

# THE DATA MANAGEMENT SYSTEM FOR THE SHIPBOARD AUTOMATED METEOROLOGICAL AND OCEANOGRAPHIC SYSTEM (SAMOS) INITIATIVE

Shawn R. Smith<sup>(1)</sup>, Jacob Rettig<sup>(1)</sup>, Jeremy Rolph<sup>(1)</sup>, Jiangyi Hu<sup>(1)</sup>, Elizabeth C. Kent<sup>(2)</sup>, Eric Schulz<sup>(3)</sup>, Ruslan Verein<sup>(3)</sup>, Steven Rutz<sup>(4)</sup>, Chris Paver<sup>(4)</sup>

<sup>(1)</sup>Center for Ocean-Atmospheric Prediction Studies, Florida State University, 236 R. M. Johnson Bldg., 32306-2840 Tallahassee USA, Email: [smith@coaps.fsu.edu](mailto:smith@coaps.fsu.edu), [jrettig@coaps.fsu.edu](mailto:jrettig@coaps.fsu.edu), [rolph@coaps.fsu.edu](mailto:rolph@coaps.fsu.edu), [hu@coaps.fsu.edu](mailto:hu@coaps.fsu.edu)

<sup>(2)</sup>National Oceanography Centre, European Way, Southampton, SO14 3ZH, UK, E-mail: [eck@noc.soton.ac.uk](mailto:eck@noc.soton.ac.uk)

<sup>(3)</sup>Centre for Australian Weather and Climate Research, Bureau of Meteorology, GPO Box 1289, Melbourne, Vic 3001 Australia, Email: [e.schulz@bom.gov.au](mailto:e.schulz@bom.gov.au), [r.verein@bom.gov.au](mailto:r.verein@bom.gov.au)

<sup>(4)</sup>National Oceanographic Data Center, NOAA/NODC E/OCI, 1315 East-West Highway, Silver Spring, MD 20910-3282, USA, Email: [Steven.Rutz@noaa.gov](mailto:Steven.Rutz@noaa.gov), [Christopher.Paver@noaa.gov](mailto:Christopher.Paver@noaa.gov)

## ABSTRACT

The Shipboard Automated Meteorological and Oceanographic System (SAMOS) initiative focuses on improving the quality of, and access to, surface marine meteorological and oceanographic data collected *in-situ* by automated instrumentation on research vessels (RVs). The data management system, from collection to archive, is described along with the benefits of the initiative, providing a model for future underway data collection. Recommendations are made for expanding the current initiative to the wider international community.

## 1. INTRODUCTION

A SAMOS (Shipboard Automated Meteorological and Oceanographic System) is typically 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. Measurements are recorded at high-temporal sampling rates (typically 1 minute or less). A SAMOS comprises scientific instrumentation deployed by the RV (Research Vessel) operator. The instruments are not provided by the SAMOS initiative.

The SAMOS initiative is complementary to the JCOMM<sup>1</sup> Voluntary Observing Ship (VOS) Scheme [7] and some vessels contribute to both programs. Traditional VOS report observations at one to six hour intervals, much less frequently than SAMOS, to meet objectives including the initialization of numerical weather prediction (NWP) models and providing input to marine climate data sets. VOS achieves higher spatial coverage than SAMOS by using a wider range of vessel types—some instrumented with automatic systems and

some with basic instrumentation supplemented with visual observations. Growing interest in climate variability and the requirement to quantify air-sea fluxes have placed unprecedented observational requirements on marine meteorology. New user communities [11] are requesting high-temporal sampling of the wide range of oceanic and atmospheric data collected via SAMOS, including coverage outside shipping lanes and in extreme environments. Some observations from sophisticated air-sea flux instrumentation are available, but these sensors are more expensive than conventional automated meteorological sensors, still under development for polar oceans, and currently are deployed only on limited research cruises [4]. SAMOS currently partners with groups deploying these flux systems to provide shipboard comparisons to the instrumentation used to collect SAMOS observations. SAMOS is a pragmatic solution, meeting the requirements for accuracy and temporal and spatial coverage by taking advantage of the existing fleet of suitably equipped RVs.

SAMOS data management provides a ship-to-shore-to-user data pathway. Daily packages of one-minute interval SAMOS data are sent to a data assembly center (DAC) at the Florida State University via e-mail attachment. Broadband satellite communication facilitates this transfer as near as possible to 0000 UTC (Universal Time, Coordinated) daily. A preliminary version of the SAMOS data is made available via web services within five minutes of receipt. The preliminary data undergo common formatting, metadata enhancement, and automated quality control (QC). Visual inspection and further scientific QC result in intermediate and research-quality SAMOS products that are nominally distributed with a 10-day delay from the original data collection date. All data and metadata are version controlled and tracked using structured query language (SQL) databases. All data are distributed free of charge and proprietary holds through the web (<http://www.coaps.fsu.edu/RVSMDC/html/data.shtml>)

---

<sup>1</sup> JCOMM is the Joint World Meteorological Organization (WMO)/Intergovernmental Oceanographic Commission (IOC) Technical Commission for Oceanography and Marine Meteorology.

and long-term archiving occurs at the US National Oceanographic Data Center (NODC).

