ICES Journal of Marine Science

ICES Journal of Marine Science (2017), doi:10.1093/icesjms/fsx179

Causal drivers of Barents Sea capelin (*Mallotus villosus*) population dynamics on different time scales

Hiroko K. Solvang[1]*, Sam Subbey[1], and Anna S. J. Frank[2]

[1]*Institute of Marine Research, PB-1870, N-5817 Bergen, Norway*
[2]*School of Pharmacy, University of Oslo, Oslo, Norway*

**Corresponding author: tel: +47 48 22 04 78; fax: +47 55 23 85 31; e-mail: hiroko.solvang@imr.no*

Solvang, H., Subbey, S., and Frank, A. S. J. 2017. Causal drivers of Barents Sea capelin (*Mallotus villosus*) population dynamics on different time scales. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsx179.

Received 18 September 2016; revised 14 July 2017; accepted 18 July 2017.

The dynamics of marine populations are usually forced by biotic and abiotic factors occurring at different intensity levels and time scales. Deriving the time frame within which each factor has a causal influence is important for predicting population trajectories. This paper presents a statistical methodology for establishing (i) the strength of causal coupling between population dynamics and environmental (biotic and abiotic) factors, and (ii) the time scales over which causal covariates have significant influence on the population dynamics. The methodology is based on combining a multivariate autoregressive model fit to data (to determine causal direction) with a quantification of the RPC of covariates in frequency domain (to quantify the strength of connection). The methodology is applied to test the existence of causal coupling between the capelin biomass and a selected number of covariates identified in the literature.

Keywords: Barents Sea, capelin, causality, climate, fisheries dynamics, MAR model, modelling, predator-prey, prediction, relative power contribution, time series, uncertainty.

Introduction

Fisheries time series data are usually highly variable because the observation data are forced by stochastic processes, which are characterized by time delays, jumps and spikes, and several other non-stationary mechanisms occurring on different time, frequency and intensity scales. According to Sundelöf *et al.* (2013), the mechanisms may be classified as being due to (i) environmental (including anthropogenic) forcing, (ii) species interactions, and (iii) internal processes (e.g. density dependent regulation of recruitment). The degree of variability observed in fisheries time series is dictated by the intensity and degree of alignment (both spatially and temporally) of these forcing mechanisms. For instance, shift in productivity has been postulated as an underlying mechanism for bursts in populations and stock recruitment (Munch and Kottas, 2009). But it is when conditions align spatio-temporally (e.g. temperature, match with prey, low cannibalism), that there is an appropriate response in recruitment, e.g. spiked recruitment (Solari *et al.*, 1997). Delayed or feedback effects resulting from strong population pulses (e.g. spiked recruitment) are manifested in the population dynamics several years afterwards (Skjoldal, 2004), particularly when they result in strong density-dependent responses, where recruitment of subsequent year classes (at the appropriate lag) is depressed (Caley *et al.*, 1996) or when they change the behaviour of the stock (Huse *et al.*, 2010).

For systems (e.g. marine ecosystems) consisting of multiple non-stationary processes, causal relationships among several components of the system will act with varying intensity, duration, and time scales. For instance, while temperature may affect fish behaviour in general, the egg and larval stages are known to be more sensitive to temperature change. The degree to which temperature change affects fish is therefore dependent on the timing (stage-dependent), intensity (level of variability), and duration (see e.g. Fuiman and Werner, 2009). Thus, while the primary

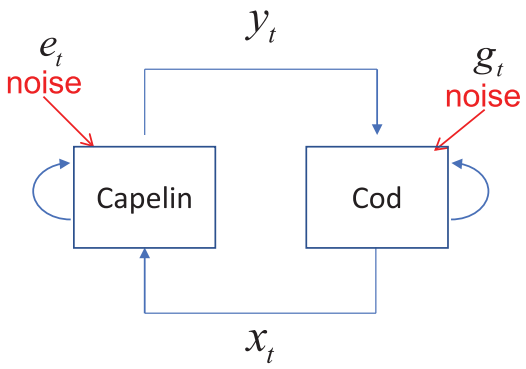

Figure 1. A simple (illustrative) feedback system involving capelin ($y(t)$) and cod ($x(t)$).

goal might be to understand feedback (causal) mechanisms, it is also noteworthy to quantify when such mechanisms are important, as well as their relative strength and duration.

For illustration, we consider the simplest case shown in Figure 1 (see Akaike and Nakagawa, 1989)—a system composed of two subsystems representing for instance the prey fish species capelin ($y(t)$) and its predator fish cod ($x(t)$). The dynamic of each subsystem is further perturbed by individual noises, $e(t)$ and $g(t)$, respectively for the capelin and cod subsystems.

Analysing the system in Figure 1 is particular challenging because $e(t)$ produces an effect on $x(t)$ through the capelin subsystem. Thus $x(t)$ and $e(t)$ are generally not uncorrelated (vice versa for $y(t)$ and $g(t)$) even if they are assumed to be statistically independent (Akaike and Nakagawa, 1989). Addressing this challenge, Akaike (1968) proposed a practical method to analyse such feedback systems by utilizing a multivariate autoregressive (MAR) model. Using a MAR representation, the noise and variable become independent, which is a precondition for spectral analysis using Fourier transformation (Akaike and Nakagawa, 1989). The power spectrum estimated by the AR coefficients and variance-covariance matrix can be expressed as a sum of the relative contributions from the individual variables. It is then possible to evaluate the degree of influence by individual variables by looking at each source contribution, and fluctuations over time, to the power spectrum. This is referred to as the Akaike's relative power contribution (RPC) (Akaike, 1968). The RPC shows the strength of causal relationship among multiple variables based on partitioning the power spectral density of an optimal autoregressive model. The RPC has been widely applied to many practical problems (see Akaike and Nakagawa, 1989; Akaike and Kitagawa, 2012), and has been referred to as *Akaike Causality* (Wong and Ozaki, 2007). Although the well-known causality concept by Granger (1969) applies only to bivariate systems, the RPC is broadly applicable to multivariate systems. It is worth mentioning that the original PRC idea assumes independence among the noises. In practice however, high correlations between noise components often occur. Hence Tanokura and Kitagawa (2004) proposed an extended power contribution approach in detecting the mutual influences involving cross-correlated noises of the variables.

The aim of this article is to demonstrate the use of the MAR model as a practical statistical tool for characterizing the causal coupling roles at various time scales, and of different candidate covariates in marine systems. A novelty of this paper lies in the combined use of the MAR modelling framework and the RPC concept, to establish causal links and causal directions between species driven by an environmental forcing (temperature).

