INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2010)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2225

A comparison of nine monthly air-sea flux products

Shawn R. Smith,[a]* Paul J. Hughes[a,b] and Mark A. Bourassa[a,b]
[a] *Center for Ocean-Atmospheric Prediction Studies, The Florida State University, Tallahassee, FL, USA*
[b] *Department of Meteorology, The Florida State University, Tallahassee, FL, USA*

ABSTRACT: A comparison is conducted between nine monthly turbulent air–sea flux products. The analysis includes *in situ*-based (Florida State University fluxes, FSU3 and National Oceanography Centre, NOC), satellite-based (Hamburg Ocean-Atmosphere Parameters from Satellite data, HOAPS2 and the French Research Institute for Exploitation of the Sea, IFREMER), hybrid (Goddard Satellite-based Surface Turbulent Fluxes, GSSTF2 and objectively analysed fluxes, OAFLUX), and reanalysis (National Centers for Environmental Prediction, NCEPR2, Japanese 25-year reanalysis, JRA, and European Centre for Medium Range Weather Forecasts reanalysis, ERA-40) products. Objectives include documenting the varying analysis methodologies and quantifying the differences and similarities between the nine products. Recommendations are made for developers of future flux products and to guide users to select products most suitable for their application.

The comparison examines turbulent fluxes of heat and momentum along with the forcing variables (air temperature, wind speed, humidity, and ocean skin temperature) that are necessary to estimate turbulent fluxes. The wide range of turbulent flux parameterisations, sampling patterns, and averaging techniques within the products are described, including some of the difficulties product differences pose when trying to compare or apply the individual products. Global comparisons of monthly means tend to reveal similar spatial patterns in latent heat flux (LHF) and sensible heat flux (SHF) for the nine products; however, the magnitudes and patterns of variability (expressed as maps of standard deviations) are widely different. Basin scale and regional analysis further reveals large differences in the products (in some cases the interquartile ranges (IQRs) do not overlap for different products), but also reveals potential sources of the differences. For example, some of the variations in LHF can be explained by large differences in the distribution of specific humidity between the products. As a final analysis, we examine how each product represents the variations in turbulent fluxes in the equatorial Pacific (EP) Ocean. This analysis provides an example of how the choice of a flux product, and understanding the strengths and weaknesses of that product, can alter research findings. Copyright © 2010 Royal Meteorological Society

KEY WORDS air–sea interaction; surface fluxes; turbulent fluxes; *in situ* fluxes; satellite fluxes; reanalyses

Received 29 June 2009; Revised 9 April 2010; Accepted 15 August 2010

1. Introduction

Nine readily available monthly air–sea flux products are compared on global, basin-wide, and regional spatial scales. The comparison is motivated by the continued need for high-quality estimates of the air–sea fluxes to support interannual to decadal climate studies, assess global and regional heat budgets, and provide forcing fields for both ocean and atmospheric models. Both operational and research users continue to seek the 'best' estimate of air–sea fluxes for their applications. At present, many users are working with older, outdated flux products that have been shown to have severe limitations (Bony *et al*., 1997; Putman *et al*., 2000; Smith *et al*., 2001; Chou *et al*., 2003; Kubota *et al*., 2003). In fact, several of the flux products considered in this manuscript have been updated since this work was completed; however, the community will continue to use the nine analysed products so that an assessment of their characteristics and limitations is valuable. We show by example that research results can be affected by the user's choice of a flux product.

One objective of our comparison is to quantify the magnitude of the differences and note similarities between the nine products. This cannot be achieved without a full understanding of the analysis methods used to create each product (refer Section 2). At present, there is no direct global standard of comparison for monthly air–sea flux estimates. Some point comparisons have been made to mooring observations (Bourras, 2006), but these are limited in both time and space (e.g. very few moorings in high latitude locations), and typically use bulk algorithms to determine fluxes rather than comparing to direct turbulent flux measurements (Yelland *et al*., 2009). Since a 'true' flux value cannot be identified, the authors instead look at the individual surface turbulent flux (stress, sensible heat, and latent heat) products to determine where agreements or disagreements occur not only in the fluxes but also in the forcing variables used to derive those fluxes. A second objective is to identify which forcing variable might be responsible for the differences noted in the nine flux products. Our third objective is to provide those involved

* Correspondence to: Shawn R. Smith, COAPS, The Florida State University, Tallahassee, FL, 32306-2840, USA.
E-mail: smith@coaps.fsu.edu

with the development of flux products with a way forward, whereby we will identify where agreement exists among these community products and note where additional research effort is needed. Finally, we will comment on the need of users to know what flux product is 'best' and will show that the 'best' varies with the objective of the user's application.

Air–sea fluxes are by definition rates of exchange, per unit surface area, between the ocean and the atmosphere. The stress ($\boldsymbol{\tau}$) is the flux of horizontal momentum (imparted by the wind on the ocean). The evaporative moisture flux would be the rate, per unit area, at which moisture is transferred from the ocean to the air. The latent heat flux (LHF) is related to the moisture flux: it is the rate (per unit area) at which energy associated with the phase change of water is transferred from the ocean to the atmosphere. Similarly, the sensible heat flux (SHF) is the rate at which thermal energy (associated with heating, but without a phase change) is transferred from the ocean to the atmosphere. In the Tropics, the LHF is typically an order of magnitude greater than the SHF; however, in polar oceans the SHF can dominate. Flux products are typically derived using the bulk formulae

$$\boldsymbol{\tau} = \rho C_{\mathrm{D}} \boldsymbol{\Psi} \tag{1}$$

$$\mathrm{SHF} = \rho c_{\mathrm{p}} C_{\mathrm{H}} \left(T_{\mathrm{skin}} - T_{\mathrm{air}}\right) w \tag{2}$$

$$\mathrm{LHF} = \rho L_{\mathrm{v}} C_{\mathrm{E}} \left(q_{\mathrm{sfc}} - q_{\mathrm{air}}\right) w \tag{3}$$

where ρ is the density of moist air, c_{p} is the specific heat capacity of air at constant pressure, L_{v} is the latent heat of vaporisation, C_{D} is the drag coefficient, C_{H} is the heat transfer coefficient, and C_{E} is the moisture transfer coefficient. The variables w, T_{air}, and q_{air} are the scalar wind speed, temperature, and specific humidity of the air at a reference height above the ocean surface. The variables T_{skin} and q_{sfc} are the sea surface temperature (SST) and the surface saturation-specific humidity corresponding to the T_{skin}. In practice, T_{skin} is usually approximated by the bulk or foundation temperature, and models are often tuned for this approximation rather than the actual skin temperature. The variable $\boldsymbol{\Psi}$ is the vector pseudostress. The zonal ($\boldsymbol{\Psi}_x$) and meridional ($\boldsymbol{\Psi}_y$) components of the pseudostress are defined as

$$\boldsymbol{\Psi}_x = uw \tag{4}$$

$$\boldsymbol{\Psi}_y = vw \tag{5}$$

where u and v represent the zonal and meridional components of the wind vector.

Air–sea flux products are derived through a variety of methods and using input data with different spatial and temporal scales. Herein we compare fluxes estimated by numerical weather prediction (NWP) reanalysis models to those derived only from *in situ* or satellite observations, and to hybrid products combining one or more of these data sources. On a monthly scale, the reanalysis products (Kalnay *et al*., 1996; Kanamitsu *et al*., 2002; Uppala *et al*., 2005; Onogi *et al*., 2007) have the advantage of longer time series than satellite-derived fields and the addition of upper-air fields for 3D analysis. However, early reanalysis products (Kalnay *et al*., 1996) are noted to have a poor handling of the wind field in equatorial regions (Putman *et al*., 2000), as well as large biases in heat fluxes (Bony *et al*., 1997; Smith *et al*., 2001). The flux parameterisations used in these reanalyses are not consistent with state-of-the-art flux models (Brunke *et al*., 2002; Curry *et al*., 2004). Satellite-derived products (Grassl *et al*., 2000; Pegion *et al*., 2000) are temporally limited; thus, monthly satellite products are of limited use for studies of phenomena with long (i.e. decadal) periods. Satellite products are more likely to have similar characteristics over time; however, statistics do change with changes in the satellite observing system. For applications that require long-term monthly surface fluxes (and related variables), research quality climatologies based on *in situ* observations are available (Da Silva *et al*., 1994; Josey *et al*., 1998; Bourassa *et al*., 2005). *In situ* products are limited by spatial and temporal sampling issues associated with irregular ship and buoy observations and by inherent observational biases (Gulev *et al*., 2007a, 2007b; Risien and Chelton, 2008; Thomas *et al*., 2008). Hybrid products (Chou *et al*., 2003; Yu and Weller, 2007) try to combine the strengths of several *in situ* or satellite data sets and NWP analyses.

Several recent comparisons have focused on the LHF (Chou *et al*., 2003; Kubota *et al*., 2003; Bourras, 2006). Kubota *et al*. (2003) and Chou *et al*. (2003) both take the approach of differencing the product developed by their group with one or more products from other researchers. This approach tends to favor the product developed 'in-house' as the benchmark, making conclusions relative to that product. By comparison, Bourras (2006) evaluated five satellite-derived flux products using point data from buoys. In some ways, this approach seems a more fair comparison; however, in several cases, the satellite products are tuned to or in other ways are affected by the mooring data during their production (a point acknowledged by Bourras, 2006). This does raise some questions about the independence of the comparison. In addition, the spatial and temporal representativeness of point measurements complicates their comparison to satellite data. Overall, Bourras (2006) found that different satellite products had good fits to individual buoy data sets [Bourras–Eymard–Liu fluxes are well matched to Tropical Atmosphere Ocean (TAO) moorings]; however, strong variations existed when a single satellite product was compared to moorings in different ocean regions (e.g. tropical *vs* mid-latitude moorings). The differences were found to be a function of the vertical distribution of atmospheric moisture, a quantity that is still difficult to measure from space.

Throughout our comparisons, no one product is singled out as a standard of comparison and our goal is to draw conclusions beyond simply what is different between the products. To aid flux product developers, we highlight differences in the fluxes and the forcing variables, and emphasise strengths and weaknesses applicable to research activities. For the user community, we provide an example of the impact of using these different flux products to evaluate the interannual variability in the tropical Pacific Ocean. This example highlights how the choice of flux products must be done with care to accurately achieve ones research goals.

2. Data products

Nine monthly air–sea flux products are compared and shown in Table I. The products are divided into four broad categories based on the input data source or analysis method: *in situ*, satellite, reanalysis, or hybrid. The varying data sources and analysis methods result in each product being available for different periods. In addition, several parameters of interest for our comparison were not available for some products (Table I) at the time of the analysis herein. To facilitate the comparisons, several products are interpolated or averaged to a common 1° latitude by 1° longitude grid with each grid box centered at 0.5°. A land mask matching the 1° grid (Figure 1) is developed based on a 1/12° topography data set (NGDC, 1988). The mask excludes any 1° ocean grid determined to be ice-covered based on a quarter degree remote sensing systems special sensor microwave/imager (SSM/I) product (Wentz, 1997) for any month from July 1987 through December 2006. A 1° grid cell is considered to be ice if two or more of the available 0.25 cells (16 total) contained within 1° cell are flagged as ice in the SSM/I product. The mask also expands the land/ice area by two grid cells over the oceans, eliminating differences between products solely due to different treatment of grid cells that are part land or ice and part water. The mask also excludes regional seas (e.g. the Gulf of Mexico and the Mediterranean Sea).

2.1. *In situ*

Two products derived only from *in situ* ship and buoy observations are version 1.1 of the National Oceanography Centre (NOC, previously known as the SOC climatology, Josey *et al*., 1998) and the third version of the Florida State University fluxes (FSU3). Both products are derived from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS; Woodruff *et al*., 1987). The NOC fluxes are based on release 1a of ICOADS; the FSU3 are constructed from release 2.2 of ICOADS. Release 2.2 is a minor update to release 2.1 (Worley *et al*., 2005) and contains additional delayed-mode ship and buoy observations compared to release 1a (Woodruff *et al*., 1993). Additional data in version 2.2 is a source of differences between the NOC and FSU3 products.

The methods to derive the NOC and FSU3 have similarities and notable differences. For both products, a series of bias corrections are applied to the individual marine data before the creation of the flux fields. The bias corrections are not consistent between these products; however, these differences are believed to be small compared to difference due to spatial averaging/interpolation. The algorithm for the NOC heat fluxes and wind stress is a combination of Smith (1980, 1988). A successive correction method is used to develop the monthly NOC flux fields (Josey *et al*., 1998). The flux algorithm used for the FSU3 is based on the method of Bourassa *et al*. (2005). The FSU3 fields are the result of a variational method (direct minimisation) with separate objectively determined weights for ship, mooring, and drifter data (Pegion *et al*., 2000; Bourassa *et al*., 2005). Additional differences include smoothing of fluxes (both products), smoothing of the forcing variables (FSU3), and the calculation of monthly fluxes from either the average of individual flux calculations (NOC) or the monthly averaged input data (FSU3).

In situ-based flux products are limited by the spatial and temporal sampling provided by ships and buoys. The spatial pattern of ship data is concentrated on preferred ship routes, and fixed buoys are limited to the Tropics and coastal regions. The Southern Hemisphere has particularly poor sampling, resulting in the FSU3 developers' decision not to produce fields south of 30 °S. In addition, the sampling network is constantly evolving. Ships are getting larger and observing heights for winds, moisture, and air temperature (T_{air}) are increasing (Thomas *et al*., 2008). The FSU3 adjust ship measurements to a constant estimated ship height that is time invariant, while height adjustments for the NOC use time-varying instrument heights or a constant value when observation heights are not available. Both methods impart artificial trends in the analyses. The comparison herein has helped developers of the NOC and FSU3 to understand the substantial shortcomings related to sampling and to propose corrections.

Several past and recent *in situ* products are not included in this analysis. The monthly flux climatology by Da Silva *et al*., (1994) has been excluded; the product is available only from 1945 to 1993, limiting the temporal overlap with the analysed satellite and hybrid products (Table I). In addition, previous comparisons (Chou *et al*., 2003; Kubota *et al*., 2003) have shown Da Silva *et al*. (1994) to have severe limitations related to sampling patterns (particularly in the southern ocean (SO) where the product has very little variability), and some versions of the product have biases related to assumptions regarding ocean budgets (Da Silva *et al*., 1994). Also excluded is the 2009 update of the NOC-Southampton *in situ* analysis (Berry and Kent, 2009).

2.2. Satellite

Two satellite products, the French Research Institute for Exploitation of the Sea (IFREMER) fluxes and the second version of the Hamburg Ocean-Atmosphere Parameters from Satellite data (HOAPS2), are included in this comparison. According to A. Bentamy (personal communication, 2007), the IFREMER surface turbulent fluxes are created using physical properties of active and passive satellite instrument measurements, empirical and

Table I. The nine flux products compared, including the product type, original spatial grid, and temporal period (month/year) when the monthly fluxes are available.

