SMA memo 149

Optimal Smoothing Algorithm for Water Vapor Monitor Data

Martina C. Wiedner, June 2002

1 Abstract

The Sub-Millimeter Array is equipped with two 183 GHz Water Vapor Monitors (WVM) (currently mounted on antenna 4 and 5), which perform phase correction. Phase correction is necessary at high frequencies or long baselines to correct phase fluctuations introduced by differential changes in the amount of atmospheric water vapor above the individual antennas. The refractive index and therefore the optical path depends on the column of water vapor. Uncorrected path length errors produce phase errors in interferometers. The radiometers measure and compute the differences in the amount of precipitable water vapor above each antenna, and calculate the phase shift this difference in vapor will introduce. The phase shifts deduced from the WVM measurements will be used to correct the interferometric phase. Obviously, the success of the phase correction will largely depend on the accuracy and the noise of the water vapor measurements. This memo investigates data reduction methods that minimize the noise in the WVM data. Five different smoothing algorithms are tested and the best method is selected.

2 Introduction

Ideally, the voltage detected by the radiometer would be a direct measure of the emitted brightness temperature. However, two types of instrumental noise affecting the measurement:

1. Thermal noise due to the receiver temperature
2. Gain variations in the mixer and the amplifiers.

(For measurements and a detailed discussion of the instrumental noise of the radiometers see SMA memo 148 "Noise Characterization and Allan Variances of Water Vapor Monitors".) As we are taking differences in the amount of precipitable water vapor (PWV) we need an extremely high accuracy of 10^{-3} to 10^{-4} in each individual measurement. This means we need a low (or constant) receiver temperature and small gain variations. Since the WVMs are not cooled, the receiver temperature is high at around 2000 K, but given the bandwidth of 500 MHz, the root mean square (rms) thermal noise is only 0.5 K during the typical 0.4 sec integration. However, given this high a receiver temperature the main constraint on the accuracy are the gain fluctuations in the mixer and the amplifiers, which are estimated to be about 0.05% rms on second time scales and an additional drift of about 2% per hour, which corresponds to 1 K rms and 40 K per hour drift using a 2000 K receiver temperature.

Since the gain stability is marginal for our accuracy requirements, we calibrate the WVM with a hot and a warm calibration load every second. (In a calibration cycle of 1 s the instruments spends 0.2 s on each load and 0.4 s on the sky, which was replaced by a liquid nitrogen load in the measurements described below. The remaining 0.2 s of data are rejected while the calibration mirror moves between the loads.) Heating circuits ensure that the calibration load temperatures of the WVMs are constant, which, according to the temperature sensors on the loads, have an rms of 20 mK, negligible in our experiment. If we integrate for even shorter times, the gain changes can be tracked on shorter time scales, but the error due to thermal noise increases. From the Allan variance measurements (SMA memo 148) we learned that the thermal noise dominates the noise budget over averaging times of 4 s or less, for longer averaging times the gain variations become noticeable. Therefore averaging over 4 s was used in most of the smoothing methods described below.

Interestingly, not only the averaging time, but also where the smoothing is applied has a considerable impact on the noise: E.g., we can average the voltages on the hot and cold load over 4 s and calculate the averaged gain and receiver temperature or, inversely, we can calculate gain and receiver temperature for every 1 s data point and then average gain and receiver temperature. We will see that this second method reduces the noise substantially compared to the first. Furthermore, we use different window functions: e.g., we can take the arithmetic average, or multiply the data by a weight, which decreases as the time difference between the point in question and the point to be averaged increases.

In the next sections the experimental set-up and the smoothing methods are described. The results using the different smoothing methods are discussed in section 5 and a summary follows in section 6.

