NGI Automated Quality Control/Assurance

Jeremy J. Rolph, Shawn R. Smith, and Mark A. Bourassa

Center for Ocean-Atmospheric Prediction Studies
The Florida State University, Tallahassee, FL 32306-2840, USA
smith@coaps.fsu.edu

http://samos.coaps.fsu.edu/

1. Introduction

Since 1994, the Center for Ocean-Atmospheric Prediction Studies has been a leader in the development and implementation of quality-control procedures for automated marine meteorological observations. Our current focus is on underway observations collected on oceanographic research vessels, although we have previous experience with moorings, drifting buoys, and fixed island stations. Through the Shipboard Automated Meteorological and Oceanographic System (SAMOS) initiative, we have developed a fully automated quality control system. This system ingests data logged by the shipboard measurement system, formats the data, augments the files with critical metadata, and then pushes the data through multiple automated QC checks. Additional visual inspection is used to identify problems in the shipboard data, but we will focus herein on the automated QC.

In 2005, the SAMOS data center agreed to provide data management services for meteorological, air-sea flux, and oceanographic data to be collected on Air Force towers in the Gulf of Mexico. We plan to apply the automated SAMOS data quality procedures to the meteorological observations from the towers. Oceanographic and flux data will also be incorporated in the automated QC in the future, although the data will initially be available in the data stream. Quality controlled data will be made available to the public in near-real time.

2. Data Flow

Meteorological Data Flow

- K-Tower is located at 29°39.97' N, 84°22.19' W or 16.8 miles from the FSU Marine Lab (Figure 7)
- One minute samples transmitted every 10 minutes via line of sight RF to FSU Marine Lab
- Data pulled from Marine Lab to COAPS data center
- Automated data formatting and QC done at COAPS
- QC'd data available via COAPS web server (on the hour)

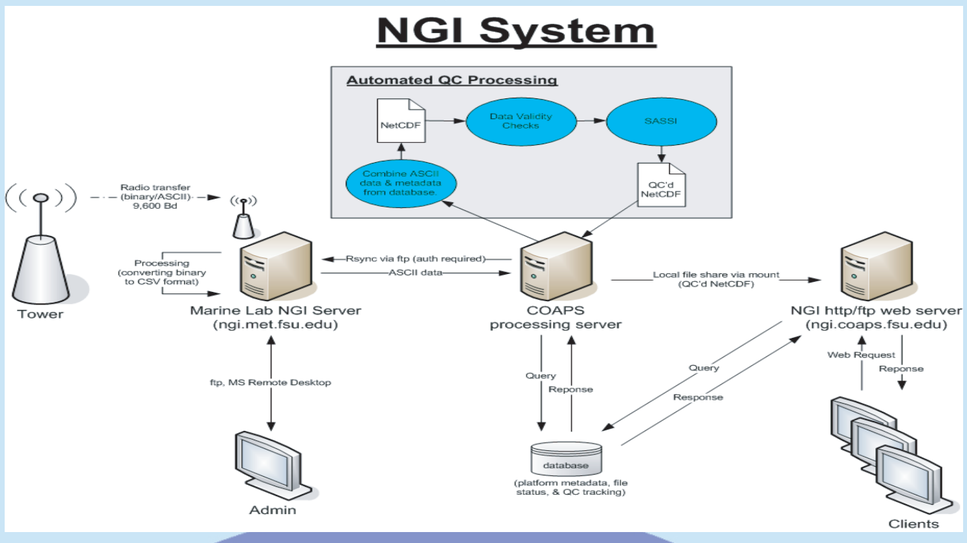

Fig. 1. Schematic of NGI data flow

3. Automated Quality Control System

Philosophy is to flag (not remove) suspect data at the parametric level.
Pre-process checks for realistic ranges, time sequence, etc.
New statistical spike/step flagging tool (SASSI)
Order of precedence for preprocessing flags