	NOC	FSU3	IFREMER	HOAPS2	NCEPR2	JRA	ERA-40	OAFLUX	GSSTF2
Type	I	I	S	S	R	R	R	H	H
Original grid	$1^\circ \times 1^\circ$	$1^\circ \times 1^\circ$	$1^\circ \times 1^\circ$	$0.5^\circ \times 0.5^\circ$	Gaussian (T62, 194×94)	(T106 L40)	$1.125^\circ \times 1.125^\circ$	$1^\circ \times 1^\circ$	$1^\circ \times 1^\circ$
Period	January 1980–December 2005	January 1978–December 2004	March 1992–October 2006	July 1987–December 2002	January 1979–December 2005	January 1979–December 2005	September 1957–August 2002	January 1981–December 2002	July 1987–December 2000
LHF	3	3	3	3	3	3	3	3	3
SHF	3	3	3	3	3	3	3	3	3
Stress (x, y)	3	3	3	–	3	3	3	–	3
Wind speed	3	3	3	3	–	–	–	–	3
Wind (u, v)	–	3	3	–	3	3	3	–	–
T_{air}	3	3	3	–	3	3	3	–	3
q_{air}	3	3	3	3	3	3	3	–	–
SST	3	3	3	3	3	3	3	–	–

I, *in situ*; S, Satellite; R, Reanalysis; and H, Hybrid.
[a] Parameters available for present comparison are marked (√) for each product.

inverse models relating satellite observations and surface parameters, and objective analysis merging various satellite estimates. The winds and LHFs are mainly derived using observations from the scatterometers onboard the European remote sensing satellites (ERS-1 and ERS-2), the NASA scatterometer (NSCAT) onboard ADEOS-1, the SeaWinds scatterometer onboard QuikSCAT, and radiometers (SSM/I) onboard the Defense Meteorological Satellite Program (F10, F11, F13, F14, and F15) spacecraft. The methods used to derive the input for the IFREMER turbulent fluxes and the bulk formulation for LHF are described in Bentamy *et al*. (2003). The IFREMER T_{air} is assumed to be equal to SST minus 1.25, where SST is the daily Reynolds optimum interpolation product. Surface humidity is assumed to be 98% of the saturation value, and the atmosphere humidity is determined from a multiple linear regression to SSM/I brightness temperatures similar to the model of Schultz *et al*. (1993, 1997). Information on a heat transfer coefficient and a drag coefficient is not provided. One cautionary note is that scatterometer rain flags were not applied; consequently, stresses are particularly poor in regions with frequent rain (e.g. Intertropical Convergence Zone, ITCZ). The IFREMER method does not account for atmospheric stratification. The quality of the derived surface winds and LHFs was investigated through comprehensive comparisons with buoy and ship estimates (Bentamy *et al*., 2003); however, some caution should be applied to these results because the buoy data were strongly assimilated in the background fields (European Centre for Medium Range Weather Forecasts, ECMWF reanalysis) used in the variational method.

The HOAPS2 product (based on the original HOAPS; Grassl *et al*., 2000) is derived using techniques similar to the IFREMER product. According to C. Klepp (personal communication, 2008), the forcing parameters (wind speed, specific humidity, and T_{skin}) are retrieved from the satellite observations (i.e. brightness temperature, backscatter, etc.). From these parameters, LHF and SHF are calculated using the COARE algorithm (version 3.0)(Fairall *et al*., 1996). Note that HOAPS2 does not provide wind stress or T_{air} values, while these are available from the IFREMER product (Table I). The wind stress is not provided because it is not directly needed for surface energy budgets, and because of the limited sampling and different calibrations of scatterometer data sets (C. Klepp, personal communication, 2008). The T_{air} was not provided because the accuracy was considered insufficient for many applications. The HOAPS2 product used in this study originated on a 0.5° latitude by 0.5° longitude grid and had a monthly time step. A simple four-box average is used to transfer the HOAPS2 to the 1° grid used for the comparison. When one or more of the four points in a 1° box are missing, the 1° grid value is set to missing.

Bourras (2006) described several other satellite products not included in our comparison. Two of the products used by Bourras (2006), HOAPS2, and the Goddard

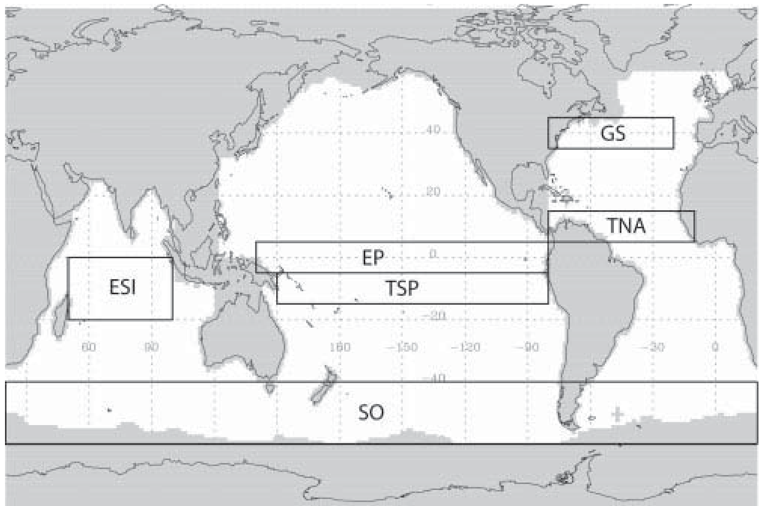

Figure 1. One-degree land (gray) *versus* water (white) mask used when comparing the flux products. Regional comparisons (black boxes) are conducted for the Gulf stream, tropical north Atlantic, equatorial Pacific, tropical south Pacific, equatorial south Indian, and southern ocean.

satellite-based surface turbulent fluxes (GSSTF2) (Section 2.4) overlap with our comparison project, allowing some comparison of our respective results. In addition, a third HOAPS product now available is not examined herein. The variations in HOAPS3 are mostly in precipitation, with very small changes in SST and hence in heat fluxes.

2.3. Reanalysis

Three reanalysis products are included in the comparison: the Japanese 25-year reanalysis (JRA; Onogi *et al*., 2007), the National Centers for Environmental Prediction–Department of Energy Atmospheric Model Intercomparison Project reanalysis (NCEPR2; Kanamitsu *et al*., 2002), and the ECMWF second generation reanalysis (ERA-40; Uppala *et al*., 2005). Each reanalysis is constructed by running a static version of its respective 3D variational (3DVAR) assimilation method and applying this model to a differing set of input observations. For example, ERA-40 merged the raw data used for NCEPR2 with data already at ECMWF, but ERA-40 also includes additional satellite observations not included in NCEPR2. The forcing variables (winds, T_{air}, and humidity) from ERA-40 and JRA compared herein are not derived from the 3DVAR assimilation, but from a 2D optimum interpolation scheme. This approach does not allow information from the upper levels to impact the ERA-40 and JRA surface analysis, a detriment when the upper analysis is accurate and a strength when the upper analysis is poor. The weaknesses probably outweigh the strengths in the Northern Hemisphere mid-latitudes, and vice versa in the Southern Hemisphere mid-latitudes.

The monthly air–sea fluxes in all three reanalyses are derived from the 6-h integrated fluxes produced by the 3DVAR method. The ECMWF oceanic surface turbulent fluxes (A. Beljaars, personal communication, 2007) are calculated in a manner largely dependent on stability (Paulson, 1970; Holtslag and De Bruin, 1988) and roughness length parameterisations. The roughness length parameterisations are similar to Smith's (1988; his Equation 7) and Beljaars (1995), except that Charnock's (1955) constant is dependent on wave characteristics (Janssen, 1989). One shortcoming, resulting in overestimation of the LHF, is the treatment of near-surface humidity as saturated for the T_{skin}. A 98% saturation is more realistic for typical oceanic surface values of salinity. The JRA oceanic surface turbulent fluxes are determined in a qualitatively similar manner except for the use of the stability dependence of Louis *et al*., (1982), which uses a bulk Richardson number to approximate the Monin–Obukhov scaling used for the ECMWF reanalysis. This approach is more computationally efficient, at the expense of some accuracy. Furthermore, the JRA value of Charnock's constant is fixed at 0.020, a compromise between typical open ocean swell values and wind driven waves. The NCEPR2 surface turbulent flux parameterisations (Long, 1986, 1990) also use a bulk Richardson number to approximate the Monin–Obukhov length scale (L). The stability parameterisations are those of Dyer (1974) and Hicks (1976) for unstable stratification and that of Long (1986) for stable stratification, which is similar to that of Nickerson and Smiley (1975). The great difference is the roughness length parameterisation, which assumes that the roughness length (for momentum, temperature, and moisture) is equal to 0.02 L. Long (1986, 1990) examined the sensitivity to the value of roughness length and found that errors due to this approximation were probably within the bounds due to uncertainty in the stability parameterisation; however, the values of roughness length used in their analysis were much larger than those found in current parameterisations, which could alter Long's conclusion regarding accuracy. The calculation of the near-surface humidity is not described in available JRA and NCEPR2 documentation.

All the three monthly reanalysis products are bilinearly interpolated from their native grid (spectral for JRA and NCEPR2, fixed 1.125° for ERA-40) to the 1° grid used for the comparisons. The use of linear interpolation, rather than a curve-fitting method, was determined to be sufficient for our analyses. No spurious values are apparent in strong gradient regions where a curve-fitting method might be more appropriate.

Several reanalyses are excluded from the comparison. Kanamitsu *et al*. (2002) stated that the NCEPR2 was a follow-on project to the initial NCEP–National Center for Atmospheric Research (NCEPR1) reanalysis (Kalnay *et al*., 1996) that fixes the human processing errors discovered in the NCEPR1. This comment, along with previous studies showing large biases in NCEPR1 heat fluxes (Bony *et al*., 1997; Smith *et al*., 2001), results in the exclusion of the NCEPR1. In addition, we have not examined the plethora of recently released reanalysis products.

The descriptions of the three reanalysis products revealed several characteristics worth noting regarding the reanalysis flux fields. Both NCEPR2 and ERA-40 use a smoothed orography that reduces the Gibbs phenomena (model ringing) noted over oceans near high topography (which was especially notable in the NCEPR1). Gibbs phenomena are noted to exist in the water vapor-related variables in the JRA (Onogi *et al*., 2007). Preliminary evaluations of the NCEPR2 found an overestimation of the outgoing longwave radiation over the tropical warm pool and also found that the upper-level tropical moisture is drier than the NCEPR1. The NCEPR2 problems are related to the new boundary layer formulation, which might have improved the precipitation, but inadvertently worsened the radiation budget (Kanamitsu *et al*., 2002). In the ERA-40, ship winds are assimilated at measured height when known; otherwise, they are assumed to be at 25 m. As noted in Section 2.1, a long-term trend existing in anemometer heights might add a small bias to the fluxes and forcing parameters in ERA-40 over the oceans; however, the low weight of ship observations (as compared to radiosonde and satellite data) in ERA-40 might render this bias undetectable. Finally, the ERA-40 analyses are found to be more moist over the tropical oceans due to the assimilation of satellite data (Andersson *et al*., 2004), which produces excessive tropical oceanic precipitation and likely impacts the LHF.

2.4. Hybrid

Hybrid products include the second version of the GSSTF2 (Chou *et al*., 2003) and the Woods Hole Oceanographic Institution's objectively analysed fluxes (OAFLUX; Yu and Weller, 2007). The term hybrid is chosen because both products objectively combine forcing variables from NWP reanalyses and satellite observations. The GSSTF2 combines surface wind and humidity from SSM/I with wind direction and air and sea temperatures from the NCEPR1 to obtain LHF, SHF, and stress using a bulk flux algorithm by Chou (1993). The heat and moisture roughness lengths are based on the Liu-Katsaros Businger model, and the momentum roughness length is similar to that used for the JRA, except that the value of Charnock's constant is 0.0144 (more typical of swell). The OAFLUX is derived from wind speed, sea and T_{air}, and humidity from NCEPR2, NCEPR1, and ERA-40, combined with satellite retrievals of wind speed and humidity from SSM/I and the SeaWinds scatterometer on QuikSCAT, and sea temperature from Advanced Very High Resolution Radiometer (AVHRR), Tropical Rainfall Measuring Mission Microwave Imager (TMI), Advanced Microwave Scanning Radiometer - E. The OAFLUX method synthesises the satellite observations and NWP outputs using a variational objective analysis to determine the best fit for the four independent variables (wind speed, air and sea temperature, and specific humidity). Earlier versions of the OAFLUX product incorrectly treated satellite winds, which are calibrated to equivalent neutral winds (Verschell *et al*., 1999), as actual winds (there are stability dependent differences), resulting in regional biases in wind speed. One clear benefit of this product comparison has been several improvements of the OAFLUX product (e.g. the use of NCEPR2 instead of NCEPR1 when NCEPR2 is available). According to L. Yu (personal communication, 2009), the OAFLUX fluxes are computed using the COARE version 3.0 bulk flux algorithm (Fairall *et al*., 2003) applied to daily averages of the input data (e.g. six hourly) NWP input. The OAFLUX product is under frequent, undocumented development and modification without version control. When this comparison was completed, only fields of LHF and SHF were available (Table I), but new releases now include the forcing variables (not discussed herein).

These hybrid products attempt to synthesise information from a variety of sources. Although they can take advantage of the complete spatial and temporal coverage of the reanalysis products, they are then subject to the internal biases of these products, as well as large random errors in the SO. For example, NCEPR1 has been shown to underestimate the wind speed in the Tropics (Putman *et al*., 2000) and Smith *et al*. (2001) revealed a cold bias in NCEPR1 compared to research ship observations. As noted above, the Gibbs phenomenon is prominent in the NCEPR1 and this propagates into the GSSTF2 and the early versions of the OAFLUX product.

2.5. Variations in flux algorithms

The nine products in the comparison employ eight different flux algorithms (OAFLUX and IFREMER use COARE version 3.0). The consequences of using different algorithms include regional biases related to (1) sea state, (2) density stratification (stability), which is largely driven by air–sea temperature differences, and (3) roughness length parameterisations in the case of the NCEPR2 product. Although differences in the algorithms contribute, in part, to the differences between the flux fields, it is likely that biases in the inputs to these flux algorithms are at least as important as the differences in parameterisation. Quantifying the differences in the algorithms would require using a common set of forcing data and is beyond the scope of this manuscript.

2.6. Averaging methods

The monthly flux products are derived using different temporal averaging methods. The NWP products are averaged over 6 h of model computations. The HOAPS, IFREMER, and GSSTF2 products calculate the fluxes after determining a daily average of the flux-related variables. The OAFLUX product creates daily objective analyses, which are used in the flux calculation. Similarly, the NOC product is based on bulk fluxes estimated in roughly six-hourly steps; however, these estimates are created only from platforms where there are observations of all the variables needed to determine the fluxes. The monthly averaged fluxes are then objectively interpolated. It has been suggested that the sampling errors in such an approach would require great smoothing (Gulev *et al*., 2007a, 2007b). The FSU3 objective interpolation greatly reduces the errors related to the sampling pattern by applying the bulk flux formulas to the monthly averaged meteorological variables (i.e. the classical method; Esbensen and Reynolds, 1981). The classical method, in contrast to the sampling method, implicitly neglects the effects of variability on time scales shorter than the averaging period and, thus, can introduce biases in the monthly mean flux estimates (Esbensen and Reynolds, 1981; Simmonds and Dix, 1989; Gulev, 1994; Josey *et al*., 1995; Zhang, 1995; Esbensen and McPhaden, 1996; Gulev, 1997). The FSU3 approach works very well for stress and pseudostress; however, it results in regional biases in surface energy fluxes. In the Atlantic basin, there are enough observations that these biases could be determined and removed from a product; however, that is unlikely to be the case in other basins. The calculation of fluxes over these different time scales (six-hourly, daily, and monthly) results in differences that are particularly large during the winter, in the mid-latitudes, and near strong surface temperature gradients (e.g. western boundary currents). In such regions, this source of error could exceed the problems mentioned above (e.g. inhomogeneous surface sampling, unknown observing practices and reference heights, variations in flux algorithms and satellite retrieval methods).

3. Product comparison

The nine flux products are compared for a common period from March 1992 through December 2000. The comparisons begin at the global scale. Regional analyses examine product variability in the Atlantic, Pacific, Indian, and Southern Oceans using zonal distributions of heat fluxes, wind stress, and the forcing variables for the fluxes. Zonal analyses at multiple quantiles reveal differences (and similarities) in the median and tails of the monthly LHF, SHF, and wind stress distributions. Regions (Figure 1) with large differences in the zonal quantiles are further examined using box and whisker plots which reveal the forcing variables in the region that are primarily responsible for differences in the heat fluxes.

3.1. Global comparison

The overall mean LHF from March 1992 to December 2000 generally shows that all products have similar large-scale spatial patterns, but large differences in magnitude (Figure 2(a)). LHF is generally maximised in the subtropical highs and over western boundary currents, with the NCEPR2 having the highest (monthly averaged) maxima (>220 Wm^{-2}). Similar to Bourras (2006), we find the peak magnitudes to be greater in the GSSTF2 than in the HOAPS2. The LHF maxima are generally lowest in the *in situ*-based products (FSU3 for the Indian and Atlantic Oceans where this product is available and NOC for the Pacific). In general, the LHF in the FSU3 is less than the other eight products. This is believed to be due to systematically larger values of q_{air} in the FSU3, reducing the air–sea moisture gradient. This bias in humidity is largely consistent with the adjustments made in a new NOC product (L. Kent, personal communication, 2008; Berry and Kent, 2009). The JRA product also has relatively large values of q_{air}, but retains large LHFs because the value of near-surface humidity is overestimated.

