LIMNOLOGY and OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 22, 2024, 840–852

doi: 10.1002/lom3.10641

Optimizing observational arrays for biogeochemistry in the tropical Pacific by estimating correlation lengths

Winnie U. Chu,[1]* **Matthew R. Mazloff**,[2] **Ariane Verdy**,[2] **Sarah G. Purkey**,[2] **Bruce D. Cornuelle**[2]

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA
[2]Scripps Institution of Oceanography, La Jolla, California, USA

Abstract

Global climate change has impacted ocean biogeochemistry and physical dynamics, causing increases in acidity and temperature, among other phenomena. These changes can lead to deleterious effects on marine ecosystems and communities that rely on these ecosystems for their livelihoods. To better quantify these changes, an array of floats fitted with biogeochemical sensors (BGC-Argo) is being deployed throughout the ocean. This paper presents an algorithm for deriving a deployment strategy that maximizes the information captured by each float. The process involves using a model solution as a proxy for the true ocean state and carrying out an iterative process to identify optimal float deployment locations for constraining the model variance. As an example, we use the algorithm to optimize the array for observing ocean surface dissolved carbon dioxide concentrations (pCO_2) in a region of strong air–sea gas exchange currently being targeted for BGC-Argo float deployment. We conclude that 54% of the pCO_2 variability in the analysis region could be sampled by an array of 50 Argo floats deployed in specified locations. This implies a relatively coarse average spacing, though we find the optimal spacing is nonuniform, with a denser sampling being required in the eastern equatorial Pacific. We also show that this method could be applied to determine the optimal float deployment along ship tracks, matching the logistics of real float deployment. We envision this software package to be a helpful resource in ocean observational design anywhere in the global oceans.

The net uptake of atmospheric carbon dioxide (CO_2) by the global ocean plays a significant role in mitigating the effects of anthropogenic climate change (Landschützer et al. 2014). Yet variability and trends in CO_2 uptake of the regional oceans are still not well understood (Gruber et al. 2023). Previous studies on air–sea CO_2 flux, derived from the partial pressure of CO_2 (pCO_2) in the surface water, share a consensus that more long-term and high-resolution observational data are necessary to better assess how the ocean's carbon uptake is changing (Fay and McKinley 2013).

The tropical Pacific Ocean is an especially critical location to constrain carbon dioxide measurements because it is one of the largest sources of regional CO_2 outgassing (Chavez et al. 1999). It is also a region that experiences highly variable biogeochemical and physical events, resulting from the convergence of multiple climatic and oceanographic phenomena like El Niño Southern Oscillation (ENSO) and Oxygen Minimum Zones (areas which have little to no concentration of oxygen) (Smith et al. 2019). Data platforms relying on autonomous floats, stationary moorings, gliders, and direct measurements observe surface pCO_2 to quantify the ocean's carbon budget. In particular, a Southern Ocean pilot array of BGC-Argo floats with the ability to infer pCO_2 from pH measurements has demonstrated the ability to constrain CO_2 fluxes from profiling floats (Gray et al. 2018; Bushinsky et al. 2019). Recently, to meet the data needs of researchers working on the carbon budget, the scientific community has called for the expansion of the BGC-Argo array to target the tropical Pacific.

Previous studies on observing system design, by Mazloff et al. (2018) and Tardif et al. (2022), have used analyses of correlation scales and covariance, respectively, to inform the observation of carbon, heat, and sea surface temperature in the Southern Ocean. We utilize a combination of these methods and a model solution to optimize the observation strategy of surface carbon dioxide (pCO_2) for the tropical Pacific Ocean. Due to recent efforts to expand the BGC-Argo array, our method is applied to float-based observations in hopes of guiding float deployments in the near future. However, we envision this technique being adapted for a broader range of observational platforms that exist or will be developed in the tropical Pacific region. This paper is a first step

*Correspondence: winniechu@princeton.edu

toward more robust observational coverage of pCO_2 that is informed by statistics.

The metric for validating the effectiveness of the observing system is the estimated uncertainty covariance for an objective mapping reconstruction (Bretherton et al. 1976). In other words, we determine where to observe to minimize the expected misfits between the model and an objective mapping reconstruction of the model. To the extent that the model is a proxy for real ocean variability, this method then yields locations from which to measure that minimize reconstruction and data differences. We make four key simplifying assumptions: (1) Argo floats remain at the deployment location, (2) the ocean statistics are stationary, (3) the model statistics used in our calculations are accurate, and (4) pCO_2 statistics are Gaussian. Future work will improve the algorithm by addressing these assumptions. In particular, Chamberlain et al. (2023) suggest a method to eliminate the first assumption by including float movement in the algorithm.

Other factors in the pCO_2 measurement process may contribute to the effectiveness of the observing system as well. Since BGC-Argo floats infer pCO_2 from pH, they are subject to sources of uncertainties beyond sensor performance and carbonate system equilibrium constants, including methods to infer other water properties and alkalinity in particular (Williams et al. 2017). Given that these sources of uncertainty pertain less to the observational array design and more to the instrument itself, we do not consider these uncertainties in our preliminary assessment of the design algorithm. However, in our concluding remarks, we do speculate on how these uncertainties may affect the algorithm's observational array design.

A primary motivation for this paper is the documentation and sharing of the observing system design algorithm. BGC-Argo implementation strategies tend to assume a uniform float distribution with the number of floats suggested by the median correlation lengths of variables of interest (Biogeochemical-Argo Planning Group 2016). These strategies omit spatial details in the sampling distribution that could potentially decrease the number of floats needed. An additional consideration is that floats are deployed from ships-of-opportunity, a constraint that is explicitly accounted for in our method to increase the realism of the algorithm's proposed designs. Given the high cost of purchasing, deploying, and analyzing the performance of ocean observing platforms, the incentives for algorithms like the one presented here are high. It is vital to determine observing strategies that maximize information gained while minimizing resource costs. While this paper focuses on pCO_2 for its relevance to the global carbon budget and climate change, the method itself can be applied to other variables of interest such as oxygen (O_2) or nitrate (NO_3), with modifications to account for variations in depth. We explain the method in full detail below. The data and float deployment planning software are available at the link provided under the "Data Availability" section.

Methods

Modeling framework

Our observing system design method could draw from observed ocean statistics if they were sufficient (e.g., as one may gain from satellite). Here we use statistics from a general circulation model as a proxy for the true ocean. Though this increases the uncertainty in our conclusions, it still provides value in guiding observational strategies, and the same technique could be used with other data sources. The Massachusetts Institute of Technology General Circulation Model (MITgcm; Campin et al. 2022) simulation we use is coupled to a biogeochemical model evolved from the Biogeochemistry with Light, Iron, Nutrients and Gases (BLING) model (Galbraith et al. 2010) in a regional domain extending from 26°S to 30°N and from 104°E to 68°W. The model grid spacing is 1/6° in latitude and longitude, which is sufficient for resolving the mesoscale pCO_2 variability. There are 51 vertical levels with a resolution of 5 m near the surface and gradually increasing with depth.

The simulation covers the period 2004 through 2019, with hourly atmospheric forcing from the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis version 5 (ERA5; Hersbach et al. 2018). Lateral boundary conditions for the dynamical state variables at the north, south, and west are from the Hybrid Coordinate Ocean Model with Navy Coupled Ocean Data Assimilation (HYCOM/NCODA; http://hycom.org). Vertical mixing is parameterized by the KPP formulation of Large et al. (1994). The physical model setup is similar to the one described in Verdy et al. (2017), with increased horizontal resolution and additional Leith viscosity parameterization (Leith 1996).

The model includes a full description of the carbon system, adapted from the MITgcm simple biogeochemistry model (Dutkiewicz et al. 2006). pCO_2 is diagnosed at every time step from dissolved inorganic carbon (DIC) and alkalinity. Atmospheric pCO_2 concentration is prescribed at the Mauna Loa values (Keeling et al. 2001) while air–sea fluxes of carbon are parameterized following Wanninkhof (1992). Future work should test the sensitivity to other carbon flux parameterizations.

DIC is taken up by biological production, which is limited by light, nitrate, phosphate, and iron, and produced by remineralization of organic matter. Alkalinity is changed by uptake and remineralization of nitrate, as well as production and dissolution of calcium carbonate. Freshwater fluxes affect the surface concentration of DIC and alkalinity. More details on the biogeochemical model can be found in Verdy and Mazloff (2017).