The Integrated Marine Observing System (IMOS; <http://www.imos.org.au>) is an initiative to observe the oceans around Australia [1]. One component of the system, the “IMOS underway ship flux project” (hereafter referred to as IMOS), is modeled on SAMOS and obtains routine meteorological and surface-ocean observations from two Australian RVs. In addition to running a parallel system to SAMOS, IMOS also

contributes to SAMOS the first observations from vessels not operated by the United States (US).

The SAMOS data management system has been operational since 2005. To date, 21 RVs have been recruited from the US National Oceanic and Atmospheric Administration (NOAA), the Woods Hole Oceanographic Institution (WHOI), the US Coast Guard (USCG), the US Antarctic Program, and IMOS. Data coverage is primarily the Atlantic and Pacific Oceans, with some high-latitude observations (Fig. 1).

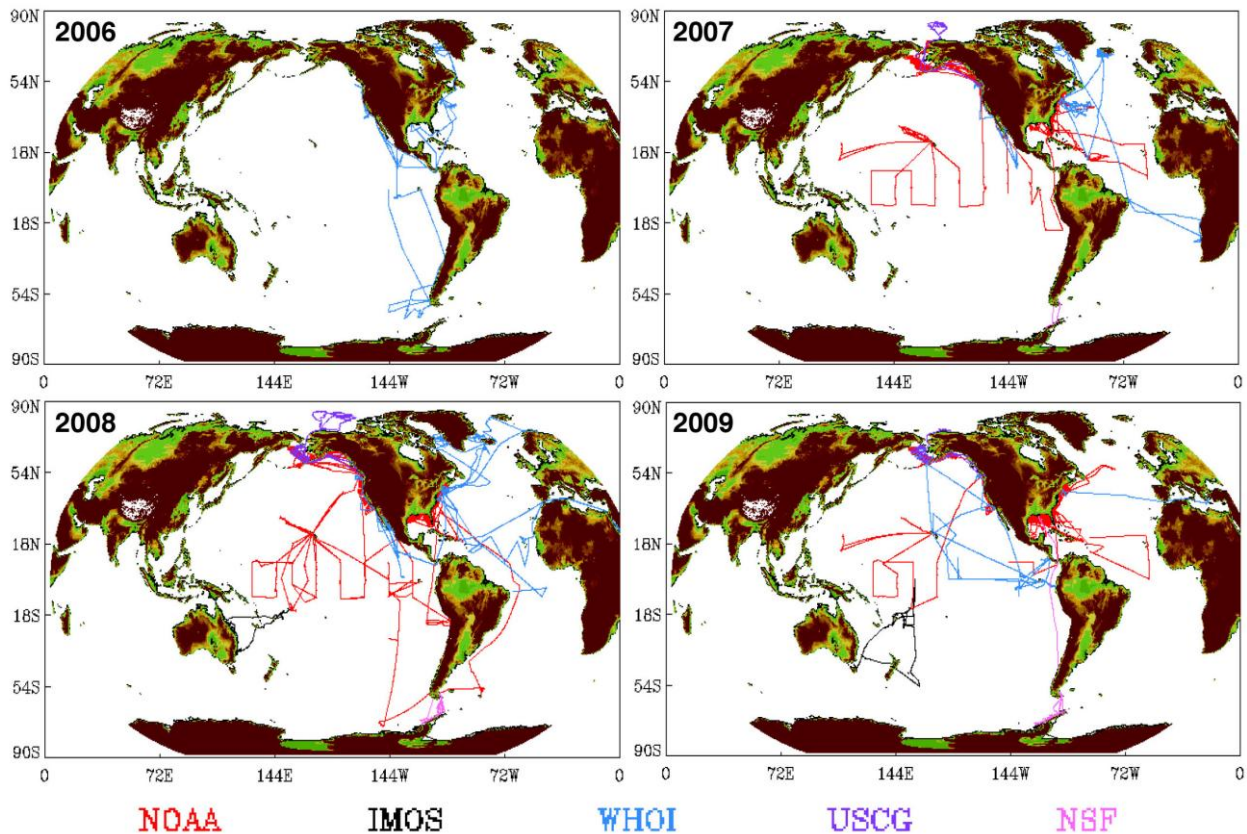


Figure 1. Cruise tracks of vessels providing SAMOS observations from 1 Jan 2006 through 20 April 2009. Tracks are color coded by vessel operator or national program.

## 2. STEWARDSHIP AND SOCIETAL BENEFIT

Scientific objectives of SAMOS include (1) creating quality estimates of the heat, moisture, momentum, and radiation fluxes at the air-sea interface; (2) improving our understanding of the biases and uncertainties in global air-sea fluxes; (3) benchmarking new satellite and model products; and (4) providing high-quality observations to support modeling activities (e.g. reanalysis) and global climate programs. Examples of scientific application of SAMOS and other underway observations from ships are provided in [11].

