SAMOS air-sea fluxes: 2005–2014

S. R. Smith[1]*, N. Lopez[1,2] and M. A. Bourassa[1,3]

[1]Center for Ocean-Atmospheric Prediction Studies, The Florida State University, Tallahassee, FL, USA
[2]Department of Scientific Computing, The Florida State University, Tallahassee, FL, USA
[3]Earth, Ocean, and Atmospheric Science Dept., The Florida State University, Tallahassee, FL, USA

*Correspondence: Shawn R. Smith, Center for Ocean-Atmospheric Prediction Studies, The Florida State University, P. O. Box 3062741, Tallahassee, FL, 32306-2741, USA, E-mail: smith@coaps.fsu.edu

Bulk turbulent heat and momentum fluxes are derived using 1-minute interval data collected by the Shipboard Automated Meteorological and Oceanographic System (SAMOS) initiative. The fluxes are provided along cruise track lines for individual research vessels (RVs) and are derived using three widely accepted air-sea flux algorithms. SAMOS data collected by 19 RVs between 2005 and 2014 are used to create the dataset. The data are concentrated in the oceans around North America, but select data are available from most ocean basins.

Introduction

In support of the air-sea exchange community, the Shipboard Automated Meteorological and Oceanographic System (SAMOS) initiative (http://samos.coaps.fsu.edu) has calculated bulk turbulent heat and momentum fluxes using a quality-controlled archive of underway observations from research vessels (RVs). The dataset includes 1-min interval latent and sensible heat flux, wind stress, and height-adjusted (10 m) wind speed, specific humidity, and potential temperature along RV cruise tracks for the period 2005–2014. Developing this product is motivated by the need for high-quality fluxes to support validation and development of numerical models and a number of air-sea

flux products (including *in situ*, satellite, and blended analyses).

1. Developing the SAMOS fluxes

The flux dataset is derived using quality-controlled marine meteorological observations collected by the SAMOS initiative. The approach is to select observations from the intermediate-quality SAMOS product, which has undergone a suite of automated quality control tests (Section 1.1). Only SAMOS data that pass these quality tests are used to calculate the bulk turbulent air-sea fluxes using three different algorithms (Section 1.2) that are widely accepted by the flux community. Criteria for flux processing are described

in Section 1.3 and an overview of the data product is given in Section 1.4.

1.1. SAMOS dataset

Since 2005, the SAMOS initiative has been collecting 1-min average navigational, meteorological, and oceanographic observations (derived from higher frequency, several *per* minute up to 1 Hz, sensor measurements) from select RVs. These averages are produced at 1-min intervals and delivered in daily ship-to-shore email messages to meet the needs of a diverse research and operational community. A SAMOS consists of a computerized data logging system that continuously records navigation (ship position, course, speed, and heading), meteorological (winds, air temperature, pressure, moisture, rainfall, and radiation), and near ocean surface (sea temperature and salinity) parameters while a vessel is underway. The authors note that scientific instrumentation providing data to the SAMOS initiative are purchased, deployed, maintained, and operated by the RV home institution. Instruments are not provided by the SAMOS initiative.

Parameters typically provided by vessels contributing to SAMOS are listed in Table 1 (we exclude radiation, rainfall, and other parameters that are only provided by select vessels). Deriving bulk air-sea fluxes requires measurements of air and sea temperature, atmospheric moisture content, atmospheric pressure, and the wind speed (and direction to determine wind stress components). Each of these parameters is typically measured by a vessel contributing to SAMOS, and in some cases, independent measurements from multiple sensors are available for a single parameter. For SAMOS, moisture content is measured using a relative humidity (RH) sensor – the most commonly deployed moisture sensor type on research ships. All of the required parameters are directly measured by sensors on the ship, with the exception of the wind direction and speed. Winds measured on a moving vessel, known as relative winds, contain a contribution to the measured wind imparted by the ship motion. This motion influence is removed using the vessel speed over the ground, course over the ground, and heading to derive a true (Earth-relative) wind speed and direction. It is these true winds that are used in the flux calculations herein. Smith *et al.* (1999) provides detailed information on the calculation of true winds. The true winds are derived by the vessel operators prior to transmission of their 1-min averaged data to the SAMOS data centre. Each of these six parameters is used as input to the flux algorithms described in Section 1.2 with some exceptions as described in Section 1.3.

The flow of these six parameters (and other parameters not used to create the fluxes) from the vessel to the SAMOS data centre (Figure 1) begins with the operator sending at 0000 UTC via e-mail all 1-min data records from the previous day to the Marine Data Center at the Florida State University. SAMOS uses a custom key:value paired comma-separated value format for data transmission which is encoded by each operator using their vessel's data acquisition software. Once received, these observations are processed into a standard network common data form (netCDF), augmented with detailed ship and instrumental metadata, quality controlled, and distributed to the user community. Primary users of the data include satellite algorithm and product developers, numerical modellers, and researchers in the air-sea flux community.

SAMOS data quality control begins with verifying that the original file came from a recruited vessel and is in the proper key:value format. Once verified, the data are converted to SI units (if necessary), checked

Table 1. Limits outside of which SAMOS flags the listed parameters with a bounds (B) flag.

