

Sub-kilometer length scales in coastal waters

Shelley M. Blackwell^{a,*}, Mark A. Moline^a, Andrew Schaffner^b,
Thomas Garrison^b, Grace Chang^c

^a*Biological Sciences Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA*

^b*Statistics Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA*

^c*Ocean Physics Lab, University of California Santa Barbara, Santa Barbara, CA 93107, USA*

Abstract

Patchiness or spatial variability is ubiquitous in marine systems. With increasing anthropogenic impacts to coastal resources and coastal systems being disproportionately large contributors to ocean productivity, identifying the spatial scales of this patchiness, particularly in coastal waters, is of critical importance to understand coastal ecosystem dynamics. The current work focuses on fine scale structure in three coastal regions. More specifically, we utilize variogram analyses to identify sub-kilometer scales of variability in biological and physical parameters measured by an autonomous underwater vehicle (AUV) in the Mid-Atlantic Bight, Monterey Bay, and in San Luis Obispo Bay between 2001 and 2004. Critical scales of variability in density, turbidity, fluorescence, and bioluminescence are examined as a function of depth and distance offshore. Furthermore, the effects of undersampling are assessed using predictive error analysis. Results indicate the presence of scales of variability ranging from 10s to 100s of meters and provide valuable insight for sampling design and resource allocation for future studies.

Keywords: . Phytoplankton; AUV; Spatial scales; Patchiness; Coastal oceanography; Sampling; LEO-15; Mid-Atlantic Bight; Monterey Bay; San Luis Obispo Bay

1. Introduction

Variability or patchiness of physical and biological parameters is a ubiquitous characteristic across the world's oceans. Considerable attention has been given to determine scales of biological and physical oceanographic processes and to identify the drivers of plankton patchiness. These questions have been addressed using a variety of different methods, including spectral analysis (Denman and Platt, 1976; Losee et al., 1989; Lovejoy et al., 2001; Washburn et al., 1998; Wiebe et al., 1996), autocorrelation or autocovariance functions (Chang et al., 2002; Mackas, 1984; Yu et al., 2002) correlograms or variograms (Dustan and Pinckney, 1989; Mackas, 1984; Yoder et al., 1987), wavelet analysis (Deutschman et al., 1993; Machu et al., 1999; Charria et al., 2003), and multifractal analysis (Seuront et al., 1996, 1999). Better understanding of the scales of variability can be used to help identify the physical

and biological processes structuring biomass distribution and community structure, and help distinguish which processes are responsible at different scales (Chang et al., 2002; Cunningham et al., 2003; Dustan and Pinckney, 1989; Seliger et al., 1981). Previous research addressing these questions in the open ocean has examined scales of variability of current patterns (Flagg and Kim, 1998), of hydrography in the Faroes frontal region (Aranuvachapun et al., 1997), and of the spatial distribution of zooplankton on Georges Bank (Wiebe et al., 1996). Washburn et al. (1998) found through coherence analysis of the scales of chlorophyll fluorescence (FL), salinity, and attenuation at 490 nm in the North Atlantic that phytoplankton distributions at larger scales (horizontal wavelengths > 7 km) were controlled by eddy advection and that non-conservative processes of phytoplankton growth, grazing, and sinking controlled distributions at smaller scales.

In addition, scale information can be used to more effectively design oceanographic sampling regimes (Bellingham and Zhang, 2005), and is necessary for the refining of models of plankton dynamics. A study on the southern British Columbia shelf examined scales of phytoplankton and zooplankton community composition and phytoplankton and zooplankton biomass in cross-shelf and alongshore directions (Mackas, 1984). Findings from this study revealed that higher resolution sampling was required when sampling across bathymetric contours than when sampling parallel to the bathymetric contours and that higher resolution sampling regimes were required to resolve community composition than biomass patterns. Although these results are spatially and temporally specific, they have important implications for the allocation of resources during future studies.

Mackas et al. (1985) defined the observed plankton patchiness of the oceans as variability on scales from 10 m to 100 km in the horizontal dimension and vertically from 0.1 to 50 m. In the 1970s and 1980s a number of studies focused on biomass patchiness in coastal waters (Denman and Powell, 1984; Haury et al., 1978; Mackas et al., 1985). However, due to sampling constraints in terms of resolution, lack of synoptic or semi-synoptic measurements and spatial and temporal aliasing these studies focused on larger scale variability within coastal regions. More recently studies by Bissett et al. (2004), Cunningham et al. (2003), Yu et al. (2002), and Chang et al. (2002) had higher resolution and were nearer to synoptic measurements, and were thus able to look at variability on smaller spatial scales. As coastal waters make up only a fraction of the world's total oceans, yet represent up to 30% of the ocean's productivity and more than 90% of global fish catch (Holligan and Reiners, 1992), our ability to understand plankton variability and the drivers of that variability across the entire range of spatial scales is of critical importance. Although sampling capabilities have improved dramatically over the last decade, variability at the smaller end of the defined range remains to be examined.

In an effort to fill this gap in our current understanding of scales of variability this study focuses on sub-kilometer scales of variability in coastal waters. To this end continuous physical, optical, and biological data were collected at three locations by two different autonomous underwater vehicles (AUVs). This study uses variogram analysis of high resolution observations to identify spatial scales of variability and is broken down into four main analyses. First, an examination is made of scales of variability along a single transect at multiple depths. Second, comparisons are made between scales identified at one depth versus the entire water column or a section thereof. Third, comparisons are made of scales identified as a function of distance offshore. Finally, predictive error analysis calculates the error associated with data collected at degraded resolutions, effectively providing a means to assess the impact of undersampling.

2. Methods

2.1. Instrumentation

Data for this study were collected during 19 deployments by two Remote Environmental Monitoring Units (REMUS) AUVs. The REMUS is a propeller-driven AUV equipped with upward and downward looking RD Instruments acoustic Doppler current profilers, and an Ocean Sensor OS-200 CTD (Moline et al., 2005). The remaining payloads of the vehicles are customizable depending on the scientific application. The first vehicle, Boomerang I (BI), was equipped with a marine bioluminescence (BL) bathyphotometer (Herren et al., 2005), that had Seapoint fluorometer and turbidity (TB) sensors incorporated into the interface section (Fig. 1(A); Moline et al., 2005). The second vehicle, Boomerang II (BII), carried multi-spectral Satlantic OCR-507R and 507I radiance and irradiance sensors and a Wetlabs ECO FL/TB sensor (Fig. 1(B); Moline et al., 2007).

2.2. Data collection

Nineteen AUV deployments took place over the course of 3 years in both the Atlantic and Pacific Oceans. As BL in marine organisms varies diurnally, all deployments with BI took place after 21:00 local time to ensure maximal BL signal (Moline et al., 2001). The first deployment took place at the Long term Ecosystem Observatory (LEO-15) region off the coast of New Jersey on 7/27/01 (Fig. 1(C)). During this deployment BI ran repeatedly along a 1 km transect at 1 m depth intervals from 2 to 12 m. BL, TB, and FL data were collected at a rate of 1 Hz. CTD data were recorded at 0.25 Hz.

The next set of deployments took place in Monterey Bay, CA (Fig. 1(D)) between 8/20/2002 and 8/25/2002 (4 deployments) and 8/9/2003 and 8/18/2003 (9 deployments). During the 2002 deployments BI ran 20 km offshore along transect A (Fig. 1(D)) undulating between 3 and 40 m depth at 8° and returned at a constant depth of 15 m on 8/20/02 and at 20 m on 8/22/2002, 8/24/2002, and 8/25/2002. During 2003, the vehicle ran offshore along transect A, and returned onshore along transect E, undulating between and 3 and 40 m. During the 2002 and 2003 deployments when the bottom depth was less than 40 m, the vehicle undulated between 3 m depth and 3 m altitude off the bottom. During the 2002 and 2003 deployments biooptical and CTD data were merged by time stamp at 1 Hz, which yielded a linear sampling resolution of ~2 m.