The SAMOS initiative evolved from the World Ocean Circulation Experiment (WOCE) DAC for marine meteorology. During WOCE, accessing meteorological data from RVs required personally contacting each vessel or chief scientist to obtain their underway data. This would result in delays of months to years between data collection and its accessibility for the user community. In many cases, meteorological and surface-ocean data for cruises were permanently lost. The SAMOS initiative therefore aims to capture and document a full record of underway meteorological and surface-oceanographic data from all cruises conducted

by participating SAMOS vessels. These data are freely and openly distributed to benefit the widest possible community. This requires automatic tracking of SAMOS observations and metadata from collection to archive. The SAMOS data management system includes both automated and manual QC methods, providing a value-added product to the community. This quality analysis includes two-way communication between the instrument technicians on SAMOS vessels and DAC personnel. Benefits of this communication include (1) the prompt notification of sensor malfunctions by the DAC directly to the vessel, reducing the period when incorrect data are collected; (2) the provision of DAC advice on sensor types and locations for proper exposure to the environment [2], improving data accuracy; and (3) the notification by technicians to the DAC of specific events that impact sensor operation (e.g. flow water system failures, shutdowns, or relocations), improving data quality. Visual QC identifies vessel-specific problems (e.g. improper sensor ventilation and radiometer shading) and has proven effective for developing advanced automated QC techniques. Extensive QC, both near real time and slightly delayed visual inspection, increases the return

on investments made by governments and institutions deploying expensive instrumentation on their RVs.

Expansion of SAMOS data stewardship activities will provide further benefits in coming years. Several current SAMOS vessels operate in remote oceans far outside routine shipping lanes, providing important data for developing satellite retrieval algorithms and constraining the range of air-sea fluxes in extreme environments. Recruitment of vessels operating in remote, under-sampled oceans (e.g. South Atlantic, South Pacific, and Southern Oceans) is a priority. New recruitments may initially bring in data of varying quality; however, routine implementation of SAMOS data and metadata standards, monitoring data quality, and providing operators best practice guidelines (e.g. [2] and [13]) and constructive feedback will improve data quality over time.

### 3. SAMOS DATA MANAGEMENT SYSTEM

#### 3.1. Data processing by SAMOS DAC

SAMOS observations originate with the computer logging system on participating vessels (Fig. 2). At the

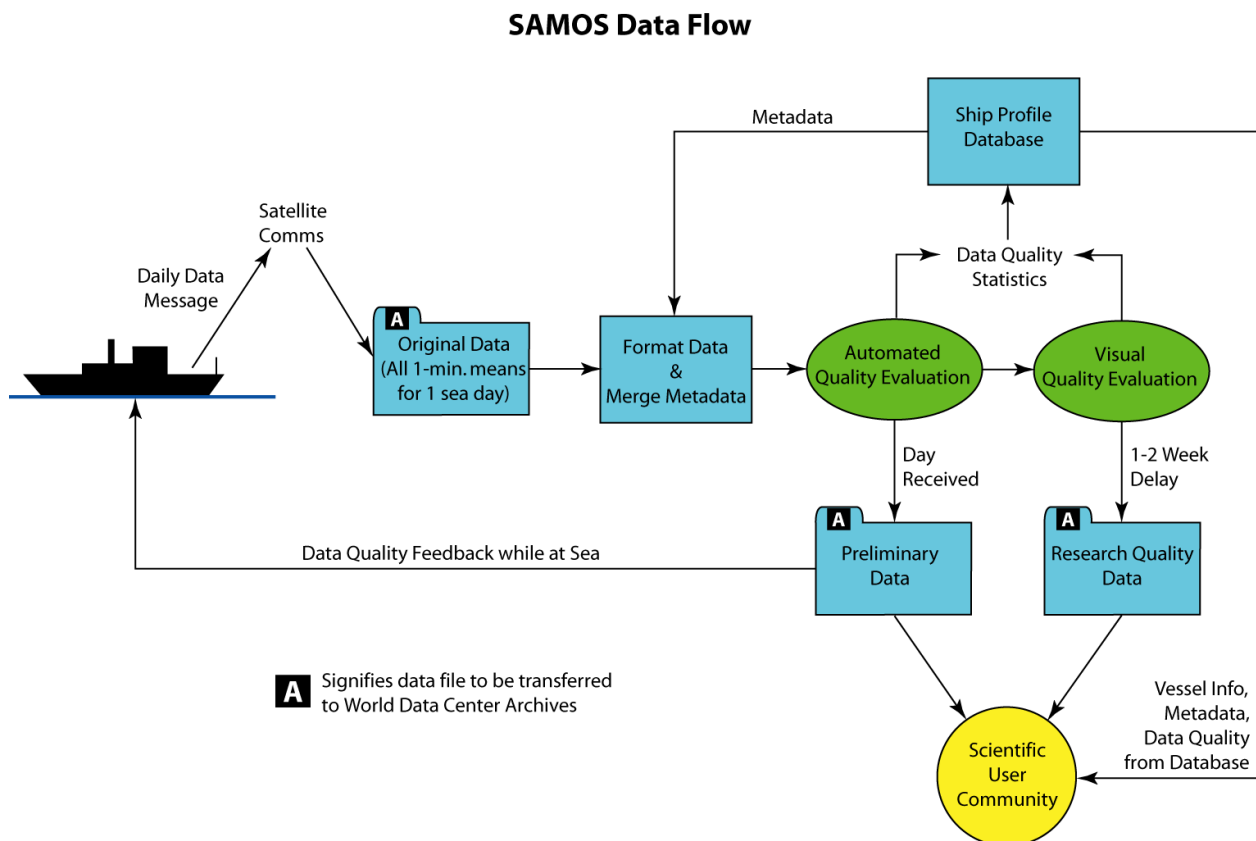


Figure 2. Overall flow of SAMOS data management system.

