ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 22, NO. 5, 2005, 677–684

Evaluation of Mid-Depth Currents of NCEP Reanalysis Data in the Tropical Pacific Using ARGO Float Position Information

XIE Jiping*[1,2] (谢基平), ZHU Jiang[1] (朱 江), XU Li [1] (许 黎), and GUO Pinwen[2] (郭品文)

[1]*Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing* 100029

[2]*Department of Meteorology, Nanjing University of Information Science & Technology, Nanjing* 210044

(Received 11 October 2004; revised 13 June 2005)

ABSTRACT

The global project of the Array for Real-time Geostrophic Oceanography (ARGO) provides a unique opportunity to observe the absolute velocity in mid-depths of the world oceans. A total of 1597 velocity vectors at 1000 (2000) db in the tropical Pacific derived from the ARGO float position information during the period November 2001 to October 2004 are used to evaluate the intermediate currents of the National Centers for Environmental Prediction reanalysis. To derive reliable velocity information from ARGO float trajectory points, a rigorous quality control scheme is applied, and by virtue of a correction method for reducing the drift error on the surface in obtaining the velocity vectors, their relative errors are less than 25%. Based on the comparisons from the quantitative velocity vectors and from the space-time average currents, some substantial discrepancies are revealed. The first is that the velocities of the reanalysis at mid-depths except near the equator are underestimated relative to the observed velocities by the floats. The average speed difference between NCEP and ARGO values ranges from about $-2.3\,\mathrm{cm\,s^{-1}}$ to $-1.8\,\mathrm{cm\,s^{-1}}$. The second is that the velocity difference between the ocean model and the observations at 2000 dB seems smaller than that at 1000 dB. The third is that the zonal flow in the reanalysis is too dominant so that some eddies could not be simulated, such as the cyclonic eddy to the east of 160°E between 20°N and 30°N at 2000 dB. In addition, it is noticeable that many floats parking at 1000 dB cannot acquire credible mid-depth velocities due to the time information of their end of ascent (start of descent) on the surface in the trajectory files. Thus, relying on default times of parking, descent and ascent in the metadata files gravely confines their application to measuring mid-depth currents.

Key words: ARGO floats, mid-depth ocean currents, reanalysis, quantitative comparison, mean flow

1. Introduction

The reanalysis dataset (Ji et al., 1995; Kalnay et al., 1996) offered by the National Centers for Environmental Prediction (NCEP) has been applied widely in relevant multidisciplinary research. Along with increasingly extensive observations assimilated in the NCEP reanalysis, the products have become one of the most important datasets for global ocean and climate variation studies. To provide credibility for being adopted in various research works, it is essential for the veracity of the analysis products to be evaluated. However the evaluation studies have been restricted for NCEP ocean reanalysis products due to limited sources of independent observations. Ji and Smith (1995) compared temperature fields. Enfield and Harris (1995) compared sea surface height from the NCEP ocean reanalysis with tide gauge records around the tropical Pacific. In recent years, more evaluation studies have focused on the current (e.g. Acero-Schertzer et al., 1997; Lagerloef et al., 1999). But they only evaluated the near-surface flow of the NCEP analyses of the tropical Pacific Ocean using current observations obtained from satellite-tracked drifting buoys or derived from altimeter. The products, ability to portray intermediate currents of ocean circulation in detail will be helpful to comprehend thermohaline currents and to improve the ability of prognosticating reasonable global climate changes. Davis (1998) analyzed the flow field at 1000 dB using the Autonomous Lagrangian Circulation Explorer (ALACE) float-derived 25-day velocities in the World Ocean Circulation Ex-

*E-mail: xiejp@mail.iap.ac.cn

periment. But in his paper, he paid more attention discussing the Southern Hemisphere ocean currents and focused on climatological current characteristics from those floats.

The project of the Array for Real-time Geostrophic Oceanography (ARGO) aims at building a global array of 3000 free-drifting profiling floats that can measure the temperature and salinity of the upper 2000 m of the ocean. Around the Pacific, ARGO float deployment formally began in the year 2000. Since then, the quality and quantity of ARGO data has grown up. In this paper, we compare the current velocities derived from the ARGO floats at mid-depth including 1000 dB and 2000 dB, with their counterparts in NCEP reanalysis in the tropical Pacific.

