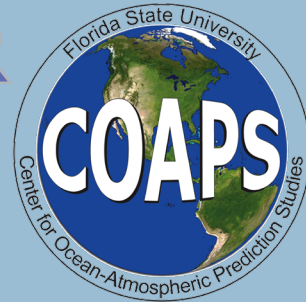


# MARINE DATA CENTER

## FLORIDA STATE UNIVERSITY

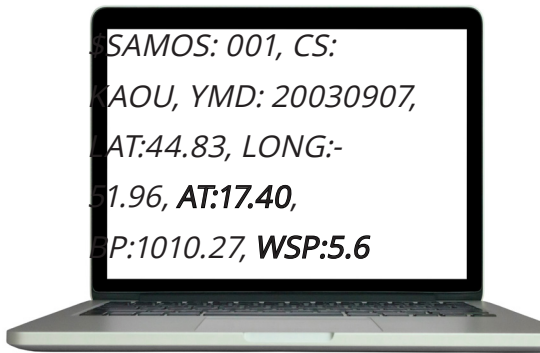
CREATED BY DANIEL SIEGEL, COAPS STUDENT  
COMMUNICATION ASSISTANT  
DECEMBER 2023



# WHY DO WE NEED METADATA?

## WHAT IS METADATA?

Metadata is essentially **data about data**. Take a look at this piece of SAMOS data. Can you understand what it means?



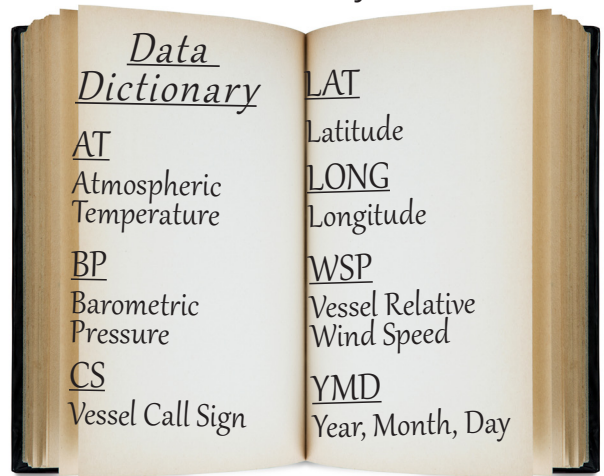
As it stands right now, this data isn't **usable**. Oceanographic data isn't useful to a researcher unless they know **what** it is, **where** and **when** it was gathered, **who** gathered it, and **how** they did so. That's where metadata comes in!



## DATA DICTIONARY

Metadata works like a **"data dictionary."** Without it, data is just a bunch of letters and numbers. Metadata tells a researcher what a data point is and what it means: What did it measure? What unit is it in? How was this measurement gathered?

When a researcher downloads a SAMOS data file, they don't just get the string of letters and numbers on the left. They also get metadata that explains what the data means! That crazy string of letters and numbers might make more sense if you know what "AT" and "WSP" mean. Our metadata helps users read our data—they can always consult their dictionary!



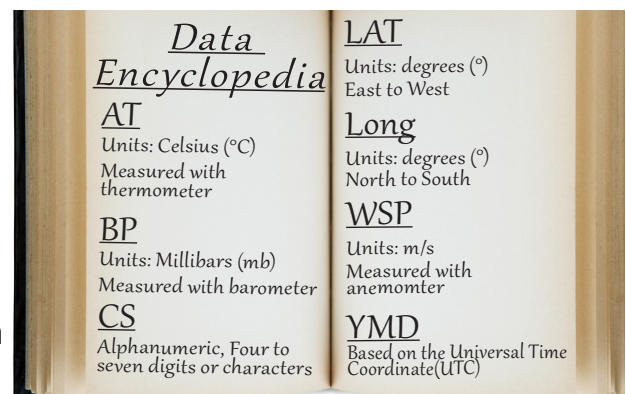
## DATA ENCYCLOPEDIA

A dictionary is useful, but sometimes you need more than just a definition. Is the air temperature 17.40 degrees Fahrenheit, or is it 17.40 degrees Celsius? Is the wind speed 5.6 miles per hour, or 5.6 meters per second? These and other questions can be answered by metadata.

Metadata is more than just a data dictionary. It's a **data encyclopedia!**

SAMOS metadata doesn't just define "WSP" as "vessel relative wind speed." It can also tell you how this speed was calculated and how it is being reported.

If a researcher wasn't sure whether a ship was reporting its course in degrees clockwise towards true north or counterclockwise, they could end up think the ship is going in a completely different direction than it really is!



# DATA INSTRUCTION MANUAL

Metadata doesn't just help researchers understand what data points mean. More than just a dictionary and encyclopedia, metadata can work as a data instruction manual. Metadata tell researchers how to use data.

