Dynamic factor analysis to estimate common trends in fisheries time series

A.F. Zuur, I.D. Tuck, and N. Bailey

Abstract: Dynamic factor analysis (DFA) is a technique used to detect common patterns in a set of time series and relationships between these series and explanatory variables. Although DFA is used widely in econometric and psychological fields, it has not been used in fisheries and aquatic sciences to the best of our knowledge. To make the technique more widely accessible, an introductory guide for DFA, at an intermediate level, is presented in this paper. A case study is presented. The analysis of 13 landings-per-unit-effort series for *Nephrops* around northern Europe identified three common trends for 12 of the series, with one series being poorly fitted, but no relationships with the North Atlantic Oscillation (NAO) or sea surface temperature were found. The 12 series could be divided into six groups based on factor loadings from the three trends.

Résumé : L'analyse factorielle dynamique (DFA) est une technique qui permet de détecter les structures communes dans des séries temporelles, ainsi que les relations entre les séries et les variables explicatives. Bien qu'utilisée régulièrement en économétrie et en psychologie, la méthode n'a pas été employée, au meilleur de notre connaissance, dans les domaines des pêches et des sciences aquatiques. On trouvera ici un guide d'introduction à la DFA, à un niveau intermédiaire, qui rendra la méthodologie plus généralement accessible. Nous présentons une étude de cas qui consiste en l'analyse de 13 séries de débarquements de *Nephrops* par unité d'effort de pêche sur les côtes de l'Europe du Nord. Il y a trois tendances communes à douze des séries; une des séries s'ajuste mal; il n'existe aucune relation ni avec l'oscillation nord-atlantique (NAO), ni avec la température de surface de la mer. Les 12 séries se divisent en six groupes d'après le poids des facteurs dans les trois tendances.

[Traduit par la Rédaction]

Introduction

To analyse data sets containing relatively large numbers of response variables, many researchers apply dimension-reduction techniques like principal component analysis (PCA), factor analysis, correspondence analysis (CA), or multidimensional scaling. In these techniques, a measure of similarity between the response variables is either explicitly or implicitly defined and a low dimensional graphical representation (e.g., a biplot) of these similarities is presented. If there are also explanatory variables available, various techniques can be used that take into account these variables, e.g., redundancy analysis and canonical correspondence analysis (CCA). The two latter techniques are basically a PCA and CA, respectively, in which the axes are restricted to linear combinations of the explanatory variables. Birks et al. (1994) compiled a list with more than 400 applications of dimension-reduction techniques in biological and related fields. The widespread use of these techniques is probably due to simplicity of interpretation of the results (via ordination diagrams), availability of software, and lack of other good and easy-to-use statistical tools. None of the techniques, however, is designed to analyse time series. Although it is possible to apply PCA to time-series data and connect consecutive points in time with each other, interpretation of the results is likely to be difficult. Indeed, these techniques do not take account of time in any way; if the order of time in the input data matrix is changed, the same results are obtained.

Dynamic factor analysis (DFA) is a dimension-reduction technique especially designed for time-series data. It has been used in econometric (Harvey 1989) and psychological fields (Molenaar 1985; Molenaar et al. 1992) since the mid-1980s. DFA can be used to model short, nonstationary time series in terms of common patterns and explanatory variables. For example, DFA can indicate whether there are any underlying common patterns in the N time series, whether there are interactions between the response variables, and what the effects of explanatory variables are.

The mathematics underlying DFA are rather complex and full details can be found in Zuur et al. (2003). Here, we explain in simple terms what dynamic factor analysis is, how to interpret its results, and what it can and cannot do. A time-series data set of fishery information for the Norway lobster, *Nephrops norvegicus*, is used to illustrate DFA. The data set consists of annual landings-per-unit-effort (LPUE)

Received 16 July 2002. Accepted 26 February 2003. Published on the NRC Research Press Web site at http://cjfas.nrc.ca on 10 June 2003.
J16893

A.F. Zuur,[1] I.D. Tuck, and N. Bailey. Statistics Group, Fisheries Research Services, Marine Laboratory, P.O. Box 101, Victoria Road, Aberdeen, AB11 9DB, U.K.

[1]Corresponding author (e-mail: a.zuur@marlab.ac.uk).

 doi: 10.1139/F03-030

collected from 13 stocks in northern European waters. Results of a univariate analysis on some of the time series were presented in Afonso-Dias (1997). Results presented in this paper were obtained with the software package Brodgar (www.brodgar.com).

Dynamic factor analysis

DFA is a multivariate time-series analysis technique used to estimate underlying common patterns in a set of time series. These patterns can be common trends, common seasonal effects, or common cycles. Because of the annual nature and relatively short length of most of the biological time-series data sets (15–25 years), only models with common trends are used in this paper. The time series are modelled in terms of (*i*) a linear combination of common trends, (*ii*) explanatory variables, (*iii*) a level parameter, and (*iv*) a noise component. We first consider a model with M common trends and noise only. The mathematical formulation for this model is as follows:

$$y_{it} = z_{i1}\alpha_{1t} + z_{i2}\alpha_{2t} + \ldots + z_{iM}\alpha_{Mt} + e_{it}$$

where y_{it} is the value of the ith time series at time t, α_{jt} is the jth common trend, z_{ij} is the factor loading, and e_{it} is noise. In matrix notation, this can be written as

$$\text{(1)} \qquad \mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{e}_t$$

where $\mathbf{y}_t$ is a $N \times 1$ vector containing the values of the N time series at time t, $\boldsymbol{\alpha}_t$ represent the values of the M common trends at time t, and $\mathbf{e}_t$ is a $N \times 1$ noise component, which is assumed to be normally distributed with mean 0 and covariance matrix $\mathbf{R}$. The $N \times M$ matrix $\mathbf{Z}$ contains the factor loadings and determines the exact form of the linear combinations of the common trends. By comparing factor loadings with each other, it can be inferred which common trends are important to a particular response variable and which group of response variables are related to the same common trend. The trends represent the underlying common patterns over time. Mathematically, they are modelled as

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \mathbf{f}_t$$

where $\mathbf{f}_t \sim N(0,\mathbf{Q})$, $\mathbf{Q}$ is a diagonal error covariance matrix, and $\mathbf{f}_t$ is independent of $\mathbf{e}_t$. Hence, the jth trend at time t is equal to the jth trend at time t – 1 plus a contribution of the noise component. If the corresponding diagonal element of $\mathbf{Q}$ is relatively small, then the contribution of the error component is likely to be small for all t, and the jth trend will be a smooth curve. If it is relatively large, then the jth trend will show more variation. Hence, the trends are smoothing functions over time and are independent of each other. To allow each linear combination of common trends to move up or down, a constant level parameter $\mathbf{c}$ of dimension $N \times 1$ is used. This results in $\mathrm{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{c} + \mathbf{e}_t$. To include explanatory variables, let $\mathbf{x}_t$ be a vector containing the values of the L explanatory variables at time t and $\mathbf{D}$ be an $N \times L$ matrix containing regression coefficients. Effects of explanatory variables are modelled as in linear regression:

$$\text{(2)} \qquad \mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{c} + \mathbf{D}\mathbf{x}_t + \mathbf{e}_t$$

Note that in eq. 2, $\boldsymbol{\alpha}_t$ represents hypothetical variables, and $\mathbf{x}_t$, real variables. Hence, $\boldsymbol{\alpha}_t$ is the information shared by a set of response variables that cannot be explained by the measured explanatory variables. Obviously, a good fit of a model with no latent variables and only explanatory variables is preferable because interpretation is easier.

