Eddy Detection Inverted from Argo Profiles to Surface Altimetry

Xiaoyan Chen,[a] Graham D. Quartly,[c] and Ge Chen[a,b]

[a] *Frontiers Science Center for Deep Ocean Multispheres and Earth System, School of Marine Technology, Ocean University of China, Qingdao, China*
[b] *Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, China*
[c] *Plymouth Marine Laboratory, Plymouth, United Kingdom*

ABSTRACT: Argo floats are widely used to characterize vertical structures of ocean eddies, yet their capability to invert sea surface features of eddies, especially those overlooked by available altimeters, has not been explored. In this paper, we propose an "interior-to-surface" inversion algorithm to effectively expand the capacity of eddy detection by estimating altimeter-missed eddies' surface attributes from their Argo-derived potential density anomaly profiles, given that the interior property and surface signature of eddies are highly correlated. An altimeter-calibrated machine learning ensemble is employed for the inversion training based on the joint altimeter–Argo eddy data and shows promising performance with mean absolute errors of 5.4 km, 0.5 cm, and 14.3 $cm^2\,s^{-2}$ for eddy radius, amplitude, and kinetic energy, respectively. Then, the trained ensemble model is applied to independently invert the properties of eddies captured by an Argo-alone detection scheme, which yields high spatiotemporal consistency with their altimeter-captured counterparts. In particular, a portion of Argo-alone eddies is ~25% smaller than altimeter-derived ones, indicating Argo's unique capability of profiling weaker submesoscale eddies. Sea surface temperature and chlorophyll data are further applied to validate the reliability of eddies identified and characterized by the Argo-only algorithm. This new methodology effectively complements that of altimetry in eddy detection and can be expanded to estimate other physical/biochemical eddy variables from a variety of in situ observations.

SIGNIFICANCE STATEMENT: Despite thousands of eddies being routinely identified on a daily basis, it has been recognized that a substantial portion of eddies may still be missed due to inadequate sampling of altimeter constellations. Taking advantage of eddy's correlation between surface and interior, a considerable number of eddies are discovered for the first time through an Argo-based eddy identification scheme. Here, we propose a new methodology to independently infer these recaptured eddies' surface properties from their vertical signals through an "interior-to-surface" inversion process. The inferred eddy properties are verified by the spatiotemporal consistency with those derived from altimetry. Since Argo is capable of profiling smaller and weaker eddies, the proposed methodology significantly complements and expands that of altimetry in eddy observation.

KEYWORDS: Ocean; Eddies; In situ oceanic observations; Satellite observations

1. Introduction

Oceanic eddies are the dominant signal in the global ocean and play a crucial role in diverse physical processes (Chelton et al. 2011a,b; Dong et al. 2014; Keppler et al. 2018). These structures, through their surface signature, have been characterized and widely investigated using sea level anomaly (SLA) maps retrieved from satellite altimetry observations (e.g., Chelton et al. 2011a; Chen et al. 2022; Faghmous et al. 2015; Mason et al. 2014; Pegliasco et al. 2022; Tian et al. 2020). However, it has recently been recognized that the spatiotemporal undersampling of present-day satellite-derived sea level fields cannot capture most of the eddies (Amores et al. 2018, 2019), making it a challenge to gather full observations and characterizations of eddies in the global ocean.

The abundance of Argo profiling data allows vertical characteristics of mesoscale eddies to be systematically revealed (e.g., Chaigneau et al. 2011; de Marez et al. 2019; Sun et al. 2017; Zhang et al. 2013). The vertical dynamics of eddies are capable of trapping ocean water and generating movements inside their interior cores, causing significant seawater temperature/density anomalies which would map onto the sea surface and create a unique eddy footprint (e.g., Barceló-Llull et al. 2017; X. Chen et al. 2021b; Gaube et al. 2013, 2014). Previous studies have indicated that the correlation between the ocean surface and the interior structure has the potential to facilitate the reconstruction of the three-dimensional upper-ocean circulation structures (e.g., Qiu et al. 2018). However, the identification and inversion of eddies based on vertical structure have not been extensively explored in existing literature. The existing studies have mainly focused on the so-called surface-to-interior reconstruction by aligning Argo measures with sea surface signals, while the ability of back-projecting sea surface features through the Argo-derived vertical features is ignored. A direct diagnosis of the eddy signal based on Argo profiling data would be an ideal way to compensate

for the sparse sampling of the altimeters, and it also gives an innovative perspective to identify eddies through in situ float observations. To do so, an Argo-based independent eddy identification scheme has first been proposed by G. Chen et al. (2021a) and further explored by X. Chen et al. (2021a) and Huang et al. (2021). Over 230 000 "missed eddies" are efficiently recaptured by this verified Argo-alone scheme, a large portion of which resides in the tropical ocean where the traditional altimeter technology appears to be less effective (Cao et al. 2022; Pujol et al. 2016). However, it is recognized that the Argo-alone criteria can only locate the onboard "eddy points," but none of these studies tackle the question of the backward inversion of the surface attributes of these reidentified eddies. Many needs exist for an Argo-based independent eddy detection scheme to provide essential information on eddy surface properties, so as to realize a complete and reliable Argo-only approach for eddy observation.

The objective of this study is to establish an "interior-to-surface" inversion to estimate eddy surface properties of radius, amplitude, and kinetic energy from their Argo-derived interior signatures of potential density anomalies (PDAs). This methodology will be effective as long as the relationship between these eddy surface characteristics and the vertical potential density profile features is unique, significant, and robust. To obtain a precise description of the functional connections between the interior configurations and different surface features for individual eddies, a novel machine learning model named ensemble learning (EL) is utilized given its significant advantages of generalization ability and tolerance to data noise (Heinermann and Kramer 2016; Mendes-Moreira et al. 2012). The rest of this paper is structured as follows: section 2 describes the altimeter, Argo, and sea surface temperature data used in this study. Details of the algorithm for altimeter- and Argo-based eddy detection are presented in sections 3a–3c. In section 3d, we correlate the vertical signals of eddies with the corresponding sea surface features to build the EL inversion model and verify the accuracy based on the joint altimeter–Argo dataset. In section 4, the trained EL model is applied to Argo-only identified eddies to give their sea surface information, and further analysis of the characteristics of Argo-only eddies is provided. Finally, section 5 gives our conclusions.

2. Data

a. Altimetry and Argo data

The delayed-time altimeter products used in this study are generated by the Copernicus Marine Environment Monitoring Service (CMEMS) from a combination of TOPEX/Poseidon, *ERS-1/2*, *Envisat*, *Jason-1/2/3*, *Sentinel-3A*, *HY-2A*, *SARAL/Altika*, *CryoSat-2*, and *GFO* missions (CMEMS 2019). Specifically, the "all-satellite" merged daily SLA fields with a spatial resolution of $1/4° \times 1/4°$ are employed for eddy identification and tracking, spanning 27 years from January 1993 to January 2019 (available at https://data.marine.copernicus.eu/product/SEALEVEL_GLO_PHY_L4_MY_008_047/). A subdataset from January 2002 to January 2019 is used for simultaneous analysis with Argo measurements.

Seventeen years of Argo data between January 2002 and 2019 are collected and made freely available by the International Argo Program and the national programs that contribute to it. The Argo profiles with subsurface temperature/salinity (T/S) measurements (down to a depth of 2000 m) are provided by the Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) at ftp://ftp.ifremer.fr/ifremer/argo (Argo 2024). Currently, Argo is collecting ~12 000 data profiles each month (~400 profiles per day), spanning the global ocean with a nominal spacing of every 3° of longitude and latitude. This greatly exceeds the amount of data that can be collected from below the ocean surface by any other method. The quality control and processing of Argo data are conducted automatically by the Coriolis Center, and only profiles flagged as "good" or "probably good" are downloaded. However, some of these preprocessed profiles may still be suspicious (Chaigneau et al. 2011). Therefore, further quality controls are implemented to ensure the high quality of the retained Argo data. This involves only using profiles with both measurements shallower than 10 m and deeper than 1000 m, with a minimum of 30 valid data points within the 0–1000-m depth range. Finally, these retained profiles are linearly interpolated onto regular vertical levels from the surface down to 1000 m at an interval of 1 m. The PDA is computed based on the International Thermodynamic Equation of Seawater 2010 from Argo T/S profiles, with a total of 1 876 687 profiles globally.

b. Sea surface temperature anomaly data

Seventeen years (1 January 2002–12 January 2019) of the daily Optimum Interpolation Sea Surface Temperature (OISST) anomaly products used in this analysis are provided by the Physical Sciences Laboratory of the National Oceanic and Atmospheric Administration (NOAA) (available at https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html). The OISST is a long-term climate data record that incorporates observations from different platforms (e.g., satellites, ships, and buoys) into a regular global grid in 1/4° spatial resolution. Detailed information can be found in Banzon et al. (2016).

c. Drifter buoys

Seventeen years (1 January 2002–12 January 2019) of 6-h interpolated ocean current data derived from drifting buoys are used in this work. These data (available at ftp://ftp.aoml.noaa.gov/pub/phod/buoydata/) are obtained at a depth of 15 m and released by the Drifter Data Assembly Center at the Atlantic Oceanographic and Meteorological Laboratory of the NOAA. The detailed descriptions can be found in Lumpkin and Centurioni (2019).

3. Methodology

a. Altimeter-based eddy identification and characterization

A four-step scheme has been established for identifying eddies from altimeter-derived SLA fields (hereafter ALT eddy) based on our earlier works by Liu et al. (2016) and optimized

by Tian et al. (2020). First, individual SLA fields are spatially high-pass filtered to remove coherent signals with wavelength scales larger than 20° of longitude by 10° of latitude using a Gaussian filter, which is proven to be very effective at removing steric heating and cooling effects, as well as other large-scale variability (Chelton et al. 2011a). A visual comparison of the high-pass-filtered SLA field on 2 February 2019 with the global SLA map (not shown) aligns well with the results presented by Chelton et al. (2011a) in their Fig. 1 (bottom). Second, the filtered SLA fields are segmented into 40 regular blocks with a meridional/zonal spacing of 45°/36° (approximately 5000/4000 km near the equator and 850/4000 km near the poles) for the parallel eddy detection, which has been proven to significantly reduce the computational load compared to previous SLA-based methods at the global scale (Liu et al. 2016). These SLA blocks are searched simultaneously for eddy seeds (i.e., local extreme points in each 5×5 grid cell) as candidate eddies. Then, SLA contours are computed using a 0.25-cm interval to extract the eddy boundaries. Finally, all blocks are merged seamlessly into a global map with duplicated eddies eliminated, and altimeter-identified eddy dataset has been created at daily resolution for a period of 27 years in the global ocean. In this work, the eddy amplitude, kinetic energy, and radius that represent both eddy dynamics and scale have been selected as typical eddy surface characteristics. The SLA-derived eddy amplitude (A) and eddy kinetic energy (EKE) can be obtained via

$$A = |h_{\text{seed}} - h_0|, \tag{1a}$$

$$\text{EKE} = \frac{1}{N} \sum_1^N \left[\frac{1}{2}(U_g^2 + V_g^2) \right], \tag{1b}$$

where h_{seed} is the SLA value at the eddy seed, h_0 is the SLA value of the outermost closed SLA contour that defines the circumference of the eddy, and N is the number of data points located inside the eddy contour. Note that only eddies with amplitude larger than 1 cm are retained. The terms U_g and V_g are the zonal and meridional geostrophic velocity anomalies, and the eddy radius R is defined as the radius of the least squares fitting circle corresponding to the outermost enclosed SLA contour. Although eddies identified based on altimeter observations are widely acknowledged as the most effective approach, limitations persist in current satellite altimetry for eddy sensing. On the one hand, the spatial (~80–400 km) scales of the merged SLA product will not allow numerous small eddy signals to be captured, particularly in the tropical ocean (Pujol et al. 2016; Ballarotta et al. 2019). Meanwhile, eddies characterized by a strong ageostrophic nature, with spatial scales much smaller than the Rossby radius of deformation (Chen and Chen 2023), pose a challenge for present-day satellite technology due to the weakening of geostrophic balance, upon which altimeter measurement relies. On the other hand, the inadequate spatial resolution and revisit time of the present-day altimeter constellations are unable to resolve the smaller features (e.g., Amores et al. 2019).

b. Calculation of the potential density anomaly

Given that density is a composite parameter reliant on temperature and salinity, it serves as a comprehensive means to assess the vertical structure induced by eddies. Based on the Argo T/S profiles, the potential density profiles are computed based on the International Thermodynamic Equation of Seawater 2010 using the Gibbs Seawater Oceanographic Toolbox (McDougall et al. 2009). Climatological potential density profiles are derived from T/S data acquired from the CSIRO Atlas of Regional Seas (accessible at https://www.marine.csiro.au/~dunn/cars2009/index.html), which offers the 3D climatological value of T/S data on a 0.5° grid with daily interpolation (Ridgway et al. 2002). Subsequently, the PDA is computed by subtracting the collocated (in terms of both space and time) climatological mean of potential density, with this procedure being repeated for all individual Argo profiles on a global scale.

c. Argo-based eddy identification

To locate altimeter-missed eddies, an Argo-based eddy identification algorithm was proposed by G. Chen et al. (2021a) using Argo-derived PDA vertical profiles (hereafter Argo-only eddy). The main idea is to set up vertical criteria for the identification of anticyclonic eddies (AEs) and cyclonic eddies (CEs) through a training–learning process using concurrent altimeter–Argo measurements and then scan all Argo profiles using the criteria to locate missing eddies. The detailed process mainly consists of the following two steps.

1) VERTICAL CRITERIA ESTABLISHMENT IN COMBINATION WITH ALTIMETER-IDENTIFIED EDDIES

It is understood that the vertical structure of eddies exhibits distinct characteristics in various geographical regions (e.g., Fig. 11 in G. Chen et al. 2021a). Considering the spatial variabilities of the eddy vertical features, the global domain was divided into 5° grids, and vertical criteria were established within each grid using concurrent altimeter–Argo measurements. It is noteworthy that the size of these grids was carefully determined taking into consideration the data richness of the Argo dataset and the precision of the established criteria. Taking a specific grid in the northwest Pacific Ocean (150°–155°E, 30°–35°N) as an example, the process of criteria establishment in this context can be summarized in two primary stages.