Initial and lateral boundary conditions for biogeochemical tracers are derived from the Global Ocean Data Analysis Project (GLODAPv2; Key et al. 2015; Lauvset et al. 2016), the World Ocean Atlas (WOA18; Garcia et al. 2019*a*,*b*), and a coarse global BLING model simulation.

Simulated fields are compared to the climatological products as well as in situ pCO_2 observations from the Surface Ocean CO_2 Atlas Database (SOCATv6; Bakker et al. 2016). Although the agreement is good overall, there are some mean biases in the model. In particular, near the South American coast, modeled surface oxygen is high and modeled surface nutrients are low compared to WOA18. This is possibly due to insufficient coastal upwelling in the model and should be kept in mind when interpreting the conclusions.

Surface pCO_2 compares qualitatively well with the neural network-based pCO_2 climatology product of Landschützer et al. (2020), though it is biased low in the eastern equatorial Pacific and high in the western and central parts of the basin (Fig. 1). The time variability of simulated pCO_2 is compared to Tropical Atmosphere Ocean (TAO) mooring observations (Sutton et al. 2003) and the Landschützer product in Fig. 2a,b. The model captures aspects of the observed variability, such as the amplitude of the seasonal cycle and the drop in pCO_2 during El Niño events in 2009/2010 and 2015/2016. Sea surface temperature variability is also captured well by the model (Fig. 2c,d).

Further discussion of model fidelity is beyond the scope of this paper, which aims to present a technique for using model statistics to inform float deployments. The broad utility of the method we present means that it can be used with a different model or model runs or observation-derived statistics if available. The results of the analysis may differ, since each statistical inference will have its own set of strengths and weaknesses in representing the variability of CO_2 and other fields. Given that this model produces a realistic representation of pCO_2 that is comparable to observations and other products, we can demonstrate a reasonable use of the float deployment algorithm with this model data.

Statistical analysis

The model represents the current baseline state and not an equilibrated climatological state. Upon initialization, there are some adjustment processes that occur. To avoid spurious statistics, we remove the first 12 months of the pCO_2 time series before analysis. In addition, we remove the annual, semi-annual, 4-month, and 3-month harmonics using a least squares fit. The harmonic cycles are assumed to be due to seasonal changes in temperature and biology. This preprocessing method extracts seasonal cycles so that our analysis of float deployment strategies can be targeted towards irregular distributions of pCO_2 that are independent of regularly occurring seasonal patterns.

Figure 3a shows the total variance of the unprocessed pCO_2 signal, the same signal whose time series is plotted at (0°N, 170°W) and (0°N, 124°W) in Fig. 2. As expected, there is larger variance along the coasts due to runoff, upwelling, biology, and other coastal processes that can impact pCO_2.

To justify our choice to remove the harmonic cycles, we use Fig. 3b to highlight areas where a large percentage of the total variance can be attributed to the annual seasonal cycle (white and light blue regions). While these areas are worth noting, our goal is to observe anomalous pCO_2 signals that cannot be explained by seasonal cycles alone (dark blue regions). The mean percent variance explained by the annual cycle is 50.43%. By contrast, the mean percent variance explained by the semi-annual cycle, 4-month cycle, and 3-month cycle is 3.31%, 0.56%, and 0.18%, respectively, so they are not shown.

Upon removing the 12-, 6-, 4-, and 3-month harmonics, we generate a residual pCO_2 signal which exposes regions whose variance is largely independent of seasonal cycles (Fig. 3c). In

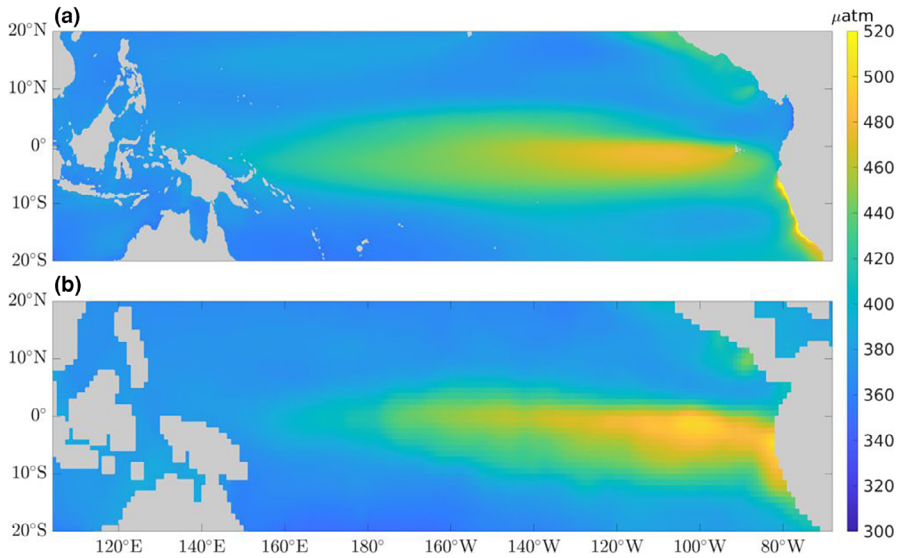

Fig. 1. Mean surface pCO_2 during 2004–2019 in the model (**a**) and in the Landschützer et al. (2020) product (**b**). Gray areas are land and the Gulf of Mexico in (**a**) is also gray.

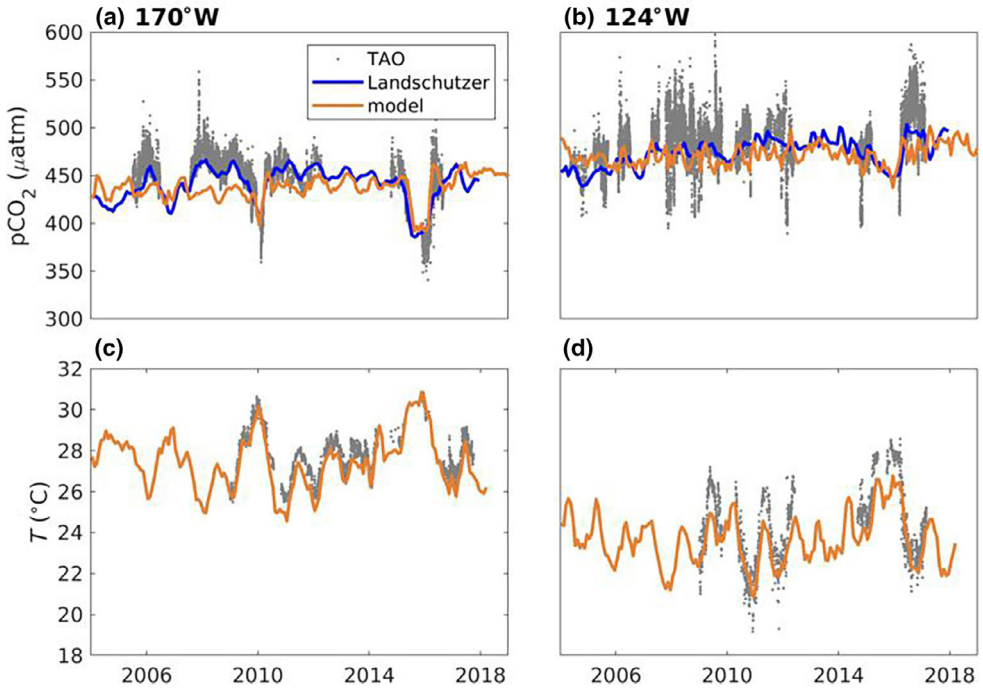

Fig. 2. Surface pCO_2 at (**a**) (0°N, 170°W) and (**b**) (0°N, 124°W) in the model (monthly-averaged), orange; in the Landschützer et al. (2020) monthly product, blue; from TAO moorings, gray. (**c**) Sea surface temperature at (0°N, 170°W) and (**d**) 20 m depth temperature at (0°N, 124°W) in the model and from TAO.

Fig. 3c, areas that are of greater interest to us are bright yellow. They are areas with high variance in the postprocessed signal, meaning pCO_2 trends there are not attributable to seasonal cycles. Instead, they are likely to be driven by yet-to-be-characterized forces. For our goal of studying the large-scale spatial dynamics of pCO_2 variability, we can make a preliminary statement that the yellow regions in Fig. 3c are where float deployment efforts should be focused.

