Radio Frequency Interference Characterization and Detection in L-band Microwave Radiometry

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Mustafa Aksoy

Graduate Program in Electrical and Computer Engineering

The Ohio State University

2015

Dissertation Committee:
Professor Joel T. Johnson, Advisor
Professor Fernando Teixeira
Professor Chi-Chih Chen
Abstract

Radio Frequency Interference (RFI) is a major issue in microwave radiometry and prevents correct estimation of geophysical parameters via remote sensing. This problem is reported even in the protected portion of the L-band (1400-1427MHz) which is allocated for only remote sensing of Earth from space. RFI contamination in radiometric measurements and the methods to mitigate it have previously been discussed in the literature. On the other hand, a comprehensive characterization of the RFI environment and an optimal RFI detection procedure which combines multiple RFI detection algorithms to effectively operate in that environment have yet to be presented. This dissertation aims to fill this gap for L-band microwave remote sensing research efforts. First, the RFI problem in microwave radiometry and previously developed RFI detection algorithms and their applications in current microwave radiometers are reported. Then, the L-band RFI environment is characterized in terms of its temporal, spectral, spatial, and statistical properties using space-borne and air-borne measurements from European Space Agency (ESA) and National Aeronautics and Space Administration (NASA) missions as well as local air-borne campaigns. It is demonstrated that RFI is a global problem, and its temporal, spectral and statistical properties may change significantly. Thus, classical RFI detection algorithms based on certain assumptions on these properties are insufficient to resolve the RFI problem and a more sophisticated approach is needed. This dissertation introduces
NASA’s Soil Moisture and Active Passive (SMAP) radiometer which was launched on January 31, 2015 as one of the first radiometers which implements such a multifaceted RFI detection technique. SMAP’s comprehensive multi-domain RFI detection approach is summarized and analyzed in terms of its performance under different RFI exposure scenarios using pre-launch and post-launch RFI studies. Finally, several improvements to the SMAP baseline algorithm, and future investigations to obtain a more efficient RFI mitigation are discussed.
Dedication

This dissertation is dedicated to my family.
Acknowledgments

I would like to thank my advisor, Professor Joel T. Johnson, and all ElectroScience Laboratory community for their constant help and supervision throughout my Ph.D. studies. I would also acknowledge Professor Fernando Teixeira, Professor Chi-Chih Chen, and Professor Yuji Kodama for serving on my Ph.D. defense committee. Finally, I would like to thank NASA Goddard Space Flight Center and Jet Propulsion Laboratory for providing the necessary data and documents for this study.
Vita

2005 ................................................................................................. Ankara Science High School, Turkey
2010 ................................................................................................. B.S. Electrical and Electronics Engineering, Bilkent University, Turkey
2014 ................................................................................................. M.S. Electrical and Computer Engineering, The Ohio State University
2010 to present ................................................................................ Graduate Research Associate, ElectroScience Laboratory, Department of Electrical and Computer Engineering, The Ohio State University

Publications

Journal Publications


**Conference Publications**


**Fields of Study**

Major Field: Electrical and Computer Engineering
# Table of Contents

Abstract ................................................................................................................................. ii

Dedication .............................................................................................................................. iv

Acknowledgments ................................................................................................................ v

Vita......................................................................................................................................... vi

Table of Contents ................................................................................................................... ix

List of Tables .......................................................................................................................... xiii

List of Figures ......................................................................................................................... xiv

Chapter 1: Introduction ......................................................................................................... 1

1.1 Microwave Radiometry ................................................................................................. 1

1.2 Radio Frequency Interference (RFI) Problem in Microwave Radiometry ............... 3

1.3 RFI Detection and Mitigation Algorithms used in Microwave Radiometry ........ 4

1.3.1 Time Domain Pulse Blanking ................................................................................... 4

1.3.2 Cross-Frequency Detection ..................................................................................... 5

1.3.3 Kurtosis and Other Normality Tests ......................................................................... 6

1.3.4 Polarimetry .............................................................................................................. 8

1.3.5 Spatial RFI Detection .............................................................................................. 8
1.3.6 Angular RFI Detection................................................................. 10

1.4 Summary of the Thesis.............................................................. 11

Chapter 2: Characterization of L-band RFI Environment ................. 12

2.1 SMOS and Aquarius Overview..................................................... 13

2.2 A Study of SMOS RFI over North America................................. 17

2.2.1 Artifact Removal for SMOS Dataset........................................... 23

2.2.2 Maps of 3\textsuperscript{rd} and 4\textsuperscript{th} Stokes Parameters.................. 28

2.2.3 SMOS CCDFs........................................................................... 30

2.3 Comparative Analysis of RFI in SMOS and Aquarius Measurements...... 33

2.3.1 Relationships between RFI in SMOS and Aquarius.................... 34

2.3.2 SMOS and Aquarius Spatial Statistics for Large RFI Sources.......... 40

2.3.3 Match-up Studies of Low Level RFI Sources............................ 43

2.3.4 Statistical Analysis of Low-Level RFI Match-ups ...................... 51

2.4 SMAP Validation Experiment 2012 ............................................ 55

2.4.1 SMAPVEX12 Campaign Overview ........................................... 57

2.4.2 PALS Radiometer and Digital Backend .................................... 60

2.4.3 Data Measurement and Calibration ......................................... 63

2.4.4 RFI Detection........................................................................... 70

2.4.5 Observed RFI Properties....................................................... 75
2.4.6 RFI Statistics ................................................................. 84

2.5 Summary and Discussion .................................................. 87

Chapter 3: Soil Moisture Active Passive (SMAP) as a Revolutionary RFI Detection and Mitigation Approach ................................................................. 89

3.1 SMAP Mission ..................................................................... 90

3.2 SMAP Radiometer ............................................................... 92

3.3 RFI Detection and Mitigation in SMAP ................................. 95

3.3.1 Pulse Detection ............................................................... 95

3.3.2 Cross Frequency Detection ............................................... 96

3.3.3 Kurtosis ......................................................................... 96

3.3.4 Polarimetry ................................................................. 97

3.4 Revisions in the SMAP Radiometer Timing and RFI Detection Algorithms ................................................................. 98

3.4.1 Revisions in Radiometer Timing ........................................ 99

3.4.2 Theoretical Standard Deviation Calculations ...................... 101

3.4.3 Double-Sided Detectors .................................................. 104

3.5 Performance of the SMAP RFI Detection and Mitigation Algorithms ................................................................. 107

3.5.1 The Comprehensive RFI Test .......................................... 108

3.5.2 Performance of the RFI Detection Algorithms: Theory vs Experiment ................................................................. 110

3.6 SMAP in Orbit: Analysis of Initial SMAP Data ...................... 119
3.7 Possible Improvements to SMAP RFI Detection and Mitigation Algorithms. 135

3.7.1 Alternative MPD Approaches to Reduce Data Loss.............................. 135

3.7.2 Updates in the Kurtosis Detector .................................................... 149

3.8 Summary and Discussion........................................................................... 167

Chapter 4: Conclusions .................................................................................. 169

Bibliography .................................................................................................. 173
List of Tables

Table 2.1: Incidence angles and footprint sizes for the three Aquarius beam .................. 16

Table 2.2: Footprint size for the three Aquarius beams and received RFI power ratios between SMOS and Aquarius for each Aquarius incidence angle ...................... 35

Table 2.3: Flight dates and durations for RFI data sampling (Primary campaign)........... 58

Table 2.4: Flight dates and durations for RFI data sampling (Denver flights).............. 59

Table 3.1: Measurements Collected by SMAP Radiometer in the Original Time Scheme (Black) and after Revisions (Red in Parenthesis)................................. 100

Table 3.2: RFI signals injected to the SMAP radiometer during the test .................... 109
List of Figures

Figure 1.1: Pulse Blanking and Cross-Frequency RFI Detection................................. 6

Figure 1.2: RFI can be detected (left) in spatial domain and (right) using polarimetry........................................................................................................... 9

Figure 1.3: RFI corrupted and clean brightness temperatures measured by the SMOS radiometer versus incidence angle. ......................................................... 10

Figure 2.1: (Left) Sample SMOS snapshot over the Iberian Peninsula. (Right) Portions of SMOS snapshot corresponding to Aquarius observation angles. ........................................................................................................... 15

Figure 2.2: Aquarius swath consisting of three beams .............................................. 17

Figure 2.3: (Upper) Max-hold and (lower) mean SMOS brightness temperatures (in K) in horizontal polarization over North America (all incidence angles) for November 2011 (both ascending and descending) ...................... 18

Figure 2.4: (Upper) Max-hold and (lower) mean SMOS brightness temperatures (in K) in horizontal polarization discarding RFI flagged measurements over North America (all incidence angles) for November 2011 (both ascending and descending) ........................................................................ 20

Figure 2.5: Level-1B images illustrating sidelobes of RFI sources affecting SMOS measurements (this figure is only an example and shows RFI sources independent from the data set used in this study) ....................................................... 22
Figure 2.6: (Left column) Level 1B and (right column) corresponding Level 1C SMOS data showing artifact corruption in (upper row) X and (lower row) Y polarizations.

Figure 2.7: Pulse detection over boresight accuracy error levels in artifact removal.

Figure 2.8: Swath reduction as the final step in artifact-removal algorithm.

Figure 2.9: (Upper) Max-hold and (lower) mean SMOS brightness temperatures (in K) in horizontal polarization discarding RFI flagged measurements and after artifact removal over North America (all incidence angles) for November 2011 (both ascending and descending).

Figure 2.10: (Upper) Max-hold and (lower) mean values of C parameter after artifact removal over North America (all incidence angles) for November 2011 (both ascending and descending).

Figure 2.11: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures in horizontal polarization.

Figure 2.12: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures in vertical polarization.

Figure 2.13: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures for the absolute value of 3rd Stokes Parameter.

Figure 2.14: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures for the absolute value of 4th Stokes Parameter.

Figure 2.15: Received RFI power for SMOS and Aquarius over Cincinnati.
Figure 2.16: SMOS (left) and Aquarius Beam 1 (right) overpasses over the RFI source in Cincinnati on May 3, 2012. For each SMOS snapshot measurements were taken from a fixed footprint (shown in white circle) where incidence angle and azimuth angle were $28.7^\circ$ and $37.3^\circ$. Red lines on the snapshots are constant incidence and azimuth angle lines. Aquarius on the other hand only provides incidence angle information and the boresight brightness temperature value for each footprint, thus all measurements for Beam 1 where the incidence angle is $28.7^\circ$ were taken into consideration. The position of the RFI source is shown in both SMOS and Aquarius swaths. .......................................................... 38

Figure 2.17: SMOS H pol Max-hold brightness temperature map (in K) for original (top) and non-RFI flagged (bottom) data. .......................................................... 41

Figure 2.18: Aquarius H pol Max-hold brightness temperature map (in K) for original (top) and RFI filtered (bottom) data. All beams are combined. .................. 42

Figure 2.19: Locations of “Low Level” RFI observations in Aquarius.................. 44

Figure 2.20: 138 Aquarius H-Pol “low-level” RFI Locations in South America............................................................................................................................. 45

Figure 2.21: Horizontally Polarized SMOS brightness temperature (in K) over “Low-Level” RFI locations vs Incidence Angle.............................................. 46

Figure 2.22: SMOS brightness temperatures (in K) over “Low-Level” RFI locations vs Incidence Angle for (left) 3rd Stokes Parameter and (right) 4th Stokes Parameter................................................................................................. 47

Figure 2.23: The Location of the Low Level RFI Source detected by Aquarius on May 6, 2012.......................................................... 48

Figure 2.24: Detected Source in Aquarius Brightness Temperature Data (in K) on May 6, 2012. The Measurements that correspond to the RFI source are circled.................................................. 48

xvi
Figure 2.25: SMOS overpass (in K) of the RFI source from Figure 2.23 on May 1, 2012. The measurement that corresponds to the RFI source is encircled.

Figure 2.26: SMOS measurements (in K) over the region of the RFI source in Figure 2.23 vs incidence angle on May 1, 2012. The outlier is encircled.

Figure 2.27: SMOS measurements (in K) over another RFI source at lat:-37.1, lon:-59.6 vs incidence angle on May 1, 2012. Outliers are again encircled.

Figure 2.28: Locations of SMOS measurements as a function of incidence angle that include more than 10 samples (green and red) and the ones that have less than 20K standard deviation (green).

Figure 2.29: Scatter Plots of SMOS and scaled Aquarius RFI (in K) in South America. These plots illustrate RFI detected in SMOS measurements using the detection algorithm (with 2 different thresholds) vs. RFI measured by Aquarius scaled according to Table 2.2. Red lines indicate where these two match each other according to the theory suggested in part 2.3.1.

Figure 2.30: Flight Route on June 7, 2012 during the Primary Campaign. The aircraft flew at two altitude levels: ~3500ft (red) and ~1000ft (blue).

Figure 2.31: Flight Routes on July 22, 2012 (green) and July 23, 2012 (black) over the Denver, CO metropolitan area.

Figure 2.32: Twin Otter Aircraft and System Electronics: (top left) the Aircraft, (top right) PALS Antenna mounted at the bottom of the aircraft, (bottom left and right) System Electronics on the aircraft.

Figure 2.33: Spectra of the PALS IF Signal (left) and the digital backend Input Signal (right) in theory. Since digital backend signal occupies the 6th Nyquist zone of the ADC spectrally reversed, 15MHz and 35MHz corresponds to 1.42GHz and 1.4GHz RF frequencies respectively.
Figure 2.34: Uncalibrated Power Spectrogram (top) and Real and Imaginary Kurtosis Spectra (bottom) examples generated using single V-pol digital backend file corresponding July 22, 2012 16:40:16 UTC. .................................................. 64

Figure 2.35: Antenna and Calibration States in V-pol PALS measurement cycle. ................................................................. 67

Figure 2.36: (Top left) V-pol raw power vs frequency over a single 350μs integration period and (top right) brightness temperature values after calibration. (Bottom left) Gain and (bottom right) offset values for the corresponding 1-second file. ................................................................. 68

Figure 2.37: Brightness Temperature Spectrogram obtained after the calibration of the power spectrogram in Figure 2.34.................................................. 69

Figure 2.38: Digital Backend Data Processing Steps. ................................................................. 70

Figure 2.39: RFI Detection Algorithm over Spectrograms. .................................................. 72

Figure 2.40: A Tb Spectrogram (top left), the same spectrogram after smoothing (top right), RFI flags after thresholding (bottom left) and Grouped RFI bins as discrete signals with distinct colors (bottom right) where each color corresponds one RFI signal. .................................................. 74

Figure 2.41: Kurtosis Spectra for the Power Spectrum shown in Figure 2.40. Red lines indicate the threshold values.................................................. 74

Figure 2.42: Narrowband Continuous RFI signal. Note that the continuous signal is not visible in time domain whereas the Kurtosis value less than 3 indicates the strong continuous RFI in Kurtosis Spectra.................................................. 76

Figure 2.43: Narrowband Pulsed RFI signal. Pulses can be observed in the spectrogram and in time domain. Higher Kurtosis values also imply pulsed type interference in the higher end of the 3-dB RF passband.................................................. 77
Figure 2.44: Chirp RFI signal. Interference is invisible in time and Kurtosis domains. RFI can be said to be a chirped pulse signal with changing frequency.......................................................................................................................... 77

Figure 2.45: Wideband Pulsed RFI signal. Strong wideband pulses can be seen in the spectrogram and in time domain. Kurtosis values are also larger than 3 for all frequencies although the threshold is biased due to the severe corruption.................................................................................................................................. 78

Figure 2.46: Wideband Continuous RFI signal. RFI can be seen in the upper half of the spectrogram. In Kurtosis domain the signal is not detectable whereas it appears as an increase in the background noise in time domain. ...................... 78

Figure 2.47: V-pol (left) and H-pol (right) Brightness Temperature vs Time and Frequency over Denver, CO on July 22, 2012.................................................................................................................. 80

Figure 2.48: V-pol (left) and H-pol (right) Brightness Temperature vs Time and Frequency over Denver, CO on July 23, 2012.................................................................................................................. 80

Figure 2.49: V-pol (left) and H-pol (right) Brightness Temperature vs Time and Frequency over Winnipeg Croplands on July 19, 2012 for an hour flight........................................................................................................................................... 81

Figure 2.50: V-pol (top) and H-pol (bottom) real (right) and imaginary (left) kurtosis vs time and frequency over Denver, CO on July 22, 2012. ........................................... 82

Figure 2.51: V-pol (top) and H-pol (bottom) real (right) and imaginary (left) kurtosis vs time and frequency over Denver, CO on July 23, 2012. ........................................... 83

Figure 2.52: Bandwidth vs Duration plots for H-pol (top) and V-pol (bottom) RFI observed over Denver, CO. Red lines indicate the minimum detectable duration as RFI detection was performed over spectrograms where the time resolution was 350μs.............................................................................................................. 85
Figure 2.53: CCDF plot of detected RFI over Denver, CO during SMAPVEX12 campaign. .............................................................. 86

Figure 3.1: A picture of the SMAP instrument in orbit [53] ................................................................. 91

Figure 3.2: SMAP Mission Science Goals and Instrument Features [53] ......................... 92

Figure 3.3: SMAP Radiometer Block Diagram ................................................................. 93

Figure 3.4: SMAP Polyphase Filter Bank Frequency Response ............................................. 94

Figure 3.5: Illustration of the original radiometer timing sequence of one SMAP footprint. Each footprint contained 11 packets for data measurement and 1 packet of calibration. Each packet consists of 4 PRIs................................................................. 94

Figure 3.6: An example SMAP footprint with the original timing scheme. Subband Data (11x1.2ms) and Fullband Data (44x350µs) are shown with pulse blanking and fullband kurtosis thresholds indicated by dashed red lines. The final maximum probability of detection flags are also displayed where red bins indicate RFI corruption. .............................................................................................. 98

Figure 3.7: Current radiometer timing after revisions showing one SMAP radiometer footprint with 8 packets for data measurement and 4 packets for calibration. .............................................................................................. 100

Figure 3.8: Another SMAP footprint with the updated radiometer timing scheme. One footprint is 32x350µs in fullband domain and 8x1.4ms in subband domain. Notice that RFI in the footprint is detected by the MPD algorithm as described in Section 3.3. .............................................................................................. 101

Figure 3.9: Standard deviations of SMAP subband measurements computed with empirical and theory based approaches at (left) 1.2ms and (right) 9.6ms resolution .............................................................................................. 103
Figure 3.10: An illustration of a Gaussian distribution of RFI free temperature samples with a mean of 290 K. The threshold cutoff line represents the mean (μ) of the lowest 90% of the data plus a number (β) of standard deviations (σ). The non-flagged temperature samples now have a mean that is biased low, represented by the green line. 

Figure 3.11: Fullband pulse detector one-sided (left column) vs two-sided (right column) showing bias induced (top row) and percentage of false alarms (bottom row) for original (blue) and formula based (red) standard deviation computations.

Figure 3.12: ROC performance results for the cross-frequency of 9.6ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2 ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against CW RFI at increasing true power levels (in Kelvin).

Figure 3.13: ROC performance results for the cross-frequency of 9.6ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2 ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against Pulsed RFI with 2µs pulse width.

Figure 3.14: ROC performance results for the cross-frequency of 9.6ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2 ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against Pulsed RFI with 51µs pulse width.

