This appendix accompanies the article

Cod and climate: effect of the North Atlantic Oscillation on recruitment in the North Atlantic

Leif Christian Stige[1], Geir Ottersen[2,3], Keith Brander[3], Kung-Sik Chan[4], Nils Chr. Stenseth[1,5,6,*]

[1]Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway

[2]Institute of Marine Research, PO Box 1870, 5817 Bergen, Norway

[3]International Council for the Exploration of the Sea, Palægade 2-4, 1261 Copenhagen K, Denmark

[4]Department of Statistics and Actuarial Science, University of Iowa, 263 Schaeffer Hall, Iowa City, Iowa 52242, USA

[5]Department of Coastal Zone Studies, Institute of Marine Research, Flødevigen Marine Research Station, 4817 His, Norway

[6]*Present address:* Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway

*Corresponding author. Email: n.c.stenseth@bio.uio.no

Marine Ecology Progress Series 325:227–241 (2006)

The appendix includes some supplementary text on the robustness of the obtained results and on the methodology used, plots showing time series of spawning stock biomass (SSB) and recruitment for each stock (Fig. A1), residual diagnostics (Figs. A2-A3), combined effects of SSB and year (Figs. A4-A5), interaction effects between SSB and climate (Fig. A6), and results if SSB effects were set to zero for 7 stocks not displaying any R-SSB relationship (Figs. A7-A9), and tables with results when auto-regressive and year effects are excluded from the model (Table A1) or SSB effects were set to zero for 7 stocks not displaying any R-SSB relationship (Table A2).

SUPPLEMENTARY TEXT

Robustness of results to the inclusion of SSB effects for stocks for which no detectable association between recruitment and SSB exists

It has been argued that using recruitment per spawning stock biomass (R/SSB) as response when no significant relation between R and SSB exists may mask possible recruitment-environment relationships (Cardinale & Hjelm 2006). Although we used $\log_e$(R/SSB) as response, we do not think this argument applies to our analyses. We think so because SSB was included among the predictors, and the model thus allowed modelling near-zero responses to SSB (as is evident for example from Fig. A5). Nonetheless, to ascertain the robustness of the results, we performed an analysis in which SSB effects were excluded from the model for the stocks for which no detectable association between R and SSB existed. We also decided to omit effects of year for stocks for which no detectable association between R and year existed.

To perform this analysis, we modified the final model (corresponding to Eq. 4) as follows:

$$\log_e(R_{i,j}) = a_i + k_i \log_e(SSB_{i,j}) + b_i\, SSB_{i,j} + c_i \log_e(R_{i,j-1}) + d_i(\mathrm{Year}) + e\, NAO_j + f(\mathrm{Long}_i,\mathrm{Lat}_i)\, NAO_j + g(\mathrm{Year}_j,\mathrm{Long}_i,\mathrm{Lat}_i)\, NAO_j + \_{i,j}. \qquad \text{(A1)}$$

For stocks displaying an R-SSB relationship, the constant k was fixed at 1 and the constant b was estimated from the data, modelling a Ricker-type response. For stocks that did not display any R-SSB relationship, both the constants k and b were set to 0. Similarly, the function d was set to zero for stocks that did not display any effect of year. To select which stocks should have SSB effects, we first substituted the two SSB-terms with a stock-specific smooth effect of $\log_e$(SSB) (in a model not including NAO or year effects). The smooth effects were selected "automatically" by adding a shrinkage term to the roughness penalty. The penalty was estimated by minimizing the GCV. If the true function was linear for a given stock, the shrinkage term would dominate the penalty. And if the true function was identically zero, a high penalty would likely be selected, inducing high shrinkage that might shrink the function to a zero function (estimated degrees of freedom close to 0; c.f. the help manual of gam.selection in the mgcv library of R developed by Wood 2001). For seven stocks the estimated degrees of freedom were less than 0.1 (Gulf of Maine, S Grand Bank, Flemish Cap, S Gulf St. Lawrence, N Gulf St. Lawrence, Baltic E, Faroe). These stocks thus displayed no relation between R and SSB, wherefore we set their values of k and b in Eq. A5 to zero. We then determined which year effects to exclude, by adding a shrinkage term to the penalty of the smooth effect of year (d in Eq. A1). All stocks displayed effects of year (estimated df $\geq$ 0.5), wherefore the d term was retained for all stocks. Finally, we added NAO effects to the model.

We found that results regarding the effects of the NAO were qualitatively similar whether or not the SSB effects were excluded from the model for the seven stocks that did not display any R-SSB relationship (Table A2, Figs. A7-A9). We therefore judged the results to be robust to the above-mentioned potential problem.

Generalized Cross Validation (GCV)

The method of cross validation (CV) consists of setting aside a data case, predicting its response value by the model fitted to the other data cases, computing the squared predictive error, and then repeating this procedure for each data case to obtain the average sum of squared predictive error:

$$CV = \frac{1}{n}\sum (y_i - \hat{y}_i^{(-i)})^2 \tag{A2}$$

where $\hat{y}_i^{(-i)}$ is the predicted value of the *i*th data case from the model fitted to all data except the *i*th data case. The calculation can be speeded up by using the formula:

$$CV = \frac{1}{n}\sum (y_i - \hat{y}_i)^2 / (1 - a_i)^2 \tag{A3}$$

where a_i is the *i*th diagonal element of the hat (influence) matrix, **H**:

$$\mathbf{H} = \mathbf{X}\,(\mathbf{X'X})^{-1}\,\mathbf{X'} \tag{A4}$$

where **X** is the vector of predictor variables. Each of the a_i elements measures the distance between a data point and the centroid of the *X*-space, and thus the influence of the data case.

