Evaluating existing ocean glider sampling strategies for submesoscale dynamics

Article

Published Version

Patmore, R. D. ORCID: https://orcid.org/0000-0002-5571-9229, Ferreira, D. ORCID: https://orcid.org/0000-0003-3243-9774, Marshall, D. P., du Plessis, M. D., Brearley, J. A. and Swaart, S. (2024) Evaluating existing ocean glider sampling strategies for submesoscale dynamics. Journal of Atmospheric and Oceanic Technology, 41 (7). pp. 647-663. ISSN 1520-0426 doi: 10.1175/JTECH-D-23-0055.1 Available at https://centaur.reading.ac.uk/116391/

To link to this article DOI: http://dx.doi.org/10.1175/JTECH-D-23-0055.1

Publisher: American Meteorological Society

Evaluating Existing Ocean Glider Sampling Strategies for Submesoscale Dynamics

RYAN D. PATMORE,[a,b] DAVID FERREIRA,[a] DAVID P. MARSHALL,[c] MARCEL D. DU PLESSIS,[d]
J. ALEXANDER BREARLEY,[e] AND SEBASTIAAN SWART[d]

[a] *Department of Meteorology, University of Reading, Reading, United Kingdom*
[b] *National Oceanography Centre, Liverpool, United Kingdom*
[c] *Department of Physics, University of Oxford, Oxford, United Kingdom*
[d] *Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden*
[e] *British Antarctic Survey, Cambridge, United Kingdom*

(Manuscript received 18 April 2023, in final form 25 April 2024, accepted 7 May 2024)

ABSTRACT: Mixing in the upper ocean is important for biological production and the transfer of heat and carbon between the atmosphere and deep ocean, properties commonly targeted by observational campaigns using ocean gliders. We assess the reliability of ocean gliders to obtain a robust statistical representation of submesoscale variability in the ocean mixed layer of the Weddell Sea. A 1/48° regional simulation of the Southern Ocean is sampled with virtual "bow-tie" glider deployments, which are then compared against the reference model output. Sampling biases of lateral buoyancy gradients associated with the arbitrary alignment between glider paths and fronts are formally quantified, and the magnitude of the biases is comparable to observational estimates, with a mean error of 52%. The sampling bias leaves errors in the retrieved distribution of buoyancy gradients largely insensitive to deployment length and the deployment of additional gliders. Notable sensitivity to these choices emerges when the biases are removed by sampling perpendicular to fronts at all times. Detecting seasonal change in the magnitude of buoyancy gradients is sensitive to the glider-orientation sampling bias but the change in variance is not. We evaluate the impact of reducing the number of dives and climbs in an observational campaign and find that small reductions in the number of dive–climb pairs have a limited effect on the results. Lastly, examining the sensitivity of the sampling bias to path orientation indicates that the bias is not dependent on the direction of travel in our deep ocean study site.

SIGNIFICANCE STATEMENT: Recent observational campaigns have focused on using autonomous vehicles to better understand processes responsible for mixing in the surface region of the ocean. There exists uncertainty around how effective these missions are at returning reliable and representative information. This study seeks to quantify the performance of existing strategies in observing mixing processes, and we confirm that strategies are biased to underestimate indicators of mixing. Furthermore, compensating for the bias by increasing the number of resources or changing the manner in which resources are used has limited reward. Our findings are important for decision-making during the planning phase of an observational campaign and display that further innovations are required to account for the sampling bias.

KEYWORDS: Oceanic mixed layer; In situ oceanic observations; Sampling

1. Introduction

Autonomous underwater vehicles, such as ocean gliders, have become a central component of many observational campaigns, having an advantage over more traditional observational techniques due to their endurance and relatively low cost. In recent years, there has been a particular interest in utilizing gliders to better understand submesoscale processes in the surface ocean mixed layer, which emerge at $\mathcal{O}(1)$ km and $\mathcal{O}(1)$ day, and their associated seasonality (Thompson et al. 2016; du Plessis et al. 2017; Viglione et al. 2018; du Plessis et al. 2019; Swart et al. 2020; Dove et al. 2021; Giddy et al. 2021). Submesoscale features are associated with a wide range of important dynamical processes, with a propensity to generate elevated vertical velocities (Taylor and Thompson 2023). Dynamical instabilities are a key source of vertical velocities at this scale (e.g., Fox-Kemper et al. 2008). Understanding the impact of submesoscale instabilities is important for estimating entrainment into the mixed layer and therefore the transfer of tracers, such as heat and carbon, between the atmosphere and the deep ocean (Boccaletti et al. 2007).

A prerequisite of submesoscale instabilities is lateral buoyancy gradients, which typically derive from frontogenesis associated with mesoscale stirring or spatial/temporal variations in surface buoyancy fluxes (McWilliams 2016). For example, meltwater production along the edge of retreating sea ice can lead to horizontal contrasts in buoyancy that promote instabilities (Lu et al. 2015; Horvat et al. 2016; Manucharyan and Thompson 2017; Biddle and Swart 2020). There are varying estimates regarding the accuracy of capturing lateral buoyancy gradients, and errors emerge that are attributed to the misalignment between glider paths and fronts (Thompson et al. 2016). These errors are often labeled a sampling bias

---

Denotes content that is immediately available upon publication as open access.

Corresponding author: Ryan D. Patmore, r.d.patmore@reading.ac.uk

DOI: 10.1175/JTECH-D-23-0055.1

and are an important consideration when interpreting results from glider observations. The sampling bias leads to a consistent underestimation of buoyancy gradients, which is currently estimated from observations and theoretical evaluation to range between 51% and 71% (Thompson et al. 2016; du Plessis et al. 2019; Swart et al. 2020).

A typical glider deployment in submesoscale field experiments consists of conducting repeat transects to capture the time evolution of scalar quantities at a particular location. These transects are able to return quasi-vertical profiles of the water column by periodically diving to a specified depth. Despite the numerous gliders deployed in this manner, choices in experimental design have received limited evaluation. The merits of many existing practices are yet to be fully explored, particularly with respect to the sampling bias.

One consideration when designing an observational program is the level of resource required. There are choices to be made on the length of deployment, and it is also becoming increasingly possible to deploy multiple gliders simultaneously. Aside from the sampling bias, uncertainty in observations is associated with the temporal and spatial coverages of data retrieved. Collecting more data will tend to reduce uncertainty in terms of both the spatial distribution and seasonal change for variables such as buoyancy gradients, but there exists no quantification of this uncertainty reduction. Furthermore, previous assessments of glider sampling strategies investigate the glider speed, dive angle, and sampling frequency (e.g., Rudnick and Cole 2011; Steinberg and Eriksen 2020). These evaluations focus on decisions related to individual dives. The merits of particular dive and transect patterns remain unknown.

To enable effective evidence-based decision-making, there is a need to both quantify the performance of existing strategies and compare these strategies with alternative options. The purpose of this study is to evaluate the performance of existing glider sampling strategies for targeting physical parameters such as buoyancy gradients and explore the sensitivity of the results to commonplace choices of experimental design. We investigate the ability of glider sampling to capture buoyancy gradients and their seasonality, while assessing the impact of altering the configuration of glider transects. Our focus is on the upper ocean because that is where gliders are typically deployed, but we acknowledge that other settings have also been identified as important regions for submesoscale fields (e.g., the bottom boundary layer; Wenegrat et al. 2018). The investigation follows that of Steinberg and Eriksen (2020) by emulating gliders within model simulations, sampling the model like a glider, and comparing the resulting data against the full model dataset.

In restricting the analysis to buoyancy gradients and other metrics of density, this study omits the assessment of other quantities relevant for observing features of submesoscale dynamics, such as horizontal velocity. Together with stratification and lateral buoyancy gradients, velocities are needed to categorize the potential for gravitational, symmetric, and inertial instabilities by mapping Ertel's potential vorticity (Thomas et al. 2013), and evaluation of this kind can be done using gliders (Thompson et al. 2016; Meunier et al. 2018; du Plessis et al. 2019). Glider estimates of absolute velocities are not obtained directly and are usually derived from other proxy measurements, which are outlined in section 5b. Results in this present study are applicable to a subset of these methods that rely on horizontal buoyancy gradients. For example, the thermal wind relation is often used to infer the vertical shear, and where this is the case, conclusion drawn from our analysis of buoyancy gradients will directly apply to estimates of velocities.

Unlike in the real ocean, the full model dataset provides an exact knowledge of the expected result against which to compare the glider sampling. One caveat of not sampling the real ocean is that uncertainty exists around transferability to real ocean conditions. For example, glider observations are known to alias high-frequency variability down to turbulent scales (Rudnick and Cole 2011; Steinberg and Eriksen 2020). The model used in this study does not capture aliasing associated with tides or turbulent dynamics because the resolution is limited to the submesoscale and tidal forcing is not implemented. Despite the model being inherently different from the observed ocean, we rationalize that it adequately represents properties of the real ocean that are of primary importance to meso/submesoscale dynamics (e.g., frontal features) to the extent that conclusions drawn from this study can be expected to apply to the observed ocean. This assertion is supported by previous studies that demonstrate the emergence of submesoscale filamentation in models, provided an adequate choice of resolution (Rosso et al. 2014; Brannigan et al. 2015; Brannigan 2016; Brannigan et al. 2017; Bachman et al. 2017). The model resolution employed in this study is approximately 1 km in the horizontal, which we consider mesoscale resolving and submesoscale permitting.

This paper is set out as follows. Section 2 outlines the numerical configurations, the process for sampling the model, and statistical methods used in this paper. Section 3 describes the broad features of the model output. Section 4 presents the results of our analysis. Section 5 discusses the remaining questions related to glider deployment, adding context to the results section. Finally, section 6 provides a concluding discussion and recommendations.

2. Methods

a. Buoyancy

The analysis in this study makes frequent use of buoyancy gradients, ∇b. Throughout, this term refers specifically to lateral buoyancy gradients, where buoyancy is defined as

$$b = g(1 - \rho_\theta/\rho_0), \tag{1}$$

and $g = 9.81$ m s^{-2} is the gravitational acceleration, ρ_θ is the potential density, and $\rho_0 = 1026$ kg m^{-3} is the reference density.

b. The model

Our investigation focuses on sampling output from a model simulation using glider trajectories, comparing the sampled data against the model results. The model simulations were carried out as a part of the Southern Ocean Carbon and Heat

Impact on Climate (SO-CHIC) project (Sallée et al. 2023) to evaluate seasonal-scale changes in buoyancy and use the Nucleus for European Modelling of the Ocean (NEMO) framework, version 4.0.4. Simulations cover an 8° × 8° extent in the horizontal, centered on 60°S and 0° (see Fig. 1). The modeled region is in the northeastern part of the Weddell Sea and lies south of the Antarctic Circumpolar Current.

The model is a regional configuration that runs as a one-way nest of the National Oceanography Centre Marine Systems Modelling Group ORCA12 global simulation (Marzocchi et al. 2015) hosted on Joint Analysis System Meeting Infrastructure Needs (JASMIN), with lateral boundary conditions extracted from 5-day means of the global ORCA12 fields. Lateral boundary forcing is used for temperature, salinity, horizontal velocities, and sea surface height. The model includes a representation of sea ice using SI^3 (Vancoppenolle et al. 2023). Sea ice is restored at the lateral boundaries using sea ice thickness, sea ice concentration, snow thickness, and sea ice temperature fields. Surface boundary forcing is identical to the global simulation, and the model is forced at the surface using DRAKKAR 5.2 reanalysis fields (Brodeau et al. 2010). This forcing dataset is a development on the ERA-Interim reanalysis that includes surface variables required by the NEMO bulk formulae. Surface forcing uses the bulk formula with the NCAR algorithm (Large and Yeager 2004), and forcing occurs on two time scales. Wind, temperature, humidity, and mean sea level pressure are applied on 3-h time scales. Short- and longwave radiation, precipitation, and snowfall are applied on 24-h time scales. Tidal forcing is omitted.

