Numerical Simulations to Project Argo Float Positions in the Middepth and Deep Southwest Pacific

TIANYU WANG

*State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences,
Guangzhou, and University of Chinese Academy of Sciences, Beijing, China, and Scripps Institution of Oceanography,
University of California, San Diego, La Jolla, California*

SARAH T. GILLE, MATTHEW R. MAZLOFF, AND NATHALIE V. ZILBERMAN

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California

YAN DU

*State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences,
Guangzhou, and University of Chinese Academy of Sciences, Beijing, China*

ABSTRACT

Argo float trajectories are simulated in the southwest Pacific basin (25°–45°S, 170°E–165°W) using velocity fields from a 1/12° Southern Ocean model and a Lagrangian particle tracking model programmed to represent the vertical motions of profiling Argo floats. The system is applied to simulate both core Argo floats (typically parked at 1000-m depth and profiling to 2000-m depth) and Deep Argo floats (parked 500 m above the seafloor). The goal is to estimate probability density functions (PDFs) predicting future float positions. Differences are expected in the trajectory statistics, largely because of limitations in the temporal and spatial resolution of the model fields and uncertainties associated with a random walk component included in the particle advection scheme to represent this unresolved variability. Nonetheless, the core Argo float displacements over ~100-day time intervals are mostly consistent with the derived PDFs, particularly in regions with stable midlayer flows. For the Deep Argo floats, which are released into the open ocean and parked near the bottom, the simulations predict an average total displacement of less than 50 km within 100 days, in good agreement with the Deep Argo floats deployed as part of a pilot study. The study explores both the representativeness and the predictability of float displacements, with an aim to contribute to planning for the float observing system.

1. Introduction

The global implementation of the core Argo array has transformed oceanography. At present, Argo operates more than 3500 floats. Every 10 days, each float transmits a profile of temperature, salinity, and pressure, typically extending over a 2000-m depth range (Roemmich et al. 1998, 2009; Roemmich and Gilson 2009). In between profiling, the floats drift at a "parking depth" 1000 m below the ocean surface. The displacements between one profile and the next provide a measure of ocean currents at 1000-m depth and have been used for a number of analyses (e.g., Davis and

Zenk 2001; Park et al. 2005; Chu et al. 2007; Lebedev et al. 2007; Ollitrault and Rannou 2013; Park and Kim 2013; Gray and Riser 2014).

There are two related challenges associated with understanding float displacements at the parking depth. First, since the ocean is vast and eddying and varies on all temporal and spatial scales, how representative are the trajectories of a finite number of floats? Second, since floats move with the currents, how well can we project where they will sample, and how do we optimize a float deployment (and reprogramming) scheme to best sample the ocean currents of a particular region?

Deep Argo will extend core Argo profiles to the seafloor. Planners for Deep Argo envision an array of 1228 deep floats at 5° latitude × 5° longitude spacing, profiling

---

Corresponding author: Sarah T. Gille, sgille@ucsd.edu

temperature, salinity, and pressure over the full water column every 15 days (Johnson and Lyman 2014). The Deep Argo program is currently in the early stages of its implementation phase.

The objective of this study is to make use of a numerical model to explore both the representativeness and the predictability of float displacements. We use a 1/12°-resolution hydrostatic version of the MIT General Circulation Model (MITgcm; http://mitgcm.org; Marshall et al. 1997), configured for the Southern Ocean. We simulate the behavior of Argo floats using a Lagrangian float-tracking model with the model output fields. By deploying a large ensemble of floats within the model, we are able to construct probability density functions (PDFs) of simulated-float trajectory end points. We adopt a two-pronged approach. First, we evaluate the performance of the simulated floats by comparing them with core Argo trajectories, which are numerous enough to allow a statistical ensemble approach. Second, we apply the methodology to individual Deep Argo floats, to test the possibility of using model trajectories to evaluate deployment plans for single floats.

The methodology presented could be applied for any time scale, and for this work we consider a seasonal scale of 100 days. On 100-day time scales, float operators often need to consider the probability that a float will enter an ice-covered region, a shallow region, or a region with scientific, societal, or political importance. With this knowledge changes to the float's programming can be made. For example, a float that is about to enter a sea ice–covered zone could have its sampling frequency reprogrammed to conserve sufficient battery life to report again once it emerges from the ice. Predictions of 100 days are obviously short for informing long-term deployment planning. However, one may make the assumption that the ocean statistics are linear and interpret the derived PDFs as being analogous to transition matrices, $\mathbf{T}$, that can be combined to yield longer predictions, as McAdam and van Sebille (2018) have shown for surface drifter trajectories. For example, a prediction for 1100 days (i.e., about 3 years) could be obtained by projecting the release point into the transition matrix 11 times (i.e., $\mathbf{T}^{11}$).

The paper is organized as follows: Section 2 introduces the Argo simulation system, including the 1/12°-resolution Southern Ocean simulation, the Lagrangian particle tracking model, and a subgrid-scale parameterization process for the simulated Argo floats. Section 3 presents results from the core Argo simulation system, with a focus on the PDFs of float displacements. Within our test region northeast of New Zealand, we will show that the simulated-float PDFs are, to leading order,

consistent with the trajectories of core Argo floats. Section 4 presents the simulations for Deep Argo floats, as well as the comparison between measured Deep Argo float displacements and model predictions. Finally, section 5 summarizes the results and assesses shortcomings of Argo simulations in regions with strong background velocities or with numerous eddies. We also identify priorities for further improvements.

2. The Argo simulation system

a. The 1/12° Southern Ocean simulation

We focus on the southwest Pacific basin (25°–45°S, 170°E–165°W; see Fig. 1), where the first two Deep Sounding Oceanographic Lagrangian Observer (SOLO) floats were deployed (Zilberman and Maze 2015). This region of the Southern Ocean is a gateway for cold, deep Antarctic Bottom Water (AABW) to enter the Pacific basin via the Pacific deep western boundary current (DWBC; e.g., Warren 1981; McCave and Carter 1997; Whitworth et al. 1999; Zilberman et al. 2013; Mazloff and Boening 2016).

To simulate mesoscale motions of Argo floats within the region, we use the MITgcm configured for the Southern Ocean at 1/12° horizontal resolution. We analyze the year 2009, with initial conditions spun up from the Southern Ocean State Estimate (SOSE; Mazloff et al. 2010) and atmospheric state prescribed from ERA-Interim (Dee et al. 2011). The model has 104 vertical levels, unevenly distributed between the surface and 6000 m. It uses a Mercator projection poleward of 30°S and telescopes in resolution from 30°S to the equator. The equatorial open boundary condition is prescribed from the Hybrid Coordinate Ocean Model (HYCOM; Chassignet et al. 2007). As shown in Fig. 1, the strongest currents at 1000 m roughly follow the continental shelf, a consequence of potential vorticity constraints (e.g., LaCasce 2000; Gille 2003). This topographic steering is well captured in the 1/12° Southern Ocean simulation.

The comparison of the model-simulated sea surface height (SSH) field to the AVISO-mapped satellite altimetry product (Fig. 2, left panels) gives a qualitative assessment of the surface circulation and its variability. The simulation period of 21 November 2008–31 December 2009 is relatively short, such that the mean of both products over this period shows significant mesoscale structure. Both also show a large-scale rise in SSH of about 0.4 m from the south to the north. A strong meridional gradient associated with zonal currents is apparent south of about 34°S. However, no notably strong zonal flow is shown in the 1000-m currents
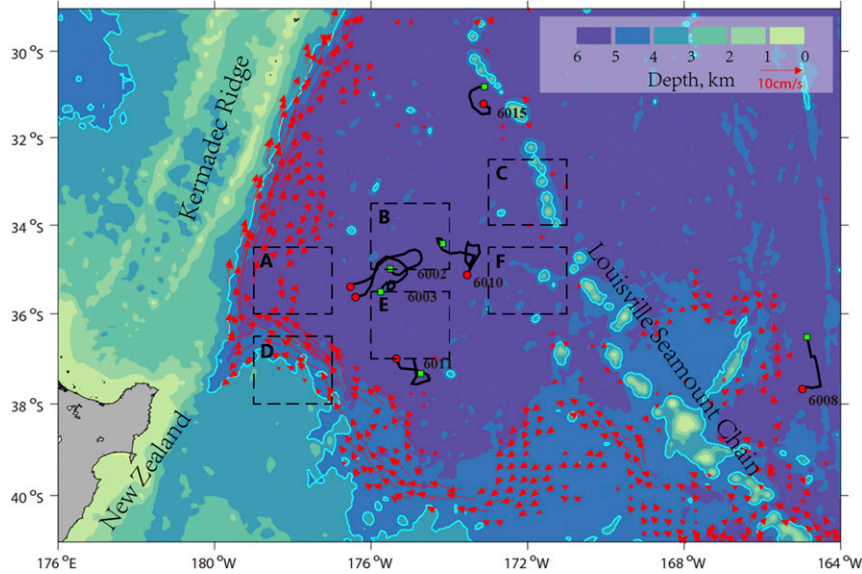
FIG. 1. Bottom topography of the southwest Pacific basin from the 1/12°-resolution model: the 4000-m isobaths (light-blue solid lines); the trajectories of the two Deep SOLO prototypes 6002 and 6003 from June 2014 to October 2014 (black lines); and the Deep SOLO floats 6008, 6010, 6011, and 6015 from August 2016 to February 2017 (details see Table 1). We use ∼100-day segments from these floats, with start positions indicated by red circles and end positions by green rectangles. Red vectors indicate annually averaged velocity at 1000-m depth (shown only when the mean speed exceeds 2 cm s^{-1}). The subdomains are marked (dashed lines): A (34.5°–36°S, 179°–177°W), B (33.5°–35°S, 176°–174°W), C (32.5°–34°S, 173°–171°W), D (36.5°–38°S, 179°–177°W), E (35.5°–37°S, 176°–174°W), and F (34.5°–36°S, 173°–171°W).

