

IMPROVED SUPERDARN RADAR SIGNAL
PROCESSING: A FIRST PRINCIPLES STATISTICAL
APPROACH FOR RELIABLE MEASUREMENT
UNCERTAINTIES AND ENHANCED DATA PRODUCTS

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy
in the Department of Physics and Engineering Physics
University of Saskatchewan
Saskatoon

By
Ashton Seth Reimer

©Ashton Seth Reimer, February 2018. All rights reserved.

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Physics and Engineering Physics
163 Physics Building, 116 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 5E2
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan
S7N 5C9
Canada

ABSTRACT

Ground-based radar systems are the best way to continuously monitor medium-to-large-scale features of the near-Earth space environment on a global scale. The Super Dual Auroral Radar Network (SuperDARN) radars are used to image the high-latitude ionospheric plasma circulation, which is produced by magnetosphere-ionosphere coupling processes generated by the interaction of both the solar and terrestrial magnetic fields. While investigating ways to expand the usable data products of SuperDARN to include electron density inferred using a multiple-frequency technique, it was determined that SuperDARN error estimates were lacking sufficient rigour.

The method to calculate SuperDARN parameters was developed approximately 25 years ago when available computing resources were significantly less powerful, which required a number of simplifications to ensure both valid data and reasonable processing time. This resulted in very conservative criteria being applied to ensure valid data, but at the expense of both rigorous error analysis and the elimination of some otherwise valid data. With access to modern computing resources, the SuperDARN data processing methodology can be modernized to provide proper error estimates for the SuperDARN parameters (power, drift velocity, width).

This research has resulted in 3 publications, which are presented here as Chapters 5, 6, and 7. The error analysis started with a first principles analysis of the self-clutter generated by the multiple-pulse technique that is used to probe the ionosphere (Chapter 5). Next, the statistical properties of voltage fluctuations as measured by SuperDARN were studied and the variance of these measurements were derived (Chapter 6). Finally, the statistical error analysis was propagated to the standard SuperDARN data products using a new First-Principles Fitting Methodology (Chapter 7). These results can be applied to all previously recorded SuperDARN data and have shown a practical increase in data of >50%. This has significant impact on the SuperDARN and space science communities with respect to, for example, global convection maps and their use in global modelling efforts. These results also enable quantitative experiment design facilitating research into using SuperDARN to provide electron density measurements, with a preliminary investigation using the new SuperDARN

fitting methodology presented in Chapter 8.

ACKNOWLEDGEMENTS

I am forever grateful to my supervisor Dr. Glenn Hussey for his support, guidance, and patience. He has always gone above and beyond to support me in all my research endeavours. I am also grateful to Dr. Kathryn McWilliams, who has provided a tremendous amount of support and guidance in uncountable ways. Both Dr. Hussey and Dr. McWilliams provided me with summer research opportunities that inspired me to pursue graduate studies.

I received support and advice from many faculty: Drs. George Sofko, Alexei Cheviakov, J.-P. St.-Maurice, Tom Steele, and fellow graduate students: Gareth Perry, Matthew Wessel, Devin Huyghebaert, Sarah Purdy, Seth Dueck, Fred Sage, Jason Ho, and Lindsay Goodwin, at the University of Saskatchewan.

Thank you to engineers Kevin Krieger, Marci Detwiller, and Keith Kotyk who spent a lot of time building and testing radar operation modes for me.

Thank you to Cindy Jelinski for routinely going above and beyond to support graduate students. She makes the Institute for Space and Atmospheric Studies feel like home.

Thank you to the Natural Sciences and Engineering Research Council (NSERC), the Canadian Space Agency (CSA), Canadian Foundation for Innovation, and Government of Saskatchewan for providing funding to do this research. Thank you also to the Department of Physics and Engineering Physics for providing supplementary funding via scholarships and thank you for the opportunity to teach.

My research relied heavily upon open-source software, in particular the Python programming language and packages Numpy, Matplotlib, Scipy, SymPy, Emcee, and DaViT-py.

DEDICATION

For my loving partner Eleanor. For my loving parents, Laurie and Syd. For my loving sisters, Miranda, Serena, and Chanelle.

CONTENTS

Permission to Use	i
Abstract	ii
Acknowledgements	iv
Dedication	v
Contents	vi
List of Tables	x
List of Figures	xi
List of Abbreviations	xviii
1 Introduction	1
1.1 Thesis Outline	3
2 Geospace Physics	6
2.1 The Sun	7
2.2 The Solar Wind	10
2.3 The Magnetosphere	13
2.3.1 Magnetic Reconnection	15
2.3.2 Magnetospheric Regions	17
2.4 The Ionosphere	19
2.4.1 Plasma Waves	22
2.4.2 Currents	23
2.5 Magnetosphere-Ionosphere Coupling	25
2.6 Ionospheric Irregularities	27
2.7 Summary	29
3 The Super Dual Auroral Radar Network	31
3.1 Introduction	31
3.2 SuperDARN Radar Principles	33
3.2.1 Radar Principles	37
3.2.2 Pulse Sequence	38
3.2.3 Doppler Shift From Quadrature	41
3.2.4 Phased Antenna Arrays	42
3.2.5 Interferometry	45
3.3 Radiophysics	46
3.3.1 Coherent Scatter	46

3.3.2	Refraction	48
3.4	Standard Data Products	51
3.4.1	A New Density Data Product	54
4	Radar Signal Processing: General Concepts and Techniques	57
4.1	Probability and Random Variables	58
4.1.1	Probability Density and Statistics	58
4.1.2	Gaussian Random Variables	60
4.1.3	Statistical Estimation	61
4.2	Radar Waveforms and Measurements	63
4.2.1	Signal, Noise, and Clutter	65
4.2.2	Matched Filtering	66
4.2.3	Waveform Design	67
4.3	Regression Techniques	72
4.3.1	Least-Squares Regression	73
4.3.2	Linear Least-Squares Regression	74
4.3.3	Non-Linear Least-Squares Regression	75
4.3.4	Confidence intervals	76
4.4	SuperDARN Signal Processing	77
4.4.1	Current SuperDARN Signal Processing	78
5	Self-Clutter Estimation	80
5.1	Background	80
5.1.1	Published Manuscript	82
5.2	Abstract	85
5.3	Introduction	85
5.3.1	Power and ACF Estimators	87
5.4	Self-Clutter Estimation	89
5.4.1	Self-Clutter Power in a Voltage Sample	93
5.4.2	Constructing the Self-Clutter Estimator	94
5.4.3	Generalized Self-Clutter Estimator	97
5.4.4	Maximal Self-Clutter Estimators	99
5.5	Testing and Discussion	101
5.5.1	Error Estimates	106
5.6	Conclusions	107
5.7	Acknowledgments	109
6	SuperDARN Auto-Correlation Function Statistics	110
6.1	Background	110
6.2	Published Manuscript	110
6.3	Abstract	113
6.4	Introduction	113
6.5	Mean ACF Estimate Statistics	117
6.5.1	Voltage Sample Statistics	118
6.5.2	Mean ACF Component Distributions	120

6.6	Analysis and Discussion	123
6.6.1	Comparison with Monte Carlo	123
6.6.2	Comparison with Experimental ACFs	128
6.6.3	Comparison to Previous Work	129
6.7	Conclusions	131
6.8	Acknowledgments	134
7	Self-Consistent Fitted Parameter Errors for SuperDARN	135
7.1	Background	135
7.2	Published Manuscript	136
7.3	Abstract	138
7.4	Introduction	138
7.4.1	SuperDARN Measurements	140
7.4.2	Noise and Clutter	141
7.4.3	Fitting for Ionospheric Parameters	143
7.5	Variance of ACF Lag Estimates	144
7.5.1	Ad Hoc Variance Estimates	145
7.5.2	Variance Estimates from First Principles	146
7.6	Estimating Self-Clutter	147
7.7	Current SuperDARN Fitting Software	150
7.7.1	Challenges of Fitting the Magnitude and Phase	151
7.8	A First-Principles Fitting Methodology	154
7.8.1	Fitted-Parameter Errors	157
7.9	Synthetic ACF Tests of FPFM-LMFIT2	158
7.10	Radar Data Analysis	162
7.10.1	Saskatoon SuperDARN Radar Analysis	169
7.10.2	FITACF Velocity	170
7.11	Summary	171
7.12	Acknowledgments	175
8	Conclusions and Future Applications	176
8.1	Summary of Contributions	176
8.1.1	Self-Clutter Estimation	177
8.1.2	SuperDARN Auto-Correlation Function Statistics	179
8.1.3	Self-Consistent Fitted Parameter Errors for SuperDARN	181
8.2	Progress on Ensuing Work	183
8.2.1	Optimal SuperDARN Pulse Sequence Investigations	184
8.2.2	Inferred Electron Density as a Standard SuperDARN Data Product	190
8.2.3	The Multiple-Incoherent-Pulse Technique	196
	References	200
A	Distributions Derived From Gaussian Random Variables	214
B	Self-Clutter Estimator Derivations	215
B.1	Expected Maximum Magnitude of Self-Clutter	215

