J-OFURO3

Official Document

Quality Check System for J-OFURO3

Hiroyuki Tomita[*1]
Tsutomu Hihara[*2]

[*1] Institute for Space-Earth Environmental Research, Nagoya University
[*2] JAMSTEC, Application Lab

V1.0E (Dec-01, 2017)

Document author: Hiroyuki Tomita and Tsutomu Hihara

Document ID: J-OFURO3_DOC_006

Document version: V1.0E (2017.12.11)

Contents

Abstract

A system for investigating accuracy and reliability of satellite-derived air-sea fluxes and their related physical parameters in third generation data set in Japanese Ocean Flux Data Sets with Use of Remote Sensing Observations third generation data set (J-OFURO3) was developed as Quality Check System (QCS). In order to conduct systematic verification for gridded data sets, QCS encompasses in situ data set and the program code for verification. As a result, verification can be conducted with only simple setting, the results can be confirmed by a web browser. The current QCS contains 814991 days in situ data obtained from 178 buoys located in the world oceans during 1972-2014. By using this system, it is easy to confirm the difference in quality of data sets from previous version, to compare multiple data sets with same benchmark.

1. Introduction

In order to investigate the accuracy and reliability of satellite-derived surface flux data, comparison with in situ observation data is indispensable. Along with developing J-OFURO3 data set, we also developed a system for checking the accuracy and reliability of dataset: Quality Check System (QCS). This document aims to provide detailed technical information about QCS. Furthermore, we demonstrate some verification works for J-OFURO3 using QCS.

2. QCS overview

A conceptual diagram of QCS is shown in Figure 1. QCS is a system for verification of the gridded air-sea flux data set and a set of programs and scripts that were developed based on general programming language, drawing and database software. The user can add the dataset to be verified and can set various conditions for comparison. QCS semi-automatically verifies the dataset based on the setting. Most of the results of verification can be checked from a web browser with many graphical outputs.

Table 1 shows the specification of QCS including verification items. Verification is carried out by confirming the average and standard deviation fields of the target dataset and by comparing with in situ observations included the system. Many items are designed so that contents can be expanded flexibly. For example, it is possible to add or limit the target area.

QCS stores in situ observation data obtained from surface buoys that are moored at various places in Pacific, Atlantic and Indian Oceans as in situ observation data. Data processing such as basic QC and time averaging for the buoy data were done, and now more that 300 thousands days of data by 138 buoys over the global oceans can be used in QCS.

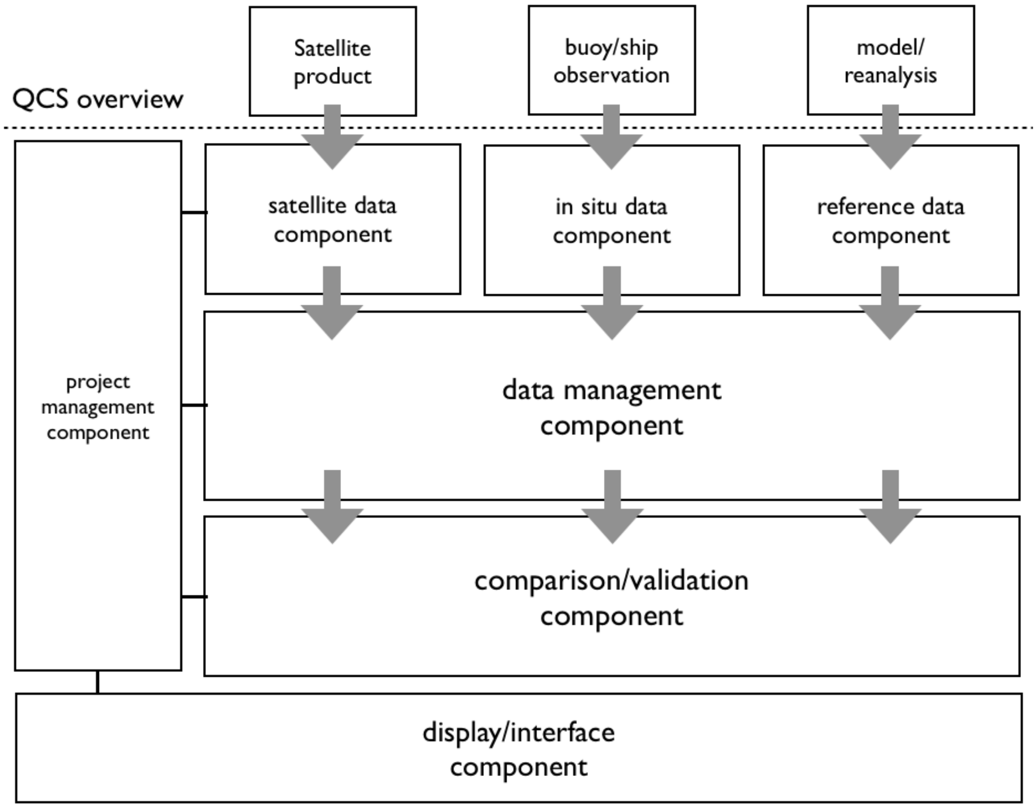


Figure 1. Schematic overview of QCS

Table 1. Specification of QCS

Item	Content	Remarks
Target data format	netCDF	Yearly
Output	HTML files figures: image files (PS or GIF) statistics: text or CSV files	Accessible using web browser
Verification items	Average and standard deviation fields, time series	Performed for each target regions
Target regions	Global, around Japan, Japan Sea	Can be added
In situ observations	Surface moored 138 buoys	Can be added
Temporal resolution for verification	Daily or monthly mean	
Temporal period	1972-2014 (it depends on variables)	Can be expanded
Target variables	Sea surface temperature, surface wind speed, humidity, air temperature, latent heat flux and sensible heat flux	Meteorological parameters are assumed as 10m height value.
Verification items for the comparison with in situ observation	Scatter diagram, time-series Statistics: average, standard deviation, RMS difference, bias, and correlation coefficient	Performed for each buoy and each year
Comparison mode	Normal / inter-comparison modes	See Figure 3

3. Detail

3.1 Preparation

Before starting the verification using QCS, it is necessary to give basic information to QCS on the verification. Necessary items and examples of setting values are summarized in Table 2. These settings need to be described in a text file.

