

# Winter distribution of blue crab *Callinectes sapidus* in Chesapeake Bay: application and cross-validation of a two-stage generalized additive model

Olaf P. Jensen<sup>1,3,\*</sup>, Ralf Seppelt<sup>2,4</sup>, Thomas J. Miller<sup>1</sup>, Laurie J. Bauer<sup>1</sup>

<sup>1</sup>University of Maryland Center for Environmental Science Chesapeake Biological Laboratory, PO Box 38, 1 Williams St., Solomons, Maryland 20688, USA

<sup>2</sup>Technical University Braunschweig, Institute of Geoecology, Department for Environmental System Analysis, Langer Kamp 19c, 38106 Braunschweig, Germany

<sup>3</sup>Present address: University of Wisconsin Center for Limnology, 680 N. Park St., Madison, Wisconsin 53706, USA

<sup>4</sup>Present address: UFZ Centre for Environmental Research Leipzig, Department for Applied Landscape Ecology, Permoserstraße 15, 04318 Leipzig, Germany

**ABSTRACT:** We present a 2-stage generalized additive model (GAM) of the distribution of mature female blue crab *Callinectes sapidus* in Chesapeake Bay based on data from a fishery-independent winter dredge survey. The distribution and abundance of blue crabs was modeled as a flexible function of depth, salinity, water temperature, distance from the Bay mouth, distance from submerged aquatic vegetation (SAV), and bottom slope for each of the 13 yr of data available. Depth, salinity, temperature, and distance from the Bay mouth were found to be the most important environmental determinants of mature female blue crab distributions. The response curves for these variables displayed patterns that are consistent with laboratory and field studies of blue crab/habitat relationships. The generality of the habitat models was assessed using intra- and inter-annual cross-validation. Although the models generally performed well in cross-validation, some years showed unique habitat relationships that were not well predicted by models from other years. Such variability may be overlooked in habitat suitability models derived from data collected over short time periods.

**KEY WORDS:** Blue crab · Chesapeake Bay · GAM · Cross-validation · Habitat suitability model

—Resale or republication not permitted without written consent of the publisher—

## INTRODUCTION

Blue crab *Callinectes sapidus* ranges along the Atlantic coast of the American continent from Brazil to Canada. Evidence does not support the existence of distinct genetic populations, but functional sub-populations are recognized with only limited exchange between them (McMillen-Jackson et al. 1994). Over the species' wide latitudinal range, individual sub-populations can experience markedly different environments. Temperature is likely the key environmental parameter causing the variation observed in life history schedules (Smith 1997). Central to temperature's role is the existence of a physiological minimum

temperature ( $T_{\min}$ ) close to 10°C, below which molting, and hence growth, ceases (Brylawski & Miller 2003). As temperatures increase above  $T_{\min}$  the period between molts shortens, and thus overall rates of growth increase. In particular, the proportion of the year during which temperatures are above  $T_{\min}$  is an important determinant of the life history pattern expressed. In the Gulf of Mexico, average bottom water temperature is above  $T_{\min}$  throughout the year, and crabs grow sufficiently fast so that they can mature, reproduce, and enter the commercial fishery in a single year. In contrast, water temperatures in mid-latitudes are unfavorable between late November and late April. During this period, crabs enter a dor-

\*Email: ojensen@wisc.edu

mant phase during which they bury into the sediments. Thus in mid-latitude populations, growth and maturation occur in different years so that individuals take a minimum of 18 to 24 mo to complete their life cycle.

Because overwintering blue crabs in Chesapeake Bay do not feed and are unlikely to experience significant predation, bioenergetic costs are likely to play a dominant role in determining overwintering survival. Laboratory studies (Tagatz 1969, McKenzie 1970) have shown that salinity and temperature interact with thermal tolerances dependant on both salinity and acclimation temperature. These results suggest that salinity and temperature, as well as factors such as depth, which might serve to limit temperature fluctuations, may be important in determining choice of overwintering habitat; however, no studies to date have examined the extent to which winter distributions of blue crab reflect differences in these variables.

The blue crab population in the Chesapeake Bay supports the single largest blue crab fishery in North America. Assessments of this stock indicate recent declines in both abundance and landings (Miller & Houde 1998, Rugolo et al. 1998, Lipcius & Stockhausen 2002) despite efforts to reduce fishing mortality. Winter distributions in the Bay are an area of recent concern for several reasons. Most directly, estimates of abundance and rates of exploitation of blue crab in Chesapeake Bay, on which stock assessments have been based, have been derived from a baywide, fishery-independent winter dredge survey (WDS) conducted between December and March (Sharov et al. 2003). Also making the winter distribution of crabs important is concern over the vulnerability of spawning females in a winter dredge fishery (Miller 2003), the efficacy of a dispersal corridor (Lipcius et al. 2001), and a combined marine protected area and dispersal corridor (Lipcius et al. 2003) that has recently been established in the Virginia (southern) portion of Chesapeake Bay. Thus, the ability to predict blue crab winter distribution has become desirable.

Generalized additive models (GAMs) provide a flexible non-parametric or semi-parametric framework to model the relationship between a response and one or more predictor variables (Hastie & Tibshirani 1990). GAMs do not require the distributional assumptions of traditional parametric approaches and provide the ability to fit flexible non-linear response curves to individual predictor variables. In GAMs, the response variable is assumed to be the sum of separate individual functions of each of the predictor variables with a link function appropriate to the distribution of the response variable (e.g. a Poisson link function is often specified for count data). In the more familiar generalized linear models (GLMs), these individual functions of the predictor variables are linear. In GAMs, the individual

functions may be linear or may be non-parametric smoothers such as regression splines, which are better-suited to modeling many common biological response curves such as threshold functions. Different functions may be specified for each predictor variable, allowing for response curves that are specific to the individual predictors.

The use of GAMs to model organism/habitat relationships increased following publication of Hastie & Tibshirani's (1990) book and Swartzman et al.'s (1992) application of this technique to model groundfish in the Bering Sea. GAMs have since become widely used in marine sciences to predict abundance and identify important habitats (e.g. Swartzman et al. 1995, Maravelias et al. 2000a) and to model stock-recruitment relationships (Cardinale & Arrhenius 2000).

Two-stage GAMs are an extension of the basic structure in which the response variable is modeled first as a binomial variable (presence/absence or yes/no) and secondly the non-zero observations (presence or yes) are modeled as a continuous or count variable, usually with a Gaussian or Poisson distribution respectively. The 2 stages may then be combined multiplicatively to yield an overall prediction (Barry & Welsh 2002). This approach is particularly useful in modeling aquatic organisms, for which, because of their patchy distributions, survey catches are often zero-inflated (Maravelias 1999). Two-stage GAMs have been used in fisheries to improve estimates of various stock assessment indices (e.g. Borchers et al. 1997, Piet 2002) and to model salmon feeding and growth (Rand 2002).

However, despite the widespread use of GAMs, studies have yet to examine their ability to find general relationships that are valid beyond the particular data set or year modeled. A risk of using highly flexible, data-driven methods is that the resulting predictive models may fit the modeled data so specifically that they may have little predictive power when applied to other data sets. The underlying goal of most habitat modeling studies is not simply to describe the trends in the modeled data, but also to produce predictions valid in other years/locations. Ideally, GAMs should produce an understanding of the functional relationship between an organism and various components of its environment. Cross-validation, by testing the ability of models based on one data set to accurately predict values in another, is a useful means of assessing the generality of a model.

Here we fit 2-stage GAMs for each of 13 yr of data from the WDS to determine the environmental variables that regulate winter distribution of mature female blue crab in Chesapeake Bay. Subsequently, we use cross-validation to assess the ability of models developed from data collected in 1 yr to predict the distribution and abundance of crabs in other years.

## MATERIALS AND METHODS

We modeled the distribution and abundance of mature female blue crab in Chesapeake Bay. Mature females were chosen as the focus of this study because of their greater per capita contribution to future generations and because current management strategies, including the lower Bay spawning sanctuary, are focused specifically on their protection.

**Data.** The WDS has been conducted annually between December and March since the winter of 1989 to 1990. Full details of the survey design are provided in Sharov et al. (2003), and are only summarized here. Survey years will henceforth be referred to by the year in which the survey was completed, e.g. the first survey is the 1990 survey. Stratification and sample size in the first 3 yr of the survey were different than in the remaining years, but except for this change the survey has been conducted consistently throughout the period of record. From 1993 to the present, 1255 to 1599 stratified random stations were sampled within 3 region-based strata. During the period from 1990 to 1992, there were more strata and generally fewer (867 to 1395) samples. A typical distribution of station locations and densities of mature female blue crabs is shown in Fig. 1. One minute tows of a 1.83 m wide crab dredge were made at each station. The length of each tow was determined by either Loran-C or a differential global positioning system (DGPS). All crabs greater than 15 mm carapace width were measured, sexed, and enumerated. Additionally, environmental parameters were measured at each station. Depletion experiments (Zhang et al. 1993, Vølstad et al. 2000), in which the same area was dredged repeatedly, were conducted yearly since 1992 to determine the fraction of blue crabs sampled by a single dredge tow, i.e. the catchability coefficient ( $q$ ). Based on these experiments, Vølstad et al. (2000) applied an exponential depletion model to estimate vessel and year specific catchability coefficients that were used to transform catch at each station into an estimate of absolute abundance.

**Environmental variables.** Six environmental variables were chosen for consideration in the GAMs based on availability and known or plausible roles in influencing blue crab distributions. Depth was measured at each WDS site. The 5 remaining variables, salinity, water temperature, distance from the Bay mouth, distance from the nearest submerged aquatic vegetation (SAV), and bottom slope, were derived using data from other sources and a geographic information system (GIS).

