Using Delta-Generalized Additive Models to Predict Spatial Distributions and Population Abundance of Juvenile Pink Shrimp in Tampa Bay, Florida

Author: Rubec, Peter J.

Source: Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science, 8(8) : 232-243

Published By: American Fisheries Society

URL: https://doi.org/10.1080/19425120.2015.1084408

Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science 8:232–243, 2016
Published with license by the American Fisheries Society
ISSN: 1942-5120 online
DOI: 10.1080/19425120.2015.1084408

SPECIAL SECTION: SPATIAL ANALYSIS, MAPPING, AND MANAGEMENT OF MARINE FISHERIES

Using Delta-Generalized Additive Models to Predict Spatial Distributions and Population Abundance of Juvenile Pink Shrimp in Tampa Bay, Florida

Peter J. Rubec*
Florida Fish and Wildlife Conservation Commission, 100 8th Avenue Southeast, St. Petersburg, Florida 33701, USA

Richard Kiltie and Erin Leone
Florida Fish and Wildlife Conservation Commission, 1105 Southwest Williston Road, Gainesville, Florida 32601, USA

Richard O. Flamm, Luke McEachron, and Christi Santi
Florida Fish and Wildlife Conservation Commission, 100 8th Avenue Southeast, St. Petersburg, Florida 33701, USA

Abstract

In this article, we present an approach based on generalized additive models (GAMs) to predict species' distributions and abundance in Florida estuaries with habitat suitability modeling. Environmental data gathered by fisheries-independent monitoring in Tampa Bay from 1998 to 2008 were interpolated to create seasonal habitat maps for temperature, salinity, and dissolved oxygen and annual maps for depth and bottom type. We used delta-GAM models assuming either zero-adjusted gamma or beta-inflated-at-zero distributions to predict catch per unit effort (CPUE) from five habitat variables plus gear type for each estuarine species by life stage and season. Bottom type and gear type were treated as categorical predictors with reference parameterization. Three spline-fitting procedures (the penalized B-spline, cubic smoothing spline, and restricted cubic spline) were applied to the continuous predictors. Two additive, linear submodels on the log and logistic scales were used to predict CPUEs >0 and CPUEs = 0, respectively, across environmental gradients. The best overall model among those estimated was identified based on the lowest Akaike information criterion. A stepwise routine was used to omit continuous predictors that had little predictive power. The model developed was then applied to interpolated habitat data to predict CPUEs across the estuary using GIS. The statistical models, coupled with the use of GIS, provide a method for predicting spatial distributions and population numbers of estuarine species' life stages. An example is presented for juvenile pink shrimp *Farfantepenaeus duorarum* during the summer in Tampa Bay, Florida.

Subject editor: Donald Noakes, Vancouver Island University, Nanaimo, British Columbia

*Corresponding author: peter.rubec@myfwc.com
Received February 10, 2015; accepted August 11, 2015

Many studies have been published which predict the spatial distributions of aquatic species (Guisan and Zimmerman 2000; Elith and Leathwick 2009; Planque et al. 2011; Le Pape et al. 2014). Most of these studies used presence or presence/absence modeling. Fewer studies have modeled and mapped the abundance of aquatic species. Our approach used habitat suitability models (HSMs) with georeferenced catch, effort, and environmental data derived from fisheries-independent monitoring (FIM) to predict the spatial distributions and abundance of estuarine species.

A large proportion of zero values is a common characteristic of environmental and ecological studies (Fletcher et al. 2005; Martin et al. 2005). To account for a high proportion of zero catches in fisheries data sets, scientists have developed models that split the data into two components: probability of zero occurrence and positive CPUE (Pennington 1983; Stefánsson 1996; Fletcher et al. 2005). These have been variously labeled conditional, hurdle, two-step, and delta models (Maunder and Punt 2004). In some studies, positive CPUEs were shown to conform to a Gaussian distribution after being log transformed (Pennington 1983; Le Pape et al. 2003). Hence, they have both a binomial component and a lognormal component and were originally called delta models. Stefánsson (1996) proposed that the delta distribution would be better named the delta-lognormal distribution, as it also is feasible to use a delta-gamma distribution.

With the delta-type modeling approach, the positive values are fitted by a generalized linear model (GLM) or a generalized additive model (GAM), and the probabilities of observing zero values are fitted by a GLM or a GAM for a binomial distribution (Stefánsson 1996; Ye et al. 2001). Although combining two submodels complicates model interpretation, delta-lognormal GLM models have been widely used to estimate bycatch and interannual indices of abundance (Maunder and Punt 2004).

Myers and Pepin (1990) provided evidence that the lognormal distribution does not always provide the best fit to abundance data having an excess of zeros. Some positive skewness remains after log transformation because the mode is often the smallest data category available, implying that the continuous lognormal distribution is not an adequate approximation of what are discrete data. Research indicates that lognormal models are very sensitive to deviations from lognormality (Syrjala 2000). The delta-lognormal estimator of the mean was shown to be positively biased. Firth (1988) found that the gamma distribution is more efficient than the lognormal one under reciprocal misspecification.

Delta-lognormal GLMs are suitable for analyzing complex data structures such as Gaussian-distributed dependent variables. Delta-gamma GAMs, however, have been reported to perform better in situations in which relationships are nonlinear (Li et al. 2011a, 2011b). Consequently, we developed delta-gamma GAMs and delta-beta GAMs during the present study.

The goal of the study was to develop statistically sound methods to model and map the spatial distributions and relative abundances of estuarine species of fish and shrimp. Analyses were conducted on 11 estuarine species with 22 life stages for four seasons to produce 87 predicted HSM maps. An example illustrating the analytical methods applied is presented for juvenile pink shrimp *Farfantepenaeus duorarum* during summer in Tampa Bay, Florida.

METHODS

Tampa Bay is the largest estuary in Florida (Lewis and Estevez 1988; TBNEP 1996). The bay spans 103,627 ha of open water and receives drainage from a 569,950-ha watershed. The estuary extends 56 km from north to south and is about 16.1 km wide near the mouth. The maximum water depth is 17.4 m, and the mean depth is 3.4 m. Benthic habitats in Tampa Bay include seagrass, mangrove, salt marsh, bare bottom, hard bottom, and oyster reef (Wolfe and Drew 1990). Small areas with hard bottom and oyster reefs have not been accurately mapped. More than 80% of the bay has sand bottom. The area analyzed for the present study in the bay and tributary rivers encompassed 102,000 ha.

The FIM sampling program, which is associated with the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute (FWRI), has conducted intensive sampling in Tampa Bay since 1989 (Matheson et al. 2005; FWRI 2013). Sampling is also conducted up tributaries to the bay, including the Alafia, Little Manatee, Manatee, and Braden rivers. Monthly stratified-random sampling by the FWRI FIM program has been conducted using a 21.3-m offshore circular bag seine (gear code 20), 21.3-m boat bag seine (23), 183-m haul seine (160), and 6.1-m otter trawl (300). The catches for each of the four gear types are characterized by a few very large catches and a large proportion of zero catches for various species (55–98%).

