# Nonlinear dynamic features and co-predictability of the Georges Bank fish community

**Hui Liu[1,6,\*], Michael J. Fogarty[1], Sarah M. Glaser[2,7], Irit Altman[3],**
**Chih-hao Hsieh[4], Les Kaufman[5], Andrew A. Rosenberg[3], George Sugihara[2]**

[1]NOAA/NMFS, Northeast Fisheries Science Center, Woods Hole, Massachusetts 02543, USA

[2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA

[3]University of New Hampshire, Ocean Process Analysis Lab, Durham, New Hampshire 03824, USA

[4]Institute of Oceanography and Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei 106, Taiwan

[5]Boston University Marine Program, 5 Cummington Street, Boston, Massachusetts 02215, USA

[6]*Present address:* Department of Marine Biology, Texas A&M University at Galveston, Texas 77553, USA

[7]*Present address:* College of William and Mary, Williamsburg, Virginia 23185, USA

ABSTRACT: We examined evidence for nonlinear dynamics in fishery-independent survey data for an assemblage of 26 fish species on Georges Bank spanning the period 1963 to 2008. We used nonlinear time series analysis to determine (1) the presence of nonlinear dynamics in fish populations on Georges Bank, (2) the minimum number of dimensions required to effectively describe system dynamics, (3) the strength of patterns of co-predictability among all possible pairs of fish species, and (4) identification of groups of species characterized by similar dynamics. Here, nonlinear behavior refers to non-equilibrium dynamics, including chaos. The population trajectories of all 26 species exhibited strong density-dependent feedback as indicated by a Partial Rate Correlation Function analysis. Significant evidence of complex dynamical behavior was found for approximately 1 in 5 species. Low dimensionality for many of the individual series was identified, suggesting that for a given level of predictability, this system can be represented by a relatively small number of critically important ecological variables. Further, we found high levels of co-predictability among pairwise combinations of individual species. We identified 4 major species groups sharing similar dynamic features on the basis of patterns of co-predictability, and explored potential mechanisms for interpreting the groupings in terms of trophic interactions and life history characteristics.

KEY WORDS: Nonlinear system dynamics · Co-predictability · Fish population · Community

## INTRODUCTION

Nonlinear population processes, ecological interactions, environmental forcing, and anthropogenic impacts can individually and collectively generate dramatic patterns of variability in marine ecosystems (Steele & Henderson 1984, Dixon et al. 1999, Hsieh et al. 2005, Anderson et al. 2008). Recognition of the consequences of this broad spectrum of critical processes affecting marine systems, and the possibility of synergistic interactions among them, has led to calls for dedicated research into the structure, function, and resilience of exploited marine systems (Mangel & Levin 2005, Murawski et al. 2010). An extensive catalogue of multispecies and ecosystem modeling approaches has now been applied in fishery management (Plagányi 2007). These models incorporate varying levels of mechanistic detail in specifying key elements of ecosystem structure and function. In many instances, the choice of model structure and complexity reflects a tradeoff between realism and accuracy due to the potential for in-

creased uncertainty in parameter estimation in more complex models (Yodzis 1998, Fulton et al. 2003). Uncertainty in specifying the critical structural elements to be included in ecosystem models, however, remains a dominant source of risk and uncertainty for management. An alternative and complementary approach is to use a class of phenomenological models fit to time series data using nonparametric methods without the requirement of specifying a particular model structure.

Here, we explored the utility of a flexible class of nonlinear time series models to assess the complexity of these systems (Sugihara & May 1990, Sugihara 1994) and to develop short-term forecasts for a multispecies assemblage of fish species. Application of these nonparametric models avoids the dilemma of specifying a particular model form while allowing considerable flexibility in representing nonlinear ecological processes. This general approach has been successfully applied to marine ecosystem dynamics (Dixon et al. 1999, Hsieh et al. 2005, Anderson et al. 2008, Glaser et al. 2011). These simpler models can often outperform more complicated mechanistic models in forecast skill. We applied the concept of co-predictability to increase forecast skill associated with relatively short time series data, and to examine possible associations between fish species in the Georges Bank system. For nonlinear systems, correlations between system components may be weak or undetectable even if they are dynamically coupled because parametric correlation analysis will only detect linear associations. Co-predictability provides an alternative quantitative measure of model forecast skill and is particularly useful for nonlinear systems. We examined pairwise combinations of fish species to assess whether the observed trajectory of 1 species can be used to model and predict the trajectory of a second species. High co-predictability between species indicates significant dynamic similarity.

Georges Bank is a highly productive submarine plateau located off the coast of New England (USA). Its distinctive bathymetric and oceanographic characteristics result in the development of seasonally important anti-cyclonic circulation with important implications for the productivity of the system related both to water mass retention and water column mixing (Townsend et al. 2006). Georges Bank has been the focus of commercial fishing activity off the southern New England coast for several centuries (German 1987). Intensive harvesting pressure has caused profound impacts on the composition, abundance, and production of the Georges Bank fish community,

with enduring effects on the overall variability of production potential and harvestable biomass in this system (Fogarty & Murawski 1998). In this study, we applied nonlinear time series analysis to fisheries-independent biomass estimates for economically and ecologically important members of the fish community on Georges Bank. We hypothesized that sympatric marine fish populations tend to be co-predictable, but not necessarily correlated, because they may share common life history and dynamical characteristics, may interact with each other, and may respond to common underlying forcing in a multiplicative rather than an intrinsically linear way. We note that not all co-predictable species are ecologically interacting, and therefore this is not a test for direct interactions. For example, species which do not interact, but are driven by the same external forcing, may exhibit similar dynamics and thus high co-predictability. Here, our objectives included detecting the dynamical features of this system (dimensionality and nonlinearity), assessing increases in forecast skill achieved by taking advantage of co-predictability, and categorizing the number of dynamically equivalent ecological components in the system as measured by co-predictability.

## METHODS

### Survey biomass index data

The Northeast Fisheries Science Center (NEFSC) has conducted fall bottom trawl surveys off the Northeast US Continental Shelf since 1963 (Grosslein 1969, Azarovitz 1981). Briefly, a stratified random design is employed using a proportional allocation scheme. Vessel and gear characteristics have been standardized throughout the survey period. All species caught are weighed and measured. Biological samples for age determination, food habits, maturity, and pathology are taken according to a length-stratified cluster sampling design. Additional physical measurements are collected at each station.