Correlation lengths in the tropical Pacific

The first step in our two-part method for float deployment strategy is to determine the optimal spacing between floats. This can be achieved through the estimation of spatial correlation lengths, which are characteristic lengths that indicate the area over which the observed signal at one location is representative of its surroundings. In terms of resource efficiency, correlation lengths are a useful and simple indicator of the float density needed to observe pCO_2, with the added consideration that high float density is more costly and less likely to be implemented.

Before the correlation lengths can be estimated, the correlation between the residual pCO_2 signal at pairs of model points in the tropical Pacific is calculated. The axes of the model are set such that x is in the longitudinal (zonal) direction and y is in the latitudinal (meridional) direction. Each arbitrary grid point k has model coordinates of (x_k, y_k) and a pCO_2 time series p_k. Relative to point k, surrounding points have locations $(x_k + \Delta x, y_k + \Delta y)$ where $(\Delta x, \Delta y)$ are some specified distance away from grid point k. We calculate the correlation between the pCO_2 time series at the central grid point and the surrounding points with no time lag (time lag = 0). We assume we do not need the entire analysis domain to determine the correlation structure and reduce computational expense by defining "surrounding" as being limited to a maximum of 240 grid points away in the x-direction (i.e., 40° away in longitude) and a maximum of 72 grid points away in the y-direction (i.e., 12° away in latitude).

An example of the correlation between a point of interest and its surroundings—a correlation map—is shown in Fig. 4. Note that each point of interest in the model will have a distinct correlation map. The locations in the map are colored by the value of the correlation coefficient between its pCO_2 time series and the time series of the central grid point k, which is located at (0, 0) on the correlation map and (0°N, 160°E) in the ocean. As seen in Fig. 4, the correlation generally decreases with distance from the central point k.

Once correlation maps have been generated for all the model grid points, we estimate the correlation lengths from each map as a means of reporting spatial variability. This allows us to condense the information from each correlation map into three numbers (semi-major axis length, semi-minor axis length, and angle) representing the scales in two dimensions. We estimate correlation lengths by performing a two-dimensional Gaussian fit to each of the correlation maps. We

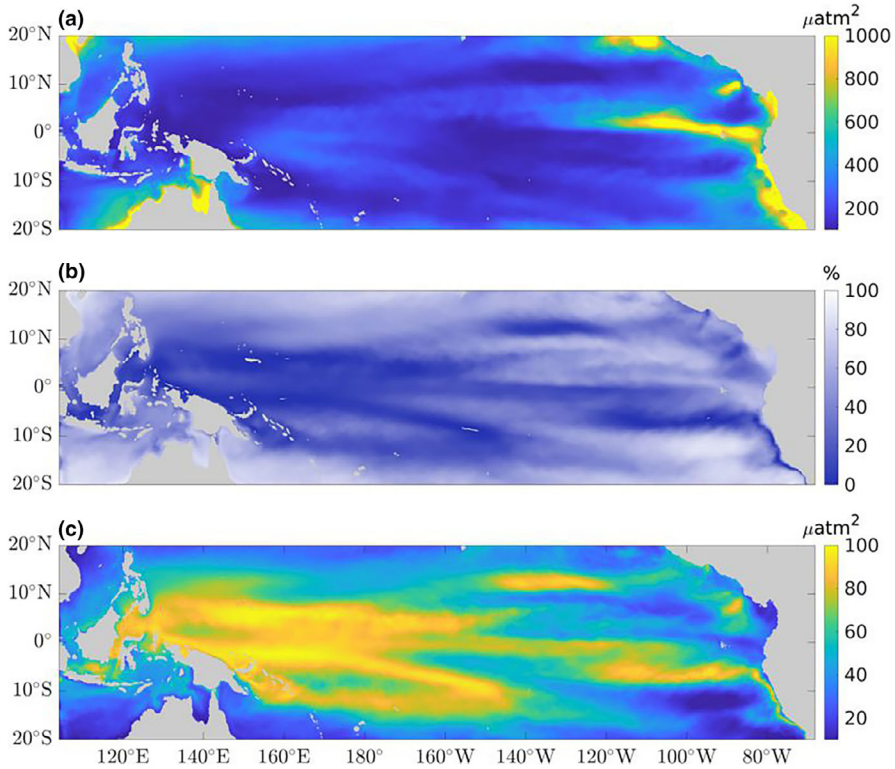

Fig. 3. (**a**) Total variance of the preprocessed pCO_2 signal in μatm^2. Dark blue areas have low variance while bright yellow areas have high variance. Gray regions are land. (**b**) Percent variance explained by the annual pCO_2 cycle. Dark blue areas have low variance explained by the annual cycle alone. (**c**) Residual variance after seasonal cycles are removed in μatm^2. Dark blue areas have low residual pCO_2 variance after seasonal cycles (12-, 6-, 4-, and 3-month cycles) were removed from the original signal. Bright yellow areas have high residual pCO_2 variance and are targets for more observations. The color scale for the residual variance is an order of magnitude smaller than for the original variance.

chose a Gaussian fit because it is intuitive and relatively easier to compute compared to more complicated geometries. The contour of the Gaussian fit at the e-folding scale (when correlation $r \approx 0.37$) is an ellipse, whose semi-major and semi-minor axes lengths are what we here define as the correlation lengths (Fig. 4). The semi-major and semi-minor axes are well suited for an estimation of spatial scales in the approximately meridional and zonal directions.

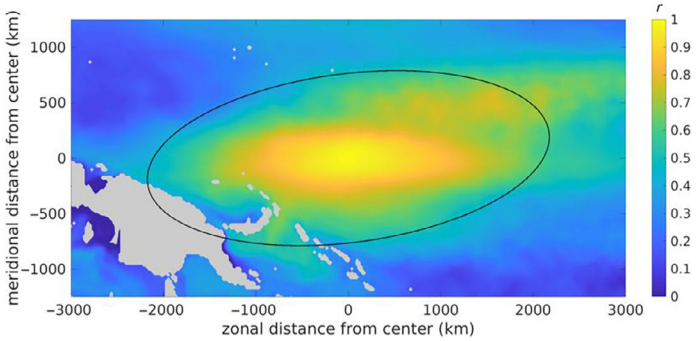

Fig. 4. pCO_2 correlation map centered at (0°N, 160°E). The plot shows the region of data used to make the fit. The e-folding contour of the Gaussian fit is annotated in black. Bright yellow areas indicate a high correlation coefficient for the pCO_2 time series in that location with the central location. Green–blue areas indicate a low correlation coefficient. Gray regions are land.

To perform the Gaussian fitting to each correlation map, we first assume that the map can be described with a Gaussian structure:

$$g_k = e^{-a_1\Delta x^2 - a_2\Delta x\Delta y - a_3\Delta y^2}. \quad (1)$$

Let $\mathbf{a} = [a_1, a_2, a_3]$ be the parameters of the optimal Gaussian fit to the correlation map. We can then estimate the parameters of the optimal Gaussian fit, $\mathbf{a}$, using a regularized least squares regression (Wunsch 1996) as stated below:

$$\mathbf{a} = \mathbf{a}_p + \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{P}^{-1}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{c} - \mathbf{H}\mathbf{a}_p\right), \quad (2)$$

where $\mathbf{a}_p$ is the a priori guess for the fitting parameters, $\mathbf{H}$ is our fitting function, and $\mathbf{R}$ is a matrix representing the uncertainty covariance of the correlation field. Similarly, $\mathbf{P}$ is a matrix representing uncertainty in $\mathbf{a}$ and based on our experience is assumed to be diagonal and isotropic with an initial e-folding value of 283 km in the zonal direction and 163 km in the meridional direction. $\mathbf{c}$ is the negative log of the correlation map unfolded to be a vector.

We find that making an a priori guess, $\mathbf{a}_p$, for the fitting parameters greatly improves the success of the algorithm since there is significant uncertainty in $\mathbf{a}$. Our method to estimate this prior involves dividing the region within the correlation map into 20° sectors (in polar coordinates) and finding the point in each sector where the correlation first falls below 0.8. We then calculate the distance between the central grid point and the point where $r < 0.8$ in each sector. The distance averaged between azimuth sectors at −40° to 40° and 140° to −140° (in polar coordinates with 0° being eastward) becomes the initial guess for the longitudinal (zonal) correlation length. The average distance over the sectors between 40° to 140° and −40° to −140° is our initial guess for the latitudinal (meridional) correlation length.

