An Improved Near-Surface Specific Humidity and Air Temperature Climatology for the SSM/I Satellite Period

XIANGZE JIN AND LISAN YU

Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

DARREN L. JACKSON

Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/Earth System Research Laboratory, Boulder, Colorado

GARY A. WICK

NOAA/Earth System Research Laboratory, Boulder, Colorado

(Manuscript received 12 April 2014, in final form 10 November 2014)

ABSTRACT

A near-surface specific humidity (Qa) and air temperature (Ta) climatology on daily and 0.25° grids was constructed by the objectively analyzed air–sea fluxes (OAFlux) project by objectively merging two recent satellite-derived high-resolution analyses, the OAFlux existing 1° analysis, and atmospheric reanalyses. The two satellite products include the multi-instrument microwave regression (MIMR) Qa and Ta analysis and the Goddard Satellite-Based Surface Turbulent Fluxes, version 3 (GSSTF3), Qa analysis. This study assesses the degree of improvement made by OAFlux using buoy time series measurements at 137 locations and a global empirical orthogonal function (EOF) analysis. There are a total of 130 855 collocated daily values for Qa and 283 012 collocated daily values for Ta in the buoy evaluation. It is found that OAFlux Qa has a mean difference close to 0 and a root-mean-square (RMS) difference of 0.73 g kg^{-1}, and Ta has a mean difference of −0.03°C and an RMS difference of 0.45°C. OAFlux shows no major systematic bias with respect to buoy measurements over all buoy locations except for the vicinity of the Gulf Stream boundary current, where the RMS difference exceeds 1.8°C in Ta and 1.2 g kg^{-1} in Qa. The buoy evaluation indicates that OAFlux represents an improvement over MIMR and GSSTF3. The global EOF-based intercomparison analysis indicates that OAFlux has a similar spatial–temporal variability pattern with that of three atmospheric reanalyses including MERRA, NCEP-1, and ERA-Interim, but that it differs from GSSTF3 and the Climate Forecast System Reanalysis (CFSR).

1. Introduction

Latent and sensible heat exchanges at the air–sea interface play a key role in the coupled air–sea interactions on various temporal and spatial scales. These fluxes are commonly estimated using the bulk formulas, which link turbulent fluxes to macroscale near-surface meteorological observables, such as air temperature (Ta), specific humidity (Qa), sea surface temperature (SST), and wind speed (Liu et al. 1979; Fairall et al. 2003). These flux-related variables are obtainable from three major sources: marine surface weather reports from the Voluntary Observing Ship (VOS) program, satellite observations, and atmospheric reanalysis and operational models. VOS observations have good accuracy and long time series but poor global coverage, as observations are concentrated along ship routes (Josey 2001; Gulev et al. 2007). In comparison with ship observations, satellite observations have the capability of providing global coverage at higher spatial and temporal resolutions, albeit with limited sampling in space and time depending on the satellite's orbit and sensor and with a shorter time record.

Several satellite-based heat flux products have been developed over the past decades, such as the Goddard Satellite-Based Surface Turbulent Fluxes (GSSTF)

Corresponding author address: Dr. Xiangze Jin, Mail Stop 21, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, MA 02543-1050.
E-mail: xjin@whoi.edu

DOI: 10.1175/JTECH-D-14-00080.1

(Chou et al. 2003; Shie et al. 2012), the Japanese Ocean Flux Data Sets with Use of Remote Sensing Observations (J-OFURO) (Kubota et al. 2002), the Institut Français de Recherche et l'Exploitation pour la Mer (IFREMER) fluxes (Bentamy et al. 2003), the Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data (HOAPS) (Andersson et al. 2011), and the SeaFlux turbulent flux dataset (Curry et al. 2004). However, these products differ considerably from each other (Gao et al. 2013). A large portion of the errors in these products is found to be associated with uncertainties in the near-surface Qa and Ta, as these near-surface atmosphere properties cannot be directly sensed from satellites, and retrieval algorithms are very different and all have uncertainties (Curry et al. 2004; Jackson et al. 2006; Jackson and Wick 2010; Roberts et al. 2010). The impact of biases of Qa and Ta on heat fluxes depends primarily on wind speed. For example, a wet bias of $1\,g\,kg^{-1}$ Qa would underestimate the latent heat flux by about $38\,W\,m^{-2}$, whereas a warm bias of 1°C Ta would reduce the sensible heat flux by about $15\,W\,m^{-2}$ at $10\,m\,s^{-1}$ wind speed (Fig. 1).

The differences in satellite-based Qa from these products are primarily due to retrieval algorithms. Schulz et al. (1993) developed a model to estimate the bottom-layer precipitable water from the temperature brightness (TB) measured by the Special Sensor Microwave Imager (SSM/I), and then related TB linearly to Qa. Based on Schulz's model, Schlüssel et al. (1995) did direct regression between TB and Qa to avoid error propagation. Bentamy et al. (2003) further updated Schulz et al.'s (1993, 1997) regression coefficients with improved training data. In a study by Jackson et al. (2006), Qa and Ta were derived by combining observations from SSM/I and the Advanced Microwave Sounding Unit-A (AMSU-A). This multisensory satellite approach helps to improve the accuracy of the retrievals in comparison with that from the single-sensor approach (Jackson et al. 2006).

In contrast to Qa, there is currently no standard approach for estimating Ta using SSM/I (Roberts et al. 2010). For instance, HOAPS estimates Ta from SST using the assumption of 80% humidity (Liu 1988) and 1°C air–sea temperature difference. GSSTF employs the NCEP–DOE AMIP-II Reanalysis (NCEP–DOE reanalysis) instead of deriving it from satellite retrievals (Chou et al. 2003; Shie et al. 2012). SeaFlux utilizes a nonlinear neutral network that was trained with in situ observations to match up with SSM/I (Roberts et al. 2010). Jackson et al. (2006) derived Ta using the same multi-instrument approach as for Qa, and Jackson and Wick (2010) implemented another approach that can retrieve Ta from satellite SST and AMSU-A.

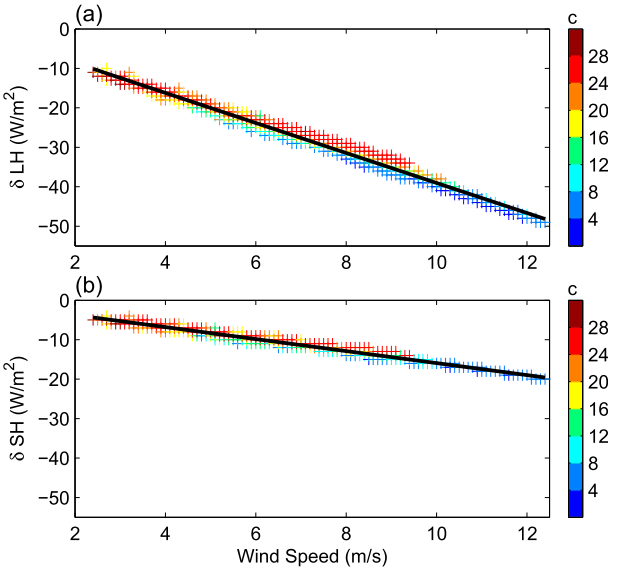

FIG. 1. (a) Scatterplots for wind speed vs change in latent heat (LH) flux due to a wet bias in Qa of $1.0\,g\,kg^{-1}$. The flux was computed using the COARE algorithm. The flux-related surface meteorological variables, including Ta, Qa, wind speed, SST, and pressure, are from the OAFlux 1° climatology (1988–2010) over the global ice-free oceans. The black solid line represents linear regression of the heat flux anomaly on wind speed, with a slope of −3.8. The colors represent the values of the corresponding SSTs. (b) As in (a), but for wind speed vs change in sensible heat (SH) flux due to a warm bias in Ta of 1.0°C, with a slope of −1.5.

The OAFlux is a synthesis analysis that integrates satellite retrievals and atmospheric reanalyses to find an optimal combination in a least squares sense (Yu and Weller 2007). The synthesis can reduce errors in input data sources and produce an estimate that has the minimum error variance. The OAFlux project has been providing two to three online updates per year for the 1° gridded global turbulent heat fluxes, as well as the flux-related variables including Qa and Ta, encompassing the past five decades from 1958 onward. In recent years, efforts have been devoted to constructing a higher-resolution (0.25°) global analysis by taking advantage of several recent achievements in satellite-based products that were made by our own group and other groups, including a 0.25° 12-sensor merged vector wind analysis (1987 onward) (Yu and Jin 2012), Qa and Ta from the multi-instrument microwave regression (MIMR) products (1999–2010) (Jackson et al. 2009), Qa from GSSTF3 (1987–2008) (Shie et al. 2012), and SST from the NOAA Advanced Microwave Scanning Radiometer–Advanced Very High Resolution Radiometer (AMSR–AVHRR) and AVHRR-only optimum interpolation analysis (Reynolds et al. 2007). A preliminary analysis of the high-resolution (HR) OAFlux turbulent latent and sensible heat fluxes was conducted in the eddy-rich Gulf

TABLE 1. List of products used in this study, including horizontal resolution and algorithms.

	Spatial resolution	Ta and Qa algorithms
CFSR	T382 (0.313°)	Ta and Qa are not explicitly assimilated
NCEP-1	T63 (1.875°)	Ta and Qa are not explicitly assimilated
MERRA	0.5° (lat), 0.667° (lon)	3D-Var analyses
ERA-Interim	T255 (0.703°)	Optimal interpolation of data from ships and buoys
MIMR	0.25°	Utilized AMSU-A and SSM/I microwave to determine Ta and Qa from a linear regression
GSSTF3	0.25°	Regresses Qa directly onto SSM/I TB with improved training dataset, plus EOF method
OAFlux-0.25°	0.25°	Objective synthesis

Stream region (Jin and Yu 2013), which found that OAFlux HR analysis clearly outperforms the atmospheric reanalyses in both latent and sensible heat fluxes in comparison with buoy measurements.

The need for high-resolution Qa and Ta with improved accuracy is vital for improvement of satellite-based heat fluxes (Curry et al. 2004). Jin and Yu (2013) indicated that in comparison with the 1° OAFlux analysis, the HR OAFlux has smaller root-mean-square (RMS) differences (RMSD) of ~0.15 g kg^{-1} in Qa and ~0.47°C in Ta against buoy measurements in the Gulf Stream region. Incorporating the MIMR satellite retrievals into the OAFlux synthesis is the key to improvement in both Qa and Ta. In the present study, we further evaluate the HR OAFlux Qa and Ta (1988–2010) over the global ice-free oceans. An intercomparison is also carried out between the HR OAFlux, MIMR, GSSTF3, and four atmospheric reanalyses (Table 1). Two approaches were used for evaluation. One is to assess the degree of improvement made to OAFlux using buoy time series measurements at 137 locations. The other approach is to use an empirical orthogonal function (EOF)-based intercomparison analysis of the seven products to ascertain the consistency of spatial–temporal variability on a basin scale.

This paper is organized as follows. Section 2 provides a general description of data products, including MIMR, GSSTF3, four atmospheric reanalyses, and the buoy measurements. Detailed descriptions of the OAFlux synthesis and a comparison of the climatology between OAFlux and the two satellite-based products are given in section 3. Section 4 shows the results of buoy evaluation. Section 5 presents the global EOF-based intercomparison analysis. The summary and conclusions are included in section 6.

2. Data description

a. *Qa from GSSTF3*

There have been several updates to the GSSTF products. We used the newly developed GSSTF3 (Shie et al. 2012) in the OAFlux synthesis. As in previous versions, the GSSTF3 Qa is statistically retrieved from the SSM/I TB, while Ta is taken from the NCEP–DOE reanalysis (Kanamitsu et al. 2002). In contrast to previous versions that derived Qa based on the bottom-layer precipitable water and the total precipitable water using an EOF method (Chou et al. 1995), GSSTF3 (Shie et al. 2012) adopted the one-step approach that regresses Qa directly onto TB (Schlüssel et al. 1995; Bentamy et al. 2003). The EOF algorithm is still retained in the updated algorithm. In both GSSTF3 and its preceding GSSTF2c, a corrected/improved set of SSM/I TB was used and that reduced a temporal trend post-year 2000 in the globally averaged latent heat flux, which was mainly due to the temporal variation–drifting (decreasing) of the earth incidence angle of the SSM/I satellites (Shie 2010a,b). The GSSTF3 is on a 0.25° grid, covering the period July 1987–December 2008. A validation against a total of 22 samples from in situ observations indicates the mean bias for Qa is 0.25 g kg^{-1} with the RMS difference of 1.11 g kg^{-1} (Shie et al. 2012).

b. *Qa and Ta from MIMR*

The MIMR utilized AMSU-A and SSM/I microwave to determine Ta and Qa from a linear regression (Jackson et al. 2006). The inclusion of the AMSU-A 52.8-Ghz channel, which has a peak weighting in the lower troposphere, plays a key role in improving the Ta and Qa retrievals. The retrievals were further improved by refinements to the regression formula, the training dataset, and the collocation procedure (Jackson et al. 2009). The training data originate from research vessels that are described in more detail in Jackson et al. (2009). No buoy data were used in the training data. While SSM/I data have been available since 1987, the MIMR data began in 1998, since AMSU-A first came available on *NOAA-15* at that time. Independent validation indicates an RMS difference of 1.59 g kg^{-1} for Qa (Jackson et al. 2009) and 1.55°C for Ta (Jackson and Wick 2010). The data of Ta and Qa used in this study have a spatial resolution of 0.25°, covering the global oceans from 70°S to 70°N for the period 1999–2010.