In this study, we conducted time series analysis of biomass indices for 26 fish species (Table 1), collected during the NEFSC fall bottom trawl survey in the period of 1963 to 2008 on Georges Bank (Fig. 1). Relative abundance data (weight) were corrected for vessel and/or gear differences in the time series on a per-tow basis as required. The stratified mean (expanded) weight-per-tow was then calculated for each species of each survey. Prior to nonlinear analyses, fish abundance data were first-differenced ($\Delta x = x_t - x_{t-1}$,

Table 1. Fish species from the Northeast Fisheries Science Center fall bottom survey on Georges Bank. E: exploited; U: unexploited

| Group | Common name | Scientific name | Family | Code | Exploitation status |
|---|---|---|---|---|---|
| Gadids | Atlantic cod | *Gadus morhua* | Gadidae | AC | E |
| | Haddock | *Melanogrammus aeglefinus* | Gadidae | HD | E |
| | Pollock | *Pollachius virens* | Gadidae | PO | E |
| | Silver hake | *Merluccius bilinearis* | Gadidae | SH | E |
| | Red hake | *Urophycis chuss* | Gadidae | RH | E |
| | White hake | *Urophycis tenuis* | Gadidae | WH | E |
| Flatfishes | Yellowtail flounder | *Limanda ferruginea* | Paralichthyidae | YF | E |
| | Winter flounder | *Pseudopleuronectes americanus* | Pleuronectidae | WT | E |
| | American plaice | *Hippoglossoides platessoides* | Pleuronectidae | AP | E |
| | Fourspot flounder | *Paralichtheys oblongus* | Paralichthyidae | FF | E |
| | Witch flounder | *Glyptocephalus cynoglossus* | Pleuronectidae | WF | E |
| | Windowpane flounder | *Scophthalmus aquosus* | Scophthalmidae | WP | E |
| Other ground | Ocean pout | *Macrozoarces americanus* | Zoarcidae | OP | U |
| | Goosefish | *Lophius americanus* | Lophiidae | GO | E |
| | Longhorn sculpin | *Myoxocephalus octodecemspinosus* | Cottidae | LS | U |
| | Cunner | *Tautogolabrus adspersus* | Labridae | CU | U |
| | Acadian redfish | *Sebastes fasciatus* | Sebastidae | AR | U |
| Skates | Little skate | *Leucoraja erinacea* | Rajidae | LK | E |
| | Winter skate | *Leucoraja ocellata* | Rajidae | WS | E |
| | Thorny skate | *Amblyraja radiata* | Rajidae | TS | E |
| | Barndoor skate | *Dipturus laevis* | Rajidae | BS | U |
| | Smooth skate | *Malacoraja senta* | Rajidae | SS | E |
| Pelagics | Atlantic mackerel | *Scomber scombrus* | Scombridae | AM | E |
| | Butterfish | *Peprilus triacanthus* | Stromateidae | BF | E |
| | Atlantic herring | *Clupea harengus* | Clupeidae | AH | E |
| Dogfish | Spiny dogfish | *Squalus acanthias* | Squalidae | DF | E |

where $t$ is time) and then standardized (mean = 0, standard deviation, SD = 1) to remove secular trends and to normalize the time series data. To avoid spurious dynamics caused by over-differencing, we conducted



Fig. 1. Georges Bank eco-region, with survey samples from the shaded strata included in the analysis

a Kwiatkowski-Phillips-Schmidt-Shin stationary test for each individual time series (Kwiatkowski et al. 1992) available in an R package (tseries). Prior to first-differencing, 16 of 26 series were non-stationary. However, after first-differencing, all time series were stationary at the $p < 0.05$ level. Compared to non-differenced time series, first-differencing the 10 stationary time series did not cause significant changes in our assessment of nonlinearity and dimensionality. Therefore, further nonlinear analyses were based on the first-differenced normalized data for all species. Fluctuations in time series data are shown in Fig. S1 in the supplement (www.int-res.com/articles/suppl/m464p195_supp.pdf), and no apparent autocorrelation patterns existed after first differencing and standardization (Fig. S2 in the supplement).

## Nonlinear analysis

Endogenous nonlinear processes of some form are a critical pre-requisite for complex dynamical behavior. Nonlinear population models have been extensively employed in ecological models to represent
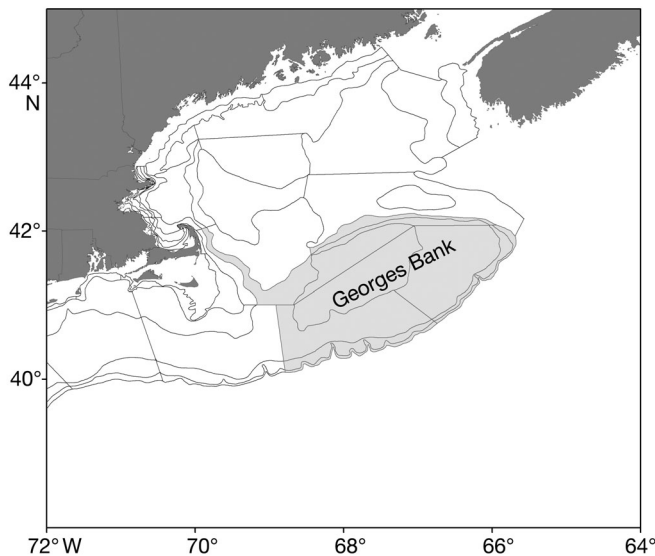
density-dependent dynamics. Many of these deter-ministic models have globally stable equilibrium points. A broad spectrum of discrete-time population models can also exhibit complex dynamics ranging from oscillatory dynamics to chaos (for a review, see Kot 2001). Systems of coupled differential equations of order ≥3 can also exhibit complex dynamics. In the following, we reserve the term nonlinear dynamics to refer specifically to dynamic complexity. As a pre-lude, we begin by testing for density-dependence and then turn to testing for evidence of complex dynamical behavior in the observed time series.

The analysis of many ecological time series has focused on estimation of Lyapunov exponents to test for complex dynamics (e.g. Ellner & Turchin 1995). However, with inevitable process noise, the meaning of such estimates becomes unclear (Sugihara 1994). Here we applied an alternative approach that com-pares forecast skill of linear and nonlinear models to judge the relative importance of nonlinear dynamics in the systems examined (Sugihara & May 1990, Sugi-hara 1994). The linear case is taken as the null model. The methods involve a nonparametric modeling scheme and are well established and validated in many areas of science, but because they are not yet broadly known in marine ecology, we provide a de-tailed account of the methodology[1].

## Density dependence

Density-dependent feedback can give rise to non-linear dynamics in biological populations. We tested for evidence of density-dependent feedback in the individual time series data using a Partial Rate Correlation Function (PRCF) analysis (Berryman & Turchin 2001, 2003, Royer & Fromentin 2006). We calculated the PRCF for the first-order finite differ-ence ($\log N_t - \log N_{t-1}$) against the lagged biomass indexes ($\log N_{t-1}$, $\log N_{t-2}$, …etc.) of time series for 26 fish populations on Georges Bank. The significance of PRCF was tested by Bartlett's criterion $\pm 2/\sqrt{n}$, where $n$ is the length of the time series at lag 1 (Box & Jenkins 1976). Because density dependence (broadly defined) is a necessary but insufficient condition for complex dynamics, we next test for non-equilibrium dynamics explicitly.

--------

[1]Readers seeking to apply the method are encouraged to contact the Sugihara Lab (email: gsugihara@ucsd.edu) for guidance and access to software. An introduction to the basic simplex method is available at http://deepeco.ucsd.edu/simplex/

## Nonlinear dynamics

To test for evidence of complex nonlinear dynamics in the Georges Bank fish population estimates, we build on 2 critical elements: (1) identification of the dimensionality of the (unobserved) underlying eco-logical system based on past values of a restricted number of measured variables, and (2) application of a nonlinear time series analysis to estimate the pres-ence of nonlinearity in the empirical time series. A key component of the approach throughout is the use of out-of-sample forecast skill as the determinant of the 'best' model. For the case of a single time series, the general approach involves splitting the series and building the model using one portion of the data (re-ferred to as the library set) while reserving the with-held portion to test the model (the prediction set). We used this approach to estimate dimensionality and classify nonlinearity in the individual time series. We also employed a related approach by which models built from all observations of 1 species were used to predict a second species. This is referred to as co-prediction and was done for all pairwise combina-tions of species.

