NASA/TM–2019-219036

Statistical Analysis of Aquarius Radiometer Radio Frequency Interference (RFI) and Implications on RFI Detection and Mitigation

Paolo de Matthaeis and David M. Le Vine

July 2019

NASA STI Program ... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM–2019-219036

Statistical Analysis of Aquarius Radiometer Radio Frequency Interference (RFI) and Implications on RFI Detection and Mitigation

Paolo de Matthaeis
Goddard Space Flight Center, Greenbelt, MD

David M. Le Vine
Goddard Space Flight Center, Greenbelt, MD

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, MD 20771

July 2019

Level of Review: This material has been technically reviewed by technical management.

Available from

NASA STI Program
Mail Stop 148
NASA’s Langley Research Center
Hampton, VA 23681-2199

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
703-605-6000

Statistical Analysis of Aquarius Radiometer Radio Frequency Interference (RFI) and Implications on RFI Detection and Mitigation

Paolo de Matthaeis
David M. Le Vine

1 Introduction

The Aquarius/SAC-D mission operated between August 2011 and June 2015 with the main goal of providing global estimates of sea surface salinity (SSS). It comprised both active and passive microwave sensors operating at L-band to observe the same surface area almost simultaneously. Measurements from both instruments underwent subsequent filtering to mitigate the effect of Radio Frequency Interference (RFI). This report describes the analysis of statistics of RFI in samples acquired by the Aquarius radiometers, and its results could be used to improve the performance of the interference detection algorithm.

The two main problems in RFI detection are false alarms and missed detection. When the detection algorithm is made very sensitive in order to detect RFI at very low level, many samples that are not affected by interference are instead erroneously flagged as they were. These samples are commonly referred to as false alarms and their percentage over the total number of samples is called false alarm rate. In addition, RFI cannot always be detected, especially when the level of the interfering signals is comparable to that of the natural surface emission, resulting in missed detection. The two issues of false alarms and missed detection cannot be solved at the same time, because if the detection algorithm is made less sensitive in order to avoid false alarms, the number of missed detections will increase.

In particular, the second problem is the most serious and still unresolved problem, and it still not clear what impact undetected RFI has on the science retrievel algorithm (i.e,. in this case, the estimation of SSS). The first part of this reports present an investigation of the general statistical properties of RFI in the measurements. In the second part of the report, a study was performed here to assess the amount of missed detection in certain ocean region where low-level interference is a problem. The approach is to estimate the statistics of RFI starting from the acquired antenna temperature samples. The statistics are then used to simulate RFI and study its detection by the RFI filter. The results of this simulation are compared with the RFI actually detected in the acquired samples to provide an estimate of the level of missed detection. As a potential application of these findings, the RFI algorithm parameters could be adapted to detect low level of RFI in regions where its presence has been documented.

An overview of the RFI detection and RFI mitigation approach is given in Section 2. An analysis of RFI statistical properties follows in Section 3 and includes the estimation of the RFI histograms. Then, the technique proposed to estimate missed detection is described in Section 4. Finally, Section 5 gives an example of estimation of RFI statistical distribution and missed detection. Conclusions are reported in Section 6.

2 Overview of Aquarius RFI Detection and Mitigation

The Aquarius algorithm assumes a signal with a normal distribution in the RFI-free environment and detects RFI by searching for outliers that exceed a threshold. The basic quantities used in the decision process are the mean value (dynamic) of the acquisitions within a moving averaging window surrounding the sample under test and the standard deviation (fixed) chosen to identify an outlier [1, 2].

2.1 Acquisition Scheme and Data Structure

The three radiometers, one for each antenna, operate in parallel. Each of them performs a measurement every 10 ms. In nominal operation, during 120 ms (one subcycle) each radiometer collects five samples looking into the antenna followed by five samples devoted to the calibration sources. The calibration sources are two noise diodes (ND) and a Dicke load (DL) [9]. One cycle of a radiometer measurement contains 12 subcycles. Figure 2.1 shows how the samples corresponding to one antenna temperature value are organized. Within each 120 ms subcycle, the first and second and the third and fourth 10 ms acquisitions are added together to produce two equivalent 20 ms samples. As a results, each subcycle contains five antenna measurements, also called short accumulations, recorded at the start of the cycle and 20 ms, 40 ms, 50 ms, and 60 ms afterwards, and five calibration (non-antenna) measurements.

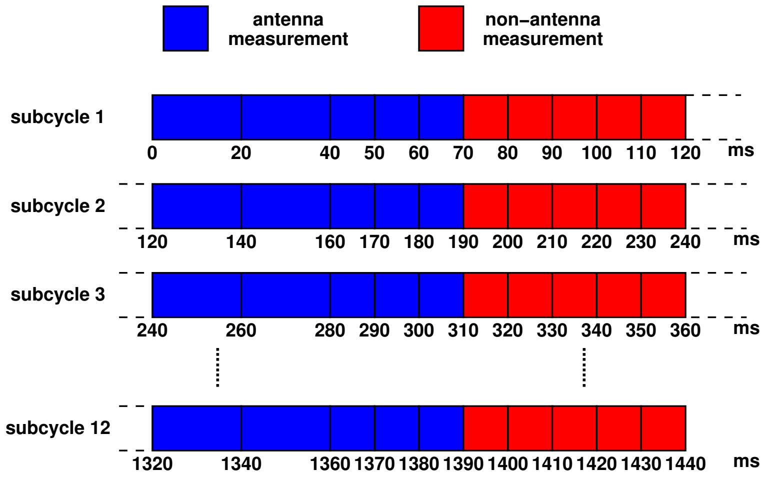


Figure 2.1: *Structure of a measurement cycle.*

The measurements acquired looking at the noise diodes and Dicke load in eight different configurations are averaged over longer periods (1.44 s) for the use of internal calibration and are called long accumulations. RFI detection and mitigation needs long accumulations in order to determine calibration gains and offsets.

Before being fed to the RFI Detection Algorithm, the string of data shown in Figure 2.1 undergoes some pre-processing. The first two short accumulations, which represent 20 ms of data, are divided by 2 and counted twice. In addition, zeros are applied when no antenna measurements were taken, such as during the intervals of internal calibration, as shown in Figure 2.2 (left). Therefore, for each subcycle of 120 ms, the result is a string of seven 10-ms spaced values $s_n (n = 1, 2, \ldots, 7)$ representing the short accumulations, followed by five zeros, $s_n (n = 8, 9, \ldots, 12)$, representing the intervals devoted to calibration. A correction for the non-linearity of the receiver is also applied to these samples $s_n (n = 1, 2, \ldots, 7)$. Since there are twelve subcycles in each 1.44 s block, a total of $7 \times 12 = 84$ non-zero 10 ms antenna samples are taken for each data point. All the strings are lined up in temporal order to form a continuous stream of values that are then used by the RFI detection algorithm.

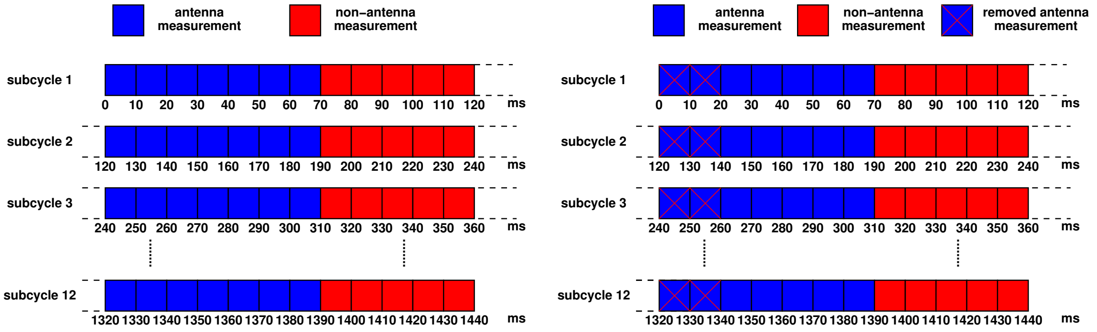


Figure 2.2: *Equivalent data string with 10 ms samples, with all measurements (left) and samples corresponding to the 20 ms first short accumulation removed (right).*

After launch, it was noticed that the first short accumulation (20 ms sample) had values inconsistents with the other short accumulations. Therefore, as a precaution, it has been excluded from the subsequent data processing. As a results, the data that are used for RFI detection algorithm are as shown in Figure 2.2 (right), with the two equivalent 10 ms samples corresponding to the first short accumulation of each subcycle being replaced by zeros. Only the samples $s_n (n = 3, 4, 5, 6, 7)$ are used for each subcycle, for a total of $5 \times 12 = 60$ non-zero 10 ms antenna samples per 1.44 s cycle.

2.2 RFI Detection and Mitigation Algorithm

The raw antenna measurements in Figure 2.2 are processed to limit the effect of Radio Frequency Interference (RFI) and converted to antenna temperatures. The flow diagram in Figure 2.3 illustrates the various stages required to accomplish this task.

The detection algorithm identifies individual samples of the antenna temperature (short accumulations) that deviate significantly from the average value of nearby samples. Mitigation is accomplished in subsequent processing steps by excluding corrupted samples before averaging them to yield the antenna temperatures. Some stages also require calibration gains and offsets that are computed based on the long accumulations. Individual samples

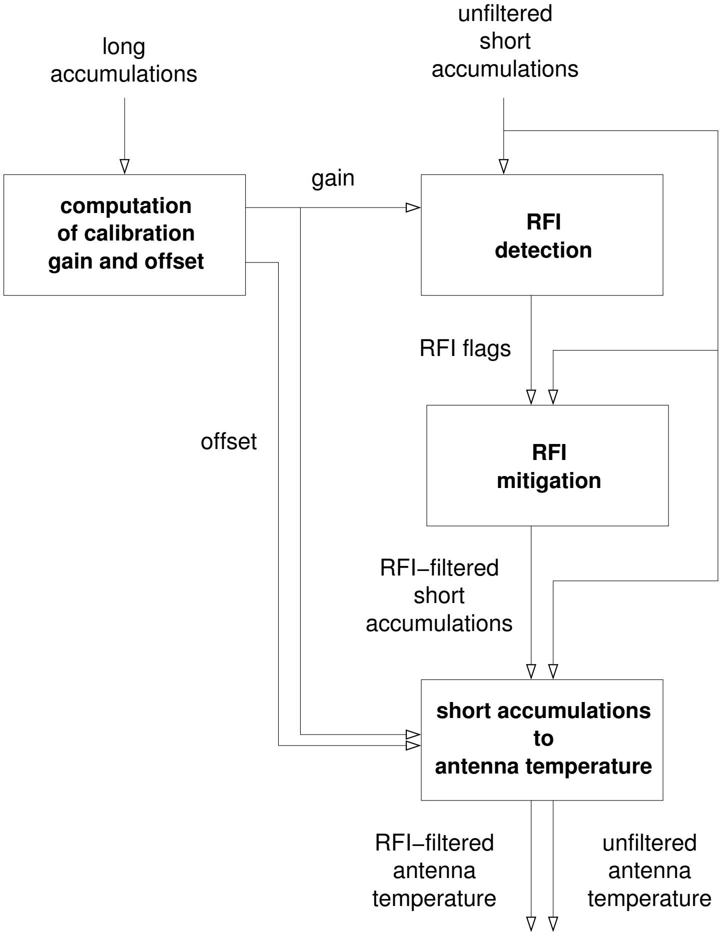


Figure 2.3: *Flow diagram for detection and mitigation of Aquarius radiometer RFI.*

of the correlated noise diode counts are also flagged as corrupted RFI if they are near an antenna temperature sample that has been flagged.

The algorithm is performed independently for each radiometer channel. All input and output data and dynamic auxiliary data are processed independently for each polarization. Accordingly, independent versions of all static auxiliary data files are maintained for each polarization and each radiometer. The following sections describe the input data and explain the different stages of the flow diagram in Figure 2.4.

The steps in the RFI detection algorithm are shown in the flow chart of Figure 2.4. The test for presence of RFI is applied to each valid sample (i.e. each value that corresponds to an antenna measurement) s_k, therein referred to as sample under test, in the data stream. The algorithm consists of the following steps.

Step 1. Selection of samples to be averaged together to estimate local mean

A set S_n consisting of $2\,W_m$ samples surrounding the sample under test is extracted from the data stream:

$$S_n = \{s_k \,|\, k \in K_n\}, \text{where } K_n = \{k = n - W_m, \ldots, n + W_m; k \neq n\} \tag{2.1}$$

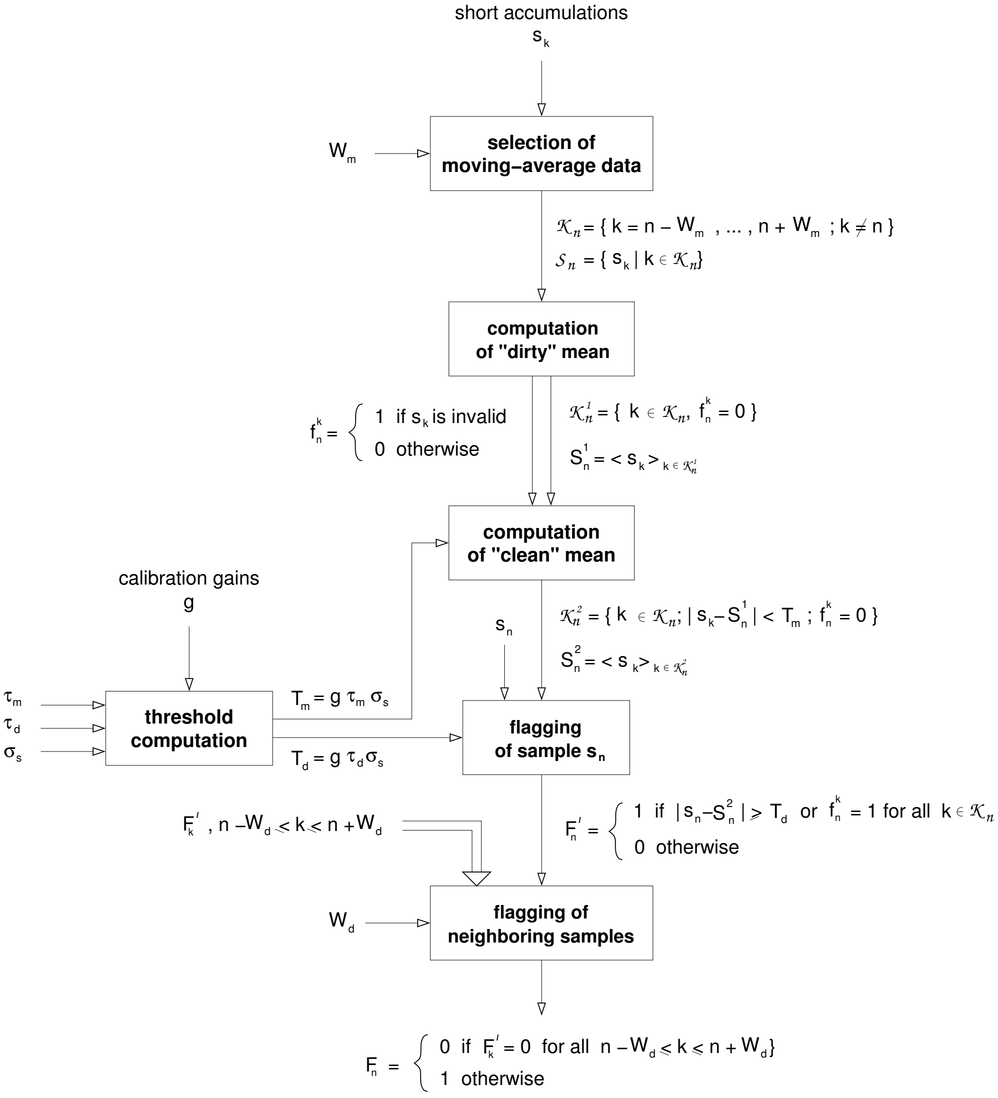


Figure 2.4: *Flow diagram for Aquarius radiometer RFI detection algorithm.*

This set will be used to estimate the local mean value of the short accumulations. However, because the samples are not uniformly spaced, the actual number of short accumulations falling within this time interval will vary. The local mean running average window, W_m is selected from a static table based on the location of the center of the antenna footprint for the sample under test. This is done in order to have the possibility of adapting the value of W_m to the characteristics of the observed surface, e.g., land or water. The W_m table is gridded in 1° increments of latitude and longitude. The exact value of the footprint location is rounded to the nearest latitude and longitude.

Step 2. Computation of dirty mean

A dirty mean is computed using all s_k in S_n, excluding samples that have been previously flagged as invalid, i.e.,

$$S_n^1 =< s_k >, \text{with } k \in K_n^1 \tag{2.2}$$

where K_n^1 are the indexes of samples in S_n not previously flagged as invalid.

Step 3. Computation of clean mean

A clean (in the sense of being free of outliers) mean S_n^2 is computed using only those samples s_k that satisfly the condition

$$|s_k - S_n^1| < T_m, \tag{2.3}$$

i.e, which differ from the dirty mean S_n^1 less than T_m in absolute value. The threshold T_m is given by

$$T_m = \tau_m \, \sigma_s \, g \tag{2.4}$$

Therefore, the expression for the clean mean is

$$S_n^2 =< s_k >, \text{with } k \in K_n^2 \tag{2.5}$$

where K_n^2 are the indexes of samples in S_n not previously flagged as invalid and that satisfy the inequality above. The clean mean is the basis for deciding if the sample being examined, s_n, is RFI.

Step 4. Testing samples for presence of RFI

The sample s_n is compared to the clean mean. If it differs from this mean by more than a preset threshold, Td, in absolute value, i.e.,

$$|s_k - S_n^2| < T_d, \tag{2.6}$$

then it is considered to be RFI. The threshold T_d is given by

$$T_d = \tau_d \, \sigma_s \, g \tag{2.7}$$

This intermediate RFI flag for the sample s_n is defined as

$$F_n' = \begin{cases} 0 & \text{if } |s_k - S_n^2| \geqslant T_d \\ 1 & \text{if } |s_k - S_n^2| < T_d \end{cases} \tag{2.8}$$

Step 5. Flagging samples on the neighborhood of RFI-flagged samples

Samples within $\mp W_d$ steps on either side of s_n (counting zeros) are considered tainted and also flagged as RFI. $W_d = 2$ in the current algorithm (corresponding to version v5.0) of the L2 data products), which means that all samples between $s_n \mp 2$ are considered to be RFI. Thus, the final RFI flag F_n for the sample s_n is defined as

$$F_n = \begin{cases} 0 & \text{if } F_n' = 0 \text{ for all } k \text{ such that } n - W_d \leqslant k \leqslant n + W_d \\ 1 & \text{otherwise} \end{cases} \tag{2.9}$$

2.3 RFI Detection Algorithm Parameters

The RFI detection algorithm uses five independent static (time-independent) parameters: σ_s, tau_m, τ_d, W_m and W_d. The parameter σ_s varies with polarization and radiometer beam, and its values are chosen to be close the standard deviation of the antenna temperatures measured over RFI-free ocean. The other four parameters do not depend on the polarization and beam, but can be tuned based on geographical location. The actual values are selected from static tables according to the location of the center of the antenna footprint for the sample under test. The τ_m, τ_d, W_m and W_d tables are gridded in 1°increments of latitude and longitude, and the exact value of the footprint location is rounded to the nearest latitude and longitude. However, in the current implementation of the algorithm, their values are constant with latitude/longitude and given by

$$\tau_m = 1.5,\ \tau_d = 4.0,\ W_m = 20,\ W_d = 2 \tag{2.10}$$

The values of σ_s currently being used in the algorithm can have two different values and are listed in Table 2.1. Land and ocean are defined based on the UCAR land-sea mask.

	σ_s over ocean				σ_s over land and sea ice			
Beam	V	+90°	–90°	H	V	+90°	–90°	H
Inner	0.558	0.551	0.540	0.532	0.720	0.731	0.725	0.695
Middle	0.543	0.562	0.548	0.538	0.707	0.726	0.737	0.709
Outer	0.552	0.548	0.554	0.546	0.720	0.763	0.740	0.717

Table 2.1: *Values of parameter σ_s (V = vertical polarization, H = horizontal polarization).*

2.4 RFI Mitigation

RFI mitigation is accomplished by removing all short accumulations flagged as corrupted by RFI in the conversion from counts to antenna temperature values. All presumed RFI-free short accumulations within a 1.44 s cycle are averaged together and converted to antenna temperatures using the gain and offset,

$$T_F = \frac{1}{g}\left[\left(\frac{1}{N_F}\sum_{n|F_n=0} s_n\right) - o\right] \tag{2.11}$$

where N_F is the number of short accumulation non flagged as corrupted by RFI in the measurement cycle under consideration. The unfiltered antenna temperature is also computed for reference,

$$T_A = \frac{1}{g}\left[\left(\frac{1}{N_F}\sum_{n} s_n\right) - o\right] \tag{2.12}$$

where N is the total number of short accumulation in the measurement cycle under consideration. The calibration gains g and offsets o are computed using the long accumulations and have different values dipending on the particular radiometer and polarization.

2.5 General Definitions and Considerations on Detected RFI

One measure of the amount of detected RFI for a particular 1.44 s data point is the *percentage of detected RFI*, defined as the fraction of acquired 10-ma antenna samples that are flagged as corrupted by RFI within the 1.44 s data point, expressed in %, i.e.,

$$P_{RFI} = 100 \frac{N_{RFI}}{N_{TOT}} \tag{2.13}$$

where N_{RFI} is the number of measurements identified as RFI and N_{TOT} is the total number of measurements averaged to obtain a data point.

As an example, Figure 2.5 is a global map of P_{RFI} over the entire duration of the Aquarius mission. It is an average over 197 values corresponding to the 7-day cycles of the mission, regridded to 0.5° latitude-longitude resolution. Since RFI can be directional and thus affect ascending and descending tracks differently, it is useful to produce ascending- and descending-only maps such as those shown in Figure 2.6.

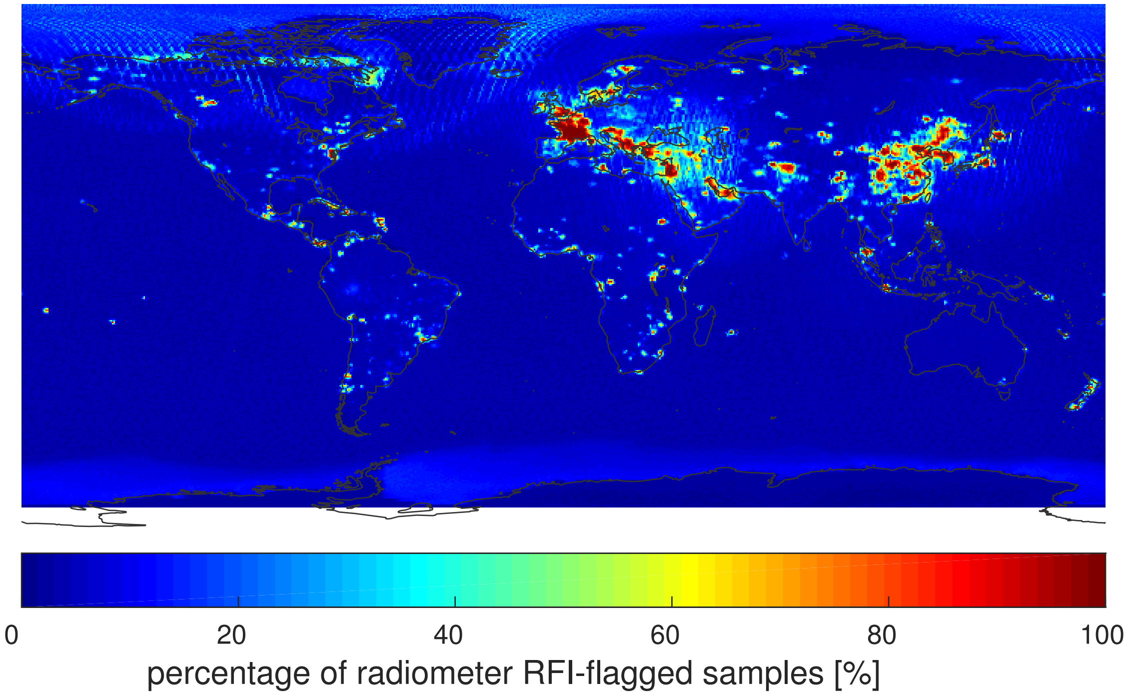


Figure 2.5: *Global map of RFI detected by the Aquarius radiometer RFI, averaged over the entire mission.*

A number of small geographic areas have been identified for further analysis. They include three ocean regions and four land regions assumed to be RFI-free and one ocean region affected by RFI. Their location is shown in Figure 2.7.

