Global Change Biology

Warming temperatures and smaller body sizes: synchronous changes in growth of North Sea fishes

Warming temperatures and smaller body sizes: synchronous changes in growth of North Sea fishes

Running head: Warming temperatures and smaller fish sizes

Alan R. Baudron[1*], Coby L. Needle[2], Adriaan D. Rijnsdorp[3] and C. Tara Marshall[1]

[1] Institute of Biological and Environmental Sciences, University of Aberdeen, Tillydrone Avenue, Aberdeen, AB24 2TZ, Scotland, UK.

[2] Marine Scotland - Science, Marine Laboratory, PO Box 101, 375 Victoria Road, Aberdeen, AB11 9DB, Scotland, UK.

[3] IMARES, Institute of Marine Resources, and Ecosystem Studies and Aquaculture and Fisheries Group, Haringkade 1, 1976 CP IJmuiden, The Netherlands.

[*]Author to whom correspondence should be addressed

E-mail: alan.baudron@abdn.ac.uk, Telephone: +44 (0)1224-272648, Fax: +44 (0)1224-272396

Key words: climate change, temperature size rule, ectotherms, fish growth, von Bertalanffy, Dynamic Factor Analysis, fisheries.

OPINION

Abstract

Decreasing body size has been proposed as a universal response to increasing temperatures. The physiology behind the response is well established for ectotherms inhabiting aquatic environments: higher temperatures decrease the aerobic capacity of individuals giving smaller body sizes a fitness advantage through reduced risk of oxygen deprivation. However, empirical evidence of this response at the scale of communities and ecosystems is lacking for marine fish species. Here we show that over a 40-year period six of the eight commercial fish species in the North Sea examined underwent a synchronous reduction in asymptotic body size that coincided with a 1-2ºC increase in water temperature. Smaller body sizes decreased the yield-per-recruit of these stocks by an average of 23%. Although it is not possible to ascribe these phenotypic changes unequivocally to temperature, four aspects support this interpretation: (i) the synchronous trend was detected across species varying in their life history and life style, (ii) the decrease coincided with the period of increasing temperature, (iii) the direction of the phenotypic change is consistent with physiological knowledge and (iv) no synchrony was detected in other species-specific factors potentially impacting growth. Our findings support a recent model-derived prediction that fish size will shrink in response to climate-induced changes in temperature and oxygen. The smaller body sizes being projected for the future are already detectable in the North Sea.

Introduction

Aquatic environments pose inherent challenges for ectothermic organisms respiring underwater (Graham, 2006). Warming temperatures compound these challenges by increasing anabolic oxygen demand while decreasing oxygen solubility. Any imbalance between oxygen demand and oxygen supply will constrain aerobic scope thereby impairing individual performance (Pörtner & Knust, 2007). In warming environments, smaller-sized individuals are better able to balance demand and uptake because of their larger surface area to volume ratio (Pauly, 2010). These physiological constraints lead to the expectation that individuals experiencing higher temperatures will have smaller body sizes, an outcome known as the temperature-size rule (TSR) (Atkinson, 1994). The physiological basis underpinning the TSR (Pörtner & Knust, 2007; Forster *et al.*, 2011) combined with cross-taxa support (Gardner *et al.*, 2011; Forster *et al.*, 2012; Edeline *et al.*, 2013) has led to smaller body size being proposed as a universal outcome of warming temperatures (Daufresne *et al.*, 2009). In marine ecosystems which include a high proportion of ectothermic species, the implications of the TSR are profound. A recent simulation integrating this ecophysiological understanding with temperature projections predicted that by 2050 the assemblage-averaged maximum body weight of fish species would shrink by 14-24% globally due to the combined impacts of smaller-sized species replacing larger-sized species and the TSR (Cheung *et al.*, 2013). This conclusion, which garnered global press coverage upon publication, has been challenged on the grounds that the scale and the speed of the change are not credible (Brander *et al.*, *in press*). Criticism of the projection model was refuted by the authors (Cheung *et al.*, *in press*). This debate highlights the need for an ecosystem-level test of whether body sizes of fishes have synchronously decreased in regional seas that have undergone warming. While the importance of TSR has been shown in laboratory conditions (Forster *et al.*, 2012),

empirical evidence is needed (Ohlberger, 2013). Statistical analyses of long term data series are required to provide the most direct evidence for changes in body size caused by climate warming (Daufresne *et al.*, 2009).

Over the past 30 years water temperatures in the North Sea have increased by 0.2-0.6 °C per decade with the rates of warming being rapid relative to other regional seas (Belkin, 2009). During this period, declining body sizes have been observed in haddock (*Melanogrammus aeglefinus*) (Baudron *et al.*, 2011), herring (*Clupea harengus*) (Brunel & Dickey-Collas, 2010) and plaice (*Pleuronectes platessa*) (van Walraven *et al.*, 2010). The fact that three species differing in their life histories, trophodynamics and vertical distribution in the water column (Supplementary Table S1) exhibited smaller body sizes concomitant with a warming environment is consistent with the claim that the TSR is a universal response. A complication in establishing direct causality between warming temperatures and decreasing body sizes in commercial stocks is that size-selective fishing mortality may select for genotypes affecting growth (Enberg *et al.*, 2012) and reduction in body size could therefore be the result of non-random genetic selection. Furthermore, commercial species experience particularly large fluctuations in abundance that could introduce variability in growth rates via density-dependent competition for resources (Taylor & Stefánsson, 1999). Unlike temperature, it is difficult to see how these two factors could impact growth uniformly across species. The scale and speed of an evolutionary response would be unique to each stock given that the pattern and degree of selection varies across stocks and stocks differ in the life history traits (e.g., age at maturity) that determine how quickly a phenotypic trait evolves (Supplementary Table S1). Similarly, the mechanisms responsible for generating density-dependent growth would also likely be species-specific given the variety in habitat and diet of North Sea fish species (Supplementary Table S1). Although evolutionary and density-dependent changes in

4

growth seem unlikely to vary synchronously across species both factors must be examined

for synchronicity. If no synchronous trends across species are detectable in these two factors

then logically they cannot be responsible for generating phenotypic changes that are

synchronous across species. Establishing that declines in body size are, firstly, synchronous

across a range of species, and secondly, concurrent with temperature would strongly imply

the "omnibus" effect of temperature.

Our aim was to test whether the North Sea fish assemblage exhibited synchronous declines in

asymptotic body size that were concurrent with increases in temperature, and consistent with

TSR. We used the von Bertalanffy growth function (VBGF) (Pauly, 2010) to estimate L_∞, the

asymptotic body length, on a cohort-by-cohort basis for eight North Sea fish species for

which age and size data were available at least annually over the past four decades. Statistical

analyses were performed to test for a synchronous trend in L_∞ across species and compare

this trend with the temperature trend. Annually resolved indices of fishing mortality and

density were also examined for synchronicity across species to determine the potential for

these factors to generate a synchronous trend in L_∞. The implications for yield were examined

for species exhibiting a synchronous component in variability of L_∞.

Methods

Sea bottom temperatures (SBT) from 1970 to 2008 were obtained from the NORWECOM model (Skogen & Søiland, 1998) which gives monthly mean values averaged over 0.25 degree latitude by 0.5 degree longitude rectangles. SBT values were averaged per roundfish area (Supplementary Fig. S1a) and per year in order to match the spatial and temporal resolution of the biological data. SBT time series showed the same trend in northern (areas 1 and 2) and southern (areas 5 and 6) regions with a ca. 3°C gradient (Supplementary Fig. S1b). Age-length keys (ALKs) generated from data collected during annual International Bottom Trawl Surveys (IBTS) of the North Sea are available for demersal and pelagic species from the DATRAS database (http://datras.ices.dk) maintained by the International Council for the Exploration of the Sea (ICES). ALKs for benthic flatfish species are estimated from a combination of commercial samples, survey samples and otolith back-calculations (Rijnsdorp *et al.*, 2010). Our analysis was restricted to commercial species having long time series of otolith-derived age estimates which are essential for modelling growth. Demersal (distributed nearer bottom) species were haddock, cod *(Gadus morhua)*, whiting *(Merlangius merlangus)* and Norway pout *(Trisopterus esmarkii)*. Pelagic (distributed nearer surface) species were herring and sprat *(Sprattus sprattus)*. Benthic (distributed on the bottom) flatfish species were plaice and sole (*Solea solea*). Collectively, these eight species (hereafter referred to as stocks) span a range of habitats, body sizes, and life history traits that are representative of the North Sea fish assemblage.