B.2	VSE	215
B.3	MVSE and MPSE	217
C	Miscellaneous Derivations	220
C.1	Derivation of Probability Distribution Functions of the Magnitude and Phase of the Mean Auto-Correlation Function	220
C.2	Reduced Correlation in the Presence of Noise and Clutter	222
D	Reference Materials	223
D.1	SuperDARN IQDAT File Contents	223
D.1.1	Scalars	223
D.1.2	Arrays	224
D.2	SuperDARN RAWACF File Contents	225
D.2.1	Scalars	225
D.2.2	Arrays	226
D.3	SuperDARN FITACF File Contents	226
D.3.1	Scalars	226
D.3.2	Arrays	228
E	Copyright Information	230

LIST OF TABLES

2.1	Solar wind parameters typically observed at a distance of 1 AU from the Sun. Parameters values obtained from <i>Baumjohann and Treumann</i> (1997) and <i>Brekke</i> (2013).	11
2.2	A summary of typical plasma and magnetic field parameter values for regions of the magnetosphere shown in Figure 2.5 (<i>Kivelson and Russell</i> , 1995). The magnetic field strength in the plasmasphere can be approximated using the magnitude of Equation 2.9, and is therefore altitude dependent.	19
3.1	Northern Hemisphere SuperDARN radars as of September 2017. Schefferville is no longer operational.	34
3.2	Southern Hemisphere SuperDARN radars as of September 2017. Falkland Islands is no longer operational.	35
5.1	Extracted Signal ACF Compared with Actual Signal ACF: A quantitative comparison of the absolute value of the difference between the real and imaginary components for all lags of the extracted signal ACF and the simulated actual ACF depicted in Figure 5.5 (the top middle and top right panels, respectively). The differences are compared with the relative error defined by Equation 5.20.	103

LIST OF FIGURES

2.1	A diagram of the layers of Sun. The photosphere is between the chromosphere and convective zone. The outer layers of the solar atmosphere (chromosphere and corona) are also indicated. $R_S = 6.96 \times 10^8\text{m}$ is the radius of the Sun. See text for more details. Reproduced from Figure 11.2 of <i>Gombosi</i> (1998). .	8
2.2	A picture of a sunspot obtained with Hinode's Solar Optical Telescope on December 13, 2006. The dark appearance of the sunspot is due to the decreased luminosity of the sunspot since it is cooler than the surrounding plasma on the surface of the Sun. Courtesy NASA/JAXA (http://www.nasa.gov/mission_pages/solar-b/solar_022.html).	9
2.3	The magnetosphere of the Earth modelled using the Tsyganenko 1995 model (<i>Tsyganenko</i> , 1995). At distances less than $5 R_E$, the magnetic field is approximately dipolar and can be accurately modelled using Equation 2.9. Figure generated using DaViT-py (https://github.com/vtsuperdarn/davitpy). .	15
2.4	The outer magnetosphere of the Earth and magnetospheric regions seen from the noon-midnight meridian. The incident solar wind forms a bow shock as it interacts with the magnetosphere of a planet with a magnetic field. A balance between solar wind and magnetic field pressures occurs at the magnetopause. The Chapman-Ferraro current flows in this region of pressure balance (current coming out of the page on the sunward side of the magnetosphere). The magnetosheath is a region of shocked plasma between the bow shock and magnetopause. The magnetotail is caused by the solar wind stretching out the magnetosphere downstream. Magnetospheric currents are indicated by circles filled with dots or Xs and are discussed in more detail in the text. Figure 9.1 from <i>Kivelson and Russell</i> (1995).	17
2.5	The locations of magnetospheric plasma layers and magnetospheric currents of the Earth are shown as viewed from the noon-midnight meridian plane. The Chapman-Ferraro current is shown in Figure 2.4 and lies sunward of the LLBL in this figure. Layer characteristics are discussed in the text. Figure 10.4 from <i>Kivelson and Russell</i> (1995).	18
2.6	The altitude profile for a variety of atoms and molecules in both the neutral atmosphere and ionosphere of the Earth at ionospheric altitudes. Note that in the E region, NO^+ and O_2^+ molecular ions dominate and in the F region, O^+ atomic ions are the major ionic species (<i>Kivelson and Russell</i> , 1995). . .	20

2.7	Figure illustrating magnetic convection as seen from above the northern hemisphere. The oval dashed curves illustrate the flow pattern of magnetic field lines from the dayside (12) to nightside (24) and back around to the dayside that results when the IMF is dominated by $B_z < 0$ (more information found in Section 2.3). E_{DD} is the Dawn-Dusk electric field generated by magnetic convection from 12 (noon) to 24 (midnight) over the polar cap (small circle inside largest circle). The region 1 current (R1), the polarization electric field (E_H), the eastward-electrojet (J_H) and plasma flow velocity (v_p) in the F region of the ionosphere are also indicated. The bold arrowed arc following the convection flow pattern in the dusk sector (18) indicates the eastward electrojet.	26
2.8	An illustration of the required conditions for GDI growth ($\gamma > 0$). Adapted from Figure 2 in <i>Keskinen and Ossakow</i> (1983).	28
3.1	The FOVs of all active and inactive SuperDARN radars in the Northern (left) and Southern (right) hemispheres. The FOVs here were modelled starting at 180 km from the radar and using 75 45 km range gates.	32
3.2	The Saskatoon SuperDARN radar during the 20 December 2016 geomagnetic storm. All 4 antennas of the interferometer array can be seen in the foreground and 10 of the main array antennas can be seen in the background, along with the shed housing the transmitters, receiver, and control systems.	36
3.3	An example 3 pulse multiple-pulse sequence (top panel) and the 8 pulse katscan multiple-pulse sequence used by SuperDARN radars (bottom panel). In the bottom panel, the horizontal dashed line indicates the location of range gate 75 at a distance of 3500 km.	40
3.4	The half power beam width of a linear antenna array consisting of N equally spaced antennas.	43
3.5	The azimuthal gain pattern for a phased antenna array consisting of 16 antennas spaced 15 meters apart for two different SuperDARN beam directions. See text for details.	44
3.6	A simulation of the expected ray paths for the Saskatoon SuperDARN radar on beam 11 with a transmit frequency of 9 MHz. The grey lines indicate ray paths with white lines crossing the rays indicating the along ray distance in 200 km increments. The magenta lines indicate magnetic field lines. The thick black lines indicate regions where the rays are within one degree of aspect conditions required for coherent backscatter. The background colour indicates the electron density. The ray trace was performed using the Virginia Tech SuperDARN online Ray-Tracing Tool.	50
3.7	Data from the Canadian SuperDARN radars collected on 25 April 2014, both before (left) and during (right) an X1.3 solar flare from sunspot region 12035. The top of each plot shows the Canadian SuperDARN radars FOVs and backscatter, while the bottom of each plot indicates the X-ray flux incident on the Earth as measured by the GOES 15 satellite.	52