3 Experimental details

On 16th July 2000 (early morning Hawaiian time) we took data samples lasting just under 1 hour with the WVM originally built for the CSO. To obtain a constant cold load temperature, a Styrofoam bucket filled with liquid nitrogen was tied in front of the WVM at the CSO. There was 1 large piece of eccosorb in the bucket, which behaves like a black body at millimeter frequencies. The eccosorb was tilted at about 40 deg with respect to the WVM beam. The humidity was high and a thin layer of frost formed on the outside of the bucket over the integration time of half an hour, starting from the bottom, where the WVM was unlikely to see it, but climbing up. The nitrogen level in the bucket decreased, but probably the WVM beam looked at the submerged (at least 90%) part of the absorber. According to the literature (Gerthsen, 2000) liquid nitrogen boils at 73.5 K at 625 mbar pressure. The walls of the Styrofoam bucket, through which the beam passes, are assumed to be 100% transparent (which is probably not entirely true). The temperature (T_{hot}) of the hot calibration load was held at 373 K, the warm calibration load temperature (T_{warm}) at 308 K. However, a temperature 372 K was used for T_{hot} to be consistent with calculations done by Richard Hills using the JCMT's smoothing method. (A 1 K higher hot load temperature will not have any effect on the rms in T_{N2} of the data and nor on our conclusions, but simply shift the calculated average N2 temperature from 81 K to 77 K.)

4 Smoothing methods

Every second the radiometers record a voltage when looking at the hot load (hot(t)), one when looking at the warm load (warm(t)) and a third when looking at the liquid nitrogen load (N2(t)). In addition the brightness temperatures of the hot load (T_{hot}) and the warm load (T_{warm}) are known and assumed constant with time. These 5 numbers are used to compute the brightness temperature of the liquid nitrogen (T_{N2}).

The gain (in units of $Kelvin/voltage$) is defined as

$$gain = (T_{hot} - T_{warm})/(hot - warm). \tag{1}$$

The receiver temperature (T_{rx}) is calculated from the averaged hot and warm load reading:

$$T_{rx} = (hot + warm)/2 * gain - (T_{hot} + T_{warm})/2. \tag{2}$$

Therefore, the liquid nitrogen brightness temperature (T_{N2}) seen by the radiometer is

$$T_{N2} = N2(t) * gain(t) - T_{rx}(t). \tag{3}$$

Assuming that the liquid nitrogen as well as the calibration loads have constant physical temperatures, the radiometer should see a constant brightness temperature of the liquid nitrogen and the reduction method that gives the liquid nitrogen brightness temperatures with the minimum root mean square (rms) error is best.

1. **Smoothing of hot and warm load measurements (blue)**

   Hot and warm load voltages are averaged over 4 seconds to calculate T_{N2} (this time scale was obtained by looking at the Allan variance).

2. **Spike removal plus sum and difference of load measurements are averaged (black)**

   First any spikes are removed from the data. Then the sum of the hot and warm load voltages are smoothed with a full width half maximum of 32s and triangular weighting. The difference in voltages (hot-warm) is Hanning smoothed, i.e. three consecutive measurements are added with weights of 0.25, 0.5, 0.25:

$$\begin{aligned} 0.25 * (hot(t-1) - warm(t-1)) \quad &+ \quad 0.5 * (hot(t) - warm(t)) \\ &+ \quad 0.25 * (hot(t+1) - warm(t+1)) \end{aligned} \tag{4}$$

   From the sum and the difference of the voltages the brightness temperature of liquid nitrogen is calculated.

3. **Sum and difference of load measurements are averaged (pink)**

   It is the same method as the previous method, except that no spikes are removed.
   (See on cfa0 home/mwiedner/wvm/software/smoothing1.c)

4. **JCMT method (green)**

   This method as well as (2) and (3) was developed by Richard Hills and is used for the WVM at the James Clark Maxwell Telescope (JCMT) (program: wvmmon.for on JCMT vax [wvm.realtime.new]). It uses "exponential" smoothing where h+w and h-w denote the sum and differences of the hot (hot) and warm load voltages (warm), respectively:

$$h + w = \sum_{i=-30}^{30} [hot(t+i) + warm(t+i)] * (1.0 - 0.6) * (0.6)^{|i|} \tag{5}$$

$$h - w = \sum_{i=-30}^{30} [hot(t+i) - warm(t+i)] * (1.0 - 0.9) * (0.9)^{|i|} \tag{6}$$

   Thirty samples (i.e. appr. 30 sec) to the left and the right of time t are averaged. The difference is smoothed over about 4 times longer time intervals than the sum.

