

# Stability in marine fish communities

Richard J. Bell<sup>1,\*</sup>, Michael J. Fogarty<sup>2</sup>, Jeremy S. Collie<sup>1</sup>

<sup>1</sup>Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island 02882, USA

<sup>2</sup>Northeast Fisheries Science Center, National Marine Fisheries Service, Woods Hole, Massachusetts 02543, USA

**ABSTRACT:** Fisheries-independent surveys sample major portions of marine systems and can be considered proxies for many important ecosystem components. A striking feature of many independent trawl-survey datasets is the relative stability of total biomass over several decades, despite overfishing of several commercial species and marked changes in the relative abundance of different species. This relative stability suggests a pattern of species replacements related to compensatory dynamics in the total production of fish biomass. In contrast, sustained declines in overall biomass are evident in other ecosystems. We used trawl-survey data to analyze the mechanisms that regulate stability across 19 marine ecosystems around the world. A linearized, multi-species Gompertz model within a state space framework was fitted to the survey data to quantify interactions among diet-based functional groups and external drivers. The total biomass in 7 of the 19 systems lacked uni- or multidirectional trends and was stationary. The stable ecosystems were associated with lower levels of aggregate community fishing pressure enabling differential rates of growth, immigration and emigration among the different populations to be the important factors regulating community dynamics. Higher levels of fishing pressure tended to synchronize the community response to the external driver resulting in large changes in total biomass. Evidence for interactions among functional groups was relatively weak; however, density dependence at the functional group level suggested within-group compensation as an important stabilizing mechanism.

**KEY WORDS:** Marine fish communities · Stability · Trawl-survey data · State space models · Gompertz model

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

## INTRODUCTION

Population dynamics are driven by intra- and inter-specific interactions, environmental forcing and demographic stochasticity (Holling 1973, Quinn & Deriso 1999, Loreau & Mazancourt 2008). The variability over time is largely a function of the environmental forcing and demographic stochasticity as amplified or tempered by intra- and interspecific interactions (Ives et al. 1999, 2000, 2003, Gonzalez & Loreau 2009). Individual populations may respond dramatically to environmental forcing and the community impact as manifested by the sum of the individual populations can elicit a range of responses. Populations may respond to external forcing in a sim-

ilar manner, amplifying the signal and synchronizing the individual trends, leading to large changes in the community. In contrast, differential responses in growth, immigration and emigration among the suite of populations within a community can compensate for one another, thereby damping the strength of the external forcing. Differential rates result in asynchronous population trajectories and relatively stable higher-level ecosystem properties (total biomass, production, metabolism, etc.) (Holling 1973, Patten 1975, Ernest & Brown 2001, Ives & Cardinale 2004).

Ecosystems with higher biodiversity have been shown to be more stable over time and better able to provide consistent ecosystem goods and services (Tilman 1996, Lehman & Tilman 2000, Tilman et al.

\*Corresponding author: rich.bell@noaa.gov

2006, Worm et al. 2006). Functional redundancy (Walker 1992, Fischer et al. 2001) or functional complementarity (Rosenfeld 2002) in a community ensures that multiple species perform similar ecosystem functions, such as primary production, nutrient cycling or grazing. The loss of certain taxa due to perturbations is offset by gains in formerly competitively inferior or more stress-tolerant species (Ives et al. 2000). Species cannot occupy identical niches, but instead perform the same functional role across a gradient of conditions (dissolved oxygen range, temperature tolerances, disturbance levels) (Rosenfeld 2002). High-diversity communities are more likely to contain species or functional groups that cover the full gradient of potential conditions, thus making them more stable by being able to withstand a range of perturbations. In addition to biodiversity, the actual make up of the species or functional groups within a community can also be important. Dominant taxa have been found to be critical structuring components of an ecosystem and essential to maintaining ecosystem function and stability (O'Connor & Crowe 2005, Sasaki & Lauenroth 2011).

Fisheries-independent surveys sample major portions of marine systems and can be considered proxies for many important ecosystem components. A striking feature of many independent trawl-survey datasets is the relative stability of biomass over several decades, despite overfishing of several commercial species and marked changes in the relative abundance of different species (Fogarty & Murawski 1998, Frank et al. 2005). The relative stability suggests a pattern of species replacements related to compensatory dynamics in the total production of fish biomass. In contrast, sustained declines in overall biomass are evident in other systems. As large-scale environmental forcing (Anderson & Piatt 1999, Perry et al. 2005, Nye et al. 2009) and fishing (Worm et al. 2009) continue to reshape ecosystems, it is important to understand the mechanisms that determine temporal stability—whether the decline in certain functional groups will be offset by increases in others or whether all species groups will respond in synchrony to external drivers. In certain areas, long-term studies have found that particular functional groups have declined while others have increased (Frank et al. 2005, Collie et al. 2008). Our goal was to understand the mechanisms regulating the temporal stability of heavily fished marine ecosystems by comparing the response of functional groups across 19 fisheries-independent trawl surveys from around the world. We analyzed diet-based functional groups instead of individual species to facilitate broad-scale compar-

isons across a range of ecosystems and diversity levels.

Stability has a range of meanings in ecology and has been defined in a number of different ways (reviewed in Lehman & Tilman 2000). We first determined the stability of each ecosystem by examining the variability of total biomass in each trawl survey and the presence of uni- or multidirectional trends in total biomass. We then calculated a range of community metrics that are commonly linked with stability. We determined the biodiversity, dominance (Simpson's evenness), compensation and fishing pressure for each survey. Finally, we quantified the interactions among functional groups and potential external drivers with a vector autoregressive model within a state space framework (Ives et al. 2003, Holmes & Ward 2012).

Compensation is well founded in theoretical ecology and is grounded by some careful empirical studies (reviewed in Gonzalez & Loreau (2009)); however, in general, compensation has been difficult to detect with real world data (Houlahan et al. 2007). Based on the variance-ratio test using abundance and not biomass, species typically covaried together, suggesting that individual populations exhibited similar responses to broad-scale environmental factors and thus the community was largely driven by these external forcings. Compensation was evident between particular species and groups, but these interactions were not an overarching driver of community dynamics.

We went beyond the simple variance-ratio test and took a more in-depth look at the among functional group interactions within each system in combination with external drivers. A multispecies Gompertz population model was fit to the functional group biomass within each trawl survey (Ives et al. 2003). The model was fit as an empirical vector autoregressive model within a state space framework that explicitly took account of the time series nature of the data. The state space structure allows the incorporation of both process error and measurement error, which is important when working with trawl-survey data. Trawl-survey data contain high levels of measurement error, which varies in magnitude among surveys. The state space framework accounts for the measurement error across the different datasets allowing them to be compared. Ives et al. (2003) originally developed the model to examine community stability among freshwater plankton communities subject to perturbations. We expanded the interpretation of the parameters to examine compensation. The empirical model quan-

tified the intraspecific interactions within functional groups (density dependence), the interspecific interactions among functional groups and the external drivers. Two environmental factors and an ecosystem-level exploitation rate were included as external drivers. Traditional measures of compensation such as the variance-ratio test and the Loreau & Mazancourt (2008) synchrony index were also calculated for comparison.

## METHODS

Annual estimates for fish and invertebrate biomass were gathered from trawl-survey data around the globe (Table 1 & Fig. S1 in the Supplement at [www.int-res.com/articles/suppl/m504p221\\_supp.pdf](http://www.int-res.com/articles/suppl/m504p221_supp.pdf)). Each survey spanned at least 18 yr and had data for a minimum of 11 yr. The majority of surveys were much longer, with a mean survey length of 34 yr. For a detailed description of the datasets see Branch et al. (2010). The surveys generally recorded all the vertebrate and some invertebrate taxa that were caught and all data were considered in the analysis. Taxa that were not well represented by the data in the trawl surveys were removed with a persistence plot, which provided an objective way to remove taxa across the different surveys (Genner et al. 2004). The log of the mean biomass of each taxon over the survey was plotted against the number of years it was

caught in the survey. An inflection point was calculated from a third-order polynomial fit to the data. Taxa that were to the right of the inflection point were recorded in the survey over multiple years, had higher mean annual biomass and often a lower annual variance. Taxa that were present in more years than the inflection point were considered well represented by the trawl survey and were included in the analysis. Trawl-survey data were not corrected for catchabilities, because the catchabilities were not known for the majority of taxa in the majority of surveys.

Resource limitation is an important component of compensation. As diet is a fundamental way in which resources are partitioned in the marine environment, the taxa in each trawl survey were divided into 4 diet-derived functional groups based on previous studies: demersal benthivores, demersal piscivores, pelagic piscivores and pelagic planktivores (Greenstreet et al. 1997, Garrison & Link 2000, Heath 2005, Steele et al. 2007, Auster & Link 2009). Years with missing biomass values at the functional group level were filled in with the mean of the biomass of the year prior and year after the missing year. The Alaskan Fisheries Science Center Aleutian Islands and Gulf of Alaska surveys were conducted every other year or every third year. The missing years in these 2 surveys were filled in the same manner, such that roughly half the observed biomass values were interpolated (Table 1).

Table 1. The fisheries-independent surveys used in the analyses

Survey	Min	Max	No. years surveyed	Span (yr)	No. taxa	No. vertebrates	No. invertebrates
Eastern Bering Sea	1982	2011	30	30	86	86	0
Aleutian Islands	1980	2010	11	31	38	38	0
Gulf of Alaska Small Mesh	1972	2007	36	36	24	13	11
Gulf of Alaska	1984	2011	12	28	79	79	0
New Zealand Chatham Rise	1992	2011	20	20	78	77	1
South African West Coast	1985	2003	15	19	88	72	16
Celtic Sea	1987	2004	18	18	49	49	0
Irish Sea	1988	2007	20	20	19	19	0
North Sea	1980	2007	28	28	48	48	0
Mid-Atlantic Bight	1967	2007	41	41	40	37	3
Georges Bank	1963	2007	45	45	37	34	3
GSO-Whale Rock	1959	2011	53	53	25	18	7
GSO-Fox Island	1959	2011	53	53	25	18	7
Gulf of Maine	1963	2007	45	45	31	28	3
Scotian Shelf	1970	2006	36	37	49	48	1
Southern Gulf of St. Lawrence	1971	2007	37	37	45	45	0
Northern Gulf of St. Lawrence	1990	2007	18	18	10	10	0
St. Pierre Bank	1951	1995	40	45	23	23	0
Grand Banks	1952	1995	41	44	16	16	0

### Stability

Three measures of stability were calculated to assess variability and trends in the total biomass of each survey. The temporal stability index (TSI) relates the total biomass to the variance–covariance matrix (Lehman & Tilman 2000).

$$TSI = \frac{\sum_s \sum_t biomass_{t,s}}{\sqrt{\sum_s variance_s + \sum_s covariance_s}} \quad (1)$$

where  $t$  is time and  $s$  is species or taxon.