3.8	A SuperDARN convection map for the Northern hemisphere on 10 December 2011 from 8:38–8:40 UT. The contours indicate lines of equal electric potential. The vectors indicate the F region plasma velocity. This map was generated using Virginia Tech SuperDARN online Interactive Map Potential Plotting.	53
5.1	An estimated ACF derived from voltage samples measured by the Clyde River SuperDARN radar using the 8 pulse multiple-pulse sequence illustrated in the Farley Diagram in Figure 5.4. The blue curve and the green curve indicate the real and imaginary parts of the estimated ACF, respectively. The missing points are caused by missing lags in the multiple-pulse sequence.	90
5.2	A Farley diagram (c.f. Figure 1 in <i>Farley</i> (1972)) showing an example of a multiple-pulse sequence using 3 pulses with $\tau = 2.4$ ms, $t_p = 300$ μ s, $N_p = 3$, and PTAB = $[0\tau, 1\tau, 3\tau]$. The grey filled rectangles indicate transmission pulses and the rectangles with blue outlines indicate received voltage samples (from scattered signals returning to the radar). The red diamonds indicate signal backscattered from a repeated range of interest, in contrast to the black diamonds indicating signal backscattered from non-repeated unwanted ranges.	91
5.3	The delay ambiguity function for the pulse-sequence in Figure 5.2.	92
5.4	The 8-pulse multiple-pulse sequence <i>katscan</i> with $\tau = 1.8$ ms, $t_p = 300$ μ s, $N_p = 8$. The dashed line indicates the location of the range of interest. Near the x-axis, the black vertical lines at the base of positive sloped lines indicate transmission pulses and the blue vertical lines at the base of negative sloped lines indicate received samples. Intersections of the slanted black lines that do not lie on the dashed line indicate where signal is received from interfering ranges.	96
5.5	Plots comparing the use of the VSE on simulated radar measurements for estimating self-clutter originating from ground echoes (top panels) and for self-clutter originating from ionospheric echoes (bottom panels). For the top panels, all interfering ranges were filled with ground echoes (targets with a velocity of 0 m/s and a spectral width of 0 m/s) and for the bottom panels, all interfering ranges were filled with ionospheric echoes (targets with a velocity of 500 m/s and a spectral width of 100 m/s).	104
5.6	A comparison between measured SuperDARN radar ACF data both without (left column) and with (right column) the self-clutter estimate of Equation 5.12 removed. The ACF was recorded on 14 July, 2014 at 6:01:07 UT in gate 25 and beam 2. The top row of panels show ACF phase and the bottom row show ACF power. The triangular data points indicate data that is currently discarded by the self-clutter identification algorithm of <i>Ponomarenko and Waters</i> (2006) (the current technique used by SuperDARN to remove data contaminated by self-clutter).	105
5.7	Distributions of the difference in fitted velocities and fitted spectral widths between fitted SuperDARN data and VSE subtracted fitted SuperDARN data from the Saskatoon SuperDARN radar from 10 July, 2014 at 10:00 UT until 14 July, 2014 at 10:00 UT. The full-width half-maximum (FWHM) is included in the top right of each panel.	106

6.1	Plots of the real (top left), imaginary (top right), magnitude (bottom left), and phase (bottom right) of a mean ACF estimated from measurements by the Saskatoon SuperDARN radar, for beam 0 range gate 20, on 14 July, 2014 at 7:42 UT.	115
6.2	The distributions of the real (top) and imaginary (bottom) components of complex voltage samples measured using the Saskatoon SuperDARN radar (red dots), for beam 0 range gate 25, on 14 July, 2014 at 7:42 UT. The grey lines indicate distributions derived from Monte Carlo simulations of the data using Equation 6.7.	119
6.3	A comparison of the Monte Carlo generated marginal distributions for lag 5 of the mean ACF with the respective analytic expressions for the distributions of the components of the mean ACF given by Equations 6.13, C.3, and C.13. For a velocity of 300 m/s and a spectral width of 200 m/s.	124
6.4	A comparison of the variance of the analytic (red dashed curves) and Monte Carlo simulated (solid blue curves) real (left column) and imaginary (right column) component distributions. A power of 1830000 (arb), a velocity of 300 m/s, and spectral widths of 10 m/s (top), 200 m/s (middle), and 1000 m/s (bottom) were used.	126
6.5	A comparison of the variance of the analytic (red dashed curves) and Monte Carlo simulated (solid blue curves) magnitude (left column) and phase (right column) component marginal distributions. A power of 1830000 (arb), a velocity of 300 m/s, and spectral widths of 10 m/s (top), 200 m/s (middle), and 1000 m/s (bottom) were used.	127
6.6	A comparison of the ACF model given by Equation 6.13 with an experimentally measured SuperDARN mean ACF. See text for details.	129
6.7	A plot of the probability distribution function for the phase of the mean ACF given by <i>Woodman and Hagfors</i> (1969) (blue solid curve) and Equation C.3 (red dashed curve). The distributions were made using $\rho = 0.5$, $K = 30$, and $\phi = \pi/4$	131
6.8	A comparison of the variance of the phase distribution given by Equation 6.15 (blue solid curve) and that derived from 6.13 (red dashed curve). The plots were made using $K = 10$, and $\phi = \pi/4$	132
7.1	The real, imaginary, magnitude, and phase components of a typical ACF estimated from voltage data obtained with the Clyde River SuperDARN radar. The error bars were calculated using the methodology described in <i>Reimer et al.</i> (2016). The blue solid line and dots show the ACF data and the red dashed line indicates the curve of the best least-squares error-weighted fit. The red dashed curves are plotted using Equation 7.3 with best fit parameters of $P = 9194.75$ arb, $w_d = 95.8$ m/s, and $v_d = -413.9$ m/s.	142
7.2	An illustration of the self-clutter filtering criterion of <i>Ponomarenko and Waters</i> (2006). Black triangles indicate lags that satisfy Equation 7.10.	149

7.3	This figure illustrates the difficulties associated with unwrapping the phase of ACF data in FITACF (top panels) versus fitting to the real and imaginary components of the ACF data in FPFM-LMFIT2 (bottom panels), both with and without strong self-clutter in the ACF data. The black line in each panel indicates the expected one-to-one trend of fitted velocities versus simulated velocities. Each dot represents the mean fitted velocity obtained from 1,000 simulated ACFs. In green dots in the top right panel, show fitted velocities from a version of FITACF without the self-clutter criterion, which illustrates that the phase unwrapping is significantly more challenging without using the self-clutter criterion. In the top panels, the orange dots represent mean fitted velocities obtained using the standard unmodified version of FITACF. The bottom panels show that fitting the real and imaginary components of the ACF, as performed by FPFM-LMFIT2, results in correctly fitted velocities, even with strong self-clutter present.	153
7.4	An example curve of least-squares χ^2 as a function of velocity (v_d , a fitted parameter). Note that the curve reaches a minimum value at $\chi^2 = 2.7$ corresponding to $v_d = -413.9$ m/s (red dot). The CCBCL method (see text for details) is also illustrated for a $\Delta\chi^2 \approx 9$. Tracing the curve from χ^2_{min} to $\chi^2_{min} + \Delta\chi^2$ results in bounds for a 3σ confidence limit (black dots intersected by dashed lines). On the left side of the fitted value, the bound extends to $\Delta v_{left} = 63.4$ m/s ($v_{left} = -477.3$ m/s) and on the right side it extends to $\Delta v_{right} = 63.4$ m/s ($v_{right} = -350.5$ m/s), resulting in the best fit $v_d = -413.9$ m/s ± 63.4 m/s.	156
7.5	A comparison of the fitted velocity error (dashed lines) and true error (solid lines) for synthetic ACF data fitted by FPFM-LMFIT2. The synthetic data used to make these plots did not include self-clutter (the ideal dataset). The left panel shows fitted velocity error versus simulated velocity for synthetic ACF data with spectral widths of 50 m/s, 250 m/s, and 500 m/s corresponding to the blue, red, and green lines, respectively. The middle panel shows fitted velocity error versus simulated spectral width for synthetic ACF data with velocities of 50 m/s, 250 m/s, and 500 m/s corresponding to the blue, red, and green lines, respectively. The right panel shows fitted velocity error versus simulated signal to noise ratio for simulated ACF data with a velocity of 250 m/s and a spectral width of 250 m/s.	161
7.6	The same comparison as plotted in Figure 7.5, but using synthetic data with strong self-clutter (the pathological dataset).	163
7.7	A RTI plot comparing fitted velocity data produced by FITACF (top panel) and by FPFM-LMFIT2 (bottom panel) for 8 hours using ACF data from the Saskatoon SuperDARN radar data. The velocity data are plotted with the range away from the radar is plotted along the y-axis and the UT time is plotted in x-axis. The colour represents the magnitude of the fitted velocity. Ground scatter has not been highlighted or removed and is just included in the plot.	166