5. **Smoothing gain and T_{rx} (red)**

   This is a slightly modified version of the method suggested by Oliver Lay. The receiver temperature is smoothed exponentially:

$$T_{rx} = \sum_{i=-\infty}^{\infty} \frac{warm[t+i] * T_{hot} - hot[t+i]) * T_{warm}}{(hot[t+i] - warm[t+i])} * exp(-0.002 * i) \tag{7}$$

   h+w is smoothed like in the "Mauna Kea" method. (In the original version a box average of h+w with 10 samples had been taken.) Here the gain is calculated as

$$gain = (2 * T_{rx} + T_{hot} + T_{warm})/(hot + warm) \tag{8}$$

5 Discussion of data

For channels 1, 2 and 3 of the WVM Figs. 1, 2 and 3 show gain, T_{rx} and T_{N2} using the 5 different smoothing methods discussed above. In the last panel T_{N2} is plotted with offsets to facilitate the comparison of the methods. Below the graph the rms of T_{rx} is listed, as well as the rms after subtracting the best fitting straight line (rms-lin). Fig. 4 contains additional data registered by the radiometer: the physical temperatures of the RF-components (Gunn, mixer, first amplifier and PLL), of the IF-components (second amplifier, second down-converter, filter, detectors), of the Hot and of the Warm load. All of these temperatures are measured with temperature sensors. Unfortunately, the data contains many spikes, probably due to computer software. None of these temperature sensor read-outs are used in any of the smoothing calculations.

The conclusions drawn from these experiments are as follows: Method (1) is definitively the worst method.

Methods (2), (3) and (4) show very similar results, suggesting that the spike removal and the weighting (triangular versus exponential) have little influence on the quality of the smoothing.

Method (5), smoothing T_{rx} for a long time (500 s), is best, as it does not have any short term oscillations (of about 2 min). However, the resultant T_{N2} contains a linear slope. It is not clear, whether (i) the temperature gradient is real, e.g. due to the formation of ice on the outside of the N2 bucket, (ii) is due to some instrumental drift within the radiometer or (iii), less likely, is due to the load temperatures changing. However, the first hypothesis seems most likely, since the rise in T_{N2} does not seem well correlated to the shape of the rise in physical temperatures of the RF or IF plate nor to the physical load temperatures as seen in Fig. 4. Hypothesis (i), however, cannot explain easily why the gradient is smaller in channel 2 than in channel 1 and 3. Whatever the origin of this slowly drifting brightness temperature gradient it will hardly affect phase correction (performed every few seconds), and any slow drift can be removed by observing a calibrator ever $\sim$20 minutes.

6 Summary

Method (5), which smoothes the receiver temperature T_{rx} over a long time ($\sim$500 s) and the sum of the hot and warm load temperatures over shorter time ($<$ 30 s), minimizes the instrumental noise and gain variations.