SAMOS parameter (abbreviation)	Lower bound	Upper bound	Units
Latitude (lat)	−90	90	Degrees North
Longitude (lon)	0	359.9999	Degrees East
Speed over ground (PL_SPD)	0	15	ms^{-1}
Course over ground (PL_CRS)	0	360	Degrees
Heading (PL_HD)	0	360	Degrees
True wind direction (DIR)	0	360	Degrees
True wind speed (SPD)	0	40	ms^{-1}
Pressure (P)	950	1050	hPa
Relative humidity (RH)	0	100	Percent
Air temperature (T)	−30 (polar)	15 (polar)	°C
	−10 (mid-latitude)	40 (mid-latitude)	°C
	10 (tropical)	40 (tropical)	°C
Sea temperature (TS)	−2 (polar)	15 (polar)	°C
	−2 (mid-latitude)	30 (mid-latitude)	°C
	15 (tropical)	35 (tropical)	°C

Air and sea temperatures apply latitude-dependent boundaries in polar (−60 to −90 or 60 to 90 degree N), mid-latitude (−30 to −60 or 30 to 60 degree N), and tropical (−30 to 30 degree N) bands. Ranges are designed to flag 'likely' errors, but do include some realistic values. For example, pressure can dip to 880 hPa in a hurricane, but the likelihood of a ship being at that location is extremely low).
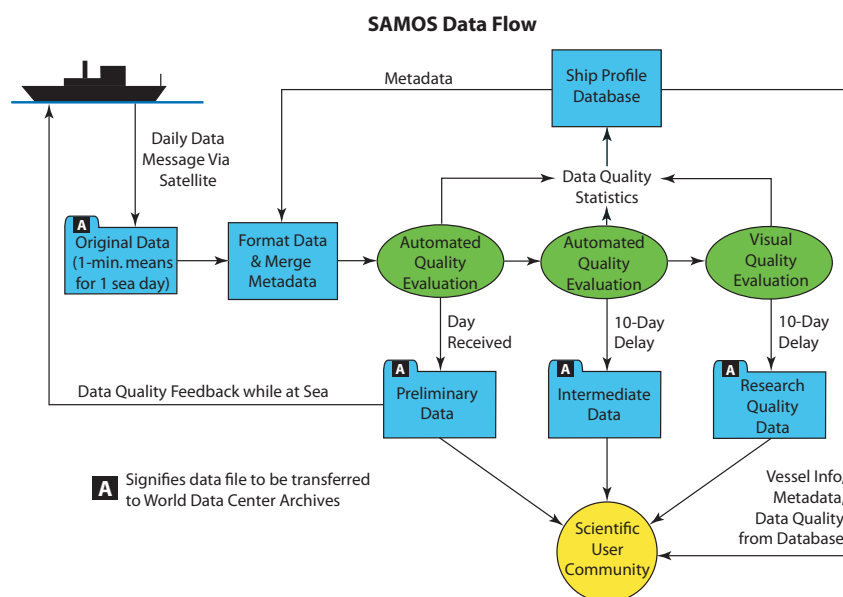
SAMOS Data Flow

Figure 1. Flow of 1-min sampling rate SAMOS observations from the vessel, through the Marine Data Center, and on to the archives and user community.

for temporal sequence, and blended with ship and instrumental metadata (e.g. instrument height, units, sensor make/model) from the SAMOS database. This first netCDF version of the observations undergoes automated evaluation to apply flags to the data. SAMOS uses a hierarchical parametric A–Z quality control scheme (e.g. each value can have only one flag; see http://samos.coaps.fsu.edu/html/samos_quality_flag.php). Initial tests verify that (1) the vessel is positioned over water by comparing the vessel latitude and longitude to a 2-min gridded global relief dataset (ETOPO2; U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, 2006; flag=L), (2) the vessel speed between sequential positions as calculated on a great circle arc is not greater than 15 m s^{-1} (a realistic speed for a RV; flag=F), and (3) the observations are within realistic physical limits (Table 1, flag=B). The pressure, air and sea temperature, wind speed, and RH are also flagged when they exceed $\pm 4\sigma$ from a monthly climatology (da Silva *et al.*, 1994; flag=G). The climatology test also uses a minimum standard deviation threshold in data-sparse areas (e.g. Southern Ocean) where da Silva *et al.* (1994) has unrealistically small standard deviations. The final automated quality tests ensure that the relationship of air temperature $\geq$ wet-bulb temperature $\geq$ dew point temperature is not violated (flag=D; although this test is not commonly used in SAMOS because moisture data is primarily measured as RH) and that true winds are properly calculated – using the reported vessel course over ground, speed over ground, heading, and relative wind direction and speed to recalculate the true wind values according to Smith *et al.* (1999) and flag (E) the reported true winds when the speed (direction) differs by more than 2.5 m s^{-1} (20°). This entire

process occurs within one to three minutes of the e-mail arriving at FSU.

On a 10-day delay from the observation date, intermediate files are automatically created by merging all preliminary files received for a given observation day. This delay allows for receipt of delayed or corrected files from the RV. The file merge takes into account temporal duplicates between multiple files. Duplicates are resolved through a series of tests that first determine whether the data values are exact or different. When they differ, the first test retains the value with the 'best' preliminary QC flag. Best flag hierarchy for position data (latitude, longitude) is Z > F > L and for other parameters (sea temperature, humidity, etc.) is Z > G > E > B > D, where Z is the flag used for data that do not fail any QC tests. If the flags on the data values are identical, the second duplicate resolution test compares the values in question to the 30-min mean centred on the duplicate time, retaining the value closest to the mean. Failure to resolve the duplicate at this stage results in all duplicate values being removed for the time in question and the situation being stored in a processing log (a compromise to allow automation of the merge process).

Version 2.0 (Version 1.0 of our flux product was developmental and was never released for public application.) of the SAMOS fluxes is derived from these intermediate-level data files. Although an additional set of visual QC is conducted for some RVs (Figure 1), the authors decided to use a common level of QC for this product release. Future releases may include visually QC'd (research quality) data.

