SSM/I and SSMIS Quality Control

CSU Technical Report

Wesley Berg and Nereida Rodriguez-Alvarez

July 2013

http://rain.atmos.colostate.edu/FCDR/

TABLE of CONTENTS

LIST of FIGURES

ACRONYMS AND ABBREVIATIONS

Acronym or Abbreviation	**Meaning**
CDR	Climate Data Record
CSU	Colorado State University
FCDR	Fundamental Climate Data Record
RSS	Remote Sensing Systems
SSM/I	Special Sensor Microwave Imager
T_a	Antenna Temperature
T_b	Brightness Temperature

1. Introduction

There are six quality control steps, of which the first corrects certain T_a data, and the subsequent five set quality flags where quality issues are found and sets the corresponding data fields to indicate missing data for those issues classified as serious.

2. Antenna Temperature Correction from Counts Analysis (SSM/I only)

This section describes the main procedure to correct antenna temperatures from counts analysis for SSM/I data.

2.1 Detection of spikes in counts data

2.1.1 Errors in antenna temperatures

By comparing the brightness temperatures (T_b) generated by Remote Sensing Systems (RSS) with the ones generated by the Colorado State University (CSU) some differences were detected. The conversion between T_b and the antenna temperatures (T_a) cannot produce the differences observed. It indicates that RSS is correcting the T_a at a prior stage. The analysis described in this report was performed for each SSM/I channel. Images for 19V are shown as examples, so where T_a and T_b occur in image labels or captions they refer to the 19V channel.

Figure 1 shows a comparison between RSS and CSU mean T_a where the mean is over all points in a scan.

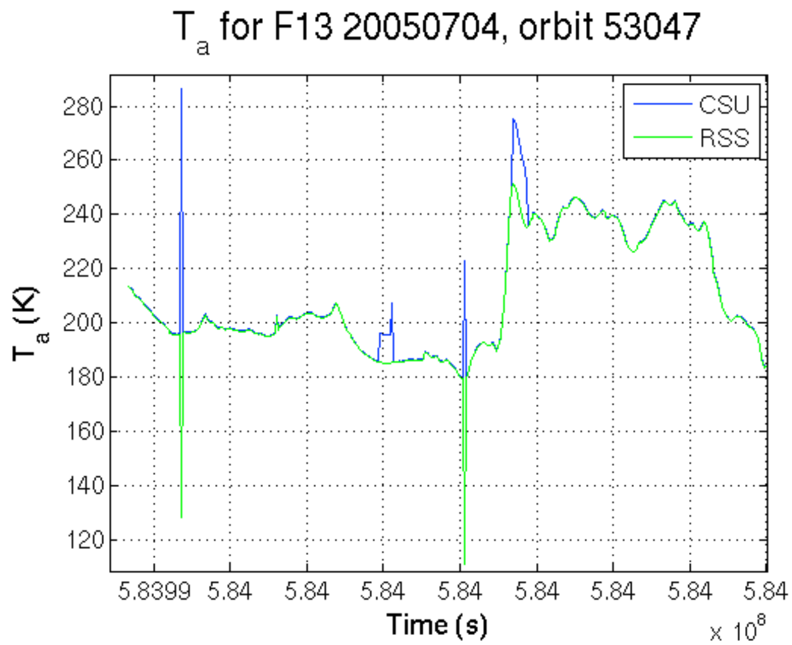


Figure 1: Comparison between RSS mean T_a and CSU mean T_a.

Figure 1 shows some differences between the two mean T_as. RSS mean T_a has a couple of spikes that CSU mean T_a also shows, but CSU mean T_a has two particular events that are not present in RSS mean T_a (Figure 2).

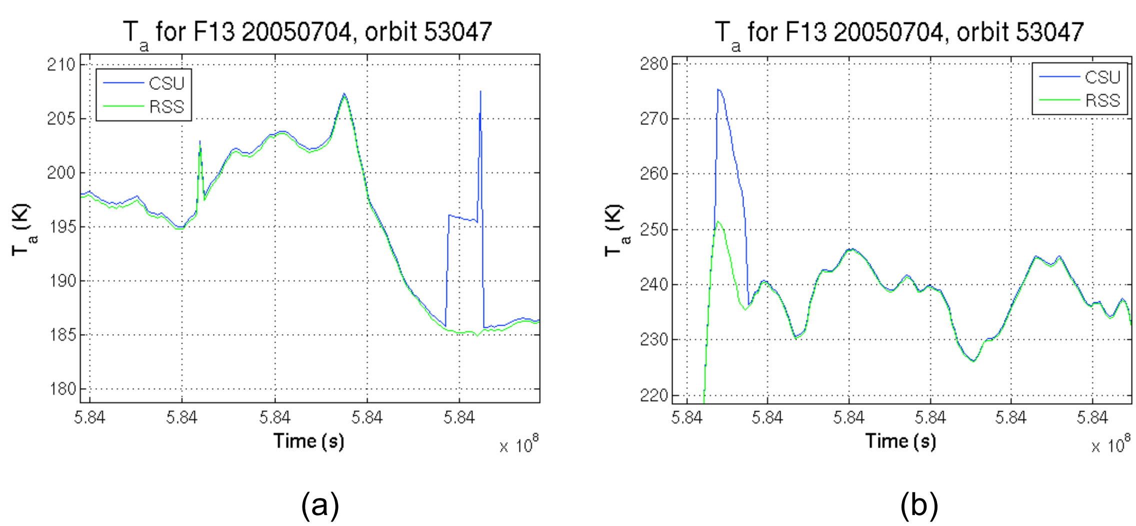


Figure 2: Zoom over the two conflictive areas on Figure1.

Figure 2 shows the zoom over the two erroneous areas in Figure 1. These errors look as jumps during a certain number of scans. In Figure 2a a spike is observed in addition to this temporary jump in the scans.

2.1.2 BASE files analysis

In order to find the focus of the problem the BASE files were analyzed. T_a provided in BASE files are generated following eqn.1, Colton (1999):

$$T_a = \left(\frac{\overline{T}_{WL} - T_C}{\overline{V}_{WL} - \overline{V}_C} \right) \cdot \left(V_a - \overline{V}_C \right) + T_C \tag{1}$$

where V_a is the scene radiometric measurement, $\overline{V}_C$ and $\overline{V}_{WL}$ are the cold space observations and the warm-load calibration target averaged for each scan, respectively, T_C is the effective cosmic background temperature and $\overline{T}_{WL}$ is the average of three precision thermometric measurements of the warm-load target. Inverting eqn.1 V_a are obtained.

If there is an error generated in T_a it should come from $\overline{V}_C$, $\overline{V}_{WL}$, V_a or $\overline{T}_{WL}$, which are the measurements that come directly from the hardware and communication system.

$\overline{V}_C$, $\overline{V}_{WL}$ or $\overline{T}_{WL}$ are provided in BASE files, but V_a need to be inverted from T_a already computed.

Eqn. 1 can be rewritten as eqn.2:

$$T_a = slope \cdot V_a + offset \tag{2}$$

where:

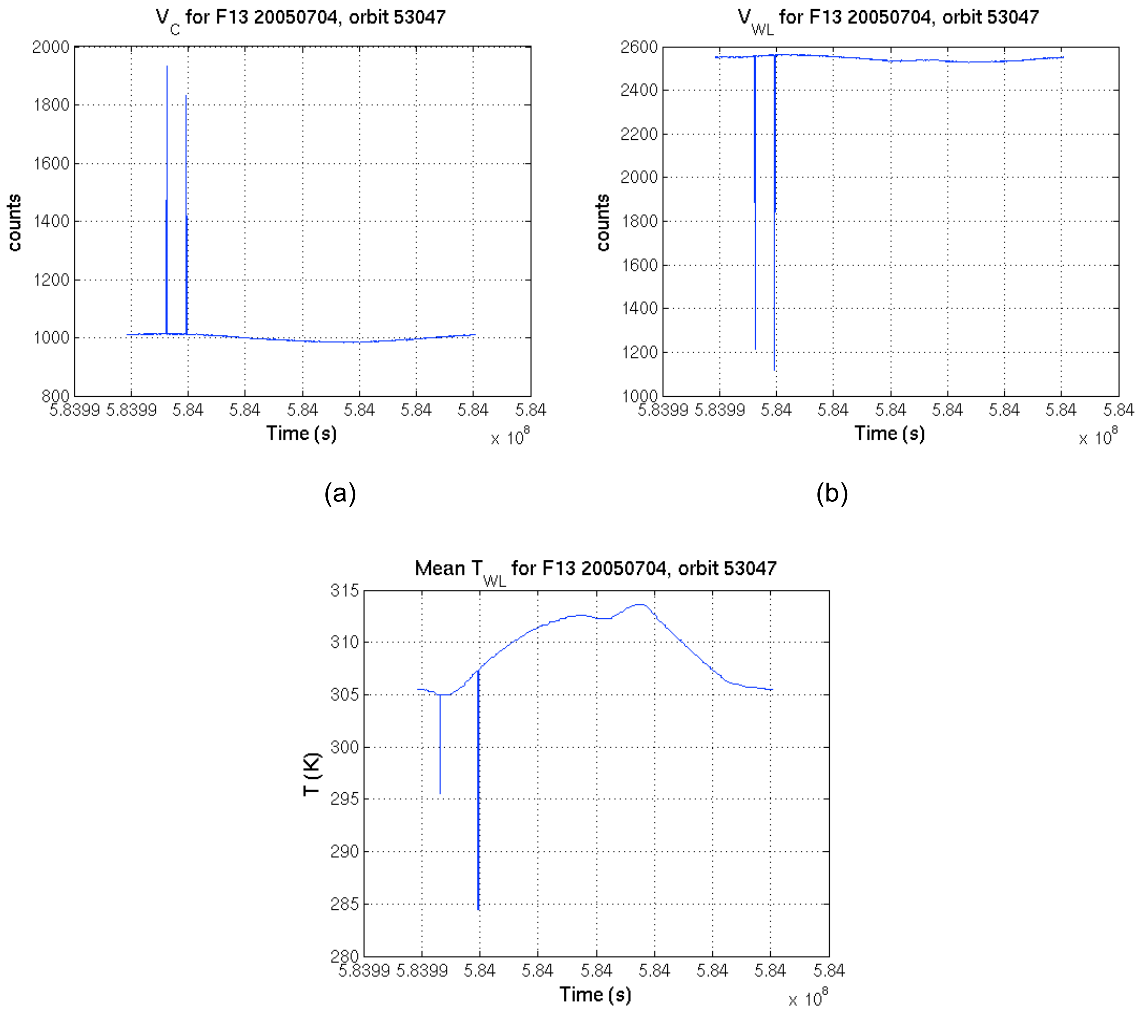
$$slope = \left(\frac{\overline{T}_{WL} - T_C}{\overline{V}_{WL} - \overline{V}_C} \right) \tag{3}$$

