New York Bight Indicator Report 2020
Part 1: Indicator Development and Selection

MOU #AM10560 NYS DEC & SUNY Stony Brook

Report contributors: Kurt Heim, Senior Researcher; Janet Nye, Associate Professor; Lesley Thorne, Assistant Professor; Joe Warren, Associate Professor; Charlie Flagg, Research Professor; Tyler Menz, Carbonate Chemistry Technician

Contact Principal Investigator, Janet A. Nye with questions or comments
janet.nye@stonybrook.edu

1. Background

Ecosystem Based Management (EBM) is "an integrated approach to management that considers the entire ecosystem, including humans" when making management and policy decisions (McLeod et al. 2005, Rosenberg and McLeod 2005). EBM is a deliberate movement away from the traditional management approach of considering only one species or one activity in isolation.

Figure 1: Location and extent of the New York Bight ecosystem.

In 2017, New York State released its Ocean Action Plan (OAP), a bold blueprint to address the many environmental stressors on the New York Bight ecosystem and to produce guidance for developing a multi-disciplinary, adaptive approach to managing coastal and offshore waters of the New York Bight (Figure 1). The OAP highlights the need to regularly assess the status of marine resources, the interactions amongst a broad range of taxa, and those environmental factors that affect them. Monitoring ecosystem health and the services provided by the ecosystem is crucial for assessing and understanding the current state of the ocean. Action 33 of the NY OAP states a need to "Develop an ocean indicators system for the New York Bight". As such, in late 2017, the New York State Department of Conservation (NYS DEC) initiated an interdisciplinary, multi-trophic level ocean monitoring in the New York Bight in order to provide information on the status of New York pelagic resources to managers. A primary objective of this work was to develop a system of indicators that will allow resource managers to better evaluate risks posed by human activities and natural processes; to track the effectiveness of management actions taken to maintain ecosystems in a healthy, productive and resilient condition; to anticipate the effects of climate variability and climate change; and to better inform decision making regionally and locally. This report documents the work to establish this indicator system and provides a baseline for evaluating conservation efforts in ocean waters.

The work presented here started in 2016 with a workshop funded by NYS DEC and led by researchers at the School of Marine and Atmospheric Sciences (SOMAS). The workshop served to scope scientists in the region on the most important indicators to develop and helped to inform the selection of indicators presented here. Here we present the progress in the development of indicators for the New York Bight from 2018-2020.

2. Why developing a system of indicators specific to the New York Bight is essential

EBM is not a new concept and was formally defined in the early 2000s (McLeod and Leslie 2009). Globally, many institutions have been working to implement EBM since that time. As part of implementing EBM, indicator development and ecosystem assessments (EAs) have been an integral part of this process. Indicators are quantitative measurements that represent key attributes of interest. When a vetted suite of indicators is assembled, it can be used to assess ecosystem status, drivers of change, and performance of management actions (Rice and Rochet 2005; Shin and Shannon 2010; Shin et al. 2010). In the United States, NOAA has taken the lead in ecosystem assessment for the broad regions of the ocean it manages and is rapidly moving towards full Integrated Ecosystem Assessments (Levin et al. 2009) for all large marine ecosystems in the US. For the Northeast US, the first EA was completed in 2009 (EcoAp 2009) and now the NEFSC releases an annual EA for the New England Region and the Mid-Atlantic Bight (MAB). The New York Bight lies just south of the New England region and makes up the northern portion of the Mid-Atlantic Bight, which also includes the continental shelf all the way down to Cape Hatteras, NC. Therefore, to meet the specific goals of the OAP and the needs of the people of New York state, indicators specific to the New York Bight, an important transition area between New England and Mid-Atlantic are critically needed.

As we developed indicators specific to the NYB, we also critically assessed the need to develop indicators in the Northeast US for different spatial units. A detailed draft manuscript of this analysis can be found in Appendix 1 and is summarized here. We systematically divided the Northeast US ecosystem in 31 different ways, covering spatial extents from 250,000 km^2 (the whole Northeast US) to 20,000 km^2 (ten subunits). The same 22 indicators were calculated for each unit, assessed for trends, and evaluated as 31 independent EAs. We found that the detected signals of ecosystem change depended on the way the ecosystem was defined or subdivided. Some indicators displayed similar trends among the different spatial units (i.e., a consistent spatial pattern in trend), while others showed very different trends among spatial units. Most commonly, a strong local trend, occurring only in a small region, could be perceived as widespread when data were aggregated over large regions giving a false impression of the spatial region over which an ecosystem trend was occurring. Conversely, an important local trend can be masked by aggregating data over too large on area.

This study in general showed that the trends depicted by indicators (e.g., increasing, decreasing, stable) will change as the boundaries defining the 'ecosystem' or 'study region' change. Thus, developing indicators specific to the NYB is well-justified. This study emphasizes the importance of examining indicators at different spatial scales, especially when they represent processes with high spatial heterogeneity. Therefore, the ecosystem indicator system developed for New York state is complimented by existing efforts by NOAA to produce EAs for the Mid-Atlantic and New England. Considering indicators at multiple scales is particularly important because management decisions occur at multiple spatial scales (e.g. county, state, regional, federal) with different goals and objectives relevant at different scales (Heim et al. in review).

3. A Framework is needed to organize and prioritize indicators

After a review of existing datasets in the NYB and other indicator-based ecosystem assessments (Appendix 2), it became clear that a framework was needed to organize and select indicators for the NY Bight Ocean Indicators project. This initial data exploration revealed a large number of potential indicators (~ 200) for the NYB using datasets from a range of fields (oceanography, fisheries biology, sociology, etc.). This review was guided by regular input from DEC, examination of indicators used by NEFSC in Ecosystem Status Reports (Link et al. 2002; EcoAP 2009, 2011; NEFSC 2017, 2018, 2019), and the subject matter expertise of SOMAS collaborators. Two important conclusions from this report and following discussions emerged, that both highlighted the need for an organization system to move forward.

First, some indicators are straightforward to calculate whereas others take an exceptional amount of work and subject matter expertise. For instance, calculating seasonal sea surface temperature for a given region (i.e., the New York State focus area, Figure 1 of the OAP) is a relatively simple task using satellite derived sea surface temperature data (Reynolds et al. 2007). In contrast, other indicators of interest to DEC such as the seasonal development of the Mid-Atlantic cold pool, or spatial habitat use of marine mammals, would take extensive time and effort to develop. Prioritizing specific indicator needs is necessary to allocate team member effort to best meet the needs of management.

Second, there is a pressing need for an organizational framework to use ecosystem indicators in a meaningful way. Even if we could develop all possible indicators of interest, interpreting and presenting them to the public in a meaningful way must be tractable. We could not possibly present 200+ indicators. While the first report was indeed useful to explore data availability and potential indicators, it represented a large data gathering effort without a well-defined understanding of exactly what was needed. Indicators serve a variety of specific purposes in the context of environmental management and EBM (Turnhout et al. 2007), and so choosing a suite of indicators should be done to reflect a specific purpose. Moreover, it is generally agreed that 'less is more' when it comes to the number of indicators to use (Link 2010).

Thus, there is an explicit need to (1) develop an organization and prioritization system for indicators and (2) develop a framework to reduce the quantity of indicators to a smaller subset that still adequately portrays the important components of interest in the ecosystem (Rice and Rochet 2005; Kershner et al. 2010). In light of these needs for a framework to guide further work, a literature review was conducted to explore options.

4. Review of example frameworks

Existing indicator systems and the published scientific literature were reviewed to explore options for an organization framework to guide the NYB ocean indicators system. Existing indicator systems reviewed included the Puget Sound Partnership Sound Indicators, the Long Island Sound Study Indicators, NEFSCs approach to EBM (nationally, and specifically in the Mid-Atlantic Bight), the Ocean Health Index, and the European Union's Marine Strategy Framework Directive.

A commonality of these systems is that each uses discrete categories to organize indicators, either based on ecosystem components or ecosystem objectives (or both). Ecosystem components are important socio-ecological characteristics of the ecosystem (CMP 2007), while objectives are explicit statements for a desired ecosystem state or ecosystem service. The Long Island Sound Study indicators framework (https://longislandsoundstudy.net) is one that revolves predominantly around specified ecosystem objectives and targets. An example target is to 'measurably reduce the area of hypoxia in the Long Island Sound from pre-2000 Dissolved Oxygen TMDL averages…". This clearly defined quantitative target provides the context for a specific indicator need, which can be developed and subsequently used to track progress. With a suite of such targets, filling in indicators linked to each target is relatively straightforward.

An alternative framework is used by the Puget Sound Partnership that instead links indicators to specified processes that are recognized as being important to ecosystem structure and function (Kershner et al. 2011; WSAS 2012). In this approach, a conceptual model of ecosystem dynamics is developed a-priori, including a hierarchy of components to capture a complete picture of the ecosystem. Indicators are developed to represent each of these processes, which provides a means to better understand the complex interactions that lead to desired outcomes for the ecosystem.

The NEFSC develops indicators in the context of an integrated ecosystem assessment (Levin et al. 2009; 2013) that involves both stated ecosystem objectives as well as a conceptual model of the ecosystem. An integrated ecosystem assessment is a structured framework to implement EBM that involves discrete steps in a feedback loop (i.e., 1: Define EBM goals and targets, 2: Develop indicators, 3: Assess ecosystem, 4: Analyze uncertainty and risk, 5: Evaluate strategies). These steps are taken to identify appropriate management actions, which can be assessed for effectiveness using the indicators. Each step is appropriately updated based on new information learned and changing stakeholder needs. The indicators chosen for the Mid-Atlantic Bight are organized explicitly around a conceptual ecosystem model, while collectively addressing a series of stated ecosystem objectives (e.g., optimize revenue, optimize recreation) based on existing federal regulatory mandates (DePiper et al. 2017). In this sense, indicators are not linked one-to-one to objectives. Instead, a variety of indicators may be pertinent to a given objective. This approach is comprehensive and involves a large team of scientists to administer, which highlights another important component of this review.

The overall time commitment by scientists and funding level of these programs is substantial and is important to consider as we develop a manageable framework for our own use. The Long Island Sound study operates with a 14-million-dollar annual budget. The NEFSC status report is supported by a full-time staff of 8, plus 3 contractors or post-docs, and each year their main product (an ESR) involves contributions from over 30 people. The Puget Sound Partnership was funded at 269 million dollars from 2006 to 2016. As we contemplated a framework for moving forward with our work, it was important to explicitly consider that most programs reviewed involve a substantially larger team of researchers and level of funding. An approach was needed that drew from the ‘best practices’ outlined in these examples and literature reviewed (e.g. Kershner et al. 2011; Rice and Rochet 2006) while also being realistic in terms of our workload capacity.

5. Selected framework for NYB indicator system

Our review of indicator systems revealed that there is no single ‘right answer’ as to how an indicators framework should be developed. Since our work was motivated primarily by Goal 1 of the NY Ocean Action Plan to “Ensure the ecological integrity of the ocean ecosystem” with a subgoal of “evaluate the ecological integrity of the ocean ecosystem of New York” we choose to take an approach similar to the Puget Sound Partnership to develop an indicator system that broadly assesses ecosystem health of the NYB. First, we developed a very simple model of the ecosystem as is recommended in many publications (CMP 2007; Kershner et al. 2011; WSAS 2012) to ensure important processes in the ecosystem are not overlooked. Moreover, by expanding our indicator framework beyond only 'status indicators' linked to explicit targets, we may gain a deeper understanding of linkages and the *suite of processes* that interact to generate a specific outcome. This cannot be accomplished by simply measuring ecosystem outcomes in relation to a few specific targets. Second, we also had several sources of information to use in constructing a conceptual model of the NYB. As the NYB is entirely contained within the Mid-Atlantic Bight, studied in detail by the NEFSC, we could use models already developed for this ecosystem to inform our work. We could also rely on existing efforts specifically targeting the NYB where ecosystem components were discussed and ranked (e.g., 2016 NYB indicators workshop), as well as other well-established ecosystem models used in indicator systems globally (see below). This approach also provided an efficient way to get the indicator system started.

We therefore chose to organize indicators around a simple conceptual model of the NYB. The following sections provide summaries of how we accomplished this first task, as well as subsequent steps of choosing appropriate indicators to represent each ecosystem component. We broke down the process into four steps and worked through each over the course of several months (Figure 2). The steps involved (1) developing a conceptual model with key ecosystem components (2) organizing a list of potential indicators for each ecosystem component (3) developing selection criteria for indicators and (4) ranking indicators to determine which indicators were best suited for inclusion in our final ocean indicator suite. This system provided an efficient way to initiate a NYB ocean indicator system, but it should be considered a 'work in progress'. Indicators can be added and subtracted as we learn more about the New York Bight. Each of the steps taken is documented below.

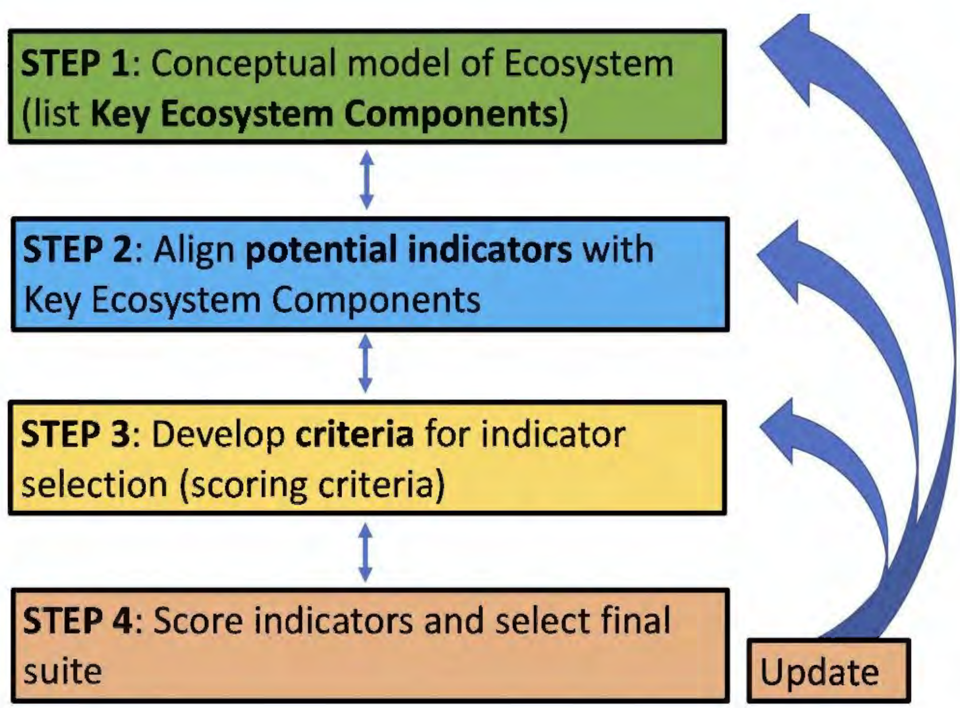


Figure 2. Steps taken to select an ocean indicator suite for the New York Bight.

6. Indicator organization and selection overview

STEP 1:

We developed a simple conceptual model that listed key ecosystem components of the NYB and selected indicators within this framework. This working version was developed by comparing (1) the list of 18 priority indicators/categories that DEC provided (**Table 1**); (2) the conceptual model of the Mid-Atlantic Bight used by NOAA; (3) the results of the 2016 indicators workshop held by DEC and SOMAS and; (4) the categories used in the European Union's Marine Strategy Framework Directive (MSFD). The NOAA model for the MAB is based on a quantitative ecosystem model and is quite comprehensive. The MAB model is an important guide for our efforts because (1) it covers main features of importance in the MAB and (2) the NYB is entirely within the MAB so this model should be applicable to our focal area. The 2016 indicators workshop held by SOMAS and DEC included a vote on indicator categories, and thus it is important to consider the broader scientific expertise in the region. Finally, the MSFD is a policy that legally requires members of the EU to develop indicators for the purpose of assessing marine ecosystem health. The MSFD outlines 11 directives that correspond with components of the ecosystem (i.e., biodiversity, food webs, contaminants). A wealth of academic literature exists related to these directives and supporting indicators, so considering these 11 categories in our model provided further insight.

Table 1. Priority indicator categories for DEC (received via email on June 4th 2020).

Category	Item
Environment	Water temperature (surface and at depth)
	pH and carbonate chemistry (inorganic carbon)
	Eutrophication (N, P, Si, DO)
	Waves, currents at depth, wind speed/direction
	Phytoplankton Chyl-a (satellite) and at depth zooplankton
Higher order	Regime shift indicators
	Climate change indicators
	Relatable to regional fisheries indicators
Humans	Ecosystem services
Marine community	Squid
	Sandlance
	Atlantic herring
	Menhaden
	Biological mid-shelf species biomass
	Biological mid-shelf species abundance
	Biological mid-shelf species distribution
	Diversity (need to include turtles and seabirds)
	Marine mammal body condition, behavior indicators

The working version brings these considerations together under three main headers (Environment, Marine Community, and People) and a depiction of comparisons is presented in Figures 2 - 4. The components are logically sorted into Key Ecosystem Components and Higher Order Properties.

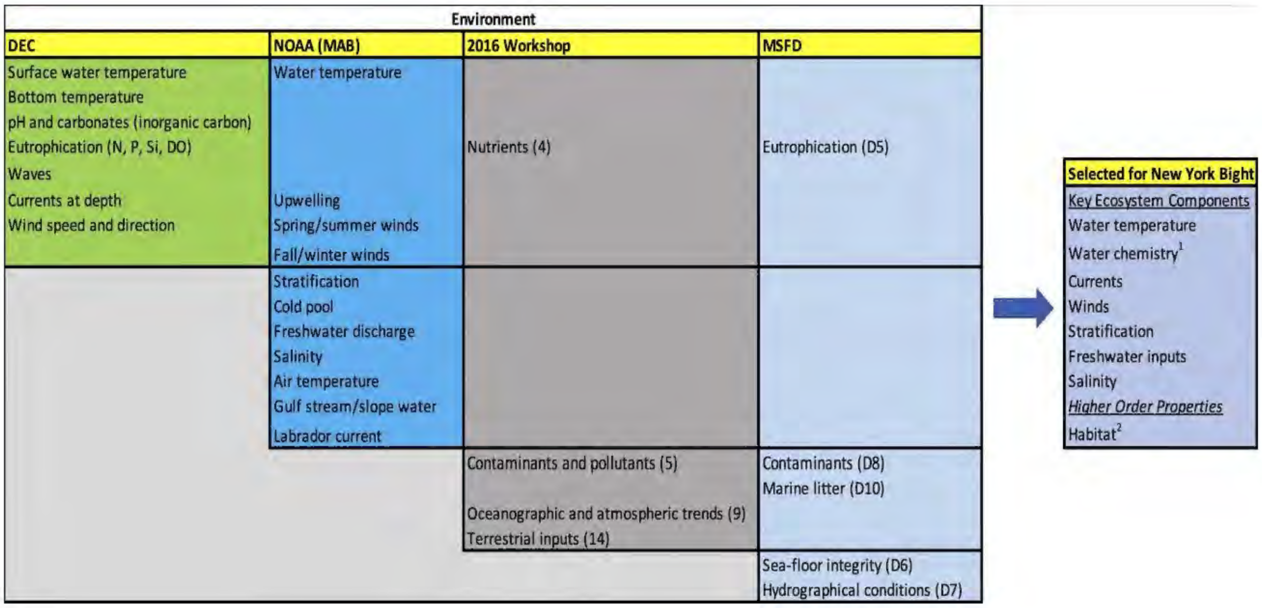


Figure 3. Comparison of DEC priorities for indicators, NOAA's components in a Mid-Atlantic Bight Ecosystem model, ranked priorities form the 2016 NYB indicators workshop, and relevant categories from the Marine Strategy Framework Directive.

Environment - The working version suggests 7 Key Ecosystem Components and 1 Higher Order Property (Figure 3). The final version covers all the priorities of the DEC and most of the categories in the MAB model. A category for "water chemistry" includes carbonate chemistry that the MAB ecosystem status report does not. Contaminants (D8) and Marine litter (D10) from the MSFD were not included here and are put into the "Humans" category (Figure 5); although they are part of the environment, they are anthropogenic modifications to the environment. Inevitably, there will be overlap in some components, especially higher order properties that may integrate across all three major categories (Environment, Marine Community, and Humans).

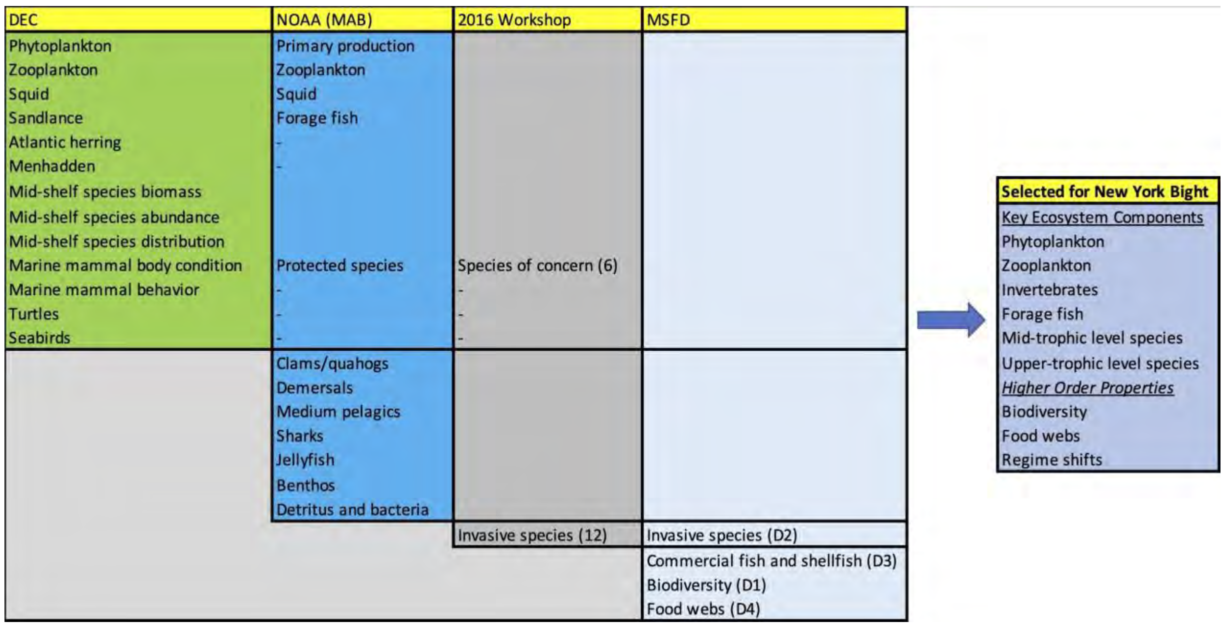


Figure 4. Comparison of DEC priorities for indicators, NOAA's components in a Mid-Atlantic Bight Ecosystem model, ranked priorities form the 2016 NYB indicators workshop, and relevant categories from the Marine Strategy Framework Directive.

Marine community - This major category includes living things and their trophic relationships. There are many ways that these indicators could have been organized but we believe this simple (and broad) approach **(**Figure 4) provides good coverage of the biological community, reduces redundancies, and provides adequate flexibility. The key ecosystem components are loosely based on trophic level. At this stage, we did not pick specific species to represent upper trophic levels and mid-trophic levels.

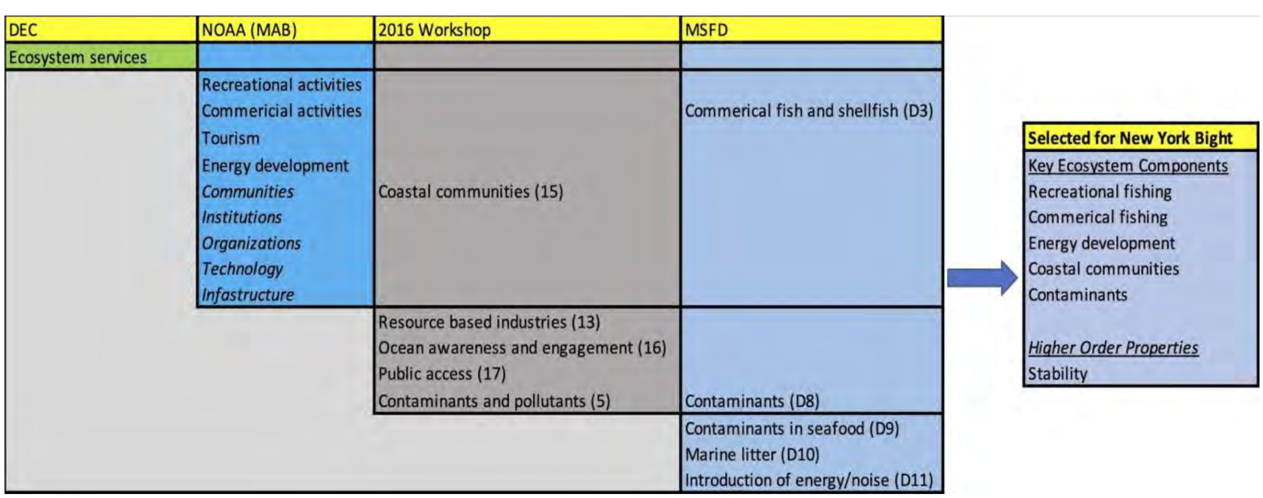


Figure 5. Comparison of DEC priorities for indicators, NOAA's components in a Mid-Atlantic Bight Ecosystem model, ranked priorities form the 2016 NYB indicators workshop, and relevant categories from the Marine Strategy Framework Directive.

Humans - This category brings together components related to ecosystem services provided to humans as well as the pressures that these activities exert on the ecosystem. Many of the ecosystem services identified are contained within the MAB effort (Figure 5). Stability is an emergent property of an ecosystem that is a measure of its resilience and is a unique indicator of the NYB indicator system.

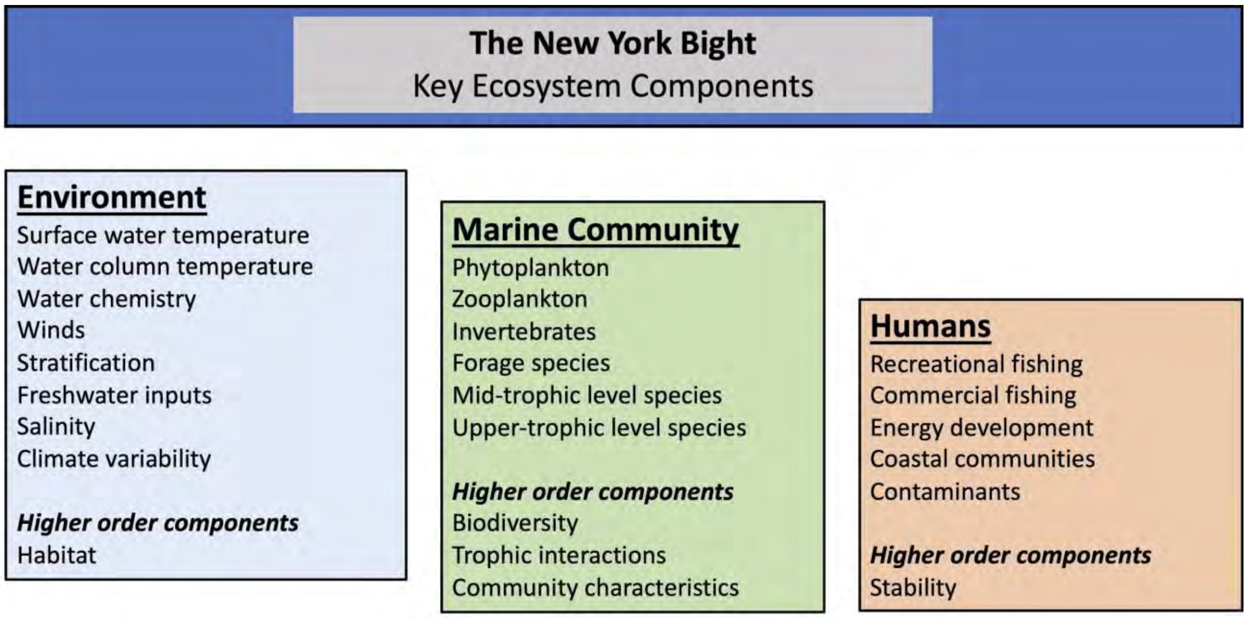


Figure 6: The categories within which to evaluate the best indicators of status for each ecosystem component. In Step 2, scientists and managers voted on the indicator that thought best represented ecosystem status based on established criteria.

After the initial indicators categories were selected, the final set of ecosystem components (Fig 6) differed from the original version (Fig 3 - 5) because updates were made as we went through step 2, 3, and 4 and discussed the results. Notable changes for each category include that we did not develop any indicators of currents per se. We classified the position of the north wall of the Gulf Stream as a mode of climate variability along with the PDO, NAO, AMO and ENSO. For the marine community indicators, we also changed the "Food Webs" category to "Trophic Interactions" to more accurately name these indicators. Without an ecosystem model, it is difficult to evaluate food webs. We did not do any regime shift analysis, although there are several statistical routines to do this so we changed 'Regime Shifts" to "Community Characteristics". A regime shift analysis of all indicators could be completed in the future.

STEP 2:

In Step 2, for each broad category identified above, we organized all of the potential indicators presented in the Indicators Report V1 (Appendix 2) and new indicators made possible by NYOS cruise data in the context of the Key Ecosystem Components. For example, a Key Ecosystem

Component is "water temperature", and in step 2 we described all of the potential indicators that could be used to represent water temperature in the NYB (Table 2).

Table 2: An example of the large number of possible indicators for the "water temperature" category. Most of these indicators are based on mean values, but other statistical properties of each indicator could be used such as the variance, maximum or minimum values, or the rate of change (slope). Thus, the number of indicators to evaluate quickly increases even within one data source.

Potential water temperature indicators
1. Average annual sea surface temperature (OISST)
2. Seasonal mean sea surface temperature (Winter, spring, summer fall with OISST)
3. Long term sea surface temperature (ERSST)
4. Number of marine heatwaves
5. Average heatwave duration
6. Average heatwave intensity
7. Longest marine heatwave
8. Marine heatwave intensity
9. Marine heatwave total days
10. Surface temperature anomaly (in situ)

STEP 3:

There are many papers in the academic literature that attempt to answer the question of how to choose indicators and what criteria should be used (Garcia et al. 2000; Piet and Jennings 2005; Rice and Rochet 2005; Rochet and Rice 2005; Greenstreet and Rogers 2006; Methratta and Link 2006; Samhouri et al. 2009; Shin et al. 2010; Kershner et al. 2011; Queirós et al. 2016; Tam et al. 2017; Otto et al. 2018). Step 3 includes developing an agreed upon set of criteria for selecting indicators (e.g., data quality, responsiveness to known pressure, relevance to OAP goals). We aimed to have at least one indicator for each of the Key Ecosystem Components, but in many cases, several would be selected to accurately describe ecosystem dynamics and status. We suggested eight total screening criteria to our team of scientists and managers that are based primarily on (Rice and Rochet 2005; Queirós et al. 2016; Otto et al. 2018) and agreed upon a total of 6 set of criteria to rank indicators. Criteria 1- 5 (bolded below) were ultimately used to rank indicators, but Criteria 6-8 were not used.

1. **Intelligible and easy to understand** (Otto et al. 2018)
   a. An indicator scores high on this criterion if it is simple and easy to understand. It scores lower if it is complex and not immediately intelligible with regards to what it means. *E.g., Mean summer water temperature is simple; Annual average of Marine Heatwave intensity is slightly more complex and harder to understand.*
   b. Scored by all

2. Relevance to DEC management goals and priorities (Otto et al. 2018)

a. An indicator scores highly on this criterion if it aligns with an Ocean Action Plan priority or provides guidance for DEC with regards to a management issue of interest.

b. Scored by DEC managers only

3. Scientific basis (Queirós et al. 2016)

a. There is evidence that the indicator under consideration serves as a good proxy to represent changes in the Key Ecosystem Component it represents and/or this indicator is also used in other Ecosystem Monitoring programs (e.g., Precedence).

b. Scored by all

4. Ecosystem Relevance (Queirós et al. 2016)

a. "There is evidence linking the indicator to (a) ecosystem level processes and function (the non-anthropocentric perspective; e.g., indicators of processes undertaken by keystone species could be particularly relevant); and/or (b) ecosystem services (the anthropocentric perspective, i.e., societal relevance)"

b. Scored by all

5. Quality of sampling method (Queirós et al. 2016)

a. "The indicator is concrete/measurable, accurate, precise or repeatable. Concreteness/measurability refers to whether the indicator can be quantitatively assessed. Accuracy refers to the closeness of an estimate of an indicator to the true value of the indicator. Precision refers to the degree of concordance among a number of estimates for the same population and repeatability to the degree of concordance among estimates obtained by different"

b. Scored by all

6. Existing and ongoing monitoring data (Queirós et al. 2016)

a. Indicators score highly in this category if they are based on existing and ongoing data sources that are immediately available for use (i.e., real time updates).

7. Public awareness (Rice and Rochet 2005)

a. Indicators that are perceived to be interesting to the public are scored highly. (e.g., Striped Bass recreational harvest is interesting, Atlantic croaker recreational harvest might be less interesting).

8. Distinctiveness relative to Mid-Atlantic Bight (Heim et al. in review)

a. An indicator scores highly if the time series generated for the NYB differs from the time series generated for the Mid-Atlantic Bight. This serves to prioritize indicators that represent a distinct feature of the NYB and are not redundant with Ecosystem Indicators reported by NOAA.

b. This is a quantitative criteria and as such will not involve participant scoring

STEP 4:

Finally, a small working group of ecologists, biological oceanographers and physical oceanographers at SOMAS scored the indicators relative to the criteria arrived upon in Step 3 so that highest scoring indicators can be identified and selected for the final suite (Step 4). Starting with the first category, "Environment", there were 41 indicators that were ranked first by the SOMAS scientists and then the 21 most highly ranked indicators were passed to NYS DEC managers to evaluate (Table 3).

Table 3: Potential indicators for the environment category. The first column contains the indicators that were ranked highest by SOMAS scientists and the second column are those that were eliminated before sending to DEC managers for ranking. All indicators were discussed after the ranking process for a final selection of indicators.

Passed to NYS DEC	Eliminated before sending to NYS DEC
1. Seasonal mean sea surface temperature (Winter, spring, summer fall with OISST)	1. Average annual sea surface temperature (OISST)
2. Marine heatwave total days	2. Number of marine heatwaves
3. Surface temperature anomaly (in situ)	3. Average heatwave duration
4. Bottom temperature anomaly	4. Longest marine heatwave
5. Mid-Atlantic cold pool duration	5. Marine heatwave intensity
6. Mid-Atlantic cold pool volume	6. Start of spring
7. Seasonal N and P concentration	7. Duration of summer
8. Mean seasonal bottom pH	8. Warm core ring frequency in NYB
9. Mean seasonal surface pH	9. Gulf Stream meanders
10. Mean seasonal bottom aragonite saturation	10. Long term seasonal sea surface temperature (ERSST)
11. Mean surface pCO2	11. Mid-Atlantic cold pool temperature
12. Mean surface aragonite concentration	12. Mean seasonal surface aragonite
13. Dissolved oxygen of bottom waters	13. Buffering capacity of surface waters
14. Position of the Gulf Stream	14. Mid-Atlantic cold pool stratification
15. Mean wind stress, mean wind speed and direction (by month or season)	15. Pacific decadal oscillation (PDO)
16. Mean number of storms	16. El Nino Southern Oscillation (ENSO)
17. Stratification anomaly	
18. Mean (or maximum) Hudson River flow	
19. Center timing of spring (or fall) flow	
20. Surface (or bottom) water salinity	
21. North Atlantic Oscillation	
22. Atlantic Multidecadal Oscillation	
23. Global concentration of carbon dioxide	
24. Location of 20C isotherm	
25. Number of days NYB averages above 20^{o}C	

The scores given by scientists and managers were not correlated (Figure 7), indicating that it is important to have the perspectives of both scientists and managers in developing an indicator system. In general, managers ranked indicators more highly than scientists at SOMAS.

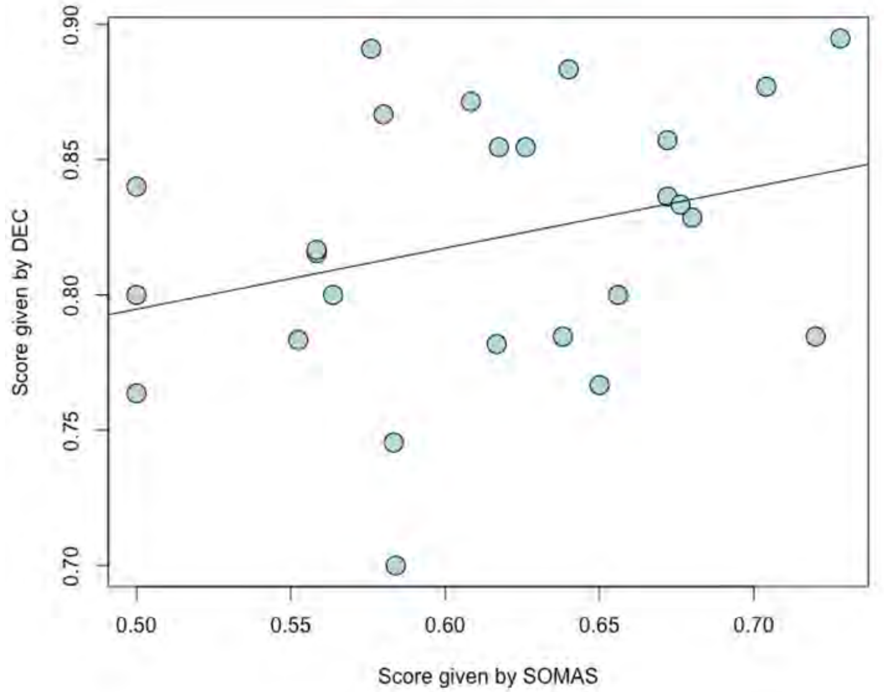


Figure 7: Lack of correlation between marine scientists and managers at NYS DEC. The Pearson correlation coefficient was ρ=0.303, p-value=0.13

A similar process occurred with the marine community indicators (Table 4), the only difference being that we sent the complete list of all indicators to DEC with an explanation as to why SOMAS scientists excluded some and not others. DEC then ranked the indicators as a team. There were so few indicators for the human community that we did not cull any indicators before sending to DEC to evaluate.

Table 4: Potential indicators for the Marine Community category. The first column contains the indicators that were ranked highest by SOMAS scientists and the second column are those that were eliminated before sending to DEC managers for ranking.

Passed to NYS DEC	Eliminated before sending to NYS DEC
1. Monthly mean surface chlorophyll concentration	1. Annual mean surface chlorophyll concentration
2. Copepod size index	2. Peak date of surface chlorophyll
3. Bottom temperature anomaly	3. Zooplankton biovolume
4. Calanus finmarchicus abundance	4. *Pseudocalanus spp*, abundance
5. Copepod lipid content	5. *Centropages hamatus* abundance
6. American lobster biomass	6. *Centrapages typicus* abundance
7. Surfclam biomass (SOMAS/DEC survey)	7. *Temora longicornis* abundance
8. Sea scallop biomass (NOAA scallop survey)	8. Sea Scallop biomass (bottom trawl survey)
9. Jonah crab biomass	9. Atlantic herring biomass
10. Longfin squid biomass	10. Atlantic mackerel biomass
11. Northern shortfin squid biomass	11. Butterfish biomass
12. Forage fish biomass	12. Chub mackerel biomass
13. Menhaden biomass	13. Scup biomass
14. Biomass of 4 feeding guilds	14. Ratio of Pelagic to demersal fish
15. Total trawl biomass	15. Fish condition (NOAA data)
16. Black sea bass biomass	16. Fish condition (BIA on Seawolf)
17. Summer flounder biomass	17. Longhorn sculpin biomass
18. Ratio of reef to non-reef benthic fish (e.g. black sea bass vs. summer flounder)	18. Number of large whale strandings
19. Ratio of southern to northern species (e.g. winter flounder vs. summer flounder biomass	19. Common merganser counts
20. Ratio of warm to cold water species (e.g. Spiny dogfish vs. smooth dogfish biomass)	20. Long-tailed duck counts
21. Humpback whale body condition	21. Swan counts
22. Humpback whale foraging behavior	22. Least tern breeding pairs on Long Island
23. Seabirds	23. Relative abundance of tropical vs. temperate odontocete species
24. Cetacean species richness	
25. Zooplankton species richness	
26. Fish species richness	
27. Overlap of marine mammals with prey	
28. Spatial correlation of acoustic backscatter from zooplankton and fish	
29. Average number of fish aggregations (Seawolf)	
30. Isotopic analysis of key mid and upper level species	
31. Average trophic level of fish community	
32. Temperature preference of the fish community	

After indicators were selected independently by SOMAS scientists and DEC managers, we met as a group and finalized a list of 40 indicators with which to proceed. Most indicators have been developed (24) or are under development (4), but many indicators (12) were proposed during the process (Table 5). These include some measure of the Mid-Atlantic Cold pool, an area of remnant cold bottom water that is extremely important for many species in the New York Bight. Calculating a satisfactory indicator has been problematic because bottom temperature data is not available at high spatial and temporal resolution. Water chemistry indicators have also been problematic because high density data has not been available until recently. There is very little nutrient data for the Mid-Atlantic Bight. Indicators of bottom dissolved oxygen, surface carbonate chemistry (pH, pCO_2, and aragonite) and humpback whale body condition are now being developed, being made possible with sampling on the RV Seawolf by SOMAS. Vessel density and speed indicators in NYB are being developed by SOMAS staff as another indicator of human activity. Other proposed indicators are possible to develop in the near term with existing data. Other indicators that are needed but would require additional effort, funding or cooperation with other researchers are listed in Table 6.

Table 5: The final list of indicators of ocean indicators for the NYB in 2020

Key Ecosystem Component	**Indicator**	**Status**
Surface Water Temperature	1. Seasonal mean sea surface temperature	Completed
	2. Marine Heatwave Days	Completed
Water Column Temperature	3. Bottom Temperature Anomaly	Completed
	4. Cold Pool Volume	**Proposed**
	5. Cold Pool Duration	**Proposed**
Water Chemistry	6. Bottom Dissolved Oxygen	**Under development**
	7. Surface pH, pCO2, aragonite	**Under development**
Winds	8. Mean Wind Stress	**Proposed**
	9. Number of Large Storms	**Proposed**
Stratification	10. Stratification anomaly	Completed
Freshwater Input	11. Hudson River Flow	Completed
Salinity	12. Surface Water Salinity	Completed
	13. Bottom Water Salinity	Completed
Climate Variability	14. Global Carbon Dioxide	Completed
Habitat	15. Lobster thermal habitat	**Proposed**
	16. Location of 20 C isotherm	**Proposed**

Phytoplankton	17. Monthly mean surface chlorophyll concentration	Completed
Zooplankton	18. *Calanus finmarchicus* abundance	Completed
	19. Copepod size index	Completed
Invertebrates	20. American lobster & Jonah crab abundance	Completed
Forage species	21. Longfin squid biomass	Completed
	22. Northern shortfin squid biomass	Completed
	23. Forage species biomass	Completed
Mid Trophic level species	24. Feeding groups (NOAA)	Completed
	25. Total trawl biomass	Completed
	26. Black sea bass biomass	Completed
	27. Summer Flounder biomass	Completed
	28. Northern vs. southern species biomass	**Proposed**
	29. Structure oriented species biomass	**Proposed**
Upper trophic level species	30. Humpback whale body condition	**Under development**
Biodiversity	31. Fish species richness	**Proposed**
Trophic Interactions	32. Average trophic level of fish community	**Proposed**
Community Characteristics	33. Temperature preference of fish community	**Proposed**
Recreational Fishing	34. Total recreational harvest in NY Bight	Completed
	35. Recreational effort (# angler trips)	**Proposed**
Commercial Fishing	36. Commercial fishing landing (weight)	Completed
	37. Commercial landings value in NY (USD)	Completed
Coastal Communities	38. Human population of Long Island	Completed
	39. Vessel density and speed in NYB	**Under development**
	40. Sea level risk for LI communities	Completed

Table 6: Indicators to be developed in the future either with additional funding, in coordination with other funded projects and/or with reallocation of effort in current projects:

1. Copepod lipid content using collections from the RV Seawolf offshore monitoring effort. Some effort to do the species ID may have to be reduced to do the lipid content analysis.
2. Surfclam abundance or biomass from Bob Cerrato's DEC-funded project to survey surf clams aboard the RV Seawolf
3. Sea scallop abundance or biomass from NOAA NEFSC scallop survey
4. Menhaden biomass to be obtained with acoustics in nearshore habitat with additional funding
5. One or two more trophic interaction indicators as Objective 6 develops and evolves
6. Seabird species - data sources need to be decided
7. Reliance and engagement in recreational/commercial fisheries using data and assistance from NOAA NEFSC
8. Develop socioeconomic indicators working with Shinnecock nation and NOAA NEFSC
9. Wind Energy produced in NY Bight - total used up space compared to available habitat (not leases, but actual WEAs)
10. Contaminants levels of PCB, Hg, etc. in fish working with other funded researchers at SOMAS
11. Estimate ecosystem stability, maturity and resilience of the NYB ecosystem models building off the mass-balance model being developed by Nye and Frisk for the MidAtlantic Bight
12. Other Environmental indicators

7. Recommendations and next steps

The structured process to arrive on a set of indicators was lengthy, but the list should be viewed as fluid. Indicators can be added, subtracted, or modified over time. The next step is to evaluate which of the proposed indicators are possible to develop over the next year given the available data and expertise. Subject area experts may need to be invited into the process to develop indicators. As stated previously, some indicators might be very easy to develop, while others may require intensive model development and vetting by the scientific community.

7. References

Conservation Measures Partnership. 2007. Open Standards for the Practice of Conservation, Version 2.0. 40 p.

DePiper, G.S., Gaichas, S.K., Lucey, S.M., Da Silva, P.P., Anderson, M.R., Breeze, H., Bundy, A., Clay, P.M., Fay, G., Gamble, R.J., Gregory, R.S., Fratantoni, P.S., Johnson, C.L., Koen- Alonso, M., Kleisner, K.M., Olson, J., Perretti, C.T., Pepin, P., Phelan, F., Saba, V.S., Smith, L.A., Tam, J.C., Templeman, N.D., and Wildermuth, R.P. 2017. Operationalizing integrated ecosystem assessments within a multidisciplinary team: Lessons learned from a worked example. ICES J. Mar. Sci. 74(8): 2076–2086. doi:10.1093/icesjms/fsx038.

EcoAP (Ecosystem Assessment Program). 2009. Ecosystem Status Report For the Northeast U.S. Continental Shelf Large Marine Ecosystem. Northeast Fisheries Science Center Reference Document, 09-11: 34.

EcoAP (Ecosystem Assessment Program). 2011. Ecosystem Status Report For the Northeast Shelf Large Marine Ecosystem-2011. Northeast Fisheries Science Center Reference Document, 12-07: 38.

Garcia, S., D. Staples, and J. Chesson. 2000. The FAO guidelines for the development and use of indicators for sustainable development of marine capture fisheries and an Australian example of their application. Ocean & Coastal Management 43:537-556.

Greenstreet, S. P., and S. I. Rogers. 2006. Indicators of the health of the North Sea fish community: identifying reference levels for an ecosystem approach to management. ICES Journal of Marine Science 63:573-593.

Heim, K. C., L. H. Thorne, J. D. Warren, J. S. Link, and J. A. Nye. in review. Marine ecosystem indicators are sensitive to ecosystem boundaries and spatial scale. Ecological Indicators.

Kershner, J., J. F. Samhouri, C. A. James, and P. S. Levin. 2011. Selecting indicator portfolios for marine species and food webs: A Puget Sound case study. PLoS ONE 6(10).

Levin, P. S., M. J. Fogarty, S. A. Murawski, and D. Fluharty. 2009. Integrated Ecosystem Assessments: Developing the Scientific Basis for Ecosystem-Based Management of the Ocean. PLoS Biol 7:e1000014.

Link, J. S., Brodziak, J. K. T., editors, and Brodziak, J. K. T., Dow, D. D., Edwards, S. F., Fabrizio, M. C., Fogarty, M. J., Hart, D. Jossi, J. W., Kane, J., Lang, K. L., Legault, C. M., Link, J. S., MacLean, S. A., Mountain, D. G., Olson, J., Overholtz, W. J., Palka, D.L., Smith, T. D., contributors. 2002. Status of the Northeast U. S. Continental Shelf Ecosystem: a report of the Northeast Fisheries Science Center's Ecosystem Status Working Group. Northeast Fish. Sci. Cent. Ref. Doc. 02-11; 245 p.

Link, J.S. 2010. Ecosystem-based fisheries management: Confronting tradeoffs. In Ecosystem-Based Fisheries Management: Confronting Tradeoffs. doi:10.1017/CBO9780511667091.

McLeod K.L., Lubchenco J., Palumbi S.R., Rosenberg A.A.. 2005. Scientific consensus statement on marine ecosystembased management. Signed by 219 academic scientists and policy experts with relevant expertise and published by the Communication Partnership for Science and the Sea.

Methratta, E.T., and Link, J.S. 2006. Evaluation of quantitative indicators for marine fish communities. Ecological Indicators 6(3): 575–588. doi:10.1016/j.ecolind.2005.08.022.
NEFSC. 2017. State of the Ecosystem—Mid-Atlantic Bight. Northeast Fisheries Science Center, Woods Hole, MA.
NEFSC 2018. State of the ecosystem 2019: Mid-Atlantic. 2018. Retrieved from https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports
NEFSC 2019. State of the ecosystem 2019: Mid-Atlantic. 2019. Retrieved from https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports
Otto, S. A., M. Kadin, M. Casini, M. A. Torres, and T. Blenckner. 2018. A quantitative framework for selecting and validating food web indicators. Ecological Indicators 84:619-631.
Piet, G. J., and S. Jennings. 2005. Response of potential fish community indicators to fishing. ICES Journal of Marine Science 62:214-225.
Queirós, A. M., J. A. Strong, K. Mazik, J. Carstensen, J. Bruun, P. J. Somerfield, A. Bruhn, S. Ciavatta, E. Flo, and N. Bizsel. 2016. An objective framework to test the quality of candidate indicators of good environmental status. Frontiers in Marine Science 3:73.
Rice, J. C., and M. J. Rochet. 2005. A framework for selecting a suite of indicators for fisheries management. ICES Journal of Marine Science 62(3):516–527.
Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., and Schlax, M.G. 2007. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. J. Clim. 20(22): 5473–5496. doi:10.1175/2007JCLI1824.1.
Rosenberg, A. A., and K. L. McLeod. 2005. Implementing ecosystem-based approaches to management for the conservation of ecosystem services: Politics and socio-economics of ecosystem-based management of marine resources. Marine ecology. Progress series 300:271-274.
Rochet, M.-J., and J. C. Rice. 2005. Do explicit criteria help in selecting indicators for ecosystem-based fisheries management? ICES Journal of Marine Science 62:528-539.
Samhouri, J. F., P. S. Levin, and C. J. Harvey. 2009. Quantitative Evaluation of Marine Ecosystem Indicator Performance Using Food Web Models. Ecosystems 12:1283-1298.
Shin, Y.-J., and L. J. Shannon. 2009. Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 1. The IndiSeas project. ICES Journal of Marine Science 67:686-691.
Shin, Y.-J., L. J. Shannon, A. Bundy, M. Coll, K. Aydin, N. Bez, J. L. Blanchard, M. d. F. Borges, I. Diallo, E. Diaz, J. J. Heymans, L. Hill, E. Johannesen, D. Jouffre, S. Kifani, P. Labrosse, J. S. Link, S. Mackinson, H. Masski, C. Möllmann, S. Neira, H. Ojaveer, K. ould Mohammed Abdallahi, I. Perry, D. Thiao, D. Yemane, and P. M. Cury. 2010. Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 2. Setting the scene. ICES Journal of Marine Science: Journal du Conseil 67:692-716.
Tam, J. C., J. S. Link, A. G. Rossberg, S. I. Rogers, P. S. Levin, M.-J. Rochet, A. Bundy, A. Belgrano, S. Libralato, and M. Tomczak. 2017. Towards ecosystem-based management: identifying operational food-web indicators for marine ecosystems. ICES Journal of Marine Science 74:2040-2052.
Turnhout, E., M. Hisschemöller, and H. Eijsackers. 2007. Ecological indicators: Between the two fires of science and policy. Ecological Indicators 7(2):215–228.

WSAS (Washington State Academy of Sciences). 2012. Sound Indicators: A review for the Pugest Sound Partnership. Accessed at: https://www.eopugetsound.org/articles/sound-indicators-review-puget-sound-partnership

New York Bight Indicator Report 2020
Part 2: Indicator Analysis

MOU #AM10560 NYS DEC & SUNY Stony Brook

Report contributors: Kurt Heim, Senior Researcher; Janet Nye, Associate Professor; Lesley Thorne, Assistant Professor; Joe Warren, Associate Professor; Charlie Flagg, Research Professor; Tyler Menz, Carbonate Chemistry Technician

Contact Principal Investigator, Janet A. Nye with questions or comments
janet.nye@stonybrook.edu

1. Executive Summary

- Surface and bottom waters of the New York Bight (NYB) are warming.
- Freshwater output from the Hudson River has increased steadily since the 1960s.
- Declines in chlorophyll in April, November, December and January coincide with a large increase in the number of marine heatwave days.
- Trends in upper trophic levels were highly nonlinear and variable between spring and fall sampling periods.
- Benthic invertebrates including American lobster, Jonah crab, black sea bass and summer flounder have increased in the last decade, but summer flounder have recently declined
- Benthivores, piscivores, total survey trawl biomass, and commercial landings declined in the 1960s and have not recovered to historic levels in the NYB since.
- Numbers of fish caught by recreational fishing has increased while commercial catch has decreased.
- The coastal communities on the south shore of Long Island as well as Queens and King counties are the most vulnerable counties in NY to sea level rise.

2. Report structure

Below we present the subset of indicators selected as described in Part 1 of this report. For every indicator, the time series of observations is fit with a linear and nonlinear trend. The linear trends were fit with simple linear regression and significance determined with standard techniques. The nonlinear trends were fit with generalized additive models (GAMs) (Hastie and Tibshirani 1987). If the trend is sufficiently linear, the fitted function of the GAM will also be a linear function. To objectively determine time periods with statistically significant trends in the indicators, we performed an analysis of the first derivatives of the fitted GAM functions (Curtis and Simpson 2014, Santibáñez et al. 2018). Throughout the document, consistent symbology is used where dashed lines through time series indicate no statistical significance, but a solid line indicates statistical significance. In the GAM-fitted models, red solid lines indicate significantly increasing and blue solid lines indicate significantly decreasing trends. All files used to create these figures and steps are documented in R Markdown files (i.e., how to access data, calculate indicator, save in standard format) and are available upon request.

3. Environment

Seasonal mean sea surface temperature

Sea surface temperature has increased in all four season in the New York Bight as indicated by satellite data extending back to the early 1980s, although not statistically significant in winter (Figure 1). The most rapid increase has occurred in summer and fall consistent with studies of the east coast of the US (Thomas et al. 2016). The increase in temperature has been most rapid from 2000 to present (Figure 2).