The original data received from the vessel and all three levels of SAMOS-quality processed files are submitted to and available from the National Centers for Environmental Information – Maryland (NCEI-MD;

1.2. Flux algorithms

In the air-sea interaction community, questions remain regarding the parameterizations of fluxes, particularly related to wave influences and stability parameterization; therefore, height-adjusted input variables and flux estimates are provided using three different algorithms, so users can select fluxes derived from the algorithm that best suits their needs. Users may also compare the three algorithms available in the SAMOS flux product, since nearly every flux observation is represented by the three different algorithms.

1.2.1. Smith, 1988 (S88)

Often used as the flux 'standard' within the air-sea interaction community (e.g. modellers, flux product developers), Smith (1988) provides surface layer coefficients related to surface roughness and boundary layer stratification to determine profiles (for wind, potential temperature, and humidity), surface wind stress, and surface turbulent heat flux in typical open ocean conditions. Smith (1988) parameterization includes turbulent transport due to smooth surface and a Charnock wind parameterization for gravity waves, with a value of Charnock's constant tuned to open ocean conditions with older seas. This assumption is made because wave data were rarely available, and there were insufficient data to determine the dependency on waves other than seen in the Charnock parameterization.

The stability parameterizations are those published in Smith (1988). These are based on Monin–Obukhov similarity theory. These parameterizations (one for stable conditions and one for unstable conditions) modify the profiles of mean wind, potential temperature, and humidity. They also modify the fluxes – increasing fluxes for unstable conditions and decreasing them for stable conditions. There have been considerable improvements that are applied in the other two flux parameterizations.

1.2.2. SAMOS variant of BVW Flux (B12)

This version of the Bourassa-Vincent-Woods (BVW) parameterization is an algorithm originally published by Bourassa *et al.* (1999) and then adjusted as described in Bourassa (2006) and Zheng *et al.* (2013). The algorithm is known for additional turbulent transport due to capillary (ripple) waves, which makes it better for lower wind speeds and an improved handling of sea state in the 2006 version. The Zheng *et al.* (2013) adjustment changed the surface roughness parameterization to one with a smooth transition from calm to rough surfaces. Since SAMOS vessels do not report wave data, we use roughness for gravity waves identical to the value in S88 with a Charnock constant equal to 0.012. B12 has the ability to consider swell moving in any direction, but we do not take

advantage of this feature because of the lack of wave data.

The stability parameterizations are published in Bourassa *et al.* (1999) and are slightly different than those in S88 for most conditions, but considerably different for highly stable conditions.

1.2.3. COARE 3.5 (c35)

This algorithm by Edson *et al.* (2013) is an enhancement of an algorithm first produced for the Coupled Ocean-Atmosphere Response Experiment (COARE). Data collected from four oceanic field experiments were used to improve the original COARE parameterizations (Fairall *et al.*, 1996, 2003) of the surface roughness and drag coefficient across the ocean. Wind stress estimates, in particular, were significantly improved for wind speeds greater than 13 m s^{-1}. Edson *et al.* (2013) showed that the inverse wave age varies nearly linearly with wind speeds up to 25 m s^{-1}; therefore, the algorithm does not require wave data, and assumes that waves and wind are moving in the same direction.

The COARE stability parameterizations are unchanged; they are slightly different than those in S88 and B12 for unstable conditions, and identical to BVW for stable conditions. One difference from S88 and B12 is that the code is assumed to iterate to a solution in three passes. This is usually a very good assumption, but will occasionally cause differences from other models.

1.3. Producing the SAMOS flux

A flow chart of the flux file processing is shown in Figure 2. Every flux file (one for each algorithm) corresponds to the data present in a single intermediate
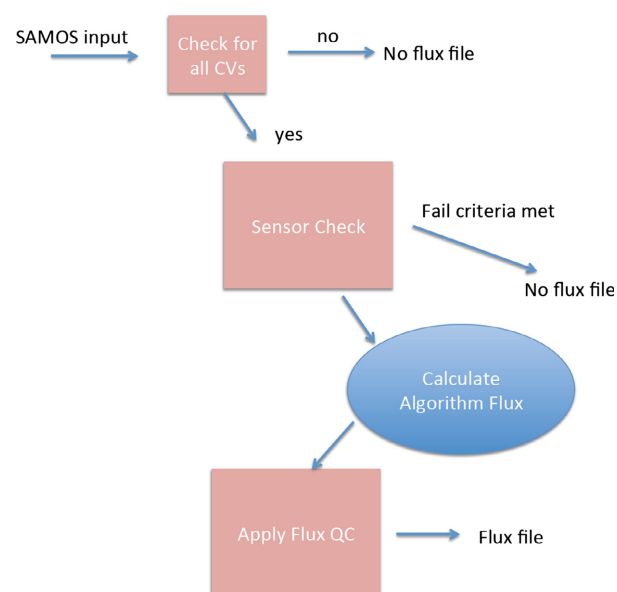

Figure 2. Major components of flux file processing. Not having all critical variables (CVs) or the necessary height metadata (sensor check) results in no flux file creation.

SAMOS file. However, not every SAMOS file has an associated flux file. Approximately 40% of all available SAMOS files have associated flux files. The need for multiple state variables to be present at a given time for flux calculation is the main reason for the small sample. Another challenge is the lack of metadata about the RV instrument heights. The specific research ships and the years for which fluxes have been derived for release 2.0 of the SAMOS fluxes are provided in Table 2. The flux files are designed to have a similar netCDF structure to the original SAMOS files.

