Can climate explain temporal trends in king mackerel (*Scomberomorus cavalla*) catch-per-unit-effort and landings?

Harford, W.J, Sagarese, S.R., Nuttall, M.A., Karnauskas, M., Liu, H., Lauretta, M., Schirripa, M. & Walter, J.F.

SEDAR38-AW-04

20 March 2014
Updates 14 July 2014

Please cite this document as:

Harford, W.J, S.R. Sagarese, M.A. Nuttall, M.A. Karnauskas, H. Liu, M. Lauretta, M. Schirripa, and J.F. Walter. 2014. Can climate explain temporal trends in king mackerel (*Scomberomorus cavalla*) catch-per-unit-effort and landings?. SEDAR38-AW-04. SEDAR, North Charleston, SC. 29 pp.

Can climate explain temporal trends in king mackerel (*Scomberomorus cavalla*) catch-per-unit-effort and landings?

July 14, 2014

Harford[1], W.J, Sagarese[1], S.R., Nuttall[2], M.A., Karnauskas[3], M., Liu[4], H., Lauretta[3], M., Schirripa[3], M. & Walter[3], J.F.

[1]Cooperative Institute for Marine and Atmospheric Studies, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Cswy, Miami, Fl, 33149

[2]Department of Marine Biology & Fisheries, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Cswy, Miami, Fl, 33149

[3]NOAA National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division, 75 Virginia Beach Dr, Miami, Fl, 33149

[4]Marine Biology Department, Texas A&M University at Galveston, Galveston, Texas, 77553

1. Introduction

Habitat associations of marine fishes are known to result from trade-offs between predation, food availability, or metabolic demands. Accordingly, environmental (e.g. temperature, salinity) and ecological factors (e.g. prey abundance) are known to be correlated with fish distributions (Gratwicke et al. 2006, MacRae and Cowan Jr 2010, McKinney et al. 2012). Investigating these relationships falls within the realm of fisheries oceanography, where the majority of studies have focused on commercially important species including large pelagics such as tunas (Andrade and Garcia 1999, Block et al. 1997, Schick et al. 2004) and groundfish including Atlantic cod (*Gadus morhua*) (Begg and Marteinsdottir 2002, Castonguay et al. 1999, Hedger et al. 2004). Over the last decade, the importance of incorporating ecosystem considerations into fisheries stock assessments has been reinforced by updated management measures (MSFCMA 2007). These considerations include calls to understand whether and how fluctuations in stock dynamics are environmentally driven.

Calls to include environmental drivers in stock assessment procedures have been made by managing partners and stakeholders alike. And indeed, determining if fluctuations in fish stocks arise from natural events and (or) fishing has long been debated in fisheries management (Thompson-Burkenroad debate; Hilborn & Walters 1992, Planque et al. 2010). Given the reliance of stock assessments on catch-per-unit-effort (CPUE) indices, it is of particular interest to separate natural and human-induced influences on CPUE fluctuations. In practice, natural sources of the fluctuations in biological time series, like CPUE, can be difficult to detect. Challenges in identifying environmental influences on CPUE can occur when sampling is not randomized (i.e. is concentrated in fished areas), is not representative of the geographic domain of the stock, or when observation does not occur at scales relevant to ecological investigation

(Wiens 1989, Rotherham et al. 2007, Johnson et al. 2012). In addition, non-linearity of biological systems can confound the use of traditional linear statistical methods for identifying key environmental drivers (Myers 1998, Sugihara et al. 2012). Highly variable and non-linear dynamics tend to be pervasive in biological systems and can be further magnified where natural populations are subject to harvesting (Ricker 1954, Anderson et al. 2008, Quetglas et al. 2012, Glaser et al. 2013).

In this study, we examined evidence of environmental forcing in fisheries-dependent CPUE indices of king mackerel (*Scomberomorus cavalla*). As a highly migratory species, king mackerel are distributed throughout US waters of the South Atlantic and Gulf of Mexico. Assessment of king mackerel is conducted separately for the South Atlantic and the Gulf of Mexico and accounts for mixing of these stocks off south Florida during winter (SEDAR 2009). Seasonal migrations are supported by correlations between catch per unit effort and water temperature in the eastern Gulf of Mexico and Atlantic waters (SEDAR 2009). Furthermore, potential links have been proposed between catch rates and ecosystem conditions, including water clarity, upwelling, temperature, and baitfish presence (Barile 2013, Wall et al. 2009). Our analyses emphasized the consideration that environment-CPUE interactions could be highly non-linear and statistical methods were chosen that had the flexibility to capture linear or non-linear relationships. Using the approaches of generalized additive modeling, boosted regression trees, and state-space reconstruction, we examined evidence of the influence of environmental conditions on landings and CPUE of the South Atlantic king mackerel stock.

2. Methods

2.1 King mackerel data

The Coastal Fisheries Logbook Program has monitored federally permitted commercial vessels in the SE US Atlantic since 1993. These censuses have spatially and temporally explicit observations of fishery catch and effort for a variety of commercially important stocks. For generalized additive modeling and boosted regression tree analyses, effort variables included region fished (Gulf of Mexico, South Atlantic, Mixing Zone), area, year, month, gear type (hook and line or trolling), number of lines (numgear), number of hooks per line, hours fished, number of crew, days away from port, and target species. Target species was treated as a categorical variable representing the degree to which king mackerel were targeted on a given fishing trip as indexed by the percentage of king mackerel in the total catch (0-25% = 1, 25-50% = 2, 50-75% = 3, and 75-100% = 4). Two response variables were obtained from this dataset: presence-absence (1=present, 0=absent) and positive CPUE (i.e. catch > 0) with:

$$CPUE = \frac{king\ mackerel\ caught\ (pounds)}{numgear * number\ of\ hooks * hours\ fished} = \frac{lbs}{\#Hooks * \#\ Hours} \tag{1}$$

where the variables are defined above. Spatial resolution within this dataset is expressed as area fished, which generally consists of 1° by 1° (Fig. 1). For the state-space reconstruction analyses, logbook CPUE and landings time series were calculated as monthly anomalies for the period of 1993 to 2012 and were standardized to reflect the spatial domain of the South Atlantic stock.

2.2 Environmental data

Sea surface temperature (SST) data were obtained for various locations in the South Atlantic and Gulf of Mexico between 1993 and 2013 using AVHRR Pathfinder Version 5.2 (PFV5.2) data, obtained from the US National Oceanographic Data Center and GHRSST (http://pathfinder.nodc.noaa.gov). The PFV5.2 data are an updated version of the Pathfinder Version 5.0 and 5.1 collection described in Casey et al. (2010). This dataset provided observations for a number of environmental variables with a spatial resolution of 4 km^2. For generalized additive modeling and boosted regression tree analyses, weekly SST values were obtained to account for missing observations (i.e. due to cloud cover). CPUE observations were then assigned a latitude and longitude based on the center of each area's spatial extent. Using these spatial reference points and the date fished, each CPUE observation was assigned a weekly SST measurement.

