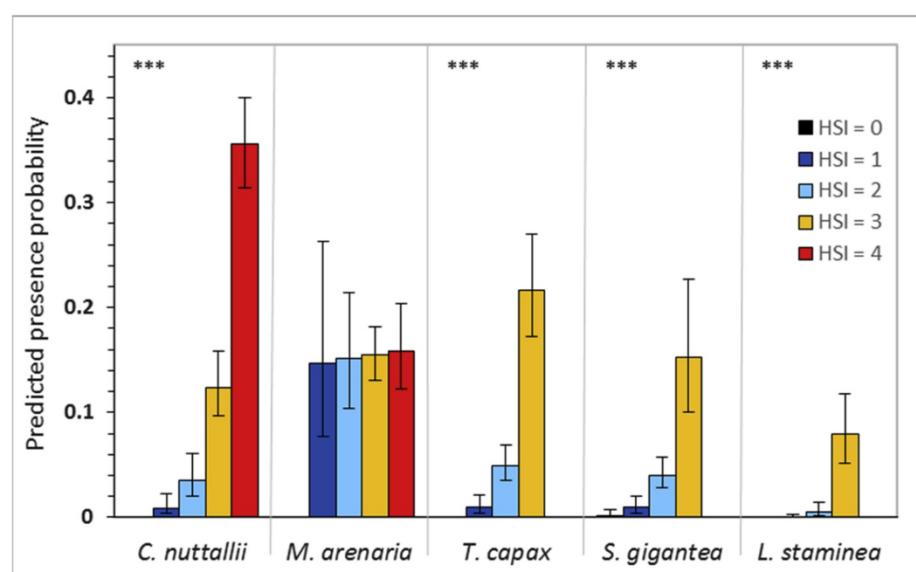


Fig. 5. Logistic regression results (predicted presence probabilities with 90% confidence intervals) used to validate FISBHE for bay-clam species in Tillamook Bay, OR. Significance level (*** $p \leq 0.001$) indicates a correspondence between HSI classes (from 0 to 4) and predicted presence probabilities for a given species. HSI classes that are not represented for a species were not predicted to occur for that particular species in Tillamook Bay.

results that misrepresent the extent of a population, as we believe may be a contributing factor to the *M. arenaria* results in Tillamook Bay. Although validation points are spread throughout Tillamook Bay, the lowest density of points occurs in the upper half of the estuary where *M. arenaria* are uniquely suited to survive. This disproportionate distribution or stratification of validation points likely leads to a lower confidence in the validation results from the upper portion of the estuary, a potential pitfall of utilizing existing datasets. However, supplemental data collection could produce a validation dataset or fill data gaps with much less effort than a comprehensive bay-wide study to quantify shellfish species distributions or to develop bivalve-habitat relationships.

Transferability was an important consideration in the development of FISBHE because we anticipated the utility of the approach would increase as the scope of its predictive ability broadened. We successfully applied this modeling framework to five bay-clam species within two PNW estuaries. However, characterization of habitat sensitivities for other bivalve species can potentially generate HSI maps for a wide range of estuarine species, given that the following assumptions are met: sensitivities are derived from literature review, target species is an estuarine filter feeder, and no exclusionary interspecific competition exists between modeled species. Misapplication of FISBHE or misinterpretation of the HSI results would likely result in inaccurate or unsupported results. For instance, applying the framework for a particular species to an estuary outside of its native range does not imply that species is present within that estuary. If there is reason to believe that viable introduction of non-native species is possible, however, then FISBHE could provide users exploratory insight into locations that would most likely be occupied by that species.

FISBHE is limited to a spatial representation of the post-larval fundamental niche for each bay-clam species because the model inputs were developed on general life-history traits. Traditional distribution models (e.g., generalized linear models [GLMs] or generalized additive models [GAMs]) are based on empirical relationships between organism occurrence data and environmental predictor variables (Fielding and Bell, 1997; Guisan and Zimmermann, 2000; Austin, 2002; Boyce et al., 2002; Elith and Graham, 2009), and thus provide estimates of the realized niche (Guisan et al., 2002; Ysebaert et al., 2002; Raybaud et al., 2015). We used independently determined sensitivity limits mainly from regional literature to set suitability thresholds within each habitat variable, which allowed us to use independent, existing bivalve datasets for validation. These limits served as simple (binary) versions of the suitability functions or response curves derived by GLMs or GAMs. Estimation of bivalve-habitat relationships solely from literature lacked the accuracy and precision for curvilinear relationships (i.e., continuous suitability functions) required for many traditional HSI models. Few studies have taken the approach of using existing independent suitability functions (Lauver et al., 2002; Vincenzi et al., 2006), and fewer have used existing literature to develop these species-habitat relationships (Blandford et al., 2013).

In many estuaries, environmental and habitat data are collected by various organizations for a variety of reasons, but complementary bivalve data do not often exist. Our framework has the capability to produce HSI estimates without any estuary-specific information about bivalve occurrence. Validation, of course, would not be possible without sufficient estuary-specific bivalve data. In Appendix 1, we provided an additional example in which FISBHE was applied to such an estuary (i.e., Willapa Bay, WA, 348 km²) to generate habitat (Fig. A1.7) and HSI maps (Fig. A1.8) that we were unable to validate due to insufficient bivalve distribution data (Fig. A1.9). The density of available validation points for Willapa Bay was approximately 0.51 points km⁻²; whereas Tillamook and Yaquina bays had approximate densities of 18.7 and 49.4 points km⁻², respectively. Models that utilize organism occurrence data in model-fitting procedures can have difficulty in accurately discerning between areas of use and non-use; although the presence of an organism is typically indicative of ‘use’, the absence of

an organism does not necessarily preclude ‘use’ (Fielding and Bell, 1997; Austin, 2002; Boyce et al., 2002). Oftentimes, these studies also split datasets into calibration and evaluation datasets – called cross-validation or data-splitting (Power, 1993) – to construct and test proposed models (Fielding and Bell, 1997; Guisan and Zimmermann, 2000; Ysebaert et al., 2002; Hirzel et al., 2006; Lauria et al., 2015). Potential issues with data-splitting include spatial sorting bias (Hijmans, 2012), smaller calibration datasets (Fielding and Bell, 1997), and circularity (Boyce et al., 2002). Our approach avoids those problems by not requiring bivalve data for model calibration, and by using independent bivalve datasets to validate the HSI models.

Probabilistic suitability or distribution models built from extensive sets of field-sampled bivalve data and associated habitat data are likely to provide the most comprehensive and accurate representation of bay-clam populations within sampled areas of a given bay. This methodology is time-intensive and costly, however, and its performance is not indicative of transferability due to the limitations of predictor variable coefficients (Power, 1993; Guisan et al., 2002; Lauria et al., 2015). Misapplication of overly flexible, or highly sensitive, models can lead to overfitting, which limits their generalizability and transferability to other locations (Vaughan and Ormerod, 2005; Randin et al., 2006). Although inappropriate to identify causal mechanisms of bivalve presence or estimate population abundance, our ordinal HSI framework provides the benefit of generalizability by using suitability thresholds developed from the collection of literature for a geographic area.

