OUTPACE long duration stations: physical variability, context of biogeochemical sampling, and evaluation of sampling strategy

Alain de Verneil[1], Louise Rousselet[1], Andrea M. Doglioli[1], Anne A. Petrenko[1], Christophe Maes[2], Pascale Bouruet-Aubertot[3], and Thierry Moutin[1]

[1]Aix-Marseille University, MIO, CNRS/INSU, Marseille, France
[2]Univ. Brest, Ifremer, CNRS, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS), IUEM, F-29280, Brest, France
[3]LOCEAN, UMR7167, Université Pierre et Marie Curie, 75252 Paris, France

Correspondence to: Alain de Verneil (alain.de-verneil@mio.osupytheas.fr)

Abstract. Research cruises to quantify biogeochemical fluxes in the ocean require taking measurements at stations lasting at least several days. A popular experimental design is the quasi-Lagrangian drifter, with in situ incubations that follow the flow of water over time. The ship then tracks the drifter because subsequent measurements are supposed to remain in the same water environment. An outstanding question is how to best determine whether this is true. During the Oligotrophy to UlTra-oligotrophy PACific Experiment (OUTPACE) cruise, from 18 February to 3 April 2015 in the western tropical South Pacific, three separate stations of long duration (five days) over the upper 500 m were conducted in this quasi-Lagrangian sampling scheme. Here we present physical data to provide context for these three stations and to assess whether the sampling strategy worked, i.e. that indeed a single body of water was sampled. After analyzing tracer variability and local water circulation at each station, we argue that, while no realization will be truly Lagrangian, all long duration three stations were conducted in their own sufficiently self-similar physical environment during OUTPACE. By directly addressing the concerns raised by these quasi-Lagrangian sampling platforms, a protocol of best practices can begin to be formulated so that future research campaigns include the complementary datasets and analyses presented here to verify the appropriate use of the drifter platform.

1 Introduction

Biogeochemical cycles dictate the global distribution and fluxes of the chemical elements. Quantifying the mechanisms that mediate the various forms key elements take in these cycles, especially in the midst of ongoing climate change in the ocean, is vital to understanding the future evolution of the Earth system (Falkowski et al., 2000; Davidson and Janssens, 2006; Gruber and Galloway, 2008). Considering the wide diversity of environments where biogeochemical processes take place, it is not surprising that each sub-discipline has its own challenges with regards to collecting and processing samples. The sampling protocols put in place thus need to ensure the mechanisms of interest are isolated and put into their proper context.

In the world's surface oceans, a dominant difficulty is the medium itself: water. Sampling in a fluid that is always liable to move normally requires that one of two approaches be taken. In the first approach, a geographic location is chosen and then repeatedly sampled. This produces an Eulerian perspective, and this methodology is employed by definition at permanent mooring platforms. Set geographic locations are also often used to define time series or recurrent sampling locations, for

example stations ALOHA, BATS, CalCOFI, DYFAMED, and PAPA (Karl and Lukas, 1996; Schroeder and Stommel, 1969; Steinberg et al., 2001; Bograd et al., 2003; Marty et al., 2002; Freeland, 2007). These sites can be combined into world-wide networks and initiatives such as OceanSITES (Send et al., 2010). While this strategy makes no attempt to actually follow a given water parcel, if currents are relatively weak during a single field campaign then the variability due to advection

5   can be ignored. Unfortunately, the spatio-temporal scales of shipboard station sampling (in time, days to weeks; in space, 1-100 km) overlap with a multitude of physical phenomena ubiquitously found in the ocean, ranging from internal waves, submesoscale turbulence, up to mesoscale eddies (d'Ovidio et al., 2015). All of these motions can easily transport water such that instantaneously observed temperature, salinity, and by extension the organisms and chemical environments mediating biogeochemical processes, are markedly different from some mean value or state.

10   One way to rectify physical displacements is the second sampling approach, namely to follow the water during ongoing experiments. This approach creates a Lagrangian point of view. A common implementation of this strategy is with quasi-Lagrangian drifting moorings (Landry et al., 2009; Moutin et al., 2012). These drifters are structured so that a vertical line with sampling devices (e.g. incubation bottles and/or sediment traps) drifts along with the flow. This approach has been in routine use for decades across the globe; some examples of French campaigns known to the authors include the OLIPAC (1994),

15   PROSOPE (1999), BIOSOPE (2004), and BOUM (2008) experiments (all data and metadata accessible from LEFE-CYBER).

Naturally, the question arises whether the trajectory undertaken by the drifting mooring in the quasi-Lagrangian approach accurately represents the water movements at each of the sampling sites. If the drifter is successful in following the water, then indeed a single biogeochemical setting will have been sampled; if not successful, then the risk grows that a different environment has been brought in via advection. In practical terms, ensuring the success of the quasi-Lagrangian strategy leads

20   to a two-step process. First, the selection of sampling sites needs to be carefully considered. Sites should be relatively uniform to provide room for error should the drifting mooring wander too far from the true water displacement. Prior to the advent of satellite oceanography, structures harboring enhanced gradients such as eddies and fronts were difficult to detect before sampling. As a result, the risk always existed that sites could be chosen close to these structures, putting the drifter's mission into jeopardy. In recent years, the incorporation of near-real-time satellite data having become routine minimizes this (in the

25   list above, since the BOUM 2008 campaign; Moutin et al., 2012). The second step, after deployment of the drifting mooring, consists of an independent, post-cruise validation of the data, which is the focus of the present study.

The Oligotrophy to UlTra-oligotrophy PACific Experiment (OUTPACE) cruise provided an opportunity to assess the success of the quasi-Lagrangian sampling strategy. Conducted from 18 February to 3 April 2015 in the western tropical South Pacific (WTSP), one of the goals of OUTPACE was to assess the contribution of nitrogen fixation as a biogeochemical process to

30   the biological carbon pump (Moutin et al., 2017). During the cruise, three long duration (LD) stations employed the quasi-Lagrangian strategy. In the subsequent discourse regarding these stations, we proceed as follows. Sect. 2 describes how the drifting mooring was deployed, how concurrent data were collected, and the analyses undertaken to answer our central question of whether we sampled a single environment. We then present the data and results in Sect. 3, followed by a discussion in Sect. 4. The paper finishes in Sect. 5 with a summary of our recommendations regarding future implementations of this sampling

35   strategy.

2 Materials and Methods

In this section, we begin by describing the manner in which the three LD stations were conducted during the OUTPACE cruise. We then outline the different sources of data presented in this study and how they were processed. Additionally, we detail the analyses needed to answer our central question regarding sampling in a coherent environment.

5 ### 2.1 Sampling strategy

The OUTPACE cruise occurred aboard the RV *L'Atalante* from 18 February to 3 April in late austral summer, starting in New Caledonia and finishing in Tahiti, traveling over 4000 km. Stations were conducted in a mostly zonal transect traveling west to east, with the ship track averaging near 19° S. The three LD stations, entitled LDA, LDB, and LDC, and lasting 5 days each, were designed to resolve a regional zonal gradient in oligotrophy, the existence of which is reflected in the surface chlorophyll-
10 *a* (chl-*a*) data (Fig. 1a). As described in the introductory article of this special issue (Moutin et al., 2017), site selection for the LD stations involved identifying physical structures by use of the SPASSO software package (http://www.mio.univ-amu.fr/SPASSO/) using near-real-time satellite imagery, altimetry, and Lagrangian diagnostics (Doglioli et al., 2013; d'Ovidio et al., 2015; Petrenko et al., 2017).

Before starting each LD station, surface velocity program (SVP; Lumpkin and Pazos, 2007) drifters were deployed adjacent
15 to the site. Two quasi-Lagrangian drifting moorings were deployed during the OUTPACE LD stations with surface floats. The first drifting mooring, hereafter referred to as the SedTrap Drifter, had a 'holey sock' attached at 15 m depth. It was followed actively by the ship and is the emphasis of this study. It had three sediment traps (Technicap PPS5/4) fixed at 150, 250, and 500 m depth, along with onboard conductivity-temperature-depth (CTD) sensors and current meters, described below in Sects. 2.3.1 and 2.4.1, respectively. It was deployed at the beginning of each station and was left in the water until the station's completion.
20 The second drifter, referred to as the Production Line, housed in situ incubation platforms for measuring primary production, nitrogen fixation, oxygen, and other biogeochemical measurements (see Moutin and Bonnet, 2015, for more documentation). The Production Line was redeployed on a daily basis close to the SedTrap Drifter. After 5 days, the SedTrap Drifter was recovered, and the LD station completed. Occasions when the exact implementation of this general strategy was not realized will be mentioned in following sections for the relevant measurements. A summary of time duration for each data source can
25 be found in Table 1.

Between the LD stations, 15 short duration (SD) stations lasting approximately 8 h each were interspersed along the ship's trajectory in roughly equidistant sections. Among the measurements made, CTD casts from SD stations will figure into the validation process in this study. Most casts (both LD and SD) were at least 200 m, with at least one 2000 m cast for all stations. These casts were conducted with the same CTD rosette platform described more fully below in Sect. 2.3.1.
30 Throughout the cruise, surface conductivity-temperature measurements from the themosalinograph (TSG) and currents from shipboard acoustic Doppler current profilers (SADCP) were collected. Their processing is described in Sects. 2.3.1 and 2.4, respectively.

2.2 Remote Sensing Data

Satellite measurements of sea surface temperature (SST), surface chl-*a*, sea surface height and its associated geostrophic currents and diagnostics, also used in the LD site selection phase (Sect. 2.1), provided temporal context for the LD stations. All processed satellite data were provided by CLS with support from CNES. SST was derived from a combination of AQUA/MODIS,

5 TERRA/MODIS, METOP-A/AVHRR, METOP-B/AVHRR sensors, with the product produced being a weighted mean spanning 5 days (inclusive) previous to the date in question. Weighting was greater for more recent data. Similar to SST, chl-*a* was a 5 day weighted mean produced by the Suomi/NPP/VIIRS sensor. The SST and chl-*a* products had a $0.02°$ resolution, equivalent to ~ 2km. These satellite products spanned from 1 December 2014 to 15 May 2015.

The temporal fluctuations of SST and surface chl-*a* were determined by producing timeseries of both variables within a

10 given spatial range surrounding the starting position of the three LD stations. Square regions of sides 120 km long, centered at each LD station, were chosen as the boundaries. Satellite pixels falling within this region were used to create a probability distribution function. The 120 km square size was chosen because 60 km is a typical size of the Rossby radius of deformation for the region (Chelton et al., 1998).

Local surface currents derived from altimetry were also provided by CLS with support from CNES. These data come from

15 the Jason-2, Saral-AltiKa, Cryosat-2, and HY-2A missions, cover a domain from $140°$ E to $220°$ E, and $30°$ S to the equator, covering the yearlong period of June 2014 to May 2015. The velocity grid had a $\frac{1}{8}°$ resolution, applying the FES2014 tidal model and CNES_CLS_2015 mean sea surface. Ekman effects due to wind were also added using ECMWF ERA INTERIM model output.

2.3 Hydrographic data and methodology

20 ### 2.3.1 CTD data sources

The shipboard CTD employed during OUTPACE was a Seabird SBE 9+ CTD-rosette, with two CTDs installed. Data from each cast were calibrated and processed post-cruise using Sea-Bird Electronics software into 1 m bins. All CTD data from other instruments mentioned later were likewise processed using Sea-Bird Electronics software. Final absolute salinity (S_A), conservative temperature (C_T), and potential density (σ_θ) were calculated using the TEOS-10 standard (McDougall and Barker,

25 2011). In total, over 200 CTD casts were performed during OUTPACE. Most SD stations had three or four casts, except for SD13, which had time for only one cast owing to a medical emergency. The LDA, LDB, and LDC stations had 46, 47, and 46 casts, respectively, each approximately 3 h apart. During LDA, the two drifting moorings accidentally collided and, due to the time necessary to disentangle them, there is a gap of 9 h in the timeseries. The majority of CTD casts were to 200 m depth, with at least one 2000 m cast per station. Mixed layer depth was determined using de Boyer Montégut et al. (2004)'s method,

30 by finding the depth where density has changed more than 0.3 kg m^{-3} from a reference value, which was chosen to be the value at 10 m depth. The 10 m reference was chosen because post-processed CTD casts did not always include the surface.

Surface C_T, S_A, and σ_θ for the entire cruise were provided from a Seabird SBE 21 SeaCAT Thermosalinograph (TSG), with SBE 38 thermometer using the ship's continuous surface water intake. Subsequent to post-cruise processing of TSG data

as detailed in Alory et al. (2015), the timeseries was available in 2 min intervals. Additionally, the SedTrap Drifter had on board six SBE 37 Microcat CTDs. Their depths, as determined by mean in situ pressure, were 14, 55, 88, 105, 137, and 197 m. These instruments yielded data every 5 min during their deployments. As mentioned in the previous paragraph, during LDA, the SedTrap Drifter tangled with the Production Line, and so the data presented here from LDA came from its re-deployment

5   until the end of LDA. No gap in the data occurred for LDB or LDC.