The first is eliminating "abnormal eddy" profiles. In principle, PDA profiles within AEs (CEs) should exhibit negative (positive) signatures. However, observations have revealed a subset of AEs (CEs) identified by altimetry that exhibit positive (negative) PDA values, contradicting the expected eddy polarities. For instance, Fig. 1 illustrates a PDA profile represented by the green line located within an altimeter-identified CE, yet it conspicuously presents a substantial negative PDA signal. Hence, it is necessary to exclude these abnormal eddies before formulating the criteria. To do so, we calculate the integral areas a_1 and a_2 of AE and CE PDA profiles ρ_1 and ρ_2 between arbitrary depth pairs h_1, h_2 and h_3, h_4 iteratively in steps of 10 m,
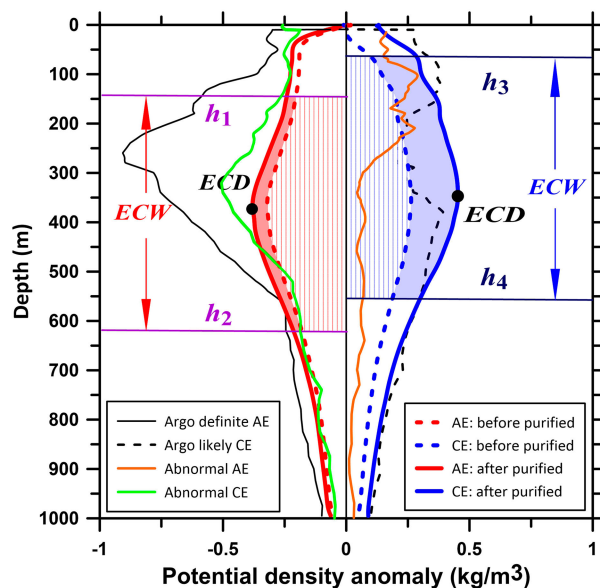
FIG. 1. Schematic illustration of the algorithm for the establishment of Argo-based eddy identification criterion. The dashed red (blue) line is the grid-averaged PDA profile of AEs (CEs) before the elimination of abnormal eddies, and the solid red (blue) line is the grid-averaged PDA profile for AEs (CEs) after the elimination of abnormal eddies. The orange (green) line is the PDA profile of individual abnormal AE (CE). The solid and dashed black lines are two typical PDA profiles outside altimeter-identified eddies but are identified by the Argo criteria as a definite AE and a likely CE, respectively.

$$a_1 = \int_{h_1}^{h_2} \rho_1(h)dh \, (0 < h_1 < h_2 < 1000 \text{ m}), \quad (2a)$$

$$a_2 = \int_{h_3}^{h_4} \rho_2(h)dh \, (0 < h_3 < h_4 < 1000 \text{ m}). \quad (2b)$$

The Argo-determined eddy polarity is classified as AE or CE if $a_1 < 0$ or $a_2 > 0$. If Argo-determined polarity is inconsistent with the polarity identified by the altimeter (e.g., an altimeter-identified AE with $a_1 > 0$), it is considered a polarity-inconsistent abnormal eddy. We systematically examine diverse depth ranges to assess the presence of abnormal eddies within each specific depth interval. When the minimum count of the abnormal eddies is obtained, the corresponding depth pairs, denoted as h_1 and h_2 (h_3 and h_4) for AE and CE, are decided as the boundaries defining the vertical core of the eddy. This choice is made because it provides the most representative indication of the eddy vertical characteristics. As presented in Fig. 1, the depth difference between h_1 and h_2 (h_3 and h_4) is defined as the eddy core width (ECW). Subsequently, eddy profiles displaying abnormal signatures are systematically removed, resulting in a "purified" dataset of altimeter–Argo collocations. The solid and dashed profiles in red and blue in Fig. 1 represent the averaged PDA profiles before and after the elimination of abnormal eddies, respectively. This procedure ensures that the refined eddies, which will serve as the basis for establishing the criteria for Argo-based eddy

identification, exhibit a distinctive vertical PDA structure that aligns with the polarity as determined by altimeter data.

Second, we proceed to determine the depth and intensity of the eddy core. Taking advantage of the purified Argo dataset, we calculate two mean PDA profiles by averaging the purified individual AEs and CEs, respectively. Subsequently, a seven-point running filter is applied iteratively until a single peak appears in the smoothed mean profiles (see the black dots in Fig. 1). The depth corresponding to the PDA extremum is identified as the eddy core depth (ECD), and the integral area within the eddy core width is defined as the eddy core intensity (ECI). These quantifiable attributes serve as fundamental criteria in the context of Argo-based eddy identification.

2) EDDY IDENTIFICATION FROM ARGO PROFILES ACCORDING TO ALTIMETER-CALIBRATED CRITERIA

Based on the criterion established in the previous step, the vertical features of each out-of-eddy PDA profile are examined in the following two steps. First, a seven-point running filter is applied to the outside-eddy profile until it manifests a single core structure, and then, the depth at which the extreme is located is assessed to determine whether it falls into the defined eddy core width. If so, its eddy core intensity is checked. Suppose the eddy core intensity of the averaged PDA profiles of altimeter-identified eddies before and after the purification for AE and CE is b_1, b_2 (indicated as the vertical line filled area in Fig. 1) and B_1, B_2 (indicated as the vertical line filled plus shaded filled area in Fig. 1), respectively. For a given outside-eddy PDA profile denoted as ρ_0, its integrated PDA areas associated with AE (A_1) and CE (A_2) within the eddy core width are computed:

$$A_1 = \int_{h_1}^{h_2} \rho_0(h)dh, \quad (3a)$$

$$A_2 = \int_{h_3}^{h_4} \rho_0(h)dh. \quad (3b)$$

If $A_1 > B_1$ or $A_2 > B_2$, then the Argo profile has a significantly strong vertical structural signal within the range of the eddy core and therefore is considered a profile inside a definite AE or CE (as an example of a definite AE profile in the solid black line in Fig. 1). If $b_1 < A_1 < B_1$ or $b_2 < A_2 < B_2$, then the Argo profile has a relatively strong vertical structural signal within the range of eddy core and therefore is considered as a profile inside a likely AE or CE (as an example of likely CE profile in the dashed black line in Fig. 1).

It should be noted that the above Argo-based eddy process is divided into four seasons to better accommodate the time-varying nature. A more detailed description of the Argo-based eddy detection algorithm as well as the results analysis can be found in G. Chen et al. (2021a). After this altimetry-calibrated training–learning process, the Argo-only detection increases the number of eddies by 22.5% for the global average. The geographical distributions of Argo profiles in altimeter and Argo-only identified eddies are shown in Fig. 2. The concentration of Argo-only eddies appears to peak in the
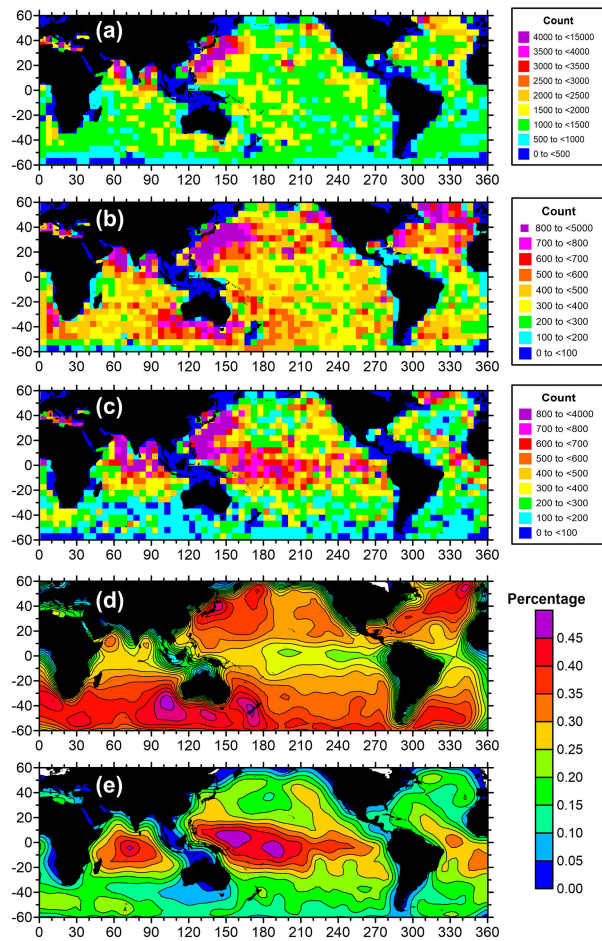
FIG. 2. Number of Argo profiles per 5 × 5 cell during 1 Jan 2002–12 Jan 2019 for (a) the global ocean, (b) inside altimeter-identified eddies, and (c) inside Argo-identified eddies. (d),(e) The frequencies of occurrence corresponding to (b) and (c) with respect to (a) all Argo profiles within the same cell.

tropical ocean, where altimeter eddy identification is known to be ineffective (Fig. 2e). In mid- and high latitudes, even with the strong eddy detection capability of the altimeter, Argo retains the capacity to identify a significant portion of eddies that are missed by the network of altimeters. Moreover, it is essential to clarify that while both this study and prior research have emphasized that a significant proportion of eddies are missed by altimeters (e.g., Amores et al. 2018), it remains an indisputable fact that altimeter-based methods are the most proficient means of eddy detection. Given the limited spatial coverage of Argo data, the eddy identification method using Argo profiles, on the other hand, represents a valuable addition to eddy identification. This method proves particularly advantageous in overcoming the spatial resolution limitations of altimeters, especially at lower latitudes where the altimeter's detection capabilities are relatively limited. Importantly, the Argo-based eddy detection also introduces novel perspectives and methods for in situ real-time observation of eddies.

d. Argo-based eddy property inversion using altimeter-calibrated ensemble learning

An interior-to-surface inversion is performed to systematically estimate typical properties (i.e., radius amplitude and kinetic energy) for these recaptured Argo-only eddies using an altimeter-calibrated ensemble learning model. We first derive the vertical features of PDA as the functional relationship with eddy surface features at a given location through the ensemble learning process and then independently infer all Argo-identified eddy features outside altimetrically derived ones according to their PDA signatures. To intuitively visualize the eddy three-dimensional structure and its intrinsic vertical/surface correlation, we show two normalized altimeter-identified AEs classified based on the vertical eddy core intensity in Figs. 3a and 3b. It can be clearly observed that a stronger eddy vertical signal is accompanied by a larger surface radius, amplitude, and geostrophic velocity. This correlation can be particularly well reflected by the functional curves of the eddy core intensity with respect to the corresponding eddy radius, amplitude, and kinetic energy, as illustrated in Figs. 3c–e. Assessing the correlation coefficient between the two shows that the eddy vertical core intensity correlates with the surface radius, amplitude, and kinetic energy of up to 0.745, 0.893, and 0.848, respectively. These reliable interior–surface relations provide essential support to the proposed idea of an independent Argo-based eddy characterization.

The detailed procedure for the Argo-based eddy property inversion algorithm is presented in a flowchart in Fig. 4. The process involves three main steps to derive the Argo-only eddy properties:

(i) Model training and testing, conducted using collocated altimeter-identified eddies and Argo profiles spanning the period from 2002 to 2009. The data are randomly sampled and divided into a training (60%) dataset and a testing (40%) dataset. This process aims to combine the eddy properties derived by the altimeter and the associated vertical PDA features derived from Argo observations to develop and test the EL model to achieve the most optimal inversion accuracy.

(ii) Model validation based on collocated altimeter-identified eddies and Argo profiles for the period 2010 to 2019. This step involves applying the model to subsequent years to validate its effectiveness. In this process, the properties of the altimeter-identified eddies are considered as the ground truth. By applying the model developed in step (i) to the vertical PDA features of this subset of eddies and obtaining the inverted properties, the model's reliability can be evaluated (refer to section 4a for a detailed discussion).

(iii) Inversion of properties for eddies cannot be detected by the altimeter but are identified through Argo PDA profiles (referred to as Argo-only eddies, as detailed in section 4b). Since these eddies are not observable by the altimeter, the reliability of the inversion results is further verified by integrating other observations, such as sea surface temperature anomaly (SSTA), chlorophyll anomaly (CHLA), and drifter data.
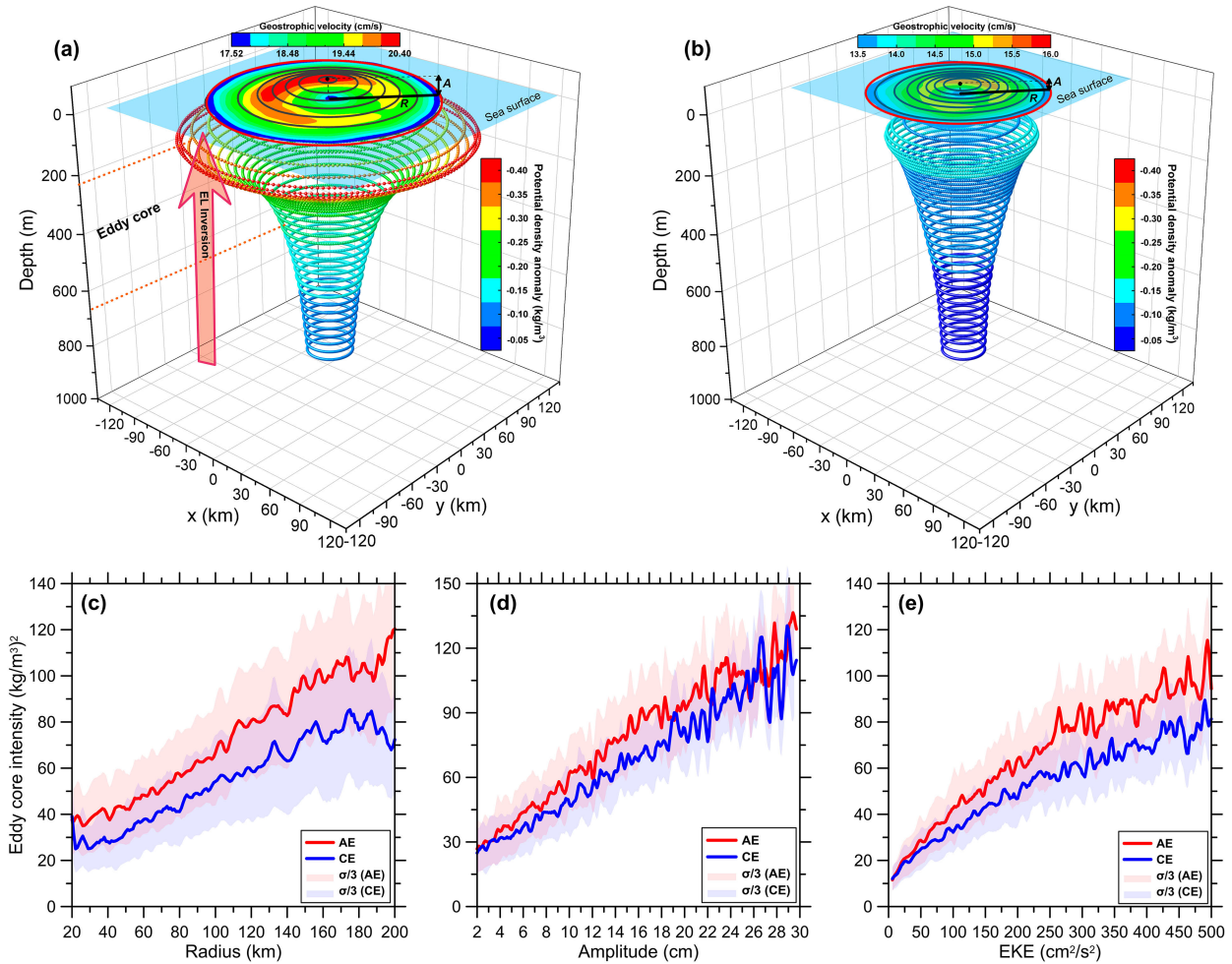
FIG. 3. (a) The global averaged vertical structure and the corresponding sea surface signature for AEs with vertical ECI larger than the global mean ECI. The gray contour and the colored map on the sea surface represent, respectively, the normalized eddy amplitude and geostrophic velocity within the eddy boundary derived from SLA fields, and the red boundary represents the eddy's mean least squares fitted circle boundary. The vertical colored lines represent the averaged PDA derived from Argo profiles. (b) As in (a), but for eddies with vertical ECI smaller than the global mean ECI. (c)–(e) ECI as a function of the eddy surface properties (radius, amplitude, and EKE) derived from the altimeter-identified eddy dataset. The red (blue) lines are for AE (CE), and the shadows around the color lines indicate the standard deviations of the corresponding eddy properties.