The grid uses the standard ORCA12 mesh in the vertical with 75 levels of increasing thickness with depth. The horizontal resolution is 1/48°, which is approximately 1 km at this latitude. Finite computational resources necessitate a compromise between resolution and domain size. The resolution choice balances the requirement to permit submesoscale dynamics, while maintaining a domain size that can accommodate a representative distribution of eddies and fronts. The vertical coordinates are z^*, with partial cell representation. The z^* coordinates distribute variations in sea surface height across all levels of the vertical grid, and the time-varying cell boundaries are referenced to an initial fixed grid (Adcroft and Campin 2004). The partial cell method is a mechanism for creating smoother topographic transitions with z levels, where the ocean is permitted to occupy a fraction of a cell (Adcroft et al. 1997).

The simulation is initially run for 1 year beginning in January 2012, spinning up from the 1/12° initial conditions. The model is then reinitialized, running between 9 December 2012 and 31 March 2013, with diagnostics output as 3-hourly snapshots. Data storage constraints led to a trade-off between the frequency of outputs and the length of the model time series. Outputting the fields for several months at this frequency allows for the investigation of seasonal transitions, while capturing the evolution of fronts. The restart period with 3-hourly snapshots is chosen to coincide with sea ice retreat, so that we can examine the seasonal changes in buoyancy gradients that emerge at this time.

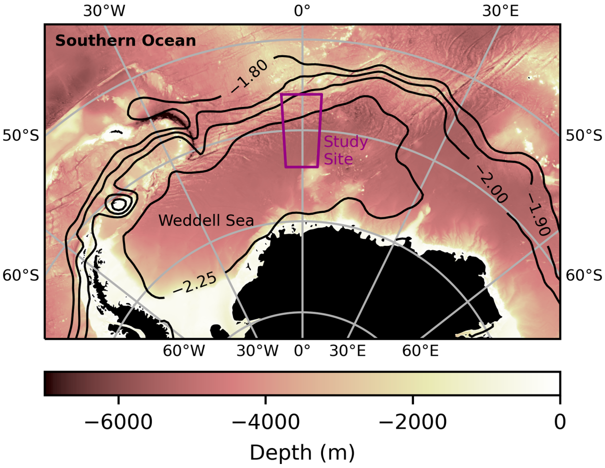

FIG. 1. Southern Ocean bathymetry with the study site highlighted in purple. Black contours show the mean dynamic ocean topography averaged between 2011 and 2020 from Armitage et al. (2018), with labels representing the height (m) of each contour. A portion of the ACC is indicated by levels −1.8, −1.9, and −2.0 m, whereas the Weddell Gyre is represented by the −2.25-m contour.

c. Sampling the model

The model is sampled using the exact sampling positions of the Robotic Observations and Modeling of the Marginal Ice Zone (ROAM-MIZ) observational campaign that took place during the 2018/19 austral summer (see Swart et al. 2020; Giddy et al. 2021). The glider conducted a regular series of climb–dive pairs (Fig. 2b) and traveled along a "bow-tie" pattern in the horizontal (Fig. 2a). The bow-tie has a long history in deep-water settings (>1000 m) relative to the emergence of glider technology, with early adoption in 2005 (Nicholson et al. 2008). In addition to the bow-tie pattern, the field campaign undertook two large transects orientated in the meridional direction to capture large-scale features. These large meridional transects are not a common feature of field campaigns that adopt a bow-tie sampling, and we therefore do not incorporate them into our analysis.

The modeled data are first mapped onto the glider positions in terms of space and time. The process extracts the latitude, longitude, depth, and time from each individual sample of the glider data and uses linear interpolation to generate one realization of a glider deployment. The process is repeated 100 times, each time changing the starting position of the glider in a stochastic manner. The virtual deployments are not limited to the area of open ocean. They can travel through, and be initiated in, regions of sea ice. The size and pattern of the bow-tie remain the same for each glider deployment. This dataset of 100 glider deployments is what allows us to build the statistical picture of sampling strategies.

Each glider deployment within the model is processed using the same methods as Giddy et al. (2021) for consistency. The samples are first linearly interpolated onto a uniform grid with a 1-m spacing in the vertical and 1 km in the horizontal. Following interpolation, the path is split into transects according to the bow-tie sections (see Fig. 2b). Our primary interest is in assessing

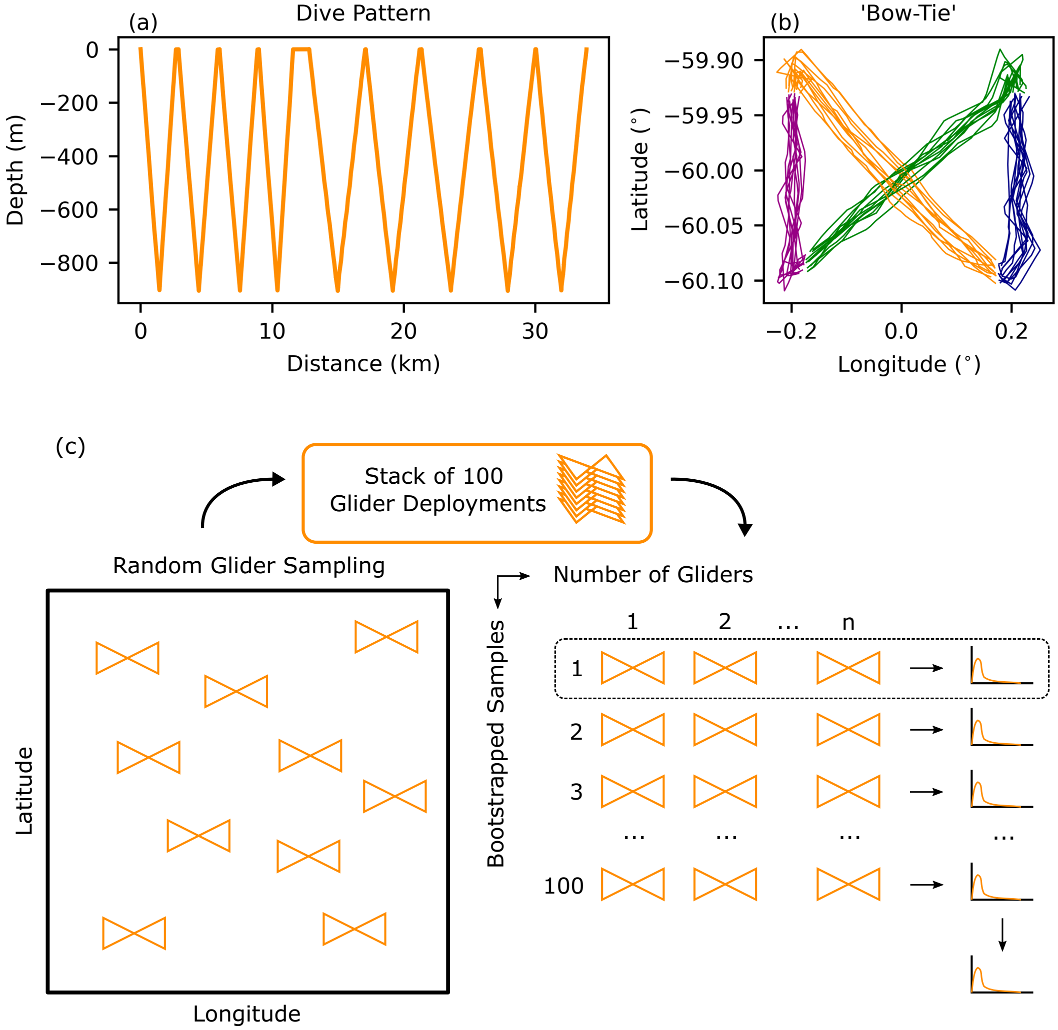

FIG. 2. Glider sampling pattern displaying the exact sampling locations used during the ROAM-MIZ observational campaign (Swart et al. 2020; Giddy et al. 2021). (a) A vertical section of the glider path showing the dive–climb pattern. This is a subset of the full path for illustration, and the glider covered a total lateral distance of 2074 km. (b) The glider path in the horizontal taking the form of a bow-tie sampling strategy. Colors in (b) represent the transects that the glider data are split into during postprocessing. (c) Schematic of the bootstrapping calculation. The model output is repeatedly sampled 100 times with a bow-tie glider sampling pattern, using randomized starting locations but initiated at the same time for each "deployment." A bootstrap sample is then formed by randomly selecting n glider deployments—with replacement—from the stack of 100 deployments, where n represents the number of gliders simultaneously sampling. The process is repeated 100 times, and a statistic is calculated for each bootstrap sample, for example, a PDF. A bootstrap mean and bootstrap uncertainty are then calculated by finding the mean, lower decile, and upper decile of the statistics calculated for each bootstrap sample.

the ability of gliders to capture mixed layer dynamics. The simulated spatial mean mixed layer depth is initially 49 m on 9 December 2012, before steadily rising over time to 93 m on 31 March 2013. Full-depth glider profiles are sampled from the model, but our analysis is restricted to 10-m depth, ensuring that samples remain within the mixed layer—which has a depth greater than 10 m for more than 99% of the model data.

d. *Statistics*

One focus of this paper is to understand the sensitivity of glider sampling accuracy to the number of gliders deployed and length of each deployment. This has parallels with the field of large ensemble climate modeling, where it is important to estimate sufficient ensemble sizes required to identify different climate responses. Methods in this analysis are derived from techniques used in the field of large ensembles (e.g., Milinski et al. 2020).

To quantify the impact of varying the number of gliders deployed at any one time, we use a form of bootstrapping. Figure 2c provides a diagram of this process. Through initializing sampling of the model in randomized positions, we have a pool of 100 glider deployments. From this pool, n gliders are selected at random and the data are combined into what we term as a *bootstrapped sample*. We generate 100 bootstrapped samples, where the selection of glider deployments is done with replacement, which means each bootstrapped sample draws from the full pool of 100 glider deployments. From each bootstrapped sample, a chosen statistic is calculated. The (un)certainty of that statistic is then evaluated by calculating the upper and lower deciles across all bootstrapped samples, which is termed the *bootstrap uncertainty*. The choice of n represents the number of gliders simultaneously sampling at any one time, and we vary this parameter in our analysis.

In addition to quantifying the impact of the number of gliders deployed, we also investigate the effect of varying the length of deployment at the same time. This is done by splitting the time series of each bootstrapped sample into chunks of a chosen time period (i.e., weekly chunks), calculating the bootstrap uncertainty for all time periods, and then taking the mean across the resulting set of bootstrap uncertainties.

3. Hydrography

Our analysis diagnoses the time evolution and distribution of the model hydrography by glider sampling. We therefore provide an overview of the features that characterize the output from the simulations.

