Status of Radiometer RFI Work and Missed Detection Analysis

Paolo de Matthaeis, Yan Soldo and David Le Vine

Completed Improvements

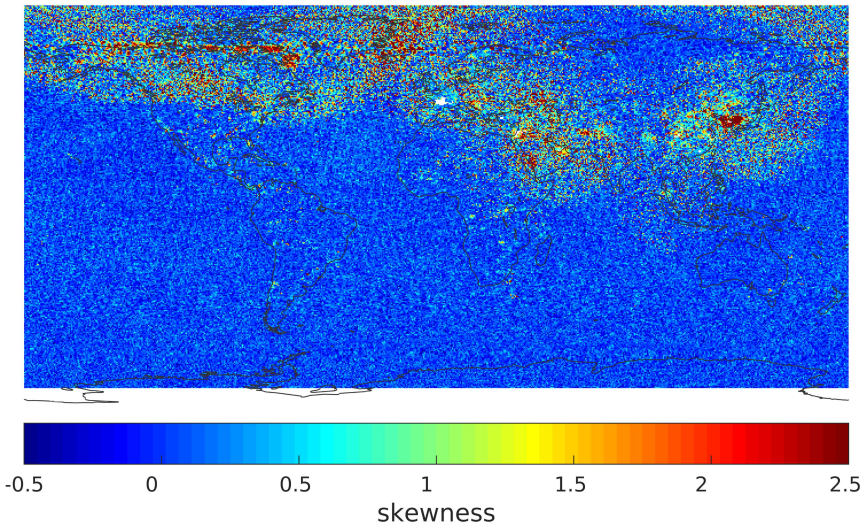
- tuning of RFI algorithm σ_s parameter to achieve the same level of false alarm rate over all radiometers and polarization channels
- adding geographical dependence for σ_s to have the same false alarm rate over land and ocean
- introducing new flag for unrealistically high levels of antenna temperature

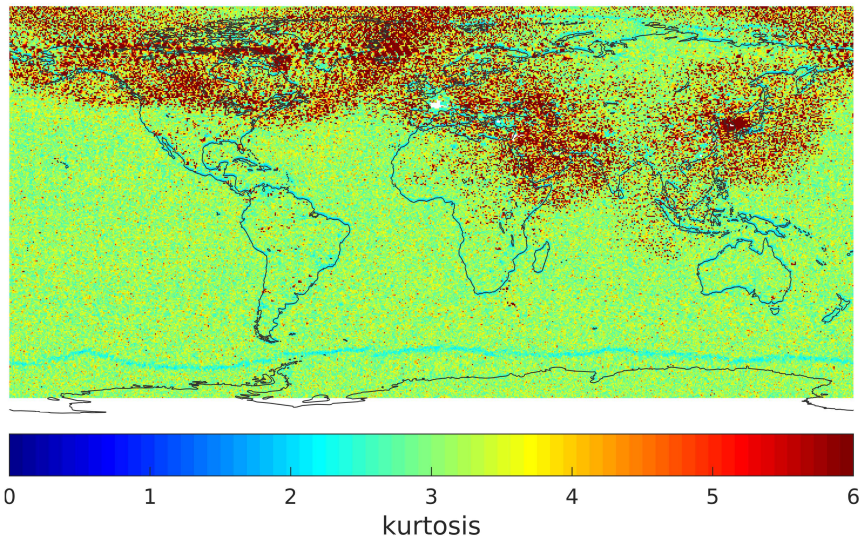
Potential New Improvements

- additional RFI flags based on higher order statistics
- tuning of other RFI algorithm parameters (W_m, τ_m, W_d, τ_d) to reduce missed detection and/or false alarm rate
- any modifications should not significantly change antenna temperature (salinity), except for improving data quality, over RFI-free regions

Skewness and Kurtosis for 1.44 s Averages

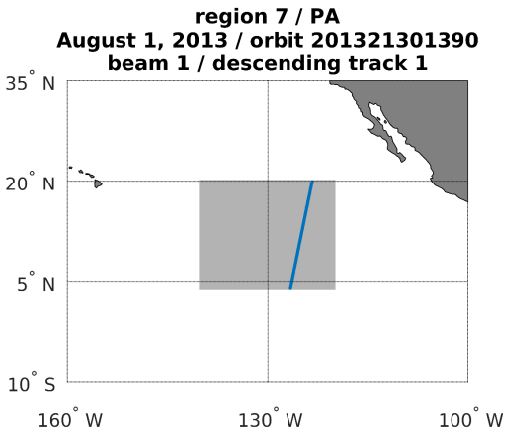
- skewness and kurtosis can be computed for every 1.44 s data point based on the 10 ms samples used in the averages
- they are very good indicators of RFI presence