(Fig. 1), implying that this flow is significantly surface intensified.

The standard deviation of the two SSH products shows a similar structure (Fig. 2, right panels). High variance occurs off New Zealand, extends northward, and then reaches eastward. Like the mean, the structure is patchy as a result of the relative brevity of the time series. Despite variability differences at specific locations, both the model and the AVISO product have roughly equal levels of SSH variability.

b. Lagrangian particle tracking model

Argo trajectories are simulated using an Argo module of the offline particle tracking model Octopus (http://github.com/jinbow/Octopus/; see e.g., Tamsitt et al. 2017; van Sebille et al. 2018). The model integrates Argo trajectories using daily average horizontal velocities from the 1/12°-resolution model and vertical velocities derived from real Argo floats. Float trajectories are integrated via a Lagrangian advection scheme: $\partial \mathbf{X}_i / \partial t = \mathbf{V}_i$, where $\mathbf{X}_i$ is the (vector) position of the ith particle and $\mathbf{V}_i$ is the velocity vector mapped to the numerical float position using three-dimensional linear interpolation.

In addition to the model velocities, Octopus also includes a stochastic velocity component to account

for the float displacement caused by processes (such as turbulence) that are unresolved by the 1/12° model. The subgrid-scale processes are parameterized by $\Delta \mathbf{X}_i = (2K\delta t)^{-2}\omega(t)$, where K represents the diffusivity tensor $(\kappa_x, \kappa_y, \kappa_z)$, ω represents a Wiener process (stochastic random walk displacements) with unit variance, and δt denotes the time step for the Lagrangian integration (e.g., van Sebille et al. 2018). The particle trajectory in discrete form becomes

$$\mathbf{X}_i^{n+1} = \mathbf{X}_i^n + \mathbf{V}_i \delta t + \sqrt{2K\delta t}\,\omega(t),$$

where n represents the time step number. In practice, the random number generator should be carefully chosen, because not all numbers are suitable for use in random walk models (Hunter et al. 1993). Here we implement the normal random number generator algorithm described by Kinderman and Monahan (1977), with a horizontal diffusivity of $1 \, m^2 \, s^{-1}$, which is consistent with the diffusivities of the 1/12°-resolution Southern Ocean simulation. An assessment of 100-day trajectory autocorrelations with varying diffusivity found the trajectories to be rather insensitive to the parameterization for values less than or equal
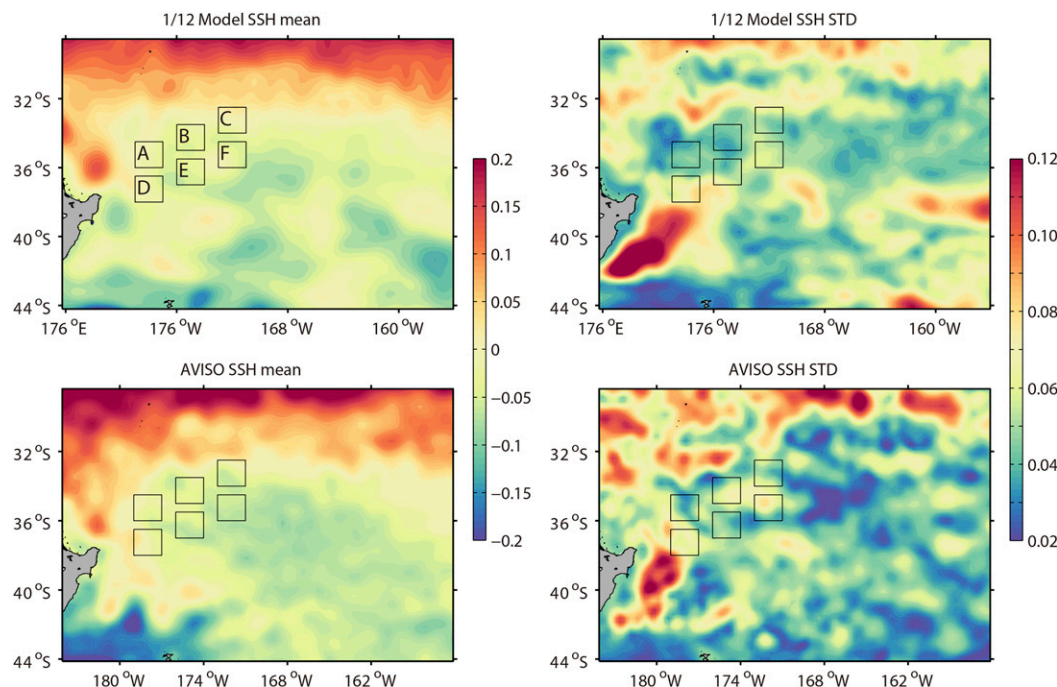
FIG. 2. The 21 Nov 2008–31 Dec 2009 mean SSH (m) from (top left) the 1/12° model and (bottom left) the AVISO-mapped product. The standard deviation of SSH from (top right) the 1/12° model and (bottom right) the AVISO-mapped product.

to $1 \, \mathrm{m^2 \, s^{-1}}$. For values of 10, 100, and $1000 \, \mathrm{m^2 \, s^{-1}}$, the autocorrelations for the trajectories dropped to approximately 0.9, 0.6, and 0.35, respectively (not shown). More details of the Lagrangian particle tracking model are available from the Octopus website, and the rationale for the random scheme is discussed in the next sections.

c. Simulated Argo parameters

1) THE MEASUREMENT CYCLE

Our objective is to simulate float behavior. The 100-day time scale of our simulations allows for significant float displacement while remaining short enough to permit a large number of comparisons between simulated and observed in situ trajectories.

Simulated core Argo floats are configured to follow the same cycle as the real floats, as defined in Fig. 3a, including an initial descent to 1000 m, parking time at 1000 m, a further descent to 2000 m, then ascent from 2000-m depth to the surface, and finally an interval of time at the ocean surface to transmit data via satellite. For the period from 2004 to 2015, cycle timings for the delayed-mode Argo data were obtained from the global data centers (ftp://usgodae.org/pub/outgoing/argo; ftp://ftp.ifremer.fr/ifremer/argo) within the southwest Pacific basin. Ocean currents are nonuniform in the vertical,

which affects the true trajectories and how we interpret Argo displacements.

The position of a float can be obtained only when the float is at the ocean surface and communicating with satellites. Core Argo floats communicate in one of two ways. Most of the Argo floats (356 of 411 in our domain) use the Argos satellite system. The remaining floats (55) use the Iridium (or Orbcomm) satellite system for data transmission and GPS for positioning. The main difference between the two systems is in their transmission strategy: Argos floats repeatedly emit all of their messages over a sufficiently long time period to enable reliable acquisition by a small number of orbiting satellites. Therefore, they spend an average of 13.4 h at the surface during each cycle and 228.9 h below the surface. In contrast, messages sent by Iridium floats can be received from anywhere on the planet by one of the 66 satellites in the Iridium constellation. Thus, the Argo floats using Iridium stay at the surface less than half an hour. Most Iridium floats are configured to profile every 10 days, with total cycle times comparable to Argos floats, though some of the Iridium floats in the southwest Pacific operated on a shorter total cycle time.

In the simulations we evaluate both types of floats. The Argo clock schedules are configured to approximate an "average" float, with cycle times and parking depths representing the weighted average of cycle times
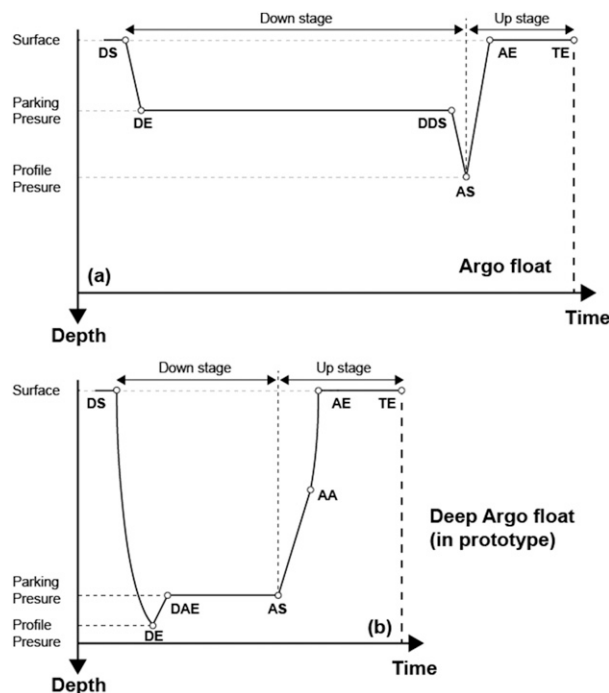
FIG. 3. Schematics of one float cycle for (a) core Argo floats and (b) Deep Argo floats during the test phase analyzed for this study. A core Argo cycle includes descent, parking, deep profiling, and surface telemetry. The schematic indicates the times when descent starts (DS), descent ends (DE), deep descent starts (DDS), ascent starts (AS), ascent ends (AE), and transmission ends (TE). The definitions are adopted from Ollitrault and Rannou (2013). A Deep Argo cycle (here using the Deep SOLO configuration) contains similar phases but with a slight change in the sequence, since the float first descends from DS to the ocean bottom (DE) and then rises to a parking pressure (DAE) for several days before rising to the surface, starting from AS. Ascent accelerates at approximately 2000-m depth (AA) to minimize drift in the upper 2000 m, and the float reaches the surface at AE.