Table 2. Necessary items and examples of setting values for QCS

Main parameter	Description	Format / option	Example setting
MAIN	ID for verification	QCS_ID_nnnn	QCS_ID_0001
INTCOMP	Inter-comparison mode switch	ON/OFF	OFF
TINC	Temporal resolution of target data	DAILY/MONTHLY	DAILY
YEAR	Year (multiple years are acceptable)	YYYY[-YYYY]	2002-2013
VARIABLE	Target variable	LHF/SHF/WND/UWND/VWND/QS/QA/TA/SST	LHF
Sub parameter (settings for each target variable)			
SUB	Sub ID for verification	01～99	01
NAM	Name of target data set	character strings	J-FURO3
VAR	Variable name	LHF/SHF/WND/UWND/VWND/QS/QA/TA/SST	LHF
FILE	File path	With four digit year expressed in YYYY	/data/J-OFURO3_LHF_HR_DAILY_V1.0_YYYY.nc

3.2 Target data

The target dataset to be verified is a gridded data with netCDF format. The verification works can be carried out efficiently if netCDF files are organized by year. The temporal resolution of the target dataset needs to be either the daily mean or the monthly mean.

3.3 In situ data

QCS stores in situ observation data obtained from surface buoys that are moored at various places in Pacific, Atlantic and Indian Oceans. The positions of the buoys stored in QCS are shown in Figure 2a. The data providers of each buoy observations are listed in Table 3. QCS uses high temporal resolution data (hourly or less interval) for surface meteorological parameters (wind speed, air temperature, air humidity, sea level pressure, sea surface temperature) provided by each data provider except for JMA. QCS uses 3 hourly data for JMA, because JMA buoy provides 3 hourly data for their surface moored buoys. Details of the data processing of buoy data will be described in the next section.

Table 3 Buoy data provider and number of buoy

Data Provider	Buoy or data name	Number of buoy	URL
JAMSTEC	JKEO	1	http://www.jamstec.go.jp/iorgc/ocorp/ktsfg/data/jkeo/JKEOdata.htm
JMA	JMA Data Report of Oceanographic Observations Special Issue (moored ocean data buoy)	6	http://www.data.jma.go.jp/gmd/kaiyou/db/vessel_obs/data-report/html/buoy/buoy_NoS2_e.html
NDBC	Historical NDBC Data	54	http://www.ndbc.noaa.gov/historical_data.shtml
NOAA PMEL	Global Tropical Moored Buoy Array : TAO	55	https://www.pmel.noaa.gov/tao/drupal/disdel/
NOAA PMEL	Global Tropical Moored Buoy Array : RAMA	24	https://www.pmel.noaa.gov/tao/drupal/disdel/
NOAA PMEL	Global Tropical Moored Buoy Array : PIRATA	21	https://www.pmel.noaa.gov/tao/drupal/disdel/
JAMSTEC NOAA PMEL	Global Tropical Moored Buoy Array : TRITON	12	https://www.pmel.noaa.gov/tao/drupal/disdel/
NOAA PMEL	Ocean Climate Station: ARC, KEO, Papa	3	https://www.pmel.noaa.gov/ocs/data/disdel/
WHOI	Stratus, SOFS	2	http://uop.whoi.edu/ReferenceDataSets/index.html
	Total	178	

3.4 Data processing for in situ data

The buoy data is provided in various forms, but QCS eventually processes it into the daily and monthly averages in same forms. In this subsection, we describe the procedure of data processing.

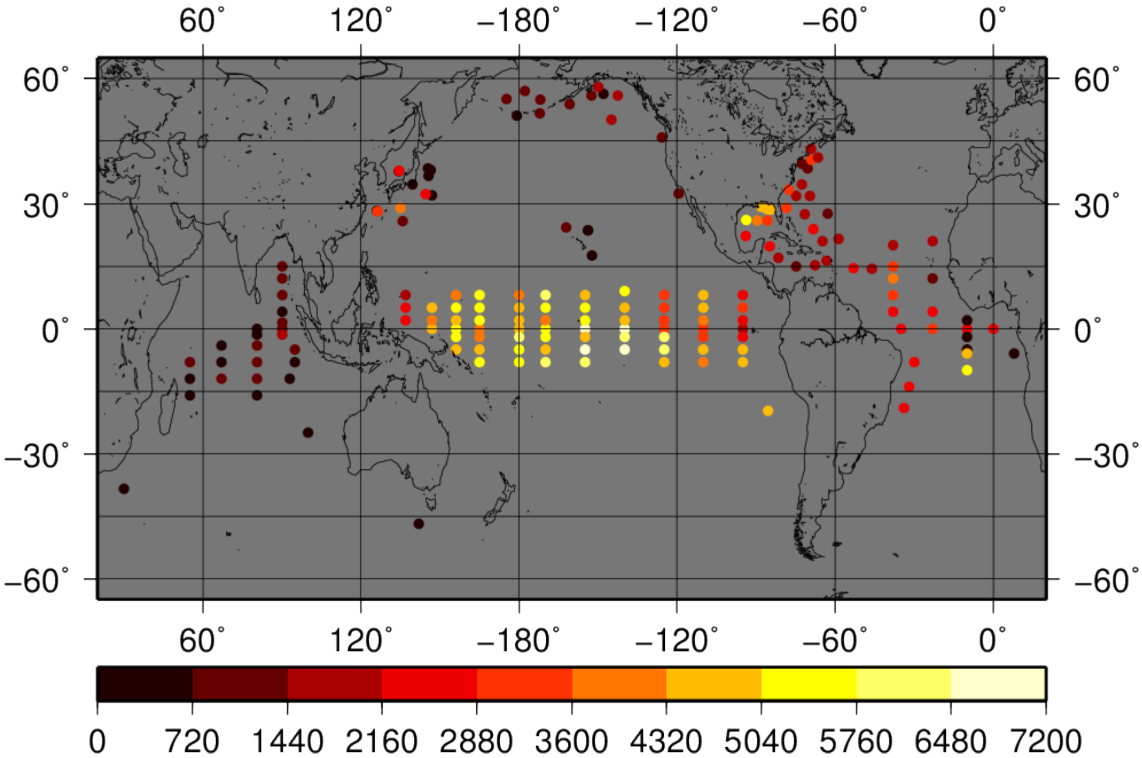
First, basic quality control is performed on the obtained hourly (3 hourly for JMA buoys) in situ buoy data. In the case of every several minutes data, hourly average is calculated then. Turbulent heat flux (latent and sensible heat flux) is then calculated by the bulk flux algorithm, COARE 3.0 (Fairall et al. 2003). The COARE 3.0 requires a set of data of sea surface temperature, surface wind speed, air temperature, humidity and sea level pressure and their observed height to calculate the fluxes. As outputs, the calculated latent and sensible heat fluxes, surface meteorological parameter converted to value at 10m height are stored. No corrections for skin temperature has been made in QCS. The daily and monthly mean values are calculated from the stored hourly data. Figure 2 shows the number of processed daily mean data.