2.3.2   Tracer analysis

The central question in this article is whether for each LD deployment the quasi-Lagrangian drifting mooring stayed within a similar physical environment. For our purposes, the physical environment is defined by the water's specific C_T and S_A (T-S) values. Generally, over the upper 200 m of the water column, the depth range of most of our CTD casts, a given profile of

10 T-S values will vary along a curve (Stommel, 1962). This reflects how each profile is made up of increasingly denser layers over depth, each with distinct histories. In some sense, these layers could be considered their own physical micro-environment; however, since the biogeochemical measurements of OUTPACE spanned the entire euphotic zone, the ensemble of these layers, the entire T-S profile, had to be combined to represent the physical environment.

One way to contextualize T-S profiles is to compare them with previous measurements or climatologies of the region. What

15 occurred for OUTPACE, however, was that the profiles of T-S values did not necessarily coincide with the World Ocean Atlas or CSIRO Atlas of Regional Seas (supplementary material, Figs. S1-S2; Boyer et al.,2013; Ridgway et al.,2002). When measures of climatological variability were available, the envelope of min/max T-S values were large enough to preclude distinguishing between different stations during OUTPACE. Another, more standard, method for breaking down T-S measurements involves identifying linear combinations of previously defined water masses associated with specific T-S values (Mackas et al., 1987;

20 Poole and Tomczak, 1999). While this method provides a least-squares best estimate of constituent water types, it requires a priori knowledge and sufficiently distinct, conservative parameters to define these water masses. The method also does not provide a clear basis by which similar profiles will be considered different (e.g., if two profiles have 5% and 10%, respectively, of a given water mass, does this signify a sufficiently different physical environment?). Moreover, the associated water mass eigenvector basis precludes capturing additional, previously unobserved water masses.

25 Given these concerns, and the need to work within the dataset of a single cruise, another approach is needed to condense T-S variability so that physical environments can be distinguished. As mentioned earlier, each density layer in a given profile has its distinctive T-S value, and together these layers constitute the physical environment. Assuming that these density layers were not subject to strong forcing such as diapycnal mixing events or atmospheric effects, their values should have remained constant until isopycnal exchange or diffusion could occur over longer timescales. Treating these density layers as separate

30 entities, variations of T-S along isopycnal surfaces can provide an approach to distinguish physical environments, the goal of our analysis.

Using density as one variable, another is needed to fully describe a water parcel's characteristics, ideally one which is independent of density. Spice, a variable constructed from T-S, is well suited for this purpose. Spice is defined such that hot and salty water is 'spicy', a convention dating to Munk (1981). In the formulation proposed by Flament (2002), its isopleths are

everywhere perpendicular to isopycnals, and it effectively both encapsulates and accentuates T-S variability at a given density into a single value. Therefore, in our analysis, spice variability in a given density layer was used to to determine whether a physical environment changed during OUTPACE.

The statistical determination of whether the physical environment changed during the OUTPACE LD stations relied on a two-step protocol: first, the determination of a spatial scale, R_Z, over which the LD station's physical environment was the same, and second, evaluation of whether that scale was surpassed by the SedTrap Drifter, in situ currents, or the SVP drifters. In order to find the R_Z of the physical environment, a baseline profile of spice-density values was defined for each LD station using the most abundant data source, namely the timeseries of CTD casts spaced every ∼3 h. Independent spice measurements were compared to the baseline using the traditional Z-score for a normally distributed variable. The relationship between Z-score and the horizontal distance between observations was used to determine R_Z, the spatial scale beyond which the physical environment changed. The second step, determining whether R_Z was crossed during each LD station, is detailed in Sect. 2.4.

The statistical baseline at each LD station was established by collecting spice observations into density bins of width 0.1 kg m^{-3}, ranging over 1021.5-1026.3 kg m^{-3} (bin centers). For each density bin, a mean S_{LD} and standard error $SErr_{LD}$ was calculated, retaining only values in the bins where at least 50 observations were available. The means and standard errors in the baseline were used to compare independent spice observations as summarized by a Z-Score, following the general formula

$$Z(\rho) = \frac{S - S_{LD}}{SErr_{LD}} \tag{1}$$

where, for a specific density layer ρ, S is the independent spice observation. Variations of this formula were used depending on the dataset in question. Typically, if multiple observations were available in a given density bin, the mean value was used for S, similar to the process used in defining the baseline S_{LD}. When a dataset spanned multiple isopycnals, the distribution of Z-scores was summarized again by the mean of their absolute magnitude. This two-step averaging was done, as opposed to averaging all the Z-scores in a profile at once, so that individual density layers were on equal footing. The assumption applied in this analysis is that isopycnal layers are independent of each other, and represent new information from different populations. If one layer had significantly more observations than another, a straightforward mean would bias the Z-score toward the value of the layer with more values.

To calculate the Z-score of other LD CTD rosette timeseries, the reference baseline values were used to calculate Z-scores for each density layer, and the mean of this distribution of Z-scores reported. Since there are two possible Z-scores for each LD pair (depending on which is chosen as the reference for the $SErr_{LD}$), the largest of the two was retained. At an SD station, with usually 3-4 CTD profiles, for each density bin the mean Z-score was used across CTD casts, and the mean of the Z-score profile reported. Z-scores from the SedTrap Drifter's six onboard Microcat CTDs are calculated with a 15 min window before and after each CTD cast used in the baseline. Since the SedTrap Drifter data represents the smallest distance scales (ranging 0.3 to 5.6 km for the entire cruise), instead of calculating a Z-score with S_{LD} being the station mean, the matching individual cast data was used. This choice was made to better test whether sampling close to the SedTrap Drifter results in greater variability or not. TSG Z-scores were calculated with data within a 500 km radius of the LD station position, both before and after sampling.

Due to ship transiting, Z-scores were not calculated for a given time period, but rather split into distance bins. For the first 100 km from the LD station position, the bin width was 10 km, and then became 20 km. Since the TSG represents near-surface values and spans few density bins, the mean of each layer's Z-scores was retained as opposed to taking a second mean to combine layers.

5   The spatial scale R_Z for each LD station was found through fitting both linear and exponential models to the Z-score vs distance relationship. Due to the differences between TSG and SD/LD Z-score distributions, various fits were applied to the entire dataset as well as subsets. A conservative version of R_Z was ultimately estimated using the SedTrap Drifter and TSG datasets. The distances at which Z = 2 (α=0.5 rejection level) for both linear and exponential models were averaged together, weighted by their coefficient of determination, or r^2, value. The definition of r^2 used was

$$ r^2 = 1 - \frac{\sum_i^N (Z_i - \hat{Z}_i(x))^2}{\sum_i^N (Z_i - \bar{Z})^2} \tag{2} $$

where Z_i is the Z-score, $\hat{Z}_i$ is the modeled Z-score, and $\bar{Z}$ is the Z-score mean.

A natural spatial scale to serve as a useful reference to the fitted distance is the first Rossby radius of deformation, approximated via Wentzel-Kramers-Brillouin (WKB) method by Chelton et al. (1998) as

$$ R_D = \frac{1}{\pi f} \int\limits_{-H}^{0} N(z) dz \tag{3} $$

15   where f is the local coriolis parameter and $N(z)$ is the depth-dependent Brunt-Väisälä frequency. R_D was calculated for each LD station using the deepest cast available: 2000 m casts for LDA and B, and a deep 5000 m cast for LDC. N was calculated with centered differences of the 1-m binned density profiles.

2.4   Velocities and drifter positions

The in situ velocities for each LD station were derived from the shipboard acoustic Doppler current profilers (SADCP), two
20   Ocean Surveyors at 150 kHz and 38 kHz. Timeseries data for the SADCPs were post-processed using the CASCADE software package (Le Bot et al., 2011; Lherminier et al., 2007) and binned into 2 min intervals. The 150 kHz SADCP provided a depth resolution of 8 m, with bins starting from 20 m, and reliable data coverage down to 200 m depth. Since the SedTrap Drifter had sediment traps extending down to 500 m depth, the 38 kHz data was also used, albeit with reduced depth resolution of 24 m bins, extending from 52 m down to 1000 m. Additional in situ velocities came from six Nortek AQUADOPP current meters,
25   positioned at 11, 55, 88, 105, 135, and 198 m on the SedTrap Drifter. The post-processed AQUADOPP timeseries provided observations every 5 min.

Velocities were integrated using a first-order Euler method to calculate the theoretical trajectories of water subsequent to the beginning of each LD station. Since the object of these calculations was to see whether water could have surpassed a critical spatial scale, for each dataset the maximal amount of time was given for the time integration. SADCP timeseries spanned
30   between the first and last CTD of the LD station, using the ship position as the initial position. The AQUADOPP integrations spanned the entirety of valid data and used the corresponding SedTrap Drifter satellite fix for an initial position.

To compare the integrated velocity positions with the realized positions of the SedTrap drifter and SVP drifters, GPS positioning was achieved by use of Iridium telemetry. Positions were successfully found for LDA before and after the SedTrap Drifter's re-deployment, along with all of LDB. During LDC, the battery of the positioning antenna ran out and so the timeseries for LDC positions of the SedTrap Drifter was shortened. Since only the initial position is needed for the velocity integration, the AQUADOPP integration was continued beyond this positioning failure until the SedTrap Drifter was recovered. Positions of the SVP drifters deployed at each station were successfully retrieved for all three LD stations. Satellite fixes were available spaced about 1 h apart for both datasets. Both SedTrap Drifter and SVP positions were interpolated to hourly timeseries. SVP positions were used to compute relative dispersion (supplementary Fig. S3) using the definition for N particles (LaCasce, 2008),

$$D(t) = \frac{1}{2N(N-1)} \sum_{i \neq j}^{N} [x_i(t) - x_j(t)]^2 \tag{4}$$

where N here is the total number of SVP drifters, and x the timeseries of the drifter i's x, y position.

3 Results

3.1 Satellite data and temporal context

The regional distributions of SST and surface chl-a as seen during the OUTPACE cruise are shown in Fig. 1. The data in Fig. 1 are weighted means, with the weight being the inverse square of the ship's daily distance to each pixel. A north-south meridional gradient was found in SST, with warmer water near the equator ($\sim 30^\circ$C) and cooler water poleward ($\sim 25^\circ$C). This gradient was uninterrupted for the duration of the OUTPACE cruise. Due to the zonal shiptrack the surface thermal conditions observed by the ship during OUTPACE were relatively homogeneous. While a north-south gradient was found in SST, the opposite was found in chl-a. Chl-a values were around 0.3 mg m^{-3} in the western portion of the domain, west of 190° E. Stations LDA and LDB were in this region, with LDB positioned inside a bloom with values near 1 mg m^{-3}. More details concerning the LDB bloom can be found in de Verneil et al. (2017). Chl-a values dropped precipitously, over an order of magnitude to 0.03 mg m^{-3}, just east of LDB near LDC. The low value of chl-a was indicative of the South Pacific Gyre (SPG; Claustre et al., 2008).

Since SST was relatively unchanging during OUTPACE, Fig. 2 provides zoomed-in views of the chl-a data for the three LD stations, with domains chosen to include the nearest SD stations. The spacing of the SD stations was relatively regular along the OUTPACE transect. In Fig. 2a the enhanced chl-a was distributed evenly inside the domain. In Fig. 2b, the chl-a was concentrated in the aforementioned bloom, with values higher than those seen in Fig. 2a near LDA. By contrast, waters outside the bloom had chl-a values lower than in Fig. 2a. The low chl-a values near LDC in Fig. 2c were typical of the SPG.

The timeseries of chl-a and SST for the three stations are shown in Fig. 3. Comparing the three LD stations, a few patterns emerge. SST showed similar trends across the three LD stations. All stations experienced warming trends from December 2014 to mid-March 2015, consistent with summer heating. The lack of data from cloud cover sometimes led to abrupt drops in the

distribution of daily SST shown. The timing of maximum temperature, and the magnitude of that warming, however, did differ between LD stations. A rapid heating in December 2014 occurred around LDA's position, which then slowly continued until early March 2015, at which point temperatures began to drop. Towards the end of sampling at station LDA another warming event occurred. The overall evolution in LDA's temperature during the period shown, from ∼26 to 30° C, represented a 4°

5 change. LDB showed a slight cooling in December 2014, but this may have been an artifact of cloud cover. Station sampling for LDB occurred immediately after the maximum heating, though the values seen at LDB were slightly warmer than at LDA. The maxima in temperature for LDA and LDB seemed to overlap in time, in early March 2015. LDB's change in temperature, from ∼27 to 30° C, was a 3° change. LDC had the smallest change in temperature, from ∼27 to 29 °C for a 2° change. Sampling for LDC coincided with the warmest period observed in the satellite data, in late March 2015.

10 The timing of temperature maxima is important to note for biological reasons, since N_2 fixation by *Trichodesmium* spp. is known to occur in warm, stratified waters (specifically, a ∼25° C threshold, White et al., 2007), and one of the goals of OUTPACE was to observe this biogeochemical process. Since SST was above 25° C for all stations throughout this period, the thermal conditions during OUTPACE would not have limited N_2 fixation.

