# Scripps Argo Trajectory-Based Velocity Product: Global Estimates of Absolute Velocity Derived from Core, Biogeochemical, and Deep Argo Float Trajectories at Parking Depth

N. V. ZILBERMAN,[a] M. SCANDERBEG,[a] A. R. GRAY,[b] AND P. R. OKE[c]

[a] *Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California*
[b] *School of Oceanography, University of Washington, Seattle, Washington*
[c] *Commonwealth Scientific and Industrial Research Organisation, Hobart, Tasmania, Australia*

(Manuscript received 15 June 2022, in final form 7 December 2022)

ABSTRACT: Global estimates of absolute velocities can be derived from Argo float trajectories during drift at parking depth. A new velocity dataset developed and maintained at Scripps Institution of Oceanography is presented based on all Core, Biogeochemical, and Deep Argo float trajectories collected between 2001 and 2020. Discrepancies between velocity estimates from the Scripps dataset and other existing products including YoMaHa and ANDRO are associated with quality control criteria, as well as selected parking depth and cycle time. In the Scripps product, over 1.3 million velocity estimates are used to reconstruct a time-mean velocity field for the 800–1200 dbar layer at 1° horizontal resolution. This dataset provides a benchmark to evaluate the veracity of the BRAN2020 reanalysis in representing the observed variability of absolute velocities and offers a compelling opportunity for improved characterization and representation in forecast and reanalysis systems.

SIGNIFICANCE STATEMENT: The aim of this study is to provide observation-based estimates of the large-scale, subsurface ocean circulation. We exploit the drift of autonomous profiling floats to carefully isolate the inferred circulation at the parking depth, and combine observations from over 11 000 floats, sampling between 2001 and 2020, to deliver a new dataset with unprecedented accuracy. The new estimates of subsurface currents are suitable for assessing global models, reanalyses, and forecasts, and for constraining ocean circulation in data-assimilating models.

KEYWORDS: Ocean; Large-scale motions; Trajectories; Oceanic profilers

## 1. Introduction

Large-scale ocean currents carry and redistribute heat, freshwater, carbon, and nutrients across ocean basins and therefore have an outsized impact on air–sea heat exchange, sea level variability, ocean uptake of anthropogenic carbon dioxide, and biological productivity (Frajka-Williams et al. 2019). The technological development of satellite-based sensors, in combination with high-frequency radar and surface drifter measurements, has significantly improved our understanding of the temporal variability and spatial structure of surface ocean dynamics (Villas Bôas et al. 2019). Although progress in characterizing the circulation beneath the surface has been achieved through the integration of surface and interior measurements, including those from repeat hydrography, Argo floats, ocean moorings, and high-resolution expendable bathythermographs (Davis et al. 2019), it remains challenging to directly observe the subsurface ocean circulation on large scales.

To leading order, the large-scale ocean circulation is well described by geostrophic balance, the equilibrium between the Coriolis and pressure gradient forces. Together with the hydrostatic relationship, this balance can be used to infer the vertical shear of horizontal geostrophic velocity from the density field (Gill 1982, 215–216). While this shear is frequently combined with an assumption of a level of no horizontal motion at a fixed neutral density surface or the deepest common depth (e.g., Zilberman et al. 2014; Fu et al. 2020), ambiguity in the choice of such a reference level often leads to uncertainties in the orientation and amplitude of the absolute oceanic currents (Beal and Bryden 1999; Frajka-Williams et al. 2018). A more accurate way to define the patterns and strength of the large-scale ocean circulation is to reference the geostrophic shear using measured absolute velocities at a single depth or pressure surface, i.e., a level of known motion (Gray and Riser 2014; Ollitrault and De Verdiere 2014).

Argo float trajectories can be used to estimate subsurface absolute velocities by considering the distance of drift at the parking depth and the duration between beginning and end of drift (Davis et al. 1992). The global Argo array, with about 3900 currently active floats, constitutes the main source of temperature and salinity profiles in the upper-2000-m ocean, and of absolute velocity measurements at parking depth (Jayne et al. 2017). The ongoing expansion of Argo, called OneArgo, includes increased numbers of Core Argo floats measuring temperature and salinity in the equatorial regions, marginal seas, high latitudes, and western boundary currents, compared to the original 3° × 3° sampling array; Biogeochemical (BGC) Argo floats that collect BGC observations in the

*Corresponding author*: Nathalie Zilberman, nzilberman@ucsd.edu

DOI: 10.1175/JTECH-D-22-0065.1

upper-2000-m ocean; and Deep Argo floats capable of profiling to 4000 m or 6000 m depth (Roemmich et al. 2019). While most Core and BGC Argo floats drift at parking depth of 1000 m, a large number of Deep Argo floats park within a few meters off the bottom to keep floats close to their deployment location (Zilberman et al. 2020), while others are set to park at the core of a deep water mass in order to track the deep water pathways of the meridional overturning circulation (Racapé et al. 2019).

The primary challenge in estimating subsurface absolute velocities from Argo float trajectories is that a float's locations are known only while at the sea surface and when connection between the float and a satellite positioning system is established (Wong et al. 2020). When computing trajectory-based velocities, this fact leads to uncertainties that arise from a number of sources, including unmeasured advection by subsurface currents, tides, and eddies during float ascent and descent; untracked float drift at the surface between the end of ascent and first position fix and between the last position fix and start of descent; and missing positions in sea ice covered regions (Davis et al. 2001; Park et al. 2005; Chamberlain et al. 2018). Most Argo floats deployed since 2011, equivalent to two-thirds of the current array, are equipped with a high bandwidth two-way communication system, while the remaining older floats have Système Argos communication devices. Among the Argo floats equipped with two-way communication system that are presently active, 99.8% use Iridium and 0.2% use BeiDou. The switch from Système Argos to a two-way communication system has led to a shortening of the time at the surface from 8–12 h to 15–30 min, thereby reducing float drift at the surface and improving the accuracy of float positioning (Wong et al. 2020). Velocity errors due to float drift at the surface are more significant for floats transmitting via Système Argos than the two-way telecommunication system as they stay longer at the surface by design. Careful quality control on float timing and locations of surfacing and diving are required to minimize uncertainties in subsurface absolute velocity estimates (Wong et al. 2020).

The Argo Data Management System provides Argo profiles, trajectories, metadata, and technical data that are freely distributed in near–real time (NRT) within 12 h, and high-quality delayed mode (DM) profile data within 12 months. Quality control procedures are regularly implemented in the Argo data stream to enable improved correction of data biases and errors (Wong et al. 2020). In 2014, the Argo data system went through a major format change to adapt to new BGC data and also added new features and attributes to the metadata and trajectory data (Argo Data Management Team 2019). The updated trajectory file format, v3.1, included improved information regarding cycle timing and temperature, salinity, and pressure measurements at parking depth. A "cookbook" was developed to provide guidance to the Data Assembly Centers (DACs) on how to consistently create real-time trajectory files for the various Argo float types and is regularly updated on the Argo website (Scanderbeg et al. 2019). The development of more consistent and detailed trajectory files now allows for more accurate estimates of velocities from Argo float trajectories.

Comparisons between Argo-based subsurface absolute velocities are needed for early identification and correction of error in the trajectory dataset. Argo trajectory products that have been generated include YoMaHa07 (Lebedev et al. 2007), Argo New Displacements Rannou and Ollitrault (ANDRO; Ollitrault et al. 2020; Ollitrault and Rannou 2013), and Absolute Geostrophic Velocities from Argo (AGVA; Gray and Riser 2014). The AGVA product is based on Argo trajectories collected between 2004 and 2010. The YoMaHa07 dataset initially covered 2001–07 and has been updated monthly since 2007. The ANDRO product was created in 2010 and has been updated periodically, with the last update in 2020. Gridded products have been made available to facilitate the use of subsurface velocity measurements to the users. These include the 1° gridded YoMaHa07 velocities called G-YoMaHa (Katsumata and Yoshinari 2010), the 3° binned ANDRO trajectory, and the 1° AGVA mapped product. While YoMaHa07 and ANDRO use first and last surface fixes to compute subsurface velocities, in the case of Système Argos communication system, the AGVA product uses extrapolation at the surface from a least squares fit to a linear background flow and inertial current to determine times and locations of the float surfacing and sinking, following Park et al. (2005).

Here, we introduce a new velocity dataset based on Argo float trajectories, which has been developed at Scripps Institution of Oceanography. The main differences between the Scripps Argo velocity product and the ANDRO and YoMaHa datasets described above are that the Scripps product extensively applies thorough quality control checks to all float locations and times, only uses trajectory files with the updated 2014 or later format, and provides observed parking depths instead of relying on preprogrammed values. It is expected that these advanced quality checks contribute to significantly improving the overall quality of the trajectory-based velocity dataset. The Scripps Argo velocity dataset will be updated every 6 months and made freely available via the UCSD library (Zilberman et al. 2022), the Argo website (https://argo.ucsd.edu/data/argo-data-products/velocity-products/), and the Argovis web app and database [https://argovis.colorado.edu, recently upgraded from the app version described in Tucker et al. (2020)]. The methods for quality control of the trajectory data, computation of the velocity estimates, and gridding of the data are explained in section 2. Section 3 compares the velocity estimates from the Scripps dataset with the ANDRO and YoMaHa07 products, assesses the impact of correcting for surface drift in older floats that use Système Argos, and evaluates the gridded mean velocity fields from the Scripps dataset against a global ocean reanalysis. Conclusions follow in section 4.