Species may be sensitive to a wide range of environmental factors, but their presence or absence may be reliably estimated with a subset of highly influential variables. Fortunately for species with a long history of recreational or commercial value, their sensitivities to key environmental variables are often well-characterized in the ecological and fisheries literature. If variables that limit species survival or distribution are also frequently measured in the environment, then the approach we portray here can be used to produce HSI estimates for these species. Potential users should consider the realistic tradeoff between accuracy and cost within the context of intended model applications before implementing this framework. The primary advantages of FISBHE include its basis in the ecology of the organism, its low cost, and its simplicity to explain. Perhaps the advantage that offers the most valuable future direction for this research is the potential to estimate vulnerability of this valued ecosystem good to changes in habitat quality and distribution. As suggested by Boyce et al. (2002), we built FISBHE on ecological principles and validated the output with field data, which allowed the framework to be more robust from an exploratory, rather than explanatory, standpoint.

Acknowledgements

The authors would like to acknowledge Steve Ferraro, Faith Cole, Melanie Frazier, and Cheryl Brown (EPA), as well as Brett Dumbauld (U.S. Dept. of Agriculture) and Tony D'Andrea (Oregon Dept. of Fish and Wildlife) for providing access to environmental and infaunal data. Pat Clinton (EPA) provided valuable technical assistance with GIS software. We thank Jim Markwiese (EPA), Susan Yee (EPA), Stephen Pacella (EPA), Steve Rumrill (Oregon Dept. of Fish and Wildlife), and four anonymous reviewers for providing invaluable feedback on initial manuscript drafts. This research was supported in part by an appointment to the Research Participation Program for the U.S. Environmental Protection Agency, Office of Research and Development, administered by the Oak Ridge Institute for Science and Education through an interagency agreement (#92429801) between the U.S. Department of Energy and EPA. The remainder of this research was funded by the EPA. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecss.2019.02.009>.

Appendix 1

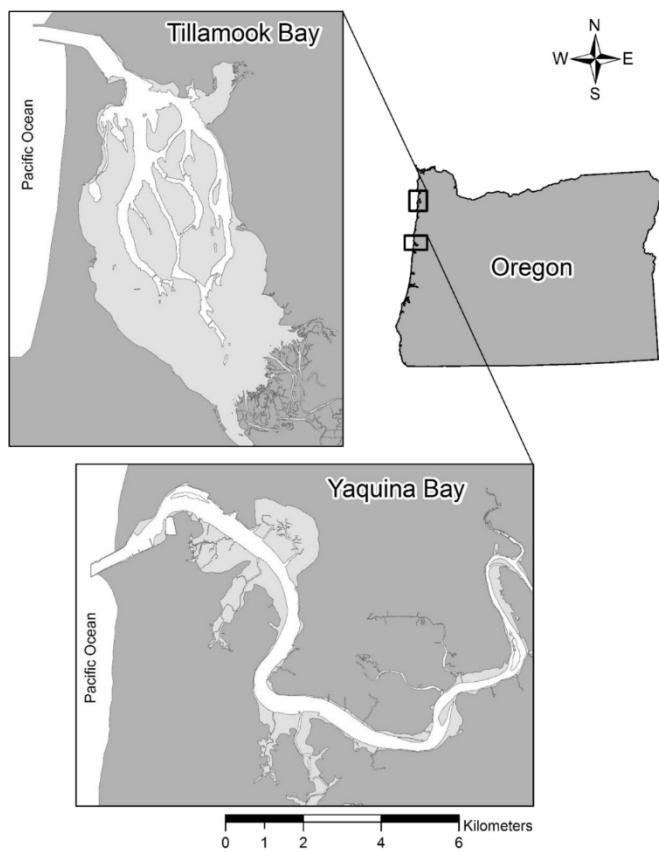


Fig. A1.1. Study area maps of Tillamook and Yaquina bays, OR.

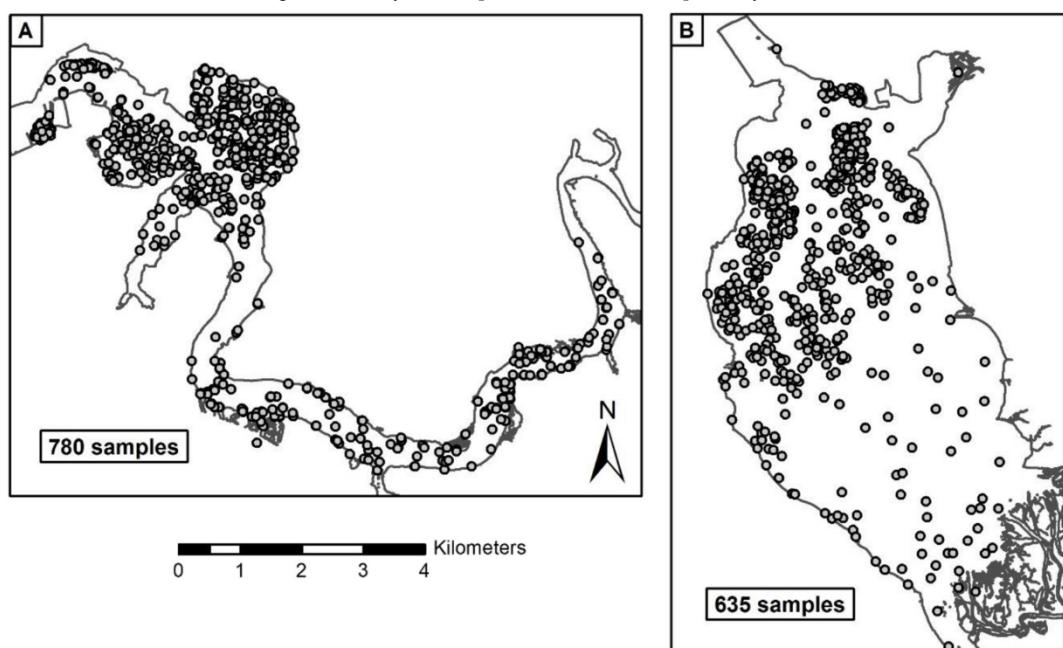


Fig. A1.2. Bay-clam sample locations in [A] Yaquina Bay, OR and [B] Tillamook Bay, OR.