To estimate $\mathbf{R}$, we assume that noise increases with distance from the central grid point k. This is because we assume that high correlations across long spatial distances are more likely spurious, rather than being driven by the same physical or biogeochemical dynamics. We choose $\mathbf{R}$ based on our initial estimate for $\mathbf{a}_p$ by assuming the correlation uncertainty is 0.01 at k and exponentially grows to 1 with e-folding length scales given by $\mathbf{a}_p$.

As stated, the least squares regression method is sensitive to the a priori guess $\mathbf{a}_p$, which may be different from the actual correlation field. Thus, after estimating the first $\mathbf{a}$, we set this as a new $\mathbf{a}_p$ and perform the regression iteratively until the output, $\mathbf{a}$, and the prior, $\mathbf{a}_p$, are less than 1% different or until the program has iterated six times, at which point we assume that the fit does not converge. However, unlike the first iteration of regression, the subsequent iterations assume an isotropic uncertainty, $\mathbf{P}$, with an e-folding value of 28 km in the zonal direction and 16 km in the meridional direction (one order of magnitude lower than the uncertainty of the first iteration).

Once the iterations are complete, we perform a principal axis rotation which diagonalizes the covariance matrix of the original Gaussian function (Eq. 1). The covariance matrices corresponding to the original and rotated Gaussian are

$$\text{original covariance matrix} \begin{bmatrix} a_1 & \frac{1}{2}a_2 \\ \frac{1}{2}a_2 & a_3 \end{bmatrix} \rightarrow \text{after rotation} \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}, \tag{3}$$

where the diagonal elements of the rotated covariance matrix are the semi-major and semi-minor correlation lengths. The rotated Gaussian parameter $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]$ has the corresponding Gaussian equation:

$$g_k = e^{-\alpha_1 \Delta x'^2 - \alpha_2 \Delta y'^2} \tag{4}$$

and is equivalent to the original Gaussian. In this form, however, we must specify the angle (θ) by which the semi-major and semi-minor axes have been rotated, since $\Delta x'$ and $\Delta y'$ are distances written in terms of the rotated axes x' and y'. The angle is found between the eigenvector corresponding to the largest eigenvalue (the semi-major length) and the x-axis using the four-quadrant inverse tangent (*atan2* function in Matlab).

The Gaussian parameters of the final iteration are reported as correlation lengths, with $\alpha_{1,\text{final}}$ being the semi-major correlation length and $\alpha_{2,\text{final}}$ being the semi-minor correlation length. We also report the angle of rotation, θ, where $\theta = 0°$ is defined as the alignment of the semi-major axis to the orientation of the equator (i.e., the zonal direction). Because we performed a least squares regression, we also calculate the variance of the residuals weighted by their uncertainty, $\mathbf{R}$, which informs the goodness of fit to the correlations. Hence, the Gaussian fitting process summarizes a complicated correlation map into four characteristics that we show below: the semi-major correlation length (Fig. 5a), semi-minor correlation length (Fig. 5b), fraction of variance explained by the fit (Fig. 5c), and the angle of rotation (not shown because 90% of the model points have a predominantly zonally oriented semi-major axis with $-10° < \theta < 10°$). These characteristics are intended to be more easily communicated and interpreted for efficient decision-making of float deployment. As stated previously, a Gaussian fit was chosen for ease of communication, although other models, which may have more characteristic parameters, could be used instead and potentially generate a better fit to the correlation distributions of pCO_2. Nevertheless, the fact that the fit converges for 99.97% of the model points and yields reasonable residuals (Fig. 5c; average $r^2 = 0.85$) is validation that a Gaussian function is a reasonable choice.

Figure 2 shows examples of the temporal signals setting the correlation length scales. We are assessing where these signals are in phase with their surroundings. The observations reveal significantly higher frequency signals, and we anticipate these to have shorter correlation lengths. If one were able to make this correlation length calculation with observations, an option would be to temporally filter the higher frequencies to better focus on the time scales the observing system is aiming to constrain. In essence, this is what our model is doing, as well as the data preprocessing step where we filter out the seasonal harmonics. Another option would be to fit the correlation maps with multiple Gaussians, each with priors that target the different dynamical scales. Figure 6 demonstrates what the algorithm currently fits when faced with multiple correlation scales, along with further discussion.

Float deployment algorithm

Optimal float deployment

Ideal locations for float deployment should best observe the pCO_2 signal over the tropical Pacific basin by maximizing the constrained variance. This section describes the algorithm used to identify the ideal float locations recursively. In the

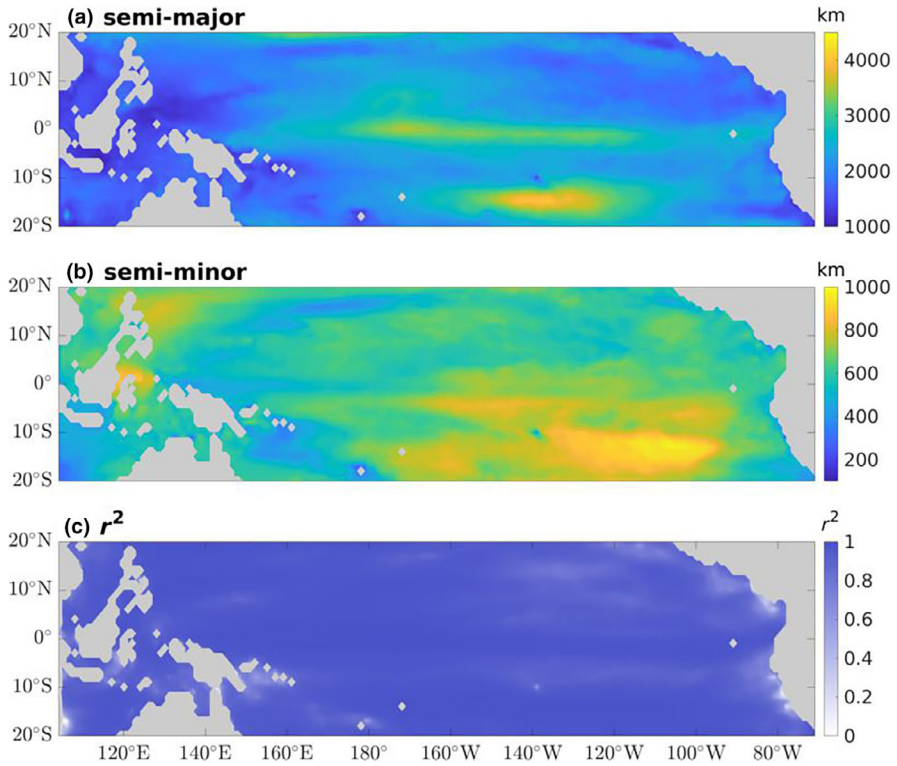

Fig. 5. (**a**) Color represents the semi-major correlation lengths of pCO_2. Semi-major lengths are typically oriented along the longitudinal (zonal) direction, representing the variability of pCO_2 to the east and west. Gray areas are land. (**b**) Color represents the semi-minor correlation length of pCO_2. Semi-minor lengths are typically oriented along the latitudinal (meridional) direction, representing the variability of pCO_2 to the north and south. (**c**) Fraction of variance explained by the fit weighted by uncertainty covariance, **R**, as defined in Eq. 2. Dark blue areas have higher r^2 values, indicating greater confidence in the estimated correlation lengths. Lighter blue/white areas have lower r^2 values, where there is lower confidence in the correlation lengths estimated for those areas.

interest of conserving resources, we also consider the accumulated total estimated variance reduction and the spatially averaged normalized estimated variance reduction, which we here call percent coverage, to determine when the information collected by an additional float may not be worth the deployment expense.

