5A.5 IMPROVEMENTS IN GOES/GOES-R WATER VAPOR CALIBRATION AND VALIDATION

Seth I. Gutman* and Daniel L. Birkenheuer
NOAA Earth System Research Laboratory, Boulder CO

1. INTRODUCTION

The calibration and validation (Cal/Val) of GOES-R moisture products will be challenging primarily for two reasons. The first is that temperature and humidity retrievals using the limited number of infrared channels on the Advanced Baseline Imager (ABI) will require the use of a numerical weather prediction model to provide the first guess for profile estimation. The second is that the relative absence of independent observations with known and/or verifiable accuracy and precision limits our ability to validate the observations and products derived there-from. Of course, these problems are not unique to GOES-R, and so improvements in our ability to calibrate, validate, and verify satellite and other upper-atmospheric observations are important in numerous areas, especially climate monitoring, but also in weather forecasting and research.

Over the last several years, NOAA's Earth System Research Laboratory (ESRL) has investigated the error characteristics of current GOES and GOES-R proxy total precipitable water (TPW) products by comparing them with TPW retrieved from Global Positioning System observations. In the process, we have compiled evidence of systematic errors in GOES TPW products as well as in the operational numerical weather prediction (NWP) models providing the first guess for these retrievals. This has raised some concern, especially in the climate community which is increasingly dependent on satellite observations to monitor climate change and verify climate model predictions (Birkenheuer and Gutman, 2010).

This paper summarizes our major findings and makes specific recommendations to improve both existing GOES water vapor products and those derived from next generation of U.S. environmental satellite sensors in geostationary and polar Earth orbits.

2. CALIBRATION AND VALIDATION ISSUES

2.1 Need for Objective Standards

According to the Bureau International des Poids et Mesures, calibration is a process of determining an instrument's response to known inputs and validation is the provision of objective evidence that a given item fulfils specified requirements that are adequate for an intended use (BIPM, 2008). In general, we think of calibration as applying to instruments and validation applying to retrievals or derived products. To accomplish calibration and/or validation with certainty, these activities should be carried out with respect to standards whose accuracy and precision are known and can be verified. This after all is the basis of Climate Reference Network and similar terrestrial climate observing systems.

In the past few years, the satellite (Ohring et al., 2005) and upper-air climate observing communities (WMO, 2007) have begun to address the need for objective criteria to verify the accuracy of observations and take tangible steps to implement them (NRC, 2007; WMO, 2009).

This paper presents an alternative methodology for verifying the validity of on-orbit or vicarious calibration of satellite sensors that is fully consistent with the approach of the GCOS Reference Upper-Air Network. Additionally, it provides an independent and thus unbiased way of evaluating the accuracy of data assimilation and prediction systems for certain parameters such as upper-air moisture.

2.2 GNSS Measurements as a Cal/Val Standard

GPS Meteorology (Figure 1) uses GPS or other Global Navigation Satellite System (GNSS) receivers to estimate the total refractivity of the lower atmosphere above a fixed site. Using the theory and techniques described by Bevis et al. 1992, Duan et al. 1996, and Gutman et al. 2004, it is possible to retrieve total TPW in the troposphere with high accuracy under all weather conditions. TPW is measured by other atmospheric observing systems including satellites, balloon-borne sounding systems and aircraft because it plays an important role in weather and climate process. In addition to the fact that GNSS operates under all weather conditions, even in the presence of clouds and precipitation that adversely impact other observing systems, GNSS has attributes that make it unique among all atmospheric observing systems.

Smith and Weintraub (1953) described the refractivity of the electrically neutral (non-dispersive) atmosphere in terms of temperature, pressure and water vapor along the paths of radio signals through the troposphere. Using the GNSS signal, the accuracy with which we are able to estimate tropospherically-induced signal delays depend on (1) the accuracy of the atomic clocks used to measure the time of flight of the radio signals and (2) our ability to model systematic errors in the positions of the satellites in space and receivers on the ground.

