Distributed Oceanographic Match-Up Service (DOMS)
Translation Specification: SAMOS In Situ Data
Version 3, August 2018

Adam Stallard, Shawn R. Smith, and Jocelyn Elya

Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, FL
Contact: smith@coaps.fsu.edu

Revision History

Version 1 – Preliminary documentation created during DOMS prototype development.

Version 2 – Translation specification in use for the first published DOMS prototype (April 2017).

Version 3 – Specification modified to include mapping of SAMOS quality control flags to the International Oceanographic Data Exchange (IODE) flag scheme. This mapping is being implemented during the integration of DOMS into the OceanWorks data analytics platform at JPL. New quality fields are added to store separated IODE quality flags for wind speed and wind components along with the original SAMOS A-Z quality flags for wind speed and direction.

Introduction

To make the DOMS matchup output meaningful and easy to use the collaborating partners must standardize their data within the DOMS prototype. This document defines the translation of SAMOS data from their native netCDF files on the SAMOS THREDDS server into Apache Solr (the in-situ indexing approach chosen for DOMS).

All SAMOS data is extracted from netCDF files on a THREDDS server, using the following non-standard libraries for python (v3.7):

- numpy (for fast data operations)
- netCDF4 (for working with netCDF datasets)
- siphon (for interacting with THREDDS)
- solrpy3 (for populating solr index directly from python)

The code that performs this data translation is located within the “doms” python package developed by Adam Stallard at COAPS. The package resides on github, and is accessible by COAPS personnel and a few members of the wider DOMS team.

Modifications to the code to support version 3.0. of the specification were made by Homer McMillan and Jocelyn Elya.

Date and time:

Convert SAMOS netCDF time convention value to ISO 8601.
SAMOS netCDF field: time
SAMOS netCDF description: number of minutes since January 1st, 1980.
ISO 8601 form: YYYY-MM-DDThh:mm:ssZ
CF name = time

Latitude:

No conversion
SAMOS netCDF field: lat
SAMOS netCDF description: decimal degrees with +North and -South
Precision: .0001f
CF name = latitude

Longitude:

Convert from SAMOS convention (0.00 to 359.99) to DOMS convention (-179.99 to 180.00)
SAMOS netCDF field: lon
SAMOS netCDF description: decimal degrees (-180 West to +180 East)
Precision: .0001f
CF name = longitude

Platform:

Set static value for all SAMOS data

DOMS Index Code	DOMS Description
1	ship

Device:

SAMOS data does not carry device information. However, this may be added in a future update to the SAMOS dataset.

DOMS Index Code	DOMS Description
0	unknown

Mission:

Set = 1, SAMOS

Quality:

SAMOS netCDF variable: flag
SAMOS netCDF dimensions: time, qcindex
Note: qcindex is an attribute associated with every variable for which quality information is stored. The qcindex is used as the 2nd dimension's index into the flag variable.

*_quality = IODE Flag as noted in the SAMOS data quality flag mapping (Table A1; Appendix 1 for details), where * represents an indexed physical parameter (see below).

*_SAMOS_quality = original alphabetic SAMOS flag directly from the "flag" array, where * represents an indexed physical parameter (see below)

Records with date, time, latitude, or longitude that have ANY flag other than Z, are excluded from DOMS entirely.

Sea_water_temperature:

No conversion.
SAMOS netCDF variable: TS
SAMOS netCDF description: Measured temperature in degrees Celsius
Precision: .01f
Value provided when available. A SAMOS value of -9999 or -8888 (indicating "missing" or "special") is set to null in DOMS.
CF name = sea_water_temperature

Sea_water_temperature_depth:

SAMOS netCDF variable: TS
SAMOS netCDF attribute: height
If available, the value is recorded and the sign is reversed to represent positive DOMS depth. A SAMOS value of -9999 or -8888 (indicating "missing" or "special") is set to null in DOMS.

Sea_water_temperature_quality:

IODE numeric flag for sea water temperature. See "quality" (above).

Sea_water_temperature_SAMOS_quality:

SAMOS alphabetic flag for sea water temperature. See “quality” (above).

Sea_water_salinity:

No conversion.
SAMOS netCDF variable: SSPS
SAMOS netCDF description: Calculated observation measured as PSU
Value provided when available. A SAMOS value of -9999 or -8888 (indicating “missing” or “special”) is set to null in DOMS.
CF name = sea_water_salinity

Sea_water_salinity_depth:

SAMOS netCDF variable: SSPS
SAMOS netCDF attribute: height

If available, the value is recorded and the sign is reversed to represent positive DOMS depth. Value provided when available. A SAMOS value of -9999 or -8888 (indicating “missing” or “special”) is set to null in DOMS.

Sea_water_salinty_quality:

IODE numeric flag for sea water salinity. See “quality” (above).

Sea_water_salinty_SAMOS_quality:

SAMOS alphabetic flag for sea water salinity. See “quality” (above).