2. Data

2.1 *NCEP ocean current data*

This study uses the NCEP Pacific Ocean Analysis data products (Ji et al., 1995; Behringer et al., 1998), which were provided by the Climate Diagnostics Center, Boulder, Colorado, USA, from their web site at http://www.cdc.noaa.gov/. The utilized analysis data contains weekly means for u and v components of ocean currents in the various areas (spatial coverage: 45°N–35°S, 122.25°E–71.25°W) at multiple depths, with a zonal resolution of 1.5° and a meridional resolution of 1.0°. They span the period from 14 October 2001 through 13 November 2004, altogether 161 weeks. In order to compare to the ARGO-derived velocities, the NCEP velocity fields were first interpolated using the Lagrangian scheme in the vertical, which then resulted in velocity component fields at some horizontal level. Then, the velocity fields, which represent the weekly mean, were interpolated linearly to match the time of each ARGO velocity that mostly represents about a 10-day mean. The velocity fields were then horizontally interpolated to ARGO velocity locations.

2.2 *ARGO data*

In an attempt to obtain the ARGO velocity as precisely as possible, the whole ARGO dataset used includes profile data, metadata, trajectory data and technical data, which are available in delayed form from the global data centers (GDACs) from their website (http://www.usgodae.fnmoc.navy.mil/).

An ARGO float drifts for a number of years in the ocean. It continuously performs measurement cycles. Each cycle lasts about 10 days and can be divided into 4 phases: a descent from the surface to a defined pressure (parking pressure), a subsurface drift, an ascending profile with measurements, and a surface drift with data transmission to a communication satellite (see Fig. 1). Obviously there are some series of position and time about an ARGO float when it drifts on the surface. The drifting velocity at parking pressure is calculated according to the surface position of the float where descent begins and where surfacing first occurs (Ichikawa et al., 2001; Davis et al., 1992; Davis and Zenk, 2001).

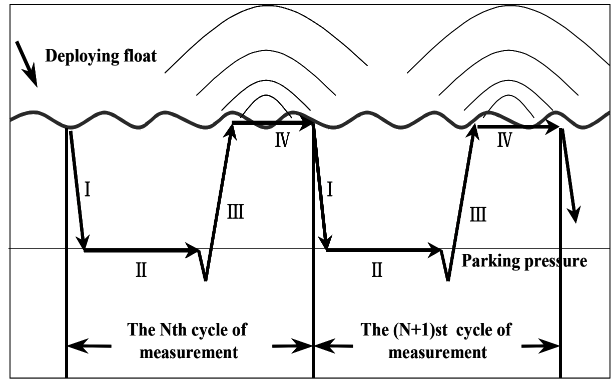

Fig. 1. Schematic of the measurement cycle of an ARGO float which includes four phases: (I) descending, (II) parking, (III) ascending, and (IV) drifting.

3. Estimation of ARGO velocity

3.1 *Estimating velocities from ARGO data*

As the ARGO floats are considered to be moving along with the surrounding water, the time and the position of the float while it is on the surface decided by the ARGO system can be used to estimate the current velocity at the float's parking depth. It is inevitable that the drifting velocity at parking depth may be overestimated by the surface drifting. Moreover the vertical current shear may also affect the estimated velocity. Therefore Ichikawa et al. (2001) estimated the error of drifting velocity for an ARGO float at parking depth in ideal conditions. In addition, they presented the estimated error of current velocity originating from three parts, as follows:

$$\varepsilon = \sqrt{2 \times [\varepsilon_0^2 + (\varepsilon_1 + \varepsilon_2)^2]}\,, \quad (1)$$

where ε_0 is related to the location by the ARGO satellites flying over the float, ε_1 is caused by drifting on the surface before (or after) the position determination can be evaluated, and ε_2 is due to drifting during ascent and decent. Using the outcomes of positioning by the ARGO system would result in a 10%–25% overestimation of the current velocity at parking depth. Of course, their estimation of the mid-depth velocity is not good enough.

3.2 *ARGO data quality control*

The data files from 2202 floats can be downloaded from GDACs, simultaneously containing profile data, metadata and trajectory data. Before 2001, ARGO deployments were sparse. We choose the time period from November 2001 to October 2004. For each float, there are four time/location recordings in a cycle of descending/ascending: the time/location of reaching the sea surface, the time/location just before descent, the time/location upon reaching the surface again, and the time/location just before the next descent. We used the second and the third time/locations to calculate the velocity at its parking depth.

