Global Air-Sea Fluxes of Heat, Freshwater, and Momentum: Energy Budget Closure and Unanswered Questions

Lisan Yu

Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA; email: lyu@whoi.edu

Keywords

global-ocean energy budget, global-ocean freshwater budget, air–sea heat flux, air–sea freshwater flux, wind stress, air–sea buoy measurements, atmospheric reanalysis, satellite observations

Abstract

The ocean interacts with the atmosphere via interfacial exchanges of momentum, heat (via radiation and convection), and fresh water (via evaporation and precipitation). These fluxes, or exchanges, constitute the oceansurface energy and water budgets and define the ocean’s role in Earth’s climate and its variability on both short and long timescales. However, direct flux measurements are available only at limited locations. Air–sea fluxes are commonly estimated from bulk flux parameterization using flux-related near-surface meteorological variables (winds, sea and air temperatures, and humidity) that are available from buoys, ships, satellite remote sensing, numerical weather prediction models, and/or a combination of any of these sources. Uncertainties in parameterization-based flux estimates are large, and when they are integrated over the ocean basins, they cause a large imbalance in the global-ocean budgets. Despite the significant progress that has been made in quantifying surface fluxes in the past 30
years, achieving a global closure of ocean-surface energy and water budgets remains a challenge for flux products constructed from all data sources. This review provides a personal perspective on three questions: First, to what extent can time-series measurements from air–sea buoys be used as benchmarks for accuracy and reliability in the context of the budget closures? Second, what is the dominant source of uncertainties for surface flux products, the flux-related variables or the bulk flux algorithms? And third, given the coupling between the energy and water cycles, precipitation and surface radiation can act as twin budget constraints—are the community-standard precipitation and surface radiation products pairwise compatible?

1. INTRODUCTION

The ocean’s role in climate is manifested in its ability to transport heat poleward and regulate climate variability through exchange of heat, fresh water, and momentum with the atmosphere (e.g., Trenberth & Caron 2001, Wunsch 2005, Stephens et al. 2012, Wild et al. 2013). The fluxes, or exchange, at the air–sea interface are fundamental processes for keeping the global climate system in balance with the incoming insolation at Earth’s surface (Loeb et al. 2012, Trenberth et al. 2014). They are also a primary conduit for coupling and feedback between the ocean and atmosphere on a broad range of scales, from synoptic weather events to regional and global circulation systems (e.g., Drennan et al. 2007, Føre et al. 2012, Gulev & Belyaev 2012, Drijfhout et al. 2014, Soloviev et al. 2014). Uncertainties in air–sea fluxes challenge our ability to understand how the ocean interacts with the atmosphere to influence the climate patterns worldwide, and how the interaction can be represented in Earth system models to improve the prediction of extreme weather events at long lead times. Air–sea flux products with not only high quality but also continuous and consistent climate records are sought to serve the needs of ocean
and climate communities for the characterization, attribution, and modeling of weather and
climate variability in the atmosphere and ocean (e.g., WGASF 2000, Curry et al. 2004, Fairall et
al. 2010, Gulev et al. 2010).

Significant progress has been made in the past four decades in understanding and
measuring the turbulent motions near the air–sea boundary (e.g., breaking waves, turbulence, sea
spray, rain, and surface films) and their cumulative effects on the rates of transports of
heat, moisture, and momentum across the interface (e.g., Louis 1979, Large & Pond 1981, Andreas
direct covariance (or eddy correlation) technique (Crawford et al. 1993) has so far been the only
established means for direct flux measurements at sea (e.g., Edson et al. 1998, Landwehr et al.
2015). However, direct flux measurements are currently available only at a limited number of
locations for limited durations, because the measurements of vertical winds as well as
temperature and humidity fluctuations need to be conducted on specially designed ships or buoys
to minimize the effects of flow distortion and turbulent injection induced by the moving
platforms. Air–sea fluxes in numerical models and global data products are computed from flux
parameterizations that link the microscale turbulent transfers to easily measured macroscale
quantities such as near-surface wind, humidity, and temperature. Sophisticated parameterizations
have been developed, including the inertial-dissipation method, which infers surface fluxes from
spectral characteristics of the inertial subrange (Fairall & Larsen 1986); the mean flux-profile
method, which utilizes the empirical relationships between surface fluxes and mean profiles
(gradients) of observed quantities in the surface layer (Paulson et al. 1972, Blanc 1983); and the
bulk aerodynamic method, which employs the Monin–Obukhov similarity theory
(Monin & Obukhov 1954, Garratt 1977, Large & Pond 1981). The bulk approach provides scaling
relationships between surface fluxes and profiles of mean variables in the surface layer, and it
determines the transfer coefficients from either empirically derived flux profiles (Liu et al. 1979)
or direct covariance experiments (Fairall et al. 1996, 2003; Edson et al. 2013).