Figure 3a shows a snapshot in late December of potential temperature and the Rossby number, $\mathrm{Ro} = \zeta/f$, where ζ is the relative vorticity and f is the Coriolis parameter. The vertical structure shows a thin layer of cold (and fresh) water near the surface, a warm (and salty) intermediate layer of Circumpolar Deep Water below, and cooler water at depth. There is a prevalence of mesoscale eddies, and some submesoscale features are also observed in the southern half of the domain.

Mixed layer instabilities associated with submesoscale features occur where Ro is $\mathcal{O}(1)$ and at spatial scales near the internal Rossby radius of deformation (Boccaletti et al. 2007):

$$L_d = \frac{NH}{f}, \tag{2}$$

where H is the mixed layer depth and $N^2 = -db/dz$ is the Brunt–Väisälä frequency. Observations from the modeled region estimate $L_d = 2 \pm 0.5$ km (Swart et al. 2020). Since the model resolution is ≈1 km and the magnitude of the Rossby number is mostly below $\mathcal{O}(1)$ (Fig. 3a), we suggest that submesoscale fronts/eddies are permitted rather than resolved. This means that submesoscale features are simulated, but their scale is close to the grid size and their evolution is dampened or augmented by parameterized momentum/scalar diffusion.

In Fig. 3b, the distribution in the magnitudes of the buoyancy gradients is similar between the zonal and meridional components, suggesting that there is no significant directional bias. The domain is far enough south of the Antarctic Circumpolar Current (ACC) for frontal systems to be relatively isotropic. As a result, we expect the qualitative results from this study to apply generally and the quantitative results to be transferable to regions away from strong jets, which have been the focus of several observational campaigns targeting submesoscale flows over the last decade (e.g., Thompson et al. 2016; Buckingham et al. 2016; du Plessis et al. 2017, 2019; Giddy et al. 2021). The analysis in this paper uses the Euclidean norm of the buoyancy gradients to capture the cross-front gradients:

$$|\nabla b| = \left[\left(\frac{db}{dx}\right)^2 + \left(\frac{db}{dy}\right)^2\right]^{1/2}, \tag{3}$$

which has a different distribution to the zonal and meridional components (Fig. 3b).

The snapshot in Fig. 3a is taken during an early part of the simulation that contains sea ice in the southern half of the domain and the submesoscale features are coincident with this region of sea ice cover (not shown). The simulation covers a period of seasonal change that exhibits a full retreat of sea ice in the south. Simulating the seasonal transition enables the evaluation of glider sampling in capturing seasonal change, which will be discussed in section 4c.

Figure 3d shows that sea ice presence is restricted to the south of the domain. During the first 4 weeks of output, there is a progressive reduction in this sea ice cover (Fig. 3d) and the associated melting leads to an elevated net downward freshwater flux at the ocean surface (Fig. 3e). Freshwater fluxes are smaller in magnitude in the north of the domain and vary on shorter time scales. This variability is dominated by synoptic-scale atmospheric events. Once sea ice is no longer present, the freshwater flux varies in the south on shorter time scales and in phase with the rest of the domain.

As common at high latitudes, the buoyancy gradients are dominated by variations in salinity in this region rather than temperature, in part, because of the low thermal expansion coefficient at low temperatures (Roquet et al. 2022). As a result, the observed sea ice loss and associated surface freshwater fluxes influence the mixed layer buoyancy gradients, with the northern and southern halves of the domain showing a heterogeneous evolution of buoyancy gradients over time (Fig. 3c). The northern half of the domain has a relatively uniform magnitude in buoyancy gradients over time, whereas in the southern half, the buoyancy gradients are larger and increase further as the sea ice melts. Following the complete loss of sea ice, the regional differences reduce and the magnitude of buoyancy gradients converges. However, this convergence occurs with a lag and the elevated buoyancy gradients that stem from the sea ice loss persist for approximately 3 weeks after the sea ice has gone.

4. Results

a. *Identifying the buoyancy gradient sampling bias*

For gliders to capture the full magnitude of buoyancy gradients, they must travel orthogonal to the gradients at all times. Since buoyancy gradients are sampled in the along-track direction, the measured gradient aligns at an angle to

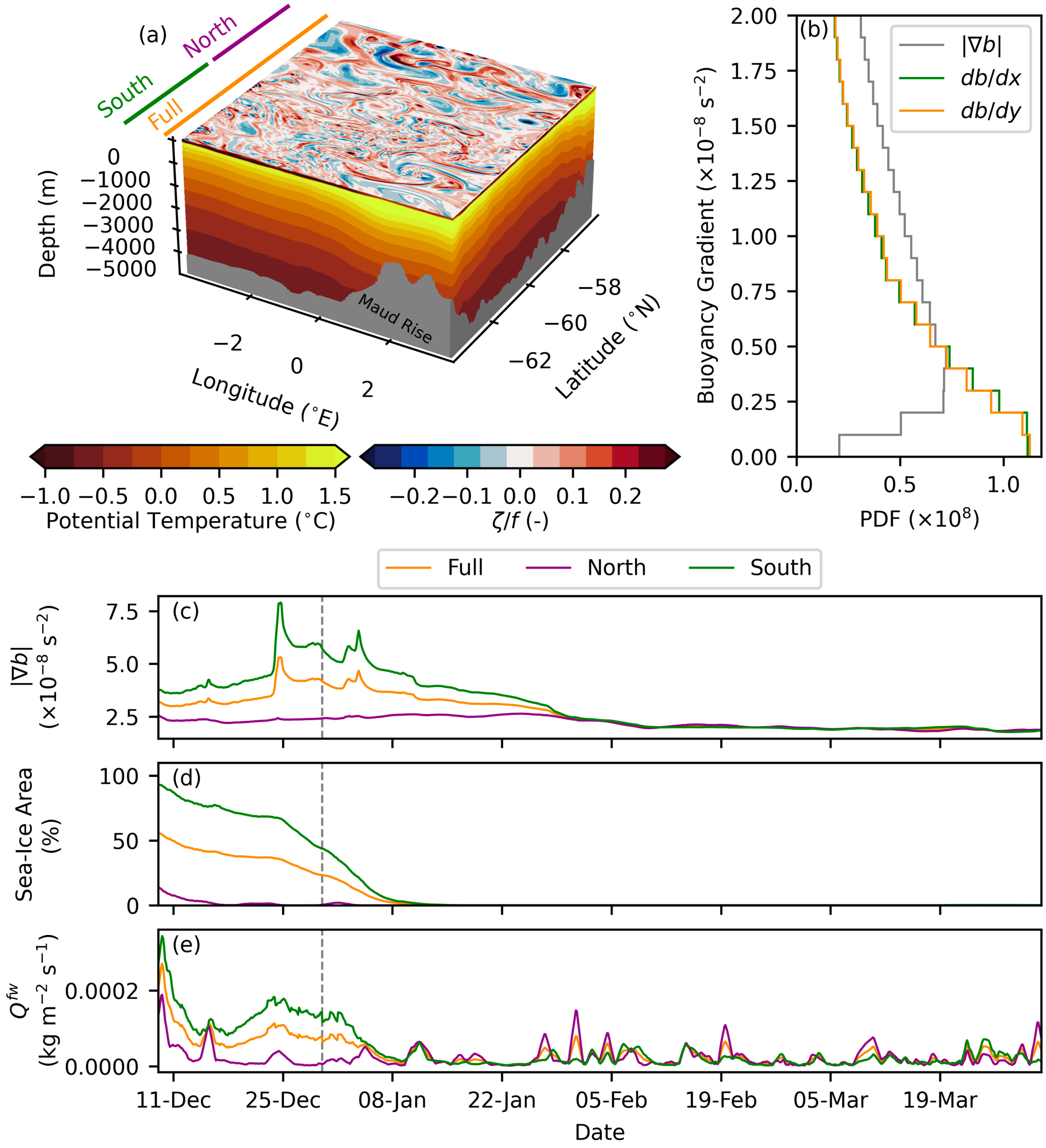

FIG. 3. Hydrography of the 1/48° NEMO simulation. (a) Snapshot of the model on 30 Dec 2012 showing the Rossby number ζ/f at the surface and potential temperature on the vertical faces. Colored lines demarcate the regions presented in (c)–(e). Gray shading shows the bottom topography. (b) PDF of the zonal, meridional, and Euclidean norm buoyancy gradients at a 10-m depth. The lower panels show the time series of (c) Euclidean norm buoyancy gradients at a 10-m depth, (d) sea ice area, and (e) net downward flux of freshwater at the sea surface. Lines in (c)–(e) are horizontally averaged over the full (orange), the northern (purple), and the southern (green) portions of the domain. The dashed vertical lines in (c)–(e) mark the snapshot for (a).

the tangent vector of the gradient. The misalignment of the across-front and along-track directions leads to a systematic observational bias in the magnitude of buoyancy gradients sampled by gliders, resulting in an underestimation of the true value. Here, we identify the sampling bias, comparing the modeled buoyancy gradients with data derived from the virtual glider deployments.

Previous assessments propose that the sampling bias equates to a 51%–71% error (see Table 1). These evaluations for the observational bias stem from either a simple analytical approach (Thompson et al. 2016) or calculating the error from observations by estimating the glider path and front directions (du Plessis et al. 2019; Swart et al. 2020). In the latter approach, uncertainties in evaluating both glider path and front directions make the accuracy of these estimates unclear. Sampling methods provided here make it possible to calculate the along-track and across-front buoyancy gradients and recover explicit estimates of the observational bias. We find the difference between the along-track and across-front buoyancy gradients for each sample within a glider deployment and

TABLE 1. Estimation of the observational bias in the observed magnitude of buoyancy gradients due to the along-track glider path not aligning perpendicular to ocean fronts.

Mean error (%)	Standard deviation	Source	Method
52	35%	This study	Explicit calculations from the model
51		Swart et al. (2020)	Estimates from observations
64		du Plessis et al. (2019)	Estimates from observations
71		Thompson et al. (2016)	Theoretical evaluation

then evaluate the mean and standard deviation across all samples. We find that the results remain the same whether this is done for a single deployment or if we bootstrap across multiple deployments. The across-front buoyancy gradients are calculated using (3), evaluating the Euclidean norm of the modeled buoyancy gradients. Gradients are calculated at the grid scale (≈1 km), avoiding issues of smoothing that would occur over larger distances.

Our estimates show that gliders underestimate the magnitude of buoyancy gradients by 52% on average (Table 1), which compares well with the previous estimate of 51% by Swart et al. (2020) and is slightly smaller than the 64% error from du Plessis et al. (2019). The standard deviation in our model derived estimate is 35%, showing the error can vary by a large amount during a single deployment. We speculate that this is a reflection of the spatial and temporal variations in front orientation associated with eddying features.

We now investigate this sampling bias in the context of deployment choices. We test the sensitivity of the bias to two aspects of the glider sampling: 1) the length of time a glider is deployed and 2) the number of gliders sampled at any one time. Deployment lengths of 1 week, 2 weeks, 3 weeks, and 3.5 months are investigated. We choose to present 1–3 weeks because the results showed the greatest sensitivity over this range. The 3.5 months is then the longest possible deployment, given the time span of the high-frequency model outputs. In order to incorporate the full time series, the time series is split by the deployment length, statistics are calculated for each chunk of time, and an average is applied across the chunks.