However, the velocities calculated this way could have large errors without careful quality control (QC). These errors could come from various sources: floats may fail to descend and spend their time at the surface; errors from equipment and signal transmission, etc. Our QC measures can be described by the following guidelines: (1) Position-accuracy-flag checking in which all concerned flags of quality on position must be used to identify possible outliers. This is from the point of view of reducing the ε_0 term in (1); (2) Sensor test: if there is a significant omission of records in temperature and salinity profile data within a cycle (e.g., 50%), then the cycle is skipped; (3) A valid cycle test in which the parking pressure in the trajectory data is compared to the maximal pressure measured in the profile data. If the latter is less than 90% of the former, this record will be neglected; (4) A pressure test in which the maximal pressure measured in the profile data is compared to the deepest pressure specified in the metadata. Then if the discrepancy between the two pressures is over 10% relative to the deepest pressure, the cycle should be eliminated; (5) A maximum velocity of 60 cm s^{-1} is set. Any calculated velocity larger than this value will be regarded as a gross error.

3.3 *ARGO data correction*

Although by the QC process, most of the velocities containing gross errors should be checked out, the mid-depth velocity by ARGO positions may be withal contaminated for the drifting on the surface (as in Davis et al., 1992). In addition, every location of an ARGO float on the surface is comprised of a location error when the satellite passes over the float.

In order to obtain as precise velocities in mid-depths as possible, we rely upon an optimum analysis method based upon the principle of least squares, which is detailed in Xie (2005). Using a sequence of surface time-position information recorded in the ARGO files, we not only corrected the effect of the float drifting on the surface affecting the mean velocity, but we were also able to draw out these mid-depth velocities at 1000 (2000) dB with relative errors less than 25%.

From the floats in the tropical Pacific during November 2001 through October 2004 after these methods are applied, we obtained 1597 velocity vectors at the two depths as shown in Fig. 2. Obviously, in the figure, the velocities at 1000 dB are sparse, only having 392 observations, which mainly results from the following two facts. First, many floats parked at 1000 dB cannot provide the time in the trajectory when ending (starting) their ascent (descent) in the measurement

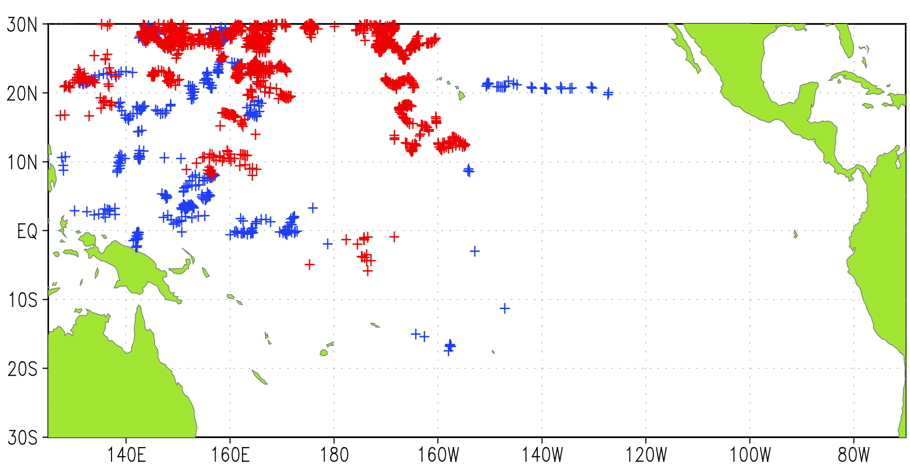

Fig. 2. Descent positions in valid point pairs by which we are able to calculate the relative error of the mid-depth velocities as less than 25% during November 2001 to October 2004. The red crosses denote 2000 dB floats, the blue denote 1000 dB floats.