Both the spatial and amplitude patterns vary between the products for SHF (Figure 2(c)). Overall, the JRA has larger magnitudes over most of the globe compared to the other products. The two satellite products (IFREMER and HOAPS2) have a northwest-to-southeast oriented maximum in the Pacific (west of South America) and a similar feature in the south Atlantic. These features are not apparent in any of the other products. These maxima might be associated with the south Pacific and south Atlantic convergence zones (SPCZ and SACZ), which might not be well represented in the *in situ* (likely due to poor sampling), reanalysis, or blended products. On the other hand, the satellite products both estimate T_{air} based on either SST or q_{air} and limitations of these retrievals may be resulting in the patterns over the south Pacific and south Atlantic. The lack of *in situ* observations in these regions limits our ability to validate the satellite retrievals.

Monthly standard deviations for the heat fluxes (Figure 2(b) and (d)) reveal patterns of variability that are largely due to sampling and natural variability. For the NOC product, a pattern of primary ship tracks exists with standard deviations being low where observational density is high (e.g. the X pattern in the Indian Ocean represents the two primary ship tracks across this basin). The analysis approach of the FSU3 largely eliminates the appearance of ship tracks in the monthly standard deviations.

The mean and standard deviations for the nine products also show clear variations in the heat fluxes over the primary western boundary currents. The mean magnitude of the LHF (SHF) over the Gulf stream (GS) ranges from 160 Wm^{-2} (35 Wm^{-2}) in the satellite products to well over 220 Wm^{-2} (50 Wm^{-2}) in the NCEPR2. The two *in situ* products show surprisingly different magnitudes for SHF over the GS, with the FSU3 having SHF >50 Wm^{-2}. The reanalysis products (and the OAFLUX which is dependent on the reanalyses) show the largest LHF

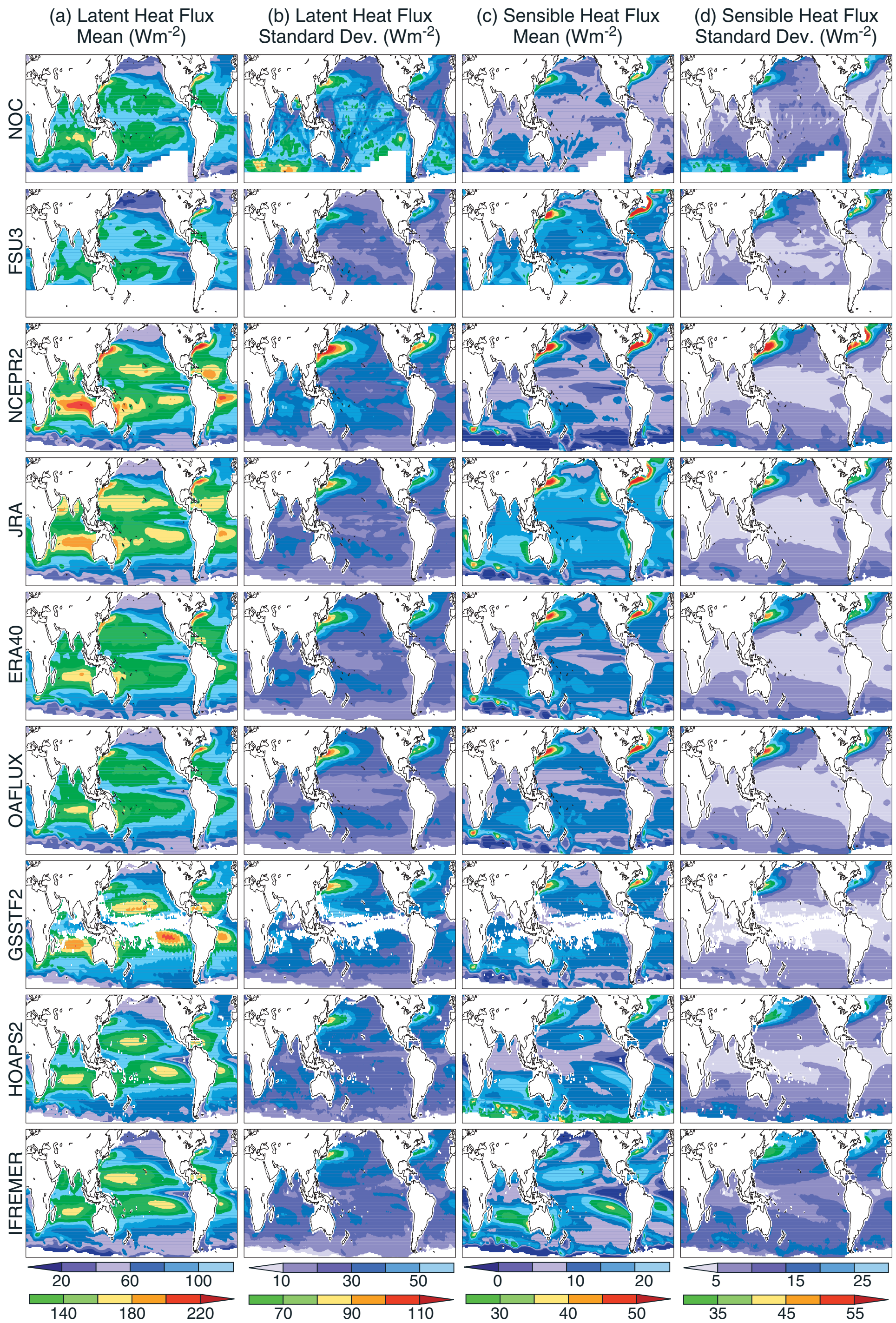

Figure 2. Average (a) latent and (c) sensible heat flux and their corresponding standard deviations (b and d) for the common period from March 1992 through December 2000. The magnitudes (Wm^{-2}) for each of the nine products are shown in the color bars. White areas on the maps represent 1° cells that are either defined as land (refer Figure 1) or had one or more missing months in the common period. The GSSTF2 had several months with missing data near the equator due to variations in satellite coverage and the FSU3 is not produced for the southern ocean. The Gibbs phenomenon is notable in the GSSTF2 means.

and SHF magnitudes over the Kuroshio and these high LHF and SHF extend further to the east than in the other products. The standard deviations for the reanalysis products also reveal a broader spread to the LHF and SHF distributions in the region of the Kuroshio.

The divergence of the wind is an important quantity for understanding global atmospheric circulation. Only five products provided the component winds necessary to compute the divergence. The mean divergence for the common period (Figure 3(a)) shows similar spatial

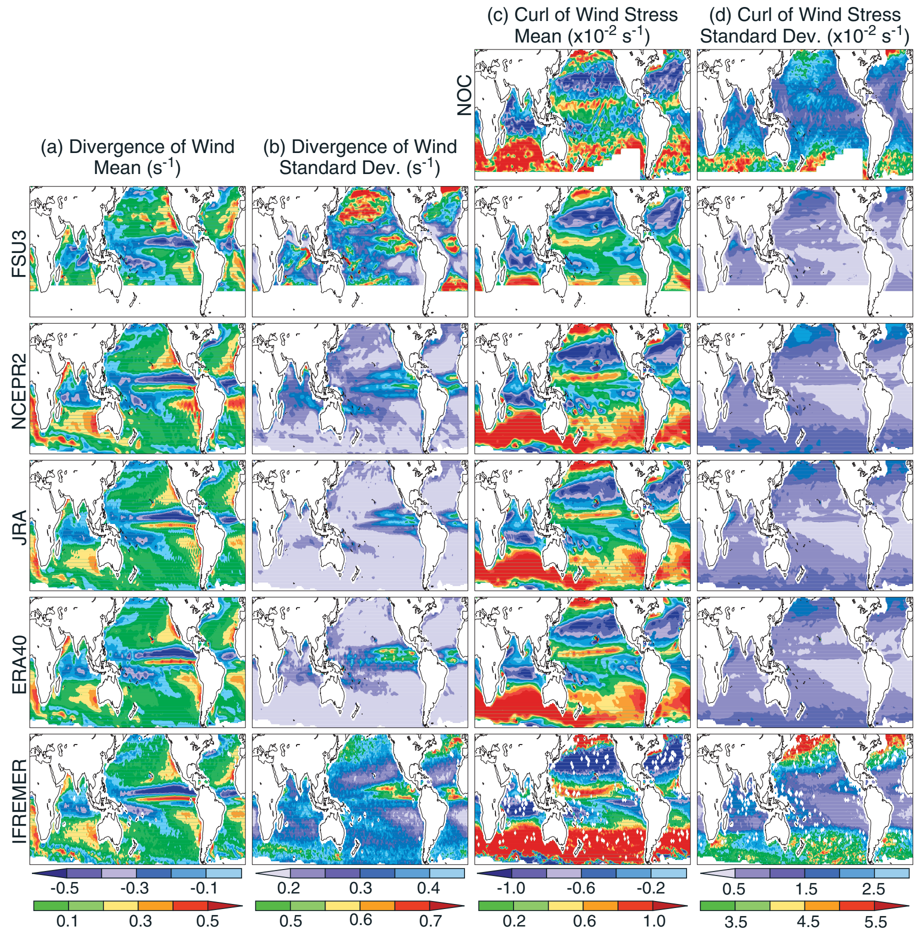

Figure 3. Same as Figure 2, except for the wind divergence (a and b) and wind stress curl (c and d). HOAPS2, OAFLUX, and NOC (divergence only) did not provide the necessary parameters to determine the wind divergence and curl of the wind stress. GSSTF2 was omitted due to the missing data in the Tropics.

patterns in the five products. The equatorial divergence in the Pacific is strongest in the IFREMER product. The standard deviation of the divergence (Figure 3(b)) shows larger values in the southwest Pacific and southwest Atlantic Oceans in the IFREMER, as compared to the three reanalysis products, which might indicate that the IFREMER is better resolving the variability in divergence/convergence associated with extratropical cyclone activity along the SPCZ and SACZ. Alternatively, the IFREMER product has been found to poorly filter out rain-related problems in wind vectors from NSCAT and QuikSCAT, which results in erroneous wind patterns (A. Bentamy, personal communication, 2008). The standard deviation of the divergence also reveals a series of north-to-southoriented bands along the equatorial Pacific (EP) for the ERA-40, corresponding to the locations of the TAO moorings, which are not present in divergence from the other four products.

The curl of the wind stress is of particular importance for forcing the ocean circulation. We derived the curl of the wind stress from six of the products and, although the spatial patterns are similar, the magnitudes and noise characteristics are largely different (Figure 3(c)). The NOC product is adversely affected by the variations in spatial sampling of the *in situ* data, resulting in a patchy curl field with many artificial gradients. The FSU3, which has a curl-based constraint in the objective algorithm, provides a smoother spatial pattern for the curl of the wind stress that is spatially similar to the three reanalysis products, but differing in magnitude. The standard deviation of the curl (Figure 3(d)) exceeds 0.02 s^{-1} poleward of roughly 30° latitude in the IFREMER and NOC products. These large standard deviations are not present in the other four products. The key contribution to these differences is likely the much greater noise in the NOC and IFREMER products. This noise is larger in areas of greater natural variability and, in the case of NOC, in areas of more variable sampling, both of which are consistent with the above-mentioned pattern.

3.2. Zonal distributions

The zonal medians of the latent and SHFs in the Atlantic, Indian, and Pacific Oceans show both similarities and differences in the nine products (Figures 4 and 5). Overall, the JRA and NCEPR2 (FSU3) tend to have the

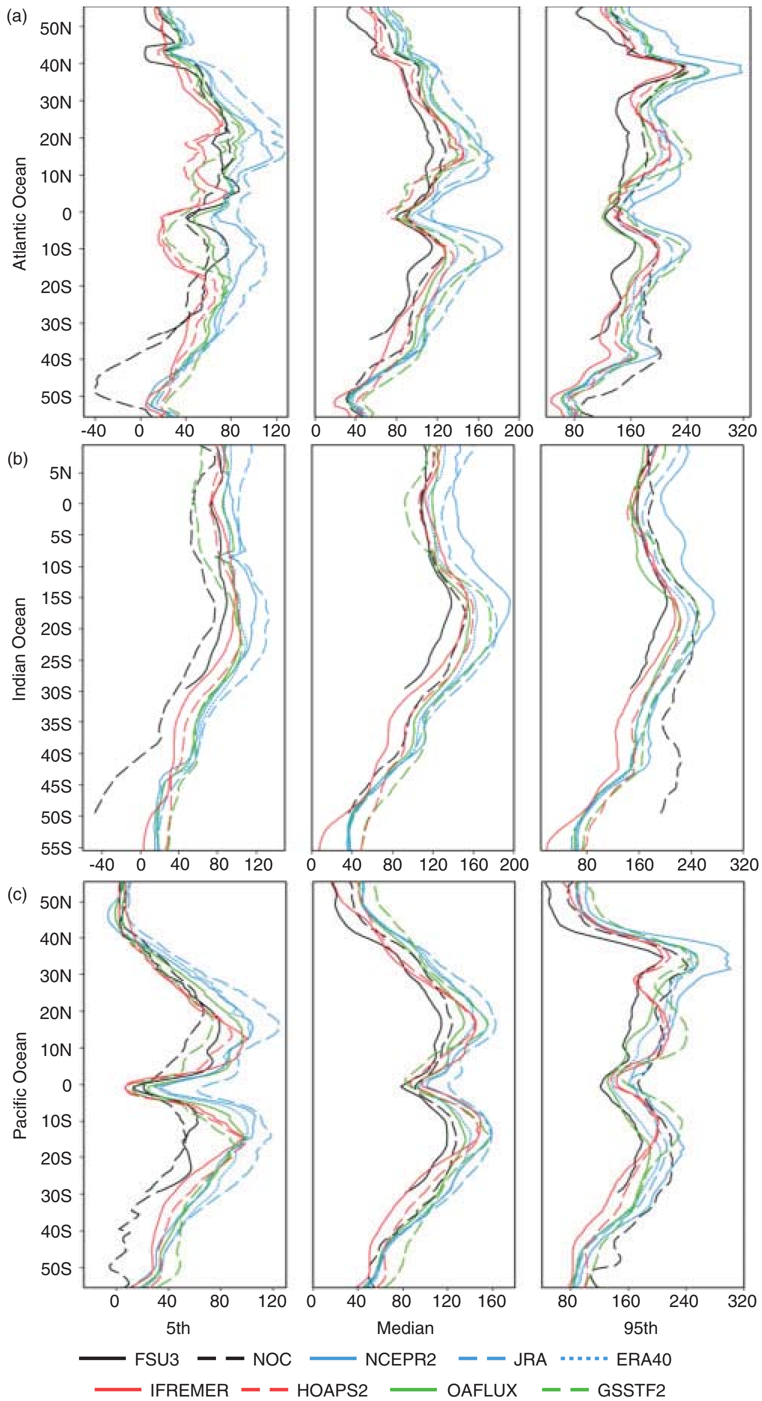

Figure 4. The value at the 5th, median, and 95th percentiles for the zonal distribution of latent heat flux (Wm^{-2}) for the (a) Atlantic, (b) Indian, and (c) Pacific Oceans. Percentiles are determined from the distribution of 1° grid values along each latitude for all months from March 1992 to December 2000. Colours denote *in situ* (black), reanalysis (blue), satellite (red), and hybrid (green) products, respectively. Note that the x-axis scales vary from basin to basin.

highest (lowest) median LHF magnitude at most latitudes in all three basins (Figure 4). The overall median in SHF reveals the JRA (NCEPR and NOC) to have the highest (lowest) magnitudes (Figure 5) in all basins. The products all show a relative minimum in LHF near the equator with the largest spread in median LHF between the products occurring in the tropical latitudes, north and south of the equator, in all basins (Figure 4). This spread maximises at $\sim$85 Wm^{-2} near 10°N and 10°S in the Atlantic, $\sim$60 Wm^{-2} near 17°S in the Indian Ocean, and $\sim$50 Wm^{-2} near 15°N and 15°S in the Pacific. Noteworthy is the median GSSTF2 LHF in the Indian Ocean, which

