

# Physical and biogeochemical spatial scales of variability in the East Australian Current separation from shelf glider measurements

Amandine Schaeffer<sup>1</sup>, Moninya Roughan<sup>1</sup>, Emlyn Jones<sup>2</sup>, and Dana White<sup>1</sup>

<sup>1</sup>Coastal and Regional Oceanography Lab, School of Mathematics and Statistics, University of New South Wales, Sydney NSW, 2052, Australia

<sup>2</sup>CSIRO Oceans and Atmosphere, Hobart, Tasmania, Australia

*Correspondence to:* Amandine Schaeffer (a.schaeffer@unsw.edu.au)

**Abstract.** In contrast to physical processes, biogeochemical processes are inherently patchy in the ocean, which affects both the observational sampling strategy and the representativeness of sparse measurements in data assimilating models. In situ observations from multiple glider deployments are analyzed to characterize spatial scales of variability in both physical and biogeochemical properties, 5 using an empirical statistical model. We find that decorrelation ranges are strongly dependent on the balance between local dynamics and mesoscale forcing. The shortest horizontal (5-10 km) and vertical (45 m) decorrelation ranges are for chlorophyll-a fluorescence. Whereas those variables that are a function of regional ocean and atmosphere dynamics (temperature and dissolved oxygen) result in anisotropic patterns with longer ranges along (28-37 km) than across the shelf (8-19 km). Variables 10 affected by coastal processes (salinity and colored dissolved organic matter) have an isotropic range similar to the baroclinic Rossby radius (10-15 km).

## 1 Introduction

At the interface between oceanic and coastal processes, continental shelf regions are characterized by complex dynamics resulting from the interaction between different water masses at smaller spatial 15 scales than the open ocean (Yoder et al., 1987). While wind, topography or density driven processes mostly influence the mixing and advection of the physical characteristics (temperature and salinity) of the shelf water masses, locally acting ecological processes are also determinant for biogeochemistry (Ballantyne et al., 2011). In particular, the numerous mechanisms driving phytoplankton distributions have been studied for many years, and highlight the complexity of these interactions 20 (Martin, 2003). Biogeochemical (BGC) processes operate over a wide range of scales and thus need to be considered separately when investigating the dominant length scales of variability for the shelf water's properties (Pan et al., 2014).

The continental shelf off southeastern Australia (between 29 and 34°S) is relatively narrow, between 16 and 70 km (mean of 37 km) from the coastline to the 200 m isobath. The dynamics on the shelf are influenced both by local coastal processes and the episodic intrusion of the large scale East Australian Current (EAC) and its eddies (Fig. 1, Schaeffer et al. (2013, 2014a)). The EAC is the western branch of the subtropical gyre in the South Pacific. It is a warm and dynamic poleward flowing current, encroaching on the continental shelf of southeastern Australia between around 18°S (Ridgeway and Godfrey, 1994) and usually 30.7 - 32.4°S (Cetina-Heredia et al., 2014) where it bifurcates eastward, forming the Tasman Front. Further south, eddies are shed (Everett et al., 2012), leading to high variability in the velocity field and water masses on the shelf (Schaeffer et al., 2014b; Schaeffer and Roughan, 2015).

Previous studies have highlighted the high spatial heterogeneity of physical (Oke et al., 2008; Schaeffer and Roughan, 2015) and biochemical (Hassler et al., 2011) variables on this narrow shelf. De-correlation time scales were quantified from *in situ* mooring observations at 30° and 34°S (Roughan et al., 2013), being of the order of hours for cross-shelf velocity to days and weeks for along-shelf flow and temperature, respectively. However, spatial scales of variability, which are essential for data assimilating models, have not been quantified.

Here we quantify for the first time the spatial scales of variability of both the physical and the BGC characteristics of the shelf water masses in the highly dynamic EAC separation zone. We use hydrographic measurements from 23 glider deployments along the coast (Section 2) to understand the variability amongst physical and BGC properties, the spatial anisotropy and the unresolved variance in the rich dataset (Section 3). Finally the results are discussed in the context of their applicability to modelling and data assimilation, where the perennial issue of relating point based measurements to model solutions is discussed (Section 4).

## 2 Methods

### 50 2.1 The Dataset

Ocean gliders are autonomous underwater vehicles which change their buoyancy to dive through the water column. Without propulsion, this vertical motion is transformed into horizontal momentum using the vehicle's wings, while its pitch controls the forward motion. During the resulting vertical sawtooth pattern through the water column, a wealth of scientific observations are recorded and analysed here. Physical and BGC measurements from 23 ocean glider deployments along the southeastern coast of Australia are used in this study. The glider missions span all seasons over 6 years, between 2008 - 2014, including results from both shallow-diving Slocum (<200 m) and deep-diving Seaglider (<1000 m) vehicles. The gliders were typically deployed at 29.4°S although some were

deployed as far south as 33°S (Fig. 1 and Schaeffer and Roughan (2015)). Missions range from 2-3 weeks to three months depending on the vehicle. The horizontal displacement between two dives increases with the depth of the dive, with median over ground distances from 130 m (for dives in 25 - 50 m of water) to 1100 m (in 150 - 200 m of water). The vertical resolution of observations is <2 m due to the fast sampling frequency.

Scientific measurements include depth, temperature and salinity (from a Seabird-CTD), dissolved oxygen (DO, from Aanderaa or Seabird Oxygen sensors), and optical parameters, chlorophyll-a fluorescence (excited / emitted wavelengths: 470 / 695 nm), colored dissolved organic matter (CDOM, excited / emitted wavelengths: 370 / 460 nm) and backscatter coefficient at 650-700 nm (from a WETLabs optical sensor).

Quality control for physical parameters (temperature and salinity) and DO are conducted following ARGO standards (Wong et al., 2014), including a salinity spike correction due to the use of unpumped CTDs in early deployments. For bio-optical parameters, quality control is more challenging due to the instrument bio-fouling and the high temporal and spatial variability of the measurements. Sensors are calibrated approximately every 2 years. To check for sensor drift, performance tests are undertaken using purple and black solid standards pre-, post-deployment and also after cleaning the sensor from bio-fouling. These tests enable the identification and flagging of suspect measurements. A global range test is also conducted with a valid fluorescence maximum set to  $50 \text{ mg m}^{-3}$ , similar to ARGO standards (Claustre, 2011). A valid regional maximum for CDOM is defined, based on all the shelf glider deployments, as the mean plus 10 times the standard deviation (= 8.0 ppb) to remove high outliers (reaching 250 ppb).

