A Comparison of START-08 Avionics and Research *In Situ* Temperatures to Radiosondes

Julie Haggerty[1] and MJ Mahoney[2]

[1]NSF NCAR
[2]JPL

The calibration of the NSF/NCAR GV (NGV) Microwave Temperature Profiler (MTP) is simplified if an in situ temperature measurement made during START-08 is first calibrated against radiosondes launched near the NGV flight track. This calibration is then transferred to the MTP data when it is processed. We have used four different tests to show that the raw ATHR2 temperature is in agreement with radiosondes. This note discusses the tests we performed to come to this conclusion.

1 Introduction

Because radiosondes (RAOBs) provide well-characterized temperature profiles of the Earth's atmosphere and because the START-08 research flights passed near many RAOB launch sites, RAOBs were used to calibrate the Microwave Temperature Profiler (MTP) measurements made during START-08. The JPL MTP group has used this calibration technique for many years with excellent results. For example, on NASA DC-8 field campaigns over a period of several years we found that the DADS/ICATS outside air temperature measurement had a consistent warm bias of ~0.8 K [1]. No one was responsible for this measurement from the Navigation Data Recorder, so the calibration was never changed. We have also been on field campaigns where we have used preliminary PI-led temperature measurements in the calibration process, and come up with a temperature calibration consistent with the final archived PI-led temperature data. For example, we found that during the CRAVE field campaign in Costa Rica aboard the NASA WB-57 that the Dryden Flight Research Center MMS preliminary temperature measurement (when compared to radiosondes) had a warm bias of +0.44 K (Tmms = Traob + (0.44 ± 0.22) K for 17 comparisons. When the final MMS data was submitted to the CRAVE ESPO archive, the PI stated that "Tstatic is lowered by some 0.4 to 0.5 K compared to the preliminary data", which was exactly in agreement with the MTP assessment of the preliminary MMS data [2].

Performing these outside air temperature (OAT) comparisons against radiosondes is a tedious process because you have to check for pressure altitude changes, pitch and roll changes, regions with large lapse rate (i.e., avoid tropospheric data), temperature variability (check soundings before and after flyby, and avoid comparisons near the tropopause.), etc. Initially these comparisons were done by hand and a single comparison would take a day to perform! When I (MJ) began doing these comparisons, I visited the originator (Bruce Gary, MTP PI emeritus) of the technique in Arizona and learned all the steps involved. I then wrote code to perform these steps, and ultimately automated all the steps. That said the setup process is still tedious – but worth it.

After a field campaign a program is run to quickly identify all the potential radiosonde comparison launch sites that were flown by. The program records on a list the time of closest approach and distance for each site. Using this list, radiosondes launched 12 and 24 hours before and after a flyby are compared for variability in temperature structure. Those with variability are generally removed from the list. Next, since radiosondes seldom ascend beyond 30-35 km, an estimate of the lapse rate above the burst altitude to facilitate retrieval coefficient calculations, and this information is added to the list of potential radiosonde comparisons. For START-08 we had nearly 100 potential comparisons – that's why this setup procedure is so tedious. However, once the setup is done, a hundred comparisons

can be done in seconds while taking into account adjustable editing criteria such as allowable altitude and attitude changes, or the averaging of MTP profiles to reduce their noise.
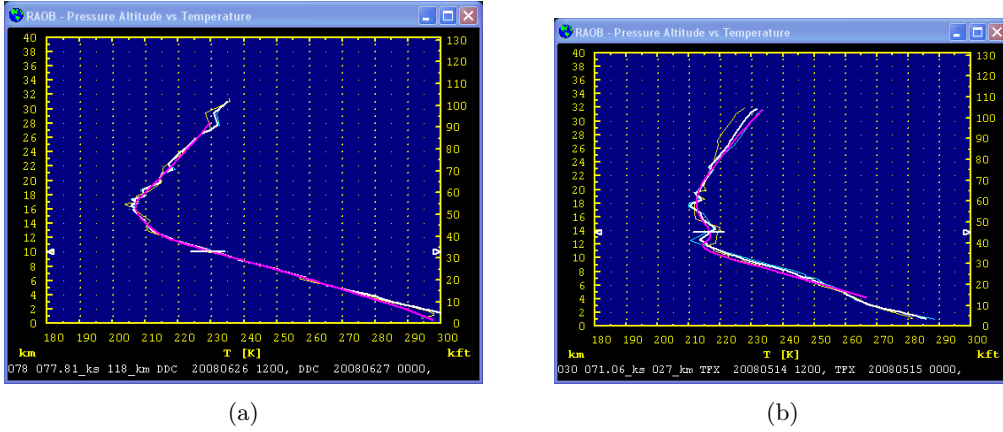