As summarized in Table 1, a total of 14 features are extracted as modeling variables to fully characterize the eddy vertical characteristics. On the one hand, the longitude, latitude, and month that reflect the spatial and temporal nature of the PDA profiles are incorporated. On the other hand, eddy representative vertical features, including ECI, eddy core extremum (ECE), ECD, and ECW as defined in section 3c, are incorporated. Furthermore, the eddy PDA features, encompassing integral area and extreme values at different depth ranges, are included to more comprehensively reflect the vertical characteristics of the eddy profiles. In addition, it is worth mentioning that the eddy-induced pycnocline displacements are chosen to be a useful feature given that the pycnocline displacements are significantly correlated with eddy properties (X. Chen et al. 2021b). The pycnocline is defined as the deepest pronounced extremum depth of the Brunt–Väisälä frequency N^2 profile,

and the displacements are computed by subtracting the pycnocline depth within eddies from the monthly climatology at the corresponding position (refer to X. Chen et al. 2021b for specific details). Figure 5 shows the relative importance of different vertical features for the inversion of eddy sea surface features of radius, amplitude, and kinetic energy. Comparing Figs. 5a and 5b, one can find that the relative importance distributions of AEs and CEs are rather consistent, indicating there is no significant eddy polarity preference for the proposed algorithm. However, a large difference exists between the relatively important vertical features for the inversion of different eddy surface parameters. For instance, the inversion of eddy kinetic energy is most sensitive to its vertical ECI, while the ECE is more important for eddy radius inversion. As for the eddy amplitude, the most sensitive features for both AE and CE are the extreme value within 600–1000 m of the eddy's PDA
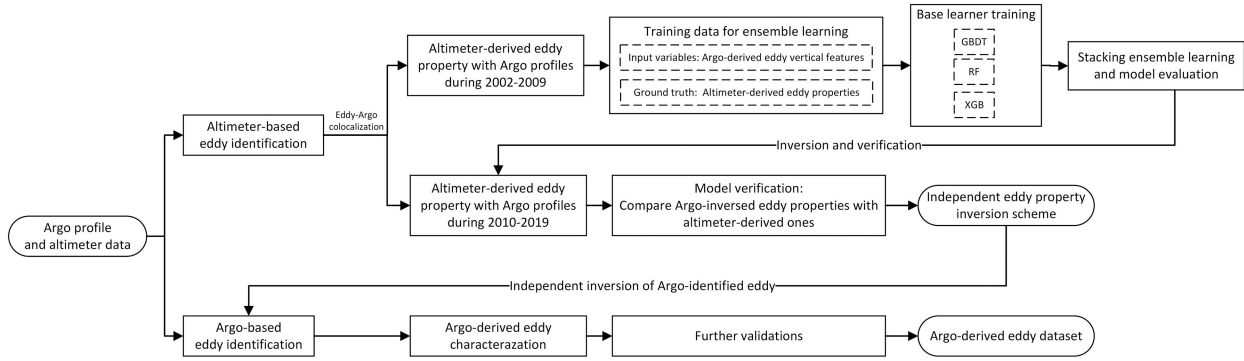
FIG. 4. A flowchart of the Argo-based eddy property inversion.

profile (E3). This could be partly attributed to the relatively robust relationship between eddy amplitude and the deep ocean layer under the potential vorticity conservation (X. Chen et al. 2021b). The sensitivity of different eddy surface features to different vertical features can be further evidenced in their vertical distributions of PDA profiles (see Fig. 11). This complex interior–surface nonlinear relationship makes machine learning an ideal approach to implementing the interior-to-surface modeling process.

A machine learning ensemble is employed to extract the eddy's vertical-surface functional relationship and conduct the interior-to-surface inversion. The multimodel ensemble learning has gained popularity due to its ability to generalize results and reduce model variance by combining multilearners. This approach has demonstrated notable advantages in estimating and predicting various oceanographic phenomena, including SST trends, subsurface thermal structures, precipitation, and coastal engineering (Kim and Lee 2022; Qi et al. 2023; Weisheimer et al. 2009; Xu et al. 2020). While it has been proven to enhance prediction reliability and reduce model uncertainty, it does come with certain limitations. On the one hand, it demands significant computational resources for training multiple models and combining their predictions. On the other hand, the interpretability of the EL model is comparatively limited, as the combination of predictions from multiple models complicates the understanding of the underlying decision-making process in comparison with a single model. Despite these limitations, ensemble learning remains a valuable and widely used approach in machine learning, particularly in situations where improved predictive performance is a priority. We experimentally compare different regression algorithms composed to ensembles (e.g., the support vector machine, the k-nearest neighbors, and the decision tree), and the gradient boosting decision tree (GBDT), the random forest (RF), and the extreme gradient boosting (XGB) are adopted as the individual base learners given their better performance. The mean absolute error (MAE) and the determination coefficient R^2 of these base learners in inverting different eddy properties are shown in Table 2. Then, the stacking method S is chosen to combine these base learners to a metalearner (Breiman 1996). Let $x_i = [v_1, v_2, \ldots, v_{14}]$ be an input vertical feature vector, and the ensemble learning inversion can be described in Eq. (4) as

$$f_p(x_i) = S_p\{[\text{GBDT}_p(x_i), \text{RF}_p(x_i), \text{XGB}_p(x_i)]\},$$
$$p = (R, A, \text{EKE}), \qquad (4)$$

where $f_p(x_i)$ is the inverted eddy properties. It should be pointed out that the Bayesian optimization, which uses a

TABLE 1. Eddy vertical features that are used for the EL inversion model.

| Feature	Description
Location: latitude (Lat)	Location (latitude) of the eddy
Location: longitude (Lon)	Location (longitude) of the eddy
Time: month (Mon)	The month in which the eddy is identified
ECI	Integral area within the depth range of the eddy core
ECE	Extreme value in the eddy core
ECD	The depth corresponding to the extreme value in the eddy core
ECW	The upper and lower boundary depth range of the eddy core
Integral area_1 (IA1)	Integral area within 20–300 m of eddy's PDA profile
Integral area_2 (IA2)	Integral area within 300–600 m of eddy's PDA profile
Integral area_3 (IA3)	Integral area within 600–1000 m of eddy's PDA profile
Extremum_1 (E1)	Extreme value within 20–300 m of eddy's PDA profile
Extremum_2 (E2)	Extreme value within 300–600 m of eddy's PDA profile
Extremum_3 (E3)	Extreme value within 600–1000 m of eddy's PDA profile
Pycnocline depth displacement (DPA)	The displacement of the pycnocline inside the eddy
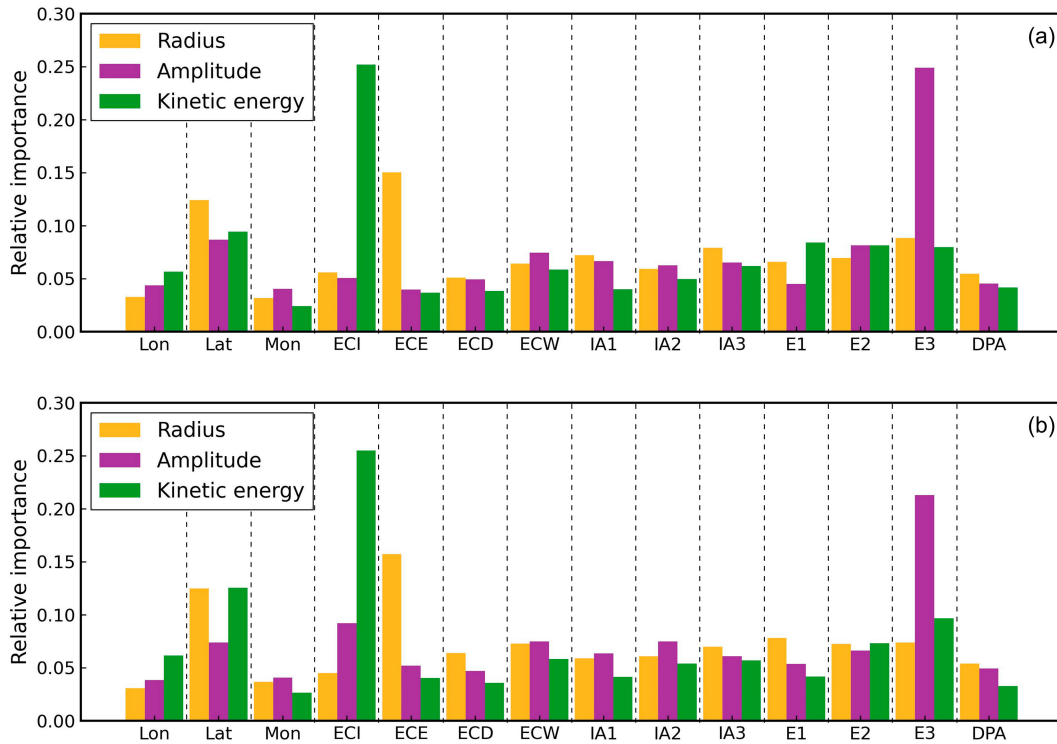
FIG. 5. Histogram of the relative importance of the eddy vertical PDA features for the inversion of sea surface properties for (a) AEs and (b) CEs. The yellow, purple, and green bars represent eddy radius, eddy amplitude, and EKE, respectively. Detailed descriptions of the eddy vertical characteristic features on the *x* axis are presented in Table 1.

small number of iterations and has a rapid running speed, is applied to each base learner to obtain the best configuration of the hyperparameters. After the Bayesian optimization, these base learners are ensembled using stacking, and the R^2 (MAE) of the EL model in inferring eddy radius, amplitude, and kinetic energy reaches the maximum (minimum) about 0.89 (5.44 km), 0.91 (0.54 cm), and 0.92 (14.29 cm^2 s^{-2}), respectively (see details in Table 2). Furthermore, bootstrap confidence interval tests are conducted to verify the stability of each EL model. Each eddy property is repeatedly inverted 1000 times. As indicated in the rightmost column of Table 2, a narrow confidence interval is obtained for each eddy parameter. Specifically, the mean confidence interval widths for the

eddy radius, amplitude, and kinetic energy are approximately 0.8 km, 0.1 cm, and 2.5 cm^2 s^{-2}, respectively, underscoring the model's robust stability.

4. Results and discussion

a. Comparison of inverted eddy properties with altimeter-derived estimates

The established EL model is first validated by inverting the surface properties of altimeter-identified eddy from the allocated Argo PDA profiles during 2010–19 (Fig. 6). The left column of Fig. 6 represents eddy radius, amplitude, and kinetic energy estimated by the altimeter-derived SLA fields, while

TABLE 2. The inversion accuracy of the EL model and base learners for eddy surface properties. The bold column indicates the MAE of the EL model, which can be found to generally outperform the base learner in inferring each eddy parameter. Note. AE = anticyclonic eddy, CE = cyclonic eddy, EL = ensemble learning, GBDT = gradient boosting decision tree, RF = random forest, and XGB = extreme gradient boosting.

| Eddy parameters	Eddy type	MAE				R^2				Bootstrap confidence interval ($\alpha = 0.05$)
		EL	GBDT	RF	XGB	EL	GBDT	RF	XGB	
Radius (km)	AE	**5.01**	5.04	6.09	6.63	0.90	0.90	0.84	0.89	(4.60, 5.38)
	CE	**5.86**	8.27	5.84	6.75	0.88	0.87	0.84	0.88	(5.36, 6.19)
Amplitude (cm)	AE	**0.54**	0.53	1.36	0.60	0.91	0.91	0.89	0.90	(0.47, 0.57)
	CE	**0.53**	0.59	1.12	0.55	0.91	0.91	0.88	0.91	(0.47, 0.59)
EKE (cm^2 s^{-2})	AE	**11.80**	15.65	21.36	16.21	0.94	0.93	0.90	0.92	(10.29, 12.43)
	CE	**16.77**	16.91	21.28	17.14	0.91	0.86	0.90	0.91	(15.37, 18.19)
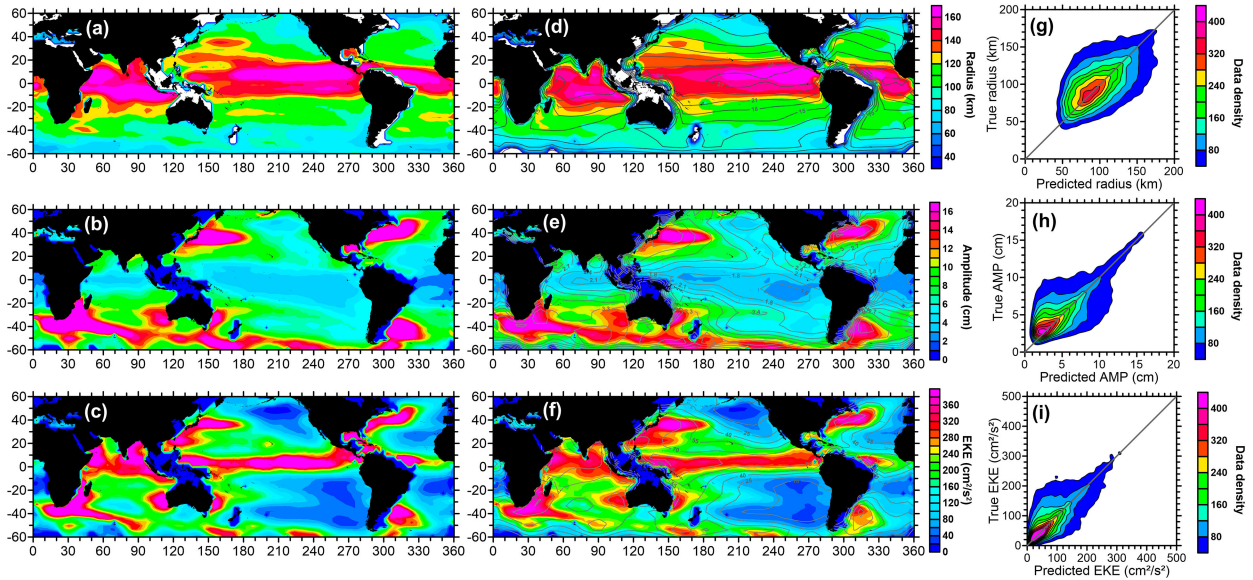
FIG. 6. (a)–(c) Global distributions of eddy properties calculated based on the SLA fields for altimeter-identified eddies during 2010–19. (d)–(f) Global distributions of eddy properties inverted from the Argo profiles based on the EL model for altimeter-identified eddies during 2010–19. The MAE between the SLA-derived and EL-derived eddy properties is overlaid as gray contours on each map. (g)–(i) Data density map of eddy properties estimated by the EL inversion as a function of SLA-derived eddy properties. From top to bottom, rows show eddy radius, eddy amplitude, and EKE.