1.3.1. Obtaining input

Several atmospheric and oceanic variables *must* be provided in the original SAMOS netCDF file to enable creation of a flux dataset. For the software implementation, we define critical variables (CVs) as:

- air temperature (T)
- sea temperature (TS)
- relative humidity
- wind speed (SPD)

Absence of these variables makes flux calculation impossible, so if any CV is missing in a given SAMOS intermediate file, no flux file is generated. RH is the most common moisture parameter measured on SAMOS vessels, so we limit the CV to include RH (as opposed to wet-bulb or dew point temperature) in version 2.0 of the fluxes. Pressure benefits the accuracy of the flux calculation, and wind direction allows for the generation of 10-m wind and wind stress component outputs; so, these two variables are also considered inputs.

When no acceptable pressure value is available, but the other CVs exist, then a default pressure of 1013 hPa is used. Differencing latent heat flux calculated from SAMOS data using measured pressure *versus* the default of 1013 hPa resulted in a standard deviation of 0.09 W m^{-2} and a maximum absolute difference of 17.13 W m^{-2}. The implication is that using the default pressure in cases when no pressure exists will result in only small biases in the latent heat flux values. Sensible heat fluxes also have only a small bias, since values would be scaled by the square root of the actual pressure divided by the 1013 hPa approximation. Since these flux values are appropriately flagged (see Section 1.3.3), the user should be able to reject them if desired.

If wind direction is not available from a given SAMOS intermediate file, then the 10-m wind components and wind stress components are set to missing (−9999), but the stress magnitude, wind speed, and all other output variable data are included in the flux file.

It is also important to note that TS from vessels is typically measured at a depth of a few metres, but the B12 flux parameterization takes skin temperature as input. No adjustments are made to TS before this sea surface temperature is applied in the B12 model. The C35 flux allows the user to classify the input TS as a

Table 2. Subset of SAMOS research vessels for which air-sea fluxes are derived.

Research vessel name	Call sign	Data provider	Years
Atlantis	KAQP	WHOI	2005–2014
David Starr Jordan	WTDK	NOAA	2008
Atlantic Explorer	WDC9417	BIOS	2010–2014
Aurora Australis	VNAA	IMOS	2008–2013
Bell M. Shimada	WTED	NOAA	2012–2014
Delaware 2	KNBD	NOAA	2009–2010
Falkor	ZCYL5	SOI	2013–2014
Healy	NEPP	USCG	2008–2007, 2011–2013
Henry Bigelow	WTDF	NOAA	2009–2014
Knorr	KCEJ	WHOI	2005–2010, 2012–2014
Lawrence M. Gould	WCX7445	OPP	2007–2014
Melville	WECB	SIO	2013–2014
Miller Freeman	WTDM	NOAA	2010
Robert G. Sproul	WSQ2674	SIO	2013–2014
Roger Revelle	KAOU	SIO	2013–2014
Southern Surveyor	VLHJ	IMOS	2008–2013
Tangaroa	ZMFR	IMOS	2011–2014
Thomas Jefferson	WTEA	NOAA	2012–2014
Thomas G. Thompson	KTDQ	UW	2013–2014

These vessels all had required data and metadata (specifically instrument heights/depths) for the flux algorithms. The call sign is used in the flux file names to identify the vessel. Within the years listed, the actual dates with flux data will vary depending on days at sea for the vessel and completeness of the data provided to the Marine Data Center. SAMOS thanks the following institutions for providing data from the vessels listed above: BIOS – Bermuda Institute of Ocean Sciences, IMOS – Australian Integrated Marine Observing System, NOAA – National Oceanic and Atmospheric Administration: Office of Marine and Aviation Operations, OPP – Contractors for National Science Foundation Office of Polar Programs, SIO – Scripps Institution of Oceanography, SOI – Schmidt Ocean Institute, USCG – United States Coast Guard, UW – University of Washington, and WHOI – Woods Hole Oceanographic Institution.

bulk surface temperature, so this setting was used for C35.

Other inputs are passed to the subroutines that are not typically measured by SAMOS ships; a description of these parameters is presented in Table 3. These values often alter the output of the flux calculation and may help the user decide which algorithm best fits their needs.

Many of the RVs also report a single variable on multiple sensors. Table 4 lists the criteria to choose between multiple sensors on the basis of combination of rules regarding SAMOS metadata heights and quality control flags. A flux file might not be created simply because of lack of instrument height metadata.

1.3.2. Processing to output

Input data for S88 and B12 algorithms are run through a Modularized Flux Testbed (MFT) developed at COAPS (Moroni, 2008). Since C35 is a newer algorithm, input data are sent through a C function created using MATLAB's 'Coder' software. For access to the original MATLAB code, contact Jim Edson at james.edson@uconn.edu. Each output netCDF SAMOS flux file contains time, latitude, and longitude, and the variables listed in Table 5. The C35 subroutine does not output momentum roughness length or Monin–Obhukov length scale height, so these variables are always set to missing (−9999) in the C35 flux file.

1.3.3. Flux quality control

The authors consider 'acceptable' flags on the SAMOS data input to the flux algorithms to be G, acknowledging that the statistical test sometimes flags realistic extreme values as a result of uncertainties in the climatology, or Z. To be consistent with the design of the SAMOS quality control flags, flux quality control flags are also parametric, but use a one-digit number instead of a letter. These numbers are based on input value uncertainties and do not represent any statistical comparison between the flux data and climatological flux values. The original time, latitude, and longitude flags are copied to the flux files, and the remainder of

Table 3. Additional input values required to reproduce the output in the SAMOS flux files.