In between December 2014-January 2015, the region around LDA had higher chl-*a* concentrations than LDB. The period

15 between February and May 2015 showed a remarkable increase in chl-*a* near the LDB site. This was due to advection of the surface bloom, which subsequently collapsed and advected away (de Verneil et al., 2017). Near LDC, chl-*a* was systematically low, a reflection of the goals of OUTPACE to sample in the SPG.

3.2 In situ tracers

The hydrographic variability during the three LD stations and surrounding SD stations are shown in the T-S diagrams of Fig.

20 4. All three stations followed a general pattern, where surface water near the 1022 kg m^{-3} isopycnal and 29° C temperature (though LDB had warmer surface water, Fig. 4b) dropped in temperature and increased in salinity until a subsurface salinity maximum near the 1025 kg m^{-3} isopycnal. The salinity maximum reflects sampling along the latitudinal gradient of the high salinity tongue of the South Pacific (Kessler, 1999). The surface water in LDA (Fig. 4a) showed a bifurcation, where the heating event at the end of LDA sampling seen in satellite data was also manifest. For LDA, neighboring stations SD2, 3, and

25 4 largely overlapped with the LDA profile. SD3, the station closest to LDA, almost entirely overlapped the LD profile, except for a subsurface salinity deviation below the 1024.5 kg m^{-3} isopycnal. SD2 and SD4 showed greater deviations, with SD4 being saltier than LDA for almost the entire profile. Similar overlaps occurred with LDB and its surrounding stations, SD12 and SD13 (Fig. 4b). SD12 showed lower salinity near the surface, with a kink in salinity at the 1025 kg m^{-3} isopycnal. The salinity offsets of SD4 and SD12 at depth are within climatological variability (Figs. S1 and S2). SD13 had similar surface

30 structure to LDB, but higher salinities from the 1023.5 to 1025 kg m^{-3} isopycnal. The LDC, SD13, and SD14 (Fig. 4c) profiles nearly entirely overlapped except near the surface when the SD stations were at first less salty at the surface and then became more salty. Additionally, the saltier nature of LDC relative to LDA and LDB, especially between 1024 and 1025 kg m^{-3}, was visible. The variability in T-S values between stations was within the range seen in the climatology of the region (Figs. S1 and S2).

The LD statistical baselines of spice in density space, with means and intervals of two standard errors, are shown in Fig. 5. These standard error intervals, representing the inherent variability in the baseline, show the values wherein a Z-score of ≤ 2 was achieved. LDB and LDC overlapped for essentially their entire profiles. LDB is missing observations near the surface due to the intense stratification which left several density bins with less than 50 observations, the threshold used in the spice

5   analysis. LDA was noticeably less spicy than the other two LD stations for density less than 1024 kg m^{-3}. In the 1024-1025 kg m^{-3} range, where LDC had high variability in spice, some density layers showed LDC differing from LDA and LDB by more than two standard error intervals. At the highest densities at depth, all three LD stations overlapped.

The Z-scores calculated from the other datasets with regard to the LD statistical baseline are plotted versus their distance from the LD station in Fig. 6. SedTrap Drifter Z-scores, centered on the left hand side of Fig. 6, showed a cloud of Z-scores

10   that does not depend on distance. Overall, these Z-Scores were below 2. The surface TSG data began to show a relationship between Z-score and distance. Starting near 15 km, Z-scores were at ≤ 2. The LDA TSG Z-scores remained below 2 until ~ 100 km distance, after which point the Z-scores quickly increased with distance until they reached the highest observed values of all the datasets near the 500 km cut-off. The Z-scores from the TSG for LDB increased faster, passing 2 between 40 and 50 km and continually increasing thereafter. The LDC TSG Z-scores started out high, having surpassed a Z-score of 2 between

15   30 and 40 km, and continued to increase in the same manner as the LDB TSG data. One should note, however, a large number of LDC TSG Z-scores below and near 2 in the 100 to 300 km range. SD station Z-scores replicated the trends seen in the TSG data, with increasing Z over greater distances. The only SD station within the canonical Rossby radius, SD3 near LDA, had a Z less than 2. Comparing between how the SD stations match up with LD baselines, LDA differed the most from the other stations, given its enhanced Z-scores. The SD Z-scores for LDB and LDC were lower by comparison, and were largely

20   grouped together, though they too displayed an increasing trend in Z-score with distance. The three LD station pairs, LDA-B, LDA-C, and LDB-C, also reflected a trend of increasing Z-score with distance. The closest station pair, LDB-C, had a Z-score less than 2. The other pairings, LDA-B and LDA-C, showed greater differences with Z-scores ~ 10.

The r^2 values for the model fits, and the corresponding distances where the models reach Z=2, are summarized in Table 2. Model fits were the lowest when including all the data. The 'TSG+SedTrap' Drifter model fit was the highest, with the r^2 being

25   in the top two for all model fits across all stations. The distances where the model fits cross Z=2 vary greatly depending upon the dataset. Negative distances occur for most fits with 'all data', possibly due to the influence of the TSG and SD/LD data having high Z-scores at two separate spatial scales, and that the model fits were not constrained to cross the origin. Negative and small distances occur with the 'TSG only' fits, again possibly due to the lack of constraint to cross the origin. Model fits for 'all but TSG' data result in large cut-off distances (all ≥ 400 km). This is possibly a result of Z-scores remaining low for some

30   SD stations, and the lack of data to introduce variability at the spatial scales between the SedTrap Drifter and SD/LD stations (normally spanned by the TSG). The 'TSG+SedTrap' model fits resulted in the most consistent Z=2 distances, between 30-90 km. This minimum of variability is possibly due to the SedTrap Drifter data forcing the model fits to be low at small scales, and that the large-scale TSG values are at spatial scales immediately adjacent to the SedTrap Drifter (TSG values begin at 15 km, as opposed to around 200 km for most SD/LD data points). Due to the fact that the model fits were consistently high, and

35   the relative lack of variability in the cut-off distances, the Z=2 distances for the 'TSG+SedTrap' models were chosen as the

R_Z spatial scale over which the physical environment was the same. Since there was no noticeable preference in r^2 for linear or exponential models, the cut-off distance is the mean weighted by r^2, resulting in R_Z values of: 66 km for LDA, 47 km for LDB, and 54 km for LDC. Therefore, with determination of R_Z, we have completed the first step in our protocol to evaluate whether the physical environment changed during the LD stations.

5   ## 3.3   Velocities and lagrangian trajectories

The second step in our protocol requires the analysis of water velocities and their associated trajectories. Timeseries of the 38 kHz SADCP and AQUADOPP data are presented in Fig. 7. The LDA timeseries of SADCP u and v components (Fig. 7a,d) showed strong near-inertial oscillations in the upper 200 m, with velocities reaching magnitudes of 0.6 m s^{-1}. A weaker tidal component was also present in this layer: below 200 m, vertical columns of alternating velocity sign indicated the semi-diurnal

10   tide. These tidal signatures were also the dominant signal in the LDB and LDC timeseries (Fig. 7b-c,e-f). The mixed layer, which, for most of the cruise, was ≤20 m, was not resolved by either SADCP. So, the near-surface velocities were only captured by the 11 m AQUADOPP and the SVP drifters drogued at 15 m. Comparing the 55 m AQUADOPP timeseries with the 52 m SADCP, the two data sources displayed similar trends for LDA. The strong near-inertial oscillations led to correlations between the AQUADOPP and 38 kHz timeseries of 0.75 and 0.76 for the u and v components, respectively. During LDB

15   and LDC, the weaker currents did not correlate as well, leading to u,v correlations of -0.0137, -0.0554 (LDB), and 0.30, 0.37 (LDC), respectively. For comparison, the 150 kHz 52 m timeseries produced u,v correlations with the AQUADOPP of 0.83, 0.80 (LDA), 0.00, 0.02 (LDB), and 0.68,0.68 (LDC). Vector correlations using the method of Crosby et al. (1993) for the three timeseries (not reported) similarly showed a maximum for LDA, minimum near-zero for LDB, and low values for LDC. These differences likely result from higher frequency fluctuations of the currents, at the inertial and tidal frequencies. The fact that a

20   higher correlation is obtained at LDA is probably partly the consequence of the larger spatial scales of the near-inertial signal dominant at LDA compared to the baroclinic tidal signal, e.g. resulting from the dispersion relation (Alford et al., 2016). These oscillations, and their implications for turbulent mixing, are analyzed in greater detail in Bouruet-Aubertot et al. (this issue).

The disagreement between the two velocity data sources had an impact on the integrated trajectories. Take, for example, a closer inspection of the SADCP and AQUADOPP during LDA, which had the strongest currents. The initial positions of the

25   ship and the SedTrap Drifter were 1.46 km apart. After 3 days and 2 hours, the AQUADOPP integration had traveled 67.75 km, the SADCP 60.71 km, with a final separation of 10.89 km. The result was a positional drift of ∼3 km day^{-1}, or an average increase in position difference of 147 m for each km traveled. A similar analysis for the LDB timeseries, with weaker currents but essentially no correlations over 4 days and 15 hours, resulted in a positional drift of 3.19 km day^{-1}, with an increased position difference of 318 m for each km traveled. Thus, a lower correlation timeseries, but with lower magnitudes, resulted in

30   similar misfit in the integrated trajectory.

The trajectories of the integrated velocities, as well as observations of SedTrap Drifter and SVP positions, are presented in Fig. 8. The average altimetry-derived currents suggested there should be recirculation around the positions of LDA and LDC, whereas LDB had a mean northward flow (Fig. 8 a-c). The SedTrap Drifter trajectory for LDA did not follow the surface altimetry currents and their anticyclonic flow, but instead underwent several oscillations while cruising in a west-

northwest direction (Fig. 8a). The SVP drifters for LDA (Fig. 8g), while also undergoing oscillatory loops, instead drifted to the south. The 38 kHz SADCP velocities showed a transition over depth with shallow water moving south-southwest, but with increasing depth the trajectories flowed northwest in a similar fashion to the SedTrap Drifter. During LDB, the SedTrap Drifter went north-northeast, in agreement with the altimetry currents (Fig. 8b). The SVP drifters moved in a similar fashion, north-

5   northeast, though they ended up undergoing more oscillations and eventually advected more eastward (Fig. 8h). The 38 kHZ SADCP velocities demonstrated that shallow depths flowed east like the SVP drifters, but with depth this advection swung to a more northerly direction (Fig. 8e). The LDC SedTrap positioning was relatively uninformative, since the Iridium satellite fix was unavailable for the second half of the LD station and so showed little displacement (Fig. 8c). The SVP drifters for LDC (Fig. 8i), similar to LDB, underwent several oscillations and were advected the farthest, moving in a southeast direction. The

10 SADCP data showed a shallow flow to the east, similar to the SVPs, but with depth the majority of trajectories oscillated near the station and even flowed southwest.

For all the LD stations, the SedTrap Drifter and integrated velocities stayed within a radius of R_Z centered at the LD starting position. The SVP drifters for LDA and LDB also stayed within the R_Z radius. The SedTrap Drifter, integrated velocities, and SVP drifters also stayed within a Rossby radius R_D of the starting positions of LDA and LDB. R_D was smaller than R_Z for

15 LDA, almost equivalent in size for LDB, and larger for LDC. For LDC, though the SedTrap Drifter and integrated velocities stayed within both R_Z and R_D, the SVP drifters reached R_Z but not R_D at the end of five days. As a result, according to our protocol they may have moved into a new physical environment, and biogeochemical observations at the end of LDC may need to be examined in closer detail for changes associated with a change in water mass. However, for the greater majority of LDC, this should not be a concern. Thus, we conclude that LDA and LDB sampling was conducted in a single physical environment,

20 whereas LDC was mostly in a single environment except perhaps at the very end.

4   Discussion

4.1   Tracer analysis and spatial scale determination

The main goal of this study is to determine whether the quasi-Lagrangian sampling strategy during the LD stations of OUT-PACE was successful, namely by remaining within a single physical environment. The motivation behind this exercise is to

25 independently evaluate if the biogeochemical measurements of OUTPACE represented a single biogeochemical milieu, rather than the advection of the SedTrap Drifter into a different area. Evaluation of the strategy is grounded in the variability of T-S and analysis of water velocities.