As mentioned above, COARE 3.0 requires several physical parameters to calculate turbulent heat flux. In order to maintain the consistency of the number of data between calculated turbulent heat fluxes and physical parameters, QCS stores output data only when turbulent heat flux is calculated. In other words, in situ data when the turbulent heat flux is not calculated are not used for verification. The verification performed with the set of data in this way is called "FLUX mode" and is a basic verification in QCS. However, exceptionally data of sea surface temperature is stored even if turbulent flux can not be calculated. This is because there are many buoys that historically observe only sea surface temperature. The verification performed with the set of this data is called "SST mode". As a result, in the verification of sea surface temperature there are two modes: FLUX and SST modes, therefore users can distinguish when confirming the results. Table 4 lists examples of the two modes for several input patterns. Pattern 1 is a case that there is no missing in the input physical parameters. Pattern 2 is a case that there is a missing, in which case turbulent heat fluxes are not calculated and outputs for other physical parameters are also missing. Pattern 3 is also a case with missing input physical parameters but data of sea surface temperature is available. In this case, output of sea surface temperature is missing in "FLUX mode", but the output is not missing in "SST mode". If you want to conduct the verification of sea surface temperature data with as much data as possible, ignoring consistency with turbulent heat flux, we recommend referring to the results of "SST mode".

Table 4. Example of FLUX and SST modes

		Sea surface temperature	Surface wind	Humidity	Air temperature	Turbulent heat flux
Pattern 1	Input	o	o	o	o	–
FLUX mode	Output	o	o	o	o	o
SST mode	Output	o	o	o	o	o
Pattern 2	Input	x	o	o	o	–
FLUX mode	Output	x	x	x	x	x
SST mode	Output	x	x	x	x	x
Pattern 3	Input	o	x	o	o	–
FLUX mode	Output	x	x	x	x	x
SST mode	Output	o	x	x	x	x

(a)



(b)

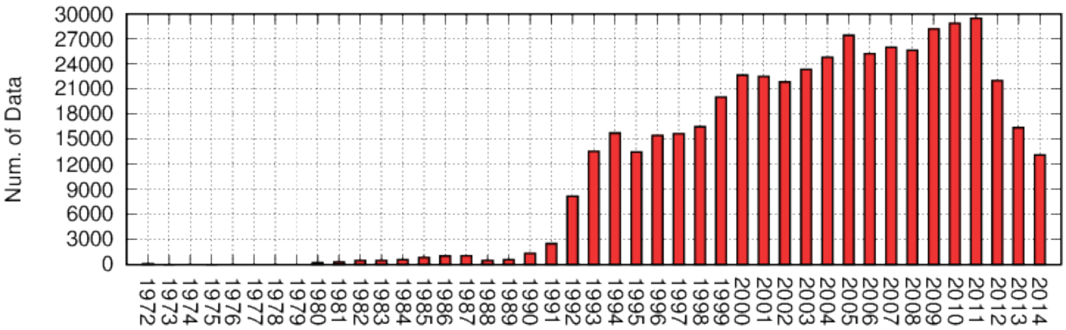


Figure 2. Number of daily mean data stored in QCS. (a) spatial distribution for the period of 1972-2014 and (b) yearly time series in the global ocean. Counting data number is based on flux mode in QCS.

3.5 Data match-up

In comparison between target gridded data and in situ buoy data, what criteria are to be associated with each other is an important aspect in the verification. QCS calculates the representative value at in situ observation from target gridded dataset by spatial interpolation on the same day and compares it with in situ data. The spatial interpolation value is calculated by two-dimensional (latitude and longitude) spatial linear interpolation method.

3.6 Comparison mode

There are two modes of comparison with in situ observation in QCS: normal mode and inter-comparison mode. Figure 3 shows the conceptual difference between two comparison modes. When evaluating a single target dataset, QCS provides exactly same results regardless of which mode is used. However, when dealing with multiple target datasets, it should be noted the difference. If you want to compare multiple datasets to fair, we recommend using inter-comparison mode. On the other hand, if you want to compare individual datasets with in situ observation data that is matched independently, we recommend QCS with normal mode.