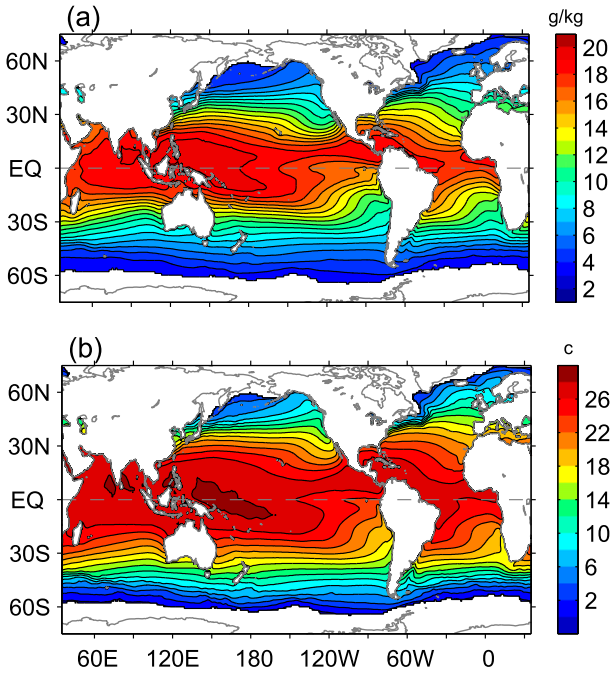

FIG. 2. Annual mean OAFlux (a) Qa and (b) Ta from 1988 to 2010.

c. *Qa and Ta from the reanalyses*

We made use of four reanalysis products in the study, including the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim) (Dee et al. 2011), NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al. 2011), the National Centers for Environmental Prediction (NCEP), the Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010), and the first-generation reanalysis from the NCEP–National Center for Atmospheric Research (NCAR) reanalysis (NCEP-1) (Kalnay et al. 1996). The horizontal resolution in the latest reanalysis products ranges between 0.313° (T382) and 0.703° (T255), which is clearly an improvement over the early reanalysis from NCEP at 1.875° (T63) resolution (Table 1). MERRA used a three-dimensional variational data assimilation (3D-Var) analysis algorithm and made extensive use of satellite radiance information and ground observations, including temperature and humidity from ships and buoys. Unlike MERRA, the near-surface (2 m) Qa and Ta were not explicitly assimilated in both NCEP and CFSR. In contrast, ERA-Interim postprocessed the ship and buoy observations into their outputs using an optimal interpolation scheme. Note that MERRA, ERA-Interim, and NCEP-1 data were used in the OAFlux synthesis, while CFSR data are not synthesized.

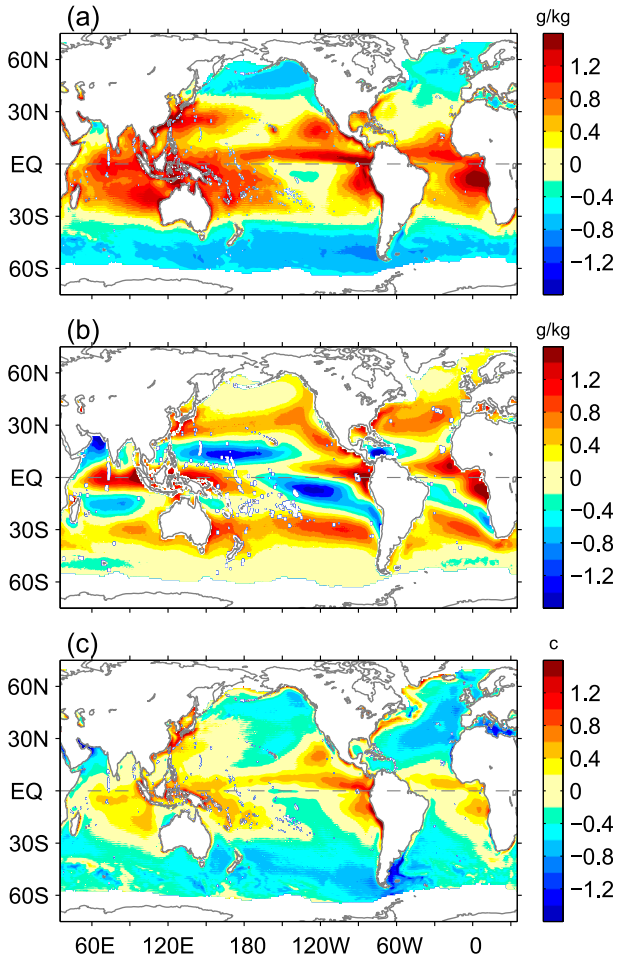

FIG. 3. (a) Mean difference of MIMR minus OAFlux in Qa from 1999 to 2010. (b) Mean difference of GSSTF3 minus OAFlux in Qa for the period 1988–2008. (c) As in (a), but for Ta.

d. *In situ buoy measurements*

The validation datasets used in this study include the buoy measurements acquired from the Tropical Atmosphere Ocean/Triangle Trans-Ocean Buoy Network (TAO/TRITON) buoy array in the tropical Pacific (McPhaden et al. 1998); the Research Moored Array for Africa–Asian–Australian Monsoon Analysis and Prediction (RAMA) in the tropical Indian Ocean (McPhaden et al. 2009); the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA) (Bourlès et al. 2008); the moored buoys at the Kuroshio Extension Observatory (KEO) (Cronin et al. 2010), the climate station Papa (Kamphaus et al. 2008), and the National Data Buoy Center (NDBC) buoys (http://www.ndbc.noaa.gov/) in the northern North Pacific; two buoys in the Southern Ocean, including the Agulhas Return Current (ARC) buoy that was located southeast of the tip of Africa (http://www.pmel.noaa.gov/OCS/ARC/)

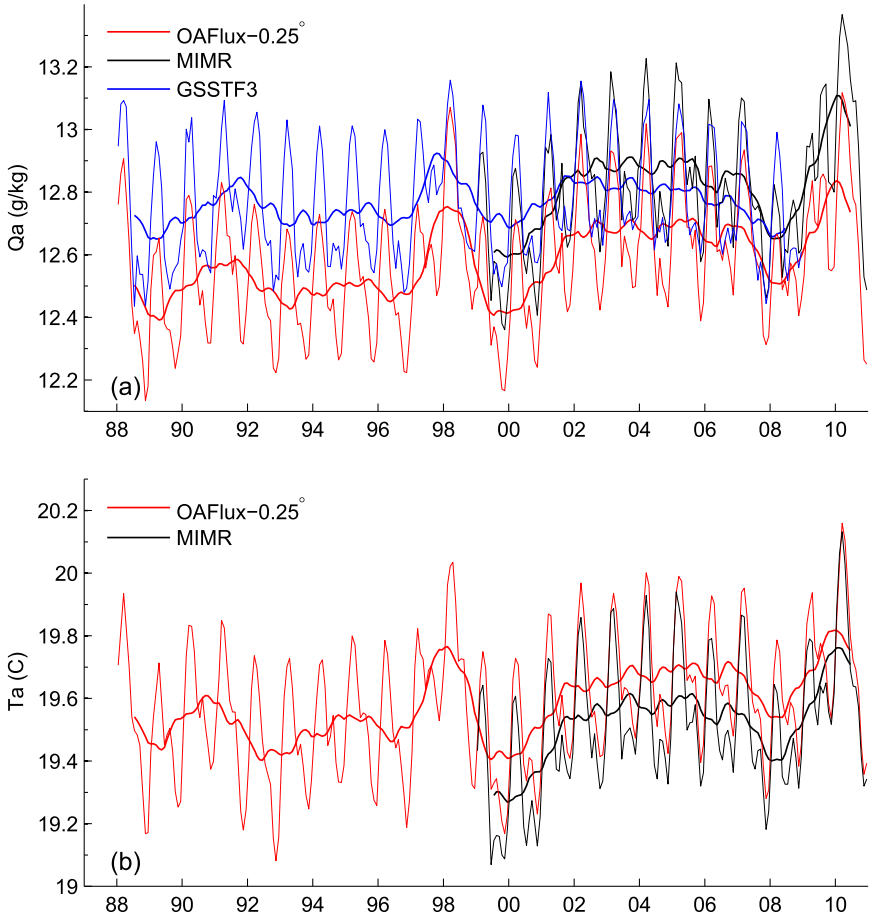

FIG. 4. Time series of monthly-mean (a) Qa and (b) Ta from OAFlux (red), MIMR (black), and GSSTF3 (blue, Qa only), averaged over the global ice-free ocean (from 65°S to 65°N). The thick lines represent the time series with a 13-month running mean.

and the Southern Ocean Flux Station (SOFS) buoy that was deployed south of Australia (http://uop.whoi.edu/projects/SOFS/); and 19 archived/active moored buoys deployed by WHOI at flux reference sites and the sites selected for targeted field programs. The active WHOI buoys include a site at 20°S, 85°W under the stratus cloud deck off northern Chile (Stratus); the Northwest Tropical Atlantic Station (NTAS) at 15°N, 51°W; and a site north of Hawaii near the WHOI Hawaii Ocean Time Series (WHOTS) site. The archived WHOI buoys include the Arabian Sea Experiment (Arabian Sea) at 15.5°N, 61.5°E; the Acoustic Surface Reverberation Experiment (ASREX) at 49.2°N, 131.9°W and 33.9°N, 69.7°W; the Coastal Mixing and Optics Experiment (CMO) at 40.5°N, 70.5°W; the COARE at 1.8°S, 156.0°E; the Marine Light–Mixed Layer Experiment 1991 (MLML91) at 59.5°N, 20.8°W; the Pan American Climate Study (PACS) at 2.8°S, 124.7°W and 10.0°N, 125.4°W; the Severe Environment Surface Mooring (SESMOOR) at 42.5°N, 61.2°W; the Shelf Mixed Layer Experiment (SMILE) at 38.7°N, 123.5°W; and the Subduction Experiment (Subduction) in the subtropical Atlantic. All the WHOI buoy data are available online (at http://uop.whoi.edu). There are a total of 122 buoy time series available for the period 1999–2010 and 137 buoy time series available for the period 1988–2010. The density of buoys is greatest in the tropical regions.

The WHOI buoys are equipped with the Improved Meteorological Instruments (IMET) system or the Air–Sea Interaction Meteorology (ASIMET) system (Weller and Anderson 1996). The three tropical arrays carry the Autonomous Temperature Line Acquisition System (ATLAS) (McPhaden et al. 1998) or ASIMET. The estimated IMET daily mean errors are 1% (3% in low wind) and 0.1°C (more in low wind) for relative humidity and Ta, respectively (Colbo and Weller 2009). Note that the 1% accuracy in relative humidity corresponds to the accuracy of Qa of about 0.03 (at high latitudes) to $0.23\,g\,kg^{-1}$ (at warm pool). The expected errors for the ATLAS/TRITON instrument are about 2% and 0.1°C,

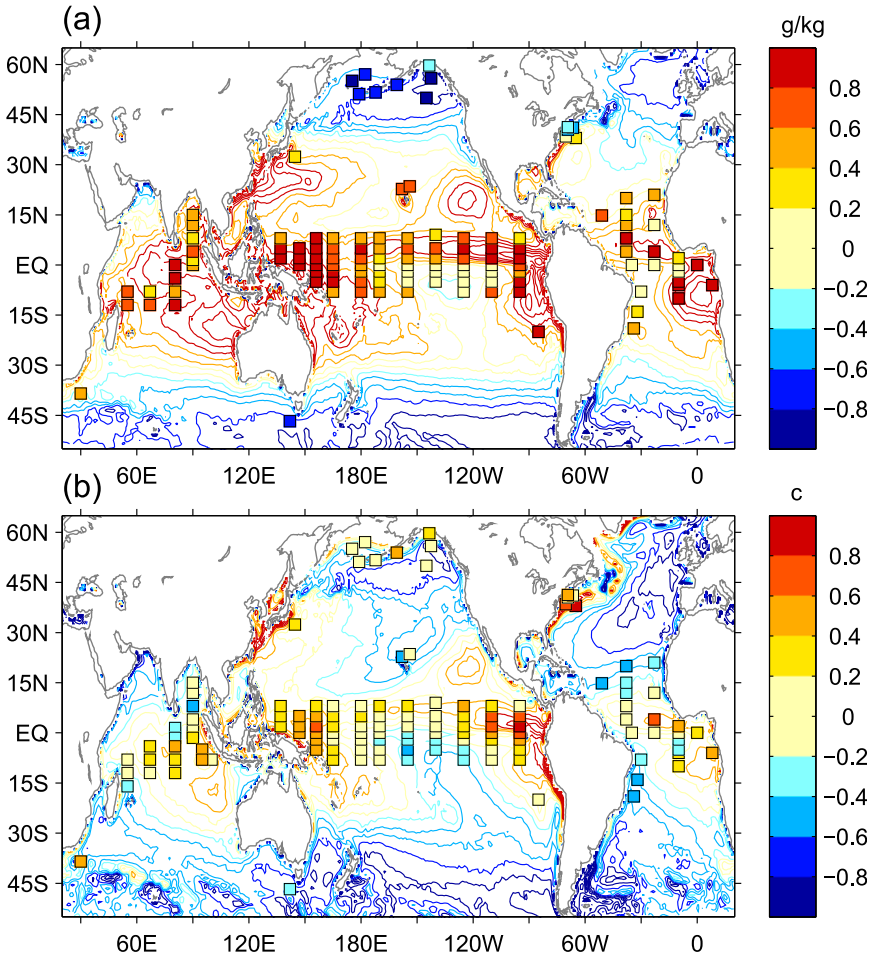

FIG. 5. Mean difference in (a) Qa and (b) Ta of MIMR minus buoy (squares) over 122 buoy locations from 1999 to 2010, of which 102 buoys are from the combined TAO/TRITON, RAMA and PIRATA arrays over the tropical oceans. Contours are the mean difference of MIMR minus 1° OAFlux over the same period. Warm colors indicate positive bias (i.e., the satellite retrieval is overestimated), and cold colors indicate negative bias (i.e., the satellite retrieval is underestimated).

whereas the error for NDBC is about 3% and 1.0°C, for relative humidity and Ta, respectively.