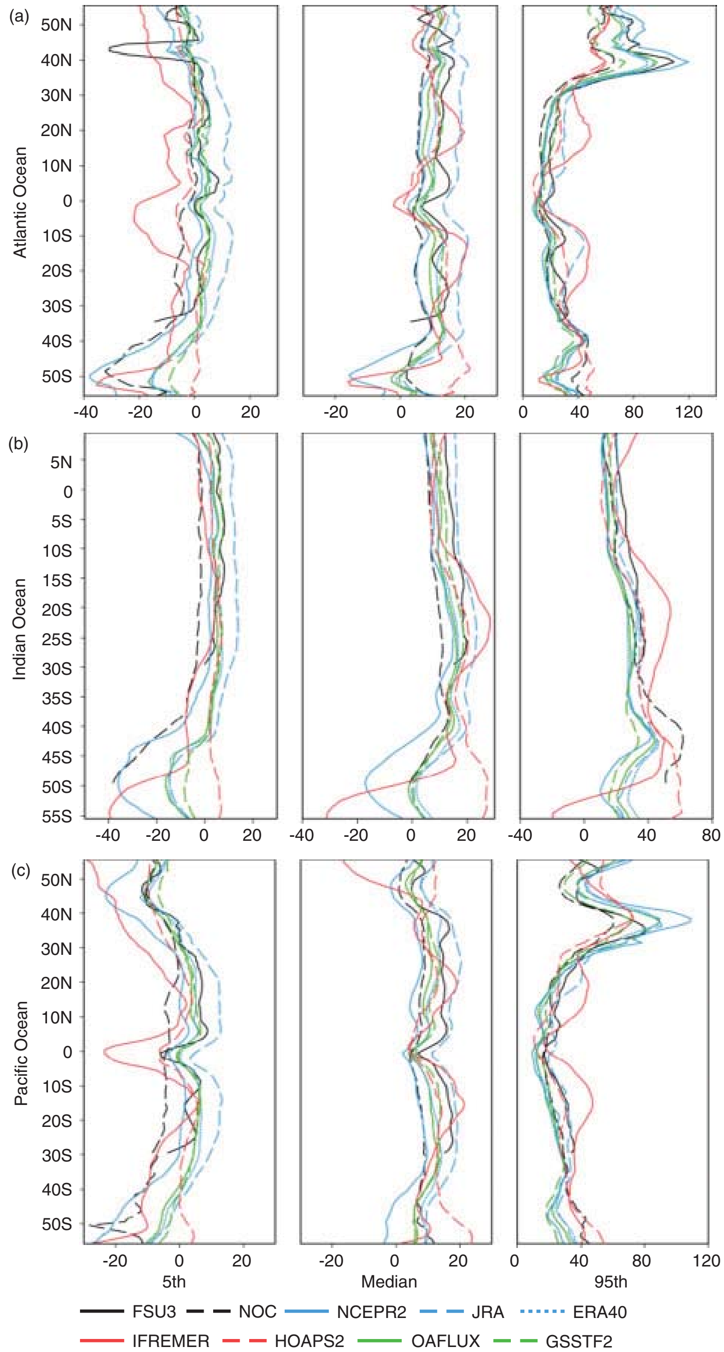

Figure 5. Same as Figure 4, except for sensible heat flux (Wm^{-2}). Note that the x-axis scales vary from basin to basin.

crosses from being the lowest LHF (100 Wm^{-2}) product at the equator to the third highest LHF (180 Wm^{-2}) at 20 °S (Figure 4(b)). Similar swings in the LHF values are not prominent in the other products. In fact, the FSU3 shows only an ∼30 Wm^{-2} difference from the equator to 20 °S. The median SHF remains fairly constant within each product at varying latitudes (Figure 5). Exceptions include relative minima at the equator and an increase in spread between SHF in the high southern latitudes (south of ∼35°S) due to large variations in the forcing variables in the SO (refer Section 3.3). The median SHF from IFREMER also differs from the other eight products due to large swings in magnitude in all ocean basins (e.g. in the Atlantic from ∼20 Wm^{-2} around 20°N and 20 °S to −5 Wm^{-2} at the equator and ∼5 Wm^{-2} at 40°N).

Examining the upper and lower tails of the heat flux distributions also reveals regional differences. A larger spread in 95th percentile LHF and SHF between the products is present at ∼40 °N in both the Atlantic and

Pacific Oceans. This signal represents the position of the GS and Kuroshio, respectively. The NCEPR2 (NCEPR2 and FSU3) exhibit the highest LHF (SHF) at the 95th percentile for the GS (Figures 4(a) and 5(a)), while the FSU3 also exhibits distinctly lower 5th percentile LHF and SHF values in the same region. Although the authors anticipate more variability near these western boundary currents, the differences between the products are larger than expected. Other notable differences in the tails of the heat flux distributions occur in the SO (south of ∼35 °S) and in the tropical oceans at ∼10° north and south of the equator. Regional differences are examined in detail in Section 3.3.

The wind stress for the nine products is very similar at the median, with some larger differences occurring in the tails of the distribution. The median meridional component of the wind stress (Figure 6) is surprisingly consistent between the products at all latitudes (the spread between products rarely exceeds 0.02 Nm^{-2}). The spread in the median zonal component of the wind stress (Figure 7) is equally small in the equatorial regions (∼0.05 Nm^{-2}), but increases south of 30 °S. The IFREMER zonal wind stress is an outlier poleward of 30° in all basins, with median values ∼0.1 Nm^{-2} larger than the other eight products. In the tails of the zonal wind stress, the IFREMER reveals a broader spread from the 5th to the 95th percentiles than any of the other products north of 30°N. South of 40 °S, the IFREMER 95th percentile exceeds the other products, but has stronger westerly component than the other products at the 5th percentile (in contrast to the more easterly component in the Northern Hemisphere). For the meridional wind stress, the NOC clearly is the outlier in the high southern latitudes, likely due to poor spatial sampling by ships in the region.

3.3. Regional comparisons

The zonal distributions of the fluxes revealed regions in each basin where the nine products exhibited broader spread. Six regions, marked in Figure 1, are examined using box and whisker analyses. The goal is to identify which forcing variables contribute to the differences noted in the fluxes. Note that the comparisons are limited in some cases due to some parameters not being available for each product (Table I). In addition, the NOC and FSU3 acknowledge the lack of data in the SO, so they are not included in SO comparisons.

The T_{air} distributions are examined for six products (Figure 8). Overall, the median values are largely similar, being within 1 °C for all products in the four tropical regions (tropical north Atlantic (TNA), equatorial south Indian (ESI), tropical south Pacific (TSP), equatorial Pacific (EP); Figure 8(b)–(e)). The medians in the GS region (Figure 8(a)) are tightly clustered, with the exception of the IFREMER, which is shifted toward larger T_{air} values. The NCEPR2 is the outlier in the SO (Figure 8(f)), with median values shifted ∼3 °C lower than the other three T_{air} products available for the region. The most notable differences in the T_{air} in all regions occur in the spread of the distributions. The IFREMER and NCEPR2 have the largest interquartile range (IQR) in several regions (TNA, ESI, EP; Figure 8(b), (c), and (e) respectively).

The sea temperature (T_{skin}) is available for seven of the products (Figure 9) and exhibits largely similar distributions within each of the six regions. The similarity is expected to an extent since most flux products are derived from Reynold's SST products. With the exception of the SO (Figure 9(f)), where the NCEPR2 and HOAPS2 are notably lower than the other products, the regional differences between T_{skin} are subtle. The FSU3 tends to be shifted slightly to larger T_{skin} values in the five regions where it is available (Figure 9(a)–(e)), and HOAPS2 T_{skin} distributions are shifted to slightly lower values in all regions. In the GS, TNA, and SO (Figure 9(a), (b), and (f)), the NCEPR2 exhibits a larger spread between the 10th and 90th percentile and substantially lower 5th and 1st percentile values (becoming unrealistically negative due to T^{skin} being over ice in the SO) than the other products.

The near-surface specific humidity (q_{air}) exhibits notable differences between the eight products for which it is available (Figure 10). The FSU3 q_{air} distribution is shifted toward higher values, with the median q_{air} being 1.5–2.0 g kg^{-1} greater than most other products in the five regions where it is available. Only the IFREMER in the GS region exhibits a similar q_{air} distribution to the FSU3 (Figure 10(a)). In the four tropical regions (Figure 10(b)–(e)), the FSU3 median q_{air} exceeds the 75th percentile value of several other products. Of the three reanalysis products, the JRA q_{air} tends to be shifted toward higher values. In the Tropics (Figure 10(b)–(e)), the NCEPR2 and ERA-40 also have the smallest IQR compared to the other products. The q_{air} distributions in the SO (Figure 10(f)) are separated into two distinct groups with the JRA, ERA-40, and IFREMER having similar distributions shifted toward higher q_{air} as compared to the NCEPR2, GSSTF2, and HOAPS2.

Examining the regional distributions of the SHF (Figure 11) reveals substantial differences in the nine products. One striking difference is the shifts in the SHF distributions exhibited by the three reanalysis products in the tropical regions (Figure 10(b)–(e)). Within these regions, the IQR of the SHF for the reanalyses barely overlap, with the JRA tending to have the highest (e.g. 15–20 Wm^{-2} in the TSP; Figure 11(d)) and the NCEPR2 the lowest (e.g. 3–9 Wm^{-2} in the TSP; Figure 11(d)) SHF. In the Tropics, the shifts in the reanalysis SHF match the shifts in the $T_{\mathrm{skin}} - T_{\mathrm{air}}$ (Figure 12(b)–(e)). The JRA exhibits greater differences between the $T_{\mathrm{skin}} - T_{\mathrm{air}}$, with the JRA IQR being fully higher than the IQR for the NCEPR2 in the TNA, ESI, and TSP regions (Figure 12(b)–(d)). The lack of substantial differences in the reanalysis wind speeds (not shown, but consistent with zonal wind stress in Figure 7) or T_{skin} (Figure 9(b)–(e)) implies that differences in T_{air}

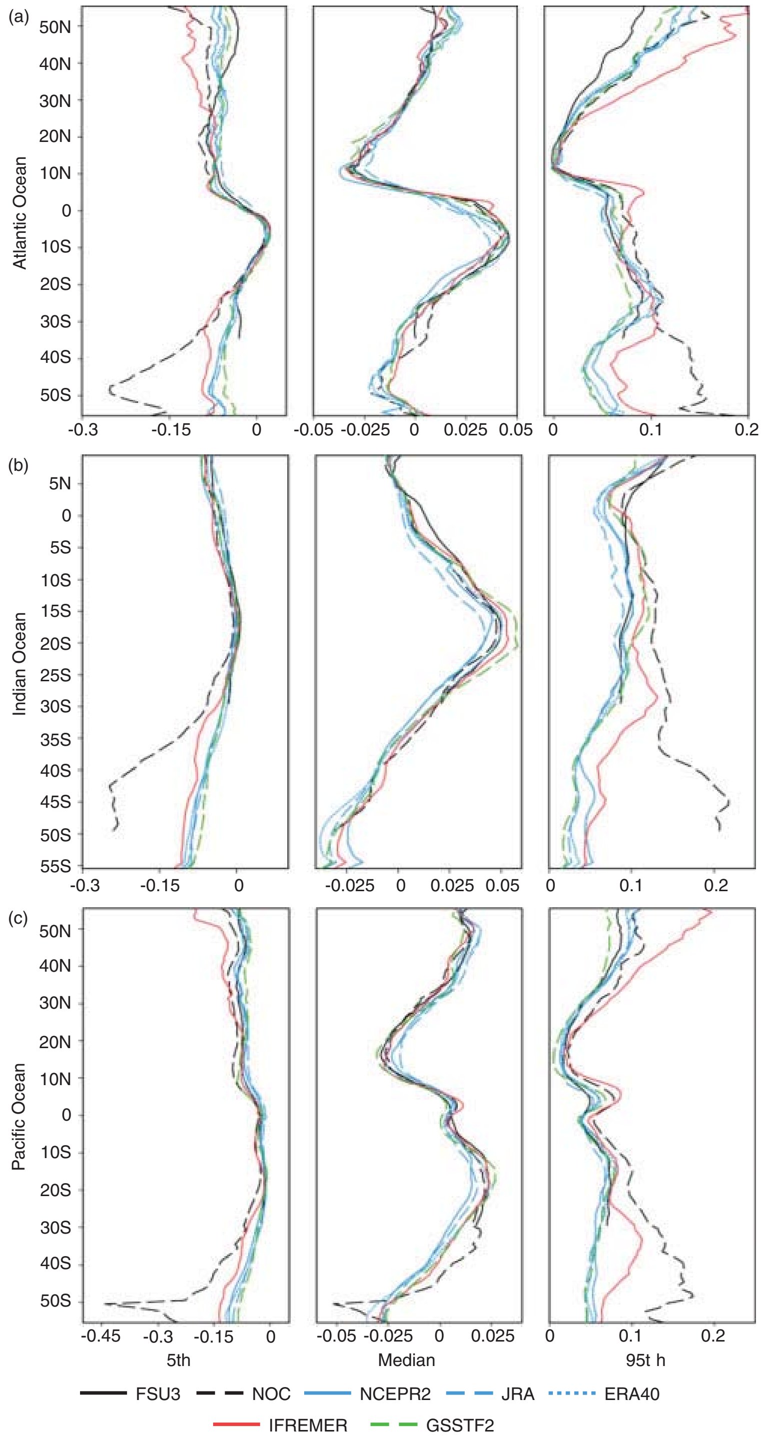

Figure 6. Same as Figure 4, except for meridional wind stress (Nm^{-2}). Note that the x-axis scales vary from basin to basin.

(Figure 8(b)–(e)) or some other factor (probably the flux algorithms) are responsible for the substantial differences in SHF in the reanalyses.

The large spread in the IFREMER SHF distributions is another notable difference in the products. Excluding the GS region, the IQR for the IFREMER is nearly double that of the other products. This increase in spread might be the result of a larger spread in IFREMER $T_{\text{skin}} - T_{\text{air}}$ (Figure 12) combined with increased spread in the distribution of the scalar wind speed (not shown, but also exhibited in the zonal wind stress; Figure 7) relative to the other products. The temperature and wind influence on the IFREMER SHF is also notable in the skew toward high SHF in the ESI (Figure 11(c)), where

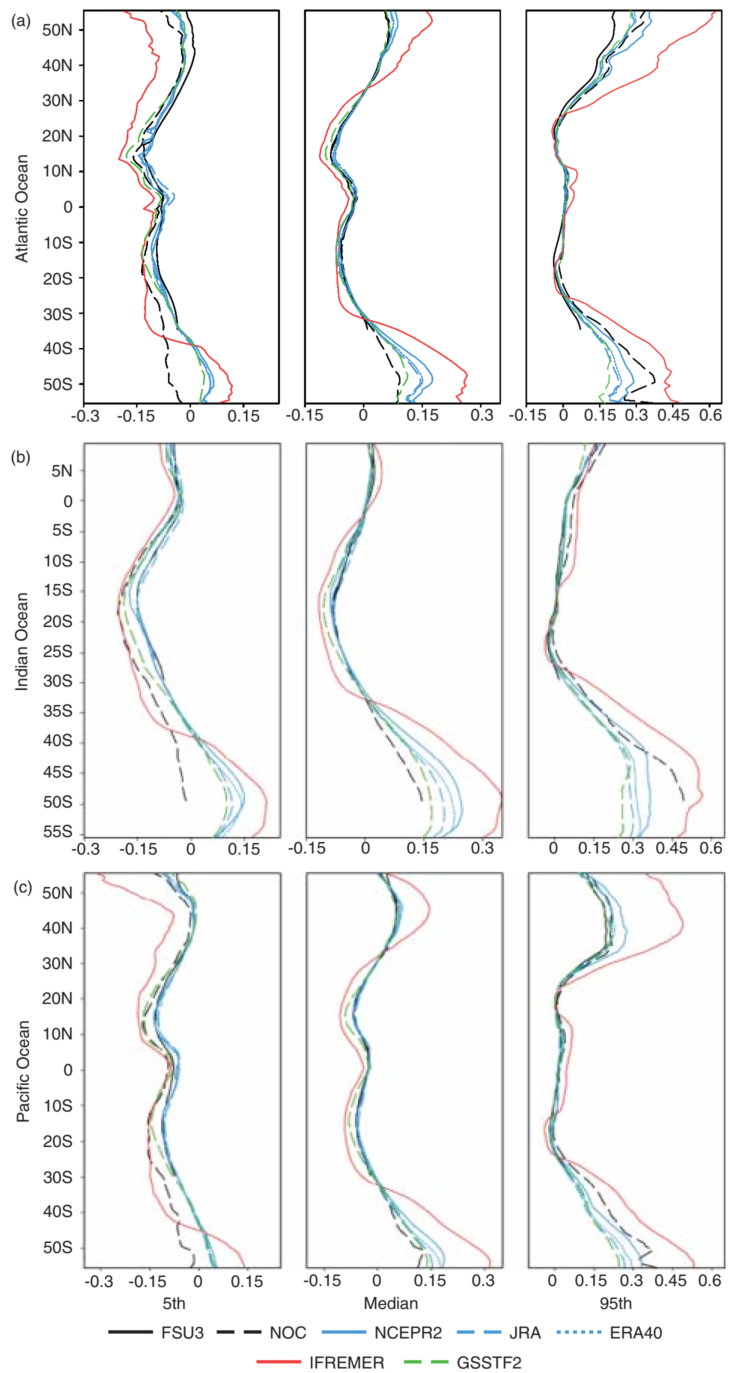

Figure 7. Same as Figure 4, except for zonal wind stress (Nm^{-2}). Note that the x-axis scales vary from basin to basin.

the 99th percentile of the IFREMER SHF is nearly 20 Wm^{-2} greater than the other products in the region. The IFREMER T_{air} in the ESI (Figure 8(c)) exhibits a skew toward lower values, which would result in a skew of the $T_{skin} - T_{air}$ (Figure 12(c)) and SHF toward higher values since the T_{skin} for the IFREMER is consistent with the other products (Figure 9(c)).