7.8	A RTI plot comparing fitted SNR data produced by FITACF (top panel) and by FPFM-LMFIT2 (bottom panel) for 8 hours using ACF data from the Saskatoon SuperDARN radar data. The SNR data are plotted with the range away from the radar is plotted along the y-axis and the UT time is plotted in x-axis. The colour represents the magnitude of the fitted SNR. Ground scatter has not been highlighted or removed and is just included in the plot.	167
7.9	A RTI plot comparing fitted velocity errors produced by FITACF (top panel) and by FPFM-LMFIT2 (bottom panel) for 8 hours using ACF data from the Saskatoon SuperDARN radar data. The velocity errors are plotted with the range away from the radar is plotted along the y-axis and the UT time is plotted in x-axis. The colour represents the magnitude of the fitted velocity. Ground scatter has not been highlighted or removed and is just included in the plot.	168
7.10	Distributions of the additional data produced by FPFM-LMFIT2 compared to FITACF for a statistical comparison of 8 years of Saskatoon SuperDARN data. The additional data is plotted as a function of velocity error and binned in to three SNR ranges. Each line plotted corresponds to a different range of SNR values. The blue, green, and red lines corresponds to a SNR values below 3 dB, from 3 to 20 dB, and above 20 dB.	170
7.11	A RTI plot of the fitted velocity obtained using FITACF (top panel) and FPFM-LMFIT2 (bottom panel) for 20 minutes of ACF data from the Rankin Inlet SuperDARN radar, on February 22, 2014.	172
8.1	A contour plot of Equation 8.1. Choosing combinations of τ and transmission frequency (f_{Tx}) to the left side of a contour of MUD guarantee a $MUD \geq$ the MUD of the contour. The hatched area shows the region in which SuperDARN radars currently operate: $1.8 \text{ ms} \leq \tau \leq 2.4 \text{ ms}$ and $8 \text{ MHz} \leq f_{Tx} \leq 20 \text{ MHz}$	186
8.2	The normalized velocity distribution for SuperDARN radars in the polar cap. The velocity bins are 25 m/s in size. N is the total number of points in the plot. $N_{5\sigma_L}$ and $N_{5\sigma_G}$ indicate the amount of data points within 5 standard deviations of the mean for the Laplace fit and the Gaussian fit, respectively.	187
8.3	The normalized velocity distribution for SuperDARN radars in the high-latitudes. The velocity bins are 25 m/s in size. N is the total number of points in the plot. $N_{5\sigma_L}$ and $N_{5\sigma_G}$ indicate the amount of data points within 5 standard deviations of the mean for the Laplace fit and the Gaussian fit, respectively.	188
8.4	The normalized velocity distribution for SuperDARN radars in the mid-latitudes. The velocity bins are 25 m/s in size. N is the total number of points in the plot. $N_{5\sigma_L}$ and $N_{5\sigma_G}$ indicate the amount of data points within 5 standard deviations of the mean for the Laplace fit and the Gaussian fit, respectively.	189
8.5	A plot of the average difference in ACF lag variance as calculated using Equation 8.2 between multiple-pulse sequences using $\tau = 1.8 \text{ ms}$ and $\tau = 2.4 \text{ ms}$. Improvement in error is indicated by red and deterioration is indicated by blue.	190
8.6	The posterior distribution functions for fitted power, velocity, spectral width, and plasma frequency resulting from a Bayesian fit of simulated data.	192

8.7	An illustration of the overlap of the fields of view of both RISR-C and the Clyde River SuperDARN radar during the 3 March, 2016 experiment.	194
8.8	A plot comparing the electron density measured along beam 42 of RISR-C at different altitudes with the Clyde River inferred electron density.	195
8.9	A Farley diagram of a multiple-pulse sequence using 3 pulses with $\tau = 2.4$ ms, $t_p = 300 \mu s$, $N_p = 3$, and $PTAB = [0\tau, 1\tau, 3\tau]$. The grey filled rectangles indicate transmission pulses and the rectangles with blue outlines indicate received voltage samples (from scattered signals returning to the radar). The red diamonds indicate signal backscattered from a repeated range of interest, in contrast to the black diamonds indicating signal backscattered from non-repeated unwanted ranges.	197

LIST OF ABBREVIATIONS

ACF	Auto-Correlation Function
CCBCL	Constant Chi-square Boundary as Confidence Limits
CSA	Canadian Space Agency
CSR	Coherent Scatter Radar
FPPM	First-Principles Fitting Methodology
FWHM	Full Width at Half Maximum
GDI	Gradient Drift Instability
HF	High Frequency
HL	High-Latitude
HPBW	Half-Power Beam Width
IMF	Interplanetary Magnetic Field
ISR	Incoherent Scatter Radar
LLBL	Low-Latitude Boundary Layer
MIPT	Multiple-Incoherent-Pulse Technique
ML	Mid-Latitude
MPOP	Multiple-Pulse Orthogonal Polarization
MPSE	Maximal Power-based Self-clutter Estimator
MSE	Mean-Square Error
MUD	Maximum Unambiguous Doppler
MUR	Maximum Unambiguous Range
MVSE	Maximal Voltage-based Self-clutter Estimator
NSERC	Natural Sciences and Engineering Research Council
PDF	Probability Distribution Function
PRT	Pulse Repetition Time
PSBL	Plasmasheet Boundary Layer
RADAR	Radio Detection and Ranging
RISR-C	Resolute Bay Incoherent Scatter Radar Canada
RTI	Range Time Intensity
SCR	Signal to Clutter Ratio
SMPRF	Simultaneous Multiple Pulse Repetition Frequency
SNR	Signal to Noise Ratio
SuperDARN	Super Dual Auroral Radar Network
VSE	Voltage-based Self-clutter Estimator

CHAPTER 1

INTRODUCTION

RAdio **D**etection **A**nd **R**anging, more commonly known as radar, is a powerful remote sensing technique used for both scientific and technological purposes. While radar systems are strictly constrained to using radio waves (and typically only 3 MHz to 300 GHz (e.g. *Richards et al.*, 2010)), the principles of radar are also used by lidar systems, which use visible light, and sonar systems, which use acoustic waves (*Skolnik*, 1980). A few examples of technologies that depend on radar include radar speed guns utilized by law enforcement officers to catch speeding drivers, Tesla electric cars with radar sensors for “autopilot” autonomous driving, and the reusable first stage of the SpaceX Falcon 9 rocket which uses a radar altimeter in combination with several other sensors to land after returning from space (*Fisher*, 1992; *The Tesla Team*, 2016; *Musk*, 2013). The earliest radars were moving target indicator radars, where radio waves were used to determine whether a metallic object was present or not and to discern its Doppler shift (*Levanon and Mozeson*, 1988). With the development of more sophisticated signal processing techniques, improved radar waveform design, and inexpensive computational resources, modern radars can be designed to accurately detect and resolve both the location and Doppler shift of a target of interest.

In the last several decades, radars have become invaluable and powerful tools for diagnosing the characteristics of the plasma in the upper atmosphere of the Earth. Two types of radar systems that are used for plasma measurements are incoherent scatter radars (ISRs) and coherent scatter radars (CSRs). My thesis discusses the development of an improved signal processing methodology for the CSRs that are used in the Super Dual Auroral Radar Network (SuperDARN) (*Greenwald et al.*, 1995), based on some of the techniques employed in ISR signal processing. The motivation for this topic arose from the unsuccessful analysis of a SuperDARN experiment that was designed to infer plasma density with minute-scale time

resolution using the frequency shifting technique of *Gillies et al.* (2011). It was determined that the experiment was unsuccessful because it had not been designed to produce velocity measurements with sufficiently small velocity uncertainty with minute-scale time resolution. This was needed for this experiment because the formula for electron density provided by *Gillies et al.* (2011) is very sensitive to the velocity uncertainty. Additional research and analysis revealed that the signal processing methodology for SuperDARN radars have several approximations and ad hoc conditions that in general prevents efficient and accurate estimation of both the velocity and the uncertainty in the velocity measured by SuperDARN radars, at least partially due to limited computing power in the past. Using the existing signal processing regime, it was not possible to quantitatively design an appropriate experiment for minute-scale resolution. Therefore this PhD dissertation is devoted to quantifying the uncertainty in the SuperDARN signal processing chain with the end goal of producing an improved methodology capable of producing accurate and reliable error estimates for SuperDARN velocity measurements. This research is applicable beyond SuperDARN velocity measurements, applying to other SuperDARN parameters such as power and spectral width and any parameters derived from them.