Figures 4a–d show the distribution of buoyancy gradient magnitudes as represented by the virtual glider deployments, sampling in the along-track direction, compared against the distribution for the model as a whole. The results show that both increasing the number of gliders deployed and the length of time sampled act to reduce uncertainty around the sampling, but the along-track glider sampling does not converge onto the distribution of buoyancy gradients provided by the model. The prevalence of large buoyancy gradients is underestimated, while it is overestimated for small buoyancy gradients (Figs. 4a–d). The convergence toward a distribution means, to some extent, the sampling bias may limit the benefits from using additional resources.

Figures 4e–h explicitly isolate the impact of sampling at an angle to the buoyancy gradients by calculating the distribution of buoyancy gradients that would arise from gliders sampling perpendicular to buoyancy gradients (across-front). In comparison to along-track sampling, there is a marked change in the alignment between the glider sampling and the model data and reductions in uncertainty result in convergence of the glider estimates toward the modeled probability density function (PDF). This provides indication that a sizeable proportion of errors in the distribution of buoyancy gradients shown in Figs. 4a–d stem from the sampling bias, which is a result of not sampling across gradients. The remaining bias arises in accordance with the quantity of samples retrieved, and this will be explored further in the subsequent subsection.

b. Quantifying the buoyancy gradient sampling bias

We have identified that our results reproduce the sampling bias of buoyancy gradients reported by existing studies. While this sampling bias remains evident regardless of the quantity of samples taken, the quantitative impact that an increase in deployment length and the number of gliders deployed have on recovering the model mean remains unclear, particularly for the along-track sampling. We now seek to quantify the outcome of these choices in deployment. Figures 5a–h show the distribution of the root-mean-square error (RMSE) of the buoyancy gradient for varying number of gliders and deployment lengths. The RMSE is calculated for each bootstrapped sample against the model mean presented in Fig. 4. The results presented are then the mean across all bootstrap-sampled RMSEs.

The largest errors in the along-track sampling occur at the largest and smallest buoyancy gradients (Figs. 5a–d). The RMSE is particularly large at gradients smaller than 4×10^{-9} s^{-2} where errors exceed 500% (not shown) These errors are associated with the differences in the distributions between the glider samples and model mean presented in Figs. 4a–d, which arise from the consistent underestimation of gradients with the along-track sampling. For 1-, 2-, and 3-week deployments, increasing the number of gliders reduces the RMSE for all buoyancy gradients. As the deployment length increases, the impact of increasing the number of gliders reduces and there is virtually no sensitivity to the number of gliders when sampling for 3.5 months. The length of deployment also has no impact when 20 gliders are used. In summary, this indicates that if a particular threshold in deployment length is met, no further gains in estimating the distribution of buoyancy gradients will arise from increasing the number of gliders and the equivalent holds for sensitivity to deployment length.

When sampling across-front, the largest errors are also found at the largest and smallest buoyancy gradients, but the magnitude of errors at small buoyancy gradients is much reduced (Figs. 5e–h). Furthermore, unlike the along-track sampling, we find that errors are sensitive to glider number and

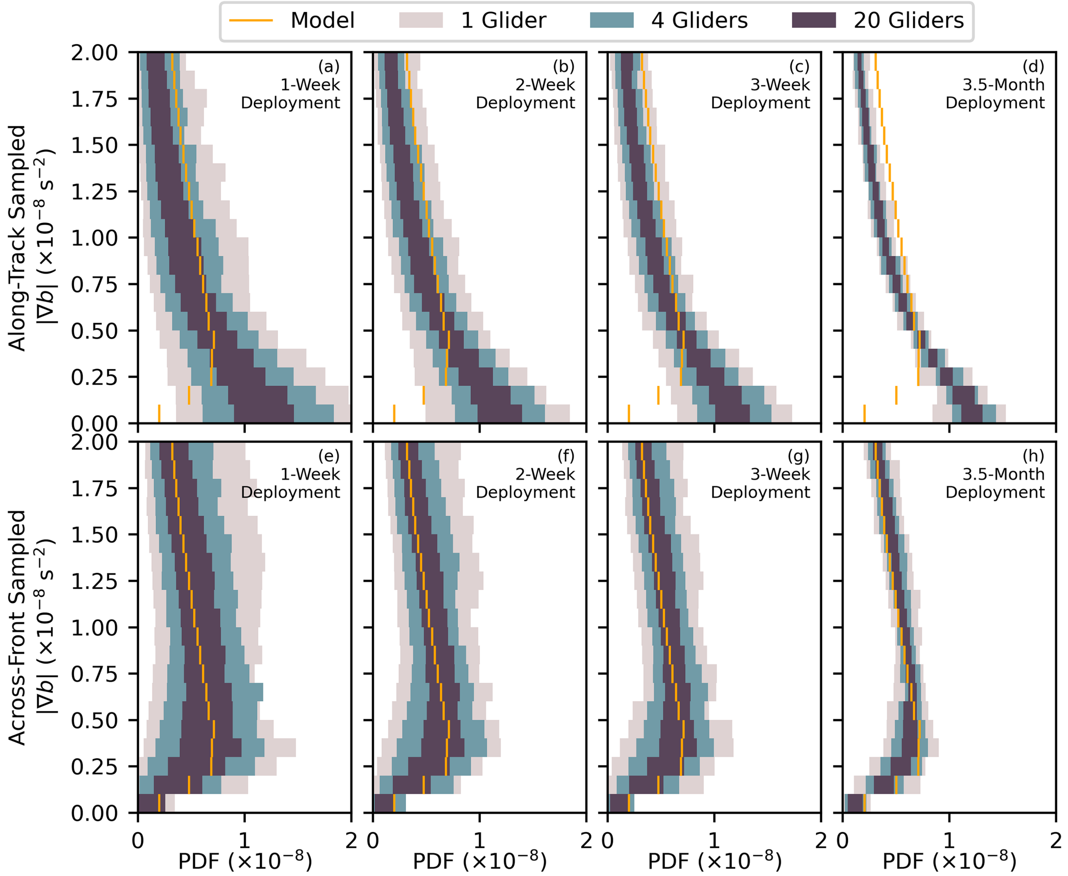

FIG. 4. PDF of buoyancy gradients at a 10-m depth from the full model (orange bars) compared against the bootstrapped sampling of the model using 1 (gray), 4 (teal), and 20 (dark purple) simultaneous glider deployments. Shading represents the interdecile range of the bootstrapped samples. Buoyancy gradients for the glider samples are calculated in (top) the along-track direction and (bottom) the cross-front direction by taking the vector norm of the buoyancy gradient at matching positions. Model curves in orange are identical between rows. Columns display the effect varying the glider deployment between (a),(e) 1 week, (b),(f) 2 weeks, (c),(g) 3 weeks, and (d),(h) 3.5 months.

deployment length across all choices presented, with errors reducing for all increases in both.

As a synopsis, in Figs. 5i and 5j, we present RMSE taken as a mean across all buoyancy gradients for all glider numbers and deployment lengths investigated. For the along-track sampling, there is reduced sensitivity to errors with increased deployment length or number of gliders, with convergence on an error of 60% (Fig. 5i). Transitioning from 1 to 5 glider deployments reduces the error when sampling for 1, 2, or 3 weeks, but when sampling for the full 3.5 months, there are no gains from increasing the number of gliders. For the along-track sampling, the largest errors are confined to buoyancy gradients that are smaller than 4×10^{-9} s^{-2}. Repeating the analysis with these buoyancy gradient bins removed for the 3.5-month deployment shows a translation of the error to a value below 40% for all glider numbers. Figure 5j shows that the RMSE error is significantly reduced for the across-front sampling in comparison to that for the along-track, reaching below 20% for some options. Although reductions occur for all deployment choices, the RMSE is sensitive to all increases in the glider number and deployment length, with reductions in RMSE most pronounced when transitioning between 1–5 gliders and 1–2-week deployments.

c. *Seasonal change*

Since the data we present exhibit a strong signature of seasonality in the upper ocean (see Fig. 3), we can investigate how glider sampling captures seasonal change. We present results for the southern half of the domain so as to not conflate the differences in seasonality between the north and the south of our model described in section 3. Figure 6a shows the change in buoyancy gradients over time for the complete model data compared to that measured by differing number of simultaneous glider deployments. For the full model domain, the magnitudes and spread in buoyancy gradients peak at the end of December and reduce over the months thereafter, coincident with seasonal sea ice melt. The glider data evolve in a similar manner over time, and the glider sampling is able to capture the change in buoyancy gradients over time. The buoyancy gradients for the glider data are in the along-track direction for consistency with observations. As a result, gliders systematically underestimate the magnitude in

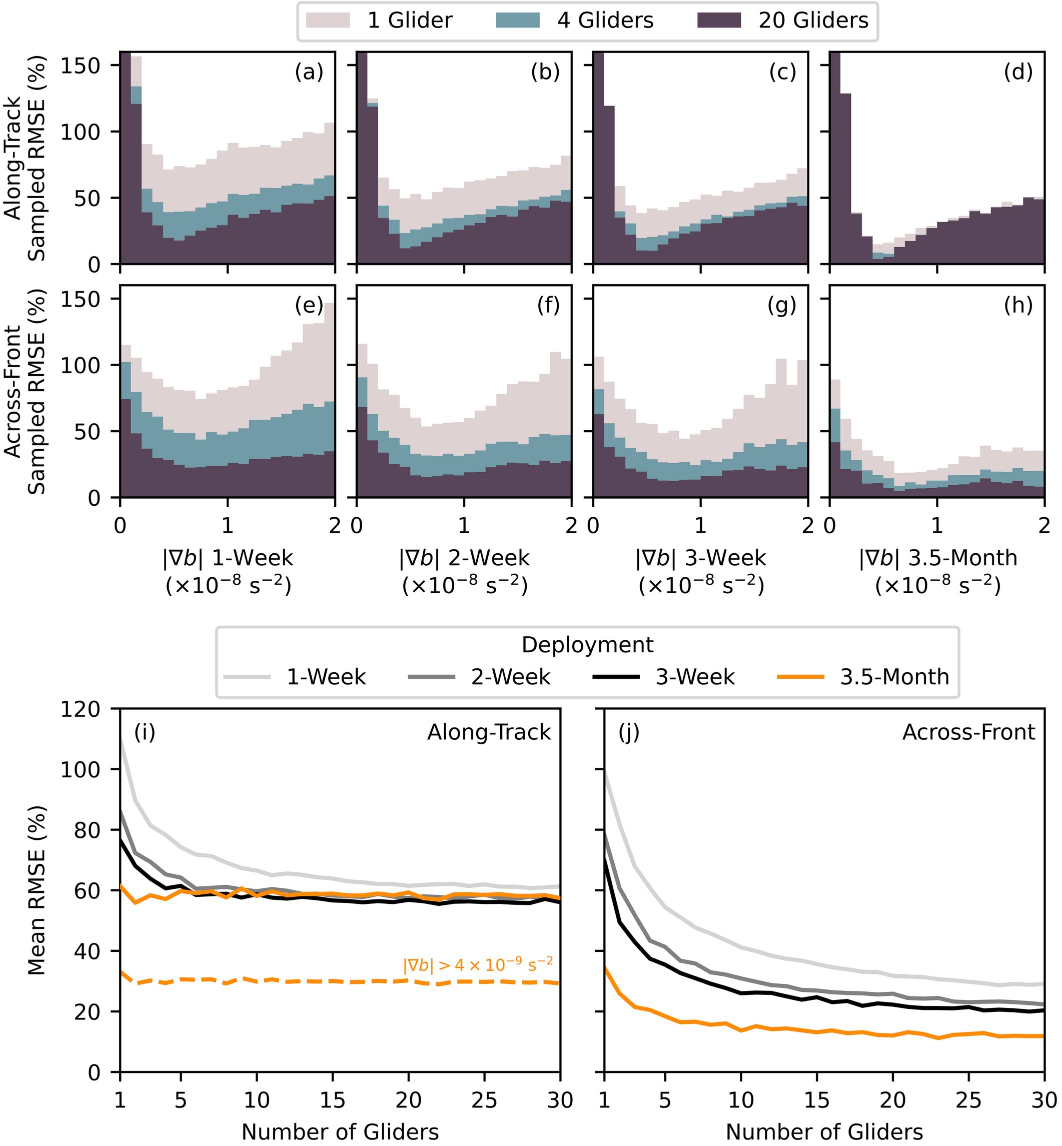

FIG. 5. RMSE of the bootstrapped glider samples compared with the model mean for the PDF of buoyancy gradients. Each individual bootstrapped sample and the model mean are associated with shading and the orange bars in Fig. 4. Columns and rows of (a)–(h) are arranged as in Fig. 4. The lower panels show mean RMSE, averaged across all buoyancy gradient bins shown in (a)–(h) for (i) along-track and (j) across-front glider sampling. The line color corresponds to the deployment time. The dashed line in (i) shows the 3.5-month deployment where the mean is restricted to buoyancy gradients with magnitudes greater than 4×10^{-9} s^{-2}.

