Data Used in Global Ocean Modeling and Data Assimilation Systems at NOAA/NCEP

Shastri Paturi, Yan Hao, Jong Kim
IMSG @ NOAA/NWS/NCEP/EMC, College Park, MD 20740
Email: Shastri.Paturi@noaa.gov

The National Centers for Environmental Prediction (NCEP) develops and delivers operational environmental forecasting systems to its partners within NOAA (e.g., NOS, NHC, etc.) and externally (e.g., US Coast Guard). These operational systems are driven by large earth observing systems that measure various parameters. Observation data sets are received by NCEP Central Operations (NCO) through the Global Telecommunication System (GTS) in BUFR format and from the NOAA/NESDIS operational server (PDA - Production Distribution and Access) in netCDF format and are saved on the dcom server and in individual data tanks based on the source of the measurement (e.g., BATHY, TESAC, SUBPFL, SHIPS, etc.). BUFR decoders are used to read the data from the BUFR file using the pneumonics of certain types of data header file [1].

The marine data observations in BUFR format consist of in-situ temperature and salinity and sea-ice concentration and observations in netCDF format consist of satellite sea surface temperature (SST), sea surface salinity (SSS) and absolute dynamic topography (ADT). These data are then converted into model specific formats to be ingested by each model's data assimilation system. This involves model specific data conversion efforts and data storage problems. Requirements for observation files and I/O handling involved in different modeling and data assimilation workflows are incredibly diverse. Creating a common software system for organizing and storing the vast amounts of observation data is highly desirable to maintain current and future operational forecast systems in a sustainable way.

As part of a modernisation effort of the ocean forecasting systems under the umbrella of the NOAA unified forecast system (UFS) program, a data unification project has been started with the Joint Effort for Data Assimilation Integration (JEDI) to establish a model agnostic method of sharing observation data and exchanging modeling and data assimilation results. The Interface for Observation Data Access (IODA) is a subsystem of JEDI that handles data processing and provides for a common data format in netCDF. This allows for the long-term storage of data and the creation of historical databases. The underlying structure of the IODA format is to represent the variables (e.g., temperature, salinity, etc.) in columns and the locations in rows. Metadata tables are associated with each axis of the data tables and the location metadata hold the values describing each location, and which are appropriate for each observation type (e.g., latitude, longitude). Actual data values are contained in the third dimension of the IODA data table: observation values, observation error, quality control flags, and simulated observation values of H(x) at different stages of the data assimilation process. The python-based IODA converters for all the marine surface and profile observation data types described above have been successfully developed and merged into the JEDI repository.

A 40-year historical IODA-based database of the marine observations were collected for the 1^{o} global reanalysis experiment of the NOAA-NCEP Next Generation Global Ocean Data Assimilation System (NG-GODAS), described in Jong-Kim et al., (2021). The database covers the period 1979-2020. Table 1 describes each parameter, source, datatype and period of data availability. The goal of this project is to create an expandable database of quality controlled observations for 40 years to be used for development, experimentation and reanalysis. These archived observational datasets will be used to develop the necessary observation quality control schemes through the JEDI Unified Forward Operator (UFO), which does the comparison between the observations and the forecast model through H(x) calculation (Honeyager *et al.*, 2020). Figure 1 shows an example of the JEDI-based UFO application result from the AVHRR SST data sets in the NG-GODAS reanalysis experiment.