Of all types of parameterizations, the bulk aerodynamic parameterization is and will
continue to be significant for air–sea flux estimation due to its easy applicability. The required
input information of near-surface meteorology is routinely available from voluntary observing
ships (VOSs), satellite remote sensing, and numerical weather prediction models. The algorithm
developed during the Tropical Ocean–Global Atmosphere (TOGA) Coupled Ocean–Atmosphere
Response Experiment (COARE) (Fairall et al. 1996, 2003; Edson et al. 2013) represents the state
of the art in accuracy (Brunke et al. 2003) and has been used widely in constructing global air–
sea flux gridded products using satellite and ship observations.

Using bulk parameterization, one can approximate surface turbulent momentum, heat,
and freshwater fluxes as

\[
\tau_x = \rho c_d u (U - U_s) \tag{1}
\]

\[
\tau_y = \rho c_d v (U - U_s) \tag{2}
\]

\[
LH = \rho L_v c_a (U - U_s) (q_s - q_a) \tag{3}
\]

\[
SH = \rho c_p c_h (U - U_s) (T_s - T_a) \tag{4}
\]

\[
E = LH / \rho_w L_v \tag{5}
\]

where \(\tau_x\) and \(\tau_y\) are the respective zonal and meridional wind stress components, \(LH\) the latent
heat flux, \(SH\) the sensible heat flux, and \(E\) the moisture flux. The input variables for calculating
the fluxes (1)-(5) are the zonal \((u)\), meridional wind \((v)\) components, and wind speed \((U)\) at a
reference height, the ocean-surface current velocity \((U_s)\) that is usually small, sea-surface
temperature \((SST, T_s)\), the potential air temperature \((T_a)\) and specific humidity \((q_a)\) at a reference
height, and saturation specific humidity ($q_s$) as a function of $T_s$ and sea level pressure. The other constants are $\rho$ the air density, $\rho_w$ the sea-water density, $L_v$ the latent heat of vaporization that is expressed as $L_v = (2.501 - 0.00237 \times T_s) \times 1.0^6$, and $c_p$ the isobaric specific heat. The turbulent transfer coefficients, $c_{d}$, $c_{e}$, and $c_{h}$, depend on wind speed, atmospheric stability, measurement height, surface roughness, surface wave height, and wave age (e.g., Charnock 1955; Drennan et al. 2003; Andreas et al., 2008; Edson et al. 2013). Bulk flux algorithms differ from each other mainly in how roughness length is parameterized under various wind speeds. Significant uncertainties in these coefficients still remain (Zeng et al. 1998; Brunke et al. 2002), particularly under very weak wind ($U < 4 \text{ m s}^{-1}$) (e.g. Chang and Grossman 2007) or storm force ($U > 24 \text{ m s}^{-1}$) conditions (e.g. Powell et al. 2003; Andreas et al. 2008).

Air–sea exchange at the ocean surface comes not only in the form of turbulent fluxes by evaporation (LH) and conduction (SH) but also by means of radiative fluxes by shortwave and longwave radiation. Evaporation releases not only latent heat but also water vapor (see Eqs. (3) and (5)). Because of the large amount of latent heat exchange during phase change to liquid water (approximately $2.5 \times 10^6 \text{ J kg}^{-1}$ if SST effect is small), the transport of water vapor is regarded as the energy transport. Therefore, the water cycle is closely linked to the energy cycle, with the atmospheric circulation acting as the linchpin connecting the atmosphere and the ocean. The energy (hereafter denoted by $Q_{net}$) and freshwater (hereafter $FW$) budgets over the global ocean surface are expressed as:

$$Q_{net} = SW - LW - LH - SH \quad (6)$$

$$FW = P - E + R \quad (7)$$

where $SW$ is the net downward shortwave radiation, $LW$ the net upward longwave radiation, $P$ the precipitation, and $R$ the river runoff. Energy and water budgets are conserved quantities, and
so $Q_{net}$ and $FW$ must be close to zero when integrated over the global ocean on annual and long-term mean basis. However, all parameterization-based flux products, constructed from either ship reports or satellite observations, do not include the ice-covered Polar Regions due to the lack of reliable observations. In this regard, the globally averaged mean represents a mean over the global ice-free open ocean rather than the entire global ocean, and so, the long-term mean average of $Q_{net}$ should not be closed exactly zero but within 2 – 3 Wm$^{-2}$ (Serreze et al. 2007; Bengtsson et al. 2013).

