Slave to the rhythm: how large-scale climate cycles trigger herring (*Clupea harengus*) regeneration in the North Sea

Joachim P. Gröger, Gordon H. Kruse, and Norbert Rohlf

Gröger, J. P., Kruse, G. H., and Rohlf, N. 2010. Slave to the rhythm: how large-scale climate cycles trigger herring (*Clupea harengus*) regeneration in the North Sea. – ICES Journal of Marine Science, 67: 454–465.

Understanding the causes of variability in the recruitment of marine fish stocks has been the "holy grail" of fisheries scientists for more than 100 years. Currently, debate is ongoing about the functionality and performance of traditional stock–recruitment functions used during stock assessments. Additionally, the European Commission requires European fishery scientists to apply the ecosystem approach to fisheries in part by integrating environmental knowledge into stock assessments and forecasts. Motivated to understand better the recent years of reproductive failures of commercially valuable North Sea herring, we studied large-scale climate changes in the North Atlantic Ocean and their potential effects on stock regeneration. Applying traffic light plots and time-series (TS) analyses, it was possible not only to explain the most recent reproductive failures, but also to reconstruct the full TS of recruitment from climate cycles, indexed by the North Atlantic Oscillation and the Atlantic Multidecadal Oscillation. A prognostic model was developed to provide predictions of herring stock changes several years in advance, allowing recruitment forecasts to be incorporated easily into risk assessments and management strategy evaluations, to promote a sustainable herring fishery in the North Sea. Insights gained from the analysis permit reinterpretation of the sharp decline in the North Sea herring stocks in the 1970s.

Keywords: AMO, Atlantic herring, climate cycles, cross-correlations, NAO, North Sea, recruitment, time-series analysis.

Received 4 September 2008; accepted 16 October 2009; advance access publication 29 November 2009.

J. P. Gröger and N. Rohlf: Institute of Sea Fisheries, Federal Research Institute for Rural Areas, Forestry and Fisheries, D-22767 Hamburg, Germany. G. H. Kruse: School of Fisheries and Ocean Sciences, University of Alaska Fairbanks, 17101 Point Lena Loop Road, Juneau, AK 99801, USA. Correspondence to J. P. Gröger: tel: +49 40 38905 266; fax: +49 40 38905 263; e-mail: joachim.groeger@t-online.de.

Introduction

European fishery managers strive to maximize landings (or economic value) on a sustainable and precautionary basis. An important constraint is the limited understanding of the typically fluctuating regeneration process of exploited fish populations and how this process interacts with exogenous factors. Rothschild and Shannon (2004) said that "Multi-decadal fluctuations in fish-population abundance . . . are often dramatic in magnitude. . . . Understanding the variability in fish populations related to regime shifts is complicated because the abundance of fish populations is driven by both environmental forcing and fishing. . . . New insights into the cause . . . will be valuable because managers will be able to adjust fishing effort to match the productivity of the ocean environment".

Here, we explore linkages between fish population dynamics and environmental cycles on a global multidecadal scale in a manner that both reproduces emergent properties of the system, such as regime shifts, and that facilitates implementation of the results into stock assessments and fishery management. Motivated to understand the recent years of reproductive failure of commercially valuable North Sea herring (*Clupea harengus*), we studied large-scale climate changes in the North Atlantic Ocean and their potential effects on herring stock regeneration. Therefore, the main objective of our study was to estimate the extent to which North Sea herring recruitment can be reconstructed over the period 1960–2006 using the indices of large-scale climate forcing, so allowing the development of an integrated prognostic recruitment model to improve management procedures. A second objective was to contribute to an ongoing debate on the functionality and performance of conventional stock–recruitment functions, such as Ricker and Beverton and Holt (B/H) curves (Beverton and Holt, 1957), or segmented regressions, typically used in stock assessments to set quotas and to forecast future fish production. Recruits (R) are herring offspring that join the adult herring stocks and are fished after hatching, metamorphosis, and rearing in nursery areas (Figure 1). Not only do stock–recruitment curves often fit the observations poorly, but failure to incorporate environmental effects into the analysis could lead to erroneous forecasts upon which management decisions are based. This is one of the main reasons the European Commission requires European fishery scientists to put the ecosystem approach into practice (CEC, 2001).

North Sea autumn-spawning (NSAS) herring consist of four stock components (Shetland, Buchan, Banks, and Downs), which co-mingle most of the year except at spawning time, which begins in late August around the Shetland and Orkney Islands and continues until January in the southern North Sea (Figure 1). Given extensive annual fluctuations in larval mortality, a variable number of larvae survive their drift to nursery areas in shallow waters near the coast (Heath *et al.*, 1989).

Recently, NSAS herring experienced a period of poor recruitment (Figure 2), in which adults produced large numbers of

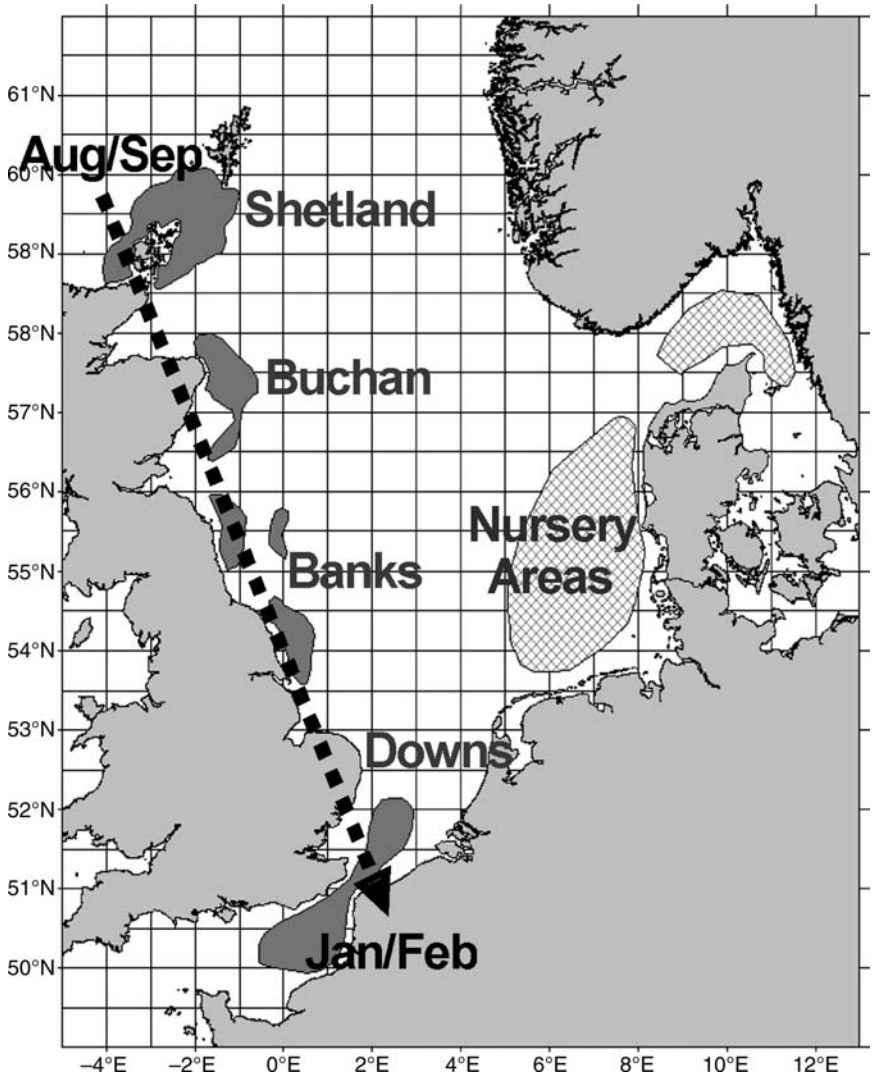


Figure 1. Schematic map of important hatching sites (dark shaded) and nursery areas (cross-hatched) of NSAS herring. Timing of spawning activity ranges from August/September to January/February, and it progresses from north to south.

