9.2 The NOAA PROducts Validation System (NPROVS) and Archive Summary System (NARCSS)

Tony Reale*
NOAA/NESDIS, Washington DC

Bomin Sun, Michael Pettey and Frank Tilley
IM Systems Group, Rockville, MD

1. INTRODUCTION

During the past 30 years of NOAA operational polar satellites, the problem of providing reliable and consistent monitoring and scientific validation of operational measurements and derived satellite soundings has been addressed through the compilation and analysis of collocated satellite and radiosonde observation datasets. The NOAA PROducts Validation System (NPROVS) (Pettey *et al.* 2009), recently deployed at the Center for Satellite Applications and Research (STAR), centralized the routine compilation of satellite and radiosonde collocation datasets among the multiple satellite derived sounding product systems operated by NOAA, including respective observation screening. These datasets have also proved useful for characterizing respective platform performance and for computing coefficients in support (tuning) of derived satellite sounding algorithms (Reale and Tilley 2009)

The following report presents an outline of NPROVS and results demonstrating strategies for:

- satellite sounding validation
- screening
- platform error characterization

Results are generated using the Environmental Data Graphical Evaluation (EDGE) interface which includes basic utilities for:

- display of collocation global distributions
- profile display and statistical analysis
- orbital product display

Corresponding author address:
Tony Reale
NOAA/NESDIS/STAR, Suitland, MD
email: tony.reale@noaa.gov

The report concludes with an example of longer term trend plots for the recently developed NPROVS ARChive Summary System (NARCSS).

2. NPROVS

Satellite derived sounding products are routinely produced by NOAA for a number of satellite platforms including GOES, NOAA-18, MetOp, NASA–EOS-AIRS and DMSP and a number of processing approaches including operational Advanced TIROS Operational Vertical Sounder (ATOVS) (Reale et al. 2008), Microwave Integrated Retrieval System (MIRS) (Boukabara *et al.* 2007) and hyper-spectral sounder approaches for AIRS and MetOp-IASI (Goldberg *et al.* 2003). Although not currently used at most of the NWP centers, derived soundings remain a mainstay of NOAA ground processing systems and may yet play a key role as an efficient data compression mechanism for assimilating hyper-spectral observations and in climate.

Figure-1 shows a schematic diagram of NPROVS and multiple satellite platforms and processing suites, including NWP, that are routinely collocated with the ground-truth (mainly radiosondes) observations. Collocations are processed daily and archived daily. NPOESS EDR Proxy indicates pending collaborations with NOAAA Integrated program Office Government Resource for Algorithm Verification, Independent Testing and Evaluation (GRAVITE) protocols in conjunction with NPOESS.

Figures 2 and 3 show examples of the global distribution of radiosondes (2) and an individual set of collocated radiosonde and satellite locations in the vicinity of Wollops Island, Va., as compiled by NPROVS during January 2009.

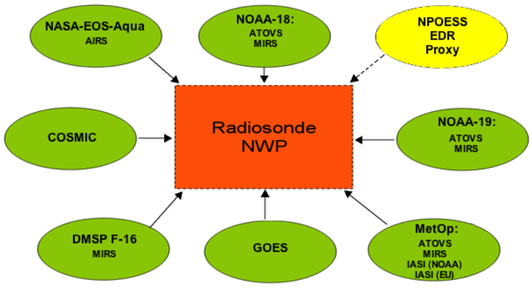

Figure 1: Diagram of current NPROVS satellite data (green) access, future data (yellow) and collocation with ground truth (red).

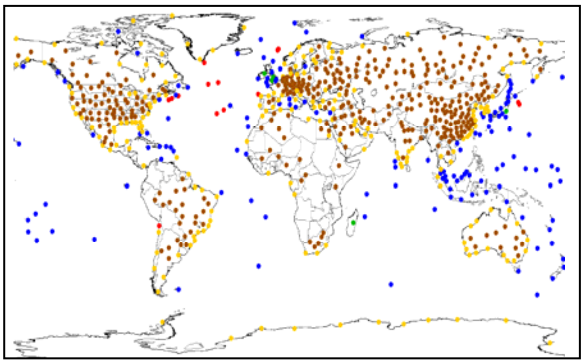

Figure 2: Global location of radiosondes collocated with at least one satellite observation platform for a 2-day period during January 2009; colors indicate the terrain flag of the radiosonde (red, ship; brown, land; yellow, coast; blue, island and green, inland island)