Although surface salinity and temperature were measured at each WDS site, the more relevant measurements for describing blue crab winter habitat

choice are the bottom salinity and temperature at the time when they bury into the sediment. For this reason, temperature and salinity used in this analysis were interpolated from Chesapeake Bay Water Quality Monitoring Program data (Chesapeake Bay Program 1993). December bottom temperature and salinity maps were produced for each year using data collected at 99 to 123 sites per year. The data were first spatially detrended in order to satisfy the assumption of first-order stationarity (Cressie 1993). Detrending was conducted using linear regression with northing, easting, and northing  $\times$  easting interaction terms. Variogram modeling and ordinary kriging were conducted on the

1998 Density of mature females

- (log # 1000m<sup>-2</sup>)
- 0.00
  - 0.01–4.00
  - 4.01–7.44

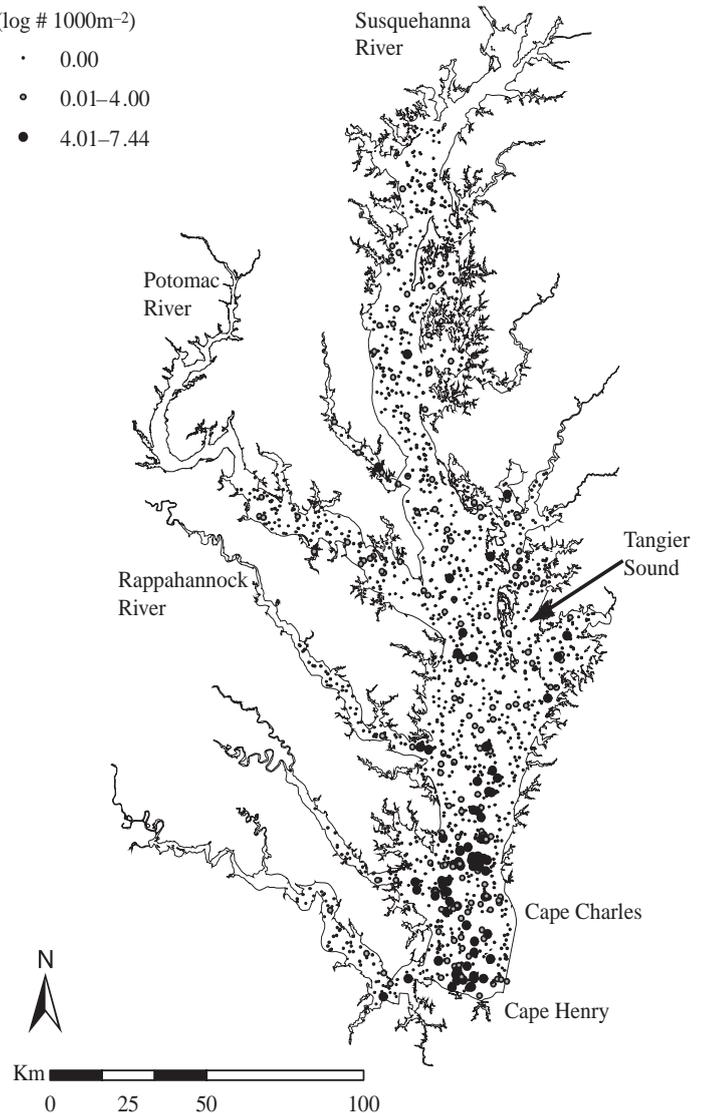


Fig. 1. *Callinectes sapidus*. Density of mature female blue crabs measured at winter dredge survey stations during the 1998 survey

residuals before adding the trend back to the kriged predictions. Gaussian, spherical, and exponential variogram models were fit to empirical variograms using non-linear least squares (SAS, NLIN procedure) and the best fitting model (lowest mean squared error) was used for kriging (SAS, KRIGE2D procedure). Interpolated bottom temperature and salinity were mapped in ArcView v8.3 and maps from the previous December were used to assign values to each WDS site.

Distance from the Bay mouth was calculated along the shortest through-water path between the dredge start point and a point in the mouth of the Bay midway between Cape Henry and Cape Charles. This distance was calculated in ArcView v8.3 using a customized script based on the lowest-cost path function and a raster map of the Bay with a resolution of 250 m (Jensen 2004). This variable was chosen based on a preference by mature females for higher salinity water in the lower Bay waters where their offspring may be more easily transported offshore (Johnson 1995). Distance from SAV was chosen as a potentially important environmental variable because of known affinities by blue crabs for SAV (Orth et al. 1996) during the spring and summer and the hypothesis that mature females may choose the nearest suitable winter habitat. In agreement with this hypothesis, SAV distributions from the previous summer were used, e.g. 1989 SAV distributions were used to predict 1990 (i.e. winter 1989 to 1990) crab distributions. Distance from the nearest SAV beds was calculated using maps of annual Chesapeake Bay SAV distributions derived from aerial photography (Orth et al. 2001). Distance from SAV was calculated as the straight-line distance and was log transformed in order to conform to a normal distribution.

Schaffner & Diaz (1988) found significant differences in the abundance of blue crabs among different sediment types and bottom morphologies (shallow spits and shoals, basins, and channels). Maps of Chesapeake Bay bottom type are not of sufficient spatial and temporal resolution to allow us to incorporate bottom type directly into our analyses. Accordingly, we used bottom slope as a proxy for benthic habitat type. This approach differentiates channel (high slope) from basin (low slope) regions and may also reflect differences in sediment type assuming that finer sediments are more likely to accumulate in low slope areas. Bottom slope was calculated from a high-resolution (30 m) bathymetric digital elevation layer (National Oceanic and Atmospheric Administration 1998) in ArcView v8.3. Bottom slope was log transformed and multiplied by a factor of 10 in order to conform to a normal distribution on a similar scale as the other environmental variables.

**Two-stage generalized additive models.** Two-stage GAMs were used to describe the relationship between

mature female blue crab density and the 6 environmental variables. Models were developed independently for each year of data using a randomly selected training subset representing 75% of the data in an individual year (650 to 1199 stations). The remaining 25% of the data were reserved for cross-validation. In the first stage of the analysis, presence or absence of mature female crabs was modeled using a logistic model with a binomial error distribution and a logit link function. In this stage, the estimated probability of crab occurrence ( $\hat{p}$ ) at any site was modeled as an additive function of the 6 environmental variables:  $D$  = water depth (m),  $M$  = distance to the Bay mouth (km),  $V$  = distance (km) to SAV beds,  $S$  = salinity (ppt),  $B$  = bottom slope, and  $T$  = water temperature ( $^{\circ}\text{C}$ ), given by:

$$\hat{p} = s(D) + s(M) + s(V) + s(S) + s(B) + s(T) \quad (1)$$

where the  $s$ 's are unique regression spline functions for each environmental variable. Penalized regression splines (Wood & Augustin 2002) were fitted using the *mgcv* (v1.0-9) package for R v1.9.1.

In the second stage of the model, log transformed mature female blue crab density ( $\# 1000 \text{ m}^{-2}$ ) of only those stations at which at least one mature female crab was caught was modeled as a function of environmental covariates with a Gaussian error distribution. The model equation was:

$$\ln(\hat{\mu}) = s(D) + s(M) + s(V) + s(S) + s(B) + s(T) \quad (2)$$

where  $\hat{\mu}$  is the predicted density of mature female blue crabs given occurrence, and the other variables are as given above. Subsequently, the predicted log abundance,  $\ln(\hat{y})$ , at a given location was calculated as the product of Stage I and Stage II (Barry & Welsh 2002):

$$\ln(\hat{y}) = \hat{p} \times \ln(\hat{\mu}) \quad (3)$$

The flexibility of the response curves was optimized using an iterative method that rewards model fit and penalizes degrees of freedom (Wood 2000). This approach avoids the subjectivity inherent in choosing degrees of freedom *a priori* and ensures that the models provide the best fit with the fewest degrees of freedom. An initial full model containing all 6 variables was simplified by removing insignificant variables (backward elimination) until all remaining variables were significant ( $\alpha = 0.05$ ). All possible 2-variable interactions using the remaining variables were then added to the model, and the model was again pared down to only significant terms. Non-significant main effect terms (i.e. a single response variable with no interaction) were retained, however, if they were also part of a significant interaction. In 2 instances, the model-fitting algorithm would not converge when the degrees of freedom for an interaction term were not

fixed. In these cases, a range of plausible degrees of freedom were given, and the model with the highest adjusted  $R^2$  was selected.

**Model fit, cross-validation, and mapping.** Receiver operating characteristic (ROC) curves were used to assess the fit and generality of Stage I (presence/absence) models. Although ROC curves are commonly used to assess logistic regression models (Hosmer & Lemeshow 2000) and have been used to assess habitat models developed through logistic regression (e.g. Bonn & Schroder 2001, Scholten et. al. 2003, Gibson et al. 2004), they can also be applied to any model that produces estimates of  $p$  (the probability of presence) for a binomially distributed response variable. ROC curves are simply a plot of sensitivity (the fraction of correctly predicted presences) against  $1 - \text{specificity}$  (the fraction of correctly predicted absences) with changing critical values of  $p$  ( $p_{\text{crit}}$ , the probability above which presence is predicted). An ROC curve for a model with no discriminatory power is simply a straight line with a slope of one, i.e. as  $p_{\text{crit}}$  changes, any increase in sensitivity is offset by an equivalent loss of specificity. ROC curves are used here to calculate the area under the ROC curve (AUC, a measure of discriminatory power),  $p_{\text{opt}}$  (the value of  $p$  which results in the highest percentage of correct predictions), and  $p_{\text{fair}}$  (the value of  $p$  for which sensitivity and specificity are equal). AUC is a threshold-independent (i.e. it does not depend on a specified  $p_{\text{crit}}$ ) summary statistic that ranges from 0 (no discriminatory power) to 1 (perfect discriminatory power) and has been previously used to assess the generality of logistic regression-based habitat models (Bonn & Schroder 2001). Although criteria for evaluating AUC values are to some extent arbitrary, Hosmer & Lemeshow (2000) suggest the following cut-offs: 0.7 to 0.8 acceptable, 0.8 to 0.9 excellent,  $>0.9$  outstanding.

Cross-validation was also used to assess the transferability of the combined models (the product of Stage I and Stage II) fitted to training data sets to a separate test data set from the same year (i.e. intra-annual cross-validation) or to data from another year (i.e. inter-annual cross-validation). The predictive ability of each combined model was assessed by regressing predicted values on the observed values. The resulting least-squares correlation coefficient was used as an index of model performance. We tested 2 hypotheses: (1) Models fitted to a training data set perform better (i.e. higher  $r$ ) on the training data than on independent test data from the same year. (2) Models perform better in intra-annual cross-validation

than when applied to data from other years (inter-annual cross validation).

To test these hypotheses, the Fisher (1915) transformation was used to normalize the cross-validation correlation coefficients ( $r$ ). The first hypothesis was tested using a paired  $t$ -test of the transformed correlation coefficients and the second was tested using a  $t$ -test for 2 samples with equal variance. To aid interpretation of the results of the cross-validation analyses we compared all individual models to the grand mean (including inter- and intra-annual) cross-validation  $r$ . We calculated standardized normal deviates:

$$z = \frac{(r_{i,j} - \bar{r})}{s} \quad (4)$$

where  $r_{i,j}$  is the Fisher (1915) transformed coefficient of determination for predictions from the model year  $i$ , applied to observed year  $j$ ,  $\bar{r}$  is the grand mean, and  $s$  the sample standard deviation of  $\bar{r}$ .

In order to visualize predicted mature female blue crab distributions, predictions from the most general model (i.e. the model with the highest mean cross-validation  $r$ -squared value), were mapped for Stage I, Stage II, and the combined model. Predictions were made for  $1 \times 1$  km grid cells based on the values of the predictor variables for each cell. Values of the dynamic predictor variables (temperature, salinity, and distance from SAV) used in mapping were the values within each grid cell for the summer (distance from SAV) or December (temperature and salinity) preceding the most general model.

## RESULTS

### Model development

Significant correlations were present between many pairs of explanatory variables (Table 1). Most notably, there was a strong and negative correlation ( $r = -0.64$ ) between salinity and distance from the Bay mouth.