Note that the buoy measures relative humidity instead of Qa. The latter was calculated in terms of Ta, surface pressure, and relative humidity. Buoy Ta and relative humidity sensors are usually deployed at 2–4-m height and measurements are made at a sample rate of 1–10 min depending upon the design of instruments. For consistency in comparison, buoy measurements were adjusted to 2-m height using the COARE algorithm (Fairall et al. 2003) and then were averaged to daily values.

3. The Qa and Ta from the OAFlux synthesis

The methodology of the OAFlux synthesis is based on the Gauss–Markov statistical estimation theorem. That is, when combining data in a linear fashion, the linear least squares estimator is the most efficient estimator (Daley 1991). In the case of the OAFlux flux analysis, the theorem led to the formulation of a least squares problem based on available satellite retrievals and reanalysis model outputs (Yu and Weller 2007). The construction of the 0.25°-gridded Qa and Ta used the same methodology, that is, merging of MIMR (1999–2010), GSSTF3 (July 1987–December 2000), the OAFlux existing 1° analysis, and three atmospheric reanalyses, including ERA-Interim, MERRA, and NCEP-1. The GSSTF3 (2001–08) was not utilized in the synthesis. Note that the optimality of the solution is dependent on the weights that theoretically are inversely proportional to the respective error of the input datasets. Since the lack of error information for the input datasets limits our ability

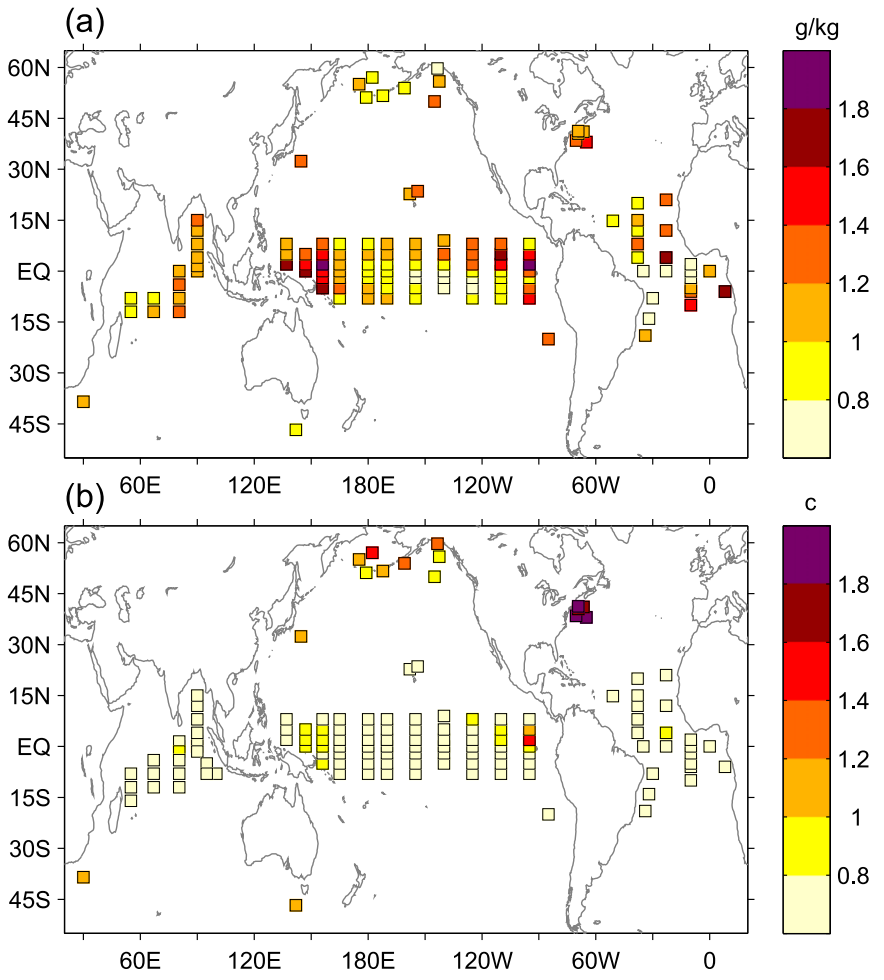

FIG. 6. RMS difference in (a) Qa and (b) Ta of MIMR minus buoys over 122 buoy locations from 1999 to 2010.

to assign "true" weights, the weights were in fact determined from the buoy-based evaluation on each input dataset. All the weights were constant, due to a lack of sufficient in situ measurements to define the latitude dependence of errors. The buoy-based evaluation was established from 137 buoy time series, of which 115 time series were from the tropical (30°S–30°N) moored array system. This indicates that the evaluation may be sufficient to characterize the error statistics of warm and wet conditions, but it has a limitation to provide relevant reference for cold and dry conditions. We point out that although buoy measurements are not directly used in the OAFlux product, they are used to determine the weights, and therefore we could expect OAFlux to get an overall better agreement in Qa and Ta with the buoy measurements.

Since all the input satellite retrievals are at 10 m above the ocean surface, these retrievals were adjusted to a height of 2 m for the synthesis. The 0.25° OAFlux wind speed (Yu and Jin 2012) and Reynolds OISST (Reynolds et al. 2007), as well as the COARE algorithm, were used for the height adjustment. A quality control was applied to reject some isolated unrealistic pairs of MIMR Qa and Ta in the tropical oceans. Those pairs of isolated extreme values could be removed once the Ta retrieval departs from the 1° OAFlux analysis by more than 6°C and results in larger air–sea temperature difference. About 0.001% pairs of Qa and Ta were rejected for being outside the limits.

The mean daily coverage of the satellite-based data for global ice-free oceans was about 55% in 1988, when there was only one SSM/I sensor, and it gradually increased to about 78% by 1996. The coverage is relatively lower in the tropical oceans than that at midlatitudes because the orbital geometry tends to create more gaps at low latitudes. Clearly there is a need to fill in gaps of missing data in order to complete the daily global field. As complete coverage is achieved every 2–3 days,

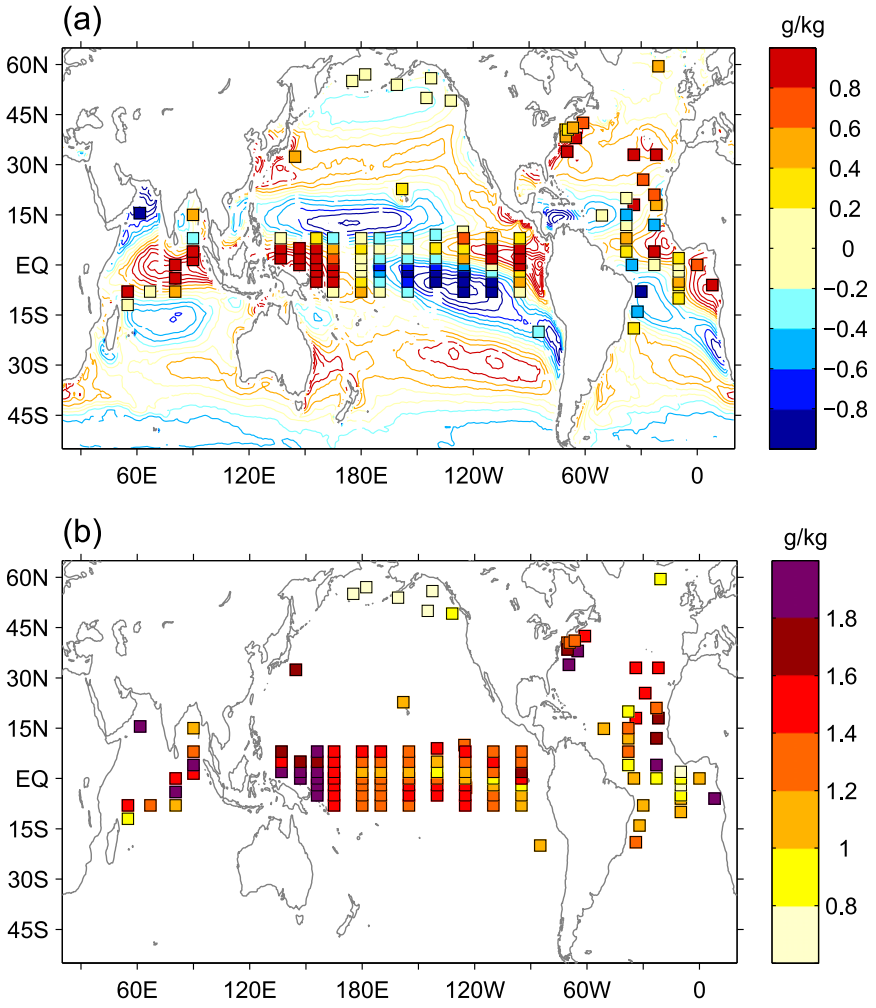

FIG. 7. (a) Mean difference in Qa of GSSTF3 minus buoys (squares) over 125 buoy locations from 1988 to 2008. Contours are the mean difference of GSSTF3 minus 1° OAFlux over the same time period. (b) RMS difference of GSSTF3 Qa minus buoys for the same period.

a temporal interpolation of the satellite retrievals was utilized to fill the gap. The interpolation was performed using the 1° OAFlux historical data as the reference to determine the direction of interpolation.

Theoretically, the synthesis process tends to cancel out errors in input datasets if the data have no systematic errors. A buoy validation indicates that both MIMR and GSSTF3 have a major systematic bias against the buoy observations. These details will be discussed further in section 4. A climatological monthly-mean adjustment was applied to both MIMR and GSSTF3 before they were merged into the synthesis. The adjustment was based on the 1° OAFlux, which has demonstrated to have overall good agreement with buoy observations (Yu et al. 2008). To do the mean adjustments, the climatology monthly mean of satellite data was averaged to 1° grids to calculate the difference against that of the 1° OAFlux and then the difference was interpolated back into the 0.25° grids. The fine structures and the temporal variability of satellite retrievals are retained. In fact, the HR OAFlux product is able to depict sharp oceanic fronts (Jin and Yu 2013). We have now completed the 0.25° OAFlux daily analysis from July 1987 to December 2010.

The annual mean OAFlux Qa and Ta are shown in Fig. 2. These mean fields are constructed over the 23-yr (1988–2010) analysis period. The two variables show a similar global distribution: higher values are in the tropical regions with the maxima over the Pacific warm pool, and the values decrease poleward. The sharpest fronts are observed in the Gulf Stream region and the southern oceans.

Figure 3 displays the mean difference of the satellite retrievals minus OAFlux over the overlapping period

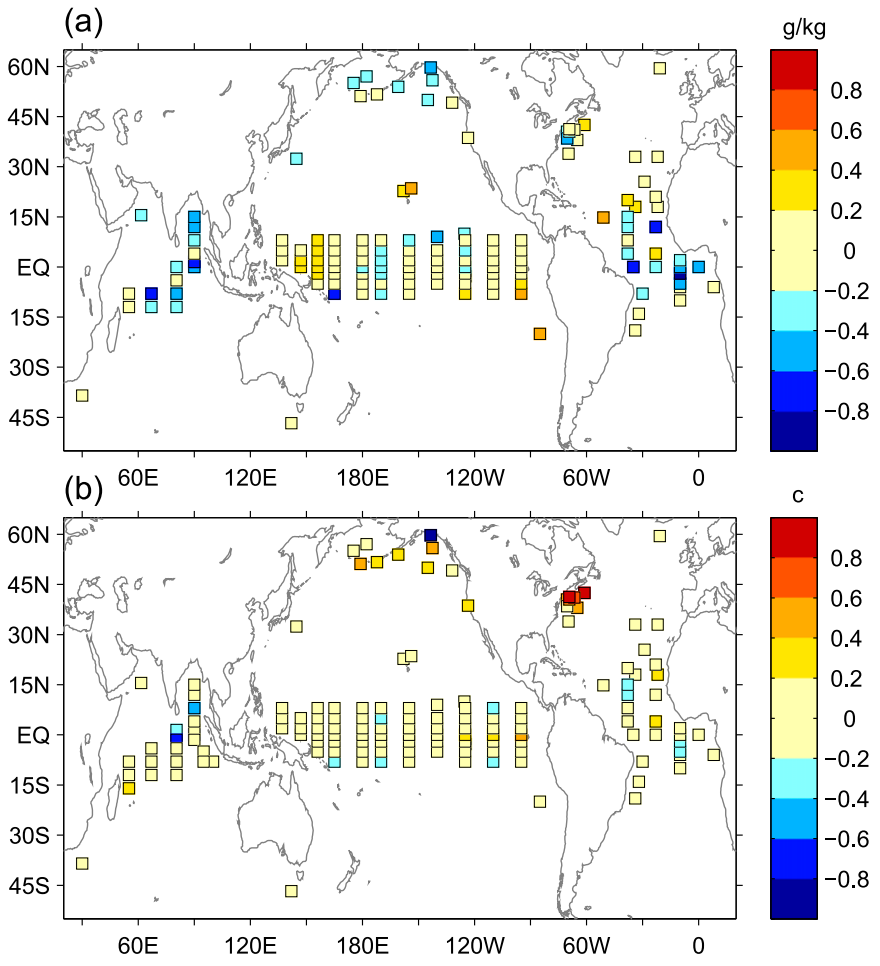

FIG. 8. Mean difference in (a) Qa and (b) Ta of OAFlux minus buoys at 137 buoy locations, from 1988 to 2010.