For the state-space reconstruction analyses, SST was calculated as standardized monthly anomalies for 1993 to 2012. We also gathered monthly North Atlantic Oscillation (NAO) anomalies from 1993 to 2012 from NOAA National Weather Service, Climate Prediction Center. NAO was included in our analysis as a coarse indicator of climatic and oceanographic variability. NAO is a measure of atmospheric variation in sea level barometric pressure and has been identified as a proxy (i.e. NAO may be correlated with more proximate drivers) for environmental drivers of dish population dynamics (Hurrell 1995, Hurrell & Deser 2010).

2.3 Generalized additive modeling

The influence of SST on the distribution of king mackerel was explored using generalized additive models (GAMs) (Hastie and Tibshirani 1990, Wood 2006). GAMs are a semi-parametric extension of the generalized linear model (GLM) and commonly model the spatial distributions of fishes (Damalas et al. 2010, Guisan et al. 2002, Leathwick et al. 2006). While GLMs use a linear predictor to define the relationship between the response and explanatory variables, GAMs utilize a smoothing function (Wintle et al. 2005). A GAM is defined by the following relationship:

$$(2) \qquad Y = \alpha + \sum f_i \, (X_i) + \varepsilon$$

where the response variable, Y, is a function of the sum of the explanatory variable(s) (X_i) altered by a nonparametric smoothing function (f_i), an intercept term (α), and an error measure (ε). GAMs can include either nonparametric and (or) linear terms for predictors, therefore uncovering structure in the data that might otherwise have been missed with usual linear assumptions (Hastie and Tibshirani 1990, Wood 2006). While GAMs are less restrictive than parametric techniques in assumptions about the underlying statistical distribution of data, they assume additivity of the functions and smoothness of the components (Guisan et al. 2002, Maravelias 1999, Maravelias et al. 2000).

The delta-lognormal approach was employed to first predict the probability of king mackerel occurrence and then their abundance given that at least one individual was observed. Trends were described for all regions combined and for each area separately. The occurrence model employed a logit link function and a binomial error distribution whereas the abundance model used a log link function and a negative binomial error distribution. GAMs were built using forward-selection and initially considered all predictor variables other than SST. The order in which individual predictor variables were added to GAMs and ultimate stopping points (in which no more predictor variables were included) was determined using the percent deviance explained, Akaike's Information Criteria (AIC), and compiled scree plots (% deviance explained vs. number of included predictors). With the optimal model obtained (of those excluding SST), the relative significance/importance of SST was then tested by adding SST to this model.

2.4 Boosted regression tree analysis

Boosted regression tree analysis joins traditional regression methods with boosting by combining a large number of simple decision trees into a single model, resulting in strong predictive performance and good description of modeled relationships (Elith et al. 2008). Inclusion of a random component within the model improves accuracy, increases speed, and reduces overfitting (Elith et al. 2008, Friedman 2002). Two parameters are required: the (1) learning rate (*lr*) which determines the contribution of each tree to the growing model and (2) tree complexity (*tc*) which controls whether interactions are fitted (*tc* = 1 for additive model, *tc* = 5 for inclusion of five-way interactions, etc.). The best model was determined by the combination of *lr*, *tc*, and the number of trees (*nt*) that minimized prediction error. All BRT analyses were run in R using the 'gbm' package supplemented with functions from Elith et al. (2008). The optimal combination was validated by ten-fold cross validation with a training dataset (75% of observations randomly allocated) to determine the optimal number of trees. Model performance was assessed by the predictive deviance between the test dataset (remaining 25% of observations) and predicted values (De'ath 2007, Elith et al. 2008, Froeschke et al. 2010, Leathwick et al. 2008) and through examination of the area under the receiver operator characteristic (ROC) curve (Hanley and McNeil 1982). The relative importance of each variable is determined by averaging the number of times a variable is selected for splitting and the squared improvement resulting from these splits, with higher percentages corresponding to stronger influences on the response variable (Froeschke et al. 2010).

2.5 Non-parametric state-space reconstruction

It is generally recognized that non-linear processes and external forcing can produce complex fluctuations in fish population dynamics (Glaser et al. 2013). Parametric models can be used to test whether observations are consistent with assumed functional relationships. Alternatively, non-parametric time series models can be used to test whether relationships exist between variables without reliance on functional forms. State-space reconstruction (SSR) techniques make no assumptions about functional relationships. Instead, they rely on the structure of the observed data to make forecasts, and are therefore flexible to accommodate linear and non-linear system behaviors (Glaser et al. 2011, Perretti et al. 2013). System behavior is reconstructed based on Takens (1981) theorem, which shows that a multidimensional system can be reconstructed through time-delayed coordinate embedding of a univariate time series. Essentially this means that a univariate time series will not only contain information about its intrinsic dynamics, but will also contain information about unobserved influential variables (Takens 1981, Deyle & Sugihara 2011). Time-delayed coordinate embedding involves transforming a univariate time series into a set of time-delayed vectors:

$$(3)\quad \mathbf{Y}_t = \left[y_t, y_{t-\tau}, y_{t-2\tau}, \ldots, y_{t-(E-1)\tau} \right],$$

where y is a univariate time series, t is the time index, τ is the time lag, and E is the embedding dimension. The embedding dimension is the minimum number of time-delayed coordinates of y that are needed to reconstruct the state space and serves as an index of system complexity as higher dimensional embedding will be needed to resolve attractors comprised of many interacting processes (Sugihara & May 1990, Glaser et al. 2013). These time-delayed vectors form the axis coordinates of the reconstructed state space. Reconstruction reflects the trajectories of the original state space, thus capturing properties of the original state space attractor (Takens

1981, Sugihara & May 1990, Deyle & Sugihara 2011). The SSR techniques that we applied specifically refer to those developed by Sugihara and May (1990) and Sugihara (1994). The simplex projection model was used to identify the best embedding dimension for the time series (Sugihara & May 1990). This embedding dimension was then used in a technique called sequentially weighted global linear maps (S-maps), which assess time series non-linearity (Sugihara 1994). Both of these modeling approaches produce forecasts. Prior to analysis, CPUE time series were processed by taking first differences ($\Delta y = y_t - y_{t-1}$) and normalizing (i.e. mean = 0, standard deviation = 1). All SSR analyses were carried out using monthly time steps.