To set the stage, consider a generalized represen-tation of an ecological variable (e.g. population size or biomass) as a function of its past values:

$$X_t = g\{X_{t-L}, X_{t-2L}, X_{t-3L}, ...X_{t-DL}\} \qquad (1)$$

where $X$ is the variable of interest, $g\{\cdot\}$ is an unspeci-fied function, $L$ is a specified lag period, and $D$ is the number of past values needed to capture the dynam-ics of the system. Difference equations of this general form have been extensively used in ecology to model population processes. May (1976) provided a cata-logue of models for the case $L = D = 1$, and Turchin (1990) explored the use of second-order difference equations ($L = 2$; $D = 1$) to assess delayed density dependence in single-species population models. If the model is intrinsically linear, we have a simple autoregressive model of the type used extensively for forecasting (Box & Jenkins 1976).

## Lagged coordinate embeddings

The modeling approach we used is nonparametric in the sense that there is no specified functional form. Rather, our approach involves interpreting the time series as having been produced from dynamics that reflect motion on an attractor (see Fig. 2). We address the inverse problem and reconstruct the attractor (and the coordinate space it occupies) from the time
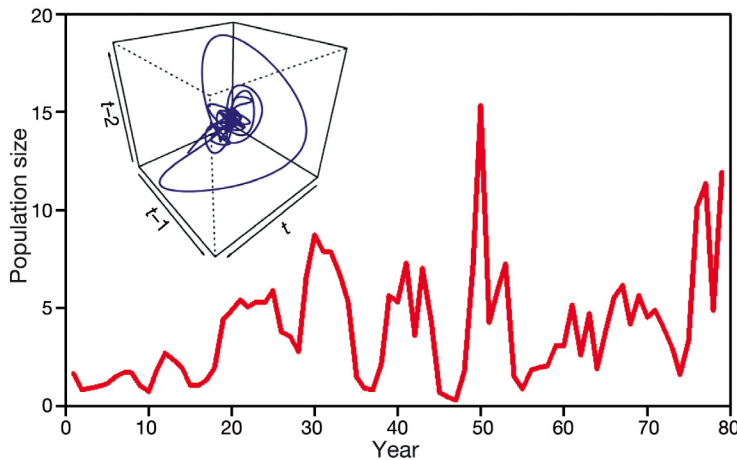
Fig. 2. Example of a hypothetical time series and corresponding 3-dimensional embedding for this series

series. Our task is to use an observed time series to attempt a reconstruction of key elements of the system as a whole. Although the true dimensionality of the system is unknown and potentially very high, it can be shown that an understanding of the number of key dimensions affecting the variable(s) of interest can be determined from the histories recorded in the time series. In practice, the effective dimensionality required of the system to achieve significant predictability is often relatively low (Sugihara & May 1990, Sugihara 1994).

For a hypothetical 3-species system with no exogenous forcing factors, the dimensionality of the system would be 3 and we construct a 3-dimensional phase space with each of the species arrayed along one of the axes. We can further track the trajectory (temporal evolution) of each species in this phase space. If there is a globally stable equilibrium point for each of the 3 species, the trajectories will converge to a set of 3 points. These points are attractors in the phase space. More complicated dynamics (e.g. limit cycles, chaotic processes) will trace different geometries (attractors) in the phase space (see below).

In a remarkable theorem, Takens (1981) showed that a time-delayed coordinate system such as Eq. (1) contains information not only about the measured variable itself but also about a set of unobserved but related variables (in our context, other species, environmental variables, etc.). In effect, the broader dimensions of the system can be captured by a more limited set of observations on one or more variables if the system is principally driven by deterministic processes. Takens' key insight was that the basic properties of the phase space for the overall system

can be reconstructed by examining the time-delayed structure of just 1 of the variables. Deyle & Sugihara (2011) provided a generalization of Takens' approach to take advantage of the case where multiple time series from the same system are available for analysis.

The resulting geometric shape in phase space is directly related to the true underlying attractor, giving a 'shadow' attractor. An example of translating a time series into a phase space representation by plotting the time-delayed coordinates is shown in Fig. 2. From Eq. (1), the dimensionality of the system, $D$, can be estimated by the embedding dimension ($E$). To do this, we transform the time series into a vector space comprising the time-delayed coordinate system. All methods of nonlinear time series analysis used to assess dynamic complexity employ time-delay coordinate systems (or embeddings) as the foundation for determining the dimensionality of the system (for a readable account, see Kaplan & Glass 1995).

The embedding dimension (or minimum dimensionality of the full system) is the minimum number of lagged variables needed to unfold the attractor so that trajectories do not cross. Higher dimensional embeddings can also fully resolve the dynamics, but they introduce greater uncertainty due to contamination of nearby points in the higher dimensional embeddings with points whose earlier coordinates are close, but whose recent coordinates are distant (Sugihara & May 1990).

In the following, we first used a simple nearest-neighbor forecasting algorithm to provide an estimate of the minimum dimensionality of the system ($E$). We sequentially identified points in the phase space and tracked these and their nearest neighbors in a forward projection. These nearby points represented similar past events on the shadow attractor in the phase space. For a chaotic system, these points will diverge exponentially with time; however, for short forecast periods, they will retain a connection with the recent past. We defined nearest neighbors in Euclidean phase space using the conventional representation:

$$\min \|X_t - X_i\| \qquad (2)$$

where $\|y\|$ is a suitable vector norm for variable $y$. Our identification of nearest neighbors depends critically on the embedding dimension. The set of $E + 1$ nearest neighbors is the minimum number of points needed to uniquely define the location of $X_t$ in $E$-dimensional space. These points define a simplex.

For a 2-dimensional system, we required a minimum of 3 points to define the vertices of the simplex (Fig. 3). Higher dimensional embeddings result in a more complex simplex structure. For a 3-dimensional system, the simplex is a tetrahedron (pyramid) and so on. We sequentially increased $E$ to a maximum of 10 and determined the 'best' embedding dimension based on the forecast skill (described below) of each model. Limitations on the length of the available series precluded examination of $E$ greater than 10. To generate a forecast, we used an exponentially weighted average of the values given by the position of nearest neighbor vectors $y_i$ at time $T + 1$:

$$\hat{y}_t = \sum_{j=0}^{E} w_j x(j) \qquad (3)$$

where the weighting $w_j$ is determined by the location of $X_t$ relative to the nearest neighbor vectors at $T$ and applied at $T + 1$. Thus, the forecast is made based on the assumption that, when the 'best' $E$ is selected to reconstruct the attractor, the trajectory of the prediction vector at $T + 1$ can be calculated based on the trajectories of the nearest library vectors on the attractor at time $T$. Recall that here we only varied the number of dimensions to include in the analysis and then calculated the resulting mean forecast based on the average of the points defining the simplex. The choice of the embedding dimension producing the highest forecast skill was selected as 'best' for this simplex projection. We defined forecast skill, or $\rho$, as the correlation coefficient between the observed and model-generated values of each $\hat{y}_t$.
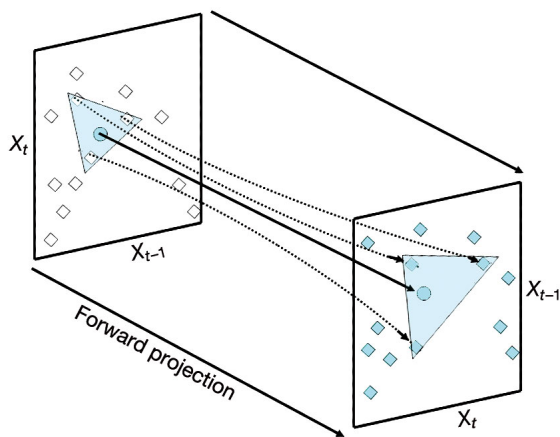