Metadata might tell a researcher that a given measurement has a 0.3% margin of error. For some statistical analyses, they might not require any more precision than this—this data might be usable. But if they need results which are more precise than this, they should know to use a different data source.

Additionally, some data measurements can be collected in multiple ways or from multiple places on a ship. Knowing where an instrument is or how it works can change how a scientist uses its data, as they might have to adjust for certain factors to ensure accuracy.

# WHY DO WE NEED METADATA?

Created by: Daniel Siegel

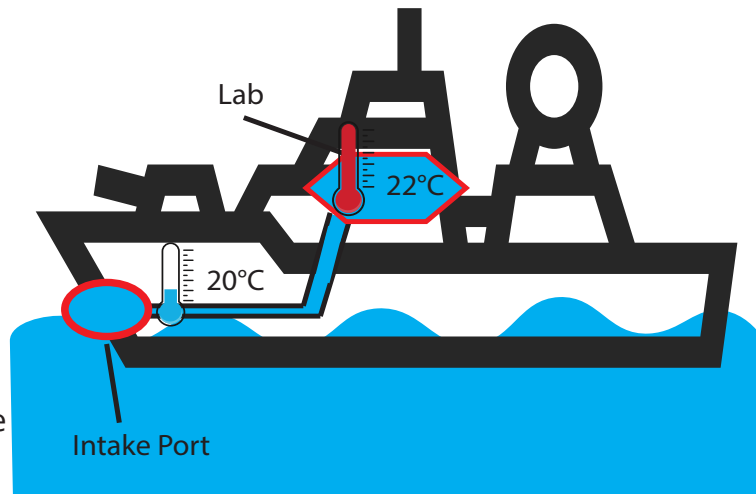


## INSTRUMENT PLACEMENT

### Thermometers (Temperature measurements)

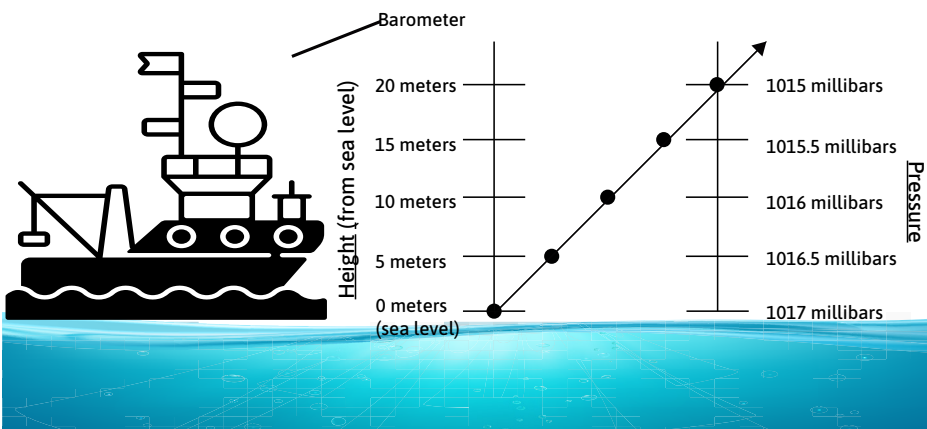
Some ships gather sea water temperature with a thermometer directly placed in the water. This measurement will be very similar to the temperature of the water in the area the vessel is sailing.

Sometimes, though, ships pull water from the ocean inside to a lab before taking temperature measurements. This measurement might be affected by the heating or cooling systems in the ship. If a ship is in the Arctic, the temperature of water in the ocean outside will likely be colder than water that has traveled through the inside of the vessel, whose engine and heating systems will warm the water up as it travels to the lab.



## DATA ADJUSTMENTS

### Barometers (Pressure measurements)



Pressure is sometimes adjusted to sea level. Other times it isn't. Knowing whether pressure is or isn't adjusted is important, because pressure decreases each meter the barometer (pressure sensor) is above sea level. These small inconsistencies can potentially impact the predictions a weather or climate model makes. Metadata can tell a user whether a pressure measurement is or is not adjusted to sea level. Metadata can also indicate how far above sea level the barometer was located so that researchers can adjust for these differences themselves.

# DATA AUTOCORRECT AND DATA PROOFREADERS



# WHY DO WE NEED METADATA?

Created by: Daniel Siegel

At SAMOS, we don't only present metadata already provided by the people or instruments which measured it. Our computerized Quality Control (QC) flagging system adds metadata to the data, helping users select the data they need and use it properly. We create new metadata in the form of QC flags and attach it to recorded data to ensure they are as specific, precise, and usable as possible. This QC system works as a **data autocorrect**.



