The Naval Research Laboratory Ocean Surface Flux (NFLUX) System: Satellite-Based Turbulent Heat Flux Products

JACKIE C. MAY AND CLARK ROWLEY

Naval Research Laboratory, Stennis Space Center, Mississippi

NEIL VAN DE VOORDE

Vencore, Inc., Stennis Space Center, Mississippi

(Manuscript received 10 July 2015, in final form 2 October 2015)

ABSTRACT

The Naval Research Laboratory ocean surface flux (NFLUX) system provides satellite-based surface state parameter and surface turbulent heat flux fields operationally over the global ocean. These products are presented as an alternative to using numerical weather prediction models—namely, the U.S. Navy Global Environmental Model (NAVGEM)—to provide the surface forcing to operational ocean models. NFLUX utilizes satellite sensor data records from the Special Sensor Microwave Imager/Sounder (SSMIS), the Advanced Microwave Sounding Unit-A (AMSU-A), the Advanced Technology Microwave Sounder (ATMS), and the Advanced Microwave Scanning Radiometer-2 (AMSR-2) sensors as well as satellite environmental data records from WindSat, the Advanced Scatterometers (ASCAT), and the Oceansat scatterometer (OSCAT). The satellite data are processed and translated into estimates of surface specific humidity, surface air temperature, and 10-m scalar wind speed. Two-dimensional variational analyses of quality-controlled satellite data, in combination with an atmospheric-model field, form global gridded surface state parameter fields. Bulk formulas are then applied to produce surface turbulent heat flux fields. Six-hourly analysis fields are created from 1 January 2013 through 31 December 2013. These fields are examined and validated against in situ data and NAVGEM. Overall, the NFLUX fields have a smaller bias, lower or similar root-mean-square error, and increased skill score relative to those of NAVGEM.

1. Introduction

Accurate representation of surface heat fluxes at the air–sea interface is an important aspect of atmospheric, oceanic, and coupled air–sea forecast modeling. The total ocean surface heat exchange is determined by the solar radiative flux, longwave radiative flux, latent heat flux (LHF), and sensible heat flux (SHF). These fluxes strongly influence the ocean mixed layer and sonic-layer depths as well as the stability and convection in the atmospheric boundary layer. Modeling for both short time scales, such as operational forecasting, and long time scales, such as climate studies, is highly dependent on the forcing from surface heat fluxes. This study focuses on the surface turbulent heat fluxes: LHF and SHF. The radiative fluxes will be the focus of a future paper.

The surface turbulent heat fluxes can change rapidly with time and space. The most dramatic changes typically occur in areas with sharp sea surface temperature (SST) gradients or strong winds. These regions include strong ocean currents, such as the Kuroshio and Gulf Stream (Small et al. 2008; Xu et al. 2011); eddies, hurricanes, tropical storms (Drennan et al. 2007); and cold-air outbreaks (Jensen et al. 2011). In the ideal situation, the turbulent heat flux fields used in global ocean modeling must be able to accurately represent the wide range of possible atmospheric and oceanic conditions, including the more extreme scenarios.

Direct surface turbulent heat flux measurements are often costly and limited to dedicated research missions (Yelland et al. 2009). Therefore, surface turbulent heat

Naval Research Laboratory Contribution Number NRL/JA/7320-15-2570.

Corresponding author address: Jackie C. May, Naval Research Laboratory, Bldg. 1009, Stennis Space Center, MS 39529.
E-mail: jackie.may@nrlssc.navy.mil

DOI: 10.1175/JAMC-D-15-0187.1

TABLE 1. Existing satellite-based global gridded surface turbulent heat flux products.

Dataset	Grid resolution	Temporal resolution	Time period
GSSTF v3	0.25° × 0.25°	Daily	Jul 1987–Dec 2008
HOAPS v3.2	0.50° × 0.50°	6-hourly, monthly	Jul 1987–Dec 2008
IFREMER v3	0.25° × 0.25°	Daily	Oct 1999–Nov 2009
J-OFURO v2	1.00° × 1.00°	Daily, monthly	Jan 1988–Dec 2008
OAFlux v3	1.00° × 1.00°	Daily, monthly	Jan 1985–Jul 2014
SeaFlux v1	0.25° × 0.25°	3 hourly	Jan 1998–Dec 2007

fluxes are typically estimated from bulk algorithms with inputs of near-surface atmospheric-model state parameters including specific humidity QA, air temperature TA, SST, and 10-m scalar wind speed U_{10}. The bulk formulas for LHF and SHF are expressed as follows:

$$\mathrm{LHF} = \rho C_E L_V |\mathbf{U}_{10} - \mathbf{U}_{\mathrm{sfc}}|(q_{\mathrm{sfc}} - \mathrm{QA}) \quad \text{and} \quad (1)$$

$$\mathrm{SHF} = \rho C_H c_p |\mathbf{U}_{10} - \mathbf{U}_{\mathrm{sfc}}|(\mathrm{SST} - \mathrm{TA}), \quad (2)$$

where ρ is the air density, C_E is the moisture transfer coefficient, L_V is the latent heat of vaporization, $\mathbf{U}_{10}$ is the vector wind at 10 m, $\mathbf{U}_{\mathrm{sfc}}$ is the vector surface current, q_{sfc} is the surface saturation specific humidity, C_H is the heat transfer coefficient, and c_p is the specific heat capacity of air. The input surface state parameters can be obtained from a variety of sources. In situ (ship and buoy) observations can provide long time series at specific locations, but the spatial coverage is very sparse, particularly in the Southern Hemisphere. Satellite observations offer much better spatial sampling at scales larger than the sensor footprint. Geostationary satellites have high temporal sampling but do not provide global coverage. Polar-orbiting satellites provide global coverage, but the temporal sampling of a given area can be relatively infrequent—from hourly to daily depending on the type and number of satellites that are being used. Because of the lack of adequate temporal and spatial coverage, the in situ and satellite observations by themselves are not able to provide the surface forcing fields for global modeling.

The alternatives to using individual in situ or satellite observations to obtain the input surface state parameters for global modeling are to use either numerical weather prediction (NWP) reanalysis products or satellite-based products. NWP reanalysis products typically assimilate at least some in situ and satellite data and can offer high temporal and spatial coverage, but there may be additional unassimilated surface observations, the flux parameterizations used within the NWP reanalysis products may not be consistent with current flux models, and there are often large regional biases (Curry et al. 2004; Smith et al. 2011). Satellite-based products produce global fields of the surface state parameters and then apply bulk algorithms to estimate the surface turbulent heat fluxes. Some satellite-based products assimilate model data and/or in situ data in addition to the satellite data. These products offer high temporal and spatial coverage and are more likely to have similar characteristics over time than the NWP reanalysis products, but they are also subject to errors in the satellite retrieval algorithms and accuracy of the surface state parameters (Smith et al. 2011).

Several satellite-based global gridded surface turbulent heat flux products are now available from various sources. Information on the most recent versions of each of the existing products is presented in Table 1. The products include the Goddard Satellite-based Surface Turbulent Fluxes, version 3 (GSSTF v3; Shie et al. 2009; Shie 2012); the Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data, version 3.2 (HOAPS v3.2; Andersson et al. 2010; Fennig et al. 2012); the French Research Institute for Exploitation of the Sea, version 3 (IFREMER v3; Bentamy et al. 2013); the Japanese Ocean Flux Dataset with Use of Remote Sensing Observations, version 2 (J-OFURO v2; Tomita et al. 2010); the objectively analyzed air–sea fluxes, version 3 (OAFlux v3; Yu and Weller 2007); and SeaFlux version 1 (v1; Clayson et al. 2015, 2014, manuscript submitted to *Int. J. Climatol.*). Each of the products has a relatively long time series and high spatial resolution, which is suitable for climate studies. None of these products are available in an operational or near-real-time mode, however; the surface fluxes are produced in a delayed mode. Because of the lack of real-time satellite-based products, the surface forcing fields used in operational ocean models currently come from NWP-model fields. In particular, the primary source of forcing for U.S. Navy global ocean models is the Navy Global Environmental Model (NAVGEM; Hogan et al. 2014).

This study discusses the development of an operational satellite-based global turbulent heat flux product, the Naval Research Laboratory (NRL) ocean surface flux (NFLUX) system. The NFLUX system is

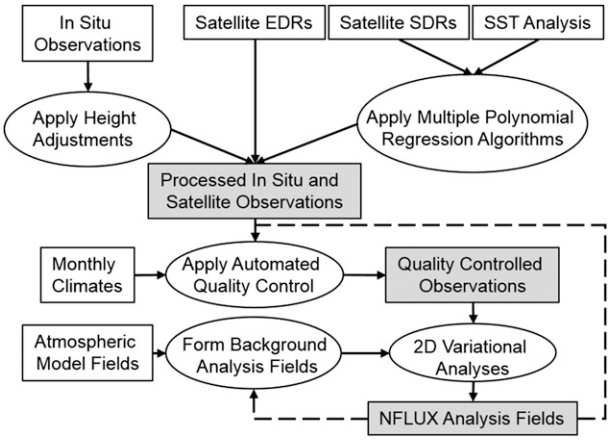

FIG. 1. The NFLUX system data-flow schematic. Unshaded rectangles signify the input datasets, ovals signify major internal processes, and shaded rectangles signify the end product from each primary component of the NFLUX system. The dashed arrow lines indicate that the end product from one cycle is used in the following cycle.

designed to provide near-real-time gridded surface flux products that preserve the temporal and spatial resolution of the NWP model while using all available satellite observations of the air–sea interface. In addition to using the NFLUX fields as an alternative to NWP fields to provide surface forcing to operational models, these satellite-based products can also be used to asses and monitor NWP products. The NFLUX surface state parameter and turbulent heat flux fields will be presented and validated against in situ observations and will be compared with NAVGEM fields.

2. The NFLUX system

NFLUX is an end-to-end data processing and assimilation system consisting of three primary components: data processing, quality control, and 2D variational analysis. A schematic of the NFLUX dataflow is presented in Fig. 1. Bulk algorithms are applied to the analysis fields to produce global gridded surface turbulent heat flux estimates.

a. Data processing

The first component of the NFLUX system processes satellite environmental data records (EDRs), sensor data records (SDRs), and in situ observations into surface state parameter estimates. All satellite data are processed at the swath level. A summary of the satellite and in situ data processed in the NFLUX system is provided in Table 2.