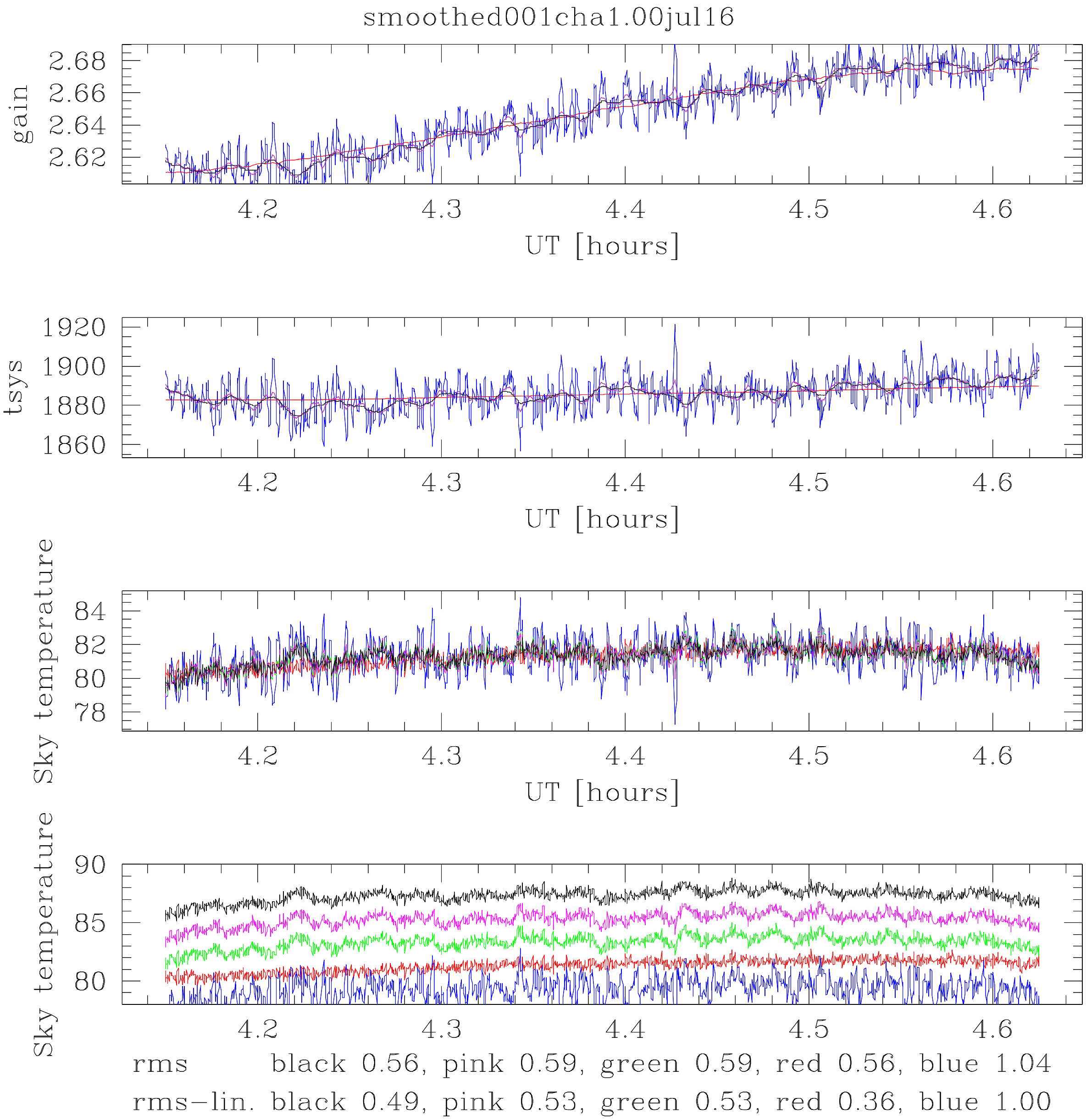

Fig 1. Gain, receiver temperature and brightness temperature measured by channel 1. When in the "sky" position the radiometer looks at a bucket of liquid nitrogen. Blue: method 1), Black: method 2), Pink: method 3), Green: method 4), Red: method 5).