The generalized cross validation is calculated by replacing the inflation factors a_i by their average value, $\overline{a}$:

$$GCV = \frac{1}{n}\sum (y_i - \hat{y}_i)^2 / (1 - \overline{a})^2 \tag{A5}$$

Besides being computationally simpler, the GCV has the advantage compared to the CV that it is invariant with respect to orthogonal transformations of the space to which the response vector and the covariate vectors belong (Wahba 1990).

The GCV of a model is thus a proxy for the model's out-of-sample predictive mean squared error. A model with lower GCV has more explanatory power, and hence is preferred, compared to a model with higher GCV. The GCV is analogous to Akaike's Information

Criterion (AIC, Akaike 1974) in that it aims at optimising the trade-off between the numbers of parameters in a model and the goodness-of-fit of the model.

In our analyses the GCV criterion was used for two purposes: (i) find the optimal roughness of each smooth term in a generalized additive model (Wood 2001, 2004), and (ii) find the optimal model structure, e.g. when comparing models with different predictor variables.

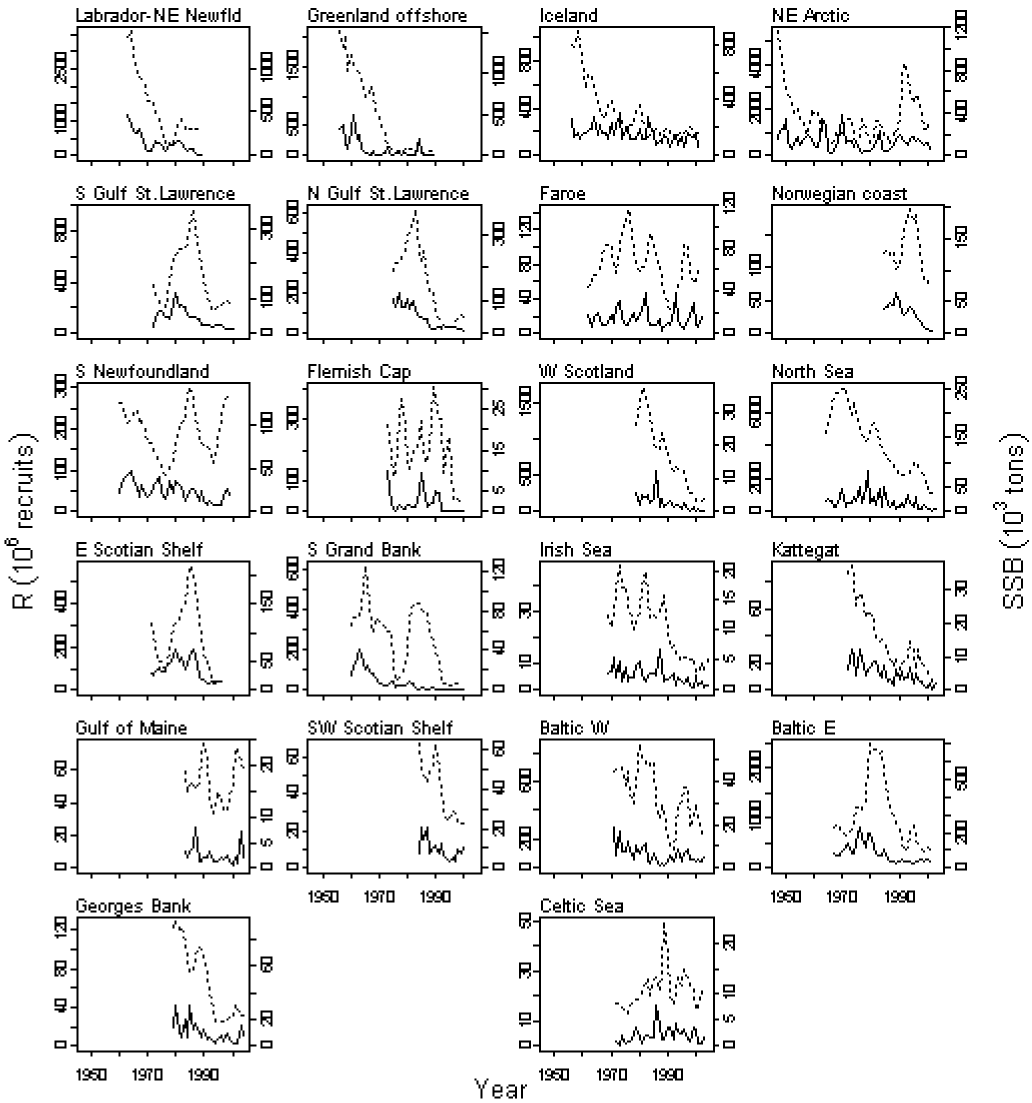

Fig. A1. Estimated annual number of recruits (R; whole lines) and spawning stock biomass (SSB; stippled lines) for each cod stock. For R, year refers to the spawning year. See Table 2 in main text for data sources.