Prescreener

- time sequencing/duplicates
- realistic range check (per parameter)
- latitude dependent climatology check (per parameter) using da Silva, et. al, 1994 (FIgure 2)
- T ≥ Tw ≥ Td relationship check (Figure 2)
- true wind recomputation test
- platform position/speed* (Figure 3)
- land check* (Figure 3)
  - * - moving platforms only

SASSI

- Spike and Stair Step Indicator (Figure 4)
- requires minimum number of points to calculate differences (~3000 or about 2 days of data)
- calculates statistics in moving window for given parameters in 2 identical passes
  - -one pass for suspect observations (differences of 1 time step)
  - -second pass to determine spike vs. step using rules (differences of 2 time steps)
- statistics centered on day of observations being processed
- utilizes specified parameter-specific standard deviation factor and minimum threshold values
- statistics utilized in determining spikes and steps

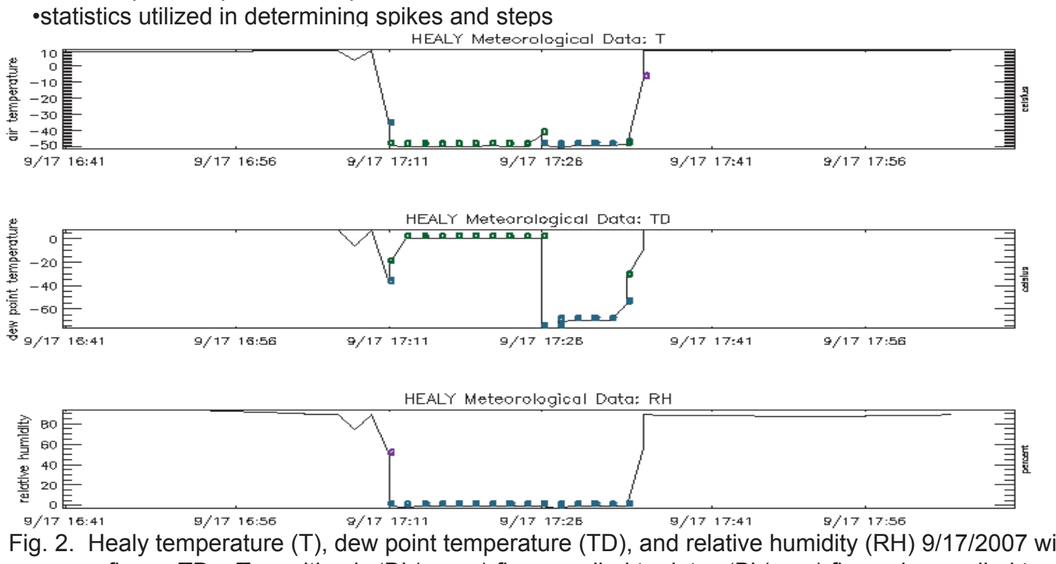

Fig. 2. Healy temperature (T), dew point temperature (TD), and relative humidity (RH) 9/17/2007 with prescreener flags. TD ≥ T resulting in 'D' (green) flags applied to data. 'B' (grey) flags also applied to all variables for data out of realistic bounds, i.e., relative humidity has negative values. T and RH also have values > 4 std. dev. from climatology resulting in 'G' flags (purple).

