OPEN ACCESS

Comparative analyses of surplus production dynamics of functional feeding groups across 12 northern hemisphere marine ecosystems

Sean M. Lucey[1],*, Adam M. Cook[2], Jennifer L. Boldt[3], Jason S. Link[1], Timothy E. Essington[4], Thomas J. Miller[5]

[1]National Marine Fisheries Service, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, Massachusetts 02543, USA

[2]Fisheries and Oceans Canada, Bedford Institute of Oceanography, PO Box 1006, Dartmouth, Nova Scotia B2Y 4A2, Canada

[3]Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T 6N7, Canada

[4]School of Aquatic and Fishery Sciences, University of Washington, UW Box 355020, Seattle, Washington 98195, USA

[5]Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, PO Box 38, Solomons, Maryland 20688, USA

ABSTRACT: Progress on ecosystem approaches to fisheries management requires comparative studies with standardized methods that incorporate readily available data. This precludes complex ecosystem models in favor of simpler models such as surplus production models. Surplus production models for individual species can provide estimates of common biological reference points such as maximum sustainable yield (MSY) and the corresponding stock biomass level (B_{MSY}). For ecosystem approaches, summing multiple single-species surplus production models to estimate the ecosystem MSY ignores potential biological and fishery interactions among species. Improved estimates of ecosystem-level MSY can be obtained by aggregating species, thereby accounting for known interactions among species. Here, we fit surplus production models to 3 different types of aggregations for 12 northern hemisphere marine ecosystems. Aggregations were based on habitat (benthic/pelagic), foraging guild (planktivore/zoopivore/benthivore/piscivore) and size class (small/medium/large). The objectives of this work were to explore, compare and contrast model outputs across the various types of aggregations and among ecosystems. We found that regardless of the type of aggregation, aggregate production never exceeded 6 t km^{-2} and was generally less than 3 t km^{-2}. Patterns of production varied among ecosystems with no particular pattern with respect to ocean basin, latitude or component species. Aggregated surplus production models can provide biological reference points that are familiar to fishery managers and can be used to set overall removals with respect to aggregate group as long as less productive stocks are protected.

KEY WORDS: Biological reference point · Ecosystem · Fisheries · Maximum sustainable yield · Species aggregation · Surplus production · Trophodynamic

INTRODUCTION

There is a growing emphasis in using comparative ecosystem studies to foster broader, ecosystem approaches to fisheries management (Murawski et al. 2010). Recognized as a way to provide ad hoc replication that is otherwise impossible owing to the large spatial scale and overall complexity of marine ecosystems, comparative studies have been at the core of international programs such as Global Ocean Ecosystem Dynamics (GLOBEC), European Research on Ocean Ecosystems under Anthropogenic and Natural Forcings (EUR-OCEANS), Indicators for the Seas (IndiSeas) and Comparative Analysis of Marine

Ecosystem Organization (CAMEO) (Megrey et al. 2009). Comparative studies have been used to gain insights on individual species, such as Atlantic cod *Gadus morhua* via ICES/GLOBEC's Cod and Climate program (e.g. Brander 1995, Planque & Fredou 1999, Dutil & Brander 2003, Ratz & Lloret 2003, Drinkwater 2005), and groups of species such as small pelagics (Checkley et al. 2009), as well as whole ecosystems (e.g. Hunt & Drinkwater 2005, Shannon et al. 2008, Gaichas et al. 2009, Link et al. 2009). Comparing similar ecosystems allows us to improve our understanding of and draw generalizations about ecosystem structure (ICES 2001, Megrey et al. 2009, Murawski et al. 2010). The implementation of these generalizations in the form of decision support tools (Kangas et al. 2008) will be an important step in supporting ecosystem-based fisheries management (EBFM; Murawski et al. 2010).

In order to facilitate comparative studies, approaches must be standardized using methods that can be readily applied to existing data. This precludes using complex ecosystem models, which, although powerful, are usually tailored to specific ecosystems. Models of lower complexity can often outperform more complicated models in terms of forecast ability (Costanza & Sklar 1985, Walters 1986, Fulton et al. 2003, Fogarty et al. this volume). Therefore, a preferred solution would be to use simpler models such as surplus production models as the foundation for comparative ecosystem studies. Surplus production models relate a population's production to its current size accounting for fishery removals. This approach relies on readily available data (biomass and landings) and is relatively robust to breach of assumptions (Restrepo et al. 1999). Although surplus production models have their strengths (e.g. limited data demands, ability to generate biological reference points) and weaknesses (e.g. lack of population structure, limitations on time lags), they have generally been deemed useful in fisheries science (Ludwig & Walters 1985, 1989, NRC 1998).

When considering production at the ecosystem level, exploring aggregated properties of biotic communities has value. Current single species reference points are derived from assessments that consider each species in isolation from the ecosystem; however, individual species catches cannot be considered independently in multispecies fisheries as a result of both biological and fishery interactions (McHugh 1959, Murawski 1984, 1991). When such interactions are ignored, the sum of single species maximum sustainable yields (MSYs) is often greater than that sustainable by the ecosystem. Yet, evalua-

tions of overall potential yield from multispecies fisheries assessments suggest that system- or aggregate-level MSY is generally less than the sum of the individual species MSYs (Pope 1975, Brown et al. 1976, May et al. 1979, NEFSC 2008). By aggregating the species within ecosystems into functional groups or aggregate system biomasses, reasonable multi-species equivalents to single species reference points can be obtained (Mueter & Megrey 2006, Sparholt & Cook 2010).

There are multiple reasons for using an aggregated approach to evaluate production at the ecosystem level. First, the energy available from lower trophic levels is limited and shared by the entire suite of living marine resources (LMR), including all fishes and invertebrates. That is, the production available to LMR for any given area of the ocean is constrained by lower trophic level production (Pauly & Christensen 1995, Pauly et al. 1998, 2002). Second, because fish stocks have different productivities, it is often difficult to simultaneously attain single stock objectives in multispecies fisheries (May et al. 1979). Third, there are biological and/or technological interactions that may not always be directly accounted for in single species assessments (Pope 1975, 1979, Fukuda 1976, May et al. 1979, Mayo et al. 1992). However, aggregate models account for all of these interactions without having to explicitly estimate them as in multispecies models (Hollowed et al. 2000, Livingston & Jurado-Molina 2000). In addition, the aggregate approach provides reference points that are familiar to both fishery scientists and LMR managers, but in an ecosystem context.

