CGMSXXX USA-WP-24
Agenda Item: II.2
Discussed in WG-2

Intercalibration of Geostationary Satellites with a Polar-Orbiting Satellite

Another approach for intercalibrating the geostationary radiances using one polar-orbiting sensor as a reference. This is another USA response to Action 29.26

Action Requested: None

INTER-CALIBRATION OF GEOSTATIONARY SATELLITES WITH A POLAR ORBITING SATELLITE

Xiangqian Wu
NOAA/NESDIS/ORA/CRAD/SIT
Camp Springs, Md

ABSTRACT:

A group of scientists envisioned the potential the international constellation of geostationary meteorological satellites has to offer to monitor rapidly changing phenomena in the tropical and subtropical regions, such as the life cycle of convection or the diurnal variation of various parameters. To realize that potential, one has to inter-calibrate the measurements made by these satellites. In this report of work in progress, some methods and issues of inter-satellite calibration were considered, and a method using a polar orbiting satellite to provide a common ground was chosen. Detailed procedure of data collection and quality control were presented, including the application of recursive regression. Preliminary results from this work are being used in compilation of a one-year dataset to study the diurnal variation of upper tropospheric water vapor.

1. INTRODUCTION

Five geostationary meteorological satellites are currently operating above the equator at 0°E (METEOSAT-7), 63°E (METEOSAT-5), 140°E (GMS-5), 135°W (GOES-10), and 75°W (GOES-8). From the advantageous points of those geostationary orbits, these satellites have the potential of providing global and nearly continuous monitoring of the surface and atmosphere in the tropical and subtropical regions. Such capacity could have strong positive impact on the observation and understanding of rapidly changing phenomena in the region, for example the life cycle of convection or the diurnal variation of various parameters.

In order to utilize the full potential of this constellation of geostationary satellites, however, one has to address the issue that these satellites were manufactured by different countries and operated by different agencies. The characteristics of the instruments are far from identical, and the measurements are not strictly comparable. A procedure therefore is necessary to inter-calibrate the measurements made by these satellites and bring them to a common ground.

Various methods exist for inter-satellite calibration, ranging from the more theoretical approach based on forward modeling using individual spectral response functions, to the more empirical approach based on comparison of overlapping observations. It is recognized that the central issue of inter-satellite calibration is that, with different spectral response functions, sensors often observe different “targets” even if they are perfectly co-located. For atmospheric channels such as the 6.7 μm water vapor channel, different sensors are sensitive to different portions (vertically) of the atmosphere. For surface channels such as the visible channel, different sensors are sensitive to different surface characteristics. The crucial step of inter-satellite calibration is thus to carefully choose the dataset that best represent the measurements to be inter-calibrated, in terms of atmospheric and surface conditions, seasons, geographic locations, observing conditions, and so forth.

In this work, we use the polar orbiting satellite NOAA-14 that flew under all the geostationary satellites to inter-calibrate the geostationary satellite measurements over a period of more than one year. The strategy is to find the measurements by the polar orbiting satellite that are concurrent and co-located with those from a geostationary satellite, study their differences, and apply that knowledge to estimate what the polar orbiting satellite would measure using measurements from that geostationary satellite. The major advantage of this approach is that it is based on samplings that best represent the measurements to be inter-calibrated. This method also directly converts geostationary satellite measurements to a common standard and avoids possible uncertainty in spectral response function.

2. DATA COLLECTION

2.1. Selection of Polar Orbiting Satellite and Instrument

Data for this comparison were collected from September 1999 to January 2002 (except for METEOSAT-5 data collection that started in October 2001), during which period NOAA-14 was the primary polar orbiting satellite. The geostationary (GEO) satellite measurements are made by imaging instruments that typically have broad spectral band and fine spatial resolution. Ideally, measurements by an imaging instrument onboard NOAA-14, such as the Advanced Very High Resolution Radiometer (AVHRR), should be used for inter-comparison. However, the AVHRR does not have the water vapor band at 6.7 μm as all geostationary satellites do, so the High-resolution InfraRed Sounder (HIRS) was used instead.