The ability to close the energy and freshwater budgets at the ocean surface has become a test of the accuracy of gridded flux products (Isemer et al 1989; Josey et al 1999; Fairall et al. 2010; Gulev et al. 2010; Yu et al. 2013; Von Schuckmann et al. 2016; Liu et al. 2017; Valdivieso et al. 2017). This review is to provide an integrative view of leading issues that challenge the parameterization-based flux products in achieving the energy and freshwater budget closures.

2. Energy and Freshwater Budget Closures and Leading Issues

2.1 Leading issues

Flux products are known to have large uncertainties that stem from both the uncertainties in flux-related variables ($u$, $v$, $U$, $q_a$, $T_a$, and $T_s$) and the uncertainties in estimates of transfer coefficients ($c_d$, $c_e$, and $c_h$) in the bulk flux algorithms (Isemer et al. 1989; Josey et al. 1999; Brunk et al. 2003; Valdivieso et al. 2017). Satellite observations represent major improvements over VOS observations owing to their unprecedented sampling frequencies, spatial resolution, and truly global coverage. Nonetheless, space-borne sensors cannot resolve the thermal quantities, $T_a$ and $q_a$, at a few meters above the surface, because the measured radiation is emitted from relatively thick atmospheric layers rather than from single levels (Simonot and
Gautier, 1989; Schulz et al. 1993). A common approach is to retrieve $T_a$ and $q_a$ from satellite observed total column-integrated water vapor using in situ measurements as reference (Liu 1988; Schlüssel et al. 1995), but the empirically-based retrieval algorithm may overly simplify the dependence of the vertical distribution of water vapor content on atmospheric stability and the advection of the large-scale circulation (Esbensen et al. 1993). There are substantial biases in $T_a$ and $q_a$ retrievals that are regime dependent (Yu and Jin 2018), and these biases have been the leading source of error for satellite-based flux products (Curry et al. 2004; Jackson et al. 2006; Prytherch et al. 2015).

The accuracy requirement for $Q_{net}$ is $10 \, \text{Wm}^{-2}$ for flux application on monthly-to-seasonal timescales (WCRP 1989; Webster and Lukas 1992; WGASF, 2000; Weller et al. 2004; Bradley and Fairall, 2007). If the goal is to detect long-term trends from a background of natural variability, the accuracy requirement is at least one order of magnitude higher, at $O(1 \, \text{W m}^{-2})$ for $Q_{net}$ and $O(1 \, \text{cm yr}^{-1})$ for $FW$ (Hansen et al. 2005; Levitus et al. 2005). Parameterization-based flux products all have difficulty closing the ocean heat budget within the above limits. Ship-based climatological analyses show mean heat gains of $\sim 30 \, \text{W m}^{-2}$ or greater by the ocean (Isemer et al. 1989; Large et al. 1997; Josey et al. 1999), and satellite-based products have a similar degree of imbalance (Liu et al. 2017). Some assumed that the imbalance is caused by errors in various flux formulae, which can be corrected by proportional adjustment of the flux components (Isemer et al. 1989; da Silva et al. 1995; Large and Yeager 2009), while some suggested that the significant source of error may come from various regional biases in flux-related variables. These biases may arise from the undersampling of extreme conditions in regions such as the high latitudes and the western boundary currents (Josey et al. 1999), uncorrected biases in $T_a$ and $q_a$ (Jin et al. 2015), etc. Hence, the unbalanced flux products are...
often adjusted by using inverse analysis (Isemer et al. 1989) with hydrographic heat transport constraints to close the global-ocean energy budget (Grist and Josey 2003). More recently, attempts are made to determine an unbiased $Q_{\text{net}}$ from combining satellite-based net radiation at the top of the atmosphere (Rad$_{\text{TOA}}$) and the divergence of vertically integrated horizontal atmospheric energy transports, using the global mean Rad$_{\text{TOA}}$ from the Clouds and the Earth’s Radiant Energy System–Energy Balanced and Filled product (CERES-EBAF; Loeb et al. 2009) that is anchored to estimates of global mean ocean heat storage.