Quantifying the uncertainty in observations and predictions is a task of great importance to any type of measurement, whether for applied research or theoretical prediction. Fundamentally, uncertainty characterizes the extent of a lack of exact knowledge. Without error bars, the precision of a theoretical prediction cannot be readily known and therefore the prediction cannot necessarily be considered valid or useful. In engineering and physics, unquantified uncertainty may pose a significant danger to public safety or determine whether a project is successful or not. In science, experimental observations that do not include a quantification of the error in the measurement cannot reliably be interpreted nor utilized with theoretical models. In general, without knowing the accuracy of experiment measurements, the precision of theoretical predictions based on the measurements cannot be determined.

There are several sources of uncertainty: measurement error, systematic error, natural variation, inherent randomness, model uncertainty, and subjective judgement (*Regan et al.*, 2002). Measurement errors result from fundamental limitations or problems with measure-

ment equipment. This can include error caused by improper operation of the instrument¹. For example, measurement error in radar systems can be caused by ambiguity limitations of the transmitted waveform. Systematic errors are caused by biases in the measurement equipment or sampling procedures. For example, a fast method² of estimating power from K complex voltage samples involves calculating the square of the sum of the magnitudes of the real and imaginary components of the voltage samples, but this method introduces a $(\pi/4 - 1/2)/K$ systematic bias in the estimated power (*Farley and Hagfors, 2006*). Uncertainty caused by natural variations are due to the evolution of a physical system in space and time. It is often assumed that the mean and variance of a radar target do not change in the time it takes to sample the target, but this is not always a valid assumption. If a physical system is inherently random, measurements cannot be predicted and are characterized by a probability distribution, resulting in statistical sampling uncertainty. Model uncertainty is caused by usage of physical models that only include variables of largest significance, which might not always accurately model experimental observations. Finally, uncertainty may be unintentionally introduced due to subjective judgement while interpreting observational results. Naively discarding data so that observations agree with preconceived conclusions or existing models is a simple way to cause uncertainty due to subjective judgement³.

1.1 Thesis Outline

This thesis has been written in the manuscript style, where Chapters 5, 6, and 7 consist of manuscripts published in the American Geophysical Union journal, *Radio Science*. Each chapter includes a preamble that describes the significance of the work in the context of the remainder of the thesis.

The objective of this thesis is to develop a new data processing methodology for SuperDARN radars that improves upon current data processing techniques and provides a robust and reliable estimate of the measurement errors. An improved accounting of SuperDARN

¹The author is reminded of the epithet “Kjellmar User Error” used to refer to the measurement errors caused by improper operation of the EISCAT Svalbard Incoherent Scatter Radar by a certain individual.

²This method was originally used in the early days of ISR (in the 1970s) because it was computationally inexpensive.

³This is also referred to as being statistically dishonest (e.g. *Madden, 1982*).

measurement errors will enable custom experiment design and the development of new SuperDARN data products, such as electron density inferred using the frequency shifting technique.

The first few chapters of this thesis provide a general overview of geospace plasma physics and ionospheric radar concepts relevant to SuperDARN. Chapter 2 provides an overview of the geophysical context within which SuperDARN radars operate. Several important physical phenomena in the geospace system are discussed to provide context and understanding of the significant contributions to space physics research made by SuperDARN. Chapter 3 discusses SuperDARN with a short overview on the current status of the network. General radar principles relevant to SuperDARN are discussed with a focus on phased antenna array theory and radiophysics, as well as radio wave propagation in the context of magnetoionic theory. A description of the common SuperDARN data products is also included. Chapter 4 presents a survey of the signal processing and probability theory principles that were essential to accomplish the significantly improved SuperDARN data analysis performed in this dissertation. The chapter also includes a discussion of the statistical techniques and numerical methods utilized to process and fit SuperDARN data.

Chapters 5, 6 and 7 contain the core of the research of this thesis. They also include detailed discussions about results and the significance of the results to SuperDARN signal processing. Chapter 5 presents three novel self-clutter estimators that can be used to estimate the interference in SuperDARN signals due to self-clutter. It includes a discussion of how self-clutter contributes to estimation uncertainty. Chapter 6 presents the statistical properties of SuperDARN measurements, including how to use self-clutter estimates to accurately obtain error estimates required to accurately fit SuperDARN data. Specifically, the statistical characteristics of the autocorrelation functions estimated from SuperDARN voltage signal measurements are derived using knowledge about the statistical properties of SuperDARN voltage samples and estimator theory. Without knowledge of the statistical characteristics of the autocorrelation function, exact knowledge of the uncertainty in SuperDARN fitted data is not possible. Chapter 7 presents an improved fitting methodology for processing SuperDARN data. The fitting methodology builds upon the results and techniques developed in Chapters 5 and 6. The new methodology provides accurate error bars for SuperDARN parameters, such as power, spectral width, and velocity. A detailed analysis

using both synthetic and real data compares the current SuperDARN signal processing with the new fitting methodology.

Chapter 8 presents a summary of the results of this thesis, including a discussion of the progress towards achieving the thesis objectives. The chapter concludes with a discussion of future work, including presentation of progress already made towards further improvements to SuperDARN. These improvements include an experiment design methodology, preliminary results obtained from comparing electron density measurements inferred by SuperDARN with electron density measurements by the Resolute Bay Incoherent Scatter Radar, and an improved multiple-pulse sequence transmission technique.

CHAPTER 2

GEOSPACE PHYSICS

The field of geospace physics involves the study of the solar-terrestrial environment, which extends from the upper atmosphere of the Earth to the photosphere of the Sun, and details a plethora of physical processes spanning a multitude of spatial and temporal scales (*Hargreaves*, 1995b). More generally, space physics (of which geospace physics is a subset) encapsulates all plasma that exists throughout the universe. Geospace is dominated by plasma and magnetic fields of both terrestrial and solar origins. Plasma in the solar system is involved in numerous processes, including nuclear fusion in the core of the Sun, coronal mass ejections and solar flares on the surface of the Sun, the supersonic solar wind that fills interplanetary space, and the terrestrial aurora. Fundamentally, each of these processes involve energy and momentum transfer, which gives rise to the dynamics we observe. At the center of the solar system, the Sun emits tremendous amounts of energy as electromagnetic radiation, particles, and magnetic fields. The electromagnetic radiation alone amounts to $3.86 \times 10^{26} \text{ Js}^{-1}$ (*Kivelson and Russell*, 1995), of which only half of a billionth (5×10^{-10}) is intercepted by the Earth. The highest-energy photons (ultraviolet light, x-rays, and gamma rays) are absorbed by and ionize the upper atmosphere of the Earth, producing the ionosphere. Energy is also transferred from the solar atmosphere into the ionosphere via the solar wind: a supersonic stream of magnetically dominated diffuse plasma that flows from the Sun. The solar wind couples the magnetosphere of the Sun to the magnetosphere of the Earth. A myriad of phenomena are observed in both the magnetosphere and the ionosphere of the Earth, but here we will only introduce processes that provide a general context for the physical processes studied by SuperDARN. Fundamentally these processes arise due to energy originating from the Sun, so it is appropriate to start our discussion there. Next we will discuss the solar wind and how it interacts with objects in the solar system. We will focus specifically on

the interaction between the solar wind and the geomagnetic field of the Earth. Finally, we will discuss the ionosphere of the Earth and the general magnetosphere-ionospheric coupling processes that are responsible for the ionospheric plasma convection pattern measured by SuperDARN radars. Only processes essential for obtaining a general understanding of the significance of SuperDARN to geospace physics research will be presented.