ALKs give the number of sampled fish in a given 1-cm length class that were assigned through otolith reading to age t in year y. For IBTS surveys conducted in quarter 1 age values remained as integers (t) while ages used for quarter 2, 3 and 4 were $t + 0.25$, $t + 0.5$ and $t +$

0.75, respectively. Because ALKs are generated through length-stratified sampling of the catch, they do not accurately represent the true length distribution-at-age. To correct for this bias, ALKs for each ICES roundfish area were raised by the catch-per-unit-effort per length class for the area except for the two flatfish species where ALKs were raised by length distributions instead. To account for the strong spatial gradient in temperature (Supplementary Fig. S1), data for cod, whiting and herring were split into northern and southern sub-stocks and the raised ALKs were combined for ICES roundfish areas 1 and 2 and for areas 5 and 6 to represent the northern and southern North Sea, respectively. Haddock and Norway pout are found in the north, whereas sprat are found in the south. Plaice and sole are found in the south but were split by sex to account for known differential growth that gives rise to large phenotypic differences between males and females (Rijnsdorp *et al.*, 2010). Splitting gave a total of 13 sub-stocks and accounted for known sources of variation in temperature (by region) and growth (by gender) that would otherwise have confounded the analysis. For each sub-stock the growth of a cohort spawned in year *y* was modelled by fitting the VBGF to the length distribution-at-age represented by the raised ALKs:

$$L_t = L_\infty (1 - e^{-K(t-t_0)})$$

where L_t is the length (cm) at age *t*, *K* is the Brody growth parameter (year[-1]), and t_0 is the hypothetical age (year) at length equal to 0. As cohort-specific values of L_∞ and *K* are negatively correlated (Pauly, 2010) (Supplementary Fig. S2) examining temporal variation in one of the two parameters is sufficient to describe growth. Examining variation in L_∞ by cohort assumes that the growth trajectory of a cohort is established in the early stages of life, an assumption supported experimentally (Forster *et al.*, 2011; Scott & Johnston, 2012). For each sub-stock the VBGF was fit for cohorts 1970 to 2008; cohorts 1970 to 2001 were modelled with data for ages 1 to 10 while cohorts 2002 to 2008 were modelled with nine to three years of data. Outliers in the L_∞ time series (> 2 * average or L_∞ having standard errors

>10% of the value) were omitted from subsequent analysis (see Supplementary Table S2).

For all other cohorts, estimated parameters (L_∞, K and t_0) had p-values <0.05. A Dynamic

Factor Analysis (DFA) (Zuur *et al.*, 2003) was used to objectively estimate common trends in

the L_∞ time series of the sub-stocks considered. DFA is a multivariate extension of structural

time series which can analyse short, non-stationary time series containing missing values.

The aim of DFA is to model as few common trends as possible while giving a reasonable

model fit. Prior to inclusion in the DFA model, L_∞ time series for all sub-stocks were

standardized by subtracting the mean and dividing by the standard error (Zuur *et al.*, 2003).

The 13 time series were each modelled as a combination of common trends (x), factor

loadings (Z) plus some offset (a) as follows:

$$L_{\infty,s}(t) = Z_{1,s}x_{1,t} + ... + Z_{i,s}x_{i,t} + a_s + \varepsilon_{s,t}$$

where $\varepsilon_t \sim$ MVN(0, R) with MVN standing for Multivariate Normal and R being the error

covariance matrix, s is the considered sub-stock and i is the number of common trends. The

magnitude and sign of Z indicate to what extent the common trends are related to the original

times series. DFA models with 1 to 6 common trends and with either a diagonal and equal or

a diagonal and unequal error covariance matrix were tested. The best of the twelve candidate

models was selected using the Akaike's information criterion (AIC). Correlation tests

between the predominant common trend (Trend 1) and SBT were performed using different

time windows for temperature impacts: SBT experienced at age 0 (no lag), at age 1 (lag 1), at

age 2 (lag 2), during the first two years of life (average SBT from age 0 to 1) and during the

first three years of life (average SBT from age 0 to 2). Correlation tests between Trend 1 and

the common trend in density were performed to test for intra-cohort (no lag), inter-cohort (lag

1 and lag 2), and cumulative (sum of densities at lags 0 to 1, and lags 0 to 2) density-

dependent growth. Since, the purpose of these tests was to assess the correlation between

underlying trends rather than short-term, high frequency variations, the P-values presented do

not account for autocorrelation as this would have involved detrending the time series. Correlations were summarized by Pearson product-moment correlation coefficients. As multiple tests were conducted, a sequential Bonferroni correction was applied to adjust the level of significance of the multiple inferences.

To control for synchronicity in fishing mortality and density across species, the DFA and correlation tests were repeated for both factors. Average fishing mortalities were obtained from ICES (http://www.ices.dk/) 2012 assessment reports for the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) and the Herring Assessment Working Group (HAWG). Assessment data were available from 1970 to 2011 for all species apart from whiting, Norway pout and sprat which assessments began respectively in 1990, 1983 and 1991. For stocks distributed across the northern and southern North Sea (whiting and herring), a survey-based assessment (SURBA) model (Beare *et al.*, 2005) was used to obtain local estimates of total mortality in order to capture spatial gradients in fishing pressure. Assuming a constant natural mortality, total mortality times series for these two stocks were used as proxies for fishing mortalities. Abundance at age 1 indices were used as a proxy for density (no sex-specific abundance index were available for plaice and sole sub-stocks). For stocks distributed across the northern and southern North Sea, the survey abundance at age 1 indices were split by area. For other sub-stocks, XSA abundance at age 1 indices given in the 2012 assessment reports were used. For Norway pout North and sole South, recruitment time series from the assessment were used as no abundance at age 1 indices were available. Data were available from 1970 to 2011 for all species apart from whiting, Norway pout and sprat for which indices were available from 1990, 1983 and 1984 respectively.

Comparative yield-per-recruit analyses were performed as a proxy for changes in yield prior to and after changes in individual body size. Yield-per-recruit was approximated by simulating a fishery on a single cohort composed of eleven year classes (from age 0 to age 10), and with an initial recruitment (abundance at age 0) of 10000 individuals. 5-year mean of L_∞ and K prior and after changes in growth were used to compute length-at-age values, assuming t_0=0 for all sub-stocks. Length values were converted into weights-at-age using length-weight relationships obtained from Marine Scotland and IMARES. Fishing mortality-at-age was estimated by the mean over the last three historical years (2009 to 2011) while natural mortality-at-age and proportion mature-at-age were assumed to be constant. For plaice and sole the fishing mortality was assumed to be equal for both sexes. All estimates were obtained from the 2012 assessment reports. The cumulative contribution to yield of the successive year classes of the cohort were then summed and divided by the original number of recruit to obtain approximations of yield-per-recruit.

All statistical analyses were performed using the R software (version 2.15.1; http://www.r-project.org/). DFA was performed using the MARSS package (Holmes *et al.*, 2012).

Results and discussion

A decrease in L_∞ (expressed as difference between average 1973-1977 and average 1993-1997) of 29%, 13%, 29%, 10%, 19%, 16%, 13%, 1% and 12% (average 16%) was observed for haddock North, whiting North, whiting South, herring North, Norway pout North, sprat South, male sole South, female sole South and male plaice South, respectively (Fig. 1c-f and h-l). These nine sub-stocks also exhibited narrow 95% confidence intervals (95%CI) around L_∞ estimates. The four remaining sub-stocks (cod North, cod South, herring South and female plaice South) showed divergent trends in L_∞ (Fig. 1a-b, g and m). Cod North, cod South and herring South time series showed an increase in L_∞ but included high proportions of outliers (Supplementary Table S2). The wide 95%CI for the two cod sub-stocks indicate that the VBGF was a poor fit to the data and the high (>200 cm) values of L_∞ reflect near-linear growth rather than the asymptotic growth assumed by the VBGF. Herring South showed a sudden decline in the late 1970s (Fig. 1f) while L_∞ for female plaice South showed an increase from 1970 to 1990 followed by a recent decrease (Fig. 1m).

The best model identified by DFA to describe temporal variation in L_∞ included two common trends (Supplementary Table S3). Fitted values for all sub-stocks showed that the model succeeded in describing the overall trends in L_∞ (Fig. 1). The DFA model captured the decrease in L_∞ for the nine sub-stocks identified and exhibited narrow 95%CI apart from the most recent cohorts which have fewer sampled age classes (Fig. 1c-f and h-l). The first common trend (Trend 1) showed a steep decline from 1978 to 1993 after which it remained stable (Fig. 2a). The second common trend (Trend 2) increased slightly until the early 1990s after which it showed a sharp decline (Fig. 2c). A DFA model fitted with one common trend only returned a trend almost identical to Trend 1 (Fig. 2a), which indicates that Trend 1 is

11

predominant in describing the common trend in L_∞ (Zuur *et al.*, 2003). The nine sub-stocks showing a decrease in L_∞ from the mid-1970s to the mid-1990s (Fig. 1) were all positively related to Trend 1 (Fig. 2b). Six of these sub-stocks (haddock North, whiting North, whiting South, herring North, Norway pout North and male sole South) had similar factor loading values (Z) on Trend 1, indicating that the stock-specific trends in L_∞ were equally well described by Trend 1 (Fig. 2b). These sub-stocks also showed small Zs on Trend 2 (Fig. 2d). Female sole South and male plaice South, although positively related to Trend 1, showed larger Zs on Trend 2 (Fig. 2d).