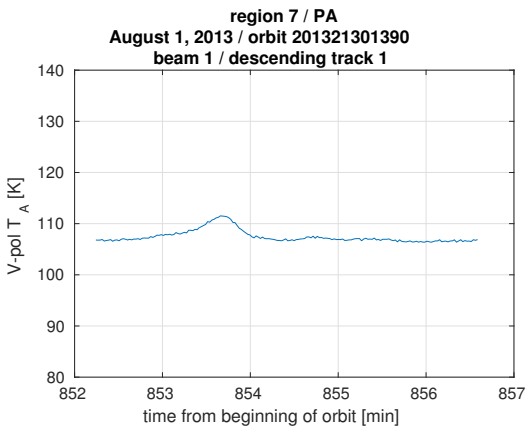


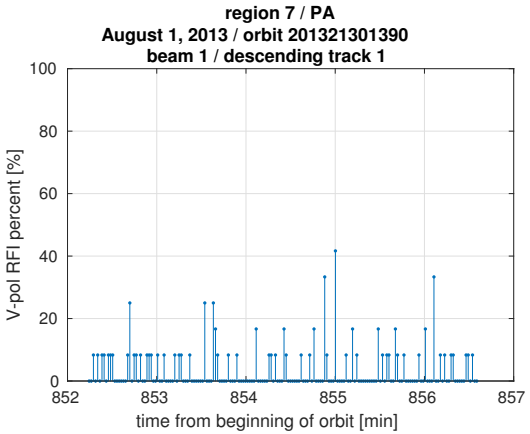


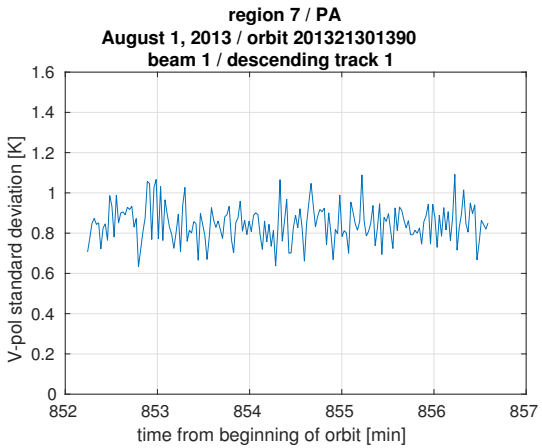
Skewness and Kurtosis for 1.44 s Averages

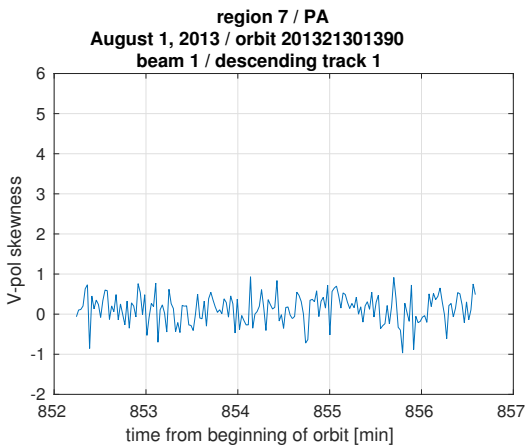
- abnormal values of skewness and kurtosis do not always correspond to detected RFI, suggesting that some RFI is being missed
- including skewness and kurtosis in the L2 products would help identify missed RFI corresponding to these cases

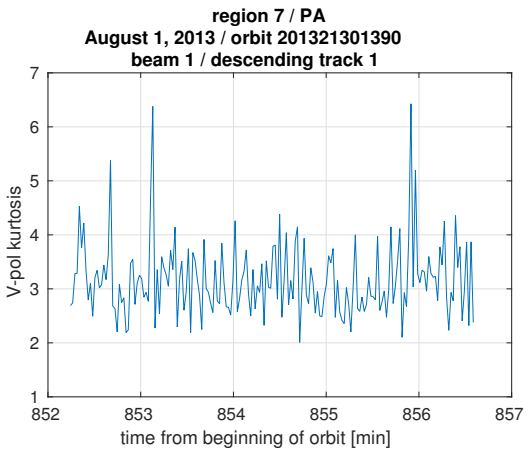












Missed RFI Detection Analysis

- low-level RFI undetected by the RFI algorithm is a still unresolved issue and its effect on salinity retrieval has not been quantified yet
- an attempt to estimate the level of missed detection has been made by comparing statistics of measurements in low-RFI and RFI-free regions
- once the rate of missed detection is known, the RFI algorithm parameters can be tuned to reduce it