Before the spice analysis was conducted, the initial context of SST and chl-*a* variability at the surface in space and time was provided by satellite products. At the regional scale of the WTSP, the LD stations were roughly positioned within the zonal

30 gradient of chl-*a* and meridional gradient in SST (Fig. 1). The gradients of surface chl-*a* around the LD stations were minimized in relation to the regional-scale gradients, partly by design in the process of choosing station locations (Fig. 2). The temporal trends of SST and chl-*a* largely reflected the seasonal cycle: chl-*a* was decreasing at the end of the summer in the MA and low values dominated in the SPG; SST reached its peak due to late summer timing (Fig. 3). The timing of temperature maxima is

important to note, as mentioned in Sect. 3.1, since N_2 fixation by *Trichodesmium* spp. is known to occur in warm waters, and one of the goals of OUTPACE was to observe this biogeochemical process. While these satellite data are sufficient to identify large-scale structures or temporal trends, the LD stations by design were in regions where it is difficult to judge whether the SedTrap Drifter stayed in the same water mass from these surface data alone. In light of this, these data are insufficient for our

5 needs in validating the quasi-Lagrangian strategy. The use of remote sensing data to help identify small-scale structures during OUTPACE outside of the LD stations is further explored in Rousselet et al. (this issue).

The depth-resolved in situ T-S data (Fig. 4) helped to capture some of the variability present during the LD and surrounding SD stations. The T-S structures showed consistent values during LD stations with deviations observed in the neighboring SD stations. As with the satellite data, the in situ data, although informative, provide qualitative interpretation. The traditional

10 quantitative methods for breaking down tracer data require identified water masses (Mackas et al., 1987; Poole and Tomczak, 1999). In some regions of the ocean, these methods are difficult to apply. Sometimes well-defined water masses are not found or difficult to distinguish. Also, maybe the full complement of tracer data (dissolved O_2, nutrients, etc.) cannot be used (like in this study) because they are liable to rapidly change in the euphotic zone due to biogeochemical processes. Additionally, local mesoscale activity can contribute to variability, as has been seen in the WTSP (Rousselet et al., 2016). In that study,

15 O_2 measurements were the distinguishing tracer, which we are precluded from using. As a result of all these factors, another quantitative method that works within the dataset of a single cruise in the WTSP is needed.

The quantitative approach used in this study leverages the large quantity of in situ T-S data available from multiple platforms to condense the physical variability present in the WTSP during OUTPACE over 4000 km distance during austral summer 2015. In order to do this, the statistical baseline in spice was defined (Fig. 5). In effect, as opposed to an absolute measure

20 (i.e. specific water mass determination), this provides a relative measure of variability that can be used. The spice baselines for all the LD stations yielded similar relationships between distance and variability (summarized as Z-scores) when compared to complementary datasets (Fig. 6). Overall, variability increased with distance, as one would logically expect. The rapid increase in Z-score above 2 (using an $\alpha = 0.5$ criterion) led to the determination of the scale R_Z to define a cut-off point at which the physical environment began to change. These distances were of the same order of magnitude as the Rossby radius R_D.

25 The limitation in this approach is that the measure is now dependent on how applicable a statistical baseline is to the given dataset, in our case the LD CTD timeseries. At some level, this has to be a subjective decision based upon the data at hand. For OUTPACE, creating the statistical baseline appears most reasonable for stations LDA and LDB due to the concentration of data over the ensemble of profiles (Fig. 4a-b). During LDC, it is clear that the mid-water spice variation, mostly due to salinity, may make this assumption problematic. Unless there is a reason to a priori remove such data (e.g. if there were visible

30 changes in the deep chl-*a* maximum during LDC), then this variability must be incorporated into the baseline definition, and it will affect the Z-scores. The combination of multiple density layers, however, should provide a more robust comparison of how similar an entire T-S profile is, and not just the variability present in a single density range. A similar phenomenon where variability is introduced to the baseline occurs with the surface heating at the end of LDA. The resulting spice baseline has enlarged standard error near the surface (Fig. 5), which possibly lowered the TSG Z-scores below 2 for up to 150 km. This

produced an R_Z larger than R_D for LDA, the only of the three LD stations to do so. Measuring only the surface instead of multiple density layers, the TSG data are probably the most sensitive to this phenomenon.

The example of surface heating at LDA introducing unwanted variation brings up another assumption in our analysis: we used spice hypothesizing that there was no diapycnal forcing. Clearly, at the surface, atmospheric forcing can influence the
5   water's T-S (and spice) characteristics, and so will impact TSG measurements as well as observations in the upper mixed layer. Future applications of this method will have to take this variability into account, and perhaps make greater use of survey data to fill in the spice variability below the surface at these spatial scales. At depth, however, the greatest source of along-isopycnal gradients in T-S, i.e. density-compensated features, is mesoscale stirring (Smith and Ferrari, 2009), and so, generally, the assumption should be applicable. Mesoscale eddies and fronts are the physical structures most likely to impact biogeochemical
10   observations in a given field campaign, and it is probably the reason that both Z-scores across the board begin to increase and the R_Z's were found to be at or near the Rossby radius R_D for each LD station. Granted, if ship sampling happens to be placed immediately next to a strong eddy or filament, Z-scores could increase over much shorter distances, though this was not observed during OUTPACE. Therefore, at first order, the Rossby radius R_D serves as a default scale at which the integrated history of previous mesoscale stirring will on average manifest itself. In situ data and further analysis is needed, however, to
15   verify whether smaller-scale variability is large through determination of R_Z. One recommendation that emerges from these results is that, if future field measurements are to take place in an area devoid of obvious mesoscale structures, the Rossby radius R_D quickly calculated from a deep CTD cast may be useful in determining at which point a quasi-Lagrangian drift array should be recovered as a precautionary measure.

The choice of spice as a variable, though useful, is not a magical transformation in itself. The similarity in T-S between
20   LDB and LDC is still manifest in spice-density space (Fig. 5). In fact, by our own metrics, since stations LDB and LDC had a Z-score less than 2, left with only this measurement the two stations were indistinguishable from each other. Only through comparison with other datasets does it become clear that variability does exist between and around the two stations, namely the SD stations surrounding LDB and LDC. This observation makes clear that independent data sources are a key requirement in this method. Having enough data to span a sufficient spatial range is what will determine what differences in spice are relevant
25   or not. Indeed, having enough data can help validate whether application of the spice baseline method in itself is insightful. For instance, if the SedTrap Drifter CTD data, representing the smallest scales, had displayed larger Z-scores, then the conclusion would have been that there was large variability right next to the ship somehow missed by the CTD baseline. Instead, by having limited variability (Z≤2) that did not visibly depend on distance, the SedTrap Z-score distribution validated the spatial scales at which the LD CTD timeseries was sampled and its use in creating a statistical baseline.

30   Having considered some of the caveats and assumptions implicit in our present approach, we feel that its application for the OUTPACE campaign was warranted and subsequently validated by the consistency of the results, both in consideration of the multiple data sources concerned but also of the theory of mesoscale circulation. In our application, a conservative R_Z was used based on surface, and not depth-resolved, data at intermediate scales. Ideally, future applications of this method could hopefully use depth-resolved measurements to see when surface layers diverge as opposed to deeper ones, which might be
35   reflective of different turbulence regimes such as the submesoscale.

4.2 Integrated velocities and drifter trajectories

The analysis of the integrated in situ velocities, alongside the SedTrap Drifter and SVP positions, was used to determine whether certain parcels of water had advected far enough during each LD station's sampling to move beyond R_Z, the chosen spatial scale. The conclusion for the OUTPACE cruise is that almost all of these proxies for Lagrangian pathways did so, and

5 hence the changes observed in the upper water column during the LD stations were probably due to biogeochemical processes alone. Nevertheless, given the multiple sources of trajectories available, there is no clear interpretation as to which of them is the most 'truly Lagrangian.'

The quasi-Lagrangian drift array, spanning the top 100's of meters of the water column, with various instruments having varying degrees of drag strewn at different depths, will never be purely Lagrangian. Unless a drift array is deployed in a

10 flow that is entirely barotropic, there will likely be sources of shear in the velocity field from the mesoscale flow and from baroclinic near-inertial waves and internal tides (see Bouruet-Aubertot et al., this issue). By definition, vertical shear leads to different velocities at different depths. The net effect of these velocities results in the drift array's observed trajectory, with water at some depths moving slower or faster than the array itself. As a result of this inevitable decoupling of in situ and drifter velocities, the water passing the drifter can come from entirely different areas. Take, for example, the near-inertial

15 oscillations during LDA (Fig. 7a,d). The oscillations influenced the top 200m, so immediately the top sediment trap at 150m experienced vigorous currents whereas the bottom two at 250 and 500m did not. This complicates the interpretation of where the falling organic matter comes from. Strong vertical structuring of the phytoplankton community further complicates this picture. In oligotrophic areas, light-adapted organisms are found near the surface with dark-adapted organisms near the deep chl-a maximum (not to mention the near-surface N_2 fixation occurring during OUTPACE). During LDA, the shear resulted in

20 bands of instantaneously opposing velocities in layers about 50 m thick. Therefore, organisms sampled by a CTD near the chl-a max at 80 m would be traveling in a direction opposite to the organisms at the surface. Immediately, is seems, the expectation of being Lagrangian is lost. However, whether this is an irrevocable loss depends upon the nature of the motions: are they associated with internal waves, such as LDA, or not? If shearing is wave-induced, then after one period, water should return to its original position. Thus, sampling over at least one wave period, and within corresponding horizontal wavelengths, could

25 help to preserve the physical environment. If motions are not related to oscillatory processes, then the only recourse would be that biological communities (and the state of biogeochemical processes underway) are contiguous enough in horizontal extent that they reflect the physical spatial scales found in the tracer analysis of this study.

The potential ability of individual shearing layers to remain coherent over long timescales during OUTPACE is reflected by the SVP drifters during LDC. While they had traveled far, surpassing R_Z and near R_D by the end of sampling, they remained

30 close to each other. This is also reflected in the SVP relative dispersion timeseries (Fig. S3). Typically, relative dispersion increases in the first few days in an exponential fashion, until the drifters begin to be influenced by mesoscale structures (LaCasce, 2008). For the OUTPACE SVP drifters, the exponential phase (a linear increase in log-space) does not abate in this timeframe. Rather than an elevated shearing apart, however, the rate of exponential growth was so low that it took over a week for the relative dispersion to arrive near 1×10^8 m^2. Thus, while the SVP drifters during LDC advected past R_Z and near R_D,

their spread was less than this, indicating that the surface layer in LDC was more coherent than comparing between integrated trajectories over different depths would suggest.

In addition to the sheared layers advecting past the drift array, what determines the trajectory of the array itself is difficult to establish. During LDA, with vigorous currents in the upper 200m, the SedTrap Drifter advected north-west, mirroring the
5 deep SADCP trajectories. By contrast, during LDB, with weaker currents, the SedTrap Drifter moved in a similar fashion to the more shallow SADCP and SVP paths. The LDC telemetry, having been cut short, is not useful here. Perhaps hidden in the data is some linear combination of SADCP and SVP trajectories that can best explain why the SedTrap Drifter moved as it did. Although perhaps useful for OUTPACE, this would be fitting a model to a particular drifter configuration, and remove future flexibility in drifter design. Hence, we chose not to pursue this calculation.

10 The different trajectories taken between the AQUADOPP, SADCP, and SVP demonstrate how trying to follow a Lagrangian perspective can get quickly complicated within a few days, if not sooner (Ohlmann et al., 2017). Even measurements at similar depths between different instruments do not correspond, such as between the AQUADOPP and SADCP at 55 and 52m, respectively, for LDB. Where they do correspond, such as for LDA, this may be due to a strong signal, which is also problematic due to the shear involved. As mentioned in Sect. 3.3, the drift between AQUADOPP and SADCP during LDA and LDB were
15 comparable, despite the differing correlations, due to the fact that currents were stronger at LDA. Additionally, while both AQUADOPP and SADCP integrations started from a chosen initial position, the timeseries data reflect measurements taken on moving platforms, i.e. the SedTrap Drifter and the ship, respectively. All these complicating factors considered, the conservative route, employed here, is to consider each source of data and give it the best chance to refute the assumption that advection was weak enough that the spatial boundary determined by tracer analysis was not crossed. This procedure is the recommendation
20 the authors provide subsequent to having done this analysis: while no individual trajectory will likely be the right one, exploring enough of the variability between them should give an idea of whether other water masses are present.

5 Conclusions

The methodology applied in analyzing the tracer data for OUTPACE is reflective of a few characteristics specific to this region and to biogeochemical datasets more generally. First, in an ideal situation, the T-S analysis would determine the component
25 water masses. While this might be achievable for some well-studied regions of the world ocean, this is not generally applicable. Second, alternative well-defined water mass properties, such as dissolved oxygen, tend to be applicable only at depth, such as through the thermocline. Naturally, this is because photosynthesis, respiration, air-sea surface flux, and a panoply of other biogeochemical and physical processes occur near the surface and impact these tracers, making them not conservative. Unfortunately, we are precisely interested in analyzing the water properties of the surface ocean where these processes are most
30 intense. As a result, methodologies need to be developed so that standard T-S measurements, unaffected by biogeochemical processes, can be used to quantify the effect of the physical circulation upon the biological environment encountered in a field campaign.