Fig. 3. Example of a simplex projection for a time series embedded in 2 dimensions ($D = 2$, where $D$ is the dimensionality of the system) showing the 1-step-ahead forecast for 1 point (dot) and its nearest neighbors from a library set. For this 2-dimensional system, 3 nearest neighbors define the simplex. The shaded triangle encloses the simplex. The projection for the simplex is shown by the shaded triangle on the right-hand side

We next examined evidence for the degree of nonlinearity in the series using a simple form of nonparametric regression based on locally weighted maps, using the best $E$ estimated by simplex projection. These locally weighted maps are known as S-maps (Sugihara 1994). Previous work (Sugihara 1994) indicated that the performance of the S-map procedure is superior to simplex projection for systems with short, noisy time series. As in simplex projection, the time series is decomposed into lagged-coordinate vectors and divided into library and prediction sets. In S-map, the location of the forecast for the prediction vector in embedded space is predicted by the location of all library vectors, not only the nearest neighbors (as in simplex projection), with heavy weightings given locally to library vectors nearest to prediction points. Again, the distances (in embedded space) from the library vectors to the prediction vectors are used to generate forecasts. Specifically, for the unknown forecast $\hat{y}_t$ of prediction vector $x_t$, given library vectors $\{x_i\}$ and setting $x_t(0) \equiv 1$ for the constant term in the solution of Eq. (4):

$$\hat{y}_t = \sum_{j=0}^{E} c_t(j) x_t(j) \qquad (4)$$

where $c$ is solved using singular value decomposition as:

$$b = Ac \qquad (5)$$

where $A$ contains elements for the library (predictor) set:

$$A_{ij} = w(\|x_i - x_t\|) x_i(j) \qquad (6)$$

$b$ contains the elements for the prediction set:

$$b_i = w(\|x_i - x_t\|) y_i \qquad (7)$$

The weighting function $w$ is defined as

$$w(d) = e^{-\theta d/\bar{d}} \qquad (8)$$

where $d$ is the Euclidean distance between any library vector ($x_i$) and the prediction vector $x_t$, $\bar{d}$ is the mean distance computed from all library vectors $\{x_i\}$ to the prediction vector, and $\theta$ is a nonlinear tuning parameter that gives variable weight to library vectors when generating the forecast. When $\theta = 0$, all library vectors are given equal weight and we have a linear model (essentially, a vector autoregressive, AR, model of order $E$). For a series of nonlinear models, $\theta > 0$, and the library vectors nearest the prediction vector are weighted according to Eq. (8). Thus, a nonlinear model gives greater weight to the neighborhood immediately surrounding the prediction vector. These nearby trajectories contain more similar recent information on the attractor, resulting in higher forecast skill if the system displays nonlinear

Table 2. Dynamic features of individual species, life history, diets, and fishing status of the functional groups in the Georges Bank fish community. P-values refer to a randomization test of the decrease in forecast mean absolute error (ΔMAE; see Fig. S5 in the supplement at www.int-res.com/articles/suppl/m464p195_supp.pdf). Values in **bold** are significant at p < 0.05. *E*: best embedding dimension; D: demersal; P: pelagic. Diet information from Smith & Link (2010). Blank cells indicate no data available

| Group | Common name | Nonlinear (p) | Best *E* | Spawning | Lifespan (yr) | Mature (yr) | Habitat | Diet |
|---|---|---|---|---|---|---|---|---|
| 1 | Barndoor skate | No (0.266) | 10 | | >10 | >10 | D | Piscivorous |
| | Winter skate | Yes (**0.04**) | 8 | | ~20 | 12 | D | Piscivorous |
| | Atlantic mackerel | No (0.062) | 10 | Spring/summer | 20 | 3 | P | Planktivorous |
| | Silver hake | No (0.208) | 4 | Spring/summer | 14 | 3 | P | Shrimp/Piscivorous |
| | Ocean pout | Yes (**0.038**) | 10 | | >16 | | D | Benthivorous |
| 2 | Witch flounder | Yes (**0.001**) | 10 | Spring/summer | 30 | 5–6 | D | Benthivorous |
| | Red hake | Yes (**0.01**) | 1 | Summer/fall | 8 | ~2 | D | Benthivorous |
| | Fourspot flounder | No (0.322) | 6 | | | | D | Piscivorous/Benthivorous |
| | Spiny dogfish | No (1) | 6 | | 35–40 | | P | Piscivorous/Benthivorous |
| | Yellowtail flounder | No (0.14) | 5 | Spring/summer | 6 | 2 | D | Benthivorous/Worm |
| | Smooth skate | No (0.094) | 8 | | | | D | Benthivorous |
| | White hake | No (1) | 6 | Spring/summer | 15 | | D | Piscivorous |
| | Atlantic herring | Yes (**0.034**) | 9 | Summer/fall | 9 | 4 | P | Planktivorous |
| 3 | Haddock | No (1) | 4 | Spring | 3–7 | 1–4 | D | Benthivorous/Echinoderms |
| | Winter flounder | No (0.362) | 6 | Winter/spring | 15–20 | | D | Benthivorous/Worm |
| | Cunner | No (0.134) | 1 | | | | D | Benthivorous |
| 4 | Goosefish | No (0.052) | 3 | Spring/summer | 12–14 | 5 | D | Piscivorous |
| | Pollock | No (0.232) | 3 | Winter | 1–14 | 3–6 | P | Piscivorous/Shrimp |
| | Longhorn sculpin | No (0.386) | 3 | | | | D | Benthivorous/Piscivorous |
| | American plaice | No (0.07) | 3 | Spring | 9 | 2–3 | D | Benthivorous/Echin |
| | Atlantic cod | No (0.12) | 2 | Winter/spring | ~5 yr | 2–4 | D | Benthivorous/Piscivorous |
| | Windowpane flounder | No (1) | 2 | Summer | 8 | 3–4 | D | Benthivorous |
| | Thorny skate | No (1) | 1 | | 16 | | D | Benthivorous/Piscivorous |
| | Little skate | No (1) | 2 | | 8–12 | | D | Benthivorous |
| | Acadian redfish | Yes (**0.004**) | 1 | Fall/winter | ~15 | | D | Piscivorous |
| | Butterfish | No (0.158) | 1 | Summer | 3 | 1 | P | Planktivorous |

behavior (Sugihara 1994). We applied a correlation coefficient (ρ) and mean absolute error (MAE) between predictions and observations to represent model performance, and the decrease in forecast error (ΔMAE) to measure nonlinearity. Nonlinearity was identified by a nonparametric randomization test for a significant decrease in model error (negative ΔMAE) as θ was tuned above 0 (Hsieh & Ohman 2006), with a cutoff p-value at 0.05 significance level.