The simplex projection model is a nearest neighbor algorithm that uses the shape of the reconstructed attractor in *E*-dimensions to predict the trajectory of the time series from time *t* to time *t+1*. A detailed mathematical description of the algorithm is provided in Sugihara and May (Sugihara & May 1990). In summary, simplex projection contains only one parameter, *E*, which is sequentially varied from 1 to 10. The chosen embedding dimension determines the lengths of the coordinate vectors. To generate forecasts for the prediction vectors, Euclidean distances to all library vectors are calculated, and the *E+1* nearest neighbors in the reconstructed state space are used to calculate a forecasted value. The set of *E+1* nearest neighbors are the minimum number of vertices needed to define a unique location in *E*-dimensional space (Sugihara & May 1990, Liu et al. 2012). Forecasts are made from the *E+1* nearest neighbors as an exponentially weighted average:

$$(4)\quad \hat{y}_T = \sum_{j=0}^{E} w_j \, y_T\left(j\right),$$

where weightings, w_j, for each neighbor *j*, are determined at time *t*, and applied to forecasts of the prediction vector, $\hat{y}_T$, at *T=t+1*. This means that library coordinate vectors that are similar to the prediction vector at time *t* and are also expected to have similar states at *t+1*. For each embedding dimension, the state space of the attractor will be differently shaped, and thus, each successive embedding will produce different forecasts. Forecast skill was defined as the Pearson correlation coefficient (ρ) and the *E* resulting in the highest ρ was selected as the best embedding dimension. Coordinate vectors constructed using the best embedding dimension were then used as input to the S-maps technique.

S-maps examine evidence of non-linearity by comparing the forecast skill obtained using equivalent linear or non-linear models. A detailed account of the implementation details are described by Sugihara (1994). Briefly, forecasts are made using all library vectors, and weighting their importance based on Euclidean distances from the prediction vector. Forecasts produced by S-maps are tuned by adjusting a non-linearity parameter, θ, within the continuous range of 0 to 10. The non-linearity parameter controls how heavily library vectors that are nearest to the prediction vector are weighed relative to distant library vectors:

$$(5)\quad w(d) = \exp\left(-\theta d / \overline{d}\right).$$

Euclidean distance, *d*, is calculated between a library vector and the prediction vector, and $\overline{d}$ is the minimum distance from all library vectors. When $\theta = 0$, linear forecasts are produced, that are essentially equivalent to a vector auto-regressive (AR) model of order *E* (Glaser et al. 2011, Liu et al. 2012). When $\theta > 0$ non-linear forecasts are produced that give greater weight to library vectors that are in close proximity to the prediction vector. Again, the best S-map model is the one that produces the highest out-of-sample forecast skill.

In addition to univariate time series forecasting (Sugihara 1994), SSR techniques have been extended to the situation where more than one time series from the same system is available for analysis (Deyle & Sugihara 2011, Sugihara et al. 2012). We used a technique called convergent cross-mapping (CCM) to determine if past trends in climate could be reconstructed from past trends in CPUE or landings (Sugihara et al. 2012). When an environmental variable, X, is a driver of biological variable, Y, information about the past history of X will be contained in the past history of Y, *but not vice versa* (Sugihara et al. 2012). Put another way, fish dynamics may contain information about past climate events, but we would not expect climate to contain information about fish dynamics (Sugihara et al. 2012). CCM was implemented using the simplex projection technique. Forecast skill of CCM was quantified using out-of-sample prediction. Since our time series were short, forecast skill was evaluated using leave-one-out cross-validation (Sugihara et al. 1996, Glaser et al. 2011, 2013, Liu et al. 2014). Forecast skill was measured as the Pearson correlation (ρ) between observed and predicted values. Of particular relevance to the CCM technique is whether convergence is evident in forecasting skill. Convergence is defined as an improvement in forecast skill (ρ) as more information (in the form of time series length) is added to the analysis. If convergence is evident, adding more events to SSR will more densely populate state space, and thus, will reduce estimation error of out-of-sample events (Sugihara et al. 2012). Convergence is essential to demonstrate that environmental variable, X, is a driver of biological variable, Y.

We also used a multivariate generalization of SSR to evaluate whether climate effects could improve forecasting skill of king mackerel CPUE. Multivariate SSR was implemented by including climate predictors X in the lagged coordinates of Y (e.g. $\left[y_t, y_{t-\tau}, y_{t-2\tau}, \ldots, y_{t-(E-1)\tau}, x_t \right]$). When Y varies in a manner that is dependent on the previous states of X, then Y and X can be considered to be interacting (Deyle et al. 2013). Multivariate SSR was implementing using the S-map technique (Sugihara 1994). A test of improvement in forecast skill when a climate variable, *X*, was included in SSR was conducted by calculating the improvement in forecasting skill ($\Delta\rho$) between univariate and multivariate models. Testing occurred by shuffling *X* and then using the S-map procedure to generate forecasts, from which null $\Delta\rho$ was calculated. Null $\Delta\rho$ represents the forecasting improvement expected by chance alone. This process was repeated 1,000 times to determine the probability that observed forecast improvement was greater than chance alone, using a cutoff p-value of 0.05.

3 Results

3.1 Generalized additive modeling

SST was determined to be significant in every GAM, however the improvement in model fit was negligible. This marginal influence of SST was seen in both the relative change in percent deviance explained (Figs. 2– 5) and relatively flat fit of the smooth function in the CPUE-SST relationship (slope ≈ 0.1). The current effect of SST (considered negligible with % deviance explained) is believed to be more an issue with the dataset than representative of the real king mackerel-SST interaction in the SE US, a thought that will be expanded upon in the discussion.

3.2 Boosted regression tree analysis

The best BRT model (*lr* = 0.005, *tc* = 5) fit 3,150 trees and explained 75.1% of the deviance. This deviance explained was primarily driven by target, serving as both a main effect and influential interactions (e.g. target x area, target x hour fished), The model converged on a stable solution as evident by the predictive deviance leveling off at the lowest value as the number of trees increased. Cross validation statistics were adequate for measures of both correlation (0.884 ± 0.001) and ROC (0.97 ± 0.001). This chosen model identified target species as the most influential variable (Table 1) followed by area. Out of all 10 explanatory variables, SST was ranked 8th in relative importance.

3.3 State-space reconstruction

Cross-mapping signals between NAO anomalies and CPUE or landings were not detected (Fig. 6). Indeed, models failed to display convergent behavior when NAO was used to reconstruct CPUE or landings (Fig. 6). Some improvement in the predictive ability of NAO can be seen as the length (L) of the CPUE time series increases; however, predictive ability of NAO remained weak overall (Fig. 6, top panel, blue line). Results from multivariate SSR were generally consistent with CCM. Using NAO as an environmental driver did not significantly improve prediction skill of relative biomass or landings (Table 2). In the multivariate SSR analysis we also considered different time lags of the effects of NAO on CPUE of up to 3 months. Time lags of 2 or 3 months performed more poorly than the 1-month times lags reported in Table 2.

Similar to the results for NAO, cross mapping did not reveal interactions between SST and CPUE or landings (Fig. 7). CCM did not show convergent behavior and predictive skill (ρ) was very low. Further, multivariate SSR failed to detect significant improvement in predictive skill when SST was included in multivariate forecasting (Table 2).