Georeferenced catch, effort, and environmental data collected by FIM from 1998 to 2008 were used with the present study. The number of individuals of each species caught per gear set within standard length ranges by species life stage (early juvenile, juvenile, and adult) were summed and then transformed into CPUEs. The CPUE data were linked to month, year, latitude, longitude, gear type, and environmental variables.

Data for temperature, salinity, and dissolved oxygen were grouped and interpolated using ordinary kriging with the Spatial Analyst extension in the ArcMap 10.1 GIS to produce seasonal grids. Bathymetry data were obtained from the National Geophysical Data Center of the National Ocean Service and interpolated using inverse distance weighting to create an annual bathymetry grid. A bottom type grid was a created by combining a polygon for submerged aquatic vegetation (SAV; derived from aerial photography by the Southwest Florida Water Management District) with kriged

sand and mud data points (obtained from 2,000 grab samples collected by the Hillsborough County Environmental Protection Commission).

The grids extended up the Manatee and Little Manatee rivers to account for increased sampling up these rivers from 1998 to 2008. Each habitat grid consisted of about 4 million 15-m × 15-m cells. Environmental data points from the center of the cells for each grid were exported from the GIS.

We developed models with CPUE data and associated environmental data using zero-adjusted gamma (ZAGA) and beta-inflated-at-zero (BEINF0) mixed distributions available with the gamlss package (Rigby and Stasinopoulos 2005; Stasinopoulos and Rigby 2007; Stasinopoulos et al. 2014) in R (R Core Team 2014). Zero-adjusted mixed distributions were used because beta and gamma distributions as parameterized in gamlss do not include zeros (Rigby et al. 2013). Before any models were estimated, all of the observed CPUEs (including positive and zero CPUEs) were divided by the maximum CPUE in the sample data plus a small constant (0.01), so that the modeled CPUEs were ≥ 0 and <1.0. Predicted CPUEs were later converted back to the original scale of the data.

The ZAGA model is essentially a delta-gamma GAM, since splines are fit separately to positive CPUE (mu) data and the probability of zero occurrence (nu) data within a season across environmental gradients. Similarly, BEINF0 can be considered a delta-beta GAM (Ospina and Ferrari 2010). Potentially, there are four continuous predictors (temperature, salinity, dissolved oxygen, and bathymetry) and two categorical predictors (bottom type and gear type). The main advantage of the gamlss models is that two multiple regressions (on log and logistic "link" scales for mu and nu, respectively) are created for either the ZAGA distribution or the BEINF0 distribution within the same program.

Categorical gear effects were modeled by first choosing a reference gear type to be represented by the intercept, then determining coefficients for the effects of each nonreference gear type relative to the reference type. As the reference, we generally used the gear type with the greatest mean CPUE in the modeled data. Parameter estimates for the effects of the nonreference gear types correspond to gear correction (GC) factors. Bottom type was treated in a similar manner.

Initially, the penalized B-spline (pbs) transformation was used to choose the best-fitting degrees of freedom (Eilers and Marx 1996) for each continuous predictor in the mu and nu sides of either the BEINF0 or ZAGA models. Then, restricted cubic splines (rcs) and cubic smoothing splines (css) were computed based on transformations of each continuous predictor with degrees of freedom equal to those identified by the pbs transformation. Because the rcs and css functions lack the ability to optimize their degrees of complexity for individual covariates while the pbs transformation has that ability, the latter was applied initially. But once degrees of freedom were chosen for each covariate by the pbs function, either the css or the rcs function was selected because the latter two functions have the potential to perform analyses more rapidly. The relative goodness of fit of the six models resulting from combining the two distributions and three spline transformations was assessed using the Akaike information criterion (AIC; Akaike 1973).

The absolute model fit was evaluated by detrended transformed Owen's plots (Owen 1995) of the normalized quantile residuals (Dunn and Smyth 1996; A. Djennad, R. Rigby, M. Stasinopoulos, and V. Voudouris, London Metropolitan University, unpublished data). Owen's plots are more helpful than traditional quantile–quantile residual plots (routinely used to evaluate normality of a distribution, especially residuals) because they include confidence intervals (CIs) that indicate acceptable normal residual fits if the fitted line for each Owen's plot is approximately horizontal. When the fitted lines in the Owen's plots were bent considerably or exceeded the upper and lower CI bounds, we reran the model after either deleting a gear type or changing the df. A new set of Owen's plots were then generated. If they indicated an acceptable horizontal fit, that model was selected for subsequent usage.

After a best model (either ZAGA or BEINF0) was identified using the full set of available predictors, stepwise model reduction was applied using AIC as the selection criterion (Venables and Ripley 2002). The predictors retained on the mu side of the final model often differed from the predictors retained on the nu side. The AIC is known to identify more-complex final models than other criteria, such as the Bayesian information criterion, when it is used as the criterion for stepwise reduction (Venables and Ripley 2002; Stasinopoulos et al. 2008).

The degrees of freedom in these models represent the degree of smoothing of the splines (Stasinopoulos et al. 2008). With the initial full model output, df = 2 represents a straight line on the transformed (link) scale, while df > 2 represents a curvilinear relationship on the transformed scale. Since the goal was to create fitted curvilinear suitability functions (usually with a convexity representing an intermediate habitat preference) across each environmental gradient, spline transformations with df = 2 resulting from the initial pbs run were changed to df = 3. To smooth the splines, the fitted pbs transformations with df > 5 were changed to either df = 4 or df = 5. These changes usually improved the fitted model, since the AIC for the final model was less than that for the initial model.

Spline transformations were fitted to gear-corrected CPUEs within each model. Bootstrapping (500×) was used to compute CIs for back-transformed fits across environmental gradients for (1) positive GC-CPUEs (mu), (2) the probability of zero occurrence (nu), and (3) predicted mean GC-CPUEs across each environmental gradient. Mean CPUE point estimates were predicted as 1 minus the probability of encountering a zero (from the nu side) times the predicted positive CPUE (from the mu side).

After the best delta-type GAM was determined using FIM data (about 2,200 samples from 1998 to 2008), the model was

applied to the environmental data derived from the habitat grids to predict CPUEs for the reference gear. The predicted GC-CPUE data were then imported into the GIS. The ArcMap Spatial Analyst extension was used to create a predicted GC-CPUE grid across Tampa Bay for each species life stage within the selected season. An HSM grid was then produced by partitioning the predicted GC-CPUE grid into four zones having approximately the same number of cells (equal areas). This second grid, with zones categorized 1–4 by means of different colors, constitutes the predicted HSM map. By multiplying the mean number of animals per square meter by the zonal area, population numbers were estimated for each species life stage by season.

Two verification tests were applied. The first test was used to determine whether independent FIM data collected from 1989 to 1997 agreed with the predicted HSM map. The second verification test, using FIM data collected from 1998 to 2008, was conducted to determine whether the observed data that went into the model agreed with the predicted HSM map.

Gear-corrected CPUEs were computed in the model from the observed FIM data for both time periods. Then the observed GC-CPUE data points for these time periods were separately overlaid onto the four zones in the HSM grid. Increasing mean GC-CPUEs across HSM zones (low, moderate, high, and optimum) indicate spatial agreement between the observed mean GC-CPUEs and the predicted mean GC-CPUEs represented by the HSM map.