the buoyancy gradients. Increasing the number of gliders not only reduces the uncertainty in the data but also results in a convergence to a lower magnitude of buoyancy gradients in comparison to the model median (Fig. 6a).

To understand the implications of the differences that appear in the time series, we present the detected change in the magnitude and variability of buoyancy gradients over time (Figs. 6b,c). For each bootstrapped sample, we calculate a 1-week temporal mean and standard deviation for two time periods ending on 1 January 2013 and 1 March 2013, find the difference between each time, and present the spread in the differences. We then replicate the process for the median, upper decile, and lower deciles of the buoyancy gradients in the model. The orange markers in Fig. 6a provide an illustration of differences for the model median.

With one glider, the uncertainty is large for the estimates of seasonal changes in the temporal mean (Fig. 6b), which is associated with the broad range of estimates of the buoyancy at any one time. Some estimates using one glider are in opposition to the modeled trend and return a negative change, indicating an increase in buoyancy gradients over time. Increasing the number of gliders deployed reduces the uncertainty and

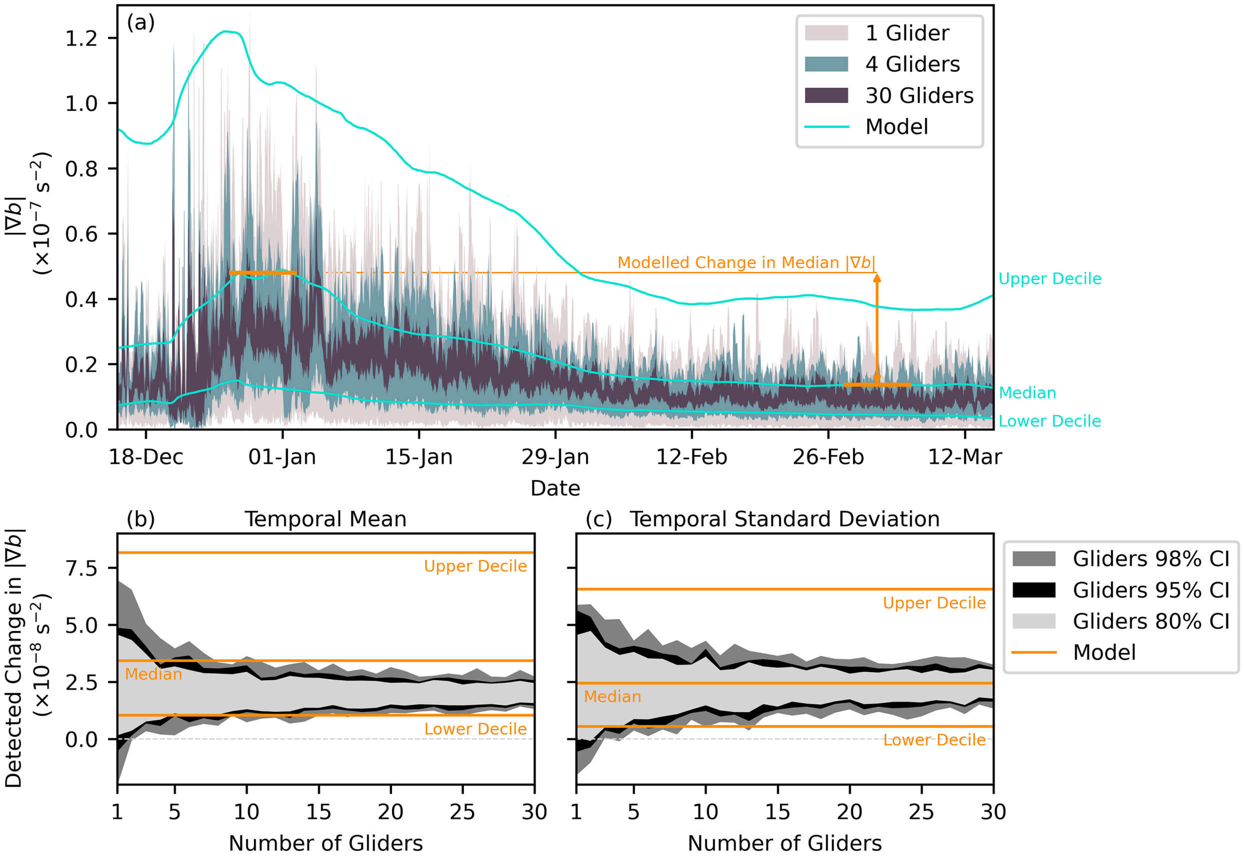

FIG. 6. Detection of seasonal change in buoyancy gradients using the along-track glider samples for differing numbers of gliders deployed. (a) Time series of buoyancy gradients observed across the model domain compared against glider samples. The model statistics show the upper decile, median, and lower decile taken laterally at a 10-depth of the 7-day rolling mean. (b) Detected change in buoyancy gradients between temporal means taken between two 7-day means ending on 1 Jan 2013 and 1 Mar 2013. (c) As in (b), but using a temporal standard deviation instead of a mean. Orange in (a) demonstrates an example of the difference calculation. Horizontal orange lines in (b) and (c) represent the change observed in the model, and shading displays the spread in glider-observed values for the 80% (light gray), 95% (black), and 98% (dark gray) confidence intervals. All data presented are restricted to the southern half of the domain.

the estimates converge. Using two or more gliders, the lower bound for the 98% confidence interval becomes positive. Comparing the change detected by the gliders against the modeled change, we see that the gliders tend to underestimate the seasonal change in the weekly temporal mean. The buoyancy gradients are consistently underestimated throughout the time series, returning reduced magnitudes in comparison to the model (Fig. 6a). Thus, the underestimation of seasonal change in the temporal mean (Fig. 6b) is likely to stem from taking a difference between two values that are themselves underestimates.

The temporal standard deviation represents the temporal variability associated with passing fronts and eddies. The change in standard deviation over time is (mostly) positive for the modeled change and the detected change from glider sampling, indicating a general reduction in temporal variability over time (Fig. 6c). As with the temporal mean, the standard deviation reveals a reduction in the uncertainty of the glider estimates with increasing numbers of gliders. However, the reduction with each additional glider appears to be more gradual. A further difference with the temporal mean is that the change in temporal standard deviation is accurately estimated by the glider sampling. The accuracy in recovering the standard deviation indicates that the temporal variability is less dependent on the sampling bias discussed in section 4a, and capturing a change in gradients over time is not so conditional on the direction that the glider travels.

d. Dives and climbs

Glider sampling uses a dive–climb pattern in the vertical to capture the depth structure of the ocean (see Fig. 2). Increasing demands on the mission length have generated questions regarding the most efficient use of dives. For example, in section 4b, we showed that in some circumstances, increasing the deployment length can improve the certainty of the data. Reducing the number of dives during a deployment has the potential to save battery and thus extend the scope of a single glider campaign, but this is at the detriment of data resolution. In Fig. 7, we present the wavenumber power spectral density of potential density when using the full reference glider path

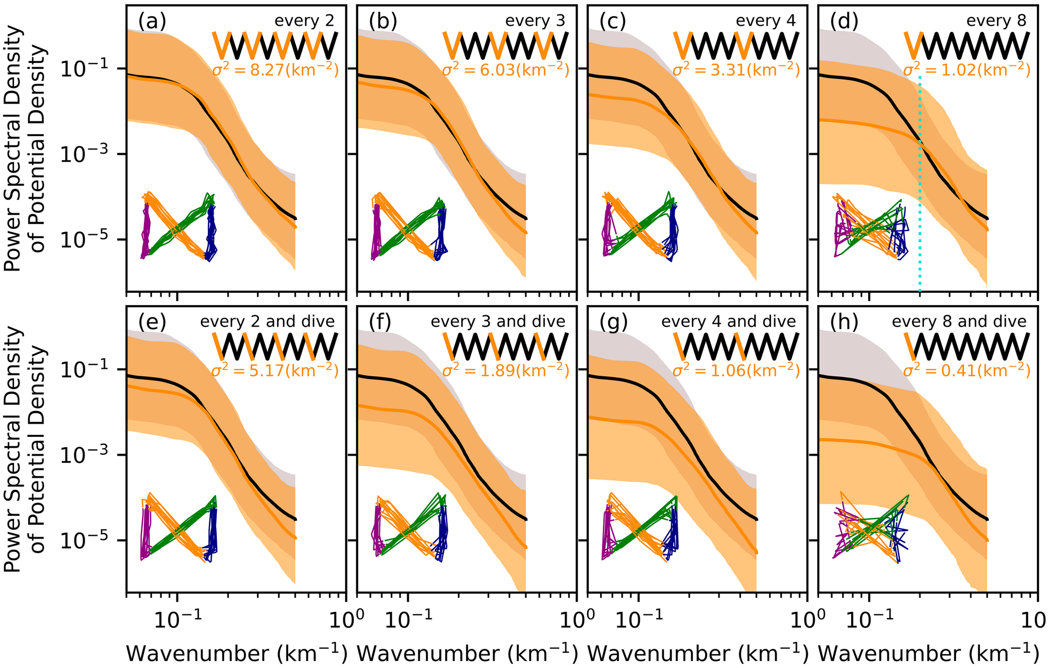

FIG. 7. Wavenumber power spectral density of the potential density ρ_θ in the upper ocean for differing dive–climb sampling patterns at a 10-m depth. Each column shows the effect of reducing sampling (orange) relative to the reference case (black), sampling (a),(e) every two, (b),(f) every three, (c),(g) every four, and (d),(h) every eight dive–climb pairs. The rows compare the effect of retaining both the (top) dive and climb vs (bottom) sampling the dive only. The resulting interpolated lateral path associated with each sampling strategy is displayed in the bottom left of each panel. The signal variance σ^2 for reduced sampling, defined as the integral of the power spectral density, is displayed in the top right of each panel for comparison against the reference case variance of 8.88 km^{-2}. The dashed turquoise line in (e) signifies the 5-km horizontal scale.

compared to reducing the number of dives and climbs, exploring the impacts of a chosen dive pattern on the spatial scales recovered. Unlike the probability density function (Fig. 4), power spectrum has the advantage that the distribution is absolute. A loss of signal at any particular frequency does not influence the magnitude at other frequencies.