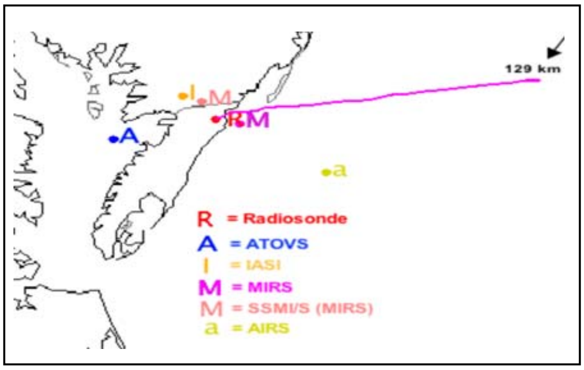

Figure 3: Example of individual set of collocated radiosonde (red) and respective satellite products (other colors) and the associated drift (pink) of the radiosonde during flight in vicinity of Ocean City, Md. on January 2, 2009

Approximately 1000 collocations (a radiosonde with at least one collocated satellite) are processed daily. The criteria for a candidate collocation are:

- radiosonde temperature and moisture profile extend at least 5 km without gaps
- satellite within 6 hours and 250 km
- single "closest" satellite retained

It is interesting to note the spatial drift of the radiosonde (pink) easily exceeds the spatial domain of the collocated observations. Conventional collocation datasets are compiled using the location and time of the radiosonde at the surface. Available drift parameters (radiosonde and satellites) are retained within NPROVS; their impact on validation is discussed in Section 3.2.

The compilation of radiosondes includes specialized testing of the radiosondes extending at least 5 km. Tests include:

- superadiabatic layer(s)
- tropopause within limits
- supersaturated level(s)
- moisture profile score
- temperature inversion(s)

Flags indicating one or more of the above occurrences are retained on the output radiosonde file. Of particular interest are tests for H_20 vapor changes and subsequent impacts on validation.

Figure 4 summarizes the moisture testing and associated results.

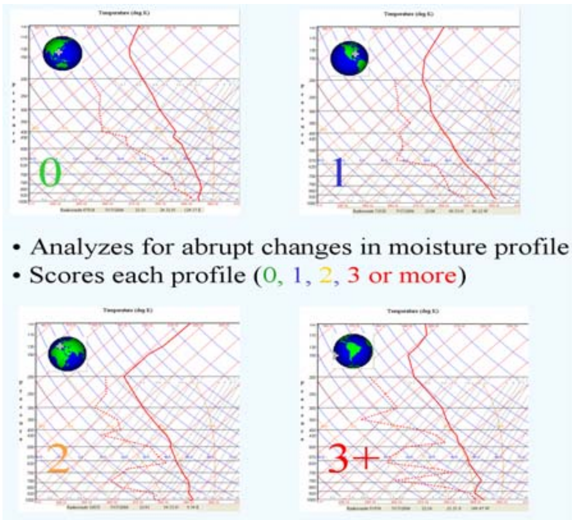

Figure 4: Raob moisture profile (dashed) scores

These results represent the end product of a series of tests which analyze the degree of deviation from a monotonically decreasing H_20 vapor mixing ratio profile. Moisture profiles exhibiting essentially monotonic decreases with height have low scores. Moisture profiles exhibiting multiple layers for which the H_20

vapor mixing ratio exhibits abrupt changes (increasing followed by decreasing ratios) have progressively higher scores.

The impact of the moisture profile score on validation is shown in Figure 5

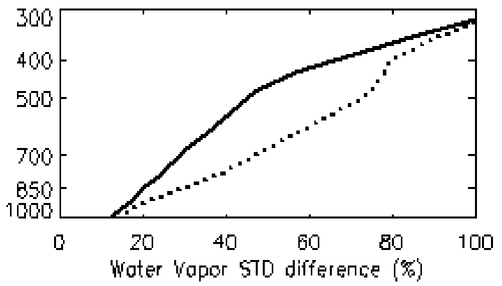

Figure 5: Raob-minus-NWP H_20 vapor fractional (%) differences for Raobs with moisture scores 0,1 (solid) and 2 or more (dashed).

