A Comparison of Atmospheric Reanalysis Surface Products over the Ocean and Implications for Uncertainties in Air–Sea Boundary Forcing

AYAN H. CHAUDHURI AND RUI M. PONTE

Atmospheric and Environmental Research, Lexington, Massachusetts

GAEL FORGET AND PATRICK HEIMBACH

Massachusetts Institute of Technology, Cambridge, Massachusetts

(Manuscript received 9 February 2012, in final form 14 June 2012)

ABSTRACT

This paper investigates the uncertainties related to atmospheric fields from reanalysis products used in forcing ocean models. Four reanalysis products, namely from 1) the interim ECMWF Re-Analysis (ERA-Interim), 2) version 2 of the Common Reference Ocean–Ice Experiments (CORE2), 3) the 25-Year Japanese Reanalysis Project (JRA-25), and 4) NCEP–NCAR, are evaluated against satellite-derived observations for eight different fields (zonal and meridional winds, precipitation, specific humidity, continental discharge, surface air temperature, and downwelling longwave and shortwave radiation fluxes). No single product is found to agree better in *all* fields with satellite-derived observations. Reanalysis products are mostly comparable to each other because of their similar physical assumptions and assimilation of common observations. Adjusted atmospheric fields from the Estimating the Circulation and Climate of the Ocean (ECCO) optimizations are also in agreement with other reanalysis products. Time-mean and time-variable errors are estimated separately and mapped globally in space, based on 14-day average fields to focus on monthly to interannual periods. Time-variable errors are larger in comparison to the signal than time-mean errors for most fields, thus justifying the need to separate them for studying uncertainties as well as formulating optimization procedures. Precipitation and wind stress fields show significant time-mean and time-variable errors whereas downwelling radiation, air temperature, and humidity fields show small time-mean errors but large time-variable errors, particularly in the tropics. Uncertainties based on evaluating multiple products presented here are considerably larger than uncertainties based on single product pairs.

1. Introduction

Investigations of the ocean state are limited by the paucity of observations in space and time. These limitations are in part mitigated by ocean general circulation models (OGCMs), which aid in interpreting ocean observations, deciphering ocean dynamics, and assessing contributions of different processes (McWilliams 1996). Moreover, models are one of the few available frameworks for making projections to past and future ocean conditions. To force OGCMs, surface atmospheric state variables are used to infer air–sea fluxes via boundary layer schemes. These atmospheric conditions are usually derived from reanalysis products (Kalnay et al. 1996; Uppala et al. 2005; Onogi et al. 2005; Dee et al. 2011) that assimilate meteorological observations (satellite and in situ) into numerical weather prediction models.

Surface fluxes of momentum, heat, and freshwater can be applied to OGCMs by different techniques. A common OGCM boundary condition is to prescribe wind stress for the momentum equation, net heat flux for the thermal equation, and net (virtual) salt flux, corresponding to the exchange of freshwater associated with evaporation, precipitation, and runoff for the salinity equation (Cox and Bryan 1984). However, uncertain measurements and complex cancelling processes such as evaporation, precipitation, ice formation and melting, and river runoff can create large discrepancies in the salinity balance. The heat balance is similarly limited by uncertain observations in turbulent (sensible and latent) and radiative (shortwave and longwave) heat fluxes.

Corresponding author address: Ayan H. Chaudhuri, Atmospheric and Environmental Research, 131 Hartwell Ave., Lexington, MA 02421.
E-mail: achaudhu@aer.com

DOI: 10.1175/JCLI-D-12-00090.1

These imbalances are usually overcome by restoring both temperature and salinity surface values to observed climatology (Haney 1971; Han 1984), at time scales ranging from days to months (e.g., Cox and Bryan 1984; Danabasoglu and McWilliams 1995). For simulations concerning natural variability dynamics, alternative mixed conditions are employed for tracers in OGCMs. In such cases temperature is restored and a salinity flux is specified by diagnosing a preliminary solution with fully restoring conditions (e.g., Bryan 1986).

A drawback to restoration is that the resulting surface tracer values do not differ largely from climatology. Furthermore, such formulations exclude the possibility of having nonzero flux coincident with correct model surface tracer fields (Large et al. 1997). An alternative approach discussed by Large and Pond (1981) uses bulk aerodynamic formulas to calculate surface flux boundary conditions. The bulk air–sea flux formulation derives momentum from surface wind speed observations at 10 m (U_{10}). The surface thermal flux has four components. 1) First, specific turbulent heat flux is estimated from U_{10}, air temperature at 2 m (or 10 m), and model sea surface temperature (SST). 2) Latent turbulent heat flux, which is a function of evaporation, is obtained from specific humidity at 2 m (or 10 m), U_{10}, and SST. 3) Shortwave radiative heat flux is estimated from albedo and downwelling surface shortwave radiation, and finally 4) longwave radiative heat flux is calculated as a function of surface downwelling longwave radiation and SST. The net surface freshwater flux is the difference between evaporation and precipitation with the addition of freshwater continental discharge. Variations of the bulk flux formulation can also be used; for example, wind stresses could be directly prescribed to the momentum equation (instead of U_{10}), while thermal and freshwater surface boundary conditions could be prescribed in bulk manner (Stammer et al. 2002; Wunsch and Heimbach 2006). Because of nonlinearity in the bulk formulas, Large et al. (1997) recommend the use of atmospheric fields sampled at high frequencies.

Atmospheric reanalyses use numerical models to synthesize observations distributed irregularly in space and time into regularly gridded meteorological dataset at high frequencies. The commonly available reanalysis products such as the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR; Kistler et al. 2001), European Centre for Medium-Range Weather Forecasts (ECMWF; Uppala et al. 2005) and interim ECMWF Re-Analysis (ERA-Interim; Dee et al. 2011) products, and the Japanese 25-Year Re-Analysis (JRA-25; Onogi et al. 2005) have stimulated ocean modeling studies not otherwise possible. However, a combination of imperfect models and data result in individual biases in these reanalysis fields. Several studies show significant biases in tropical wind stress and its divergence (Milliff et al. 1999), near-surface humidity (Sun et al. 2003), polar temperatures (Drobot et al. 2006), bulk flux algorithms (Brunke et al. 2003), precipitation (Smith et al. 2001; Nicolas and Bromwich 2011), clouds in the South Pacific convergence zone (Wang and McPhaden 2001), and the sea surface albedo (Zhang et al. 1995). These biases have been corrected in several ways. Large and Yeager (2004) created a new dataset called the Common Ocean-Ice Reference Experiment (CORE) that applies corrections to original NCEP reanalysis fields by adjusting them against a variety of satellite-based and in situ derived radiation, SST, sea ice concentration, and precipitation products. Similarly, Brodeau et al. (2010) calibrate the air–sea fluxes of the 40-yr ECMWF Re-Analysis (ERA-40) by applying corrections to the reanalysis fields from satellite-derived and in situ-derived datasets. Both Large and Yeager (2004) and Brodeau et al. (2010) verify that the corrections are consistent with a near-zero global imbalance of heat and freshwater fluxes.

An alternative method of correcting atmospheric reanalysis fields has been suggested by Stammer et al. (2004). Their methodology, developed by the ''Estimating the Circulation and Climate of the Ocean'' (ECCO) Consortium, involves synthesizing a large variety of oceanic observations over long periods to produce a dynamically consistent state of the ocean. Part of this synthesis involves adjusting the air–sea fluxes within error limits to make the model consistent with available ocean data. Stammer et al. (2004) report that adjustments to momentum, heat, and freshwater fluxes from the NCEP reanalysis show good correspondence with the CORE dataset.

While significant efforts have been made toward correcting biases in reanalysis fields, quantitative uncertainty estimates of the atmospheric state are very limited. Lucas et al. (2008) present ERA-40 uncertainties for the thermal boundary condition in the North Atlantic by taking the difference between ERA-40 fields and their corresponding CORE field. They note that their estimates do not represent the full range of errors in the ERA-40 fields as the models driving the ECMWF and CORE reanalysis use similar physical assumptions and both systems assimilate common observations as boundary conditions. Similarly, Leeuwenburgh (2005) estimates differences between ERA-40 and the 15-yr ECMWF Re-Analysis (ERA-15) for the tropical Pacific over a 2-yr period to obtain a statistical description of error characteristics in zonal and meridional surface stress, air temperature, dewpoint temperature, and shortwave radiation. Brunke et al. (2011) have estimated uncertainties

in surface turbulent fluxes by comparing reanalysis products against cruise-based fluxes. While these measurements are sparse in space and time, they provide the most accurate values for error estimation. Future advances in ocean modeling and ocean forecasting will undoubtedly benefit from global uncertainty estimates of atmospheric state fields (Brodeau et al. 2010).

In this study, our goal is to obtain best possible global ocean uncertainties in atmospheric fields used to force OGCMs. In particular, errors are derived for atmospheric forcing fields used in the ECCO modeling framework, which include wind stress, air temperature, specific humidity, continental discharge, precipitation, and downwelling longwave and shortwave radiation fluxes [surface pressure has been treated by Salstein et al. (2008)]. We also evaluate 10-m zonal and meridional wind speed, which are widely used by the ocean modeling community in bulk formulations and are important for both heat and freshwater fluxes. Rather than attempting to define the uncertainties in a particular atmospheric reanalysis product, we try to provide an upper bound on the expected uncertainties for each of the mentioned forcing fields by comparing not only several reanalyses but also some reference datasets. The data and procedures used to derive errors in the atmospheric parameters are described in section 2. Comparisons of reanalysis and ECCO-optimized atmospheric fields against satellite-based references are presented in section 3, based on 14-day averaged fields to focus on monthly and longer periods. A separate discussion of time-mean and time-variable errors is presented in section 4. Our main findings are summarized in section 5.

2. Data and methods

We consider atmospheric state fields from three reanalysis products (NCEP, ERA-Interim, and JRA-25), one reanalysis-derived product [version 2 of CORE (CORE2)], and one ECCO-optimized atmospheric solution for evaluation against corresponding satellite-derived reference fields. The five products are compared against each other and also against satellite-based data (Table 1) chosen as references because of their high temporal and spatial coverage globally. Further details about the reanalysis products and satellite-based references are presented in the appendix.

By definition, a difference between any two products provides an estimate of error contributed by both products, to the extent that their errors are uncorrelated. If a reanalysis product is compared against a weather buoy or a shipboard instrument (e.g., Brunke et al. 2011), and one can consider the in situ observation as reference (ignoring instrumentation errors), their difference would reflect the error in the reanalysis product solely, which can include a significant contribution from representation error (i.e., errors associated with physical processes that appear in observations but are not represented in the model; Lorenc 1986). Lack of in situ data in both time and space, or reference datasets more generally, limits our ability to perform such accurate analysis. We employ a more rudimentary approach in which combinations of differences are taken not only between reanalysis products and satellite-based references, but also among themselves. The largest differences among these combinations are chosen to represent the errors. This conservative approach attempts to encompass the entire range of forcing field uncertainties for products considered in this study, without accounting for the possibility of correlated errors. By taking the largest differences we essentially search for the maximum uncertainties possible across all products, without focusing on any specific one. In the case of uncorrelated errors, these estimates would provide an upper bound on uncertainty for each variable analyzed. The use of different products, including reference datasets, can also partly mitigate the possibility of underestimation of errors because of correlated errors in any single two-product comparison.