the middle column corresponds to those inferred by the EL model based on the PDA profiles. Estimates of eddy surface features from the two different dimensions give broadly similar distributions for both the eddy scale and kinetic energy. Taking amplitude as an example, it is larger in the western boundary currents (e.g., the Kuroshio Extension and the Gulf Stream) and the Antarctic Circumpolar Current (ACC), while smaller in the tropical oceans, which indicates the robust behavior and regional independency of the proposed methodology. To further verify the accuracy of the EL model, the derived parameters as a function of the SLA-derived ones are plotted in Figs. 6g–i, in which the diagonal lines represent the ideal prediction results (i.e., the predicted values are equal to the true values), showing that the majority of the estimations are distributed around the ideal lines. As summarized in Table 3, the global average of the eddy parameters obtained by the EL inversion is very close to the true values calculated by SLA fields. The global average differences between altimeter- and EL model-derived eddy parameters are 3.1 km, 0.8 cm, and 9.4 cm^2 s^{-2} in terms of eddy radius, amplitude,

and kinetic energy, further indicating a promising inversion of eddy surface parameters.

It should be pointed out that the eddy activity also has significant interannual variability, which may be related to the evolution of El Niño–Southern Oscillation (ENSO), Indian Ocean dipole, and other climate modes. For instance, Zheng et al. (2018) investigated the interannual variability of eddy kinetic energy in the subtropical southeast Indian Ocean, revealing a negative correlation between ENSO and eddy kinetic energy. They documented that ENSO modulates this variability by influencing the baroclinic instability of velocity shear. Zheng et al. (2016) identified tropical instability vortices (TIVs) in the equatorial eastern tropical Pacific Ocean using surface drifters. They found that the number of identified TIVs also exhibits substantial interannual variability, with fewer TIVs identified during El Niño events and more TIVs found during La Niña events. Therefore, ENSO and climate change introduce additional sources of variability and complexity when characterizing eddies through Argo profiles. Nevertheless, our specific investigation into the interannual variability of inversion errors in eddy parameters within the equatorial eastern tropical Pacific Ocean reveals a notable degree of stability, even during intense ENSO years. Indeed, the observed stability is logically justified since the calculation of eddy properties based on the altimeter SLA field has been high-pass filtered to remove large-scale signals. Concurrently, the Argo PDA profiles are systematically excluded from the climatological background field. Therefore, the relationship established between the two is thought to be representative of eddy characteristics to a larger extent. In addition, the augmented Dickey–Fuller (ADF) test is applied to the monthly

TABLE 3. Statistics of altimeter- and Argo-only identified eddy surface properties.

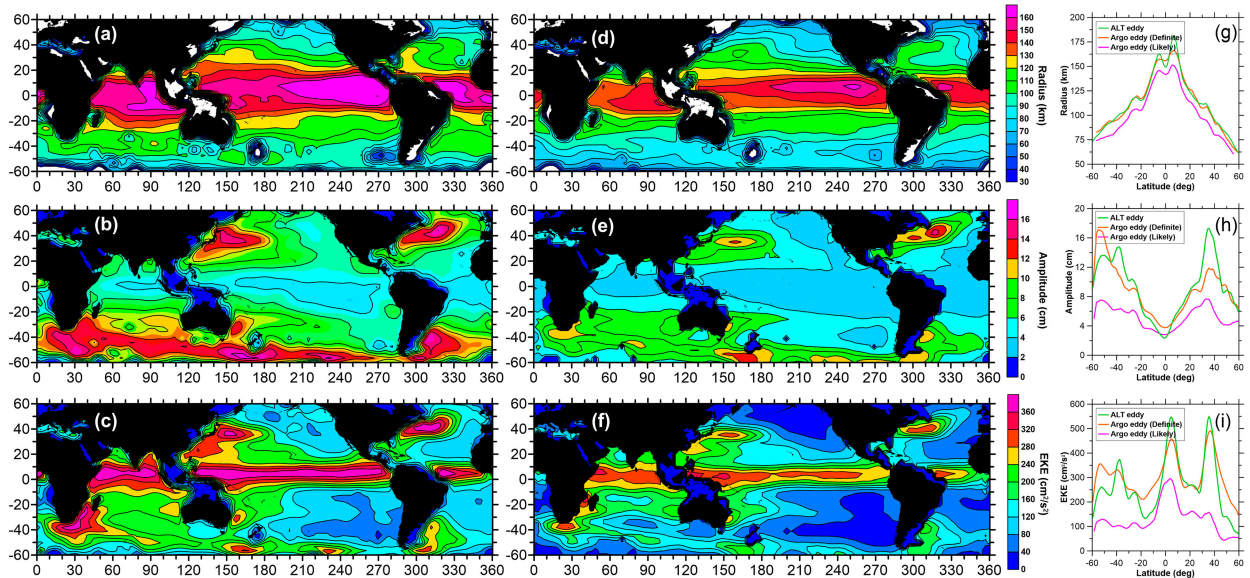| Eddy type	ALT eddy		Argo-only eddy	
			Definite	Likely
Methodology	SLA field	EL inversion	EL inversion	
Radius (km)	111.3	107.2	110.4	99.8
Amplitude (cm)	10.4	9.6	9.1	4.9
EKE (cm^2 s^{-2})	190.9	181.5	189.8	159.7

FIG. 7. (a)–(c) Global distributions of the eddy radius, amplitude, and kinetic energy derived from the Argo profiles based on the EL model for Argo-identified definite eddies. (d)–(f) As in (a)–(c), but for Argo-identified likely eddies. (g)–(i) Corresponding zonal distribution of (a)–(f). The green lines are for ALT eddies, and the orange (magenta) lines are for Argo definite (likely) eddies.

variation of the errors between the inverted eddy properties derived from the Argo profiles and their true properties derived from the SLA fields, providing a statistical assessment of the model's stability in temporal variations. The ADF test yields p values below the 1% significance level for each eddy property in both AE and CE. This outcome not only proves the robust stability of the proposed model in terms of time dimension but also emphasizes its resilience in the context of global warming and ENSO dynamics.

b. Independent inversion for Argo-identified eddy properties

In this section, the calibrated EL model is applied to independently estimate the surface properties of all the Argo-only eddies. There are two primary reasons for the failure of present-day altimeter measurements to detect all eddies. The first reason pertains to the undersampling of currents by current altimeters, as demonstrated in studies such as Amores et al. (2019) and Cao et al. (2022). The second reason is associated with the diminishing influence of geostrophic balance, particularly in the context of submesoscale ageostrophic eddies, which are too weak for detection by altimeters (e.g., Callies and Ferrari 2013). Based on the proposed Argo-only eddy identification algorithm, a large portion of missed eddies can be recaptured and exhibit significant geographical complementarity with the eddies identified by the altimeter, which exhibits an Argo-only eddy peak in the equatorial region (see Figs. 2d,e). These Argo eddies are classified into two categories during the Argo-based eddy identification procedure, definite eddies and likely eddies, according to the eddy core intensity (see section 3c for the detailed criterion of Argo-identified definite and likely eddies). The trained EL model is applied to the vertical profiles of the Argo-only eddies, and

the inverted sea surface properties of the Argo-only eddies show the same broad spatial structure with the ALT eddies (Figs. 7a–f), while the Argo-identified likely eddies are systematically smaller and weaker (Figs. 7d–f). It is estimated that the global averages of the definite (likely) Argo-only eddy radius, amplitude, and EKE are 110.4 km (99.8 km), 9.1 cm (4.9 cm), and 189.8 cm^2 s^{-2} (159.7 cm^2 s^{-2}), respectively (see Table 3). In other words, the likely Argo eddies are about 9.9%, 49.7%, and 15.9% weaker than ALT and definite Argo eddies with respect to eddy radius, amplitude, and EKE, indicating Argo's capability of profiling weaker and smaller eddies. It is worth pointing out that the Argo-only eddies exhibit the most markedly small amplitudes (down to almost half the size of those captured by the altimeter), and the height of sea surface amplitude is the main foundation for the altimeter to distinguish eddy signals. When the eddy is too "short," it is difficult to be captured by available altimeters but can be diagnosed and indicated through Argo profiles, which further reflects the complementary and extended nature of the proposed methodology for altimeter eddy observation. To highlight the surface signature differences in the altimeter and Argo identification eddies, the corresponding zonal distributions are given in Figs. 7g–i. The latitudinal variation of the definite Argo eddy features almost coincides with the ALT eddies (orange and green lines in Figs. 7g–i), reflecting the high probability of this part of eddies being within the sampling spacing of the altimeter thus not being captured. However, some notable differences exist in magnitude between ALT eddies and likely Argo eddies (magenta and green lines in Figs. 7g–i) but remain in agreement with the zonal variation trend.

In terms of the temporal variability, we examine the yearly variation of the three properties for the two eddy categories
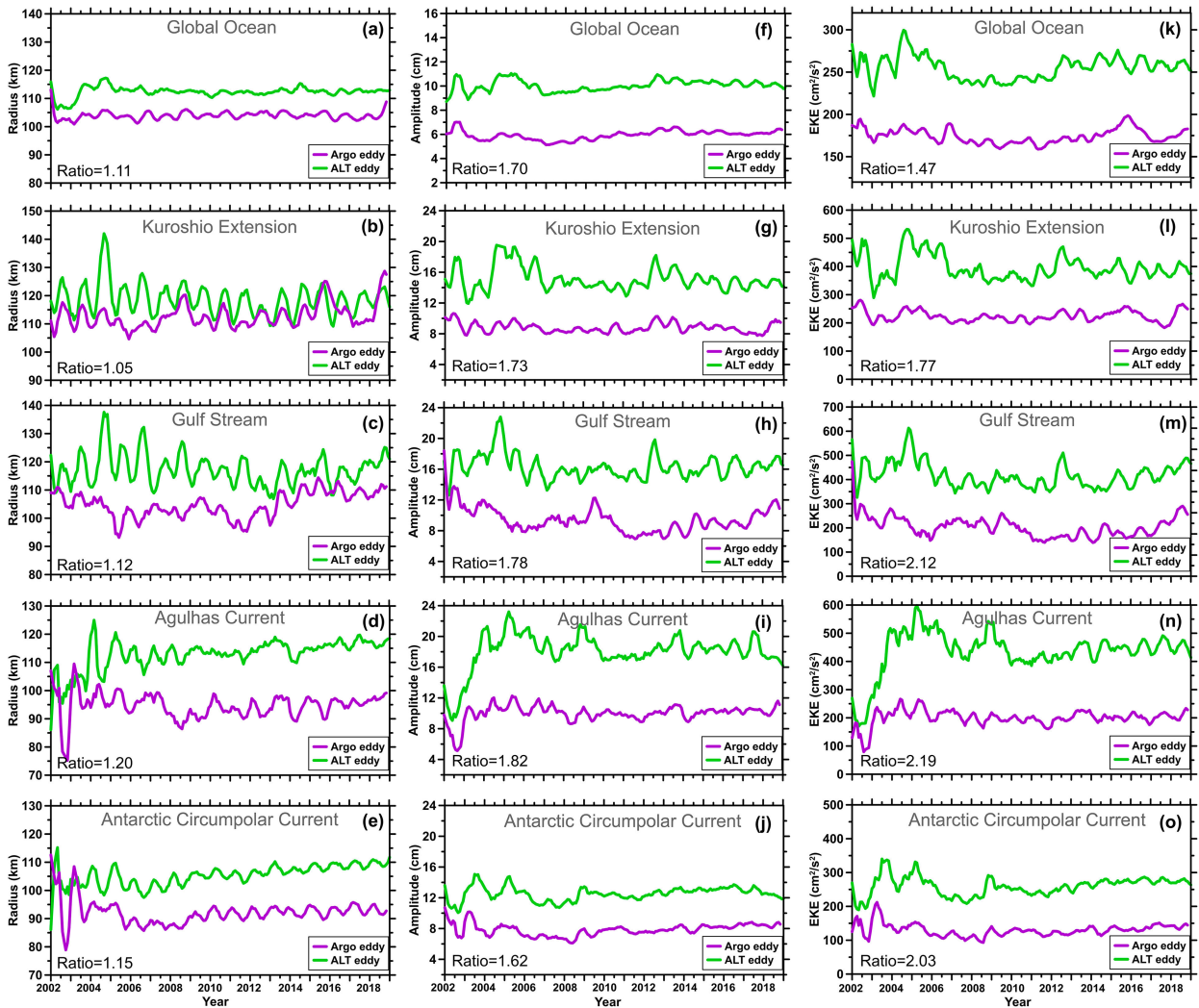
FIG. 8. Variations in monthly mean eddy properties of Argo-only eddies derived by the EL model (purple lines) and ALT eddies derived from SLA fields (green lines) in the global and typical regions from 2002 to 2019. (left) Eddy radius, (middle) amplitude, and (right) kinetic energy. From top to bottom, rows show the global ocean, the Kuroshio Extension, the Gulf Stream, the Agulhas Current, and the ACC. The ratio of the two curves is depicted in the lower-left corner of each panel.