## 2.2 Characterising spatial variability

The semivariogram approach was first introduced in geostatistics (Journel and Huijbregts, 1978) to characterize the spatial variability of a sparsely distributed dataset. It describes the average dis-similarity between measurements as a function of the distance separating them. This difference is generally small for measurements within close proximity, increasing with distance, until it does not depend on a spatial lag (decorrelated values) (Legarda and Thomas, 2007; Tortell et al., 2011).

For a variable anomaly  $Z(x)$ , the semivariogram or structure function,  $\gamma(h)$ , is defined as half the mean square difference between values at a given separation  $h$ :

$$\gamma(h) = \frac{1}{2} \frac{1}{N} \sum ([Z(x) - Z(x + h)]^2) \quad (1)$$

where the sum is over all  $N$  pairs of observations that are separated by the distance  $h$  in the  $x$  direction.

95 In order to take into account outliers in the distribution of the empirical anomalies  $Z$ , Cressie and  
 Hawkins (1980) proposed a modified estimate of the structure function which is more robust when  
 the anomaly fields deviate from being Gaussian:

$$\gamma(h) = \frac{\frac{1}{2} \left( \frac{1}{N} \sum [Z(x) - Z(x+h)]^{1/2} \right)^4}{0.457 + \frac{0.494}{N}} \quad (2)$$

100 In this equation, the power 1/2 comes from a fourth-root of  $[Z(x) - Z(x+h)]^2$  that reduces the  
 skewness in the distribution, thereby approaching a Gaussian process. The 4<sup>th</sup> square acts to correct  
 the scale and returns the same units as equation 1, while the denominator adjusts the bias resulting  
 from the whole transformation. This estimate is more robust statistically in the sense that the mean  
 can be applied to the new distribution. Compared to equation 1, the semivariogram is only slightly  
 modified for the highest lags when using the robust equation 2, but the parameters (sill, range and  
 105 nugget that are investigated in section 3) remain very similar.

110 The variables' anomalies are obtained by removing large scale patterns, resulting from the average  
 of all glider measurements over predefined bins determined by latitude and depth as in Schaeffer and  
 Roughan (2015). This three-dimensional mean state is then smoothed using a spline method before  
 being removed from each observation.

115 Both cross- and along-shelf semivariograms are calculated to investigate anisotropy, where  $h = \Delta x$   
 is the zonal distance, or  $h = \Delta y$  is the meridional distance, respectively. The cross-shelf semivariance  
 is calculated following equation 2 from measurement pairs located within  $0.1^\circ$  ( $\sim 10$  km) of  
 latitude. Similarly, the along-shelf semivariance  $\gamma(h)$  is computed using observations within  $0.1^\circ$  of  
 longitude ( $\sim 10$  km) from each other. In both cases the distance vector is discretized with intervals  
 of 500 m and the time lag between pairs is limited to one day. The semivariograms are calculated in  
 the horizontal plane at three depths: surface (0 - 5 m), mixed layer depth (MLD, 5 - 30 m, defined  
 from the average profiles), or below the MLD at 50 m. Finally, glider profiles are also used to analyse  
 vertical scales by computing  $\gamma(h)$  with  $h = \Delta z$  (intervals of 1 m).

120

The semivariance  $\gamma(h)$  is computed from the trimmed mean (20 % outliers excluded) of measurements over all glider deployments, provided there are at least 10 (5 for CDOM across the shelf, see section 3.4) different missions and more than 30 pairs for each spatial lag, to avoid seasonal bias or insignificant values. We then fit a mathematical spherical model (Doney et al., 2003) to the empirical  
 125 semivariogram in order to extract the physical characteristics of the function, following:

$$\begin{cases} \sigma_0^2 + (\sigma^2 - \sigma_0^2) \left( \frac{3}{2} \frac{h}{r} - \frac{1}{2} \left( \frac{h}{r} \right)^3 \right) & 0 < h \leq r \\ \sigma^2 & h > r \end{cases} \quad (3)$$

where  $h$  is the distance between measurements,  $\sigma^2$  is the sill,  $\sigma_0^2$  is the nugget and  $r$  the range. (These variables are described physically in the example below.) Exponential and Gaussian models (Biswas and Si, 2013) were also tested but were less adequate in terms of sum of squared error (sse) and 130 adjusted R-square statistics for the fit of the empirical semivariogram.

### 3 Results

#### 3.1 Satellite derived SST semivariogram

By way of both example and validation we calculate the cross-shelf semivariogram obtained from 135 daily satellite remote sensed sea surface temperature (SST) anomalies (Fig. 2). The spherical model (Eq. 3) is fitted to the empirical semivariance values calculated for cross-shelf lags over daily maps of SST in 2014. Only days with spatial coverage greater than 30 % of the domain are considered. The physical characteristics extracted from the model are indicated in Fig 2. The sill  $\sigma^2$  reflects the 140 constant background variability of the variable. It is reached at a specific distance, here  $r = 24$  km, which is referred to as the (decorrelation) range or the dominant length scale. For lags greater than this range, the two observations are considered randomly correlated spatially. The nugget,  $\sigma_0^2$ , is the semivariance obtained from the model at the origin. If different from 0, it implies variability at 145 shorter spatial scales than those resolved by the observations. This variability is either a) real but unresolved, or b) resulting from measurement errors. The semivariogram for SST (Fig. 2) shows very little nugget effect, showing the accuracy of the measurements and an adequate spatial resolution.

As expected, the semivariance of the SST anomaly (the annual mean was subtracted) differs with seasonality, as shown by the monthly empirical semivariograms (colored dots in Fig. 2). Austral Summer and Autumn months are characterized by a sharper increase in the SST variance with greater variability in sill, due to more pronounced spatial temperature gradients. However, the semivariogram range is similar, with dominant cross-shelf scales between 18 km and 32 km (not shown). 150 The semivariogram reaches a plateau for all months with the exception of January, suggesting a trend of longer scales (Yoder et al., 1987) and a limitation of the method.