The reanalysis products have different spatial and temporal resolutions. ERA-Interim and ECCO-adjusted ERA-Interim have the highest spatial resolution (~70 km), followed by JRA-25 (~120 km). The NCEP and CORE2 fields have the coarsest resolution (~180 km). All the reanalysis fields are regridded to the coarsest grid (i.e., NCEP) to minimize interpolation errors. The NCEP-grid masking procedure is used to identify land points, which are ignored from the analysis. The reanalysis fields,

TABLE 1. Analysis periods for different atmospheric parameters shown above are largely constrained by the availability of satellite-based data. QuikSCAT data are level 2B multi-algorithm ocean wind stress components in a 25.0-km swath grid. HOAPS is a multisatellite product consisting of all available SSM/I instruments and National Oceanographic Data Center (NODC)–Rosenstiel School of Marine and Atmospheric Science (RSMAS) Pathfinder SST data on a 0.5° global grid. ISSCP uses a 280-km equal-area grid, with 3-hourly global coverage.

Variable	Reference	Time period
U, wind stress	QuikSCAT	1999–2006
V, wind stress	QuikSCAT	1999–2006
Air temperature	Ensemble	1992–2007
Specific humidity	HOAPS	1992–2005
Precipitation	HOAPS	1992–2005
Longwave radiation	ISCCP	1992–2007
Shortwave radiation	ISCCP	1992–2007
Continental discharge (Dai et al. 2009)	Syed et al. (2010)	1994–2006

TABLE 2. Percentage of grid points contributed by each product pair to the error estimates in each atmospheric variable. Bold values denote maximum percentage for each variable. Annotations are as follows: E: ERA-Interim, J: JRA25, N: NCEP, C: CORE2, R: Reference. Refer to Table 1 for specific references.

T_m	E-J	E-N	E-C	E-R	J-N	J-C	J-R	N-C	N-R	C-R
TauX	7.01	4.97	**27.44**	2.00	6.24	2.67	9.51	7.54	17.19	15.39
TauY	12.36	8.34	3.06	4.16	7.16	2.69	5.11	12.63	**25.97**	18.46
LwDn	**73.12**	21.06	0.27	0	1.39	0.05	0	3.02	0	1.05
SwDn	9.15	3.00	1.62	0.1	2.57	16.57	0.87	**52.63**	13.46	0.01
Tair	20.83	1.15	10.28	n/a	6.00	22.04	n/a	**39.69**	n/a	n/a
Humd	11.35	1.04	0.87	0.07	3.21	9.32	3.84	**35.56**	34.47	0.22
Rain	8.52	10.34	1.73	6.04	**20.85**	3.38	10.16	13.22	12.84	12.88
T_v	E-J	E-N	E-C	E-R	J-N	J-C	J-R	N-C	N-R	C-R
TauX	0.35	2.26	**32.96**	0.89	2.30	12.36	2.10	29.23	2.06	15.44
TauY	0.39	2.97	24.9	0.66	3.18	11.67	1.42	**30.23**	4.92	19.61
LwDn	1.05	0.3	13.2	0.92	6.64	**36.69**	9.7	22.96	7.55	0.89
SwDn	0.2	13.52	0.08	0.38	24.18	0	3.39	19.6	4.3	**34.20**
Tair	24.71	13.48	19.93	n/a	5.04	**32.93**	n/a	3.86	n/a	n/a
Humd	10.53	3.56	9.39	10.73	1.61	2.96	10.72	0.53	15.02	**34.91**
Rain	0.03	0.12	4.68	0.94	0.45	12.57	2.57	15.40	20.68	**42.5**

originally provided at 6-hourly intervals, are averaged over 14-day intervals. The 14-day interval is chosen to optimize the use of high-frequency reanalysis (6 hourly) along with low-frequency satellite-based composites (daily to weekly). Furthermore, the 14-day interval corresponds to the adjustment period of atmospheric fields chosen in the ECCO framework and requiring uncertainty estimates that are consistent with that period. Many other operational OGCMs have adjustment windows of 7–14 days (Broquet et al. 2009; Moore et al. 2011) and therefore knowledge of uncertainties in air–sea flux forcing at comparable intervals would be beneficial for such optimization procedures. Similarly the satellite-derived datasets (Table 1) are also gridded onto the NCEP grid at 14-day averages. Errors are partitioned into time-mean and time-variable components to explore how systematic biases compare with random uncertainties and study their individual influence on the total error of a specific atmospheric field. Differences in time-mean and time-variable errors, if significant, would point to the need to have a separate treatment of these components when trying to correct available atmospheric fields (e.g., Stammer et al. 2004; Large and Yeager 2004).

At each grid point, for a given pair of products with time series x and y, the time-mean error is calculated as $|\overline{x} - \overline{y}|$, where the overbar denotes time averaging, and time-variable error is calculated as the standard deviation of $(x - y)$. These values are calculated for all possible pairs and maximum values are taken to represent the error at each grid point. In regions of missing data or fields without a ''reference'' dataset (e.g., air temperature), the time-mean and time-variable errors are estimated from differences only within the reanalysis fields. This methodology is particularly used for data-sparse high-latitude regions. The estimated errors are compared against the original field by evaluating the signal-to-noise ratio of each atmospheric field as follows: the time-mean and time-variable signals are represented as the maximum of all reanalysis and reference mean and standard deviation fields, respectively, at a particular grid point. We choose the maximum signal to be consistent with the maximum error estimates described above. The prominent contributors of time-mean and time-variable errors for each field are determined by initially noting the source pairs that provide the largest differences at each ocean grid point. The sum of grid points for a particular source divided by the total number of ocean NCEP grid points represents the percentage of that source in the global error map. The error sources will be discussed in section 4 and shown in Table 2.

3. Comparison of reanalysis products

A comparison of reanalysis solutions against satellite-based data is conducted to better understand their fidelity to these references and among themselves. Note that some of the satellite-based ''reference'' datasets [e.g., Quick Scatterometer (QuikSCAT)] are directly assimilated into reanalysis models and hence are not true independent references. Furthermore, these datasets have implicit errors of their own, which are neglected as part of this study. However, the high-resolution spatial and temporal coverage provided by these satellite-based datasets make them most suitable for comparison.

Atmospheric fields from three reanalyses (NCEP, ERA-Interim, and JRA-25), one reanalysis-derived

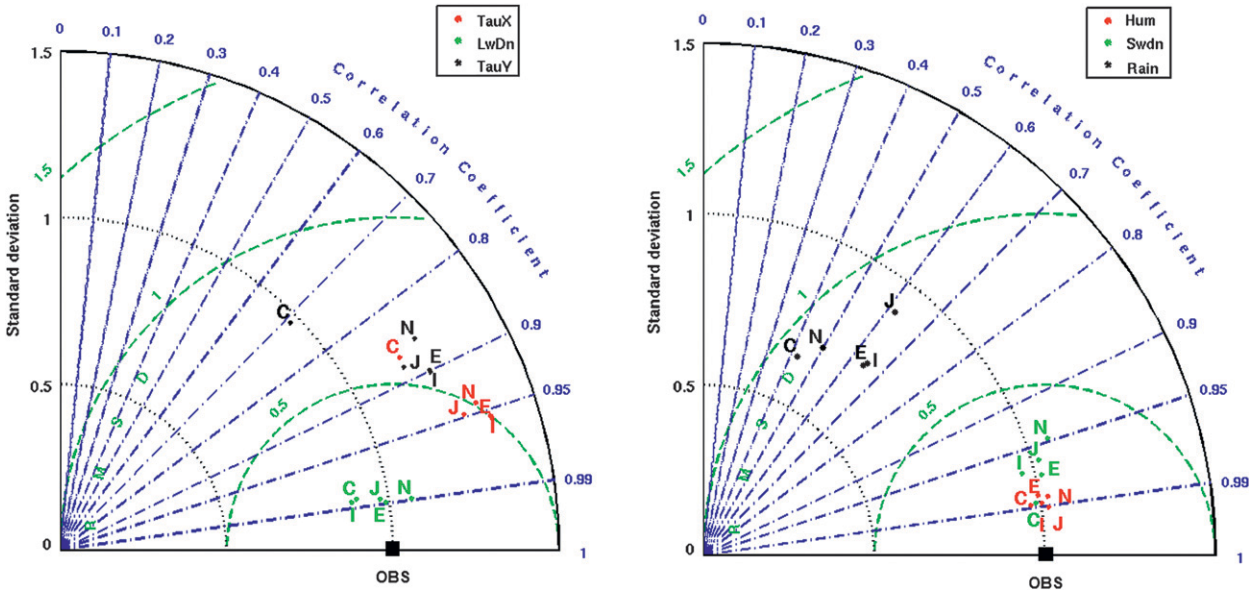

FIG. 1. Taylor diagram representing model performance (Taylor 2001). A polar coordinate system is used, with radius representing the normalized standard deviation, and angles (with respect to horizontal) representing the correlation coefficient, decreasing from 1 to 0. Green dashed lines show root-mean-square differences. (a) Comparison of satellite-based observations (black square) against zonal wind stress (TauX), meridional wind stress (TauY), and downwelling longwave radiation (LwDn) for NCEP (N), CORE2 (C), JRA-25 (J), ERA-Interim (E), and ECCO (I). (b) Comparison for precipitation (Rain), humidity (Hum), and downwelling shortwave radiation (SwDn).

product (CORE2), and the ''optimized'' ECCO product are evaluated quantitatively against respective satellite-based observations using Taylor (2001) diagrams. The Taylor diagram (Fig. 1) provides a statistical summary of how well the reanalysis products match the satellite-based observation patterns in terms of their correlation, root-mean-square (RMS) difference, and the ratio of their standard deviations. The radial distances from the origin represent the standard deviations, whereas the azimuthal positions show correlation coefficients. Simulation features that match well with observations in both amplitude and phase appear closest to the observed point in the diagram (e.g., the longwave radiation of ERA-Interim in Fig. 1a). There is also no restriction placed on the time or space domain considered. Thus we concatenate the time series at each grid point into a single vector and maintain the same order of concatenation for each dataset. The focus is on comparative statistics and not individual analysis. As the units of measurement are different for the five different fields, their statistics are nondimensionalized by normalizing, for each field, the RMS difference and the two standard deviations by the standard deviation of the corresponding observed field. This enables us to compare all the fields in one Taylor diagram.

Comparison to satellite-based observations shows that for any given variable, all reanalysis products are clustered around each other (Fig. 1). This is not surprising as most of the reanalysis solutions are driven by similar model dynamics and largely assimilate common datasets. Furthermore, assessment of individual atmospheric fields suggests that no single product seems to agree better in *all* fields with satellite-derived observations. Zonal and meridional wind stresses for ECCO, JRA-25, NCEP, and ERA-Interim display larger variability against QuikSCAT data (Fig. 1a) but compare reasonably with each other. CORE2 wind stresses originally derived by adjusting NCEP wind speeds to QuikSCAT winds expectedly show better agreement in standard deviation with QuikSCAT data. While CORE2 improves on standard deviation, it shows larger RMS differences and lower correlations in comparison to other reanalysis products (Fig. 1a). These differences are likely due to the short time period (2000–01) of analysis used to determine the biases between NCEP and QuikSCAT (Large and Yeager 2004). ECCO stresses are similar to ERA-Interim, implying small adjustments to the first-guess forcing fields in the preliminary optimization results analyzed here. Zonal and meridional wind speeds (not shown) display slightly better fidelity to QuikSCAT winds in comparison to wind stress components and have almost similar spreads amongst the reanalysis products.