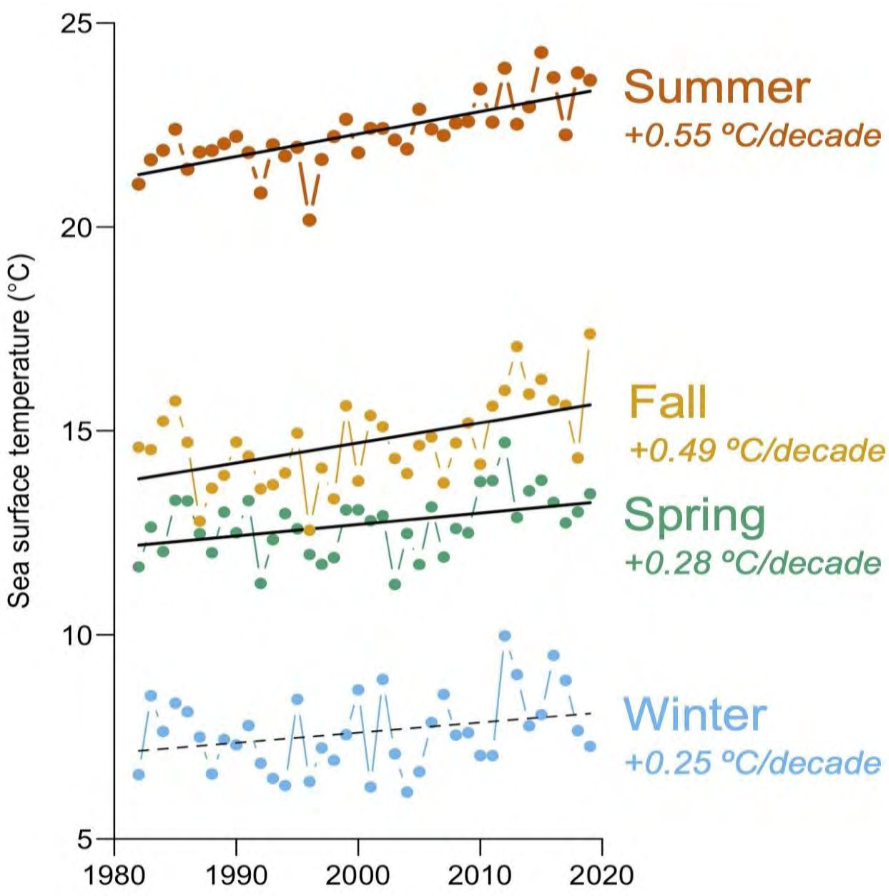


Figure 1. Seasonal mean sea surface temperature. Dashed line indicates the trend is not significant (p>0.05) and a solid line indicates a statistically significant linear trend (p<0.05). The warming trend is not statistically significant in winter, but the trend in summer, fall and spring are statistically significant with the highest rate occurring in summer.

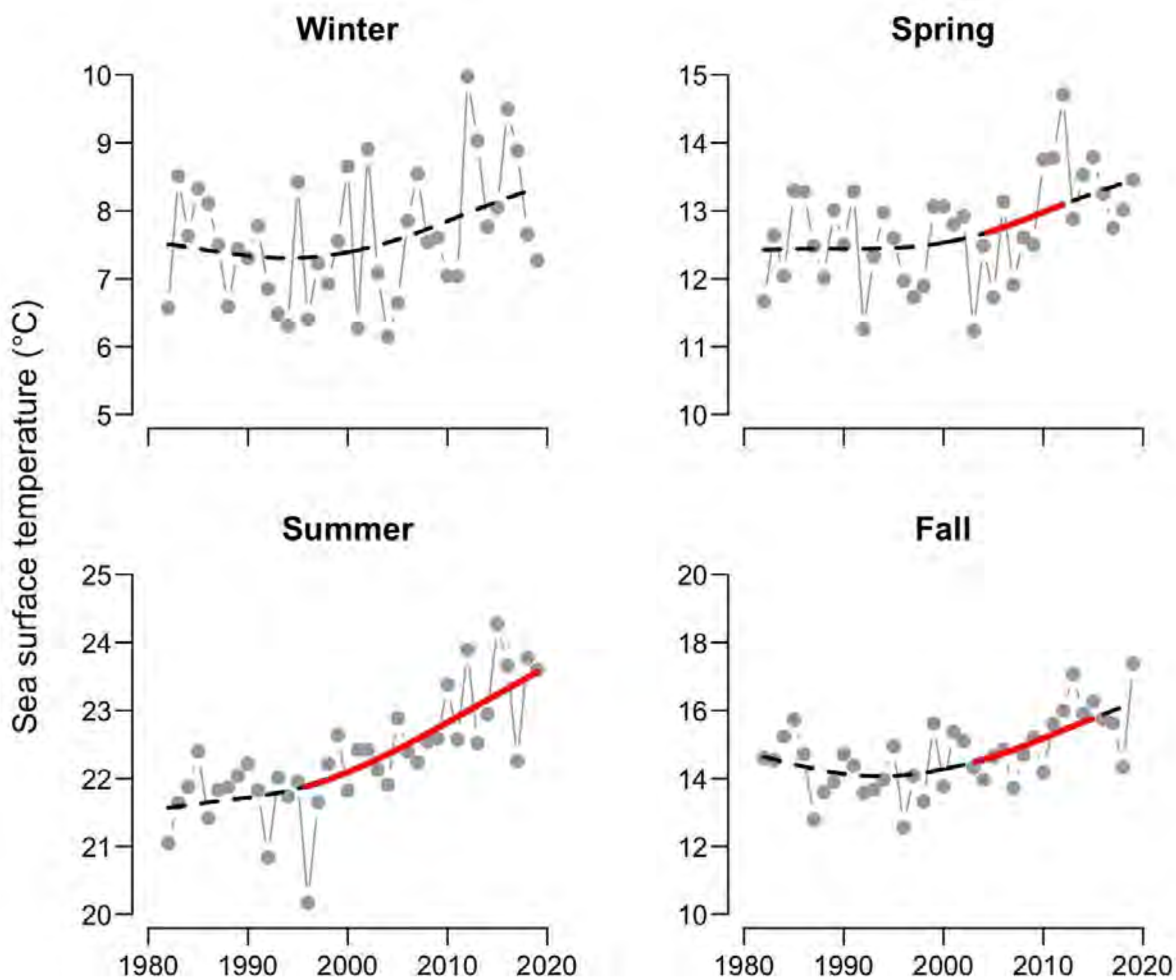


Figure 2. Seasonal mean sea surface temperature with nonlinear trends. Solid red lines indicate a significant increasing trend with blue indicating a decreasing trend. The dashed black line indicates no statistically significant trend ($p < 0.05$).

Marine heatwave days

Ocean heatwaves are extreme events where surface temperatures exceed the 90th percentile of historic average for at least 5 days (Hobday et al. 2018, Oliver et al. 2020). Ocean heatwaves have become more frequent and intense globally (Oliver et al. 2018) and the NYB is no different (Figure 3, 4). The number of heatwave days in the NYB has been 50 days or more in each of the last ten years. This means organisms are regularly experiencing temperatures much warmer than they have ever experienced in the past. Marine heatwave days increased rapidly starting in 2000 and have remained high in recent years, suggesting we have entered a new ocean regime. However, this must be fully assessed with continued monitoring of this indicator.

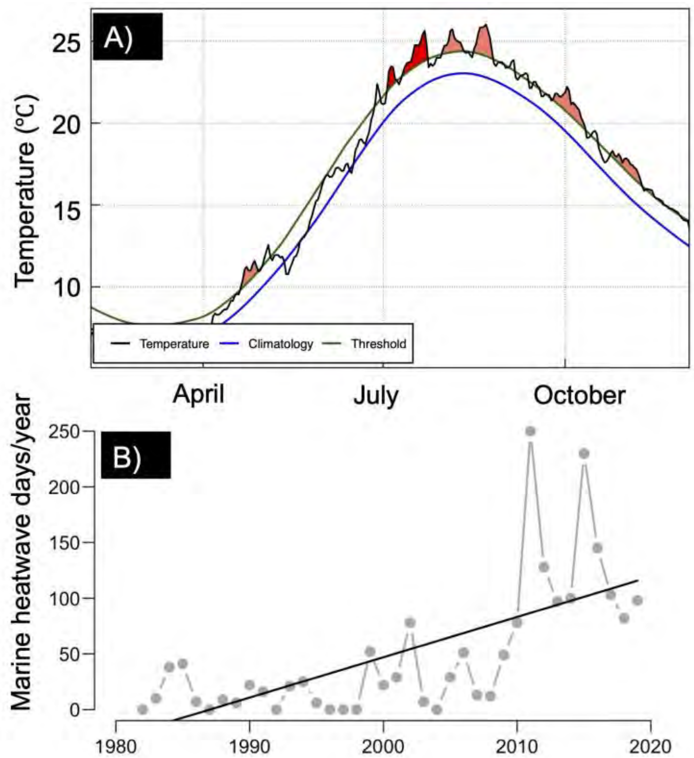


Figure 3. A) Seasonal sea surface temperature cycle for 2019 with red areas indicating marine heatwaves as an example of how marine heatwave days are calculated. B) Trend in total marine heatwave days over time. The solid line indicates a statistically significant trend ($p<0.05$).

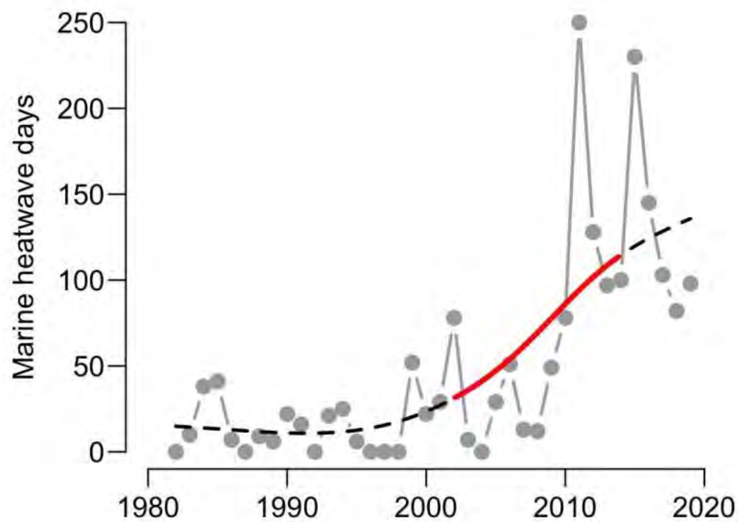


Figure 4. Marine heatwave days per year now showing the trend with a nonlinear GAM technique where red indicates a significant increasing trend and blue indicates a significant decreasing trend and the dashed line indicates no trend.

Bottom temperature anomaly

Observations of bottom temperature are more sparse than satellite-derived SST observations spatially and temporally, but reveal similar trends. Bottom temperatures have also increased in the NYB and the trend is linear when fit with both linear regression (Figure 5) and the flexible GAM approach (Figure 6). Future work will investigate seasonal differences in warming rate.

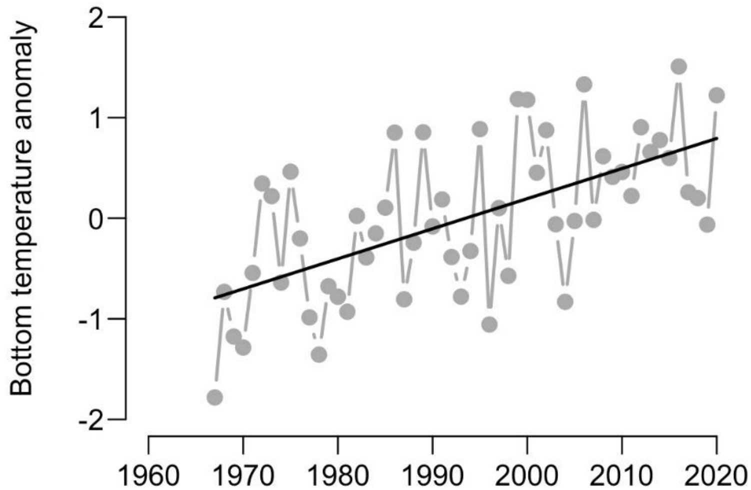


Figure 5. Bottom temperature anomaly linear fit. Dashed line indicates no statistical significance (p>0.05). Solid line indicates statistical significance (p<0.05).

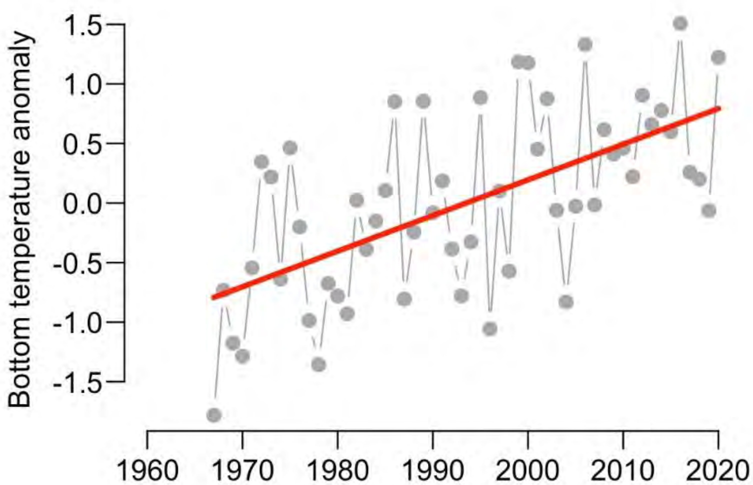


Figure 6. Bottom temperature anomaly GAM where red indicates a significant increasing trend and blue indicates a significant decreasing trend and the dashed line indicates no trend.

Stratification anomaly

Stratification was calculated using in situ data from CTD casts so spatial and temporal coverage is relatively limited. Stratification was highly variable from year to year and there was no statistical trend over time (Figure 7). Further investigation into this variable including seasonality may reveal important changes that influence phytoplankton blooms over time.

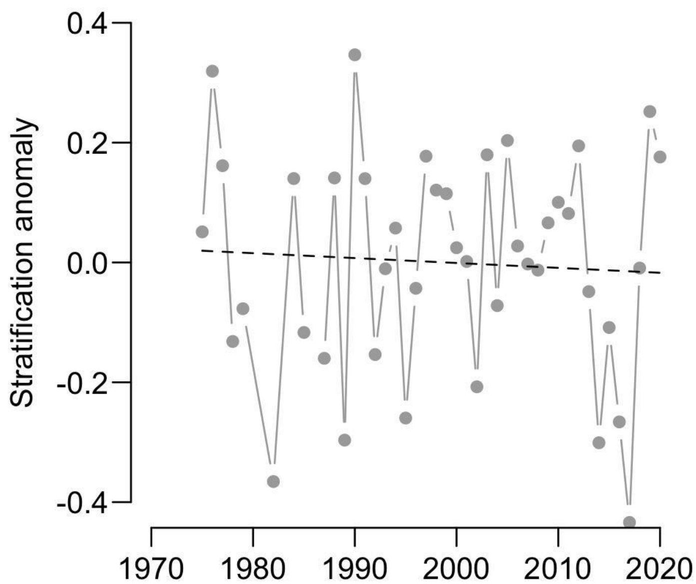


Figure 7. Stratification anomaly linear fit. Dashed line indicates no statistical significance (p>0.05) and the solid line indicates statistical significance(p<0.05). The GAM fit for this indicator was also linear and not statistically significant so not shown.

Freshwater inputs

The river discharge from the Hudson River, the major source of freshwater in the NYB, has increased significantly over time. This may be a result of the observed increase in the magnitude and frequency of heavy precipitation in the Northeast US (Mallakpour and Villarini 2017). Higher river flow will bring more nutrients and sediment to the NYB and in turn impact coastal eutrophication, dissolved oxygen, carbonate chemistry and phytoplankton bloom dynamics.

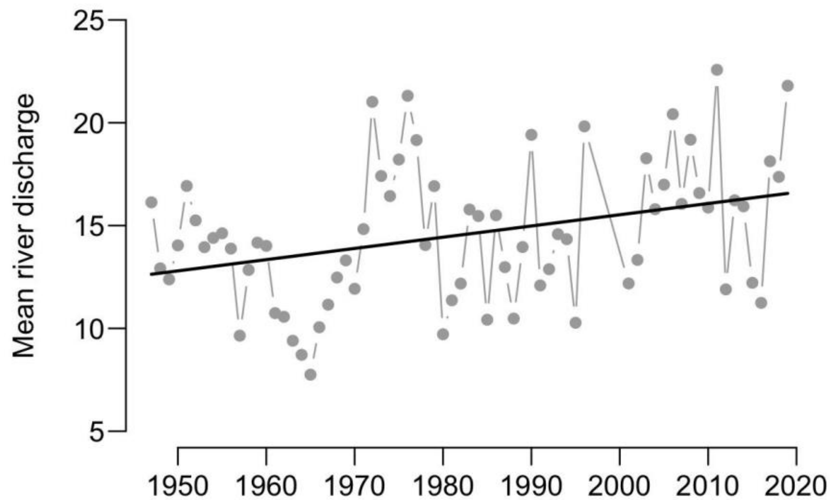


Figure 8. River discharge of the Hudson River has increased over time. Dashed line indicates no statistical significance ($p>0.05$). Solid line indicates statistical significance ($p<0.05$). Units are thousands of CFS (cubic feet/second * 1000). The GAM fit is also linear and statistically significant so is not shown.

Surface and bottom water salinity

Salinity is an important indicator of changes in ocean circulation and salinity along with temperature can be used to identify water masses. In the NYB, water originating from the Gulf Stream, including water core rings, is warmer and saltier while water originating from upstream Labrador slope water is colder and less saline. These oceanic waters can also be differentiated from freshwater systems by salinity. Surface and bottom water salinity have not changed significantly over time (Figure 9). There are some notable years where salinity was anomalously "fresh" around 1999. From 2015-2017, the NYB was much "saltier" which is likely associated with warm core ring activity (Gangopadhyay et al. 2019, Gawarkiewicz et al. 2019).

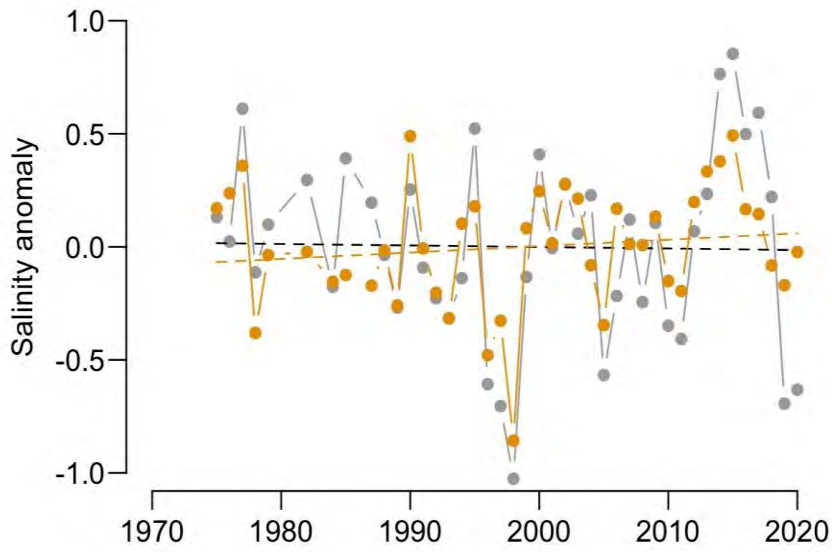


Figure 9. Surface and bottom water salinity anomalies are similar to each other and highly variable. There is no trend over time. Dashed line indicates no statistical significance ($p>0.05$). Solid line indicates statistical significance ($p<0.05$).

Global Carbon Dioxide

Although many time series of atmospheric carbon dioxide exist, the time series from Mauno Loa is the longest and most consistent time series available in the US (Figure 10). Atmospheric carbon dioxide has increased exponentially since the dawn of the Industrial Revolution.

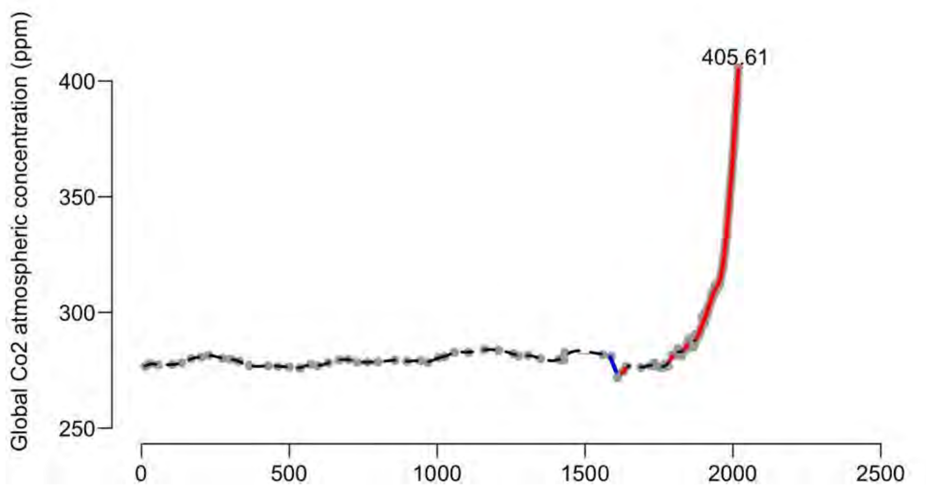


Figure 10. Global atmospheric CO_2 from Mauno Loa time series with trend diagnosed with a GAM. Red indicates a significant increasing trend and blue indicates a significant decreasing trend and the dashed line indicates no significant trend ($p>0.05$). Solid line indicates statistical significance ($p<0.05$).

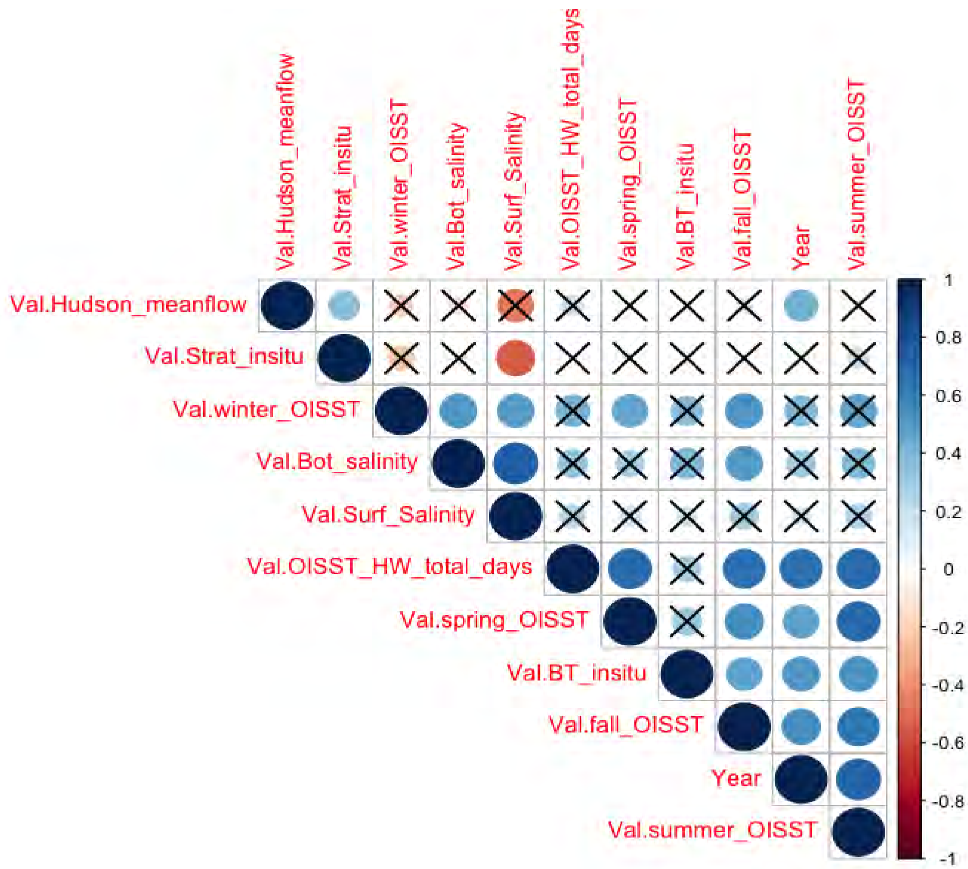


Figure 11. Correlations between environment indicators with X indicating non-significant correlations among all environment indicators presented above. Blue indicates a positive correlation while red indicates a negative correlation, the magnitude of which is given in the scale bar to the right.

Environmental indicator highlights

We calculated the Pearson correlation coefficients among all environmental indicators (Figure 11). As described in Part 1 of this report, we attempted to reduce the number of indicators in part by eliminating those that were highly correlated and would thus yield the same information. As indicated by the number of non-significant correlations we were successful in this goal (Figure 11). All the temperature indicators are positively correlated with each other as indicated by the blue symbols in the lower right hand corner of the correlation matrix and all illustrate the warming of NYB waters. Stratification (Val.Strat_insitu) was negatively correlated with surface salinity suggesting that increases in stratification are associated with decreases in surface salinity. This is consistent with physical principles and our calculation of stratification based on both temperature and salinity. The negative correlation between stratification and salinity as well as the positive correlation between river flow and stratification suggests that freshwater input may drive changes in stratification in the NYB, but these dynamics warrant further investigation.

4. Marine Community

Monthly mean surface chlorophyll

Mean surface chlorophyll was highly variable from year to year. Of note, chlorophyll levels declined late in the time series in April, November, December and January. The low chlorophyll levels in November and December could be related to the warmer surface temperatures in the fall in recent years, but the causes must be further evaluated.

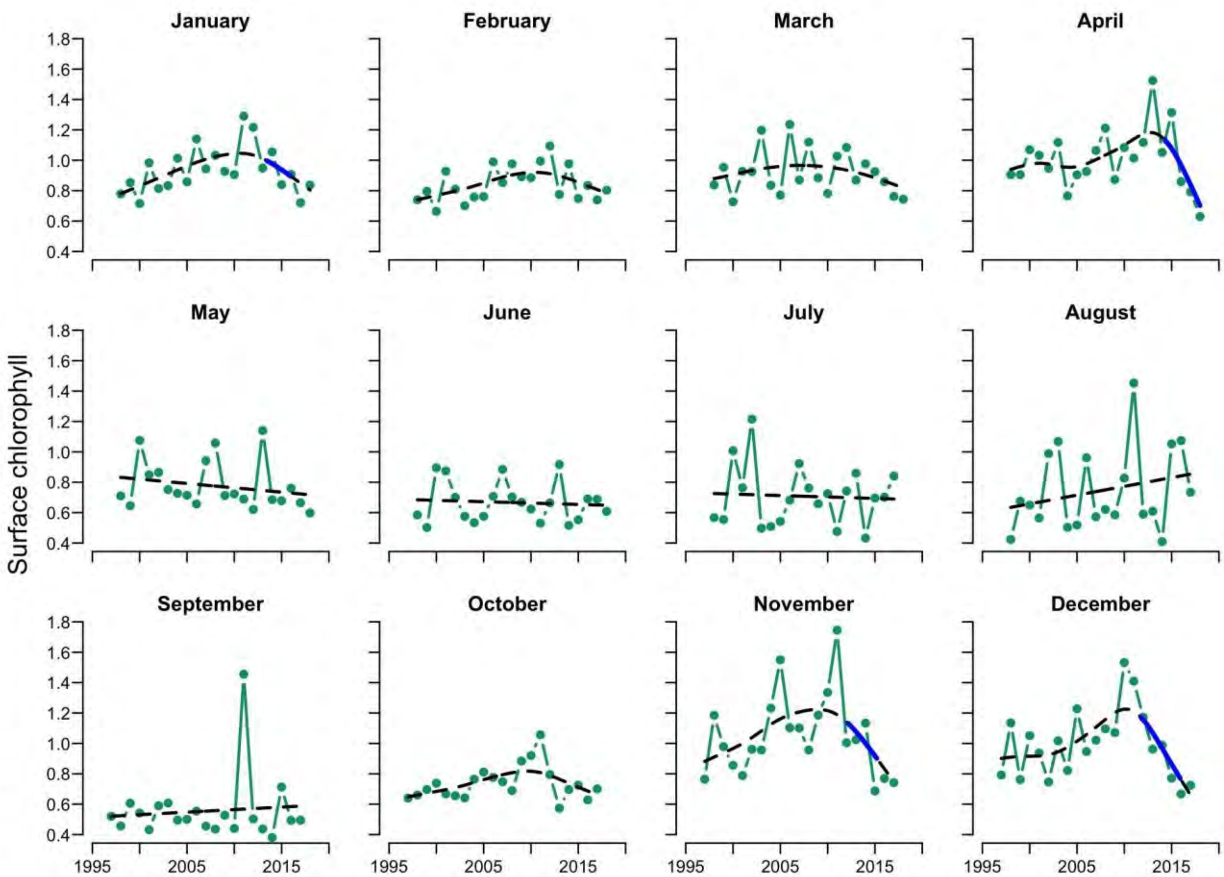


Figure 12. Monthly Chlorophyll using GAM approach where red indicates a significant increasing trend and blue indicates a significant decreasing trend and the dashed line indicates no trend.

Calanus finmarchicus abundance and Copepod size index

Zooplankton assemblages are important to monitor because with their limited mobility they are excellent indicators of water masses (Johnson et al. 2011). Zooplankton dynamics are also critical in determining the flow of energy in the food web that ultimately determines productivity of fish stocks (Pershing et al. 2005, Stock et al. 2017) and marine mammal populations (Greene and Pershing 2004, Pendleton et al. 2009, Pershing et al. 2009). *Calanus finmarchicus* is the most abundant in the Northwest Atlantic and is at its southern range limit in the NYB (Greene et al. 2003). Its abundance is largely, but not completely driving the trend in small to large copepods. When *Calanus finmarchicus* has been at low abundances the ratio increases, but when abundance was at its highest the ratio decreases (Figure 13). However, there were some years that where *Calanus finmarchicus* was unusually low (1991) that corresponded to a high copepod ratio. A lower ratio of small to large zooplankton likely leads to more fisheries productivity (Stock et al. 2017). However, some species of larval fish in the Mid-Atlantic prefer specific species of zooplankton (Friedland et al. 2013, Llopiz 2013)

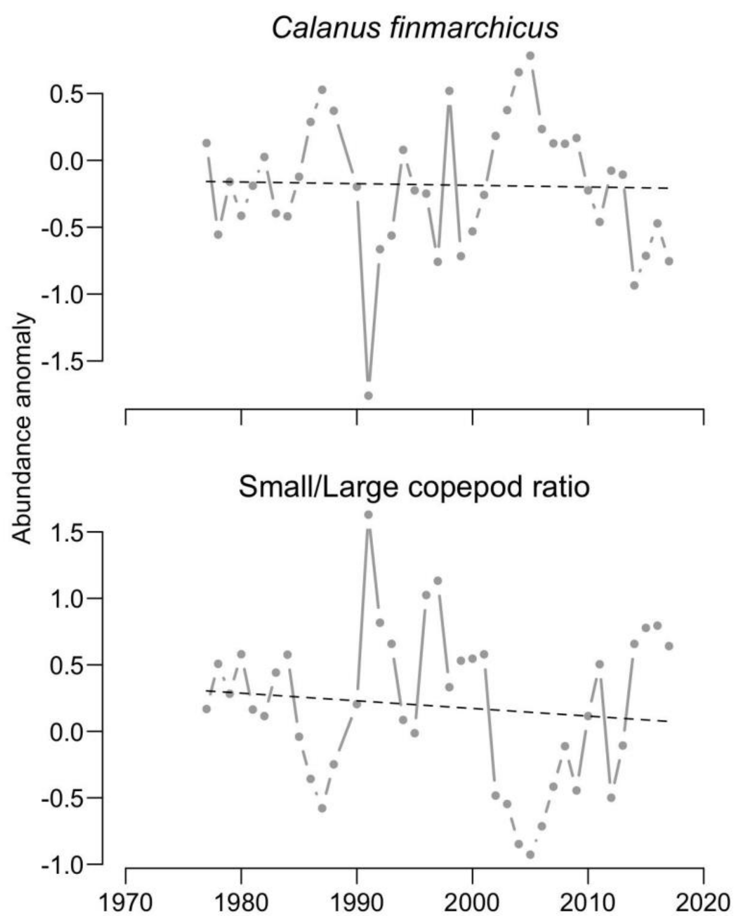


Figure 13. Zooplankton anomalies of *Calanus finmarchicus*, the largest and most abundant copepod in the Northwest Atlantic. Dashed line indicates no statistical significance (p>0.05) and the solid line indicates a statistically signficance linear trend (p<0.05).

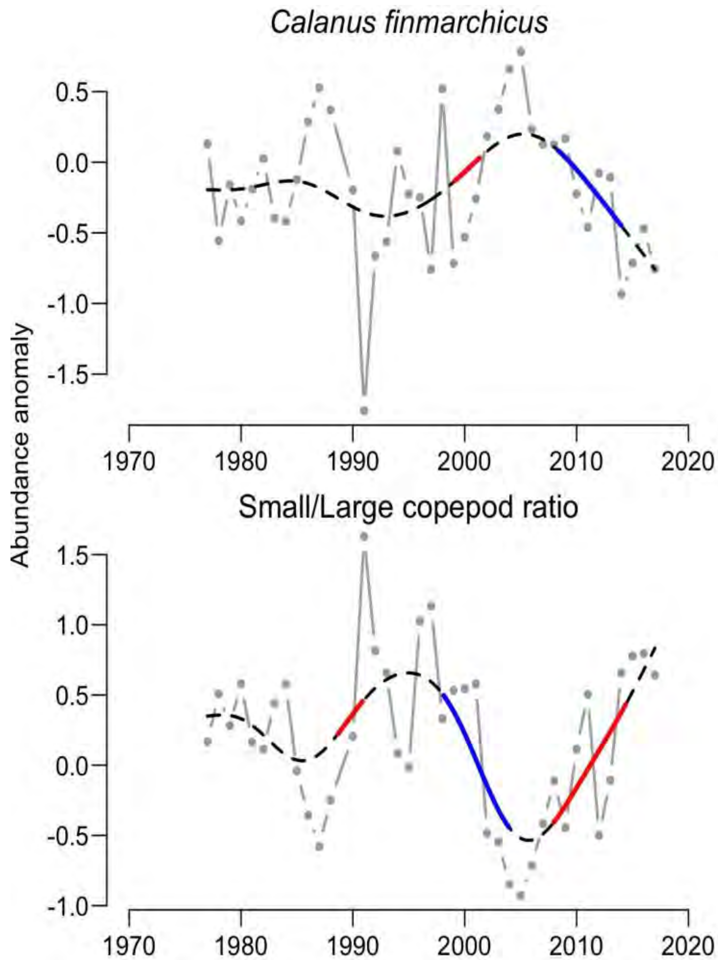


Figure 14. Zooplankton anomaly GAM fit where red indicates a significant increasing trend and blue indicates a significant decreasing trend and the dashed line indicates no trend.

American lobster biomass and Jonah crab biomass

American lobster amd Jonah crab were chosen as indicators of invertebrate fisheries in the NYB. Where lobster once supported a thriving fishery, they have since declined (Pearce and Balcom 2005). Jonah crab has become a more important fishery in many regions of the Northwest Atlantic. Peak abundances for both species occurred in the 1970s as indicated by both the fall and spring NEFSC bottom trawl survey (Figure 15). Both seem to have increased in recent years. However, the slight increase in American lobster was not evident in spring and not statistically significant in the fall and thus, not a robust trend (Figure 15).

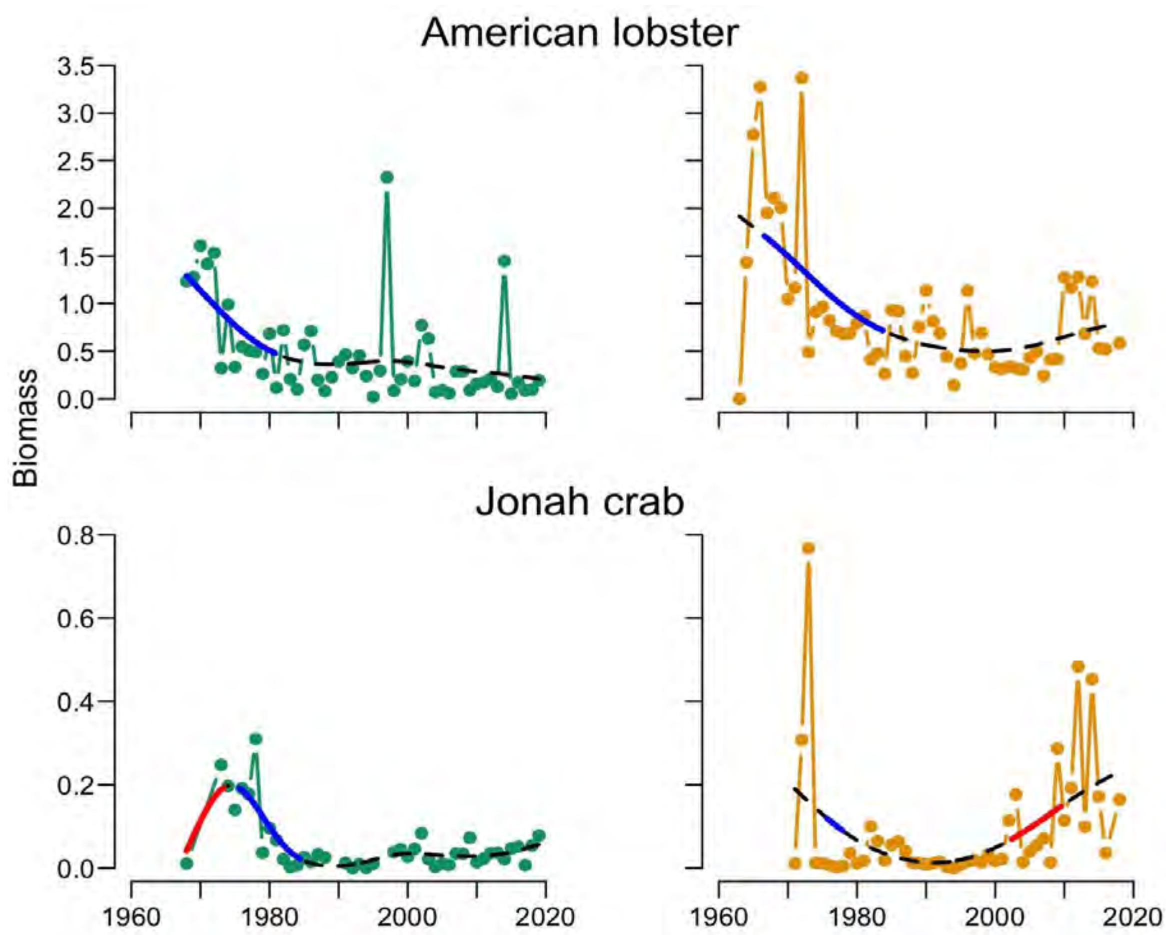


Figure 15. Lobster and Jonah crab biomass in the spring (green) and fall (orange). Solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

Forage species biomass

Forage fish are an important component of the MidAtlantic Bight food web (Link et al. 2008) and squids make up some of the most important species in New York state. When all forage species are combined, there does not appear to be any strong linear change over time in either the spring or fall surveys (Figure 16). However, when evaluated with nonlinear fitting techniques there was a significant increase in the fall survey at the beginning of the time series (Figure 17). Similarly, there was not a significant change in Northern shortfin squid over time (Figure 16), but there were some large increases identified when using nonlinear techniques (Figure 17). Interestingly, there is a decrease in longfin squid in the spring and an increase in the fall, that suggests some interesting dynamics in this species. Although these squid species do not appear to have shifted their spatial distribution (Thorne and Nye in review) as other fish species have

(Nye et al. 2009), the Mid-Atlantic State of the Ecosystem report suggests that their energy density has been consistently lower 2017-2018 compared to the 1980s and 1990s.

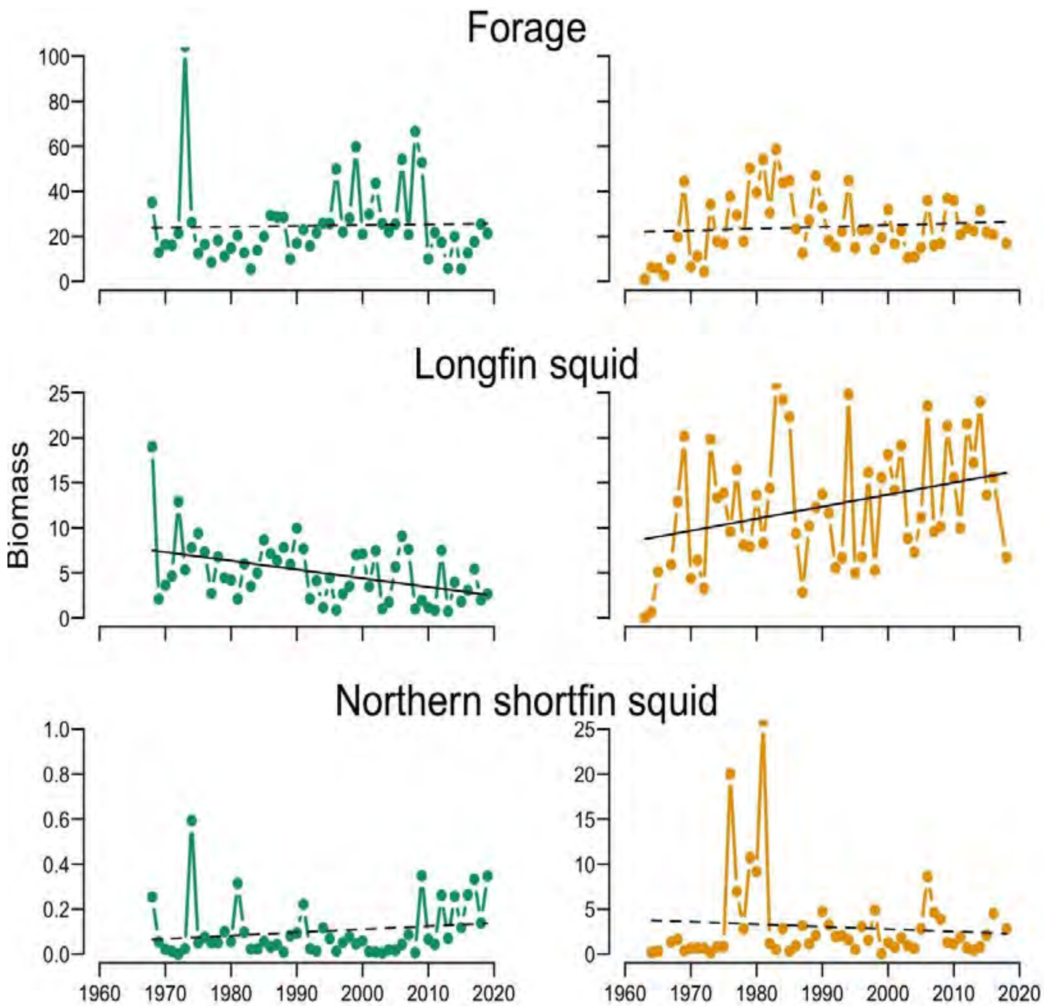


Figure 16. Forage species indicators in the spring (green) and fall (orange). Dashed line indicates no linear trend (p>0.05). Solid line indicates statistically significant linear trend (p<0.05).

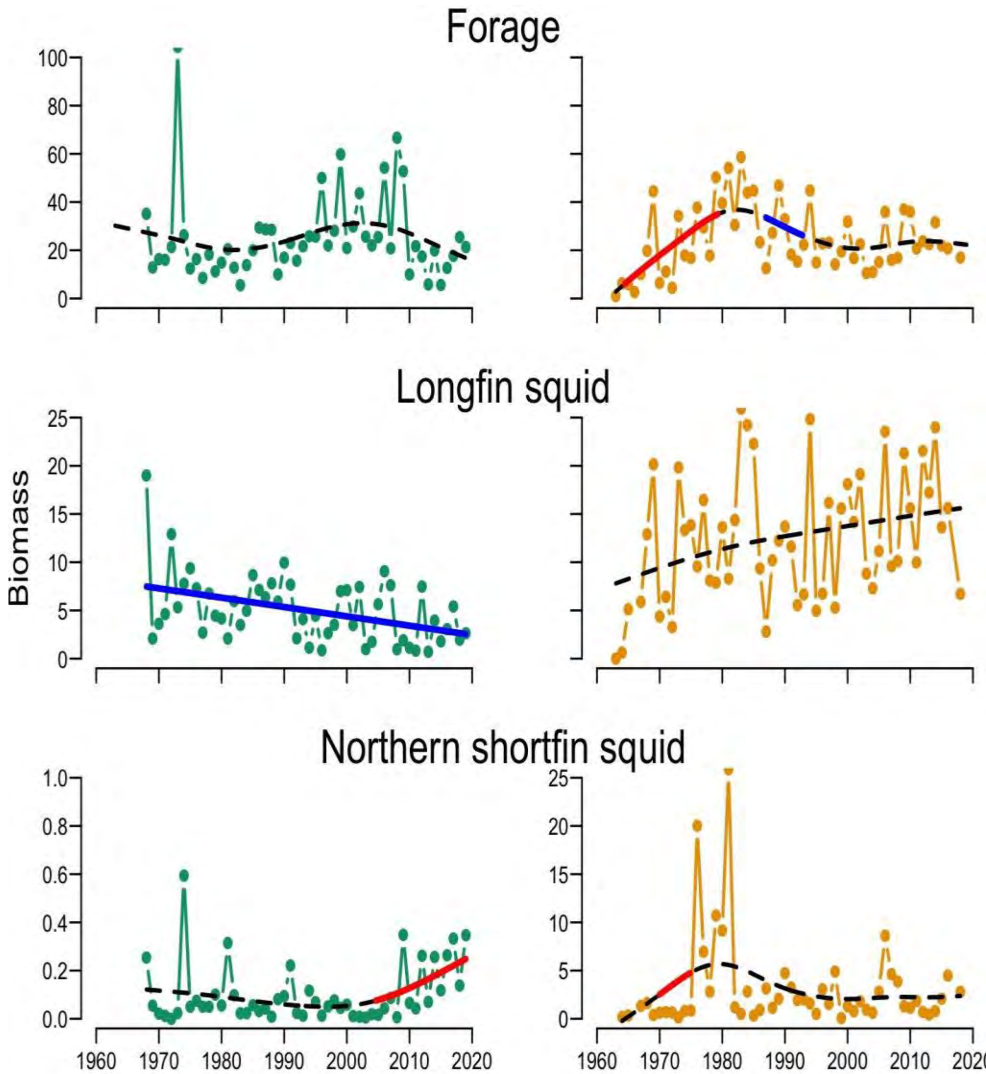


Figure 17. Forage species trends using GAM approach in the spring (green) and fall (orange) where solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

Feeding groups and fish biomass

It is intractable to meaningfully summarize the biomass times series for every species found in the NYB, thus we grouped these fish and macroinvertebrates by their trophic niches (Garrison and Link 2000). Trawl biomass has significantly declined in the fall primarily driven by a large decline in the 1960s, but this trend was not evident in the spring (Figure 18, 19). Benthic organisms have increased linearly while benthivores have decreased in the fall (Figure 18,19). No linear trends were statistically significant for planktivores, but clearly there was a peak in abundance in the early 1980s that then leveled out in the 1990s to present as evident by nonlinear fits (Figure 19). There was no trend in piscivores in the spring, but a decline in the fall, again, driven by high abundances in the past (Figure 18, 19).

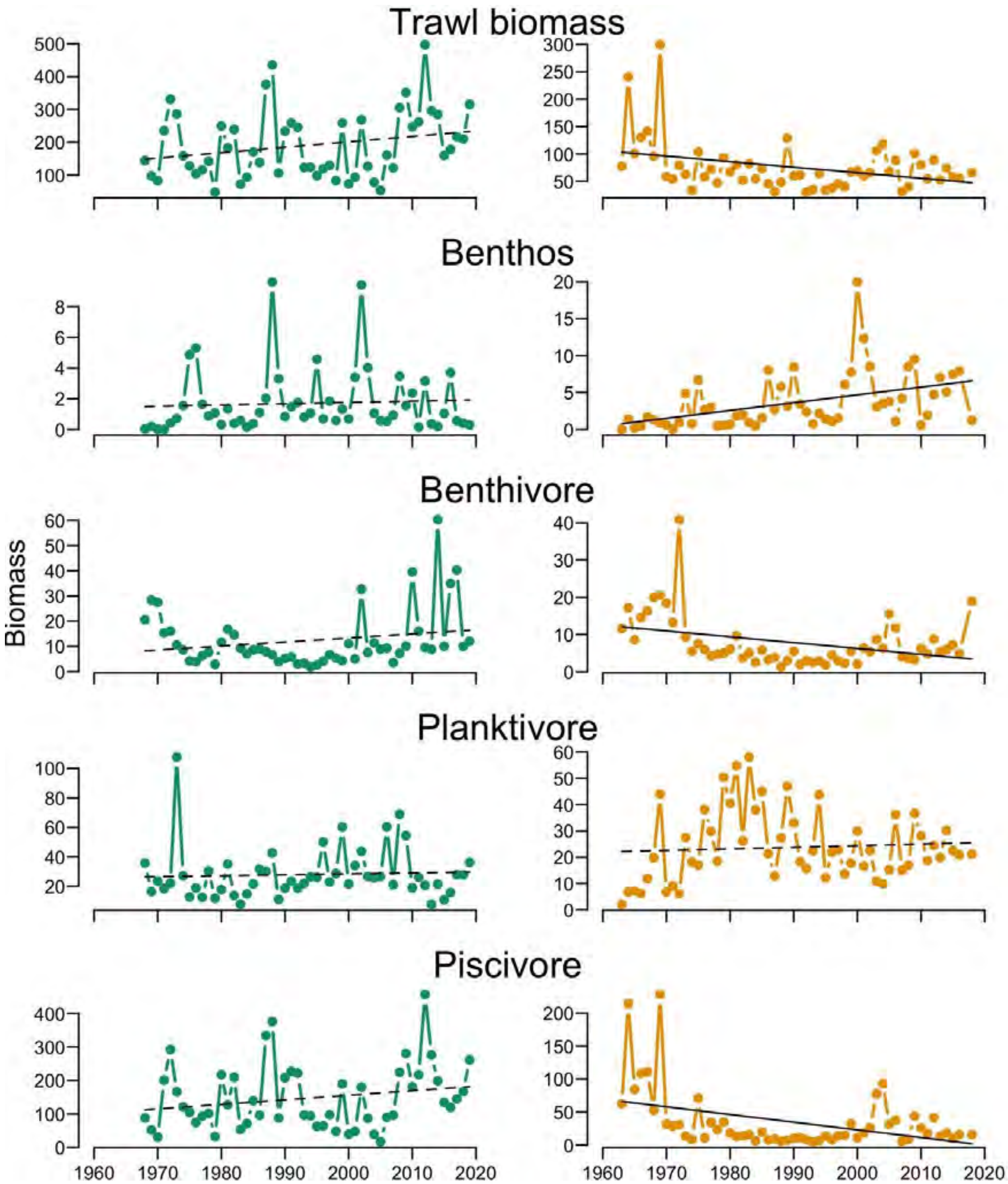


Figure 18. Mid-trophic level taxonomic groups assessed with linear regression in the spring (green) and fall (orange). The dotted line means not statistically significant (p>0.05) and solid line is statistically significant (p<0.05).

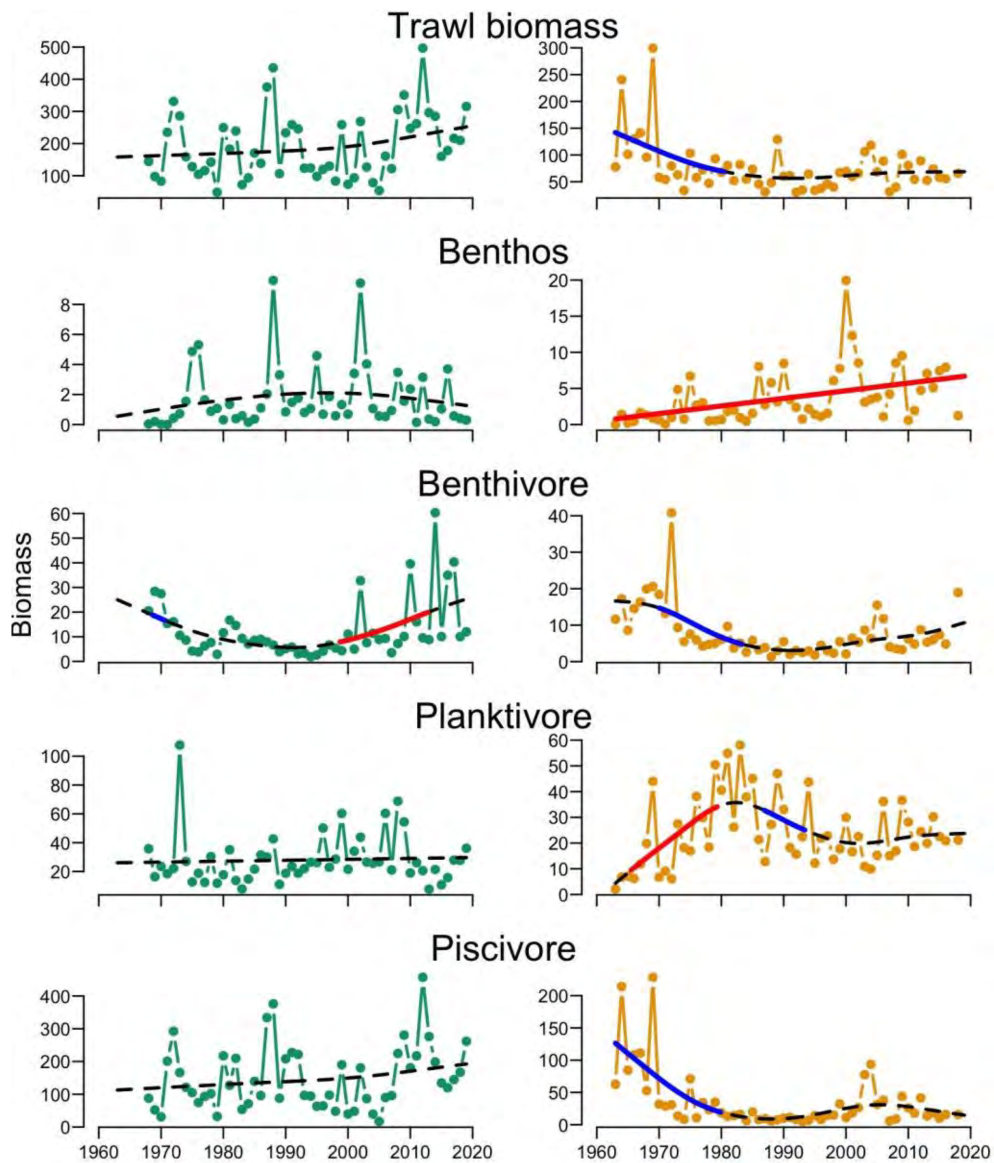


Figure 19. Mid trophic level taxanomic groups using GAM approach in the spring (green) and fall (orange) where solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

Summer Flounder Biomass and Black Sea Bass biomass

Summer flounder and black sea bass support some of the most important commercial and recreational fisheries in New York. Both species have increased from population lows in the 1980s (Figure 20, 21). However, summer flounder has declines in recent years (Figure 21) consistent with stock assessments (Terceiro 2015).

