Breaking down the climate effects on cod recruitment by principal component analysis and canonical correlation

Paolo Sirabella\textsuperscript{1}, Alessandro Giuliani\textsuperscript{2}, Alfredo Colosimo\textsuperscript{1,*}, Joachim W. Dippner\textsuperscript{3}

\textsuperscript{1}University of Rome ‘La Sapienza’, Department of Biochemical Sciences - P.le A. Moro 5, 00185 Roma, Italy
\textsuperscript{2}Istituto Superiore di Sanitá - TCE Laboratory, V.le Regina Elena, 199, 00161 Roma, Italy
\textsuperscript{3}Institut für Ostseeforschung Warnemünde, Seestr. 15, 18119 Rostock, Germany

ABSTRACT: The pattern of temporal correlations between cod recruitment and sea temperature, in conjunction with the climate variability of atmospheric pressure anomalies (NAO index) was investigated by means of a combined use of principal component analysis (PCA) and canonical correlation analysis (CCA), using time series collected in the area surrounding the Kola peninsula (Barents Sea) and in the North Sea. The proposed data analysis strategy, namely to carry out a PCA of the temperature, cod recruitment and NAO time series followed by a CCA between the component spaces of all the possible data sets couples (recruitment vs temperature, recruitment vs NAO and NAO vs temperature), allowed us to sketch a general model of correlation between climate and cod recruitment dynamics. Two independent effects of temperature variability on cod recruitment emerged for the Kola region, pointing to the existence of at least 2 different mechanisms of comparable importance by which temperature may affect cod recruitment. In the North Sea the situation is somewhat simpler, and the data are compatible with only 1 major interaction mechanism. Moreover, the general effect of temperature on cod recruitment was opposite in the 2 regions: direct correlation for the Barents Sea, inverse correlation for the North Sea. This is probably due to the existence of an optimal temperature regime for cod recruitment lying in between the ‘cold’ Barents Sea and the ‘warm’ North Sea.

KEY WORDS: Cod recruitment · Climate variability · Descriptive ecology · Multivariate analysis

INTRODUCTION

Variability in marine ecosystems appears in time and space and is influenced by physical forcing, biological interactions and anthropogenic signals. An important question in order to understand interannual variability in regional seas is how much of the observed interannual and interdecadal variability can be directly addressed to climatic variability (Cushing & Dickson 1976). Studies on the impact of climate change and climate variability to the marine ecosystem exist for the northwest Atlantic (Frank et al. 1990) and for the west coast of the United States, where the influence of the 1982/83 El Niño event on phyto- and zooplankton and benthic communities has been investigated (Avaria & Muñoz 1987, Carrasco & Santander 1987, Lubchenco et al. 1993, Peterson et al. 1993). An important result of these studies is that with increasing temperature a shift occurs in the species composition from cold to warm water adapted species. A similar ecosystem response to climate variability has been detected for the North Atlantic. Fromentin & Planque (1996) correlated the abundance of 2 Calanus species to the NAO index, and Taylor (1995) demonstrated the existence of a correlation between the Gulf Stream index and the zooplankton in the NE Atlantic Ocean.

*Corresponding author. E-mail: colosimo@caspur.it

© Inter-Research 2001
As for the specific influence of temperature on cod populations, Saetersdal & Loeng (1987) found that strong year classes of the Arcto-Norwegian cod tended to occur in the warm years and proposed, as a causal mechanism, that in warm periods the gonad production increased, resulting in better recruitment. Ellertsen et al. (1989) using a virtual population analysis (VPA) approach concluded that, since year class strength was always low in cold years, while in warm years both good and bad year classes occurred, a high temperature is a necessary but not sufficient condition for the formation of strong year classes. As a causal mechanism, these authors indicated the oscillations in production of Calanus nauplii, the main prey item for cod larvae. Other authors (Sundby 1998) proposed a link between recruitment and climate variability via a temperature effect on growth rate.

A brief comment to these studies, in agreement with Sundby (1998), may be that more than 1 mechanism is probably in operation at the same time, thus giving rise to the establishment of complex nonlinear relationships between climate variability and cod recruitment. In such a situation, any proposed mechanism should rely on as robust as possible correlation structures among different data sets. This fact implies that any approach able to provide ecologists with (1) reliable estimates of the minimal number of independent components playing a role in the studied phenomenon and (2) useful clues to experimentally validate the sketched hypotheses, without imposing any predefined functional model to the data, should be welcome.