Downwelling longwave radiation fields from reanalysis and ECCO-optimized solutions show close correspondence and high correlations when compared

against satellite-based International Satellite Cloud Climatology Product (ISCCP) data (Fig. 1a). Longwave radiation is one of the few fields that compares well for all products and the reference. Assessment of downwelling shortwave radiation reanalysis fields (Fig. 1b) also shows good correspondence with ISCCP data. It should be noted that CORE2 shortwave radiation is derived from ISCCP and hence shows best agreement, but Large and Yeager (2004) actually adjust the ISCCP field between 50°S to 40°N because of regional biases in the equatorial Atlantic and Pacific data between ISCCP and mooring data. The largest discrepancies between reanalysis and reference are seen in the precipitation field. Significant disparities occur in both standard deviations and rms differences of reanalysis products against the Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data (HOAPS) dataset (Fig. 1b). The spread of behavior between the reanalysis products is comparatively larger in precipitation than other fields. The smallest discrepancies are observed in the humidity field, which displays good correspondence with the HOAPS dataset for all the products. The good fidelity in humidity is likely due to its close correspondence to near-surface air temperature, which in turn is largely modulated by SST in the open oceans (Chou et al. 2000). Since SST measurements are readily available during the ''satellite-era'' (including our study period) and applied in most reanalyses products, humidity estimates are mostly similar. Broadly, most of the products evaluated in Fig. 1 show reasonable correspondence with each other; however, large discrepancies are seen for wind stress and precipitation fields in comparison to satellite-based references. The implications of these patterns in relation to the uncertainties associated with the various surface atmospheric fields are discussed in the next section.

4. Error estimates

In this section we will present error estimates of various atmospheric parameters used in forcing OGCMs and discuss which pairs of products yield the largest differences and contribute to the estimated uncertainties. Given the close proximity of the ECCO-adjusted ERA-Interim atmospheric fields to the original ERA-Interim fields (Fig. 1), the ECCO fields are not used in the analysis.

a. Wind stress and wind speed

The wind stress errors are computed for the period 1999–2006, which is chosen to coincide with the data available from QuikSCAT and CORE2. The mean zonal wind stress from the four reanalysis products appears to be stronger than QuikSCAT in most basins (Fig. A1). Risien and Chelton (2008) compare QuikSCAT against NCEP climatologies and suggest that coarser resolutions in reanalysis products allow them to resolve only large-scale features (>500 km). The dynamically important small-scale wind stress features are not resolved accurately. Further evidence for their importance was found in a regional eddy-permitting state estimates of the Southern Ocean (Mazloff et al. 2010), which produced small-scale wind stress adjustments. Milliff and Morzel (2001) also show spectral ringing in NCEP reanalysis fields globally. Xu and Scott (2008) note that ignoring ocean current dependence in the wind stress calculation artificially increases global wind power by 32% in strong current regimes. Ocean circulation effects are accounted for in scatterometer products but not in reanalysis products. Furthermore, Risien and Chelton (2008) suggest that wind stress fields in the various products are dependent on the specific formulation of drag coefficient used in the bulk aerodynamic formulas. Time-mean errors in zonal wind stress (Fig. 2a) are largest (>0.05 N m^{-2}) in the Southern Ocean. Smaller time-mean errors (<0.04 N m^{-2}) are seen across all basins within 10°–30°S, 10°–30°N, and 40°–70°N. Similar patterns are also noted by Large and Yeager (2009) for the CORE2 dataset. The large errors in the Southern Ocean seem to be within the 30%–50% range of the signal (Fig. 3a), which is comparable to signal-to-noise ratios of regions with smaller errors such as the equatorial basins. Thus regions with large magnitudes in wind stress have large time-mean errors (e.g., the Southern Ocean) but their noise level is similar to other regions with small errors. Time-variable errors for zonal wind stress (Fig. 2b) are <0.04 N m^{-2} in the low and midlatitudes up to 40°, but larger (>0.05 N m^{-2}) poleward of 40° for the reasons described earlier. The QuikSCAT data coverage is limited to 80°S–80°N; hence, errors at the highest latitude regions are based solely on largest differences in reanalysis products. A large portion of the Arctic Ocean shows smaller errors than subtropical regions because the reanalysis products are closer to each other in magnitude and variability. Also the presence of sea ice at higher latitudes likely has a reducing effect on time-variable errors However, signal-to-noise ratios of time-variable zonal stress are <2 globally (Fig. 3b). Thus, while zonal wind stress time-variable errors exhibit considerable spatial variability (Fig. 2b), noise levels are generally large (>50%) irrespective of error magnitudes (Fig. 3b).

Errors in meridional wind stress are comparatively smaller than in zonal wind stress (Fig. 2). The largest errors (>0.03 N m^{-2}) in the time mean (Fig. 2c) occur in the upwelling favorable regions such as the western coasts of North and South America, southwest Africa,

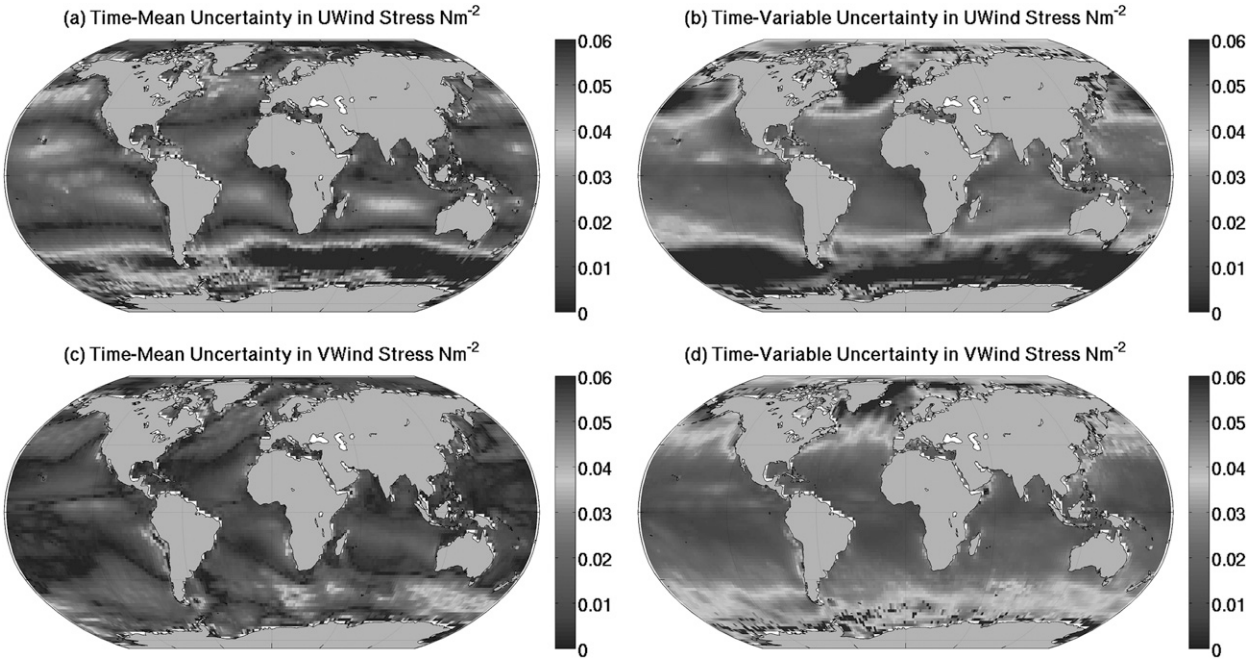

FIG. 2. Global wind stress uncertainty estimates in N m^{-2}. (a) Time-mean uncertainties in zonal wind stress. (b) Time-variable uncertainties in zonal wind stress. (c) Time-mean uncertainties in meridional wind stress. (d) Time-variable uncertainties in meridional wind stress. The gray scale is the same for both zonal and meridional wind stress components.

and Australia. Large errors are also seen east and west of Greenland due to the enhancement of wind speed by nearly 50% for these regions in the CORE2 dataset (Large and Yeager 2009). Again, the errors associated with the higher latitudes might be larger than depicted, but lack of data prevents more conservative values. Time-mean signal-to-noise ratios are generally high in low error regions and low in high error regions (Fig. 3c). The largest time-variable errors (>0.05 N m^{-2}) occur poleward of 40° in both hemispheres (Fig. 2d). Specifically, regions east and west of Greenland show large time-variable uncertainties. As with zonal stress, the noise levels in Fig. 3d are high (>50%) in most ocean basins. From Figs. 2 and 3, time-variable wind stress errors are mostly greater than time-mean errors in both absolute magnitude and in relation to the respective signals. Regarding the source of the errors (Table 2), the most values for time-variable (33%) and time-mean (27%) errors in zonal wind stress come from differences between ERA and CORE2; for meridional wind stress, the most time-mean (26%) and time-variable (30%) values originate from differences between NCEP and QuikSCAT and NCEP and CORE2, respectively.

The magnitudes of wind speed at 10 m also vary considerably among the various products and thus contribute toward differences in wind stress (Fig. 4). Considering first zonal wind speeds, time-mean errors are large (>1 m s^{-1}) in the equatorial regions, particularly in the intertropical convergence zone (ITCZ) and also in the Southern Ocean (Fig. 4a). Most of the subtropical basins show small time-mean errors (Fig. 4a). These results are in agreement with Wallcraft et al. (2009), who found similar biases and suggest that NCEP reanalysis have weak skill in the subtropical regions. Large time-variable errors are seen in the extratropical regions and the Southern Ocean (Fig. 4b). Broad coherence in spatial patterns exists between zonal wind stress (Fig. 1) and wind speed (Fig. 4) time-variable errors; however, the northern tropical regions such as the tropical eastern North Pacific are notable exceptions. Meridional wind speed show large time-mean errors (Fig. 4c) in upwelling favorable regions similar to meridional wind stress (Fig. 1c); however, large errors are also seen in the northeast Pacific trade wind region. Time-variable errors in meridional wind speed (Fig. 4d) are large in the extratropical and higher latitudes, similar to zonal wind speed (Fig. 4b). Time-variable errors are mostly larger than time-mean error for both zonal and meridional wind speeds (Fig. 4). The noise for zonal wind speed time-mean errors (Fig. 3o) is small (<20%) in most regions except the subpolar and higher latitudes and the tropical Pacific and Indian oceans. Time-variable errors have relatively higher noise levels (Fig. 3p) apart from some regions such as the Arctic, Arabian Sea, and Bay of