RFI Statistics and Missed Detection

- statistical distribution of RFI level is estimated using actual Aquarius data
- simulated RFI levels are generated using these RFI statistics
- simulated 10 ms antenna temperature samples without and with simulated RFI are run through the RFI filter to estimate missed detection

Estimation of RFI Histograms: Theory

- in general, at 10-ms level,

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

  where
  - $T_{A,obs}$ = observed antenna temperature
  - $T_{A,exp}$ = expected antenna temperature
  - $T_{A,RFI}$ = contribution to antenna temperature from RFI
  - N = contribution to antenna temperature from system noise

- for RFI-free measurements, $T_{A,RFI} = 0$, therefore:

  $$T_{A,obs}[\text{no RFI}] = T_{A,exp}[\text{no RFI}] + N[\text{no RFI}]$$
  $$\Longrightarrow\ N[\text{no RFI}] = T_{A,obs}[\text{no RFI}] - T_{A,exp}[\text{no RFI}]$$

- for RFI-affected measurements:

  $$T_{A,obs}[\text{RFI}] = T_{A,exp}[\text{RFI}] + T_{A,RFI} + N[\text{RFI}]$$
  $$\Longrightarrow\ T_{A,RFI} = T_{A,obs}[\text{RFI}] - T_{A,exp}[\text{RFI}] - N[\text{RFI}]$$

- assuming $N[\text{RFI}] = N[\text{no RFI}]$:

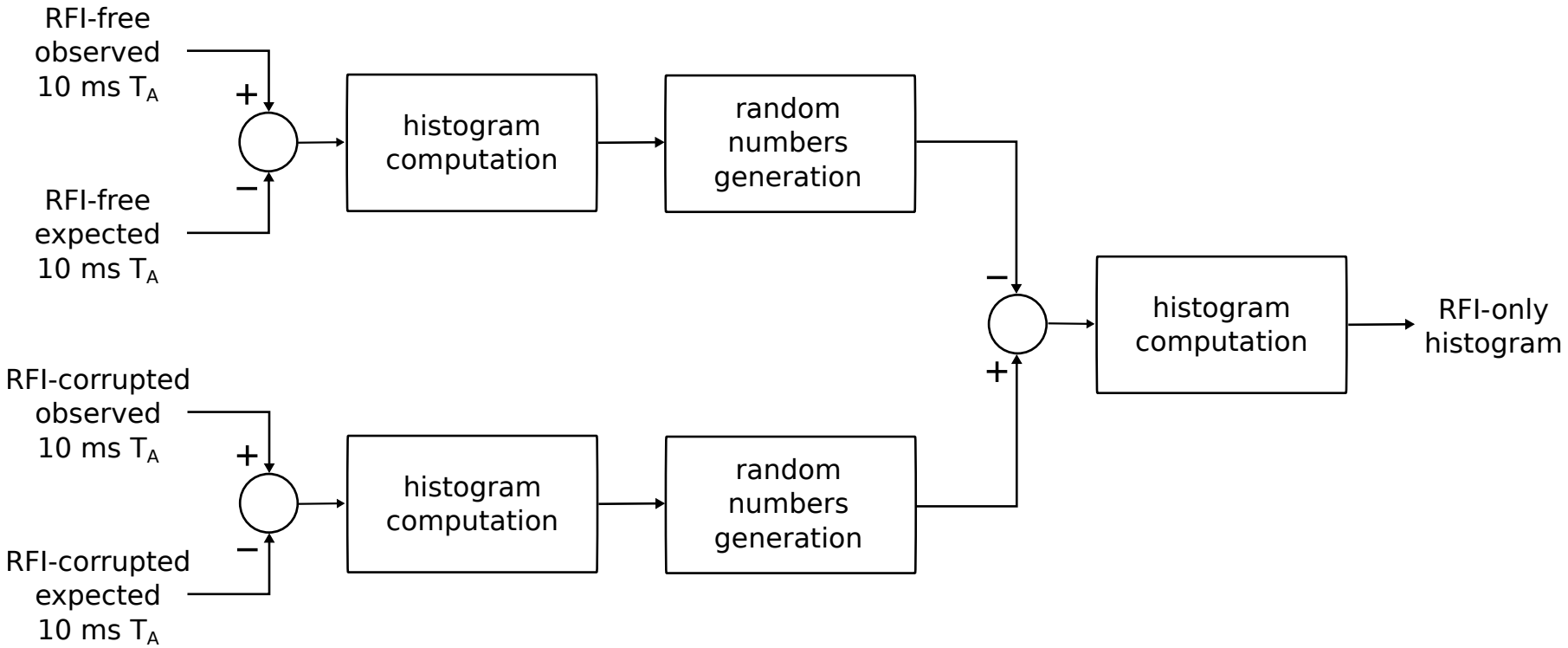
  $$T_{A,RFI} = \left(T_{A,obs}[\text{RFI}] - T_{A,exp}[\text{RFI}]\right) - \left(T_{A,obs}[\text{no RFI}] - T_{A,exp}[\text{no RFI}]\right)$$