$$offset = \left(\frac{\overline{T}_{WL} - T_C}{\overline{V}_{WL} - \overline{V}_C} \right) \cdot \left(-\overline{V}_C \right) + T_C = -slope \cdot \overline{V}_C + T_C \tag{4}$$

Inverting eqn.2 V_a are obtained.

$$V_a = \left(\frac{T_a - offset}{slope} \right) \tag{5}$$

First the $\overline{V}_C$, $\overline{V}_{WL}$, and $\overline{T}_{WL}$ have been analyzed (Figure 3).



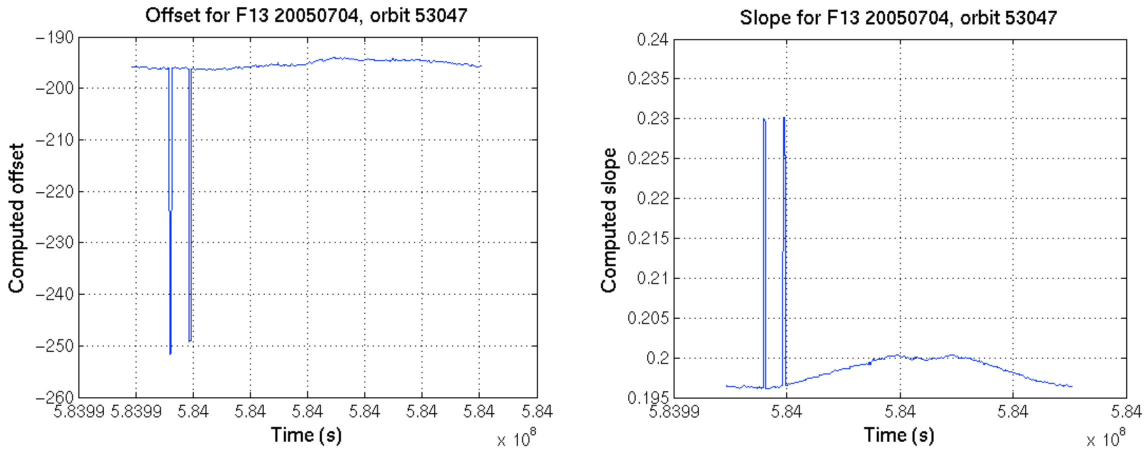
(a) (b)

(c)

Figure 3: BASE files values for (a) $\overline{V}_C$, (b) $\overline{V}_{WL}$, and (c) $\overline{T}_{WL}$.

As shown in Figure 3, there are some spikes either positive or negative in the three basic parameters from which T_a is obtained. The original V_a is not stored in the BASE files, so the only way to go back is applying eqn. 5.

So that prior to correcting these $\overline{V}_C$, $\overline{V}_{WL}$ and $\overline{T}_{WL}$ and recomputing the offset and slope values, we need to use the stored offset and slopes (computed without correcting the spikes) in order to go back to V_a from T_a.



(a) (b)

Figure 4: Slope and offset stored in BASE files used to compute T_a from V_a.

Note that the slope and offset were computed with the $\overline{V}_C$, $\overline{V}_{WL}$ and $\overline{T}_{WL}$ containing spikes (following eqn.3 and eqn.4) and then averaged using a 10 samples window. So that, the spikes in $\overline{V}_C$, $\overline{V}_{WL}$ and $\overline{T}_{WL}$ become jumps of 10 scans for the offset and slope computations.

If those original values for the offset and the slope are used in eqn. 5, we go back to the original V_a.

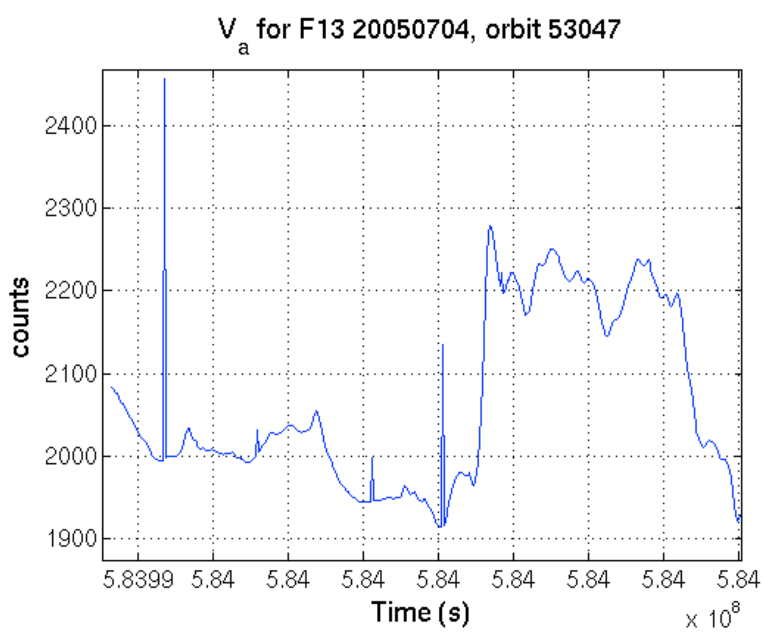


Figure 5: Mean V_a recomputed using the original offset and slope values and the T_a in BASE files.

As it can be seen in Figure 5 the jumps observed in Figure 1 are gone and just some spikes remain there. That means that at this stage we don't have auto-induced errors and only those coming from real hardware or communication system problems remain.

2.1.3 Correcting BASE files data

The $\overline{V}_C$, $\overline{V}_{WL}$, V_a, and $\overline{T}_{WL}$ have been corrected by applying a spike-detection and interpolation algorithm written in C. The result is shown in Figure 6.