and parking depths for the 411 floats in the region: the average total cycle duration is 242.7 h, including 208.7 h of parking time, 21.8 h ascending and descending, and 12.2 h at the surface. For the simulations we identify final float positions after exactly 100 days, even if the simulated float has not reached the surface. Simulated Deep Argo floats are configured to mimic Deep SOLO floats built at Scripps Institution of Oceanography and deployed in the southwest Pacific Ocean. This includes two Deep Argo floats deployed northeast of New Zealand in June 2014 and four standard Deep Argo floats deployed in August 2016. We use approximately 100-day trajectories from these floats, as shown in Fig. 1. Floats descend from the surface at a linear damped rate (from $0.2 \, \mathrm{m \, s^{-1}}$ at the surface to about $0.03 \, \mathrm{m \, s^{-1}}$ near the bottom, as shown in Fig. 4a). Floats ascend in two stages: first from the deepest profile

pressure (near the seafloor) to the parking depth (500 m above the seafloor) and later, at the end of the cycle, from the parking depth to the surface (Fig. 3b). The ascent rate is $0.055 \, \mathrm{m \, s^{-1}}$, nearly constant below 2000-m depth, and it increases linearly above 2000 m (Fig. 4b). The descent and ascent times depend on the ocean depth. For the two Deep Argo floats deployed in 2014, in order to facilitate a rapid assessment of float and CTD sensor performance, short cycle times were used (i.e., for a 5000-m profile, the cycle duration of float is 4.5 days long, including a 2-day parking period). The floats deployed in 2016 had longer cycle times (typically 10 days in total with an 8.3-day parking phase). Each deep float is simulated separately, so the cycle time, parking time, and surface time for each float are individually configured (see Table 1).

2) QUANTITY OF SIMULATED ARGO FLOATS

To maximize the statistical samples using a model of limited duration, we carry out repeated simulated-float releases, separated by time intervals large enough to produce statistically independent trajectories. Figure 5 shows temporal autocorrelations for both midlayer (1000 m) and deep-layer (5000 m) velocities. To obtain a time scale, we fit the autocorrelations to a Gaussian function: $R(x) = \exp(-x^2/2a^2)$. The decorrelation time scales are about 5 days for midlayer velocities (corresponding to core Argo floats) and 20 days for deep-layer velocities. In light of the decorrelation time scales, simulated core floats are released at 10-day intervals starting on day 1, and with the last release on day 271. Thus, there are 28 independent release times during the ~380 days of the model run. Deep Argo floats have longer decorrelation times, and we release at 40-day intervals, with the last at day 281, yielding eight independent release times for deep Argo.

In general, increasing the number of simulated floats should lead to more robust statistics (although results could also be sensitive to interannual variability, which is not included in our 380-day model output time period). We carried out an experiment in which we released a varying number of numerical floats in order to determine how many floats are needed to obtain well-converged, stable float statistics. As a convergence test, we quantify the number of bins on a $1/12° \times 1/12°$ grid containing particles after 100 model days. This quantity asymptotes once a sufficient number of particles are released. In the southwest Pacific basin, results in Fig. 6 show that 10000–15000 particles are sufficient to yield robust statistics for the distribution of possible final float positions. We obtained similar results for both core Argo (red line in Fig. 6) and Deep Argo floats (blue line in Fig. 6).
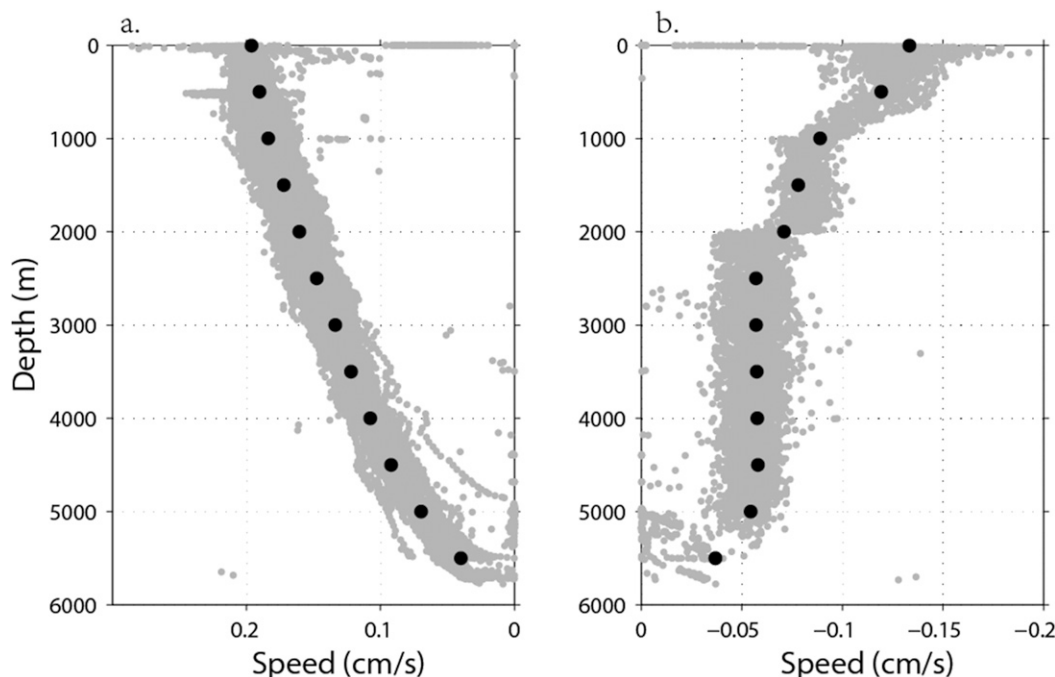
FIG. 4. Rates of (a) descent and (b) ascent as a function of depth derived from Deep Argo float 6002. The gray dots are the original data, and the black dots indicate the averaged values, which we used in Octopus.

The statistics shown in Fig. 6 are applicable for both core and Deep Argo investigations, although the two depth ranges require different release strategies. The simulated core Argo floats were released on a 40 × 25 grid in the analysis domain, making them readily comparable to the large ensemble of core Argo floats in the ocean. In contrast, for Deep Argo, at each location of true float releases we simulate a cluster of 1000 floats at 10-day intervals or 3000 floats at 40-day intervals. Because the Deep Argo floats studied here park near the bottom, they travel more slowly than core Argo floats and spread across fewer bins. The 28 start dates at 10-day intervals resulted in 28 000 (100 day) float trajectories, which should be sufficient to yield converged statistics. For the 40-day intervals, the eight start dates imply 24 000 (100 day) float trajectories.

3. Core Argo simulation

a. Regional circulation and subdomains

As described in section 2, in the southwest Pacific basin, AABW is transported in a DWBC (e.g., Warren 1981; McCave and Carter 1997; Whitworth et al. 1999; Zilberman et al. 2013). The complex western boundary current system makes the study region more energetic to the west. Six geographic subdomains are identified in order to examine the sensitivity of the results to specific geographic constraints from topography, the mean flow, or other processes. Each of these subdomains has to be large enough to capture an adequate number of true float trajectories for statistical comparison but spatially restricted enough to represent a single-flow regime. With this in mind we chose the domains shown in Fig. 1:

TABLE 1. Deep Argo float information. As discussed in the text, date ranges are selected to be as close to 100 days as possible, given actual float surfacing times. We do not make use of the initial testing phase right after deployment, when floats were operated with more rapid cycle times.

| Float No.	Dates	Lon (°W)	Lat (°S)	Cycle hours	Parking hours	Surface hours
6002 (prototype)	10 Jul 2014–19 Oct 2014	176.5–175.5	35.4–35.0	86	48	0.25
6003 (prototype)	11 Jul 2014–14 Oct 2014	176.1–175.7	35.5–35.4	86	48	0.25
6008	26 Aug 2016–4 Dec 2016	164.9–164.8	37.7–36.9	240	200	0.25
6010	26 Oct 2016–3 Feb 2017	173.5–174.2	35.1–34.4	240	200	0.25
6011	8 Jul 2016–16 Oct 2016	175.4–174.8	37.0–37.2	240	200	0.25
6015	4 Sep 2016–13 Dec 2016	173.1–173.1	31.2–30.8	240	200	0.25
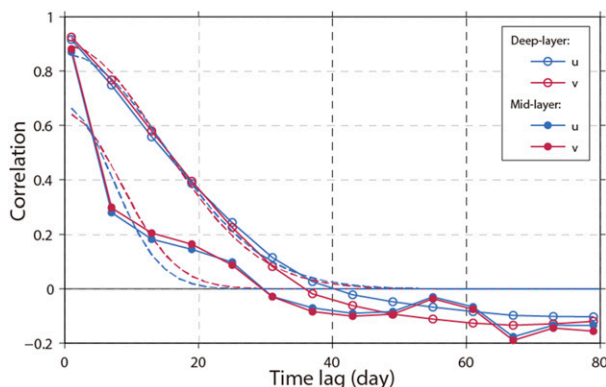
FIG. 5. Velocity autocorrelation functions computed from regional averaged velocities over the southwest Pacific basin (25°–45°S, 170°E–165°W; see Fig. 1) in the 1/12° Southern Ocean simulation. Shown are autocorrelations for zonal components (*u*) (red lines with dots) and meridional components (*v*) (blue lines), and Gaussian fits to the correlation functions (thin dashed lines).
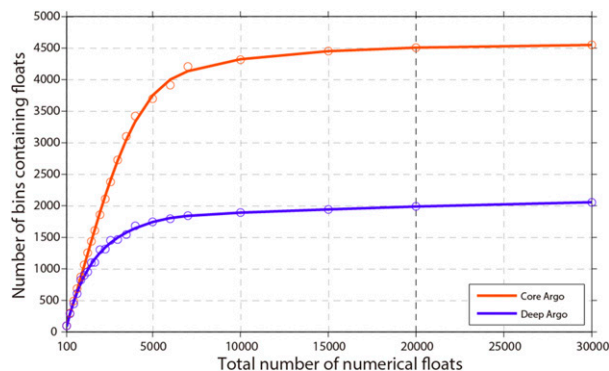

FIG. 6. A convergence test as a function of the number of simulated floats for core Argo (red) and Deep Argo (blue). The *y* axis indicates the number of $1/12° \times 1/12°$-resolution bins containing simulated floats at the end of 100 days. Each value is an average from five float release experiments conducted at a fixed location (36°S, 177°W) within the southwest Pacific basin. The circles represent results from the original numerical calculations, and the lines indicate polynomial fits.

