Journal of Marine Systems
October 2016, Volume 162, Pages 18-28
http://dx.doi.org/10.1016/j.jmarsys.2016.03.013
http://archimer.ifremer.fr/doc/00323/43464/
© 2016 Published by Elsevier B.V.

Achimer
http://archimer.ifremer.fr

High frequency mesozooplankton monitoring: Can imaging systems and automated sample analysis help us describe and interpret changes in zooplankton community composition and size structure — An example from a coastal site

Romagnan Jean-Baptiste [1, 5, *], Aldammam Lama [2], Gasparini Stéphane [1, 2], Nival Paul [1, 2], Aubert Anaïs [4], Jamet Jean Louis [3], Stemmann Lars [1, 2, *]

[1] Sorbonne Universités, UPMC Univ Paris 06, UMR 7093, LOV, Observatoire Océanologique, F-06230 Villefranche-sur-mer, France
[2] CNRS, UMR 7093, LOV, Observatoire Océanologique, F-06230 Villefranche-sur-mer, France
[3] Université de Toulon, Laboratoire EBMA-PROTEE EA 3819, BP 20132, 83957 LA Garde Cedex, France
[4] Aix-Marseille Université, UM110 CNRS/INSU, IRD, Mediterranean Institute of Oceanography (MIO), IRD, 13288 Marseille Cedex 09, France
[5] Ifremer, Unité Ecologie et Modèle pour l'Halieutique, BP 21105, 44311 Nantes Cedex 03, France

* Corresponding authors : Jean-Baptiste Romagnan, email addresses : romagnan@obs-vlr.fr ; jean.baptiste.romagnan@ifremer.fr ; Lars Stemmann, email address : stemmann@obs-vlr.fr

Abstract :

The present work aims to show that high throughput imaging systems can be useful to estimate mesozooplankton community size and taxonomic descriptors that can be the base for consistent large scale monitoring of plankton communities. Such monitoring is required by the European Marine Strategy Framework Directive (MSFD) in order to ensure the Good Environmental Status (GES) of European coastal and offshore marine ecosystems. Time and cost-effective, automatic, techniques are of high interest in this context. An imaging-based protocol has been applied to a high frequency time series (every second day between April 2003 to April 2004 on average) of zooplankton obtained in a coastal site of the NW Mediterranean Sea, Villefranche Bay. One hundred eighty four mesozooplankton net collected samples were analysed with a Zooscan and an associated semi-automatic classification technique. The constitution of a learning set designed to maximize copepod identification with more than 10,000 objects enabled the automatic sorting of copepods with an accuracy of 91% (true positives) and a contamination of 14% (false positives). Twenty seven samples were then chosen from the total copepod time series for detailed visual sorting of copepods after automatic identification. This method enabled the description of the dynamics of two well-known copepod species, *Centropages typicus* and *Temora stylifera*, and 7 other taxonomically broader copepod groups, in terms of size, biovolume and abundance–size distributions (size spectra). Also, total copepod size spectra underwent significant changes during the sampling period. These changes could be partially related to changes in the copepod assemblage taxonomic composition and size distributions. This study shows that the use of

Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site.

high throughput imaging systems is of great interest to extract relevant coarse (*i.e.* total abundance, size structure) and detailed (*i.e.* selected species dynamics) descriptors of zooplankton dynamics. Innovative zooplankton analyses are therefore proposed and open the way for further development of zooplankton community indicators of changes

Highlights

► A high frequency time series of mesozooplankton net collected samples were analysed with a Zooscan and an associated semi-automatic classification technique over a year. Samples were collected on average every second day for a total of 184 samples. ► The use high of throughput imaging systems is of great interest to extract relevant descriptors of zooplankton dynamics based on coarse and detailed (species) taxonomic resolution and organisms' body size. Innovative zooplankton analyses combining size and taxonomy open the way for further development of zooplankton community indicators of changes. ► The combined analysis of simple size and taxonomic descriptors at high frequency could then be considered in the future as a relevant method to define monitoring indicators of mesozooplankton communities in the framework of European wide coastal systems as imposed to member states by the MFSD

Keywords : Mesozooplankton, Copepod, Size distributions, High frequency, Automatic classification, Zooscan, MSFD

Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site.

ACCEPTED MANUSCRIPT

Introduction

Initiatives to monitor marine coastal areas have increased in the last decades, as anthropogenic perturbations on coastal systems have become increasingly more evident (ICES report, 2014 ; Large *et al.*, 2015). The Marine Strategy Framework Directive (MSFD: 2008/56/EC) is considered to be the environmental pillar of the Integrated Maritime Policy which aims at the Good Environmental Status (GES) of European coastal and offshore marine waters by 2020. This European regulation framework considers the marine system through an ecosystem approach. Its goal to achieve GES requires thus a good understanding of the different compartments of the marine system, both abiotic and biotic. Pelagic habitats and their associated communities, more notably the large heterotrophic zooplankton (mesozooplankton and macroplankton), are scarcely studied in the context of a European wide monitoring of marine waters as considered by the MFSD. Adequate monitoring of these organisms requires consistency in sampling and analysis in order to enable the constitution of comprehensive time-series and comparisons between regions. Such consistency does not exist yet at the European scale. The ICES Working Group on Zooplankton Ecology 2014 (ICES WGZE report, 2014) report shows that a lot of coastal and offshore zooplankton time series monitoring programs exist in European waters. However, almost each program uses either a different methodology for sampling (*i.e.* sampling gear mesh size, frequency, depth) or a different sample analysis (*i.e.* binocular examination or imaging). In addition, the objectives of the monitoring may vary from one place to another (*e.g.* one monitor jellyfishes while the other may focus on a single copepod genus or species). This methodological heterogeneity may prevent scientists from addressing European scale marine ecosystems' changes in the context of current environmental changes. A good example of what kind of program should be set up exists for the open sea: the Continuous plankton Recorder (CPR) time-series (England) which is based on continuous funding and exists since the 1930's (Richardson *et al.*, 2004). Today, progress is needed to monitor meso- to macro- zooplankton communities both in coastal and open seas and at reasonable cost, in a consistent manner, across European member states. New methods should provide descriptive parameters comparable over large spatial and temporal scales. Better monitoring of meso- to macro- zooplankton is also critical to understand planktonic systems dynamics (Mitra & Davis, 2010 ; Romagnan *et al.*, 2015). Among mesozooplanktonic organisms, copepods usually make up most of the abundance and biomass, therefore they are cornerstone components of marine ecosystems and deserve particular attention in the framework of monitoring programs.

ACCEPTED MANUSCRIPT

Many of the planktonic processes take place at spatial and temporal scales that are difficult to detect by traditional binocular examination of net collected samples. Estimation under the binocular is time consuming, and is not suitable for high frequency monitoring. Alternatively, optical instruments such as the Optical Plankton Counter have been used widely for *in situ* measurements or the analyses of net samples in the last 20 years (Heath *et al.*, 1999 ; Basedow *et al.*, 2010 ; Vandromme *et al.*, 2014). Optical counters provide information on total concentration of suspended particles but their ability to distinguish between living and non-living particles is limited (Gonzalez-Quiros and Checkley, 2006 ; Stemmann and Boss, 2012a ; Vandromme *et al.*, 2014). More recently, imaging techniques followed by semi-automatic classification (automatic classification of imaged objects followed by expert validation) have been applied to *in situ* studies of zooplankton (Benfield *et al.*, 2007 ; Stemmann *et al.*, 2008 ; Forest *et al.*, 2012), or to the analyses of net samples in time series or spatial surveys (Garcia-Comas *et al.*, 2011 ; Lebourges-Dhaussy *et al.*, 2014 ; Romagnan *et al.*, 2015).

Imaging techniques can decrease the time of analysis although they cannot achieve the taxonomic resolution obtained under microscopes (Gorsky *et al.*, 2010). They were developed to analyse lots of zooplankton samples in reasonable time and therefore may be useful to achieve relevant monitoring of mesozooplankton in coastal systems. Imaging techniques provide reliable and consistent estimates of object's size that can be used to compute mesozooplankton size distributions, biomass proxies to understand community dynamics and processes (Ye *et al.*, 2013 ; Garcia-Comas *et al.*, 2014 ; Romagnan *et al.*, 2015). Size distributions have long been used as proxies for estimating food web structure and dynamics (Plat and Denman, 1978 ; Zhou, 2006) and descriptors of the flow of matter within planktonic communities (San Martin *et al.*, 2006). More recently, size distributions have also shown to be good predictors of copepods diel vertical migration behaviour (Ohman and Romagnan, 2015). Such descriptors could be the base for the development of indicators in the context of the MSFD at the scale of member-states conventions such as the North-East Atlantic Convention (OSPAR) or the Baltic-Sea one (HELCOM) (*i.e.* abundance ratio of key copepod species, spatial distribution of invasive copepod species, biomass ratio between planktonic crustacean and jellyfishes, etc.). Consider taxonomical resolutions broader than the species (*i.e.* Plankton Ecological Categories, PECs, Romagnan *et al.*, 2015, or Plankton Functional Types, PFTs, LeQuéré *et al.*, 2005) is also necessary to develop comparable and common descriptors between member-states that could in addition be used in modelling studies. The use of imaging techniques enables the consistency of taxonomical analyses, and offers the

ACCEPTED MANUSCRIPT

simultaneous measurement of size of a large number of organisms making these techniques good candidates for deployment in the framework of large scale monitoring.