Higher levels of stability are the result of greater biomass per a given level of variance and covariance, lower variance of each species for a given level of total biomass or lower covariance among species for a given level of total biomass. Covariance among species decreases as the positive correlations among species decrease or negative correlations increase.

Three simple linear models were fit to the natural log of total biomass to determine whether the series exhibited uni- or multidirectional trends. The natural log of total biomass ( $X = \ln[\text{total biomass}]$ ) was fit with an intercept-only model,  $X = \beta_0$  where  $\beta_0$  is the  $y$  intercept, a linear regression (first-order polynomial)  $X = \beta_0 + \beta_1 \times \text{year}$  where  $\beta_1$  is the slope, and a second-order polynomial  $X = \beta_0 + \beta_1 \times \text{year} + \beta_2 \times \text{year}^2$  where  $\beta_2$  is the quadratic term. The 3 models were compared with Akaike's information criterion (AIC) to determine which model provided the best fit with the least amount of complexity, regardless of the significance of the higher-order terms. The natural log of the functional group biomass was analyzed in the same manner.

Each survey was then tested for stationarity as a final measure of stability. The natural log of total biomass was tested for a unit root with the augmented Dickey-Fuller test (ADF) (Brocklebank & Dickey 2003).

$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \delta_1 \Delta X_{t-1} + \dots + \delta_{p-1} \Delta X_{t-p+1} + \epsilon_t \quad (2)$$

In the equation,  $\Delta X_t$  is the change in total biomass from one year to the next ( $\Delta X_t = X_{t+1} - X_t$ ),  $\alpha$  is an intercept term (drift),  $\beta t$  is a linear trend term,  $p$  is the number of augmented lags and the  $\delta$  terms are differenced lag terms. The null hypothesis is that the non-differenced first lag term equals zero ( $\gamma = 0$ ). Surveys in which  $\gamma < 0$  rejected the null hypothesis that the time series has a unit root and were considered stationary. The log of total biomass was over-fit to the ADF equation (fit with too many lags). Lag terms and the trend term were removed based on backwards elimination and the overall model AIC (Brocklebank

& Dickey 2003). The appropriate model (correct number of lags, with or without trend term) was tested for a unit root in SAS.

### Community metrics

Compensation is often examined with the variance-ratio test (Klug et al. 2000, Ernest & Brown 2001, Houlihan et al. 2007, Solow & Duplisea 2007). It has been used in a number of studies, but it may not be the best measure of compensation. It is a relatively simple metric that is useful, but it defines compensation as a simple binary response (that is, either  $<1$ , compensation, or  $>1$ , not compensation). There is no gradient, which is more typical of natural systems (Loreau & Mazancourt 2008). It does not take into account the time series nature of the data. The value is rarely tested to determine if it is significantly  $<1$  (Solow & Duplisea 2007). The null hypothesis is that non-differentiated species competing for a similar resource would have negative covariance, which may not be true (Loreau & Mazancourt 2008) and there are limits on the magnitude of the negative covariance given that the variance of the sum must be  $\text{var}(\sum x_i) \geq 0$  (Brown et al. 2004).

The variance-ratio test and the Loreau & Mazancourt (2008) synchrony index were calculated as simple metrics of compensation for comparison (Schluter 1984, Klug et al. 2000, Loreau & Mazancourt 2008, Gonzalez & Loreau 2009). Both relate the variability of the whole to the variability of the individual time series. The variance-ratio test was calculated as the variance of the total biomass divided by the sum of the variances of the biomass of the individual functional groups, i.e.  $\text{var}(\sum x_i) / \sum \text{var}(x_i)$ . If the variance-ratio test is  $<1$ , the variability of the whole is considered less than the variability of the individual time series and compensation is an important driver of community dynamics. If the ratio is  $>1$ , compensation may still exist, but external factors are considered to be the major driver of community dynamics. Loreau & Mazancourt (2008) developed a similar metric to examine the synchrony among time series. It is calculated as the variance of the sum of the individual time series divided by the standard deviations of the individual time series summed together and then squared:

$$SI = \frac{\text{var}(\sum x_i)}{(\sum \sigma_i)^2} \quad (3)$$

The metric ranges from 0 (perfect asynchrony) to 1 (perfect synchrony) providing a gradient instead of a simple cut off at 1.

Biodiversity was calculated with the Shannon-Wiener diversity index

$$H = -\sum p_i \cdot \ln(p_i) \quad (4)$$

and the Simpson diversity index for each trawl survey:

$$D = 1 - \sum p_i^2 \quad (5)$$

Simpson's evenness was calculated to determine the distribution of biomass among taxa and thus the degree to which each system was dominated by individual taxa:

$$E = \frac{1}{\sum p_i^2 \cdot S} \quad (6)$$

In all 3 indices,  $p_i = s_i / \sum s_i$ , where  $s_i$  is the biomass for each species, and  $S$  is the number of taxa (Smith & Wilson 1996). The relationship between TSI and each community metric was tested with a linear regression (Sasaki & Lauenroth 2011).

The community fishing pressure index (*FPI*) for each trawl survey was compiled from stock assessments with the RAM Legacy database (Ricard et al. 2011). The biomass and catch of stocks that spanned more than one area were partitioned based on the area of the trawl survey divided by the total area covered by the stock. To compare across surveys, they were standardized according to the exploitation level that would result in maximum sustainable yield ( $U_{msy}$ ) for each stock. The community *FPI* was calculated as the sum of the annual catch ( $C$ ) of stock  $i$  over the annual biomass of stock  $i$  divided by the  $U_{msy}$  of stock  $i$  weighted by the biomass of stock  $i$ .  $U_{msy}$  for each stock was from a Schaefer model fit.

$$FPI_t = \sum_{i=1}^n \frac{\left(\frac{C_{it}}{B_{it}}\right) \cdot B_{it}}{U_{msy\ i}} \cdot \frac{B_{it}}{\sum_{i=1}^n B_{it}} \quad (7)$$

The mean of the community fishing pressure index ( $\overline{FPI}_t$ ) provided a starting point to compare exploitation across ecosystems.

### Gompertz population model

A linear, multispecies Gompertz population model was fit to the log biomass of the functional groups of each trawl survey within a state space framework (Fogarty & Brodziak 1994, Ives et al. 1999, 2003, Lindegren et al. 2009, Holmes & Ward 2012). The log biomass of each functional group was z-transformed (subtract the mean and divide by the standard deviation). In the nonlinear, single-species Gompertz model, the biomass at time  $t+1$  is equal to the biomass at time  $t$  times the exponential intrinsic growth

rate ( $a$ ), which is dampened by a density-dependent term ( $b$ ).

$$bio_{t+1} = bio_t \cdot e^{a+(b-1) \cdot \ln(bio_t)} \quad (8)$$

The equation can be made linear by taking the natural log of both sides. The term  $\ln(bio)$  is replaced with  $X$  for simplification. The model becomes an autoregressive lag-one model such that  $X$  at time  $t+1$  is a linear function of  $X$  at time  $t$ . Population dynamics occur that manifest themselves over a time delay of greater than one year (Fogarty & Brodziak 1994); however, fitting and interpreting higher-order models, particularly with short time series, can be problematic. First-order autoregressive models can approximate more complex models and provide interpretable results (Ives 1995, Ives et al. 2003). The lag-one autoregressive model provides a simplified tool to analyze complex communities.

$$X_{t+1} = a + bX_t + \varepsilon \quad (9)$$

External drivers such as fishing and the environment are added by including a parameter  $c$  and the appropriate time series  $U$ .

$$X_{t+1} = a + bX_t + cU_t + \varepsilon \quad (10)$$

The model can be expanded to a multispecies model by changing the parameters to matrices. The model then becomes a vector autoregressive model in which the density-dependent terms (diagonal of the  $B$  matrix), the among group or interspecific interaction terms (off diagonal of the  $B$  matrix) and the external drivers ( $C$  matrix) are estimated within a state space framework.

$$X_{t+1} = A + BX_t + CU_t + w_t, \text{ where } w_t \sim MVN(0, Q) \quad (11)$$

$$Y_t = ZX_t + f + v_t, \text{ where } v_t \sim MVN(0, R) \quad (12)$$

The  $B$  matrix represents the interactions among the 4 diet-based functional groups and the  $C$  matrix represents the interaction among the functional groups and the external drivers (2 levels of environmental factors and 1 aggregate exploitation rate). In the observation equation (Eq. 12),  $Y_t$  is the observed biomass,  $X_t$  is the true state,  $Z$  is the identity matrix and  $f$  is set to 0. The state space framework allows the incorporation of both process error and measurement error within the optimization, enabling better quantification of the true information within the trawl-survey datasets. The error terms ( $w_t$  and  $v_t$ ) are multivariate, normally distributed random variables with mean zero, process error variance  $Q$  and measurement error variance  $R$ . Models were fit via maximum likeli-

hood with a Kalman filter in the freeware package MARSS in R (Holmes & Ward 2012).

The parameters in the  $B$  matrix represent the potential direction and strength of the intra- and inter-interactions among functional groups. The diagonal of the  $B$  matrix is the density-dependent term for each functional group. Density dependence indicates self regulation at the functional group level. The off-diagonal terms could take the form of competition (negative–negative interaction between 2 functional groups), predator–prey dynamics (negative–positive interaction between 2 functional groups) and mutual benefit (positive–positive interaction between 2 functional groups). Compensation is typically considered the outcome of competition (Gonzalez & Loreau 2009); however, predator–prey interactions where the cycles are out of synchrony could also result in compensation. We examined all of the interactions in relation to stability.

Two levels of environmental proxies were used as potential external drivers to capture different time scales within the Gompertz population model. Annual change was captured with variations in sea surface temperature. The North Atlantic Oscillation (Hurrell & Deser 2009, NCAR 2011), Pacific Decadal Oscillation (JISAO 2011) and the Southern Oscillation Index (NWS 2011) were included to capture semi-decadal to decadal time scale changes in the respective ocean basins (see Text S1 in the Supplement at [www.int-res.com/articles/suppl/m504p221\\_supp.pdf](http://www.int-res.com/articles/suppl/m504p221_supp.pdf) for sources).

The aggregate exploitation rate for each trawl survey used in the Gompertz population model was compiled from stock assessments with the RAM Legacy database (Ricard et al. 2011). Relatively few stock assessments were available for many of the Canadian surveys and the aggregate exploitation rate was calculated from catch data and trawl-survey biomass (The Scotian Shelf, Pierre Bank [NAFO 3Ps], the Southern Grand Banks [NAFO 3NO] and the Northern and Southern Gulf of St. Lawrence [NAFO 4RS, 4T]).