∗ *Corresponding author address: Seth I. Gutman, NOAA Earth System Research Laboratory, 325 Broadway, R/GSD7, Boulder, CO 80305-3328; e-mail: seth.i.gutman@noaa.gov*

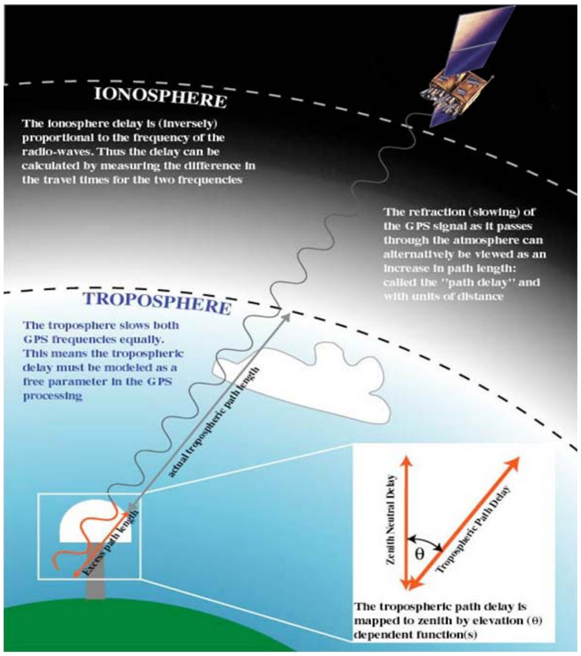

Fig 1. Estimating the total refractivity of the atmosphere by measuring GPS signal delays. Modified from an illustration by Steven M. Businger and used by permission.

Since the accuracy of the atomic clocks used in GNSS are constantly improving, the number of GNSS satellites in orbit are increasing, the number of GNSS receivers on the ground and their distribution over the planet are also increasing, and the models used to estimate positions in space and time are improving, it seems reasonable to assume that even if multipath-induced errors at GNSS sites and electronic noise and receiver-dependent errors do not get worse than they presently are, the accuracy of any parameter estimated elements should similarly improve with time. The GNSS error budget for the pseudorange observable is presented in (1).

$$P = R + c\,(\Delta T - \Delta t) + \Delta\, ion + \Delta\, trop + \Delta\, multi + \varepsilon \qquad (1)$$

where:
P = measured pseudorange
R = the geometric range to the satellite
ΔT and Δt = errors in the receiver and satellite clocks
$\Delta\, ion$ = ionospheric signal delays
$\Delta\, trop$ = tropospheric signal delays
$\Delta\, multi$ = errors introduced by multipath
ε = receiver noise.

In practice, the tropospheric signal delay is modeled as a free parameter in the estimation of antenna position. The strength of GNSS for atmospheric remote sensing is that as long as there are four or more satellites in view at one time, the estimate of the tropospheric signal delay is over not under-determined as it is for many other remote sensing observations using, for example, radiometric techniques.

As a consequence we should (at least in principle) be able to use GNSS to independently check the accuracy of any observing system that makes a column-integrated measurement that is transformable into units of length as long as we can define a suitable transfer function that allows us to map an observation of TPW (and possibly other parameters that are transformable into units of length) to a GNSS observable.

There are two recent examples of applications of this hypothesis. In the first, Gutman et al., 2005 demonstrated that it was possible to use GNSS observations made in close proximity to upper-air sites to detect erroneous rawinsonde soundings in near real-time with high probability of detection and low false alarm rate.

In the second, McMillin et al. (2007) used GNSS TPW estimates to reduce systematic errors in rawinsonde moisture soundings. The method used to do this was described by Turner et al. 2003 and involved (1) comparison of radiances calculated from raw rawinsonde temperature and moisture soundings with radiances measured by the AIRS radiometer aboard the Aqua spacecraft, and (2) bias-correction of rawinsonde moisture soundings and comparison of the corrected rawinsonde-derived radiance estimates with the AIRS observations. In almost all cases, the differences between the AIRS observations and the corrected rawinsonde estimates were smaller than the uncorrected comparisons. This provided an estimate of the possible accuracy of the AIRS radiometer by providing a verifiable comparison with a completely independent observation.

Interestingly, one of the recommendations made by McMillin et al. 2007 was for GNSS water vapor observing systems (known in the U.S.A. as GPS-Met and in Europe as E-GVAP) be added to all GCOS GUAN sites with the goal of providing independent Cal/Val references for satellite IR water vapor estimates over land and improved global rawinsonde data quality for climate studies and assimilation into global numerical weather models.