end of each observation-day, a ship-specific set of one-minute interval meteorological, oceanographic, and navigational parameters (Tab. 1) are bundled into a single ASCII file. The majority of vessels are exchanging data via a flexible “key: value” paired format designed by the SAMOS DAC. The SAMOS exchange format is preferred over standard SHIP or BUFR (Binary Universal Form for the Representation of meteorological data) coding used on the Global Telecommunication System (GTS) because it allows a wider range of parameters to be transmitted (e.g.

radiation, PAR, salinity, etc.) in a single message format and at higher precision. The daily file is attached to an e-mail (possibly compressed) and sent via satellite communication from the vessel directly to the SAMOS DAC. Nominally, daily files are transmitted shortly after 0000 UTC for each day at sea. The entire process onboard ship, up to e-mail receipt by the DAC, is designed and implemented by each individual vessel operator. For example, NOAA uses their Scientific Computer System software, while WHOI uses their Caliope software.

Primary	Secondary
Observation time (UTC)	Vessel pitch, roll, heave
Latitude and Longitude	Photosynthetically active radiation (PAR)
Ship speed over ground	Ultraviolet radiation
Ship course over ground	Net radiation
Ship heading	Visibility
Ship speed over water (fore-aft and along beam components)	Ceiling (i.e., height of lowest clouds)
Ship-relative wind speed and direction (as measured by anemometer)	Fluorescence
Earth-relative (true) wind speed and direction	Dissolved oxygen
Atmospheric pressure	Radiometric sea surface temperature
Air temperature	Swell and wave heights and directions
Humidity (wet-bulb/dew-point temperature, relative/specific humidity)	
Precipitation (or rain rate)	
Shortwave atmospheric radiation	
Longwave atmospheric radiation	
Sea temperature	
Salinity	
Conductivity	

*Table 1. Primary and secondary parameters requested as part of routine data acquisition from vessels contributing to the SAMOS DAC. One or more recruited vessels provide the highlighted parameters. The remaining secondary parameters are possible to measure with automated sensors, but at present are not provided routinely to the DAC.*

Preliminary SAMOS processing at the DAC begins with each attached file being extracted from the e-mail. Files are verified to ensure they are from a recruited vessel and in the expected format. Data are automatically converted from the received ASCII (American Standard Code for Information Interchange) to a SAMOS network common data form (netCDF). This process converts original units to the international system of units, ensures the data are in temporal sequence, and blends the observations with metadata from the ship profile database. After conversion, the observations undergo an automated quality evaluation. This initial evaluation verifies (1) the vessel is positioned over water, (2) the vessel speed between sequential positions is realistic, and (3) the observations are within realistic physical limits. Some parameters (e.g. pressure, temperature, wind speed, etc.) are flagged when they exceed  $\pm 4\sigma$  from a climatology [3], with a minimum standard deviation assigned in data sparse areas with unrealistically small standard deviations. Other tests

ensure that the relationship between air, wet-bulb, and dew-point temperatures is physical and true winds are properly calculated. This entire process occurs within 3-5 minutes of the e-mail arriving at the DAC.

After the preliminary data are processed, a trained analyst reviews all the previous day’s observations. The analyst ensures that data anticipated from recruited vessels are received in a timely manner and identifies gross malfunctions and problems. When problems occur, the analyst promptly notifies the vessel at sea. This notification greatly benefits the shipboard technicians and often results in quick repairs to systems.

Ten days after each observation day, intermediate files are automatically created by merging all preliminary files for a given observation day, allowing for receipt of delayed or corrected files from the vessel. 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. If the flags on the data values are identical, the second duplicate resolution test compares the values in question to the 30-minute mean centered 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 at the time in question and the situation stored in a processing log (a compromise to allow automation of the merge process).

After the merge is complete, an automated Spike And Stair Step Indicator (SASSI) is run to identify steps, spikes, and excessive noise in the data time series. One advantage of this approach is that these flags tend to cluster for conditions (e.g. ship-relative wind directions) adversely impacted by flow distortion. The range of suspect directions can easily be determined (for each instrument) from several days of observations, allowing automatic flagging in the future. A change in the directional distribution of these flags is also an indicator of changes in location of instruments or objects on the ship’s deck.

SASSI calculates the absolute value of first differences, and sorts them to determine the geophysical value associated with a user-defined percentile (typically the 99<sup>th</sup> percentile). The extreme 1% of first differences are not used in the calculation of a standard deviation of ‘typical’ values. The threshold used for determining suspect differences in one- or two-time steps is equal to this standard deviation multiplied by an analyst-determined scale factor. A minimum value of the threshold is set equal to the precision of the instrument; consequently, this threshold can also be used to identify periods of unrealistically low variability (e.g. when an instrument malfunctions). Both of the observations associated with a suspect difference are flagged. The mean values of unflagged observations before and after a series of flagged observations are used to determine if the flags are associated with a step, a spike, or excessive noise.

The final QC procedure uses the SAMOS VIsual Data Assessment Tool (SVIDAT), a graphical user interface developed by the DAC using the Interactive Data Language (IDL). A trained data analyst employs SVIDAT to verify all the automated QC and to add/delete/modify QC flags. One unique feature allows the analyst to flag interesting features in the data (e.g. fronts, extreme pressure minima, etc.). Visual quality evaluation results in research-quality files. These files are nominally produced on a 10-day delay once the intermediate files are available. In practice, the delay can be longer due to resource and time limitations for the data quality analysts.