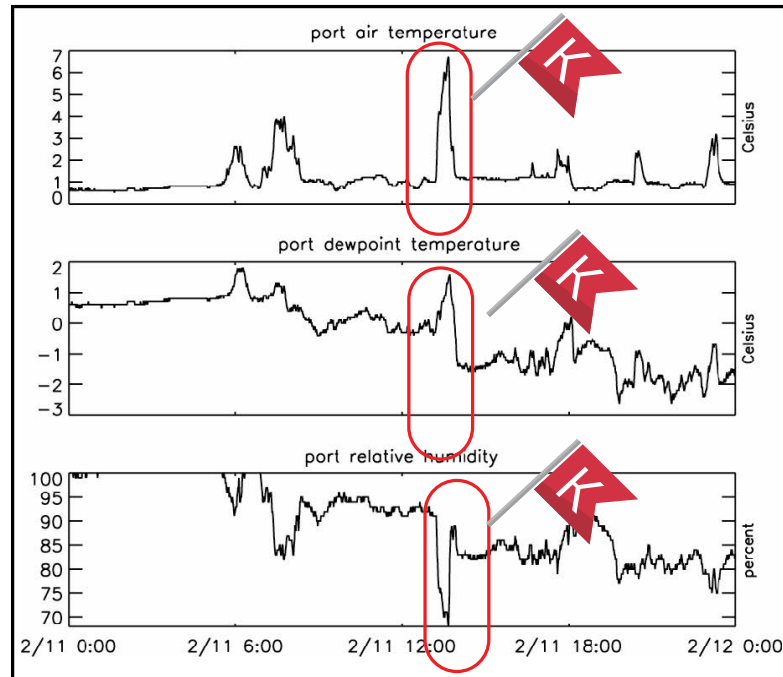
We don't stop at this automatic system. We also have a team of data analysts who review data flagged by the computer system—these are **data proofreaders**. Our computer system notices when data points stick out—our analysts will add or modify flags for each data value, using metadata to warn a user that something might be wrong.

## SUSPECT DATA



In the graphs to the right, there are big spikes in "air temperature" and "dewpoint temperature," and a big dip in "port relative humidity."

These kinds of spikes and dips often happen when sensors get contaminated by a ship's exhaust. Metadata can warn a user that this data is likely inaccurate. SAMOS QC systems would attach a "K" flag to these spikes and dips, telling a user "Suspect data: use with caution."



## INTERESTING FEATURE METADATA

Not all flags are bad! SAMOS also flags data which is correct and valid but has been influenced by events such as hurricanes, typhoons, convective events, or frontal passages. Our data quality proofreaders will add an "I" flag to these extreme data points, along with a note explaining the event that is being reflected.

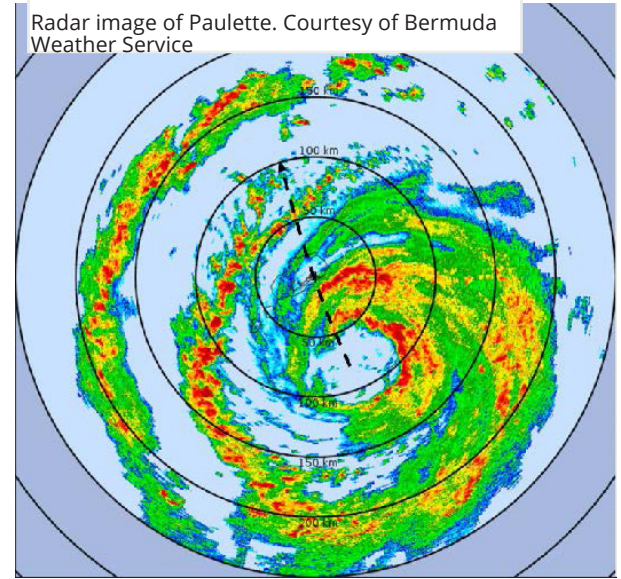
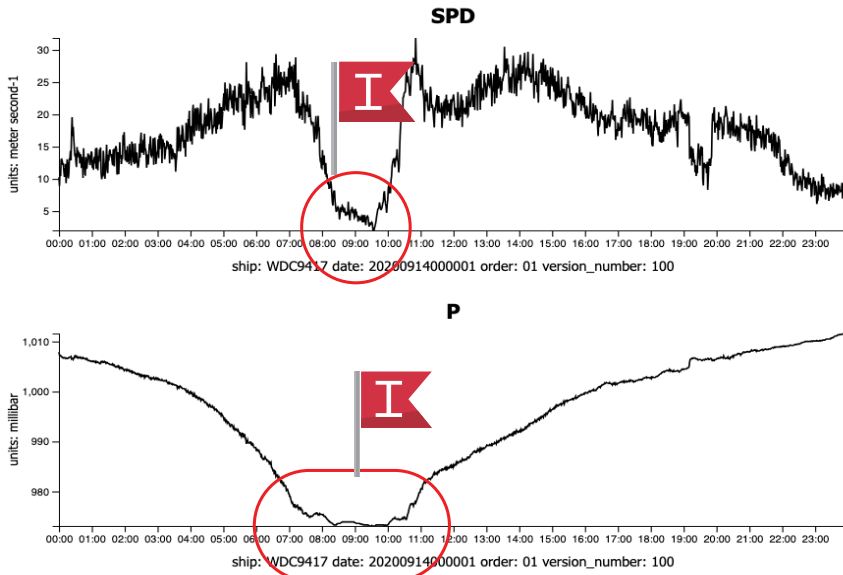
Often it is important to study extreme or unusual events. "I" flags help researchers pinpoint these events in our data sets and differentiate them from data spikes caused by instrument malfunctions.