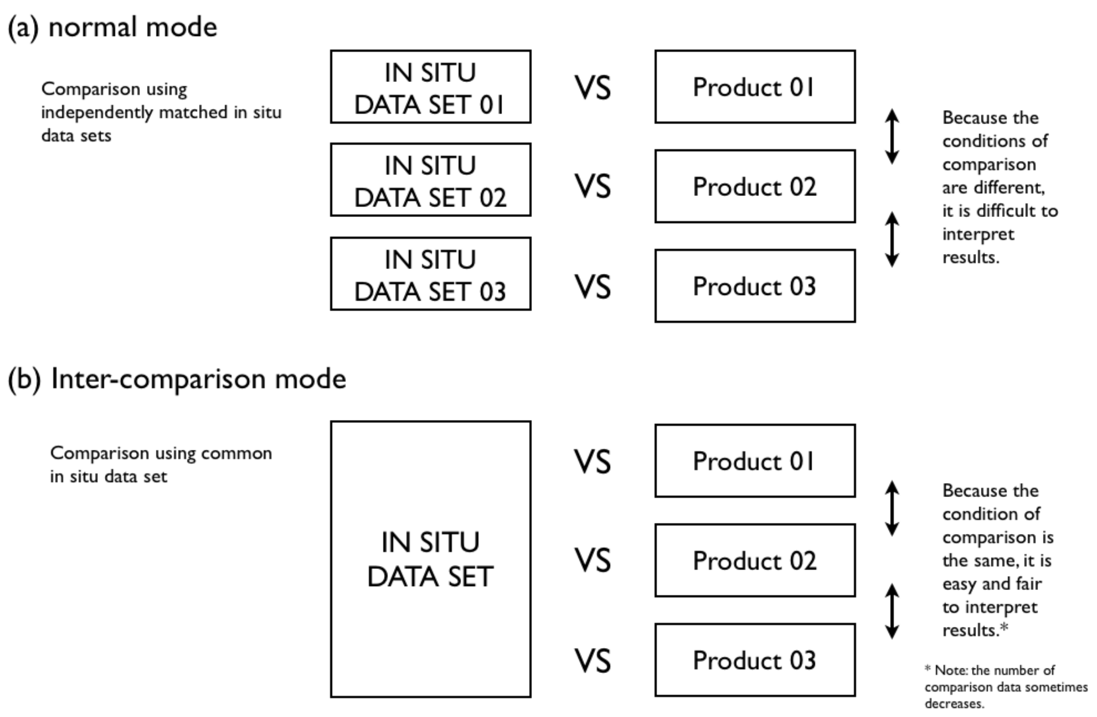
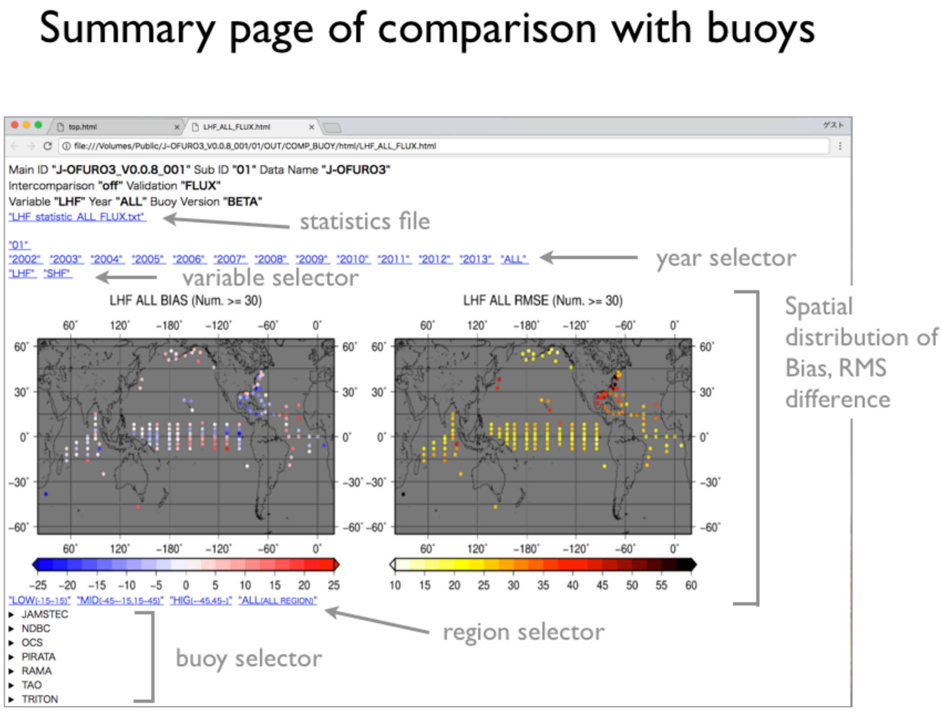


Figure 3. Differences between two comparison modes: normal and inter-comparison modes.

3.7 Outputs

Actual results of verification of the target dataset using QCS are numerical data files and figures. All of results are related by project ID and aggregated into html files. Therefore, results can be confirmed using a web browser. Figure 4 shows examples of display of verification results using QCS.

(a)



(b)

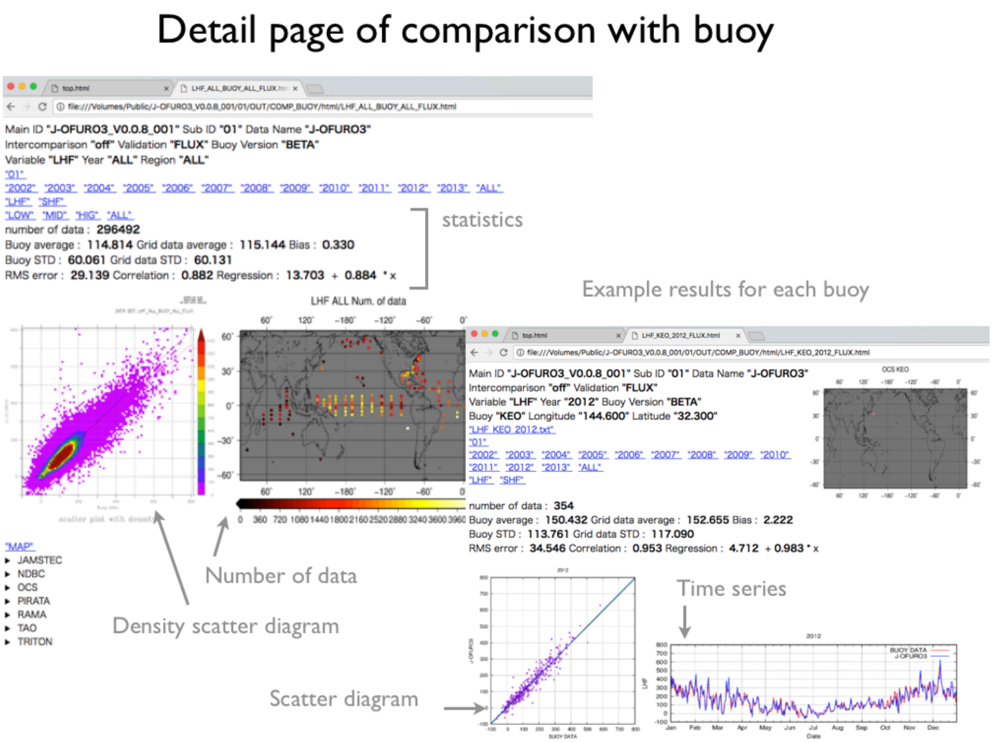


Figure 4. Examples of results of verification using QCS

4. Demonstration

In this chapter, we show the results of verification of J-OFURO3 V1.0 dataset and the inter-comparison among several global products as a demonstration showing what information can be obtained from QCS.

4.1 Verification of J-OFURO3 LHF and SHF

The target dataset for this verification is J-OFURO3 V1.0, with a daily mean and 0.25 degree gridded latent and sensible heat fluxes (LHF and SHF) data sets in 2002-2013. The verification using QCS was implemented as inter-comparison mode: off. Figure 5 shows the spatial distribution of number of daily mean data during the period. Although the number of surface moored buoys and their number of data varies depending on the year and buoy location, the number of daily mean data for 296492 days from over 100 buoys located in the world oceans was used for the verification. Figure 6 shows density scatter diagrams by the comparison with all buoy data. Further more, the density scatter diagrams divided into low-, mid-, and high latitude regions are shown in Figure 7. Table 5 summarizes the comparison statistics for the each comparison. Figure 8 shows the distribution of statistical values. From this figure, you can see the distribution of the bias (J-OFURO3 minus buoy) and RMS difference calculated at each buoy location.