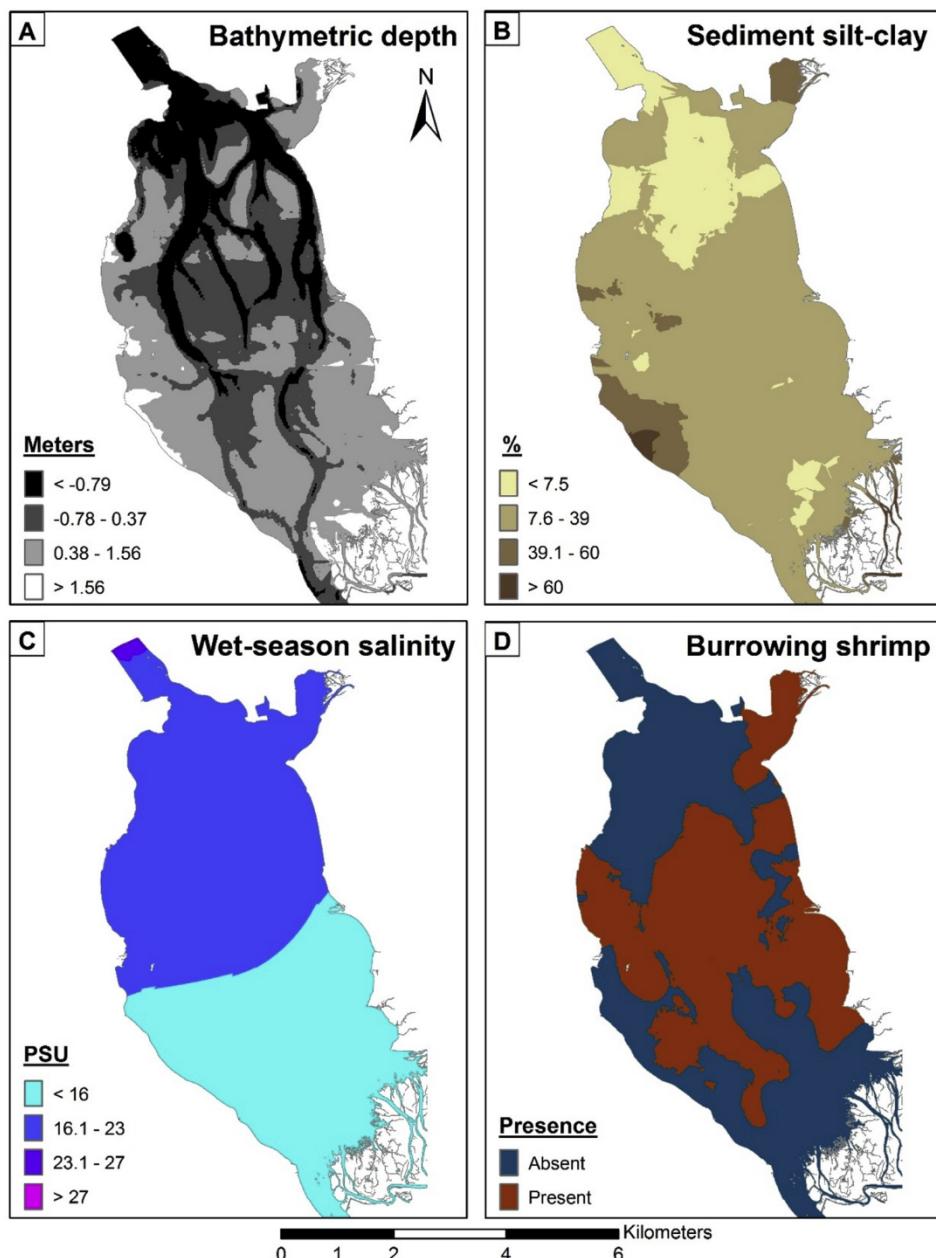


Fig. A1.3. Estuary-wide estimations (for Tillamook Bay, OR) of the input habitat variables used in FISBHE: [A] bathymetric depth (m; MLLW); [B] sediment silt-clay (%); [C] wet-season salinity (PSU); [D] burrowing shrimp (presence). Estimation surfaces were interpolated using geostatistical kriging.

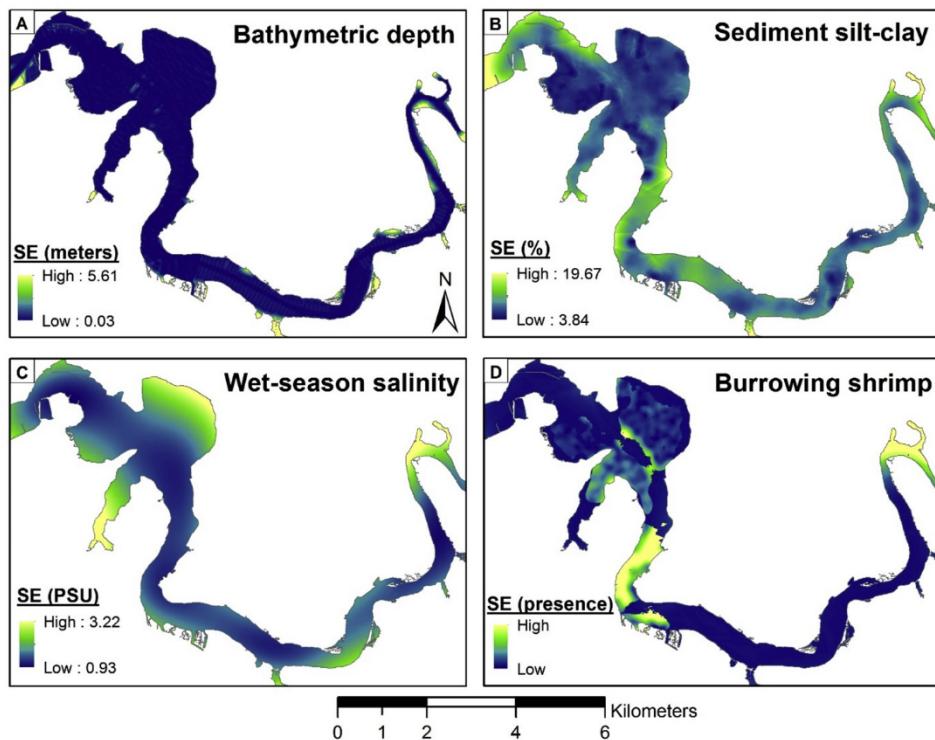


Fig. A1.4. Uncertainty (\pm SE) associated with habitat-variable interpolations in Yaquina Bay, OR: [A] bathymetric depth (m; MLLW); [B] sediment silt-clay (%); [C] wet-season salinity (PSU); [D] burrowing shrimp (presence probability).

