Shifts in North Sea forage fish productivity and potential fisheries yield

Worsøe Clausen, Lotte; Rindorf, Anna; van Deurs, Mikael; Dickey-Collas, Mark; Hintzen, Niels T.

Published in:
Journal of Applied Ecology

Link to article, DOI:
10.1111/1365-2664.13038

Publication date:
2018

Document Version
Peer reviewed version

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
TITLE: Shifts in North Sea forage fish productivity and potential fisheries yield

AUTHORS: Lotte Worsøe Clausen\textsuperscript{1,2}, Anna Rindorf\textsuperscript{2}, Mikael van Deurs\textsuperscript{2*}, Mark Dickey-Collas\textsuperscript{1,2} and Niels T. Hintzen\textsuperscript{3}

AUTHOR AFFILIATION:

\textsuperscript{1}International Council for the Exploration of the Sea (ICES), H. C. Andersens Boulevard 44-46, DK 1553 Copenhagen V, Denmark.

\textsuperscript{2}National Institute for Aquatic resources, Technical University of Denmark, Kemitorvet 1, DK-2800 Lyngby, Denmark.

\textsuperscript{3}Wageningen Marine Research, PO Box 68, 1970 AB IJmuiden, The Netherlands.

* Corresponding author: Mikael van Deurs, National Institute for Aquatic resources, Technical University of Denmark, Kemitorvet 1, DK-2800 Lyngby, Denmark; mvd@aqua.dtu.dk

RUNNING TITLE: Fish stock productivity and fisheries management
(1) Forage fish populations support large scale fisheries and are key components of marine ecosystems across the world, linking secondary production to higher trophic levels. While climate-induced changes in the North Sea zooplankton community are described and documented in the literature, the associated bottom-up effects and consequences for fisheries remain largely unidentified.

(2) We investigated the temporal development in forage fish productivity and the associated influence on fisheries yield of herring, sprat, Norway pout, and sandeel in the North Sea. Using
principal component analysis, 40 years of recruitment success and growth proxies were analysed to reveal changes in productivity and patterns of synchrony across stocks (i.e. functional complementarity). The relationship between forage fish production and *Calanus finmarchicus* (an indicator of climate change) was also analysed. A population model was used to demonstrate how observed shifts in productivity affected total forage fish biomass and fisheries yield.

(3) The productivity of North Sea forage fish changed around 1993 from a higher average productivity to lower average productivity. During the higher productivity period, stocks displayed a co-variance structure indicative of functional complementarity. *C. finmarchicus* was positively correlated to forage fish recruitment, however, for growth, the direction of the response differed between species and time periods. Maximum Sustainable Yield (MSY) and the associated fishing mortality ($F_{msy}$) decreased by 33-68% and 26-64%, respectively, between the higher and lower productivity periods.

(4) *Synthesis and applications.* The present results demonstrate that fisheries reference points for short-lived planktivorous species are highly dynamic and respond rapidly to changes in system productivity. Furthermore, from an ecosystem-based fisheries management perspective, a link between functional complementarity and productivity, indicate that ecosystem resilience may decline with productivity. Based on this, we advise that system productivity, perhaps monitored as forage fish growth, becomes an integral part of management reference points; in both single species and ecosystem contexts. However, to retain social license of biological advice when fish catch opportunities are reduced, it is crucial that shifts in productivity are thoroughly documented and made apparent to managers and stakeholders.
KEY WORDS: Maximum sustainable yield; fisheries management; bottom-up effects; Portfolio effect; functional complementarity; small pelagic fisheries, forage fish, system productivity, fisheries yield; recruitment

INTRODUCTION

Changes in ecosystems are often gradual and smooth and follows for example large-scale climatological fluctuations. However, sudden rapid changes are not uncommon (Scheffer et al. 2001), and the detection of such shifts is pivotal in responsive ecosystem based management (Pikitch et al 2004; Casini et al. 2010; King et al. 2015). In the North Sea, the zooplankton community underwent considerable changes in the late 1980s and early 1990s, with the decline of the subarctic calanoid copepod *Calanus finmarchicus* being one of the most well documented examples (Beaugrand et al. 2003; Pitois and Fox 2006; Defriez et al. 2016). While several authors have interpreted these changes as evidence of a major system wide regime-shift (Beaugrand 2003; McQuatters-Gollop et al. 2007; Barrange et al. 2008), resulting bottom-up effects on the fish stocks are still not well understood (Reid et al. 2001, Ware and Thompson 2005).

The North Sea has a combined standing stock of commercially exploited forage fish equalling c. 7.6 million tonnes and annual yields averaging 1.5 million tonnes in the 1990s (Dickey-Collas et al. 2013; Engelhard et al. 2013). The forage fish biomass includes species such as herring (*Clupea harengus*), sprat (*Sprattus sprattus*), Norway pout (*Trisopterus esmarkii*), and lesser sandeel (*Ammodytes marinus*), which together consume c. 110 million tonnes of zooplankton
annually (derived from data in Greenstreet et al. 1997; van Deurs et al. 2013a; ICES 2015a,b).

The combination of a central position in the food chain and large responsiveness to climate change and environmental variability makes the management of forage fisheries a major challenge (Smith and Jarre 2011; Fauchald et al. 2011; Fiechter et al. 2015).