in the global ocean and specific regions (the Kuroshio Extension, the Gulf Stream, the Agulhas Current system, and the ACC) during the study period, as shown in Fig. 8. Going through the top panels for global average, an overall impression is that the Argo-only identified eddies (including definite and likely ones) have smaller values than the altimeter-identified ones, with the ratios of the parameters of altimeter- to Argo-identified eddies being 1.11, 1.70, and 1.47, for eddy radius, amplitude, and EKE, respectively. In terms of several typical strong current systems accompanied by substantial eddy activities, the parameters display a notable annual and semiannual cycle for both ALT and Argo-only eddies. For example, the two types of eddy properties in the Kuroshio Extension display interannual fluctuations under strikingly similar frequencies, with an increase in the first half of the year and a decrease in the second half. As previously reported, the eddy kinetic energy reaches a maximum in April/May and a minimum in

December/January associated with the seasonal heating/ cooling of the upper layer that modifies the vertical velocity shear (Qiu 1999), which in turn reveals the intrinsic connection between eddy's surface and vertical kinetic energy. Previous investigations have indicated that within the Gulf Stream, eddies are slightly bigger but less intense during autumn and winter and, conversely, more intense yet smaller in size during spring. The observed seasonal variability is believed to arise from a confluence of potential mechanisms involving wind stress, thermal forcing, and topographic influences (Kang and Curchitser 2013). However, it is important to note that the Argo-based eddy detection, due to its capability to identify a subset of small and weak eddies, which are more susceptible to the abovementioned external driving forces, may display seasonal variability influenced by factors that lead to trends occasionally divergent from those observed in altimeter-derived eddies. A more comprehensive examination of the underlying mechanism for

different types of eddies will be conducted in future investigations. In the Agulhas Current, high eddy kinetic energy develops during the wind/spring when encountering stronger wind stress (Zlotnicki et al. 1989). Such a consistent result confirms that the proposed interior-to-surface inversion is globally applicable and regionally guaranteed, which is therefore confident to argue that the Argo-based scheme offers an independent, complete, and robust eddy detection process in the global ocean. Moreover, we extended the application of the Argo-based eddy detection and inversion to PDA profiles from 2020 to 2023 in the whole Pacific Ocean to further verify the effectiveness (not shown). Our findings demonstrate that the proposed approach remains highly effective even in the context of global warming for recent years, thereby providing confidence to argue that the Argo-based scheme offers an independent, complete, and robust eddy detection process.

c. Further evidence of Argo-only derived eddies

To further validate the recaptured and characterized eddies obtained from the vertical Argo PDA profiles, a comparison of these eddies with those derived by the altimeter is performed in terms of both surface and vertical characteristics. It is noteworthy that the eddies identified and characterized by Argo PDA profiles examined in this section are not detectable by the altimeter; they are solely identified by the Argo-based method proposed in this study and are defined as "Argo-only eddies." On the one hand, independent verification is adopted based on other satellite measurements of the SSTA and CHLA. Given that eddies are often associated with distinctive SSTA signatures (e.g., X. Chen et al. 2021a; Dong et al. 2011), the SSTA characteristics induced by the eddies identified by the altimeter and those that are exclusively detected by Argo PDA profiles (Argo-only eddies) are examined. The eddy-entrained chlorophyll is selected as another case for the evaluation. In doing so, we selected a similar area to Gaube et al. (2013) in the eastern southern Indian Ocean and conducted a normalized composition of chlorophyll within eddies identified by both the altimeter and Argo-only methods. As shown in Figs. 9a and 9b, the SSTA distributions in the eddy coordinate system inside Argo-only eddies show a clear monopole core in the eddy center, which, although not as strong as for ALT eddies, provides a convincing demonstration of the unique advantages and high reliability of Argo-eddy recognition. The normalization process consists of dividing the displacements of the SSTA or chlorophyll features by the derived eddy radius and collating these to map the mean intraeddy variations; this is explained in more detail in Chen et al. (2019). In particular, significant SSTA features occur within the equivalent circle boundary corresponding to the Argo-only eddy radius inverted using the PDA profiles, and the SSTA signals gradually decay when exceeding the eddy boundary, which provides a further demonstration of the reliability of the Argo-inverted eddy radius. Similar to the SSTA analysis, Argo-only eddies displayed CHLA patterns akin to those of the altimeter-identified eddies (Figs. 9e–h), further providing a convincing demonstration of the unique advantages and high reliability of Argo-eddy recognition,

whereas the observed distinctions in the high CHLA distribution between Argo-only and altimeter-identified CE may be attributed to the relatively weaker nature of Argo-only identified eddies, leading to the CHLA distribution being susceptible to wind fields.

It is well established that the drifter buoys can reveal the rotational motion of eddies, giving us another effective way to validate the Argo-derived eddy signals. As an example, Figs. 10a and 10b show two drifter trajectories around the Argo-only AE and CE, whose PDA profile is present in Fig. 10c. It is apparent that the drifters moved clockwise (anticlockwise) around the Argo-only AE (CE), indicating that Argo-only eddies possess not only significant vertical signatures but also typical circulation characteristics. Then, based on the inverted eddy radius, the composite analyses are conducted on the collected drifter and Argo-only eddies from the period 2002–19. In the Northern (Southern) Hemisphere, a characteristic clockwise (anticlockwise) current field is observed within the Argo-only AE along with a strong current speed, particularly within the boundaries corresponding to the inferred eddy radius (Figs. 10d,e). The opposite behavior is observed for the Argo-only CE. There are thousands of matchups of Argo profiles with currents from surface drifters, and these can be used to give a mean circulation pattern within Argo-only eddies (Fig. 10). However, the derived field is too noisy to allow a portrayal of the mean curl $[(\partial v/\partial x) - (\partial u/\partial y)]$ within the bodies. It is further estimated that the global mean current velocity $\pm$ standard deviation within the Argo-only AE (CE) is 28.53 $\pm$ 21.46 cm s^{-1} (28.91 $\pm$ 21.09 cm s^{-1}), while for the altimeter-derived AE (CE), it is 29.30 $\pm$ 19.92 cm s^{-1} (28.77 $\pm$ 19.88 cm s^{-1}). These quantitative analyses provide further evidence that Argo-only identified eddies exhibit comparable current speeds and circulation structures to those identified by the altimeter but with larger current speed deviations, indicating that the Argo-only eddies have a certain degree of instability which makes the proposed algorithm an ideal complement to traditional eddy identification methods.

Another way to assess the validity of eddies derived by Argo alone is to compare their PDAs with those within altimeter-captured eddies. To do so, the vertical distributions of the mean PDA in ascending eddy property levels for both Argo-only eddies and ALT eddies are shown in Fig. 11. It can be clearly found that the vertical distributions derived from Argo-only eddies are strikingly similar to the altimeter-collocated results, for which the stronger and larger the eddy is, the more anomalous its PDA is likely to be. More importantly, the PDA distributions of Argo-only eddies with the variation of inferred eddy surface parameters are remarkably consistent with the PDA distributions of ALT eddies with the change in SLA-derived eddy parameters, implying that identified and characterized Argo-only eddies are as robust and stable as altimeter ones for both AEs and CEs. In addition to the global annual average, the PDA profiles of Argo-only identified eddies in different seasons and typical regions are examined, as well as regarding the position of the eddy core (i.e., surface/subsurface intensified eddies) (please see the appendix for details). As pointed out by G. Chen et al. (2021a), an ideal strategy for eddy observation is considered the joint use of Argo and altimeter techniques which are highly complementary in terms of
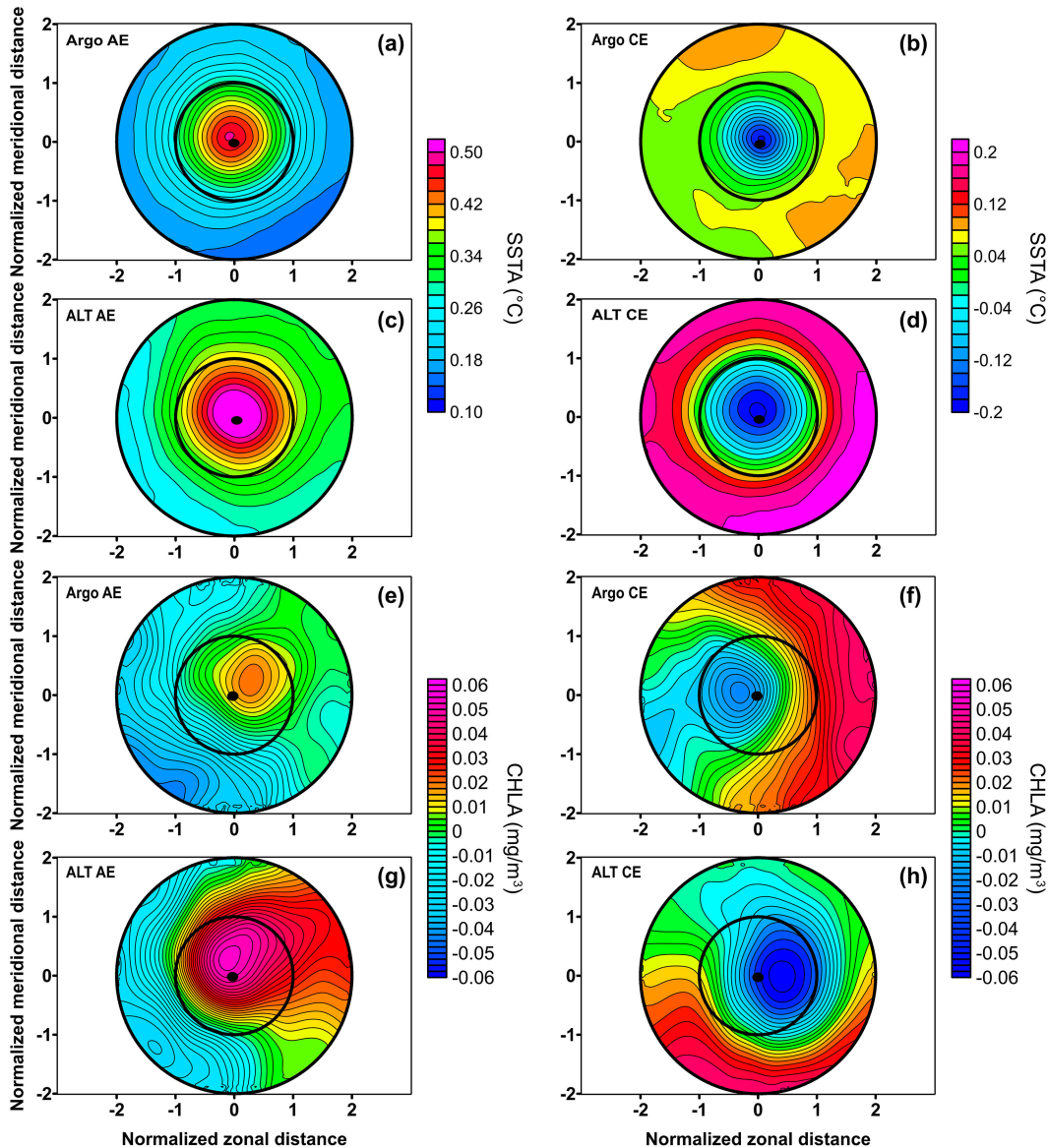
FIG. 9. (a),(b) SSTA distribution inside Argo-identified eddies after circular normalization for (a) AE and (b) CE. (c),(d) As in (a) and (b), but are derived from altimeter-identified eddies. (e)–(h) As in (a)–(d), but for the CHLA distribution in the eastern southern Indian Ocean.

revealing simultaneously the interior structure and surface manifestation.

d. Practical contribution and implication of the Argo-inverted eddy algorithm

A comprehensive understanding of the potential applications of the Argo-based eddy identification and inversion methodology is essential. Given the strong entraining capability of swirling eddies, Argo floats emerge as useful tracers of objects entrapped within eddies. For instance, Xu et al. (2016) observed the continuous entrapment of an Argo float within an anticyclonic eddy along the Kuroshio Extension for 167 days, while Ma et al. (2020) reported an Argo float persistently observing a target