The final set of deployments took place in San Luis Obispo Bay, California (Fig. 1(E)) between 6/4/2004 and 6/16/2004 (5 deployments). During these deployments, BII ran offshore on the T1 line at 1 m depth and returned along the same line, undulating between 1 and 40 m depth or 4 m altitude when depths were less than 44 m. During these

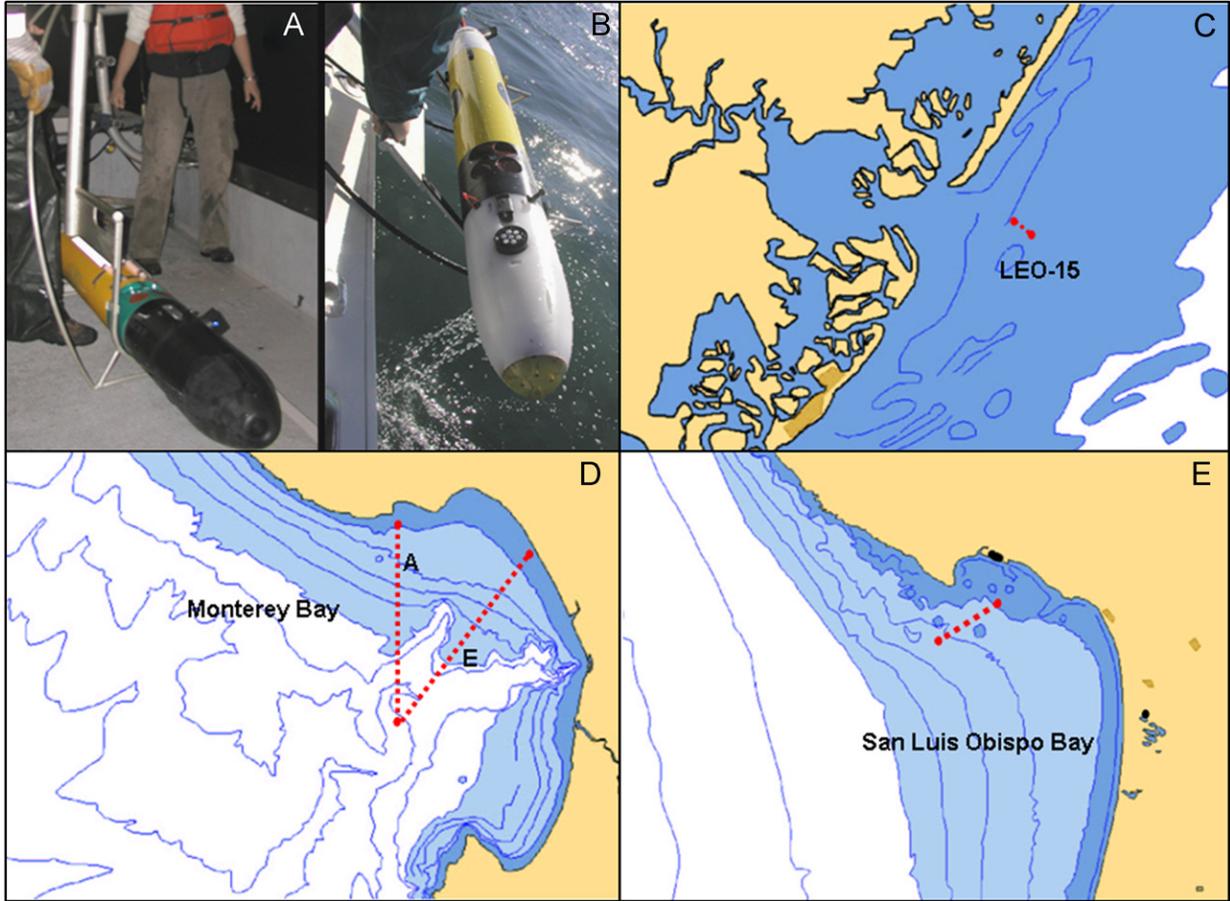


Fig. 1. (A) The REMUS autonomous underwater vehicle, Boomerang I, with bioluminescence bathyphotometer, Seapoint fluorometer, and turbidity sensors incorporated into the front end of the vehicle during deployments in Monterey Bay. (B) Boomerang II with Wetlabs ECO fluorescence/turbidity sensor and Satlantic radiance and irradiance sensors situated posterior to the standard REMUS nose section. Geographic locations and survey lines during deployments in (C) LEO-15, New Jersey, (D) Monterey Bay, CA, and (E) San Luis Obispo Bay, CA.

deployments, all data were merged at 1 Hz, giving a sampling resolution of 2 m.

2.3. Data processing and statistical analyses

Data processing and statistical analysis involved two major steps, first of which was the removal of outliers. Outliers were defined as observations whose standard deviation was greater than three times that of the 11 surrounding observations, these observations were removed from the data set and not included in the analysis (Yu et al., 2002). Outlying data made up less than 4% of the data on average and never more than 10% of the data. The second step involved either one or both of the following: variogram analysis and/or predictive error estimation, both of which are described in more detail below.

First developed by mining engineers in their studies of mineral deposits, the empirical variogram or semi-variogram provides a description of how data are correlated with distance (Huijbregts, 1975; Journel and Huijbregts, 1978; Jumars, 1978; Yost et al., 1982). The semi-variogram

function, $\hat{\gamma}(h)$, was originally defined by Matheron (1963) as half the average squared distance between points separated by a distance h and is computed as follows:

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} (z_i - z_j)^2, \quad (1)$$

where $N(h)$ is the set of all pairwise Euclidean distances $i-j = h$, $|N(h)|$ is the number of distinct pairs in $N(h)$, and z_i and z_j are data values at spatial locations i and j , respectively (Dustan and Pinckney, 1989; Kaluzny et al., 1998).

Intrinsic stationarity implies a process with a constant mean and with a variance defined only through the magnitude of the distance h . (Kaluzny et al., 1998). Preliminary analysis assessed violations of this assumption and showed that the data exhibited large scale trend. In order to satisfy the assumption of stationarity, trend was removed from the data (Legendre, 1993; Washburn et al., 1998). To accomplish this, a generalized additive model (GAM) using a linear loess smoothing function with a span of 1 km was applied to remove trend on kilometer or

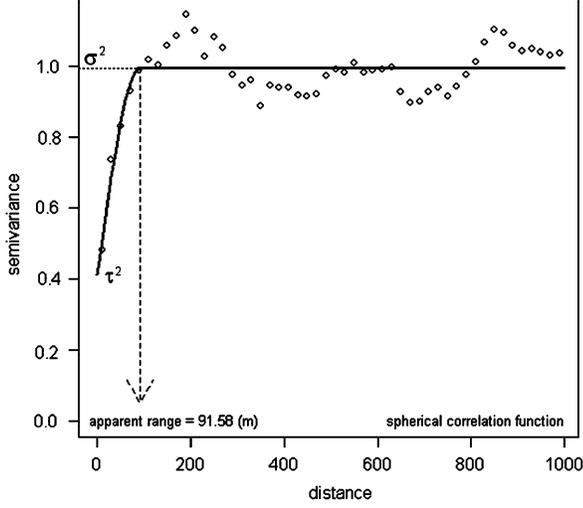


Fig. 2. Variogram of fluorescence collected at 1 m depth on June 7, 2004 in SLO Bay, CA, σ denotes the sill value, v is the nugget, and r represents the range of the variogram.

greater scales (see conclusions for further discussion on the choice of the GAM model). Subsequent variogram analysis on the underlying process was then conducted on the residuals obtained from the GAM.

In order to describe the spatial correlation structure objectively, a theoretical variogram was used to model the empirical variogram (Fig. 2). The theoretical variogram is described by several parameters, the *nugget* τ^2 , the *sill* σ^2 , and the *range* ϕ , and can be formulated as:

$$\gamma(h) = \tau^2 + \sigma^2[1 - \rho(h)], \quad (2)$$

where $\rho(h)$ is the correlation function described later. The nugget effect has a dual interpretation as either the measurement error or spatial variation on a scale smaller than the smallest distance between any two points (Diggle and Ribeiro, 2007). The sill is the value of the variogram semi-variance as h increases to infinity and represents the variability over the field of data. The range is the distance (if any) at which the data are no longer correlated.