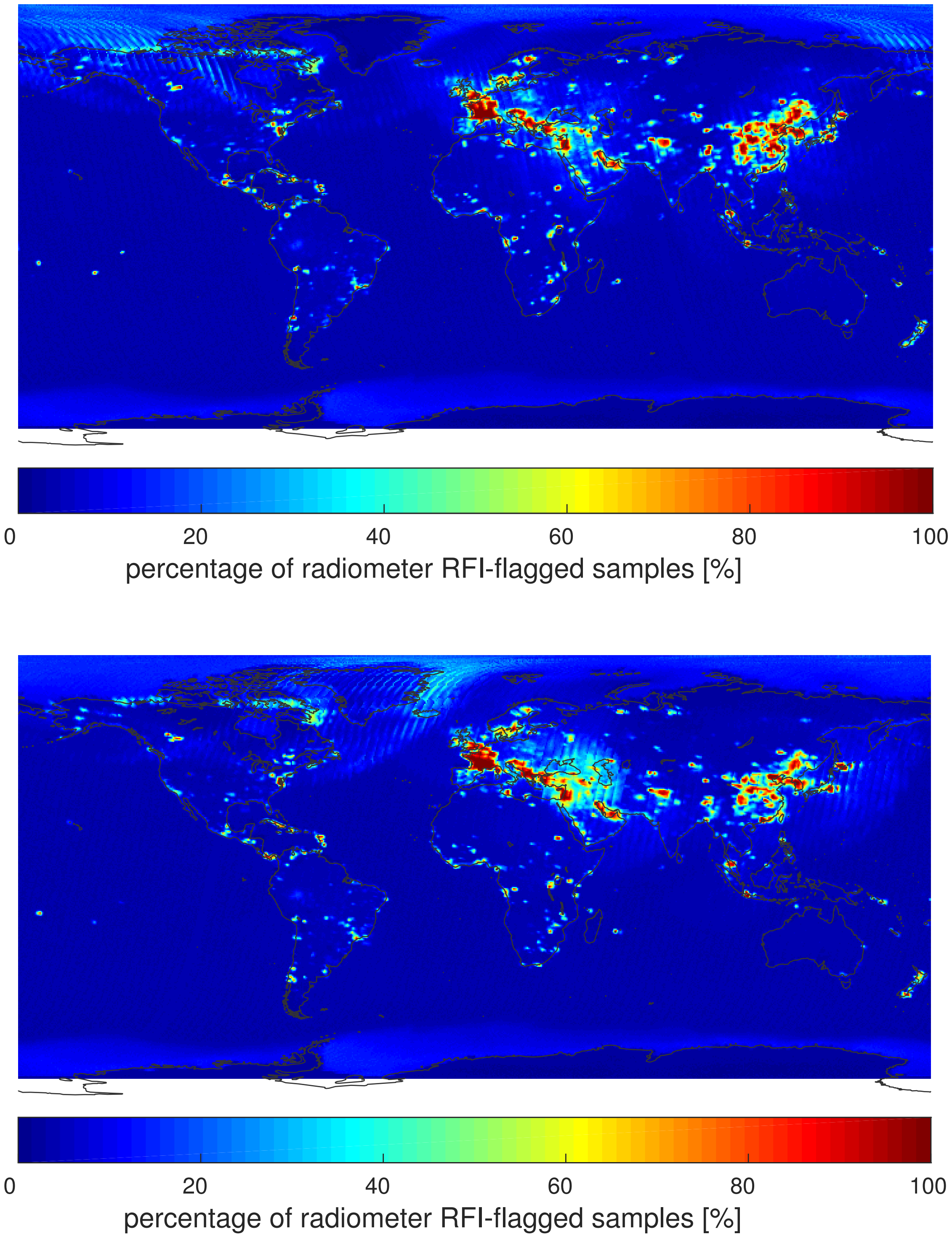


Figure 2.6: *Same as in Figure 2.5, but separately for ascending (top) and descending (bottom).*

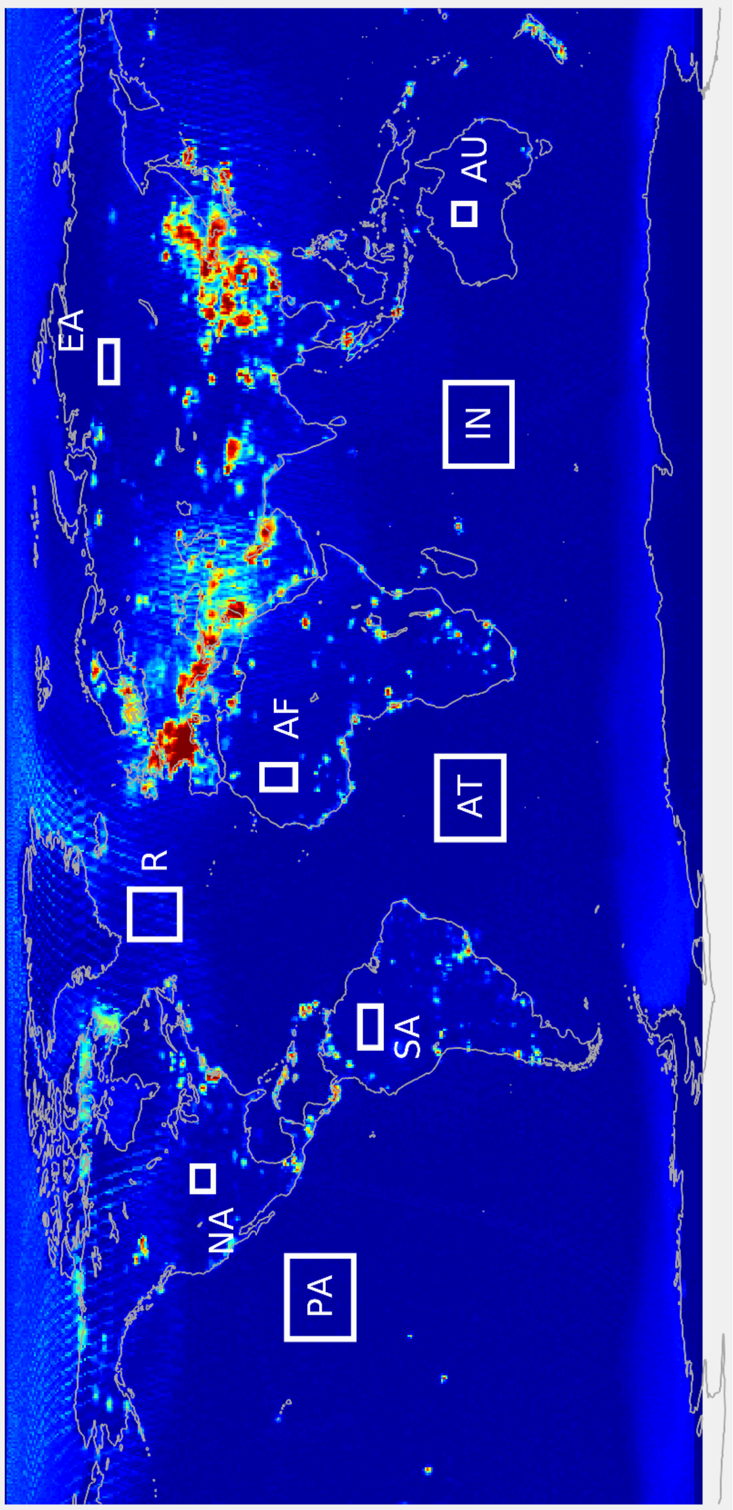


Figure 2.7: *Location of study regions.*

3 RFI Statistics

3.1 General Considerations

Since the instrument is moving along its orbit, measurements in a data stream are not acquired on the same footprint. Even though within a few 10-ms intervals the acquisition may cover partially overlapping footprints given the size of the antenna main lobes, a small variation in the signal will still be added to the statistical fluctuations of the measurements. In order to limit the effect of these physical variations in the observed target (scene), the statistics are computed within a small time interval surrounding each sample. This is the approach followed in the RFI algorithm, where $2\,W_m$ samples surrounding the sample under test are considered in the RFI-flagging process. In addition, other processing steps may introduce further errors. As a results, when trying to estimate the statistical properties of RFI, the following points should noted:

- due to the varying observed scene, the local RFI statistics depend on the size of the time window in which they are computed;
- the processing includes a discretization that introduces a small error;
- the presence of identical 10-ms equivalent samples due to the splitting of the 20-ms short accumulations can also alter the statistics.

These issues will be examined in more detail in the course of this report.

3.2 Local Moments

In this section, the moments of the 10 ms samples are computed within the 1.44 s averaging window. The definitions of mean μ, standard deviation σ, skewness γ and kurtosis κ are as follows:

$$\mu = \mathrm{E}[Q] = \frac{1}{N} \sum_n q_n \tag{3.1}$$

$$\sigma = \mathrm{E}[(Q-\mu)^2]) = \frac{1}{N} \sum_n (q_n - \mu)^2 \tag{3.2}$$

$$\gamma = \frac{\mathrm{E}[(Q-\mu)^3]}{(\mathrm{E}[(S-\mu)^2])^{3/2}} = \frac{\mathrm{E}\left[(Q-\mu)^3\right]}{\sigma^3} = \frac{1}{\sigma^3}\,\frac{1}{N} \sum_n (q_n - \mu)^3 \tag{3.3}$$

$$\kappa = \frac{\mathrm{E}[(Q-\mu)^4]}{(\mathrm{E}[(Q-\mu)^2])^2} = \frac{\mathrm{E}\left[(Q-\mu)^4\right]}{\sigma^4} = \frac{1}{\sigma^4}\,\frac{1}{N} \sum_n (q_n - \mu)^4 \tag{3.4}$$

Q is the random variable associated with the 10-ms samples s_n corrected for the gain and offset

$$q_n = \frac{s_n - o_n}{g_n} \tag{3.5}$$

where g_n and o_n are the gain and offset for 10-ms sample s_n. With this transformation, q_n is an antenna temperature espressed in Kelvin, and 3.1-3.4 are the local moments of the antenna temperature withing the 1.44 s averaging window.

As a result of choosing an averaging window equal to 1.44 s, the mean μ is the same as the antenna temperature T_A or T_F, depending on whether the RFI-flagged samples are included or not, thus only the local standard deviation σ, skewness γ and kurtosis κ are further examined here.

Figure 3.1 shows global maps of σ, γ and κ before RFI mitigation for the period August 1-14, 2013. Values of skewness close to zero and of kurtosis close to three indicate a Gaussian behaviour, while other values would indicate presence of RFI. Areas with high standard deviation appear to correlate well with areas of non-zero skewness and kurtosis different from three, however the latter two seem to be more sensitive and better indicators of RFI. Figure 3.2 shows the same three moments after RFI mitigation. Note that areas in white correspond to 100% RFI, where all the samples where removed. Most values in these maps indicate a Gaussian behaviour in the remaining samples, and suggest that RFI detection and mitigation is effective.

Histograms corresponding to the global maps in Figures 3.1 and 3.2 are shown in Figure 3.3. The standard deviation curves have two peaks for ocean and land. After RFI mitigation, the peaks shift toward lower values since the samples with a higher standard deviation due to RFI have been removed. The skewness has a peak slightly above zero, and RFI mitigation does not significantly change its position. The peaks of kurtosis are centered slightly below three, with RFI mitigation adding a small shift to the left. This difference with what would be expected from Gaussian-distributed samples is most likely due to discretization and is consistent with what has been reported by [Misra (add reference)].

Figure 3.4 shows separate histograms for ocean and land before RFI mitigation. While the peaks of the standard deviation differ for ocean and land, this is not the case for skewness and kurtosis, suggesting that these two parameters may be used as a better indicator of RFI than standard deviation given their relative insensitivity to type of observed scene.

For each of the nine regions in Figure 2.7, the ascending and descending passes (orbit tracks) for the week August 1-7, 2013, were examined, and the variation with time of the RFI-unfiltered antenna temperature T_A, RFI percent, standard deviation, skewness and kurtosis plotted along these tracks. Plots for selected cases are shown in Figures 3.5-3.10, together with a map showing the location of the measurements. Only beam 1 tracks are considered here, but the results can be generalized for beam 2 and 3 as well. For the land regions 1 to 5, the below considerations can be made:

- for all regions the unfiltered antenna temperature T_A do not show any signs of corruption;
- the RFI percent is a slightly higher, but still below 40%, for ascending tracks 1 in regions 1 to 4 and descending track 2 in region 4;
- for ascending tracks 1 in region 1, the highest RFI percent occurences correlate with a high standard deviation and high values of skewness and kurtosis.

For the ocean regions 6 to 9, the following can be observed:

- except for RFI region R1, the unfiltered antenna temperature T_A do not seen affected by RFI;

- the supposedly RFI-free regions 6 to 3 have all very low RFI percent and reasonable values of standard deviation and skewness, but in several cases there are a few spikes above 5 in the kurtosis, which often do not correspond to changes in RFI percent, standard deviation and skewness;
- in the RFI region R1, all parameters indicate the presence of RFI, particularly in thre descending tracks, however not always there is a clear correspondence between high RFI percent and anomalous levels of skewness and kurtosis.

Figures 3.11-3.16 show the histograms of standard deviation, skewness and kurtosis for the same tracks in Figures 3.5-3.10. It can be observed that when RFI is not present, skewness is normally between -1 and 1, and kurtosis values do not exceeed 6, while where there is RFI skewness can exceed 6, and kurtosis values be as high as almost 40 as it can be seen in Figure 3.16.

Table 3.1 lists the mean and standard deviation values of the standard deviation, skewness and kurtosis within the orbit segments contained within the regions defined in Figure 2.7. For the ocean regions with no RFI (6 to 9), the average value of the skewness is around 0.1 with ∼0.3-0.4 standard deviation and that of the kurtosis is around 3.1 with ∼0.6 standard deviation. All quantities here have been computed by excluding outliers falling beyond a threshold of three times the standard deviation to exclude any sample clearly affected by RFI.

As explained in Section 3.2, the two first short accumulations in the acquisition scheme correspond to measurements made over a 20 ms interval rather than 10 ms as the other short accumulations. As a results, these samples may have slightly different statistics. In order to investigate this hypothesis, the plots of standard deviation, skewness and kurtosis have also been generated without using the short accumulation SA2 (short accumulation SA1 had already been excluded as also mentioned in in Section). The results are illustrated in Figure 3.17 for region 6, ascending track 6. The left column show the standard deviation, skewness and kurtosis in different colors with and without accounting for the short accumulation SA2, and the right column the difference between including SA2 or not (i.e., blue curve minus red curve). Surprisingly, removing SA2 increases the standard deviation σ and bring the kurtosis κ below 3, while the skewness γ approaches zero. It is not clear why this happens, but this is likely an effect of having a smaller number of samples in the calculations of these local moments. Table 3.2 is the equivalent of Table 3.1, but the values in it are computed without SA2.

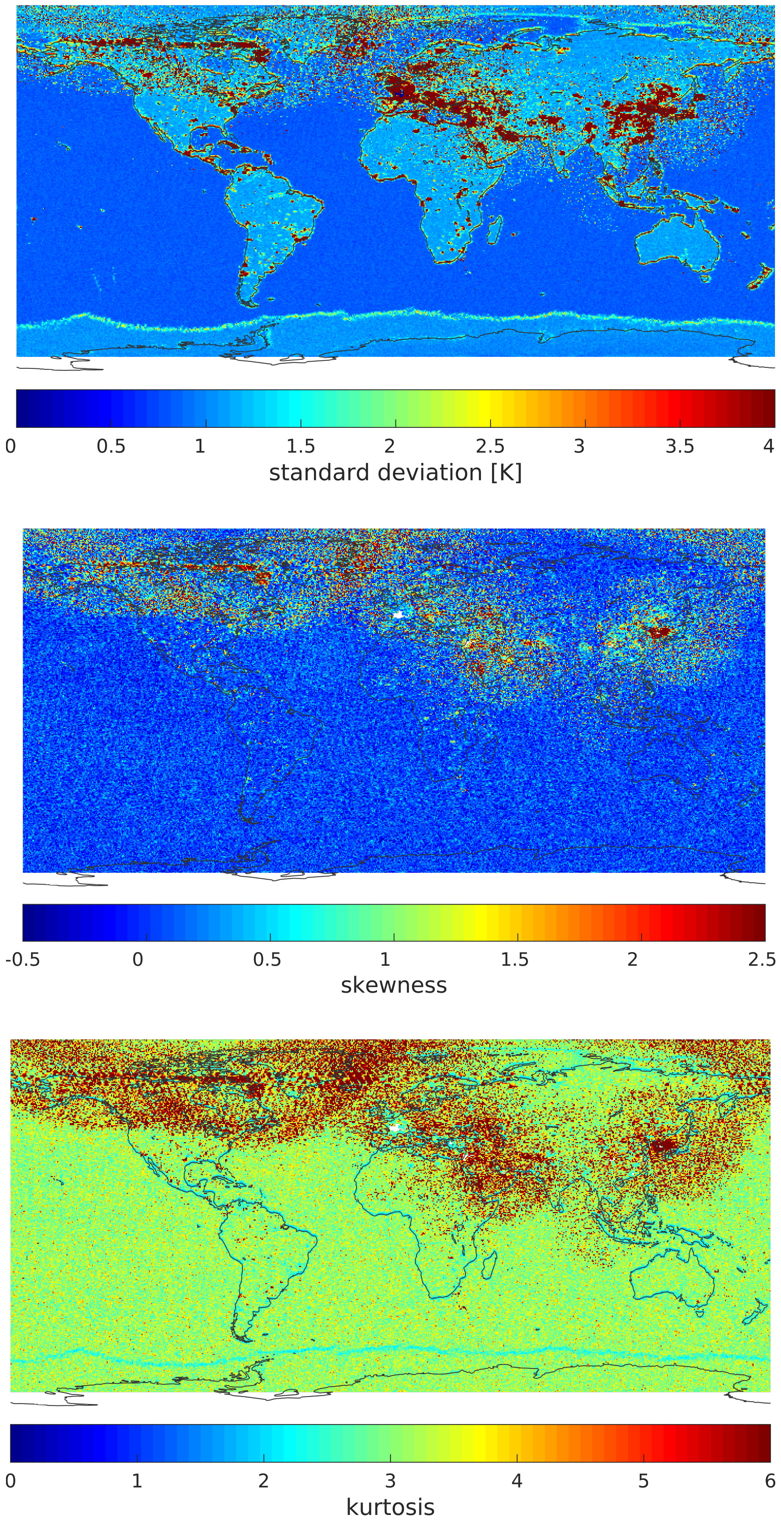


Figure 3.1: *Standard deviation (top), skewness (middle) and kurtosis (bottom) before RFI mitigation.*

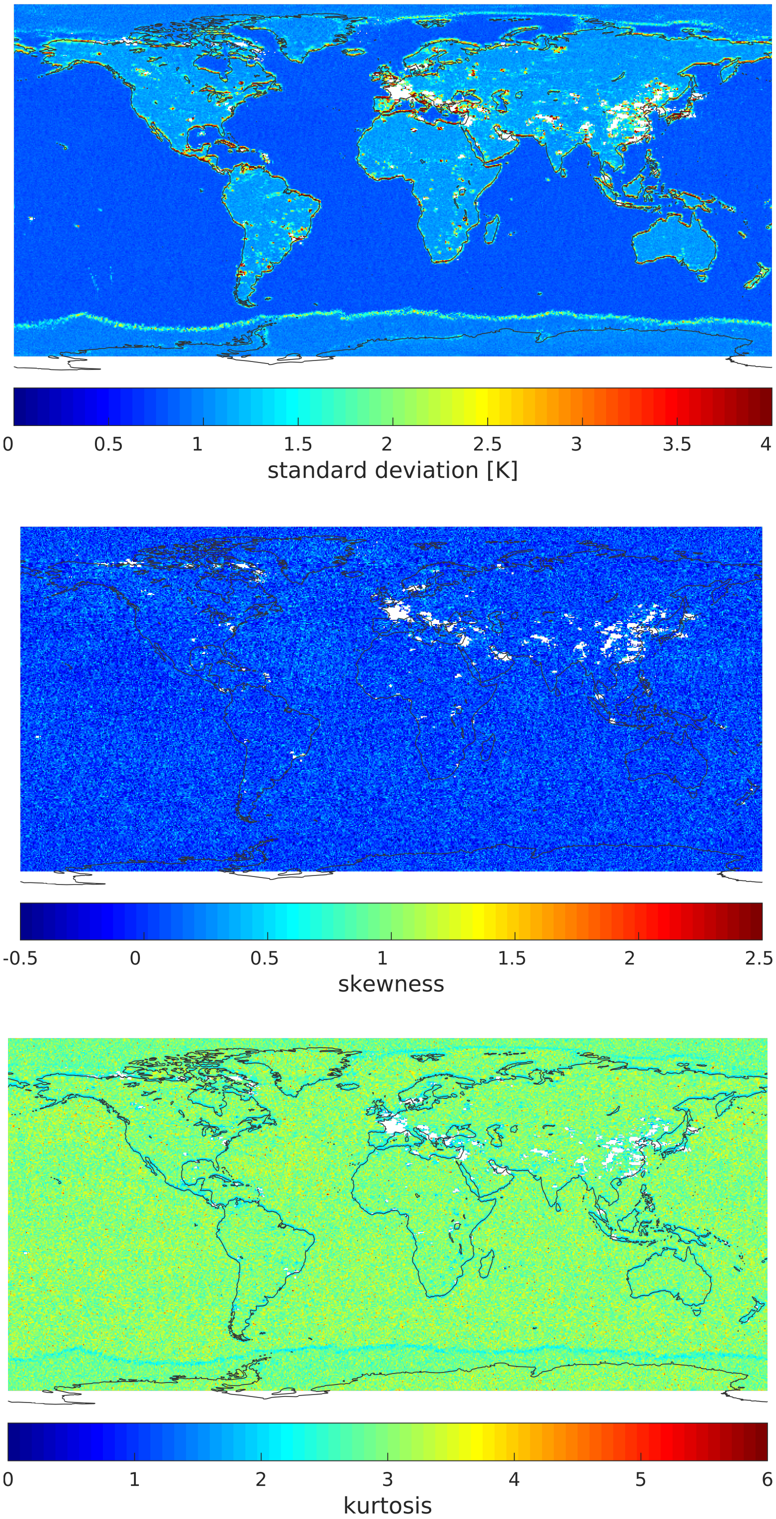


Figure 3.2: *Standard deviation (top), skewness (middle) and kurtosis (bottom) after RFI mitigation.*