The goal of this work is to demonstrate the capacity of high throughput imaging systems for the analysis of net collected mesozooplankton in the framework of a monitoring program. We hypothesize that imaging systems and associated automatic identification techniques are appropriate methods to depict complex copepod assemblages' high frequency temporal dynamics. We intend to underline the potential of imaging for the development of mesozooplanktonic assemblages dynamics indicators based on taxa at different taxonomic resolutions (*i.e.* from coarse taxonomic categories to species) and copepod size distributions. This study presents the analysis of a high frequency time series of large mesozooplanktonic copepods using the Zooscan. We focused on large copepods owing to their abundance, ease of collection and relatively good accurate automatic sorting. We emphasize the possibility to combine mainstream and innovative size distribution analyses with more traditional taxonomic data obtained through imaging to depict copepod assemblages' dynamics as well as two species dynamics, *Centropages typicus* and *Temora stylifera*. We chose to analyse specifically these two species because they are thought to be dominant at the study site, easily identifiable and have been proposed to be indicator species of plankton community shifts related to environmental changes in the Mediterranean Sea (Molinero *et al.*, 2005 and references therein). Finally, an important body of literature on these two species is available to compare our results obtained from imaging with those obtained from more traditional microscopic analysis.

Methods

Study site

The sampling station (point B) is located at the entrance of the Bay of Villefranche-sur-Mer (43°41.10 N, 7°19.00 E; water depth ~85 m), on the northern coast of the Ligurian Sea (NW Mediterranean Sea). It is characterized by low chlorophyll-*a* concentrations (Bustillos-Guzman *et al.*, 1995), and is part of the "intermittent blooming" eco-region of the Mediterranean Sea (d'Ortenzio and d'Alcala, 2009). Water-column stratification varies from well mixed in winter to strongly thermally-stratified in summer, with transition periods in April-May and November-December. Deep waters from the Var Canyon (1 km offshore) can be upwelled into the bay (Nival & Corre, 1976), providing nutrient replenishment.

ACCEPTED MANUSCRIPT

Environmental data

Water temperature, salinity and fluorescence profiles and nutrients and Chl *a* discrete samples (from depths of 1, 10, 20, 30, 50 and 75 m) were obtained during the sampling period, on a weekly basis, from the "Service d'Observation Rade" (http://www.obs-vlfr.fr/Rade/) (supported by the Oceanologic Observatory of Villefranche and the SOMLIT observation network). Meteorological data (air temperature, precipitation, atmospheric pressure, irradiance and wind speed) were provided by METEO France from daily records at a meteorological station close to the study site (Cap Ferrat semaphore ~1 km from the sampling site).

Mesozooplankton sampling

Sampling of mesozooplankton consisted of vertical hauls from a depth of 75 m to the surface, carried out with a 50 µm mesh size net (57 cm opening diameter). The 50 µm net was used instead of a more traditional wp2 200 µm (*e.g.* as used in Gorsky *et al.*, 2010) to ensure the collection of copepods > 300 µm in length that may be under sampled with a wider mesh size net. The net setting did not include a flow-meter. The filtering efficiency was assumed to be sufficiently good and constant over time given the quasi oligotrophic characteristic of the sampling site. Therefore the filtered volumes were estimated by multiplying the net's mouth opening surface by the length of cable rolled out during each deployment. Zooplankton samples were immediately preserved onboard with 4 % sodium tetraborate buffered formaldehyde. Sampling was stopped during 6 weeks in the summer of 2003 and 2 weeks in winter of 2004 and no sampling was carried out during the weekends. A total of 184 samples were collected. Sampling dates are summarized in table S1.

Mesozooplankton samples analysis

Sample treatment

The samples were digitized using the Zooscan (Gorsky *et al.*, 2010). Output images were analysed by semi-automatic plankton identification (Gorsky *et al.*, 2010). The Zooscan is a waterproof flatbed scanner that generates 16 bits grey-levels high resolution images of zooplankton samples. Each image is 25.84 cm x 15.98 cm at a 2400 dpi (dot-per-inch) resolution, which entails a linear pixel size of 10.58 µm. It is suitable for accurately measuring and identifying organisms ranging from ~300 to 5000 µm. Each of the 184 samples collected with the 50 µm net was first divided into four size fractions using 300, 500

ACCEPTED MANUSCRIPT

and 1000 µm sieves: D1 for the organisms larger than 1 mm, D2 for the organisms ranging from 500 µm to 1 mm, D3 for the organisms ranging from 300 µm to 500 µm, and finally D4 for the organisms smaller than 300 µm. The D4 size fraction (< 300 µm) was not analysed in this study because the majority of the objects it contained were barely identifiable on images, at the scanning resolution used in this study. Each of the 3 remaining size fractions (D1, D2 and D3) were aliquoted with a Motoda plankton splitter to reach aliquots of ~500-1000 objects, and imaged with the Zooscan. The initial size fractionation was intended to limit the underrepresentation of large objects that would become rare after aliquoting the whole sample at once. Image analysis was done on all images using the Zooprocess (V7.18 developed during JERICO project). Images were segmented to spot the objects, and the Zooprocess enabled the measurements of 42 features associated with each object, including morphological features (*i.e.* area of the objects, major and minor axes of ellipsoid that best fit the object, fractal dimension of the objects outlines, etc.) and grey level features (*i.e.* mean, minimum and maximum values, skewness and kurtosis of grey level distributions, etc.). The recent version of Zooprocess used in this study enabled the measurement of 5 more morphological features than those presented in Gorsky *et al.*, 2010, see table S2 for more details. The measured features were used to perform automatic identification of digitized objects.

Automatic identification

Automatic identification of objects was done using the Plankton Identifier 1.2.6 software (Gasparini, 2007). A learning set was built by visually sorting 10,400 objects in 47 categories covering the diversity of mesozooplankton of the time series as well as the diversity of non-living objects and artefacts (e.g. detritic particles, fibers, out-of-focus objects, etc.). A learning set is a set of visually identified, manually classified, objects and associated features. The learning set is the basis for building a model of the manually classified plankton categories by the use of a supervised machine learning algorithm (here, the Random Forest algorithm [Breiman, 2001]) in which the classified objects' features are input. This process outputs a model of the classified categories also referred to as a classifier. The classifier efficiency was tested by cross validation (2 folds, 5 trials). Basically, a cross validation uses a part of the learning set (randomly selected objects representing half of the initial learning set) to build the classifier, while the other part is used as a test set to estimate classification errors. Cross validation provides a confusion matrix where confusions between predicted categories

ACCEPTED MANUSCRIPT

can be diagnosed. Initial plankton categories were aggregated iteratively according to successive confusion matrices results until an acceptable trade-off between the number of distinct categories and error rates was reached. It is worth noticing that all initial objects were kept in the successive learning sets (in Gorsky *et al.*, 2010, the authors mentioned that they removed underrepresented categories from the initial learning set). The evaluation of the successive learning set performances was based on the confusion matrices outputs and two performance metrics: true positive rate and false positive rate. The true positive rate (also called Recall), is the proportion of objects that is correctly classified as belonging to a category. The false positive rate (also called contamination) is the proportion of objects that is incorrectly classified as belonging to a category. The learning set used in this study have been specifically designed to maximize the prediction efficiency of copepods and comprised 8 categories: Copepods, Chaetognatha, Gelatinous Organisms, *Cavolinia*, Zoo Large, Other, Spherical Objects and Artefacts. The category 'Copepods' contained essentially Calanoid copepods, adults and juveniles; the category 'Chaetognatha' contained mixed, unidentifiable species of Chaetognatha; the category 'Gelatinous Organisms' contained both large cnidarians (*i.e.* Medusae and Siphonophores) and large Thaliaceans (*i.e.* Salpidae and Doliolidae); the category '*Cavolinia*' contained essentially *Cavolinia inflexa* individuals juveniles and adults mixed with other *Cavolinia spp* pteropods; the category 'Zoo Large' contained a mix of large planktonic crustaceans (*i.e.* decapod larvae, euphausiids, amphipods), a few annelids (*i.e. Tomopteris sp* individuals) and a few fish larvae; the category ‘Other’ contained both planktonic (*i.e.* Cladocera, Appendicularia, and Ostracods) and non-planktonic objects (*i.e.* detritus, fecal pellets, fibres and pieces of organisms) that could not be automatically sorted out efficiently as single categories; the category 'Spherical Objects' contained both planktonic (*i.e.* single cell protists and crustacean eggs) and non-planktonic (*i.e.* air bubbles and lipidic drops) objects; the 'Artefacts' category contained unrecognizable objects that could not be allocated to any other category, essentially out of focus objects. The learning set performances are summarized in table 1. This method resulted in a time series of 184 mesozooplankton samples predicted (*i.e.* automatically identified) in 8 categories (12 less than in Gorsky *et al.*, 2010, and designed to maximize the identification of copepods), with each individual's size precisely measured. Approximately 250,000 mesozooplanktonic organisms were identified from more than 500,000 digitized objects, the remaining objects being detritus, out of focus or artefact objects.

Semi-automatic identification of Copepods

ACCEPTED MANUSCRIPT

In addition to automatic identification, 27 samples were further sorted by expert visual examination of individual objects images after the initial prediction (*i.e.*, semi-automatic method, see Gorsky *et al.*, 2010). These samples were chosen to represent most of the copepods time series variability (2-3 samples per month) and were associated to major peaks and troughs in the copepod time series. Individual images enabled the sorting of Copepods from these 27 samples in 9 groups (*i.e. Centropages typicus* adults and copepodids, *Temora stylifera* adults and copepodids, *Acartia* spp, *Oithona* spp, other calanoids, other cyclopoids and harpacticoids). These 9 groups were chosen based on their numerical abundance and the ease to recognize group members from images. More than 30,000 copepods from the 27 samples were identified. A strong effort was set to identify *Centropages typicus* adults and copepodids and *Temora stylifera* adults and copepodids because they were identified as environmental variability indicator species in the Ligurian Sea (Molinero *et al.*, 2005).