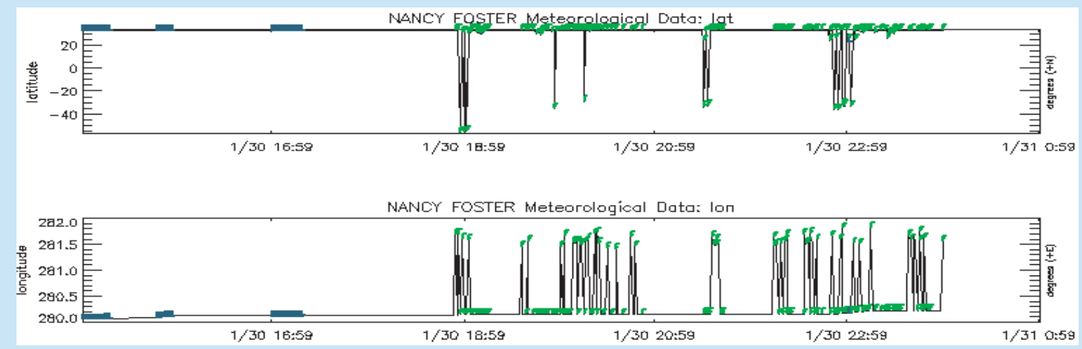

Fig. 3. Latitude and longitude data from the Nancy Foster with 'L' (grey) and 'F' (green) flags applied by the prescreener for values over land and unrealistic platform movement 1/30/2008.

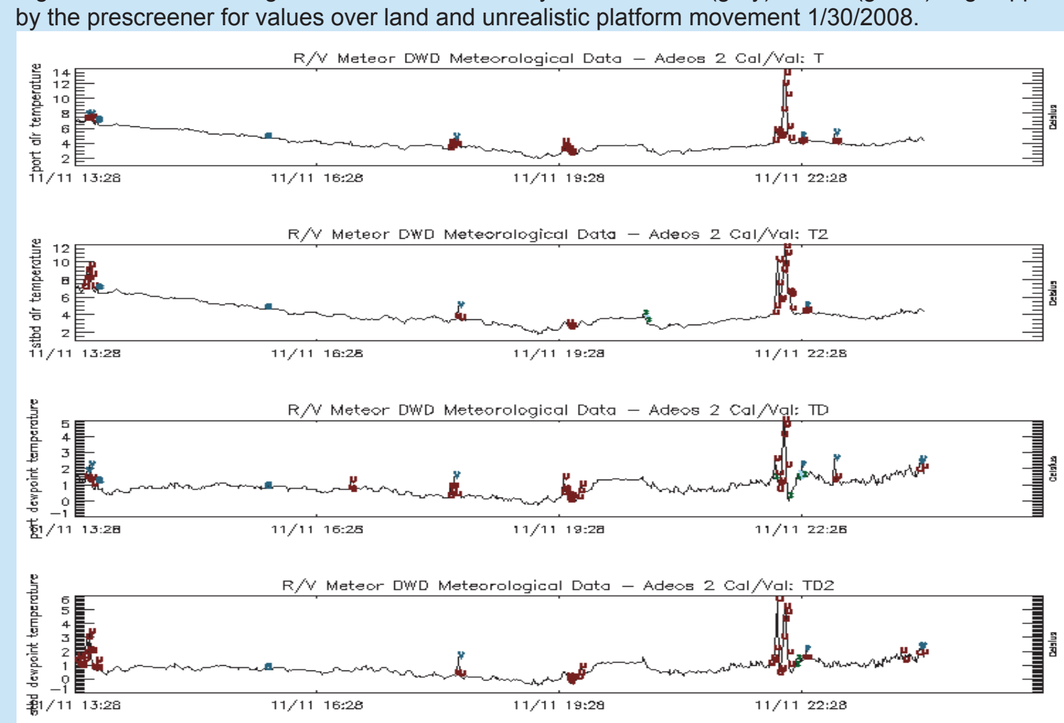

Fig. 4. Time series of port (T, TD) and starboard (T2, TD2) temperature and dew point temperature from the R/V Meteor after processing through SASSI. Flags (U, V) added to data indicating suspect (U - brown) and spike (V - grey) data. Data contaminated by stack exhaust.