Examining the SHF distributions for the SO (Figure 11(f)) reveals substantial differences that were anticipated by the authors. There is no overlap in the IQR between the highest (HOAPS2) and lowest (NCEPR2) products. The difference in medians between these products is nearly 25 Wm^{-2} and the majority of the IQR for the NCEPR2 falls in the range of negative SHF. $T_{skin} - T_{air}$ for the NCEPR2 are shifted toward lower values (Figure 12(f)) and this might contribute to the low SHF. Assessing the high SHF for the HOAPS2 is not possible as this product does not provide T_{air}.

Regional differences in LHF also exhibit substantial differences between the nine flux products (Figure 13).

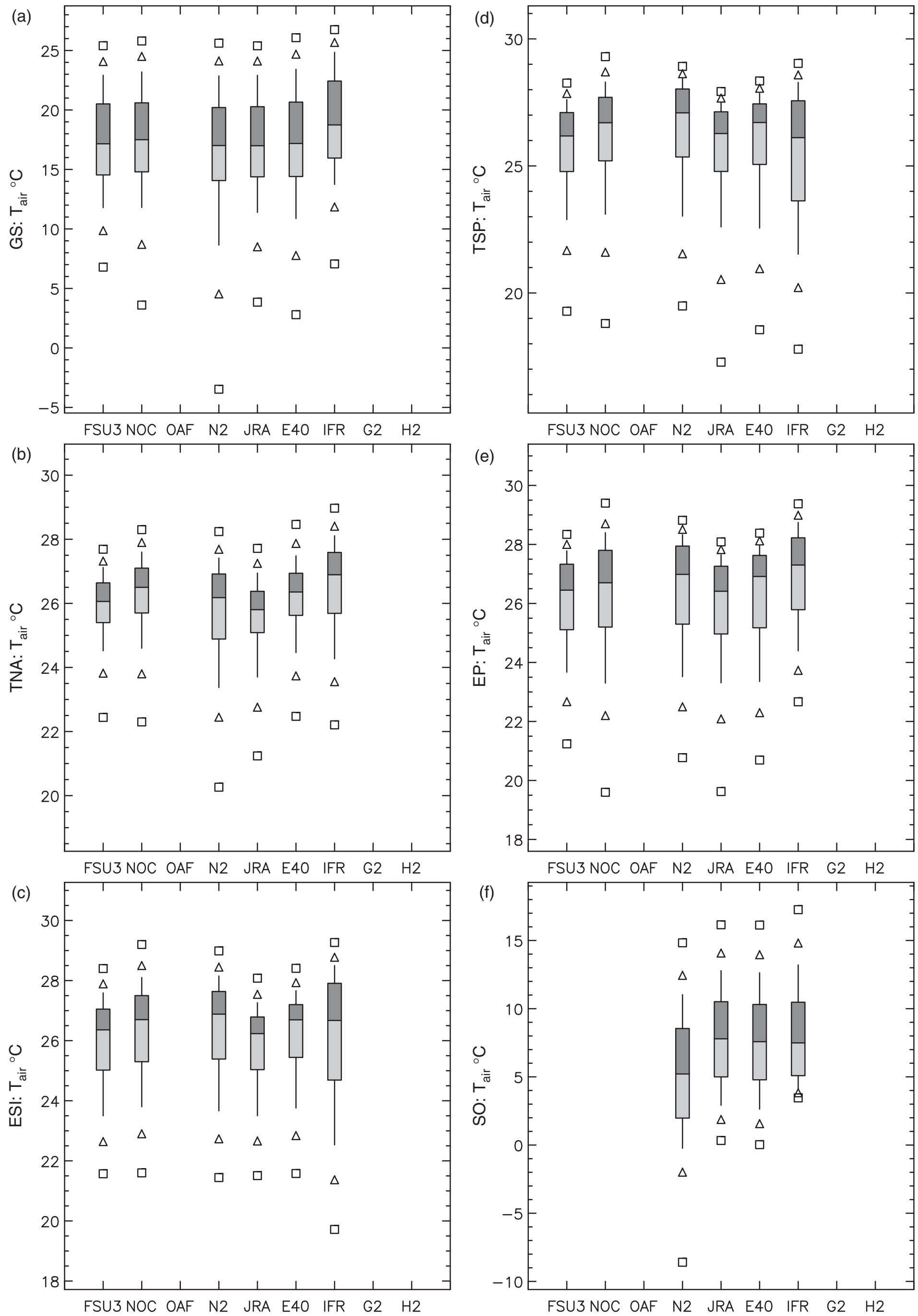

Figure 8. The distribution of air temperature for the (a) Gulf stream, (b) tropical North Atlantic Ocean, (c) southern equatorial Indian Ocean, (d) tropical South Pacific Ocean, (e) equatorial Pacific Ocean, and (f) circumpolar southern ocean. The upper and lower ends of the boxes are drawn at the 75th and 25th quartiles, respectively, and the bar through the box is drawn at the median. The whiskers extend from the quartiles to the 90th and 10th percentiles. The triangles (squares) represent the 95th and 5th (99th and 1st) percentiles, respectively. OAF, N2, E40, IFR, G2, and H2 mark the OAFLUX, NCEPR2, ERA-40, IFREMER, GSSTF2, and HOAPS2 products.

The JRA has the highest median LHF in four regions (GS, TNA, TSP, and EP; Figure 13(a), (b), (d), and (e)). The NCEPR2 LHF distribution also tends to be shifted toward higher values, most notably in the ESI (Figure 13(c)). In the TNA (Figure 13(b)), median LHF values for the JRA and NCEPR2 are 20 $\mathrm{Wm^{-2}}$ and 25 $\mathrm{Wm^{-2}}$ higher than the ERA-40, respectively. The NCEPR2 q_{air} distribution is shifted to lower values (Figure 10(b)) that might increase $q_{\mathrm{sfc}} - q_{\mathrm{air}}$, resulting in higher LHF. Unlike the $T_{\mathrm{skin}} - T_{\mathrm{air}}$, the authors are unable to accurately calculate $q_{\mathrm{sfc}} - q_{\mathrm{air}}$ from monthly means due to nonlinearities with respect to temperature, so the following interpretations using $q_{\mathrm{sfc}} - q_{\mathrm{air}}$ are qualitative. For the JRA, the values of q_{air} are unlikely to explain the

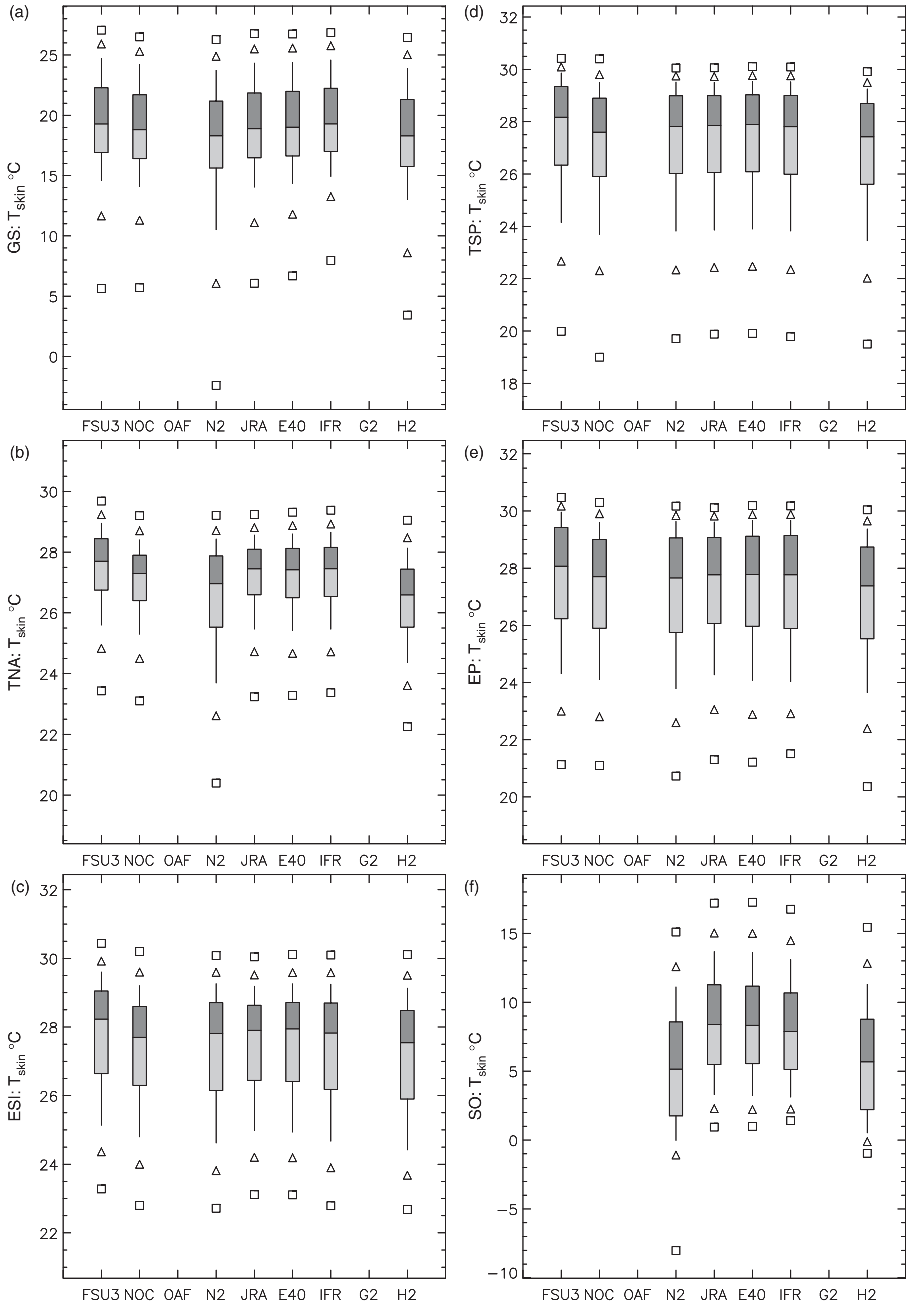

Figure 9. Same as Figure 8, except for sea surface temperature.

larger values of LHF; the greater momentum roughness length (larger value of Charnock's constant) contributes to this bias, but it is not clear if it is the dominant factor.

The FSU3 LHF distribution is shifted toward lower fluxes in the five regions it is available (Figure 13(a)–(e)); however, in the GS region (Figure 13(a)), the FSU distribution is similar to the IFREMER and HOAPS2. The specific humidity distribution for each of these products (Figure 10(a)) is shifted to higher values relative to the other five products (OAFLUX did not provide q_{air}), which might partially account for the lower LHF values due to a reduction in $q_{sfc} - q_{air}$. As is noted above, the FSU3q_{air} distribution is consistently shifted toward higher values (Figure 10(a)–(e)), with medians typically 0.5–1.0 g kg^{-1} higher than the other available products, which contributes to lower $q_{sfc} - q_{air}$ values resulting in lower LHF.

The spread of LHF distributions is relatively wide for the NOC, NCEPR2, and GSSTF2 in the tropical regions (Figure 13(b)–(e)). For example, the IQR in the TSP ranges from 25 Wm^{-2} for OAFLUX to 60 Wm^{-2} and 75 Wm^{-2} for the NOC and GSSTF2, respectively (Figure 13(d)). The three reanalysis products

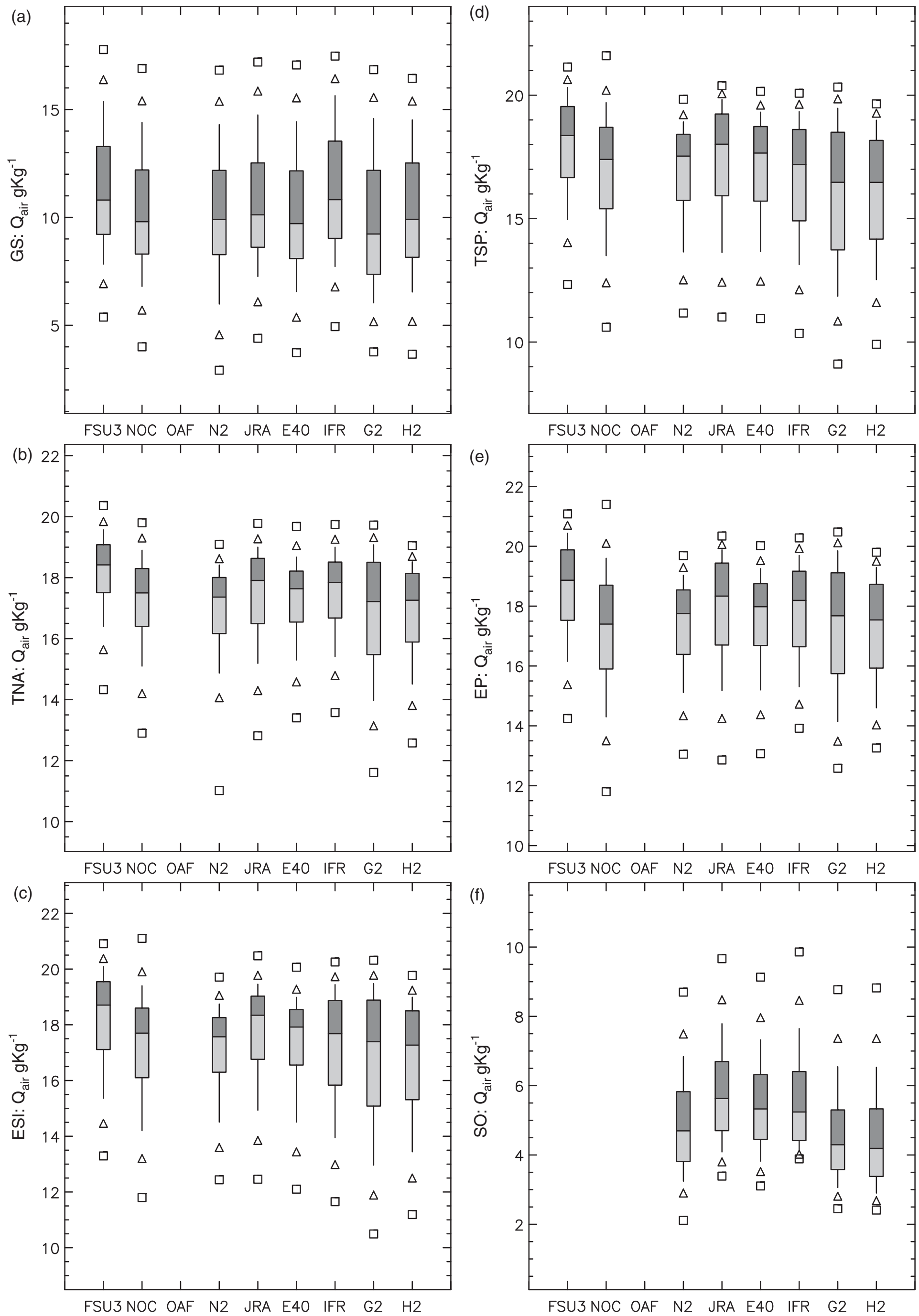

Figure 10. Same as Figure 8, except for specific humidity.

also exhibit differing IQR in the Tropics. The source of these differences is not readily apparent. In the TSP, the T_{skin} distributions for six of the products (Figure 9(d)) are nearly identical (this is expected since most use a version of Reynold's SST), but both the OAFLUX and GSSTF2 lack a T_{skin} product. In the TSP (as well as the TNA, ESI, and EP), the spread of q_{air} is largest for the GSSTF2 (Figure 10(b)–(e)) compared to the other products with q_{air}, which might partially explain the larger spread of LHF in these regions. The spread of the wind speed is largest for the NOC (not shown) in several tropical regions, which might partially explain the larger IQR in the NOC's LHF distribution.