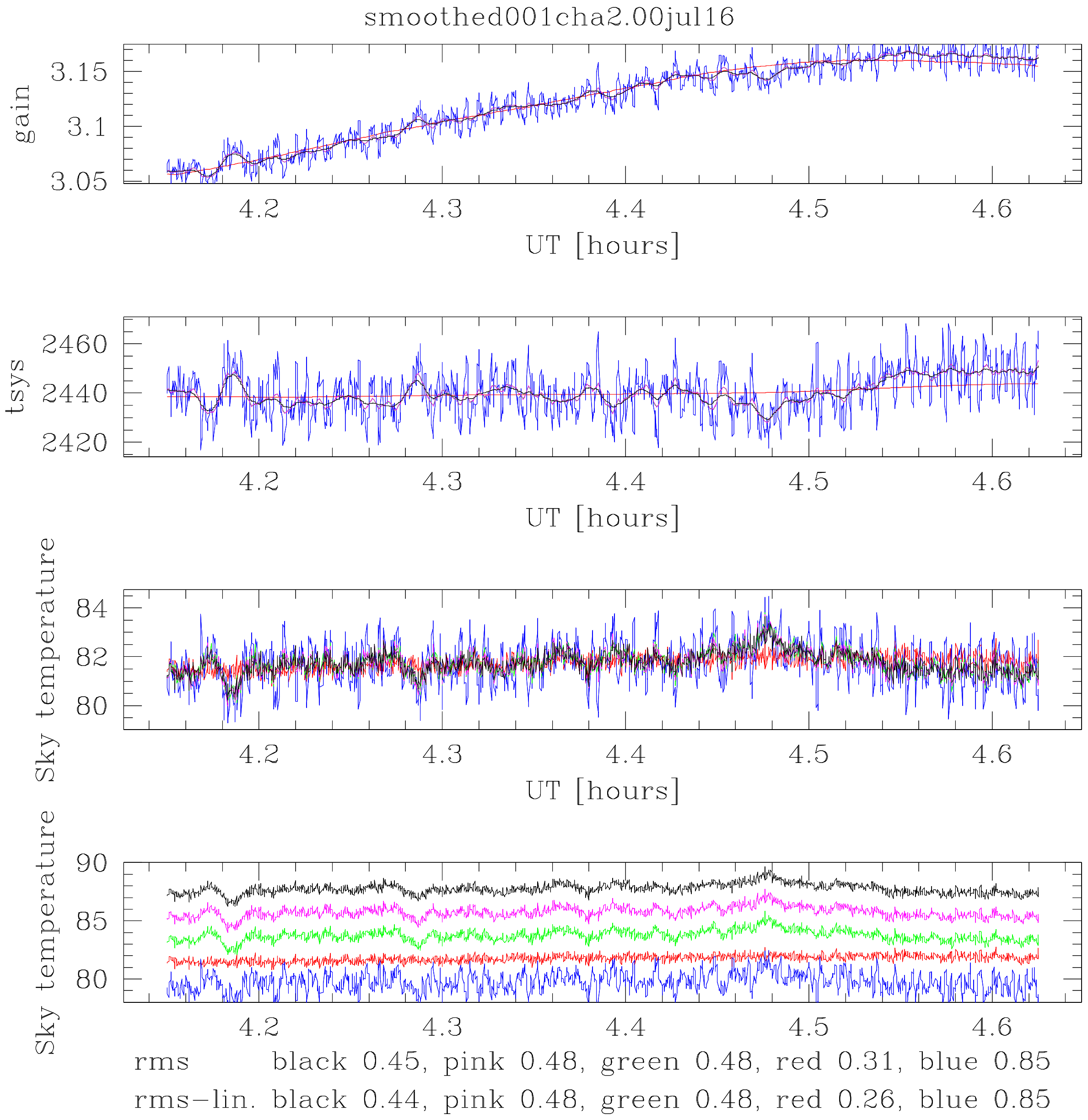

Fig 2. Gain, receiver temperature and brightness temperature measured by channel 2. When in the "sky" position the radiometer looks at a bucket of liquid nitrogen. Blue: method 1), Black: method 2), Pink: method 3), Green: method 4), Red: method 5).