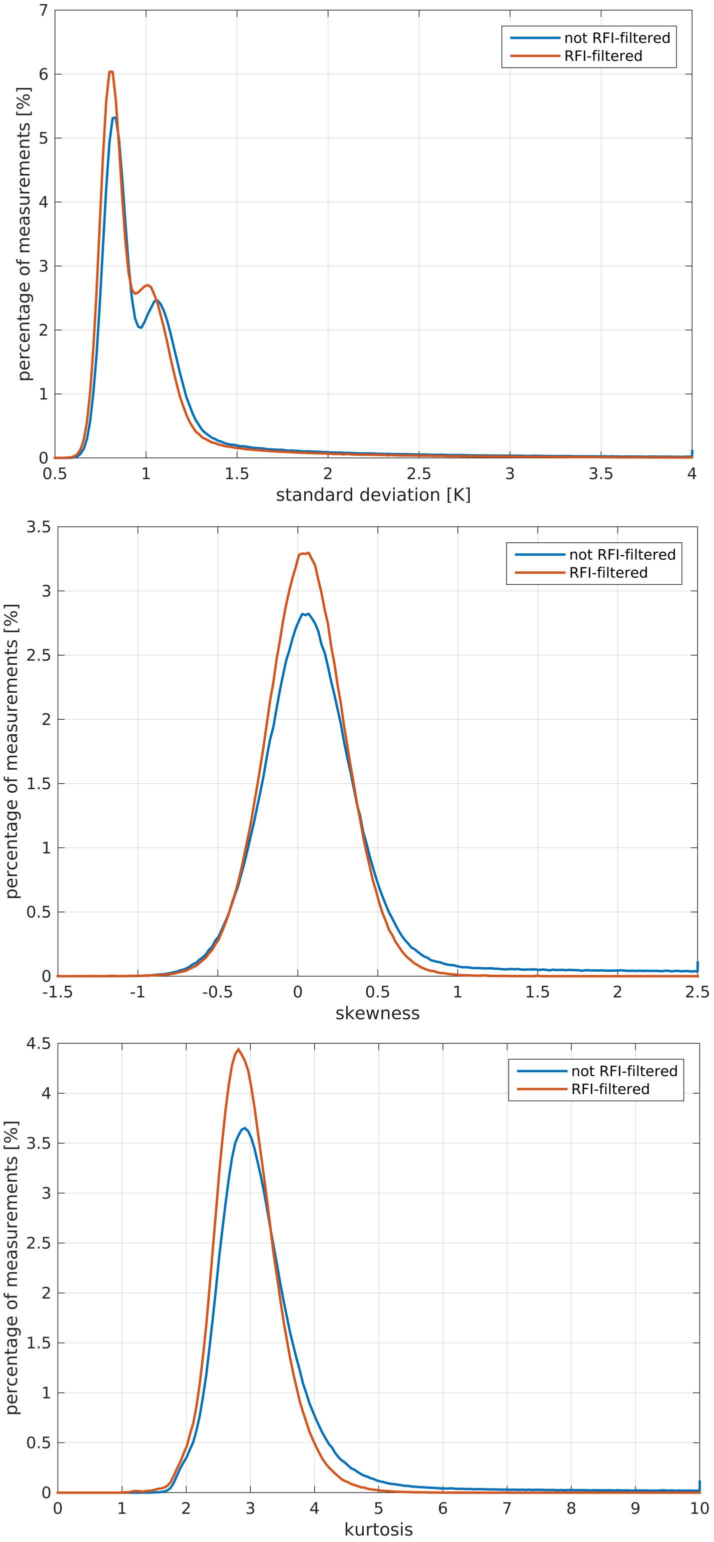


Figure 3.3: *Histograms of standard deviation (top), skewness (middle) and kurtosis (bottom) before and after RFI mitigation.*

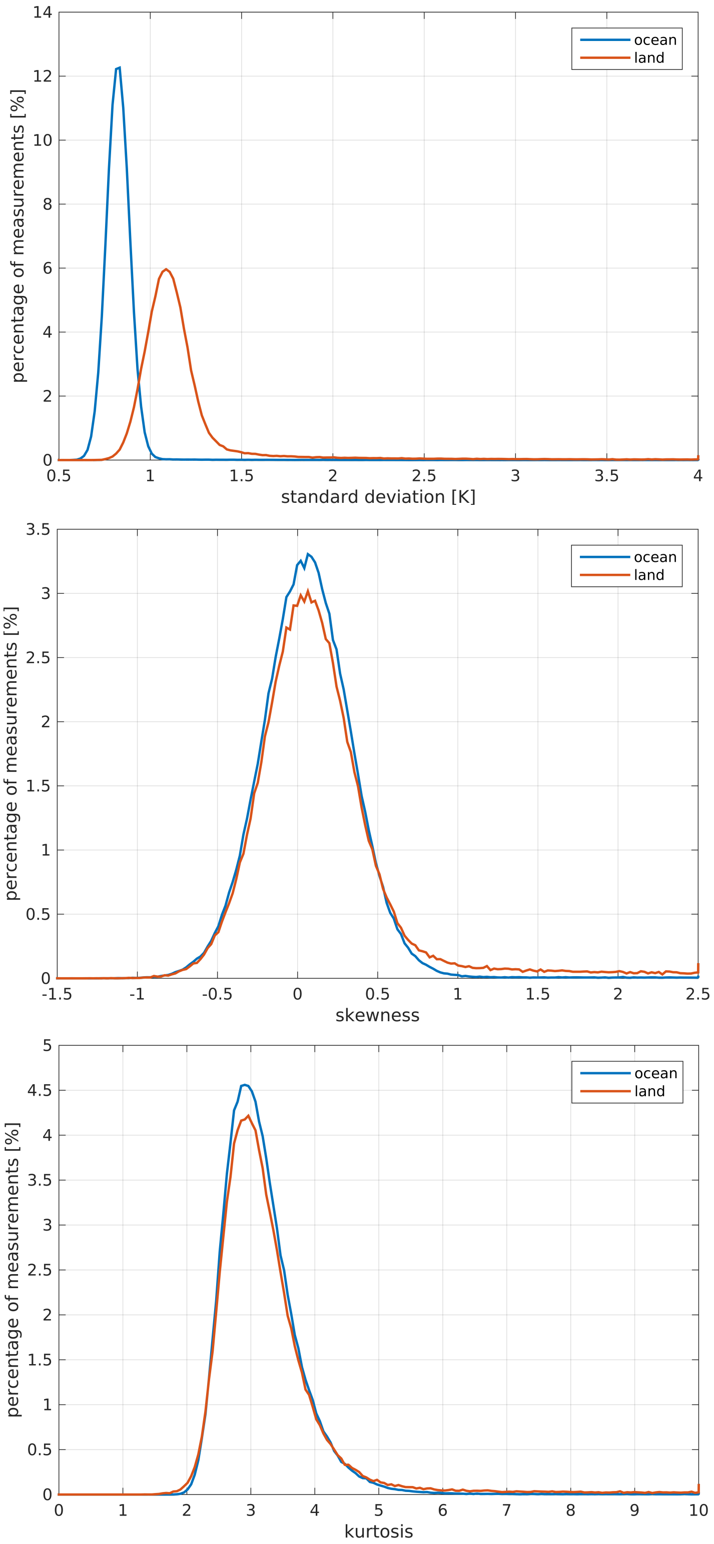
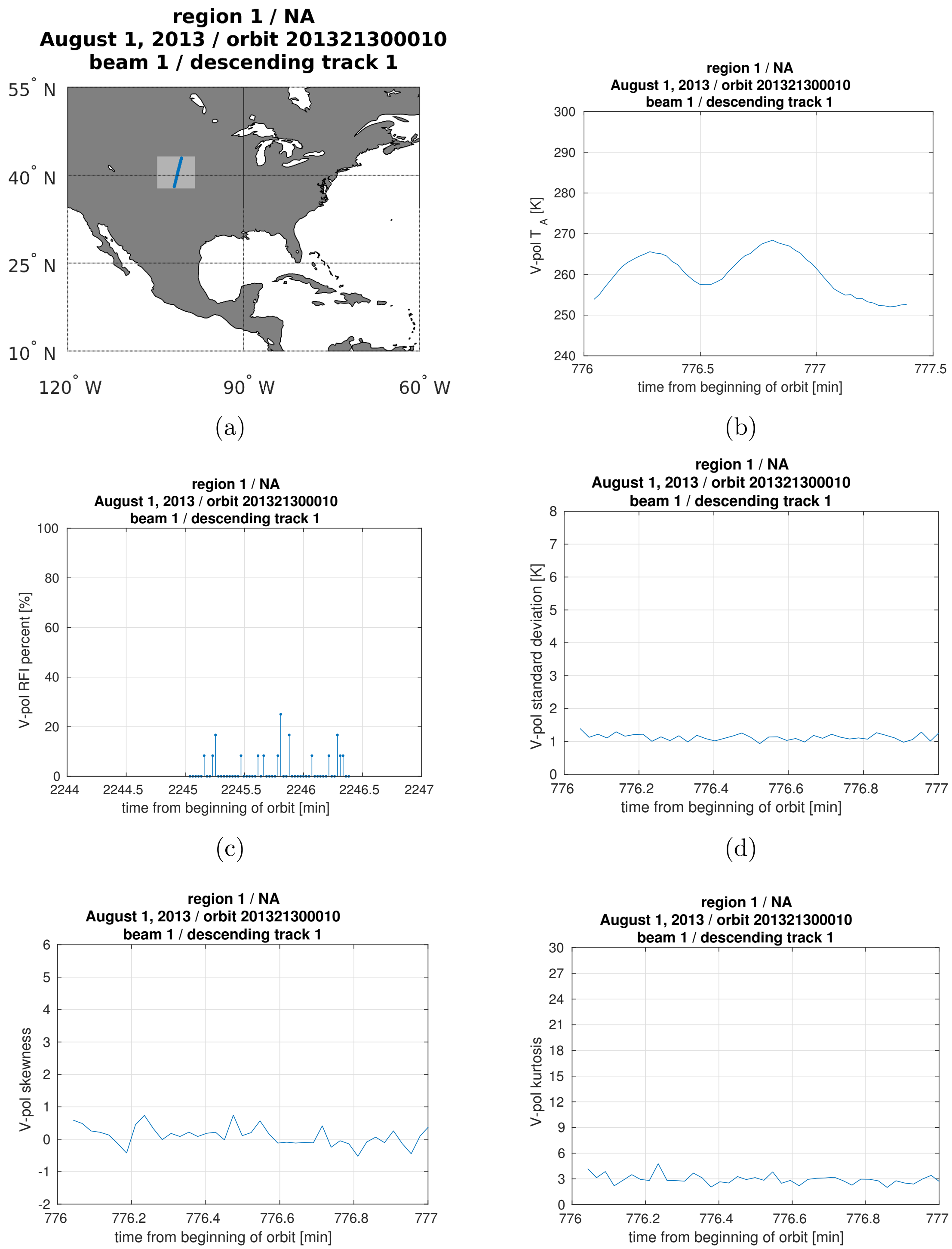
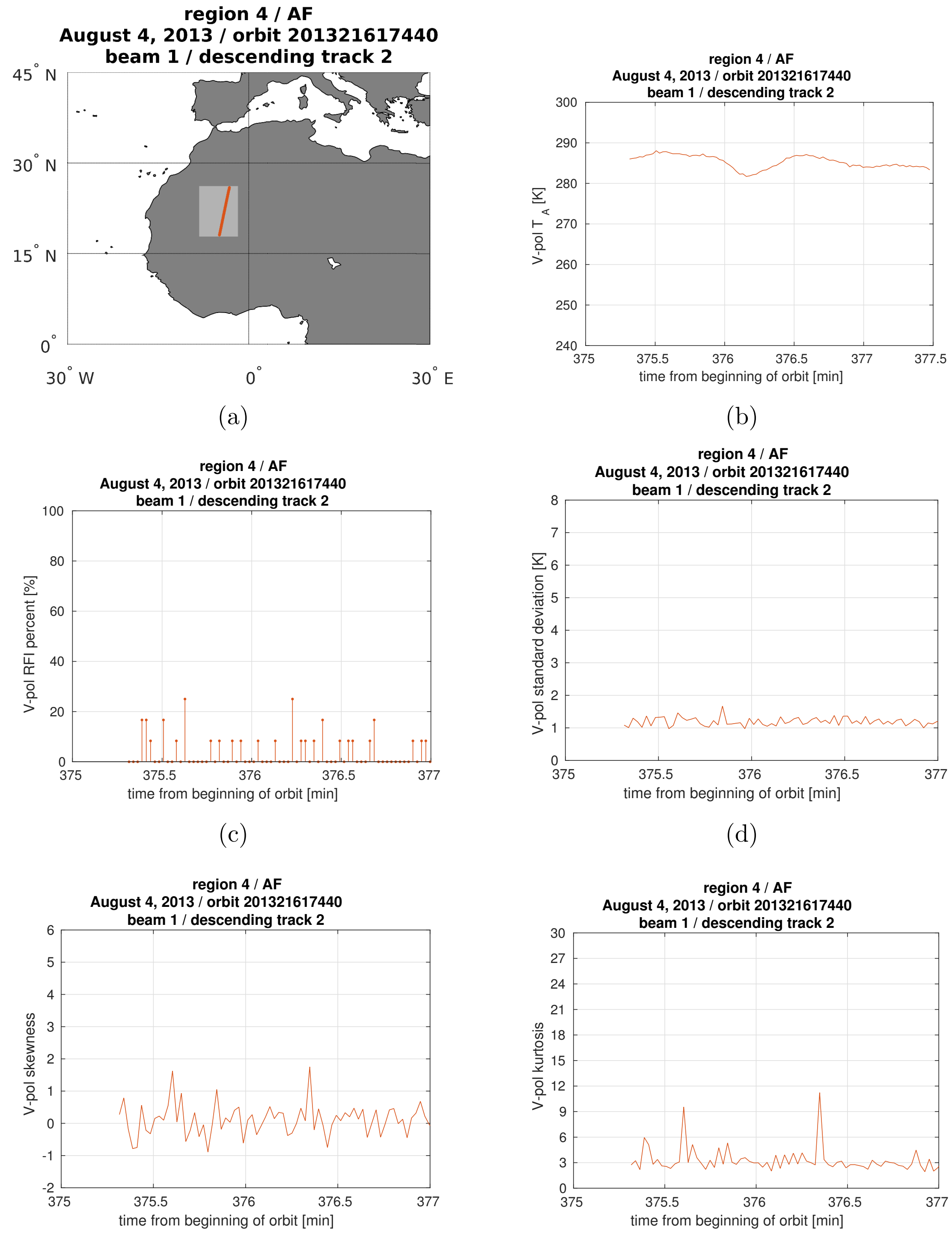


Figure 3.4: *Histograms of standard deviation (top), skewness (middle) and kurtosis (bottom) for ocean- and land-only.*



(a) (b) (c) (d) (e) (f)

Figure 3.5: *Local moments of 10 ms samples in 1.44 s averaging window as function of time for descending orbit track 1 in region 1: (a) track location; (b) antenna temperature T_A; (c) RFI percent; (d) standard deviation; (e) skewness; (f) kurtosis.*



(a) (b)

(c) (d)

(e) (f)

Figure 3.6: *Local moments of 10 ms samples in 1.44 s averaging window as function of time for descending orbit track 2 in region 4: (a) track location; (b) antenna temperature* T_A*; (c) RFI percent; (d) standard deviation; (e) skewness; (f) kurtosis.*

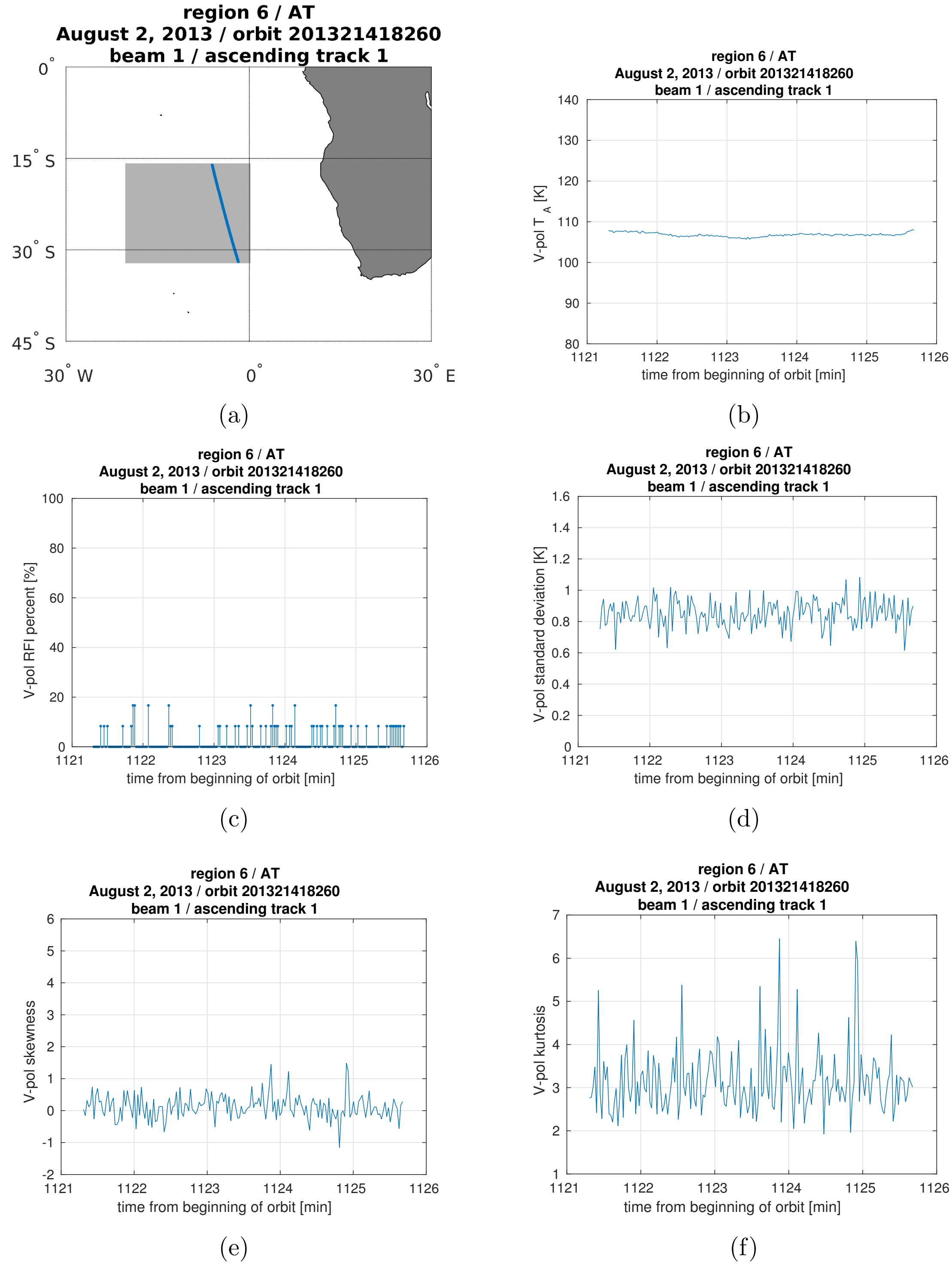
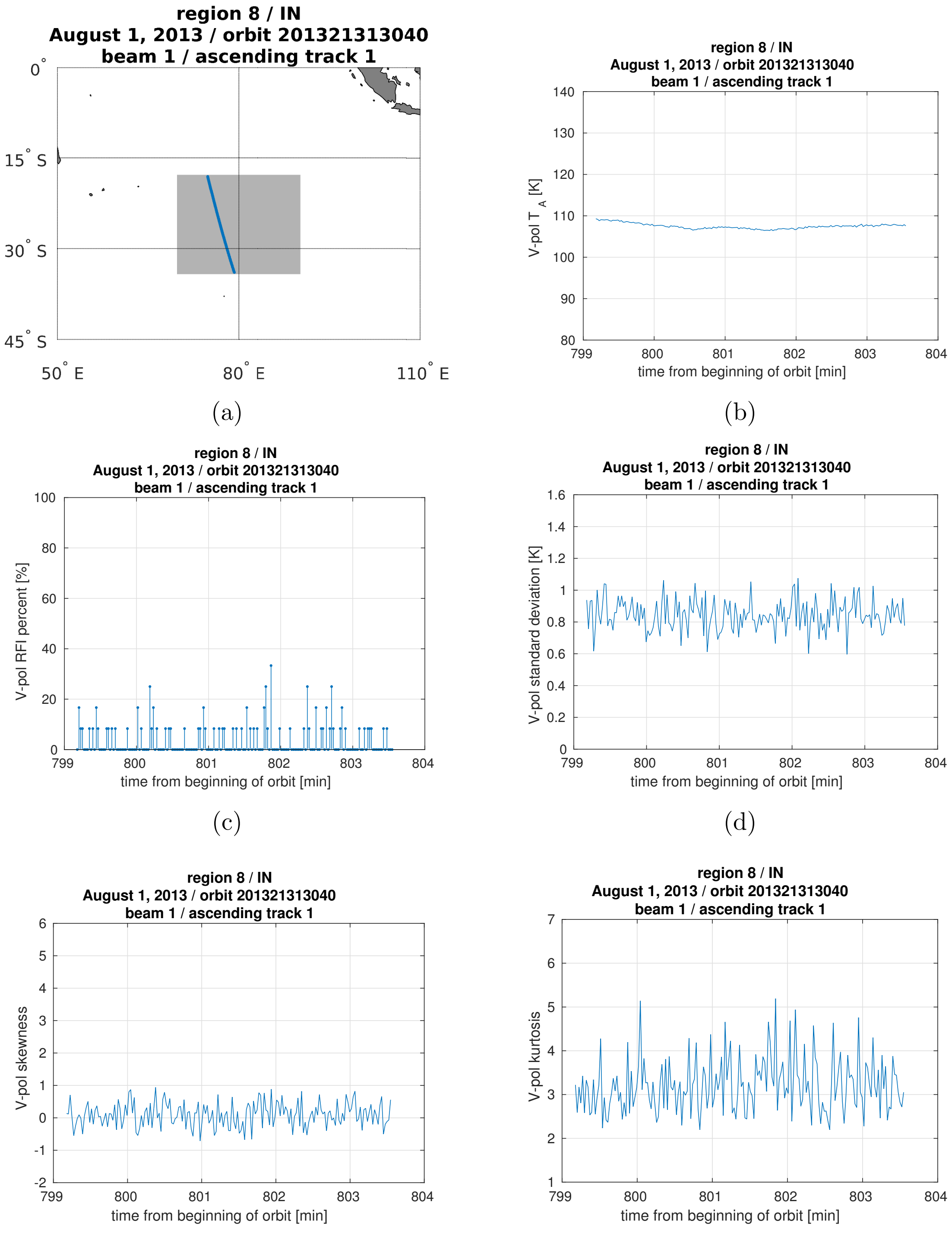
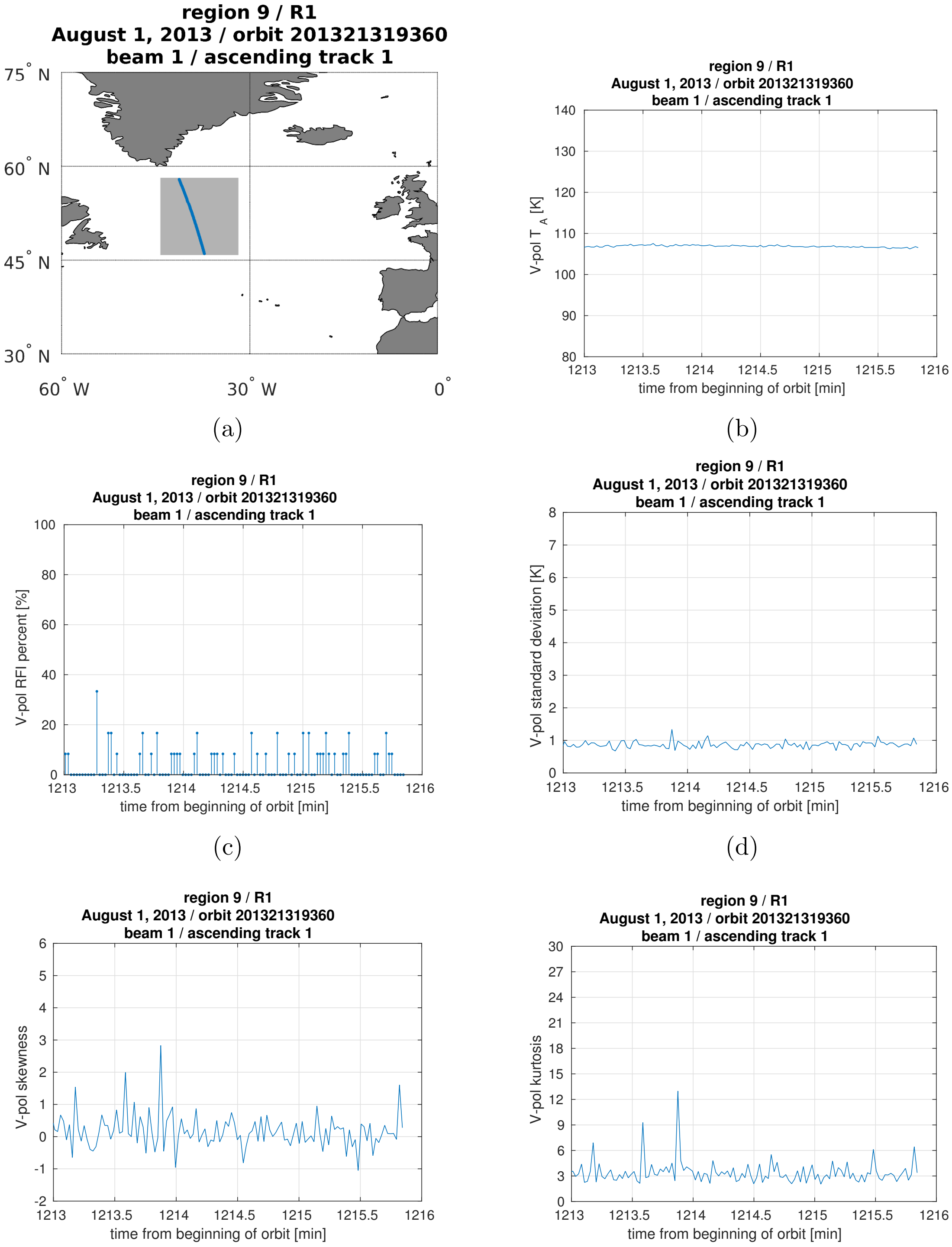