The multiple in situ data sources compiled during the LD stations of the OUTPACE cruise allowed for the determination of whether the ship, and its associated quasi-Lagrangian drifting array, sampled the same physical environment. The procedure used to do this consists of two steps. First, the T-S variability from a single source is transformed into spice-density coordinates and used to establish an adequate statistical baseline. Comparison of independent T-S data to the baseline is used to calculate

5   Z-scores and establish a spatial scale R_Z beyond which the T-S differences amount to a change in the physical environment. For the OUTPACE cruise, this scale was found to be close to the Rossby radius R_D, in the 45-65 km range. The second step is to then use all available data regarding currents and drifter positions to evaluate whether any water parcel could ostensibly have traveled farther than R_Z. During OUTPACE, we conclude that largely this did not occur.

The methodology used in this study provides a framework wherein readily available T-S data can be used to answer the
10 same question (whether a single physical environment was sampled following a quasi-Lagrangian drifting mooring) for other oceanographic cruises. More traditional methods, such as absolute water mass determination, or using alternative tracers such as dissolved oxygen, require prior knowledge of a given region or are not applicable in the euphotic zone where biogeochemical measurements are made. While sampling in a Lagrangian manner is preferable to not attempting to follow a water parcel at all, the inevitable failure to be truly Lagrangian with these platforms should be recognized so that experiments are not allowed
15 to either last too long or be deployed in an inappropriate flow regime. Regarding the use of this methodology, we give a few recommendations for future cruise sampling:

–   Maximize use of remote sensing data during the cruise to identify possible mesoscale features to either avoid or sample inside of. This can be achieved with software such as SPASSO (Petrenko et al., 2017).

–   Upon arrival at the selected site, a deep CTD cast below the thermocline can be used to quickly calculate the local R_D
20   Rossby radius in real time and produce a rough estimate for maximum spatial scale.

–   Before and after each station, sample with a surveying instrument such as ISIIS, SWIMS, SeaSoar, or MVP beyond R_D to get depth-resolved data at intermediate scales.

–   If possible, mount CTDs and current meters on the quasi-Lagrangian drifting array (perhaps a sediment trap that does not need to be removed constantly). Multiple independent observations over a large range of spatial scales are essential
25   to calculate robust R_Z estimates.

–   Deploy surface drifters to compensate for the lack of SADCP observations near the surface and to provide spatial context beyond the research vessel.

References

Alford, M. H., MacKinnon, J. A., Simmons, H. L., and Nash, J. D.: Near-inertial internal gravity waves in the ocean, Annual review of marine science, 8, 95–123, 2016.

Alory, G., Delcroix, T., Téchiné, P., Diverrès, D., Varillon, D., Cravatte, S., Gouriou, Y., Grelet, J., Jacquin, S., Kestenare, É., et al.: The French contribution to the voluntary observing ships network of sea surface salinity, Deep Sea Research Part I: Oceanographic Research Papers, 105, 1–18, 2015.

Bograd, S. J., Checkley, D. A., and Wooster, W. S.: CalCOFI: A half century of physical, chemical, and biological research in the California Current System, 2003.

Bouruet-Aubertot, P., Cuyper, Y., Le Goff, H., Rougier, G., de Verneil, A., Doglioli, A., Picheral, M., Yohia, C., Caffin, M., Lefèvre, D., Petrenko, A., and Moutin, T.: Longitudinal contrast in small scale turbulence along 20° S in the Pacific Ocean: origin and impact on biogeochemical fluxes, Biogeosciences, This issue, this issue.

Boyer, T., Antonov, J., Baranova, O., Coleman, C., Garcia, H., Grodsky, A., Johnson, D., Locarnini, R., Mishonov, A., O'Brien, T., Paver, C., Reagan, J., Seidov, D., Smolyar, I., and Zweng, M.: World Ocean Databse 2013, NOAA Atlas, doi:10.7289/V5NZ85MT, 2013.

Chelton, D. B., Deszoeke, R. A., Schlax, M. G., El Naggar, K., and Siwertz, N.: Geographical variability of the first baroclinic Rossby radius of deformation, Journal of Physical Oceanography, 28, 433–460, 1998.

Claustre, H., Sciandra, A., and Vaulot, D.: Introduction to the special section bio-optical and biogeochemical conditions in the South East Pacific in late 2004: the BIOSOPE program, Biogeosciences, 5, 679–691, doi:10.5194/bg-5-679-2008, https://www.biogeosciences.net/5/679/2008/, 2008.

Crosby, D., Breaker, L., and Gemmill, W.: A proposed definition for vector correlation in geophysics: Theory and application, Journal of Atmospheric and Oceanic Technology, 10, 355–367, 1993.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, Nature, 440, 165, 2006.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, Journal of Geophysical Research: Oceans, 109, 2004.

de Verneil, A., Rousselet, L., Doglioli, A. M., Petrenko, A. A., and Moutin, T.: The fate of a southwest Pacific bloom: gauging the impact of submesoscale vs. mesoscale circulation on biological gradients in the subtropics, Biogeosciences, 14, 3471, 2017.

Doglioli, A. M., Nencioli, F., Petrenko, A. A., Rougier, G., Fuda, J.-L., and Grima, N.: A software package and hardware tools for in situ experiments in a Lagrangian reference frame, Journal of Atmospheric and Oceanic Technology, 30, 1940–1950, 2013.

d'Ovidio, F., Della Penna, A., Trull, T. W., Nencioli, F., Pujol, M.-I., Rio, M.-H., Park, Y.-H., Cotté, C., Zhou, M., and Blain, S.: The biogeochemical structuring role of horizontal stirring: Lagrangian perspectives on iron delivery downstream of the Kerguelen plateau, Biogeosciences, 12, 5567–5581, 2015.

Falkowski, P., Scholes, R., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Högberg, P., Linder, S., et al.: The global carbon cycle: a test of our knowledge of earth as a system, science, 290, 291–296, 2000.

Flament, P.: A state variable for characterizing water masses and their diffusive stability: spiciness, Progress in Oceanography, 54, 493–501, 2002.

Freeland, H.: A short history of Ocean Station Papa and Line P, Progress in Oceanography, 75, 120–125, 2007.

Gruber, N. and Galloway, J. N.: An Earth-system perspective of the global nitrogen cycle, Nature, 451, 293, 2008.

Karl, D. M. and Lukas, R.: The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation, Deep Sea Research Part II: Topical Studies in Oceanography, 43, 129–156, 1996.

Kessler, W. S.: Interannual variability of the subsurface high salinity tongue south of the equator at 165 E, Journal of physical oceanography, 29, 2038–2049, 1999.

5   LaCasce, J.: Statistics from Lagrangian observations, Progress in Oceanography, 77, 1–29, 2008.

Landry, M. R., Ohman, M. D., Goericke, R., Stukel, M. R., and Tsyrklevich, K.: Lagrangian studies of phytoplankton growth and grazing relationships in a coastal upwelling ecosystem off Southern California, Progress in Oceanography, 83, 208–216, 2009.

Le Bot, P., Kermabon, C., Lherminier, P., and Gaillard, F.: CASCADE V6. 1: Logiciel de validation et de visualisation des mesures ADCP de coque, 2011.

10   LEFE-CYBER: http://www.obs-vlfr.fr/proof/cruises.php.

Lherminier, P., Mercier, H., Gourcuff, C., Alvarez, M., Bacon, S., and Kermabon, C.: Transports across the 2002 Greenland-Portugal Ovide section and comparison with 1997, Journal of Geophysical Research: Oceans, 112, 2007.

Lumpkin, R. and Pazos, M.: Measuring surface currents with Surface Velocity Program drifters: the instrument, its data, and some recent results, Lagrangian analysis and prediction of coastal and ocean dynamics, pp. 39–67, 2007.

15   Mackas, D. L., Denman, K. L., and Bennett, A. F.: Least squares multiple tracer analysis of water mass composition, Journal of Geophysical Research: Oceans, 92, 2907–2918, 1987.

Marty, J.-C., Chiavérini, J., Pizay, M.-D., and Avril, B.: Seasonal and interannual dynamics of nutrients and phytoplankton pigments in the western Mediterranean Sea at the DYFAMED time-series station (1991–1999), Deep Sea Research Part II: Topical Studies in Oceanography, 49, 1965–1985, 2002.

20   McDougall, T. J. and Barker, P. M.: Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox, SCOR/IAPSO WG, 127, 1–28, 2011.

Moutin, T. and Bonnet, S.: OUTPACE cruise, RV L'Atalante, doi:10.17600/15000900, 2015.

Moutin, T., Van Wambeke, F., and Prieur, L.: Introduction to the Biogeochemistry from the Oligotrophic to the Ultraoligotrophic Mediterranean (BOUM) experiment, Biogeosciences, 9, 3817, 2012.

25   Moutin, T., Doglioli, A. M., de Verneil, A., and Bonnet, S.: Preface: The Oligotrophy to the UlTra-oligotrophy PACific Experiment (OUTPACE cruise, 18 February to 3 April 2015), Biogeosciences, 14, 3207–3220, doi:10.5194/bg-14-3207-2017, https://www.biogeosciences.net/14/3207/2017/, 2017.

Munk, W.: Internal waves and small-scale processes, Evolution of physical oceanography, 291, 1981.

Ohlmann, J., Molemaker, M., Baschek, B., Holt, B., Marmorino, G., and Smith, G.: Drifter observations of submesoscale flow kinematics in

30   the coastal ocean, Geophysical Research Letters, 44, 330–337, 2017.

Petrenko, A. A., Doglioli, A. M., Nencioli, F., Kersalé, M., Hu, Z., and d'Ovidio, F.: A review of the LATEX project: mesoscale to submesoscale processes in a coastal environment, Ocean Dynamics, 67, 513–533, 2017.

Poole, R. and Tomczak, M.: Optimum multiparameter analysis of the water mass structure in the Atlantic Ocean thermocline, Deep Sea Research Part I: Oceanographic Research Papers, 46, 1895–1921, 1999.

35   Ridgway, K., Dunn, J., and Wilkin, J.: Ocean interpolation by four-dimensional least squares - Application to the waters around Australia, J. Atmos. Ocean. Tech., 19, 1357–1375, 2002.

Rousselet, L., Doglioli, A., Maes, C., Blanke, B., and Petrenko, A.: Impacts of mesoscale activity on the water masses and circulation in the Coral Sea, Journal of Geophysical Research: Oceans, 121, 7277–7289, 2016.

Rousselet, L., deVerneil, A., Doglioli, A., Petrenko, A., Duhamel, S., Maes, C., and Blanke, B.: Overview of the large to submesoscale surface circulations and their implications on biogeochemical/biological horizontal distribution during the OUTPACE cruise (SouthWest Pacific), Biogeosciences, this issue.

Schroeder, E. and Stommel, H.: How representative is the series of Panulirus stations of monthly mean conditions off Bermuda?, Progress in Oceanography, 5, 31 – 40, doi:http://dx.doi.org/10.1016/0079-6611(69)90025-1, http://www.sciencedirect.com/science/article/pii/0079661169900251, 1969.

Send, U., Weller, R. A., Wallace, D., Chavez, F., Lampitt, R., Dickey, T., Honda, M., Nittis, K., Lukas, R., McPhaden, M., et al.: OceanSITES, 2010.

Smith, K. S. and Ferrari, R.: The production and dissipation of compensated thermohaline variance by mesoscale stirring, Journal of Physical Oceanography, 39, 2477–2501, 2009.

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., and Knap, A. H.: Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry, Deep Sea Research Part II: Topical Studies in Oceanography, 48, 1405–1447, 2001.

Stommel, H.: On the cause of the temperature-salinity curve in the ocean, Proceedings of the National Academy of Sciences, 48, 764–766, 1962.

White, A. E., Spitz, Y. H., and Letelier, R. M.: What factors are driving summer phytoplankton blooms in the North Pacific Subtropical Gyre?, Journal of Geophysical Research: Oceans, 112, 2007.

Biogeosciences
Discussions

Open Access

EGU

Table 1. Start and stop times for the timeseries data in each LD station of OUTPACE. Times expressed in DD/MM/YYYY HH:MM:SS (GMT) format. When multiple instruments were deployed or multiple discrete observations made, this is also noted.