Here we explore, compare and contrast production model outputs across both various aggregation schemes and multiple ecosystems. This work is part of a hierarchy of studies in which surplus production models were used to undertake comparative ecosystem studies. Holsman et al. (2012, in this Theme Section) present simple surplus production models of cod and herring to examine the potential for biophysical, trophodynamic and exploitative drivers to explain patterns in production. In contrast, Bundy et al. (2012, in this Theme Section) conducted an examination of full system-level aggregate production using surplus production modeling. Here we aggregated species using 3 different aggregation types — habitat, feeding guild and size class — thereby examining processes at a resolution intermediate between Holsman et al. (2012) and Bundy et al. (2012). The aggregation types define functional roles without regard to taxonomy, highlighting important trophodynamic and functional ecological groups and habitat repre-

sentations within the ecosystems (Werner & Gilliam 1984, Ross 1986, Hawkins & Macmahon 1989, Austen et al. 1994, Fogarty & Murawski 1998, Piet et al. 1999, Garrison & Link 2000, Brose et al. 2006). Species within these aggregations should share similar life history traits and respond similarly to perturbations to the system. Interactions within these aggregations tend to result in compensatory dynamics, which are more stable than individual species dynamics (e.g. Duplisea & Blanchard 2005, Auster & Link 2009). Because of this, management based on single species dynamics may not have the desired results and, hence, the need for decision support tools based on broader ecosystem dynamics merits exploration (Walters & Kitchell 2001, Walters et al. 2008).

MATERIALS AND METHODS

The best available information for landings data and biomass estimates were assembled for all commercial species across 12 northern hemisphere ecosystems (Link et al. 2010 and are described in further detail in Fu et al. 2012, in this Theme Section). The ecosystems investigated were the Baltic Sea, Barents Sea, eastern Bering Sea, eastern Scotian Shelf, Georges Bank, Gulf of Alaska, Gulf of Maine, Hecate Strait, North Sea, Norwegian Sea, southern Gulf of St. Lawrence and western Scotian Shelf (see Fig. 2 in Link et al. 2012, in this Theme Section). The longest available time series for both landings and biomass from each ecosystem were used for the analyses

(Table 1). Biomass estimates were acquired from stock assessments, fishery independent surveys corrected for catchability (q) or minimum trawlable swept-area estimates from fisheries independent surveys, depending on the ecosystem and species. Though very infrequent, data for years with missing biomass estimates were linearly interpolated (Fu et al. 2012).

Data were combined based on 3 different species aggregation types within each ecosystem. The first type of aggregation was based on the habitat in which species primarily spend their time foraging: demersal or pelagic habitats. The second type of aggregation was based on specific diet information and species were grouped according to their feeding guild. Originally, there were 7 feeding guilds identified for this project (Link et al. 2010, Fu et al. 2012); however, only 4 of those had sufficient numbers of species across multiple ecosystems for analyses here: planktivores, zoopivores (those animals that consumed zooplankton, shrimp and fish), benthivores and piscivores. The final type of aggregation accounted for size-dependent differences in trophic position. Fish species (excluding invertebrates) were aggregated into 3 size class groups based on the 30th and 70th percentiles of the cumulative frequency histogram of maximum fish lengths across all species. The 3 size categories were: 'small' (<55 cm), 'medium' (≥55 cm and <100 cm) and 'large' (>100 cm).

Aggregate annual surplus production (ASP) was calculated for each aggregate group (agg) and ecosystem (j) as:

Table 1. The 12 northern hemisphere ecosystems analyzed in this study. Length of the time series for each aggregation per ecosystem is listed as well as the physical area of the ecosystem. nd: no data for that aggregation within that ecosystem. Ecosystems are: Baltic Sea (BALT), Barents Sea (BS), eastern Bering Sea (EBS), eastern Scotian Shelf (ESS), Georges Bank (GB), Gulf of Alaska (GoA), Gulf of Maine (GoM), Hecate Strait (HS), North Sea (NORT), Norwegian Sea (NS), southern Gulf of St. Lawrence (GSL), and western Scotian Shelf (WSS)

Region	Area (km²)	Demersal	Pelagic	Planktivore	Zoopivore	Benthivore	Piscivore	Small	Medium	Large
BALT	211 069	1974–2008	1974–2008	1974–2008	nd	nd	1974–2008	1974–2008	nd	1974–2008
BS	747 893	1946–2009	1950–2009	1950–2009	1964–2009	1950–2009	1946–2009	1950–2009	1986–2009	1946–2009
EBS	430 829	1954–2009	1964–2009	1964–2009	nd	1954–2009	1976–2009	1954–2009	1964–2009	1977–2009
ESS	113 704	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008
GB	42 154	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009
GoA	238 439	1950–2009	1950–2009	1950–2009	1960–2009	1977–2009	1950–2009	1961–2009	1961–2009	1950–2009
GoM	76 483	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009	1963–2009
HS	23 501	1984–2009	1951–2009	1951–2009	1984–2009	1984–2009	1984–2009	1951–2009	1984–2009	1984–2009
NORT	609 748	1963–2007	1963–2007	1963–2007	nd	1963–2007	1963–2007	1963–2007	1963–2007	1963–2007
NS	728 331	nd	1950–2009	1950–2009	nd	nd	nd	1950–2009	1972–2009	nd
GSL	74 137	1971–2009	1971–2009	1971–2009	1971–2009	1971–2009	1971–2009	1971–2009	1971–2009	1971–2009
WSS	73 344	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008	1970–2008

$$ASP_{\text{agg},j,t} = B_{\text{agg},j,t+1} - B_{\text{agg},j,t} + C_{\text{agg},j,t} \qquad (1)$$

where $B_{\text{agg},j,t}$ is the total biomass of all species within aggregation in ecosystem j for year t and $C_{\text{agg},j,t}$ is the corresponding total catch. We examined the relationship between ASP and annual biomass by plotting $ASP_{\text{agg},j,t}$ against $B_{\text{agg},j,t}$. We fit both a null model and a Graham-Schaefer surplus production model. The null model assumed that aggregate annual surplus production was linearly related to the aggregate biomass:

$$ASP_t = \beta B_t \qquad (2)$$

where the intercept is 0 (no production at no biomass) and β is the slope of the relationship. While the surplus production model (hereafter the process error model) estimated surplus production as a quadratic function of biomass (Graham-Schaefer functional form with additive error; e.g. Quinn & Deriso 1998) given by:

$$ASP_t = \alpha B_t + \beta B_t^2 + \varepsilon_t \qquad (3)$$

where α and β are regression parameters. The process model assumed deterministic biomass and first-order autocorrelation error structure [$\varepsilon_t = \varphi\varepsilon_{t-1} + V_t$, where $V_t \approx N(0,\sigma^2)$]. Parameters were estimated with a generalized least squares regression using the package 'nlme' in R (v. 2.14.1, R Foundation for Statistical Computing). This model assumes that observations are made without error and that all of the error occurs in the change in population size (process error). Biological reference points (BRPs) were calculated directly from the process error model parameters; MSY = $\alpha^2/4\beta$ and $B_{\text{MSY}} = \alpha/2\beta$.