1999–2008. In comparison with OAFlux, the MIMR satellite Qa is generally overestimating at high Qa in the tropical oceans (30°S–30°N) and underestimating at low Qa in the mid–high latitudes (Fig. 3a). A large positive difference exists in the intertropical convergence zone (ITCZ), the east-equatorial Pacific, and the east-equatorial Atlantic, with the maximum at about 1.5 g kg^{-1}. The difference at high latitude is relatively small. The significant contrast between the tropics and the high latitudes seems to be related to the magnitude of the Qa. A comparison of Qa between GSSTF3 and OAFlux shows a mixture of striking positive and negative differences in low to mid-latitudes (Fig. 3b). In general, GSSTF3 is wetter than OAFlux in the midlatitudes and the equatorial regions and drier in the off-equator regions. Interestingly, this pattern is found to be similar to that of total cloud cover; that is, the wet Qa difference corresponds to the high total cloud cover, and the dry Qa difference corresponds to the low total cloud cover (not shown). The difference in Qa at high latitudes, however, is rather small. Figure 3c shows the difference in Ta between MIMR and OAFlux, of which the pattern is nearly identical to what is shown in Fig. 3a; that is, the large warm difference corresponds to the large wet difference and vice versa.

Figure 4a displays the time series of Qa from OAFlux, MIMR, and GSSTF3, averaged over global ice-free oceans (65°S–65°N) for the period 1988–2010. The thin line represents the monthly mean, while the thick line represents a 13-month running mean. The values of MIMR Qa are higher than those of OAFlux but have a similar variability during the period. The GSSTF3 Qa follows OAFlux from 1988 to 2002 with high values, but thereafter GSSTF3 decreases, while OAFlux remains steady before a dip in 2007. In fact, the GSSTF3 Qa (2001–08) was not utilized in the OAFlux synthesis for consistency of the time series.

The corresponding time series of Ta from OAFlux and MIMR is shown in Fig. 4b. The two time series are nearly in phase but vary in the mean values. Note that the OAFlux analysis of Ta (1988–98) is solely dependent

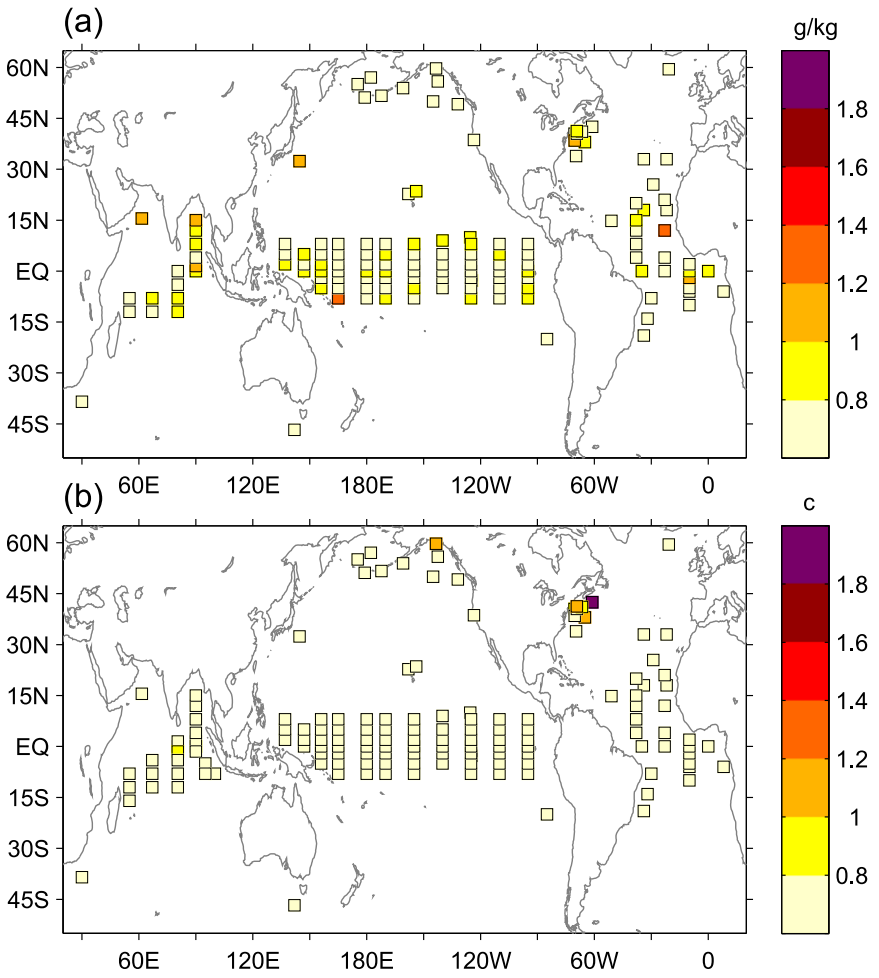

FIG. 9. As in Fig. 8, but for RMS difference.

on the reanalyses, as no satellite Ta retrieval was utilized in the synthesis.

4. Buoy evaluation

a. Satellite retrievals versus the buoy

In this study, validations were performed using collocated daily mean time series. As mentioned above, both the satellite retrievals and buoy measurements were adjusted to the 2-m height for validation. Because the data grids do not generally coincide with buoy positions, a bilinear interpolation between the four grid values surrounding the buoy location was used to obtain the satellite value at the location. If one of the grid values was missing, then the nearest-neighbor grid value among the four was selected. The satellite value would be marked as missing if none of the grid values was available.

The uncertainty associated with collocation could be caused by 1) height adjustments, 2) spatial interpolation, and 3) satellite sampling errors as the satellite retrieval does not sample the entire diurnal cycle like the buoy data. We compared the difference in Qa between a pair of 2895 collocated daily mean time series from MIMR and the Stratus buoy (20°S, 85°W) at different heights, and found the RMS difference between MIMR and the buoy was ~1.35 g kg^{-1} at 2 m compared to ~1.41 g kg^{-1} at 10 m. The difference caused by the height adjustment was relatively small. Using the same buoy time series, we subsampled the OAFlux Qa by including the MIMR missing gap to estimate the uncertainty caused by spatial interpolation, and found that the change in the RMS difference between OAFlux and the buoy is ~0.003 g kg^{-1}. We also used the 3-hourly MIMR data to match the buoy time series, and found the change in the RMS difference is ~0.03 g kg^{-1}. In conclusion, the uncertainty associated with collocation is relatively small. Note that these errors are spatially dependent and might be larger in a location with a larger diurnal cycle. It also needs to be recognized that buoy

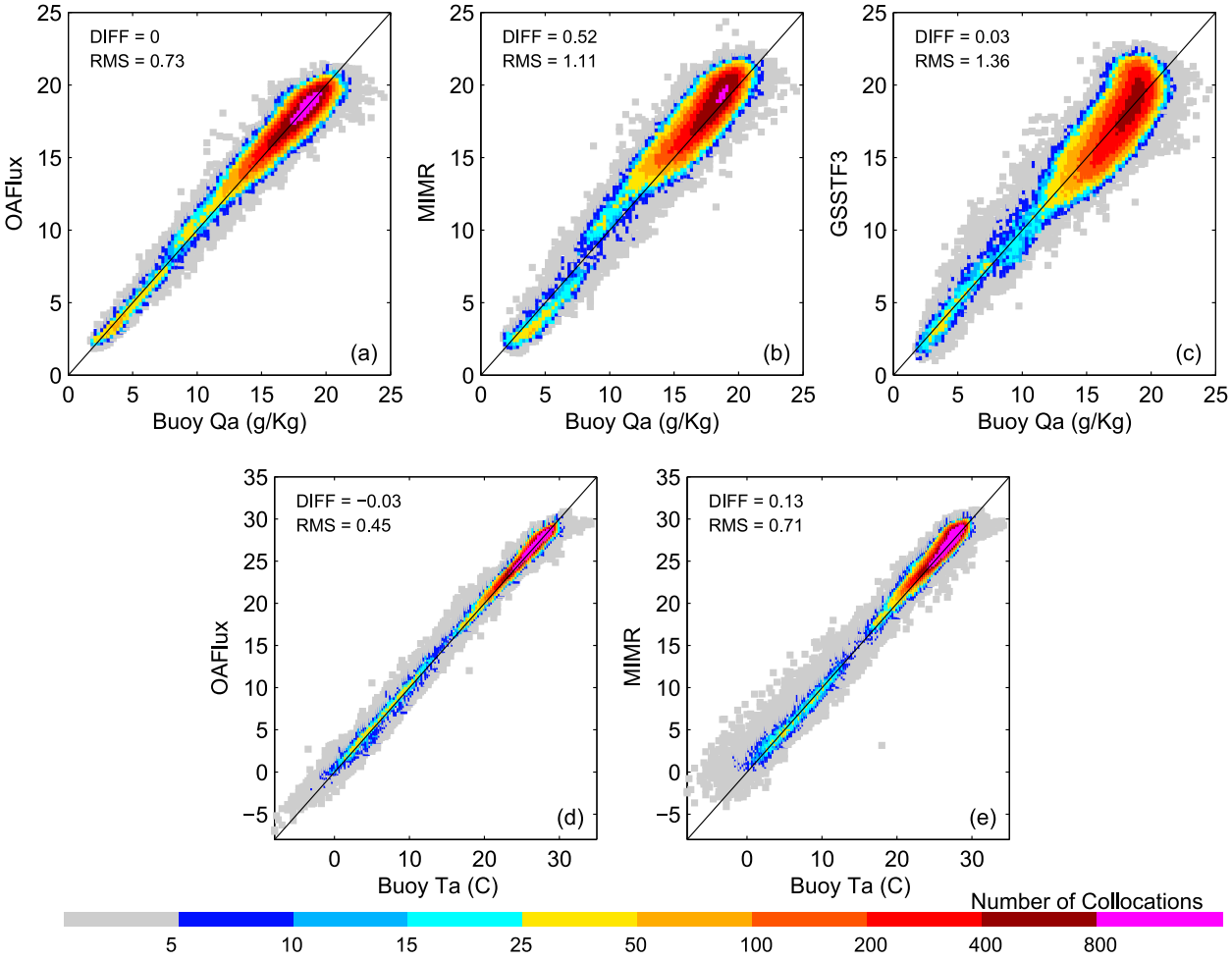

FIG. 10. (top) Scatterplots of Qa for (a) OAFlux, (b) MIMR, and (c) GSSTF3 vs the respective buoy and (bottom) scatterplots of Ta for (d) OAFlux and (e) MIMR. The plots of Qa are based on a total of 130 855 collocations among the three participating products and buoy measurements from 1999 to 2008. The plots of Ta are based on a total of 282 977 collocations between the two participating products and buoy measurements from 1999 to 2010.

measurements are single-point measurements; while the gridded product estimates at each grid point represent a gridcell average. So, any interpolation will miss the effect of scaling associated with the collocation of the gridcell average with the point measurement. Such an effect can be large in the regions where strong spatial gradients are presented.

The mean differences in Qa and Ta between MIMR and the buoy are shown in Fig. 5. The pattern of the mean difference in Qa shows that warm-colored square points predominate in the tropical oceans, while cold-colored points predominate in the high latitudes. Obviously in comparison with the buoys, the MIMR Qa is overestimated, in particular in the Pacific–Indian warm pool and the tropical eastern Pacific and eastern Atlantic. In the northern North Pacific, however, Qa is underestimated. The maximum difference is up to $0.8\,g\,kg^{-1}$. The wet bias could underestimate the latent heat flux up to 12–15 $W\,m^{-2}$ at 4–5 $m\,s^{-1}$ wind speeds, and the dry bias could overestimate the latent heat flux up to 24–27 $W\,m^{-2}$ at 8–9 $m\,s^{-1}$ wind speeds in the tropical oceans and in the northern North Pacific. Note that the colors of the square points (which represent differences with respect to the buoy) match the colors of the contours (which represent differences to the 1° OAFlux analysis) very well, indicating that the systematic bias in MIMR Qa could be diminished by a mean value adjustment based on the 1° OAFlux analysis. As we mentioned above, such an OAFlux-based mean value adjustment was indeed applied to satellite retrievals to construct the 0.25° OAFlux analysis.

A similar pattern of the mean difference in Ta is observed in the tropical oceans; that is, the warm Ta difference corresponds to the wet Qa difference. In

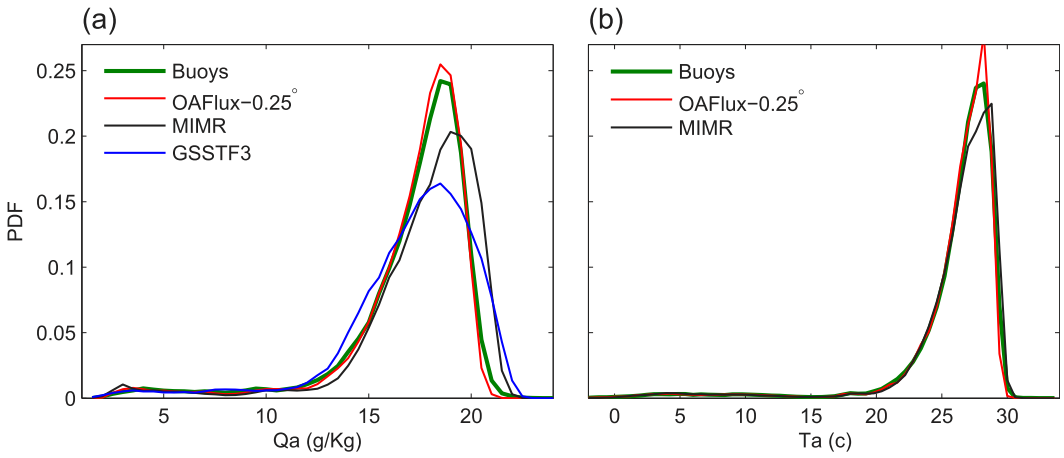

FIG. 11. Comparison of PDFs of (a) Qa and (b) Ta from the collocations between the two participating products and buoy measurements as described in Fig. 10.

general, the difference is less than 0.6°C. The most striking difference exists in the Pacific–Indian warm pool, the east-equatorial Pacific, and in the vicinity of the Gulf Stream boundary current. Compared to these extremes, the northern North Pacific is characterized by smaller differences.