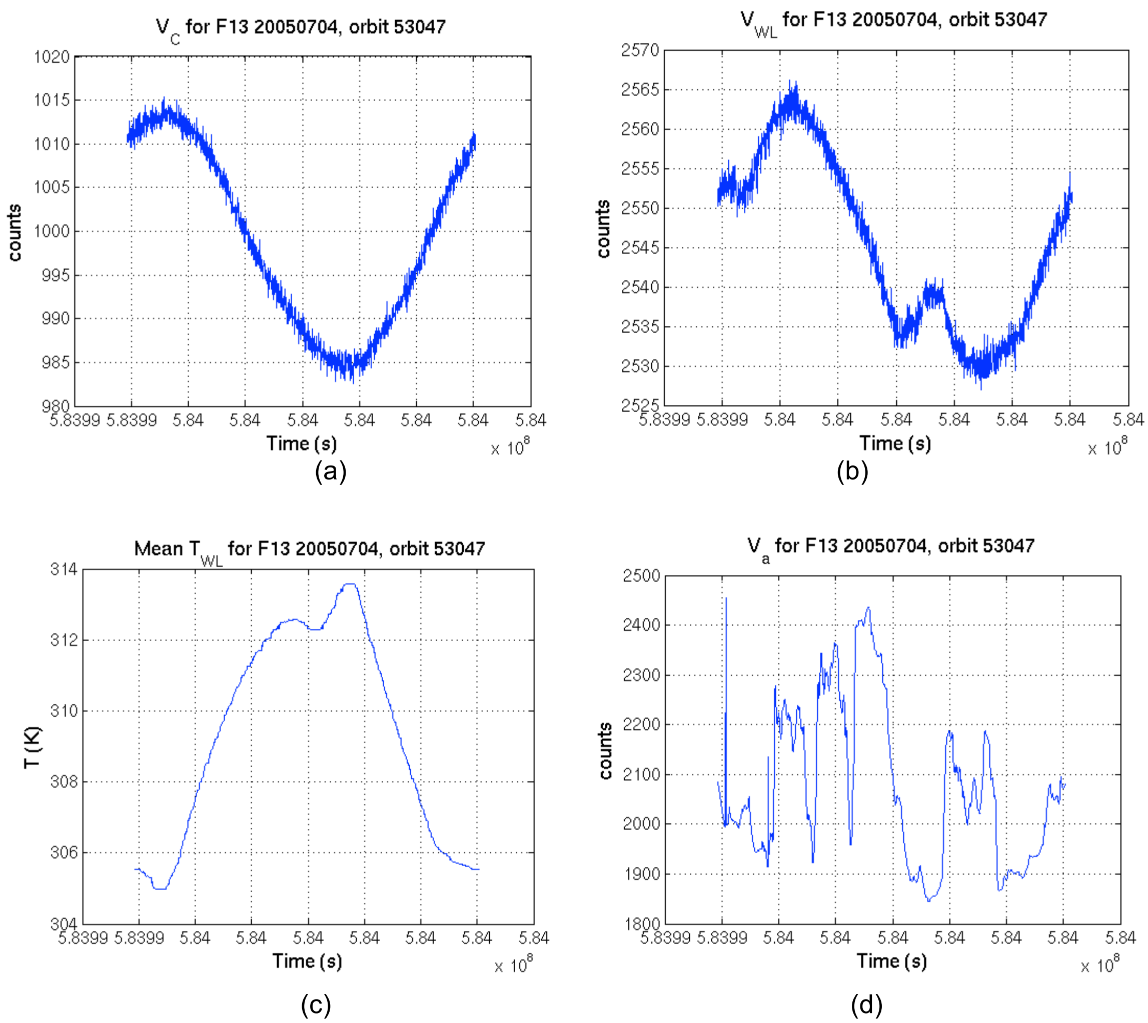
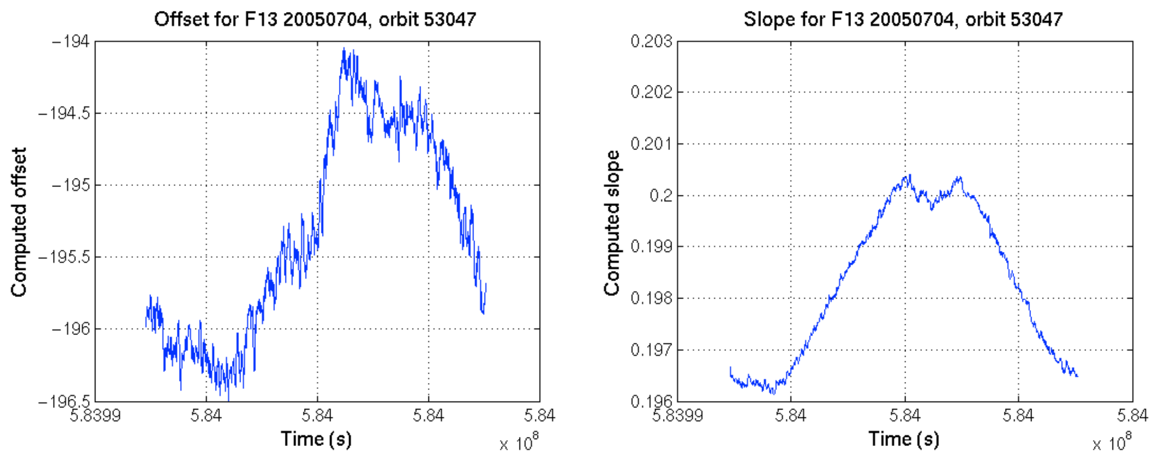


Figure 6: BASE files corrected values for (a) $\overline{V}_C$, (b) $\overline{V}_{WL}$, (c) $\overline{T}_{WL}$, and (d) mean V_a.

The spikes present in Figure 3a, b and c have been corrected as shown in Figure 6a, b, and c respectively. In Figure 6d the are still some spikes which have not been corrected since they are not in $\overline{V}_C$, $\overline{V}_{WL}$, and/or $\overline{T}_{WL}$. The reason for that is that V_a variability is high due to the observation scene, and sometimes is difficult to distinguish real variations from artificial spikes. So that, the decision is to leave those spikes in V_a data and allow to the scan check for deviations from climatology decide if those spikes are good data or should be set to missing.

2.1.4 Re-computing T_a

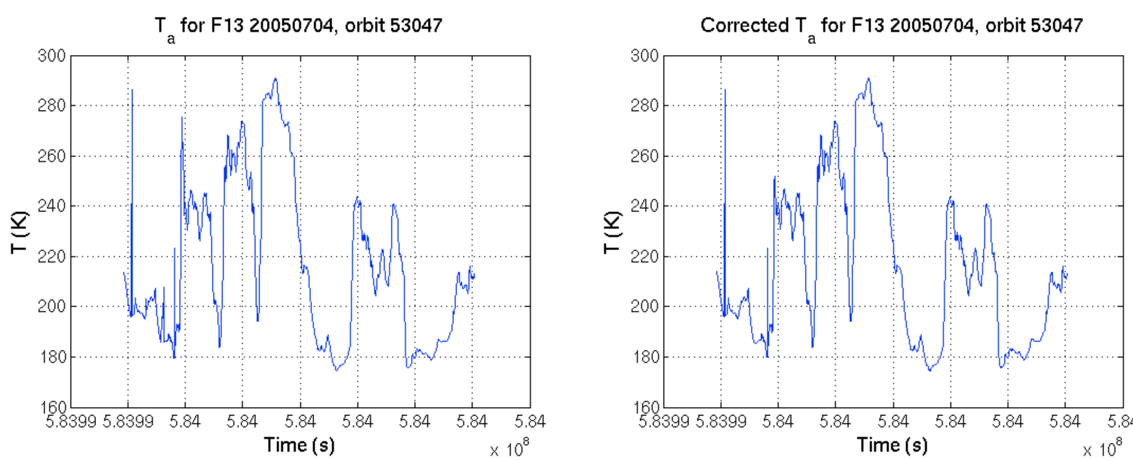
Once $\overline{V}_C$, $\overline{V}_{WL}$, V_a and $\overline{T}_{WL}$ are corrected, eqn.3 and eqn.4 are again applied to these new values getting the corrected versions of $slope$ and $offset$, respectively.



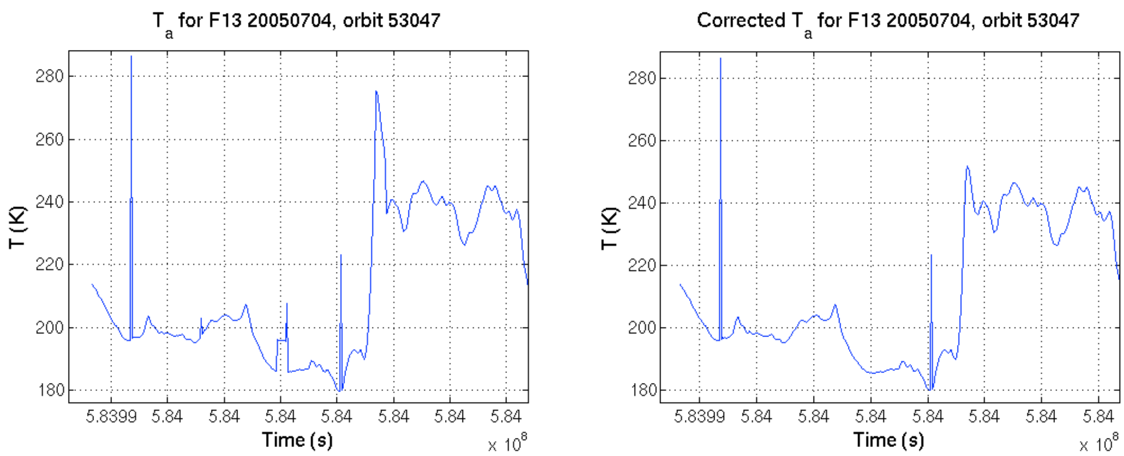
(a) (b)

Figure 7: Recomputed slope and offset.

From them and using the corrected V_a, eqn. 2 is applied to get T_a.



(a) (b)



(c) (d)

Figure 8: T_a comparison: (a) original mean T_a, (b) corrected mean T_a, (c) zoom for original mean T_a, and (d) zoom for the corrected mean T_a.

Figure 8 shows a comparison between the original mean T_a and the corrected mean T_a. As it can be seen in the zooms, Figure 8c and d, the spikes coming from the basic parameters are gone.

From T_a, T_b can be obtained following eqn.6, Colton (1999).

$$T_{b_p} = c_0^{'} \cdot T_{a_p} + c_1^{'} \cdot T_{a_q} \tag{6}$$

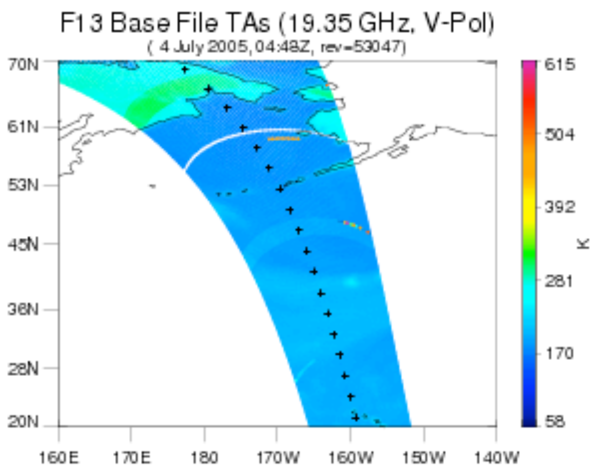
where $c_0^{'} = c_0 + c_2 + c_3$ and $c_1^{'} = c_1$ (coefficients obtained from CSU database and different for each channel and each satellite), and sub-indices p and q refers to either v or h polarizations.

In the processing to produce the FCDR, the T_a correction described in this report is part of the quality control stage. Additional quality control checks are performed and the cross-track bias correction is applied to T_a before T_a is converted to T_b. In the FCDR data set, this correction has been performed using the CSU database information and for the corrected data, the quality flag is set to a value of 14.