Cod North, cod South, and female plaice South were negatively related to Trend 1 but positively related to Trend 2 (Fig. 2b and 2d). Herring South was negatively related to Trend 2 and showed a different temporal trend (Fig. 1g). The high Z of female plaice South on Trend 2 (Fig. 2d) suggested that Trend 2 was mainly driven by this sub-stock. Because the L_∞ time series for the cod sub-stocks and herring South included many outliers (Supplementary Table S2) it is difficult to describe decadal-scale trends for these sub-stocks with confidence. Flatfish sub-stocks showed mixed trends, with males of both species conforming to Trend 1, whereas, the decrease in L_∞ was small in female sole South and absent in female plaice South. Differences between males and females in their growth response to temperature have been detected in fish (Kuparinen *et al.*, 2011) although the physiological basis is unknown.

The common trend in L_∞ represented by Trend 1 describes the synchronous component of variability in the L_∞ time series. Residual variation around Trend 1 for a given stock reflects the influence of stock-specific factors and estimation error, neither of which are beyond the scope of this study. Trend 1 showed concurrent and opposite variation with sea bottom temperature (SBT) and the temperature increase through the 1980s was concurrent with the

decline in L_∞ described by Trend 1 (Fig. 2e). This is consistent with the TSR prediction that higher temperatures result in smaller body sizes. Furthermore, this cross-stock synchronicity was detected when growth was examined by cohort suggesting that temperatures experienced early in the life of the cohort are critical to determining L_∞, a result also consistent with current physiological understanding of growth (Forster *et al.*, 2011; Scott & Johnston, 2012). Trend 1 was significantly, negatively correlated ($p \leq 0.05$) with the average monthly SBT experienced at age 0, age 1, age 2 as well as during the first two years and the first three years of life (Table 1). The nine sub-stocks exhibiting the synchronous decline in L_∞ described by Trend 1 represent six species with different life-history characteristics and asynchronous trends in fishing mortalities (Supplementary Fig. S3 and S5) ruling out the possibility that non-random genetic selection is responsible for inducing cross-stock synchronicity in L_∞. Although a common trend in density was detected by DFA, it was not equally strongly supported by all nine sub-stocks (Supplementary Fig. S4 and S5). No significant correlations were observed between Trend 1 in L_∞ and the common trend in density (Supplementary Table S4). Therefore, the TSR is a plausible and parsimonious explanation for the synchronous reduction in L_∞ detected in the majority of North Sea sub-stocks and species examined here.

The nine sub-stocks exhibiting a detectable degree of synchrony had fewer outliers and less uncertainty in L_∞ estimates than the four sub-stocks that did not conform to Trend 1. They included fast-growing/early-maturing species as well as slow-growing/late maturing ones and their habitats encompassed pelagic, demersal and benthic regions. This suggests a uniformity of the response to warming temperature which is consistent with current physiological (Pörtner & Knust, 2007; Forster *et al.*, 2011; Scott & Johnston, 2012) and ecological (Daufresne *et al.*, 2009; Gardner *et al.*, 2011; Sheridan & Bickford, 2011; Ohlberger, 2013)

understanding. It is not possible to infer direct causality from our analysis due to the comparative shortness of the L_∞ time series used here (<40 years) and the intrinsically "uncontrolled" nature of ecosystems. Support for inferring a causal relationship between temperature and growth would be provided if ecosystems showing strong warming consistently showed evidence of synchrony in growth rates across species while ecosystems with little to no warming were asynchronous. For example, juvenile growth rates (proportional to the VBGF parameter K) were correlated with temperature for six of eight long-lived commercial fish species in the temperate southwestern Pacific (Thresher *et al.*, 2007). The intrinsically negative correlation between K and L_∞ (Pauly, 2010) (Supplementary Fig. S2) suggests that there may have been a corresponding reduction in asymptotic body size in the Pacific species, a hypothesis that is difficult to test due to their longevity. The two studies, undertaken in temperate regions of the northern and southern hemisphere, confer a degree of verisimilitude on the inference that temperature is responsible for imposing a detectable, synchronous signal on temporal variation in individual growth rates of fish that is consistent with TSR.

If the synchronous decline in L_∞ observed in several species were driven by temperature, as the reasoning above suggests, then this study can be considered to support the prediction that the future assemblage-averaged maximum body weight of species will be substantially smaller (Cheung *et al.*, 2013). This model-derived projection of future shrinkages (14-24% smaller by 2050) reflects both temperature impacts on individual growth and biogeographic shifts towards communities having a higher proportion of smaller-sized species. Considering only impacts on individual growth resulted in 10% shrinkage (Cheung *et al.*, *in press*). Our study showed that a relatively small increase in temperature (<2 ºC) over a 40-year period was concurrent with reductions in L_∞ that were variable across species (1% to 29%) but

surprisingly large (average 16%) and comparable in scale to the Cheung *et al.* projection. The speed and scale of Cheung *et al.*'s prediction about body sizes in future oceans appear more reasonable than critics claim (Brander *et al.*, *in press*). In our opinion the projections of Cheung *et al.* have merit. In combination with our empirical observation of synchronous growth trends in North Sea fishes, it is clear that a comprehensive analysis of available data on individual growth is required.

Our results also provide empirical support for a recent simulation study showing that shrinking body sizes impact fisheries yield (Audzijonyte *et al.*, 2013). Comparing two years before and after the decline in L_∞ (1978 and 1993, respectively) shows that, under several assumptions, yield-per-recruit i.e., the catch in weight per recruit entering the fishery (Beverton & Holt, 1957) for the affected North Sea stocks decreased by 3.1% to 48.1% with an average reduction of 23.1% (Table 2). Despite the many assumptions required to estimate yield-per-recruit, it is self-evident that smaller body sizes will decrease *per capita* estimates of productivity. The magnitude of these declines seems both substantial and underappreciated relative to the well-documented impacts of fishing over recent decades (Fernandes & Cook, 2013). Given that seasonal mean surface temperatures in the North Sea are predicted to increase by 2.42-3.27ºC by the end of the century (MCCIP 2010) future synchronous reductions in yield-per-recruit are probable.

Warming temperatures are generally associated with faster growth rates (higher *K*) for temperate stocks (Thresher *et al.*, 2007; Neuheimer & Grønkjaer, 2012). However, it has not been fully appreciated that, by virtue of the negative relationship between L_∞ and *K*, the downside of fast early growth is smaller adult body size. Consequences of smaller adult body size include reduced per-capita reproductive rates (Rijnsdorp *et al.*, 2010), decreased

resilience (Hsieh *et al.*, 2006) and altered ecosystem function and services (Sheridan &

Bickford, 2011; Edeline *et al.*, 2013). If individual growth rates change directionally in

response to warming then management strategies that assume productivity can be restored to

levels observed when temperatures were cooler must be re-considered.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

References

Atkinson D (1994) Temperature and organism size – a biological law for ecototherms? *Advances in Ecological Research*, **25**, 1–58.

Audzijonyte A, Kuparinen A, Gorton R, Fulton EA (2013) Ecological consequences of body size decline in harvested fish species: positive feedback loops in trophic interactions amplify human impact. *Biology letters*, **9**, 20121103.

Baudron AR, Needle CL, Marshall CT (2011) Implications of a warming North Sea for the growth of haddock Melanogrammus aeglefinus. *Journal of Fish Biology*, **78**, 1874–1889.

Beare D, Needle C, Burns F, Reid D (2005) Using survey data independently from commercial data in stock assessment: an example using haddock in ICES Division VIa. *ICES Journal of Marine Science*, **62**, 996–1005.

Belkin IM (2009) Rapid warming of Large Marine Ecosystems. *Progress in Oceanography*, **81**, 207–213.

Beverton RJH, Holt SJ (1957) On the dynamics of exploited fish populations. *Fisheries Investigations, London, Ser. 2*. **19**: 533 pp.

Brander K, Neuheimer A, Andersen KH, Hartvig M (*in press*) Overconfidence in model projections. *ICES Journal of Marine Science*, doi:10.1093/icesjms/fst055.