Satellite EDRs provide 10-m wind speeds that are retrieved from three sensors on board four platforms: the WindSat sensor on board the *Coriolis* platform (Bettenhausen et al. 2006), the Advanced Scatterometer (ASCAT) on board the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) *MetOp-A* and *MetOp-B* platforms (Figa-Saldaña et al. 2002; Verspeek et al. 2012), and the Oceansat scatterometer (OSCAT) on board the *Oceansat-2* platform (Chakraborty et al. 2013). The OSCAT sensor failed on 20 February 2014, but the data from prior to the failure are used within NFLUX. WindSat 50-km-resolution EDRs are obtained from the Fleet Numerical Meteorology and Oceanography Center (FNMOC). ASCAT 25-km-resolution and OSCAT 50-km-resolution EDRs are obtained from the Royal Netherlands Meteorological Institute (KNMI) Ocean and Sea Ice Satellite Application Facility. The retrieved EDR wind speeds are used without modification.

Satellite SDRs are retrieved from four sensors on board 10 platforms: the Special Sensor Microwave Imager/Sounder (SSMIS) sensor on board the Defense Meteorological Satellite Program (DMSP) *F16*, *F17*, and *F18* platforms; the Advanced Microwave Sounding Unit-A (AMSU-A) sensor on board the National Oceanic and Atmospheric Administration (NOAA) Polar-Orbiting Environmental Satellite *NOAA-15*, *NOAA-18*, and *NOAA-19* platforms and the *MetOp-A* and *MetOp-B* platforms; the Advanced

TABLE 2. Satellite and in situ data used within the NFLUX system.

Sensor	Platform	Data type	Source	NFLUX parameter
AMSU-A	*NOAA-15*, *NOAA-18*, *NOAA-19*; *MetOp-A*, *MetOp-B*	SDR	CLASS	QA, TA
ATMS	*SNPP*	SDR	CLASS	QA, TA
SSMIS	DMSP *F16*, *F17*, *F18*	SDR	CLASS	QA, TA, U_{10}
AMSR-2	*GCOM-W1*	SDR	JAXA	QA, TA, U_{10}
WindSat	*Coriolis*	EDR	FNMOC	U_{10}
ASCAT	*MetOp-A*, *MetOp-B*	EDR	KNMI	U_{10}
OSCAT	*Oceansat-2*	EDR	KNMI	U_{10}
—	—	In situ	NAVOCEANO	QA, TA, U_{10}

Technology Microwave Sounder (ATMS) sensor on board the *Suomi–National Polar-Orbiting Partnership* (*SNPP*) platform; and the Advanced Microwave Scanning Radiometer-2 (AMSR-2) sensor on board the *Global Change Observation Mission 1st–Water* (*GCOM-W1*) platform. The SSMIS, AMSU-A, and ATMS level-1B SDR data are obtained from the Comprehensive Large Array-data Stewardship System (CLASS). The AMSR-2 level-1R SDR data are obtained from the *GCOM-W1* data-providing service, the Japan Aerospace Exploration Agency (JAXA). Limb corrections, which adjust the off-nadir brightness temperatures to nadirlike brightness temperatures, have been applied to the AMSU-A and ATMS SDRs. In addition, antenna pattern corrections proposed by Mo (1999) have been applied to the NOAA AMSU-A SDRs. No additional corrections have been applied to the SSMIS or AMSR-2 SDRs.

The retrieved SDRs are translated into estimates of the surface state parameters using multiple polynomial regression algorithms on the basis of a series expansion of the hyperbolic tangent function. The coefficients were developed using a bootstrapping technique with all available brightness temperature channels for a given sensor–platform combination and a background SST field. For each retrieved parameter for each sensor–platform combination, unique algorithms were developed for ascending and descending orbits, as well as clear and cloudy conditions. The SDR data processing also includes checks for valid brightness temperature ranges, rain, near-land-contaminated data, and ice-contaminated data. The background SST fields are daily SST analysis fields produced by the U.S. Naval Oceanographic Office (NAVOCEANO) at 0.1° resolution. AMSU-A and ATMS data are used to estimate QA and TA, and SSMIS and AMSR-2 data are used to estimate QA, TA, and U_{10}. For full details on satellite SDR data processing and validation of swath-level satellite retrievals, refer to Van de Voorde et al. (2014, 2015).

In situ data are retrieved by NAVOCEANO from the World Meteorological Organization Global Telecommunication System. The in situ data include both voluntary observing ship (VOS) and buoy data. The in situ observations make direct measurements of TA and U_{10}. Depending on the reported measurements, QA is either calculated from the observed dewpoint temperature and surface pressure or from the relative humidity, air temperature, and surface pressure. Buoy QA and TA observation heights are nominally at 5 m, buoy U_{10} observation heights are nominally at 10 m, and the ship observation heights vary from 10 to 40 m (Kent et al. 2007). All in situ observations are height adjusted to a common height of 5 m for QA and TA and 10 m for U_{10} using the Coupled Ocean–Atmosphere Response Experiment (COARE) 3.0 algorithm (Fairall et al. 2003). These heights were chosen to allow the majority of in situ observations to be used without applying a height adjustment.

b. Quality control

The second component of the NFLUX system applies an automated quality control (QC) to all in situ and satellite observations. The NFLUX QC process is a variant of the Navy Coupled Ocean Data Assimilation (NCODA) system (Cummings 2005). A complex QC process is used in which each observation is subjected to a series of tests, with the final probability of error, or error likelihood, assigned on the basis of all of the QC test results. The most important QC tests in the system are the differences between the observation and the monthly climatological values and between the observation and the previous analysis field. The QC process was designed to ensure that erroneous data would be assigned a high probability of error, while extreme but still valid data would be assigned a low probability of error. Error likelihoods range from 0 to 1, with values near 1 representing a high probability of error.

To construct the monthly climate fields used in the QC process, we considered the satellite-based products listed in Table 1 and selected the SeaFlux dataset because of the high spatial and temporal resolution. Note here that the GSSTF v3 QA fields account for the Earth incidence angle (EIA) drifting effect, which causes artificial variations in the brightness temperatures (Hilburn and Shie 2011; Shie 2012), unlike the SeaFlux QA fields. By correcting the brightness temperatures for the EIA drift, the linear regression trend of the globally averaged QA is reduced from −0.2% to −0.03% per year (Shie 2012), which corresponds to approximately a $0.02\,g\,kg^{-1}$ correction per year. The total global mean QA uncertainty of the SeaFlux data is $0.57\,g\,kg^{-1}$ (Clayson et al. 2015, 2014, manuscript submitted to *Int. J. Climatol.*), which is greater than the expected error associated with the EIA drift. The SeaFlux 3-hourly global gridded products from January 1998 through December 2007 were obtained from the SeaFlux Internet site (http://seaflux.org/) and were averaged to create monthly climate fields.

c. 2D variational analysis

The third component of the NFLUX system performs 2D variational analyses of the quality-controlled satellite observations with atmospheric-model forecasts to produce gridded analysis fields of TA, QA, and U_{10}. For this study, all in situ observations have been withheld

from the analyses so that they can be used for validation. As with the QC component, the variational analysis component is also a variant of the NCODA system. This study uses NFLUX global gridded surface state parameter fields produced with a 6-hourly update cycle from 1 January 2013 through 31 December 2013. The global grid extends from 79.15°S to 79.15°N with 24-km spatial resolution.

Each 2D variational analysis is performed with a blended background, or prediction, field. The blended background field for QA and TA (U_{10}) is formed by adding the weighted average of the previous 12 (2) NFLUX analysis increment fields to the atmospheric-model forecast field (here, the NAVGEM 12-h forecast fields). The increment field is the gridded correction field estimated by the analysis. Adding previous increment fields to the forecast field acts to ''persist'' previously observed satellite-minus-model corrections.

The NFLUX analysis uses parameter-specific spatial correlation scales to define the background error, modeled as the product of a spatial correlation and a variance. The second-order autoregressive form is used as the analytical correlation structure in these results. To estimate the error correlation length scales, a time series was assembled of NAVGEM 12-h forecasts and verification analyses for each parameter, every day at 0000 and 1200 UTC for a full year (2013). At each verification time, an error field was calculated as the difference between the 12-h forecast field and the verification analysis field. At every point on a uniform 2° grid, neighboring error values from the difference field were accumulated into 50-km bins, from 50 to 500 km, and the binned covariances were fit with a Gaussian function. The characteristic scale of the Gaussian function was defined as the correlation length scale at that grid point. The time series of correlation length fields were then averaged, and a two-way smoothing filter was applied to produce the final error correlation length scales used within the NFLUX system (Fig. 2). The error variance is estimated at each analysis time as the weighted sum of the previous 10 days of successive 6-h forecast-field differences.

Each satellite retrieval algorithm has a unique observation error, as does each platform that provides EDR wind speeds. The observation errors are taken to be the root-mean-square error of the swath-level retrievals as compared with in situ data that are within 1° and 3 h of the satellite swath-level retrieval time and location. The range of observation errors for each sensor is given in Table 3. For a more detailed list of observation errors, refer to Van de Voorde et al. (2015).