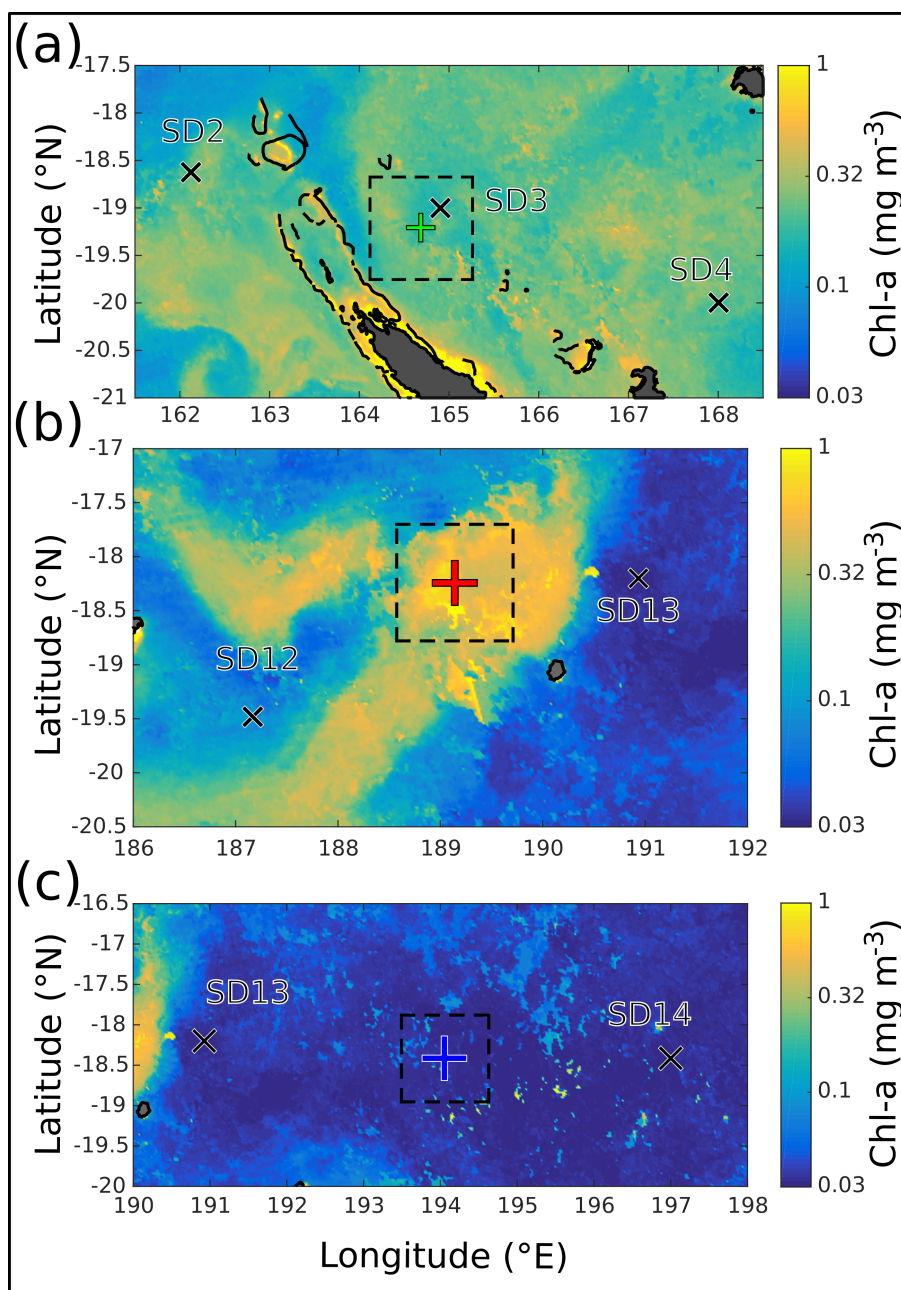| STATION				
LDA	**Instrument**	**CTD Rosette**	**SVP**	**AQUADOPP**
Longitude:	#	46 casts	3 drifters	6 deployed
123.34° E	Start	2/25/2015 14:09:18	2/25/2015 20:00:00	2/26/2015 22:40:00
Latitude:	Stop	3/02/2015 16:10:10	3/02/2015 22:00:00	3/02/2015 16:10:00
18.32° S	Duration	5d 2h 0m 52s	5d 2h	3d 17h 30m
	Instrument	**SADCP 150**	**SADCP 38**	**SedTrap Position**
	Start	2/25/2015 14:09:57	2/25/2015 14:10:26	2/25/2015 19:01:13
	Stop	3/02/2015 16:09:51	3/02/2015 16:08:38	3/02/2015 22:00:00
	Duration	5d 1h 59m 54s	5d 1h 58m 12s	5d 2h 58m 47s
LDB	**Instrument**	**CTD Rosette**	**SVP**	**AQUADOPP**
Longitude:	#	47 casts	6 drifters	6 deployed
123.34° E	Start	3/15/2015 12:04:44	3/15/2015 10:00:00	3/15/2015 23:10:00
Latitude:	Stop	3/20/2015 14:16:13	3/20/2015 23:00:00	3/20/2015 14:15:00
18.32° S	Duration	5d 2h 11m 29s	5d 13h	4d 15h 5m
	Instrument	**SADCP 150**	**SADCP 38**	**SedTrap Position**
	Start	3/16/2015 08:51:53	3/15/2015 23:06:31	3/15/2015 12:15:48
	Stop	3/20/2015 14:15:54	3/20/2015 14:14:50	3/20/2015 21:00:00
	Duration	4d 5h 24m 1s	4d 15h 8m 19s	5d 8h 44m 12s
LDC	**Instrument**	**CTD Rosette**	**SVP**	**AQUADOPP**
Longitude:	#	46 casts	4 drifters	6 deployed
123.34° E	Start	3/23/2015 23:10:57	3/23/2015 12:00:00	3/23/2015 23:25:00
Latitude:	Stop	3/28/2015 14:32:30	3/28/2015 22:00:00	3/28/2015 14:30:00
18.32° S	Duration	4d 15h 21m 33s	5d 10h	4d 15h 5m
	Instrument	**SADCP 150**	**SADCP 38**	**SedTrap Position**
	Start	3/23/2015 12:08:12	3/23/2015 12:06:55	3/23/2015 12:19:55
	Stop	3/28/2015 14:30:31	3/28/2015 14:31:39	3/26/2015 03:31:34
	Duration	5d 2h 22m 19s	5d 2h 24m 44s	2d 15h 11m 39s

Table 2. r^2 values and Z=2 distances for the linear and exponential model fits of each LD station.

STATION	TSG only	All but TSG	TSG+SedTrap	All data
LDA				
Linear				
r^2	0.79	0.85	0.81	0.13
$Dist(Z=2)$ (km)	96	437	46	-344
Exponential				
r^2	0.79	0.87	0.83	-0.32
$Dist(Z=2)$ (km)	0.3	809	85	-7.4×10^3
LDB				
Linear				
r^2	0.78	0.60	0.90	0.18
$Dist(Z=2)$ (km)	21	658	37	89
Exponential				
r^2	0.77	0.56	0.81	0.04
$Dist(Z=2)$ (km)	-30	900	57	-990
LDC				
Linear				
r^2	0.54	0.68	0.73	0.25
$Dist(Z=2)$ (km)	43	726	56	-355
Exponential				
r^2	0.50	0.72	0.67	0.10
$Dist(Z=2)$ (km)	-49	1.05×10^3	51	-3×10^3

Biogeosciences
Discussions

Open Access

EGU
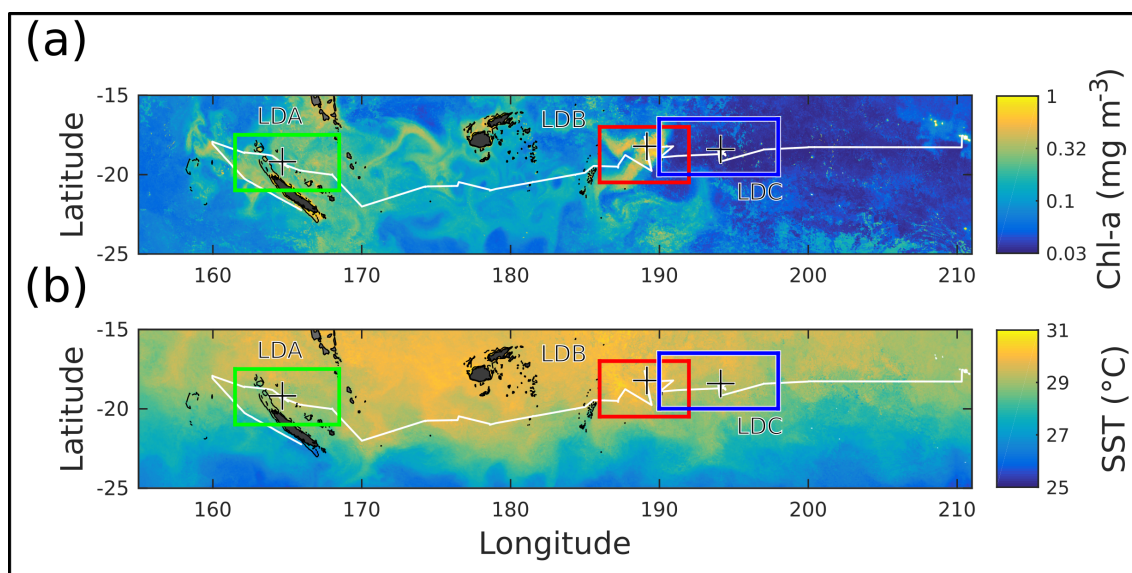
Figure 1. Satellite surface (a) chl-*a* and (b) SST for the OUTPACE cruise. Pixel data are weighted by the normalized inverse distance squared between each pixel and the RV *L'Atalante*'s daily position over the 42 days of OUTPACE. Shiptrack shown in white. LD station locations shown with black +'s. Domains used in Fig. 2 are shown by color-coded rectangles, with green for LDA, red for LDB, and blue for LDC.

