Journal of Marine Systems 142 (2015) 16–24

Contents lists available at ScienceDirect

Journal of Marine Systems

journal homepage: www.elsevier.com/locate/jmarsys

Multidecadal spatial reorganisation of plankton communities in the North East Atlantic

Victoria Harris [a,*], Sofia C. Olhede [b], Martin Edwards [c,d]

[a] Department of Biostatistics, Institute of Psychiatry, King's College London, 16 De Crespigny Park, London SE5 8AF, United Kingdom
[b] Department of Statistical Science, University College, Gower Street, London WC1E 6BT, United Kingdom
[c] Sir Alister Hardy Foundation for Ocean Science, The Laboratory, Citadel Hill, Plymouth PL1 2PB, United Kingdom
[d] Plymouth University, Marine Institute, Plymouth PL4 8AA, UK

ARTICLE INFO

Article history:
Received 2 June 2014
Received in revised form 1 September 2014
Accepted 3 September 2014
Available online 16 September 2014

Keywords:
Atlantic Multidecadal Oscillation
Copepods
Diatoms
Ecoregions
Non-stationarity
Northern Hemisphere Temperature
Plankton
Regime shift

ABSTRACT

Changes in the spatial distribution of plankton populations are thought to have a profound effect on the oceanic ecosystem across all levels. This study aims to address how the spatial distribution of different plankton assemblages has changed over a multidecadal period. The multivariate structure on the CPR dataset is analysed using a technique called sparse principal component analysis. We identify functional groups of species and show that there have been changes in the ecoregions in the North East Atlantic over a multidecadal period. This technique is data-driven and can be used to identify biologically defined ecoregions based on dominant assemblages from the dataset without relying on prior knowledge. Whilst there is a change in ecoregions across time for both zooplankton and diatoms, the nature of the changes differs for the two assemblages. For zooplankton species there has been a shift in the ecoregions towards higher latitudes, implying that cold water zooplankton have moved further into arctic waters. For species that were previously restricted to the south of the region, these have been identified with increasing frequency further north. The change in spatial distribution for different species assemblages can be attributed to different factors. For example, the primary driver of zooplankton abundance across all spatial locations appears to be temperature. It is speculated that the observed northward movement of zooplankton species is a response to rising sea surface temperatures. The abundance of diatom species is instead highly correlated with the Atlantic Multidecadal Oscillation (AMO), with the spatial patterns becoming more clearly defined in its positive phase. For the diatom species, changes may be cyclic and so mean reverting to a certain extent, but for the zooplankton, continued changes can be expected if the current warming trend continues.

1. Introduction

Over recent decades several studies have suggested that a ‘regime shift’ in the spatio-temporal behaviour of North Atlantic plankton species has occurred (Beaugrand and Reid, 2003; Beaugrand et al., 2009; Beaugrand et al., 2003; Letcher, 2009; Burkill and Reid, 2010; Philippart et al., 2011; Walther et al., 2002; Beaugrand and Kirby, 2010). This is thought to be in part due to a trend towards rising Northern Hemisphere Temperature (NHT) (Beaugrand et al., 2009). An increase in average sea surface temperature (SST) is believed to have resulted in warm water species being found further north than before, coupled with declining numbers of cold water species in the northern North Sea, see Beaugrand et al. (2008) and Kirby et al. (2007). For species typically found in colder climates, such as the copepod *Calanus finmarchicus*, the temperate North East Atlantic environment may lie at the edge of its ecological niche, making it more sensitive to changes in temperature in this region (Helaout and Beaugrand, 2009). In this study we aim to address whether this hypothesised regime shift may be observed in the spatial patterns as defined by dominant species assemblages and to identify different climate trends that might influence this change in spatial patterns. The purpose of this study is to aid in our understanding of how the marine ecosystem has changed across space in response to climate.

Large scale regions of the world's oceans can be defined based on physical features, such as ocean shelves or currents, whilst smaller scale regions can be defined based on ecological features (Spalding et al., 2007). Provinces as defined by physical features divide the world's oceans into large scale regions, e.g. the South China Sea,Mediterranean Sea, or the oceanic shelf in the North East Atlantic reaches from north of Scotland to the beyond the French coast, covering all of the North Sea and parts of the open ocean. Spalding et al. (2007) divide the world's coastal waters into 62 provinces. In this study we focus on ecoregions as defined by ecological features, i.e. the dominant functional groups of taxa. We focus on ecoregions on the scale of the North East Atlantic

---

* Corresponding author.
E-mail addresses: victoria.harris@kcl.ac.uk (V. Harris), S.olhede@ucl.ac.uk (S.C. Olhede), maed@sahfos.ac.uk (M. Edwards).

http://dx.doi.org/10.1016/j.jmarsys.2014.09.002
0924-7963/

and break this region down into three subregions. These subregions are on a similar scale to the North Sea. Spalding et al. (2007) define ecoregions as the smallest subdivision of the world's oceans that are large enough to encompass ecological or life history processes for most sedentary species. In total they divide the coastal regions of the world's oceans into 232 ecoregions, e.g. Adriatic Sea, Cortezian, Ningaloo, and Ross Sea. Ecoregions are dynamic and can shift over time, meaning that they can sometimes be difficult to define using only physical data (Spalding et al., 2007) and so in this study a method for defining regions based on the dominant species over a particular time period was advocated. Many marine organisms spend part of their lifecycle in the plankton (Hardy, 1971) and thus the spatial distribution of plankton can relate to the spatial distribution of many other marine organisms, e.g. Atlantic cod (Beaugrand and Kirby, 2010) and benthic communities (Kirby et al., 2007). Plankton might also act as indicators of changes in the marine ecosystem as a whole (Edwards, 2009).