Though some studies have applied the MAR model to ecological problems (see e.g. Ives, 1995; Hampton *et al.*, 2013), the approach adopted in this paper is absent in the ecological literature, and new to fisheries science. We present application examples were the main variable of interest is an index of species abundance (age-structured biomass), and the aim is to involve biotic and abiotic time series as candidate causal variables. The goal is to demonstrate how the methodology can be used to classify which covariates have most relevance for short, medium or long term prediction of the species dynamics. Such knowledge is important when devising management decisions on different time horizons. The article uses age-structured biomass data of Barents Sea capelin for the index, and discusses how the methodology may help improve our understanding of capelin stock dynamics.

Quantifying causality—the RPC

For the sake of simplicity, we first explain the methodology for quantifying causality using a 2D data example, and then extend the results to a more general, multi-dimensional case, which is relevant to this article. With reference to Figure 1, let the observed 2-dimensional time series be denoted by $(y(t), x(t))^{\mathrm{T}}$ for $t = 1, \ldots, N$, where $(\cdot)^{\mathrm{T}}$ denotes transposition. We assume that the data is generated by a multivariate auto-regressive process given by Equation (1):

$$\begin{cases} y(t) &= \sum_{m=1}^{M} a_{yy}(m)y(t-m) + \sum_{m=1}^{M} a_{yx}(m)x(t-m) + \epsilon_y(t), \\ x(t) &= \sum_{m=1}^{M} a_{xy}(m)y(t-m) + \sum_{m=1}^{M} a_{xx}(m)x(t-m) + \epsilon_x(t), \end{cases} \quad (1)$$

where a_{yy}, a_{yx}, a_{xy} and a_{xx} are the autoregressive (AR) coefficients, M is the AR order. The terms $\epsilon_y(t)$ and $\epsilon_x(t)$ are *i.i.d.* with mean zero and unknown variance, and result from whitening of the noise terms (in Figure 1) $e(t)$ and $g(t)$, respectively. The AR coefficients can be estimated by the ordinary least squares method or some other numerical algorithm such as the Yule-Walker method (Hamilton, 1994). The Akaike Information Criteria (AIC) (Akaike and Kitagawa, 2012) is used to determine the AR order of the best-fit model. Using this best-fit model, one obtains predictions for $x(t)$ and $y(t)$, as well as estimates of the errors $\epsilon_x(t)$ and $\epsilon_y(t)$, and associated variance-covariance matrix given by Equation (2),

$$\Sigma = \mathbb{E}[\varepsilon(t)\varepsilon(t)^{\mathrm{T}}] = \begin{pmatrix} \sigma_{yy}^2 & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx}^2 \end{pmatrix}, \quad (2)$$

where $\mathbb{E}$ is expectation, $\varepsilon(t) = [\epsilon_y(t), \epsilon_x(t)]^{\mathrm{T}}$, and the variances are represented by the diagonal elements of the matrix. The off-diagonal elements of Σ are the covariances, where $\sigma_{yx} = \sigma_{xy}$. A basic theoretical assumption of the methodology is that Σ has zero off-diagonal elements. In practical implementations (involving empirical and usually uncertain data), this requirement is considered fulfilled when the main diagonal elements of Σ are dominant. A quantitative approach for establishing this (dominant main diagonal) condition is to test for statistical significance of the off-diagonal elements (Σ_{ij}, $i \neq j$) using any statistical procedure for such tests, for instance, the Spearman's ρ (Best and Roberts, 1975) or the Kendall's τ statistic (Hollander *et al.*, 2013).

The physical interpretation of the estimated auto-regressive coefficients is obtained by considering the procedure in the frequency domain. The cross power spectra $\mathbf{P}(f)$ of the $(x(t), y(t))$ components generated by Equation (1) is given by Equation (3) (see Akaike and Nakagawa, 1989, Chapter 3).

$$\mathbf{P}(f) = \mathbf{A}(f)\Sigma\mathbf{A}^*(f),\ 0 \le f \le 0.5\Delta, \tag{3}$$

where Δ is the sampling interval, and $\mathbf{A}^*(f)$ is the complex conjugate of $\mathbf{A}(f)$, which is the frequency response defined by Equation (4):

$$\mathbf{A}(f) = \left\{\mathbf{I} - \sum_{m=1}^{M} \mathbf{A}(m)e^{-2\pi fm}\right\}^{-1}, \tag{4}$$

$$\mathbf{A}(m) = \begin{pmatrix} a_{yy}(m) & a_{yx}(m) \\ a_{xy}(m) & a_{xx}(m) \end{pmatrix}. \tag{5}$$

$\mathbf{I}$ is the identity matrix and $\mathbf{A}(m)$ is the AR coefficients matrix defined by Equation (5). The Akaike's RPC is defined by Equation (6), where A_{ji} is the ith row and jth column element of $\mathbf{A}(f)$.

$$r_{ij}(f) = \frac{|A_{ji}(f)|^2\sigma_{ii}^2}{|P_j(f)|} = \frac{|A_{ji}(f)|^2\sigma_{ii}^2}{\sum_{i=1}^{k}|A_{ji}(f)|^2\sigma_{jj}^2} \in [0,\ 1];\ i, j = 1, 2, \ldots, k. \tag{6}$$

It must be cautioned that the RPC does not carry the same statistical connotation as correlations or cross-power spectra. The value of $r_{ij}(f)$ quantifies the percentage contribution from other variables to the power spectrum of the target variable. For a given target, when the various RPCs are graphically represented in the frequency domain, the pattern of the contribution of the noise sources to the system behaviour becomes clear (Akaike and Nakagawa, 1989).

Illustrative example

We present a simple illustrative example for the feedback relationship between the biomasses for capelin and cod (see Figure 1). We assume the estimation process yielded the MAR Equation (2) model defined by Equation (7):

$$\begin{cases} y(t) &= 1.00y(t-1) - 0.37y(t-2) - 0.48x(t-1) \\ &\quad +0.54x(t-2) + \epsilon_y(t), \\ x(t) &= 0.19y(t-1) - 0.19y(t-2) \\ &\quad +1.5x(t-1) - 0.63x(t-2) + \epsilon_x(t), \end{cases} \tag{7}$$

and the variance-covariance matrix in Equation (8):

$$\Sigma = \begin{pmatrix} 0.350 & 0.001 \\ 0.001 & 0.053 \end{pmatrix}. \tag{8}$$

Using (7), we define the elements of $\mathbf{A}(m)$ in (5),

$$\mathbf{A}(1) = \begin{pmatrix} 1.00 & -0.48 \\ 0.19 & 1.50 \end{pmatrix},\quad \mathbf{A}(2) = \begin{pmatrix} -0.37 & 0.54 \\ -0.19 & -0.63 \end{pmatrix}, \tag{9}$$

where $\mathbf{A}(1)$ is the coefficient matrix for $(t-1)$ terms in Equation (7); correspondingly for $\mathbf{A}(2)$. Using the above estimates, we calculate the RPC with $k = 2$ in Equation (6), based on Equations (3–4). Figure 2 is a graphical illustration of the calculated RPC, where the x- and y-axes represent the frequency domain and the RPC ratio, respectively. The frequency scales have been converted to annual cycles, to reflect the actual sampling interval of the data. Table 1 defines notations for Figure 2, that are consistent with Equation (6).