Figure 2. Satellite chl-*a* around (a) LDA, (b) LDB, and (c) LDC. SD stations shown by black x's, land is shaded gray, and coastlines/reefs plotted in black. LD stations shown by +'s following the color code from Fig. 1. Squares with 120 km to a side plotted around each LD station to represent approximate Rossby radius R_D.
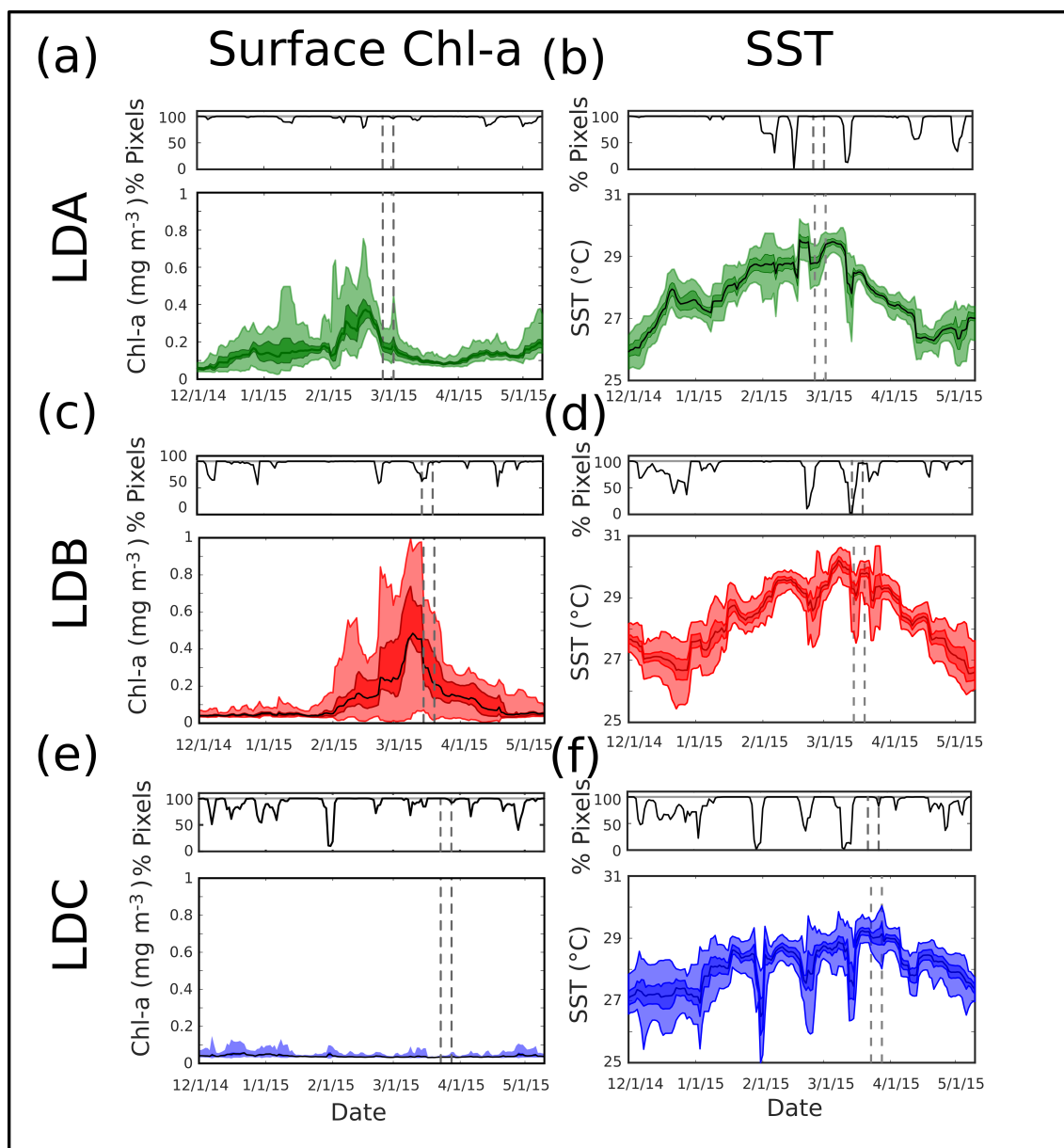
Figure 3. Timeseries of surface chl-*a* and SST, respectively, for (a,b) LDA, (c,d) LDB, and (e,f) LDC. Intervals of LD sampling shown with gray dashed lines. Mean values are plotted in black, with darker shades representing the 25-75% interval and lighter shades for 1-99%. Subpanels above each timeseries depict the % of pixels with data. All data come from within the 120 km × 120 km squares shown in Fig. 2.
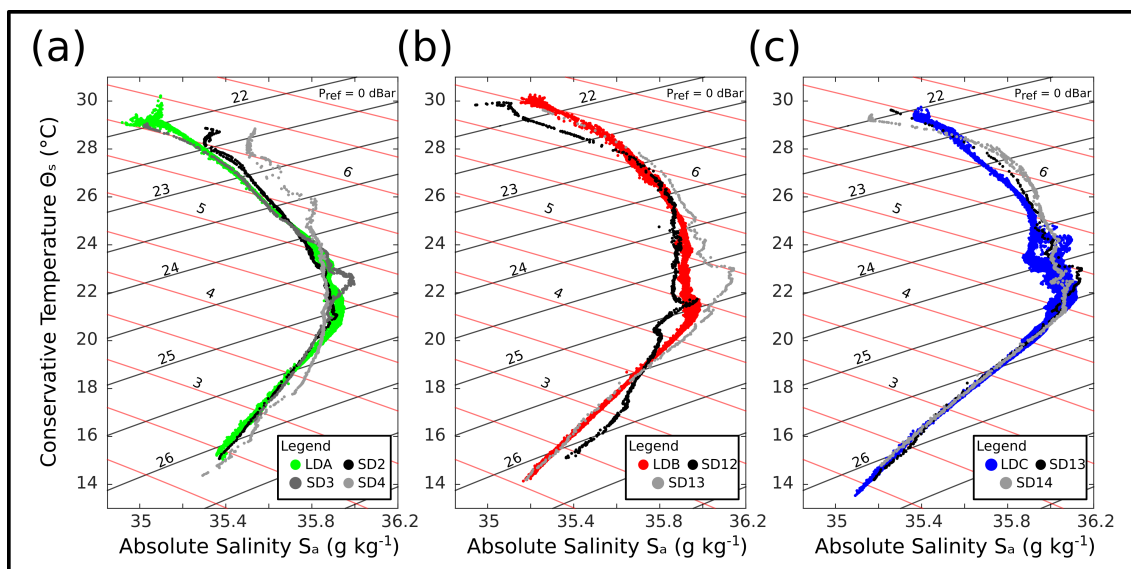
Figure 4. T-S diagrams of (a) LDA, (b) LDB, and (c) LDC and surrounding stations. LD stations are color-coded, and SD stations different shades of gray. Isopycnals are displayed in black, with isopleths of spice shown in red.
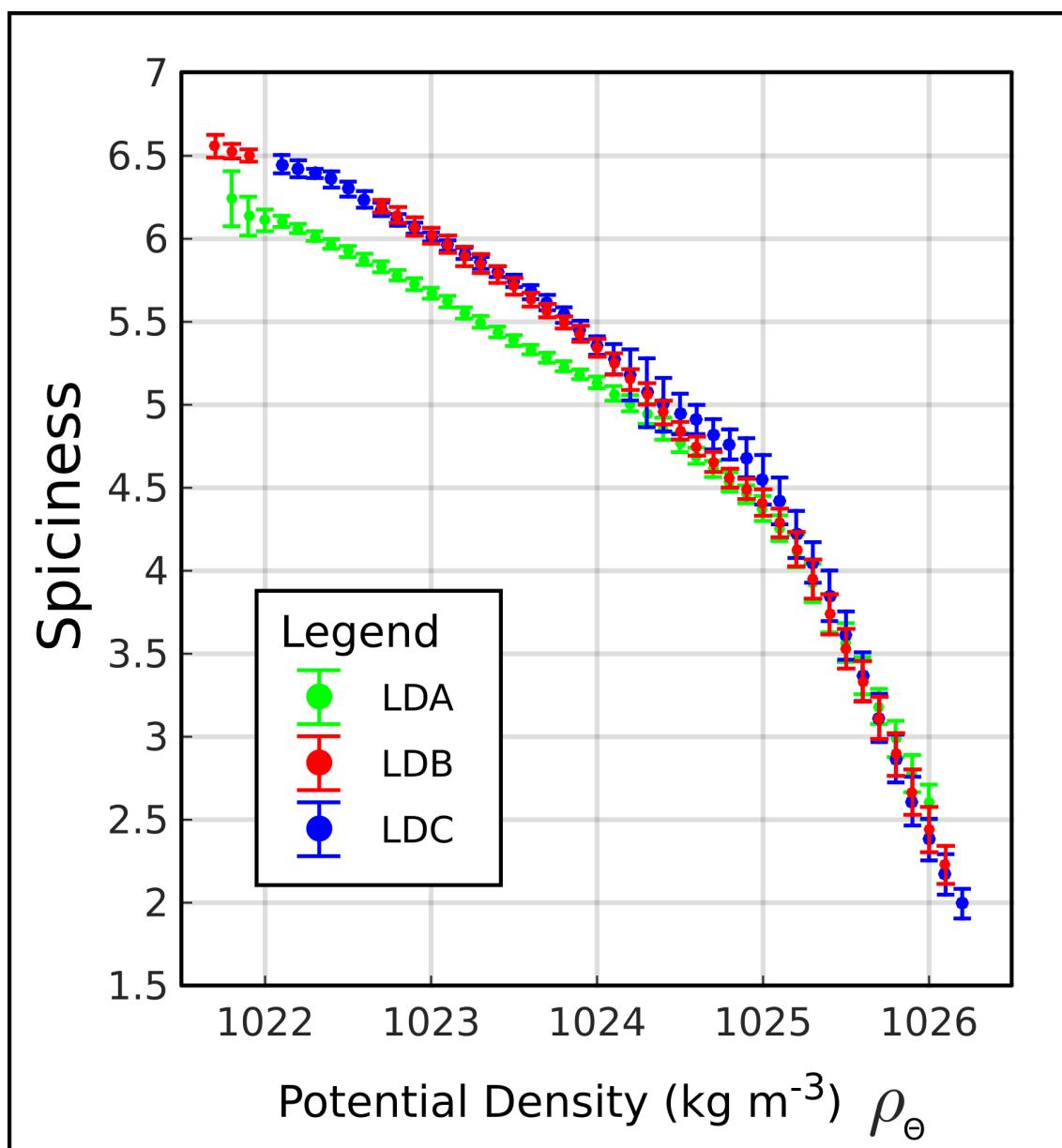
Figure 5. Statistical LD baseline of spice versus potential density. Mean values (dots) plotted with intervals of two standard error.
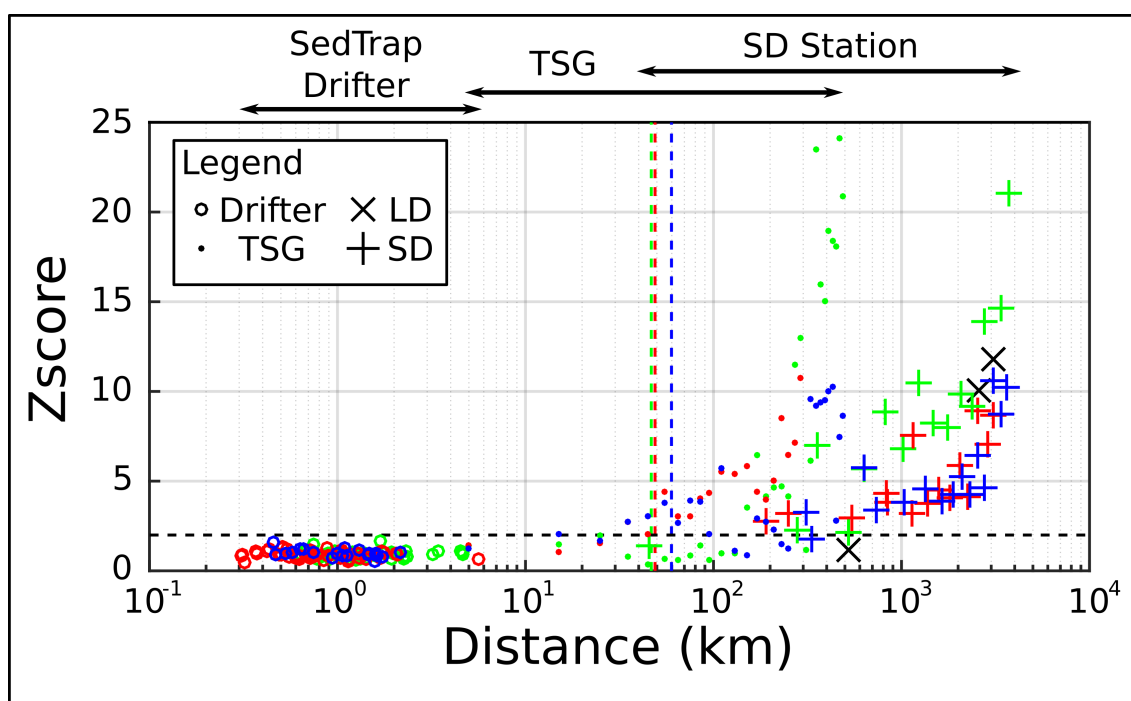
Figure 6. Z-score plotted over distance for the SedTrap Drifter (large empty circles), TSG (small filled circles), SD (+'s), and LD (x's) datasets. Colors for each symbol correspond with the same color code as previous figures, except the LD stations in black. Distance axis is log-transformed. Z = 2 reference line plotted with a black horizontal dashed line. Rossby radii R_D calculated from deep CTD casts plotted in vertical dashed lines following the color code.
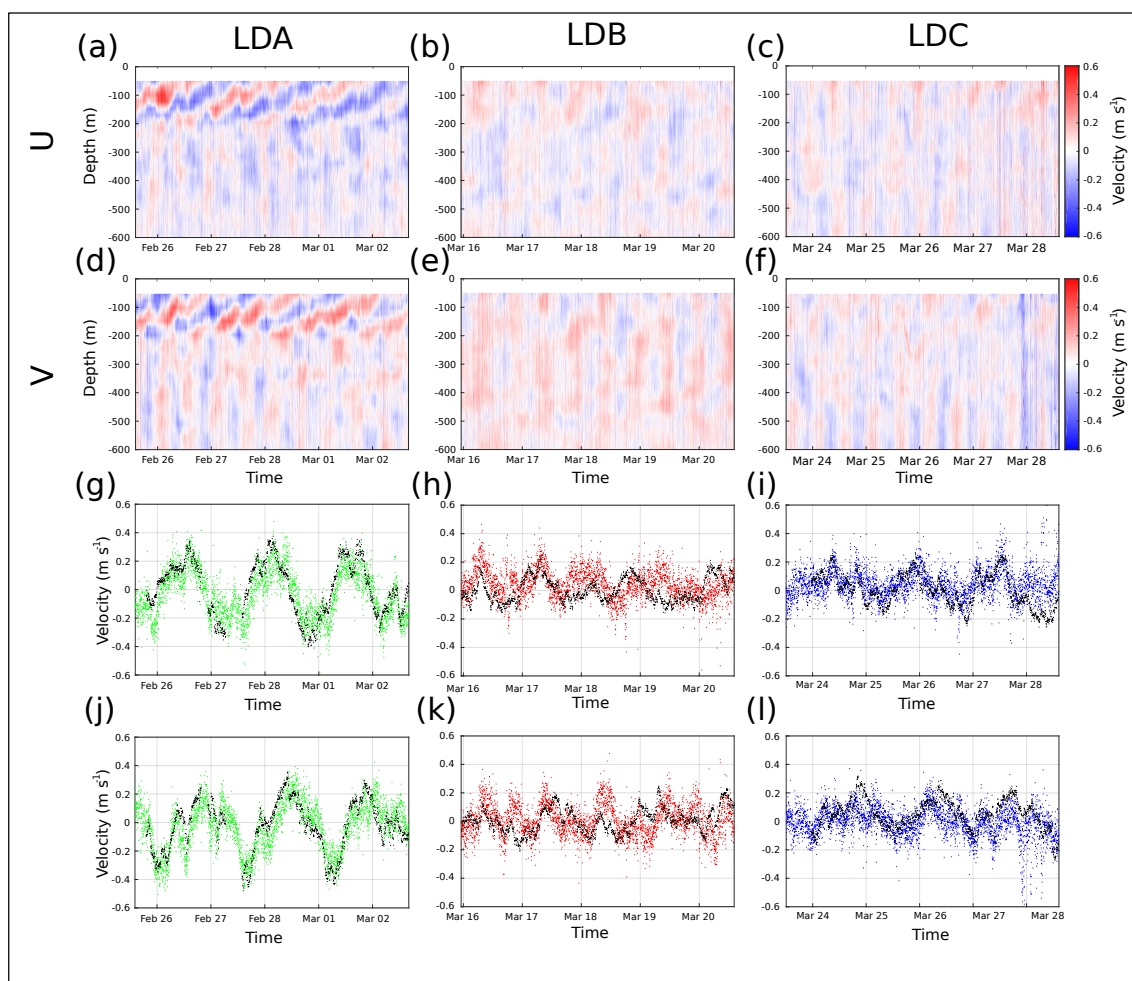