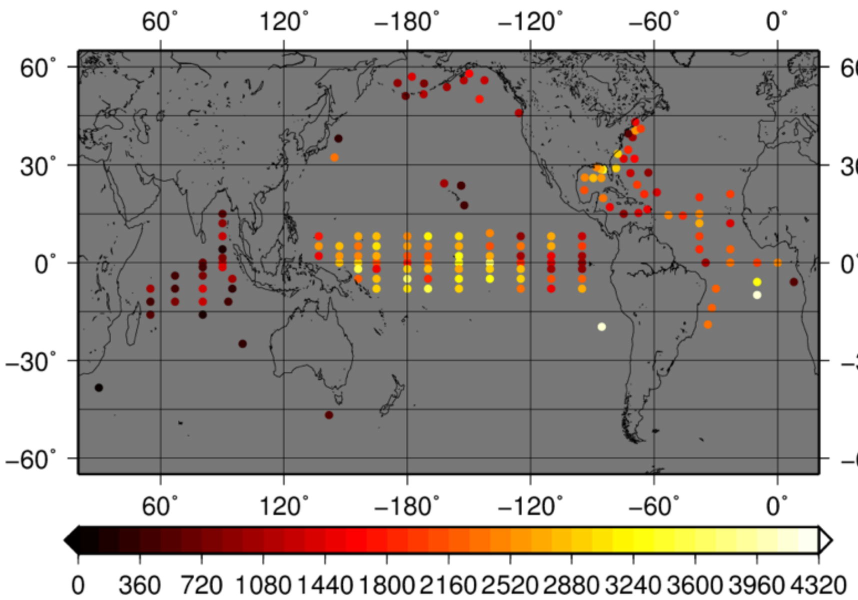


Figure 5. Spatial distribution of number of daily mean data in the verification of J-OFURO3 V1.0 during 2002-2013 using QCS.

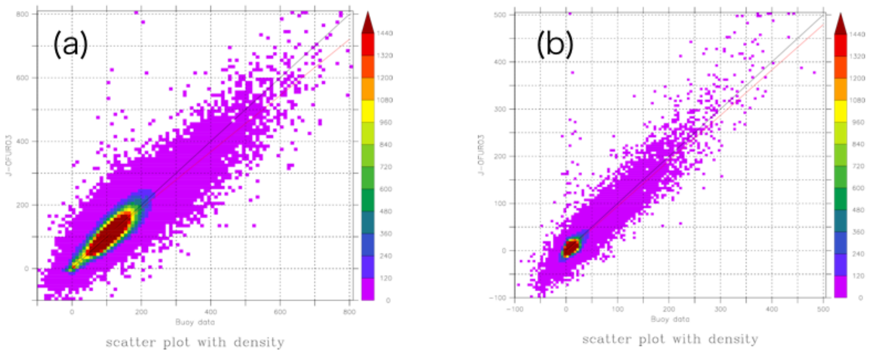


Figure 6. Density scatter-diagrams for (a) LHF and (b) SHF. X and Y axis are showing buoy and J-OFURO3 V1.0 values [W/m^2], respectively. Color means number of data.

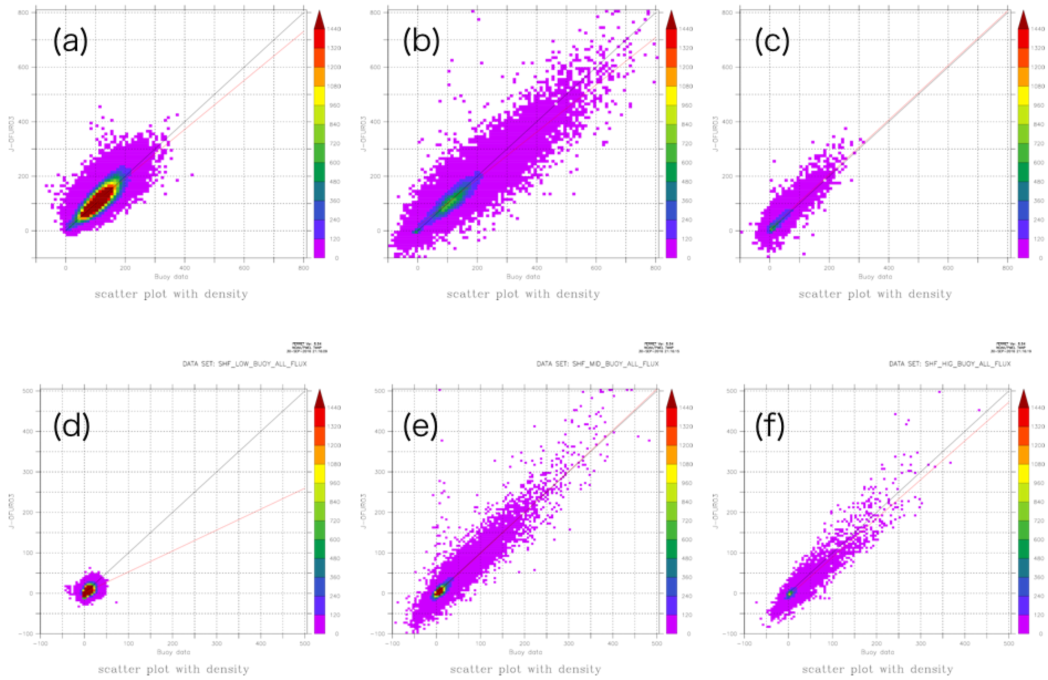


Figure 7. Density scatter-diagrams of LHF (a-c) and SHF (d-f) for each latitudinal region:(a) and (d) are for low latitude region (15S-15N). (b) and (e) are for mid latitude region (15S-45S, 15N-45N). (c) and (f) are for high-latitude region (45S-60S, 45N-60N).Axes and color meaning are same as Figure 6.