Most North Sea fisheries are presently managed using harvest control rules targeting the fisheries Maximum Sustainable Yield (MSY) (Kell et al. 2005; Mohn and Chouinard 2007). In the EU, MSY is defined as the highest yield that can be taken on average from a stock under existing average environmental conditions without significantly affecting reproduction (EU 2013). If all forage fish stocks in the North Sea show synchronous responses to climate-driven changes in the zooplankton community, this will impact the combined forage fish biomass and MSY of these stocks. While scientific single-species advice on catch opportunities will likely reflect this within a span of less than five years, the decreasing catch potential of the individual stocks may not be as easily accepted by stakeholders (Kelly et al. 2017). On the other hand, if the responses are asynchronous (i.e. inverse correlations between species), the overall biomass and yield may remain relatively stable over time, providing a basis an multi-species approach to management and likely enhancing acceptance among stakeholders. These stabilizing mechanisms are known as the portfolio effect (Tilman 1998) or functional complementarity (Lindegren et al. 2016) and have been demonstrated in several ecosystems around the world (Lluck et al. 1992; Lindegren et al. 2016).

In the present study, we analysed time series of two productivity metrics, length anomalies and recruitment success, for the major forage fish stocks in the North Sea, and tested the hypothesis that overall forage fish productivity has changed alongside with the changes in the zooplankton community described in the literature. Alternatively, some stocks have responded negatively and
others positively (i.e. functional complementarity). Lastly, we simulated stock development to
demonstrate implications for MSY and the potential forage fish biomass in the North Sea.

MATERIALS AND METHODS

Metrics of fish productivity
The analyses combined data from five major forage fish stocks in the North Sea: herring,
Norway pout, sprat, northern sandeel, and southern sandeel (the southern and northern sandeel
stock correspond to stock assessment area 1 and 3, respectively; see ICES 2015b). Growth and
recruitment success were used as metrics of fish productivity (Kell et al. 2016). A pre-analysis
revealed that annual growth effects exceeded cohort effects. Hence, growth was derived from
average annual length anomalies of consistently sampled ages, rather than as cohort specific
estimates. Length at age data for herring, sprat and Norway pout were derived from the ICES
International Bottom Trawl Survey Database (February 1967 to 2014). The surveys contributing
to this database use demersal trawls to sample the abundance of fish. The survey catch rates are
used in both herring, sprat and Norway pout stock assessments in the area, and for herring and
sprat, the results from this survey shows high consistency with data from acoustic surveys when
used to estimate abundance indices. In contrast to e.g. acoustic surveys, the sampling does not
target aggregations of fish and hence avoids possible bias if high aggregations tend to contain
e.g. smaller individuals. Total catch and length composition of the majority of species is
recorded in each haul. Haul duration generally varies between 15 and 35 minutes, and on average
two trawl hauls are conducted within each ~30×30 nm statistical rectangle in a grid covering the

5
entire North Sea, though not all rectangles were monitored prior to the 1980s. The length of the
time series differed between the three stocks as sprat and Norway pout were not aged up until
1976 and 1984, respectively.

Age-length keys were estimated using continuation-ratio logits (Rindorf and Lewy 2001;
Gerritsen et al. 2006; Rindorf et al. 2008). All fish ages were measured in years from January 1
in the year of spawning. The mean length of each stock $s$ at age $a$ in year $y$, $(l_{a,y,s})$, was
calculated by estimating average length of all fish caught in a given age group. Unfortunately,
the data are often recorded in plus-groups for older ages and hence, age segregated data for older
ages are not consistently reported.

Sandeel are not representatively caught in the trawl surveys, and therefore length at age data for
sandeel were derived from fish sampled from the commercial fishery in the second quarter,
which roughly captures the fishing season. The samples are a combination of on-board samples
(samples organized by control agencies) and port samples. Sample size is approximately one kg
taken randomly from the catch. Mean length at age in each statistical rectangle was estimated by
combining sampled length distributions with age-length keys. Length at age in a given year was
estimated as the average mean length in the commercial catches. Samples were available from
1978 to 2014 for the southern sandeel stock (except from 1990) and from 1982 to 2014 for the
northern stock. These procedures provided yearly estimates of mean length at age of northern
and southern sandeel, sprat and Norway pout of ages 2 to 4 and herring of ages 2 to 5. Fish of
age 0 and 1 were not included as these were only partially recruited to the survey sampling gear.
The estimated mean lengths at age were then used to estimate the annual mean length anomalies.
Length at ages 2+ was highly correlated between ages within years for all stocks (Pearson´s
product moment correlation: coefficients > 0.60, $p < 0.001$ for all ages and stocks). This
indicates that a common factor is influencing the length of all ages in a given year. Reducing complexity of data by computing the anomalies was therefore desirable. The mean length anomaly of fish, \( \lambda_{y,s} \), was computed as:

\[
\lambda_{y,s} = \left( \frac{l_{2,y,s} - \overline{l}_{2,s}}{\overline{l}_{2,s}} + \frac{l_{3,y,s} - \overline{l}_{3,s}}{\overline{l}_{3,s}} + \cdots + \frac{l_{\text{max},y,s} - \overline{l}_{\text{max},s}}{\overline{l}_{\text{max},s}} \right)/(\text{max} - 1),
\]

where \( l_{a,y,s} \) is the length at age \( a \) in year \( y \) of stock \( s \), \( \overline{l}_{a,s} \) is the average length at age \( a \) of stock \( s \) across all years and \( \text{max} \) denotes the oldest age included in the analysis for stock \( s \) (\( \text{max} \) is 5 for herring and 4 for other stocks; \( \text{max} - 1 \) is used in the denominator, since age 1 fish is not included in the analysis, as explained above).