eddy for over 2 months. Consequently, the capability of eddy tracking based on the Argo-only eddy detection scheme is further verified to provide a comprehensive application of the proposed methodology. Figure 12 displays two Argo trajectories within the Argo-only eddies, which escape detection and continuous tracking by the altimeter constellation but are successfully captured by Argo PDA profiles. Employing the proposed Argo-based algorithm, a westward-propagating AE is consistently tracked by Argo profiles for more than 2 years, spanning from 9 March 2003 to 6 June 2005 (Figs. 12a–c), featuring strong negative PDA signatures at around 180-m depth. A similar result can be found in a westward-propagating Argo-only CE track in the Northern Hemisphere from 25 June 2007 to 18 July 2008,
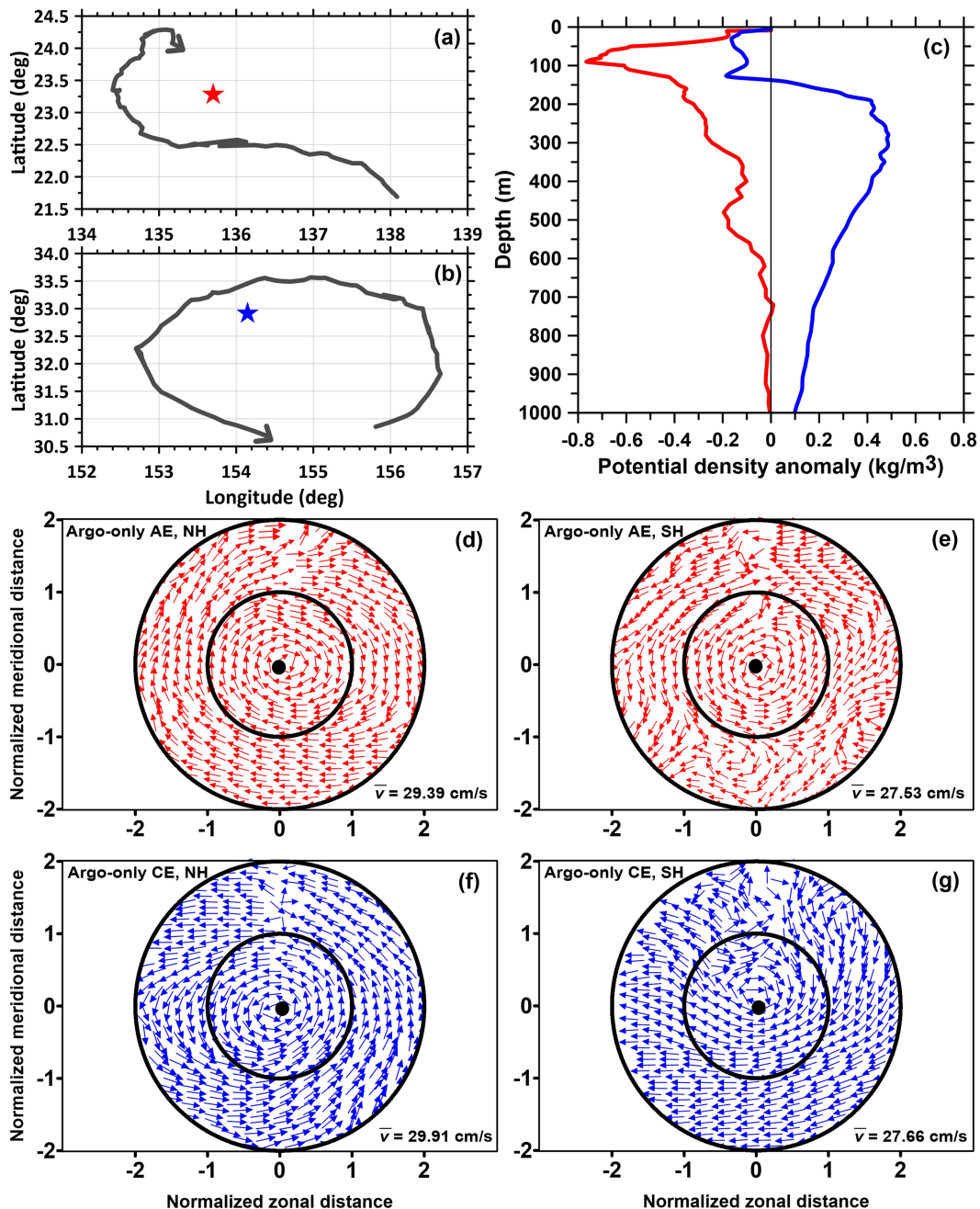
FIG. 10. Current field validation of the Argo-only derived eddies based on the drifter buoys. Argo-only identified eddies and the concurrent drifter trajectories (±10 days) for (a) AE and (b) CE. The red (blue) star marks the Argo-only AE (CE), and the gray lines with arrows represent the drifter trajectories. (c) The corresponding PDA profiles in (a) and (b). The red (blue) line is AE (CE). Normalized distributions of the current field in Argo-only AEs during 2002–19 in the (d) Northern Hemisphere and (e) Southern Hemisphere. The arrows point to the direction of the current field. The scales of the arrows represent the current velocity, with the average velocity _v labeled in the bottom right corner. (f),(g) As in (d) and (e), but for Argo-only CEs.

characterized by positive PDA values (Figs. 12d–f). The capability of eddy tracking based on the Argo-only eddy detection scheme would also be helpful to partially tackle the problem of the "ghost eddies," which is when altimeter-tracked eddies

disappear between two parallel altimeter tracks during the eddy tracking process.

Apart from the ability to consistently track eddies that are not capturable by the altimeters, the Argo-based eddy
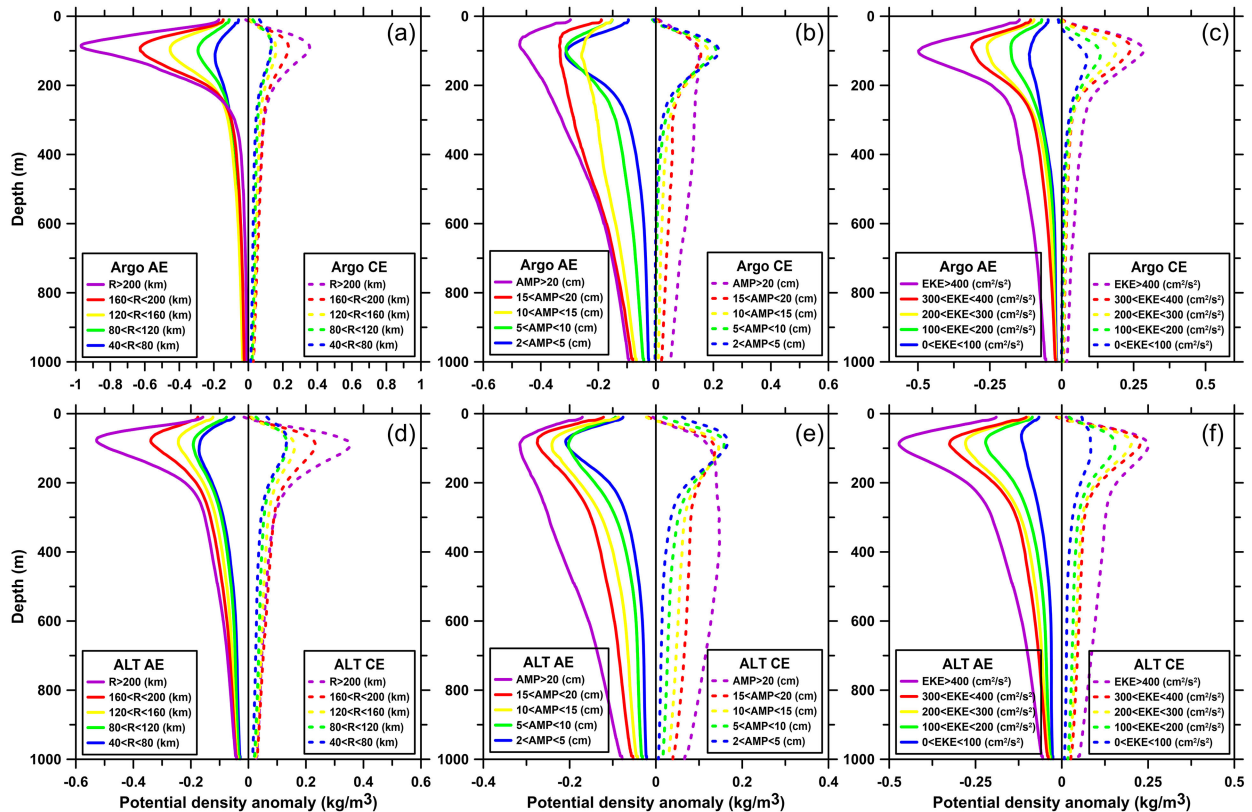
FIG. 11. Vertical distribution of mean PDA derived from Argo-identified eddies with respect to corresponding inverted (a) eddy radius, (b) eddy amplitude, and (c) eddy kinetic energy. (d)–(f) As in (a)–(c), but are derived from altimeter-identified eddies.

detection method would have significant practical implications relevant to operational ocean monitoring and biogeochemistry oceanographic research. On the one hand, the implementation of the proposed algorithm proves crucial for practical marine operations and environmental monitoring. In many instances, for ocean in situ observations, the ability to assess eddy characteristics based on in situ profiling data in real time proves highly advantageous, which will also facilitate timely strategic arrangements. More importantly, the proposed method stands out as a valuable instrument for tracking eddies, particularly in the context of the increasing use of autonomous underwater vehicles in contemporary field experiments. It is believed that the proposed methodology extends beyond its applicability solely to Argo floats and can be adapted to various observations such as CTD and XBT. Furthermore, the utility of the profiles is not confined to potential densities but can be effectively employed with temperature or salinity as well. Researchers can refer to the modeling method developed in this study to train the model using the corresponding profile data, so as to establish targeted vertical-surface connections and independently track eddies. On the other hand, given that eddies exhibit not only typical vertical structures in physical parameters such as temperature, salinity, and density but also vertical biochemistry features such as chlorophyll, dissolved oxygen, and backscatter coefficient, the inversion of eddy surface features by PDA profile features, as

presented in this paper, can provide information about the vertical structure of eddy biochemistry. This extension establishes a link between the eddy surface signal and its vertical biochemical structure, holding crucial implications for in situ observations of eddy ecological impact.

5. Conclusions

The high correlation between an eddy's surface properties and its vertical features provides a unique opportunity for global eddy characterization. Thanks to the simultaneous availability of Argo PDA profiles, and taking advantage of the machine learning method, this study develops an independent and robust Argo-based eddy property inversion scheme to evaluate eddy surface attributes from the vertical features, which significantly complement and expand the present altimeter-eddy detection and provide a special perspective of Argo-eddy observation.

The PDA profiles allocated to altimeter-identified eddies during 2002–09 are examined individually and fed into a novel ensemble learning model to extract detailed and sensitive vertical signals for eddies of both cyclonic and anticyclonic types. This altimeter-calibrated EL model has promising performance, with the average R^2 (MAE) of the eddy radius, amplitude, and kinetic energy reaching the optimal value of 0.90 (5.01 km), 0.91 (0.54 cm), and 0.94 (11.80 cm^2 s^{-2}) for AEs and 0.88 (5.86 km), 0.91 (0.53 cm), and 0.91 (16.77 cm^2 s^{-2}) for
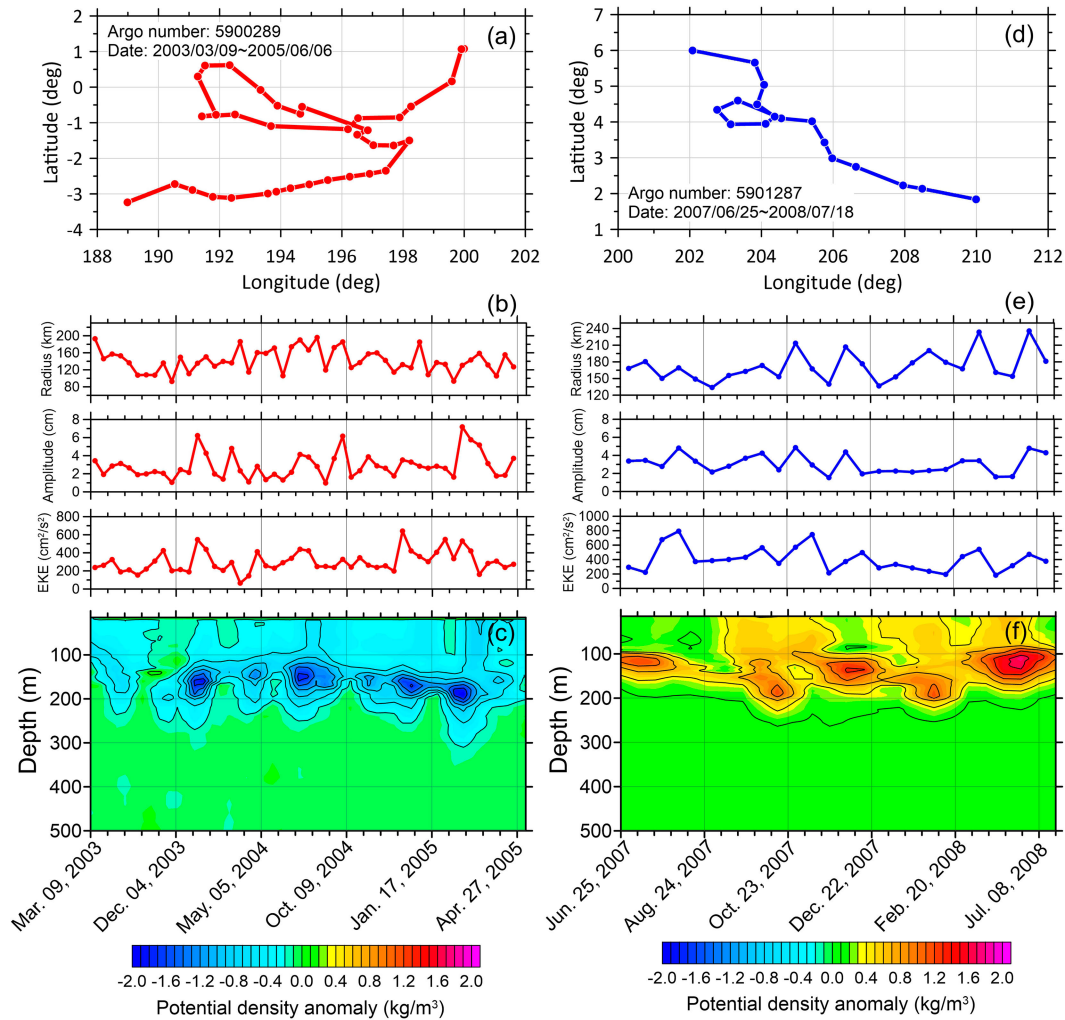
FIG. 12. (a) An Argo-only identified AE tracked by Argo float (No. 5900289) during 9 Mar 2003–6 Jun 2005. (b) The inverted eddy radius, amplitude, and kinetic energy with respect to the time series. (c) The corresponding PDA contour map with respect to the time series. (d)–(f) As in (a)–(c), but for an Argo-only identified CE during 25 Jun 2007–8 Jul 2008.

CEs, respectively. It should be noted that the eddy properties of AEs can be estimated more accurately compared with those of CEs. A plausible explanation for this phenomenon is that the concave surface sea level anomaly signals generated by cyclonic eddies may induce a convergence effect in the reflected electromagnetic waves from the sea surface (G. Chen et al. 2021a). Such multiple coherent reflections are likely to be noisier compared to AEs for altimeter observation, making the joint Argo–altimeter scheme appear to be more effective for AEs. The proposed EL inversion model is further validated by independently estimating characteristics of ALT eddies during 2010–19 and shows satisfactory accuracy when compared with those computed from SLA fields, with global mean differences of 3.1 km, 0.8 cm, and 9.4 cm² s⁻² for eddy radius, amplitude, and kinetic energy, respectively.

The EL inversion model is further applied to the eddies recaptured by Argo floats. It is demonstrated that the Argo-independent eddy detection is an effective spatial complementary and property

expansion to that of altimetry. The global patterns of the inverted three parameters of Argo-only eddies are comparable to ALT eddies. Areas filled with relatively high-energy eddies are concentrated in the western boundary currents such as the Kuroshio Extension, the Gulf Stream, and the ACC. In particular, Argo-identified likely eddies are systematically 25.2% smaller and weaker compared with ALT eddies, and the most distinct magnitude differences are observed particularly in instability flow systems and the tropical oceans. The most significant property difference between Argo- and altimeter-derived eddy is found to be in the eddy amplitude, which appears to be 0.47 times "shorter" of Argo-only eddy than the ALT eddy in terms of global mean, indicating the extended ability of the proposed Argo-based methodology for altimeter eddy observation. Meanwhile, the 2002–19 annual variations for the ALT- and Argo-derived eddy parameters are found to be essentially identical both globally and in typical regions. This spatial and temporal consistency of the two further proves the reliability of the Argo-based inversion results.