A (34.5°–36°S, 179°–177°W), B (33.5°–35°S, 176°–174°W), C (32.5°–34°S, 173°–171°W), D (36.5°–38°S, 179°–177°W), E (35.5°–37°S, 176°–174°W), and F (34.5°–36°S, 173°–171°W). Subdomains A and D are dominated by topographic steering and have the highest kinetic energy at depth, C and F have the lowest kinetic energy and flow structure, and B and E are in moderately energetic flow conditions. However, as can be seen in Fig. 2, all regions have similar levels of sea surface height variability, implying that eddies are nonnegligible in all domains.

Figure 7 shows an example of the float trajectories released in subdomain E. To make the trajectories easier to distinguish, in this demonstration of the method we advect using only the time mean flow and use only 1000 simulated floats. These are initially released on a 40×25 grid within the subdomain (black dots in the square). After 100 days, most floats have drifted significantly from their release point. The overall spread of the float array is a function of the domain size and the flow structure. We quantify the horizontal spreading of the floats using a discretized two-dimensional probability density function (2D-PDF) denoted as $P(x, y, T, \Delta L)$, where x and y indicate geographic coordinates, T represents the elapsed time, and ΔL is the grid size. The 2D-PDF is computed by counting the number of floats within a grid box centered at (x, y) and normalizing by the total number of floats and grid box area $(\Delta L)^2$ after spreading time T. Thus, for example, $P(x_0, y_0, 100 \text{ days}, 1\text{km}) = 0.6\% \text{ km}^{-2}$ means that 6 out of 1000 floats can be found within a $1 \text{ km} \times 1 \text{ km}$ grid box centered at (x_0, y_0) after 100 days.

b. Comparison between simulations and real floats

The 2D-PDFs are shown in Fig. 8. To identify real Argo float trajectories, we first find every Argo float that has ever entered a specific subdomain. Using that information we extract 100-day trajectories that originate within the subdomain. (Since real floats do not surface precisely every 10 days, we select the total number of cycles that is as close to 100 days as possible.) To expand the sample size, we break each Argo trajectory into multiple 100-day segments. We allow segments to overlap but consider them to be independent only if the starting position of each segment is separated by more than 9 km (or 1/12°) in space and 60 days in time and still originates within the subdomain. This provides 19 trajectories for subdomain A, 19 for B, 18 for C, 29 for D, 42 for E, and 15 for F. The positions of the floats at the end of ~100 days are shown as black dots in Fig. 8.

Visual examination indicates that subdomains A and D display less agreement between observations and simulations, while subdomains B, C, E, and F show reasonable agreement, with the exception of a few outliers. There are a number of possible reasons for the poorer level of agreement in subdomains A and D. For example, differences in eddy energy in the model relative to the real ocean could lead to trajectory differences, as would discrepancies in the mean velocities or errors in the model circulation. Inadequate float numbers within a specific subdomain could also lead to poor statistical convergence.

To develop a quantitative comparison of the simulated and real floats, we use one-dimensional cumulative distribution functions (CDFs) as a function of longitude
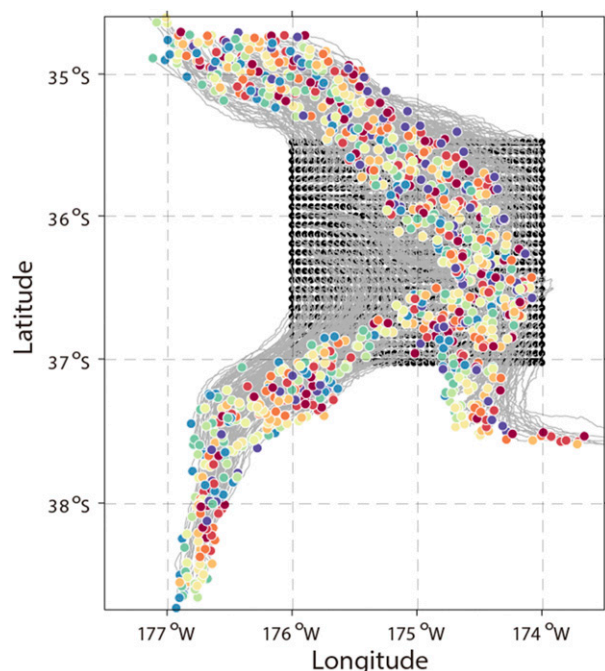
FIG. 7. Lagrangian motions of 1000 simulated core Argo floats in 100-day model simulations. Simulated Argo floats are advected by the mean current within subdomain E (shown in Fig. 1). Release locations are distributed on a 40×25 matrix, marked by black dots. The positions at the end of 100 days are marked by the randomly colored dots, and the trajectories are indicated by gray lines.

and latitude. Direct comparison of the CDFs (see Fig. 9) shows agreement between simulations and observations in subdomains B–E. As explained in the appendix, Kolmogorov–Smirnov (KS) statistics provide a formal metric to evaluate similarities between datasets based on the maximum separation between their CDFs in the zonal (D_{lon}) and meridional (D_{lat}) directions (Fig. 9). In five of the six subdomains (all except subdomain A), we find that at the 95% significance level the actual float trajectories are consistent with our derived CDFs.

We have far fewer Deep Argo floats than core Argo, so our assessment of Deep Argo performance will need to consider individual float trajectories rather than an ensemble of observed ones. In analogy with the Deep Argo approach (presented in the next section), and to highlight the constrasts between single-float and ensemble approaches, we consider six core Argo floats from subdomains D and E, selected because both domains passed the KS test. Table 2 provides detailed information about the six floats. At the start position of each of the real floats, a total of 28 000 simulated floats (1000 at each of the 28 start dates) are released in a circular area with radius $1/12°$ (i.e., ~9 km). The PDFs obtained from these point releases are expected to differ

from those obtained with an ensemble of subdomain releases. Indeed, the PDFs for the ensemble presented in Fig. 8 are smooth and continuous. In contrast, the PDFs for the point releases are more fragmented with multiple noncontiguous regions of high probability (Fig. 10). Comparing the PDFs derived by the two release methods provides context to guide interpretation of the Deep Argo point release results in the next section. The difference in the derived PDFs between Figs. 8 and 10 implies that at middepth the mesoscale motion of a single float will be harder to predict than the ensemble distribution of many floats released over a subdomain. However, deep ocean velocities are less energetic, and the results in the next section will show less fragmented PDFs that result from simulations of Deep Argo float point releases.

4. Simulation of Deep Argo floats

a. Deep Argo floats

The six Deep SOLO floats listed in Table 1 were used to assess model simulations of deep float trajectories. Each float went through a few test profiles after deployment to progressively deeper depths, in order to test float performance. For this study we removed the initial experimental stage and started our simulations and analysis with the first full-depth float cycles. The simulated Deep Argo trajectories were initiated at six fixed points, corresponding to the locations where we begin analysis of the Deep SOLO floats (see Table 1). Our goal is to ask whether the model-derived float trajectories provide trajectory statistics that can be useful in guiding planning for regional Deep Argo deployments.

b. PDFs of simulated Deep Argo floats

The Deep Argo simulation output for the six Deep Argo cases is shown as 2D-PDFs in Fig. 11. Because mesoscale variability is relatively low in the deep ocean, the distributions of simulated Argo floats are spatially compact, with statistically significant regions of the PDFs confined to regions less than $2°$ in longitude $\times 2°$ in latitude. This contrasts with the fragmented PDFs in Fig. 10 for single-float cases computed with simulated core Argo. In particular, for cases 6003, 6008, 6010, 6011, and 6015, the maximum of the PDFs $P(x, y, 100\,\text{days}, 1\,\text{km})$ is approximately $0.02\% \,\text{km}^{-2}$, meaning that after 100 days about 1 out of 100 floats can be found within the peak $1/12° \times 1/12°$ grid boxes.