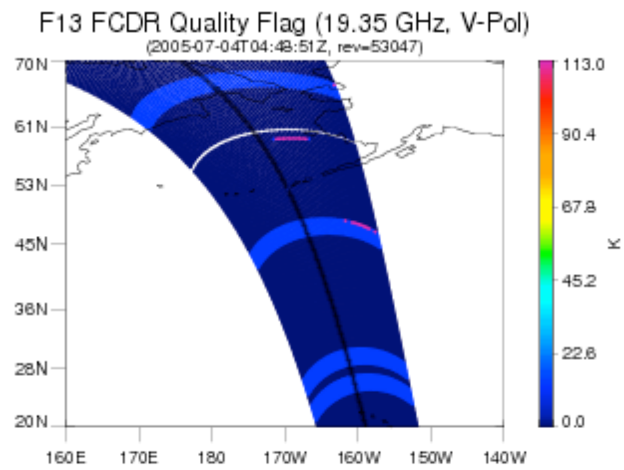
2.2 Examples

Some examples of the C code running are shown now.

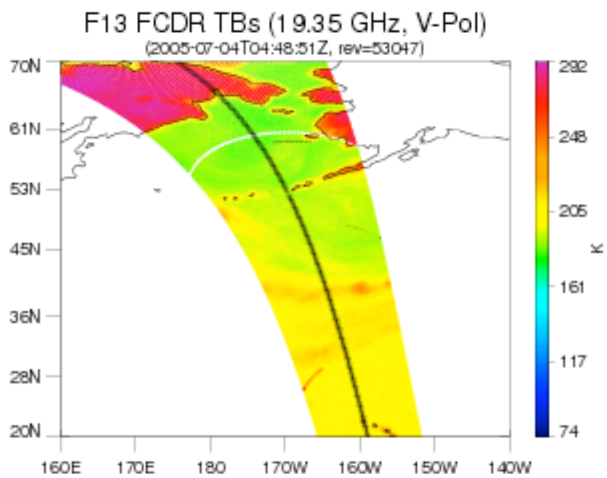
Case 1. 20050704, ORBIT: 53047



(a)



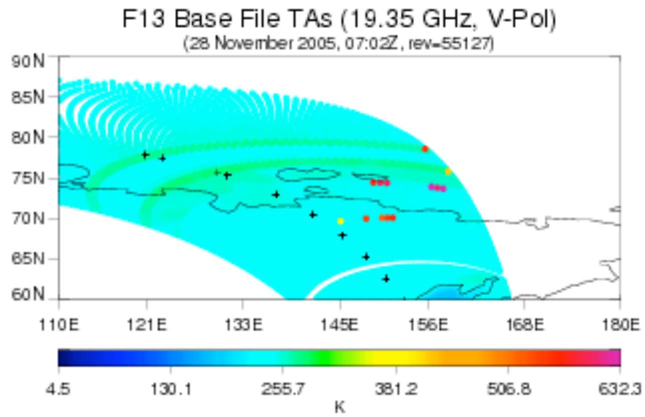
(b)



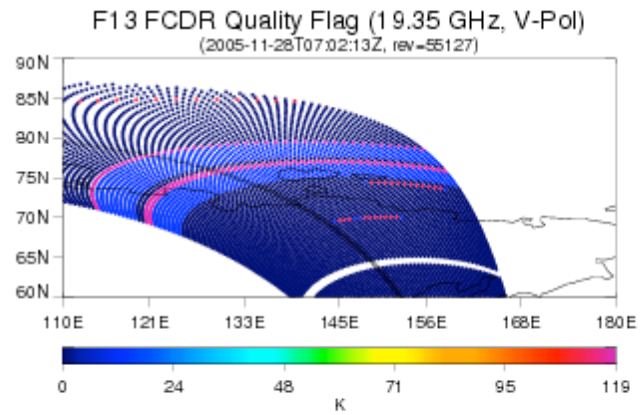
(c)

Figure 9. 19.35 GHz V-pol channel for the F13 satellite on 20050704, ORBIT: 53047. (a) original T_a, (b) scans set to missing (pink) and scans corrected (blue), and (c) corrected T_b.

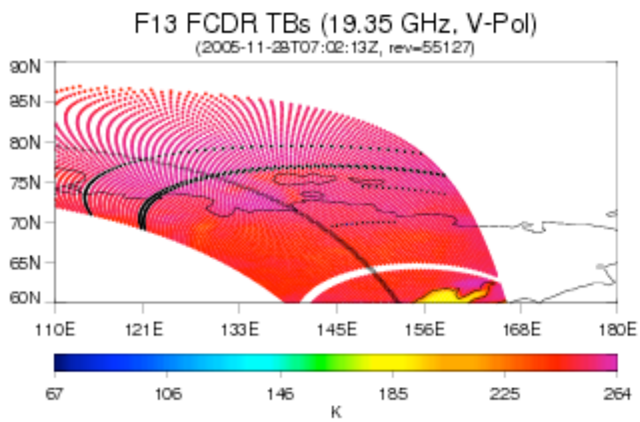
Case 2. 20051128, ORBIT: 55127



(a)



(b)



(c)

Figure 10. 19.35 GHz V-pol channel for the F13 satellite on 20051128, ORBIT: 55127. (a) original T_a , (b) scans set to missing (pink) and scans corrected (blue), and (c) corrected T_b .

Case 3. 20050617, ORBIT: 52810

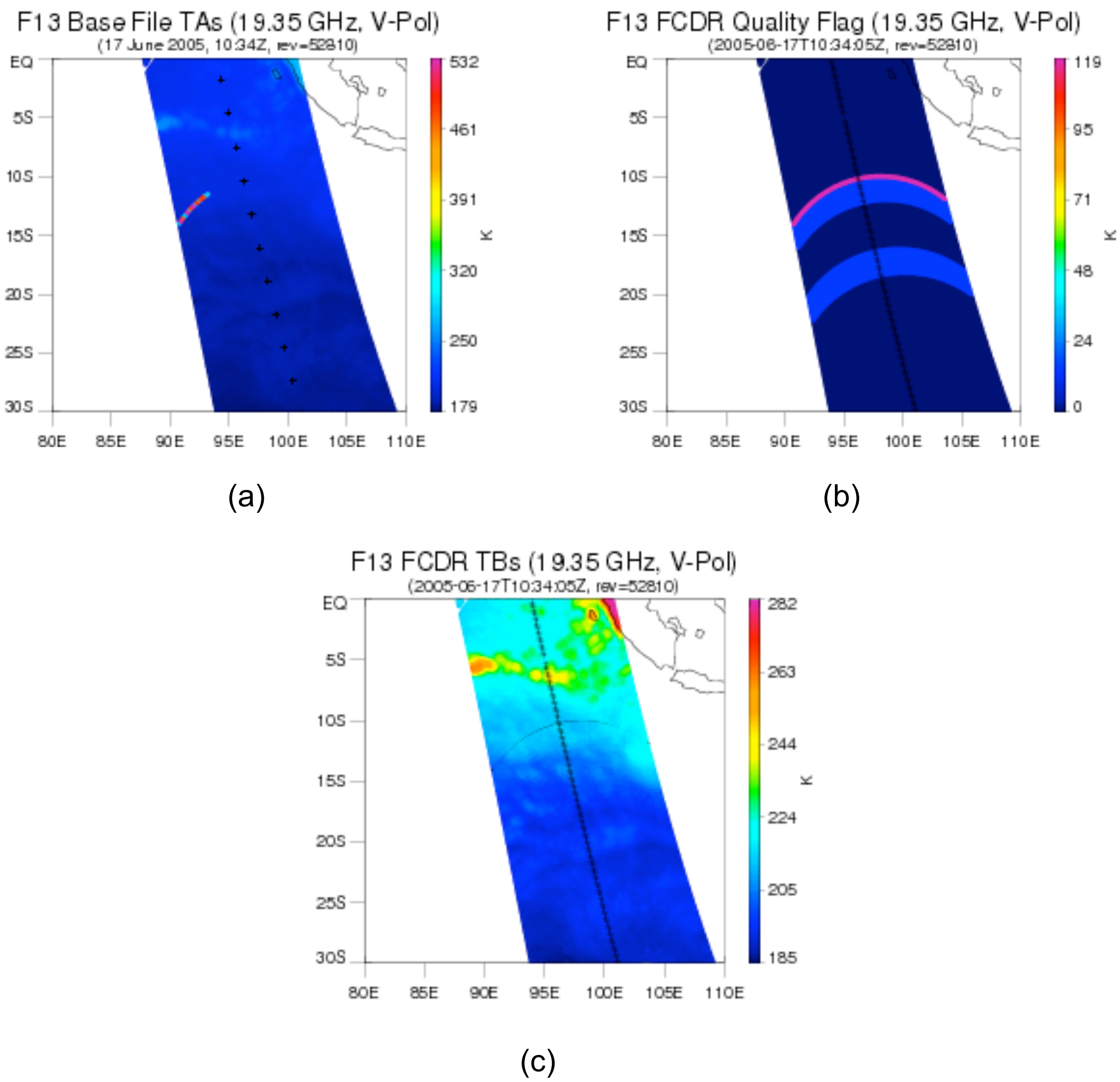
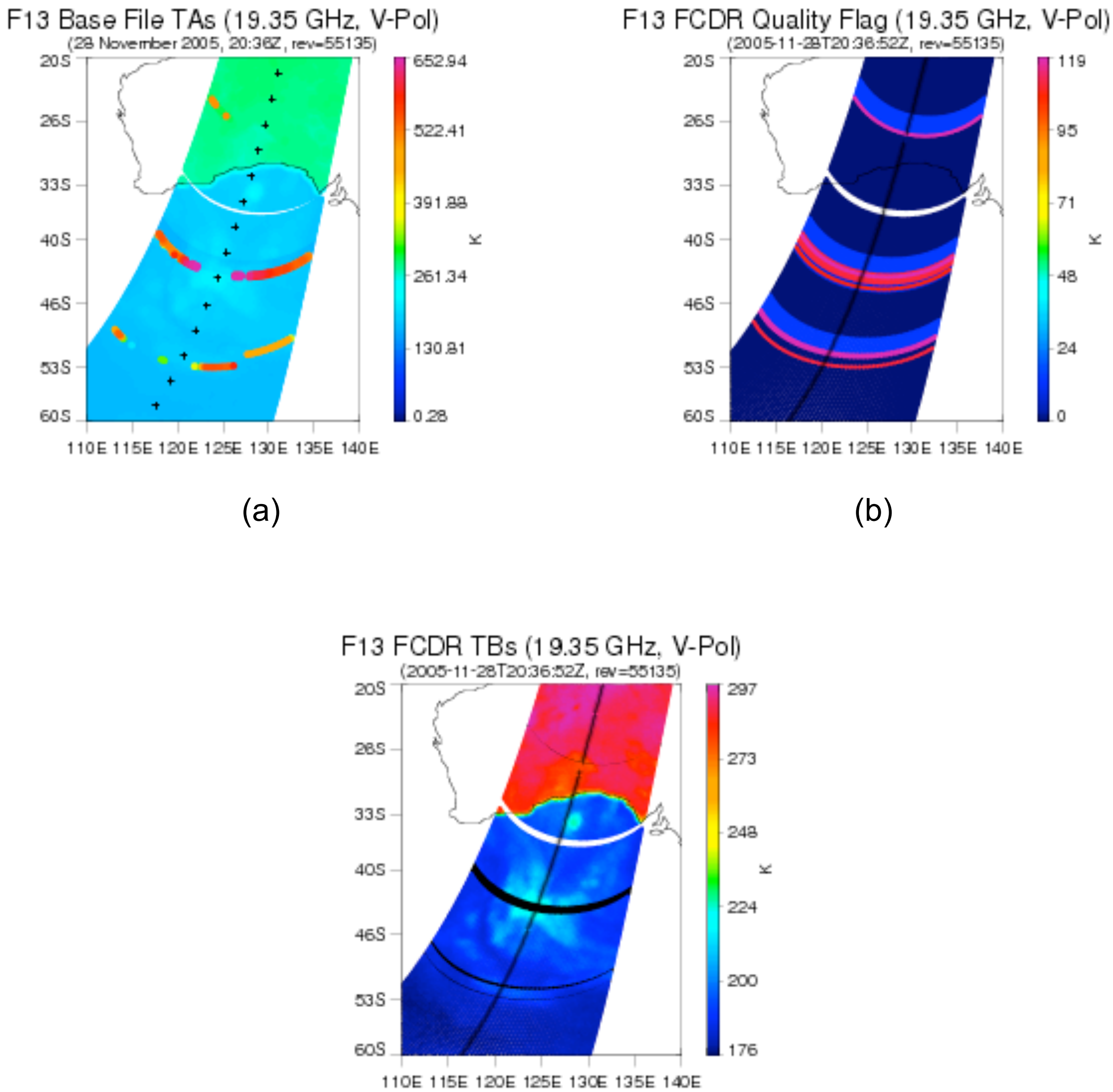