Data analysis

Biovolume

The use of imaging techniques enabled the estimation of size and biovolume of each organism in each plankton category (Vandromme *et al.*, 2012 ; Garijo & Hernandez-Leon, 2015). In this study, 'size' was measured as the length of the major axis of the ellipsoid that best fit the silhouette of the organism ('Major' feature measured by the Zooprocess, see Gorsky *et al.*, 2010). Then, the surface area of each individual object was converted into an equal area disk that was assumed to be equivalent to the cross sectional area of a sphere. The corresponding ESD (*i.e.* Equivalent Spherical Diameter) was calculated to compute the individual's biovolume following the equation:

$$Bv = 4/3\pi(ESD/2)^3 \qquad (1)$$

The biovolume of each mesozooplankton category (Bv, mm^3) was estimated by adding up the biovolumes of all individual organisms in that category, and the result was referenced to the volume of seawater sampled ($mm^3.m^{-3}$).

Size distribution

The measure of individual sizes enabled the computation of size distributions of mesozooplanktonic organisms. Plankton size distribution provides a description of the relation

ACCEPTED MANUSCRIPT

between a given plankton organism abundance and the size of individuals in the studied assemblage. This relationship can be approximated by a two-parameter power-law function:

$$n = b\mathrm{d}^{-k} \qquad (2)$$

where n is the number of individuals, b is a constant, k the slope (in log-log form), and d the particle or organism diameter. The differential plankton abundance, n = dN/dd, can be calculated from dN, the total number of objects per unit volume in a diameter range between d and d + Δd, where Δd is a small diameter increment. The slope (k) is commonly used as a descriptor of the shape of the size distribution (Stemmann and Boss, 2012). In an idealized case, slopes dynamics represent variations in the relative proportion of small vs. large individuals, *i.e.* the smaller the proportion of large individuals in the community, the steeper the slope is. Here, we sorted copepods ranging from 0.3 mm to 3 mm into 13 contiguous size bins of equal width. The binning was chosen large enough to get quantitative abundances in each size class. All the copepods larger than 3 mm were sorted in the last bin.

Hierarchical Clustering

Hierarchical Clustering was used to investigate the power of size distribution to describe Copepods temporal dynamics (agglomerative approach with flexible link [Lance & Williams, 1967], settings: α_i = 0.625, α_j = 0.625, β = -0.25, γ = 0). Prior to analysis, each sample's predicted copepods' size distribution was converted to cumulative frequency distribution. The value in each size bin (number of organism in the size bin) was divided by the total number of organisms, and multiplied by 100 to obtain frequencies by size bin. Then the successive frequencies were cumulated as a running total from the first bin to the last. The data matrix used in this analysis had 184 rows (sampling days) x 13 columns (cumulative frequencies distributions). Each of the 13 columns was a descriptor in the clustering analysis. The Kolmogorov distance was used in the analysis. This distance is defined as the absolute difference between two observed cumulative frequency distributions (Legendre and Legendre, 1984). The RST (Random Simulation Test, Guidi *et al.*, 2009) procedure was used to determine an appropriate level of partition, and a meaningful number of clusters in the size distributions time series dataset. The data formatting and analysis were performed using a Matlab 2015a custom-made code.

ACCEPTED MANUSCRIPT

Results

Environmental conditions at the studied site

The water column at the study site was thermally homogeneous during the winter (November to late March) and thermally stratified in summer with a thermocline around 25 m depth in May and June, and deeper in summer and autumn (FigureA). A rapid mixing event took place at the end of June 2003 bringing cold water to the surface. The highest surface temperatures (>25°C) were observed between in July and August and vertical stratification was disrupted during October. In 2003, chlorophyll fluorescence time series showed relatively homogeneous distributions at intermediate depths, below 40 m from April to October 2003. Chlorophyll fluorescence at the top 40 m of the water column was relatively low at that time. Two periods of high chlorophyll fluorescence appeared in summer, probably corresponding to a deep chlorophyll maximum below the thermocline (Chl *a* up to 0.76 mg Chl *a* m^{-3}, data not shown here). From autumn 2003 to spring 2004, the fluorescence signal was weak, indicating a Chl *a* depleted water column. Two water column deep pulses in fluorescence occurred in mid-February and mid-March 2004 (Figure 3B) corresponding to two consistent peaks in Chl *a* (data not shown here).

Plankton automatic classification performances

The automatic identification performances metrics (*i.e.* true positives rate and false positives rate) of each category are summarized in Table 1. The global true positives rate was 0.82, and the global false positive rate was 0.18. 'Copepods', 'Artefacts' and 'Other' were the best automatically identified categories with true positives rates of 0.91, 0.82 and 0.76, respectively. 'Zoo Large' and 'Spherical Objects' were the worst automatically identified categories with 0.23 and 0.32 true positives rates, respectively. 'Gelatinous Organisms', '*Cavolinia*' and 'Chaetognatha' were also successfully identified with true positives rate ranging from 0.67 to 0.71. The false positives rates were globally good, ranging from 0.06 to 0.40. Copepods were moderately contaminated showing a false positive rate of 0.14. The most contaminated category was ‘Zoo Large’, showing a false positive rate of 0.40.

The copepods abundances provided by the semi-automatic identification (images of individuals visually validated by a taxonomist) were compared to those of the automatic classification on 27 samples (table S1). The copepod semi-automatic abundances estimates were significantly correlated with the automatic abundances estimates (regression slope =

ACCEPTED MANUSCRIPT

0.935, 95% confident interval 0.917-0.952, n = 27, R^2 = 0.98, p < 0.001 **Error! Reference source not found.**).

High frequency dynamics of total copepods.

The seasonal evolution of the total copepod abundance estimated based on the automatic classification showed three periods of high abundances (Figure 5). The first period can be divided in two peaks of approximately one month from May 22nd, 2003 to June 25th, 2003 (maximum of ~11000 ind.m^{-3} June 12th, 2003) and from June 25th, 2003 to July 16th, 2003 (maximum of ~6500 ind.m^{-3}). The second period occurred from September 19th, 2003 to November 27th, 2003 (maximum of ~6000 ind.m^{-3} on October 07th, 2003) and the third from February 11th, 2004 to April 06th, 2004 (maximum of ~14000 ind.m^{-3}). We refer to the first, second and third zooplankton peaks as the late spring (2003), autumn (2003) and early spring (2004) peaks. In addition to the general pattern described above, short (1-2 data points) and numerous peaks and troughs are visible in the copepods time series showing the high variability of the copepods assemblage abundances over time.

High frequency dynamics of copepod size structure

Size distributions of copepods showed a common general pattern over time: decreasing copepod abundances when size increases, and a mode in abundances in the small size classes between 0.4 and 0.8 mm (Figure 6, dark red areas). During the late spring period (2003), the mode shifted gradually from ~ 0.8 mm to less than 0.5 mm, and abundances in the large size classes (*i.e.* > 1.5 mm) were low. During the 2003 autumn period, the mode of the size distribution remained around 0.5 mm but abundances of larger organisms (*i.e.* > 1.5 mm) tended to increase between October and November 2003, and decrease in December. During the 2004 period, copepods size distribution showed less prominent modes in the small size classes and more evenly distributed abundances along the size axis (Figure 6). In addition, short (< 1 month) and numerous 'peaks' (here defined as increases in abundances of large organisms) are visible in the three periods, indicating that the copepods size structure exhibits high frequency dynamics.

The RST method indicated an optimal partition level at the cut-off level 14.51 after 1000 simulations (Figure S1). This cut-off level outputs 22 clusters (Figure S2). For the sake of synthesis, we chose to partition the classification tree into four clusters (cut-off level 86.82) corresponding to four groups of size distributions.

ACCEPTED MANUSCRIPT

Group 1 (G1, red symbols, Figure 7) abundances ranged from 18.4 ind.m^{-3} to 14.10^{4} ind.m^{-3} (median = 2.4 10^{3} ind.m^{-3}). G1 size distribution slopes ranged from -6.9 to -1.2 (median = -3.1). G1 mean organisms size ranged from 0.66 mm to 0.98 mm. Group 2 (G2, blue symbols, Figure 7) abundances ranged from 1.1 10^{3} ind.m^{-3} to 1.04 10^{4} ind.m^{-3} (median = 4.5 10^{3} ind.m^{-3}). G2 size distribution slopes ranged from -4.9 to -2.8 (median = -3.4). G2 mean organisms size ranged from 0.60 mm to 0.71 mm. Group 3 (G3, green symbols, Figure 7) abundances ranged from 33.7 ind.m^{-3} to 6.2 10^{3} ind.m^{-3} (median = 996 ind.m^{-3}). G3 size distribution slopes ranged from -3.6 to -0.95 (median = -1.9). G3 mean organisms size ranged from 0.67 mm to 1.01 mm. Group 4 (G4, black symbols, Figure 7) abundances ranged from 483.2 ind.m^{-3} to 5.9 10^{3} ind.m^{-3} (median = 1.6 10^{3} ind.m^{-3}). G4 size distribution slopes ranged from -3.7 to -1.6 (median = -2.7). G4 mean organisms size ranged from 0.73 mm to 1.17 mm (Figure 7D, 7E and 7F).

Kruskal-wallis tests and subsequent pair-wise comparisons were performed to test the differences in the distributions of abundances, size distribution slopes and mean organisms size among the four clusters. There was a significant difference among the four clusters' abundances distributions (khi^{2} = 31.61, df = 3, $p < 0.001$), size distribution slopes distributions (khi^{2} = 78.29, df = 3, $p < 0.001$), and mean organisms size distributions (khi^{2} = 72.11, df = 3, $p < 0.001$). The results of the pair-wise comparisons for each of the three tests can be found in the Text S1. The copepod assemblage composition for each cluster can be found in Figure S3.