We first analyzed all 26 species to quantify embedding dimension and identify the presence of nonlinear dynamics in time series of individual species (Table 2). We identified dynamically equivalent species groupings through co-prediction analysis. As noted above, we defined co-prediction as the ability for the trajectory of 1 species to predict the trajectory of a second species. Statistically significant forecast skill, ρ, between 2 species suggested they contained similar dynamics. We thus operationally define dynamically equivalent species as 2 species that contained significant co-prediction skill, or $\rho_{co}$, calculated as cross-species co-prediction from 1 species (predictor) to another species (predictee).

We calculated pairwise co-predictions for time series data of 26 species, resulting in 676 cross-species combinations. The 676 measures of $\rho_{co}$ from simplex projection and S-map models were expressed as two 26 × 26 predictor–predictee matrices. A third 26 × 26 matrix of pairwise linear correlation coefficients was calculated using the normalized and first differenced data to compare conventional correlation approaches with the nonlinear forecasting methods.

To further explore the pairwise patterns of copredictability among 26 fish populations, we calculated the dissimilarity matrix measured as the Euclidean distance among pairwise predictions based on predictor–predictee $\rho_{co}$. We then conducted a hierarchical agglomerative cluster analysis to identify dynamically equivalent groups of predictor fish species on the basis of the dissimilarity matrix. For assessing the uncertainty in the hierarchical cluster analysis, we calculated p-values associated with each hierarchical cluster through a multi-scale bootstrap re-sampling technique (Shimodaira 2004) available in an R package (pvclust). In this method, the p-value of a cluster is a value between 0 and 1,

indicating how strongly the data support the cluster. Briefly, the pvclust package performs hierarchical cluster analysis via function hclust and automatically computes p-values for all clusters contained in the clustering of original data. It provides 2 types of p-values: approximately unbiased (au) and bootstrap probability. We used the 'au' p-value calculation because it provided a better approximation of an unbiased p-value than alternative methods.

## RESULTS

### Density dependence

The PRCF plots displayed a highly significant first-order negative density-dependent feedback for all 26 time series plus weak and relatively higher-order negative feedback for some fish species on Georges Bank (Fig. S3 in the supplement at www.int-res.com/articles/suppl/m464p195_supp.pdf). This set the stage for further examination of whether the underlying dynamics are stable or unstable through nonlinear time series analysis.

### Nonlinear time series analysis

The results of single-species analysis by simplex projection models disclosed a pattern of relatively low dimensionality for the Georges Bank fish community (Table 2, Fig. S4 in the supplement). Forecast skill ($\rho$) varied considerably among species. The S-map forecast models demonstrated evident nonlinear behaviors in 6 of 26 time series of fisheries-independent relative biomass indices based on a nonparametric randomization test of $\Delta$MAE at a 0.05 level (Fig. S5).

### Nonlinear co-predictability

Many pairwise combinations of results from the simplex projection (Fig. 4) and S-map models (Fig. 5) showed significant co-predictability. S-map model predictions tended to be stronger than simplex projection models for producing co-prediction measures
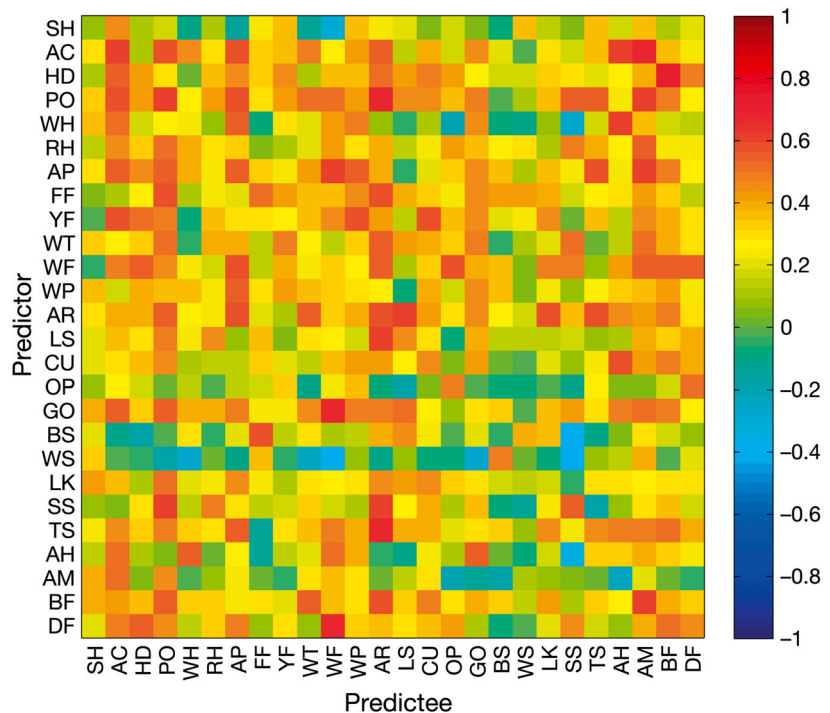


Fig. 4. Co-prediction coefficient matrix from a simplex projection model represents the dynamical connection among the Georges Bank fish community. Species codes listed in Table 1

($\rho_{co}$). In contrast, the ordinary correlation analysis revealed a different pattern of species interactions among the Georges Bank fish community. Most species appeared to have only moderate positive or negative association (Fig. 6). To interpret these results, we should note the symmetrical structure of the correlation coefficient matrix in contrast to the nonsymmetrical matrices of the co-prediction measures from simplex projection and S-map models. The co-prediction matrices calculated from the nonlinear models should not be considered equivalent to the correlation coefficients from the ordinary correlation analysis under linear assumptions of classical statistics. In particular, one should not expect identical $\rho$ for each pairwise component when swapping the direction of predictor and predictee. This is because the singular forecast for the predictee species is constructed from the multiple trajectories of the $E + 1$ nearest neighbor vectors from the predictor species; symmetry is structurally implausible.

### Dynamically equivalent units

The cluster analysis revealed 4 main dynamically equivalent groups (Table 2) with high probability
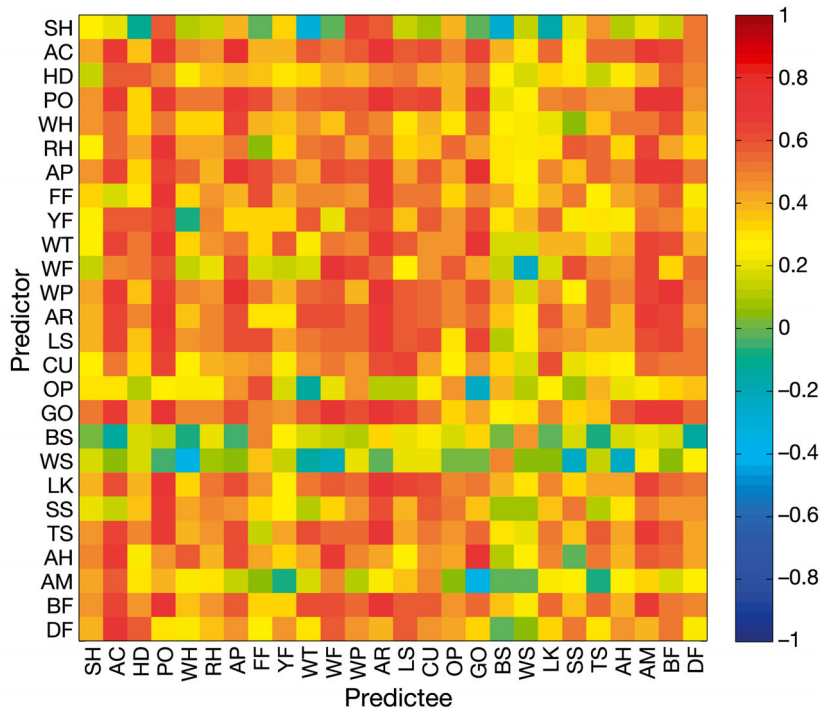
Fig. 5. Co-prediction coefficients matrix from an S-map prediction model represents the dynamical connection among the Georges Bank fish community. Species codes listed in Table 1