Figure 3.15: ROC performance results for the cross-frequency of 9.6ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against Pulsed RFI with 95µs pulse width.
Figure 3.16: AUC performance results for the cross-frequency of 9.6ms integration time, cross-frequency of 1.2ms integration time, pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against a sweep of CW, 2 us pulsed, 51us pulsed, and 95us pulsed interference. ................................................................. 119

Figure 3.17: SMAP Max-hold H-pol Brightness Temperatures between May 1, 2015 and May 8, 2015. .................................................................................................................. 121

Figure 3.18: SMAP Max-hold H-pol Brightness Temperatures between May 1, 2015 and May 8, 2015 after RFI mitigation. ............................................................... 122

Figure 3.19: SMAP Max-hold H-pol RFI between May 1, 2015 and May 8, 2015.......................................................................................................................... 123

Figure 3.20: SMAP Max-hold H-pol Brightness Temperatures over Europe between May 1, 2015 and May 8, 2015 ............................................................................. 124

Figure 3.21: SMAP Max-hold H-pol Brightness Temperatures over Europe between May 1, 2015 and May 8, 2015 after RFI mitigation ........................................... 125

Figure 3.22: SMAP Max-hold H-pol RFI over Europe between May 1, 2015 and May 8, 2015. ............................................................................................................. 126

Figure 3.23: SMAP brightness temperature measurements on April 22, 2015 over Western Europe before and after RFI mitigation. Notice the unmitigated RFI over Madrid, Spain. ............................................................................ 127

Figure 3.24: Time Series Fullband Brightness Temperatures over Madrid on April 22, 2015. Red Circles indicate flagged data by Pulse Detector....................... 128

Figure 3.25: Time Series Fullband Kurtosis values over Madrid on April 22, 2015. Red Circles indicate flagged data by Fullband Kurtosis Detector ................. 128
Figure 3.26: Time Series Fullband 3rd and 4th Stokes Parameters with Horizontal Polarization Brightness Temperatures over Madrid on April 22, 2015.......................... 129

Figure 3.27: Subband Brightness Temperatures over Madrid on April 22, 2015................................................................. 130

Figure 3.28: Subband Kurtosis values over Madrid on April 22, 2015........................... 130

Figure 3.29: Subband 3rd (left) and 4th (right) Stokes Parameters over Madrid on April 22, 2015................................................................. 131

Figure 3.30: SMAP fullband and subband Data for a RFI contaminated footprint. Red circles indicate the pulse detector flags, black crosses and circles indicate 9.6ms and 1.2ms cross frequency detector flags respectively. Channels 3-16 are completely discarded by the MPD flag. .................... 132

Figure 3.31: SMAP RFI Detection Rates by Channel ......................................................... 133

Figure 3.32: SMAP Overall RFI Detection Rates .......................................................... 134

Figure 3.33: Narrowband Pulsed RFI simulation. Only a single frequency channel includes RFI but other channels are also flagged by the baseline method................................................................. 136

Figure 3.34: Narrowband continuous wave RFI simulation. An entire footprint is lost by the baseline algorithm due to fullband Kurtosis flagging................................................................. 136

Figure 3.35: Missed Wideband RFI pulse by the Suggested RFI detection scheme................................................................. 138

Figure 3.36: Missed Narrowband Continuous wave RFI by the Suggested RFI detection scheme................................................................. 139
Figure 3.37: Mixed RFI detection Approach. Notice the reduced data loss compared with the Baseline Maximum probability of detection Algorithm. ............................ 140

Figure 3.38: Weakness of the Lowest Approach in case of the presence of multiple RFI sources. Notice that both subband and fullband products contain RFI ................................................................................................................. 141

Figure 3.39: Performance of the alternative RFI detection algorithms versus the baseline maximum probability of detection algorithm. The mixed approach is labeled as “Limited Full”................................................................................................................. 143

Figure 3.40: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “baseline” MPD approach ................................................................................................................. 144

Figure 3.41: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “only subband” MPD approach ................................................................................................................. 145

Figure 3.42: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “only fullband” MPD approach ................................................................................................................. 146

Figure 3.43: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “mixed” MPD approach ................................................................................................................. 147

Figure 3.44: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “lowest” MPD approach ................................................................................................................. 148

Figure 3.45: Probability of Detection vs False Alarm Rate for a pulse sinusoidal RFI where d=100%, λ=15 (left) and λ=20 (right). ................................................................................................................. 161
Figure 3.46: Probability of Detection vs False Alarm Rate for a pulse sinusoidal RFI where $\lambda=15$, $d=10\%$ (left) and $d=50\%$ (right).

Figure 3.47: Performance of the AND (green) and OR (red) Kurtosis detection algorithms versus the baseline (black) kurtosis algorithm.

Figure 3.48: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, percentage of data flagged by the (bottom-left) fullband and (bottom-right) subband “baseline” kurtosis algorithms.

Figure 3.49: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, percentage of data flagged by the (bottom-left) fullband and (bottom-right) subband “OR” kurtosis algorithms.

Figure 3.50: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, percentage of data flagged by the (bottom-left) fullband and (bottom-right) subband “AND” kurtosis algorithms.
Chapter 1: Introduction

This thesis describes studies to characterize the Radio Frequency Interference (RFI) problem for L-band passive microwave remote sensing, and examines current and future RFI detection and mitigation algorithms to overcome this problem.

Chapter 1 introduces the fundamental concepts of passive microwave remote sensing, the underlying reasons for the RFI problem in microwave radiometry, and RFI detection algorithms currently implemented in microwave radiometers. Thus, this section provides the motivation behind this study and prepares for a better understanding of the rest of the thesis.

1.1 Microwave Radiometry

Radiometers are passive remote sensing instruments that measure the electromagnetic radiation emitted by objects. All objects receive electromagnetic energy from their surroundings and absorb part of it; the remaining energy is reflected. Objects that absorb all of the incident energy and do not reflect any of it are known as “blackbodies”. In thermal equilibrium, objects emit all energy that they absorb. For blackbodies this radiated power (in the low frequency limit such as microwave frequencies) is formulated by Plank’s law under Rayleigh-Jeans approximation as:
\[ P = kTB \]  \hspace{1cm} \text{(1.1)}

where \( k \) is Boltzmann’s constant, \( T \) is the physical temperature of the blackbody in Kelvin, and \( B \) is the bandwidth in Hertz of the receiver that measures the power [1]. In real life there are no blackbodies, and objects do not absorb and emit all the energy they receive which means that the radiated power by them is less than (1.1). The ratio of the power radiated by the real object to the power radiated by a blackbody at the same temperature and in the same bandwidth is defined by a parameter called the “emissivity” (\( e \)):

\[ e = \frac{P_o}{P_b} \]  \hspace{1cm} \text{(1.2)}

where \( P_o \) and \( P_b \) are the power radiated by the real object and the blackbody respectively. The power emitted by the object can also be defined with another parameter called the “brightness temperature”. The brightness temperature of an object indicates the temperature of a blackbody which would radiate the same amount of electromagnetic power as that object at a certain temperature [2]. Therefore, the emissivity can also be written as the ratio of the brightness temperature of the object (\( T_b \)) to its physical temperature (\( T \)).

\[ e = \frac{T_b}{T} \]  \hspace{1cm} \text{(1.3)}

Since the emitted power depends on the reflected power, the emissivity of an object is a function of its “reflectivity”:

\[ e(f, \Theta, p) = 1 - |R(f, \Theta, p)|^2 \]  \hspace{1cm} \text{(1.4)}

where \( R \) is the Fresnel reflection coefficient as a function of frequency \( f \), incidence angle \( \Theta \), and the polarization \( p \). Note that this equation is exact only for specular reflection, and although the physics behind the phenomenon is the same, computation of the rough surface...
emissivity requires more advanced formulations based on the roughness parameters [3]. The Fresnel reflection coefficient is also a function of dielectric constant of the medium which is related to the physical properties of it such as its temperature and composition. Therefore, the emissivity of an object, and thus the amount of electromagnetic radiation from it, depends on the physical properties of the object and environmental parameters which shape it. Under certain circumstances, those properties and parameters can be retrieved since the 1960’s, and radiometers have been actively used for Earth remote sensing to retrieve the geophysical parameters of Earth to understand our planet better. The potential of observing the features of atmosphere, land, ocean and ice surfaces on global scales have led to rapid developments of different types of microwave radiometers, and these radiometers have been operating over a wide range of frequencies according to their specific purposes.

1.2 Radio Frequency Interference (RFI) Problem in Microwave Radiometry

Radiometers measure the electromagnetic power captured by their antenna but do not identify the sources of that power. Considering the fact that the electromagnetic spectrum is very densely occupied and used by radars, radios, TVs, communication systems and other applications besides microwave radiometers, this power usually comes from different sources. Therefore, for microwave radiometers that measure natural emissions from atmosphere and Earth to retrieve some of their properties, interference from other sources, i.e. Radio Frequency Interference (RFI), is a major problem which leads to
erroneous estimation of physical parameters [4]-[8]. To overcome this problem, numerous studies has been conducted to develop RFI detection and mitigation algorithms for microwave radiometry; the results have been implemented in past and currently operational remote sensing missions [9]-[18]. Section 1.3 introduces some of these algorithms.

1.3 RFI Detection and Mitigation Algorithms used in Microwave Radiometry

Results of RFI studies have demonstrated that the interference signals have distinctive features which separate them from natural geophysical emissions. This section summarizes the RFI detection algorithms which utilize this fact and are widely used in current passive remote sensing applications.

1.3.1 Time Domain Pulse Blanking

Time domain pulse blanking algorithms search for pulses in the received power concentrated in time based on the assumption that natural emissions are not pulsed signals. Received powers that exceed a certain threshold above the estimated RFI-free power level (power outlier) are flagged as RFI [19]. This threshold in general can be expressed as:

$$T = c(P_{fa}) \times \sigma$$

Where $c(P_{fa})$ is a constant that depends on the probability of false alarm and $\sigma$ is the standard deviation of estimated RFI-free power. The threshold should be set according to the accuracy and false alarm rate requirements for the radiometer. Another important parameter for pulse-blanking algorithms is the integration time of the receiver. Shorter
integration times increase noise fluctuations above the threshold level which may lead to high false alarm rates, whereas longer integration periods may lead to lower probability of detection by averaging out the RFI pulses. An optimal value for the integration time therefore should be found, and it is reported that for pulse blanking algorithms the optimal radiometer integration time is the expected value of the RFI pulse lengths [19].

Pulse blanking algorithms are developed to detect pulsed interference whereas they are inefficient for continuous type RFI. Radars usually emit pulsed signals, so pulse blanking is desirable to remove RFI from such sources.

1.3.2 Cross-Frequency Detection

Cross-Frequency algorithms are basically the same as the pulse blanking algorithms but they operate in the frequency domain and search for power outliers concentrated in frequency based on the assumption that geophysical emissions are smooth functions of frequency over a limited bandwidth. A threshold is applied to the received power at different frequencies and the outliers are flagged as RFI. Such algorithms require radiometers with multiple frequency channels or are applied after the spectrum of the received power is obtained via a Fourier transform [20]. Therefore, more complex hardware designs are required, but current developments in digital circuitry have made cross-frequency algorithms possible. The bandwidth of the frequency channels (the frequency resolution of the spectra) and the threshold values are the important parameters in cross-frequency algorithms, analogous to the integration time and threshold level in pulse blanking algorithms.
Band limited interference is the main target for cross-frequency algorithms. Thus, many continuous type interference missed by pulse-detection algorithms can be detected by cross-frequency algorithms.

Figure 1.1 depicts an example of pulse blanking and cross-frequency RFI detection.

Figure 1.1: Pulse Blanking and Cross-Frequency RFI Detection.

1.3.3 Kurtosis and Other Normality Tests

Due to the Central Limit Theorem, natural emissions, thermal noise emitted from atmosphere, land and oceans, are assumed to be Gaussian signals unlike man-made radiations. To utilize this assumption, a parameter called ‘Kurtosis’ expressed in the equation below is defined as a function of the moments of the received power:

\[ K = \frac{m_4}{m_2^2} \]  

(1.6)
where $m_n$ is the $n^{th}$ central moment of the signal. For geophysical emissions, the kurtosis estimate itself is a Gaussian random variable with a mean of 3 and standard deviation of $24/N$ where $N$ is the number of samples used to compute it, assuming $N$ is large enough. On the other hand, for human made non-Gaussian signals the kurtosis deviate from 3 significantly except for pulses with 50% duty cycle where the kurtosis is exactly 3. Thus, it is possible to eliminate RFI by setting a maximum and minimum limit for acceptable Kurtosis values for the received power [21].

Kurtosis algorithms can be applied in both the time and frequency domain since the spectrum of a normal signal is also normal. The definition of the Kurtosis for a signal in a particular frequency channel is:

$$SK = \frac{n}{(n-1)} \times \left\{ (n+1) \times \frac{m_4}{m_2^2} - 2 \right\}$$

where $n$ is the number of samples in the frequency channel used to compute Kurtosis. This new Kurtosis definition in frequency domain is called "Spectral Kurtosis" and very close to unity for Gaussian emissions and deviates rapidly for RFI (Note that if $n$ becomes very large, this formula is equivalent to the time domain Kurtosis) [22].

The kurtosis test in the time domain is very efficient for pulsed interference with small duty cycles whereas its efficiency significantly suffers for larger duty cycles and continuous interference. However, pulses with larger duty cycles in the time domain (>15%) and continuous RFI can be detected better with Spectral Kurtosis in the frequency domain, and it does not have a blind point for signals with 50% duty cycle in the time domain. Its blind point appears only if the duty cycle of the signal is 50% on that specific frequency channel (or that frequency bin in the spectra). Therefore, when it is applicable,
it is suggested to use the time and spectral domain kurtosis tests together for a more efficient RFI detection [23].

Besides Kurtosis there are other normality tests in the literature, however they are not as widely used as Kurtosis algorithms [24].

1.3.4 Polarimetry

Radiation from natural sources is usually unpolarized, and the correlation between its horizontal and vertical polarization components is negligible. Therefore, 3rd and 4th Stokes parameters [25] are expected to be very low for thermal noise emissions from Earth and atmosphere. Using this fact, several RFI detection algorithms have been developed which detects the incoming power as RFI if 3rd and 4th Stokes parameters are higher than a certain threshold [26].

1.3.5 Spatial RFI Detection

Some microwave radiometers use imaging tools for measurements or brightness temperature maps that are created in post-processing. Thus, it is possible to search for brightness temperature anomalies (hot or cold spots) in such images and maps and flag them as RFI since thermal emission from Earth is expected to be a smooth function of geographical properties of the surface [27]. The spatial resolution of the radiometer, i.e. the antenna footprint, determines the measured RFI and the accuracy of spatial RFI detection, as expressed in the density of interferers equation:
\[ T_{RFI} = \frac{EIRP}{\text{Antenna Footprint Area}} \quad (1.8) \]

where \( EIRP \) is the equivalent isotropically radiated power for the RFI source. As the resolution decreases, the measured RFI power becomes lower. Thus, very low spatial resolutions may lead to unmitigated RFI contamination especially for interference at the low-to-medium level.

Figure 1.2 shows spatial maps of vertically polarized brightness temperature and 3\(^{rd}\) Stokes parameter measured by the Windsat radiometer [8]. The “hot spots” on these maps are due to RFI as geophysical emissions do not possess strong polarimetric signature and are smooth functions of space.

Figure 1.2: RFI can be detected (left) in spatial domain and (right) using polarimetry.
1.3.6 Angular RFI Detection

Natural emissions are not radiated strongly in particular directions. Although the radiation is a function of incidence angle according to the emission theory, it is expected to be a smooth function of angle. Thus, some RFI detection algorithms examine the incoming power as a function of incidence angle and flag rapid changes in the angular domain as RFI [21]. Figure 1.3 depicts a plot of RFI corrupted and clean brightness temperatures measured by the SMOS radiometer [19] versus incidence angle. As seen from the figure, RFI corrupted measurements show a strong emission between 15° and 20° incidence angles.

Figure 1.3: RFI corrupted and clean brightness temperatures measured by the SMOS radiometer versus incidence angle.
1.4 Summary of the Thesis

This thesis first focuses on studies to characterize the RFI problem for L-band microwave radiometry and then discusses efficient ways to reduce it.

Chapter 2 analyzes measurements from space-borne and air-borne radiometers to characterize the current RFI environment in L-band. Data from the European Space Agency’s (ESA) Soil Moisture and Ocean Salinity (SMOS) and the National Aeronautics and Space Agency’s (NASA) Aquarius satellite based radiometers are examined to demonstrate the RFI problem on global scales. Then, the data measured during an airborne campaign called the SMAP Validation Experiment 2012 (SMAPVEX12) is examined to detail the temporal, spectral, and statistical properties of the L-band RFI environment. These properties are discussed to understand the requirements for efficient RFI detection and mitigation algorithms.

Chapter 3 introduces a novel RFI detection and mitigation approach implemented in NASA’s Soil Moisture Active Passive (SMAP) space-borne radiometer which was recently launched in January 2015. An overview of the radiometer itself is given and the RFI detection and mitigation procedure is summarized. The performance of SMAP RFI detection algorithms under different RFI scenarios according to the characterized RFI environment in Chapter 2 is studied in detail. This chapter also discusses possible improvements to the baseline SMAP RFI detection and mitigation approach, and demonstrates initial SMAP observations from space.

The last chapter, Chapter 4, concludes the thesis by summarizing the results and contributions of this research and future studies based on them.
Chapter 2: Characterization of L-band RFI Environment

ESA’s SMOS and NASA’s Aquarius missions utilize space-borne L-band radiometers to provide Soil Moisture and Ocean Salinity information on a global scale. Despite the fact that these radiometers operate in the protected part of the L-band spectrum (1400MHz-1427MHz) which is exclusively allocated to the earth observation remote sensing systems, their measurements have been reported to be severely contaminated by RFI [28]- [33]. Thus, analyzing SMOS and Aquarius data is helpful to understand the universal RFI problem for L-band passive remote sensing systems.

Although they provide data on a global scale, the SMOS and Aquarius radiometers lack sufficient temporal, spectral and spatial resolution to fully characterize the RFI environment. Therefore, air-borne campaigns have been conducted using more advanced radiometer and processing systems to examine the properties of the RFI signals in smaller regions near urban areas where it is possible to capture a variety of RFI signals due to the density of human-made emissions (telecommunication systems, radars etc.). The SMAP Validation Experiment 2012 (SMAPVEX12) is one of the most recent such campaigns during which an L-band radiometer was operated to collect high resolution data over urban and rural areas in the United States and Canada [34].

This chapter examines the measurements from space-borne SMOS and Aquarius radiometers as well as the measurements from SMAPVEX12 campaign to characterize the
L-band RFI environment. In Section 2.1, an overview of SMOS and Aquarius radiometers, and the RFI detection techniques implemented in them are examined. Section 2.2 presents a study of SMOS RFI over North America. Artifacts in SMOS data, removal of them, and RFI observed by the SMOS radiometer are discussed. The next section analyses SMOS and Aquarius measurements comparatively especially for low-level RFI. The global RFI environment seen by the two space-borne L-band radiometers is presented, and RFI detection techniques which can be applied in addition to their built-in algorithms for a more complete RFI mitigation are demonstrated. Section 2.4 describes the SMAPVEX12 airborne campaign. The goal of the campaign, the hardware used to collect data, signal processing and RFI detection algorithms are explained in detail. Finally, the observed RFI properties and their implications for current and future RFI detection approaches in L-band microwave radiometers are examined.

2.1 SMOS and Aquarius Overview

SMOS is an ESA mission that has been operational since November 2009. The goal of the mission is to accurately estimate soil moisture and ocean salinity on Earth’s surface. The SMOS L-band interferometric radiometer has a single frequency channel (1400-1427 MHz) and provides fully polarimetric multi-angular brightness temperature measurements averaged over a 1.2 second time interval. The multi-angular measurements are created by combining multiple “snapshots” of brightness temperatures (Figure 2.1) during an overpass of a given Earth location. The spatial resolution of a given observation varies with the
observation angle, but is nominally approximately 40kmx40km. The SMOS processor is capable of reporting brightness temperatures up to 10,000 K or more, and such brightness temperatures are observed in some situations. More details about the instrument can be found in [35]. Since there is not enough temporal and spectral resolution, pulse-detection, cross-frequency and normality based RFI detection algorithms are not applicable in SMOS. Therefore, SMOS applies a special grid on Earth surface and uses a spatial RFI detection technique that holds the statistics of brightness temperature measurements for each grid point. For each location, the probability of brightness temperature measurements exceeding a certain threshold is computed from a time series of SMOS data. The RFI detection algorithm creates an RFI mask using the locations where this probability is higher than a pre-determined value, and flags the data under this mask as RFI corrupted [27]. SMOS RFI flags based on this algorithm are inefficient unless RFI contamination is very high, in the order of ~100K [36], therefore several polarimetric and angular RFI detection algorithms have been developed for further analysis of SMOS data as post processing [21]. It is shown in [28] and [36] that the interferometric properties of the SMOS radiometer can further complicate the RFI problem. RFI artifacts caused by aliasing and/or sidelobes of large RFI sources cause RFI corruption in locations that may be widely separated from the true source location.
Aquarius is a NASA mission to estimate the ocean salinity and water circulations in world oceans. It operated between August 2011 and June 2015, and its single-frequency channel (1400-1427 GHz) radiometer provided horizontal, vertical, and 3rd Stokes brightness temperatures with three separate real aperture antenna at three different but fixed incidence angles (28.7°, 37.8°, 45.6°). Therefore, it lacked the multi-angular property of SMOS. Table 2.1 provides the spatial resolution for each Aquarius beam; the three Aquarius footprints together form the swath of Aquarius which is approximately 390km (Figure 2.2). The time resolution in Aquarius measurements is 10ms which is much finer than that of SMOS; Aquarius could utilize a simple time domain pulse blanking algorithm for RFI detection. The algorithm runs a local window over the time domain data, computes the local mean and eliminates outliers higher than the mean by a certain multiple of the standard deviation in the same window. The same procedure is repeated iteratively to remove pulse type interference. The algorithm is however inefficient against continuous interference as well as pulsed signals with pulse widths much shorter than 10ms. Details
of the instrument and the RFI detection algorithm can be found in [37] and [38].
Polarimetric detection methods are also possible with the 3rd Stokes Aquarius channel, but multi-angular strategies are inapplicable since Aquarius observed a given Earth location at only a single incidence angle at a given time. Aquarius provided both RFI contaminated and RFI free data in its science products at 1.44 second time resolution, so that detected RFI can be computed by taking the difference between them. In contrast to the very high brightness amplitudes that can be measured by SMOS, it was observed that the Aquarius processor discarded any measurements having brightness temperatures greater than \(~1000K\), and reported no corrected brightness when large RFI contributions are present. It is noted that Aquarius was also subject to RFI “artifacts” through contributions entering the radiometer through antenna sidelobes. However, the aliasing problem of SMOS was not present since a real aperture system was used.