We compared the 2 models using Akaike's information criterion with correction for small sample size (AIC$_c$; Anderson 2008). Differences in AIC$_c$ between the null and process error models (diffAIC$_c$) were calculated as:

$$\text{diffAIC}_c = \text{AIC}_{c_{\text{null}}} - \text{AIC}_{c_{\text{process}}} \qquad (4)$$

This is similar to ΔAICc, which in typical AIC notation is the difference between a candidate model and the best model, where high ΔAICc values indicate less credibility for the candidate model, with values less than 14 having at least some credibility (Anderson 2008). However, since we used the equation above, we defined diffAICc values of 2 or greater to indicate a better fit by the process error model over the null model, values of between 2 and −2 to indicate equal support for both models and values of −2 or less to indicate a better fit by the null model. Independent of model selection, overall fit of the models to the data was also determined with R^2 value and p-values of regression coefficients.

Finally, comparisons were made across ecosystems by aggregation type. To facilitate comparisons, we standardized the BRPs by the area of the ecosystem (Table 1). Observed variation within aggregate groups was quantified with coefficients of variation (CVs). We then conducted a 2-factor ANOVA to test whether there were significant differences between ecosystems or aggregate groups. Further analysis was done with Tukey's HSD test to make multiple comparisons of means between ecosystems as well as between aggregate groups. Based on these tests, we identified patterns with respect to overall aggregate production across ecosystems and relative production between aggregation types.

RESULTS

There was substantial variation in physical size and fish biomass density among the 12 ecosystems examined. There was a greater than 30 times difference in area between the smallest ecosystem (Hecate Strait, 23 501 km^2) and the largest ecosystem (Barents Sea, 747 893 km^2) (Table 1). Despite these differences, the average annual biomass estimates per unit area from 1984 to 2008 for all aggregate groups were generally ≤13 t km^{-2} (0.001 to 31.200 t km^{-2}); the one consistent exception was the eastern Bering Sea ecosystem, which had high estimates of biomass per unit area for several of the aggregate groups (Fig. 1). For most aggregate groups, the top species (or species group) contributed over 50% of the biomass per unit area (Fig. 1, striated area of bars).

Generally, the process error model was more informative than the null model. Some ecosystems lacked species in certain aggregate groups, leaving a total of 98 combinations of ecosystems and aggregate groups for which comparisons could be made. Of the 98 combinations, there was substantial evidence in favor of the process error model being a better fit in 78 instances (79.6%; Table 2). There was equal support for both models in 17 combinations (17.3%). Substantial support for the null model was found in only 3 instances (3.1%). General fit of the process error model to the data was good with a range of R^2 values from 0.031 to 0.928, while 85.7% and 73.5% of the α and β parameters were significant at the 0.05 level, respectively (Fig. 2, Table S1 in the supplement at www.int-res.com/articles/suppl/m459p219_supp.pdf). Hecate Strait on average had the lowest R^2 values while the Baltic Sea had the highest. For all eco-
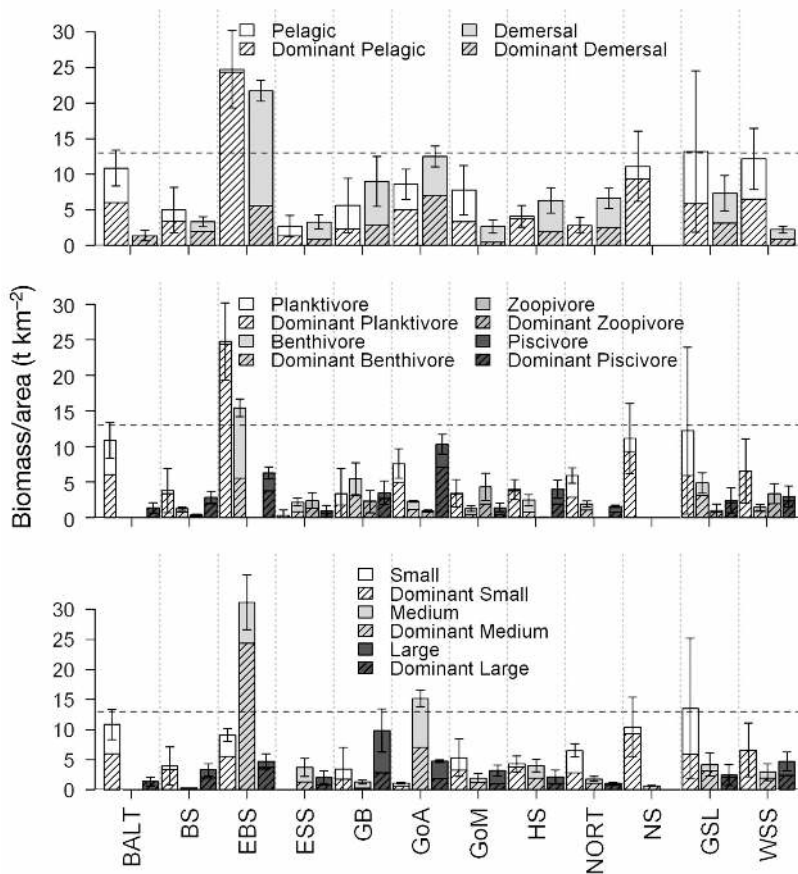
Fig. 1. Average (±SE) biomass per unit area (t km^{-2}) of each aggregation for 12 ecosystems during 1984–2008. Each bar is divided into 2 sections. The bottom striated section of each bar represents the top ranked species or species group (i.e. highest proportion of total biomass). The top section of each bar represents the rest of the species in that group. For reference purposes a horizontal line is shown at 13 t km^{-2}. See Table 1 for definition of ecosystem abbreviations

systems, these R^2 values were generally higher than similar values for the null model.

There were similarities and differences between ecosystems and aggregations with respect to BRPs estimated by the process error model (Figs. 3–5). Estimates of both MSY and B_{MSY} varied between aggregate groups (Table 3). The 2-factor ANOVA showed significant difference between ecosystems and aggregate groups (Table 4). Comparing the differences in means with Tukey's HSD test, we found that most of the differences between ecosystems arose from the 2 systems: the eastern Bering Sea and southern Gulf of St. Lawrence (Tables S2 & S3 in the supplement). Most other ecosystems were not significantly different. As expected, there were more significant differences between aggregate groups (Tables S4 & S5 in the supplement).