The name "dynamic factor analysis" suggests that the technique is related to factor analysis and this is indeed correct. The relationship becomes clear from the expression for the covariance matrix of the response variables in eq. 1, which is given by

$$\text{(3)} \qquad \mathrm{Cov}(\mathbf{y}_t) = \mathbf{Z}\mathbf{Z}^t + \mathbf{R}$$

It can be shown that $\mathbf{Q}$ can be set to the identity matrix (Zuur et al. 2003). The expression in eq. 3 is the same as in factor analysis (Krzanowski 1988). The main difference between the two techniques is that in DFA the axes (or common trends) are restricted to be smooth functions over time. Another difference concerns the covariance matrix $\mathbf{R}$. Nearly all software packages use routines for factor analysis in which $\mathbf{R}$ is modelled as a diagonal matrix. This is because it has certain advantages with respect to interpretation and diagnostic tools. However, our experience using a diagonal matrix for $\mathbf{R}$ in DFA is that it can lead to common trends that are only related to two or three response variables. Alternatively, a symmetric positive-definite matrix $\mathbf{R}$ can be used, in which case off-diagonal elements of $\mathbf{R}$ represent information in two response variables that cannot be explained by the common trends. The disadvantage of using such a covariance matrix is that the number of parameters increases drastically.

The aim of DFA is to set M, the number of common trends, as small as possible but still have a reasonable model fit. The more common trends are used, the better the fit will be, but the more parameters have to be estimated and the more information has to be interpreted. In PCA, arbitrary rules exist to decide how many axes to present (Krzanowski 1988; Jolliffe 2002). Less arbitrary rules exist for DFA, for example, using Akaike's information criterion (AIC). The AIC is a function of a measure for goodness of fit and the number of parameters in the model. It can be calculated for models containing any number of common trends, and the model containing the smallest AIC value can be selected as the most appropriate model. The AIC has received some criticism in the time-series literature because it tends to select too many autoregressive terms in autoregressive models. Alternative model selection procedures exist, for example, the Consistent AIC (CAIC) or Bayesian information criterion (BIC) (Jones 1993). Because the basic dynamic factor model does not contain autoregressive terms, AIC was used here.

Case study: *Nephrops* LPUE time series from around Europe

The Norway lobster, *Nephrops norvegicus*, is one of the most valuable lobster resources in the world, with annual landings of approximately 60 000 tonnes (t). Management advice is provided from regular analytical assessments, but at present no account is taken of variability through time in environmental influences on populations, and little effort has been expended in investigating and comparing long-term trends in different stocks. If trends in the stocks are associated with environmental series, then these may be used as predictive indices to help improve assessment performance. Given the state of *Nephrops* stocks south of the Bay of Biscay (International Council for the Exploration of the Sea (ICES) 2001), where the observation of widespread declines has led to suggestions that fishing mortality rate is not solely responsible, investigations into the effects of environmental trends can only be of benefit to assessments. This paper describes the first implementation of DFA to fisheries data time series, enabling common trends to be identified, and correlations with environmental indices to be examined. ICES *Nephrops* working groups have for some time collated LPUE data at a relatively small scale compared with fish stocks. Several sets were available from northwestern Europe.

Nephrops is a mud-burrowing decapod, only available to trawls when outside the burrow, and catch rates vary over daily and seasonal time scales in relation to burrow emergence (Chapman 1980; Tuck et al. 1997). In this case study, *Nephrops* LPUE time series from 13 sites around northern Europe (Fig. 1) are investigated. We use the term “sites” to indicate discrete areas of mud inhabited by *Nephrops*. LPUE refers to the landed component of the vessels’ catch after the small and unwanted *Nephrops* are discarded. LPUE is a measure of the total landings of *Nephrops* caught in a given time period and is used as an index of abundance when averaged over a year (ICES 1999).

Some LPUE series had values between 20 and 35 $kg \cdot h^{-1}$ and others between 5 and 10 $kg \cdot h^{-1}$. This arose partly because of differences in general levels of abundances and partly because a variety of gears with different performance are used in the various fisheries. Interpretation of factor loadings and common trends is generally easier if the response variables have approximately the same scale, and for this reason, the series were standardised (mean deleted and divided by the standard deviation). As a result, all factor loadings, common trends, and fitted values will be unitless. A time-series analysis should always start with a simple plot of the (standardised) series versus time (Fig. 2). A visual inspection of this figure suggests that the series show considerable variability between sites. Differences between the *Nephrops* stocks and fishery discarding patterns may contribute to this.

The underlying questions in this case study are (*i*) whether common patterns exist in the LPUE series and (*ii*) if there are any relationships between the LPUE series and environmental factors. Sea surface temperature (SST) is frequently implicated as a factor affecting abundances and has been measured at various places. Another possible explanatory variable is the North Atlantic Oscillation (NAO) index. This index is defined as the difference between pressure at the Azores and Iceland. The index might be seen as a proxy for SST (Becker and Pauly 1996) and for this reason we used the NAO index (winter averages) as an explanatory variable in the analyses (the index does not contain missing values). The explanatory variables were standardised. Eight DFA models (Table 1) were applied to answer the underlying questions. **R** is the covariance matrix of the noise term. For each model, different values of M can be used. As explained above, models can be compared with each other via the AIC, though care is needed here and “biological interpretation” is also important in deciding which model to select as “the best model”. The AIC (Table 2) indicated that the model containing three common trends, no explanatory variables, and a (symmetric) positive-definite covariance matrix **R** was the “best” model. Results of this model are presented below.

The first common trend (Fig. 3) shows a sharp drop followed by a slight rise in the late 1970s, remaining relatively constant after this. The second common trend (Fig. 3) increases gradually and then more sharply until 1984, decreases sharply until 1993, and then increases dramatically to the end of the series (1998). The third common trend (Fig. 3) decreases gradually to the late 1970s, increases to the mid-1980s, and then declines to the end of the series, with a stable period in the late 1980s. The factor loadings (Table 2) indicate which common trend is related to which time series. For better interpretation, a visualisation is required. One option is a two-dimensional scatter plot (e.g., plotting factor loadings of trend 1 versus trend 2), but since there are three trends, this is not an ideal graphical presentation. A better option is to plot the larger factor loadings (in absolute sense) of each trend as vertical lines (Fig. 4). Sites with factor loadings smaller than an arbitrary chosen cutoff level of 0.2, in absolute sense, were not plotted in the graph. Examining this group of greater loadings indicates that the first common trend has both positive and negative factor loadings, whereas the second and third have only positive factor loadings. Interpreting factor loadings and common trends in two or more dimensions can be as difficult as in PCA. Ideally, each common trend is related to a group of response variables characterised by a common feature, e.g., geographic proximity. However, if response variables are related to more than one common trend, interpretation can be difficult. To detect a grouping in the time series, the factor loadings can be visualised with help of a Venn diagram (Fig. 5). The data series were positioned within this diagram relative to their factor loadings for each of the trends (i.e., only Trend 1 was important for FU7, whereas both trends 1 and 2 were important for FU13). Within regions in the diagram, series with opposite loadings (positive or negative) were grouped separately (Fig. 5). An arbitrary cutoff level for loadings of 0.2 was chosen to define whether a series was associated with a trend, but series with loadings just below this level were positioned adjacent to the less important trend areas. FU5 was just below the cutoff level for trend 2, whereas FU8, FU11, and FU14 were just below the cutoff level for trend 3. Factor loadings for FU12 were low for all three trends, and this series was not included in further analysis. Having plotted the series on the Venn diagram as described above, six groups can be identified, namely group 1 (FU7), group 2 (FU3, FU4, and FU15), group 3 (FU16), group 4 (FU5 and FU6), group 5 (FU8 and FU13), and group 6 (FU9, FU11, and FU14).