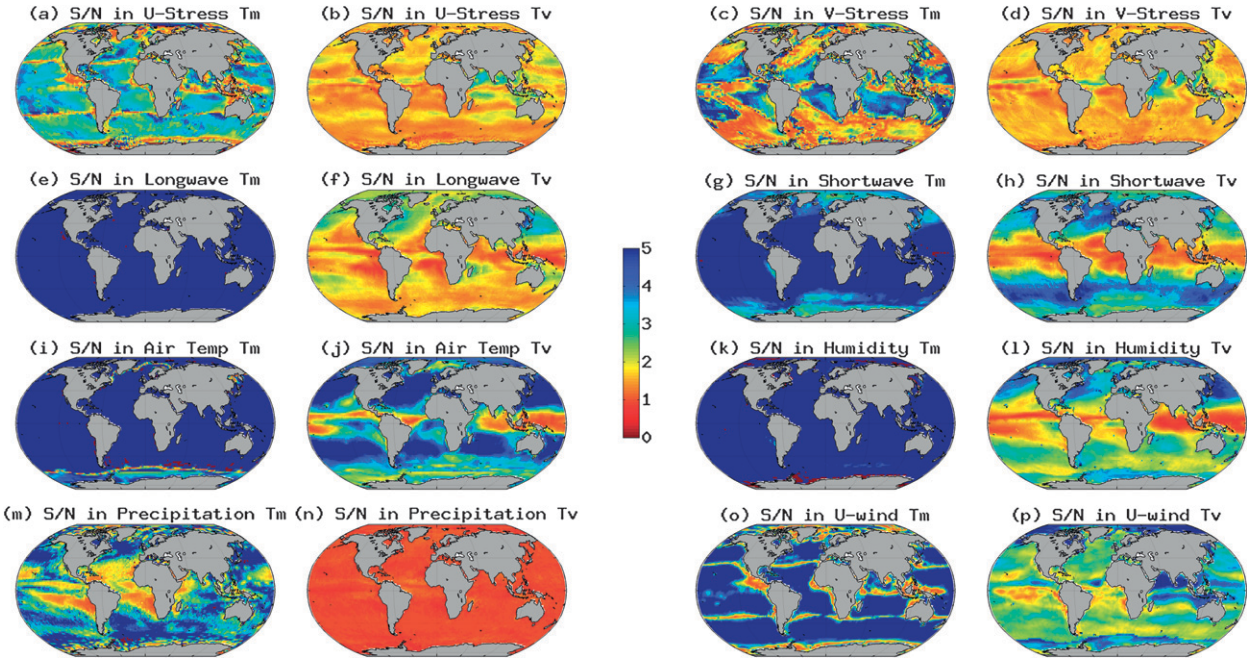

FIG. 3. Signal-to-noise ratios of (a) time-mean errors in zonal wind stress, (b) time-variable errors in zonal wind stress, (c) time-mean errors in meridional wind stress, (d) time-variable errors in meridional wind stress, (e) time-mean errors in downwelling longwave radiation, (f) time-variable errors in downwelling longwave radiation, (g) time-mean errors in shortwave radiation, (h) time-variable errors in downwelling shortwave radiation, (i) time-mean errors in air temperature, (j) time-variable errors in air temperature, (k) time-mean errors in humidity, (l) time-variable errors in humidity, (m) time-mean errors in precipitation, (n) time-variable errors in precipitation, (o) time-mean errors in zonal wind speed, and (p) time-variable errors in zonal wind speed. Meridional wind speed errors are not shown. The signal-to-noise ratios can be converted to percentage of noise (section 4) by taking the inverse of the plotted values and multiplying by 100. Thus regions showing blue (~5) can be interpreted as having 20% noise whereas regions showing red (~1) can be interpreted as having 100% or larger noise.

Bengal. Meridional wind stress signal-to-noise ratios are similar to zonal ratios and hence are not shown. In general, the noise levels in wind speeds are less than wind stress, which implies that uncertainties in other quantities such as drag coefficients and air densities amplify the uncertainties in wind stress.

b. Downwelling radiation

Downwelling radiation depends heavily on the parameterization of clouds, which is weakly resolved in radiative transfer models used in most reanalyses. Regions of high cloud cover, where low-level stratus and stratocumulus are inaccurately represented, have large systematic uncertainties. Time-mean (Fig. 5a) and time-variable (Fig. 5b) errors in downwelling longwave radiation show values >25 W m^{-2} in the Arctic and Southern Ocean regions. In the Arctic, Tjernström et al. (2008) found a systematic negative bias in downwelling radiation by comparing six regional models from the Arctic Regional Climate Model Intercomparison (ARCMIP) project against observations from the Surface Heat Budget of the Arctic Ocean (SHEBA) experiment. They conclude that some modeled cloud properties, such as the cloud water paths, are reasonable in a climatological sense, but temporal correlation of model cloud properties with observations are poor. Most models underestimate the presence of high clouds, and the modeled low clouds are too thin and displaced downward. Similar large uncertainties are also expected in the Southern Ocean (Dong et al. 2007). While high latitudes show similar spatial patterns in time-mean and time-variable errors, considerable differences occur at midlatitudes. Time-mean errors dominate in most regions, especially in the Southern Hemisphere, whereas time-variable errors are confined to coastal regions, especially upwelling regions. Signal-to-noise ratios for time-mean longwave radiation suggest noise levels $<20\%$ of the signal globally (Fig. 3e), which is relatively low compared to other fields (Fig. 3). In contrast, noise levels for time-variable term are generally $>50\%$ for most regions except the northeast Pacific region (Fig. 3f). Differences between ERA and JRA-25 contribute 73% of the time-mean error values (Table 2). For time-variable errors, most values (36%) arise from differences between CORE2 and JRA-25 (Table 2).

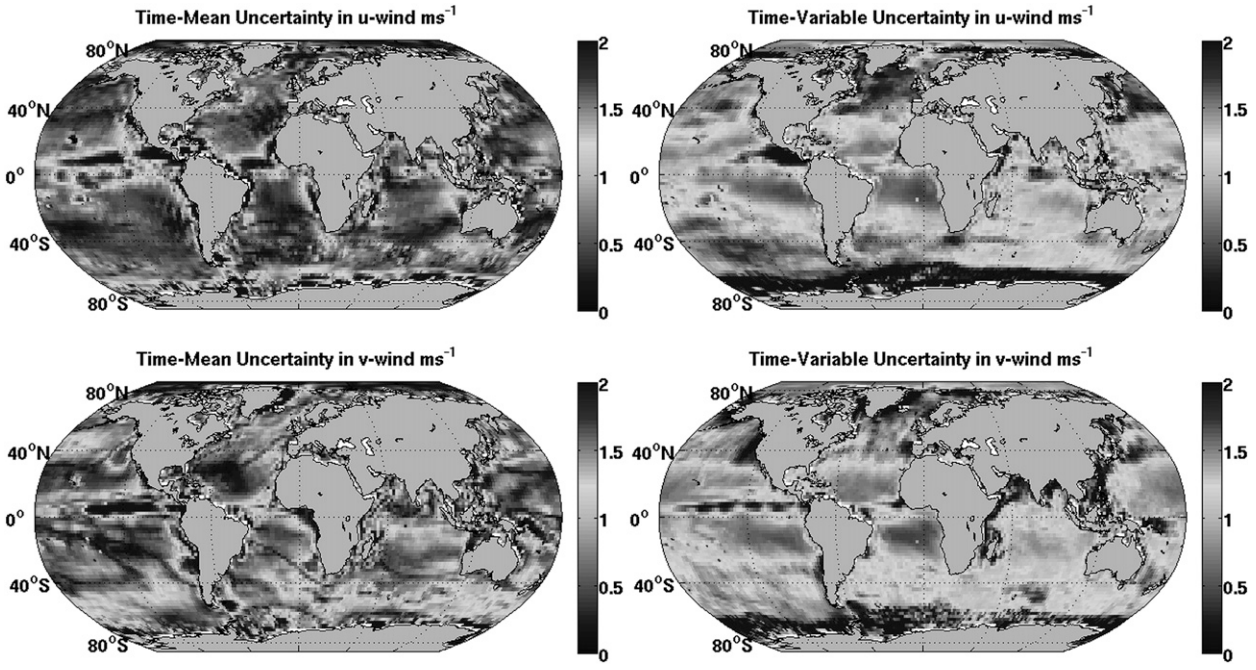

FIG. 4. Global wind speed error estimates in m s^{-1}. (a) Time-mean errors in zonal wind speed. (b) Time-variable errors in zonal wind speed. (c) Time-mean errors in meridional wind speed. (d) Time-variable errors in meridional wind speed.

In general, errors in downwelling shortwave radiation are larger than errors in longwave radiation. Time-mean errors (Fig. 5c) show values >50 W m^{-2} in the northwest and southeast Pacific, around Japan, and the Peruvian/Chilean coast respectively. Also, higher latitudes show expectedly large errors in the time mean. Within the Atlantic, large time-mean errors are seen offshore of the western African coast. The Arctic and

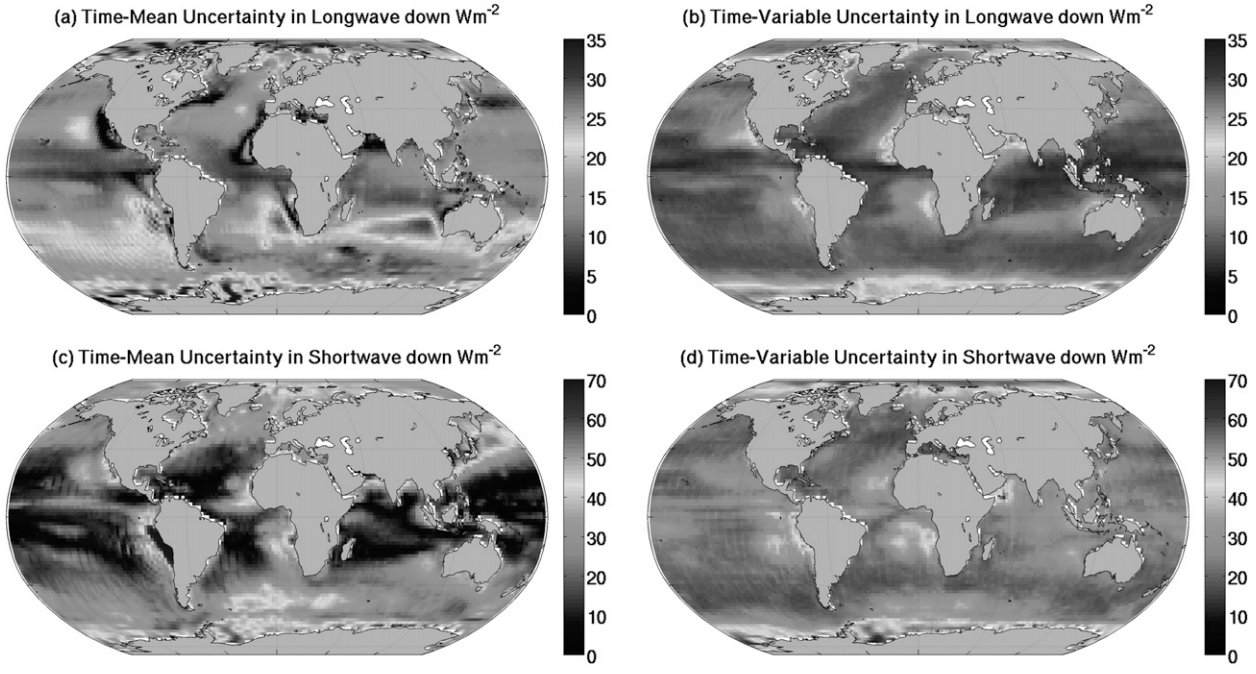

FIG. 5. Global surface downwelling radiation error estimates in W m^{-2}. (a) Time-mean errors in longwave radiation. (b) Time-variable errors in longwave radiation. (c) Time-mean errors in shortwave radiation. (d) Time-variable errors in shortwave radiation. The gray scale for shortwave radiation is double that of longwave radiation.