Figure 11: 19.35 GHz V-pol channel for the F13 satellite on 20050617, ORBIT: 52810. (a) original T_a, (b) scans set to missing (pink) and scans corrected (blue), and (c) corrected T_b.

Case 4. 20051128, ORBIT: 55135



(a) (b)

(c)

Figure 12: 19.35 GHz V-pol channel for the F13 satellite on 20051128, ORBIT: 55135. (a) original T_a , (b) scans set to missing (pink) and scans corrected (blue), and (c) corrected T_b .

2.3 Notes

2.3.1 Time jumps

An important thing to check is to make sure that the correcting process that it is being applied to T_a and so T_b, is not adding data that actually do not exist. Figure 13 shows one of this time jumps.

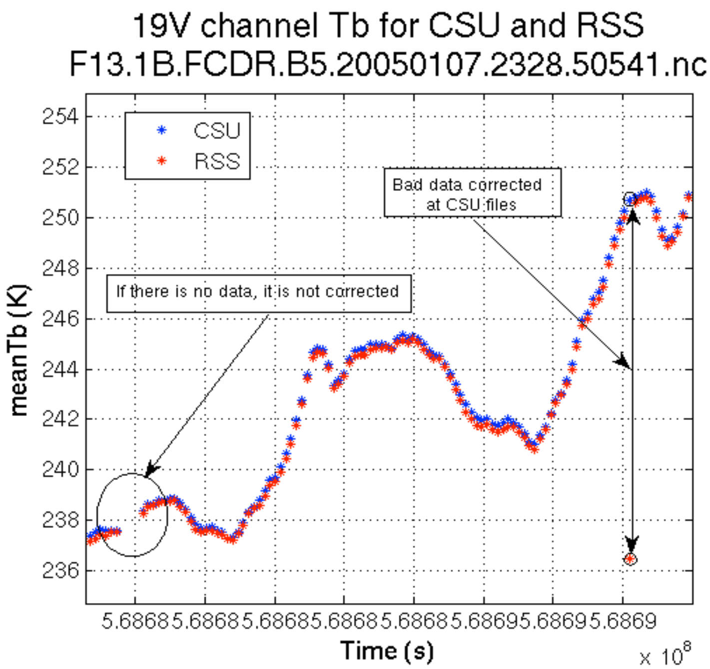


Figure 13: Time jumps are not affected by the correction process.

As it can be seen in Figure 13 when no data is present for a period of time, any sample is added to the corrected CSU T_b.

2.3.2 Other anomalies found when comparing to RSS T_b

Sometimes, when comparing T_b values from RSS and CSU some unexpected differences appear at isolated scans (some of them marked in red at Figure 14).

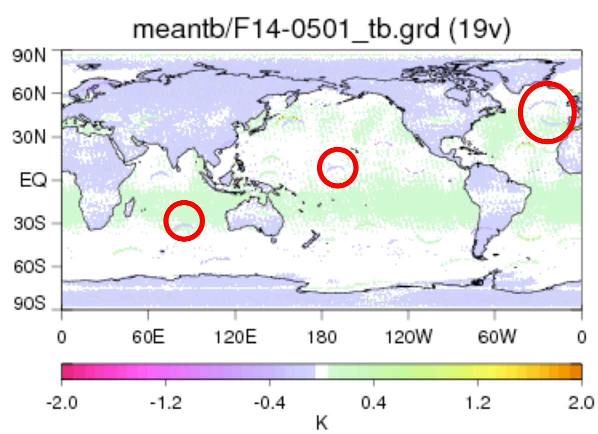


Figure 14: Difference between RSS and CSU mean T_b for satellite F14 on January 2005, channel 19V.

These scans have been analyzed for every single orbit and many of them are due to RSS T_b spikes. Figure 15 shows these events.

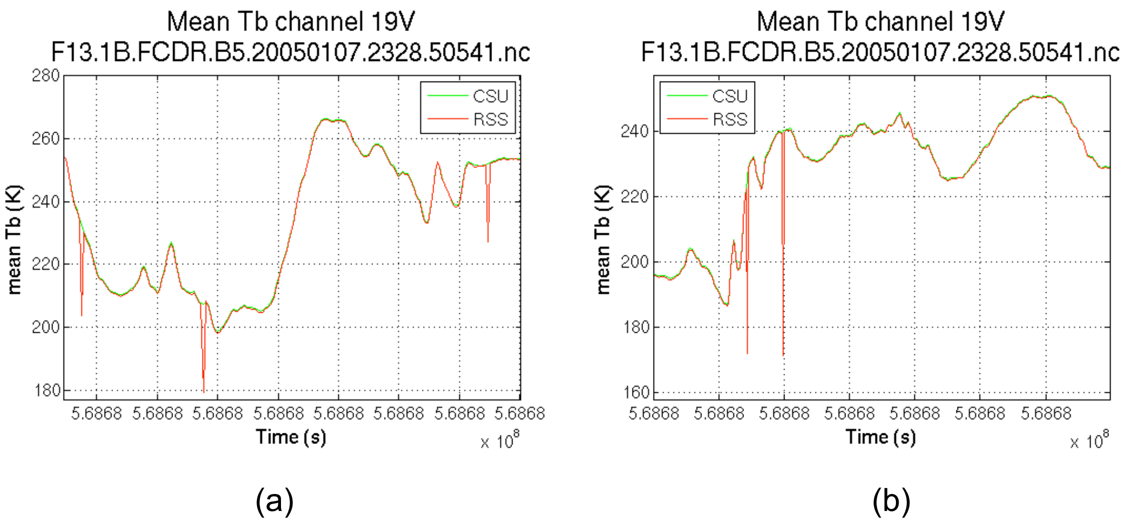


Figure 15: RSS T_b anomalies.

Probably these scans will be flagged on the scan check for deviations from climatology, but they don't seem to be flagged on RSS T_b.