The path of the glider consists of four transects. Each choice of the dive–climb configuration alters the structure of these transects through reducing the resolution in the horizontal. Each panel in Fig. 7 has an illustration of how the transects change for each configuration presented. The reductions in sampling augment the transects in a way that the paths begin to no longer represent the original bow-tie and transects are no longer distinct.

Figure 7a shows that halving the number of dive–climb pairs has no significant impact at all spatial scales and fidelity of the data is retained. As more dive–climb pairs are removed, the alignment with the full sampling patten is reduced. Sampling every eight pairs leads to a loss of information at the large scales, and this is reflected in the spectral curves (Fig. 7d). At scales smaller than 5 km, the slopes broadly agree but reduced sampling leads to a flattening of the profile above this distance. The reduced sensitivity below 5 km could be a reflection of the model resolution.

In addition to making fewer dives, the amount of time spent with sampling switched on can also influence battery demands. Figures 7e and 7f show the result of combining reduced dives with sampling only on the dive portion of a dive–climb cycle. When comparing the dive-only sampling with half the number of dives (Fig. 7e), there is minimal impact on the spectral curves, with loss of information focused at the smallest and largest spatial scales and a reduced variance of 5.17 km^{-2} compared to that of 8.88 km^{-2} for the unaltered path. Sampling in dive–climb pairs retains the horizontal resolution on small scales, regardless of the number of pairs that are removed. On the other hand, stopping sampling on the climbs results in the highest resolution being dependent on the distance between dives. As the number of dives is reduced to every three (Fig. 7f) and every four (Fig. 7g) cycles, there is a more obvious loss of fidelity in comparison to using full dive–climb pairs (Figs. 7b,c). Extending to every eight dives, differences with the full sampling method continue to increase, but the errors in the dive–climb (Fig. 7d) and dive-only methods (Fig. 7h) are comparable.

The purpose of this analysis is to demonstrate the potential benefits of sampling less frequently, and practical applications of such measures have not been accounted for in the experimental design. Reduced sampling introduces choices around glider piloting. For example, removing dive–climb pairs, while maintaining the length of deployment, results in gliders remaining at the surface for longer periods of time, increasing the risk of drift. As highlighted by Todd (2020), drift can be significant in locations with strong currents and has the potential to cause sizeable disruption to piloting. A mitigation technique is to have gliders loiter below the pycnocline, where horizontal currents can be weaker. Alternatively, the removal of dive–climb pairs could be implemented by switching sensors off at various intervals while maintaining the dive–climb pattern. This strategy would save on battery consumption without impacting the level of drift. The nuances of these decisions are not addressed in our analysis, and dive patterns analyzed here should be viewed as a basis for decision-making and not an explicit method of piloting.

e. Geometry

The bow-tie pattern is a common sampling strategy for observational campaigns that focus on the submesoscale dynamics. The path is split into four transects, two parallel and two crossing. It is currently unknown whether potential benefits can be made from adjusting this strategy. We examine this by sampling the model with a subset of the bow-tie transects and testing the full bow-tie in a different orientation to reveal if there is any bias associated with the orientation of individual transects or the full bow-tie.

Figure 8 shows the difference in the buoyancy gradients between the model and the glider samples for each choice of path depicted. The calculation is made by first finding a weekly rolling mean and standard deviation of buoyancy gradients for each glider deployment and corresponding patch of model that is sampled. The data from each patch then have a spatial mean applied. This leaves a set of 100 glider deployments and associated model patches with the single dimension of time. A difference is then calculated between each glider and corresponding model patch in the set, giving 100 error estimates for both the temporal (weekly) mean and temporal standard deviation. Figure 8 presents the mean and standard deviation of these errors.

For the temporal mean, using crossing transects as opposed to the full bow-tie has little impact on neither the mean error nor the spread in the error, with both bars appearing almost identical (Fig. 8a). The choice of parallel paths affects the spread, displaying a larger range of errors, but the mean difference remains the same. Then similarly, rotating the bow-tie increases the spread but not the mean error. The larger spread in the rotation and parallel examples indicates that the reliability of these sampling patterns may depend more on the deployment location than the cross- and bow-tie patterns. However, since the mean error remains similar across the examples, these results suggest that samples from the parallel and crossing transects do not exhibit any particular bias, and the orientation of these transects is not of specific importance when it comes to measuring buoyancy gradients at fine spatial scales. This is not necessarily the case when sampling at different spatial scales or location. We suspect that the invariance to alterations in the glider pattern is a product of the isotropic composition of dynamical features, associated with both fine-scale dynamics and the modeled location (see section 3). Were this analysis to be repeated at a location that exhibited a strong directional bias in the buoyancy gradients, different results are expected to emerge.

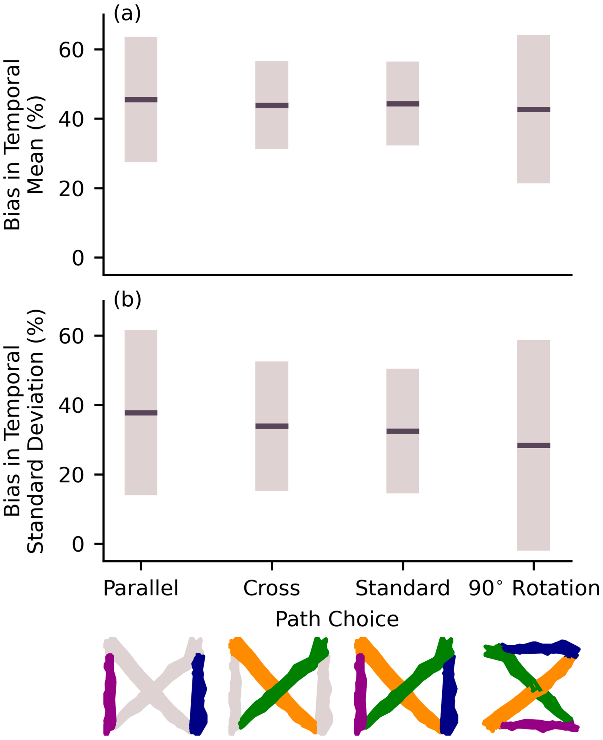

FIG. 8. The percentage difference between the modeled buoyancy gradients and that measured by the virtual glider sampling of the model data for (a) the weekly temporal mean and (b) weekly temporal standard deviation. Bars represent the spread across the 100 glider deployments, with the line showing the mean across the deployments. Each bar is for a differing choice of path configuration, which is depicted beneath.

The results are broadly similar for the temporal standard deviation (Fig. 8b), except that the mean difference is marginally smaller for the rotated orientation and larger for the parallel configuration. Aside from this, there is slightly larger spread in the differences across all choices. Overall, the outcome remains that, under these conditions, the sampling error is not particularly sensitive to the orientation of transects and/or bow-tie. Again, this result may differ in a region where there is strong directional bias in the buoyancy gradients.

5. Discussion

The focus of this investigation is on the evaluation of existing sampling strategies in recovering buoyancy gradients associated with (sub)mesoscale dynamics. We have raised the presence of a sampling bias as a core source error for glider estimates of buoyancy gradients. Questions remain on the key aspect of glider deployments such as 1) the potential of alternative path choices outside of the bow-tie, 2) the fidelity of glider-derived velocity measurements, and 3) the influence of dynamical features below the submesoscale that are responsible for aliasing of glider measurements. We will now discuss each of these points in further detail to provide further

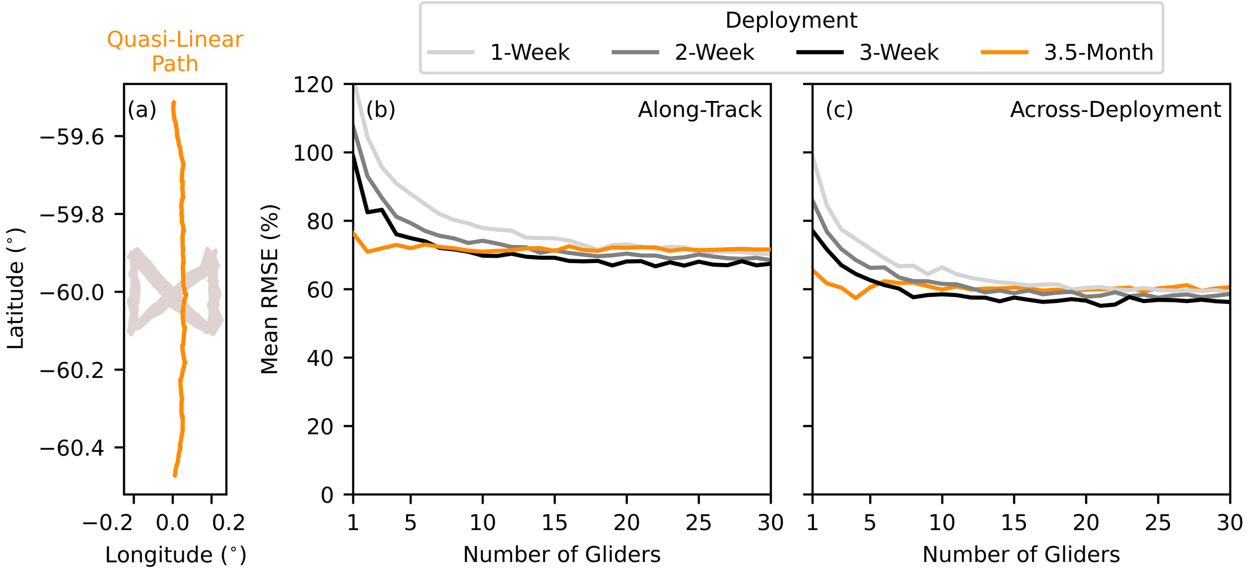

FIG. 9. RMSE of the bootstrapped glider samples compared with the model mean for the PDF of buoyancy gradients for a quasi-linear deployment path. (a) The path of the quasi-linear deployment (orange) compared to the bow-tie (gray). The right-hand panels show the RMSE averaged across all buoyancy gradient bins, where buoyancy gradients are calculated (b) along-track and (c) across two simultaneous deployments. The line color corresponds to the deployment time.

context to our analysis. In addition, we postulate possible avenues for tackling the sampling bias going forward.

a. *Examples of alternative sampling*

The evaluation has centered on bow-tie sampling, a common choice in the open ocean, but other path choices exist. Quasi-linear paths can be used as an alternative to the bow-tie in order to cover greater distances during deployments. Here, we evaluate the sampling bias associated with quasi-linear transects to provide context of the bow-tie deployment against other existing path choices.

1) QUASI-LINEAR GLIDER PATH

The ROAM-MIZ campaign used quasi-linear transects to address large-scale features (e.g., Giddy et al. 2021). An example is shown in Fig. 9a and consists of a north–south-oriented path, with slight meridional distortion due to large-scale currents. We evaluate repeat transects of this quasi-linear path such that the duration of the deployment matches that of the bow-tie for consistency.