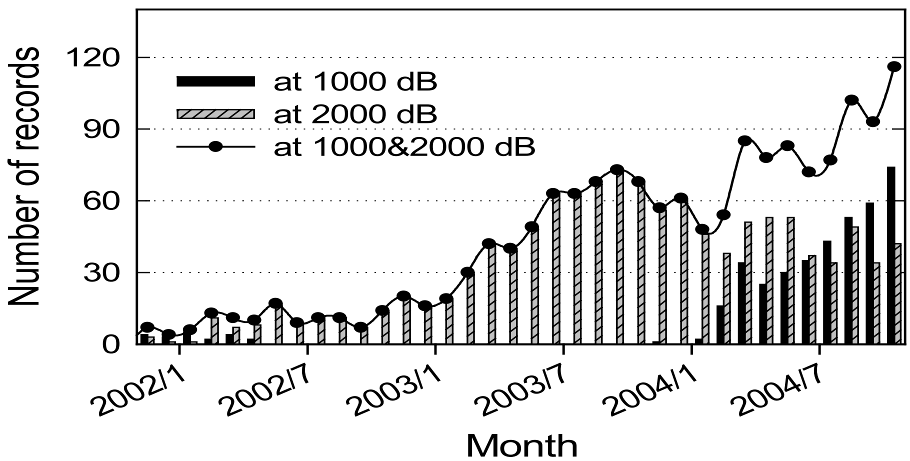

Fig. 3. The number of velocity records in the tropical Pacific along each month during November 2001 to October 2004. The black (grey) histograms represent velocity records at 1000 (2000) dB. The solid line is the total number of records for both 1000 dB and 2000 dB.

cycle. Secondly, some floats parked at 1000 dB in their metadata do not afford the time period information about the float ascent, descent and parking when submerged, so there are 1575 corrected velocities not offering the relative error estimate. Of course, the situation of floats at 2000 dB is better.

The distribution of the mid-depth velocity vectors in Fig. 2 for each month can be seen in Fig. 3. With the promotion of the ARGO project, the observations of credible velocities are increasingly enriched. It is noticeable that the 1000 dB velocities after 2004 are rapidly increased, which suggests that the ARGO floats parking at 1000 dB are more and more important in our analysis.

4. Results and discussions

4.1 *Quantitative comparison*

By interpolating at every descent position (as in Fig. 2), we can obtain the velocity depicted by NCEP, but the velocity components at mid-depths by NCEP differ from the velocities by ARGO measurements in magnitude. At 1000 dB, there are only 18 descent points where the departure of the u component between the reanalysis and the observations is within the range of 25% of the measured velocity. For the v component, there are only 10, and there are even no points at which the departures of the u and v components between the two datasets are simultaneously less than 25% of the observed velocity components. And at 2000 dB, the descent points for the u component number 79, while the v component they number only 53, but there is one point where the u and v components are within the 25% scope. Based on the above comparison, it is assured that the difference of mid-depth velocities between the two datasets is rather obvious in magnitude at every point of observation.

Thus, it is necessary to realize their statistical discrepancy with regard to these aspects at 1000 and 2000 dB (see Fig. 4), which comes from subtracting the ARGO value from the NCEP estimation at the descent position. Figure 4a shows the distribution of the deviation of the u component at 1000 and 2000 dB: at 1000 dB the spread of the difference is asymmetric and the average difference is about -1.5 cm s^{-1}, but for the u component difference at 2000 dB the spread is approximately normal, and the deviation lying between -4 cm s^{-1} and 4 cm s^{-1} is greater than 86% of all the 1205 velocity components at this depth, which is far from the distribution at 1000 dB. For the v component difference, the distributions at the two depths partly resemble each other, where at 2000 dB, the deviation between -4 cm s^{-1} and 4 cm s^{-1} is more concentrated with a value of about 89%, which is higher than the 74% value for 1000 dB in Fig. 4b. As for the mid-depth velocity magnitude at every descent point, the deviation distribution between the two datasets is obviously asymmetric in Fig. 4c, where most of the differences are negative. The average deviation of the velocity magnitude at 1000 dB is about -2.3 cm s^{-1}, but at 2000 dB it is about -1.8 cm s^{-1}. In Fig. 4d, the distributions of the direction deviation at 1000 dB and 2000 dB are almost homogeneous from $-180°$ to $180°$, especially at 2000 dB.