Estimation of RFI Histograms: Methodology

- the contribution of RFI to the antenna temperature is given by

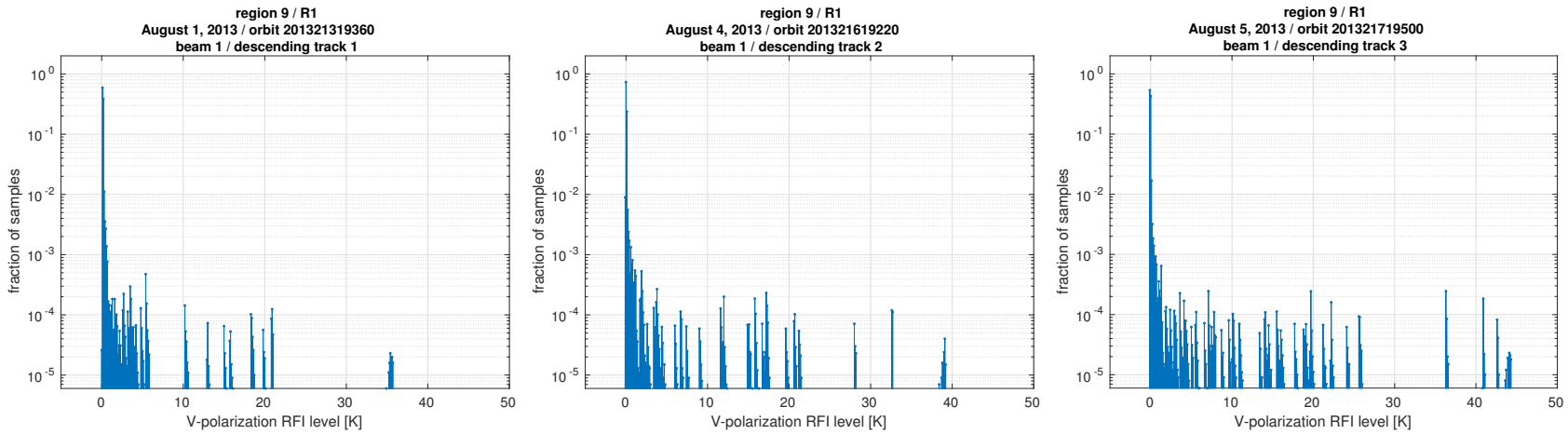
$$T_{A,RFI} = \left(T_{A,obs}[\text{RFI}] - T_{A,exp}[\text{RFI}]\right) - \left(T_{A,obs}[\text{no RFI}] - T_{A,exp}[\text{no RFI}]\right)$$

- the difference between the two terms in this expression is calculated in statistical terms following the flow chart below