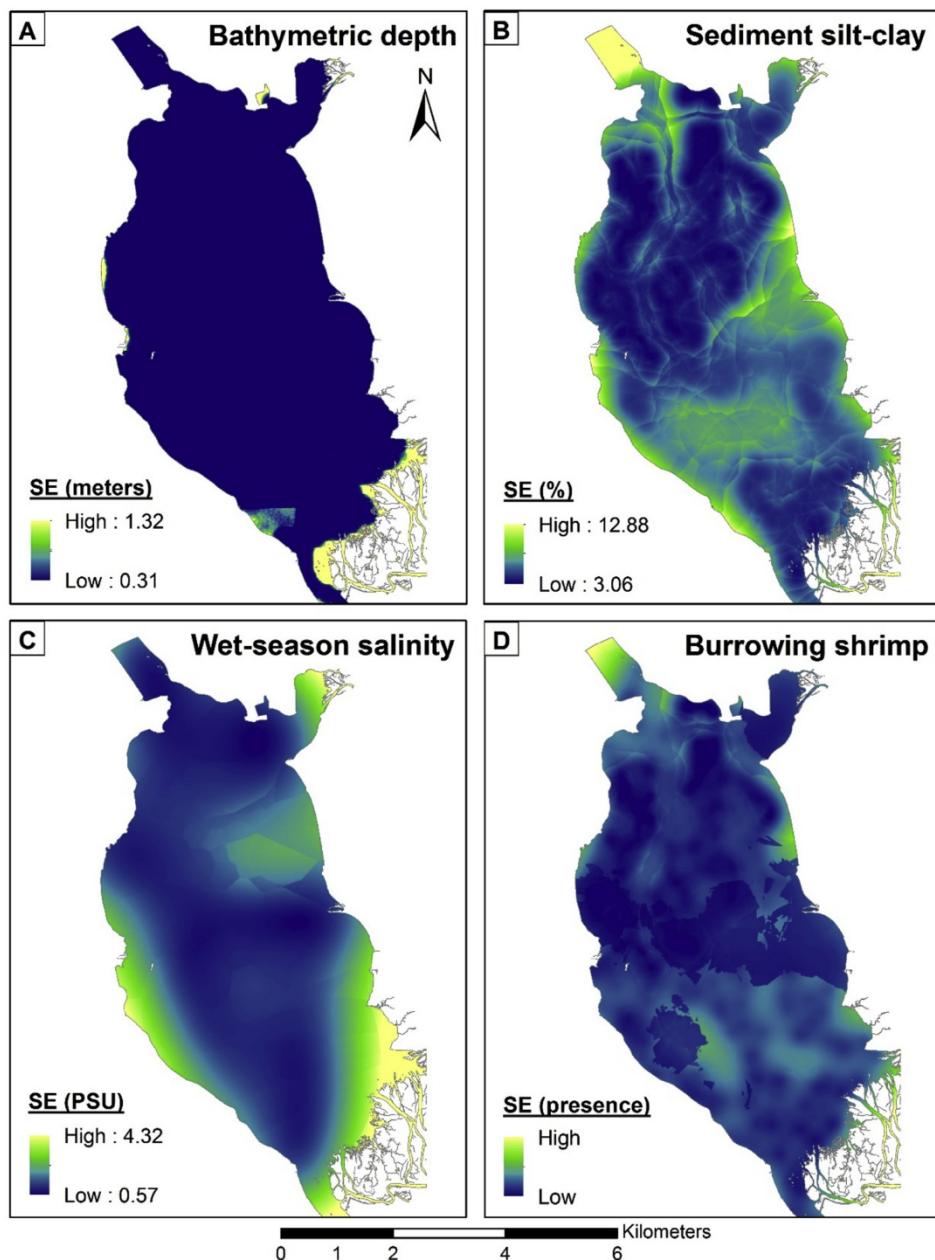


Fig. A1.5. Uncertainty (\pm SE) associated with habitat-variable interpolations in Tillamook Bay, OR: [A] bathymetric depth (m; MLLW); [B] sediment silt-clay (%); [C] wet-season salinity (PSU); [D] burrowing shrimp (presence probability).

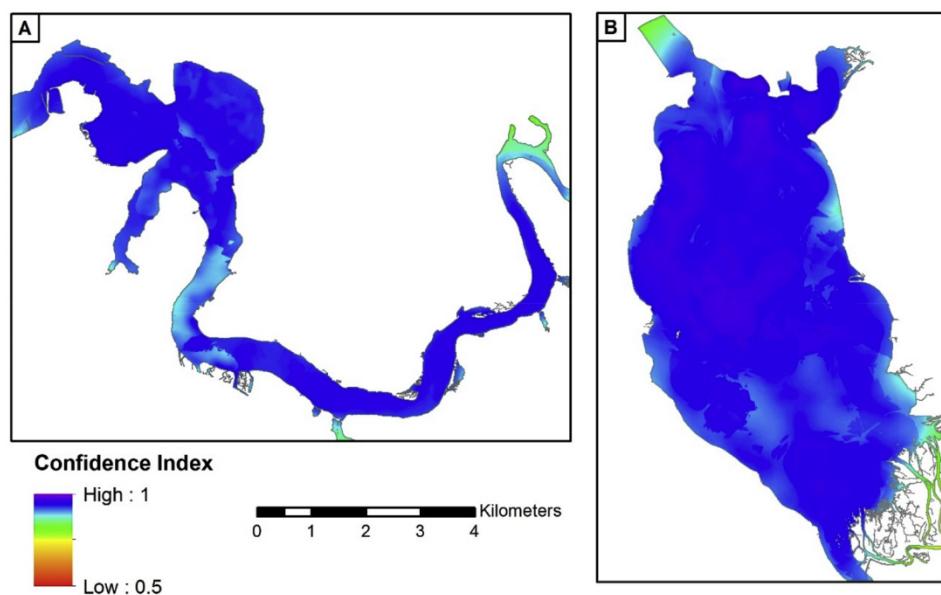


Fig. A1.6. Confidence index values illustrating high confidence in FISBHE results throughout (A) Yaquina Bay, OR and (B) Tillamook Bay, OR. Interpretation of these values is solely descriptive and cannot be extended to formal statistical inference.

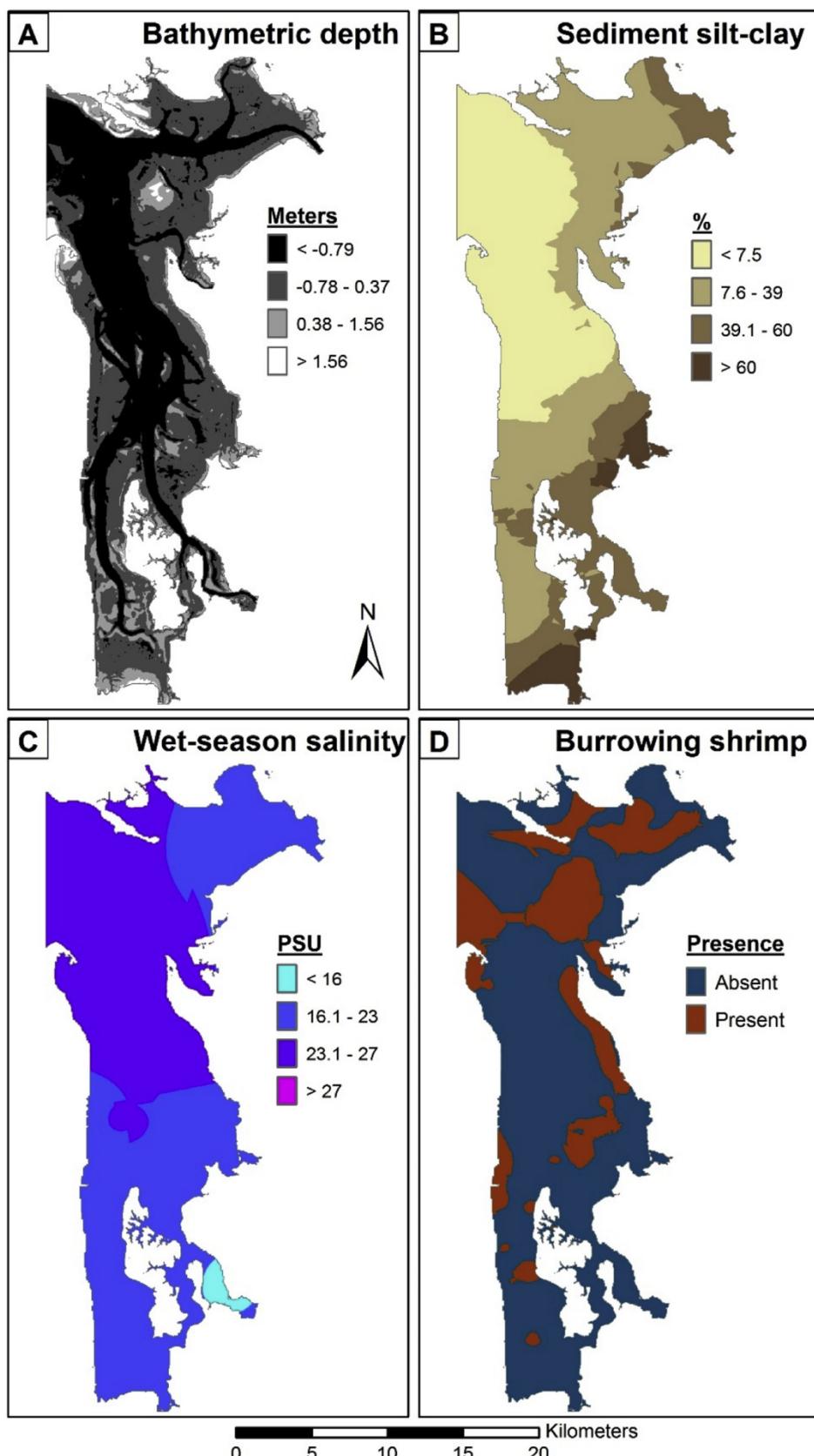


Fig. A1.7. Estuary-wide estimations (for Willapa Bay, WA) of the input habitat variables used in FISBHE: [A] bathymetric depth (m; MLLW); [B] sediment silt-clay (%); [C] wet-season salinity (PSU); [D] burrowing shrimp (presence). Estimation surfaces were interpolated using geostatistical kriging.