The aggregate exploitation rate was the sum of the annual catch for all stocks available within the survey area over the sum of the annual biomass of the same stocks ( $i$  is taxon and  $t$  is year).

$$U_t = \frac{\sum_{i=1}^n C_{it}}{\sum_{i=1}^n B_{it}} \quad (13)$$

Only stocks with both catch and biomass from stock assessments were included. For the Canadian

surveys, the catch included all available catch (all species) from the North Atlantic Fisheries Organization (NAFO) and the biomass included the total biomass from the trawl surveys. The catch of herring was removed from total catch in the Southern Gulf of St. Lawrence (NAFO 4T). Eigenvalue stability analysis was conducted on the results of the Gompertz population model and simulated  $B$  matrices to determine the effect of different levels of interactions on overall stability. Simulations were run for varying levels of density dependence as well as varying numbers and intensity of interaction among functional groups (see the Supplement).

## RESULTS

The temporal variation in biomass exhibited a range of patterns across the trawl surveys (Fig. 1). The mean proportion of biomass in the different functional groups varied by system; however, the majority of the surveys were dominated by demersal organisms (benthivores and demersal piscivores) (Figs. 2A & S2). The Celtic Sea, New Zealand and Alaskan surveys had the highest proportion of pelagic biomass, while the Irish Sea and North American east coast had the highest proportion of demersals. The majority of the surveys on the east coast of North America did not have any taxa that were considered pelagic piscivores. All the data in this study were from demersal trawl surveys. Pelagic species are caught in demersal trawl surveys, but are typically underrepresented due to low catchability. While pelagic diversity tends to be low in the trawl datasets, the pelagic biomass can be quite high.

### Stability

The trawl surveys exhibited a range of values for the TSI with the Eastern Bering Sea ranking as the most stable and the Northern Gulf of St. Lawrence as the least stable. Ecosystems that were geographically close to each other and shared similar taxa did not exhibit similar stability properties (Fig. 2B).

The 3 linear models of increasing complexity were used to determine whether total biomass exhibited a uni- or multidirectional trend. The total biomass in the Gulf of Maine survey initially declined and then increased in later years. It was best fit with a second-order polynomial. The Aleutian Islands, Georges Bank, Gulf of Alaska, Pierre Bank, Southern Grand Banks, Whale Rock and Southern Gulf of St.

Lawrence surveys exhibited a positive or negative trend over the time series. The Eastern Bering Sea, Gulf of Alaska small mesh survey, Celtic Sea, Northern Gulf of St. Lawrence, Fox Island, North Sea, Irish Sea, Mid-Atlantic Bight, New Zealand Chatham

Rise, South African West Coast and Scotian Shelf surveys were best fit with an intercept-only model, suggesting that over the entire time series these surveys did not exhibit a statistical uni- or multidirectional trend. Visual inspection, however, suggested that the

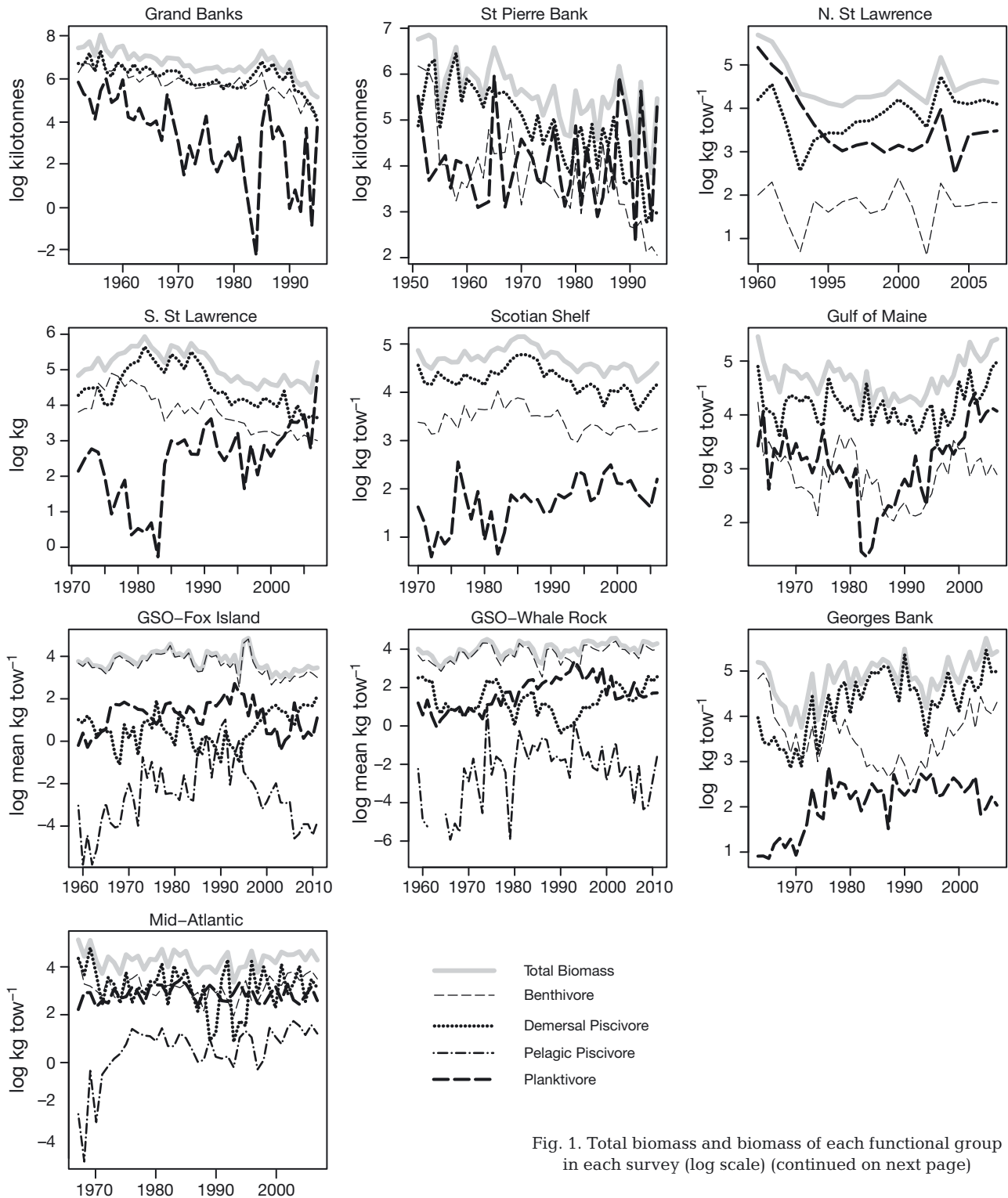


Fig. 1. Total biomass and biomass of each functional group in each survey (log scale) (continued on next page)

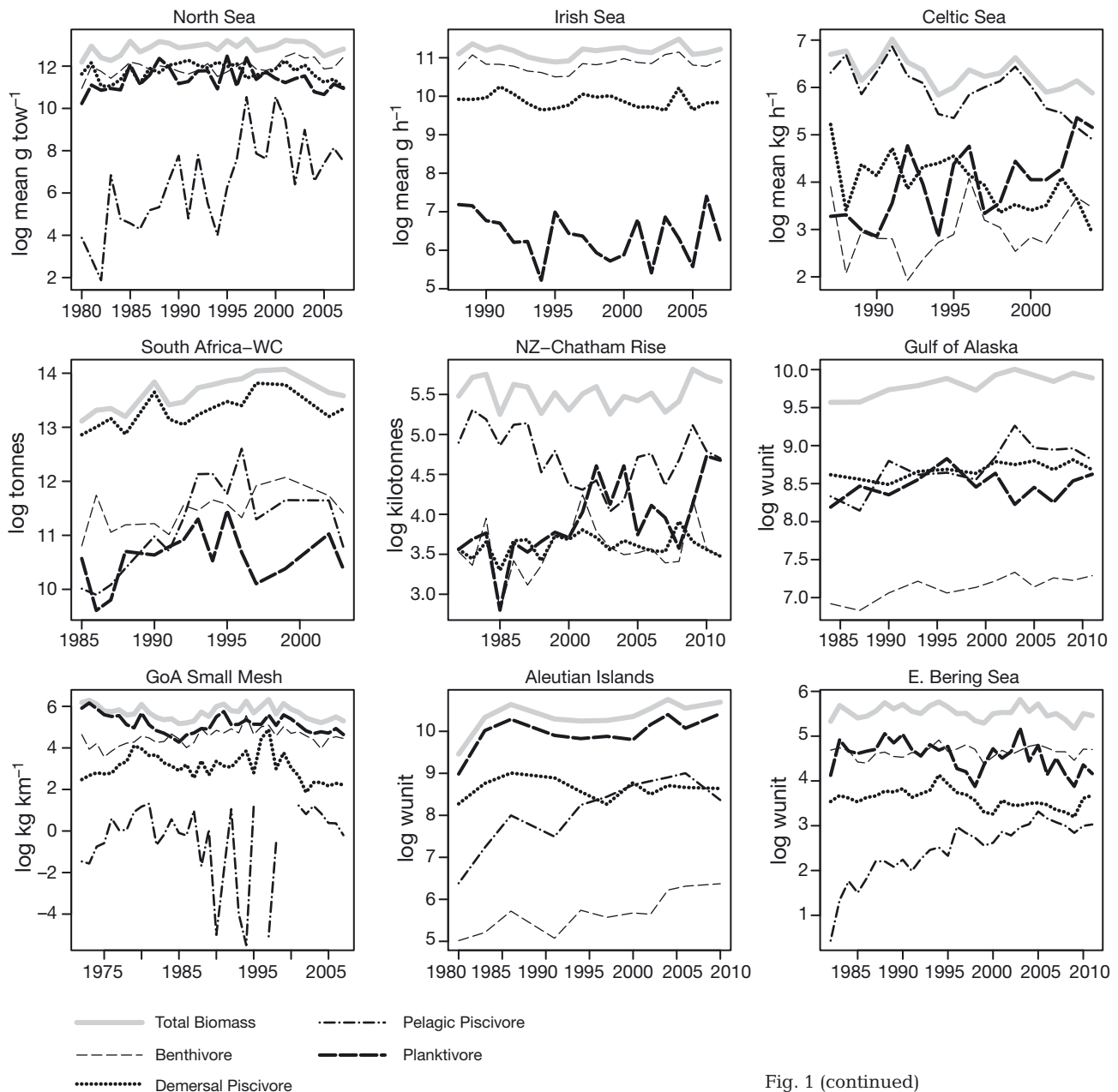


Fig. 1 (continued)

intercept-only model might not be the best fit for the Celtic Sea, Northern Gulf of St. Lawrence, South African West Coast and Scotian Shelf surveys.