2.1 The Sun

“The Sun is the nearest star; a glowing sphere of gas shining because of its heat, like a red hot poker.”

-Carl Sagan, “Carl Sagan’s Cosmos”

The Sun provides the energy needed to sustain all life on Earth, without which we would not survive for long. Yet in the context of the galaxy and observable universe, the Sun is merely an ordinary G2V spectral class star. Composed mostly of hydrogen ($\sim 90\%$) and helium ($\sim 10\%$), the structure of the interior of the Sun consists of several different layers that vary in both density and temperature (*Kivelson and Russell, 1995*). As illustrated in Figure 2.1, these layers are the core, the radiative zone, the convective zone, the photosphere, the chromosphere, and the corona (*Gombosi, 1998*). Compared with a solid body, such as the Earth, the photosphere is analogous to the “surface” of the Sun and the “atmosphere” of the Sun consists of the chromosphere and the corona. The energy output of the Sun originates from the core, where the density is approximately 150 times greater than that of water and the temperature is approximately 15×10^6 K (*Brekke, 2012*). In the early 1900s, it was believed that energy of the Sun was produced by the adiabatic compression of the gas as the massive gravity of the Sun compressed it. This was known as the contraction hypothesis¹. In the early twentieth century, contraction theory was discarded when it was discovered that nuclear fusion was the cause (*Eddington, 1920*). Subsequent work by others led to an understanding of important fusion reactions, like the proton-proton chain and the

¹Eddington on the contraction hypothesis: “Only the inertia of tradition keeps the contraction hypothesis alive - or, rather, not alive, but an unburied corpse.” (*Eddington, 1920*)

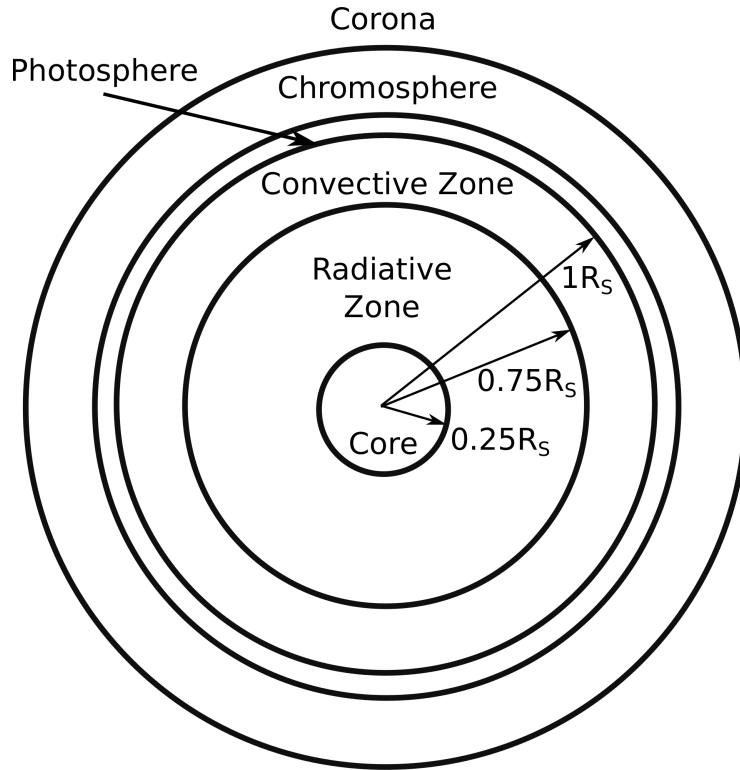


Figure 2.1: A diagram of the layers of Sun. The photosphere is between the chromosphere and convective zone. The outer layers of the solar atmosphere (chromosphere and corona) are also indicated. $R_S = 6.96 \times 10^8 \text{m}$ is the radius of the Sun. See text for more details. Reproduced from Figure 11.2 of *Gombosi* (1998).

carbon-nitrogen-oxygen cycle, that fuse hydrogen into helium and also synthesize heavier elements (*Burbidge et al.*, 1957).

The photon energies emitted by the Sun extend well beyond the visible spectrum, ranging from radio waves to gamma rays. In the core, gamma ray photons are produced by fusion reactions and slowly travel outward until they reach the radiative zone. The gamma rays are thought to be absorbed and re-emitted multiple times as they propagate away from the core such that by the time the gamma rays have reached the convection zone and the photosphere, they have been converted primarily into heat. Due to this continuous process of absorption and re-emission, it can take as long as 2×10^5 years for a photon produced in the core to reach the convective zone (*Brekke*, 2012), however the total time scale of energy transport from the core to the surface is 3×10^7 years (*Stix*, 2003). In the convection zone the density has decreased enough that outward energy transfer is dominated by convection (diffusion *and* advection) processes instead of radiation.

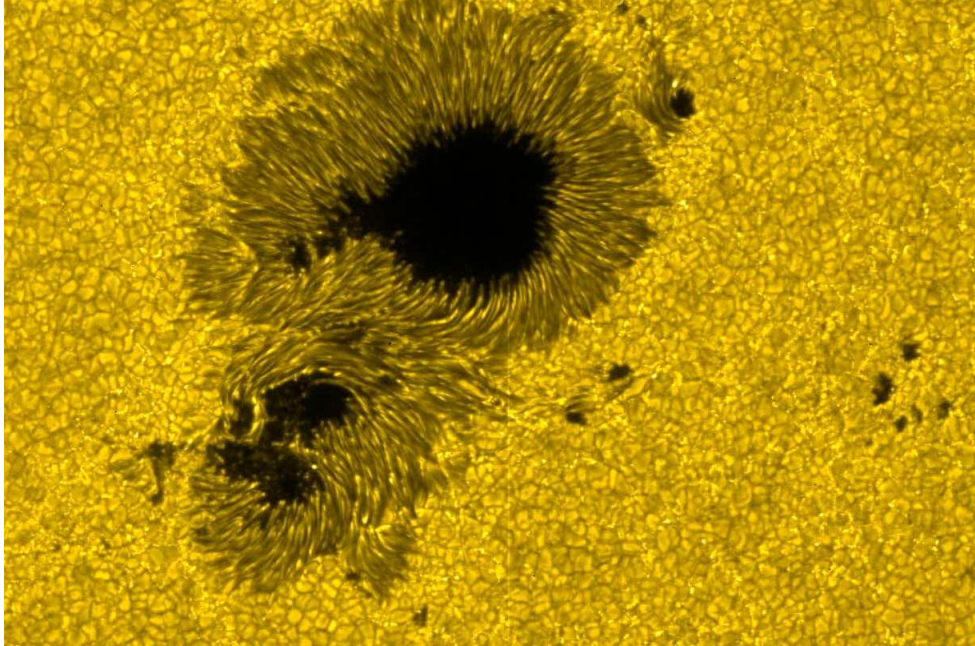


Figure 2.2: A picture of a sunspot obtained with Hinode’s Solar Optical Telescope on December 13, 2006. The dark appearance of the sunspot is due to the decreased luminosity of the sunspot since it is cooler than the surrounding plasma on the surface of the Sun. Courtesy NASA/JAXA (http://www.nasa.gov/mission_pages/solar-b/solar_022.html).

Observations show that the period of rotation of the Sun is on average 27 days, but the rotation period is a function of latitude. This phenomenon is known as differential rotation and is a consequence of plasma motion in the convection zone redistributing angular momentum (*Brekke, 2013*). Sunspots are a stellar phenomenon produced in the convection zone due to differential rotation. It is thought that differential rotation induces “stresses” on magnetic flux tubes inside the Sun that produces a concentrated magnetic field. Regions of strengthened magnetic field locally slow convective heat transport resulting in a number of sunspots appearing on the surface of the Sun. As shown in Figure 2.2, sunspots are cooler than the surrounding surface and thus appear darker due to their reduced luminosity. Sunspots are not a permanent feature observed on the Sun, as they release tremendous energy as they grow and fade (*Kivelson and Russell, 1995*).