As seen, the radiosonde and NWP agreement improves on the order of 20% to 40% in the middle troposphere for moisture scores of 0 or 1. This underscores the potential impacts for satellite validation (and tuning) since most sensors cannot unambiguously discern moisture structures corresponding to higher scores. Furthermore, the possibility that the radiosonde profile contains errors also increases for higher scores.

Similar impacts (not shown) are also observed for profiles exhibiting temperature inversions, particularly those exceeding 2 km in depth.

NPROVS, and in particular the EDGE analytical interface, also keeps track of the respective satellite observations and sounding profile quality control (QC) indicators as provided for each respective platform. The satellite QC is not considered when compiling collocations but can significantly impact respective validation results as discussed in Section 3.

NPROVS collocation datasets are compiled daily and are processed into weekly and monthly datasets for more meaningful statistical validation and archive.

3. RESULTS

3.1 Validation and Screening Strategies

The EDGE statistical interface provides options for applying QC information to select collocations for validation. QC parameters are available for the respective satellites and ground truth radiosondes. Figures 6a through 7b illustrate the impacts of various sampling strategies with respect to satellites and radiosonde QC. The period of record is a 7-day period in January 2009 for which the total sample size of candidate radiosondes for collocation was about 7000.

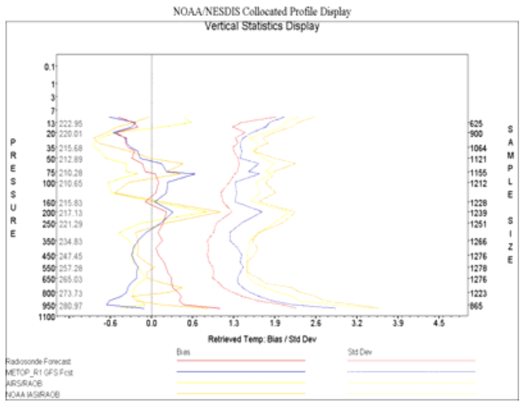

Figure 6a: AIRS and IASI Sat-minus-Raob mean and standard deviation differences for temperature at 500mb using Sat products which passed QC.

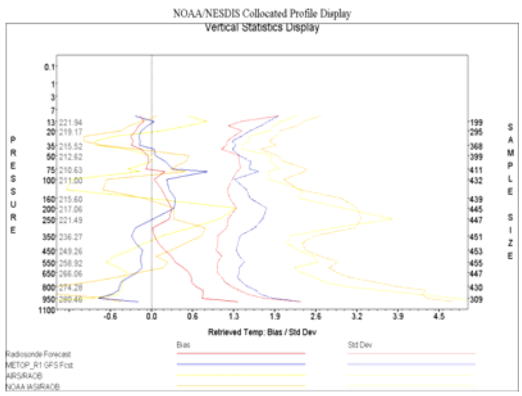

Figure 6b: AIRS and IASI Sat-minus-Raob mean and standard deviation differences for 500mb temperature using Sat which failed QC.

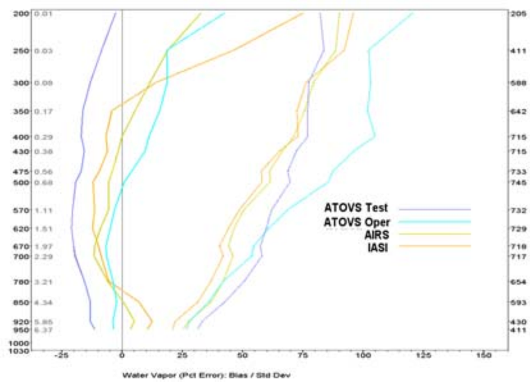

Figure 7a: ATOVS and hyper-spectral Sat-minus-Raob H_20 fraction (%) mean and standard deviation differences using Raobs with H_20 profile scores of 0 or 1.