## 2. Method

### a. *Quality control of the trajectory data in the Scripps product*

The Scripps Argo velocity product is based on quality-controlled trajectory data processed by the Argo DACs and made publicly available at the Argo Global Data Assembly Centers (GDACs; http://doi.org/10.17882/42182#76230). The

netCDF Scripps velocity dataset contains velocity estimates from drift at parking pressure, surfacing and diving times and positions, float identification numbers, cycle number, pressure and temperature at parking, positioning system, number of surface fixes for each cycle, and information on which sensors the float carries, which mission the float is part of (Core, BGC, or Deep), and whether the velocity was based on extrapolated positions. The Scripps product information is mostly similar to what is available for YoMaHa07 and ANDRO, but specifies only measured pressure and temperature at parking, provides velocities based on extrapolated surface positions for cycles using Système Argos with at least six good surface fixes, and includes a flag indicating if the float is a Core, BGC, or Deep model.

Out of 15 970 floats examined, 11 736 (73.5%) were included in the Scripps product. Excluded from the dataset are (i) floats that do not have a trajectory file on the Argo GDAC (606 floats, 3.8%), (ii) floats in trajectory file formats prior to v3.1 that have missing timing information and pressure measurements (2722 floats, 17.0%), and (iii) floats with fewer than three cycles with good positions or drift pressures (906 floats, 5.7%).

All float data go through a quality control procedure that first selects only data with a quality control flag of "1" in the trajectory files and then identifies bad surface positions. Procedures to estimate surface times and positions are described in section 2a(1), followed by a description of the quality control procedures applied to the cycle times [section 2a(2)], and drift pressures and temperatures [section 2a(3)]. After all quality control criteria have been applied, a float must have at least three good surface positions and drift pressures within the 100–6200 dbar range to be included in the Scripps product.

#### 1) SURFACE TIMES AND POSITIONS

Determining times and locations of the float's surfacing and sinking depends on the type of communication system and the float model. For trajectory cycles in delayed mode, adjusted times (JULD_ADJUSTED) and positions (POSITION) are used only if they have good ("1") quality control flags (JULD_ADJUSTED_QC and POSITION_QC, respectively). No additional quality control is applied to surface times and positions for these cycles. The quality control procedures applied to surface times and positions for real-time cycles are described next.

All newer floats with two-way communication systems are equipped with high-quality GPS positioning with accuracy of 10 m, which is 10–100 times more accurate than Système Argos. If the GPS fails to get a position fix, a lower accuracy Iridium fix is sometimes provided, if the Iridium data are sent through the Short Burst Data (SBD) method. In this product, surface intervals with single Iridium positions are not included due to their low accuracy (with errors of up to 5 km). However, when both Iridium and GPS positions are included in one surface interval, all positions are quality controlled according to the GPS quality control process described next. For most floats with GPS positioning (e.g., SOLO-II, Apex APF9 after July 2014, Navis, Apex APF11), there is often only one high-quality position and time fix; that fix is used for times and locations of float surfacing and sinking. Most SOLO-II floats are programmed to obtain two fixes, upon arrival at the surface and prior to descent; the first fix is assigned as the rise time and location, and the last fix is assigned as the fall time and location. When APEX floats have difficulty sending their data to the Iridium satellites, perhaps due to the rough sea state, they are programmed to stop transmitting to Iridium, obtain a new GPS fix, and then try transmitting to Iridium again. This process can be repeated several times, which results in multiple GPS positions per surface interval. Any time a float receives multiple GPS fixes per surface interval, the first fix is assigned as the rise time and location and the last fix is assigned as the fall time and location.

To perform further quality control on the GPS location fixes, speeds were calculated between the last GPS fix in the current cycle ($N$) and the first GPS fix in the next cycle ($N + 1$). Then, each subsurface speed ($N$ to $N + 1$) was compared to the mean of four independent neighboring subsurface speeds ($N - 3$ to $N - 2$, $N - 2$ to $N - 1$, $N + 2$ to $N + 3$, and $N + 3$ to $N + 4$). Unless the cycle is near the beginning or end of the float lifetime, GPS-based subsurface speed is compared with two preceding and two following subsurface speeds, and without using the same positions in any of the other speed calculations to ensure that the values are independent. The immediately preceding and following subsurface speeds are not used in the mean comparison to prevent using one of the potentially bad positions from the current subsurface speed in a neighboring subsurface speed. For subsurface speeds occurring near the beginning or end of the float lifetime, the four neighboring cycles are not split evenly between the immediately preceding and following subsurface speeds. For example, for the first subsurface speed of the float lifetime ($N$ to $N + 1$), the four neighboring speeds are from the following current cycles ($N + 2$ to $N + 3$, $N + 3$ to $N + 4$, $N + 4$ to $N + 5$, and $N + 5$ to $N + 6$). For the second subsurface speeds of the float lifetime ($N + 1$ to $N + 2$), the four neighboring speeds are from the following current cycles ($N + 3$ to $N + 4$, $N + 4$ to $N + 5$, $N + 5$ to $N + 6$, and $N + 6$ to $N + 7$). For the third subsurface speed ($N + 2$ to $N + 3$), one is from the first cycle ($N$ to $N + 1$) and the other three are from the following cycles ($N + 4$ to $N + 5$, $N + 5$ to $N + 6$, and $N + 6$ to $N + 7$). Likewise, for the final subsurface speed ($N$ to $N - 1$), the four neighboring speeds are all from the preceding cycles ($N - 2$ to $N - 3$, $N - 3$ to $N - 4$, $N - 4$ to $N - 5$, and $N - 5$ to $N - 6$). For the second to last subsurface speed ($N - 1$ to $N - 2$), the four neighboring speeds are from the preceding cycles ($N - 3$ to $N - 4$, $N - 4$ to $N - 5$, $N - 5$ to $N - 6$, and $N - 6$ to $N - 7$). For the third to last subsurface speed ($N - 2$ to $N - 3$), one speed is from the last cycle ($N$ to $N - 1$) and the other three are from cycles preceding it ($N - 4$ to $N - 5$, $N - 5$ to $N - 6$, and $N - 6$ to $N - 7$).

Cycles with speeds that exceed the average surrounding speeds by a factor of 5 are investigated as possibly indicative of an erroneous position far away from surrounding positions. The choice of a factor of 5 is admittedly subjective. Using a factor of 4 would lead to twice the number of bad positions and is thus less efficient at removing bad data than is using a factor of 5. On the other hand, using a factor larger than 5

would lead to detecting similar (within 10%) numbers of bad positions as the real-time Argo QC process and thus does not significantly improve on the real-time quality control. There are two situations that can occur when high speeds are identified: in the first situation both surface intervals have only one GPS fix, and in the second situation one or both surface intervals have multiple GPS fixes.

The first situation has two possible scenarios. In the first scenario where two consecutive speeds are higher than the average, the position used in both speed calculations is considered erroneous and not used. The velocity is then calculated using the previous good position before and the next good position after the identified erroneous position. In the second scenario where there is only one bad speed higher than the average, because identifying the erroneous position is challenging, that speed is removed from the product.

The second situation has three possible scenarios: (i) the surface interval of the current cycle ($N$) has one position and the surface interval of the next cycle ($N + 1$) has multiple positions, (ii) the surface interval of the current cycle ($N$) has multiple positions and the surface interval of the next cycle ($N + 1$) has one position, or (iii) both surface intervals have multiple positions. To determine which position(s) is erroneous, two distances in degree units are calculated and compared. The first distance is between the last latitude in the current cycle ($N$) and the first latitude in the next cycle ($N + 1$) (cycle difference) and the second difference is between latitudes on the surface interval(s) for which there are multiple positions (surface difference). This surface difference is possible only for cycles with multiple positions in one surface interval. A 0.1° distance threshold is set for drift at the surface based on the fact that a float drifting at the surface for at most one hour with a nominal speed of 2 m $s^{-1}$ would be displaced about 0.1°. A 2° threshold was chosen for maximum drift at parking depth and corresponds to float displacement at a nominal speed of 0.25 m $s^{-1}$ for 10 days. Note that these GPS positions have already been identified as problematic by comparison with the mean of nearby speeds, so this threshold is used to determine which of the positions is erroneous. Using a higher threshold for maximum drift at parking depth (up to 4°) did not significantly change the number of erroneous GPS positions compared to 2°.

Positions marked as erroneous are identified for each scenario presented above. Then cycle differences and surface differences are compared to their respective thresholds, 2° and 0.1°, respectively, and used to determine which position is bad. In cases where either cycle or surface difference is large compared to the threshold, the corresponding bad position is removed from the product. For cases where both cycle and surface differences are larger than thresholds, the corresponding speed is removed from the product. This resulted in about 5500 positions being identified as bad. The reader is referred to the GPS QC method document on the Scripps Argo trajectory-based velocity product DOI web page (Zilberman et al. 2022) for an in-depth description of how to assess erroneous GPS positions.

For floats with communication Système Argos, the surface locations and times are first sorted to be in chronological order and then further quality controlled to remove bad fixes using a method developed by Japan Agency for Marine-Earth Science and Technology (JAMSTEC) and documented by Nakamura et al. (2008). The JAMSTEC method is an optional real-time quality control test for Argos floats described in the Argo Quality Control Manual for CTD and trajectory data (Wong et al. 2021), which is applied here to all Argos files. Once the surface positions and times are quality controlled, the locations of surfacing and sinking are extrapolated by assuming that the surface currents have both linear velocity and inertial velocity components as defined in Park et al. (2004). Positions with bad quality control flags (i.e., not equal to "1") are not used in the extrapolation. The Park et al. (2005) method involves a least squares fit for each surface interval using all the positions available and weighted by the accuracy of those fixes. The least squares fit is then extrapolated to the times of surfacing and sinking in order to estimate the positions corresponding to the float arriving at the surface and leaving the surface. If a surface interval has less than six position fixes, the extrapolation cannot be achieved and only a transmitted velocity, based on transmitted times and locations, is calculated. For these cycles, the first surface position is assigned as the rise location and the last surface position is assigned as the fall location, as is done in the YoMaHa07 and ANDRO products. The "EXTRAP_FLAG" variable of the Scripps product indicates if the velocity was calculated using extrapolated positions and times (value equals to 1) or using transmitted positions and times (value equals to 0).