Table 1. Pearson correlation coefficients between all pairs of environmental variables. Significant correlations ( $p < 0.05$ ) are shown in **bold**. All other correlations are insignificant ( $p > 0.05$ )

	Distance ( <i>M</i> )	Salinity ( <i>S</i> )	Depth ( <i>D</i> )	Temp. ( <i>T</i> )	Slope ( <i>B</i> )	SAV ( <i>V</i> )
Distance ( <i>M</i> )	1					
Salinity ( <i>S</i> )	<b>-0.637</b>	1				
Depth ( <i>D</i> )	<b>-0.038</b>	<b>0.192</b>	1			
Temp. ( <i>T</i> )	<b>-0.091</b>	<b>0.341</b>	<b>0.118</b>	1		
Slope ( <i>B</i> )	<b>0.107</b>	<b>-0.073</b>	<b>0.284</b>	-0.002	1	
SAV ( <i>V</i> )	-0.010	<b>-0.133</b>	<b>0.139</b>	<b>0.079</b>	<b>-0.055</b>	1

Moderately strong correlations occurred between salinity and temperature ( $r = 0.34$ ), and between depth and bottom slope ( $r = 0.28$ ). Although the correlations among the explanatory variables were often statistically significant, even the 2 most strongly correlated variables (salinity and distance from the Bay mouth) do not overlap entirely as salinity patterns are strongly influenced by freshwater flow from the western shore tributaries, which, combined with the Coriolis effect, results in a pattern of lower salinities in the western region of the Chesapeake Bay. Colinearity among the explanatory variables was not deemed sufficient to drop variables from the full models, but will be considered in the interpretation of model selection results.

All 6 explanatory variables were included as either significant main effects or in interaction terms in at least 3 of the final models; however, no variable occurred in all models (Table 2). Distance from the Bay mouth and depth were the most commonly included variables. In Stage I, distance from the Bay mouth was significant in 9 out of 13 yr and depth was

significant in all years. Distance from Bay mouth appeared in 10 out of 13 yr for Stage II models, while depth was included in 5 Stage II models. Water temperature also appeared commonly in Stage I, occurring in 9 out of 13 yr, but was only found to be significant in 2 of the Stage II models. Salinity was included in 8 yr for Stage I and in 2 yr for Stage II. Importantly, salinity was often included in Stage I models when distance from the Bay mouth was not. The remaining 2 variables, bottom slope and distance from SAV, were not often found to be significant in either model stage.

Penalized regression spline fits of individual environmental factors to blue crab density varied from simple linear functions to highly complex curves. We present response curves for the most commonly included variables, distance from the Bay mouth (Fig. 2), salinity (Fig. 3), depth (Fig. 4), and temperature (Fig. 5), for all years in which they were included as significant main effect terms only, i.e. not in interaction. All other response curves are available at: <http://hjort.cbl.umces.edu/crabs/GAM.html>

Table 2. *Callinectes sapidus*. Model selection results for (a) Stage I (presence/absence) and (b) Stage II (abundance) GAMs. Significance test p-values are given for the explanatory variables distance from Bay mouth, salinity, depth, temperature, bottom slope, distance from SAV, and interaction terms. Terms that were not significant (ns,  $p > 0.05$ ) were dropped from the model unless they were involved in a significant interaction (I). Degrees of freedom (df) were fixed for terms in **bold**. The adjusted  $R^2$  and percent of deviance explained are also given for each model

Year	N	df	Distance (M)	Salinity (S)	Depth (D)	Temp. (T)	Slope (B)	SAV (V)	Interaction		$R^2$ adj	Deviance explained
									I	II		
(a)												
1990	650	25.4	ns	<0.001	I	I	ns	ns	<i>D/T</i> 0.027	ns	0.274	31.5
1991	723	27.8	I	ns	I	ns	0.048	ns	<i>M/D</i> <0.001	ns	0.246	26.1
1992	1046	10.3	ns	<0.001	<0.001	ns	ns	ns	ns	ns	0.221	22
1993	941	22.7	ns	I	I	<0.001	ns	0.044	<i>S/D</i> <0.001	ns	0.255	27.1
1994	1071	37.1	I	I	<0.001	ns	ns	0.036	<i>M/S</i> 0.036	<i>M/V</i> 0.030	0.225	28.1
1995	1199	18.5	ns	I	<0.001	I	ns	ns	<b><i>S/T</i> 0.005</b>	ns	0.097	17.7
1996	1187	32.5	<0.001	I	I	<0.001	ns	I	<i>S/V</i> 0.001	<i>S/D</i> <0.001	0.255	27.6
1997	1193	16.1	I	<0.001	I	<0.001	ns	ns	<i>M/D</i> <0.001	ns	0.189	23.1
1998	1181	29.2	I	ns	I	<0.001	ns	ns	<i>M/D</i> <0.001	ns	0.267	28.3
1999	1139	22.4	<0.001	ns	I	I	ns	ns	<i>D/T</i> 0.014	ns	0.197	27.2
2000	1133	17.3	<0.001	ns	<0.001	<0.001	ns	ns	ns	ns	0.27	28.6
2001	1167	25.0	I	ns	I	<0.001	ns	ns	<b><i>M/D</i> 0.002</b>	ns	0.328	38.4
2002	1148	18.9	I	0.012	I	ns	ns	0.018	<i>M/D</i> 0.002	ns	0.215	29.8
(b)												
1990	91	3.0	ns	ns	ns	0.004	ns	0.007	ns	ns	0.136	16.5
1991	161	9.0	<0.001	ns	0.002	ns	0.002	0.047	ns	ns	0.31	34.9
1992	197	2.0	<0.001	ns	0.003	ns	ns	ns	ns	ns	0.116	12.5
1993	166	2.2	<0.001	ns	ns	ns	ns	ns	ns	ns	0.145	15.6
1994	150	6.3	<0.001	ns	ns	ns	ns	ns	ns	ns	0.265	29.6
1995	88	11.0	0.002	ns	ns	ns	0.047	ns	ns	ns	0.219	31.8
1996	204	10.8	<0.001	ns	<0.001	ns	ns	ns	ns	ns	0.359	39.3
1997	149	12.0	I	ns	I	ns	ns	0.037	<i>M/D</i> 0.002	ns	0.457	50.1
1998	185	7.8	0.022	<0.001	0.007	ns	ns	ns	ns	ns	0.393	41.9
1999	102	4.9	<0.001	ns	ns	0.002	ns	ns	ns	ns	0.318	35.1
2000	193	2.6	ns	<0.001	ns	ns	ns	ns	ns	ns	0.182	19.3
2001	116	1.0	<0.001	ns	ns	ns	ns	ns	ns	ns	0.136	14.4
2002	100	2.3	ns	ns	ns	ns	ns	0.024	ns	ns	0.077	9.8

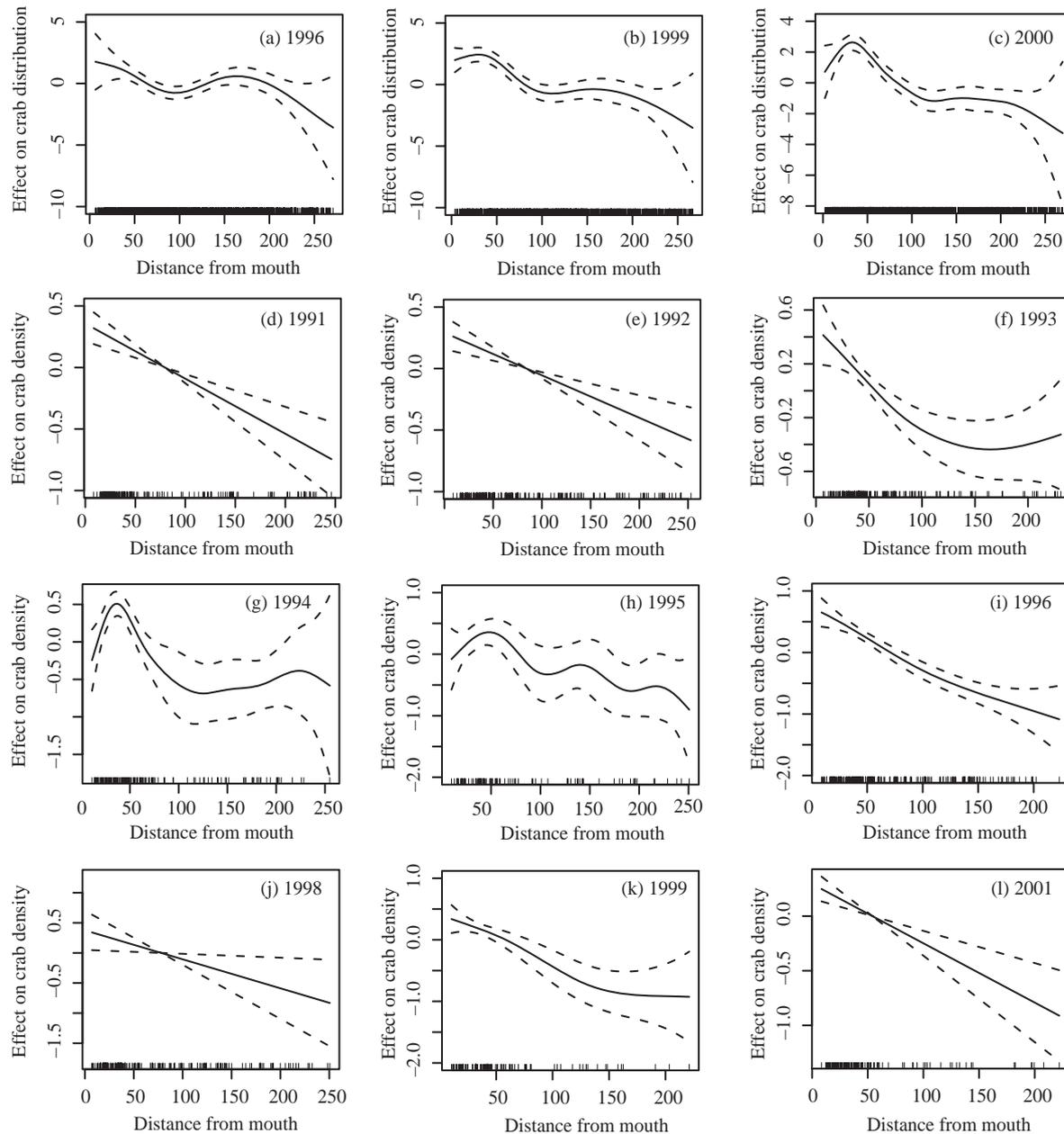


Fig. 2. *Callinectes sapidus*. Conditional regression spline smooths of distance from Bay mouth (km) on the probability of mature female blue crab presence for (a) 1996, (b) 1999, and (c) 2000 and density (ind.  $1000\text{ m}^{-2}$ ) given presence for (d–i) 1991 to 1996, (j) 1998, (k) 1999, and (l) 2001. Smooths are shown only for those years in which the variables were significant ( $p < 0.05$ ) and not included in an interaction term. The y-axis is the normalized effect of the variable; rugplot on the x-axis represents the number of observations; dashed lines are the  $\pm 2$  SE confidence belts

Relationships between crab distribution and abundance and distance from the Bay mouth showed 2 dominant patterns. A linear decline in crab density with increasing distance from the Bay mouth was seen in 4 of the 7 Stage II models examined (Fig. 2d,e,j,l). The second pattern, a maximum at approximately 25 to 50 km, was observed in 2 Stage I (Fig. 2b,c) and 2 Stage II models (Fig. 2g,h). The shape of these re-

sponse curves at greater distances from the Bay mouth was highly variable, reflecting the relatively rare catch of mature female blue crabs in the upper Bay, and some curves (Fig. 2c,f,g,k) suggest that the decline in crab density may level off beyond 100 km from the Bay mouth.