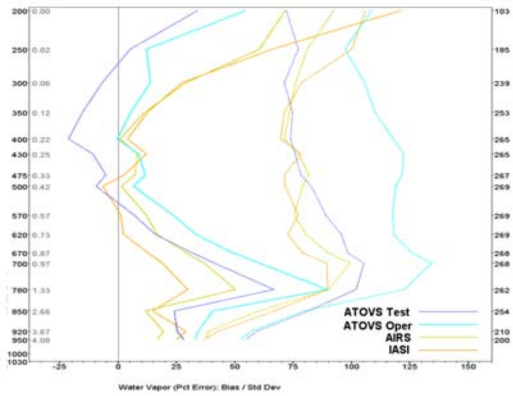

Figure 7b: ATOVS and hyper-spectral satellite-minus-raob $H_2$0 vapor fractional (%) mean and standard deviation differences using raobs with $H_2$0 profile scores of 2 or more.

Figures 6a and 6b show examples of the impact of satellite QC for collocations containing hyper-spectral AIRS and IASI soundings for validating temperature (and first guess). The vertical pressure scale ranges from 1000 mb to 10 mb and the horizontal axis ranges from -1.5K to 4.5 K.

Figures 7a and 7b show examples of the impact of radiosonde $H_2$0 vapor tests (section2) for collocations containing ATOVS and hyper-spectral soundings that passed their respective QC. The vertical pressure scale is from 1000 mb to 200 mb and the horizontal axis from -25% to 150%.

As seen, the impact of increased QC reduces the satellite-minus-radiosonde differences but also reduces the sample size (yield). In figure 6a, requiring collocations to contain both AIRS and IASI soundings for which both passed their respective QC resulted in a sample reduction (originally about 6000) to approximately 1300 (almost 80%). By comparison, the sample of collocated AIRS and IASI products which both failed QC is about 450 but as can be seen the satellite-minus-radiosonde differences are significantly increased.

Figure 7a illustrates the cumulative impact of combined satellite and radiosonde QC requirements for validation. In this case, requiring that collocations contain AIRS, IASI and ATOVS products which all passed QC and for which the radiosonde moisture score was 0 or 1 reduced the sample to about 750 (almost 90%). Figure 7b is similar to 7a but only includes collocations which all passed their respective satellite QC but were collocated with radiosondes with moisture scores of 2 or more. As seen, the satellite-minus-radiosonde differences are significantly higher, attesting to the value of the moisture profile scores; the sample is relatively small (approximately 250).

The point is that care is needed when comparing different satellites to insure compatible QC constraints while retaining adequate sample size for a meaningful validation. This begins with assuring that the respective QC protocols among the satellite systems are compatible, which they are not, and for which NPROVS is a good source of feedback.

3.2 Platform Performance and Sensitivity

One of the potential strengths of the NPROVS collocation dataset is that it can provide feedback to data providers (and managers) concerning the sensitivities and relative performance of the respective satellite and ground truth data platforms, particularly those secured over a continuous long-term record.

Figure 8 illustrates an example using the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) radio occultation (RO) observations to monitor radiosonde upper troposphere relative humidity (RH) for different radiosonde instrument types. Shown are histograms of upper troposphere (300 mb) Radiosonde-minus-COSMIC mean relative humidity differences segregated by specific radiosonde type groupings (Sun *et al.* 2009). Blue indicates daytime differences and gray nighttime differences.

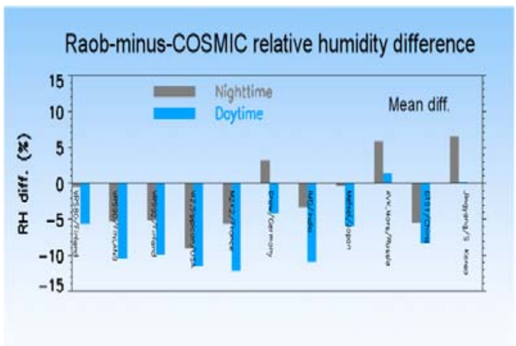

Figure 8: Histogram of COSMIC-minus-Raob daytime versus nighttime upper tropospheric RH for a 6-month period in 2008 for radiosonde instrument type groupings (Sun and Reale 2009)

Results indicate an overall dry bias for the radiosondes (lower relative humidity) except for selected radiosonde types over Russia and that the bias is generally greater during the day than at night. Normally, studies of this nature are obtained through intensive and expensive research field experiments but using NPROVS are achieved through relatively inexpensive data compilation and archive. Results agree with previous publications from such experiments (Wang and Zhang 2008).