G1 was mainly observed in Spring and Autumn 2003 (85% of the cases). G2 (cluster with highest abundances, steepest slopes and smallest individuals) was observed mostly in Spring 2003 (85% of the cases). G3 was observed only in Autumn 2003 and early Spring 2004. G4 (cluster with larger organisms) was observed mostly during Spring 2003 (72% of the cases). Slopes did not show a clear seasonal cycle (Figure 7B). They tend to be steeper (higher proportion of small individuals) from Spring to Autumn 2003 and flatter (higher proportion of larger individuals) from Winter 2003 to spring 2004.

Nine copepods groups: temporal dynamics of taxa and size

The proportion of each copepod category is given in Table 2. In term of abundance, ‘other calanoids’ was the most important category (23.4 to 78.5% of total copepods abundance) followed by *Oithona* (3.6 to 50.7%) and *Acartia* (5.8 to 32.1%). *Centropages typicus* represented generally a lower proportion of total copepods (0.5-20.5%) with an

ACCEPTED MANUSCRIPT

exception on May 20, 2003 were it represented up to 31% (combining copepodids and adults) of total copepods abundances. *Temora stylifera* was rare in comparison (never more than 6.8%). Other cyclopoids and harpacticoids never exceeded 9.7% and 12.2% of the total abundance, respectively. When converted to biovolume, the proportions varied in favour of *Centropages typicus* and *Temora stylifera*. These two species (combining copepodids and adults) reached up to 42.8% (May 20th, 2003) and 18 % (November 26th, 2003) of total copepod biovolume, respectively. On the contrary, the biovolume proportion of *Oithona* and *Acartia*, 30 and 23% respectively, was less than their abundance respective proportion.

Centropages typicus copepodids and adults were present throughout the time series. They both showed their highest abundances in spring 2003 but adults became less abundant from autumn 2003 to spring 2004 whereas copepodids showed a second period of moderately high abundances in Spring 2004. In late spring 2003 adults were more abundant than juveniles. Altogether, this species' abundances ranged from 3 to 150 ind.m^{-3} with the lowest abundances in October 2003 and highest in May 2003. Size of *Centropages typicus* copepodids ranged from 0.5 mm to 1.2 mm and tended to decrease over the period of the study, whereas adults ranged from 0.8 mm to 1.5 mm, and tended to increase in size (Figure 8A and B).

Temora stylifera adults were present only during the spring and autumn of 2003 while copepodids were observed during the whole time series. Adults showed their highest abundances in autumn 2003 and almost disappeared in 2004 whereas copepodids were more abundant in spring 2003. Altogether, this species' abundances ranged from ~1 to 100 ind.m^{-3} with the lowest concentration in May 2003 and highest in October 2003. Size of *Temora stylifera* copepodids varied from 0.4 mm to 1 mm while adults' size ranged from 1 mm to 1.7 mm (Figure 8C and D).

Acartia spp. corresponded mainly to *Acartia clausi* and contained both adults and copepodids. They were present during the whole time series, and showed their higher abundances in spring 2003. Their size ranged from 0.4 to 1.2 mm. This category's abundance-size distribution is similar to that of adults *T. stylifera*, but with the dominant sizes being shifted toward the smaller size classes.

Oithona spp. contained both adults and copepodids of at least 3 species (*O. similis, O. plumifera, O. nana*). They were present at relatively constant abundances over the whole time series except during April 2003 where they were almost absent. Their size ranged from 0.4 to 0.7 mm. This category's size distribution tended to shift towards larger individuals from spring 2003 to spring 2004.

ACCEPTED MANUSCRIPT

The category “other calanoids” showed high abundance during the whole time series. Its abundance pattern followed that of total copepods (Figures 4 & 5). They ranged in size from 0.4 to more than 2 mm. This category's abundance - size distribution shows a slight shift towards a greater proportion of large individuals over time, especially between autumn 2003 and spring 2004.

The category “other cyclopoids” was present at high abundances in the spring and autumn of 2003 but showed low abundance in the spring of 2004. “Other cyclopoids” ranged in size from 0.4 to 1 mm. This category's abundance - size distribution showed a rapid shift from a population dominated by midsize organisms in May 2003 to a population dominated by small individuals from June 2003 to spring 2004.

The category "Harpacticoids" was present at high abundances in spring and autumn 2003 and decreased to low abundance levels in 2004. They ranged in size from 0.4 to 1.1 mm. This category's abundance - size distribution showed a shift from a population dominated by small individuals in spring 2003 to a population dominated by larger, midsized, individuals in autumn 2003. The population shifted back to a population dominated by small individuals in spring 2004.

Discussion

Imaging methods for the analysis of mesozooplankton

In the past years, numerous imaging systems, both bench-top and in situ, have been developed to foster plankton observation at high frequency and across plankton size range (nano- to microplankton by Imaging FlowCytoBot, IFCB, [Olson and Sosik, 2007]; FlowCAM, [Sieracki *et al.*, 1998]), mesozooplankton by Visual Plankton Recorder, VPR, [Davis *et al.*, 2005] and by Zooscan (Gorsky *et al.*, 2010), large gelatinous macroplankton and ichtyoplankton by Underwater Vision Profiler, UVP, [Stemmann *et al.*, 2008; Picheral *et al.* 2010] and In Situ Ichtyoplankton Imaging System, ISIIS, [Cowen and Guigand, 2008]). The first main issue in analysing mesozooplankton quantitatively is sampling (or imaging) a volume of water large enough to provide quantitative estimates of abundances and size distributions. Mesozooplankton quantitative analysis requires a sampling effort that is counted in cubic meters which is barely achieved by *in situ* instruments. The VPR, during a typical cast, can image volumes up to several hundreds of litres (Davis *et al.*, 2005). The UVP can

ACCEPTED MANUSCRIPT

image up to 18 cubic meters of water per 1000 m profile but its pixel size ~60 µm) do not enable the accurate identification of organisms smaller than 0.5 cm (Picheral *et al.*, 2010). The ISIIS enable the sampling of several hundreds of cubic meters of water per deployment but its pixel size (>70 µm) do not enable accurate identification of organisms smaller than a few millimetres (Cowen and Guigand, 2008). Net sampling followed by bench-top digitization may provide the best trade-off between quantitative sampling and image resolution to perform automated analysis of mesozooplankton. This advantage may be minored by the lower spatial resolution achieved by net sampling. However, the number of organisms identified is suitable for calculating quantitative estimates (Figure 2) on taxonomic categories (this study ; Vandromme *et al.*, 2011 ; Garcia-Comas *et al.*, 2011 ; Romagnan *et al.*, 2015) as well as for deriving quantitative size distributions of plankton, excluding non-living particles (This study, Marcolin *et al.*, 2013 ; Garcia-Comas *et al.*, 2014). The second main issue regarding imaging of mesozooplankton is the computing capacity and time needed to analyse large amount of images, and get planktonic data (excluding detritus and particulate matter). ISIIS, UVP and VPR generate thousands of raw images per deployment, and hundreds of Gb of raw images that needs to be processed. After processing, there can be millions of objects (including plankton, detritus and imaging artefacts) from which to extract relevant planktonic data. Handling such datasets is not easy and requires computing power, time to compute the images, and time to build predictors that can efficiently separate plankton from other particles, in order to either work on the predicted planktonic data, or to facilitate the expert taxonomic validation. The Zooscan technique presented here enables the analysis (measurements and identification) of ~3000 objects per sample per day. The total analysis time can be divided in the following steps (a) sample preparation (size fractionation and splitting): 1 h; (b) scanning: 20 min per fraction; (c) processing images: 20 min per scan; (d) prediction using an existing Learning Set: 5 min per sample; (e) optional validation in 8 plankton categories: 1000 - 1500 objects per hour. It follows that the technique presented in this paper enable the quantitative monitoring of mesozooplankton from a coarse taxonomic perspective, and a size distribution perspective, at frequencies higher than one observation per week. Our case study is intended to underline the feasibility of this statement, and develop further the idea by showing size resolved species level high frequency mesozooplankton data obtained by this technique.

Case study: the Bay of Villefranche Copepod community

ACCEPTED MANUSCRIPT

The Bay of Villefranche-sur-Mer is a coastal area monitored since the 70's (Molinero *et al.*, 2005 ; Licandro *et al.*, 2006). Weekly frequency hydrographic monitoring has shown that short-term events, such as upwelling, mixing or stratification are important triggers for plankton dynamics (Bustillo Guzman *et al.*, 1995 ; Romagnan *et al.*, 2015). The sampling site is characterized by exceptional zooplankton diversity, *i.e.* more than 160 species of copepods (Djordjevic, 1963 ; Seguin, 1981). However, no studies, to our knowledge, have addressed the monitoring of the full copepod community because traditional taxonomic counting under microscope requires extensive taxonomic skills and knowledge and is long and costly in the framework of a monitoring program. All studies using traditional methods have focused on a few target species or large plankton categories mostly to depict annual cycles or long term changes disregarding short term events (Molinero *et al.*, 2005; Licandro *et al.*, 2006). Other studies on these same time series using high throughput imaging systems have confirmed the previously observed changes, showing their potential for the study of long temporal scales. The present study shows how these methods can also help to monitor plankton dynamics at smaller time scales in coastal systems impacted by short term events such as mixing events, upwelling pulses or stratification. This study combines the analysis of coarse copepod community descriptors (*i.e.* total abundance and size distribution) obtained automatically, along with more detailed semi-automatic taxonomic sorting of copepods down to species (*i.e.* *Centropages typicus* adults and copepodids, *Temora stylifera* adults and copepodids, *Acartia* spp, *Oithona* spp., other calanoids, other cyclopoids and harpacticoids).