We note that the power contribution from cod to capelin (Figure 2a) is largest at a cycle slightly longer than 16 years, while the power contribution from capelin to cod (Figure 2b), is significantly largest at around 8 years. In this particular case, the power contribution from capelin to cod is in total larger than the power contribution from cod to capelin. In other words, the (signal) driver of the capelin biomass dynamics has a more regulative effect on cod, than the effect of the cod biomass driver on capelin. Another representation of the RPC is the use of heatmaps (Figure 2c and d). The application of 2D heatmaps are especially attractive when dealing with multi-dimensional data since they are intuitive, and ease making inference on the RPC of all data sources. Heatmaps will be used to illustrate the RPC for the particular application considered in this paper, as it involves five different time series datasets.

Application to case study—Barents Sea capelin

Capelin in the Barents Sea is a short-lived (1–4 years) pelagic species, that is considered to be the most important pelagic fish stock in the Barents Sea (Gjøsæter and Ushakov, 2003). It is the main diet for Northeast Arctic cod (Bogstad and Mehl, 1997; Gjøsæter *et al.*, 2009) and juvenile herring (Gjøsæter and Bogstad, 1998; Hallfredsson and Pedersen, 2009). Several marine mammals (e.g. harp seals, humpback whales, minke whales), seabirds, kittiwakes and guillemots are also known to prey on capelin. Capelin recruitment is thought to be mainly regulated by the degree of juvenile herring predation on capelin larvae (Carscadden *et al.*, 2013; Gjøsæter *et al.*, 2015) and predation by Northeast Arctic cod (Gjøsæter *et al.*, 2015). Both biotic (food supply—copepods, euphausiids, and hyperiids) and abiotic (ambient temperature) have been reported to affect capelin feeding, condition factor and distribution (see Orlova *et al.*, 2004). Drastic changes in stock size have occurred in the last three decades, with three stock collapses in 1985–1989, 1993–1997, and 2003–2006 (Gjøsæter *et al.*, 2009). It has however, been difficult to unravel the causes of these variations. Though the literature contains a number of possible explanations, they fail to explicitly explain the observed capelin dynamics over the years. For instance, capelin is known to overlap spatially with cod and herring at different stages of its life history (Huse and Gjøsæter, 1999) with young herring (mainly age groups 1 and 2) preying on juvenile capelin. Gjøsæter and Bogstad (1998) therefore suggest that the abundance of herring leads to recruitment failure and eventual collapse. Hjermann *et al.* (2009) argue that high abundance of young herring is a necessary but not sufficient condition for recruitment failure of the capelin stock. With annual capelin consumption in the order of several thousand tonnes, cod (mainly 3–6 year of age) predation

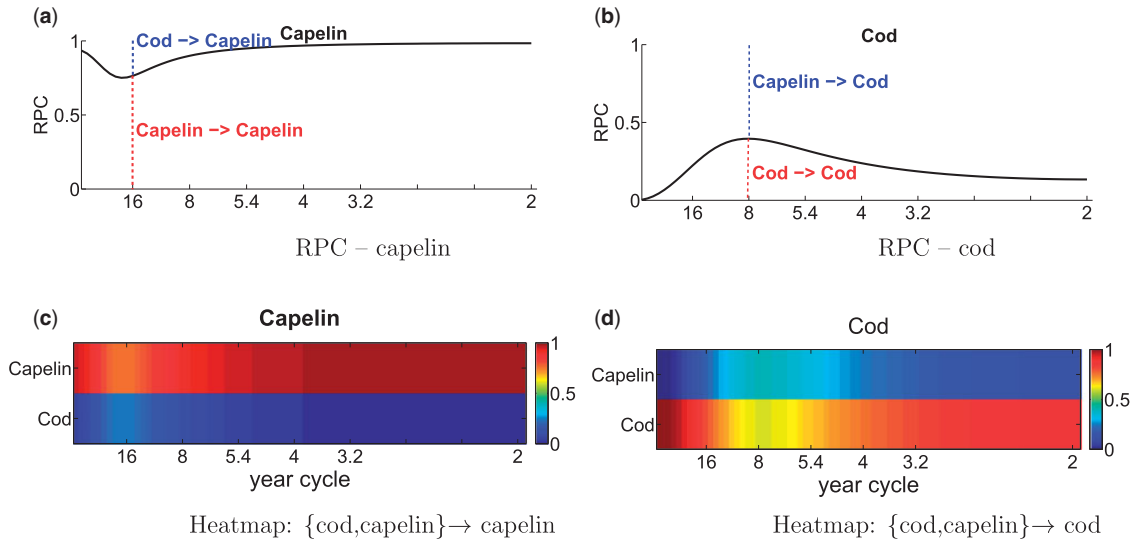

Figure 2. The RPC and heatmap representations for the example capelin-cod system. In figure (a), the lengths of the blue and red stippled lines are the RPC values for cod →capelin (RPC ≈ 0.2) and capelin →capelin (RPC ≈ 0.8), respectively, at a 16-year cycle. An analogous interpretation applies to figure (b), at an 8-year cycle. Observe the correspondence between the max/min points in the RPC curves (a and b) and the change in colour intensity at corresponding year cycle in the heatmaps (c and d). The heatmaps represent the contribution from {cod, capelin} to capelin (c), and from {cod,capelin} to cod (d).

Table 1. Consistency between Figure 2 and $r_{ij}(f)$ notation in Equation (6).

i	J		$r_{ij}(f)$	
1	1	capelin	→	capelin
1	2	cod	→	capelin
2	1	capelin	→	cod
2	2	cod	→	cod

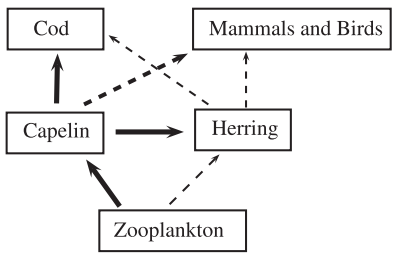

Figure 3. A simplified foodweb (biomass flow) representation of the Barents Sea ecosystem, redrawn after Hjermann *et al.* (2004). The thickness of each arrow (from prey to predator) indicates the perceived importance of the pathway. This manuscript focuses only on pathways involving fully drawn arrows. The zooplankton consists mainly of copepods, krill and amphipods.

on capelin has been suggested as another potential cause of capelin stock collapse (Dolgov, 2002).