The sensitivity of collocated observations with respect to spatial and temporal differences is also a topic of interest. Emerging principles for the Global Climate Observing System (GCOS) Reference Upper Air Network (GRUAN) (World Meteorological Organization 2008) have questioned the importance of synchronized satellite and ground truth observations from GRUAN sites for climate. Preliminary studies using NPROVS estimate spatial and temporal sensitivities on the order of 0.1K per hour and 0.5K per 100 km on the real-time weather scale (Sun 2009). Studies to determine such sensitivities on the climate scale using significantly longer periods than for this study are pending.

Figure 9 illustrates the impact of including the drift parameters for radiosondes (see figure 3) and COSMIC profiles when compiling collocated observations for computing vertical accuracy statistics. Radiosonde time and location are “conventionally” defined at the surface and for COSMIC at the “occultation point” (typically between 700 mb and 500 mb). As seen, accounting for drift (solid red and blue curves), which can vary up to 200 km and 3 hours for the radiosonde and 200 km for the COSMIC can reduce RMS errors 10% or more.

Figure 10 shows mean (left) and standard deviation (right) statistics for NOAA NWP and five independently processed sets of temperature soundings from the ATOVS Operation, ATOVS Test, MIRS, AIRS, and IASI systems versus the collocated radiosondes. The sample is a common denominator subset of satellite observations over sea for which the recommended QC parameters for each system are adhered, an almost 95% reduction of the original sample.

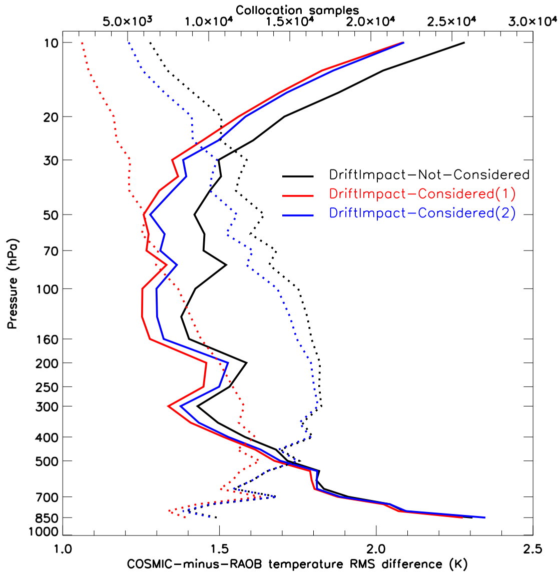

Figure 9: Reduction in Raob-minus-COSMIC RMS (solid) when drift parameters are used (red and blue) to compile collocation samples (dashed) at each vertical level (Sun et al. 2009) with sample size along the top axis and pressure (1000 mb to 10 mb) along the left axis.

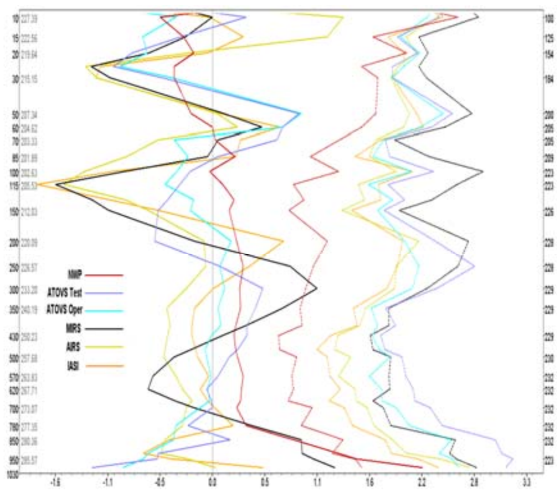

Figure 10: Mean and standard deviation differences for NOAA NWP (red) and five (5) independent suites of temperature soundings-minus-Raob for the period February 8-14, 2009 using common denominator sample (right axis) over sea with pressure (1000 mb to 10 mb) along the left axis.