(a) (b)

Figure 1: Images showing superimposed temperature profiles for the RAOB before NGV flyby (yellow), for the RAOB after NGV flyby (cyan), the temporally interpolated RAOB at NGV closest approach (white), and the uncalibrated MTP retrieved temperature profile at closest approach (magenta). (a) shows no temporal variability near flight level, while (b) shows substantial temporal variability, especially near the tropopause. Such comparisons are to be avoided.

As shown in Figure 1, images showing the before and after soundings, their temporal interpolation to the flyby time, and the average MTP profile are generated for each comparison, so that they can be examined visually for temperature structure variability, high lapse rate and closeness to the tropopause.

Although the MTP and RAOB comparisons could be done directly as suggested above, the MTP gain calibration (which ultimately determines the MTP temperature calibration) is simplified if an in situ outside air temperature (**Tnav**) measurement is available from the aircraft. Following this approach, the MTP gain is adjusted to get the best possible agreement between the MTP temperature at flight level (**Tmtp**) and **Toat**. Both **Tmtp** and **Tnav** are compared to the radiosonde temperature at flight level (**Traob**) to determine what corrections, if any, are needed to make **Tmtp** and **Tnav** agree with **Traob**. We also check to see if the corrections have a pressure altitude dependence.

Table 1: Temperature measurements that are available for START-08

NCAR Temperature Parameter	Parameter Name	Comment
AT_A	Tavi	Ambient Temperature from the Air Data Computer (ADC)
ATFR	Tres	Ambient Temperature - Fast Response (Unheated Probe, Left, Research Temperature)
ATHR1	Tres2	Ambient Temperature - Heated Probe Right #1 (Research Temperature)
ATHR2	Tres3	Ambient Temperature - Heated Probe Right #2 (Research Temperature)
-	Tmtp	MTP-derived outside air temperature
-	Traob	RAOB-derived outside air temperature

2 Available *In Situ* Temperatures

There are four separate temperature probes on the NGV; these are the first four entries in Table 1. One (**Tavi**) is from the Air Data Computer (ADC), and the other three are "research" temperatures: one is from an unheated probe on the left side of the aircraft (Tres=AT_A) and the other two are

from heated probes on the right side of the aircraft (Tres2=ATHR1 and Tres3=ATHR2). We will use the generic notation **Tnav** for any of these temperatures. During START-08 the temperature in the 1 Hz IWG data was **Tres** (ATFR). When performing the temperature calibration, we are also interested in the pressure altitude because it sometimes happens that the temperature calibration is pressure altitude dependent. The pressure altitude (**Zavi**) used by the ADC is NCAR parameter **PALT_A**; the pressure altitude (Zres) associated with the "research" temperatures can be derived from the Corrected Static Pressure, Fuselage (NCAR parameter **PSFC**) using standard equations [3].

2.1   Temperature Calibration Techniques

Four different techniques were used to assess the accuracy of the temperatures measured on the NGV:

1. Compare the four aircraft temperatures to RAOB flight level temperatures at flight level.

2. Examine the behaviour of **Tnav-Traob** as a function of a temperature threshold.

3. Examine the behaviour of **Tnav-Traob** as a function of the range (i.e., distance) from the RAOB launch sites.

4. Examine the behaviour of **Tnav-Traob** when measurements near the tropopause are excluded (to avoid temperature structure variability).