The program in R allows the computation of CIs (1) about the Owen's plots, (2) about the fitted splines, (3) at the FIM GC-CPUE sampling points, and (4) across the estuary using bootstrapping.

RESULTS

Panel A of Figure 1 shows an Owen's plot associated with the css BEINF0 model for juvenile pink shrimp created using four gear types: a 21.3-m offshore circular seine (20), a 21.3-m boat-set seine (23), a 183-m haul seine (160), and a 6.1-m otter trawl (300). The predicted line and the upper and lower CI bounds deviate markedly from the horizontal in the upper quantiles where there were fewer observations. Panel (B) shows an Owen's plot associated with a BEINF0 model for the same species life stage with gear type 23 removed but using gear types 20, 160, and 300. The predicted line and the upper and lower CI bounds show less deviation from the horizontal. After making changes to the degrees of freedom for the mu and nu sides of the input and using gear types 20, 160, and 300, the model switched from BEINF0 to ZAGA (Figure 1C). The adjusted degrees of freedom improved the fit of the Owen's plot, resulting in less deviation from the horizontal and a lower AIC value (which dropped from −1,164.82 to −1,423.68).

Table 1 presents parameter estimates, standard errors, *t*-values, and significance levels for the linear effects of the covariates from the reduced css model on the link (log or logistic) scales. Most of the factors in the table are significant on both the mu and nu sides of the model. The statistical output indicates that the significant environmental factors influencing the density of juvenile pink shrimp during the summer on the mu side are salinity, depth, bottom mud, and dissolved oxygen. On the nu side, the significant factors are temperature, depth, salinity, bottom SAV, and dissolved oxygen.

The fitted spline plots give more complete representations of the covariate effects. Figures 2 and 3 illustrate back-transformed fits for positive GC-CPUEs and the probability of zero occurrence. The fits for the GC-CPUEs represent the habitat affinities across the different environmental gradients (Figure 4). These relationships indicate affinities by juvenile pink shrimp in Tampa Bay for moderate salinities (10–20 ‰), higher temperatures (30–32°C), moderate dissolved oxygen levels (4.5–5.5 mg/L), a shallow depth range (0–2 m), and submerged aquatic vegetation (SAV).

The HSM map for juvenile pink shrimp in summer (Figure 5) indicates that they were most abundant in shallow water with submerged aquatic vegetation at moderate salinities. Predicted mean GC-CPUEs, the number of 15-m × 15-m cells by HSM zone, zonal areas (m^2), and estimated population numbers by HSM zone are presented in Table 2. The total number of juvenile pink shrimp estimated to be present in Tampa Bay in summer is 255,496,040. The estimates are long-term averages for the summer derived from FIM sampling from 1998 through 2008.

Verification tests for two time periods are presented in Table 3. The mean GC-CPUEs for 1989–1997 were derived from FIM data collected with gears 20 and 300. The mean GC-CPUEs for 1998–2008 were derived using data collected with gears 20, 160, and 300. In both cases, the observed mean GC-CPUEs increase across the HSM zones. One-way ANOVAs indicate that the differences between the means are highly significant for each time period

DISCUSSION

The present study developed CPUE-based HSM maps for estuarine species in Tampa Bay using either delta-gamma or delta-beta GAMs. While we found studies that used a delta-lognormal or delta-gamma GLM to predict the spatial distributions of marine species, we did not find any fishery studies that used a delta-gamma or delta-beta GAM to create such maps.

Back-transformed splines show habitat affinities for juvenile pink shrimp in Tampa Bay (Figure 4) that generally agree with the scientific literature. Williams (1955) found that the maximum densities of juveniles in estuaries in North Carolina

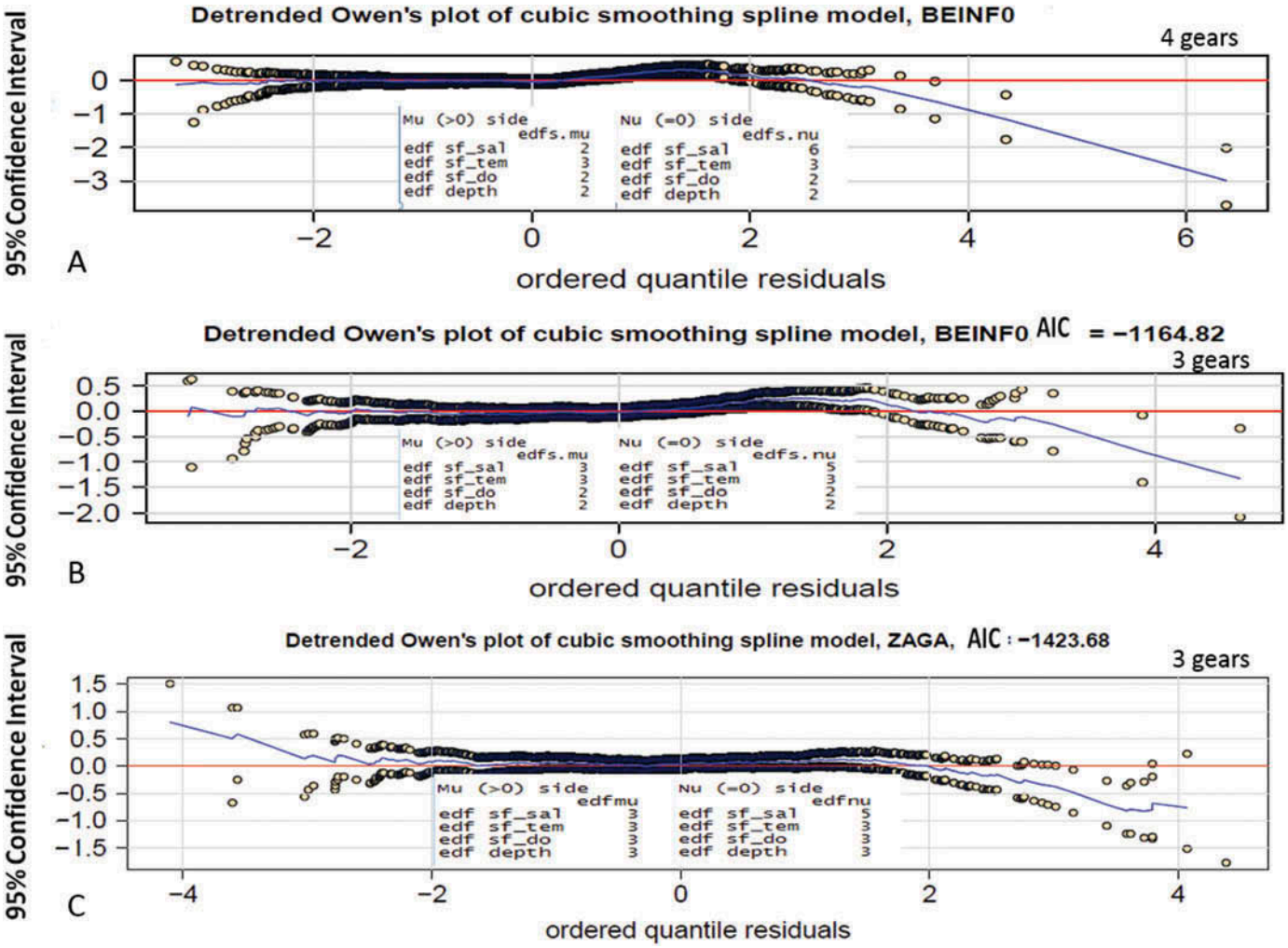

FIGURE 1. Owen's plots for juvenile pink shrimp in summer derived using the cubic smoothing spline option. Panel **(A)** shows the Owen's plot derived using gears 20, 23, 160, and 300; panel **(B)** shows the Owen's plot and initial df after deleting gear 23; and panel **(C)** shows the Owen's plot for gears 20, 160, and 300 after the df were changed. The Owen's plot has normalized the residuals if the predicted line (blue) is close to being horizontal (red line). Open circles represent upper and lower CI bounds for the predicted line. Abbreviations are as follows: edf (effective degrees of freedom) = df, sf_sal = surface salinity, sf_tem = surface temperature, sf_do = surface dissolved oxygen.