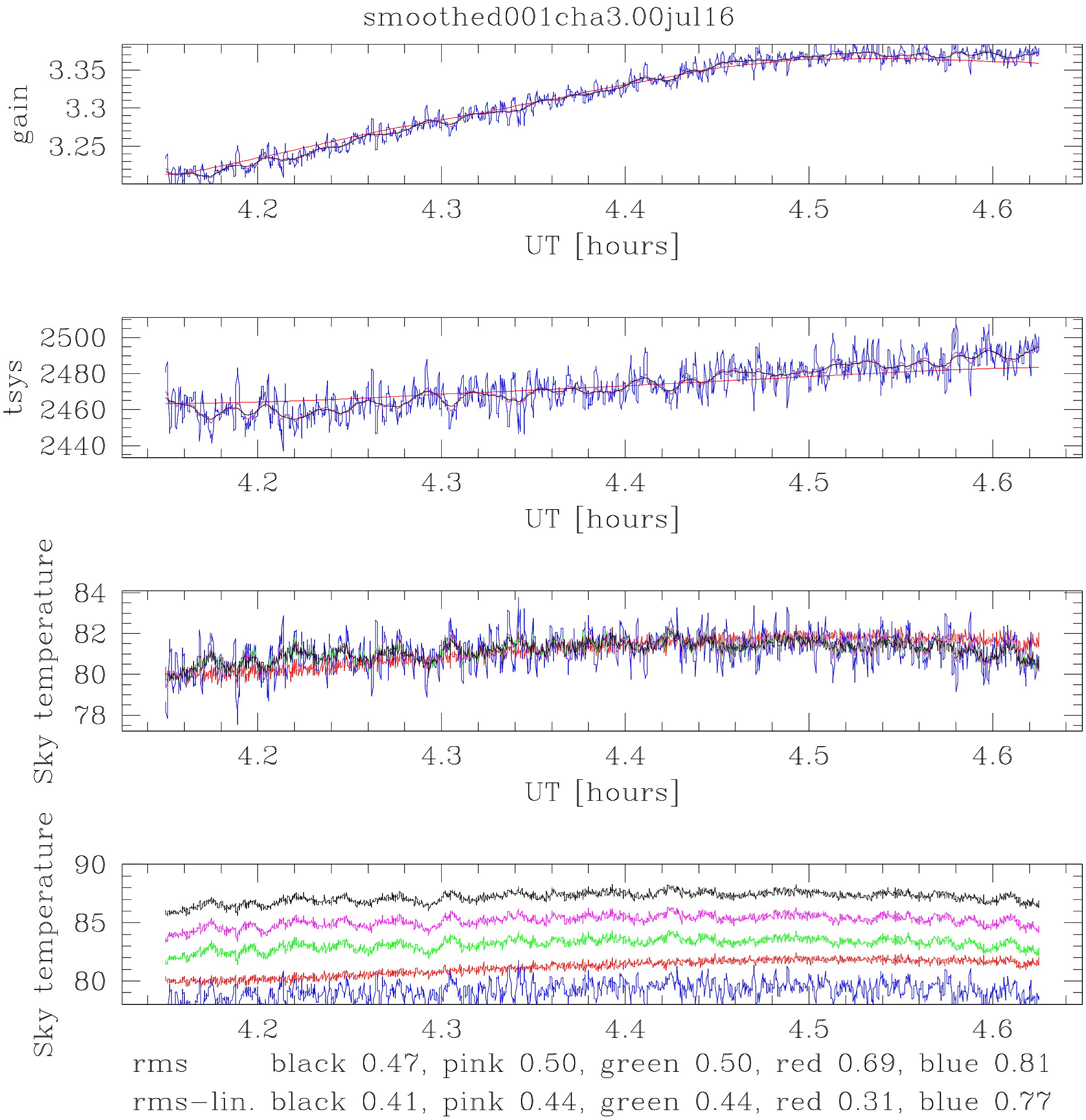

Fig 3. Gain, receiver temperature and sky brightness temperature measured by channel 3. When in the "sky" position the radiometer looks at a bucket of liquid nitrogen. Blue: method 1), Black: method 2), Pink: method 3), Green: method 4), Red: method 5).