Some Argo float models, including PROVOR, ARVOR, SOLO-II, Apex APF9 built after July 2014, Navis, and Apex APF11, send back their ascent end time (AET) and descent start time (DST). SOLO floats (prior to the SOLO-II model) do not send back this timing information but have a programmed period of time at the surface that is constant. For these SOLO floats, the time elapsed from the first to last message is subtracted from the total surface time. The remaining time is split in half and one-half is subtracted from the first message time to determine the AET. The other half is added to the last message time to determine the DST. For old APEX floats (APF8 and APF9 prior to July 2014), the AET corresponds to the JULD_ASCENT_END_TIME in the trajectory file. If that variable contains a fill value, then the JULD_TRANSMISSION_START_TIME in the trajectory file that is approximately 10 min after the ascent end time, is used instead. The DST is determined based on last message times (LMT), following the procedure from Park et al. (2005) that is also described in appendix B of the DAC Trajectory Cookbook (Scanderbeg et al. 2019). Each old APEX float's surface times, including AET and estimated DST, are visually examined; 76 old APEX floats were rejected from the final Scripps product due to anomalous AET values or other file format errors.

#### 2) FLOAT CYCLE TIMES

While most Argo floats have a nominal cycle time of 10 days, some floats are programmed to profile at higher sampling frequency during part of their mission in order to capture short-

lived events such as hurricanes and biogeochemical processes. Other floats, when unable to report data in the case of bad satellite transmission at the surface, grounding, or ice cover at the surface, report positions at longer intervals than 10 days. Following Gray and Riser (2014), the Scripps product only uses subsurface drifts of 5–25 days to calculate geostrophic velocity estimates at parking depth. This time window is chosen to eliminate noise from smaller-scale fluctuations and to optimize sampling of the large-scale signal. Most velocities (99%) are derived from subsurface drifts of 5–11 days. A total of 332 643 velocities derived from float cycles outside of the 5–25-day window were excluded from the analysis, leaving 1 456 876 velocities inside the target time window. This includes 322 320 velocity estimates from cycles shorter than 5 days, and 10 323 velocities from cycles longer than 25 days. About 25% of the cycles shorter than 5 days were between a few hours and one day, which is a result of many floats doing an initial short cycle prior to their regular 10-day cycle. Another 58% were between 4 and 5 days, which corresponds to about 1800 floats with shorter programmed cycle times. Cycles that last longer than 25 days usually result from missing positions in between regular 10-day cycles. The positions may be missing due to poor transmission or because the positions were marked as bad.

Accurate cycle times are especially important for APEX Argos floats to estimate the DST [see section 2a(1)]. To check if the cycle time is accurate, the cycle number is plotted for each float as a function of all surface times relative to the first cycle. If the cycle time is accurate and there is no clock drift, all surface times should be lined up as seen in Fig. 7 in appendix B in the Argo DAC Trajectory Cookbook (Scanderbeg et al. 2019). When the cycle time is incorrect or the float's clock is severely drifting, the surface times do not systematically line up, as seen in Fig. 13 in appendix B (Scanderbeg et al. 2019). In the situation where the cycle time is wrong in the metadata file, a range of new cycle times is tested based on cycle times, deployment times, and locations from other similar float types. In general, a more accurate cycle time can be visually identified from plotting cycle numbers as a function of surface times relative to the first cycle. In this analysis, cycle time was adjusted for about 1400 APEX Argos floats.

Cycle times showing large differences from one cycle to the next usually occur when the wrong cycle number is applied, causing too much time to elapse between cycles. In this situation, the float cycles before and after the jump must be processed separately to determine the descent start times as described in section 2a(1).

#### 3) Drift pressures and temperatures

Pressure and temperature measurements are taken at different times during drift, when most floats are parked at 1000 m depth, depending on the float model. The pressure and temperature measurements during drift are identified by the "MEASUREMENT_CODE" variable in the Argo trajectory file. In general, the older floats using Système Argos return only average measurements due to data transmission limitations. APEX Argos floats send the average pressure and temperature during drift, while SOLO floats send two averages: one for the first half of drift and one for the second half of drift. Most of the newer float types that use higher bandwidth communications send back a series of measurements taken during drift. Here, drift pressure and temperature are determined for each cycle based on the information available in the trajectory file. When selecting pressure and temperature measurements in the trajectory file, adjusted values are used if available, and quality control flags are applied, according to the Argo User's Manual recommendations (Argo Data Management Team 2021). If several pressure measurements are available during drift, they are averaged. If no pressure is available during drift, that cycle is removed from the Scripps dataset. The pressure measurements are subject only to the basic quality control of excluding values that are negative, values larger than programmed maximum profiling depth (e.g., >2200 dbar for 2000 dbar floats, >4200 dbar for 4000 dbar Deep Argo floats, and >6200 dbar for 6000 dbar Deep Argo floats), and values shallower than 100 dbar.

#### 4) Velocity calculations

Two different types of velocities were calculated based on quality-controlled float locations and timing. The first was based on the extrapolated positions and times for Argos float cycles with at least six surface fixes. The second was based on the transmitted positions and times for Argos float cycles with less than six surface fixes, and for Iridium floats. Extrapolated velocity is calculated using the difference between the extrapolated surface descent position of a cycle and the extrapolated surface arrival position of the next cycle, divided by the difference in seconds between the extrapolated surface descent time and the extrapolated surface arrival time in the next cycle. There are 858 062 extrapolated velocities calculated in the Scripps Argo velocity dataset. Transmitted velocity is calculated in a similar manner to YoMaHa07 and ANDRO, using the difference between the final transmitted surface position of a cycle and the first surface transmitted position of the next cycle, divided by the difference in seconds between the final transmitted surface time and the first transmitted surface time in the next cycle. There are 1 456 876 transmitted velocities calculated in the Scripps Argo velocity dataset.

Two types of nongridded velocity estimate NetCDF files are made available in the Scripps product: one including a drift pressure range between 100 and 6200 dbar, and the other including only velocities in the targeted drifting pressure range of 1000 ± 200 dbar. Both of these files include a flag that indicates whether an extrapolated velocity is available (EXTRAP_FLAG = 1) or if only the transmitted velocity was calculated (EXTRAP_FLAG = 0). Results presented in section 3 are based on velocity estimates between 800 and 1200 dbar.

#### 5) Computation of the gridded velocity field

A time-mean gridded velocity estimate was computed at 1° horizontal resolution, using all velocities within the 1000 ± 200 dbar pressure range. Extrapolated velocities were included when possible; otherwise, transmitted velocities were used. To determine the time-mean gridded estimate, a

multiscale, local-window mapping method was used, based on Gray and Riser (2015) and Kuusela and Stein (2018). For each grid point, all data within a 250 km radius were first used to compute a large-scale, time-invariant velocity estimate via a linear least squares fit to second-order polynomials that varied with latitude and longitude. The resulting large-scale velocity field was then interpolated to the locations of the individual velocity estimates and subtracted from those data to produce small-scale velocity anomalies. At each grid point, all anomalies within a 500 km radius were used to compute a small-scale, time-varying velocity estimate at 10-day frequency, via Gaussian process regression. Total time-mean velocity fields were next produced by summing the temporally constant large-scale gridded field with the temporal average of the small-scale gridded field. Finally, using the time-mean velocity estimates computed from the sum of the large-scale and small-scale signals, a geostrophic streamfunction was calculated for the global domain following the method outlined in Li et al. (2006).

The gridding technique used in this study, also known as objective mapping or kriging as well as Gaussian process regression, requires modeling the autocovariance in the field of interest. The small-scale velocity field being mapped here is assumed to be in geostrophic balance, as in Davis (2005) and Gray and Riser (2014). The zonal and meridional components of the horizontal velocity, $\mathbf{u} = (u, v, 0)$, are thus related through a geostrophic streamfunction $\Psi$:

$$\mathbf{u} = -\frac{1}{f}\hat{\mathbf{k}} \times \nabla\Psi,$$

where $f$ is the Coriolis parameter and $\hat{\mathbf{k}}$ is the unit vector in the vertical direction. Using this relationship, the auto- and cross covariances of the zonal and meridional components of the velocity field can be calculated from derivatives of the autocovariance of the underlying geostrophic streamfunction (Bretherton et al. 1976). Following Park et al. (2022), an anisotropic spatiotemporal Matern function of order 3/2 was used to model the autocovariance of $\Psi$, and for each grid point maximum likelihood estimation was used to determine, directly from the observations, the signal and noise variances as well as the decorrelation scales corresponding to zonal, meridional, and temporal distances. Because the assumption of geostrophy breaks down as $f$ goes to zero, gridded velocities are not computed at the equator.

One of the advantages of gridding data with Gaussian process regression is that this technique produces a quantitative uncertainty value associated with each individual estimate (Bretherton et al. 1976). The value of the uncertainty at any particular grid point depends on a combination of the distribution of the nearby observations in space and time, the amount of variance in both the signal and noise at that point, and the spatiotemporal decorrelation scales. For the time-mean gridded velocity field, total (squared) mapping errors were calculated as the mean of all the individual squared error estimates for the time-varying field divided by the effective degrees of freedom, which varied spatially. The quadratic formula of Bretherton et al. (1999) was used to compute the effective degrees of freedom from the temporal decorrelation of the autocovariance function, determined at each grid point as described above.

### b. Characteristics of the reanalysis product

The ocean reanalysis used in this study is the 2020 version of the Bluelink Reanalysis (BRAN) called BRAN2020 (Chamberlain et al. 2021). This product is near global (65°S–65°N), eddy resolving, and spans January 1993 to December 2020. BRAN2020 uses a multiscale data-assimilation approach that constrains sequentially the broad scales and mesoscale structure of the ocean state. The reanalysis assimilates conventional observation platforms (e.g., Argo, XBT, satellite altimetry, and satellite SST) used in prior versions of BRAN that are combined with updated sources including observations from marine mammals, moorings, and shipborne surveys. Here, we use the monthly mean velocity fields between January 2001 and December 2020 for consistency with the Scripps trajectory product.