The data collected in the Barents Sea (see Fig. 1) by the Russian Kola peninsula hydrographic station (Bochkov 1982, Tereshchenko 1996, Ådlandsvik & Loeng 1991) present a unique opportunity to couple over a sufficiently long time (around 50 points on a yearly basis) high-quality information on both sea temperature and Arcto-Norwegian cod recruitment. For comparative and checking purposes, the corresponding data relative to the North Sea have been analyzed in parallel, to put in light possible differences/similarities between 2 independent geographical and ecological situations (Fig. 1). In both regions, however, the correlation between cod recruitment and temperature time series did not reach statistical significance (Pearson r = 0.28 for the Barents Sea and −0.27 for the North Sea). Cod recruitment and sea temperature data relative to the above regions were supplemented by large-scale climate information represented by the time series of the variability of the North Atlantic Oscillation (NAO), the most dominant mode of atmospheric variation in the whole North Atlantic area (Walker & Bliss 1932, Barnston & Liezvey 1987, Kushnir & Wallace 1989, Jones et al. 1997). The NAO is the only atmospheric mode that is robustly present in every month of the year. It is most pronounced in amplitude and areal coverage during winter, and it accounts for more than one third of the total variance in sea level pressure (Marshall & Kushnir 1997). During high (> +1) NAO winters the surface geostrophic zonal winds over Europe are over 8 m s⁻¹ stronger than during low (< −1) NAO winters, and, consequently, during high NAO winters the moderating influence of the ocean results in unusually warmer winter temperatures in Europe (Hurrell 1995).

With the aim of clarifying the possible ‘correlation channels’ linking climate variability and cod recruitment by means of 2 purely descriptive and model-independent statistical tools, principal component analysis (PCA) and canonical correlation analysis (CCA), we were able to detect the links between climatic and biological processes in the above-mentioned regional seas. Thus, an empirical and preliminary
model of the significant correlations between global- and local-scale physical forcing and cod populations starts to emerge (O’Brien et al. 2000).

MATERIALS AND METHODS

Data. Concerning the North Sea, the sea surface temperature has been derived from the GISST data bank V. 2.2 (http://www.badc.rl.ac.uk) in the form of year averages of monthly recorded data (on a 1° × 1° grid). Six randomly distributed locations over the whole North Sea region (53° N, 2° E; 55° N, 2° E; 56° N, 3° E; 56° N, 5° E; 57° N, 4° E; 57° N, 0° E; 55° N, 4° E) were taken and the corresponding time series averaged on a year basis. The very high pairwise correlation between the time series (average Pearson r = 0.97) insures that they convey practically identical information. Recruitment data are from the public-domain ICES data bank (courtesy of Dr Keith Brandt). Concerning the Barents Sea, we used data in the form of 3 yr running means for the 1943–1994 period of: (1) year-class strength of the Arcto-Norwegian cod and (2) sea-temperatures averaged over 0 to 200 m depth in the Kola section, along the 33°30’E meridian from 70°30’N to 72°30’N, as reported by Sundby (1998).

As for the NAO index, it was defined as the difference between normalized sea-level pressure anomalies during winter between Ponta Delgada in the Azores and Akureyri in Iceland (Lamb & Peppler 1987); we used the values reported by Sundby (1998).

All the time series (except temperatures) were analyzed in their standardized and nondimensional form, i.e. expressing in standard deviation units the differences from the mean.

Principal component analysis. Each time series was subjected to a time delay embedding and the resulting multivariate matrix analyzed by means of PCA. Time delay embedding is a very common method in dynamics (Broomhead & King 1986, Zbilut & Webber 1992) and consists of shifting the series of a constant lag (equal to 1 yr in our case) n times, thus generating a collection of n surrogate series corresponding to delayed copies of the original one (with delays going from 1 to n). The resultant n + 1 rank matrix, having as columns the original series plus the n surrogates, is then submitted to a PCA in order to extract the principal modes which determine the temporal variability of the series (Ghil & Vautard 1991).