2.2. Measurements Co-location

Attempt to co-locate the HIRS and GEO data is limited to within 45 degree from GEO's nadir such that the maximum GEO zenith angle is about 60°, similar to that for HIRS. The HIRS and GEO data are considered concurrent if the relevant scanning lines were started within 15 minutes from each other. These data are then spatially co-located within the accuracy of operational navigation. A 5-by-5 pixel array of GEO measurements is used to match the HIRS field-of-view (FOV). Although actual GEO FOV varies, this is the best overall compromise.

2.3. Alignment of Viewing Geometry

Even if the HIRS and a GEO view the same location at the same time, the measurements may differ if the location is viewed with different zenith angle. This is particularly true for the 6.7 μm channel. If the atmosphere is not horizontally homogeneous, even the direction of view matters.

To avoid these confusions, the difference of HIRS' and GEO's zenith angles must be limited. Since it is the optical path within the atmosphere that must be similar between the HIRS and GEO, it was required that the difference between the secant of the two zenith angles, which is proportional to the optical path, is less than 0.05. This allows rather large difference in zenith angles around the nadir (~18°) but very small difference in zenith angles away from the nadir (~2° when zenith angle is 45°). It was also required that, if both zenith angles are larger than 5°, i.e. none can be regarded as nadir view, the relative azimuth angle between HIRS and GEO must be less than 30°. These ensure that the views by HIRS and GEO are adequately aligned to minimize the difference in measurements due to viewing geometry.

2.4. Homogeneity of FOV

As mentioned before, an array of GEO pixels was used to match a single HIRS FOV. This match is not perfect for all GEO's. In addition, navigation errors in either system, nominally a fraction of FOV, may further increase the magnitude of mismatch in space. To alleviate this problem, it was required that the standard deviation of the 5-by-5 GEO pixels for the 11 μm channel be less than 1°K. It is hoped that by avoiding scenes of large spatial variation, the difference in measurements due to spatial mismatch will also be reduced.

3. RECURSIVE REGRESSION

Scatter diagram of the co-located NOAA-14 HIRS and GMS-5 for the 11 μm channel is displayed in the upper left panel of Fig. 1, which reveals two features: (a) Even after all the careful preparations, some of the co-located pairs are undoubtedly affected by error (outliers). Interestingly, most outliers suggest that the low earth orbiting (LEO) NOAA-14 measurements were contaminated by clouds. This is plausible, since HIRS has larger FOV that is more vulnerable to partial cloud contamination. (b) Less obvious but as important, the two sets of measurements do not simply differ by a constant. The difference also seems to depend on the scene temperature (non-unity slope).