occurred at 23–28°C. Pink shrimp exhibit different degrees of preference for salinity at different stages of their life cycle. Hildebrand (1955) reported that juveniles exhibited a preference for salinities of 20‰ or more. As they grow, they move into deeper, more saline water (Williams 1955). The pink shrimp simulation model of Browder et al. (2002) predicts they will be less abundant at salinities <20‰. However, Browder and Robblee (2009) noted a high abundance of juveniles at salinities <12‰ in Whitewater Bay, situated in southwest Florida, which did not agree with models for their growth and survival. Juvenile shrimp are found in estuarine areas with seagrass, and it has been suggested that the distribution of seagrass influences their geographic distribution (Costello and Allen 1970).

Booth (1998, 2000) noted that although there is a strong spatial component in fisheries-dependent and fisheries-independent data, spatial components are often ignored in stock assessments. Nishida and Booth (2001) discuss georeferenced fish resource assessment and the need for space-based fisheries management. Stock assessors should incorporate the spatial component of these data into their stock assessment framework.

Government survey vessels are often used to collect georeferenced catch, effort, and environmental data. Fisheries-independent data have been used to estimate fish population abundance using the stratified mean method (Berg et al. 2014). This is usually done by expanding arithmetic mean CPUEs for the area swept by a trawl to

TABLE 1. Parameter estimates obtained for juvenile pink shrimp in summer. The test statistics are for the linear parametric parts of the covariate effects. The reference bottom type was sand, and the reference gear type was 20; DO = dissolved oxygen and SAV = submerged aquatic vegetation. All spline treatments were cubic smoothing splines with 3 df except for salinity (nu), for which there were 5 df.

Covariate	Estimate	SE	*t*-value	*P*
Mu parameter coefficients				
Intercept	−1.2410	3.5020	−0.3545	0.723
Spline (salinity)	−0.0654	0.0149	−4.3950	<0.0001
Spline (temperature)	0.0651	0.1197	0.5435	0.5868
Spline (DO)	−0.3410	0.1337	−2.5500	0.0109
Spline (depth)	−0.4094	0.0706	−5.7990	<0.0001
Bottom mud	0.6378	0.2897	−2.2020	0.0278
Bottom SAV	0.2682	0.1755	1.5280	0.1267
Gear type 160	−4.3220	0.4969	−8.6990	<0.0001
Gear type 300	−1.5330	0.2308	−6.6440	<0.0001
Nu parameter coefficients				
Intercept	19.5800	3.3410	5.8590	<0.0001
Spline (salinity)	0.0366	0.0121	3.0330	0.0024
Spline (temperature)	−0.7208	0.1129	−6.3820	<0.0001
Spline (DO)	0.2995	0.1379	2.1720	0.0300
Spline (depth)	0.1951	0.0630	3.1000	0.0020
Bottom mud	−0.3534	0.2568	−1.3760	0.1689
Bottom SAV	−0.4597	0.1528	−3.0080	0.0027
Gear type 160	3.6230	0.3445	10.5100	<0.0001
Gear type 300	−0.5117	0.1841	−2.7790	0.0055

the area of each depth stratum and then summing across all strata for the total area sampled. Most stock assessments, however, are conducted primarily using fishery-dependent monitoring data gathered onshore after the fish have been landed. Often the only information used for stock assessments derived from the FIM surveys are the interannual CPUE time series used to tune population models.

The usual procedure for the creation of interannual CPUE time series with delta-type GLM models is to use the year effects as indices of abundance (Maunder and Punt 2004). Berg et al. (2014) stated that this is not possible for delta-type GAM models and that instead one must integrate the fitted abundance surface to obtain the CPUE index. However, he qualified this by stating that if one divides the survey area into subareas and at least one trawl haul is taken in each area, the number of grid cells corresponds approximately to the number of hauls in the actual surveys and using more grid points does not significantly change the resulting index of abundance.

Only a few studies have constructed delta-type models using georeferenced catch, effort, and environmental data for estimating interannual CPUE time series (Li et al. 2011b). Muller et al. (2015) has used catch, effort, and environmental data with delta-gamma GLM models to estimate interannual CPUE time series for Common Snook *Centropomus undecimalis* from FWRI FIM data collected in Tampa Bay and other Florida estuaries.

Drexler and Ainesworth (2013) and Grüss et al. (2014) developed GAMs to support the creation of CPUE-based distribution maps of adult pink shrimp using SEAMAP trawl survey data and oceanographic data from the northern Gulf of Mexico. Predicted CPUE maps were created that also included coastal areas in the southern Gulf of Mexico and off Cuba. The studies by Jensen et al. (2005), Li et al. (2011a, 2011b), and Grüss et al. (2014) had access to georeferenced catch, effort, and environmental data, but they did not attempt to estimate population abundance directly from FIM data using delta-type models.

The predicted GC-CPUE grid for juvenile pink shrimp in Tampa Bay created in the present study provides the information needed to estimate their population abundance. This can be accomplished in a manner similar to the stratified mean method by multiplying the mean GC-CPUEs within HSM zones by the area of each zone and summing the estimated population numbers across the four zones. Alternatively, CPUEs can be summed across all cells in the entire estuary or in a delineated portion to estimate relative population numbers (or biomass). We believe that delta-type models used in conjunction with GIS are a reliable and statistically defensible means of spatially deriving population estimates for estuarine and marine species from georeferenced sampling data.