#### 3.2 Sill: *In situ* spatial variance

155 Semivariance values from glider measurements are analysed based on the values of the sill in each of the semivariograms shown in Fig. 3. Temperature, dissolved oxygen (DO), and to a lesser extent colored dissolved organic matter (CDOM) and salinity, are characterised by a greater variance in the vertical than in the horizontal (see the different y-axis). In contrast, chlorophyll-a fluorescence shows comparable variability in all directions. 160 Focusing on horizontal sills (Fig. 3 middle and left), the highest variance for salinity and CDOM

occurs at the surface in agreement with the influence of riverine input. The cross-shelf sill for DO is greater at 50 m than at the surface, suggesting more spatial variability due to bio-physical processes (remineralisation, respiration or bottom water uplift) than resulting from gas exchange with the atmosphere. Chlorophyll-a fluorescence shows little variance at 50 m depth due to light limitation 165 preventing biological activity. The highest horizontal sill for temperature appears below the MLD along the shelf, in agreement with the large latitudinal gradients in bottom temperature evidenced by Schaeffer and Roughan (2015). The surface temperature sill is smaller when measured by the gliders (Fig. 3) than by satellite (Fig. 2), possibly due to different measurement depth (*in situ* 0-5 m 170 versus skin SST), or seasonality, as glider deployments are more numerous in winter. Nevertheless, the cross-shelf dominant length scales are in good agreement in the two datasets, with ranges of 25 and 19 km, respectively.

### 3.3 Range: *In situ* scales of variability

Cross-shelf, along-shelf and vertical ranges from the semivariograms are presented in Fig. 3 and 175 summarized in Table 1. Spatial scales highlight different directional patterns between the parameters. Horizontal scales for salinity and CDOM are 9-15 km, 5-10 km for chlorophyll-a fluorescence, similar across and along the shelf. Mean temperature scales across the shelf are 18-19 km at the surface and in the MLD, only 14 km at 50 m. Scales found along the shelf are greater, being 28-29 km and 37 km, respectively. This directional anisotropy for temperature is in agreement with 180 the geometry of the shelf and the influence of the EAC at the shelf break (Fig. 1). Schaeffer and Roughan (2015) and Oke et al. (2008) both evidenced greater temperature gradients across than along the shelf, based on satellite, model and glider datasets. This directional anisotropy is also evident in density (not shown), which has been shown to be mostly temperature driven (Schaeffer et al., 185 2014b), and even more intensified for DO. While DO is characterized by dominant cross-shelf scales similar to salinity and CDOM (8 - 15 km), the along-shelf spatial variability seems to be linked to the shallow EAC watermass, resulting in decorrelation scales of 27 - 35 km (surface and MLD) similar to temperature.

Chlorophyll-a fluorescence has the smallest characteristic length scales both across and along the shelf, but also in the vertical. Measurements of fluorescence are decorrelated for depth lags greater 190 than 46 m, in agreement with shallow (near surface) chlorophyll blooms. Vertical length scales for DO and CDOM (57 - 58 m), are less than those for temperature and salinity (62 m and 66 m, respectively).

The second peak in semivariance (at 80 - 100 m for temperature, salinity and DO, Fig. 3, right) indicates an anti-correlation for these lags (Legaard and Thomas, 2007). Negative correlation coefficients reaching -0.6 were previously observed from moored autumnal temperature observations in 195 100 m water depth at 30°S (Roughan et al., 2013) and attributed to simultaneous heating source in

the surface layers and cooling at depth due to EAC encroachments and slope water uplift. Our results suggest that these current-driven uplifts are associated with a signature in salinity and DO .

200 **3.4 Nugget: *In situ* unresolved variance**

The fraction of resolved and unresolved variance is estimated from the semivariogram parameters, the sill and nugget, respectively. A nugget occurs when the difference between the two closest measurements is greater than 0, and can be seen at the origin of the semivariogram.

Overall, the high density glider observations capture most of the spatial ocean variability. The advantage of this sampling strategy is that nearly all the vertical variance is resolved for most of the parameters (ratio  $\sigma_0^2/\sigma^2 \sim 0 - 3 \%$ , Table 1) due to the high sampling frequency of the gliders compared to their vertical displacement velocity. The only exception is for CDOM with the nugget being 24% of the total variance (Fig. 3 and Table 1).

Horizontal variability is well resolved for temperature and salinity with ratios  $\sigma_0^2/\sigma^2 \leq 10 \%$  across the shelf, mostly  $\leq 14 \%$  along the shelf. Nuggets for BGC parameters are higher, reaching 27 % of the sill. While high nuggets for fluorescence and DO can be attributed to horizontal sub-scale unresolved biological activity, CDOM datasets might also suffer from measurement errors and quality control issues, as suggested by the high nugget effect in the vertical, the large outliers and the larger amount of cross-shelf lags necessary for the successful fit of a mathematical model (see section 2 and Fig. 3 e).

## 4 Discussion

This study combines in situ measurements from multiple glider deployments between 2008 - 2014 on the southeastern Australian continental shelf, to provide insight into the surface and sub-surface structure of the water-mass dynamics, including the influence of the EAC, upwelling and freshwater inputs. Analysis of length scale dependent variability demonstrates that much of the spatial variance in physical and BGC parameters typically occurs at scales ranging 5 km for chlorophyll-a fluorescence to  $\sim 35$  km for along-shelf temperature.

In this study the length scales were averaged from data obtained over 2 degrees of latitude, however we expect more regional variability resulting from the different latitudinal regimes evidenced by Schaeffer and Roughan (2015), driven by the meso-scale circulation. In addition, we expect that spatial scales may vary seasonally, particularly in the biological parameters. This will be tested when we have sufficient data in each season.