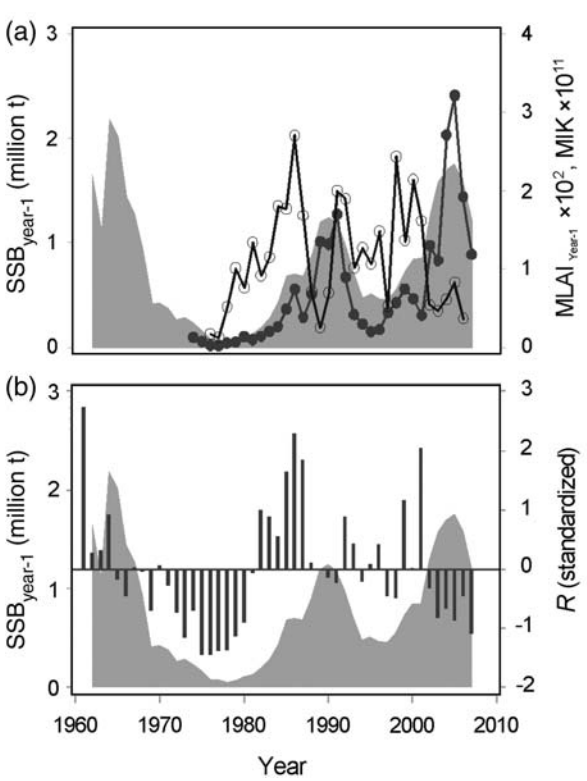
eggs and larvae [MLAI (multiplicative larval abundance index); ICES, 2007] over the years 2000–2006, but very few survived to maturity during the overwintering period, September–February/March [MIK (Methot Isaacs–Kidd) index; ICES, 2007]. We hypothesize that (i) most fluctuations in herring recruitment are driven by variability in large-scale climatic factors during the overwintering period, and (ii) the periodicity of these climatic factors is responsible for the strong cyclical pattern in recruitment.

Methods

Biological and environmental data

To test our hypotheses, we obtained time-series (TS) of two different larval surveys, the MLAI and the MIK (ICES, 2007), from the Herring Assessment Working Group (HAWG) of the International Council for the Exploration of the Sea (ICES). The number of recruits (fish aged 1 year), *R*, and spawning-stock biomass (SSB) are estimated annually by the HAWG using integrated catch analysis (ICES, 2007). Two climate indices were used representing large-scale processes that may influence the recruitment of many fish stocks in the North Atlantic Ocean: the North Atlantic Oscillation (NAO) and the Atlantic Multidecadal Oscillation (AMO). These climate proxies are standardized and may be considered latent factors that average out and "bundle" more regionalized effects that are difficult to compare individually and inter-regionally, which helps to avoid potential "bathtub" problems generated by redundancy, multi-collinearity, or error inflation (Fahrmeir *et al.*, 1996).

Whereas the NAO is an atmospheric sea-level pressure (SLP) anomaly based on the difference in the normalized SLP between Iceland and either the Azores or Portugal (Rogers, 1984; Tunberg and Nelson, 1998; Stenseth *et al.*, 2004), the AMO is an index of long-term sea surface temperature (SST) in the North Atlantic Ocean (Enfield *et al.*, 2001). Both the NAO and the AMO are updated monthly and may be averaged over different periods. Given the focus on the overwintering period, in both cases, we used the winter indices from December to March. For the NAO, we used the mean normalized values of the difference in the normalized SLP among Lisbon, Portugal, and Stykkisholmur/



(c)

Year	R	MIK	MLAI	SSB
1976	20	20	20	20
1977	20	20	20	20
1978	20	20	20	20
1979	20	40	20	20
1980	20	40	20	20
1981	60	60	20	20
1982	80	40	40	20
1983	80	60	40	40
1984	80	80	40	40
1985	100	80	60	40
1986	100	100	40	40
1987	100	80	60	60
1988	60	40	80	80
1989	60	20	80	80
1990	60	40	100	80
1991	40	100	60	60
1992	80	100	40	40
1993	80	40	40	40
1994	40	60	40	40
1995	60	60	40	40
1996	80	80	40	40
1997	40	20	40	40
1998	40	100	60	40
1999	100	60	60	60
2000	60	100	40	60
2001	100	80	80	80
2002	40	20	80	100
2003	20	20	100	100
2004	40	20	100	100
2005	20	40	100	100
2006	40	20	80	80

Figure 2. Herring TS plots: (a) SSB (grey shaded), MLAI (as relative abundance, line and dots), and MIK index (in numbers, line and open circles) over the period 1961–2006; (b) herring recruitment *R* (standardized numbers, needle bars; Gröger *et al.*, 2001) and SSB (grey shaded) for the period 1961–2006; (c) traffic light plot of the normally distributed 20th percentiles of *R*, MLAI, MIK, and SSB from 1976 to 2006. The cells display the quintiles in which the data fall for that year from white (lowest quintile) to black (highest quintile): "20" stands for $0 < x \leq 20\%$, "40" for $20 < x \leq 40\%$, "60" for $40 < x \leq 60\%$, "80" for $60 < x \leq 80\%$, and "100" for $80 < x \leq 100\%$.

Reykjavík, Iceland (http://www.cgd.ucar.edu/cas/jhurrell/indices.html/), and for the AMO, we used data generated from the unsmoothed version of the index from 1856 to present (http://www.esrl.noaa.gov/psd/data/correlation/amon.us.long.data). For detail on the NAO, see Hurrell and van Loon (1997) and Hurrell and Dickson (2004), and for detail on the AMO, see Enfield *et al.* (2001) and http://www.esrl.noaa.gov/psd//data/timeseries/AMO/.

Data analyses

The influence of biotic and climatic factors on recruitment processes was evaluated by considering TS indices of small (MLAI) and large (MIK) larvae (Figure 2a). Annual estimates of *R* were used as the biological response variable, and estimated SSB was taken as one of the potential biotic factors influencing recruitment. For comparison, we fitted three conventional *R*–SSB models (Ricker curve, B/H relationship, segmented regression) to the data. We compared their diagnostics with that of a linearized Ricker curve extended by climate as a potential abiotic factor (winter AMO and winter NAO), as well as that of an ARIMAX model (transfer function) based on the technique of autoregressive-integrated moving averages that also utilizes exogenous climate information. Additionally, other TS techniques, such as spectral analysis (SA) and cross-correlations, were employed to specify the ARIMAX model correctly.

Fitting R–SSB models

All three conventional recruitment models assume positive density-dependence at low stock size. However, after reaching its maximum, the Ricker curve assumes a strong negative density-dependence with increasing SSB, whereas the other two functions assume no density-dependence either after asymptotically reaching a plateau (B/H) or after linearly reaching a breakpoint (segmented regression).

To test the hypothesis whether SSB solely affects *R* or in combination with other exogenous factors, such as climate, we compared the fit diagnostics of the three conventional *R*–SSB models with that of a simple linearized version of the Ricker curve extended by winter AMO and winter NAO (perhaps lagged):

$$\ln\left(\frac{R_t}{\mathrm{SSB}_{t-1}}\right) = \ln(b_1) + b_2\,\mathrm{SSB}_{t-1} + b_3\,\text{winter AMO}_{t-\text{lag1}} + b_4\,\text{winter NAO}_{t-\text{lag2}}, \quad (1)$$

where t is the time, b_1, b_2, b_3, and b_4 the constants, and lag1 and lag2 the time-lags for the winter AMO and the winter NAO, respectively. We implemented a version of Equation (1) that is corrected for autocorrelation using stationary TS and appropriate lags for winter AMO and NAO, making use of information from cross-correlations (see below).

Time-series analysis

TS models are designed to forecast the future states of the system. In contrast to simple (linear or non-linear) regression, TS analysis attempts to identify systematic, self-contained information in the data, such as serial correlation and periodicity, usually without using further external information. Working with TS requires

careful treatment of the data in ways that differ from other types of data. Therefore, we employed three sequential TS analyses:

(i) Cross-correlation analysis—to detect potential delayed external effects, we first cross-correlated the recruitment data with incrementally lagged climate data (i.e. winter NAO and winter AMO);

(ii) SA—we then studied the cyclicity of recruitment and of the (delayed) external factors detected to be influential, and overlaid these with each other to look for corresponding time-trajectories;

(iii) ARIMA modelling—given the information from cross-correlation and SA, we finally integrated the delayed external factors and patterns identified into one ARIMAX recruitment model to be used for prognostic forecasts.

The five steps of model identification are (i) stabilizing and prewhitening the TS, (ii) process identification (influencing factors, lag-structures), (iii) estimation of model parameters, (iv) diagnostics, and (v) prediction (forecasting). For details of the three methods in a marine biological context, see Gröger and Rumohr (2006) and Gröger *et al.* (2007a). For a purely statistical description, see for instance Box and Jenkins (1970), Schlittgen (2001), and Schlittgen and Streitberg (2001). Analyses were performed with the SAS procedures PROC REG, PROC AUTOREG, PROC ARIMA, PROC SPECTRA, and PROC NLIN version 9.2.