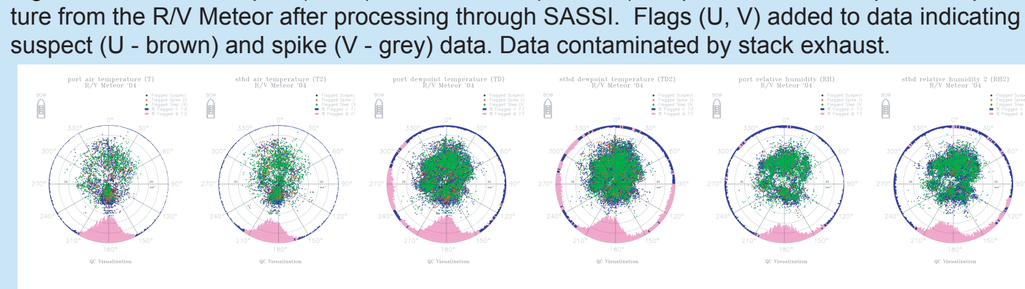

Fig. 5. Port and starboard air temperature, dew point temperature and relative humidity data from R/V Meteor after SASSI processing. SASSI flags plotted relative to the platform-relative wind direction and speed. Results indicate flow from stern of vessel has negative impact on parameters given as reflected in time series Fig. 4.

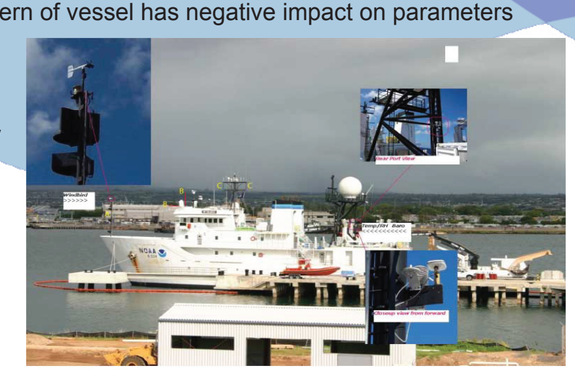

Fig. 6. Sensor placement and knowledge of sensor placement is critical in collecting quality data. Photo is of NOAA R/V Hi'lalakai and insets are sensors. Temperature and relative humidity data contamination take place due to sensor locations near stack, attached to black structure when the platform-relative winds are from the starboard-bow (±45°).

4. Summary

Automated QC procedures enable the capability of identifying data issues such as flow distortion and heating problems.

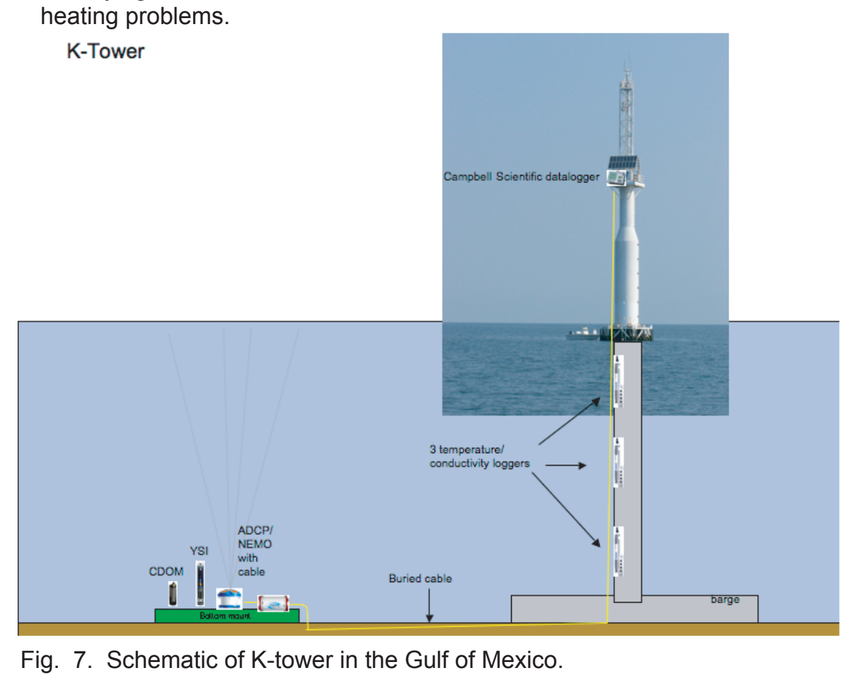

Fig. 7. Schematic of K-tower in the Gulf of Mexico.

Acknowledgements

The NGI data center thanks NOAA and the Northern Gulf of Mexico Cooperative Institute for their funding and support.