Brunel T, Dickey-Collas M (2010) Effects of temperature and population density on von Bertalanffy growth parameters in Atlantic herring: a macro-ecological analysis. *Marine Ecology Progress Series*, **405**, 15–28.

Cheung WWL, Sarmiento JL, Dunne J, *et al.* (2013) Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. *Nature Climate Change*, **3**, 254–258.

Cheung WWL, Pauly D, Sarmiento J (*in press*) How to make progress in projecting climate change impacts. *ICES Journal of Marine Science*, doi:10.1093/icesjms/fst133.

Daufresne M, Lengfellner K, Sommer U (2009) Global warming benefits the small in aquatic ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*. **106**, 12788–12793.

Edeline E, Lacroix G, Delire C, Poulet N, Legendre S (2013) Ecological emergence of thermal clines in body size. *Global change biology*, doi: 10.1111/gcb.12299.

Enberg K, Jørgensen C, Dunlop ES, *et al.* (2012) Fishing-induced evolution of growth: concepts, mechanisms and the empirical evidence. *Marine Ecology*, **33**, 1–25.

Fernandes PG, Cook R (2013) Reversal of fish stock decline. *Current Biology*, **23**, 1432–1437.

Forster J, Hirst AG, Atkinson D (2011) How do organisms change size with changing temperature? The importance of reproductive method and ontogenetic timing. *Functional Ecology*, **25**, 1024–1031.

Forster J, Hirst AG, Atkinson D (2012) Warming-induced reductions in body size are greater in aquatic than terrestrial species. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 19310–19314.

Gardner JL, Peters A, Kearney MR, Joseph L, Heinsohn R (2011) Declining body size: a third universal response to warming? *Trends in ecology & evolution*, **26**, 285–91.

Graham JB (2006) The physiology of fishes (eds Evans DH and Clairborne JB) pp. 85–117. Taylor and Francis.

Holmes EE, Ward EJ, Wills K (2012) MARSS□: Multivariate Autoregressive State-space Models for Analyzing Time-series Data. *The R Journal*, **4/1**.

Hsieh C-H, Reiss CS, Hunter JR, Beddington JR, May RM, Sugihara G (2006) Fishing elevates variability in the abundance of exploited species. *Nature*, **443**, 859–862.

Kuparinen A, Cano JM, Loehr J, Herczeg G, Gonda A, Merilä J (2011) Fish age at maturation is influenced by temperature independently of growth. *Oecologia*, **167**, 435–443.

MCCIP (2010). Marine Climate Change Impacts Annual Report Card 2010–2011 (eds Baxter JM, Buckley PJ, and Wallace CJ) Summary Report, MCCIP, Lowestoft, 12 pp.

Neuheimer AB, Grønkjaer P (2012) Climate effects on size-at-age: growth in warming waters compensates for earlier maturity in an exploited marine fish. *Global Change Biology*, **18**, 1812–1822.

Ohlberger J (2013) Climate warming and ectotherm body size - from individual physiology to community ecology. *Functional Ecology*, **27**, 991–1001.

Pauly D (2010) Gasping fish and panting squids: oxygen, temperature and the growth of water-breathing animals. International Ecology Institute.

Pörtner HO, Knust R (2007) Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science (New York, N.Y.)*, **315**, 95–7.

Rijnsdorp AD, van Damme CJG, Witthames PR (2010) Implications of fisheries-induced changes in stock structure and reproductive potential for stock recovery of a sex-dimorphic species, North Sea plaice. *ICES Journal of Marine Science*, **67**, 1931–1938.

Scott GR, Johnston IA (2012) Temperature during embryonic development has persistent effects on thermal acclimation capacity in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 14247–14252.

Sheridan JA, Bickford D (2011) Shrinking body size as an ecological response to climate change. *Nature Climate Change*, **1**, 401–406.

Skogen MD, Søiland HA (1998) User's guide to NORWECOM v2.0. The NORWegian ECOlogical Model system. Tech.report, Fisken og Havet 18/98, Institute of Marine Research, Pb.1870, NO-5024 Bergen, 42pp.

Taylor L, Stefánsson G (1999) Growth and Maturation of Haddock (Melanogrammus aeglefinus) in Icelandic Waters. *Journal of Northwest Atlantic Fishery Science*, **25**, 101–114.

Thresher RE, Koslow JA, Morison AK, Smith DC (2007) Depth-mediated reversal of the effects of climate change on long-term growth rates of exploited marine fish. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 7461–7465.

van Walraven L, Mollet FM, van Damme CJG, Rijnsdorp AD (2010) Fisheries-induced evolution in growth, maturation and reproductive investment of the sexually dimorphic North Sea plaice (Pleuronectes platessa L.). *Journal of Sea Research*, **64**, 85–93.

Zuur AF, Fryer RJ, Jolliffe IT, Dekker R, Beukema JJ (2003) Estimating common trends in multivariate time series using dynamic factor analysis. *Environmetrics*, **14**, 665–685.

Supporting Information legends

Table S1. Data availability, distribution and life history traits of the species considered in the analysis. Mean length-at-age 1 was calculated from the age-length keys obtained from the DATRAS database (http://datras.ices.dk/Home/Default.aspx), except for plaice and sole which calculated length-at-age 1 using the von Bertalanffy equation. A50 (age at 50% maturity) values were averaged across time series and were estimated by fitting maturity ogives to sex maturity age-length keys obtained from the DATRAS database except for plaice where proportion of mature-at-age values were used. For sole, the value 2.5 was chosen as A_{50} is reached between age 2 and age 3. The selectivity (age classes targeted by the fishery) values were obtained from ICES assessment working group reports and correspond to the age ranges used to estimate the average fishing mortality (F). Main preys were estimated from diet data given in Pinnegar *et al.* (2011) and Greenstreet (1996).

Table S2. Summary table of the cohorts considered as outliers for each sub-stock, with their L_∞ values and associated standard errors (S.E.). Unrealistically high values of L_∞ reflect growth trajectories that are more linear than asymptotic.

Table S3. Selection table of candidate models tested in the Dynamic Factor analysis including log-likelihood, Akaike criterion (AIC) and the difference (Δ_{AIC}) between the AIC of the considered model and the best candidate model (minimum AIC observed).

Table S4 Estimated correlations between Trend1 and the trend in density for the sub-stocks related to Trend 1, with their corresponding *P*-values. The lag included in the estimation of the density is indicated (see Methods). Significance was adjusted by a sequential Bonferroni

correction: the ordered *P*-values were compared with the inequality, $P_i \leq \alpha(1 + k - i)^{-1}$, where

α is the confidence level to test for significance (0.05), *K* is the number of correlation tests

carried out and *i* is the rank of the correlation considered. Correlations for which the

inequality is met are significant (*).

Figure S1. a: the ICES standard roundfish areas for the North Sea used for the International

Bottom Trawl Surveys. b: Overall average annual sea bottom temperature (thick continuous

line) between the average of roundfish areas 1 and 2 (lower continuous line) and the average

of roundfish areas 5 and 6 (upper continuous line). The two lower dashed lines correspond to

areas 1 and 2, the two upper dashed lines to areas 5 and 6.

Figure S2. Log-scaled relationships between the *K* and L_∞ parameters for the sub-stocks

considered in the analysis (triangles: cod, straight crosses: haddock, circles: whiting, squares:

herring, diagonal crosses: Norway pout, stars: sprat, F and M: female and male plaice, f and

m (in grey): female and male sole). Filled symbols stand for sub-stock in northern North Sea,

empty symbols for sub-stocks in the southern North Sea. Lines correspond to linear models

fitted to the data points.

Figure S3. Fishing mortality (filled circles) for the sub-stocks related to Trend 1 (fishing

mortality was assumed to be equal for both male and female sole South) plotted along the

fitted values from the best Dynamic Factor Analysis model (line) and their corresponding

95% confidence intervals (a: haddock North, b: Norway pout North, c: Sprat South, d: plaice

South, e: sole South, f: whiting North, g: whiting South, h: herring North).

Figure S4. Abundance at age 1 indices (filled circles) used as a proxy for density for the sub-stocks related to Trend 1 (for both plaice and sole sub-stocks the abundance index stands for the males and females together as no sex-specific abundance index were available) plotted along the fitted values from the best Dynamic Factor Analysis model (line) and their corresponding 95% confidence intervals (a: haddock North, b: Norway pout North, c: Sprat South, d: plaice South, e: sole South, f: whiting North, g: whiting South, h: herring North).