3. IDENTIFICATION OF DIFFERENCES IN GOES-EAST AND GOES-WEST TPW PRODUCTS

This section compares and contrasts water vapor estimates made by NOAA at more than 400 GNSS sites distributed over the contiguous 48 United States (Figure 2) with TPW products derived from the GOES-East and West satellites every hour. Three facts are of paramount importance in this discussion:

1) estimates of TPW derived from GNSS signal delays are completely independent of TPW estimates derived from GOES sounder radiances;
2) retrievals of TPW made by different GPS receivers in close proximity (meters apart) agree at the sub millimeter TPW level;

3) characteristics of the infrared sounders aboard the GOES-8 through GOES-14 spacecraft from which derived water vapor products were created and used in this study are essentially identical (Schmit et al., 2009).

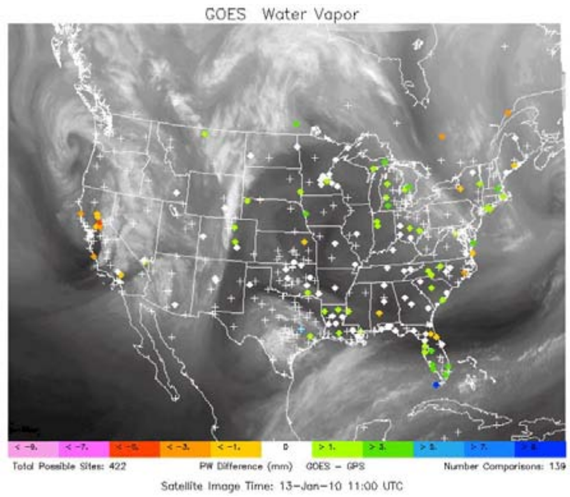

Fig 2. GOES water vapor image with locations of GPS sites providing dual frequency carrier phase observations to NOAA. The color of the dots represents the amount of TPW retrieved from the measured GNSS signal delays. Black dots identify stations that did not provide observations at the time of this image.

Obvious systematic differences between GNSS TPW estimates and operational GOES-derived values are apparent in Figure 3 (GPS vs GOES-East) and Figure 4 (GPS vs GOES-West). In both figures, the top panel is bias (GOES minus GPS); the middle panel is RMS difference (GOES minus GPS); and the bottom panel indicates the number of sites used in the calculation of these statistics every hour. The red line in all cases is the 5-day running average of these values. Here are some of the major differences:

- GOES-East TPW has a strong positive bias during the warm months that is not apparent in the GOES-West comparisons;
- GOES-East is wetter than GNSS most of the time regardless of season;
- RMS differences between both satellites increase during the warm months although GOES-East exhibits more variability than GOES-West;
- Because the number and distribution of GPS sites-East of the Rocky Mountains far exceeds those to the west, the number of GOES-East comparisons every hour are about twice that of GOES-West.
- The number of collocated GOES-East & West comparisons is very limited.

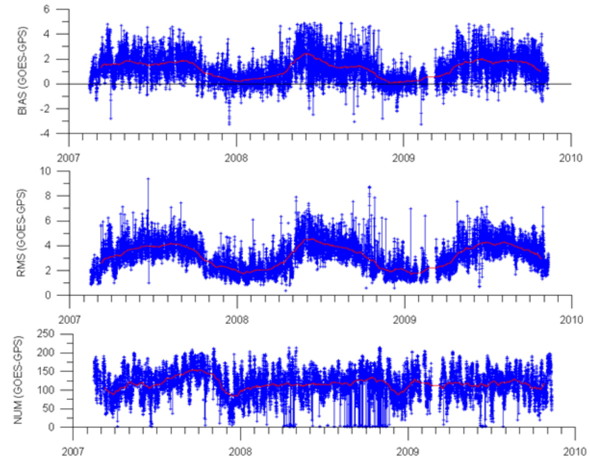

Fig 3. Operational *GOES-East* minus GPS TPW at sites in the domain shown in Figure 2 over about 1,023 days between 2007 and 2010. The average difference is 1.452 mm (GOES-East > GPS) and the RMS is 3.244 mm.