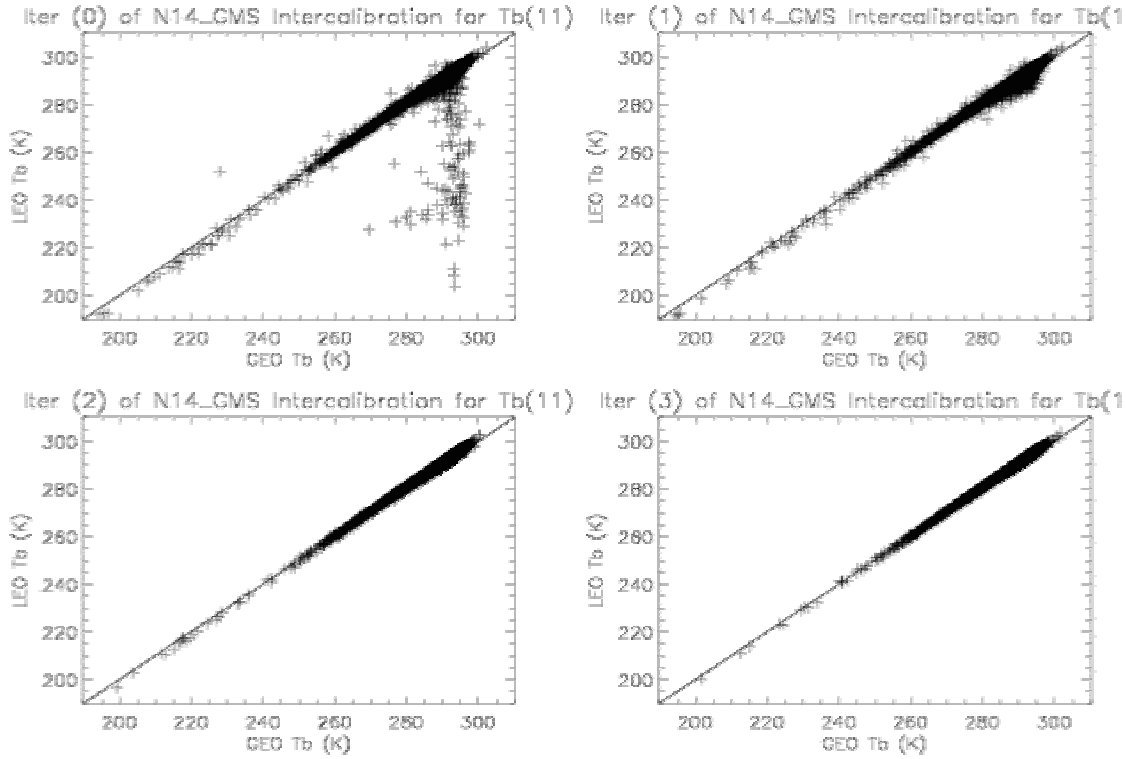

Figure 1: Scatter diagrams of the co-located NOAA-14 HIRS and GMS-5 for the 11 μm infrared window channel for the original data (upper left) and for the remaining data after removing outliers based on the first, second, and third regression (upper right, lower left, and lower right, respectively). The diagonal lines in these plots show that a linear regression based on these data may have non-unity slope, as well as non-zero intercept. Only 1/10 of randomly selected data were plotted in each diagram.

It was tempting to identify and remove outliers by imposing that the HIRS and GEO measurements be within certain range, however it is not practical to choose appropriate range objectively. To add to the complexity, the HIRS and GEO measurements are truly consistently different in certain circumstances (Fig. 2). Instead, a recursive regression is used to determine and confirm non-unity slope, as well as to identify and remove outliers. The premise of this method is that there are few outliers in the sample that do not behave as the majority in a statistically predictable way. The NOAA-14 and GMS-5 inter-calibration for the 11 μm channel is used to illustrate the procedure.