Data on recruitment (\( R \)) and spawning stock biomass (\( S \)) were derived from stock assessments (ICES 2015a,b). Recruitment success was estimated as \( \ln(R/S) \) and the annual anomaly of all-time series was estimated as described for length at age. However, to avoid erroneous conclusion, the type of analysis was repeated with residuals from a Ricker relationship (see online supporting information).

**Statistical analysis of data time series**

Data were analysed both within each stock and across stocks. Analyses carried out within each stock served to identify time trends in length anomaly and recruitment success, and was conducted using Pearson’s product moment correlation tests. The across stock analyses focused on the more general correlation structure and was conducted using principle components analyses (PCA) in line with the approach described in Lindegren et al. (2016).
Initial spectral analyses showed periodicities in data. Data were therefore also analysed as three year running averages to differentiate between correlation patterns resulting from multi-annual trends and correlation patterns resulting from year-to-year fluctuations. Applying running averages, rather than the underlying annual values, can potentially introduce an artificial autocorrelation in the data reducing the effective number of degrees of freedom below the number of observations. To remove this source of bias the degrees of freedom was reduced in tests in accordance with the method described in Pyper and Peterman (1999).

As PCA estimates a common correlation structure throughout the data set, we analysed whether there were significant changes in the correlation structure over time. The analysis was used to identify the most likely single year breakpoint in the time series (across stocks). The breakpoint analysis was performed by estimating the parameters in the following model:

\[ X[t] \sim N(Z[t], I \times s^2) \]

Where the observation at time \( t \) is \( X[t] = (x[t1], \ldots, x[t5]) \), with \( x[t1] \) being the observed productivity value for stock 1 at time \( t \) and so forth. \( Z[t] = (z[t1], \ldots, z[t5]) \) denotes the underlying process describing productivity and \( I \) is a matrix with 1 in the diagonal and 0 elsewhere, and \( s^2 \) is the variance parameter. The underlying process model for \( Z \) is assumed to be a multivariate random walk with a common correlation parameter \( \rho_1 \) for all pairs up to time \( \tau \) after which point the common correlation parameter is \( \rho_2 \). The log likelihood of each possible breakpoint \( \tau \) is then evaluated to determine if there is a clear change point in the correlation structure.

Lastly, a PCA analysis for each productivity metric was conducted. Based on the results from the breakpoint analysis described above, the PCA for length anomalies was split into the period
before and including 1993 and the period after 1993. The contribution of the different stocks to
the first and second principal component (PC1 and PC2, respectively) and the time trends of PC1
and PC2 were examined.

Zooplankton

To investigate if trends in forage fish productivity were correlated to changes in the zooplankton,
the PC1 from all three PCAs was related to an annual *Calanus finmarchicus* North Sea index
using Pearson’s product moment correlation tests. *C. finmarchicus* abundance (numbers m$^{-3}$) was
derived from the continuous plankton recorder survey (CPR, Batten et al. 2003; Reid et al. 2003)
and provided by the Sir Alister Hardy Foundation for Ocean Science, SAHFOS (Richardson et
al. 2006; Stevens and Johns 2017), where it was delivered in the CPR standard format of
monthly values for each of the four CPR standard areas within the North Sea (c1, c2, d1, and d2).
The annual index was subsequently calculated by first deriving the median value across all
months and then the median across all areas.

The effect of fish stock productivity on the fisheries and potential forage fish biomass

In order to demonstrate how changes in recruitment and growth impact MSY, the associated
mortality $F_{msy}$ and the potential forage fish spawning stock biomass (i.e. spawning stock biomass
at a fishing mortality of zero), we simulated stock dynamics for the period before and after 1993
(see estimation of time series breakpoints above). Calculations were based on a dual simulation
model where the yield resulting from a given fishing mortality is determined based on a
'perceived stock' and subsequently removed from a 'true' stock, which receives a stochastic
annual recruitment. The model framework is in accordance with the models described in Butterworth and Punt (1999) and Butterworth (2007) and ICES (2013). F_{msy} is defined as the fishing mortality resulting in the highest long-term average yield that can be obtained while at the same time ensuring that probability of the spawning stock (S) falling below the lower limit biomass reference point, represented by B_{lim}, does not exceed 5% (ICES 2014). Simulations of stocks development when managed according to F_{msy} were carried out for a period of twenty years and repeated 1000 times. The probability of achieving a spawning stock below B_{lim} (adopted from the ICES reports (ICES 2015a,b)) was calculated as the number of simulated years with S<B_{lim} divided by the total number of simulated years. To demonstrate the effects of reduced productivity on the potential combined forage fish biomass, we also simulated the potential combined forage fish spawning stock before and after 1993 in a no-fishing scenario (F=0). Simulations were carried out for each stock for the two periods identified in the breakpoint analysis described above (1983-1993 and 1994-2013). Period-specific stock-recruitment relationships and mean weight at age were calculated from values in the ICES reports for the respective stocks (ICES 2015a,b). For simplicity, we applied stock-recruitment relationships of the hockey-stick type (Barrowman and Myers 2000) with time invariant breakpoints at B_{lim}. Hence, only the upper level of the hockey stick and mean weight at age varied between the two periods (before and after 1993). As a simplification and to avoid making prejudiced assumptions, correlations between stocks were not accounted for. Hence, the focus of the simulations were merely to establish the overall change between the two periods defined above. A more detailed description of the model and input data can be found in the online supporting information.
RESULTS

Time trends in productivity within stocks

All stocks displayed decreasing trends in production, reflected in either length anomalies, recruitment success or both. Length anomalies showed a significant declining trend in four out of five stocks, sprat being the exception (Table 1, Fig. 1). Using 3-year running averages, significant trends were detected for herring and northern sandeel, although all stocks but southern sandeel showed a tendency toward lower recruitment success late in the time series (Table 1, Fig. 2). Overall these results demonstrate a shift from a period of high forage fish productivity (across all species and stocks) to a period of considerably lower productivity.