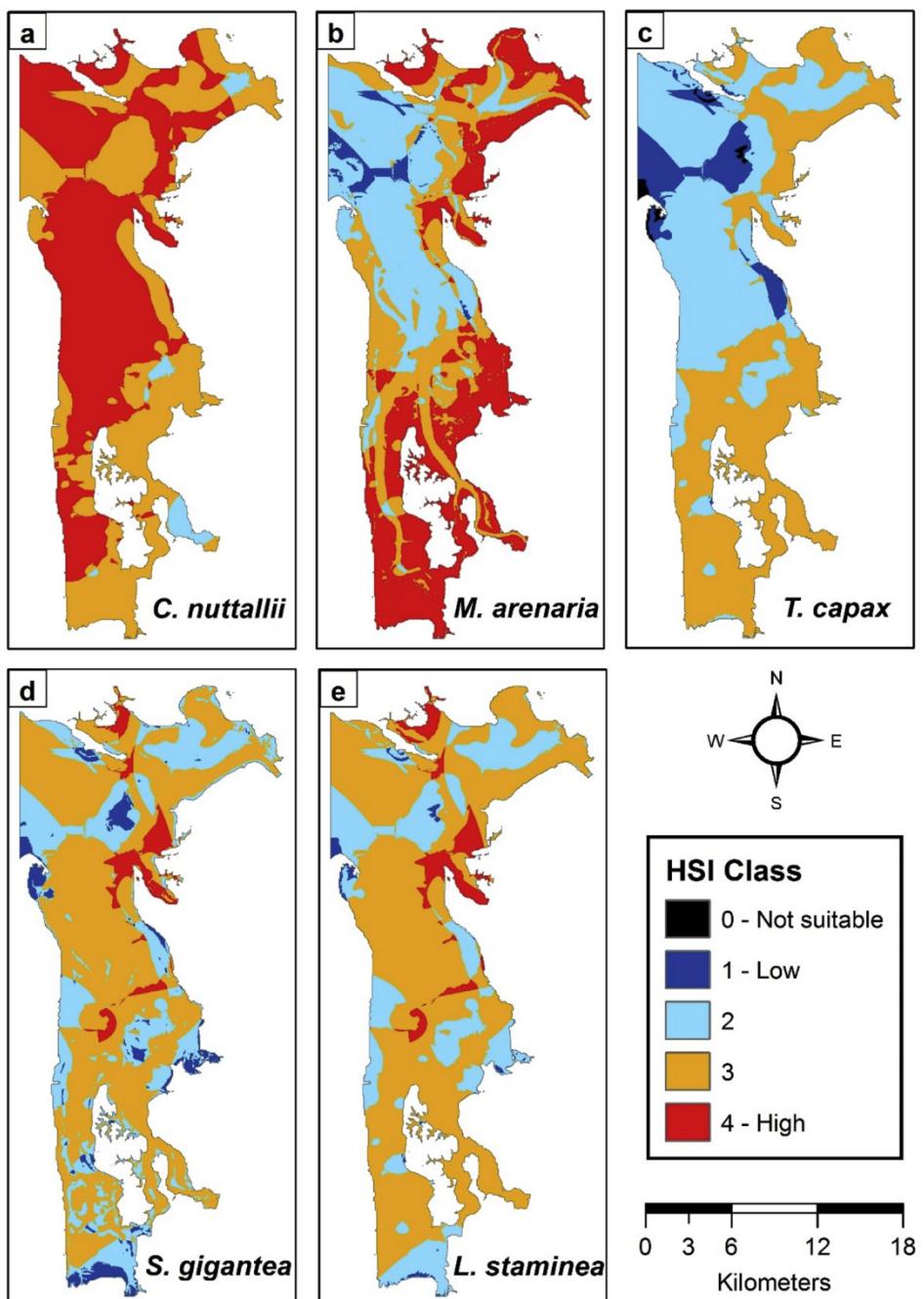


Fig. A1.8. Results from FISBHE in Willapa Bay, WA; estuary-wide HSI predictions for each species considered (a – *Clinocardium nuttallii*, b – *Mya arenaria*, c – *Tresus capax*, d – *Saxidomus gigantea*, e – *Leukoma staminea*). Predicted habitat suitability increases from HSI 0–4 (black to red).

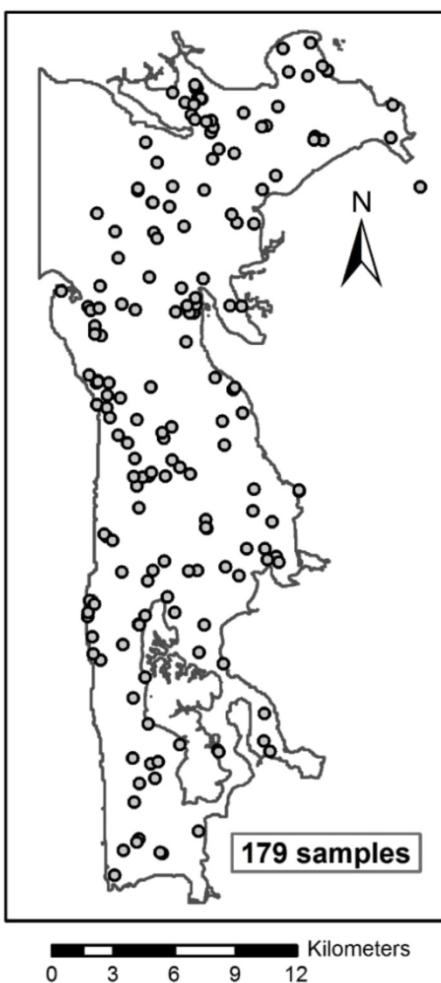


Fig. A1.9. Bay-clam sample locations in Willapa Bay, WA.

Table A1.1

Attributes of the geospatial point datasets used to interpolate habitat variables in FISBHE for Yaquina and Tillamook estuaries. Data sources: Environmental Protection Agency (EPA), Oregon Dept. of Fish and Wildlife (ODFW), Oregon Dept. of Environmental Quality (DEQ), National Ocean Service (NOS), U.S. Army Corps of Engineers (USACE), and U.S. Dept. of Agriculture (USDA).