4 Discussion

At the temporal and spatial scales considered in this investigation, environmental forcing (specifically, by SST, NAO, AMO) was not evident in CPUE of king mackerel. Each of the modeling approaches consistently failed to detect coarse scale influences of SST on logbook CPUE. Non-parametric time series modeling also failed to reveal evidence of NAO as a driver of logbook CPUE. SST and NAO as predictors of landings were also explored, which provided an alternative test of whether a signature of climate-driven changes in king mackerel availability was evident in fishery data. However, climate-driven effects on landings were not detected in our analyses.

Despite the lack of influence of SST in the king mackerel GAMs, as evident by the marginal improvement in percent deviance explained after adding SST (Figs. 2 – 5), this analysis did show SST as a significant predictor of nominal king mackerel CPUE ($p < 0.001$). Indeed, this significance was expected for a fish that apparently prefers the 20ºC isotherm throughout its global range (Valeiras and Abad 2006). While this significance may be a factor of the large sample size of the king mackerel logbook dataset, it may also reflect the spatial distribution of fishing effort. Specifically, area fished (the *area* predictor variable) is likely to be dependent on the spatial distribution of the targeted stock (*target*). If this stock's spatial distribution is then dependent on the availability of preferential conditions (SST, NAO, etc), then area fished may also be biased towards these preferred conditions. As a result, this analysis was thought to underestimate the relative importance of SST; the magnitude of the predicted king mackerel-SST effect was believed inaccurate to account for any effect of fisher targeting behavior. In addition, the weak effect of SST could be a consequence of poor precision of CPUE estimates or the coarse spatial scale at which the analysis was conducted.

In applying non-parametric time series modeling, it is important to recognize the limitations of CCM and multivariate SSR in the context of interpreting the study results. The permutation test used in multivariate SSR evaluated empirical evidence of interactions between climate variables and king mackerel CPUE. Significance of the permutation test would reveal such an interaction; however, non-significance may be insufficient to reject the possibility that such an interaction exists. Deyle et al. (2013) applied a similar test to the situation where the driving variable X is stochastic. However, if X cannot be considered stochastic (i.e. the climate variable contains a deterministic signal), then the meaning of non-significance is less clear owing to mathematical theory on which the analyses are constructed (i.e. Takens 1981). Conversely, CCM is useful for demonstration of an interaction between X and Y, even when X is deterministic.

When the analyses conducted in this study are viewed together, they provide some guidance for future research. The analysis using GAMs revealed a significant contribution of SST to prediction of CPUE, but the additional complexity of including SST produced a very minor improvement in fit. Whether higher spatial resolution could reveal effects of environmental drivers on the distribution of king mackerel remains an open question. Similarly, non-parametric modeling, CCM in particular, revealed convergent behavior when CPUE was used to reconstruct NAO time series. Again however, the predictive skill of this relationship was quite weak. This result could reflect observation error or could reflect a mismatch between the temporal and/or spatial scales of the data and the scales at which environmental variables truly influence king mackerel distribution. In addition, NAO fluctuations themselves may or may not be a causative agent for relative biomass fluctuations; rather, the weak signal associated with

NAO could reflect associations with more important environmental drivers that are correlated with NAO.

These findings highlight the importance of simultaneous collection of *in situ* information of environmental conditions and the associated levels of catch and effort. This approach could help to resolve issues with data resolution and are especially important for regions such as the Gulf of Mexico Large Marine Ecosystem, in which management is challenged by numerous and complex trophic linkages that still require quantification (Walters et al 2008). To enhance management of king mackerel, further research is needed for: (1) habitat selection of king mackerel and how the environment influences their distributions; (2) important ecological interactions (e.g. SA king mackerel and croaker (P. Barile, pers comm.)); and (3) development of ecological indices (i.e. primary production and prey availability) for incorporation into stock assessments of king mackerel. Much work remains in elucidating species dynamics within the Gulf of Mexico, especially when compared to its temperate northeast US counterpart, which possesses one of the largest time series of species distributions, abundance, and trophic dynamics collected in a consistent manner.

Effective fishery management depends upon reliable indices of relative biomass/abundance, a case made famous by Atlantic cod (*Gadus morhua*) management where biased estimates led to stock collapse (Hutchings 1996). Targeting and identifying drivers of stock dynamics is one approach for moving towards reduction of bias in perceived stock trends. Here, we presented a variety of statistical approaches to examine the effects of climate variables on temporal and spatial trends in king mackerel CPUE. Habitat modeling enables environmental and ecological influences on seasonal distributions of marine stocks to be elucidated (Stoner et al. 2001), and thus, can reduce bias in stock size estimation when important drivers are identified. Non-parametric time series modeling has proven similarly useful for short-term forecasting of complex non-linear dynamics (Glaser et al. 2011, Liu et al. 2012, 2014) and for identifying non-linear relationships between biological and environmental variables (Sugihara et al. 2012, Deyle et al. 2013). Thus, we believe that methodological approaches like those we present in this study can enhance management and conservation of stocks by helping to identify appropriate ecosystem considerations to incorporate into stock assessments.

5 References

Anderson CNK, Hsieh C, Sandin SA, Hewitt R, Hollowed A, Beddington J, May RM, Sugihara G (2008) Why fishing magnifies fluctuations in fish abundance. Nature 452:835–839

Andrade HA, Garcia CAE (1999) Skipjack tuna fishery in relation to sea surface temperature off the southern Brazilian coast. Fish Oceanogr 8(4):245-254

Barile PJ (2013) Analysis of environmental factors affecting king mackerel landings along the east coast of Florida.SEDAR38-DW-07, North Charleston, SC 14 pp

Begg GA, Marteinsdottir G (2002) Environmental and stock effects on spatial distribution and abundance of mature cod Gadus morhua. Mar Ecol Prog Ser 229:245-262

Block B, Keen J, Castillo B, Dewar H and others (1997) Environmental preferences of yellowfin tuna (Thunnus albacares) at the northern extent of its range. Mar Biol 130(1):119-132

Casey KS, Brandon TB, Cornillon P, Evans R (2010) The past, present, and future of the AVHRR Pathfinder SST program. In: Oceanography from Space,p 273-287. Springer

Castonguay M, Rollet C, Fréchet A, Gagnon P, Gilbert D, Brêthes JC (1999) Distribution changes of Atlantic cod (Gadus morhua L.) in the northern Gulf of St. Lawrence in relation to an oceanic cooling ICES J Mar Sci 56:333-344

Damalas D, Maravelias CD, Katsanevakis S, Karageorgis AP, Papaconstantinou C (2010) Seasonal abundance of non-commercial demersal fish in the eastern Mediterranean Sea in relation to hydrographic and sediment characteristics. Estuar Coast Shelf Sci 89(1):107-118

De'ath G (2007) Boosted trees for ecological modeling and prediction. Ecology 88(1):243-251

Deyle ER, Fogarty M, Hsieh C -h., Kaufman L, MacCall AD, Munch SB, Perretti CT, Ye H, Sugihara G (2013) Predicting climate effects on Pacific sardine. Proc Natl Acad Sci 110:6430–6435

Deyle ER, Sugihara G (2011) Generalized Theorems for Nonlinear State Space Reconstruction. PLoS ONE 6:e18295