Size and taxa combination obtained by imaging as potential indicators of Copepods dynamics

The copepod size distributions slopes showed a seasonal evolution with a higher proportion of small individuals from late May to late October 2003 (steep slopes) and a higher proportion of large individuals from November 2003 to May 2004 (flat slopes). These changes in slopes reflect what is observed in the predicted “all copepods” assemblage (Figure 4) and in the more taxonomically detailed analysis showed in Figure 8. In the meantime, the clustering analysis shows that G1 is distributed throughout the time series and overlaps with the three other clusters. On the contrary, G2, G3 and G4 are almost completely segregated in time, following the sequence: G4 (Spring 2003), G2 (Summer to Autumn 2003) and G3 (Winter 2003 to Spring 2004).

Interestingly, size distribution slopes seem to be de-coupled from mean organisms sizes, except for G2 (steepest slopes and smallest mean organisms sizes). This decoupling is

ACCEPTED MANUSCRIPT

hypothesized to reflect the taxonomic changes as well as the developmental changes (adults, recruitment of larvae, copepodids stages) in the copepod assemblage. The mean assemblages compositions by cluster shows that the proportion of small copepods categories (*Oithona* spp and other cyclopoids) and copepodids (*C. typicus* and *T. stylifera*) increase between G1 and G2 (Figure S3 A and B) along with an increase in abundance from G1 to G2 (Figure 7D). The size structure of *Centropages typicus* shifted from a dominance of large adults (May 2003) to dominance of copepodids (June 2003). *C. typicus* life cycle is likely to significantly influence the total Copepods assemblage size distribution because adults and juveniles together can make up to 36 % of the total biovolume during a short period. For example, the sharp reduction of the total copepod mean size after the June 2003 is associated to a rapid decrease in *C. typicus* (adults and juveniles) concentration and to an increase in abundance of smaller calanoids (Figure 8A and B). It is likely that these were juveniles calanoids. These results suggest that the temporal evolution of the copepod size distributions results from change in the taxonomic composition of the assemblage as well as changes in the age structure of dominant species.

The RST method (Guidi *et al.*, 2009) underlined that the copepod size distributions observed in this study rather represent a continuum (Figure 7C) than clear clusters (Figure S1 & S2). Altogether, these observations suggest that the synthesis of a plankton community structure by the estimation of its size distribution slope only may oversimplify the description of the community. The calculation of size distribution slopes assumes a linear relationship between abundances and sizes of organisms (Platt and Denman, 1978). This assumption may not be always relevant because the linear relation is not always significant in plankton size distributions, and secondary spectra 'waves' (*i.e.* peaks after the modal size, in large size classes) can impair the significance of the linear model (Garcia-Comas *et al.*, 2014). Using all the information contained in the size structure rather than deriving a single, not always significant parameter such as the spectral slope, may help to detect more subtle changes in a plankton community (here exemplified with the copepod assemblage shifts in composition and mean size). Therefore, we suggest that the analysis of plankton size distributions slopes could be completed by other methods that take into account the full size distribution, such as size distribution clustering, in combination with taxonomic information to develop potential indicators of zooplankton community dynamics, in the framework of a high frequency monitoring program.

ACCEPTED MANUSCRIPT

Semi automatic analysis enables the detection of *C. typicus* and *T. stylifera* life cycles features

The range of concentration observed for *C. typicus* and *T. stylifera* was in the range of what previous studies have found in the same area or in the Gulf of Lion (Razouls, 1973 ; Halsband-Lenk *et al.*, 2004). *C. typicus and T. stylifera* showed noteworthy differences in their temporal patterns. *C. typicus* occurred throughout the year, with its highest abundances in spring 2003 whereas *T. stylifera* showed up at the end of Spring 2003 and developed in a narrower temporal window (Autumn 2003). This pattern is consistent with what has been observed at the scale of the Mediterranean Sea (Mazzocchi *et al.*, 2007), and more locally in the Gulf of Naples (Di Capua and Mazzocchi, 2004) and in the Bay of Villefranche (Halsband-Lenk *et al.*, 2001). Because of their short live cycle (Halsband-Lenk *et al.*, 2004), their presence over cross-seasonal time scales suggests they reproduce continuously throughout a favourable season. The successive abundance maxima depicted in Figure 8A-D suggest that spring conditions favour recruitment of *C. typicus* while autumn conditions favour that of *T. stylifera*. The successive abundance peaks of *C. typicus* may correspond to distinct generations, which is in agreement with previous studies in the Mediterranean (Halsband-Lenk *et al.*, 2001; Ianora *et al.*, 2007). *T. stylifera* adults showed two major peaks while its copepodids showed four distinct periods of high abundances suggesting that the number of generations may be equal or greater than four, in agreement with previous studies (Halsband-Lenk *et al.*, 2001 ; Halsband-Lenk *et al.*, 2004). However, *T. stylifera* almost disappeared during Spring 2003, suggesting that the number of generations may vary from year to year. The smallest *C. typicus* individuals appeared a few weeks after the highest temperature at the end of summer or the beginning of autumn, and the largest were generally observed in winter (Halsband-Lenk *et al.*, 2001). Although adults of *C. typicus* decreased in abundance to low levels from Spring to Winter 2003, their size distribution shifted towards larger size. Also, the smallest *C. typicus* copepodids were observed during the warm season, from June 2003 to October 2003 and the largest from February to April (Figure 8A, B). This is in agreement with previous observations and suggests that, at the study site, *C. typicus* winter generations are larger than spring generations.

Conclusions

The use of a high throughput imaging system to analyse zooplankton samples from a high frequency time series (184 samples over a year, equal to sampling every second day) enabled the description of the temporal dynamics of the total copepod assemblage size distributions

ACCEPTED MANUSCRIPT

and of two well-known copepods species dynamics (*Centropages typicus* and *Temora stylifera*). Subtle known seasonal changes in individual abundance-size dynamics of *Centropages typicus* and *Temora stylifera* were detected. The combined analysis of simple size and taxonomic descriptors at high frequency could then be considered in the future as a relevant method in the frame of the monitoring of zooplankton for the MSFD and for defining associated derived necessary indicators.. This conclusion is applicable to the coastal North-Western Mediterranean but similarly designed studies should be performed in other systems in order to assess the potential of such high throughput methodology for a larger application at the European level in the frameof the MSFD.

Acknowledgements

We thank the SOMLIT observational program and the Service d'Observation Rade for the hydrological data. We are also thankful to the following colleagues for their help: Corinne Desnos, Amanda Elineau and Marc Picheral. We finally thank the service RADEZOO at the Observatoire Océanographique de Villefranche-sur-Mer (OOV) and the crews of the NO Sagitta and NO Velelle for collection and the VIP3 platform for image analysis http://rade.obs-vlfr.fr/RadeZoo/). This project has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 123456, JERICO. The CNRS/UPMC Chair VISION supports Lars Stemmann.

References

Basedow, S.L., Tande, K.S., Zhou, M., 2010. Biovolume spectrum theories applied: spatial patterns of trophic levels within a mesozooplankton community at the polar front. J Plankton Res. 32(8):1105-19.

Benfield, M.C., Grosjean, P., Culverhouse, P.F., Irigoien, X., Sieracki, M.E., Lopez-Urrutia, A., Dam, H.G., Hu, Q., Davis, C.S., Hansen, A., Pilskaln, C.H., Riseman, E.M., Schultz, H., Utgof, P.E., Gorsky, G., 2007. RAPID: Research on Automated Plankton Identification, Oceanography 20.

Breiman, L., 2001. Random forests. Machine Learning 45, 5-32.

Bustillos-Guzman, J., Claustre, H. & Marty, J. C. 1995 Specific phytoplankton signatures and their relationship to hydrographic conditions in the coastal Northwestern Mediterranean sea. Marine Ecology Progress Series 124, 247–258. (doi:10.3354/meps124247)

Cowen, R.K., Guigand, C.M., 2008. In situ ichthyoplankton imaging system (ISIIS): system design and preliminary results. Limnol Oceanogr Methods. 6(2):126-32.

Davis, C.S., Thwaites, F.T., Gallager, S.M., Hu, Q. 2005. A three-axis fast-tow digital Video Plankton Recorder for rapid surveys of plankton taxa and hydrography. Limnol Oceanogr Methods.3(2):59-74.

ACCEPTED MANUSCRIPT

Di Capua, I., Mazzocchi, M.G., 2004. Population structure of the copepods Centropages typicus and Temora stylifera in different environmental conditions. Ices Journal of Marine Science 61, 632-644.

Djordjevic, M., 1963. Observations sur les copépodes pélagique en rade de Villefranche de mars à août 1962. Rapport Procès Verbal Réunion du CIESM 17, 575-590.

D'Ortenzio, F., Ribera d'Alcalà, M., 2009. On the trophic regimes of the Mediterranean Sea: a satellite analysis. Biogeosciences. 6(2):139-48.

Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L., Babin, M., 2012. Size distribution of particles and zooplankton across the shelf-basin system in southeast Beaufort Sea: combined results from an Underwater Vision Profiler and vertical net tows. Biogeosciences 9, 1301-1320.

Garcia-Comas, C., Chang, C.-Y., Ye, L., Sastri, A.R., Lee, Y.-C., Gong, G.-C., Hsieh, C.-h., 2014. Mesozooplankton size structure in response to environmental conditions in the East China Sea: How much does size spectra theory fit empirical data of a dynamic coastal area? Prog. Oceanogr. 121, 141-157.

Garcia-Comas, C., Stemmann, L., Ibanez, F., Berline, L., Mazzocchi, M.G., Gasparini, S., Picheral, M., Gorsky, G., 2011. Zooplankton long-term changes in the NW Mediterranean Sea: Decadal periodicity forced by winter hydrographic conditions related to large-scale atmospheric changes? Journal of Marine Systems 87, 216-226.

Garijo, J.C., Hernández-León, S., 2015. The use of an image-based approach for the assessment of zooplankton physiological rates: a comparison with enzymatic methods. J Plankton Res. fbv056.