Habitat variable	Estuary	Samples (points)	Date range	Sources
Bathymetric depth	Yaquina	54,295	1953–2015	ODFW, USACE, NOS
	Tillamook	49,084	1995	USACE
Sediment silt-clay	Yaquina	982	1996–2005	EPA
	Tillamook	627	1999–2011	EPA, ODFW
Mean wet-season salinity	Yaquina	97	1960–2006	DEQ
	Tillamook	22	1960–2007	DEQ
Burrowing shrimp	Yaquina	1883	1996–2012	EPA, ODFW, USDA
	Tillamook	751	1999–2011	EPA, ODFW

Appendix 2

Text A2.1. Detailed information on the establishment of habitat variables and sensitivities

In our review of the literature, we sought to determine several habitat variables that not only had a substantial influence the survival or distribution of bay-clam species, but could also be modeled geospatially with existing datasets. To achieve our desired outcome, these habitat variables could not be ephemeral (e.g., macroalgae blooms), but rather consistent components of the ecosystem. Temperature and dissolved oxygen fluctuations can obviously affect the survival of bay-clams (Lewis and DeWitt, 2017), but we categorize these more as ephemeral events that are difficult to predict, not habitat variables suitable for spatial and temporal interpolation.

We quickly determined that definitive sensitivity ranges within habitat variables do not exist for the bay-clam species of interest, as that would require multiple replicated studies that performed coupled field-sampling and laboratory experiments aimed at specific sensitivity measures, which simply do not exist. What do exist, however, are a handful of scientific literature that speak to general species ‘preferences’ or associations for each habitat variable of interest (i.e., bathymetric depth, sediment silt-clay %, low salinity, and the concurrence of burrowing shrimp). These associations

are presented as full or partial (e.g., optimal) ranges for multiple measures that each can imply sensitivity (e.g., presence, density, mortality, survival, growth), either qualitatively or quantitatively – highlighting the disparities among the sensitivity literature. We split the value range of each habitat variable within our framework into four subcategories to allow for a binary approach to assigning species suitability; this method also did not require a high degree of precision in determining species sensitivity ranges. We broadly defined the sensitivity range for a given species x habitat variable as the range (bounded by estimated minimum and maximum thresholds) in which a species commonly inhabits.

We examined these sources collectively to develop what we believed to be the ‘average’ sensitivity range for each species x habitat variable combination (Tables A2.1 – A2.4). This determination required some discretion on our part as marine ecologists familiar with the species and habitats. In making these decisions, we considered: the quantity of sources, overlap between the suggested ranges, the type of ranges, the measures used, the type of studies, whether the data were quantitative, location of the study, and alignment with our category thresholds. We limited the subjective decisions by establishing several guidelines: first, when the suggested sensitivity ranges (of all sources) for a species overlapped within a value category of a variable, that variable category was deemed suitable for that species; second, sensitivity ranges that were listed as optimal or ‘preferential’ were always included within our sensitivity range; third, quantitative data were favored over qualitative data; and lastly, PNW sources were favored over sources from elsewhere.

Table A2.1

Literature sources used to estimate the average bathymetric depth sensitivity range (to be used in FISBHE) for each species (orange bar). Listed with each source are fields indicating the cited sensitivity range (gray bar), accompanying measure of sensitivity, whether the cited range was full or partial, optimal or survivable, and whether the information was qualitative or quantitative.

		Bathymetric depth - MLLW (m)				Measure	Full range	Partial range	Optimal	Qual.	Quant.
		Subtidal	Low-intertidal	Mid-intertidal	High-intertidal						
	Source:	< -30	-0.79	0.4	1.56	≥3.11					
<i>C. nuttallii</i>	Endreny and Sylvia 2006						Presence	X		X	X
	Gallucci and Gallucci 1982						High density		X	X	X
	Gaumer and McCrae 1990						Presence	X			X
	Rudy et al. 2013						Presence	X			X
	Ratti 1977						Presence	X			X
	Coan et al. 2000						Presence	X	X	X	X
	Bernard 1983						Presence	X		X	X
	CBRAT						Presence	X		X	X
HSI model											
<i>M. arenaria</i>	Dethier 2006						Presence		X	X	X
	Endreny and Sylvia 2006						Density	X			X
	Gaumer and McCrae 1990						Presence	X			X
	Newell et al. 1986						Presence	X			X
	CBRAT						Presence	X		X	X
	HSI model										
<i>T. capax</i>	Dethier 2006						Presence	X		X	X
	Endreny and Sylvia 2006						Density	X			X
	Gaumer and McCrae 1990						Density	X			X
	Lauzier et al. 1998						Density	X		X	X
	White and Orr 2011						Abundance		X	X	X
	Zhang and Campbell 2002						Presence	X		X	X
	Bernard 1983						Presence	X		X	X
	CBRAT						Presence	X		X	X
HSI model											
<i>S. gigantea</i>	Dethier 2006						Presence		X	X	X
	Endreny and Sylvia 2006						Density	X			X
	Gaumer and McCrae 1990						Presence	X			X
	Bernard 1983						Presence	X		X	X
	CBRAT						Presence	X		X	X
	HSI model										
<i>L. staminea</i>	Chew and Ma 1987						Presence	X			X
	Dethier 2006						Presence	X		X	X
	Endreny and Sylvia 2006						Density	X			X
	Gaumer and McCrae 1990						Density	X			X
	Rudy et al. 2013						Presence	X			X
	Bernard 1983						Presence	X		X	X
	CBRAT						Presence	X		X	X
	HSI model										

Table A2.2

Literature sources used to estimate the average sediment silt-clay sensitivity range (to be used in FISBHE) for each species (orange bar). Listed with each source are fields indicating the cited sensitivity range (gray bar), accompanying measure of sensitivity, whether the cited range was full or partial, optimal or survivable, and whether the information was qualitative or quantitative.

		Sediment silt-clay (%)				Measure	Full range	Partial range	Optimal	Qual.	Quant.
		Sand	Sand/Gravel/Mud	Sand/Mud	Mud						
	Source:	0	7.5	39	60						
<i>C. nuttallii</i>	Dethier 2006					Presence	X			X	
	Gallucci and Gallucci 1982					Presence	X			X	
	Gaumer and McCrae 1990					Presence		X	X	X	
	Rudy et al. 2013					Presence				X	
	Coan et al. 2000					Presence		X	X	X	
	CBRAT					Presence	X				X
	HSI model										
<i>M. arenaria</i>	Dethier 2006					Presence	X			X	
	Gaumer and McCrae 1990					Presence		X	X	X	
	Newell et al. 1986					Presence	X			X	
	Rudy et al. 2013					Presence	X			X	
	CBRAT					Presence	X				X
	HSI model										
<i>T. capax</i>	Dethier 2006					Presence	X			X	
	Endreny and Sylvia 2006					Density	X		X	X	
	Gaumer and McCrae 1990					Density		X	X	X	
	Lauzier et al. 1998					Presence	X			X	
	White and Orr 2011					Density	X			X	
	CBRAT					Presence		X			X
<i>S. gigantea</i>	Dethier 2006					Presence		X	X	X	
	Endreny and Sylvia 2006					Density	X		X	X	
	Gaumer and McCrae 1990					Presence		X	X	X	
	Rudy et al. 2013					Presence	X			X	
	CBRAT					Presence	X				X
	HSI model										
<i>L. staminea</i>	Chew and Ma 1987					Presence					
	Dethier 2006					Presence		X	X	X	
	Endreny and Sylvia 2006					Presence	X			X	
	Gaumer and McCrae 1990					Density	X	X	X	X	
	Rudy et al. 2013					Presence		X	X	X	X
	CBRAT					Presence	X				X
HSI model											

Table A2.3

Literature sources used to estimate the average salinity sensitivity range (to be used in FISBHE) for each species (orange bar). Listed with each source are fields indicating the cited sensitivity range (gray bar), accompanying measure of sensitivity, whether the cited range was full or partial, optimal or survivable, and whether the information was qualitative or quantitative.