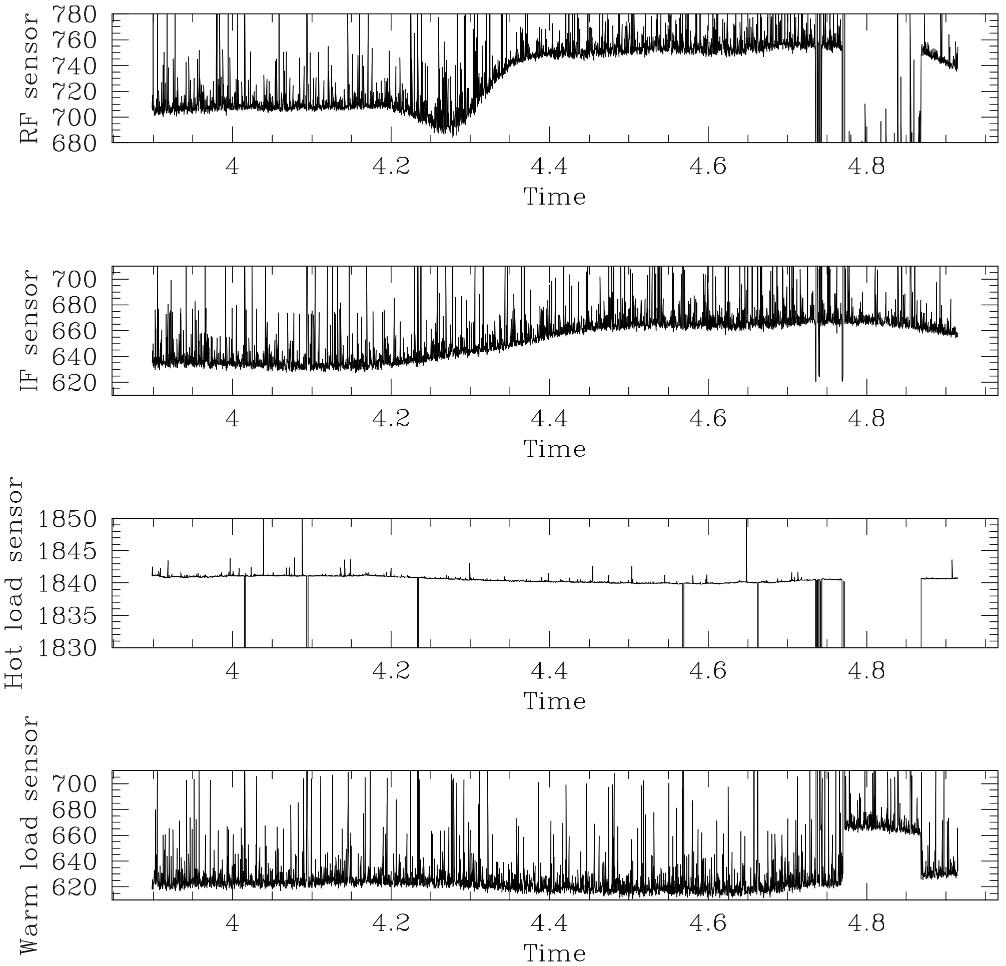

Fig 4. Temperature sensor readings from top to bottom: RF plate, IF plate, hot load, warm load. 20 counts are approximately equivalent to 1 K. The spikes are likely to be artifacts of the counter board. The sudden change in count rate around 4.8 hours UT seems unphysical too, but its origin is unclear. None of these temperature sensor readings are used in any of the smoothing calculations.