<table>
<thead>
<tr>
<th>Incidence Angle</th>
<th>Beam 1</th>
<th>Beam 2</th>
<th>Beam 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>28.7°</td>
<td>37.8°</td>
<td>45.6°</td>
<td></td>
</tr>
<tr>
<td>Footprint Size</td>
<td>76kmx94km</td>
<td>84kmx120km</td>
<td>96kmx156km</td>
</tr>
</tbody>
</table>

Table 2.1: Incidence angles and footprint sizes for the three Aquarius beams
2.2 A Study of SMOS RFI over North America

SMOS measurements over North America were examined as the first step to characterize the L-band RFI environment. This choice was made because RFI contamination in North America appears to be localized at certain urban areas which is easy to analyze. Moreover, it is more low level than in other regions in the world and such low-level RFI is hard to detect by the algorithms described in Section 1.3 and of key importance in RFI environment characterization.

SMOS fully polarimetric Level 1C data [39] was used to generate “max-hold” and mean brightness temperature maps for all incidence angles using a 0.25°x0.25° grid over North America for November 2011. Figure 2.3 provides initial max-hold (i.e. the maximum brightness temperature at any incidence angle within a grid cell is recorded over the month)
and mean (the average of all measurements recorded in this time period) maps for horizontal polarization, where all of the original SMOS dataset is used without any modifications except the correction for Faraday and geometric rotations to obtain brightness temperatures in the Earth reference frame.

Figure 2.3: (Upper) Max-hold and (lower) mean SMOS brightness temperatures (in K) in horizontal polarization over North America (all incidence angles) for November 2011 (both ascending and descending).
The max-hold map clearly shows evidence of RFI corruption, including some apparent lines of high brightness temperature measurements in Northern US regions [40]. The SMOS Level 1C dataset has an RFI flag and the flag consists of 3 bits for each measurement [33]. These bits indicate whether the measurement is an output of a corrupted snapshot, corresponds to a grid point in the vicinity of an RFI source, or is affected by large RFI source contributions. Figure 2.4 displays the same data as Figure 2.3 with all RFI flagged data discarded.
Figure 2.4: (Upper) Max-hold and (lower) mean SMOS brightness temperatures (in K) in horizontal polarization discarding RFI flagged measurements over North America (all incidence angles) for November 2011 (both ascending and descending).
The difference between the two figures indicates that the RFI flags in the SMOS dataset detect some, but not all, RFI artifacts. Some high brightness temperature lines still remain as artifacts and RFI contamination is again apparent.

An analysis of the source of these linear shapes was performed using SMOS Level 1B data (full snapshots including aliased regions). The antenna geometry of SMOS results in a hexagonal total field of view at Level 1B [40], as shown in Figure 2.5, and the spacing between SMOS antennas creates aliasing in the regions bounded by the dashed lines. Figure 2.5 also illustrates how aliased sources outside the extended alias-free field of view (EAF-FOV) can nevertheless create sidelobe contributions within the EAF-FOV. This can be particularly problematic for sources on the Earth’s horizon (the arc of sources in the upper portion of the figure). Such sources are not within the EAF-FOV but are frequently present at Level 1B due to the observing geometry, prevalence of ground-based sources directed toward the horizon, and apparent over-the-horizon propagation effects (likely atmospheric ducting.)
Figure 2.5: Level-1B images illustrating sidelobes of RFI sources affecting SMOS measurements (this figure is only an example and shows RFI sources independent from the data set used in this study).

Figure 2.6 provides an additional example of how such sources can cause Level 1C artifacts. The left plots in Figure 2.6 are Level 1B snapshots in “X” (upper) and “Y” (lower) polarizations (priori to correction to the Earth referenced frame in this example), while the right plots are the corresponding Level 1C data. In this case, sources on the Earth horizon that are far northward of the Level 1C region are aliased into the lower left portion of the 1B snapshot and appear in the corresponding Level 1C region. Several detailed Level 1B / 1C matchup studies showed this behavior to be the source of the linear features in the max-hold map of Figure 2.3 and Figure 2.4.
2.2.1 Artifact Removal for SMOS Dataset

The aliasing effects previously discussed occur primarily on the edges of the snapshots due to horizon sources, which are in the forward path of the SMOS swath [33]. In this study, the area of interest is North America, so larger corruption occurs on ascending orbits due to the presence of strong sources along the “distant early warning line” in Alaska and Northern Canada; these sources readily produce artifacts at much lower latitudes. An initial artifact removal strategy simply eliminates ascending observations from the dataset. Note that although such an approach can be acceptable for RFI studies as performed here, it is likely not acceptable for science applications.
Severe artifact corruption can also occur in cases with excessively large sources within the remaining snapshots. To address these cases, a pulse detection algorithm was developed in which the boresight accuracy error level [40], a quantity proportional to the average brightness of the snapshot and available in Level 1C SMOS data, was used to flag bad snapshots. The algorithm has two steps to flag outliers. First, snapshot boresight accuracy error is averaged over 5 snapshots, and snapshots in this window with boresight errors more than 0.01K larger than the mean are flagged. A 2nd order polynomial fitting of the remaining boresight accuracies is then applied in a window of 30 snapshots, and snapshots that do not have accuracy level within 0.01K of the fitting polynomial are flagged. This approach is similar in concept to that used at Level 1A in [29], but differs in that it is applied at Level 1C. Parameters in this detection algorithm were determined empirically after several trials to minimize the artifacts in SMOS data.

Figure 2.7 depicts an example from Nov. 15, 2011, in which the boresight accuracy is plotted versus scaled snapshot number (time). The boresight accuracy is found to vary slowly in time (with local variations << than 0.1 K) as various geophysical scenes are encountered. Bad snapshots flagged by the first stage algorithm are marked in green, while those flagged in the second step are marked in red. Detected snapshots in both the first and the second steps are discarded from further processing. Parameters of the algorithm can be varied to be more aggressive in artifact removal versus more conservative in retaining true RFI sources. The results to follow are examples using a more aggressive approach to eliminate artifacts, but likely eliminate some true RFI sources as well.
During the study, it was also observed that outer parts of the SMOS swath are more susceptible to artifact pollution due to the proximity of aliases as shown earlier. It is noted that such aliases affect both the alias free (AF) and extended alias free fields of view, so that the AF vs. EAF-FOV flags in the dataset are not sufficient to remove potential artifacts. Thus, a final step reduced the SMOS swath by retaining only measurements whose accuracy error levels were within 10% of the boresight accuracy error level. Figure 2.8 shows an example snapshot and the portion retained by this procedure.
Figure 2.8: Swath reduction as the final step in artifact-removal algorithm.

Figure 2.9 provides horizontal polarization brightness temperature maps after artifact removal. Although a significant amount of data that may include real RFI effects was discarded by the artifact removal algorithm (some points have no remaining observations for the month for the conservative parameters and fine grid utilized), comparing Figure 2.9 and Figures 2.3 and 2.4, it is seen that the process reduces the presence of obvious artifacts significantly, so that the remaining sources appear more “point-like” in the max-hold image.
Figure 2.9: (Upper) Max-hold and (lower) mean SMOS brightness temperatures (in K) in horizontal polarization discarding RFI flagged measurements and after artifact removal over North America (all incidence angles) for November 2011 (both ascending and descending).
2.2.2 Maps of 3rd and 4th Stokes Parameters

Artifact free brightness temperature maps can then be utilized to locate “true” RFI sources for further analysis. It is known that polarimetric properties of SMOS measurements can also be advantageous for locating low level sources [31], [8]. In RFI free land observations, 3rd and 4th Stokes parameters of the thermal emission are expected to be small, so that high brightness temperatures in 3rd and 4th Stokes parameters can be good indicators of RFI effects. Detection performance however is limited by the relatively high noise level of SMOS polarimetric observations, as well as variations of this noise level across the swath.

A map of the polarimetric parameter

\[ C = \left( T_u^2 + T_d^2 \right)^{1/2} \]  

(2.1)

where \( T_u \) and \( T_d \) are the 3rd and 4th Stokes parameters is illustrated in Figure 2.10. These results still show artifacts and swath dependent effects that remain, but also show some moderate level point-like RFI sources. The detection of “true” low-level RFI sources remains challenging, but is being pursued using polarimetric quantities as well as spatial and multi-angular methods.
Figure 2.10: (Upper) Max-hold and (lower) mean values of C parameter after artifact removal over North America (all incidence angles) for November 2011 (both ascending and descending).
2.2.3 SMOS CCDFs

Complementary cumulative distribution functions (CCDF) of the measured brightness temperatures provide information on the fraction of SMOS observations exceeding a specified brightness temperature value, and can be utilized as tools to describe the brightness temperature distribution (and partially the RFI distribution) over the region of interest. Figures 2.11 through 2.14 provide CCDF plots for the original SMOS data in horizontal (Figure 2.11), vertical (Figure 2.12), absolute values of 3rd (Figure 2.13) and 4th (Figure 2.14) Stokes parameters, compiled from the Nov 2011 North American dataset. CCDF results after RFI flagged measurements are discarded and after artifacts are cleared are also illustrated. Although, during artifact clearance process, a significant portion of the data was discarded including possible true large RFI sources, the focus here is low level RFI contamination which is usually the case over North America.
Figure 2.11: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures in horizontal polarization.

Figure 2.12: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures in vertical polarization.
Figure 2.13: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures for the absolute value of 3rd Stokes Parameter.

Figure 2.14: CCDF plots for (blue) original, (green) RFI-free-according-to-SMOS flags, and (red) artifact-removed SMOS brightness temperatures for the absolute value of 4th Stokes Parameter.
The “long tails” of the original CCDF plots indicate significant RFI and artifact corruption in all polarizations by large RFI sources. For instance, 0.001% of the SMOS measurements in horizontal polarization are greater than 2000K which cannot be explained as geophysical emission.

CCDFs using data after RFI flagged measurements are discarded show that the large source corruption is apparently significantly reduced. However the maps from previous parts have shown that these results still contain some RFI and artifact contributions.

CCDFs after artifact clearance appear similar to the non-RFI-flagged case, except for the 3rd Stokes parameter. Apparently artifact corruption of this quantity remains significant so that the clearance approach has an increased influence. A detailed study of the influence of the artifact removal strategies showed the importance of the swath reduction step as compared to the existing SMOS RFI flagging. These results again confirm the increased susceptibility to RFI of non-central portions of the SMOS swath. Continued studies are attempting to utilize additional source detection methods to determine a CCDF of moderate-to-low level RFI in the SMOS dataset, although this is a challenging task due to the difficulty of detecting low level sources in SMOS measurements.

2.3 Comparative Analysis of RFI in SMOS and Aquarius Measurements

Properties of SMOS and Aquarius instruments and RFI detection algorithms utilized by them have been discussed in Sections 2.1 and 2.2. SMOS provides full
polarimetric multi-angular observations but lacks a powerful RFI detection capability whereas Aquarius implemented a pulse blanking RFI detection algorithm but observed Earth at fixed incidence angles. Therefore, a comparative analysis of SMOS and Aquarius data combines the capabilities of two instruments and leads to a better understanding of L-band RFI properties, especially at low RFI levels.

2.3.1 Relationships between RFI in SMOS and Aquarius

The properties of SMOS and Aquarius enable some expectations about RFI corruption to be determined. For an RFI source located in the radiometer footprint emitting a given effective isotropic radiated power (EIRP, Watts) in the direction of a satellite, the RFI observed by the satellite ($T_{RFI}$ in K) nominally follows the “density of interferers” equation [41]:

$$T_{RFI} \propto \frac{EIRP}{Footprint\ Area}$$  

Therefore SMOS and Aquarius observations of an identical source nominally should show amplitudes proportional to the footprint area. Because all Aquarius beams have larger footprint areas than the nominal 40kmx40km of SMOS, the Aquarius observed RFI typically should be smaller than that measured by SMOS. Table 2.2 also illustrates the footprint area ratio in each of the three Aquarius beams; note the “true” SMOS footprint size (which varies with observation angle) was used to compute these ratios. It is recognized that these scale factors should occur only on average, and that several factors (such as the particular source waveform, how it is encountered in a particular radiometer
measurement, any azimuthal rotation of the source, etc.) can cause an individual observation to deviate from these expectations. These scale factors indicate in general that “low level” RFI detected by Aquarius should be ~5-6 times larger in amplitude for SMOS and therefore potentially more detectable.

<table>
<thead>
<tr>
<th>Incidence Angle</th>
<th>Beam 1</th>
<th>Beam 2</th>
<th>Beam 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>28.7°</td>
<td>37.8°</td>
<td>45.6°</td>
</tr>
<tr>
<td>Footprint Size</td>
<td>76kmx94km</td>
<td>84kmx120km</td>
<td>96kmx156km</td>
</tr>
<tr>
<td>Received RFI Power Ratio (SMOS/Aquarius)</td>
<td>4.96</td>
<td>5.45</td>
<td>6.22</td>
</tr>
</tbody>
</table>

Table 2.2: Footprint size for the three Aquarius beams and received RFI power ratios between SMOS and Aquarius for each Aquarius incidence angle

In some situations, statistics of RFI occurrence are of interest; for example information on the percentage of the time that specific RFI amplitude is exceeded globally. Such information can be computed for both Aquarius and SMOS, although SMOS information is available only for very large source amplitudes. Comparisons of such information from Aquarius and SMOS must also recognize differences in the coverage and resolutions of these instruments. For example, the larger footprints of Aquarius cause a specific RFI source to remain in the footprint for a longer time interval as compared to SMOS. More Aquarius data (i.e. number of 1.44 second observations) will therefore be RFI contaminated for one overpass as compared to SMOS for a single RFI source. On the
other hand, in some cases Aquarius may count separate RFI sources as single sources due
to its poor spatial resolution whereas SMOS can successfully resolve them. Thus the
percentage of visible RFI in Aquarius data reduces compared to SMOS.

One approach for developing a “scale” factor for RFI occurrence between SMOS
and Aquarius is to compare the number of measurements for which a given Earth location
remains within the satellite footprint. This approach is considered in what follows, but
remains under examination due to the influence of the differences in coverage for each
satellite. Differences in the influence of sidelobes and artifacts also contribute to
uncertainties in approaches for scaling occurrence information, as well as the differing
procedures for retaining or discarding large brightness observations (i.e. those of 1000K or
greater) between the two missions.

A last important point influencing RFI contributions arises from differences in the
SMOS and Aquarius instrument passbands. A wider nature for the Aquarius passband
(deemed acceptable given the Aquarius focus on oceanic measurements) as compared to
SMOS has been reported. RFI contamination from sources in adjacent bands may therefore
be significantly larger for Aquarius than for SMOS [42].

To provide a more specific example of SMOS and Aquarius relationships, Figure
2.15 plots SMOS and Aquarius measurements (Beam 1, horizontal polarization) of an RFI
source in Cincinnati, OH, USA, both acquired on May 3, 2012. The ‘x’ and ‘o’ markers
are SMOS Level 1C and Aquarius Level 2 measurements, respectively. Because these data
were not acquired simultaneously, the time axis of the figure is shifted with respect to the
time of the maximum brightness temperature for this location during a single overpass (the
peak of the SMOS measurements was observed at 11:04 UTC whereas this time for Aquarius Beam 1 was 23:26 UTC. In Figure 2.15 they are both shown at the origin to compare RFI received by the two instruments consistently). Datapoints are indicated at the relative times from this maximum of the horizontal polarization measurements for each instrument. SMOS data further represents a time series “overpass” (as opposed to an angular series) because a time series of measurements having incidence and azimuth angles fixed at $28.7^0$ and $37.3^0$, respectively, is used. The RFI source is apparent in both datasets, and shows larger amplitude in SMOS as compared to Aquarius, as described previously. Figure 2.16 demonstrates the overpass geometries of both SMOS and Aquarius for this source.

Figure 2.15: Received RFI power for SMOS and Aquarius over Cincinnati
Figure 2.16: SMOS (left) and Aquarius Beam 1 (right) overpasses over the RFI source in Cincinnati on May 3, 2012. For each SMOS snapshot measurements were taken from a fixed footprint (shown in white circle) where incidence angle and azimuth angle were 28.7° and 37.3°. Red lines on the snapshots are constant incidence and azimuth angle lines. Aquarius on the other hand only provides incidence angle information and the boresight brightness temperature value for each footprint, thus all measurements for Beam 1 where the incidence angle is 28.7° were taken into consideration. The position of the RFI source is shown in both SMOS and Aquarius swaths.
A fit to this dataset was attempted based on the SMOS/Aquarius relationships described previously, assuming that the RFI source amplitude is identical in the two measurements. First, “RFI-free” brightnesses were estimated by fitting a $16^{th}$ order polynomial to measurements outside $+/-$ 10 seconds of the peak brightness location (shown as dashed lines in Figure 2.15). A “true” RFI amplitude was then estimated as the difference between the peak SMOS brightness and the “RFI-free” curve at the peak location. This “true” RFI amplitude for SMOS was then scaled into a “true” RFI amplitude for Aquarius using the amplitude ratio for Beam 1 from Table 2.2. The “widths” of the RFI response in time was then estimated by dividing the respective footprint along track sizes by the known respective satellite velocities (i.e. the nominal duration over which a point on Earth remains within the footprint.) A Gaussian function of time with the known widths in time and peak values were then added to the RFI-free values to create predicted SMOS and Aquarius RFI responses (solid curves in Figure 2.15.)

The results show a reasonable qualitative agreement with the overpass measurements, but large differences are also observed in some situations. In particular, the peak Aquarius RFI value is underestimated by the scaling procedure. This example highlights the fact that many factors can influence particular RFI measurements (as discussed previously) beyond the basic expectations presented in this section. However the example also shows that insights from these basic expectations (e.g. SMOS RFI levels for a particular source being typically larger than corresponding Aquarius levels) are applicable qualitatively if not precisely.
2.3.2 SMOS and Aquarius Spatial Statistics for Large RFI Sources

Although low to moderate amplitude RFI is the focus of this study, a brief examination of the properties of large RFI sources is also informative, particularly with regard to the influence of artifacts in studies of lower level contributions. For this purpose, “max-hold” maps for May 2012 (i.e. the largest brightness temperature encountered during May 2012 in a given grid point is plotted) compiled from SMOS Level 1C and Aquarius Level 2 data are illustrated in Figures 2.17 (SMOS) and 2.18 (Aquarius). Aquarius data from all three incidence angles is included in Figure 2.18 in order to achieve global coverage. For correspondence, the SMOS image is compiled using only incidence angles within $1^\circ$ of the three Aquarius incidence angles. Figure 2.17 also contains a map in which data flagged as containing RFI by the SMOS processor is discarded when computing the max-hold. Correspondingly, Figure 2.18 includes a map created using the RFI mitigated brightness temperature available from Aquarius. It is noted that in general the RFI obvious as “hot-spots” in any of these maps must have large amplitude (on the order of 100K or more) in order to be distinguishable.
Figure 2.17: SMOS H pol Max-hold brightness temperature map (in K) for original (top) and non-RFI flagged (bottom) data.
Figure 2.18: Aquarius H pol Max-hold brightness temperature map (in K) for original (top) and RFI filtered (bottom) data. All beams are combined.

SMOS and Aquarius max-hold maps without RFI exclusion or mitigation have qualitatively similar characteristics in terms of the location of strong sources, both those that are widespread in space (e.g. Europe and China), as well as those that are more localized in space (e.g. Africa, North America, South America.) Differences in spatial resolutions are also immediately apparent, with localized sources being resolved much more finely by SMOS than by Aquarius. Note that the Aquarius dataset includes
observations over oceans (whereas the SMOS dataset used in this study does not); the Aquarius dataset shows some evidence of at-sea RFI sources.

The influence of the finer SMOS spatial resolution is not apparent in the “widespread” RFI regions (particularly Europe), where Aquarius RFI appears to be much more spatially resolved than that of SMOS. Artifact effects are particularly pronounced for SMOS in these regions, and likely are the source of the widespread corruption observed [36]. In addition, it is observed from the Aquarius dataset used that Aquarius discards measurements of brightness temperatures greater than ~1000K, and this fact also influences these differences.