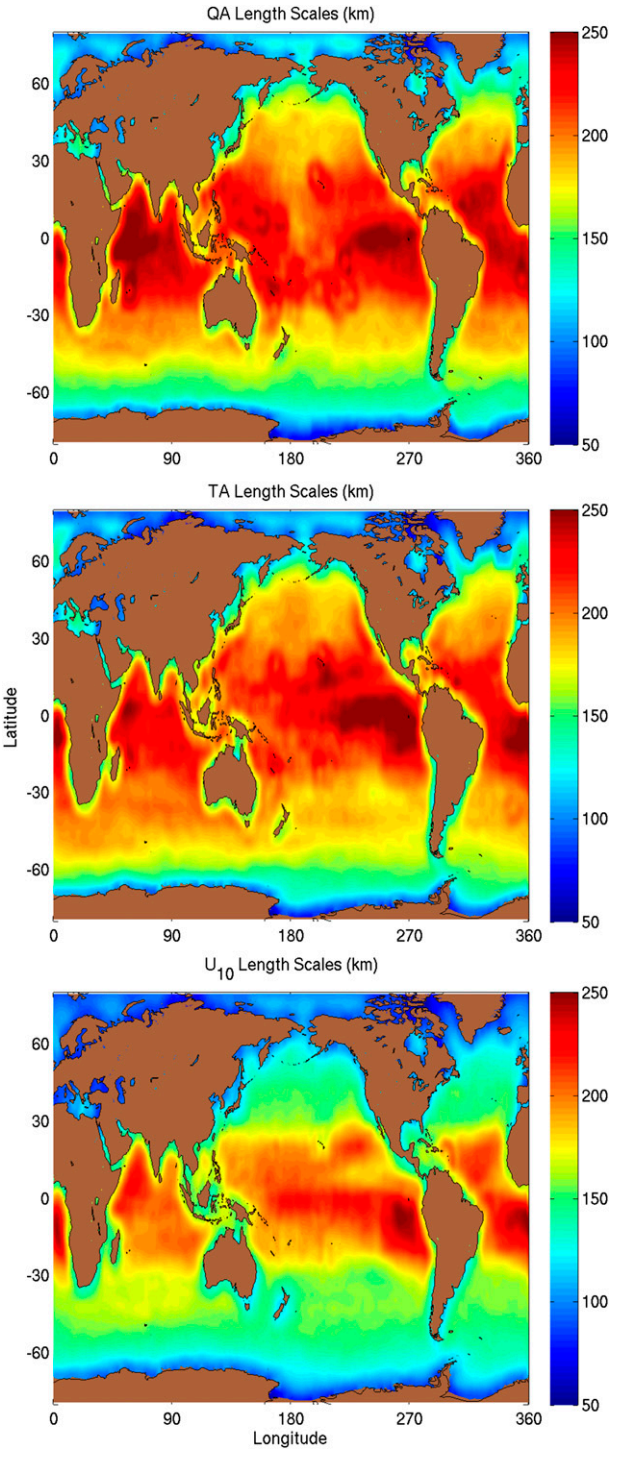

FIG. 2. NFLUX horizontal length scales (km) for (top) specific humidity, (middle) air temperature, and (bottom) wind speed.

All available satellite observations with a QC value of 0.90 or less and an observation time within 3 h of the NFLUX analysis time are assimilated into the analyses. As a means of thinning the data for computational

TABLE 3. Range of satellite observation errors by sensor.

Sensor	QA (g kg^{-1})	TA (°C)	U_{10} (m s^{-1})
AMSU-A	1.47–2.33	1.28–3.17	—
ATMS	1.19–1.59	1.34–1.86	—
SSMIS	1.22–1.86	1.50–1.87	1.45–2.85
AMSR-2	1.16–1.53	1.27–1.84	1.17–2.43
WindSat	—	—	1.69
ASCAT	—	—	1.25–1.34
OSCAT	—	—	1.24

efficiency, ''super observations'' are created for each surface state parameter by averaging the input satellite observations within bins that are 3 times the size of the global analysis grid-mesh interval. For locations and times at which there are no satellite data available, the blended background field is effectively persisted.

d. Flux fields

The NFLUX system uses the COARE 3.0 bulk algorithm (Fairall et al. 1996, 2003) to produce global stability-dependent surface latent and sensible heat flux estimates. This algorithm was chosen on the basis of previous studies by Brunke et al. (2003) and Iwasaki et al. (2010) that compared various bulk formulas and determined that the COARE 3.0 algorithm had the best overall performance. In the current study, we do include the warm-layer and cool-skin calculations. The velocity roughness length is specified as Charnock's expression; there is no wave dependence. Fairall et al. (2003) found that the COARE 3.0 algorithm is accurate to within 5% for wind speeds of 0–10 m s^{-1} and 10% for wind speeds of 10–20 m s^{-1}.

The inputs to the COARE 3.0 algorithm include NFLUX QA, TA, and U_{10} gridded analysis fields and NAVGEM SST forecast fields. The U_{sfc} term is set to zero. Because this study compares NFLUX with NAVGEM and the NFLUX system does not provide SST analysis fields, the NAVGEM SST fields were used so that the results highlight the differences in the NFLUX and NAVGEM LHF and SHF fields attributable to the QA, TA, and U_{10} fields. Here, we use the convention that positive values represent an upward heat flux from the ocean to the atmosphere and that negative values represent a downward heat flux from the atmosphere to the ocean.

In addition to the heat fluxes, ocean models also use the surface momentum flux. The momentum flux can be estimated using the COARE 3.0 bulk algorithms and vector wind stress. The current NFLUX system produces a scalar wind speed and not a vector wind speed, which limits the application of the bulk algorithms to the turbulent heat flux estimations; therefore, the momentum flux is not computed as part of the NFLUX system.

3. NAVGEM

NAVGEM is a global atmospheric forecast and data assimilation system that uses the NRL Atmospheric Variational Data Assimilation System–Accelerated Representer (NAVDAS-AR), an operational four-dimensional weak-constraint variational data assimilation system (Xu et al. 2005; Rosmond and Xu 2006). The assimilation of new data occurs every 6 h, and full forecasts are run every 12 h; forecast fields are output every 3 h. For each analysis, the SST field is obtained from the FNMOC. The SST analysis field is used as a lower boundary condition in NAVGEM and is held constant throughout the forecast (Hogan et al. 2014). NAVGEM does assimilate in situ data, as well as data from some of the sensors that are used within NFLUX. Microwave sounding radiances from the AMSU-A and SSMIS sensors are assimilated, but the QA and TA retrievals are not. ASCAT, SSMIS, and WindSat U_{10} data are also assimilated, but the scaling factors, the background fields, the forming of super observations, and other parameters that are used to form the analysis fields are different between NAVGEM and NFLUX.

The NAVGEM fields used in this study include the surface air temperature, surface specific humidity, 10-m wind speed, and sea surface temperature fields. These fields are retrieved from the operational product distribution on a uniform 0.5° application grid and are used without modification. We calculate turbulent heat fluxes from the NAVGEM retrieved products using the same COARE 3.0 bulk algorithms that are used to calculate the NFLUX turbulent fluxes. As with the NFLUX turbulent heat flux calculations, the U_{sfc} term is set to zero.

4. Comparisons with in situ observations

One year of 0000 and 1200 UTC NFLUX global gridded surface state analysis fields and surface turbulent heat flux fields are compared with NAVGEM 12-h forecast fields and are validated against in situ observations. Because NAVGEM assimilates in situ observations, the 12-h forecast fields, rather than the analysis fields, are used for comparison to ensure that the in situ observations that are used for validation have not yet been assimilated.

The statistical metrics used to evaluate the gridded products in this study are mean error (ME), standard deviation (SD), root-mean-square error (RMSE), correlation coefficient squared R^2, and an overall skill score (SS). The error statistics are defined using

$$\mathrm{ME} = \overline{Y} - \overline{X}, \tag{3}$$

$$\mathrm{SD} = \left\{ \frac{1}{n} \sum_{i=1}^{n} [(Y_i - \overline{Y}) - (X_i - \overline{X})]^2 \right\}^{1/2}, \tag{4}$$

$$\mathrm{RMSE} = \left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - X_i)^2 \right]^{1/2}, \tag{5}$$

$$R^2 = \left[\frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) / (\sigma_X \sigma_Y) \right]^2, \quad \text{and} \tag{6}$$

$$\mathrm{SS} = 1 - \frac{\mathrm{MSE}(Y, X)}{\mathrm{MSE}(C, X)}, \tag{7}$$

where X_i are the in situ observations, Y_i are the NFLUX or NAVGEM analysis field values, σ_X and σ_Y are the corresponding standard deviations, and the overbar represents a simple average. MSE(Y, X) is the mean square error of the NFLUX or NAVGEM analysis field values relative to in situ data, and MSE(C, X) is the mean square error of monthly climatological values relative to in situ data.

The ME measures the overall mean bias, SD is the standard deviation of the difference between the in situ and model data, RMSE measures the absolute error between the in situ and model data, and R^2 is a measure of the linear association between the model and the observation. The SS is a nondimensional measure of the skill of a model relative to using climatological values (Murphy 1988, 1995). In this study, the same SeaFlux monthly climatological fields used in the QC are used as the reference climatological fields in the SS calculation. A positive SS represents improved skill, and a negative SS represents reduced skill. An SS of 1.0 is perfect, and an SS of 0.0 represents no difference in skill between using the model and using monthly climatological values.

The in situ observations of QA and TA (U_{10}) used for validation have been height adjusted to 5-m (10 m) reference heights, as discussed in section 2a, and have been through the automated QC but have not been assimilated into the NFLUX analyses. The in situ observations of LHF and SHF used for validation have been calculated using the COARE 3.0 bulk algorithms, with the U_{sfc} term set to zero. VOS and buoy in situ observations that contain valid measurements of QA, TA, U_{10}, and SST are used to estimate the turbulent heat fluxes. VOS SST observations are known to have large uncertainty that is due to different measuring methods (Berry and Kent 2011). By only using buoys that report the four components used to calculate the turbulent fluxes, however, the in situ coverage is very sparse. Including both VOS and buoy observations provide adequate spatial coverage and observations that sample a wide range of atmospheric conditions.

In situ observations assigned an error likelihood estimate of 0.90 or less by the QC process and observed within 3 h of the analysis time are used to create the matchup validation dataset. For each verification analysis time, each in situ observation call sign or identification number (ID) must be unique. If multiple call signs or IDs exist, only the observation nearest in time to the verification analysis time is used. To reduce coastal boundary effects in the matchup dataset, the in situ validation observations must also be located greater than 111 km from land. The global in situ validation dataset consists of 128 086 QA observations, 199 944 TA observations, 194 649 U_{10} observations, and 81 510 turbulent heat flux observations.

a. Surface state parameters

Comparisons of the NFLUX and NAVGEM state parameter fields versus the in situ data are presented in Fig. 3, and the corresponding error statistics are provided in Table 4. NFLUX and NAVGEM both have a strong linear relationship with the in situ data for QA and TA. The NFLUX and NAVGEM comparisons for U_{10} also have a linear relationship with the in situ data, but with much more scatter. The large scatter in U_{10} is primarily due to comparing a model wind speed, which represents the wind speed over a large area (24 km in this study), with a single point observation.