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77(4):802-813

Friedman JH (2002) Stochastic gradient boosting. Computational Statistics and Data Analysis 38:367-378

Froeschke J, Stunz GW, Wildhaber ML (2010) Environmental influences on the occurrence of coastal sharks in estuarine waters. Mar Ecol Prog Ser 407:279-292

Glaser SM, Fogarty MJ, Liu H, Altman I, Hsieh C-H, Kaufman L, MacCall AD, Rosenberg AA, Ye H, Sugihara G (2013) Complex dynamics may limit prediction in marine fisheries. Fish Fish:1–18

Glaser SM, Ye H, Maunder M, MacCall A, Fogarty M, Sugihara G (2011) Detecting and forecasting complex nonlinear dynamics in spatially structured catch-per-unit-effort time series for North Pacific albacore (Thunnus alalunga). Can J Fish Aquat Sci 68:400–412

Gratwicke B, Petrovic C, Speight M (2006) Fish distribution and ontogenetic habitat preferences in non-estuarine lagoons and adjacent reefs. Environ Biol Fishes 76(2-4):191-210

Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol Model 157(2-3):89-100 characteristic (ROC) curve. Radiology 143:29-36

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating

Hastie TJ, Tibshirani RJ (1990) Generalized Additive Models. Chapman & Hall, New York, NY

Hedger R, McKenzie E, Heath M, Wright P and others (2004) Analysis of the spatial distributions of mature cod (Gadus morhua) and haddock (Melanogrammus aeglefinus) abundance in the North Sea (1980-1999) using generalised additive models. Fish Res 70(1):17-25

Hilborn R, Walters C (1992) Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York, N.Y.

Hurrell JW (1995) Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation. Science 269:676–679

Hurrell JW, Deser C (2010) North Atlantic climate variability: The role of the North Atlantic Oscillation. J Mar Syst 79:231–244

Hutchings JA (1996) Spatial and temporal variation in the density of northern cod and a review of hypotheses for the stock's collapse. Can J Fish Aquat Sci 53(5):943-962

Johnson AF, Jenkins SR, Hiddink JG, Hinz H (2012) Linking temperate demersal fish species to habitat: scales, patterns and future directions. Fish Fish:no–no

Leathwick JR, Elith J, Chadderton WL, Rowe D, Hastie T (2008) Dispersal, disturbance and the contrasting biogeographies of New Zealand's diadromous and non-diadromous fish species. J Biogeogr 35(8):1481-1497

Leathwick JR, Elith J, Hastie T (2006) Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. Ecol Model 199(2):188-196

Liu H, Fogarty M, Glaser S, Altman I, Hsieh C, Kaufman L, Rosenberg A, Sugihara G (2012) Nonlinear dynamic features and co-predictability of the Georges Bank fish community. Mar Ecol Prog Ser 464:195–207

Liu H, Fogarty MJ, Hare JA, Hsieh C, Glaser SM, Ye H, Deyle E, Sugihara G (2014) Modeling dynamic interactions and coherence between marine zooplankton and fishes linked to environmental variability. J Mar Syst 131:120–129

MacRae PS, Cowan Jr JH (2010) Habitat Preferences of Spotted Seatrout, Cynoscion nebulosus, in Coastal Louisiana: A Step Towards Informing Spatial Management in Estuarine Ecosystems. Open Fish Science Journal 3:154-163

Maravelias CD (1999) Habitat selection and clustering of a pelagic fish: effects of topography and bathymetry on species dynamics. Can J Fish Aquat Sci 56(3):437-450

Maravelias CD, Reid DG, Swartzman G (2000) Modelling spatio-temporal effects of environment on Atlantic herring, Clupea harengus. Environ Biol Fishes 58:157-172

McKinney J, Hoffmayer E, Wu W, Fulford R, Hendon J (2012) Feeding habitat of the whale shark Rhincodon typus in the northern Gulf of Mexico determined using species distribution modelling. Mar Ecol Prog Ser 458:199-211

MSFCMA (2007) Magnuson-Stevens Fishery Conservation and Management Act. U.S. Department of Commerce NOaAA, National Marine Fisheries Service

Myers RA (1998) When Do Environment–recruitment Correlations Work? Rev Fish Biol Fish 8:285–305

Perretti CT, Munch SB, Sugihara G (2013) Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data. Proc Natl Acad Sci 110:5253–5257

Planque B, Fromentin J-M, Cury P, Drinkwater KF, Jennings S, Perry RI, Kifani S (2010) How does fishing alter marine populations and ecosystems sensitivity to climate? J Mar Syst 79:403–417

Quetglas A, Ordines F, Hidalgo M, Monserrat S, Ruiz S, Amores Á, Moranta J, Massutí E (2012) Synchronous combined effects of fishing and climate within a demersal community. ICES J Mar Sci J Cons:fss181

Ricker WE (1954) Stock and Recruitment. J Fish Res Board Can 11:559–623

Rotherham D, Underwood AJ, Chapman MG, Gray CA (2007) A strategy for developing scientific sampling tools for fishery-independent surveys of estuarine fish in New South Wales, Australia. ICES J Mar Sci J Cons 64:1512–1516

Schick R, Goldstein J, Lutcavage M (2004) Bluefin tuna (Thunnus thynnus) distribution in relation to sea surface temperature fronts in the Gulf of Maine (1994–96). Fish Oceanogr 13(4):225-238

SEDAR (2009) SEDAR 16: South Atlantic and Gulf of Mexico King Mackerel. North Charleston, SC 484 pp

Stoner AW, Manderson JP, Pessutti JP (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

Sugihara (1994) Nonlinear forecasting for the classification of natural time series. Philos Trans R Soc Lond Ser -Math Phys Eng Sci 348:477–495

Sugihara G, Allan W, Sobel D, Allan KD (1996) Nonlinear control of heart rate variability in human infants. Proc Natl Acad Sci 93:2608–2613

Sugihara G, May RM (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature 344:734–741

Sugihara G, May R, Ye H, Hsieh C -h., Deyle E, Fogarty M, Munch S (2012) Detecting Causality in Complex Ecosystems. Science 338:496–500

Takens F (1981) Detecting strange attractors in turbulence. In: Rand D, Young L-S (eds) Dynamical Systems and Turbulence, Warwick 1980. Springer Berlin Heidelberg, p 366–381

Valeiras J, Abad E (2006) Chapter 2.1.10.4 King Mackerel: Description of King Mackerel.8 pp

Wall CC, Muller-Karget FE, Roffer MA (2009) Linkages between environmental conditions and recreational king mackerel (Scomberomorus cavalla) catch off west-central Florida. Fish Oceanogr 18(3):185-199

Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3:385–397

Wintle BA, Elith J, Potts JM (2005) Fauna habitat modelling and mapping: a review and case study in the Lower Hunter Central Coast region of NSW. Austral Ecol 30(7):719-738

Wood SN (2006) Generalized Additive Models: an Introduction with R. Chapman & Hall/CRC, Boca Raton, FL

Table 1. Relative importance of explanatory variables identified as influential in describing the occurrence of king mackerel within boosted regression tree analysis.