As a final measure of stability, the surveys were tested for stationarity (Table 2). Several surveys were stationary, rejecting the null hypothesis that they contained a unit root. The Eastern Bering Sea, Gulf of Alaska small mesh survey, Fox Island, North Sea, Irish Sea, Mid-Atlantic Bight, and New Zealand Chatham Rise surveys were stationary ( $\alpha = 0.1$  for the Irish Sea). These surveys were considered the most stable over the time period. Georges Bank, Southern

Grand Banks and Whale Rock were only stationary after detrending and were not considered stable ( $\alpha = 0.1$  for the Southern Grand Banks). The surveys broke down into 4 categories (Table 3).

The 7 surveys that were considered stable had no trend in total biomass and were stationary. A few appeared to show some change over time, such as the Gulf of Alaska small mesh survey, but were time invariant in the mean and variance. The surveys designated as undetermined typically had a significant uni- or multidirectional trend, but had very similar AIC values for the intercept-only model (Fig. 1). The



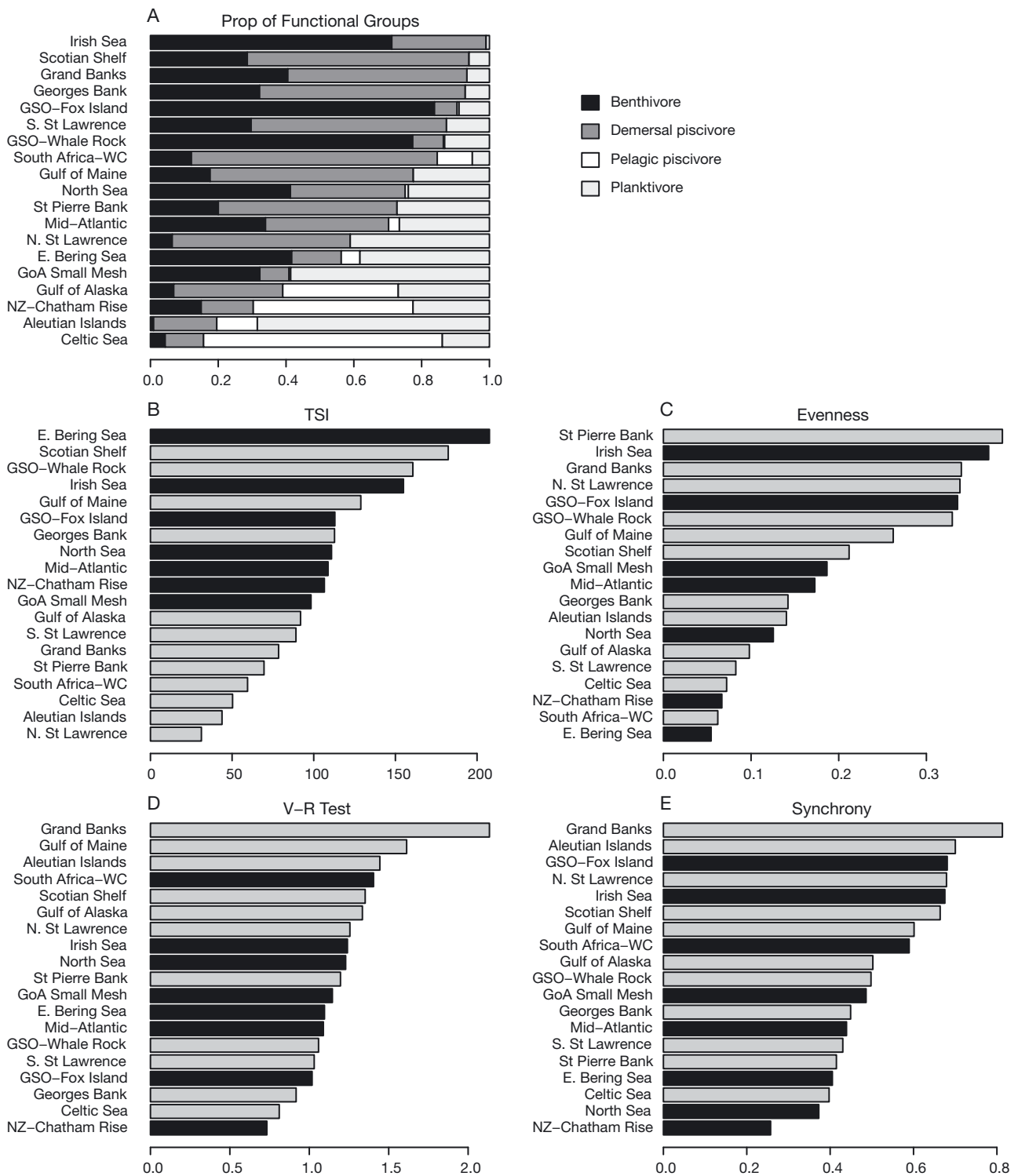


Fig. 2. (A) Mean proportion of functional group biomass over the length of the survey. Demersal fish are indicated with darker colors and pelagic fish with lighter colors. A number of surveys did not have any pelagic piscivores. The black bars in the following 4 subplots (B–E) are the surveys that were stationary and considered stable; the grey bars were the surveys that were not stationary. (B) The temporal stability index (TSI) (Lehman & Tilman 2000). (C) Simpson’s evenness. (D) Variance-ratio test on the total biomass and functional group biomass of the 19 surveys. (E) A metric of synchrony among time series that is similar to the variance-ratio test (Loreau & Mazancourt 2008). The metric scales between 0 and 1, with 0 representing complete asynchrony among time series and 1 complete synchrony

Table 2. Results of the augmented Dickey-Fuller test for stationarity on the log of total biomass with the appropriate lags, drift term (intercept) and trend term if needed. Surveys without a trend that were significant ( $p < 0.05$ ; **bold** text) rejected the null hypothesis that the time series has a unit root and were stationary. Surveys with a significant trend ( $p < 0.05$ ) were stationary after detrending

Survey	Type	Lag	$\tau$	$p < \tau$
<b>Eastern Bering Sea</b>	Drift	0	-3.92	<b>0.0053</b>
Aleutian Islands	Trend	4	-2.21	0.4624
<b>Gulf of Alaska Small Mesh</b>	Drift	0	-2.97	<b>0.0472</b>
Gulf of Alaska	Trend	4	-0.51	0.9752
<b>New Zealand Chatham Rise</b>	Drift	0	-4.00	<b>0.0070</b>
South African West Coast	Drift	0	-1.88	0.3353
Celtic Sea	Drift	0	-2.40	0.1572
<b>Irish Sea</b>	Drift	0	-2.70	<b>0.0921</b>
<b>North Sea</b>	Drift	0	-4.35	<b>0.0019</b>
<b>Mid-Atlantic Bight</b>	Drift	0	-6.30	<b>0.0002</b>
Georges Bank	Trend	0	-4.37	0.0061
GSO-Whale Rock	Trend	0	-4.29	0.0067
<b>GSO-Fox Island</b>	Drift	0	-3.56	<b>0.0097</b>
Gulf of Maine	Drift	1	-1.38	0.5840
Scotian Shelf	Drift	0	-2.05	0.2663
Southern Gulf of St. Lawrence	Trend	0	-2.69	0.2470
Northern Gulf of St. Lawrence	Trend	0	-2.62	0.2763
St. Pierre Bank	Trend	0	-5.86	0.0001
Grand Banks	Trend	0	-3.10	0.1187

Celtic Sea and South African West Coast had significant slopes in total biomass and in some of the functional groups, but fit the intercept-only model equivalently. They did not meet the definition of stationary and thus were not considered stable. The Scotian Shelf biomass increased to a peak in the 1980s before declining to a somewhat steady level and had a significant fit to a second-order polynomial. The peak only covered about half of the time series and an intercept-only model fit equally well. It was clear, however, that there had been a significant change in the total biomass over the time series and it was not stationary. The Northern Gulf of St. Lawrence had a sharp decline in the early part of the time series and then total biomass was relatively flat. None of the

models fit this survey particularly well and it was not stationary.

The linear model fits of the logged functional group biomass varied by group and survey (Fig. 3). When the competing models had AIC values that were <2.5 apart, the simpler model (model with fewer terms) was selected. Many of the trawl surveys that had stable total biomass were also stable for the majority of their functional groups. The stable Irish Sea, New Zealand Chatham Rise and Gulf of Alaska small mesh surveys were flat (intercept-only model) for all functional groups, while others were flat for all but one functional group. For the stable Eastern Bering Sea, Mid-Atlantic Bight and others, the one functional group that exhibited a trend constituted a small portion of the total biomass (Fig. 2A). The pelagic piscivores in the Mid-Atlantic Bight increased over the early part of the time series, but made up <10% of the total biomass. In contrast, trawl surveys that had trends in total biomass often had trends in functional group biomass. Benthivores and planktivores on Georges Bank decreased and then increased over the time series (second-order polynomial), while demersal piscivores had an increasing trend for the entire period. The Southern Grand Banks also had trends in functional group biomass and in total biomass. The functional group trends or lack of trends in some surveys, however, did not match the stability category of the total biomass.

The Aleutian Islands and Gulf of Alaska surveys were not conducted annually. The interpolated time series for the 2 surveys most likely reduced their variability compared with their true variability. This slightly limits the ability to compare these surveys along with the annual surveys across metrics that include variability such as the TSI. The interpolated and non-interpolated time series for the 2 surveys, however, exhibited significant trends in total biomass and were therefore not stationary. The trend would most likely have persisted even if there were data during the missing years and the non-stationary status of the 2 time series would be retained.