Figure 3.7: *Local moments of 10 ms samples in 1.44 s averaging window as function of time for ascending orbit track 1 in region 6: (a) track location; (b) antenna temperature T_A; (c) RFI percent; (d) standard deviation; (e) skewness; (f) kurtosis.*



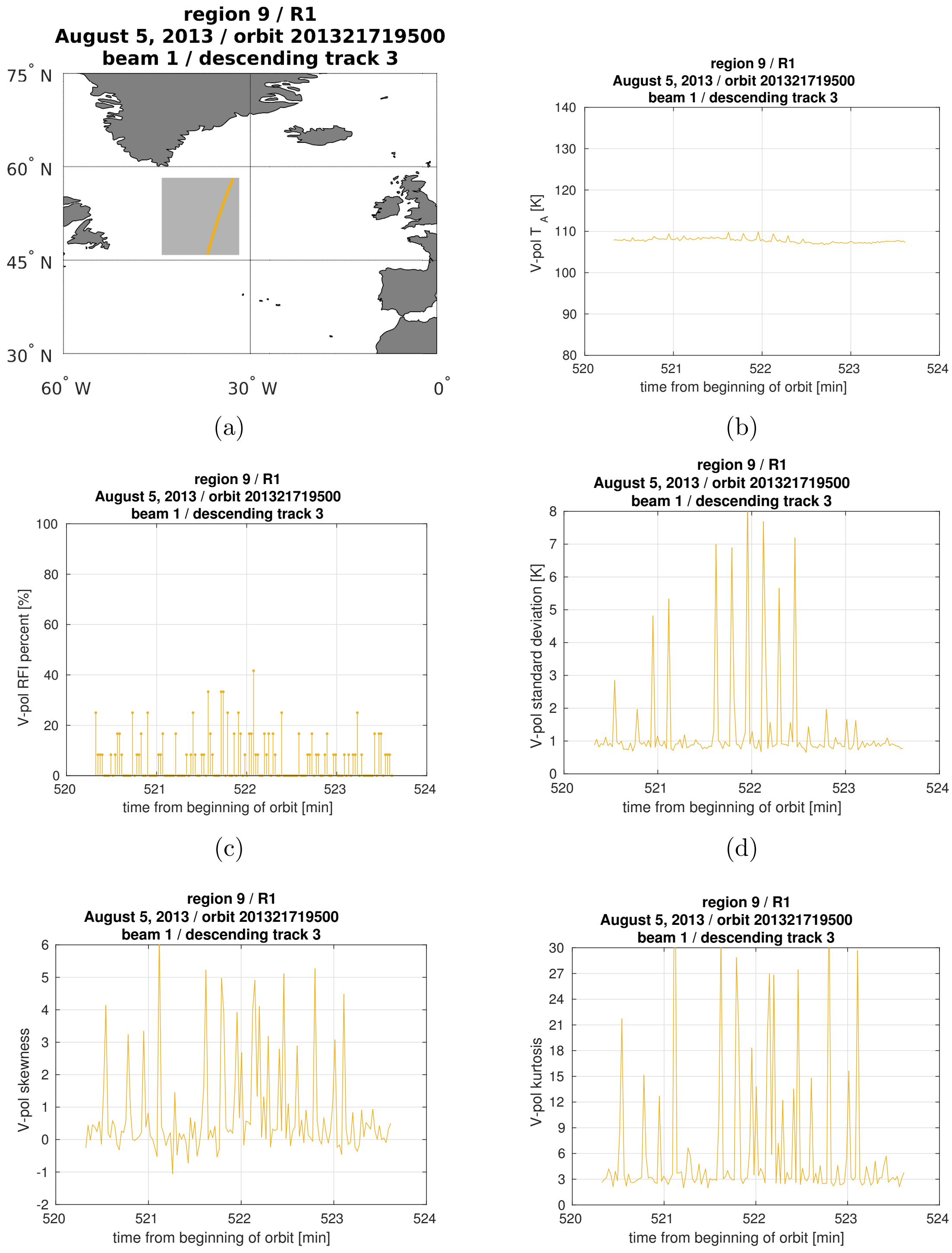
(a) (b) (c) (d) (e) (f)

Figure 3.8: *Local moments of 10 ms samples in 1.44 s averaging window as function of time for ascending orbit track 1 in region 8: (a) track location; (b) antenna temperature T_A; (c) RFI percent; (d) standard deviation; (e) skewness; (f) kurtosis.*



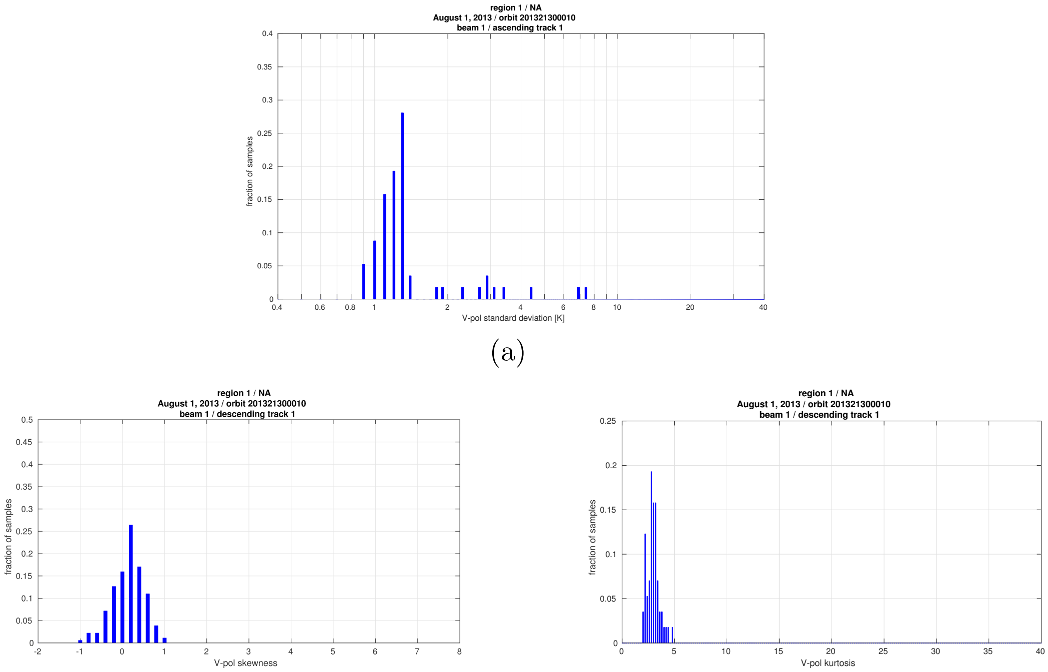
(a) (b) (c) (d) (e) (f)

Figure 3.9: *Local moments of 10 ms samples in 1.44 s averaging window as function of time for ascending orbit track 1 in region 9: (a) track location; (b) antenna temperature T_A; (c) RFI percent; (d) standard deviation; (e) skewness; (f) kurtosis.*



(a) (b) (c) (d) (e) (f)

Figure 3.10: *Local moments of 10 ms samples in 1.44 s averaging window as function of time for descending orbit track 3 in region 9: (a) track location; (b) antenna temperature* T_A*; (c) RFI percent; (d) standard deviation; (e) skewness; (f) kurtosis.*

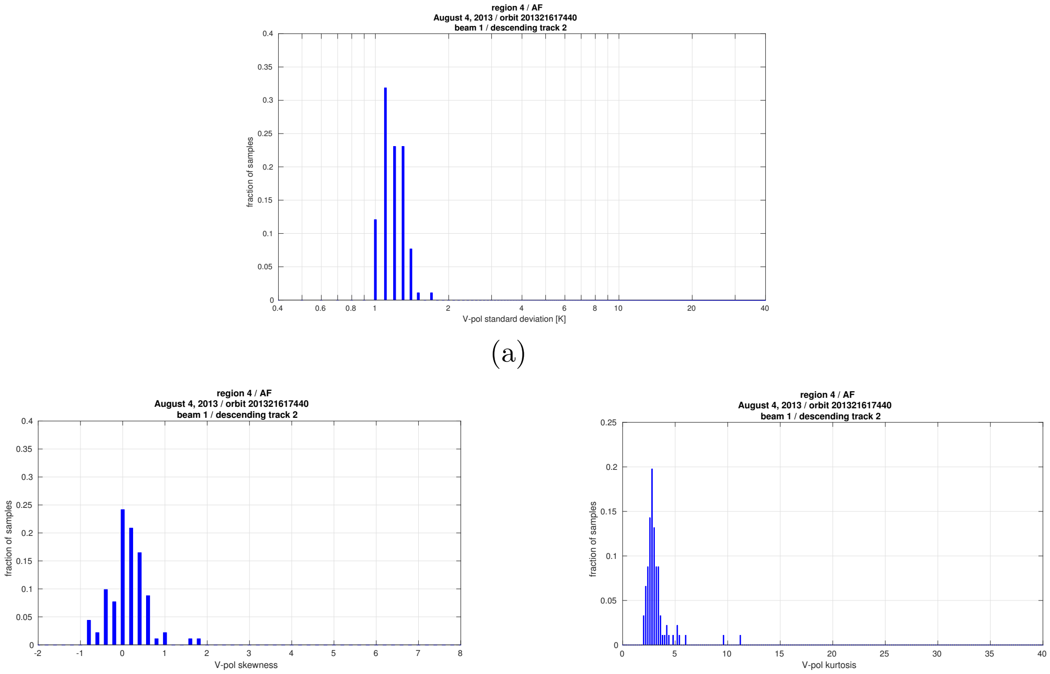


(a)

(b)

(c)

Figure 3.11: *Histograms of (a) standard deviation, (b) skewness and (c) kurtosis for descending orbit track 1 in region 1.*

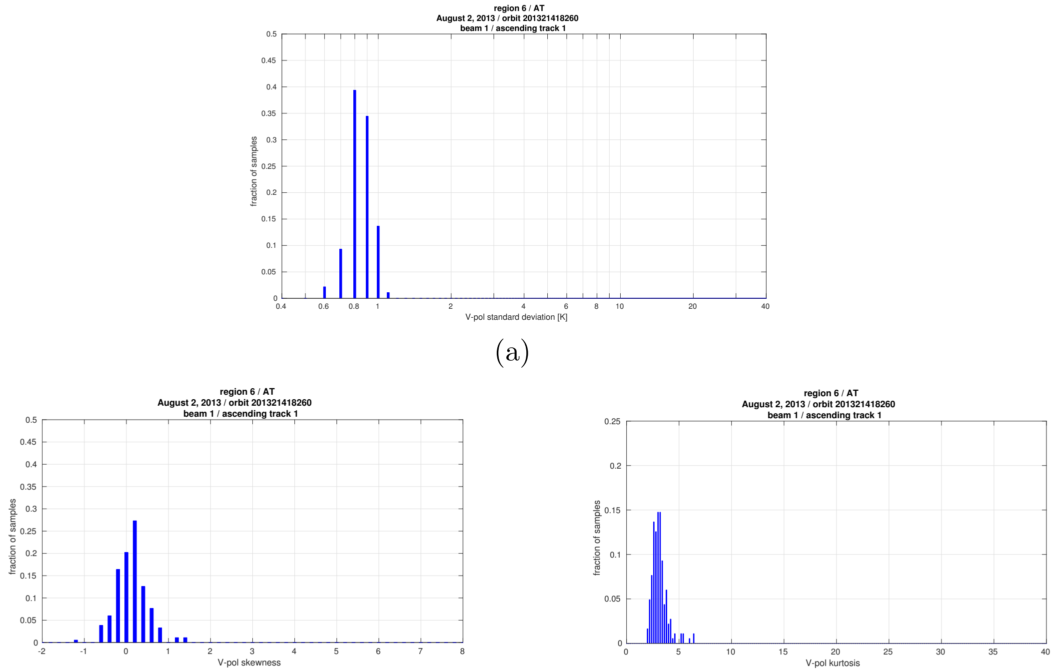


(a)

(b)

(c)

Figure 3.12: *Histograms of (a) standard deviation, (b) skewness and (c) kurtosis for descending orbit track 2 in region 4.*

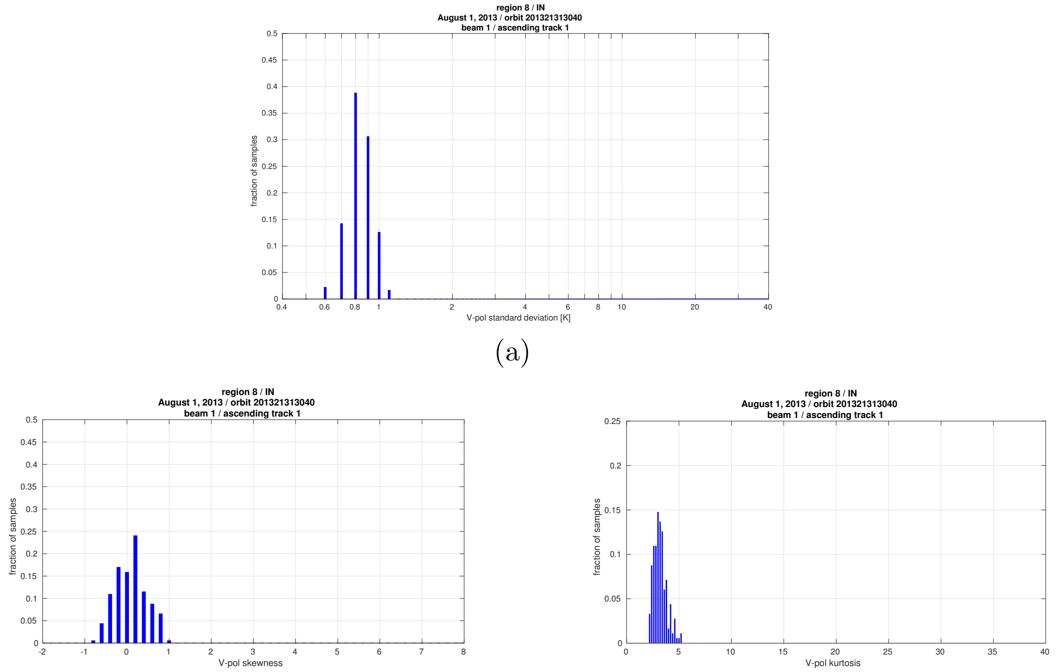


(a)

(b)

(c)

Figure 3.13: *Histograms of (a) standard deviation, (b) skewness and (c) kurtosis for ascending orbit track 1 in region 6.*

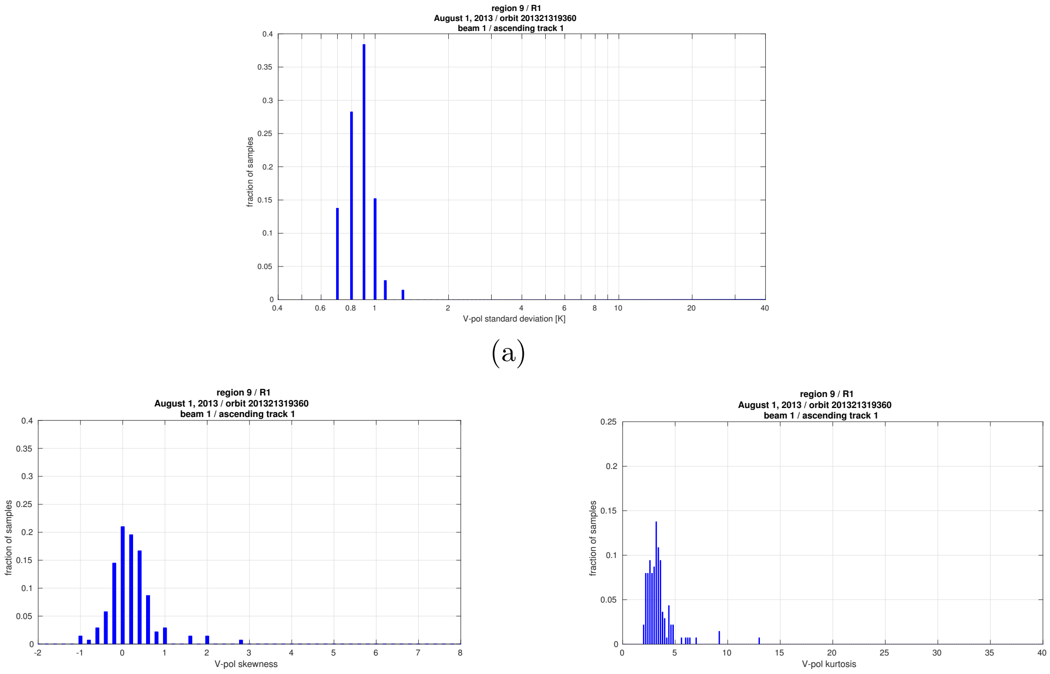


(a)

(b)

(c)

Figure 3.14: *Histograms of (a) standard deviation, (b) skewness and (c) kurtosis for ascending orbit track 1 in region 8.*



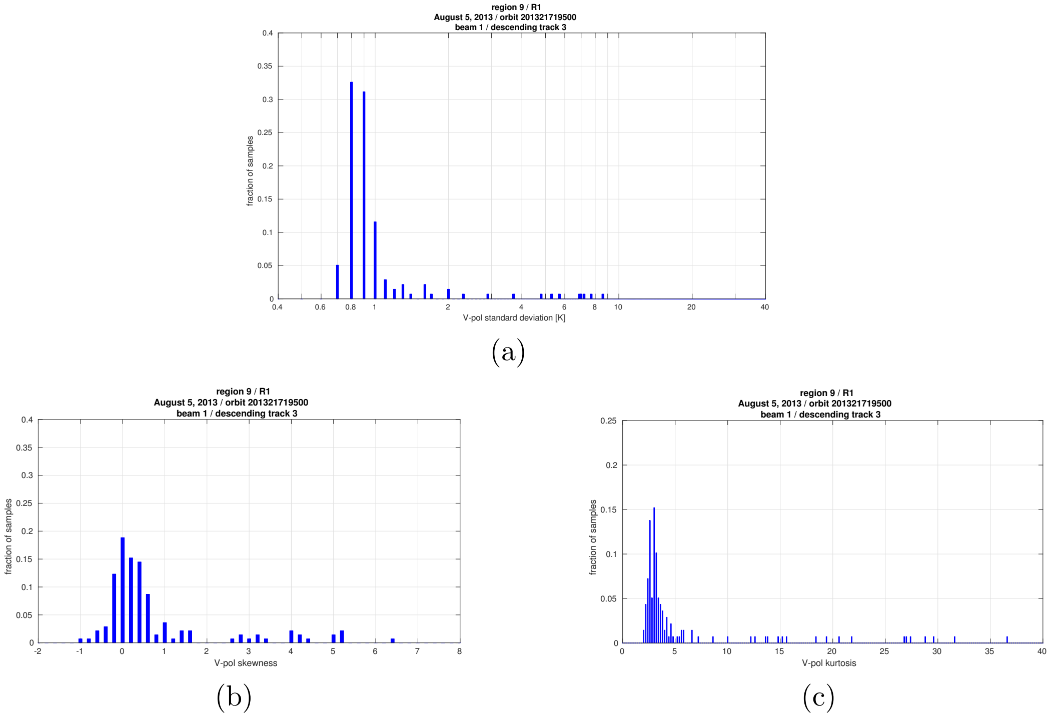
region 9 / R1
August 1, 2013 / orbit 201321319360
beam 1 / ascending track 1

(a)

(b)

(c)

Figure 3.15: *Histograms of (a) standard deviation, (b) skewness and (c) kurtosis for ascending orbit track 1 in region 9.*



region 9 / R1
August 5, 2013 / orbit 201321719500
beam 1 / descending track 3

(a)

(b)

(c)