Consequently from the above comparison, there is an obvious discrepancy between the reanalysis and the observations regarding the velocity vectors at 1000 and

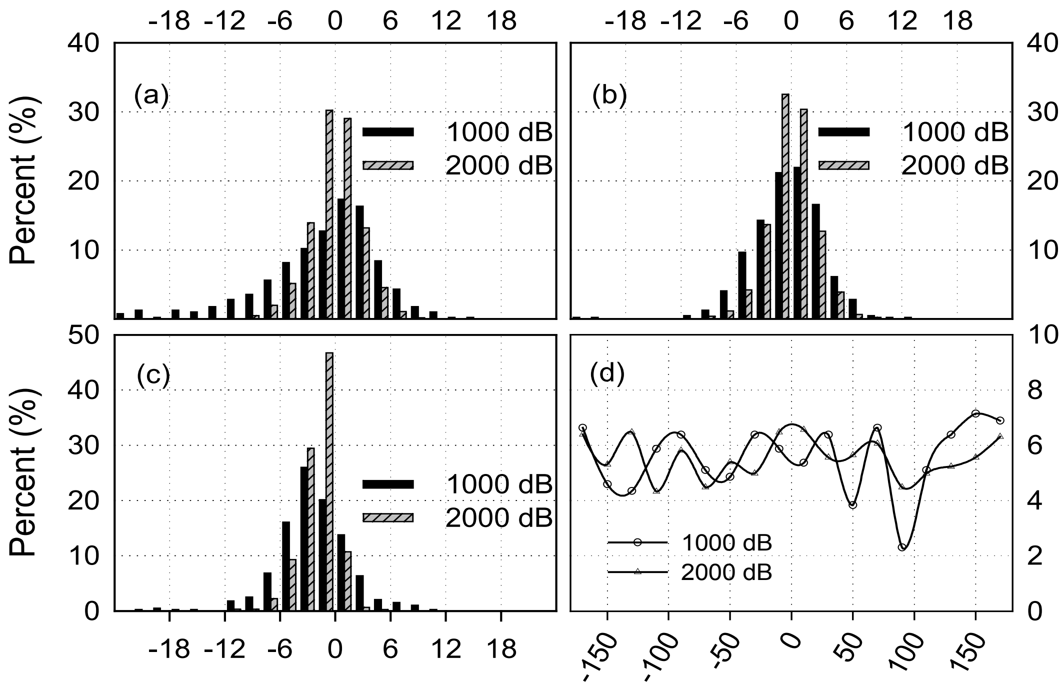

Fig. 4. Frequency distribution of the deviations between the reanalysis and the measurements at 1000 and 2000 dB. (a) u component (units: cm s^{-1}). (b) v component (units: cm s^{-1}). (c) magnitude of velocity vector (units: cm s^{-1}). (d) the direction of velocity vector (units: °), where a negative value means the velocity direction by NCEP is clockwise relative to the velocity direction by ARGO.

2000 dB. On average, the NCEP velocities are underestimated by 2.3 (or 1.8) cm s^{-1} at 1000 (2000) dB, and the distributions of their $u(v)$ component differences suggest that the reanalysis velocities at 2000 dB differ less with the observation than those at 1000 dB. In addition, the direction deviations of velocity vectors at 1000 and 2000 dB seem random.

4.2 *Mean flow comparison*

Of course, the above quantitative comparisons are mainly based on the point of Lagrange velocity, thus in order to explore the circulation deficiency by the reanalysis at mid-depth, it is necessary to compare their Eulerian currents. Because all the trajectory points are inhomogeneous in space as in Fig. 2, and since they are not simultaneously observed, it is difficult to directly compare the NCEP and ARGO currents for their sparse density. Toner et al. (2001) presented that with appropriate coverage, drifter data could provide accurate basin-scale reconstruction of Eulerian velocity fields. With enough data, one can obtain the mean flow fields and their variances which are statistically rigorous (e.g., Davis, 1998; Chapman et al., 2003).

So, in this paper, the mid-depth velocities at the descending point derived from the ARGO position, span from November 2001 to October 2004, and are averaged in 3° (lon) ×2° (lat) bins. Obviously, different boundaries, namely the different modes under which the boxes are distributed, would result in some distinctions about the average currents (see Fig. 5). It is in the nature of things in the acquisition of the detailed mid-depth currents from the finite observation velocities by ARGO positions for this to happen. Besides that, the number of velocity samples ought to be enough in valid bins for the averaging error so that for velocities at 1000 dB, the first grid (the most western and southern grid) is at 30°S, 123°E and the smalest number is 7 in each bin, but at 2000 dB, the first grid is at 30°S, 122°E and the minimum number in each bin is 11.

Only boxes containing seven or more velocity vectors are shown at 1000 dB, and the maximum number of records in these boxes was 15 in Fig. 6. The base of the arrows denotes the center of each box. Obviously by this means, we are virtually able to not only avoid all the debate over currents in the regions containing observations, but also to reduce the sampling error.