From Table 4 it is seen that NFLUX (NAVGEM) has an overall positive (negative) bias for each parameter. For QA and U_{10}, NFLUX has a smaller absolute value for the bias, lower RMSE, and increased SS relative to NAVGEM. For TA, NFLUX has a smaller absolute value for the bias, similar RMSE, and similar SS relative to NAVGEM. To examine the errors further, the total number of in situ observations for each parameter was sorted and divided into 25 equally populated bins. The bias and RMSE were then calculated for each bin (Fig. 4). The general trends of the bias and RMSE for each parameter are similar between NFLUX and NAVGEM, with the NFLUX bias showing a positive offset from the NAVGEM bias throughout the sampled range.

The NFLUX QA comparisons have a slightly higher RMSE at low specific humidities and a lower RMSE at high specific humidities relative to NAVGEM. As QA increases from approximately 17 g kg^{-1}, NAVGEM has an increasing RMSE and increasing negative bias, which is likely related to the model assimilation of satellite radiances reaching saturation. NFLUX shows a similar but less pronounced pattern. The improved results for NFLUX seen at high specific humidities are attributed to the multiple polynomial regression algorithms used in the satellite retrievals, which are able to more accurately

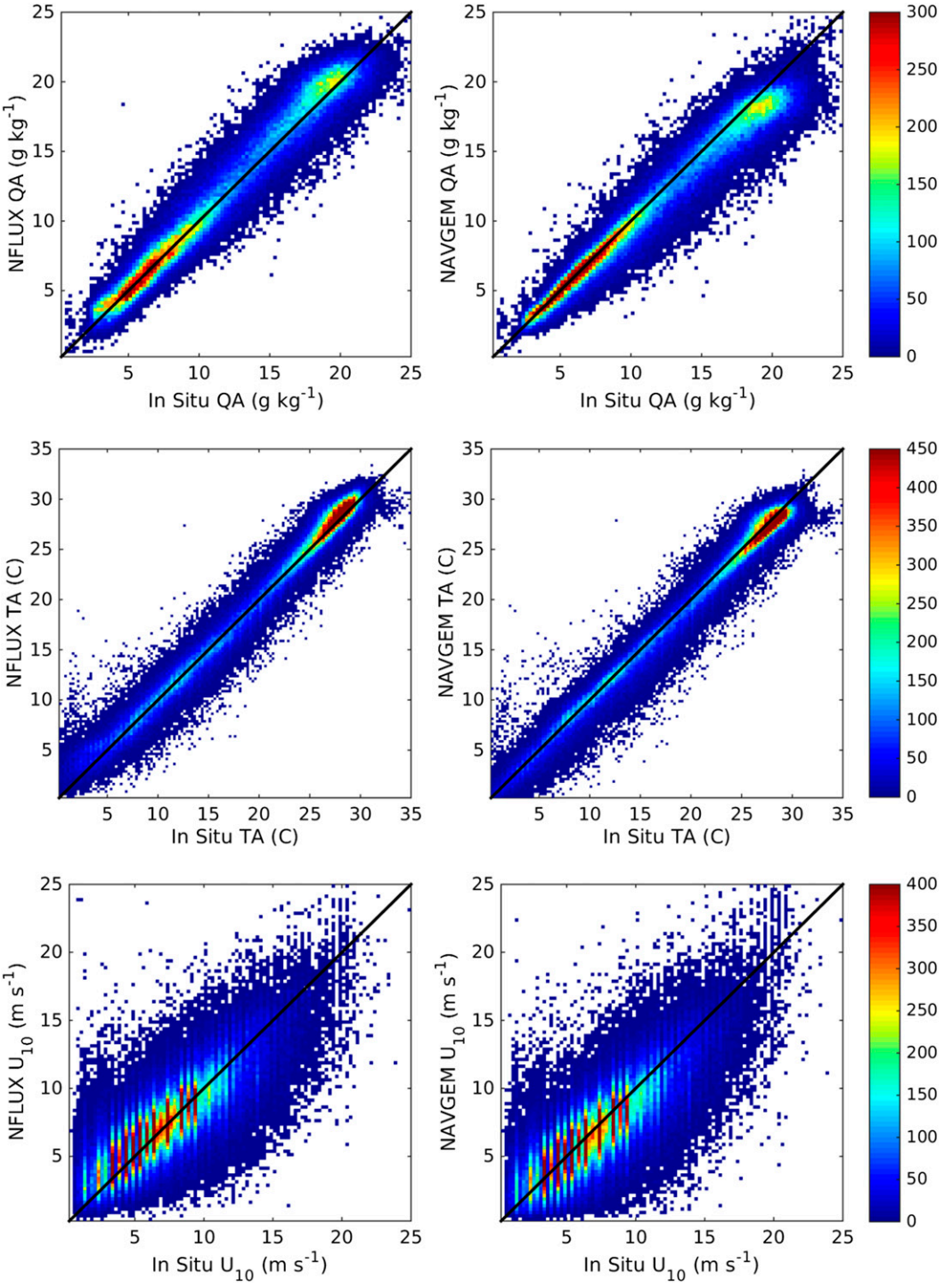

FIG. 3. (left) NFLUX and (right) NAVGEM (top) specific humidity, (middle) air temperature, and (bottom) wind speed vs in situ observations. The plots represent 2D histograms, with the colors showing the number of observations within each 0.25-unit square bin.

TABLE 4. NFLUX and NAVGEM specific humidity, air temperature, and wind speed error statistics relative to in situ data. The ME, SD, and RMSE specific humidity errors are reported in grams per kilogram, air temperature errors are reported in degrees Celsius, and wind speed errors are reported in meters per second. A negative ME indicates an underestimation by the analysis.

	ME	SD	RMSE	R^2	SS
NFLUX QA	0.3274	1.2019	1.2457	0.9578	0.3742
NAVGEM QA	−0.4896	1.1798	1.2773	0.9579	0.3420
NFLUX TA	0.2408	1.2216	1.2451	0.9789	0.4989
NAVGEM TA	−0.3043	1.2117	1.2494	0.9788	0.4955
NFLUX U_{10}	0.2142	2.0563	2.0674	0.6442	0.5480
NAVGEM U_{10}	−0.3258	2.1421	2.1668	0.6258	0.5035

estimate high values (Van de Voorde et al. 2015). The in situ specific humidity matchups that are greater than or equal to 17 g kg^{-1} (which account for 25% of the total matchups) relative to NFLUX (relative to NAVGEM) have an overall bias of 0.3994 (−1.2209) g kg^{-1} and an overall RMSE of 1.4261 (1.8038) g kg^{-1}. For in situ matchups that are less than 17 g kg^{-1} relative to NFLUX (relative to NAVGEM), the overall bias is 0.3030 (−0.2408) g kg^{-1} and the overall RMSE is 1.1781 (1.0391) g kg^{-1}.

Both the NFLUX and NAVGEM TA comparisons have similar RMSE values, with NFLUX showing a slightly lower RMSE from approximately 13° to 23°C. The large positive bias and high RMSE seen in NFLUX for low air temperatures are attributed to the satellite retrievals. At low temperatures, 1) larger scatter is seen in the satellite brightness temperatures and the SSTs and 2) the satellite brightness temperatures are generally less sensitive to temperature changes (Jackson and Wick 2010). If in situ air temperatures of less than 5°C (which account for 6% of the total matchups) are eliminated from the matchups, NFLUX (NAVGEM) has an overall bias of 0.1820°C (−0.3292°C) and an overall RMSE of 1.1912°C (1.2421°C).

The NFLUX U_{10} comparisons have a higher RMSE at low wind speeds and a lower RMSE at high wind speeds relative to NAVGEM. Both NFLUX and NAVGEM have a high positive bias and high RMSE at low wind speeds. For scatterometer retrievals, the uncertainty in positive scalar quantities is largest at low wind speeds; therefore, there is often a low wind cutoff applied at 3 m s^{-1} (Freilich and Dunbar 1999; Bourassa et al. 2003). The SSMI and SSMIS satellite retrievals exhibit similar degraded performance at low wind speeds (Goodberlet et al. 1989; Bentamy et al. 1999; Van de Voorde et al. 2015). At this time, NFLUX does not apply a low wind cutoff for the satellite wind speed assimilation. Retrievals of wind speed at low values are less reliable and degrade the overall NFLUX performance in these regions. NAVGEM and, to a lesser extent, NFLUX have an increased negative bias and increased RMSE as wind speed increases. This pattern has also been documented in other studies such as Bentamy et al. (1999) and Roberts et al. (2010). The improvement NFLUX

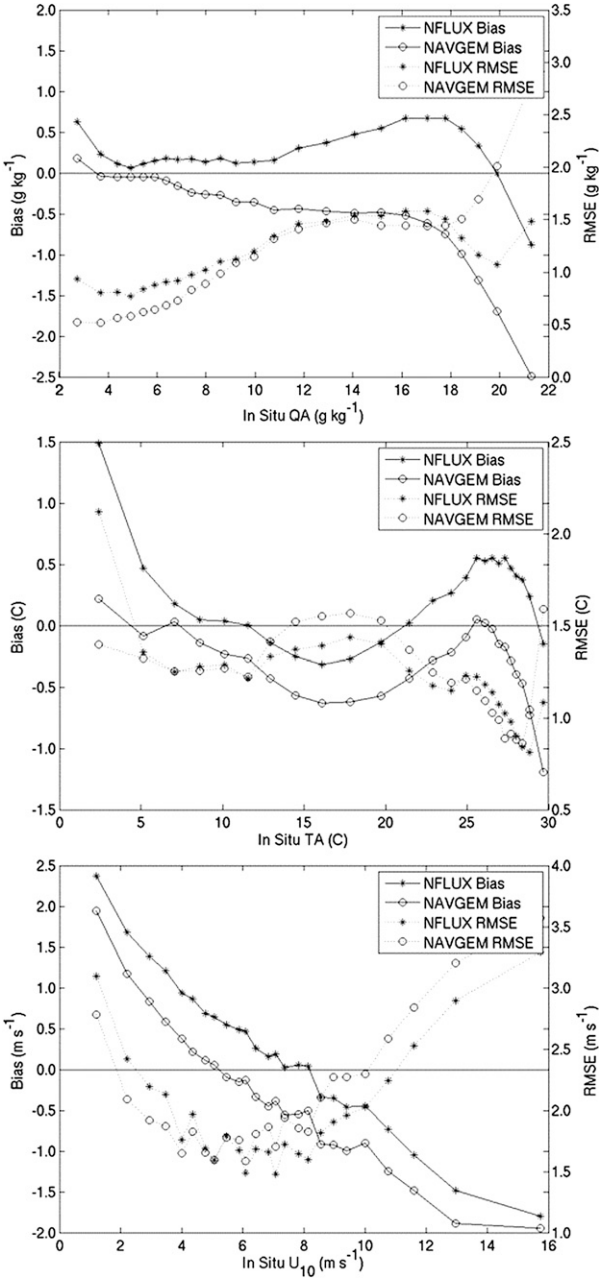

FIG. 4. NFLUX and NAVGEM bias and RMSE statistics computed for the observed range of (top) specific humidity, (middle) air temperature, and (bottom) wind speed. The in situ observations were sorted and divided into 25 equally populated bins for each surface state parameter.