All preliminary, intermediate, and research-quality data are distributed via web (<http://samos.coaps.fsu.edu>), ftp, and OPeNDAP servers. The final step of the data flow is archiving at NODC (Sect. 3.2). At present, only SAMOS observations from US vessels are archived by NODC.

The SAMOS DAC maintains a ship profile database containing extensive metadata for each vessel. Ship-specific metadata generally follow the conventions of the VOSCLim (Voluntary Observing Ship Climate) Project [6], while an extensive list of instrument-specific metadata is also required (e.g. 3-dimensional instrument position, units, calibration date, instrument type, averaging methods, etc.). Digital photos and vessel schematics are collected to aid with quality evaluation and to allow recommendations to vessel operators to improve sensor exposure. QC flags are updated in both the netCDF files and the database after each QC procedure. Additionally, hourly records are extracted from the one-minute netCDF files and stored in the database. This allows users to refine data searches on our web interface by data location and quality.

The DAC uses database technology to ensure robust file tracking and version control. Each original data file received is tracked and linked to subsequent quality-processed data files via a unique file version number. Version numbers, vessel call signs, and the observation date are stored in the database and can be used to build links to uniquely named data files and processing log files through the use of a structured file-naming convention.

### **3.2. Long-term preservation of SAMOS data**

NODC and the co-located World Data Center for Oceanography, Silver Spring, preserve and disseminate SAMOS and other ocean data in their Ocean Archive System (OAS). Archiving allows a wide community of users access to these data independent of the SAMOS initiative. To enable long-term preservation and dissemination of the SAMOS data with the OAS, which follows the Reference Model for an Open Archival Information System (OAIS) [8], SAMOS (the producer in the OAIS) and NODC (the archive) generate Submission, Archival, and Dissemination Information Packages (SIP, AIP, and DIP). Each SIP, generated by SAMOS, includes files that contain the original, preliminary, and research-quality data and metadata (e.g. file naming and format descriptions), and a message-digest algorithm 5 (MD5) checksum for each file to ensure their integrity. SAMOS copies each SIP to a server for NODC to pull. After pulling an SIP, each consisting of data collected during a month from one ship, NODC generates an AIP. Each AIP contains all the data and metadata of its associated SIP, plus information that NODC adds to manage the AIP and to

facilitate the dissemination of the data. For example, NODC adds an accession number (for tracking an AIP) and a Federal Geographic Data Committee standard-metadata record. After each AIP is completed, NODC makes it available online via two types of DIPs: the public may download individual files in the AIP—each file has a unique URL—or the entire AIP in one “tarball” file. In addition, users may find all the SAMOS AIPs by searching for SAMOS under “Contributing projects” on the OAS at <http://www.nodc.noaa.gov/Archive/Search>.

### 3.3. Adoption of SAMOS practices by IMOS

IMOS, which has adopted significant portions of SAMOS practices, could provide a model for other national observing programs. The requirements for shipboard observation sampling frequency, data reduction, and telemetry are as specified by SAMOS, allowing the creation of daily files containing one-minute averaged observations. All instruments and ships have comprehensive metadata records that are routinely updated and archived. The SAMOS netCDF has been adopted. The IMOS QC system generally follows the SAMOS procedures and adopts the SAMOS QC flag system. The theory used in SASSI is implemented in IMOS. The SVIDAT is written in the IDL language and was easily ported to the IMOS computing infrastructure. Coupled with the adoption of SAMOS file formats, implementation of SVIDAT is straightforward. The consistent processed file format and naming conventions across IMOS and SAMOS allow for seamless machine and human interpretation of these files, and later merging, comparison, archiving, and processing.

There are differences between IMOS and SAMOS due to project science objectives, individual preferences, and computing platforms. The SAMOS data processing code was not designed for portability to other computing platforms. The only code that has been adopted completely is the SVIDAT. A number of subroutines and fragments of the automated QC code have been used in the IMOS system. The data flow and file versioning differ. The automated QC system varies from SAMOS in a number of subtle ways: the order and precedence of the QC checks and flags have been changed; the sanity and climatology checks are performed as two separate tests; different climatology test sets are used; different ranges are used for the sanity test; the land test is based on a higher resolution land-sea mask; and ‘fuzziness’ at the coast is allowed to account for small errors in masks, position reporting and precision. The general meaning of the QC flag set is consistent with SAMOS, but the source of the QC processing and associated differences must be recognized by the data user to enable correct interpretation. Metadata standards are comparable, but platform and calibration metadata history is stored and

managed based on systems at the Australian Bureau of Meteorology. A dedicated IMOS DAC—eMarine Information Infrastructure (eMII)—manages data archiving and distribution. eMII has different data format requirements than SAMOS; however, the Australian Bureau of Meteorology maintains a parallel archive in SAMOS format.

## 4. THE NEXT DECADE

The SAMOS data management system is a model for the stewardship of automated underway meteorological and surface-oceanographic data. The SAMOS initiative has developed procedures for data and metadata, including acquisition, QC, distribution, and archiving, to meet the needs of a wide user community. The use of common data stewardship methods and structures support the aggregation and integration of SAMOS observations with other components of the ocean observing system. The SAMOS model for data stewardship, which has been adopted by the IMOS project, might provide a structure for evolving international efforts to manage underway observations.