Measuring cross-correlations and model performance

In contrast to simple correlation, cross-correlation functions (CCFs) require two treatments of the data before they can be cross-correlated to avoid bias: (i) make both TS stationary, and (ii) prewhiten both TS. These two treatments change the association of the two variables to be compared; the first detrends the data, and the second filters the data by removing autocorrelation. This is necessary to preclude false signals of correspondence that can result simply from the sequential order of the data that do not reflect a true underlying relationship among the two variables. In our case, we carried out the first step by differencing the TS, converting the data from absolute values to sequential changes in time (rates). The second step involved removing autocorrelation, as identified, by applying specific autoregressive models to the TS.

In cases of non-linearity, it is difficult to quantify how good the model is in absolute terms: the best model can only be judged in relation to others. Hence, for each model, $r_{\text{performance}}$ was estimated as the coefficient of correlation between predicted and observed values in an approach similar to that of Kruse and Tyler (1989). As this approach does not account for varying degrees of freedom, we used a bias-corrected version of Akaike's information criterion (AIC) to select the best model (Hurvich and Tsai, 2008):

$$\text{AICC}_k = \ln(\sigma_k^2) + \frac{2 \times k}{(n - k - 2)}, \tag{2}$$

in which σ_k^2 is the residual variance, n the length of the TS, and k the number of parameters estimated. Because the AICC is larger than the AIC and has the tendency to select models with fewer estimated parameters, it is often used in situations where the sample size is small. As σ_k^2 is not independent of the number of estimated parameters (Schlittgen, 2001), it is normally estimated as $\text{MSE}_k = \text{SSE}/(n - k)$, in which MSE is the mean squared error, SSE the sum of squared residuals, n the sample size (length of the complete TS), and k the number of parameters estimated (Schlittgen and Streitberg, 2001). However, the maximum likelihood estimate of MSE assuming $\sigma_k^2 = \sigma^2$ was used. The SAS procedures PROC CORR, PROC REG, PROC NLIN, and PROC ARIMA version 9.2 were applied to calculate most of the measures above.

Statistical significance and standardization of recruitment values

We generally set our significance level to $\alpha = 0.05$. In cases of non-linear models fitted numerically, the tests were based on the asymptotic assumptions requiring large sample sizes ($n > 30$). For distributional tests of normality, we chose a higher α-value of 0.1, increasing the power $(1 - \beta)$ of the test (with β being the type II error and depending on α) to reject the null hypothesis of a normal distribution of residuals (Hartung, 2002).

To better reveal the inherent periodicity in R (as in Figure 2b), we standardized recruitment observations by

$$R_{\text{standard}} = \frac{R - \mu}{\sigma}, \tag{3}$$

where μ is the mean and σ the standard deviation of R, respectively. Standardization was conducted using the SAS procedure PROC STANDARD version 9.2.

Results

Conventional recruitment models

All three models and their parameter estimates were significant (all $p < 0.05$). Except for R in the years 1981–1983, the null hypothesis of positive density-dependence associated with low SSB is supported; negative density-dependence at high SSB values is supported by most observations except R in 1962, 1964, and 1965 (Figure 3a). However, none of the models fitted the data well over the full range (Table 1). All models tended to underestimate R at low values of SSB, to overestimate R at intermediate values of SSB, and to underestimate R again at high SSB (Figure 3b); such systematic bias is characteristic of a mis-specified model. Coefficients of determination ($r^2_{\text{performance}}$) indicate that all three models explain only 22–26% of the variability in observed R (Table 1). Further, model parameter estimates are highly correlated, as the symmetrical off-diagonal values of the approximate correlation matrices show (segmented regression −0.96; Ricker 0.89; B/H 0.75).

Changes in the winter NAO and winter AMO patterns

Traffic light plots of quintiles of monthly values provide a simple visual method to observe systematic changes in seasonal and annual patterns in NAO and AMO over the period 1900–2007 (Figure 4). When values of both indices are normalized by month across all 107 years to give a mean of 0 and a standard deviation of 1, a systematic (wave) pattern with two cooling periods (dark blue, light blue, and green) are evident for AMO, approximately 1903–1925 and 1965–1979 (AMO panel of Figure 4a). Transitions between cool and warm periods tend to be smooth, making it difficult to identify clear breakpoints between warm and cool periods. Nonetheless, it appears that the first cool period is much longer (22 years) than the second (14 years). In contrast, large and small NAO values appear to be more evenly distributed without obvious systematic patterns (NAO panel of Figure 4a). Yet, when the quintiles are based on normalized data for each year across all 12 months, a clear systematic change in

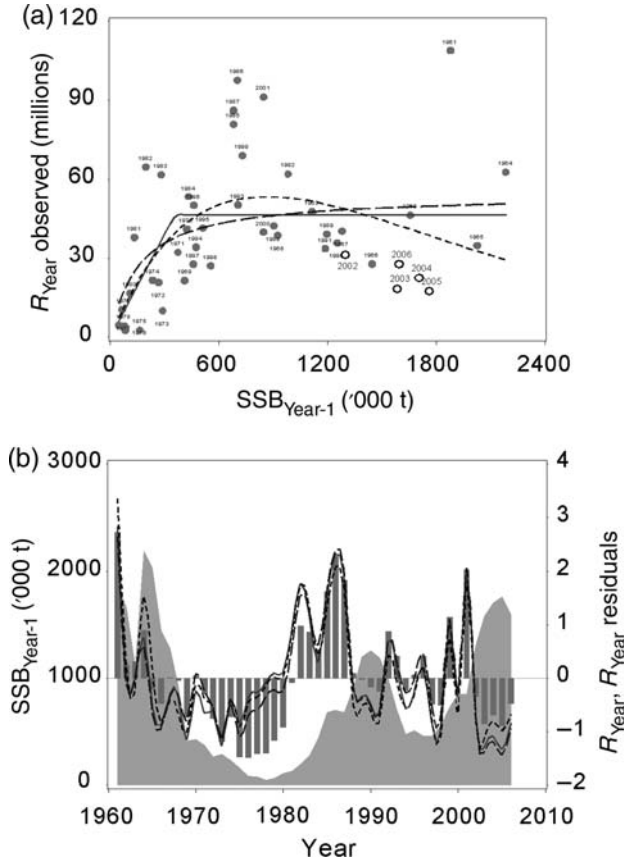


Figure 3. (a) Observed recruitment *R* (millions, dots) plotted against observed SSB (t). The three curves represent the fitted segmented regression (continuous line with increasing and horizontal components), the Ricker curve (short dashed line), and the B/H curve (long dashed line). Values of reduced recruitment in the most recent years (2002–2006) are indicated as open dots (point labels enlarged, lower right). (b) TS plot of observed *R* (standardized numbers, needle bars; Gröger *et al.*, 2001), SSB (grey shaded), and estimated residual *R* from the three stock–recruit models (segmented regression, continuous line; Ricker, short dashed line; B/H, long dashed line) for the period 1960–2006.

Table 1. Parameter estimates and quality-of-fit information from fits of stock–recruitment models.

	Model parameters		Model diagnostics	
Model	b_1	b_2	$r_{performance}$	**AICC**
Segmented regression	130.7	356 091	0.49	33.9384
Ricker	167.4	1.154E−6	0.51	33.9069
Beverton and Holt	0.002 85	54 137 400	0.47	33.9607

All three models and the single parameters associated with them were asymptotically significant at a 5% level.

the patterns is only evident for the NAO during the last 18 or 19 years, with a tendency to low values during the period August–November and high values from December to February (NAO panel of Figure 4b), indicating a shift in the NAO. However, there is also some similarity to the first period of the 20th century (NAO panel of Figure 4b). In contrast, large and small AMO values appear to be more evenly distributed without obvious systematic patterns (AMO panel of Figure 4b).

Cross-correlation and spectral analyses

After detrending and pre-whitening of the TS, we examined the CCFs between *R* and winter AMO, and between *R* and winter NAO, with lags of up to 10 years (≈25% of the total timespan) to detect potential climatic effects during the overwintering period. Given different autocorrelation structures, pre-whitening was handled slightly differently for the two climate variables. For winter NAO, first-, second-, and fourth-order autoregressive components were fitted to both the detrended NAO and *R* dataseries. For winter AMO, only first- and second-order autoregressive components were fitted to both the detrended AMO and *R* dataseries; no moving average components were significant in both cases. Bartlett confidence intervals were used to set confidence limits (Schlittgen, 2001). The results indicate that winter AMO and NAO exert their greatest influence on *R* at lags of 3 and 5 years, respectively (Figure 5a and b).