The six different correlation functions used to model the theoretical variogram from the empirical data have characteristic differences in their behavior for small distances, $h \approx 0$ and how quickly $\rho(h)$ approaches the sill with increasing distances h . The six different correlation functions that were considered are

$$\text{exponential : } \rho(h) = \exp\left(-\frac{h}{\phi}\right), \quad (3)$$

$$\text{Gaussian : } \rho(h) = \exp\left[\left(-\frac{h}{\phi}\right)^2\right], \quad (4)$$

$$\text{spherical : } \rho(h) = \begin{cases} 1 - 1.5\left(\frac{h}{\phi}\right) + 0.5\left(\frac{h}{\phi}\right)^3 & \text{if } h < \phi, \\ \sigma^2 & \text{otherwise,} \end{cases} \quad (5)$$

$$\text{cubic : } \rho(h) = \begin{cases} 7\left(\frac{h}{\phi}\right)^2 - 8.75\left(\frac{h}{\phi}\right)^3 + 3.5\left(\frac{h}{\phi}\right)^5 \\ = 0.75\left(\frac{h}{\phi}\right)^7 & \text{if } h < \phi, \\ \sigma^2 & \text{otherwise,} \end{cases} \quad (6)$$

$$\text{circular : } \rho(h) = \begin{cases} 1 - 2\frac{\theta\sqrt{1-\theta^2} + \sin^{-1}\theta}{\pi} \\ \text{where } \theta = \min\left(\frac{h}{\phi}, 1\right), \\ \sigma^2 & \text{otherwise,} \end{cases} \quad (7)$$

$$\text{wave : } \rho(h) = \left(\frac{\phi}{h}\right) \sin\left(\frac{h}{\phi}\right). \quad (8)$$

Note that the exponential and Gaussian functions approach the sill asymptotically, whereas the wave function's amplitude decreases around the sill. When $\rho(h)$ approaches the sill in this manner, as h increases, the range is undefined. In order to maintain consistency in the range defined by each of the different models, this study instead calculates the apparent range. The apparent range is defined as the distance h_0 at which $\rho(h_0) = 0.05$ or equivalently when $\gamma(h_0) = \tau^2 + 0.95\sigma^2$ (Diggle and Ribeiro, 2007). For the spherical, cubic, and circular functions, the range parameter is the apparent range.

Theoretical variograms were fit to the six candidate models and the best was chosen based on the minimum loss value determined by a weighted least squares (WLS) algorithm that minimizes the difference between the empirical variogram and the theoretical variogram. The loss function used is the sum of the number of pairs of observations at a distance h times the squared difference between the values of the empirical and theoretical variogram at that distance:

$$LOSS(\theta) = \sum_{N(h)} N(h)[\hat{\gamma}(h) - \gamma(\theta)]^2. \quad (9)$$

Once the best model was identified, the apparent range was computed.

Statistical analyses were carried out using R and the GeoR statistical package. The apparent range of the theoretical variogram as defined by the above analyses will, from here onwards, be referred to as the critical scale of variability (CSV) and are specifically length scales identified on a sub-kilometer level.

To quantify the loss of spatial variance in oceanographic measurements when high resolution measurements are not made, predictive error estimations were made. To this end, AUV observations were subsequently re-sampled at degraded resolutions ranging from 30 m to 1.8 km at intervals of 40 m. Sub-sampled data sets were then used to estimate observations at the originally sampled resolution of 2 m employing a simple linear interpolation for estimations. Percent errors are reported as the percent error of the observed range (PER) of any particular

parameter according to the equation:

$$PER = 100 \times \frac{\sum_{N(i,j)} |x_e - x_{i,j}| / (\max x_{i,j} - \min x_{i,j})}{N(x_{i,j})}. \quad (10)$$

Actual error (AE) was calculated according to the following equation:

$$AE = \frac{\sum_{N(i,j)} |x_e - x_{i,j}|}{N(x_{i,j})}, \quad (11)$$

where x_e is the interpolated estimate, x_o the observed value at the location i, j , and $N(x_{i,j})$ is the total number of observations. Predictive error analyses were made from data collected in Monterey Bay during 2003 and in San Luis Obispo Bay during 2004.

3. Results and discussion

3.1. Critical scales of variability as a function of depth

Results from the 1 km transect data collected in the LEO-15 study region on July 27, 2001 demonstrated a high degree of variability in the scales of FL identified at different depths, ranging from a minimum CSV of 23 m at 12 m depth to a maximum of 170 m at 8 m depth (Fig. 3). For consistency, CSVs in Fig. 3 were calculated using the spherical model because it was the most common best fit among the different depths. When data were analyzed allowing the model to vary depending on the best fit and loss function the results did not deviate significantly. The largest CSVs were located above and below the FL maxima, with the largest layered on top of the pycnocline. The smallest CSV occurred within the bottom 2 m of the

water column in what appears to be a bottom nepheloid layer (TB data not shown; Boss et al., 2004). The transition zone just above the peak biomass was also characterized by a small CSV. The layering of peak biomass on top of the pycnocline is consistent with findings that showed peak chlorophyll biomass in the LEO-15 region that tracked the maximum density gradient (Moline et al., 2004). During summer months the LEO-15 region is characterized by alternating upwelling and downwelling conditions, which vary depending on the intensity and duration of forcing events (Glenn et al., 2004). The distribution of phytoplankton depends considerably upon these forcing events, with phytoplankton biomass being concentrated on top of the pycnocline during periods of episodic upwelling and exhibiting a more even distribution throughout the water column when the pycnocline is deeper during downwelling conditions (Moline et al., 2004).

Our deployment occurred during highly stratified conditions typical of the early onset of an upwelling event (Munchow and Chant, 2000), where physical and biological parameters were distributed in a near horizontal plane. This condition is conducive to identify horizontal CSVs as a function of depth, however as upwelling progresses the physical environment will tend toward baroclinic conditions (Rao et al., 2005; Hamilton and Rattray, 1978; O'Brien and Hurlburt, 1972), with vertical displacement of isopycnals variable relative to horizontal distance offshore (Moline et al., 2004; Glenn et al., 2004).

Depending on the environmental conditions of interest, it is likely that length scales determined by data collected at only one depth (or a series of depths) will not adequately provide the information needed to characterize the inherent spatial variability of the system. The order of magnitude difference between the scales identified near the bottom

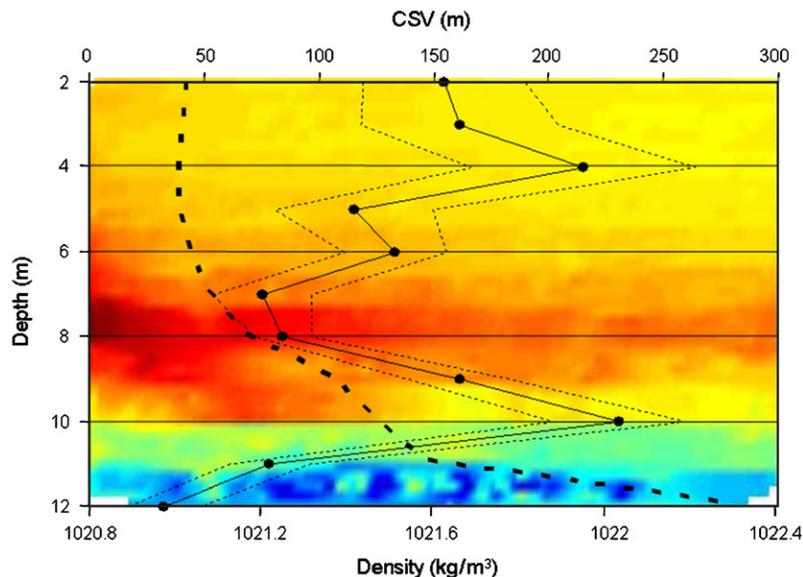


Fig. 3. Fluorescence collected along a 1 km transect in the LEO-15 study region between 2 and 12 m at 1 m depth intervals, data are shown in relative fluorescence units with red indicative of high levels of fluorescence and blue indicative of lower fluorescence. Overlaid are the mean densities for each depth along the transect (dashed line), CSVs of fluorescence (filled black circles), and the standard error for the CSVs (dotted line).

and just above the pycnocline identifies the problems associated with the traditional coarsely spaced vertical sampling nearshore. This variability among depths makes it difficult to determine the optimal horizontal sampling resolution, and suggests the need for alternative approaches if one is interested in determining optimal sampling resolution for a semi-synoptic view of a slice or volume of water. Variability also provides valuable information on designing sampling strategies for characterizing transition zones, suggesting that increased resolution is necessary for characterization of these highly dynamic regions. The characterization of thin layers, for example, requires greater resolution in both the vertical and horizontal dimensions (Dekshenieks et al., 2001).

