ORCAS merge data product readme v. 20170316

The ORCAS instrumentation is described at https://www.eol.ucar.edu/content/orcas-aircraft-instruments and instrument specific readme files are available at the ORCAS data access page http://data.eol.ucar.edu/master_list/?project=ORCAS. The types of data collected in the project include 1) high-rate (~ 1 Hz) in situ measurements from many instruments, 2) in situ gas-chromatography (GC) measurements on samples collected over a short (~ 30 sec) period of time but spaced by several minutes from the TOGA instrument, 3) laboratory measurements on whole air samples collected in pressurized metal canisters over short (~ 10-20 sec) periods of time by the AWAS sampler, and 4) laboratory measurements on whole air samples collected in continuously flowing glass flasks with longer (~ 30 sec) volume mixing times by the Medusa sampler. The wide range of representative sample times and the unique nature of each measurement system necessitate several different merged data products.

1. File format

All merge products are flat ascii tables, space delimited, with variable names in the first row. Missing values are given as "NA". The data columns are grouped by source and include:

a. **Date/time variables:** Year, flt (flight number), DOY (day of year), and UTC (UTC seconds since midnight on the day of takeoff). Flight numbers of -3, -2, -1, and 0 correspond to TF01, TF02, FF01, and FF02; flight numbers 1-19 correspond to RF01-RF19, and flight number 20 corresponds to FF03.

b. **RAF variables:** measurements from the suite of RAF supported sensors are described in the ORCAS Project Manager report and RAF bulletins listed here https://www.eol.ucar.edu/node/11647

c. **Science payload variables:** measurements from the science payload have the instrument name appended as _AO2, _NOAA, _QCLS, _TOGA, _AWAS, or _CLH and are described in the instrument specific readme files at http://data.eol.ucar.edu/master_list/?project=ORCAS

d. **Forecast model output:** output from the forecast configuration of the Community Earth System Model (CESM) run during the campaign is listed with _FMPD appended to the variable name. The model simulates tracers responsive to different surface fluxes. The tracers with ocean only fluxes have _OCN in the variable name, land only components include _LND in the variable name, and fossil-fuel components include _FFF in the variable name. CO2_FMPD_SUM is the sum of the individually transported ocean, land, and fossil CO_2 components. CO2_FMPD_DIF is the difference between CO2_FMPD transported as total CO_2 and CO2_FMPD_SUM. ocean_air_FMPD is the modelled

fraction of air that made most recent surface contact with ocean rather than land. socn_air_FMPD is the modelled fraction of air that made most recent surface contact with the Southern Ocean.

e. **Profile numbers:** The merge products include an "n.prof" variable sequentially numbering all 215 vertical profile maneuvers in the campaign. A value of 0 for n.prof indicates times between profiles.

f. **.X "recommended" variables:** The merge products include the variables CO2.X, CH4.X, CO.X, and H2O.X, which represent the recommended values to use in the case of multiple instruments measuring the same gas.

   - CO2.X is a copy of CO2_NOAA for all flights where CO2_NOAA is available, CO2_QCLS on TF01, and CO2_AO2 on FF01 and FF02.
   - CH4.X is a copy of CH4_NOAA for all flights where CH4_NOAA is available, and CH4_QCLS for the first half of TF01.
   - CO.X is a copy of CO_QCLS for all flights where CO_QCLS is available, and CO_NOAA for the first 1/3 of RF01 and all of flights RF05, RF10, RF13, RF15, RF16, and FF03.
   - H2O.X is a copy of VMR_VXL for all flights.