Figure 3 (redrawn after Hjermann *et al.*, 2004) represents a simplified foodweb of the Barents Sea, showing capelin (focal species) and its link to both lower and higher trophic level species. Given its central role, the effects of capelin collapse have been registered both downwards and upwards in the Barents Sea food web (Gjøsæter *et al.*, 2009).

The data

Based on Figure 3, we define the biotic dataset by the annual biomasses of capelin of ages 1–4, the total annual biomass of cod and herring, and the krill biomass density in the Barents Sea. The average August temperature taken from the Bird Island–Bear Island section was used as abiotic driver in the analysis. This temperature series has been reported to be correlated with the temperature of inflowing Atlantic waters into the Barents Sea, which in turn, influences the distribution of pre-juvenile (0-group) capelin, Gundersen (1993). The data are taken from the database of the Working Group on the Integrated Assessments of the Barents Sea (see e.g. ICES, 2016). Figure 4 shows all variables used in the analysis.

For the observations shown in Figure 4, we derive the 5D time series vector $\boldsymbol{D} \equiv$ (capelin*, cod, krill, herring, temperature) where capelin* represents age-dependent capelin biomass at either age 1–4. We use capelin biomass data from 1972 to 2015, and associated covariate data within the same time range.

Of the data ensemble used, temperature is the only data that can be considered as raw, with measurement precision being ∼10% of a degree Celcius. The biotic data is either processed (upscaled or averaged) survey information (e.g. krill density) or estimates derived from stock assessment models (biomass of capelin, herring, and cod). Unfortunately, the abiotic data comes with no measure of precision. Figure 4d shows a spike in the herring biomass after 1982. The spike (in 1983) corresponds to an extraordinary year class of the stock after a long collapse period, where the spawning-stock biomass went from almost zero (between 1970 and 1982), to about 0.5 million tonnes in 1983 (see Toresen and Østvedt, 2000; Røttingen and Tjelmeland, 2003).

Observation data for capelin of age 4+ (see Figure 4) is usually sparse and unreliable because Barents Sea capelin usually spawns at 3 years, and then dies (Gjøsæter *et al.*, 2002). We shall therefore restrict the analysis and discussion of simulation results to cover

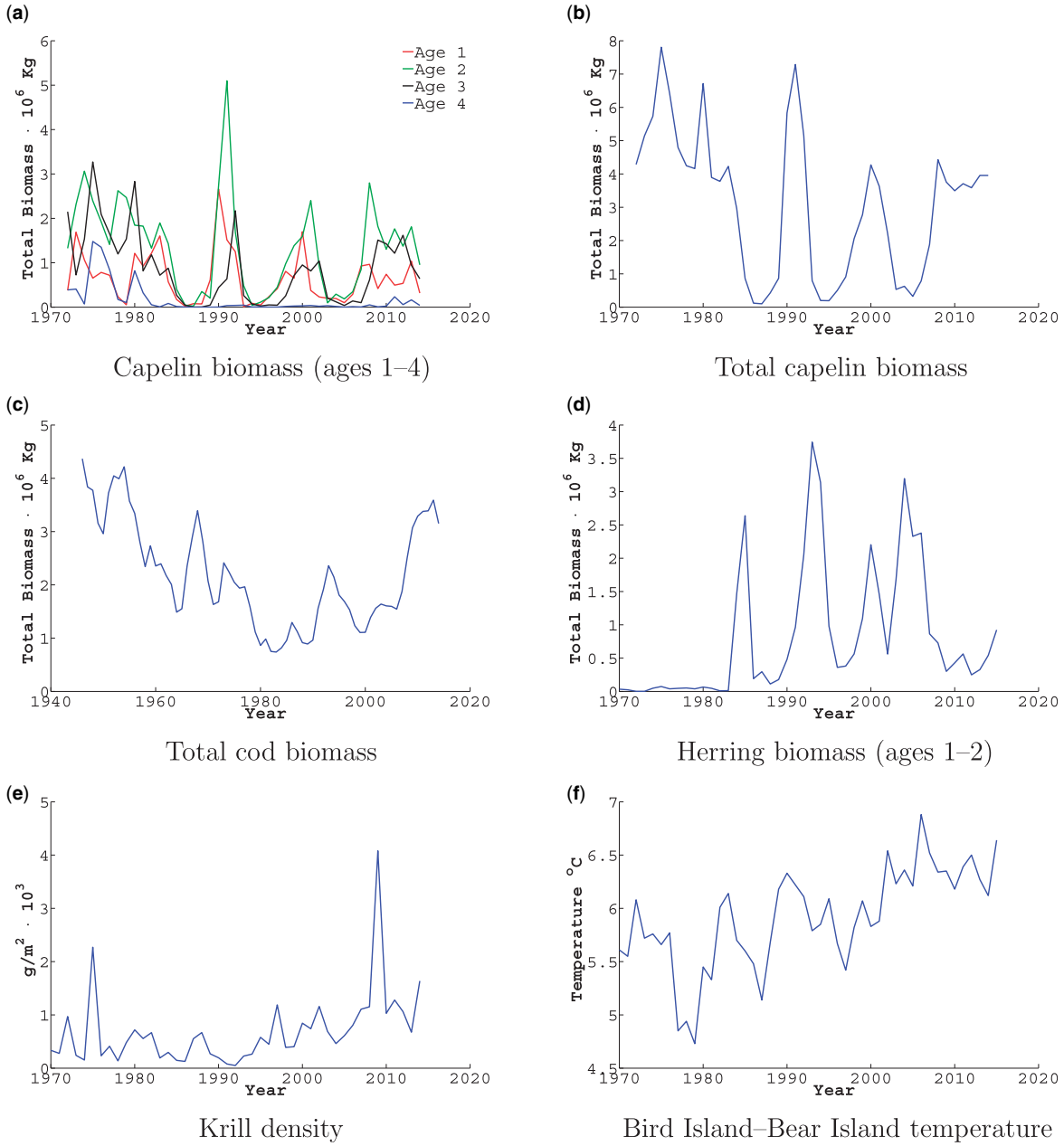

Figure 4. The biotic and abiotic data.

capelin in the age range 1–3 only, even though age-4 results will be shown for the sake of completeness.

Results and discussion

The application of the MAR model in this article is justifiable since there were no strong nonlinearities in the data. Had such nonlinearities existed however, data transformation (through e.g. log-transformation) would have been required, in addition to window shifting so that the MAR model applies to a fixed data length (see e.g. Francis *et al.*, 2014).

We applied a MAR model to $\boldsymbol{D}$, where the model coefficients were calculated using the Burg algorithm (see e.g. Schlögl, 2006), and the best-fit model was selected within AR order 1–5. Based

Table 2. Results for AIC and model order for biomass groups.

Autoregression order	Biomass age group			
	1	2	3	4
0	146.36	157.66	161.04	147.31
1	0.00	3.87	6.50	0.00
2	18.56	0.00	0.00	4.85
3	19.74	13.80	10.56	18.87
4	36.20	17.21	28.70	39.01
5	29.59	19.09	49.84	58.34

For each column under the biomass groups, the values show deviations from the minimum AIC values for that particular column.