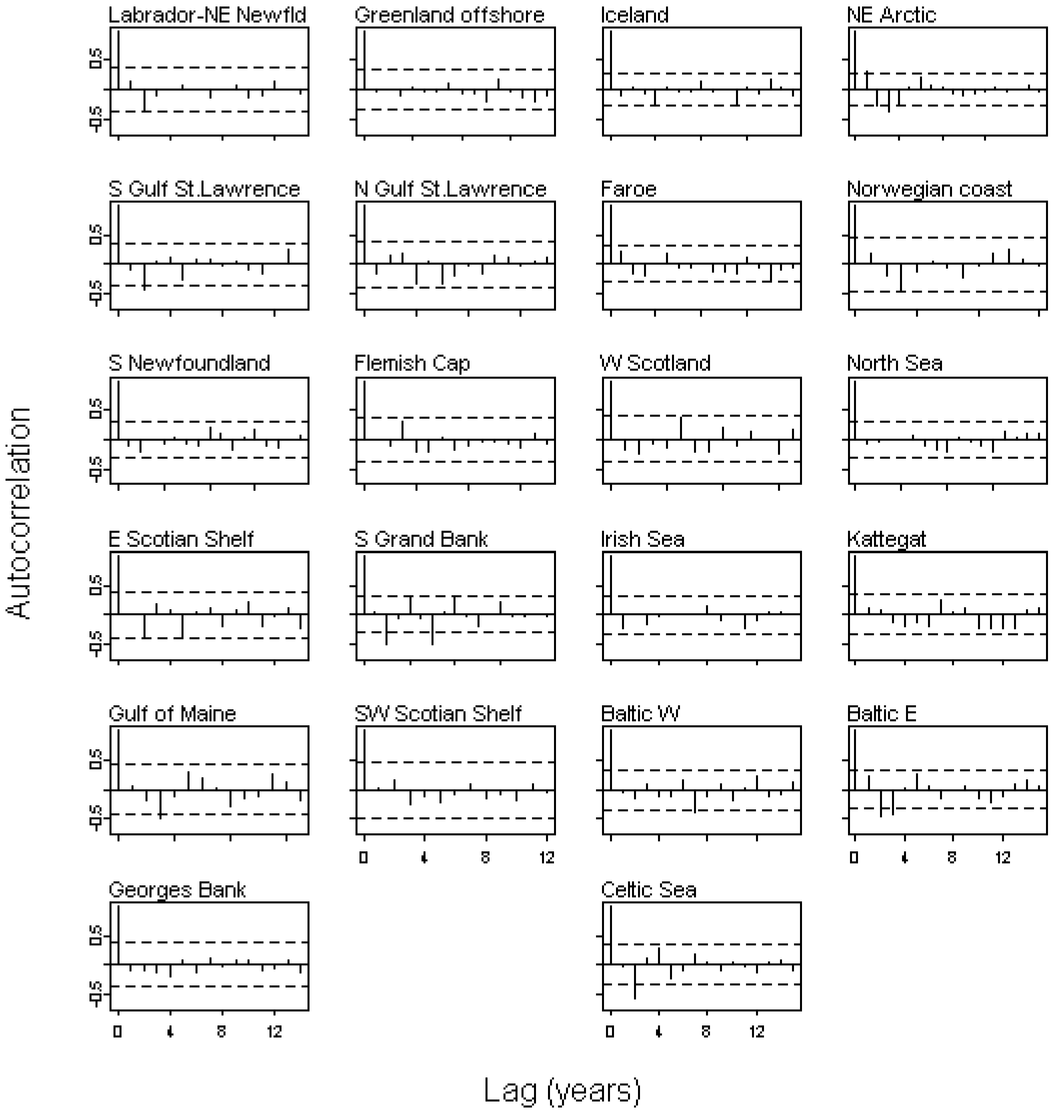

Fig. A2. Stock-specific autocorrelation functions of residuals from the final cod recruitment model (Eq. 3). Stippled lines: 95% confidence intervals.