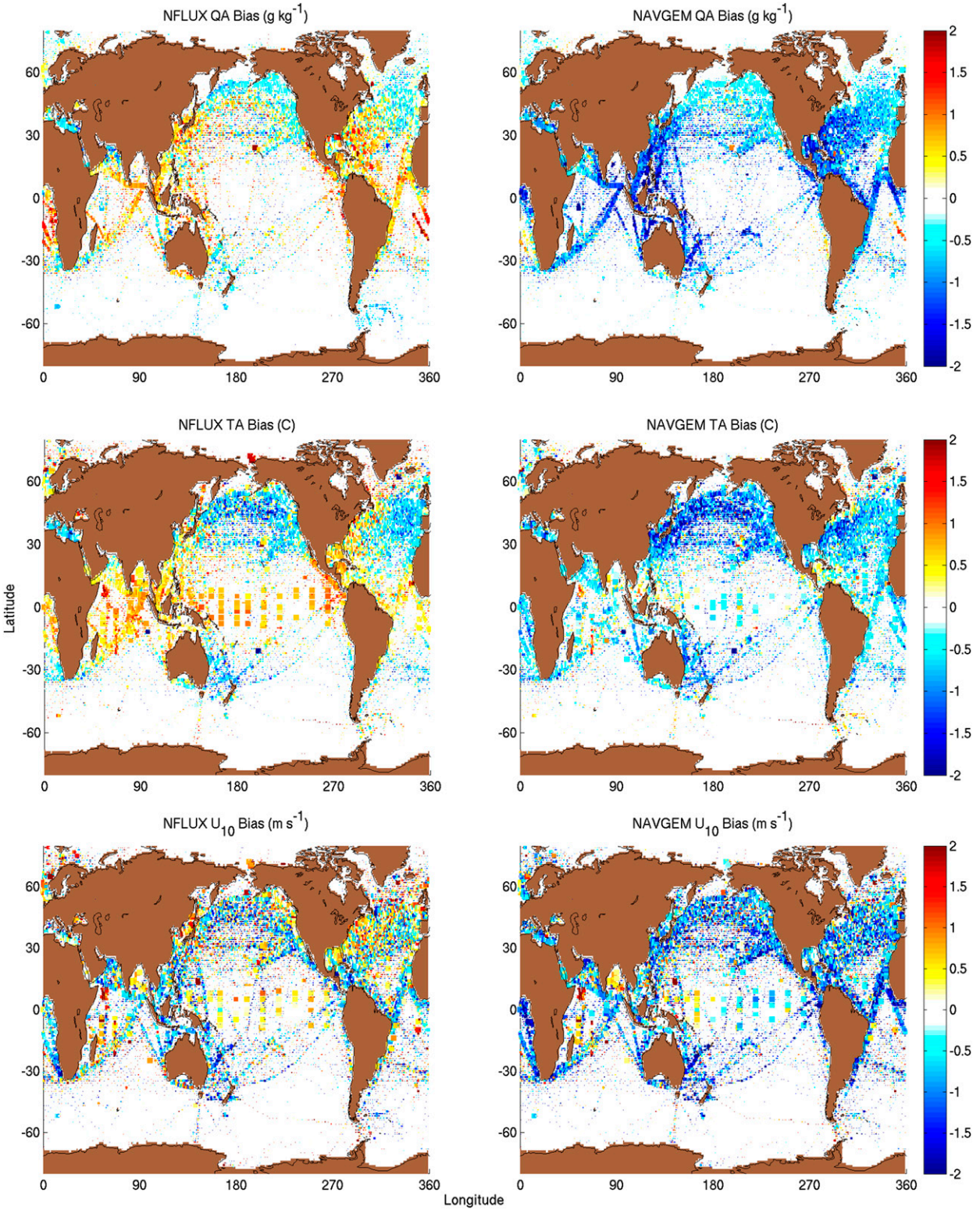

FIG. 5. (left) NFLUX and (right) NAVGEM 1-yr-average (top) specific humidity, (middle) air temperature, and (bottom) wind speed bias relative to in situ observations. The sizes of the squares represent the number of observations in each grid box, ranging from 5 to 50.

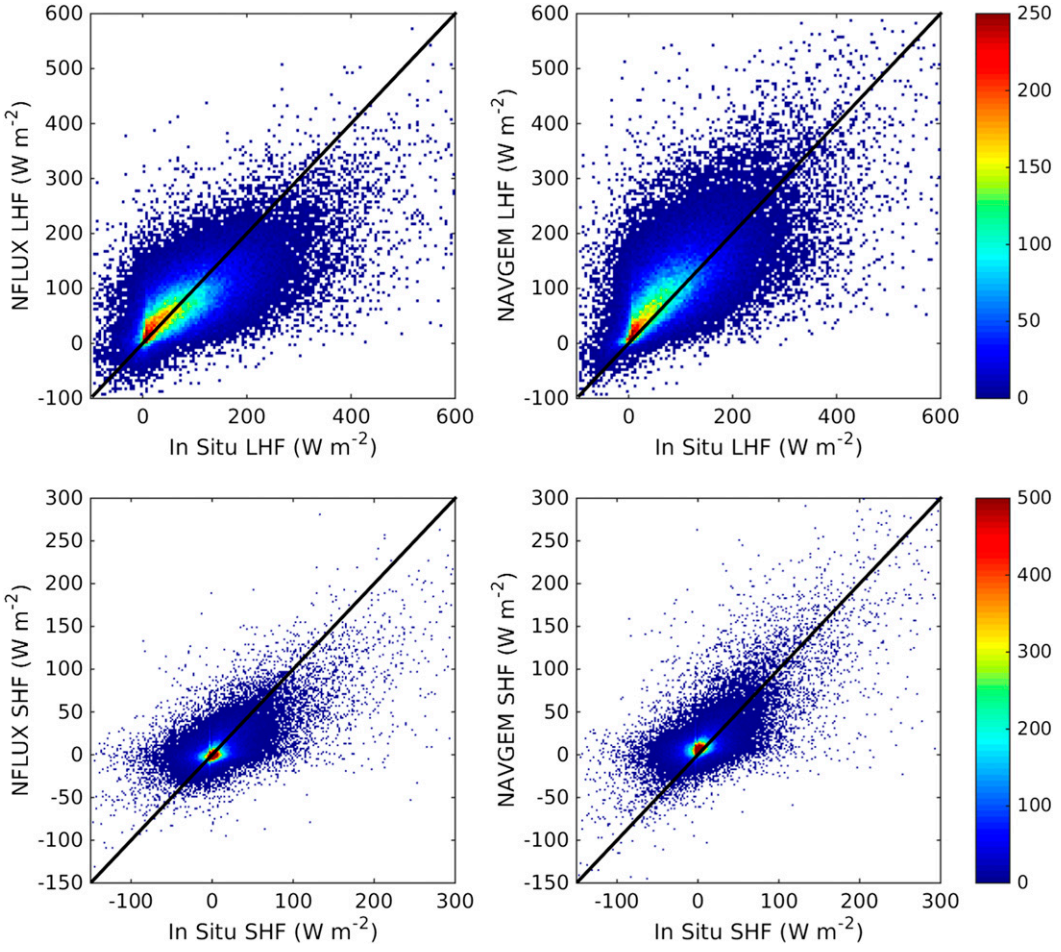

FIG. 6. (left) NFLUX and (right) NAVGEM (top) latent heat flux and (bottom) sensible heat flux vs in situ observations. The latent heat flux (sensible heat flux) plots represent 2D histograms, with the colors showing the number of observations within each 5 (2) W m^{-2} square bin.

shows over NAVGEM at high wind speeds is a result of the use of multiple polynomial regression algorithms in the satellite retrievals, similar to the results seen in QA. In situ wind speed matchups less than or equal to 3 m s^{-1} account for 11% of the total matchups, and in situ wind speed matchups greater than 10 m s^{-1} account for 18% of the total matchups. For in situ matchups that are less than or equal to 3 m s^{-1} relative to NFLUX (NAVGEM), the overall bias is 1.8678 (1.3832) m s^{-1} and the overall RMSE is 2.6536 m s^{-1} (2.3499 m s^{-1}). For in situ matchups greater than 3 m s^{-1} and less than or equal to 10 m s^{-1} relative to NFLUX (NAVGEM), the overall bias is 0.3175 m s^{-1} (−0.2701 m s^{-1}) and the overall RMSE is 1.7564 m s^{-1} (1.8647 m s^{-1}). For in situ matchups that are greater than or equal to 10 m s^{-1} relative to NFLUX (NAVGEM), the overall bias is −1.2120 m s^{-1} (−1.5942 m s^{-1}) and the overall RMSE is 2.7164 m s^{-1} (3.0168 m s^{-1}).