3.2. Sampling critical scales of variability in coastal regions

In an effort to develop a means to determine CSVs for a volume of water, comparisons were made between the CSVs estimated from AUV transect data collected in undulating modes versus those collected at a single depth. Observations of each mode were compared for the same horizontal transects in both Monterey and SLO Bay (Fig. 4). Paired *t*-tests of the CSV identified by variogram analysis of undulating versus single-depth modes showed there was a significant difference for all parameters in SLO Bay ($p < 0.05$, $n = 5$) and for density ($p < 0.05$, $n = 4$) and FL ($p < 0.05$, $n = 4$) in Monterey Bay. Data were paired per survey transect to ensure that comparisons of undulating versus one depth mode were made within the same volume of water. The CSVs of undulating data from SLO Bay were 120 m (± 10 m) shorter on average than the CSVs of one-depth data; in Monterey Bay the CSVs of undulating data were on average 70 m (± 20 m) shorter than one-depth data. The average CSV for undulating and one-depth data ranged from 50 to 190 m, depending on the location and parameter measured.

Significantly shorter CSVs of undulating versus one-depth data in SLO Bay, suggest that there was significant vertical variability in the study area and more variability at depth versus 1 m. During spring months, SLO Bay is a region of intense upwelling, with cycles of phytoplankton

accumulation/growth and senescence lasting approximately 2 weeks on average (M. Tognazzini, personal communication). Evidence of phytoplankton accumulation and growth and stratified conditions is especially apparent on 6/04/2004 and then again on 6/14/2004 and 6/16/2004 (data not shown). This was likely a key factor in contributing to the vertical variability and difference between one-depth and undulating data in SLO Bay. Monterey Bay is also a region characterized by strong spring and summer upwelling. During the time of this study (August 2001) the upwelling centers were at least 100 km away from the AUV deployment region (Shulman et al., 2005), nevertheless significant vertical variability existed in both density and FL. Interestingly there was no significant difference between one-depth and undulating CSVs for TB.

Undulating data have been treated in a variety of manners when used to examine the question of scales. In many cases observations were averaged per undulation or half undulation (Chang et al., 2002; Ondercin et al., 1995; Washburn et al., 1998). During a study in the North Atlantic by Washburn et al. (1998), data were collected by a towed undulating platform. The horizontal distance traveled per undulation was 1.1 km, undulating between 10 and 140 m depth and the ascent/descent angle was 13° . Observations in each half of the undulation were averaged above the mixed layer, which resulted in an average horizontal resolution of 550 m. This methodology results in reduced horizontal resolution and a smoothing of the vertical variability, especially when the rate of vertical ascent/descent is slow and significant horizontal distance is traveled with each undulation. However, in the case of an open ocean study, reduced resolution does not pose the problem that it does in studies of coastal waters. In contrast, the undulating data collected in this coastal study maintained measurements in the vertical and horizontal dimensions of 0.18 and 2 m resolutions, respectively.

In a different study (Yu et al., 2002), a Laplacian gridding algorithm was applied in order to interpolate the data onto $x-z$ grids; data were subsequently vertically averaged for the upper and lower halves of the water column, leading to two single horizontal transects of the

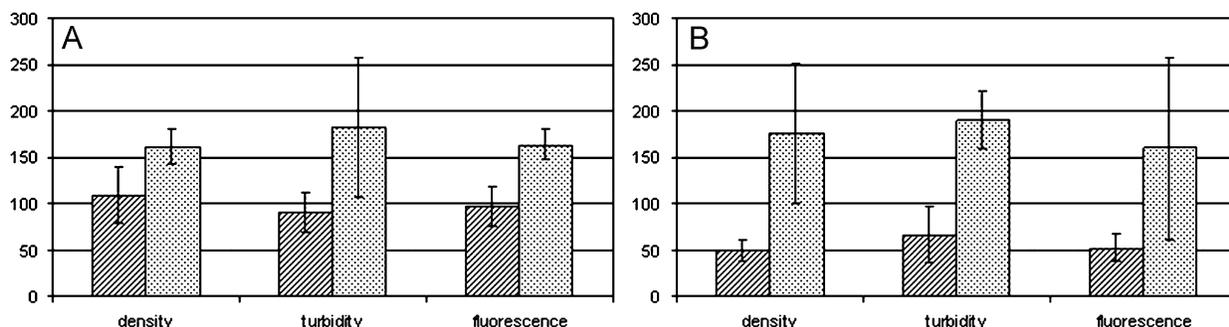


Fig. 4. CSVs for parameters measured in (A) Monterey Bay, August 2001 and in (B) SLO Bay, June 2004. Dotted bars represent data collected at one depth while hashed bars represent data collected in an undulating pattern.

water column. This approach has the benefit of efficiency by employing an undulating survey strategy while at the same time maintaining highly resolved depth and distance information. However, interpolation often introduces artifacts into the data set, which are not actually characteristics of the original measurements and can cause errors in analysis outcomes. Furthermore it reduces the depth information from a continuum into two discrete depths.

The observed difference between undulating and single-depth sampling modes illustrates that undulating data, particularly at times of high vertical variability, may lead to smaller CSVs than single-depth data. Moreover these results highlight how a system's inherent variability can be over- or underestimated by sampling at only one depth and underscore how analysis of undulating data provides a better estimation of spacing requirements for traditional shipboard profiling sampling regimes. The ability to maintain resolution and incorporate both vertical and horizontal dimensions into the length scale estimates represents a valuable and efficient approach for identifying CSVs.

3.3. Critical scales of variability as a function of distance offshore

Previous work has shown that physical and biological spatial and temporal variabilities increase with increasing proximity to shore (Bissett et al., 2004; Chang et al., 2002; Lovejoy et al., 2001; Yoder et al., 1987). This phenomenon has been referred to often in Stommel diagrams (Stommel, 1963; Lovejoy et al., 2001; Dickey, 1991) and on broad scales, has been accepted as common knowledge. However, there are relatively few studies closer to shore that have addressed this same question on the smaller scales. One such study examining spatial scales by Bissett et al. (2004) showed the optimal sampling distance decreased by as much as two orders of magnitude as one moved from 20 km offshore to within 5 km of shore in the remotely sensed ocean color PHILLS II image. This study determined that at distances offshore greater than ~20 km the optimal sampling distance approached 5 km and that within 5 km the optimal sampling distance ranged 50–200 m.

In order to determine if this is a consistent trend in the very near shore, if the trend is observed in other parameters and to better design sampling strategies for coastal waters, the next section of this study examined how distance offshore related to the CSV. More specifically, the five 5 km transects from SLO Bay were broken down into four 1.25 km sections. Data were detrended according to the methodology described in Section 2.3 before breaking down each of the five transects into sub-sections. Subsequent to splitting the transect, variogram analysis was carried out on each of the four 1.25 km sections for observations of density, TB, and chlorophyll a FL. Density, TB, and FL data from the 2003 deployments in

Monterey Bay were broken down into four 5 km sections and analyzed likewise.