Figure S5. The common trends (black line) identified by the best-fitting Dynamic Factor Analysis to describe temporal variation in fishing mortality model (panels a and c) and density (panel e) for the eight sub-stocks that were positively related to Trend 1 (grey line) and their corresponding factor loadings for each sub-stock (panels b, d and f respectively). For fishing mortality, the best model identified by DFA included more than one trend suggesting that there is no synchrony in the fishing mortality time series for these eight sub-stocks. For density, although the best model identified by DFA included a single trend, haddock North and sprat South did not conform to it while whiting North showed the highest factor loadings of all sub-stocks, suggesting that the trend was mainly driven by this sub-stock only.

Tables

Table 1 Estimated correlations between sea bottom temperature SBT and Trend 1 and their corresponding *P*-values. The time period included in the estimation of mean temperature is indicated (see Methods). Significance was adjusted by a sequential Bonferroni correction: the ordered *P*-values were compared with the inequality, $P_i \le \alpha(1 + k - i)^{-1}$, where α is the confidence level to test for significance (0.05), *K* is the number of correlation tests carried out and *i* is the rank of the correlation considered. Correlations for which the inequality is met are significant (*).

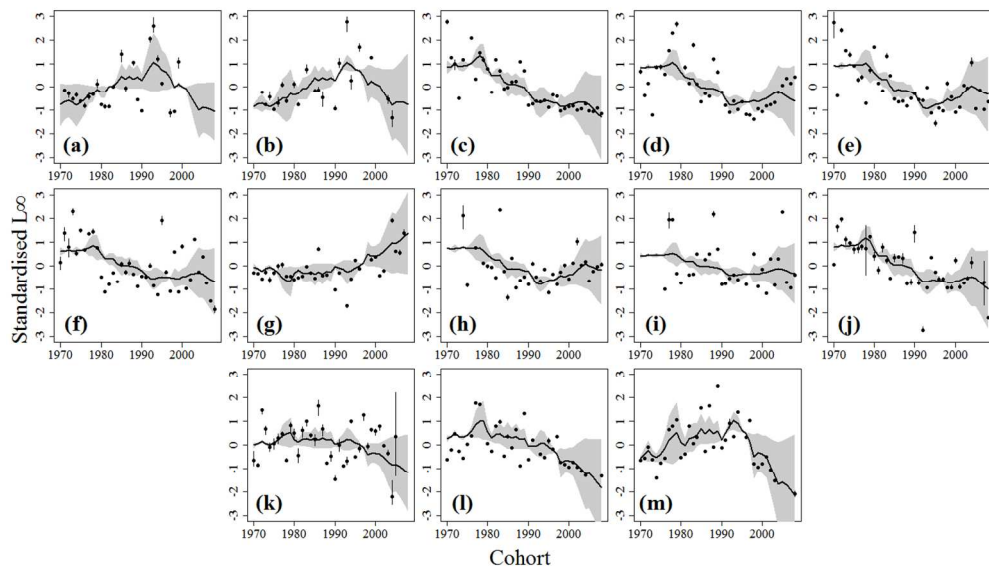| Time period	Correlation	*p*-value	$\alpha(1 + k - i)^{-1}$
0 to 2 years	-0.54	0.00064	0.010*
0 to 1 years	-0.49	0.00182	0.013*
2 years	-0.49	0.00200	0.017*
1 year	-0.45	0.00510	0.025*
0 years	-0.43	0.00640	0.050*

Table 2 Yield-per-recruit (YPR) values (kg) prior (1978) and after (1993) the observed decline in L_∞, with corresponding individual yield loss in value (kg) and percentage. 1977 and 1997 were years in which the standardized common Trend 1 reached its maximum and minimum values prior and after the decline in L_∞. For sole and plaice the low natural mortality estimates (0.1 at all ages) resulted in high YPR values for these two species compared to other species.

Sub-stock	YPR 1978	YPR 1997	Individual yield loss	% loss
Haddock North	0.00473	0.00290	0.00183	38.7
Whiting North	0.00089	0.00086	0.00003	3.1
Whiting South	0.00116	0.00060	0.00056	48.1
Herring North	0.00514	0.00450	0.00063	12.3
N. Pout North	0.00171	0.00133	0.00038	22.2
Sprat South	0.00075	0.00072	0.00003	4.0
Sole male South	0.10458	0.08600	0.01858	17.8
Sole female South	0.14949	0.12571	0.02377	15.9
Plaice male South	0.12375	0.06664	0.05711	46.2
Average				23.1

Figure legends

Figure 1 Standardized L_∞ time series for the thirteen sub-stocks considered in the analysis (filled circles) with their 95% confidence intervals (vertical segments), along with the fitted values from the selected Dynamic Factor Analysis model (line) and their corresponding confidence intervals (shaded areas). a) cod North; b) cod South; c) haddock North; d) whiting North; e) whiting South; f) herring North; g) herring South; h) Norway pout North; i) sprat South; j) male sole South; k) female sole South; l) male plaice South; m) female plaice South.

Figure 2 Common trends given by the best candidate model (panels a and c) to describe L_∞ time variations over time and the corresponding factor loadings for the thirteen sub-stocks (panels b and d). In panel a the grey line corresponds to the common trend given by a model fitted with one common trend. In panel e the Trend 1 is plotted along with the sea bottom temperature (grey line) averaged across the roundfish areas 1, 2, 5 and 6.
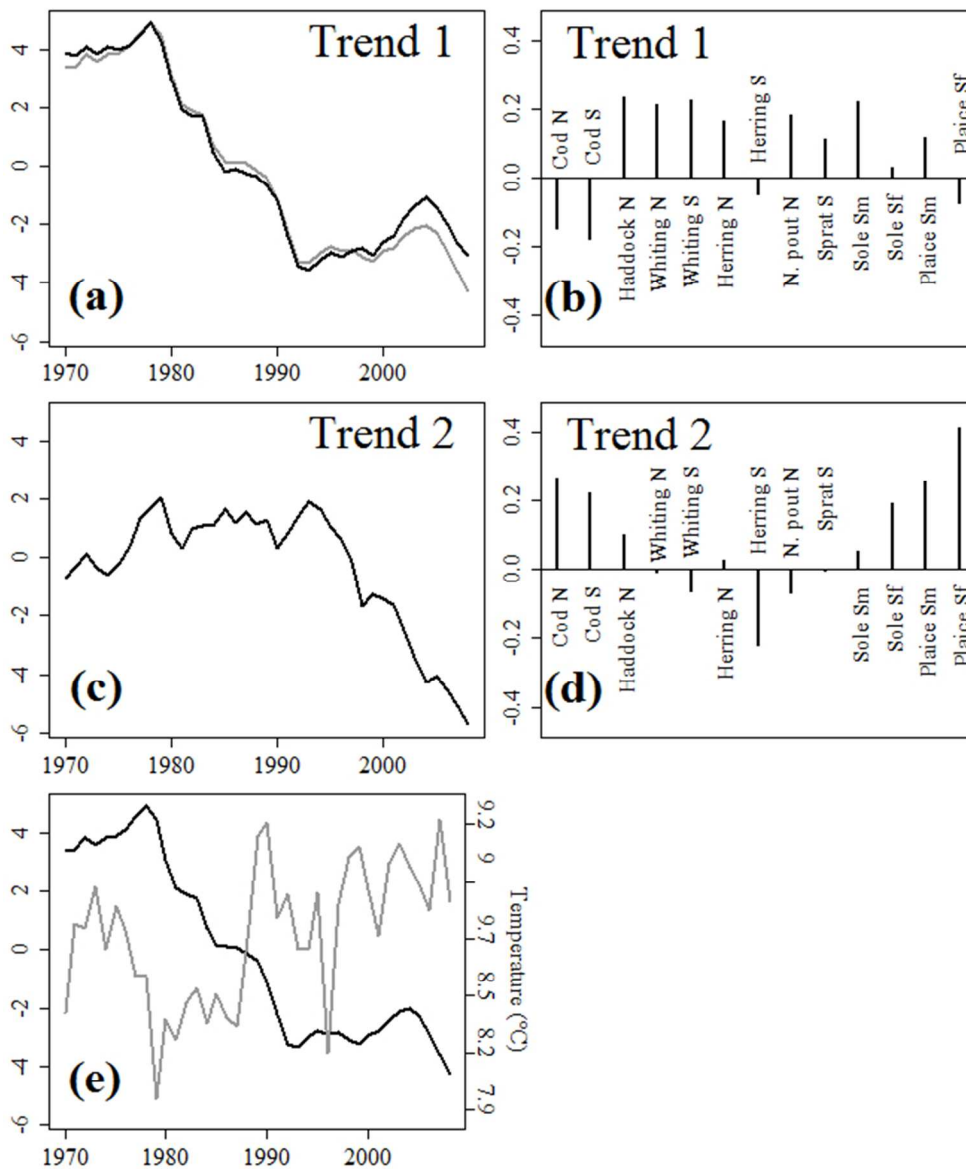
508x288mm (72 x 72 DPI)

237x282mm (72 x 72 DPI)

Supporting Information

Table S1 Data availability, distribution and life history traits of the species considered in the analysis. Mean length-at-age 1 was calculated from the age-length keys obtained from the DATRAS database (http://datras.ices.dk/Home/Default.aspx), except for plaice and sole which calculated length-at-age 1 using the von Bertalanffy equation. A_{50} (age at 50% maturity) values were averaged across time series and were estimated by fitting maturity ogives to sex maturity age-length keys obtained from the DATRAS database except for plaice where proportion of mature-at-age values were used. For sole, the value 2.5 was chosen as A_{50} is reached between age 2 and age 3. The selectivity (age classes targeted by the fishery) values were obtained from ICES assessment working group reports and correspond to the age ranges used to estimate the average fishing mortality (F). Main preys were estimated from diet data given in Pinnegar *et al.* (2011) and Greenstreet (1996).