The corresponding RMS difference between MIMR Qa and the buoy is overall larger than 1.0 g kg^{-1} (Fig. 6a). The most striking difference appears in the warm pool, the east-equatorial Pacific, and at various sites in the tropical Atlantic, with the maximum exceeding 1.6 g kg^{-1}. On the other hand, the largest RMS difference of 1.8°C for Ta occurs in the vicinity of the Gulf Stream boundary current (Fig. 6b). The RMS difference in the northern North Pacific is also significant, despite that the mean difference is small in that area. In the tropical oceans, except for the Pacific warm pool and the east equatorial Pacific, the RMS difference is less than 0.8°C.

Figure 7a shows the mean difference in Qa between GSSTF3 and the buoy and between GSSTF3 and the 1° OAFlux. The striking positive difference (exceeding 0.8 g kg^{-1}) between GSSTF3 and the buoy appears in the Pacific–Indian warm pool, the east-equatorial Pacific and eastern Atlantic, and the subtropical Atlantic. A large negative difference appears in the cold tongue region, the Arabian Sea, and the western Atlantic off the coast of Brazil. As might be expected, the pattern of the difference between GSSFT3 and the 1° OAFlux (which is represented by contours) is nearly the same as the one between GSSTF3 and the 0.25° OAFlux, as the latter was constructed under the constraint of the 1° OAFlux-based mean adjustment.

The corresponding RMS difference in Qa is found to be overall larger than 1.2 g kg^{-1}, except in the tropical Atlantic and northern North Pacific (Fig. 7b). The Pacific–Indian warm-pool, northeastern tropical Atlantic, Arabian Sea, Kuroshio Extension, and Gulf Stream regions, and southeast of the tropical Atlantic off the coast of Africa show the greatest RMS difference (>1.8°C).

b. *OAFlux versus the buoy*

Jin and Yu (2013) evaluated the OAFlux 0.25° flux products in resolving the air–sea exchange in the eddy-rich Gulf Stream region. Two approaches were used for evaluation: one was point-to-point validation based on six moored buoys in the region, and the other was basin-scale statistical analysis in terms of wavenumber spectra

TABLE 2. Comparison of four moments for buoys, OAFlux, MIMR, and GSSTF3, where S is skewness and K is kurtosis. There are a total of 130 855 collocations of daily mean Qa among the three participating products and buoy measurements from 1999 to 2008, and a total of 283 012 collocations of daily mean Ta between the two participating products and buoy measurements from 1999 to 2010.

	Ta				Qa			
	Mean(°C)	Std dev (°C)	S	K	Mean (g kg^{-1})	Std dev (g kg^{-1})	S	K
Buoys	25.94	19.15	−3.66	18.35	17.10	9.32	−2.31	9.59
OAFlux	25.91	18.87	−3.73	18.76	17.10	8.90	−2.53	10.77
MIMR	26.06	19.02	−3.58	17.64	17.62	10.23	−2.42	10.47
GSSTF3					17.14	10.70	−1.79	7.62

and the probability density function (PDF). The six buoys used in the validation include one that was deployed at 36°N, 65°W, close to the central location of the climatological maximum of turbulent heat fluxes, during the Climate and Ocean: Variability, Predictability and Change (CLIVAR) Mode Water Dynamic Experiment (CLIMODE) (Marshall et al. 2009; Weller et al. 2012); and another five offshore moored stations owned and maintained by NOAA's NDBC. There were a total of 4288 product/buoy collocations during the 3-yr period from 2005 to 2007. Results indicate that the mean differences in Qa and Ta with respect to the buoy measurements are $-0.06\,g\,kg^{-1}$ and 0.35°C, with an RMS difference of $0.77\,g\,kg^{-1}$ and 0.97°C, respectively. A further analysis of the time series over the CLIMODE buoy reveals that the Qa bias accounts for 11.7% of the variance of the latent heat flux bias, whereas the Ta bias accounts for 26.2% of the variance of sensitive heat flux bias. Large differences in the latent and sensible heat fluxes are primarily due to a mismatch in the SST between gridded data and point measurements associated with a highly variable current.

In this study, we evaluate the buoy comparison for Qa and Ta for the global oceans. Figure 8 shows the mean difference in Qa and Ta between OAFlux and the buoy over 137 buoy sites for the period 1988–2010. Despite that a large, dry Qa difference (exceeding $0.8\,g\,kg^{-1}$) is observed at various locations, across the board, OAFlux represents a major improvement over both MIMR and GSSTF3 with respect to buoy measurements. The OAFlux Ta is also well produced, in particular in the tropical oceans, where no obvious mean bias is apparent. Major differences were observed in the vicinity of the Gulf Stream boundary current where OAFlux overestimates Ta by up to more than 0.8°C, which can underestimate the sensible heat flux by $14\,W\,m^{-2}$ at wind speed at $12\,m\,s^{-1}$, compared to the three buoys, including CLIMODE, NDBC station 44018, and SESMOOR. The SESMOOR buoy was deployed in the winter of 1988/89 at 42.5°N, 61.2°W for the Experiment on Rapidly Intensifying Cyclones over the Atlantic at a position about 300 km southeast of Halifax, Nova Scotia, Canada. The NDBC 44018 was at 41.3°N, 69.3°W off Cape Cod, Massachusetts. All three buoys were located within the region of energetic mesoscale and synoptic variability. In fact, the largest RMS difference in Ta is evident in this region (Fig. 9). For example, the RMS difference is ~1.9°C at the SESMOOR site, where the standard deviation of the measured Ta is ~5.4°C. In contrast, the RMS difference is generally less than 0.6°C, while the standard deviation of the buoy Ta is less than 2.0°C in the tropical oceans. The RMS difference of Qa is overall less than $0.8\,g\,kg^{-1}$. A large difference exists at various locations but no major significant systematic bias against the buoy is observed.

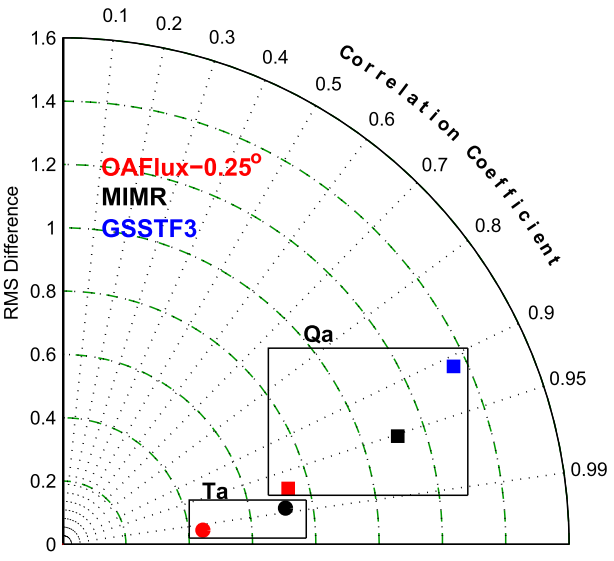

FIG. 12. Taylor diagram showing two statistical properties of the Qa (circles) and Ta (squares) comparison: the cc and RMS of the differences of the products vs buoys. A total of 130 855 collocations were used for Qa from OAFlux, MIMR, and GSSTF3, and a total of 283 012 collocations for Ta from OAFlux and MIMR.

c. Comparison of statistics

Comparisons of the collocated OAFlux, MIMR, GSSTF3, and the buoy daily mean Qa are shown in Figs. 10a–c. There are a total of 130 855 collocations

TABLE 3. Statistics of buoy evaluation for OAFlux, MIMR, and GSSTF3. There are a total of 130 855 collocations of daily mean Qa among the three participating products and buoy measurements from 1999 to 2008, and a total of 283 012 collocations of daily mean Ta between the two participating products and buoy measurements from 1999 to 2010. Three statistical properties are listed, including mean difference (Diff).

	Ta			Qa		
	Diff (°C)	RMSD (°C)	cc (0–1)	Diff ($g\,kg^{-1}$)	RMSD ($g\,kg^{-1}$)	cc (0–1)
OAFlux	−0.03	0.45	0.99	−0.00	0.73	0.97
MIMR	0.13	0.71	0.99	0.52	1.11	0.95
GSSTF3				0.03	1.36	0.91

TABLE 4. Estimated biases in LH and SH caused by the biases in Qa and Ta against buoy measurements for the collocated data from buoys, OAFlux, MIMR, GSSTF3, ERA-Interim, MERRA, CFSR, and NCEP-1.

	LH		SH	
	Diff ($W m^{-2}$)	RMSD ($W m^{-2}$)	Diff ($W m^{-2}$)	RMSD ($W m^{-2}$)
OAFlux	−0.2	18.6	0.4	4.8
MIMR	−11.7	27.7	−0.5	8.3
GSSTF3	4.9	33.3	—	—
ERA-Interim	14.9	22.2	2.3	5.2
MERRA	1.7	19.8	−0.3	5.6
CFSR	11.5	20.8	−0.7	5.1
NCEP-1	−7.3	26.1	2.0	8.4

among the three participating products and buoy measurements across 111 buoy sites from 1999 to 2008. Using the buoy measurements as a reference, the RMS difference for OAFlux is about 0.73 $g kg^{-1}$ accompanied by a mean difference near zero. The scatterplot of OAFlux versus the buoy depicts a nearly symmetric distribution about a perfect-fit line across the full range of Qa, except for a band around 10 $g kg^{-1}$, which was identified from the time series as being over the Stratus buoy, where the mean buoy Qa is about 10.5 $g kg^{-1}$ and the OAFlux overestimates by 0.46 $g kg^{-1}$.

In comparison with the buoy, MIMR is systematically drier at low Qa (i.e., the center core of the distribution shifts to below the perfect-fit line) and wetter at high Qa (i.e., the core of the distribution shifts to above the perfect-fit line) (Fig. 10b). The mean difference between MIMR and the buoy is about 0.52 $g kg^{-1}$, with an RMS difference of 1.11 $g kg^{-1}$.

The scatterplot of GSSTF3 versus the buoy exhibits a relatively large amount of scatter. The center core of the distribution departs from the perfect-fit line toward a more positive bias direction with increasing Qa. Despite that the mean difference is small (~0.03 $g kg^{-1}$), the RMS difference (~1.36 $g kg^{-1}$) is larger in comparison with those of OAFlux and MIMR.

Comparisons of the collocated OAFlux, MIMR, and the buoy daily mean Ta are shown in Figs. 10d,e. There are a total of 283 012 collocations for the period from 1999 to 2010. In comparison with MIMR, OAFlux shows a better linear relationship with respect to the buoy. The mean difference for OAFlux is about −0.03°C with an RMS difference of 0.45°C, compared to the mean difference of 0.13°C with an RMS difference of 0.71°C for MIMR.

Figure 11 shows PDFs of the collocated Qa and Ta from the buoys and three satellite-based datasets. The values for four moments for buoys, OAFlux, MIMR, and GSSTF3 are summarized in Table 2. It is evident that OAFlux and buoy Qa are near the same distribution. The distribution is not Gaussian in nature; instead, it is highly skewed to high Qa, as expected because most of the buoys are located in the tropical oceans. Table 2 shows that the respective values of skewness and kurtosis are quite similar for both buoy and MIMR, despite that MIMR shifted in the location from the buoy due to the overall wet bias. In contrast, GSSTF3 tends to be more stretched out toward normal distribution. The distributions of Ta agree very well among the buoy, OAFlux, and MIMR.

To summarize the comparisons of the statistics among OAFlux, MIMR, and GSSTF3 with respect to the buoy observations, a Taylor diagram displaying the correlation coefficients (cc) and the RMS difference between the three products and the buoy is shown in Fig. 12. The statistics of the mean difference, RMS difference, and cc is summarized in Table 3. It is evident that OAFlux is the

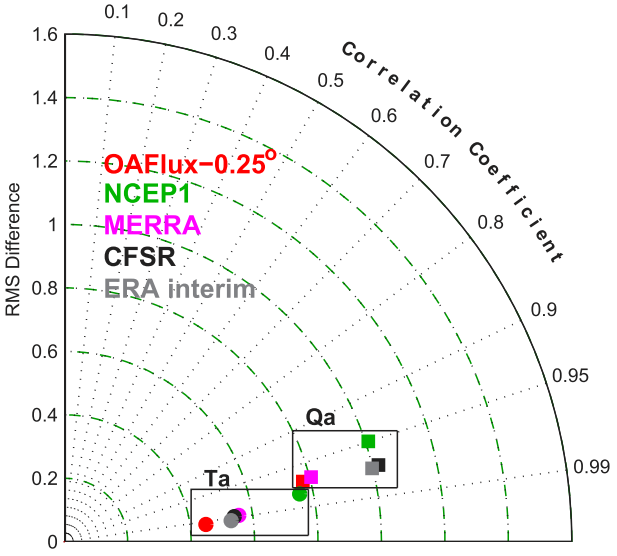

FIG. 13. Taylor diagram showing two statistical properties of the Qa (squares) and Ta (circles) comparison: the cc and RMS of the differences of the products vs buoys. A total of 273 957 for Qa and 437 011 for Ta collocations were used from OAFlux, NCEP-1, MERRA, CFSR, and ERA-Interim.