Mean values for both MSY and B_{MSY} were larger for the pelagic aggregate group than the demersal aggregate group with values ~1.5 and 1.3 times larger, respectively (Table 3). Estimates of MSY had approximately the same amount of variation for the pelagic and demersal aggregate groups (CVs, 62.9 and 62.2%, respectively; Table 3). There was slightly more vari-

Table 2. Values of diffAIC$_c$ (AIC$_{c null}$ − AIC$_{c process}$; see text for explanation). Negative values indicate where the null model fit the data better than the process model. Values that are not shaded indicate combinations for which the process error model had substantially more support than the null model. Light grey boxes show where there is equal support for both models (−2 < diffAIC$_c$ < 2). Dark grey boxes show where there is substantial support for the null model over the process error model. nd: data missing for individual levels of aggregation within ecosystems. See Table 1 for definition of ecosystem abbreviations

Aggregate group	Ecosystem											
	BALT	BS	EBS	ESS	GB	GoA	GoM	HS	NORT	NS	GSL	WSS
Pelagic	0.7	7.4	12.2	9.9	12.2	13.5	9.3	−1.6	44.3	20.4	14.5	−0.4
Demersal	4.0	20.9	20.2	−1.1	17.8	5.7	10.8	4.4	3.5	nd	5.9	51.8
Planktivore	0.7	4.4	12.2	10.8	13.8	8.8	8.8	−1.7	13.6	20.4	13.3	27.2
Benthivore	nd	16.2	18.9	10.5	13.0	2.3	12.2	2.5	6.3	nd	5.4	23.7
Zoopivore	nd	6.5	nd	11.0	16.4	0.8	17.7	−2.7	nd	nd	4.6	14.2
Piscivore	4.0	29.4	−0.5	6.4	16.5	1.6	14.2	4.8	−1.4	nd	0.7	27.3
Small	0.7	4.8	11.5	−2.9	13.9	−1.2	7.3	−1.8	14.6	20.3	14.5	25.7
Medium	nd	8.9	13.5	13.4	27.6	29.8	15.0	5.1	6.2	27.6	2.8	12.6
Large	4.0	23.2	30.3	−0.1	23.6	−3.7	16.6	1.1	7.9	nd	0.7	31.5
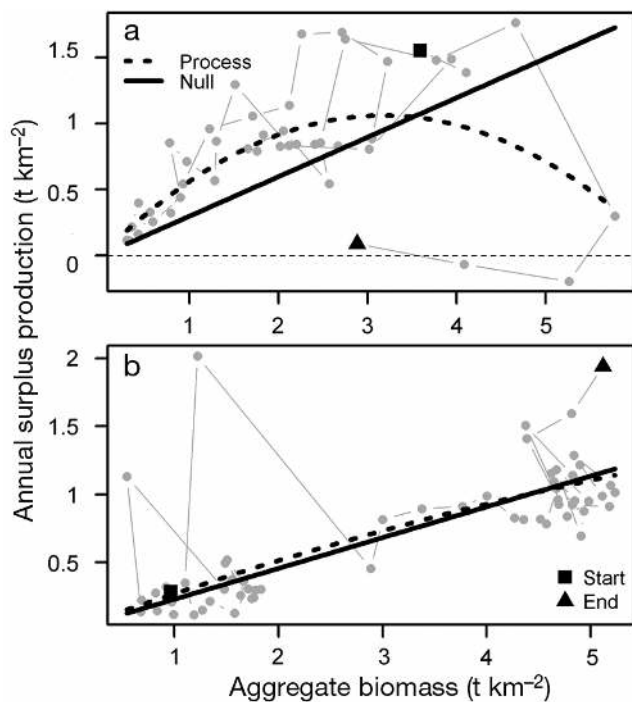
Fig. 2. Examples of the fit by the models to the data. (a) Example where the process error model (thick dashed line) fits the data well (North Sea pelagic aggregate group). This occurs for the majority of the aggregate groups across the ecosystems. (b) Example of where the null model (solid line) fits the data well (Gulf of Alaska 'large' aggregate group). This occurred in only 3 aggregate groups. The thin dashed line shows where annual surplus production equals 0
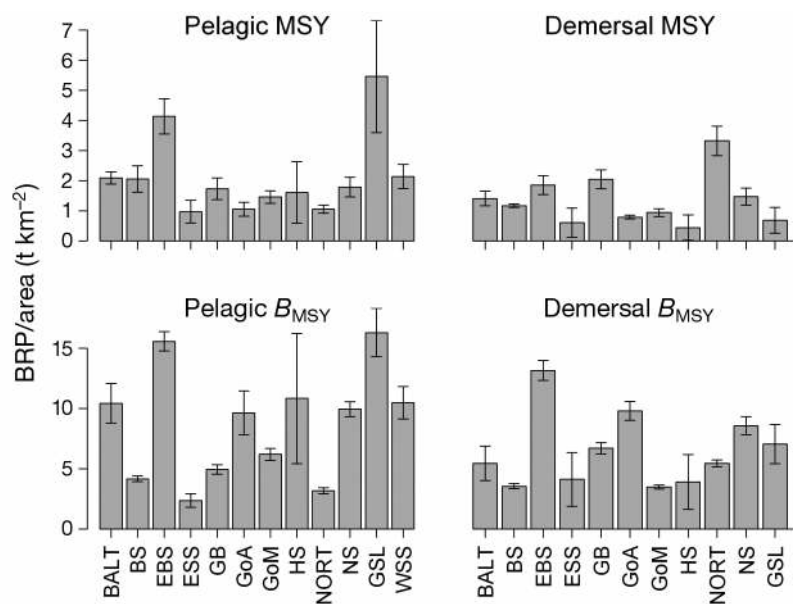

Fig. 3. Area-corrected maximum sustainable yield (MSY) and biomass at maximum sustainable yield (B_{MSY}) derived from the process error model for the habitat aggregation type by ecosystems. See Table 1 for definition of ecosystem abbreviations

ation in the pelagic estimate of B_{MSY} (CV, 52.7%) than the demersal estimate (CV, 46.9%) (Table 3). Estimates of MSY for the pelagic aggregate groups ranged from 0.97 to 5.46 t km^{-2}, while estimates of B_{MSY} ranged from 2.36 to 16.31 t km^{-2} (Table 3). The highest estimates for MSY and B_{MSY} for the pelagic aggregate group were from the eastern Bering Sea and southern Gulf of St. Lawrence (Fig. 3). For both systems, the pelagic productivity was much larger than production from their demersal aggregate groups (2.6 and 4.2 times larger for the eastern Bering Sea and southern Gulf of St. Lawrence, respectively). In contrast, estimates of MSY for the demersal aggregate groups ranged from 0.45 to 3.32 t km^{-2} while estimates of B_{MSY} ranged from 3.49 to 13.16 t km^{-2} (Table 3). The most productive demersal aggregate group was from the North Sea ecosystem. The North Sea's demersal aggregate group was 3.1 times more productive than its pelagic aggregate group (Fig. 3). However, unlike the pelagic aggregate group, the ecosystem with the highest estimate of B_{MSY} (eastern Bering Sea) was not the most productive (Fig. 3).