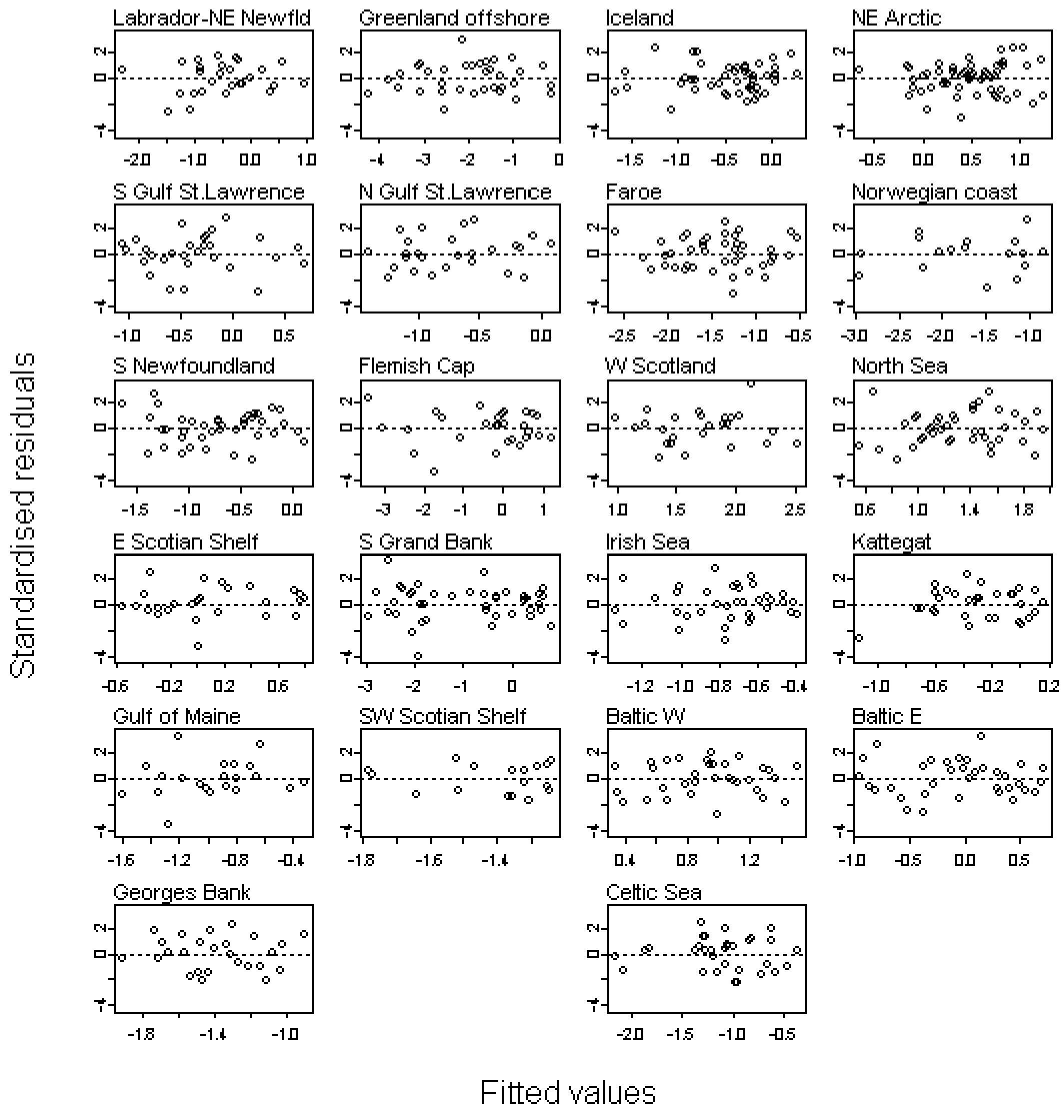

Fig. A3. Residuals and fitted values from the final cod recruitment model (Eq. 3). The residuals are weighted by an iteratively determined function of stock and spawning stock biomass.

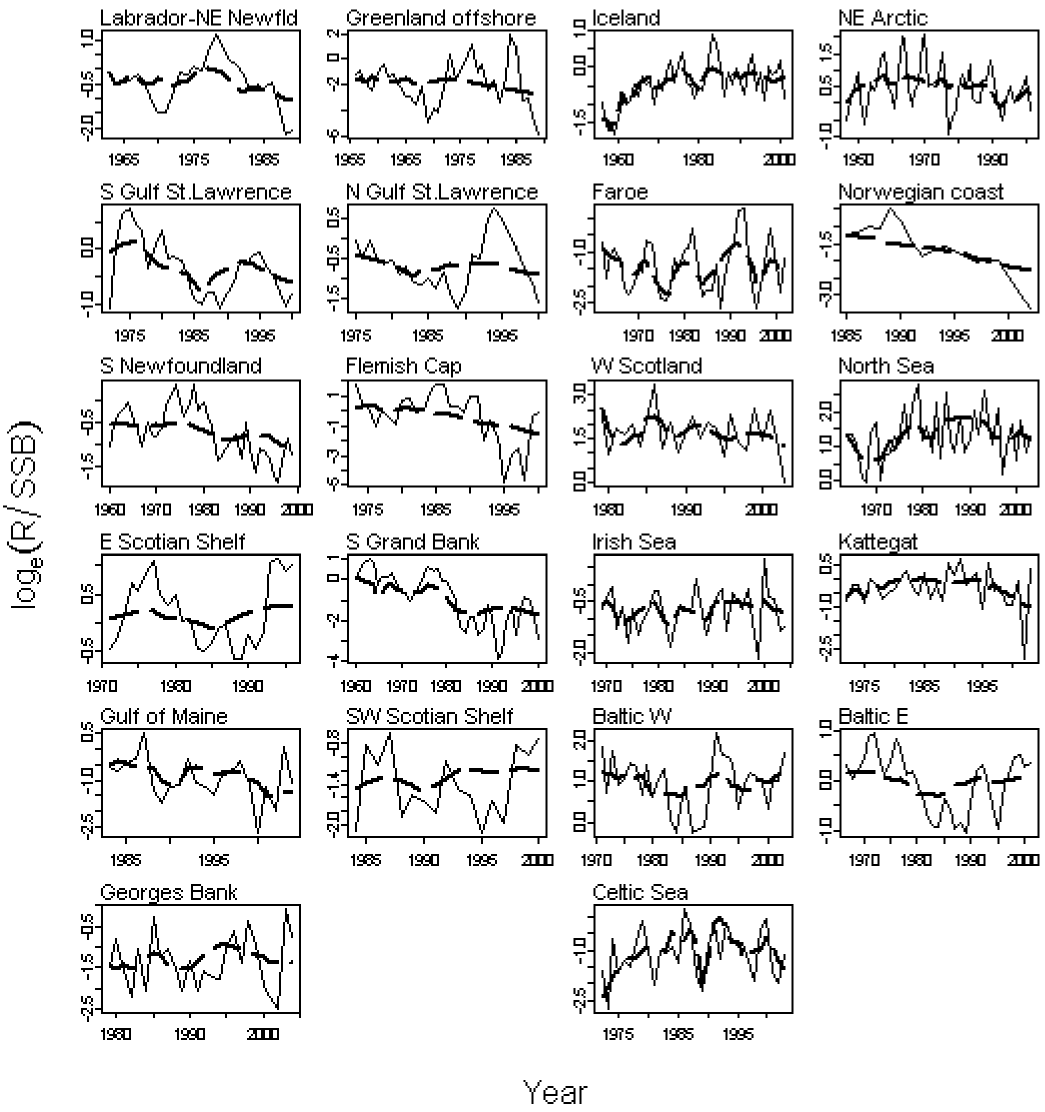

Fig. A4. The combined effects of year and spawning stock biomass on recruitment ($\log_e$[R/SSB]). R is annual number of recruits and SSB is spawning stock biomass. Whole, thin lines: observed time trends. Bold, broken lines: the predicted combined effect of year and SSB in a generalized additive model (Eq. 2) at NAO (North Atlantic Oscillation Index) = 0.