Co-variation patterns in productivity across stocks

The across stock co-variation structure in the length anomaly data changed significantly over time, with the most likely breakpoint being in 1993. In contrast, the co-variation structure of recruitment success remained the same throughout the time series. The first principal component (PC1) for length anomaly up to 1993 explained most of the variation (PC1: 80%, PC2: 11%) (Table 2), hence, the across stock correlation structure was strong. Up till 1993 PC1 was positively correlated to length anomaly of sprat and the two sandeel stocks and negatively correlated to herring and Norway pout.

After 1993, the across stock correlation structure in the length anomaly data was weaker, as indicated by the relatively low proportion of variation explained by PC1 (PC1: 37%, PC2: 31%) (Table 2). PC1 was now negatively correlated to length anomaly of sprat and the northern
sandeel stock but positively correlated to herring, Norway pout and the southern sandeel stock. Hence, a high value of PC1 prior to 1993 corresponds to high growth of sprat and sandeel of both stocks and a low growth of herring and Norway pout whereas a high value of PC1 after 1993 corresponds to low growth of sprat and northern sandeel and a high growth of herring, Norway pout and southern sandeel. To make the results from the two periods more intuitively comparable, we used PC1 in analyses of the period before 1993. Using this method, a high value of PC1 corresponds to a high growth of herring and Norway pout and a low growth of sprat and northern sandeel in both periods.

PC1 for recruitment success explained 42% of the total variation and PC2 added another 33%. PC1 was positively correlated to the recruitment success of all stocks, although the correlation with southern sandeel was weak (Table 2). PC2 was positively correlated to southern sandeel and negatively correlated to sprat, indicating that years of high recruitment success for one of these stocks matched up with years of poor recruitment success for the other. PC1 showed an overall significant decreasing trend over time (Pearson’s product moment correlation: correlation -0.76, p=0.002), indicating a general decrease in forage fish recruitment success. For further information about correlations between individual stocks can be found in Table S3 in online supporting information.

Zooplankton

The annual *C. finmarchicus* index was positively correlated to PC1 for recruitment (c = 0.41, t = 2.32, df = 26, p = 0.029 ), PC1 for growth (i.e. length anomalies) in the late period (after 1993) (c = 0.51, t = 2.54, df = 18, p = 0.021), and PC1 for growth in the early period (up till 1993) (c = 0.59, t = -2.52, df = 12, p = 0.027) (Fig. 3). Hence, a higher density of *C. finmarchicus* was
consistent with higher growth of herring and Norway pout and lower growth of sprat and northern sandeel in both periods.

Productivity and maximum sustainable yield
Weight at age in the catch decreased by 13% on average across age classes and stocks between the early period and the late period (before and after 1993), and mean recruitment dropped nearly 30% on average (Table 3). These reductions led to a decrease in MSY and $F_{MSY}$ of around 47% and 46%, respectively (ranging from 26 to 69%, Table 3). In comparison the reduction in the potential forage fish spawning stock biomass (given a fishing mortality of zero) amounted to c. 25% (Fig. 4). The greater decrease in MSY is caused by the necessity to sustain the stock above the unchanged biomass limit reference point $B_{lim}$.

DISCUSSION
Our results revealed that a shift in North Sea forage fish productivity (reflected in recruitment to the stocks or the growth of individuals) occurred in succession with the well documented shifts in the plankton community (Pitois and Fox 2006; McQuatters-Gollop et al. 2007). The year of 1993 represented the statistically most likely single-year breakpoint between a period of higher correlation in forage fish productivity and a subsequent lower correlation. The shift in productivity around 1993 was consistent across stocks. Hence, the functional complementarity hypothesis was not supported over the entire time period. However, in the period before 1993, productivity (represented by growth) was inversely correlated between two different subsets of forage fish stocks (sprat and sandeel vs. herring and Norway pout), indicating
that functional complementarity may have existed when productivity was high. Different species within the same functional group, respond in different ways to environmental drivers (Chavez et al. 2003; Lindegren et al. 2016). The stocks experience different prey fields, prefer different prey (Last 1987; Albert 1994; Raab et al. 2012; van Deurs et al. 2013b), and may be expected to display different responses to changes in food availability (Defriez et al. 2006). Prior to 1993, years favouring growth of herring and Norway pout were different from those favouring sandeel and sprat, creating a synchrony among some segments of the forage fish community and an asynchrony among others. Thus in periods with high forage fish productivity, the different spatial distributions and life histories may result in divergent growth patterns, facilitating overall stability in the forage fish community (i.e. functional complementarity). However, below a certain threshold, the functional complementarity was no longer be sustained (i.e. after 1993).

The simulated population dynamics revealed that the reduction in forage fish productivity around 1993 roughly halved the maximum sustainable yield (MSY) and $F_{\text{msy}}$ (the fishing mortality associated with MSY) and reduced the potential forage fish spawning stock biomass by c. 25%. This drop in productivity is comparable to a study conducted on North Sea cod (Kell et al. 2016). However, Shephard et al. (2014) found that the forage fish biomass in the North Sea has remained relatively stable between 1983 and 2012, it may therefore be so that productivity changes amongst the forage fish stocks after all have been successfully counteracted by appropriate management action (Dickey-Collas et al. 2014). Recruitment and spawning stock is often poorly correlated (Szuwalski et al. 2015; Kell et al. 2016). Hence, we chose to use a simple hockey stick relationship for the simulations, rather than making assumptions about more complicated relationships (Chen and Irvine 2001). Another simplification made here was that
reference points were kept constant across the period of change, as opposed to having dynamic reference points (Punt et al. 2013).