Figure 3.16: *Histograms of (a) standard deviation, (b) skewness and (c) kurtosis for descending orbit track 3 in region 9.*

		σ [K]		γ		κ	
region	orbit track #	mean	standard dev.	mean	standard dev.	mean	standard dev.
1 - NA	asc 1	1.45	0.68	0.43	1.02	4.48	4.46
1 - NA	asc 2	1.10	0.14	0.02	0.39	3.30	0.72
1 - NA	desc 1	1.11	0.11	0.09	0.32	2.93	0.53
2 - SA	asc 1	1.22	0.40	0.12	0.44	3.56	3.01
2 - SA	asc 2	1.20	0.18	-0.01	0.37	3.01	0.61
2 - SA	asc 3	1.12	0.13	0.08	0.41	3.10	0.72
2 - SA	desc 1	1.15	0.14	0.07	0.33	3.08	0.58
2 - SA	desc 2	1.11	0.10	0.02	0.35	3.09	0.66
2 - SA	desc 3	1.12	0.12	0.05	0.30	3.01	0.63
3 - EA	asc 1	1.19	0.13	-0.01	0.29	2.89	0.53
3 - EA	asc 2	1.19	0.14	0.01	0.26	2.74	0.45
3 - EA	asc 3	1.12	0.12	0.16	0.35	3.12	0.74
3 - EA	desc 1	1.09	0.12	-0.00	0.37	3.30	0.69
3 - EA	desc 2	1.06	0.12	0.10	0.39	3.10	0.59
3 - EA	desc 3	1.08	0.12	0.07	0.33	3.22	0.68
4 - AF	asc 1	1.14	0.12	0.10	0.38	3.09	0.75
4 - AF	asc 2	1.16	0.12	-0.02	0.38	3.13	0.65
4 - AF	desc 1	1.17	0.11	0.07	0.43	3.06	0.61
4 - AF	desc 2	1.19	0.11	0.09	0.38	3.02	0.73
5 - AU	asc 1	1.21	0.17	0.02	0.31	3.11	0.58
5 - AU	asc 2	1.12	0.10	0.08	0.30	3.03	0.56
5 - AU	desc 1	1.08	0.11	0.08	0.40	3.34	0.72
6 - AT	asc 1	0.85	0.09	0.11	0.34	3.08	0.61
6 - AT	asc 2	0.85	0.08	0.11	0.34	3.10	0.61
6 - AT	asc 3	0.83	0.08	0.13	0.34	3.08	0.63
6 - AT	asc 4	0.86	0.08	0.13	0.29	3.04	0.64
6 - AT	asc 5	0.86	0.09	0.11	0.36	3.06	0.68
6 - AT	asc 6	0.86	0.08	0.13	0.35	3.10	0.63
6 - AT	desc 1	0.83	0.08	0.09	0.38	3.06	0.55
6 - AT	desc 2	0.82	0.09	0.15	0.35	3.04	0.65
6 - AT	desc 3	0.83	0.08	0.12	0.33	3.11	0.68
6 - AT	desc 4	0.83	0.08	0.09	0.35	3.09	0.62
6 - AT	desc 5	0.85	0.09	0.12	0.38	3.05	0.64
7 - PA	asc 1	0.87	0.08	0.09	0.37	3.13	0.68
7 - PA	asc 2	0.86	0.09	0.14	0.34	3.01	0.59
7 - PA	asc 3	0.84	0.09	0.10	0.36	3.11	0.69
7 - PA	asc 4	0.88	0.09	0.08	0.32	3.09	0.65
7 - PA	asc 5	0.85	0.09	0.06	0.34	3.05	0.63
7 - PA	asc 6	0.87	0.09	0.11	0.39	3.16	0.80
7 - PA	desc 1	0.85	0.09	0.13	0.36	3.11	0.57
7 - PA	desc 2	0.84	0.08	0.13	0.35	3.04	0.59
7 - PA	desc 3	0.85	0.10	0.07	0.33	3.10	0.67
7 - PA	desc 4	0.87	0.08	0.13	0.36	2.99	0.62
7 - PA	desc 5	0.86	0.09	0.06	0.31	3.02	0.63
7 - PA	desc 6	0.85	0.08	0.15	0.35	3.02	0.63
8 - IN	asc 1	0.84	0.10	0.10	0.37	3.18	0.59
8 - IN	asc 2	0.86	0.09	0.08	0.34	3.04	0.64
8 - IN	asc 3	0.86	0.10	0.06	0.36	3.08	0.64
8 - IN	asc 4	0.85	0.09	0.02	0.37	3.04	0.64
8 - IN	asc 5	0.86	0.09	0.02	0.35	3.02	0.60
8 - IN	desc 1	0.84	0.08	0.09	0.37	3.09	0.65
8 - IN	desc 2	0.86	0.09	0.12	0.33	3.03	0.62
8 - IN	desc 3	0.85	0.09	0.10	0.37	3.18	0.61
8 - IN	desc 4	0.86	0.08	0.09	0.33	3.10	0.66
8 - IN	desc 5	0.86	0.08	0.12	0.36	3.12	0.62
9 - R	asc 1	0.87	0.10	0.15	0.42	3.26	0.88
9 - R	asc 2	0.87	0.10	0.24	0.46	3.48	1.17
9 - R	asc 3	0.88	0.17	0.25	0.39	3.15	0.85
9 - R	desc 1	0.92	0.22	0.40	0.77	3.92	2.88
9 - R	desc 2	0.98	0.48	0.41	0.88	4.06	3.07
9 - R	desc 3	1.04	0.63	0.62	1.21	4.98	5.09

Table 3.1: *Mean and standard deviation of standard deviation σ, skewness γ and kurtosis for κ for all the segments of orbit tracks contained within the regions defined in Figure 2.7.*

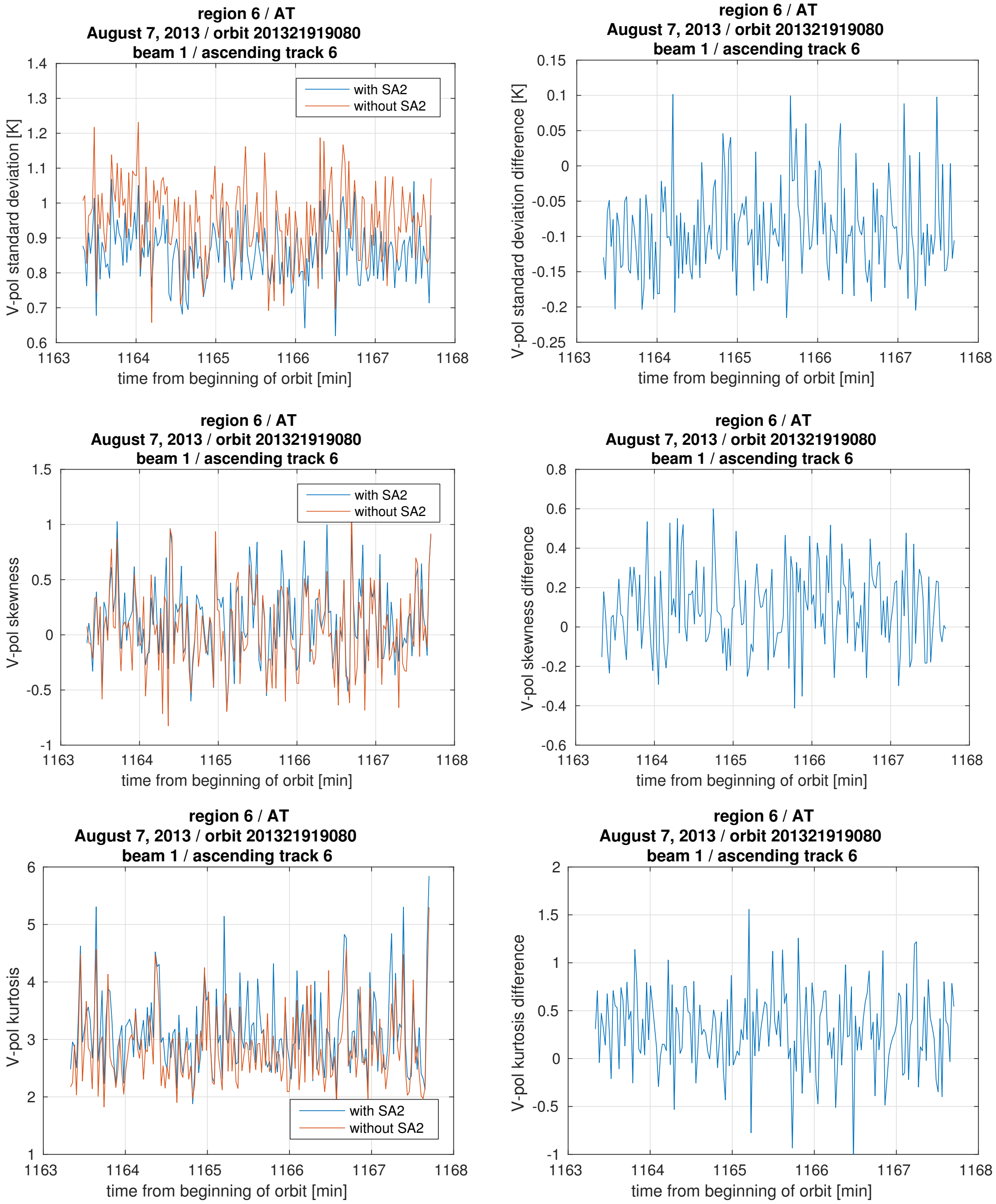


Figure 3.17: *Effect of removing not using short accumulation SA2 in local moment computation: (top) standard deviation, (middle) skewness and (bottom) kurtosis for ascending orbit track 6 in region 6.*

		σ [K]		γ		κ	
region	orbit track #	mean	standard dev.	mean	standard dev.	mean	standard dev.
1 - NA	asc 1	1.61	0.77	0.34	0.87	3.70	2.85
1 - NA	asc 2	1.24	0.17	-0.05	0.37	2.93	0.63
1 - NA	desc 1	1.22	0.12	0.05	0.34	2.63	0.43
2 - SA	asc 1	1.35	0.53	-0.00	0.42	2.88	0.74
2 - SA	asc 2	1.31	0.21	-0.02	0.37	2.80	0.49
2 - SA	asc 3	1.24	0.16	0.02	0.43	2.92	0.75
2 - SA	desc 1	1.28	0.17	-0.03	0.36	2.74	0.49
2 - SA	desc 2	1.23	0.13	-0.01	0.33	2.74	0.54
2 - SA	desc 3	1.25	0.16	0.02	0.34	2.74	0.58
3 - EA	asc 1	1.31	0.16	-0.10	0.32	2.68	0.55
3 - EA	asc 2	1.30	0.18	-0.02	0.25	2.45	0.42
3 - EA	asc 3	1.24	0.16	0.09	0.29	2.70	0.57
3 - EA	desc 1	1.22	0.17	-0.03	0.40	3.00	0.62
3 - EA	desc 2	1.16	0.15	0.05	0.42	2.94	0.72
3 - EA	desc 3	1.20	0.17	-0.02	0.32	2.86	0.56
4 - AF	asc 1	1.26	0.15	0.06	0.40	2.80	0.73
4 - AF	asc 2	1.29	0.16	-0.06	0.33	2.84	0.64
4 - AF	desc 1	1.28	0.16	0.02	0.36	2.71	0.56
4 - AF	desc 2	1.29	0.16	0.07	0.39	2.76	0.60
5 - AU	asc 1	1.33	0.18	0.00	0.34	2.87	0.50
5 - AU	asc 2	1.26	0.15	0.05	0.33	2.73	0.54
5 - AU	desc 1	1.22	0.14	0.02	0.38	2.92	0.65
6 - AT	asc 1	0.93	0.11	0.03	0.36	2.80	0.47
6 - AT	asc 2	0.94	0.11	0.05	0.33	2.77	0.60
6 - AT	asc 3	0.92	0.10	0.04	0.36	2.80	0.56
6 - AT	asc 4	0.96	0.10	0.06	0.28	2.68	0.54
6 - AT	asc 5	0.95	0.11	0.04	0.36	2.80	0.60
6 - AT	asc 6	0.95	0.11	0.04	0.36	2.83	0.61
6 - AT	desc 1	0.92	0.11	0.02	0.38	2.78	0.58
6 - AT	desc 2	0.91	0.12	0.07	0.37	2.74	0.54
6 - AT	desc 3	0.92	0.10	0.02	0.31	2.74	0.51
6 - AT	desc 4	0.92	0.09	0.03	0.35	2.83	0.54
6 - AT	desc 5	0.93	0.11	0.05	0.37	2.76	0.55
7 - PA	asc 1	0.97	0.11	0.04	0.39	2.79	0.54
7 - PA	asc 2	0.95	0.12	0.06	0.34	2.71	0.55
7 - PA	asc 3	0.93	0.11	0.02	0.38	2.81	0.61
7 - PA	asc 4	0.97	0.12	-0.03	0.34	2.79	0.56
7 - PA	asc 5	0.95	0.12	-0.01	0.35	2.77	0.58
7 - PA	asc 6	0.97	0.11	0.06	0.38	2.75	0.66
7 - PA	desc 1	0.94	0.12	0.07	0.34	2.75	0.46
7 - PA	desc 2	0.94	0.11	0.05	0.39	2.81	0.62
7 - PA	desc 3	0.94	0.12	-0.02	0.39	2.86	0.60
7 - PA	desc 4	0.96	0.11	0.05	0.36	2.70	0.54
7 - PA	desc 5	0.95	0.11	0.02	0.30	2.72	0.54
7 - PA	desc 6	0.94	0.11	0.08	0.34	2.74	0.54
8 - IN	asc 1	0.93	0.11	0.03	0.38	2.88	0.56
8 - IN	asc 2	0.96	0.11	0.00	0.35	2.75	0.55
8 - IN	asc 3	0.95	0.11	-0.02	0.37	2.83	0.59
8 - IN	asc 4	0.94	0.12	-0.04	0.35	2.81	0.65
8 - IN	asc 5	0.95	0.12	-0.05	0.35	2.76	0.58
8 - IN	desc 1	0.93	0.10	0.03	0.36	2.79	0.58
8 - IN	desc 2	0.94	0.11	0.04	0.33	2.78	0.57
8 - IN	desc 3	0.95	0.12	0.03	0.38	2.86	0.55
8 - IN	desc 4	0.96	0.11	0.04	0.34	2.79	0.55
8 - IN	desc 5	0.96	0.10	0.06	0.34	2.79	0.53
9 - R	asc 1	0.95	0.12	0.08	0.40	2.81	0.59
9 - R	asc 2	0.96	0.12	0.13	0.43	3.05	0.96
9 - R	asc 3	0.95	0.12	0.14	0.35	2.81	0.71
9 - R	desc 1	0.99	0.24	0.22	0.60	3.16	1.55
9 - R	desc 2	1.04	0.38	0.23	0.67	3.43	2.16
9 - R	desc 3	1.13	0.67	0.42	1.05	3.99	3.52

Table 3.2: *Mean and standard deviation of standard deviation σ, skewness γ and kurtosis for κ as in Table reftab:3.2, but computed excluding short accumulation SA2.*

3.3 Histograms and Probability Density Functions

3.3.1 Motivation

In this section, an analysis of the histograms and probability density functions of the acquired 10 ms samples will be carried out. The ultimate goal is to find a way to estimate the statistics of low-level RFI still present after RFI mitigation and remove it.

The antenna temperature T_A measured when the main beam is centered in the direction (θ_o, ϕ_o) is given by

$$T_{A,obs}(\theta_o, \phi_o) = \int_{FOV} G(\theta, \phi) T_B(\theta, \phi) d\Omega \tag{3.6}$$

where G is the antenna gain, T_B is the brightness temperature and the integration is performed over the field of view (FOV) or visible disk centered in the direction (θ_o, ϕ_o). Introducing an expected brightness temperature component $T_{B,exp}$, a contribution from RFI $T_{B,RFI}$ and the radiometer system noise N,so that equation (3.6) becomes

$$T_{A,obs}(\theta_o, \phi_o) = \int_{FOV} G(\theta, \phi) \left[T_{B,exp}(\theta, \phi) + T_{B,RFI}(\theta, \phi)\right] d\Omega + N \tag{3.7}$$

or in terms on antenna temperatures

$$T_{A,obs} = T_{A,exp} + T_{A,RFI} + N \tag{3.8}$$

where

$$T_{A,exp} = \int_{FOV} G(\theta, \phi) T_{B,exp}(\theta, \phi) d\Omega \tag{3.9}$$

and the RFI contribution is

$$T_{A,RFI} = \int_{FOV} G(\theta, \phi) T_{B,RFI}(\theta, \phi) d\Omega \tag{3.10}$$

For an RFI-free region, $T_{A,RFI} = 0$, and equation (3.8) can be solved for N:

$$N[\text{no RFI}] = T_{A,obs}[\text{no RFI}] - T_{A,exp}[\text{no RFI}] \qquad \text{for RFI-free regions only} \tag{3.11}$$

while when RFI is present

$$T_{A,RFI} = (T_{A,obs}[\text{RFI}] - T_{A,exp}[\text{RFI}]) - N[\text{RFI}] \tag{3.12}$$

Since N is the same in RFI-free and RFI-affected regions, then $N[\text{RFI}] = N[\text{no RFI}]$ and equation (3.11) can be substituted in equation (3.12):

$$T_{A,RFI} = (T_{A,obs}[\text{RFI}] - T_{A,exp}[\text{RFI}]) - (T_{A,obs}[\text{no RFI}] - T_{A,exp}[\text{no RFI}]) \tag{3.13}$$

In order to evaluate equation (3.13), two set of 10 ms samples are needed, one acquired over an RFI free region and the other over an RFI-affected region. The computation of for the RFI-affected region is performed based on equation (3.13) by following the flow diagram in Figure 3.18. The result comes in the form of a set of RFI samples after the summation circle on the left, that can then be used to generate an RFI histogram.

This analysis requires 10 ms simulated samples to be statistically compared to the observed samples. These can be constructed starting from the 1.44 s expected values of the antenna temperature, as explanied in the following subsection.

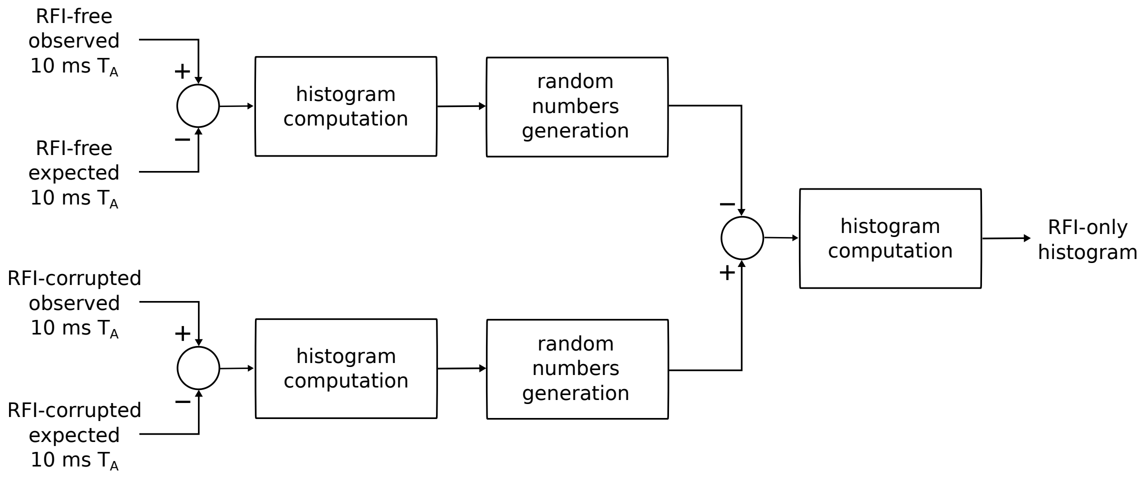


Figure 3.18: *Flow diagram for computation of RFI histograms.*

3.3.2 Simulation of expected antenna temperature samples

The expected antenna temperature levels at 1.44 s time resolution are provided in the Aquarius L2 products. However, they are reliable only on the ocean surface, therefore the analysis of the simulated 10 ms expected antenna temperature will be limited to regions over ocean. In order to generate the corresponding 10 ms samples, one has to assume that the area observed betwen the 1.44 s interval changes in such a way that its brightness temperature varies linearly between points except for changes caused by thermal noise. This is obviously not completely true, and rapid variation in the observed footprint are likely to increase the satandard deviation of the measurements, but the assumption should be reasonably valid over the ocean.

The simulated expected antenna temperature samples are therefore generated using the following steps, as also illustrated in Figure 3.20:

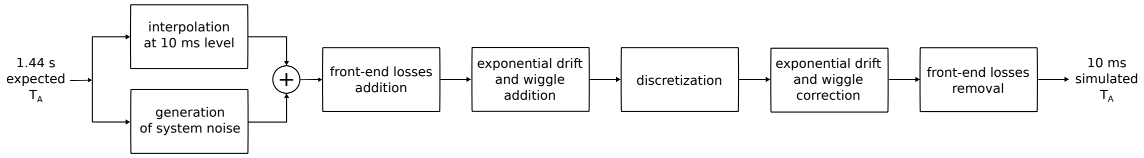


Figure 3.19: *Simulation of expected antenna temperature 10 ms samples from expected antenna temperature 1.44 s values.*

1. The expected T_A at L2 level is linearly interpolated between 1.44 s points to obtain corresponding 10 ms points, as shown in Figure 3.20(a).

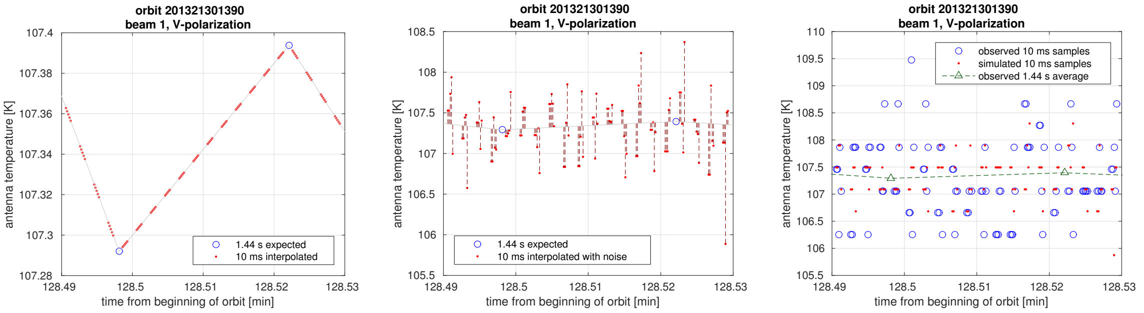
2. Random Gaussian noise is added to each interpolated 10 ms value. Within each 1.44 s window, this noise is assumed to have zero mean and standard deviation

$$\sigma_{exp} = \frac{T_A + T_{rec}}{\sqrt{B\tau}} \tag{3.14}$$

where $T_{rec} = 74.6$ K, $\sqrt{B\tau} = 474.34$ K and T_A is the expected value of the antenna temperature at 1.44 s. An example of the result of this operation is illustrated in Figure 3.20(b);

3. The receiver front-end losses and the error caused by the exponential drift and wiggles are added to the 10 ms samples.
4. The resulting values are discretized.
5. In order to have values comparable to the observed 10 ms samples, front-end losses, exponential drift and wiggles are removed from the discretized 10 ms samples.