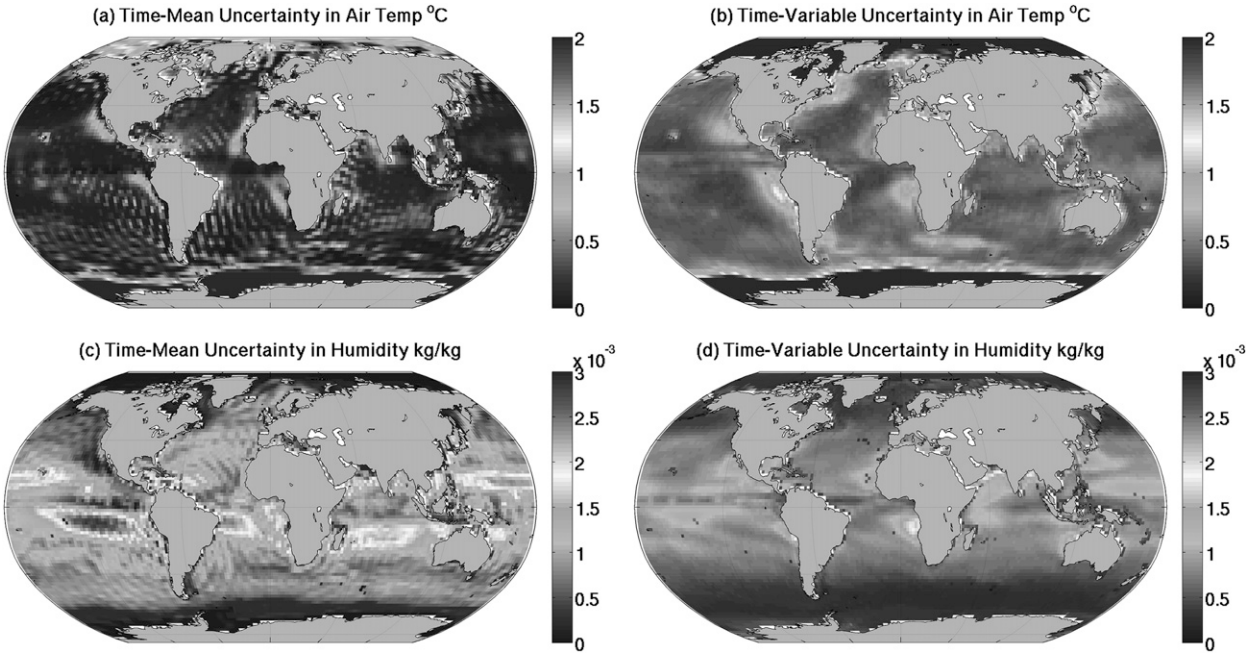

FIG. 6. Global air temperature and humidity uncertainty estimates at 2 m in °C and kg kg^{-1}, respectively. (a) Time-mean uncertainties in air temperature. (b) Time-variable uncertainties in air temperature. (c) Time-mean uncertainties in humidity. (d) Time-variable uncertainties in humidity.

the Southern Ocean display the largest time-variable errors (Fig. 5d), likely due to lack of data and knowledge of sea ice behavior in these regions. Signal-to-noise ratios for time-mean fields (Fig. 3g) show relatively low noise levels (<20%) in most basins except the high latitudes. However, the time-variable component (Fig. 3h) shows large noise (>50%) in most tropical regions. Global time-mean errors in shortwave radiation are dominated spatially by differences between NCEP and CORE2 (52%; Table 2). Differences between CORE2 and ISCCP contribute the most to the time-variable errors (34%) and are likely due to the 5% reduction in the ISCCP insolation made by Large and Yeager (2004) between 50°S and 30°N.

c. *Air temperature and humidity*

Air temperature is one of the fields where a global reference data is not available. In this case the largest differences between the four reanalysis products are taken to compute errors. The largest errors in both the time mean (Fig. 6a) and time variable (Fig. 6b) occur at high latitudes of both hemispheres. Rigor et al. (2000) report uncertainties in the range of 3°–5°C based on comparing NCEP and ECMWF reanalysis air temperatures with data from Arctic drift stations during 1979–97. If we take the RMS of time-mean and time-variable air temperatures errors for the Arctic, our estimates show an uncertainty of 3° to 4°C, similar to Rigor et al. (2002). Similarly, the Southern Ocean shows large errors (>2°C) in both time-mean and time-variable components. Most of the errors in the time mean for the Southern Ocean come from large differences between the NCEP and CORE2 products. These errors are likely due to a persistent cold bias in NCEP reanalysis air temperatures in comparison to weather station and drifting buoy data from Antarctica (Large and Yeager 2009). Large temperature errors in polar regions can lead to errors in water mass formation rates, which are critical in determining global ocean deep water properties and overturning circulation rates. Errors in low and midlatitudes are relatively small (<1°C). Problems related to topography truncation, as seen from the short-scale spotty patterns in many regions, also contribute to uncertainties (Fig. 6a). Signal-to-noise ratios for time-mean errors show low noise (<20%) for most regions (Fig. 3i), but considerably higher noise values in the tropics are present for time-variable fields (Fig. 3j).

Time-mean (Fig. 6c) and time-variable errors (Fig. 6d) in humidity are largest in tropical regions. Prominent errors (>0.002 kg kg^{-1}) in time mean are observed in the central basins of most tropical regions. Also, coastal regions along South America, southwest Africa, Brazil, and southern Asia show significant errors. These errors can be attributed to the moist bias introduced in the reanalysis fields over the tropics due to assimilation of satellite data (Brodeau et al. 2010). Comparison of Tropical

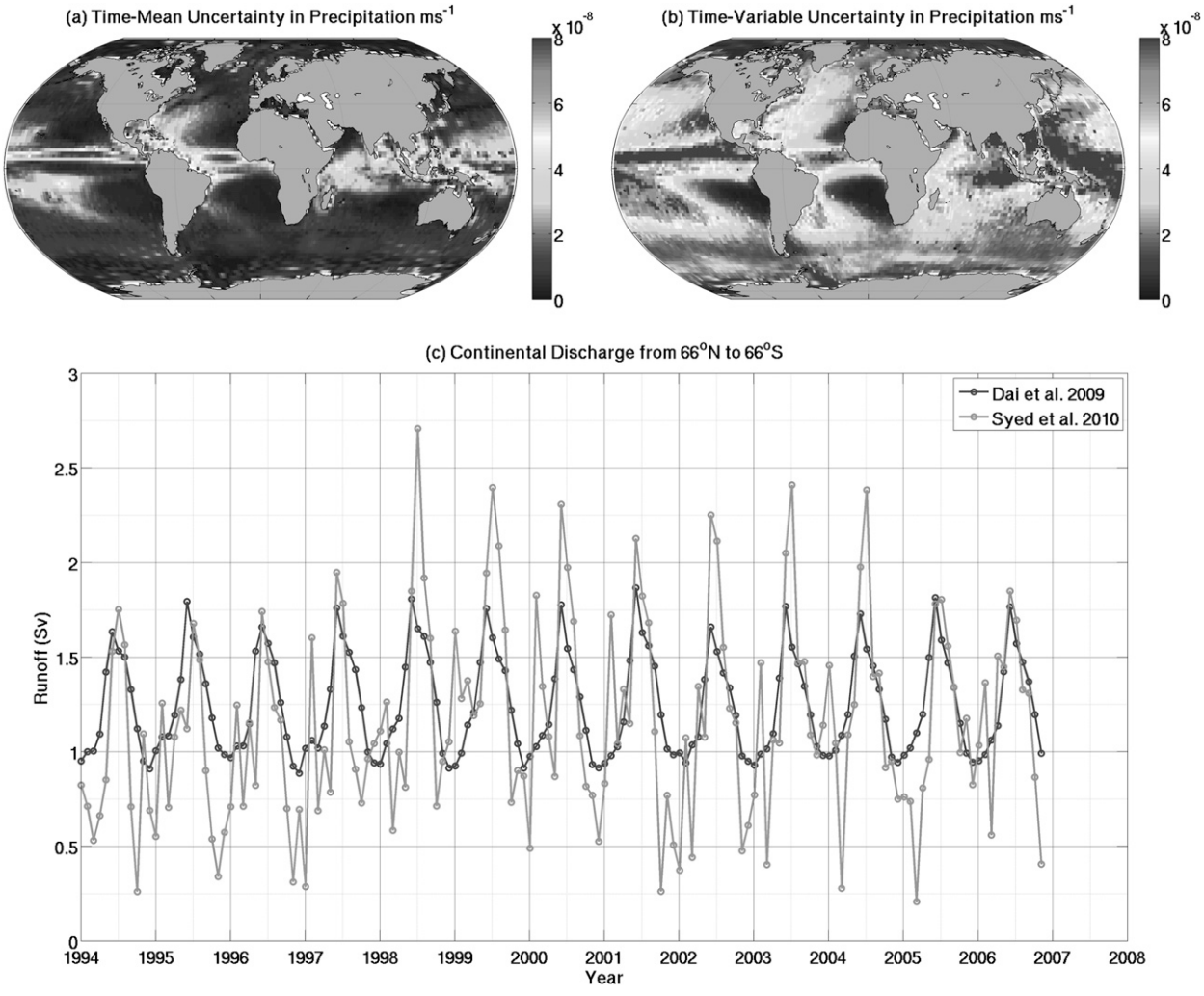

FIG. 7. (a) Time-mean and (b) time-variable precipitation uncertainties in m s^{-1}. (c) Estimates of continental discharge from Dai et al. (2009) database (dark gray) and Syed et al. (2010) global continental discharge (light gray) from 1994 to 2006 between 66°N and 66°S.

Atmosphere–Ocean (TAO) moorings with NCEP humidity in the western tropical Pacific shows a year-round positive bias of 2%–3% (Wang and McPhaden 2001). This bias is likely due to shortcomings in the boundary layer dynamics of NCEP reanalysis, which has a tendency for excessive evaporation (Smith et al. 2001). Furthermore, Kent et al. (1993) shows that ship-based humidity measurements ingested into the NCEP model have a moist bias. Other reanalysis products could possibly be affected by the same shortcomings. At high latitudes, cold air constrains humidity to very low values and hence the errors are relatively lower. Time-mean (Fig. 6c) errors in humidity are larger than time-variable errors (Fig. 6d). However, when compared against the signal, time-variable fields (Fig. 3k) are noisier than time-mean fields (Fig. 3l), with time-mean fields showing relatively low noise (<20%) globally (Figs. 3k,l).

d. Precipitation and continental discharge

Similar to radiation and humidity, precipitation estimates depend heavily on the representation of cloud cover, which is the weakest feature of reanalysis models (Taylor 2000). Both time-mean (Fig. 7a) and time-variable (Fig. 7b) errors are largest in the tropical regions, particularly in the ITCZ, but time-variable errors are also substantial in some of the extratropical basins of the Pacific and Atlantic. Brodeau et al. (2010) suggest that excess tropical precipitation in ERA solutions result from assimilation of ''contaminated'' satellite data due to the influence of the Mt. Pinatubo disturbance. Uppala et al. (2005) report that volcanic aerosols from the Mt. Pinatubo eruption in 1991 were misinterpreted as increased moisture in the High Resolution Infrared Sounder (HIRS) infrared radiance data. Similar errors

are expected to occur in other reanalysis products. Signal-to-noise ratios for time-mean fields show large noise (>50%) in the tropical Atlantic and Pacific basins and lower elsewhere (Fig. 3m). In contrast, time-variable fields (Fig. 3n) are very noisy globally (>80%) and in fact are the noisiest of all the variables examined in Fig. 3. Spurious trends and inhomogeneities as a result of changes in the observing system appear particularly strong in the Southern Ocean (Bromwich et al. 2011).

The ocean's density, and hence its pressure and circulation, are strongly affected by the transport of water across the ocean boundaries via continental discharge (Carton 1991). Furthermore continental discharge also provides minerals, nutrients, and contaminants to the oceans, thus affecting biogeochemical processes in the ocean. Since most OGCMs presently have formulations to include continental discharge as runoff in their solutions (Griffies et al. 2005; Danabasoglu et al. 2006), uncertainties in discharge are considered here although they are not directly obtained from reanalysis products. Errors in continental runoff are difficult to estimate as multiple observations at the same site are not available globally in space and time. A comparison between Dai et al. (2009) and Syed et al. (2010) runoff estimates between 66°N and 66°S show similar patterns in seasonal variability with peak continental discharge in Northern Hemisphere summer and low discharge in the winter (Fig. 7c). The annual signals of both datasets show good correspondence in phase but differ in amplitude. Syed et al. (2010) show larger month-to-month variations (Fig. 7c). From the differences of the monthly time series in Fig. 7c, estimated time-mean errors are 0.8 Sv ($1\ \text{Sv} \equiv 10^6\ \text{m}^3\ \text{s}^{-1}$) and time-variable errors are 0.24 Sv. In comparison to the signal, time-mean errors are ~10% and time-variable errors are ~45%, thus indicating large uncertainties present in continental discharge estimates.