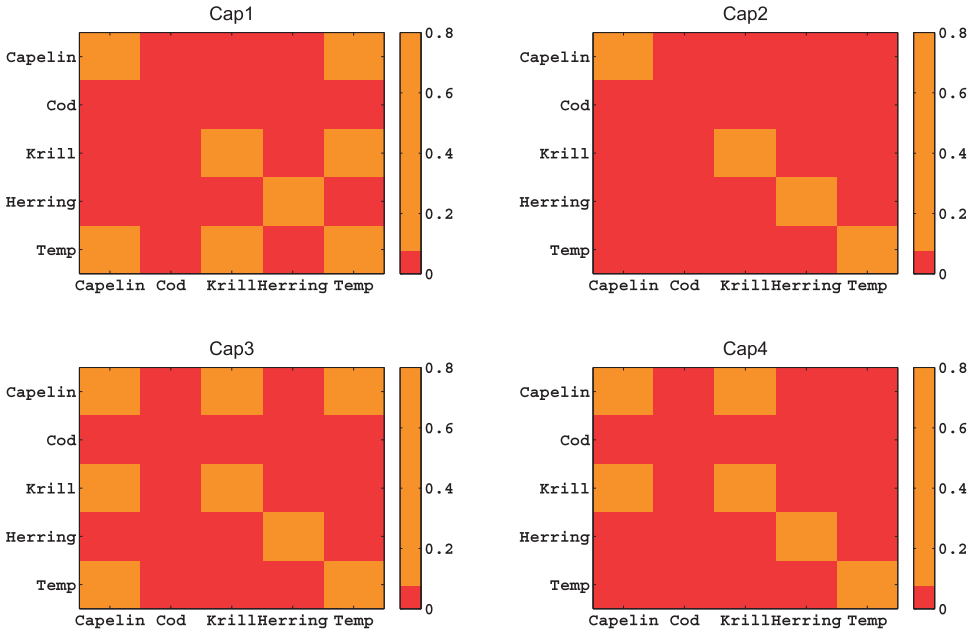

Figure 5. Variance-covariance matrices of covariates in time domain.

on the estimated AR coefficients, the prediction was obtained by Equation (1), and the variance-covariance matrix Σ was obtained using the prediction error. It must be mentioned that the (full) general MAR model, where each time series variable $(x(t), y(t))$ has feedback to each other time series variable, was applied on $\boldsymbol{D}$. However, it is inconceivable that fish will have direct effect on temperature. There are two options in addressing this particular case:

(1) define the MAR model such that coefficients of terms quantifying fish effect on temperature (Fish $\Rightarrow$ Temperature) is set to zero, or

(2) allow the framework (including AIC-base model choice) to determine the best model.

The latter option (unconstrained coefficients) allows us to test the robustness of our framework, and in particular, its ability to work in situations where there are no clear guidelines about improbable causal flow directions. We have opted for this option, and based on the AIC, evaluated the robustness of our numerical framework.

A table of varying model orders and AIC differences is presented in Table 2.

The AR order corresponding to zero AIC value is indicative of the best fit model. The results from this table establish the model order. Next, we adopt option 1 above, in determining the structure of the causal flow model. Figure 5 are the generated heatmaps of the variance–covariance matrices.

We follow the proposed statistical procedure to determine whether the autocovariance (main diagonal) components of the variance-covariance are dominant. Accordingly, we calculated the correlation (Spearman's ρ) matrix of $\epsilon(t)$, and associated p-values (see Table 3). We evaluate the statistical significance of the elements in the correlation matrix at a significance level $\alpha = 5\%$. The results in Table 3 show that with exception of $\boldsymbol{P}_{1,5}$ for age-1 capelin (p-value $= 0.024$), $\boldsymbol{P}_{1,2}$ for age-2 capelin (p-value $= 0.013$) and $\boldsymbol{P}_{1,3}$ for age-3 capelin (p-value $= 0.004$), we can assume the off-diagonal elements to be insignificant. Since the number of samples is <50, we also applied Kendall's τ statistic. We obtined results that were consistent with those obtained using the Spearmans ρ. Given that each age-group only one of the off-diagonal elements failed our significance test, we can assume (within the constraints of uncertainty) that the fundamental assumption of the methodology is satisfied. On the other hand, for cases where several off-diagonal elements appear to be significant, one may apply a more complicate model, such as the extended power contribution model by Tanokura and Kitagawa (2004). Further discussion of the results will concentrate on the RPC heatmaps.

Figure 6 shows heatmaps of the RPC of capelin (auto-contribution), cod, krill, herring, and temperature. The y-axis lists the contributing variables, while the x-axis indicate annual cycles. The strength of a contribution is linked to the heatmap color intensity, defined by the colorbar. We first present point-wise summary of the results, where the emphasis is on the RPC to the capelin biomass from all other sources (i.e. {cod,krill,herring,temperature}$\rightarrow$ Capelin) in Figure 6. Next, we discuss the model results in light of existing knowledge from the literature.

Summary of results: RPCs to capelin biomass

The analysis shows strong auto-contribution of capelin to its own dynamics, with multiple periodicities lasting from between 2 and 5 years. This means the effect of a strong capelin year-class will be expected to last for up to ~5-years. This observation applies reliably to capelin of ages 1–3. The direct influence of cod on age-1 capelin dynamics appears to be insignificant. The cod dynamics appears to have most influence on ages 2 and 3 capelin, but over different time scales. For age-2 capelin, the cod influence is registered after 3 years, while the effect on age-3 capelin is at least, 9 years. Our analysis (see the heatmaps) indicates a generally weak link between krill and capelin biomass dynamics. The herring power contribution to capelin dynamics is most significant on age-1 capelin, with a 2–5 year cycle, and this influence can last

Table 3. *P*-value matrix ***P***, and associated Spearmaned ρ statistic matrix for capelin.