2. Merge products

a. **ORCASall.merge10.tbl:** All high-rate in situ data, averaged by 10 second bins. The "UTC" variable corresponds to the mean of the 10 sec interval used for averaging. We recommend using this file for analyses of high rate species, as time matching and sensor precision are better at 10 sec than 1-Hz, without significant loss of information. The n.prof column gives medians of the 1-Hz n.prof values to avoid non integer values at profile edges.

b. **ORCASall.merge.tbl:** All high-rate in situ data, at 1 Hz resolution. This file is provided should users require 1 Hz data, but care should be taken as it is generally not possible to match inlet tubing delays at this time resolution for instruments located in the cabin.

c. **ORCASall.merge60.tbl, ORCASall.merge60n.tbl, and ORCASall.merge60sd.tbl:** All high-rate in situ data, averaged by 60 second bins. These files may be of a more manageable size for some users, and with better precision from averaging for many species, however with a nominal GV ascent/descent rate of 500 m per minute, some fine atmospheric structure will be lost in the averaging. The n.prof column gives medians of the 1-Hz n.prof values to avoid non integer values at profile edges. ORCASall.merge60n.tbl gives the number of 1-Hz values used for each variable within each minute while ORCASall.merge60sd.tbl gives the standard deviation of 1-Hz values for each variable within each minute. ORCASall.merge60wt.tbl and ORCASall.merge60sd.tbl may be useful for filtering ORCASall.merge60.tbl to exclude data points with very little or highly variable corresponding 1-Hz data.

d. **ORCASall.mergeTOGA.tbl, ORCASall.mergeTOGAwt.tbl, and ORCASall.mergeTOGAsd.tbl:** All high-rate in situ data, averaged between the start and stop times of each discrete TOGA sample. Values in columns to the right of the n.prof column come directly from the individual flight TOGA files, while values in columns to the left of the n.prof column are means of 1-Hz data. The n.prof column gives medians of the 1-Hz n.prof values to avoid non integer values at profile edges. Values of -888 indicate concentrations below the instrumental limit of detection. ORCASall.mergeTOGAwt.tbl gives the fraction of 1-Hz values available for each variable within each TOGA sample period while ORCASall.mergeTOGAsd.tbl gives the standard deviation of 1-Hz values for each variable within each TOGA sample period. ORCASall.mergeTOGAwt.tbl and ORCASall.mergeTOGAsd.tbl may be useful for filtering ORCASall.mergeTOGA.tbl to exclude data points with very little or highly variable corresponding 1-Hz data.

e. **ORCASall.mergeAWAS.tbl, ORCASall.mergeAWASwt.tbl, and ORCASall.mergeAWASsd.tbl:** All high-rate in situ data, averaged between the start and stop times of each discrete AWAS sample. Values in columns to the right of the n.prof column come directly from the AWAS data file, while values in columns to the left of the n.prof column are means of 1-Hz data. The n.prof column gives medians of the 1-Hz n.prof values to avoid non integer values at profile edges. Values of -8888 indicate concentrations below the instrumental limit of detection. ORCASall.mergeAWASwt.tbl gives the fraction of 1-Hz values available for each variable within each AWAS sample period while ORCASall.mergeAWASsd.tbl gives the standard deviation of 1-Hz values for each variable within each AWAS sample period. ORCASall.mergeAWASwt.tbl and ORCASall.mergeAWASsd.tbl may be useful for filtering ORCASall.mergeAWAS.tbl to exclude data points with very little or highly variable corresponding 1-Hz data.

f. **ORCASall.mergeMED.tbl:** All high-rate in situ data, weighted and averaged according to the Medusa flask sampling kernel for each flask. The Medusa flask sampler collects air into glass flasks that are continuously purged at a controlled flow and pressure with a mixing time of around 30 sec and a representative sampling kernel that drops off exponentially with a several-minute tail. Values in columns to the left of the UTC column come directly from the individual flight Medusa files, with the exception that MED_CO2 has been adjusted for differences between the Scripps O_2 Laboratory and WMO X2007 CO_2 scales, on the basis of cylinders measured by both Scripps and NOAA. The equation for this adjustment is:

SIO_to_WMO = - 0.00001796 * MED_CO2^2 + 0.0101942 * MED_CO2 - 1.083

Where MED_CO2 is from the individual flight Medusa files on the Scripps O_2 Laboratory CO_2 scale, and the adjustment at 400 ppm equals +0.12 ppm. Values from the UTC column to the n.prof column are weighted means of 1-Hz data. Columns to the right of the n.prof column with the variable name prefix “wt.” give the weighted fraction of 1-Hz data available for each variable within each Medusa sample averaging period. These

weight variables may be useful for filtering the mean variables to exclude data points with very little corresponding 1-Hz data.