## 3. Results and discussion

The global large-scale circulation at 1000 ± 200 dbar derived from the Scripps dataset of trajectory-based velocities is first presented, followed by a comparison of the Scripps product with the ANDRO and the YoMaHa07 datasets. The impact of using extrapolated instead of transmitted positions for floats with telecommunication Système Argos is next described. The time-mean velocity field based on the Scripps dataset is then compared to velocities from the BRAN2020 reanalysis.

### a. Circulation at parking depth based on the Scripps product

Maps of gridded absolute velocities, computed at 1° horizontal resolution from the 1 316 506 trajectory-based velocity estimates at 1000 ± 200 dbar contained in the Scripps product, reveal the spatial structure of the large-scale flow in that depth range (Figs. 1a,d). Outside of the equatorial region, the standard deviations computed from the time-varying, small-scale mapped estimate are comparable for the zonal (Fig. 1b) and meridional (Fig. 1e) components of the velocity field. The spatial patterns in the mapping errors [calculated following the method described in section 2a(5)] are also largely similar, although the magnitudes roughly scale with the size of the velocity in each direction (Figs. 1c,f). The direction and strength of the time-mean geostrophic flow at 1000 ± 200 dbar is illustrated by the geostrophic streamfunction (Fig. 2).

Main features of the large-scale ocean circulation at 1000 ± 200 dbar are apparent in the gridded Scripps velocity fields (Figs. 1 and 2), including the boundary currents and gyres (Wunsch and Heimbach 2013; Qu et al. 2019) and the Antarctic Circumpolar Current (ACC; Rintoul and da Silva 2019). The strongest velocities (>0.02 m $s^{-1}$) are seen in the boundary currents forming the northern and western limbs of the subpolar gyres, in the midlatitude western boundary currents and recirculation zones, and in the ACC. In regions with strong variability (Figs. 1b,e), the mapping errors are large but only represent a limited (<20%) fraction of the signal (Figs. 1c,f).

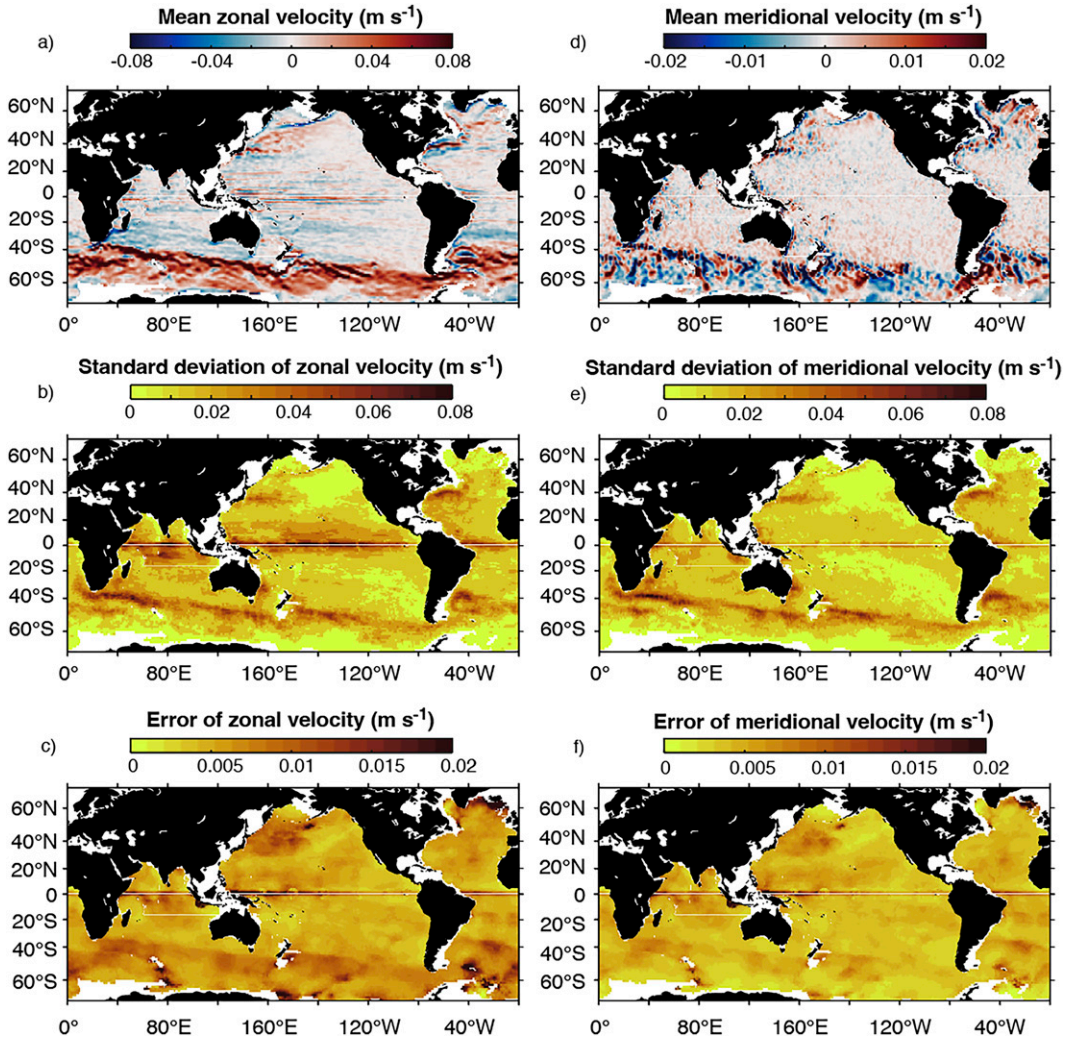


FIG. 1. Maps of (a) zonal velocity, (b) zonal velocity standard deviation, (c) zonal velocity error, (d) meridional velocity, (e) meridional velocity standard deviation, and (f) meridional velocity error. All are from the gridded Scripps Argo dataset between 2001 and 2020 and for 1000 ± 200 dbar. White areas indicate places where gridded velocity estimates were not computed: along the equator, the marginal seas, grid boxes shallower than 1000 m [based on bathymetry from Amante and Eakins (2009)], and ice-covered regions with sparse data.

Increased velocity (>0.01 m $s^{-1}$) error in the northwest Pacific stems from the combination of reduced sampling (Fig. 3) and higher variability (Figs. 1b,e). The largest mapping error values (>0.015 m $s^{-1}$), which are found close to shallow regions particularly in the subpolar North Atlantic basin and localized areas in the Southern Ocean and south of the Aleutian Ridge, are mainly attributable to long decorrelation time scales together with lower float densities and greater variance. In these regions where the mapping errors represent a significant fraction (>50%) of the signal, the gridded trajectory-based velocity estimates should be interpreted with caution.

### b. *Comparison of Scripps transmitted velocities with YoMaHa and ANDRO*

Three main differences in quality control procedures distinguish the Scripps, YoMaHa, and ANDRO products: (i) the YoMaHa product relies only on programmed parking depth from the metadata files, the ANDRO product uses programmed parking depth from the metadata files only when pressure measurements are unavailable in the trajectory file, while Scripps only uses pressure measurements taken during drift in the trajectory files; (ii) the Scripps product uses trajectories from 5- to 25-day cycles and floats with at least three cycles, while the ANDRO and YoMaHa products apply no restrictions to the cycle time and number of cycles per float; and (iii) the ANDRO and YoMaHa products are derived from the entire historical Argo trajectory dataset, while Scripps uses only trajectory data upgraded to the new 2014 format (v3.1) that includes increased documentation of the float location and time for each cycle.

Other differences in the Scripps product compared to ANDRO and YoMaHa products reside in the use of QC position flags and positional QC [described in section 2a(1)], data

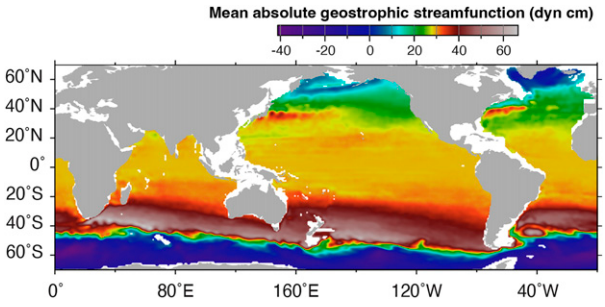


FIG. 2. Map of streamfunction from the gridded Scripps Argo dataset between 2001 and 2020 and for 1000 ± 200 dbar.

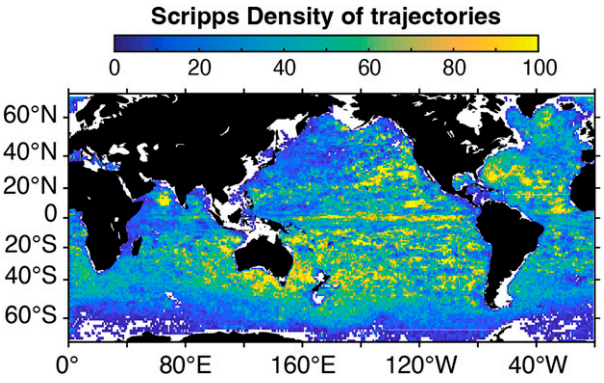


FIG. 3. Map of density of observation per 1° bins from the Scripps Argo dataset between 2001 and 2020 and for 1000 ± 200 dbar.

source, and grounded cycles. The Scripps product takes advantage of available delayed mode trajectory files, which were not available when YoMaHa and ANDRO products were created. This means that quality control flags applied by delayed mode experts were used in the Scripps product but may not have been available or identified for YoMaHa and ANDRO. The JAMSTEC method for real-time quality control of Argos floats developed by Nakamura et al. (2008) is used for the Scripps and ANDRO products but not in YoMaHa. In the ANDRO product, some surface positions that were missing in the trajectory files on the GDACs were recovered by decoding raw data collected from individual DACs that were not publicly available (Ollitrault and Rannou 2013). In contrast, only decoded trajectory data directly available from the GDACs are used in the Scripps and YoMaHa products. Grounding flags are applied in the ANDRO product, but they are not utilized in Scripps or YoMaHa. Overall, YoMaHa has 1 387 930 velocities, Scripps has 1 316 506 velocities, and ANDRO has 911 317 velocities in the 2001–20 time period and in the 800–1200 dbar pressure range.