Datatype	Satellite/sensor	Source	Availability	Data Level
Sea Surface Temperature	NOAA/AVHRR	ESA-CCI ftp://anon-ftp.ceda.ac.uk/neodc/esacci/sst/data/CDR_v2/AVHRR/L3U/v2.	1981-2015	L3U
	NOAA/AVHRR	pathfinder ftp://ftp.nodc.noaa.gov/pub/data.nodc/pathfinder/Version5.3/L3C	1981-2020	L3C
	NOAA/AVHRR	NOAA/NESDIS	2002-2018	L3U
	METOPA/AVHRR	ftp://ftp.star.nesdis.noaa.gov/pub/socd2/coastwatch/sst/ran/avhrr_gac		
	GCOM-W/AMSR2	https://podaac-tools.jpl.nasa.gov/drive/files/allData/ghrsst/data/GDS2/L3U/AMSR2	2004-2020	L3U
	Coriolis/WindSat	https://podaac-tools.jpl.nasa.gov/drive/files/allData/ghrsst/data/GDS2/L3U/WindSat	2004-2020	L3U
	Global Precipitation Measurement (GPM) /GMI	https://podaac-tools.jpl.nasa.gov/drive/files/allData/ghrsst/data/GDS2/L3U/GMI	2014-2020	L3U
	Suomi NPP/VIIRS-NPP		2012-2020	L3U
		ftp://ftp.star.nesdis.noaa.gov/pub/socd2/coastwatch/sst/ran/viirs/npp/l3u	2012-2015; 2018-2020	
		https://podaac-tools.jpl.nasa.gov/drive/files/allData/ghrsst/data/GDS2/L3U/VIIRS_NPP/OSPO/v2.61	2016-2017	
In-situ Temperature & Salinity	ARGO; buoys (drifting & moored); CTD; MBT; XBT; ship	WOD https://data.nodc.noaa.gov/woa/WOD/YEARLY	1979-2020	
Sea Surface Salinity	SMAP	NASA/JPL https://podaac-opendap.jpl.nasa.gov/opendap/hyrax/allData/smap/L2/JPL/V4.3	2015-2020	L2
		https://podaac-opendap.jpl.nasa.gov/opendap/hyrax/allData/smap/L2/RSS/V4/SCI		
	SMOS	ESA https://smos.esa-dias.eo.esa.int	2010-2020	L2
ICE Concentration	DMSP/n07; f08; f11; f13; f17	NSIDC https://www.ncei.noaa.gov/data/sea-ice-concentration/access	1979-2015	L3
	DMSP/SSMR; SSMI; SSMIS (285, 286 GHZ)	L1b BUFR converted to L2 nc files*	2003-2020	L2
ICE freeboard	cryosat	ftp://science-pds.cryosat.esa.int/SIR_GDR	2010-2020	
Absolute dynamic Topography	cryosat-2; Jason-1,2; Altika; Sentinel 3A,3B;	NOAA/NESDIS	1993-2020	
	TOPEX/POSEIDON; Envisat	ftp://ftp.star.nesdis.noaa.gov/pub/sod/lsa/rads/adt		

SMAP: Soil Moisture Active Passive
SMOS: Soil Moisture Ocean Salinity
* NOAA-NEP-EMC C++ converter provided by Robert Grumbine
DMSP: Defense Meteorological Satellite Program
JPL: Jet Propulsion Laboratory
NSIDC: National Snow and Ice Data Center
WOD: World Ocean Database
ESA/CCI: European Space Agency/Climate Change Initiative

Table 1: 40 year marine observation datasets archived in JEDI-IODA format.

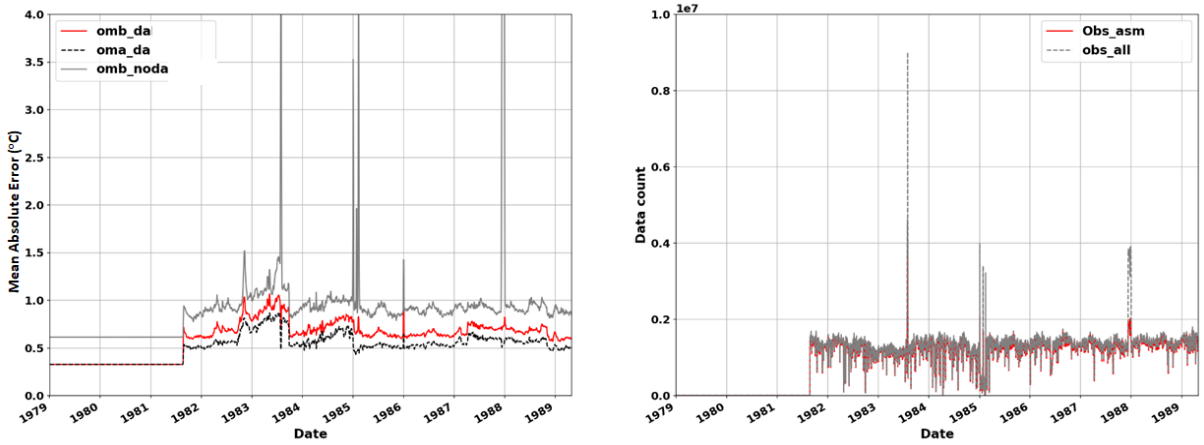

Figure 1. Left panel is time series statistics of the satellite SST observation-analysis (oma) and observation-background (omb). Right panel is data counts for assimilation and no assimilation. 1deg experiment. AVHRR ESA-CCI data sets are applied for the time period 1979~1990.

References

Honeyager, R., Herbener, S., Zhang, X., Shlyaeva, A., and Trémolet, Y., 2020: Observations in the Joint Effort for Data assimilation Integration (JEDI) - UFO and IODA. JCSDA Quarterly, 66, Winter 2020.

Jong Kim, Yi-Cheng Teng, Guillaume Vernieres, Travis Sluka, Shastri Paturi, Yan Hao, Denise Worthen, Bin Li, Jun Wang, JieShun Zhu, Hae-Cheol Kim, Daryl Kleist, NOAA-NCEP Next Generation Global Ocean Data Assimilation System (NG-GODAS), Blue Book, 2021.

[1] https://emc.ncep.noaa.gov/emc/pages/infrastructure/bufrlib.php