Figure 11 illustrates trend statistics from the recently deployed NARCCS for the period January through July, 2009.

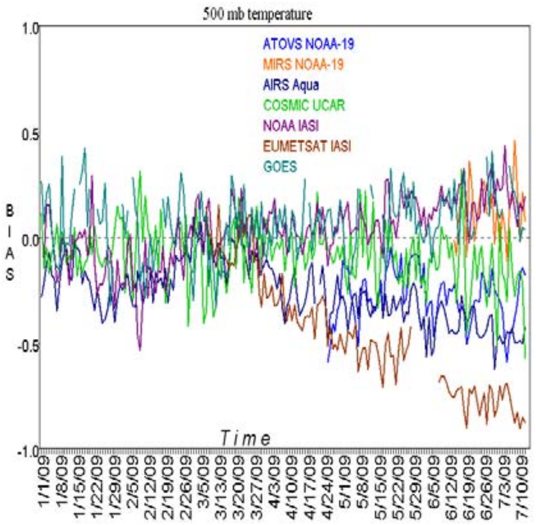

Figure 11: Trend plots of mean Sat-minus-Raob 500mb temperature (+/- 1.0K) for denoted Sat systems during January to July 2009

As seen most systems collect within +/- 0.5K except for EUMETSAT IASI which gradually drifts out of this range over time.

The integrated use of NPROVS collocation datasets and associated EDGE analytical interface provides an effective validation, troubleshooting and product monitoring function at STAR.

5. SUMMARY

NPROVS provides NOAA STAR with a centralized validation protocol for the routine monitoring and inter-comparing derived atmospheric weather products from polar orbiting and GOES environmental satellites. This is primarily achieved through the compilation and analysis of collocated radiosonde, NWP and independently processed satellite product systems; currently 19 operational and experimental products systems are included. NPROVS compiles collocations on a daily basis with all collocations routinely archived at STAR. NPROVS includes a variety of analytical interface and sampling options (EDGE) including satellite and Raob QC, space and time windows, terrain designation, individual and common denominator sampling, radiosonde instrument type selection, regional (ie GOES Conus) designation and more. Analysis on real-time weather (daily, weekly) and climate scales (monthly, seasonal, annual) are facilitated.

6. REFERENCES

Pettey, M., B. Sun and A. Reale, 2009: The NOAA product (integrated) validation system (NPROVS) and environmental data graphical evaluation (EDGE) interface part-2: system. *25th Conference on International Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography and Hydrology,* 89th AMS Annual Meeting, Phoenix, AZ, 11-15 Jan.

Sun, B., A. Reale and D. Hunt, 2009: The use of Radiosonde-COSMIC collocations to identify radiosonde type characteristics and quantify the impacts of collocation mismatch on validation. *13th Conference on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface (IOAS-AOLS),* 89th AMS Annual Meeting, Phoenix, AZ, 11-15 Jan.

Reale, A. and F. Tilley, 2009: ATOVS derived soundings using NOAA products validation system (NPROVS) datasets for computing first guess and sensor bias adjustments independent of NWP. *16th Conference on satellite Meteorology,* 89th AMS Annual Meeting, Phoenix, AZ, 11-15 Jan.

Reale, A., F. Tilley, M. Ferguson and A. Allegrino, 2008: NOAA operational sounding products for ATOVS. *IJRS,* **29,** (16), 4615-4651

Wang, J., and L. Zhang, 2008: Systematic errors in global radiosonde precipitable water data from comparisons with ground-based GPS measurements. *Journal of Climate*, **21**, 2218-2238.

Boukabara, S. A., F. Weng and Q. Liu, 2007: Passive microwave remote sensing of extreme weather events using NOAA-18 AMSUA and MHS. *IEEE Trans. on Geoscience and Remote Sensing*, July 2007. **45**, (7), 2228-2246.

Goldberg, M.D., Y. Qu, L.M. McMillin, W.W. Wolf, L. Zhou and M. Divakarla, 2003. AIRS near-real-time products and algorithms in support of operational weather prediction. *IEEE Trans. Geosci. Remote Sens.*, 41, 379-389.

World Meteorological Organization, 2008: Report of the GCOS reference upper-air network meeting. *GCOS-121*, WMO/TD No. 1435.