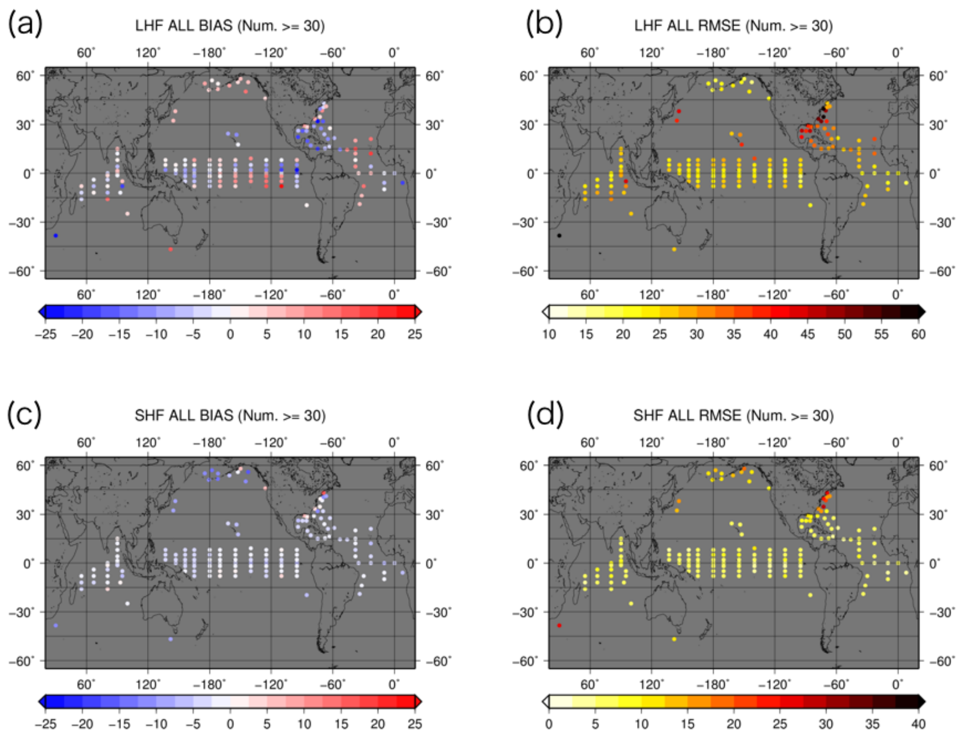


Figure 8 Spatial distributions of comparison statistics for LHF (a-b) and SHF (c-d). (a) and (c) are bias (J-OFURO3 minus buoy). (b) and (d) are RMS difference. Units in W/m^2. The statistics are shown for the buoy that total data number is larger than 30.

Table 5.

4. 2 Inter-comparison of global LHF and SHF datasets

In addition to the verification of individual data set as shown Section 4.1, QCS can also conduct inter-comparison among multiple data sets. The target data sets in the demonstration are five kinds of global data sets including J-OFURO3 V1.0. Table 6 summarizes the information of each data set. The temporal period for the verification is one year of 2008. The verification using QCS was implemented as inter-comparison mode: on. Figure 9 shows number of in situ buoy data and their spatial distribution.

Table 6. Target data sets used in inter-comparison using QCS

Data set name	Data source	Spatial grid size	Version	Reference
J-OFURO3	satellite	0.25 deg.	V1.0	Tomita et al. (in preparation)
J-OFURO2	satellite	0.25 deg.	HF004	Tomita et al., 2010
GSSTF3	satellite	0.25 deg.	v3	Shie, 2012
IFREMER	satellite	0.25 deg.	v3	Bentamy et al., 2013
OAFlux	satellite and atmospheric reanalysis	1 deg.	v3	Yu and Weller, 2007

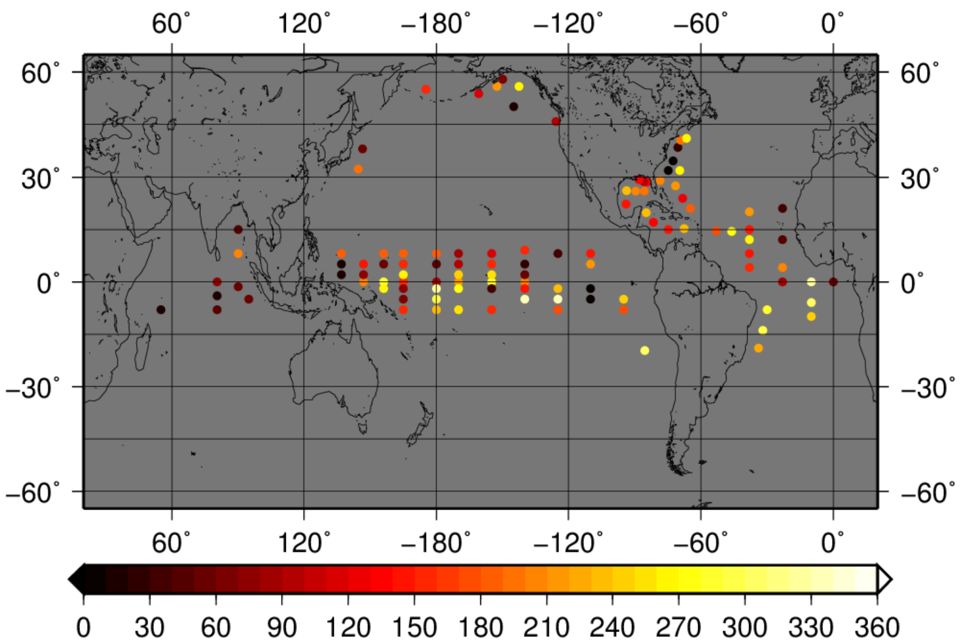


Figure 9. Spatial distribution of number of daily mean data in the inter-comparison during 2008 using QCS.

Figure 10 shows density scatter-diagrams using all buoy data. Differences in rough characteristics of each target data set can be depicted. Furthermore, Figures 11-13 are for three latitudinal zones: low-, mid-, and high-latitude zones. Table 7 summarizes comparison statistics for the inter-comparison obtained from QCS. From these figures and table, it is possible to investigate the difference in characteristics of each target data set by the region.

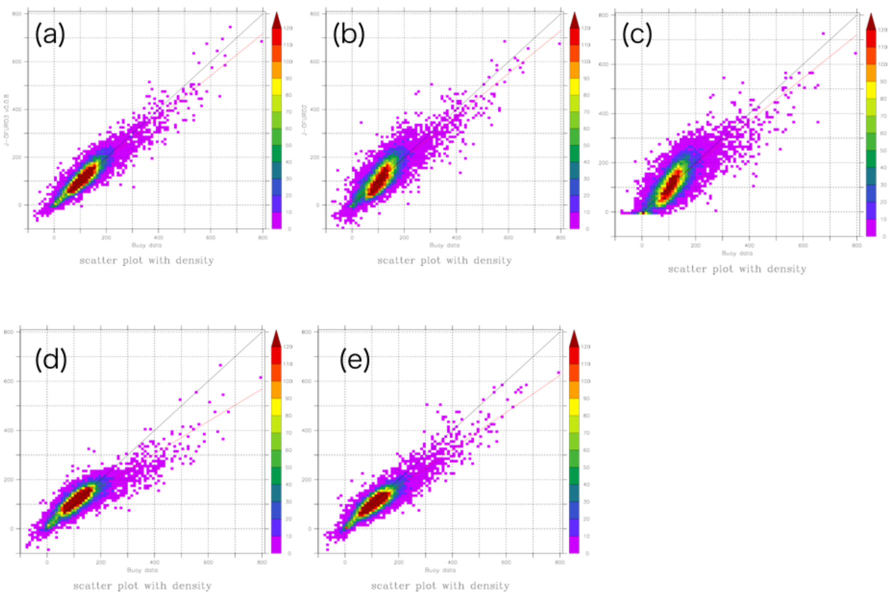


Figure 10. Density scatter-diagrams for each target product: (a) J-OFURO3, (b) J-OFURO2, (c) GSSTF3, (d) IFREMER, and (e) OAFlux. Axes and color meanings are same as Figure 6.