Variable	Relative Influence (%)
target	93.66
area	2.35
away	1.06
year	0.83
hour fished	0.63
month	0.49
numgear	0.40
SST	0.37
effort	0.12
crew	0.06
gear	0.04

Table 2. Improvement in predictive skill ($\Delta\rho$) when climate variables are included in multivariate SSR forecasts of king mackerel relative biomass and landings. The multivariate predictive ability, ρ multiSSR , is the correlation between observed and predicted time series (relative biomass or landings). Permutation testing was used to test whether calculate forecast improvement ($\Delta\rho$) was greater than chance alone, using 0.05 as a cutoff p-value.

Index	ρ multiSSR	$\Delta\rho$	p-value
North Atlantic oscillation			
CPUE	0.526	0.035	0.098
Landings	0.613	0.021	0.115
Sea surface temperature			
CPUE	0.483	-0.011	0.089
Landings	0.589	0.004	0.145

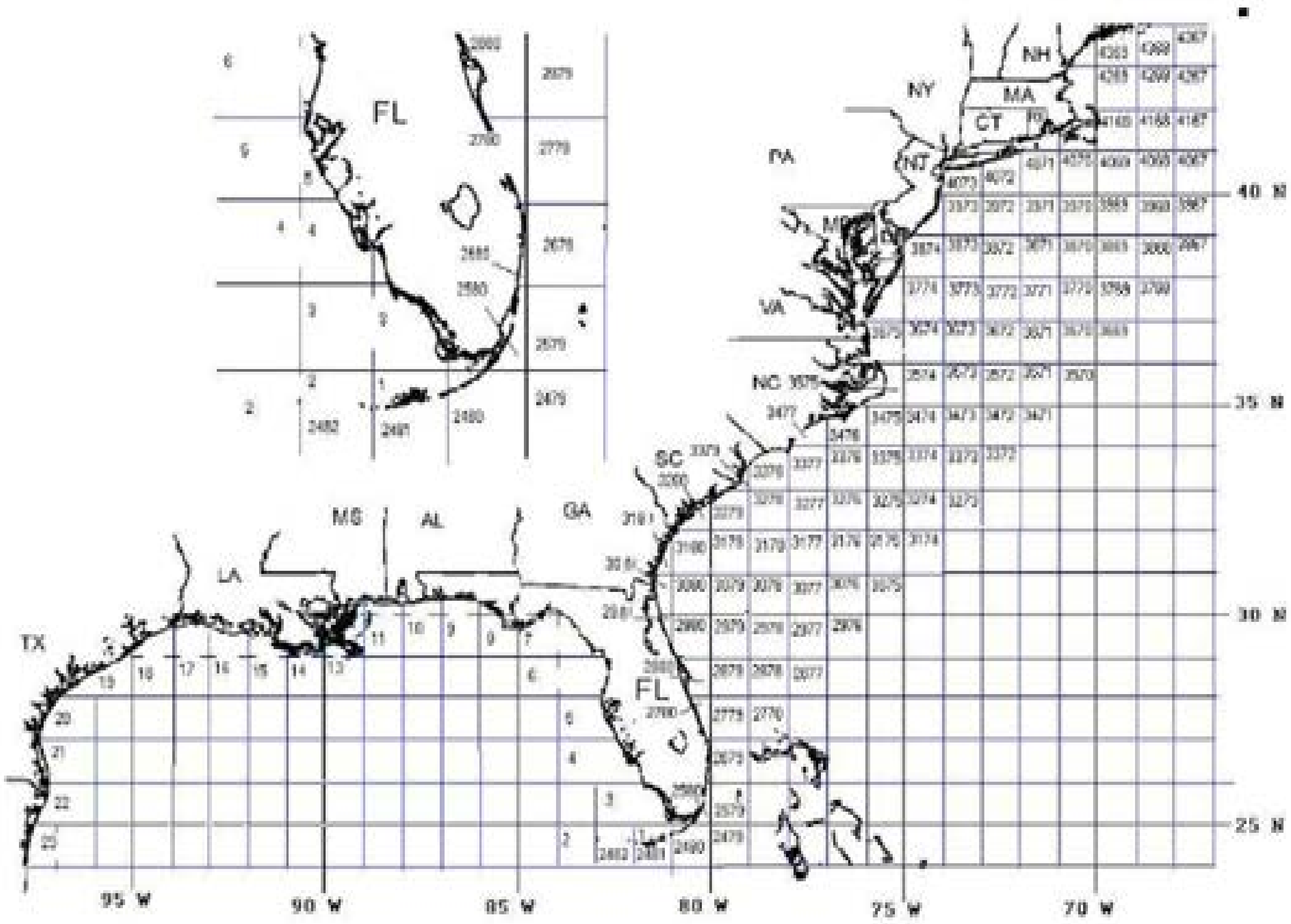

Figure 1**.** Statistical areas used to classify area fished in the NOAA commercial logbook database.

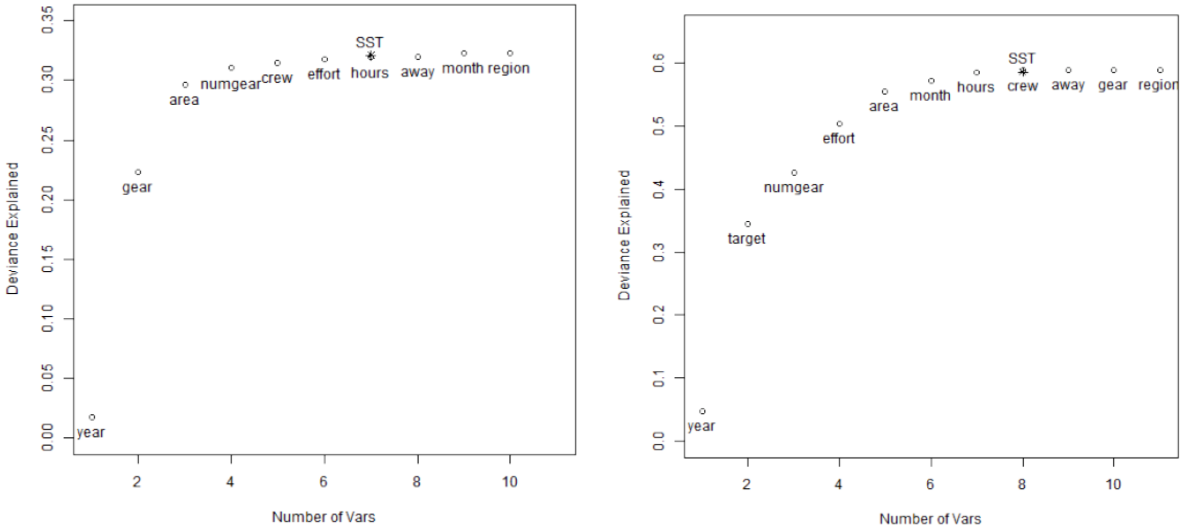

Figure 2. Scree plots representing the fit of the presence-absence (left) and positive CPUE (right) GAMs to the entire king mackerel CPUE dataset (all regions). The star represents where the SST variable was tested and the associated difference in percent deviance explained.