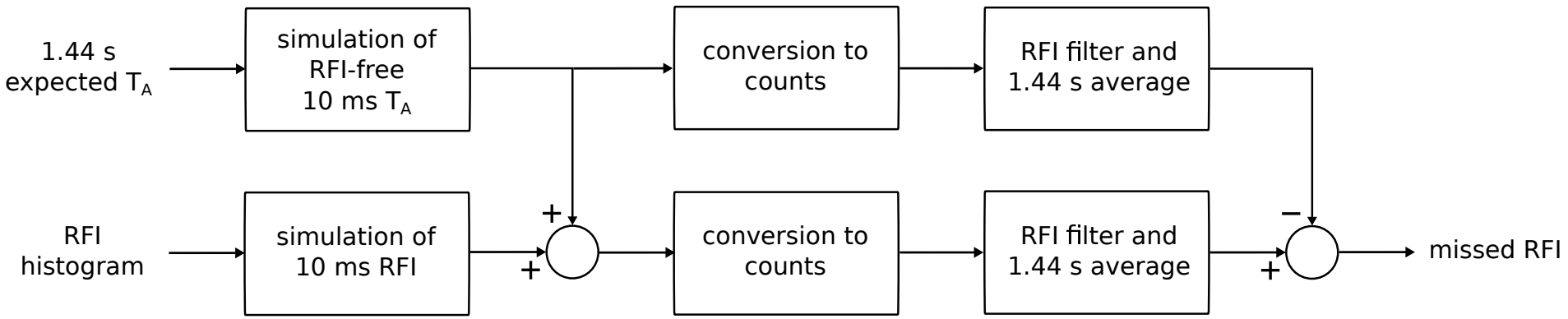
Estimated RFI Histograms

- a region in the North Atlantic ocean south-west of Iceland is considered as case study
- the RFI environment in this region changed in October 2013 following an operating frequency adjustment in Northen Canada radars
- measurements acquired before October 2013 are assumed corrupted by RFI, while measurements after October 2013 are considered RFI-free
- statistics are considered for each data point over a number of cycles (i.e., cycles 102 to 117 for RFI-corrupted samples and cycles 119 to 123 for RFI-free samples)



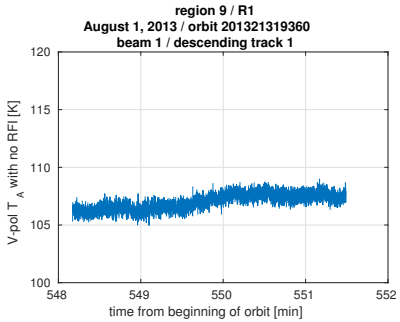
Estimation of Missed Detection

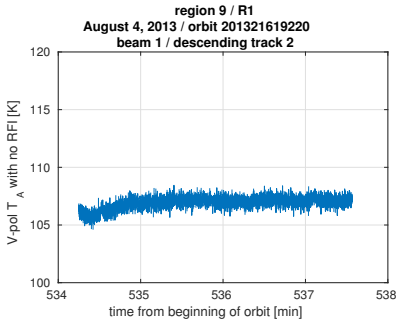
- expected RFI-free antenna temperature is simulated at 10 ms intervals
- random levels of RFI at 10 ms intervals are generated using the RFI histograms and added to expected RFI-free antenna temperature to simulate RFI-affected 10 ms antenna temperature
- after conversion to counts, both RFI-free and RFI-affected simulated samples are processed by the RFI filter
- the difference between the filtered RFI-affected and the RFI-free filtered antenna temperature represents the missed RFI