Our analysis suggested a positive relationship between C. finmarchicus and forage fish recruitment. The importance of C. finmarchicus eggs and nauplii as food for fish larvae has been highlighted in relation to sandeel and cod (Arnott and Ruxton 2002; Beaugrand et al. 2003, van Deurs et al. 2009). For sprat and Norway pout, recruitment success has not formerly been linked to C. finmarchicus. However, sprat recruitment is reportedly positively correlated to temperature (Baumann et al. 2006). Hence, variation in temperature in the southern North Sea could potentially explain the inverse patterns in recruitment success of sandeel in the southern area and sprat. The substantial downward shift in key elements of the plankton (Pitois and Fox 2006; McQuatters-Gollop et al. 2007) and stock production across a range of zooplanktivorous species (results presented here), seem to indicate a declining primary production (Ware et al. 2005). However, other studies point to changes in zooplankton species composition as the cause of reduced fish growth (Beaugrand et al. 2003; van Deurs et al. 2015). Factors such as spawning site diversity or the frequency of unlikely events governing extreme recruitment outburst (Pepin 2015; Lowerre-Barbieri et al. 2016) have also been proposed as drivers of variation in fish stock production. However, such mechanisms would not be expected to influence growth as found in the present study.

With the growing interest in implementing ecosystem based fisheries management (Pikitch et al. 2004), the regime-shift subject is gradually becoming an integral part of the scientific endeavors in fisheries biology and management (Anderson and Piatt 1999; Barange et al. 2008, King et al. 2014). Several studies have argued that the North Sea underwent a major regime-shift in the late 1980s, manifested first as changes in primary and secondary production (Reid et al. 2001;
McQuatters-Gollop et al. 2007). This is just a few years earlier than the break point identified in the present study and therefore credible that the shift in forage fish production is the consequence of such a regime-shift. However, based on the time series length presented here and the large inter-annual fluctuations in weight-at-age and recruitment, that even after converting to three years running averages were still distinctive for some stocks, it cannot be concluded if an abrupt regime-shift like transition from one state to another has truly taken place, or if the change we observed reflects climatic oscillations on a multi-decadal time-scale, as reported for anchovy-sardine systems (Chavez et al. 2003). Such differentiation may, however, be worth pursuing in the future, as it would be a major asset in models of resource fluctuations and execution of ecosystem based management actions (King et al. 2014).

Here we have shown that what appears to have been climate induced changes in productivity in the North Sea had substantial effects on MSY and F_{msy}. While, similar findings have previously been presented for cod, this is the first study to show how productivity in an entire forage fish community can shift in less than a decade, reducing the sustainable catches opportunities by a factor of two across all major forage fish stocks (Kell et al. 2005; Mohn and Chouinard 2007; Morgan et al. 2014). In addition, this study reveals a link between productivity and functional complementarity among stocks (the portfolio effect), supporting one of the recommendations put forward in Dickey-Collas et al. (2014), that a forage fish community cannot be managed as an overall resilient pool of biomass where one species take over the niche when another is in decline. Hence, we advocate against having a long-term maximum on the total removal of forage fish (i.e. a community MSY), which is in contrast to what has been suggested for some other systems (e.g. Lluch et al. 1992). At least, system productivity ought to be integrated into any long-term management plan and fisheries reference points should be kept dynamic, irrespective
of whether it is made in a single species or multi species context. Furthermore, it is here demonstrated how already existing time series of forage fish growth and reproductive output can potentially be used as indicators of system productivity, and last but not least, to ensure social license, when catch opportunities are reduced as a result of declining productivity, documentation and communication to stakeholders is critical (Kelly et al. 2017).

ACKNOWLEDGEMENTS

This study received funding from (1) the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement MYFISH number 289257 (AR, LWC, NH), (2) COFASP ERA-NET project, which has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 321553 (GOFORIT project), and by the national funding agency of Denmark (Innovationsfonden) (MVD), (3) the Danish council of independent research DFF-4002-00114 (MVD), and (4) the European Fisheries and Maritime Fund (33113-B-15-002, Minestry of Environment and Food in Denmark) (MVD, LWC).

Authors’ contributions

Author Contributions statement: – All authors were involved in conceiving the ideas and designing methodology; AR and MVD collected and analysed the data; LWC, AR and MVD led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.
Data accessibility

Data are available in the Dryad Digital Repository doi:10.5061/dryad.tq1f7

REFERENCES


van Deurs, M., Christensen, A. and Rindorf, A. (2013a) Patchy zooplankton grazing and high energy conversion efficiency – Ecological implications of sandeel behavior and strategy. *Marine Ecology Progress Series* 487, 123-133


Additional Supporting Information may be found online in the supporting information tab for this article.
Table 1. Time trends. Pearson’s product moment coefficients and p-values (in parenthesis) are provided for the correlations of time versus recruitment success and length anomalies, respectively. Tests was made for both annual values and 3-year averages (see material and methods). Significant correlations are in bold font.