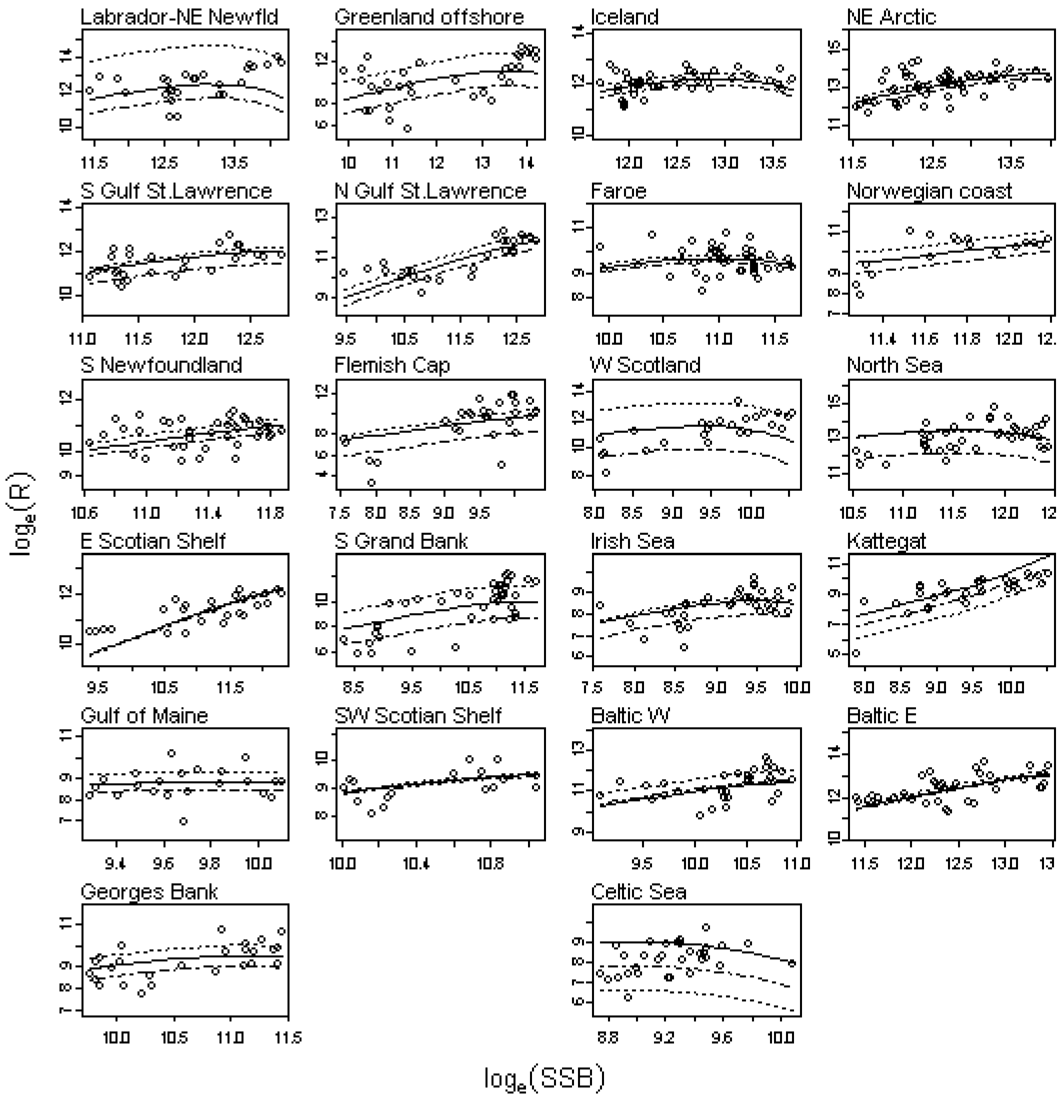

Fig. A5. The effects of spawning stock biomass (SSB) and year on annual numbers of recruits (R). Stippled, whole and broken lines are, respectively, predictions for the first, middle and last year with data for each stock. The predictions are from a generalized additive model (Eq. 2) at NAO (North Atlantic Oscillation Index) = 0. See Fig. 3 in the main text for the predicted year effects on recruitment when SSB is fixed.