Fig. 6. Correlation coefficient matrix of the Georges Bank fish community. Species codes listed in Table 1

based on the dissimilarity of co-predictability ($\rho_{co}$; Fig. 7). The first group included barndoor skate, winter skate, silver hake, ocean pout, and Atlantic mackerel at a probability of 96%. The second group contained haddock, winter flounder, and cunner with 89% probability. Witch flounder, red hake, fourspot flounder, spiny dogfish, smooth skate, white hake, and Atlantic herring comprised the third group. All other species were grouped in the fourth cluster. These groupings differ from classifications one would expect based solely on taxonomic affinity. Here, groupings emerged based on similarity in their co-predictability (dynamic equivalence) for other species, and thus, these groupings reflect a combination of similarities in life history and dynamical characteristics, possible biotic interactions, and common natural and anthropogenic forcing.

# DISCUSSION

## Density-dependent dynamics

Understanding dynamic features of biological populations (i.e. nonlinearity and dimensionality) is a critical step for modeling and forecasting complex systems (Sugihara & May 1990, Sugihara 1994). Analysis of ecological time series provides clues for understanding the underlying causal structure driving the observed dynamics. PRCF is a useful tool for detecting density-dependent structure in ecological time series (Berryman & Turchin 2001). In general, negative feedback of the first order usually represents intraspecific interactions, while negative feedback of the second order often results from interactions between species, particularly between consumers and their resources (Royama 1992, Berryman 1999, Turchin 2003). In the present study, a significant first-order negative feedback was found for all 26
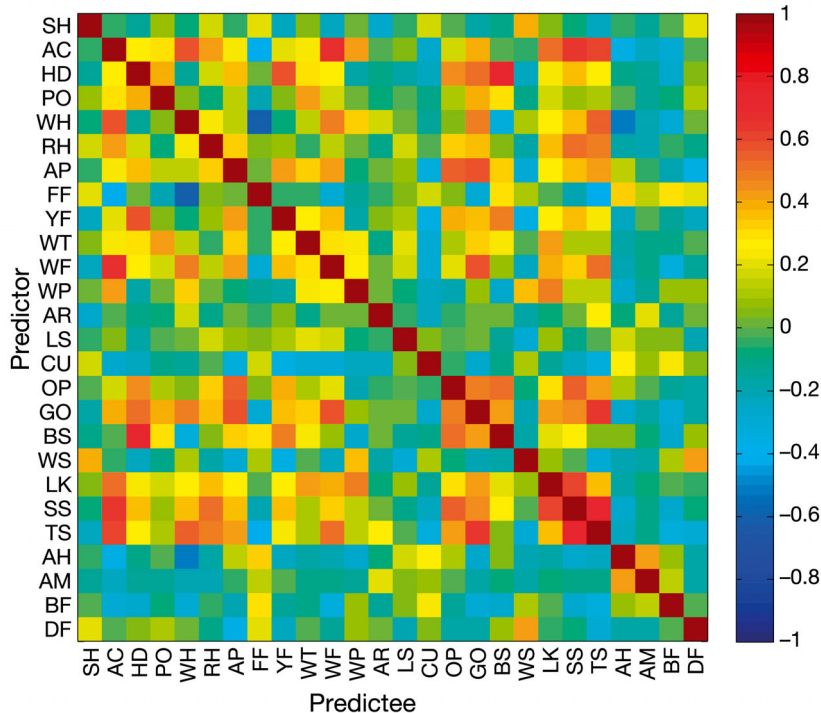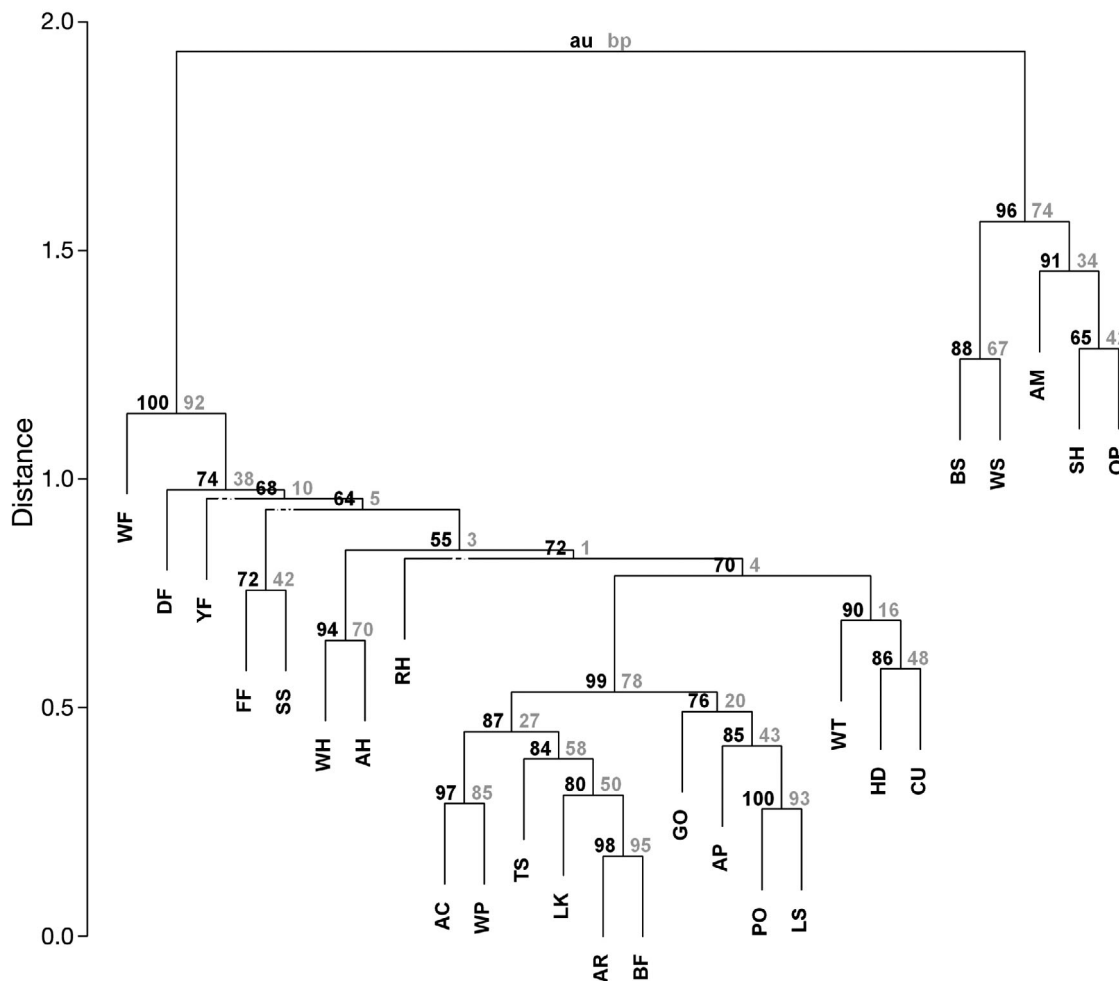
Fig. 7. Dendrogram for finding the dynamically equivalent groups among 26 fish populations at Georges Bank based on their co-predictability computed from the S-map models. Species codes listed in Table 1. au: approximately unbiased probability; bp: bootstrap probability

species. This suggests that intraspecific density-dependent regulation is common among the species considered in this analysis.