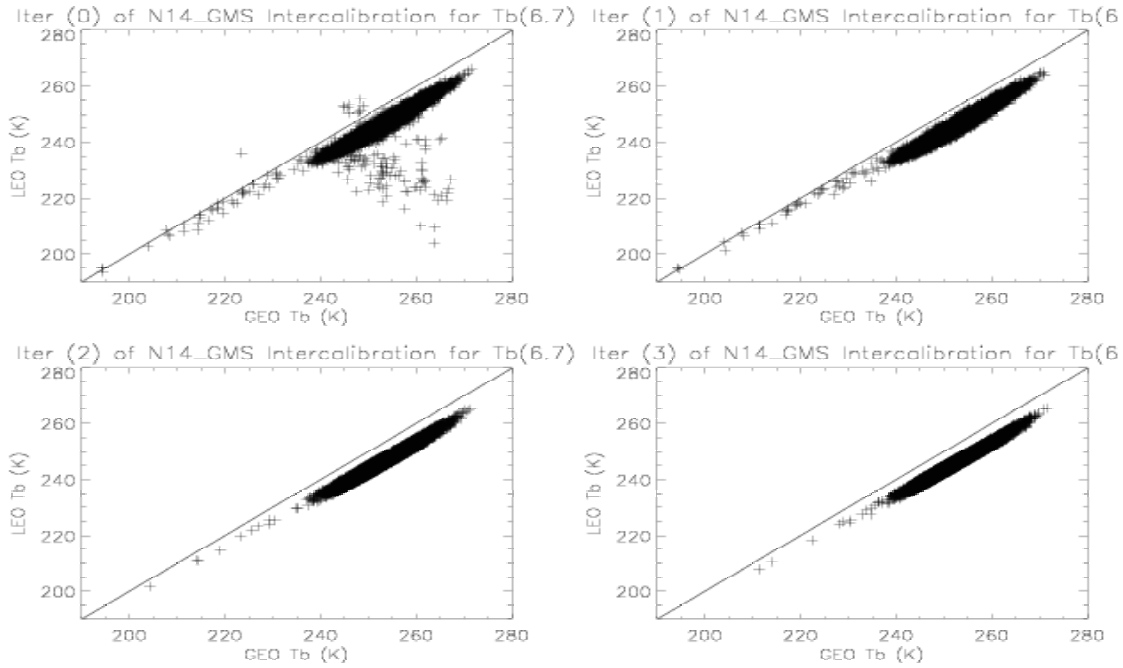

Figure 2: As Fig. 1, but for the 6.7 μm water vapor channel.

First, Fig. 1(a) suggests a linear regression:

(1) $$est\_HIRS = a + b * GEO$$

where *GEO* is $T_b(11)$ from GMS-5, *est_HIRS* is regression estimate of HIRS $T_b(11)$, and *a* & *b* are regression coefficients. Results of this regression are reported in the first row in Table 1, where the first column is the order of iteration (to be explained later), the second column is the number of samples used in regression, and *a* & *b* columns are derived parameters for use in Eq. (1). Some related statistics are also reported here: S_b is the standard deviation for *b*, a measure of its confidence interval; *F* is the F-test for goodness of fit; ρ is the linear correlation coefficient; and σ is the standard deviation of residual. Together, these last four parameters measure different aspects of goodness of fit.

After the initial regression, a second regression is performed, but those samples whose residual is larger than twice of σ are not admitted. This process is then repeated, and the results from the first few iterations are summarized in Table 1 and the corresponding scatter plots were presented in Fig. 1. The last column of Table 1 is the ratio of current sample size to the sample size of previous iteration.

Table 1: Results of recursive regression for 11 μm channel inter-calibration on NOAA-14 and GMS-5

Iter.	*N*	*a*	*b*	S_b	*F*	ρ	σ	*Ratio*
1	157308	-2.45	1.0094	.00138	0.5584	.8794	3.81	-
2	156024	-4.43	1.0171	.00030	11.418	.9932	0.83	0.992
3	149635	-3.98	1.0159	.00022	20.771	.9964	0.56	0.959
4	141337	-3.63	1.0148	.00020	26.154	.9973	0.47	0.945
5	135028	-3.38	1.0140	.00019	29.665	.9977	0.43	0.955
6	130602	-3.28	1.0136	.00018	33.362	.9980	0.41	0.967

To justify the application of recursive regression, one notes that Table 1 shows consistently high value of linear correlation coefficient throughout iterations, thus Eq. (1) is appropriate. Visual

inspection of Fig. 1(a) can be deceptive, as it shows a lot of outliers. This is because most of the well-behaved data are overlaid on top of each other. Indeed, the last column of Table 1 shows that 99.2% of samples in Iteration 1 were used in Iteration 2, yet comparison of the upper two panels of Fig. 1 shows that most obvious outliers have been removed from the second regression. It is thus concluded that recursive regression is applicable here.