Gasparini, S. 2007 Plankton Identifier : a software for automatic recognition of planktonic organisms. http://www.obs-vlfr.fr/~gaspari/Plankton_Identifier/index.php

Guidi, L., Ibanez, F., Calcagno, V., Beaugrand, G., 2009. A new procedure to optimize the selection of groups in a classification tree: Applications for ecological data. Ecological Modelling. 220(4):451-61.

Gonzalez-Quiros, R., Checkley, D.M., 2006. Occurrence of fragile particles inferred from optical plankton counters used in situ and to analyze net samples collected simultaneously. Journal of Geophysical Research-Oceans 111, -.

Gorsky, G., Ohman, M.D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., Cawood, A., Pesant, S., Garcia-Comas, C., Prejger, F., 2010. Digital zooplankton image analysis using the ZooScan integrated system. Journal of Plankton Research 32, 285-303.

Halsband-Lenk, C., Carlotti, F., Greve, W., 2004. Life-history strategies of calanoid congeners under two different climate regimes: a comparison. Ices Journal of Marine Science 61, 709-720.

Halsband-Lenk, C., Nival, S., Carlotti, F., Hirche, H.J., 2001. Seasonal cycles of egg production of two planktonic copepods, Centropages typicus and Temora stylifera, in the north-western Mediterranean Sea. Journal of Plankton Research 23, 597-609.

Heath, M.R., Dunn, J., Fraser, J.G., Hay, S.J., Madden, H., 1999. Field calibration of the Optical Plankton Counter with respect to Calanus finmarchicus. Fisheries Oceanography [Fish. Oceanogr.] 8, 13-24.

Ianora, A., Miralto, A., Halsband-Lenk, C., 2007. Reproduction, hatching success, and early naupliar survival in Centropages typicus. Prog. Oceanogr. 72, 195-213.

Lance, G.N., Williams, W.T., 1967.A General Theory of Classificatory Sorting Strategies 1. Hierarchical Systems. The Computer Journal. 9(4):373-80.

Large, S.I., Fay, G., Friedland, K.D., Link, J.S., 2015. Critical points in ecosystem responses to fishing and environmental pressures. Mar Ecol Prog Ser 521, 1-17.

Lebourges-Dhaussy, A., Huggett, J., Ockhuis, S., Roudaut, G., Josse, E., Verheye, H., 2014. Zooplankton size and distribution within mesoscale structures in the Mozambique Channel: A comparative approach using the

ACCEPTED MANUSCRIPT

TAPS acoustic profiler, a multiple net sampler and ZooScan image analysis. Deep-Sea Research Part Ii-Topical Studies in Oceanography 100, 136-152.

Legendre, L., Legendre, P., 1984. Ecologie numérique 1. Le traitement multiple des données écologiques. Masson, Québec.

LeQuéré, C.L., Harrison, S.P., Colin Prentice, I., Buitenhuis, E.T., Aumont, O., Bopp, L., et al., 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. Global Change Biology. 11(11):2016-40.

Licandro, P., Ibanez, F. & Etienne, M. 2006 Long-term fluctuations (1974-1999) of the salps Thalia democratica and Salpa fusiformis in the northwestern Mediterranean Sea: Relationships with hydroclimatic variability. Limnology and Oceanography 51, 1832–1848.

Marcolin, C.d.R., Schultes, S., Jackson, G.A., Lopes, R.M., 2013. Plankton and seston size spectra estimated by the LOPC and Zoo Scan in the Abrolhos Bank ecosystem (SE Atlantic). Continental Shelf Research 70, 74-87.

Martin, E.S., Irigoien, X., Harris, R.P., Lopez-Urrutia, A., Zubkov, M.V., Heywood, J.L., 2006. Variation in the transfer of energy in marine plankton along a productivity gradient in the Atlantic Ocean. Limnology and Oceanography 51, 2084-2091.

Mazzocchi, M.G., Christou, E.D., Di Capua, I., de Puelles, M.L.F., Fonda-Umani, S., Molinero, J.C., Nival, P., Siokou-Frangou, I., 2007. Temporal variability of Centropages typicus in the Mediterranean Sea over seasonal-to-decadal scales. Prog. Oceanogr. 72, 214-232.

Mitra, A., Davis, C., 2010. Defining the « to » in end-to-end models. Progress in Oceanography. 84(1–2):39-42.

Molinero, J.C., Ibanez, F., Souissi, S., Chifflet, M., Nival, P., 2005. Phenological changes in the Northwestern Mediterranean copepods Centropages typicus and Temora stylifera linked to climate forcing. Oecologia 145, 640-649.

Nival, P. & Corre, M. 1976 Annual Variation of Surface Hydrology in Bay of Villefranche-Sur-Mer. Annales de l'Institut Oceanographique 52, 57–78.

Olson, R.J., Sosik, H.M., 2007. A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot. Limnol Oceanogr Methods. 5(6):195-203.

Ohman, M.D.O., Romagnan, J.-B. 2015. Nonlinear effects of body size and optical attenuation on Diel Vertical Migration by zooplankton. Limnol Oceanogr, *in press.*

Picheral, M., Guidi, L., Stemmann, L., Karl, D.M., Iddaoud, G., Gorsky, G., 2010. The Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. Limnol Oceanogr Methods. 8(9):462-73.

Platt, T. & Denman, K. 1978 The structure of pelagic marine ecosystems. Journal du Conseil International pour l'Exploration de la Mer 173, 60–65.

Razouls, C., 1973. Quantitative Annual Variations of 2 Dominant Copepod Plankton Species, Centropages-Typicus and Temora-Stylifera from Banyuls Region - Biological Cycles and Production Estimates .1. Methodology and Demonstration of Quantitative Importance of These 2 Species through Relationship to Copepod Group. Cahiers De Biologie Marine 14, 361-390.

Razouls, C., Guiness, C., 1973. Varaitions annuelles quantitatives de deux especes dominantes de copepodes planctoniques, Centropages typicus et Temora stylifera, de la région de Banyuls sur Mer.II Variations dimensionnelles et mesures de la croissance. Cahier de Biologie Marine 14, 413-427.

Richardson, A.J., John, E.H., Irigoien, X., Harris, R.P., Hays, G.C., 2004. How well does the Continuous Plankton Recorder (CPR) sample zooplankton? A comparison with the Longhurst Hardy Plankton Recorder (LHPR) in the northeast Atlantic. Deep Sea Research Part I: Oceanographic Research Papers. 51(9):1283-94.

ACCEPTED MANUSCRIPT

Romagnan, J.-B., Legendre, L., Guidi, L., Jamet, J.-L., Jamet, D., Mousseau, L., Pedrotti, M.-L., Picheral, M., Gorsky, G., Sardet, C., Stemmann, L., 2015. Comprehensive Model of Annual Plankton Succession Based on the Whole-Plankton Time Series Approach. Plos One 10.

San Martin, E., Harris, R. P. & Irigoien, X. juillet Latitudinal variation in plankton size spectra in the Atlantic Ocean. Deep Sea Research Part II-Topical Studies in Oceanography 53, 1560–1572. (doi:10.1016/j.dsr2.2006.05.006)

Schultes, S., Sourisseau, M., Le Masson, E., Lunven, M., Marie, L., 2013. Influence of physical forcing on mesozooplankton communities at the Ushant tidal front. Journal of Marine Systems 109, S191-S202.

Seguin, G., 1981. Dynamique des copépodes pélagiques en rade de Villefranche-sur-Mer à partir de prélèvements quotidiens (année 1972). Bilan quantitatif et écologique. Oceanologica Acta. 4, 405-414.

Sieracki, C.K., Sieracki, M.E., Yentsch, C.S., 1998. An imaging-in-flow system for automated analysis of marine microplankton. Mar Ecol Prog Ser. 168:285-96.

Stemmann, L., Boss, E., 2012. Plankton and Particle Size and Packaging: From Determining Optical Properties to Driving the Biological Pump, in: Carlson, C.A.G.S.J. (Ed.), Annu Rev Mar Sci, pp. 263-290.

Stemmann, L., Youngbluth, M., Robert, K., Hosia, A., Picheral, M., Paterson, H., Ibanez, F., Guidi, L., Lombard, F., Gorsky, G., 2008. Global zoogeography of fragile macrozooplankton in the upper 100-1000 m inferred from the underwater video profiler. Ices Journal of Marine Science 65, 433-442.

Vandromme, P., Nogueira, E., Huret, M., Lopez-Urrutia, A., Gonzalez-Nuevo Gonzalez, G., Sourisseau, M., Petitgas, P., 2014. Springtime zooplankton size structure over the continental shelf of the Bay of Biscay. Ocean Science 10, 821-835.

Vandromme, P., Stemmann, L., Garcìa-Comas, C., Berline, L., Sun, X., Gorsky, G., 2012. Assessing biases in computing size spectra of automatically classified zooplankton from imaging systems: A case study with the ZooScan integrated system. Methods in Oceanography. 1–2:3-21.

Vandromme, P., Stemmann, L., Berline, L., Gasparini, S., Mousseau, L., Prejger, F., Passafiume, O., Guarini, J.M., Gorsky, G., 2011. Inter-annual fluctuations of zooplankton communities in the Bay of Villefranche-sur-mer from 1995 to 2005 (Northern Ligurian Sea, France). Biogeosciences 8, 3143-3158.

Ye, L., Chang, C.-Y., Garcia-Comas, C., Gong, G.-C., Hsieh, C.-h., 2013. Increasing zooplankton size diversity enhances the strength of top-down control on phytoplankton through diet niche partitioning. Journal of Animal Ecology 82, 1052-1060.