3. Check BASE file (SSM/I and SSMIS)

Some quality control checks were done during the creation of the BASE files including checks for erroneous pixel geolocation and/or large variance from climatology for multiple data scans. This routine sets the associated FCDR quality flag and sets the affected output Tb to missing.

4. Check Geolocation (SSM/I and SSMIS)

For data where the original pixel location, given by latitude and longitude, is more than 100 kilometers away from the computed location, the Ta data are set to missing and a quality flag set to indicate a geolocation issue.

5. Check Sensor (SSM/I and equivalent SSMIS channels)

Data is set to missing and the quality flag is set for known sensor issues as determined from documented issues and data monitoring. For example, dates corresponding to the failure of first the 85v and then the 85h channels on the F08 SSM/I were determined from an analysis of the data and hardcoded into this routine. The details of which satellites and orbits/dates have known sensor issues are hard-coded into the function that performs this check rather than read from a file.

6. Check Climatology (SSM/I and equivalent SSMIS channels)

For each channel, scans where a significant fraction of pixels differ from the climatological mean values (determined as described in section 3.4.5.2) by more than 3 standard deviations are flagged. Two levels of climatology checks are identified. A climatology warning flag is set for scans near the threshold (within 5%) and the Ta data are retained. For scans exceeding this threshold, an error flag value is set in the corresponding quality flag and the Ta data is set to missing.

Figures 16 and 17 show examples of bad data that must be identified and removed. In Figure 16 the data is significantly displaced from the reported pixel geolocation, which shows the southern tip of South America misplaced to the North by a substantial amount. Figure 17 shows a case in which there are interleaved unphysical brightness temperatures over the entire orbit, only part of which is shown here. The magnitude and frequency of these erroneous data can have significant implications for climate applications.

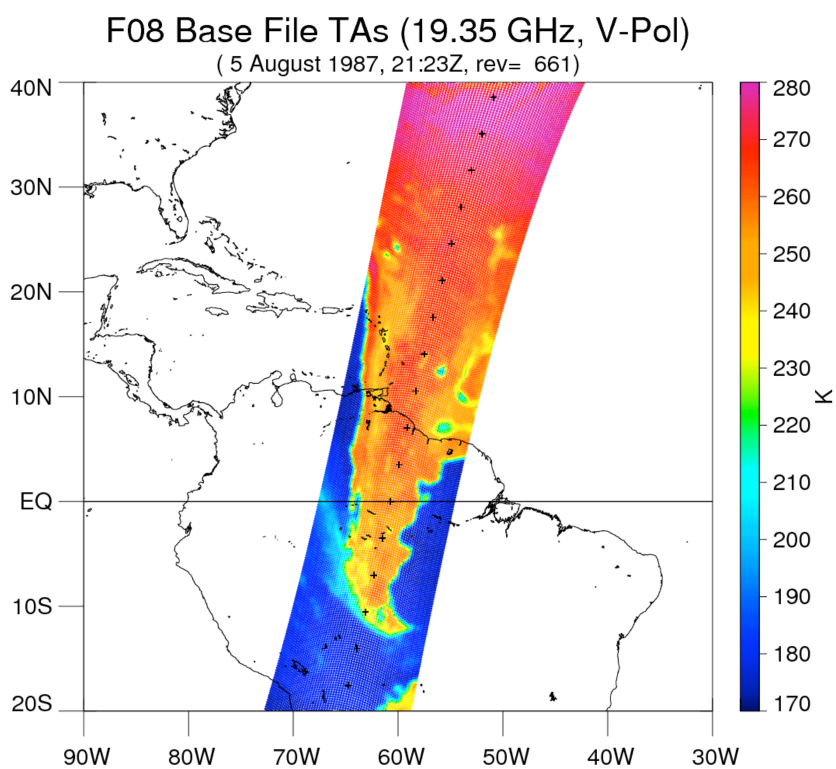


Figure 16: Mislocated data due to erroneous scan time values

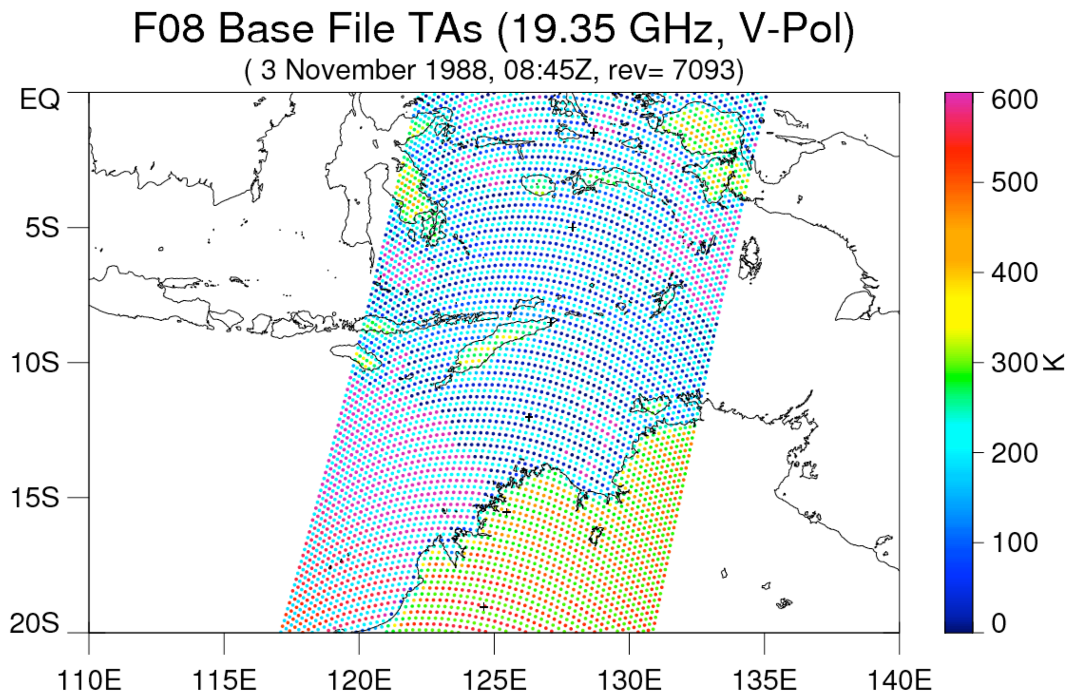


Figure 17: Erroneous Tb values interleaved with good data. The corresponding scan time values are slightly offset from the good scans, however, as seen from the color table most of the bad scans have unphysical Tb values.

To identify sections of orbits such as those shown in Figure 16 that are misplaced or significantly deviate from climatological conditions, a test was developed that compares antenna temperatures to the mean climatological values. An example is in Figure 18 with the original Ta shown in the top panel for the 19 GHz vertically polarized channel. The middle panel shows differences from climatological values in units of standard deviations from the mean. Pixels more than 3 standard deviations from the mean are shaded in grey.

A moving window of width 1500 seconds is used to identify sections with a large number of scans than are more than 3 sigma from the climatological values. The panel on the bottom of figure 18 shows the number of pixels, out of a maximum of 64, outside of 3 sigma. The red line indicates the region of data flagged by the quality control check and set to missing in the output FCDR data file. This QC test was developed by analyzing a large number of cases such as the one shown here and was found to effectively flag sections of data with a large number of misplaced scans, while retaining data containing large meteorological events. A secondary screen to identify shorter sections of misplaced or bad data with unphysical characteristics is used in combination with this broad window test.