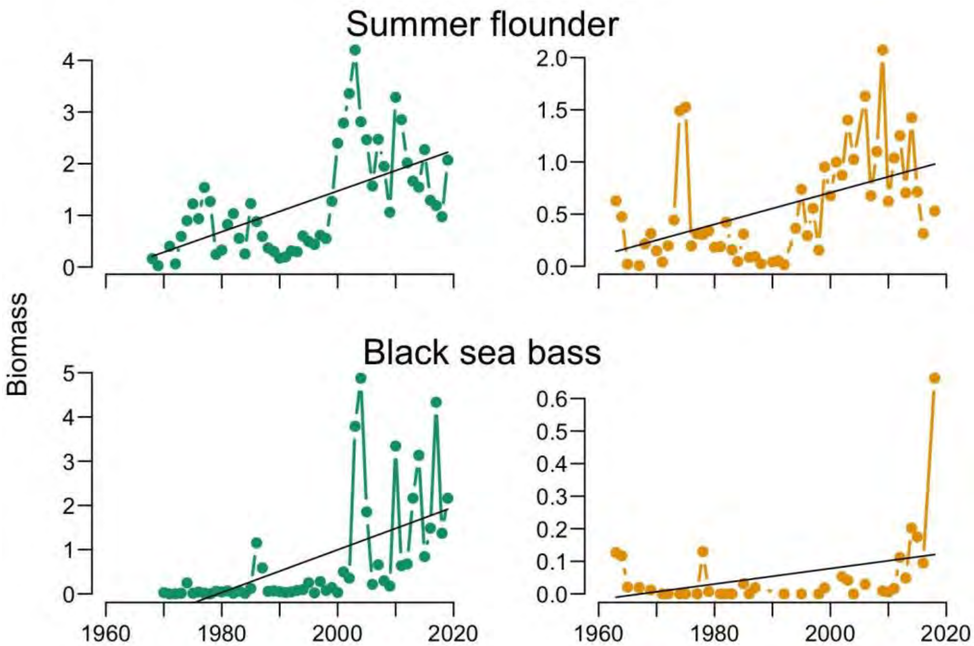


Figure 20. Summer flounder and black sea bass biomass assessed with linear regression in the spring (green) and fall (orange). The dashed line means not statistically significant (p>0.05) and solid line is statistically significant (p<0.05).

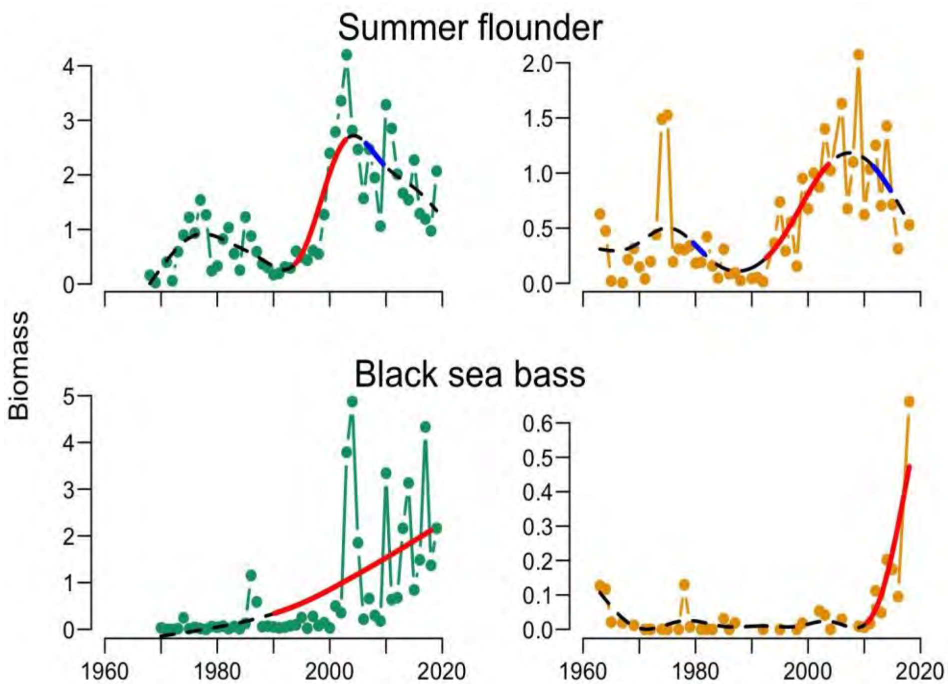


Figure 21. Summer flounder and black sea bass biomass using GAM approach in the spring (green) and fall (orange) Solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

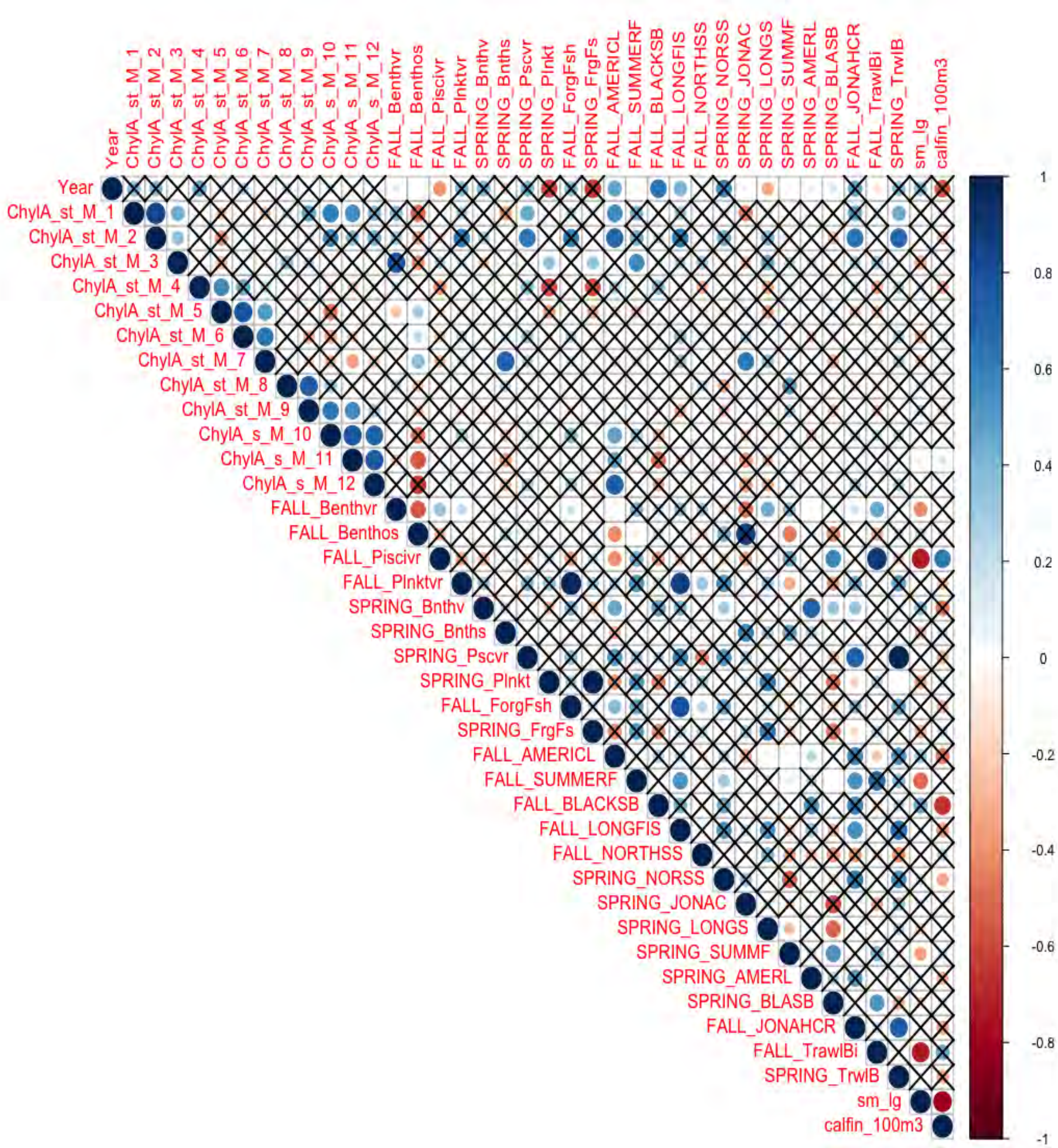


Figure 22. Pairwise correlations amongst the marine community indicators. X means no correlation

Summary for Marine community

Most of the indicators for the marine community are not correlated as indicated by the X in the above correlation matrix (Figure 22). Redundancy of similar indicators was less of an issue when selecting indicators for the marine community because there were fewer sources of data for this taxonomic level than for temperature variables. However, as might be expected, chlorophyl concentration in adjacent months were positively correlated. Benthic biomass was positively correlated with spring chlorophyl so our indicators may be picking up on the relationship between spring phytoplankton productivity and subsequent fall recruitment of benthic organisms (Figure 22). Many indicators were positively correlated because the same observations went into both indicators. For instance, the spring forage fish biomass was highly correlated with planktivores because many of the same organisms fit both of those categories. In general, the trends for the marine community were more nuanced than the environmental indices, were nonlinear and even differed between spring and fall surveys. Differences in biomass between fall and spring surveys reflect changes in migration phenology of different species.

5. Humans

Fishing is one of the many ecosystem services that the NYB marine ecosystem provides. While the total number of fish harvested by recreational fishers has decreased over time, the number of fish caught and released has increased (Figure 23, 24). The total number of fish caught, both harvested and released by recreational fishers has increased over time (Figure 25, 26). However, much of this increase is due to an increase in catch and release activity.

Recreational harvest (total number all species)

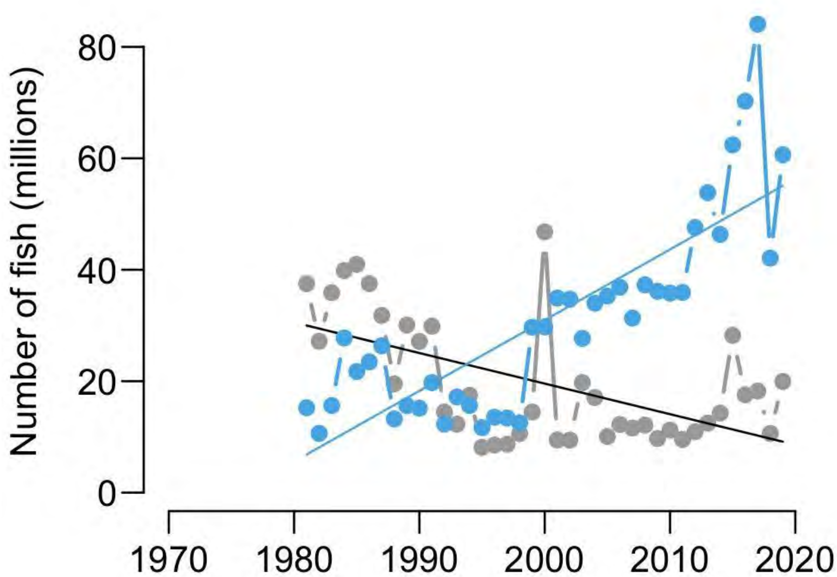


Figure 23. Number of fish in recreational fishery released alive (catch and release; light blue) has increased over time while the recreational harvest has decreased (gray) over time. A dotted line means not statistically significant (p>0.05) and solid line is statistically significant (p<0.05).

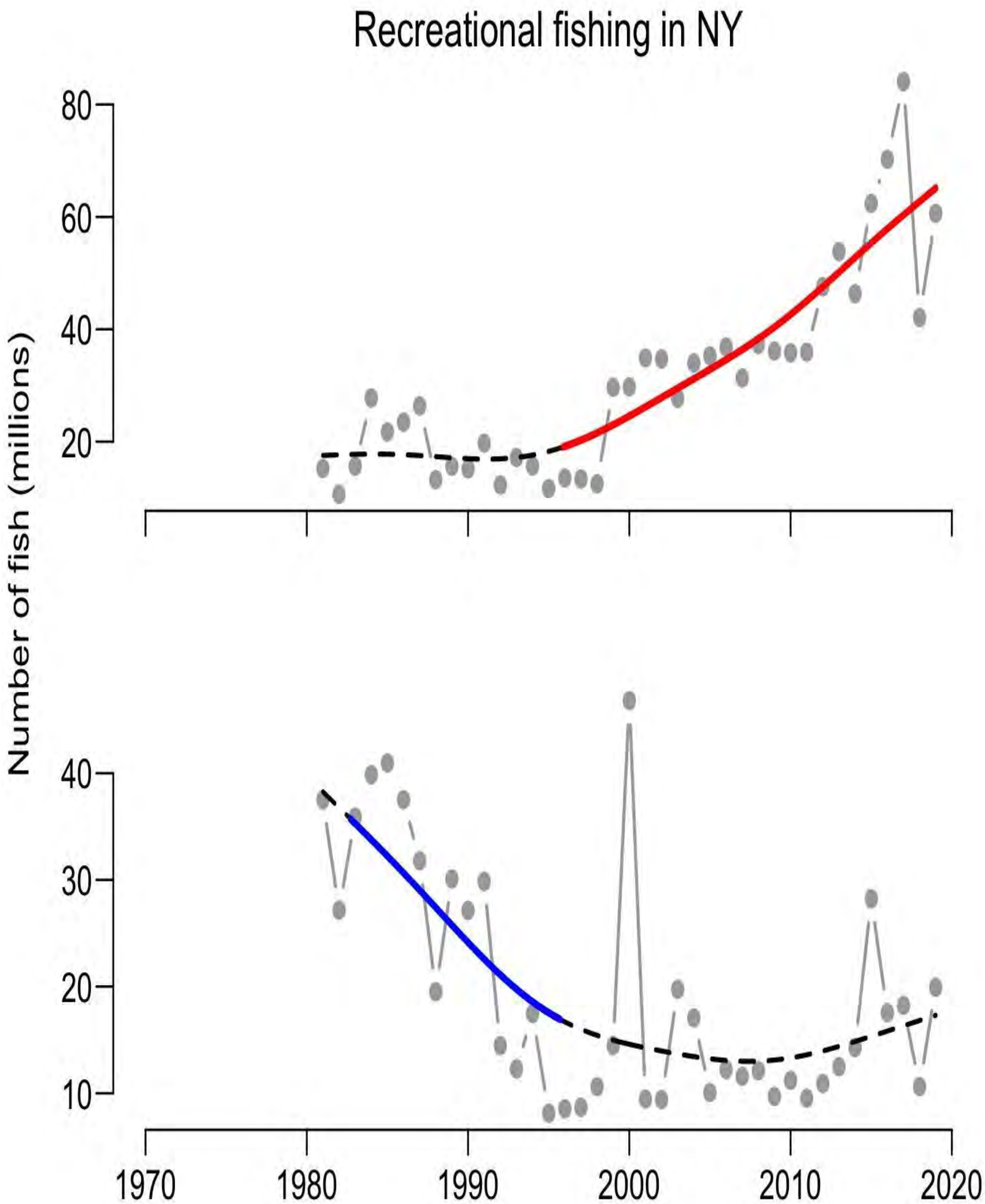


Figure 24. Number of fish in recreational fishery released alive (catch and release) using GAM approach where solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

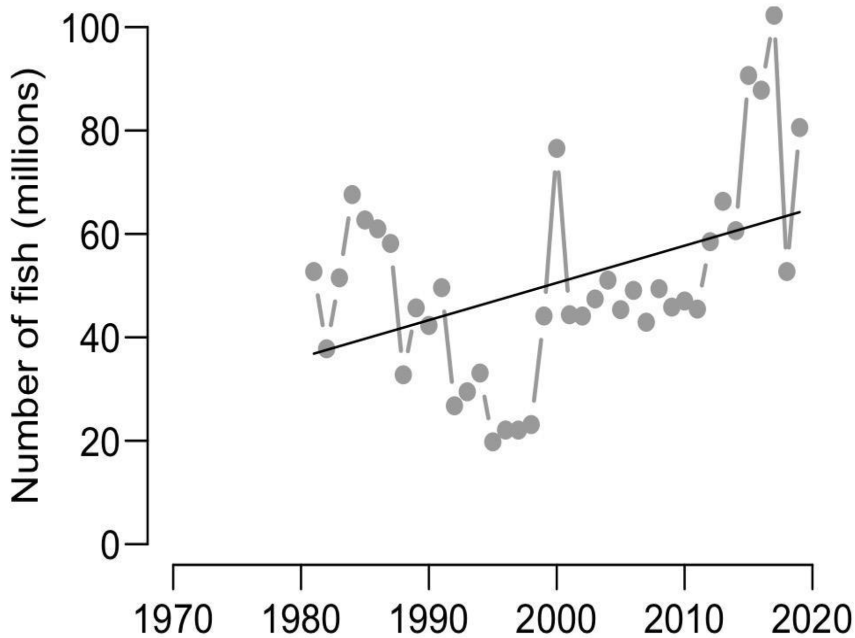


Figure 25. Total Fish Caught (harvested + released) by recreational fishers. The dotted line means not statistically significant (p>0.05) and solid line is statistically significant (p<0.05).

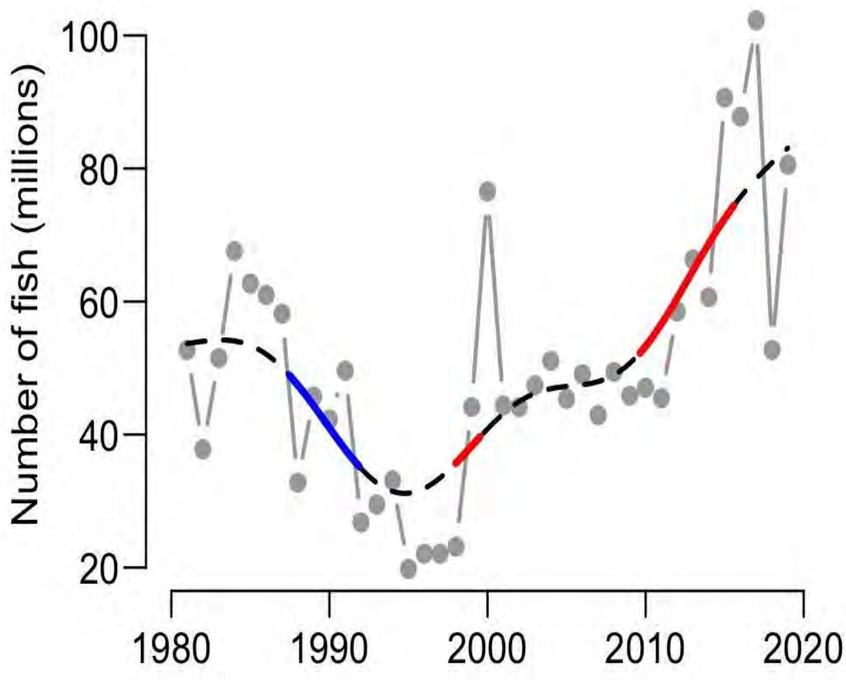


Figure 26. Recreational catch (harvested + released) using GAM approach. Solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

Commercial fishing landings and Commercial value in NY

Commercial landings statistics go back to the 1940s and commercial catch has declined primarily with a step decrease in the 1960s. However, the total landings in US dollars has not declined as quickly (Figure 27).

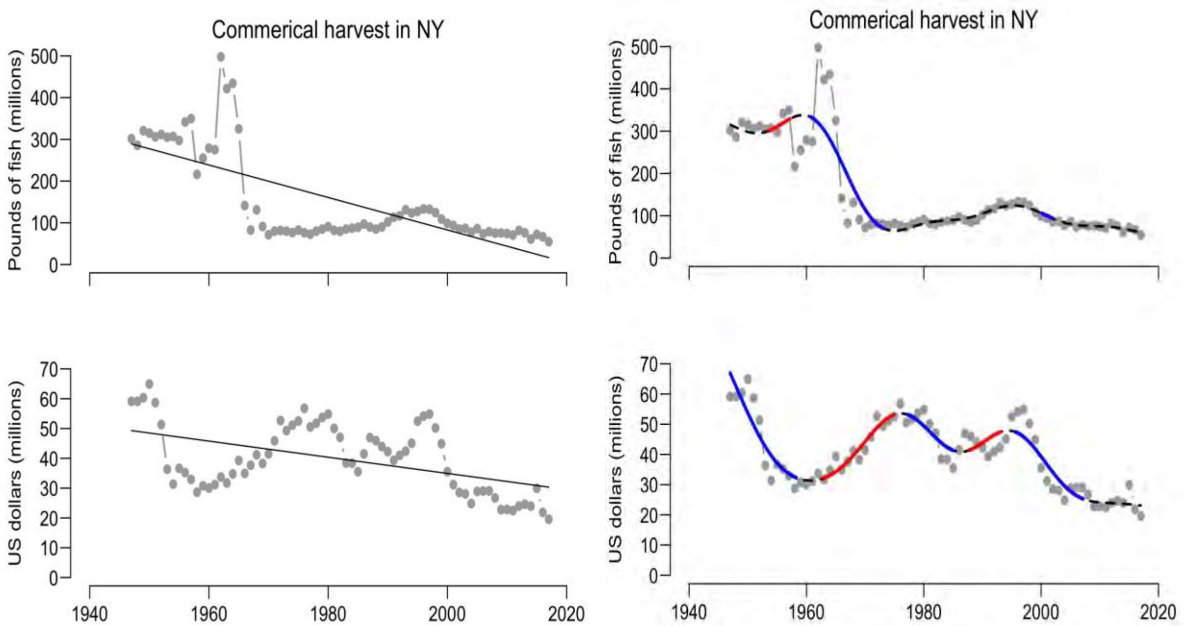


Figure 27. Commercial harvest in NY in pounds of fish and in US dollars has declined since a peak in the 1960s shown with linear trends (left panels) and GAM trends (right panels). Solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

Human population trends

The populations of Suffolk and Nassau counties, the two counties in closest proximity to the NYBs have increased dramatically over time but has leveled off since 2010.

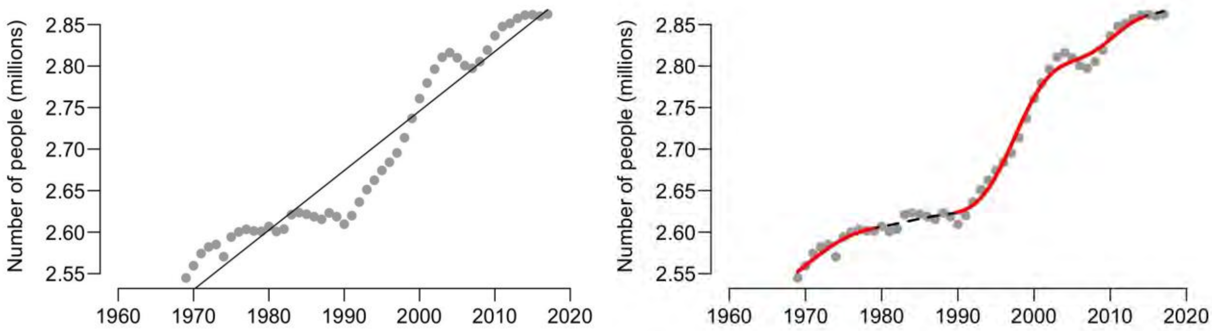


Figure 28. Humans of Suffolk and Nassau county with linear (left panel). The dotted line means not statistically significant (p>0.05) and solid line is statistically significant (p<0.05). The right panel is the GAM (right panel) trend Solid red lines indicate a significant increasing trend, the blue indicates a significant decreasing trend and the dashed line indicates no trend.

Figure 29. Sea level risk map for LI communities

Sea level risk can be estimated in a variety of ways. We show here a sea level risk indicator by coastal community developed by the NOAA NEFSC Social Science branch. It is calculated using the land area at elevations from one to six feet above mean higher high water within each community boundary using coastal elevation data developed by the NOAA High-resolution digital elevation models (DEMs; 5 to 10-meter horizontal resolution) from the National Elevation Dataset for coastal areas of the U.S. The estimates are adjusted for variations in local tide levels and a 'bathtub' approach is used. The communities with the highest exposure to sea level rise are Queens and King County and the south shore of Long Island.

6. Conclusions

The increases in water temperature, marine heatwaves and river flow that we have documented are consistent with the effects projected for the Northeast US due to climate change (IPCC 2013). Stronger stratification and an associated decrease in primary productivity has also been generally predicted as a result of climate change in general. We did not detect any trend in stratification over time in the NYB, but we did see a recent decline in chlorophyll in the Fall and in April since 2010. While not a direct measurement of primary productivity, the decline in chlorophyll suggests weaker blooms in both the spring and summer coincident with the sharp increase in temperature and marine heatwave days. Upper trophic level production is generally higher when large diatoms and large zooplankton dominate the base of the food chain (Stock et al. 2017). Thus, an important indicator of food web dynamics is the ratio of small to large zooplankton. As waters warm large northern-affiliated zooplankton are expected to decrease in abundance and frequency while smaller zooplankton are expected to become more dominant. Indeed, *Calanus finmarchicus*, the largest and historically most abundant copepod in the Northeast US has also declined in abundance and smaller copepods have become relatively more abundant. How this affects fish production and protected species is the subject of future work.

The trends in phytoplankton and zooplankton appear to be reflective of the environmental conditions, but the trends in upper trophic levels were highly nonlinear and variable between spring and fall surveys. Starting in the 1970s, there was a decrease in American lobster trawl survey biomass which coincides with the known decline and crash of the lobster fishery in New York in the 1990s (Pearce and Balcom 2005). Jonah crabs and other benthic organisms also decreased at around the same time as lobster, but Jonah crab has had a recent uptick since 2010. There was no apparent trend in forage fish biomass or in shortfin squid in the NYB, but an increase in longfin squid in the fall and a decrease in the spring. There was an overall decrease in total biomass of fish and macroinvertebrates in the trawl survey and a strong increase in benthic macroinvertebrates since the 1960s. For fish classified by trophic guild, there was a concomitant decrease in benthivore and piscivores and an increase in planktivores over the 50+ year times series likely associated with historic overfishing. There has been an increase in summer flounder and black sea bass since the 1990s, but there has been a recent decrease in summer flounder. The total number of fish caught (harvested and released) in NY state has increased with an increase in recreational catch and a decrease in commercial harvest. An increase in recreational fishing has been observed in many regions of the US (Coleman et al. 2004, Ihde et al. 2011).

The population of Suffolk and Nassau county, the counties most closely associated with the NYB ocean habitat, increased dramatically since the 1970s, but that growth has recently leveled out. The coastal communities on the south shore of Long Island as well as Queens and King counties are the most vulnerable to sea level rise.

7. References

Coleman, F. C., W. F. Figueira, J. S. Ueland, and L. B. Crowder. 2004. The impact of United States recreational fisheries on marine fish populations. Science **305**:1958-1959.

Curtis, C. J., and G. L. Simpson. 2014. Trends in bulk deposition of acidity in the UK, 1988–2007, assessed using additive models. Ecological Indicators **37**:274-286.

Friedland, K. D., J. Kane, J. A. Hare, R. G. Lough, P. S. Fratantoni, M. J. Fogarty, and J. A. Nye. 2013. Thermal habitat constraints on zooplankton species associated with Atlantic cod (Gadus morhua) on the US Northeast Continental Shelf. Progress in Oceanography **116**:1-13.

Gangopadhyay, A., G. Gawarkiewicz, E. N. S. Silva, M. Monim, and J. Clark. 2019. An observed Regime Shift in the formation of Warm core Rings from the Gulf Stream. Scientific Reports **9**:1-9.

Garrison, L. P., and J. S. Link. 2000. Dietary guild structure of the fish community in the Northeast United States continental shelf ecosystem. Marine Ecology Progress Series **202**:231-240.

Gawarkiewicz, G., K. Chen, J. Forsyth, F. Bahr, A. M. Mercer, A. Ellertson, P. Fratantoni, H. Seim, S. Haines, and L. Han. 2019. Characteristics of an Advective Marine Heatwave in the Middle Atlantic Bight in Early 2017. Frontiers in Marine Science **6**.

Greene, C. H., and A. J. Pershing. 2004. Climate and the conservation biology of North Atlantic right whales: the right whale at the wrong time? Frontiers in Ecology and the Environment **2**:29-34.

Greene, C. H., A. J. Pershing, A. Conversi, B. Planque, C. Hannah, D. Sameoto, E. Head, P. C. Smith, P. C. Reid, and J. Jossi. 2003. Trans-Atlantic responses of *Calanus finmarchicus* populations to basin-scale forcing associated with the North Atlantic Oscillation. Progress In Oceanography: Achievements of the continuous plankton recorder survey and a vision for its future **58**:301-312.

Hastie, T., and R. Tibshirani. 1987. Generalized additive models: some applications. Journal of the American statistical association **82**:371-386.

Hobday, A. J., E. C. Oliver, A. S. Gupta, J. A. Benthuysen, M. T. Burrows, M. G. Donat, N. J. Holbrook, P. J. Moore, M. S. Thomsen, and T. Wernberg. 2018. Categorizing and naming marine heatwaves. Oceanography **31**:162-173.

Ihde, T. F., M. J. Wilberg, D. A. Loewensteiner, D. H. Secor, and T. J. Miller. 2011. The increasing importance of marine recreational fishing in the US: challenges for management. Fisheries Research **108**:268-276.

IPCC. 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. New York, NY USA.

Johnson, C. L., J. A. Runge, K. A. Curtis, E. G. Durbin, J. A. Hare, L. S. Incze, J. S. Link, G. D. Melvin, T. D. O'Brien, and L. Van Guelpen. 2011. Biodiversity and ecosystem function in the Gulf of Maine: pattern and role of zooplankton and pelagic nekton. PLOS One **6**:e16491.

Link, J., W. Overholtz, J. O'Reilly, J. Green, D. Dow, D. PAlka, C. Legault, J. Vitaliano, V. Guida, M. Fogarty, J. Brodziak, L. Methratta, W. Stockhausen, L. Col, and C. Griswold. 2008. The Northeast US continental shelf Energy Modeling and Analysis exercise

(EMAX): Ecological network model development and basic ecosystem metrics. Journal of Marine Systems **74**:453-474.
Llopiz, J. K. 2013. Latitudinal and taxonomic patterns in the feeding ecologies of fish larvae: A literature synthesis. Journal of Marine Systems **109**:69-77.
Mallakpour, I., and G. Villarini. 2017. Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. Theoretical and Applied Climatology **130**:345-363.
Nye, J. A., J. S. Link, J. A. Hare, and W. J. Overholtz. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Marine Ecology Progress Series **393**:111-129.
Oliver, E. C. J., J. A. Benthuysen, S. Darmaraki, M. G. Donat, A. J. Hobday, N. J. Holbrook, R. W. Schlegel, and A. S. Gupta. 2020. Marine Heatwaves. Annual review of marine science.
Oliver, E. C. J., M. G. Donat, M. T. Burrows, P. J. Moore, D. A. Smale, L. V. Alexander, J. A. Benthuysen, M. Feng, A. Sen Gupta, A. J. Hobday, N. J. Holbrook, S. E. Perkins-Kirkpatrick, H. A. Scannell, S. C. Straub, and T. Wernberg. 2018. Longer and more frequent marine heatwaves over the past century. Nature Communications **9**:1324.
Pearce, J., and N. Balcom. 2005. The 1999 Long Island Sound lobster mortality event: Findings of the comprehensive research initiative. Journal of Shellfish Research **24**:691-697.
Pendleton, D. E., A. J. Pershing, M. W. Brown, C. A. Mayo, R. D. Kenney, N. R. Record, and T. V. Cole. 2009. Regional-scale mean copepod concentration indicates relative abundance of North Atlantic right whales. Marine Ecology Progress Series **378**:211-225.
Pershing, A. J., C. H. Greene, J. W. Jossi, L. O'Brien, J. K. T. Brodziak, and B. A. Bailey. 2005. Interdecadal variability in the Gulf of Maine zooplankton community, with potential impacts on fish recruitment. ICES Journal of Marine Science: Journal du Conseil **62**:1511-1523.
Pershing, A. J., N. R. Record, B. C. Monger, C. A. Mayo, M. W. Brown, T. V. N. Cole, R. D. Kenney, D. E. Pendleton, and L. A. Woodward. 2009. Model-based estimates of the right whale habitat use in the Gulf of Maine. Marine Ecology Progress Series **378**:245-257.
Santibáñez, P. A., O. J. Maselli, M. C. Greenwood, M. M. Grieman, E. S. Saltzman, J. R. McConnell, and J. C. Priscu. 2018. Prokaryotes in the WAIS Divide ice core reflect source and transport changes between Last Glacial Maximum and the early Holocene. Global Change Biology **24**:2182-2197.
Stock, C. A., J. G. John, R. R. Rykaczewski, R. G. Asch, W. W. L. Cheung, J. P. Dunne, K. D. Friedland, V. W. Y. Lam, J. L. Sarmiento, and R. A. Watson. 2017. Reconciling fisheries catch and ocean productivity. Proceedings of the National Academy of Sciences.
Terceiro, M. 2015. Stock Assessment update of Summer flounder for 2015. US Department of Commerce, Woods Hole, MA.
Thomas, A., R. Weatherbee, A. Pershing, K. Mills, J. Nye, and M. Henderson. 2016. SST Anomalies, Seasonality and Phenology Changes on the NE US and Scotian Shelf.*in* AGU Fall Meeting Abstracts.

Appendix 1: New York Bight Indicator Report 2020

Marine ecosystem indicators are sensitive to ecosystem boundaries and spatial scale

Kurt C. Heim[1], Lesley H. Thorne[1], Joseph D. Warren[1], Jason S. Link[2], Janet A. Nye[1]

[1]School of Marine and Atmospheric Science, Stony Brook University, Stony Brook, NY, USA

[2]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Office of the Assistant Administrator, Woods Hole, MA, USA

Keywords
Ecosystem indicators; Problem of scale; Spatial scale; Ecosystem Based Management; Dynamic Ocean Management; Marine Strategy Framework Directive; Marine Ecosystems

Abstract

Time series indicators are widely used in ecosystem-based management. A suite of indicators is typically calculated for a static region or subregions and presented in an ecosystem assessment (EA). These are used to guide management decisions or determine good environmental status (GES) related to the Marine Strategy Framework Directive. Yet, few studies have examined how the spatial scale of an EA influences indicator behavior. We explore this question using the Northwest Atlantic continental shelf ecosystem (USA). We systematically divided the ecosystem in 31 different ways, covering spatial extents from 250,000 km^2 (a single ecosystem unit) to 20,000 km^2 (ten subunits). The same 22 indicators were calculated for each unit, assessed for trends, and evaluated as 31 independent EAs. We found that the detected signals of ecosystem change depended on the way the ecosystem was defined or subdivided. A single EA for the whole region differed by 23% (in terms of the 22 indicator trends) relative to ones for spatially nested 120,000 km^2 subunits, and by up to 36% for EAs at smaller scales. Indicator trend disagreement occurred because (1, most common) a local trend propagated and was perceived as widespread, (2, common) a local trend was masked by aggregating data over a large region, or (3, least common) a local trend switched direction when considered at a broader scale. Yet, there was variation among indicators in their scale sensitivity related to trophic level. Indicators of temperature, chlorophyll-a, and zooplankton were spatially coherent: trends portrayed were similar regardless of scale. Mid-trophic level indicators (fish and invertebrates) showed more spatial variation in trends. We also compared trend magnitude and indicator values to spatial extent and found relationships consistent with scaling theory. Indicators at broad scales produced subdued trends and values relative to indicators developed at smaller spatial scales, which often portrayed 'hotspots' of local abundance or strong trend. Our results imply that subsequent uses of indicators (e.g., determining GES, risk assessments, management decisions) are also sensitive to ecosystem delineation and scale. Ultimately, indicators and EAs would be wisely evaluated at multiple spatial scales and complimented with spatially explicit analysis to reflect the hierarchical structure of ecosystems. One scale is not best, rather we gain a new level of understanding at each scale examined, that can contribute to management decisions in a multiscale governance framework characterized by goals and objectives relevant at different scales.

1. Introduction

Integrating information across disciplines in marine science is often facilitated by indicators, which we define here as quantitative measurements that represent key attributes of interest. When a vetted suite of indicators is assembled, it can be used to assess ecosystem status, drivers of change, and performance of management actions (Rice and Rochet 2005; Shin and Shannon 2010; Shin et al. 2010). Indicators have come to play a central role in implementing Ecosystem Based Management (EBM) and an Ecosystem Approach to Fisheries Management (EAFM) and are widely considered an important bridge making ecosystem science accessible to policy makers (Turnhout et al. 2007). Indeed, some governments now explicitly require indicator assessments through formal policy (e.g., determining Good Environmental Status [GES] in the European Union) and others have turned to indicator suites to fulfill or guide a variety of interconnected regulatory mandates (Levin et al. 2013; Foran et al. 2016).

Given their increasingly prominent role in ecosystem science and management of marine systems, a large body of research has focused on how to select the best indicators (Rice and Rochet 2005; Queirós et al. 2016; Tam et al. 2017; Otto et al. 2018). Other studies provide advice on how to use indicators in the context of marine ecosystem management (Large et al. 2013, 2015; Levin and Möllmann 2015; Burthe et al. 2016; Bal et al. 2018; Gaichas et al. 2018). In the US, much of this work is in support of regional integrated ecosystem assessments (IEAs) by the National Marine Fisheries Services (NMFS) (Levin et al. 2013; Harvey et al. 2017), and in the European Union, much of the work supports the Marine Strategy Framework Directive and assessing GES by member countries (Borja et al. 2016). This body of work emphasizes indicator selection, aggregation, interpretation, and use for management, yet the sensitivity of commonly used indicators to spatial scale is a relatively unexplored but very important topic.

To generate *any* quantitative indicator for testing and eventual use, one must first define a spatial region (or regions) over which to collect or summarize data. Regions for which indicators time series are calculated may be the entire ecosystem or different ecosystem subunits (e.g., ecological production units in the US or assessment regions of regional seas in the European Union). In any case, scientists and managers must decide (1) how large of a spatial extent is appropriate to define the ecosystem (2) how and where should the exact boundaries be drawn (3) should ecosystem subunits be defined (4) how and where should subunit boundaries occur. These decisions should ideally integrate an understanding of spatial structure in key ecosystem processes with jurisdictional boundaries and policies (Levin et al. 2013). This latter consideration is a major challenge (Crowder et al. 2006). What works politically might make little ecological sense (or vice versa). Regardless of the reasoning or methods used to draw boundaries, any boundary system is likely to have some impact on how indicators behave (**Figure 1**). Moreover, since indicators are central to many next steps in EBM or EAFM (e.g., ecosystem assessment, risk assessment, management strategy evaluation; Bunnefeld et al. 2011; Gaichas et al. 2018), the sensitivity of indicators to spatial scale could be manifested across the entire management process.

Ecological theory and a broader recognition of spatial structuring in ecosystems all but guarantee some indicators are sensitive to spatial scale. The 'problem of scale' (*sensu* Levin et al. 1992) is that processes can show different trends or relationships depending on the focal region of a study (i.e., the spatial extent) and the spatial units used to sample or analyze data (i.e., the spatial grain). For example, based on a small number of independently studied breeding populations a simple 'sea ice hypothesis' posited that some penguin species benefitted from sea ice decline (Fraser et al. 1992). Yet a spatially explicit metapopulation analysis, including 70 populations, showed little support for the hypothesis (Lynch et al. 2012). Changing extent (e.g., one or two populations vs. 70) revealed a single population misrepresented regional dynamics. In another example, Rose and Leggett (1990) demonstrate that the correlation between a predator and its prey is scale dependent; at large sampling grain predators and prey show strong positive correlation (co-occur in space), but at fine sampling grain predator avoidance is detected and correlations become negative. The main lesson is that when processes are spatially structured, choices related to extent and grain of analysis can markedly influence results (Levin 1992). And since advances in spatial statistics, geographic information systems, and genomics continue to reveal that spatial structure is pervasive in marine ecosystems (e.g., marine heatwaves, Oliver et al. 2018; zooplankton, Morse et al. 2017; fish, Ciannelli et al. 2008; Bradbury et al. 2013; penguins, Lynch et al. 2012; humans, Colburn et al. 2016), exploring the consequences of spatial structure to indicator behavior is a pressing need.

We conceptually illustrate some potential issues related to defining boundaries as they pertain to indicators (**Figure 1**). For a process operating systematically across a large spatial domain it would matter little how the boundaries of an ecosystem were defined (**Figure 1A**). All treatments of any data subset would reveal the same trend. Such an indicator might be considered robust to ecosystem delineation or the spatial scale of analysis. However, if there is spatial heterogeneity in a trend (i.e., a north to south gradient, an east to west gradient, or patchy populations and metapopulation dynamics, **Figure 1B - D**) then how one defined the spatial boundaries of the ecosystem will alter the perceived trend. If the five processes are considered at the same time, we find that any particular boundary system chosen might work well for some processes, but not others.

Here, we recreate the conceptual analysis portrayed in **Figure 1** with actual data from the Northeast United States Continental Shelf (NES) large marine ecosystem (Sherman 1991) to explore indicator scale sensitivity. Specifically, our objectives were to (i) determine how changing the spatial extent and region covered by an ecosystem assessment (EA) alters one's perception of trends (ii) to examine which indicators are most sensitive to spatial scale and (iii) explore the relationship between grain, trend strength, and values of indicators. The NES is an ideal region for this exploration because there is abundant long-term data, and a great deal of empirical work on specific ecosystem processes (i.e., fisheries, zooplankton, oceanography) and EAs to provide context for our exploratory results (Link et al. 2002; EcoAP 2009, 2011; NEFSC 2017a, 2017b, 2018a, 2018b). Our goal therefore is not to elucidate novel ecological or oceanographic trends in the NES, re-define or evaluate how boundaries are currently defined in the NES, or develop new and thoroughly vetted indicators for application. Rather, our study is intended to highlight issues to be aware of when developing EAs that use indicators. The potential issues related to scale highlighted in this study will apply to any scenario where boundaries are drawn and indicators are calculated, and so should be globally applicable in both marine and terrestrial systems.

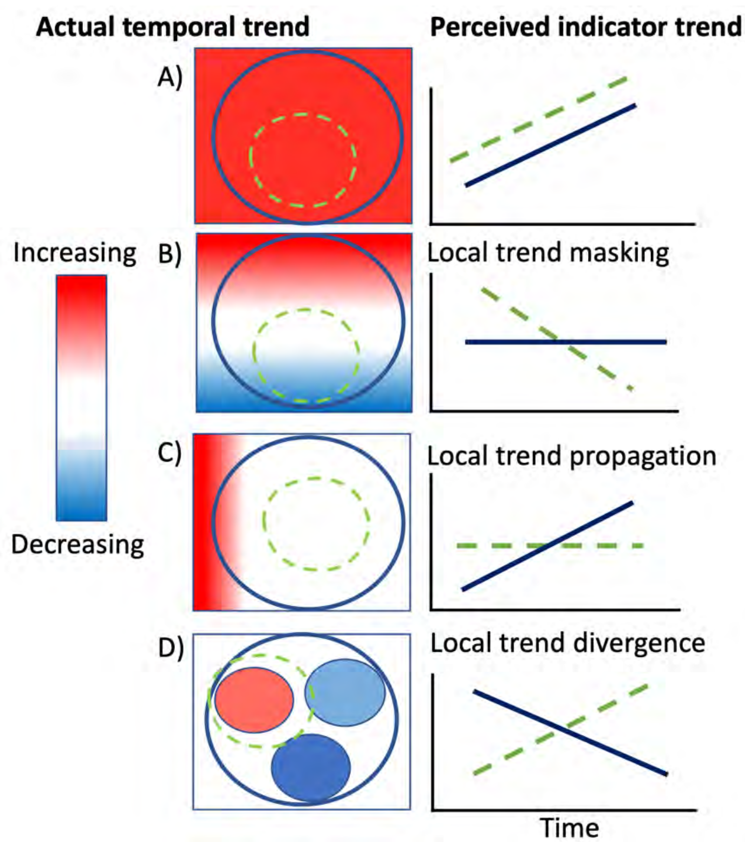


Figure 1. Hypothetical scenarios of how the spatial extent and boundaries influence indicator trends. A 'true' spatial pattern of temporal trend exists in each domain (left boxes), but the perception of trend is scale dependent. Scenarios include (A) consistent trends perceived at all spatial scales; (B) when no trend is perceived at a large extent, but strong local trend(s) exist within the spatial domain; (C) when a trend is perceived at large extent, but driven by a highly localized trend; (D) when a trend is present at large extent, but a different trend is perceived at a local scale within this region.

2. Methods

2.1. Study area background and existing spatial delineations

The NES is a large marine ecosystem (Sherman 1991) that covers a spatial area of 250,000 km^2 extending from Cape Hatteras (USA, North Carolina) in the South to Novia Scotia (CA) in the North (34 - 45 °N, **Figure 2A**). Fisheries are currently managed from a single-species perspective by two regional fisheries management councils, yet there is growing interest and commitment to implement EBFM (Levin et al. 2013). To this end, an IEA has been adopted as the cornerstone of the federal approach to integrating ecosystem information to inform management decisions (Levin et al. 2013; Harvey et al. 2017). Indicators are a foundational component of the IEA and are regularly calculated for four distinct spatial regions within the NES (Lucey and Fogarty 2013). These include the Gulf of Maine (69,000 km^2), the Scotian Shelf (31,000 km^2), Georges Bank (58,000 km^2), and the Mid-Atlantic Bight (126 km^2). These are considered, at present, the appropriate spatial scale to conduct an IEA (DePiper et al. 2017). Indicators used are multidisciplinary time-series contributed by over 30 scientists each year, which are assembled, analyzed with a consistent trend fitting technique, and presented in an EA report. This provided a good model with well-developed indicators, based on extensive survey data, to explore our research questions.

2.2. General approach

We created a suite of 31 assessment region boundaries that can be thought of in a similar way as Ecological Production Units, or in the context of the MSFD, these are akin to subregions within a regional sea (**Figure 2A**). We then developed the same 22 indicator time series for each of these units (31 units x 22 indicators = 682 total, **Table 1**). Each set of indicators is treated as an independent assessment (i.e., a mini EA). Conceptually, we follow the procedure shown in **Figure 1** but with real data, then explore the resulting time series in the context of our research questions.

2.3. Defining ecosystem subunit boundaries

We used the region covered by the NOAA bottom trawl survey offshore strata, that are regularly sampled, to define the total geographic extent of the NES for this study (**Figure 2A**). This survey covers the entire NES with 350 - 400 trawl stations sampled annually (Azarovitz 1981). Next the NES was split in half at the latitude which gave two polygons of exactly the same spatial area, resulting in a northern unit and a southern unit of ~ 120,000 km^2 each. Subsequent ecosystem splits were made within the northern and southern units by splitting each into 2, 3, 4, and 5 units using a more complicated procedure (**Figure 2A**, *supplementary materials*). In short, subsequent splits were done in an automated manner that maximized evenness in sampling density of the survey data used for indicator calculations. Spatial units are an average of 245,000 (n = 1), 123,000 (n = 2), 59,000 (n = 4), 39,000 (n = 6), 29,000 (n = 8), and 23,000 (n =10) km^2, though these numbers are rounded in figures and subsequent text for simplicity. These boundaries are naïve to the inherent biophysical structure present in the ecosystem, and existing governance structure. This was done to focus the results on broader issues related to scale sensitivity, rather than nuances of how this particular ecosystem is, or should be defined in practice.

2.4. Indicators

We used a total of 22 indicators to represent each spatial unit. Five described the physical environment, eight represented lower trophic level processes, and nine represented mid trophic level processes. We required indicators with long-time series known to be useful for describing the ecosystem, so we used indicators commonly appearing in federal EAs for the NES. All but two of our final 22 indicators (Spiny Dogfish and Silver Hake) have been used in previous federal EA reports and have been vetted and selected from much broader lists of potential indicators (Link et al. 2002; Methratta and Link 2006; EcoAP 2009, 2011; NEFSC 2018a, 2018b, 2019a, 2019b). Spiny dogfish and Silver Hake are abundant and important mid trophic level taxa that also provided interesting contrasts of spatial variation in trends (Nye et al. 2011). The methods used for all zooplankton indicators and those based on the NEFSC trawl survey (**Table 1**) are all identical to those from NEFSC (2019a) so are not described in detail in the main text (*supplementary materials*). All indicators end at 2018 or 2017 and have variable start dates (**Table 1**). The minimum length for a time series was 37 years and the maximum was 48.

2.3.1. Indicators using World Ocean Data

Indicators for *in situ* bottom temperature, surface temperature, stratification, and chlorophyll-a (in situ) were developed with data from the World Ocean Database, which includes contributions from various agencies (including all NMFS surveys) and researchers. There were 58,554 CTD casts available for temperature and stratification indicators, and 17,379 casts that contained chlorophyll-a measurements. Because these represent a broad collection of datasets (i.e., CTD casts are not random in space or time) we used generalized additive models (GAMs) to develop an annual index that accounted for spatial unevenness and the month the CTD casts were taken (**Table 1**, *supplementary materials*). The GAMs were fit using the "mgcv" package in the R statistical programming environment (Wood 2001; R Development Core Team 2019). We used the default thin plate regression spline option (bs = "tp") and restricted maximum likelihood to fit the models (method = "REML"). Models were first run including all data for the NES to assess fit and identify appropriate transformations before running for data subset to each spatial unit. The year coefficient from the model was used as the indicator (**Table 1**). Similar methods are regularly used to standardize fisheries sampling data and assess chlorophyll trends (Maunder and Punt 2004; Boyce et al. 2010).

2.3.2. Indicators using satellite data

We used the NOAA Optimum Interpolation ¼ degree Daily Sea Surface Temperature Analysis (OISST) to calculate an indicator of summer sea surface temperature (SST) and a count of days classified as Marine Heatwaves (MHWs) per year (Reynolds et al. 2007; Hobday et al. 2016). For the summer SST indicator, we averaged SST measures occurring within a spatial unit by year including only data from July to September (day-of-year 182 to 273). For the MHW indicator, we first built a single time series of average daily temperatures within a spatial unit. This was then used to detect marine heatwave events according to methods described in Hobday et al. (2016). We defined a MHW as an anomalously warm water event (> 90^{th} percentile based on a climatology from 1981 to 2012) that lasted for more than 5 days. The indicator represents a count of all days per year, that were classified as MHW days. An indicator of chlorophyll-a was developed using satellite data from the Ocean Colour - Climate Change Initiative dataset (4.0), a recent data product that merges data from several satellite missions (Sathyendranath et al. 2019). We averaged daily chlorophyll-a estimates, occurring within each spatial unit annually to produce the indicator.

2.4. Data analysis

The resulting 682 time series (31 spatial units x 22 indicators) were assessed for a monotonic trend with a Mann-Kendall non-parametric test (Mann 1945; Kendall 1957). A Bonferroni correction was included to account for multiple testing within each scale by dividing alpha (0.05) by the number of tests done. For example, for indicators done at the 20,000 km^2 scale (i.e., $1/10^{th}$ of the NES) we used an alpha of 0.05/10, for the 30,000 km^2 we used 0.05/8. Theil-sen slopes and intercepts were also calculated for each time series (Sen 1968; Theil 1992). All time series were independently scaled and centered (subtracting mean and dividing by standard deviation) prior to analysis to focus this analysis on trends, rather than absolute values.

Table 1. Indicators used in this study, all indicators run from the start year (shown) to 2018 unless otherwise noted. Abbreviations are as follows: GAM, generalized additive model; BT, bottom temperature; lat, latitude; lon, longitude; mon, month; yr, year; WOD, world ocean database; SST, sea surface temperature; Strat, stratification defined as the difference in water density between the surface and at 50 m; yday, day of year in numeric form; OISST, optimally interpolated sea surface temperature; Chyl, chlorophyll-a; const., a small constant number; OCCI, Ocean Colour Climate Change Initiative; EcoMon, Ecosystem Monitoring survey by the North East Fisheries Science Center (NEFSC). Indicators in bold are calculated with the exact methods described in NEFSC (2019a) and data using the NEFSC trawl survey are from the spring sampling period.

Category	Indicator	Description	Data	Start
Physical	Bottom temperature	log(BT+1) ~ s(lat, lon) + s(mon) + yr	WOD	1975
	Surface temperature	SST ~ s(lat, lon) + s(mon) + yr	-	1975
	Stratification	log(Strat) ~ s(lat, lon) + s(mon) + yr	-	1975
	Summer SST	Mean (yday 182 - 273)	OISST	1981
	Marine heatwave days	Extensive analysis	-	1981
Lower TL	Chlorophyll-a	log(Chyl + const.) ~ s(lat, lon) + s(mon) + yr	WOD	1980
	Chlorophyll-a (sat)	Annual mean Chyl	OCCI	1997
	Zooplankton volume	Annual volumetric anomaly	EcoMon	1977*
	Centropages typicus	Annual abundance anomaly	-	-
	Pseudocalanaus spp.	-	-	-
	Temora longicornis	-	-	-
	Calanus finmarchicus	-	-	-
	Small/large copepod	-	-	-
Mid TL	**Trawl biomass**	Biomass tow^{-1}	NEFSC	1970
	Benthivore	Biomass tow^{-1} 42 benthivore sp.	-	-
	Benthos	Biomass tow^{-1} 9 benthos sp.	-	-
	Piscivore	Biomass tow^{-1} 63 piscivore sp.	-	-
	Planktivore	Biomass tow^{-1} 20 planktivore sp.	-	-
	Sea Scallop	Biomass tow^{-1} Sea Scallop	-	-
	American Lobster	Biomass tow^{-1} American Lobster	-	-
	Silver Hake	Biomass tow^{-1} Silver Hake	-	-
	Spiny Dogfish	Biomass tow^{-1} Spiny Dogfish	-	-

*End year for indicators using EcoMon data is 2017.

To examine how the overall perception of trends change with the extent and region of an EA (*objective i*), we compared one simulated EA (i.e., all 22 trend results for a unit) to another EA that was nested within it. How well does a broad scale EA represent the same trends evaluated at a more local scale? Trends were classified as significantly increasing, decreasing, or not-trending based on the Mann-Kendall test, and then assessed for consistency. For example, if trawl biomass was increasing in the 250,000 km^2 unit but not trending in a nested spatial unit, this is a classified as a disagreement. We counted the number of disagreements for each comparison and report this as a % of the total comparisons. We also tallied disagreements according to their cause (**Figure 1B-D**). Masking is when a trend is *absent* at a broad extent but *present* within a nested unit; propagation is when a trend is *present* at a broad extent but *absent* within a nested unit; divergence is when a trend is *present* at a broad extent *and* a nested unit but the direction of trends differed.

In addition to visualizations, we calculated several metrics to assess which indicator trends were most sensitive to scale (*objective ii*). The first was the variance of the 31 Theil-Sen slopes for a given indicator. Low variance implies the monotonic trendline is consistent, regardless of the spatial unit considered. High variance means the slope of the trend changes depending on the spatial region considered. We also calculated the average pairwise Pearson correlation coefficient between all 31 units, for a given indicator (*Ave* r_p). High *Ave* r_p means the raw indicator time series data are consistent across space (i.e., spatially coherent), and this metric makes no assumption about the presence or absence of a trend (Östman et al. 2017). Lastly, for a subset of indicators (foraging groups) we graphically depict the relationship between indicator absolute values, indicator time series trend, and spatial extent of the EA (*Objective iii*).