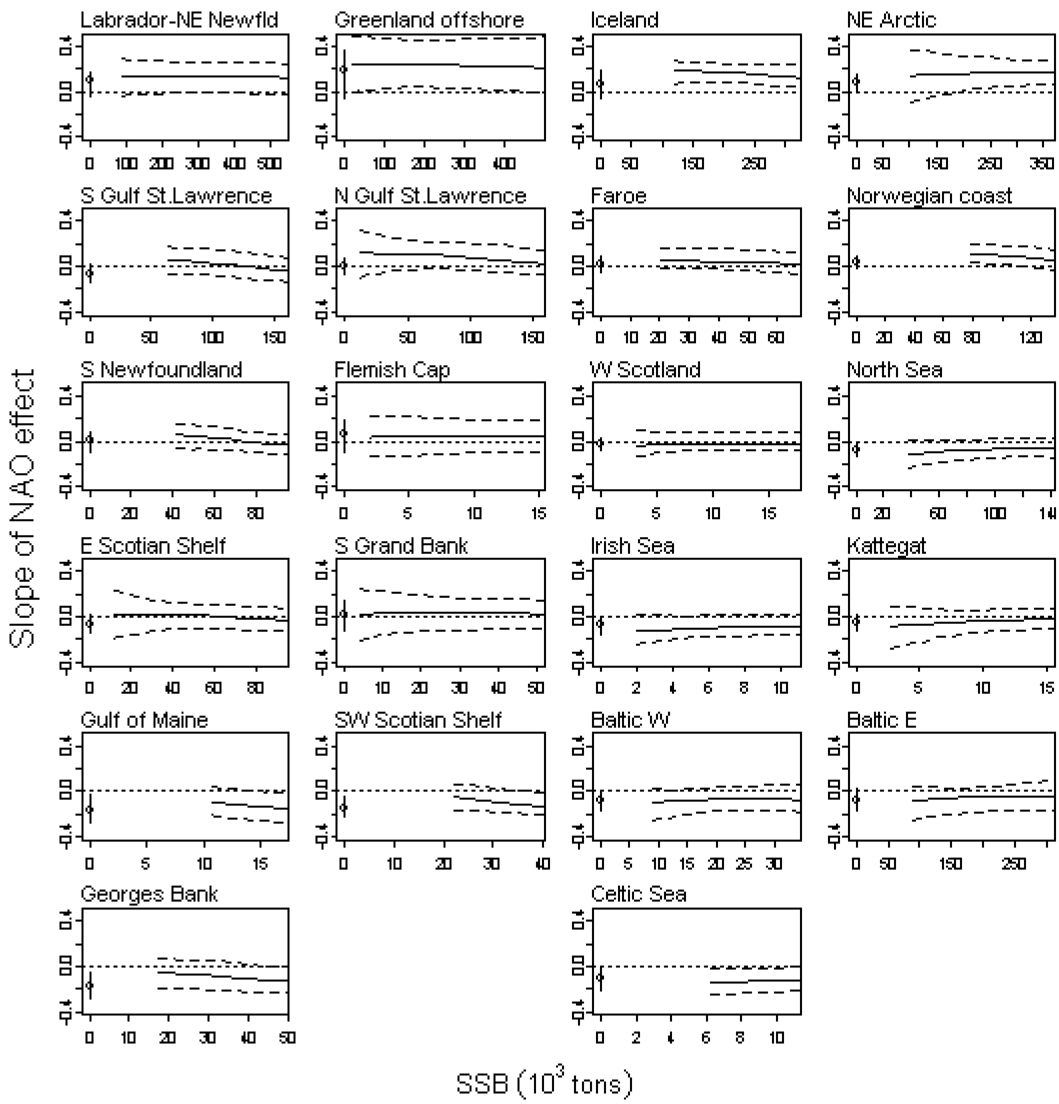

Fig. A6. Interaction between the effects of spawning stock biomass and the North Atlantic Oscillation (NAO) on cod recruitment. Each panel shows the estimated slope of a linear effect of NAO on $\log_e$(R/SSB), where R is annual number of recruits and SSB is spawning stock biomass. Whole and broken lines: predictions and 95% bootstrap confidence bands from a model in which the NAO effect depends on SSB (Eq. 4). Points and bars: predictions and 95% bootstrap confidence limits from a model in which the NAO effect is fixed (Eq. 2; Fig. 4).