		Cod	Haddock	Whiting	Herring	Norway pout	Sprat	Plaice	Sole
Time period	Time series	1971 - 2011	1970 - 2011	1970 - 2011	1970 - 2011	1972 - 2011	1972 - 2011	1970 - 2011	1970 - 2011
Region	North	*	*	*	*	*			
	South	*		*	*		*	*	*
Lifestyle	Demersal	*	*	*		*			
	Pelagic				*		*		
	Benthic							*	*
Body size	Length-at-age 1 (mm)	229	194	181	156	125	98	124	103
Maturity	A_{50}	3.2	2	1.3	2.6	1.1	1.0	2.87	2.5
Selectivity	F	2 - 4	2 - 4	2 - 5	2 - 6	1 - 2	1 - 2	2 - 6	2 - 6
Diet	Main prey	Norway pout Sandeel	Sandeel Benthos	Copepods Sandeel	Copepods	Krill Copepods	Copepods Fish eggs	Benthos	Benthos

1

Table S2. Summary table of the cohorts considered as outliers for each sub-stock, with their L_∞ values and associated standard errors (S.E.). Unrealistically high values of L_∞ reflect growth trajectories that are more linear than asymptotic.

Sub-stock	Cohort	L_∞	S.E.
Cod North	1984	144.15	22.21
	1987	201.43	32.54
	1991	358.03	36.63
	1996	499.24	75.95
	2000	214.24	30.58
	2001	190.90	38.39
	2002	176.35	33.65
	2003	161.43	23.16
	2004	409.59	325.06
	2005	192.89	23.77
	2006	1508.37	0.54
	2007	775.07	978.26
	2008	753.37	0.40
Cod South	1971	142.77	16.39
	1979	360.00	55.28
	1982	168.17	17.39
	1984	165.66	21.66
	1988	251.69	29.54
	1989	171.30	45.43
	1992	4597.72	2.05
	1995	245.78	28.18
	1997	1579.30	0.56
	1998	152.11	18.73
	2000	262.16	70.70
	2001	1255.88	0.29
	2002	144.82	17.50
	2005	483.12	423.22
	2006	2111.77	0.46
	2007	1413.29	0.35
	2008	79.45	27.70
Haddock North	1975	132.33	13.81
Whiting North	1980	705.58	0.63
Herring South	1989	85.97	11.99
	1992	407.45	0.38
	1997	166.15	13.69
	2003	261.36	0.57
	2008	271.79	0.28

2

Norway pout North	1976	44.55	2.01
Sprat South	1981	161.84	0.70
	1985	18.87	0.62
	1995	63.44	0.76
	1997	-322.35	1.11
Sole male South	2005	43.17	4.45
	2006	81.56	1.37
Sole female South	2006	234.83	1.12
	2007	195.84	0.69
Plaice male South	1992	93.06	11.72
	2005	169.96	3.02
	2006	96.24	1.98
	2007	52.55	13.51
Plaice female South	1995	384.98	69.40
	2004	530.15	2.45
	2005	690.01	1.61
	2006	566.30	0.80
	2007	364.75	0.52

3

Table S3. Selection table of candidate models tested in the Dynamic Factor analysis including log-likelihood, Akaike criterion (AIC) and the difference (Δ_{AIC}) between the AIC of the considered model and the best candidate model (minimum AIC observed).

Error covariance matrix	Number of trends	Log-likelihood	AIC	Δ_{AIC}
diagonal and equal	2	-542.67	1140.75	0.00
diagonal and unequal	3	-516.89	1144.33	3.59
diagonal and equal	3	-531.79	1144.58	3.84
diagonal and unequal	2	-533.06	1149.51	8.76
diagonal and equal	1	-563.05	1155.09	14.35
diagonal and equal	4	-525.32	1156.16	15.41
diagonal and unequal	4	-509.85	1156.34	15.59
diagonal and unequal	1	-554.61	1164.62	23.87
diagonal and unequal	5	-502.42	1166.14	25.40
diagonal and equal	5	-523.51	1175.69	34.94
diagonal and unequal	6	-500.78	1185.79	45.05
diagonal and equal	6	-523.25	1196.69	55.94

Table S4 Estimated correlations between Trend1 and the trend in density for the sub-stocks related to Trend 1, with their corresponding *P*-values. The lag included in the estimation of the density is indicated (see Methods). Significance was adjusted by a sequential Bonferroni correction: the ordered *P*-values were compared with the inequality, $P_i \leq \alpha(1 + k - i)^{-1}$, where α is the confidence level to test for significance (0.05), K is the number of correlation tests carried out and i is the rank of the correlation considered. Correlations for which the inequality is met are significant (*).

Lag	Correlation	*p*-value	$\alpha(1 + k - i)^{-1}$
2 years	-0.31	0.06193	0.010
0 to 2 years	-0.26	0.11790	0.013
1 year	-0.23	0.15790	0.017
0 to 1 year	-0.18	0.27780	0.025
0 year	-0.15	0.36450	0.050

5

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
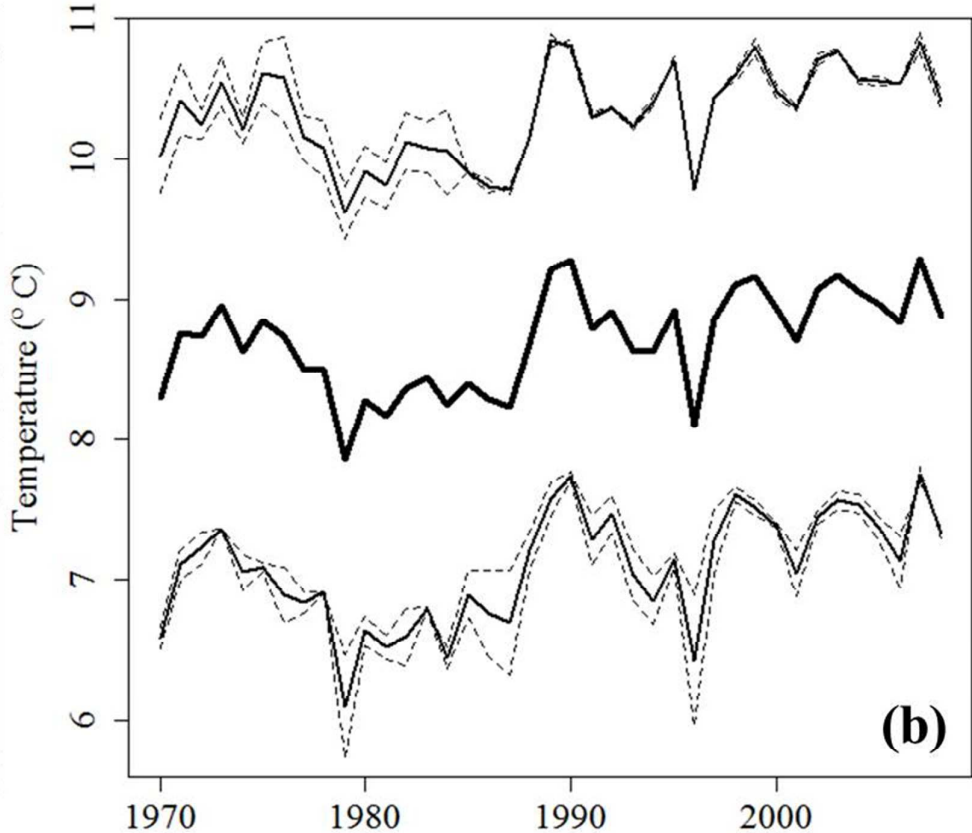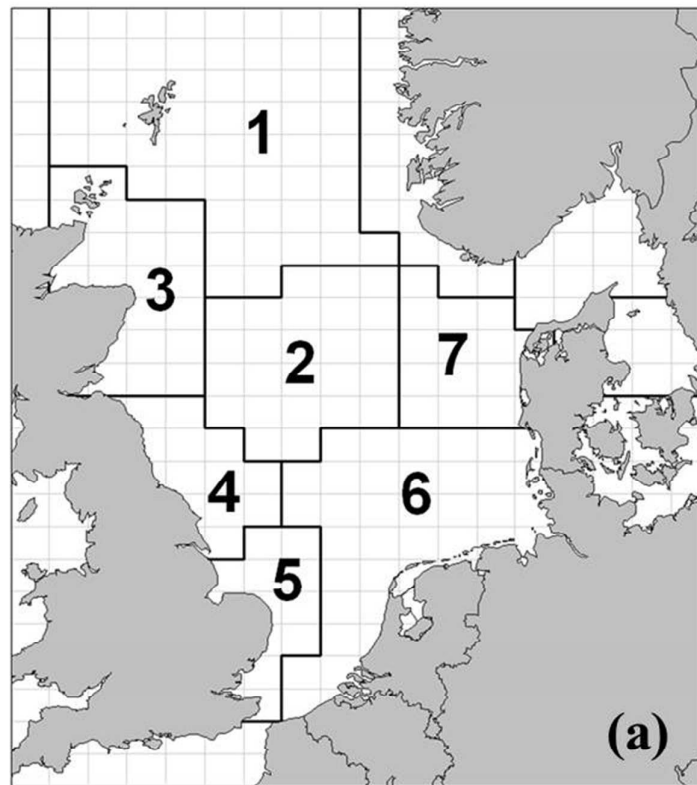18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