Although density-dependent feedback was found in the fish community on Georges Bank, we should note that the driving processes behind the observations remain hidden. In general, negative feedback operates through competition for environmental resources (food or living space), higher death rates due to starvation, cannibalism, predation, or disease. We found complex nonlinear dynamics in about one-fifth of the fish populations on Georges Bank. We note that observational error and stochasticity can obscure underlying nonlinear dynamics but that the methods used in this study have proven relatively robust in the face of noise (e.g. Sugihara & May 1990).

## Nonlinear dynamics of the Georges Bank fish community

Our findings clearly indicate that the dynamics of this fish assemblage are generally characterized by low dimensionality, which is consistent with findings from studies of time series data of various marine fisheries populations in other ecosystems from the Great Barrier Reef, Australia, to the west Pacific coast of North America and the Mediterranean (Dixon et al. 1999, Hsieh et al. 2005, Royer & Fromentin 2006, Anderson et al. 2008, Glaser et al. 2011). Low dimensionality implies that the system can be further modeled with few variables to achieve relatively high forecast skill and explored to illuminate the key processes driving the system using a multivariate embedding technique (see Dixon et al. 1999). The

S-map modeling approach provides a useful way to identify nonlinearity in a stationary time series by the difference in forecast skill achieved with a nonlinear model versus a linear model. For complex systems with higher dimensionality and strong nonlinear signals, S-map usually produces higher forecast skill (Sugihara 1994). This might explain why the S-map approach tended to be more successful than the simplex projection model for our time series data. These data are characterized by low dimensionality and apparent nonlinear signatures.

We note that temporal and spatial scales are critical issues in terms of model forecast skill (Sugihara 1994, Hsieh et al. 2008, Glaser et al. 2011). In the present study, the moderate time series length (~45 yr) and relatively broad spatial coverage of the study area might lessen the possibility of detecting nonlinear behaviors, as spatial averaging tends to mask local nonlinearity and accentuate linear autoregressive signals. Extending length through concatenating time series data (Hsieh et al. 2008), or analysis of spatially structured data (Glaser et al. 2011), can improve forecast skill and the detection of nonlinear dynamics by effectively increasing sample size. In any event, the spatial and temporal scales employed in this analysis are well suited for management applications on Georges Bank. Given the temporal length and spatial resolution of this study, results from the simplex projection and S-map forecast models in this study showed promise for identifying nonlinearity and complexity within the fish community on Georges Bank (see Figs. S4 & S5 in the supplement at www.int-res.com/articles/suppl/m464p195_supp.pdf). Successful development of short-term forecasts can be quite useful in management for quota setting and other applications.

## Co-predictability and dynamically equivalent units

By definition, co-predictability indicates that an explanatory variable or a collection of explanatory variables can be used to improve out-of-sample prediction of the trajectory of another dynamically dependent variable. In this study, we examined dynamical equivalence among species within the fish community on Georges Bank. Co-predictability as a quantitative measure of model forecast skill is particularly useful for nonlinear systems where normal correlations between system components may be weak or undetectable even if they are dynamically coupled. In stable linear systems, correlation should be sufficient to identify links between coupled components. However, in nonlinear systems, correlation analysis and linear models are not sufficient to detect dynamic linkages. Co-prediction, by design, is able to accommodate both linear and nonlinear linkages between components and can identify underlying dynamical connections that correlation or other linear approaches may fail to detect. Thus, co-predictability can be used for finding relevant sets of system components to objectively characterize marine ecosystems that share common dynamical features. Co-prediction alone, however, is not sufficient to identify causal connections among species.

We identified 4 dynamically equivalent groups of fish species based on a cluster analysis using the co-predictability results. In general, the grouped species tend to share more or less similar dynamical features (i.e. dimensionality and nonlinearity), and life history traits and feeding habits (Table 2). Recent hypotheses have proposed that ecologically equivalent species sharing analogous topological attributes in a system space can be predicted from the time series of other equivalent species using the simplex projection and S-map techniques (Hsieh et al. 2008). Conversely, significant co-predictability in the time series data of the Georges Bank fish community derived from these nonlinear analyses presumably indicates that these fish species likely come from dynamically equivalent groups or ones responding to common underlying mechanisms such as environmental and fishing-related forcing.

Ecological dynamics of communities reflect more than the sum of their parts (Mangel & Levin 2005). To interpret the groupings of dynamically equivalent units classified based on co-predictability, we note that our findings are not explicitly based on either trophodynamic or habitat analyses. However, trophodynamic mechanisms may underlie some of the observed patterns of co-predictability in our analyses because species in this study with relatively high co-predictability often showed similar feeding patterns or reflected predator–prey interactions (Table 2). For example, dynamically grouping winter skate and barndoor skate together is consistent with our knowledge based on life history characteristics, spatial distributions, and diet. Both are benthic invertebrate feeders and piscivores with relatively consistent diet structure over spatial and seasonal time scale scales (Smith & Link 2010). Overlap in feeding habits for species within the dynamically equivalent groups also exists for other groups. Many gadid species included in this study are known to prey extensively on pelagic species such as herring, mackerel, and butterfish (Grosslein et al. 1980), while mackerel and

herring can also be voracious predators on early life stages of groundfishes (Pepin et al. 1987). Therefore, behind commonly observed negative correlations among pelagics, gadids, and flatfishes hides a quite different picture of dynamic connection among them (see Figs. 4–6). We note that most fish species in the Georges Bank system tend to be generalist carnivores with moderate dietary overlap, and many species also show strong ontogenetic shifts in diet (Bowman & Michael 1984, Garrison & Link 2000a,b).

A generalized competitive release due to the removal of historically dominant groundfishes by fishing activities may have caused increases in historically underexploited species on Georges Bank (Fogarty & Murawski 1998). We can expect that species subject to similar external forcing mechanisms exhibit common features in their dynamics. A diffuse competitive release may have explained the observed patterns of negative correlations but positive co-predictions in the Georges Bank fish community. For example, co-predictability among species of skates, flatfishes, and gadids may be related to competitive release following the extensive removal of flatfishes and gadids in the 1970s. Skates are considered potential competitors of flatfishes and gadids and share relatively high levels of diet similarity with haddock, yellowtail flounder, and winter flounder (Grosslein et al. 1980). We note that, in general, diet overlap within the fish community of the Northeast Continental Shelf Large Marine Ecosystem (NES LME) is moderate (Smith & Link 2010), and if competitive release is important, it likely operates in a diffuse manner at a broad system level.

Intrinsic population processes, environmental variability, and harvesting are 3 major causes of nonlinear fluctuations and dynamic complexity in marine fisheries populations. In this study, we have identified negative density-dependent feedback as a potentially important mechanism; however, we did not explicitly address the effects of environmental variation and fishing pressure in understanding system dynamics. Fishing pressure can magnify the nonlinear fluctuations in fish abundance (Anderson et al. 2008), while climate variability can cause alterations in the spatial distributions of fish populations (Nye et al. 2009). These 2 processes tend to act synergistically on the variability of fish population dynamics (Hare et al. 2010, Planque et al. 2010).