The remaining question is when the recursion should stop. Subjectively, it appears that the second regression was necessary, as it removed most obvious outliers, changed regression coefficients *a* & *b* significantly, and improved fitting dramatically. After the third or arguably the second regression, however, all parameters changed only incrementally and the scatter plots remained basically the same. So the recursion should stop after the third or second regression.

An objective criterion might be derived from the last column of Table 1. Without compelling evidence otherwise, the residual error of fit should be normally distributed after outliers are removed. One characteristics of normally distributed population is that approximately 4.6% of all samples lie outside of two sigma from mean. Table 1 shows that in the original dataset, only 0.8% of samples have residual error larger than two sigma, which concurs that the residual error was not normally distributed then (and those 0.8% of samples are indeed outliers). After those outliers were removed and regression was repeated, about 4.1% of sample residuals are larger than two sigma, which means that the residual error is now approximately normally distributed, or the outliers have become insignificant in the sample. So the recursion could stop after the second regression.

In this study, the end of recursion is decided somehow subjectively, using objective criterion as an aid. In this case, for example, it was felt that outliers still exist after the second regression [ref. upper right panel of Fig. 1], although they may be insignificant because of the very large number of "normal" samples. On the other hand, removing these outliers (and together with a few thousands of potentially "normal" samples) should not hurt. So the results from the third regression were taken as final results.

4. RESULTS AND DISCUSSION

Recursive regression was performed for all HIRS/GEO co-located measurements, results were summarized in Table 2 and plotted in Fig. 3 & 4. Table 2 is similar to Table 1, except rows are final results for different inter-calibrations, instead of intermediate results for different iterations. Also, the last column of Table 1 becomes irrelevant in Table 2.

The *N* column in Table 2 suggests that sample size is sufficiently large for all regressions, including the HIRS/METEOSAT-5 inter-calibration that started late. Figs. 1-4 show that the samples are reasonably well distributed over the range of expected T_b's. Therefore the statistical inference derived from these data sets should be credible.

The slopes of regressions (*b* column) are generally non-unity, though some are very close to unity. Their standard deviations (S_b column) indicate that the *b* values are accurate to the third or higher digit after the decimal point. These non-unity *b* values agree with intuition gained from inspection of Figs. 1-4.

The columns of S_b, *F*, and ρ are measures of goodness of fit, and these parameters are internally consistent. The fit is generally better for GOES, worse for GMS-5, and worst for METEOSAT-5. This is consistent with the sample size.