Results of this analysis demonstrated a trend of increased CSV with increased distance from shore in TB and FL in SLO Bay. The greatest difference in magnitude was observed in TB. The increase was not a linear trend, though the most noticeable increase corresponded to the most offshore section of the transect which was located outside the breakwall of the bay (Fig. 1(E)). CSVs for the onshore section were significantly shorter than those of the most offshore section in TB and FL (Tukey's pairwise comparison $p < 0.05$). The CSVs of density of the two nearshore sections were smaller than the offshore sections but were only significant with a 90% confidence interval. CSVs of TB for the offshore section were significantly greater than the first three sections (Tukey's pairwise comparison $p < 0.05$) and the CSVs of FL for the first two sections were significantly smaller those for the offshore two sections (Tukey's pairwise comparison $p < 0.05$). The CSVs for density, TB, and FL from the near shore section ranged from ~30 to ~50 m, with the CSV of the most offshore section being at least one-third greater than the near shore section, ranging from ~60 to ~80 m (Fig. 5). These results are comparable to the optimal sampling distance of ocean color parameters identified by Bissett et al. (2004) for 50–200 m within 5 km of shore.

Measurements of density made at 1 m depth along the transect in SLO Bay on 4/2004 in SLO Bay did not arrive at a range in the variogram analysis in the two furthest offshore 1.25 km sections of the transect, indicating there were no characteristic scales of variability in density on a sub-kilometer scale. It is likely that characteristic scales of variability in density on these days were greater than 1 km and in order to identify these scales the maximum distance of the variogram would need to be extended.

Interestingly, density and TB data from Monterey Bay did not show this same increase with distance offshore (Figs. 5(D) and (E)). CSVs ranged from ~50 to ~90 m and exhibited no characteristic pattern or trend. There was however a significant difference in CSVs of FL in which the most onshore section was significantly shorter than the offshore sections of the transect (Tukey's pairwise comparison, $p < 0.05$). One possible explanation as to why we did not observe the same increased trend in Monterey Bay was because the entire transect was located within the bay and the dynamics and circulation patterns within the bay are known to be highly variable, particularly during times of upwelling (Drake et al., 2005). The 9 deployments in Monterey Bay coincided with the transition from a strong upwelling event to gradual relaxation (Shulman et al., 2005).

3.4. Predictive error analysis—effects of undersampling

Based on the above results it is apparent that there is a high degree of sub-kilometer scale variability in near shore coastal regions. However, to what degree is this variability

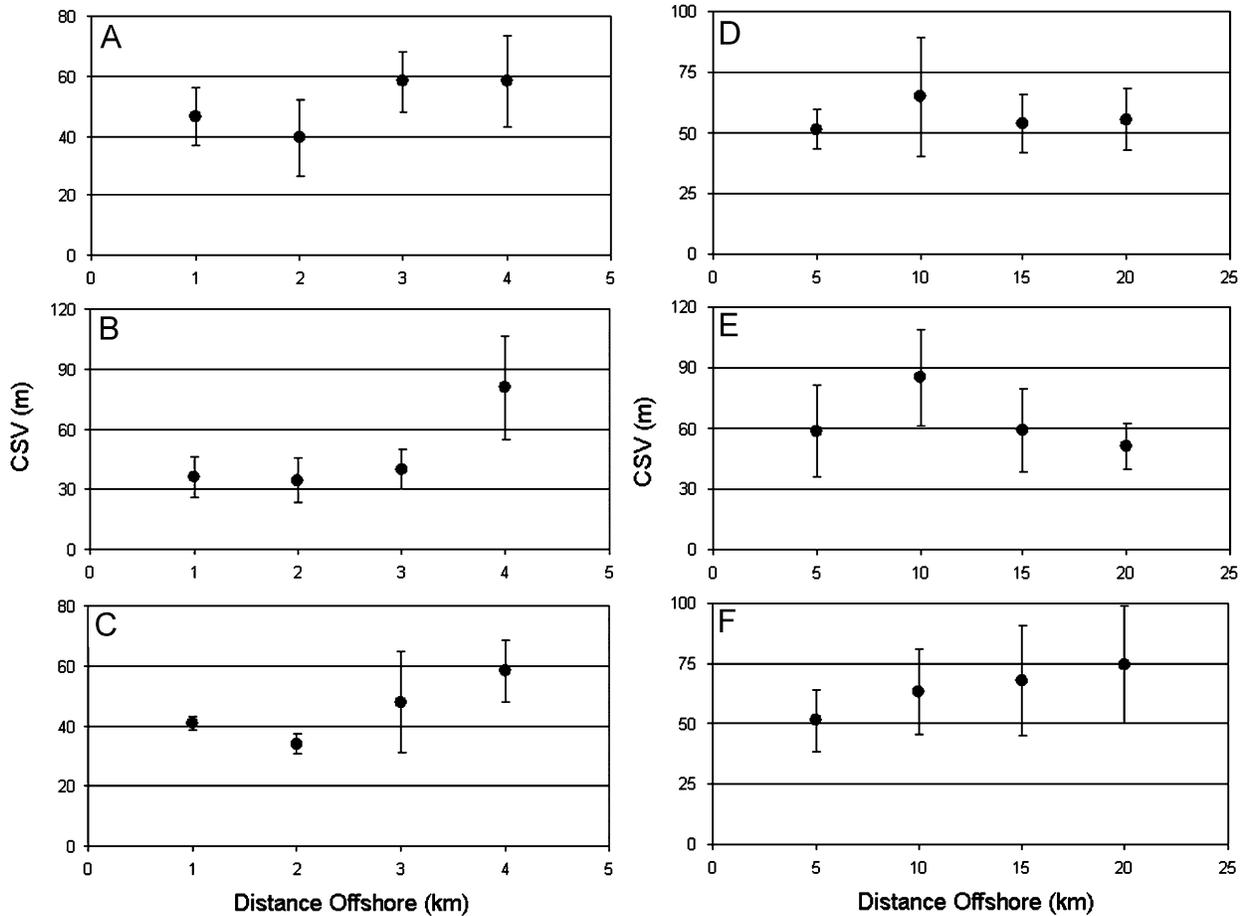


Fig. 5. CSVs in SLO Bay as a function of distance offshore for (A) density, (B) turbidity, and (C) chlorophyll a fluorescence and in Monterey Bay for (D) density, (E) turbidity, and (F) chlorophyll a fluorescence.

important in estimating nearshore dynamics? More specifically how can this knowledge be integrated in the future to more effectively and efficiently design oceanographic sampling strategies? Predictive error analysis, as described in Section 2.3, addressed these questions by calculating the PER and *AE* involved in linearly interpolating subsequently lower and lower resolution data. Maximum and minimum sampling resolutions for all parameters in SLO Bay and Monterey Bay were 30 and 1800 m, respectively. As would be expected, in both locations decreased resolution resulted in an increase in predictive error (Figs. 6 and 7). In SLO Bay the predictive error for density was 3% at the maximum resolution and increased to 17% at ~400 m resolution, the resolution at which the exponential curve reached the asymptote. At the asymptote the *AE* predicted was ~0.15 kg/m³ (Fig. 6(A)). Predictive error in TB ranged from 3% to 9% at ~450 m resolution with a maximum *AE* at the asymptote of 0.3RTU (Fig. 6(B)). FL ranged from 4% to 15% at a resolution of 600 m, the *AE* predicted at the asymptote was ~1.25RFU (Fig. 6(C)). In Monterey Bay the predictive error for density, BL, and TB reached their asymptotes at ~450 m. The percent errors at the asymptote, representative of the maximum estimated

error, for density, BL, and TB were 13.5%, 10%, and 10% respectively, *AEs* at the asymptote were 0.29 kg/m³, $1.0 + e^{10}$ (photons/s), and 0.055RTU, respectively. Error in FL ranged from 1.75% at 30 m resolution to 12% at 300 m resolution. The *AE* at the asymptote was ~0.37RFU.