<table>
<thead>
<tr>
<th>Stock</th>
<th>Recruitment success</th>
<th>Length anomaly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trend in annual values</td>
<td></td>
</tr>
<tr>
<td>Herring</td>
<td>-0.21 (0.092)</td>
<td>-0.45 (0.001)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
<td>-0.75 (0.009)</td>
</tr>
<tr>
<td>n</td>
<td>67</td>
<td>49</td>
</tr>
<tr>
<td>Norway pout</td>
<td>-0.14 (0.4616)</td>
<td>-0.51 (&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
<td>-0.47 (0.0616)</td>
</tr>
<tr>
<td>n</td>
<td>30</td>
<td>41</td>
</tr>
<tr>
<td>Southern sandeel</td>
<td>Trend in annual values</td>
<td>0.02 (0.894)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
<td>0.10 (0.688)</td>
</tr>
<tr>
<td>n</td>
<td>31</td>
<td>37</td>
</tr>
<tr>
<td>Northern sandeel</td>
<td>Trend</td>
<td>-0.20 (0.282)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
<td>-0.52 (0.037)</td>
</tr>
<tr>
<td>n</td>
<td>31</td>
<td>37</td>
</tr>
<tr>
<td>Sprat</td>
<td>Trend in annual values</td>
<td>-0.12 (0.468)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
<td>-0.41 (0.161)</td>
</tr>
<tr>
<td>n</td>
<td>40</td>
<td>41</td>
</tr>
</tbody>
</table>
Table 2. Principal component contributions of different stocks. Contribution indicates the correlation between the species specific metric and the principal component. Large contributions are present where the principal component closely follows the species specific metric, small contributions indicate that the principal component does not reflect variation in the species specific metric.

<table>
<thead>
<tr>
<th>Productivity metric</th>
<th>Stock</th>
<th>PC 1 contribution</th>
<th>PC 2 contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recruitment success</strong></td>
<td>Herring</td>
<td>0.44</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>Norway pout</td>
<td>0.55</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Sandeel (southern)</td>
<td>0.11</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Sandeel (northern)</td>
<td>0.61</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>Sprat</td>
<td>0.35</td>
<td>-0.59</td>
</tr>
<tr>
<td><strong>Length anomaly before 1993</strong></td>
<td>Herring</td>
<td>-0.48</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>Norway pout</td>
<td>-0.40</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Sandeel (southern)</td>
<td>0.49</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>Sandeel (northern)</td>
<td>0.40</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>Sprat</td>
<td>0.46</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Length anomaly after 1993</strong></td>
<td>Herring</td>
<td>0.56</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Norway pout</td>
<td>0.21</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Sandeel (southern)</td>
<td>0.46</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Sandeel (northern)</td>
<td>-0.51</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Sprat</td>
<td>-0.42</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Table 3. Relative change in weight at age, recruitment, MSY, and $F_{msy}$ for each stock. Values are given as % change between the period before 1993 and the period after 1993 (minus indicate a negative change). Data on mean weight at age and recruitment were adopted from the stock assessment reported by ICES and are the same data applied in the calculations of MSY and $F_{msy}$ (see material and methods).

<table>
<thead>
<tr>
<th>Stock</th>
<th>Weight at age</th>
<th>Recruitment</th>
<th>MSY</th>
<th>Fmsy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandeel (southern)</td>
<td>-26%</td>
<td>-23%</td>
<td>-56%</td>
<td>-50%</td>
</tr>
<tr>
<td>Sandeel (northern)</td>
<td>-24%</td>
<td>-52%</td>
<td>-68%</td>
<td>-64%</td>
</tr>
<tr>
<td>Sprat</td>
<td>-6%</td>
<td>-9%</td>
<td>-34%</td>
<td>-48%</td>
</tr>
<tr>
<td>Herring</td>
<td>-3%</td>
<td>-27%</td>
<td>-33%</td>
<td>-26%</td>
</tr>
<tr>
<td>Norway pout</td>
<td>-8%</td>
<td>-29%</td>
<td>-45%</td>
<td>-41%</td>
</tr>
</tbody>
</table>
Fig. 1. Development in length analomy (solid line) of herring (a), sprat (b), Norway pout (c), southern sandeel (d), northern sandeel (e) and all stocks together (f). Broken lines show averages before and after 1993.

Fig. 2. Development in recruitment success (solid line) of herring (a), sprat (b), Norway pout (c), southern sandeel (d), northern sandeel (e) and all stocks together (f). Broken lines show averages before and after 1993.

Fig. 3. Correlations between *Calanus finmarchicus* and the first principal component of recruitment (a) and growth (b,c). Since the PCA analysis of growth was divided into an early period and a late period (see Material and Methods), the correlation analyses were also divided in two (b,c). All correlations were significant according to Pearson's product-moment correlation test (see Results).

Fig. 4. Model simulations of how the combined forage fish spawning stock biomass would have developed even without fishing (F=0) (i.e. the potential forage fish biomass). Black line is the mean of one thousand simulations and grey lines represents the standard deviations. Hence, the decline is caused by the introduction of the observed changes in stock recruitment and weight at age in the stocks into the model (see table 4). The grey field defines the transition zone between the high productivity period and the low productivity period.
Supplementary material