The PCA method allows us to dissect out independent contributions (the extracted modes are orthogonal by construction) to the observed trajectory and is particularly suited for the study of a relatively short and noisy time series, being as it is not constrained by stationarity and/or harmonic constraints (Vautard et al. 1992). Figs. 2 to 4 report the first 3 components of each series under study, together with the reconstruction of the original ones, as linear combinations of the main components extracted from them. It is evident from the figures that the extracted components are both trend-like and quasi-harmonic (actually, they appear in the form of sine-cosine associated pairs of components) and allow an almost complete reconstruction of the original series.

It is worth noting that the components were extracted in order of explained variance, with the ‘signal’ components appearing first and the noisy part of the information relegated to the latter components (Broomhead & King 1986). This implies that focussing the attention on the first few components will emphasize the rule-obeying portion of the available information and reinforce the heuristic power of the analysis. The component scores are based on a correlation matrix and are expressed in standardized units.

Canonical correlation analysis. CCA is a technique for analyzing relationships between 2 sets of variables, in our case the recruitment-temperature, recruitment-NAO and NAO-temperature couples. Each set can contain several variables, and CCA calculates a linear combination from each set, called a canonical variable, such that the correlation between 2 canonical variables is maximized (Lebart et al. 1984, Rencher 1992). This correlation between 2 canonical variables is the first canonical correlation. CCA goes on by finding a second pair of canonical variables, independent of the first one, that produces the second highest correlation coefficient. The process of constructing canonical variables continues until the number of pairs of canonical variables equals the number of variables in the smaller of the 2 sets.

Each canonical variable is orthogonal to all the others of either set, except for the corresponding one in the other set. This permits the identification of canonical correlations as independent ‘correlation channels’ between the 2 sets, indicating distinct mechanisms through which the linkage between sets takes place (Rencher 1992).

Data analysis strategy. Information from fields of ‘predictors’ and ‘predictands’ are related to each other in the following way: First, the principal components (known as empirical orthogonal functions, or EOFs; Preisendorfer 1988) of the predictors and predictands are computed, so that most of the available information is concentrated into a smaller number of new, mutually independent dimensions (leading eigenmodes). Second, a CCA is carried out for the leading eigenmodes of the predictor-predictand couple in order to highlight the possible correlation channels linking the 2 data fields.
Each time series has been expressed here by means of its first 3 components, to keep the dimensionality of the system low and to avoid chance correlations due to overdetermination problems. The correlation between biological (cod recruitment) and physical (temperature, NAO index) data sets was established by computing the correlation between the corresponding component spaces. A very similar approach was adopted in a recent paper by Li & Kafatos (2000).

RESULTS

Principal component analysis

Due to the different length of the available series, the Kola and NAO series were submitted to an 8-dimensional embedding, while the North Sea series underwent a 6-dimensional embedding. Both the temperature and cod recruitment time series relative to the Barents Sea region (Fig. 2) were quite well reconstructed by the 3-component solution that together explained around 80% of the total variability for both temperature and cod recruitment.

The first temperature mode (t1) is a strongly anharmonic mode, while t2 and t3 constitute a quasi-harmonic sine/cosine pair with a circa-decennial period (Vautard et al. 1992, Li & Kafatos 2000). The first cod population mode is strongly anharmonic and is mainly driven by a sudden drop of cod population during the first few years of the 1970s, followed by a partial recovery. The second and third components have a strict antiphase relation of the sine/cosine type and probably derive from 2 associated oscillators of 5 and 10 yr. Even the North Sea series are well described (Fig. 3) by a 3-component solution of both the temperature and cod recruitment series (around 80% of variance explained). The first 2 temperature series (t1, t2) constitute a sine/cosine pair with periods of approximately 8 yr, while t3 is strongly anharmonic.

The first component of North Sea cod recruitment (cod1) is constituted by a strong trend-like mode which describes a decline in cod recruitment beginning in the late 1970s that followed the gadoid outburst of the 1960s (Daan et al. 1994). Cod2 is a relative high-frequency oscillation (ca 2 yr) of cod population. This quasi-biennial oscillation exists not only in the recruitment of North Sea cod, but also in haddock and western mackerel (Dippner 1997b). In contrast, cod3 has a very erratic profile.