For each of these parameters, sampling at resolutions higher than where the curve reaches the asymptote will yield more accurate estimation of parameters in regions not directly sampled. However, after the curve reaches the asymptote, there is little to no difference in the associated error. In a similar effort, Bellingham and Zhang (2005) examined how increased coverage rates of AUVs affected the error field for a specified survey duration and survey resolution. They found that increased coverage rates, achieved by adding propeller driven AUVs to the original fleet of gliders surveying a given area, decreased the error field by as much as 10–20%. Information regarding the effectiveness and efficiency of a given sampling strategy and resolution is critical in terms of deciding how to allocate resources, while undersampling results in an unacceptable level of error, oversampling wastes resources that are often difficult to obtain.

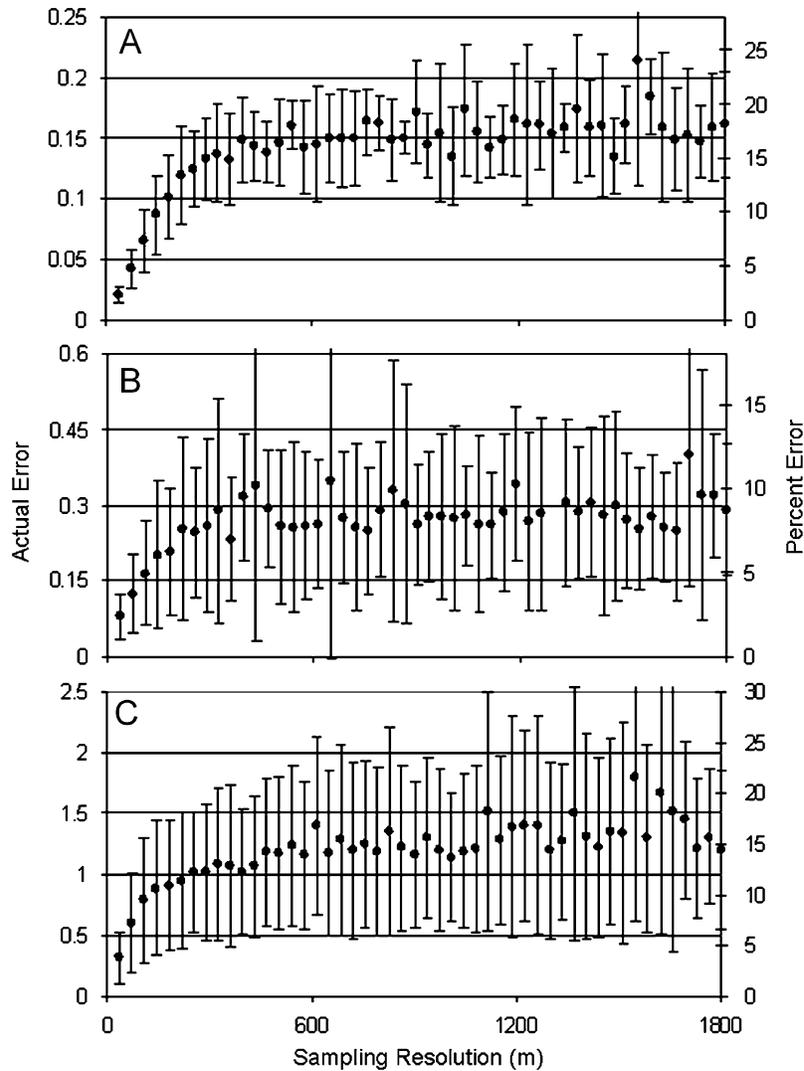


Fig. 6. Predictive error analysis as a function of sampling resolution for (A) density (kg/m^3), (B) turbidity (RTU), and (C) fluorescence (RFU), collected in SLO Bay during June 2004. Actual error in respective units represented on left-hand Y-axis, percent error on right-hand Y-axis. Black dots represent the average of five deployments with standard deviation bars in both positive and negative directions.

4. Conclusions

Autocovariance and autocorrelation functions as well as variograms and correlograms are referred to as structure functions and require the assumption of stationarity (Legendre, 1993), that is, the mean and variance of the parameter over the study area must have constant and finite values and not depend on their position in the study area. In order to satisfy this assumption, trend must be removed from the data. This study focused on sub-kilometer scales of variability and in order to accomplish this, trend was removed at the scale of 1 km. Davis (1993) states that non-stationarity may be removed by fitting regional means or by trend surface analysis. Diggle and Ribeiro (2007) discuss first and second order trend removals. There is, however, no standard for trend removal in the literature and unless very clearly specified, the

ambiguity in detrending methods creates difficulty in making meaningful interpretations of the results of scale studies. Because many processes in coastal waters operate on scales of less than 1 km such as Langmuir circulation, coastal jets, convergence and divergence zones, strong stratification layers, and biological thin layers, our focus was to examine sub-kilometer variability. This study applied a first order GAM model with a loess span of 1 km in order to remove linear trend at scales equal to and greater than 1 km. However, as with any detrending method, there is the possibility that variogram estimates may be affected.

During this study, first and second order models were considered; however, variability on sub-kilometer scales was often still masked by larger scale trends, which rendered it impossible to examine whether significant correlation structure existed on sub-kilometer scales. For

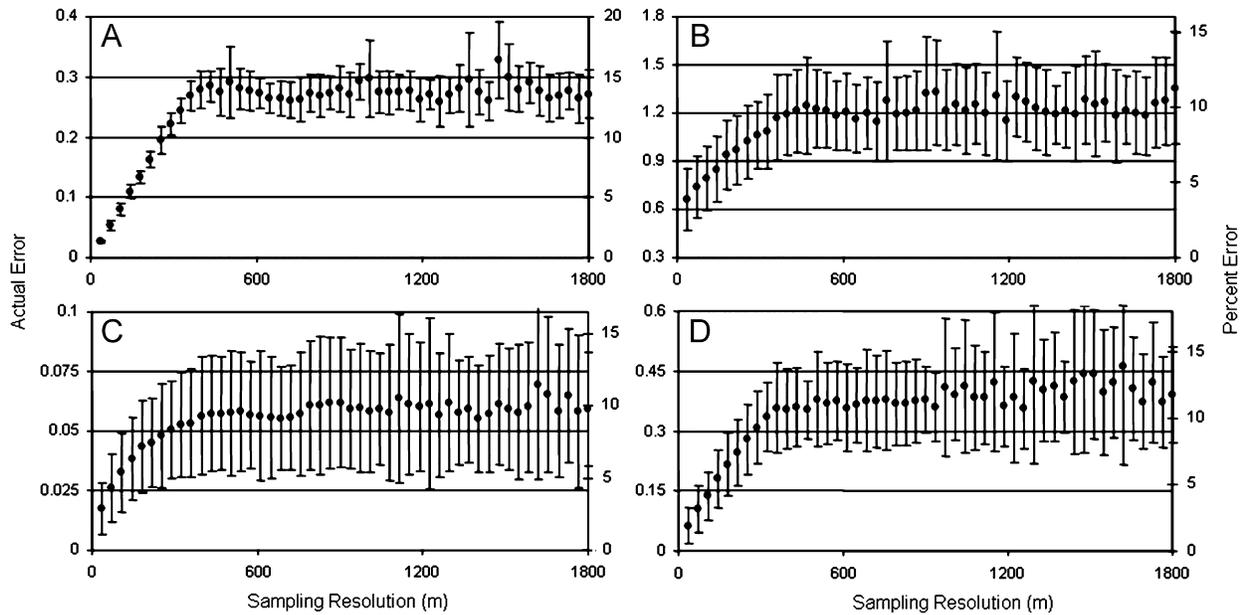


Fig. 7. Predictive error analysis as a function of sampling resolution for (A) density (kg/m^3), (B) log bioluminescence ($\text{photon L}^{-1} \text{s}^{-1}$), (C) turbidity (RTU), and (D) chlorophyll a fluorescence (RFU), collected in Monterey Bay, CA during August 2004. Actual error in respective units represented on left-hand Y-axis and percent error on right-hand Y-axis. Black dots represent the average of five deployments with standard deviation bars in both positive and negative directions.

these reasons, although it may reduce small scale variability, the application of the linear GAM model was most appropriate. The proportion of variability present at these small scales is usually less than the variability present over larger spatial scales. This is evident in both the semi-variances computed in the variograms and in the results of the predictive error analysis of this study.