Similar patterns were observed in the other 2 aggregation types, although CVs for the majority of the aggregate groups within the feeding guild and size aggregation types showed more variation than with the habitat aggregation type (Figs. 3–5, Table 3). Ranges for aggregate groups usually associated with 'forage' fish (pelagic, planktivore and 'small') were similar (Table 3). The 'forage' fish aggregations were also more productive than the other aggregate groups (Table 3). Based on Tukey's HSD test, differences in means were significantly different for all 3 'forage' fish aggregate groups against the benthivore aggregate group. Means were also significantly different for the pelagic and planktivore aggregate groups against the zoopivore, piscivore, medium and large aggregate groups (Tables S4 & S5 in the supplement).

There were also similarities with regards to relative differences in production as well as dominate ecosystems. The greatest variability among functional groups was observed in the feeding guild aggregation type, of which the planktivore aggregate group's mean MSY was 2 to 5 times greater than the other feeding groups (Table 3). Similar to the pelagic aggregate group, the east-
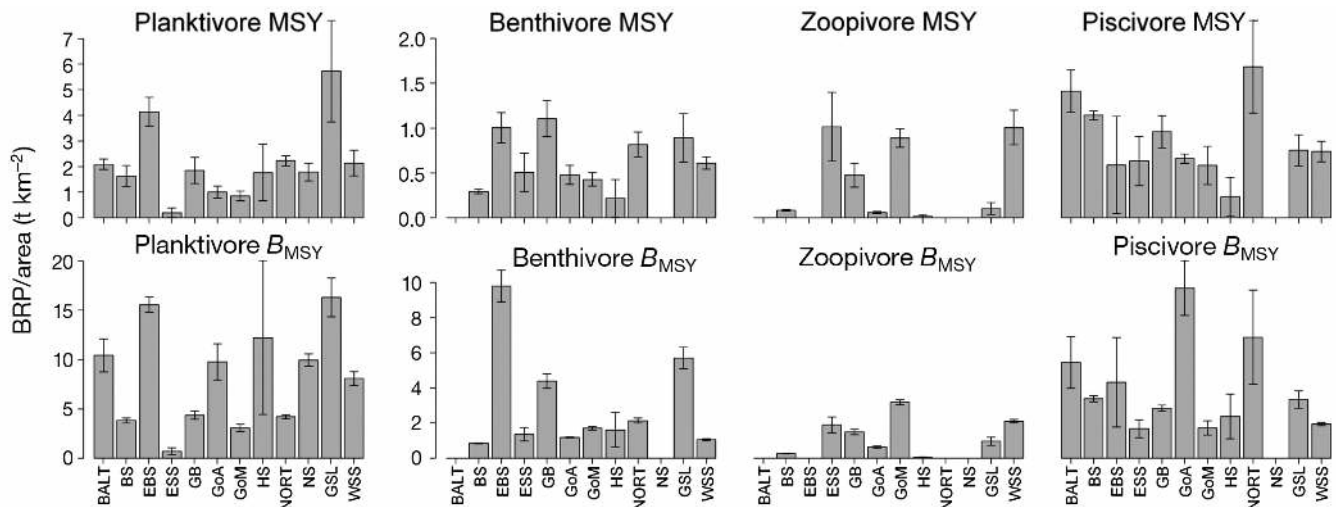
Fig. 4. Area-corrected maximum sustainable yield (MSY) and biomass at maximum sustainable yield (B_{MSY}) derived from the process error model for the feeding guild aggregation type by ecosystems. Note the different scale for the planktivore aggregate group than the other 3 aggregate groups. See Table 1 for definition of ecosystem abbreviations
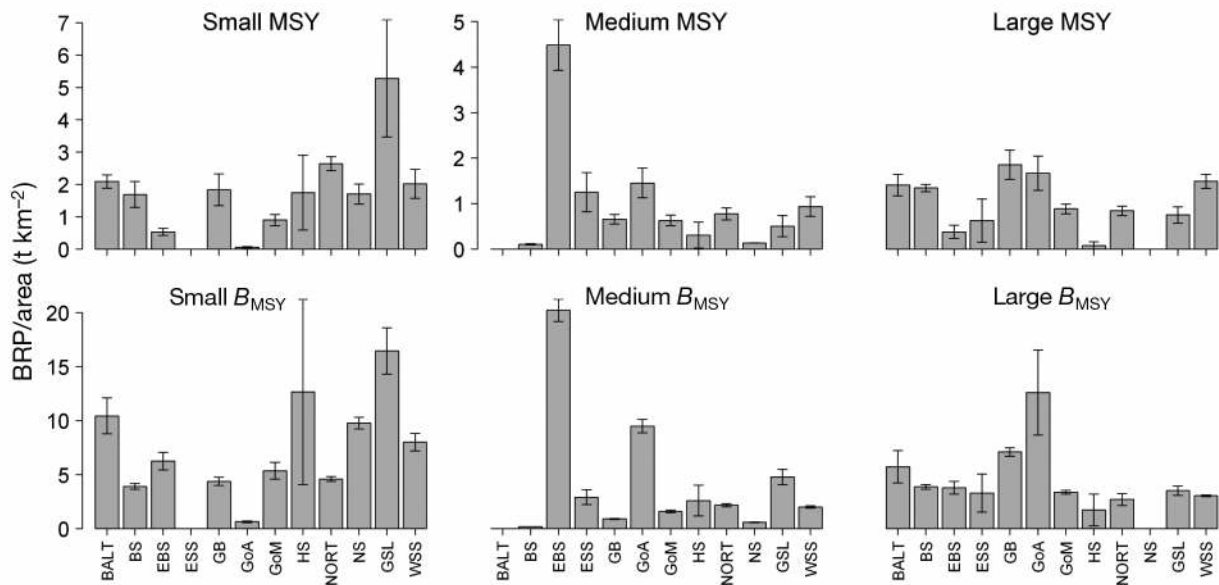

Fig. 5. Area-corrected maximum sustainable yield (MSY) and biomass at maximum sustainable yield (B_{MSY}) derived from the process error model for the size class aggregation type by ecosystems. See Table 1 for definition of ecosystem abbreviations

ern Bering Sea and southern Gulf of St. Lawrence were the most productive (Figs. 4 & 5). For both systems, the most productive aggregate group within the feeding guild aggregation type was the planktivore aggregate group (Fig. 4). However, for the size aggregation, production was higher within the 'medium' size aggregate group for eastern Bering Sea and not the 'small' aggregate group (Fig. 5).