Figure 9b shows the error associated with sampling in the along-track direction using quasi-linear transects. Error is larger for the quasi-linear deployment, converging roughly 10% higher than the estimates for the bow-tie method. In common with the bow-tie, increasing the number of gliders deployed at any one time has little impact on the error unless the deployment length is short (i.e., 1 week). In summary, bow-tie sampling performs better than a quasi-linear deployment at recovering the buoyancy gradient field. We speculate that the bow-tie outperforms the quasi-linear path in this context because a feature can be observed from multiple orientations during its life cycle with this method, while the long-return period (longer than the life cycle of submesoscale features) of the quasi-linear path does not allow resampling.

2) ACROSS-DEPLOYMENT SAMPLING

The along-track sampling is biased toward underestimating buoyancy gradients due to an offset between sampling direction and front orientation. Another potential source bias stems from uncertainty associated with calculating spatial gradients from density measurements offset in time. Quasi-linear deployments provide the opportunity to remove the time offset in gradient calculations by using simultaneous deployments to calculate gradients between the instruments. Using two parallel quasi-linear paths separated by 1/12°, we demonstrate the differences in error when calculating buoyancy gradients across glider deployments in place of sampling along-track.

Figure 9c shows that gradients taken across two simultaneous deployments contribute to a reduction in the error of approximately 10%, but this reduction is smaller than that associated with the across-front sampling (Figs. 5i,j). Despite these modest improvements associated with across-deployment gradient calculations, the benefits of this strategy could be sensitive to deployment choices, such as the distance between the gliders. Further investigation into sensitivities of this sampling strategy is required to fully appreciate the potential benefits of this strategy, especially given that the simultaneous deployment of gliders is easier to implement than capturing across-front calculations. This, however, is beyond the scope of our investigation.

b. *Horizontal velocities*

Absolute horizontal velocities are not directly observable from gliders due to their moving position, and they are derived using methods of varying complexity. The most basic derivation is to calculate the depth-averaged velocity from changes in glider position between dives. To retrieve the vertical structure of horizontal velocities, estimates are required of vertical shear, which are then referenced to the depth-averaged velocity (Todd et al. 2017). There are two primary ways of retrieving vertical shear: it is 1) derived from an upward-looking acoustic Doppler current profiler or 2) calculated from profiles of temperature and salinity using the thermal wind relation. Our investigation does not focus on horizontal velocities, but aspects of our analysis overlap with velocities that are derived from thermal wind because of the reliance on horizontal buoyancy gradients. Our assessment on errors associated with the sampling bias will transfer directly to thermal wind-derived velocities.

c. *Spatial scales and aliasing*

Numerical models are selective of spatial and temporal scales and do not capture the full range of processes. Finite computational resources demand a decision between domain size and resolution. Processes that occur on a finer scale to the resolution are parameterized. Models can also be process selective, and features such as tides can be included or not based on decisions made during configuration. Raw observational data are not scale or process selective, leading to aliasing of high-frequency variability that results from fine-scale turbulent processes or tidal forcing (Rudnick and Cole 2011). Aliasing of features omitted from simulations or ones that occur below spatial and temporal scales of models is not represented by virtual glider deployments of model output. This benefits the analysis by isolating scales of interest and quantifying the uncertainty of sampling processes these scales. The drawback is a loss of fidelity with real sampling, in that sampling errors associated with turbulent scales (or tides) are missing from the uncertainty estimates. It is possible for models to represent aliasing due to features such as tides, as demonstrated by Steinberg and Eriksen (2020). However, capturing aliasing associated with fine-scale features presents a greater challenge due to the resolution required. With a model resolution of approximately 1 km in the horizontal, here we target processes at the submesoscale and above. Were we to capture the full range of scales in this present study, we postulate that uncertainty for each sampling strategy would increase due to the additional background noise augmenting the sampling of lateral buoyancy gradients.

d. *Addressing the sampling bias*

Our results confirm the presence of a sampling bias in ocean glider observations related to the misalignment of glider tracks with frontal features. One method for reducing the impact of this sampling bias is to devise a protocol for targeting fronts using satellite altimetry. Precedence for this is presented by Martin et al. (2009) for the purpose of tracking a mesoscale eddy in near–real time, where glider piloting is adapted according to the interpretation of sea level anomaly maps retrieved twice a week. Extending this method to submesoscale features is made difficult by the short time and spatial scales at which the dynamics evolve. Emerging programs such as the Surface Water and Ocean Topography (SWOT; Morrow et al. 2019) mission may address the spatial component by targeting a 250-m resolution, but with repeats occurring on a 21-day cycle, accounting for temporal variation will remain a challenge. Recent attention on submesoscale dynamics has led to the development of alternative techniques, tailored for identifying the orientation of submesoscale fronts. It has been demonstrated that Lagrangian diagnosis of the surface strain field can be used as a measure of submesoscale frontogenesis due to the correlation between surface strain and lateral buoyancy gradients (Archer et al. 2020; Siegelman et al. 2020). We are yet to see such techniques applied to guiding glider piloting, but future campaigns may benefit from exploring the merits of tracking the strain fields remotely for this purpose.

6. Conclusions

Glider observations are important for understanding submesoscale dynamics. The endurance and autonomy of gliders results in the ability to sample at high spatial resolution without being constrained to ship time. Their high spatial resolution sampling facilitates the recovery of spatial gradients that enable the evaluation of key diagnostics for submesoscale dynamics. However, glider sampling has its limitations. While time series are continuous, the position of samples is constantly evolving and so too is the field that is being observed. Observing eddies and instabilities has the potential to be biased by the sampling strategy that is used. For example, a sampling bias has been identified due to the alignment of glider paths with respect to the tangent of horizontal gradients. The main result of this paper has been to evaluate the uncertainty in observations that arises from existing glider sampling strategies and to identify the most efficient use of resources.

Previous estimates of the sampling bias that results from the inability to sample in the across-gradient direction are 71% (Thompson et al. 2016), 64% (du Plessis et al. 2019), and 51% (Swart et al. 2020). Our results align with the most recent of these estimates, showing a 52% error in the mean across individual samples. When viewing the error across multiple deployments using bootstrapping (see section 2d), the bias skews the distribution of buoyancy gradients retrieved toward smaller gradients. The largest bias occurs in locations coincident with weak fronts or a well-mixed surface layer, where lateral buoyancy gradients are smaller than 4×10^{-9} s^{-2}.

We have quantified the impact of increasing the number of gliders deployed at any one time and altering the length of deployment in the context of this bias. Uncertainty is reduced as the quantity of samples increases, e.g., deploying more gliders for a longer amount of time. However, the mean errors in the distribution of buoyancy gradients, summarized in Table 2, are largely insensitive to both choices. The along-track sampling shows a floor error of 60%. For short deployments (1–3 weeks), errors exceed this floor and are reduced by using multiple gliders, but there are a few improvements beyond using five gliders. On the other hand, when deploying gliders for

TABLE 2. RMSE of glider sampling compared to model for the along-track sampling, with the across-front sampling in parentheses. This is subset of data presented in Fig. 5, displaying the realistic range of choices.

	Deployment length			
No. of gliders	1 week	2 weeks	3 weeks	3.5 months
One glider	110 (99)	86 (78)	76 (70)	61 (34)
Two gliders	90 (82)	72 (61)	68 (49)	56 (26)
Three gliders	81 (68)	69 (52)	64 (43)	58 (21)
Four gliders	78 (61)	65 (43)	61 (37)	57 (20)
Five gliders	74 (54)	64 (41)	61 (35)	60 (18)

several months, the floor in the error is met and there appears to be no benefit of increasing the number of gliders. This result is a product of the along-track sampling. When repeated for samples that align at tangents to the gradients, not only do errors dramatically decrease but they are also sensitive to all choices of the glider number and the length of sampling. When calculating buoyancy gradients from along-track samples, the mean distribution across deployments is biased toward small buoyancy gradients and does not align with the model distribution. Therefore, reducing uncertainty by increasing the number of samples does not reduce the mean error. With existing sampling strategies, from the perspective of efficiency, we recommend using the minimum resources required to achieve an error floor of 60% shown in Table 2.

The underestimation of buoyancy gradients due to the sampling bias influences the accuracy of estimating seasonal change. We have shown that uncertainties in estimates reduce as more gliders are deployed but these converge to detect a smaller seasonal change than is observed in the model. Since the magnitude of buoyancy gradients is underestimated at all times, the change in the magnitude is also underestimated. For changes in the temporal standard deviation (variability), the result is different. Gliders do well at capturing the seasonality in short-term variability, and estimates are improved with an increased number of gliders. Although the experiments have been configured for the marginal ice zone, the results can be applied more broadly to all locations with seasonality.

We have shown that increasing the length of glider deployments can increase the certainty of the data retrieved. The length of glider deployments can be limited by battery life, and we have investigated methods for conserving battery during deployment. Gliders profile in the vertical by periodically diving and climbing along their path. Reducing the number of dives or the length of time that sensors are utilized is one method for lowering battery demands. Examining the impact of these choices on the quality of the data retrieved, we have found that halving the number of dive–climb pairs has a negligible negative impact on capturing the dynamics. Then, reducing the number of dives by three- or fourfold, the data begin to deteriorate. We have tested the impact of only sampling on the dive of a dive–climb pair, in addition to removing dive–climb pairs. We find a minor difference when halving the dives with the dive-only strategy, but removing full dive–climb pairs is superior upon further reductions. The results indicate that future campaigns may be able to extend battery life without the loss of determining the buoyancy gradients by reducing the frequency of dives. The practical implementation of reduced dives remains to be determined, and questions exist around piloting in the downtime between profiles.

Lastly, we have analyzed the potential sensitivity of the sampling bias to each individual transect and the orientation of the bow-tie. We find that the underestimation of buoyancy gradients is not dependent on the choice of the glider's lateral path. We speculate that this is the result of the chaotic and isotropic character of the eddy field.

We find that the sampling bias associated with glider paths aligning at an angle to the buoyancy gradients is a leading source of errors. Benefits of campaign design choices remain limited unless the bias is reduced, and future work could focus on innovations aimed at reducing the sampling bias. Glider deployments commonly recover buoyancy gradients in order to parameterize metrics associated with submesoscale dynamics such as heat fluxes associated with restratification by baroclinic instability (e.g., Thompson et al. 2016; du Plessis et al. 2019; Giddy et al. 2021). There remains a knowledge gap on the level of accuracy provided by parameterizations derived from glider measurements. As suggested by a reviewer, methods developed in this paper could be extended to evaluate the implementation-parameterized metrics in this context and a future investigation could be established to address this knowledge gap.

Acknowledgments. This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement 821001. Md. P. is supported by the European Union Marie Skłodowska-Curie individual fellowship under project ID 101032683. A. B. is supported by the ENCORE programme (NERC Grant NE/V013254/1). S. S. is supported by Wallenberg Academy Fellowship (WAF 2015.0186) and the Swedish Research Council (VR 2019-04400).

Data availability statement. Data associated with this paper can be found on Zenodo. Digital object identifiers are available for the processed data (https://doi.org/10.5281/zenodo.10498363), the NEMO configuration files (https://doi.org/10.5281/zenodo.7817589), and the code for processing the output and constructing the configuration (https://doi.org/10.5281/zenodo.10498451).

REFERENCES

Adcroft, A., and J.-M. Campin, 2004: Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modell.*, **7**, 269–284, https://doi.org/10.1016/j.ocemod.2003.09.003.