To avoid misinterpretation, it should be noted that CCFs cannot be compared with Pearson's correlation coefficients or with $r_{performance}$ measures. Such an invalid comparison would be misleading because the magnitude of the latter two may be artificially inflated by serial correlation reducing the effective degrees of freedom (Pyper and Peterman, 1998) and may be biased by non-stationarity. In contrast, CCFs are based on stationary (detrended) and pre-whitened (autocorrelation-free) TS. For this reason, Pearson's correlation coefficients or $r_{performance}$ measures tend to yield larger values than CCFs.

Fluctuations in observed and fitted *R* and fitted winter AMO using SA techniques (wave decomposition) provide some interesting insights into the nature of the apparent relationship between SST and *R*. A strong correspondence between winter AMO and *R* persists over the years 1960–1981; high (low) values of *R* are associated with high (low) values of winter AMO. However, after about 1982, this strong association becomes much weaker. Although the cause of this change in relationship is not certain, we suspect that it results from a switch in the dominance of winter NAO over winter AMO effects. A decomposition of the winter NAO signal into its wave components using SA (similar to the winter AMO decomposition) revealed that the long-wave component has an inflection point around 1982, with lower values before 1982 and higher values thereafter (Figure 5c). Although atmospheric SLP anomalies do not directly affect herring recruitment, oceanographic and ecological processes associated with westerlies typical of the positive phase of the NAO may dominate only when the winter NAO is strongly positive, so trumping winter AMO effects.

ARIMAX modelling

Given our findings from cross-correlation and SA, and the underlying first-order autoregressive process in the recruitment residuals, we combined *R* with the lagged winter NAO and AMO by setting up an ARIMAX model (Figure 6). The 5-year-lagged winter NAO and the 3-year-lagged winter AMO were not correlated ($r = -0.14$, $p = 0.37$), so multi-collinearity between input variables was not an issue. Therefore, we determined the following integrated ARIMAX model based on stationary *R* (i.e. with differentiation order $d = 1$) and with the help of Table 2 as a generalized transfer function, with AR order $p = 0$, MA order $q = 0$, and winter NAO and AMO as exogenous variables (both pre-whitened and made stationary) lagged by 5 and

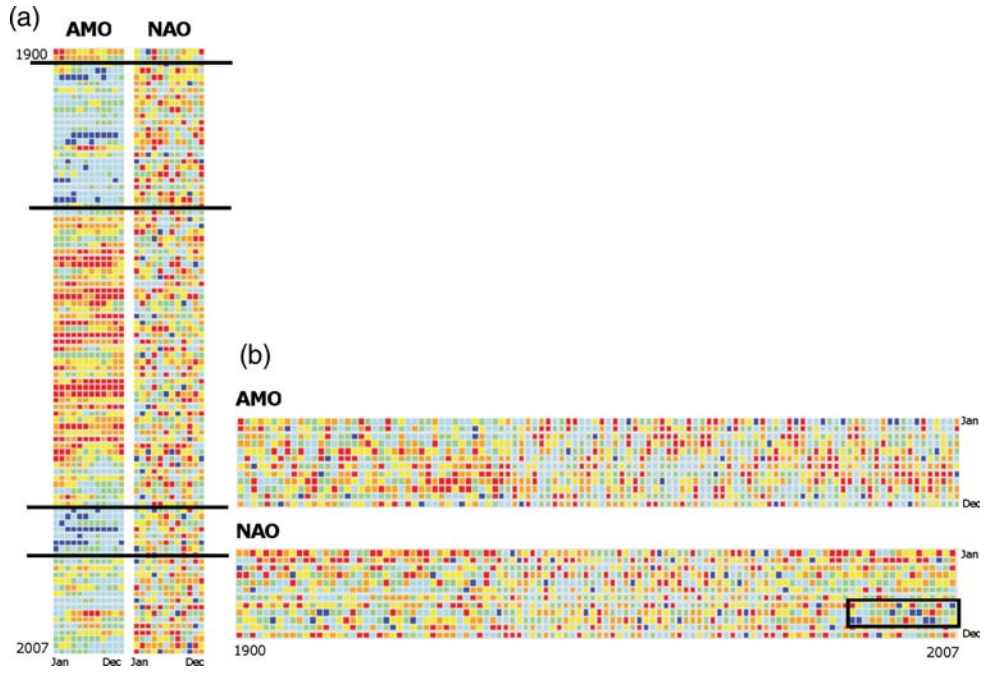


Figure 4. Traffic light plots of the normally distributed 20th percentiles of AMO and NAO from 1900 to 2007 and months January–December. The colour of the cells display the quintiles (x) in which the data fall for that year and month, respectively: dark blue represents $x =$ 0%, light blue means $0 < x \leq 20\%$, light green means $20 < x \leq 40\%$, yellow means $40 < x \leq 60\%$, orange means $60 < x \leq 80\%$, and red means $80 < x \leq 100\%$. (a) Data normalized for each month across all 107 years; horizontal lines separate cooling from warming periods. (b) Data normalized for each year across all 12 months; the black rectangle indicates a higher concentration of seasonally very low NAO quintiles during late summer/autumn when comparing the last two decades with all other previous decades back to 1900.

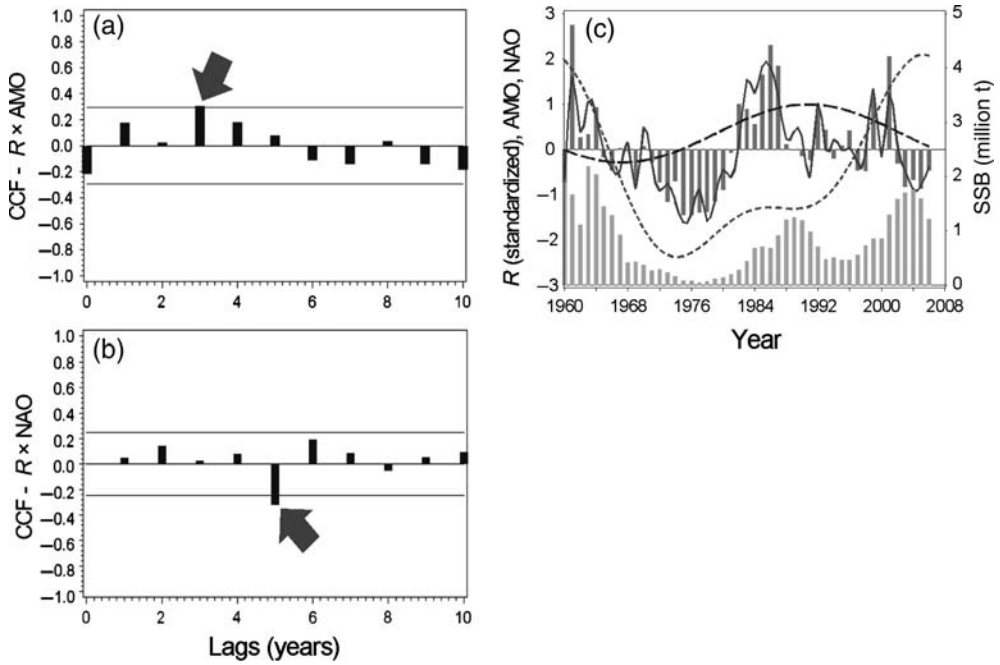


Figure 5. CCFs with 95% confidence bounds (horizontal lines) of R up to a lag of 10 years with (a) winter AMO and (b) winter NAO after detrending and pre-whitening the TS. Arrows mark significant CCFs (exceeding the confidence bounds) at a lag of 3 years in (a) and at a lag of 5 years in (b). (c) Time-trajectories of observed R (standardized numbers, dark grey bars; Gröger *et al.*, 2001), fitted R (cosine–sine SA model, continuous line), observed SSB (million tonnes, light grey bars), fitted winter AMO ($\times 10^{-1}$, cosine–sine SA model, short dashed line), and fitted winter NAO ($\times 10^{-1}$, cosine–sine SA model, long dashed line).