Despite much progress since the work by Isemer et al. (1989) and Josey et al. (1999), the inability to close the ocean heat budget remains a common problem in present parameterization-based products that are largely constructed from satellite observations. Among a number of fundamental issues that are yet to be answered, the following three are most critical. First, all flux products have been rested on the assumption that a good comparison with high-quality independent measurements from air-sea buoys warrants accuracy and reliability. Then, why can’t the energy budget be closed even though flux products are in good agreement with buoy measurements? Second, there seems to be a consensus that the primary source of the energy budget imbalance is the underestimation of $LH$ by about 15%, using the inverse flux adjustment analysis (Isemer et al. 1989; Grist and Josey 2003) and the vertically integrated energy budget adjustment (Liu et al. 2017). Is the underestimation caused solely by biases in flux-related variables (such as $q_a$)? Or does the bulk flux parameterization also play a role? Thirdly, the ocean energy and freshwater budgets are connected through $LH$ (Eqs. (6)-(7)), suggesting that the amount by which $LH$ needs to be adjusted to close the energy budget can potentially be constrained using the ocean freshwater budget. Nowadays the surface radiation product from CERES-EBAF (Kato et al. 2013; Loeb et al. 2018) and the precipitation product from the Global
Precipitation Climatology Project (GCPM) (Adler et al. 2003) have become community standard products. Can they be paired to help diagnose the leading sources of uncertainties in parameterization-based turbulent flux products? The three issues are reviewed below.

2.2 Flux products

Different products use different bulk formulae. Satellite-derived flux products (e.g., Chou et al. 1995; Kubota et al. 2002; Roberts et al. 2010; Andersson et al. 2011; Bentamy et al. 2013; Yu and Jin 2014; 2018) are all established from the COARE bulk flux algorithms (Fairall et al. 1996; 2003; Edson et al. 2013). The ship-based turbulent flux climatology compiled by the National Oceanographic Centre (NOC) (Josey et al., 1999; Berry and Kent 2011) is computed from Smith (1988) algorithm. Atmospheric reanalyses have their own bulk parameterization schemes (Kalnay et al. 1996; Kanamitsu et al. Saha et al. 2010; Dee et al, 2011; Rienecker et al. 2011; Kobayashi et al. 2015; Molod et al., 2015). Surface flux products differ from each other because input data sources (satellite, VOS reports, and NWP models) have uncertainties arising from at least one of the deficiencies: incomplete global coverage, indirect satellite retrievals, systematic bias, and random error. Surface flux products are also sensitive to the choice of algorithms (e.g., Webster and Lukas 1992; Miller et al. 1992; Zeng et al. 1998; Brunke et al. 2003).

The Objectively Analyzed air–sea Fluxes (OAFlux) project at the Woods Hole Oceanographic Institution (WHOI) has been through two phases of flux product development. The first phase led to a 1°-gridded turbulent heat and moisture (i.e. \( L_H \), \( S_H \), and \( E \)) flux analysis (hereafter OAFlux-1x1), with \( q_a \) and \( T_a \) determined from objective synthesis of satellite-derived retrievals and atmospheric reanalyses and \( U \) from multiple satellite sensors (Yu and Weller 2007;
Yu et al. (2008). The second phase of development has focused on constructing high-resolution (HR; 0.25°-gridded), full-range (i.e., \( LH, SH, E, \tau_x \) and \( \tau_y \)) turbulent flux products (hereafter OAFX-HR), with flux-related variables determined solely from satellite retrievals (Jin and Yu 2013; Yu and Jin 2014; 2018). Compared to OAFX current 1°-gridded analysis (hereafter OAFX-1x1; Yu and Weller 2007; Yu et al. 2008), OAFX-HR has made improvements in three main aspects: spatial resolution, \( q_a \) and \( T_a \) estimates, and the inclusion of momentum fluxes. The improvement leads to an increase of \( LH+SH \) by ~ 8 Wm\(^{-2}\), but disappointingly, it does not lead to an energy budget closure. When combined with CERES EBAF surface radiation (SW-LW), OAFX-1x1 \( LH+SH \) produces a mean heat gain of ~ 25 Wm\(^{-2}\) over the global ocean while OAFX-HR \( LH+SH \) has a gain of ~ 17 Wm\(^{-2}\). Since CERES EBAF has been adjusted to balance the Earth’s energy budget, the imbalance is once again pointed to as-yet uncorrected bias in OAFX-HR. From the viewpoint of the flux variable estimation, the argument is not convincing. The OAFX-HR satellite-derived variables, \( q_a, T_a \) and \( U \), have thoroughly validated with in situ time series measurements at more than 120 locations. The mean biases relative to buoy measurements are – 0.34 g kg\(^{-1}\) for \( q_a \) (i.e. a dry bias), – 0.08°C for \( T_a \) (i.e., a slight cold bias), and – 0.13 m s\(^{-1}\) for \( U \) (i.e., a weak bias) (Yu and Jin 2012; 2018). A simple error diagnosis of the bulk formula for \( LH \) and \( SH \), assuming a mean wind speed of 7 m s\(^{-1}\), suggests that the adjustment of 17 W m\(^{-2}\) imbalance requires the mean state of the near-surface air to be either further dried up by 0.74 g kg\(^{-1}\) or cooled down by 0.46°C. The magnitude of adjustment is way beyond the product accuracy defined by buoy evaluation.