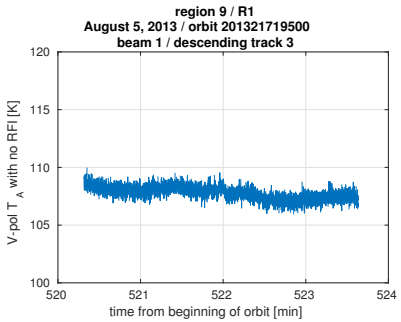


Simulated Antenna Temperature and RFI

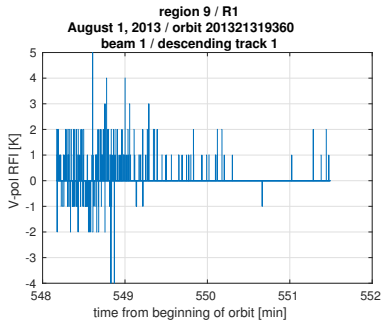
Simulated 10 ms T_A without RFI

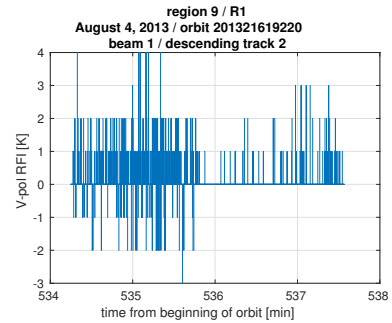


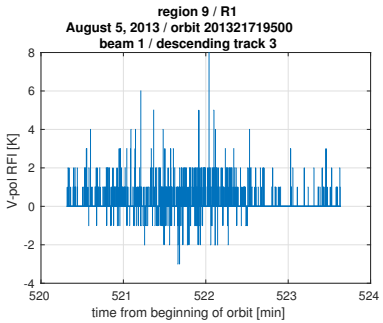




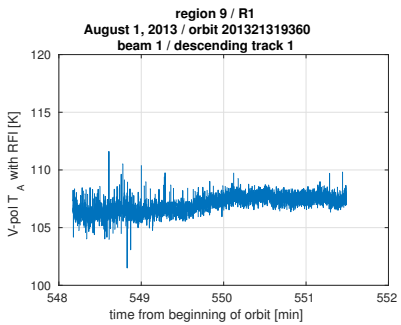
Simulated RFI

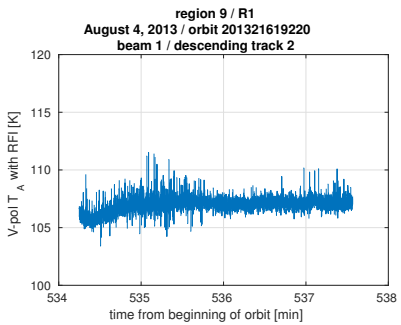


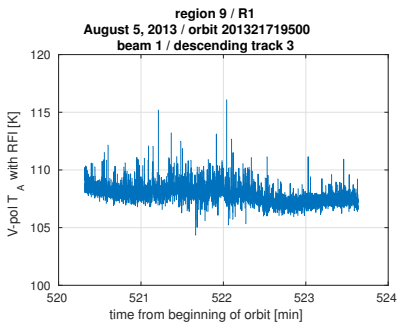




Simulated 10 ms T_A with RFI

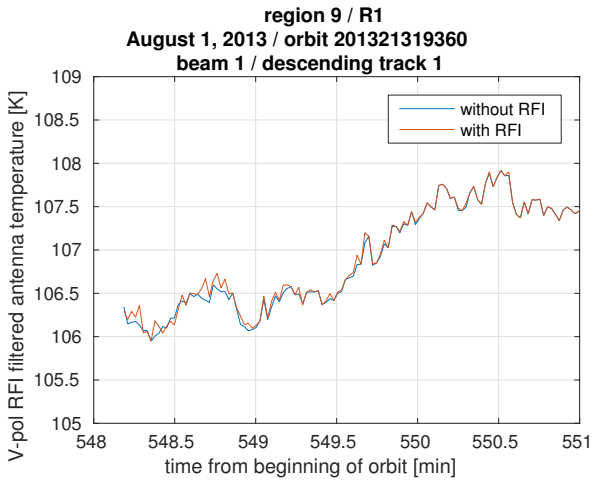


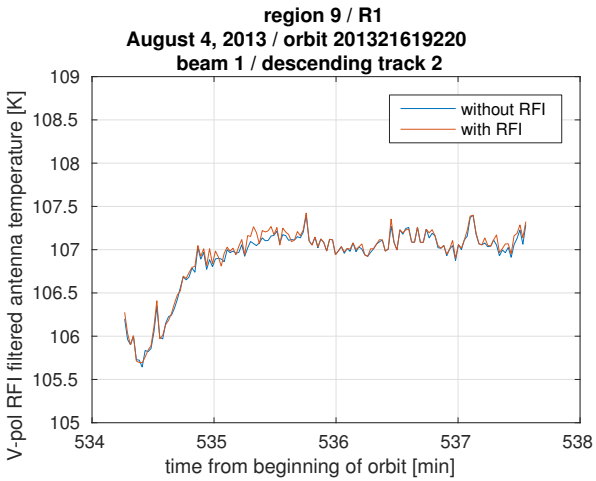


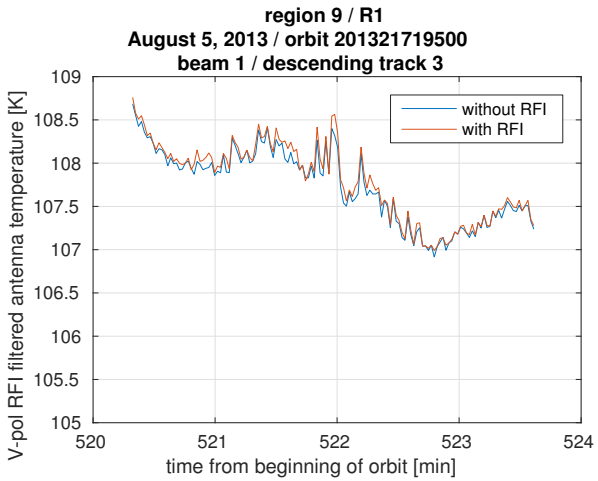


Results of Missed RFI Analysis

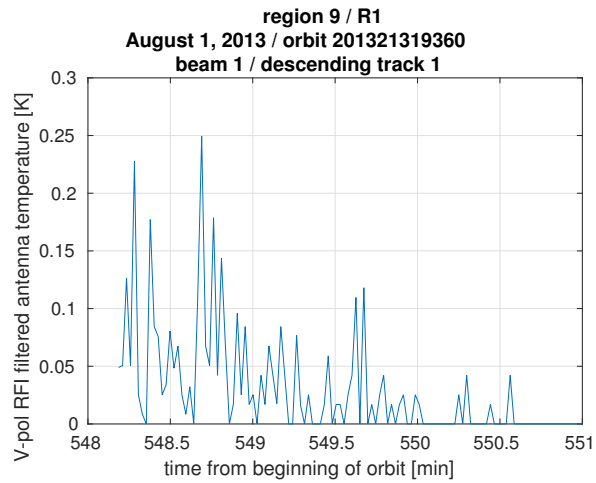
Simulated 1.44 s T_A, with and without RFI, after filtering