### Relevance to management

Traditional fisheries management has focused on a single-species approach designed to maximize the long-term sustainable yield of target species. As progress toward implementing ecosystem-based fisheries management unfolds, development of tools to evaluate aspects of ecosystem structure becomes increasingly important. The initial motivation of this work was to assess possible approaches to complement and augment traditional ecosystem modeling approaches. A lack of understanding of ecosystem structure and function limits our ability to develop successful ecosystem-based management plans and to conduct effective ecosystem comparisons in the oceans (Megrey et al. 2009, Murawski et al. 2010). This study represents an effort to overcome these obstacles, particularly the issue of uncertainty of the appropriate functional forms to be included in multispecies/ecosystem models. Our results indicate that the effective dimensionality of the fish community in the Georges Bank system is relatively low. Further, evidence of co-predictability between fish populations derived in these nonlinear time series models points to important inter-relationships or common underlying forcing mechanisms influencing species within the fish assemblage. The traditional approaches to defining species assemblages in marine ecosystems (e.g. by size, taxonomic affinity, etc.) can be expanded to include consideration of the underlying dynamical features of the system components.

### LITERATURE CITED

Anderson CNK, Hsieh CH, Sandin SA, Hewitt R and others (2008) Why fishing magnifies fluctuations in fish abundance. Nature 452:835–839

Azarovitz TR (1981) A brief historical review of the Woods Hole Laboratory trawl survey time series. Can Spec Publ Fish Aquat Sci 58:62–67

Berryman A (1999) Principles of population dynamics and their application. Stanley Thornes, Cheltenham

Berryman A, Turchin P (2001) Identifying the density-dependent structure underlying ecological time series. Oikos 92:265–270

Bowman RE, Michael WL (1984) Food of seventeen species of Northwest Atlantic fish. NOAA Tech Memo NMFS-F/NEC–28. Northeast Fisheries Science Center, Woods Hole, MA

Box GEP, Jenkins GM (1976) Time series analysis: forecasting and control. Holden Day, Oakland, CA

Deyle ER, Sugihara G (2011) Generalized theorems for nonlinear state space reconstruction. PLoS ONE 6:e18295

Dixon PA, Milicich MJ, Sugihara G (1999) Episodic fluctuations in larval supply. Science 283:1528–1530

Ellner SP, Turchin P (1995) Chaos in a noisy world: new methods and evidence from time-series analysis. Am Nat 145:343–375

Fogarty MJ, Murawski SA (1998) Large-scale distribution and the structure of marine systems: fishery impacts on Georges Bank. Ecol Appl 8(Suppl):s6–s22

Fulton EA, Smith ADM, Johnson CR (2003) Effect of complexity on marine ecosystem models. Mar Ecol Prog Ser 253:1–16

Garrison LP, Link JS (2000a) Fishing effects on spatial distribution and trophic guild structure of the fish community in the Georges Bank region. ICES J Mar Sci 57:723–730

Garrison LP, Link JS (2000b) Dietary guild structure of the fish community in the Northeast United States continental shelf ecosystem. Mar Ecol Prog Ser 202:231–240

German AW (1987) History of the early fisheries: 1720–1930. In: Backus RH (ed) Georges Bank. MIT Press, Cambridge, MA, p 409–424

Glaser SM, Ye H, Maunder M, MacCall A, Fogarty MJ, 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

Grosslein MD (1969) Groundfish survey program of BFC Woods Hole. Commer Fish Rev 31:22–35

Grosslein MD, Langton R, Sissenwine MP (1980) Recent fluctuations in pelagic fish stocks in the Northwest Atlantic, Georges Bank region in relation to species interactions. Rapp PV Reùn Cons Int Explor Mer 177:374–405

Hare JA, Alexander MA, Fogarty MJ, Williams EH, Scott JD (2010) Forecasting the dynamics of coastal fishery species using a coupled climate-population model. Ecol Appl 20:452–464

Hsieh CH, Ohman MD (2006) Biological responses to environmental forcing: the linear tracking window hypothesis. Ecology 87:1932–1938

Hsieh CH, Glaser SM, Lucas AJ, Sugihara G (2005) Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. Nature 435:336–340

Hsieh CH, Anderson CNK, Sugihara G (2008) Extending nonlinear analysis to short ecological time series. Am Nat 171:71–80

Kaplan D, Glass L (1995) Understanding nonlinear dynamics. Springer-Verlag, New York, NY

Kot M (2001) Elements of mathematical ecology. Cambridge University Press, Cambridge

Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. J Econom 54:159–178

Mangel M, Levin P (2005) Regime, phase and paradigm shifts: making community ecology the basic science for fisheries. Philos Trans R Soc Lond B Biol Sci 360:95–105

May RM (1976) Simple mathematical models with very complicated dynamics. Nature 261:459–467

Megrey BA, Link JS, Hunt GL Jr, Moksness E (2009) Comparative marine ecosystem analysis: applications, opportunities, and lessons learned. Prog Oceanogr 81: 2–9

Murawski SA, Steele JH, Taylor P, Fogarty MJ, Sissenwine MP, Ford M, Suchman C (2010) Why compare marine ecosystems? ICES J Mar Sci 67:1–9

Nye JA, Link JS, Hare JA, Overholtz WJ (2009) Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Mar Ecol Prog Ser 393:111–129

Pepin PS, Pearre S Jr, Koslow JA (1987) Predation on larval fish by Atlantic mackerel, *Scomber scombrus*, with a comparison of predation by zooplankton. Can J Fish Aquat Sci 44:2012–2018

Plagányi ÉE (2007) Models for an ecosystem approach to fisheries. FAO Fish Tech Pap 477. FAO, Rome

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

Royama T (1992) Analytical population dynamics. Chapman & Hall, London

Royer F, Fromentin JM (2006) Recurrent and density-dependent patterns in long-term fluctuations of Atlantic bluefin tuna trap catches. Mar Ecol Prog Ser 319:237–249

Shimodaira H (2004) Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. Ann Stat 32:2616–2641

Smith BE, Link JS (2010) The trophic dynamics of 50 finfish and 2 squid species on the Northeast U.S. continental shelf. NOAA Tech Memo NMFS-NE-216. Northeast Fisheries Science Center, Woods Hole, MA

Steele JH, Henderson EW (1984) Modeling long-term fluctuations in fish stocks. Science 224:985–987

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

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

Takens F (1981) Detecting strange attractors in turbulence. Lect Notes Math 898:366–381

Townsend DW, Thomas AC, Mayer LM, Thomas M, Quinlan J (2006) Oceanography of the Northwest Atlantic Continental Shelf. In: Robinson AR, Brink KH (eds) The Sea, Vol 14A. Harvard University Press, Cambridge, MA, p 119–168

Turchin P (1990) Rarity of density dependence or population regulation with lags? Nature 344:660–663

Turchin P (2003) Complex population dynamics: a theoretical/empirical synthesis. Princeton University Press, Princeton, NJ

Yodzis P (1998) Local trophodynamics and the interaction of marine mammals and fisheries in the Benguela ecosystem. J Anim Ecol 67:635–658