Capelin age	***P* (*p*-value matrix)**	**Spearmanatrρ matrix**
1	$\begin{bmatrix} - & 0.848 & 0.868 & 0.841 & 0.024 \\ & - & 0.481 & 0.800 & 0.768 \\ & & - & 0.919 & 0.207 \\ & & & - & 0.716 \\ & & & & - \end{bmatrix}$	$\begin{bmatrix} 1 & -0.164 & -0.178 & -0.159 & 0.313 \\ & 1 & 0.008 & -0.134 & -0.117 \\ & & 1 & -0.222 & 0.130 \\ & & & 1 & -0.091 \\ & & & & 1 \end{bmatrix}$
2	$\begin{bmatrix} - & 0.013 & 0.751 & 0.975 & 0.582 \\ & - & 0.259 & 0.984 & 0.712 \\ & & - & 0.729 & 0.158 \\ & & & - & 0.473 \\ & & & & - \end{bmatrix}$	$\begin{bmatrix} 1 & 0.352 & -0.110 & -0.314 & -0.033 \\ & 1 & 0.105 & -0.341 & -0.091 \\ & & 1 & -0.099 & 0.162 \\ & & & 1 & 0.011 \\ & & & & 1 \end{bmatrix}$
3	$\begin{bmatrix} - & 0.084 & 0.004 & 0.923 & 0.155 \\ & - & 0.224 & 0.997 & 0.470 \\ & & - & 0.903 & 0.409 \\ & & & - & 0.804 \\ & & & & - \end{bmatrix}$	$\begin{bmatrix} 1 & 0.222 & 0.411 & -0.230 & 0.165 \\ & 1 & 0.123 & -0.436 & 0.012 \\ & & 1 & -0.209 & 0.038 \\ & & & 1 & -0.138 \\ & & & & 1 \end{bmatrix}$
4	$\begin{bmatrix} - & 0.280 & 0.131 & 0.342 & 0.611 \\ & - & 0.519 & 0.717 & 0.757 \\ & & - & 0.929 & 0.424 \\ & & & - & 0.840 \\ & & & & - \end{bmatrix}$	$\begin{bmatrix} 1 & 0.093 & 0.179 & 0.065 & -0.045 \\ & 1 & -0.007 & -0.092 & -0.111 \\ & & 1 & -0.233 & 0.031 \\ & & & 1 & -0.158 \\ & & & & 1 \end{bmatrix}$

In the *p*-value matrix, we use the notation '–' to represent *p*-values $< 2.2e\text{-}16$.

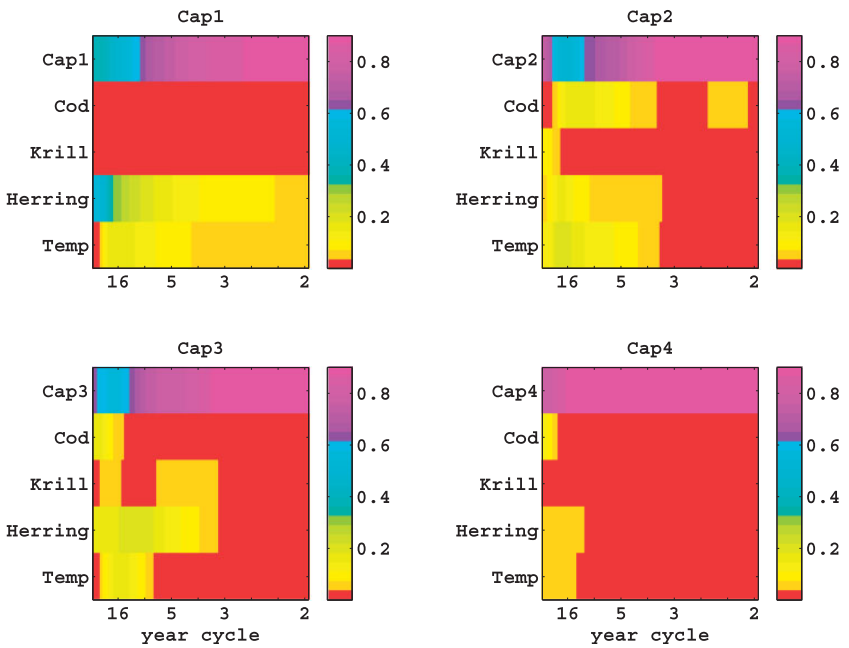

Figure 6. Heatmaps of the RPC in frequency domain for capelin. In general Cap(a) refers to capelin of age a. Each row corresponds to one RPC, $r_{aj}(f)$, from covariate j on the vertical axis to Cap(a).

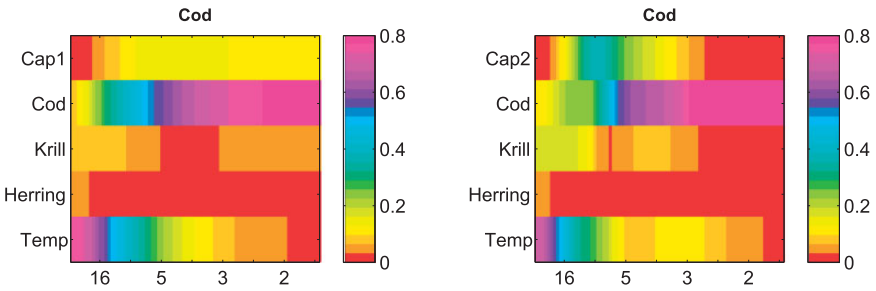

Figure 7. Heatmaps of the RPC in frequency domain for cod.

over several overlapping cohorts, resulting in a strong 16-year periodicity influence (60%), see Figure 6—cap1. The effect of temperature seem to be strongest on age-1 capelin, and decreases in influence with increasing capelin age.

Discussion

According to the heatmaps, the dynamics of the capelin biomass appears to be predominantly autonomous (RPCs of over 80%—see the strength of the capelin → capelin RPC for all ages in Figure 6), with a cycle of at least 5 years. This periodicity cycle also coincides with observed stock fluctuations, with capelin collapse being reported during the periods 1985–1989, 1993–1997, 2003–2007 (Gjøsæter *et al.*, 2015). Yndestad and Stene (2002) reported fluctuations of the capelin stock—which they referred to as stochastic resonance, with a 6.2 year cycle—as a natural environmental adaptation, and optimal strategy for growth and survival. Our results on the autonomous dynamics of the capelin biomass and its periodicity are therefore consistent with the literature, and with empirical observations.

Our results show that of all the age groups, age-2 capelin appears to be most affected by cod predation, occurring at cycles between 2 and 3 years, and more than 3 years. This is consistent with (Hjermann *et al.*, 2004), where age-2 capelin is reported to be most affected, especially during years of high cod predation. Hamre (2002) reported (a) temperature cycles of 8–15 years in the Barents Sea, coincident with strong year classes of herring and cod, and that (b) the abundance of immature cod determines the mortality of maturing capelin, i.e. capelin at around age-3. Put together, we should expect immature cod to influence the dynamics of age-3 capelin at the same frequency as the occurrence of the strong age classes of cod, i.e. between 8 and 15 years, which is consistent with the results from our analysis. With regards to herring, the strong effect on young capelin (age 1) is supported by the fact that herring preys principally on capelin larvae, with the presence of young herring (1–2 years old) being associated with low capelin recruitment; even close to zero recruitment in some years (Huse and Toresen, 2000; Hjermann *et al.*, 2004). The direct effect of herring on ages 2 and 3 capelin are either weak or insignificant. This result fits well with the herring ecology as they feed on younger stages of capelin.