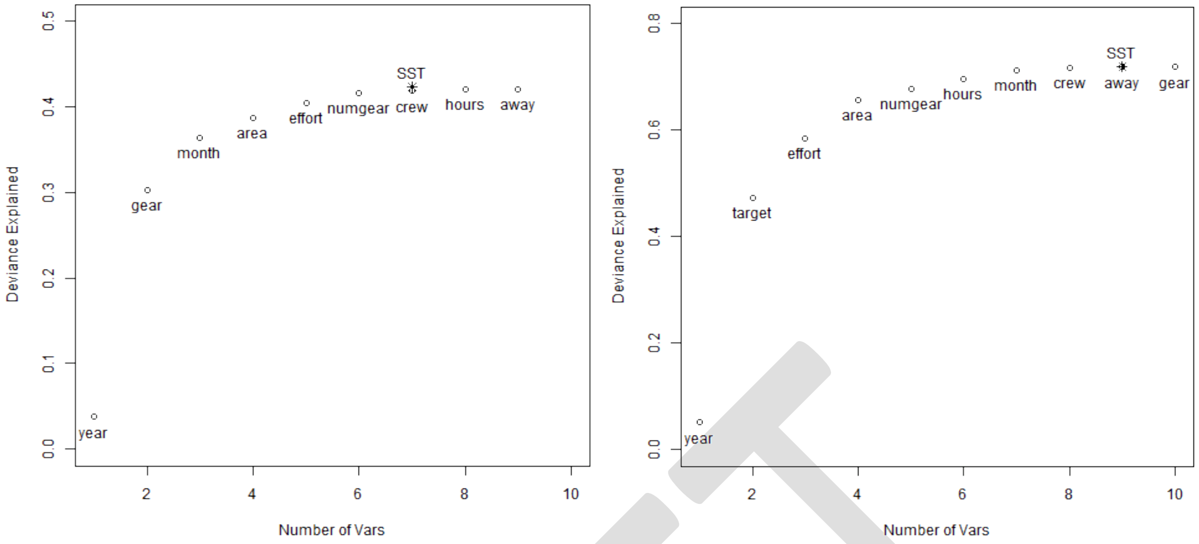

Figure 3. Scree plots representing the fit of the presence-absence (left) and positive CPUE (right) GAMs to the king mackerel CPUE dataset subset for the Gulf of Mexico. The star represents where the SST variable was tested and the associated difference in percent deviance explained.

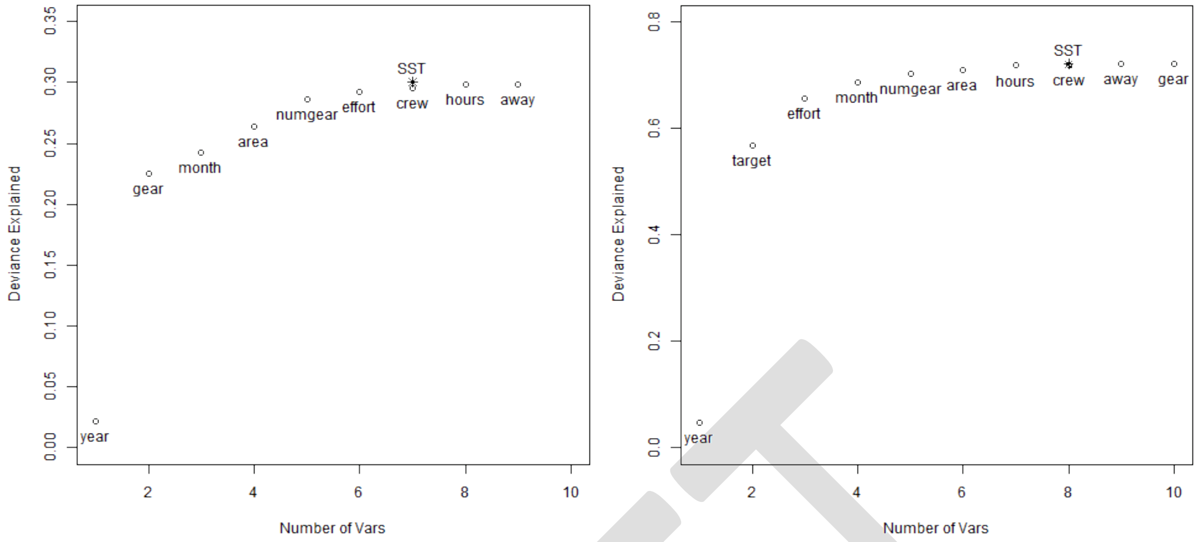

Figure 4. Scree plots representing the fit of the presence-absence (left) and positive CPUE (right) GAMs to the king mackerel CPUE dataset subset for the South Atlantic. The star represents where the SST variable was tested and the associated difference in percent deviance explained.

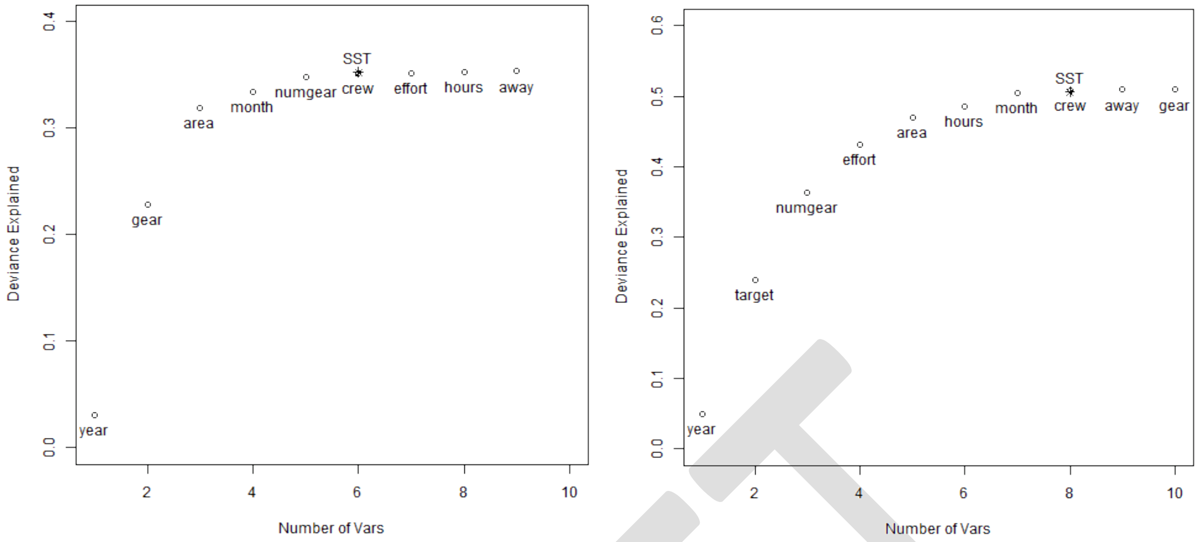

Figure 5**.** Scree plots representing the fit of the presence-absence (left) and positive CPUE (right) GAMs to the king mackerel CPUE dataset subset for the Mixing Zone. The star represents where the SST variable was tested and the associated difference in percent deviance explained.