5. Discussion and summary

Atmospheric forcing field uncertainties in this study are obtained by comparing multiple products. In contrast, several studies estimate errors by taking differences between two specific products. Lucas et al. (2008) compare ERA-40 and CORE fields, Leeuwenburgh (2005) calculates differences between ERA-40 and ERA-15, and Alves and Robert (2005) use differences between NCEP and ERA-40 to estimate uncertainties in their studies. Similarly, preliminary uncertainties derived from differences between NCEP and ERA-Interim at 14-day averages (at similar frequencies described in section 2) are currently being used in the ECCO optimization procedure. These uncertainties are needed as a priori variance estimates in the cost function term that penalizes too large atmospheric adjustments. To evaluate differences in error estimates obtained from using two-product and multiproduct approaches, we have compared uncertainties based on differences between ERA-Interim and NCEP fields (ERA–NCEP) with error estimates obtained from taking the maximum of differences between ERA and all other products discussed in this study.

For simplicity, we consider total uncertainties by taking the RMS of the time-mean and time-variable errors. Figure 8 shows comparisons for the case of zonal wind stress and precipitation. Uncertainties based on ERA–NCEP differences show relatively small errors, with prevalent errors confined along the continental margins (Fig. 8a). In contrast, errors obtained from ERA compared to multiple products show considerably larger errors over most ocean basins poleward of 40° (Fig. 8b). Similarly for precipitation, the ERA–NCEP errors (Fig. 8c) are much smaller than errors derived from multiproduct comparisons (Fig. 8d). However both estimates show similar spatial patterns, with largest errors in regions equatorward of 40°. Comparisons of ERA–NCEP errors with multiproduct derived error estimates for humidity, downwelling longwave and shortwave radiation, and meridional wind stress fields (not shown) display similar characteristics. Only air temperature errors in ERA–NCEP and multiproduct derived estimates show good correspondence in both magnitude and spatial patterns. The disparity in estimates can be attributed to evaluating multiple sources of errors in a multiproduct approach. Moreover, a two-product approach could be severely biased if the errors from the two sources are strongly correlated. The use of different products, including reanalysis and satellite-based reference datasets, moderates the possibility of error biases due to correlated errors in any single two-product comparison. Table 2 further shows that no product pairs dominate error estimates in most of the fields and choosing any pair can result in significant differences in error estimates. Uncertainty accuracy could be improved by excluding products that have known errors in certain regions (e.g., topography truncation errors in NCEP solutions could be omitted from the uncertainty analysis for these regions).

Large errors in atmospheric forcing fields discussed in this study have several implications on ocean state estimates. For example, wind stress errors in the subtropical and subpolar regions of the North Pacific and North Atlantic (Fig. 2) can create large uncertainties in the circulation patterns of these regions, including prevailing currents such as the Gulf Stream–North Atlantic, Greenland–Labrador, and Kuroshio–Oyashio systems.

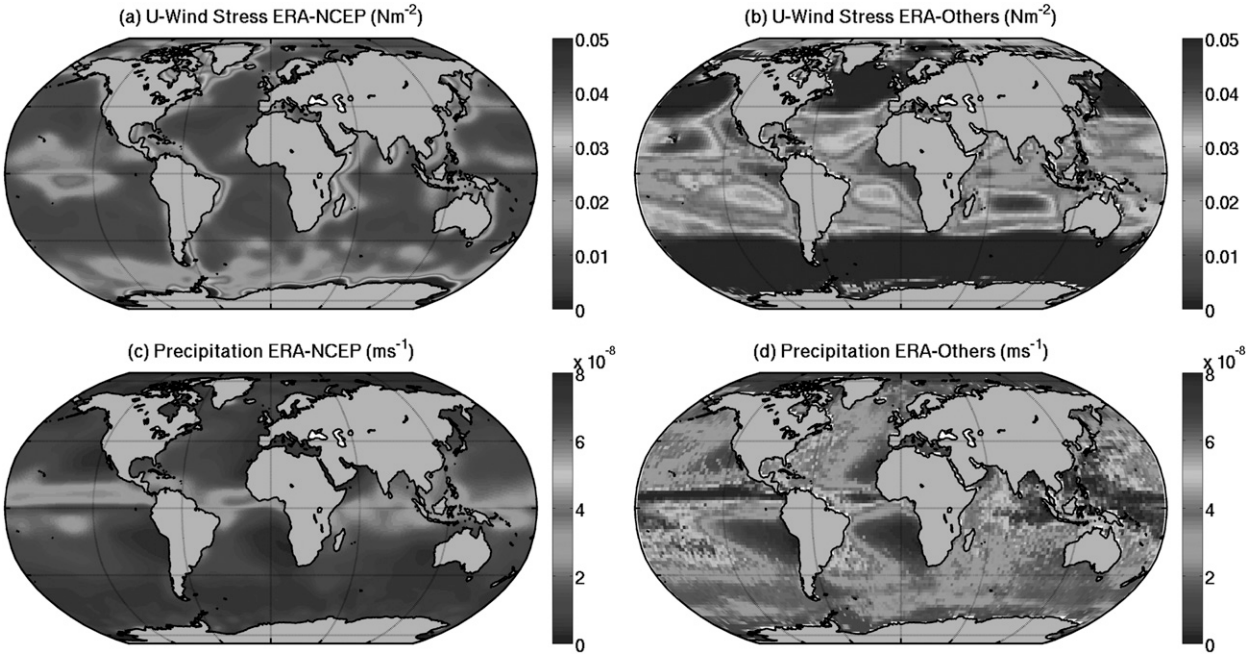

FIG. 8. Comparison of two-product and multiproduct error estimates. (a) ERA–NCEP zonal wind stress error. (b) ERA–multiple products zonal wind stress error. (c) ERA–NCEP precipitation error. (d) ERA–multiple products precipitation error.

Considerable errors in upwelling regions seen in shortwave and longwave fields (Fig. 5) can likely affect solution accuracies in the most ecologically active regions of the world. Moreover, for sea ice regions, a 1 W m^{-2} flux imbalance equates to 10 cm of ice melt in a year, which represents a significant fraction of the ice budget (Bourassa et al. 2013). Substantial errors in air temperature and humidity within the tropics will create biases in the heat and salinity budgets of these regions (Fig. 6). For example, Brodeau et al. (2010) compared solutions of the Atlantic meridional overturning circulation (AMOC) by forcing an OGCM with the CORE dataset and a bias-corrected ERA-40 dataset. They found that the strength of the AMOC was weaker by 2 Sv in comparison to both corrected ERA-dataset and in situ observations. They noted that weaker AMOC in the CORE solution was due to warm and moist biases in near surface air temperature and humidity fields respectively, which decreased buoyancy loss in deep convection regions and subsequently affected mixed layer depths and AMOC strength. Significant uncertainties in precipitation create large salinity imbalances in ocean models and particularly affect halocline-driven polar ocean solutions. Large uncertainties in surface atmospheric forcing fields further complicate the task of optimizing internal OGCM model parameters such as vertical mixing and diffusion.

In summary, we present estimates of uncertainty in surface atmospheric state fields commonly used to compute air–sea fluxes. Uncertainties are derived from comparisons between four reanalysis and satellite-based estimates of forcing fields. In the case of uncorrelated errors, these estimates provide an upper bound on uncertainty for each respective surface field considered in this study. Solutions from reanalysis products are mostly comparable because of similar physical assumptions and assimilation of common observations. The reanalysis products show large discrepancies particularly for wind stress and precipitation fields in comparison to satellite-based references. No reanalysis product shows best agreement for *all* fields when compared against satellite-based references. Large differences between time-mean and time-variable errors in most atmospheric fields justify the need to separate them for studying uncertainties as well as enhancing optimization procedures. Precipitation and wind stress fields display large time-mean and time-variable uncertainties when compared to the signal. Adjusted atmospheric fields from ECCO optimized solutions remain within the cluster spanned by the reanalyses. Errors obtained by evaluating multiple products in this study are significantly larger than errors estimated from a single pair of products. While this study estimates uncertainties in global atmospheric forcing fields, the next phase would involve quantifying the impacts of these atmospheric forcing field errors on ocean state solutions.

Acknowledgments. This study was supported by NSF Grant ARC-1022733. The authors thank Dr. Carl Wunsch

(MIT) for his comments and suggestions and Charmaine King (MIT) for data support. We also acknowledge Dr. James Famiglietti (UC Irvine) and Dr. Tajdarul Syed (ISM Dhanbad) for providing their continental discharge data and four anonymous reviewers for their insightful comments.

APPENDIX

Reanalysis- and Satellite-Derived Datasets

a. ERA-Interim

The ERA-Interim reanalysis is the most recent generation of ECMWF reanalysis (Simmons et al. 2006; Dee et al. 2011). The ECMWF-Interim model run covers the period 1979–present and has a horizontal resolution of T255 (~70 km) with 60 vertical layers. The model is an enhancement from the lower-resolution ERA-40 reanalysis (Uppala et al. 2005). It uses a four-dimensional variational data assimilation (4D-Var) technique to ingest data from in situ measurements as well as satellite-derived sources. These include water vapor and surface wind speeds from the Special Sensor Microwave Imager (SSM/I); ocean wave height and surface wind from European remote sensing satellites (*ERS-1* and *-2*), clear-sky radiances and upper-level winds from Meteosat-2, Global Ozone Monitoring Experiment (GOME) ozone profiles, total column ozone from the Total Ozone Mapping Spectroradiometer (TOMS), radio occultation measurements from the Challenging Minisatellite Payload (CHAMP), the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC), and the Gravity Recovery and Climate Experiment (GRACE). Improvements in data assimilation scheme, radiative transfer scheme, and model physics are made in ERA-Interim in comparison to the previous-generation ERA-40 product. The 6-hourly dataset of all atmospheric state variables is obtained from Research Data Archive (RDA) maintained by the Computational and Information Systems Laboratory (CISL) at NCAR. ERA-Interim-derived global mean zonal wind stress from 1999 to 2006 (Fig. A1) show strong positive stresses in the subtropical regions [~(30° to 60°)] in both hemispheres and negative stresses in lower-latitude tropical regions (<30°).

b. JRA-25

The Japanese 25-Year Reanalysis (JRA-25) is the first long-term reanalysis program undertaken in Asia (Onogi et al. 2005, 2007). JRA-25 uses the Japan Meteorological Agency (JMA) model and forecasting system to estimate atmospheric conditions from 1979 to present. The global model has a spatial resolution of T106 (~120 km) and 40 vertical levels. JRA-25 uses a 3D variational assimilation scheme to ingest data from conventional and satellite-derived data sources similar to NCEP and ERA-Interim. In addition, JRA-25 also assimilates data from wind profile retrievals surrounding tropical cyclones (TCR), SSM/I snow coverage, digitized Chinese snow depth data, and reprocessed atmospheric motion vectors (AMV). The 6-hourly dataset of all JRA-25 atmospheric state fields is downloaded from RDA/CISL. JRA-25-derived global mean zonal wind stresses from 1999 to 2006 (Fig. A1) show spatial patterns similar to those in ERA-Interim.

c. NCEP–NCAR

The NCEP reanalysis used here is the NCEP–NCAR reanalysis (Kistler et al. 2001). The NCEP model is run from 1948 to present and has a horizontal resolution of T62 (~180 km) with 28 vertical layers. The model ingests data from ships, aircrafts, buoys, wind profilers, radiosondes, and dropsondes. It also assimilates satellite-derived data for winds and radiances. Both in situ and remotely sensed data are assimilated into the model using spectral statistical interpolation (Kalnay et al. 1996). The 6-hourly dataset of all NCEP atmospheric state fields is acquired from the Earth System Research Laboratory, Physical Sciences Division (ESRL PSD), Boulder, Colorado, from their website (http://www.esrl.noaa.gov/psd/). NCEP-derived global mean zonal wind stress from 1999 to 2006 shows good correspondence with wind stress profiles of ERA-Interim and JRA-25 (Fig. A1); however, stronger negative wind stress estimates are seen at higher latitudes of Southern Hemisphere.