Zhou, M. 2006 What determines the slope of a plankton biomass spectrum? Journal of Plankton Research 28, 437–448. (doi:10.1093/plankt/fbi119)

Figures captions

Figure 1: Location of the sampling site, Point B in the coastal Ligurian sea, and bathymetric map of the bay of Villefranche (modified from Romagnan *et al.*, 2015)

Figure 2: Exemple images of the 9 copepod categories determined using the semi-automatic classification method (prediction followed by expert visual examination). A) *Centropages typicus* copepodids, B) *Centropages typicus* adults, C) *Temora stylifera* copepodids, D) *Temora stylifera* adults, E) *Acartia spp*, F) *Oithona spp*, G) other calanoids, H) other cyclopoids and I) Harpacticoids.

Figure 3: Time series (2003/2004) of A) temperature and B) chlorophyll fluorescence of the water column near the sampling point.

ACCEPTED MANUSCRIPT

Figure 4: Corrélation of Copepods abundance estimates using the automatic classification method and the semi automatic classification method. The dashed line represent the 1:1 line. (slope = 0.935, 95% Confident Interval [0.917-0.952], R^2=0.98, $p < 0.001$, n=27).

Figure 5: Time series of total copepods abundance. The time series extended from April, 28th 2003 to April, 27th 2004 (n = 184). The red line shows the value of the third quartile of the Copepods abundances distribution (Q3 = 3313 ind.m^{-3}).

Figure 6: Time series of the automatically classified copepod size distributions from April, 28th 2003 to April, 27th 2004 (n = 184). The color scale represent the abundance of copepods in each size bin (log10 ind.m^{-3}). The dotted vertical line separates years 2003 and 2004.

Figure 7: A) Time series of copepods abundance showing the temporal distributions of the four clusters. B) Size distributions plotted as cumulative frequency and colored according to their groups determined by the hierarchical clustering analysis , C) time series of size distributions slope (same color code as above). D) distributions of the abundances in the 4 clusters, E) distributions of size distribution slopes in the 4 clusters, F) distributions of mean sizes in the 4 clusters. D-F: the red horizontal bar represent the median, the extension of the blue box represent the interquartile range, the whiskers represent the 5-95% extension of the distributions. Outliers are depicted as red crosses.

Figure 8: Time series of abundance - size distributions of the 9 copepod categories determined using the semi-automatic classification method (prediction followed by expert visual examination). The 27 samples that were visually identified and counted are shown by the white dots on top of each plot. A) *Centropages typicus* copepodids, B) *Centropages typicus* adults, C) *Temora stylifera* copepodids, D) *Temora stylifera* adults, E) *Acartia spp*, F) *Oithona spp*, G) other calanoids, H) other cyclopoids and I) Harpacticoids.

Tables and tables captions

Table 1: Eight categories learning set performances.

	# of objects	True positives rate	False positives rate
Copepods	5284	0.91	0.14
Chaetognatha	240	0.67	0.14
Gelatinous Organisms	335	0.71	0.25
Zoo_Large	299	0.23	0.40
Cavolinia	83	0.68	0.28
Other	3285	0.76	0.24
Spherical Objects	39	0.32	0.06
Artefacts	835	0.82	0.11
Total	10400	0.82	0.18

Table 2: Contributions of the different copepods categories, in terms of abundance and biovolume, to the total copepod assemblage. The medians, minima and maxima have been calculated using the 27 samples analysed with semi automatic classification method.

Copepod category	**Contribution (% of abundance)**			**Contribution (% of biovolume)**		
	Med.	**Min**	**Max**	**Med.**	**Min**	**Max**

ACCEPTED MANUSCRIPT

***Centropages Typicus* copepodids**	2.5	0.5	20.5	2.9	0.6	21.2
***Centropages Typicus* adults**	0.8	0.1	9.8	7.3	1.1	21.6
***Temora stylifera* copepodids**	0.7	0.2	6.8	0.7	0.1	12.5
***Temora stylifera* adults.**	0.1	0.0	1.5	0.4	0	5.5
Acartia spp	12.8	5.8	32.1	10.6	4.6	27.0
Oithona spp	12.2	3.6	50.7	6.1	0.8	29
other calanoids	59.8	23.4	78.5	62.3	23.1	85
other cyclopoids	4.4	0.5	9.7	2.7	0.4	12.5
Harpacticoids	1.8	0.0	12.2	0.6	0	9.5

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figures

Figure 2

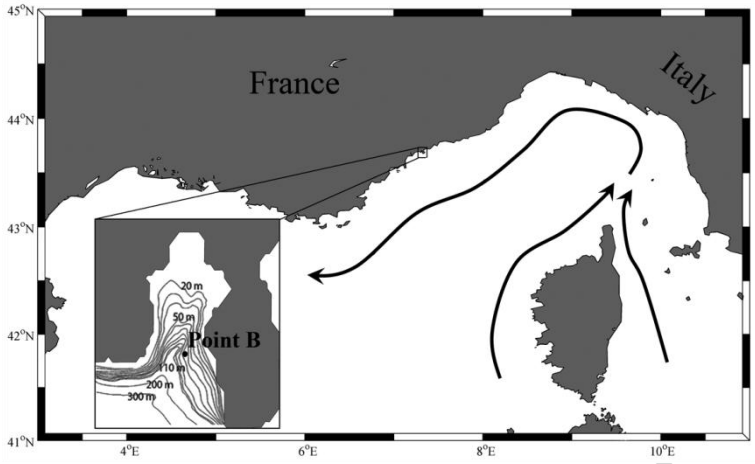


Figure 2

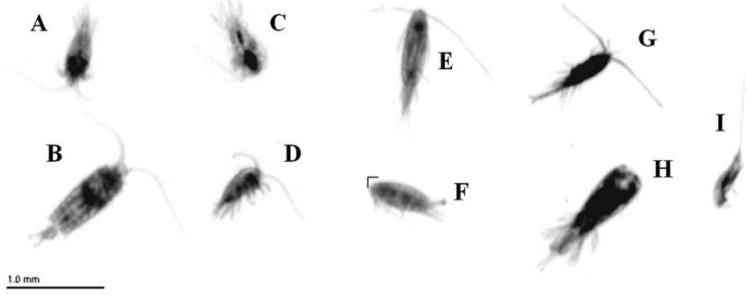
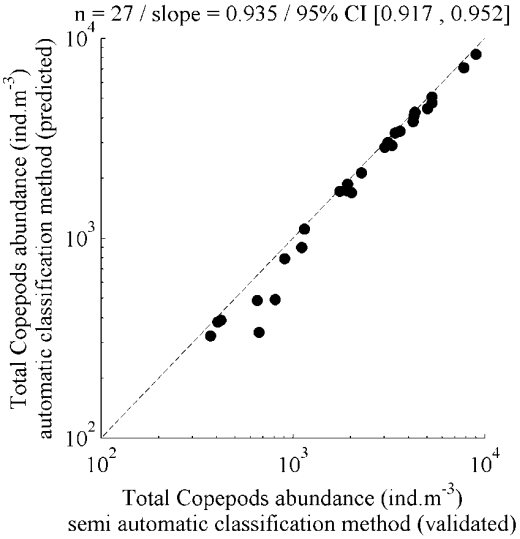


Figure 3



ACCEPTED MANUSCRIPT

Figure 4

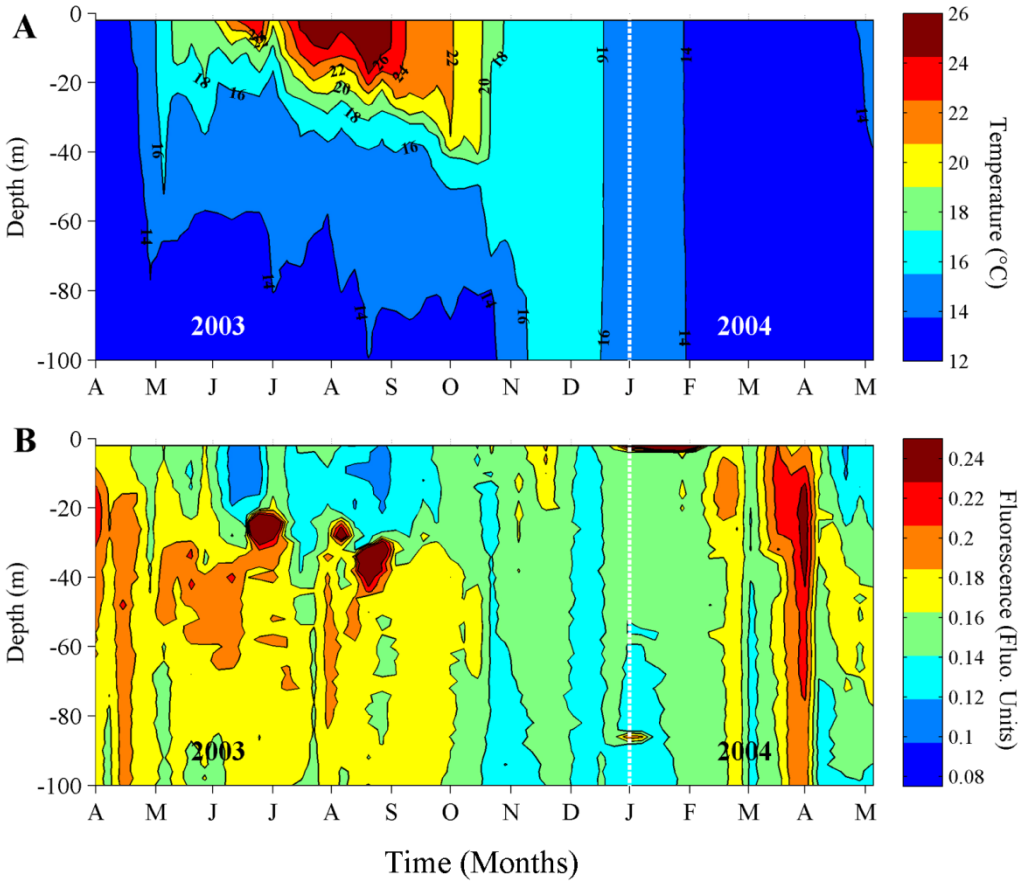
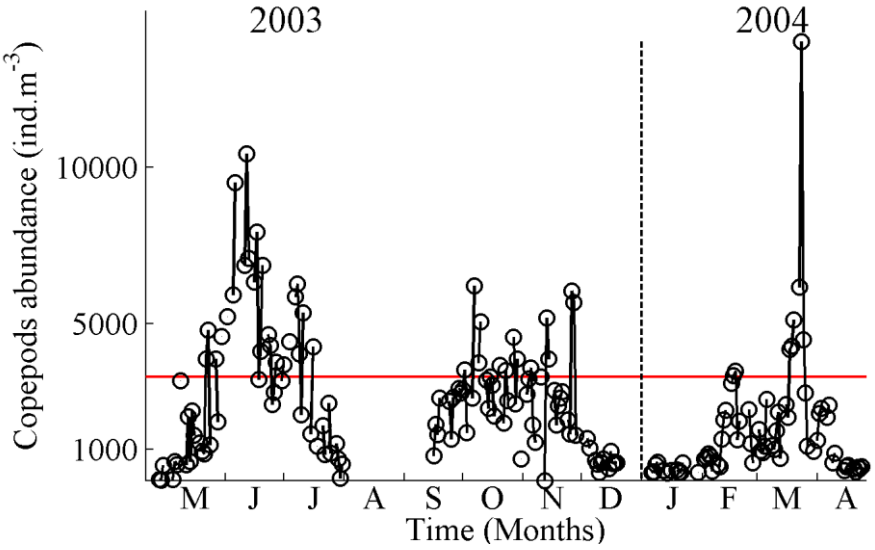