Groups 1, 2, and 3 are associated with only one common trend, whereas groups 4 and 5 are influenced by trends 1 and 2 (but with opposite loadings for trend 1), and Group 6 is influenced by trends 1 and 3. FU8 in group 5 also has negative factor loadings slightly below the cutoff level for Trend 3, but because it has greater loadings on trends 1 and 2, the 3rd trend has not been considered. The biological interpretations of these groupings are discussed later.

Fig. 1. Locations of the sites used in the case study in northern Europe.

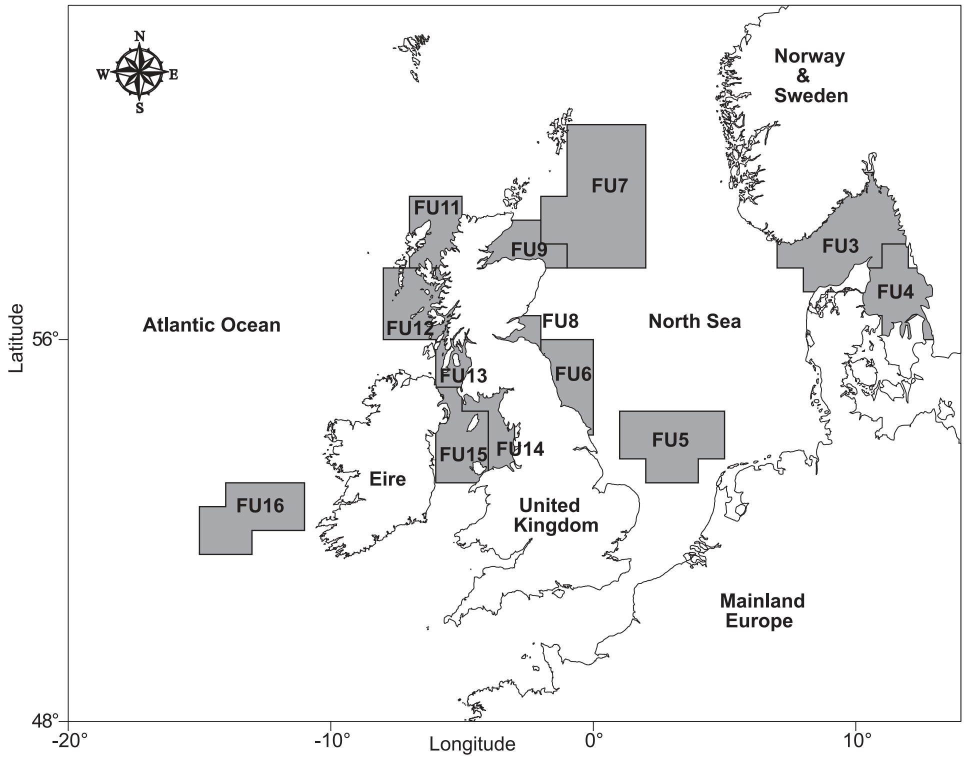
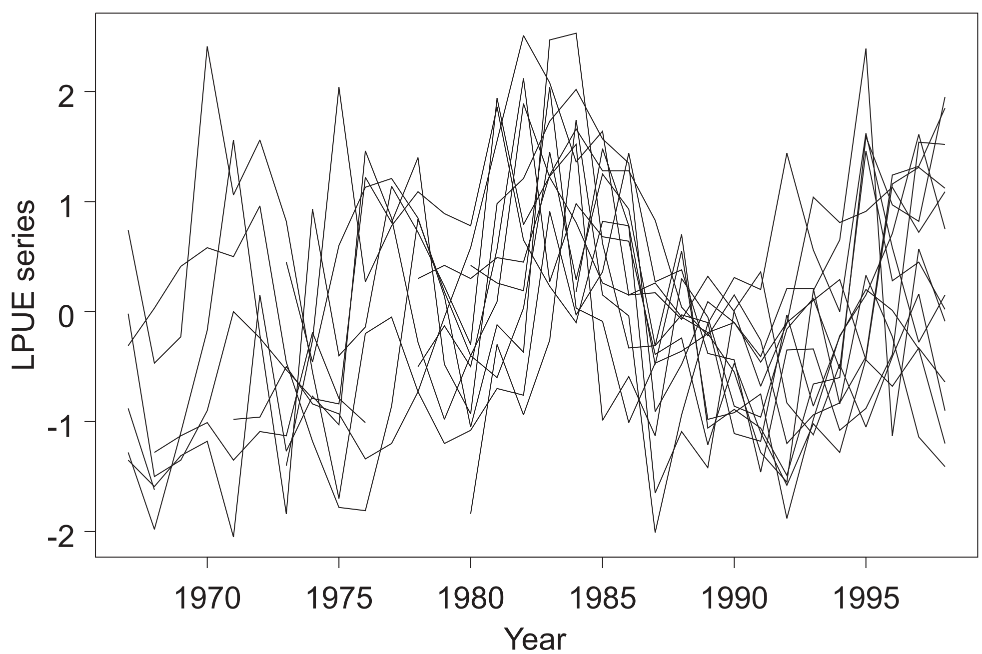


Fig. 2. Standardised *Nephrops* landings-per-unit-effort (LPUE) series measured at various sites in European waters. Each line represents the time series at a site. The series are unitless.



The fitted values (Fig. 6) suggest that a few years of the FU7, FU12, and FU5 series are not fitted particularly well. Alternatively, the ratio of the sum of squared observed totals ($\Sigma_t y_{it}^2$) and the sum of squared residuals ($\Sigma_t e_{it}^2$) can be calculated for each site (Table 2). A high ratio is an indication that the particular series, or a few years of the series, are not fitted well. Results confirm that a few years of the series FU7, FU12, and FU5 were not fitted well. Factor loadings for FU12 are very low for all common trends (Table 2). Although the first common trend factor loading for FU7 is high and greater than the cutoff for FU5, these data series only extend back over the time when the first common trend remains relatively stable, and this may be why DFA finds the series fit to the trend. The 10 remaining series are fitted

Table 1. Dynamic factor models applied to the *Nephrops* landings-per-unit-effort time series from around Europe.

	Model	Covariance matrix **R**
1	Data = M common trends + noise	Diagonal
2	Data = M common trends + noise	Positive definite
3	Data = M common trends + SST + noise	Diagonal
4	Data = M common trends + SST + noise	Positive definite
5	Data = M common trends + NAO + noise	Diagonal
6	Data = M common trends + NAO + noise	Positive definite
7	data = M common trends + SST + NAO + noise	Diagonal
8	data = M common trends + SST + NAO + noise	Positive definite

Note: SST, sea surface temperature; NAO, North Atlantic Oscillation index.

adequately by DFA. The contribution of each common trend to the fitted values can be visualised for all 13 LPUE time series (Fig. 7). The dotted line represents the first common trend plus a constant level parameter ($z_{i1}\alpha_{1t} + c_i$) for each response variable. The thin line shows the effects of the first two common trends and constant level parameter ($z_{i1}\alpha_{1t} + z_{i2}\alpha_{2tv} + c_i$) for each response variable, and the thick line represents the fitted values ($z_{i1}\alpha_{1t} + z_{i2}\alpha_{2t} + z_{i3}\alpha_{3t} + c_i$). By comparing these lines, one can identify the effects of each of the common trends. For example, the dotted line for FU3 is straight, and the other lines are nearly identical. This indicates that the first common trend has no effect on the series, the second common trend has the greatest influence, and the addition of the third has a small effect, i.e., this series is predominantly determined by the second common trend.