Table 3. The surveys were broken down into 4 categories based on the stability analysis

Second-order polynomial	Trending	Stable	Undetermined
Gulf of Maine	Southern Gulf of St. Lawrence	GSO-Fox Island	Scotian Shelf
	St. Pierre Bank	North Sea	Northern Gulf of St. Lawrence
	Grand Banks	Irish Sea	Celtic Sea
	GSO-Whale Rock	Mid-Atlantic Bight	South African West Coast
	Georges Bank	Eastern Bering Sea	
	Aleutian Islands	New Zealand Chatham Rise	
	Gulf of Alaska	Gulf of Alaska Small Mesh	

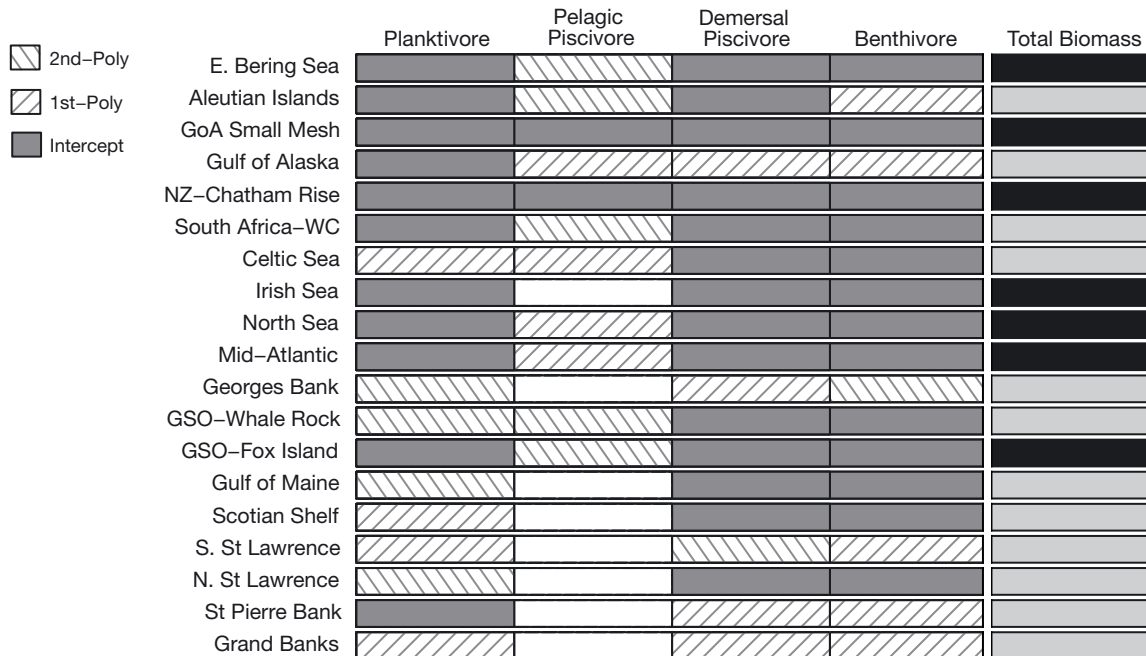


Fig. 3. Best-fit linear model based on Akaike's information criterion for the functional groups in each trawl survey (intercept-only model, linear regression—first-order polynomial and second-order polynomial). Surveys without pelagic piscivores have blank bars. The black bars in the total biomass column are the surveys that were considered stable

### Community metrics

#### Diversity and dominance

The Shannon-Wiener and Simpson diversity indices exhibited similar patterns and were highly correlated (adjusted  $R^2 = 0.75$ ,  $p < 0.0001$ ). The Scotian Shelf survey had the highest diversity and the Northern Gulf of St. Lawrence had the lowest. The Northern Gulf of St. Lawrence survey included only 10 species that were well represented over the time series. Again areas of similar geographical locations did not exhibit similar diversity patterns.

The values for all the surveys were relatively low for Simpson's evenness, ranging from 0.05 to 0.39 (Fig. 2C). The total biomass in each survey was largely dominated by a small number of taxa. Surveys with an evenness value  $< 0.1$  typically had 1 or 2 taxa that constituted a large proportion of the total biomass, while surveys with higher values were less influenced by the variability of a single taxon. The stable surveys were spread across the range of values for diversity and evenness without a clear pattern.

#### Compensation metrics

The majority of surveys had a variance-ratio value  $> 1$ , indicating that external factors were the main

drivers of community dynamics and not compensation (Fig. 2D). The New Zealand Chatham Rise, Celtic Sea and Georges Bank surveys were the only surveys  $< 1$ , suggesting compensation. When examined with the phase-scrambled bootstrap technique of Solow & Duplisea (2007) to test for significance (not shown), none of the variance-ratio tests indicated compensation. All of the surveys were above their critical value indicating that they were  $> 1$ . The stable surveys exhibited a range of values for the variance-ratio with no relationship between the two.

The surveys exhibited a range of values for the synchrony index applied to biomass from the least synchronous, the New Zealand Chatham Rise survey at 0.26, to the most synchronous, the southern Grand Banks at 0.81 (Fig. 2E). The stable surveys were interspersed along the gradient. The 2 surveys at the extremes of the variance-ratio test and synchrony index were the same, but the order of the remainder of the surveys was often quite different between the 2 metrics.

Stability as measured by the TSI was regressed against the measures of diversity, evenness and compensation. The TSI exhibited a positive relationship with both diversity indices, Shannon's (adjusted  $R^2 = 0.14$ ,  $p = 0.06$ ) and Simpson's (adjusted  $R^2 = 0.18$ ,  $p = 0.04$ ). The TSI showed no relationships with Simpson's evenness or any of the compensation metrics (Fig. S3). The TSI tended to be higher for the station-

ary systems; however, some systems with strong uni- or multidirectional trends also had high TSI values (Fig. 2B).

### Fishing pressure index

The mean of the fishing pressure index ( $\overline{FPI}_t$ ) ranged from 0.38 to 1.57 (Fig. 4). While the mean does not provide information on the changes in exploitation over the time series, it does provide a starting point for comparisons among surveys. Values  $<1$  represent systems in which the mean exploitation is below the value that would produce the maximum sustainable yield and is often considered a reference point for overfishing. The fishing pressure is on a system-wide scale and is sensitive to the assessments available. While certain individual species may be under or over exploited, the mean  $FPI$  is an average exploitation across the entire community. The systems below 1 were below 1 for the entire time series of the trawl survey (Fig. S4). The South African West Coast survey was the one exception and was  $>1$  for 3 early years out of a total of 23 yr. All the surveys with  $\overline{FPI}_t < 1$  were considered stable or were increasing (Gulf of Alaska,

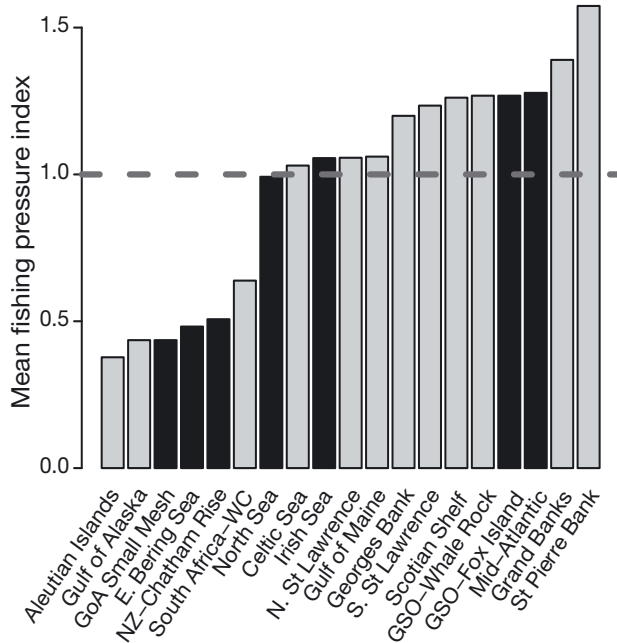


Fig. 4. Standardized mean fishing pressure index for each survey over the time series based on stock assessments. Different numbers of assessments were used due to availability (Eastern Bering Sea had 19 while some Canadian surveys had 1 or 2). Black bars represent surveys that were stable over the time series

Aleutian Islands, South African West Coast). Five surveys had mean exploitation rates at  $\sim 1$ . The  $FPI_t$  for the 3 European surveys varied around 1 for the length of the trawl surveys, while the Northern Gulf of St. Lawrence and the Gulf of Maine had very high exploitation rates during portions of the time series and then dropped in later years. The remaining surveys were all  $>1$ , indicating that they had been heavily exploited. In all of these systems, exploitation rates had dropped in the later part of the time series, but had been high. In many of the Canadian surveys, the fishing pressure index was based only on a small number of assessments (cod, haddock and sometimes redfish); however, these taxa represented the bulk of the biomass and the majority of the landings. The  $FPI_t$  for the Canadian surveys from assessments was quite similar to the aggregate exploitation rate from the NAFO catch data (see the Supplement for further discussion).

### Gompertz population model

The linearized, multispecies Gompertz model fit the data well and the parameter estimates were reasonable, except for the Northern Gulf of St. Lawrence. In most surveys the predicted time series were very similar to the observed time series (Figs. S6 & S7). Time series models such as the vector autoregressive model typically follow the observed data closely. The fit of the predicted biomass was also directly related to the magnitude of the measurement error. The Kalman filter predicts the value of the biomass based on the value of the previous predicted biomass, the current observed biomass in combination with the magnitude of the process and measurement error. If the error terms are small, the predicted estimate is very close to the observed estimate. Demersal piscivores typically had very low measurement error suggesting that they were relatively evenly distributed over the entire area of the trawl surveys. The predicted biomass of demersal piscivores in turn followed the observed biomass closely. In contrast, planktivores typically had relatively high error terms and much smoother predicted biomass estimates. The state space model was able to incorporate the measurement error in the trawl surveys and provided better estimates of the biomass over time.

The actual model for each functional group typically reduced to a few explanatory parameters (Fig. 5). The density-dependence terms (diagonal of the  $B$  matrix) at the functional group level and the interactions between functional groups, when re-