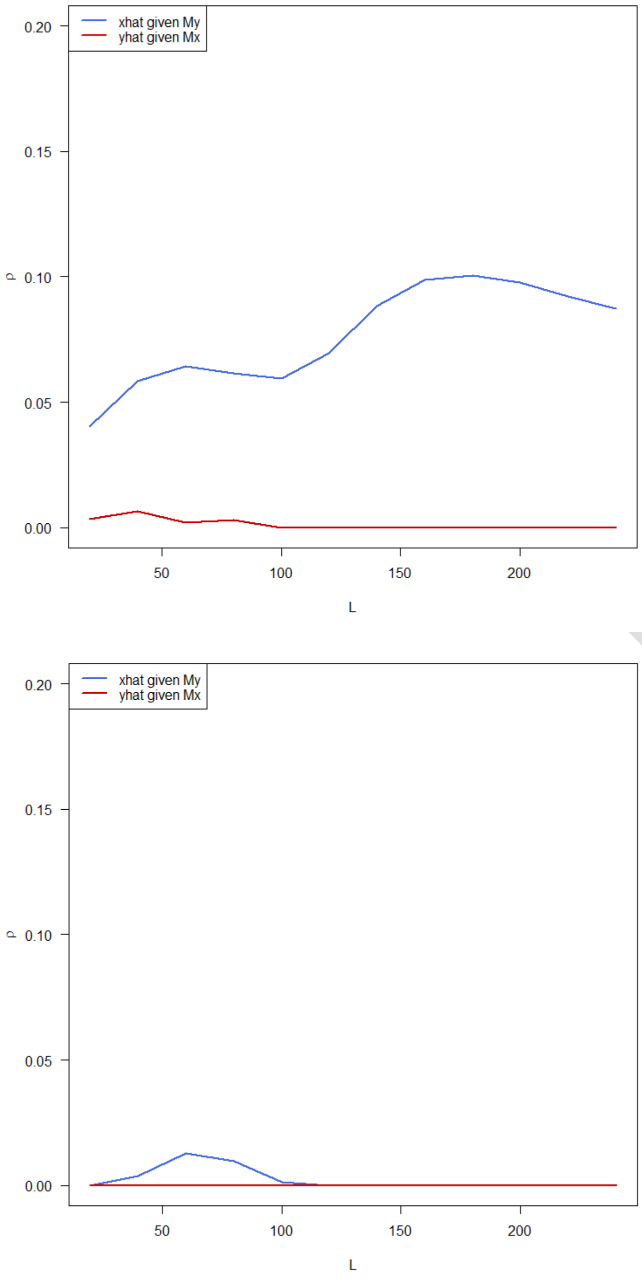

Figure 6. Convergent cross-mapping of monthly north Atlantic oscillation anomaly (NAO) and king mackerel time series. Top panel: blue line shows how prediction skill (ρ) of NAO (*xhat*) improves when state-space reconstruction (*My*) of CPUE is based on sequentially longer time series length (L). Bottom panel: blue line shows no improvement in NAO prediction as the length of the landing time series (*My*) is increased. As expected, red lines in both panels show that past climate events (*Mx*) contain no information about king mackerel time series (*yhat*).

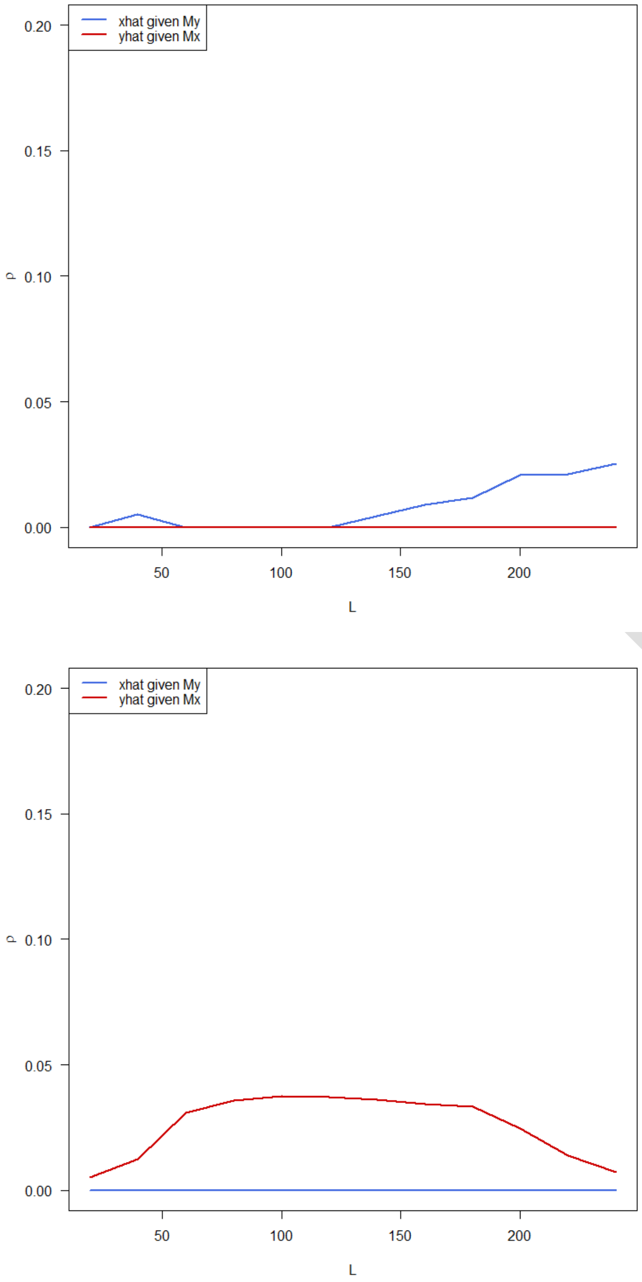

Figure 7. Convergent cross-mapping of monthly sea surface temperature anomaly (SST) and king mackerel time series. Top panel: blue line shows how prediction skill (ρ) of SST (*xhat*) improves slightly when state-space reconstruction (*My*) of CPUE is based on sequentially longer time series length (L). Bottom panel: blue line shows no improvement in SST prediction as the length of the landing time series (*My*) is increased. As expected, red lines in both panels show that past climate events (*Mx*) contain no information about king mackerel time series (y*hat*).