Variable	Units	S88	B12	C35	Description
Flux Model Parameter	None	1	0	NA	Setting in MFT
Convective Parameter[1]	None	0	0.6	1.25	Bulk parameterization of air-sea fluxes for TOGA COARE
Surface Relative Humidity[2]	None (fraction)	1.00	0.98	1.0	
Wave Age of Dominant Waves[3]	None	43.64	43.64	NA	Used to determine the Charnock (1955) parameter. C55 uses a wind speed dependent Charnock parameter
Roughness Length Parameterization		0	0	NA	Bourassa *et al.* (1999)
Capillary Wave Roughness included	None	No	Yes	No	Bourassa *et al.* (1999)
Atmospheric Stability Option	None	1	0	NA1	0 = BVW parameterizations; 1 = Smith (1988) parameterizations: C35 is described in Fairall *et al.* (1996)
Equivalent Neutral Winds Parameter	None	0	0	NA	No modification to calculation of 10 m winds
Cool Skin Adjustment	None	None	None	Yes	
Net Energy[4]	W m^{-2}	0	0	25.0	Flux through the air/sea interface
Downward Shortwave[4]	W m^{-2}	0	0	150.0	
Downward Longwave[4]	W m^{-2}	0	0	370.0	
Latitude[5]	Degrees	0	0	45	
PBL Height	Metres	0	0	600	Height of Boundary Layer

The S88 and B12 algorithms are run from a Modularized Flux Testbed (MFT; Moroni (2008); www.coaps.fsu.edu/~bourassa/MFT_html/MFT_docs.php) that allows many different concepts to be tested in isolation or in combination. Similarly, the C35 algorithm uses additional assumed input in calculating fluxes.

[1]Smith (1988) does not assume convection, so a convective parameter of 0.0 is applied.

[2]The input surface humidity for the B12 parameterizations has been chosen to be 98% of saturation to account for the influence of salt in the water, but the other two algorithms assume 100% saturation.

[3]The MFT has an option that allows the Charnock parameter to be determined from the wave age, and a wave age of 43.64 results in a value of the Charnock parameter that is consistent with the original S88 algorithm.

[4]The radiation flux values in the C35 model are used in the conversion between a bulk SST and a skin SST.

[5]Latitude is used to calculate a slightly more accurate value of the gravitational constant for the C35 model, but instead of using latitude measured by the vessel at each minute, a default value of 45° is applied.

Table 4. The rules for choosing which sensor on the RV will be used for flux calculation.

Variable	Sensor names	Description of choice	Fail criteria
Pressure	P, P2, P3	Sensor with most number of good flags	None. If no pressure sensor exists, default pressure of 1013 mb is used
Temperature (Temp)	T, T2, T3	Sensor with most number of Temp + Rel_hum acceptable flags that also has height metadata[1]	Height metadata not present for any temperature or relative humidity sensors
Relative Humidity (Rel_hum)	RH, RH2, RH3	Sensor with most number of Temp + Rel_hum acceptable flags that also has height metadata[1]	Height metadata not present for any temperature or relative humidity sensors
Wind Speed (Wind)	SPD, SPD2, SPD3	Sensor with most number of Wind + Dir acceptable flags that also has height metadata[2]	Height metadata not present for any wind speed/direction sensors
Wind Direction (Dir)	DIR, DIR2, DIR3	Sensor with most number of Wind + Dir acceptable flags that also has height metadata[2]	None. If no wind direction sensor exists, wind components will be set to MISSING in the file
Sea Surface Temperature (SST)	TS, TS2, TS3	Sensor with most number of acceptable flags	No SST sensor is present

Not all RVs have multiple sensors, but for those that do, the flux code supports up to 3 sensors. 'Acceptable' flags are considered 'G' or 'Z' flags from the SAMOS flagging scheme, while 'Z' represents good data that did not fail any automated QC tests. If the 'Fail Criteria' is met on any of the Critical Variables (CVs are highlighted in bold), then no flux file is created.

[1]Rel_hum height metadata can be used in place of Temp metadata (and vice versa), since the sensors are mounted within the same instrument.

[2]Dir metadata can be used in place of Wind metadata (and vice versa).

Table 5. List of flux output variables available in algorithm flux files.

Variable	Name in file	Algorithms
Latent heat flux	hfls	S88, B12, C35
Sensible heat flux	hfss	S88, B12, C35
Wind stress magnitude	tau	S88, B12, C35
Zonal wind stress	tauu	S88, B12, C35
Meridional wind stress	tauv	S88, B12, C35
Zonal wind speed at 10 m	uas	S88, B12, C35
Meridional wind speed at 10 m	vas	S88, B12, C35
Wind speed magnitude at 10 m	u_10	S88, B12, C35
Potential[1] temperature at 10 m	t_10	S88, B12, C35
Specific humidity at 10 m	q_10	S88, B12, C35
Monin–Obhukhov length scale height at 10 m	z_over_L	S88, B12
Momentum roughness length	zo_m	S88, B12

All zonal and meridional wind components may be set to a missing value (-9999) if wind direction is not present.