Investigations were made on the effects of detrending using data from the 2003 deployments in Monterey Bay. Comparisons were made between FL, TB, salinity, temperature, and density analyzed using a 1 km span for the loess smoothing of the GAM versus a loess smooth across the entire 20 km transect. Interestingly there was less than a 20 m increase in the CSV identified when smoothing occurred across the entire transect. Although these results are specific temporally and spatially to the local conditions, they provide support for the methodology applied in this study and suggest that it is effective in identifying sub-kilometer variability when characteristic scales exist. As previously mentioned there were times when no apparent range was achieved, suggesting the absence of sub-kilometer variability with significant variability occurring at larger scales. This occurred with FL in SLO Bay and with BL in Monterey Bay on 3 days when the data was smoothed across the entire transect.

As in any given scale analysis there is the possibility of frequency aliasing, this study is no exception. However, the smallest characteristic scale was identified at 23 m. One would expect that if frequency aliasing were a problem due to variability at higher spatial frequencies than the 2 m sampling resolution, characteristic scales would have been identified between 23 and 2 m.

While structure functions are useful in determining inherent scales of variability they do not provide ample information in order to recreate a representation of the spatial structure. A review by Martin (2003) on phytoplankton patchiness suggests multi-fractals are well suited for analyzing parameters with inhomogeneous distributions marked by sharp fluctuations or intermittency, characteristics of physical and biological oceanographic parameters. While variogram analysis was sufficient for the goals of this study, multifractal analysis represents a more comprehensive methodology for examining variability and recreating spatial structure and should be examined in future studies on scale.

Ranging from tens to hundreds of meters, this study identified significant variability on a sub-kilometer level in physical, optical, and biological parameters. There was significant variability between CSVs identified at different depths for FL off the coast of New Jersey, and as a result this study presented the analysis of undulating data as a means to describe sub-kilometer variability for a slice or volume of water. In addition this study showed that CSVs were greater in the furthest offshore section of the transect for FL, TB, and density in SLO Bay, but that intermediate sections were not significantly different. With the exception of FL, which showed a significant increase in CSVs between the most nearshore section and the most offshore section of the transect, no significant trend was observed in Monterey Bay. Finally, results of the predictive error analysis showed that on average the maximum sampling error ranged between 10% and 18%, and the asymptote of the predictive error curve occurred on average between 200 and 300 m.

Keeping in mind that these results are representative of limited temporal and geographic applicability, they illustrate the presence of significant sub-kilometer variability present in near shore regions and underscore the importance of high resolution sampling. Spatial complexity in coastal regions is in itself variable, however without highly resolved data sets it is difficult to identify the processes that determine the scales of variability. As such, AUVs, gliders, and/or towed undulators are effective platforms to provide this initial high resolution information on which sampling strategies for process studies can then be based.

Acknowledgments

We would like to thank a number of people for their help with AUV deployments, Jessica Connolly, Matt Oliver, Ian Robbins, Mike Sauer, Oscar Schofield, Jeff Sevardjian, Mark Tognazzini, and the crew of the R/V *Paragon*. Thanks also to the LEO-15 2001 summer field team, the AOSN II, and Tom Moylan for logistical support. This work was supported by the Office of Naval Research (N00014-00-1-0570 and N00014-03-1-0341 to M. Moline).

References

- Aranuvachapun, S., Maskell, S.J., Griffiths, C., 1997. Spatial length scale analysis of data from observations and a model of the Iceland Faroes Front. *Ocean Engineering* 24, 513–530.
- Bellingham, J.G., Zhang, Y., 2005. Observing processes that vary in time and space with heterogeneous mobile networks. In: *Proceedings of the International Workshop on Underwater Robotics for sustainable management of Marine Ecosystems and Environmental Monitoring*, Genoa, Italy.
- Bissett, W.P., Arnone, R.A., Davis, C.O., Dickey, T.D., Dye, D., Kohler, D.D.R., Gould Jr., R.W., 2004. From meters to kilometers: a look at ocean-color scales of variability, spatial coherence, and the need for fine-scale remote sensing in coastal ocean optics. *Oceanography* 17 (2), 32–43.
- Boss, E., Pegau, W.S., Lee, M., Twardowski, M., Shybanov, E., Korotaev, G., Baratange, F., 2004. Particulate backscattering ratio at LEO 15 and its use to study particle composition and distribution. *Journal of Geophysical Research* 109 (C01014).
- Chang, G.C., Dickey, T.D., Schofield, O., Weidemann, A.D., Boss, E., Moline, M.A., Glenn, S.M., 2002. Nearshore physical forcing of bio-optical parameters in the New York Bight. *Journal of Geophysical Research* 107 (C9), 3133.
- Charria, G., Melin, F., Dadou, I., Radenac, M.H., Garçon, V., 2003. Rossby wave and ocean color: the cells uplifting hypothesis in the South Atlantic Subtropical Convergence Zone. *Geophysical Research Letters* 30 (3), 1125.
- Cunningham, A., Mckee, D., Craig, S., Tarran, G., Widdicombe, C., 2003. Fine-scale variability in phytoplankton community structure and inherent optical properties measured from an autonomous underwater vehicle. *Journal of Marine Systems* 43, 51–59.
- Davis, F.W., 1993. Introduction to spatial statistics. In: Levin, S.A., Powell, T.M., Steele, J.H. (Eds.), *Patch Dynamics*. Springer, Berlin, Heidelberg, pp. 16–26.
- DeKshenicks, M.M., Donaghay, P.L., Sullivan, J.M., Rines, J.E.B., Osborn, T.R., Twardoski, M.S., 2001. Temporal and spatial occurrence of thin phytoplankton layers in relation to physical processes. *Marine Ecology Progress Series* 223, 61–71.
- Denman, K.L., Platt, T., 1976. The variance spectrum of phytoplankton and temperature in a turbulent ocean. *Journal of Marine Research* 34, 593–601.
- Denman, K.L., Powell, T.M., 1984. Effects of physical processes on planktonic ecosystems in the coastal ocean. *Oceanographic Marine Biology Annual Review* 22, 125–168.
- Deutschman, D.H., Bradshaw, G.A., Childress, W.M., Daly, K.L., Grunbaum, D., Pascual, M., Schumaker, N.H., Wu, J., 1993. Mechanisms of patch formation. In: Levin, S.A., Powell, T.M., Steele, J.H. (Eds.), *Patch Dynamics*. Springer, Berlin, Heidelberg, pp. 184–209.
- Dickey, T., 1991. The emergence of concurrent high resolution physical and bio-optical measurements in the upper ocean and their applications. *Reviews of Geophysics* 29 (3), 383–413.
- Diggle, P.J., Ribeiro, P., 2007. *Model-based Geostatistics*. Springer, New York.
- Drake, P.T., McManus, M.A., Storlazzi, C.D., 2005. Local wind forcing of the Monterey Bay area inner shelf. *Continental Shelf Research* 25, 397–417.
- Dustan, P., Pinckney Jr., J.L., 1989. Tidally induced estuarine phytoplankton patchiness. *Limnology and Oceanography* 34 (2), 410–419.
- Flagg, C.A., Kim, H.S., 1998. Upper ocean currents in the northern Arabian Sea from shipboard ADCP measurements collected during the 1994–1996 US JGOFS and ONR Programs. *Deep-Sea Research II* 45, 1917–1959.
- Glenn, S., Arnone, R., Bergmann, T., Bissett, W.P., Crowley, M., Cullen, J., Gryzmski, J., Haidvogel, D., Kohut, J., Moline, M., Oliver, M., Orrico, C., Sherrell, R., Song, T., Weidemann, A., Chant, R., Schofield, O., 2004. Biogeochemical impact of summertime coastal upwelling on the New Jersey Shelf. *Journal of Geophysical Research* 109, C12S02.
- Hamilton, P., Rattray, M., 1978. A numerical model of depth dependent, wind driven upwelling circulation on a continental shelf. *Journal of Physical Oceanography* 8, 437–457.
- Haury, L.R., McGowan, J.A., Wiebe, P.H., 1978. Patterns and processes in the time-space scales of plankton distributions. In: Steele, J.H. (Ed.), *Spatial Pattern in Planktonic Communities*. Plenum Press, New York, pp. 277–327.
- Herren, C.M., Haddock, S.H.D., Johnson, C., Orrico, C.M., Moline, M.A., Case, J.F., 2005. A multi-platform bathyphotometer for fine-scale, coastal bioluminescence research. *Limnology and Oceanography: Methods* 3, 247–262.
- Holligan, P.M., Reiners, W.A., 1992. Predicting the responses of the coastal zone to global change. *Advances in Ecological Research* 22, 211–255.
- Huijbregts, C.J., 1975. Regionalized variables and quantitative analysis of spatial data. In: Davis, J.C., McCullagh, M.J. (Eds.), *Display and Analysis of Spatial Data*. Wiley, London, pp. 38–53.
- Journel, A.G., Huijbregts, C.J., 1978. *Mining Geostatistics*. Academic Press, London.
- Jumars, P.A., 1978. Spatial autocorrelation with RUM (Remote Underwater Manipulator): vertical and horizontal structure of a bathyal benthic community. *Deep-Sea Research* 25, 589–604.
- Kaluzny, S.P., Vega, S.C., Cardoso, T.P., Shelly, A.A., 1998. *S+ Spatial-Stats: User's Manual for Windows and UNIX*. Springer, New York.
- Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm. *Ecology* 74 (6), 1659–1673.
- Losee, J., Richter, K., Lieberman, S., Lapota, D., 1989. Bioluminescence: spatial statistics in the North Atlantic. *Deep-Sea Research* 36 (5), 783–801.
- Lovejoy, S., Currie, W.J.S., Claeroboudt, T.Y., Bourget, E., Roff, J.C., Schertzer, D., 2001. Universal multifractals and ocean patchiness: phytoplankton, physical fields and coastal heterogeneity. *Journal of Plankton Research* 23 (2), 117–141.
- Machu, E., Ferret, B., Garçon, V., 1999. Phytoplankton pigment distribution from SeaWiFS data in the subtropical convergence zone south of Africa: a wavelet analysis. *Geophysical Research Letters* 26, 1469–1472.