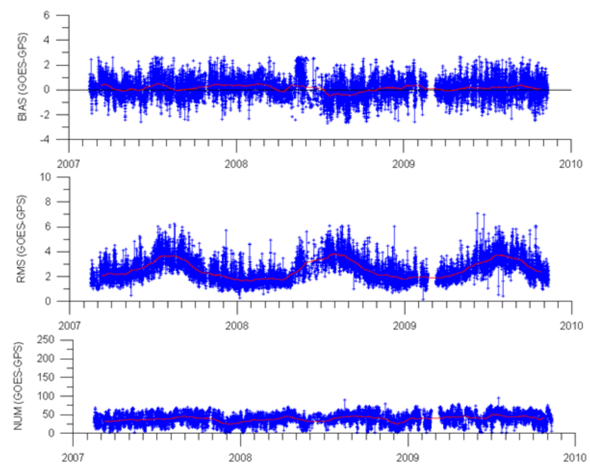

Fig 4. Operational *GOES-West* minus GPS TPW at sites in the domain shown in Figure 2 over about 1,023 days between 2007 and 2010. The average difference is 0.299 mm (GOES-West > GPS) and the RMS is 2.522 mm.

Another way to assess the systematic differences between GPS TPW and the GOES satellites products presented above is to plot the bias versus RMS differences for both systems.

If both observing systems have small systematic and random errors, we expect the result to plot as a nearly symmetric cluster of points distributed near the origin (bias = 0, RMS differences = 0). Figures 5 and 6 present the actual comparisons using the data presented above. Note the asymmetry in the GOES-East comparisons (Fig 5) caused by the GOES-East wet bias and the tendency for the GOES-East TPW bias and RMS differences to be temporally correlated.

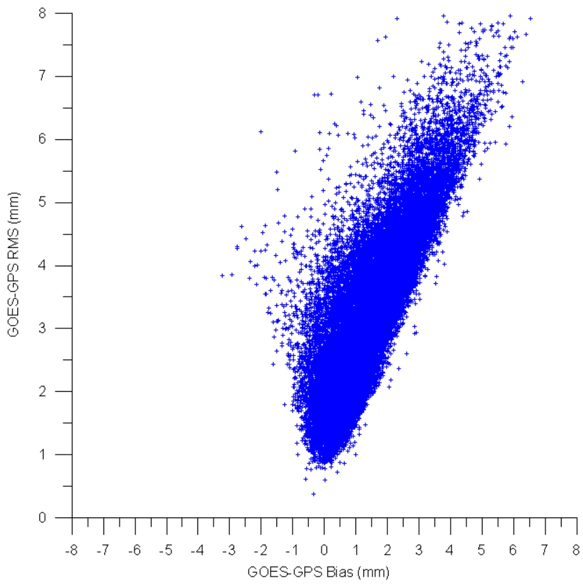

Fig 5. Plot of operational *GOES-East* TPW bias on the abscissa versus RMS differences on the ordinate. This plot presents 28,524 comparisons between GPS and the operational GOES-East TPW product used in Figure 3 above.

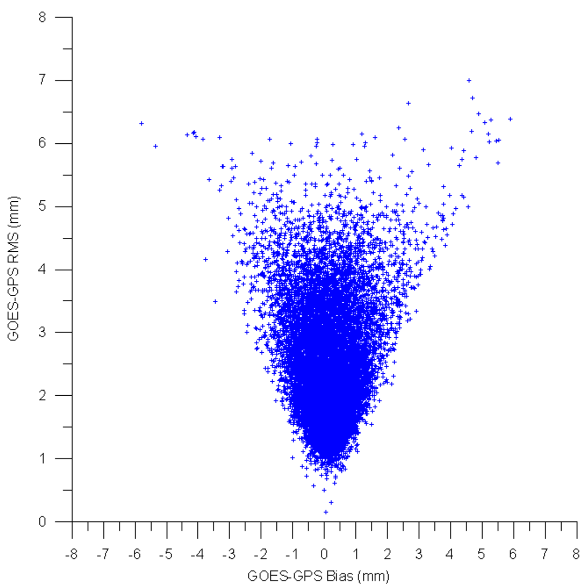

Fig 6. Plot of operational *GOES-West* TPW bias on the abscissa versus RMS differences on the ordinate. This plot presents 24,554 comparisons between GPS and the operational GOES-West TPW product used in Figure 4 above.

4. IMPROVEMENTS IN GOES TPW PRODUCTS

4.1 Emperical Corrections

Improvements in GOES TPW products can be achieved in several ways. An empirical approach was taken by Birkenheuer et al. 2008 that involved empirical modeling of the diurnal differences between operational GOES-East and GOES-West TPW products and GPS estimates, and applying these corrections to each GOES satellite TPW product to minimize systematic diurnal errors. An extension of this technique to account for and remove systematic regional and seasonal variations is thought to be straight forward. However, since this approach is not physically based, its operational implementation is clearly not desirable.