[1]Temperature adjusted adiabatically to 10 m from instrument height based potential temperature profile.

the variables use numbered flags according to the following system:

- Flux variables (including wind stress magnitude)
  - Z: All CVs had acceptable flags and the actual pressure is used for flux calculation
  - 0: All CVs had acceptable flags, but the default pressure is used
  - 1: Exactly 1 input CV had an unacceptable flag
  - 2: Exactly 2 input CVs had an unacceptable flag
  - 3: Exactly 3 input CVs had an unacceptable flag
  - 4: Exactly 4 input CVs had an unacceptable flag
- Wind stress and 10-m wind components
  - 5: input wind direction and/or wind speed had an unacceptable flag

The application of a one-digit number flag to an observation implies that the flux value exists and can be adopted at the user's discretion. On the other hand, some values in a flux file are assigned a missing value (-9999) with a corresponding 'Z' flag in the flag array (missing values are considered to be good quality – no corresponding missing value flag is associated with the flux value). The presence of missing values in flux files is triggered by the appearance of a missing value for one or more CVs in the original SAMOS file. When a CV sensor stops reporting (i.e. when a ship reaches a port), a flux cannot be computed for these time stamps.

We recommend the use of flux data with a '0' or 'Z' flag; caution should be exercised with data containing a different flag. Additionally, the user should filter out missing values when using the flux files.

1.4. Product overview

The SAMOS flux product is derived from data collected along select RV cruise tracks, so the spatial coverage varies over the global oceans (Figure 3). The distribution of latent heat flux values from the Bourassa algorithm (Figure 3) is representative of all the parameters and algorithms provided in version 2.0 of the SAMOS fluxes. Densities are highest around North America since the majority of the vessels recruited by SAMOS are U.S. operated. Flux values exist from the tropics to the polar regions, with Arctic Ocean sampling primarily from the *RV Healy* and Southern Ocean sampling from the *RVs Lawrence M. Gould* and *Aurora Australis* (the latter courtesy of the Australian Integrated Marine Observing System). Gaps in data coverage exist in the western north Pacific, the central south Pacific, the Indian, and the south Atlantic oceans. Despite the gaps in coverage, the SAMOS fluxes provide data across a range of ocean environments.

Distributions of the primary heat, wind stress, and 10-m adjusted parameters provided by the SAMOS flux product (Figure 4) are all within expected ranges, with the primary differences being related to the subtle variations in the three flux algorithms. For input values used to derive the fluxes, potential air temperature (median value 15.6°C for S88 and B12; 15.3°C for C35) and specific humidity (median value of 8.4 g kg^{-1} for all three products) are nearly identical across the interquartile range (Figure 4(c) and (d)). Wind speed at 10 m (Figure 4(e)) is shifted towards slightly higher values for C35 (median = 6.1 m s^{-1}) *versus* the other two algorithms (6.0 m s^{-1} for S88 and B12). The latent heat flux distribution for C35 is shifted towards lower values at all percentiles as compared to S88 and B35 (Figure 4(a)). B12 also shows a wider spread of IQR for sensible heat flux as compared to S88 and C35 (Figure 4(b)) and has a median shifted towards higher values (6.9 W m^{-2} for B12 *vs* 5.4 W m^{-2} and 4.5 W m^{-2} for S88 and C35 respectively). Overall differences in the wind stress magnitudes for the three products is negligible, though C35 does have a slightly larger skew towards larger stress in the upper tail of the distribution (Figure 4(f)). Wind stress components are also available for all three flux products (plots not shown).

2. Dataset location and format

Both the flux dataset and source SAMOS data files are contained in a hierarchical directory structure (Figure 5). At the lowest level, the netCDF flux files have the following nomenclature:

  CCCC_YYYYMMDDv2XXZZfluxS88.nc
  CCCC_YYYYMMDDv2XXZZfluxB12.nc
  CCCC_YYYYMMDDv2XXZZfluxC35.nc

Here, CCCC denotes the ship's call sign (4–7 alphanumeric characters), YYYYMMDD denotes the date, and XX and ZZ are, respectively, the version and order number of the intermediate SAMOS file used to create the flux.
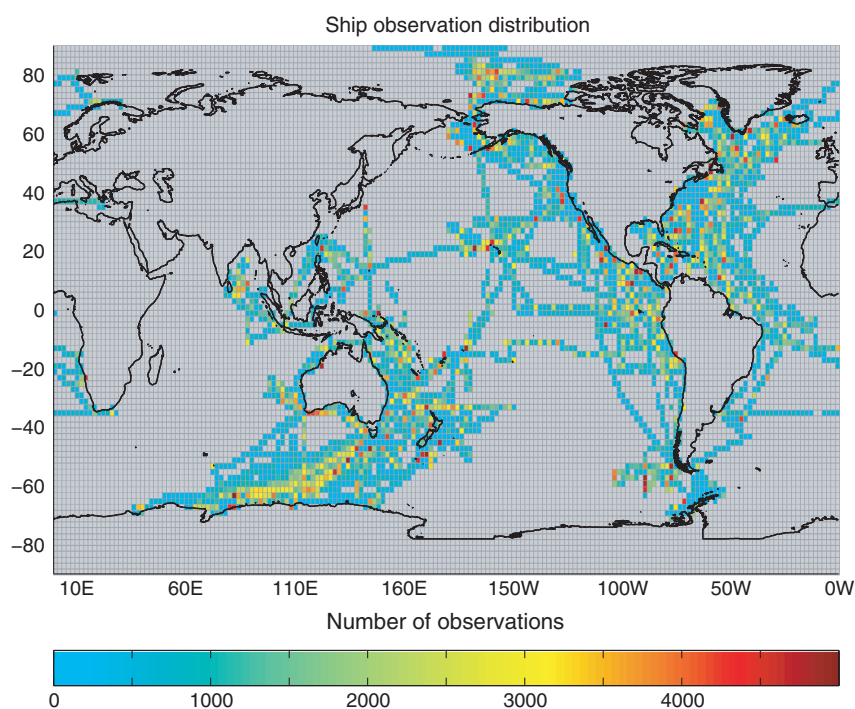