2.1.1   Technique 1: Compare Tnav to Traob

This technique is basically the same as the one we used to study the accuracy of **Tnav** during the NGV T-Rex field campaign[4]. Although the MTP did not fly on the NGV during T-Rex, the extensive MTP data analysis software developed to do radiosonde comparisons could be readily modified to facilitate comparisons. We did these comparisons because there were known inconsistencies between the Tnav temperature measurements. Comparisons with radiosondes allow a completely independent assessment of the measurement accuracy. And we had very good statistics because not only could we do comparisons on the transit flights from Jeffco to near Independence, CA, but Leeds University also launched frequent soundings from Independence. For the T-Rex RAOB comparisons, we only compared **Tavi** and **Tres** to **Traob**. For START-08 we compared all four NGV temperatures (Tavi, Tres, Tres2, and Tres3) to **Traob**. These temperatures, as well as many other parameters, were written every second into ASCII text files that the data analysis software could read. Without any editing (see following techniques), we found the following very robust results for N = 61 comparisons:

Table 2: Temperature biases found for the four NGV temperature measurements compared to nearby radiosondes.

$$Tavi - Traob = 1.74 \pm 0.21K$$

$$Tres - Traob = 0.58 \pm\ 0.19K$$

$$Tres2 - Traob = 0.43 \pm 0.20K$$

$$Tres3 - Traob = -0.03 \pm 0.19K$$

While the population standard deviation for each of these four comparisons was nearly 2 K in each case, the formal error on the temperature bias is only ~0.20 K. Note that the avionics temperature **Tavi** has the poorest accuracy. Within the errors, this result is identical to what we found during T-Rex[4]. However, during T-Rex we found that the temperature bias for **Tavi** was pressure altitude dependent, being nearly 2 K at 12.5 km and about 1 K at 8.5 K. Since the average altitude for the START-08 comparisons was 12.0 km, the 1.74 K bias compared to **Traob** is completely consistent with the T-Rex results at 12.5 km. On the other hand, the T-Rex research temperature (**Tres = ATFR**) showed a temperature error of 2.5-3.0 K, while during START-08 the error had dropped to 0.58 K. It

is unclear what this means, since we don't know whether recovery factors and other corrections to the "research" temperatures were changed between the two campaigns.

When MTP temperature profiles are compared to RAOBs, the comparisons are made not just at flight level, but at all altitudes, since this gives us a handle on how good the retrievals are away from flight level. (The MTP retrieval accuracy depends on the distance from the aircraft because more information is available near flight level than further away.) Based on the results shown Table 2, we decided to substitute the **Tres3** (=ATHR2) temperature and **Zres** pressure altitude (based on NCAR parameter **PSFC**) into the IWG line recorded by the MTP during flight, since it shows essentially no bias with respect to **Traob**.

Using Technique 1 for all of the possible **Tnav** and **Traob** temperature comparisons, we obtained **Tnav-Traob** biases with a formal error of < 0.21 K (see Table 2). It is insightful to investigate the impact of eliminating some of the comparisons because the temperature errors are too large, because the RAOB launch site was too far from the airplane, or because the comparisons were made near the tropopause with known temperature variability (based on comparing before and after flyby radiosondes). The next three techniques explore these options.

2.1.2 Technique 2: Behaviour under Temperature Threshold Changes for Accepting Comparisons

In order to assess the accuracy of the aircraft temperature measurement (Tnav = Tres3 = ATHR2) compared to radiosondes, objective criteria have to be developed for accepting or rejecting a particular comparison. Figure 2 shows one such criteria. Here we ask the question: If temperature comparisons are excluded because the error exceeds some temperature threshold, does this tell us anything useful? If data are excluded because of a large temperature threshold, you might argue that this is justified because more distant comparisons might result in larger errors. (Figure 1 in Reference [5] that this is actually the case by looking at the temperature structure function at different flight altitudes.)
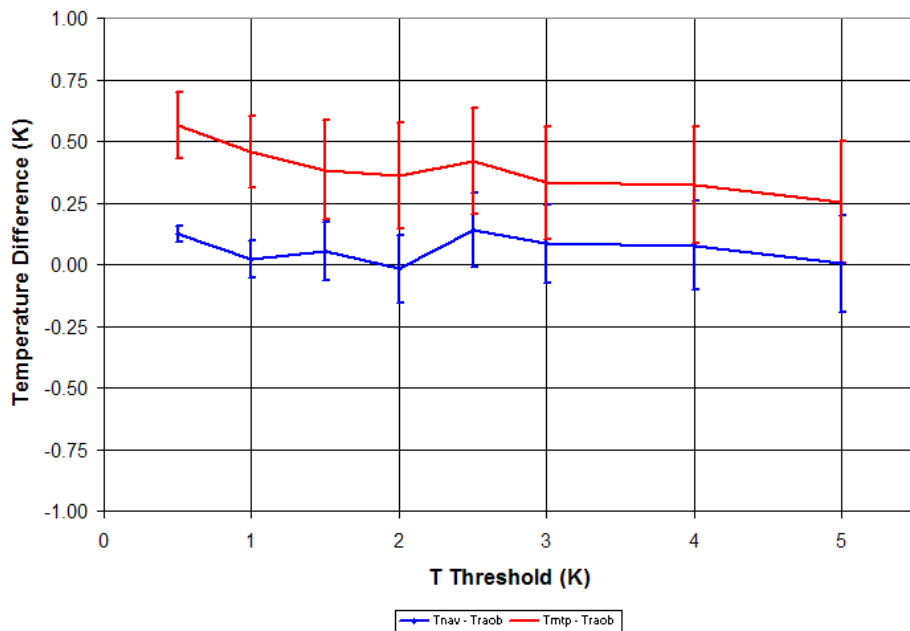