Uncertainty in bulk flux algorithm is the only stone left unturned in our pursuit of surface energy budget closure. When comparing the two versions of OAFX products with atmospheric reanalyses, the influence of bulk algorithms on surface flux estimates is evident. Hence, there is a
need for understanding the uncertainties in both flux-related variables and bulk algorithms to
gain a complete understanding of the cause of surface budget imbalance. Since satellite-derived
products are all produced from COARE version 3 (v3), differences between products reflect the
differences between variable estimation which have been characterized by several comparison
studies (Betamy et al. 2017). To narrow down the scope of this review, we limit the discussion to
9 atmospheric reanalyses, 2 OAFlux products, and the ship-based NOC, and use CERES and
GPCP as budget constraints (Table 1).

The OAFlux-HR full-range turbulent flux products can be combined with CERES and
GPCP to provide a complete description of ocean-surface heat, freshwater, and momentum
fluxes. The annual-mean fields of $Q_{net}$ from CERES and OAFlux-HR, $E-P$ from OAFlux-HR and
GPCP, and wind stress vector and wind stress curl (i.e. $\partial \tau_y/\partial x - \partial \tau_x/\partial y$) from OAFlux-HR in 2014
(Figure 1). Consistent with the climatological mean patterns (e.g. Josey et al. 2013), the tropical
ocean is the primary region of atmospheric heat and freshwater input to the ocean and the
subtropical ocean, particularly the western boundary current (WBC) regime, is the region of
oceanic heat and freshwater transfer to the atmosphere. In the Northern Hemisphere, cyclonic
(positive) wind stress curl drives an upward Ekman pumping and upwelling, while anticyclonic
(negative) wind stress curl drives Ekman suction and downwelling. In the Southern Hemisphere,
the effects are opposite with cyclonic (positive) wind stress curl denoting downwelling and
anticyclonic (negative) wind stress curl upwelling. Although CERES SW and LW are 1° gridded
and GPCP precipitation 2.5° gridded, the high-resolution advantage of OAFlux-HR in depicting
the fine structure of frontal-scale air-sea exchanges is seen in the WBC regimes.

2.3 Differences in bulk parameterization algorithms
Unlike OAFlux-1x1 that is constructed from COARE v3, the OAFlux-HR flux fields in Figure 1 are computed from an updated COARE bulk flux algorithm, version 4 (Edson et al. 2010; 2012; 2013). COARE v4 (Jim Edson, personal communication) has focused on improving turbulent transfer coefficients, particularly, $c_e$ and $c_h$ for LH and SH. In COARE v3, the coefficients for LH and SH are identical, assuming similarity in the transfer of heat and mass. In COARE v4, LH and SH are modeled with separate formulae and validated with direct flux measurements from field programs. The $c_e$ estimate in the two algorithms exhibits the same overall characteristics of a minimum around wind speed at 3 – 4 ms$^{-1}$; after that, $c_e$ in COARE v4 increases to a maximum around wind speed at 12 ms$^{-1}$ before falling off at higher winds, while $c_e$ in COARE v3 shows a near-linear increase with wind speed. In the following, OAFlux-HR computed from COARE v3 is denoted OAFlux-HR3 and that from COARE v4 is OAFlux-HR4.

The zonal averages of the annual-mean LH+SH fields in 2014 from OAFlux-HR3, -HR4, and OAFlux-1x1 and their differences (Figures 2a-b) show that the three products differ most at low and mid latitudes. The differences between OAFlux-HR3 and -HR4 reflect the change induced by COARE algorithms, and v4 produces stronger LH+SH at all latitudes with maximum differences of ~20 Wm$^{-2}$ at 30–40° latitudes north and south. The latter are the locations of strong turbulent heat loss associated with WBCs. The differences between OAFlux-1x1 and -HR3 reflect the change made in flux variables due to resolution change and the use of satellite-only input data source, and the improvement leads to an averaged increase of ~10 Wm$^{-2}$ for the latitudes between 40°S and 40°N. In general, COARE v3 is a weaker algorithm compared to v4.