Krill is known to be most important for planktivorous capelin of age 2–3 years old, and the lack of relationship for age 1 is supported by Dalpadado and Skjoldal (1996). The literature shows that a strong reduction in stock size of older capelin (between 1984 and 1987) was followed by an increase in biomass of the two main krill species in the Barents Sea (*Thysanoessa inermis* and *Thysanoessa longicaudata*), and a decrease in abundance and biomass of krill could be linked to the rapid growth of the capelin stock up to 1991 (Dalpadado and Skjoldal, 1996). The general weak effect for all ages, can be explained by the fact that the capelin-krill interaction is has a stronger top-down, rather than a bottom-up effect. This inference is strongly supported by the literature (see e.g. Skjoldal and Rey, 1989; Dalpadado and Skjoldal, 1996b; Dalpadado *et al.*, 2001; Baum and Worm, 2009).

In general, the literature reports a weak, direct climatic influence on the dynamics of capelin (Ozhigin and Luka, 1985). The apparently stronger temperature effect on age-1 capelin may be explained by the fact that the temperature series used has been reported to correlate with the temperature of inflowing Atlantic waters into the Barents Sea. These waters in turn, influence the distribution and survival of pre-juvenile (0-group) capelin (Gundersen, 1993), and subsequently, the biomass of age-1 capelin.

Ability to detect causal direction

In general, most models for inferring causality between an observed pair of observations are based on the assumption that one of the observation sets in the causal-effect pair is measured accurately, see e.g. Janzing *et al.* (2012). Observations with measurement errors for both the input and output of a natural system are common, and complicate the task of determining a causal direction (Zhang and Luo, 2014). Figure 7 demonstrates the ability of the methodology to deduce directionally dependent RPCs (distinguishing between $A \rightarrow B$ and $B \rightarrow A$). This characteristic translates into the ability of the methodology to quantify multi-directional causal links. As an illustration, we found no contribution from cod to the dynamics of age-1 capelin in Figure 6. However, Figure 7 shows that age-1 capelin has significant RPC to the total cod biomass dynamics in 2–5 year cycle. Furthermore, for age-2 capelin, a comparison of the heatmaps in Figure 7 and Figure 6 shows that a clear distinction between the dynamics capelin → cod, and cod → capelin can be made.

Conclusions

We have presented a statistical method for analysing the feedback relationship among system observations, and applied it to investigate how the biomass dynamics of Barents Sea capelin is mediated by other biological species and temperature. In fisheries science, it is usual that the time series data available are short (see Schnute, 2004) and thus prone to aliasing. Our analysis is incapable of identifying mediating effects on a time scale shorter than two years. This limitation is linked to the Nyquist-Shannon sampling theorem (Shannon, 1949; Jerri, 1977), which states that *the sampling frequency should be at least twice the highest frequency contained in the signal.* The goal of the Nyquist-Shannon theorem

is to avoid aliasing, i.e. when a signal is discretely sampled at a rate that is insufficient to capture the changes in the signal.

Funding

This research has received funding from the following IMR Research Programmes—MarPro-PROVEN (Project no. 14412), the Barents Sea Programme (Management Strategies for the Barents Sea, Project no. 84126) and Reduced Uncertainty in Stock Assessments (REDUS, Project no. 14809-01).

References

Akaike, H. 1968. On the use of a linear model for the identification of feedback systems. Annals of the Institute of Statistical Mathematics, 20: 425–439.

Akaike, H., and Kitagawa, G. 2012. The Practice of Time Series Analysis. Springer Science & Business Media, New York.

Akaike, H., and Nakagawa, T. 1989. Statistical Analysis and control of Dynamic Systems, Vol. 2. Kluwer Academic Publishers, Netherlands.

Baum, J. K., and Worm, B. 2009. Cascading top-down effects of changing oceanic predator abundances. Journal of Animal Ecology, 78: 699–714.

Best, D., and Roberts, D. 1975. Algorithm as 89: the upper tail probabilities of spearman's rho. Journal of the Royal Statistical Society. Series C (Applied Statistics), 24: 377–379.

Bogstad, B., and Mehl, S. 1997. Interactions between cod (gadus morhua) and its prey species in the barents sea. forage fishes in marine ecosystems. *In* Proceedings of the International Symposium on the Role of Forage Fishes in Marine Ecosystems, no. 97-01 in Alaska Sea Grant College Program Report, pp. 591–615. University of Alaska Fairbanks.

Caley, M., Carr, M., Hixon, M., Hughes, T., Jones, G., and Menge, B. 1996. Recruitment and the local dynamics of open marine populations. Annual Review of Ecology and Systematics, 27: 477–500.

Carscadden, J. E., Gjøsæter, H., and Vilhjálmsson, H. 2013. Recruitment in the Barents Sea, Icelandic, and eastern Newfoundland/Labrador capelin (Mallotus villosus) stocks. Progress in Oceanography, 114: 84–96.

Dalpadado, P., Borkner, N., Bogstad, B., and Mehl, S. 2001. Distribution of themisto (amphipoda) spp. in the barents sea and predator-prey interactions. ICES Journal of Marine Science: Journal du Conseil, 58: 876–895.

Dalpadado, P., and Skjoldal, H. J. 1996. Abundance, maturity and growth of the krill species *Thysanoessa inermis* and *T. longicaudata* in the Barents Sea. Marine Ecology Progress Series, 144: 175–183.

Dolgov, A. 2002. The role of capelin (Mallotus villosus) in the food-web of the Barents Sea. ICES Journal of Marine Science, 59: 1034–1045.

Francis, T. B., Wolkovich, E. M., Scheuerell, M. D., Katz, S. L., Holmes, E. E., and Hampton, S. E. 2014. Shifting regimes and changing interactions in the lake washington, usa, plankton community from 1962–1994. PloS One, 9: e110363.

Fuiman, L. A., and Werner, R. G. 2009. Fishery Science: The Unique Contributions of Early Life Stages. John Wiley & Sons, UK.

Gjøsæter, H., and Bogstad, B. 1998. Effects of the presence of herring (*Clupea harengus*) on the stock-recruitment relationship of Barents Sea capelin (*Mallotus villosus*). Fisheries Research, 38: 57–71.

Gjøsæter, H., Bogstad, B., and Tjelmeland, S. 2002. Assessment methodology for Barents Sea capelin, Mallotus villosus (Müller). ICES Journal of Marine Science, 59: 1086–1095.

Gjøsæter, H., Bogstad, B., and Tjelmeland, S. 2009. Ecosystem effects of the three capelin stock collapses in the barents sea. Marine Biology Research, 5: 40–53.

Gjøsæter, H., Hallfredsson, E. H., Mikkelsen, N., Bogstad, B., and Pedersen, T. 2015. Predation on early life stages is decisive for year-class strength in the Barents Sea capelin (Mallotus villosus) stock. ICES Journal of Marine Science, 73: 182–195.