The authors envision three options for expanding participation by other countries in the SAMOS data management system:

- Establish a distributed network of regional centers (e.g. SAMOS North America, Euro-SAMOS, Australian-SAMOS, etc.) to conduct data stewardship activities. This approach would require replication of the data management system at these centers using mutually agreed upon formats, metadata, procedures, and conventions. Current SAMOS data management tools would have to be ported for use by the centers. Regional centers would provide data through a distributed, but interconnected, data server system, and one or two world archive centers would house definitive copies of the SAMOS contributions from all regional centers. The ability to provide funding via existing national programs is one advantage of the regional centers. In addition, the centers might be housed at existing national oceanographic, weather, or data centers.
- Establish two international DACs, similar to the JCOMM structures used both by Argo [9], and by VOS via two Global Collecting Centers (GCCs, [14]). This approach would provide redundancy for data preservation and validation of operational procedures. This option has the advantage (over option 1) of being easier to manage (only two entities to ensure compatibility and convergence towards data and metadata standards). The two DACs could also have slightly different roles for the overall SAMOS data management to maximize

efficiency. The challenge with this approach is funding a multinational effort at only two national institutions. Some level of resources would be needed from participating countries to support the twin SAMOS DACs.

- Expand the existing SAMOS DAC to provide data stewardship for a multinational fleet. This approach would ensure uniform application of the SAMOS data management model to all SAMOS observations. The challenge would be funding a multinational effort at a single national institution.

All three options might benefit from the development of new working relationships or agreements at the JCOMM level, and in particular—due to increasing technological similarities developing between VOS and SAMOS data—the potential for closer cooperation and synergies with a planned modernization of the two GCCs [14].

Additional coordination will be needed between the SAMOS DAC and groups developing data discovery and integration tools. Several activities are underway to improve metadata acquisition (e.g. SensorML (OpenGIS (Open Geodata Interoperability Specification) Sensor Model Language)). The SAMOS DAC is working towards using the climate and forecast (CF) conventions for netCDF files [5] to ensure interoperability with other ocean data sources (e.g. OceanSITES [10]). As part of determining the best international framework for SAMOS data management, the international community will have to reconcile the variety of QC systems for underway oceanographic (e.g. Global Ocean Surface Underway Data [GOSUD] project) and meteorological data to ensure uniformity across national systems.

The focus of data acquisition activities in the next decade should involve obtaining routine SAMOS observations from a multinational fleet. There is a growing need for underway oceanographic and meteorological observations in coastal zones, regional seas, and remote oceans. Several areas of high scientific value were identified at the 2<sup>nd</sup> Joint GOSUD/SAMOS Workshop in 2008 [12], including the highly variable Southern Ocean, regions of low and high salinity variability, and regions of strong vertical and horizontal gradients. The ocean observing community should engage all countries, including developing nations, with capable vessels to consider contributing to the SAMOS initiative. In the past few years, the SAMOS DAC has approached or has been approached by several countries (including Brazil, Columbia, Greece, Germany, and France) interested in contributing observations and these collaborations need to be resourced and strengthened.

Another highly desirable activity will be to ensure that SAMOS observations from RVs reach climate archives (e.g. the International Comprehensive Ocean Atmosphere Data Set, ICOADS, [16]). VOS (and many other types of *in-situ* marine) data are contributed to ICOADS either through real-time data exchange on the GTS or through delayed-mode data transfer [14]. It is critical that SAMOS observations be clearly identified in climate archives so they can be excluded for archive comparison and model validation studies.

Delayed-mode transfer of WOCE meteorological data has occurred from the DAC to ICOADS [16] using the International Maritime Meteorological Archive (IMMA; [14]) format. IMMA supports embedded metadata and can hold more precise data than is possible using current JCOMM delayed-mode exchange formats. The DAC created hourly subsamples from the higher-frequency WOCE data and a similar approach for SAMOS would add ~175,000 hourly observations to ICOADS from the 21 ships contributing to the DAC between 2005-2009. It is expected that the modernization of data flow through the GCCs [14] will facilitate and formalize the integration of both SAMOS and VOS data into the WMO Integrated Global Observing Systems (WIGOS), the WMO Information System (WIS), and ultimately into climate datasets including ICOADS.

A SAMOS does not typically contribute to the exchange of data on the GTS in support of real-time applications (e.g. NWP), although some ships carrying SAMOS do participate in the VOS scheme using a distinct set of instruments provided by their national VOS operator. A barrier to participation can be that many RV operators are unfamiliar with operational user needs and in some cases may be reluctant to release underway oceanographic and meteorological data due to ownership concerns among chief scientists. The latter view is rarely a problem for meteorological data and views are changing for some underway-oceanographic data. Still there needs to be more promotion of the importance of contributing to the VOS scheme and a focus on establishing national policies on the release of SAMOS-type observations.