Response curves for salinities below 15 to 20 ppt were characterized by lower probabilities of pres-

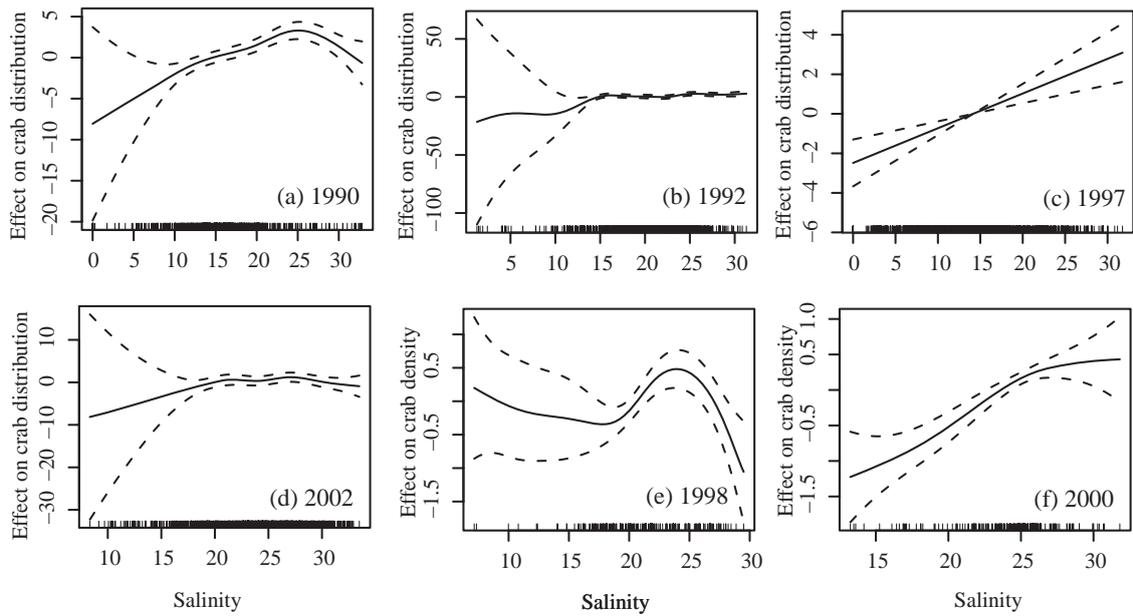


Fig. 3. *Callinectes sapidus*. Conditional regression spline smooths of salinity (ppt) on the probability of mature female blue crab presence for (a) 1990, (b) 1992, (c) 1997, and (d) 2002 and density (ind. 1000 m<sup>-2</sup>) given presence for (e) 1998 and (f) 2000. Smooths are shown only for those years in which the variables were significant ( $p < 0.05$ ) and not included in an interaction term. The y-axis is the normalized effect of the variable; rugplot on the x-axis represents the number of observations; dashed lines are the  $\pm 2$  standard error confidence belts

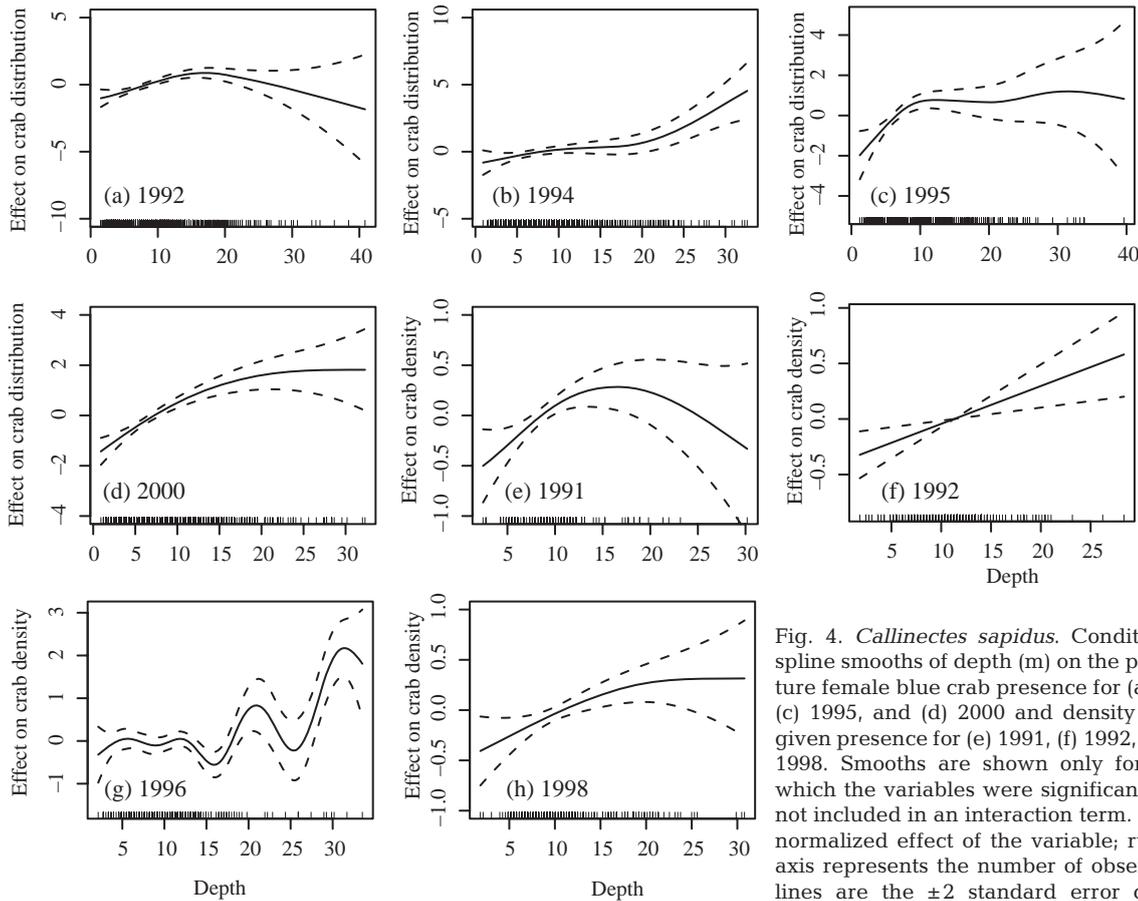


Fig. 4. *Callinectes sapidus*. Conditional regression spline smooths of depth (m) on the probability of mature female blue crab presence for (a) 1992, (b) 1994, (c) 1995, and (d) 2000 and density (ind. 1000 m<sup>-2</sup>) given presence for (e) 1991, (f) 1992, (g) 1996, and (h) 1998. Smooths are shown only for those years in which the variables were significant ( $p < 0.05$ ) and not included in an interaction term. The y-axis is the normalized effect of the variable; rugplot on the x-axis represents the number of observations; dashed lines are the  $\pm 2$  standard error confidence belts

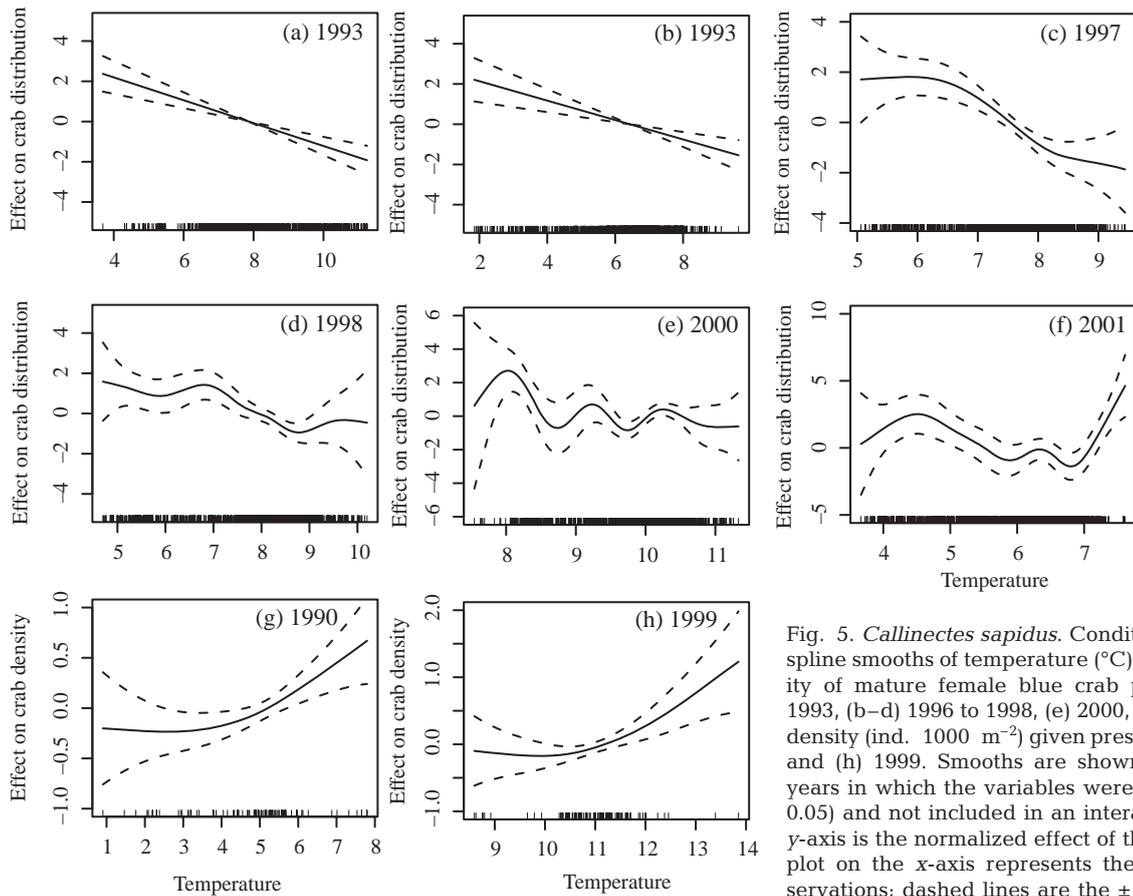


Fig. 5. *Callinectes sapidus*. Conditional regression spline smooths of temperature ( $^{\circ}\text{C}$ ) on the probability of mature female blue crab presence for (a) 1993, (b–d) 1996 to 1998, (e) 2000, and (f) 2001 and density (ind.  $1000\text{ m}^{-2}$ ) given presence for (g) 1990 and (h) 1999. Smooths are shown only for those years in which the variables were significant ( $p < 0.05$ ) and not included in an interaction term. The y-axis is the normalized effect of the variable; rug-plot on the x-axis represents the number of observations; dashed lines are the  $\pm 2$  standard error confidence belts

ence and lower abundance given presence as well as extreme variability due to the smaller number of samples at low salinity. Some curves (Fig. 3a,e,f) showed a maximum or a leveling off at approximately 25 ppt.