Figure 5



ACCEPTED MANUSCRIPT

Figure 6

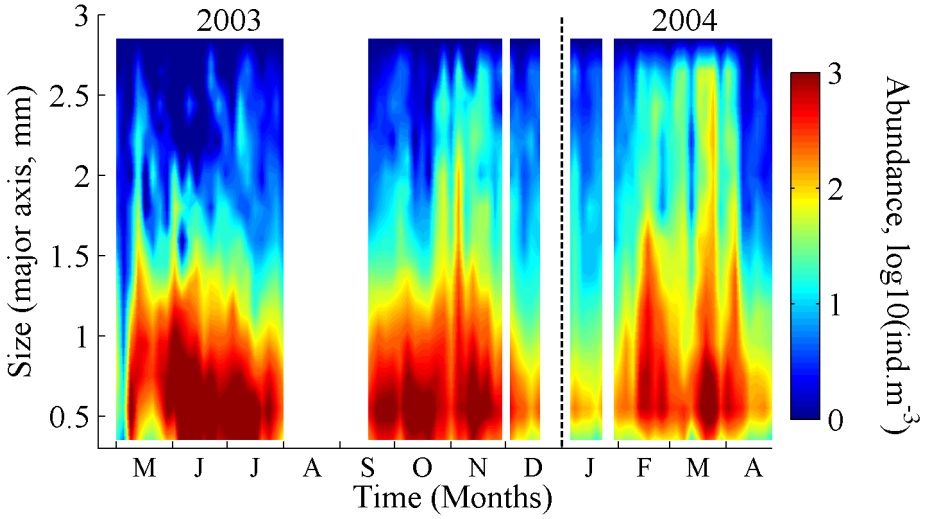
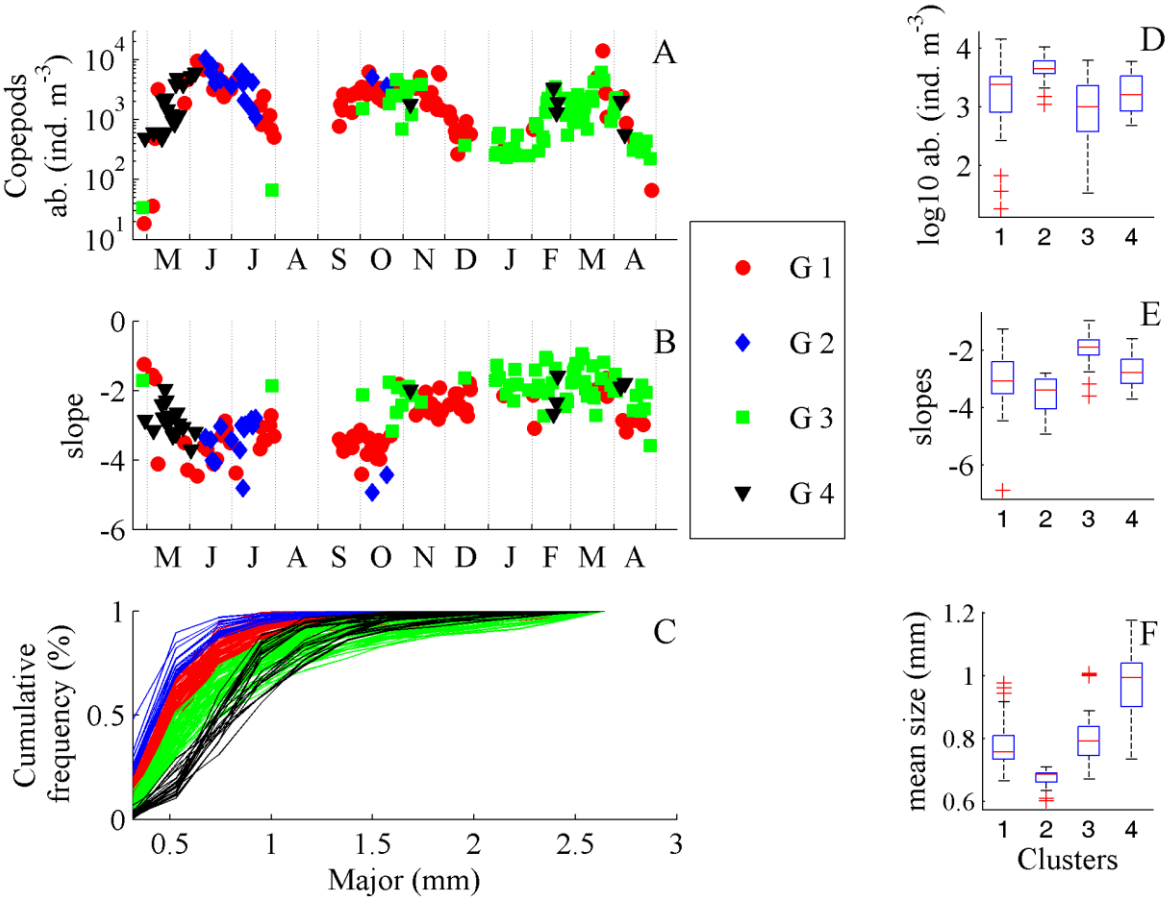
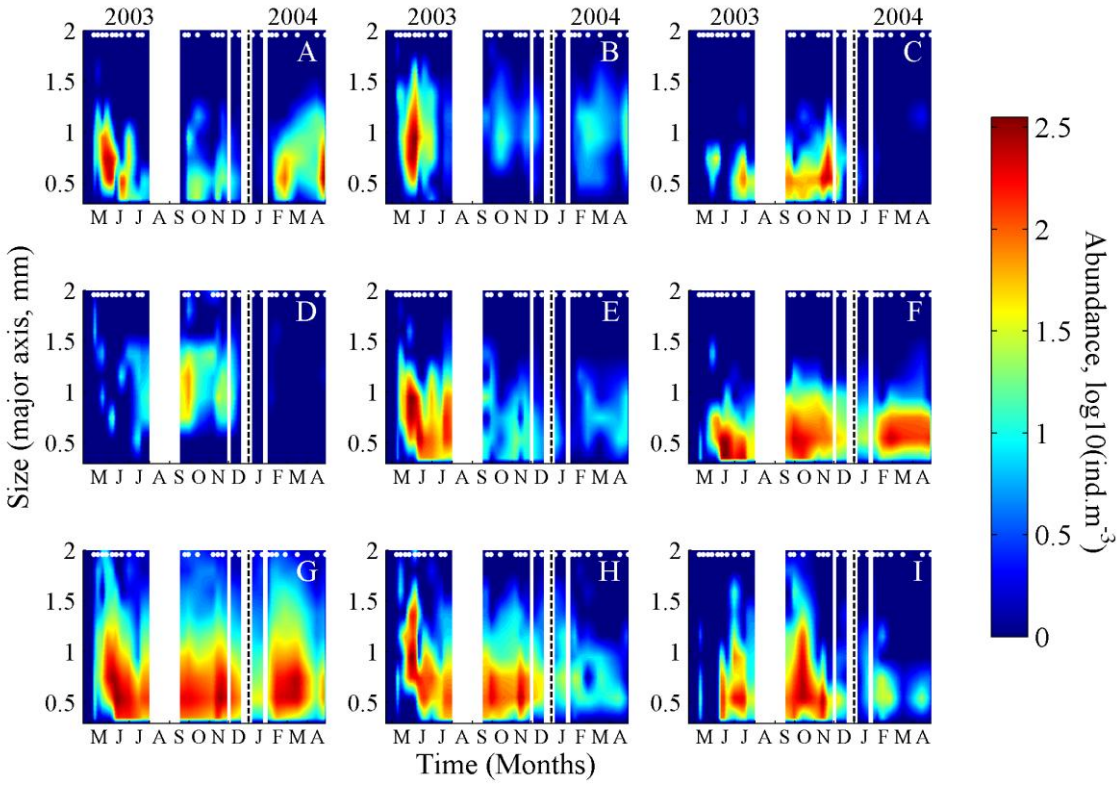


Figure 7



ACCEPTED MANUSCRIPT

Figure 8



ACCEPTED