Variation among northeast Atlantic regions in the responses of zooplankton to climate change: Not all areas follow the same path

N. McGinty, A.M. Power, M.P. Johnson *

School of Natural Sciences, Martin Ryan Institute, NUI Galway, University Road, Galway, Ireland

A B S T R A C T

Broad scale climate forcing can interact with local environmental processes to affect the observed ecological phenomena. This causes potential problems of over-extrapolation for results from a limited number of sites or the averaging out of region-specific responses if data from too wide an area are combined. In this study, an area similar in extent to the Celtic Biscay Large Marine Ecosystem, but including off-shelf areas, was partitioned using clustering of satellite chlorophyll (chl-a) measurements. The resulting clusters were used to define areas over which to combine copepod data from the Continuous Plankton Recorder. Following filtering due to data limitations, nine regions were defined with sufficient records for analysis. These regions were consistent with known oceanographic structure in the study area. Off-shelf regions showed a progressively later timing in the seasonal peak of chl-a measurements moving northwards. Generalised additive models were used to estimate seasonal and multiannual signals in the adult and juvenile stages of Calanus finmarchicus, C. helgolandicus and the Paracalanus–Pseudocalanus group. Associations between variables (sea surface temperature (SST), phenology and annual abundance) differed among taxonomic groups, but even within taxonomic groups, relationships were not consistent across regions. For example, in the deep waters off Spain and Portugal the annual abundance of Calanus finmarchicus has a weak positive association with SST, in contrast to the pattern in most other regions. The regions defined in this study provide an objective basis for investigations into the long term dynamics of plankton populations and suggest suitable sub regions for deriving pelagic system indicators.

1. Introduction

Climate change interacts with processes at different scales to affect ecosystem function and the distribution and abundance of individual species (Overland et al., 2010; Philippart et al., 2011-this issue). The interactions between processes at different scales can produce divergent results in the same response variable. For example, large-scale declines in the northern species *Calanus finmarchicus* (Gunnerus) (such as Planque and Fromentin, 1996) exist alongside divergent abundance falls and increases in adjoining areas (Beare et al., 2002a; Beare et al., 2002b). Specifying the scale (both geographical extent and grain size) can help clarify the key processes in a study system. Terrestrial ecosystems have been defined based on a hierarchical scale domain ranging from the global >10,000 km down to micro <10 m scales (Willis and Whittaker, 2002). These scale domains have analogues in ocean systems where continental scales (2000–10,000 km) correspond to the extent of ocean basins while regional scales (200–2000 km) correspond to the different water masses within ocean basins. Little work addressing regional scale zooplankton dynamics has been done so far in the offshore waters and continental shelf areas around Ireland; despite these setting the scene for interpretation of ecosystem changes documented in inshore areas e.g. in the English Channel (Southward et al., 1995; Hawkins et al., 2003), North Sea (Clarke et al., 2003) and German Bight (Wiltshire et al., 2010). The present study aims to define appropriate regional scales at which to study zooplankton dynamics in the offshore waters and shelf areas around Ireland. Our approach is to use Chlorophyll a (*chl-a*), as a proxy for autotrophic activity, to define coherent regions that will form the basis of our analysis.

The significance of regional variability in ecosystem structure and response has been recognised in environmental policy. For example, a fundamental component of the Ecosystem Approach to Fisheries, which forms a policy objective of international commitments to the Convention on Biological Diversity, is the need for scientific descriptions of ecosystem scales to guide appropriate management scales (Garcia et al., 2003). Two major descriptions of ecosystem scale are available for the North East Atlantic. Longhurst (1998) defined a Northeast Atlantic Shelves Province (NECS) ranging from northern Spain to the edge of the Faroe Shetland channel in the North and as far East as the Baltic Sea. This area is approximately the same as is delineated by the three Large Marine Ecosystems (LMEs) of the Baltic Sea, North Sea and Celtic-Biscay Shelf (http://www.lme.noaa.gov/). Longhurst (1998) recognised that

subdivisions based on oceanographic criteria (primarily the positions of fronts) may be appropriate in certain circumstances. The current study examines the evidence for such coherent regions at levels below the Province and LME extent, given that the CPR analysis of Beare et al. (2002b) suggests divergence among areas as small as 3 degrees latitude by 8.75 degrees longitude.

A number of decisions are required when partitioning data within spatially extensive datasets such as the Continuous Plankton Recorder (CPR) database (Beare et al., 2003). Attempts to impose a standard grain size have included using standardised statistical areas (Colebrook, 1960, 1975; Robinson, 1970), whereby data were first aggregated into 1×0.5 degree squares. The size and position of these areas ensured that enough samples were available for subsequent statistical analyses. Interpolation methods using inverse distance weighting (Beaugrand et al., 2001) and kriging (Planque and Fromentin, 1996) have produced regularly-spaced sample pixels of 1×1 degree latitude and longitude for the North Atlantic. Interpolated data were subsequently used to document pronounced changes in zooplankton community structure of the North Atlantic during recent decades, such as large poleward shifts of warm-water assemblages (Beaugrand et al., 2002a). However, interpolation from the CPR may not be an appropriate way to define regions. For example, in the standardised area C3 (Irish Sea), there are approximately 5.2 samples month^{-1} over an area ~60,000 km^2. Interpolation schemes have included the use of six or more neighbours with a search radius of 250 nm (Beaugrand et al., 2002b), which would make it necessary to select samples across prominent fronts such as the Celtic Deep Front (Pingree et al., 1982). This approach is problematic for planktonic organisms because in the process of sampling across strong physical barriers, one could move from areas of high to low biological productivity within only a few km (Molinero et al., 2008). Zooplankton assemblages including chaetognaths (Pierrot-Bults, 2008), salps and krill larvae (Molinero et al., 2008) and copepods (Berasategui et al., 2006) may be strongly influenced by frontal structures. Interpolating data across such features can mean that the particular characteristics of a distinct hydrographical boundary could be lost through smoothing.