Figure 3. Distribution of 1-min latent heat fluxes from the Bourassa flux algorithm. Values are binned in two-degree latitude by two-degree longitude grids with magnitude coded according to the colour bar. Values are assigned to a bin only when the position data and the latent heat flux value are flagged as good (e.g. flags = Z or 0). The plot is representative of the flux data coverage for all variables and flux algorithms. Grey boxes indicate no available flux values.
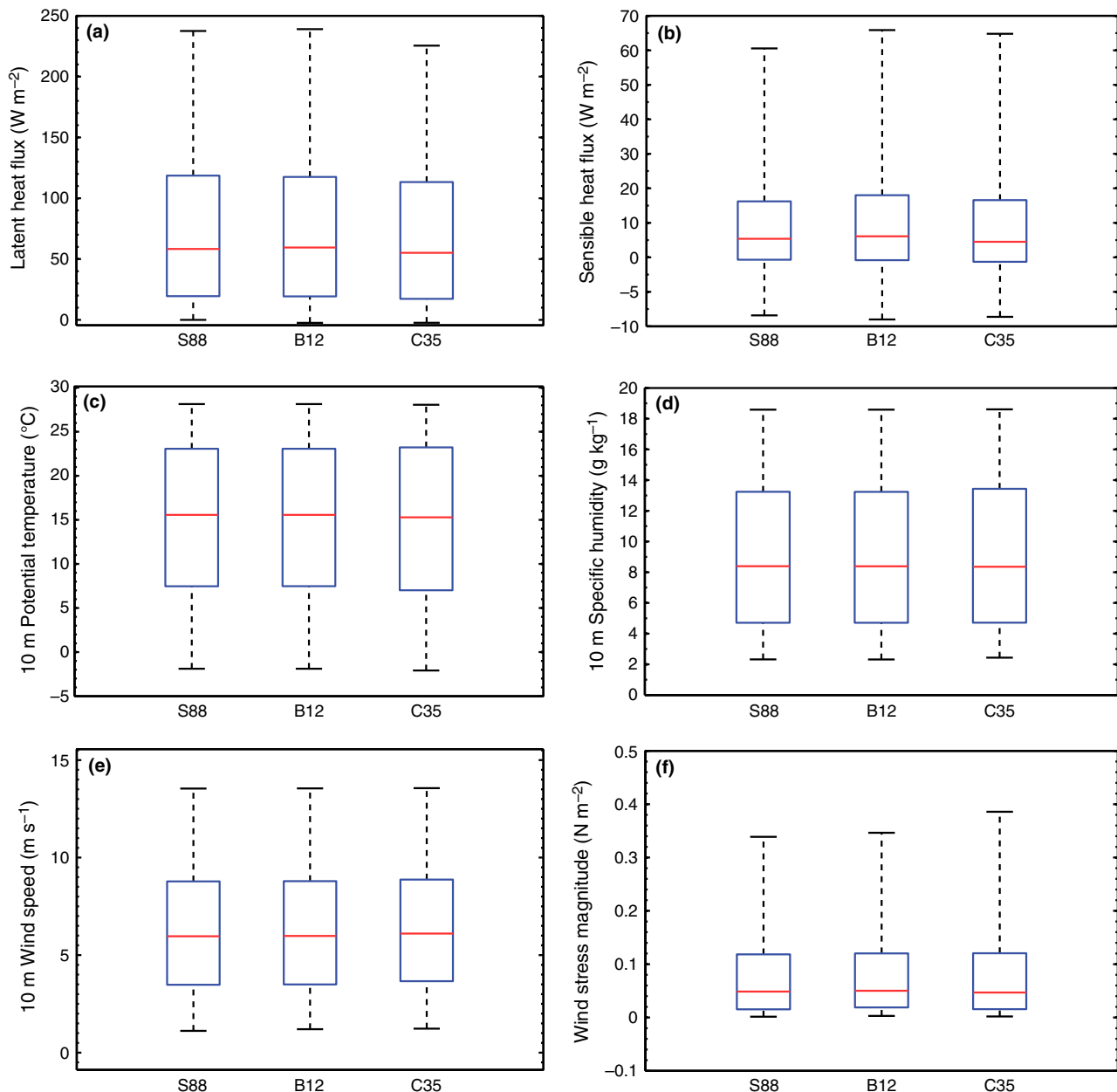
Figure 4. Box plots showing the distribution of (a) latent heat flux, (b) sensible heat flux, (c) 10-m potential temperature, (d) 10-m specific humidity, (e) 10-m wind speed, and (f) wind stress magnitude. Distributions are created using only data values that are flagged as good (e.g. flags = Z or 0). The distributions include data from all ships and years listed in Table 2. The plots display the median (red line), interquartile range (blue box), and the 5th and 95th percentile values (dashed whiskers). Missing and flagged SAMOS observations resulted in a 10% (11%) loss of data for calculating heat fluxes, 10-m adjusted values, and the wind stress magnitude for S88 and B12 (C35).

Version 2.0 of the SAMOS fluxes, covering 2005–2014, are available from the Research Data Archive of the National Center for Atmospheric Research (NCAR) (http://dx.doi.org/10.5065/D6930R70). Subdirectories organize the data by vessel call sign and year with the file naming structure as described above.