Estimation of MSY and $F_{MSY}$

The model tracks cohort development influenced by an age- and season-specific natural mortality ($M$) and fishing mortality ($F$) (according to the resolution applied in the corresponding ICES stock assessments). Stock numbers ($N$) are derived for discrete age groups ($a_i \in [a_0, a_1 \ldots a_{max}]$) and updated at the beginning of each time step ($t_i \in [t_1, t_2, \ldots t_{max}]$ (i.e. $t_{max}$ equals 4 if the stock is analyzed in quarterly time steps) in each year ($y_i$):

$$N_{a_0, t_1, y_1} = \begin{cases} \frac{N_{\text{initiate}, t_1}}{1} \times f(S_{y_1})^{e^{\frac{\sigma_R}{2}}} \times \text{NORM}(0,1), & t_1 = t_1 \land a_0 = a_0 \land y_i > y_1 \\ \frac{N_{a_1, t_{max}, y_{i-1}}}{1} \times e^{-M_{a_1} - F_{a_1}}, & t_i = t_1 \land a_i > a_0 \land y_i > y_1 \\ \frac{N_{a_{i-1}, t_i}}{1} \times e^{-M_{a_{i-1}} - F_{a_{i-1}}}, & t_i > t_1 \land a_i > a_0 \land y_i > y_1 \end{cases}$$

The youngest age group is age-0 ($a_0$). $a_0$ in $t_1$ signifies the recruitment ($R$) of young fish to the stock and is modelled as a function of spawning stock biomass ($S$). $a_{max}$ is a plus-group and covers all age groups at or above the oldest age group for which accurate data is available. $N$ is the stock number for any given age group ($a_i$), season ($t_i$), and simulation year ($y_i$). Each simulation year starts at the time of spawning, in accordance with the biology of these stocks. $N_{\text{initiate}}$ is the initial stock numbers required to initiate the simulations. $f(S)$ is the stock-recruitment function, in this case a hockey-stick function (with the horizontal bar equal to the geometric mean recruitment for the specific stock and period (as previously described) and the brake point equal to $B_{lim}$), to which a log-normal distributed error term is added to simulate environmental variation ($\sigma_R = 0.35$ was applied for all stocks). $S$ at the onset of $y_i$ is calculated from stock numbers at age, the mean weight, and proportion mature. Proportion mature were the same independent of the period and corresponded to the values applied by ICES in the current short-term projections (see ICES 2015a,b). $F$ in the model is composed of a multiplier and a
constant age- and season-specific exploitation pattern. $B_{\text{lim}}$, $M$, proportion mature, and the exploitation pattern were the same independent of the period and were adopted from ICES (2015a,b). $F$ and $M$ was implemented in the population using Pope’s approximation (Myers and Cadigan 1995), which was applied with ten sequential iteration cycles for each season and age group. The associated catch in tons was derived by summing up the numbers removed by $F$ in each sequential iteration cycle and from each age group and season multiplied by the age- and season-specific mean weights.

The ‘perceived’ stock is modelled at the beginning of the simulation year ($t_1$) as the true value plus an error term to mimic uncertainty in stock number estimates. The following equation is used to generate the ‘perceived’ stock numbers:

$$N'_{t_1|\text{max}t_1;S} = \text{NORM}(N_{t_1|\text{max}t_1;S}, \sigma_N)$$

where $\sigma_N$ is the virtual “stock assessment uncertainty” of stock number estimates ($\sigma_N = 0.3$ was applied for all stocks and both periods).

The shifts in $F_{\text{msy}}$ reported in the Table 3 in the main text is driven solely by shifts in mean weight at age and recruitment (i.e. the upper level of the $S$-$R$ hockey stick), which can be derived from data reported in ICES (2015a,b). As an example of data input, Table S1 provides the the upper level of the $S$-$R$ hockey stick from before and after 1993, calculated as the geometric mean recruitment in years where $S$ was above $B_{\text{lim}}$, and fig. S1 depicts the $S$-$R$ scatter plots adopted from the respective ICES assessment working group reports ICES (2015a,b).

$F_{\text{msy}}$ (i.e. the $F$ that leads to a probability $P$ of 0.05 of falling below $B_{\text{lim}}$) was identified by calculating $P(S<B_{\text{lim}})$ for a range of different $F$ values and for each stock and period, respectively (fig. S2).

Model assumptions were kept at a minimum to homogenize the individual sub models, such that the results reported in table 3 in the main text arrived only from the stock specific differences in weight at age and recruitment, before and after 1993.
Analyses of time trend and common patterns in recruitment residuals from a Ricker relationship

The methods used here are identical to the analysis given in the main manuscript with the exception that the recruitment indicator used was the annual residual from a Ricker relationship between spawning stock biomass ($S$) and recruitment ($R$) (Ricker 1954), rather than $\ln(R/S)$. The fitted stock recruitment relationships are shown in fig. S3. The temporal development in the residuals can be seen in fig. S4.

All stocks displayed decreasing trends in recruitment anomalies similarly to the analyses conducted using $\ln(R/S)$ in the main text (Table S2), and also when using 3-year running averages (Table S2).

PC1 for recruitment success explained 42% of the total variation and PC2 added another 28%. Similarly to the analyses conducted with $\ln(R/S)$, PC1 was positively correlated to the recruitment success of all stocks, although the correlation with southern sandeel was weak (Table S3).

PC2 was positively correlated to southern sandeel and negatively correlated to sprat, indicating that years of high recruitment success for one of these stocks matched up with years of poor recruitment success for the other. PC1 showed a significant decreasing trend over time (Pearson’s product moment correlation: correlation -0.76, $p<0.0001$), indicating a general decrease in forage fish recruitment success.