After rigid error control, the resulting Eulerian velocities from ARGO positions at 1000 dB in the tropical Pacific are all concentrated between 5°S and 30°N. First, it is notable in Fig. 6b that the currents from NCEP are mostly underestimated except south of 5°N

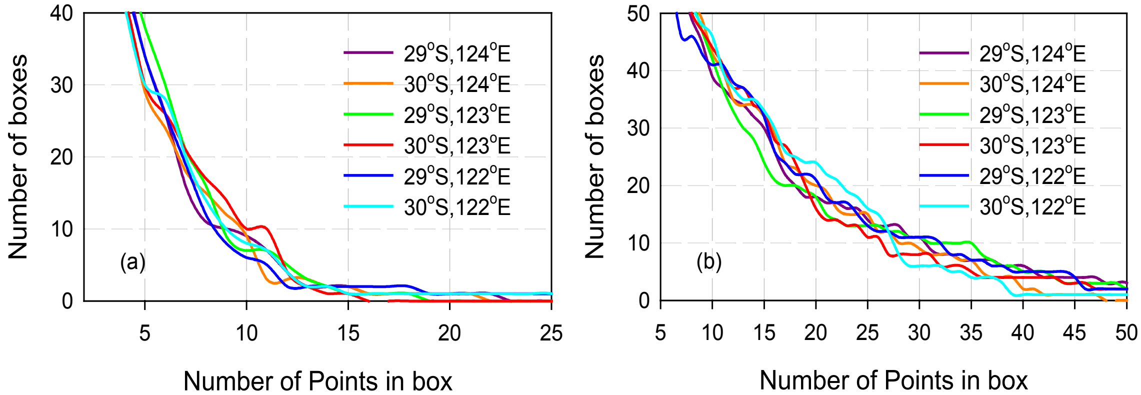

Fig. 5. The number of valid boxes versus the number of velocity points in each box at different grid boxes where the most western (southern) center of the bins is respectively 122°, 123°, and 124°E (for 30° and 29°S). (a) for 1000 dB; (b) for 2000 dB.

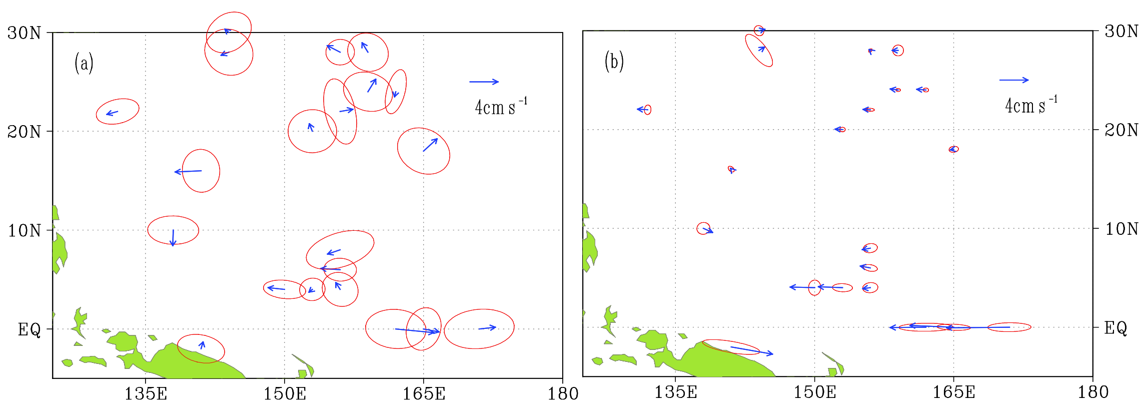

Fig. 6. Mean flow field based on 3°(lon)×2°(lat) bins together with standard deviation ellipses at 1000 dB in which the blue vectors are average current and the red ellipses represent their variance. The u (v) component variance is the long (short) axis. Data are only shown where more than six data points occur in each box. Note the scales are 4 cm s^{-1}. (a): observations by ARGO; (b): reanalysis by NCEP.

where the flow is westward and appears considerably zonally increased with values over 10 cm s^{-1}, but in Fig. 6a the Equatorial Intermediate Current (EIC) from 160°E to 175°E is eastward with values almost as high as 8 cm s^{-1}. Secondly, a distinctly different current shown up in the east of 150°E between 20°N and 30°N, where the observed current is evidently cyclonic whereas the reanalysis current is unanimously westward with values less than 2 cm s^{-1}. Moreover, between 135°E and 150°E, the two zonal currents around the north of 10°N are in opposite directions. However, between 5°N and 10°N, the zonal currents from the two datasets are remarkably consistent to the east of 150°E. Of course, the ARGO current shows greater variance than the reanalysis as a whole.