		Wet-season salinity (PSU)				Measure	Full range	Partial range	Optimal	Qual.	Quant.
Source:		Fresh	16	23	27	Marine	≥ 32.5				
<i>C. nuttallii</i>	Endreny and Sylvia 2006							High density	X	X	X
	Ratti 1977							Presence	X		X
	Liu et al. 2010							Expt.	X		X
	Bernard 1983							Mortality & growth	X		X
	CBRAT							Presence	X		X
	HSI model										
<i>M. arenaria</i>	Dethier 2006							Presence	X	X	X
	Endreny and Sylvia 2006							Presence	X		X
	Gaumer and McCrae 1990							Density	X		X
	Newell et al. 1986							Presence	X		X
	Rudy et al. 2013							Presence	X		X
	CBRAT							Presence	X		X
<i>T. capax</i>	Lauzier et al. 1998							Presence	X		X
	Bernard 1983							Mortality & growth	X		X
	CBRAT							Presence	X		X
	HSI model										
<i>S. gigantea</i>	Dethier 2006							Slow growth	X		X
	Bernard 1983							Presence	X		X
	CBRAT							Presence	X		X
	HSI model										
<i>L. staminea</i>	Chew and Ma 1987							Tolerance	X		X
	Dethier 2006							Tolerance	X		X
	Endreny and Sylvia 2006							Density	X	X	X
	Rudy et al. 2013							Presence	X		X
	Bernard 1983							Mortality & growth	X		X
	CBRAT							Presence	X		X
HSI model											

Table A2.4

Literature sources used to estimate the average burrowing shrimp sensitivity range (to be used in FISBHE) for each species (orange bar). Listed with each source are fields indicating the cited sensitivity range (gray bar), accompanying measure of sensitivity, whether the cited range was full or partial, optimal or survivable, and whether the information was qualitative or quantitative.

		Burrowing shrimp		Measure	Full range	Partial range	Optimal	Qual.	Quant.
Source:		Absent	Present						
<i>C. nuttallii</i>	Gaumer and McCrae 1990			Density	X	X	X		
	Hancock et al. 1979				X	X	X		
	HSI model								
<i>M. arenaria</i>	Gaumer and McCrae 1990			Density	X	X	X		
	Hancock et al. 1979				X	X	X		
	HSI model								
<i>T. capax</i>	Gaumer and McCrae 1990			Density	X	X	X		
	Hancock et al. 1979				X	X	X		
	HSI model								
<i>S. gigantea</i>	Gaumer and McCrae 1990			Density	X	X	X		
	Hancock et al. 1979				X	X	X		
	HSI model								
<i>L. staminea</i>	Gaumer and McCrae 1990			Density	X	X	X		
	Hancock et al. 1979				X	X	X		
	HSI model								