TABLE 5. Statistics of buoy evaluation for OAFlux, ERA-Interim, MERRA, CFSR, and NCEP-1. There are a total of 273 957 and 437 011 collocations from the five participating products and buoy measurements from 1988 to 2009.

	Ta			Qa		
	Diff (°C)	RMSD (°C)	cc (0–1)	Diff (g kg^{-1})	RMSD (g kg^{-1})	cc (0–1)
OAFlux	−0.04	0.45	0.99	−0.04	0.78	0.97
ERA-Interim	−0.25	0.53	0.99	−0.67	1.00	0.97
MERRA	0.15	0.56	0.99	−0.03	0.80	0.97
CFSR	0.03	0.54	0.99	−0.66	1.02	0.97
NCEP-1	−0.18	0.76	0.98	0.17	1.01	0.95

best in both Qa and Ta among the three datasets, and GSSTF3 is less favored with regard to the buoy comparison. The estimated biases in the latent and sensible heat fluxes caused by the biases in OAFlux Qa and Ta are −0.2 and 0.4 W m^{-2}, with RMS differences of 18.6 and 4.8 W m^{-2}, respectively. Overall, wet bias of MIMR can result in underestimates of latent heat flux by 11.7 W m^{-2} with an RMS difference of 27.3 W m^{-2}. The estimated RMS difference in latent heat flux caused by the bias of GSSTF3 Qa is 33.3 W m^{-2}, which is the largest among the three datasets (Table 4).

In this study, the buoy validation has focused on OAFlux and the two satellite-based products. The validation for the four atmospheric reanalyses is therefore simply summarized by descriptive statistics given in Fig. 13 and Table 5. The statistics for OAFlux are also listed as a reference. There are a total of 273 957 and 437 011 collocations from OAFlux, NCEP-1, MERRA,

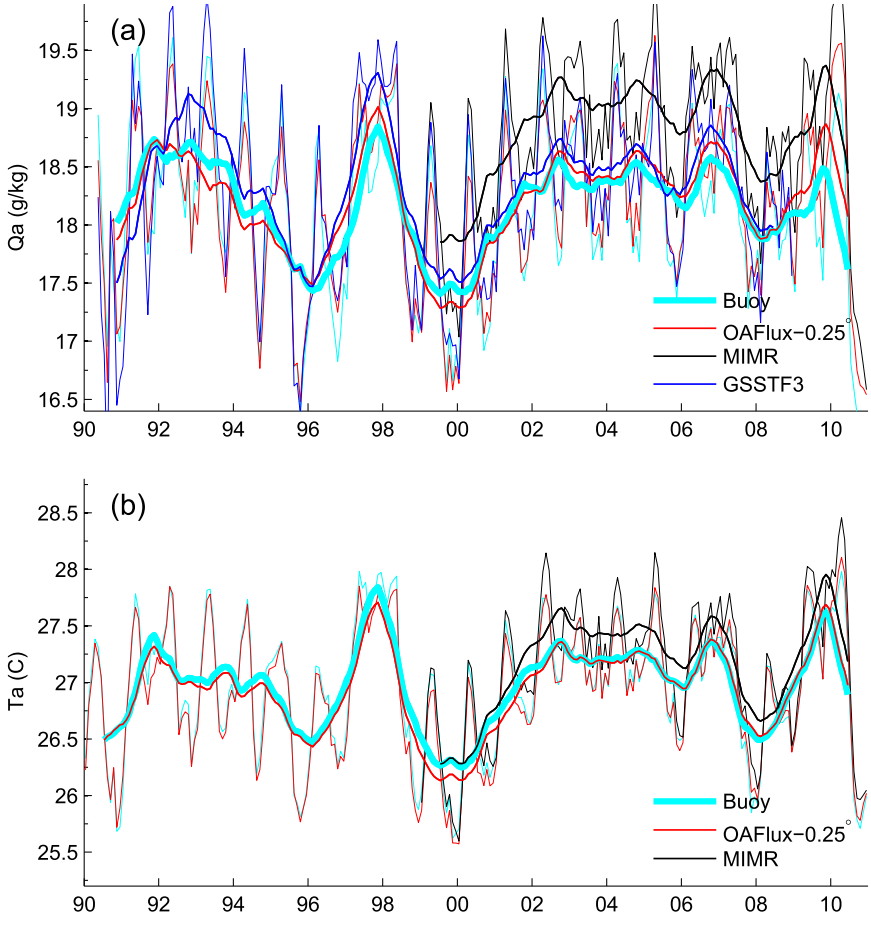

FIG. 14. (a) Time series of monthly-mean Qa from the buoy (cyan), OAFlux (red), MIMR (black), and GSSTF3 (blue), averaged over the TAO array. The thick lines represent the time series with a 13-month running mean. (b) As in (a), but for Ta. GSSTF3 Ta is not included.

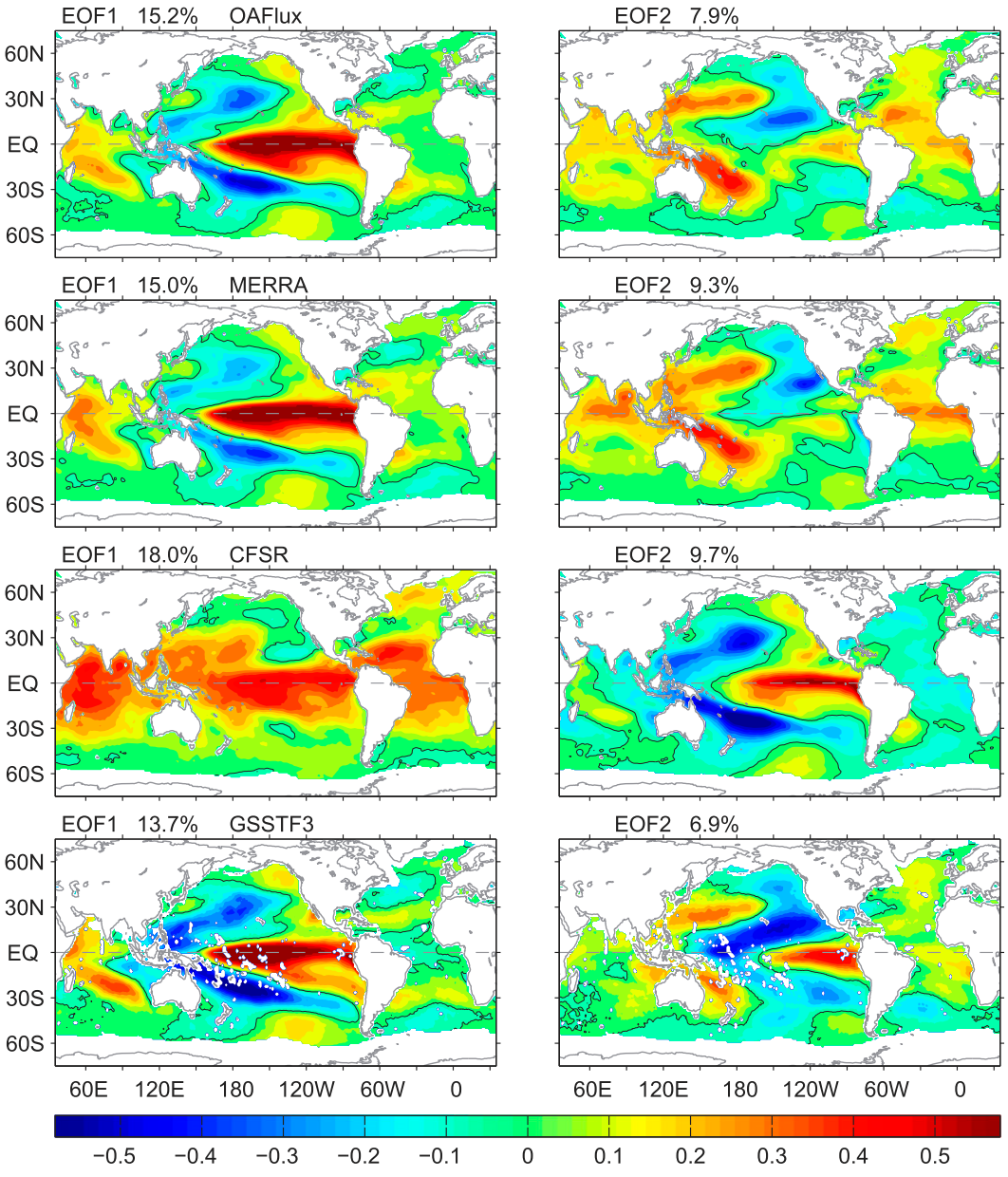

FIG. 15. First two EOF patterns—(left) EOF1 and (right) EOF2—of monthly-mean anomalies of Qa for (top to bottom) OAFlux, MERRA, CFSR, and GSSTF3, from 1988 to 2008.

CFSR, and ERA-Interim for Qa and Ta, respectively, over the period 1988–2009. Among the five products, OAFlux shows the best agreement with the buoy in both Qa and Ta. On the other hand, the MERRA Qa stands out as the best among the four reanalyses and, in fact, is very close to OAFlux. Note that the validation in the Gulf Stream region indicated that MERRA is too smooth to resolve small-scale variability in Ta and Qa (Jin and Yu 2013). Clearly, MERRA gets quite good agreement against measured data outside the eddy regions. The validation gives similar statistics for CFSR and ERA-Interim, which is slightly better than MERRA for Ta. NCEP-1 shows the largest difference in Ta and Qa against the buoy observations among the four reanalyses. Note that in comparison with the satellite retrievals, which did not assimilate buoy observations, the reanalyses show a better agreement with buoy observations of Ta and Qa.

d. Time series at TAO buoy array

The TAO buoys were first deployed in the early 1980s (McPhaden et al. 1998). The number of buoys was about

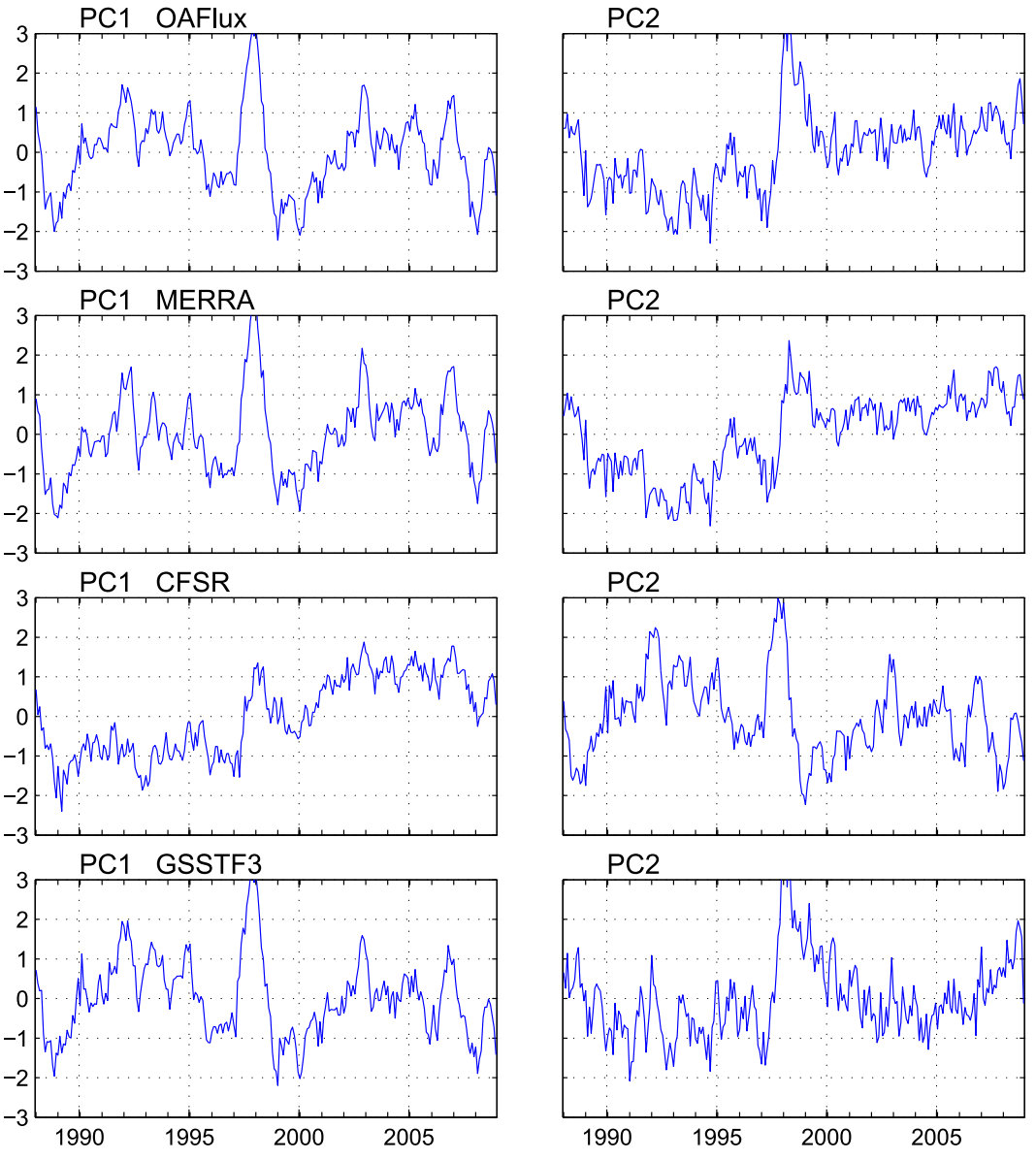

FIG. 16. Corresponding normalized PCs of the two leading modes shown in Fig. 15.