The RFI excluded or mitigated plots in Figures 2.17 and 2.18 both show a reduced, but still significant, level of RFI corruption. SMOS results in these figures in particular illustrate the limitations of the currently available L1C RFI flag for low-level RFI sources, which is primarily focused on removing the largest sources only. The Aquarius RFI mitigated data clearly removes many large sources, but small regions surrounding these sources remain even after application of the “pulse” detection strategy. Both these results show that further improvements in RFI detection and mitigation for both instruments are of interest.

2.3.3 Match-up Studies of Low Level RFI Sources

In order to compare “low-level” RFI contributions, Aquarius data having a detected RFI level between 1K and 10K was selected from the May 2012 data files. This RFI dataset was generated by subtracting the Aquarius RFI-mitigated measurements from
unmitigated Aquarius measurements (both are provided by Aquarius). Figure 2.19 illustrates the locations of these measurements. There are 210277 low-level RFI measurements, 55.67% and 44.33% of these points in horizontal and vertical polarizations, respectively. The widespread nature of the resulting points suggests that many result from sidelobe contributions of larger sources. However some situations show evidence of more localized points, particularly in South America. Subsequent analyses therefore focus on the 6656 Aquarius “low-level” measurements in South America, 66.84% of which were detected in horizontal polarization. The locations of these H-pol low-level RFI were examined, and measurements obtained within 0.5 degrees latitude and longitude of each other were considered a single “RFI location”. The resulting 138 distinct low level RFI locations are shown in Figure 2.20.

Figure 2.19: Locations of “Low Level” RFI observations in Aquarius.
For the spatial locations in South America containing H-Pol low-level Aquarius RFI, SMOS data within 0.5 degree latitude and longitude were selected during May 2012. Of these 73724 data points in 550 functions of angles associated with overpasses above RFI locations, 11% were flagged as containing RFI, suggesting that the SMOS dataset can include low to moderate level RFI corruption that is not flagged.

Figure 2.21 plots all 550 resulting H-pol May 2012 SMOS curves as a function of incidence angle for this dataset. Although the wide resulting geophysical scatter makes low-level RFI information difficult to observe, clear evidence of some RFI sources
localized in angle (and in some cases near the Aquarius incidence angles) shows that an RFI detection strategy based on angular information may be beneficial in some cases.

![Figure 2.21: Horizontally Polarized SMOS brightness temperature (in K) over “Low-Level” RFI locations vs Incidence Angle.](image)

Similar plots are shown in Figure 2.22 for the SMOS 3rd and 4th Stokes parameters. Large measurements are also observed in some cases, but examinations of specific points that included large polarimetric returns showed relatively smooth functions of angle even in such cases. SMOS 3rd and 4th Stokes parameter measurements in general show a trend of increasing amplitudes as the incidence angle increases and as the angle with respect to the SMOS boresight increases. Such effects make the detection of low to moderate amplitude sources very challenging in polarimetric measurements alone.
Figure 2.22: SMOS brightness temperatures (in K) over “Low-Level” RFI locations vs Incidence Angle for (left) 3rd Stokes Parameter and (right) 4th Stokes Parameter.

As a more specific example, a low-level source detected at lat:-34.4, lon:-59.79 (Figure 2.23) in Aquarius Beam 1 was selected from the RFI dataset defined above. Aquarius measurements on May 6, 2012 reported 1.76 K H-pol RFI amplitude. The Aquarius time history corresponding to this overpass is illustrated in Figure 2.24, with brightness temperatures in the left plot and the detected RFI level on the right. The circled points correspond to the source of interest; note the earlier larger RFI level reported does not correspond to this source. A SMOS overpass from May 1\textsuperscript{st}, 2012 is plotted in Figure 2.25 (using a “time series” procedure similar to that used in Figure 2.15), again with the point corresponding to the RFI source circled. This point was not flagged by the SMOS RFI flags. Although this point shows a somewhat larger brightness temperature than nearby points in time, the RFI amplitude observed is comparable to the geophysical and instrument variation, making this source difficult to detect in the time series SMOS data.
Figure 2.23: The Location of the Low Level RFI Source detected by Aquarius on May 6, 2012.

Figure 2.24: Detected Source in Aquarius Brightness Temperature Data (in K) on May 6, 2012. The Measurements that correspond to the RFI source are circled.
Figure 2.25: SMOS overpass (in K) of the RFI source from Figure 2.23 on May 1, 2012. The measurement that corresponds to the RFI source is encircled.

Figure 2.26 depicts SMOS measurements in horizontal polarization from May 1\(^{st}\), 2012 over the same source as a function of incidence angle. The relatively smooth behavior of these measurements matches reasonable geophysical expectations, and suggests a search for outliers versus angle to detect RFI. In this example, the apparent outlier at angle 27.5\(^{0}\) (roughly matching the angle of Aquarius beam 1) indicates that a search vs. angle can be successful at detecting low-level RFI. However considering the amplitudes of random fluctuations around the 2\(^{nd}\) order fit that describes the general pattern, this outlier would likely lie at the margin of detectability given the geophysical and instrument-induced variations in the SMOS data.

If a 2\(^{nd}\) order polynomial (as suggested in [30]) is fit to the measurements to represent the geophysical dependence, the outlier appears to be approximately 8K above the value of the polynomial at that angle. This value corresponds to 0.61 standard
deviations with respect to the 2\textsuperscript{nd} order polynomial fit to the SMOS data. Recalling the theory explained in part 2.3.1, it is known that the observed RFI power in SMOS should be 4.96 times higher than the RFI observed by Aquarius for Beam 1. In this case the observed 8K SMOS RFI level corresponds to 1.76K in Aquarius, a ratio of 4.54:1.

Figure 2.26: SMOS measurements (in K) over the region of the RFI source in Figure 2.23 vs incidence angle on May 1, 2012. The outlier is encircled.

Figure 2.27 similarly depicts the SMOS angular response for another RFI source detected on May 1, 2012 by this algorithm (again not flagged by the SMOS RFI flag). Outliers are circled with red at incidence angles of 45.97\(^0\) and 46.35\(^0\) (near the Aquarius beam 3 incidence angle). The latter outlier appears to be 47.4 K above the polynomial fit. This source is located at lat:-37.1, lon:-59.6, south of the source depicted in Figure 2.23. The maximum RFI observed by Aquarius Beam 3 for this region is 8.7 K on May 18, 2012,
a factor of 5.44 less than that observed by SMOS, as compared to the predicted ratio of 6.22 for Aquarius Beam 3 as shown in Table 2.2.

Figure 2.27: SMOS measurements (in K) over another RFI source at lat: -37.1, lon: -59.6 vs incidence angle on May 1, 2012. Outliers are again encircled.

2.3.4 Statistical Analysis of Low-Level RFI Match-ups

A larger set of “low level RFI” comparisons was created using similar 2\textsuperscript{nd} order polynomial fits applied to the 232 individual functions of angle shown in Figure 2.21 that have more than 10 points (i.e. angles) in the curve. Additionally, observations for which the standard deviation of the function of angle from the 2\textsuperscript{nd} order polynomial fit was greater than 20 K were also discarded. This step was used to eliminate measurements near oceans, rivers and mountainous regions where the angular fitting procedure is less reliable. Figure
2.28 indicates the locations of the 86 angular functions that have more than 10 measurements and fit standard deviation smaller than 20K. A threshold of 0.6 standard deviations of the each of the remaining 86 functions of incidence angle above the polynomial fit was used to flag RFI corrupted angles similar to previous 2 specific RFI sources. 44.67% of the measurements in these functions were flagged. On the other hand flagged measurements were counted as RFI comparable to Aquarius RFI only if the corresponding incidence angle was within 1.5 degrees of an Aquarius incidence angle (19.7% of the flagged measurements), where the maximum difference between the flagged measurements and the polynomial fit was noted as detected RFI in SMOS. The results showed that 94.19% of the curves had such RFI. This fraction is much higher than that detected by the existing SMOS flags (only approximately 11%) but the relatively low threshold utilized (to enhance sensitivity) would likely result in a high false alarm rate if used in practice. Figure 2.29 is a scatter plot of these low-level RFI sources for each Aquarius incidence angles after Aquarius RFI was scaled according to “density of interferers” equation as shown in Table 2.2. When the threshold was increased to 2.5 standard deviations, 11.63% of SMOS functions appeared to have RFI at Aquarius incidence angles which is approximately similar to existing SMOS flags performance over 138 RFI locations in South America. Figure 2.29 also depicts the RFI amplitude comparison for this threshold after scaling. As expected when the threshold increases, most of the low-level RFI detected in SMOS is not retained.
Figure 2.28: Locations of SMOS measurements as a function of incidence angle that include more than 10 samples (green and red) and the ones that have less than 20K standard deviation (green).

Scatter plots in Figure 2.29 indicate that some RFI sources were observed by SMOS and Aquarius as described by the simple qualitative “density of interferers” equation. On the other hand a significant portion of RFI sources violate this approach and appear larger in either SMOS or scaled Aquarius measurements. The differences in SMOS and Aquarius instruments such as their differing passband filters and the differences due to the RFI source itself such as the particular source waveform, how it is encountered in a particular
radiometer measurement, any azimuthal rotation of the source, etc. could be the source of these differences, as discussed in part 2.3.1. In addition, the Aquarius pulse detection algorithms are insensitive to continuous wave RFI sources whereas the SMOS angular domain algorithm can detect such sources. This situation would produce larger RFI detections in SMOS than the scaled RFI detected by Aquarius. On the other hand since Aquarius has a wider passband than SMOS has, the sources emitting in adjacent frequency bands may appear larger in scaled Aquarius RFI than in SMOS.
These plots illustrate RFI detected in SMOS measurements using the detection algorithm (with 2 different thresholds) vs. RFI measured by Aquarius scaled according to Table 2.2. Red lines indicate where these two match each other according to the theory suggested in part 2.3.1.

### 2.4 SMAP Validation Experiment 2012

The National Aeronautics and Space Agency’s (NASA) Soil Moisture Active Passive (SMAP) mission was launched in January 2015 to measure soil moisture on a
global scale via microwave remote sensing. SMAP utilizes both a scatterometer and a radiometer making it capable of providing soil moisture information with high resolution and high accuracy. The SMAP radiometer has 16 frequency channels and operates in the protected 1400-1427 MHz frequency band that is allocated for geophysical remote sensing applications in which any other electromagnetic emission is prohibited. However, as shown in the previous sections, measurements in this protected portion of L-band are contaminated by radio frequency interference. Thus, RFI detection and mitigation is also a crucial aspect of the SMAP mission to ensure reliable soil moisture measurements. SMAP uses a comprehensive RFI detection algorithm which is a combination of pulse detection, cross frequency, kurtosis and polarimetric techniques [43] which will be discussed in the next chapter.

In order to develop an optimal RFI detection algorithm for SMAP, the temporal, spectral, and statistical properties of RFI signals should be well known. Data from currently operational spaceborne microwave radiometers such as SMOS and Aquarius provide RFI information globally but their data lack the necessary temporal, spatial and spectral resolution to fully characterize the RFI environment as discussed previously. Thus, several airborne campaigns such as SMAPVEX08, SMAPE-1, SMAPE-2, SMAPE-3, and SMAPVEX12 have been used to collect additional RFI environment information [44]- [45] and SMAP Validation Experiment 2012 (SMAPVEX12) conducted in summer 2012 is one recent airborne campaign.


2.4.1 SMAPVEX12 Campaign Overview

The SMAPVEX12 campaign can be separated into two parts. The primary campaign took place near Winnipeg, Canada over six weeks between June 6 and July 19, 2012. The main goal of the primary campaign was to collect ground and airborne data simultaneously to test and validate the SMAP soil moisture retrieval algorithms. Details about the scientific background of this part of the experiment and participants are provided in [46]. In the second part of SMAPVEX12, the aircraft with the same configuration flew over the Denver, CO metropolitan area for two days after the primary campaign and collected L-band data particularly for RFI detection and mitigation studies.

The primary campaign was conducted over different types of croplands, forests, and pasture fields near Winnipeg, Canada. These regions were selected to test the ability of SMAP’s soil moisture retrieval algorithms to handle different vegetation and surface roughness types. Being relatively rural, data obtained from measurements over these regions were not expected to have serious RFI contamination. RFI sampling data was obtained on three days shown in Table 2.3 during the primary campaign; only these flights of the primary campaign are considered in this thesis. Figure 2.30 demonstrates the flight route for data collected on June 7, 2012. Flight routes may slightly change for other days but cover generally the same soil moisture focus sites.
Table 2.3: Flight dates and durations for RFI data sampling (Primary campaign)

<table>
<thead>
<tr>
<th>Flight Date</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 7, 2012</td>
<td>5.5 hours</td>
</tr>
<tr>
<td>July 17, 2012</td>
<td>4.4 hours</td>
</tr>
<tr>
<td>July 19, 2012</td>
<td>4 hours</td>
</tr>
</tbody>
</table>

Figure 2.30: Flight Route on June 7, 2012 during the Primary Campaign. The aircraft flew at two altitude levels: ~3500ft (red) and ~1000ft (blue).

As mentioned above, to study the RFI problem more comprehensively, two days of flight over the Denver, CO metropolitan area were added to the primary SMAPVEX12 campaign. This urban area was selected to provide a more diverse RFI environment. Dates
and durations of these flights and the flight routes are given in Table 2.4 and Figure 2.31, respectively.

<table>
<thead>
<tr>
<th>Flight Date</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 22, 2012</td>
<td>47 min</td>
</tr>
<tr>
<td>July 23, 2012</td>
<td>44 min</td>
</tr>
</tbody>
</table>

Table 2.4: Flight dates and durations for RFI data sampling (Denver flights)

Figure 2.31: Flight Routes on July 22, 2012 (green) and July 23, 2012 (black) over the Denver, CO metropolitan area.
2.4.2 PALS Radiometer and Digital Backend

The Passive/Active L-band sensor (PALS) [47] has been developed by Jet Propulsion Laboratory (JPL) and deployed in several airborne campaigns including SMAPVEX12. Although PALS has both a radiometer and radar, only radiometer data is considered in this thesis. The PALS radiometer has two channels for horizontal and vertical polarizations. Measured RF signals are filtered to occupy a 3-dB bandwidth of 1.4-1.42 GHz. The filtered RF signals are downconverted to 190-210 MHz 3-dB IF bandwidth for further signal processing.

One PALS measurement cycle lasts 50.4 ms and consists of twelve 4.2 ms observation states of which seven are antenna looks and the remaining five are calibration states. Each of these 4.2 ms states is further divided into twelve 350 μs intervals where integration is done over 300 μs for the radiometer and 50μs for the active part of the receiver (scatterometer). This measurement cycle, therefore results in a 50% duty cycle for radiometer antenna observations (7x12x300μs= 25.2ms). More information on PALS measurement timing can be found in [44].

The PALS antenna is a dual-polarized L-band patch array antenna with a two-sided 3-dB beamwidth of 20°. Throughout the campaign it was mounted on the aircraft in such a way that 40° incidence angle was obtained on the earth surface as shown in Figure 2.32.
During the campaign a direct sampling digital backend developed under a research and technology demonstrator program by JPL measured PALS IF signals (190-210MHz) in both polarization channels. The digital backend utilized the CASPER Berkley developed IBOB (Interconnect Break-Out Board) FPGA-base processing board to implement the high rate direct sampling detector [48]. This backend converts the analog pre-detected signals measured by PALS at a rate of 75 MSPS and 12 bit resolution continuously. The digitized output is stored in a digital backend PC via 2 10GigE interfaces. The 190-210 MHz IF
frequency band is in the 6th Nyquist zone of the ADC (187.5-225 MHz) and occupies the 15-35 MHz (spectrally reversed) frequency interval in this zone. Figure 2.33 shows the theoretical spectrum of the PALS IF signal and the spectrum of the signal stored by the digital backend. The resulting dataset has 13.3 ns time resolution, and raw samples are stored separately for horizontal and vertical polarizations in multiple files where each file corresponds to one second of measurement time. This unprocessed data with such fine time resolution enables the application of a variety of RFI detection algorithms to characterize the RFI environment. It is noted that previous studies have reported initial results from this campaign [49], but this study is the first to utilize calibrated data and to explore more detailed analysis of observed RFI signal properties.

Figure 2.33: Spectra of the PALS IF Signal (left) and the digital backend Input Signal (right) in theory. Since digital backend signal occupies the 6th Nyquist zone of the ADC spectrally reversed, 15MHz and 35MHz corresponds to 1.42GHz and 1.4GHz RF frequencies respectively.
2.4.3 Data Measurement and Calibration

Measured digital backend data was first processed to create power spectrograms and kurtosis spectra to analyze the received signals in time and frequency domains simultaneously. Power spectrograms were then calibrated to obtain brightness temperature information resolved in time and frequency. After calibration, RFI detection algorithms were applied to detect and characterize RFI.

In each one second data file, digital backend data was divided into 1024-sample “frames” (~13.65μs), and an FFT operation was performed for each frame. The resulting spectra occupy 1397.5-1435 MHz at a resolution of ~0.07 MHz which is smaller than the bandwidth of typical RFI signals. To obtain power spectrograms, the absolute square of spectra was taken after the FFT operation and the results were averaged over the corresponding 350μs PALS radiometer integration period. In addition to these power spectrograms, two kurtosis spectra were created for each one second data file. To compute the kurtosis spectra, after the FFT operation, the kurtosis of the real and imaginary parts of the FFT output was computed using the all antenna measurements in each frequency bin. Since each 1024-sample FFT corresponds to 13.65 μs and only seven of twelve PALS states are for antenna measurements, kurtosis computations were done using approximately \[ 1s/(13.65\mu s) \times [7/12] \approx 42750 \] samples. Figure 2.34 depicts raw power and kurtosis spectrograms for a single file as examples. High power stripes in the power spectrograms coincide with the calibration states of PALS, and the non-uniform ~20MHz passband of the radiometer can be seen. High kurtosis levels at 1.42GHz (which exclude the calibration state data) show a non-gaussian behavior of the signal at this frequency due to a weak
interference which cannot be observed in the uncalibrated power spectrogram. Since the PALS filters reject any signals outside the 1.4-1.42 GHz bandwidth as seen from the spectrograms, only this bandwidth was considered in what follows.

Figure 2.34: Uncalibrated Power Spectrogram (top) and Real and Imaginary Kurtosis Spectra (bottom) examples generated using single V-pol digital backend file corresponding July 22, 2012 16:40:16 UTC.
The order of the measurement states of PALS is given in [44] as:

1. Reference load
2. Reference + H-pol Noise Diode
3. Reference + HV-pol Noise Diode
4. Reference + V-pol Noise Diode
5. Common Noise Diode
6. 7 Antenna States

where each state lasts 4.2ms and consists of twelve 350μs integration time periods. Using this sequence combined with reference load and noise diode temperature information provided by JPL, calibration of power spectrograms was possible in each frequency channel. The measurement system provided only limited state information (marking only the antenna observations but not other states), thus the separation of measurements into antenna and calibration states was performed in post processing. First, the Reference + Noise Diode states were detected using a power threshold as they correspond to the highest power levels. Other states are determined according to the known order of the state sequence and the provided antenna state markers. Then calibration gain and offset values and the brightness temperature for each frequency bin in each 1 second data file were computed as:

\[ G = \frac{T_{\text{noise}}}{ndt - \text{ref}} \]
\[ O = T_{\text{ref}} - (G \times \text{ref}) \]
\[ T_b = G \times P + O \]  \hspace{1cm} (2.3)

65
where \( P \) is the raw power, \( T_{ref} \) and \( T_{noise} \) are the known reference load and noise diode temperatures, and \( ndt \) and \( ref \) represent the mean power values of the Reference load and Reference + Noise Diode states for the corresponding frequency bin. Figure 2.35 demonstrates the order of the antenna and calibration states for V-pol digital backend measurements (fullband power). Figure 2.36 then shows plots of raw V-pol digital backend power in the 1.4-1.42 GHz band measured over a single 350μs period at 15:54:54 UTC, computed gain and offset values for the corresponding 1-second file, and the resulting calibrated data (brightness temperature). The variation in brightness temperature is significant (~150K) and resembles the NEDT value for 350μs time and 0.07MHz frequency resolution \( \text{NEDT} = T_{sys}/B\tau \approx 150K \) for \( \tau=350\mu s \), \( B=0.07\text{MHz} \) and \( T_{sys}=750K \). Integration over the entire bandwidth would reduce this deviation to a level below 10K.
Figure 2.35: Antenna and Calibration States in V-pol PALS measurement cycle.
Figure 2.36: (Top left) V-pol raw power vs frequency over a single 350μs integration period and (top right) brightness temperature values after calibration. (Bottom left) Gain and (bottom right) offset values for the corresponding 1-second file.

Calibration was performed for each frequency channel in the power spectrograms, and brightness temperature values were obtained versus time and frequency. Calibration in separate frequency channels also removes the non-uniformness in the passband of the radiometer. For kurtosis spectra no calibration was necessary, and calibration states are simply discarded before kurtosis computations. Figure 2.37 shows the same spectrogram depicted in Figure 2.34 after calibration. Notice that only the 1.4-1.42GHz bandwidth is
shown and the spectrogram is averaged over 4.2ms (duration of one PALS state) to reduce noise.