Wind_speed:

No conversion.
SAMOS netCDF variable: SPD
SAMOS netCDF description: Earth relative wind speed calculated as m/s based on relative wind, heading, and GPS SOG and COG
Precision: .1f
Value provided when available. A SAMOS value of -9999 or -8888 (indicating “missing” or “special”) is set to null in DOMS.
CF Name = wind_speed

Wind_speed_quality:

IODE numeric flag for wind speed. See “quality” (above).

Eastward_wind and Northward_wind:

SAMOS netCDF variable: SPD, DIR
SAMOS netCDF description: See wind speed for SPD. DIR is calculated as degrees clockwise from true north.
Precision (DIR): 1.f

Value provided when available. A SAMOS value of -9999 or -8888 (indicating “missing” or “special”) is set to null in DOMS.
If both SPD and DIR are available (values not -9999 or -8888), calculate vector components to derive the following:

eastward_wind, as positive east, with precision = 0.1 m/s
northward_wind, as positive north, with precision = 0.1 m/s

regardless of the value of the wind_speed_SAMOS_quality or wind_direction_SAMOS_quality (see below).

The conversion equation written in python:

```
Import numpy as np
def wind(dir, spd):
    '''Given a direction and speed, return vector components'''
    if dir == None or spd == None:
       return None, None
    dtor = np.pi / 180.0 # decimal to radians
    mdir = 270.0 - float(dir)
    if mdir <= 0.0:
       mdir += 360.0
    else:
       mdir -= 360.0
    u = float(spd) * np.cos(mdir * dtor)
    v = float(spd) * np.sin(mdir * dtor)
    return u, v
```

If any of SPD or DIR are missing, these values are omitted.

Note that a SAMOS value of -9999 or -8888 is interpreted as “missing” or “special” and is left out of DOMS. In this case, this applies when either SPD or DIR have these values.

CF Name = eastward_wind, northward_wind

Wind_component_quality:

IODE numeric flag that applies to both eastward_wind and northward_wind values. See quality (above).

In this case, the IODE flag is assigned using on the “worst” SAMOS flag associated with the wind speed (SPD) and direction (DIR) at a given time. E.g., if SPD flag is Z, but DIR flag is J – this flag is set to IODE flag 4 according to the mapping in table A1.

Wind_speed_SAMOS_quality:

SAMOS alphabetic flag for wind speed. See quality (above).

Wind_direction_SAMOS_quality:

SAMOS alphabetic flag for wind direction. See quality (above).

Wind_depth:

SAMOS netCDF variable: SPD, DIR
SAMOS netCDF attribute: height
The value of height is set using heights recorded SPD and DIR, with the height for SPD taking precedence over that of DIR.
If available, the value is recorded and the sign is reversed to represent negative DOMS depth convention for sensors located above sea level. A SAMOS value of -9999 or -8888 (indicating “missing” or “special”) is set to null in DOMS.
Wind_depth applies to the wind_speed, eastward_wind, and northward_wind values.

Meta:

SAMOS specific metadata string created for DOMS, not internal to SAMOS.
Has the following pattern:
{Call sign}_{YYYYMMDD}v{version number/order number}_{index value}
Example: "KAOU_20131222v20001_0766" represents a record from the ship with callsign “KAOU”, on 2013, December 22, using the intermediate quality (v200) first order (01) datafile, located at time-index 766 within the file.
CF Name = meta

Provenance:

The source of each record is populated using a THREDDS url with the appropriate netCDF dataset record and index value associated to each record.

Appendix 1: SAMOS to IODE flag mapping

Mapping from the alphabetic quality flags used by the SAMOS initiative (Smith et al. 2018) to the IODE primary flag scheme (Table A2; Paris 2013), selected for use by the OceanWorks project, is outlined in Table A1. The rationale for some mapping decisions is provided.

Table A1: Definitions of the alphabetic flags used in the SAMOS quality control procedures and a proposed mapping to the IODE standard.

IODE Flag	SAMOS Flag	Definition
4	**B**	Original data were out of a physically realistic range bounds outlined.
4	**D**	Data failed the T>=Tw>=Td test. In the free atmosphere, the value of the temperature is always greater than or equal to the wet-bulb temperature, which in turn is always greater than or equal to the dew point temperature.
3	**E**	Data failed the resultant wind re-computation check. When the data set includes the platform's heading, course over the ground, and speed over the ground along with platform relative wind speed and direction, a program re-computes the earth relative wind speed and direction. A failed test occurs when the difference between the reported and re-computed wind direction is >20 (or >2.5 m/s for wind speed).
3	**F**	Platform velocity unrealistic. Determined by comparing sequential latitude and longitude positions.
3	**G**	Data are greater than 4 standard deviations from the climatological means (da Silva et al. 1994). The test is only applied to pressure, temperature, sea temperature, relative humidity, and wind speed data.
3	**H**	Discontinuity (step) found in the data. Flags assigned to the maximum and minimum points in the discontinuity.
1	**I**	Interesting feature found in the data. Examples include: hurricanes passing stations, sharp seawater temperature gradients, strong convective events, etc.
4	**J**	Data are of poor quality by visual inspection, DO NOT USE.
3	**K**	Data suspect/use with caution – this flag applies when the data look to have obvious errors, but no specific reason for the error can be determined.
4	**L**	Vessel position over land based on reported latitude and longitude.
4	**M**	Known instrument malfunction.
3	**N**	Signifies that the data were collected while the vessel was in port. Typically these data, though realistic, are significantly different from open ocean conditions.
3	**Q**	Questionable – observation reported as questionable/uncertain in consultation with vessel operator (use with caution).
4	**S**	Spike in the data. Usually one or two sequential data values (sometimes up to 4 values) that are drastically out of the current data trend. Spikes for many reasons including power surges, typos, data logging problems, lightning strikes, etc.
1	**Z**	Data passed evaluation.