15 in 1988, and then increased rapidly to more than 60 by mid-1992. These buoy time series of ~20 yr provide a valuable reference for validating the consistency of estimates of Qa and Ta over the satellite era. Figure 14a shows the time series of monthly-mean Qa from the buoy, OAFlux, MIMR, and GSSTF3, averaged over the TAO array. The thick lines represent the 13-month running means. The three satellite-based products are sampled with the same spatial and temporal coverage as the buoy. The OAFlux Qa time series tracks the buoy very well, whereas MIMR follows the buoy but overestimates Qa by ~0.6 g kg^{-1}. Despite that the mean differences between GSSTF3 and the TAO buoy are considerably large at each individual buoy location, the time series of GSSTF3, which represents a group mean of Qa across the array at each time step, matches the time series of the buoy very well, except that the GSSTF3 Qa is overestimated in the early 1990s, so that it shows a slight downward tendency compared to the buoy time series. The time series of Ta depicts a good consistency between OAFlux and the buoy throughout the entire analysis period (Fig. 14b). The values of MIMR Qa are very close to the buoy for the period 1999–2001 and show a steady overestimation of 0.2°–0.3°C against the buoy time series.

5. EOF-based intercomparison

The buoy validation shows that OAFlux represents an improvement over the satellite retrievals and the four atmospheric reanalyses in Ta and Qa. It should be noted that the buoy validation was based on an uneven spatial–temporal sampling and was performed at limited locations, so it is not sufficient to provide an integrated perspective over the global basin scale. To evaluate the consistency of spatial–temporal variations of the estimated OAFlux Qa and Ta over the global domain, an intercomparison among OAFlux, MIMR, GSSTF3, and the four reanalyses was carried out using an EOF analysis.

Figure 15 shows the leading two EOFs of the monthly-mean anomalies of Qa for the time period 1988–2008 for OAFlux, MERRA, CFSR, and GSSTF3, while that for ERA-Interim and NCEP-1 are very similar to MERRA and are not shown. The corresponding principal components (PCs) are shown in Fig. 16. The monthly-mean anomaly was computed by removing the monthly-mean climatology for the period 1988–2008. It is clear that OAFlux and MERRA have similar patterns, whereas CFSR and GSSTF3 are considerably different from each other and from others. The leading EOF of OAFlux over the global oceans is the El Niño–Southern Oscillation (ENSO) mode, indicating clearly the influence of SST on Qa. The spatial pattern associated with the warm phase of ENSO exhibits positive Qa anomalies that are related to extra heat and evaporation across the east tropical Pacific, and negative Qa anomalies in the west and over the North and South Pacific. This single mode accounts for 15.2% of nonseasonal total variance of Qa over the global oceans for over 21 yr. The second EOF mode (~7.9% of total variance) represents a decadal variability with an upward trend in Qa since 1993 embedded with intense interannual variability. The 1997/98 El Niño influence is evident. An opposite polarity is observed in the North Pacific, where the Qa increases in the northwest but decreases in the northeast over the same period.

In contrast to OAFlux, the leading mode of CFSR represents a nearly uniform pattern over the tropical oceans associated with a rapid rise in Qa. This mode accounts for 18.0% of the total variance. The second EOF mode of CFSR (~9.7% of total variance) correlates with ENSO, but it differs from the EOF mode of OAFlux in the tropical Atlantic, where the CFSR Qa anomalies are negative rather than positive during ENSO's warm-phase years.

The leading EOF mode of GSSTF3 (~13.7% of total variance) is similar to that of OAFlux over the global ocean, except for the tropical Atlantic. The major difference between GSSFT3 and OAFlux is in the second mode; for example, GSSTF3 shows significant negative anomalies in the west equatorial Pacific and a different variability in the PCs from 2000 to 2005. As a result, GSSTF3 has a different decadal variability in the basin-averaged Qa compared to that indicated by OAFlux.

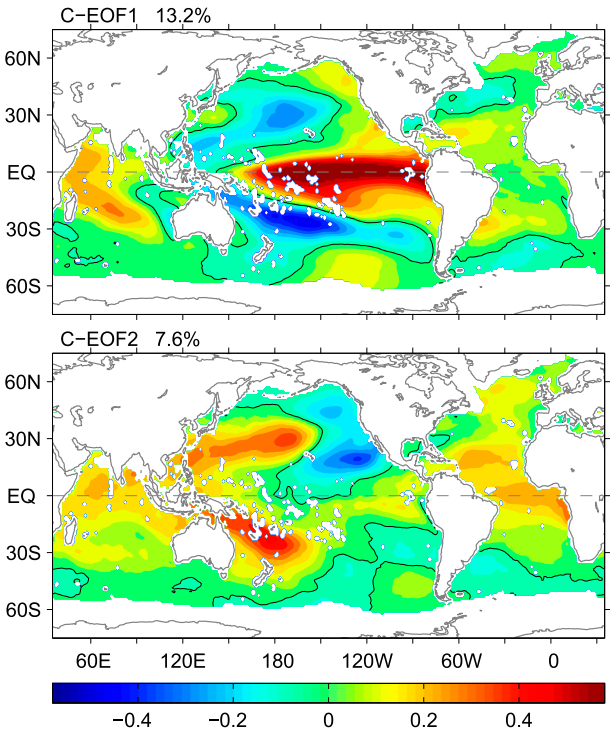

FIG. 17. (top) First and (bottom) second common EOFs for the combined Qa from six datasets, i.e., OAFlux, GSSTF3, CFSR, MERRA, ERA-Interim, and NCEP-1. The unit is $g\,kg^{-1}$ per standard deviation of the corresponding PC.

To ensure common features among the different products, we used the technique of common EOF analysis (Barnett 1999). This technique combines OAFlux, GSSTF3, and the four reanalyses into a single dataset, of which the data on common grids are combined along the time axis, and an EOF analysis is applied to the combined dataset. The two leading common EOF modes are very similar to the respective individual OAFlux EOFs (Fig. 17).

Consistency in spatial–temporal variability patterns among OAFlux and reanalyses, except for CFSR, which was not utilized in the synthesis, might be expected. In fact, the impact of GSSTF3 on the OAFlux synthesis from 1988 to 2000 was compromised by merging with the reanalyses. On the other hand, MIMR (1999–2010) has very two similar leading EOF patterns as the reanalyses used in the synthesis (not shown).

Figure 18 shows the two leading EOF modes of the monthly-mean anomalies of Ta for OAFlux, MERRA, and CFSR from 1988 to 2008. The corresponding PCs

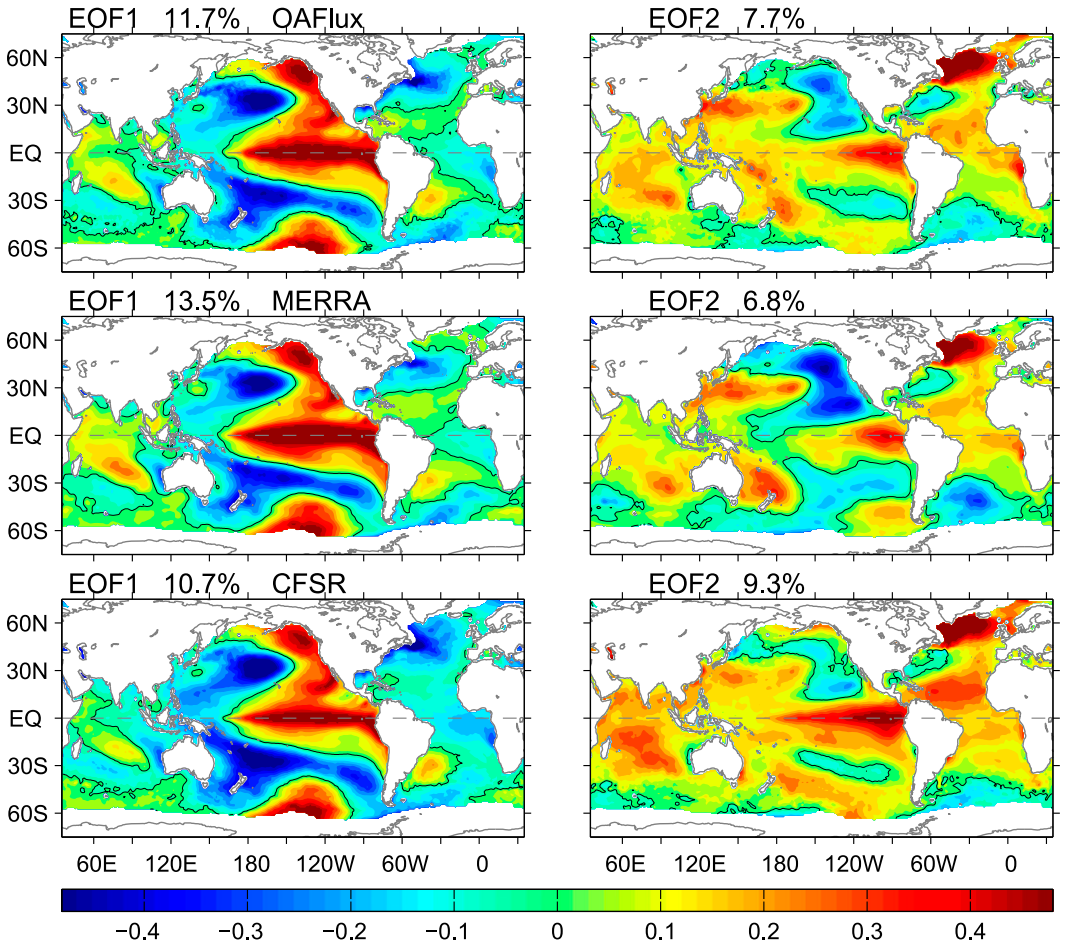

FIG. 18. As in Fig. 15, but for Ta for (top to bottom) OAFlux, MERRA, and CFSR.

are shown in Fig. 19. The EOF modes for ERA-Interim and NCEP-1 are very similar to that of MERRA and are not shown. The leading EOF mode correlates with ENSO but depicts a large anomaly in the northeast North Pacific. The second EOF mode is associated with basin-scale warming. Note that although the first EOF mode of CFSR (~10.7% of total variance) is the ENSO mode, it is statistically mixed up with its second EOF mode (~9.3% of variability) in terms of the North et al. (1982) criterion. Moreover, it differs slightly from OAFlux and MERRA; for example, the anomalies in the tropical Atlantic are negative overall and are out of phase with the anomalies in the eastern tropical Pacific. On the other hand, OAFlux and MERRA have very similar EOF patterns.

6. Summary and conclusions

This paper used 137 buoy time series as a benchmark to assess a daily, 0.25° gridded global ocean near-surface Qa and Ta developed by the OAFlux. The construction of the Qa and Ta used the same methodology for the 1° OAFlux analysis, that is, merging of two satellite-based datasets provided by MIMR (1999–2010) and GSSTF3 (July 1987–December 2000), the 1° OAFlux analysis, and the three atmospheric reanalyses. An intercomparison between OAFlux and the two satellite-based products was performed based on the total 139 053 collocated daily mean data for Qa and the total 294 238 collocated daily mean data for Ta over archived buoy sites deployed from 1999 to 2010. The buoy comparison shows that OAFlux has a lower mean difference and a smaller RMS difference in both Qa and Ta in comparison with MIMR and GSSTF3. The RMS difference in Qa for OAFlux is about 0.73 g kg^{-1}, compared to 1.11 and 1.36 g kg^{-1} for MIMR and GSSTF3, respectively. The RMS difference in Ta for OAFlux is about 0.45°C, compared to 0.71°C for MIMR. The GSSTF3 Ta was taken from the NCEP–DOE analysis and therefore was not included in this study.

No major systematic bias between OAFlux and the buoy was observed across all selected buoy locations, except in the vicinity of the Gulf Stream boundary

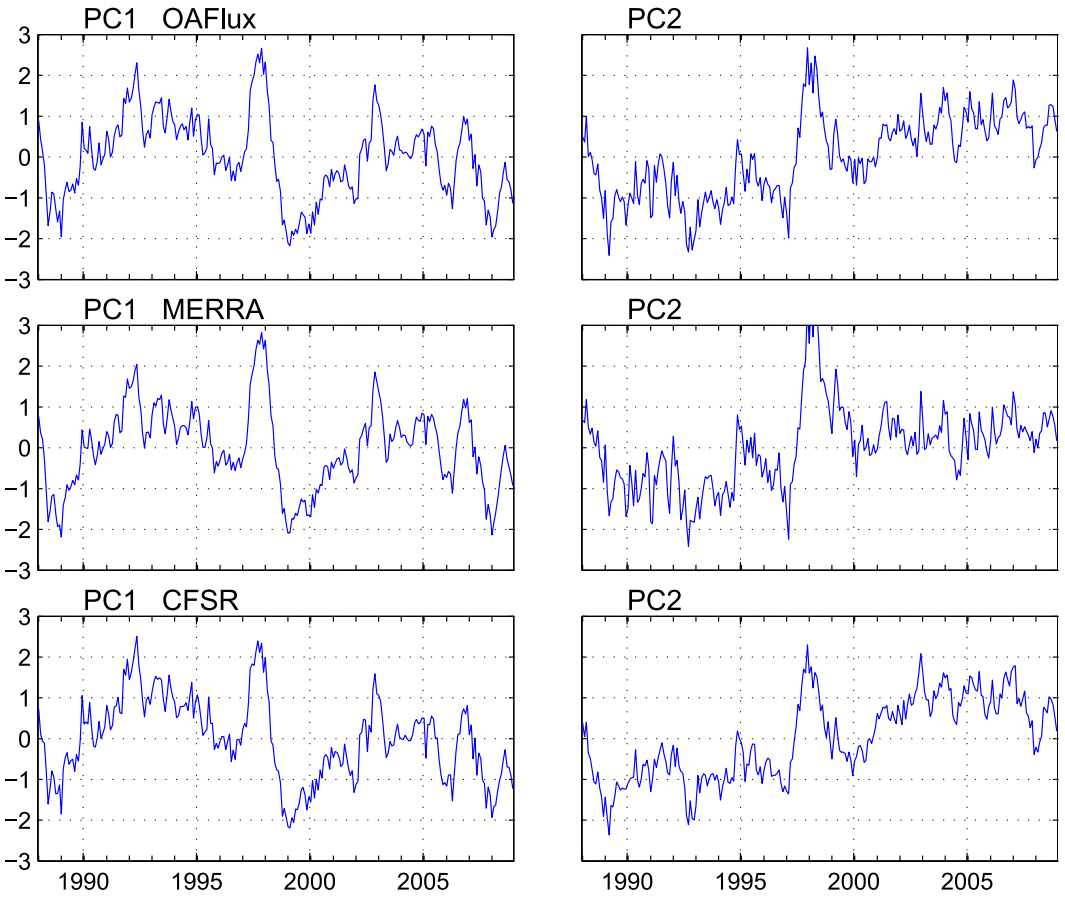

FIG. 19. Corresponding normalized PCs of the two leading modes shown in Fig. 18.

current, where the OAFlux overestimated Ta by more than 0.8°C, with a maximum RMS difference exceeding 1.8°C. On the other hand, MIMR overestimates both Qa and Ta over the tropical oceans but underestimates Qa in the northern North Pacific. A comparison of Qa between GSSTF3 and the buoy shows a mixture of striking positive and negative difference in the low to mid-latitudes. The corresponding RMS difference in Qa is overall larger than 1.2 g kg^{-1}.