The off-diagonal elements of the covariance matrix **R** represent interactions between response variables that are not captured by the common trends or explanatory variables. It is interesting to inspect this matrix in the hope of finding a pattern in the joint interactions. The estimated elements can be presented in a table, but if the number of response variables is large, interpretation can be difficult. An alternative is to visualise the covariance matrix **R** using multidimensional scaling (MDS). This requires that **R** be transformed into a dissimilarity matrix. Krzanowski (1988) suggests transforming correlations into dissimilarity coefficients by

$$d(y_1,y_2) = \sqrt{2(1 - \mathrm{cor}(y_1, y_2))}$$

where $d(y_1,y_2)$ represent the dissimilarity between y_1 and y_2. The disadvantage of this transformation is that a large negative correlation is classified as “not very similar”. By using absolute correlations, both large positive and negative correlations are labelled as “similar”. MDS was applied on the resulting dissimilarity matrix (Fig. 8). Points close to each other have a high (absolute) correlation and correspond to response variables sharing a certain amount of information not explained by the common trends and explanatory variables. Results indicate that this is the case for the series FU4, FU5, and FU15. The same conclusion can be made for (*i*) FU3 and FU14, (*ii*) FU8 and FU9, and (*iii*) FU7, FU11, and FU13.

Examining the grouping of the series, biological interpretation can be made. Group 1 includes only FU7 (Fladen Ground) and is influenced solely by Trend 1, with negative factor loadings. The Fladen Ground is a large stock that has a relatively short data series over a period in which effort and landings have increased rapidly. The unique nature of this fishery may explain its position in a group of one. Group 2 includes FU3, FU4, and FU15 and is associated with the second common trend with positive factor loadings. It is perhaps not surprising that the first two show a similar trend because they are geographically adjacent (Skagerrak and Kattegat, respectively), but the third (Irish Sea West) is more puzzling. However, this data series is short compared with the others, and if the full time scale had been available, different associations may have been illustrated. Group 3 includes only FU16 (Porcupine Bank) and is associated with the third common trend with positive factor loading. The Porcupine Bank is the most westerly and isolated of the areas considered in this analysis, and this may explain its isolation in this group. It is also the only deepwater stock used in this analysis. Group 4 includes FU5 and FU6 and is a combination of the first and second common trends, with negative and positive factor loadings, respectively. Although these stocks (Botney Gut and Farn Deeps) are not adjacent, they are both in the central North Sea and so might be expected to follow similar trends. Group 5 includes FU8 and FU13 and is also a combination of the first and second common trends, both with positive factor loadings. These areas (Firth of Forth and Firth of Clyde, respectively) are on opposite coasts of Scotland, and so similarity on the basis of geographic proximity is not plausible, but both areas have an estuarine influence, and it may be that this feature that links them. Group 6 includes FU9, FU11, and FU14 and is determined by the first and third common trend, with negative factor loadings. FU14 (Irish Sea East) and FU11 (North Minch) are on the west coast of the U.K., whereas FU9 (Moray Firth) is in the North Sea. However, the location of the Moray Firth in the North Sea is probably that most influenced by the west coast water. Overall, the series associated with the second common trend appear to have a North Sea bias, whereas those associated with the third common trend may have some link with the west coast water mass. The identification of “regional” common trends and the similarity in series between some geographically adjacent sites suggests that environmental factors acting over spatial scales larger than individual sites may be influencing *Nephrops* abundance.

Although the AIC indicated that models containing the NAO index and SST were not optimal, this does not mean that these variables are not related to the LPUE series in any way. By adding an explanatory variable to the model, 13 extra parameters have to be estimated (one for each time series). If the explanatory variable is only related to a few LPUE series, then the AIC might indicate that this explanatory variable does not result in an overall model improvement. Another way of detecting effects of explanatory variables is by calculating canonical correlations. These are the cross correlations between the estimated common trends and explanatory variables. Techniques such as discriminant analysis (Huberty 1994)

Table 2. Estimated factor loadings and sum of squared measured totals ($\Sigma_t y_{it}^2$), sum of squared residual totals ($\Sigma_t e_{it}^2$), and ratio of these two sums of squares.

	Factor loadings			Sum of squares		
Site	Trend 1	Trend 2	Trend 3	$\Sigma_t y_{it}^2$	$\Sigma_t e_{is}^2$	$\Sigma_t e_{is}^2/\Sigma_t y_{is}^2$
FU3	0.00	**0.40**	0.04	28.98	3.55	0.12
FU4	–0.08	**0.45**	0.08	20.02	3.66	0.18
FU5	**–0.27**	*0.17*	0.07	17.99	13.91	0.73
FU6	**–0.23**	**0.24**	0.01	30.98	11.48	0.37
FU7	**–0.61**	–0.02	–0.04	18.00	12.02	0.67
FU8	**0.34**	**0.20**	*–0.16*	31.03	16.17	0.52
FU9	**–0.23**	0.00	**0.29**	31.02	12.83	0.41
FU11	**–0.34**	0.06	*0.17*	27.03	15.95	0.59
FU12	*–0.18*	0.13	0.06	30.96	22.26	0.72
FU13	**0.34**	**0.34**	–0.06	31.02	9.52	0.31
FU14	**–0.45**	0.04	*0.18*	25.02	14.35	0.57
FU15	0.11	**0.40**	0.04	13.96	3.31	0.24
FU16	0.02	0.14	**0.36**	27.00	2.48	0.09

Note: Factor loadings in bold are above the cutoff of 0.2 in absolute value, whereas those in italics are marginally below (see text). Factor loadings are unitless. The first common trend shows a sharp drop followed by a slight rise, remaining relatively constant after this. The second common trend increases dramatically to the end of the series. The third common trend decreases gradually to the late 1970s, increases to the mid-1980s, and then declines to the end of the series, with a stable period in the late 1980s.

Fig. 3. Common trends for the *Nephrops* landings-per-unit-effort (LPUE) series obtained by the model containing three common trends (*a*, *b*, *c*) and a symmetric, nondiagonal matrix. Common trends are unitless.