Detailed information about correlations between individual stocks can be found in Table S4. This information were not included in the main document.
Table S1. Upper levels of the S-R hockey sticks before and after 1993. Only recruitment values from years where spawning stock size was above \( B_{\text{lim}} \) was included in the geometric means. Numbers are calculated from data \((B_{\text{lim}}, S, \text{and } R)\) reported in ICES (2015a,b).
Table S2. Time trends. Pearson’s product moment coefficients and p-values (in parenthesis) are provided for the correlations of time versus recruitment anomalies. Tests were made for both annual values and 3-year averages (see material and methods). Significant correlations are in bold font.

<table>
<thead>
<tr>
<th>Stock</th>
<th>Recruitment success</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trend in annual values</td>
</tr>
<tr>
<td>Herring</td>
<td>-0.49 (0.0012)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
</tr>
<tr>
<td></td>
<td>-0.60 (&lt;0.0001)</td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>40</td>
</tr>
<tr>
<td>Norway pout</td>
<td>-0.13 (0.5019)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
</tr>
<tr>
<td></td>
<td>-0.42 (0.0244)</td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Southern sandeel</td>
<td>-0.26 (0.1632)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
</tr>
<tr>
<td></td>
<td>-0.46 (0.0114)</td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Northern sandeel</td>
<td>-0.24 (0.1949)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
</tr>
<tr>
<td></td>
<td>-0.55 (0.0019)</td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Sprat</td>
<td>-0.33 (0.0384)</td>
</tr>
<tr>
<td></td>
<td>Trend in 3-year running mean</td>
</tr>
<tr>
<td></td>
<td>-0.52 (0.0007)</td>
</tr>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>40</td>
</tr>
</tbody>
</table>
Table S3. Principal component contributions of different stocks. Contribution indicates the correlation between the species specific metric and the principal component. Large contributions are present where the principal component closely follows the species specific metric, small contributions indicate that the principal component does not reflect variation in the species specific metric.

<table>
<thead>
<tr>
<th>Productivity metric</th>
<th>Stock</th>
<th>PC 1 contribution</th>
<th>PC 2 contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruitment success</td>
<td>Herring</td>
<td>0.55</td>
<td>-0.11</td>
</tr>
<tr>
<td></td>
<td>Norway pout</td>
<td>0.32</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Sandeel (southern)</td>
<td>0.53</td>
<td>-0.45</td>
</tr>
<tr>
<td></td>
<td>Sandeel (northern)</td>
<td>0.54</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Sprat</td>
<td>0.12</td>
<td>0.69</td>
</tr>
</tbody>
</table>
Table S4. Correlations between stock. Pearson’s product moment coefficients and p-values (in parenthesis) are provided for all combinations of stocks and for recruitment success and length anomalies, respectively. Significant correlations are in bold font.

<table>
<thead>
<tr>
<th>Productivity metric</th>
<th>Stock</th>
<th>Norway pout</th>
<th>Sandeel (southern)</th>
<th>Sandeel (northern)</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recruitment success (3 year averages)</strong></td>
<td>Herring</td>
<td>-0.01 (0.973)</td>
<td>0.00 (0.993)</td>
<td>0.30 (0.106)</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Norway pout</td>
<td></td>
<td></td>
<td>0.39 (0.034)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sandeel (southern)</td>
<td></td>
<td></td>
<td>0.49 (0.006)</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Sandeel (northern)</td>
<td></td>
<td></td>
<td>0.54 (0.002)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Length anomaly before 1993</strong></td>
<td>Herring</td>
<td>0.35 (0.136)</td>
<td>-0.77 (&lt;0.001)</td>
<td>-0.10 (0.701)</td>
<td>-0.67</td>
</tr>
<tr>
<td></td>
<td>Norway pout</td>
<td></td>
<td>-0.56 (0.020)</td>
<td>0.02 (0.932)</td>
<td>-0.22</td>
</tr>
<tr>
<td></td>
<td>Sandeel (southern)</td>
<td></td>
<td></td>
<td>0.27 (0.318)</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Sandeel (northern)</td>
<td></td>
<td></td>
<td></td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Length anomaly after 1993</strong></td>
<td>Herring</td>
<td>-0.25 (0.270)</td>
<td>0.33 (0.151)</td>
<td>0.04 (0.855)</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>Norway pout</td>
<td></td>
<td>-0.05 (0.830)</td>
<td>-0.21 (0.378)</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Sandeel (southern)</td>
<td></td>
<td></td>
<td>0.09 (0.715)</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>Sandeel (northern)</td>
<td></td>
<td></td>
<td></td>
<td>0.52</td>
</tr>
</tbody>
</table>
Figure S1. Scatter plots of spawning stock biomass ($S$ [tons]) versus recruitment ($R$ [millions]) for each stock, respectively; Herring (a), Norway Pout (b), Northern sandeel (c), Southern sandeel (d), and sprat (e). Before 1993 (red) and after 1993 (blue). $B_{lim}$ is represented by a vertical green dashed line. Data was taken from the respective ICES assessment working group reports ICES (2015a,b)
Figure S2. The probability $P$ of falling below $B_{lim}$ calculated for a range of $F$ values and for each stock and period, respectively; Herring (a), Norway Pout (b), Northern sandeel (c), Southern sandeel (d), and sprat (e). Before 1993 (red curve) and after 1993 (blue curve). Green dashed line defines $P(S < B_{lim}) = 0.05$ and $F_{msy}$ is where the green dashed line intercepts with the curves.
Figure S3. Scatter plots of spawning stock biomass ($S$) and recruitment ($R$) for each stock. Red points represent fitted values using a Ricker stock-recruitment model (Ricker 1954).
Figure S4. Time series of residuals from figure S1 for each stock (observed $R$ – predicted $R$).