Since we have a limited number of Deep Argo trajectories, the KS test cannot be applied. Therefore, we assess the accuracy of PDFs on the basis of visual comparisons with the trajectories of real floats. The results,
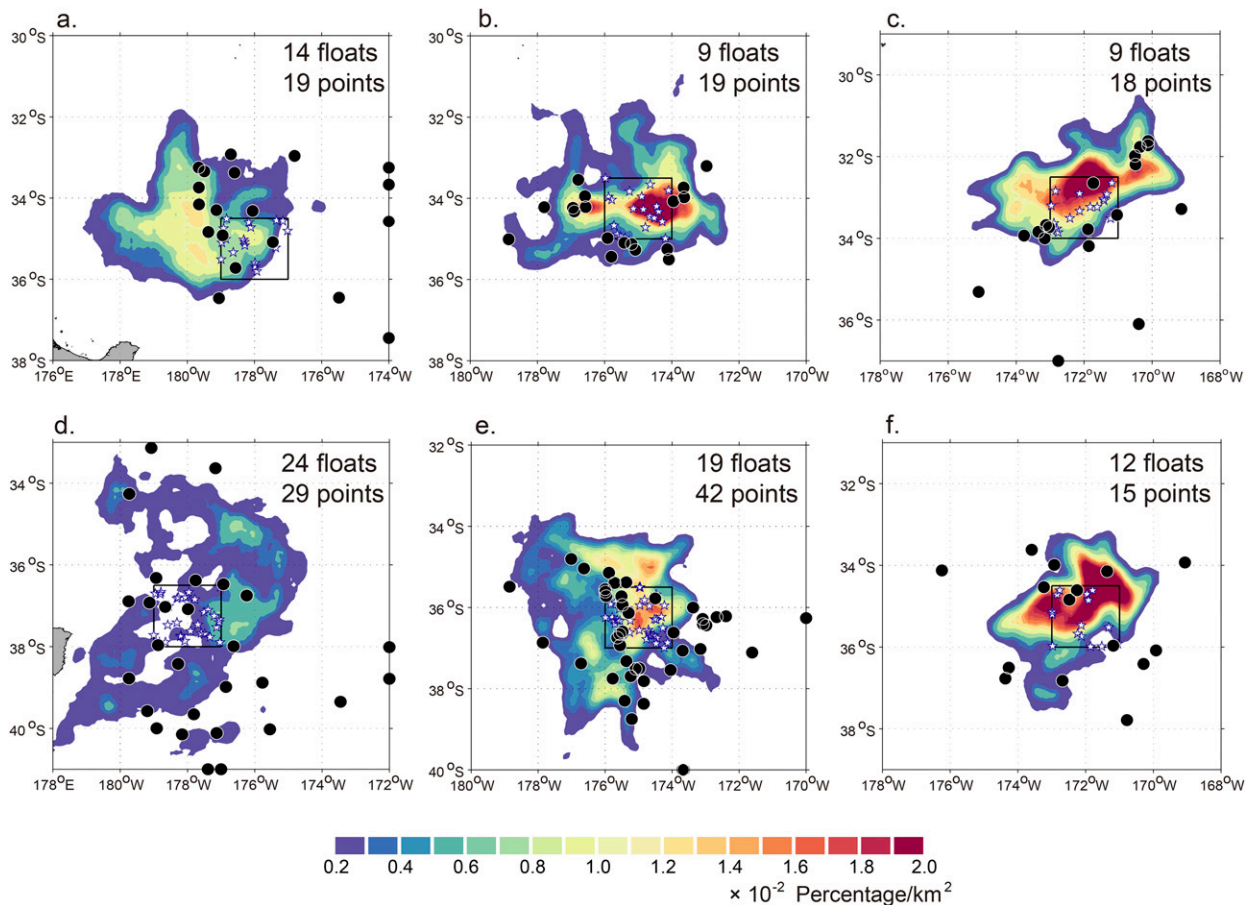
FIG. 8. The 2D-PDFs of the simulated core Argo floats at the end of 100 days in color. The starting positions of the real core Argo floats are marked by open stars, and end locations are indicated by closed circles. Panel letters (a)–(f) correspond to the boxes A–F in Fig. 1.

shown in Fig. 11, indicate that for cases 6003, 6008, and 6015 the final float positions (black dots) are within the nonnegligible regions of the PDFs, implying that the predictions are consistent with the observations. In the remaining cases (6002, 6010 and 6011), we find only weak agreement. Regardless, the Deep Argo locations and the significant PDF clouds differ by less than $1/2°$ (i.e., ~ 40 km), and this geographic proximity suggests that simulations can provide relevant information about the slow rate of displacement in many parts of the deep ocean to inform Deep Argo deployment plans.

c. Model-derived trajectories within the full domain

The six Deep Argo floats considered in Fig. 11 were deployed in the interior of the Pacific Ocean, where deep currents are relatively quiescent, offshore of the DWBC system at the Tonga–Kermadec Ridge (Whitworth et al. 1999). To explore the representativeness of these floats, we released simulated core Argo

and Deep Argo floats on a regular grid throughout our whole analysis domain, including regions with strong topographically steered deep currents (red vectors in Fig. 1). Figure 12 shows the resulting simulated Argo trajectories. Compared with core Argo (Fig. 12a), Deep Argo floats (Fig. 12b) experience minimal displacements from the initial release points. For the floats released in the ocean interior, the mean displacement is less than 50 km. In contrast floats released in the DWBC region and near topography (Fig. 1) show displacements greater than 200 km. At middepth, core Argo float displacements show more anisotropic spreading, with mean displacements ranging from 100 to 200 km.

Because Deep Argo floats in the quiescent ocean interior do not spread much, they do not readily capture the full range of variability in deep ocean currents. If a global Deep Argo program aimed to obtain representative deep ocean velocity statistics, then care would be needed to deploy floats within a full range of deep ocean dynamical regimes.
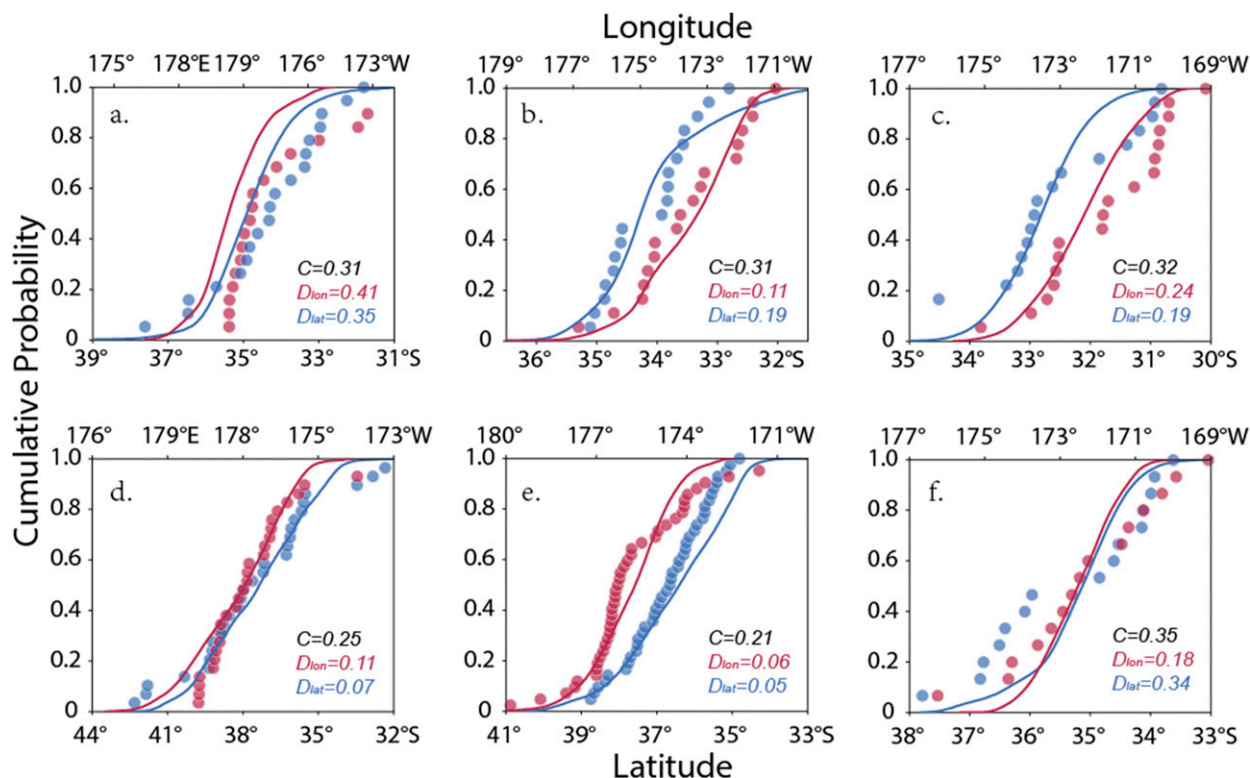
FIG. 9. CDFs of the numerical particles (curves) and the core Argo floats (circles) at the end of ~100-day trajectories. The CDFs are calculated along longitude (red) and latitude (blue). Panels (a)–(f) correspond to subdomains A–F in Fig. 1. The Kolmogorov–Smirnov test statistics (C, D_{lon}, D_{lat}; see appendix) are indicated in each panel.

5. Uncertainties of the simulation system

The random walk scheme implemented in the Lagrangian tracking model is necessary to account for the mixing resulting from unrepresented processes in the model velocities. As stated above, we set the horizontal diffusivity to $1\,\mathrm{m^2\,s^{-1}}$, which is consistent with the explicit diffusivity parameterization of the 1/12°-resolution Southern Ocean simulation. Material line stretching in turbulent flows predicts that two neighboring fluid particles should move apart exponentially with time (Batchelor 1952). This means that as the numerical integration proceeds, uncertainties in simulated Argo float positions can grow exponentially. When ocean currents are included, the stochastic portion of the float displacement is expected to be enhanced and can be thought of as an amplified random walk.

To evaluate this, we deploy 28 000 simulated core Argo floats and 28 000 simulated Deep Argo floats released at a fixed location (36°S, 177°W; see Fig. 13) within the southwest Pacific basin and estimate the amplified random walk scale via a bootstrapping method. The mean amplified random walk scale is defined by the root-mean-square distance between each float and the center of mass of the ensemble of floats:

TABLE 2. Case study core Argo trajectories. Analysis durations are selected to be as close to 100 days as possible, given actual float cycle times.