3. Results

3.1. The degree of inconsistencies between ecosystem assessments at different scales

The number of significant indicator trends varied within each scale considered indicating that some parts of the ecosystem were changing more so than others (**Figure 2B**). Treating the ecosystem as a single unit (extent = 250,000 km^2) resulted in 12 out of 22 indicators with significant trends (55%); splitting it in two showed that the north had 13 (59%) significant trends whereas the south had only 10 (45%). Further disaggregation at progressively smaller spatial extents revealed anywhere from 5 to 12 significant trends. This means that one way of defining a region for an assessment would show only 23% of indicators are trending, yet a different way of defining it would show 59% of indicators are trending. Comparisons across scales demonstrate a pattern where more significant trends are detected at larger spatial extents (**Figure 2B**). Without the Bonferroni correction, this trend was weaker, but still visually present (**Figure 2C**).

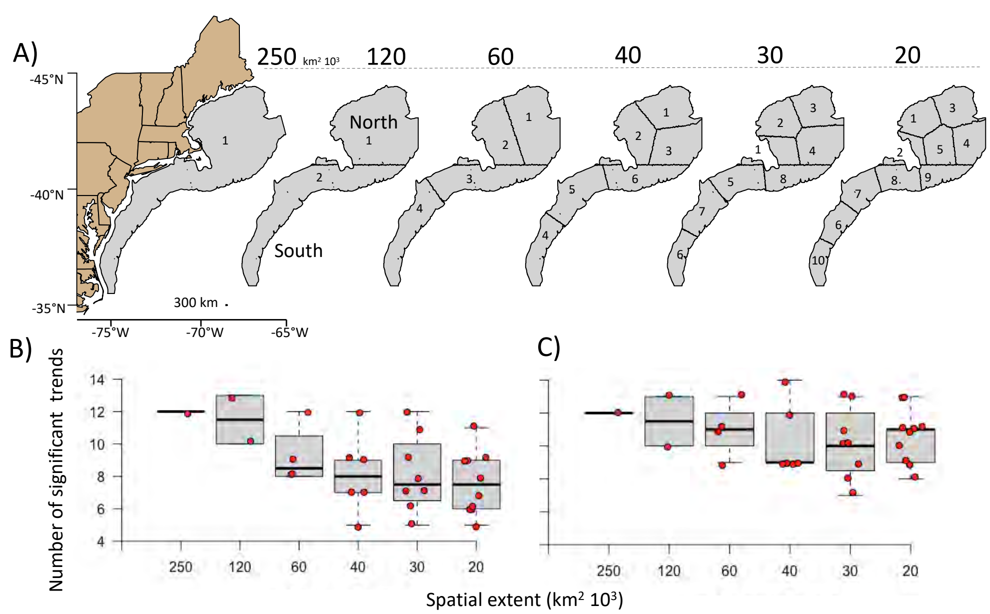


Figure 2. The Northeast US Atlantic Continental Shelf Large Marine Ecosystem, split into 31 units (**A**) with the average spatial extents of ecosystem units shown (above). Panel B shows the number of significant indicator trends (out of 22 tested) within each spatial unit assessed with a Mann-Kendall test and a Bonferroni correction. Each point corresponds to a unit in panel A. The same results using a consistent alpha of 0.05 is shown in panel C.

Comparisons of indicator trends among spatial units showed inconsistencies of between 4% and 36% (**Figure 3**; **Table 2**). These comparisons relate the trend of an indicator (increasing, decreasing, stable) calculated at a broad extent to the *same indicator* calculated at a smaller and nested scale (do results match?). For example, if we treat the NES as a single region and calculate indicators and then compare these to the indicators for the northern unit, 5 out of the 22 trend results were inconsistent (23%). The whole NES version at 250,000 km^2 was also 23% inconsistent with the southern unit. Comparing the 250,000 km^2 unit to the ten nested 20,000 km^2 units gave a mean inconsistency of 30% (range 23% to 36%), implying a single EA would misrepresent local trends by as much as 36%. Overall there appeared to be a trend where mean and maximum inconsistencies increased when the 250,000 km^2 was compared to units of progressively smaller spatial extent (**Figure 3A, B**), but this trend flattened once a scale of 40,000 km^2 was reached.

The southern 120,000 km^2 scale EA described nested local trends better than the northern version did, but there was great variation within the comparisons (**Figure 3**). Mean inconsistencies in the north were 23% for most scales, while they were closer to 15% for most comparisons in the south. Yet, the observed variation within a given spatial scale means that the 120,000 km^2 units reflected trends occurring in some nested spatial units quite well (**Figure 3C**) while they poorly matched those occurring in other nested units (**Figure 3B**). Like the comparisons with the 250,000 km^2 units, mean and maximum inconsistencies appeared to increase when comparisons were made with units of smaller spatial extent (**Figure 3A, B**).

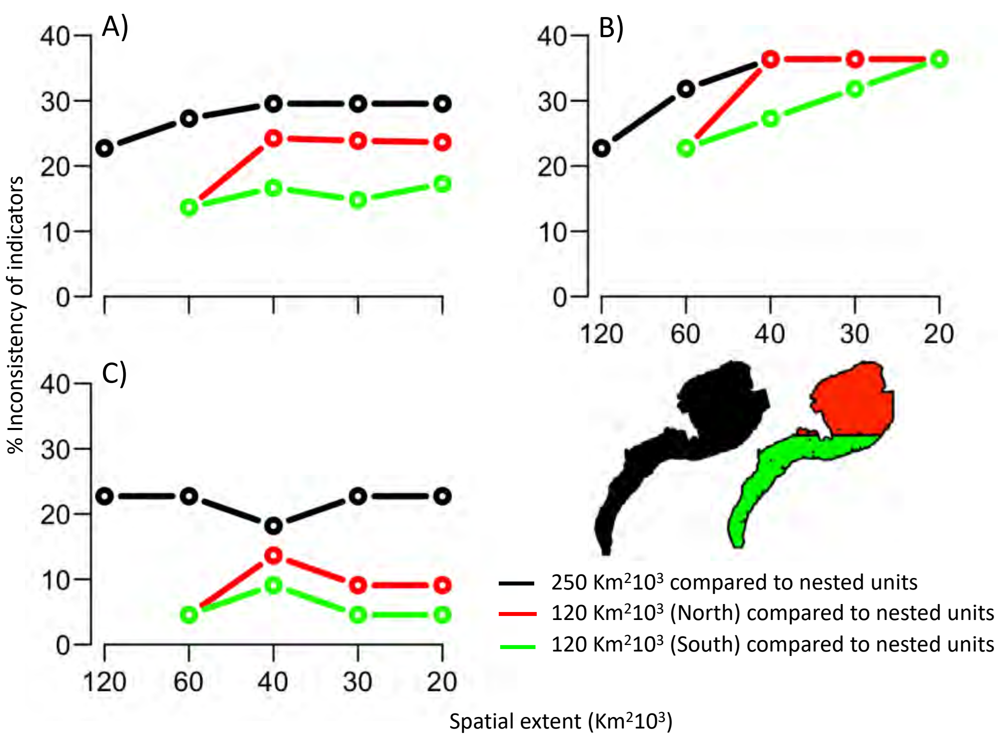


Figure 3. Percent inconsistency of indicator trends when comparing those calculated for one spatial unit (250,000 km^2, 120,000 km^2 north, 120,000 km^2 south; colored lines) to nested spatial units in terms of the mean inconsistency (A), maximum inconsistency (B), and minimum inconsistency (C). Percent inconsistency is defined as the number of time series trends that differed, divided by the total number compared (n = 22).

Table 2. Indicator consistency and the reason for inconsistencies. The comparison column shows which units are being compared. For example, 250:120 designates that indicator trends (n = 22) for the whole ecosystem version (250,000 km^2) were compared to the same indicator trends calculated for the 120,000 km^2 units (i.e., a northern and southern unit, **Figure 2**). Percent inconsistency is defined as the number of time series trends that differed, divided by the total number compared. Inconsistency (%) is followed by the minimum and maximum observed in a comparison. Total inconsistent (n) is followed by the total number of comparisons made. The proportion of total inconsistencies due to three causes as depicted in **Figure 1** is also shown.

Comparison	Inconsistency (%)	Total inconsistent (n)	Propagation (%)	Masking (%)	Divergence (%)
250:120	23 (23 - 23)	10/44	40	40	20
250:60	27 (23 - 32)	24/88	63	29	8
250:40	30 (18 - 36)	39/132	72	23	5
260:30	30 (23 - 36)	52/176	71	21	8
250:20	30 (23 - 36)	65/220	75	18	6
120(N):60	14 (5 - 23)	6/44	100	0	0
120(S):60	14 (5 - 23)	6/44	67	33	0
120(N):40	24 (14 - 36)	16/66	94	6	0
120(S):40	17 (9 - 27)	11/66	73	27	0
120(N):30	24 (9 - 36)	21/88	95	5	0
120(S):30	15 (5 - 32)	13/88	77	23	0
120 (N):20	24 (9 - 36)	26/110	100	0	0
120(S):20	17 (5 - 36)	19/110	79	21	0

3.2. Common causes for cross-scale inconsistencies

Local trend propagation (**Figure 1C**) was the dominant cause for cross-scale indicator inconsistencies, accounting for between 40% and 100% of them relative to the total for each comparison (**Table 2**). For example, trawl biomass increased when considered at the whole ecosystem scale but only increased in one of the two nested units at the 120,000 km^2 scale (the south) (**Figure 4C**). The strong trend in the southern unit thus 'propagated' to the next spatial extent, creating an inconsistent result between the 250,000 km^2 unit (increasing) and the northern unit (no trend). Further disaggregation of the southern unit revealed this increasing trend was not widespread (increasing in only 2/10 units, **Figure 4C**). In this case, 8 units exhibit no trend and are thus inconsistent with the increasing trend portrayed by the 250,000 km^2 unit. Other examples of local trend propagation in **Figure 4** include *Temora longicornis* abundance (increased in northern unit but not in some nested subunits) and piscivore biomass (increased in southern unit but not in some nested subunits).

Local trend masking (**Figure 1B**) accounted for between 0% and 40% of inconsistencies (**Table 2**) and was the second most common cause of cross-scale inconsistencies. The indicator of *Temora longicornis* abundance (a copepod) presents two examples. First, there was no trend detected at whole ecosystem scale, but there was an increasing trend occurring in the spatially nested northern unit (**Figure 4B**). This trend failed to propagate to the next spatial extent, and thus was masked by aggregation of data across a large region where no trend was occurring. Additionally, an increasing trend was detected at the southernmost unit in the 60,000 km^2 scale, and since this trend also did not propagate to units of larger extents, this created inconsistencies.

Other notable examples of masking include Silver Hake (**Figure 4F**), *Pseudocalanus spp.* abundance (**Figure S5**), and benthivore biomass (**Figure S6**).

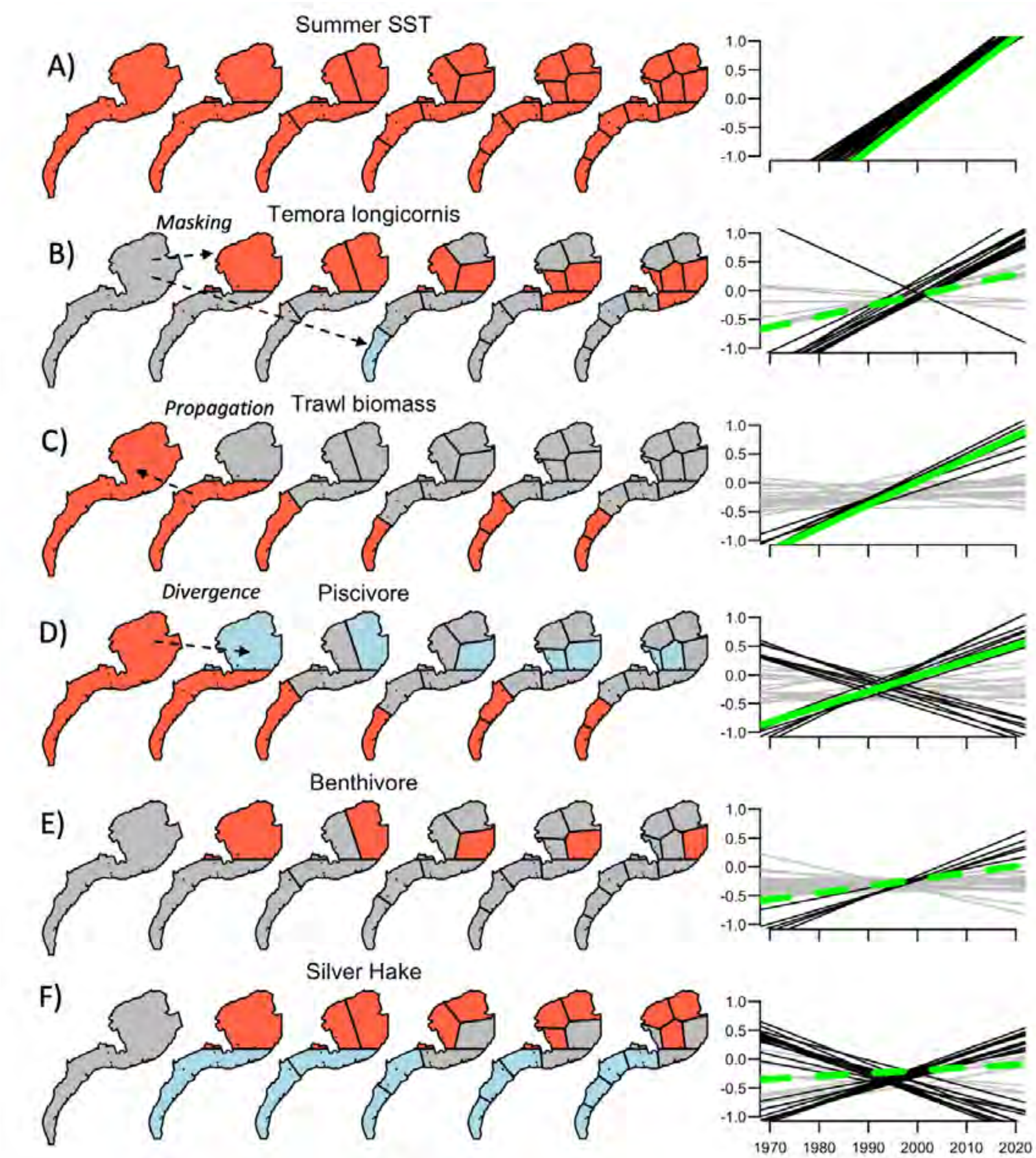


Figure 4. Perceived trends in five indicators (A - E) based on 31 different ways of dividing the NES into units for an ecosystem status report at six spatial scales. The leftmost panel is the trend that would be perceived if the ecosystem as treated as a single unit (245,000 km^2), other panels in a row show progressive splitting, recalculation of the indicator, and re-analysis of the time series. Units are coded based on results of Mann-Kendall test for a monotonic trend on the indicator time series (red = increasing, blue = decreasing, grey = no trend). The farthest right panel shows the linear fits to the time series (31 lines appear in each panel, one corresponding to each spatial unit). Significant trends are shown in black, non-significant trends are in grey, and the trend for the whole ecosystem version (245,000 km^2) is green with a solid line if significant and a dashed line if not.

Local trend divergence (**Figure 1D**) accounted for between 6% and 20% of inconsistencies when comparing the 250,000 km^2 scale indicators to ones for nested units but did not occur at all when comparing northern and southern units to ones nested within them (**Table 2**). One example is piscivore biomass, which would be perceived as increasing from the whole ecosystem version (250,000 km^2) but after disaggregation, we find the northern unit is actually decreasing (**Figure 4D**). Similarly, the whole ecosystem indicator gives a perception of an increasing trend in American Lobster biomass, while the southern unit portrays a strongly decreasing trend (**Figure S7**).

3.3. Scale sensitivity of indicators and spatial coherence (Ave r^p)

We used two quantitative metrics to describe scale sensitivity of specific indicators and found that *Ave* r_p provided a more intuitive and useful description than Theil-Sen slope variability (**Table 3**). The two metrics were not correlated ($p = 0.46$, $r = -0.16$) which means that indicators with high spatial coherence of the raw time series (represented by *Ave* r_p) do not necessarily have more consistent monotonic trends. This is not intuitive, but we note that the Theil-Sen slope variability included 31 slope estimates (for each indicator), that were in many cases representing non-significant trends. Many slope estimates, for time series without a trend, lead to erratic slope estimates and thus high Theil-Sen slope variation. This occurred even when the time series, as determined by *Ave* r_p and visual inspection of the data, were quite coherent across space. We therefore favored *Ave* r_p to describe which indicators were most sensitive to scale but report both (**Table 3**).

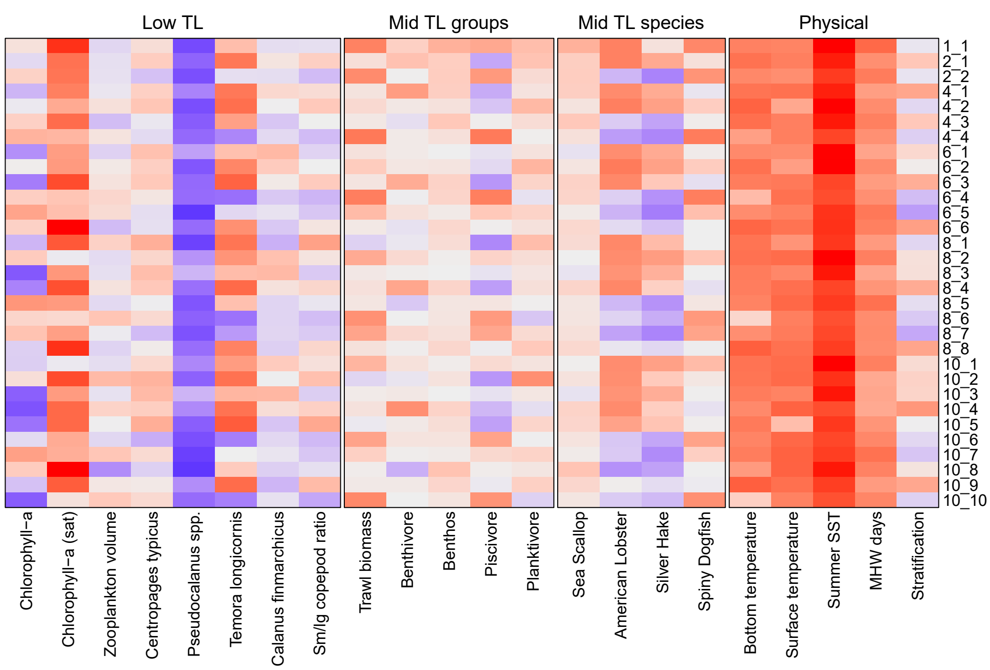


Figure 5. Variation in Theil-Sen slopes of indicator time series (columns) for 31 different spatial units (rows, see **Figure 2**). Slope values are depicted as shaded from dark red (strong increase) to dark blue (strong decrease). The TL used in headers refers to trophic level.

In general, scale sensitivity appeared to follow a pattern where physical indicators of temperature were not sensitive, lower trophic level indicators were intermediately sensitive, and mid trophic level indicators were most sensitive (**Table 3**). No matter how the ecosystem was subdivided, temperature indicators were increasing (i.e., the large block of red on the right, **Figure 5**), among spatial unit variability of Theil-Sen slopes was low, and *Ave* r_p was high (**Table 3**). All Indicators of chlorophyll-a had high spatial coherence (0.71 - 0.76), indicators for zooplankton had intermediate coherence (0.51 - 0.66), and those for trawl survey indicators (Mid-TL) had the lowest spatial coherence (0.18 - 0.54, **Table 3**). An outlier to the apparent pattern of physical indicators being more coherent was that water column stratification was not spatially coherent, having an *Ave* r_p of 0.27.

Table 3. Indicators and metrics of their behavior including *Ave* r_p (a measure of spatial coherence), Slope variability (variability of Theil-Sen slopes fit to time series for the 31 units), increasing (number of significant increasing monotonic trends), and decreasing (number of decreasing monotonic trends). Rows are sorted according to *Ave* r_p within each category and color scales applied to highlight variation within columns.

Category	Indicator	Ave r_p	Slope variability	Increasing	Decreasing
Physical	Summer SST	0.81	2.57	31	0
	MHW days	0.8	3.47	30	0
	Surface temperature	0.67	5.09	21	0
	Bottom temperature	0.66	10.02	23	0
	Stratification	0.27	25.51	0	0
Lower TL (chyl-a)	Chlorophyll	0.76	59.08	0	2
	Chlorophyll (sat)	0.71	42.61	0	0
Lower TL (zooplankton)	Sm/lg copepod ratio	0.66	19.02	0	0
	Zooplankton volume	0.52	11.27	0	0
	Pseudocalanus spp.	0.71	10.38	0	24
	Centropages typicus	0.61	13.74	0	0
	Temora longicornis	0.55	92.59	12	1
	Calanus finmarchicus	0.51	12.76	0	0
Mid TL (groups)	Benthivore	0.54	13.52	5	0
	Trawl biomass	0.3	24.00	8	0
	Benthos	0.28	3.01	14	0
	Planktivore	0.24	9.14	7	0
	Piscivore	0.18	47.55	8	7
Mid TL (species)	Silver Hake	0.3	58.52	11	12
	American Lobster	0.26	69.20	16	7
	Sea Scallop	0.26	3.66	12	0
	Spiny Dogfish	0.2	23.65	11	0

3.6. Patterns between spatial extent, trend magnitude, and absolute indicator values

We found a greater range of trend magnitudes and absolute values of indicators when smaller and smaller subunits were used (**Figure 6**). In simple terms this means that localized 'hotspots' of trend or abundance were portrayed in full strength when indicators were calculated for a small area, but when indicators were generated for bigger regions these ‘hotspots’ were subdued by inclusion of areas with no trend (or contrasting trend) and lower abundance. The same could be said for ‘cold spots’ (e.g., areas of notably decreasing trend or low abundance).

A notable example is the pattern depicted by Benthivore biomass, which is increasing significantly in only one of ten 20,000 $km^2 10^3$ units (in the North East, in what is called the Scotian Shelf) and lead to widespread trend propagation (**Figure 4E).** When trends are compared to spatial extent (**Figure 6**) there is *always a single unit* that has a trend far higher than others. Any unit that by chance happened to include this local area of increasing abundance ended up being driven by this local trend. The mean of all trends (dashed line) is quite close to the trend that would be depicted by a single indicator for the whole region. Similarly, the mean values of indicators (kg tow^{-1}) show greater variation with decreasing spatial extent. For example, some units had a high density of benthivores (max 95 kg tow^{-1}) whereas other units have very few benthivores (min = 7 kg tow^{-1}).

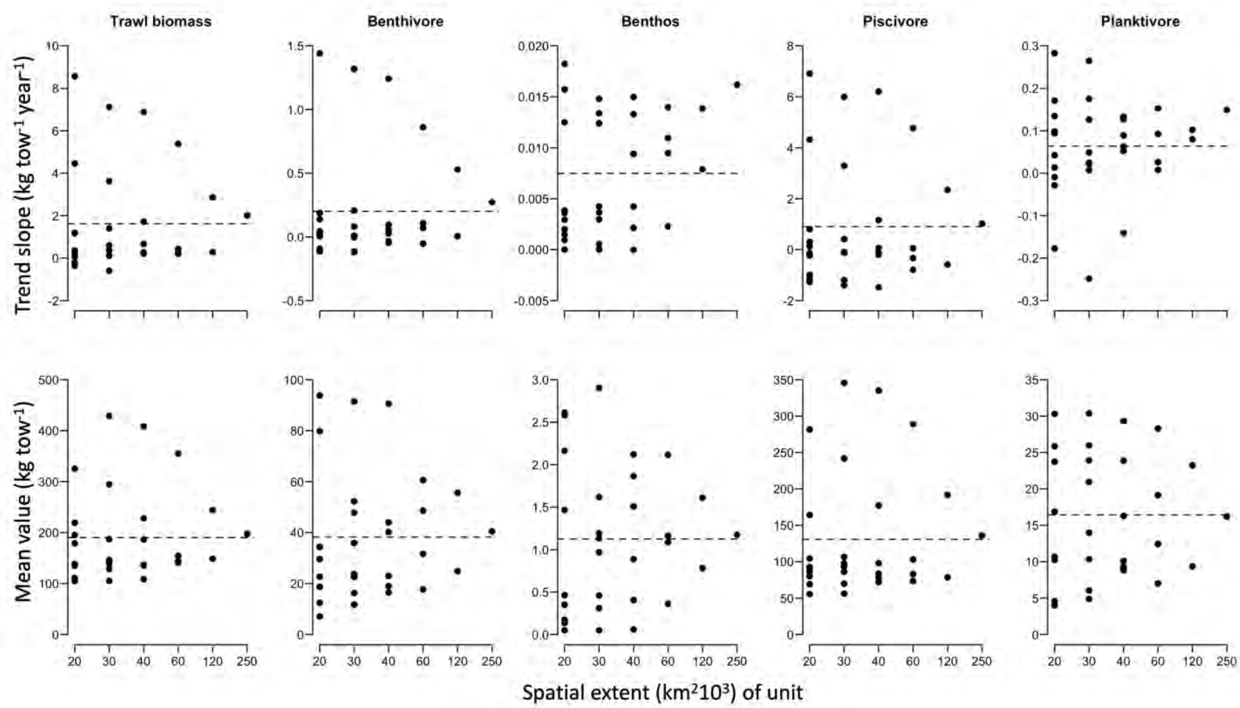


Figure 6. Trends (Theil-Sen slopes) and mean indicator values (1970 - 2018) of aggregate species groups examined at multiple spatial scales. Points correspond to each spatial unit in **Figure 2A**.

4. Discussion

We provide the first evaluation to our knowledge of how boundaries and ecosystem subunit delineation can influence an entire EA and found the perception of the ecosystem, in terms of 22 indicators, changed by as much as 38% as a function of scale. This result emphasizes the importance of a preliminary step in developing *any* ecological or ecosystem indicator: Defining spatial scale and boundaries. Since the simple question “is there a trend?’ is so strongly influenced by this step, which should be made early in the indicator development process, we expect many subsequent uses of indicators are also sensitive to scale. Thus, the scale that indicators are developed should be chosen carefully, and the consequences of different boundary systems explored to determine how they might influence indicator behavior and uses in management.

4.1. Degree and common causes for inconsistencies

When preparing an EA or assessing good environmental status under the MSFD, there are endless options for defining spatial boundaries and subunits (Queirós et al. 2016; Otto et al. 2018), and our results demonstrate that the specifics of these boundaries can markedly change the overall perception of the ecosystem. We found that the detected signals of ecosystem change depend on the way the ecosystem is defined. At the broadest spatial extent, the NES was treated as a single spatial unit to develop indicators, which is common in some research articles (Large et al. 2013) and EAs in the region (Fogarty et al. 2012). When we compared this whole ecosystem version of the 22 indicators to those for northern and southern subunits, results were 23% inconsistent (in both comparisons). Assuming the north and south units (each 120,000 km^2) represent regional dynamics well, we think of this as a 23% erosion of a potential signal due to averaging spatial variation. When the whole ecosystem version of indicators was compared to those for progressively smaller subunits, inconsistencies increased and reached a maximum of 36%. The main message is that a single EA will provide a big picture view of an ecosystem but can mis-represent local dynamics. The details of these inconsistencies (i.e. **Figure 1B-D**) could have different implications for management.

The most common case for cross-scale inconsistencies was 'local trend propagation': Change is occurring, but not across the entire assessment region (sensu **Figure 1C**). One management implication of local trend propagation is that a prescribed treatment for a problem is applied to an unnecessarily broad region. For example, if a broad scale indicator (e.g., the whole NES) portrays excessive fisheries bycatch, widespread fisheries closures could cause unnecessary economic losses if the bycatch was actually driven by a localized area or fishery gear type. This issue (and others) underly criticisms of static spatial management that promising tools, like dynamic ocean management, seek to resolve (Maxwell et al. 2015; Dunn et al. 2016; Hazen et al. 2018). In another scenario, localized trends in biodiversity, a notoriously scale-dependent and patchy ecosystem property (Peterson et al. 1998), could drive an undesirable trend. Under the MSFD, such an instance could produce low scores for GES for the biodiversity descriptor if spatial variation was not adequately recognized (Uusitalo et al. 2016).

The next most common cause for inconsistencies was local trend masking (sensu **Figure 1B**). The critical point in these examples is that because of the chosen boundaries used, we fail to detect a trend. A well understood example is given by Silver Hake, which is managed as two discrete stocks and is experiencing a poleward range shift (Nye et al. 2009, 2011). Because of these regionally divergent trends, we would not perceive a trend when aggregating to the whole NES when in fact important changes are happening (Link et al. 2011). Other instances of local trend masking are likely common in practice across the breadth of indicator applications. Fishing pressure may appear stable even as local fishing pressure intensifies (Preciado et al. 2019), broad scale indicators of pollution or eutrophication could appear stable as point sources vary (Gergel 2005), or cultural service indicators could be insensitive to temporal change if societal interest shifted spatially within the study extent (Richards and Friess 2015). In the NES few of our specific examples of masking are surprising because of an extensive body of research. Yet, this issue could be more problematic in poorly understood ecosystems where such a background knowledge is lacking.

The least common, but perhaps most consequential issue was when a broad scale indicator displayed a significant trend, but a nested spatial region showed a significant trend in the opposite direction (sensu **Figure 1D**). This is exemplified by the American Lobster example. Again, we know that this species is experiencing a dramatic increase in the north and a decrease in the south of the NES (Steneck and Wahle 2013). Our results demonstrate how this northern trend overwhelms the declines in the south, leading to a perception of increasing trends when aggregated to the whole NES. In terms of decision making, this issue could be especially problematic. Imagine a very simple proposed management action based on a trigger (if indicator X is increasing, do A: if decreasing, do B); we might be doing A for a large region where locally the right approach is actually B. Although this was not a common finding in our case study, it could be a common pattern in other systems or with other indicators.

4.2. What indicator trends are most sensitive to scale?

Our results suggest a pattern where temperature indicators trends were the least sensitive to scale, and then indicators of the marine community varied predictably along trophic levels (i.e., phytoplankton, zooplankton, fish and invertebrates). These results are largely consistent with theory and field-specific research. For example, climatic variation and global warming operate across relatively broad spatial scales (Belkin 2009). Although there is variability in rates of warming in the NES, as well as marine heatwave frequency and duration (Pershing et al. 2015; Oliver et al. 2018), nearly any way of dividing the NES will still reveal an increasing trend. In our study, temperature and heatwave indicators displayed the highest spatial coherence in the time series (Ave r_p 0.66 - 0.81). Patterns in chlorophyll were similarly insensitive to scale. Although we detected no long-term trends in chlorophyll, both the satellite derived and in-situ based indicator time series were highly coherent across spatial units (Ave r_p = 0.76, 0.71). Most zooplankton indicators were quite coherent across spatial subunits, but we did detect some divergent trends and lower coherence in some indicators (Ave r_p as low as 0.51 for *Calanus finmarchicus*). This is consistent with recent studies demonstrating regional variation in zooplankton dynamics in the NES within the four Ecological Production Units defined by NMFS (Morse et al. 2017). Lastly, mid-trophic level indicators (fish and invertebrates) displayed the lowest spatial coherence and highest variation in trends, consistent with Östman et al. (2017) that found strong spatial coherence in physical pressures, but far less for indicators of the coastal fish community in the Baltic Sea. Whether this apparent pattern in indicator scale sensitivity is supported by other indicators, in other ecosystems, is a topic worth further investigation. In any case, this highlights again a fundamental challenge when assessing ecosystems; a spatial scale that is effective for capturing spatial variation in one indicator might be inappropriate for another.

4.3. A generalized relationship between scale, indicator trends, and values

The pattern of heteroscedasticity observed when comparing indicator trends and mean values to spatial extent (e.g., triangular pattern in **Figure 6**) is noteworthy and consistent with expectations from scaling theory. The six spatial breaks used in our study are akin to changing grain with extent held constant and Weins' (1989) predictions are applicable. When data are compiled for the entire NES to produce a single annual mean, all spatial variation is lost to averaging (e.g., maximal within subunit variance, **Figure 7**) and this produced intermediate trends and absolute values of indicators. At the smallest grain, 'hotspots' of local abundance and variability in temporal trends emerge (e.g., maximal between subunit variance, **Figure 7**). These patterns arose by chance in our analysis, since subunits delineation was naïve to any spatial structure in the data. Had we based subunits on the spatial structure of indicator datasets, this variation may have been even more evident. For example, had we developed boundaries that matched distinct ecoregions (Lucey and Fogarty 2013) or followed bathymetric features (e.g., on the continental shelf vs. shelf break), we would be maximizing within subunit similarity and exemplifying between subunit variation (**Figure 7**).

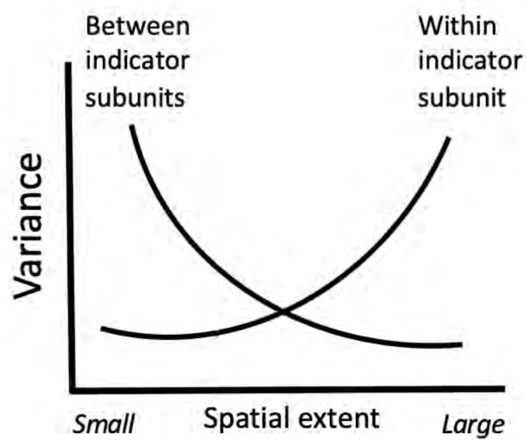


Figure 7. Weins (1989) depiction of how variance changes with scale applied to indicators. Consider an indicator calculated for a single large spatial unit (large extent) versus the same indicator calculated for each of multiple smaller subunits (small extent). At a large extent, subunit values (or trends of indicator time series) are more similar to one another (low variance), but within subunit spatial variance is high and lost to averaging across a large region. As subunit extent is decreased, variance between subunits increases and within subunit spatial variance decreases.

This consistent result suggests that a variety of spatial statistics will be useful to more deeply explore scale sensitivity and characteristic spatial scales of various ecosystem indicators. It should be noted that our work is not directly a spatial analysis; we repeatedly developed a simple EA using permutations of different boundary systems and conducted trend analysis in each case independently. This process matched how EAs are often conducted in practice (i.e., develop boundaries, calculate indicators, analyze indicators) to explore realistic issues that could arise, yet application of spatial tools could be quite useful. For instance, semivariograms, a tool to explore the distance and strength of autocorrelation (Sokal and Oden 1978) could be applied to either (1) raw survey data values or (2) mean values of indicators developed at different spatial extents to inform appropriate indicator scales. While usually applied to measured data values (e.g., biomass of a trawl sample, temperature at a point; e.g., Ciannnelli et al. 2008) these could also be applied to spatially referenced ecosystem trends. This would address the question; at what spatial scales do indicator *trends* display autocorrelation, rather than the scales at which indicator *values* display autocorrelation. Exploring these two questions together, for a suite of indicators, would be highly informative. The relevant point to glean from our results is that if there is high spatial variation in an underlying process, broad scale indicators will produce subdued trends and values relative to indicators developed at smaller spatial scales. Whether capturing this variation is necessary will depend on the intended purpose and objectives for developing indicators in the first place.

4.4. Implications for marine ecosystem assessments and management

We have stepped back to a preliminary step (i.e., done before any subsequent analysis or decisions can be made *using* indicators) and found that different choices related to spatial scale and ecosystem boundaries can dramatically alter the perception of trends. Had we carried this analysis through subsequent steps in a management framework, we would undoubtedly have found that they too were sensitive to scale. Aside from broader awareness of this potential issue when using and interpreting indicators, we provide several recommendations to better incorporate spatial variation into indicator-based assessments.

First, spatial analysis should be done to inform the boundary systems used for indicator development and to compliment the simple, but integrative perspective provided by whole ecosystem assessments. By using an arbitrary boundary system, we showed that boundaries are quite important, and thus it is wise to delineate ecosystems and subunits based on the spatial structure of the ecosystem. For example, in the NES a spatial clustering analysis of physiographic, oceanographic, and primary production datasets was conducted by NMFS to define the aforementioned ecological production units (Lucey and Fogarty 2013). Although NMFS recognizes limitations of these static boundaries (e.g., dependence on variables used, jurisdictional issues, and openness of the system, Gamble et al. 2016) they are ecologically meaningful and have proven useful for structured Ecosystem Assessments. Also, while these integrated assessments may not portray spatial variation in all indicators, they are informed by and complimented with rigorous spatial analysis of survey datasets. Supporting mechanistic analyses help identify spatial variation that could lead to otherwise undetected instances of propagation, masking, and divergence (sensu Figure 1, Nye et al. 2011; Link et al. 2011; Kleisner et al. 2017). Moreover, datasets are publicly available and so can be evaluated (as we have done) by other scientists in the context of more topic-specific research.

Our results also suggest scale considerations should be included in indicator selection routines, which often consider responsiveness to a known pressure as a desirable feature (Otto et al. 2018; Shin et al. 2018). We expect relationships between indicators and pressures are also scale dependent. For example, the strength of predator-prey relationships or the negative impacts of bottom trawling on food webs, can weaken or strengthen depending on the spatial scale considered (Frank and Leggett 1981; Levin 1992; Preciado et al. 2019). Thus, if a pressure operates and influences a process of interest at fine spatial scales, then indicators would be most responsive to this pressure if assessed at finer resolution (Preciado et al. 2019). An indicator may be judged as poor for myriad reasons, but one reason is simply that it was not developed at a characteristic spatial scale. While not considered explicitly here, we recommend that the influence of spatial scale in 'indicator vetting' routines be examined in more detail.

For ecosystem indicators to be useful, they should be developed at scales matching those needed to make management decisions as well as those matching ecological processes. In most cases, goals and objectives should be defined for an EBM or EBFM framework and these should guide the scales at which indicators are developed. This could involve the same organization developing indicators a variety of scales, dependent on a weighted consideration of (1) the scales that specific ecosystem objectives and management needs and (2) the scale that the underlying processes operate at. For example, to provide ecosystem context for fisheries managers in a single-stock framework, NMFS annually re-calculates a suite of indicators (taken from the EAs) for regions matching the stock boundaries. Thus, indicators already developed in the EA process are spatially customized to match stock structure of managed species in the NES, that do not always not match Ecological Production Units used for EAs.

Yet, since the governance structure in many ecosystems is multiscale (i.e., federal, state, local user groups), there is also value in developing nested assessments of a large marine ecosystems to facilitate decision making by different groups (Steneck and Wilson 2010). A state or local government may have a hard time making meaningful decisions when using indicators generated by federal decision makers that cover much broader regions (and are developed to address broader scale objectives). Likewise, federal agencies will find excessive subunits at small spatial extents to be unwieldy or uninformative relative to their own objectives. Thus, the idea that indicators are sensitive to scale (i.e., tell a different story, depending on the boundaries used) is not necessarily a problem, but an incentive to learn about ecosystems more deeply. Such a framework of 'nested' EA reports is also demonstrated in the NES by two regional reports by NMFS (Mid-Atlantic and Northeast), and several local assessments of connected systems (e.g., Long Island Sound Study, LISCCMP 2015; and the Chesapeake Bay Program Hershner et al. 2007). Collectively, these can support integrated multiscale governance of ecosystems, but will require coordination across different levels of government (Steneck and Wilson 2010) to avoid already complicated issues related to sector-based management (Crowder et al. 2006).

4.5. Conclusions

Indicators are sensitive to spatial scale, but some indicators are more sensitive than others. The implications of this fact will vary depending on the intended use of the indicator and the indicator suite. Deeper exploration of indicator trends, at multiple scales, allows for the recognition of masking, propagation, and asynchrony that may be occurring within a spatial domain of interest. That these cross-scale inconsistencies occur is not a problem in itself, but they are important to recognize when making decisions regarding the management of ecosystems.

5. Acknowledgements

We thank the NEFSC Ecosystem Dynamics and Assessment branch for making available computer code and data to reproduce indicators appearing in their reports. We acknowledge the efforts of all those at the NEFSC who have contributed to the planning and execution of the bottom trawl, plankton, scallop and many other surveys. We thank the NEFSC staff for their dedicated work in auditing and maintaining these databases, and thoughtful feedback on the ideas and results presented here. This work was funded by the NYS DEC MOU # 10560

6. References

Azarovitz, T. R. 1981. A brief historical review of the Woods Hole laboratory trawl survey time series. – Canadian Special Publication of Fisheries and Aquatic Sciences. 58: 62–67.

Bal, P., A. I. T. Tulloch, P. F. E. Addison, E. McDonald-Madden, and J. R. Rhodes. 2018. Selecting indicator species for biodiversity management. Frontiers in Ecology and the Environment 16(10):589–598.

Belkin, I. M. 2009. Rapid warming of Large Marine Ecosystems. Progress in Oceanography 81(1–4):207–213. Elsevier Ltd.

Borja, A., M. Elliott, J. H. Andersen, T. Berg, J. Carstensen, B. S. Halpern, A.-S. Heiskanen, S. Korpinen, J. S. S. Lowndes, G. Martin, and N. Rodriguez-Ezpeleta. 2016. Overview of Integrative Assessment of Marine Systems: The Ecosystem Approach in Practice. Frontiers in Marine Science 3(March):1–20.

Boyce, D. G., M. R. Lewis, and B. Worm. 2010. Global phytoplankton decline over the past century. Nature 466(7306):591–596. Nature Publishing Group.

Bradbury, I. R., S. Hubert, B. Higgins, S. Bowman, T. Borza, I. G. Paterson, P. V. R. Snelgrove, C. J. Morris, R. S. Gregory, D. Hardie, J. A. Hutchings, D. E. Ruzzante, C. T. Taggart, and P. Bentzen. 2013. Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. Evolutionary Applications 6(3):450–461.

Bunnefeld, N., E. Hoshino, and E. J. Milner-Gulland. 2011. Management strategy evaluation: A powerful tool for conservation? Trends in Ecology and Evolution 26(9):441–447.

Burthe, S. J., P. A. Henrys, E. B. Mackay, B. M. Spears, R. Campbell, L. Carvalho, B. Dudley, I. D. M. Gunn, D. G. Johns, S. C. Maberly, L. May, M. A. Newell, S. Wanless, I. J. Winfield, S. J. Thackeray, and F. Daunt. 2016. Do early warning indicators consistently predict nonlinear change in long-term ecological data? Journal of Applied Ecology 53(3):666–676.

Ciannelli, L., P. Fauchald, K. S. Chan, V. N. Agostini, and G. E. Dingsør. 2008. Spatial fisheries ecology: Recent progress and future prospects. Journal of Marine Systems 71(3–4):223–236.

Colburn, L. L., M. Jepson, C. Weng, T. Seara, J. Weiss, and J. A. Hare. 2016. Indicators of climate change and social vulnerability in fishing dependent communities along the Eastern and Gulf Coasts of the United States. Marine Policy 74:323–333. Elsevier.

Crowder, L. B., G. Osherenko, O. R. Young, S. Airamé, E. A. Norse, N. Baron, J. C. Day, F. Douvere, C. N. Ehler, B. S. Halpern, S. J. Langdon, K. L. McLeod, J. C. Ogden, R. E. Peach, A. A. Rosenberg, and J. A. Wilson. 2006. Resolving mismatches in U.S. ocean governance. Science 313(5787):617–618.

DePiper, G. S., S. K. Gaichas, S. M. Lucey, P. P. Da Silva, M. R. Anderson, H. Breeze, A. Bundy, P. M. Clay, G. Fay, R. J. Gamble, R. S. Gregory, P. S. Fratantoni, C. L. Johnson, M. Koen-Alonso, K. M. Kleisner, J. Olson, C. T. Perretti, P. Pepin, F. Phelan, V. S. Saba, L. A. Smith, J. C. Tam, N. D. Templeman, and R. P. Wildermuth. 2017. Operationalizing integrated ecosystem assessments within a multidisciplinary team: Lessons learned from a worked example. ICES Journal of Marine Science 74(8):2076–2086.

Dunn, D. C., S. M. Maxwell, A. M. Boustany, and P. N. Halpin. 2016. Dynamic ocean management increases the efficiency and efficacy of fisheries management. Proceedings of the National Academy of Sciences of the United States of America 113(3):668–673.

EcoAP (Ecosystem Assessment Program). 2009. Ecosystem Status Report For the Northeast U.S. Continental Shelf Large Marine Ecosystem. Northeast Fisheries Science Center Reference Document, 09-11: 34.

EcoAP (Ecosystem Assessment Program). 2011. Ecosystem Status Report For the Northeast

Shelf Large Marine Ecosystem-2011. Northeast Fisheries Science Center Reference Document, 12-07: 38.

Foran, C. M., J. S. Link, W. S. Patrick, L. Sharpe, M. D. Wood, and I. Linkov. 2016. Relating mandates in the united states for managing the ocean to ecosystem goods and services demonstrates broad but varied coverage. Frontiers in Marine Science 3(FEB).

Frank, K. T., and W. C. Leggett. 1981. Wind Regulation of Emergence Times and Early Larval Survival in Capelin (Mallotus villosus). Can. J. Fish. Aquat. Sci. 38(m):215–223.

Fraser, W. R., W. Z. Trivelpiece, D. G. Ainley, and S. G. Trivelpiece. 1992. Increases in Antarctic penguin populations: reduced competition with whales or a loss of sea ice due to environmental warming? Polar Biology 11:525–531.

Gamble R., Fogarty M., Keith, C. 2016. Ecological Production Units for the Northeast U.S. Continental Shelf. NEFSC Ecosystem and Climate Science Review, Woods Hole, MA. Accessed at https://www.integratedecosystemassessment.noaa.gov/sites/default/files/pdf/ne-ecological-production-units-paper.pdf

Gaichas, S. K., G. S. DePiper, R. J. Seagraves, B. W. Muffley, M. G. Sabo, L. L. Colburn, and A. J. Loftus. 2018. Implementing Ecosystem Approaches to Fishery Management: Risk Assessment in the US Mid-Atlantic. Frontiers in Marine Science 5(November):1–25.

Gergel, S. E. 2005. Spatial and non-spatial factors: When do they affect landscape indicators of watershed loading? Landscape Ecology 20(2):177–189.

Harvey, C. J., C. R. Kelble, and F. B. Schwing. 2017. Implementing “the IEA”: Using integrated ecosystem assessment frameworks, programs, and applications in support of operationalizing ecosystem-based management. ICES Journal of Marine Science 74(1):398–405.

Hazen, E. L., K. L. Scales, S. M. Maxwell, D. K. Briscoe, H. Welch, S. J. Bograd, H. Bailey, S. R. Benson, T. Eguchi, H. Dewar, S. Kohin, D. P. Costa, L. B. Crowder, and R. L. Lewison. 2018. A dynamic ocean management tool to reduce bycatch and support sustainable fisheries. Science Advances 4(5):1–8.

Hershner, C., K. Havens, D. M. Bilkovic, and D. Wardrop. 2007. Assessment of Chesapeake Bay Program selection and use of indicators. EcoHealth 4(2):187–193.

Hobday, A. J., L. V. Alexander, S. E. Perkins, D. A. Smale, S. C. Straub, E. C. J. Oliver, J. A. Benthuysen, M. T. Burrows, M. G. Donat, M. Feng, N. J. Holbrook, P. J. Moore, H. A. Scannell, A. Sen Gupta, and T. Wernberg. 2016. A hierarchical approach to defining marine heatwaves. Progress in Oceanography 141:227–238. Elsevier Ltd.

Kendall, M. G. 1957. Rank Correlation Methods. Biometrika.

Kleisner, K. M., M. J. Fogarty, S. McGee, J. A. Hare, S. Moret, C. T. Perretti, and V. S. Saba. 2017. Marine species distribution shifts on the U.S. Northeast Continental Shelf under continued ocean warming. Progress in Oceanography 153:24–36. Elsevier Ltd.

Large, S. I., G. Fay, K. D. Friedland, and J. S. Link. 2013. Defining trends and thresholds in responses of ecological indicators to fishing and environmental pressures. ICES Journal of Marine Science 70(4):755–767.

Large, S. I., G. Fay, K. D. Friedland, and J. S. Link. 2015. Quantifying patterns of change in marine ecosystem response to multiple pressures. PLoS ONE 10(3):1–15.

Levin, P. S., C. R. Kelble, R. L. Shuford, C. Ainsworth, Y. DeReynier, R. Dunsmore, M. J. Fogarty, K. Holsman, E. A. Howell, M. E. Monaco, S. A. Oakes, and F. Werner. 2013. assessments : a US perspective. ICES Journal of Marine Science 71:1198–1204.

Levin, P. S., and C. Möllmann. 2015. Marine Ecosystem regime shifts: Challenges and opportunities for ecosystem-based management. Philosophical Transactions of the Royal Society B: Biological Sciences 370(1659):1–8.

Levin, S. A. 1992. the Problem of Pattern and Scale in Ecology. Ecology 73(6):1943–1967.

Link, J. S., J. K. . Brodziak, S. F. Edwards, W. J. Overholtz, D. Mountain, J. W. Jossi, T. D. Smith, and M. J. Fogarty. 2002. Marine ecosystem assessment in a fisheries management context. Canadian Journal of Fisheries and Aquatic Sciences 59(9):1429–1440.

Link, J. S., Brodziak, J. K. T., editors, and Brodziak, J. K. T., Dow, D. D., Edwards, S. F., Fabrizio, M. C., Fogarty, M. J., Hart, D. Jossi, J. W., Kane, J., Lang, K. L., Legault, C. M., Link, J. S., MacLean, S. A., Mountain, D. G., Olson, J., Overholtz, W. J., Palka, D.L., Smith, T. D., contributors. 2002. Status of the Northeast U. S. Continental Shelf Ecosystem: a report of the Northeast Fisheries Science Center's Ecosystem Status Working Group. Northeast Fish. Sci. Cent. Ref. Doc. 02-11; 245 p.Link, J. S., A. Bundy, W. J. Overholtz, N. Shackell, J. Manderson, D. Duplisea, J. Hare, M. Koen-Alonso, and K. D. Friedland. 2011a. Ecosystem-based fisheries management in the Northwest Atlantic. Fish and Fisheries 12(2):152–170.

Link, J. S., J. A. Nye, and J. A. Hare. 2011. Guidelines for incorporating fish distribution shifts into a fisheries management context. Fish and Fisheries 12: 461–469.

LISCCMP 2015 (Long Island Sound Comprehensive Conservation and Management Plan 2015). Accessed at https://longislandsoundstudy.net/wp-content/uploads/2015/09/CCMP_LowRes_Hyperlink_singles.pdf

Lucey, S. M., and M. J. Fogarty. 2013. Operational fisheries in New England: Linking current fishing patterns to proposed ecological production units. Fisheries Research 141:3–12.

Lynch, H. J., R. Naveen, P. N. Trathan, and W. F. Fagan. 2012. Spatially integrated assessment reveals widespread changes in penguin populations on the Antarctic Peninsula. Ecology 93(6):1367–1377.

Mann, Henry, B. 1945. Nonparametric Tests Against Trend. Econometrica 13(3):245–259.

Maunder, M. N., and A. E. Punt. 2004. Standardizing catch and effort data: A review of recent approaches. Fisheries Research 70(2-3 SPEC. ISS.):141–159.

Maxwell, S. M., E. L. Hazen, R. L. Lewison, D. C. Dunn, H. Bailey, S. J. Bograd, D. K. Briscoe, S. Fossette, A. J. Hobday, M. Bennett, S. Benson, M. R. Caldwell, D. P. Costa, H. Dewar, T. Eguchi, L. Hazen, S. Kohin, T. Sippel, and L. B. Crowder. 2015. Dynamic ocean management: Defining and conceptualizing real-time management of the ocean. Marine Policy 58:42–50.

Methratta, E. T., and J. S. Link. 2006. Evaluation of quantitative indicators for marine fish communities. Ecological Indicators 6(3):575–588.

Morse, R. E., K. D. Friedland, D. Tommasi, C. Stock, and J. Nye. 2017. Distinct zooplankton regime shift patterns across ecoregions of the U.S. Northeast continental shelf Large Marine Ecosystem. Journal of Marine Systems 165:77–91. Elsevier B.V.

NEFSC. 2017a. State of the Ecosystem—Mid-Atlantic Bight. Northeast Fisheries Science Center, Woods Hole, MA.

NEFSC. 2017b. State of the Ecosystem—Gulf of Maine and Georges Bank. Northeast Fisheries Science Center, Woods Hole, MA.

NEFSC (2018a). State of the ecosystem 2019: Mid-Atlantic. 2019. Retrieved from https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports

NEFSC (2018b). State of the ecosystem 2019: New England. 2019. Retrieved from https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports

NEFSC (2019a). State of the ecosystem 2019: Mid-Atlantic. 2019. Retrieved from https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports

NEFSC (2019b). State of the ecosystem 2019: New England. 2019. Retrieved from https://www.integratedecosystemassessment.noaa.gov/regions/northeast/reports

Nye, J. A., T. M. Joyce, Y. O. Kwon, and J. S. Link. 2011. Silver hake tracks changes in Northwest Atlantic circulation. Nature Communications 2(1):412–416. Nature Publishing Group.

Nye, J. A., J. S. Link, J. A. Hare, and W. J. Overholtz. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Marine Ecology Progress Series 393:111–129.

Oliver, E. C. J., M. G. Donat, M. T. Burrows, P. J. Moore, D. A. Smale, L. V. Alexander, J. A. Benthuysen, M. Feng, A. Sen Gupta, A. J. Hobday, N. J. Holbrook, S. E. Perkins-Kirkpatrick, H. A. Scannell, S. C. Straub, and T. Wernberg. 2018. Longer and more frequent marine heatwaves over the past century. Nature Communications 9(1):1–12. Springer US.

Östman, Ö., A. Lingman, L. Bergström, and J. Olsson. 2017. Temporal development and spatial scale of coastal fish indicators in reference ecosystems: hydroclimate and anthropogenic drivers. Journal of Applied Ecology 54(2):557–566.

Otto, S. A., M. Kadin, M. Casini, M. A. Torres, and T. Blenckner. 2018. A quantitative framework for selecting and validating food web indicators. Ecological Indicators 84(March 2017):619–631. Elsevier.

Pershing, A. J., M. A. Alexander, C. M. Hernandez, L. A. Kerr, A. Le Bris, K. E. Mills, J. A. Nye, N. R. Record, H. A. Scannell, J. D. Scott, G. D. Sherwood, and A. C. Thomas. 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science 350(6262):809–812.

Peterson, G., C. R. Allen, and C. S. Holling. 1998. Ecological resilience, biodiversity, and scale. Ecosystems 1(1):6–18.

Preciado, I., N. L. Arroyo, J. M. González-Irusta, L. López-López, A. Punzón, I. Muñoz, and A. Serrano. 2019. Small-scale spatial variations of trawling impact on food web structure. Ecological Indicators 98(June 2018):442–452. Elsevier.

Queirós, A. M., J. A. Strong, K. Mazik, J. Carstensen, J. Bruun, P. J. Somerfield, A. Bruhn, S. Ciavatta, E. Flo, N. Bizsel, M. Özaydinli, R. Chuševė, I. Muxika, H. Nygård, N. Papadopoulou, M. Pantazi, and D. Krause-Jensen. 2016. An Objective Framework to Test the Quality of Candidate Indicators of Good Environmental Status. Frontiers in Marine Science 3(May).

R Development Core Team. 2019. R: A Language and Environment for Statistical Computing.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax. 2007. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. Journal of Climate 20(22):5473–5496.

Rice, J. C., and M. J. Rochet. 2005. A framework for selecting a suite of indicators for fisheries management. ICES Journal of Marine Science 62(3):516–527.

Richards, D. R., and D. A. Friess. 2015. A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. Ecological Indicators 53:187–195. Elsevier Ltd.