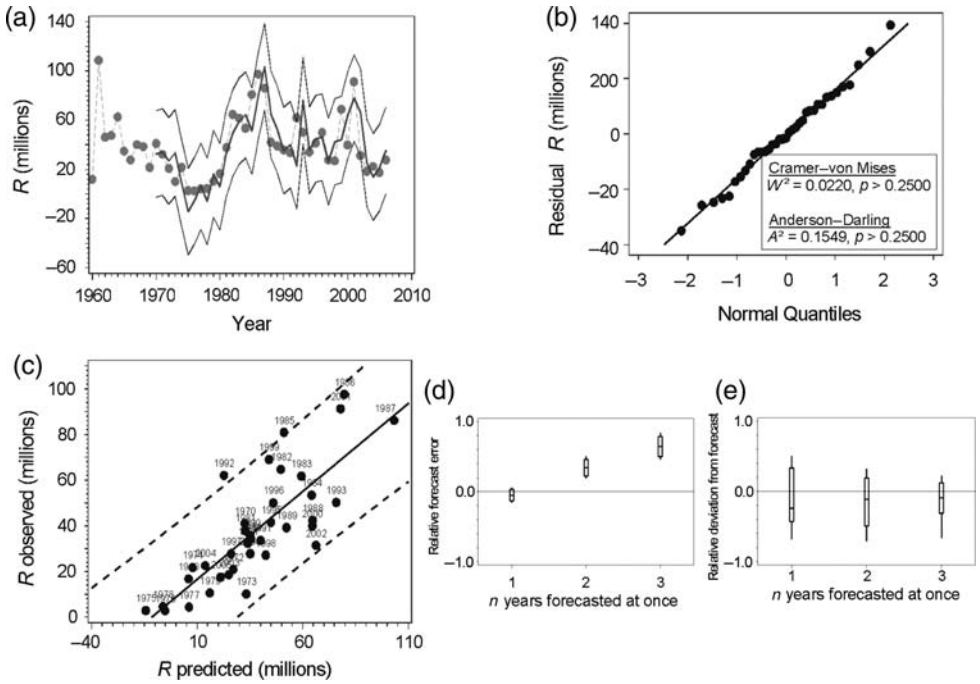


Figure 6. ARIMAX modelling. (a) Observed (dots connected by broken line) and fitted (thick continuous line) *R* values with forecast intervals (thin continuous lines, $\alpha = 0.05$). Owing to the lag structure of the model, the forecasts start in 1970. Values for 2005 and 2006 were omitted from the analysis to allow forecasting them for cross-validation. (b) Q–Q plot of residual values with test results. (c) Observed against predicted *R*. (d) Box plots of relative forecast errors against the number of years forecast simultaneously. (e) Box plots of relative deviations of forecast values from true values against the number of years forecast simultaneously.

Table 2. The ACF and PACF patterns relative to the type of process.

Process	ACF	PACF
AR	Tails off exponentially or in sine-waves	Drops off after lag *p*
MA	Drops off after lag *q*	Tails off exponentially or in sine-waves
ARMA	Tails off exponentially or in sine-waves	Tails off exponentially or in sine-waves

3 years, respectively:

$$(1-B)R_t = \mu + \frac{\omega_1}{(1-\delta_{1,1}B-\delta_{1,2}B^2-\delta_{1,3}B^4)}(1-B)\mathrm{NAO}_{t-5} + \frac{\omega_2}{(1-\delta_{2,1}B-\delta_{2,2}B^2)}(1-B)\mathrm{AMO}_{t-3}+\varepsilon_t, \quad (4)$$

where $B^l = y_{t-l}$ is a shift parameter of order l, and $(1-B)^d = (y_t - y_{t-d})$ is a differentiation parameter of order d. The estimated parameter values are $\mu = -984\,756$, $\omega_1 = -3\,615\,901$, $\omega_2 = 46\,403\,321$, $\delta_{1,1} = -5\,458\,727$, $\delta_{1,2} = -2\,599\,460$, $\delta_{1,3} = -2\,314\,394$, $\delta_{2,1} = -41\,147\,565$, and $\delta_{2,2} = -38\,037\,078$. By including the two delayed climate effects as input series, the first-order serial correlation of the pure *R* has been removed as indicated by a generalized Durbin/Watson test ($DW = 2.68$, $p_{DW} << 0.05$), such that the residuals (and hence the error term ϵ_t) now appear as white noise and normal (Figure 6b), making it unnecessary to fit an extra ARIMA model to them. This also indicates that no other exogenous systematic process is involved. In addition, all other model restrictions (e.g. homoscedasticity, the smallest information criterion AICC) are also fulfilled.

The good fit (Figure 6a) is demonstrated by two features. First, the correlation between predicted and observed *R* values (Figure 6c) is strong ($r_{performance} = 0.80$, $p << 0.05$, AICC = 33.4960, $n_{observations} = 47$, $n_{residuals} = 37$, $k = 10$), with no values exceeding the 95% forecast intervals (Figure 6a). Second, to determine whether the forecasts of our model are stable and plausible, we performed a diagnostic forecast experiment. We did this by successively omitting observations at the end of the TS until a total of 10 years of observations had been reached. After each step, we performed a 1-, 2-, and 3-year prognosis in which we simultaneously forecast either 1, 2, or 3 years ahead; simultaneously forecasting more than 3 years ahead is not possible given the specific lag structure of the ARIMAX model. We estimated the deviations of the forecast values from the known true values as well as the forecast error, both expressed as a percentage of the observed value (for details on the concept of forecast error and how to calculate it, see Pindyck and Rubinfeld, 1991). The forecast error is used to construct the forecast interval and among others contains two major elements: the residual variance and the distance between the centre (year) of the data and the year of the forecast. Therefore, the forecast error increases the further one forecasts into the future as the second component increases which the first component remains constant.

This experiment shows that the relative forecast errors are highly dependent on the number of years being forecast simultaneously; the size of the forecast error is reflected by the size of the two-sided 95% forecast intervals in Figure 6a. Forecast errors for a 1-year prognosis range from −16 to +6% over the 10 iterations with an average of around −5% relative to the forecast

error (being 17 960 801) of the full observation period 1960–2006. In contrast, forecast errors for a 2-year prognosis range from 19 to 50% with an average of ~35%, and those for a 3-year forecast range from 46 to 84% with an average of ~64%. In fact, only the relative forecast error of the 1-year ahead prognosis did not significantly differ from 0 (F-test; $p > 0.05$; see the 95% box plots in Figure 6d), whereas 2- and 3-year averages did ($p << 0.05$). In contrast to the relative forecast error, the actual forecast values did not differ significantly from the true values in all three cases (all three vertical confidence bars include zero in Figure 6e); moreover, the three averages of actual forecast values did not differ significantly from each other ($p >> 0.05$). Interestingly, the 3-year prognosis seemed to perform slightly better than the 1- or 2-year forecasts because the expected value was closer to zero.

Extended Ricker curve

To explore the effects of an interrelated SSB–climate effect, we evaluated a linearized Ricker curve extended by two linear terms related to winter AMO and winter NAO in a manner similar to Equation (1). The Ricker curve was selected because it gave the best quality-of-fit results among the conventional R–SSB curves, and it also took into account a potential density effect (Figure 3). Unsurprisingly, because the response variable was changed from R to $\ln(R/\mathrm{SSB})$, the CCF analyses revealed significant spikes at different lags (lag 6 for both winter NAO and winter AMO) than for the ARIMAX model for both climate indices. Whereas the performance of the extended Ricker curve was better than that of the conventional R–SSB models, it performed somewhat more poorly than the ARIMAX model ($r_{\mathrm{performance}} = 0.74$, $p << 0.05$, AICC = 33.7532, $n_{\mathrm{observations}} = 47$, $n_{\mathrm{residuals}} = 38$, $k = 9$). There was also no obvious signal between SSB and either winter NAO or winter AMO, as revealed by the CCFs between SSB and climate ($p > 0.05$).

Comparisons among alternative models

In summary, the performance measures for the five R–SSB models were as listed below.

(i) B/H model: $r_{\mathrm{perfomance}} = 0.47$, AICC = 33.9607;

(ii) Ricker model: $r_{\mathrm{perfomance}} = 0.51$, AICC = 33.9069;

(iii) Segmented regression model: $r_{\mathrm{perfomance}} = 0.49$, AICC = 33.9384;

(iv) ARIMAX [Equation (4)]: $r_{\mathrm{perfomance}} = 0.80$, AICC = 33.4960;

(v) Extended Ricker model [Equation (1)]: $r_{\mathrm{perfomance}} = 0.74$, AICC = 33.7532.