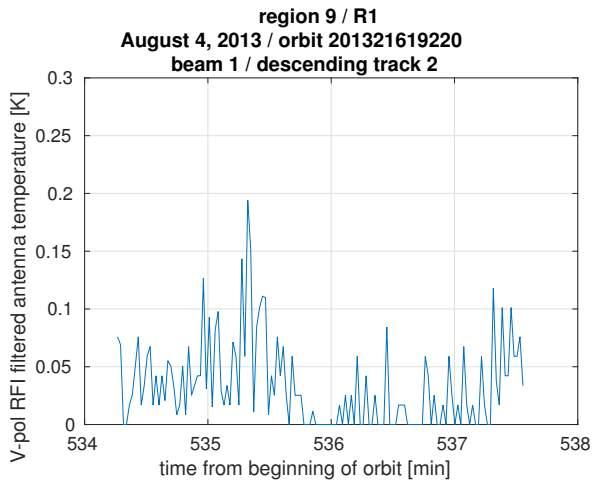


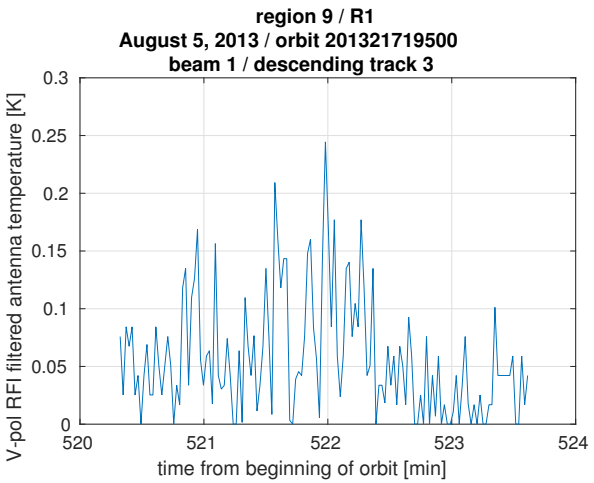




Estimated Missed RFI







Discussion

- results are preliminary, but RFI filter appears to miss a lot of low-level RFI under $\sim$0.2 K $\Rightarrow$ threshold adjustment needed?
- high rate of missed detection may be due to particular RFI modeling (no pulse behavior considered)
- procedure applied only to region where RFI disappeared in October 2013, more validation needed before attempting to use it globally
- future refinements:
  - use more cycles to increase accuracy of estimated histograms
  - introduce pulse behavior in RFI modeling
  - generalize to all ocean regions affected by RFI