LHF distributions in the SO (Figure 13(f)) deviate from the patterns found in other regions. In the SO, the GSSTF2 (NCEPR2, IFREMER, and HOAPS2) median and distribution are shifted toward higher (lower) values

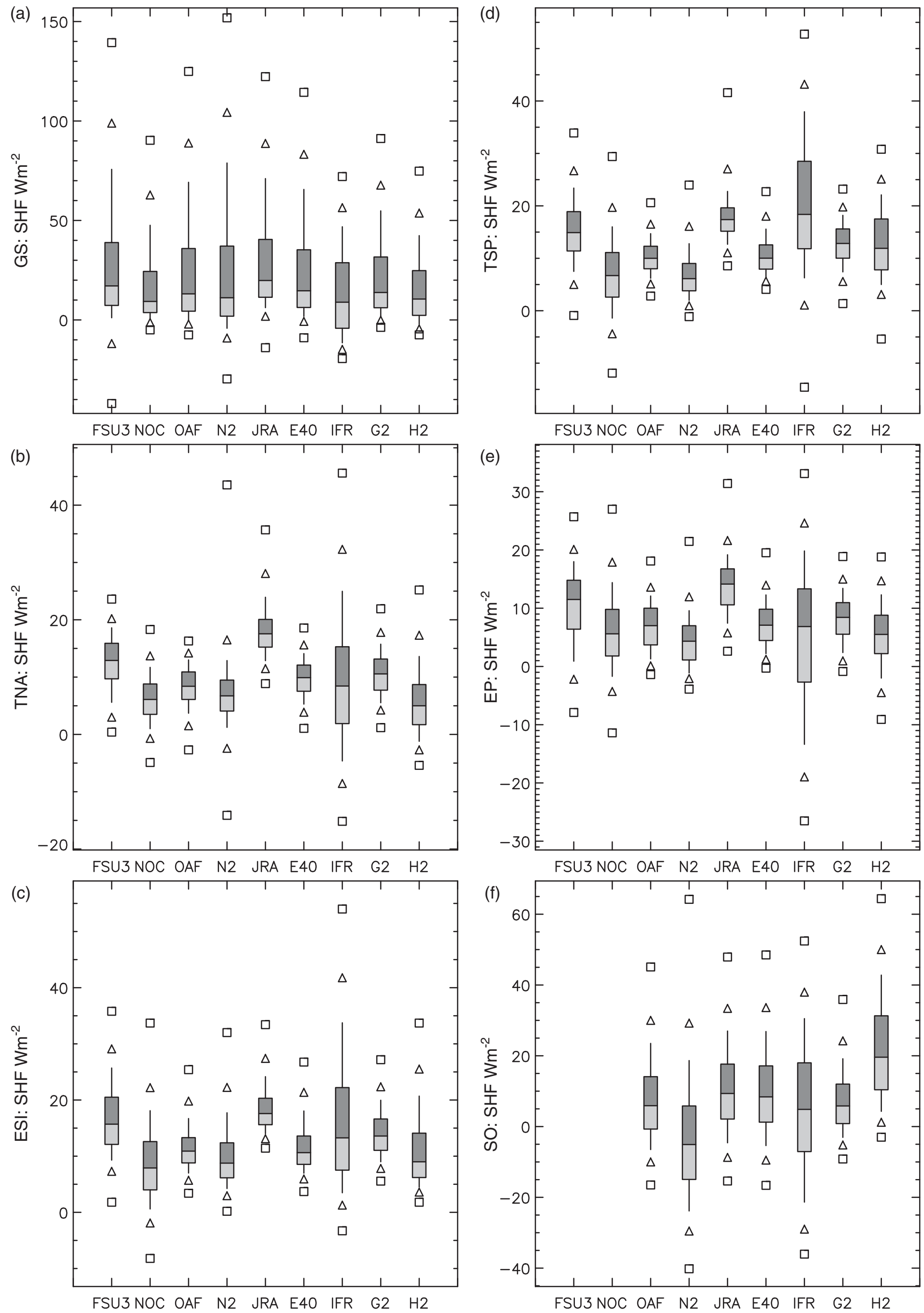

Figure 11. Same as Figure 8, except for sensible heat flux.

(Figure 13(f)). This shift is not completely explained by the $q_{\mathrm{sfc}} - q_{\mathrm{air}}$ relationship. The NCEPR2 q_{air} distribution is shifted toward low values but is similar to the GSSTF2 and the HOAPS2 distributions (Figure 10(f)). Both the NCEPR2 and HOAPS2 also have T_{skin} distributions that are shifted toward lower T_{skin} compared to other products (Figure 9(f)), which might partially explain the lower LHF for these products due to reduced $q_{\mathrm{sfc}} - q_{\mathrm{air}}$. The GSSTF2 q_{air} distribution is shifted toward lower values; consequently, the LHF distribution is shifted toward relatively higher values. The lack of T_{skin} for the GSSTF2 prevents even a qualitative assessment of the role low q_{air} might play in the high LHF.

4. Application example: ENSO variability

We have shown the global and regional variations between the nine turbulent flux products. The concern for many users is how much the differences between flux

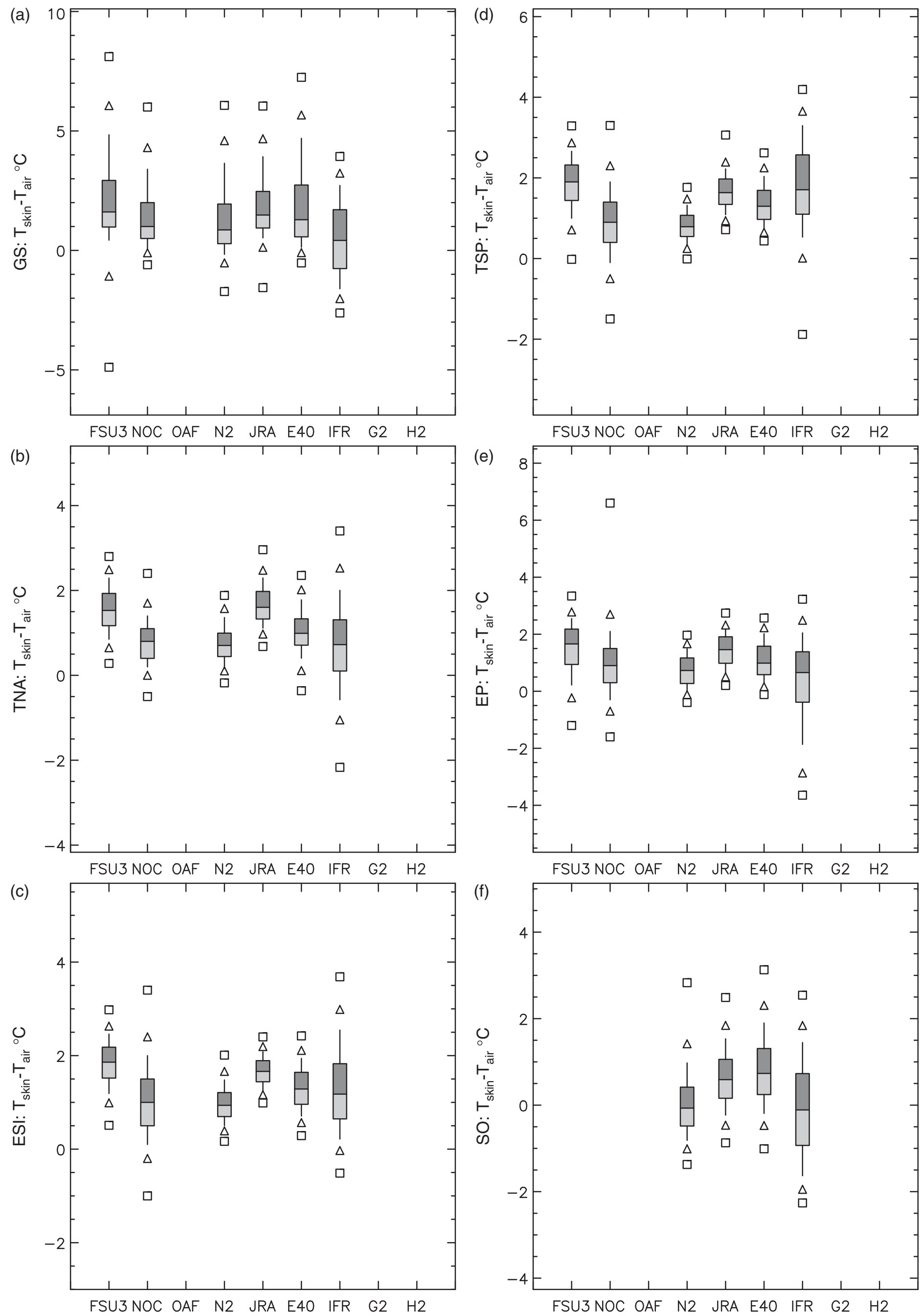

Figure 12. Same as Figure 8, except for sea minus air temperature (T_{air}) difference.

products might impact their specific research activities. Herein we provide an example focusing on the tropical Pacific and El Niño–Southern Oscillation (ENSO). ENSO is known to influence both atmospheric and oceanic conditions over the globe (Glantz, 1996; Smith *et al*., 1998). The interaction between the atmosphere and ocean in the ENSO region occurs through the air–sea fluxes. The nine available products exhibit large differences in the strength and location of the heat flux anomalies associated with warm (El Niño) and cold (La Niña) ENSO phases during the period 1978–2004.

Anomalies of the heat fluxes and forcing variables are examined using Hovmüller diagrams in the ENSO region of the tropical Pacific (5°N–5°S, 180°–80°W). Each product's anomaly is computed by subtracting that product's monthly mean for the period 1993–2000 from the individual monthly values. No spatial or temporal smoothing was applied to the Hovmüller diagrams: some products (particularly the NOC) exhibit relatively noisy

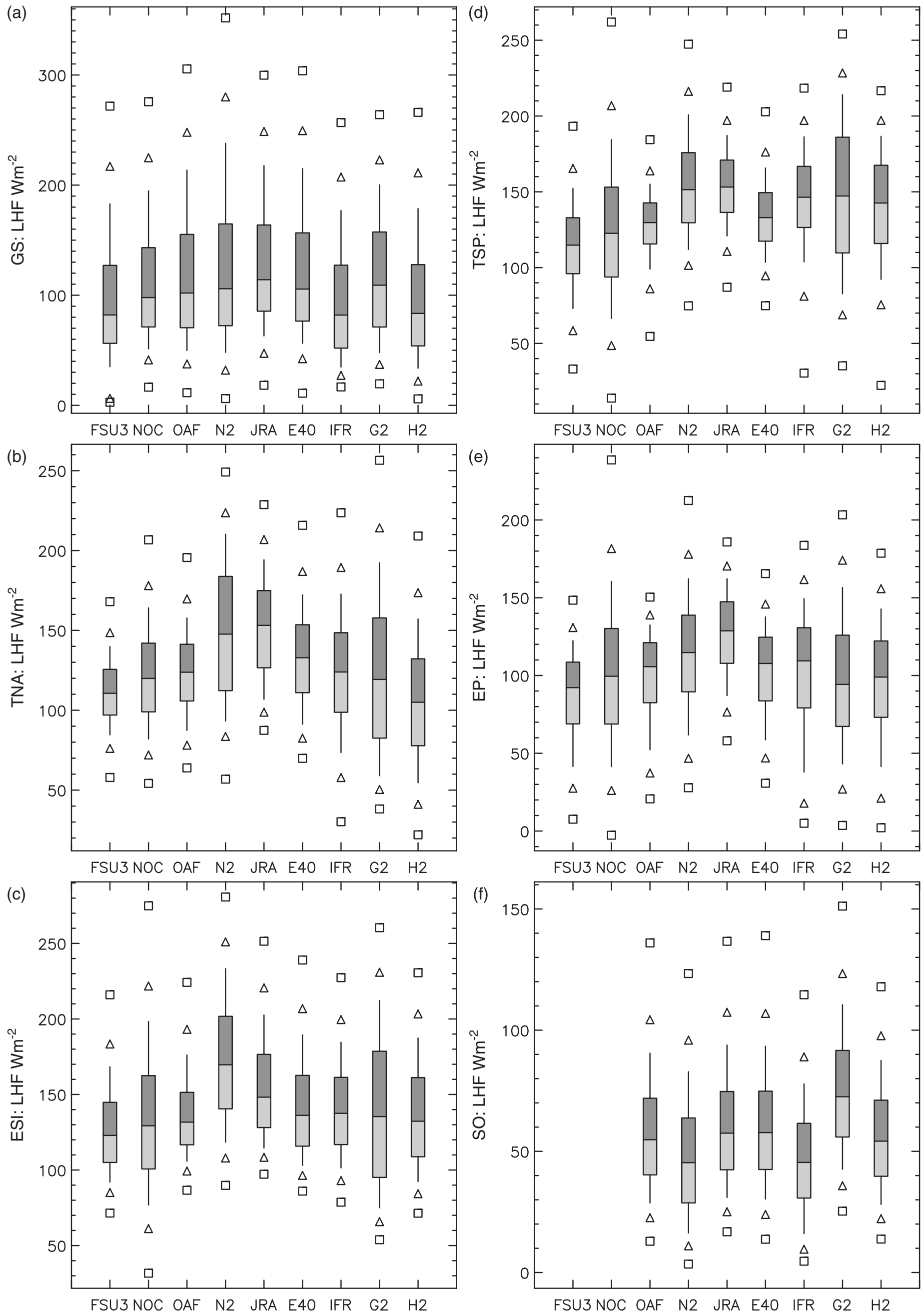

Figure 13. Same as Figure 8, except for latent heat flux.

patterns. The six products that provided T_{skin} (Table I) all had similar maximum and minimum values associated with the major El Niño (warm anomalies) and La Niña (cool anomalies) events during the period 1978–2006 (not shown). There was good agreement in the T_{skin} anomaly magnitudes between the reanalysis, satellite, and *in situ*-based products. Major El Niños occurred from 1982 to 1983 and 1997 to 1998, while La Niñas occurred from 1987 to 1988 and an extended period from 1999 to 2000.