Figure S1. a: the ICES standard roundfish areas for the North Sea used for the International Bottom Trawl Surveys. b: Overall average annual sea bottom temperature (thick continuous line) between the average of roundfish areas 1 and 2 (lower continuous line) and the average of roundfish areas 5 and 6 (upper continuous line). The two lower dashed lines correspond to areas 1 and 2, the two upper dashed lines to areas 5 and 6.
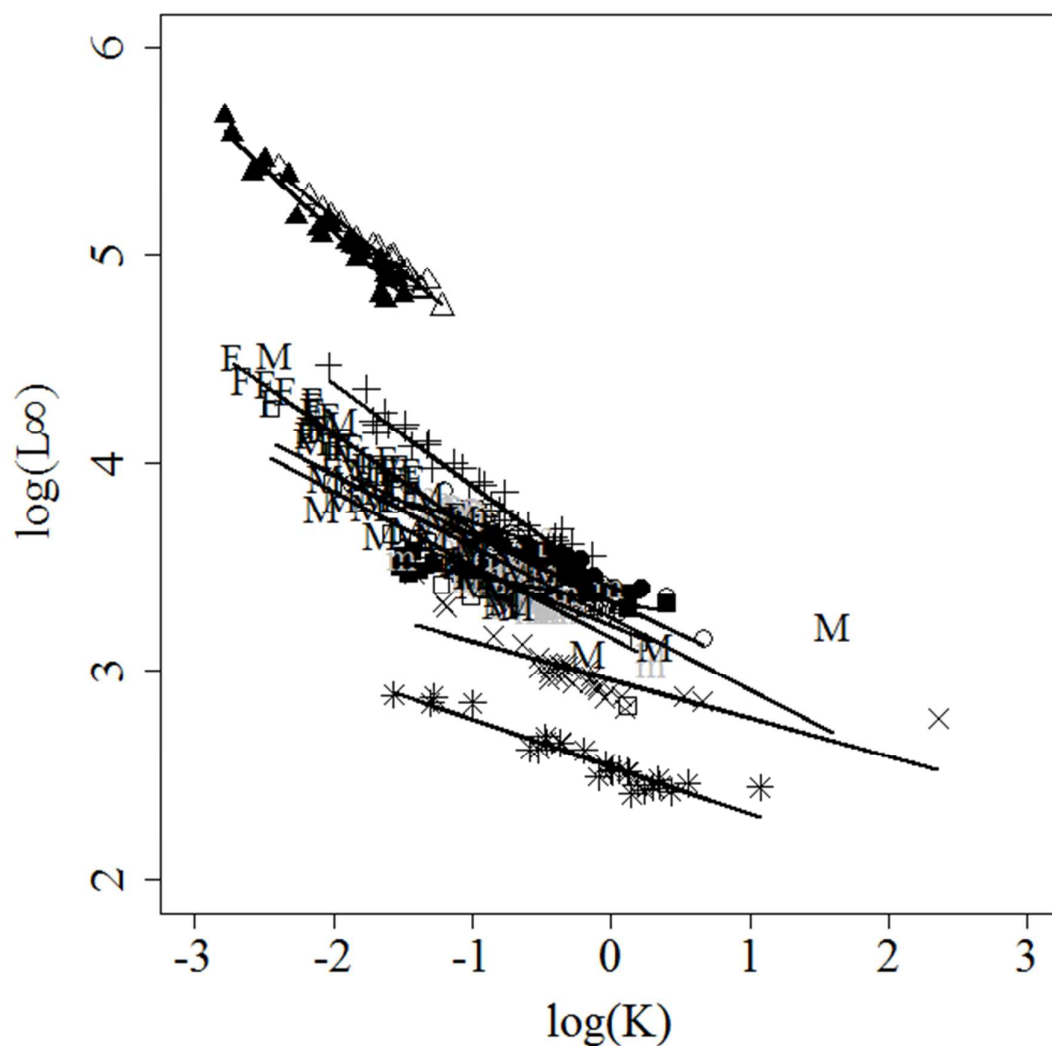
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