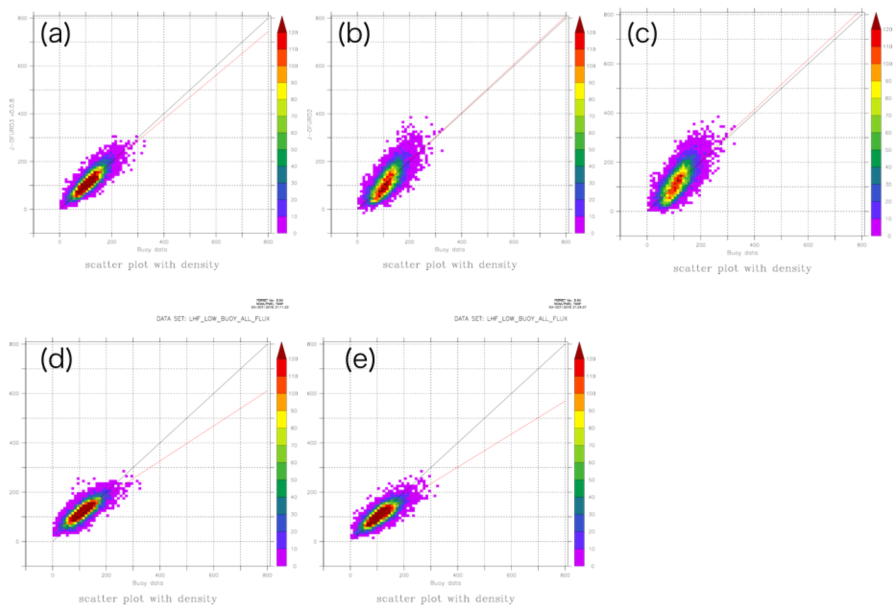


Figure 11. Same as Fig.10 except for low latitude region

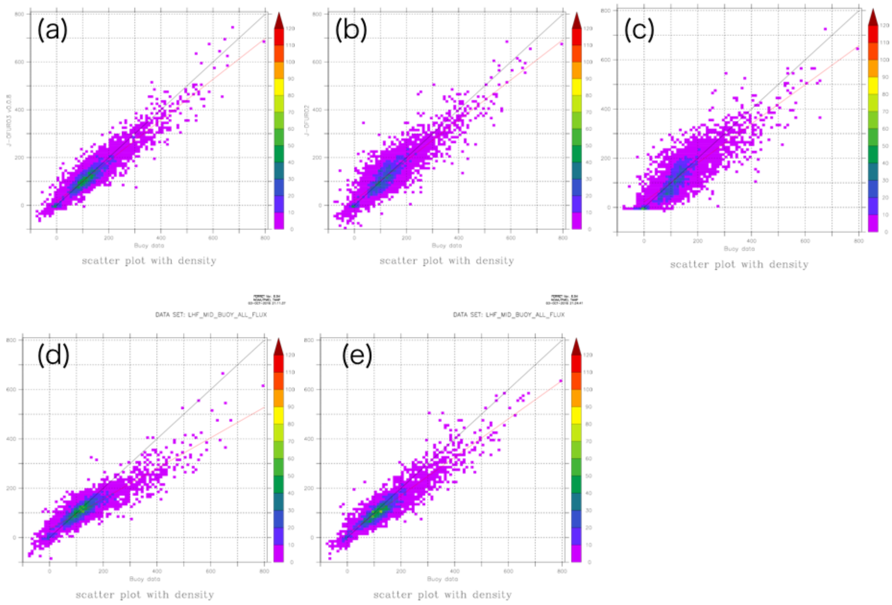


Figure 12. Same as Fig.10 except for mid latitude region

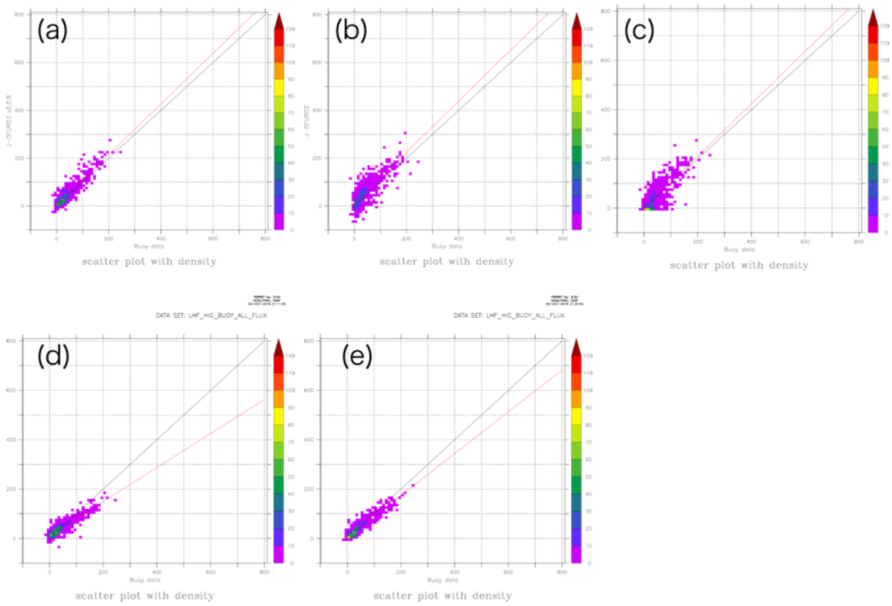


Figure 13. Same as Fig.10 except for high latitude region

Table 7. Inter-comparison statistics obtained from QCS

	Target data	Buoy Ave. [W/m^2]	Ave. [W/m^2]	Buoy Std. [W/m^2]	Std. [W/m^2]	BIAS [W/m^2]	RMS [W/m^2]	CC	NUM
All	J-OFURO3	115.65	116.77	61.44	59.74	1.12	26.82	0.90	17064
	J-OFURO2		117.65		66.80	2.00	38.60	0.82	
	GSSTF3		120.04		68.57	4.39	43.16	0.79	
	IFREMER		117.19		47.97	1.53	33.01	0.85	
	OAFlux		109.15		52.04	−6.50	28.44	0.89	
High Lat.	J-OFURO3	40.20	46.23	37.72	44.24	6.03	19.18	0.90	919
	J-OFURO2		54.77		51.22	14.58	31.71	0.79	
	GSSTF3		41.59		50.75	1.39	31.18	0.79	
	IFREMER		39.84		29.25	−0.36	17.93	0.89	
	OAFlux		39.43		34.54	−0.77	14.44	0.92	
Mid Lat.	J-OFURO3	136.56	131.66	93.58	86.09	−4.90	34.84	0.93	4253
	J-OFURO2		134.46		90.04	−2.10	45.93	0.88	
	GSSTF3		130.72		89.09	−5.84	52.78	0.83	
	IFREMER		117.36		64.89	−19.20	46.09	0.89	
	OAFlux		121.25		78.60	−15.31	36.76	0.92	
Low Lat.	J-OFURO3	114.01	116.90	39.70	43.08	2.89	23.48	0.84	11892
	J-OFURO2		116.50		53.79	2.49	35.88	0.75	
	GSSTF3		122.29		56.47	8.28	39.35	0.72	
	IFREMER		123.10		35.11	9.10	23.82	0.80	
	OAFlux		110.21		34.00	−3.80	24.97	0.78	

5. Discussion

5.1 Advantage

We, here, discuss the advantages of using QCS for flux product verification. We think that there are the following three advantages: 1) Easy operation, 2) Unified benchmark framework, and 3) Several modes specialized for verification.

1) Easy operation

One of the advantages of this system is that various works for the verification can be easily performed. It is not necessary to write the program code to conduct and the verification can be performed with only simple setting. Since figures and tables obtained from QCS as the results of verification are aggregated into html files, the results can be confirmed by a web browser. These improvements of convenience will contribute to encouraging implementation of more elaborate data verification and releasing data developers from routine works for the verification.

2) Unified benchmark framework

By systemizing various things on verification such as processing and editing in situ observation data, QCS provides a framework of verification including a unified benchmark dataset for verification. This is considered the most important aspect to characterize QCS. By having verification framework with unified benchmark dataset, inter-comparison of multiple data sets, confirming changes in quality due to update of data version, etc. are conducted using the same criteria. It is expected to be used for verification of other flux products of satellite-derived products, atmospheric reanalysis, and outputs from ocean-atmosphere coupled models such as CMIP5.

3) Several modes specialized for verification

The inter-comparison mode is one of the functions to characterize QCS. In order to fairly evaluate multiple data sets on the same criteria, it is necessary to align the number of in situ data used for verification. Moreover, the FLUX mode ensures the consistency of the physical input parameters related with turbulent heat fluxes.

5.2 Limits of QCS and future task