LHF anomalies exhibit large differences between the products (Figure 14). During the El Niño events in 1982–1983 and 1997–1998, the LHF anomalies are greater than 40 Wm^{-2} around 100°–140 °W in all products, except the JRA. Overall the JRA has very weak signals of the El Niños (Figure 14(b)). The warm anomalies in the LHF during the 1997–1998 El Niño are much larger (>60 Wm^{-2}) and large anomalies extend farther to the east (to near 90 °W) in the two satellite products (HOAPS2 and IFREMER). This eastward extension is associated with these products having lower specific humidity anomalies than the non-satellite products 1.5–2.0 g kg^{-1} *vs* 2.5–3.0 g kg^{-1} in the same region during the El Niño events (Figure 15(e) and (f)). The

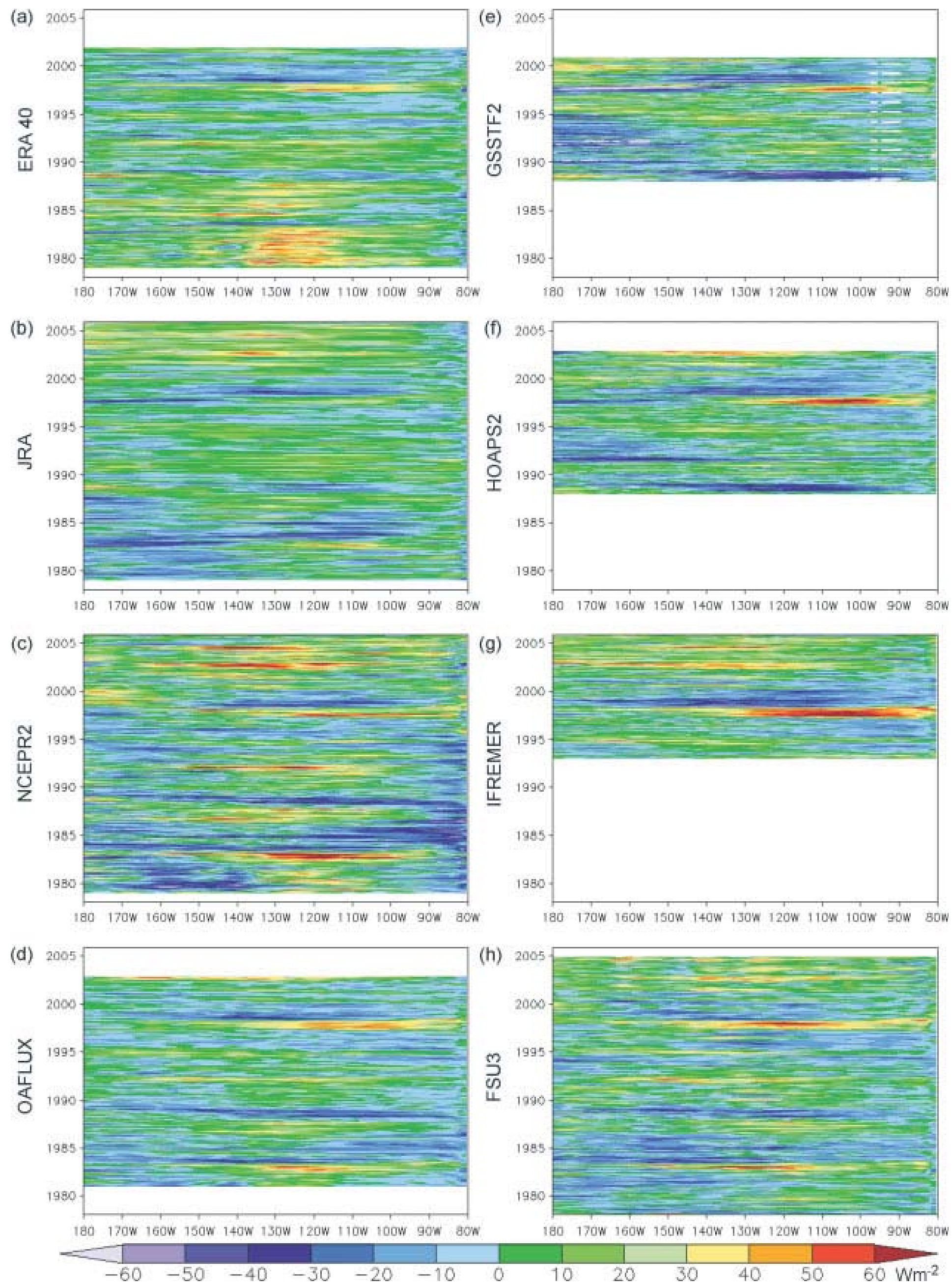

Figure 14. Meridional average equatorial (5°N–5°S) latent heat flux anomalies for (a) ERA-40, (b) JRA, (c) NCEPR2, (d) OAFLUX, (e) GSSTF2, (f) HOAPS2, (g) IFREMER, and (h) FSU3. Anomalies are computed whenever a product has a complete year of data from 1978 to 2005 by removing the overall monthly mean for the period 1992–2000. No zonal smoothing has been applied.

authors hypothesise that the satellite products have difficulty estimating near-surface humidity off the west coast of Peru due to the persistent stratus clouds in the region. The region has an incredibly sharp transition in q_{air} and T_{air} between the boundary layer and the free atmosphere (C. Fairall, personal communication, 2008). Sharp vertical gradients pose challenges for satellite retrievals of q_{air} and T_{air} (G. Wick, personal communication, 2008). This hypothesis is in agreement with Bourras (2006), who similarly noted the difficulty with satellite retrievals of q_{air} and the adverse impact on satellite LHF products in the Tropics. As a result, the examined satellite LHF products might not be as applicable to ENSO studies near the west coast of South America.

A similar situation occurs with the SHF. The SHF for the OAFLUX, GSSTF2, HOAPS, and reanalysis products (Figure 16) show consistent anomaly magnitudes and spatial patterns, with the FSU3 exhibiting larger amplitude in the pattern. The IFREMER is the outlier with SHF anomalies in the eastern Pacific (80°–120°W) reaching >12 Wm^{-2} during the 1997–1998 El Niño (Figure 16(f)). T_{air} anomalies for the IFREMER in this

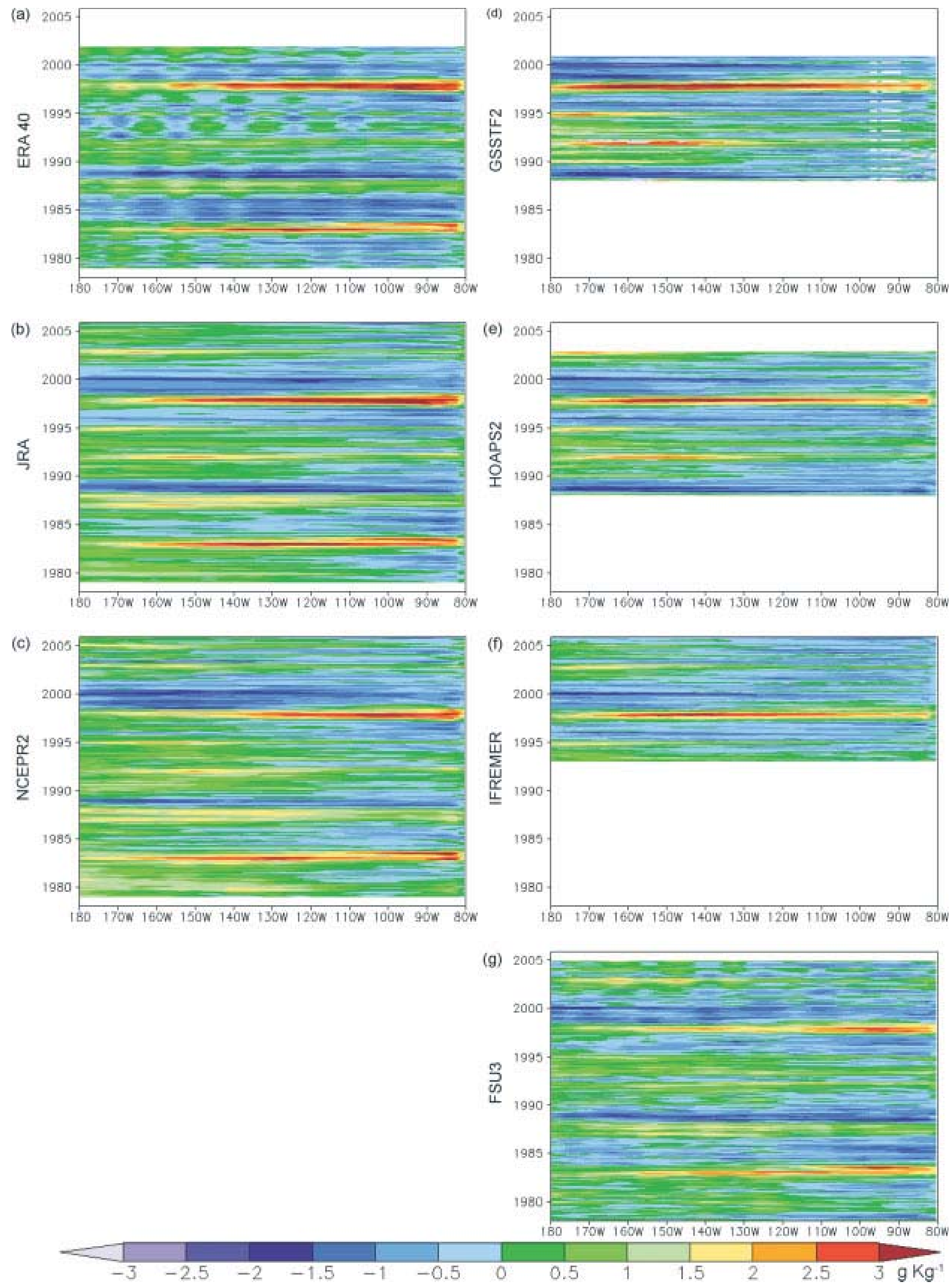

Figure 15. Meridional average equatorial (5°N–5°S) specific humidity anomalies for (a) ERA-40, (b) JRA, (c) NCEPR2, (d) GSSTF2, (e) HOAPS2, (f) IFREMER, and (g) FSU3. Anomalies are constructed using the same method as Figure 14.

region (Figure 17(d)) average 2 °C, while the reanalyses and the FSU3 show temperature anomalies >3 °C. As noted above, all products have similar T_{skin} anomalies, so the reduction in T_{air} for the IFREMER increases the $T_{\mathrm{skin}} - T_{\mathrm{air}}$, resulting in higher SHF. The low T_{air} in the IFREMER might again be associated with difficulties in retrieving T_{air} in the stratus cloud region off Peru.

Other noteworthy features in the tropical Pacific are the negative LHF and SHF anomalies between 80 °W and 90 °W in the NCEPR2 prior to 1997 (Figures 14(c) and 16(c)) and the non-physical maxima or minima in q_{air} oriented along longitude lines in the ERA-40 (Figure 15(a)). Although the source of the SHF and LHF anomalies along the South American coast cannot be verified, they may be the result of a change in satellite observations being assimilated into NCEPR2. These negative heat flux anomalies are absent in all other products with the exception of the OAFLUX, which uses the NCEPR as an input. As for the anomaly pattern in the ERA-40 q_{air}, the authors suspect that it is related to assimilation of data from the TAO mooring array.

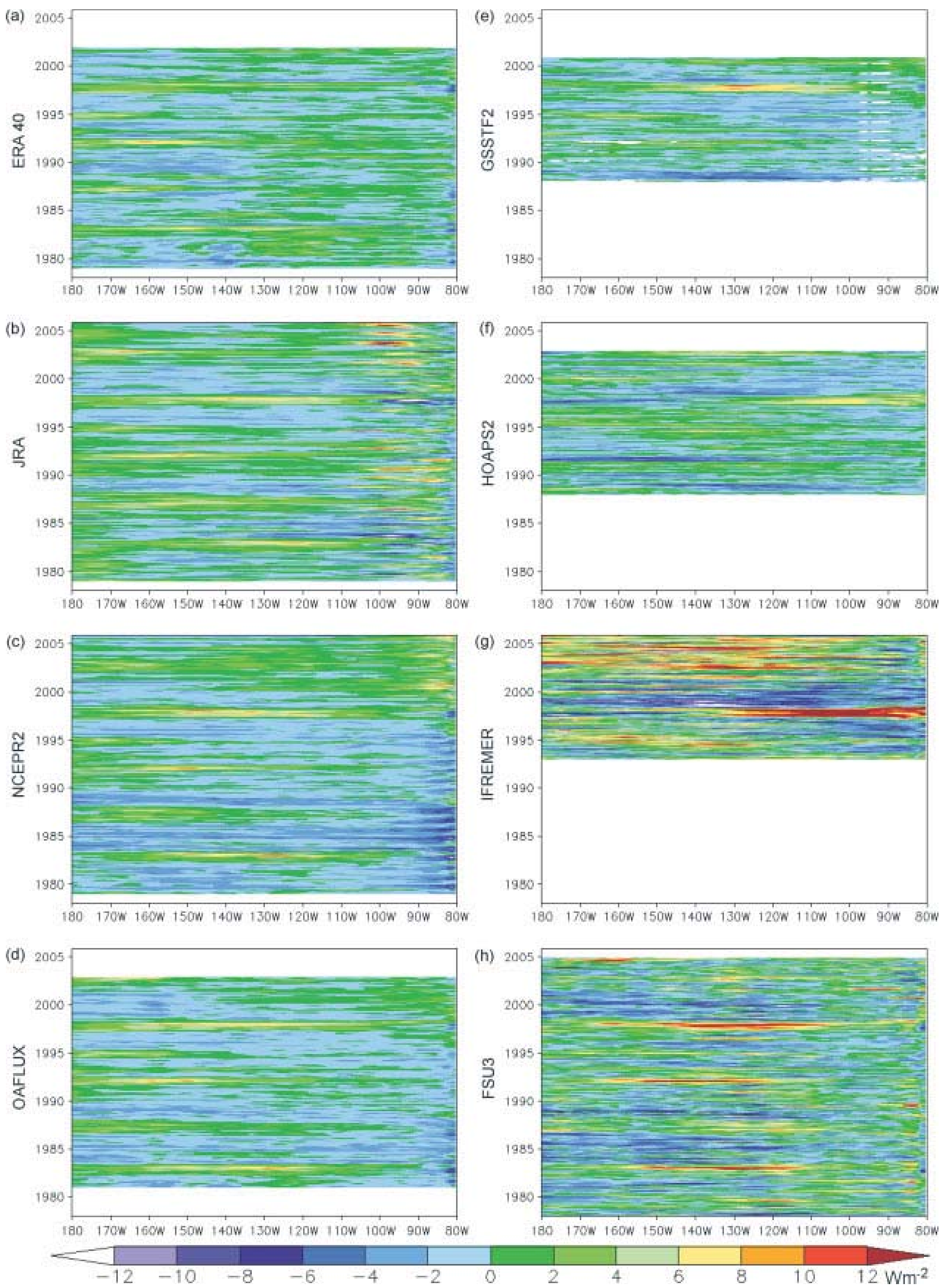

Figure 16. Same as Figure 14, except for sensible heat flux anomalies.