An intercomparison between OAFlux and the four reanalyses indicates that OAFlux has the best agreement with the buoy in both Qa and Ta. Among the four reanalyses, MERRA is found to agree better in Qa with buoy observations. CFSR and ERA-Interim are comparable in terms of the validation in both Qa and Ta. The comparison of the global mean EOF analysis indicates that OAFlux has a similar spatial–temporal variability pattern with that of MERRA, NCEP-1, and ERA-Interim, and differs from CFSR and GSSTF3.

Acknowledgments. This study was supported by the NOAA Ocean Climate Observation (OCO) program under Grant NA09OAR4320129. The OAFlux 1° analysis is available from the project website (http://oaflux.whoi.edu/). The MIMR near-surface air temperature and specific humidity are available from the NOAA Earth System Research Laboratory (ftp://ftp1.esrl.noaa.gov). The GSSTF3 specific humidity data were obtained from NASA (ftp://measures.gsfc.nasa.gov/data/s4pa/GSSTF/GSSTF.3/). The WHOI and SOFS buoy measurements were downloaded (http://uop.whoi.edu/projects/), as were the NDBC buoy measurements (http://www.ndbc.noaa.gov/) and the TAO/TRITON, PIRATA and RAMA data (http://www.pmel.noaa.gov/tao). KEO, PAPA, and ARC data were obtained from the NOAA Pacific Marine Environmental Laboratory. The ERA-Interim, CFSR, and NCEP reanalyses were downloaded from the NCAR Research Data Archive (http://rda.ucar.edu). MERRA data were downloaded from the Global Modeling and Assimilation Office (GMAO) and the GES DISC (ftp://goldsmr2.sci.gsfc.nasa.gov).

REFERENCES

Andersson, A., C. Klepp, K. Fennig, S. Bakan, H. Grassl, and J. Schulz, 2011: Evaluation of HOAPS-3 ocean surface

freshwater flux components. *J. Appl. Meteor. Climatol.,* **50,** 379–398, doi:10.1175/2010JAMC2341.1.

Barnett, T. P., 1999: Comparison of near-surface air temperature variability in 11 coupled global climate models. *J. Climate,* **12,** 511–518, doi:10.1175/1520-0442(1999)012<0511:CONSAT>2.0.CO;2.

Bentamy, A., K. B. Katsaros, A. M. Mestas-Nuñez, W. M. Drennan, E. B. Forde, and H. Roquet, 2003: Satellite estimates of wind speed and latent heat flux over the global oceans. *J. Climate,* **16,** 637–656, doi:10.1175/1520-0442(2003)016<0637:SEOWSA>2.0.CO;2.

Bourlès, B., and Coauthors, 2008: The PIRATA program: History, accomplishments and future directions. *Bull. Amer. Meteor. Soc.,* **89,** 1111–1125, doi:10.1175/2008BAMS2462.1.

Chou, S.-H., R. M. Atlas, C.-L. Shie, and J. Ardizzone, 1995: Estimates of surface humidity and latent heat fluxes over oceans from SSM/I data. *Mon. Wea. Rev.,* **123,** 2405–2425, doi:10.1175/1520-0493(1995)123<2405:EOSHAL>2.0.CO;2.

——, E. Nelkin, J. Ardizzone, R. M. Atlas, and C.-L. Shie, 2003: Surface turbulent heat and momentum fluxes over global oceans based on the Goddard satellite retrievals, version 2 (GSSTF2). *J. Climate,* **16,** 3256–3273, doi:10.1175/1520-0442(2003)016<3256:STHAMF>2.0.CO;2.

Colbo, K., and R. A. Weller, 2009: The accuracy of the IMET sensor package in the subtropics. *J. Atmos. Oceanic Technol.,* **26,** 1867–1890, doi:10.1175/2009JTECHO667.1.

Cronin, M. F., and Coauthors, 2010: Monitoring ocean-atmosphere interactions in western boundary current extensions. *Proceedings of the OceanObs'09: Sustained Ocean Observations and Information for Society,* J. Hall, D. E. Harrison, and D. Stammer, Eds., Vol. 2, ESA Publ. WPP-306, doi:10.5270/OceanObs09.cwp.2.

Curry, J. A., and Coauthors, 2004: SEAFLUX. *Bull. Amer. Meteor. Soc.,* **85,** 409–424, doi:10.1175/BAMS-85-3-409.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

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

Fairall, C. W., E. F. Bradley, J. E. Hare, A. A. Grachev, and J. B. Edson, 2003: Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm. *J. Climate,* **16,** 571–591, doi:10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2.

Gao, S., L. S. Chiu, and C.-L. Shie, 2013: Trends and variations of ocean surface latent heat flux: Results from GSSTF2c data set. *Geophys. Res. Lett.,* **40,** 380–385, doi:10.1029/2012GL054620.

Gulev, S. K., T. Jung, and E. Ruprecht, 2007: Estimation of the impact of sampling errors in the VOS observations on air–sea fluxes. Part I: Uncertainties in climate means. *J. Climate,* **20,** 279–301, doi:10.1175/JCLI4010.1.

Jackson, D. L., and G. A. Wick, 2010: Near-surface air temperature retrieval derived from AMSU-A and sea surface temperature observations. *J. Atmos. Oceanic Technol.,* **27,** 1769–1776, doi:10.1175/2010JTECHA1414.1.

——, ——, and J. J. Bates, 2006: Near-surface retrieval of air temperature and specific humidity using multisensor microwave satellite observations. *J. Geophys. Res.,* **111,** D10306, doi:10.1029/2005JD006431.

——, ——, and F. R. Robertson, 2009: Improved multisensor approach to satellite-retrieved near-surface specific humidity observations. *J. Geophys. Res.,* **114,** D16303, doi:10.1029/2008JD011341.

Jin, X., and L. Yu, 2013: Assessing high-resolution analysis of surface heat fluxes in the Gulf Stream region. *J. Geophys. Res.,* **118,** 5353–5375, doi:10.1002/jgrc.20386.

Josey, S. A., 2001: A comparison of ECMWF, NCEP–NCAR, and SOC surface heat fluxes with moored buoy measurements in the subduction region of the northeast Atlantic. *J. Climate,* **14,** 1780–1789, doi:10.1175/1520-0442(2001)014<1780:ACOENN>2.0.CO;2.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kamphaus, R., M. Cronin, C. Sabine, S. Emerson, C. Meinig, and M. Robert, 2008: New surface mooring at Station Papa monitors climate. *PICES Press,* No. 2, PICES Secretariat, Sidney, BC, Canada, 26–27.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, J. Potter, and M. Fiorino, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83,** 1631–1643, doi:10.1175/BAMS-83-11-1631.

Kubota, M., N. Iwasaka, S. Kizu, M. Konda, and K. Kutsuwada, 2002: Japanese Ocean Flux Data Sets with Use of Remote Sensing Observations (J-OFURO). *J. Oceanogr.,* **58,** 213–225, doi:10.1023/A:1015845321836.

Liu, W. T., 1988: Moisture and latent heat flux variabilities in the tropical Pacific derived from satellite data. *J. Geophys. Res.,* **93,** 6749–6760, doi:10.1029/JC093iC06p06749.

——, K. B. Katsaros, and J. A. Businger, 1979: Bulk parameterization of the air–sea exchange of heat and water vapor including the molecular constraints at the interface. *J. Atmos. Sci.,* **36,** 1722–1735, doi:10.1175/1520-0469(1979)036<1722:BPOASE>2.0.CO;2.

Marshall, J., and Coauthors, 2009: The CLIMODE Field Campaign: Observing the cycle of convection and restratification over the Gulf Stream. *Bull. Amer. Meteor. Soc.,* **90,** 1337–1350, doi:10.1175/2009BAMS2706.1 .

McPhaden, M. J., and Coauthors, 1998: The Tropical Ocean-Global Atmosphere (TOGA) observing system: A decade of progress. *J. Geophys. Res.,* **103,** 14 169–14 240, doi:10.1029/97JC02906.

——, and Coauthors, 2009: RAMA: The Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction. *Bull. Amer. Meteor. Soc.,* **90,** 459–480, doi:10.1175/2008BAMS2608.1.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.,* **110,** 699–706, doi:10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution blended analyses for sea surface temperature. *J. Climate,* **20,** 5473–5496, doi:10.1175/2007JCLI1824.1.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* **24,** 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Roberts, J. B., C. A. Clayson, F. R. Robertson, and D. L. Jackson, 2010: Predicting near-surface atmospheric variables from Special Sensor Microwave/Imager using neural networks with a first-guess approach. *J. Geophys. Res.,* **115,** D19113, doi:10.1029/2009JD013099.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.,* **91,** 1015–1057, doi:10.1175/2010BAMS3001.1.

Schlüssel, P., L. Schanz, and G. Englisch, 1995: Retrieval of latent heat flux and longwave irradiance at the sea surface

from SSM/I and AVHRR measurements. *Adv. Space Res.,* **16,** 107–115, doi:10.1016/0273-1177(95)00389-V.

Schulz, J., P. Schlüssel, and J. Grassl, 1993: Water vapor in the atmospheric boundary layer over oceans from SSM/I measurements. *Int. J. Remote Sens.,* **14,** 2773–2789, doi:10.1080/01431169308904308.

——, J. Meywerk, S. Ewald, and P. Schlüssel, 1997: Evaluation of satellite-derived latent heat fluxes. *J. Climate,* **10,** 2782–2795, doi:10.1175/1520-0442(1997)010<2782:EOSDLH>2.0.CO;2.

Shie, C.-L., 2010a: A recently revived dataset of satellite-based global air-sea surface turbulent fluxes (GSSTF2b)—Features and applications. *17th Conf. on Satellite Meteorology and Oceanography,* Annapolis, MD, Amer. Meteor. Soc., J1.1. [Available online at https://ams.confex.com/ams/17Air17Sat9Coas/techprogram/paper_174182.htm.]

——, 2010b: Science background for the reprocessing and Goddard Satellite-based Surface Turbulent Fluxes (GSSTF2b) data set for global water and energy cycle research. NASA GES DISC, Version 1, 18 pp. [Available from online at ftp://meso-a.gsfc.nasa.gov/pub/shieftp/fluxdocu/gsstf2b/Science_of_the_data.GSSTF2b.pdf.]

——, K. Hilburn, L. S. Chiu, R. Adler, I.-I. Lin, E. Nelkin, J. Ardizzone, and S. Gao, 2012: Goddard satellite-based surface turbulent fluxes, daily grid, version 3. A. Savtchenko, Ed., Goddard Earth Science Data and Information Services Center, doi:10.5067/MEASURES/GSSTF/DATA301.

Weller, R. A., and S. P. Anderson, 1996: Temporal variability and mean values of the surface meteorology and air–sea fluxes in the western equatorial Pacific warm pool during TOGA COARE. *J. Climate,* **9,** 1959–1990, doi:10.1175/1520-0442(1996)009<1959:SMAASF>2.0.CO;2.

——, P. B. Sebastien, J. Lord, J. D. Ware, and J. B. Edson, 2012: A surface mooring for air–sea interaction research in the Gulf Stream. Part I: Mooring design and instrumentation. *J. Atmos. Oceanic Technol.,* **29,** 1363–1376, doi:10.1175/JTECH-D-12-00060.1.

Yu, L., and R. A. Weller, 2007: Objectively analyzed air–sea heat fluxes for the global ice-free oceans (1981–2005). *Bull. Amer. Meteor. Soc.,* **88,** 527–539, doi:10.1175/BAMS-88-4-527.

——, and X. Jin, 2012: Buoy perspective of a high-resolution global ocean vector wind analysis constructed from passive radiometers and active scatterometers (1987–present). *J. Geophys. Res.,* **117,** C11013, doi:10.1029/2012JC008069.

——, ——, and R. A. Weller, 2008: Multidecade global flux datasets from the Objectively Analyzed Air-Sea Fluxes (OAFlux) Project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables. Woods Hole Oceanographic Institution OAFlux Project Tech. Rep. OA-2008-01, 64 pp.