Given the reasonable results obtained in this study, the proposed idea of independent eddy surface signature inversion is expected to be effective and efficient for the vertical profiles obtained by various in situ observation floats [e.g., the Biogeochemical Argo (BGC-Argo), the expendable bathythermographs, and the subsurface mooring array] to further derive other eddy dynamics and biochemical characteristics. It can be independently determined in near–real time whether there exist eddies, as well as the corresponding eddy features, which will provide a significant technical methodology and open a new window for in situ eddy observations. Future exploration based on this feasibility study could be done to improve the algorithm of the eddy tracking in combination with Argo profiling observations (e.g., ghost eddy problem with temporary disappearance of eddies during eddy tracking) and hopefully allow a more accurate prediction of eddy trajectory based on the vertical profiling data.

Data availability statement. The datasets used to support this study are available as follows. The altimeter all-sat merged sea level anomaly data can be downloaded from https://data.marine.copernicus.eu/product/SEALEVEL_GLO_PHY_L4_MY_008_047/description. The altimeter-derived eddy dataset can be obtained from https://data.casearth.cn/en/sdo/detail/5fa668ad1f4600005e005ba3 (Data ID: XDA19090202). The original Argo data are provided by IFREMER at ftp://ftp.ifremer.fr/ifremer/argo (Argo 2022). The daily Optimum Interpolation Sea Surface Temperature (OISST) products are provided by the NOAA (available at https://www.ncei.noaa.gov/products/optimum-interpolation-sst). The near-surface CHLA data are provided by MODIS and can be downloaded from http://oceandata.sci.gsfc.nasa.gov/. The drifter-derived current data are provided by the Atlantic Oceanographic and Meteorological Laboratory of NOAA (available at ftp://ftp.aoml.noaa.gov/pub/phod/buoydata/). The experimental environment is performed on a computer with Windows 10, and Python is the primary programming language. In the spirit of reproducibility, the Python code will be released online in conjunction with the publication of the paper.

APPENDIX

Further Validation of the Vertical Structures of Argo-Only Eddies

To provide a comprehensive verification of the Argo-only identified eddies, we further examine their vertical structures in terms of seasonal variations, typical regions, and eddy core positions (i.e., surface- or subsurface-intensified eddies). Results show that Argo-only eddies have distinct vertical signatures in different seasons and correlate significantly with the inverted eddy radius for both anticyclonic and cyclonic eddies, indicating that the proposed algorithm maintains seasonal stability and time-dependent robustness (Fig. A1). Argo-only eddies display unique vertical features specific to different typical regions which consistently align with the vertical features of altimeter-identified eddies (Fig. A2). Furthermore, it is understood that eddies could be categorized into two types according to the depth of extreme values of PDA: surface intensified with an eddy core near the sea surface or subsurface intensified with an eddy core below the surface (Assassi et al. 2016; de Marez et al. 2019). Figure A3 illustrates Argo's capacity to identify and characterize both types of eddies whose extreme values of PDA are located at ~50 m (Fig. A3a) and below ~100 m (Fig. A3b), with surface properties inverted by the PDA profiles of both eddies intuitively linked to their vertical structure. Moreover, it is not only on a global scale but also within a representative region, such as the Arabian Sea, and is comparable with previous studies (e.g., de Marez et al. 2019). These results provide a robust and multifaceted demonstration of the effectiveness of our proposed algorithm, thereby confirming the reliability of Argo-only eddy vertical structures in diverse contexts.
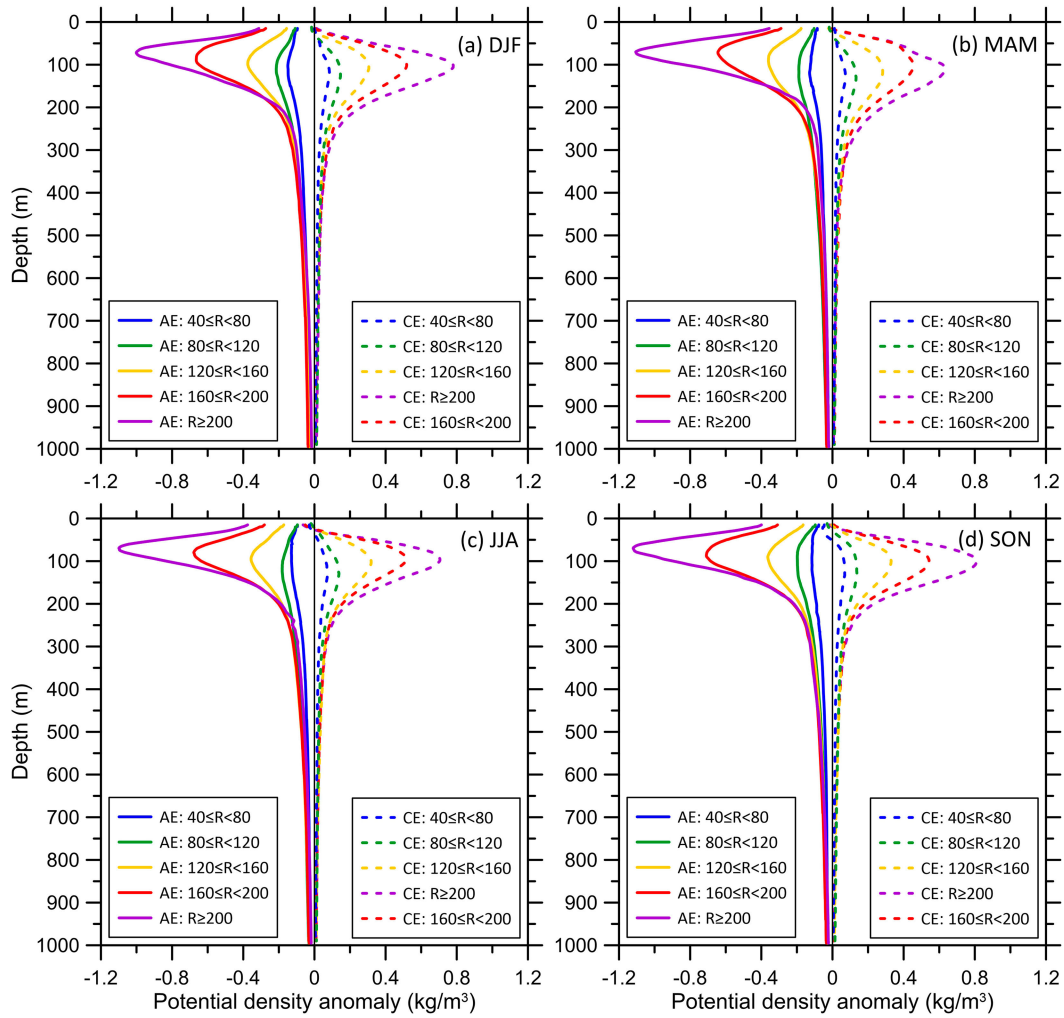
FIG. A1. Seasonal variations of PDA for Argo-only identified AEs (solid lines) and CEs (dashed lines) with respect to inverted eddy radius.
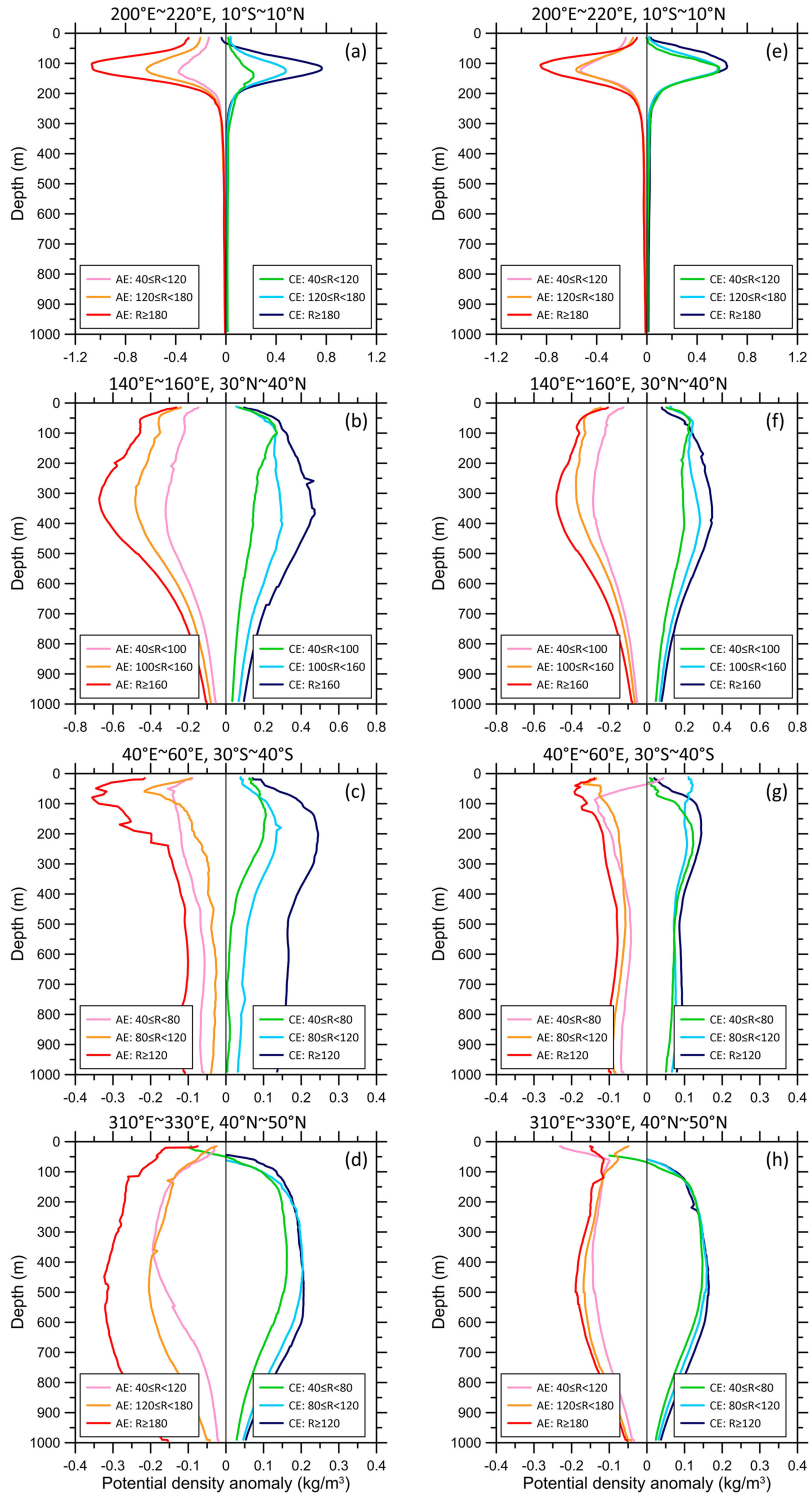
FIG. A2. Vertical distribution of mean PDA derived from Argo-only identified eddies with respect to corresponding inverted eddy radius in (a) the equatorial Pacific, (b) the northwest Pacific, (c) the southern Indian Ocean, and (d) the North Atlantic. (e)–(h) As in (a)–(d), but for altimeter-identified eddies.
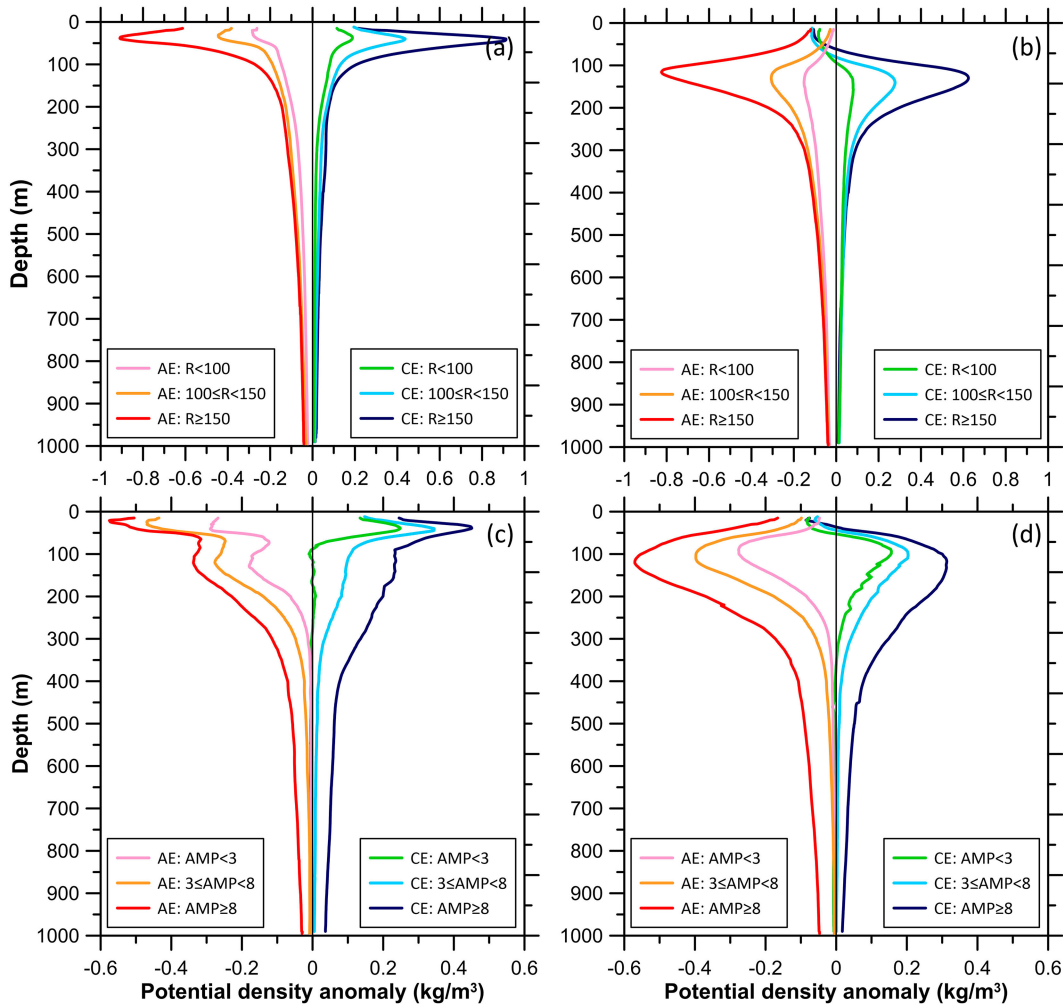
FIG. A3. (a) Vertical distributions of mean PDA derived from Argo-identified surface-intensified eddies with respect to inverted eddy radius for the global ocean. (b) As in (a), but for the subsurface-intensified eddies. (c) Vertical distributions of mean PDA derived from Argo-identified surface-intensified eddies with respect to inverted eddy amplitude for the Arabian Sea (5°–25°N, 30°–80°E). (d) As in (c), but for the subsurface-intensified eddies.