230 As for all statistics, limitations arise from the amount of data used (especially along the shelf where the data density is smaller) and contamination of the dataset (for instance CDOM). In geo-statistics, uneven spatial distribution of the observations over the analyzed area can be a limitation as well but remains difficult to quantify. The major advantage of the semivariogram method used is that it can be applied to sparse dataset like glider observations, as opposed to spatial autocorrelations for  
235 instance. It allows objective comparison of interesting parameters (range, sill, nugget) for different variables, directions and depths. In this study, the results compare well when using different statistical fits, and are consistent with expected outcomes based on previous knowledge of local dynamics and related studies in other regions.

#### 4.1 Related studies

240 From a global analysis of satellite derived surface data, Doney et al. (2003) found comparable small-scale variability for biology and physics. However, they were not able to characterize scales  $< 15$  km based on the satellite products used. Here we find that BGC distribution occurs predominantly at submesoscales (5 - 14 km for chlorophyll-a, CDOM), while scales for temperature are larger (14 - 37 km). These short scales of variability for BGC are in agreement with the effect of nutrient cycling,  
245 reproductive rate and community interaction (e.g. grazing pressure from zooplankton) that can lead to patches of 5-10 km (Ballantyne et al., 2011; Denman et al., 1977; Goebel et al., 2014).

According to Mahadevan and Campbell (2002), the fine scale patchy distribution of phytoplankton is linked to the short characteristic time in response to disturbance in their concentration, as opposed to the longer time for temperature to adjust to external forcing. We find temperature horizontal scales  
250 (14 - 37 km) that are of the same order of magnitude as over the Malvinas current region, derived from SST (20 - 47 km, Tandeo et al. (2014)) or over the Middle Atlantic Bight from *in situ* glider observations (10 - 35 km, Todd et al. (2013)). The anisotropic shape of the temperature variance is consistent with a highly dynamic circulation (Tandeo et al., 2014), here driven by the EAC, characterized by a greater signature in temperature than in salinity.

255 Spatial variability in salinity is predominantly isotropic and similar to CDOM with decorrelation length scales of 9 - 15 km, corresponding to the first Rossby baroclinic radius of deformation (12 - 15 km based on local moored observation, Schaeffer et al. (2014b)), and high surface variance, suggesting a predominant influence of coastal processes and river input.

#### 260 4.2 Drivers of variability in a modelling perspective

Assuming that there is no first order feedback from the biology to the physics, we can think of the physics variables  $\mathbf{X} = T, S$  (temperature and salinity) being a function of internal dynamics ( $I$ , e.g. mixing), atmospheric forcing ( $A$ ), coastal buoyancy forcing arising from river discharge ( $R$ ), friction due to shallow bathymetry ( $F$ ) and open ocean forcing (e.g. tidal, geostrophy) and water masses ( $O$ ).

265 Therefore the state of the model at some spatial location "s" at time  $t$  is given by:

$$\mathbf{X}(s, t) = f(I, A, R, F, O) \quad (4)$$

where  $f(I, A, R, F, O)$  for the physical variables can be solved numerically in various hydrodynamic models.

For the state variable of temperature, we assume that there is little effect from river input in this  
270 region (e.g. water coming in is about the same temperature as the surface layer), while the effect from coastal processes is large for salinity. Therefore Eq. 4 simplifies to:

$$T(s, t) = f(I, A, F, O) \quad (5)$$

$$S(s, t) = f(I, A, R, F, O) \quad (6)$$

275 Given that both T and S are subjected to the same advection and diffusion equations, but differ only in the source/sink and boundary terms of  $f(A)$ ,  $f(R)$  and  $f(O)$ , those are the major drivers for the difference in the along shelf sills and differences in the nugget. Salinity varies over shorter length scales due to river input and the markedly different freshwater inputs from various catchment sizes along the coast. Whereas temperature is largely controlled by the regional scale EAC forcing and the  
280 relatively smooth atmospheric forcing applied which varies over spatial scales of 50 km or more.

A similar approach can be applied to the BGC variables, but  $f(I)$  is more complicated as it includes the turnover of biomass/nutrients between different plankton functional types or nutrient pools. But ultimately, one would expect  $f(I)$  to introduce variability at scales equal to or less than those seen in salinity. This hypothesis is supported by the ranges reported in the chlorophyll-a fluorescence and  
285 CDOM variables, which are biologically derived. However, as CDOM can also be introduced into the coastal ocean via river plumes and has a similar sill structure to salinity, we suggest that the CDOM measured by the glider is largely due to river discharge. The DO distribution in the surface layer is largely a function of air-sea exchange rather than primary production and will have similar variability to temperature due to the forcing mechanism. However, below the mixed layer, DO is  
290 function of the remineralisation rate and also vertical mixing/exchange with surface water, explaining the shorter decorrelation range in DO found below the mixed layer.

### 4.3 Observing system design

The length scales calculated here can be used to guide the design of ocean observing systems, in particular to answer questions related to the observation density needed to resolve along and cross-shore variability in both the physical and biological parameters. The temperature anisotropy in our results, consistent with findings of Oke and Sakov (2012) and Jones et al. (2015), shows that the required  
295

observation density will vary along and across the shelf. Thus high resolution cross-shelf mooring or glider lines every Y km are more useful than simply a glider endurance line or equally spaced 300 moorings. The distance Y can be initially derived from satellite observations, or determined after a number of glider missions. In contrast, the understanding of BGC variability, characterized by short isotropic length scales, will require high spatial resolution observations (e.g. gliders) to determine the representativeness of the measurements.

#### 305 4.4 Data assimilation

There are a variety of data assimilation systems based upon two broad approaches, ensemble methods (e.g. Oke et al. (2008), Jones et al. (2012)) and variational methods, that minimize a cost function (e.g. Moore (2011)). Regardless of the approach used, assumptions are made about the spatial footprint of an observation, for which a key parameter is the decorrelation length scale. Within the 310 ensemble (e.g. Oke et al. (2008)) and hybrid (Pan et al., 2011) data assimilation approaches, covariance localization (Sakov and Bertino, 2011) is used to increase the rank of the background error covariance matrix. The anisotropic (along-shelf and cross-shelf) ranges presented in this study and method used to derive them, allow for the direct calibration of the decorrelation scales enforced within most data assimilation systems that are currently in use. Additionally, estimates of how these 315 decorrelation scales vary in time is also available (e.g. Figure 2), suggesting that an optimally tuned data assimilation system should allow for temporal variation in the localization or provide an assessment of the temporal variability of the ensemble from an an Ensemble Kalman Filter (EnKF) system.