The NFLUX and NAVGEM global distributions of the bias relative to in situ data for each surface state parameter are presented in Fig. 5. The biases are computed over the given year and are binned to a uniform 1.0° grid. NFLUX QA shows the largest positive biases (too moist) in the tropics, the Kuroshio in the western Pacific Ocean, and the Gulf Stream in the western Atlantic Ocean. NFLUX QA shows a slight negative bias throughout the midlatitudes. The NAVGEM QA bias is almost exclusively negative (too dry), with the largest negative biases matching the areas with the largest positive biases that are seen with NFLUX QA. NFLUX TA shows positive biases (too warm) throughout the tropics and Gulf Stream with negative biases (too cold) throughout the midlatitudes. NAVGEM TA shows mostly a negative bias (too cold) pattern globally, with the largest negative biases seen in the midlatitudes. The NFLUX U_{10} global bias pattern

TABLE 5. NFLUX and NAVGEM latent heat flux and sensible heat flux error statistics relative to in situ data. The ME, SD, and RMSE turbulent heat flux errors are reported in watts per meter squared. A negative ME indicates an underestimation by the analysis.

	ME	SD	RMSE	R^2	SS
NFLUX LHF	−9.1749	59.6318	60.3331	0.4532	0.3223
NAVGEM LHF	16.6872	62.6359	64.8203	0.4705	0.2177
NFLUX SHF	−0.8003	24.0492	24.0624	0.3861	0.3899
NAVGEM SHF	3.6264	24.1525	24.4231	0.4460	0.3715

is less uniform than either the QA or TA bias pattern, but there is a general overestimation throughout the tropics and western boundary currents and an underestimation in the midlatitudes. The NAVGEM U_{10} global bias pattern shows a general uniform underestimation, with an area of overestimation seen in the western tropics.

b. *Surface turbulent heat fluxes*

Comparisons of the NFLUX and NAVGEM calculated surface turbulent heat flux fields versus the in situ data are presented in Fig. 6, with the corresponding error statistics provided in Table 5. Each of the NFLUX and NAVGEM flux comparisons have a general linear relationship with the in situ data. As LHF or SHF increases, NFLUX exhibits an increasing underestimation in the surface flux, whereas NAVGEM seems to better estimate the higher surface fluxes. NFLUX (NAVGEM) has an overall negative (positive) bias relative to the in situ data. For both LHF and SHF, NFLUX has a smaller absolute bias, lower RMSE, and increased SS relative to NAVGEM.

For a more detailed comparison, the in situ flux values were sorted and split into 25 equally populated bins. The bias and RMSE were then calculated for each bin (Fig. 7). The overall trends for the bias and RMSE for each parameter are similar between NFLUX and NAVGEM, with the NFLUX bias showing an increasing negative offset relative to the NAVGEM bias. NFLUX has a lower RMSE for LHF relative to NAVGEM until approximately 150 W m^{-2}, and then NFLUX has a higher RMSE. NFLUX has a similar RMSE for SHF throughout the sampled range when compared with NAVGEM.

The NFLUX and NAVGEM global distributions of the bias relative to in situ data for LHF and SHF are presented in Fig. 8. The biases are computed over the given year and are binned to a uniform 1.0° grid. For both LHF and SHF, NFLUX shows an overall positive bias (too much flux leaving the ocean) throughout the midlatitudes and a negative bias (too much flux entering the ocean) in the tropics and Gulf of Mexico regions. NAVGEM shows an overall positive bias globally for LHF. For SHF, NAVGEM generally shows a positive bias throughout the midlatitudes and a negative bias in the tropics.

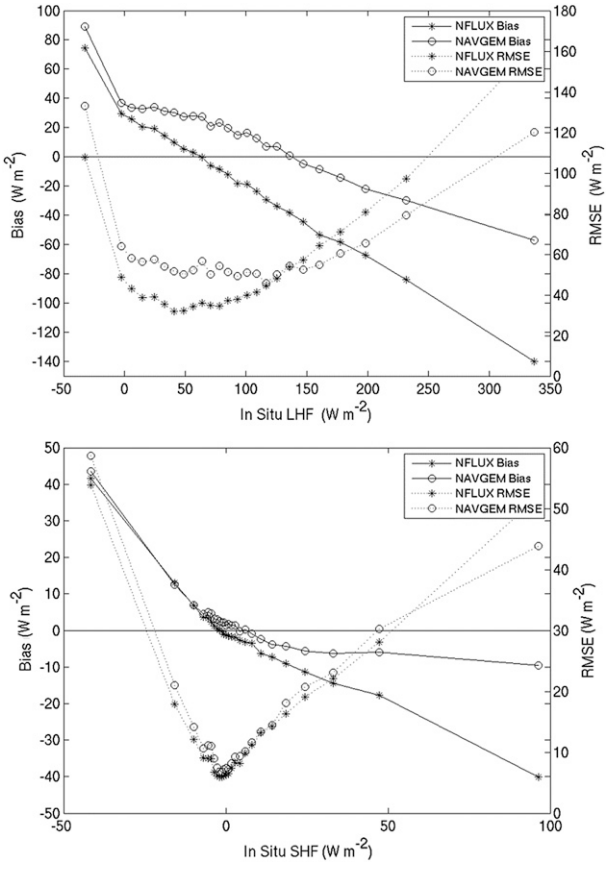

FIG. 7. NFLUX and NAVGEM bias and RMSE statistics computed for the observed range of (top) latent heat flux and (bottom) sensible heat flux. The in situ observations were sorted and evenly divided into 25 equally populated bins for each surface turbulent heat flux parameter.

5. Discussion

Large-scale biases in the state parameter fields, such as those seen in Fig. 5, are largely due to errors in the satellite retrievals. As discussed previously, NFLUX developed four unique algorithms for each retrieved parameter on each sensor–platform combination for ascending versus descending orbits and clear versus cloudy skies. As expected, some of the algorithms perform better than others. Each of the retrieval algorithms also uses a daily SST as a predictor. Jackson and Wick (2010) discuss how satellite retrievals that use SST as a predictor are less likely to capture large SST–TA differences because the retrieval will draw toward the SST

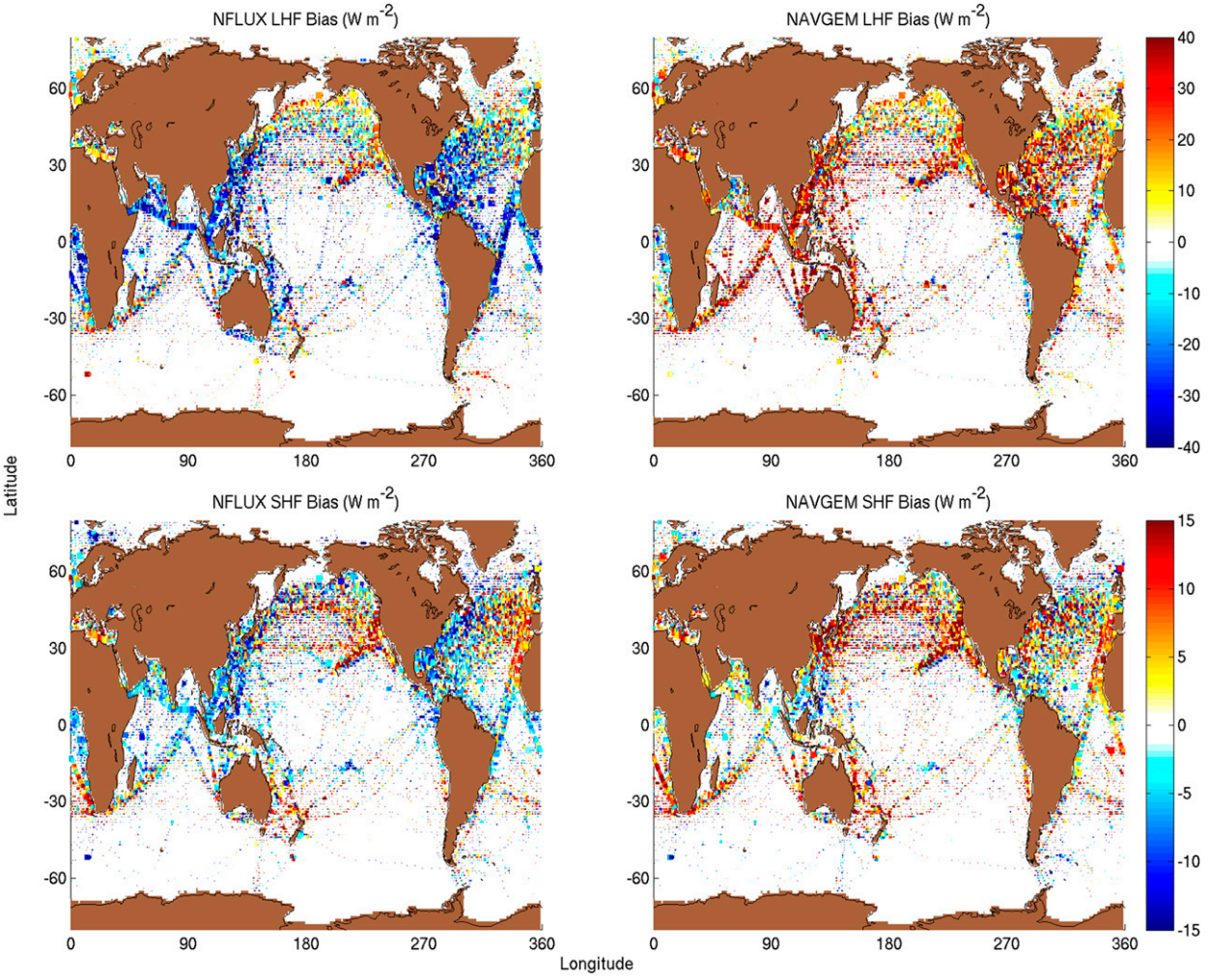

FIG. 8. (left) NFLUX and (right) NAVGEM 1-yr-average (top) latent heat flux and (bottom) sensible heat flux bias relative to in situ observations. The sizes of the squares represent the number of observations in each grid box, ranging from 5 to 50.

value, shallow layers of cold or warm air are difficult for satellite observations to detect, and the large satellite footprint will smooth out small temperature gradients. Even with these issues, satellite retrievals that include SST as a predictor have improved overall results over satellite retrievals that do not include SST as a predictor (Jackson and Wick 2010; Roberts et al. 2010; Van de Voorde et al. 2014).