- Mackas, D.L., 1984. Spatial autocorrelation of plankton community composition in a continental shelf ecosystem. *Limnology and Oceanography* 29 (3), 451–457.
- Mackas, D.L., Denman, K.L., Abbott, M.R., 1985. Plankton patchiness: biology in the physical vernacular. *Bulletin of Marine Science* 37 (2), 652–674.
- Matheron, G., 1963. Principles of geostatistics. *Economic Geology* 58, 1246–1266.
- Martin, A.P., 2003. Phytoplankton patchiness: the role of lateral stirring and mixing. *Progress in Oceanography* 57, 125–174.
- Moline, M.A., Blackwell, S.M., Chant, R., Oliver, M.J., Bergmann, T., Glenn, S., Schofield, O.M.E., 2004. Episodic physical forcing and the structure of phytoplankton in the coastal waters of New Jersey. *Journal of Geophysical Research* 109, C12S05.
- Moline, M.A., Heine, E., Case, J.F., Herren, C., Schofield, O., 2001. Spatial and temporal variability of bioluminescence potential in coastal regions. In: Case, J.F., Herring, P.J., Haddock, S.H.D., Kricka, L.J., Stanley, P.E. (Eds.), *Bioluminescence and Chemiluminescence 2000*. World Scientific, Singapore, pp. 123–126.
- Moline, M.A., Blackwell, S.M., Von Alt, C., Allen, B., Austin, T., Case, J., Forrester, N., Goldsborough, R., Purcell, M., Stokey, R., 2005. Remote environmental monitoring units: an autonomous vehicle for characterizing coastal environments. *Journal of Atmospheric and Oceanic Technology* 22, 1797–1808.
- Moline, M.A., Woodruff, D.L., Evans, N.R., 2007. Optical delineation of benthic habitat using an autonomous underwater vehicle. *Journal of Field Robotics* 24 (6), 461–471.
- Munchow, A.N., Chant, R.J., 2000. Kinematics of inner shelf motions during the summer stratified season off new Jersey. *Journal of Physical Oceanography* 30, 247–268.
- O'Brien, J.J., Hurlburt, H.E., 1972. A numerical model of coastal upwelling. *Journal of Physical Oceanography* 2, 14–26.
- Ondercin, D.G., Atkinson, C.A., Kiefer, D.A., 1995. The distribution of bioluminescence and Chl during the late summer in the North Atlantic: maps and a predictive model. *Journal of Geophysical Research* 100, 6575–6590.
- Rao, A.D., Joshi, M., Baby, S.V., 2005. A three-dimensional numerical model of coastal upwelling along the west coast of India. *Mathematical and Computer Modeling* 41, 177–195.
- Shulman, I., McGillicuddy Jr., D.J., Moline, M.A., Haddock, S.H.D., Kindle, J.C., Nechaev, D., Phelps, M.W., 2005. Bioluminescence intensity modeling and sampling strategy optimization. *Journal of Atmospheric and Oceanic Technology* 22, 1267–1281.
- Seliger, H.H., McKinley, K.R., Biggley, W.H., Rivkin, R.B., Aspden, K.R.H., 1981. Phytoplankton patchiness and frontal regions. *Marine Biology* 61, 119–131.
- Seuront, L., Schmitt, F., Lagadeuc, Y., Schertzer, D., Lovejoy, S., Frontier, S., 1996. Multifractal analysis of phytoplankton biomass and temperature in the ocean. *Geophysical Research Letters* 23 (24), 3591–3594.
- Seuront, L., Schmitt, F., Schertzer, D., Lovejoy, S., 1999. Universal multifractal analysis as a tool to characterize multiscale intermittent patterns: example of phytoplankton distribution in turbulent coastal waters. *Journal of Plankton Research* 21, 877–922.
- Stommel, H., 1963. Varieties of oceanographic experience. *Science* 139 (3555), 572–576.
- Washburn, L., Emery, B.M., Jones, B.H., Ondercin, D.G., 1998. Eddy stirring and phytoplankton patchiness in the subarctic North Atlantic in late summer. *Deep-Sea Research* 45, 1411–1439.
- Wiebe, P.H., Mountain, D.G., Stanton, T.K., Greene, C.H., Lough, G., Kaartvedt, S., Dawson, J., Copley, N., 1996. Acoustical study of the spatial distribution of plankton on Georges Bank and the relationship between volume backscattering strength and taxonomic composition of the plankton. *Deep-Sea Research II* 43 (7–8), 1971–2001.
- Yoder, J.A., McClain, C.A., Blanton, J.O., Oey, L., 1987. Spatial scales in CZCS-chlorophyll imagery of the southeastern US continental shelf. *Limnology and Oceanography* 32 (4), 929–941.
- Yost, R.S., Uehara, G., Fox, R.L., 1982. Geostatistical analysis of soil chemical properties of large land areas. 1. Semi-variograms. *Soil Science Society of America Journal* 46, 028–1037.
- Yu, X., Dickey, T., Bellingham, J., Manov, D., Streitlien, K., 2002. The application of autonomous underwater vehicles for interdisciplinary measurements in Massachusetts and Cape Cod Bays. *Continental Shelf Research* 22, 2225–2245.