Our algorithm estimates which locations constrain the most information given our pCO_2 model statistics. The objective function we want to maximize is that of the constrained variance, which minimizes the estimated error variance (see Eq. 10 in Bretherton et al. 1976). The equation for variance constrained (**VC**) from a float deployed at location k is

$$\mathbf{VC}_k = \left(\mathbf{PLH}_k^T\right)\left(\mathbf{H}_k\mathbf{PH}_k^T + \mathbf{R}\right)^{-1}\mathbf{H}_k\mathbf{LP}, \tag{5}$$

where we aim to minimize the trace of this matrix over model points with depth exceeding 2000 m as Argo floats only operate in deep waters. **P** and **R** are not the same as used in the "Methods" section. Here, **P** represents the full covariance of pCO_2. $\mathbf{H}_k$ represents where the ocean is sampled (i.e., at point k), and **R** is the assumed uncertainty on the float and is here taken to be negligible. The matrix **L** is a symmetric distance scaling matrix using the correlation lengths that were previously discussed. **L** is 1 at observation points and thus left out of the denominator term $\mathbf{H}_k\mathbf{PH}_k{}^T$, which represents the data variance. **L** goes to 0 for points far from k. Since we are interested in the trace and **R** is small, this is approximately the summed localized covariance of all model points with the k point squared normalized by the variance at point k.

Initially the covariance **P** is determined from the residual pCO_2 time series (i.e., the variance shown in Fig. 3c). Locations that are not land and have a depth greater than 2000 m are contenders for float deployment. All potential locations are tested to find the location k that captures the greatest percentage of this signal (i.e., has the greatest $\mathbf{VC}_k$). We record this position with the greatest variance constrained, and that location becomes saved as the new float location f. When this step happens, we say that a new float has been deployed. A new residual pCO_2 timeseries, $\mathbf{p}'$, is calculated at every point to represent the unconstrained pCO_2 signal that remains after the field is constrained at f,

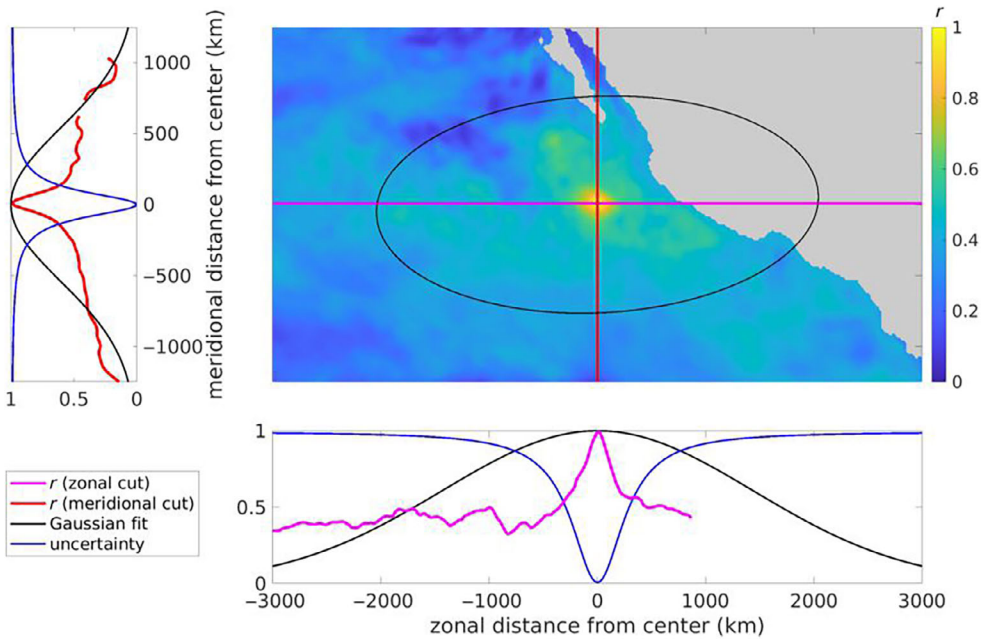

Fig. 6. Non-Gaussian distribution in pCO_2 correlation map. Correlation coefficients between the pCO_2 time series at (18°N, 110°W) and surrounding model points are shown on the upper right. The contour of the Gaussian fit at the e-folding scale is the black ellipse. Bright yellow areas indicate a high correlation while green-blue areas indicate a low correlation coefficient. Gray regions are land. The magenta line in the upper right plot corresponds to a zonal cross-section of the correlation coefficient shown in the bottom right. The red line in the upper right plot corresponds to the meridional cross-section shown in the upper left. The Gaussian fit is annotated in black for the upper left and bottom right plots. The observation uncertainty variance values on the diagonal of **R** are also annotated in blue.

$$\mathbf{p}_l' = \mathbf{p}_l - \gamma \mathbf{p}_f, \tag{6}$$

where l represents model points, and γ is a regression coefficient enabling removal of the pCO_2 covariance with location f. We determine γ from the covariance between each model location l and the float location f divided by the variance of the float location f:

$$\gamma = \frac{\mathbf{p}_f \cdot \mathbf{p}_l}{\mathbf{p}_f \cdot \mathbf{p}_f}. \tag{7}$$

This renders the new residual pCO_2 time series, $\mathbf{p}_l'$, to be orthogonal with the signal correlated with deployed floats. We then move to the next iteration and repeat this procedure to determine the new **VC** for all k locations and determine the next float location f.

Deployment along ship paths

One practical extension of the recursive method further restricts the possible float locations (f) to some prespecified ship path. This constraint takes inspiration from real-world float deployment, where floats are deployed along ship tracks rather than location-by-location. For simplicity, the ship track is represented by a line between two waypoints, although ship tracks are not perfectly straight in reality. The variance constrained (**VC**) is only calculated for proposed deployment locations along the cruise path, but the general methodology remains the same. The location with the highest gain is the chosen float location (f) for each iteration.

Metrics for observing system evaluation

Our motivation is to recommend the ideal locations for deploying floats in the tropical Pacific Ocean. While it is not required for the algorithm to function, the recommendation process is made easier with visualization of the float locations and the remaining model uncertainty (i.e., unconstrained variance) in pCO_2, which is represented by the diagonal of the normalized mapping error (**NME**). The calculation of **NME** for N floats deployed is:

$$\mathbf{NME}_N = 1 - (\mathbf{VC})\mathbf{P}^{-1} \tag{8}$$

Since floats are selected iteratively, N grows with time and **NME** must be recalculated when N changes. However, here the covariance **P** is the initial covariance (i.e., determined from the residual pCO_2 timeseries shown in Fig. 3c) and does not change with N. This is so that we can properly observe how the floats have constrained the full pCO_2 model field, rather than the modified model field which removes the influence of float locations.

NME is the normalized uncertainty of pCO_2 concentration at each model location if our observations were restricted to

that of the chosen float locations, as they would be in the real world. (One caveat is that floats drift from their original deployment location, which is not considered within the algorithm.) **NME** represents how well the float array in each iteration would constrain the pCO_2 signal of the overall tropical Pacific Ocean. In a perfect world where we could deploy M floats for every single model location, **NME** would be 0 everywhere, meaning we can observe the pCO_2 levels at every location with absolute certainty (assuming there is no pCO_2 measurement uncertainty).

Another metric of success in our float deployment is the percent coverage (c_p) which describes how much observational coverage of the model is achieved by the floats deployed. Percent coverage is calculated from the average **NME** and can be thought of as a record of the total information gain from our float array. The equation for percent coverage is given below:

$$c_p = 100 - 100 \times \sum_{l=1}^{M} \frac{\mathbf{NME}_l}{M} \tag{9}$$

where M is the total number of model points.

Correlation lengths

Areas with longer correlation lengths are generally recommended for float deployment since floats there can capture a greater spatial extent of pCO_2 patterns. However, these areas may also have low variance, which creates a trade-off since we are looking to minimize covariance. Ideally, we want long scales and high variance. As a reminder, Fig. 3c shows the residual variance of pCO_2 we hope to resolve with our float deployment strategy.

The semi-major and semi-minor correlation lengths generated by the Gaussian fitting are shown in Fig. 5a,b with 1° longitude and latitude grid spacing. The average angle of the Gaussian fit is 0.18°, so semi-major correlation lengths typically align with the longitudinal (zonal) direction and semi-minor correlation lengths align with the latitudinal (meridional) direction. This pattern is expected since properties affecting pCO_2, such as temperature and biology, tend to vary more in the meridional direction than zonal. Thus, as a general rule, floats should be deployed more closely together in the latitudinal direction rather than longitudinal. The longer zonal lengths are also in agreement with a correlation length analysis of the Southern Ocean (Mazloff et al. 2018). Regions inconsistent with this pattern are found near coastlines and reflect the influence of nearby land masses and meridional currents on pCO_2 variability (Kuhn et al. 2019).