The reconstruction of the NAO index by a 3-component solution (Fig. 4) is fairly good (70% of explained variability), if one takes into account the relatively long time span considered. The reconstructed NAO index appears as the result of the convolution of a ca 18 yr cycle (first component, n1) with 2 ca 8 to 9 yr cycles (second and third components, n2 and n3). The first cycle is consistent with the turnaround time of the sub-polar gyre in the North Atlantic, whereas the second cycle represents the travelling time of the Rossby wave in the North Atlantic (Hurrell 1995, Marshall & Kushnir 1997). Such longer cycles show a clear tendency towards increased amplitude during the last 30 yr, in accordance with reports concerning the increase of average temperatures in those areas. This 1990s warm period, in combination with the 1930s warm and the 1960s cold period, could be connected to the 65 to 70 yr oscillation in the global climate system (Schlesinger & Ramankutty 1994). A similar period of 80 yr has been detected by Appenzeller et al. (1998) in Greenland ice cores.

Canonical correlation analysis

Table 1 shows the result of the CCA carried out over the 3 main components of the studied time series. The existence of highly significant correlations between cod recruitment and temperature emerges in both the Barents Sea and the North Sea. In the Barents Sea region, the correlation takes place along 2 distinct and independent channels, as indicated by the 2 significant canonical correlations between cod recruitment and temperature. Only 1 significant correlation between cod recruitment and temperature scored in the North Sea. The NAO and temperature spaces are fully interconnected (very high and significant canonical correlations) in both areas, thus confirming, from an empirical data-driven point of view, the influence of this global-scale climate mode on local temperature dynamics (Dippner 1997a). Despite the strong link existing between NAO and temperature fields, cod recruitment and NAO spaces did not show any significant canonical correlation. This result shows that temperature is probably the climate signal that is ‘observable’ directly acting on (sensed by) cod recruitment and is transforming large-scale atmospheric variability (NAO) to a local scale.

The Pearson correlation coefficients between the principal components and the canonical variables, which allowed us to assign a ‘shape’ to the scored correlations, are reported in Table 2. It is important to note how the first canonical correlations of Barents Sea and North Sea point in opposite directions, thus implying a negative relation between temperature increase and cod recruitment in North Sea as compared to a positive effect of the increasing temperature in Barents Sea. This observation confirms, in a temporal dimension, previous results obtained from spatial variability studies (Dippner & Ottersen 2001) and points to the existence of an optimal temperature range for cod recruit-
ment in between the Barents Sea and North Sea temperature regimes.

As for Barents Sea, the first canonical correlation is driven by the first component of both temperature and cod recruitment (correlation coefficients = 0.97 and 0.98 with the respective canonical variables), indicating a generally linear ‘average’ effect exerted by temperature on cod recruitment. The second canonical correlation is driven by the second cod recruitment component (correlation coefficient = 0.97) and by the

Fig. 2. Temperature and cod recruitment dynamics in the Barents Sea. The 4 upper panels refer to temperature. t1–t3 correspond to the first 3 principal components of the time series; the fourth panel reports the reconstruction of the original temperature series (solid circles) by means of its principal components estimate (open circles). The 4 bottom panels refer to cod recruitment and have been drawn according to the same criteria. The fraction of total variability explained by cod1, cod2 and cod3 were, respectively, 33.6, 28.8 and 27.0 %. The corresponding values for t1, t2 and t3 were 42.8, 33.6 and 16.0 %, respectively.
third temperature component (correlation coefficient = 0.92). Looking at the graphs of these 2 modes in Fig. 2, it appears that such correlations point to more subtle and nonlinear effects of temperature variability on cod recruitment. As shown by Dippner & Ottersen (2001) in a similar investigation, no climate signal can be detected in total stock biomass, spawning stock biomass, catch, or survival index. Therefore, spawning stock biomass was not considered in our investigation, because the signals of climate variability and of fishing pressure cannot be separated from a single time series.