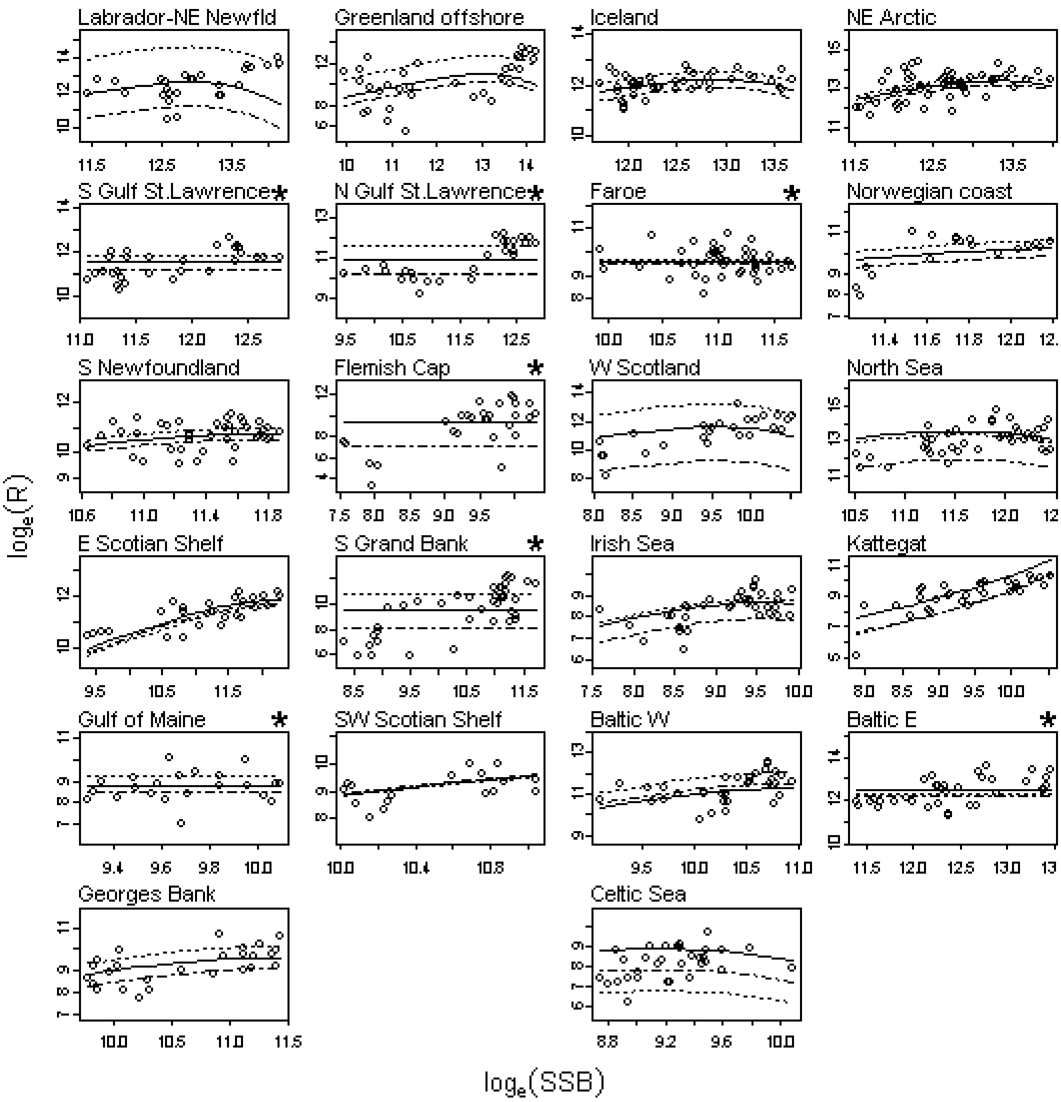

Fig. A7. The effects of spawning stock biomass (SSB) and year on annual numbers of recruits (R). Stippled, whole and broken lines are, respectively, predictions for the first, middle and last year with data for each stock. The predictions are from a generalized additive model in which SSB effects were set to zero for 7 stocks not displaying any R-SSB relationship (marked by asterisks). See Supplementary text for details. The figure is analogous to Fig. A5.

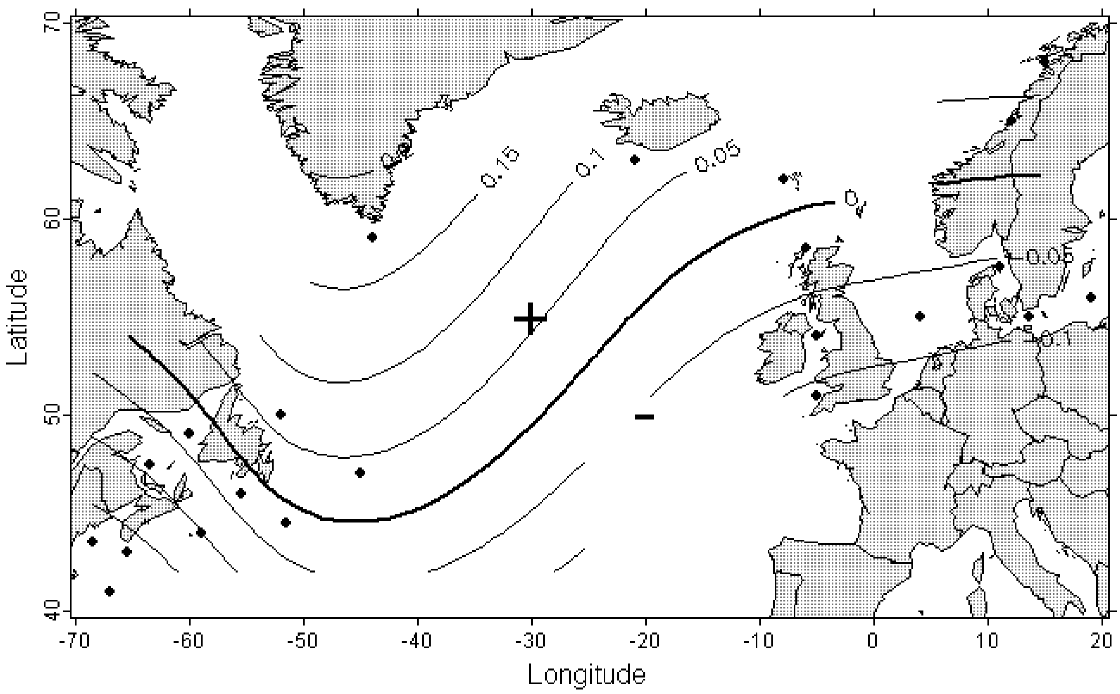

Fig. A8. Spatial pattern of the effect of the North Atlantic Oscillation (NAO) on cod recruitment. The isoclines represent the slope of a linear effect of NAO on $\log_e(R)$, where R is annual number of recruits. SSB is spawning stock biomass. The plot is based on a model in which SSB effects were set to zero for 7 stocks not displaying any R-SSB relationship (see Supplementary text), and is analogous to Fig. 4.