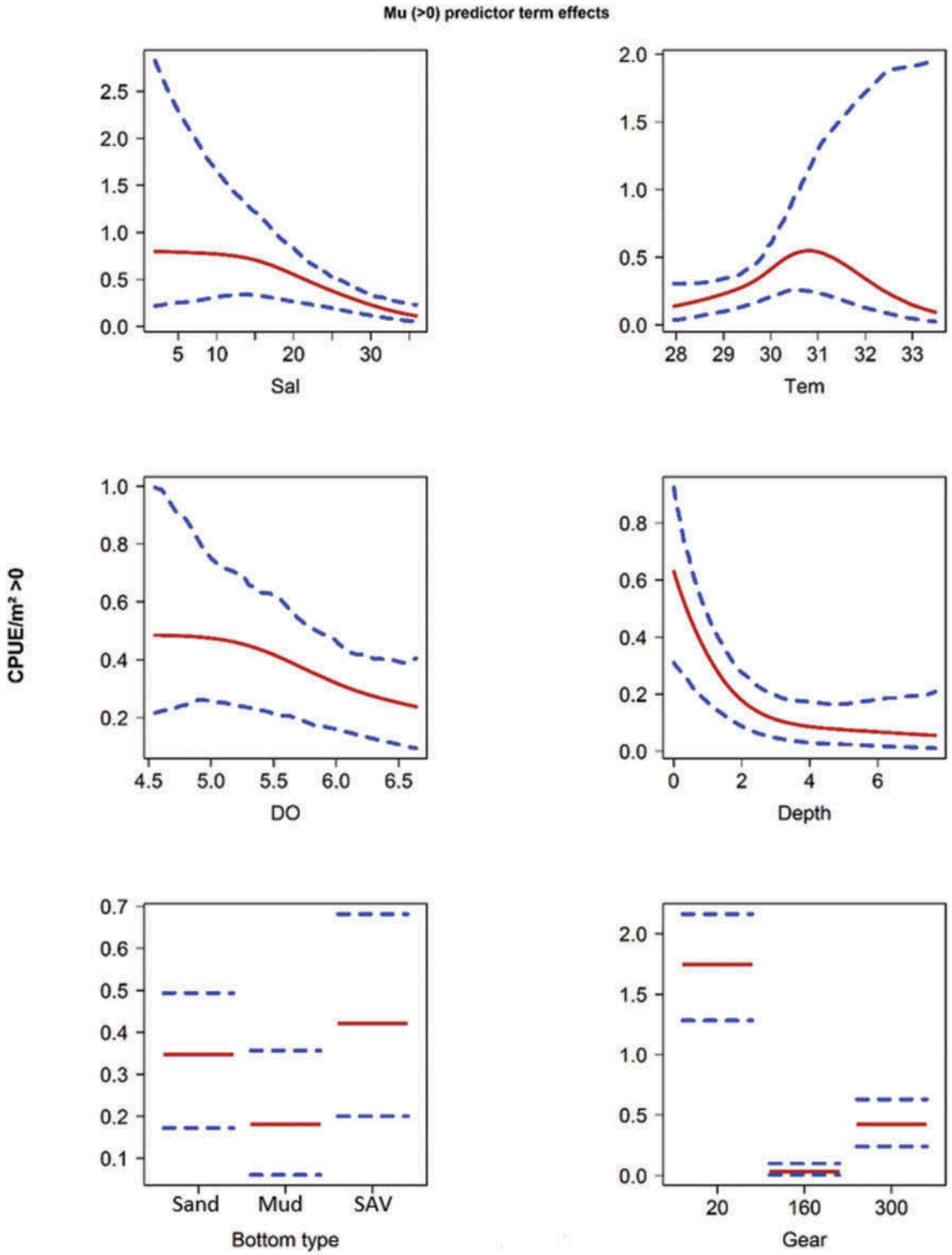

FIGURE 2. Fitted back-transformed CPUE > 0 spline function plots on the mu side of the reduced model for juvenile pink shrimp in summer. Red lines are fitted positive CPUEs for the mu side. Dashed blue lines indicate upper and lower CI bounds.

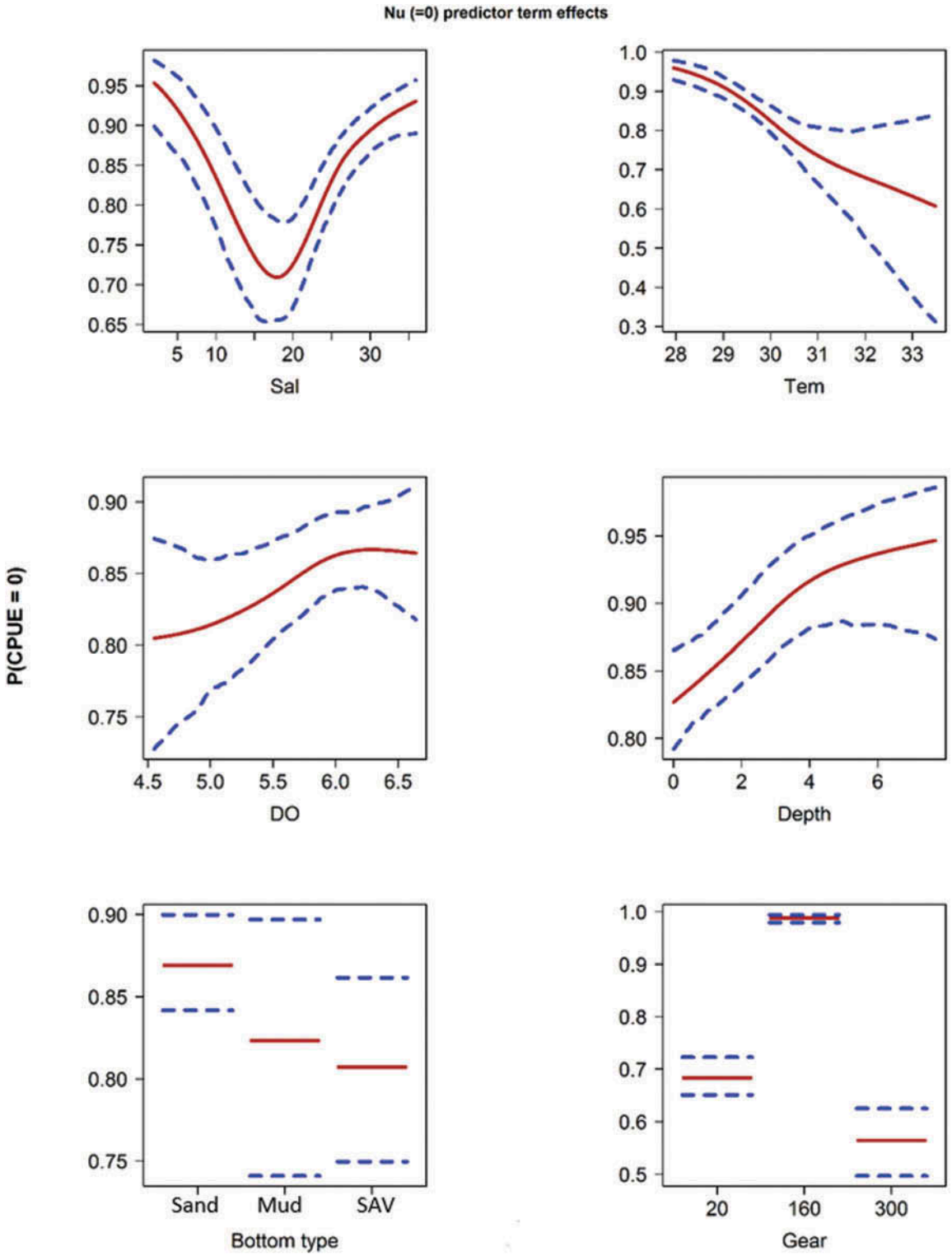

FIGURE 3. Fitted back-transformed CPUE = 0 spline function plots on the nu side of the reduced model for juvenile pink shrimp in summer. Red lines are fitted frequency of occurrence of zeros for the nu side. Dashed blue lines indicate upper and lower CI bounds.