Since extrapolated trajectories are not provided in the ANDRO and YoMaHa datasets, when comparing these three products, only transmitted trajectories from the Scripps product are considered. Between 2001 and 2020, there are 226 698 velocities from the YoMaHa product, 16% of the dataset, that are not identical to the Scripps product (Table 1). Out of this amount, 34% have cycle times less than 5 days (Fig. 4a); 31% show starting or ending position differences due to additional QC in the Scripps product (Fig. 4b); 22% correspond to trajectories not in the Scripps product because (i) there is no drift pressure in the trajectory file, (ii) the drift pressure is less than 100 dbar or greater than 2200 dbar for 2000 dbar floats or greater than 4200 or 6200 dbar for 4000 and 6000 dbar floats, or (iii) the drift pressure is determined to be bad [as described in section 2a(3); Fig. 4c]; 8% are measured drift pressures outside of the Scripps's targeted 800–1200 dbar range; 2% are floats with less than three cycles; 1% are in the old trajectory format; 1% are cycle times greater than 25 days; and the remaining 1% are likely due to applying delayed-mode QC flags and a combination of the aforementioned causes (cycle length, additional position and pressure QC in the Scripps product). In the YoMaHa product, trajectories with different starting or ending positions with Scripps are distributed pretty homogeneously globally (Fig. 4b). In contrast, trajectories from short (<5 days) cycles are mostly located in the Kuroshio and north Indian Ocean (Fig. 4a), while most trajectories with different parking pressure than the Scripps product are located off the coast of Alaska and in the northwest Atlantic basin (Fig. 4c).

Between 2001 and 2020, there are 221 833 velocities, 24% of the ANDRO product, that are not identical to Scripps (Table 1). Out of this amount, 45% are floats in the old, pre-2014 trajectory format (Fig. 5a); 24% show differences in starting and ending positions from additional QC in the Scripps product (Fig. 5b); 23% have cycle times shorter than 5 days (Fig. 5c); 3% are floats with less than three cycles; 3% are anomalous trajectories not in the Scripps product because (i) there is no drift pressure in the trajectory file, (ii) the drift pressure is less than 100 dbar or greater than 2200 dbar for 2000 dbar floats, or greater than 4200 dbar or 6200 dbar for 4000 and 6000 dbar floats, or (iii) the drift pressure is determined to be bad [as described in section 2a(3)]; and 3% are either recovered raw data only available in the ANDRO product, floats with cycles longer than 25 days, floats with measured parking pressure outside of Scripps's targeted 800–1200 dbar range (cycles outside of the 800–1200 dbar range are provided in the 100–6200 dbar Scripps product, see section 2a), or newly available delayed mode data. High densities of ANDRO trajectories in the old format are located in the northwest Pacific, east Indian, and northeast Atlantic Oceans (Fig. 5a). As in the YoMaHa product, ANDRO trajectories with different starting or ending positions than Scripps are mainly homogeneously distributed (Figs. 4b and 5b), and ANDRO trajectories from short (<5 days) cycles are mostly located in the Kuroshio region and in the north Indian Ocean (Figs. 4a and 5c).

### c. *Comparison between transmitted and extrapolated fields from Scripps*

Argo floats with telecommunication Système Argos may drift at the surface for up to 4 h between arrival at the surface and first position fix, and between last position fix and start of descent. In energetic current and eddy fields, untracked changes in Argos float positioning at the surface prior to first transmission to satellite and after last transmission may have

TABLE 1. Number of velocity data from the ANDRO and YoMaHa products that are not included in the Scripps product. Causes of discrepancies between the Scripps and ANDRO and YoMaHa products are indicated.

| | ANDRO velocities not in Scripps | YoMaHa velocities not in Scripps |
|---|---|---|
| Cycles < 5 days | 50 101 | 77 679 |
| Cycles > 25 days | 583 | 1936 |
| No pressure in trajectory file or bad pressure as identified by Scripps | 5575 | 49 518 |
| Pressure outside of the 800–1200 dbar range of the Scripps product | 1899 | 17 835 |
| Old trajectory format | 99 329 | 1473 |
| Different starting/ending positions due to additional Scripps QC | 54 184 | 69 882 |
| Less than three cycles per float | 5577 | 4521 |
| Applied DMQC flags or combination of other causes | 4585 | 3854 |
| Total number | 221 833 | 226 698 |
| Percentage of dataset | 24% | 16% |

significant impact on subsurface velocity estimates based on trajectories. This is less of an issue for two-way communication system floats that, by design, spend only 15–30 min at the surface and for which trajectories are estimated using positions and times at first and last transmission. A method for estimating surfacing and diving positions and times for Argos cycles in the Scripps product is described in section 2a(1). In strong currents where the flow does not reverse with depth, the extrapolated distance between surfacing and diving is expected to be smaller than the transmitted distance, leading to slower extrapolated velocity estimates. In contrast, if the mean flow reverses between the surface and parking depth, the extrapolated distance between surfacing and diving would be larger than the transmitted distance, resulting in higher extrapolated velocities. Of course, in reality, complex three-dimensional flows that are highly variable in space and time mean that either case is possible.

A pairwise comparison of all velocity estimates computed using extrapolated positions with the corresponding estimates computed using transmitted positions reveals a mean absolute difference in magnitude of $<0.002$ m s$^{-1}$, which is statistically significant ($p < 0.01$) but still well within any reasonable uncertainty range for an individual trajectory-based velocity estimate. However, for more than 1.3% of the velocities considered (amounting to >10 000 data points), the difference between the transmitted and extrapolated speeds was greater than one-half of their mean. In this group of velocity estimates, extrapolated speeds were, on average, greater than their transmitted counterparts, with a mean absolute difference of 0.0125 m s$^{-1}$.

To further assess the impact of using extrapolated positions as compared to transmitted positions on the gridded velocities, time-averaged gridded fields were calculated for the period 2001–10 using transmitted velocities only and compared

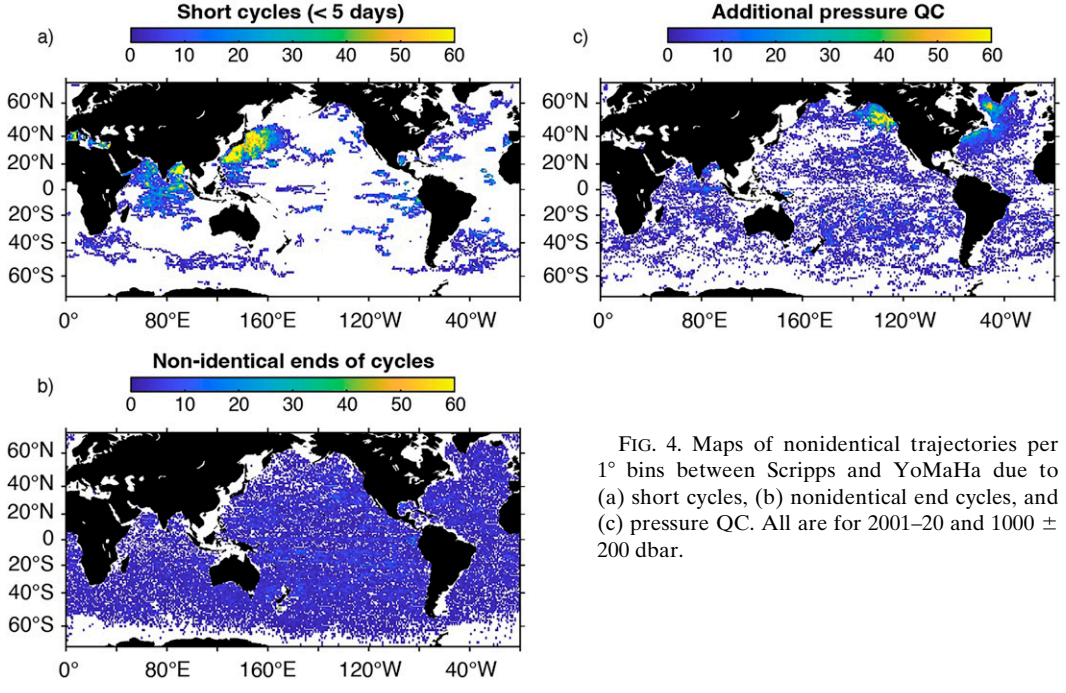


FIG. 4. Maps of nonidentical trajectories per 1° bins between Scripps and YoMaHa due to (a) short cycles, (b) nonidentical end cycles, and (c) pressure QC. All are for 2001–20 and 1000 ± 200 dbar.