d. CORE2

The Coordinated Ocean Reference Experiment (CORE) is a dataset developed by Large and Yeager (2004). They take fields from NCEP reanalysis for near surface wind, air temperature, specific humidity, density, and data from satellite-based instruments for radiation, SST, sea ice concentration, and precipitation as their base dataset. This dataset resolved on T62 grid (~200 km) is then corrected by comparison with scatterometer winds, near-shore surface stations, and ocean buoy and ship data. The corrected data are then used together with the historical SST records to reproduce the air–sea fluxes of momentum, sensible and latent heat, and evaporation using bulk formulas. The major adjustments are a general increase in wind speed, decrease in humidity, and 5% reduction in tropical solar radiation (Large and Yeager 2009; Griffies et al. 2009). Fields from CORE version 2, which has improvements in continental runoff and air temperature and corrections to errors found in version 1, are used in this study (Griffies

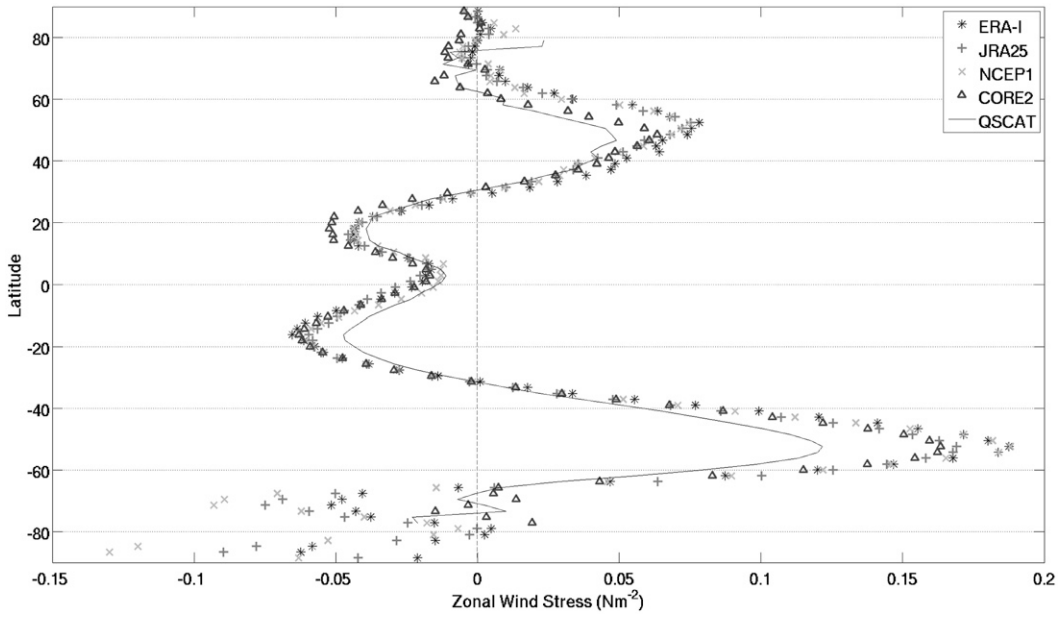

FIG. A1. Zonally averaged profiles of zonal wind stress from 1999–2006 for ERA-Interim, JRA-25, NCEP1, CORE2, and QuikSCAT.

et al. 2012). The air temperature and humidity provided at 10 m in CORE2 are readjusted to 2 m for this study to match other reanalysis products. The adjustment from 10 m to 2 m is done by using the bulk algorithm described by Large and Yeager (2004). While the CORE2 dataset is a mix of reanalysis and satellite-derived data, for the purposes of error estimation we consider it as a reanalysis product. The 6-hourly CORE2 dataset of winds, specific humidity, and air temperature, daily dataset of radiation, and monthly dataset of precipitation and continental discharge are downloaded from http://data1.gfdl.noaa.gov/. CORE2-derived global mean zonal wind stress profiles for the period between 1999 and 2006 show stronger negative wind stress estimates and weaker positive wind stress estimates for most of the Northern Hemisphere. In the Southern Hemisphere, the CORE2 wind stress show good correspondence with all the other reanalysis datasets presented in Fig. A1 and closest agreement with QuikSCAT data at higher latitudes.

e. ECCO

The ocean state estimate used in this study is the new-generation fully global ECCO version 4 product (G. Forget et al. 2012, unpublished manuscript), which takes atmospheric fields from ERA-Interim as its initial forcing state. The forcing fields together with initial conditions and model parameters such as mixing and dissipation form the control vector. The ECCO framework computes model-data misfits using a cost function (Wunsch and Heimbach 2007) with typical formulation as described in Stammer et al. (2002). As part of an iterative optimization procedure, the control vector (including atmospheric forcing) is adjusted to reduce the model-data misfits and minimize the cost function. The adjustment period for ECCO is 14 days. The minimization itself is dependent on the error covariances (or weights) for each dataset supplied to the cost function. Ideally a full error covariance matrix should be used, but in practice only diagonal error estimates are provided (Stammer et al. 2002). Here we use adjusted ERA-Interim atmospheric fields from a preliminary solution of an optimization in progress, mainly to make a first assessment of typical changes in atmospheric fields required to best match available oceanic observations, and how these changes compare to other atmospheric products. Although the optimization is not fully converged yet, the magnitudes of adjustments are not expected to change substantially as the optimization converges.

f. Satellite-based datasets

Zonal and meridional ocean wind stresses are derived from the SeaWinds scatterometer on board the QuikSCAT satellite. The QuikSCAT satellite samples nearly 90% of the world's ocean, measuring the surface roughness of the ocean from which neutral stability wind at 10-m height is retrieved (Schlax et al. 2001) daily since 1999 at 25-km spatial resolution. Comparison of scatterometer data against buoy observations suggests accuracies within 0.75 m s^{-1} in along-track and 1.5 m s^{-1} in cross-track directions (Chelton and Freilich 2005). The 10-m

winds are then converted into wind stress by bulk aerodynamic formulas (Risien and Chelton 2008). The zonal and meridional wind stress analysis for this study is conducted from 1999 to 2006 to coincide with the data available from QuikSCAT. Daily QuikSCAT wind stress fields are acquired from http://podaac.jpl.nasa.gov. The global temporal mean of the scatterometer-derived zonal wind stress shows considerable differences in comparison to the four reanalysis products (Fig. A1) with weaker stresses seen globally. A detailed analysis is presented in section 4.

Downwelling longwave and shortwave radiative fluxes from the reanalysis products are compared against satellite-derived fluxes from the International Satellite Cloud Climatology Product (ISCCP-FD) (Zhang et al. 2004), which are available at 280-km resolution. The FD version is the newest dataset from ISCCP that reduces noticeable errors from previous versions by using an advanced National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) radiative transfer model, improved ISCCP cloud climatology, and more ancillary datasets. Improvements are also made in the treatment of ice clouds, aerosol climatology, water vapor profiles, land surface albedos, and emissivities (Zhang et al. 2004). We analyze downwelling radiative flux fields from 1992 to 2007 for our study. The 3-hourly ISCCP data are obtained from the International Satellite Cloud Climatology Project web site (http://isccp.giss.nasa.gov) maintained by the ISCCP research group at the NASA Goddard Institute for Space Studies, New York, NY.

Precipitation and specific humidity reanalysis data are compared against satellite-derived observations from Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data (HOAPS; Andersson et al. 2010). Near-surface specific humidity is derived from SSM/I (Bentamy et al. 2003) using an improved version of the Schultz model (Schultz et al. 1993). The precipitation values are derived from a neural network–based precipitation algorithm that takes SSM/I data as input (Andersson et al. 2010). The HOAPS product does not account for ice-covered regions and thus large portions of the polar region are missing in the dataset. It must be noted that other global products such as Global Precipitation Climatology Project (GPCP) that composite satellite and rain gauge data (Huffman et al. 1997) and Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) that combines satellite, rain gauge, and atmospheric forecast models (Xie and Arkin 1997) are also available. Bosilovich et al. (2008) have used these datasets to compare against each other as well as against precipitation data from reanalysis products. However, we chose HOAPS for our study because of its specific focus on ocean precipitation. Specific humidity and precipitation are analyzed from 1992 to 2005. The 5-day averages of HOAPS-G version 3 data are obtained from the Climate and Environmental Data Retrieval and Archive website (http://cera-www.dkrz.de/CERA/) at the Max Planck Institute for Meteorology at 0.5° (~60 km) resolution.

No long-term global satellite-based estimates of lower atmosphere air temperature are currently available (Curry et al. 2004). The Atmospheric Infrared Sounder (AIRS) is a recent remote sensor that measures near-surface air temperature. However, the currently available length of processed data is considered too short for analysis in our study. We thus create an ensemble of the air temperature at 2 m from the four reanalysis products for the period of 1992 to 2007. The reanalysis ensemble mean air temperature is considered as the reference.

Very few global analyses of continental discharge exist that can quantify variations and changes in global freshwater discharge from land into the oceans, partly because of a lack of reliable data (Peel and McMahon 2006). A global dataset of historical monthly streamflow at the farthest downstream stations for the world's 925 largest ocean-reaching rivers has been created by Dai et al. (2009). They use gauge records from 80% of global ocean-draining land areas. Gaps in gauge data are filled by estimates from a land surface model. This dataset has been included into CORE2 at monthly intervals from 1948 to 2007. Another attempt to quantify global continental discharge has been undertaken by Syed et al. (2010) for the period 1994–2006. They have used 13 years of satellite-derived precipitation, evaporation, and altimeter-based sea level data to estimate freshwater discharge. Their formulation evaluates contributions of sea level changes, precipitation, and evaporation data to the global ocean mass balance. The residual mass is equated to continental discharge. However, their estimates are limited to 66° latitude in both hemispheres and thus exclude the Arctic and ice-covered regions. We subset the Dai et al. (2009) dataset from 66°N to 66°S to spatially match with Syed et al. (2010). While mapping temporal errors in space is not possible, we calculate global continental discharge errors by comparing time series of the two datasets.

REFERENCES

Alves, O., and C. Robert, 2005: Tropical Pacific Ocean model error covariances from Monte Carlo simulations. *Quart. J. Roy. Meteor. Soc.,* **131,** 3643–3658.

Andersson, A., K. Fennig, C. Klepp, S. Bakan, H. Graßl, and J. Schulz, 2010: The Hamburg Ocean–Atmosphere Parameters and Fluxes from Satellite Data—HOAPS-3. *Earth Syst. Sci. Data,* **2,** 215–234.

Bentamy, A., K. B. Katsaros, A. M. Mestas-Nuñez, W. M. Drennan, E. B. Forde, and H. Roquet, 2003: Satellite estimates of wind speed and latent heat flux over the global oceans. *J. Climate,* **16,** 637–656.

Bosilovich, M. G., J. Chen, F. R. Robertson, and R. F. Adler, 2008: Evaluation of global precipitation in reanalyses. *J. Appl. Meteor. Climatol.,* **47,** 2279–2299.

Bourassa, M. A, and Coauthors, 2013: High-latitude ocean and sea-ice surface fluxes: Challenges for climate research. *Bull. Amer. Meteor. Soc.,* in press.

Brodeau, L., B. Bernard, A. M. Treguier, T. Penduff, and S. Gulev, 2010: An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modell.,* **31,** 88–104.

Bromwich, D. H., J. P. Nicolas, and A. J. Monaghan, 2011: An assessment of precipitation changes over Antarctica and the Southern Ocean since 1989 in contemporary global reanalyses. *J. Climate,* **24,** 4189–4209.