Only the AICC values of models (i)–(iv) can be compared directly because they are based on the same dependent variable, namely R. Model (v), where the endogenous variable is $\ln(R/\mathrm{SSB})$, had to be back-transformed first to yield absolute recruitment values and to allow a full comparison of the AICC values because they are not scale-invariant. Based on the inspection of both performance values, it is clear that the ARIMAX model is superior not only to the conventional R–SSB models (i)–(iii) but also to the extended Ricker model (v). Improvement resulting from the inclusion of climate information while at the same time removing SSB as a potential factor is apparent. Therefore, recruitment forcing seems to be mainly climate-driven rather than SSB-related. The best model to predict future recruitment is hence model (iv) because it requires neither independent forecasts of another exogenous variable, such as SSB, nor back-transformation as long as the prognosis does not exceed 3 years ahead. Hence, the lag structure of the ARIMAX model (climate effects delayed by 3 and 5 years, respectively) restricts forecasts to 3 years ahead of time.

Discussion

Although it is intuitively appealing to assume that reproductive success depends on parental biomass, for North Sea herring the superior performance of the ARIMAX recruitment model over those models involving SSB indicates that climate was the primary determinant of herring recruitment over the period considered. This is not to imply that there is no stock effect, but rather that climate effects appear to dominate to the degree that a stock effect cannot be detected statistically. Conflicting evidence for reduced R at high SSB stems largely from recent (2002–2006) observations (Figure 3a), which resulted in statistically non-significant, better measures of fit for the Ricker curve (Table 1). Interestingly, during the summer acoustic survey in 2007, we observed cannibalism of large larvae by adult herring. This could be related to our observations of unusually long residence times of post-spawning adult herring in shallow waters (nursery areas, German Wadden Sea) during that survey, which may further explain the decline in the MIK index while the MLAI index increased in recent years (Figure 2a and c). Cannibalism provides a potential mechanism for a density-dependent effect at high levels of SSB. However, it appears to operate only in years when adult herring have extended residence in shallow water, which is not generally the case (Bailey and Steele, 1992). As a result, there is conflicting evidence for density-dependence in the stock–recruit relationship at high SSB values, based on data from the 1960s vs. data from the 2000s (Figure 3a). We speculate that climate variability causes interannual changes in prey availability and therefore the need for adult herring to resort to cannibalism in any one year. Directed field research is necessary to understand better the apparent ephemeral role of cannibalism in North Sea herring.

The analysis of potential SSB effects on R involves several non-trivial technical issues that warrant discussion. For instance, analysis of R solely by ARIMAX modelling may ignore an inherent SSB effect. Yet, a failure to remove a potential SSB effect from R is only problematic if SSB is independently influenced by climate (e.g. winter NAO and winter AMO) and functions as a mediator. There may be two ways to handle this situation: either directly to take into account a significant SSB effect by including SSB or to remove an inherent SSB effect from R. Plots of R on SSB reveal that this relation is usually non-linear, so the inclusion of SSB linearly into an ARIMAX model would not make sense. Therefore, we used a linearized approach instead, based on the extended Ricker curve [Equation (1)]. Alternatively, within the context of an ARIMAX model, the potential SSB influence may be removed by fitting a conventional R–SSB model leading to

$$\text{Residuals} = f(\text{climatic factors}), \tag{5}$$

in which the residuals stem from one of the three conventional R–SSB models. These residuals (R_{residual}) would then be used to fit an ARIMAX model. However, there are three reasons not to do this. First, if SSB is not linked directly to a time-trend being removed, then removing the SSB effect from R could make even stationary recruitment dataseries non-stationary; this would seriously

violate the assumptions of TS methods. Second, for large uncertainty in the stock–recruit relationship, none of the conventional models would fit well; this is currently the debate regarding conventional R–SSB models, as mentioned in the Introduction. Given this, it would be poor statistical practice to calculate residuals by removing a poorly fitted model from observed data because this would erode the quality of the recruitment data to be explained by exogenous factors. Third, our goal was to find a prognostic model of the type

$$\text{Recruitment} = f\,(\text{climatic factors}), \tag{6}$$

which can forecast future recruitment for prognostic purposes and not forecast residuals as in Equation (5), which cannot be converted into real recruitment values owing to the lack of a valid R–SSB model.

Our finding of non-statistically significant effects of SSB on R differs from those of Nash and Dickey-Collas (2005), who found that SSB was related to recruitment in 80% of years. It should be noted that our study makes use of a much longer, more complete TS (46 vs. 22 years) and includes modelled recruitment (VPA), providing more information than the survey data alone, as used by Nash and Dickey-Collas (2005). Before a full comparison can be made between these studies, it would be necessary to re-run the Nash and Dickey-Collas (2005) approach by including all years back to 1960 and by taking into account autocorrelation as we have done.

Perhaps the most striking feature of the SSB and R time-series is the decline in the 1960s and 1970s. The fishery certainly contributed because fishing mortality is estimated to have been very high then, perhaps accelerating a decline into a stock collapse (Bailey and Steele, 1992). However, the historical decline in North Sea herring is widely viewed today as resulting solely from overfishing. For overfishing to explain this decline fully, there needs to be clear causality in the relationship between R and SSB. However, not only does our analysis reveal uncertainty in this relationship, we identified a significant reverse causality $\text{SSB} = f(R)$, because additional cross-correlation experiments revealed significant peaks at negative lags of SSB with delays of 2 and 3 years, indicating a feedback situation (Bailey and Steele, 1992; Schlittgen, 2001). Rather, two findings indicate that year-class strength is largely determined at the larval stage, specifically after yolk-sac absorption. First, the abundance of small yolk-sac larvae is linearly correlated with SSB (Gröger *et al.*, 2001; ICES, 2007). Second, among our cross-correlations of MLAI and MIK indices with winter (December–March) values of AMO and NAO, only the MIK index showed significant spikes at exactly the same lags of winter NAO and winter AMO as for R. Taken together, this implies to us that abundances of eggs and small larvae are mainly determined by adult herring biomass, whereas the abundance of post yolk-sac larvae is mainly determined by environmental factors linked to climate.

We reach the same conclusion when we focus on the most recent decline of 2001–2006, which cannot be explained by high fishing mortality (from 2001 to 2006, all values of F-at-age-1 were <0.07, and all values of F averaged over ages 2–6 were <0.31; ICES, 2007). Additionally, the more recent decline cannot be explained by parent stock size, which remained relatively high. Another observation is that in years such as 1989–1990 and 2000–2005, the production of eggs and small larvae (MLAI) and SSB were quite high (Figure 2a), but the resultant abundances of larger larvae (MIK; Figure 2a) and recruitment were poor (Figure 2b). Moreover, we see similar trends of 16–24 mm larvae that are still on the hatching grounds compared with trends of large (~30 mm) larvae indexed by the MIK in the following spring. Collectively, these historical and recent observations suggest to us that recruitment is likely driven primarily by external factors, which most likely operate shortly after the yolk-sac stage and continue through the overwintering period until larger larvae are seen in early spring (MIK).

Cross-correlations suggest that environmental conditions related to winter NAO and winter AMO are largely responsible for the variability in R. In this case, lags of 3 (AMO) and 5 years (NAO) imply delayed effects of physical forcing on herring recruitment about which one can only speculate. These time-lags indicate to us that both physical processes are not directly influencing the larvae, but are instead operating via intermediate ecological processes. Such time-lags could develop from a combination of delayed development of ocean features, such as ocean currents and properties of the water mass (e.g. temperature and salinity), in response to NAO- and AMO-related physical forcing, or sequenced ecological processes involving food-chain effects. It is logical that the indirect influence of the AMO is faster because it indexes a large-scale sea surface effect, whereas the indirect influence of the NAO is slower because it indexes a large-scale atmospheric effect.

The NAO and AMO indices were chosen for this analysis because they reflect large-scale climate and ocean dynamics, thereby increasing the utility of our approach for other fish stocks in the North Atlantic. Decadal or multidecadal shifts in AMO and the recent seasonal shifts in high AMO SSTs and low NAO values during autumn reflect large-scale changes that likely translate into regional dynamics of ocean currents and water properties (e.g. temperature). These directly or indirectly interact with early life stages of fish stocks, including herring. One of the more well-known examples of this is the Russell Cycle, which describes major regime shifts in fish fauna in the English Channel associated with systematic changes in the NAO, ocean circulation, and zooplankton species composition and abundance (Cushing and Dickson, 1976; Southward, 1980). Whereas regime shifts provide extreme examples of the coupled response of marine ecosystems to climate change, regional changes in climate and oceanography are often more subtle; and individual fish stocks may respond to climate indirectly through the population dynamics of their prey organisms, so leading to more gradual and delayed responses.