In the case of the North Sea, the canonical correlation between cod recruitment and temperature is driven by the first component on the side of biological

![Graphs of temperature and cod recruitment dynamics in the North Sea.](image)

Fig. 3. Temperature and cod recruitment dynamics in the North Sea. The various panels contain the identical type of information as the corresponding panels in Fig. 2. The fraction of the total variability explained by cod1, cod2 and cod3 were, respectively, 23.9, 19.9, and 14.7%. The corresponding values for t1, t2 and t3 were 33.9, 28.8 and 17.4%, respectively.
variables (correlation coefficient = -0.99) and by a linear positive combination of all 3 modes on the temperature side. A straightforward interpretation is that the first component of cod recruitment (cod1) mirrors the effect exerted by the temperature (as a whole) on cod recruitment, thus providing a quantitative estimate of the relative importance of temperature in the dynamics of cod recruitment, which we estimated to be between 14.1 and 23.9%. The former value (14.1%) is the product of the percentage of the variability explained by cod1 (23.9%, Fig. 3) and the square (0.59) of the canonical correlation (0.77, see Table 1). It corresponds to the fraction of the variance explained by cod1 that can be modeled by the temperature. The latter value (23.9%) holds if one assumes cod1 as totally driven by temperature and implies that the lack of perfect correlation between cod1 and temperature is only due to measurement errors. In any case, we can safely estimate the effect exerted by temperature on cod population as being between these 2 extremes.

Table 1. Canonical correlation coefficients, together with their significance values, relative to the pairs of canonical variables for all the between-fields comparisons. Temp = temperature in °C; Recr = recruitment and NAO = North Atlantic Oscillation. The statistically significant correlations are underlined; the significance value under the column headings indicates the global significance of the comparison provided by the most conservative among a panel of tests, namely Wilk’s lambda, Pillai’s trace, Otelling-Lawley and Royce max. root (suggested by Rao 1973) for the multivariate linear regression. As for the significance of single regressors, the estimate is based on Rao’s modified approximate F statistic (Rao 1973).

Table 2. Temperature versus cod recruitment: correlation coefficients between principal components and canonical variables which attained a statistical significance for the temperature-cod recruitment relationships. Large absolute values of correlation point to a dominant role exerted by the corresponding components in driving the correlations.
DISCUSSION

A major problem with looking for a rationale in ecological data is the design of plausible causal relationships between biological and environmental variables. Throughout the whole North Atlantic region, the amplitude of interannual-to-decadal-scale sea temperature fluctuations is very relevant and is accompanied by fluctuations in other important parameters such as wind, turbulence and light conditions, which altogether influence basic biological processes linked to plankton production and trophic transfer. Many interpretations of observable trends may be proposed at different levels of complexity, from those involving physiological considerations to those based upon trophic cascade models (Saetersdal & Loeng 1987, Ellertsen et al. 1989).

In the North Sea, interannual and interdecadal variability has been observed in hydrography as well as in biology. Three strong signals have been observed in long-term variability: the so-called Russel cycle (Russel 1973), the gadoid outburst (Cushing 1982) and the great salinity anomaly (GSA) (Dickson et al. 1988). The observed Russel cycle started in the mid-1920s: a system of herring and macrozooplankton in the western English Channel was replaced within a few years by one of pilchards and smaller zooplankton (Russel 1973). This and other profound ecosystem changes were reversed between 1965 and 1979, when more northwesterly winds prevailed again and the previous faunal composition returned (Cushing 1978, 1996). The reason for the gadoid outburst in the 1960s is still unclear and it may be linked to optimal environmental conditions combined with excellent growth rates and good year class strength (Daan et al. 1994). Between 1977 and 1978 the GSA, a huge body of water with extremely low salinity and very low temperature, entered the North Sea (Dickson et al. 1988); this had a major influence on cod recruitment dynamics. The ‘global magnitude’ of this effect is similar in the different areas, but the underlying causal routes are probably different, involving 2 independent mechanisms in the Barents Sea region and apparently only 1 main mechanism in the North Sea. A similar conclusion was reached for haddock and herring recruitment data collected in the North Sea (results not shown). The obtained results fit well to the results of Dippner (1997b) and Svendsen et al. (1995) for the North Sea, as well as to the results of Ottersen & Sundby (1995) and Dippner & Ottersen (2001) for the Barents Sea.