Table A2: Primary level IODE quality control flag definitions.

Value	Primary-level flag short name	Definition
1	Good	Passed documented required QC tests
2	Not evaluated, not available or unknown	Used for data when no QC test performed or the information on quality is not available
3	Questionable/suspect	Failed non-critical documented metric or subjective test(s)
4	Bad	Failed critical documented QC test(s) or as assigned by the data provider

9	Missing data	Used as place holder when data are missing

Notes: No SAMOS data are mapped to IODE flag 2 as all SAMOS data values undergo minimum automated QC tests. When the value for an indexed parameter (sea water temperature, sea water salinity, or winds) is missing, the corresponding IODE flag is set to 9.

Rationale:

The decision on how to map individual SAMOS flags to IODE was reviewed by the SAMOS data quality analysts and project investigators. Most mappings were easy to agree upon, but there are a few SAMOS flags whose meaning and usage falls in between IODE suspect (3) and bad (4) flags. The final decision on the mapping was weighted heavily on the input from the SAMOS visual data quality evaluator's expertise. We note that the IODE flags are designed to support data filtering at a very broad level. Users are encouraged to examine the original SAMOS flags when they have questions regarding their specific application.

The true wind test (E flag) is mapped to suspect (3) in the IODE system for two reasons. First there is always uncertainty in the recalculated true wind value created by the SAMOS software since we may not know exactly which relative wind and navigation parameters (ship speed, course, and heading) were used by the operator to derive their reported true wind. Second, the thresholds used by the test for direction and speed mismatch may be too stringent (or not strict enough) depending on the user's application. This test was always designed as a warning that there may be problems with the true winds, but not as a marker that the true winds are obviously incorrect.

The SAMOS unrealistic ship velocity (F) flag has been mapped to suspect (3) in the IODE system. A great deal of the F flags result either from coarse resolution of or averaging issues within the position data. Habitually the SAMOS data analyst has not removed these F-flags, as the reason for F-flags is less clear than for vessel over land (L) flags in tight waterways, for example. Granted there are plenty of cases that the position data are obviously bad, but the analyst leans towards expecting the majority of F flags are probably not truly bad data. The user must be aware that there will be obviously bad position data with F-flags, but since this flag is rarely applied (~0.1% of the data in 2017), the conservative approach was to consider F-flags to be suspect, not always bad.

In port data (N flag) are also considered suspect (IODE category 3), since there are cases where this flag is assigned when the vessel was known to be in drydock. Although perhaps that still leaves the N-flagged position data themselves as "good...", the analyst notes the N flag on the position data is more an indicator about the quality of the associated meteorological or oceanographic data. When in port, these parameters are almost always different from open ocean conditions and frequently are suspect because the operator is either running test on the sensor suite, cleaning sensors, etc.

Spikes in the data (S-flag) are mapped to bad (4) in the IODE system. Although spikes can occur as part of real weather events (e.g., changes in winds or pressure associated with a thunderstorm), the vast majority (~95%) of the S flags applied by the data analysts represent

truly bad data. The analyst also notes that not all spikes that are flagged are “drastically” out of the norm for the time series. In fact, they often may be quite small but obvious, for example in the case of sea temperature data where it's clear there's air intrusion, especially for vessels already known to have a shallow intake (or for cases of repeated, obvious electrical interference). Here again, though, maybe it's a question of quantity. Spikes also tend to be infrequent when looking at the overall data picture (~0.4% of the data in 2017), so even if some good values are filtered by setting the IODE flag to bad (4), the loss of data values would be small.

References

Paris. Intergovernmental Oceanographic Commission of UNESCO. 2013. Ocean Data Standards, Vol.3: Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine Meteorological Data. (IOC Manuals and Guides, 54, Vol. 3.) 12 pp. (English.) (IOC/2013/MG/54-3)

Smith, S. R., K. Briggs, M. A. Bourassa, J. Elya, and C. R. Paver, 2018: Shipboard automated meteorological and oceanographic system data archive: 2005–2017. *Geosci Data J.*, **5**, 73–86. https://doi.org/10.1002/gdj3.59