——, C. Hill, and J. Marshall, 1997: Representation of topography by shaved cells in a height coordinate ocean model. *Mon. Wea. Rev.*, **125**, 2293–2315, https://doi.org/10.1175/1520-0493(1997)125<2293:ROTBSC>2.0.CO;2.

Archer, M., A. Schaeffer, S. Keating, M. Roughan, R. Holmes, and L. Siegelman, 2020: Observations of submesoscale variability and frontal subduction within the mesoscale eddy field of the Tasman Sea. *J. Phys. Oceanogr.*, **50**, 1509–1529, https://doi.org/10.1175/JPO-D-19-0131.1.

Armitage, T. W. K., R. Kwok, A. F. Thompson, and G. Cunningham, 2018: Dynamic topography and sea level anomalies of the Southern Ocean: Variability and teleconnections. *J. Geophys. Res. Oceans*, **123**, 613–630, https://doi.org/10.1002/2017JC013534.

Bachman, S. D., J. R. Taylor, K. A. Adams, and P. J. Hosegood, 2017: Mesoscale and submesoscale effects on mixed layer depth in the Southern Ocean. *J. Phys. Oceanogr.*, **47**, 2173–2188, https://doi.org/10.1175/JPO-D-17-0034.1.

Biddle, L. C., and S. Swart, 2020: The observed seasonal cycle of submesoscale processes in the Antarctic marginal ice zone. *J. Geophys. Res. Oceans*, **125**, e2019JC015587, https://doi.org/10.1029/2019JC015587.

Boccaletti, G., R. Ferrari, and B. Fox-Kemper, 2007: Mixed layer instabilities and restratification. *J. Phys. Oceanogr.*, **37**, 2228–2250, https://doi.org/10.1175/JPO3101.1.

Brannigan, L., 2016: Intense submesoscale upwelling in anticyclonic eddies. *Geophys. Res. Lett.*, **43**, 3360–3369, https://doi.org/10.1002/2016GL067926.

——, D. P. Marshall, A. Naveira-Garabato, and A. J. G. Nurser, 2015: The seasonal cycle of submesoscale flows. *Ocean Modell.*, **92**, 69–84, https://doi.org/10.1016/j.ocemod.2015.05.002.

——, ——, A. C. Naveira Garabato, A. J. G. Nurser, and J. Kaiser, 2017: Submesoscale instabilities in mesoscale eddies. *J. Phys. Oceanogr.*, **47**, 3061–3085, https://doi.org/10.1175/JPO-D-16-0178.1.

Brodeau, L., B. Barnier, A.-M. Treguier, T. Penduff, and S. Gulev, 2010: An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modell.*, **31**, 88–104, https://doi.org/10.1016/j.ocemod.2009.10.005.

Buckingham, C. E., and Coauthors, 2016: Seasonality of submesoscale flows in the ocean surface boundary layer. *Geophys. Res. Lett.*, **43**, 2118–2126, https://doi.org/10.1002/2016GL068009.

Dove, L. A., A. F. Thompson, D. Balwada, and A. R. Gray, 2021: Observational evidence of ventilation hotspots in the Southern Ocean. *J. Geophys. Res. Oceans*, **126**, e2021JC017178, https://doi.org/10.1029/2021JC017178.

du Plessis, M., S. Swart, I. J. Ansorge, and A. Mahadevan, 2017: Submesoscale processes promote seasonal restratification in the Subantarctic Ocean. *J. Geophys. Res. Oceans*, **122**, 2960–2975, https://doi.org/10.1002/2016JC012494.

——, ——, ——, ——, and A. F. Thompson, 2019: Southern Ocean seasonal restratification delayed by submesoscale wind–front interactions. *J. Phys. Oceanogr.*, **49**, 1035–1053, https://doi.org/10.1175/JPO-D-18-0136.1.

Fox-Kemper, B., R. Ferrari, and R. Hallberg, 2008: Parameterization of mixed layer eddies. Part I: Theory and diagnosis. *J. Phys. Oceanogr.*, **38**, 1145–1165, https://doi.org/10.1175/2007JPO3792.1.

Giddy, I., S. Swart, M. du Plessis, A. F. Thompson, and S.-A. Nicholson, 2021: Stirring of sea-ice meltwater enhances submesoscale fronts in the Southern Ocean. *J. Geophys. Res. Oceans*, **126**, e2020JC016814, https://doi.org/10.1029/2020JC016814.

Horvat, C., E. Tziperman, and J.-M. Campin, 2016: Interaction of sea ice floe size, ocean eddies, and sea ice melting. *Geophys. Res. Lett.*, **43**, 8083–8090, https://doi.org/10.1002/2016GL069742.

Large, W. G., and S. G. Yeager, 2004: Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies. NCAR Tech. Note NCAR/TN-460+STR, 105 pp., https://doi.org/10.5065/D6KK98Q6.

Lu, K., T. Weingartner, S. Danielson, P. Winsor, E. Dobbins, K. Martini, and H. Statscewich, 2015: Lateral mixing across ice meltwater fronts of the Chukchi Sea shelf. *Geophys. Res. Lett.*, **42**, 6754–6761, https://doi.org/10.1002/2015GL064967.

Manucharyan, G. E., and A. F. Thompson, 2017: Submesoscale sea ice-ocean interactions in marginal ice zones. *J. Geophys. Res. Oceans*, **122**, 9455–9475, https://doi.org/10.1002/2017JC012895.

Martin, J. P., C. M. Lee, C. C. Eriksen, C. Ladd, and N. B. Kachel, 2009: Glider observations of kinematics in a Gulf of Alaska eddy. *J. Geophys. Res.*, **114**, C12021, https://doi.org/10.1029/2008JC005231.

Marzocchi, A., J. J.-M. Hirschi, N. P. Holliday, S. A. Cunningham, A. T. Blaker, and A. C. Coward, 2015: The North Atlantic subpolar circulation in an eddy-resolving global ocean model. *J. Mar. Syst.*, **142**, 126–143, https://doi.org/10.1016/j.jmarsys.2014.10.007.

McWilliams, J. C., 2016: Submesoscale currents in the ocean. *Proc. Roy. Soc.*, **472A**, 20160117, https://doi.org/10.1098/rspa.2016.0117.

Meunier, T., and Coauthors, 2018: Intrathermocline eddies embedded within an anticyclonic vortex ring. *Geophys. Res. Lett.*, **45**, 7624–7633, https://doi.org/10.1029/2018GL077527.

Milinski, S., N. Maher, and D. Olonscheck, 2020: How large does a large ensemble need to be? *Earth Syst. Dyn.*, **11**, 885–901, https://doi.org/10.5194/esd-11-885-2020.

Morrow, R., and Coauthors, 2019: Global observations of fine-scale ocean surface topography with the Surface Water and Ocean Topography (SWOT) mission. *Front. Mar. Sci.*, **6**, 232, https://doi.org/10.3389/fmars.2019.00232.

Nicholson, D., S. Emerson, and C. C. Eriksen, 2008: Net community production in the deep euphotic zone of the subtropical North Pacific gyre from glider surveys. *Limnol. Oceanogr.*, **53**, 2226–2236, https://doi.org/10.4319/lo.2008.53.5_part_2.2226.

Roquet, F., D. Ferreira, R. Caneill, D. Schlesinger, and G. Madec, 2022: Unique thermal expansion properties of water key to the formation of sea ice on Earth. *Sci. Adv.*, **8**, eabq0793, https://doi.org/10.1126/sciadv.abq0793.

Rosso, I., A. M. Hogg, P. G. Strutton, A. E. Kiss, R. Matear, A. Klocker, and E. van Sebille, 2014: Vertical transport in the ocean due to sub-mesoscale structures: Impacts in the Kerguelen region. *Ocean Modell.*, **80**, 10–23, https://doi.org/10.1016/j.ocemod.2014.05.001.

Rudnick, D. L., and S. T. Cole, 2011: On sampling the ocean using underwater gliders. *J. Geophys. Res.*, **116**, C08010, https://doi.org/10.1029/2010JC006849.

Sallée, J. B., and Coauthors, 2023: Southern Ocean carbon and heat impact on climate. *Philos. Trans. Roy. Soc.*, **A381**, 20220056, https://doi.org/10.1098/rsta.2022.0056.

Siegelman, L., P. Klein, P. Rivière, A. F. Thompson, H. S. Torres, M. Flexas, and D. Menemenlis, 2020: Enhanced upward heat transport at deep submesoscale ocean fronts. *Nat. Geosci.*, **13**, 50–55, https://doi.org/10.1038/s41561-019-0489-1.

Steinberg, J. M., and C. C. Eriksen, 2020: Glider sampling simulations in high-resolution ocean models. *J. Atmos. Oceanic Technol.*, **37**, 975–992, https://doi.org/10.1175/JTECH-D-19-0200.1.

Swart, S., M. D. du Plessis, A. F. Thompson, L. C. Biddle, I. Giddy, T. Linders, M. Mohrmann, and S.-A. Nicholson, 2020: Submesoscale fronts in the Antarctic marginal ice zone and their response to wind forcing. *Geophys. Res. Lett.*, **47**, e2019GL086649, https://doi.org/10.1029/2019GL086649.

Taylor, J. R., and A. F. Thompson, 2023: Submesoscale dynamics in the upper ocean. *Annu. Rev. Fluid Mech.*, **55**, 103–127, https://doi.org/10.1146/annurev-fluid-031422-095147.

Thomas, L. N., J. R. Taylor, R. Ferrari, and T. M. Joyce, 2013: Symmetric instability in the Gulf Stream. *Deep-Sea Res. II*, **91**, 96–110, https://doi.org/10.1016/j.dsr2.2013.02.025.

Thompson, A. F., A. Lazar, C. Buckingham, A. C. Naveira Garabato, G. M. Damerell, and K. J. Heywood, 2016: Open-ocean submesoscale motions: A full seasonal cycle of mixed layer instabilities from gliders. *J. Phys. Oceanogr.*, **46**, 1285–1307, https://doi.org/10.1175/JPO-D-15-0170.1.

Todd, R. E., 2020: Export of middle Atlantic bight shelf waters near Cape Hatteras from two years of underwater glider observations. *J. Geophys. Res. Oceans*, **125**, e2019JC016006, https://doi.org/10.1029/2019JC016006.

——, D. L. Rudnick, J. T. Sherman, W. B. Owens, and L. George, 2017: Absolute velocity estimates from autonomous underwater gliders equipped with Doppler current profilers. *J. Atmos. Oceanic Technol.*, **34**, 309–333, https://doi.org/10.1175/JTECH-D-16-0156.1.

Vancoppenolle, M., and Coauthors, 2023: SI3, the NEMO Sea Ice Engine. Zenodo, accessed 13 January 2023, https://doi.org/10.5281/zenodo.7534900.

Viglione, G. A., A. F. Thompson, M. M. Flexas, J. Sprintall, and S. Swart, 2018: Abrupt transitions in submesoscale structure in southern Drake Passage: Glider observations and model results. *J. Phys. Oceanogr.*, **48**, 2011–2027, https://doi.org/10.1175/JPO-D-17-0192.1.

Wenegrat, J. O., J. Callies, and L. N. Thomas, 2018: Submesoscale baroclinic instability in the bottom boundary layer. *J. Phys. Oceanogr.*, **48**, 2571–2592, https://doi.org/10.1175/JPO-D-17-0264.1.