Despite physical differences among the ecosystems, estimated BRPs appear to be independent of ecosystem size. Specifically, the Barents Sea, which was the largest ecosystem at ~748 000 km^2, had BRPs that were generally intermediate to other systems. Whereas the relatively small ecosystem, Georges Bank, had high demersal area-corrected MSY, which may be related to its shallow average depth and rel-

Table 3. Measures of central tendencies for estimated biological reference points obtained for the aggregate groups. Values were derived using the process model. Ecosystems without data for an aggregate group were excluded from calculations

Aggregate group	Biological reference point									
	MSY_{mean}	MSY_{median}	MSY_{SD}	MSY_{range}	MSY_{CV}	$B_{MSY_{mean}}$	$B_{MSY_{median}}$	$B_{MSY_{SD}}$	$B_{MSY_{range}}$	$B_{MSY_{CV}}$
Pelagic	2.13	1.76	1.34	0.97–5.46	62.9	8.68	9.80	4.57	2.36–16.31	52.7
Demersal	1.34	1.17	0.83	0.45–3.32	62.2	6.48	5.45	3.04	3.49–13.16	46.9
Planktivore	2.12	1.82	1.48	0.20–5.73	70.1	8.21	8.93	5.04	0.71–16.33	61.3
Benthivore	0.63	0.56	0.30	0.22–1.11	48.0	2.98	1.66	2.87	0.85–9.79	96.2
Zoopivore	0.46	0.29	0.45	0.01–1.01	98.9	1.33	1.23	1.05	0.04–3.18	79.3
Piscivore	0.85	0.74	0.42	0.23–1.68	48.8	3.96	3.34	2.49	1.67–9.67	62.9
Small	1.71	1.73	1.40	0.01–5.28	81.9	6.88	5.81	4.83	0.01–16.46	70.2
Medium	1.02	0.66	1.22	0.11–4.48	119.7	4.10	2.06	5.57	0.18–19.17	135.8
Large	1.03	0.88	0.56	0.08–1.86	54.7	4.38	3.34	2.87	1.65–11.96	65.5

Table 4. Results of 2-factor ANOVAs for MSY and B_{MSY}. The 2 factors were ecosystem and aggregate group (see Table 1). ***Factor significant at p < 0.001

	df	SS	MS	F-value	Pr(>F)
MSY					
Ecosystem	11	34.59	3.144	5.040	8.04×10^7 ***
Agg. group	8	37.63	4.704	7.540	1.22×10^8 ***
Residuals	176	109.79	0.624		
B_{MSY}					
Ecosystem	11	653.3	59.39	5.689	1.20×10^6 ***
Agg. group	8	433.0	54.12	5.184	3.33×10^5 ***
Residuals	78	814.3	10.44		

atively high primary production (O'Reilly & Zetlin 1998). There did not appear to be a trend in BRPs with regard to ocean basin or latitude (data not shown).

DISCUSSION

We have shown that, with a few exceptions, estimated aggregate group BRPs are relatively invariant among ecosystems when data are aggregated to reflect habitat, trophodynamic or allometric affinities. While aggregate production varied across aggregation types and ecosystems, MSY never exceeded 6 t km^{-2} and was generally less than 3 t km^{-2}. The comparative ecosystem approach is important for revealing commonalities and differences across ecosystems. Yet, to date, the bulk of scientific literature on ecosystems is composed of 'within ecosystem' comparisons, which focus on comparing different sections of various time series (Murawski et al. 2010). 'Within ecosystem' comparisons are helpful in determining covariation among changes in species pro-

duction dynamics and regime shifts within particular ecosystems but do not often identify broader, overarching ecosystem patterns and functions. Our study revealed insights from a broader intersystem comparison (Murawski et al. 2010) within the temperate northern hemisphere ecosystems. Although not replicates in traditional statistical parlance, all of our ecosystems are similar in that they are highly productive and have sustained long-term fisheries removals (Fu et al. 2012). Through applying surplus production models to several types of functional aggregations, we can draw some generalities that are applicable to a wide range of ecosystems.

Previous studies have demonstrated that developing estimates of aggregate ecosystem MSY is a more conservative management approach than summing estimates from multiple single-species models (Pope 1975, Brown et al. 1976, May et al. 1979, Collie & Gislason 2001, Walters et al. 2005, Mueter & Megrey 2006, Sparholt & Cook 2010) This occurs because within any given aggregation, it is assumed that as individual species compete with one another resources become limited and each species cannot be maintained at carrying capacity simultaneously (Gamble & Link 2009). In addition, the functional aggregate group is less annually variable than individual species as perturbations in population dynamics are smoothed. Moreover, Bundy et al. (2012) show that aggregating at the full system level is even more conservative (1 to 5 t km^{-2}) than aggregating at the functional group level, as presented here.

There was no apparent pattern in aggregate productivity with respect to ocean basin, region or component species. For many different aggregation types the southern Gulf of St. Lawrence and the eastern Bering Sea were the most productive ecosystems. The southern Gulf of St. Lawrence was located within the Atlantic basin at the mid-range of our

study geographically while the eastern Bering Sea was the most northerly of the Pacific basin eco-systems examined. The high production from both can be attributed to the pelagic and planktivore aggregate groups. But there were differences in the most productive size class aggregation, with the southern Gulf of St. Lawrence dominated by the 'small' size class and the eastern Bering Sea domi-nated by the 'medium' size class. Both the southern Gulf of St. Lawrence and eastern Bering Sea are rel-atively shallow with seasonal ice cover; however, the southern Gulf of St. Lawrence is an inland sea con-taining a mixture of estuarine and marine, as well as subtropical to subarctic species (Fu et al. 2012).

We found that aggregations typically associated with 'forage fish' (pelagic, planktivore and 'small') were consistently more productive than other aggre-gation groups. 'Forage fish' play a key ecological role within marine ecosystems, transferring energy from primary production to upper trophic levels (Pikitch et al. 2012). When dealing with single species manage-ment, BRPs may be overestimated with respect to MSY or underestimated with respect to B_{MSY} without explicitly accounting for this higher natural mortality (Tyrrell et al. 2011). However, as noted previously, by aggregating species we indirectly account for these interactions while simultaneously dampening single species variability; thus, aggregate 'forage' BRPs may be a useful alternative. It seems reasonable that special attention should be afforded these aggregate 'forage' groups as trophic pressures may increase as predator stock sizes continue to rebuild (Overholtz et al. 2008, Smith et al. 2011).

To utilize information from comparative studies, outputs should be familiar to managers. Therefore, we provide BRP estimates derived from aggregates rather than single species. In fisheries science, BRPs are benchmarks from which to assess the status of fish stocks. In single-species assessment, they are usually based on fishing mortality or biomass levels. As we transition to EBFM, the concept of reference points will need to be expanded to include other eco-logical indicators and multivariate indices (Link 2005). However, a first step is to apply more familiar BRPs, such as MSY and B_{MSY}, in a context broader than single species such as the aggregate reference points presented here. While we assume that LMR managers could interpret other metrics, the familiar-ity with more traditional BRPs may aid in the transi-tion to other EBFM reference points. In addition to familiarity, the value of aggregated metrics can be high. A major criticism of MSY is its appropriateness as a management goal (Punt & Smith 2001). This is based on the static nature of MSY, which ignores nat-ural fluctuations in species abundance. By aggregat-ing species into similar life-history categories some of this natural fluctuation may be dampened. This could lead to more robust yet conservative estimates of pro-ductivity.

