Potential Atmospheric Reanalyses for Improving Research Vessel Data Quality Control

Lauryn Fox, Shawn R. Smith

Florida State University – Center of Oceanic-Atmospheric Prediction Studies

Abstract

Having access to long-term, widespread, and accurate weather data is a staple of climate studies. The purpose of the Shipboard Automated Meteorological Oceanographic Systems project is to publish quality-assured weather data over the oceans to provide accurate values to a diverse research community. SAMOS receives data from 44 research vessels on which it runs automated quality control processes, which places flags on unlikely values. The values that are a standard of plausibility come from a climatology, which comes from a historic atmospheric analysis. The current one used for the SAMOS project by Da Silva, only uses in situ measurements and is becoming outdated compared to current technology, so a new reanalysis could be a better asset to the SAMOS quality control process. Each atmospheric reanalysis has its benefits and limitations for the sake of the project, so each must be tested and compared. The ERA5 reanalysis seems to be a promising option that is worth comparing the results of a climatology test between it and the existing Da Silva climatology. By assigning flags to the implausible values when comparing a year of SAMOS observations to the ERA5 climatology, and examining the results of the same year with the current SAMOS climatology, the two can be evaluated and compared.

Background

Literature review on the documentation for the several reanalysis products reveals the differences between them and assures whether the product provides the attributes that are necessary for SAMOS. See chart 1.

	CFSR [5]	JRA-55 [6]	ERA5 [2]	CERA-20C [8]	20CR [7]	Da Silva
Spatial	0.25°x0.25° near equator 0.5°x0.5° near poles	T319L60 Spectral ~55km Top layer 0.1hPa	0.25°x0.25° (atm) 0.5°x0.5° (ocean waves)	125km	1°x1°	1°x1°
Time range	1979-2017	1957-2022	1959-present	1900-2010	1836-2015	1945-1989
Data Format	netCDF	netCDF	netCDF	netCDF	netCDF	netCDF
Frequency	Sub-daily Monthly	Sub-daily Monthly	12-hour	3 hours Monthly means	3-hourly Daily Monthly	Monthly
SAMOS Variables						All SAMOS concerning variables
Strength	-Accounts for changes in trace gases (CO_2) -Assimilates satellite radiances	-Variational bias correction on satellite data -Option to compare assimilated data to non-assimilated data	Has strong temporal consistency for times that comes from the assimilation of satellite data.	Provides a yearly report detailing the quality and the errors of the system. Could maybe be used as a reference.	Has the most historical data	
Weakness	-Not well evaluated - The reanalysis can be "uncoupled" due to ocean-atmosphere interactions being noted but not accounted for	-Warm bias in the upper troposphere varying with time -Dry bias in the upper and middle troposphere	There can be biases and unrealistic predictions in areas where data is sparse.	Time range for data stops earlier than others.	-Warm bias in cold months -Sea ice issues where it is less observed	Does not assimilate satellite data and is a mean of each variable rather than an assimilation of how the variables may affect eachother.

Chart 1: Comparison between reanalyses products

Introduction

An atmospheric reanalysis is an assimilation of historic observational data that spans across a period of time. The ERA 5 reanalysis' spatial distribution is a grid system of 0.25° latitude x 0.25° longitude over land, and 0.5°x0.5° over the ocean, and these grids separate data to keep values relevant to location; when a SAMOS vessel's coordinates fall within one grid square, that is the square that will be referred to for quality control.

The ERA 5 assimilates historic data starting from 1959 with the modern observational data, which is both in situ and satellite, and publishes an accurate monthly mean from the hourly reanalysis available 2-3 months after each month, which is used for the climatology. The climatology sets the standard for likely values in one grid square, for the sake of the SAMOS project, a value must be within four standard deviations of the monthly climatology to be accepted, otherwise, it is flagged. What makes this reanalysis an eligible candidate, is its publication of accurate monthly means, and a large list of variables including most SAMOS concerning variables, and ones that could be a beneficial addition to the current SAMOS variable list. The outlook for improvement of the QC test from the Da Silva climatology is the finer grid squares and the assimilation of satellite data that are a part of the ERA 5 product.