Figure 2: Tnav-Traob as a function of temperature threshold.

Conversely, if data are excluded because of a small temperature threshold, you might argue that this is "cooking the books." We are sympathetic to this concern! And will say more in a moment. So how can this process be made objective? Throwing out errors greater than 4 or 5 K makes some sense, but throwing out small errors is harder to justify. However, suppose that there is a real temperature difference (**Tnav-Traob**) of 2 K. In that case, there would likely be very few comparisons < 1 K or < 0.5 K, and the error on the bias would be large. On the other hand, if the real bias is 2 K, there would likely be more comparisons at this bias, therefore better statistics, and the error on the bias

would be small. What normally happens is that when you get down to throwing out small errors, the error bars on the errors become much larger (unless the real bias is indeed small) due to the diminishing number of comparisons. Markedly, in Figure 2 this did not happen, and in fact the error on the bias becomes smallest (0.03 K) for the < 0.5 k threshold. This is yet further evidence that **Tres3-Traob** is close to 0.0 K. If we accept the comparisons between the 0.5 K and 3 K thresholds, we find that **Tres3-Traob** = 0.07 ± 0.11 K. Another remarkable result in Figure 2 is that the **Tmtp-Traob** bias is also fairly constant and tracks **Tnav-Traob**. This is good looking data!

2.1.3 Technique 3: Behaviour under Range Threshold Changes

Under Technique 2 we mentioned that larger range might cause larger temperature errors, and reference [5] concurs with this. So another approach is to remove comparisons based on range. This is shown in Figure 3. While Technique 2 showed a slight warm bias for **Tnav-Traob**, this technique shows a slight cool bias. In addition the errors on the biases are larger than for Technique 2. Using all the measurements we find: **Tnav-Traob** = -0.09 ± 0.23 K. As was the case for Technique 2, **Tmtp-Traob** is warmer (and by about the same amount) than **Tnav-Traob** except at the shortest range.
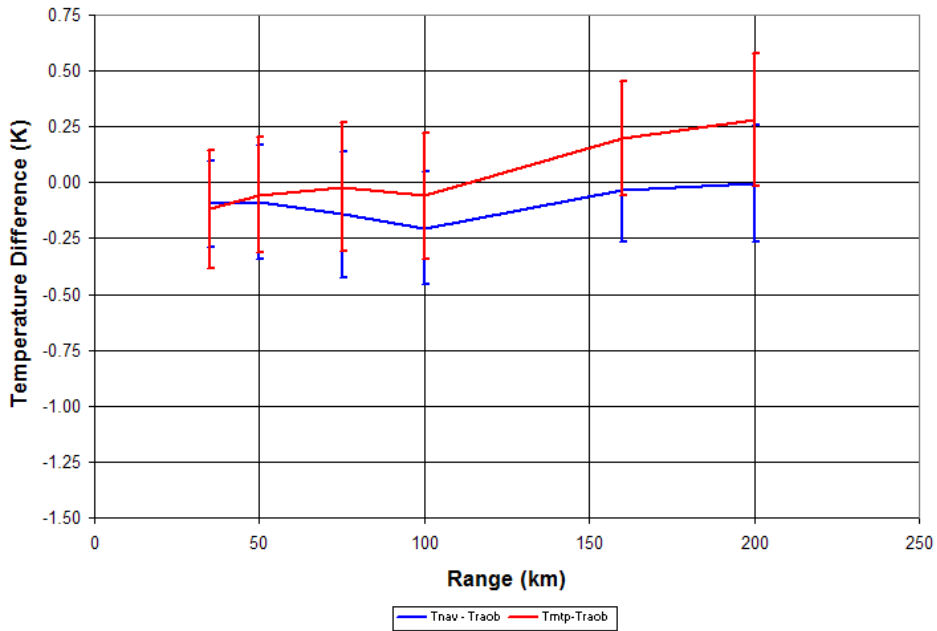