4.2 Physical Retrievals

Rather than the direct approach described above, the approach being taken by the University of Wisconsin's Space Science Engineering Center (SSEC) in Madison, WI is to identify the source of and then mitigate systematic errors identified in the GOES-GPS TPW comparison product. This activity started in November 2007 and is an ongoing collaboration between NESDIS, OAR and UW scientists that started in November 2007 and uses the GOES-GPS TPW comparisons generated by ESRL to 1) identify potential problems and 2) evaluate the response of the TPW product to changes in the retrieval algorithm.

Figure 7 is identical to Figure 3 except that the TPW comparisons use algorithms that are under testing and evaluation by SSEC in preparation for the launch of GOES-R.

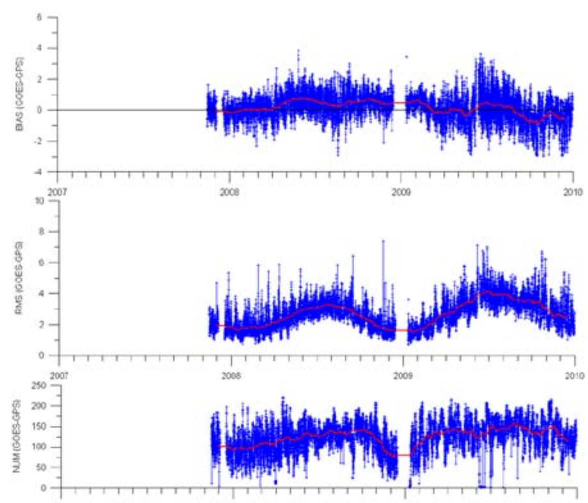

Fig 7. Experimental *GOES-East* minus GPS TPW over about 780 days between 2007 and 2010. The average difference is 0.149 mm (GOES-East > GPS) and the RMS is 2.681 mm.

The apparent improvement in the experimental GOES-East TPW product over the existing operational product is impressive: average bias reduced by an order of magnitude and RMSs reduced by almost 20%. In fact, the experimental GOES-East TPW comparisons have much in common with the GOES-West comparisons, which is what we would expect from functionally and physically equivalent sounders.

Figure 8 plots SSEC GOES-East bias versus RMS, confirming the improvement.

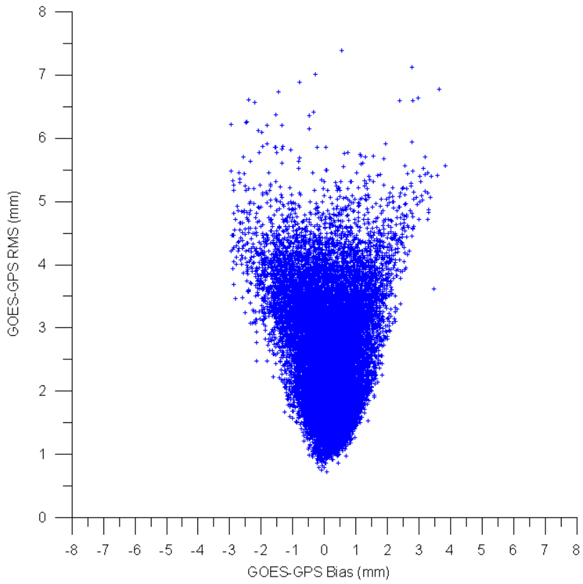

Fig 8. Plot of SSEC *GOES-East* TPW bias on the abscissa versus RMS differences on the ordinate. This plot shows 16,450 comparisons between GPS and the experimental GOES-East TPW product used in Figure 7 above.

4.3 *Numerical Weather Prediction Models*

When you have a limited number of channels with which to retrieve physical parameters such as PW from radiometers, it is common practice to use a numerical weather prediction model to provide an estimate of the vertical profile of that parameter as the first guess in an iterative solution. This is the case in the GOES retrievals of TPW, and will also be the case with the Advanced Baseline Imager aboard the GOES-R series spacecraft.

The operational weather prediction model used by NOAA for this purpose is the Global Forecast System (GFS) described in Section 2 of Yang et al., 2006. The GFS does not currently assimilate ground-based GPS-Met observations, so GFS analyses and predictions are thought to be completely independent of ground-base GPS observations.