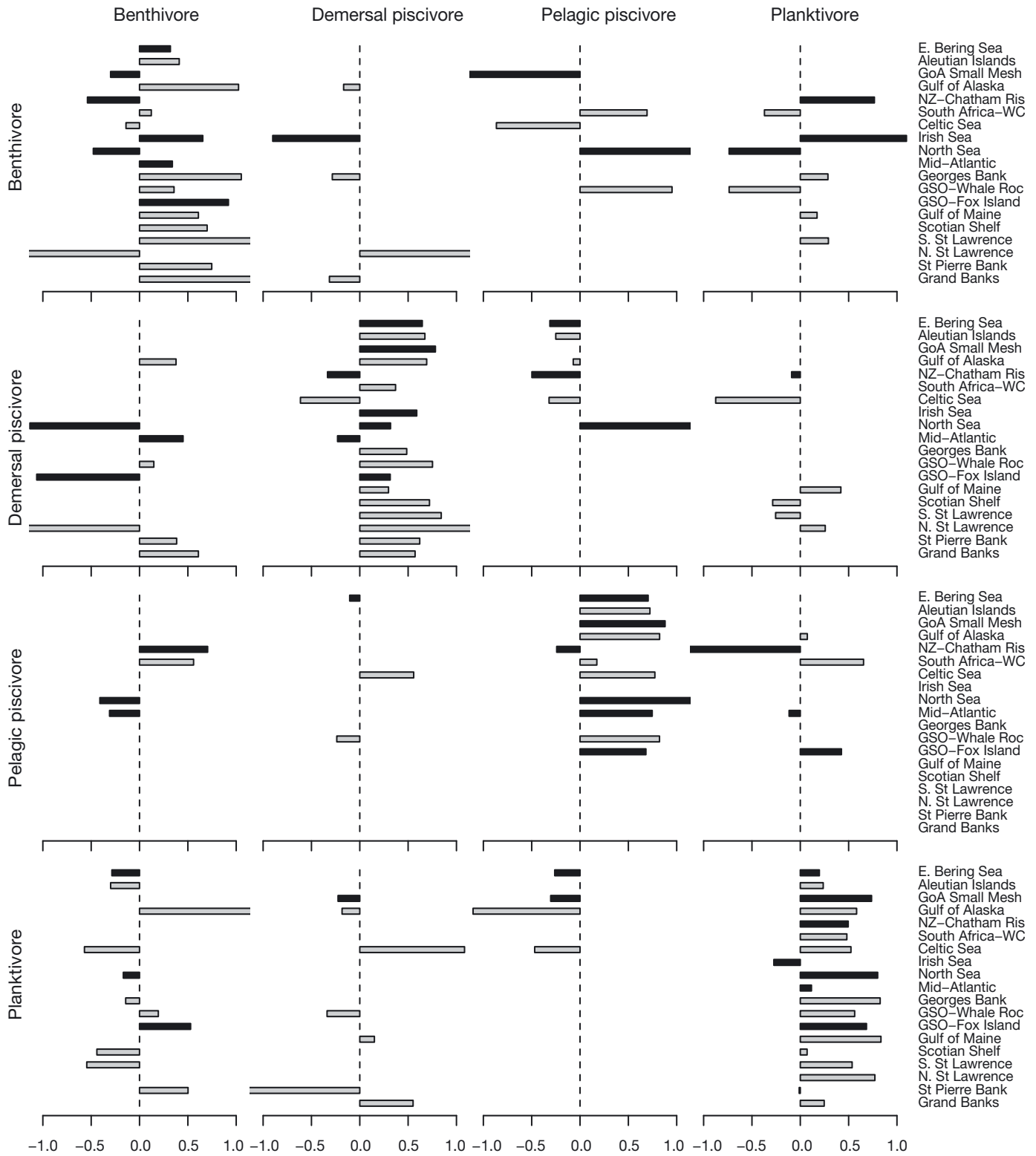


Fig. 5. Parameter estimates for the interaction matrix ( $B$  matrix). Each bar corresponds to the parameter estimate for a particular interaction for each survey as listed on the right. The density-dependent effects (e.g. planktivore-planktivore interaction) are the subfigures along the diagonal. For example, the first bar on the top left is the parameter estimate for benthivores as a function of benthivores on E. Bering Sea. The rows are a function of the functional groups in the column headings. For example, in 5 surveys demersal piscivores are a negative function of pelagic piscivores

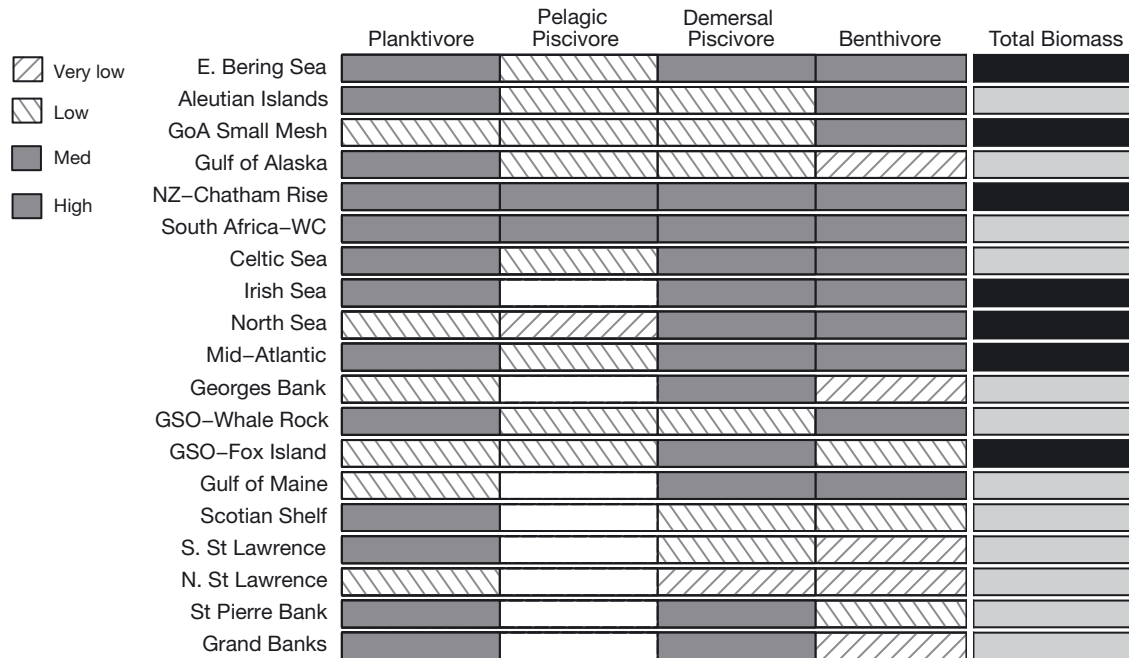


Fig. 6. Density-dependent terms from the Gompertz models for each survey. Surveys without pelagic piscivores have blank bars. The black bars in the total biomass column are the surveys that were considered stable

tained, were highly variable. The external drivers (the environment and exploitation) also varied in strength among surveys.

The South African West Coast survey and the Mid-Atlantic Bight exhibited high levels of density dependence (low parameter estimates along the diagonal of the  $B$  matrix), while Georges Bank, the Southern Gulf of St. Lawrence and the Northern Gulf of St. Lawrence had moderate to low levels of density dependence (Fig. 6). High density dependence results in a high level of self regulation and could be an indication of within-group compensation. A number of stable surveys, such as the Mid-Atlantic Bight, exhibited both high density dependence in their functional groups and lacked uni- or multidirectional trends. The pelagic piscivores of the Mid-Atlantic Bight had low density dependence and highly variable biomass, but constituted a very small proportion of the system and their biomass would have limited influence on the variability of the total biomass. Similarly, the 2 dominant functional groups in the stable North Sea survey (benthivores and demersal piscivores) had high density-dependence terms and stable biomass resulting in stable system-level biomass. The biomass of pelagic piscivores was highly variable and had low density dependence, but made up a small proportion of the total biomass.

Certain surveys had negative interaction terms, suggesting that particular functional groups compen-

sated for others over the course of the time series, but they were not ubiquitous (Fig. 1). On Georges Bank, benthivores had a negative parameter estimate for demersal piscivores, suggesting that the demersal piscivores compensated for the decline in benthivores, particularly during the 1980s. On the New Zealand Chatham Rise, the Gompertz model suggested that planktivores compensated for pelagic piscivores when they dipped in the 2000s. In the Southern Gulf of St. Lawrence and Scotian Shelf, there were negative interactions between planktivores and the demersal taxa (benthivores and demersal piscivores). The most common interaction among all the surveys was planktivores increasing when benthivores decreased. The majority of interactions were one way, however, and there were fewer estimates for benthivores as a function of planktivores within the same survey. The Celtic Sea stood out as the only survey with interaction parameter estimates that were numerous and largely negative. Overall, however, there were few interactions among functional groups.

Based on the simulated interaction matrices, systems with high density dependence and weak or absent interactions were the most stable over time (lowest eigenvalues) (Fig. S8). Almost all of the simulated systems with weak interactions, regardless of the strength of the density-dependence terms, had eigenvalues within the unit circle. For the systems

with weak interactions, only those with interaction terms for every functional group were at the limit of the unit circle or above it. Systems with strong interactions varied considerably in stability. Those with 2 or less interactions were consistently within the unit circle, but higher numbers of strong interactions led to instability. Strong interactions among functional groups could only be sustained in systems with high density dependence.

All the eigenvalues of the  $B$  matrix from the Gompertz models fit to observed data were  $<1$ , except for the Northern Gulf of St. Lawrence (Fig. S9). This indicates that under steady-state conditions all the surveys, but one, would move toward their stationary distribution. The values are relatively high, however, suggesting that they would be sensitive to small perturbations. The eigenvalues do not correlate with any of the measures of stability.

## DISCUSSION

Despite major commercial fishing operations taking place in every system sampled by the fisheries-independent trawl surveys, a number of areas had relatively stable biomass over their time series. The surveys exhibited a large range in variability, but in 7 of the 19 communities analyzed, total biomass did not exhibit a uni- or multidirectional trend. The ecosystems with lower levels of fishing pressure and higher levels of density dependence were the most stable. Community metrics, such as compensation among functional groups and evenness, however, did not explain the patterns in stability.

All surveys had eigenvalues within the unit circle except for the Northern Gulf of St. Lawrence, indicating that under steady-state conditions the surveys would be stable. Trends in total biomass, therefore, strongly suggested the role of external drivers in determining the community dynamics. Fishing is a major external driver in marine systems, removing huge amounts of biomass, and when unchecked can severely deplete target populations (Jennings 2007, Worm et al. 2009). The stability of each trawl survey was strongly related to its mean fishing pressure index. Surveys with low levels of community exploitation were stable or increased over the time series, while those with high levels of community exploitation typically had a history of substantial declines.

The time series of fishing pressure for systems that were below their community-level  $U_{msy}$  were below that level for almost their entire survey record. Exploitation was low enough that it did not synchro-

nize the community dynamics and species-level interactions and environmental factors regulated the system. The biomass of the major functional groups was relatively flat and the total biomass was stable or increasing. Surveys whose mean mean fishing pressure index was above the community-level  $U_{msy}$  typically had high exploitation during the early years of the survey that decreased to near or below the community-level  $U_{msy}$  in the later years. The fishing levels were high enough that target species, and in some cases potential compensating species, were removed. Alternative functional groups that typically made up a small proportion of total biomass, such as planktivores in the Southern Gulf of St. Lawrence and Scotian Shelf, exhibited some compensation, but it was not enough to offset the removals. The external driver was large enough that it overwhelmed the species interactions and synchronized the trends in populations. As exploitation levels declined in the later years, the functional groups responded to the change in the external forcing, but continued to covary together. The synchronized increase in the later part of the time series is evident in total biomass and functional groups in Georges Bank and the Gulf of Maine. The biomass of the Southern Grand Banks has also increased since the 1990s; however, this analysis only had the data until 1995 when the trawl survey changed gears. The European surveys were roughly at their community-level  $U_{msy}$  and fluctuated around it over the survey record without exhibiting a significant trend. The total biomass in these surveys appeared to have equilibrated with the level of fishing and varied about a mean level.