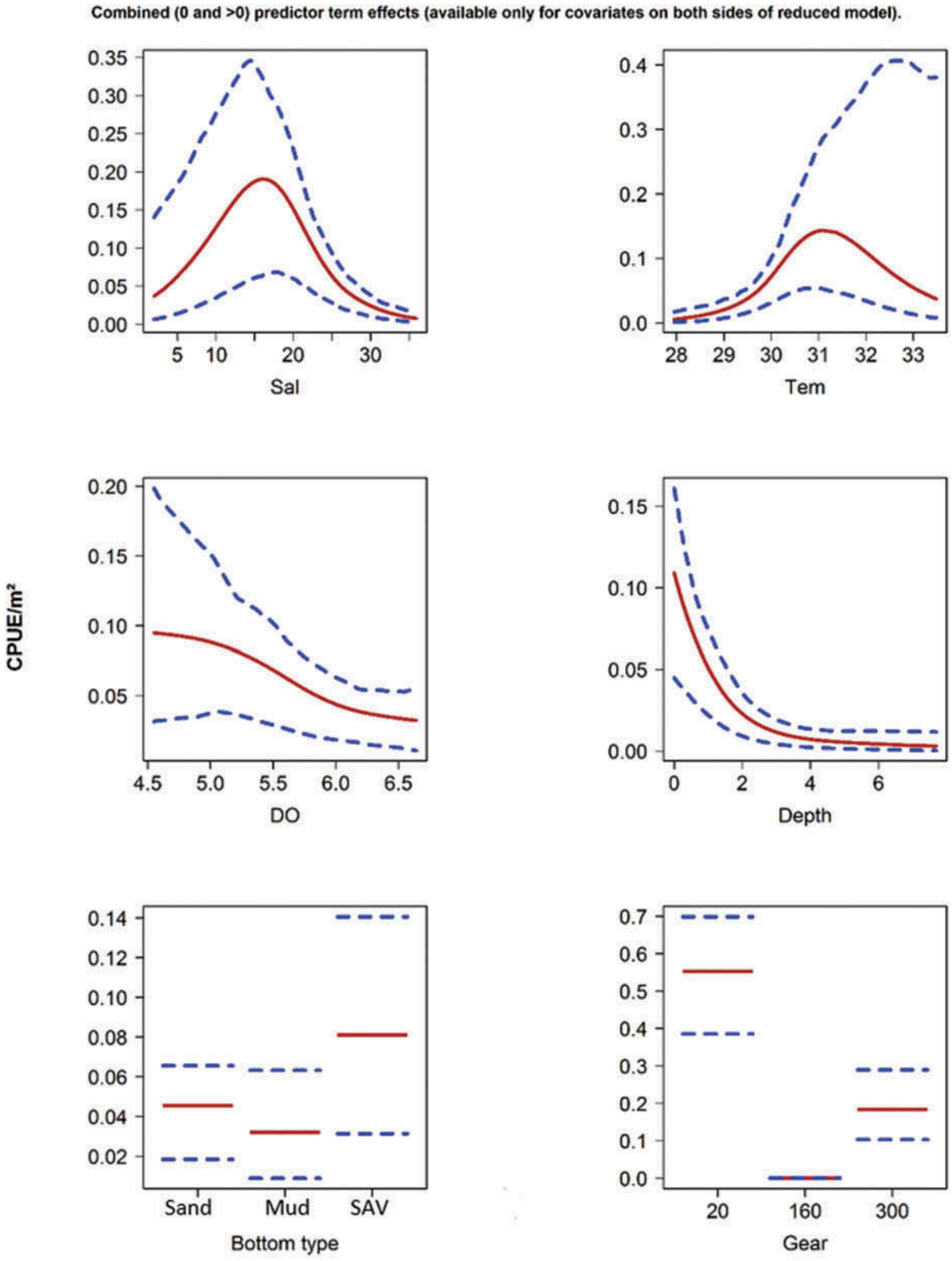

FIGURE 4. Fitted back-transformed overall CPUE (number/m^2) function plots for the reduced model for juvenile pink shrimp in summer. Red lines represent CPUEs (no/m^2) for mu × nu. Dashed blue lines indicate upper and lower CI bounds.

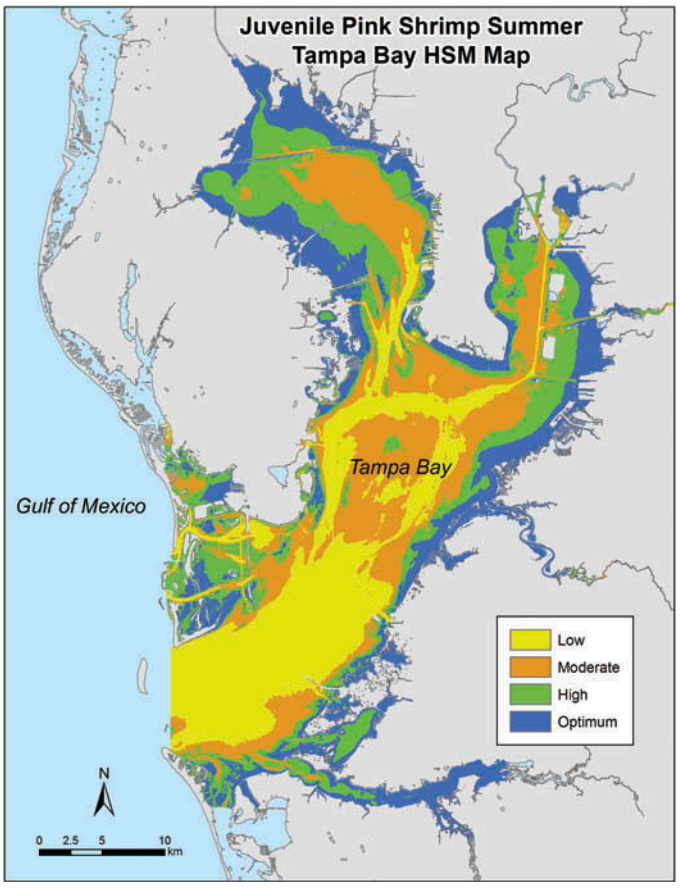

FIGURE 5. Habitat suitability model map with predicted CPUE zones (number/m^2) for juvenile pink shrimp in summer in Tampa Bay based on FIM sampling from 1998 to 2008.

We estimated the population abundance of juvenile pink shrimp in summer using habitat grids and suitability functions derived from 11 years of FIM sampling data. Since only about 400 FIM samples were collected by season each year, the samples were insufficient to interpolate temperature, salinity, and dissolved oxygen seasonally within years. Consequently, we did not attempt to estimate population numbers on an annual basis. Jensen et al. (2005) used delta-lognormal GAMs to model the distributions and map the abundance of female blue crabs *Callinectes sapidus* in Chesapeake Bay during the winter on an annual basis for 13 years (1990–2002). This was possible because their FIM program had large sample sizes, ranging from 867 to 1,599 each year.