It should be noted that the GFS currently assimilates refractivity estimates derived from satellite radio occultation (RO) measurements (Cucurull and Derber, 2008). These measurements are made using the same GPS radio signals that are used by ESRL to estimate TPW using ground-based techniques referenced previously. The errors associated with estimates of signal delay or TPW derived from ground based measurements or bending angle or refractivity estimates and temperature, pressure and water vapor retrievals derived from space-based measurements are independent and uncorrelated with the following proviso: correlated errors will undoubtedly occur if problems with the signals broadcast by the GNSS satellites or mis-modeling of their orbits are undetected and these data are not excluded from analysis and/or direct assimilation.

Figure 9 compares the TPW analyses from the GFS model with GPS TPW retrievals over the same period (2007-2010). Note the change in vertical scale from the previous timeseries.

Averaged over this period, GFS has a dry bias with respect to GPS of about -0.022 mm and the RMS difference is 2.6 mm. Two features of interest are the dry bias during the warm seasons and the positive slope (+3.2 mm/decade) of the GFS-GPS TPW differences. The large GFS decadal trend may be caused by model changes over this period (George Ohring, personnel communication). If this is the case, it should disappear when we compare the GPS estimates with the GFS reanalysis since the latter uses a fixed software configuration.

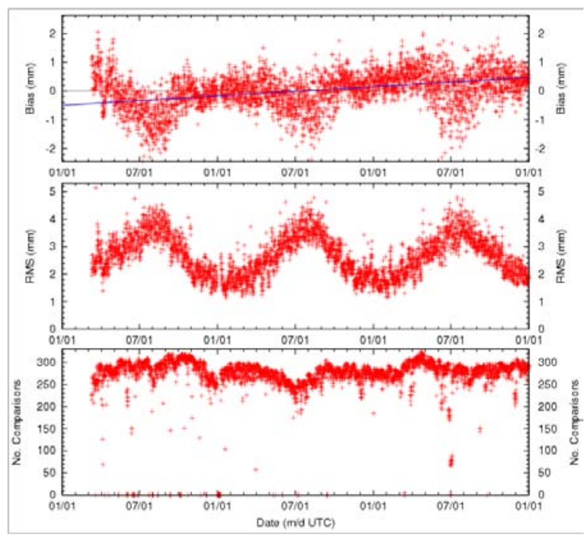

Fig 9. Operational GFS minus GPS TPW at sites in the domain shown in Figure 2.

The question of what impact systematic errors in the GFS analysis have on GOES and other satellite retrievals, and the use of these data in climate studies (Li et al. 2008; Seo et al. 2005) is posed but not answered in this paper.

5. GOES-R TPW SIMULATION

The spectral bands of the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Aqua and Terra spacecraft have similar counterparts to those on the Advanced Baseline Imager (ABI) aboard GOES-R. Using a data set prepared by CIMMS (Gunshor et al. 2008), TPW products derived from MODIS radiances were used by us to evaluate the probable characteristics of GOES-R TPW retrievals.

Figure 10 compares one year of MODIS-derived TPW with GPS estimates at the sites identified in Figure 2. Note the change in vertical scale from similar previous figures.

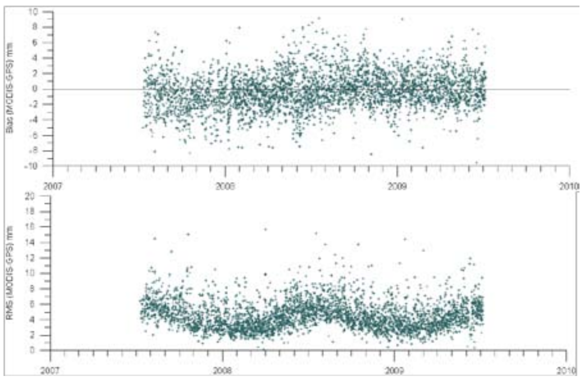

Fig 10. MODIS/ABI proxy minus GPS TPW over 841 days between 2007 and 2009. The average difference is -0.278 mm (GPS > MODIS/ABI Proxy) and the RMS is 4.497 mm: about twice that of GOES

Figure 11 plots MODIS/ABI proxy TPW bias on the abscissa versus RMS differences on the ordinate. This plot shows only 3727 comparisons between MODIS/ABI and GPS. Note the change in vertical and horizontal scale from previous satellite-GPS comparisons.

While the MODIS-GPS TPW bias is small, the almost two-fold increase in the scatter of the MODIS-derived estimates about the mean compared to the other comparisons is troubling.