Figure 2.37: Brightness Temperature Spectrogram obtained after the calibration of the power spectrogram in Figure 2.34.

Figure 2.38 demonstrates a flowchart that summarizes the digital backend data processing steps explained so far starting from the raw data to the RFI detection algorithms described in the next section.
2.4.4 RFI Detection

To detect RFI contamination in the stored data, thresholding algorithms were implemented on the brightness temperature spectrograms and kurtosis spectra. RFI detection over spectrograms is similar to the methods explained in [50] and [51].

First, each brightness temperature spectrogram was smoothed using a median filter. A median filter with a size of 8x8 time & frequency bins was used. Although this step decreases the sensitivity of the algorithm to pulsed and narrowband sources it was done to reduce noise in the spectrograms defined by NEDT value given in the previous section.
Next, RFI was detected using a thresholding algorithm. The threshold for brightness temperature spectrograms was computed as the mean value of the lowest 90% of bins in the spectrogram plus three times the standard deviation of the measurements in those bins. For kurtosis spectra, no smoothing was performed and a two sided threshold was set as three times the measured standard deviation of the closest 90% of bins to three as the RFI free data was expected to have its kurtosis value three [17]. Such thresholding resulted in a false alarm rate approximately 1.8%. Note that the use of the ~350μs integrated spectrograms in this process does not fully exploit the time resolution available from the digital backend measurements (13.3 ns) in resolving short pulsed interference. The use of the ~350μs integrations represents a compromise between time resolution and the need to reduce data volume in the digital backend processing. Future work may explore the detection of pulsed sources at finer time resolutions.

Bins in which the 350μs integrated brightness temperature values were above the threshold values were flagged as corrupted in each spectrogram. Then flags from the consecutive 1-second spectrograms were combined for longer term analysis. Finally, neighboring flagged bins were grouped using the MATLAB function ‘bwlable’ to be represented as single RFI signals occupying multiple time and frequency bins. This function labels each ‘connected component’ formed of non-zero entries with a distinct integer value. For example, if there are two flagged bins in the same frequency channel in [0μs, 350μs] and [350μs, 700μs] intervals, they are treated as a single RFI signature between 0μs and 700μs at that frequency and labeled with a unique integer (assuming no other neighboring bins are flagged). Note that although this grouping algorithm is an
efficient way to distinguish separate RFI signals and store their duration and bandwidth information, in some cases the algorithm may erroneously combine separate signals that are closely spaced or overlapping in time or frequency, resulting in errors in the output spectral and temporal properties of the RFI signal observed. However such errors are expected to be moderate and acceptable given the goal of obtaining general RFI environment properties in this work.

Kurtosis RFI detection was done at 1-second time resolution and was mainly used to extract information about the characteristics of the RFI signal detected by the spectrogram algorithm as this algorithm is very sensitive to pulsed interference which leads to kurtosis values greater than 3 whereas continuous RFI signals result in kurtosis values smaller than 3 [17] assuming single source contamination. Figure 2.39 shows a flowchart that summarizes the spectrogram RFI detection algorithm. Figure 2.40 and Figure 2.41 demonstrate the RFI detection algorithms over an example digital backend brightness temperature spectrogram and kurtosis spectra.

![RFI Detection Algorithm over Spectrograms](image)

Figure 2.39: RFI Detection Algorithm over Spectrograms.

Figure 2.40 depicts two major RFI contaminations at 1.4GHz and 1.413GHz. Other flagged bins are localized in time and frequency and are possibly due to false alarms.
Although at 350μs resolution these RFI signals appear to be continuous in time, their amplitudes change with time. Thus, kurtosis spectra in Figure 2.41 indicate a pulse type behavior (kurtosis is larger than 3) for the signal at 1.413GHz. On the other hand, no kurtosis detections occur for the RFI signal at 1.4GHz. It is reported that Kurtosis algorithm is less sensitive to weaker continuous wave signals and pulsed signals duty cycle of which is close to 50% [17]. These factors might be the reasons for such no kurtosis detection cases as we are unable to extract more information over the brightness temperature spectrograms without analyzing the original high resolution digital backend data. If the high resolution data were examined, the “blind spot” of the kurtosis algorithm for 50% duty cycle could be removed by using other normality tests such as Shapiro-Wilk, as indicated in [52].
Figure 2.40: A Tb Spectrogram (top left), the same spectrogram after smoothing (top right), RFI flags after thresholding (bottom left) and Grouped RFI bins as discrete signals with distinct colors (bottom right) where each color corresponds one RFI signal.

Figure 2.41: Kurtosis Spectra for the Power Spectrum shown in Figure 2.40. Red lines indicate the threshold values.
2.4.5 Observed RFI Properties

RFI signals with various properties were observed during the campaign especially during the flights over Denver, CO. Unlike the urban Denver region, data from the primary campaign showed a relatively RFI free environment as expected. Figures 2.42-2.46 show brightness temperature spectrograms, kurtosis spectra (with kurtosis thresholds in red) and raw antenna power plots (at 13.65μs resolution) of some RFI signals with different features in terms of their duration, bandwidth and amplitude observed on July 22 over Denver, CO. Note that brightness temperature spectrograms are shown after smoothing to demonstrate the RFI signature better.

Observed RFI signals during the airborne campaign can be classified in terms of their bandwidth as narrowband and wideband, and their duration as pulsed and continuous signals.

Figure 2.42 displays a narrowband continuous signal detected over Denver, CO on July 22, 2012 at 16:37:32 UTC. The power spectrogram demonstrates this signal at 1.404GHz. The signal cannot be detected in time domain as a continuous signal, however it appears in kurtosis plots as the kurtosis is much below the detection threshold at this frequency.
Figure 2.42: Narrowband Continuous RFI signal. Note that the continuous signal is not visible in time domain whereas the Kurtosis value less than 3 indicates the strong continuous RFI in Kurtosis Spectra.

Figure 2.43 shows another narrowband but short duration RFI signature observed on the same day at 16:38:32 UTC. This pulsed signal has a larger bandwidth and is easily detectable in time domain and kurtosis analyses as these RFI detection algorithms performs well for pulsed RFI.

During the campaign RFI with changing frequency was observed as well. Figure 2.44 depicts an example of a chirped RFI signal detected again on July 22 at 16:32:30 UTC over Denver, CO. The power spectrogram shows frequency change from 1.412GHz to 1.42GHz within half a second. The time domain plot indicates pulses with different amplitudes whereas Kurtosis plots show no interference signature. The reason might be duty cycle of the pulses when kurtosis was computed for 1 second in corresponding frequency. It is documented that kurtosis is not sensitive enough to detect pulses with high duty cycles [17].
higher end of the 3-dB RF passband.

Figure 2.44: Chirp RFI signal. Interference is invisible in time and Kurtosis domains. RFI can be said to be a chirped pulse signal with changing frequency.

In addition to narrowband signals, wideband RFI was also encountered during the SMAPVEX12 campaign. Figures 2.45 and 2.46 show some examples of such signals. Figure 2.45 depicts a severe pulsed type wideband RFI contamination observed on July 22 at 16:28:47 UTC. Both the power spectrogram and the time domain data clearly
demonstrate the signal. Kurtosis analysis also leads to kurtosis values higher than three, although the kurtosis threshold is biased due to the fact that RFI occupies the entire 1.4-1.42GHz band.

Figure 2.45: Wideband Pulsed RFI signal. Strong wideband pulses can be seen in the spectrogram and in time domain. Kurtosis values are also larger than 3 for all frequencies although the threshold is biased due to the severe corruption.

Figure 2.46: Wideband Continuous RFI signal. RFI can be seen in the upper half of the spectrogram. In Kurtosis domain the signal is not detectable whereas it appears as an increase in the background noise in time domain.
The RFI signal shown in Figure 2.46 is both wideband and continuous. The power spectrogram and the time domain plot depict a relatively weak increase in background power as the RFI signature. However the interference is not visible in kurtosis plots. Wideband continuous RFI signals are noise like signals and difficult to detect using time and frequency domain algorithms as well as statistical approaches like kurtosis unless the RFI amplitude is high enough to distinguish it from the geophysical emissions.

As mentioned in the previous section, it is also desirable to combine one second data files and analyze the RFI environment for longer time periods and larger regions. The Twin Otter aircraft flew with a speed of 50m/s and its antenna produced a footprint size which varies with flight altitude but was usually larger than 500m [46]. Therefore, combining one second data files and integrating them over 1s does not lead to a serious information loss since each footprint was observed typically for more than ten seconds. Such integration also provides more stable and long term RFI analysis. Kurtosis spectra computed for each 1-s file can also be combined and matched with integrated brightness temperature spectrograms. Figures 2.47 and 2.48 depict such long term brightness temperature spectrograms for the Denver flights. Figure 2.49 displays same information for one day of the primary campaign.
Figure 2.47: V-pol (left) and H-pol (right) Brightness Temperature vs Time and Frequency over Denver, CO on July 22, 2012.

Figure 2.48: V-pol (left) and H-pol (right) Brightness Temperature vs Time and Frequency over Denver, CO on July 23, 2012.
These long term spectrograms provide more information about the RFI environment. As shown in Figures 2.47, 2.48 and 2.49, RFI corruption with various bandwidths and durations was observed especially during the Denver flights. Data measured over the croplands near Winnipeg were not severely contaminated by RFI as expected. The majority of RFI signals seem to be narrowband or pulse type sources although full identification of RFI sources requires on site investigation. Leakage from adjacent frequency bands is also obvious as strong RFI signatures can be seen around 1.4GHz and 1.42GHz in both polarizations during the Denver flights. Furthermore, there are some continuous interference signals which are present in all measurements at certain frequencies (1.41GHz, 1.4156GHz, 1.4192GHz), so it was suspected that these signals were due to internal emissions of the radiometer, digital backend, or other onboard system.
electronics. Such signals were not included in the analysis of RFI statistics presented in the following section.

Figure 2.50 and 2.51 show the kurtosis spectrograms generated by combining kurtosis spectra for each 1-second files in the same time intervals shown in Figures 2.47 and 2.48.

Figure 2.50: V-pol (top) and H-pol (bottom) real (right) and imaginary (left) kurtosis vs time and frequency over Denver, CO on July 22, 2012.
Kurtosis spectrograms indicate more pulse type RFI signatures that are less detectable in the brightness temperature spectrograms especially in horizontal polarization. The reason for kurtosis analysis to be more sensitive than the brightness temperature spectrograms was discussed in Section 2.4.4. However they are not very efficient in continuous wave interference detection. Furthermore, it can be noticed that although the real and imaginary kurtosis spectrograms are similar, they are not exactly the same which
means after the FFT operation both real and imaginary parts of the spectra carry RFI information.

2.4.6 RFI Statistics

The spectrogram based RFI detection algorithm described in Section 2.4.4 flags time and frequency bins where RFI is detected, and groups the neighboring flags to attempt to identify individual RFI sources. Thus, it is possible to classify these RFI in terms of their duration and bandwidth by measuring the maximum extent of each signal in time and frequency. Such classification, though imprecise as explained previously, can be very useful since the performance of common time and frequency domain based RFI detection algorithms depend on these parameters. Figure 2.52 depicts the duration versus bandwidth plots for detected RFI during the entire SMAPVEX12 campaign.

Figure 2.52 shows that the majority of the observed RFI is either short-duration or narrowband type interference. Such interference is detectable with classical pulse detection, kurtosis and cross frequency detection algorithms [14]. On the other hand there are some cases where the observed interference is relatively wideband and near continuous. This type of RFI is the main concern for the design of RFI detection and mitigation algorithms since they are noise like signals, and may not be easily detected in time and frequency domains.
Figure 2.52: Bandwidth vs Duration plots for H-pol (top) and V-pol (bottom) RFI observed over Denver, CO. Red lines indicate the minimum detectable duration as RFI detection was performed over spectrograms where the time resolution was 350μs.

The amplitude of RFI signals is another important parameter to characterize the RFI environment. To compute the interference amplitude, in each 1-second data file the difference between the average of the spectrogram before and after RFI flagging was taken. For RFI signals observed in more than one 1-second spectrograms, the difference was computed for each of those separately and the maximum difference was considered as the amplitude of the signal. Such amplitude statistics were reported in [44] for the previous
SMAP validation experiment campaign called SMAPVEX08 conducted in 2008. Figure 2.53 plots the complementary cumulative distribution functions (CCDF) for the observed RFI (integrated over 2 seconds) during both campaigns. The results demonstrate that a wide range of RFI amplitudes was observed during both campaigns. Thus RFI detection and mitigation algorithms implemented in L-band microwave radiometers should be sensitive to low, medium and high levels of RFI. Moreover the results from two campaigns are not exactly the same but similar to each other although they were conducted at different times, over different locations and they utilized different types of RFI detection algorithms to identify RFI signals.

Figure 2.53: CCDF plot of detected RFI over Denver, CO during SMAPVEX12 campaign.
2.5 Summary and Discussion

In this chapter measurements from two space-borne (SMOS and Aquarius) and one air-borne (PALS at SMAPVEX12) L-band radiometer are examined to characterize the L-band RFI environment.

In section 2.2, it is shown that SMOS measurements are severely affected by RFI sources and their associated artifacts. These artifacts appear to arise at least in part from the presence of RFI sources in the aliased portion of the SMOS field of view, especially sources located on the Earth horizon and its aliases. An artifact clearance algorithm was proposed for reduction of these effects in SMOS observations of North America. It was found that the elimination of ascending orbits is effective in reducing artifacts in North America, and that existing SMOS RFI flags may not capture all RFI contributions, as evident in the third Stokes parameter results of Figure 2.13. The benefit of eliminating non-central portions of the SMOS swath is also illustrated for this quantity. An initial examination of polarimetric properties of the SMOS data following artifact removal is also provided. CCDFs for brightness temperatures are presented as statistical properties of SMOS measurements before and after artifact clearance. Results of the analyses has shown that once artifacts are cleared from SMOS dataset, moderate and large RFI sources can be distinguished via spatial analysis but low level RFI sources are still hard to detect.

Section 2.3 shows some of the similarities and differences of Aquarius and SMOS with respect to RFI corruption, and highlights the challenges for both instruments in continuing to improve algorithms for the detection and mitigation of low to moderate level RFI sources. Evidence is provided to show that expectations for relationships in RFI levels
derived from the “density of interferers” equation are reasonable, although many effects can lead to differences in RFI amplitudes observed by SMOS and Aquarius in a given measurement. Evidence of RFI remaining in both SMOS and Aquarius datasets after RFI mitigation (in the case of Aquarius) or exclusion (in the case of SMOS) is also provided. Low-level RFI matchup studies using the Aquarius pulse detection algorithm as the source of “true” low-level RFI information shows that multi-angular approaches with SMOS have some promise for detecting low to moderate level sources in SMOS data. However such methods will remain limited by geophysical and instrument-induced errors, so that low to moderate level RFI detection will remain very challenging for SMOS in particular.

In the final section of the chapter, an overview of the SMAPVEX12 airborne campaign and associated RFI observations has been presented. The direct sampling digital backend operated with JPL’s PALS radiometer provided raw data with very high time resolution. This dataset enabled characterization of the RFI environment both in rural and urban areas. The results show differing characteristics of RFI in these areas in terms of variety and occurrence rate. Denver flights (urban areas) in particular indicate RFI to be a more serious problem, with properties that vary significantly. Therefore, the need for comprehensive RFI detection and mitigation algorithms for microwave radiometers operating in this band remains apparent. Methods based on only one feature of the interference signal such as pulse-detection alone or cross-frequency detection alone may not be sufficient for elimination of the range of RFI encountered. The SMAP RFI detection algorithm, which will be discussed in Chapter 3, has been developed as a combination of different RFI detection techniques based on this fact [43].
Chapter 3: Soil Moisture Active Passive (SMAP) as a Revolutionary RFI Detection and Mitigation Approach

Chapter 2 describes the characteristics of the L-band RFI environment analyzing the space-borne and air-borne radiometer datasets. It is shown that the RFI is indeed a complicated problem for L-band radiometers as its temporal, spectral, spatial and statistical properties may vary significantly. Simple RFI detection and mitigation algorithms which assume certain RFI properties, therefore, are insufficient to mitigate RFI in passive remote sensing measurements, and a more comprehensive approach is needed. This chapter introduces NASA’s recently launched L-band Soil Moisture Active Passive (SMAP) radiometer and its RFI detection and mitigation algorithms as the first example of such efforts.

The chapter begins with an overview of SMAP mission in Section 3.1. Then, in sections 3.2 and 3.3, the SMAP radiometer and its RFI detection and mitigation approach are presented. Section 3.4 discusses some revisions made in the original SMAP radiometer timing and RFI detection and mitigation algorithms before its launch and their impacts on SMAP brightness temperature products. A description of the comprehensive RFI test conducted at NASA GSFC and theoretical and experimental performance analysis of SMAP RFI detection and mitigation algorithms based on this test are explained in Section 3.5. After its launch on January 31, 2015 and subsequent system controls in orbit, SMAP radiometer began providing data from space since March 2015. Section 3.6 demonstrates
some initial SMAP data from space and examines the performance of the RFI detection algorithms. Finally, possible improvements and alternative ways to the current SMAP RFI detection algorithms are discussed in Section 3.7.

3.1 SMAP Mission

NASA’s Soil Moisture Active Passive (SMAP) mission is an Earth observation satellite launched to measure the soil moisture present at the Earth’s land surface. Measurement of soil moisture will improve the estimation of water, energy and carbon cycles in our ecosystem. Soil moisture is also an important parameter for drought monitoring and flood assessments, therefore SMAP will provide significant social and economic benefits [53].

SMAP utilizes a radar and a radiometer operating in the protected portion of the L-band, sharing a rotating 6-meter reflector mesh antenna to provide high resolution and high accuracy soil moisture measurements which covers the globe every three days [53]. In this dissertation, the SMAP radiometer will be the main focus.
Figure 3.1: A picture of the SMAP instrument in orbit [53]

The spatial resolution of SMAP radiometer measurements is 40 km (3-dB antenna footprint). The radiometer measures the brightness temperature of Earth’s surface in all four Stokes parameters (Horizontal and vertical polarizations as well as 3rd and 4th Stokes parameters) from its 6 am – 6 pm sun synchronous orbit at a fixed incidence angle of 40°. While SMAP was being designed, considering already operational L-band space-borne missions like SMOS and Aquarius, it was expected that the radiometer would suffer from RFI from out-of-band and spurious emissions, thus a digital backend has been developed as a part of the SMAP radiometer to detect and mitigate any RFI in SMAP measurements.
A block diagram of the SMAP radiometer is demonstrated in Figure 3.3. A significant aspect of the SMAP radiometer is its utilization of a digital backend that enables high resolution data sampling, power, and statistics computations as well as frequency sub-channels.
The bandwidth of the SMAP radiometer IF signal is 24MHz centered at 120MHz. This signal is first downconverted to 24MHz, digitized and sampled at 96MSPS rate. Then, the signal is downconverted once more to baseband to produce a complex signal centered at 0MHz with 24MHz bandwidth. This baseband signal is called “fullband data.” Subsequently the fullband data is separated into 16x1.5MHz channels using a polyphase filter (Figure 3.4) and “subband data” is obtained. Statistics and cross-polarization correlations of both fullband and subband data are computed for kurtosis and polarimetric RFI detection. Later, the data is calibrated and integrated. The nominal SMAP integration time is 300μs for fullband data and 1.2ms for subband data (although in orbit these parameters vary with spacecraft altitude). The original SMAP timing sequence used 44 fullband and 11 sub-band integration periods to form one SMAP footprint as shown in Figure 3.5. Recent revisions in SMAP measurement timing will be discussed in section 3.4.
Figure 3.4: SMAP Polyphase Filter Bank Frequency Response

Figure 3.5: Illustration of the original radiometer timing sequence of one SMAP footprint. Each footprint contained 11 packets for data measurement and 1 packet of calibration. Each packet consists of 4 PRIs.
3.3 RFI Detection and Mitigation in SMAP

Utilizing its advanced digital backend, the SMAP radiometer implements time domain pulse detection, frequency domain cross frequency detection, polarimetric detection algorithms as well as kurtosis tests to its fullband and subband data to detect and mitigate RFI in its measurements. Each detector operates independently and produces RFI flags indicating the corrupted data. Then, independent flags from each detector are combined with a logical “OR” operator to maximize the probability of detection. The resulting RFI flags are called “Maximum Probability of Detection (MPD)” flags. Fullband and subband flags are combined in such a way that if a fullband flag is set high, then all 16 subband channels which include that time interval are also flagged as RFI contaminated. The individual SMAP RFI detection algorithms are summarized in sections 3.3.1-3.3.4. Figure 3.6 then shows an example SMAP footprint data and RFI detection and mitigation algorithms operating on it. The revisions made in SMAP RFI detection algorithms before the SMAP launch will be explained in Section 3.4.