Conceivable uses of such aggregated metrics war-rant consideration. We could envision them being adopted in the following way. First, aggregate BRPs should be used as overall caps to fishery removals with respect to aggregate groups. This proposed usage conforms to the recent paradigm of using MSY as a limit rather than a target and has had some precedence in some ecosystems (e.g. Witherell et al. 2000). Aggregated metrics can also balance yield and biodiversity objectives (Worm et al. 2009, Gaichas et al. this volume). Species within aggregate groups should generally have similar life histories; however, some are more productive than others. To optimize biodiversity objectives, removals should be distributed within the aggregation with some consid-eration of individual species productivity. This way, lower productive stocks will be safeguarded against overfishing (Mueter & Megrey 2006). This also cre-ates a bridge between single species assessments and multispecies assessment that can better inform managers of potential risks of particular manage-ment decisions (Mace 2001).

As fishery management transitions towards an eco-system approach, comparative studies like this one will provide useful insights. Comparative studies allow for ad hoc replication between similar eco-systems. The aggregation schemes we employed were based on various criteria, but there does not ap-pear to be a difference in using habitat, feeding guild or size class. Future choices regarding the best ag-gregation to use for developing decision support tools will certainly need to be based on individual ecosystem history and species composition as well as the management question being addressed, but we assert that the general patterns we observed should be informative for such selections. The logical next step would be to move to the final level of compara-tive hierarchy outline by Murawski et al. (2010). That would be a global comparison between dissimilar ecosystems. By comparing the contrast between eco-systems from temperate marine ecosystems to coral reef ecosystems or even terrestrial ecosystems we can begin to answer the broadest of ecosystem ques-tions related to biodiversity, variability and produc-tivity (Murawski et al. 2010). This work represents a useful step to that end and ultimately a global syn-thesis of the determinants of fisheries production.

LITERATURE CITED

Anderson DR (2008) Model based inference in the life sciences: a primer on evidence. Springer, New York, NY

Austen DJ, Bayley PB, Menzel BW (1994) Importance of the guild concept to fisheries research and management. Fisheries 19(6):12–18

Auster PJ, Link JS (2009) Compensation and recovery of feeding guilds in a northwest Atlantic shelf fish community. Mar Ecol Prog Ser 382:163–172

Brander KM (1995) The effect of temperature on growth of Atlantic cod (*Gadus morhua* L). ICES J Mar Sci 52:1–10

Brose U, Jonsson T, Berlow EL, Warren P and others (2006) Consumer-resource body-size relationships in natural food webs. Ecology 87:2411–2417

Brown BE, Brennan JA, Grosslein MD, Heyerdahl EG, Hennemuth RC (1976) The effect of fishing on the marine finfish biomass in the Northwest Atlantic from the Gulf of Maine to Cape Hatteras. Int Comm Northwest Atl Fish Res Bull 12:49–68

Bundy A, Bohaboy EC, Hjermann DO, Mueter FJ, Fu C, Link JS (2012) Common patterns, common drivers: comparative analysis of aggregate surplus production across ecosystems. Mar Ecol Prog Ser 459:203–218

Checkley D, Roy C, Alheit J, Oozeki Y (eds) (2009) Climate change and small pelagic fish. Cambridge University Press, Cambridge

Collie JS, Gislason H (2001) Biological reference points for fish stocks in a multispecies context. Can J Fish Aquat Sci 58:2167–2176

Costanza R, Sklar FH (1985) Articulation, accuracy, and effectiveness of ecological models: a review of freshwater wetland applications. Ecol Model 27:45–68

Drinkwater KF (2005) The response of Atlantic cod (*Gadus morhua*) to future climate change. ICES J Mar Sci 62: 1327–1337

Duplisea DE, Blanchard F (2005) Relating species and community dynamics in an heavily exploited marine fish community. Ecosystems 8:899–910

Dutil JD, Brander K (2003) Comparing productivity of North Atlantic cod (*Gadus morhua*) stocks and limits to growth production. Fish Oceanogr 12:502–512

Fogarty MJ, Murawski SA (1998) Large-scale disturbance and the structure of marine systems: fishery impacts on Georges Bank. Ecol Appl 8(Suppl):S6–S22

Fogarty MJ, Overholtz WJ, Link JS (2012) Aggregate surplus production models for demersal fishery resources of the Gulf of Maine. Mar Ecol Prog Ser 459:247–258

Fu C, Gaichas S, Link JS, Bundy A and others (2012) Relative importance of fisheries, trophodynamic and environmental drivers in a series of marine ecosystems. Mar Ecol Prog Ser 459:169–184

Fukuda Y (1976) A note on yield allocation in multi-species fisheries. ICNAF Res Bull 12:83–87

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

Gaichas S, Skaret G, Falk-Petersen J, Link JS and others (2009) A comparison of community and trophic structure in five marine ecosystems based on energy budgets and system metrics. Prog Oceanogr 81:47–62

Gaichas S, Gamble R, Fogarty M, Benoît H and others (2012) Assembly rules for aggregate-species production models: simulations in support of management strategy evaluation. Mar Ecol Prog Ser 459:275–292

Gamble RJ, Link JS (2009) Analyzing the tradeoffs among ecological and fishing effects on an example fish community: a multispecies (fisheries) production model. Ecol Model 220:2570–2582

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

Hawkins CP, Macmahon JA (1989) Guilds: the multiple meanings of a concept. Annu Rev Entomol 34:423–451

Hollowed AB, Bax N, Beamish R, Collie J and others (2000) Are multispecies models an improvement on single-species models for measuring fishing impacts on marine ecosystems? ICES J Mar Sci 57:707–719

Holsman KK, Essington T, Miller TJ, Koen-Alonso M, Stockhausen WJ (2012) Comparative analysis of cod and herring production dynamics across 13 northern hemisphere marine ecosystems. Mar Ecol Prog Ser 459:231–246

Hunt GL, Drinkwater KF (eds) (2005) Ecosystem studies of sub-Arctic seas (ESSAS) science plan. Global Ocean Ecosyst Dynam (GLOBEC) Rep No. 19

ICES (International Council for the Exploration of the Sea) (2001) Report of the planning group on comparing the structure of marine ecosystems in the ICES Area, Copenhagen, Denmark, 26–29 June 2001. ICES Document CM 2001/G: 03. ICES, Copenhagen

Kangas AJ, Kangas J, Kurttila M (2008) Decision support for forest management. Springer Ser Manag For Ecosyst Vol 16. Springer, Berlin