Figure 7. Timeseries of 38 kHz SADCP u and v components over depth for, respectively, (a,d) LDA, (b,e) LDB, and (c,f) LDC. Comparison of 55 m Aquadopp u and v components with 52 m SADCP u and v for, respectively, (g,j) LDA, (h,k) LDB, and (i,l) LDC. Aquadopp measurements shown in black, SADCP measurements follow the color code. Units are in ms^{-1}.
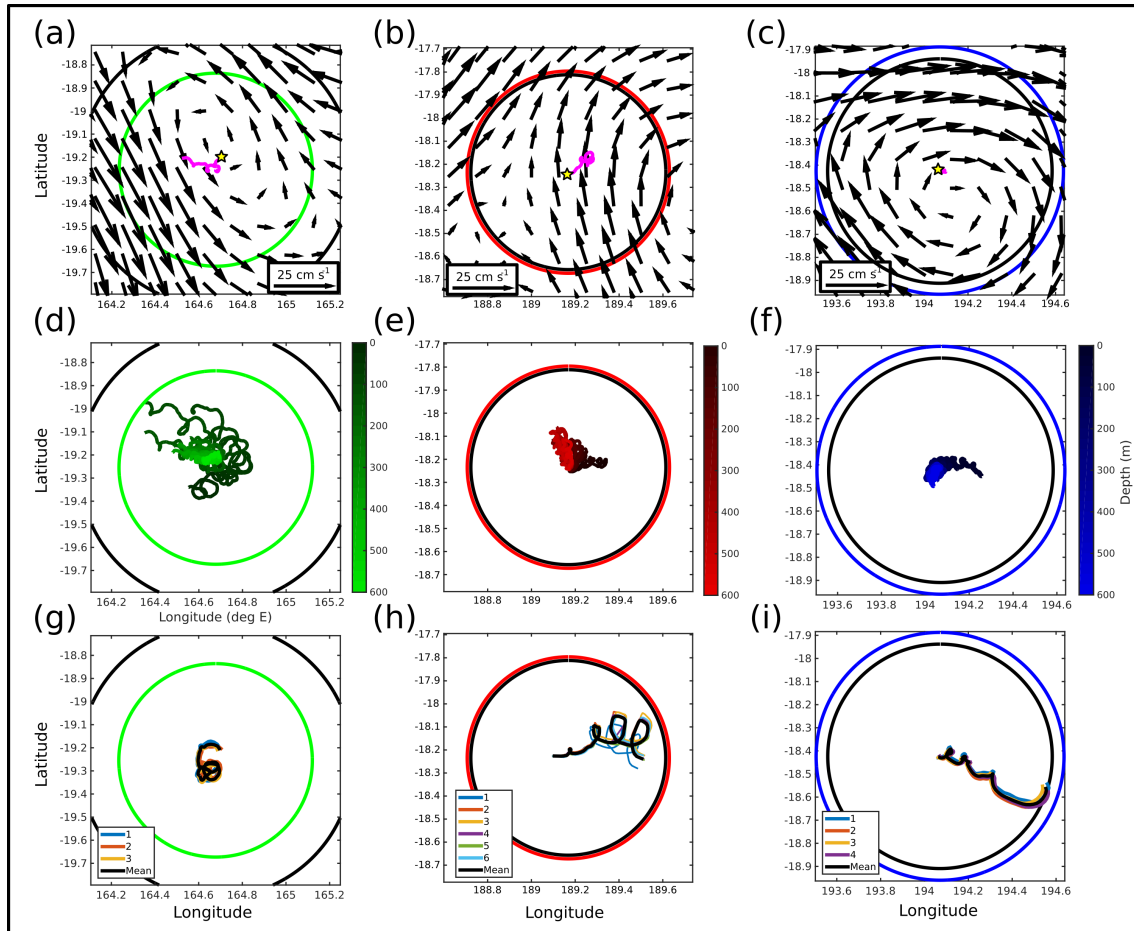
Figure 8. SedTrap Drifter Trajectories, SADCP 38kHz hodographs, and SVP trajectories for, respectively, (a,d,g) LDA, (b,e,h) LDB, and (c,f,i) LDC. Rossby radii R_D plotted following the color code, and spatial scale found through tracer analysis traced in black. SedTrap Drifter Trajectories plotted in magenta, with starting positions shown with a yellow star, superimposed on time-averaged altimetry-derived surface currents provided by CLS/CNES. Mean SVP drifter position plotted with black line. Domains plotted are the 120 km $\times$ 120 km squares shown in Fig. 2
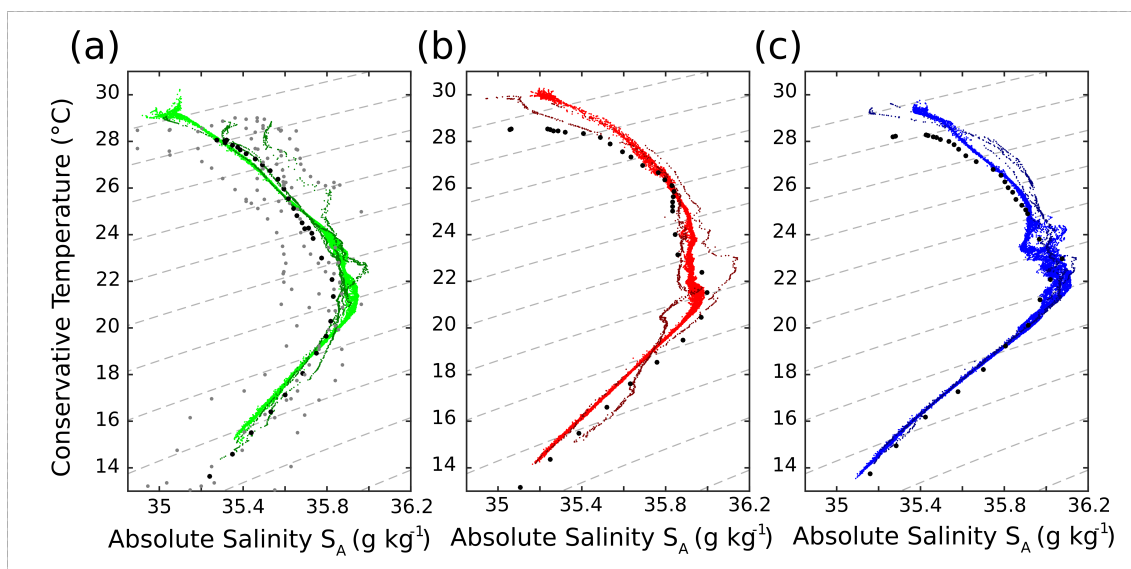
Figure 9. Supplementary Figure 1. T-S Diagrams of LD stations and neighboring SD stations, as in Fig. 4, with World Ocean Atlas 2013 v2 climatological profiles at the LD station positions superimposed. Climatological monthly means shown in black, with extrema of T-S shown in gray, as calculated by +/- 2 standard errors. Insufficient data at LDB and LDC preclude calculating these errors.
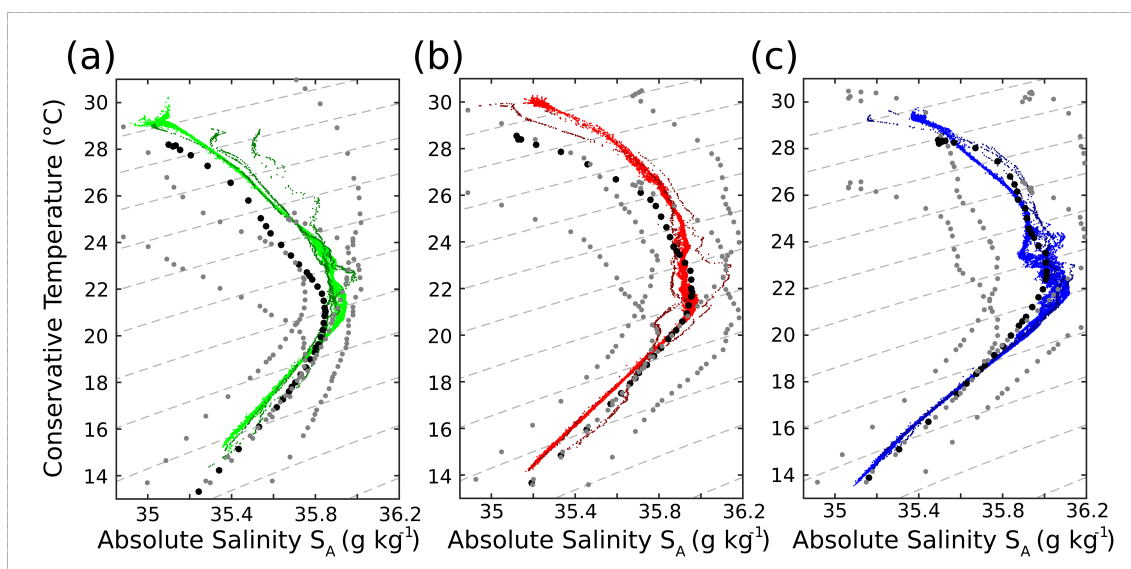
Figure 10. Supplementary Figure 2. T-S Diagrams like in Fig. S1, but with the CSIRO Atlast of Regional Seas (CARS) climatology.
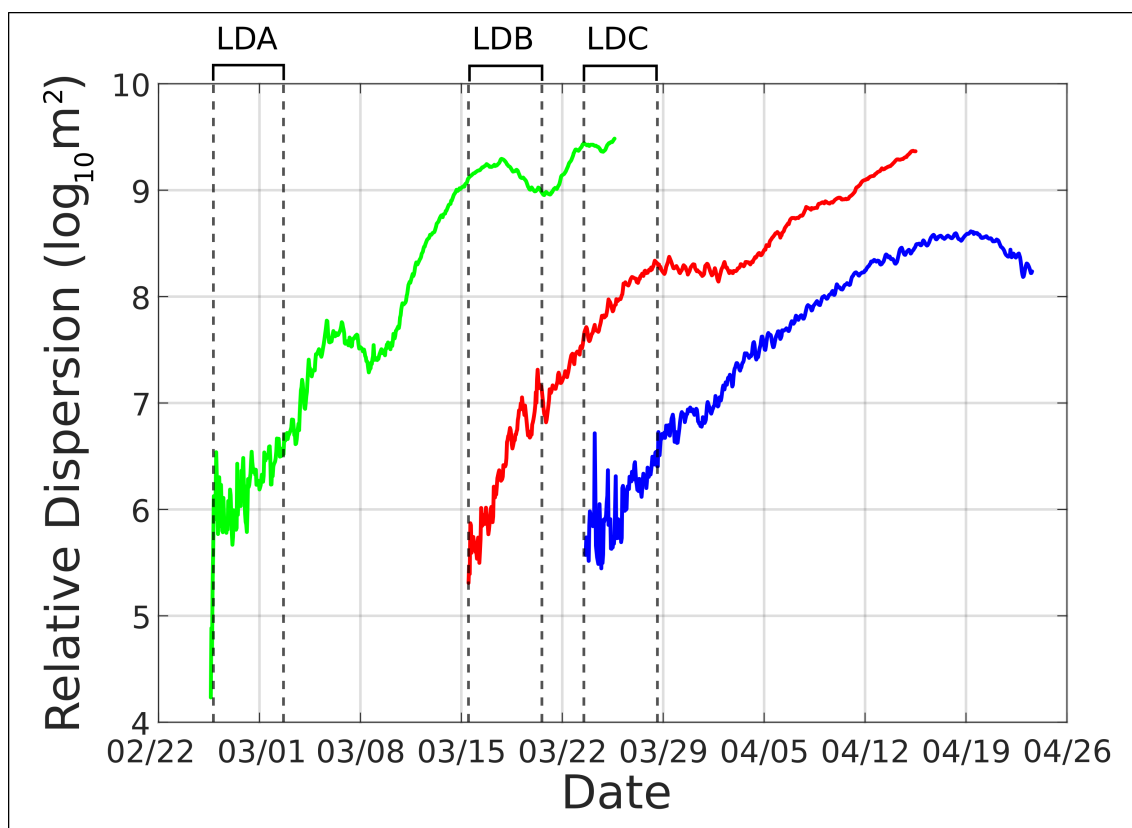
Figure 11. Supplementary Figure 3. Timeseries of SVP drifter relative dispersion during LD stations. Sampling start and end times demarcated by dashed gray lines. Data for the first month after SVP deployments are shown.