The results from this study also allow us to partly answer the question of how to relate a point 320 based observation with the output from a numerical model, which assumes the average concentration of a variable within a model cell  $X_{mod}$ . If we take a Bayesian view stating that we observe some true state variable with error (e.g. Parslow et al. (2013)), this can be written as:

$$X_{obs} = X_{true} + \epsilon_m + \epsilon_v \quad (7)$$

where  $X_{obs}$  is the observed variable,  $X_{true}$  is the true unknown value of the variable,  $\epsilon_m$  is the instrument 325 error and  $\epsilon_v$  is the sampling error due to unresolvable small scale variability. The observation is then related to the modelled variable by:

$$X_{mod} = X_{obs} + \epsilon_r \quad (8)$$

where  $\epsilon_r$  is typically referred to as the representation error (Oke and Sakov, 2008) associated with difference in kind (e.g. measuring fluorescence, but modelling biomass), or averaging across a model 330 grid cell that contains a point measurement.

Assuming  $\epsilon_m$  is known from calibration studies, results of studies like that presented here allow us

to explore the characteristics of  $\epsilon_v$  and  $\epsilon_r$ . For a particular variable, we can assume that the nugget is approximately equal to  $\epsilon_v$  and given a-priori information about a model grid, the spherical model applied to the semivariogram can then also be used to provide an empirical estimate for  $\epsilon_r$ .

335

To this end, the results of this study allow us to characterise the length scales of the physical and BGC properties on the shelf and relate variability to the dynamical drivers, but additionally, the methodology developed here can be directly used to improve observing system design, and to tune key data assimilation parameters that are presently poorly understood.

340 *Acknowledgements.* IMOS is an initiative of the Australian Government being conducted as part of the National Collaborative Research Infrastructure Strategy. Assistance with logistical and technical support for this project has been provided by ANFOG - Australian Facility for Ocean Gliders and the NSW-IMOS glider team. All datasets are freely available at <http://www.aodn.org.au>.

## References

345 Ballantyne, F., O. M. Schofield, and S. Levin, 2011: The emergence of regularity and variability in marine ecosystems: the combined role of physics, chemistry and biology. *Scientia Marina*, **75** (4).

Biswas, A. and B. C. Si, 2013: *Model Averaging for Semivariogram Model Parameters, Advances in Agrophysical Research*, chap. 4, 81–96. doi:10.5772/52339.

Cetina-Heredia, P., M. Roughan, E. van Sebille, and M. A. Coleman, 2014: Long-term trends in the east australian current separation latitude and eddy driven transport. *Journal of Geophysical Research: Oceans*, **119**, 4351–4366, doi:10.1002/2014JC010071.

Claustre, H., 2011: Bio-optical sensors on argo floats. reports of the international ocean-colour coordinating group, no. 11, ioccg. Tech. rep., IOCCG, Dartmouth, Canada.

Cressie, N. and D. Hawkins, 1980: Robust Estimation Of The Variogram. *Journal Of The International Association For Mathematical Geology*, **12** (2), 115–125, doi:10.1007/BF01035243.

355 Denman, K., A. Okubo, and T. Platt, 1977: The chlorophyll fluctuation spectrum in the sea. *Limnology and Oceanography*, **22** (6), 1033–1038, doi:10.4319/lo.1977.22.6.1033.

Doney, S. C., D. M. Glover, S. J. McCue, and M. Fuentes, 2003: Mesoscale variability of sea-viewing wide field-of-view sensor (seawifs) satellite ocean color: Global patterns and spatial scales. *Journal of Geophysical Research: Oceans*, **108** (C2), 3024, doi:10.1029/2001JC000843.

360 Everett, J. D., M. E. Baird, P. R. Oke, and I. M. Suthers, 2012: An avenue of eddies: Quantifying the biophysical properties of mesoscale eddies in the Tasman Sea. *Geophys. Res. Lett.*, **39**, L16608, doi:10.1029/2012GL053091.

Goebel, N. L., S. Frolov, and C. A. Edwards, 2014: Complementary use of wave glider and satellite measurements: Description of spatial decorrelation scales in chl-a fluorescence across the pacific basin. *Methods in Oceanography*, **10** (0), 90 – 103, doi:<http://dx.doi.org/10.1016/j.mio.2014.07.001>.

Hassler, C., J. Djajadikarta, M. Doblin, J. Everett, and P. Thompson, 2011: Characterisation of water masses and phytoplankton nutrient limitation in the east australian current separation zone during spring 2008. *Deep Sea Research Part II: Topical Studies in Oceanography*, **58** (5), 664 – 677, doi:10.1016/j.dsr2.2010.06.008.

370 Oke, P. R., F. Rizwi and L. M. Murray, 2008: Assimilation of glider and mooring data into a coastal ocean model. *Ocean Modelling*, **47**, 1 – 13, doi:10.1016/j.ocemod.2011.12.009.

Jones, E. M., M. A. Doblin, R. Matear, and E. King, 2015: Assessing and evaluating the ocean-colour footprint of a regional observing system. *Journal of Marine Systems*, **143** (0), 49 – 61, doi:<http://dx.doi.org/10.1016/j.jmarsys.2014.10.012>.

375 Journel, A. G. and C. J. Huijbregts, 1978: *Mining geostatistics*. Academic Press London ; New York, x, 600

Leggaard, K. R. and A. C. Thomas, 2007: Spatial patterns of intraseasonal variability of chlorophyll and sea surface temperature in the california current. *Journal of Geophysical Research: Oceans*, **112** (C9), 006, doi:10.1029/2007JC004097.

Mahadevan, A. and J. W. Campbell, 2002: Biogeochemical patchiness at the sea surface. *Geophysical Research Letters*, **29** (19), 32–1–32–4, doi:10.1029/2001GL014116.