Semi-major correlation lengths range from 1000 to 4500 km, while semi-minor lengths are an order of magnitude shorter, ranging from 100 to 1000 km. The mean semi-major length is 2156 km and the mean semi-minor length is 638 km. These length scales are consistent, to first order, with those found in Yasunaka et al. (2019) (13° ≈ 1300 km semi-major length; 6° ≈ 600 km semi-minor length). The similarity between our model-derived length scales and the observation-based length scales of Yasunaka et al. (2019) justifies our assumption that the model statistics represent the real world. We do note that our calculated semi-major lengths are longer than Yasunaka et al. (2019). Further analysis should be done to identify the cause of this discrepancy.

We also observe that regions with the longest correlation lengths are somewhat inconsistent between the semi-major (Fig. 5a) and semi-minor (Fig. 5b) axes. Regions with the longest semi-major lengths were found near the equator and along 15°S. In contrast, regions with the longest semi-minor lengths were disproportionately located in a region between 10 to 20°S and 100 to 140°W.

Some outliers fell outside the reported ranges of semi-major and semi-minor lengths, but they were typically near or in areas where the Gaussian fitting to correlation maps did not converge to a set of parameters. The points of no convergence comprise 0.0266% of the total area analyzed so they do not significantly change our results, but further investigation is needed to determine why the correlation in some regions is not well described by a single Gaussian. We hypothesize that these regions may be located near boundaries between different dynamical regimes or land, which can cause the correlation map to look non-Gaussian. One example of a coastal location with a non-Gaussian correlation map is located at (18°N, 110°E) and shown in Fig. 6. The initially prescribed uncertainty (blue lines in Fig. 6) captures the small region of high correlation near the central point. The Gaussian fit attempts to capture both the smaller high correlation region and larger scales of medium correlation, indicating multiple dynamical scales as we alluded to in our discussion of the pCO_2 time series (Fig. 2). This fit suggests our prescribed uncertainty does not grow fast enough away from the center to adequately dampen out the remote influence in this region. Future work to adjust the algorithm could involve better accounting for uncertainty (e.g., giving the blue line a steeper fall-off) or enhancing the model by using a sum of multiple Gaussians.

Simulated float deployments

Given a hypothetical budget of 100 floats to deploy across the tropical Pacific Ocean, our covariance algorithm for float deployment found that approximately 67% of the variance can be constrained. The percent coverage corresponding to 25, 50, 75, and 100 floats is listed in Fig. 7 alongside the normalized mapping error corresponding to each number of floats (i.e., iterations of the float algorithm). In the more northern and southern parts of the tropical Pacific, floats are more sparsely deployed, reflecting the longer spatial scales we found from analysis of the semi-major and semi-minor correlation lengths. The equatorial band (approximately 5°S to 5°N) requires a much greater density of floats deployed, especially

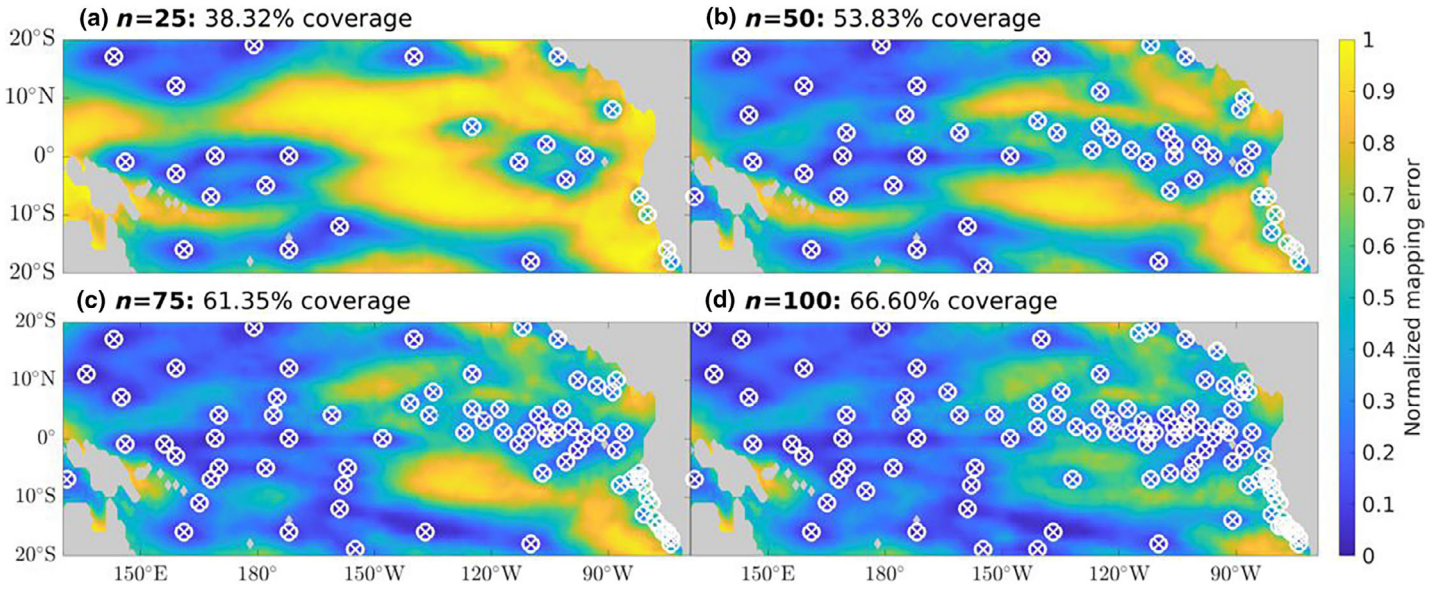

Fig. 7. Mapping error of pCO_2 with 25, 50, 75, and 100 floats deployed. White ⊗s are locations where floats were deployed by the algorithm. Blue areas are well-constrained by the total array of floats deployed at that iteration. Yellow areas are unconstrained or less constrained by the deployed floats. Gray areas are land.

in the eastern equatorial region. The coastline of South America also contains many deployed floats, which may be due to physical processes like upwelling contributing large-scale pCO_2 variability (Lefevre et al. 2002). Notably, there are also patches in the open ocean devoid of floats, such as the areas around (10°N, 140°W) and (10°S, 110°W). However, the overall low mapping error (blue) of Fig. 7 means that the model is well constrained by an array of 100 floats. We find that any additional floats would not contribute much new information (Fig. 8). This is evident from the marginal gain in percent coverage as the number of floats increases above 50.

Recently, the Tropical Pacific Observing System (TPOS) committee has recommended 124 BGC-Argo floats to be deployed between 10°S and 10°N (Kessler et al. 2021). While this recommendation was primarily intended to constrain dissolved oxygen (O_2) in the tropics, these proposed float deployments would likely be effective at constraining pCO_2 as well. Our results assume the ability to select any particular deployment location and that the model statistics are realistic. In addition, our results assume that floats (or deployed platforms) remain stationary at the deployment location, which is untrue in the real world, as discussed in Mazloff et al. (2018) and Chamberlain et al. (2023). Thus, in practice, a recommended deployment of 124 floats is reasonably consistent with our findings.

Simulated float deployment along ship paths

To better accommodate the logistics of float deployment, an extension of our analysis uses a similar float-finding

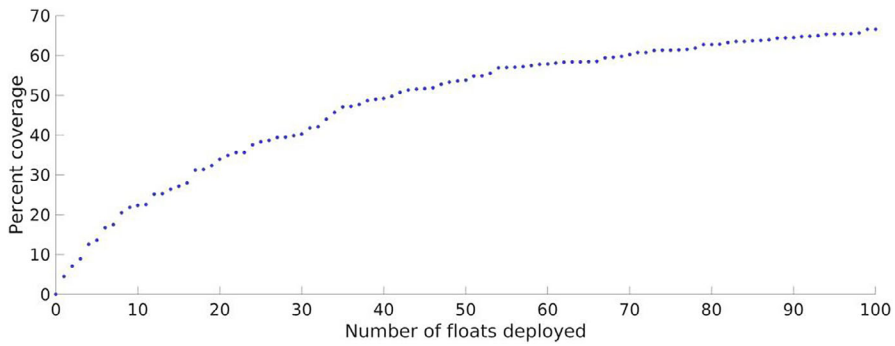

Fig. 8. Percent coverage as a function of number of floats deployed. Each point represents a float deployed in the nth iteration of the algorithm, where n is between 0 and 100. Percent coverage (c_p) indicates the percentage of area in the tropical Pacific Ocean constrained by the floats deployed. Percent coverage asymptotes at approximately 65%, by inspection.