The Mid-Atlantic Bight and Fox Island communities both had high fishing pressure indices and were stable. These 2 surveys do not appear to follow the pattern as described above, but this is most likely due to a simple lack of data rather than an anomaly. A few stocks from these surveys, which constitute a large amount of biomass, had unusable assessments or were not in the RAM Legacy assessment database (Ricard et al. 2011) and are thus missing from the analysis. These stocks would most likely bring down the fishing pressure index. Butterfish and Loligo squid are 2 of the top 3 species in terms of biomass in the Mid-Atlantic Bight and neither could be included in the fishing pressure index. Both are short-lived species (1 to 2 yr life history) that can potentially breed more than once a year and are difficult to assess. Biological reference points were not available because of the difficulties in assessment. Consumption by predators, however, far outweighs removals due to fishing and both are considered lightly fished

(NEFSC 2010, 2011). In addition, non-commercial stocks are also important in some areas, but do not have assessments. Cancer crabs make up a large proportion of the total biomass in the Fox Island survey, but are not assessed. The inclusion of these stocks would bring down the mean fishing pressure index.

Despite stable survey biomass in certain systems, overfishing of particular species or stocks may still be occurring. Aggregate community-level fishing pressure says little about an individual species exploitation. The analysis does not imply that systems with low community exploitation or stable biomass result in stable biomass for individual species or that all species are doing well. In many theoretical analyses, stable community-level biomass often results in less stable individual species biomass (May 1974a, Lehman & Tilman 2000).

Based on the Gompertz model simulations, systems with high density dependence and weak (or absent) interactions were the most stable over time. Food webs with weak interactions ensure that predators are not reliant on a single species and competition is not so strong as to drive inferior species to extinction (May 1974a, McCann et al. 1998, Closs et al. 1999, Garrison & Link 2000, Link 2002). Predators with many weak links are generalists and do not rely on a single prey species. They have the ability to switch between available prey and thus maintain their own population despite environmental factors that may favor one prey species over another (May 1974a, Closs et al. 1999). Weak competitive interactions allow numerous species to coexist within similar niche spaces without any of them being driven to local extinction by the competitively dominant species. In contrast, systems with large numbers of strong interactions can lead to highly unstable biomass and potentially chaotic dynamics. Numerous modeling studies have shown that as the number of linkages in a food web increases, the strength of the interactions must decrease to persist (May 1974a,b, McCann et al. 1998, Closs et al. 1999, Link 2002). Strong interactions could only be sustained in simulations with high levels of self regulation and limited numbers of interactions. Strong interactions make the entire community reliant on a limited number of species or functional groups. If that functional group declines, the entire system will decline with it. In the Peruvian upwelling system, sardines and anchovies are weakly density dependent and the links between them and higher trophic levels (predatory fish, marine mammals and birds) are very tight. The impacts of El Niño Southern Oscillation severely affect the pelagic planktivores, which cascades up the food

web causing large changes in total system biomass (Pauly & Tsukayama 1987).

The Gompertz model parameter estimates for the real survey data did not fit into the neat categories developed for the simulations. Most trawl surveys had density-dependent terms that ranged from low to high and had both strong and weak interaction terms. The relatively high eigenvalues just below 1 for most surveys reflect the range of parameter estimates. Overall, however, there were few interactions among functional groups (off diagonal terms in the *B* matrix) and few consistent patterns, suggesting that at the functional group level there was little evidence for compensation as the main regulator of system dynamics. The variance-ratio test and the Loreau & Mazancourt (2008) synchrony index also showed little evidence for compensation among functional groups. In some surveys, however, there was evidence of negative interactions among functional groups and strong density dependence within functional groups.

In the standard Gompertz population groups model, the *b* term represents the density dependence of a single species: the feedback relationship between the number of individuals and the available resources. Density dependence at the functional group level (the diagonal of the *B* matrix) represents the same relationship except it relates to the different taxa that constitute a functional group and the available resources. Ecosystems with high levels of density dependence typically had no trends in the functional groups that comprised the majority of their biomass, resulting in community-level stability. The within-group dynamics led to self regulation at the functional group level and relatively stable biomass. The within-group dynamics could be competition for resources, which suggests within-functional-group compensation. In contrast, surveys with strong trends, such as the Southern Gulf of St. Lawrence and Scotian Shelf, exhibited low density dependence in their major functional groups (benthivores and demersal piscivores) and the biomass fluctuated over time.

Many studies have found a strong link between diversity and stability in terrestrial ecosystems (Tilman 1996, Lehman & Tilman 2000, Tilman et al. 2006). In the marine systems examined, diversity had a significant or near significant relationship to the variability of the communities (TSI), but was not a good predictor of stability as indicated by stationarity. The analysis supported the idea that increased diversity results in increased temporal stability; however, trawl-survey data may be difficult to use with the TSI. The measurement error that is part of a fisheries-independent trawl survey is implicitly included



in the calculation of the TSI. The measurement error increases the variance and decreases the TSI, but it is not uniform across surveys and varies with the survey design and gear. We therefore focused on long-term trends in total biomass as measured with the augmented Dickey-Fuller test and analyses that accounted for measurement error to avoid the year-to-year variability, which could cause problems with the TSI.

Compensation is the mechanism that enables community-level stability despite variation in external factors. While compensation is well founded on theoretical grounds, it is difficult to find with empirical data (Houlahan et al. 2007, Gonzalez & Loreau, 2009). The lack of compensation among functional groups may be due to the coarse scale with which we aggregated the taxa. Houlahan et al. (2007) found that data aggregated at higher levels may have obscured detectability. We followed the broad definitions of Heath (2005) and Greenstreet et al. (1997) for each functional group to facilitate comparisons across a range of trawl surveys from around the world and to examine changes in energy pathways. While we used the sample areas of the trawl surveys to define the different ecosystems, the survey boundaries did not cover the full range of all the species sampled. This is a potential issue for all studies that use trawl-survey data and we attempted to minimize it by including only those species whose main distribution fell within the survey areas and were well represented in the trawl data. In addition, the relatively short time series limited the number of parameters we could estimate with the vector autoregressive state space models and therefore required aggregation. The aggregated data allowed us to smooth over missing data. However, the 4 diet-based functional groups may not recognize the detailed niche partitioning in the diverse ecosystems represented. While compensation appeared to be a major structuring agent within functional groups, a finer scale approach may yield additional insight.

The length of the trawl-survey time series also limited our ability to examine interactions that might take place over a time delay of more than 1 year. Interactions occur that manifest themselves at longer time scales (Fogarty & Brodziak 1994); however, fitting and interpreting higher-order models, particularly with short time series, can be problematic. First-order vector autoregressive models can approximate more complex models and provide interpretable results (Ives 1995, Ives et al. 2003). The lag-one vector autoregressive models provide simplified tools to analyze complex communities.

The stability of each trawl survey is closely linked to higher levels of density dependence—within-group compensation—and low levels of exploitation. The classification of surveys as stable, however, is also related to the time period of the fishery that is captured by the survey. The surveys started at different times in relation to the development of the fishery and thus record different periods. Over the past few decades, the majority of the surveys have been relatively stable or increasing; however, surveys that began in the 1950s and 1960s, such as those on the east coast of North America, capture an initial decline in biomass. The IndiSeas working group has found similar patterns (Bundy et al. 2010). The east coast of the United States experienced heavy fishing by foreign fleets until exclusive economic zones were imposed in the mid-1970s and then by national fleets until the Magnuson-Stevens Act in the 1990s (Collette & Klein-MacPhee 2002). Georges Bank and the Gulf of Maine surveys began in 1963 and captured a period of biomass depletion by mechanized fisheries followed by an increase over the past 2 decades. The North Sea survey began in 1980, well after the start of industrial fishing, and potentially missed an initial fishing down period. A reconstruction of the North Sea ecosystem from 1880 suggests that the total biomass is less than half of what it was a century ago (Mackinson 2002, Mackinson & Daskalov 2007). Industrial fishing after World War II dramatically reduced the fishery. There are also systems in which management policies were imposed towards the beginning of the fishery and total biomass was not fished down (Eastern Bering Sea, New Zealand Chatham Rise). The timing of the fisheries-independent trawl surveys is important, but it does not change the outcome of the study. The exploitation history was included, so if surveys had started earlier and potentially captured biomass depletion, they would have also included higher levels of exploitation. While the North Sea may be depleted compared with historical levels, over the past 30 yr, fishing has been at a level that maintains a relatively constant total biomass. The systems with lower levels of aggregate exploitation allowed high levels of within-group compensation and were more stable.

## CONCLUSION

Fisheries-independent trawl surveys provide valuable insight into the dynamics of marine ecosystems. In systems with higher levels of exploitation, the intra- and interspecific interactions were over-

whelmed, synchronizing the populations and leading to large changes in total biomass. Lower levels of fishing pressure enabled the differential rates of growth, immigration and emigration among the different populations to be the important factors regulating community dynamics, resulting in more stable total biomass. The stable systems showed higher levels of density dependence, suggesting that strong self regulation within functional groups mediated through compensation maintained total biomass. Systems with stable total biomass typically had functional groups that were relatively flat over the time series and exhibited little evidence for compensation among functional groups.

*Acknowledgements.* We thank the countless individuals and institutions who spent years collecting and processing the fisheries-independent data and environmental data (Canadian Department of Fisheries and Oceans; University of Rhode Island, USA; Centre for Environment, Fisheries and Aquaculture Science, UK; National Marine Fisheries Service, USA; Alaskan Department of Fish and Game, USA; National Institute of Water and Atmospheric Research, New Zealand; Marine and Coastal Management, South Africa). We gratefully acknowledge D. Ricard, T. Branch, C. Minto, B. Smith, S. Gaichas, K. Aydin, M. Baker, M. Dunn, M. Palomares, E. Holmes, E. Ward and R. Leslie. This work was supported by the Fisheries and the Environment (FATE) program through the National Marine Fisheries Service. Three anonymous reviewers improved the manuscript.