TABLE 2. Estimated population numbers derived from the predicted HSM map.

HSM zone	Mean GC-CPUE	Cell count	Zonal area	Estimated population
Low	0.0131	1,061,493	238,836,000	3,128,752
Moderate	0.0390	1,081,831	243,412,000	9,493,068
High	0.1328	1,020,807	229,682,000	30,501,770
Optimum	0.9131	1,033,707	232,584,000	212,372,450
Total				255,496,040

TABLE 3. Verification tests to determine whether mean GC-CPUEs increase across HSM zones for two time periods.

	HSM zone				
Period	Low	Moderate	High	Optimum	*P*
1989–1997	0.0945	0.0619	0.1624	0.5879	<0.0001
1998–2008	0.0154	0.0412	0.1600	1.0825	<0.0001

Assuming that the number of FIM samples collected using random-stratified sampling in Tampa Bay is adequate to estimate annual CPUE time series, it may be possible to create delta-type models and annually predict population numbers of estuarine fish and shrimp species within seasons across years in the estuary. Suitability functions could be used with habitat grids that do not change between years (depth and bottom type) and seasonal grids that do change (temperature, salinity, current speed, and current direction derived from circulation modeling) to create predicted annual CPUE grids within seasons. The predicted interannual CPUE grids could be scaled to the CPUE time series and interannual population estimates derived from the adjusted CPUE grids. More research is required to determine whether this approach can be used to estimate interannual trends in population numbers to support management of the fisheries.

Analyses using delta-type GAM models provide information concerning the habitat requirements of estuarine species of fish and shrimp for various life stages and seasons. They can be used to help determine essential fish habitat (Rubec et al. 1998, 1999; Le Pape et al. 2003; Trimoreau et al. 2013) and habitat areas of particular concern for fisheries management (Rosenberg et al. 2000); to support ecosystem-based fisheries management (EPAP 1999; Carocci et al. 2009); to support inclusion of fish habitat information in fisheries ecosystem plans by U.S. fishery management councils (GMFMC 2005); to determine critical habitats for threatened and endangered species (Scott et al. 2006); to support oil spill response and natural resource damage assessment of areas impacted by chemical spills (French-McCay 2009); and to support coastal zone planning and management. The main products are a predicted CPUE grid with CPUEs computed in numbers per square meter, an HSM map created within the GIS, and tabular estimates of population abundance.

ACKNOWLEDGMENTS

We thank Robert Muller of the Florida Fish and Wildlife Conservation Commission (FWC) as well as Nicholas Farmer and Joan Browder of the National Marine Fisheries Service for reviewing the manuscript. We also appreciate the insightful comments provided by two anonymous reviewers. Editing by Bland Crowder of FWC helped to improve the manuscript.

REFERENCES

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. Pages 267–281 *in* B. Petrov and F. Csaki, editors. Proceedings of the second international symposium on information theory. Akademiai Kiado, Budapest.

Berg, C. W., A. Nielson, and K. Kristensen. 2014. Evaluation of alternative age-based methods for estimating relative abundance from survey data in relation to assessment models. Fisheries Research 151:91–99.

Booth, A. J. 1998. Spatial analysis of fish distribution and abundance patterns: a GIS approach. Pages 719–740 *in* F. Funk, T. J. Quinn II, J. Heifetz, J. N. Ianelli, J. E. Powers, J. F. Schweigert, P. J. Sullivan, and C.-I. Zhang, editors. Fishery stock assessment models. University of Alaska Sea Grant College Program, Publication AF-SG-98-01, Fairbanks.

Booth, A. J. 2000. Incorporating the spatial component of fisheries data into stock assessment models. ICES Journal of Marine Science 57:858–865.

Browder, J. A., and M. B. Robblee. 2009. Pink shrimp as an indicator for restoration of Everglades ecosystems. Ecological Indicators 9 (Supplement):S17–S28.

Browder, J. A., Z. Zein-Eldin, M. M. Criales, M. B. Robblee, and T. I. Jackson. 2002. Dynamics of pink shrimp recruitment in relation to Florida Bay salinity and temperature. Estuaries 25:1335–1371.

Carocci, F., G. Bianchi, P. Eastwood, and G. Meaden. 2009. Geographic information systems to support the ecosystem approach to fisheries. FAO (Food and Agriculture Organization of the United Nations) Technical Report 532.

Costello, T. J., and D. M. Allen. 1970. Synopsis of biological data on the pink shrimp *Penaeus duorarum* Burkenroad, 1939. FAO (Food and Agriculture Organization of the United Nations) Fisheries Report 57:1499–1538.

Drexler, M., and C. H. Ainsworth. 2013. Generalized additive models used to predict species abundance in the Gulf of Mexico: an ecosystem modeling tool. PLoS (Public Library of Science) ONE [online serial] 8(5):e64458.

Dunn, P. K., and G. K. Smyth. 1996. Randomized quantile residuals. Journal of Computational and Graphical Statistics 5:236–244.

Eilers, P. H. C., and B. D. Marx. 1996. Flexible smoothing with B-splines and penalties. Statistical Science 11:89–121.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40:677–697.

EPAP (Ecosystem Principles Advisory Panel). 1999. Ecosystem-based fishery management. Report to Congress. Ecosystem Principles Advisory Panel, Seattle.

Firth, D. 1988. Multiplicative errors: log-normal or gamma? Journal Research Statistical Society Series B Methodological 50:266–268.

Fletcher, D., D. MacKenzie, and E. Villouta. 2005. Modelling skewed data with many zeros: a simple approach combining ordinary and logistic regression. Environmental and Ecological Statistics 12:45–54.

French-McCay, D. 2009. State-of-the-art and research needs for oil spill impact assessment modeling. Pages 601–653 *in* Proceedings of the 32nd AMOP technical seminar on environmental contamination and response. Environment Canada, Emergencies Science Division, Ottawa.

FWRI (Fish and Wildlife Research Institute). 2013. Fisheries-independent monitoring program 2013 annual data summary report. Florida Fish and Wildlife Conservation Commission, In-House Report IHR2014-003, St. Petersburg.

GMFMC (Gulf of Mexico Fisheries Management Council). 2005. Final generic amendment number 3 for addressing essential fish habitat requirements, habitat areas of particular concern, and adverse effects of fishing in the following fishery management plans of the Gulf of Mexico (GOM): shrimp fishery of the GOM, United States waters, Red Drum fishery of the GOM, reef fish fishery of the GOM, coastal migratory pelagic resources (mackerels) in the GOM and South Atlantic, stone crab fishery of the GOM, spiny lobster fishery of the GOM, spiny lobster in the GOM and South Atlantic, coral and coral reefs of the GOM. GMFMC, Tampa, Florida.

Grüss, A., M. Drexler, and C. H. Ainsworth. 2014. Using delta generalized additive models to produce distribution maps for spatially explicit ecosystem models. Fisheries Research 159:11–24.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modeling 135:147–186.

Hildebrand, H. H. 1955. A study of the fauna of the pink shrimp grounds in the Gulf of Campeche. Publications of the Institute of Marine Science University of Texas 4:169–232.