Sathyendranath, S., R. J. W. Brewin, C. Brockmann, V. Brotas, B. Calton, A. Chuprin, P. Cipollini, A. B. Couto, J. Dingle, R. Doerffer, C. Donlon, M. Dowell, A. Farman, M. Grant,

S. Groom, A. Horseman, T. Jackson, H. Krasemann, S. Lavender, V. Martinez-Vicente, C. Mazeran, F. Mélin, T. S. Moore, D. Müller, P. Regner, S. Roy, C. J. Steele, F. Steinmetz, J. Swinton, M. Taberner, A. Thompson, A. Valente, M. Zühlke, V. E. Brando, H. Feng, G. Feldman, B. A. Franz, R. Frouin, R. W. Gould, S. B. Hooker, M. Kahru, S. Kratzer, B. G. Mitchell, F. E. Muller-Karger, H. M. Sosik, K. J. Voss, J. Werdell, and T. Platt. 2019. An ocean-colour time series for use in climate studies: The experience of the ocean-colour climate change initiative (OC-CCI). Sensors (Switzerland) 19(19).

Sen, P. K. 1968. Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association, 63: 1379–1389.

Sherman, K. 1991. The large marine ecosystem concept: research and management strategy for living marine resources. Ecological Applications 1(4):349–360.

Shin, Y. J., A. Bundy, L. J. Shannon, M. Simier, M. Coll, E. A. Fulton, J. S. Link, D. Jouffre, H. Ojaveer, S. MacKinson, J. J. Heymans, and T. Raid. 2010. Can simple be useful and reliable? Using ecological indicators to represent and compare the states of marine ecosystems. ICES Journal of Marine Science 67(4):717–731.

Shin, Y. J., J. E. Houle, E. Akoglu, J. L. Blanchard, A. Bundy, M. Coll, H. Demarcq, C. Fu, E. A. Fulton, J. J. Heymans, B. Salihoglu, L. Shannon, M. Sporcic, and L. Velez. 2018. The specificity of marine ecological indicators to fishing in the face of environmental change: A multi-model evaluation. Ecological Indicators 89(February):317–326.

Shin, Y. J., and L. J. Shannon. 2010. Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 1. the indiSeas project. ICES Journal of Marine Science 67(4):686–691.

Sokal, R. R., and N. L. Oden. 1978. Spatial autocorrelation in biology 2: some biological implications and four examples of evolutionary and ecological interest. Biological Journal of the Linnean Society 10:229-249.

Steneck, R. S., and R. A. Wahle. 2013. American lobster dynamics in a brave new ocean. Canadian Journal of Fisheries and Aquatic Sciences 70(11):1612–1624.

Steneck, R. S., and J. A. Wilson. 2010. A fisheries play in an ecosystem theater: Challenges of managing ecological and social drivers of Marine fisheries at multiple spatial scales. Bulletin of Marine Science 86(2):387–411.

Tam, J. C., J. S. Link, A. G. Rossberg, S. I. Rogers, P. S. Levin, M. J. Rochet, A. Bundy, A. Belgrano, S. Libralato, M. Tomczak, K. Van De Wolfshaar, F. Pranovi, E. Gorokhova, S. I. Large, N. Niquil, S. P. R. Greenstreet, J. N. Druon, J. Lesutiene, M. Johansen, I. Preciado, J. Patricio, A. Palialexis, P. Tett, G. O. Johansen, J. Houle, and A. Rindorf. 2017. Towards ecosystem-based management: Identifying operational food-web indicators for marine ecosystems. ICES Journal of Marine Science 74(7):2040–2052.

Theil, H. 1992. A rank-invariant method of linear and polynomial regression analysis. In Henri Theil's Contributions to Economics and Econometrics, pp. 345–381. Ed. by B. Raj and J. Koerts. Springer, Dordrecht.

Turnhout, E., M. Hisschemöller, and H. Eijsackers. 2007. Ecological indicators: Between the two fires of science and policy. Ecological Indicators 7(2):215–228.

Uusitalo, L., H. Blanchet, J. H. Andersen, O. Beauchard, T. Berg, S. Bianchelli, A. Cantafaro, J. Carstensen, L. Carugati, S. Cochrane, R. Danovaro, A. S. Heiskanen, V. Karvinen, S. Moncheva, C. Murray, J. M. Neto, H. Nygård, M. Pantazi, N. Papadopoulou, N. Simboura, G. Srebaliene, M. C. Uyarra, and A. Borja. 2016. Indicator-based assessment of marine biological diversity-lessons from 10 case studies across the european seas. Frontiers in

Marine Science 3(SEP).
Wood, S. N. 2001. mgcv: GAMs and generalized ridge regression for R. R News.

A review and exploration of potential ecosystem indicators for the New York Bight using previously existing datasets

January 15, 2020

Contributors: Kurt Heim, Janet Nye, Lesley Thorne, Joseph Warren, Eleanor Haywood

Summary

- This report presents indicators developed for the NYB (using previously existing datasets) that can become part of a future DEC "Ocean Indicator System" (OAP item #33)
- All steps necessary to reproduce each indicator are documented in separate files that will be made available to project collaborators. These show:
  - Where and how to access datasets/who to contact
  - How to prepare data for analysis
  - Computer code to analyze the data and produce indicators
- Many datasets were identified and accessed to calculate nearly 200 indicator time series covering social, economic, recreational fishing, commercial fishing, lower trophic level biota, mid-trophic level biota, and upper trophic level biota in the NYB.
  - Instructions are stored in 18 ".R" or ".Rmd" files
  - In some instances, generating indicators is possible with existing data, but would require substantial expertise or time to complete. These are noted in the report and can be prioritized for further work.
- Here we have produced a long list of indicators that can be developed with previously existing datasets; it is necessary to prioritize those for use, and identify others to pursue
  - Final indicator selection will depend on the desired audience, format, intended use, and goals of the "Ocean Indicators System" (OAP #33)
  - Discussion of these topics will be needed to guide future efforts
- This result focuses on time series indicators; additional work (guided by discussions and needs of DEC) can include
  - Spatial analysis of these datasets to show 'hotspots' and spatial variation within NYB
  - Discussion of indicators made possible by ongoing offshore survey work

Introduction

This report addresses **objective 1** of the MOU (AM10560) between the School of Marine and Atmospheric Sciences (SoMAS) and the New York Department of Environmental Conservation (DEC). This objective is to examine existing datasets that will be useful for developing indicators for the NY Bight (NYB) ecosystem and identify data gaps. Objective 1 is linked to some specific items in the Ocean Action Plan (OAP) including:

Item #32: Establish a baseline ocean monitoring system for the NYB
Item #33: Develop an ocean indicators system for the NYB
Item #34: Publish a State of the Ocean Report (that includes ocean indicators)

The work was performed primarily by Kurt Heim (SoMAS post-doc) with assistance from SoMAS PIs and project collaborators. Additionally, Kurt met regularly (~ once monthly, July - October 2019) with the DEC to discuss priority datasets and indicators to focus on. An outline for this report was reviewed during these discussions. The report includes:

Part 1: A brief review of existing ecosystem monitoring efforts in the region
Part 2: A description of how work was documented (data access, analysis, storage)
Part 3: Presentation of key datasets identified, and indicators developed

What are indicators?

Ecosystem indicators are measurements of key ecosystem components over time that describe the physical, biological, and socioeconomic status of the system. These standardized metrics help scientists and managers to (1) understand natural variability (2) determine what drives this variability (3) quantify human services provided by ecosystems and, importantly, (4) recognize how human activities (i.e., fishing, development, climate change) impact ecosystems. Thus, ecosystem indicators are a fundamental component of Ecosystem Based Management (Link 2010) as they serve as important metrics fundamental to decision making. Indicators are also an important way to provide information to stakeholders.

The figure below shows some of the important considerations for developing an ocean indicators system (OAP item #33). In this report we focus on step 1(gathering potential indicators with previously existing data), which is recommended in Link (2013) and was inclusive as possible. As the NY Offshore Survey is in its infancy, time series indicators cannot yet be generated, but these data provide many potential options. The 2019 annual report summarizes these datasets. As this project moves forward, it will be important to discuss what other potential indicators should be explored (based on existing data), how to reduce the ‘long laundry list of indicators’ presented here to those of greatest relevance, and how these will ultimately be used and presented and to what audience. Development and use of indicator systems is an adaptive process so as lessons are learned, new datasets developed, and needs/interests of stakeholders and managers change the indicator system will be updated.

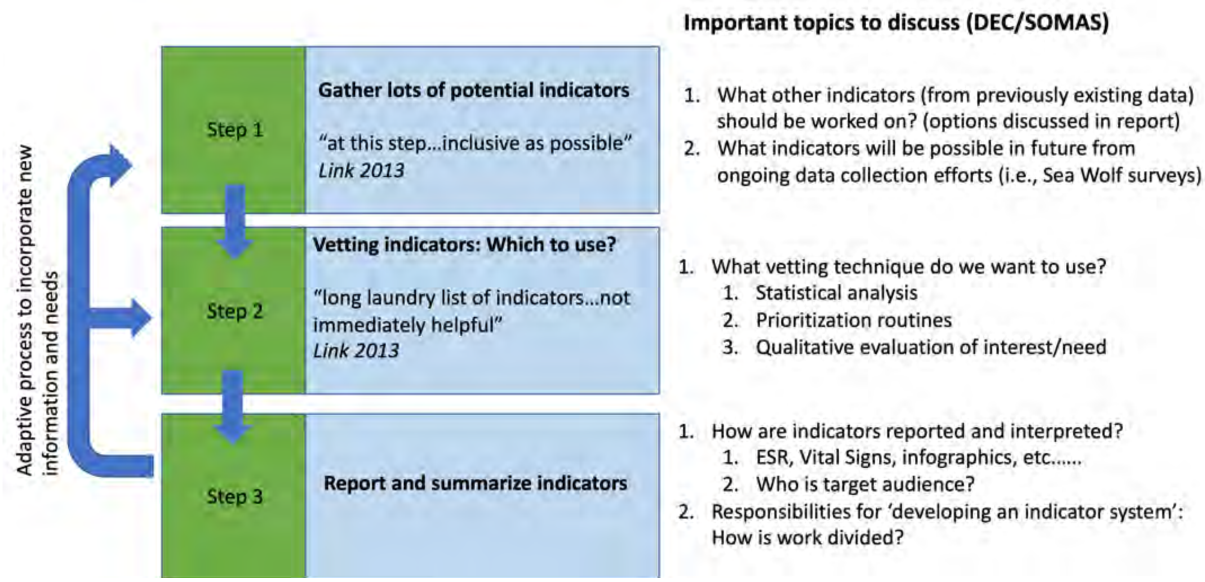


Figure I. Steps in developing an indicator system; we are currently at step 1 (gathering potential indicators).

Part 1: Ecosystem reporting in the Mid-Atlantic region

Summary

- NOAAs ecosystem status report (ESR) for the Mid-Atlantic Bight provides a roadmap to useful data sources and many potential indicators
- 30 people contribute annually to NOAAs ESR; a DEC ocean indicator system (OAP #33) will likely need to be less comprehensive
- The Long Island Sound Study (LISS) "ecosystem targets and supporting indicators" system is briefly reviewed and contrasted with NOAAs ESR
- Indicator systems reviewed are interdisciplinary collaborations; contributions of final indicator time-series are generally provided by a variety of subject matter experts

As it pertains to OAP item #33 (develop an ocean indicator system), there are several reasons why it is important to know what other agencies are doing. We can (1) use their datasets and (2) review their end products to guide the scope and detail of our work. Knowledge of other efforts is also important to (3) make sure our work is not duplicative (both in terms of reporting, scope, and data collection efforts).

Mid-Atlantic Bight Ecosystem Status Report by NOAA

The Northeast Fisheries Science Center (NEFSC) has been developing a suite of indicators for the Mid-Atlantic region for nearly 20 years and provides a good starting point to identify key datasets and useful indicators for the NYB (Methratta and Link 2006). The NEFSC uses several long-established monitoring programs (a spring and fall bottom trawl survey for groundfish, regular ecosystem monitoring cruises [EcoMon], scallop surveys, etc.) as key data sources for their description of the ecosystem. They also incorporate other datasets (satellite data, socio-economic data, etc.). They divide the Northeast US Continental Shelf Large Marine Ecosystem (NES) into four sub-regions for reporting (Figure 1.1). The NYB is in the Mid-Atlantic Bight.

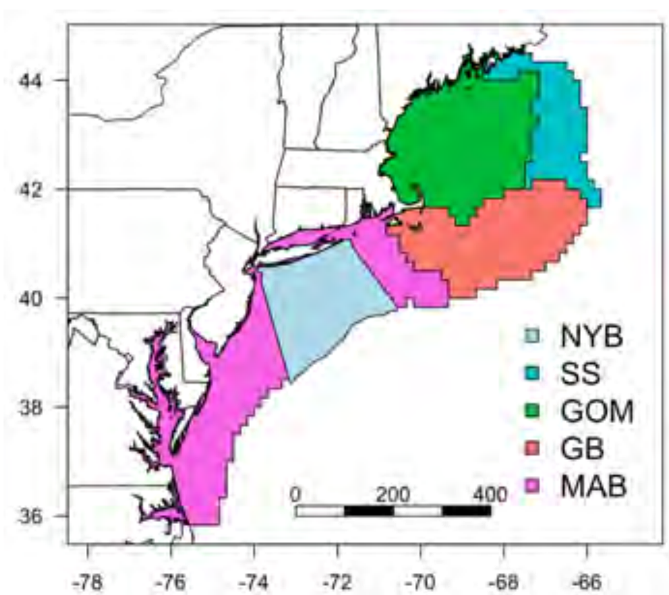


Figure 1.1. The New York Bight (NYB) study area and the Ecological Production Units (EPUs) defined by NOAA including the Scotian Shelf (SS), Gulf of Maine (GOM), Georges Bank (GB), and the Mid-Atlantic Bight (MAB).

In their reports, NOAA has transitioned from a very long and comprehensive presentation of indicators (254 page report in 2002) to a more concise Ecosystem Status Report (ESR) that was most recently 28 pages (in the 2019 report) and included about ~70 indicators (**Table 1.1**). The indicators used by NOAA have changed over time, highlighting that ecosystem indicators are adapted to meet changing needs of stakeholders, managers, and policy makers (Figure I; Link 2010). NOAA incorporates feedback annually to improve the ESR for decision makers and stakeholders. A complete list of indicators in the 2019 report is presented in **Table A1**.

Table 1.1. A brief review of NOAAs Ecosystem Status Reports for the Northeast US Shelf Large Marine Ecosystem.

Report title	Year	Contributors	Pages
Status of the Northeast US continental shelf ecosystem	2002	17	254
Ecosystem Status Report For the Northeast U.S. Continental Shelf Large Marine Ecosystem	2009	31	40
Ecosystem Status Report for the Northeast Shelf Large Marine Ecosystem - 2011	2011	27	39
State of the Ecosystem - Mid-Atlantic	2017	-	16
State of the Ecosystem - Mid-Atlantic Bight	2018	31	23
State of the Ecosystem 2019: Mid-Atlantic	2019	31	28

This ESR is part of an Integrated Ecosystem Assessment (IEA) framework to incorporate ecosystem considerations into management (DePiper et al. 2017; Gaichas et al. 2018). Their IEA and ESR work is supported by a full time staff of 8, plus 3 contractors or post-docs. Each year their updated ESR involves contributions by around 30 people. The work is structured such that different subject matter experts are 'in charge' of developing and maintaining updated indicators to contribute to the ESR each year. This workflow can be reviewed at their website and technical documentation site.

https://www.integratedecosystemassessment.noaa.gov

Long Island Sound Study and Ecosystem reporting

The Long Island Sound Study (LISS) does not cover the same geographic scope of our work, but provides an example of a more localized ecosystem indicator program (relative to the entire Mid-Atlantic) at an earlier stage of development than NOAAs. LISS is a large, well-funded EPA led program (~$14 million in 2019) that presents a series of ecosystem indicators on their website and in annual reports. Perhaps more so than the NOAA ESR, the LISS indicator system focuses on very specific indicators that relate to specified and quantitative management targets for the ecosystem. Because LISS also is a large source of funding for restoration projects, indicators are used to clearly define progress and quantity benefits of funded projects. They define "Ecosystem Targets" under the following categories

- Clean waters and healthy watershed (7 targets + indicators)
- Thriving habitats and abundant wildlife (7 target + indicators)
- Sustainable and resilient communities (6 targets + indicators)
- Climate change (5 indicators to detect effects of climate change)

The LISS website lists many indicators and their data sources, and who is responsible for maintaining the indicator. Like NOAAs indicator system, many contributors provide the final indicator time series for use in reporting and presentation on the website. Although these datasets are generally not applicable (i.e., are for the Long Island Sound, not the NYB), many of the social and economic datasets could be used in a DEC indicator program. For example, an indicator used is “number of beach closures and advisories” which could be calculated for NY oceanfront beaches and used an indicator. The LISS calculates this for beaches on the Long Island Sound only.

Other datasets of interest are discussed under their respective data type category (**Part 3.1–3.6**). The website for the LISS indicator system can be found at the following link.

http://longislandsoundstudy.net/research-monitoring/liss-ecosystem-targets-and-supporting-indicators/all-ecosystem-targets-and-indicators/

Part 2: Workflow and documentation (data access, analysis, storage)

Summary

- For indicators of interest we searched for the highest quality datasets available
- When data was accessible and appropriate, indicators were calculated and workflow documented in R markdown files
- Final indicator time series are stored in a standard format in ".csv" files

The approach taken for Objective 1 was to (1) consider an indicator of interest (e.g., surface water temperature, large whale strandings, zooplankton biovolume) then (2) identify if datasets were available to calculate it and if so (3) calculate the time series for the NYB and document all steps taken. We focused on indicators of interest to the DEC and ones used elsewhere to describe large marine ecosystems (LMEs). When data were available but extensive expertise or work were required to generate an indicator, these are noted in the report (Part 3) as potential indicators requiring further attention.

We did not generate a complete list or survey all existing datasets for the NYB. Rather, we focused on indicators of interest and then asked if it was possible to calculate those with existing data. There is a tremendous amount of data available in the region, and a complete synthesis of these data would require more personnel and experts in a variety of fields. Moreover, existing data for the region have already been reviewed at length and reported by NYSERDA in the NY State Offshore Wind Master Plan documents. There are studies reviewing (1) fish and fisheries (202 pages), birds and bats (142 pg), marine mammals and sea turtles (164 pg), and shipping and navigation (84 pg) available online.

https://www.nyserda.ny.gov/All-Programs/Programs/Offshore-Wind/Marine-Ecosystems

Indicators that were calculated with existing data are discussed in detail in Part 3 of this report. Part 3 also discusses instances where data exists, but further work is needed, and also what indicators of interest could not be calculated because of data gaps.

Documentation of data access, analysis, and storage

We calculated many indicators for the NYB, and documented all steps involved including where to get the data, how to get it, how to process it, how to analyze it to generate a time series, and how to store it. These 'instructions' are stored in a series of R markdown files, and listed in **Appendix 1**. Markdown files integrate text, computer code, and figures seamlessly and can be viewed in R Studio or output to word, .html (i.e., viewed in a web browser), or pdf files. In the future, these documents can be used to regularly update the indicators; one would simply need to follow the instructions to get an updated data set and re-run the script.

The workflow involves finding and accessing a data source; this is described in detail and web-links are provided in the R markdown files. Once the analysis is performed the data are written to disk in a standard format that is described in **Table 2.1**. This is a “long format” where a particular indicator is accessed by querying “Variable” column.

Of primary interest is the NYB, but for comparative purposes the indicators were calculated at a variety of spatial scales. T**able 2.1** shows the standard data format for final series, the **ScaleType** column describes the spatial scale that applies to the indicator. For example it may be *global* (whole ecosystem, or whole world), *State* (calculated for each US State), *LME* (the mid-Atlantic bight [MAB], Southern New England [SNE], Gulf of Maine [GME], Grand Banks[GB], or the New York Bight [NYB]), *stratum* (each stratum of the NMFS bottom trawl survey within the North East Atlantic Shelf Large Marine Ecosystem [NES LME), or *county* (counties within states). The **ScaleName** describes the name of the scale unit (i.e., the name of the state, county, or strata); the ScaleName2 describes the next higher up ScaleName (i.e., if the indicator is calculated at **ScaleType** of *county* then the **ScaleType2** column would indicate what *state* the county is in.

Table 2.1. Description of fields in final formatted dataset

Column name	Description	Example
Year	Year	1990
Variable	Abbreviated name of indicator	Winter_OISST
Val	Value of indicator for that year	14.9
ScaleType	Scale that calculation was done	Global, LME, County, State
ScaleName	Scale unit name	MAB, NYB, Suffolk County, NY
ScaleName2	Secondary scale unit name	MAB, GOM
ScaleName3	Tertiary scale unit name	MAB, GOM
N	Sample size used for annual calculation	100

Part 3: Key datasets, indicators developed, and potential indicators

3.1: Physical oceanographic indicators

Summary

- Many datasets exist, but are scatted across sources and in myriad formats
- The World Ocean Dataset (WOD) is a useful data integration platform for in situ measurements throughout the water column but is limited in spatial and temporal coverage
- Satellite datasets are ideal for indicator development because of high spatial and temporal resolution, but are limited to surface conditions
- 36 indicators are calculated (30 for NYB, 6 broad-scale drivers)
- Surface and bottom temperatures are warming
- Marine heatwaves are increasing (more, longer, more intense)
- Stratification appears to be decreasing (but not significantly) in recent years
- Many indicators are redundant based on correlation analysis
- Additional indicators and datasets are discussed

Key datasets identified and data coverage

Physical oceanography datasets are numerous, but scattered across different agencies, user groups, universities, private entities, and data repositories. Moreover, these data are in varied formats (i.e., csv, netCDF, shp, mat, Rda). For the purpose of regular use in a DEC/SoMAS indicator project, tracking down many scattered datasets to update indicators (annually, or every few years) is not likely practical. We therefore focused on datasets that can be easily accessed, are updated regularly, and/or already integrate datasets from multiple sources. We focus on (1) available indices (e.g., NAO, AMO, PDO) provided from various website (2) data available from the World Ocean Dataset (WOD) platform and (3) satellite derived SST (OISST). Other datasets used but not discussed at length include river discharge data from the U.S. Geological Society (USGS) and buoy data from NOAA.

World Ocean Dataset coverage

Data were accessed through the WODselect website to gather all available high resolution CTD data (see below, other datasets are available) for the Northeast US Continental Shelf LME (n = 78,359 casts). The spatial and temporal coverage of the resulting CTD dataset is shown for the last 25 years (**Figure 3.1.1, Figure 3.1.2**). Out of 78,359 CTD casts made in the NES since 1961 a total of 8,243 of them have been in the NYB study area (10%) (**Table 3.1.1**) The last five years have a substantially lower number of casts than the preceding 20 years, because of the lag time between when researchers take measurements, quality control the data, and then upload it.

The CTD dataset is only one of many data types available from the WODselect website, others included are shown in T**able 3.1.1 and 3.1.2**. These estimates were based on a query using geographic coordinates in the WODselect website on September 16th, 2019 and should be considered approximate.

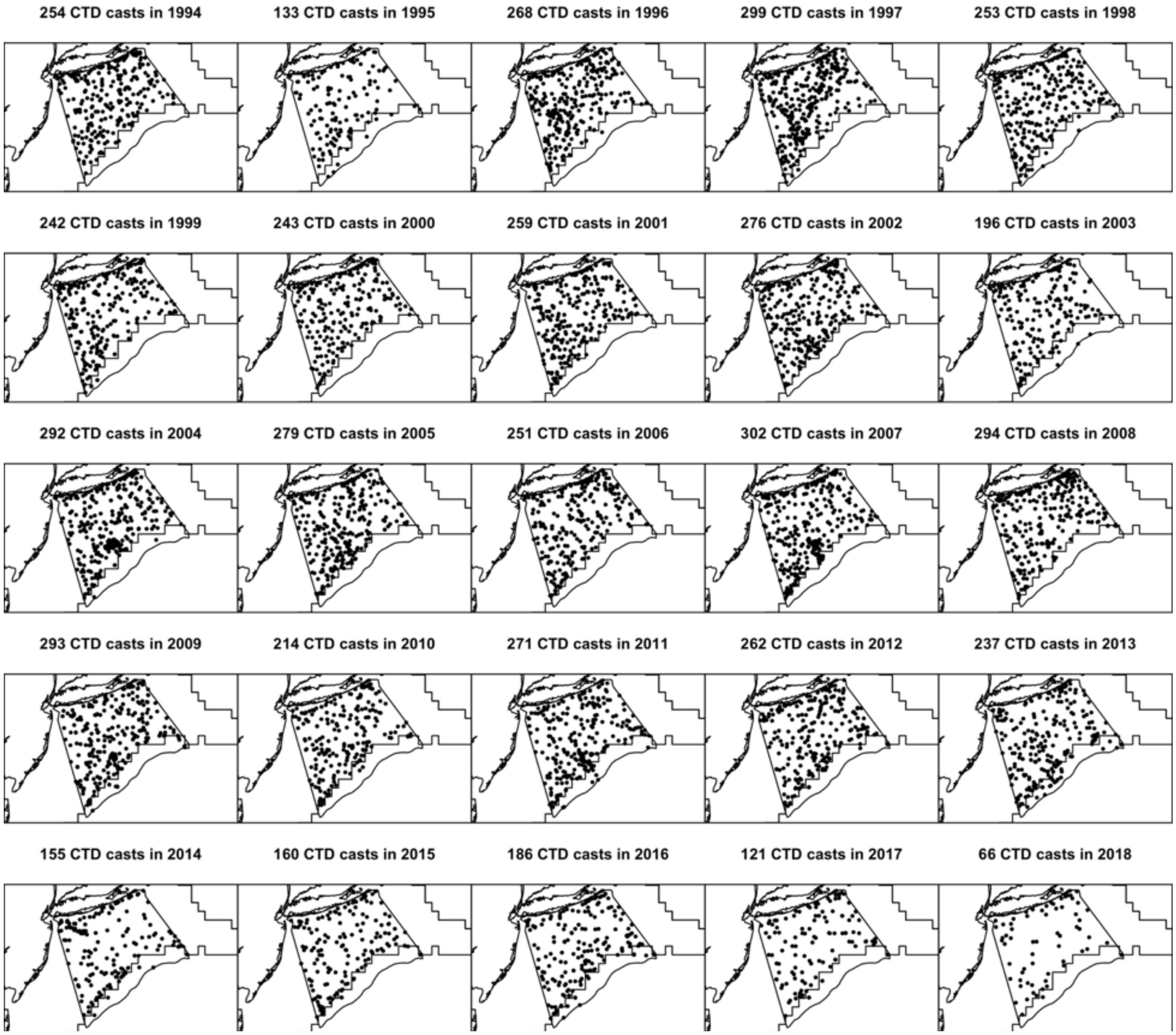


Figure 3.1.1. The spatial distribution of CTD casts present in the WOD dataset from 1994 to 2018. The NY Bight region and the Mid-Atlantic Bight (MAB) Ecological Production Unit boundaries are shown.

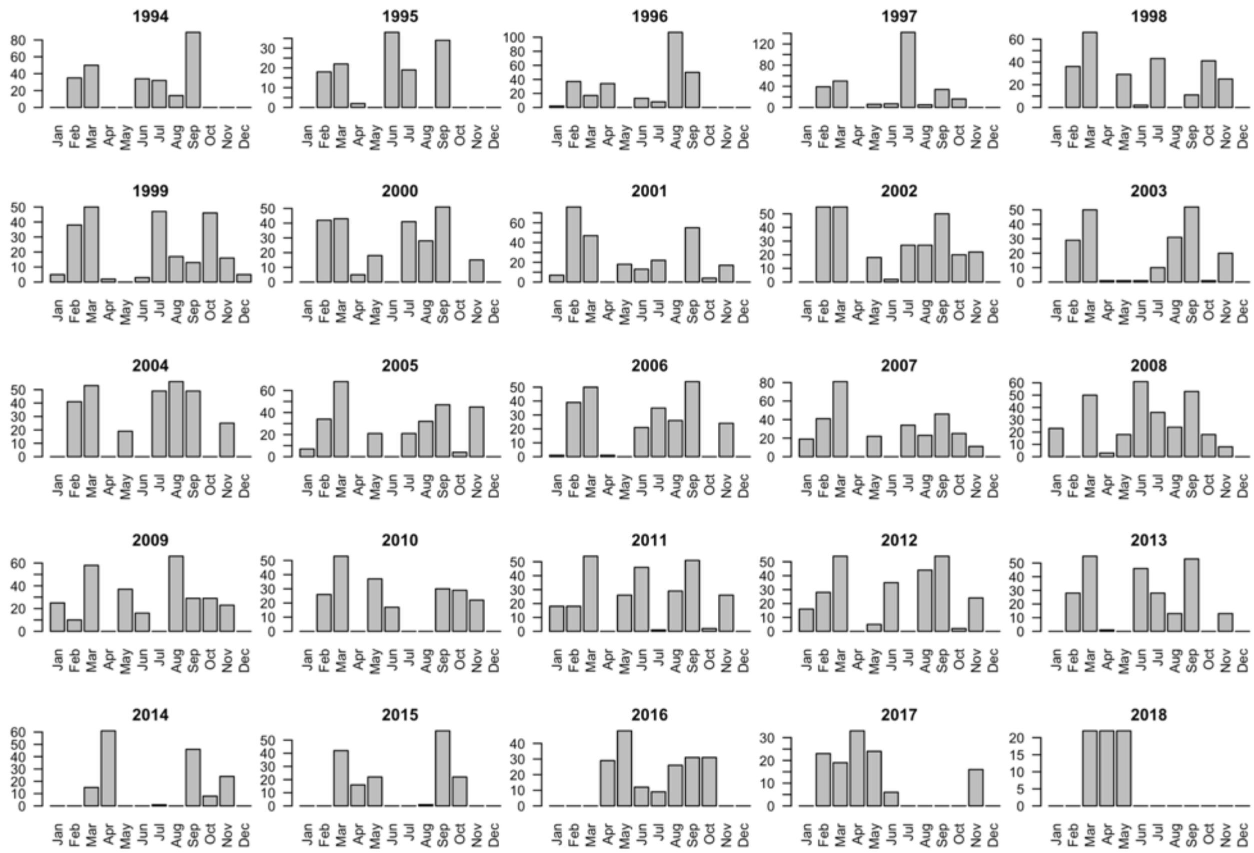


Figure 3.1.2. The temporal distribution of CTD casts present in the WOD dataset from 1994 to 2018.

Table 3.1.1. The number of casts available in the North East US Continental Shelf Large marine ecosystem (NES) and the New York Bight (NYB) from the World Ocean Database (WOD) as of September 16th, 2019.

Data type	Casts NES[1]	Casts NYB[2]	Explanation
OSD casts	122324	20875	Ocean Station Data [Bottle, low resolution CTD/XCTD, plankton data]
CTD casts	78359	11917	High Resolution CTD/XCTD
XBT casts	110683	22483	Expendable Bathythermograph
MBT casts	191683	28330	Mechanical Bathythermographs [includes Digital Bathythermograph, μBT]
PFL casts	6320	338	Profiling Floats
DRB casts	1	0	Drifting Buoys
MRB casts	0	0	Moored Buoys [TAO, PIRATA, others]
APB casts	1452	0	Autonomous Pinniped Bathythermographs
UOR casts	41489	0	Undulating Oceanographic Recorder [Towed CTD]
SUR cruises	708	33	Surface-Only [Bucket, Thermosalinograph]
GLD casts	115063	27829	Glider data
Total casts	668082	111805	

[1]The boundaries used to define the NES for this inquiry were N = 45.00, S = 35.00, E = -64.00, W = -78.00
[2]The boundaries used to define the NYB for this inquiry were N = 41.0755, S = 38.43636, E = -70.56894, W = -73.95522

Table 3.1.2. Other physical variables available in the World Ocean Database Portal accessible through WODselect.

Variable	Units	Where Found?
Temperature	°C	OSD, CTD, MBT, XBT, SUR, APB, MRB, PFL, UOR, DRB, GLD
Salinity	unitless	OSD, CTD, SUR, APB, MRB, PFL, UOR, DRB, GLD
Oxygen	µmol/kg	OSD, CTD, PFL, UOR, GLD, DRB
Phosphate	µmol/kg	OSD
Silicate	µmol/kg	OSD
Nitrate and Nitrate+Nitrite	µmol/kg	OSD, PFL
pH	unitless	OSD, SUR, PFL
Chlorophyll	µg/l	OSD, CTD, SUR, UOR
Plankton	multiple	OSD
Alkalinity	meq/l	OSD, SUR
Partial Pressure of Carbon Dioxide	µatm	OSD, SUR
Dissolved Inorganic Carbon	mM	OSD
Transmissivity	1/m	CTD, PFL
Pressure	dbar	OSD, CTD, UOR, GLD, PFL
Air temperature	°C	SUR
CO_2 warming	°C	SUR
CO_2 atmosphere	ppm	SUR
Air pressure	mbar	SUR
Tritium	TU	OSD
Helium	nmol/kg	OSD
Delta Helium-3	%	OSD
Delta Carbon-14	‰	OSD
Delta Carbon-13	‰	OSD
Argon	nmol/kg	OSD
Neon	nmol/kg	OSD
Chlorofluorocarbon 11 (CFC 11)	pmol/kg	OSD
Chlorofluorocarbon 12 (CFC 12)	pmol/kg	OSD
Chlorofluorocarbon 113 (CFC 113)	pmol/kg	OSD
Delta Oxygen-18	‰	OSD

NOTE: Table data is found in DEC_indicators/datasetinfo/WOD_dataQueryTable.xlsx

Optimally interpolated sea surface temperature data

The optimally interpolated sea surface temperature (OISST) dataset (Reynolds et al. 2007) is useful for indicator development because of high spatial and temporal data availability (**Figure 3.1.3**). It dates back to 1981 and provides daily estimates of SST for 74 grid cell locations within the NYB. This data product can be used for simple summary calculations, or more complex analysis such as marine heatwaves or assessing long term trends in seasonal warming (Henderson et al. 2017).

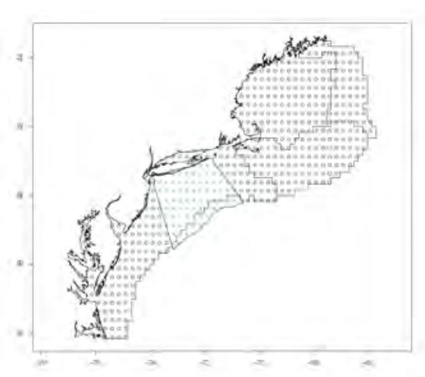


Figure 3.1.3. The grid points where daily sea surface temperature estimates are made in the optimally interpolated sea surface temperature dataset (OISST). A total of 74 fall within the New York Bight, 212 fall within the Mid-Atlantic Bight, and 479 fall within the North East US continental shelf large marine ecosystem (NES).

NYB indicator overview-physical conditions

We developed 30 indicators to represent the physical state of the NYB and acquired six indices to represent broad drivers of oceanographic conditions (e.g., NAO, AMO, Co2 concentration) shown in **Table 3.1.3**. The 30 indicators for the NYB were also calculated for the MAB and NES for comparative purposes. All steps are documented in R Markdown files (i.e., how to access data, calculate indicator, save in standard format). Recent trends (i.e., since 2010) as determined by fitted GAMs and evaluation of significant periods of change via first derivatives (details in appendix at end of this document), are shown for all physical indicators in **Table 3.1.4**. Notably, there are no significant decreasing trends, but many increasing trends. Several of interest are described in more detail below.

Table 3.1.3. Physical indicators calculated for New York Bight. First six (bold) are not specific to NYB.

Indicator	Description	Date range
NAO	North Atlantic Oscillation	1856 - 2018
AM0	Atlantic Multidecadal Oscillation	1824 - 2018
Co2	Global Co2 concentration	13 - 2018
GSI	Index of Gulf Stream Position	1954 - 2018
ENSO	El Nino/Southern Oscillation	1979 - 2018
PDO	Pacific Decadal Oscillation	1854 - 2018
Winter OISST	Satellite SST, Winter mean [yday(1:90)]	1981 - 2018
Spring OISST	Satellite SST, Spring mean [yday(91:181)]	1981 - 2018
Summer OISST	Summer SST, Summer mean [yday(182:273)]	1981 - 2018
Fall OISST	Fall SST, Fall mean [yday(274:365)]	1981 - 2018
Marine heatwaves (20 indices)	Indices of marine heatwaves (n = 20)	1981 - 2018
SST (in-situ)	SST index from in-situ measurements	1975 - 2018
Bottom temperature (in-situ)	Bottom temperature index from in-situ measurements	1975 - 2018
Bottom salinity	Bottom salinity index from in-situ measurements	1975 - 2018
Surface salinity	Surface salinity index from in-situ measurements	1975 - 2018
Stratification index	Stratification index from in-situ measurements	1975 - 2018
River flow	Mean and total river discharge (Hudson River)	1947 - 2018

Table 3.1.4. Trends in indicators assessed with linear regression (direction indicated by arrows for significant relationships, r-squared shown in last column). The non-linear trend (since 2010) shows whether there is a significant increasing, decreasing, or no trend (with arrows) as of 2010 determined by GAM fit.

Indicator	Linear trend	Non-linear trend (since 2010)	Linear r-squared
AMO			0
NAO			0.001
Co2	↑	↑	0.327
GSI	↑	↑	0.49
ENSO_meiv2			0.084
PDO_ERSSTV5			0.003
fall_OISST	↑	↑	0.324
spring_OISST	↑		0.126
summer_OISST	↑	↑	0.546
winter_OISST			0.077
OISST_HW_count	↑	↑	0.467
OISST_HW_duration	↑	↑	0.288
OISST_HW_duration_max	↑	↑	0.248
OISST_HW_intensity_mean	↑		0.18
OISST_HW_intensity_max	↑	↑	0.295
OISST_HW_intensity_max_max	↑	↑	0.331
OISST_HW_intensity_var	↑	↑	0.24
OISST_HW_intensity_cumulative	↑	↑	0.306
OISST_HW_intensity_mean_relThresh	↑	↑	0.236
OISST_HW_intensity_max_relThresh	↑	↑	0.319
OISST_HW_intensity_var_relThresh	↑	↑	0.225
OISST_HW_intensity_cumulative_relThresh	↑	↑	0.338
OISST_HW_intensity_mean_abs			0.103
OISST_HW_intensity_max_abs			0.133
OISST_HW_intensity_var_abs	↑	↑	0.223
OISST_HW_intensity_cumulative_abs	↑	↑	0.337
OISST_HW_rate_onset			0.06
OISST_HW_rate_decline			0.01
OISST_HW_total_days	↑	↑	0.4
OISST_HW_total_icum	↑	↑	0.366
BT_index	↑		0.205
ST_index	↑	↑	0.295
bot_sal			0.016
surf_sal		↑	0.023
strat_index			0.025
Hudson_MeanDischarge	↑	↑	0.09
Hudson_TotalDischarge	↑	↑	0.09
Delaware_TotalDischarge			0.003
Delaware_MeanDischarge			0.003

Broad scale drivers

Trends in broad scale drivers are shown with non-linear (GAM) trends fitted; colored areas show regions of significant increases or decreases in the time series (**Figure 3.1.4**). These are well-known to be important in the NES and NYB, and so are not discussed at length.

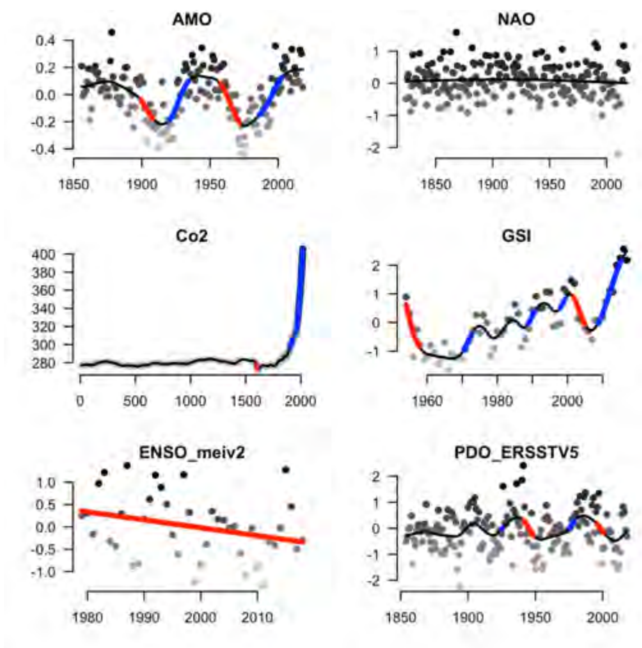


Figure 3.1.4. Broad scale drivers and fitted trends.

Satellite derived sea surface temperature

Winter OISST, Spring OISST, Summer OISST, Fall OISST, and river flow are simple summary statistics of the OISST datasets.

Sea surface temperature (SST) is warming in the NYB (**Figure 3.1.5, Figure 3.1.6, Table 3.1.5**). Using satellite derived data subset to the NYB analyzed with linear regression, summer SST is increasing by an average of 0.05 C°/year or 0.58 C°/decade, and this is consistent with warming in the NES as a whole (0.05 C°/year) and the Mid-Atlantic Bight (0.05 C°/year). All regressions were significant ($p < 0.05$) except for those on winter SST (**Table 3.1.5**). Non-linear GAMs fit to the NYB trendlines show similar, but more nuanced patterns (**Figure 3.1.6**). The edf values represent the complexity of the relationship, when edf is 1 it approaches a linear relationship, when it is higher than 1 it is an increasingly more complex fit. The trend in summer warming is essentially linear (edf = 1.49), while the fall trend shows no trend until about year 2000 after which it becomes significantly increasing. Only a slight period of warming was detected in spring SST, and none for the winter trendline.

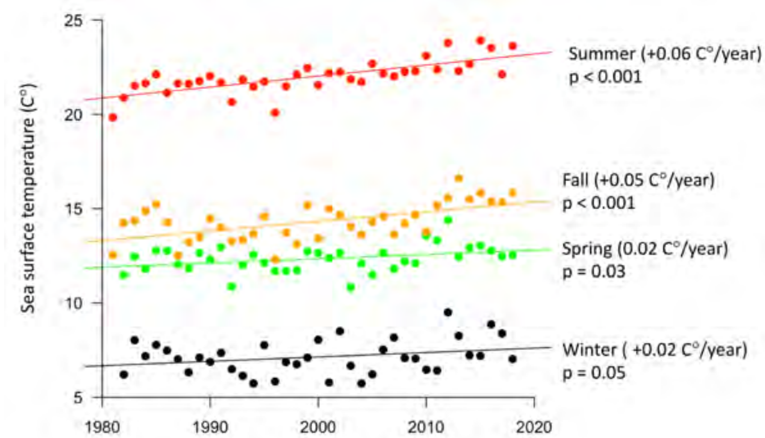


Figure 3.1.5. Linear trends in sea surface temperature in the New York Bight.

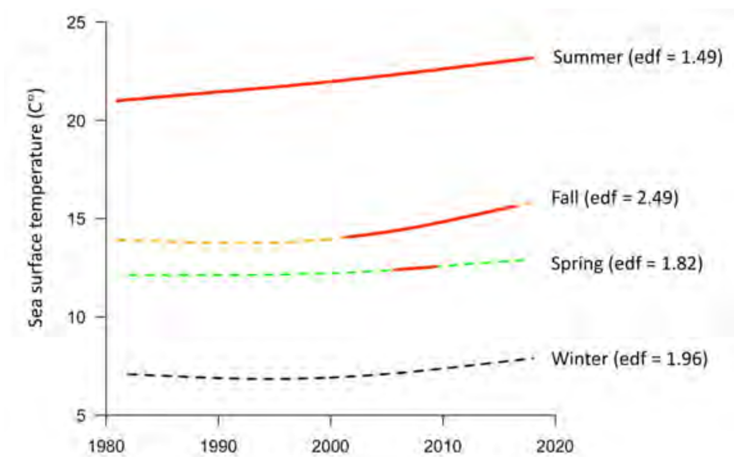


Figure 3.1.6. Non-linear trends in sea surface temperature in the New York Bight.

Table 3.1.5. Linear regression slopes (x 10) to show estimated SST increases per decade (C°) by season and region. Increases per decade are similar across regions; all regressions were significant except those in winter (coefficients shown in bold).

Season	NES	MAB	NYB
Winter	**0.18**	**0.12**	**0.23**
Spring	0.28	0.29	0.23
Summer	0.59	0.54	0.58
Fall	0.50	0.46	0.51

Marine heatwaves

The marine heatwave indicators (n = 20, **Table 3.1.6**) are generated using published and automated analytical methods (Hobday et al. 2016) applied to a daily-average SST dataset for the NYB (i.e., daily average of all cells in the NYB, **Figure 1.1.3**). A marine heatwave is a 'discrete anomalously warm water event in a particular location'; in our analysis we use recommendations of Hobday et al. (2016) to define 'prolonged' as 5 or more days and 'anomalously warm' as a daily temperature that exceeds the 90th percentile based on a 30-year climatology (1981 to 2011). By analyzing the NYB daily average SST data, the 'particular location' is the NYB. This analysis could also be done by defining ‘particular location’ as the individual cell (where SST estimates are made) within the NYB, this would result in 74 estimates of the 20 MHW indices and would allow for spatial comparisons of heatwaves at locations within the NYB. Definitions of MHW statistics are in **table 3.1.6**.

Figure 3.1.7 shows an example MHW that occurred in 2011 and 2012 and helps to show how a MHW event is defined and detected. When the measured temperature increases above the MHW threshold (the blue line is the long-term average, the green line shows the 90th percentile threshold) and stays above this level for five consecutive days, this is classified as a MHW event. **Figure 3.1.7** shows a MHW event for the NYB that lasted 182 days. The other panels (B – D) show statistics by year; all are increasing.

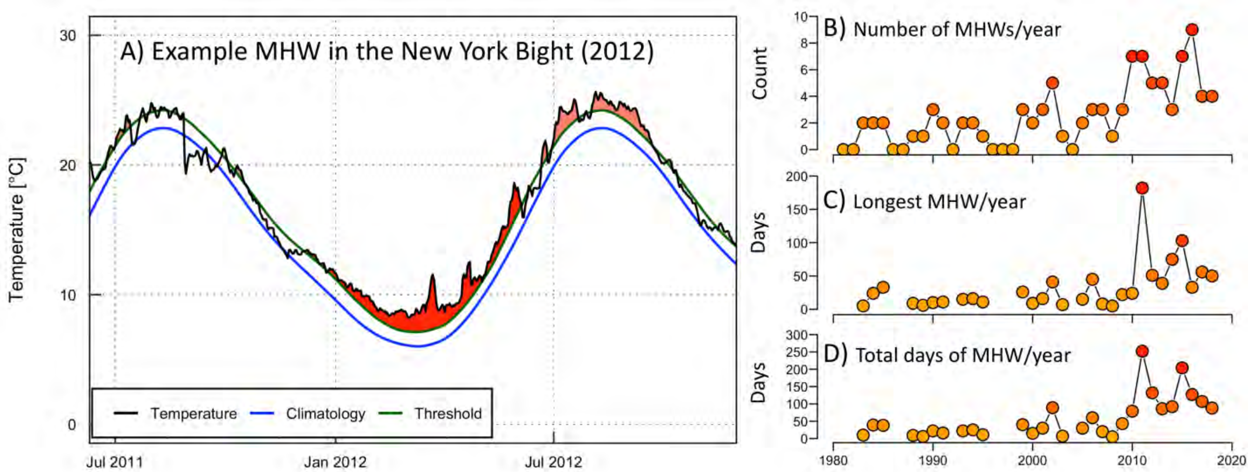


Figure 3.1.7. Marine heatwaves in New York Bight showing the number of heatwaves (A) the annual average duration of MHW events (B), the longest MHW per year (C), and the total number of days per year classified as MHW days (D). All regressions were significant and the slopes are shown (multiplied by ten) to convey the decadal increase in the MHW statistic.

Table 3.1.6. Definition of marine heatwave statistics for the New York Bight. Note that these are copied directly from https://github.com/robwschlegel/heatwaveR

	Description
year	The year over which the metrics were averaged.
count	The number of events per year.
duration	The average duration of events per year [days].
duration_max	The maximum duration of an event in each year [days].
intensity_mean	The average event "mean intensity" in each year [deg. C].
intensity_max	The average event "maximum (peak) intensity" in each year [deg. C].
intensity_max_max	The maximum event "maximum (peak) intensity" in each year [deg. C].
intensity_var	The average event "intensity variability" in each year [deg. C].
intensity_cumulative	The average event "cumulative intensity" in each year [deg. C x days].
rate_onset	Average event onset rate in each year [deg. C / days].
rate_decline	Average event decline rate in each year [deg. C / days].
total_days	Total number of events days in each year [days].
total_icum	Total cumulative intensity over all events in each year [deg. C x days].

Note: intensity_max_relThresh, intensity_mean_relThresh, intensity_var_relThresh, and intensity_cumulative_relThresh are as above except relative to the threshold (e.g., 90th percentile) rather than the seasonal climatology
Note: intensity_max_abs, intensity_mean_abs, intensity_var_abs, and intensity_cumulative_abs are as above except as absolute magnitudes rather than relative to the seasonal climatology or threshold

Bottom temperature index

A bottom temperature index for the NYB was generated from the WOD dataset. Indicators using data acquired from WOD involve some analysis that should be reviewed/discussed prior to use in reports or any final documents. Our approach was to use GAMs to extract a fixed effect of year while accounting for potentially non-linear effects of other variables on observed measurements (e.g., latitude, day of year, depth). The year effect (i.e., the coefficient from the GAM) was used as the index. We validated the approach with the following comparison.

We compared a bottom temperature index published by NOAA for the MAB to the index generated using our GAM approach applied to all WOD data for the MAB. Assuming the NOAA version is accurate, our estimate should be similar to this estimate. NOAAs methods are described here https://noaa-edab.github.io/tech-doc/bottom-temperatures.html and are an 'extensive analysis; not yet published'. Overall, our BT index matched the NOAA version well (r = 0.68) and the time series were visually similar (**Figure 3.1.8**). Two years in the early part of the time series did not match well and should investigated. Given the method appears to be valid for the MAB, we proceeded to use it to estimate a BT index for the NYB.

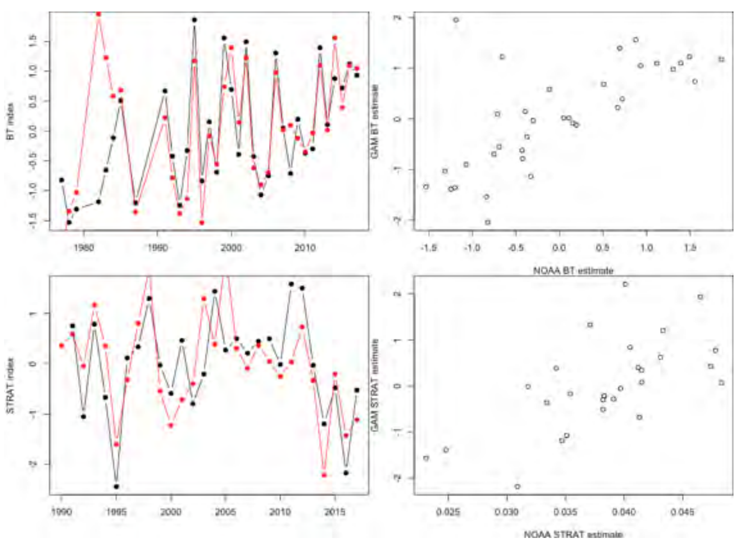


Figure 3.1.8. Validation of bottom temperature index (top two panels) and stratification index (bottom two panels) methodology. The top two panes show the comparison of bottom temperature index published by NOAA for the MAP to one made using our GAM approach (left, red is our estimate, black is NOAAs estimate). The right panel shows their correlation to one another (r = 0.68, but also shows the two years that did not match well). The bottom panes show NOAAs stratification index compared to a stratification index done here using the GAM approach.

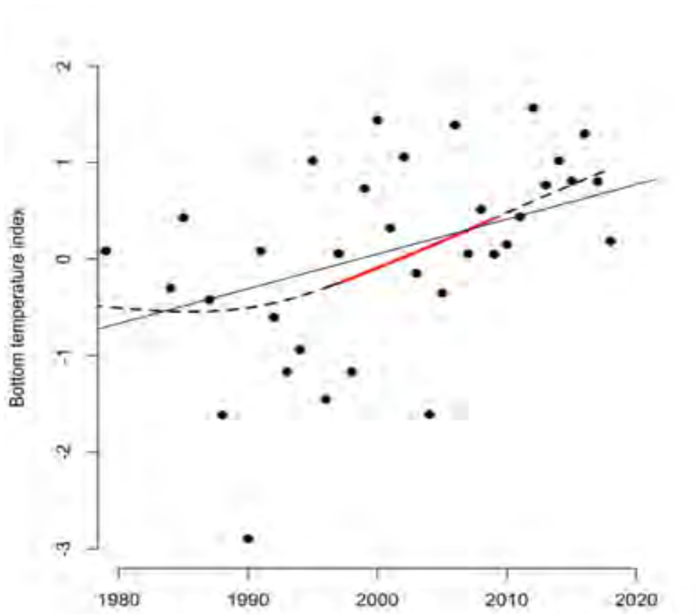


Figure 3.1.9. Bottom temperature index for the NYB. The associated linear regression (straight line) was significant; the GAM fit shows a period of increase but that the trend is no longer significant (since 2009) due to a recent decline over the last two years.

Stratification index

Stratification indices were produced for the NYB following the same method for generating a bottom temperature index. Methods are described in detail in the associated R script. As with the BT index, we validated the method worked by running it for the WOD dataset in the MAB and comparing it to a published NOAA estimate of stratification for the MAB (**Figure 3.1.8**). The trends were similar and correlated (r = 0.65), again, suggesting this is an adequate method.

Stratification appears to be decreasing in recent years, but this change is not statistically significant with linear or non-linear fits (**Figure 1.1.10**). Notable outliers will need to be investigated.

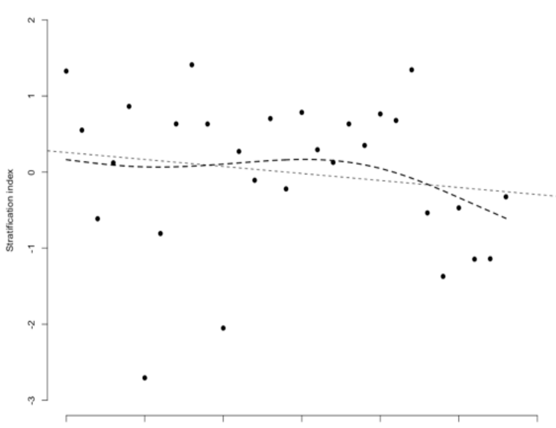


Figure 3.1.10. Stratification index for the New York Bight. The linear and non-linear fits were not significant.

__NOTE__: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

Hudson river flow

River flow is easy to access and work with using the R package "waterdata". **Figure 3.1.11** shows the mean annual flow (in cubic meters per second, CMS) of the Hudson River at green island (USGS gage id 01358000). Based on linear regression, there is an increasing trend ($p < 0.05$) and the coefficient suggest an increase of 1.31 CMS per year or 13 per decade.