3.3.1 Pulse Detection

SMAP’s pulse detection algorithm searches for measured horizontally and vertically polarized fullband powers above those produced by a normal statistical behavior [44] and flags them as RFI in each 44-sample SMAP footprint. Detection occurs when:

\[ T_A - m \geq \beta_{td}\sigma_{td} \]  \hspace{1cm} (3.1)
where \( m \) is the mean of the measurements except the highest 10% to avoid bias in case of RFI contamination, \( \sigma_{td} \) is the standard deviation of those measurements, and \( \beta_{td} \) is an adjustable parameter to set the false alarm rate. Nominally \( \beta_{td} \) is set to be 3 [43]. This algorithm is also referred to as asynchronous pulse blanking since no periodic properties of the RFI are assumed.

### 3.3.2 Cross Frequency Detection

SMAP’s cross frequency detection algorithm is similar to pulse detection but applied in frequency domain over 16 channels at 1.2ms and 15.4ms (i.e. the footprint) time resolution. The mean and the standard deviation of samples to calculate the detector thresholds are computed omitting the largest \( N \) channels to avoid RFI bias. Nominally \( N \) is set to be 2. Subband channels adjacent to those flagged as RFI corrupted are also flagged as corrupted and discarded by the SMAP cross frequency algorithms.

### 3.3.3 Kurtosis

SMAP uses the kurtosis test [43] in both fullband and subband channels to detect RFI in time and frequency domains. Detection occurs when:

\[
|K - K_{\text{nom}}| > \beta_K \sigma_K
\]

where \( K_{\text{nom}} \) and \( \sigma_K \) are predetermined nominal kurtosis and standard deviation values determined in pre-launch tests. \( \beta_K \) is an adjustable parameter to set the false alarm rate and is nominally set to be 3. Kurtosis of in-phase and quadrature radiometer channels are
computed separately and averaged before thresholding shown in (3.2) is applied in both subband and fullband SMAP footprints.

3.3.4 Polarimetry

As mentioned in Chapter 1, for natural emissions from Earth, the 3rd and 4th stokes parameters are assumed to be small [8], thus SMAP flags data as RFI in both fullband and subband channels when:

$$|T_{3,4} - T_{\text{nom}}| \geq \beta_{3,4} \sigma_{3,4}$$

(3.3)

Where $T_{\text{nom}}$ is a predermined nominal value for 3rd and 4th stokes parameter, $\sigma_{3,4}$ is the standard deviation of measured 3rd and 4th stokes parameters. Predetermined nominal values and $\beta_{3,4}$ are again adjustable parameters to set the false alarm rate. $\beta_{3,4}$ is nominally set to be 3. Note because SMAP polarimetric detections are applied prior to Faraday rotation corrections, the threshold for the 3rd Stokes parameter must be set to accommodate Faraday rotation induced contributions. Currently the polarimetric detectors are off in SMAP radiometer.
Figure 3.6: An example SMAP footprint with the original timing scheme. Subband Data (11x1.2ms) and Fullband Data (44x350µs) are shown with pulse blanking and fullband kurtosis thresholds indicated by dashed red lines. The final maximum probability of detection flags are also displayed where red bins indicate RFI corruption.

3.4 Revisions in the SMAP Radiometer Timing and RFI Detection Algorithms

After the initial design reported in [43] and summarized in the previous sections, SMAP radiometer timing and RFI detection algorithms have been updated for a better calibration of the measurements and to reduce biases in RFI free brightness temperature products. Two updates were made to the pulse and cross-frequency detection algorithms. In the modified algorithms, first, the standard deviation of the footprint (for pulse and 9.6ms cross-frequency detection) or one subband time interval (for 1.2ms cross-frequency
detection) are calculated using theoretically, and second, a “two sided” detection is used as opposed to the previous “one sided” approach as formulated in (3.1).

3.4.1 Revisions in Radiometer Timing

The radiometer-timing sequence was updated from 11x1.2msec (i.e. 11 subband and 44 fullband products) to 8x1.2msec (8 subband and 32 fullband products) of scene observation to allow increasing from 1 to 4 packets for calibration within one footprint as shown in Figure 3.7. The packets themselves are the same as the ones in the original timing and correspond to 1.4ms intervals and 1.2ms radiometer integration periods. Table 3.1 summarizes the number of measurements for each polarization in the subband and fullband channels of SMAP radiometer for one footprint in the original and updated timing sequences.

With the updated timing scheme, each footprint has more calibration data, therefore, a more stable calibration is possible. On the other hand, the reduced number of antenna measurements impacts RFI detection algorithms that attempt to estimate the scene temperature, due to the shorter integration time available.

The results to be shown in what follows all utilize the 8x1.2msec timing sequence.
Figure 3.7: Current radiometer timing after revisions showing one SMAP radiometer footprint with 8 packets for data measurement and 4 packets for calibration.

Table 3.1: Measurements Collected by SMAP Radiometer in the Original Time Scheme (Black) and after Revisions (Red in Parenthesis).
Figure 3.8 below demonstrates another SMAP footprint in fullband and subband domains with RFI flags produced by the MPD algorithm after the revisions in radiometer timing.

Figure 3.8: Another SMAP footprint with the updated radiometer timing scheme. One footprint is 32x350µs in fullband domain and 8x1.4ms in subband domain. Notice that RFI in the footprint is detected by the MPD algorithm as described in Section 3.3.

3.4.2 Theoretical Standard Deviation Calculations

Previously, for pulse detection over a footprint, the standard deviation, $\sigma_{td}(t)$, needed in (3.1) for setting the detector threshold was calculated empirically from the lowest 90% of samples acquired during a footprint. The finite number of samples used in this estimation resulted in a noisy estimate of the signal standard deviation, and therefore
causing a widely varying detection threshold from footprint to footprint. In contrast, a theoretical “formula based” approach computes $\sigma_{td}(t)$ using a model for the standard deviation:

$$\sigma(t) = \frac{T_{rec} + m(t)}{\sqrt{BW \cdot \tau}}$$

(3.4)

where $T_{rec}$ is the radiometer receiver temperature; $m(t)$ is the mean of measured samples in the time window excluding the largest and the smallest 5% (in total 10% is excluded, see the following discussion on double sided detectors vs single sided detectors); $BW$ is the radiometer bandwidth of 24 MHz; and $\tau$ is the integration time. Also, to avoid biasing $\sigma(t)$ should there be persistent RFI in the footprint, values of $m(t)$ exceeding 400K are replaced with 400K.

Similarly, the cross-frequency algorithm calculates the standard deviation, $\sigma$, used in setting its threshold from the data itself in the original design. Again the formula based approach uses a model for the standard deviation:

$$\sigma = \frac{T_{rec} + m(f)}{\frac{BW}{\sqrt{16 \cdot \tau}}}$$

(3.5)

where $T_{rec}$ is again the radiometer receiver temperature; $m(f)$ is the mean of measured samples in the time window excluding the channels with the largest and the smallest measurement (2 channels are excluded); $BW$ is the radiometer bandwidth of 24 MHz; and $\tau$ is the integration time. A maximum value of 320K for $m(f)$ is set for the cross-frequency detector. This value is chosen less than that of the pulse detector due to the ability of the cross-frequency detector to resolve interference in frequency.
SMAP subband data measured during the comprehensive RFI test (see Section 3.5) were used to estimate the standard deviations using the original (i.e. from the data itself) and the formula based approaches at 1.2ms and 9.6ms (i.e. 8x1.2ms) time resolutions. The original approach computes the standard deviations of the observed antenna temperatures in 14 frequency channels (excluding the lowest and the highest channels). Figure 3.9 compares the original (labeled “experimental”) and formula based (labeled “theory”) estimates. The fluctuations caused by the small number of samples used in the original method are apparent, and the widely varying threshold can cause excessive false alarms and/or missed detections in the associated detection algorithm. Use of the formula-based method corrects these issues and provides a more stable threshold. The formula based method is therefore applied for the pulse and cross frequency detectors in what follows.

Figure 3.9: Standard deviations of SMAP subband measurements computed with empirical and theory based approaches at (left) 1.2ms and (right) 9.6ms resolution.
3.4.3 **Double-Sided Detectors**

In theory, brightness temperatures in the subband and fullband channels of SMAP radiometer following integration can be assumed to follow Gaussian distributions. The original pulse and cross detection algorithms described in Section 3.3 detect and remove only samples that exceeded the “mean” by a specified number of standard deviations. The “mean” estimate was computed by excluding only the highest 10% of the data (or the highest 2 of 16 frequency channels.) Figure 3.10 illustrates this concept: a Gaussian histogram of simulated temperature samples with a mean of 290 K. The pulse and cross-frequency detection equations are applied to the distribution. Choosing a threshold value, $\beta$, of 3, the resultant threshold cutoff line is shown with a dashed red line. All fullband or subband products above 313 K are flagged and removed. Following this mitigation, the mean of the remaining samples is below 290 K (green dashed line). As a result, a single-sided detector causes the mean to bias low following mitigation when no RFI is present. This bias can be modeled and removed in subsequent processing, but can also be removed by replacing the “one sided” detectors used previously with “two sided” detectors.
Figure 3.10: An illustration of a Gaussian distribution of RFI free temperature samples with a mean of 290 K. The threshold cutoff line represents the mean (μ) of the lowest 90% of the data plus a number (β) of standard deviations (σ). The non-flagged temperature samples now have a mean that is biased low, represented by the green line.

A “two-sided” detector detects and removes samples both above and below a number of standard deviations from the “mean”. A “two-sided” process is also applied to estimate the “mean” by excluding the 5% highest and 5% lowest data when performing the average over the remainder. Due to its symmetric nature, the two-sided detector produces no bias for RFI free measurements. Note, however, that the two-sided detector will produce approximately twice as high a false alarm rate as compared to the one-sided detector if the same β value is used. Thus, a study was conducted using the RFI free SMAP data measured during the RFI test to examine one sided versus two sided pulse detectors. The mitigation-
induced bias and false alarm percentages (all samples flagged are false alarms since no RFI was injected in these tests) in each case are shown in Figure 3.11, for one-sided (left) and two-sided (right) detectors, and also using the “original” (labeled “NEDT est”) and “formula-based” (labeled “NEDT pred”) methods for determining the threshold. Results show the absolute value of the bias induced for the one-sided detector (upper left) is clearly higher than that of the two-sided detector (upper right), as expected. Results also show a significant increase in false alarm rate when switching to a two-sided detector (lower right) under the original threshold setting. However, using the formula based standard deviation calculation method described in Section 3.4.2 significantly decreases the false alarm rate to an acceptable level even with the two-sided detection approach.

Due to a desire to simplify the calibration biases and eliminate the need to correct any mitigation induced biases, the SMAP pulse and cross frequency detection algorithms were modified to be two-sided. Note the kurtosis and polarimetric detectors already were operated in a two-sided fashion.
Figure 3.11: Fullband pulse detector one-sided (left column) vs two-sided (right column) showing bias induced (top row) and percentage of false alarms (bottom row) for original (blue) and formula based (red) standard deviation computations.

3.5 Performance of the SMAP RFI Detection and Mitigation Algorithms

So far in this chapter, the SMAP mission, the SMAP radiometer and the RFI detection and mitigation algorithms implemented for the SMAP radiometer are presented.
This section examines the performance of these RFI detection algorithms, both experimentally and theoretically.

### 3.5.1 The Comprehensive RFI Test

To test the performance of SMAP RFI detection algorithms, a comprehensive test was conducted at NASA Goddard Space Flight Center between April 6, 2013 and April 20, 2013 with the SMAP radiometer in thermal vacuum conditions. The objectives were (1) to verify that SMAP can detect and mitigate pulsed and continuous wave RFI, (2) to determine the performance of RFI algorithms against low and high level RFI, and (3) to determine the performance against multiple RFI modulation schemes. During the test, known RFI signals were injected to the radiometer front end for a certain period of time (~30 seconds) followed by RFI free states lasting for the same amount of time. Therefore, RFI detection performance could be evaluated by observing the difference between SMAP products after RFI mitigation for RFI injected states and SMAP products for RFI free states with and without mitigation, as described in [43]. As RFI signals, continuous wave, pulsed, and communication signals with different frequencies, amplitudes, and modulation schemes were injected to the radiometer. This study focuses on SMAP performance against continuous wave and pulsed RFI signals. Table 3.2 demonstrates the RFI signals used in the RFI test with their properties (type, frequency, amplitude, pulse width etc.).
As seen in Table 3.2, besides pulsed and continuous wave RFI signals (first two rows), known radar signals (3<sup>rd</sup> column), different communication signals (4<sup>th</sup> column), multiple RFI sources (5<sup>th</sup> column) as well as RFI signals measured during the SMAPVEX12 airborne campaign reported in [34], [49], and [54] were injected to the SMAP radiometer. RFI detection performance against these types of signals will be examined in future studies.
3.5.2 Performance of the RFI Detection Algorithms: Theory vs Experiment

The performance of the updated RFI detection algorithms was evaluated using measurements done during the RFI test. In this study, the primary method used for evaluation is the Receiver Operating Characteristic (ROC) curve and the associated area under the curve (AUC) value calculated for each ROC curve. ROC curves indicate the performance of a detector by showing the probability of detection as a function of false alarm rate [55]. By varying a detector’s threshold over a large range, the associated probability of detection ($p_d$) and false alarm ($p_{fa}$) can be computed and plotted. Since infinite and zero thresholds mean no detection and one hundred percent detection respectively, all ROC curves start from ($p_{fa}, p_d$)=($0, 0$) and end at ($p_{fa}, p_d$)=($1, 1$). The AUC is calculated by taking the integral of the ROC curve over the false alarm rate [56]. A “random guessing” detector has equal probabilities of false alarm and detection, so that AUC=1/2. Since a good detector maximizes the probability of detection even for small false alarm rates, higher AUC values mean higher probabilities of detection over the same false alarm rates and thus better detectors. The ideal case is to obtain $p_d=1$ for all false alarm rates, so that AUC=1.

To plot ROC curves and compute AUC parameters against each RFI signal, updated RFI detection algorithms were applied to the measured data by changing the detector thresholds (changing $\beta$) in post-processing. A detection for a detector is defined as any pixel in the 8x16 bin subband or 1x32 bin fullband footprint being flagged by that detector. False alarm rates were computed by averaging the detection rates over 30 second RFI free
measurement periods and the corresponding probability of detection rates were calculated over the adjacent 30 second RFI-injected measurement periods.

ROC curves were generated for five different RFI baseline detectors including the cross-frequency for integration times 1.2ms (1 packet integration), 9.6ms (8 packet integration), pulse detection, kurtosis fullband, and kurtosis subband. The polarization detectors were excluded from this analysis due to the symmetrical injection during the test of RFI into both the vertical and horizontal ports (an unrealistic RFI scenario). Figure 3.12 illustrates measured ROC curves for continuous wave (CW) RFI signals at 1412.75 MHz with varying true power levels ranging from 0 Kelvin to ~4 Kelvin (note the “true RFI levels” may differ for the subband to fullband products due to the separate calibration of the two, but very close values were selected). As seen from the results, when the true power level starts increasing, the cross-frequency algorithms begin detecting the source at moderate power levels. This is expected due to the nature of the RFI injected, as cross frequency algorithms are more sensitive to single tone continuous wave interference than kurtosis and pulse detectors [14]. Cross frequency detection at 9.6ms time resolution is observed to be the most effective algorithm in these cases since averaging over time reduces the noise fluctuations before detection. These results show that continuous wave RFI should be detectable at a level < 3K by the SMAP radiometer, and even at < 2K if false alarm rates of ~20% are tolerated.
Figure 3.12: ROC performance results for the cross-frequency of 9.6 ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2 ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against CW RFI at increasing true power levels (in Kelvin).

Figures 3.13, 3.14 and 3.15 illustrate ROC curves for the same detectors for pulsed-wave injected RFI with pulse width values of 2\(\mu\)s, 51\(\mu\)s, and 95\(\mu\)s at 1412.75 MHz with similar power levels from \(~0.5\)K to \(~4\)K. Pulses of 2\(\mu\)s width were injected at 365Hz pulse repetition frequency (see Table 3.2) thus each footprint has approximately 4 pulses, and each pulse falls in a different time bin. These narrow pulses localized in short time and
narrow frequency intervals cause relatively big power spikes even for low power RFI cases, and are detected by almost all detectors even at power levels < 1 K. The kurtosis algorithms are more sensitive to such low duty cycle pulsed RFI, but all algorithms achieve good performance as the RFI power level increases.

Figure 3.13: ROC performance results for the cross-frequency of 9.6ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2 ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against Pulsed RFI with 2µs pulse width.
51µs and 95µs pulses were injected at 748 Hz and 1 kHz pulse repetition rates respectively during the RFI test. The higher PRF’s for these sources caused each footprint to contain 8 or 11 pulses, respectively. Therefore, the 32 fullband products of a footprint contain RFI in more than 25% of the samples. This significant corruption means that the mean and standard deviation calculations of the fullband detector are biased because only the largest and lowest 5% of the data are discarded in the threshold computation. As a result the performance of the fullband detectors diminishes compared with the 2µs pulsed RFI case. The subband kurtosis algorithm also shows reduced performance because the kurtosis algorithm is insensitive to RFI with higher duty cycles. Since the RFI is still localized in frequency, and is in effect becoming more continuous as the pulse width increases, the cross frequency algorithms still perform well.

Figures 3.12-3.15 also include results from the combined maximum probability of detection (MPD) approach of the SMAP radiometer. The MPD performance can differ from that of the best individual detector because it experiences a higher false alarm rate when it combines multiple detector outputs. This is a direct result of the MPD algorithm where flags from individual algorithms are combined with a logical OR operator. Other ways to decrease the false alarm rate while keeping the current interference detection performance of the SMAP radiometer are currently being investigated and will be discussed in Section 3.7.
Figure 3.14: ROC performance results for the cross-frequency of 9.6ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2 ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against Pulsed RFI with 51µs pulse width.
Figure 3.15: ROC performance results for the cross-frequency of 9.6ms integration time (i.e. 8 packet resolution), cross-frequency of 1.2ms integration time = (i.e. 1 packet resolution), pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against Pulsed RFI with 95μs pulse width.

ROC curve examples show RFI detection performance against specific RFI sources. The AUC can be computed for each of these ROC curves, and results then plotted as a function, for example, of the RFI source amplitude. Figure 3.16 demonstrates such measured AUC values versus true RFI amplitudes for the continuous wave, 2μs, 51μs and 95μs pulsed RFI cases, including sources at a variety of center frequencies throughout the
SMAP frequency band. As expected, higher true RFI levels lead to better detection performances (larger AUC values), and different detectors behave similarly to the results seen in the previous section. The results show the cross frequency algorithms always to perform well. This is due to the fact that all the RFI considered in these examples consists is a modulated single frequency sinusoidal signal. The results in general confirm the ability of the SMAP RFI detection and mitigation algorithms to detect RFI at levels of 1-2 Kelvin.

It is also possible to compute AUC values against these RFI signals theoretically under certain assumptions, as described in [57] and [58]. To compute AUC values theoretically, it was assumed:

1. that the 300µs integration intervals for fullband and 1.2ms integration intervals for subband detectors could not contain more than one RFI pulse, and that RFI pulses lied completely within one integration period.

2. that in each 1.2ms subband integration period, any RFI pulse, if present, existed exactly at the center frequency of one of the 16 frequency channels.

3. that the integration period was sufficient for the large sample limit of the Kurtosis distribution to be applied [17].

Under these assumptions, 300µs fullband and 1.2ms subband RFI free measurements were regarded as central chi-squared distributions whereas RFI contaminated ones were considered to be non-central chi-squared distributions, with the non-centrality parameter determined by the properties of the RFI signal as explained in [57]- [58]. The kurtosis mean and standard deviation for RFI free and sinusoidal RFI contaminated measurements was determined following [17]. Theoretical results for the
AUC for the corresponding sources are included in Figure 3.16 and compared with the measured. It is observed that theoretical AUC values are typically slightly higher than the empirical results. This difference can be explained by the violation of the above assumptions during the RFI test. RFI injection was not synchronized with SMAP radiometer timing, thus RFI pulses are frequently split into more than one integration period. Also, RFI frequencies do not align perfectly with SMAP subband channel frequencies. These two facts degrade the measured performance because RFI power is less detectable when split into multiple time/frequency bins instead of concentrated in a single bin. Moreover kurtosis detectors also show significant differences from the model since the number of measurement samples used to compute to kurtosis is not high enough to completely satisfy the central limit theorem conditions assumed in [17]. The 300µs and 1.2ms integration intervals include 28800 and 7200 samples for Nyquist sampling of a 24 MHz bandwidth, respectively. In general, however, the theoretical and empirical results match each other well given the inherent assumptions in the comparison.
Figure 3.16: AUC performance results for the cross-frequency of 9.6ms integration time, cross-frequency of 1.2ms integration time, pulsed detection, kurtosis fullband, and kurtosis subband detection algorithms against a sweep of CW, 2 us pulsed, 51us pulsed, and 95us pulsed interference.