Lagged effects involving the winter NAO are not limited to herring. For instance, the recruitment dynamics of pikeperch (*Sander lucioperca*) in a Baltic lagoon lagged changes in the winter NAO by 5 years, and temperature and salinity by 3–5 years (Gröger *et al.*, 2007a). Preliminary analyses indicated that these lagged effects were probably attributable to cycles of the prey of pikeperch (Gröger *et al.*, 2007a). An identical delay of 5 years involving the winter NAO was found for benthic organisms in the Baltic Sea (Gröger and Rumohr, 2006). In a more general context, Post (2004) distinguished between three fundamentally different mechanisms of population time-lag responses, including influences of (i) atmospheric processes, (ii) species life history, and (iii) trophic interactions. These are discussed below.

(i) Atmospheric processes—Lagged population responses to large-scale climatic variability may arise when the proximal abiotic factor influencing the population dynamics is itself

correlated with regional atmospheric processes at some time in the past (Post, 2004). Physical mechanisms functioning between global proxies and regional variables are complex. Depending on the region, the leading pattern of regional SST anomalies in the North Atlantic or the Gulf Stream is related directly to anomalous air–sea fluxes associated with the NAO with lags between 1 and 3 years (Taylor and Stephens, 1998; Watanabe and Kimoto, 2000; Conversi *et al.*, 2001; Marshall *et al.*, 2001).

(ii) Species life history—Climate-driven effects usually act only on specific components of the life cycle and could be manifest as lagged effects through reproduction (Post, 2004). In our case, winter NAO and winter AMO indices (lagged or non-lagged) do not show any statistically significant correlations with herring SSB. Ecologically this makes sense, because adult herring are expected to be less vulnerable to changes in their environment than early life stages. For instance, adults may swim away from unfavourable conditions, such as very cold temperatures perhaps associated with poor feeding conditions (Jacobsson and Østvedt, 1996; Maravelias, 1997). Of the early life stages, inability to avoid adverse conditions is most acute for the egg stage, during which anoxia events can cause mass mortality (Morrison *et al.*, 1991). With increasing body size of herring, resistance to starvation increases. The early life-history stages of herring are most critical for year-class success.

(iii) Trophic interactions—Interactions among species may also produce time-lags (Post, 2004). Because climate factors, such as atmospheric pressure, are unlikely to produce significant direct effects on recruitment (in our case, the CCFs between *R* and NAO as well AMO at lag 0 are insignificant), a plausible trophic linkage involves larval herring prey (e.g. the nauplii of *Calanus finmarchicus*). In the North Sea, zooplankton species composition and abundance are linked to increased (warm) or decreased (cold) inflows of nutrient-rich oceanic water from the North Atlantic (Beaugrand *et al.*, 2002; Reid *et al.*, 2003). Periods of enhanced inflows are associated with increased westerlies associated with changes in wind patterns and increased wind strength, phenomena highly correlated with the NAO and SSTs (Beaugrand *et al.*, 2002; Reid *et al.*, 2003). In the Gulf of Maine, *C. finmarchicus* abundance in summer was positively related to winter SST 2 years earlier or to the NAO index 4 years earlier (Conversi *et al.*, 2001).

In addition to the abundance and species composition of their prey, the timing of prey availability may be critical for larval feeding success and survival. For instance, the point of no return (the point at which fish become too weak to feed and recover) is just 5–6 d for newly hatched larvae and increases to 15 d ~90 d after hatching, when larvae are larger (Blaxter and Ehrlich, 1974). The sensitivity of first-feeding larvae to prey availability and quality led to Cushing's (1969, 1995) match–mismatch hypothesis, which posits that when the spring bloom is poorly matched with the relatively fixed spawning time of fish, poor survival will lead to a weak year class. This hypothesis may be supported by the fact that the marginal increase in North Sea SST during the past decade led to a significant (~1 month) delay of the spring phytoplankton bloom and a subsequent change in the zooplankton species composition, including the nauplii upon which herring larvae feed (Wiltshire and Bryan, 2004; see also Nash and Dickey-Collas, 2005). Climate effects on herring prey may, however, be twofold, either operating directly on their physiology (e.g. temperature regulation of growth and reproduction) or operating indirectly via oceanographic processes. Additional analyses (not shown here) suggest that SST in February and March is the proximal environmental variable responsible for the climate connection.

An alternative mechanism was proposed by Corten (1986), who suggested that climate-driven changes in North Sea circulation disrupted the advection of herring larvae from spawning to nursery grounds, contributing to the recruitment failures of the 1970s. Windstress drives ocean circulation patterns in the North Sea during winter; winds vary greatly, alternating between periods of strong southerly or southeasterly winds lasting for days to weeks, to westerly flows more typical of that time of year (Bartsch *et al.*, 1989). Therefore, although advection modelling showed that eastward transport was the dominant mode, patterns of larval dispersal vary annually depending on meteorological conditions (Bartsch *et al.*, 1989). Still, the relative role of larval advection on the annual variability in North Sea herring recruitment remains unknown.

Although details of the underlying operative mechanisms linking the NAO and AMO to herring recruitment remain to be resolved, an ideal feature of the ARIMAX model [Equation (4)] is that it can be used easily for short-term prediction as part of North Sea herring assessments. We simply need to replace the conventional *R*–SSB function (currently either a segmented regression or a Ricker curve) by our ARIMAX model and leave the remaining equations untouched (for an illustration, see Gröger *et al.*, 2007b). Forecasting the recruitment for the scenarios is then straightforward because the input values for NAO and AMO are known 5 and 3 years back, respectively, so do not need to be forecast. The shorter lag of 3 years fixes the maximum length of the planning horizon at 3 years. As an example, suppose we want to forecast the recruitment for the years 2010–2012. To predict R_{2010}, we need only insert the known values for NAO_{2005} and AMO_{2007} into Equation (4), for 2011 those of NAO_{2006} and AMO_{2008}, and for 2012 those of NAO_{2007} and AMO_{2009}. In contrast to this, for prognostic purposes, all other approaches [models (i)–(iii) and (v) as specified above] require independently forecast values of SSB. However, forecast error in SSB leads to additional forecast error in the recruitment model. Moreover, Equation (1) requires some back-transformation to obtain real *R* values from the predicted (forecast) $\ln(R/\mathrm{SSB})$ values and does not account for autocorrelation in the residuals; this most likely results in biased parameter estimates. The resulting short-term predictions from the ARIMAX model appear from the prognostic indicators to be rather reliable (Figure 6d and e).

Although our best model does not incorporate an SSB effect on *R* in the particular case of North Sea herring, this should not in any way be misconstrued from a fisheries management perspective. Evidence for density-dependence (i.e. a reduced frequency of good year classes) at low stock sizes strongly implies the need to avoid overfishing. Our results (see SA and Figure 5c) imply that initial declines in *R* in the 1960s and 1970s stemmed from environmental causes, but that the complicit effects of overfishing converted the SSB decline into stock collapse (Bailey and Steele, 1992). Stocks driven to such low levels, as in the 1970s, increase the risk of delayed recovery and prolonged periods of drastic

management measures, such as fisheries closures. Therefore, when the stock regeneration process is seemingly compromised, regardless of cause (e.g. overfishing, periods of poor recruitment attributable to unfavourable environmental conditions, or increases in natural mortality), it is incumbent on fishery managers to reduce the total mortality by reducing the exploitation rate (i.e. fishing mortality F). Management strategy evaluations (Kell *et al.*, 2005; A'mar *et al.*, 2008) provide an approach to evaluate quantitatively the trade-offs among alternative harvest strategies to sustain North Sea herring stocks and fisheries under the combined effects of fishing and climate fluctuations.

Acknowledgements

We are very grateful for the comments we received from anonymous reviewers, which greatly improved the paper. Also, we thank the editor for devoting so much time in helping us improve the presentation.

References

A'mar, Z. T., Punt, A. E., and Dorn, M. W. 2008. The management strategy evaluation approach and the fishery for walleye pollock in the Gulf of Alaska. *In* Resiliency of Gadid Stocks to Fishing and Climate Change, pp. 317–346. Ed. by G. H. Kruse, K. Drinkwater, J. N. Ianelli, J. S. Link, D. L. Stram, V. Wespestad, and D. Woodby. Alaska Sea Grant, University of Alaska Fairbanks, AK-SG-08-01, Fairbanks, Alaska, USA.