According to Van de Voorde et al. (2015), the TA descending (nighttime) retrievals generally have improved performance relative to the ascending (daytime) retrievals, with similar performance between the clear and cloudy algorithms. This agrees with the findings from Roberts et al. (2010), who discussed that TA retrievals rely more on the SST input, with only modest improvement when using separate clear- and cloudy-sky algorithms. The improved TA retrievals seen during nighttime could be an indicator that using a diurnally varying SST, instead of a daily SST, would improve the TA estimates. This would most significantly affect areas with large diurnal SSTs, such as the tropics, which are seen to have a large positive bias (Fig. 5).

From Van de Voorde et al. (2015), the QA cloudy-sky and U_{10} clear-sky algorithms generally have improved performance, with similar performance between the ascending and descending algorithms. This result also agrees with the findings from Roberts et al. (2010), who discussed that QA and U_{10} retrievals rely more on the microwave measurements that are affected by cloud liquid water, with small improvements seen when SST is included as an input. NFLUX uses a cloud liquid water threshold value of 0.025 mm for separating cloudy versus clear skies. From the QA and U_{10} biases seen in

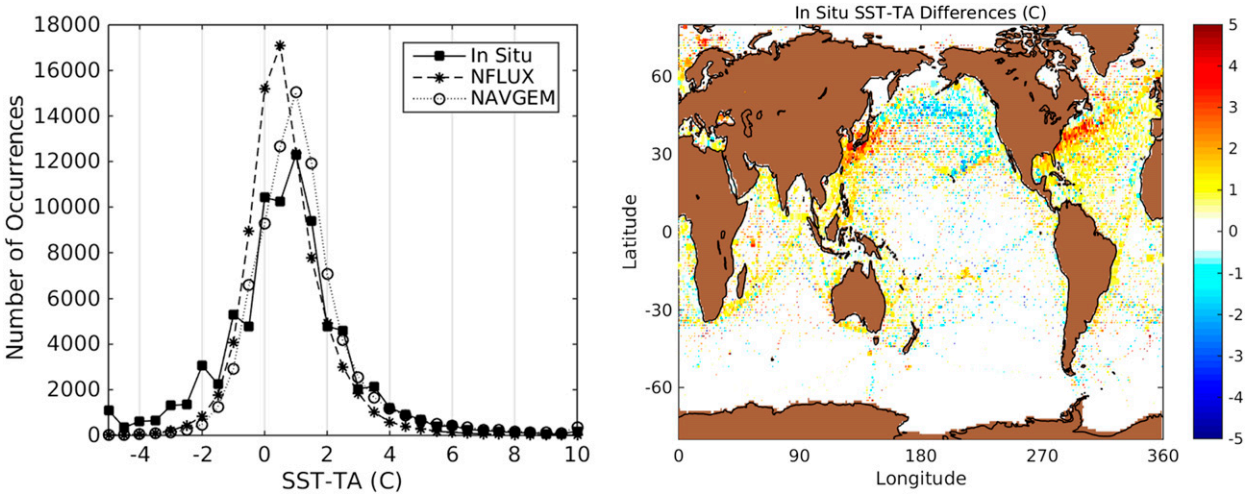

FIG. 9. (left) Histogram of the in situ, NFLUX, and NAVGEM SST–TA differences used in calculating the surface turbulent heat fluxes. Bins have a width of 0.5°C. (right) In situ 1-yr-average SST–TA differences. The sizes of the squares represent the number of observation in each grid box, ranging from 5 to 50.

Fig. 5, a reassessment of the cloud liquid water threshold value should be considered. This would likely have the largest impact on areas with significant cloud coverage, such as the intertropical convergence zone and areas of low pressure. Similar to the TA retrievals, there is also expected to be a slight improvement when including diurnal SSTs, as opposed to daily SSTs, in the QA and U_{10} satellite retrievals.

Several recent studies (Bourras 2006; Brunke et al. 2011; Smith et al. 2011) have performed comparisons of the different versions of the existing satellite-based datasets and have shown that the differences in the QA and TA fields are the main contributors to differences in the LHF and SHF fields. This relationship is easily seen in this study by comparing the global QA and TA bias patterns from Fig. 5 with the global flux bias patterns from Fig. 8. The regions with too much (not enough) moisture or heat are associated with regions of too much flux leaving (entering) the ocean. Therefore, it is expected that by improving the state parameter retrievals, the corresponding LHF and SHF estimates would also improve.

Although the QA and TA fields are the main contributors to differences in the flux fields, it is actually the differences in the surface and near-surface state parameters (q_{sfc}–QA and SST–TA) that are used in estimating the surface fluxes. The q_{sfc} term in the COARE 3.0 algorithms is calculated using the saturated vapor pressure at the SST. With the strong dependence on SST in both LHF and SHF estimates, SST should be considered with respect to the surface fluxes as well. The left panel of Fig. 9 shows a histogram of the SST–TA differences for the in situ data in comparison with those used in NFLUX and NAVGEM. NAVGEM is seen to better represent the number of occurrences throughout the sampled stability range than does NFLUX. The global distribution of the in situ SST–TA differences is presented in the right panel of Fig. 9. The SST–TA differences are computed over the given year and are binned to a uniform 1.0° grid. The largest areas of SST–TA differences, greater than 3°C, are associated with the western boundary currents.

The in situ SST–TA differences are divided into five bins, and the bias and RMSE are calculated and

TABLE 6. NFLUX and NAVGEM specific humidity and latent heat flux errors relative to in situ data, split into SST–TA difference bins.

		QA bias (g kg^{-1})		QA RMSE (g kg^{-1})		LHF bias (W m^{-2})		LHF RMSE (W m^{-2})	
SST–TA (°C)	N	NFLUX	NAVGEM	NFLUX	NAVGEM	NFLUX	NAVGEM	NFLUX	NAVGEM
From −5 to −1	10 784	0.0588	−0.6378	1.3531	1.5060	27.8152	42.5576	55.4896	72.2622
From −1 to 0	12 210	−0.0322	−0.7450	1.2026	1.4251	11.1612	28.2728	44.3939	57.1837
0–1	26 045	0.2353	−0.8726	1.2440	1.5385	−9.3059	17.0984	48.9121	55.8690
1–3	23 611	0.3104	−0.9494	1.2923	1.6025	−25.1870	4.6896	60.6867	59.2398
3–10	7746	0.3566	−1.0881	1.3355	1.6899	−50.8749	−9.7836	97.6058	89.0471

TABLE 7. NFLUX and NAVGEM air temperature and sensible heat flux errors relative to in situ data, split into SST–TA difference bins.

		TA bias (°C)		TA RMSE (°C)		SHF bias (W m^{-2})		SHF RMSE (W m^{-2})	
SST–TA (°C)	*N*	NFLUX	NAVGEM	NFLUX	NAVGEM	NFLUX	NAVGEM	NFLUX	NAVGEM
From −5 to −1	10 784	−0.9799	−1.3309	1.8778	2.2245	20.7018	21.2051	29.5924	30.8920
From −1 to 0	12 210	−0.3026	−0.5929	1.2089	1.3860	8.1070	9.2464	15.7426	18.3128
0–1	26 045	0.0073	−0.4089	1.0027	1.1334	0.6107	3.0941	10.6038	12.6976
1–3	23 611	0.3790	−0.2171	1.1509	1.1999	−8.2693	−3.5275	16.4467	17.9793
3–10	7746	0.9988	−0.4946	1.8748	1.8533	−28.8488	−9.5418	43.5407	40.4704

included for QA and LHF in Table 6 and for TA and SHF in Table 7. NFLUX QA has improved error statistics relative to NAVGEM for each of the stability bins. NFLUX LHF, TA, and SHF have improved error statistics relative to NAVGEM for the first three bins of SST–TA differences: from −5° to −1°C, from −1° to 0°C, and 0°–1°C. From the 1°–3°C bin to the 3°–10°C bin, NFLUX shows increasing degradation in the error statistics relative to NAVGEM. It was previously mentioned that the QA retrievals do not depend greatly on the SST input, which is a fact that is consistent with NFLUX QA having improved error statistics relative to NAVGEM QA for each stability bin. The other parameters do in fact show a larger dependence on the input SST.

As discussed previously, the same SST fields from NAVGEM are used in both the NFLUX and NAVGEM turbulent heat flux calculations to highlight the differences from the QA, TA, and U_{10} fields. The NAVGEM SST fields do not have diurnal variability, which, depending on the time of day, can dramatically change the SST–TA differences and in turn the LHF and SHF results. Similar to the discussion for the satellite retrievals, it is expected that including a diurnally varying SST in the turbulent heat flux calculations would decrease the associated biases.

6. Conclusions

NFLUX is a new data processing and analysis system (built on the existing NCODA system) that provides satellite-based global gridded surface state parameter fields in near–real time. Surface turbulent heat flux fields are then estimated from the COARE 3.0 bulk algorithms using the satellite-based surface state parameters as input. This study uses NAVGEM to provide the atmospheric background and SST fields, but the NFLUX system can use atmospheric forecasts from other sources. NFLUX is designed to be an alternative to using NWP-model fields, such as NAVGEM, to provide the surface forcing for operational ocean models and to provide a mechanism for using satellite and in situ observations of the air–sea interface to assess and monitor the NWP products.

The NFLUX products compare well to both in situ data and the NAVGEM forecast fields. Overall, NFLUX has a lower absolute bias, lower or similar RMSE, and increased skill score for the surface state parameters relative to NAVGEM. The multiple polynomial regression algorithms used in the satellite retrievals are shown to better estimate high surface state values than does NAVGEM, but NFLUX has degraded results relative to NAVGEM at low temperatures (<5°C, which accounts for 6% of the total matchups) and low wind speeds (<3 m s^{-1}, which accounts for 11% of the total matchups) because of known issues in satellite retrievals. The NFLUX LHF and SHF comparisons show a lower overall absolute bias, lower overall RMSE, and an increased overall skill score relative to NAVGEM. At large SST–TA differences, NFLUX has degraded error statistics relative to NAVGEM as a result of including daily SST as a predictor in the satellite retrievals and flux estimates.