Another challenge is the need for carefully tracked instruments and associated metadata. The instruments used by SAMOS should meet the requirements of the VOS scheme, including regular calibration traceable to standards, but procedures need to be formally established. It should also be possible to provide the required metadata for the VOS scheme [7], but a mechanism needs to be developed for transfer of metadata from the SAMOS DAC to JCOMM, requiring co-operation between the DAC and national meteorological services. In practice, the SAMOS

metadata requirements are modeled on the VOS scheme and it should be possible for the SAMOS DAC to work with the national meteorological services to ensure the proper VOS recruitment forms are complete without unnecessary duplication.

Once all the instruments and procedures are approved, SAMOS observations could contribute to VOS. Data reports would be constructed at one- to six-hour intervals from subsamples of the high-resolution SAMOS data. These reports could come directly from the RV via INMARSAT (International Maritime Satellite) (or analogous methods) or be generated by the SAMOS DAC on a routine schedule. It might be challenging to meet the timeliness requirements for NWP if the latter route is chosen.

Finally, the input of manual elements of reports is highly desirable, but difficult to get people to do in practice when most of the observation is automated. There is a continued scientific need for both instrumental and visual observations of weather, precipitation, wave heights, clouds, and radiation [11].

In addition to ensuring that a subsample of SAMOS observations contribute to the VOS scheme, an opportunity will exist in the next decade to expand high-temporal frequency sampling from VOS that are currently deploying automated weather systems (AWS). Increasingly countries (e.g. France, Canada) are using AWS on their recruited VOS to produce hourly observations in support of national weather service operations. Although these AWS may not collect all the parameters listed in Tab. 1, they use WMO certified instruments and typically are capable of collecting observations at higher sampling rates. Making these SAMOS-type observations available would greatly expand data coverage. These data would likely have to be stored onboard the vessel until it reaches port, since the AWS-equipped VOS may not have the capability or resources to move these data to shore in real-time. A delayed-mode process could be envisioned whereby these rapid-rate AWS data are either (1) physically collected from the vessel at regularly scheduled port stops by a port meteorological officer or technician or (2) are automatically downloaded through a wireless connection with the vessel (when it is within range of a suitable shore site). The latter option would require some technology development.

## 5. NEW INITIATIVES

The SAMOS DAC is partnering with the Rolling deck To Repository (R2R; <http://www.rvdata.us/overview>) project. Funded by the National Science Foundation, R2R is developing a protocol for transferring all underway data (navigation, meteorology, oceanographic, seismic, bathymetry, etc) collected on

US University National Oceanographic Laboratory System (UNOLS) RVs to a central shore repository. Currently, the only UNOLS vessels contributing to the SAMOS DAC are those operated by WHOI. The focus of the R2R is capturing all these data at the end of each planned cruise; however, the SAMOS DAC will be developing a real-time component to transfer a subset of meteorological and surface-oceanographic data from ship to shore. The data will be transferred at the full observational resolution for the specified sensor (in some cases up to 1Hz samples) on a yet to be determined transfer schedule. The transfer protocol will take full advantage of the evolving broadband satellite communication technology. The SAMOS DAC will then develop a modified data management system that will pull the real-time data from the R2R and process it into SAMOS-formatted observations (identical to our current records).

The DAC is planning to develop an automated ship-to-shore metadata transfer protocol in collaboration with the R2R, UNOLS, NOAA, and any other interested RV operators. The primary hurdle will be defining a metadata transfer format. The SAMOS DAC envisions examining ongoing efforts (e.g. SensorML) and is willing to work with the international community to produce a workable standard.

Finally, as the ocean observing system expands beyond ships for mobile surface-ocean monitoring, the developers of the SAMOS data management system welcome operators of autonomous surface vessels (ASVs) to consider our protocols for their own data management. ASVs are becoming viable [11] and a plan should be developed in the coming decade to ensure the careful management of ASV underway observations.

## 6. RECOMMENDATIONS

The following recommendations and suggestions for implementation are made to advance the stewardship of SAMOS observations from RVs, merchant ships, and possibly future ASVs:

1. Conduct a forum with representatives from all interested nations and organizations to consider further expanding SAMOS activities to the international community. The forum should examine the three alternatives for SAMOS data management centers outlined herein, discuss the strengths and limitations of the current SAMOS data management system, and address the need for sustained funding to capitalize on national investments in shipboard instrumentation. The potential for developing new linkages with the VOS GCCs as part of their upcoming modernization should be actively explored in this context.



2. Engage developing nations in obtaining new observations within their territorial waters and establish procedures to ensure easy access to SAMOS observations for use by these nations.
3. Continue to integrate SAMOS activities with operational and climate data collection activities through enhanced links with JCOMM, the GCCs, national VOS operators, and ICOADS.
4. Focus additional resources on collecting SAMOS data in regions of high scientific value. Engage the scientific community to further define these high-priority regions.

Implementing these recommendations requires an open dialog between interested nations and organizations. One potential venue is a planned joint GTSP (Global Temperature Salinity Profile Program) /GOSUD/SAMOS workshop to be held in Oostende, Belgium in the spring of 2010. This would provide an opportunity to bring together stakeholders to discuss the future data management structure for SAMOS and to forge partnerships with international marine *in-situ* observational programs. Linkages between SAMOS and VOS will also continue to be strengthened via SAMOS collaboration with the JCOMM Ship Observation Team. Expanded international participation is welcome at future SAMOS workshops, which bring together the shipboard observers, technical experts, data managers, and scientific data users. We challenge the scientific community to provide input on the highest priority regions for marine atmospheric and surface-oceanic observations to provide guidance to SAMOS plans to recruit additional research and merchant vessels.