3.6 SMAP in Orbit: Analysis of Initial SMAP Data

As mentioned before, SMAP was launched on January 31, 2015. After system check-ups, the instrument was turned on and SMAP started providing data from space. This section presents the initial SMAP radiometer data provided by NASA JPL. Although
the data has not been validated yet, it is still possible to examine the RFI environment observed by SMAP and performance of its RFI detection and mitigation approach using it.

SMAP data (with the version code (CRID (Composite Release ID)) 11580_001) for May 1, 2015 – May 8, 2015 time period has been analyzed and plotted on 0.25°x0.25° gridded maps. Although data in all four Stokes parameters have been examined, only the horizontal polarization data is presented in this work. Other polarizations have similar characteristics.

Figure 3.17 shows the “max-hold” horizontal polarization brightness temperature map in which the maximum brightness temperature from a week of SMAP data is displayed in each grid cell. The map shows significant RFI corruption over Europe and Asia, as well as some localized RFI sources over Africa and Americas. Thus, like SMOS and Aquarius, RFI is a serious threat for SMAP mission success and should be removed from its science products.
Figure 3.17: SMAP Max-hold H-pol Brightness Temperatures between May 1, 2015 and May 8, 2015.

Figure 3.18 shows the same brightness temperature map after RFI removal by the SMAP RFI detection and mitigation algorithms described in earlier sections. It can be seen that the majority of RFI corruption is cleared although there are some regions where RFI is still present at a lower level. Figure 3.19 is a map showing the max-hold RFI power in logarithmic scale generated by taking the difference between brightness temperature values before and after RFI mitigation.
Figure 3.18: SMAP Max-hold H-pol Brightness Temperatures between May 1, 2015 and May 8, 2015 after RFI mitigation.
As seen from Figure 3.19, RFI is present at all levels in SMAP brightness temperature data, however significant and widespread RFI which may exceed 100K is corrupting the SMAP measurements over Europe and Asia. Moreover, some of those high level RFI cannot be removed by the RFI detection algorithms. Therefore, a more detailed analysis of Europe and Asia is necessary. Figures 3.20, 3.21 and 3.22 depict the similar brightness temperature maps before and after RFI mitigation as well as the detected RFI map respectively for Europe.
Figure 3.20: SMAP Max-hold H-pol Brightness Temperatures over Europe between May 1, 2015 and May 8, 2015.
Figure 3.21: SMAP Max-hold H-pol Brightness Temperatures over Europe between May 1, 2015 and May 8, 2015 after RFI mitigation.
Although RFI is widespread and at significant levels, the comprehensive RFI detection algorithms implemented by the SMAP processor are able to remove most of it. However, especially over large urban areas like London, Madrid and Moscow, RFI is still present even after the RFI mitigation. Figure 3.23 shows maps of brightness temperatures measured by the SMAP radiometer when it flew over Madrid on April 22, 2015, both before and after RFI mitigation where the undetected RFI is obvious over Madrid, Spain. Time series fullband brightness temperature, kurtosis, and 3\textsuperscript{rd} and 4\textsuperscript{th} Stokes parameter values as well as the corresponding RFI flags near Madrid are also plotted in Figures 3.24, 3.25 and 3.26 respectively using a relative time axis where the time axis begins when the
boresight of the SMAP antenna is 0.5° degree away from the Madrid city center at 40.5° N, 3.5° W.

Figure 3.23: SMAP brightness temperature measurements on April 22, 2015 over Western Europe before and after RFI mitigation. Notice the unmitigated RFI over Madrid, Spain.

Time series brightness temperature plot demonstrates periodic high power levels which last for 6-7 footprints. This implies a continuous interference as the rotating SMAP antenna looking at the Madrid city center as the pulse detection algorithm is insensitive and doesn’t flag the measurements. Kurtosis plots also indicate kurtosis values much greater than 3 at some moments during these interference periods which implies pulsed RFI [17]. On the other hand, 3rd and 4th Stokes parameters are rarely sensitive to the RFI corruption as seen from Figure 3.26.
Figure 3.24: Time Series Fullband Brightness Temperatures over Madrid on April 22, 2015. Red Circles indicate flagged data by Pulse Detector.

Figure 3.25: Time Series Fullband Kurtosis values over Madrid on April 22, 2015. Red Circles indicate flagged data by Fullband Kurtosis Detector.
Figure 3.26: Time Series Fullband 3rd and 4th Stokes Parameters with Horizontal Polarization Brightness Temperatures over Madrid on April 22, 2015.

Figures 3.27, 3.28 and 3.29 demonstrate the subband brightness temperature, kurtosis and 3rd and 4th Stokes parameters for the same time period over Madrid. Long and periodic high power measurements, high kurtosis values from time to time and insensitive 3rd and 4th Stokes parameters can also be seen here. The important additional information that the subband data provides is the fact that the interference occupies the entire SMAP bandwidth.
Figure 3.27: Subband Brightness Temperatures over Madrid on April 22, 2015.

Figure 3.28: Subband Kurtosis values over Madrid on April 22, 2015.
Figure 3.29: Subband 3rd (left) and 4th (right) Stokes Parameters over Madrid on April 22, 2015.

Figure 3.30 below shows a sample footprint taken during one of the RFI contamination periods. Fullband and Subband brightness temperature and kurtosis values are plotted with RFI detector flags. The remaining subband data after RFI mitigation is also demonstrated. It can be seen that RFI which cannot be detected and mitigated by SMAP is continuous and wideband. Since it is continuous, pulse detection and kurtosis algorithms are insensitive [13] and since it is wideband, cross-frequency algorithms are unable to detect it properly [20]. The threshold of the cross-frequency detectors are biased due to the widespread corruption, and all of the footprint but 2 channels are flagged by double sided detectors.
Detected RFI is also analyzed in terms of their frequencies. Figure 3.31 depicts the rate at which measurements are flagged as RFI by the SMAP maximum probability of detection algorithm. It can be seen from the figure that the at each channel RFI detection rates are unique indicating RFI sources at different frequencies. However RFI in the 7th, 8th and 9th SMAP channels are more significant than other channels. These channels are reported to be at the edges of the SMAP 1400-1424MHz band due to the frequency mapping, therefore more susceptible to the emissions from the adjacent unprotected frequency bands.
Figure 3.31: SMAP RFI Detection Rates by Channel.
Figure 3.32 shows the total RFI detection rates combining all channels. The false alarm rate (FAR) for SMAP RFI detection algorithms can be estimated to be ~5% looking at the detection rates over oceans where no RFI is expected. Over the high RFI contamination regions in Europe and Asia, the detection rate is much higher. FAR may be increased after the polarimetric detectors are turned on, but since the thresholds are tunable as well, the science team may adjust it according to their science goals.

Figure 3.32: SMAP Overall RFI Detection Rates.
3.7 Possible Improvements to SMAP RFI Detection and Mitigation Algorithms

This section provides possible modifications to the baseline SMAP RFI detection and mitigation algorithms based on the results of the L-band RFI environment characterization studies.

3.7.1 Alternative MPD Approaches to Reduce Data Loss

In Chapter 2, it has been shown that narrowband pulsed and narrowband continuous sources are likely to represent the majority of RFI sources to be encountered by SMAP. The baseline SMAP maximum probability of detection algorithm may lead to unnecessary data loss under certain circumstances for such sources. In case of short-duration narrowband pulsed RFI or strong narrowband continuous wave RFI, the fullband pulse and kurtosis flags can be set high (Figures 3.33 and 3.34). As a result, all corresponding subband channels are also considered to be RFI contaminated and flagged although only a few channels are really corrupted.
Figure 3.33: Narrowband Pulsed RFI simulation. Only a single frequency channel includes RFI but other channels are also flagged by the baseline method.

Figure 3.34: Narrowband continuous wave RFI simulation. An entire footprint is lost by the baseline algorithm due to fullband Kurtosis flagging.
Besides this high data loss problem for narrowband pulsed and strong continuous wave RFI cases, the fine time resolution available in fullband products is also not used optimally by the baseline approach. Therefore, some alternative methods have been suggested below.

### 3.7.1.1 Use Only Subband Data and RFI flags

This method suggests using only subband RFI detection algorithms and computing the final brightness temperature product by averaging the subband data not flagged by the subband RFI detection algorithms. This method eliminates the high data loss in case of narrowband RFI contamination since the subband RFI detection algorithms only flags the corrupted channels. However, since they are insensitive to any wideband RFI, the SMAP products will be biased in case there is a wideband interference in the environment affecting several SMAP subband channels. Such an example is shown in Figure 3.35.
3.7.1.2 Use Only Fullband Data and RFI flags

Another approach can be using only fullband data and fullband RFI detection algorithms to compute the final SMAP brightness temperature product. The advantage of this approach is utilizing the finer time resolution (300µs) of the SMAP radiometer to efficiently detect pulsed RFI. However, this approach is not sensitive to weak continuous RFI contamination for which pulse detection and fullband kurtosis algorithms are not very efficient. An example of such a case is demonstrated in Figure 3.36.

Figure 3.35: Missed Wideband RFI pulse by the Suggested RFI detection scheme.
Figure 3.36: Missed Narrowband Continuous wave RFI by the Suggested RFI detection scheme.

### 3.7.1.3 A Mixed Approach

Considering the advantages and disadvantages of the previous two approaches, a mixed method has been suggested. This method is similar to the baseline maximum probability of detection algorithm however before flagging all corresponding subband channels in case of a fullband detection, it checks the subband RFI flags in that time interval. If there is a narrowband RFI detected by the subband channels only the RFI contaminated channels are flagged. Otherwise, again all subband channels are flagged as the situation implies a wideband RFI. This approaches both reduces the data loss in case of narrowband pulsed and strong narrowband continuous RFI cases shown in Figures 3.33 and 3.34, and keep the sensitivity of the algorithm to wideband RFI. Figure 3.37 shows an example case for this approach.
3.7.1.4 The Lowest Approach

This approach separates the fullband and subband data and RFI detection algorithms and computes two brightness temperature products for each. Since the RFI is additive, the method accepts the lowest of the two brightness temperature as the RFI-free final product. This approach is preferable as it retains the finer time resolution of the SMAP radiometer and is very efficient against pulsed or narrowband RFI. However in case of multiple RFI sources some of which are only detectable by only fullband detectors and the remaining is detectible by only subband detectors, both fullband and subband products contain RFI and even the lowest of the two remains RFI biased. Figure 3.38 illustrates an example of such a case.

Figure 3.37: Mixed RFI detection Approach. Notice the reduced data loss compared with the Baseline Maximum probability of detection Algorithm.
3.7.1.5 Baseline and Alternative MPD Approaches in the Comprehensive RFI Test

The alternative MPD approaches summarized in the previous sections have been tested as post-processing using the data measured during the comprehensive RFI test explained in Section 3.5.1. Their performances vs true RFI power levels are plotted in Figure 3.39 in terms of bias after RFI mitigation, percent of data discarded, and standard deviation of the mitigated products.

Bias after RFI mitigation is the difference between the RFI corrupted brightness temperature after RFI mitigation and the RFI free brightness temperature. For a perfect RFI detector, this bias should be zero since all RFI is mitigated. If there is no detection, the bias is equal to the true RFI level. Percentage of discarded samples and records entirely lost indicate the subband (or fullband in the “fullband only” approach) measurements and entire
footprints flagged as RFI corrupted respectively. Standard deviation of the mitigated brightness temperatures are the NEDT values for the remaining footprints after RFI mitigation. As the number of remaining footprint samples decreases this value increases.

The results shown in Figure 3.39 point out that the bias after RFI mitigation is similar for all approaches except the “fullband only” approach explained in Section 3.7.1.2. RFI greater than ~2K were detected by the baseline and alternative approaches, and the bias stays less than 1K although the true RFI level increases. The failure of the “fullband only” approach was due to the fact that the majority of the input RFI were continuous wave signals to which the fullband pulse blanking and kurtosis detectors of the SMAP radiometer are insensitive. On the other hand, the alternative MPD approaches significantly reduced the data loss and no footprint was entirely lost due to the RFI detection.
Figure 3.39: Performance of the alternative RFI detection algorithms versus the baseline maximum probability of detection algorithm. The mixed approach is labeled as “Limited Full”.

3.7.1.6 Baseline and Alternative MPD Approaches in the Initial SMAP Data

It is also very informative to analyze the real SMAP measurements from its orbit with the baseline and alternative MPD approaches. Figures 3.40-3.44 depict H-pol brightness temperature maps generated by a SMAP orbit from July 1, 2015 (with CRID no 11750_001) before and after RFI mitigation, as well as the percentage of footprints flagged as RFI by the corresponding MPD approach.
Figure 3.40: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “baseline” MPD approach.
Figure 3.41: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “only subband” MPD approach.
Figure 3.42: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “only fullband” MPD approach.
Figure 3.43: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “mixed” MPD approach.
Figure 3.44: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, (bottom) percentage of data flagged by the “lowest” MPD approach.

Figures 3.40-3.44 demonstrate that the “baseline”, “mixed”, and “only subband” approaches show similar characteristics in terms of mitigated brightness temperatures whereas “only fullband” and “lowest” approaches fail to mitigate some RFI compared with the previous ones. This implies the coexistence of continuous wave, pulsed, and
narrowband RFI in the region of interest. The fullband detectors are insensitive to continuous wave RFI, and the “lowest” approach is weak in case of multiple RFI sources that cannot be detected by only subband (wideband RFI) or only fullband (continuous wave RFI) detectors as explained in section 3.7.1.4. The RFI mitigated by “only fullband” on the other hand is expected to be pulsed RFI. Therefore, the characteristics of the L-band RFI environment, RFI with various properties, explained in the previous chapter is confirmed by the SMAP observations.

In terms of data loss, it can be seen that the alternative MPD approaches significantly reduces the percentage of data flagged as RFI. “Mixed” and “only subband” approaches seem to be more suitable for this SMAP orbit as they also provide a good RFI mitigation performance.

3.7.2 Updates in the Kurtosis Detector

SMAP Kurtosis Detector which uses both in-phase and quadrature data stream in the SMAP radiometer is explained in Section 3.3. In the baseline algorithm, kurtosis values from these two channels are averaged first and then the threshold is applied. In this section, alternative ways to utilize in-phase and quadrature kurtosis values for a better RFI detection performance is examined, both theoretically and numerically.
3.7.2.1 Theory: Kurtosis in SMAP Radiometer

In RFI free environments, detected voltages in the in-phase and quadrature channels of the radiometer can be modeled as two independent Gaussian distributions with zero mean and standard deviations ($\sigma$) proportional to the input power, i.e. the scene brightness temperature. These RFI signals can be defined by their probability density functions (pdf) as below:

$$p_{nI}(v) = \frac{1}{\sigma \sqrt{2\pi}} e^{\frac{-v^2}{2\sigma^2}}$$

$$p_{nQ}(v) = \frac{1}{\sigma \sqrt{2\pi}} e^{\frac{-v^2}{2\sigma^2}}$$

(3.6)

where $nI$ and $nQ$ stand for the thermal noise in in-phase and quadrature channels. Central moments of such distributions are well defined and the kurtosis of signals with this type of distributions is given as 3 independent of the standard deviation [17].

As discussed in [17] in real systems, the number of samples, $N$, and digitization operations may result in a kurtosis value slightly different than 3. However, it is also stated that for sufficiently large values of $N$ and enough signal power these effects can be ignored.

The SMAP fullband integration time is 300$\mu$s and data is sampled at 96MSPS rate, thus the number of samples to compute the kurtosis in in-phase and quadrature channels is:

$$N = \frac{300 \mu s}{1/96 \times 10^6} = 28800$$

(3.7)
With this $N$ and enough input power, the kurtosis can be taken as 3 for RFI free cases without any significant error. Thus, for the rest of the analysis we accept this as a fact.

Now consider a pulse sinusoidal RFI signal voltage defined as:

$$RFI = R \cos(wt + \theta)$$ \hspace{1cm} (3.8)

Where $R$ is the RFI amplitude and $\theta$ is the phase of the RFI signal.

$R$ and $\theta$ are also random variables and should be modeled properly for further analysis. In the literature it has been reported that RFI amplitudes can be modeled as a Rayleigh distribution with a certain scale parameter and $\theta$ is uniformly distributed in the $[0,2\pi]$ interval [59]. This model provides that in-phase and quadrature components of the pulse sinusoidal RFI signal amplitudes are statistically independent but identically distributed Gaussian random variables with zero mean and a certain standard deviation.

The RFI signal can be decomposed into its in-phase and quadrature components as follows:

$$RFI = R \cos(wt + \theta) = R \cos \theta \cos(wt) - R \sin \theta \sin(wt)$$

$$RFI_i = R \cos \theta \cos(wt)$$ \hspace{1cm} (3.9)

$$RFI_q = -R \sin \theta \sin(wt)$$

Where $R$ is modeled as a Rayleigh distribution defined with a probability density function (pdf) as:

$$p(r) = \frac{r}{\lambda^2} e^{-r^2/2\lambda^2}$$ \hspace{1cm} (3.10)

And $\theta$ is uniform in $[0,2\pi]$:

$$\theta \sim U(0,2\pi)$$ \hspace{1cm} (3.11)
Equations (3.10) and (3.11) provide that $Rsin(\theta)$ and $Rcos(\theta)$, amplitudes of in-phase and quadrature components of the RFI signal, are independent Gaussian random variables with zero mean and equal variance.

In RFI contaminated environments, therefore, the in-phase and quadrature channels of the radiometer will have Gaussian thermal noise plus a sinusoidal signal amplitude of which is a Gaussian random variable with zero mean and a certain standard deviation. Furthermore, signals in in-phase and quadrature channels are statistically independent since both noises and the RFI components in them are independent from each other.

Using the results in [17] the kurtosis of the signals in in-phase and quadrature channels can be computed. It is given that the expected value and the standard deviation of the kurtosis (denoted as $K$) of a Gaussian noise plus sinusoidal signal is:

$$E[K] = \frac{m_4}{m_2^2} = 3 \frac{1 + 2S + \frac{1}{2d} S^2}{(1 + S)^2}$$

$$\sigma(K) = \sqrt{\frac{1}{Nm_2^4} \left[ m_8 - m_4^2 + \frac{4m_4^3}{m_2^2} - \frac{4m_4 m_6}{m_2^2} \right]}$$

(3.12)

Where $S$, the signal-to-noise ratio and the central moments $m_i$’s are defined as:
In (3.13), $d$ is the duty cycle of the sinusoidal signal, $A$ is the amplitude of it and $\sigma$ is the standard deviation of the Gaussian noise. In our case $A$ is $R\cos(\theta)$ and $-R\sin(\theta)$ for in-phase and quadrature channels, $\sigma$ is proportional to the scene brightness temperature and $d$ is variable from 0% to 100%. Notice that for RFI free cases where $S=0$, $E[K]$ becomes 3 and $\sigma(K)$ is $24/N$. For 50% duty cycle, the kurtosis is blind to RFI as well giving $E[K]=3$ independent of the RFI power.

If the number of samples to compute kurtosis, $N$, is large enough, the distribution of the Kurtosis become Gaussian as a result of the central-limit theorem [17]. For SMAP, as discussed before, this condition is satisfied and kurtosis values in in-phase and quadrature channels are normally distributed random variables. Moreover, since the data sets in these channels which the kurtosis is computed are independent, the kurtosis values of in-phase and quadrature channels themselves are statistically independent as well.

As a result, in case of a pulse sinusoidal interference, the kurtosis of in-phase and quadrature channels can be modeled as statistically independent Gaussian distributions. Corresponding probability density functions (pdf) are:

$$S = \frac{dA^2}{2\sigma^2},$$
$$m_2 = \sigma^2(1 + S),$$
$$m_4 = 3\sigma^4(1 + 2S + \frac{1}{2d}S^2),$$
$$m_6 = 5\sigma^6(3 + 9S + \frac{9}{2d}S^2 + \frac{2}{(2d)^2}S^3),$$
$$m_8 = 35\sigma^8(3 + 12S + \frac{18}{2d}S^2 + \frac{8}{(2d)^2}S^3 + \frac{1}{(2d)^3}S^4).$$
where expected values and standard deviations are given in (3.12) and (3.13) and $K_I$ and $K_Q$ stand for kurtosis in in-phase and quadrature channels respectively. Recall that $A = R \cos(\theta)$ for in-phase and $A = -R \sin(\theta)$ for quadrature channels. The cumulative distribution functions for these distributions can be written as:

$$
\begin{align*}
    p_{k_i}(k) &= \frac{1}{\sigma(K_i) \sqrt{2\pi}} e^{-\frac{(k - E(K_i))^2}{2\sigma(K_i)^2}} \\
    p_{k_q}(k) &= \frac{1}{\sigma(K_q) \sqrt{2\pi}} e^{-\frac{(k - E(K_q))^2}{2\sigma(K_q)^2}}
\end{align*}
$$

Now let’s examine the fullband kurtosis RFI detection algorithm implemented in SMAP and proposed modifications to it.