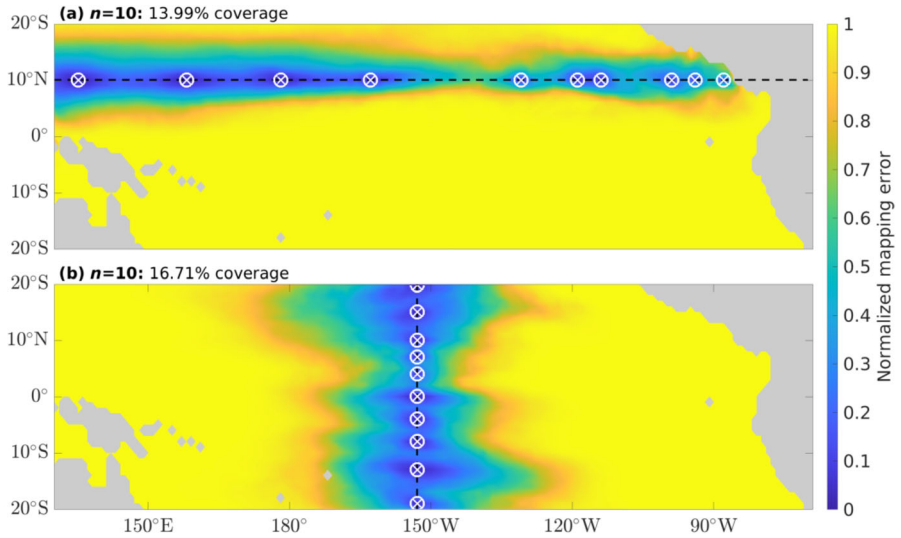

Fig. 9. Mapping error of pCO_2 using 10 floats along ship tracks P4 (top) and P16 (bottom). The track lines themselves are indicated by a black dotted line, while the algorithm-produced float locations are indicated by white ⊗s. Gray areas are land.

technique where potential areas for deployment are limited along a single linear cruise path that can be designated by the user. The results of the float deployment vary by the selection of waypoints for the beginning and end of the ship path. However, unlike the overall float deployment analysis, the "ship-path" method covers a lower percentage of the pCO_2 signal in the entire domain, limited by the possible areas for deployment. The highest percent coverage observed by test paths has been around 25%.

Figure 9 depicts the algorithm's chosen float locations along the Hydrographic Program of the International World Ocean Circulation Experiment (WOCE) track lines P4 and P16, which have waypoints of (10°N, 140°E) to (10°N, 80°W) and (20°S, 153°W) to (20°N, 153°W), respectively. Additional skill from float deployments along both paths appears to asymptote around 10 floats, with an average percent coverage of approximately 15%. The float deployment is nonuniform, showing the utility of this method for not only identifying an appropriate number of floats to deploy but also their optimal location.

Concluding remarks

An observing system design algorithm was developed to optimize locations to measure the ocean in order to maximize information return. The algorithm iteratively selects the statistically-optimal measurement locations and, as shown, can be custom-tailored to account for logistical constraints posed by deployment from ships. A two-step covariance analysis is used to estimate the optimal (1) number of floats to deploy and (2) float deployment locations to help resolve large-scale long-term patterns of ocean property variability.

The method has broad utility and may be applied to any observational platform. Here we applied it to the problem of determining pCO_2 in the tropical Pacific Ocean. We framed the problem in terms of deploying BGC-Argo floats. These floats do not currently directly measure pCO_2, but infer it from statistical relationships with the other measurements they make. We do not address the added uncertainty inherent in this inference as the structure of this uncertainty is unclear. In general, if the pCO_2 uncertainty is larger, we assume that more floats would be necessary to constrain the field since the variance explained by each float would be lower. Nevertheless, this likely would only impact our results in determining percent coverage; the recommended platform distribution reflects model statistics and should be robust.

Our recommendations come with additional caveats beyond sensor limitation. One of these caveats is the underlying assumption that Argo floats are stationary, which we know to be false. In reality, floats can drift up to hundreds of kilometers from the deployed location. For simplicity, we chose not to consider this part of float behavior in our analysis, but a possible extension of this work would be to understand and account for how the changing locations of the floats would affect the observation of pCO_2. Convergence of floats due to wind-driven advection near the ocean surface or eddy dynamics (Wang et al. 2020) could spur fewer float deployments in a region; more floats would be needed in regions where floats diverge. Sevellec et al. (2017) show that modeled float distributions before and after 10 d remain relatively uniform, which could suggest that the statistics used in our algorithm remain robust for short time scales despite the movement of floats. For longer time periods, one potential method to account for float drift would be a transition matrix (Chamberlain et al. 2023). While theoretically straightforward to implement,

this increases the computational cost of the inversion and comes with a new set of additional assumptions to explore. Thus, we have chosen to leave it for future work.

Other assumptions we make are that the ocean statistics are stationary and that the model statistics are realistic. Further studies may be conducted to assess the validity of these assumptions. Finally, as shown in Fig. 6, our assumption that pCO_2 correlations can be represented by a single Gaussian is not true for some locations in the tropical Pacific. While we acknowledge this, the simplicity of the Gaussian function makes it favorable for our algorithm and the challenge of describing as large an area as the Pacific Ocean.

Future work should validate our results by using alternative sources of model data. The sensitivity of this method can also be tested by performing this analysis with the same model at different resolutions or with different initialization parameters.

This research is a small part of the larger combined effort of researchers across the globe to efficiently obtain long-term high-resolution biogeochemical data in hopes of better understanding the ocean's response to a changing climate. The method described here is a step towards improving observational arrays through a statistics-based approach. While we motivate this paper with the goal of observing air–sea carbon dioxide flux, the major strength of this method is its potential to be applied to other biogeochemical variables. It would be interesting to compare the observational array distribution presented here to ones derived for constraining other quantities of interest (e.g., oxygen concentration). The method can also be applied using other model simulations, in other regions, and for design of different observing platforms. As state estimation tools and observational capabilities evolve, this method can be used and evolved in tandem for a more robust and comprehensive measurement system for climate variability.

Data availability statement

The float deployment algorithm and analysis code as well as the model output used here are available at https://github.com/winnie-chu-aos/float-finding-2023.

References

Bakker, D., and others. 2016. A multi-decade record of high-quality fCO_2 data in version 3 of the surface ocean CO_2 atlas (SOCAT). Earth Syst. Sci. Data **8**: 383–413. doi:10.5194/essd-8-383-2016

Biogeochemical-Argo Planning Group. 2016. The scientific rationale, design and implementation plan for a biogeochemical-Argo float array. *Ifremer*. doi:10.13155/46601

Bretherton, F. P., R. E. Davis, and C. Fandry. 1976. A technique for objective analysis and design of oceanographic experiments applied to mode-73. Deep-Sea Res. Oceanogr. Abstr. **23**: 559–582. doi:10.1016/0011-7471(76)90001-2

Bushinsky, S. M., and others. 2019. Reassessing southern ocean air–sea CO_2 flux estimates with the addition of biogeochemical float observations. Global Biogeochem. Cycles **33**: 1370–1388. doi:10.1029/2019GB006176

Campin, J.-M., and others. 2022. Mitgcm/mitgcm: checkpoint68i. *Zenodo*. doi:10.5281/zenodo.6498956

Chamberlain, P., L. D. Talley, M. Mazloff, E. van Sebille, S. T. Gille, T. Tucker, M. Scanderbeg, and P. Robbins. 2023. Using existing Argo trajectories to statistically predict future float positions with a transition matrix. J. Atmos. Ocean. Technol; **40**: 1083–1103. doi:10.1175/JTECH-D-22-0070.1

Chavez, F., P. Strutton, G. Friederich, R. Feely, G. Feldman, D. Foley, and M. McPhaden. 1999. Biological and chemical response of the equatorial pacific ocean to the 1997–98 El Niño. Science **286**: 2126–2131. doi:10.1126/science.286.5447.2126