Methods

- *To begin programming the comparisons, the netCDF files for SAMOS ships are downloaded from SAMOS (), and the ERA5 reanalysis dataset is downloaded from copernicus().
- *The SAMOS ship used for this display is the Roger Revelle, with daily files for days 01/01/2023-01/07/2023, within a coordinate range surrounding southwest South America.
- These ship files will be used for the following procedures:

- Air temperature files will be appended to an array in order of changing position of the ship
- The latitudes and longitudes will be rounded down to the nearest 0.25 and appended to their own arrays in order of changing position of the ship
- Using the January time step from the ERA5 product, each latitude and longitude will be matched to their most accurate grid cell in the ERA5 and append the matching air temperature for that coordinate cell.
- Produce a mean and a standard deviation for this range of reanalysis temperatures.
- Use these statistical values as a range for the ship's recorded air temperature values
- Compare these results with the gathered results of the operational SAMOS processing using the Da Silva climatology

Discussion

- An initial observation of the SST imitation QC test is that the ERA5 will yield more much more flags
- (Go through and write the number of flags for each)
- Each of the time series plots suggests there is a smaller accepter range of values from the ERA5 QC standards than those of the Da Silva, leaning especially toward lower temperatures. This is consistent with the histogram in figure() that accepts about only a quarter of the lower end temperatures that the Roger Revelle actually measured during this week.
- These higher temperatures could possibly be explained by the positive SST anomaly seen in figure(). Data such as this would get assimilated into the ERA5 model which could provide a benefit to the accuracy of future SST acceptable ranges.
- Although sea surface temperatures are usually only affected in equatorial pacific during a La Niña event, there is a possibility that this year's La Niña could have an impact on the recorded values.
- The monthly climatology created has suggested to not have enough variability. The standard deviation between means of Januarys over decades is quite small, so this could explain why the accepted range of 4 standard deviations from the ERA5 is so small. This does not necessarily make it more accurate than the Da Silva, possibly just not specific enough.
- Another possible contributor to the increased number of flags is the more refined grid spacing of the ERA5. As each temperature is compared to a more specific location, there is more room for error.

DAY	JAN01	JAN02	JAN03	JAN04	JAN05	JAN06
SAMOS FLAG COUNT	163	0	984	1440	504	288
ERA5 FLAG COUNT	782	1439	1271	1440	1440	1440

Chart 2: Comparison between out of range flagged values for SAMOS and ERA5 QC

```
OutOfRangeIndex=[]
OutOfRangeTemps=[]
AcceptableTemps=[]
AcceptableIndex=[]
count=0
range=stdvev(ERA.temperature.array)

for x, y in zip(RRtemperatures4, ERAtemperatures4):
    diff = x - y
    if diff > abs(range) or diff2 < -abs(range):
        OutOfRangeTemps.append(x)
        count+=1
        OutOfRangeIndex.append(Index)
    else:
        AcceptableIndex.append(Index)
        AcceptableTemps.append(x)
```