Figure 3: Tnav-Traob as a function of range threshold.

2.1.4 Technique 4: Eliminate Comparisons Made Near the Tropopause

The final technique involves examining the images of the individual RAOB comparisons and removing comparisons for which the before and after soundings show obvious temperature structure variability – especially near the tropopause. This was illustrated by Figure 1b. When we did this we were left with N = 39 comparisons, one of which had a temperature difference of -7.22 K. Using Peirce's Criterion for outliers, this comparison was removed to give **Tres3-Traob** = -0.17 ± 0.20 K (N=38)[6].

3 Conclusions

Using the four techniques described above, we found the four results shown below in Table 3. Within the errors for the bias determination, we can safely assume that **Tres3-Traob** (ATHR2-Traob) = 0.0 ± 0.2 K. The WMO reports the random errors in RAOB measurements are < 0.2 K at night and < 0.3 K during the daytime in the troposphere and stratosphere[7]. This is in agreement with finding, and is the typical accuracy that we find when doing RAOB comparisons. When there are not many radiosondes to compare to, the accuracy might degrade to 0.3 K. These comparisons provide a valuable assessment

Table 3: The results of using four techniques to determine **Tres3-Traob**.

Technique 1	$Tres3 - Traob = -0.03 \pm 0.19K$
Technique 2	$Tres3 - Traob = +0.07 \pm 0.11K$
Technique 3	$Tres3 - Traob = -0.09 \pm 0.23K$
Technique 4	$Tres3 - Traob = -0.17 \pm 0.20K$

of the accuracy of aircraft in situ temperature measurements. Although tedious, this process is often much more direct than trying to determine all the factors that might affect an aircraft temperature measurement such as the recovery factor, electronics thermal issues, leaks in pressure lines, etc. As shown in Table 2, it is difficult to get four different temperature measurements on the same aircraft during START-08 to agree with one another.

References

[1] M. J. Mahoney, "Camex-4 temperature intercomparisons," November 2002. [Online]. Available: https://wayback.archive-it.org/3544/20200922020347/https://archive.eol.ucar.edu/raf/instruments/MTP/www/missions/camex4/Science/CAMEX4_Temperature_Comparisons.html

[2] ——, "Microwave temperature profiler (mtp) measurements for crave," August 2006. [Online]. Available: https://wayback.archive-it.org/3544/20200922022018/https://archive.eol.ucar.edu/raf/instruments/MTP/www/missions/crave/science/CRAVE_InSitu_T_Comparisons.html

[3] NOAA, NASA, and U. A. Force, *U.S. Standard Atmosphere, 1976.* NOAA, 1976.

[4] M. J. Mahoney and J. Haggerty, "A comparison of t-rex avionics and research in situ temperatures to radiosondes," December 2007. [Online]. Available: https://doi.org/10.26023/41TZ-HZJH-EZ0Y

[5] M. J. Mahoney, "Steps in mtp post-campaign data analysis." [Online]. Available: https://wayback.archive-it.org/3544/20200922015435/https://archive.eol.ucar.edu/raf/instruments/MTP/www/notes/software/Step_07.html

[6] S. Ross, "Peirce's criterion for the elimination of suspect experimental data," *Journal of Engineering Technology*, Fall 2003. [Online]. Available: https://www.eol.ucar.edu/sites/default/files/files_live/private/piercescriterion.pdf

[7] R. Smout, T. Oakley, B. Pathack, and S. Krunosenko, "The wmo intercomparison of radiosonde systems: final report," WMO, Tech. Rep. WMO/TD-No. 1303, 2006.