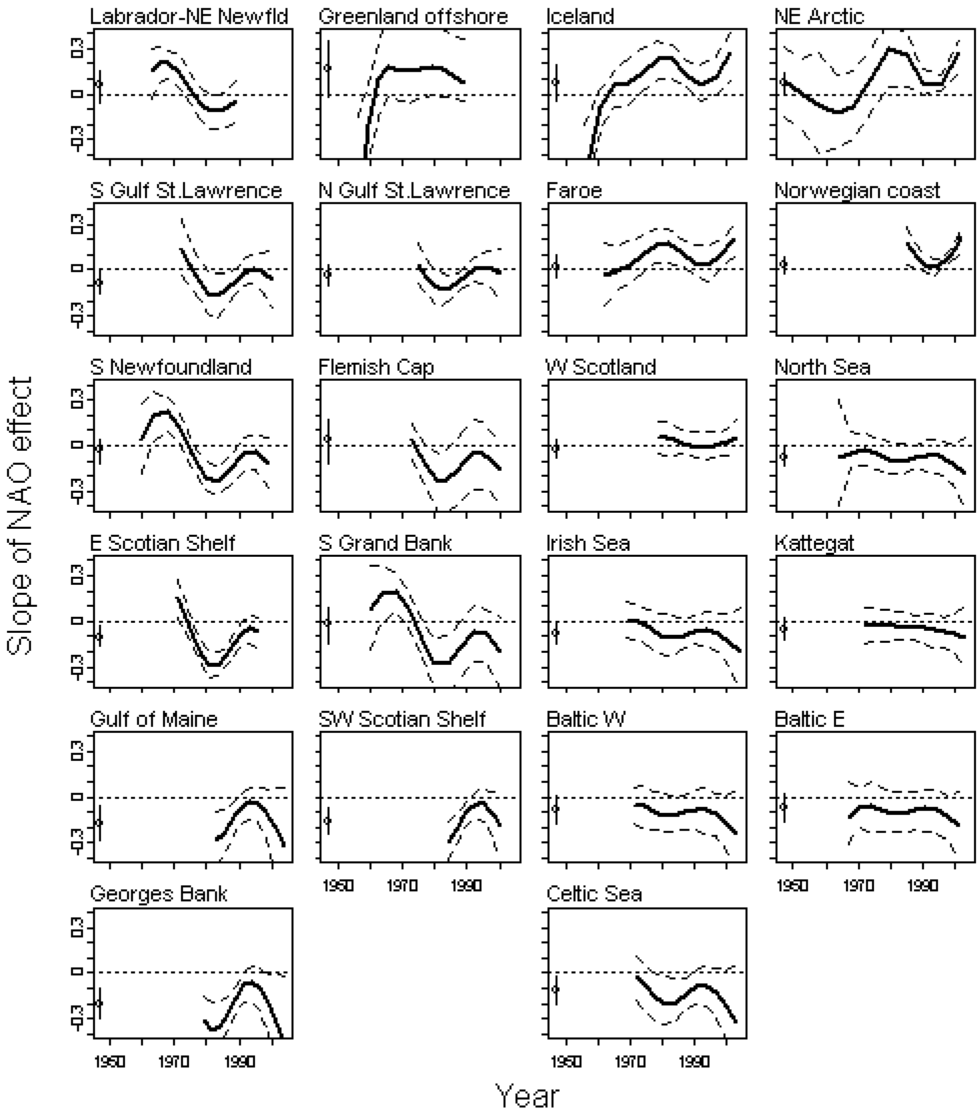

Fig. A9. Temporal change in the impact of the North Atlantic Oscillation (NAO) on cod recruitment. The lines show estimated slopes of a linear effect of NAO on $\log_e(R)$, where R is annual number of recruits. The effect of NAO is modelled to be linear for any given location and year, but the slope of the effect may vary spatially and temporally. Bold and broken lines: estimated slope and 95% bootstrap confidence bands for spatially and temporally varying NAO effect. Points and bars: estimated slope and 95% bootstrap confidence limits for spatially varying NAO effect. The plot is based on a model in which effects of spawning stock biomass (SSB) were set to zero for 7 stocks not displaying any R-SSB relationship (Eq. A1), and is analogous to Fig. 5.

Table A1. Comparison of original results (Table 3 in main text) with results obtained when auto-regressive and year effects were not included in the model ("alternative results"). See Table 3 for explanations of acronyms and terms.

	Original results		Alternative results	
Predictors	GCV	R^2	GCV	R^2
Stock	4.43	57.1%	4.43	57.1%
Spawning stock biomass	3.82	65.3%	3.82	65.3%
Auto-regressive effect	2.43	79.3%	-	-
Year effect	2.18	83.1%	-	-
Fixed NAO effect	2.14	83.8%	3.65	67.1%

Non-linear NAO effect	2.14	83.8%	3.65	67.7%
NAO effect _ SSB	2.12	84.3%	3.66	67.3%
NAO effect _ Year	2.07	84.8%	3.58	71.6%

Table A2. Comparison of original results (Table 3 in main text) with results obtained when SSB effects were set to zero for 7 stocks not displaying any R-SSB relationship ("alternative results"). See Supplementary text and Table 3 for details and explanations of acronyms and terms. Note that in the original analyses the response variable was $\log_e(R/SSB)$ while in the alternative analysis it was $\log_e(R)$. The GCV and R^2 values are therefore not directly comparable across analyses, although the relative merits of the different sub-models are.

	Original results		Alternative results	
Predictors	GCV	R^2	GCV	R^2
Stock	4.43	57.1%	5.84	79.9%
Spawning stock biomass	3.82	65.3%	4.68	83.8%
Auto-regressive effect	2.43	79.3%	2.46	92.0%
Year effect	2.18	83.1%	2.16	93.7%
Fixed NAO effect	2.14	83.8%	2.09	94.0%

Non-linear NAO effect	2.14	83.8%	2.12	93.9%
NAO effect _ SSB	2.12	84.3%	2.07	94.3%
NAO effect _ Year	2.07	84.8%	2.01	94.4%

LITERATURE CITED

Akaike H (1974) A new look at statistical model identification. IEEE Transactions on Automatic Control AU-19:716-722

Cardinale M, Hjelm J (2006) Marine fish recruitment variability and climate indices. Mar Ecol Prog Ser 309:307-309

Wahba G (1990) Spline models for observational data. SIAM, Philadelphia

Wood SN (2001) mgcv: GAMs and Generalized Ridge Regression for R. R News 1/2:20-25

Wood SN (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models. J Am Stat Ass 99:673-686