Fig1: Example of code that creates the arrays for plotting flagged values

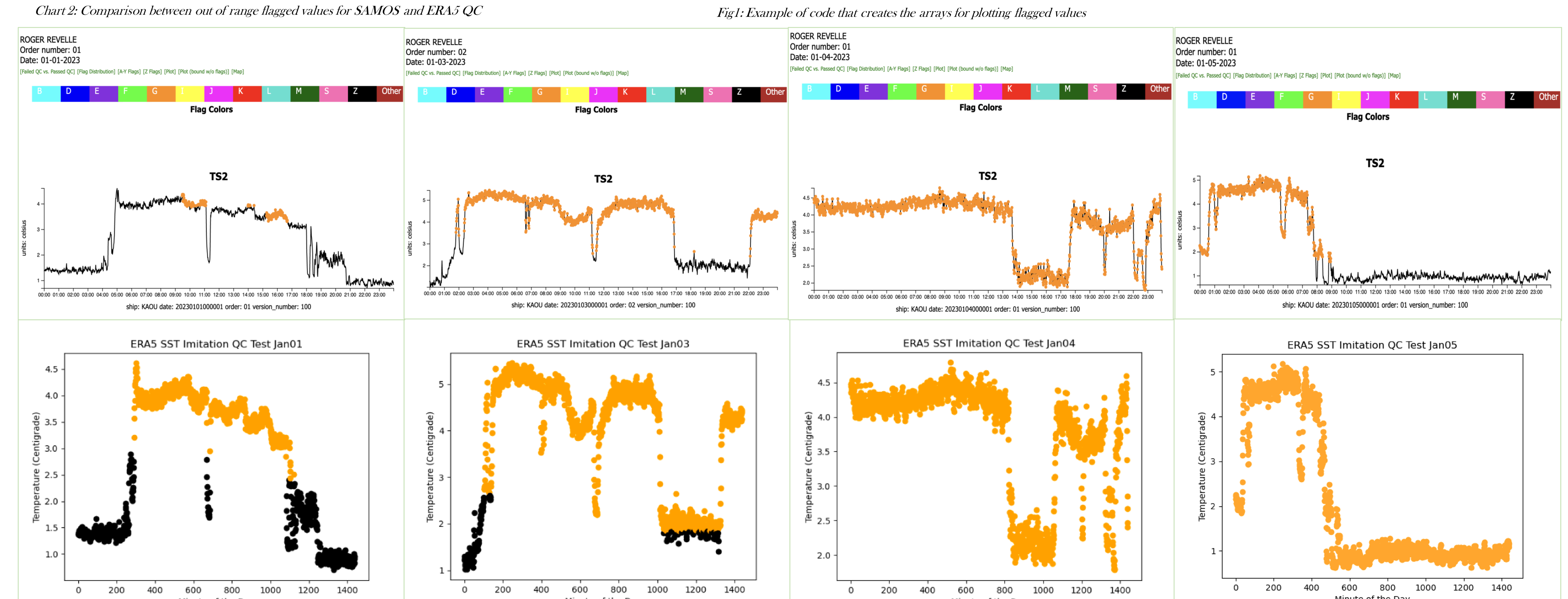

Fig3 : SAMOS Da Silva QC summary January 1, 3, 4, 5 Sea Surface Temperature compared with the ERA5 imitation QC results https://samosdev.coaps.fsu.edu/charts/index.php?ship=KAOU&id=73&date=20230103000001&order=02&history_id=681000&mode=6&fbound=1

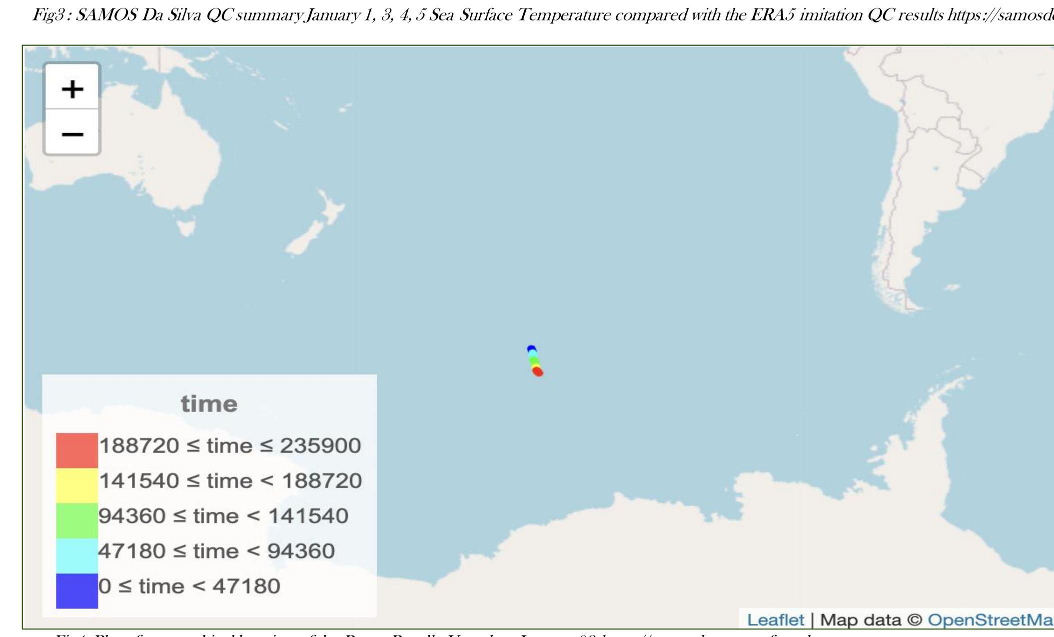

Fig4: Plot of geographical location of the Roger Revelle Vessel on January 03 https://samosdev.coaps.fsu.edu...