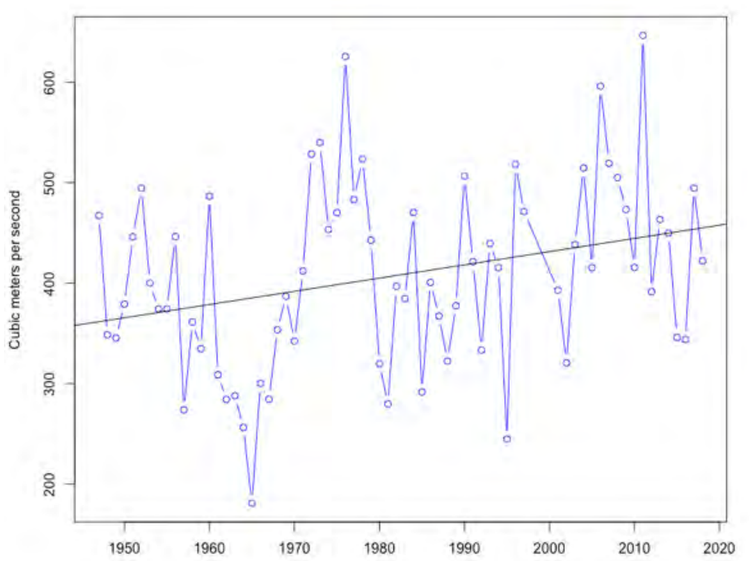


Figure 3.1.11. Mean annual flow of the Hudson River (in cubic meters per second). Note, hydrologists calculate flow statistics based on the 'water year' which is from Oct 1 to Oct 1, and the x axis here is water year. MAF has significantly increased over the period of record based on linear regression.

Relationships between variables and redundancy

Figure 3.1.12 shows correlation values between all physical indicators; clearly many are redundant. We list some points of interest below but note that all of these should be interpreted with caution and suspect given the potential for spurious correlations:

- High positive correlation between most MHW metrics
- Negative correlation between salinity indices and Hudson river flow (i.e., more water in a year is associated with lower salinity in the NYB).
- MHW intensity correlated with salinity (negative) and river flow (positive)

Figure 3.1.13 shows a PCA of the indicators (since 1990), the first two axes explain 44 and 23% of observed variability in the indicator suite. That is, the first two principle components explain a total of 67% of variability in the entire suite of indicators. The factor loadings are consistent with expectations; most are positively correlated (i.e., point the same way on the bioplot, also see **Figure 1.1.12**) expect for Hudson River discharge which corresponds with the second PC axis.

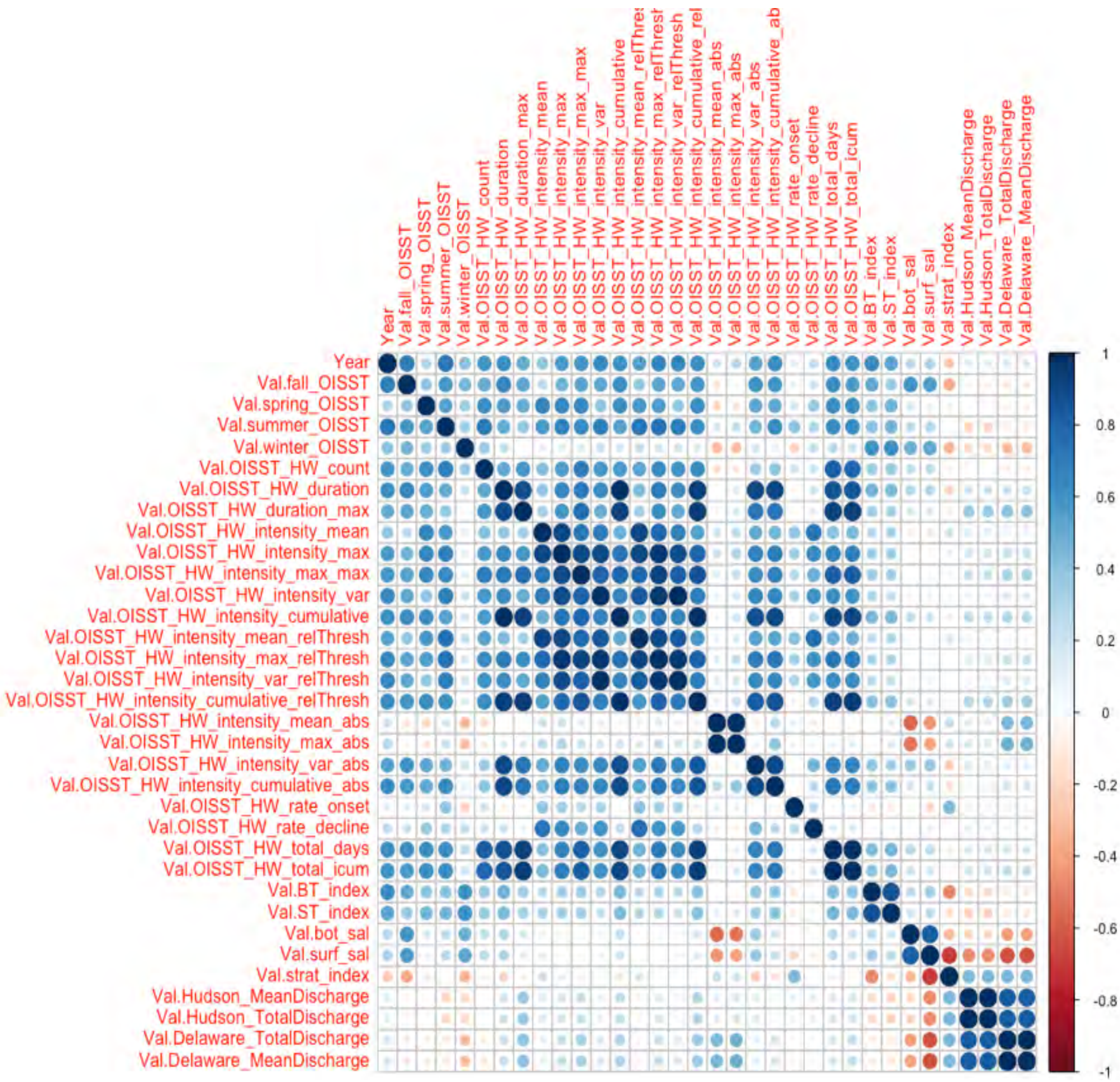


Figure 3.1.12. Correlation between indicators (positive = blue, negative = red). This figure is meant to only quickly show trends in the data and should be interpreted as preliminary and exploratory.

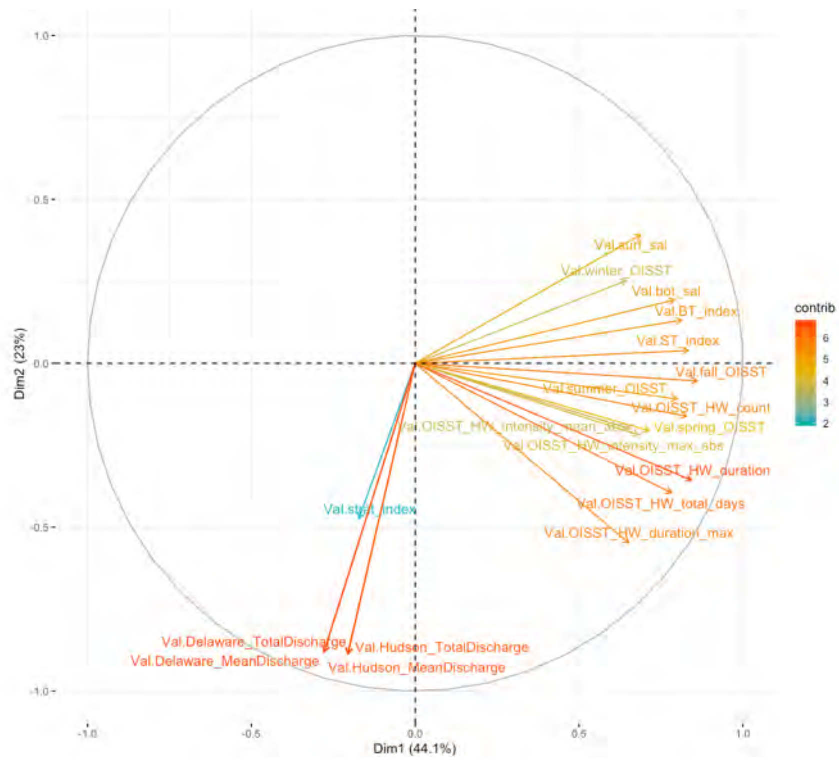


Figure 3.1.13. Principal component analysis (PCA) of indicators for the NYB (since 1990). To reduce clutter, some of the 20 Marine Heatwave indicators were excluded. This figure is meant to only quickly show trends in the data and should be interpreted as preliminary and exploratory.

<u>Other potential datasets and indicators to develop</u>

Buoy data

The buoy data from NOAA are collected at six locations offshore in the NY Bight Region (**Table 3.1.6**). Overall, these data are very easy to acquire because they can be imported directly into the R working directory with the package "rnoaa". The data are collected at fine temporal resolution (1 measurement per hour, or every several minutes) and can be used to estimate parameters representing seasonality (onset of summer, length of summer, etc.). However, there are two primary limitations. First, the data for many buoys does not go back very far. Second, in all of the data series there are a substantial number of missing observations that would need to be dealt with in the calculation of a consistent temperature indicator (e.g., via imputation).

An example of buoy data is shown in **Figure 3.1.13** which depicts temperature at the buoy offshore of Islip NY.

Additional variables are available from buoys (salinity, wave height and direction, etc.) that were not explored in depth but could potentially be used in development of indicators.

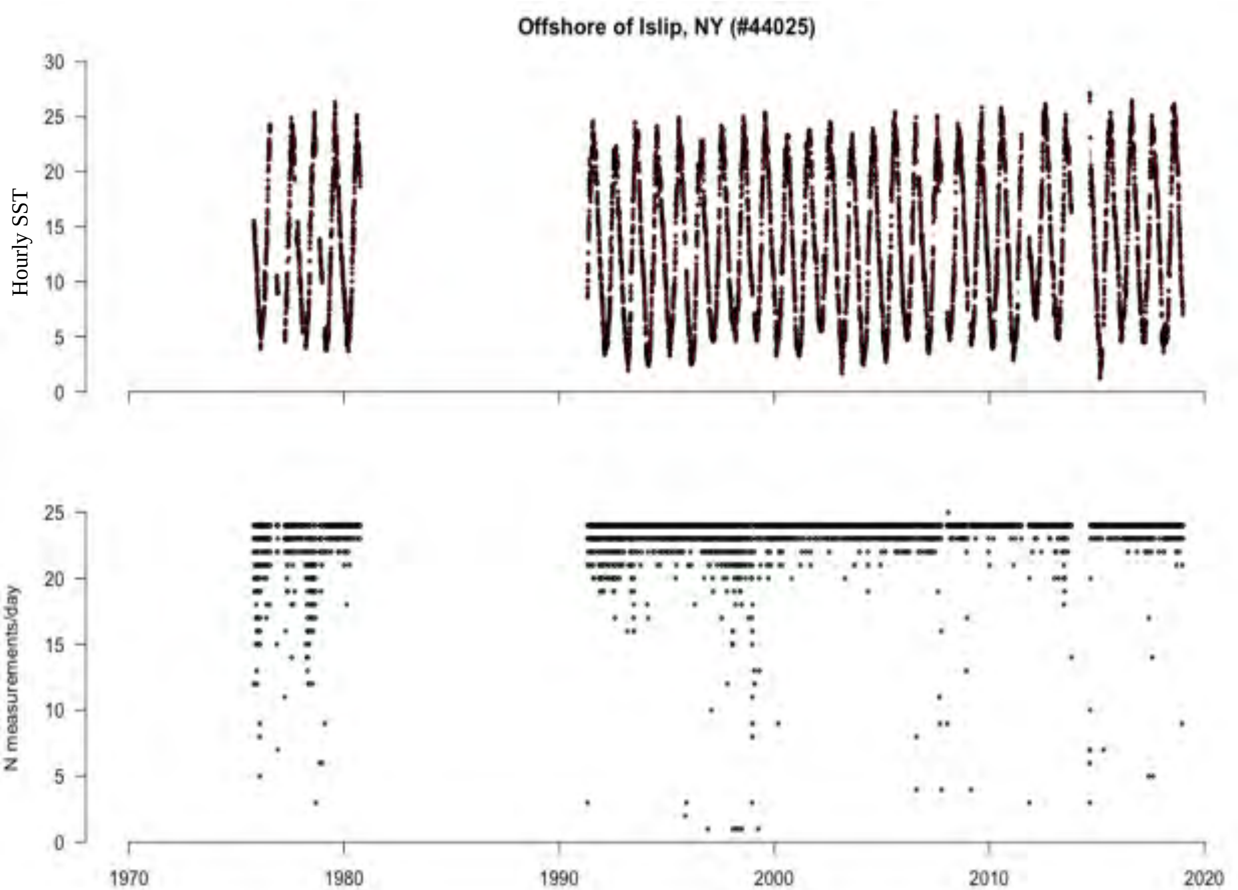


Figure. 3.1.14. Buoy observations from Buoy 44025, offshore of Islip NY. Top panel shows temperature measurements; bottom panel shows the number of measurements per day (ideally 24, but often far less which indicates missing data). Use of this time series would require dealing with missing values.

Table 1.1.6. Buoys in or near the NYB region.

Name	Location	Lat	Long	ID	Start	End
Barnegat, NJ	Offshore	39.778	73.769	44091	2008	2018
Texas Tower	Offshore (by Hudson Canyon)	39.618	72.644	44066	2009	2018
Sandy Hook, NJ	Inshore	40.467	74.009	SDHN4	2004	2018
New York Harbor Entrance	Offshore, 15 miles from Breezy Point	40.369	73.703	44065	2008	2018
Long Island	Offshore, 30 miles from Islip	40.251	73.164	44025	1975	2018
Montauk	Offshore	40.693	72.049	44017	2002	2018
Block Island	Offshore	40.969	71.127	44097	2009	2018

Contaminants/nutrients

The WOD data lists several datasets that could potentially be used to generate an index of nutrients or contaminants; these were not explored in depth. There are 20,875 Ocean Station Bottle (OSD) casts in the NYB over the period of record in the dataset (**Table 3.1.1**) and these will consist of a variety of data types (see **Table 1.1.2**) such as Phospate, Silicate, Nitrate, pH, Alkalinity, etc. The number 20,857 does not imply that each cast has each of these parameters measured; yet, the data could be further explored given a particular metric of interest.

The 2016 'indicators workshop' document also mentions a New York Bight Water Quality Monitoring Program by the EPA that has not been explored in depth.

Seasonality of warming

Indicators representing shifting seasonality could be calculated with the OISST data or buoy data. These include the start of spring, duration of summer, or similar metrics based on prior studies (Henderson et al. 2017; Staudinger et al. 2019). We began to code these but have not finished calculations. Previous work indicates that the timing of spring warming has been highly variable since the 1980s, but has occurred earlier in the last decade (**Figure 3.1.15**). The more striking finding is that fall cooling occurs later and there is a strong change over time. This is corroborated by analysis of air temperature showing that the frost-free season, or growing season, was 10 days longer in the region over the 1986-2015 period compared to 1901-1960 (Mark Lowery NYSDEC https://www.newsday.com/long-island/environment/long-island-climate-changes-1.38957682)

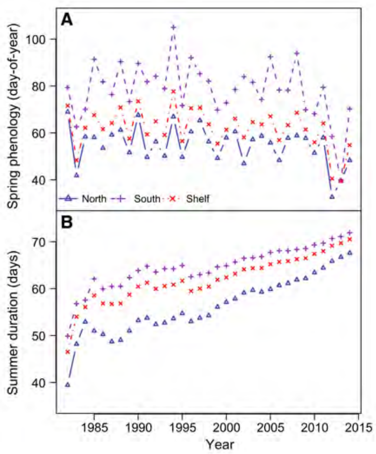


Figure 3.1.15. Trends in spring warming and duration of summer from Henderson et al. 2016.

Habitat

There were many suggestions for habitat quantity or quality indicators in the 2016 indicators workshop. For the most part these were not explored, but a measure of cold pool extent or volume may be possible (see below).

Ocean Acidification datasets

The *Ocean Carbon Data System website (OCADS)* has a data portal for ocean carbon and acidification data. Of particular interest are time-series and moorings, yet there are none listed that occur within the NYB. There are moorings with continuous pCo2 in Martha's Vineyard and the Gulf of Maine. Overall, this website does not appear to be very user friendly as it ‘hosts’ many individual datasets that would need to be independently tied together.

https://www.nodc.noaa.gov/ocads/

The *Surface Ocean Co2 Atlas (SOCAT***)** dataset is promising and ties together a variety of quality-controlled surface ocean Co2 measurements. Following the global trend, the number of pCO2 measurements in NYB has increased over time (**Figure 3.1.16, 3.1.17**). The issue with this data source is that the measurements are absent or sparse until the 2000s (and probably late 2000s) and the spatial and seasonal coverage is inconsistent until recently. We are considering how we might be able to use this data source. A benefit however, is that (similar to the WOD dataset) data are submitted by various contributors and then linked together so they can be downloaded as single files.

https://www.socat.info/index.php/data-access/

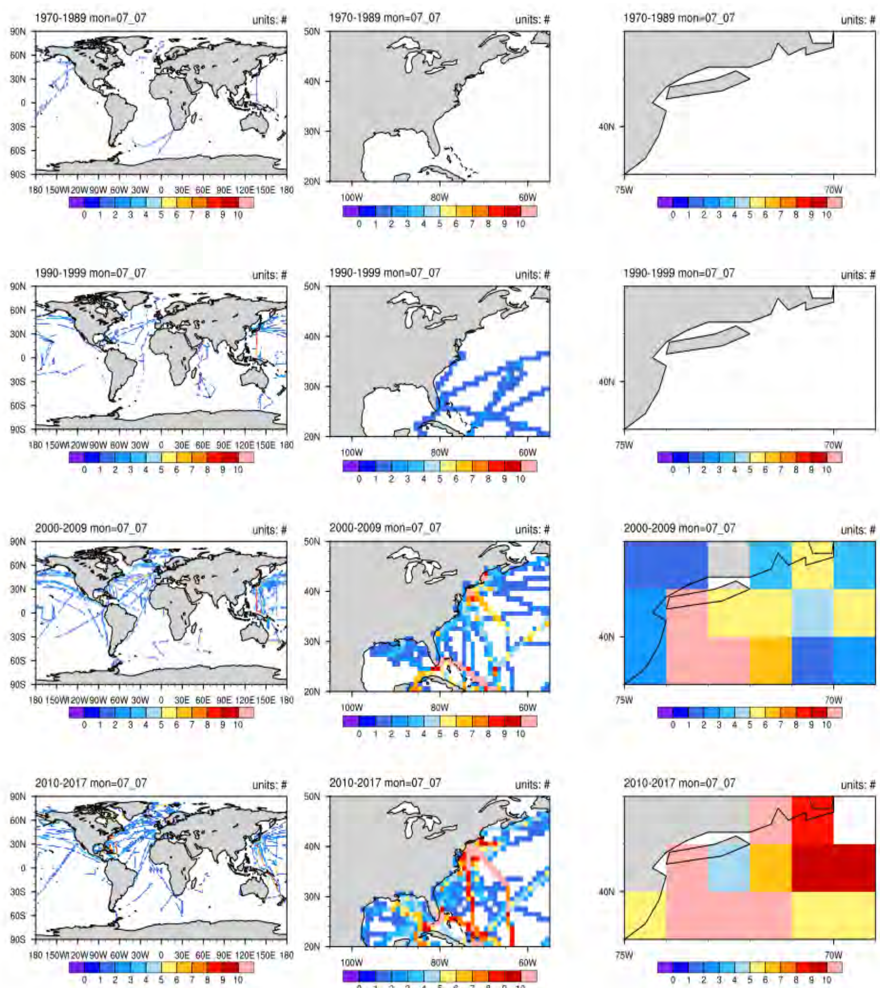


Figure 3.1.16. Number of surface pCO2 measurements in January over time in NYB. Color bar indicates the number of observations, which has increased from 1970-1989, 1990-1999, 2000-2009, 2010-2017.

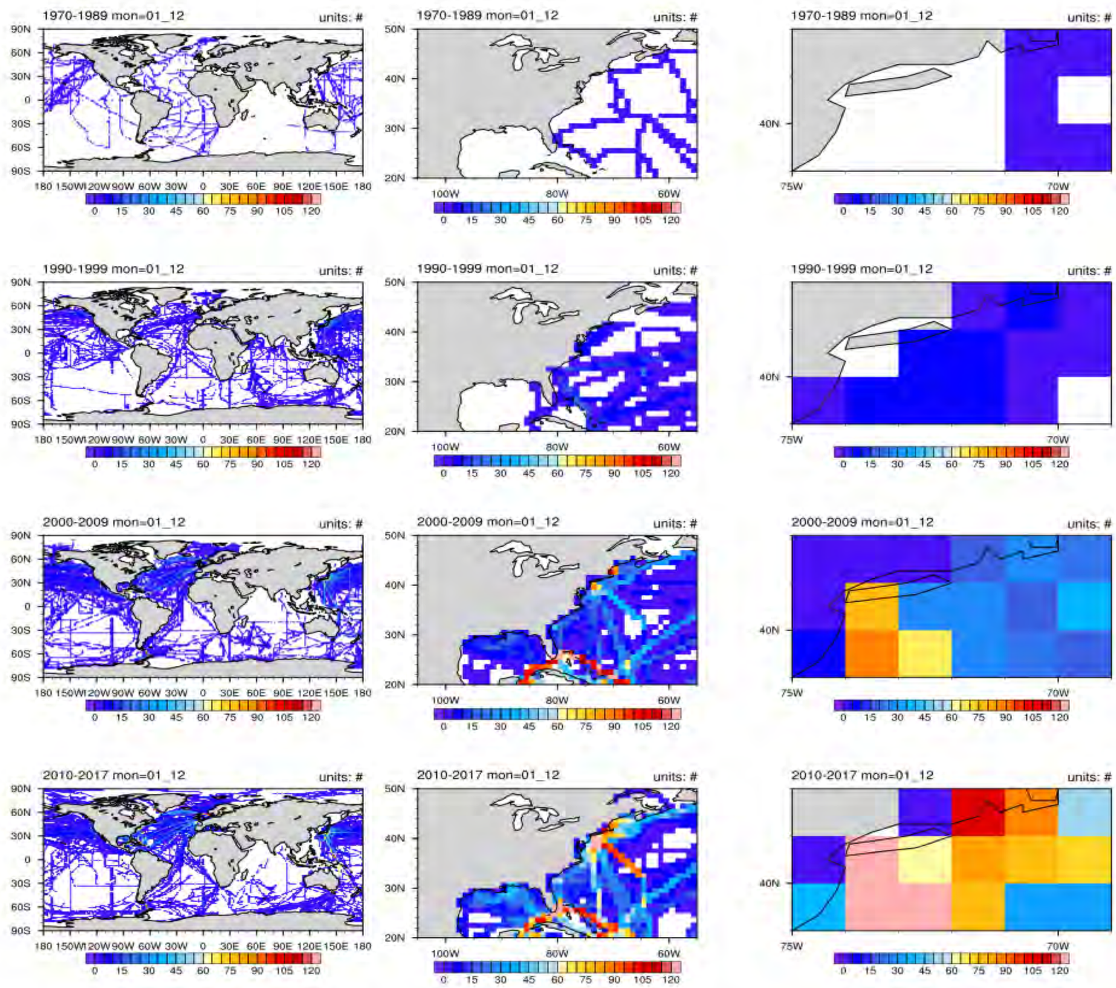


Figure 3.1.17. Number of surface pCO2 measurements in July over time in NYB. Color bar indicates the number of observations, which has increased from 1970-1989, 1990-1999, 2000-2009, 2010-2017.

Cold pool dynamics

Of interest is generating a time series to represent annual variability in the cold pool; this is a task that would require substantial work to synthesize existing data and/or incorporate new data collection efforts.

Warm core ring frequency in the NYB

Marine heatwaves could be caused by a variety of drivers; one such driver is the presence of warm core rings (WCRs). These are circulating rings of warm water that break off from the Gulf Stream. These are increasing from an average of 18 per year in 1980 to 1999, to more than 30 from 2000 to 2017 in the Mid-Atlantic region (**Figure 3.1.18**, Gangopadhyay et al. 2019).

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

Quantifying the occurrence of WCRs that enter the NYB would also be an interesting indicator, but probably redundant with MHW and SST metrics.

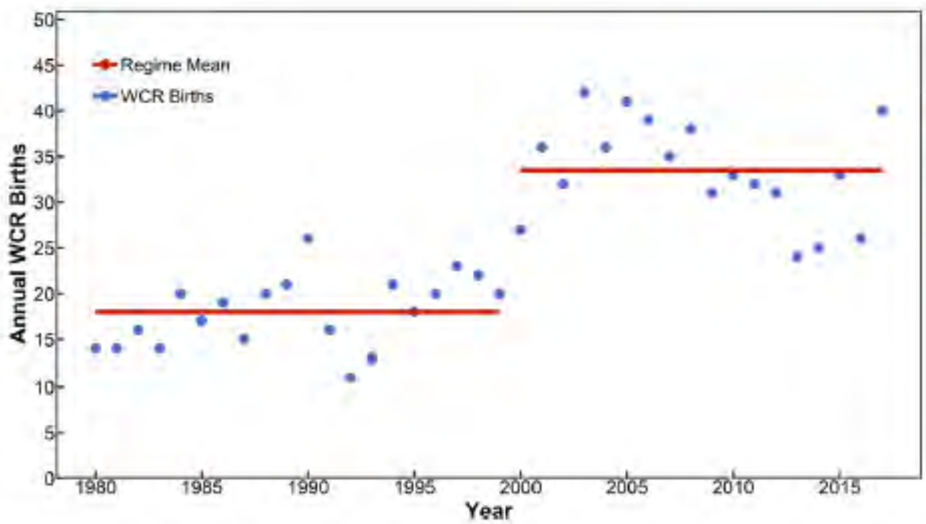


Figure 3.1.18. Frequency of warm core ring births off the Gulf Stream since 1980 in the Mid-Atlantic. Figure taken from Gangopadhyay et al. (2018).

Improved index methods

We used the WOD dataset to calculate indices of physical variables (e.g., stratification, surface salinity, bottom temperature) by fitting a GAM model to available in situ measurements. These models account for systematic variation in the measurements related to date of sampling, longitude, latitude, depth, and other variables and a fixed effect of year is used as the index over time. This method appears adequate (**Figure 3.1.8**), but could be improved. This process is essentially the same as catch standardization in fisheries (i.e., generate an index of abundance from survey data), and could potentially benefit from novel approaches in that field. Vector autoregressive spatio-temporal (VAST) models could be used for this and it would be a useful project to fit VAST models to these data and compare to other methods, like GAMs (Thorson 2019).

3.2: Social, economic, recreational, and commercial indicators

Summary

- The Marine Recreational Information Program (MRIP), Economics: National Ocean Watch (ENOW), and NOAA commercial landings data are useful and easy to access
- 16 indicators are presented; more could be developed with the acquired datasets based on DEC needs or interests
- Total recreational harvest of fish in NY is decreasing, but the number of fish caught and released (C&R) is increasing
- The MRIP dataset could provide information for anglers in NY State and could be a focus of additional work or studies (discussed)
- Confidential commercial landings data have not yet been accessed but would provide better detail than the public dataset
- Additional indicators are suggested including megawatt hours from wind energy, boat traffic from VMS, and others used by NOAA and LISS

Key datasets identified and data coverage

Population size

The National Cancer Institute (NCI) produces a data product for population size estimates that incorporates multiple sources of data from the census bureau. It is easier to work with than the data directly from Census Bureau. It is available at the county and state level and separates by age, sex, race. There are probably many other datasets that could be used to estimate population size, this is just one of them.

https://coast.noaa.gov/digitalcoast/data/demographictrends.html

Per capita income from CPS survey

There is a really nice dataset that is based on the Census Bureau survey called the "CPS" survey that is done every year. From it one can extract median household income.

Economics National Ocean Watch (ENOW) dataset

A promising dataset is the Economics: National Ocean Watch (ENOW) data source which summarizes ocean economics data. This dataset summarizes data from the Bureau of Economic Analysis (BEA) but specifically for ocean related profits and economic factors (by sector and industry at the state and county level). The sectors and industries are shown in **Table 3.2.1**. In 2016 there was a curious increase in many of the values. It appears that after 2016 a new 'industry' was added (wholesale seafood sales) and this makes the pre 2016 data not comparable to the data including and after 2016. This inconsistency would need to be addressed to continue to use this dataset.

Table 3.2.1. The Economics: National Ocean Watch (ENOW) dataset includes estimates of the number of establishments, employment, wages, GDP, and real GDP for each of the sectors and industries shown here. The columns with "not shown" indicate that the NAICS Industry subcategories are not shown for brevity.

Sector	Industry	NAICS Industry (2012 NAICS)
Living Resources	Fishing	Finfish Fishing
		Shellfish Fishing
		Other Marine Fishing
	Fish Hatcheries and Aquaculture	Finfish Farming and Harvesting
		Shellfish Farming
		Other Aquaculture
	Seafood Processing	Seafood Product Preparation and Packaging
	Seafood Markets	Fish and Seafood markets
		Fish and Seafood Merchant Wholesalers
Marine Construction	Marine Related Construction	Not shown
Marine Transportation	Deep Sea Freight	Not shown
	Marine Passenger Transportation	Not shown
	Marine Transportation Services	Not shown
	Search and Navigation Equipment	Not shown
	Warehousing	Not shown
Offshore Mineral Resources	Limestone, Sand and Gravel	Not shown
	Oil and Gas Exploration and Production	Not shown
Ship & Boat Building	Boat Building & Repair	Not shown
	Ship Building & Repair	Not shown
Tourism and Recreation	Boat Dealers	Not shown
	Eating and Drinking Places	Not shown
	Hotels and Lodging	Not shown
	Marinas	Not shown
	Recreational Vehicle Parks and Campsites	Not shown
	Scenic Water Tours	Not shown
	Sporting Goods	Not shown
	Amusement and Recreation Services	Not shown
	Zoos and Aquaria	Not shown

NOTE: For full table and information visit https://coast.noaa.gov/data/digitalcoast/pdf/enow-crosswalk-table.pdf

Marine Recreational Information Program data

The Marine Recreational Information Program data (MRIP) is the clear choice to develop indicators related to recreational fishing. This NOAA-led program uses extensive surveys and interviews to produce estimates of how many fish (by species) are caught, harvested, and released by recreational anglers. It is fairly easy to work with and access. A variety of statistics are estimated and available (**Table 3.2.2**) providing opportunities for many potential indicators (e.g., catch diversity, Shannon index, specific fish species, inshore vs. offshore catch) that could be explored for NY State.

Table 3.2.2. Variables included in the Marine Recreational Information Program (MRIP) dataset that could be used for indicators, including response variables and variables to query by or categorize data

Response variables	Variables to query by	Query levels
Effort	Fishing area (3 levels)	i.e., inshore, offshore…
Participation	Species (739 levels)	
Catch (harvest)	Fishing mode (7 levels)	i.e., shore, charter boat…
Catch (release)		
Fish Length		

Commercial fishing landings data

A commercial landings data set is available from NOAA Fisheries. This is a publicly available dataset that excludes all confidential information regarding where fish are caught (i.e., statistical fishing area). The data are summarized by where fish are landed, not where they are caught, so the dataset is not necessarily ideal. However, it is very easy to access and should be available without a doubt for the foreseeable future.

NYB indicator overview

A total of 16 indicators are calculated; however, many more could be generated with the acquired datasets. For example, there are several hundred species included in the MRIP dataset and recreational catch estimates could be generated for any one of them or any combination of them. It would also be simple to calculate indicators of effort (i.e., days fished, # of fishing trips) and this might be interesting to combine with human population size estimates for costal NY, or median household income. For example, estimated number of MRIP trips might make more sense to be expressed as a 'per capita' rate rather than an absolute rate to convey overall interest in recreational fishing in NY as the population size is growing.

Table 3.2.3. Indicators calculated for the NYB (all are done for the State of NY, but could are also calculated for counties in some cases). Note that there are fewer non-linear trends than linear trends, and sometimes linear and non-linear trends show different results (i.e., rec stripers harvest, rec stripers harvest boat, rec harvest stripers shore).

Indicator	Linear trend	Non-linear trend (since 2010)	Linear r-squared
Population size	↑	↑	0.78
Per capita income	↑	↑	0.96
LMR employment	↑	↑	0.70
LMR establishments	↑	↑	0.51
LMR GDP	↑	↑	0.59
LMR GDP percent			0.21
Rec Harvest	↓		0.35
Rec Release	↑	↑	0.66
Rec Bluefish harvest	↓		0.47
Rec Bluefish release	↑		0.23
Rec Stripers harvest	↑	↓	0.53
Rec Stripers release	↑		0.62
Striper harvest shore	↑	↓	0.19
Striper harvest boat	↑	↓	0.61
Bluefish harvest shore	↓		0.32
Bluefish harvest boat	↓		0.41
Commerical Landings KG	↓		0.53
Commerical Landings Value	↓		0.24

Table 3.2.4. Social, economic, and recreational indicators calculated for the state of NY.

Indicator	Description
Population size	Total population size
Per capita income	Median per capita income
LMR employment	# people employed in Living Marine Resources
LMR establishments	# living marine resource establishments
LMR GDP	# GDP from living marine resources
LMR GDP percent	% of GDP that is from living marine resources
Rec Harvest	Recreational harvest all species
Rec Release	Recreational release all species
Rec Bluefish harvest	Total recreational harvest bluefish
Rec Bluefish release	Total recreational release of bluefish
Rec Stripers harvest	Total recreational harvest of bluefish
Rec Stripers release	Total recreation release of bluefish
Striper harvest shore	Shore based harvest of Stripers in NY
Striper harvest boat	Private boat-based harvest of Stripers in NY
Bluefish harvest shore	Shore based harvest of Bluefish in NY
Bluefish harvest boat	Private boat-based harvest of Bluefish in NY

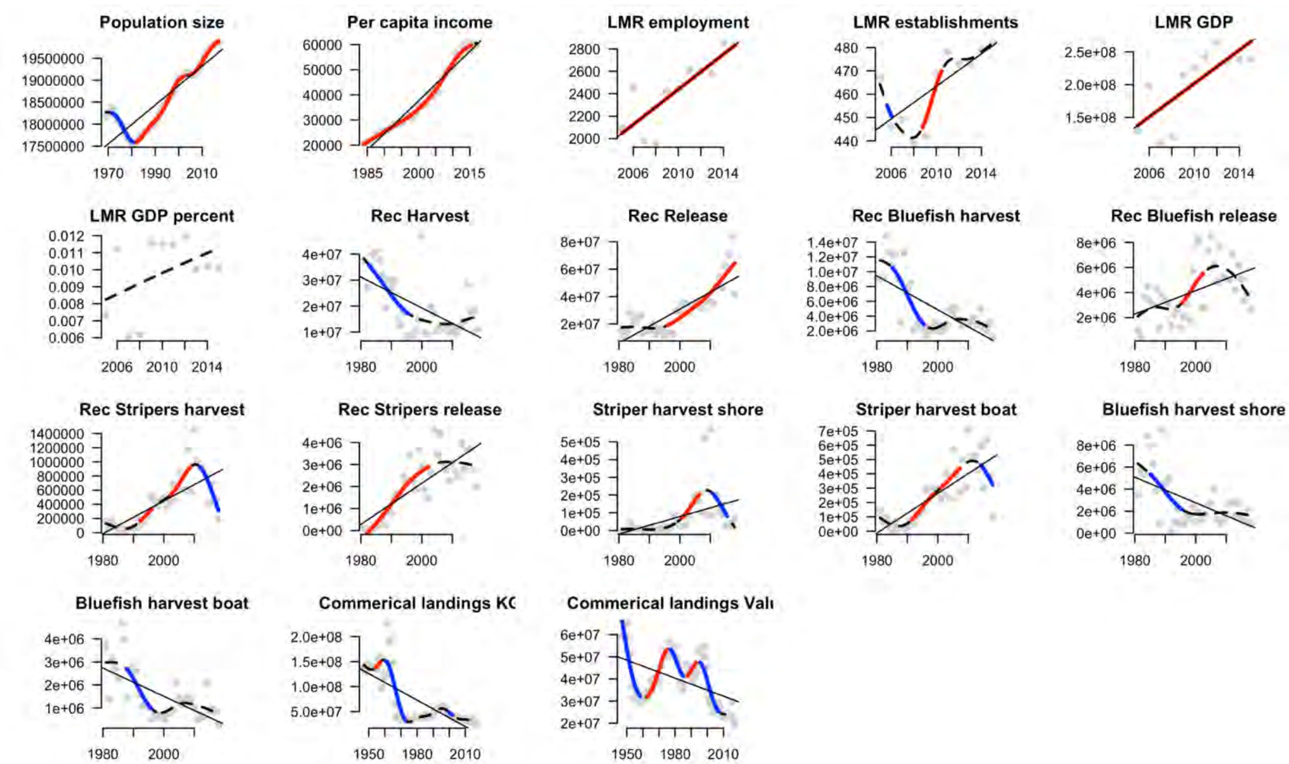


Figure 3.2.1. Time series of indicators for NY State. Solid black line is added when a significant linear trend was found; fitted GAM predictions are shown as a dashed line and filled in with red or blue when increases and decreases are statisticaly significant.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

Population size and income

Population size in NY has increased markedly over the last several decades to a most recent estimate of 19,849,399 based on the NCI dataset (**Figure 3.2.1**). Counties on Long Island show increasing but variable trends; Nassau has not grown much, but Suffolk and Bronx County have increased markedly (**Figure 3.2.2**). These figures only show examples of the dataset, there are similar estimates for all counties in NY, other states along the Atlantic Coast, and all counties across the Atlantic Coast.

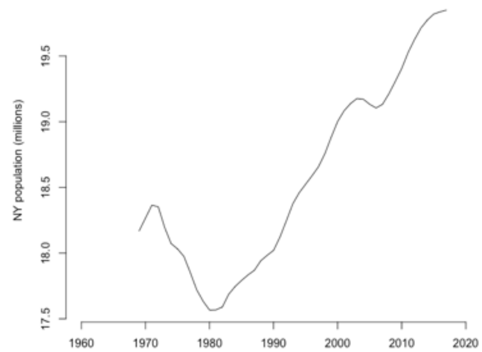


Figure 3.2.1. Human population in NY since 1970

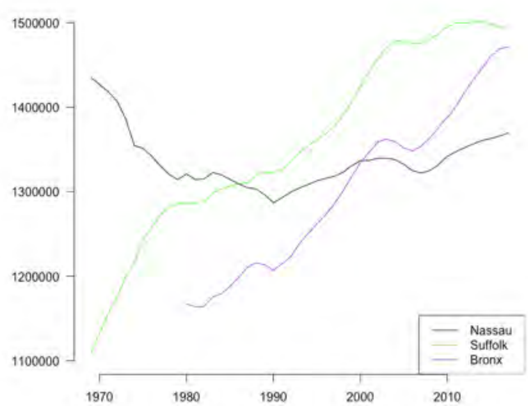


Figure 3.2.2. Human population in three counties on Long Island NY, since 1970.

ENOW datasets

To be consistent with (Link and Marshak 2019) we calculate four indicators including the number of establishments for the Living Resources sector, the number of people employed in the Living Resources sector, the real GDP for Living resources and the percent of real GDP for Living Resources/the GDP for all ocean sectors. Note however, that there are many other variables included in this dataset (**Table 3.2.1**). These are calculated at the entire NES ecosystem, States, and County scales. Trends are shown for the LMR statistics at only the NY

State level in **Figure 3.2.1**. Below are the results of the number of the LMR establishment and LMR employment over time broken down by all counties in NY state.

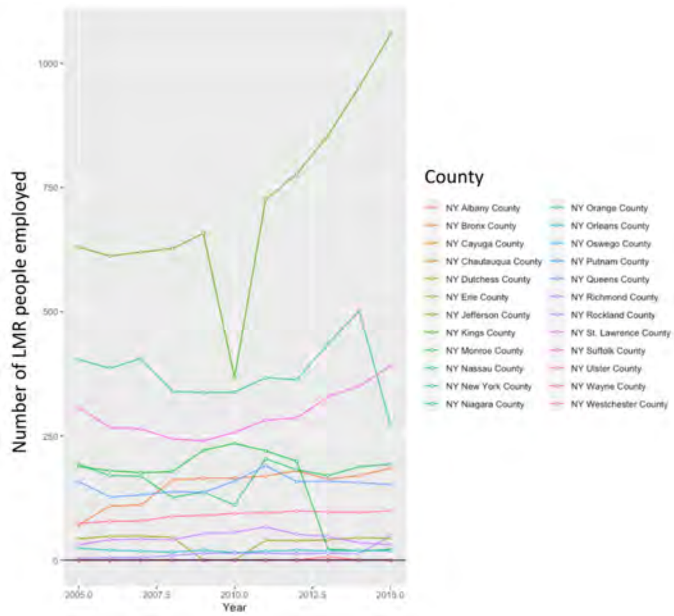


Figure 3.2.3. Trends in the number of people employed in the field of Living Marine Resources (LMR); for details of what type of jobs this includes see **Table 3.2.1.** The county with the highest overall number is Kings County. Some oddities should be investigated (i.e., dip in Kings County in 2010) prior to use.

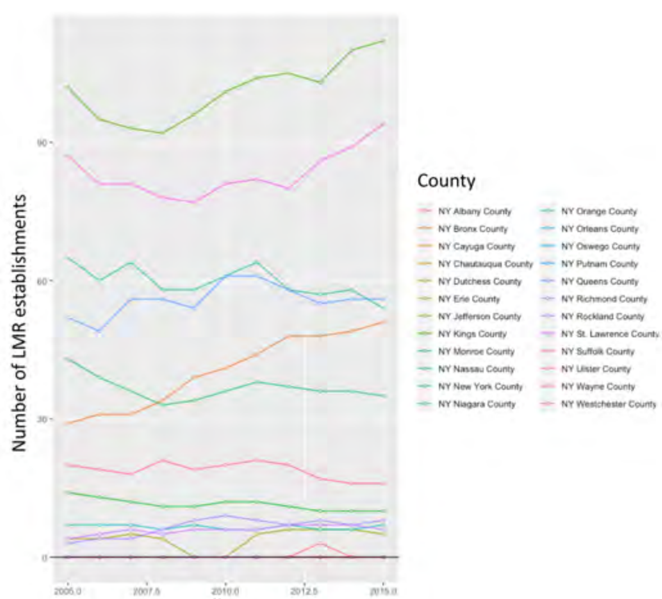


Figure 3.2.4. Trends in the number of Living Marine Resources (LMR) establishments; for details of what type of establishments this includes see **table 3.2.1.** The county with the highest overall number is Kings County and shows an increase in recent years.

Recreational fishing

Trends in recreational fishing indicators are highly nonlinear. For example, recreational harvest of Striped Bass has increased after 1990, peaked in about 2010 and has been rapidly declining since then. This is consistent with the most recent benchmark stock assessment in 2018 suggesting the stock is overfished and declining.

It is also interesting to view these data based on the mode of fishing (i.e., surf casting, personal boat/charter, party boat). For example, the estimates suggest that 2008 and 2010 were exceptional years for surf casting in NY state with estimates of nearly 600,000 fish harvested by land-based anglers (**Figure 3.2.6**). This is 700% higher than the long-term average estimate of 76,000. These data could be explored and used in many ways that are of interest to stakeholders.

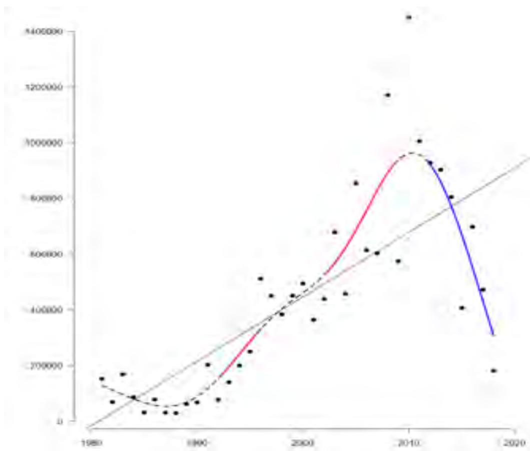


Figure 3.2.5. Total recreational harvest of Striped Bass in NY from MRIP data with a linear regression and fitted trend by GAM.

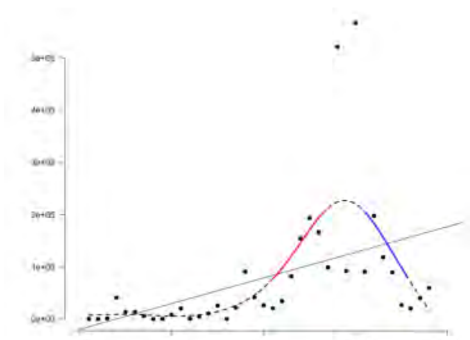


Figure 3.2.6. Recreational harvest of Striped Bass in NY by shore based anglers from MRIP data with a linear regression and fitted trend by GAM. The two extreme points are from 2008 and 2010.

The trends for Bluefish also demonstrate the importance of viewing the data with non-linear techniques. Bluefish harvest shows a negative linear trend, but the non-linear GAM reveals that the decrease mainly occurred from 1980 to 1990 and has been largely stable (i.e., no trend in the GAM) since then (**Figure 3.2.7**).

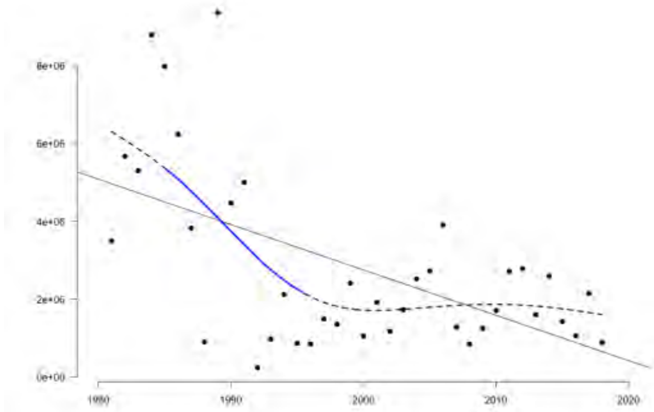


Figure 3.2.7. Trends in recreational harvest of Bluefish in NY state.

Commercial fishing

Commercial landings (total, all species) are shown in **Figure 3.2.1**. This dataset is not ideal for use as an indicator but is publicly available.

Relationships between variables and redundancy

Relationships between variables are shown in **figure 3.2.7**.

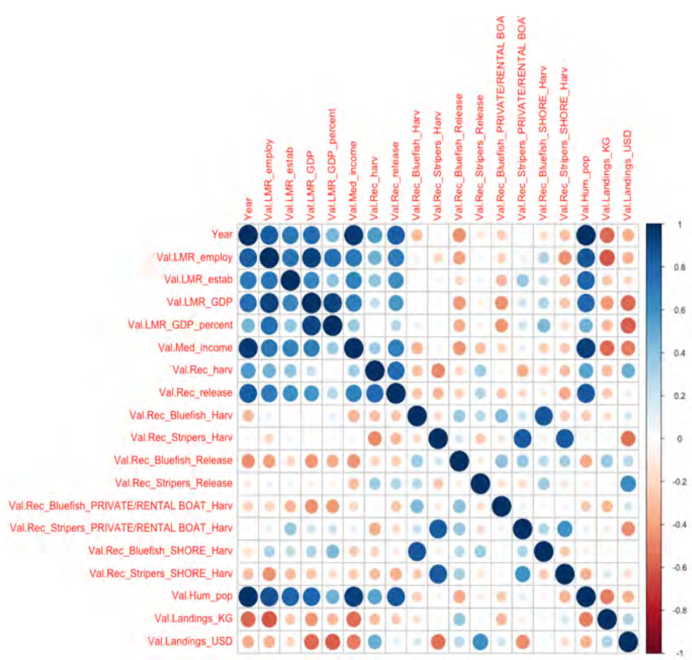


Figure 3.2.7. Correlation among variables.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

Other potential datasets and indicators to develop

Alternative commercial landings data

There are limitations to using the publicly available commercial landings data; for example, the website where the data are accessed states:

"Total landings by state include confidential data and will be accurate, but landings reported by individual species may, in some instances, be misleading due to data confidentiality."

Discussions have begun (K.H. and DEC) to acquire confidential commercial landings data, that could either be done through the DEC or by contacting NOAA for the statewide commercial landings' dataset. These could be used to better understand spatial variation in commercial fishing pressure on individual species in the NYB.

NOAA social indicators of fishing community vulnerability and resilience

A fairly recent effort by NOAA (led by Lisa Colburn) has integrated multiple data sources to develop indicators for each coastal community along the NE coast.

For example, the entire dataset for "commercial reliance and engagement" (**Figure 3.2.9**) and "recreational reliance and engagement" (**Figure 3.2.10**) are shown. These metrics incorporate a variety of information on commercial landings, permits, commercial dealers and are spatially explicit to show how different communities depend on living marine resources. The dataset also includes a 'vulnerability to sea level rise" index that used a digital elevation model (DEM) to calculate the land area of each community between one and six feet above sea level; the categorical risk rankings are shown in **Figure 3.2.11**. Lastly, **Table 3.2.4** shows the rankings for all included variables; for more information on this dataset and the methods and meaning of each indicator see the following websites.

Information on methods:

https://www.fisheries.noaa.gov/national/socioeconomics/social-indicator-supporting-information#data-series

Online map viewer of indicators and data download site:

https://www.st.nmfs.noaa.gov/data-and-tools/social-indicators/

Publication describing development of methods:

Colburn, L.L., Jepson, M., Weng, C., Seara, T., Weiss, J., and Hare, J.A. 2016. Indicators of climate change and social vulnerability in fishing dependent communities along the Eastern and Gulf Coasts of the United States. Mar. Policy **74**: 323–333. Elsevier. doi:10.1016/j.marpol.2016.04.030.

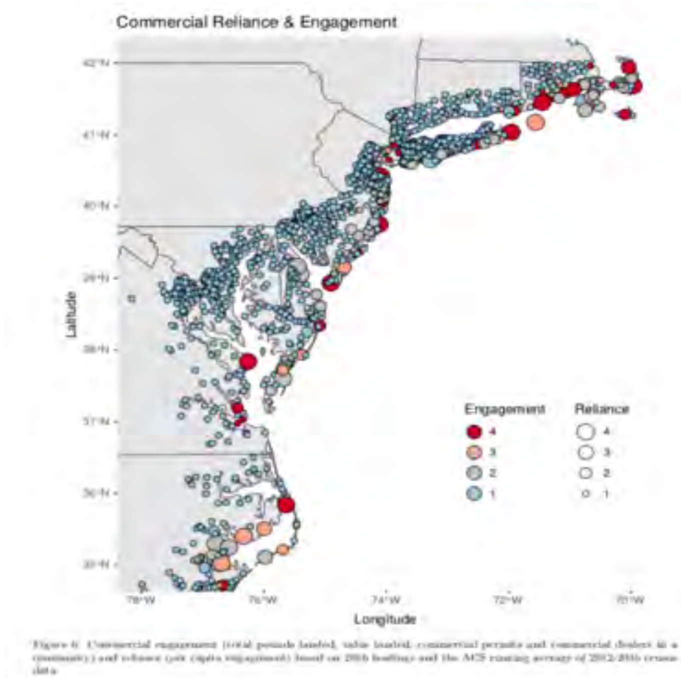


Figure 3.2.9. Directly quoted from NOAA "Commercial engagement (total pounds landed, value landed, commercial permits and commercial dealers in a community) and reliance (per capita engagement) based on 2016 landings and the ACS running average of 2012 – 2016 census data."

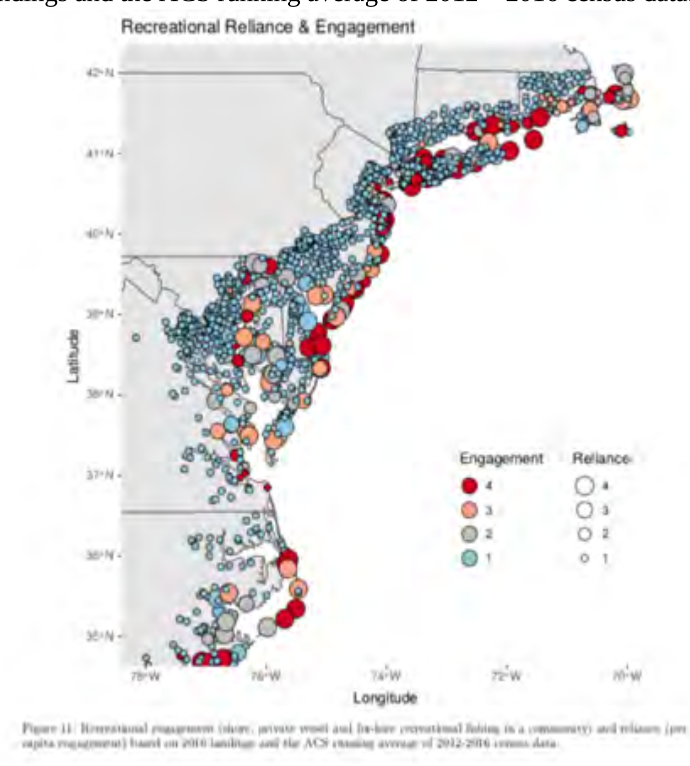


Figure 3.2.10. Directly quoted from NOAA "Recreational engagement (shore, private vessel and for-hire recreational fishing in a community) and reliance (per capita engagement) based on 2016 landings and the ACS running average of 2012 – 2016 census data."

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

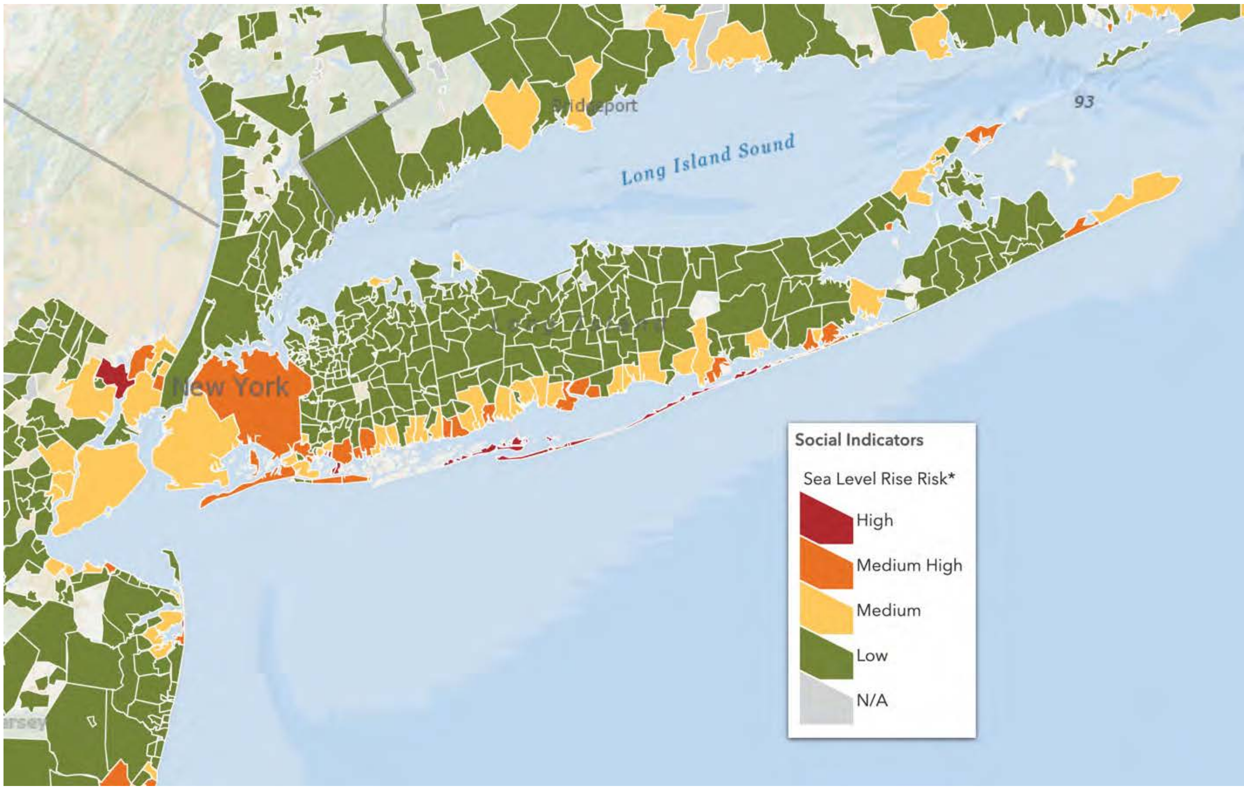


Figure 3.1.11. Indicator of sea level risk for Long Island communities.