Thus, the discrepancies of mean flow between the NCEP reanalysis and the ARGO observations at 1000 dB have the following two properties: (1) the very strong westward EIC from the NCEP reanalysis is reversed relative to the observations by ARGO, and (2) the feeble current of the reanalysis cannot depict the cyclonic current near 150°E between 20°N and 30°N, which is considerably zonal.

As for the velocities at 2000 dB, those by ARGO are mostly spread outside 10°N (see Fig. 2) so that the mean flow between 10°N and 30°N at this depth can be discussed. Furthermore, the quantity of velocity vectors outclasses that at 1000 dB so that the average current of the Eulerian velocity at 2000 dB has 11 values in each 3° (lon) ×2° (lat) bin. Only boxes containing eleven or more velocity vectors at 1000 dB are shown in Fig. 5, where the maximum number of records in these boxes is 56. There are nearly 21 boxes in which the number of velocity vectors is over 20.

In Fig. 7, it is evident that the Eulerian velocities from the reanalysis at 2000 dB are underestimated. The strongest flow from the NCEP reanalysis is only about 1 cm s^{-1} in the figure, but the flow from the

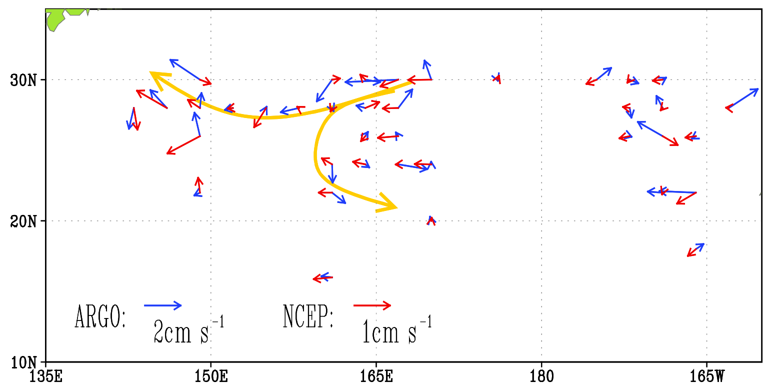

Fig. 7. Mean flow field based on 3°(lon)×2°(lat) bines at 2000 dB where the blue vectors are observations by ARGO, the red vectors are the reanalysis by NCEP, and the yellow circulations are deduced from observing the average currents. The flow can be shown only where there are more than six data points in each box. Note that the scales for the two data sources are different.

ARGO observations is near 3 cm s^{-1}. From the mean flow by the observations in the figure, we find a westward flow along 30°N from the west of 180°E, and south of 30°N, there is a bifurcation into two flows. The westward branch shows and anticyclonic tendency, and to the west of 150°E, it turns to the northwest. The other branch along 160°E first flows nearly southward until it approaches 20°N where it swerves to eastward, which suggests a cyclonic eddy. At this depth, there is a resemblance to ocean circulations in the oceanology chart by Reid (1997). However, around 165°E between 20°N and 30°N, the mean flows in the reanalysis are notably westward, which differs from the observational currents. Only the west branch can be exactly depicted by the reanalysis.

5. Summary

The activity of the global ARGO project not only creates the possibility to observe the temperature and salinity profiles in the mid- and upper-ocean in real time, but also provides a unique opportunity by which we can observe the absolute velocity at mid-depths within the world's oceans in near real-time.

In this paper, based on the principles of the operation and measurement characteristics of the ARGO floats, we perform some rigid quality control for the sake of including some gross errors in the derived mid-depth velocities. In virtue of our correction scheme for the drift error on the surface, we obtain 1597 velocity vectors in the tropical Pacific at 1000 dB and 2000 dB from the position information during the period from November 2001 to October 2004, which contain relative errors less than 25%.