In Figure 3.20(c), the simulated 10 ms samples are shown together with the observed samples. The effect of the discretization is very evident, and the resulting discretized levels of the simulated samples is consistent with those of the observed samples.



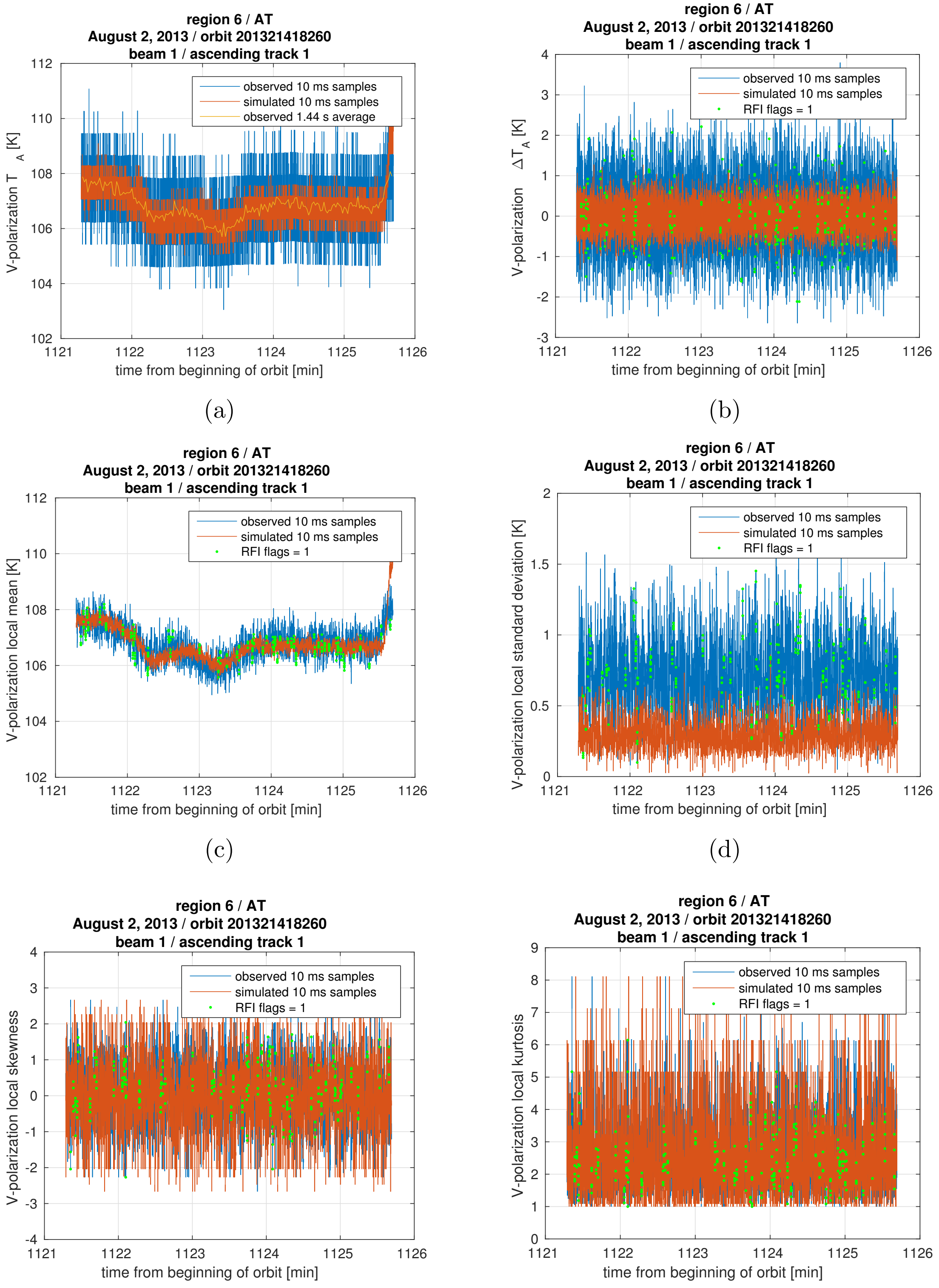
(a) (b) (c)

Figure 3.20: *Simulation of expected antenna temperature 10 ms samples: (a) interpolated 10 ms values without added noise; (b) interpolated 10 ms values with added noise.*

Plots versus time for region 6, ascending track 1 are given in Figure 3.21. Figure 3.21(a) shows the simulated 10 ms samples together with the observed 10 ms samples and the observed 1.44 s points. Figure 3.21(b) shows the simulated and observed samples minus the local mean (in the RFI algorithm windw of width W_m=20), and the points flagged as RFI in green color. Figure 3.21(c),(d),(e) and (f) show the local mean, standard devation, skewness and kurtosis, respectively, for simulated and observed samples, with points flagged as RFI in green color. The histograms corresponding to the plots in 3.21 are shown in Figure 3.22. In addition to considering the observed values (left column) and the simulated values (right column), the histograms for the simulated samples before discretazation are shown here in the middle column.

3.3.3 Example of RFI histogram estimation

In this section,the histograms of the noise N and of the RFI $T_{A,RFI}$ will be evaluated through equation (3.11) and (3.13), then missed detection will be estimated.

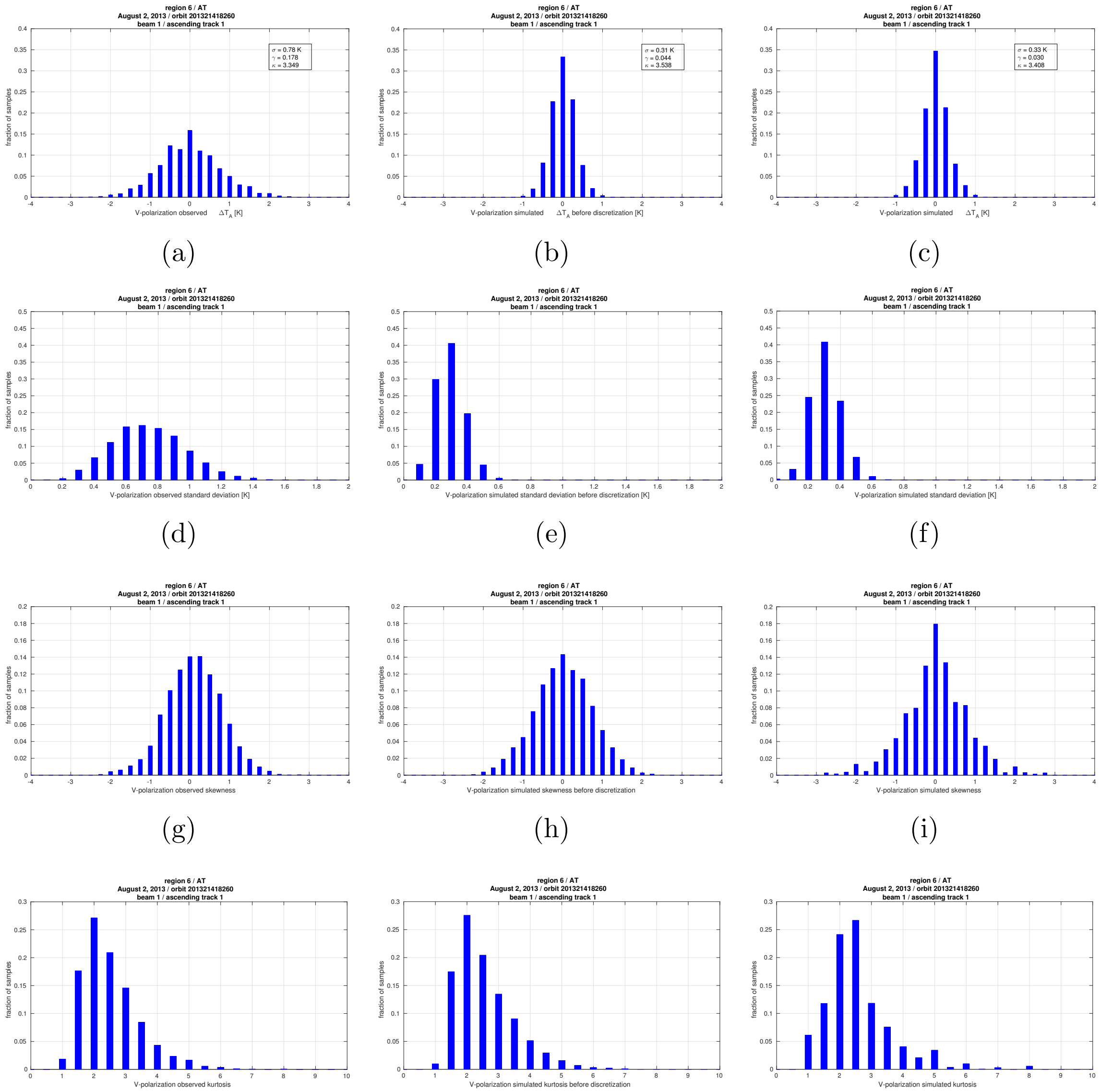


(a) (b)

(c) (d)

(e) (f)

Figure 3.21: *Ascending orbit track 1, region 6: (a) antenna temperature, (b) antenna temperature minus local mean* $W_m = 20$ *averaging window; (c) local mean; (d) local standard deviation; (e) local skewness; (f) local kurtosis.*



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Figure 3.22: *Histograms for ascending region 6, orbit track 1: (a) observed antenna temperature minus local mean in* $W_m = 20$ *averaging window; (b) simulated antenna temperature minus local mean before discretization; (c) simulated antenna temperature minus local mean after discretization; (d) local standard deviation of observed antenna temperature; (e) local standard deviation of simulated antenna temperature before discretization; (f) local standard deviation of simulated antenna temperature after discretization; (g) local skewness of observed antenna temperature; (h) local skewness of simulated antenna temperature before discretization; (i) local skewness of simulated antenna temperature after discretization; (j) local kurtosis of observed antenna temperature; (k) local kurtosis of simulated antenna temperature before discretization; (l) local kurtosis of simulated antenna temperature after discretization;*

The histograms of the observed minus expected 10 ms antenna temperature are shown in Figure 3.23 for ascending region 6, orbit track 1 before (a) and after (b) discretization. These histograms are generated using a 0.1 K bin. However, results can depend on the choice of the bin size, as illustrated in Figure 3.24. Here, the red histograms have been normalized to match the height of the blue histograms for the sake of a better visual comparison. The histograms generated using the discretized 10 ms simulated samples appear to be a “sampled“ copy of the ones generated using the 10 ms simulated samples before discretization. This suggest that estimating the histogram of the observed minus expected 10 ms antenna temperature using the 10 ms simulated samples before discretization is a better choice. Figure 3.25 shows the histograms for the remaining five tracks and for all tracks combined in plot (f). Figure 3.26 shows the corresponding histograms before discretization of the simulated samples. Finally, the overall histograms for all three RFI-free ocean regions are given in Figure 3.27. The values of standard deviation, skewness and kurtosis correspoinding to these histograms are given in Table 3.3. The values vary slighly from region to region, suggesting a small geographical dependence. Based on equation (3.11), the histograms in Figure 3.27 can be considered representative of the noise component N.

The same procedure of generating the histograms of the observed minus expected 10 ms antenna temperature is repeated over the RFI-affected region R1 and the results are illustrated in Figure 3.28. Finally, Figure 3.29 gives the RFI histograms computed using Equation 3.13 as described earlier.

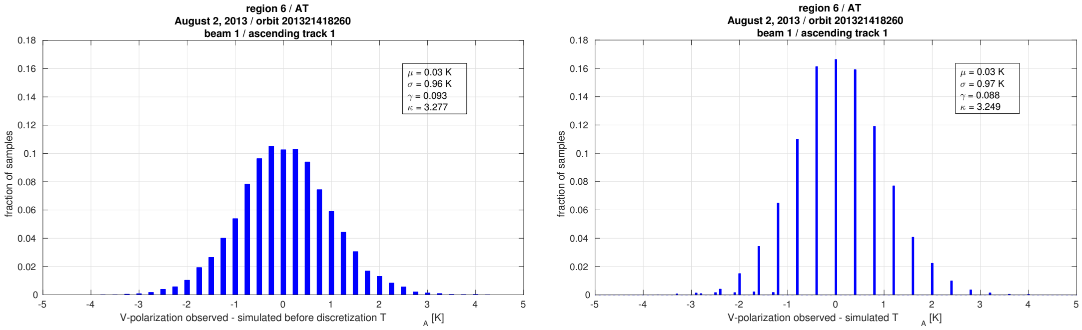


Figure 3.23: *Histograms of the observed minus expected 10 ms antenna temperature for ascending region 6, orbit track 1 before (a) and after (b) discretization, using a 0.1 K bin.*

		AT		PA		IN	
		asc	desc	asc	desc	asc	desc
standard deviation	before discretization	0.948	0.920	1.008	1.047	0.953	0.956
	after discretization	0.955	0.927	1.014	1.054	0.960	0.963
skewness	before discretization	0.102	0.110	0.072	0.308	0.053	0.112
	after discretization	0.098	0.108	0.073	0.305	0.049	0.109
kurtosis	before discretization	3.239	3.269	3.289	3.432	3.263	3.241
	after discretization	3.230	3.258	3.269	3.421	3.249	3.230

Table 3.3: *Standard deviation, skewness and kurtosis for observed minus expected 10 ms antenna temperature.*

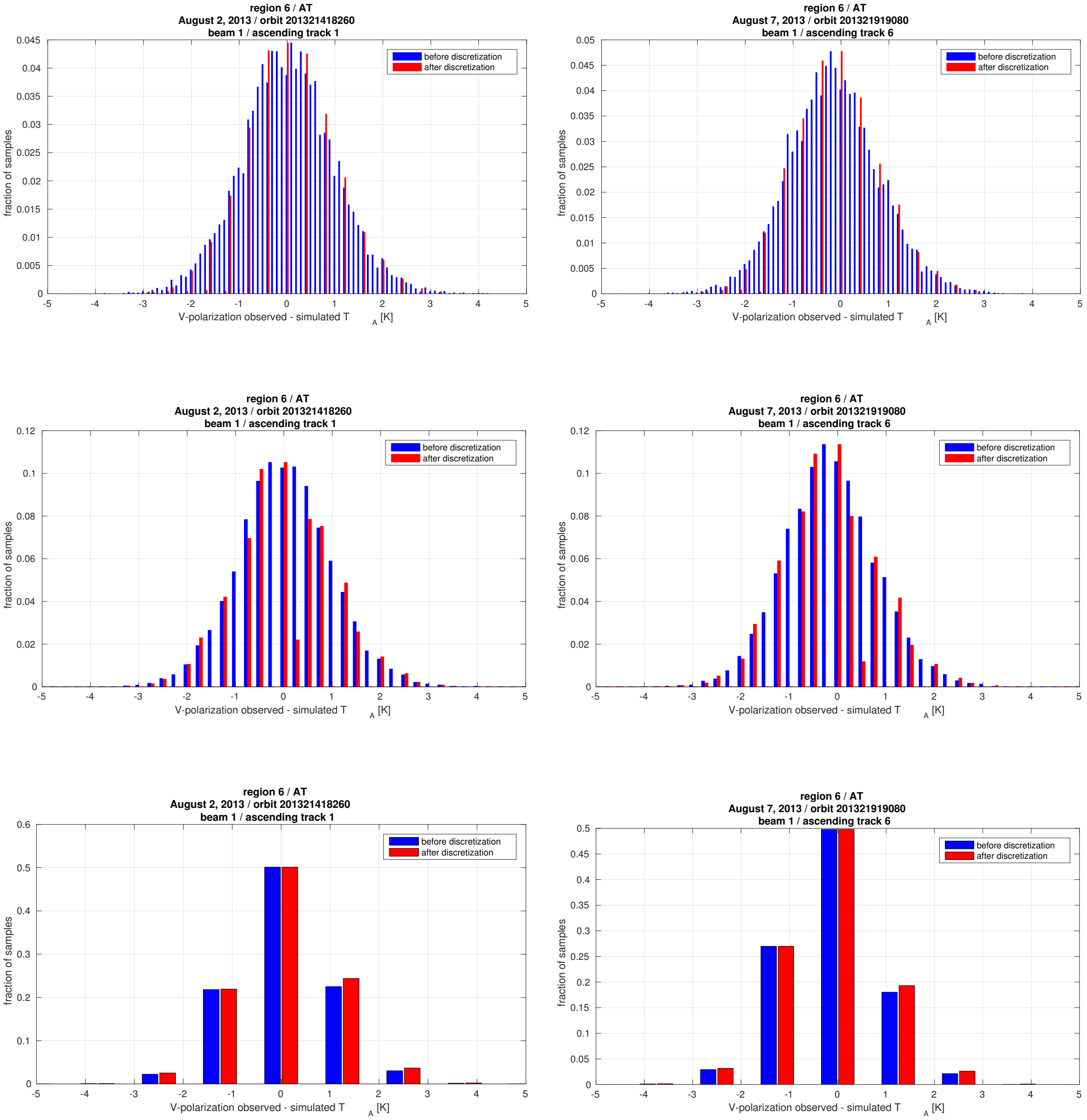
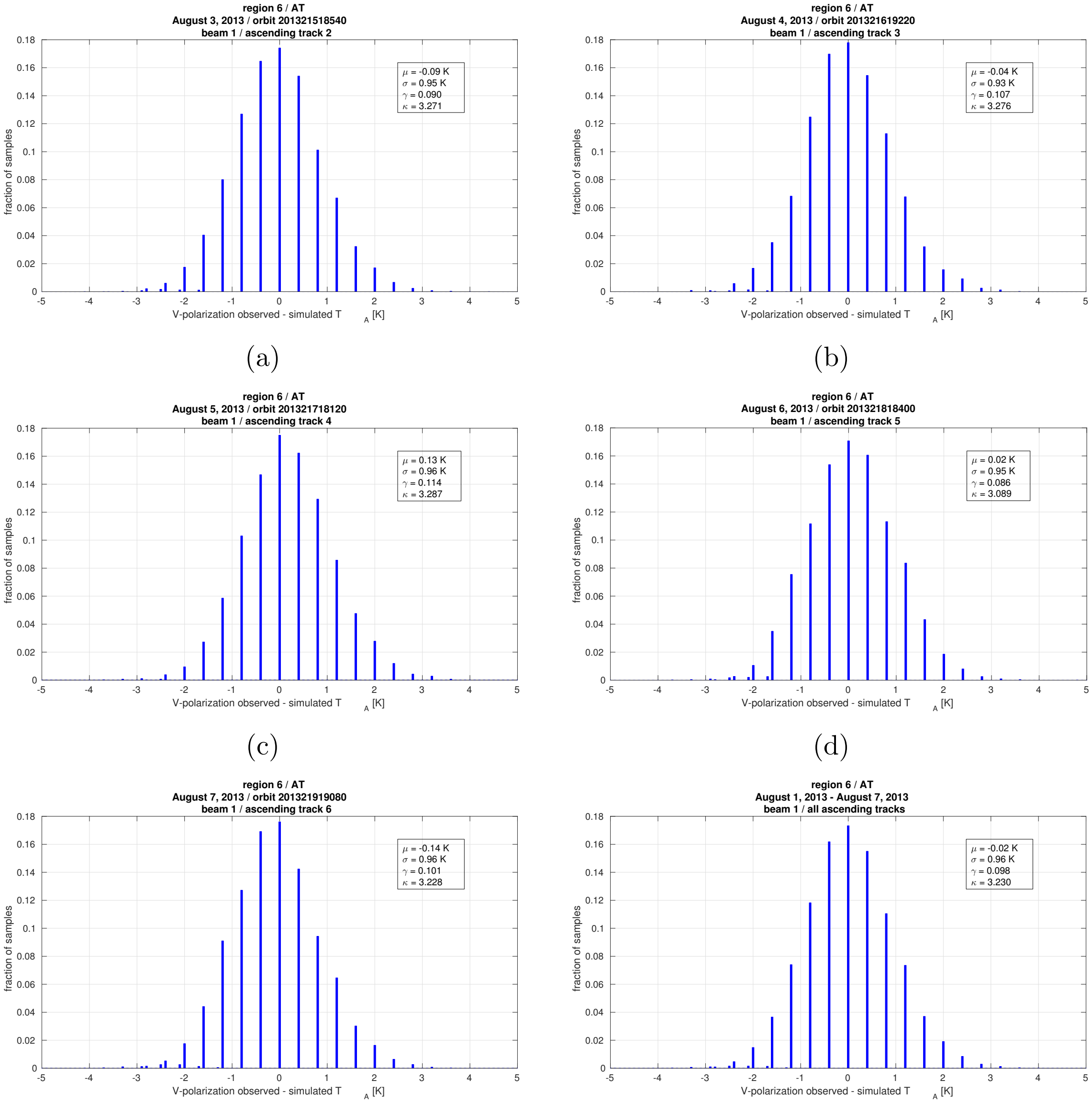
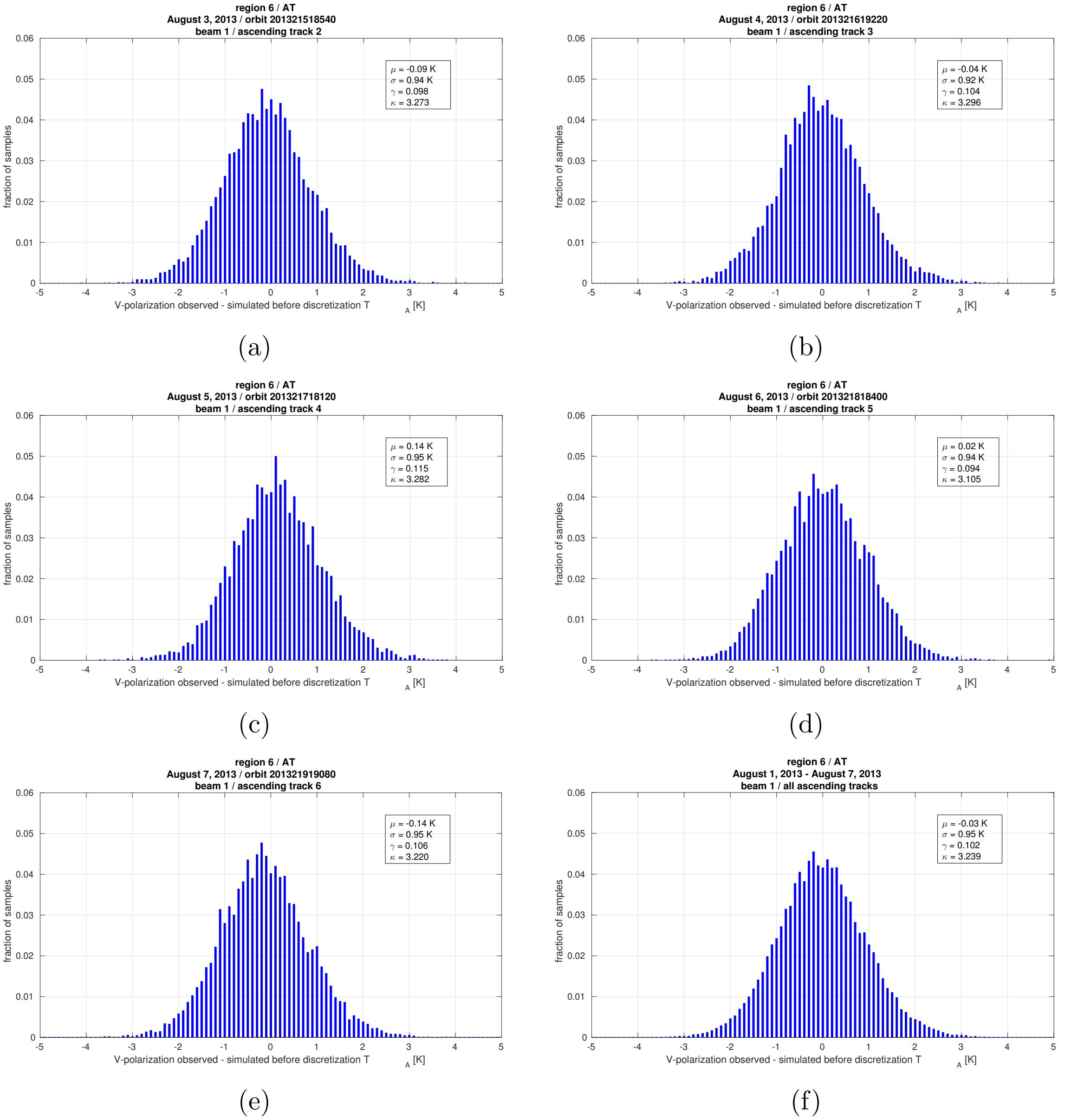