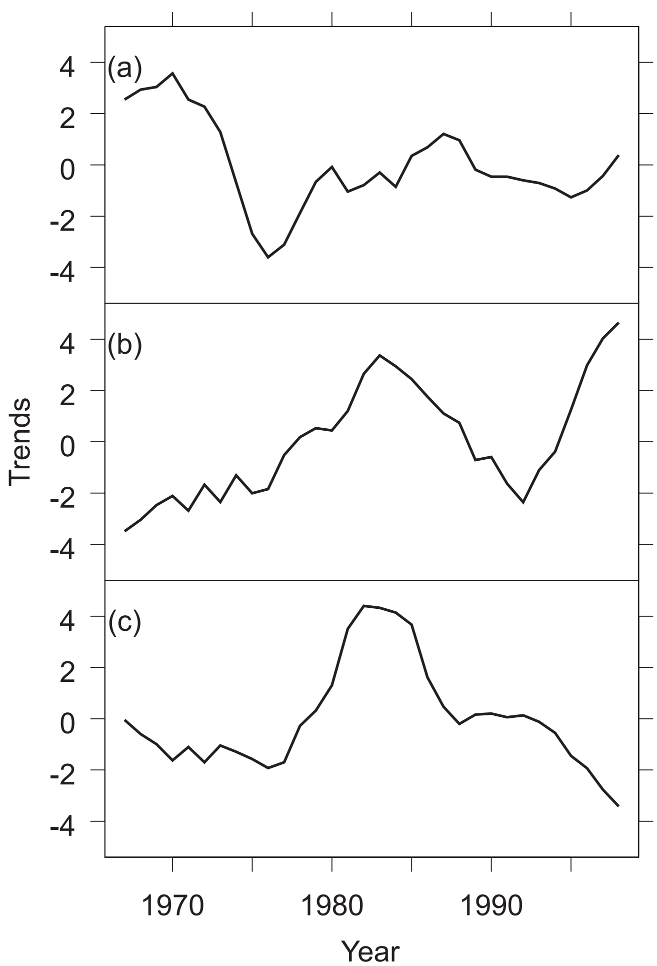


Fig. 4. Factor loadings for the *Nephrops* landings-per-unit-effort series obtained by the model containing three common trends and a symmetric, nondiagonal matrix. Factor loadings smaller than 0.2 were not plotted. Parts *a*, *b*, and *c* contain the factor loadings for trends 1, 2, and 3, respectively. Factor loadings are unitless.

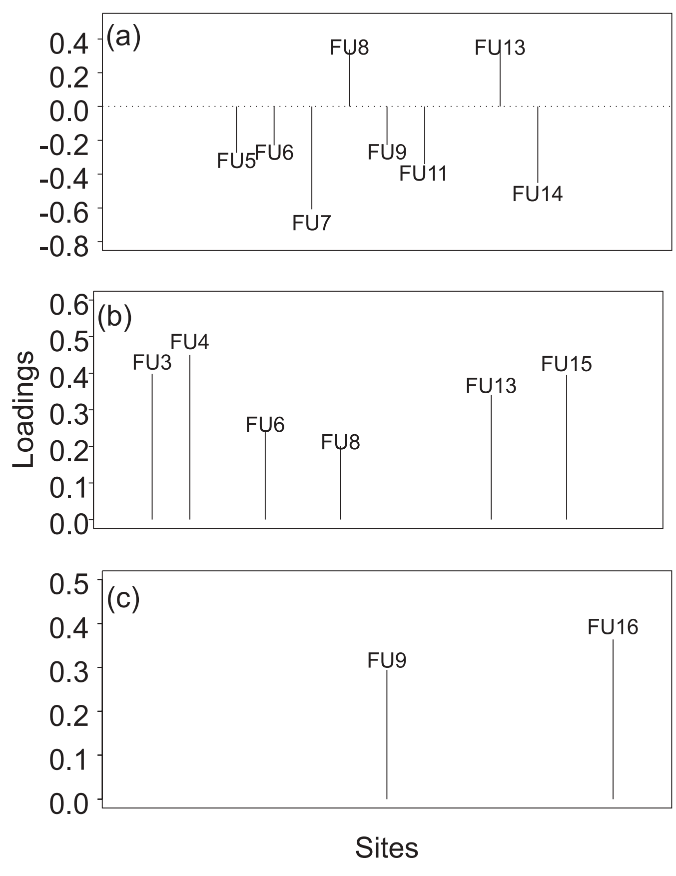
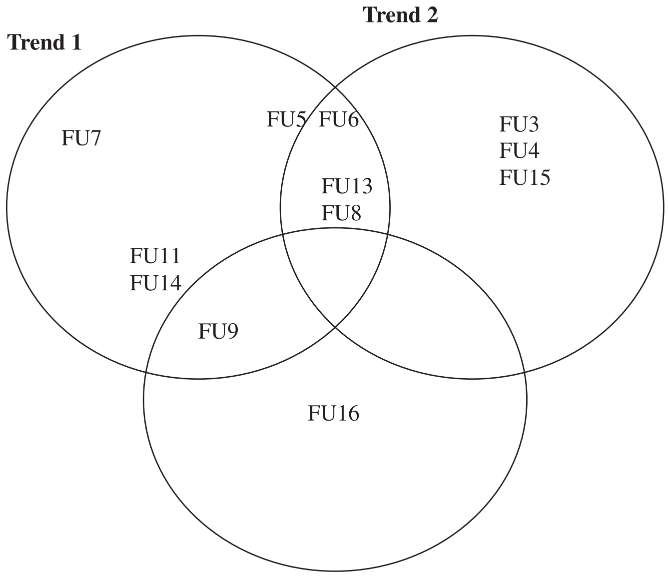


Fig. 5. Venn diagram of factor loadings.



and CCA (Ter Braak 1986) also make use of these correlations. One of the advantages of canonical correlations is that they can also be calculated for explanatory variables containing missing values. Cross correlations between the common trends and the NAO index and SST measured at various places around the U.K. coast, namely Millport (close to FU13), Fair Isle (close to FU7), and Scarborough (close to FU6), were calculated (Table 3). This table contains only cross correlations with a time lag of zero years. Results show that only 3 of the 12 estimated correlations between the common trends and four environmental variables were significant at the 5% level, and their estimated values were not convincing. The NAO index was not significantly related to any of the common trends. Cross correlations between the same variables but with time lags up to 2 years were calculated as well (Table 4). Results show that all common trends are related to at least two explanatory variables. However, the sign of the maximum time lag does not imply a causal relationship between common trends and SST or the NAO index. For example, the maximum correlation between the first common trend and the SST at Scarborough of 0.78 was obtained for a time lag of 2. This means that the SST time series at Scarborough followed the pattern of the common trend with a delay of 2 years. This might be an indication that other environmental variables not used in the analysis influence both the *Nephrops* time series and the SST time series.

Discussion

DFA applied to the European *Nephrops* time series indicated that there were three underlying common trends. Interpreting these trends was hard and involved the use of tools such as canonical correlation. Results indicated that all common trends were related to SST. However, the time lags indicated that there was no causal relationship. A speculative explanation is that there are other environmental factors, not used in the analyses, that are driving both SST and the *Nephrops* series. The dimension-reduction techniques mentioned in the Introduction would not have found these results, and as such, DFA is a very useful technique in the toolbox of scientists working with time-series data.