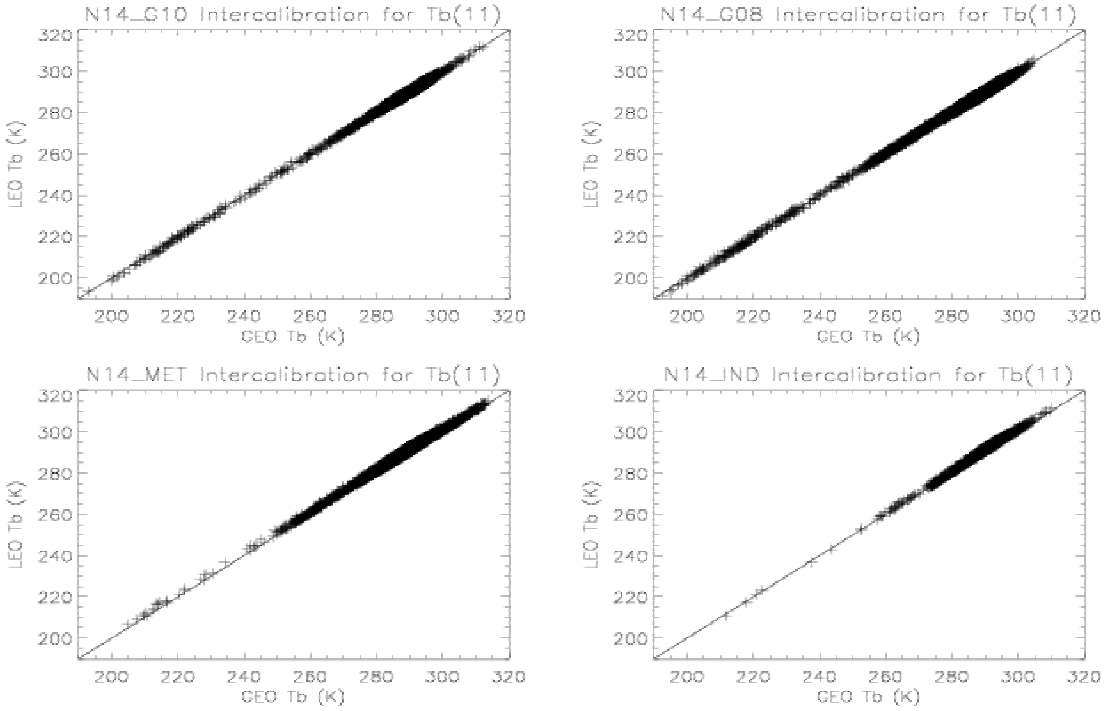

Figure 3: Scatter diagrams of the co-located NOAA-14 HIRS and GOES-10 (upper-left), GOES-8 (upper right), METEOSAT-7 (lower left), and METEOSAT-5 (lower right) for the 11 μm infrared window channel. These are the data for final iteration, or equivalently the "panel (c)" in Figs. 1 & 2. The diagonal lines in these plots show that a linear regression based on these data may have non-unit slope, as well as non-zero intercept. Only 1/10 of randomly selected data were plotted in each diagram.

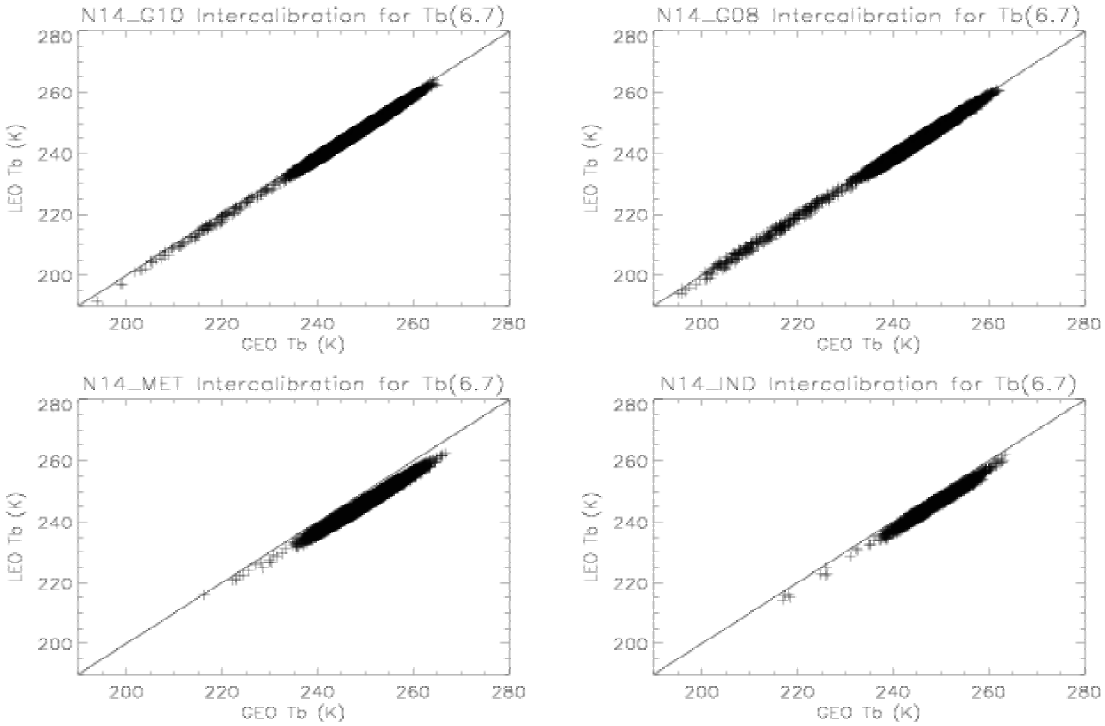

Figure 4: As Fig. 3, but for the 6.7 μm water vapor channel.