The relationship between crab abundance and depth showed a general increase in both probability of crab presence and in density given presence as depth increases. Within this generally positive trend, a maximum (Fig. 4, panels a and e) or a leveling off of the curve (Fig. 4, panels c, d, and h) was frequently observed at approximately 15 to 20 m.

Differences were apparent between the Stage I and Stage II response curves for temperature. Stage I curves (Fig. 5, panels a–f) showed a generally negative relationship between temperature and the probability of crab presence while Stage II curves (Fig. 5, panels g and h) both indicate a positive relationship between temperature and crab density given presence. Substantial interannual differences in December bottom temperatures, however, make it difficult to compare models for which temperature ranges do not overlap.

### Model fit, cross-validation, and mapping

We used the models developed on the training data in a single year to predict crab distribution and abundance for the test data for that year, and for the entire data sets for alternative years. Stage I (presence/absence) models were evaluated using ROC curves to assess model fit (Table 3a) and generality (Table 3b). The percent of correct predictions for models applied to the training data varied from 82 to 93% at  $p_{\text{opt}}$  and from 74 to 85% at  $p_{\text{fair}}$ . The AUC for the training data varied from 0.81 to 0.91. These levels are equivalent to Hosmer & Lemeshow's (2000) excellent rating. AUC values were generally lower for the cross-validation where models developed with data from 1 yr were applied to data from another year; however, all models displayed acceptable discriminatory power ( $\text{AUC} > 0.7$ ) for at least 4 other years. The Stage I models from 1997 and 1998 were the most general with  $\text{AUC} > 0.7$  for all years other than 1995. The 1995 data were well predicted only by Stage I models from 1994 and 1995.

Abundance was highly variable and more difficult to predict than distribution. Two-stage GAMs devel-

Table 3. *Callinectes sapidus*. Evaluation of Stage I (presence/absence) model fits to the training data (a) using receiver operating characteristic (ROC) curves and cross-validation of Stage I models (b). Values in (a) represent the area under the ROC curve (AUC), the critical p-values:  $p_{\text{optimum}}$  ( $p_{\text{opt}}$ ) and  $p_{\text{fair}}$ , and their sensitivity (sens.), specificity (Spec.), and percent correct predictions (% Corr.). Values in (b) represent the AUC where models developed with data from one year (columns) are applied to data from another (rows). Values on the diagonal represent intra-annual cross-validation where models developed using a training data subset are applied to the test data subset for the same year. AUC values >0.7 are shaded

(a)													
Year	AUC	Value	$p_{\text{opt}}$ Sens.	Spec.	% Corr.	Value	Sens.	$p_{\text{fair}}$	Spec.	% Corr.			
1990	0.870	0.470	0.396	0.975	89.4	0.155	0.780	0.785	78.5				
1991	0.839	0.453	0.460	0.927	82.3	0.235	0.752	0.749	75.0				
1992	0.809	0.575	0.284	0.973	84.3	0.180	0.746	0.740	74.1				
1993	0.845	0.580	0.277	0.983	85.9	0.180	0.765	0.767	76.6				
1994	0.862	0.405	0.393	0.951	87.3	0.175	0.793	0.793	79.3				
1995	0.824	0.325	0.046	0.997	92.7	0.085	0.773	0.762	76.3				
1996	0.851	0.470	0.373	0.966	86.4	0.195	0.770	0.770	77.3				
1997	0.837	0.410	0.302	0.974	89.0	0.125	0.752	0.748	74.9				
1998	0.849	0.455	0.400	0.960	87.2	0.150	0.768	0.770	77.0				
1999	0.869	0.380	0.294	0.981	91.9	0.110	0.784	0.786	78.6				
2000	0.859	0.533	0.290	0.972	85.6	0.200	0.777	0.783	78.2				
2001	0.905	0.540	0.302	0.983	91.5	0.110	0.845	0.847	84.7				
2002	0.884	0.400	0.310	0.979	92.1	0.105	0.810	0.819	81.8				
(b)													
Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
1990	0.74	0.69	0.65	0.74	0.73	0.76	0.64	0.73	0.72	0.44	0.57	0.68	0.71
1991	0.71	0.78	0.65	0.79	0.75	0.74	0.78	0.75	0.81	0.65	0.69	0.63	0.70
1992	0.58	0.76	0.77	0.72	0.78	0.74	0.74	0.71	0.79	0.79	0.79	0.72	0.78
1993	0.73	0.78	0.73	0.79	0.74	0.74	0.72	0.72	0.76	0.60	0.67	0.54	0.73
1994	0.58	0.80	0.78	0.77	0.82	0.81	0.82	0.76	0.79	0.78	0.78	0.73	0.75
1995	0.63	0.64	0.62	0.67	0.72	0.70	0.65	0.63	0.66	0.64	0.67	0.50	0.65
1996	0.66	0.69	0.64	0.65	0.67	0.62	0.72	0.70	0.71	0.57	0.62	0.69	0.69
1997	0.67	0.72	0.64	0.70	0.66	0.66	0.68	0.76	0.73	0.59	0.65	0.65	0.67
1998	0.64	0.73	0.69	0.72	0.67	0.71	0.70	0.76	0.77	0.56	0.60	0.57	0.64
1999	0.64	0.80	0.81	0.77	0.82	0.81	0.82	0.81	0.85	0.85	0.86	0.68	0.86
2000	0.73	0.77	0.79	0.74	0.80	0.76	0.74	0.77	0.79	0.83	0.84	0.71	0.79
2001	0.74	0.75	0.72	0.71	0.74	0.57	0.76	0.71	0.77	0.63	0.78	0.80	0.80
2002	0.52	0.71	0.73	0.59	0.70	0.62	0.65	0.74	0.75	0.70	0.76	0.52	0.75

oped using the 6 habitat variables included in this study explained between 13 and 36% (mean  $R^2 = 0.277$ ) of the variability in blue crab winter densities in the training data set (Table 4a). WDS samples were characterized by a large percentage (80 to 90%) of observations containing no mature female blue crabs, as well as a small number of very high-density samples. The 2-stage models showed no evidence of bias and generally predicted realistic densities but underestimated the observed variability. For example, predicted log densities from the 1998 2-stage model showed a similar mean as the survey observations, with the linear regression of observed vs. predicted falling nearly coincident with the one-to-one line, but showed lower variability, i.e. fewer low- or zero-density predictions and a lower range of predicted values (Fig. 6). Because of the relatively short tows (1 min), observed densities show a notable gap between tows with zero catches, and the lowest observed densities.

The mean  $R^2$  for the intra-annual comparison was 0.192. Results for the intra-annual cross-validation showed that there was a significant difference in model performance between test data and training data (paired  $t$ -test,  $p = 0.002$ ). The inter-annual cross-validation displayed substantial variation among years and was significantly less accurate than the intra-annual cross-validation ( $t$ -test,  $p < 0.001$ ).

The cross-validation table (Table 4b) represents the ability of a model developed with data from 1 yr (columns) to predict data from other years (rows), and it displays 2 different but related pieces of information. Examining the patterns within a column evaluates the characteristics of one model. Patterns within a row relate to the characteristics of a particular data set.

The column patterns show that apart from 1990 and 2001 all models yielded above average  $r$  values for at least 4 other years of data. This suggests that the models, though they differ in their individual parameters, do capture some general features of the blue crab

habitat preference. The 1998 model displayed above average cross-validation *r* values for all years except 1990, 1995, and 2002. The other striking feature of the column patterns is that the 1990 and 2001 models yielded below average *r* values for nearly all data sets except test data from the same year.

The row patterns offer further information about inter-annual differences. The year 1990 is well predicted (i.e. above average *r*) only by the model from the same year. The data for 1995 are simply difficult to predict with any model. The 1994 data are well predicted by models from any year other than 1990, 1997, or 2001.

Predictions from the Stage I (Fig. 7a), Stage II (Fig. 7b), and combined (Fig. 7c) models were mapped for the 1998 model since this year exhibited the greatest generality for both Stage I and the combined model. Critical *p*-values used for classifying the Stage I map were  $p_{\text{fair}} = 0.15$  and  $p_{\text{opt}} = 0.455$ . Of the samples in the 1998 training data that contained one or more mature female blue crabs, 77% occurred within

the shaded areas of Fig. 7a, and 40% occurred within the dark shaded area. Much of the mainstem Bay south of the Rappahannock River is shaded, indicating higher probability of blue crab presence. North of the Rappahannock River, shaded areas are generally restricted to the deeper mainstem channels and the channel in Tangier Sound.

Patterns of predicted density given presence (conditional density) shown in Fig. 7b are broadly similar to patterns in probability of occurrence. Higher conditional densities are predicted in the lower Bay mainstem and in deeper channels throughout the Bay. The higher conditional densities predicted in the upper reaches of western shore tributaries and the northernmost region of the mainstem are not found in the Stage I model and are apparently extrapolations of the salinity effect beyond the range of sampled salinities.

The combined model (Fig. 7c) is derived from raster multiplication of Figs. 7a & b and reflects the influence

Table 4. *Callinectes sapidus*. Cross-validation where models developed with data from one year (columns) are applied to data from another (rows). Values in (a) represent the cross-validation *r*-squared. Values on the diagonal (in **bold**) represent intra-annual cross-validation where models developed using a training data subset are applied to the test data subset for the same year. The first row of (a) represents the model fit to the training data. Values in (b) represent the *z*-score, i.e. the number of standard deviations above (shaded) or below the grand mean Fisher (1915) transformed cross-validation correlation coefficient