380 Martin, A., 2003: Phytoplankton patchiness: the role of lateral stirring and mixing. *Progress in Oceanography*, **57** (2), 125 – 174, doi:[http://dx.doi.org/10.1016/S0079-6611\(03\)00085-5](http://dx.doi.org/10.1016/S0079-6611(03)00085-5).

A. M. Moore, H. G. Arango, G. Broquet, C. Edwards, M. Veneziani, B. Powell, D. Foley, J. D. Doyle, D. Costa, P. Robinson, 2011: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part II - Performance and application to the California Current System. *Progress in Oceanography*, **91**, 50 – 73, doi:10.1016/j.pocean.2011.05.003.

385 Oke, P. and P. Sakov, 2012: Assessing the footprint of a regional ocean observing system. *Journal of Marine Systems*, **105–108 (0)**, 30 – 51, doi:10.1016/j.jmarsys.2012.05.009.

Oke, P. R., G. B. Brassington, D. A. Griffin, and A. Schiller, 2008: The bluelink ocean data assimilation system  
390 (bodas). *Ocean Modelling*, **21 (1-2)**, 46 – 70, doi:10.1016/j.ocemod.2007.11.002.

Oke, P. R. and P. Sakov, 2008: Representation Error of Oceanic Observations for Data Assimilation. *Journal of Atmospheric and Oceanic Technology*, **25**, 1004, doi:10.1175/2007JTECHO558.1.

Oke, P. R., P. Sakov, and S. P. Corney, 2007: Impacts of localisation in the enkf and enoi: experiments with a small model. *Ocean Dynamics*, **57 (1)**, 32–45.

395 Pan, C., M. Yaremchuk, D. Nechaev, and H. Ngodock, 2011: Variational assimilation of glider data in Monterey Bay. *Journal of Marine Research*, **69 (2-3)**, 331 – 346, doi:10.1357/002224011798765259.

Pan, C., L. Zheng, R. H. Weisberg, Y. Liu, and C. E. Lembke, 2014: Comparisons of different ensemble schemes for glider data assimilation on west florida shelf. *Ocean Modelling*, **81**, 13 – 24, doi:10.1016/j.ocemod.2014.06.005.

400 Parslow, J., N. Cressie, E. P. Campbell, E. Jones, and L. Murray, 2013: Bayesian learning and predictability in a stochastic nonlinear dynamical model. *Ecological applications*, **23 (4)**, 679–698.

Ridgeway, K. and J. Godfrey, 1994: Mass and heat budgets in the east australian current: a direct approach. *Journal of Geophysical Research*, **99**, 3231–3248.

Roughan, M., A. Schaeffer, and S. Kioroglou, 2013: Assessing the design of the nsw-imos moored observation  
405 array from 2008-2013: Recommendations for the future. *Oceans - San Diego, 2013*, 1–7.

Sakov, P. and L. Bertino, 2011: Relation between two common localisation methods for the EnKF. *Computational Geosciences*, **15 2**, 225–237. doi:10.1007/s10596-010-9202-6

Schaeffer, A., M. Roughan, and B. Morris, 2013: Cross-shelf dynamics in a western boundary current. implications for upwelling. *Journal of Physical Oceanography*, **43**, 1042 – 1059, doi:10.1175/JPO-D-12-0177.1.

410 Schaeffer, A., M. Roughan, and B. Morris, 2014a: Corrigendum cross-shelf dynamics in a western boundary current. implications for upwelling. *Journal of Physical Oceanography*, **44**, 2812–2813, doi:<http://dx.doi.org/10.1175/JPO-D-14-0091.1>.

Schaeffer, A., M. Roughan, and J. E. Wood, 2014b: Observed bottom boundary layer transport and uplift on the continental shelf adjacent to a western boundary current. *Journal of Geophysical Research: Oceans*, **119**,  
415 4922 – 4939, doi:10.1002/2013JC009735.

Schaeffer, A., and M. Roughan, 2015: Influence of a western boundary current on shelf dynamics and upwelling from repeat glider deployments. *Geophysical Research Letters*, **42**, 121 – 128, doi:10.1002/2014GL062260.

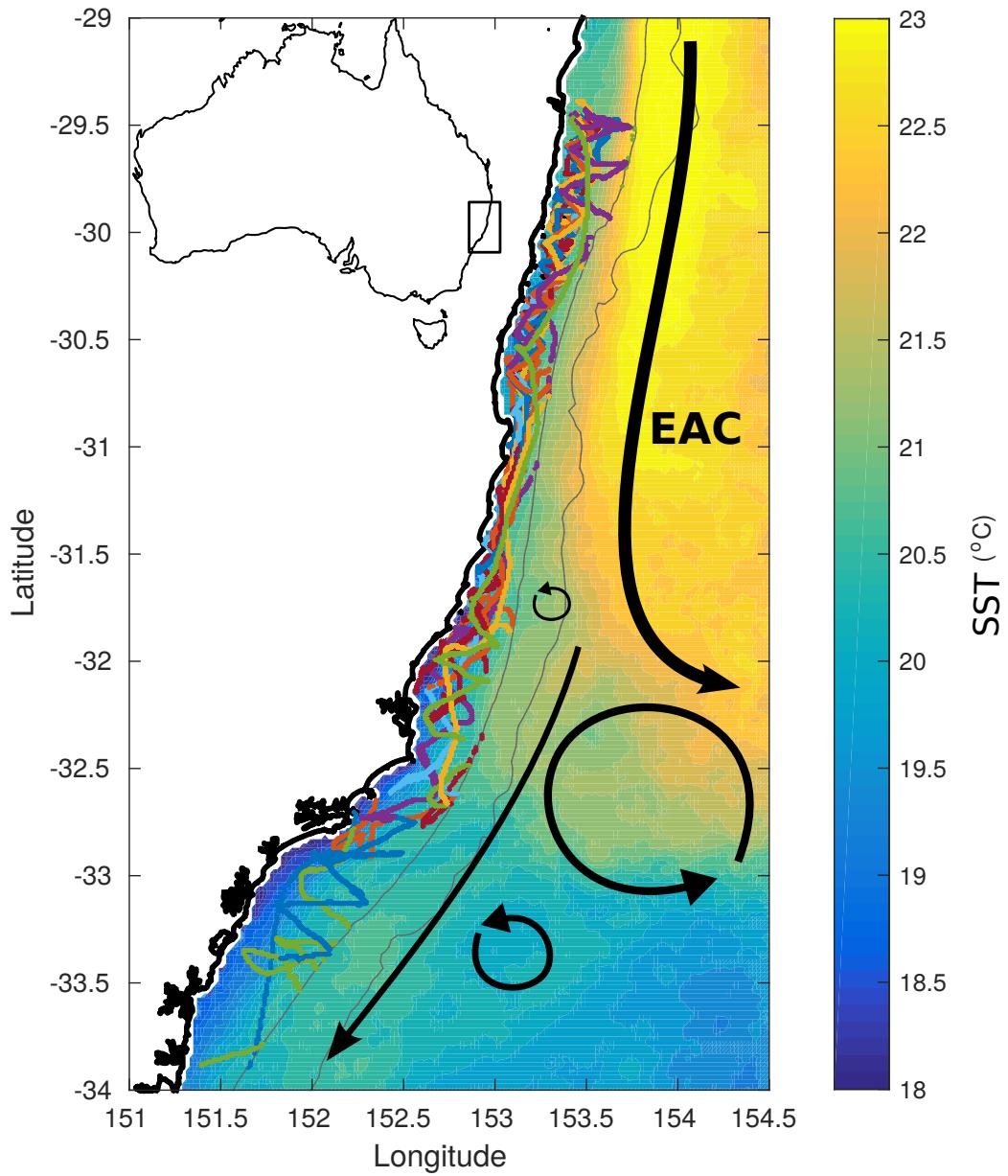
Tandeo, P., E. Autret, B. Chapron, R. Fablet, and R. Garello, 2014: {SST} spatial anisotropic covariances from metop-avhrr data. *Remote Sensing of Environment*, **141 (0)**, 144 – 148,  
420 doi:<http://dx.doi.org/10.1016/j.rse.2013.10.024>.