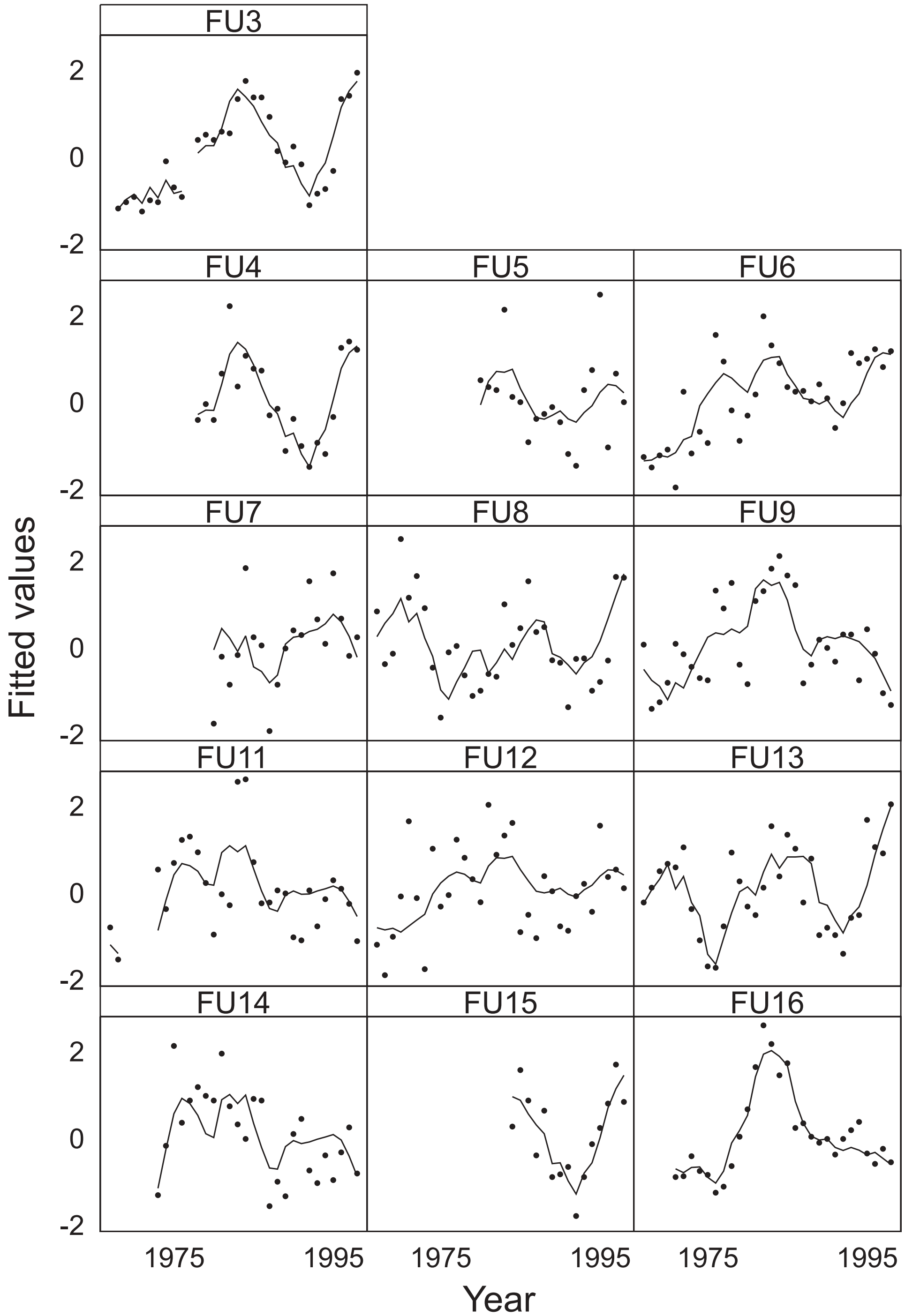
DFA is a latent variable model. The same holds for techniques such as CA and CCA (Ter Braak 1985, 1986). These techniques generate latent (hypothetical) variables, suggesting that the latent variable is an existing quantity that could have been measured. This is not always the case. Indeed, latent variables only represent a pattern present in one or more of the response variables, as modelled by the researcher via eq. 1. Tools like canonical correlations and factor loadings might be helpful for finding a plausible explanation for the common trends, but there is no guarantee that one can be found. Two other disadvantages of DFA are that it does not take into account time lags and it is based on normality. To model time lags, the model in eq. 1 can easily be extended to

$$\mathbf{y}_t = \mathbf{Z}_0\boldsymbol{\alpha}_t + \mathbf{Z}_1\boldsymbol{\alpha}_{t-1} + \mathbf{Z}_2\boldsymbol{\alpha}_{t-2} + \ldots + \mathbf{Z}_L\boldsymbol{\alpha}_{t-L} + \mathbf{e}_t$$

In these models, the response variables are modelled as a function of latent variables at time *t*, plus a time delay in these variables. If the latent variables represent a factor, e.g., temperature, such a model is plausible (see Molenaar (1985) and Molenaar et al. (1992) for a detailed discussion).

The dynamic factor model can also be seen as a regression model, and therefore it depends on the same underlying assumptions, namely normality, independence, and homogeneity of residuals. Various diagnostic tools to check the residuals for violations of these assumptions are presented in Harvey (1989), Fahrmeir and Tutz (1994), and Durbin and Koopman (2001). As in linear regression, nonnormality itself is not a serious problem. Possible reasons for nonnormality are outliers or nonlinear relationships. These might be dealt with by data transformations. An interesting approach with respect to outliers is presented in Harvey (1989), who used dummy variables to model outliers. Another reason for nonnormality might stem from the nature of the data. For example, count data with low numbers, or presence or absence data cannot be transformed into normally distributed data. For such data, the underlying dynamic factor model should be extended to Poisson or binomial distributions. The estimation procedure for DFA makes use of Kalman filtering and smoothing (Zuur et al. 2003), and Fahmeir and Tutz (1994) showed how this algorithm could be extended towards other distribution functions. Implementation of this would require extensive modification of the DFA parameter estimation procedure used in Zuur et al. (2003) and Shumway (2000) and requires further research. Heteroscedasticity of residuals can also be dealt with by data transformations (Quinn and Keough 2002). Serial correlation of the residuals is caused by an incorrect model specification in terms of the common trends and explanatory variables (Durbin and Koopman 2001). Each of the common trends is modelled as a random walk trend (Chatfield 1989; Harvey 1989). Under certain circumstances (e.g., time series with cyclic behaviour or rapid fluctuating patterns), other trend formulations might be more appropriate, for example, the local linear trend or the cyclical trend (Harvey 1989). Both can be used in the dynamic factor model, though there is no existing software capable of doing this. Another way to deal

Fig. 6. Fitted values obtained by the dynamic factor model containing three common trends and a nondiagonal matrix **R**. The lines represent the fitted values and the solid circles represent the observed values. The heading in each graph refers to the site. Fitted values are unitless.



with serial correlation is to extend the eq. 1 with an autocorrelated error structure:

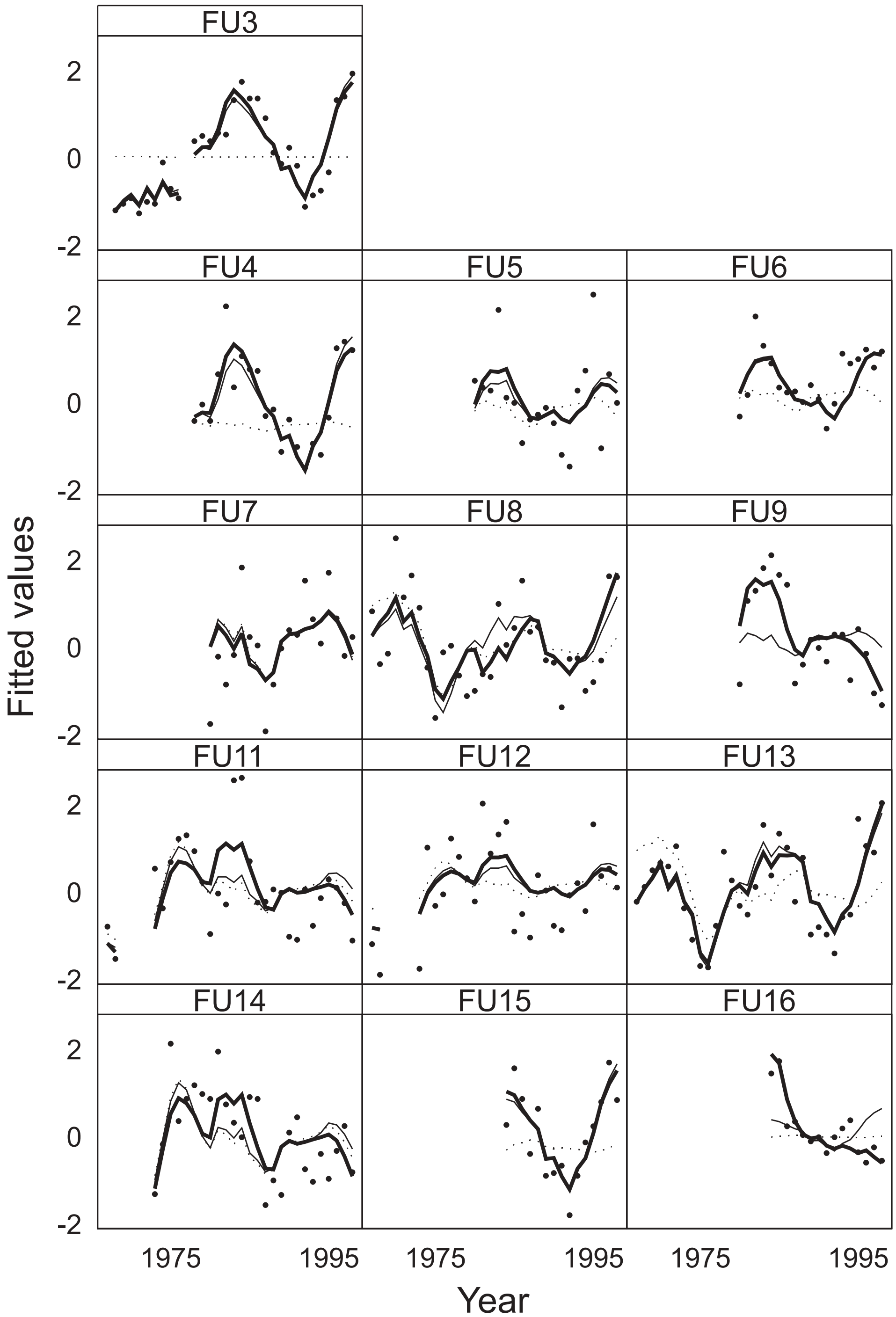
$$\mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{e}_t + \mathbf{e}_{t-1}$$

This can be further extended towards a dynamic factor model with a moving-average error structure (as in ARMA models). Alternatively, the autocorrelation mechanism can be brought into the model by using lagged observed values:

$$\mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{C}\mathbf{y}_{t-1} + \mathbf{e}_t$$

where **C** is a matrix of unknown autoregressive parameters, and further lagged observations can be added. Yet, another option to deal with serial correlation is to use a time-varying error covariance matrix **R**. However, if the diagnostic tools indicate a combination of nonnormality, heterogeneity, and serial correlation of residuals, estimated parameters and trends are biased, and further model improvements should be made.

Fig. 7. The dotted line is the curve obtained by $z_{i1}\alpha_{1t} + c_i$ and illustrates the effects of the first common trend for each landings-per-unit-effort series. The thin line represents the effects of the first two common trends ($z_{i1}\alpha_{1t} + z_{i2}\alpha_{2t} + c_i$), and the thick line is the model fit.



Just as in linear regression, violation of the independence assumption is the most serious one.

DFA was developed to analyse short, nonstationary time series with missing values and to answer the question: "what is going on?" (Zuur et al. 2003). The method is able to estimate simultaneously (*i*) the effects of explanatory variables, (*ii*) the remaining common patterns, and (*iii*) interactions between response variables. The common patterns are smoothing curves, which take account of the sequential nature of the data, and the amount of smoothing is determined by the data. This makes DFA a good alternative to smoothing methods such as LOWESS or semiparametric additive models, although results might be similar for some data sets.

Two final questions need to be addressed. (*i*) Can the method be used to predict fisheries data? (*ii*) How does it compare with other time-series techniques? As to the first question, the statistical estimation procedure uses the Kalman filter algorithm, which makes predicting future values simple; the Kalman prediction step must be run a few more times. However, if explanatory variables were used in the

Fig. 8. Multidimensional scaling applied on the transformed covariance matrix **R**. Codes refer to the sites.

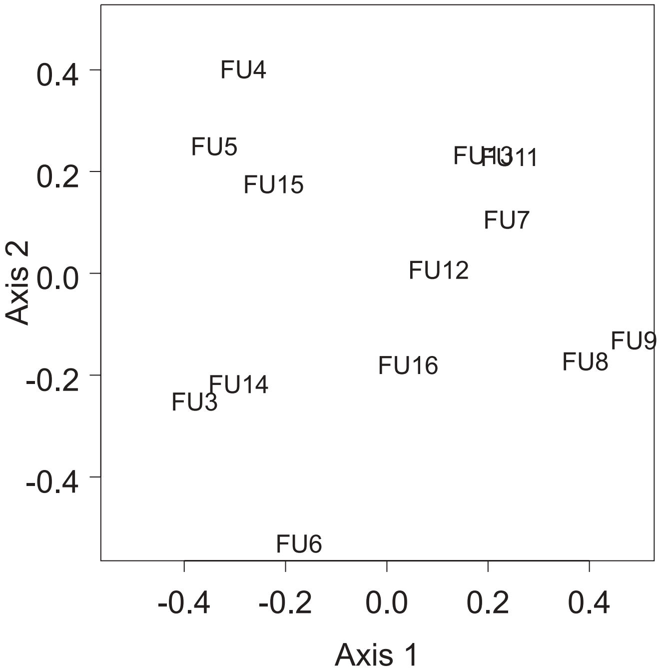


Table 3. Canonical correlations between the common trends and sea surface temperature measured at various places around the U.K. coast and North Atlantic Oscillation index (NAO).

	Millport	Fair Isle	Scarborough	NAO
Trend 1 FUs	–0.22	0.17	0.00	–0.30
Trend 2 FUs	–0.17	–0.09	**–0.43**	0.01
Trend 3 FUs	**–0.46**	**–0.51**	–0.35	0.07

Note: Values in bold refer to significant (at the 5% level) cross correlations. FU (functional unit) refers to the sites.

Table 4. Maximum (taken over time lags between –2 and 2 years) canonical correlations between the common trends and sea surface temperature series measured at various places around the U.K. coast and the North Atlantic Oscillation index (NAO).

	Millport	Fair Isle	Scarborough	NAO
Trend 1 FUs	-0.33_{-1}	0.36_{2}	$\mathbf{0.78_{2}}$	$\mathbf{-0.40_{-1}}$
Trend 2 FUs	-0.35_{-2}	$\mathbf{-0.61}_{-2}$	$\mathbf{-0.43_{0}}$	0.11_{2}
Trend 3 FUs	$\mathbf{-0.52_{1}}$	$\mathbf{-0.52_{1}}$	-0.38_{-2}	0.07_{0}

Note: Subscripts refer to time lags. Values in bold refer to significant (at the 5% level) cross correlations. Canonical correlations with a time lag k are calculated as the cross-correlation coefficient between Y_t and X_{t+k}, where Y_t is one of the variables in the rows and X_t is one of the variables in the columns. FU (functional unit) refers to the sites.