Broquet, G., C. A. Edwards, A. M. Moore, B. S. Powell, M. Veneziani, and J. D. Doyle, 2009: Application of 4D variational data assimilation to the California current system. *Dyn. Atmos. Oceans,* **48,** 69–92.

Brunke, M. A., C. W. Fairall, X. Zeng, L. Eymard, and J. A. Curry, 2003: Which bulk aerodynamic algorithms are least problematic in computing ocean surface turbulent fluxes? *J. Climate,* **16,** 619–635.

——, Z. Wang, X. Zeng, M. Bosilovich, and C.-L. Shie, 2011: An assessment of the uncertainties in ocean surface turbulent fluxes in 11 reanalysis, satellite-derived, and combined global data sets. *J. Climate,* **24,** 5469–5493.

Bryan, F., 1986: High-latitude salinity effects and interhemispheric thermohaline circulations. *Nature,* **323,** 301–304.

Carton, J. A., 1991: Effect of seasonal surface freshwater flux on sea surface temperature in the tropical Atlantic Ocean. *J. Geophys. Res.,* **96,** 12 593–12 598.

Chelton, D. B., and M. H. Freilich, 2005: Scatterometer-based assessment of 10-m wind analyses from the operational ECMWF and NCEP numerical weather prediction models. *Mon. Wea. Rev.,* **133,** 409–429.

Chou, S. H., W. Zhao, and M.-D. Chou, 2000: Surface heat budgets and sea surface temperature in the Pacific warm pool during TOGA COARE. *J. Climate,* **13,** 634–649.

Cox, M. D., and K. Bryan, 1984: A numerical model of the ventilated thermocline. *J. Phys. Oceanogr.,* **14,** 674–687.

Curry, J. A., and Coauthors, 2004: SEAFLUX. *Bull. Amer. Meteor. Soc.,* **85,** 409–424.

Dai, A., T. Qian, K. E. Trenberth, and J. D. Milliman, 2009: Changes in continental freshwater discharge from 1948 to 2004. *J. Climate,* **22,** 2773–2792.

Danabasoglu, G., and J. C. McWilliams, 1995: Sensitivity of the global ocean circulation to parameterizations of mesoscale tracer transports. *J. Climate,* **8,** 2967–2987.

——, W. G. Large, J. J. Tribbia, P. R. Gent, B. P. Briegleb, and J. C. McWilliams, 2006: Diurnal coupling in the tropical oceans of CCSM3. *J. Climate,* **19,** 2347–2365.

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

Dong, S., S. T. Gille, and J. Sprintall, 2007: An assessment of the Southern Ocean mixed layer heat budget. *J. Climate,* **20,** 4425–4442.

Drobot, S. D., J. A. Maslanik, and C. Fowler, 2006: A long-range forecast of Arctic summer sea-ice minimum extent. *Geophys. Res. Lett.,* **33,** L10501, doi:10.1029/2006GL026216.

Griffies, S. M., and Coauthors, 2005: Formulation of an ocean model for global climate simulations. *Ocean Sci.,* **1,** 45–79.

——, and Coauthors, 2009: Coordinated Ocean-Ice Reference Experiments (COREs). *Ocean Modell.,* **26,** 1–46.

——, M. Winton, B. Samuels, G. Danabasoglu, S. Yeager, S. Marsland, and H. Drange, 2012: Datasets and protocol for the CLIVAR WGOMD Coordinated Ocean–Sea Ice Reference Experiments (CORE*s*), 20 pp. [Available online at http://data1.gfdl.noaa.gov/~nnz/mom4/COREv2/doc/CORE_notes_15feb2012.pdf.]

Han, Y.-J., 1984: A numerical world ocean general circulation model. Part II: A baroclinic experiment. *Dyn. Atmos. Oceans,* **8,** 141–172.

Haney, R. L., 1971: Surface thermal boundary condition for ocean circulation models. *J. Phys. Oceanogr.,* **1,** 241–248.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation datasets. *Bull. Amer. Meteor. Soc.,* **78,** 5–20.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–431.

Kent, E., P. Taylor, B. Truscott, and J. Hopkins, 1993: The accuracy of voluntary observing ships' meteorological observations—Results of the VSOP-NA. *J. Atmos. Oceanic Technol.,* **10,** 591–608.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Large, W. G., and S. Pond, 1981: Open ocean momentum flux measurements in moderate to strong winds. *J. Phys. Oceanogr.,* **11,** 324–336.

——, and S. G. Yeager, 2004: Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies. NCAR Tech. Rep. NCAR/TN-460+STR, 105 pp.

——, and ——, 2009: The global climatology of an interannually varying air–sea flux data set. *Climate Dyn.,* **33,** 341–364.

——, G. Danabasoglu, S. C. Doney, and J. C. McWilliams, 1997: Sensitivity to surface forcing and boundary layer mixing in a global ocean model: Annual mean climatology. *J. Phys. Oceanogr.,* **27,** 2418–2447.

Leeuwenburgh, O., 2005: Assimilation of along-track altimeter data in the tropical Pacific region of a global OGCM ensemble. *Quart. J. Roy. Meteor. Soc.,* **131,** 2455–2472.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.

Lucas, M., N. Ayoub, B. Barnier, T. Penduff, and P. de Mey, 2008: Stochastic study of the temperature response of the upper ocean to uncertainties in the atmospheric forcing in an Atlantic OGCM. *Ocean Modell.,* **20,** 90–113.

Mazloff, M. R., P. Heimbach, and C. Wunsch, 2010: An eddy-permitting Southern Ocean state estimate. *J. Phys. Oceanogr.,* **40,** 880–899.

McWilliams, J. C., 1996: Modeling the oceanic general circulation. *Annu. Rev. Fluid Mech.,* **28,** 1–34.

Milliff, R. F., and J. Morzel, 2001: The global distribution of the time-average wind-stress curl from NSCAT. *J. Atmos. Sci.,* **58,** 109–131.

——, W. G. Large, J. Morzel, G. Danabasoglu, and T. M. Chin, 1999: Ocean general circulation model sensitivity to forcing from scatterometer winds. *J. Geophys. Res.,* **104,** 11 337–11 358.

Moore, A. M., H. G. Arango, G. Broquet, B. S. Powell, J. Zavala-Garay, and A. T. Weaver, 2011: The Regional Ocean Modeling System (ROMS) 4-dimensional variational

data assimilation systems. Part I—System overview and formulation. *Prog. Oceanogr.,* **91,** 34–49.

Nicolas, J. P., and D. H. Bromwich, 2011: Precipitation changes in high southern latitudes from global reanalyses: A cautionary tale. *Surv. Geophys.,* **32,** 475–494.

Onogi, K., and Coauthors, 2005: JRA-25: Japanese 25-Year Re-Analysis Project—Progress and status. *Quart. J. Roy. Meteor. Soc.,* **131,** 3259–3268.

——, and Coauthors, 2007: The JRA-25 Reanalysis. *J. Meteor. Soc. Japan,* **85,** 369–432.

Peel, M. C., and T. A. McMahon, 2006: Recent frequency component changes in interannual climate variability. *Geophys. Res. Lett.,* **33,** L16810, doi:10.1029/2006GL025670.

Rigor, I. G., R. L. Colony, and S. Martin, 2000: Variations in surface air temperature observations in the Arctic, 1979–97. *J. Climate,* **13,** 896–914.

——, J. M. Wallace, and R. L. Colony, 2002: Response of sea ice to the Arctic oscillation. *J. Climate,* **15,** 2648–2663.

Risien, C. M., and D. B. Chelton, 2008: A global climatology of surface wind and wind stress fields from eight years of QuikSCAT scatterometer data. *J. Phys. Oceanogr.,* **38,** 2379–2413.

Salstein, D. A., R. M. Ponte, and K. Cady-Pereira, 2008: Uncertainties in atmospheric surface pressure fields from global analyses. *J. Geophys. Res.,* **113,** D14107, doi:10.1029/2007JD009531.

Schlax, M. G., D. B. Chelton, and M. H. Freilich, 2001: Sampling errors in wind fields constructed from single and tandem scatterometer datasets. *J. Atmos. Oceanic Technol.,* **18,** 1014–1036.

Schultz, J., P. Schlussel, and H. Grassl, 1993: Water vapour in the atmosphere boundary layer over oceans from SSM/I measurements. *Int. J. Remote. Sens.,* **14,** 2773–2789.

Simmons, A., S. Uppala, D. Dee, and S. Kobayashi, 2006: ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter,* No. 110, ECMWF, Reading, United Kingdom, 25–35.

Smith, S. R., D. M. Legler, and K. V. Verzone, 2001: Quantifying uncertainties in NCEP reanalyses using high-quality research vessel observations. *J. Climate,* **14,** 4062–4072.

Stammer, D., and Coauthors, 2002: The global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *J. Geophys. Res.,* **107,** 3118, doi:10.1029/2001JC000888.

——, K. Ueyoshi, A. Kohl, W. G. Large, S. A. Josey, and C. Wunsch, 2004: Estimating air–sea fluxes of heat, freshwater, and momentum through global ocean data assimilation. *J. Geophys. Res.,* **109,** C05023, doi:10.1029/2003JC002082.

Sun, B., L. Yu, and R. A. Weller, 2003: Comparisons of surface meteorology and turbulent heat fluxes over the Atlantic: NWP model analyses versus moored buoy observations. *J. Climate,* **16,** 679–695.

Syed, T. H., J. S. Famiglietti, D. P. Chambers, J. K. Willis, and K. Hilburn, 2010: Satellite-based global-ocean mass balance estimates of interannual variability and emerging trends in continental freshwater discharge. *Proc. Natl. Acad. Sci. USA,* **107,** 17 916–17 921.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in single diagram. *J. Geophys. Res.,* **106,** 7183–7192.

Taylor, P. K., Ed., 2000: Intercomparison and validation of ocean–atmosphere energy flux fields: Final report of the Joint World Climate Research Program and Scientific Committee on Ocean Research working group on air–sea fluxes. World Climate Research Program Rep. WCRP-112, 303 pp.

Tjernström, M., J. Sedlar, and M. D. Shupe, 2008: How well do regional climate models reproduce radiation and clouds in the Arctic? An evaluation of ARCMIP simulations. *J. Appl. Meteor. Climatol.,* **47,** 2405–2422.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.

Wallcraft, A. J., A. B. Kara, C. N. Barron, E. J. Metzger, R. L. Pauley, and M. A. Bourassa, 2009: Comparisons of monthly mean 10-m wind speeds from satellites and NWP products over the global ocean. *J. Geophys. Res.,* **114,** D16109, doi:10.1029/2008JD011696.

Wang, W., and M. J. McPhaden, 2001: Surface layer temperature balance in the equatorial Pacific during the 1997–98 El Niño and 1998–99 La Niña. *J. Climate,* **14,** 3393–3407.

Wunsch, C., and P. Heimbach, 2006: Estimated decadal changes in the North Atlantic meridional overturning circulation and heat flux 1993–2004. *J. Phys. Oceanogr.,* **36,** 2012–2024.

——, and ——, 2007: Practical global ocean state estimation. *Physica D,* **230,** 197–208.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558.

Xu, Y., and R. B. Scott, 2008: Subtleties in forcing eddy resolving ocean models with satellite wind data. *Ocean Modell.,* **20,** 240–251.

Zhang, Y.-C., W. B. Rossow, and A. A. Lacis, 1995: Calculation of surface and top-of-atmosphere radiative fluxes from physical quantities based on ISCCP datasets: 1. Method and sensitivity to input data uncertainties. *J. Geophys. Res.,* **100,** 1149–1165.

——, ——, ——, V. Oinas, and M. I. Mishchenko, 2004: Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *J. Geophys. Res.,* **109,** D19105, doi:10.1029/2003JD004457.