#### LITERATURE CITED

- Anderson P, Piatt J (1999) Community reorganization in the Gulf of Alaska following ocean climate regime shift. *Mar Ecol Prog Ser* 189:117–123
- Auster P, Link J (2009) Compensation and recovery of feeding guilds in a northwest Atlantic shelf fish community. *Mar Ecol Prog Ser* 382:163–172
- Branch TA, Watson R, Fulton EA, Jennings S and others (2010) The trophic fingerprint of marine fisheries. *Nature* 468:431–435
- Brocklebank J, Dickey D (2003) SAS for forecasting time series, 2nd edn. SAS Institute, Cary, NC
- Brown J, Bedrick E, Ernest S, Catron J, Kelly J (2004). Constraints on negative relationships: mathematical causes and ecological consequences. In: Taper ML, Lele SR (eds) *The nature of scientific evidence: statistical, philosophical, and empirical considerations*. University of Chicago Press, Chicago, IL, p 298–308
- BSCW (2011) Bering climate: sea surface temperature. [www.beringclimate.noaa.gov/data/index.php](http://www.beringclimate.noaa.gov/data/index.php)
- Bundy A, Shannon LJ, Rochet MJ, Neira S, Shin YJ, Hill L, Aydin K (2010) The good(fish), the bad, and the ugly: a tripartite classification of ecosystem trends. *ICES J Mar Sci* 67:745–768
- Closs G, Balcombe S, Shirley M (1999) Generalist predators, interaction strength and food-web stability. *Adv Ecol Res* 28:93–126
- Collette B, Klein-MacPhee G (eds) (2002) *Bigelow and Schroeder's fishes of the Gulf of Maine*. Smithsonian Institution, Washington, DC
- Collie J, Wood A, Jeffries H (2008) Long-term shifts in the species composition of a coastal fish community. *Can J Fish Aquat Sci* 65:1352–1365
- Ernest S, Brown J (2001) Homeostasis and compensation: the role of species and resources in ecosystem stability. *Ecology* 82:2118–2132
- Fischer J, Frost T, Ives A (2001) Compensatory dynamics in zooplankton community responses to acidification: measurement and mechanisms. *Ecol Appl* 11:1060–1072
- Fogarty M, Brodziak J (1994) Multivariate time series analysis of fish community dynamics on Georges Bank. ICES C.M. 1994/P:3. International Council for the Exploration of the Sea, Copenhagen
- Fogarty MJ, Murawski SA (1998) Large-scale disturbance and the structure of marine systems: fishery impacts on Georges Bank. *Ecol Appl* 8:S6–S22
- Frank KT, Petrie B, Choi JS, Leggett WC (2005) Trophic cascades in a formerly cod-dominated ecosystem. *Science* 308:1621–1623
- Garrison L, Link J (2000) Dietary guild structure of the fish community in the Northeast United States continental shelf ecosystem. *Mar Ecol Prog Ser* 202:231–240
- Genner MJ, Sims DW, Wearmouth VJ, Southall EJ, Southward AJ, Henderson PA, Hawkins SJ (2004) Regional climatic warming drives long-term community changes of British marine fish. *Proc Biol Sci* 271:655–661
- Gonzalez A, Loreau M (2009) The causes and consequences of compensatory dynamics in ecological communities. *Annu Rev Ecol Evol Syst* 40:393–414
- Greenstreet S, Bryant A, Broekhuizen N, Hall S, Heath M (1997) Seasonal variation in the consumption of food by fish in the North Sea and implications for food web dynamics. *ICES J Mar Sci* 54:243–266
- Heath M (2005) Changes in the structure and function of the North Sea fish foodweb, 1973–2000, and the impacts of fishing and climate. *ICES J Mar Sci* 62:847–868
- Holling C (1973) Resilience and stability of ecological systems. *Annu Rev Ecol Evol Syst* 4:1–23
- Holmes E, Ward E (2012) Analysis of multivariate time-series using the MARSS package, R software, version 2.8. NOAA Fisheries, NWFSC, Seattle, WA
- Houlahan JE, Currie DJ, Cottenie K, Cumming GS and others (2007) Compensatory dynamics are rare in natural ecological communities. *Proc Natl Acad Sci USA* 104:3273–3277
- Hurrell J, Deser C (2009) North Atlantic climate variability: the role of the North Atlantic Oscillation. *J Mar Syst* 78:28–41
- Ives AR (1995) Predicting the response of populations to environmental change. *Ecology* 76:926–941
- Ives AR, Cardinale BJ (2004) Food-web interactions govern the resistance of communities after non-random extinctions. *Nature* 429:174–177
- Ives AR, Carpenter SR, Dennis B (1999) Community interaction webs and zooplankton responses to planktivory manipulations. *Ecology* 80:1405–1421
- Ives AR, Klug JL, Gross K (2000) Stability and species richness in complex communities. *Ecol Lett* 3:399–411
- Ives AR, Dennis B, Cottingham KL, Carpenter SR (2003) Estimating community stability and ecological interactions from time series data. *Ecol Monogr* 73:301–330
- Jennings S (2007) Reporting and advising on the effects of fishing. *Fish Fish* 8:269–276

- JISAO (2011) Joint Institute for the Study of the Atmosphere and Ocean at the University of Washington: Pacific Decadal Oscillation. [www.jisao.washington.edu/pdo/PDO.latest](http://www.jisao.washington.edu/pdo/PDO.latest)
- Kalnay E, Kanamitsu M, Kistler R, Collins W and others (1996) The NMC/NCAR 40-Year Reanalysis Project. *Bull Am Meteorol Soc* 77:437–471
- Klug J, Fischer J, Ives A, Dennis B (2000) Compensatory dynamics in planktonic community responses to pH perturbations. *Ecology* 81:387–398
- Krause AE, Hayes DB, Bence JR, Madenjian CP, Stedman RM (2002) Measurement error associated with surveys of fish abundance in Lake Michigan. *J Gt Lakes Res* 28: 44–51
- Lehman C, Tilman D (2000) Biodiversity, stability and productivity in competitive communities. *Am Nat* 156: 534–552
- Lindegren M, Mollmann C, Nielsen A, Stenseth NC (2009) Preventing the collapse of the Baltic cod stock through an ecosystem-based management approach. *Proc Natl Acad Sci USA* 106:14722–14727
- Link J (2002) Does food web theory work for marine ecosystems? *Mar Ecol Prog Ser* 230:1–9
- Loreau M, Mazancourt C (2008) Species synchrony and its drivers: neutral and nonneutral community dynamics in fluctuating environments. *Am Nat* 172:E48–E66
- Mackinson S (2002) Representing trophic interaction in the North Sea in the 1880s using the Ecopath mass-balance approach. In: Gunette V, Christensen D (eds) *Fisheries impacts on North Atlantic ecosystems: models and analyses*. Fisheries Centre Research Reports 9(4), p 35–98
- Mackinson S, Daskalov G (2007) An ecosystem model of the North Sea to support an ecosystem approach to fisheries management: description and parameterisation. Tech Rep, Cefas, Lowestoft
- May RM (1974a) Ecosystem patterns in randomly fluctuating environments. In: Rosen R, Snell F (eds) *Progress in theoretical biology*, Vol. 3. Academic Press, New York, NY, p 349
- May RM (1974b) How many species: some mathematical aspects of the dynamics of populations. In: Cowen J (ed) *Lectures on mathematics in the life sciences. Some mathematical questions in biology*, Vol. 6. American Mathematical Society, Providence, RI, p 65–98
- McCann K, Hastings A, Huxel G (1998) Weak trophic interactions and the balance of nature. *Nature* 395:794–798
- NCAR (2011) Climate data guide: North Atlantic Oscillation. <http://climatedataguide.ucar.edu/guidance/hurrell-north-atlantic-oscillation-nao-index-station-based>
- NEFSC (2010) 49th Northeast Regional Stock Assessment Workshop (49th SAW) Assessment Report. US Department of Commerce, Northeast Fisheries Science Centre Ref Doc 10-03. Tech Rep, National Marine Fisheries Service, Woods Hole, MA
- NEFSC (2011) 51st Northeast Regional Stock Assessment Workshop (51st SAW) Assessment Report. US Department of Commerce, Northeast Fisheries Science Centre Ref Doc 11-01. Tech Rep, National Marine Fisheries Service, Woods Hole, MA
- NWS (2011) NOAA Climate Prediction Center: Southern Oscillation Index. [www.cpc.ncep.noaa.gov/data/indices/soi](http://www.cpc.ncep.noaa.gov/data/indices/soi)
- Nye J, Link J, Hare J, Overholtz W (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
- O'Connor NE, Crowe TP (2005) Biodiversity loss and ecosystem functioning: distinguishing between number and identity of species. *Ecology* 86:1783–1796
- Patten B (1975) Ecosystem linearization: an evolutionary design problem. *Am Nat* 109:529–539
- Pauly D, Tsukayama I (eds) (1987) *The Peruvian anchoveta and its upwelling ecosystem: three decades of change*. ICLARM Stud Rev 15. Instituto del Mar del Peru (IMARPE), Callao; Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), Eschborn; International Center for Living Aquatic Resources Management (ICLARM), Manila
- Pennington M (1985) Estimating the relative abundance of fish from a series of trawl surveys. *Biometrics* 41:197–202
- Pennington M (1995) Measuring the effect of changes in catchability on the variation of marine survey abundance indices. *Fish Res* 23:301–310
- Perry AL, Low PJ, Ellis JR, Reynolds JD (2005) Climate change and distribution shifts in marine fishes. *Science* 308:1912–1915
- Quinn T, Deriso R (1999) *Quantitative fish dynamics*. Oxford University Press, New York, NY
- Ricard D, Minto C, Jensen O, Baum J (2011) Evaluating the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. *Fish Fish* 13:380–398
- Rosenfeld J (2002) Functional redundancy in ecology and conservation. *Oikos* 70:156–162
- Sasaki T, Lauenroth W (2011) Dominant species, rather than diversity, regulates temporal stability of plant communities. *Oecologia* 166:761–768
- Schluter D (1984) A variance test for detecting species associations with some example applications. *Ecology* 65: 998–1005
- Smith B, Wilson J (1996) A consumer's guide to evenness indices. *Oikos* 76:70–82
- Smith T, Reynolds R, Peterson T, Lawrimore J (2007) Improvements to NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880–2006). <https://www.ncdc.noaa.gov/ersst/>
- Solow A, Duplisea D (2007) Testing for compensation in a multi-species community. *Ecosystems* 10:1034–1038
- Steele D, Collie J, Bisagni J, Gifford D and others (2007) Balancing end-to-end budgets of the Georges Bank ecosystem. *Prog Oceanogr* 74:423–448
- Tilman D (1996) Biodiversity: population versus ecosystem stability. *Ecology* 77:350–363
- Tilman D, Reich P, Knops J (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441:629–632
- Viscido S, Holmes E (2010) Statistical modelling of communities and ecosystems using the LAMDA software tool. *Environ Model Softw* 25:1905–1908
- Walker B (1992) Biodiversity and ecological redundancy. *Conserv Biol* 6:18–23
- Worm B, Barbier EB, Beaumont N, Duffy JE and others (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314:787–790
- Worm B, Hilborn R, Baum J, Branch T and others (2009) Rebuilding global fisheries. *Science* 325:578–585
- Xue Y, Smith T, Reynolds R (2003) Interdecadal changes of 30-yr SST normals during 1871–2000. *J Clim* 16:1601–1612