Bailey, R. S., and Steele, J. H. 1992. North Sea herring fluctuations. *In* Climate Change, Climate Variability and Fisheries, pp. 213–230. Ed. by M. H. Glantz. Cambridge University Press, Cambridge, UK.

Bartsch, J., Brander, K., Heath, M., Munk, P., Richardson, K., and Svendsen, E. 1989. Modelling the advection of herring larvae in the North Sea. Nature, 340: 632–636.

Beaugrand, G., Reid, P. C., Ibañez, F., Lindley, J. A., and Edwards, M. 2002. Reorganization of North Atlantic marine copepod biodiversity and climate. Science, 296: 1692–1694.

Beverton, R. J. H., and Holt, S. J. 1957. On the dynamics of exploited fish populations. Fishery Investigations, London, Series II, 19. 533 pp.

Blaxter, J. H. S., and Ehrlich, K. F. 1974. Changes in behaviour during starvation of herring and plaice larvae. *In* The Early Life History of Fish, pp. 575–588. Ed. by J. H. S. Blaxter. Springer, Berlin.

Box, G. E. P., and Jenkins, G. M. 1970. Time Series Analysis—Forecasting and Control. Holden Day, San Francisco.

CEC. 2001. Green Paper on the future of the Common Fisheries Policy. Commission of the European Communities, Green Paper Volume 1, 2.

Conversi, A., Piontkovski, S., and Hameed, S. 2001. Seasonal and interannual dynamics of *Calanus finmarchicus* in the Gulf of Maine (Northeastern US shelf) with reference to the North Atlantic Oscillation. Deep Sea Research II, 48: 519–530.

Corten, A. 1986. On the causes of the recruitment failure of herring in the central and northern North Sea in the years 1972–1978. Journal du Conseil International pour l'Exploration de la Mer, 42: 281–294.

Cushing, D. H. 1969. The regularity of the spawning season of some fishes. Journal du Conseil International pour l'Exploration de la Mer, 33: 81–92.

Cushing, D. H. 1995. Population Production and Regulation in the Sea: a Fisheries Perspective. Cambridge University Press, Cambridge, UK.

Cushing, D. H., and Dickson, R. R. 1976. The biological response in the sea to climate changes. Advances in Marine Biology, 14: 1–122.

Enfield, D. B., Mestas-Nunez, A. M., and Trimble, P. J. 2001. The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental US. Geophysical Research Letters, 28: 2077–2080.

Fahrmeir, L., Hamerle, A., and Tutz, G. 1996. Multivariate statistische Verfahren. Walter de Gruyter, Berlin.

Gröger, J. P., Rountree, R. A., Missong, M., and Rätz, J. 2007b. A stock rebuilding algorithm featuring risk assessment and an optimization strategy of single or multispecies fisheries. ICES Journal of Marine Science, 64: 1–15.

Gröger, J. P., and Rumohr, H. 2006. Modelling and forecasting long-term dynamics of western Baltic macrobenthic fauna in relation to climate signals and environmental change. Journal Sea Research, 55: 266–277.

Gröger, J. P., Schnack, D., and Rohlf, N. 2001. Optimisation of survey design and calculation procedure for the international herring larvae survey in the North Sea. Archive of Fishery and Marine Research, 49: 103–116.

Gröger, J. P., Winkler, H., and Rountree, R. A. 2007a. Population dynamics of pikeperch (*Sander lucioperca*) and its linkage to fishery driven and climatic influences in a southern Baltic lagoon of the Darss–Zingst Bodden Chain. Journal of Fisheries Research, 84: 189–201.

Hartung, J. 2002. Statistik. Oldenbourg, München, Germany.

Heath, M., Leaver, M., Matthews, A., and Nicoll, N. 1989. Dispersion and feeding of larval herring (*Clupea harengus* L.) in the Moray Firth during September 1985. Estuarine, Coastal and Shelf Science, 28: 549–566.

Hurrell, J. W., and Dickson, R. R. 2004. Climate variability over the North Atlantic. *In* Marine Ecosystems and Climate Variability, pp. 15–46. Ed. by N. Ch. Stenseth, G. Ottersen, J. W. Hurrell, and A. Belgrano. Oxford University Press, Oxford, UK.

Hurrell, J. W., and van Loon, H. 1997. Decadal variations in climate associated with the North Atlantic Oscillation. Climate Change, 36: 301–326.

Hurvich, C. M., and Tsai, C. L. 2008. A corrected Akaike information criterion for vector autoregressive model selection. Journal of Time-series Analysis, 14: 271–279.

ICES. 2007. Report of the Herring Assessment Working Group South of 62°N (HAWG). ICES Document CM 2007/ACFM: 11.

Jacobsson, J., and Østvedt, O. J. 1996. A preliminary review of the joint investigations on the distribution of herring in the Norwegian and Icelandic Seas 1950–1970. ICES Document CM 1996/H: 14.

Kell, L. T., Pilling, G. M., Kirkwood, G. P., Pastoors, M., Mesnil, B., Korsbrekke, K., Abaunza, P., *et al.* 2005. An evaluation of the implicit management procedure used for some ICES roundfish stocks. ICES Journal of Marine Science, 62: 750–759.

Kruse, G. H., and Tyler, A. V. 1989. Exploratory simulation of English sole recruitment mechanisms. Transactions of the American Fisheries Society, 118: 101–118.

Maravelias, C. D. 1997. Trends in abundance and geographic distribution of North Sea herring in relation to environmental factors. Marine Ecology Progress Series, 159: 151–164.

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrell, J., *et al.* 2001. North Atlantic climate variability: phenomena, impacts and mechanisms. International Journal of Climatology, 21: 1863–1898.

Morrison, J. A., Napier, I. R., and Gamble, J. C. 1991. Mass mortality of herring eggs associated with a sedimenting diatom bloom. ICES Journal of Marine Science, 48: 237–245.

Nash, R. D. M., and Dickey-Collas, M. 2005. The influence of life history dynamics and environment on the determination of year class strength in North Sea herring (*Clupea harengus* L.). Fisheries Oceanography, 14: 279–291.

Pindyck, R. S., and Rubinfeld, D. L. 1991. Econometric Models and Economic Forecasts. McGraw-Hill, New York.

Post, E. 2004. Time lags in terrestrial and marine environments. *In* Marine Ecosystems and Climate Variability, pp. 165–167. Ed. by

N. Ch. Stenseth, G. Ottersen, J. W. Hurrell, and A. Belgrano. Oxford University Press, Oxford, UK.

Pyper, B. J., and Peterman, R. M. 1998. Comparison of methods to account for autocorrelation in correlation analysis of fish data. Canadian Journal of Fisheries and Aquatic Sciences, 55: 2127–2140.

Reid, P. C., Edwards, M., Beaugrand, G., Skogen, M., and Stevens, D. 2003. Periodic changes in the zooplankton of the North Sea during the twentieth century linked to oceanic inflow. Fisheries Oceanography, 12: 260–269.

Rogers, J. C. 1984. The association between the North Atlantic Oscillation and the southern oscillation in the northern hemisphere. Monthly Weather Review, 112: 435–446.

Rothschild, B. J., and Shannon, L. J. 2004. Regime shifts and fishery management. Progress in Oceanography, 60: 397–402.

Schlittgen, R. 2001. Angewandte Zeitreihenanalyse. Oldenbourg, München, Germany.

Schlittgen, R., and Streitberg, B. H. J. 2001. Zeitreihenanalyse. Oldenbourg, München, Germany.

Southward, A. J. 1980. The western English Channel—an inconstant ecosystem? Nature, 185: 361–366.

Stenseth, N. Ch., Ottersen, G., Hurrell, J. W., and Belgrano, A. 2004. Marine Ecosystems and Climate Variation. Oxford University Press, Oxford, UK.

Taylor, A. H., and Stephens, J. A. 1998. The North Atlantic Oscillation and the latitude of the Gulf Stream. Tellus, 50A: 134–142.

Tunberg, B., and Nelson, W. G. 1998. Do climatic oscillations influence cyclical patterns of soft bottom macrobenthos communities on the Swedish west coast? Marine Ecology Progress Series, 170: 85–94.

Watanabe, M., and Kimoto, M. 2000. On the persistence of decadal SST anomalies in the North Atlantic. Journal of Climate, 13: 3017–3028.

Wiltshire, K. H., and Bryan, F. J. M. 2004. The warming trend at Helgoland Roads, North Sea: phytoplankton response. Helgoland Marine Research, 58: 269–273.

doi:10.1093/icesjms/fsp259