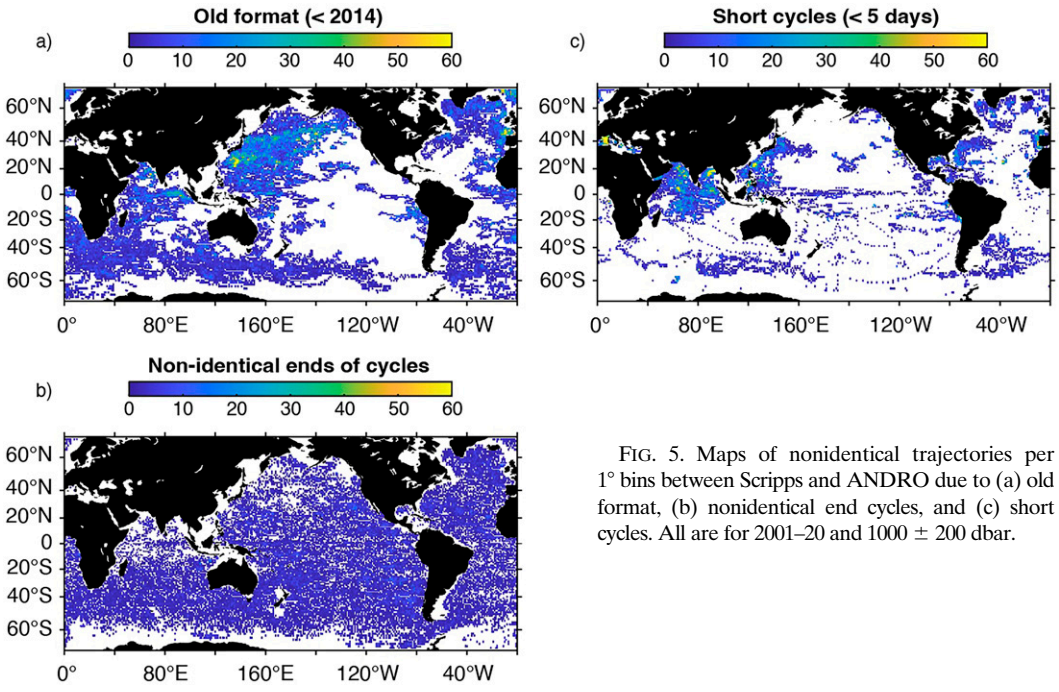


FIG. 5. Maps of nonidentical trajectories per 1° bins between Scripps and ANDRO due to (a) old format, (b) nonidentical end cycles, and (c) short cycles. All are for 2001–20 and 1000 ± 200 dbar.

to the same fields computed using extrapolated velocities (Fig. 6). During this time period, which is prior to the switch to two-way communication systems, 90% of all surface intervals considered have at least six position fixes, such that extrapolation can be achieved. Differences between extrapolated and transmitted time-mean gridded speeds are highest ($>0.004$ m s$^{-1}$) in the North Pacific western boundary current and associated recirculation region, as well as portions of the ACC, although the differences are significantly smaller than the associated mapping errors (Fig. 6). Thus, while individual velocity estimates can be substantially changed when the displacement of the float at the surface is taken into account, the impact on the 10-year-average gridded velocity fields appears minimal.

### d. *Comparison with ocean predictions from BRAN2020 reanalysis*

One of the intended applications of the Scripps velocity dataset that we present here is the assessment of ocean models, ocean reanalyses, and ocean forecasts. Additionally, the modeling community may exploit the Scripps velocity dataset to help constrain their models. To demonstrate the sort of insights the modeling community might gain from this comparison, we include here a comparison between the reanalyzed velocities from BRAN2020 and observed velocities from the Scripps product (Figs. 7a,b). The errors of reanalyzed BRAN2020 temperature and salinity between 800 and 1200 m depth are reasonably well understood. Chamberlain et al. (2021) report that between 2000 and 2020, the reanalyzed temperature and salinity fields are unbiased, with errors of about 0.2°C and 0.04 PSS-78 (their Figs. 7a and 8a). These errors are small compared to other similar reanalyses (e.g., Lellouche et al. 2021, their Fig. 1). No comprehensive assessment of the velocities from BRAN2020 has been performed.

For consistency between the gridded Scripps velocities and BRAN2020, the BRAN2020 velocities are averaged over 1° × 1° bins and set to zero in shallow (<2000 m) regions. The BRAN2020 speed estimates are also averaged between 775 and 1238 dbar, which correspond to the BRAN2020 pressure levels nearest to the float parking pressure range of 800–1200 dbar used in the Scripps product. In this comparison with BRAN2020, velocities used in the Scripps product include extrapolated velocities for Argos float cycles with at least six good surface fixes, and transmitted velocities when extrapolation cannot be achieved and for floats with two-way communication systems.

The comparisons between Scripps and BRAN2020 velocities around 1000 m depth show some systematic differences that are statistically significant (Figs. 7c,d). We attribute these differences to limitations of the data-assimilating model, particularly to the projection of assimilated data onto the model's barotropic mode, and to the horizontal displacement of resolved oceanographic features. The projection onto the barotropic mode is influenced by the model's vertical resolution (Stewart et al. 2017) and by the projection of assimilated observations onto unobserved variables (Oke et al. 2013), namely, horizontal velocities. The excessive projection of observations onto the barotropic mode is largely due to the dynamical imbalance of increments that are added to the model at each data assimilation cycle (e.g., Oke et al. 2007, 2008). Errors associated with the projection of assimilated data onto the model's barotropic mode may generate systematic differences seen in the East Australian Current, the central equatorial Pacific and Indian Oceans, and the equatorial Atlantic Ocean (Fig. 7c). The horizontal displacement of strong, narrow oceanographic features, such as the ACC in the Southern Ocean and boundary currents in the North Atlantic and along

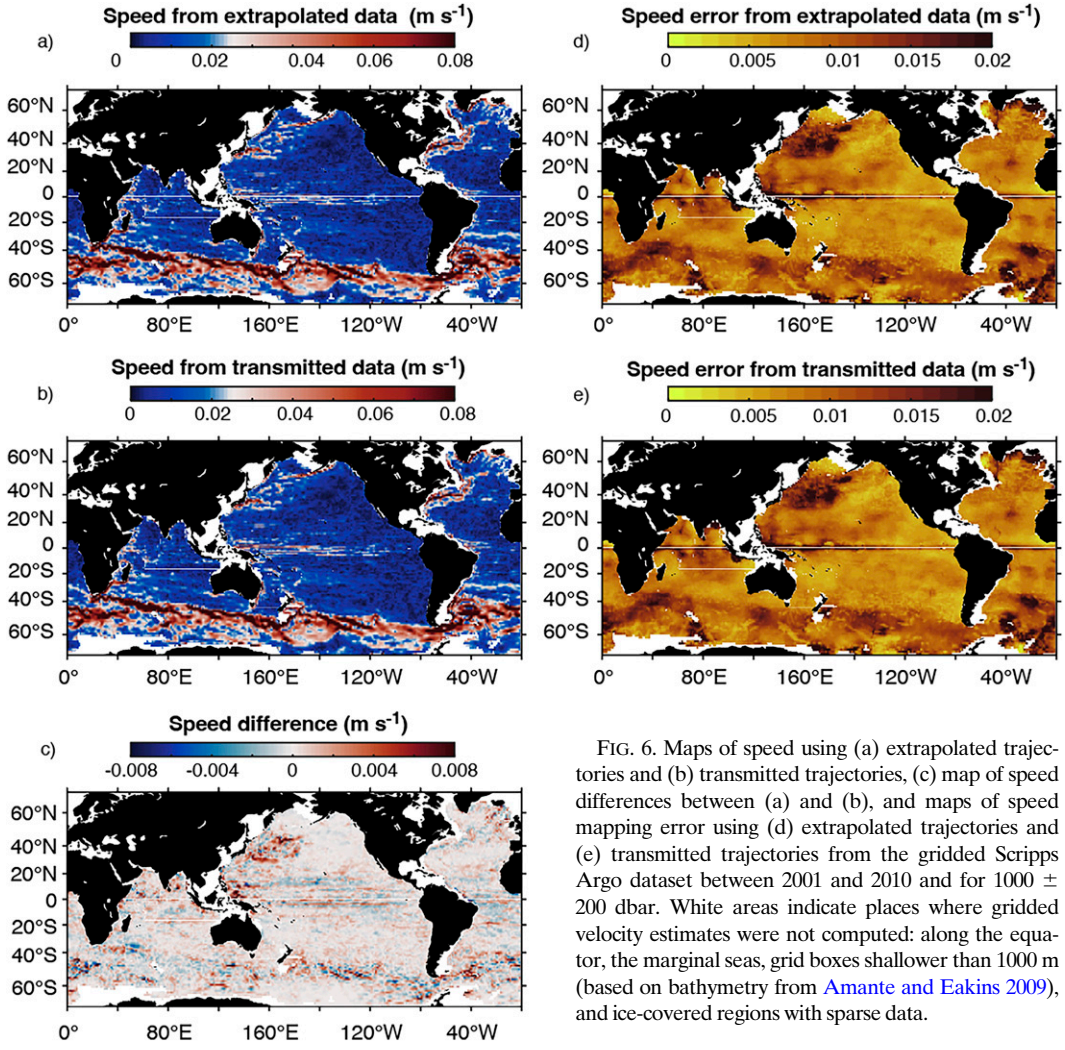


FIG. 6. Maps of speed using (a) extrapolated trajectories and (b) transmitted trajectories, (c) map of speed differences between (a) and (b), and maps of speed mapping error using (d) extrapolated trajectories and (e) transmitted trajectories from the gridded Scripps Argo dataset between 2001 and 2010 and for 1000 ± 200 dbar. White areas indicate places where gridded velocity estimates were not computed: along the equator, the marginal seas, grid boxes shallower than 1000 m (based on bathymetry from Amante and Eakins 2009), and ice-covered regions with sparse data.

the Aleutian Ridge, appears as alternating bands of strong positive and negative differences between Scripps and BRAN2020 (Fig. 7c). For example, along the path of the ACC there are many regions of large positive and negative differences that are quasi zonal. We attribute these differences to errors in the latitude of the ACC jets and filaments; if a BRAN2020 jet is too far to the north, there will be a negative quasi-zonal difference to the north, adjacent to a positive quasi-zonal difference to the south, and vice versa. These sort of systematic errors in ocean reanalyses are most likely due to errors in the assumed mean sea level field that is used to convert modeled sea level to sea level anomaly for assimilation, as discussed by Oke et al. (2013).