3. Dataset use and reuse

The SAMOS flux products are available for use/reuse without restriction to support the widest possible range of communities. Since the original SAMOS observations are collected from diverse ocean regions, often outside of normal shipping lanes, and are sampled at 1-min intervals, the bulk air-sea fluxes calculated from these data provide a unique dataset. The SAMOS fluxes are ideal for matching to satellite-derived flux estimates since the 1-min sampling interval allows more accurate collocation in space and time between the two observing platforms. Additionally, the 1-min sampling rate supports model validation by allowing integration (averaging) of the SAMOS fluxes in a manner that can mimic the integration periods used by numerical weather, climate, and ocean models.
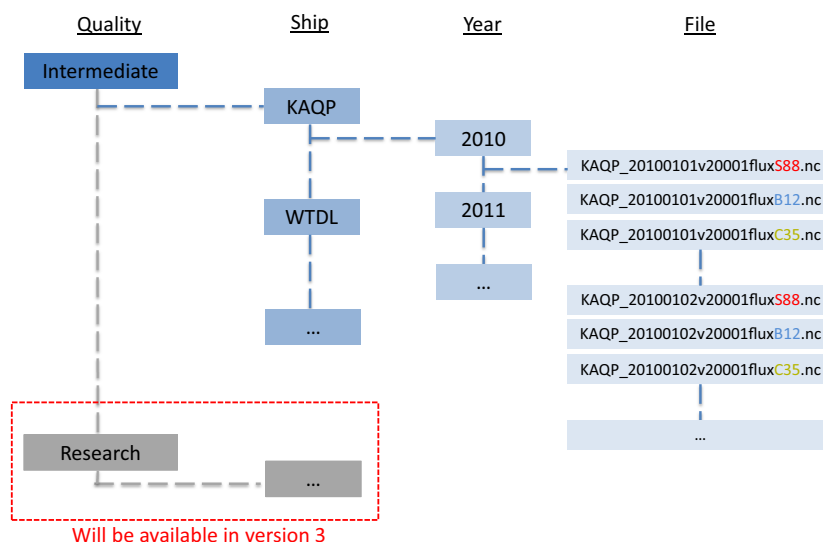
Figure 5. Directory tree showing organization of the SAMOS flux files (the source SAMOS netCDF files are stored in an identical directory tree on the SAMOS ftp server – ftp://coaps.fsu.edu/samos_pub/data/). SAMOS flux version 2.0 only uses intermediate SAMOS quality processed data, but the authors anticipate creating fluxes for the available research level quality-controlled files when version 3 of the SAMOS fluxes is released.

4. Future of the dataset

Plans for version 3.0 of the SAMOS fluxes include 'research' quality SAMOS input data and a few other code improvements that will increase the quantity and quality of the flux observations being produced. The authors anticipate that Rvflux-3.0 will have the advantage of automatic flux file creation as new SAMOS files are put into repository. The next version will use a more robust flagging system and a sensor-checking scheme that allows additional moisture variables, not just RH, to be used as an input to the flux algorithms. Since a few ships report shortwave radiation to SAMOS, including this variable for flux input is a possibility.

References

Bourassa MA. 2006. Satellite-based observations of surface turbulent stress during severe weather. *Atmosphere-ocean Interactions* **2**: 35–52.

Bourassa MA, Vincent DG, Wood WL. 1999. A flux parameterization including the effects of capillary waves and sea state. *Journal of the Atmospheric Sciences* **56**: 1123–1139.

Center for Ocean-Atmospheric Prediction Studies, The Florida State University, [Smith, S. R., N. Lopez, and M. A. Bourassa]. 2015. SAMOS Air-sea fluxes. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. doi:10.5065/D6930R70

Edson JB, Jampana V, Weller RA, Bigorre SP, Plueddemann AJ, Fairall CW, Miller SD, Mahrt L, Vickers D, Hersbach H. On the exchange of momentum over the open ocean. *Journal of Physical Oceanography* **43**: 1589–1610, doi: 10.1175/JPO-D-12-0173.1.

Fairall CW, Bradley EF, Rogers DP, Edson JB, Young GS. 1996. Bulk parameterization of air-sea fluxes for TOGA COARE. *Journal of Geophysical Research* **101**: 3747–3764.

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

Moroni DF. 2008. A global and regional diagnostic comparison of air-sea flux parameterizations during episodic events. M.S. Thesis, Florida State University, Tallhaassee, FL. Available at http://diginole.lib.fsu.edu/etd/2261/ (accessed 01 February 2016).

da Silva AM, Young CC, Levitus S. 1994. *Atlas of Surface Marine Data, Vol. 1: Algorithms and Procedures*. NOAA Atlas NESDIS 6, U. S. Department of Commerce, NOAA, 83 pp. Available from: NOAA/NODC, Customer Service, E/OC, 1315 East-West Highway, Silver Spring, MD 20910.

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

Smith SR, Bourassa MA, Sharp RJ. 1999. Establishing more truth in true winds. *Journal of Atmospheric and Oceanic Technology* **16**: 939–952.

Smith SR, Rolph JJ, Briggs K, Bourassa MA. 2009. Quality-controlled underway oceanographic and meteorological data from the Center for Ocean-Atmospheric Prediction Studies (COAPS) - Shipboard Automated Meteorological and Oceanographic System (SAMOS). *National Oceanographic Data Center, NOAA*. doi:10.7289/V5QJ7F8R.

U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, 2006. *2-minute Gridded Global Relief Data (ETOPO2v2)*. Available at http://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html (accessed 01 February 2016).

Zheng Y, Bourassa MA, Hughes PJ. 2013. Influences of sea surface temperature gradients and surface roughness changes on the motion of surface oil: a simple idealized study. *Journal of Applied Meteorology and Climatology* **52**: 1561–1575. doi:10.1175/JAMC-D-12-0211.1.