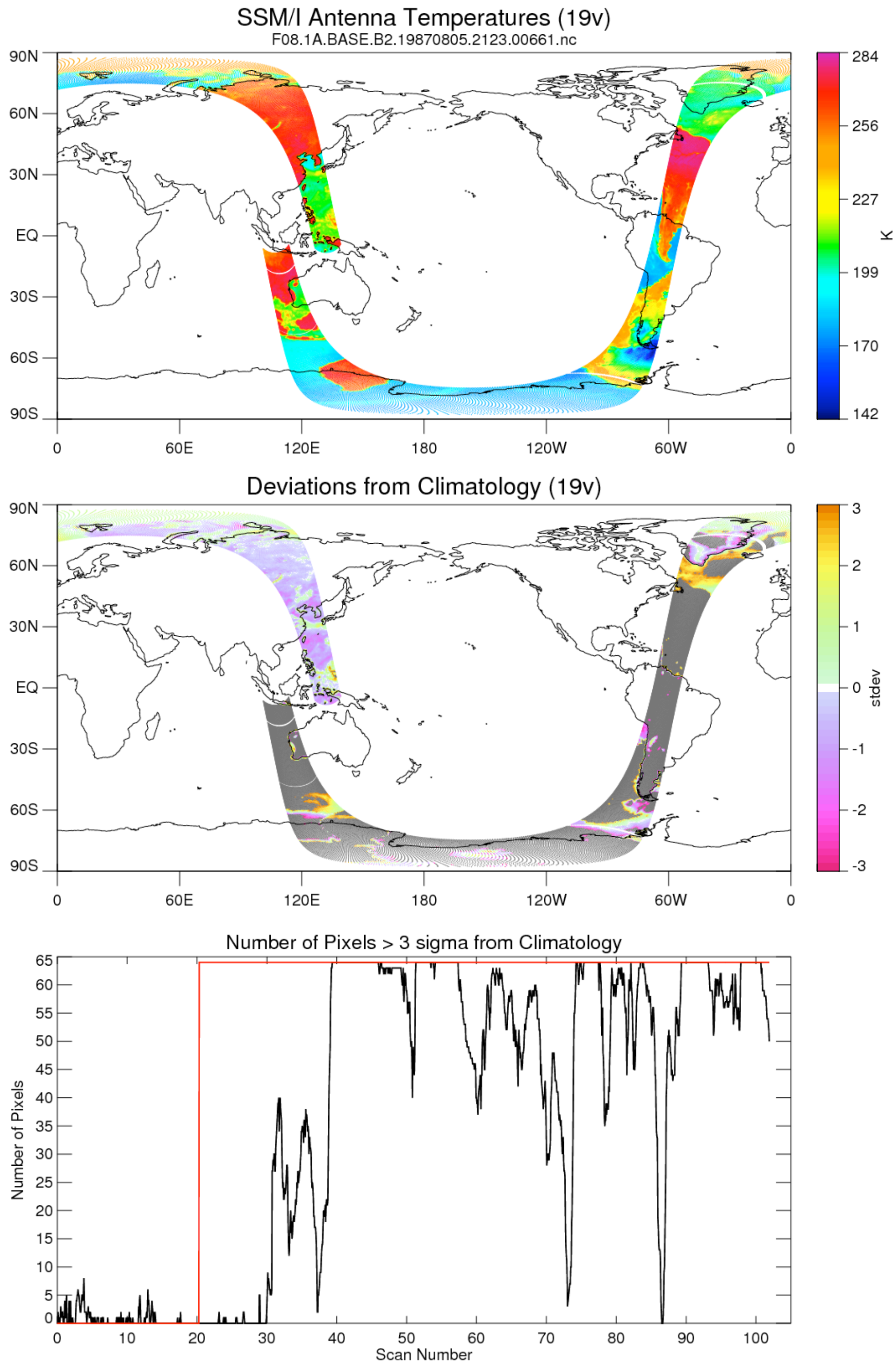


Figure 18: Top panel shows original Ta values with large chunk of scans over South America significantly displaced. The middle panel shows the deviation of the Ta values from the climatological mean in units of the standard deviation. Gray shaded scans indicate pixels exceeded 3 sigma from the mean. The bottom panel shows the resulting screen with the red indicating the scans flagged as bad.

7. Check Pixel (SSM/I and SSMIS)

Pixels where the distance between adjacent pixels along a scan is outside minimum and maximum thresholds, the Ta values are outside minimum and maximum thresholds (i.e. nonphysical), or the pixel latitude/longitude is out of range, have the appropriate quality flag set and the affected Ta are set to missing. The min/max thresholds are set in the parameter file params.h and the current values are TAMIN=50.0 and TAMAX=350.0.

8. Sensor Statistics

Monthly mean summary statistics for the series of nine SSM/I and SSMIS sensors extending from July of 1987 through December of 2012 are shown below in Figures 19-22. These statistics are based on the production of the BASE files, showing the data completeness, the number of duplicate scans removed, the number of scans flagged by the geolocation or time check, and the number of scans flagged by the climatology check. These statistics are intended to indicate an overall picture of the data availability, scan duplication, and gross scan data errors in the original TDR data for each of the nine sensors. For example, as shown in Figure 22 F08 and F10 both have significant periods where a large percentage of scans is flagged by the climatology check.