5. Automate ship-to-shore delivery of cruise, ship, and instrument metadata, since most metadata changes from cruise to cruise.

Efforts are underway within the US to develop a protocol for automating metadata transfers. These developments should take advantage of international metadata standards and allow for easy interoperability between a wide data management and user community.

6. Develop reliable and routinely updated information on upcoming RV cruises and the instruments deployed on each vessel (a recommendation also made by the hydrographic [6] and the overall underway ship observing [11] communities).
7. Ensure vessels on high-priority cruises are adequately staffed and/or in frequent communication with technical personnel to monitor and maintain the instrument systems. The at-sea technicians are critical to obtaining the highest quality observations.

The primary issue behind the latter two recommendations is providing adequate resources. Several groups are in position to help develop and maintain a comprehensive RV database (e.g. POGO, JCOMMOPS), but these efforts must be adequately funded. In addition, more interaction is needed between these groups and national efforts to collect and inventory RV cruise data and metadata (e.g. R2R in the US). Adequate shipboard or shoreside technical personnel must be maintained to ensure the quality of automated instruments on ships. Efforts should be made to retain talented personnel and to reinvigorate the network of port meteorological officers, with retraining as necessary to ensure their ability to work with modern automated instrument and communications systems.

## 7. REFERENCES

1. The Australian Integrated Marine Observing System, 2008: *J. Ocean Technology* 3(3), 80-81.
2. Bradley, F. and C. Fairall, C., 2007: A guide to making climate quality meteorological and flux measurements at sea. NOAA Tech. Memo. OAR PSD-311, NOAA/ESRL/PSD, Boulder, CO, 108 pp.
3. da Silva, A. M., C. C. Young, and S. Levitus, 1994: *Atlas of Surface Marine Data, Vol. 1: Algorithms and Procedures*. NOAA Atlas NESDIS 6, U. S. Department of Commerce, NOAA, 83pp. Available from: NOAA/NODC, Customer Service, E/OC, 1315 East-West Highway, Silver Spring, MD 20910.
4. Fairall, C. & Co-Authors (2010). "Observations to Quantify Air-Sea Fluxes and their Role in Climate Variability and Predictability" in these proceedings (Vol. 2), doi:10.5270/OceanObs09.cwp.27.
5. Hankin, S. & Co-Authors (2010). "NetCDF-CF-OPeNDAP: Standards for Ocean Data Interoperability and Object Lessons for Community Data Standards Processes" in these proceedings (Vol. 2), doi:10.5270/OceanObs09.cwp.41.
6. Hood, M. & Co-Authors (2010). "Ship-Based Repeat Hydrography: A Strategy for a Sustained Global Program." in these proceedings (Vol. 2), doi:10.5270/OceanObs09.cwp.44.
7. Kent, E. & Co-Authors (2010). "The Voluntary Observing Ship (VOS) Scheme" in these proceedings (Vol. 2), doi:10.5270/OceanObs09.cwp.48.
8. OAIS, 2002: Reference Model for an Open Archival Information System (OAIS). Recommendation for Space Data System Standards, CCSDS 650.0-B-1. Blue Book. Issue 1. Washington, D.C.: CCSDS, January 2002. [Equivalent to ISO 14721:2003.]
9. Pouliquen, S., Schmid, C., Wong, A., Guinehut, S. and Belbeoch, M., (2010). "Argo Data Management" in these proceedings (Vol. 2), doi:10.5270/OceanObs09.cwp.70.

10. Send, U. & Co-Authors (2010). "OceanSITES" in these proceedings (Vol. 2),  
doi:10.5270/OceanObs09.cwp.79.
11. Smith, S. & Co-Authors (2010). "Automated Underway Oceanic and Atmospheric Measurements from Ships" in these proceedings (Vol. 2),  
doi:10.5270/OceanObs09.cwp.82.
12. Smith, S. R., and L. Petit de la Villéon, 2008: Report of the 2nd Joint GOSUD/SAMOS Workshop. IOC Workshop Report #218, 72 pp. [Available online at [http://www.coaps.fsu.edu/RVSMDC/marine\\_workshop4/docs/report\\_v3\\_1.pdf](http://www.coaps.fsu.edu/RVSMDC/marine_workshop4/docs/report_v3_1.pdf).]
13. Smith, S. R., M. A. Bourassa, and R. J. Sharp, 1999: Establishing more truth in true winds. *J. Atmos. Oceanic Technol.* **16**(7), 939-952.
14. Woodruff, S. & Co-Authors (2010). "Surface In Situ Datasets for Marine Climatological Applications" in these proceedings (Vol. 2),  
doi:10.5270/OceanObs09.cwp.93.
15. Woodruff, S., 2007: Archival of Data Other than in IMMT Format: The International Maritime Meteorological Archive (IMMA) Format. Second Session of the JCOMM Expert Team on Marine Climatology (ETMC), Geneva, Switzerland, 26-27 March 2007, JCOMM Meeting Report No. 50, 68-101.
16. Worley, S. & Co-Authors (2010). "The Role of the International Comprehensive Ocean-Atmosphere Data Set in the Sustained Ocean Observing System" in these proceedings (Vol. 2),  
doi:10.5270/OceanObs09.cwp.94.