The occurrence of sunspots follows an approximately 11 year cycle, with the number of sunspots increasing to a maximum and then decreasing to a minimum. This cycle is known as the solar cycle. The period of time when sunspot numbers are highest is called solar maximum

and the period of time when sunspot numbers are lowest is called solar minimum. In each solar cycle the magnetic polarity of the Sun reverses so that a complete cycle of this reversal occurs with a period of approximately 22 years (*Brekke*, 2013). The 11 year solar cycle periodicity is associated with periodic changes in geomagnetic activity observed on Earth, which produces beautiful displays of the Aurora Borealis and Aurora Australis. Increasing sunspot activity also results in coronal mass ejections, solar flares, and intensification of the solar wind (*Kivelson and Russell*, 1995).

2.2 The Solar Wind

The solar wind is a collisionless magnetohydrodynamic (MHD) fluid that propagates through interplanetary space at supersonic speed. As it propagates, it carries with it the interplanetary magnetic field (IMF). Within 4 to 40 solar radii of the Sun, the solar wind is radially structured by the strong magnetic field near the Sun. The magnetic field dominates the dynamics of the plasma and makes it very stable. Further away, the solar wind becomes turbulent as the strength of the IMF weakens. Recent observations by *DeForest et al.* (2016) show evidence of flocculae, or locally dense blobs of plasma, developing as the solar wind plasma becomes unstable and plasma instabilities develop due to the weakening IMF. As it propagates outwards from the rotating Sun, the solar wind forms a Parker spiral, a type of Archimedean spiral, similar to the spiral shape of a spinning ballerina skirt (*Mursula and Hiltula*, 2003) or the spiral path water follows from a rotating lawn sprinkler (e.g. *Brekke*, 2013). A current sheet, known as the heliospheric current sheet, flows in the region where the solar wind magnetic field reverses polarity and also follows the same spiral shape.

Essentially, the solar wind is caused by the outward expansion of the atmosphere of the Sun. The expanding atmosphere of the Sun produces a “bubble” of particles and magnetic field, called the heliosphere, that reaches deep into interplanetary space. Beyond Neptune, the solar wind becomes subsonic as it interacts with interstellar medium forming a termination shock; the first step in reaching interstellar space (*Brekke*, 2013). Recently, the Voyager 1 spacecraft crossed the boundary between the heliosphere and interstellar space at a distance of approximately 120 AU (*Fuselier and Cairns*, 2015).

The heliosphere provides the solar system with shielding from cosmic rays. The amount of shielding is a function of solar activity, with maximum (minimum) shielding occurring at solar maximum (minimum). The fluctuations in this shielding may be observed indirectly as fluctuations of Carbon-14 concentrations in tree rings (*Mori, 1981*). There is also evidence that shows a correlation in lightning occurrence rates (*Scott et al., 2014*) and cloud formation rates that suggests a cause related to modulation of cosmic rays due to varying solar activity (*Enghoff et al., 2011*).

The solar wind contains a fully ionized plasma composed mainly of protons and electrons and is approximately 3% fully-ionized Helium (*Baumjohann and Treumann, 1997*). Table 2.1 shows the range of values of parameters of the solar wind. On average, the solar wind has a speed of 500 km/s, a plasma density of 5 cm^{-3} , and temperature of 12,000 K at 1 AU from the Sun. The properties of the solar wind vary strongly with the 11-year solar cycle; for example, the solar wind speed is larger on average during periods of solar minimum than it is during periods of solar maximum (*Tokumaru et al., 2010*).

Parameter	Mean	Low	High
Density (cm^{-3})	5	1	20
Speed (km/s)	500	300	1400
Ion Temperature (K)	1.2×10^5	10^4	2×10^5
Magnetic Field (nT)	5	1	15

Table 2.1: Solar wind parameters typically observed at a distance of 1 AU from the Sun. Parameters values obtained from *Baumjohann and Treumann (1997)* and *Brekke (2013)*.

The MHD description of the solar wind is valid so long as the time-scale of any change to the plasma is much smaller than the ion-cyclotron frequency and the length-scale of the dynamics being described is much larger than the gyro-radius (*Baumjohann and Treumann, 1997*). These conditions are satisfied in the solar wind. Exploring the MHD description, we can prove that the IMF is “frozen-in” to and carried by the plasma in the solar wind. We start with the generalized Ohm’s law and Maxwell’s equations. From the generalized Ohm’s law

$$\frac{\mathbf{J}}{\sigma} = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \quad (2.1)$$

where \mathbf{J} , \mathbf{E} , and \mathbf{B} are the current density, electric field, and magnetic field, respectively, in the frame of the plasma moving with velocity \mathbf{v} and σ is the conductivity of the plasma. Combining Equation 2.1 with Faraday's Law of induction

$$\nabla \times \mathbf{E} = -\frac{\delta \mathbf{B}}{\delta t} \quad (2.2)$$

yields

$$-\frac{\delta \mathbf{B}}{\delta t} = \nabla \times (\mathbf{J}/\sigma - \mathbf{v} \times \mathbf{B}). \quad (2.3)$$

Finally, Ampère's Law for a static electric field, $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$, is used to obtain

$$\frac{\delta \mathbf{B}}{\delta t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B}, \quad (2.4)$$

where μ_0 is the vacuum magnetic permeability (*Baumjohann and Treumann, 1997*). The first term in Equation 2.4 describes how the motion of the plasma changes the magnetic field and the second term describes how the magnetic field is changed by diffusion. When the conductivity is infinite, diffusion is not possible ($1/\sigma \rightarrow 0$), so the magnetic field must move with the plasma. This is made explicitly clear by considering the magnetic flux, Φ , through a surface S , represented mathematically as

$$\Phi = \int_S \mathbf{B} \cdot \hat{\mathbf{n}} dS. \quad (2.5)$$

Here we take our surface to be that of a cylindrical tube, called a magnetic flux tube. If this surface is moving with velocity \mathbf{v} , calculating the partial derivative of Equation 2.5 with respect to time results in

$$\frac{\delta \Phi}{\delta t} = \int_S \frac{\delta \mathbf{B}}{\delta t} \cdot \hat{\mathbf{n}} dS + \oint_L \mathbf{B} \cdot (\mathbf{v} \times d\boldsymbol{\ell}). \quad (2.6)$$

Using Stokes Theorem on the second term of the right-hand side of the equation yields

$$\frac{\delta \Phi}{\delta t} = \int_S \left(\frac{\delta \mathbf{B}}{\delta t} - \nabla \times (\mathbf{v} \times \mathbf{B}) \right) \cdot \hat{\mathbf{n}} dS. \quad (2.7)$$

Recalling that the conductivity of the solar wind is extremely large ($\sigma \rightarrow \infty$) and that the diffusion term in Equation 2.4 is then 0, it is clear that the right-hand side of Equation 2.7 becomes 0. This means that the magnetic flux must be constant in a flux tube that moves with a plasma that has very large conductivity. This is known as Alfvén’s Theorem, or more commonly, the frozen-in theorem (*Alfvén, 1942*). As will be discussed later, Alfvén’s Theorem also applies to high altitude terrestrial plasmas in the ionosphere of the Earth.

The solar wind interacts with objects in the solar system as it propagates through interplanetary space. Another interesting insight from the MHD description of the infinitely conducting solar wind follows from Equation 2.1. Taking the limit as the conductivity approaches infinity gives

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0, \tag{2.8}$$

which shows that in the frame of the moving plasma (i.e. $\mathbf{v} = 0$), there are no electric fields in an infinitely conducting plasma. However, outside the frame of the moving plasma, as the solar wind propagates past bodies in the solar system, such as the Earth (i.e. $\mathbf{v} \neq 0$), an electric field of $\mathbf{E} = -\mathbf{v} \times \mathbf{B}$ is induced.

We will focus primarily on the interaction of the solar wind with the magnetic field of the Earth, but the solar wind interacts with all objects in the solar system. Some of these objects have an atmosphere and a very weak magnetic field, such as comets. The solar wind interacts directly with the ionosphere of such objects, compressing the ionosphere and producing more ionization through particle impacts (*Edberg et al., 2016*). The planet Venus is also a good example of such a solar system object (*Bauer et al., 1977*). Some objects do not have an atmosphere and instead the solar wind directly interacts with the surface of the object resulting in the production of an ionosphere. The moon of the Earth is an example of such an object (*Choudhary et al., 2016*). Next the interaction between the magnetosphere of the Earth and the solar wind will be discussed.