To assess the difference between COARE v3 and the bulk flux algorithms in reanalyses, the flux-related variables from NCEP1, CFSR, ERA-interim, and MERRA were used as input to COARE v3 to compute a set of COARE v3-based reanalysis fluxes. The zonally averaged mean
differences between the original reanalysis fluxes and the COARE v3-based reanalysis fluxes in 2014 (Figures 2c-d) indicate that COARE v3 is a weak algorithm compared to the four reanalysis algorithms. The ERA-interim algorithm is the closest to COARE v3, and the differences are mostly within 5 Wm$^{-2}$ except for a 10 Wm$^{-2}$ spike at ~15°N/S. The NCEP1 algorithm has the largest departure from COARE v3, with magnitude approaching 40 Wm$^{-2}$ at subtropical latitudes. CFSR and MERRA algorithms are respectively about 8 and 12 Wm$^{-2}$ stronger at most latitudes.

2.4 Interpretation of buoy evaluation

Time series measurements from moored air-sea buoys in the global ocean serve as benchmarks for validating flux products constructed from various sources (Fairall et al. 2010; Gulev et al. 2010; Yu et al. 2013; Bentamy et al. 2017; Valdivieso et al. 2017). Despite good comparisons, none of flux products is yet able to achieve an energy budget closure if additional adjustments are not imposed (e.g. Isemer 1989; Josey et al. 1999). Two factors might be responsible for this. One is that buoy fluxes are not measured but computed (Weller et al. 2008), and the algorithm for buoy LH+SH is COARE v3. The computed buoy fluxes may not be bias free if there is uncertainty in the flux algorithm (Figure 2). The other is that the majority of buoys are deployed in the tropical warm water zone with very limited number of buoys in the vicinity of WBCs and high-latitude cold water zone (Figure 3a).

To illustrate that COARE v3-based buoy fluxes may not be a viable verification for flux products, we computed daily-mean buoy fluxes (in terms of SW-LW and LH+SH) that were acquired between 1990 and 2015 at 126 buoy locations (Figure 3a) and compared with collocated daily-mean CERES SW-LW, OAFlux-1x1 LH+SH and 6 atmospheric reanalyses (SW-LW and LH+SH). Since surface fluxes are a sensitive function of SST, we binned the
product-minus-buoy flux differences onto every 0.5°C SST grids using buoy observations. Distribution of product-minus-buoy differences with SST (Figures 3b-c) indicates that there are some exceptionally large values in a few SST regimes: low SSTs (<6°C), SSTs of 15-20°C, and very high SSTs (>30°C). The number of available buoy measurements is limited (less than 50) in these SST ranges so that the performance of flux products may not be statistically well represented. Away from these ranges, the errors in reanalysis SW-LW increase sharply for SST greater than 20°C, which corresponds to the tropical-subtropical warm water regime. Only satellite-derived CERES SW-LW is unbiased. As for error distribution in LH+SH, all reanalyses have a similar error distribution pattern: errors are smaller when SST is less than 15°C and larger when SST is greater than 20°C. Except for NCEP1 and MERRA, the errors remain more or less constant for SST between 20 – 28°C though with varying magnitude. JRA55 differs by more than 40 Wm\(^{-2}\), ERA-interim by 20 Wm\(^{-2}\), and CFSR by 17 Wm\(^{-2}\). OAFlux-1x1 is largely unbiased - but it is computed from COARE v3, the same algorithm used by buoy fluxes. Given COARE v3 is weak in comparison with reanalysis algorithms (Figures 2c-d), it is yet to be determined which is a more dominant source of uncertainty for LH+SH products, the bulk algorithm or the flux-related variables.

2.5 Differences in long-term mean fields

The standard deviations (STD) between 12 mean \(Q_{net}\) products (Table 1) averaged over the overlapping 10-year period of 2000-2010 conveys the same message that surface heat flux estimates are most uncertain in the tropical and subtropical region (Figure 4a). In the Indo-Pacific warm pool, for instance, the STD differences between products exceed 30 Wm\(^{-2}\), which are greater than the ensemble mean of the products. Zonal averages of the 10-year mean \(Q_{net}\),
SW-LW, and LH+SH (Figures 4b-d) indicate that JRA55 $Q_{net}$ is an outliner, as its LH+SH is excessively strong in the tropics. OAFlux-HR4 is in the same range as reanalysis LH+SH between 25-45°N/S, but is stronger than the reanalysis at mid latitudes because the high resolution of HR4 can better resolve the LH+SH associated with the WBC fronts.