Even with the known issues in the satellite retrievals, NFLUX shows an overall improvement over NAVGEM globally. Future work will investigate the satellite retrievals at low values and high SST–TA differences for further improvement. Using a diurnally varying SST field in the satellite retrievals, as well as the flux calculations, will also be considered. Work is also under way to develop the radiative flux components of the NFLUX system to provide a full set of surface flux measurements.

Acknowledgments. This research is funded by the Office of Naval Research Program Element 62435N under the ''propagation across the coupled turbulent interface'' and ''calibration of ocean forcing with satellite flux estimates'' Naval Research Laboratory projects. The authors thank Drs. Mark Bourassa and Carolyn Reynolds for helpful comments on an early version of this document. The 3-hourly SeaFlux products were obtained from the SeaFlux project (http://seaflux.org/). This paper has been approved for public release.

REFERENCES

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, doi:10.5194/essd-2-215-2010.

Bentamy, A., P. Queffeulou, Y. Quilfen, and K. Katsaros, 1999: Ocean surface wind fields estimated from satellite active and passive microwave instruments. *IEEE Trans. Geosci. Remote Sens.*, **37**, 2469–2486, doi:10.1109/36.789643.

——, S. A. Grodsky, K. Katsaros, A. M. Mestas-Nuñez, B. Blanke, and F. Desbiolles, 2013: Improvement in air–sea flux estimates derived from satellite observations. *Int. J. Remote Sens.*, **34**, 5243–5261, doi:10.1080/01431161.2013.787502.

Berry, D. I., and E. C. Kent, 2011: Air–sea fluxes from ICOADS: The construction of a new gridded dataset with uncertainty estimates. *Int. J. Climatol.*, **31**, 987–1001, doi:10.1002/joc.2059.

Bettenhausen, M. H., C. K. Smith, R. M. Bevilacqua, N.-Y. Wang, P. W. Gaiser, and S. Cox, 2006: A nonlinear optimization algorithm for WindSat wind vector retrievals. *IEEE Trans. Geosci. Remote Sens.*, **44**, 597–610, doi:10.1109/TGRS.2005.862504.

Bourassa, M. A., D. M. Legler, J. J. O'Brien, and S. R. Smith, 2003: SeaWinds validation with research vessels. *J. Geophys. Res.*, **108**, 3019, doi:10.1029/2001JC001028.

Bourras, D., 2006: Comparison of five satellite-derived latent heat flux products to moored buoy data. *J. Climate*, **19**, 6291–6313, doi:10.1175/JCLI3977.1.

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, doi:10.1175/1520-0442(2003)016<0619:WBAAAL>2.0.CO;2.

——, 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 datasets. *J. Climate*, **24**, 5469–5493, doi:10.1175/2011JCLI4223.1.

Chakraborty, A., R. Kumar, and A. Stoffelen, 2013: Validation of ocean surface winds from the *OCEANSAT-2* scatterometer using triple collocation. *Remote Sens. Lett.*, **4**, 84–93, doi:10.1080/2150704X.2012.693967.

Clayson, C. A., J. B. Roberts, and A. S. Bogdanoff, 2015: The SeaFlux turbulent flux dataset version 1.0 documentation. SeaFlux Data Rep., 5 pp. [Available online at http://seaflux.org/seaflux_data/DOCUMENTATION/SeaFluxDocumentationV12.pdf.]

Cummings, J. A., 2005: Operational multivariate ocean data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3583–3604, doi:10.1256/qj.05.105.

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

Drennan, W. M., J. A. Zhang, J. R. French, C. McCormick, and P. G. Black, 2007: Turbulent fluxes in the hurricane boundary layer. Part II: Latent heat flux. *J. Atmos. Sci.*, **64**, 1103–1115, doi:10.1175/JAS3889.1.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, 1996: Bulk parameterization of air-sea fluxes in TOGA COARE. *J. Geophys. Res.*, **101**, 3747–3767, doi:10.1029/95JC03205.

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

Fennig, K., A. Andersson, S. Bakan, C. Klepp, and M. Schroeder, 2012: Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite data—HOAPS 3.2—Monthly means/6-hourly composites. Satellite Application Facility on Climate Monitoring, accessed 7 May 2014, doi:10.5676/EUM_SAF_CM/HOAPS/V001.

Figa-Saldaña, J., J. J. W. Wilson, E. Attema, R. Gelsthorpe, M. R. Drinkwater, and A. Stoffelen, 2002: The Advanced Scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers. *Can. J. Remote Sens.*, **28**, 404–412, doi:10.5589/m02-035.

Freilich, M. H., and R. S. Dunbar, 1999: The accuracy of the NSCAT 1 vector winds: Comparisons with National Data Buoy Center buoys. *J. Geophys. Res.*, **104**, 11 231–11 246, doi:10.1029/1998JC900091.

Goodberlet, M. A., C. T. Swift, and J. C. Wilkerson, 1989: Remote sensing of ocean surface winds with the special sensor microwave/imager. *J. Geophys. Res.*, **94**, 14 547–14 555, doi:10.1029/JC094iC10p14547.

Hilburn, K. A., and C.-L. Shie, 2011: Decadal trends and variability in Special Sensor Microwave Imager (SSM/I) brightness temperatures and Earth incidence angle. Remote Sensing Systems Rep. 092811, 53 pp. [Available online at http://www.remss.com/papers/tech_reports/hilburn_SSMI_EIA_TB_RSS_Tech_Rpt_092811.pdf.]

Hogan, T. F., and Coauthors, 2014: The Navy Global Environmental Model. *Oceanography*, **27**, 116–125, doi:10.5670/oceanog.2014.73.

Iwasaki, S., M. Kubota, and H. Tomita, 2010: Evaluation of bulk method for satellite-derived latent heat flux. *J. Geophys. Res.*, **115**, C07007, doi:10.1029/2010JC006175.

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

Jensen, T. G., T. J. Campbell, R. A. Allard, R. J. Small, and T. A. Smith, 2011: Turbulent heat fluxes during an intense cold-air outbreak over the Kuroshio Extension region: Results from a high-resolution coupled atmosphere–ocean model. *Ocean Dyn.*, **61**, 657–674, doi:10.1007/s10236-011-0380-0.

Kent, E. C., S. D. Woodruff, and D. I. Berry, 2007: Metadata from WMO Publication No. 47 and an assessment of voluntary observing ship observation heights in ICOADS. *J. Atmos. Oceanic Technol.*, **24**, 214–234, doi:10.1175/JTECH1949.1.

Mo, T., 1999: AMSU-A antenna pattern corrections. *IEEE Trans. Geosci. Remote Sens.*, **37**, 103–112, doi:10.1109/36.739131.

Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417–2424, doi:10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2.

——, 1995: The coefficients of correlation and determination as measures of performance in forecast verification. *Wea. Forecasting*, **10**, 681–688, doi:10.1175/1520-0434(1995)010<0681:TCOCAD>2.0.CO;2.

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

Rosmond, T., and L. Xu, 2006: Development of NAVDAS-AR: Non-linear formulation and outer loop tests. *Tellus*, **58A**, 45–58, doi:10.1111/j.1600-0870.2006.00148.x.

Shie, C.-L., 2012: Science background for the reprocessing and Goddard Satellite-based Surface Turbulent Fluxes (GSSTF3) dataset for global water and energy cycle research. Goddard Earth Sciences Data and Information Services Center Rep., 22

pp. [Available online at http://disc.sci.gsfc.nasa.gov/measures/documentation/Science_of_the_data.GSSTF3.pdf.]

——, and Coauthors, 2009: A note on reviving the Goddard Satellite-based Surface Turbulent Fluxes (GSSTF) dataset. *Adv. Atmos. Sci.*, **26**, 1071–1080, doi:10.1007/s00376-009-8138-z.

Small, R. J., and Coauthors, 2008: Air–sea interaction over ocean fronts and eddies. *Dyn. Atmos. Oceans*, **45**, 274–319, doi:10.1016/j.dynatmoce.2008.01.001.

Smith, S. R., P. J. Hughes, and M. A. Bourassa, 2011: A comparison of nine monthly air–sea flux products. *Int. J. Climatol.*, **31**, 1002–1027, doi:10.1002/joc.2225.

Tomita, H., M. Kubota, M. F. Cronin, S. Iwasaki, M. Konda, and H. Ichikawa, 2010: An assessment of surface heat fluxes from J-OFURO2 at the KEO and JKEO sites. *J. Geophys. Res.*, **115**, C03018, doi:10.1029/2009JC005545.

Van de Voorde, N., J. May, and C. Rowley, 2014: Validation test report for NFLUX PRE: Validation of specific humidity, surface air temperature, and wind speed precision and accuracy for assimilation into global and regional models. NRL Rep. NRL/MR/7320–14-9523, 55 pp. [Available online at http://www7320.nrlssc.navy.mil/pubs/2014/rowley1-2014.pdf.]

——, ——, and ——, 2015: NFLUX PRE: Validation of new specific humidity, surface air temperature, and wind speed algorithms for ascending/descending directions and clear or cloudy conditions. NRL Rep. NRL/MR/7320–15-9611, 30 pp. [Available online at http://www7320.nrlssc.navy.mil/pubs/2015/may-2015.pdf.]

Verspeek, J., A. Stoffelen, A. Verhoef, and M. Portabella, 2012: Improved ASCAT wind retrieval using NWP ocean calibration. *IEEE Trans. Geosci. Remote Sens.*, **50**, 2488–2494, doi:10.1109/TGRS.2011.2180730.

Xu, H., M. Xu, S.-P. Xie, and Y. Wang, 2011: Deep atmospheric response to the spring Kuroshio over the East China Sea. *J. Climate*, **24**, 4959–4972, doi:10.1175/JCLI-D-10-05034.1.

Xu, L., T. Rosmond, and R. Daley, 2005: Development of NAVDAS-AR: Formulation and initial tests of the linear problem. *Tellus*, **57A**, 546–559, doi:10.1111/j.1600-0870.2005.00123.x.

Yelland, M. J., R. W. Pascal, P. K. Taylor, and B. I. Moat, 2009: AutoFlux: An autonomous system for the direct measurement of the air-sea fluxes of CO_2, heat and momentum. *J. Oper. Oceanogr.*, **2**, 15–23, doi:10.1080/1755876X.2009.11020105.

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