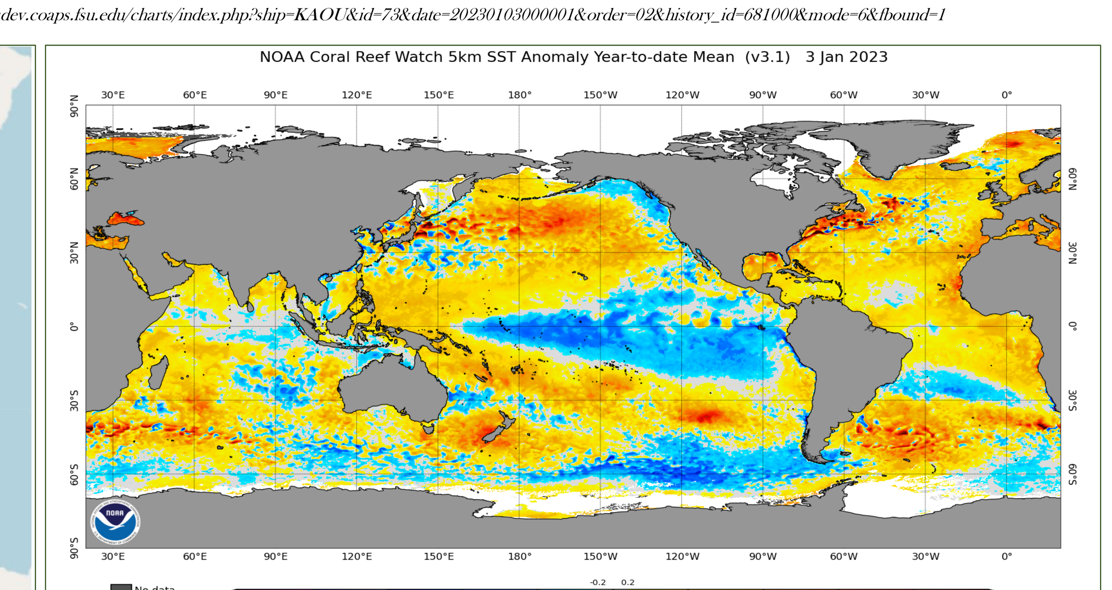

Fig5: January 3rd Sea Surface Temperature anomalies from: https://coralreefwatch.noaa.gov/product/5km/

Future Directions

- Based on the results yielded, the ERA5 has promising potential for providing more data accuracy, but more testing must be performed before adopting it to the SAMOS QC process.
- Another imitation code could be written with more time refinement from the monthly climatology to provide more variability. Perhaps a 10-day or weekly type of climatology.
- Each of the SAMOS concerning variables should also be tested to identify any biases the ERA5 model tends to have.
- A similar process should be repeated for each of the reanalysis products to determine which one has the least biases and is the most compatible product for the goals of SAMOS.