Table 3.2.4. Selected communities in NY and the associated indicator rankings. All of the available indicators are shown (14) with values for 2016; similar data are available for 355 communities in NY with unique estimates for each year between 2009 and 2016. Meaning of rankings differ for each indicator but a "1" generally means low values and a "4" is the highest value.

Community Name	New York/Manhattan	Port Jefferson	Cutchogue	Mastic	Montauk	Orient	Riverhead
Poverty	3	1	1	2	1	1	2
Labor.Force	1	1	3	1	2	4	3
Housing.Characteristics	1	1	1	1	1	1	1
Population.Composition	2	1	1	2	1	1	2
Personal.Disruption	2	1	1	2	1	1	2
Housing.Disruption	1	3	4	4	4	4	4
Retiree.Migration	1	2	3	1	3	4	3
Urban.Sprawl	4	4	3	3	3	3	3
Commercial.Engagement	3	1	1	1	4	1	1
Commercial.Reliance	1	1	1	1	4	1	1
Recreational.Engagement	1	4	1	1	4	3	1
Recreational.Reliance	1	2	1	1	4	4	1
Sea.Level.Rise.Risk	1	1	1	1	2	3	1
Storm.Surge.Risk	2	1	1	1	1	2	1

Wind Farm Megawatts generated in NY Bight

Given the increasing prevalence and role of wind farms in generating electricity for NY state, it will be important to quantify this contribution. Indeed, one of the main goals of the OAP is to support sustainable development of wind energy (i.e. human use of the ocean, Goal #2). A potential indicator is ‘Megawatts (MW) of Wind Energy’ produced in NY Bight, or ‘MW based on permitted leases’ or something similar. Another possibility is to quantify the proportion of electricity for long island that is generated by wind. Datasets to calculate these have not been explored.

Information from NY Marine Registry

For the interested public, it would be useful to present an annual count of people in the recreational marine fishing registry. Although compliance with this regulation (enacted in 2011) is unknown, stakeholders would be interested in seeing these data. A personal perspective (KH) is that anglers like to see the results of programs that they go out of their way to participate in.

Boat traffic/commercial boat traffic

Vessel monitoring system (VMS) data could be used to develop an annual index of vessel activity in the NYB or regions of the NYB. As an indicator, this could be useful to quantify the increasing risk to cetaceans or as a driver of interannual variation in ‘boat strikes’ (see part 3.4). This would also serve as an indicator of ‘human benefit’ provided by the ocean, as higher values would indicate an increasing role in marine transportation, recreation, and commercial fishing.

Social indicators from LISS applied to NY oceanfront

The LISS “Ecosystem targets and supporting indicators” system provides some additional examples (and associated datasets) that could be used for indicators in the NY Bight. These are listed below, data sources for these are listed at the LISS website

- Number of volunteers participating in beach cleanups: could be done for NY Ocean Front Beaches, currently calculated for LIS in NY
- Pounds of debris collected per mile of beach: uses data provide by American Littoral society to quantity ‘trash on the beach’
- Number of beach closures per year: Uses EPA BEACON 2.0 dataset (Beach advisor and closing on-line notification). Could be calculated for NY Ocean Front Beaches
    - https://watersgeo.epa.gov/beacon2/

Deeper exploration of MRIP dataset

The MRIP dataset is a fascinating dataset from the perspective of a recreational angler and could be explored in more detail to summarize information of interest for NY fisherman. For example, it is interesting to see which species were the “top 6 fish” in terms of number of landed fish over time (**Figure 3.2.11**). Another interesting topic to explore is the ratio of recreational harvest and mortality to commercial. This is a topic often discussed by recreational anglers (i.e., contribution of commercial to recreational mortality for species of interest). Finger pointing is a common discussion point on online fishing forums (i.e., stripersonline.com, stripers247, bassbarn.com, stripersurf.com), this indicator would be of interest no doubt.

Furthermore, from a research perspective, summarizing these datasets for the North East Atlantic could be an important research contribution. For example a recent paper describes the changes in the MRIP estimates for the Southeast United States Atlantic Ocean (Shertzer et al. 2019), and found that the recreational sector accounts for 76% of sportfish mortality. This research is a simple summary of MRIP estimates, and could be done in a similar way for the North East.

In particular, it would be interesting to determine the contribution of mortality by fishing mode (shore, private boat, party boat) and how this proportion may change depending on economic trends (i.e., income, demography, etc.) or environmental factors that would expose more migratory fish to capture in near-shore environments.

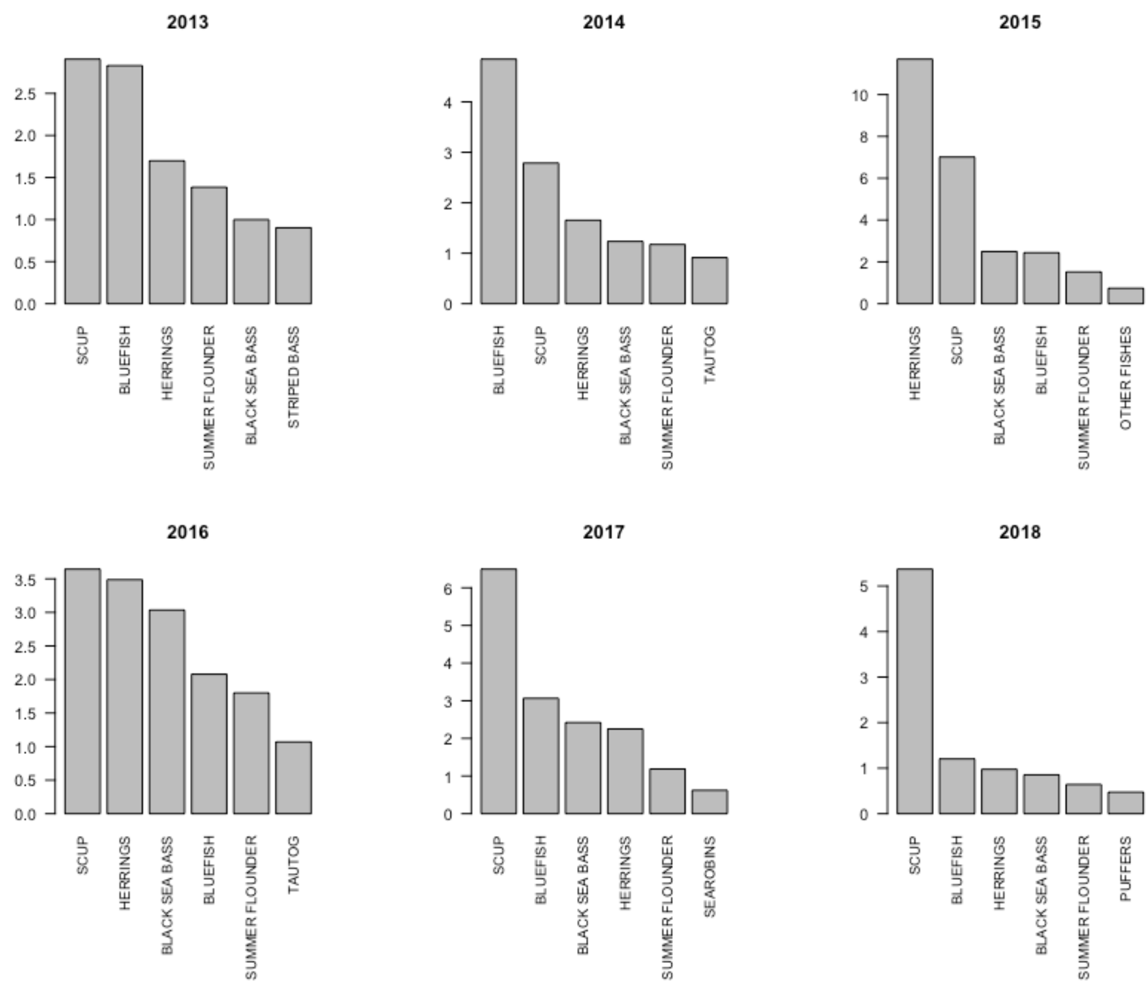


Figure 3.2.11. Top six species of fish for the last six years, by number landed by the recreational sector. Shown on the Y axis is millions of fish.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

3.3: Phytoplankton and zooplankton

Summary

- Satellite ocean color data provides continuous estimates of Chlorophyll-a; there are several datasets to choose from
- The OC-CCI merged product is used here
- NOAAs MARMAP/EcoMon survey provides good coverage of zooplankton in NYB but has caveats (limited summer sampling, processing lag time)
- Chlorophyll-a concentration began decreasing around 2010; zooplankton trends are variable but an overall index of volume in the NYB is decreasing
- The small to large copepod index indicates increases in small-bodied zooplankton, and a decrease in large copepods (*Calanus finmarchicus*) since ~ 2005
- Other indicators could use EcoMON data to explore variability in larval fish abundance
- The Continuous Plankton Recorder and COPEPOD website provide additional data

Key datasets identified and data coverage

Satellite ocean color data

Chlorophyll-a concentration on the ocean surface waters can be estimated from satellite images, and several data products are freely available. Data come from satellite 'missions' that last for a given duration of time before the satellite is decommissioned; therefore time-series are not continuous. For example, the SeaWiFS data range from 1997–2012, the MODIS range from 2002 – present, and VIIRS from 2012 – present. Given the complexity of attempting to 'blend' these datasets manually, we chose to explore a recently merged data product (the OC-CCI), which is the Ocean Color (OC) data set from the Climate Change Initiative (CCI). This dataset blends MERIS, Aqua-MODIS, SeaWiFS, and VIRS to produce a single time series.

The data set has not been thoroughly vetted for accuracy but should, in theory be more robust than any of the individual datasets used alone or in combination by simply adding the series to form a single one (*personal communication*, Kimberly Hyde).

NOAA EcoMON

NOAA conducts a regular survey called the Ecosystem Monitoring survey (EcoMON) that began in 1992, and is the longest continuous program at the NEFSC. Prior to 1992, another survey titled the Marine Resources Monitoring, Assessment and Prediction (MARMAP) survey collected plankton data. The EcoMON survey collects plankton data six times a year in the NES. The data are collected at about 120 randomly selected stations and 35 fixed stations through the NES. Information on sample sizes, spatial and temporal coverage of this survey are depicted in **figures 3.3.1, 3.3.2**, each figure showing only those data from the NYB. The data are available with a 2–3 year lag time, and fewer samples are collected in the summer than other months. About 50 samples are collected in the NYB annually.

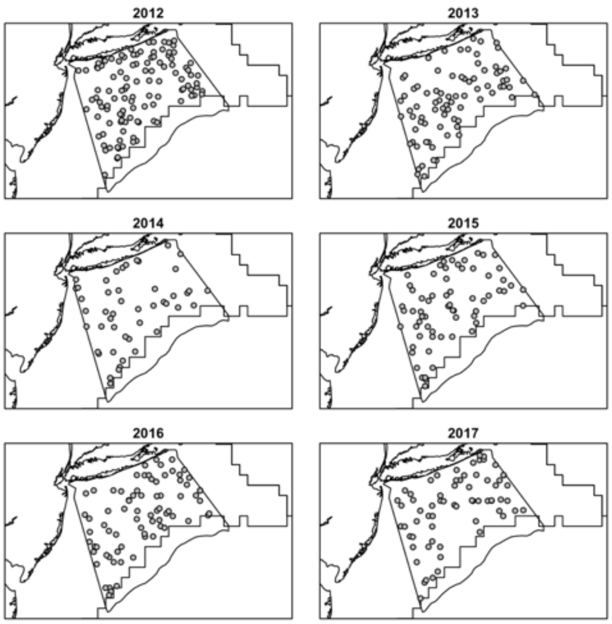


Figure 3.3.1. Spatial distribution of EcoMON samples in last six years of recent dataset.

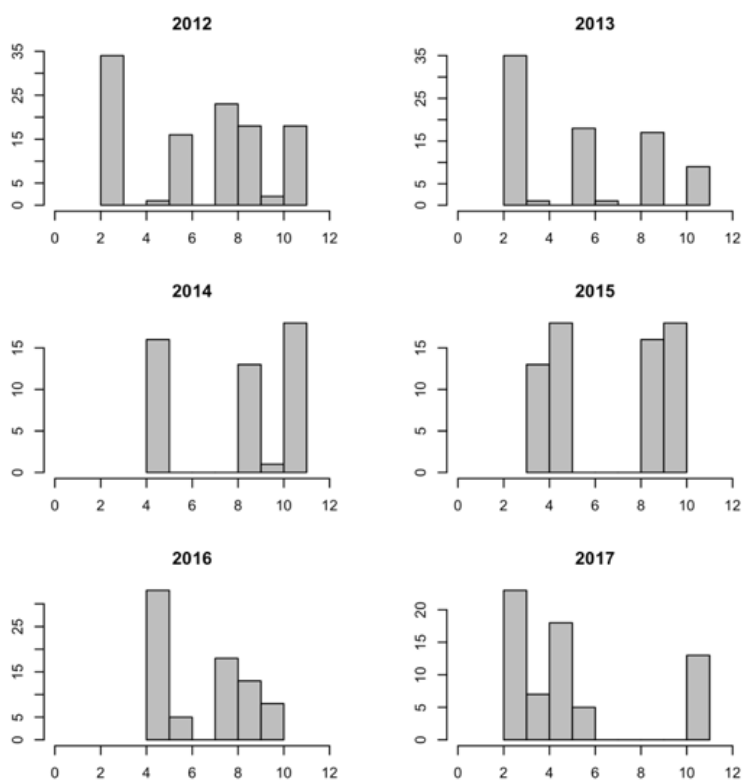


Figure 3.3.2. Temporal distribution of EcoMON samples in the NYB (1 = January, 2 = February…). Note the general lower amount of sampling occurring in the summer months (June, July August) in most years.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

NYB indicator overview

We calculated one indicator to represent chlorophyll-a concentration, and 29 indicators of zooplankton or ichthyoplankton abundance (**Table 3.3.1**). Calculations were done by subsetting datasets to all data available for the NYB and following methods validated and used by NOAA (for EcoMON) or simply averaging the Chl-a data by year. The method used for the EcoMON data generates an anomaly value for zooplankton abundance for each year and accounts for seasonal variation in sampling timing. Briefly, the entire time series is used (for a given taxa) to generate an 'expected' value for each day of the year; individual samples are then compared to this expected value to calculate the deviation from the expectation. All of these deviations are averaged to get the annual estimate. The 'small to large' copepod index compares several highly abundant small bodied taxa to *Calanus finmarchicus* (a big one). High values indicate a higher ratio of small bodied copepods. The time series for the top 29 most abundant taxa are shown (i.e., the most abundant taxa in the entire time series for the NYB). There are several hundred taxa included in the EcoMoN dataset that could also be explored, including roughly 60 larval fish species categories. The abbreviations used in figures are shown (**Table 3.3.2**). Time series with linear and non-linear trends are shown in **Table 3.3.2**. Lengthy discussion is not provided; however, two noteworthy trends are that (1) chlorophyll a appears to be decreasing in recent times as does (2) the total volume of zooplankton in ring net tows. These two trends are shown on the same plot in **Figure 3.3.3**, all trends are shown in **Tigures 3.3.4** and **3.3.5**.

Table 3.3.1. Trends in zooplankton fit with linear regression and GAMs, GAMs indicate if trends are significantly increasing or decreasing since 2010.

Indicator	Linear trend	Non-linear trend (since 2010)	Linear r-squared
Chlorophyll-a		↓	0.015
Zooplankton volume		↓	0.004
Centropages typicus			0.002
Calanus finmarchicus		↓	0.001
Pseudocalanus spp.	↓	↓	0.41
Penilia spp.			0.067
Temora longicornis			0.003
Centropages hamatus			0.054
Echinodermata	↑	↑	0.536
appendicularians	↑	↑	0.174
Paracalanus parvus	↑		0.231
Gastropoda	↑		0.251
Acartia spp.	↓		0.1
Metridia lucens		↓	0.011
Evadne spp.	↑	↑	0.291
Salpa	↑	↑	0.159
Oithona spp.	↑		0.427
Chaetognatha	↑		0.18
Hyperiidea	↑		0.417
Calanus minor	↑	↑	0.398
Clausocalanus arcuicornis	↑		0.331
Decapoda			0.073
Euphausiacea	↑		0.129
Protozoa	↑		0.211
Polychaeta	↑	↑	0.127
Pisces			0.004
Oncaea spp.	↑	↑	0.235
Corycaeidae	↑	↑	0.22
siphonophores	↑		0.228
Small to large copepod index		↑	0.002

Table 3.3.2. Indicator abbreviations and description; indicators are calculated anomaly values based on seasonal expectations for that taxa's catch over the duration of the time-series.

Indicator	Description
Chlorophyll-a	From OC-CCI blended product
volume	zooplankton displacement volume
ctyp	Centropages typicus
calfin	Calanus finmarchicus
pseudo	Pseudocalanus spp.
penilia	Penilia spp.
tlong	Temora longicornis
cham	Centropages hamatus
echino	Echinodermata
larvaceans	appendicularians
para	Paracalanus parvus
gas	Gastropoda
acarspp	Acartia spp.
mlucens	Metridia lucens
evadnespp	Evadne spp.
salps	Salpa
oithspp	Oithona spp.
chaeto	Chaetognatha
hyper	Hyperiidea
calminor	Calanus minor
clauso	Clausocalanus arcuicornis
dec	Decapoda
euph	Euphausiacea
prot	Protozoa
poly	Polychaeta
fish	Pisces
oncaea	Oncaea spp.
cory	Corycaeidae
siph	siphonophores
sm_lg	Small to large copepod index

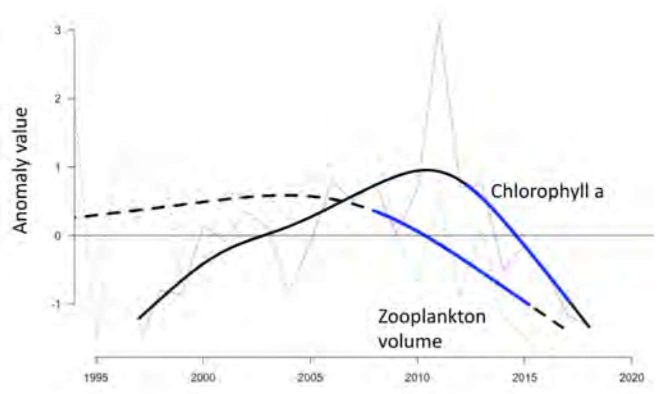


Figure 3.3.3. Trends for chlorophyll a from satellite ocean color data (solid line) and an index of total zooplankton biovolume from EcoMON and MARMAP (dashed) with fitted non-linear trends. Darker lines are GAM smoothed trends. Data include only observations within the New York Bight.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

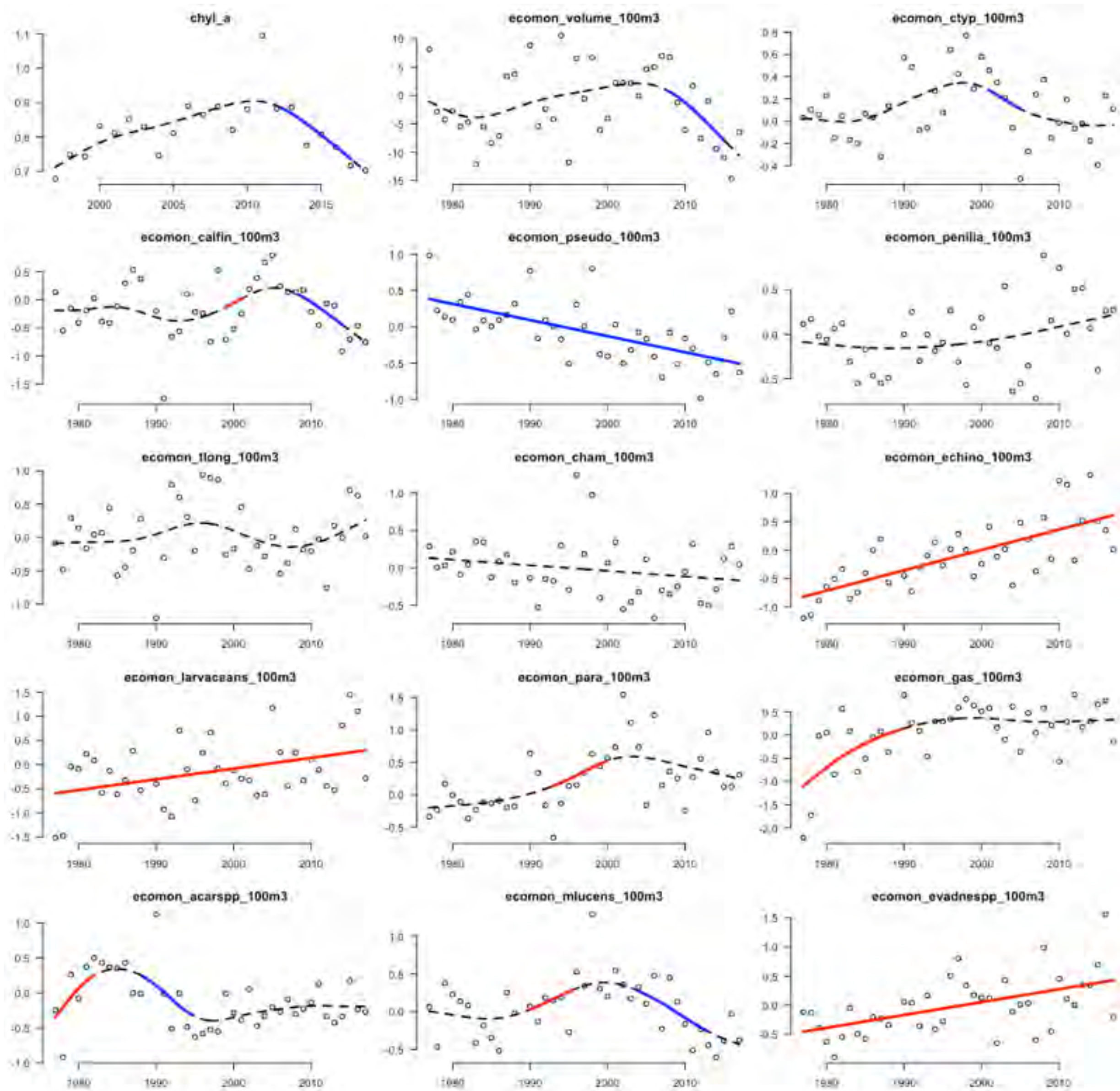


Figure 3.3.4. Time series with non-linear trends for chlorophyll a and time series generated from the EcoMON and MARMAP data. For abbreviations see table 3.3.1.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

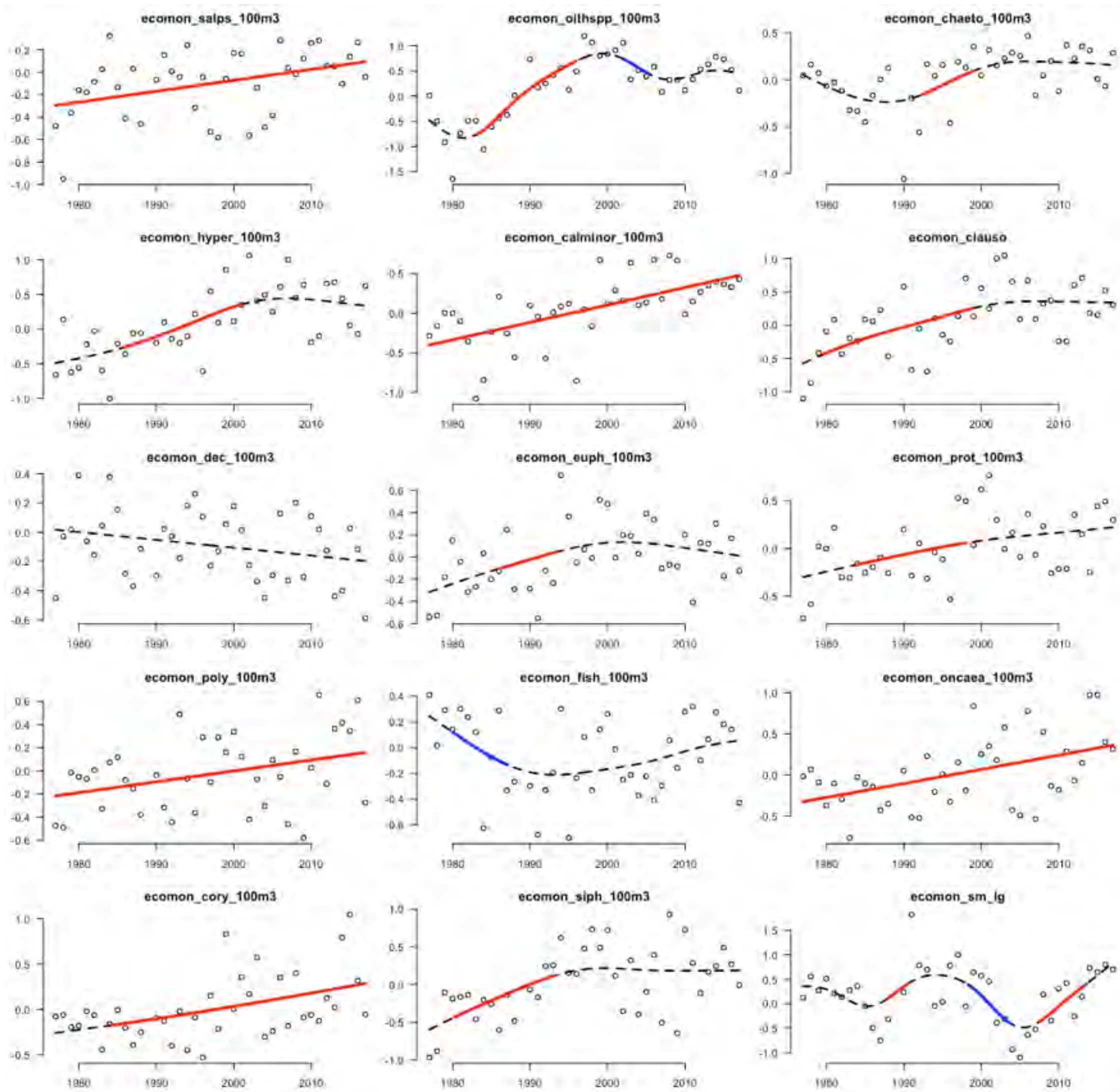


Figure 3.3.5. Time series with non-linear trends for time series generated from the EcoMON and MARMAP data. For abbreviations see table 3.3.1.

<u>Relationships between variables and redundancy</u>

The figures below show the correlations between all indicators and a PCA of the indicators (**Figure 3.3.6, 3.3.7**). These should be taken as exploratory and potentially spurious correlations, but (*potentially*) interesting results are noted

- Negative correlation between fish biovolume and *Calanus finmarchicus*
- Positive correlation between *Clausocalanus arcuicornis* and *Paracalanus parvus*
- Positive correlation between Corycaeidae and *Oncaea spp.*

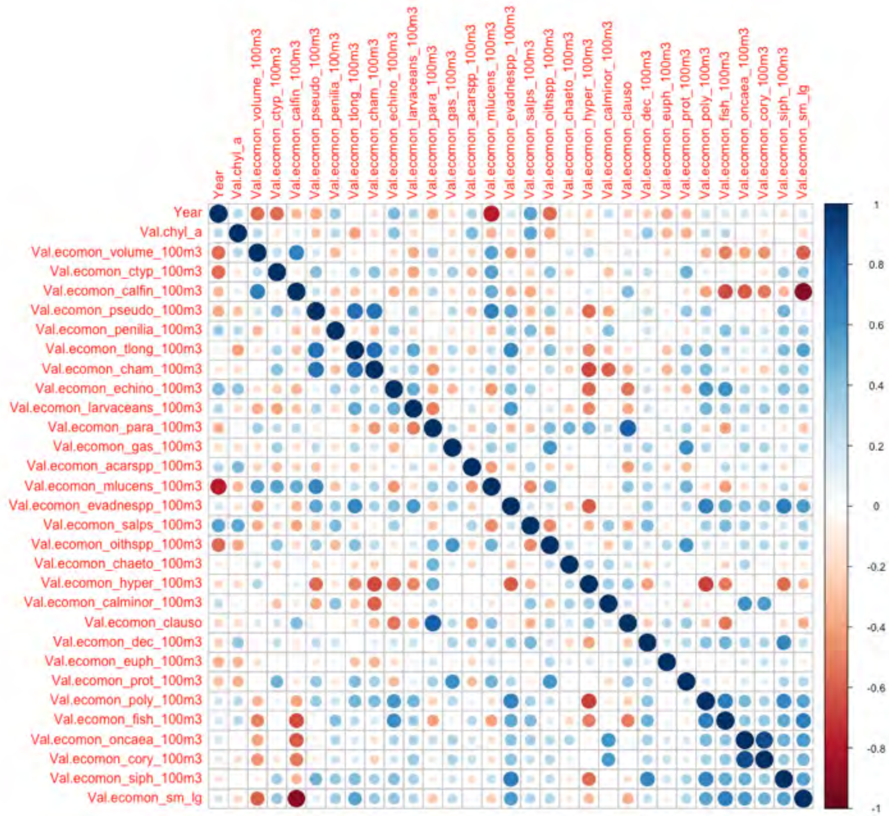


Figure 3.3.6. Correlations between indicators. Bigger and darker circles show stronger correlations.

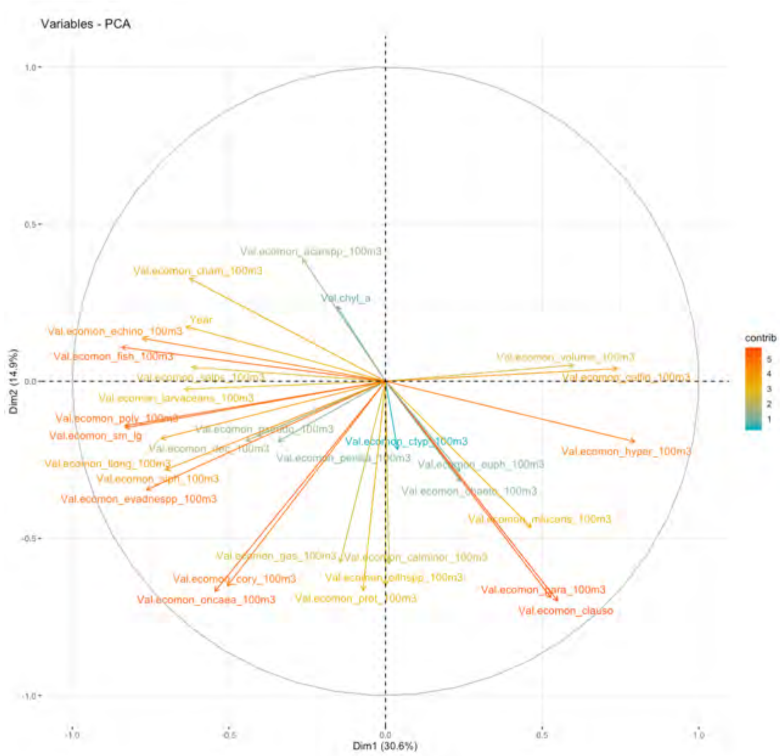


Figure 3.3.7. Principal components analysis of the dataset (including only post 2000 data)

Other potential datasets and indicators to develop

Menhaden or other larval fish from EcoMON

EcoMON data include roughly sixty fish species or taxa categories that could be further explored. For example, there is a column for *Brevoortia tyrannus* larvae (Atlantic Menhaden) that are caught in ring net tows that could be explored. However, oceanic sampling of a fish (like Menhaden) with a substantial part of its life history in estuarine waters could be problematic or a serious limitation

World Ocean Dataset

As mentioned in section 3.1, the WOD dataset could potentially be explored for other variables of interest, including chlorphyll-a. This would be a worthwhile comparison to validate whether the noted decreasing trend in chlorophyll-a from ocean color data is corroborated. Some exploratory work on this has begun; there are around 20,000 casts that take chlorophyll-a measurements present in the dataset at multiple depths.

The Continuous Plankton Recorder Survey and a microplastics indicator

The continuous plankton recorder survey provides free access to data (a data request must be submitted) and provides data that has been collected in a standard way for more than 60 years. The data also include information on microplastics (since 2004) that could be an interesting indicator that is of interest to the general public (**Figure 3.3.8**). There appears to be data collected within the NYB (**Figure 3.3.9**) but a data access request would need to be made to further explore the use of this dataset.

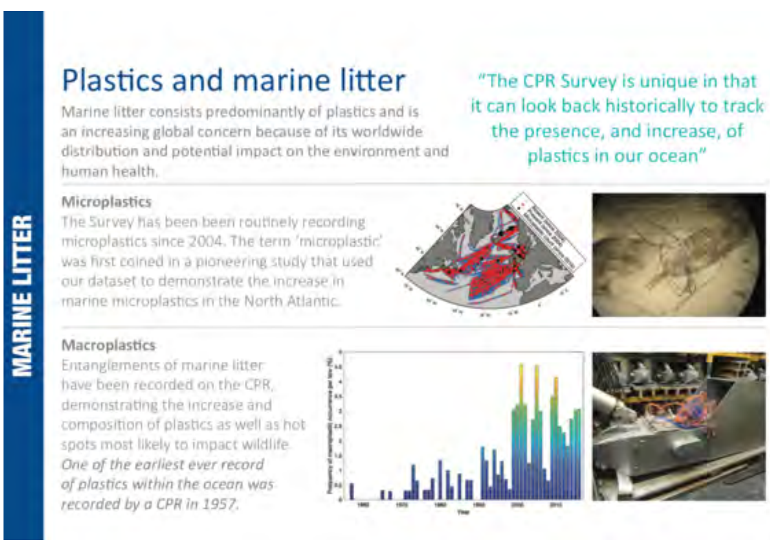


Figure 3.3.8. Infographics for the CPR survey on microplastics and marine litter.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

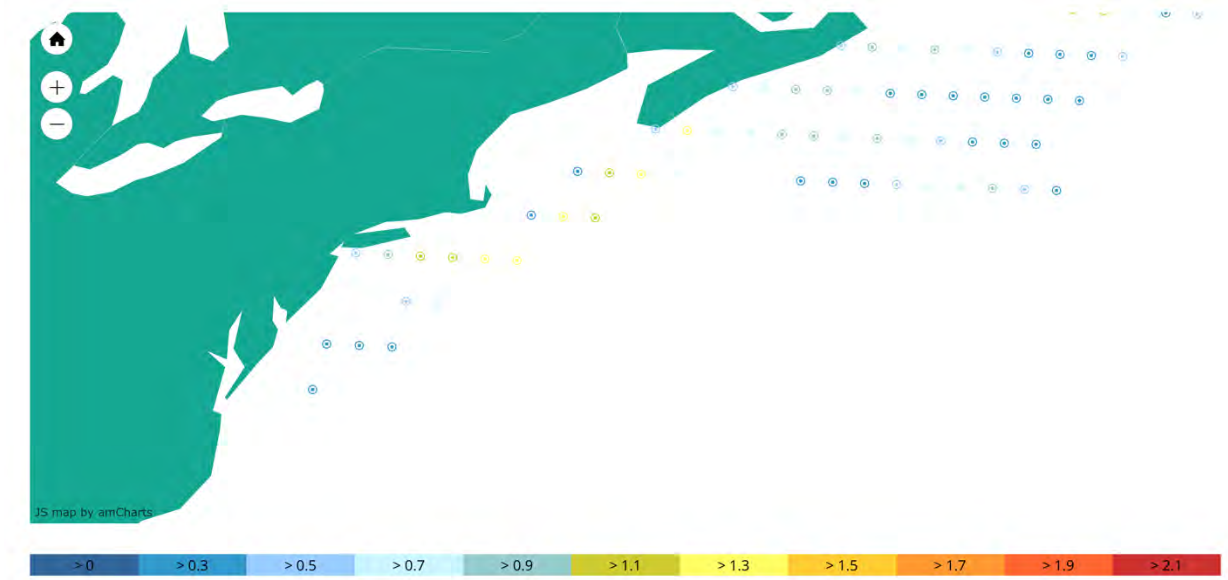


Figure 3.3.9. Continuous plankton recorder ‘samples per annum’ (not exactly sure what this means since it is not an integer) in the NYB.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

3.4: Fish and shellfish indicators

Summary

- NOAA surveys of bottom fish, scallops, clams and quahogs cover the NYB each year
- Indicators are presented for individual fish species and aggregated fish species groups (e.g., Planktivores, Benthivores, Piscivores, Forage fish)
- Other indicators using these datasets could be developed such as: warm/cold fish ratio, fish size, fish distribution, etc. These would require additional work and discussion.
- Quantitative analysis to determine if time series for the NYB differ from the MAB as a whole are being conducted
- Good indicators of mid-water column fish biomass in the NYB are not currently available
- Some mid-water acoustic data are collected during NOAA trawls and NEAMAP surveys (inshore, since 2018) but would require substantial resources to process and analyze.

Key datasets identified and data coverage

NOAA conducts three regular surveys that extend across most of the NES, and cover the NYB fairly well. These are reviewed briefly below along with the NEAMAP inshore survey and the new DEC funded NY inshore survey.

NOAA Bottom Trawl Survey

The NOAA Fall bottom trawl survey began in 1963 and has ran continuously since then, in 1968 a Spring survey was added. For a short period there were winter and summer surveys, but these were discontinued. The survey also covered inshore strata for a short time but these were dropped and now the survey only covers offshore (> 3 miles) strata. Because methods have changed (i.e., boats, nets), calibration factors are applied to catch to standardize the time series for long term trend evaluation. This standardization is done by NOAA before the data are sent to us. Details of these gear and vessel changes are discussed in the R markdown file that analyzes these data, and at the following website:

https://noaa-edab.github.io/tech-doc/survdat.html

Over the course of the survey from 1963 to present, a minimum of 47 trawl samples were collected annually within the NYB (max = 208, mean = 116). It appears that the fall survey was cut short in 2017 and was only completed in the Northern part of the NES. The timing of the survey has been relatively stable over the time period for the spring survey, but the fall survey displays a more prominent trend (**Figure 3.4.2**). Data from this survey are used to develop indices of abundance for stock assessments and used by NOAA for ecosystem indicators of functional feeding groups (see below).

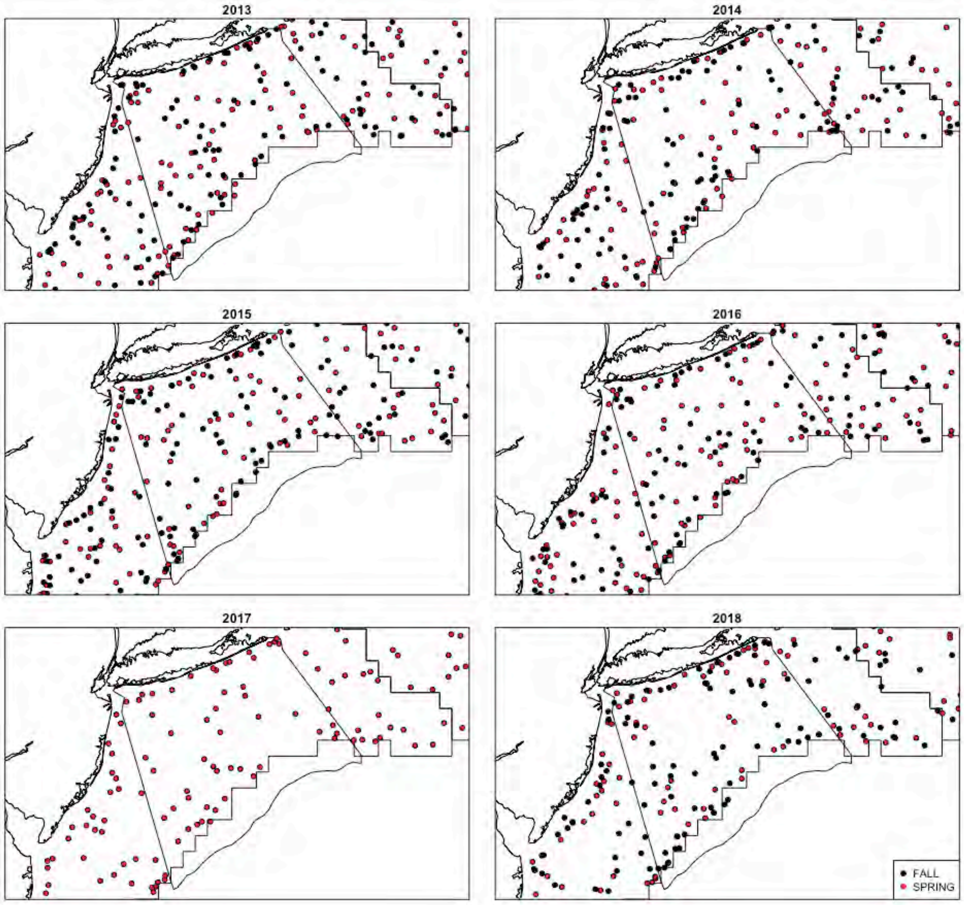


Figure 3.4.1. Distribution of samples in the New York Bight from NOAA bottom trawl survey (2013 to 2018).

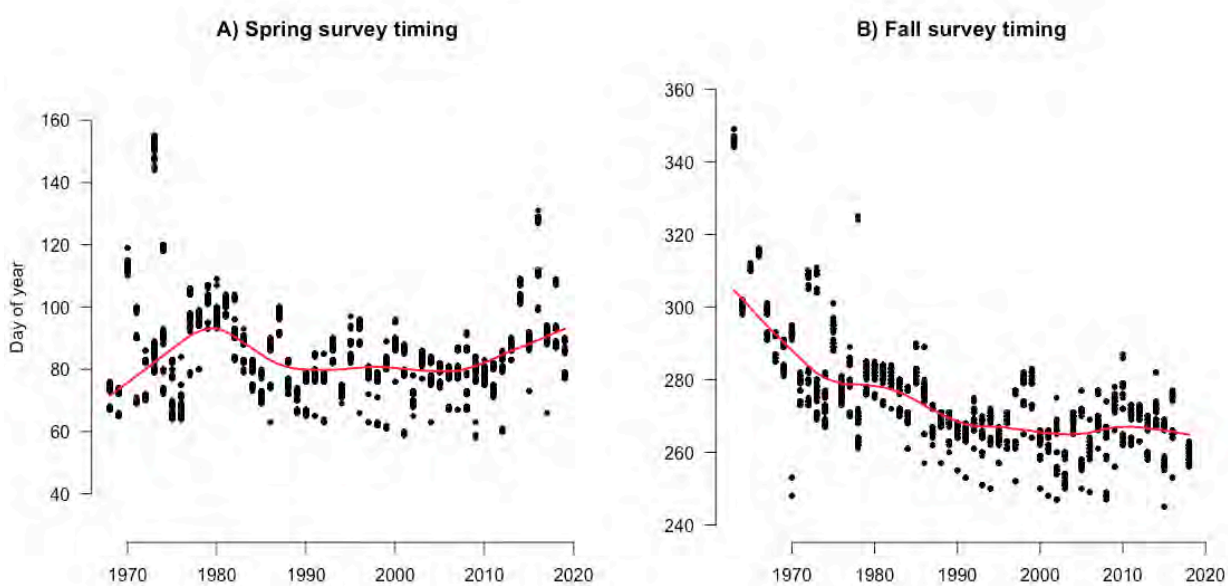


Figure 3.4.2. The timing of the spring and fall bottom trawl survey. Shown are all samples (i.e., trawls, many are overlapping) that occurred in the NYB over the duration of the survey with a loess smoother to display the trend.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

NOAA Scallop Survey

The NEFSC began a standardized Sea Scallop survey in 1980 and covers an area from Cape Hatteras to Georges Bank. The sampling is conducted with a dredge (8-ft wide New Bedford style) and a towed camera (HabCam). Several vessels have been used during the duration of the survey. This dataset was acquired but has not yet been evaluated for trends in the NYB.

https://catalog.data.gov/dataset/sea-scallop-survey

NOAA Atlantic Surfclam and Ocean Quahog Survey

The NEFSC Atlantic Surfclam and Ocean Quahog Survey covers an area from NC to Georges Bank; the survey area is only sampled in part each year with full converge of this area taking about 3 years. The survey operated with consistent methods and gear (a 5' hydraulic dredge) from 1982 to 2011, in 2012 the platform was changed and now a 13' commercial style dredge is used. The survey records catch weight, catch number, individual weight, individual length, and individual meat weight. This dataset was acquired but has not yet been evaluated for trends in the NYB.

https://catalog.data.gov/dataset/atlantic-surfclam-and-ocean-quahog-survey

NEMAP

The North East Area Monitoring and Assessment Program (NEAMAP) is a fisheries independent bottom trawl survey that has been conducted twice annually since 2006 in nearshore waters off the Atlantic Coast. It is run by the Virginia Institute of Marine Science (VIMS) at William & Mary college. The survey, covers an area that is not surveyed by NOAAs bottom trawl survey (**Figure 3.4.3**). In 2018 a Simrad ES80 bioacoustics system was added to their vessel (the F/V Darana). Their website states that a collaboration with NEFSC is working to analyze the acoustics data to develop biomass estimates of pelagic fish (see link A below). The trawl data are now regularly incorporated into stock assessments. Further information can be seen on their website (Link A).

Link A:
http://www.asmfc.org/files/Meetings/77AnnualMeeting/ISFMP_PolicyBoardPresentations_Oct2018.pdf

Link B:
https://www.vims.edu/research/departments/fisheries/programs/multispecies_fisheries_research/neamap/index.php

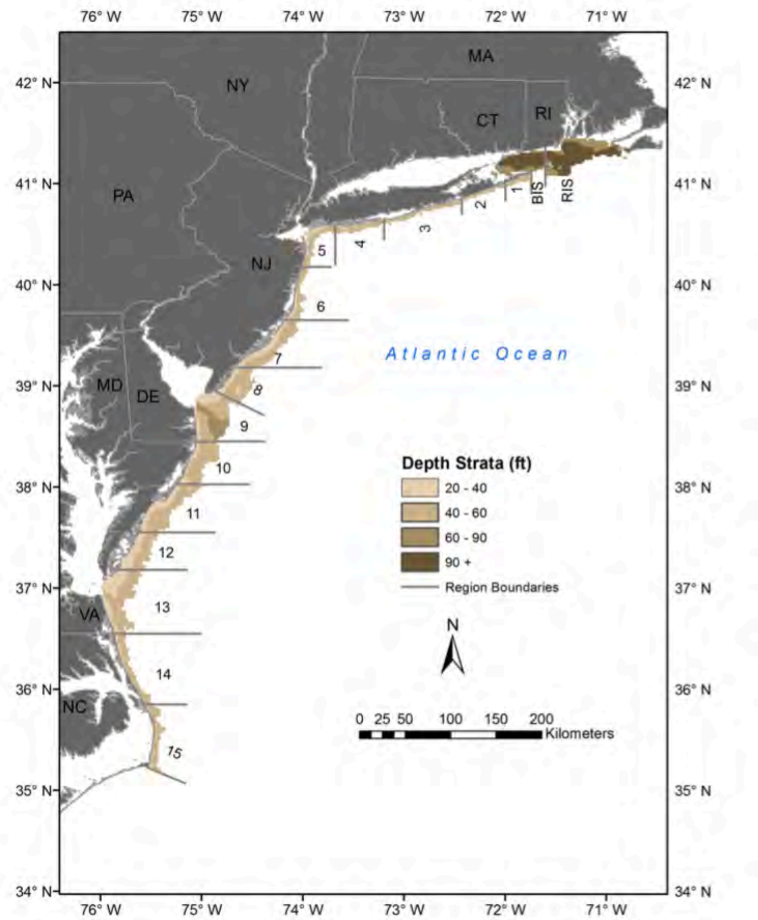


Figure 3.4.3. Sampling distribution of the NEAMAP survey

NY State Trawl

NY State DEC began an Ocean Trawl Survey in 2017 (a bottom trawl), that is conducted on the RV Seawolf. The distribution of sampling strata are shown in **figure 3.4.4**.

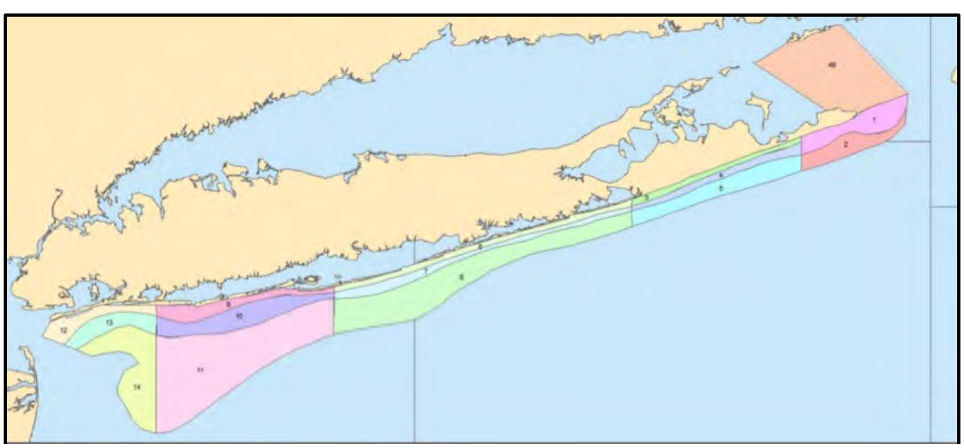


Figure 3.4.4. Sampling distribution of the NY Inshore trawl survey.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

NYB indicator overview

Indicators were calculated using the NEFSC Bottom Trawl survey. The other datasets described above were not explored in depth. We used methods described by NOAA to calculate mean biomas per tow (per year) for (1) individual fish species of interest (2) species groupings used in the 2018 ESR for the Mid-Atlantic Bight, and species groupings used in the 2011 ESR for the Mid-Atlantic Bight. The 2011 groupings included a category representing “Forage Fish” a group that DEC staff suggested to explore.

Trends were assessed with linear regression and GAMS and are presented below.

Table 3.4.1. Trends in single species indicators assessed with linear regression (direction indicated by arrows for significant relationships, r-squared shown in last column). The non-linear trend (since 2010) shows whether there is a significant increasing, decreasing, or no trend (with arrows) as of 2010 determined by GAM fit.

Indicator	Linear trend	Non-linear trend (since 2010)	Linear r-squared
AMERICAN LOBSTER FALL	⇩		0.209
AMERICAN LOBSTER SPRING	⇩		0.19
BUTTERFISH FALL			0.006
BUTTERFISH SPRING			0.041
LONGFIN SQUID FALL	⇧		0.127
LONGFIN SQUID SPRING		⇩	0.17
LONGHORN SCULPIN FALL	⇩		0.261
LONGHORN SCULPIN SPRING	⇩	⇩	0.075
NORTHERN SHORTFIN SQUID FALL			0.002
NORTHERN SHORTFIN SQUID SPRING		⇧	0.021
SCUP FALL		⇧	0.021
SCUP SPRING	⇧	⇧	0.216
SILVER HAKE FALL	⇩		0.309
SILVER HAKE SPRING	⇩		0.422
SUMMER FLOUNDER FALL	⇧		0.202
SUMMER FLOUNDER SPRING	⇧		0.367
WINTER FLOUNDER FALL	⇩		0.266
WINTER FLOUNDER SPRING			0.018

Table 3.4.2. Trends in aggregate group indicators (from 2011 ESR) assessed with linear regression (direction indicated by arrows for significant relationships, r-squared shown in last column). The non-linear trend (since 2010) shows whether there is a significant increasing, decreasing, or no trend (with arrows) as of 2010 determined by GAM fit.

Indicator	Linear trend	Non-linear trend (since 2010)	Linear r-squared
BenthicInvert FALL	⇧	⇧	0.151
BenthicInvert SPRING			0.003
Forage Fish Fall			0.013
Forage Fish Spring			0.011
Mid-Atlantic Groundfish Fall		⇧	0.003
Mid-Atlantic Groundfish Spring	⇧	⇧	0.284
Elasmobranchs FALL	⇩		0.126
Elasmobranchs SPRING			0.027
Groundfish FALL	⇩		0.418
Groundfish SPRING	⇩		0.639

Table 3.4.3. Trends in aggregate group indicators (from 2018 ESR) assessed with linear regression (direction indicated by arrows for significant relationships, r-squared shown in last column). The non-linear trend (since 2010) shows whether there is a significant increasing, decreasing, or no trend (with arrows) as of 2010 determined by GAM fit.