Concerning the Barents Sea, the influence on regional northern NE Atlantic Ocean climate of large-scale atmospheric forcing has already been suggested by Izhevskii (1964), who argued for what he called a monophasic form of fluctuation throughout the Gulf Stream system (i.e. no time lag in winter temperature fluctuations from the Florida current to the Barents Sea). More recent studies have demonstrated that much of the decadal variability in North Atlantic sea surface temperature can be explained as a local oceanic response to atmospheric variability (Deser & Blackman 1993, Battisti et al. 1995, Houghton 1996). Ådlandsvik & Loeng (1991) showed, for the period 1970–1986, that a substantial part of the variability in volume flux of Atlantic water to the Barents Sea was determined by atmospheric forcing. Changes in Barents Sea temperature have on some occasions been observed to occur simultaneously over a large area (Loeng et al. 1992). Statistical evidence for these relations has, however, been relatively weak so far, partly due to the lack of suitable time series. Our results show that there is indeed a very strong (canonical correlation = 0.90) and statistically significant link between the North Atlantic winter sea level pressure (SLP) field as indexed by the NAO and the Barents Sea temperature. This link is then ‘sensed’ by the cod recruitment through the temperature.

It is worth noting the differences between the correlation structure highlighted by our approach and the information derivable by the simple correlation coefficients between temperature and cod recruitment series. In both regions this correlation does not achieve statistical significance (Pearson r = 0.28 for Barents Sea and −0.27 for the North Sea) and impairs the discovery of the ecological role of temperature. The combined use of PCA and CCA, however, allowed us to both filter the original series from their noise components and explore the existence of complex links between biological and environmental variables. This proposed approach may provide important clues to discriminate among alternative models used to explain environment-biology links. It reveals, in fact, a number of complex and not immediately recognizable relationships among data fields, with a consequent net increase in sensitivity.

Our results indicate that sea temperatures exert a relevant influence on cod recruitment dynamics. The ‘global magnitude’ of this effect is similar in the different areas, but the underlying causal routes are probably different, involving 2 independent mechanisms in the Barents Sea region and apparently only 1 main mechanism in the North Sea. A similar conclusion was reached for haddock and herring recruitment data collected in the North Sea (results not shown). The obtained results fit well to the results of Dippner (1997b) and Svendsen et al. (1995) for the North Sea, as well as to the results of Ottersen & Sundby (1995) and Dippner & Ottersen (2001) for the Barents Sea.

Similar computations have been carried out for the Baltic Sea. However, apart from the strong correlation between NAO and sea surface temperature, in this investigation no meaningful correlation between temperature and cod recruitment could be found. One possible mechanism which could explain the long-term
variability of cod recruitment in the Baltic Sea is the climate-driven interannual variability in salinity (Zorita & Laine 2000), which has a strong influence on the vertical position of the floating cod eggs and therefore on the survival of the eggs and larvae. A further mechanism controlling interannual variability in Baltic Sea cod might be the size of the reproductive volume, a volume of water in which salinity, oxygen, and temperature conditions permit successful egg development (MacKenzie et al. 2000). Both interannual variability in salinity and variability in reproductive volume will be the subject of further investigations in understanding interannual variability in Baltic Sea cod.

Acknowledgements. We acknowledge support by the European Commission’s Environment and Climate Programme under contract number ENV4-CT97-0596 as part of the project CLICOFI, which contributes to the ELOISE thematic network. This publication has the ELOISE publication no. 218. J.W.D. was supported by the Alfred Wegener Institute in the framework of the CLICOFI project and by the mare Drei- viertel Verlag which is greatly acknowledged. Cod data for the North Sea were provided by K. Brander (ICES).

LITERATURE CITED

Izhevskii GK (1964) Forecasting of oceanological conditions and the reproduction of commercial fish. Pishcepromizdat, Moscow (translated by Israel Progr Sci Transl, Jerusalem)
Tereshchenko VV (1996) Seasonal and year-to-year variations of temperature and salinity along the Kola meridian transect. ICES CM 1996/C:11, p 24
Walker GT, Bliss EW (1932) World weather V. Mem Roy Soc 4: 53–84

Editorial responsibility: Otto Kinne (Editor), Oldendorf/Luhe, Germany

Submitted: August 18, 2000; Accepted: October 10, 2000
Proofs received from author(s): June 22, 2001