Figure 3.24: *Histograms of the observed minus expected 10 ms antenna temperature for region 6, ascending beam 1 orbit tracks 1 (left) and 6 (right), before (blue) and after (red) discretization, using a 0.1 K (top), 0.5 K (center) and 1.25 K (bottom) bin .*



(a) (b) (c) (d) (e) (f)

Figure 3.25: *Histograms of the observed minus expected 10 ms antenna temperature for region 6, ascending beam 1 orbit tracks 2 (a), 3 (b), 4 (c), 5 (d), 6 (e) and all tracks combined (f) after discretization of the simulated samples.*



region 6 / AT
August 3, 2013 / orbit 201321518540
beam 1 / ascending track 2

region 6 / AT
August 4, 2013 / orbit 201321619220
beam 1 / ascending track 3

region 6 / AT
August 5, 2013 / orbit 201321718120
beam 1 / ascending track 4

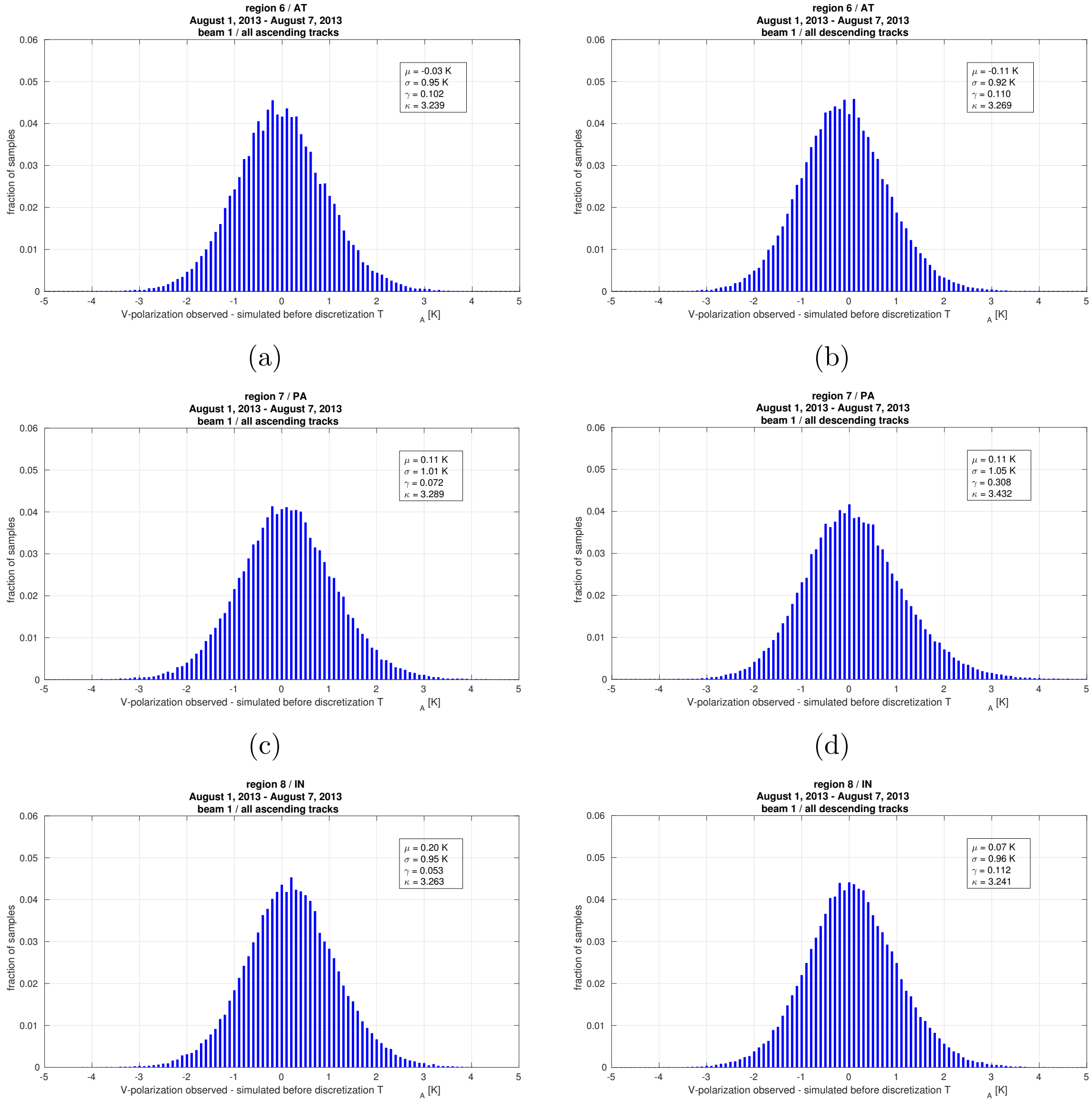
region 6 / AT
August 6, 2013 / orbit 201321818400
beam 1 / ascending track 5

region 6 / AT
August 7, 2013 / orbit 201321919080
beam 1 / ascending track 6

region 6 / AT
August 1, 2013 - August 7, 2013
beam 1 / all ascending tracks

(a) (b) (c) (d) (e) (f)

Figure 3.26: *Histograms of the observed minus expected 10 ms antenna temperature for region 6, beam 1 ascending orbit tracks 2 (a), 3 (b), 4 (c), 5 (d), 6 (e) and all tracks combined (f) before discretization of the simulated samples.*



(a) (b)

(c) (d)

(e) (f)

Figure 3.27: *Histograms of the observed minus expected 10 ms antenna temperature before discretization of the simulated samples for all RFI-free ocean regions (beam 1 only).*

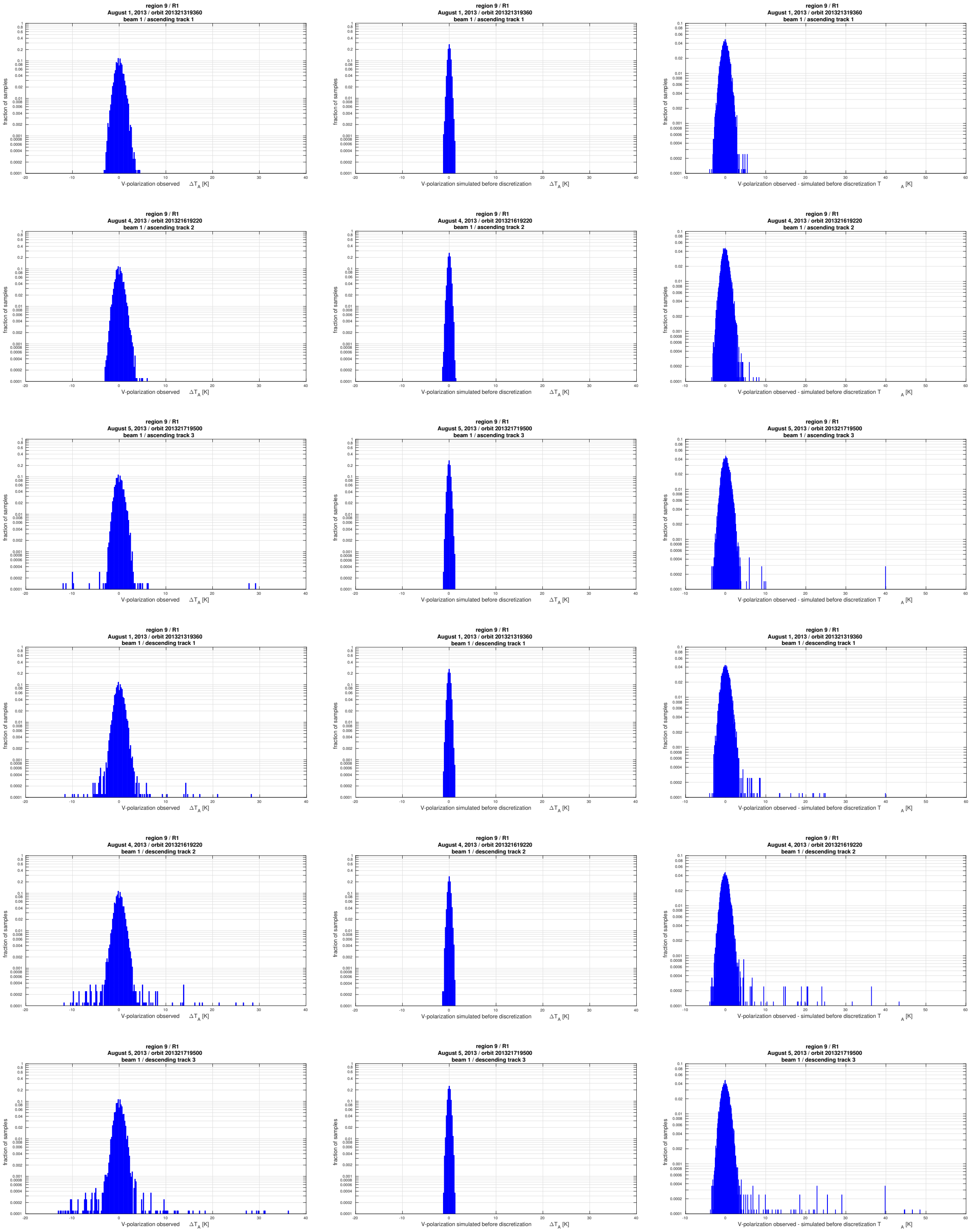


Figure 3.28: *Histograms of antenna temperature observed (left column) and simulated before discretization (center column) and of difference between observed (left column) and simulated antenna temperature (right column) for all ascending and descending beam 1 orbit tracks of region 9.*

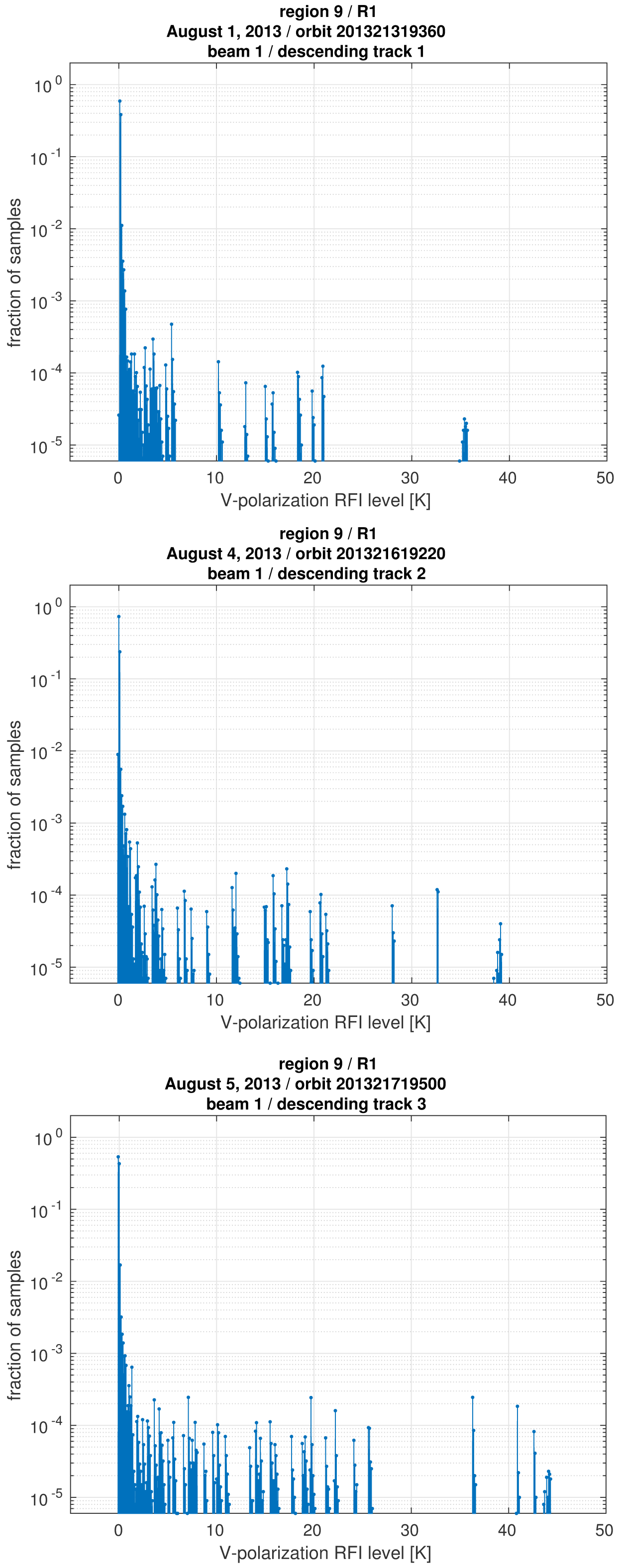


Figure 3.29: *Estimated histograms of RFI for all descending beam 1 orbit tracks of region R.*

4 Missed Detection

Once the statistical distribution of the RFI is known, it can be used to estimate missed detection. The different steps in the estimation are illustrated in Figure 4.1 and are as follows:

1. computation of simulated 10 ms samples without RFI, $T_{A,sim}$[no RFI] starting from 1.44 s values of extected antenna temperature;
2. generation of random RFI samples $T_{A,RFI}$ according to the RFI histogram;
3. computation of simulated samples with RFI, $T_{A,sim}[\text{RFI}] = T_{A,sim}[\text{no RFI}] + T_{A,RFI}$;
4. RFI filtering of simulated samples without RFI, $T_{A,sim}$[no RFI] to produce $T_{F,sim}$[no RFI];
5. RFI filtering of simulated samples with RFI, $T_{A,sim}$[RFI] to produce $T_{F,sim}$[RFI] ;
6. computation of detected RFI as $T_{A,sim}[\text{RFI}] - T_{F,sim}[\text{RFI}]$;
7. estimation of missed detection level as $T_{F,sim}[\text{RFI}] - T_{F,sim}[\text{no RFI}]$.

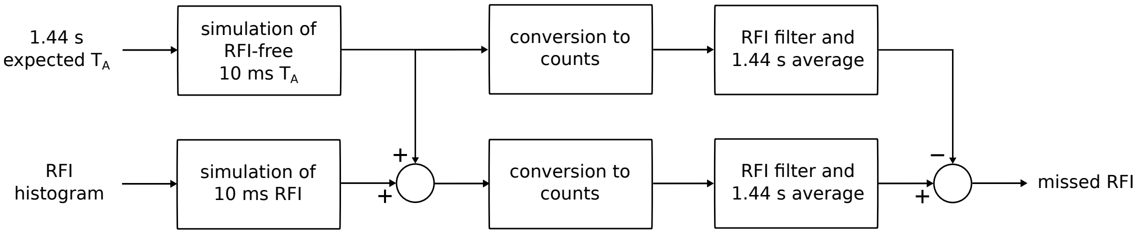
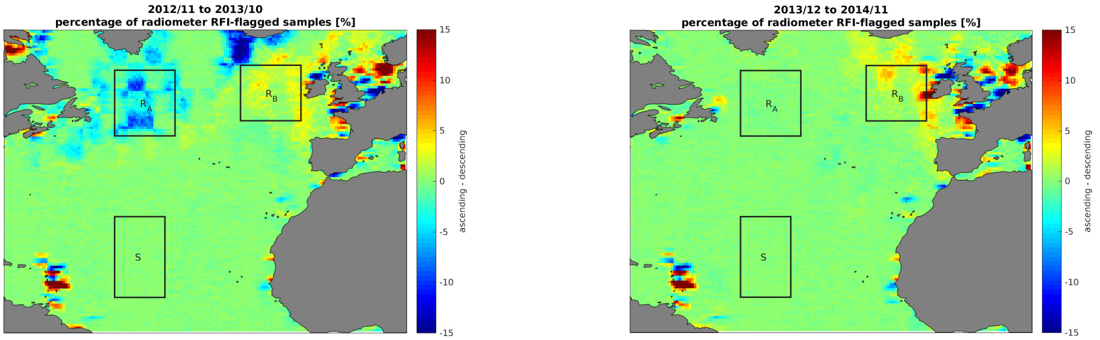


Figure 4.1: *Flow diagram for estimation of missed detection.*

5 RFI histograms and missed detection estimation over North Atlantic ocean

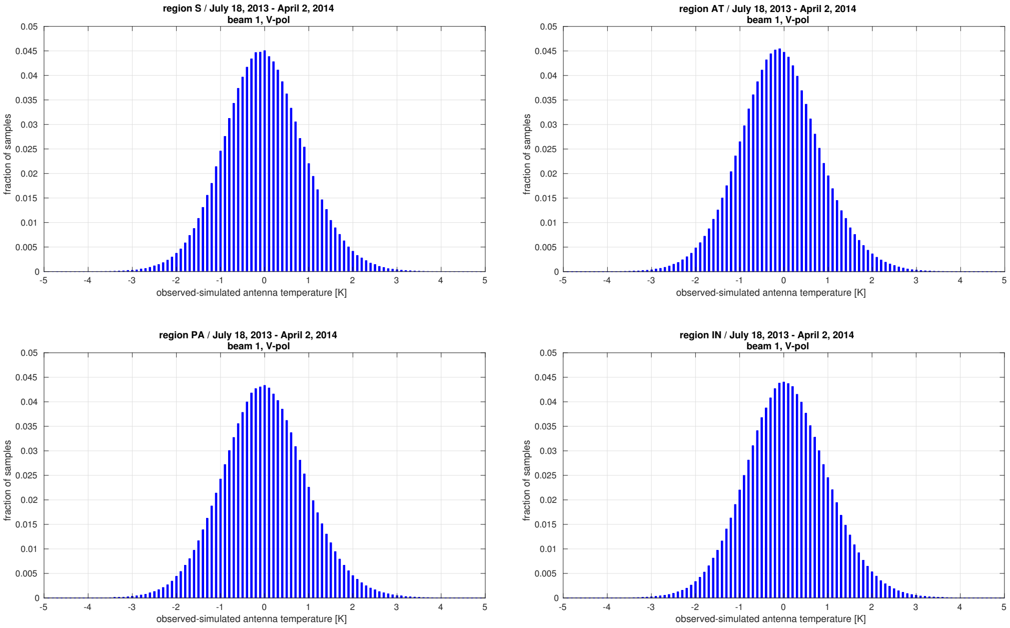
A specific case is examined now, taking advantage of the fact that RFI over part the North Atlantic disappeared during the Aquarius cycle 118 in November 2013 following a change of operating frequency in the Northern Warning System (NWS, formerly known as Distant Early Warning or DEW line) radars. The regions used in this analysis are shown in Figure 5.1 on maps of ascending/descending difference in RFI percentage for the 12 months before (left) and after (right) the disappearance of RFI from the NWS radars in November 2013. Ascending/descending differences can be considered a proxy for the presence of RFI leaking into the antenna from the side lobes and affecting the measuremnts at lower level than interference entering the main antenna beam. The region R_A roughly corresponds to the region R_1 used in Section 2, but its boundaries have been slightly modified to cover the area affected by the change of RFI environment. The blue color in this region on the left side map indicate that the RFI present before November 2013 affected mainly the descending tracks.


2012/11 to 2013/10
percentage of radiometer RFI-flagged samples [%]

2013/12 to 2014/11
percentage of radiometer RFI-flagged samples [%]

Figure 5.1: *Regions for RFI histograms and missed detection analysis.*

Histograms of observed-simulated antenna temperature $T_{A,obs} - T_{A,sim}$ have been computed for region S and compared to those of regions AT, PA and IN defined in Section 2 (see Figure 5.2) to show that S can be assumed to represent all those three regions in terms of statistics.


region S / July 18, 2013 - April 2, 2014
beam 1, V-pol

region AT / July 18, 2013 - April 2, 2014
beam 1, V-pol

region PA / July 18, 2013 - April 2, 2014
beam 1, V-pol

region IN / July 18, 2013 - April 2, 2014
beam 1, V-pol

Figure 5.2: *Comparison of histograms of for S region to histograms for AT, PA and IN regions.*

The estimation of the RFI histograms for descending tracks in region R_A is performed in two different ways:

1. the reference RFI-free data set are measurements after cycle 118 when RFI from the NWS radar was no longer present;
2. the reference RFI-free data set are measurements from region S in Figure 5.1.

Specifically, method 1. uses data from cycles 119 to 140 and method 2. uses data from cycles 96 to 140, while the RFI-corrupted data are from cycles 96 to 117. RFI in region R_A

disappeared during the weekly cycle 118, therefore data acquired in this week are excluded from the calculations for region R_A.

The following procedure has then been applied:

A. histograms are generated for each 1.44 s point using all the 10 ms sample associated with that point and with five point to before and after, resulting a total of $60 \times 11 = 660$ samples per histogram;

B. step A. is repeated for every weekly cycle;

C. the average histogram over all the weekly cycles is considered.

The histograms for V-polarization generated with these two approaches are shown in Figure 5.3, on the left column for method 1., and on the right for 2. The results obtained with the two methods are noticeably consistent with each other.

These estimated RFI histograms are used to randomly generate RFI that is then added to the simulated RFI-free 10-ms antenna temperature values. The RFI values at 10-ms time interval along with the antenna temperature values, without and with RFI, for three descending tracks within region R_A and cycle 102 are shown in the top, middle and bottom row of Figure 5.4, respectively.

Figures 5.5 to 5.7 present the result of applying this procedure to three orbit tracks within region R_A during weekly cycle 102. Specifically, in Figure 5.5 the percentage of RFI randomly generated and added to the antenna temperature before feeding it to the RFI filter is shown in blue, while the percentage of RFI actually detected by the Aquarius RFI algorithm is in red color. In the top row, RFI histograms exploiting the November 2013 change in RFI environment were used, while in the three bottom plots the reference region S was used for RFI histogram estimation. Figure 5.6 shows the RFI-filtered antenna temperature T_F along the same tracks. The blue curve is obtained without adding any RFI to the 10 ms samples, while the red curve is obtained using the same 10 ms samples with added randomly generated RFI. The red curve should always be higher than the blue courve, however in some cases this may not occur due to false alarms. The difference between the two curves is the missed detection given in Figure 5.7.

The same analysis is repeated for region R_B and the results are given in Figure 5.8. In this case, only region S could be used to estimate the RFI histograms.