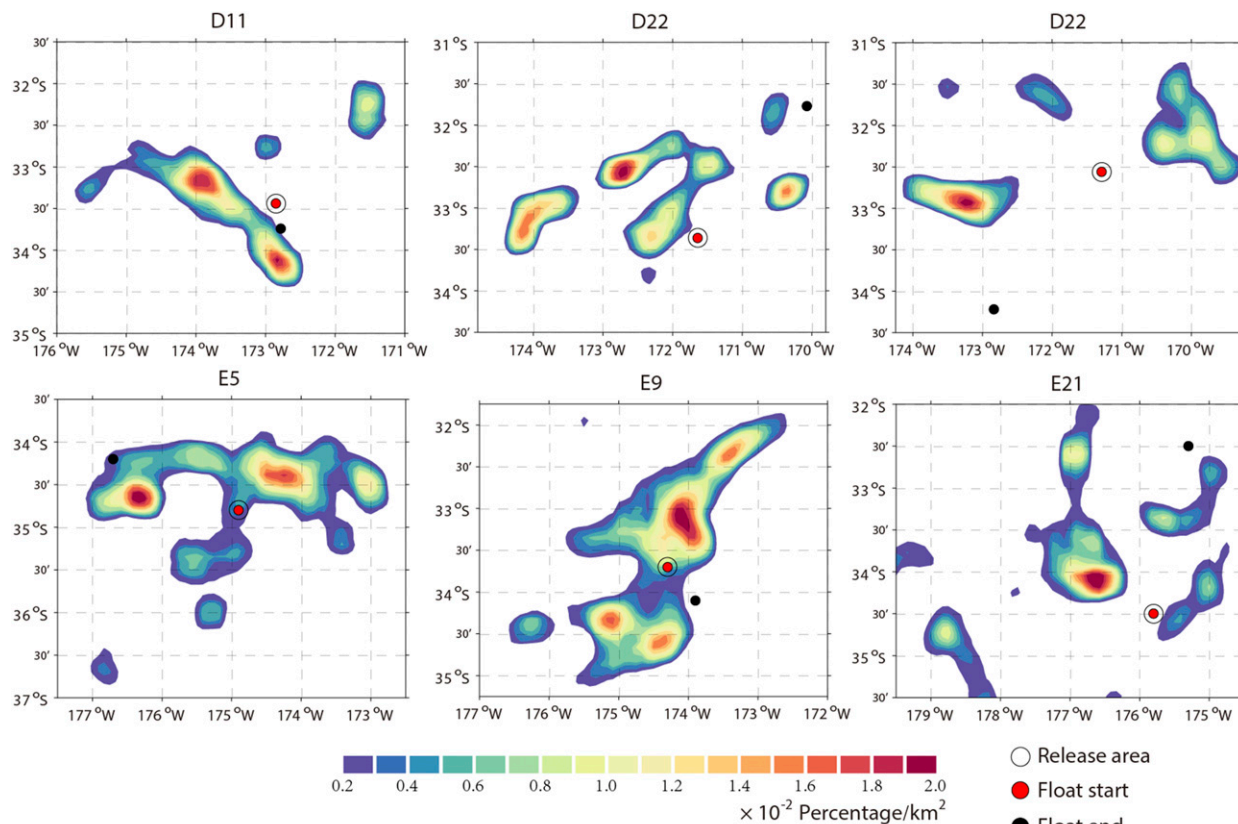| Case No.	Float series	Dates	Lon (°W)	Lat (°S)
D11	5901426	17 Nov 2007–25 Feb 2008	172.8–172.8	33.4–33.7
D22	5902176	13 Dec 2011–23 Mar 2012	171.6–170.1	33.4–31.8
D24	5903344	11 Feb 2012–19 May 2012	171.3–172.8	32.6–34.7
E9	5900401	6 Mar 2007–11 Jun 2007	173.3–173.8	33.7–34.1
E5	5901797	24 Apr 2009–1 Aug 2009	174.9–176.8	34.8–34.1
E21	5901798	13 May 2012–21 Aug 2012	175.8–175.3	34.5–32.4

FIG. 10. As in Fig. 8, but for six specific core Argo floats from domains D and E in Fig. 1. The starting positions of the real core Argo floats are marked by closed red circles, and end points are indicated by closed black circles. The release positions for simulated floats are contained within the open circle. Panel letters indicate the domain, and details for each numbered float are listed in Table 2.

$$S(t) = \sqrt{\frac{\sum\limits_{i=1}^{n} [(\mathbf{X}_i(t) - \mathbf{X}_C(t)]^2}{n}},$$

where t is the running time, n is the number of floats, $\mathbf{X}_i$ denotes the location of the ith float, and $\mathbf{X}_C$ denotes the center of mass of the ensemble of floats:

$$\mathbf{X}_C = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i.$$

The random walk can be associated with an effective mixing scale, which can be derived by advecting particles with eddies only, using zero background velocity to obtain an estimate unbiased by ocean currents.

The results, shown in Fig. 14, reveal that for time scales less than 100 days, both the original random walk (black line) and the amplified random walk (red and blue solid lines) grow exponentially, though the amplified growth is much more rapid. At the end of 100 days, the deviation associated with the original random walk is only 7 km. In contrast, when we include the effect of

ocean currents, it increases to 90 ± 20 km for simulated core Argo floats and 50 ± 10 km for Deep Argo floats. The magnitude of the amplified random walk scale is similar to the radius of the PDFs (~100 km for core Argo in Fig. 10; ~50 km for Deep Argo in Fig. 11). This suggests that the uncertainties in simulated-float end points can be thought of as stemming from the current-induced component of a random walk.

Since the release point used for these tests was selected to have strong eddy kinetic energy (EKE) both at middepth (1000 m) and near the bottom (5000 m) (see Fig. 13), the random walk statistics found here reflect the influence by eddies. The differences in eddy energy in the model relative to the real ocean could lead to trajectory differences, as would discrepancies in the mean velocities. We hypothesize that if eddy effects are suppressed by time averaging the advective velocities, then particles should experience different behavior over 100 days. To evaluate the impact of eddies, we compute the component of a random walk induced by the mean current. Here the mean is defined simply as the temporal average from the 380 days of archived model output,
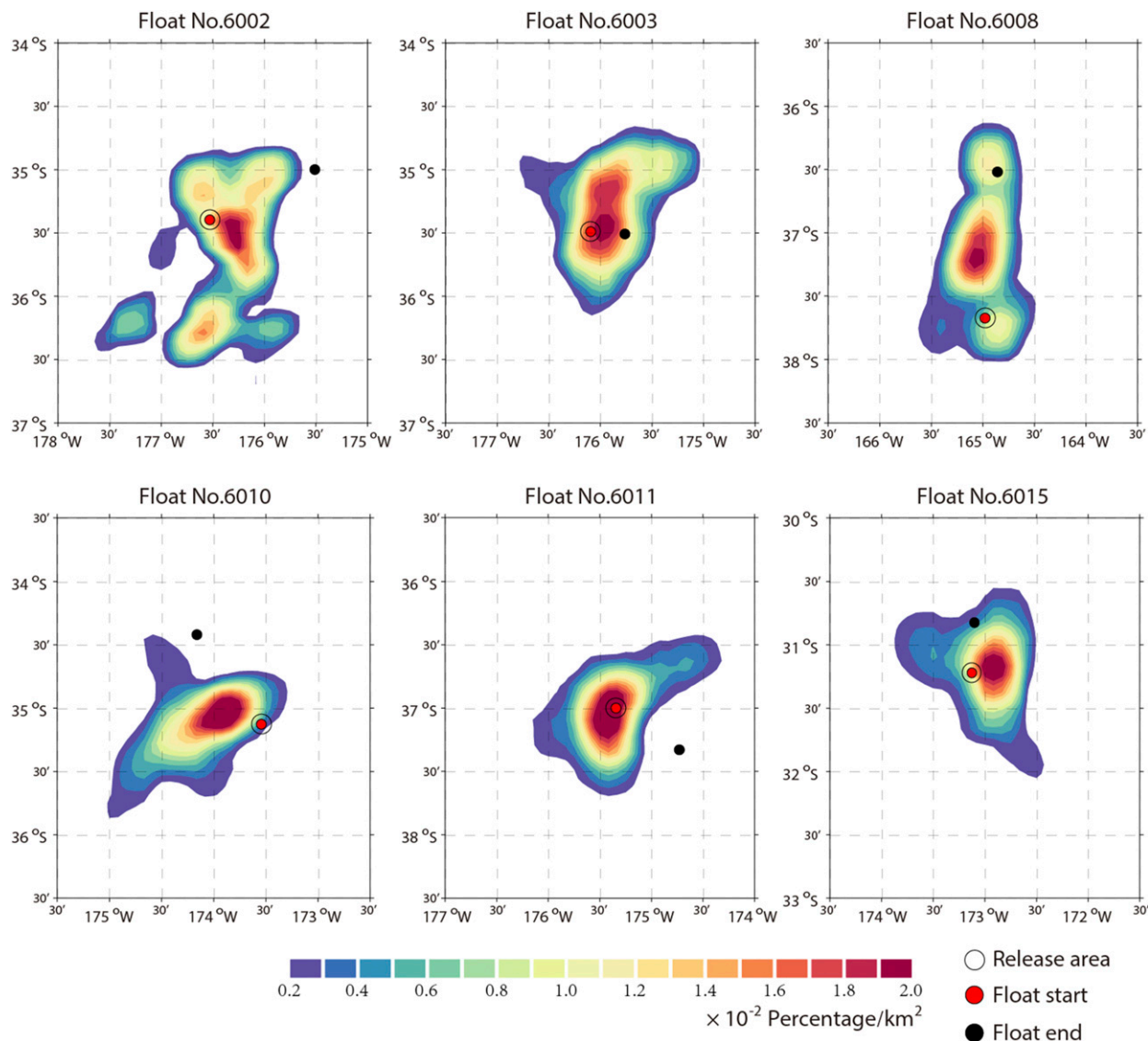
FIG. 11. As in Fig. 8, but for the six Deep Argo floats shown in Fig. 1. Panel labels indicate the float number, with the details listed in Table 1.

and a significant mesoscale structure does remain in this steady field. Nevertheless, using this means the random walk displacements reduce to 14 km for simulated core Argo floats and 12 km for Deep Argo floats (see Fig. 14), implying that within the first 100 days, transient eddies are responsible for most of the amplified random walk.