(optimal) model, one has to predict these first. To make reliable predictions for fisheries data, we advise applying DFA on at least 25–30 years of data. Using the estimated parameters and trends, one might obtain predicted values with reasonably small confidence intervals for the next 2 or 3 years, but no predictions beyond this time period should be made, as predicted confidence intervals tend to become rather large (for example, see fig. 1.6.1 in Harvey (1989)). Sjöstedt (1996) and Löfgren et al. (1993) used similar dimension-reduction techniques with the prime aim of forecasting. As to the second point, classical time-series techniques consist of autoregressive integrated moving-average models with exogenous variables (ARIMAX; Ljung 1987), Box Jenkins models (Box and Jenkins 1970), and spectral analysis (Priestley 1988). More recently, nonlinear autoregressive models (Tong 1993), wavelet analysis, and vector autoregressive models (Lütkepohl 1991; Shumway and Stoffer 2000) were successfully applied in various scientific fields. The problems with these techniques are that they are based on stationarity, need relatively long time series (wavelet analysis, spectral analysis), and cannot cope well with missing values. More importantly, these methods do not provide answers to our fundamental question, namely, what is going on? Spectral analysis and wavelet analysis give information on cyclic patterns in the data. Annual fisheries time series of only 20 years are just too short for these types of analyses. Furthermore, most fisheries time series are nonstationary. To make the time-series stationary, one can remove the trends or analyse integrated time series ($y_t - y_{t-m}$). However, for the "what is going on" question, the trends are of prime interest. In this respect, techniques based on stationarity (e.g., ARIMAX models) are of limited use for fisheries time series.

DFA is a useful technique in the toolbox of the fisheries scientist, but it should not be the only technique. Other good supplementary methods are multivariate regression trees (De Ath 2002), MAFA (Solow 1994), multivariate smoothing methods, e.g., vector GAM (Yee and Wild 1996), and fuzzy graph theory (Saila 1992). Each technique will give a certain amount of information, and together they might provide a satisfying answer to the "what is going on" question.

Up to approximately 30 time series with 10 common trends can be analysed with the algorithm described in Zuur et al. (2003) if a diagonal matrix for **R** is used. For data sets of these sizes, the algorithm is stable in the sense that the same results are obtained for different starting values. Using larger data sets means that the computing time becomes in the order of hours and the algorithm becomes unstable. We are currently extending the methodology to analyse data sets containing hundreds of time series.

The application of DFA in the case study has enabled the identification of common trends in the data. For the northern European data series, three common trends were identified with 12 series dividing into six groups (one additional series was poorly fitted by the common trends). Wider-scale common patterns may signify important environmental drivers for populations, affecting, for example, recruitment trends. Such phenomena would require more careful interpretation of trends in fisheries data as being the result of exploitation alone.

Acknowledgements

We wish to thank Rob Fryer of the Marine Laboratory in Aberdeen and Ian Jolliffe of the University of Aberdeen for valuable comments on the manuscript and Barbara Buckett (Marine Laboratory) for drawing Fig. 1.

References

Afonso Dias, M. 1997. Variability of *Nephrops norvegicus* populations in Scottish waters in relation to the sediment characteristics of the seabed. Ph.D. thesis, University of Aberdeen, Scotland.

Becker, G.A., and Pauly, M. 1996. Sea surface temperature changes in the North Sea and their causes. ICES J. Mar. Sci. **53**: 887–898.

Birks, H.J.B., Peglar, S.M., and Austin, H.A. 1994. An annotated bibliography of canonical correspondence analysis and related constrained ordination methods 1986–1993. Botanical Institute, Bergen, Norway.

Box, G.E.P., and Jenkins, G.M. 1970. Time series analysis, forecasting and control. Holden-Day, San Francisco. [Revised edition published in 1976.]

Chapman, C.J. 1980. Ecology of juvenile and adult *Nephrops*. *In* The biology and management of lobsters. Vol. II. *Edited by* J.S. Cobb and B. Phillips. Academic Press, New York. pp. 143–178.

Chatfield, C. 1989. The analysis of time series. An introduction. 4th ed. Chapman and Hall, London.

De Ath, G. 2002. Multivariate regression trees: a new technique for modelling stock–recruitment relationships. Ecology, **83**: 1105–1117.

Durbin, J., and Koopman, S.J. 2001. Time series analysis by state–space methods. Oxford University Press Inc., New York.

Fahrmeir, L., and Tutz, G. 1994. Multivariate statistical modelling based on generalized linear models. Springer-Verlag, New York.

Harvey, A.C. 1989. Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge.

Huberty, C.J. 1994. Applied discriminant analysis. John Wiley & Sons, Inc., New York.

International Council for the Exploration of the Sea. 1999. Report of the Working Group on *Nephrops* Stocks. ICES (International Council for the Exploration of the Sea) CM:1999/ACFM13.

International Council for the Exploration of the Sea. 2001. Report of the Working Group on *Nephrops* Stocks. ICES CM:2001/ACFM16.

Jolliffe, I.T. 2002. Principal component analysis. 2nd ed. Springer-Verlag, New York.

Jones, R.H. 1993. Longitudinal data with serial correlation: a state–space approach. Chapman and Hall, London.

Krzanowski, W.J. 1988. Principles of multivariate analysis: a users's perspective. Clarendon Press, Oxford.

Ljung, L. 1987. System identification: theory for the user. Prentice-Hall, New York.

Löfgren, G.K.G., Ranneby, B., and Sjöstedt, S. 1993. Forecasting the business cycle without using minimum autocorrelation factors. J. Forecast. **12**: 481–498.

Lütkepohl, H. 1991. Introduction to multiple time series analysis. Springer-Verlag, Berlin.

Molenaar, P.C.M. 1985. A dynamic factor model for the analysis of multivariate time series. Psychometrika, **50**: 181–202.

Molenaar, P.C.M., de Gooijer, J.G., and Schmitz, B. 1992. Dynamic factor analysis of nonstationary multivariate time series. Psychometrika, **57**: 333–349.

Priestley, M.B. 1988. Non-linear and non-tationary time series analysis. Academic Press, London.

Quinn, G.P., and Keough, M.J. 2002. Experimental design and data analysis for biologists. Cambridge University Press, Cambridge.

Saila, S.B. 1992. Application of fuzzy graph theory to successional analysis of a multispecies trawl fishery. Trans Am. Fish. Soc. **121**: 211–233.

Shumway, R.H. 2000. Dynamic mixed models for irregularly observed time series. Resenhas IME-USP, Vol. 4. Instituto de Matemática e Estatística, São Paulo. pp. 433–456.

Shumway, R.H., and Stoffer, D.S. 2000. Time series analysis and its applications. Springer-Verlag, Inc., New York.

Sjöstedt, S. 1996. Forecasting multiple time series using minimum/maximum autocorrelation factors. Statistical Research Report, Institute of Mathematical Statistics, University of Umeå, Sweden.

Solow, A.R. 1994. Detecting change in the composition of a multispecies community. Biometrics, **50**: 556–565.

Ter Braak, C.J.F. 1985. Correspondence analysis of incidence and abundance data: properties in terms of a unimodal response model. Biometrics, **41**: 859–873.

Ter Braak, C.J.F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology, **67**: 1167–1179.

Tong, H. 1993. Non-linear time series: a dynamical system approach. Clarendon Press, Oxford.

Tuck, I.D., Chapman, C.J., and Atkinson, R.J.A. 1997. Population biology of the Norway lobster, *Nephrops norvegicus*, in the Firth of Clyde, Scotland. I: Growth and density. ICES J. Mar. Sci. **54**: 125–135.

Yee, T.W., and Wild, C.J. 1996. Additive extensions to generalized estimation equation methods. J. R. Stat. Soc. B, **58**: 711–725.

Zuur, A.F., Fryer, R.J., Jolliffe, I.T., Dekker, R., and Beukema, J.J. 2003. Estimating common trends in multivariate time series using dynamic factor analysis. Environmetrics, **7**. In press.