(a)													
Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Training	0.299	0.315	0.255	0.296	0.270	0.130	0.314	0.241	0.346	0.232	0.305	0.360	0.239
1990	<b>0.099</b>	0.064	0.016	0.069	0.059	0.091	0.014	0.060	0.081	0.001	0.015	0.069	0.013
1991	0.028	<b>0.219</b>	0.048	0.135	0.174	0.071	0.128	0.071	0.195	0.043	0.148	0.009	0.086
1992	0.003	0.257	<b>0.291</b>	0.191	0.265	0.210	0.146	0.099	0.271	0.235	0.273	0.069	0.221
1993	0.129	0.188	0.128	<b>0.238</b>	0.113	0.089	0.113	0.090	0.165	0.022	0.047	0.031	0.125
1994	0.000	0.194	0.200	0.232	<b>0.244</b>	0.237	0.117	0.056	0.194	0.176	0.222	0.041	0.124
1995	0.010	0.032	0.040	0.063	0.095	<b>0.086</b>	0.016	0.008	0.040	0.023	0.048	0.010	0.014
1996	0.043	0.111	0.088	0.077	0.098	0.042	<b>0.168</b>	0.110	0.152	0.017	0.005	0.062	0.081
1997	0.087	0.165	0.095	0.171	0.097	0.055	0.129	<b>0.197</b>	0.172	0.024	0.097	0.054	0.110
1998	0.035	0.105	0.085	0.109	0.079	0.066	0.084	0.099	<b>0.129</b>	0.008	0.049	0.005	0.044
1999	0.000	0.155	0.236	0.157	0.139	0.174	0.108	0.088	0.176	<b>0.196</b>	0.205	0.080	0.184
2000	0.042	0.206	0.222	0.202	0.183	0.191	0.092	0.077	0.243	0.255	<b>0.311</b>	0.069	0.165
2001	0.078	0.184	0.108	0.080	0.121	0.009	0.129	0.078	0.166	0.036	0.054	<b>0.228</b>	0.087
2002	0.001	0.083	0.075	0.016	0.084	0.021	0.019	0.056	0.084	0.062	0.128	0.061	<b>0.089</b>
(b)													
Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
1990	0.04	-0.43	-1.38	-0.36	-0.52	-0.06	-1.44	-0.50	-0.20	-2.10	-1.41	-0.37	-1.46
1991	-1.08	1.35	-0.69	0.48	0.89	-0.33	0.39	-0.33	1.11	-0.79	0.62	-1.61	-0.12
1992	-1.91	1.72	2.04	1.07	1.79	1.26	0.60	0.05	1.85	1.51	1.87	-0.37	1.37
1993	0.40	1.04	0.39	1.53	0.21	-0.08	0.22	-0.07	0.80	-1.22	-0.71	-1.03	0.36
1994	-2.16	1.10	1.16	1.47	1.59	1.53	0.26	-0.57	1.10	0.92	1.38	-0.82	0.35
1995	-1.58	-0.99	-0.84	-0.45	-0.01	-0.13	-1.37	-1.66	-0.84	-1.20	-0.70	-1.58	-1.43
1996	-0.78	0.20	-0.09	-0.25	0.03	-0.80	0.83	0.18	0.67	-1.36	-1.78	-0.47	-0.19
1997	-0.11	0.80	0.00	0.86	0.02	-0.58	0.41	1.13	0.87	-1.18	0.02	-0.59	0.18
1998	-0.93	0.12	-0.13	0.17	-0.22	-0.41	-0.15	0.04	0.40	-1.66	-0.68	-1.80	-0.77
1999	-2.28	0.69	1.52	0.71	0.52	0.89	0.15	-0.10	0.92	1.12	1.21	-0.21	1.00
2000	-0.80	1.22	1.38	1.18	0.99	1.07	-0.04	-0.25	1.59	1.70	2.22	-0.36	0.80
2001	-0.24	0.99	0.15	-0.21	0.32	-1.61	0.40	-0.23	0.81	-0.91	-0.60	1.43	-0.11
2002	-2.07	-0.17	-0.28	-1.39	-0.16	-1.25	-1.31	-0.56	-0.15	-0.46	0.39	-0.48	-0.08

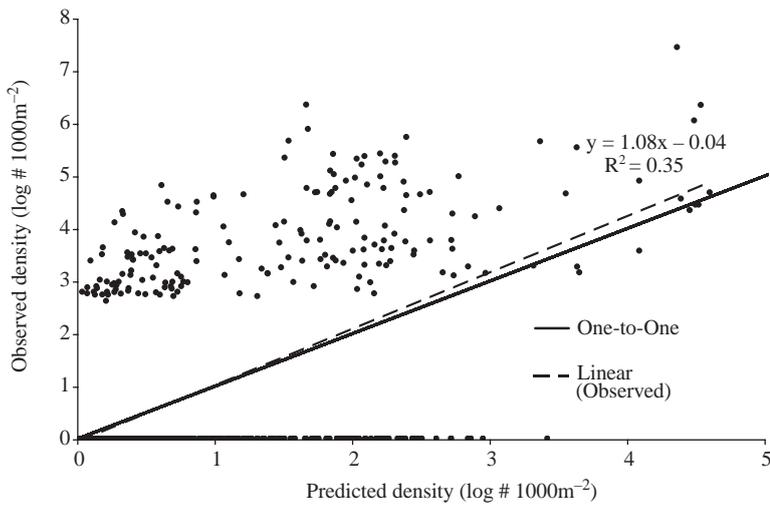


Fig. 6. *Callinectes sapidus*. Cross-validation plot showing observed vs. predicted log density (log # 1000m<sup>-2</sup>) for the 1998 training data, the one-to-one line (solid) and the linear regression line (dashed) fit to all data, including zeros. Note: multiple observations are overlaid along the x-axis near the origin

of both model stages. Highest predicted densities are found in the lower Bay mainstem and deep channels. The high conditional densities predicted in Stage II for the upper reaches of western shore tributaries and the northernmost region of the mainstem are largely nullified by the low probability of crab presence predicted for these areas in Stage I.

**DISCUSSION**

The spatial patterns of winter distribution and abundance of mature female blue crabs in Chesapeake Bay were significantly related to several environmental factors over 13 yr. Depth and distance from the Bay mouth (and their interaction) were the dominant variables for predicting both presence/absence and abundance of mature female blue crab. In-

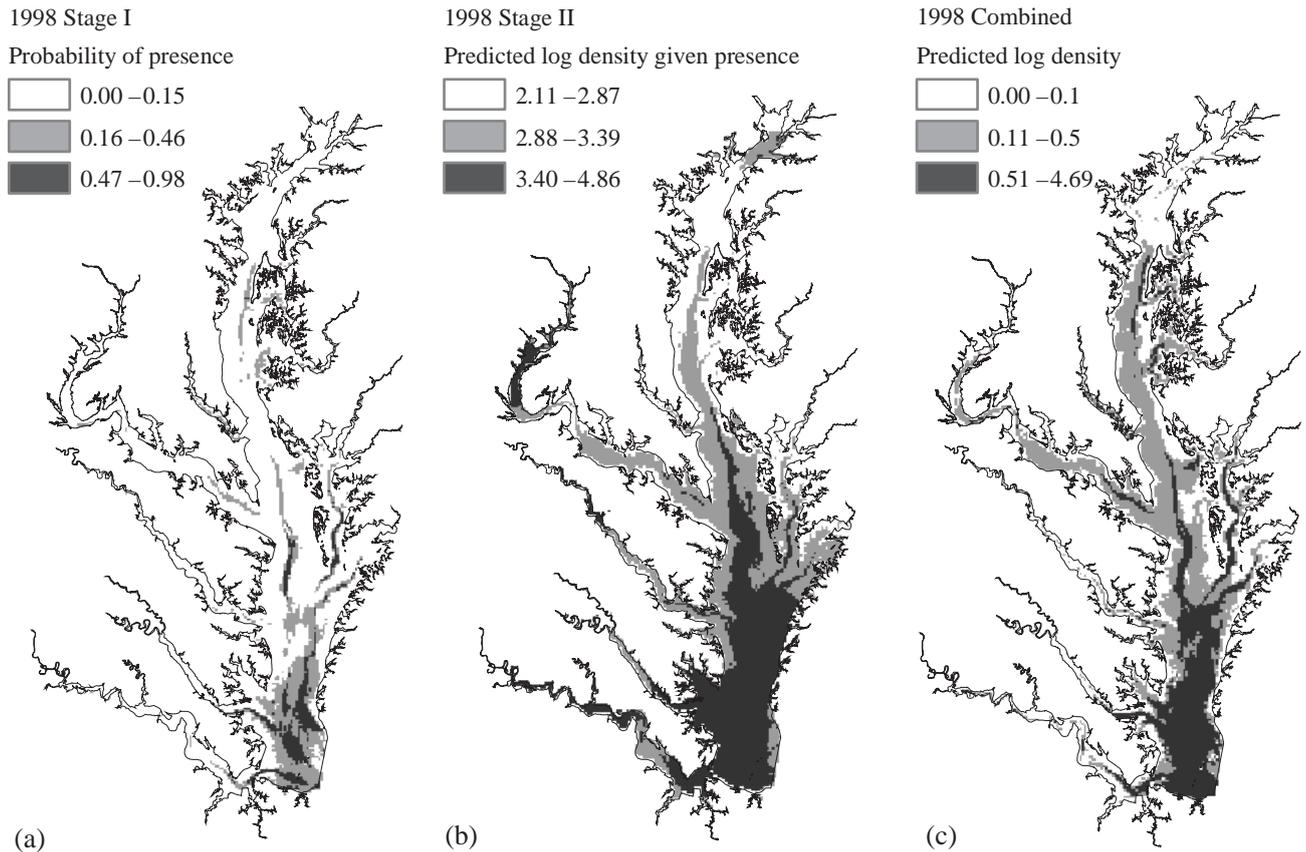


Fig. 7. *Callinectes sapidus*. (a) Predicted probability of mature female blue crab presence, (b) predicted log density (log # 1000m<sup>-2</sup>) given presence, and (c) predicted log density (log # 1000m<sup>-2</sup>) based on the product of (a) and (b) from the 1998 model

creased depth was associated with increased probability of finding crabs and increased crab abundance where they were present. Although greater depths are generally associated with lower Bay waters, depth may also have a direct effect on habitat suitability by providing some protection against rapid temperature changes due to changing air temperature. Bottom waters in shallow parts of Chesapeake Bay typically show greater daily temperature variations than do deep bottom waters (O. Jensen, unpublished analysis of Chesapeake Bay Water Quality Monitoring Program data).

Probability of crab presence and crab abundance generally decreased with distance from Bay mouth, although in some years a maximum was discernible at 25 to 50 km from the Bay mouth. The correlation between distance from the Bay mouth and other environmental variables, salinity in particular, makes it difficult to determine what, if any, direct influence distance from the Bay mouth could have on blue crabs. However, mature female blue crabs both spawn and overwinter in the lower Bay, and it is likely that effective offshore transport of newly hatched crab larvae is dependent on their release location. The role of distance from the Bay mouth in explaining abundance of mature female blue crab may be related to selection for optimum larval transport conditions or a balance between conditions favoring higher survival and those favoring reproductive success.

Salinity and temperature were also frequently found to be significant factors in determining crab distributions, although perhaps not as often as might be expected given the demonstrated effects of salinity and temperature on the bioenergetics (Guerin & Stickle 1992, Brylawski & Miller 2003), growth (Tagatz 1968, Smith 1997), and survival (Tagatz 1969) of blue crabs in the laboratory. Higher salinities were associated with higher probability of blue crab presence and higher density given presence, with a maximum at 25 ppt observed in some years. Females migrating from the upper Chesapeake Bay likely do not spawn until the season after mating (Turner et al. 2003); however, there are potential advantages to overwintering in the lower Bay. Although adult females tolerate a wide range of salinities, they may be less efficient osmoregulators at lower salinity (Tan and Van Engel 1966), and may be less tolerant of extreme temperatures at low salinity (Tagatz 1969). Laboratory studies have demonstrated that blue crab respiration increases at decreasing salinity (Engel & Eggert 1974, Guerin & Stickle 1992), thus overwintering in high salinity waters may allow females to conserve energy. In addition to having lower salinity, upper Bay waters also have greater temperature fluctuations.

Higher temperatures were associated with a lower probability of crab presence, but higher density given

presence. Blue crabs may be expected to have conflicting demands regarding temperature. Mortality rates increase sharply below 5°C (L. Bauer pers. obs.), but respiration and metabolic costs begin to increase rapidly above approximately 15°C (Brylawski & Miller 2003). Still, it is unclear why the direction of the temperature response should vary between Stage I and Stage II models.