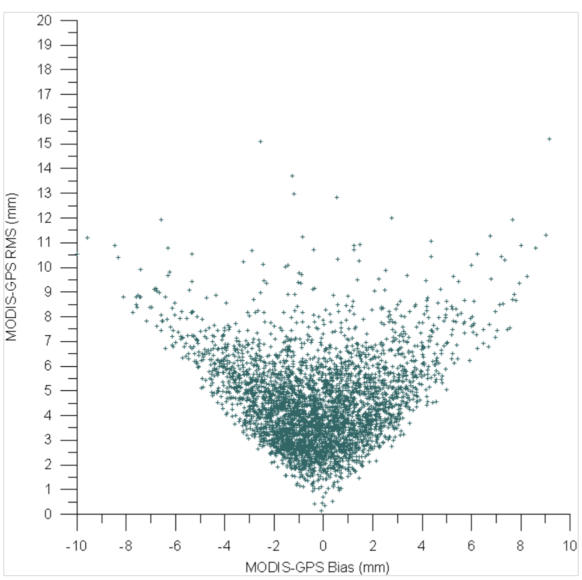

Fig 11. MODIS/ABI Proxy TPW bias on the abscissa versus RMS differences on the ordinate. This plot shows 3727 comparisons with GPS.

6. CONCLUSIONS AND RECOMMENDATIONS

The characterization of observing system errors requires comparison with independent observations of known accuracy and precision. In the absence of direct (one-to-one) comparisons, observation errors can be estimated by defining a transfer function that allows one of the measurements to be expressed in terms of units that are common to the other.

The demand that satellite observations meet the same rigorous requirements for accuracy and precision as terrestrial measurements means that the methods used to calibrate and validate remote sensing observations must be irrefutable. Finding a method to do this for any much less all observations will be a challenging but rewarding endeavor.

In this paper we have presented an argument that GNSS observations may be able to provide this level of certainty for selected observations that can be expressed in terms of units of length. Our goal is to prepare a rigorous logical argument that demonstrates this in the near future.

Until that goal is accomplished, we can say with certainty that GPS/GNSS observations provide us with a way to estimate the average refractivity above a fixed site that does not require external calibration and with accuracy that improves rather than degrades with time. Errors in modeling each element of the GNSS error budget except receiver noise and multipath are systematic and can be reduced over time with improvements in the atomic clocks, increasing the number of signals in space and receivers on the ground.

Comparison of GOES-East and West TPW products with GPS observations allowed us to (1) monitor the performance of these sensors over time and (2) identify systematic changes in the characteristics or performance of the GOES instruments virtually in real time.

Developers of experimental satellite retrieval algorithms now have access to near instantaneous feedback on changes made to any part of the satellite data processing chain using actual rather than synthetic observations.

Comparison of GPS data with the operational GFS model has revealed a heretofore unknown or undocumented systematic error in modeling the fields responsible for the estimation of integrated (total column) precipitable water, at least over CONUS. Since the GFS provides the first guess for many satellite retrievals, and is used in climate studies, it is imperative that this finding be independently verified and the problem corrected.

We recommend the following:

6a. GPS/GNSS receivers should be installed at all GCOS/GUAN and GRUAN sites as part of WIGOS.

6b. International protocols for processing GPS/GNSS data for climate applications should be established.

6c. Independent verification and validation of upper-air observations and observing systems should start now, not sometime in the indefinite future when hypothetical capabilities are developed, tested and implemented.
6d. Transition of improved satellite data processing algorithms from research to operations should be accelerated.
6e. Although the MODIS/ABI Proxy bias is relatively small, the cause of the increased scatter about the mean compared with current GOES sensors should be investigated.
6f. Comparisons between GPS and the GFS TPW analysis should be repeated using the GFS reanalysis to determine if the cause of the increase in model TPW is due to configuration changes or some other factor.

7. ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of the following individuals to our research. Kirk Holub at OAR/ESRL in Boulder, CO; Tim Schmit, Jim Nelson, Tony Schreiner, Gary Wade and Li Zhenglong at SSEC in Madison, WI; and Jaime Daniels, Americo (Rico) Allegrino at NESDIS/ORA in Suitland, MD.

8. REFERENCES

Bevis, M., S. Businger, T. Herring, C. Rocken, R. Anthes, R. Ware, 1992. GPS meteorology: remote sensing of the atmospheric water vapor using the global positioning system. J. Geophys. Res., Vol. 97, No. D14, 75-94.