6. Conclusions

This paper has described a methodology for predicting the statistics of Lagrangian motions of core Argo and Deep Argo floats. By understanding the dynamical processes that govern the trajectories of simulated Argo floats, we can gain a window into the dynamical information contained in Argo float displacements. The key feature of this methodology is that we integrate the governing equation for the particle position using both the model velocity vectors and a subgrid parameterization, simulated by a stochastic random walk scheme, to produce probability density functions (PDFs) informing future float positions. The simulated floats adhere to the same cycle timings as the real floats, with matched time intervals spent at the parking depth and at the sea surface.

Tests for the simulation method carried out using velocities from a 1/12° Southern Ocean model showed that with a minimum of 10 000–15 000 float trajectories,
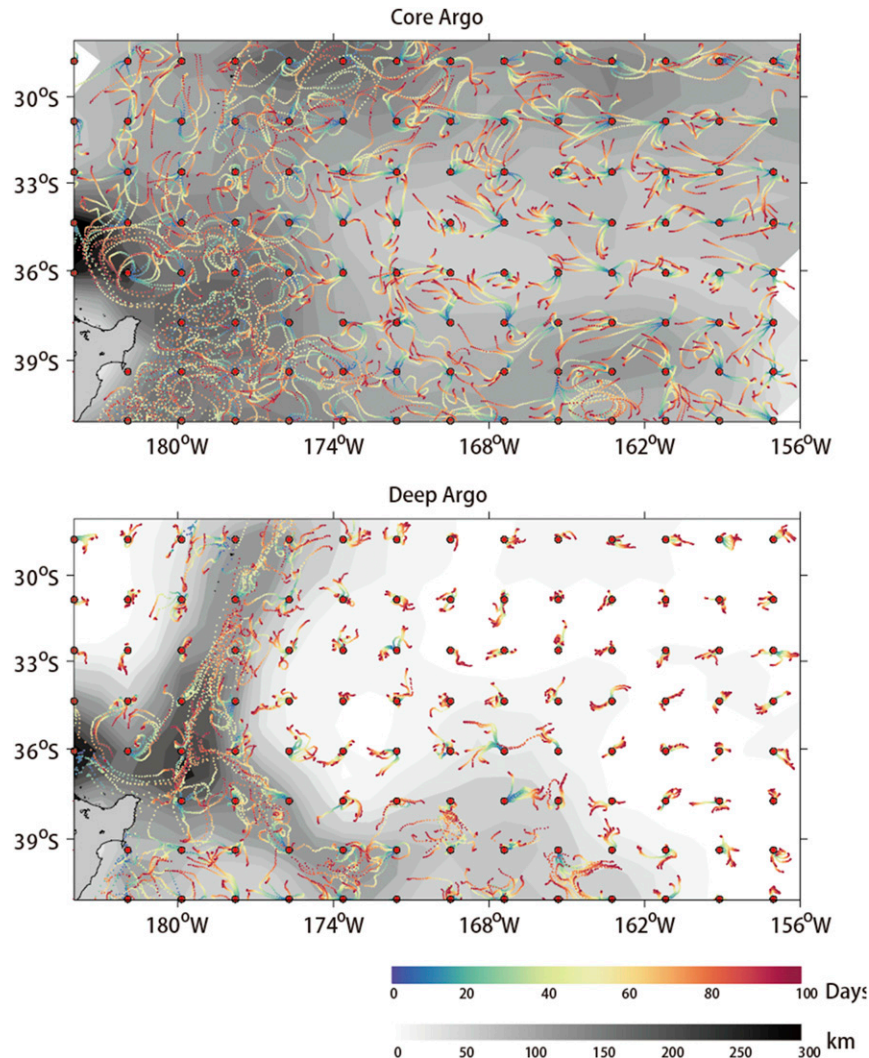
FIG. 12. The Lagrangian trajectories for (a) simulated core Argo floats and (b) simulated Deep Argo floats during 100-day model runs. The trajectories are marked with colored dots indicating time, and they are superimposed over shading representing total displacements of the trajectories. To contrast the trajectories, both panels show a set of nine subreleases at intervals of 30 days. This differs from the release method described previously (10-day intervals for core Argo and 40-day intervals for Deep Argo). During the 100-day simulations, the floats are programmed to follow the behavior of core Argo floats and Deep Argo floats (see text).

100-day float displacements converge to yield robust PDFs. In this study a total of 28 000 (100 day) simulated middepth trajectories (and 24 000 simulated deep Argo trajectories) were analyzed. In five of six subdomains, a K–S test indicates that the statistics of simulated 100-day core Argo float displacements are consistent with observed Argo float displacements. These results tell us that the numerical simulations are successful at replicating the behavior of an ensemble of 100-day trajectories from the Argo array.

For Deep Argo, a large ensemble of float trajectories is not yet available, and simulated Deep Argo floats are

compared with six Deep SOLO floats that were deployed northeast of New Zealand in 2014 and 2016. A direct comparison between the relatively high-probability regions of simulated-float PDFs and the positions of real floats implies that in three of these cases, float trajectories are successfully predicted. Though the direction was unsuccessfully predicted in the other three cases, in all six cases the distances traveled by the simulated floats and the real floats were similar. The differences between single-float trajectories and simulated PDFs are unsurprising and can be attributed to two factors. First, a large number of trajectories are
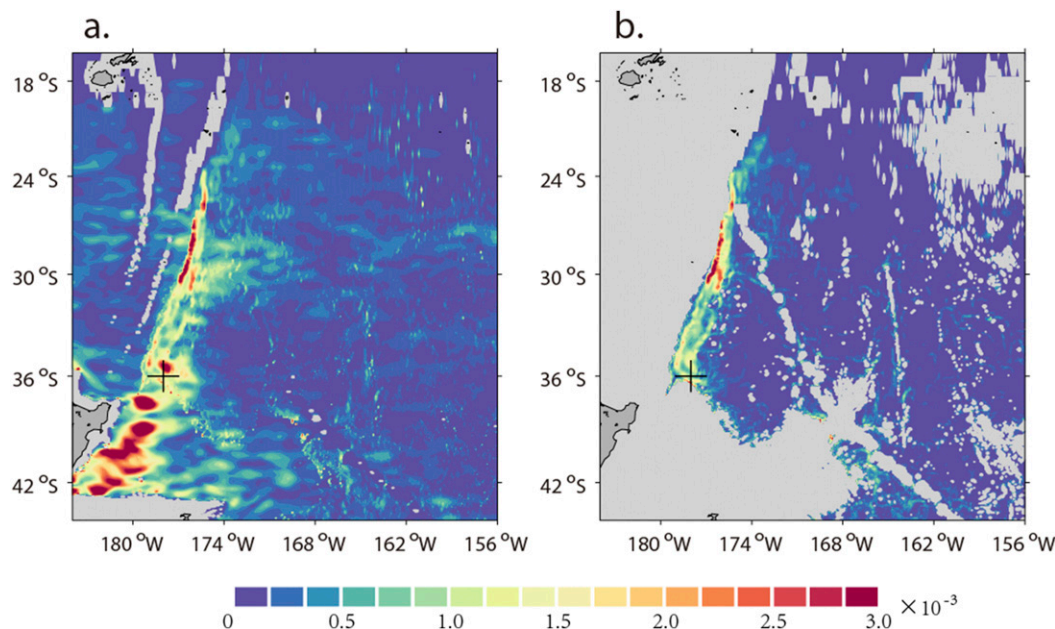
FIG. 13. Mean EKE ($m^2 s^{-2}$) at (a) 1000- and (b) 5000-m depth for 2009, based on velocities from the 1/12°
Southern Ocean simulation. The light gray shading indicates solid earth. The black cross in both panels mark the
same simulated floats release location used in Fig. 14.

needed to obtain robust PDFs, and emulating a single
float thus requires releasing a large number of simulated
floats within a small geographic region, which means
that simulated floats may not be entirely statistically
independent. Second, the model may misrepresent to-
pographic features, and deep eddies and flow structures
may differ from those of the true ocean (particularly for
phenomena that occur on scales smaller than the Rossby
radius), leading to inaccuracies in advective field
statistics.

Ocean velocities amplify float motions relative to
what would be obtained from a diffusive random walk.
A preliminary assessment of the Lagrangian tracking
model indicates that over 100 days, the length scale of
the ocean-amplified random walk is 90 ± 20 km for
simulated core Argo floats and 50 ± 10 km for Deep
Argo floats. The amplified random walk associated with
the mean current alone is only 14 km for core Argo floats
and 12 km for Deep Argo floats. This indicates that
transient eddies are the major contributor to the am-
plified random walk. Our 1/12°-resolution simulation
captures mesoscale eddies in a statistical sense; however,
it is unclear how the results would differ were we to
resolve higher spatial and temporal scales.