Dutkiewicz, S., M. J. Follows, P. Heimbach, and J. C. Marshall. 2006. Controls on ocean productivity and air–sea carbon flux: An adjoint model sensitivity study. Geophys. Res. Lett. **33**: 2–5. doi:10.1029/2005GL024987

Fay, A., and G. McKinley. 2013. Global trends in surface ocean pCO_2 from in situ data. Global Biogeochem. Cycles **27**: 541–557. doi:10.1002/gbc.20051

Galbraith, E. D., A. Gnanadesikan, J. P. Dunne, and M. R. Hiscock. 2010. Regional impacts of iron-light colimitation in a global biogeochemical model. Biogeosciences **7**: 1043–1064. doi:10.5194/bg-7-1043-2010

Garcia, H., and others. 2019*a*. World ocean atlas 2018. Vol. 4: Dissolved inorganic nutrients (phosphate, nitrate and nitrate + nitrite, silicate). NOAA Atlas NESDIS 84. https://www.ncei.noaa.gov/sites/default/files/2022-06/woa18_vol4.pdf

Garcia, H., and others. 2019*b*. World ocean atlas 2018. Vol. 3: Dissolved oxygen, apparent oxygen utilization, and dissolved oxygen saturation. NOAA Atlas NESDIS 84. https://www.ncei.noaa.gov/sites/default/files/2022-06/woa18_vol3.pdf

Gray, A. R., and others. 2018. Autonomous biogeochemical floats detect significant carbon dioxide outgassing in the high-latitude southern ocean. Geophys. Res. Lett. **45**: 9049–9057. doi:10.1029/2018GL078013

Gruber, N., D. C. Bakker, T. DeVries, L. Gregor, J. Hauck, P. Landschützer, G. A. McKinley, and J. D. Müller. 2023. Trends and variability in the ocean carbon sink. Nat. Rev. Earth Environ. **4**: 119–134. doi:10.1038/s43017-022-00381-x

Hersbach, H., and others. 2018. ERA5 hourly data on single levels from 1979 to present [Data set]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi:10.24381/cds.adbb2d47

Keeling, C. D., S. C. Piper, R. B. Bacastow, M. Wahlen, T. P. Whorf, M. Heimann, and H. A. Meijer. 2001. Exchanges of atmospheric CO_2 and $13CO_2$ with the terrestrial biosphere and oceans from 1978 to 2000. I. Global aspects. doi:10.1007/0-387-27048-5_5

Kessler, W., S. Cravatte, and L. Authors. 2021. Final report of TPOS 2020 August 2021. TPOS 2020 GOOS-268. p. 83

Available from https://tropicalpacific.org/tpos2020-project-archive/reports/

Key, R. M., and others. 2015. Global ocean data analysis project, version 2 (glodapv2). Ornl/Cdiac-162, Ndp-093. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Dept. of Energy. doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2

Kuhn, A., S. Dutkiewicz, O. Jahn, S. Clayton, T. Rynearson, M. Mazloff, and A. Barton. 2019. Temporal and spatial scales of correlation in marine phytoplankton communities. J. Geophys. Res. Oceans **124**: 9417–9438. doi:10.1029/2019JC015331

Landschützer, P., N. Gruber, D. C. Bakker, and U. Schuster. 2014. Recent variability of the global ocean carbon sink. Global Biogeochem. Cycles **28**: 927–949. doi:10.1002/2014GB004853

Landschützer, P., G. G. Laruelle, A. Roobaert, and P. Regnier. 2020. A uniform pCO_2 climatology combining open and coastal oceans. Earth Syst. Sci. Data **12**: 2537–2553. doi:10.5194/essd-12-2537-2020

Large, W. G., J. C. McWilliams, and S. C. Doney. 1994. Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. Rev. Geophys. **32**: 363–403. doi:10.1029/94RG01872

Lauvset, S. K., and others. 2016. A new global interior ocean mapped climatology: The 1×1 glodap version 2. Earth Syst. Sci. Data **8**: 325–340. doi:10.5194/essd-8-325-2016

Lefevre, N., J. Aiken, J. Rutllant, G. Daneri, S. Lavender, and T. Smyth. 2002. Observations of pCO_2 in the coastal upwelling off Chile: Spatial and temporal extrapolation using satellite data. J. Geophys. Res. Oceans **107**: 8-1–8-15. doi:10.1029/2000JC000395

Leith, C. 1996. Stochastic models of chaotic systems. Phys. D Nonlinear Phenom. **98**: 481–491. doi:10.1016/0167-2789(96)00107-8

Mazloff, M., B. Cornuelle, S. Gille, and A. Verdy. 2018. Correlation lengths for estimating the large-scale carbon and heat content of the southern ocean. J. Geophys. Res. Oceans **123**: 883–901. doi:10.1002/2017JC013408

Sevellec, F., A. Colin de Verdiére, and M. Ollitrault. 2017. Evolution of intermediate water masses based on Argo float displacements. J. Phys. Oceanogr. **47**: 1569–1586. doi:10.1175/JPO-D-16-0182.1

Smith, N., and others. 2019. Tropical pacific observing system. Front. Mar. Sci. **6**: 31. doi:10.3389/fmars.2019.00031

Sutton, A. J., C. L. Sabine, S. Musielewicz, S. Maenner Jones, C. Dietrich, R. Bott, and J. Osborne. 2003. High-resolution ocean and atmosphere pCO_2 time-series measurements from mooring WA_125W_47N in the North Pacific Ocean (NCEI Accession 0115322) [Data set]. NOAA National Centers for Environmental Information. doi:10.3334/cdiac/otg.tsm_wa_125w_47n

Tardif, R., G. J. Hakim, K. A. Bumbaco, M. A. Lazzara, K. W. Manning, D. E. Mikolajczyk, and J. G. Powers. 2022. Assessing observation network design predictions for monitoring Antarctic surface temperature. Q. J. R. Meteorol. Soc. **148**: 727–746. doi:10.1002/qj.4226

Verdy, A., B. Cornuelle, M. R. Mazloff, and D. L. Rudnick. 2017. Estimation of the tropical pacific ocean state 2010–13. J. Atmos. Oceanic Tech. **34**: 1501–1517. doi:10.1175/JTECH-D-16-0223.1

Verdy, A., and M. R. Mazloff. 2017. A data assimilating model for estimating southern ocean biogeochemistry. J. Geophys. Res. Oceans **122**: 6968–6988. doi:10.1002/2016JC012650

Wang, T., S. T. Gille, M. R. Mazloff, N. V. Zilberman, and Y. Du. 2020. Eddy-induced acceleration of Argo floats. J. Geophys. Res. Oceans **125**: e2019JC016042. doi:10.1029/2019JC016042

Wanninkhof, R. 1992. Relationship between wind speed and gas exchange. J. Geophys. Res. **97**: 7373–7382. doi:10.1029/92JC00188

Williams, N. L., and others. 2017. Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. Global Biogeochem. Cycles **31**: 591–604. doi:10.1002/2016GB005541

Wunsch, C. 1996. The ocean circulation inverse problem. Cambridge Univ. Press. doi:10.1017/cbo9780511629570

Yasunaka, S., S. Kouketsu, P. Strutton, A. Sutton, A. Murata, S. Nakaoka, and Y. Nojiri. 2019. Spatio-temporal variability of surface water pCO_2 and nutrients in the tropical pacific from 1981 to 2015. Deep-Sea Res. II Top. Stud. Oceanogr. **169**: 104680. doi:10.1016/j.dsr2.2019.104680

Acknowledgments

We thank our anonymous reviewer and the associate editor for their helpful comments, which greatly improved the paper. This work is supported by NOPP grant NA19OAR4310361, by NOAA grants NA21OAR4310257 and NA22OAR4310597, and NSF grants OPP-1936222 and OPP-2149501. This work was also supported by the 2020 Scripps Undergraduate Research Fellowship (SURF) at Scripps Institution of Oceanography, funded by the National Science Foundation, Research Experiences for Undergraduates GEO/OCE Award #1659793. A special thanks extends to Dr. Jane Teranes and her dedication to the SURF program. We gratefully acknowledge use of data made available by the TAO Project Office of NOAA/PMEL (http://www.pmel.noaa.gov/tao). The results contain modified Copernicus Climate Change Service information. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

Conflict of Interest

The authors declare they have no competing interests.

Submitted 16 December 2023
Revised 23 June 2024
Accepted 30 July 2024

Associate editor: Mike DeGrandpre