Indicator	Linear trend	Non-linear trend (since 2010)	Linear r-squared
Benthivore FALL	↓		0.134
Benthivore SPRING			0.017
Benthos FALL	↑	↑	0.204
Benthos SPRING			0.004
Piscivore FALL	↓		0.183
Piscivore SPRING			0.003
Planktivore FALL			0.005
Planktivore SPRING			0.012

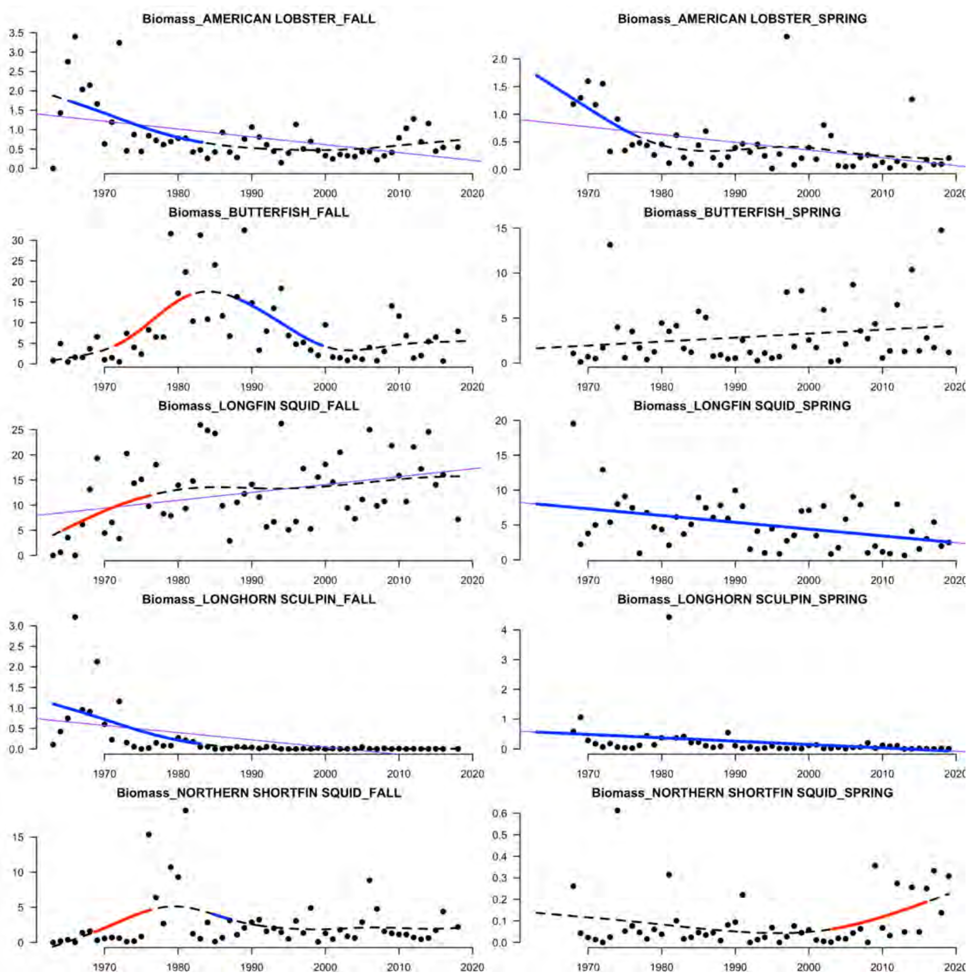


Figure 3.4.5. Indicators of individual fish species in fall and spring surveys. Time series with significant linear trends have a purple line (linear regression fit), GAM estimates are fit with a black line and significant periods of change are colored (red = increasing, blue = decreasing).

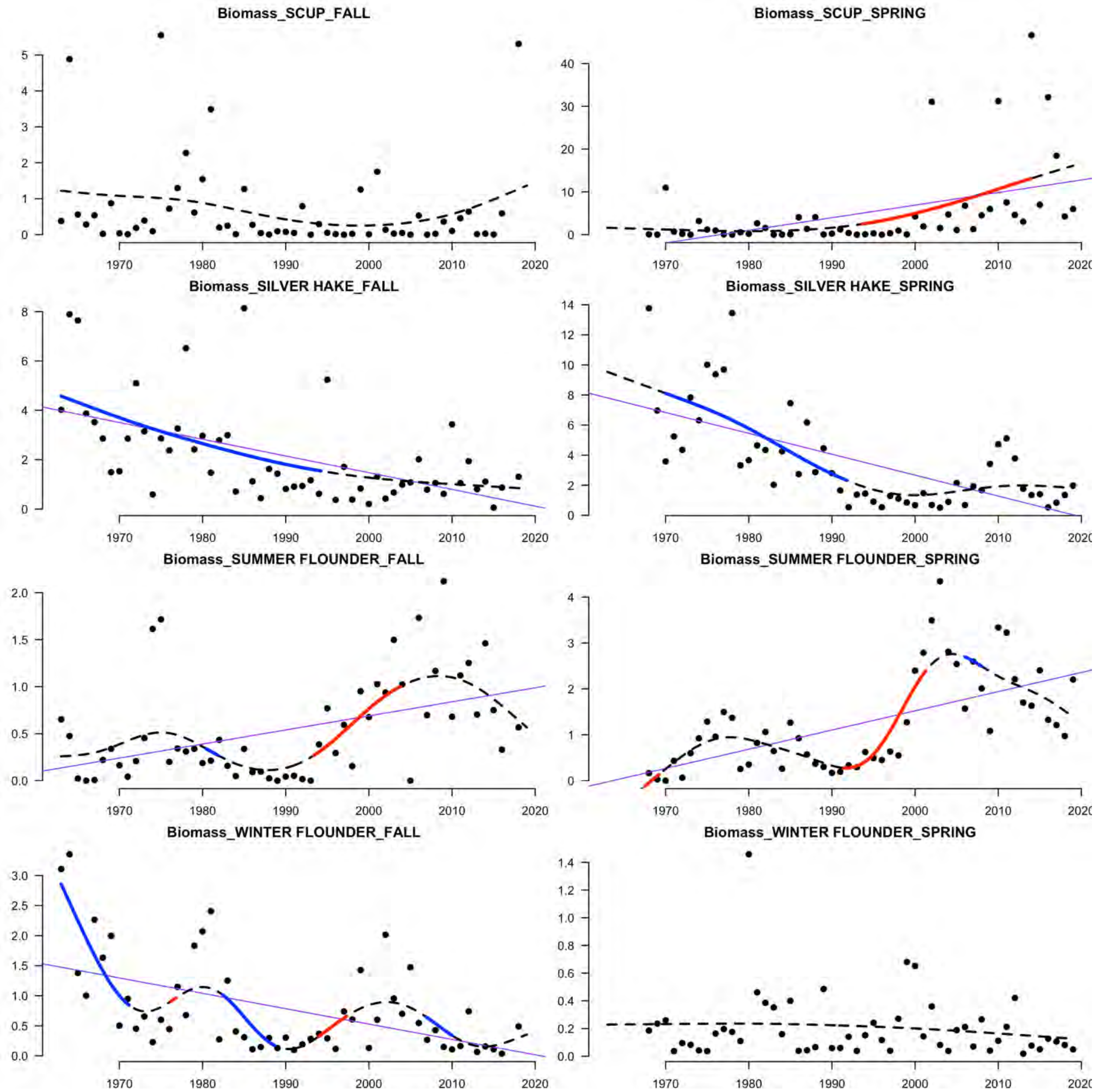


Figure 3.4.6. Indicators of individual fish species in fall and spring surveys. Time series with significant linear trends have a purple line (linear regression fit), GAM estimates are fit with a black line and significant periods of change are colored (red = increasing, blue = decreasing).

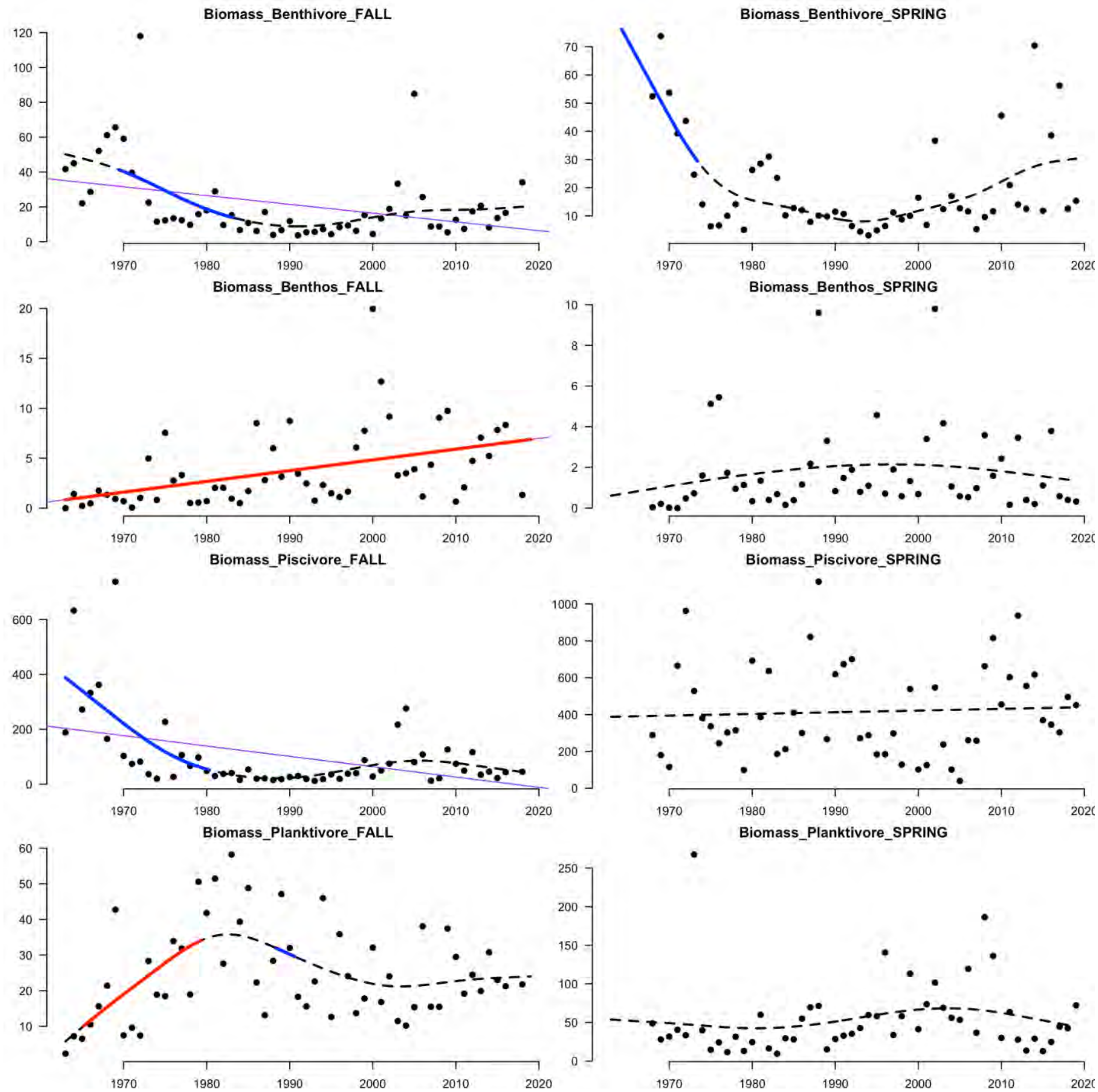


Figure 3.4.7. Indicators of individual fish species in fall and spring surveys. Time series with significant linear trends have a purple line (linear regression fit), GAM estimates are fit with a black line and significant periods of change are colored (red = increasing, blue = decreasing).

__NOTE__: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

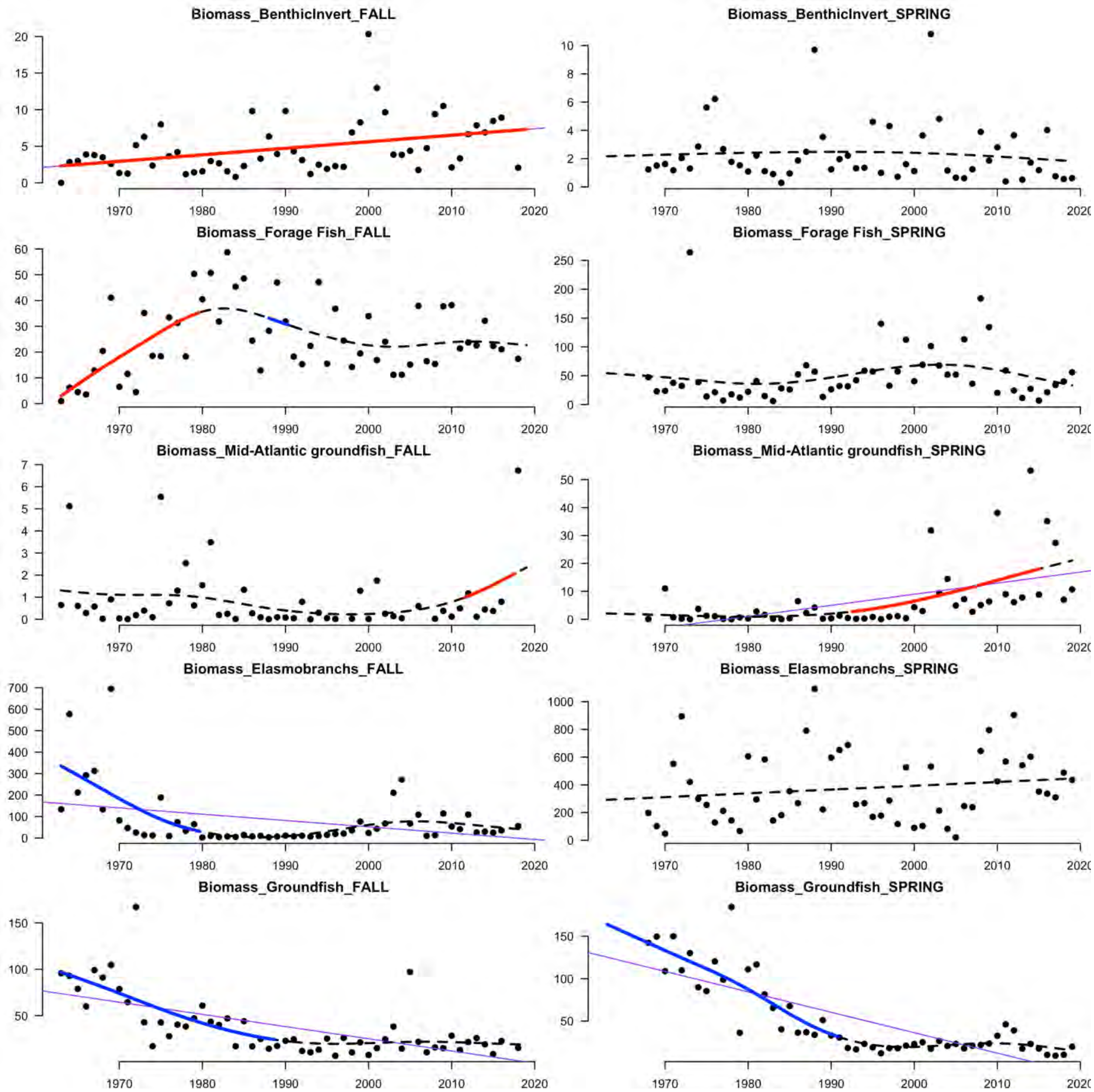


Figure 3.4.8. Indicators of individual fish species in fall and spring surveys. Time series with significant linear trends have a purple line (linear regression fit), GAM estimates are fit with a black line and significant periods of change are colored (red = increasing, blue = decreasing).

Similarity of NYB to MAB time series

NOAA makes indicator trends available for the NYB online and via an R package (ecodata) and therefore these are simple and easily accessed indicators that could be used "as is" in a DEC based assessment. However, the MAB includes an area greater than the NYB that ranges south to Cape Hatteras. Thus, one factor to consider is how the trends generated for the MAB relate to the trends that are calculated for the NYB. This is in fact an important consideration for any of the indicators presented through this entire report and should be discussed. This consideration was investigated briefly for fish indicators.

Not surprisingly, the trends for the NYB were quite similar to the trends for the MAB in many cases (**Figure 3.4.9**). **Figure 3.4.9** shows the series plotted together in four examples. In all comparisons (n = 45) the mean correlation (i.e., NYB trend compared to MAB trend) was 0.68 (sd = 0.27) and was as high as 0.98. The examples also show that although trends are often similar, they may have slightly different absolute values (i.e., forage fish where NYB is higher than MAB generally but trends are equivalent). Others show lower consistency in trends like Scup biomass in the fall, where it is showing an increase in the MAB as a whole but this is not attributable to increases in the NYB.

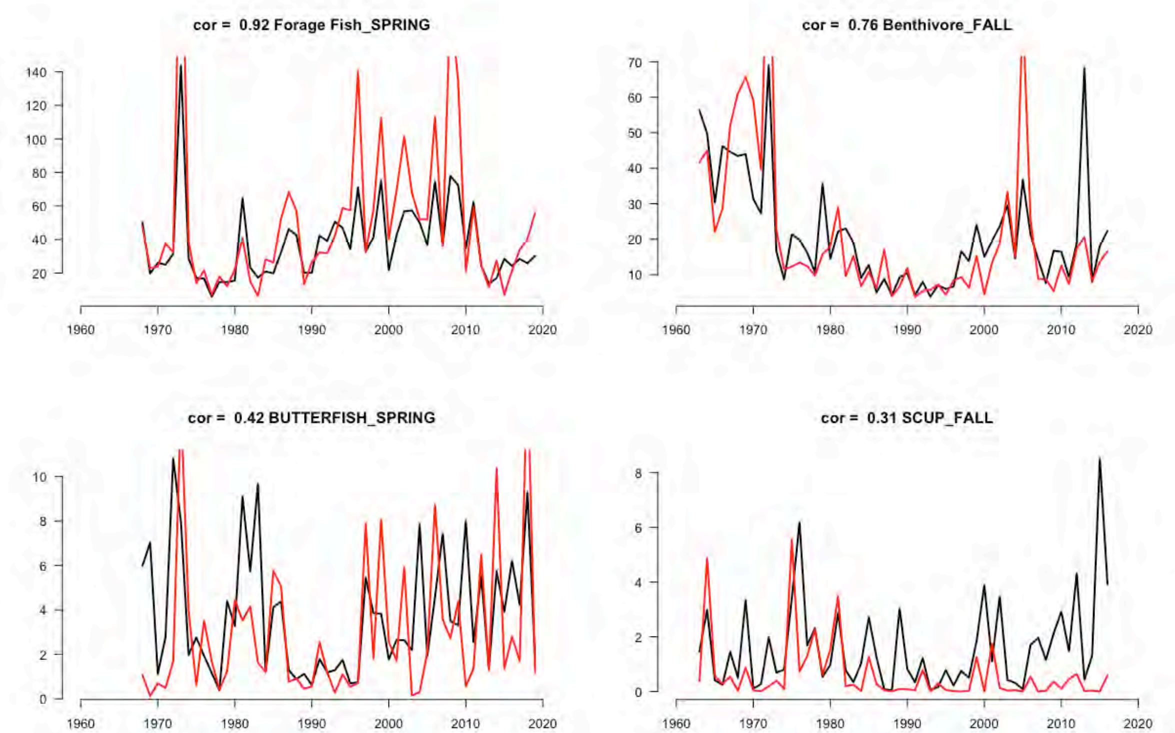


Figure 3.4.9. Comparison of indicator trends for the NYB (red) and the MAB (black). Titles show the calculated correlation.

NOTE: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

<u>Other potential datasets and indicators to develop</u>

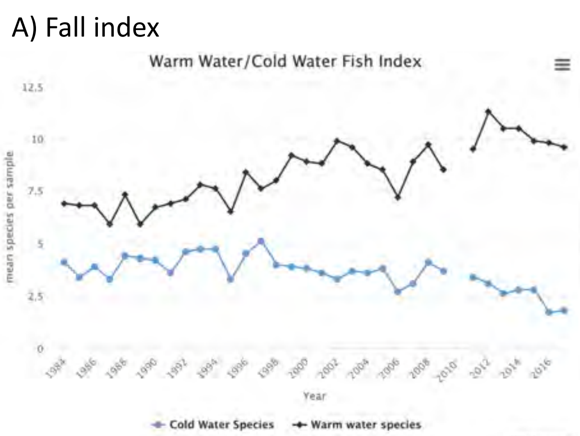
Other NOAA indicators

Any of the fish indicators used by NOAA (and based on the bottom trawl survey) could be re-worked for the NY Bight. These include total catch biomass, information on fish lengths, ratios of one group to another, ratio of large to small fish, and species composition or diversity indices.

Warm to cold fish ratio

The DEC expressed interest in a warm/cold fish ratio indicator to reflect changes in species composition that might result from climate change and warmer water temperatures. Currently the Long Island Sound Study uses a similar indicator for trawl samples in the Long Island Sound. A description of this indictor is copied below:

"Finfish species captured in the CTDEEP Long Island Sound Trawl Survey were divided into adaptation groups based on their temperature tolerance and seasonal spawning habits. Members of the cold-adapted group (numbering 33 species) prefer water temperatures below 15C (60F), tend to spawn early in the year, and are more abundant north of the Sound than south of New York. In contrast, the warm-adapted group (numbering 38 species) prefer warmer temperatures (11-22C or 50-72F), tend to spawn later, and are more abundant south of the Sound than north of Cape Cod. The index is the average number of species in each group (also called species richness) captured in spring (April, May, June) and fall (September, October) Survey samples each year from 1984 through 2014 (note no data were taken in fall 2010)."



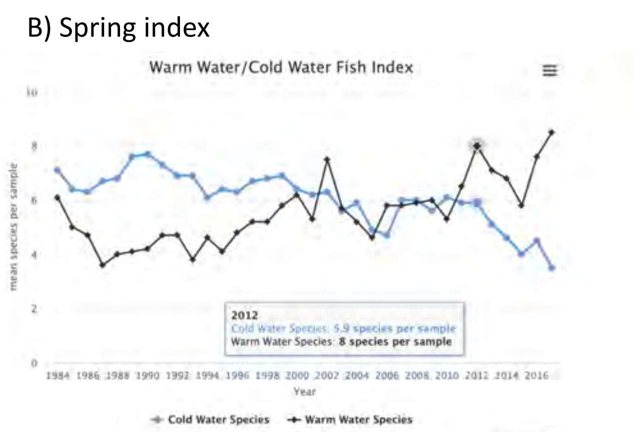


Figure 3.4.10. Cold to warm fish indicator for Long Island Sound in the Fall (left) and Spring (right).

To develop a warm/cold fish ratio indicator in the NYB we could either (1) use the same species groupings as the LISS (above, would need to contact these scientists) or (2) develop our own groupings based on existing thermal tolerance or preference data. Once this decision is made it will be easy to calculate this indicator. In the LISS version the number of species caught per tow is used, but relative biomass could also be explored.

__NOTE__: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.

NOAA acoustics

It appears that fisheries acoustics data have been collected aboard the vessel conducting the NOAA spring and fall bottom trawl surveys and others. This data can be accessed by request and could (potentially) be used to generate time series of mid-water fish and zooplankton abundance in the NYB. However, the datasets are large, probably require substantial quality control measures and time investment to work with. Although appearing in the data catalogs below, we are not aware of this data being used for indicator development, stock assessments, or publications.

Link A: Description of the acoustic operations by NOAA NEFSM
https://www.nefsc.noaa.gov/esb/acoustics/operations.html

Link B: map showing data availability of acoustics and how to access.
https://www.ngdc.noaa.gov/mgg/wcd/

Link C: Data information
https://inport.nmfs.noaa.gov/inport/item/26654

3.5: Cetaceans, coastal seabirds, and sea turtles

Summary

- Many data portals provide access to individual datasets, but few long-term data series are available to easily generate consistent abundance indicators for cetaceans, seabirds, or sea turtles
- Some annual surveys are available for birds on Long Island
- Large whale stranding data were accessed recently, and several simple indicators are shown

Key datasets identified and data coverage

For marine mammals, coastal and seabirds, and sea turtles we only show a few representative indicators for the NY Bight; this section primarily describes the general datasets that are known to exist for the NYB. There are a number of datasets that could possibly be used for a DEC indicator; yet there are important caveats. For example, there are large databases of marine mammal observations, but data are not (generally) not collected in consistent spatial and temporal sampling design that would allow for a simple and representative time series to be calculated.

In other cases, there are data available from community groups but the scientific credibility of the datasets are not tested. For example, time series of bird counts are available from the NY State Ornithological Association but how representative these are given sampling effort (i.e., number of samples per year) is not clear (**Figure 3.6.1**), and would require additional validation.

Rather than diving deep into any of these potential indicators, this section presents some insights and a review of the existing datasets encountered during the search.

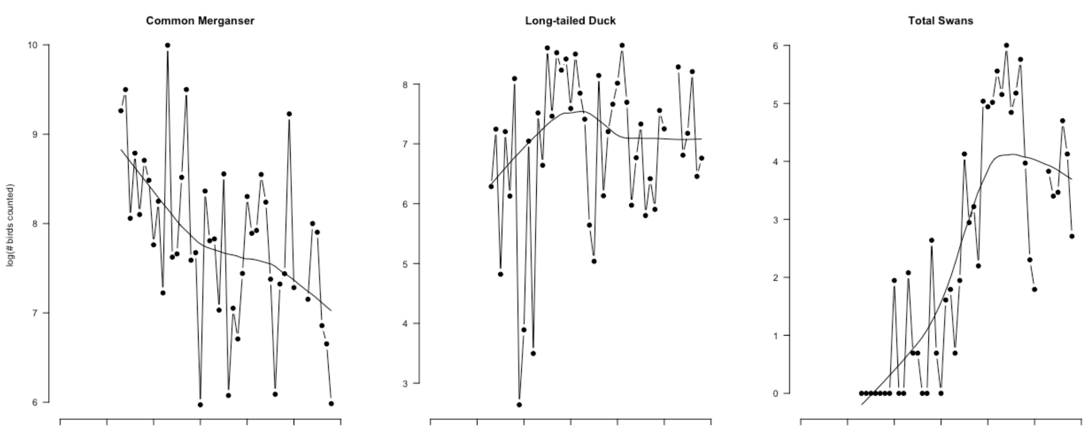


Figure 3.6.1. Bird counts from the NY State Ornithological Association on Long Island, counts are transformed (log +1) to better visualize trends, and a LOESS smother is applied with default values.

NOTE*: This report is prepared for internal use only (SoMAS, DEC) and is not intended for public dissemination.*

National and international databases of observations and telemetry data

There are many large data repositories where data is contributed by many users and hosted online (**Table 3.5.1**). These datasets are varied and often include telemetry datasets, structured surveys of limited duration (i.e., several years), or opportunistic sightings data.

Table 3.5.1. National or international databases of relevant datasets.

Data source	Type	Taxa
Ocean Biogeographic Information System (OBIS)	Obs.; Telemetry	Mammalia; Aves; Chondrichthyes; Reptilia
Movebank	Telemetry	Mammalia; Aves; Chondrichthyes; Reptilia
Global Seabird Tracking Database (Birdlife international)	Telemetry	Aves
USGS Patuxent Atlantic Offshore Seabird Dataset Catalog	Obs.	Aves
The State of the World's Sea Turtles (SWOT)	Obs.; telemetry	Reptilia (sea turtles)
BirdLife Seabird Tracking Database	Telemetry	Aves
Atlantic Offshore Seabird Dataset Catalog	Obs.	Aves
National Stranding Database	Stranding	Mammalia
Sea Turtle Rehabilitation and Necropsy Database	Stranding	Reptillia
Sea Turtle Nest Monitoring System	Nest Obs.	Reptillia
Marine Mammal Stock Assessment Reports (SARs)	Abundance	Mammalia
USFWS Mid-Winter Waterfowl Survey	Abundance	Aves
USFWS breeding bird survey	Abundance	Aves

Level 1 stranding data

A promising dataset for a useful indicator is the level A stranding data. These data include location, date, species, and cause of death (when this can be determined conclusively) and could be used to quantify the occurrence of standings on Long Island. These could also be compared to other states (i.e., proportion of total strandings on Long Island vs. coastwide).

These data would be particularly useful to include and compare with indicators of boat traffic, marine heatwaves, population trends in marine mammal species, or other potential drivers that could contribute to unusual mortality events.

This dataset was acquired in December and some preliminary plots are shown below (**Figure 3.6.2**).

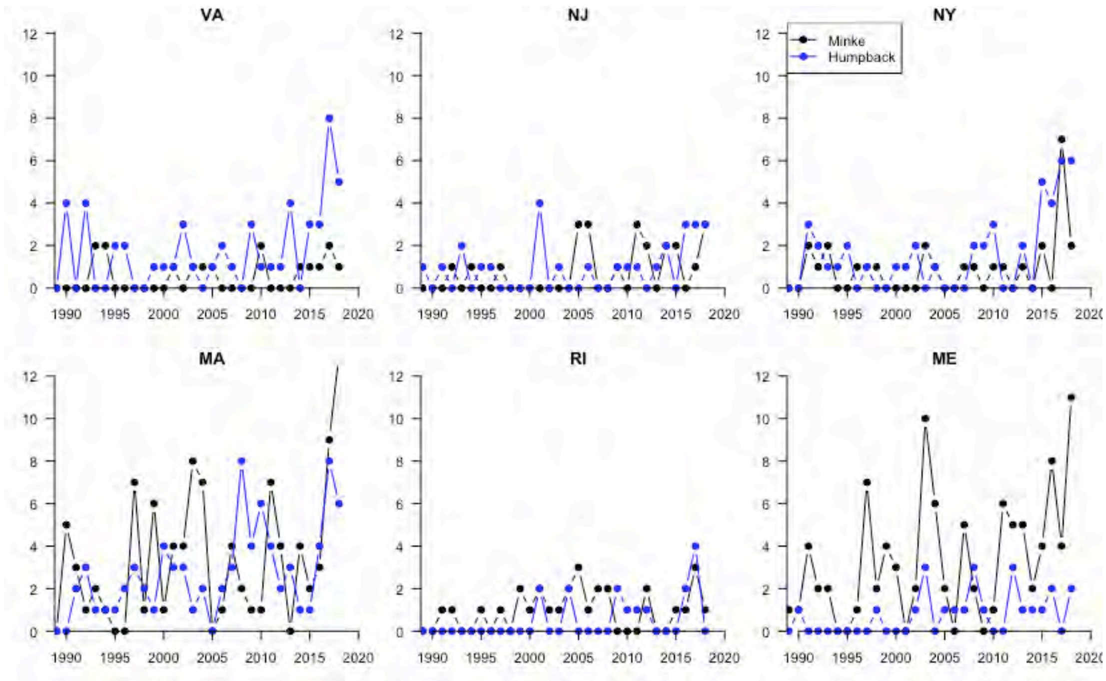


Figure 3.6.2. Count of stranding records in six states for Minke and Humpback whales

Marine Mammal Stock Assessment Reports

The annual stock assessment report by NOAA for marine mammals has many datasets that could be used; but few are specific to NY. For example, time series of abundance estimates of each stock of marine mammals are presented in this document for all species occurring in US waters, but these do not provide specific information about the occurrence of whales in the NYB. These trends could be useful though for a variety of purposes (i.e., do strandings occur more in NY as population wide trends increase?).

NY State Acoustic Survey

Cornell University is leading a study using passive acoustic monitoring for large whales in the NYB; it will consist of recorders that can identify whales and be used to generate the following data products (below is quoted directly from website link below).

1. Percent and number of days per deployment in which at least one call was heard
2. For each month and season the pattern of daily contacts
3. Spatial distribution of days within each month and season containing calls per receiver
4. Analysis of inter-annual variability of these parameters
5. Calculations of ambient noise and acoustic masking potential
6. Maps of cumulative spatial distribution of each whale species for each year and for the entire survey

The study will be conducted for only 3 years so will not provide a long-term standardized data source for indicator development unless the project is extended.

https://www.dec.ny.gov/lands/113828.html

NY State Aerial Survey

The text below was copied from the DEC website (i.e., September 2019 report by tetra tech), and describes this survey, that is set to cover 3 years. With only three years of data coverage, this will not be useful for long-term indicator development.

"Tetra Tech, Inc., in coordination with LGL Ecological Research and Aspen Helicopters, Inc. (collectively, the "survey team"), is contracted by the New York State Department of Environmental Conservation (NYSDEC), Division of Marine Resources to conduct 36 monthly line-transect aerial surveys focused on the six large whale species most likely to occur in the New York Bight. This survey report documents the survey effort and sightings from the September 2019 survey, representing the 31st of the 36 surveys scheduled to occur under this contract."

Woods Hole DMON Buoy

In 2016 a buoy was deployed outside NY harbor that runs continuously and can detect the presence of large whales. This could be a valuable data source for an indicator if deployed indefinitely and maintained to provide long-term indices. The contacts for information on this dataset are Mark Baumgartner (Woods Hole Oceanographic Institution) and Howard Rosenbaum (Wildlife Conservation Society).

http://dcs.whoi.edu/nyb0616/nyb0616.shtml

NY State Ornithological Association Bird Counts

Every January since 1955, members of the NYSOA have been counting waterfowl; of interest here are counts from Region 1 (Long Island). Results are published in the NYSOA magazine titled the Kingbird, and are also available online. Although these counts are shore based, and focus on waterfowl, and may not be an ideal scientific design, they may be of some value for use as an indicator to DEC. **Figure 3.6.1**. provides an example.

https://nybirds.org/ProjWaterfowl.htm

Long Island Colonial Waterbird Survey

There is little data regarding this survey online, yet some reports are available and suggest the data may be useful. A report by Kevin Jennings (DEC, Stony Brook) shows 237 sites surveyed by a collaboration of 30 agencies/groups, and roughly 120 people participating in this monitoring program. The data seem to go back to ~2000 and highlight an example where broader involvement of other DEC staff and resources could contribute to the Ocean Indicators System.

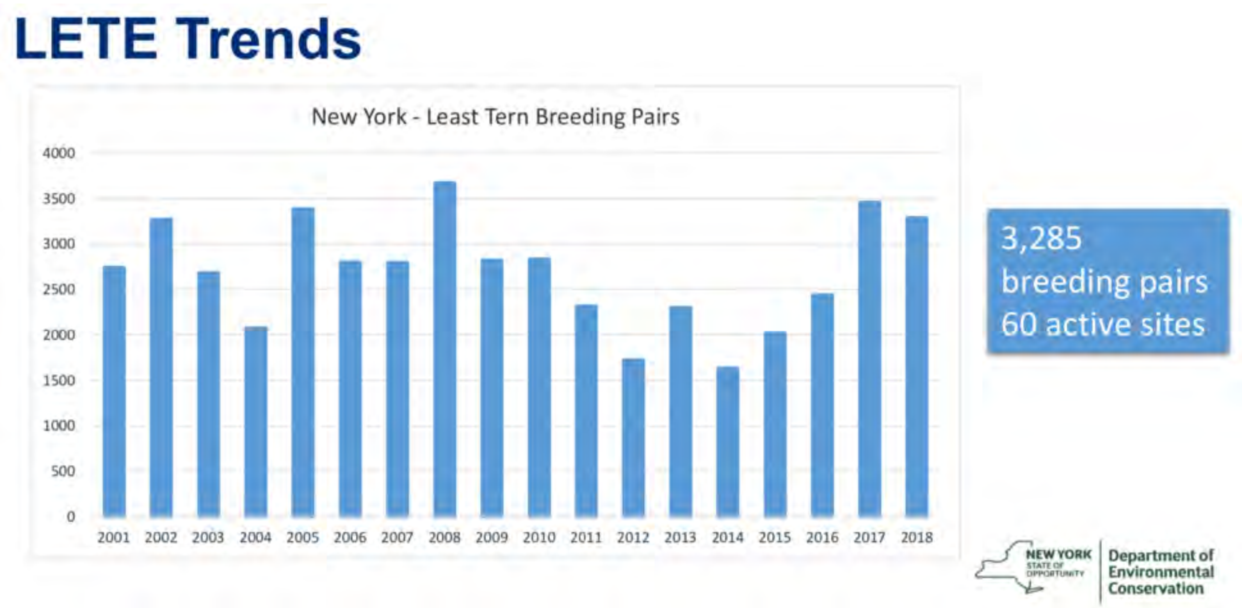


Figure 3.6.2. Least terns breeding pairs from the Colonial Waterbird Survey

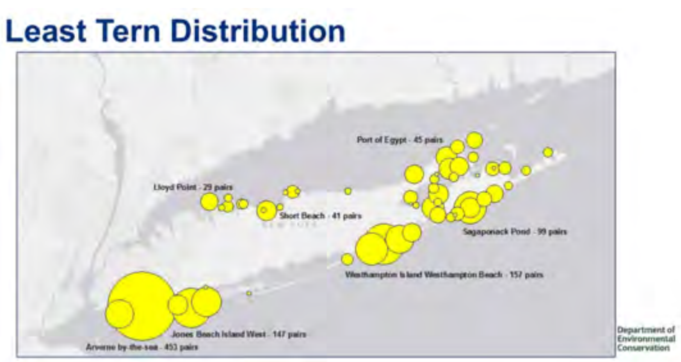


Figure 3.6.3. Least tern distribution from the Colonial Waterbird Survey

References

Curtis, C.J., and Simpson, G.L. 2014. Trends in bulk deposition of acidity in the UK, 1988-2007, assessed using additive models. Ecol. Indic. **37**(PART B): 274–286. Elsevier Ltd. doi:10.1016/j.ecolind.2012.10.023.

DePiper, G.S., Gaichas, S.K., Lucey, S.M., Da Silva, P.P., Anderson, M.R., Breeze, H., Bundy, A., Clay, P.M., Fay, G., Gamble, R.J., Gregory, R.S., Fratantoni, P.S., Johnson, C.L., Koen-Alonso, M., Kleisner, K.M., Olson, J., Perretti, C.T., Pepin, P., Phelan, F., Saba, V.S., Smith, L.A., Tam, J.C., Templeman, N.D., and Wildermuth, R.P. 2017. Operationalizing integrated ecosystem assessments within a multidisciplinary team: Lessons learned from a worked example. ICES J. Mar. Sci. **74**(8): 2076–2086. doi:10.1093/icesjms/fsx038.

Gaichas, S.K., DePiper, G.S., Seagraves, R.J., Muffley, B.W., Sabo, M.G., Colburn, L.L., and Loftus, A.J. 2018. Implementing Ecosystem Approaches to Fishery Management: Risk Assessment in the US Mid-Atlantic. Front. Mar. Sci. **5**(November): 1–25. doi:10.3389/fmars.2018.00442.

Gangopadhyay, A., Gawarkiewicz, G., Silva, E.N.S., Monim, M., and Clark, J. 2019. An Observed Regime Shift in the Formation of Warm Core Rings from the Gulf Stream. Sci. Rep. **9**(1): 1–9. doi:10.1038/s41598-019-48661-9.

Hastie, T., and Tibshirani, R. 1986. Generalized additive models. Stat. Sci. doi:10.1214/ss/1177013604.

Henderson, M.E., Mills, K.E., Thomas, A.C., Pershing, A.J., and Nye, J.A. 2017. Effects of spring onset and summer duration on fish species distribution and biomass along the Northeast United States continental shelf. Rev. Fish Biol. Fish. **27**(2): 411–424. Springer International Publishing. doi:10.1007/s11160-017-9487-9.

Hobday, A.J., Alexander, L. V., Perkins, S.E., Smale, D.A., Straub, S.C., Oliver, E.C.J., Benthuysen, J.A., Burrows, M.T., Donat, M.G., Feng, M., Holbrook, N.J., Moore, P.J., Scannell, H.A., Sen Gupta, A., and Wernberg, T. 2016. A hierarchical approach to defining marine heatwaves. Prog. Oceanogr. **141**: 227–238. Elsevier Ltd. doi:10.1016/j.pocean.2015.12.014.

Link, J.S. 2010. Ecosystem-based fisheries management: Confronting tradeoffs. *In* Ecosystem-Based Fisheries Management: Confronting Tradeoffs. doi:10.1017/CBO9780511667091.

Link, J.S., and Marshak, A.R. 2019. Characterizing and comparing marine fisheries ecosystems in the United States: determinants of success in moving toward ecosystem-based fisheries management. *In* Reviews in Fish Biology and Fisheries. Springer International Publishing. doi:10.1007/s11160-018-9544-z.

Methratta, E.T., and Link, J.S. 2006. Evaluation of quantitative indicators for marine fish communities. Ecol. Indic. **6**(3): 575–588. doi:10.1016/j.ecolind.2005.08.022.

R Development Core Team. 2013. R: A Language and Environment for Statistical Computing. Available from http://www.r-project.org/.

Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., and Schlax, M.G. 2007. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. J. Clim. **20**(22): 5473–5496. doi:10.1175/2007JCLI1824.1.

Santibáñez, P.A., Maselli, O.J., Greenwood, M.C., Grieman, M.M., Saltzman, E.S., McConnell, J.R., and Priscu, J.C. 2018. Prokaryotes in the WAIS Divide ice core reflect source and transport changes between Last Glacial Maximum and the early Holocene. Glob. Chang.

Biol. **24**(5): 2182–2197. doi:10.1111/gcb.14042.
Shertzer, K.W., Williams, E.H., Craig, J.K., Fitzpatrick, E.E., Klibansky, N., and Siegfried, K.I. 2019. Recreational sector is the dominant source of fishing mortality for oceanic fishes in the Southeast United States Atlantic Ocean. Fish. Manag. Ecol. (March): 621–629. doi:10.1111/fme.12371.
Staudinger, M.D., Mills, K.E., Stamieszkin, K., Record, N.R., Hudak, C.A., Allyn, A., Diamond, A., Friedland, K.D., Golet, W., Henderson, M.E., Hernandez, C.M., Huntington, T.G., Ji, R., Johnson, C.L., Johnson, D.S., Jordaan, A., Kocik, J., Li, Y., Liebman, M., Nichols, O.C., Pendleton, D., Richards, R.A., Robben, T., Thomas, A.C., Walsh, H.J., and Yakola, K. 2019. It's about time: A synthesis of changing phenology in the Gulf of Maine ecosystem. Fish. Oceanogr. (March): 1–34. doi:10.1111/fog.12429.
Thorson, J.T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fish. Res. **210**(October 2018): 143–161. Elsevier. doi:10.1016/j.fishres.2018.10.013.
Wood, S.N. 2001. mgcv: GAMs and generalized ridge regression for R. R News.

Appendices

Table A1. Description of scripts used in analysis.

Item	File Name	Includes	Lines	Requires
1	"GlobalPhysicalDrivers.Rmd"	6 indicies of widespread use (i.e., AMO, GSI)	228	NA
2	"Prep_OISST_dataset.Rmd"	Get Reynods OISST from server connection; save locally	155	#19
3	"OISST_means.Rmd"	Calculate seasonal mean SST indicators	188	#1
4	"OISST_heatwaves.Rmd"	Calculate MHW statistics	374	#1
5	"WOD_conversion_D50.Rmd"	Access and reformat ALL WOD CTD data	374	#19
6	"Bottom_andSurfTempGamOCT17.Rmd"	Generate in situ surface and bottom index*	369	#5
7	"Stratification_index.Rmd"	Generate stratification index*	212	#5
8	"SurfSalGamOCT2.Rmd"	Generate surface and bottom salinity index*	205	#5
9	"River_Flow.Rmd"	Get USGS data from server, caclualte indicators	129	NA
10	"GetAllBouyData_AUG6.R"	Get NOAA bouy data from server	74	NA
11	"PopIncEcon.Rmd"	Get and calculate social, rec, and commerical indciators	432	#19
12	"NOAA_ColburnDataset.R"	Summarize NOAA social vulnerability data	23	NA
13	"ChlorophyllOCT2.Rmd"	Get satellite Chyl-a data from server, process indciators**	256	#19
14	"EcoMonNOAA_NYB_SCALE.Rmd"	Generate indices from EcoMON dataset	266	#19
15	"AggregateBiomassDEC5.Rmd"	Generate indices from NOAA trawl data***	522	#19
16	"MakeWhalePlot.R"	Generates whale strandings indicators****	71	NA
17	"Bird_Script.Rmd"	Describes some bird datasets; no indicators produced	38	NA
18	"Birdy.R"	Make indicators from LI bird survey****	20	NA
19	"LabelShit.R"	Functions to stratify datasets; often used	72	NA
20	"Gam_DerivGreenWoodANDHEIM.R"	Custom functions for GAM and derivative analysis	337	NA

* Should update with new GAM method (i.e., s(lat,long) instead of s(lat) + s(lon))
** Includes options to perform analysis with MODIS, SeaWiFS, VIIRS, or OC-CCI (this is preferred)
*** Methods appear to differ from NOAA version (trends same, absolute values higher than NOAA)
****Must complte .Rmd document to describe data access

Table A2. The indicators reported in 2019 Ecosystem Status Report (ESR) for the Mid-Atlantic Bight (MAB) by NOAA.

Category	Indicator
Economic and Social	
Commercial sector	Total commercial seafood landings
	Mid-Atlantic (MAFC) managed seafood landings
	Total commercial seafood landings by feeding guild
	Total commercial revenue for region
	Revenue from MAFMC managed species
	Bennet indicator
	Commercial reliance and Engagement (spatial)
	Oyster harvest
Recreational sector	Recreational seafood harvest
	Recreational effort
	Recreational anglers
	Recreational fleet diversity
	Recreational diversity of catch
	Recreational reliance and engagement (spatial)
Protected species	
Sea turtles	**No indicator**
Right whale abundance	Time series/stock assessment
Gray seals	**No indicator**
Harbor porpoise	Harbor porpoise bycatch
Shortnose sturgeon	**No indicator**
Fish and invertebrates	
	Survey biomass by feeding guild
	NEFMC benthivores in Mid-Atlantic
	MAFMC planktivores in GOM
	MAFMC planktivores in GB
	Fish condition (coastwide)
	Small fish per large fish biomass (coastwide)
	Mid-Atlantic larval diversity
	Mid-Atlantic larval species richness
Habitat quality and ecosystem productivity	
Estuarine habitat	Chesapeake Bay Water quality goal attainment
Ocean circulation and SST	Gulf stream index
	Long term SST
	SST anomaly (Winter, Spring, Summer, Fall)
	Bottom temperature anomaly
Primary production	Chlorophyll a
	Monthly median primary production
Zooplankton	Calanus, Centropages, Pseudocalanus anomaly
	Abundance by seasons
	Small-large copepod abundance
	PP anomaly ratio

Statistical methods

Eventually, a standard statistical method will be needed to assess indicators for significant temporal trends in a State of the Ocean report. We suggest an approach using generalized additive models (GAMs) (Hastie and Tibshirani 1986) to describe potentially non-linear trends in indicator time series and identify time periods (i.e., years) of significant increasing or decreasing trends. Overall, the approach is not very complicated and lends itself to simple visual presentation that can be easily understood (i.e., see above figures with red showing decreases, blue showing increases).

In this report we fit GAMs to calculated indicators (i.e., 1 value per year) with the following formula:

$$\textit{Indicator} = \text{intercept} + \text{s}(\textit{year}) + \text{E}$$

where s(year) is a smoothing function applied to year, and E is the error term. The GAMs were fit using the "mgcv" package in the R statistical programming environment (Wood 2001; R Development Core Team 2013). We used the default thin plate regression spline option (bs = "tp") and restricted maximum likelihood to fit the models (method = "REML"). A benefit of GAM modeling is that the fitting procedure (once k is selected, see below) is

''completely automatic, i.e., no detective work is needed on the part of the statistician" (Hastie and Tibshirani 1986).

Therefore, the complexity and fit of the trendline is defined by the dataset, not the user. This is a desirable feature for the initial review and presentation of many indicator time series. As second desirable feature of GAMs is that when a trend is approximately linear, the fitted function will have an effective degree of freedom (edf) value near 1, and when it is a higher order non-linear relationship the edf values will increase (> 1). The edf value, then, is a useful metric of how complex the underlying pattern in the data is.

To determine time periods (i.e., years) statistically significant trends in the indicators, we performed an analysis of the first derivatives of the fitted GAM functions. The first derivates were calculated with the method of finite differences by using 200 evenly spaced values along the fitted trend and calculating uncertainty estimates to evaluate statistical significance ($p < 0.05$) (Curtis and Simpson 2014; Santibáñez et al. 2018). Where the slope (i.e., rate of change of the GAM function) is not-significantly different from zero the biological interpretation is that the indicator is stable, if it is above or below zero the indicator is significantly trending in the upward or downward direction. Periods of significant change can be highlighted in the plots by using different color lines (Figure A1).

One can impose a limitation to the complexity of the fit by specifying a maximum K value. This limits the 'wigliness' of the fit. The results of changing K and the p value for the significance of the trend is depicted in Figures A1 and A2. If K is constrained, oscillating patterns in the data are not well fit with the GAM (i.e., bottom row of figure A1).

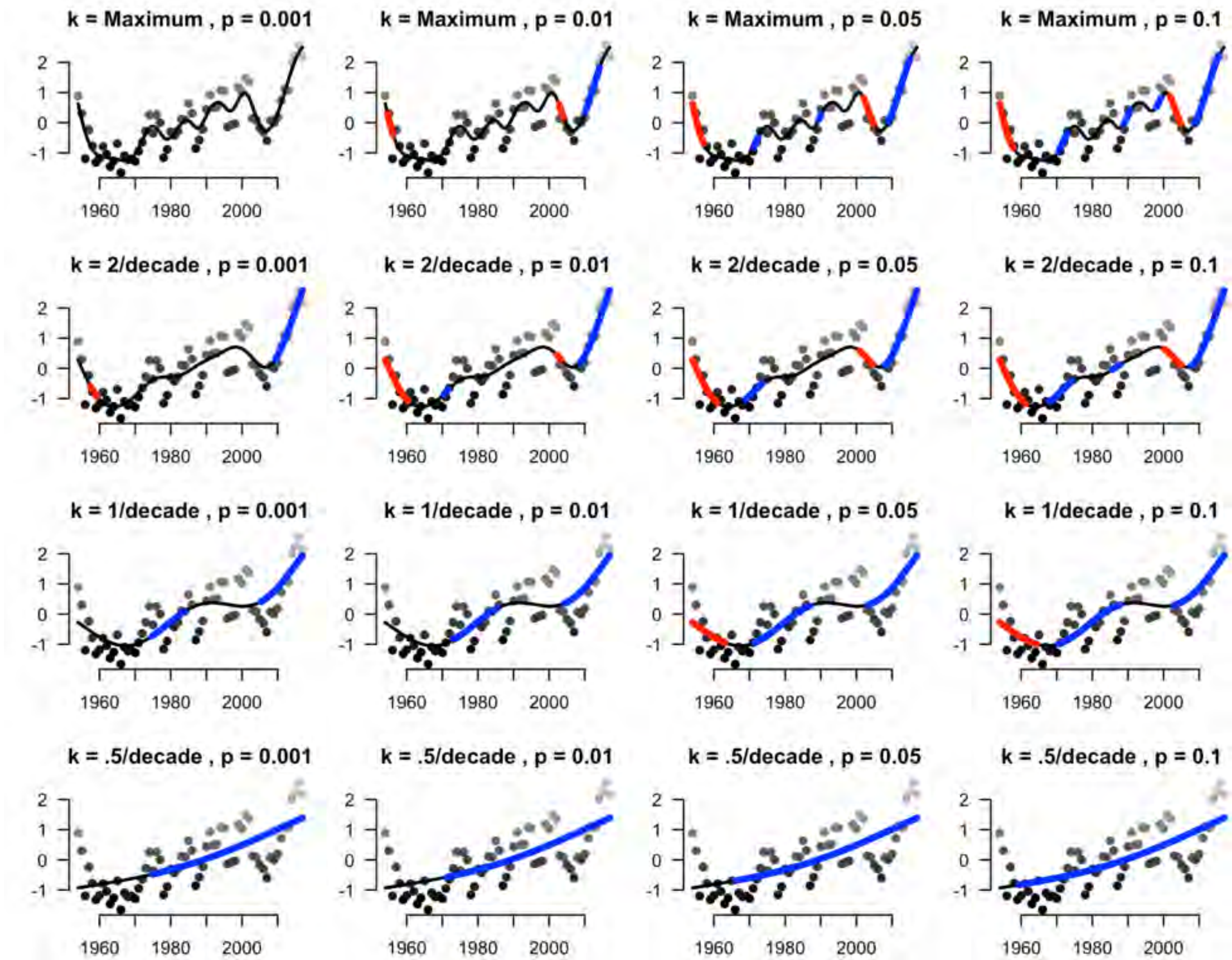


Figure A1. The AMO index shown with GAMs and significant periods of change shown in red or blue. The 25 plots show the effects of changing K and p. A maximum k is one where (potentially) as many splines can be fit as the number of years in the dataset; we also show GAMs with k set at 2 per decade (i.e., potentially one spline for every five years), 1 per decade, and 1 per two decades. Along columns, the p values are changed.

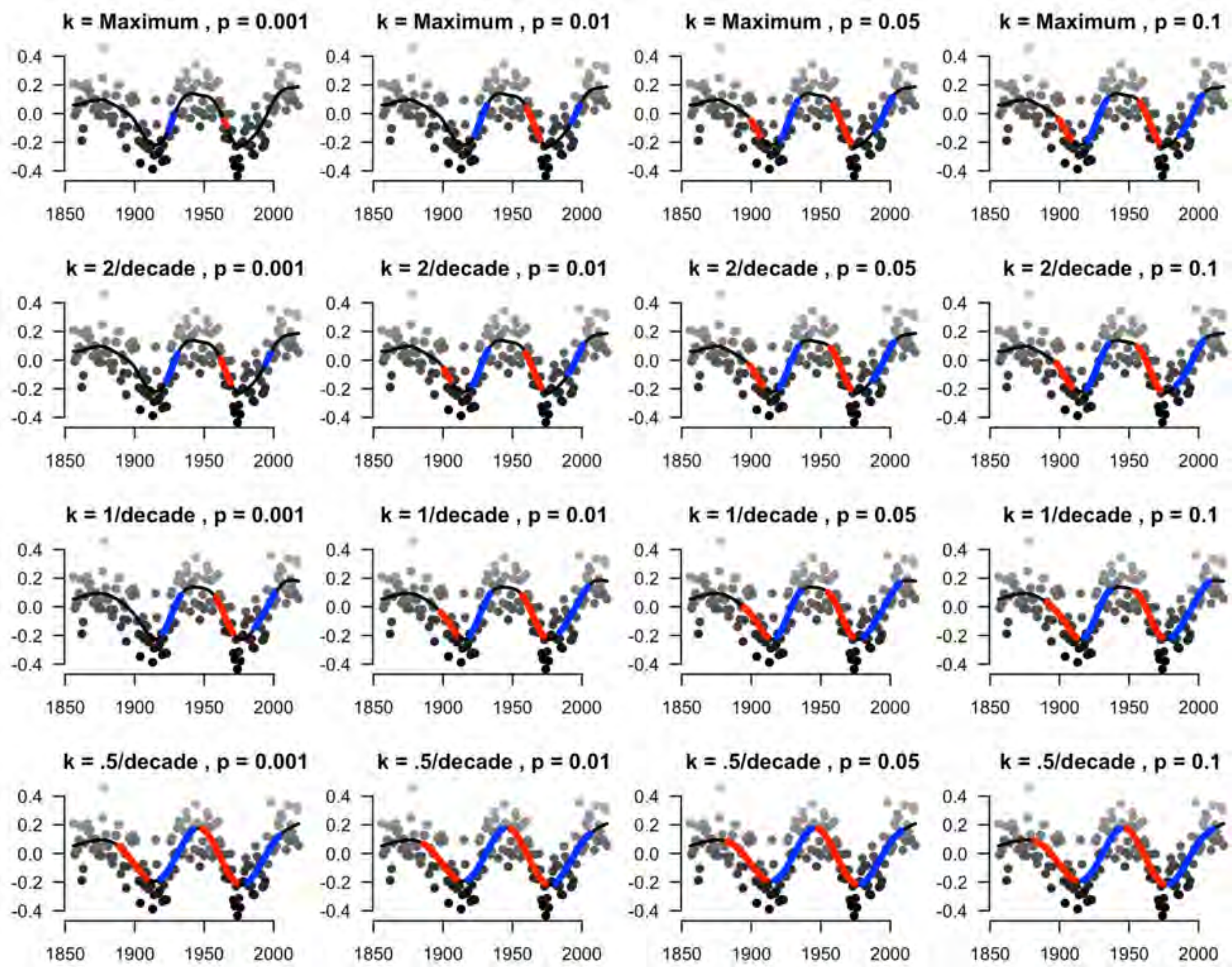


Figure A2. The AMO index shown with GAMs and significant periods of change shown in red or blue. The 25 plots show the effects of changing K and p. A maximum k is one where (potentially) as many splines can be fit as the number of years in the dataset; we also show GAMs with k set at 2 per decade (i.e., potentially one spline for every five years), 1 per decade, and 1 per two decades. Along columns, the p values are changed.