5. Conclusions and recommendations

The comparison of nine monthly turbulent air–sea flux products revealed substantial differences in both the heat fluxes and the forcing variables used to estimate the fluxes. The differences can be attributed to a combination of spatial and temporal sampling variations, the averaging methodology used to arrive at a monthly mean, the use of different flux algorithms, and the difficulties in estimating the forcing fields (e.g. bias corrections and height adjustments). In many regions, the differences in T_{air} and q_{air} between the products clearly had a greater impact than the discrepancies in wind speed or T_{skin} on the derived heat fluxes. The q_{air} for the FSU3 tended to be shifted toward higher values as compared to the other product, resulting in smaller $q_{sfc} - q_{air}$ and lower LHF. The recent reprocessing of the NOC fluxes (E. Kent, personal communication, 2008) indicates that their revised q_{air} values will exhibit means similar to the FSU3, providing the authors some confidence that the FSU3 q_{air} and subsequent LHF might be outliers for physically sound reasons. In addition, satellite retrievals of q_{air} and T_{air} continue to be problematic in regions with sharp vertical gradients. For those developing flux products,

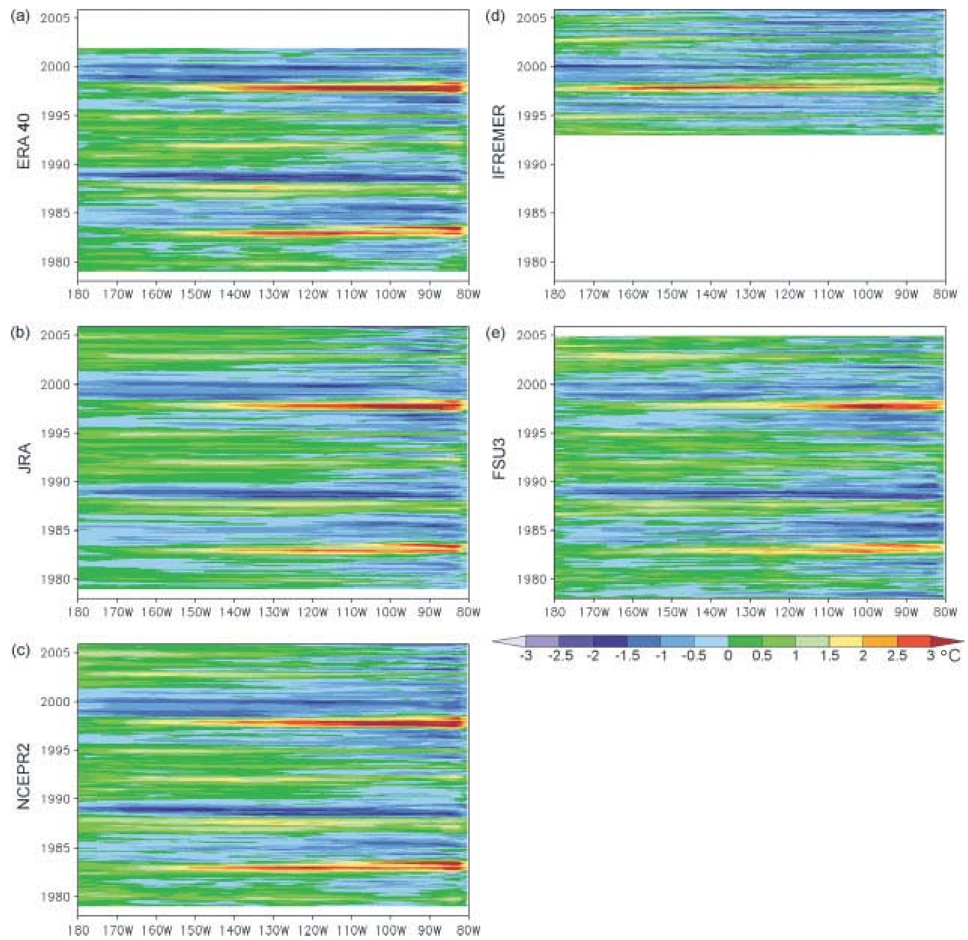

Figure 17. Meridional average equatorial (5°N–5°S) air temperature anomalies for (a) ERA-40, (b) JRA, (c) NCEPR2, (d) IFREMER, and (e) FSU3. Anomalies are constructed using the same method as Figure 14.

the authors recommend further research into improving and validating global q_{air} and T_{air} fields. The authors encourage developers to work with the SEAFLUX project and the World Climate Research Programme (WCRP) working group on surface fluxes to identify the strengths and weaknesses of the multiple algorithms currently in use, and to help develop a consensus flux algorithm. This will greatly aid the evaluation of future flux products by removing the differences inherited by the eight flux algorithms used in the studied products.

The regional analysis of the nine flux products quickly revealed that no one product is ideally suited for every application. In addition, flux products are continually undergoing improvement and new updates are released on an irregular schedule. Users should not choose a flux product based only on its ease of use or availability, without understanding whether or not the product will adequately resolve the features of interest in their research endeavor. The example for the tropical Pacific Ocean showed that if one was interested in the interannual variability of the air–sea fluxes off the coast of Peru (e.g. for ENSO variability studies), the satellite-based flux products examined herein would not be a good choice due to limitations in resolving the fluxes in the stratus cloud region. The ERA-40 is ill suited for the tropical Pacific, due to the spurious variability associated with the assimilation of the TAO buoys. As another example, if your interest was the seasonal variations associated with the western boundary currents or the track of atmospheric cyclones, you might not choose the FSU3 since the averaging approach reduces the influence of daily synoptic variability in these regions. This variability is better captured in products that derive daily fluxes and then average these daily values to determine the monthly mean. The authors hope that the details presented in this manuscript will help users of monthly turbulent flux products to select a product that is most appropriate to achieve their research objectives.

Acknowledgements

The NOAA Climate Observation Division's Office of Climate Observations (NA06OAR4310070) largely supported this analysis. The Physical Oceanography section of the National Science Foundation provided additional support (OCE0136934). The authors thank A. Beljaars, A. Bentamy, C. Fairall, M. Kanamitsu, E. Kent, C. Klepp, K. Onagi, G. Wick, and L. Yu for their input to this manuscript.

References

Andersson E, Bauer P, Beljaars A, Chevallier F, Hólm E, Janisková M, Kállberg P, Kelly G, Lopez P, McNally A, Moreau E, Simmons A, Thépaut J-N, Tompkins A. 2004. Assimilation and modeling of the hydrological cycle in the ECMWF forecasting system. *Bulletin of American Meteorological Society* **86**: 387–402.

Beljaars ACM. 1995. The parametrization of surface fluxes in large-scale models under free convection. *Quarterly Journal of the Royal Meteorological Society* **121**: 255–270.

Bentamy A, Katsaros KB, Alberto M, Drennan WM, Forde EB, Roquet H. 2003. Satellite estimates of wind speed and latent heat flux over the global oceans. *Journal of Climate* **16**: 637–656.

Berry DI, Kent EC. 2009. A new air-sea interaction gridded dataset from ICOADS with uncertainty estimates. *Bulletin of American Meteorological Society* **90**: 645–656, DOI: 10.1175/2008BAMS2639.1.

Bony S, Sud Y, Lau KM, Susskind J, Saha S. 1997. Comparison and satellite assessment of NASA/DAO and NCEP-NCAR reanalyses over tropical ocean: atmospheric hydrology and radiation. *Journal of Climate* **10**: 1441–1462.

Bourras D. 2006. Comparison of five satellite-derived latent heat flux products to moored buoy data. *Journal of Climate* **19**: 6291–6313.

Bourassa MA, Romero R, Smith SR, O'Brien JJ. 2005. A new FSU winds climatology. *Journal of Climate* **18**: 3686–3698.

Brunke MA, Zeng X, Anderson S. 2002. Uncertainties in sea surface turbulent flux algorithms and data sets. *Journal of Geophysical Research* **107**: 3142, DOI: 10.1029/2001JC000992.

Charnock H. 1955. Wind stress on a water surface. *Quarterly Journal of the Royal Meteorological Society* **81**: 639–640.

Chou S-H. 1993. A comparison of airborne eddy correlation and bulk aerodynamic methods for ocean-air turbulent fluxes during cold-air outbreaks. *Boundary-Layer Meteorology* **64**: 75–100.

Chou S-H, Nelkin E, Ardizzone J, Atlas RM, Shie C-L. 2003. Surface turbulent heat and momentum fluxes over global oceans based on the Goddard satellite retrievals, version 2 (GSSTF2). *Journal of Climate* **16**: 3256–3273.

Curry JA, Bentamy A, Bourassa MA, Bourras D, Brunke M, Castro S, Chou S, Clayson CA, Emery WJ, Eymard L, Fairall CW, Kubota M, Lin B, Perrie W, Reeder RR, Renfrew IA, Rossow WB, Schultz J, Smith SR, Webster PJ, Wick GA, Zeng X. 2004. SEAFLUX. *Bulletin of American Meteorological Society* **85**: 409–424.

Da Silva A, Young CC, Levitus S. 1994. *Algorithms and Procedures, Vol. 1, Atlas of Surface Marine Data 1994*. NOAA Atlas NESDIS 6. US Government Printing Office: Washington DC, 83 pp.

Dyer AJ. 1974. A review of flux-profile relations. *Boundary-Layer Meteorology* **7**: 363–372.

Esbensen SK, McPhaden MJ. 1996. Enhancement of tropical ocean evaporation and sensible heat flux by atmospheric mesoscale systems. *Journal of Climate* **9**: 2307–2325.

Esbensen SK, Reynolds RW. 1981. Estimating monthly averaged air-sea transfers of heat and momentum using the bulk aerodynamic method. *Journal of Physical Oceanography* **11**: 457–465.

Fairall CW, Bradley EF, Hare JE, Grachev A, Edson JB. 2003. Bulk parameterization of air-sea fluxes : updates and verification for the COARE algorithm. *Journal of Climate* **16**: 571–591.

Fairall CW, Bradley EF, Rogers DP, Edson JB, Young GS. 1996. Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment. *Journal of Geophysical Research* **101**: 3747–3764.

Glantz MH. 1996. *Currents of Change: El Niño's Impact on Climate and Society*, Cambridge University Press: Cambridge, UK; 194 pp.

Grassl H, Jost V, Schultz J, Ramesh Kumar MR, Bauer P, Schluessel P. 2000. The Hamburg Ocean-Atmosphere Parameters and Fluxes from Satellite Data (HOAPS): A Climatological Atlas of Satellite-Derived Air-Sea Interaction Parameters over the World Oceans. Report 312, ISSN 0937–1060. Max Plank Institute for Meteorology: Hamburg, 132 pp. Available from http://www.hoaps.zmaw.de [accessed May 1 2009].

Gulev SK. 1994. Influence of space-time averaging on the ocean-atmosphere exchange estimates in the north atlantic midlatitudes. *Journal of Physical Oceanography* **24**: 1236–1255.

Gulev SK. 1997. climatologically significant effects of space-time averaging in the north atlantic sea-air heat flux fields. *Journal of Climate* **10**: 2743–2763.

Gulev SK, Jung T, Ruprecht E. 2007a. Estimation of the impact of sampling errors in the VOS observations on air-sea fluxes. Part 1: uncertainties in climate means. *Journal of Climate* **20**: 279–301.

Gulev SK, Jung T, Ruprecht E. 2007b. Estimation of the impact of sampling errors in the VOS observations on air-sea fluxes. Part II: impact on trends and interannual variability. *Journal of Climate* **20**: 302–315.

Hicks BB. 1976. Wind profile relationships from the 'Wangara' experiment. *Quarterly Journal of the Royal Meteorological Society* **102**: 535–551.

Holtslag AAM, De Bruin HAR. 1988. Applied modelling of the night-time surface energy balance over land. *Journal of Applied Meteorology* **27**: 689–704.

Janssen PAEM. 1989. Wave-induced stress and the drag of airflow over sea waves. *Journal of Physical Oceanography* **19**: 745–754.

Josey SA, Kent EC, Taylor PK. 1995. Seasonal variations between sampling and classical mean turbulent heat flux estimates in the eastern North Atlantic. *Annales Geophysicae* **13**: 1054–1064.

Josey SA, Kent EC, Taylor PK. 1998. The Southampton Oceanography Centre (SOC) Ocean-Atmosphere Heat, Momentum and Freshwater Flux Atlas. James Rennell Division, Southampton Oceanography Centre, Report 6, Southampton, UK, 59 pp. Available from http://www.noc.soton.ac.uk/ooc/REFERENCES/PREPRINTS/SOC_flux_atlas.pdf [accessed May 1 2009].

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woolen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetma A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.

Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II reanalysis (R-2). *Bulletin of the American Meteorological Society* **83**: 1631–1643.

Kubota M, Kano A, Muramatsu H, Tomita H. 2003. Intercomparison of various surface latent heat flux fields. *Journal of Climate* **16**: 670–678.

Long PE Jr. 1986. *An Economical and Compatible Scheme for Parameterizing the Stable Surface Layer in the Medium-Range Forecast Model*. Office Note 321, US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, National Meteorological Center. 24 p.

Long PE Jr. 1990. *Derivation and Suggested Method of Application of Simplified Relations for Surface Fluxes in the Medium-Range Forecast Model: Unstable Case*. Office Note 356, US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, National Meteorological Center. 53 p.

Louis J-F, Tiedtke M, Geleyn J-F. 1982. *A short history of the PBL parameterization at ECMWF*. Workshop on planetary boundary layer parameterization, ECMWF: Reading, England.

NGDC. 1988. *Data Announcement 88-MGG-02, Digital relief of the Surface of the Earth*, NOAA, National Geophysical Data Center: Boulder, Colorado.

Nickerson EC, Smiley VE. 1975. Surface layer and energy budget parameterizations for mesoscale models. *Journal of Applied Meteorology* **14**: 297–300.

Onogi K, Tsutsui J, Koide H, Sakamoto M, Kobayashi S, Hatsushika H, Matsumoto T, Yamazaki N, Kamahori H, Takahashi K, Kadokura S, Wada K, Kato K, Oyama R, Ose T, Mannoji N, Taira R. 2007. The JRA-25 reanalysis. *Journal of the Meteorological Society of Japan* **85**: 369–432.

Paulson CA. 1970. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *Journal of Applied Meteorology* **9**: 857–861.

Pegion PJ, Bourassa MA, Legler DM, O'Brien JJ. 2000. Objectively derived daily "winds" from satellite scatterometer data. *Monthly Weather Review* **128**: 3150–3168.

Putman WM, Legler DM, O'Brien JJ. 2000. Interannual variability of synthesized FSU and NCEP-NCAR reanalysis pseudostress products over the Pacific Ocean. *Journal of Climate* **13**: 3003–3016.

Risien CM, Chelton DB. 2008. A global climatology of surface wind and wind stress fields from eight years of QuikSCAT scatterometer data. *Journal of Physical Oceanography* **38**: 2379–2413, DOI: 10.1175/2008JPO3881.1.

Schultz J, Meywerk J, Ewald S, Schluessel P. 1997. Evaluation of satellite-derived latent heat fluxes. *Journal of Climate* **10**: 2782–2795.

Schultz J, Schluessel P, Grassl H. 1993. Water vapor in the atmospheric boundary layer over oceans from SSM/I measurements. *The International Journal of Remote Sensing* **14**: 2773–2789.

Simmonds I, Dix M. 1989. The use of mean atmospheric parameters in the calculation of modeled mean surface heat fluxes over the world's oceans. *Journal of Physical Oceanography* **19**: 205–215.

Smith SD. 1980. Wind stress and heat flux over the ocean in gale force winds. *Journal of Physical Oceanography* **10**: 709–726, DOI: 10.1175/1520-0485(1980)010<0709:WSAHFO>2.0.CO;2.

Smith SD. 1988. Coefficients for sea surface wind stress, heat flux, and wind profiles as a function of wind speed and temperature. *Journal of Geophysical Research* **93**((C12)): 15467–15472, DOI: 10.1029/JC093iC12p15467.

Smith SR, Green PM, Leonardi AP, O'Brien JJ. 1998. Role of multiple-level tropospheric circulations in forcing ENSO winter precipitation anomalies. *Monthly Weather Review* **126**: 3102–3116.

Smith SR, Legler DM, Verzone KV. 2001. Quantifying uncertainties in NCEP reanalyses using high quality research vessel observations. *Journal of Climate* **14**: 4062–4072.

Thomas BR, Kent EC, Swail VR, Berry DI. 2008. Trends in ship wind speeds adjusted for observation method and height. *International Journal of Climatology* **28**: 747–763.

Uppala SM, Kållberg PW, Simmons AJ. and co-authors. 2005. The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.* **131**: 2961–3012.

Verschell MA, Bourassa MA, Weissman DE, O'Brien JJ. 1999. Model validation of the NASA Scatterometer winds. *Journal of Geophysical Research* **104**: 11359–11374.

Wentz FJ. 1997. A well-calibrated ocean algorithm for SSM/I. *Journal of Geophysical Research* **102**((C4)): 8703–8718.

Woodruff SD, Lubker SJ, Wolter K, Worley SJ, Elms JD. 1993. Comprehensive ocean-atmosphere dataset (COADS) release 1a: 1980–1992. *Earth System Monitor* **4**: 4–8.

Woodruff SK, Slutz RJ, Jenne RL, Steurer PM. 1987. A comprehensive ocean-atmosphere dataset. *Bulletin of the American Meteorological Society* **68**: 1239–1250.

Worley SJ, Woodruff SD, Reynolds RW, Lubker SJ, Lott N. 2005. ICOADS release 2.1 data and products. *International Journal of Climatology* **25**: 823–842.

Yelland MJ, Pascal RW, Taylor PK, Moat BI. 2009. AutoFlux: an autonomous system for the direct measurement of the air-sea fluxes of CO2, heat and momentum. *Journal of Operational Oceanography* **2**: 15–23.

Yu L, Weller RA. 2007. Objectively analyzed air-sea heat fluxes for the global ice-free oceans (1981–2005). *Bulletin of the American Meteorological Society* **88**: 527–539.

Zhang GJ. 1995. Use of monthly mean data to compute surface turbulent fluxes in the tropical Pacific. *Journal of Climate* **8**: 3084–3090.