From these examples, there appears to be a considerable amount of missed RFI. The level of missed interference rarely exceeds 0.2 K but in some cases can be as high as 0.3K. A lower limit for RFI to be detected by the filter due to threshold settings is probably around 0.2 K. If RFI is at this level is common in a particular region, it could introduce a significant error in the antenna temperature measurements.

based on before/after November 2013 comparison

based on region S

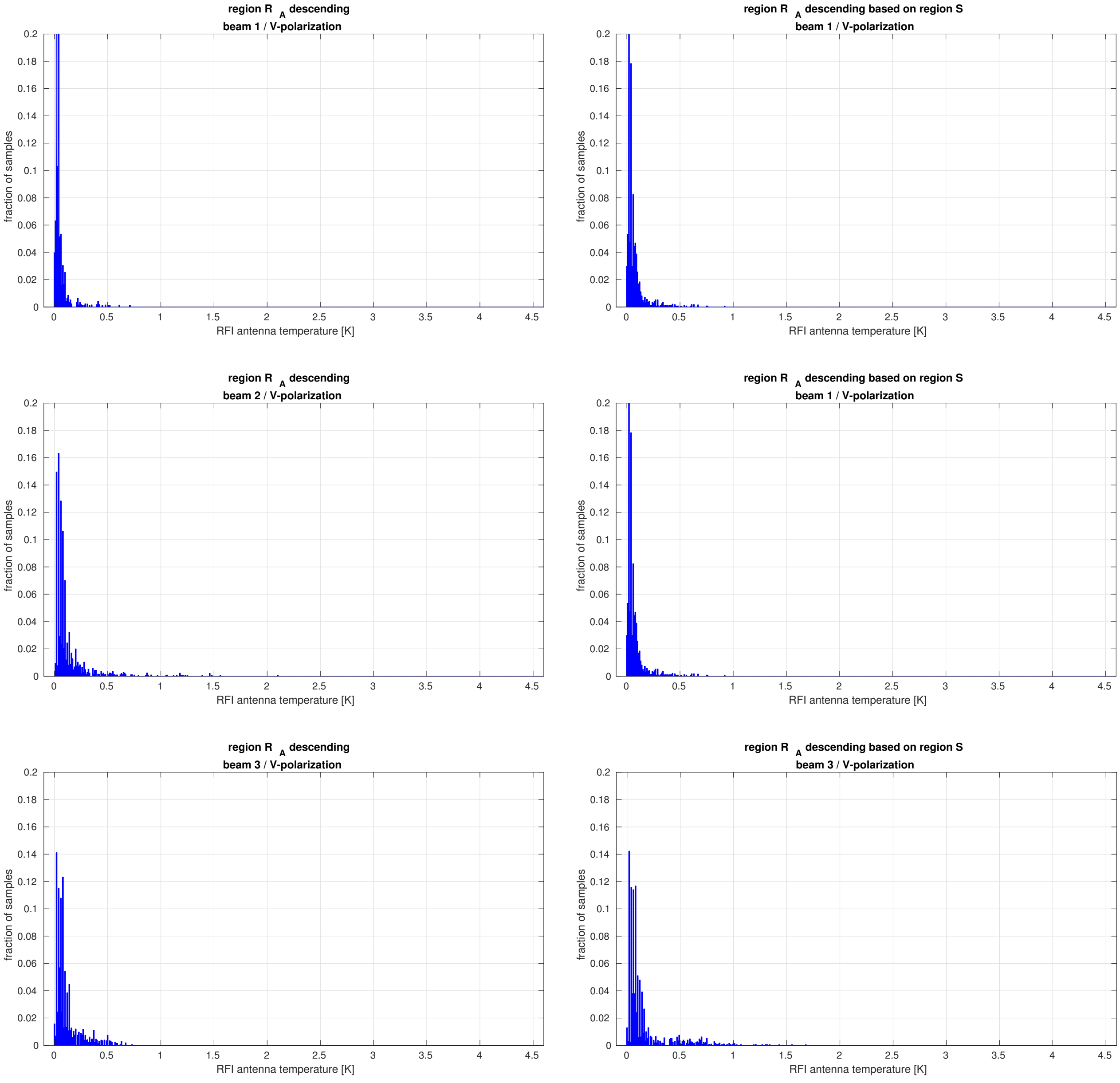


Figure 5.3: *RFI histograms for region* R_A

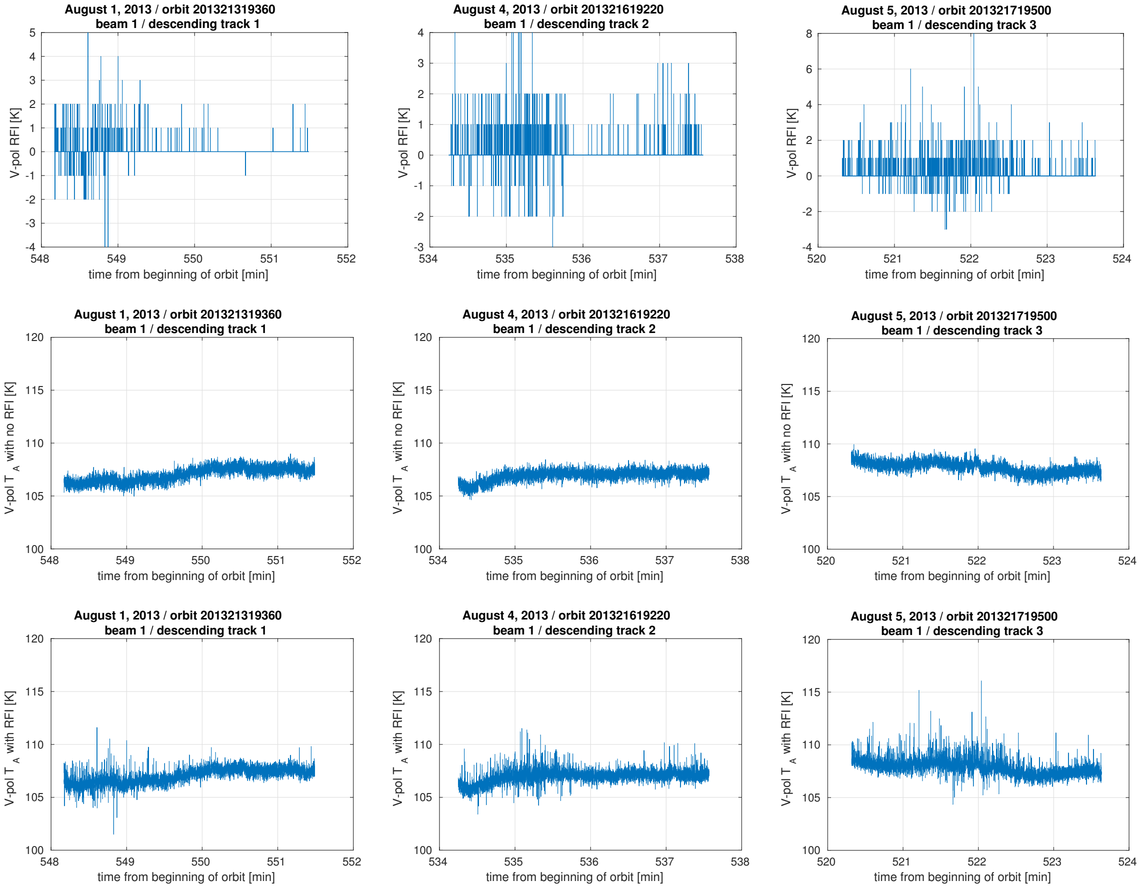
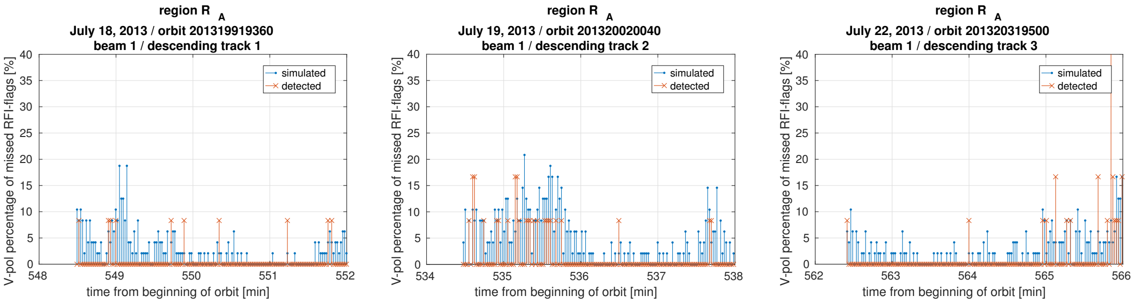


Figure 5.4: *Randomly generated RFI (top), and simulated 10-ms T_A without RFI (middle) and with RFI (bottom) for region R_A descending, cycle 102.*

based on before/after November 2013 comparison



based on region S

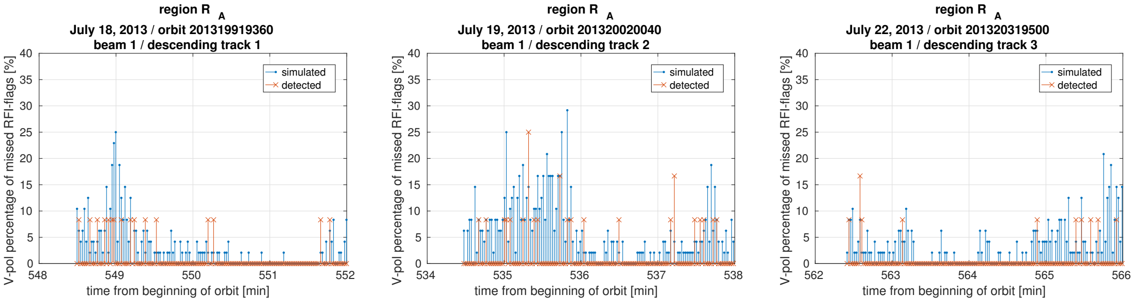
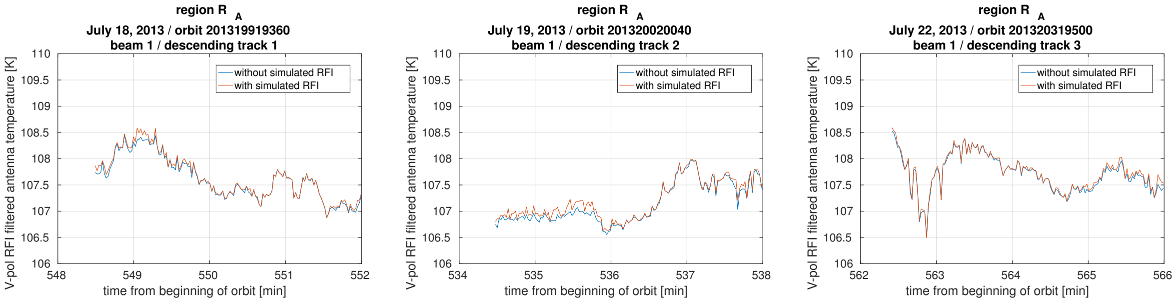


Figure 5.5: *Comparison of simulated and detected RFI for R_A region.*

based on before/after November 2013 comparison



based on region S

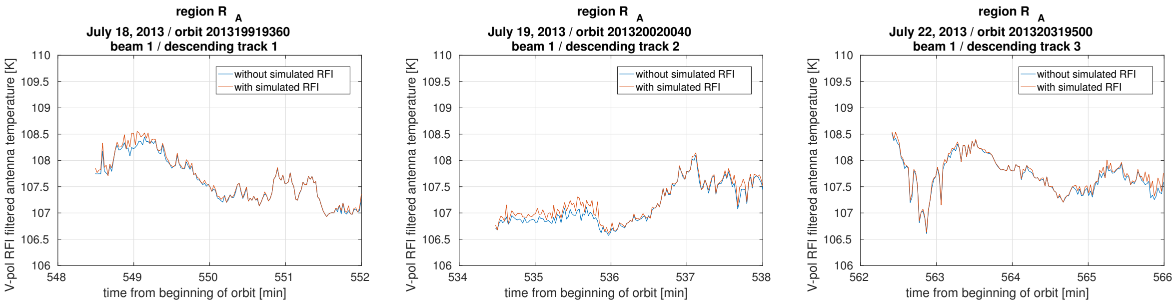
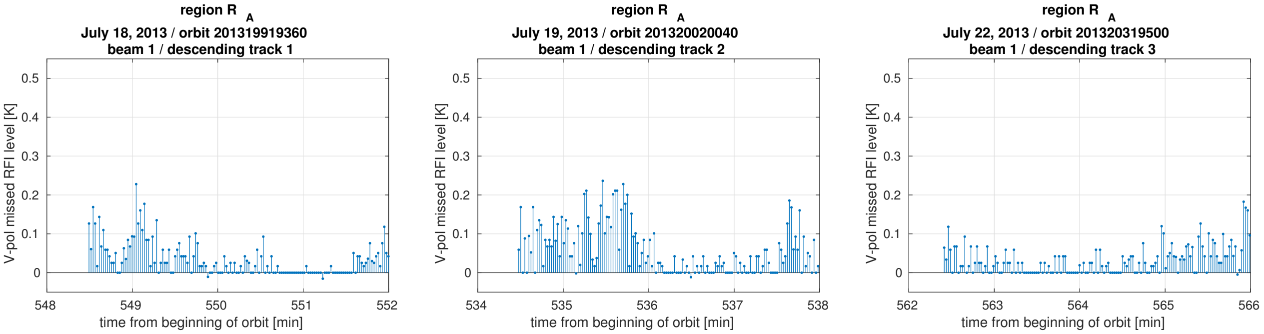


Figure 5.6: *Comparison of RFI-mitigated antenna temperature T_F for R_A region.*

based on before/after November 2013 comparison



based on region S

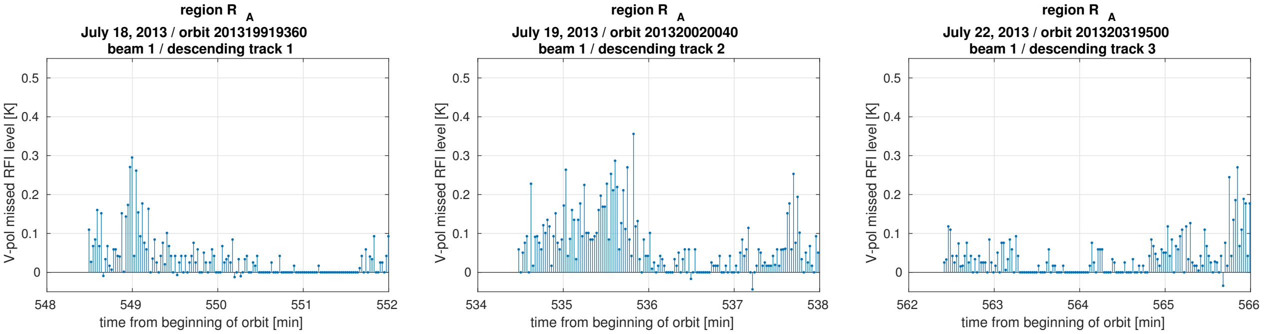


Figure 5.7: *Comparison of missed RFI percent for R_A region.*

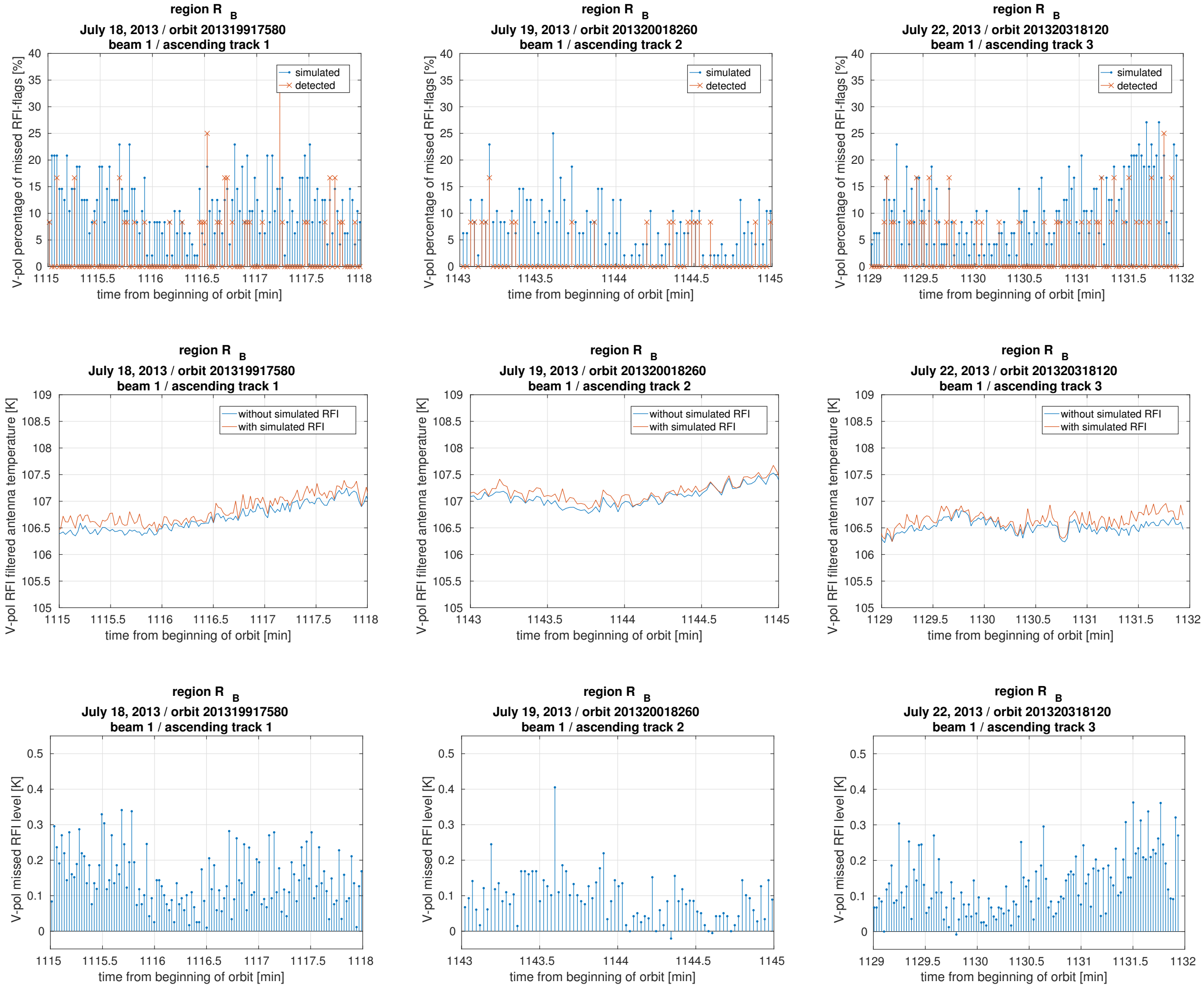


Figure 5.8: *Simulated and detected RFI (top), RFI-filtered antenna temperature T_F (middle) and missed detection (bottom) for region R_B.*

6 Conclusions

A statistical analysis of RFI in Aquarius radiometer measurements has been conducted and described in this report. Aspects that were considered include:

1. Higher moments (such as standard deviation, skewness and kurtosis) of 10 ms samples within the local 1.44 s averaging window used for L2 data generation;
2. Estimation of RFI histograms by comparison of RFI-free and RFI-affected datasets;
3. Estimation of missed RFI detection based on RFI histograms from 2.

Conclusions and recommendations for future work can be summarized as follows:

1. For higher moments analysis:
    - non-normal (i.e., implying non-Gaussian statistics) values of skewness and kurtosis do not always correspond to detected RFI, which could be an indication that some interference cause skewness and kurtosis to change but does not alter the time behavior of the RFI, resulting in missed detection by the time-based RFI filter; including skewness and kurtosis in the L2 products could help identify missed RFI corresponding to these cases, and this could be implemented through simple anomalous skewness and kurtosis flags.
2. For RFI histograms analysis:
    - the statistics of the difference between observed and expected 10 ms antenna temperature values for the reference region S match very well those of other reference ocean regions (which are already in good agreement with each other);
    - the RFI histograms obtained using reference region S and exploiting the November 2013 change in RFI environment are similar, suggesting that using a reference region is feasible;
    - the shape of estimated histograms is consistent with other studies, such as those given by Chen and Ruf in [4].
3. For missed RFI detection analysis:
    - the RFI filter appears to miss a large part of low-level RFI underapproximately 0.2 K, which is the fraction of RFI whose level is below the RFI algorithm threshold;
    - the procedure has been applied using a reference and exploiting the November 2013 change in RFI environment and the results obtained in terms of missed RFI level are similar;
    - procedure could be used globally to produce maps of missed detection that could be compared with those obtained by other methods such as integration of interference signal over antenna sidelobes;
    - once a reliable estimate of missed detection is available, it could be used for
        - better tuning of τ_m and τ_d parameters in low-RFI regions;
        - determining an average RFI bias to correct antenna temperatures for missed RFI in low-RFI region.

References

[1] D.M. Le Vine, P. de Matthaeis, C.S. Ruf, and D.D. Chen, "Aquarius RFI Detection and Mitigation Algorithm: Assessment and Examples," *I*EEE Transactions on Geoscience and Remote Sensing, vol.52, no.8, pp.4574-4584, Aug. 2014.

[2] D.M. Le Vine and P. de Matthaeis, "Aquarius Active/Passive RFI Environment at L-Band", *I*EEE Geoscience and Remote Sensing Letters, vol.11, no.10, pp.1747-1751, Oct. 2014.

[3] D.M. Le Vine, G.S.E. Lagerloef , F.R. Colomb, S.H. Yueh, and F.A. Pellerano, "Aquarius: an instrument to monitor sea surface salinity from space," *I*EEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 7, pp. 2040-2050, July 2007.

[4] G.S.E. Lagerloef, F.R. Colomb, D.M. Le Vine, F. Wentz, S.H. Yueh, C.S. Ruf, J. Lilly, J. Gunn, Y. Chao, A. deCharon, G. Feldman, and C. Swift, "The Aquarius/SACD Mission: Designed to meet the salinity remote-sensing challenge," *O*ceanography, vol. 21 no. 1, pp. 6881, March 2008.

[5] Footnote 5.340 of the Table of Frequency Allocations, *R*adio Regulations, ITU Publication, Geneva, 2016.

[6] S. Misra and C. Ruf, "Detection of Radio Frequency Interference for the Aquarius Radiometer," *I*EEE Transactions on Geoscience and Remote Sensing, vol. 46, no. 10, pp.3123-3128, October 2008.

[7] F.J. Wentz and D.M. Le Vine, Algorithm Theoretical Basis Document Aquarius Salinity Retrieval Algorithm, RSS Technical Report 082912, August 29, 2012.

[8] HYCOM, 2012: The Hybrid Coordinate Ocean Model: http://hycom.org.

[9] J. Piepmeier, Aquarius Radiometer Post-Launch Calibration, Aquarius Project Document: AQ-014-PS-0015, Feb 19, 2013.

[10] D. D. Chen and C. S. Ruf, "Adaptive Control of Undetected Radio Frequency Interference With a Spaceborne Microwave Radiometer," *I*EEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 9, pp. 4972-4984, Sept. 2015.

[11] S. Misra, R.D. De Roo, and C.S. Ruf, "An Improved Radio Frequency Interference Model: Reevaluation of the Kurtosis Detection Algorithm Performance Under Central-Limit Conditions" *I*EEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 11, pp. 4565-4574, Nov. 2012.