REFERENCES

Amores, A., G. Jordà, T. Arsouze, and J. Le Sommer, 2018: Up to what extent can we characterize ocean eddies using present-day gridded altimetric products? *J. Geophys. Res. Oceans*, **123**, 7220–7236, https://doi.org/10.1029/2018JC014140.

——, ——, and S. Monserrat, 2019: Ocean eddies in the Mediterranean Sea from satellite altimetry: Sensitivity to satellite track location. *Front. Mar. Sci.*, **6**, 703, https://doi.org/10.3389/fmars.2019.00703.

Argo, 2024: Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE, accessed 15 November 2021, https://doi.org/10.17882/42182.

Assassi, C., and Coauthors, 2016: An index to distinguish surface- and subsurface-intensified vortices from surface observations. *J. Phys. Oceanogr.*, **46**, 2529–2552, https://doi.org/10.1175/JPO-D-15-0122.1.

Ballarotta, M., and Coauthors, 2019: On the resolutions of ocean altimetry maps. *Ocean Sci.*, **15**, 1091–1109, https://doi.org/10.5194/os-15-1091-2019.

Banzon, V., T. M. Smith, T. M. Chin, C. Liu, and W. Hankins, 2016: A long-term record of blended satellite and in situ sea-surface temperature for climate monitoring, modeling and environmental studies. *Earth Syst. Sci. Data*, **8**, 165–176, https://doi.org/10.5194/essd-8-165-2016.

Barceló-Llull, B., E. Pallàs-Sanz, P. Sangrà, A. Martínez-Marrero, S. N. Estrada-Allis, and J. Arístegui, 2017: Ageostrophic secondary circulation in a subtropical intrathermocline eddy. *J. Phys. Oceanogr.*, **47**, 1107–1123, https://doi.org/10.1175/JPO-D-16-0235.1.

Breiman, L., 1996: Stacked regressions. *Mach. Learn.*, **24**, 49–64, https://doi.org/10.1007/BF00117832.

Callies, J., and R. Ferrari, 2013: Interpreting energy and tracer spectra of upper-ocean turbulence in the submesoscale range (1–200 km). *J. Phys. Oceanogr.*, **43**, 2456–2474, https://doi.org/10.1175/JPO-D-13-063.1.

Cao, C., G. Chen, and X. Wang, 2022: How mesoscale eddies masquerade as Rossby waves in merged altimetric products. *J. Geophys. Res. Oceans*, **127**, e2022JC018981, https://doi.org/10.1029/2022JC018981.

Chaigneau, A., M. L. Texier, G. Eldin, C. Grados, and O. Pizarro, 2011: Vertical structure of mesoscale eddies in the eastern South Pacific Ocean: A composite analysis from altimetry and Argo profiling floats. *J. Geophys. Res.*, **116**, C11025, https://doi.org/10.1029/2011JC007134.

Chelton, D. B., M. G. Schlax, and R. M. Samelson, 2011a: Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.*, **91**, 167–216, https://doi.org/10.1016/j.pocean.2011.01.002.

——, P. Gaube, M. G. Schlax, J. J. Early, and R. M. Samelson, 2011b: The influence of nonlinear mesoscale eddies on near-surface oceanic chlorophyll. *Science*, **334**, 328–332, https://doi.org/10.1126/science.1208897.

Chen, G., G. Han, and X. Yang, 2019: On the intrinsic shape of oceanic eddies derived from satellite altimetry. *Remote Sens. Environ.*, **228**, 75–89, https://doi.org/10.1016/j.rse.2019.04.011.

——, X. Chen, and B. Huang, 2021a: Independent eddy identification with profiling Argo as calibrated by altimetry. *J. Geophys. Res. Oceans*, **125**, e2020JC016729, https://doi.org/10.1029/2020JC016729.

——, ——, and C. Cao, 2022: Divergence and dispersion of global eddy propagation from satellite altimetry. *J. Phys. Oceanogr.*, **52**, 705–722, https://doi.org/10.1175/JPO-D-21-0122.1.

Chen, X., and G. Chen, 2023: Quantifying the degree of eddy quasi-geostrophy by generalizing Rossby deformation. *J. Geophys. Res. Oceans*, **128**, e2023JC019652, https://doi.org/10.1029/2023JC019652.

——, ——, L. Ge, B. Huang, and C. Cao, 2021a: Global oceanic eddy identification: A deep learning method from Argo profiles and altimetry data. *Front. Mar. Sci.*, **8**, 646926, https://doi.org/10.3389/fmars.2021.646926.

——, H. Li, C. Cao, and G. Chen, 2021b: Eddy-induced pycnocline depth displacement over the global ocean. *J. Mar. Syst.*, **221**, 103577, https://doi.org/10.1016/j.jmarsys.2021.103577.

CMEMS, 2019: Product user manual for sea level SLA products. Copernicus Marine Service Product User Manual CMEMS-SL-PUM-008-032-062, 39 pp., http://marine.copernicus.eu/documents/PUM/CMEMS-SL-PUM-008-032-062.pdf.

de Marez, C., P. L'Hégaret, M. Morvan, and X. Carton, 2019: On the 3D structure of eddies in the Arabian Sea. *Deep-Sea Res. I*, **150**, 103057, https://doi.org/10.1016/j.dsr.2019.06.003.

Dong, C., F. Nencioli, Y. Liu, and J. C. McWilliams, 2011: An automated approach to detect oceanic eddies from satellite remotely sensed sea surface temperature data. *IEEE Geosci. Remote Sens. Lett.*, **8**, 1055–1059, https://doi.org/10.1109/LGRS.2011.2155029.

——, J. C. McWilliams, Y. Liu, and D. Chen, 2014: Global heat and salt transports by eddy movement. *Nat. Commun.*, **5**, 3294, https://doi.org/10.1038/ncomms4294.

Faghmous, J. H., I. Frenger, Y. Yao, R. Warmka, A. Lindell, and V. Kumar, 2015: A daily global mesoscale ocean eddy dataset from satellite altimetry. *Sci. Data*, **2**, 150028, https://doi.org/10.1038/sdata.2015.28.

Gaube, P., D. B. Chelton, P. G. Strutton, and M. J. Behrenfeld, 2013: Satellite observations of chlorophyll, phytoplankton biomass, and Ekman pumping in nonlinear mesoscale eddies. *J. Geophys. Res. Oceans*, **118**, 6349–6370, https://doi.org/10.1002/2013JC009027.

——, D. J. McGillicuddy Jr., D. B. Chelton, M. J. Behrenfeld, and P. G. Strutton, 2014: Regional variations in the influence of mesoscale eddies on near-surface chlorophyll. *J. Geophys. Res. Oceans*, **119**, 8195–8220, https://doi.org/10.1002/2014JC010111.

Heinermann, J., and O. Kramer, 2016: Machine learning ensembles for wind power prediction. *Renewable Energy*, **89**, 671–679, https://doi.org/10.1016/j.renene.2015.11.073.

Huang, B., L. Ge, X. Chen, and G. Chen, 2021: Vertical structure-based classification of oceanic eddy using 3-D convolutional neural network. *IEEE Trans. Geosci. Remote Sens.*, **60**, 4203614, https://doi.org/10.1109/TGRS.2021.3103251.

Kang, D., and E. N. Curchitser, 2013: Gulf Stream eddy characteristics in a high-resolution ocean model. *J. Geophys. Res. Oceans*, **118**, 4474–4487, https://doi.org/10.1002/jgrc.20318.

Keppler, L., S. Cravatte, A. Chaigneau, C. Pegliasco, L. Gourdeau, and A. Singh, 2018: Observed characteristics and vertical structure of mesoscale eddies in the southwest tropical Pacific. *J. Geophys. Res. Oceans*, **123**, 2731–2756, https://doi.org/10.1002/2017JC013712.

Kim, T., and W. D. Lee, 2022: Review on applications of machine learning in coastal and ocean engineering. *J. Ocean Eng. Technol.*, **36**, 194–210, https://doi.org/10.26748/KSOE.2022.007.

Liu, Y., G. Chen, M. Sun, S. Liu, and F. Tian, 2016: A parallel SLA-based algorithm for global mesoscale eddy identification. *J. Atmos. Oceanic Technol.*, **33**, 2743–2754, https://doi.org/10.1175/JTECH-D-16-0033.1.

Lumpkin, R., and L. Centurioni, 2019: NOAA Global Drifter Program quality-controlled 6-hour interpolated data from ocean surface drifting buoys. NOAA National Centers for Environmental Information, accessed 6 May 2022, https://doi.org/10.25921/7ntx-z961.

Ma, C., Z. Gao, S. Li, S. Li, and G. Chen, 2020: An eddy-borne Argo float measurement experiment in the South China Sea. *Ocean Dyn.*, **70**, 1325–1338, https://doi.org/10.1007/s10236-020-01402-3.

Mason, E., A. Pascual, and J. C. McWilliams, 2014: A new sea surface height–based code for oceanic mesoscale eddy tracking. *J. Atmos. Oceanic Technol.*, **31**, 1181–1188, https://doi.org/10.1175/JTECH-D-14-00019.1.

McDougall, T. J., and Coauthors, 2009: The International Thermodynamic Equation Of Seawater—2010: Calculation and use of thermodynamic properties. Intergovernmental Oceanographic Commission, Manuals and Guides 56, 220 pp., https://www.teos-10.org/pubs/TEOS-10_Manual.pdf.

Mendes-Moreira, J., C. Soares, A. M. Jorge, and J. F. D. Sousa, 2012: Ensemble approaches for regression: A survey. *ACM Comput. Surv.*, **45**, 1–40, https://doi.org/10.1145/2379776.2379786.

Pegliasco, C., A. Delepoulle, E. Mason, R. Morrow, Y. Faugère, and G. Dibarboure, 2022: META3.1exp: A new global mesoscale eddy trajectory atlas derived from altimetry. *Earth Syst. Sci. Data*, **14**, 1087–1107, https://doi.org/10.5194/essd-14-1087-2022.

Pujol, M.-I., Y. Faugère, G. Taburet, S. Dupuy, C. Pelloquin, M. Ablain, and N. Picot, 2016: DUACS DT2014: The new multi-mission altimeter data set reprocessed over 20 years. *Ocean Sci.*, **12**, 1067–1090, https://doi.org/10.5194/os-12-1067-2016.

Qi, J., L. Zhang, B. Yin, D. Li, B. Xie, and G. Sun, 2023: Advancing ocean subsurface thermal structure estimation in the Pacific Ocean: A multi-model ensemble machine learning approach. *Dyn. Atmos. Oceans*, **104**, 101403, https://doi.org/10.1016/j.dynatmoce.2023.101403.

Qiu, B., 1999: Seasonal eddy field modulation of the North Pacific subtropical countercurrent: TOPEX/Poseidon observations and theory. *J. Phys. Oceanogr.*, **29**, 2471–2486, https://doi.org/10.1175/1520-0485(1999)029<2471:SEFMOT>2.0.CO;2.

——, S. Chen, P. Klein, J. Wang, H. Torres, L.-L. Fu, and D. Menemenlis, 2018: Seasonality in transition scale from balanced to unbalanced motions in the world ocean. *J. Phys. Oceanogr.*, **48**, 591–605, https://doi.org/10.1175/JPO-D-17-0169.1.

Ridgway, K. R., J. R. Dunn, and J. L. Wilkin, 2002: Ocean interpolation by four-dimensional weighted least squares—Application to the waters around Australasia. *J. Atmos. Oceanic Technol.*, **19**, 1357–1375, https://doi.org/10.1175/1520-0426(2002)019<1357:OIBFDW>2.0.CO;2.

Sun, W., C. Dong, R. Wang, Y. Liu, and K. Yu, 2017: Vertical structure anomalies of oceanic eddies in the Kuroshio Extension region. *J. Geophys. Res. Oceans*, **122**, 1476–1496, https://doi.org/10.1002/2016JC012226.

Tian, F., D. Wu, L. Yuan, and G. Chen, 2020: Impacts of the efficiencies of identification and tracking algorithms on the statistical properties of global mesoscale eddies using merged altimeter. *Int. J. Remote Sens.*, **41**, 2835–2860, https://doi.org/10.1080/01431161.2019.1694724.

Weisheimer, A., F. J. Doblas-Reyes, T. N. Palmer, A. Alessandri, A. Arribas, M. Déqué, and P. Rogel, 2009: ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions—Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.*, **36**, L21711, https://doi.org/10.1029/2009GL040896.

Xu, L., P. Li, S.-P. Xie, Q. Liu, C. Liu, and W. Gao, 2016: Observing mesoscale eddy effects on mode-water subduction and transport in the North Pacific. *Nat. Commun.*, **7**, 10505, https://doi.org/10.1038/ncomms10505.

——, N. Chen, X. Zhang, and Z. Chen, 2020: A data-driven multi-model ensemble for deterministic and probabilistic precipitation forecasting at seasonal scale. *Climate Dyn.*, **54**, 3355–3374, https://doi.org/10.1007/s00382-020-05173-x.

Zhang, Z., Y. Zhang, W. Wang, and R. X. Huang, 2013: Universal structure of mesoscale eddies in the ocean. *Geophys. Res. Lett.*, **40**, 3677–3681, https://doi.org/10.1002/grl.50736.

Zheng, S., M. Feng, Y. Du, X. Cheng, and J. Li, 2016: Annual and interannual variability of the tropical instability vortices in the equatorial eastern Pacific observed from Lagrangian surface drifters. *J. Climate*, **29**, 9163–9177, https://doi.org/10.1175/JCLI-D-16-0124.1.

——, ——, ——, X. Meng, and W. Yu, 2018: Interannual variability of eddy kinetic energy in the subtropical southeast Indian Ocean associated with the El Niño-Southern Oscillation. *J. Geophys. Res. Oceans*, **123**, 1048–1061, https://doi.org/10.1002/2017JC013562.

Zlotnicki, V., L.-L. Fu, and W. Patzert, 1989: Seasonal variability in global sea level observed with Geosat altimetry. *J. Geophys. Res.*, **94**, 17 959–17 969, https://doi.org/10.1029/JC094iC12p17959.