By the 1597 velocity vectors, we evaluate the intermediate currents of the NCEP reanalysis. A quantitative comparison of descent points shows that the mid-depth velocities by the reanalysis are underestimated except near the equator. On average, the speed differences between the NCEP and the ARGO values reach −2.3 (−1.8) cm s^{-1} at 1000 (2000) dB, may be partly due to the fact that the velocity at 2000 dB is feeble. The direction deviations of the velocities appear random with no evident trends. However, the comparison of the distribution of the u and v component differences also reveals that the velocity differences between the ocean model and the observations at 2000 dB seem smaller than at 1000 dB. On the other hand, the mean flow analysis suggests that there are considerable gaps between the reanalysis and the observations in the mean current at mid-depths. Due to the fact that the zonal flow in the reanalysis is too strong, some eddies cannot be simulated, such as the cyclonic eddy to the east of 160°E in Fig. 7.

In addition, through the velocity comparison at the two depths, it is noticeable that many floats parking at 1000 dB cannot be used to acquire credible mid-depth velocities. Because the time information at the end of their ascent and start of descent at the surface takes default values in the trajectory files, and since the time information about their parking, descent and ascent cannot be realized from their metadata files, their application to measuring mid-depth currents is gravely restricted.

Acknowledgments. We would like to thank Drs. Joseph L. Reid and Russ. E. Davis for providing many help, and to Dr. Zhou Guangqing for helpful advices. We wish to thank reviewers for their numerous comments and suggestions that helped to improve the original manuscript significantly. This research is supported by Natural Science Foundation of China (Contract No. 40437017 and 40225015).

REFERENCES

Acero-Schertzer, C. E., D. V. Hansen, and M. S. Swenson, 1997: Evaluation and diagnosis of surface currents in the National Centers for Environmental Prediction's ocean analyses. *J. Geophys. Res.*, **102**, 21 037–21 048.

Behringer, D. W, M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Mon. Wea. Rev.*, **126**, 1013–1021.

Chapman, P., S. F. Di Marco, R. E. Davis, and A. C. Coward, 2003: Flow at intermediate depths around Madagascar based on ALACE float trajectories. *Deep-Sea Research II*, **50**, 1957–1986.

Davis, R. E., 1998: Preliminary results from directly measuring mid-depth circulation in the tropical and South Pacific. *J. Geophys. Res.*, **103**, 24 619–24 639.

Davis, R. E., and W. Zenk, 2001: Ocean circulation and climate: Subsurface Lagrangian observation during the 1990s. *Ocean Circulation and Climate*, Siedler et al., Eds., Academic Press, San Diego, 123–139.

Davis, R. E., D. C. Webb, L. A. Regier, and J. Dufour, 1992: The Autonomous Lagrangian Circulation Explorer (ALACE). *J. Atmos. Oceanic Technol.*, **9**, 264–285.

Enfield, D. B., and J. E. Harris, 1995: A comparative study of tropical Pacific sea surface height variability: Tide gauges versus the National Meteorological Center data-assimilation ocean general circulation model, 1982–1992. *J. Geophys. Res.*, **100**, 2515–2522.

Ichikawa, Y., Y. Takatsuki, K. Mizuno, N. Shikama, and K. Takeuchi, 2001: Estimation of drifting velocity and error at parking depths for the ARGO float. *ARGO Tech. Rep. FY2001*, Japan Agency for Marine-Earth Science and Technology, 68–77.

Ji, M., A. Leetmaa, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev.*, **123**, 460–481.

Ji, M., and T. M. Smith, 1995: Ocean model response to temperature data assimilation and varying surface wind stress: Intercomparisons and implications for climate forecast. *Mon. Wea. Rev.*, **123**, 1811–1821.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.

Lagerloef, G. S. E., G. Mitchum, R. Lukas, and P. Niiler, 1999: Tropical Pacific near-surface currents estimated from altimeter, wind, and drifter data. *J. Geophys. Res.*, **104**, 23313–23326.

Reid, J. L., 1997: On the geostrophic circulation of the Pacific Ocean: Flow patterns, tracers, and transports. *Progress in Oceanography*, **39**, 263–352.

Toner, M., A. C. Poje, A. D. Kirwan Jr., C. K. R. T. Jones, B. L. Lipphardt, and C. E. Grosch, 2001: Reconstructing basin-scale Eulerian velocity fields from simulated drifter data. *J. Phys. Oceanogr.*, **31**, 1361–1376.

Xie Jiping, 2005: A study on the acquirability of mid-depth current information by the ARGO float. M. S. thesis, Department of Meteorology, Nanjing University of Information Science & Technology, 110pp. (in Chinese)