The current QCS has several limitations and challenges. The target variables are only the turbulent heat fluxes and their related parameters. It is a future task to verify shortwave, longwave radiation, net heat fluxes, and other fluxes. There is also a limit on resolution of target data set. The current QCS limits the temporal resolution of target dataset to daily or monthly values. Therefore, QCS cannot treat hourly or instantaneous satellite data. This is because in situ observation data stored in QCS is processed on only daily and monthly basis. It is necessary to maintain in situ dataset with high-resolution temporal interval (e.g., hourly) in future system. At the same time, it is also necessary to handle larger size datasets. Although the current system is adequate for basic verification of our J-OFURO3 dataset, expansion of QCS is needed for verification of higher resolution dataset. Attempts on such extensions are the next step in future QCS.

Expansion and management of the in situ data set are also major issues in future QCS. In situ data set stored in the current system is manually acquired and edited in advance. For adding new in situ data, some manual works are required again. In the future, it is necessary to have a mechanism that automatically update in situ data using an internet. In addition, it is also important to manage in situ data set acquired and edited in QCS. It is necessary to manage with the log what and when data were added. It is also needed to manage revision of in situ data set in QCS.

6. Summary

A system for investigating accuracy and reliability of satellite-derived air-sea fluxes and their related physical parameters in J-OFURO3 was developed as Quality Check System (QCS). In order to conduct systematic verification for gridded data set, QCS encompasses in situ data set and the program code for verification. As a result, the verification can be conducted with only simple setting, the results can be confirmed by a web browser. By using this system, it is easy to confirm the difference in quality of data set from previous version, to compare multiple data sets with same benchmark. As a demonstration of QCS, the some results of verification using actual gridded flux data sets was shown, and obtained figures and comparison statistics values were introduced.

Acknowledgement

This QCS was developed based on JAXA GCOM fifth research announcement “Construction of validation system for AMSR2 ocean products”. In situ buoy data in QCS are based on data released by JAMSTEC, JMA, NOAA PMEL, NDBC, and WHOI. The surface latent heat flux data sets used for the demonstration of the QCS are data released by NASA, IFREMER, and WHOI.

References

Bentamy, A., Grodsky SA, Katsaros K, Mestas-Nunez AM, Blanke B, and Desbiolles F (2013), Improvement in air-sea flux estimates derived from satellite observations, Int. J. Remote. Sens. 34(14), doi:10.1080/01431161.2013.787502.

Fairall CW, Bradley EF, Hare JE, Grachev AA, Edson JB (2003) Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. J Clim 16(4): 571-591. https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2

Shie, CL. (2012), Science background for the reprocessing and Goddard satellite-based surface turbulent fluxes (GSSTF3) data set for global water and energy cycle research, Science Document for the Distributed GSSTF3 via Goddard Earth Sciences (GES) Data and Information Services Center (DISC), 21 pp.

Tomita H, Kubota M, Cronin MF, Iwasaki S, Konda M, and Ichikawa H (2010), An assessment of surface heat fluxes from J-OFURO2 at the KEO and JKEO sites, J. Geophys. Res -Oceans, 115(C3), doi:10.1029/2009jc005545.

Yu LS, and Weller RA (2007), Objectively analyzed air-sea heat fluxes for the global ice-free oceans (1981-2005), Bull. Amer. Meteorol. Soc., 88(4), 527-539.