Birkenheuer, D. and S. Gutman, 2005: A comparison of GOES moisture-derived product and GPS-IPW data during IHOP 2002. J. Atmos. Ocean. Technol., Vol. 22, No. 11, pp. 1840–1847.

__________, S. Gutman, S. Sahm, K. Holub, 2008: Devised scheme to correct GOES bias in GOES operational total precipitable water products. OAR-GSD Tech Memo-34, 18pp.

__________ and Gutman, 2010: Detection and characterization of systematic errors in atmospheric models. 18th Conference on Applied Climatology paper 5A.4, Atlanta, GA, Jan. 17-21.

Bureau International des Poids et Mesures, 2008: International Vocabulary of Metrology – Basic and General Concepts and Associated Terms VIM, 3rd edition, JCGM 200.

Cucurull, L., and J.C. Derber, 2008: Operational Implementation of COSMIC Observations into NCEP's Global Data Assimilation System. Wea. Forecasting, 23, 702–711.

Duan, J.M., M. Bevis, P. Fang, Y. Bock, S.R. Chiswell, S. Businger, C. Rocken, F. Soldheim, R.H. Ware, T.A. Hering, and R.W. King, 1996: Remote Sensing Atmospheric Water Vapor using the Global Positioning System, J. Appl. Meteor, 35, 830-838.

Gunshor, M.M., E. Olson, J. Sieglaff, T. Greenwald, A. Huang, and J. A. Otkin, 2008: GOES-R ABI proxy data set generation at CIMSS. Fifth GOES Users' Conference Poster Session, paper P1.89, New Orleans, LA, Jan. 20-24.

Gutman, S.I., Facundo, J. and Helms, D., 2005: Quality control of radiosonde moisture observations. Ninth Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface (IOAS-AOLS), paper 11.7, San Diego, CA, Jan. 9-14.

Li, S., J. Lu, G. Huang, and K. Hu, 2008: Tropical Indian Ocean Basin Warming and East Asian Summer Monsoon: A Multiple AGCM Study. J. Climate, 21, 6080–6088.

McMillin, L. M., J. Zhao, M. K. Rama Varma Raja, S. I. Gutman, and J. G. Yoe, 2007: Radiosonde humidity corrections and potential Atmospheric Infrared Sounder moisture accuracy, J. Geophys. Res., 112, D13S90, doi:10.1029/2005JD006109.

National Research Council, 2007: Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond. National Academy Press, 428 pp.

Ohring, G., B. Wielicki, R. Spencer, B. Emery, and R. Datla, 2005: Satellite Instrument Calibration for Measuring Global Climate Change: Report of a Workshop. Bull. Amer. Meteor. Soc., 86, 1303–1313.

Schmit, T.J., J. Li, S.A. Ackerman, and J.J. Gurka, 2009: High-Spectral and High-Temporal-Resolution Infrared Measurements from Geostationary Orbit. J. Atmos. Oceanic Technol., 26, 2273–2292.

Seo K. H., J. K. E. Schemm, C. Jones, and S. Moorthi, 2005: Forecast skill of the tropical intraseasonal oscillation in the NCEP GFS dynamical extended range forecasts. Climate Dyn., 25, 265–284.

Smith, E.K. and S. Weintraub, 1953. The constants in the equation of atmospheric refractive index at radio frequencies. Proceedings of the Institute of Radio Engineers, Vol. 41, No. 8, pp. 1035-1037.

Turner, D. D., B. M. Lesht, S. A. Clough, J. C. Liljegren, H. E. Revercomb, and D. C. Tobin (2003), Dry bias variability in Vaisala RS80-H radiosondes: The ARM experience, J. Atmos. Oceanic Technol., 20, 117– 132.

World Meteorological Organization, 2007: GCOS reference upper air network (GRUAN): Justification, requirements, siting and instrumentation options, Rep. GCOS-112 (WMO/TD No. 1379), Geneva, Switzerland.

__________, 2009: GRUAN Implementation Plan 2009-2013, Rep. GCOS-134 (WMO/TD No. 1506), Geneva, Switzerland.

Yang, F., H.L. Pan, S.K. Krueger, S. Moorthi, and S.J. Lord, 2006: Evaluation of the NCEP Global Forecast System at the ARM SGP Site. Mon. Wea. Rev., 134, 3668–3690.