In this study the zooplankton and the phytoplankton (represented by diatoms) taxa are studied separately for changes in the temporal behaviour and spatial distribution. Using a method corresponding to sparse principal component analysis (related to work by d'Aspremont et al., 2007; Zou et al., 2006) we identify dominant assemblages across space and their joint functional behaviour.

We used data taken from the Continuous Plankton Recorder (CPR) survey (which is a database of monthly plankton abundances across the North Atlantic across multiple species). This study builds on the work of Beaugrand et al. (2002a, 2002b) and Beaugrand et al. (2009). Beaugrand et al. (2002a, 2002b) explore spatial reorganisation for a subset of the zooplankton, the calanoid copepods, by using hierarchical clustering methods on tables of diversity measures to find regions. Beaugrand et al. (2009) use previously defined species assemblages, which have been identified using ordination methods, and estimate the rate of geographical shift of each assemblage. This study builds upon previous work in that we focus on different groups of zooplankton taxa and we extend our analysis to diatom assemblages, which have not previously been explored using such methodology. In this study we aim to build upon previous work by adapting data-driven variable selection methods that can be used to produce summaries of this complex dataset. The richness is the CPR dataset which is one of its strengths but it is also what poses the greatest statistical challenge. The use of sparse principal component analysis in this study is novel in that it allows the selection of dominant species based on their joint functional response without prior knowledge and together with the cluster analysis can be used to produce summaries across space, time and species.

2. Methods

2.1. The dataset

2.1.1. The biological data

The ecological data was taken from the Continuous Plankton Recorder (CPR) survey, which is a database of over five hundred different species of plankton recorded from the world's oceans (Batten et al.,2003). The data is collected by volunteer merchant vessels at a monthly resolution, leading to an irregular spatial sampling (Vezzulli and Reid, 2003), which means that the data must first be interpolated onto a regularly spaced grid in space. The CPR survey has been operating since the 1930s, making it one of the longest running marine ecological datasets in the world. This makes it an excellent resource for studying the long term effects of environmental changes (Letcher, 2009; Hardy, 1971).

In this study we limited our analysis to data from 1958 till 2008, since reliable records for many species begin around this time (Vezzulli and Reid, 2003). In our analysis 31 species of zooplankton and 27 species of diatom are included. These species are selected based on pre-existing knowledge of the ecology, in particular these species are selected because they are known to be frequently occurring in the region of interest. For the diatoms the group rhizosolenia are excluded based on their atypical behaviour. Zooplankton and diatoms were analysed separately, since the differences in biomass and behaviour make it inadvisable to consider them together.

Raw data are given as abundances by longitude, latitude and date. The data were recorded in regular intervals across time but irregularly across space, due to the sampling method being reliant on the shipping routes of voluntary merchant vessels (Batten et al., 2003). A logarithmic transform, the log of the data transposed by one to avoid taking logarithms of zero abundances, was used to remove the relationship between the mean and variance of the data. Since abundances are counts they can be assumed to follow a Poisson distribution, i.e. large frequencies for relatively small counts but lower frequencies of very high counts. Under a Poisson distribution the variance is equal to the mean, which means that the variability increases as the count increases. The logarithmic transform removes this relationship and hence reduces the skew of the data. Prior to any analysis the data was transformed to a regularly spaced grid and kernel smoothing was used to estimate missing values (see Appendix S1 in Supporting material). The data is interpolated across each month separately due to seasonal variability (Hardy, 1971). Yearly averages are taken after interpolating to average out seasonal trends. The interpolation method was used to transfer the raw data on to a regular 1° by 1° grid, which extended from 20°W to 10°E and 40°N to 65° N.

2.1.2. Climate data

The climate data were taken from the Hadley Centre (http://www.metoffice.gov.uk/climate-change/resources/hadleycentre) and the National Climatic Data Centre (http://www.ncdc.noaa.gov/cdo-web/). We considered four climate indices as possible drivers of plankton abundance: the Northern Hemisphere Temperature (NHT) signal, the Atlantic Multidecadal Oscillation (AMO), the East Atlantic Pattern (EAP) and the North Atlantic Oscillation (NAO). These were chosen based on having been shown to influence plankton abundance and sea surface temperature (SST) in previous studies Harris et al. (2014); Beaugrand et al. (2000); Beaugrand and Kirby (2010); Beaugrand and Reid (2003); Cannaby and Husrevoglu (2009).We cannot, however, exclude the possibility that other variables that have not been included in the analysis might be influential.

2.2. Statistical techniques

One of our aims is to summarise the CPR data across multiple dimensions. Analysing the dataset across time, space and multiple species can be statistically complex but some work has been done in this area. Beaugrand et al. (2000) explored the spatio-temporal structure of the CPR dataset across 11 species using a method called three mode PCA (Tucker, 1966). In this method PCA is carried out across a species by using a space–time matrix, which means that the columns are functions of species and the rows are functions of both space and time; a time by species–space matrix, and a space by species–time matrix. This returns eigenvectors that are functions of species, time and space, respectively. These eigenvectors are then used to find a matrix of the interdependence between these three modes. Clustering is also carried out on three way tables in the three-mode PCA study in order to find groups of species and spatial regions.

In our analysis we explore 31 zooplankton species and 28 diatom species using a method called sparse principal component analysis, which allows us to select dominant taxa by their joint functional behaviour. This method also allows the number of individuals in each different functional group and the number of functional groups, i.e. the diversity of the species assemblages, to be calculated. Spatial regionalisation is then adaptively defined using *k*-means clustering (Jain et al., 1999), which is a method for partitioning a dataset into a pre-defined number of groups. The North Sea regime shift is hypothesised to occur around the mid-1980s (Beaugrand et al., 2009) and so regionalisation