8.1 Data completeness

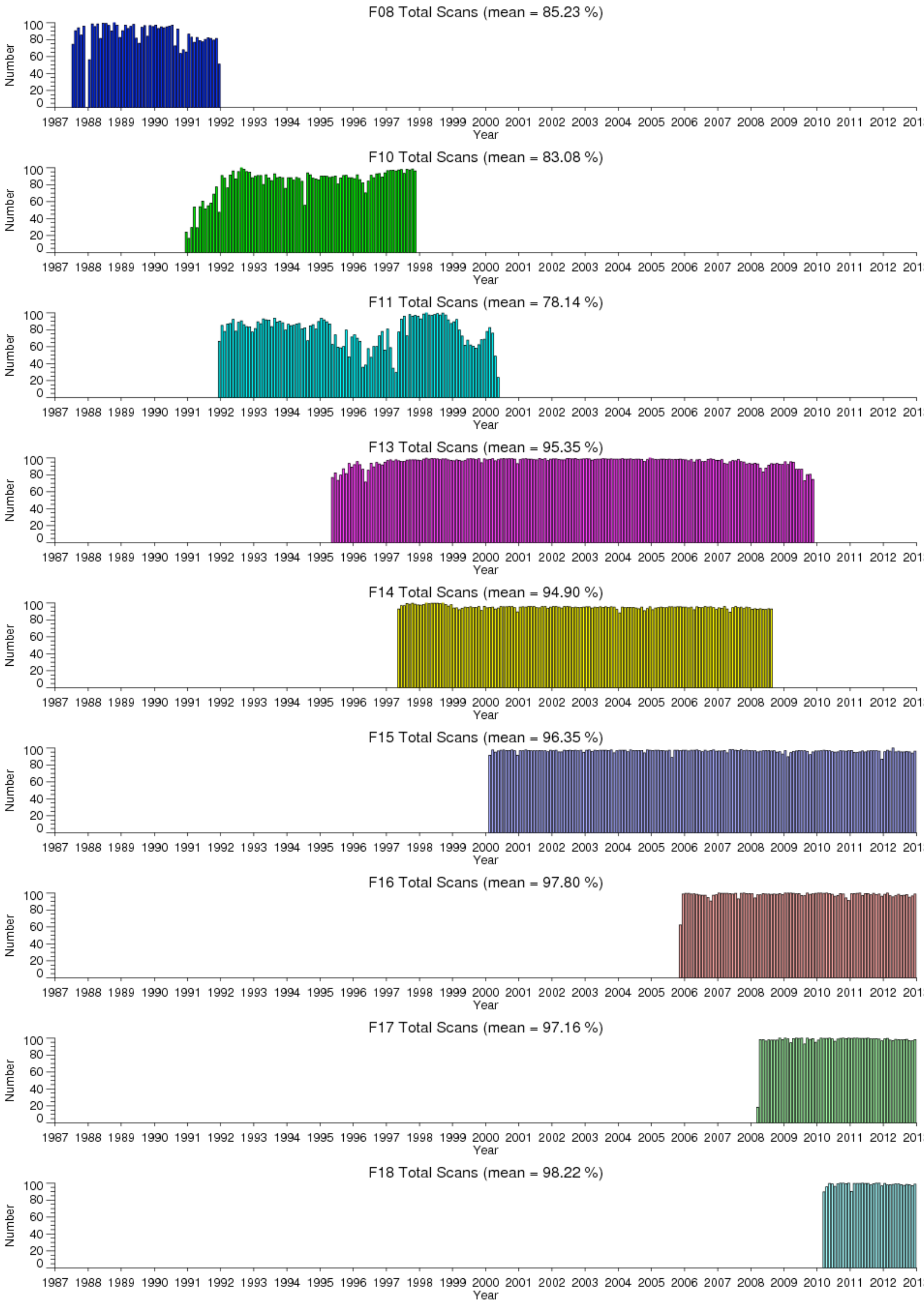


Figure 19: Monthly average percentage of scans available

8.2 Duplicate Scans

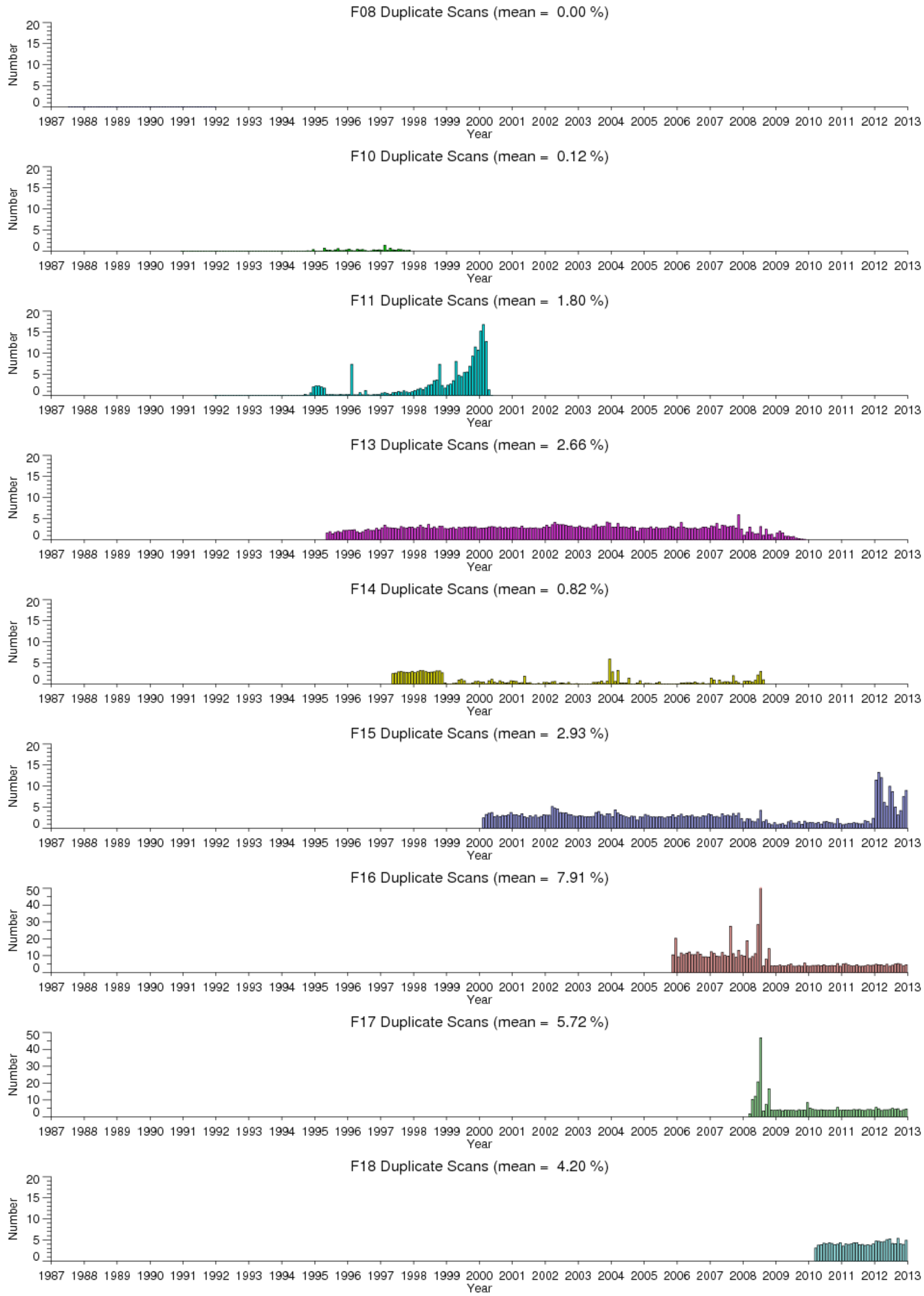


Figure 20: Monthly mean percentage of duplicate scans in original TDR data

8.3 Data flagged with bad scan time values

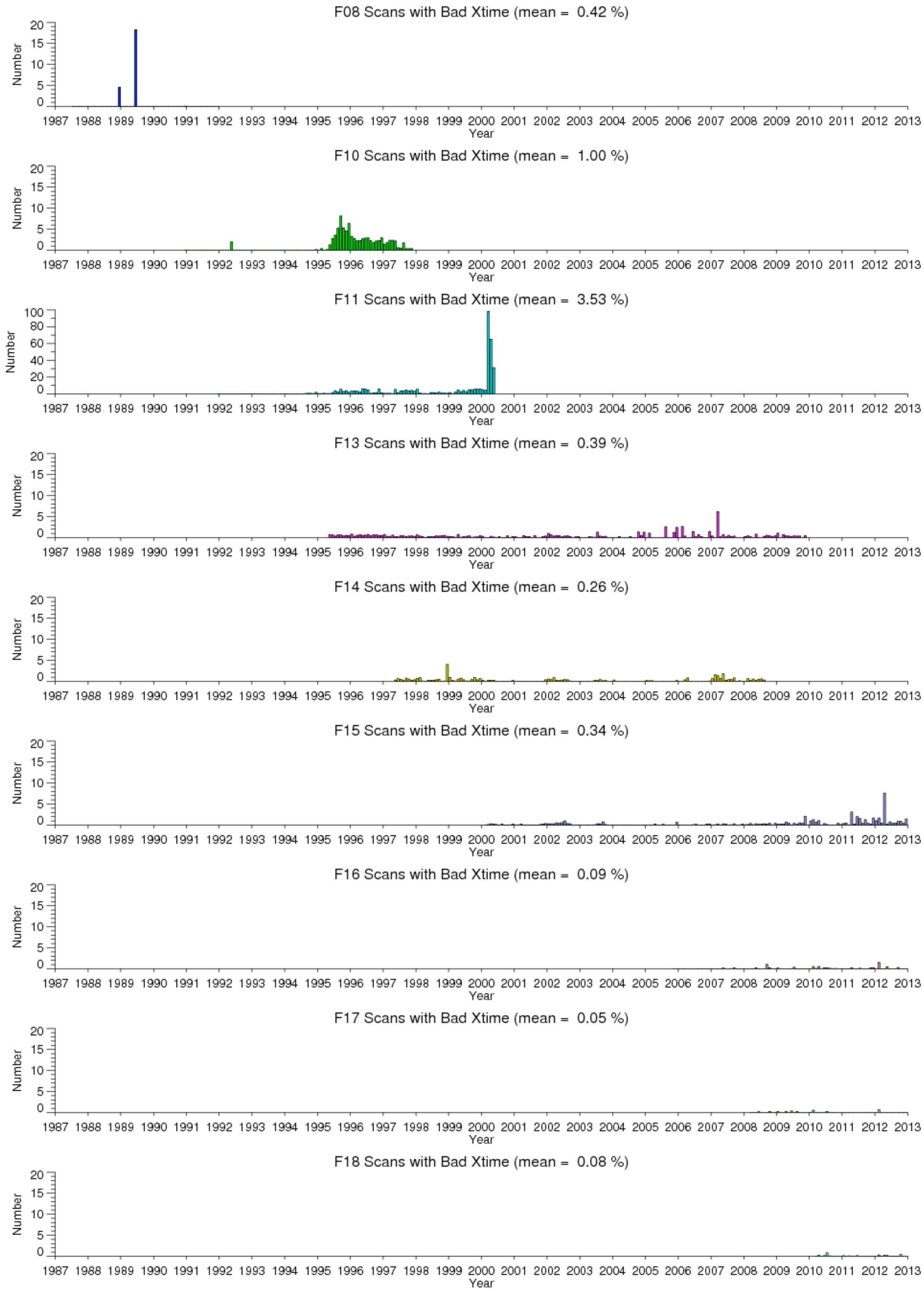


Figure 21: Monthly mean percentage of scans flagged with bad time values.

8.4 Data flagged based on climatology check

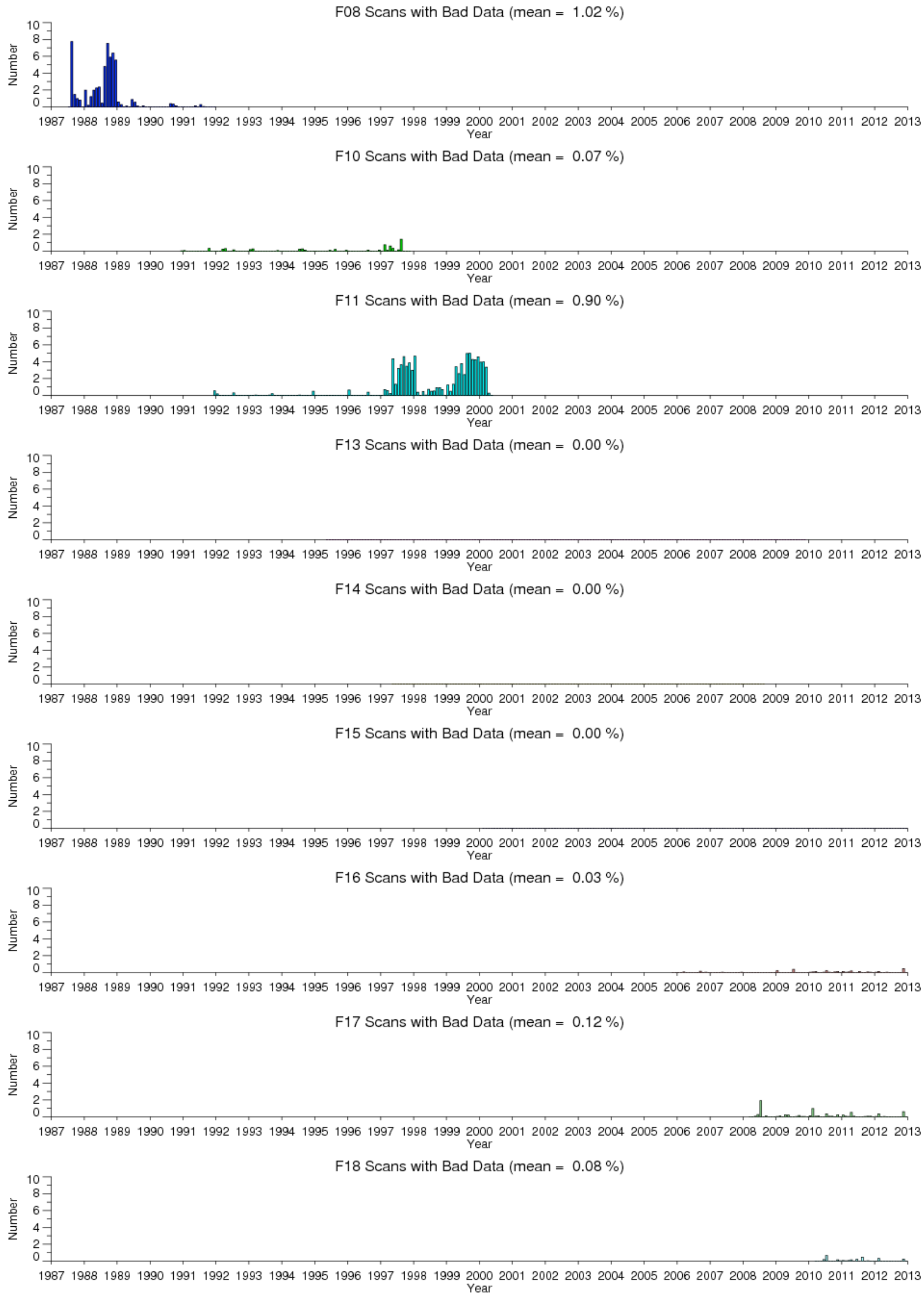


Figure 22: Monthly mean percentage of scans flagged based on climatology check.

9. References

Colton, M. C., and G. A. Poe (1999). Intersensor Calibration of DMSP SSM/I's: F-8 to F-14, 1987 - 1997. *IEEE Transactions on Geoscience and Remote Sensing,* 37 (1) : 418-439.