Jensen, O. P., R. Seppelt, T. J. Miller, and L. J. Bauer. 2005. Winter distribution of blue crab *Callinectes sapidus* in Chesapeake Bay: application and cross validation of a two-stage generalized additive model. Marine Ecology Progress Series 299:239–255.

Le Pape, O., F. Chauvet, S. Mahévas, P. Lazure, D. Guérault, and Y. Désaunay. 2003. Quantitative description of habitat suitability for the juvenile Common Sole (*Solea solea*, L.) in the Bay of Biscay (France) and the contribution of different habitats to the adult population. Journal of Sea Research 50:139–149.

Le Pape, O., J. Delavenne, and S. Vaz. 2014. Quantitative mapping of fish habitat: a useful tool to design specialized management measures and marine protected areas with fishery objectives. Ocean and Coastal Management 87:8–19.

Lewis, R. R. III, and E. D. Estevez. 1988. The ecology of Tampa Bay, Florida: an estuarine profile. U.S. Fish and Wildlife Service Biological Report 85 (7.18).

Li, Y., Y. Jiao, and Q. Hi. 2011a. Decreasing uncertainty in catch rate analyses using Delta-AdaBoost: an alternative approach in catch and bycatch analyses with high percentage of zeros. Fisheries Research 107:261–271.

Li, Y., Y. Jiao, and K. Reid. 2011b. Gill-net saturation in Lake Erie: effects of soak time and fish accumulation on catch per unit effort of Walleye and Yellow Perch. North American Journal of Fisheries Management 31:280–290.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: improving ecological inference by modeling of zero observations. Ecology Letters 8:1235–1246.

Matheson, R. Jr., R. McMichael Jr., D. Leffler, and T. MacDonald. 2005. Populations of juvenile and small-adult fishes in Tampa Bay: a decadal perspective. Pages 3–18 *in* S. F. Treat, editor. Proceedings, Tampa Bay area scientific information symposium, BASIS4. Tampa Bay National Estuary Program, St. Petersburg, Florida.

Maunder, M. N., and A. E. Punt. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70:141–159.

Muller, R. G., A. A. Trotter, and P. W. Stevens. 2015. The 2015 stock assessment update of Common Snook, *Centropomus undecimalis*. Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, IHR 2015-004, St. Petersburg.

Myers, R. A., and P. Pepin. 1990. The robustness of lognormal-based estimators of abundance. Biometrics 46:1185–1192.

Nishida, T., and A. J. Booth. 2001. Recent approaches using GIS in the spatial analysis of fish populations. Pages 19–36 *in* G. H. Kruse, N. Bez, A. Booth, M. W. Dorn, S. Hills, R. N. Lipcius, D. Pelletier, C. Roy, S. J. Smith, and D. Witherell, editors. Spatial processes and management of

marine populations. University of Alaska Sea Grant College Program, Publication AK-SG-01-03, Fairbanks.

Ospina, R., and S. L. P. Ferrari. 2010. Inflated beta distributions. Statistical Papers 51:111–126.

Owen, A. B. 1995. Nonparametric confidence bands for a distribution function. Journal of the American Statistical Association 90:516–521.

Pennington, M.1983. Efficient estimators of abundance for fish and plankton surveys. Biometrics 39:281–286.

Planque, B., C. Loots, P. Petitgas, U. Lindstrøme, and V. Sandrine. 2011. Understanding what controls the spatial distribution of fish populations using a multimodel approach. Fisheries Oceanography 20:1–17.

R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available: http://www.R-project.org/. (March 2016).

Rigby, R. A., and D. M. Stasinopoulos. 2005. Generalized additive models for location, scale, and shape. Applied Statistics 54:507–554.

Rigby, B., M. Stasinopoulos, and V. Voudouris. 2013. GAMLSS flexible regression in R. [online book]. London Metropolitan University, London. Available: http://www.gamlss.org/wp-content/uploads/2013/10/FlexibleRegression.pdf. (March 2016).

Rosenberg, A., T. E. Bigford, S. Leathery, R. L. Hill, and K. Bickers. 2000. Ecosystem approaches to fishery management through essential fish habitat. Bulletin of Marine Science 66:535–542.

Rubec, P. J., J. C. W. Bexley, H. Norris, M. S. Coyne, M. E. Monaco, S. G. Smith, and J. S. Ault. 1999. Suitability modeling to delineate habitat essential to sustainable fisheries. Pages 108–133 *in* L. R. Benaka, editor. Fish habitat: essential fish habitat and restoration. American Fisheries Society, Symposium 22, Bethesda, Maryland.

Rubec, P. J., M. S. Coyne, R. H. McMichael Jr., and M. E. Monaco. 1998. Spatial methods being developed in Florida to determine essential fish habitat. Fisheries 23(7):21–25.

Scott, J. M., D. D. Goble, and F. W. Davis. 2006. The Endangered Species Act at thirty, volume 2. Conserving biodiversity in human-dominated landscapes. Island Press, Washington, D.C.

Stasinopoulos, D. M., and R. A. Rigby. 2007. Generalized additive models for location, scale, and shape (GAMLSS) in R. Journal of Statistical Software 23:1–46.

Stasinopoulos, M., B. Rigby, and C. Akantziliotou. 2008. Instructions on how to use the gamlss package in R, 2nd edition [online book]. London Metropolitan University, London. Available: http://www.gamlss.org/wp-content/uploads/2013/01/gamlss-manual.pdf. (March 2016).

Stasinopoulos, M., M. Rigby, C. Akantziliotou, G. Heller, R. Ospina, N. Motpan, F. McElduff, V. Voudouris, M. Djennad, M. Enea, and A. Ghalanos. 2014. gamlss.dist: distributions to be used for GAMLSS modeling in R package version 4.3-0. Available: http://CRAN.R-project.org/package=gamlss.dist. (March 2016).

Stefánsson, G. 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES Journal of Marine Sciences 53:577–588.

Syrjala, S. 2000. Critique of the use of the delta distribution for the analysis of trawl survey data. ICES Journal of Marine Science 57:831–842.

TBNEP (Tampa Bay National Estuary Program). 1996. Charting the course for Tampa Bay: comprehensive conservation and management plan for Tampa Bay. TBNEP, St. Petersburg, Florida.

Trimoreau, E., B. Archambault, A. Brind'Amour, M. Lepage, J. Guitton, and O. Le Pape. 2013. A quantitative estimate of the function of soft-bottom sheltered coastal areas as essential flatfish nursery habitat. Estuarine, Coastal, and Shelf Science 133:193–205.

Venables, W. N., and B. D. Ripley. 2002. Modern applied statistics with S, 4th edition. Springer-Verlag, New York.

Williams, A. B. 1955. A survey of North Carolina shrimp distribution. Journal of Elisha Mitchell Scientific Society 71:200–207.

Wolfe, S. H., and R. D. Drew, editors. 1990. An ecological characterization of the Tampa Bay watershed. U.S. Fish and Wildlife Service, Biological Report 90(20).

Ye, Y., M. Al-Husaini, and A. Al-Baz. 2001. Use of generalized linear models to analyze catch rates having zero values: the Kuwait drift net fishery. Fisheries Research 53:151–168.