The STD differences between 11 mean $E-P$ products averaged over 2001-2010 (Figures 5a-c) are most pronounced in the tropical/subtropical regions between 30°S and 30°N. Major uncertainty is the spread in $P$ products in regions of the Intertropical Convergence Zone (ITCZ) and South Pacific Convergence Zone (SPCZ), with the satellite-based GPCP having the weakest rainfall and JRA55 the strongest rainfall. The pattern of differences suggests that reanalyses have difficulty in simulating tropical convective clouds and rainfall processes (Rosenfeld and Lensky 1998; Newman et al. 2000; Yu et al. 2017). In contrast to the STD $E-P$ pattern, the STD differences between 11 mean wind stress magnitude, $\tau$, products averaged over 2001-2010 show that large deviations are located at mid to high latitudes where winds are strong (Figure 6a). The zonal averages reveal that the spread in the products is caused primarily by the gaps between two groups, the group that assimilates satellite scatterometers (i.e. CFSR, ERA-interim, JRA55, MERRA, and MERR2, OAFlux-HR) and the group that does not (NCEP2, ERA20C, and 20CR). Winds are vectors, which explain that the zonal averages of $\tau_x$ and $\tau_y$ are not proportional to the zonal average of $\tau$ due to the sign cancellation (Figures 6b-c).

2.6 Surface budget imbalance: are CERES and GPCP compatible constraints?

Given the large uncertainties in surface flux estimates in the tropical-subtropical ocean, it is not a surprise that the surface energy and freshwater budgets determined by the mean $Q_{net}$ and $E-P$ products differ considerably between them. (Figures 7a-b). The surface energy budget
ranges from a significant ocean heat deficit of $-16 \text{ W m}^{-2}$ by JRA55 to a significant ocean heat gain of $25 \text{ W m}^{-2}$ by OAFlux-1x1. The surface freshwater budget ranges from a nearly perfect balance between $E$ and $P$ by CFSR to a large freshwater imbalance of $27 \text{ cm yr}^{-1}$ by the combined OAFlux-HR4 and GPCP. Interestingly, the product series of OAFlux affects the surface energy and freshwater budget balance in an opposite way. While the imbalance in the energy budget is reduced by the order from OAFlux-1x1, to HR3, and to HR4, the imbalance in the freshwater budget is increased in the same order.

The scatter plots between $SW-LW$ and $LH+SH$ and between $E$ and $P$ (Figures 7c-d) shed some light on how the ocean and atmospheric flux components could be partitioned to achieve balanced budgets. As stated in the Introduction, the energy budget determined from surface heat flux products is expected to achieve a closure within $2-3 \text{ W m}^{-2}$ due to the exclusion of Polar regions. This implies that $SW-LW$ should be balanced with $LH+SH$ within the limit. Surface radiative budgets from CERES, NOC, ERA-interim, and CFSR agree well with each other, but the deviations in LH+SH set the total budgets apart. OAFlux-HR4 is a far better match for CERES compared to OAFlux-1x1. However, this works opposite for the freshwater budget. On the long-term mean basis, the ocean freshwater budget should be balanced, that is $E - P - R \approx 0$ (Eq.(7)). If expressing the $E/P$ in terms of $E/P \approx (P+R)/P = 1 + R/P$, one can expect that the larger (smaller) the ratio, the more (less) continental runoff is needed to balance the water budget over the ocean. The $E/P$ ratio is found to be about 1.1 in most reanalysis products (Yu et al. 2017). OAFlux-1x1 and GPCP fall exactly on the line that delineates $E/P = 1.1$ (Figure 7d), while HR4 is significantly off.

CERES EBAF and GPCP are community standard products. The improvement made in OAFlux-HR4 improves the surface energy budget constrained by CERES but deteriorates the
surface freshwater budget constrained by GPCP. Looking at the scatter relationships (Figures 7c-d), CERES surface heat input is on the higher end and consistent with three reanalyses. By comparison, GPCP freshwater input is lowest among all reanalyses. Are GPCP and CERES pairwise compatible in terms of surface energy and freshwater budgets?

3. FUTURE PERSPECTIVES

This review presents a perspective on the imbalance in surface energy and freshwater budgets using parameterization-based flux products. Most viewpoints stem from our own decade-long research developing surface turbulent heart, moisture, and momentum fluxes from satellite observations. The inability to close the surface energy budget, despite many efforts that have been made to improve the estimates of flux-related variables, has led us to reframe our thinking and embrace questions that are still largely unanswered. Achieving globally balanced energy and freshwater budgets is a multifaceted challenge, and this review has focused on only three questions. Nonetheless, our study stresses the importance of collaborations between various groups to understand and resolve a number of discrepancies in the present-day turbulent flux estimates. These include the differences between COARE algorithms and bulk flux parameterizations used in atmospheric reanalyses, differences in flux-related variables from VOS reports, satellite observations, and atmospheric reanalysis outputs, and differences between surface radiation and precipitation products in the context of surface energy and water cycles.