Todd, R. E., G. G. Gawarkiewicz, and W. B. Owens, 2013: Horizontal scales of variability over the middle atlantic bight shelf break and continental rise from finescale observations. *J. Phys. Oceanogr.*, **43**, 222–230, doi:<http://dx.doi.org/10.1175/JPO-D-12-099.1>.

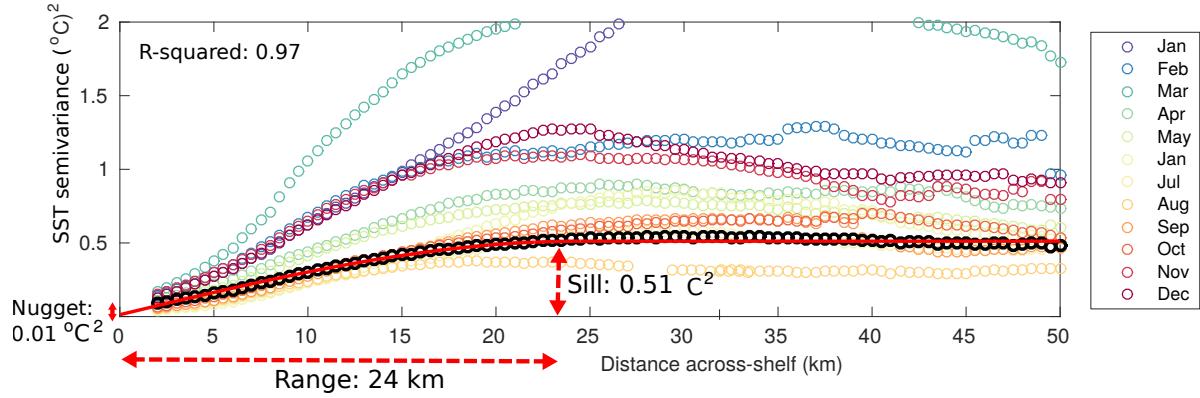
Tortell, P., C. Gueguen, M. Long, C. Payne, P. Lee, and G. DiTullio, 2011: Spatial variability and temporal  
425 dynamics of surface water pCO<sub>2</sub>, O<sub>2</sub>/Ar and dimethylsulfide in the Ross Sea, Antarctica. *Deep Sea Research Part I: Oceanographic Research Papers*, **58 (3)**, 241 – 259, doi:<http://dx.doi.org/10.1016/j.dsr.2010.12.006>.

Wong, A., R. Keeley, and T. Carval, 2014: Argo quality control manual. Tech. rep., Argo Data Management  
Team. doi:<http://dx.doi.org/10.13155/33951>.

Yoder, J. A., C. R. McClain, J. O. Blanton, and L.-Y. Oeymay, 1987: Spatial scales in czcs-chlorophyll  
430 imagery of the southeastern u.s. continental shelf. *Limnology and Oceanography*, **32 (4)**, 929–941, doi:[10.4319/lo.1987.32.4.0929](https://doi.org/10.4319/lo.1987.32.4.0929).