The Lagrangian float simulations and statistical ana-
lyses presented in this paper provide a framework for
coordinating Argo and Deep Argo to meet observing
system requirements, with a particular focus on projec-
ting the behavior of a large ensemble of floats. This
approach has the potential to be particularly valuable
both in the deep ocean, where we have little in situ in-
formation about the velocity field (Johnson et al. 2015),
and in regions where middepth Lagrangian motions can
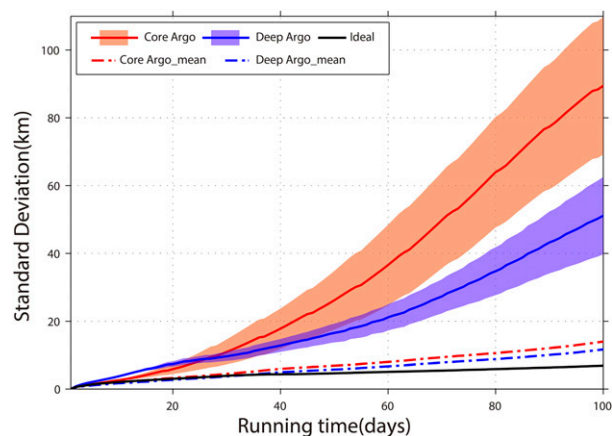be substantial: equatorial regions, marginal seas, seasonal

FIG. 14. The amplified random walk scales of simulated floats as
a function of run duration. The scale is calculated by the mean dis-
tance between each simulated float and the center of mass of the
ensemble of floats. Uncertainties, indicated by shading, are estimated
via a bootstrapping method, using 28 000 simulated floats released at
a fixed location (36°S, 177°W) within the southwest Pacific basin. The
curves are the random walk scales amplified by the full velocities
(solid) and present the components amplified by only the mean flows
(dashed). The results of the core Argo simulation (red lines) and the
Deep Argo simulation results (blue lines) with the original random
walk curve (without the ocean currents) (black lines).

ice zones, and western boundary currents. Another use of this method is for targeted campaigns. One may be interested in equipping Argo floats with additional sensors—for example, to measure microstructure (Lien et al. 2016) or biogeochemistry (Johnson and Claustre 2016)—and using them in process studies. These studies typically target a specific location, and an estimated PDF predicting float positions may be of utility.

APPENDIX

The Kolmogorov–Smirnov Test

The Kolmogorov–Smirnov test (hereinafter the KS test) provides a formal statistical metric for quantifying the probability that two datasets could be drawn from the same distribution (Stephens 1974). In our study the KS statistic is $D_{n,n'} = \sup_x |F_{s,n}(x) - F_{o,n'}(x)|$, where $F_{s,n}$ and $F_{o,n'}$ are the CDFs of the simulations and observations, respectively; sup is the supremum of the set of distances between the two CDFs; n is the number of simulated floats (28 000); and n' is the number of real floats within each subdomain (indicated in Fig. 8). The null hypothesis is rejected at level α if $D_{n,n'} > C(\alpha, n, n')$, where $C(\alpha, n, n') = c(\alpha)\sqrt{(n + n')/(nn')}$. The coefficient $c(\alpha)$ is a function of the significance level α. Working under the assumption that the CDFs of our simulated Argo floats are independent, we interpret the simulations to match the observations only when neither the meridional nor the zonal KS statistic (D_{lon}, D_{lat}) would justify rejecting the null hypothesis.

REFERENCES

Batchelor, G. K., 1952: The effect of homogeneous turbulence on material lines and surfaces. *Proc. Roy. Soc. London*, **A213**, 349–366, https://doi.org/10.1098/rspa.1952.0130.

Chassignet, E. P., H. E. Hurlburt, O. M. Smedstad, G. R. Halliwell, P. J. Hogan, A. J. Wallcraft, R. Baraille, and R. Bleck, 2007: The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *J. Mar. Syst.*, **65**, 60–83, https://doi.org/10.1016/j.jmarsys.2005.09.016.

Chu, P. C., L. M. Ivanov, O. V. Melnichenko, and N. C. Wells, 2007: On long baroclinic Rossby waves in the tropical North Atlantic observed from profiling floats. *J. Geophys. Res.*, **112**, C05032, https://doi.org/10.1029/2006JC003698.

Davis, R. E., and W. Zenk, 2001: Subsurface Lagrangian observations during the 1990s. *Ocean Circulation and Climate: Observing and Modelling the Global Ocean*, G. Siedler, J. Church and J. Gould, Eds., International Geophysics Series, Vol. 77, Academic Press, 123–139, https://doi.org/10.1016/S0074-6142(01)80115-7.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Gille, S. T., 2003: Float observations of the Southern Ocean. Part I: Estimating mean fields, bottom velocities. *J. Phys. Oceanogr.*, **33**, 1167–1181, https://doi.org/10.1175/1520-0485(2003)033<1167:FOOTSO>2.0.CO;2.

Gray, A. R., and S. C. Riser, 2014: A global analysis of Sverdrup balance using absolute geostrophic velocities from Argo. *J. Phys. Oceanogr.*, **44**, 1213–1229, https://doi.org/10.1175/JPO-D-12-0206.1.

Hunter, J., P. Craig, and H. Phillips, 1993: On the use of random walk models with spatially variable diffusivity. *J. Comput. Phys.*, **106**, 366–376, https://doi.org/10.1016/S0021-9991(83)71114-9.

Johnson, G. C., and J. M. Lyman, 2014: Oceanography: Where's the heat? *Nat. Climate Change*, **4**, 956–957, https://doi.org/10.1038/nclimate2409.

——, ——, and S. G. Purkey, 2015: Informing Deep Argo array design using Argo and full-depth hydrographic section data. *J. Atmos. Oceanic Technol.*, **32**, 2187–2198, https://doi.org/10.1175/JTECH-D-15-0139.1.

Johnson, K. S., and H. Claustre, 2016: Bringing biogeochemistry into the Argo age. *Eos*, **97**, 407–435, https://doi.org/10.1029/2016EO051467.

Kinderman, A. J., and J. F. Monahan, 1977: Computer generation of random variables using the ratio of uniform deviates. *ACM Trans. Math. Software*, **3**, 257–260, https://doi.org/10.1145/355744.355750.

LaCasce, J. H., 2000: Floats and *f/H. J. Mar. Res.*, **58**, 61–95, https://doi.org/10.1357/002224000321511205.

Lebedev, K. V., H. Yoshinari, N. A. Maximenko, and P. W. Hacker, 2007: YoMaHa'07: Velocity data assessed from trajectories of Argo floats at parking level and at the sea surface. IPRC Tech. Note 4 (2), 16 pp.

Lien, R.-C., T. B. Sanford, J. A. Carlson, and J. H. Dunlap, 2016: Autonomous microstructure EM-APEX floats. *Methods Oceanogr.*, **17**, 282–295, https://doi.org/10.1016/j.mio.2016.09.003.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.*, **102**, 5753–5766, https://doi.org/10.1029/96JC02775.

Mazloff, M. R., and C. Boening, 2016: Rapid variability of Antarctic Bottom Water transport into the Pacific Ocean inferred from GRACE. *Geophys. Res. Lett.*, **43**, 3822–3829, https://doi.org/10.1002/2016GL068474.

——, P. Heimbach, and C. Wunsch, 2010: An eddy-permitting Southern Ocean state estimate. *J. Phys. Oceanogr.*, **40**, 880–899, https://doi.org/10.1175/2009JPO4236.1.

McAdam, R., and E. van Sebille, 2018: Surface connectivity and interocean exchanges from drifter-based transition matrices. *J. Geophys. Res. Oceans*, **123**, 514–532, https://doi.org/10.1002/2017JC013363.

McCave, I. N., and L. Carter, 1997: Recent sedimentation beneath the deep western boundary current off northern New Zealand. *Deep-Sea Res. I*, **44**, 1203–1237, https://doi.org/10.1016/S0967-0637(97)00011-3.

Ollitrault, M., and J.-P. Rannou, 2013: ANDRO: An Argo-based deep displacement dataset. *J. Atmos. Oceanic Technol.*, **30**, 759–788, https://doi.org/10.1175/JTECH-D-12-00073.1.

Park, J. J., and K. Kim, 2013: Deep currents obtained from Argo float trajectories in the Japan/East Sea. *Deep-Sea Res. II*, **85**, 169–181, https://doi.org/10.1016/j.dsr2.2012.07.032.

——, ——, B. A. King, and S. C. Riser, 2005: An advanced method to estimate deep currents from profiling floats. *J. Atmos. Oceanic Technol.*, **22**, 1294–1304, https://doi.org/10.1175/JTECH1748.1.

Roemmich, D., and J. Gilson, 2009: The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Prog. Oceanogr.*, **82**, 81–100, https://doi.org/10.1016/j.pocean.2009.03.004.

——, and Coauthors, 1998: On the design and implementation of Argo—An initial plan for a global array of profiling floats. ICPO Rep. 21, GODAE Rep. 5, 32 pp.

——, and Coauthors, 2009: The Argo Program: Observing the global oceans with profiling floats. *Oceanography*, **22** (2), 34–43, https://doi.org/10.5670/oceanog.2009.36.

Stephens, M. A., 1974: EDF statistics for goodness of fit and some comparisons. *J. Amer. Stat. Assoc.*, **69**, 730–737, https://doi.org/10.1080/01621459.1974.10480196.

Tamsitt, V., and Coauthors, 2017: Spiraling pathways of global deep waters to the surface of the Southern Ocean. *Nat. Commun.*, **8**, 172, https://doi.org/10.1038/s41467-017-00197-0.

van Sebille, E., and Coauthors, 2018: Lagrangian ocean analysis: Fundamentals and practices. *Ocean Modell.*, **121**, 49–75, https://doi.org/10.1016/j.ocemod.2017.11.008.

Warren, B. A., 1981: The shallow oxygen minimum of the South Indian Ocean. *Deep-Sea Res.*, **28A**, 859–864, https://doi.org/10.1016/S0198-0149(81)80005-2.

Whitworth, T., III, B. A. Warren, W. D. Nowlin Jr., S. B. Rutz, R. D. Pillsbury, and M. I. Moore, 1999: On the deep western-boundary current in the Southwest Pacific Basin. *Prog. Oceanogr.*, **43**, 1–54, https://doi.org/10.1016/S0079-6611(99)00005-1.

Zilberman, N. V., and G. Maze, 2015: Report on the Deep Argo Implementation Workshop. Ifremer Doc. LPO-15-04, 36 pp., http://archimer.ifremer.fr/doc/00281/39238/37799.pdf.

——, D. H. Roemmich, and S. T. Gille, 2013: The mean and the time variability of the shallow meridional overturning circulation in the tropical South Pacific Ocean. *J. Climate*, **26**, 4069–4087, https://doi.org/10.1175/JCLI-D-12-00120.1.