Gjøsæter, H., and Ushakov, N. G. 2003. Capelin in the Barents Sea. *In* Proceedings of the 10th Norwegian–Russian Symposium. Ed. by Å. Bjordal, H. Gjøsæter, and S. Mehl, pp. 8–17. Bergen, Norway.

Granger, C. W. 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society, 37: 424–438.

Gundersen, A. C. 1993. Distribution of larval and 0-group capelin (Mallotus villosus) in the Barents Sea in relation to environmental factors, 1981–91. ICES. C.M. H: 15. 1–20.

Hallfredsson, E. H., and Pedersen, T. 2009. Effects of predation from juvenile herring (clupea harengus) on mortality rates of capelin (mallotus villosus) larvae. Canadian Journal of Fisheries and Aquatic Sciences, 66: 1693–1706.

Hamilton, J. D. 1994. Time Series Analysis, Vol. 2. Princeton University Press, Princeton.

Hampton, S. E., Holmes, E. E., Scheef, L. P., Scheuerell, M. D., Katz, S. L., Pendleton, D. E., and Ward, E. J. 2013. Quantifying effects of abiotic and biotic drivers on community dynamics with multivariate autoregressive (mar) models. Ecology, 94: 2663–2669.

Hamre, J. 2002. Capelin and herring as key species for the yield of cod: Results from multispecies model runs. Science Marine, 67(Suppl. 1): 315–323.

Hjermann, D. Ø., Bogstad, B., Dingsør, G. E., Gjøsæter, H., Ottersen, G., Eikeset, A. M., and Stenseth, N. C. 2009. Trophic interactions affecting a key ecosystem component: a multi-stage analysis of the recruitment of the Barents Sea capelin. ICES. CM documents; 2009/ C: 07.

Hjermann, D. Ø., Stenseth, N. C., and Ottersen, G. 2004. Indirect climatic forcing of the Barents Sea capelin: a cohort effect. Marine Ecology Progress Series, 273: 229–238.

Hollander, M., Wolfe, D. A., and Chicken, E. 2013. Nonparametric Statistical Methods. John Wiley & Sons, New Jersey.

Huse, G., Fernö, A., and Holst, J. C. 2010. Establishment of new wintering areas in herring co-occurs with peaks in the 'first time/repeat spawner' ratio. Marine Ecology Progress Series, 409: 189–198.

Huse, G., and Gjøsæter, H. 1999. A neural network approach for predicting stock abundance of the Barents Sea capelin. Sarsia, 84: 457–464.

Huse, G., and Toresen, R. 2000. Juvenile herring prey on Barents Sea capelin larvae. Sarsia, 85: 385–391.

ICES. 2016. Working Group on the Integrated Assessments of the Barents Sea (WGIBAR). ICES CM 2016/SSGIEA:04, ICES, Murmansk, Russia.

Ives, A. R. 1995. Predicting the response of populations to environmental change. Ecology, 76: 926–941.

Janzing, D., Mooij, J., Zhang, K., Lemeire, J., Zscheischler, J., Daniušis, P., Steudel, B. *et al.* 2012. Information-geometric approach to inferring causal directions. Artificial Intelligence, 182: 1–31.

Jerri, A. J. 1977. The Shannon sampling theorem - Its various extensions and applications: A tutorial review. Proceedings of the IEEE, 65: 1565–1596.

Munch, S., and Kottas, A. 2009. A Bayesian modeling approach for determining productivity regimes and their characteristics. Ecological Applications, 19: 527–537.

Orlova, E. L., Boitsov, A. D., Dolgov, A. V., Rudneva, G. B., and Nesterova, V. N. 2004. The relationship between plankton, capelin, and cod under different temperature conditions. ICES Journal of Marine Science, 62: 1281–1292.

Ozhigin, V., and Luka, G. 1985. Some peculiarities of capelin migrations depending on thermal conditions in the Barents Sea. *In* The Proceedings of the Soviet–Norwegian Symposium on the Barents Sea Capelin. Institute of Marine Research, Bergen, Norway, pp. 135–147.

Røttingen, I., and Tjelmeland, S. 2003. Evaluation of the absolute levels of acoustic estimates of the 1983 year class of Norwegian spring-spawning herring. ICES Journal of Marine Science, 60: 480–485.

Schlögl, A. 2006. A comparison of multivariate autoregressive estimators. Signal Processing, 86: 2426–2429.

Schnute, J. T. 2004. Mathematics and fisheries: Match or mismatch? *In* 2004 Conference on Differential Equations and Applications in Mathematical Biology, Nanaimo, BC, Canada [online]. Electronic Journal of Differential Equations, Conference, vol. 12, pp. 143–158.

Shannon, C. E. 1949. Communication in the presence of noise. Proceedings of the IRE, 37: 10–21.

Skjoldal, H., and Rey, F. 1989. Pelagic production and variability of the barents sea ecosystem. *In* Biomass yields and geography of large marine ecosystems, Ed. by Sherman, K. and Alexander, L. M. Westview Press Inc., Boulder, CO. pp. 241–286.

Skjoldal, H. R. 2004. Fish stocks and fisheries in relation to climate variability and exploitation. Natural resource system challenge: oceans and aquatic ecosystems. In: Encyclopedia of Life Supporting Systems (EOLSS), Developed under the auspices of the UNESCO. Eolss Publishers, Oxford, UK. http://www. eolss. net.

Solari, A. P., Martin-Gonzalez, J., and Bas, C. 1997. Stock and recruitment in Baltic cod (Gadus morhua): a new, non-linear approach. ICES Journal of Marine Science, 54: 427–443.

Sundelöf, A., Bartolino, V., Ulmestrand, M., and Cardinale, M. 2013. Multi-annual fluctuations in reconstructed historical time-series of a European lobster (Homarus gammarus) population disappear at increased exploitation levels. PloS One, 8: e58160.

Tanokura, Y., and Kitagawa, G. 2004. Power contribution analysis for multivariate time series with correlated noise sources. Advance and Application in Statistics, 4: 65–95.

Toresen, R., and Østvedt, O. J. 2000. Variation in abundance of Norwegian spring-spawning herring (Clupea harengus, Clupeidae) throughout the 20th century and the influence of climatic fluctuations. Fish and Fisheries, 1: 231–256.

Wong, K. F. K., and Ozaki, T. 2007. Akaike causality in state space. Biological Cybernetics, 97: 151–157.

Yndestad, H., and Stene, A. 2002. System dynamics of the barents sea capelin. ICES Journal of Marine Science, 59: 1155–1166.

Zhang, Y., and Luo, G. 2014. Inferring causal directions in errors-in-variables models. *In* Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Quebec City, Canada. pp. 3152–3153.

Handling editor: Anna Kuparinen