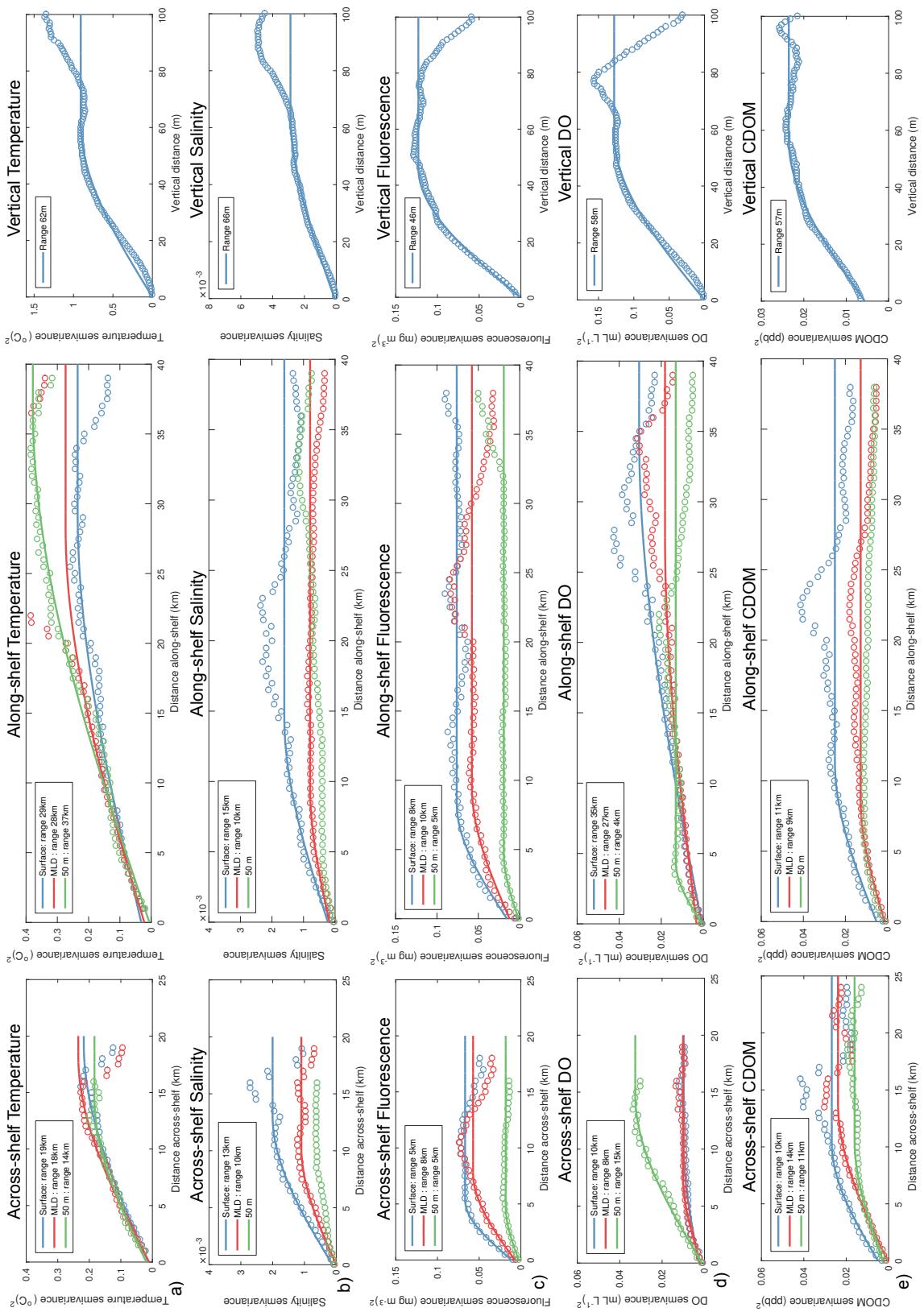
**Figure 1.** Monthly mean Sea Surface Temperature (AVHRR L3S product) over southeastern Australia for October 2014. The coastline, 200 and 2000 m isobaths are shown. Glider tracks over the shelf (depth <200 m) are indicated by colored lines. A schematic of the typical circulation is shown with the poleward flowing East Australian Current (EAC) bifurcating to the East around  $32^{\circ}\text{S}$ , its weaker extension, anticyclonic and cyclonic eddies.



**Figure 2.** Cross-shelf empirical semivariogram estimated from daily SST over the southeastern Australian shelf (depth <200 m, 29 - 34°S, AVHRR L3S product) for 2014 (black bold dots) and for each month in 2014 (colored dots). The spherical model (red line, R-squared of 0.97 for the fit) and resulting parameters (range, sill, nugget) are shown for 2014 semivariance.

		Temperature	Salinity	Fluorescence	DO	CDOM
Cross-shelf	range	19 km	13 km	5 km	10 km	10 km*
	ratio $\sigma_0^2/\sigma^2$	6 %	6 %	17%	27 %	19%
	R-squared fit	0.96	0.94	0.89	0.85	0.49
MLD	range	18 km	10 km	8 km*	8 km	14 km
	ratio $\sigma_0^2/\sigma^2$	4 %	0%	13%	18%	2%
	R-squared fit	0.92	0.83	0.58	0.89	0.96
50m	range	14 km		5 km	15 km	11 km
	ratio $\sigma_0^2/\sigma^2$	10%		15%	4%	17%
	R-squared fit	0.97		0.73	0.98	0.88
Along-shelf	range	29 km	15 km*	8 km	35 km	11 km*
	ratio $\sigma_0^2/\sigma^2$	14%	14 %	20%	2%	21%
	R-squared fit	0.93	0.52	0.83	0.90	0.56
MLD	range	28 km	10 km	10 km*	27 km	9 km*
	ratio $\sigma_0^2/\sigma^2$	8%	23%	21%	18 %	10%
	R-squared fit	0.99	0.93	0.53	0.96	0.27
50m	range	37 km		5 km	4 km*	
	ratio $\sigma_0^2/\sigma^2$	1%		8%	5%	
	R-squared fit	0.97		0.87	0.33	
Vertical	range	62 m	66 m	46 m	58 m	57 m
	ratio $\sigma_0^2/\sigma^2$	0%	3 %	1%	0%	24%
	R-squared fit	0.97	0.98	0.99	0.98	0.99

**Table 1.** Spatial scales of variability for spherical fit to semivariograms for different parameters and depths across, along the shelf and along the vertical. The range, percentage ratio of the nugget to the sill ( $\sigma_0^2/\sigma^2$ ) and R-squared for the model fit to experimental values are indicated (ranges with \* correspond to R-squared <0.7). Blanks indicate unsuccessful fit to the spherical model.



**Figure 3.** Cross-shelf (left), along-shelf (middle) and vertical (right) empirical semivariograms estimated from glider measurements of a) temperature, b) salinity, c) chlorophyll-a fluorescence, d) DO and e) CDOM. Spherical models are shown by the solid lines and the resulting spatial ranges are indicated in the insert for successful fits. Blue, red, green symbols for horizontal semivariograms correspond to surface (0 - 5 m), MLD (5 - 30 m) and 50 m measurements, respectively.