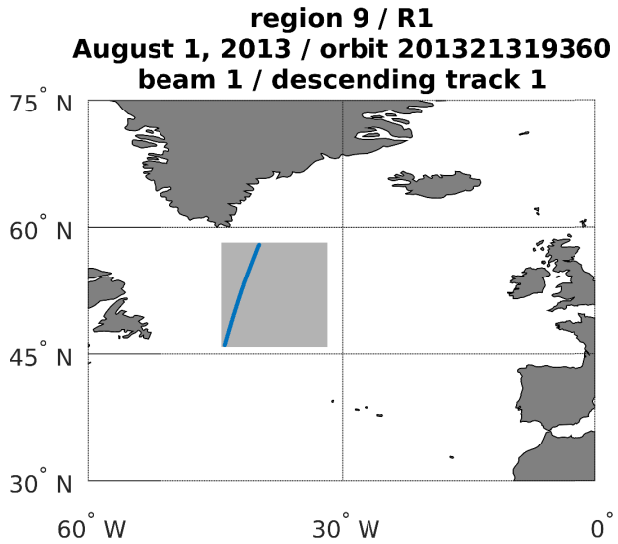
Conclusions

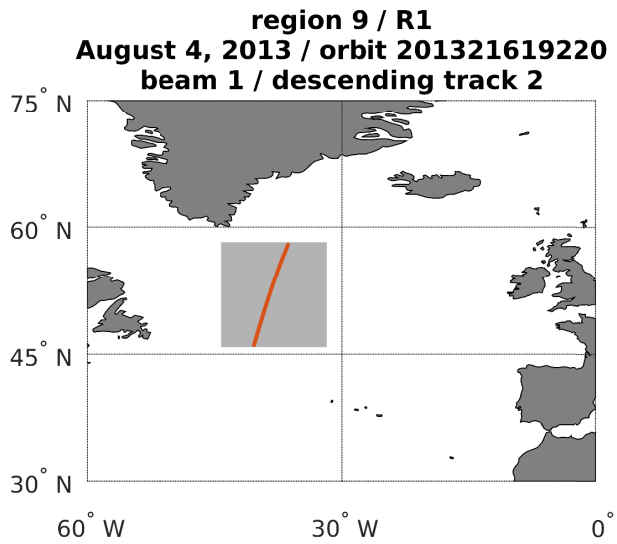
- adding values of standard deviation, skewness and kurtosis to the data product, and possibly an additional flag based on skewness and/or kurtosis, could help reduce missed RFI
- if results of estimation of missed RFI in low-RFI regions are confirmed and validated, they could have applications for:
  - 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

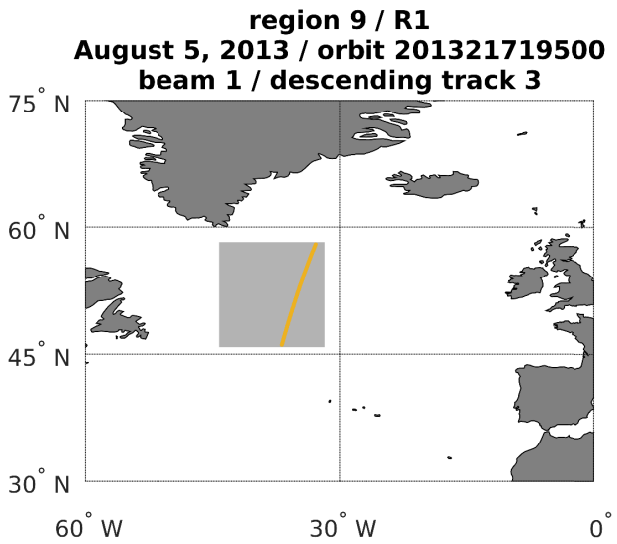
Backup Slides

- Location of RFI Regions
- Simulation of 10 ms Expected Antenna Temperature

Location of RFI Regions



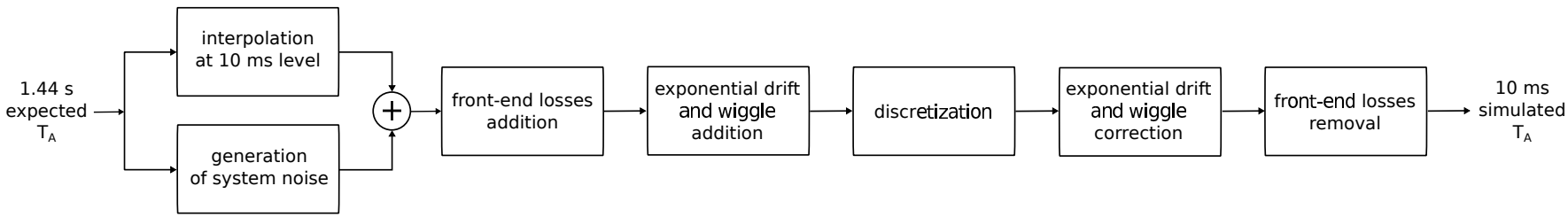




Simulation of 10 ms Expected Antenna Temperature

- the expected T_A at L2 level is linearly interpolated between 1.44 s points to obtain corresponding 10 ms points
- 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

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

where T_{rec} = 74.6 K, $\sqrt{B\tau}$ = 474.34 K

- the receiver front-end losses and the error caused by the exponential drift and wiggles are added to the 10 ms samples
- the resulting values are discretized
- 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