These points are also extreme. While they may be useful, they do not reflect typical meteorological or oceanographic patterns. Often they need to be excluded from datasets or statistical analyses. Once again, metadata helps our users understand the data they're looking at and figure out how to use it!

# OTHER EXAMPLES

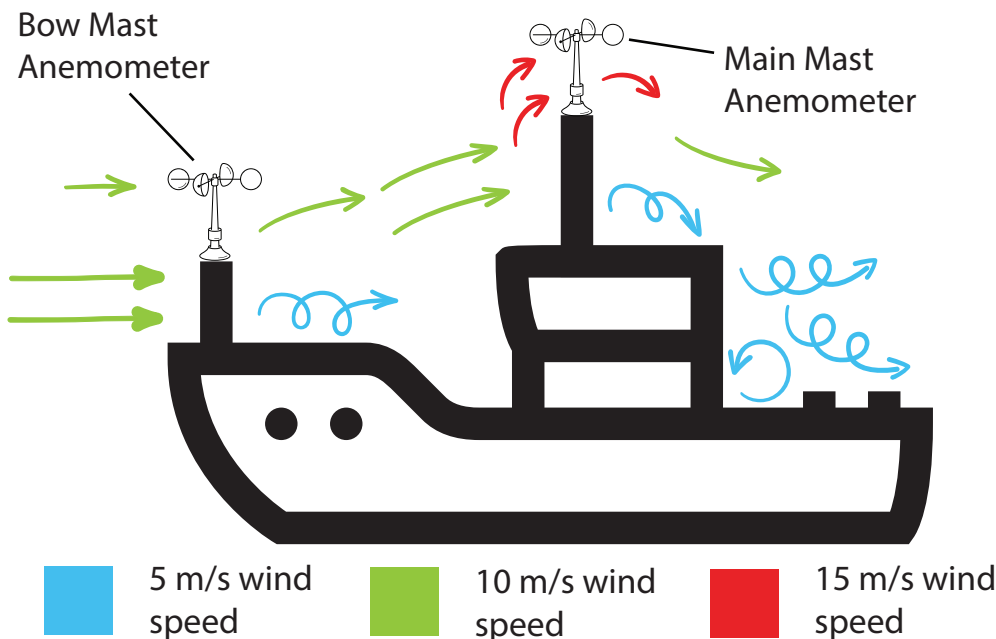
## STORM EVENTS

Hurricane Paulette passed over the island of Bermuda in 2020, with the low pressure and high winds near the eye passing over the RV *Atlantic Explorer*. The atmospheric pressure and wind speed measurements in the graph below are unusual, but they are not incorrect.



### Anemometers (Wind Speed measurements)

Wind speed measurements can also change depending on where a sensor is placed on a vessel. Air flow around a vessel can be distorted by the ship's structure. On some vessels, an instrument on a bow mast might measure a wind speed of 10 m/s whereas an instrument on a main mast would measure a speed of 15 m/s. You can see why a researcher might need to know where on a ship a measurement was taken!



### Averages

If a value represents an average, it's important to know how that was calculated. A mean can be very different from a median. Knowing the Time Center of this average is important too. Even when reporting an average value, instruments will assign a time for when the value was gathered. Users need to know if it is reporting the end of the period being averaged, the beginning, or the middle. Knowing the number of values which make up the average is also important. The more values provided, the more confidence a user will have in an averaged value.



This work was supported by NOAA's Global Ocean Monitoring and Observing Program, a subaward (191001.361472.01B) from the Northern Gulf of Mexico Cooperative Institute grant NA21OAR4320190 administered by Mississippi State University.

The COAPS Marine Data Center dispenses a wide range of products used to explore the atmosphere-ocean interface. For more information, reach out to: [mdc@coaps.fsu.edu](mailto:mdc@coaps.fsu.edu).