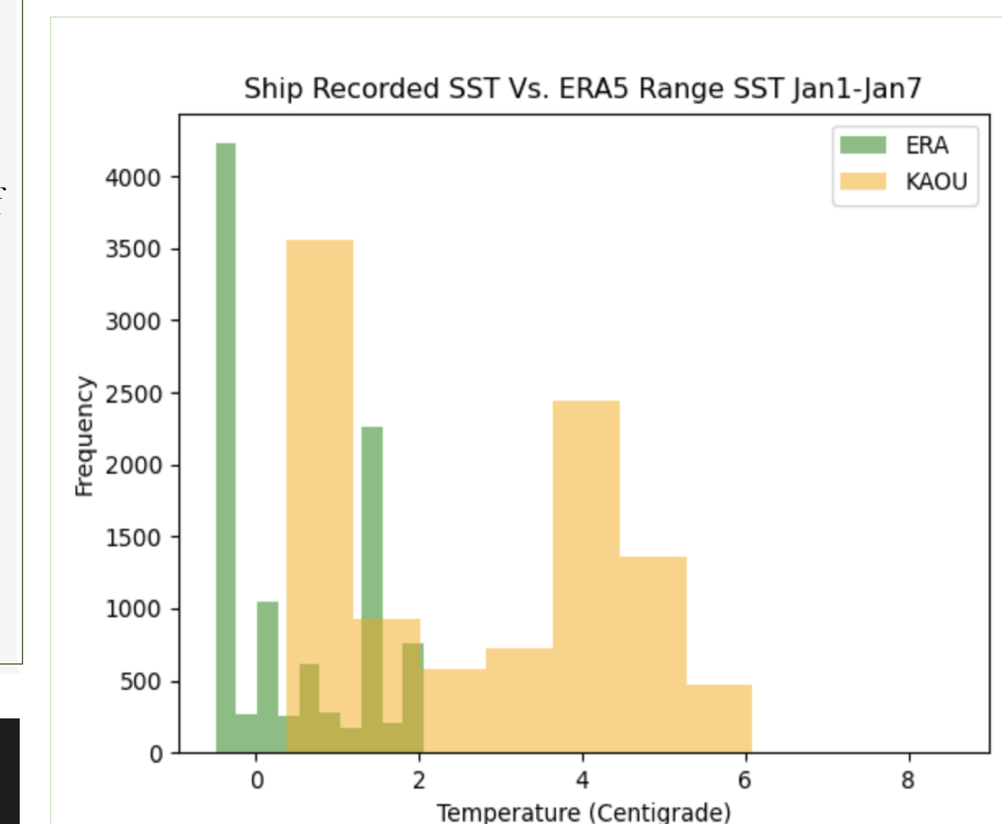

Fig2: Histogram of ship recorded SST Centigrade ERA5 mean in corresponding location

Acknowledgements

I would like to thank Shawn Smith for giving me the opportunity to be a part of the SAMOS research and guiding me through using atmospheric reanalyses.

Thank you to Marc Castells for creating the monthly mean climatology from the ERA5 dataset and helping me to write much of the code that made this possible.

References

[1] DaSilva, A.M., C.C. Young, And S. Levitus. *Atlas Of Surface Marine Data, Vol. 1: Algorithms And Procedures*. NOAA Atlas NESDIS 6, U. S. Department Of Commerce, NOAA, 1994. Print.

[2] Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2018): ERA5 hourly data on single levels from 1959 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). 10.24381/cds.adbb2d47

[3] Lindsay, R., Wensnahan, M., Schweiger, A., & Zhang, J. (2014). Evaluation of seven different atmospheric reanalysis products in the Arctic*. *Journal of Climate*, *27*(7), 2588–2606. doi:10.1175/jcli-d-13-00014.1

[4] Smith SR, Briggs K, Bourassa MA, Elya J, Paver CR. Shipboard automated meteorological and oceanographic system data archive: 2005–2017. Geosci Data J. 2018;5:73– 86.

[5] Climate Data Guide. Climate Forecast System Reanalysis (CFSR) | Climate Data Guide. (1979, January 1). Retrieved February 27, 2023, from https://climatedataguide.ucar.edu/climate-data/climate-forecast-system-reanalysis-cfsr

[6] Climate Data Guide. JRA-55 | Climate Data Guide. (1957, December 1). Retrieved February 27, 2023, from https://climatedataguide.ucar.edu/climate-data/jra-55

[7] Team, P. S. L. W. (n.d.). *The twentieth century reanalysis project*. PSL. Retrieved February 27, 2023, from https://psl.noaa.gov/data/20thC_Rean/#:~:text=20CR%20uses%20an%20ensemble%20filter,estimates%20uncertainty%20in%20that%20analysis.

[8] Laloyaux, P. (2022, August 16). *Cera-20C*. ECMWF. Retrieved February 27, 2023, from https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/cera-20c