Table 2: Regression coefficients for deriving NOAA-14 HIRS $T_b(6.7)$ and $T_b(11)$ using various geostationary satellites.

Sat.	*N*	*a*	*b*	S_b (x10^4)	F (x10^6)	ρ	σ
GMS 6.7	148253	+1.93	0.9662	4.x	4.98	0.9854	0.92
GMS 11	149635	-3.98	1.0159	2.2	20.77	0.9964	0.56
G10 6.7	296765	-3.22	1.0073	1.6	40.37	0.9963	0.49
G10 11	293657	-0.12	1.0005	1.3	54.63	0.9973	0.48
G08 6.7	359351	-0.71	0.9983	1.7	33.38	0.9947	0.57
G08 11	350245	+2.60	0.9918	1.2	67.40	0.9974	0.62
M07 6.7	199892	+5.99	0.9608	3.1	9.92	0.9901	0.68
M07 11	196389	-0.58	1.0076	2.4	18.26	0.9947	0.71
M05 6.7	34381	-0.33	0.9899	8.5	1.37	0.9876	0.75
M05 11	33984	-10.59	1.0415	5.0	4.30	0.9961	0.66

Within each satellite, the fit for the infrared window channel is better than that for the water vapor channel. A better fit does not necessarily mean a better instrument unless the NOAA-14 HIRS measurements are regarded as "perfect". From a practical point of view, however, a better fit does imply better predictability of HIRS T_b using GEO T_b. The last column (σ) is a measure how well, in absolute sense (error in degree K), one can estimate HIRS T_b from GEO T_b, but it is not necessarily a measure of goodness of fit. Look at METEOSAT-7, for example. The root-mean-square-error (RMSE) is 0.68K for the water vapor channel and 0.71K for the infrared window channel. Since the mean brightness temperatures for these two channels are 246°K and 290°K, respectively, the relative error is actually smaller for the infrared window channel, thus a better fit.

5. SUMMARY

A group of scientists envisioned the potential the international constellation of geostationary meteorological satellites has to offer to monitor rapidly changing phenomena in the tropical and subtropical regions, such as the life cycle of convection or the diurnal variation of various parameters. To realize that potential, one has to inter-calibrate the measurements made by these satellites. In this report of work in progress, some methods and issues of inter-satellite calibration were considered, and a method using a polar orbiting satellite to provide a common ground was chosen. Detailed procedure of data collection and quality control were presented, including the application of recursive regression. Preliminary results from this work are being used in compilation of a one-year dataset to study the diurnal variation of upper tropospheric water vapor.

The inter-satellite calibration reported in this note may be further pursued in several ways. As noted before, the inter-calibration can be based on radiative transfer model and sensor's spectral response function. A key issue in that approach is how sensitive is the inter-calibration to the diurnal, seasonal, and geographical variations of atmosphere, in that regard the dataset compiled in this study will be a great asset. A well designed inter-calibration based on radiative transfer theory will not only corroborate or correct the findings so far, but may also provide better inter-calibration for specific conditions such as observing geometry. Similarly, inter-calibration based on overlapping observations may also be valuable.

Acknowledgements: This work started when the author was with the Cooperative Institute for Meteorological Satellite Studies (CIMSS), Space Science and Engineering Center (SSEC), University of Wisconsin-Madison. The scientific vision and financial support provided by Paul Menzel were essential to bring this project to its current stage, as well as constant encouragements from Brian Soden, Michael Weinreb, and Jo Schmetz.