Response curves for the 2 remaining variables (not shown) are complex and do not coincide with simple biological explanations. Bottom slope and distance from SAV showed little ability to explain crab distributions or abundance. Even when these variables were determined to be significant, the response curves were highly variable and no support was provided for the hypothesis that lower slope and shorter distance from SAV represent preferred winter habitat. Such year-to-year variability in response curves may indicate that relationships to some habitat parameters are complex and dynamic or may change with changes in population size. It is also likely that some spurious relationships may be found to be significant when fitting 26 separate models.

Although correlation among environmental parameters is likely the norm rather than the exception, such dependencies must be considered when evaluating model selection results. For example, a strong correlation exists between salinity and distance from the Bay mouth. As a result, although both variables were common in the final models, it was relatively unusual for both to be included in the same model. Although efforts were made to make all variables equally likely to enter into the model (e.g. by transforming non-normally distributed variables and rescaling some variables so that all were of the same magnitude), inherent differences in variability and measurement error are still likely to influence model selection. As Håkanson & Peters (1995) have pointed out, if 2 environmental parameters are equally related to a response, it is the parameter with lower variability and lower measurement error that is most likely to be selected by the model. In this case, the static variables (depth and distance from the Bay mouth) have an advantage in that they can be measured with little error and they do not change over time. Even if individual crabs are selecting an overwintering location based on temperature and salinity at the time of burying, depth and distance from the Bay mouth may be more powerful predictors of distributions, despite not being the proximate cue, to the extent that they integrate information about salinity and temperature over the period during which crabs are selecting an overwinter location. Such questions about migration cues cannot be resolved through an empirical habitat modeling approach.

Despite the highly flexible modeling process and the inclusion of interactions among parameters, the full 2-

stage GAMs explained only a fraction (13 to 36% when applied to the training data) of the variability in crab abundance. In addition, the 2 most important variables, depth and distance from the Bay mouth, are spatially static and thus cannot explain inter-annual changes in distribution patterns. Either there exist other important environmental determinants of crab distributions than those explored here or, although habitat affinities clearly exist (as evidenced by the consistent significance of some of the environmental parameters), much of the observed variability in distribution patterns is not the result of habitat selection. If other environmental factors are guiding habitat selection, it is unclear what those factors may be. Although hypoxia is prevalent in deeper Bay waters in the summer, winter dissolved oxygen levels are sufficient for blue crabs in even the deepest Bay waters (Wang et al. 2001). The development of detailed benthic habitat maps would allow for inclusion of sediment type as a predictor variable; however, the evidence for sediment type effects on blue crab distributions is mixed. Schaffner & Diaz (1988) found significantly higher abundance of blue crabs on sediments with intermediate sand content (41–60%) based on a limited number (94) of lower Bay sample sites. In contrast, Sharov et al. (2003) reported that stratification by sediment type produced only marginal gains in precision of baywide blue crab abundance estimates based on a sample size of more than 1200 stations distributed throughout the Bay. We conclude that factors other than habitat, such as the timing of the onset of cold weather and density-dependent habitat selection (discussed below), may also be important in determining blue crab winter distributions.

Although blue crab density was difficult to predict, the broader question of determining whether a given habitat is likely to contain blue crabs or not proved more tractable. Stage I (presence/absence) models showed considerable ability to discriminate between suitable and unsuitable habitat with approximately 75 to 80% correct predictions at  $p_{\text{fair}}$ . The discriminatory power of the Stage I models was also maintained when applied to other years with an average AUC of 0.71. Indeed, the most general Stage I model, the 1998 model, yielded an AUC greater than 0.7 for all but one of the other years indicating that it provides broadly applicable predictions which could be useful for management purposes. Furthermore, the probability map for the 1998 Stage I model confirms observations that mature female blue crab catch per unit effort (CPUE) is higher in deep water (Lipcius et al. 2001) and in the lower Bay, but also predicts relatively high probabilities of occurrence in some of the deeper channels of the middle and upper Bay and Tangier Sound. Mature females are found in WDS samples at these middle and

upper Bay locations, but it is unclear whether these individuals represent crabs that failed to complete their migration to the lower Bay, as suggested by Turner et al. (2003), or if these deep middle and upper Bay channels also represent preferred overwintering habitat. One of the component variables of the 1998 Stage I model is the dynamic variable, temperature. To the extent that temperature patterns vary from year to year, the predictions of the model are also likely to shift, giving the model the ability to adapt its predictions to changing environmental conditions.

The use of a GIS in combination with habitat suitability models has become widespread as a method of visualizing and mapping the results of habitat modeling (Stoner et al. 2001), as a qualitative test of habitat model output (Zheng et al. 2002), and as a tool for measuring variables that were not or could not be measured in the field (Brown et al. 2000, Clark et al. 2003). Variables such as slope, bathymetric variance, and distance from a particular point or habitat type may provide important information about habitat suitability, but cannot be easily measured in the field. In this study, 3 of the GIS-derived variables (through-water distance from the Bay mouth, salinity, and temperature) were found to be important factors for predicting distributions while the other 2 (bottom slope and distance from SAV) were not. The ease with which such GIS-derived variables can be calculated and tested for predictive ability makes this an appealing method for exploring potential habitat suitability factors.

Maps derived from such GIS-based habitat models may be useful for siting marine protected areas and dispersal corridors, which, for blue crab in Chesapeake Bay, are currently based simply on observations of higher concentrations of adult females at greater depths (Lipcius et al. 2001) as well as non-biological concerns such as enforceability. The concentration of suitable overwintering habitat in Virginia (southern) waters, which are subject to a winter dredge fishery, raises concerns over the vulnerability of overwintering mature female blue crabs. The region of highest probability of blue crab presence (Fig. 7a) is concentrated in a relatively small region of the lower Bay. Habitat with a probability of blue crab presence  $> 0.47$  occupies an area of approximately 500 km<sup>2</sup>. This suggests that even a modestly-sized closed area could protect a large proportion of the blue crab's preferred overwintering habitat.

The cross-validation confirmed the ability of GAMs to describe general patterns, but provides a warning against naïve application of models to predict distributions in other years. Although the mean inter-annual cross-validation  $R^2$  value (0.101) was significantly lower than that for the intra-annual comparison (0.192), most models provided above average fits to

several other data sets, and the best model provided above average predictions for 10 out of 13 other years. The failures of model generality were confined primarily to 2 or 3 years. The data for 1990 are a good example. Despite the fact that the 1990 model showed a slightly above average fit to the training data ( $R^2 = 0.299$ ), the 1990 data were poorly predicted in inter-annual cross-validation with  $R^2$  values below average for all comparisons. The intra-annual cross-validation  $R^2$ , however, was approximately average. Similarly, the 2001 model showed the best fit of any model to the training data ( $R^2 = 0.360$ ) and well above average intra-annual cross-validation, yet displayed poor generality with below average inter-annual cross-validation  $R^2$  values for all comparisons. This indicates that although the strength of the response to habitat variables in 1990 and 2001 was normal, the details of that response were different to those observed in most other years. The explanation appears to lie in the unusually early and severe winters of 1989-1990 and 2000-2001, which had the 2 lowest average December temperatures observed over the 13 years of the survey. Thus, hypotheses or management actions developed from habitat models based on 1990 or 2001 data would likely not be applicable to other years. However, there was no reason *a priori* to anticipate this lack of generality from the model fits or intra-annual cross-validations. Accordingly, we caution against the application of habitat models based on a single year of data without adequate inter-annual cross-validation.

Although models were generally transferable, some years consistently defied prediction by models developed from other years' data. The data from 1995, for example, were poorly predicted by all models and had the poorest observed fit to the training data. The relationships between blue crabs and their habitat in 1995 appear to be weak or highly variable as all models fit poorly to data from this year. Despite this high variability in 1995, the mean response to environmental variables in this year appears to have been fairly typical as the 1995 model displayed above average prediction accuracy when applied to 4 other years of data.

Some of the inter-annual variation in the models and their cross-validation performance is likely related to the date of onset of cold weather and the severity of the winter. Blue crabs are restricted in their level of activity by ambient temperatures. Early onset of cold weather is thought to result in an arrested migration that may strand individual blue crabs in sub-optimal habitat. Consequently, the earlier the onset of cold weather, the less likely observed distributions reflect true habitat preference. Prolonged periods of cold weather also appear to increase the amount of winter mortality (Sharov et al. 2003). To the extent that crabs concentrate in areas where overwinter survival is

higher, severe winters may highlight the expression of existing habitat affinities by preferentially removing individuals that stray from optimal habitat. Conversely, severe winters are likely to obscure the consequences of habitat choice based on factors other than survival, for example, spawning success.

Density-dependent changes in habitat use offer an intriguing alternative explanation for inter-annual differences. MacCall's (1990) basin model predicts that at low population density all individuals will concentrate in the preferred habitats. As population density increases, the preferred habitats become full and individuals are forced to seek out alternative sub-optimal sites. Over the 13 years of the winter dredge survey, estimates of baywide mature female abundance show a more than 4-fold variation from a high of 182 million in 1991 to a low of 41 million in 2001. Interannual changes in abundance are significantly correlated with changes in the location of the center of mass of blue crab distributions (Jensen 2004). Even in years with high blue crab abundance the vast majority of samples showed densities below one crab  $m^{-2}$ , suggesting that space is not physically limiting. The potential mechanisms behind this apparent density-dependent habitat selection deserve further study.

Despite inter-annual variation and the existence of non-habitat related influences, a GAM approach offers unique insights into the factors determining winter distributions of mature female blue crabs. Environmental factors were considered not in isolation, but simultaneously and in interaction. The value of the GAM approach is that, from this collection of intercorrelated variables, it was possible to discern general patterns that persisted from year to year and to identify depth and the distance from the Bay mouth as the 2 most important environmental determinants of winter habitat. The details of these relationships and formal hypothesis tests for individual factors are more appropriately the domain of other methods.

GAMs have become widely recognized as an important tool for understanding species distributions (reviewed in Guisan et al. 2002) because they effectively address many of the statistical challenges (e.g. non-linear responses, complex interactions, and counts that are zero-inflated or otherwise problematic in their distribution) associated with field survey data. One of the concerns with using such a flexible approach is that better model fit might come at the expense of generality. Although some applications of GAMs have successfully addressed concerns regarding generality by dividing large data sets into different years and analyzing them separately (e.g. Begg & Marteinsdottir 2002) or including a year term (e.g. Maravelias et al. 2000b) in the model, and one has used inter-annual cross-validation to compare 2 separate model years

(Forney 2000), this is the first systematic test of GAM habitat model generality. Interannual differences in the structure of models that, we developed, together with their performance in cross-validation trials underscore the importance of having more than a single year 'snapshot.' Although most models performed well in cross-validation, a few years were different enough from the general pattern that they resulted in models with little ability to predict distributions in other years. Such aberrant years can provide unique insights (in this case, suggesting the importance of the timing of cold weather) and, with multiple years of data, they can be identified and effectively dealt with. In the absence of sufficient temporal scope to the data, however, habitat suitability models may be misleading.