References

- Ainsworth, J.C., D'Andrea, A.F., Vance, M., Groth, S.D., Perotti, E.A., 2014. Status of Oregon Bay Clam Fisheries, Stock Assessment, and Research. Oregon Dept. of Fish and Wildlife, Newport, OR.
- Austin, M., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol. Model.* 157, 101–118.
- Bernard, R., 1983. Physiology and the mariculture of some northeastern Pacific bivalve molluscs. *Can. Spec. Publ. Fish. Aquat. Sci.* 63, 1–24.
- Blandford, B.L., Ripy, J., Grossardt, T.H., 2013. GIS-based Expert Systems Model for Predicting Habitat Suitability of Blackside Dace in Southeastern Kentucky. Research Report No. 13- 5059. Kentucky Transportation Center, University of Kentucky, Lexington, KY.
- Boyce, M.S., Vernier, P.R., Nielsen, S.E., Schmiegelow, F.K., 2002. Evaluating resource selection functions. *Ecol. Model.* 157, 281–300.
- Burrough, P., 2001. GIS and geostatistics: essential partners for spatial analysis. *Environ. Ecol. Stat.* 8, 361–377.
- Chew, K.K., Ma, A.P., 1987. Species Profiles: Life Histories and Environmental Requirements of Coastal Fishes and Invertebrates (Pacific Northwest) - Common Littleneck Clam. U.S. Fish Wildl. Serv. Bio. Report 82(11.78). U.S. Army Corps of Engineers, Vicksburg, MS.
- Coan, E.V., Scott, P.V., Bernard, F.R., 2000. Bivalve Seashells of Western North America: Marine Bivalve Mollusks from Arctic Alaska to Baja California. Santa Barbara Museum of Natural History, Santa Barbara, CA.
- Dethier, M.N., 2006. Native Shellfish in Nearshore Ecosystems of Puget Sound. Puget Sound Nearshore Partnership Report No. 2006-04. U.S. Army Corps of Engineers, Seattle, WA.
- Dethier, M.N., Ruesink, J., Berry, H., Sprenger, A.G., 2012. Decoupling of recruitment from adult clam assemblages along an estuarine shoreline. *J. Exp. Mar. Biol.* 422, 48–54.
- Endreny, P.A., Sylvia, G., 2006. Mapping a Path to Sustainable Utilization of Bay Clam Resources in Oregon's 21st Century Estuaries: an Investigation into the Potential of Oregon's Commercial Bay Clam Fishery: Report to the Oregon Department of Fish and Wildlife and the Oregon Commercial Bay Clam Industry. MS capstone project. Oregon State University, Corvallis, OR.
- Elith, J., Graham, C.H., 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32, 66–77.
- Emmett, R., Llansó, R., Newton, J., Thom, R., Hornberger, M., Morgan, C., Levings, C., Copping, A., Fishman, P., 2000. Geographic signatures of North American west coast estuaries. *Estuaries* 23, 765–792.
- Feldman, K.L., Armstrong, D.A., Dumbauld, B.R., DeWitt, T.H., Doty, D.C., 2000. Oysters, crabs, and burrowing shrimp: review of an environmental conflict over aquatic resources and pesticide use in Washington State's (USA) coastal estuaries. *Estuaries* 23, 141–176.
- Ferraro, S.P., Cole, F.A., 2011. Ecological periodic tables for benthic macrofaunal usage of estuarine habitats in the US Pacific Northwest. *Estuar. Coast Shelf Sci.* 94, 36–47.
- Ferraro, S.P., Cole, F.A., 2012. Ecological periodic tables for benthic macrofaunal usage of estuarine habitats: insights from a case study in Tillamook Bay, Oregon, USA. *Estuarine, Coastal and Shelf Science* 102, 70–83.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24, 38–49.
- Gallucci, V.F., Gallucci, B.B., 1982. Reproduction and ecology of the hermaphroditic cockle *Clinocardium nuttallii* (Bivalvia: Cardiidae) in Garrison bay. *Mar. Ecol. Prog. Ser.* 7, 137–145.
- Gaumer, T.F., McCrae, J., 1990. Status of Bay Clams in Oregon, 1975-88. Oregon Dept. of Fish and Wildlife, Newport, OR.
- Guisan, A., Edwards, T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Model.* 157, 89–100.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8, 993–1009.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147–186.
- Hancock, D., Gaumer, T., Wilhelke, G., Robart, G., Flynn, J., 1979. Subtidal Clam Populations: distribution, Abundance, and Ecology. Oregon State University Sea Grant Program, Corvallis, OR.
- Hickey, B.M., Banas, N.S., 2003. Oceanography of the US Pacific Northwest coastal ocean and estuaries with application to coastal ecology. *Estuaries* 26, 1010–1031.
- Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93, 679–688.
- Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecol. Model.* 199, 142–152.
- Hosmer, D.W., Lemeshow, S., 2000. Interpretation of the fitted logistic regression model. In: Applied Logistic Regression, second ed. John Wiley & Sons, Inc., Hoboken, NJ, pp. 47–90.
- Lauria, V., Power, A.M., Lordan, C., Weetman, A., Johnson, M.P., 2015. Spatial transferability of habitat suitability models of *Nephrops norvegicus* among fished areas in the Northeast Atlantic: sufficiently stable for marine resource conservation? *PLoS One* 10 (2), e0117006.
- Lauver, C.L., Busby, W.H., Whistler, J.L., 2002. Testing a GIS model of habitat suitability for a declining grassland bird. *Environ. Manag.* 30, 88–97.
- Lauzier, R., Hand, C., Campbell, A., Heizer, S., 1998. A Review of the Biology and Fisheries of the Horse Clams (*Tresus capax* and *Tresus Nuttallii*). Canadian Stock Assessment Secretariat, Research Document 98/88. Fisheries and Oceans Canada, Nanaimo, B.C.
- Lee II, H., Reusser, D.A., 2014. Coastal Biodiversity Risk Analysis Tool (CBRAT). www.cbrat.org, Accessed date: 26 February 2014.
- Lenoir, S., Beaugrand, G., Lecuyer, E., 2011. Modelled spatial distribution of marine fish and projected modifications in the North Atlantic Ocean. *Glob. Chang. Biol.* 17, 115–129.
- Levins, R., 1966. The strategy of model building in population biology. *Am. Sci.* 54, 421–431.
- Lewis, N.S., DeWitt, T.H., 2017. Effect of green macroalgal blooms on the behavior, growth, and survival of cockles (*Clinocardium nuttallii*) in Pacific NW estuaries. *Mar. Ecol. Prog. Ser.* 582, 105–120.
- Liu, W., Gurney-Smith, H., Beerens, A., Pearce, C., 2010. Effects of stocking density, algal density, and temperature on growth and survival of larvae of the basket cockle, *Clinocardium nuttallii*. *Aquaculture* 299, 99–105.
- Murphy, R.C., 1985. Factors affecting the distribution of the introduced bivalve, *Mercenaria mercenaria*, in a California lagoon—the importance of bioturbation. *J. Mar. Res.* 43, 673–692.
- Newell, C.R., Hidu, H., Parsons, J., 1986. Species Profiles: Life Histories and Environmental Requirements of Coastal Fishes and Invertebrates (North Atlantic) - Softshell Clam. U.S. Fish Wildl. Serv. Bio. Report 82(11.53). U.S. Army Corps of Engineers, Vicksburg, MS.
- Olivier, F., Wotherspoon, S.J., 2005. GIS-based application of resource selection functions to the prediction of snow petrel distribution and abundance in East Antarctica: comparing models at multiple scales. *Ecol. Model.* 189, 105–129.
- Pereira, J., Itami, R., 1991. GIS-based habitat modeling using logistic multiple regression - a study of the Mt. Graham red squirrel. *Photogramm. Eng. Rem. Sens.* 57, 1475–1486.
- Peterson, C., Scheidegger, K., Komar, P., Niem, W., 1984. Sediment composition and hydrography in six high-gradient estuaries of the northwestern United States. *J. Sediment. Res.* 54, 86–97.
- Power, M., 1993. The predictive validation of ecological and environmental models. *Ecol. Model.* 68, 33–50.
- R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are niche based species distribution models transferable in space? *J. Biogeogr.* 33, 1689–1703.
- Ratti, F.D., 1977. Reproduction and Growth of the Pacific Basket-Cockle, *Clinocardium Nuttallii* Conrad, from Intertidal and Subtidal Environments of Netarts Bay. MS thesis. Oregon State University, Corvallis, OR.
- Raybaud, V., Beaugrand, G., Dewarumez, J.-M., Luczak, C., 2015. Climate-induced range shifts of the American jackknife clam *Ensis directus* in Europe. *Biol. Invasions* 17, 725–741.
- Reiss, H., Cunze, S., König, K., Neumann, H., Kröncke, I., 2011. Species distribution modelling of marine benthos: a North Sea case study. *Mar. Ecol. Prog. Ser.* 442, 71–86.
- Robinson, K.A., Ramsay, K., Lindenbaum, C., Frost, N., Moore, J., Wright, A.P., Petrey, D., 2011. Predicting the distribution of seabed biotopes in the southern Irish Sea. *Cont. Shelf Res.* 31, S120–S131.
- Rudy Jr., P., Rudy, L.H., Shanks, A., Butler, B., 2013. Oregon Estuarine Invertebrates, second ed. University of Oregon, Eugene, OR.
- Schoeman, D., Richardson, A., 2002. Investigating biotic and abiotic factors affecting the recruitment of an intertidal clam on an exposed sandy beach using a generalized additive model. *J. Exp. Mar. Biol. Ecol.* 276, 67–81.
- Shirzad, F.F., Orlando, S.P., Klein, C.J., Holliday, S.E., Warren, M.A., Monaco, M.E., 1988. National Estuarine Inventory: Supplement 1, Physical and Hydrologic Characteristics, the Oregon Estuaries. National Oceanic and Atmospheric Administration, Rockville, MD.
- Thrush, S.F., Hewitt, J.E., Norkko, A., Nicholls, P.E., Funnell, G.A., Ellis, J.I., 2003. Habitat change in estuaries: predicting broad-scale responses of intertidal macrofauna to sediment mud content. *Mar. Ecol. Prog. Ser.* 263, 101–112.
- Valavanis, V., Pierce, G., Zuur, A., Palaiaxis, A., Saveliev, A., Katara, I., Wang, J., 2008. Modelling of essential fish habitat based on remote sensing, spatial analysis and GIS. *Hydrobiologia* 612, 5–20.
- Vaughan, I., Ormerod, S., 2005. The continuing challenges of testing species distribution models. *J. Appl. Ecol.* 42, 720–730.
- Vincenzi, S., Caramori, G., Rossi, R., De Leo, G.A., 2006. A GIS-based habitat suitability model for commercial yield estimation of *Tapes philippinarum* in a Mediterranean coastal lagoon (Saccia di Goro, Italy). *Ecol. Model.* 193, 90–104.
- White, L.F., Orr, L.C., 2011. Native clams facilitate invasive species in an eelgrass bed. *Mar. Ecol. Prog. Ser.* 424, 87–95.
- Ysebaert, T., Meire, P., Herman, P.M., Verbeek, H., 2002. Macrobenthic species response surfaces along estuarine gradients: prediction by logistic regression. *Mar. Ecol. Prog. Ser.* 225, 79–95.
- Zhang, Z., Campbell, A., 2002. Application of a stochastic spawning stock biomass per recruit model for the horse clam fishery in British Columbia. *Fish. Res.* 57, 9–23.