Phytoplankton data represent, through satellite measurements, the only biological data available across wide regions with a fine grain size (Longhurst, 1998). These data integrate many of the signals relevant to the definition of regions but avoid the relative sparseness of CPR data for regions below that of LMEs. We developed a method to select, *a priori,* the regions (200–2000 km) from which to sample CPR by using satellite-derived ocean colour from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS). This dataset offers accessible high-resolution data of calibrated *chl-a* across 12 years (Vantrepotte and Melin, 2009). *Chl-a* is the principal photosynthetic pigment found in phytoplankton (Ryther and Yentsch, 1957) and this can describe much of the variance in phytoplankton abundance such as the timing of the spring bloom and changes in overall biomass (Harding et al., 2005; Raitsos et al., 2005). *Chl-a* can therefore be thought of as a proxy for variability in primary production, facilitating the clustering of regions that have similar temporal patterns (Longhurst, 1998). The close relationship between phytoplankton and copepod zooplankton (Richardson and Schoeman, 2004) implies that *chl-a* is also an appropriate measure to define relevant regions for zooplankton.

Once a series of regions was defined within an extensive study area, time series of the abundance of selected zooplankton were examined for each region, namely the copepod taxa *C. finmarchicus,* *C. helgolandicus* (Claus) and a group which includes *Paracalanus* and *Pseudocalanus*. These species are numerically important in the CPR data, are implicated as having important trophic links to fisheries (Heath and Lough, 2007) and provide useful benchmarks for comparison with CPR analyses made at different scales. Changes in phenology and average annual abundance were compared with changes in SST, to examine the effects of physical forcings, as well as to examine whether the relationships were similar or different among regions. The definition of regions and subsequent inter-regional comparisons add a different perspective to the current understanding of the links between climate and plankton dynamics, particularly as multi-regional comparisons are far less common to date than within-region analyses of zooplankton time series (Mackas and Beaugrand, 2010).

2. Materials and methods

2.1. Defining the ecosystem regions

The study area included the area of the Celtic-Biscay LME with slight overlaps into the North Sea, Iberian Coastal and Faroe Plateau LMEs. In contrast to the LMEs, waters off the shelf edge were included. ICES have also proposed eco-regions for the implementation of the ecosystem approach and Marine Strategy Framework Directive in European waters (ICES, 2004). The Celtic Seas eco-region, as defined by ICES, is contained within the study area.

Satellite-derived ocean colour data acquired from the SeaWiFS data set (assimilated *chl-a,* http://www.oceancolor.gsfc.nasa.gov) were used to partition the study area into regions with coherent phytoplankton dynamics. Due to resuspended sediments and terrestrially-derived materials, shallower coastal regions (so-called 'Case 2 waters' e.g. Alvain et al., 2005; Lee and Hu, 2006) return inflated and incorrect *chl-a* values (Harding et al., 2005) in comparison to oceanic 'Case 1 waters' (Lee and Hu, 2006). While these considerations may confound absolute estimates of phytoplankton density, influences of estuarine outflow and suspended sediments on plankton communities are well known (Tian et al., 2009). We therefore consider that the presence of yellow substance and suspended solids is information about the coastal ecosystem which will in fact add to the definition of regions.

Data for all months between 1998 and 2008 were extracted as level 3 monthly composite images (pixels of $9 \text{ km} \times 9 \text{ km}$ resolution) of *chl-a* (O'Reilly et al., 1998, 2000) and the data were $\ln(x+1)$ transformed. It became apparent that coverage within the study area decreased to $<10\%$ in the months between November and February in all years so these months were omitted from subsequent analysis. The maps were organised into a matrix where the pixels from each map were unrolled to form a column corresponding to a particular year-month, with each row consisting of values for the same pixel through time. A second filtering process was necessary on the temporal dimension, such that all rows that contained more than ten months of missing values were removed (see Beaugrand et al., 2002b for an example of a similar matrix design).

Pixel vectors (point location, separate values for each date) with similar chlorophyll patterns over time were clustered using a K-means method. This method aims to cluster *n* objects into K clusters in a *p*-dimensional space, selecting the optimum clustering by reducing the error sum of squares (E_K^2). The error sum of squares summarises the distance of each object to the centroid of the cluster to which it has been assigned. The amount of variation explained for each choice of K can be estimated using a sum of squares comparison analogous to ANOVA to generate an r^2 value (Legendre and Legendre, 1998). The total sum of squares for this calculation is calculated using the distance of each object from the global mean when K = 1. The most parsimonious number of clusters can be defined using the asymptote of r^2 values with respect to increasing values of K. Determination of the asymptote was based on the point at which the change in r^2 between different values of K fell below 5%.

K-means clustering is unsupervised, meaning that clusters can be of any size. Cluster definition based on *chl-a* was subsequently refined to ensure sufficiently large numbers of CPR samples were found within each region. The first step was to repeat the r^2-based definition of optimum K using only those clusters with over 2000 pixels. This means that small clusters, unlikely to contain sufficient CPR samples to define a time series, were excluded from the count of defined