Efforts addressing the following three aspects are particularly relevant. First, in situ air–sea measurements of fluxes-related variables, though limited in space, are indispensable for establishing benchmark accuracy for gridded flux products, as well as for maintaining long-term
stability. To advance the skills of bulk flux parameterizations, more direct flux measurements are needed. Second, cross-comparisons among hierarchy products from varied sources are useful in identifying and understanding the uncertainties in flux products. In this regard, atmospheric reanalyses are excellent tool for such study. Lastly, limitations in current in situ air–sea observing capability suggest the need to include ocean observations and ocean data-model syntheses to achieve greater consistency by balancing the regional and/or global energy and freshwater budgets.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGEMENTS

This work was supported by funding from NOAA Ocean Observing and Monitoring Division, NOAA Climate Reanalysis Task Force Team activity, and NASA Ocean Vector Wind Science Team activity. X. Jin provided programming help. J. Edson is sincerely thanked for providing the COARE programming codes.

LITERATURE CITED


Surface warming hiatus caused by increased heat uptake across multiple ocean basins,

Edson, J.B., A.A. Hinton, K.E. Prada, J.E. Hare, and C.W. Fairall, 1998: Direct Covariance Flux

L. Mahrt, D. Vickers, and H. Hersbach, 2013: On the Exchange of Momentum over the Open

Sensor Microwave Imager evaporation estimates over the global oceans. J. Geophys. Res., 98,
7081–7101

Fairall, C. W. and S. E. Larsen, 1986: Inertial Dissipation Methods and Turbulent Fluxes at the
Air Ocean Interface, Boundary-Layer Meteorol. 34, 287-301.


Fairall, C. et al. 2010. Observations to quantify air-sea fluxes and their role in climate variability
and predictability, Plenary White Paper in Proceedings of the "OceanObs'09: Sustained Ocean
Observations and Information for Society" Conference. ESA Publication WPP-306, Venice,
Italy, 21-25 Sept 2009


https://doi.org/10.1175/2011JCLI4211.1


Working Group on in situ measurements for Fluxes WCRP-23 (WMO/ TD No.304), WMO, Geneva.


Figure 1. Annual-mean (a) Qnet from CERES+OAFlux-HR4, (b) E-P from OAFlux-HR4 and GPCP, and (b) wind stress vector and wind stress curl (colors) in 2014.
Figure 2. Zonally averaged mean LH+SH in 2014. (a) 3 OAFlux products. (b) Differences between OAFlux-HR4 and HR3 versus between OAFlux-HR3 and 1x1. (c) Original NWP fluxes (thick lines) and recomputed fluxes using NWP variables and COARE v3 algorithm (thin lines). (d) Differences between NWP fluxes and COARE v3 based fluxes.
Figure 3. (a) Mean SST field in 2014 superimposed with locations of 126 buoys. (b) Distribution of product-minus-buoy differences in SW-LW with SST. (c) Distribution of product-minus-buoy differences in LH+SH with SST. (d) Number of buoy-product collocation pairs for daily-mean SW-LW. (e) Number of buoy-product collocation pairs for daily-mean LH+SH.
Figure 4. (a) Standard deviations between 12 mean Qnet products. The mean fields are constructed over the 10-year period between 2001 and 2010. Zonal averages of (b) Qnet, (c) SW-LW, and (d) LH+SH.
Figure 5. (a) Standard deviations between 11 mean E-P products. The mean fields are constructed over the 10-year period between 2001 and 2010. Zonal averages of (b) E-P, (c) E, and (d) P.
Figure 6. (a) Standard deviations between 10 mean wind stress magnitude products. The mean fields are constructed over the 10-year period between 2001 and 2010. Zonal averages of (b) wind stress magnitude, (c) zonal wind stress, and (d) meridional wind stress.
Figure 7. (a) Global-ocean mean energy (Qnet) budget. (b) Global-ocean mean freshwater (E-P) budget. (c) The ratio of mean average of SW-LW to LH+SH. (d) The ratio of mean average of $E$ to $P$. The black dash line denotes that the $E/P$ ratio equals to 1.10.