## 4. Summary and conclusions

A new subsurface velocity dataset based on Core, BGC, and Deep Argo float trajectories collected between 2001 and 2020 called the Scripps Argo trajectory-based velocity product was made freely available. Comparisons were made between the Scripps product and existing YoMaHa and ANDRO datasets for concurrent time periods. While the transmitted trajectories from the Scripps dataset are largely consistent with those from YoMaHa and ANDRO, differences arise mainly due to selected trajectory format, cycle time, and quality control of pressure at drift. Between 2001 and 2020, 16% of the YoMaHa and 24% of the ANDRO transmitted trajectories are different than the Scripps product. In addition to velocities computed from transmitted trajectories, extrapolated velocity estimates are provided in the Scripps product for float cycles using Système Argos with at least six surface fixes. Differences between transmitted and extrapolated velocity estimates are typically negligible, although significant deviations are found for a small fraction of the dataset.

Velocity fields based on Argo float drift at parking depth are rarely used for data assimilation, in contrast to Argo temperature and salinity profiles. A number of studies have recently emerged that describe encouraging progress on the use of Argo velocity data for increasing assimilative model and ocean reanalysis performance. For instance, adding subsurface velocity may

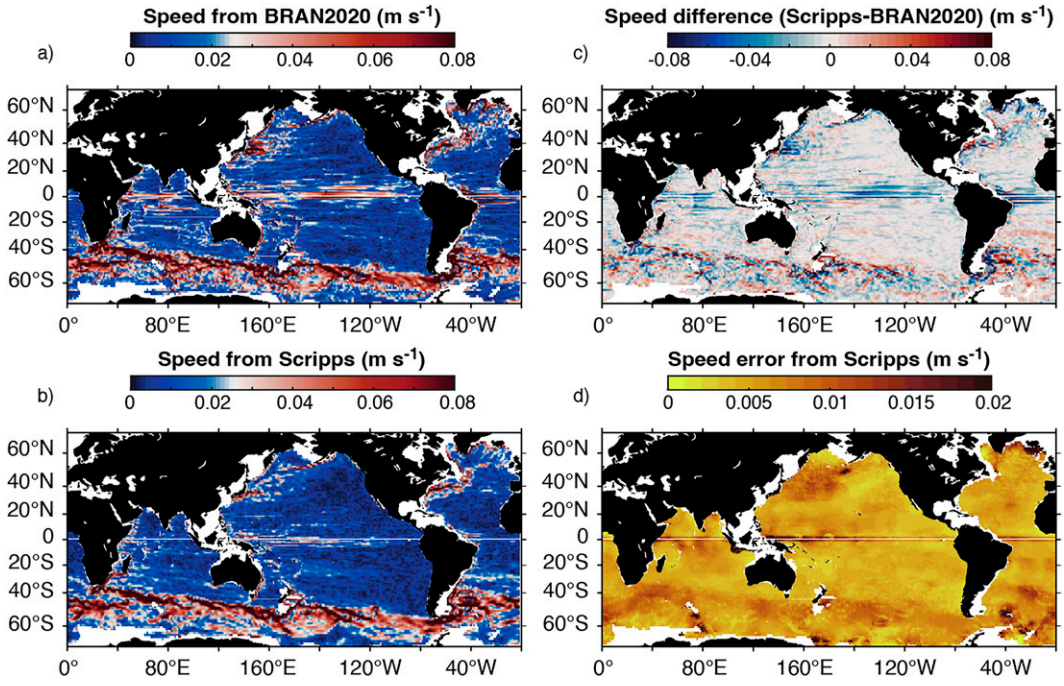


FIG. 7. Maps of (a) speed from BRAN2020 between 2001 and 2020 and for (775–1238 dbar), (b) speed from the gridded Scripps Argo dataset between 2001 and 2020 and for 1000 ± 200 dbar, (c) difference between (a) and (b), and (d) mapping error associated with speed from the gridded Scripps Argo dataset between 2001 and 2020 and for 1000 ± 200 dbar. White areas indicate places where gridded velocity estimates were not computed: along the equator, the marginal seas, grid boxes shallower than 1000 m (based on bathymetry from Amante and Eakins 2009), and ice-covered regions with sparse data.

improve velocity estimates from the Ocean Data Synthesis System in several regions including the North Pacific and North Atlantic oceans and Drake Passage (Masuda et al. 2014). Assimilation of Argo velocity in the Regional Ocean Model System contributes to adjusting the unbalanced component in the velocity increments and therefore improves the predicted velocity fields in the western South China Sea (Wang et al. 2020). Near-real-time Argo trajectory data would constitute a useful addition to observational reference datasets commonly used for intercomparison of forecasts and reanalyses, such as satellite observations of sea surface temperature and sea level anomaly, and subsurface in situ temperature and salinity measurements taken from Argo (Divakaran et al. 2015; Ryan et al. 2015; Oke et al. 2012). Some forecast and reanalysis systems impose a balance constraint when model fields are updated (e.g., King et al. 2018), while others do not (e.g., Chamberlain et al. 2021). Discrepancies between Argo float trajectory-based subsurface velocities from the Scripps product and the BRAN2020 reanalysis expose a drawback of not imposing any constraint.

We anticipate that the newly available Scripps trajectory dataset introduced here will permit modeling groups to better assess their systems, identify systematic errors, and ultimately improve the quality of their reanalyses and forecasts. The Scripps trajectory dataset provides the oceanographic community with a reliable and complete observational picture of the ocean circulation at intermediate depths. Envisioned applications range from defining connectivity between ocean basins to fundamental studies of the mean and variability of the ocean's large-scale circulation.

*Acknowledgments.* NZ and MS were supported by the NOAA Global Ocean Monitoring and Observing Program through Award NA20OAR4320278. NZ was supported by NOPP (NOAA Grant NA18OAR0110434). ARG received support from NOAA through Award NA20OAR4320271, from NASA through Grant 80NSSC19K1252, and from NSF through Grant OCE-1946578. We would like to acknowledge high-performance computing support from Cheyenne (doi: 10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. PO was supported by Australia's Integrated Marine Observing System (IMOS)—IMOS is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS). The authors thank John Gilson for his assistance with Argo quality control, Lisa Lehmann for her contribution to data processing, and Donata Giglio and William Mills for including the Scripps Institution of Oceanography Argo velocity dataset on the web app and database Argovis (https://argovis.colorado.edu), designing a new schema and application programming interface for the data, and creating a demo notebook to access the product and visualize the data (https://github.com/argovis/demo_notebooks). Donata Giglio and William Mills were supported by NSF Awards 1928305 and 2026954.

*Data availability statement.* The Scripps Institution of Oceanography Argo velocity dataset is freely available on the UCSD library (https://doi.org/10.6075/J0KD1Z35), Argo website (https://argo.ucsd.edu/data/argo-data-products/velocity-products/), and the Argovis web app and database [https://argovis.colorado.edu, recently upgraded from the app version described in Tucker et al. (2020)]. A demo notebook leveraging the new Argovis API to access and visualize the product is available at https://github.com/argovis/demo_notebooks.

## REFERENCES


Amante, C., and B. W. Eakins, 2009: ETOPO1 1 arc-minute global relief model: Procedures, data sources and analysis. NOAA Tech. Memo. NESDIS NGDC-24, 25 pp., https://doi.org/10.7289/V5C8276M.

Argo Data Management Team, 2019: Argo user's manual, version 3.3. IFREMER Doc., 109 pp., https://doi.org/10.13155/29825.

——, 2021: Argo user's manual v.3.41.1. Argo Normative Doc., 111 pp., https://doi.org/10.13155/29825.

Beal, L. M., and H. L. Bryden, 1999: The velocity and vorticity structure of the Agulhas Current at 32°S. *J. Geophys. Res.*, **104**, 5151–5176, https://doi.org/10.1029/1998JC900056.

Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, **12**, 1990–2009, https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2.

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

Chamberlain, M. A., P. R. Oke, R. A. S. Fiedler, H. M. Beggs, G. B. Brassington, and P. Divakaran, 2021: Next generation of Bluelink ocean reanalysis with multiscale data assimilation: BRAN2020. *Earth Syst. Sci. Data*, **13**, 5663–5688, https://doi.org/10.5194/essd-13-5663-2021.

Chamberlain, P. M., L. D. Talley, M. R. Mazloff, S. C. Riser, K. Speer, A. R. Gray, and A. Schwartzman, 2018: Observing the ice-covered Weddell Gyre with profiling floats: Position uncertainties and correlation statistics. *J. Geophys. Res. Oceans*, **123**, 8383–8410, https://doi.org/10.1029/2017JC012990.

Davis, R. E., 2005: Intermediate-depth circulation of the Indian and South Pacific Oceans measured by autonomous floats. *J. Phys. Oceanogr.*, **35**, 683–707, https://doi.org/10.1175/JPO2702.1.

——, D. C. Webb, L. A. Regier, and J. Dufour, 1992: The Autonomous Lagrangian Circulation Explorer (ALACE). *J. Atmos. Oceanic Technol.*, **9**, 264–285, https://doi.org/10.1175/1520-0426(1992)009<0264:TALCE>2.0.CO;2.

——, J. T. Sherman, and J. Dufour, 2001: Profiling ALACEs and other advances in autonomous subsurface floats. *J. Atmos. Oceanic Technol.*, **18**, 982–993, https://doi.org/10.1175/1520-0426(2001)018<0982:PAAOAI>2.0.CO;2.