6

Figure S2. Log-scaled relationships between the K and L_∞ parameters for the sub-stocks considered in the analysis (triangles: cod, straight crosses: haddock, circles: whiting, squares: herring, diagonal crosses: Norway pout, stars: sprat, F and M: female and male plaice, f and m (in grey): female and male sole). Filled symbols stand for sub-stock in northern North Sea, empty symbols for sub-stocks in the southern North Sea. Lines correspond to linear models fitted to the data points.
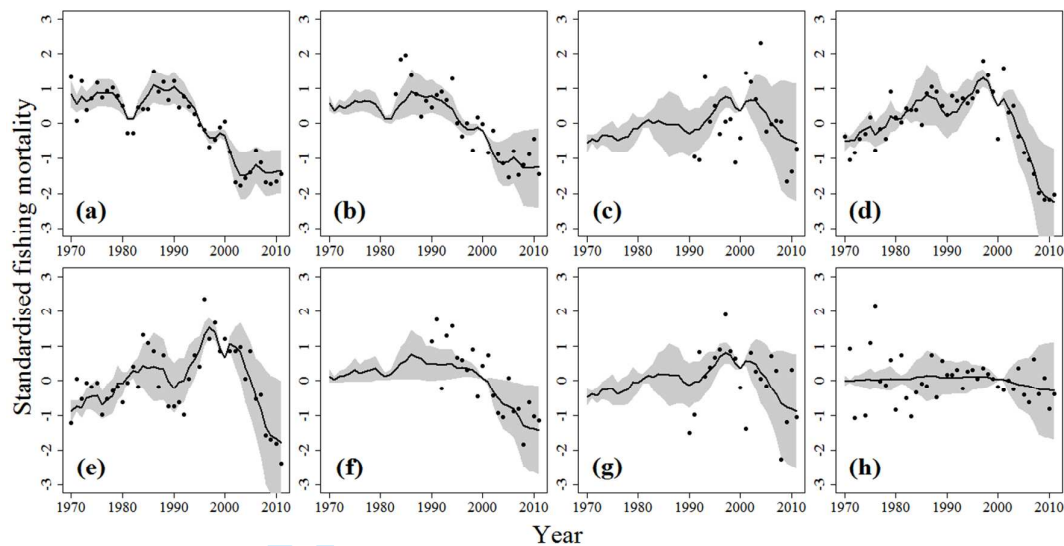
Figure S3. Fishing mortality (filled circles) for the sub-stocks related to Trend 1 (fishing mortality was assumed to be equal for both male and female sole South) plotted along the fitted values from the best Dynamic Factor Analysis model (line) and their corresponding 95% confidence intervals (a: haddock North, b: Norway pout North, c: Sprat South, d: plaice South, e: sole South, f: whiting North, g: whiting South, h: herring North). For stocks distributed across the northern and southern North Sea (whiting and herring), a survey-based assessment (SURBA) model (Beare et al. 2005) was used to obtain local estimates of total mortality in order to capture spatial gradients in fishing pressure. Assuming a constant natural mortality, total mortality times series for these two stocks were used as proxies for fishing mortalities. Fishing mortality as well as the stock weight-at-age and proportion of mature-at-age required for the SURBA model were obtained from the International Council for the Sea (http://www.ices.dk/) 2012 assessment reports for the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) and the Herring Assessment Working Group (HAWG). Assessment data were available from 1970 to 2011 for all species apart from whiting, Norway pout and sprat which assessments began respectively in 1990, 1983 and 1991.
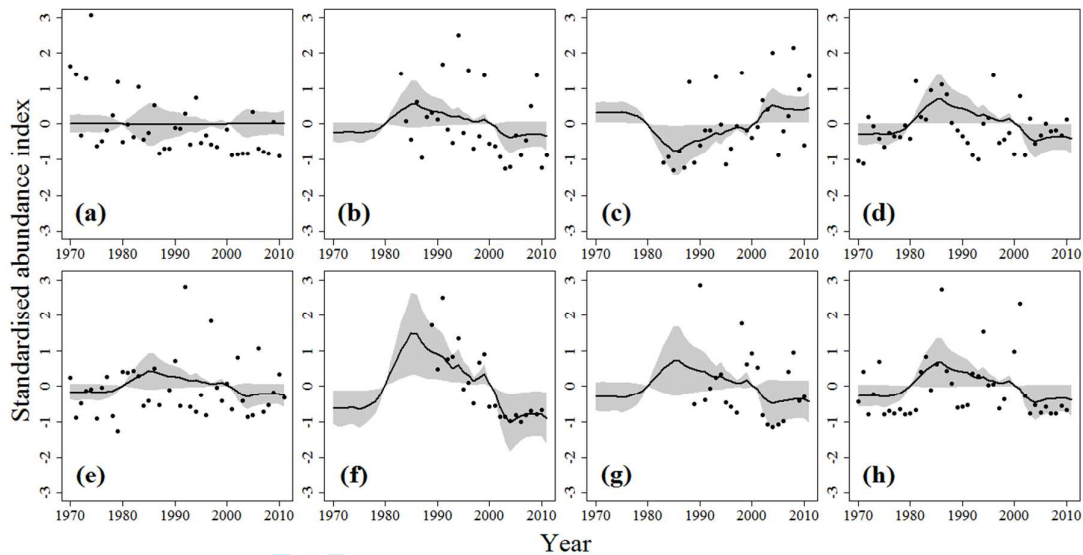
Figure S4. Abundance at age 1 indices (filled circles) used as a proxy for density for the sub-stocks related to Trend 1 (for both plaice and sole sub-stocks the abundance index stands for the males and females together as no sex-specific abundance index were available) plotted along the fitted values from the best Dynamic Factor Analysis model (line) and their corresponding 95% confidence intervals (a: haddock North, b: Norway pout North, c: Sprat South, d: plaice South, e: sole South, f: whiting North, g: whiting South, h: herring North). For stocks distributed across the northern and southern North Sea (whiting and herring), the survey abundance at age 1 indices were split by area. For other sub-stocks, XSA abundance at age 1 indices given in the International Council for the Sea (http://www.ices.dk/) 2012 assessment reports for the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) and the Herring Assessment Working Group (HAWG) were used. For Norway pout North and sole South, recruitment time series from the assessment were used as no abundance at age 1 indices were available. Data were available from 1970 to 2011 for all species apart from whiting, Norway pout and sprat for which indices were available from 1990, 1983 and 1984 respectively.
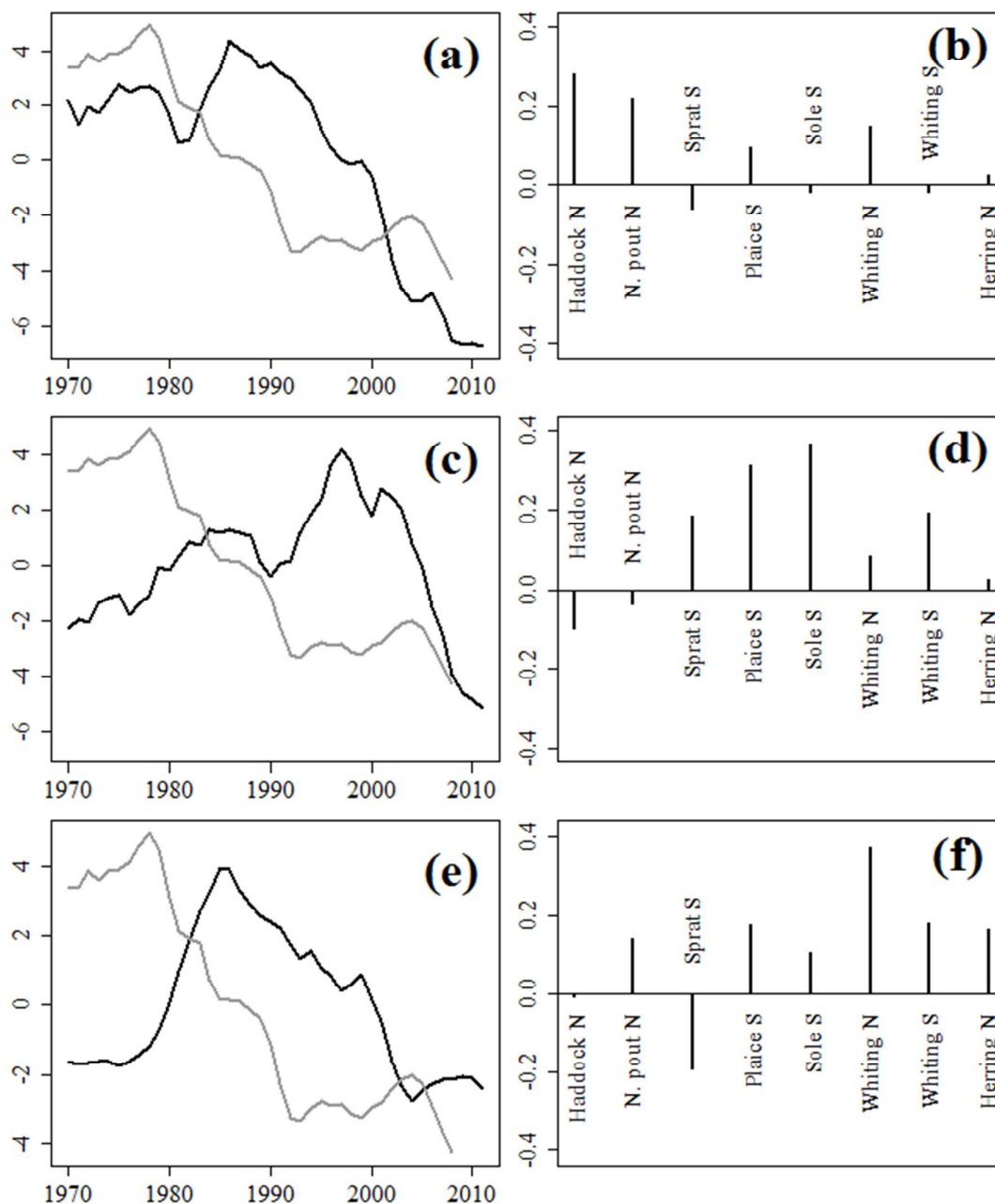
Figure S5. The common trends (black line) identified by the best-fitting Dynamic Factor Analysis to describe temporal variation in fishing mortality model (panels a and c) and density (panel e) for the eight sub-stocks that were positively related to Trend 1 (grey line) and their corresponding factor loadings for each sub-stock (panels b, d and f respectively). For fishing mortality, the best model identified by DFA included more than one trend suggesting that there is no synchrony in the fishing mortality time series for these eight sub-stocks. For density, although the best model identified by DFA included a single trend, haddock North and sprat South did not conform to it while whiting North showed the highest factor loadings of all sub-stocks, suggesting that the trend was mainly driven by this sub-stock only.

Supporting References

Beare D, Needle C, Burns F, Reid D (2005) Using survey data independently from commercial data in stock assessment: an example using haddock in ICES Division VIa. *ICES Journal of Marine Science*, **62**, 996–1005.

Greenstreet SPR (1996) Estimation of the daily consumption of food by fish in the North Sea in each quarter of the year. *Scottish Fisheries Research Report No. 55*.

ICES (2009) Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK). *ICES CM 2009/ACOM: 10.* 1028 pp.

ICES (2009) Report of the Herring Assessment Working Group for the Area South of 62 N (HAWG). *ICES CM 2009/ACOM: 03.* 638 pp.

Pinnegar JK, Platts M. (2011) DAPSTOM - An Integrated Database & Portal of Fish Stomach Records. Version 3.6. Centre for Environment, Fisheries & Aquaculture Science, Lowestoft, UK. Phase 3, Final Report, July 2011, 35pp.