*Acknowledgements.* We thank G. Davis for providing winter dredge survey data and valuable discussions, S. Wood and E. Russek-Cohen for statistical advice, and O. Richter for accommodating O.P.J. at the Institute of Geocology, Department for Environmental System Analysis, Technical University Braunschweig. M. C. Christman and E. D. Houde provided helpful comments on an earlier version of the manuscript. This work was supported by a grant to O.P.J. from the German Academic Exchange Service (DAAD), and by a fellowship from the University of Maryland Sea Grant Program. L.J.B. was supported by a fellowship from the University of Maryland Sea Grant Program. T.J.M. was funded by a grant from University of Maryland Sea Grant (R/F89 R/F93). This is contribution number 3847 from the University of Maryland Center for Environmental Science Chesapeake Biological Laboratory.

#### LITERATURE CITED

- Barry SC, Welsh AH (2002) Generalized additive modelling and zero inflated count data. *Ecol Model* 157:179–188
- Begg G, Marteinsdottir G (2002) Environmental and stock effects on spatial distribution and abundance of mature cod *Gadus morhua*. *Mar Ecol Prog Ser* 229:245–262
- Bonn A, Schroder B (2001) Habitat models and their transfer for single and multi species groups: a case study of carabids in an alluvial forest. *Ecography* 24:483–496
- Borchers DL, Buckland ST, Priede IG, Ahmadi S (1997) Improving the precision of the daily egg production method using generalized additive models. *Can J Fish Aquat Sci* 54:2727–2742
- Brown SK, Buja KR, Jury SH, Monaco ME (2000) Habitat suitability index models for eight fish and invertebrate species in Casco and Sheepscot Bays, Maine. *North Am J Fish Manage* 20:408–435
- Brylawski BJ, Miller TJ (2003) Bioenergetic modeling of the blue crab (*Callinectes sapidus*) using the fish bioenergetics (3.0) computer program. *Bull Mar Sci* 72:491–504
- Cardinale M, Arrhenius F (2000) The influence of stock structure and environmental conditions on the recruitment process of Baltic cod estimated using a generalized additive model. *Can J Fish Aquat Sci* 57:2402–2409
- Chesapeake Bay Program (1993) Guide to using Chesapeake Bay Program water quality monitoring data. CBP/TRS 78/92. Chesapeake Bay Program, Annapolis, MD
- Clark RD, Morrison W, Christensen JD, Monaco ME, Coyne MS (2003) Modeling the distribution and abundance of spotted seatrout: integration of ecology and GIS technology to support management needs. In: Bortone SA (ed) *Biology of the spotted seatrout*. CRC Press, Boca Raton, FL
- Cressie N (1993) *Statistics for spatial data*, John Wiley & Sons, New York
- Engel DW, Eggert LD (1974) The effect of salinity and sex on the respiration rates of excised gills of the blue crab, *Callinectes sapidus*. *Comp Biochem Physiol* 47A:1005–1011
- Fisher RA (1915) Frequency distributions of the values of the correlation coefficient in samples from an infinitely large population. *Biometrika* 10:507–521
- Forney KA (2000) Environmental models of cetacean abundance: reducing uncertainty in population trends. *Conserv Biol* 14:1271–1286
- Gibson LA, Wilson BA, Cahill DM, Hill J (2004) Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *J Appl Ecol* 41:213–223
- Guerin JL, Stickle WB (1992) Effect of salinity gradients on the tolerance and bioenergetics of juvenile blue crabs (*Callinectes sapidus*) from waters of different environmental salinities. *Mar Biol* 114:391–396
- Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol Model* 157:89–100
- Håkanson L, Peters RH (1995) *Predictive limnology methods for predictive modelling*. SPB Academic Publishing, Amsterdam
- Hastie TJ, Tibshirani RJ (1990) *Generalized additive models*. Chapman & Hall, London
- Hosmer DW, Lemeshow S (2000) *Applied logistic regression*. John Wiley & Sons, New York
- Jensen OJ (2004) *Spatial ecology of blue crab (Callinectes sapidus) in Chesapeake Bay*. MS thesis, University of Maryland, College Park, MD
- Johnson DR (1995) Wind forced surface currents at the entrance to Chesapeake Bay: their effect on blue crab larval dispersion and post-larval recruitment. *Bull Mar Sci* 57:726–738
- Lipcius R, Seitz RD, Goldsborough WJ, Montane M, Stockhausen WT (2001) A deepwater dispersal corridor for adult female blue crabs in Chesapeake Bay. In: Kruse GH and 8 others (eds) *Spatial processes and management of marine populations*. Alaska Sea Grant, Fairbanks, AK, p 643–666
- Lipcius RN, Stockhausen WT (2002) Concurrent decline of the spawning stock, recruitment, larval abundance, and size of the blue crab *Callinectes sapidus* in Chesapeake Bay. *Mar Ecol Prog Ser* 226:45–61
- Lipcius RN, Stockhausen WT, Seitz RD, Geer PJ (2003) Spatial dynamics and value of a marine protected area and corridor for the blue crab spawning stock in Chesapeake Bay. *Bull Mar Sci* 72:453–469
- MacCall AD (1990) *Dynamic geography of marine fish populations*. University of Washington Press, Seattle, WA
- Maravelias C (1999) Habitat selection and clustering of a pelagic fish: effects of topography and bathymetry on species dynamics. *Can J Fish Aquat Sci* 56:437–450
- Maravelias C, Reid D, Swartzman G (2000a) Seabed substrate, water depth and zooplankton as determinants of the prespawning spatial aggregation of North Atlantic herring. *Mar Ecol Prog Ser* 195:249–259
- Maravelias C, Reid D, Swartzman G (2000b) Modeling spatio-temporal effects of environment on Atlantic herring, *Clupea harengus*. *Environ Biol Fish* 58:157–172
- McKenzie MD (1970) Fluctuations in abundance of the blue crab and factors affecting mortalities. South Carolina

- Marine Resources Division Technical Report, p 45
- McMillen-Jackson AL, Bert TM, Steele P (1994) Population genetics of the blue crab *Callinectes sapidus*: modest population structuring in a background of high gene flow. *Mar Biol* 118:53–65
- Miller TJ (2003) Incorporating space into models of the Chesapeake Bay blue crab population. *Bull Mar Sci* 72: 567–588
- Miller TJ, Houde ED (1998) Blue crab target setting. Final report to the living resources subcommittee, Chesapeake Bay Program, [UMCES]CBL TS-177–99. Chesapeake Biological Laboratory, Solomons, MD
- National Oceanic and Atmospheric Administration (1998) Estuarine bathymetric digital elevation model for Chesapeake Bay. Silver Spring, MD; <http://oceanservice.noaa.gov/mapfinder/products/bathy/fgdcbathy.pdf>
- Orth RJ, Van Montfrans J, Lipcius R, Metcalf KS (1996) Utilization of seagrass habitat by the blue crab. In: Kuo J, Phillips RC, Walker DI, Kirkman H (eds) *Seagrass biology: proceedings of an international workshop*, Western Australia. Western Australian Museum, Perth, p 213–214
- Orth RJ, Wilcox DJ, Nagey LS, Tillman AL, Whiting JR (2001) Distribution of submerged aquatic vegetation in the Chesapeake Bay and coastal bays. Chesapeake Bay Program, Annapolis, MD
- Piet G (2002) Using external information and GAMs to improve catch-at-age indices for North Sea plaice and sole. *ICES J Mar Sci* 59:624–632
- Rand P (2002) Modeling feeding and growth in Gulf of Alaska sockeye salmon: implications for high-seas distribution and migration. *Mar Ecol Prog Ser* 234:265–280
- Rugolo L, Knotts K, Lange A, Crecco V (1998) Stock assessment of Chesapeake Bay blue crab (*Callinectes sapidus* Rathbun). *J Shellfish Res* 17:493–517
- Schaffner LC, Diaz RJ (1988) Distribution and abundance of overwintering blue crabs, *Callinectes sapidus*, in the lower Chesapeake Bay. *Estuaries* 11:68–72
- Scholten M, Wirtz C, Fladung E, Thiel R (2003) The modular habitat model (MHM) for the ide, *Leuciscus idus* (L.): a new method to predict the suitability of inshore habitats for fish. *J Appl Ichthyol* 19:315–329
- Sharov A, Davis G, Davis B, Lipcius R, Montane M (2003) Estimation of abundance and exploitation rate of blue crab (*Callinectes sapidus*) in Chesapeake Bay. *Bull Mar Sci* 72: 543–565
- Smith SG (1997) Models of crustacean growth dynamics. PhD dissertation, University of Maryland, College Park, MD
- Stoner A, Manderson J, Pessutti J (2001) Spatially explicit analysis of estuarine habitat for juvenile winter flounder: combining generalized additive models and geographic information systems. *Mar Ecol Prog Ser* 213:253–271
- Swartzman G, Huang C, Kaluzny S (1992) Spatial analysis of Bering Sea groundfish survey data using generalized additive models. *Can J Fish Aquat Sci* 49:1366–1378
- Swartzman G, Silverman E, Williamson N (1995) Relating trends in walleye pollock (*Theragra chalcogramma*) abundance in the Bering Sea to environmental factors. *Can J Fish Aquat Sci* 52:369–380
- Tagatz ME (1968) Biology of the blue crab, *Callinectes sapidus* Rathbun, in the St. Johns River, Florida. *Fish Bull* 67:17–33
- Tagatz ME (1969) Some relations of temperature acclimation and salinity to thermal tolerance of the blue crab, *Callinectes sapidus*. *Trans Am Fish Soc* 4:713–716
- Tan EC, Van Engel WA (1966) Osmoregulation in the adult blue crab, *Callinectes sapidus* Rathbun. *Chesapeake Sci* 7:30–35
- Turner HV, Wolcott DL, Wolcott TG, Hines AH. (2003) Post-mating behavior, intramolt growth, and onset of migration to Chesapeake Bay spawning grounds by adult female blue crabs, *Callinectes sapidus* Rathbun. *J Exp Mar Biol Ecol* 295:107–130
- Vølstad J, Sharov A, Davis G, Davis B (2000) A method for estimating dredge catching efficiency for blue crabs, *Callinectes sapidus*, in Chesapeake Bay. *Fish Bull* 98:410–420
- Wang P, Batiuk R, Linker L, Shenk G (2001) Assessment of best management practices for improvement of dissolved oxygen in Chesapeake Bay estuary. *Water Sci Tech* 44(7): 173–180
- Wood SN (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties. *J R Stat Soc Br* 62: 413–428
- Wood SN, Augustin NH (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecol Model* 157:157–177
- Zhang CI, Ault JS, Endo S (1993) Estimation of dredge sampling efficiency for blue crabs in Chesapeake Bay. *Bull Korean Fish Soc* 26:369–379
- Zheng X, Pierce GJ, Reid DG, Jolliffe IT (2002) Does the North Atlantic current affect spatial distribution of whiting? Testing environmental hypotheses using statistical and GIS techniques. *ICES J Mar Sci* 59:239–253

Editorial responsibility: Howard Browman (Associate Editor-in-Chief), Storebø, Norway

Submitted: March 4, 2004; Accepted: March 13, 2005  
Proofs received from author(s): August 8, 2005