——, and Coauthors, 2019: 100 years of progress in ocean observing systems. *A Century of Progress in Atmospheric and Related Sciences: Celebrating the American Meteorological Society Centennial*, *Meteor. Monogr.*, No. 59, Amer. Meteor. Soc., https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0014.1.

Divakaran, P., and Coauthors, 2015: GODAE Oceanview intercomparison for the Australian region. *J. Oper. Oceanogr.*, **8** (Supp1.), s112–s126, https://doi.org/10.1080/1755876X.2015.1022333.

Frajka-Williams, E., M. Lankhorst, J. Koelling, and U. Send, 2018: Coherent circulation changes in the deep North Atlantic from 16°N and 26°N transport arrays. *J. Geophys. Res. Oceans*, **123**, 3427–3443, https://doi.org/10.1029/2018JC013949.

——, and Coauthors, 2019: Atlantic meridional overturning circulation: Observed transport and variability. *Front. Mar. Sci.*, **6**, 260, https://doi.org/10.3389/fmars.2019.00260.

Fu, Y., F. Li, J. Karstensen, and C. Wang, 2020: A stable Atlantic meridional overturning circulation in a changing North Atlantic Ocean since the 1990s. *Sci. Adv.*, **6**, eabc7836, https://doi.org/10.1126/sciadv.abc7836.

Gill, A. E., 1982: *Atmosphere–Ocean Dynamics*. Academic Press, 662 pp.

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.

——, and ——, 2015: A method for multiscale optimal analysis with application to Argo data. *J. Geophys. Res. Oceans*, **120**, 4340–4356, https://doi.org/10.1002/2014JC010208.

Jayne, S. R., D. H. Roemmich, N. V. Zilberman, S. C. Riser, K. S. Johnson, G. C. Johnson, and S. R. Piotrowicz, 2017: The Argo program: Present and future. *Oceanography*, **30**(2), 18–28, https://doi.org/10.5670/oceanog.2017.213.

Katsumata, K., and H. Yoshinari, 2010: Uncertainties in global mapping of Argo drift data at the parking level. *J. Oceanogr.*, **66**, 553–569, https://doi.org/10.1007/s10872-010-0046-4.

King, R. R., J. While, M. J. Martin, D. J. Lea, B. Lemieux-Dudon, J. Waters, and E. O'Dea, 2018: Improving the initialization of the Met Office operational shelf-seas model. *Ocean Modell.*, **130**, 1–14, https://doi.org/10.1016/j.ocemod.2018.07.004.

Kuusela, M., and M. L. Stein, 2018: Locally stationary spatio-temporal interpolation of Argo profiling float data. *Proc. Roy. Soc.*, **474A**, 20180400, https://doi.org/10.1098/rspa.2018.0400.

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., https://doi.org/10.13140/RG.2.2.12820.71041.

Lellouche, J.-M., and Coauthors, 2021: The Copernicus global 1/12° oceanic and sea ice GLORYS12 reanalysis. *Front. Earth Sci.*, **9**, 698876, https://doi.org/10.3389/feart.2021.698876.

Li, Z., Y. Chao, and J. C. McWilliams, 2006: Computation of the streamfunction and velocity potential for limited and irregular domains. *Mon. Wea. Rev.*, **134**, 3384–3394, https://doi.org/10.1175/MWR3249.1.

Masuda, S., N. Sugiura, S. Osafune, and T. Doi, 2014: Improvement of ocean state estimation by assimilating mapped Argo drift data. *Sci. World J.*, **2014**, 975618, https://doi.org/10.1155/2014/975618.

Nakamura, T., N. Ogita, and T. Kobayashi, 2008: Quality control method of Argo float position data. *JAMSTEC Rep. Res. Dev.*, **7**, 11–18, https://doi.org/10.5918/jamstecr.7.11.

Oke, P. R., P. Sakov, and S. P. Corney, 2007: Impacts of localization in the EnKF and EnOI: Experiments with a small model. *Ocean Dyn.*, **57**, 32–45, https://doi.org/10.1007/s10236-006-0088-8.

——, G. B. Brassington, D. A. Griffin, and A. Schiller, 2008: The Bluelink Ocean Data Assimilation System (BODAS). *Ocean Modell.*, **21**, 46–70, https://doi.org/10.1016/j.ocemod.2007.11.002.

——, ——, J. Cummings, M. Martin, and F. Hernandez, 2012: GODAE inter-comparisons in the Tasman and Coral Seas. *J. Oper. Oceanogr.*, **5**, 11–24, https://doi.org/10.1080/1755876X.2012.11020135.

——, and Coauthors, 2013: Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3. *Ocean Modell.*, **67**, 52–70, https://doi.org/10.1016/j.ocemod.2013.03.008.

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.

——, and A. C. De Verdiere, 2014: The ocean general circulation near 1000-m depth. *J. Phys. Oceanogr.*, **44**, 384–409, https://doi.org/10.1175/JPO-D-13-030.1.

——, P. Rannou, E. Brion, C. Cabanes, A. Piron, G. Reverdin, and N. Kolodziejczyk, 2020: ANDRO: An Argo-based deep displacement dataset. SEANOE, accessed 4 March 2022, https://doi.org/10.17882/47077.

Park, B., M. Kuusela, D. Giglio, and A. Gray, 2022: Spatio-temporal local interpolation of global ocean heat transport using Argo floats: A debiased latent Gaussian process approach. *Ann. Appl. Stat.*, in press.

Park, J. J., K. Kim, and W. R. Crawford, 2004: Inertial currents estimated from surface trajectories of Argo floats. *Geophys. Res. Lett.*, **31**, L13307, https://doi.org/10.1029/2004GL020191.

——, ——, 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.

Qu, T., I. Fukumori, and R. A. Fine, 2019: Spin-up of the Southern Hemisphere super gyre. *J. Geophys. Res. Oceans*, **124**, 154–170, https://doi.org/10.1029/2018JC014391.

Racapé, V., V. Thierry, H. Mercier, and C. Cabanes, 2019: ISOW spreading and mixing as revealed by Deep Argo floats launched in the Charlie-Gibbs Fracture Zone.*J. Geophys. Res. Oceans*, **124**, 6787–6808, https://doi.org/10.1029/2019JC015040.

Rintoul, S. R., and C. E. da Silva, 2019: Antarctic circumpolar current. *Encyclopedia of Ocean Sciences*, 3rd ed. J. K. Cochran, H. J. Bokuniewicz, and P. L. Yager, Eds., Elsevier, 248–261.

Roemmich, D., and Coauthors, 2019: On the future of Argo. *Front. Mar. Sci.*, **6**, 439, https://doi.org/10.3389/fmars.2019.00439.

Ryan, A. G., and Coauthors, 2015: GODAE Oceanview class 4 forecast verification framework: Global ocean inter-comparison. *J. Oper. Oceanogr.*, **8** (Supp1.), s98–s111, https://doi.org/10.1080/1755876X.2015.1022330.

Scanderbeg, M., J. P. Rannou, J. Buck, C. Schmid, J. Gilson, and D. Swift, 2019: Argo DAC trajectory cookbook. Argo Data Management Doc., 175 pp., https://doi.org/10.13155/29824.

Stewart, K. D., A. M. Hogg, S. M. Griffies, A. P. Heerdegen, M. L. Ward, P. Spence, and M. H. England, 2017: Vertical resolution of baroclinic modes in global ocean models. *Ocean Modell.*, **113**, 50–65, https://doi.org/10.1016/j.ocemod.2017.03.012.

Tucker, T., D. Giglio, M. Scanderbeg, and S. S. P. Shen, 2020: Argovis: A web application for fast delivery, visualization, and analysis of Argo data. *J. Atmos. Oceanic Technol.*, **37**, 401–416, https://doi.org/10.1175/JTECH-D-19-0041.1.

Villas Bôas, A. B., and Coauthors, 2019: Integrated observations of global surface winds, currents, and waves: Requirements and challenges for the next decade. *Front. Mar. Sci.*, **6**, 425, https://doi.org/10.3389/fmars.2019.00425.

Wang, P., W. Zhang, H. Wang, H. Dai, and X. Wang, 2020: Assimilation of mid depth velocities from Argo floats in the western South China Sea. *J. Atmos. Oceanic Technol.*, **37**, 141–157, https://doi.org/10.1175/JTECH-D-18-0233.1.

Wong, A. P. S., and Coauthors, 2020: Argo data 1999–2019: Two million temperature-salinity profiles and subsurface velocity observations from a global array of profiling floats. *Front. Mar. Sci.*, **7**, 700, https://doi.org/10.3389/fmars.2020.00700.

——, and Coauthors, 2021: Argo quality control manual for CTD and trajectory data. Argo Data Management Team Doc., 113 pp., https://doi.org/10.13155/33951.

Wunsch, C., and P. Heimbach, 2013: Two decades of the Atlantic meridional overturning circulation: Anatomy, variations, extremes, prediction, and overcoming its limitations. *J. Climate*, **26**, 7167–7186, https://doi.org/10.1175/JCLI-D-12-00478.1.

Zilberman, N. V., D. H. Roemmich, and S. T. Gille, 2014: Meridional volume transport in the South Pacific: Mean and SAM-related variability. *J. Geophys. Res. Oceans*, **119**, 2658–2678, https://doi.org/10.1002/2013JC009688.

——, ——, and J. Gilson, 2020: Deep-ocean circulation in the southwest Pacific Ocean interior: Estimates of the mean flow and variability using Deep Argo data. *Geophys. Res. Lett.*, **47**, e2020GL088342, https://doi.org/10.1029/2020GL088342.

——, M. C. Scanderbeg, A. R. Gray, and P. R. Oke, 2022: Scripps Argo trajectory-based velocity product 2001–01 to 2020–12: In Scripps Argo trajectory-based velocity product. University of California, San Diego, Library Digital Collections, accessed 15 August 2022, https://doi.org/10.6075/J0KD1Z35.