Link JS, Gaichas S, Miller TJ, Essington T and others (2012) Synthesizing lessons learned from comparing fisheries production in 13 northern hemisphere ecosystems: emergent fundamental features. Mar Ecol Prog Ser 459: 293–302

Link JS (2005) Translating ecosystem indicators into decision criteria. ICES J Mar Sci 62:569–576

Link JS, Stockhausen WT, Skaret G, Overholtz W and others (2009) A comparison of biological trends from four marine ecosystems: synchronies, differences, and commonalities. Prog Oceanogr 81:29–46

Link JS, Megrey BA, Miller TJ, Essington T and others (2010) Comparative analysis of marine ecosystems: international production modelling workshop. Biol Lett 6:723–726

Livingston PA, Jurado-Molina J (2000) A multispecies virtual population analysis of the eastern Bering Sea. ICES J Mar Sci 57:294−299

Ludwig D, Walters CJ (1985) Are age-structured models appropriate for catch-effort data? Can J Fish Aquat Sci 42:1066−1072

Ludwig D, Walters CJ (1989) A robust method for parameter estimation from catch and effort data. Can J Fish Aquat Sci 46:137−144

Mace PM (2001) A new role for MSY in single-species and ecosystem approaches to fisheries stock assessment and management. Fish Fish 2:2−32

May RM, Beddington JR, Clark CW, Holt SJ, Laws RM (1979) Management of multispecies fisheries. Science 205:267−277

Mayo RK, Fogarty MJ, Serchuk FM (1992) Aggregate fish biomass and yield on Georges Bank, 1960−87. J Northwest Atl Fish Sci 14:59−78

McHugh JL (1959) Can we manage our Atlantic coastal fishery resources? Trans Am Fish Soc 88:105−110

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

Mueter FJ, Megrey BA (2006) Using multi-species surplus production models to estimate ecosystem-level maximum sustainable yields. Fish Res 81:189−201

Murawski SA (1984) Mixed-species yield-per-recruitment analyses accounting for technological interactions. Can J Fish Aquat Sci 41:897−916

Murawski SA (1991) Can we manage our multispecies fisheries? Fisheries 16(5):5−13

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

NEFSC (Northeast Fisheries Science Center) (2008) Assessment of 19 northeast groundfish stocks through 2007. Report of the 3rd groundfish assessment review meeting (GARM III). Ref Doc 08-15. NOAA Natl Mar Fish Serv Northeast Fish Sci Cent, Woods Hole, MA

NRC (National Research Council) (1998) Improving fish stock assessment. National Academy Press, Washington, DC

O'Reilly JE, Zetlin C (1998) Seasonal, horizontal, and vertical distribution of phytoplankton chlorophyll a in the northeast U.S. continental shelf ecosystem. NOAA Technical Report NMFS 139. US Dept of Commerce, Seattle, WA

Overholtz WJ, Jacobson LD, Link JS (2008) An ecosystem approach for assessment advice and biological reference points for the Gulf of Maine−Georges Bank Atlantic herring complex. N Am J Fish Manag 28:247−257

Pauly D, Christensen V (1995) Primary production required to sustain global fisheries. Nature 374:255−257

Pauly D, Christensen V, Dalsgaard J, Froese R, Torres F Jr (1998) Fishing down marine food webs. Science 279: 860−863

Pauly D, Christensen V, Guenette S, Pitcher TJ (2002) Towards sustainability in world fisheries. Nature 418: 689−695

Piet GJ, Pet JS, Guruge W, Vijverberg J, Densen WLTV (1999) Resource partitioning along three niche dimensions in a size-structured tropical fish assemblage. Can J Fish Aquat Sci 56:1241−1254

Pikitch E, Boersma PD, Boyd IL, Conover DO and others (2012) Little fish, big impact: managing a crucial link in ocean food webs. Lenfest Ocean Program, Washington, DC

Planque B, Fredou T (1999) Temperature and the recruitment of Atlantic cod (*Gadus morhua*). Can J Fish Aquat Sci 56:2069−2077

Pope JG (1975) The effect of biological interaction on the theory of mixed fisheries. Int Comm Northwest Atl Fish Sel Pap 1:157−162

Pope JG (1979) Stock assessment in multistock fisheries, with special reference to the trawl fishery in the Gulf of Thailand. SCS/DEV/79/19. FAO, Manila

Punt AE, Smith ADM (2001) The gospel of maximum sustainable yield in fisheries management: birth, crucifixion and reincarnation. In: Reynolds JD, Mace GM, Redford KR, Robinson JR (eds) Conservation of exploited species. Cambridge University Press, Cambridge, p 41−66

Quinn TP, Deriso RB (1998) Quantitative fish dynamics. Oxford University Press, Oxford

Ratz HJ, Lloret J (2003) Variation in fish condition between Atlantic cod (*Gadus morhua*) stocks, the effect on their productivity and management implications. Fish Res 60: 369−380

Restrepo VR, Mace PM, Serchuk FM (1999) The precautionary approach: a new paradigm or business as usual? In: NOAA Tech Memorandum NMFS-F/SPO-41. US Dep Comm, Washington, DC, p 61−70

Ross ST (1986) Resource partitioning in fish assemblages: a review of field studies. Copeia 1986:352−368

Shannon LJ, Neira S, Taylor M (2008) Comparing internal and external drivers in the southern Benguela and the southern and northern Humboldt upwelling ecosystems. Afr J Mar Sci 30:63−84

Smith ADM, Brown CJ, Bulman CM, Fulton EA and others (2011) Impacts of fishing low-trophic level species on marine ecosystems. Science 333:1147−1150

Sparholt H, Cook RM (2010) Sustainable exploitation of temperate fish stocks. Biol Lett 6:124−127

Tyrrell MC, Link JS, Moustahfid H (2011) The importance of including predation in fish population models: implications for biological reference points. Fish Res 108:1−8

Walters C (1986) Adaptive management of renewable resources. Macmillan, New York, NY

Walters C, Kitchell JF (2001) Cultivation/depensation effects on juvenile survival and recruitment: implications for the theory of fishing. Can J Fish Aquat Sci 58:39−50

Walters CJ, Christensen V, Martell SJ, Kitchell JF (2005) Possible ecosystem impacts of applying MSY policies from single-species assessment. ICES J Mar Sci 62: 558−568

Walters CJ, Hilborn R, Christensen V (2008) Surplus production dynamics in declining and recovering fish populations. Can J Fish Aquat Sci 65:2536−2551

Werner EE, Gilliam JF (1984) The ontogenetic niche and species interactions in size-structured populations. Annu Rev Ecol Syst 15:393−425

Witherell D, Pautzke C, Fluharty D (2000) An ecosystem-based approach for Alaska groundfish fisheries. ICES J Mar Sci 57:771−777

Worm B, Hilborn R, Baum JK, Branch TA and others (2009) Rebuilding global fisheries. Science 325:578−585