### 3.7.2.2 Baseline SMAP Kurtosis Algorithm:

In the baseline algorithm, kurtosis values computed in in-phase and quadrature channels are averaged and then a threshold is applied to eliminate the outliers as RFI in a product of 32 kurtosis samples. Thresholding is done according to the formula below:

$$
\begin{align*}
    P_{k_i}(k < K) &= \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{K - E(K_i)}{\sqrt{2\sigma(K_i)^2}} \right) \right] \\
    P_{k_q}(k < K) &= \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{K - E(K_q)}{\sqrt{2\sigma(K_q)^2}} \right) \right]
\end{align*}
$$
\[ -\text{Threshold} < K - K_{\text{nominal}} < \text{Threshold} \]  

(3.16)

where \( K_{\text{nominal}} \) is considered to be 3.

It is known that the average of two independent Gaussian distributions is again normally distributed. Its mean is the average of means of those two normal distributions and its variance is the sum of their variances divided by 4. Thus, if we define a new Gaussian distribution for the average of in-phase and quadrature kurtosis values its mean and standard deviation becomes:

\[
\begin{align*}
\mu_{K_{\text{baseline}}} &= \frac{1}{2} \left( E[K_I] + E[K_Q] \right) \\
\sigma_{K_{\text{baseline}}} &= \frac{1}{2} \sqrt{\left( \sigma(K_I)^2 + \sigma(K_Q)^2 \right)}
\end{align*}
\]

(3.17)

False Alarm Rate (FAR) and Probability of Detection (PD) calculations for a set of threshold values are relatively easy for a normally distributed set of samples using its cumulative distribution function (cdf).

For FAR calculations we consider RFI free case where \( E[K_I] = E[K_Q] = 3 \) and \( \sigma(K_I) = \sigma(K_Q) = 24/N \). Thus;
For PD calculations we include RFI as well. So;

\[
\mu_{K_{baseline}} = \frac{1}{2} (3 + 3) = 3
\]

\[
\sigma_{K_{baseline}} = \frac{1}{2} \sqrt{\left(\frac{24}{N}\right)^2 + \left(\frac{24}{N}\right)^2} = \frac{12\sqrt{2}}{N}
\]

\[
FAR(\text{threshold}) = P(\text{threshold}_\text{lower} < k_{baseline} < \text{threshold}_\text{upper}) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{z}{\sqrt{2}} \right) - \text{erf} \left( -\frac{z}{\sqrt{2}} \right) \right]
\]

\[
\text{threshold}_\text{upper} = \mu_{K_{baseline}} + z \sigma_{K_{baseline}} = 3 + z \frac{12\sqrt{2}}{N}
\]

\[
\text{threshold}_\text{lower} = \mu_{K_{baseline}} - z \sigma_{K_{baseline}} = 3 - z \frac{12\sqrt{2}}{N}
\]

(3.18)

For PD calculations we include RFI as well. So;

\[
\mu_{K_{baseline}} = \frac{1}{2} \left( E[K_I] + E[K_Q] \right)
\]

\[
\sigma_{K_{baseline}} = \frac{1}{2} \sqrt{\left(\sigma(K_I)^2 + \sigma(K_Q)^2\right)}
\]

\[
PD(\text{threshold}) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{\text{threshold}_\text{upper} - \mu_{K_{baseline}}}{\sigma_{K_{baseline}} \sqrt{2}} \right) - \text{erf} \left( \frac{\text{threshold}_\text{lower} - \mu_{K_{baseline}}}{\sigma_{K_{baseline}} \sqrt{2}} \right) \right]
\]

\[
\text{threshold}_\text{upper} = 3 + z \frac{12\sqrt{2}}{N}
\]

\[
\text{threshold}_\text{lower} = 3 - z \frac{12\sqrt{2}}{N}
\]

(3.19)

Where \( E[K_I] \), \( E[K_Q] \), \( \sigma(K_I) \) and \( \sigma(K_Q) \) can be computed by using (3.12) and (3.13) for certain \( R \) distributions and duty cycles.
3.7.2.3 AND Algorithm

An alternative method suggests individual thresholding in in-phase and quadrature channels followed by an AND operation between them to generate final RFI flags. In this case we can compute the FAR and PD values for each channel as we did above and then use the simple probability rule for AND operations between statistically independent random variables which says:

\[ P(A \text{ AND } B) = P(A) \times P(B) \]  \hspace{1cm} (3.20)

Thus, for FAR calculations where there is no RFI:

\[
\begin{align*}
FAR_{\text{AND}}(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) &= FAR_i(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) \times FAR_q(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) \\
FAR_i(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) &= FAR_0(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) \nonumber
\end{align*}
\]

\[
\begin{align*}
\text{threshold}_{\text{upper}} &= \mu_k + z \sigma_k = 3 + \frac{24}{N} \\
\text{threshold}_{\text{lower}} &= \mu_k - z \sigma_k = 3 - \frac{24}{N} \\
\mu_k &= \mu_{k_i} = \mu_{k_q} \\
\sigma_k &= \sigma_{k_i} = \sigma_{k_q} \\
\end{align*}
\]

\hspace{1cm} (3.21)

For PD calculations in the presence of RFI:
\[ PD_{\text{AND}}(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) = PD_I(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) \times PD_Q(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) \]

\[ PD_I(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{\text{threshold}_{\text{upper}} - E[K_i]}{\sigma(K_i)\sqrt{2}} \right) - \text{erf} \left( \frac{\text{threshold}_{\text{lower}} - E[K_i]}{\sigma(K_i)\sqrt{2}} \right) \right] \]

\[ PD_Q(\text{threshold}_{\text{lower}}, \text{threshold}_{\text{upper}}) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{\text{threshold}_{\text{upper}} - E[K_\varphi]}{\sigma(K_\varphi)\sqrt{2}} \right) - \text{erf} \left( \frac{\text{threshold}_{\text{lower}} - E[K_\varphi]}{\sigma(K_\varphi)\sqrt{2}} \right) \right] \]

\[ \text{threshold}_{\text{upper}} = 3 + \frac{24}{N} \]

\[ \text{threshold}_{\text{lower}} = 3 - \frac{24}{N} \]

(3.22)

Where \( E[K_i], E[K_\varphi], \sigma(K_i) \) and \( \sigma(K_\varphi) \) can be computed by using (3.12) and (3.13) for certain R distribution and duty cycle.

### 3.7.2.4 OR Algorithm:

The OR Algorithm is very similar to the ‘AND Algorithm’ described above with the only difference that the results of in-phase and quadrature channels are combined with an OR operation rather than an AND operation. Thus, the FAR and PD calculations for individual channels are exactly the same with AND method but this time we use OR rules of probability which says:

\[ P(A \text{ OR } B) = P(A) + P(B) - P(A \text{ AND } B) \]

(3.23)

For FAR calculations assuming there is no RFI:
\[ FAR_{OR}(threshold_{lower}, threshold_{upper}) = FAR_i(threshold_{lower}, threshold_{upper}) + FAR_q(threshold_{lower}, threshold_{upper}) - FAR_{AND}(threshold_{lower}, threshold_{upper}) \]

\[ FAR_i(threshold_{lower}, threshold_{upper}) = FAR_q(threshold_{lower}, threshold_{upper}) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{z}{\sqrt{2}} \right) - \text{erf} \left( \frac{-z}{\sqrt{2}} \right) \right] \]

\[ \text{threshold}_{upper} = \mu_K + z\sigma_K = 3 + z \frac{24}{N} \]
\[ \text{threshold}_{lower} = \mu_K - z\sigma_K = 3 - z \frac{24}{N} \]
\[ \mu_K = \mu_{K_I} = \mu_{K_Q} \]
\[ \sigma_K = \sigma_{K_I} = \sigma_{K_Q} \]

(3.24)

where \( FAR_{AND} \) can be found via (3.21)

For PD calculations in the presence of RFI:

\[ PD_{OR}(threshold_{lower}, threshold_{upper}) = PD_i(threshold_{lower}, threshold_{upper}) + PD_q(threshold_{lower}, threshold_{upper}) - PD_{AND}(threshold_{lower}, threshold_{upper}) \]

\[ PD_i(threshold_{lower}, threshold_{upper}) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{\text{threshold}_{upper} - E[K_I]}{\sigma(K_I)\sqrt{2}} \right) - \text{erf} \left( \frac{\text{threshold}_{lower} - E[K_I]}{\sigma(K_I)\sqrt{2}} \right) \right] \]

\[ PD_q(threshold_{lower}, threshold_{upper}) = 1 - \frac{1}{2} \left[ \text{erf} \left( \frac{\text{threshold}_{upper} - E[K_Q]}{\sigma(K_Q)\sqrt{2}} \right) - \text{erf} \left( \frac{\text{threshold}_{lower} - E[K_Q]}{\sigma(K_Q)\sqrt{2}} \right) \right] \]

\[ \text{threshold}_{upper} = 3 + z \frac{24}{N} \]
\[ \text{threshold}_{lower} = 3 - z \frac{24}{N} \]

(3.25)

Where \( E[K_I], E[K_Q], \sigma(K_I) \) and \( \sigma(K_Q) \) can be computed by using (3.12) and (3.13) for certain \( R \) distributions and duty cycles. \( PD_{AND} \) can be computed via (3.22).
3.7.2.5 Analytical and Numerical Comparison of Baseline, AND, and OR Algorithms

We have derived FAR and PD formulas for the baseline, AND, and OR algorithms in case of a pulse sinusoidal interference. By computing them for a range of thresholds, we can generate receiver operating characteristics (ROC) curves for different RFI amplitudes determined by the scale parameter of $R$, $\lambda$, and different duty cycles. For each $\lambda$ value 1000 realizations of $\theta$, so 1000 in-phase and quadrature decompositions are averaged. First we kept the duty cycle constant at $d=100\%$ and increased the RFI amplitude by increasing $\lambda$. Then for constant $\lambda=15$, we examined the effect of change in the duty cycle.

Results shown in Figures 3.45 and 3.46 indicate that OR method gives the best performance although the performance of the baseline algorithm is very close to it. AND method on the other hand gives the worst performance, i.e. the worst probability of detection for the same false alarm rate. We can also see the blindness of kurtosis algorithm for 50% duty cycle in Figure 3.46. It is also observed that increasing RFI amplitudes and duty cycles result in better performance.

These results can be checked with Monte-Carlo simulations. The same plots for 1000 trials where in each trial, Gaussian signals were generated in in-phase and quadrature channels, a sinusoidal signal generated by using a certain $\lambda$ was decomposed into its in-phase and quadrature components determined by a random $\theta$, and they were added to the Gaussian signals. Then the kurtosis was computed in each channel and a threshold was applied according to described methods. The results of the Monte-Carlo simulations for the same $\lambda$ and duty cycle values are also plotted in Figures 3.45 and 3.46.
Monte-Carlo simulations are very consistent with the analytical analyses and give the same results which confirm our analytical derivations. As discussed before, it should be remembered that for finite values of $N$, the mean and the variance of Gaussian distributions are also random variables which causes an error between analytical and Monte-Carlo simulations. As $N$ increases this error diminishes.

Figure 3.45: Probability of Detection vs False Alarm Rate for a pulse sinusoidal RFI where $d=100\%$, $\lambda=15$ (left) and $\lambda=20$ (right).

Figure 3.46: Probability of Detection vs False Alarm Rate for a pulse sinusoidal RFI where $\lambda=15$, $d=10\%$ (left) and $d=50\%$ (right).
3.7.2.6 Baseline, AND, and OR Algorithms in the Comprehensive RFI Test

SMAP RFI detection algorithms were tested using the data obtained during the comprehensive RFI test with the “baseline”, “AND”, and “OR” kurtosis algorithms. Figure 3.47 demonstrates the bias after RFI mitigation and data loss rates similar to Figure 3.39 in section 3.7.1.5. It can be seen that the “AND” approach is not capable of mitigating some RFI at ~1K level whereas the bias after mitigation is very low for those RFI with “OR” and “baseline” kurtosis algorithms. On the other hand, data loss rate is lower for such low level RFI with the “AND” algorithms as expected.

Figure 3.47: Performance of the AND (green) and OR (red) Kurtosis detection algorithms versus the baseline (black) kurtosis algorithm.
3.7.2.7 Baseline, AND, and OR Algorithms in the Initial SMAP Data

The SMAP orbit analyzed for alternative MPD approaches in Section 3.7.1.6 was also analyzed for the “baseline”, “AND”, and “OR” kurtosis algorithms. In Figures 3.48-3.50, H-pol brightness temperature maps before and after RFI mitigation as well as RFI flagging rates for fullband and subband kurtosis algorithms were plotted similar to Figures 3.40-3.44 in section 3.7.1.6.

It can be observed that, in terms of RFI mitigation, all three algorithms provide similar results. This is expected as in the previous sections it was demonstrated that the “baseline”, “AND”, and “OR” algorithms perform similarly except very low-level RFI. Also the RFI mitigated brightness temperatures are computed by the overall SMAP MPD algorithm which means other detectors (pulse blanking and cross frequency) may overshadow the effect of applying different kurtosis algorithms. On the other hand, the plots for the detection rates for different kurtosis algorithms depict that the “AND” algorithm resulted in the least amount of data loss due to RFI flagging compared to the “baseline” and “OR” kurtosis algorithms.
Figure 3.48: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, percentage of data flagged by the (bottom-left) fullband and (bottom-right) subband “baseline” kurtosis algorithms.
Figure 3.49: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, percentage of data flagged by the (bottom-left) fullband and (bottom-right) subband “OR” kurtosis algorithms.
Figure 3.50: H-pol brightness temperatures in K (top-left) before and (top-right) after RFI mitigation, percentage of data flagged by the (bottom-left) fullband and (bottom-right) subband “AND” kurtosis algorithms.
3.8 Summary and Discussion

This chapter introduces the SMAP mission, radiometer, and SMAP RFI detection and mitigation algorithms. SMAP radiometer is unique due to its capable digital backend and comprehensive RFI detection and mitigation algorithms. Owing to its digital backend SMAP implements multiple RFI detection algorithms sensitive to different types of RFI simultaneously in time and frequency domain which minimizes the RFI corruption in SMAP science products.

Modifications to the SMAP radiometer timing sequence and RFI detection algorithms to reduce biases and false alarm rates in SMAP radiometer products are also presented in this chapter. The Comprehensive RFI test conducted at NASA GSFC is explained, and empirical and theoretical analysis of the performance of the updated SMAP RFI detection algorithms is done using the data measured during the test with ROC curves and AUC plots. It is shown that SMAP RFI detection algorithms perform as expected and they are capable of detecting RFI as low as ~2K. Cross frequency algorithms are found effective in all RFI cases examined (due to the modulated sinusoidal sources considered) whereas pulse detection and kurtosis algorithms loses their efficiency against continuous and high duty cycle pulsed RFI cases. Theoretical studies also confirmed the empirical results to the fidelity allowed by the theoretical modeling approach. The algorithms shown are current as of the January 31, 2015 SMAP launch and will continue to be assessed using SMAP post-launch data.

An analysis of initial post-launch SMAP data from space is given in Section 3.6. It is demonstrated that SMAP, like SMOS and Aquarius, captures significant amount of RFI
as it observes Earth. RFI detection and mitigation algorithms are shown to be very effective against most of the RFI, except the fact that some RFI cannot be removed over large urban areas in Europe and Asia. Further examination reveals that RFI is wideband and continuous in these regions and SMAP algorithms are not sensitive to this type of interference.

Finally, possible improvements to the baseline SMAP RFI detection and mitigation algorithm are discussed. Alternative ways to combine individual detectors are proposed to reduce the data loss due to the baseline maximum probability of detection algorithm and different approaches to effectively utilize the in-phase and quadrature kurtosis are discussed. SMAP baseline algorithms are shown to be improved by using these alternative approaches in certain conditions.
Chapter 4: Conclusions

This dissertation summarizes the properties of the RFI signals observed by L-band space-borne and air-borne radiometers operating in the protected spectrum (1400-1427 MHz) and introduces SMAP RFI detection and mitigation algorithm as a novel and aggressive multi-domain approach to minimize RFI contamination in radiometer measurements.

The key points and contributions of this research can be summarized as below:

- RFI may lead to bias in radiometric measurements, and thus, erroneous retrieval of the geophysical parameters even in the protected frequency bands solely allocated to the passive remote sensing applications.
- Many RFI detection and mitigation algorithms have been developed and implemented in past and currently operational radiometers. However these efforts were not sufficient to remove the significant portion of the RFI corruption.
- ESA’s SMOS and NASA’s Aquarius, two space-borne radiometers that measure soil moisture and ocean salinity on Earth surfaces and operating in the protected portion of the L-band, experience significant RFI degradation on global scale. Moreover, SMOS’s interferometric nature leads to aliasing problems which further complicate the problem. A comparative analysis of SMOS and Aquarius indicates
that the RFI environment seen by the two radiometers are similar, and RFI detection is a challenging problem, especially for low level RFI due to the lack of temporal, spectral and spatial resolution of the data provided by space-borne instruments. Thus, airborne campaigns are needed to collect high resolution data.

- SMAP Validation Experiment 2012 (SMAPVEX12) is such an air-borne campaign where high resolution data was collected by an L-band radiometer operated with a digital sampling backend. Analysis of the data indicates that duration, bandwidth and amplitude of the RFI signals may vary significantly in the protected portion of the L-band. Thus, single RFI detection algorithms based on certain assumptions about these properties are ineffective. More aggressive, multi-domain approaches are required.

- NASA’s Soil Moisture Active Passive (SMAP) radiometer implements such an aggressive multi-domain RFI detection algorithm utilizing its digital backend. SMAP applies RFI detection algorithms in time, frequency, polarimetric and statistical domains simultaneously and combines them in a way that maximizes the probability of detection.

- It was shown that using theoretical standard deviation calculations for RFI detection thresholding for SMAP radiometers resulted in more stable RFI thresholds and less false alarm rate. SMAP RFI detection algorithms have been updated to use such theoretical threshold computations.
✓ One-sided SMAP pulse blanking and cross-frequency algorithms have been updated with double-sided detectors to minimize the bias due to RFI detection and mitigation in RFI free brightness temperature products.

✓ Pre-launch analyses of the SMAP RFI detection algorithms show that SMAP is capable of detection different types of RFI signals greater than ~2K. This is a big improvement in radiometry for detecting low-level RFI owing to the digital backend of SMAP radiometer.

✓ It was shown that SMAP baseline Maximum Probability of Detection approach which combines individual RFI detection algorithms in different domains might cause high and unnecessary data loss in case of a narrowband and pulsed RFI corruption. Several alternative approaches were suggested to reduce this data loss. Pre-launch and post-launch studies have shown that each alternative has its own advantages and disadvantages and can be used according to the RFI properties of the observed scene.

✓ Kurtosis algorithm implemented in the SMAP RFI detection procedure computes kurtosis separately in in-phase and quadrature radiometer channels and apply the RFI threshold to the average of the two. Instead of this “average then flag” approach, “flag then OR” and “flag then AND” approaches in which kurtosis RFI detection algorithm was applied in in-phase and quadrature channels separately, then the resulting flags were combined with OR or AND operators respectively. The baseline, “OR”, and “AND” algorithms were defined theoretically, and compared analytically, numerically and using the pre-launch and post-launch RFI
studies. It was shown that the “OR” algorithm performed slightly better than the baseline algorithm whereas the “AND” algorithm diminished the detection performance.

SMAP is one of the first radiometers with a multi-domain comprehensive RFI detection and mitigation approach. The future research in this field will be focused on integrating more RFI detection techniques to such multi-domain approaches, and combining individual domains in a smarter way to minimize the data loss while keeping the probability of detection at the maximum level. Angular (for radiometers that provide multi-angular observations) and spatial domain algorithms described in Chapter 1 can be added to SMAP’s list of detection algorithms for a more complete approach. In addition to these well-known algorithms in microwave radiometry, RFI detection algorithms based on stationarity of the geophysical emissions have been developed for radio astronomy [60] [61] [62]. Such algorithms besides the image processing algorithms used to differentiate noise from the human made signals can also be used for microwave radiometry and are being investigated [63]–[64].
Bibliography


[40] A. Gutierrez, R. Castro, and P. Vieira, "SMOS L1 processor L1c data processing model," *DEIMOS Engenharia, Lisboa, Portugal*.


[48] [Online]. Available: https://casper.berkeley.edu/wiki/IBOB.