2.3 The Magnetosphere

The region of space surrounding the Earth that is dominated by the geomagnetic field of the Earth is called the magnetosphere. The magnetosphere of the Earth shields the terres-

trial atmosphere from the solar wind. It is postulated that a magnetosphere helps reduce atmospheric erosion due to the interaction of the solar wind with the atmosphere of a planet (*Barabash et al.*, 2007). Within approximately 5 Earth radii (R_E) of the Earth, the geomagnetic field can be modelled as a dipole field (*Brekke*, 2013). The geomagnetic field of the Earth can be modelled using the dipole field equation

$$\mathbf{B} = (2Mr^{-3} \cos \theta) \hat{\mathbf{r}} + (Mr^{-3} \sin \theta) \hat{\boldsymbol{\theta}}, \quad (2.9)$$

where $M = 8 \times 10^{15} \text{ Tm}^3$ is the dipole magnetic moment of the Earth, r is the distance from the centre of the Earth and θ is the magnetic colatitude (*Kivelson and Russell*, 1995). The magnetic poles of the geomagnetic field are tilted by approximately 10° relative to the rotation axis of the Earth. Currently, the north magnetic pole is located in the Southern hemisphere and the south magnetic pole is located in the Northern hemisphere. It is interesting to note that measurements of old magnetic volcanic rock shows that the geomagnetic field can undergo polarity reversals, similar to polarity reversals on the Sun, but on time scales of the order of 10^6 years (*Brekke*, 2013).

In reality, the geomagnetic field is not a perfect dipole and is more accurately modelled using a spherical harmonic expansion that includes multi-pole contributions that distort the dipole shape (*Brekke*, 2013). This is illustrated in Figure 2.3, which shows a 2D plot of the magnetosphere of the Earth. The geomagnetic poles are geographically asymmetric: the magnetic poles do not lie along a line that passes through the centre of the Earth. For example, the north geomagnetic pole in the Southern hemisphere is more than 8.5° farther from the rotation axis of the Earth than the south geomagnetic pole is. This produces a diurnal variation in the location of magnetically conjugate points on the surface of the Earth, presenting a challenge for measuring magnetically conjugate events (*Laundal et al.*, 2016). Also, the 10° tilt of the geomagnetic axis creates time-dependent asymmetries in the electric fields in the magnetosphere, resulting in enhanced plasma flows in some regions (*Sojka et al.*, 1979).

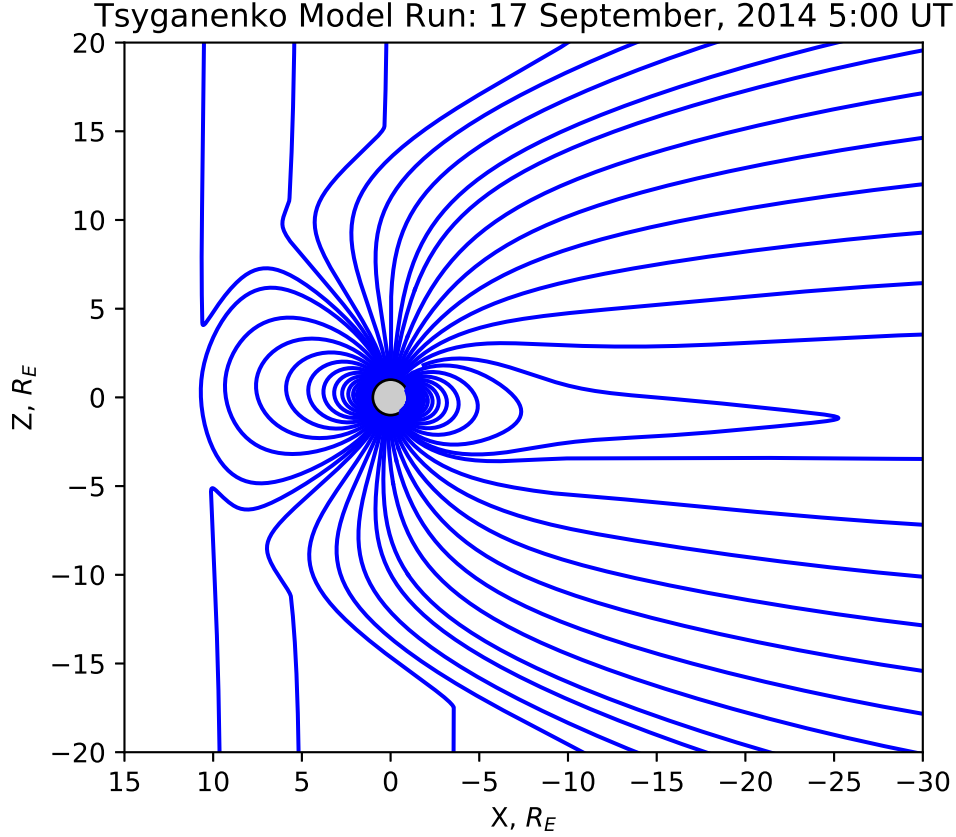


Figure 2.3: The magnetosphere of the Earth modelled using the Tsyganenko 1995 model (*Tsyganenko, 1995*). At distances less than $5 R_E$, the magnetic field is approximately dipolar and can be accurately modelled using Equation 2.9. Figure generated using DaViT-py (<https://github.com/vtsuperdarn/davitpy>).

2.3.1 Magnetic Reconnection

Interaction with the solar wind distorts the dipole shape of the geomagnetic field. In Figure 2.3, the solar wind is propagating from the left to right, so the sunward direction (dayside) is given by positive X , geographic north is in the positive Z direction, and the Earth is centred at $(0,0)$. Notice that the magnetosphere is compressed by the solar wind on the sunward side and elongated into a “tail” (called the magnetotail) on the anti-sunward side. The IMF couples to the geomagnetic field on the dayside of the magnetosphere and the solar wind continues to flow past the Earth producing an electric field that drives plasma on the newly merged geomagnetic field lines. The plasma flow stretches the geomagnetic field in the anti-

sunward direction producing the magnetotail, which is stretched out to hundreds of Earth radii. The magnetosphere can be visualized as a magnetic wind sock fluttering in a magnetic wind.

In general, the solar wind compresses the magnetic field on the dayside of the Earth such that the magnetic pressure increases until it is able to balance the dynamic pressure of the solar wind (*Brekke*, 2013). This is illustrated in Figure 2.4. The region of pressure balance between the solar wind and the magnetosphere is called the magnetopause. In the magnetopause the IMF and geomagnetic field of the Earth can couple through a process called magnetic reconnection (or merging). Through magnetic reconnection, the IMF magnetically couples the atmospheres of both the Sun and the Earth. Magnetic reconnection can occur where a large enough shear exists between the IMF and geomagnetic field (*Dungey*, 1961; *Gombosi*, 1998). In the region where magnetic reconnection occurs, the plasma dynamics are kinetically dominated and the topology of the magnetic field is reconfigured. The plasma in this region is no longer frozen-in and instead the diffusion term in Equation 2.4 becomes much more important (*Kivelson and Russell*, 1995).

By Ampère’s Law, shears between the IMF and geomagnetic field requires currents to exist. On the dayside of the magnetosphere, the induced current is called the Chapman-Ferraro current (the dayside current out of the page in Figure 2.4). In the magnetotail, the region of anti-sunward magnetosphere, a cross-tail current is induced (the nightside current out of the page in Figure 2.4). Several models of reconnection exist (*Parker and Krook*, 1956; *Sweet*, 1958; *Petschek*, 1964), but a full theoretical description of the physics of reconnection remains an open problem, especially in regards to the kinetic physics at electron-scales. Recently, a four satellite NASA mission called the Magnetosphere Multiscale Mission made measurements of a reconnection event in the magnetopause and directly observed the electron-scale physics (*Burch et al.*, 2016). By the process of reconnection, the magnetic flux on the magnetosphere of the Earth that is “closed” to the solar wind (closed flux) becomes open through reconnection becoming “open” to the solar wind on the dayside. Reconnection also occurs in the magnetotail, decoupling the IMF from the geomagnetic field. On the dayside, reconnection converts particle energy into magnetic energy and on the nightside, reconnection converts magnetic energy into particle energy.

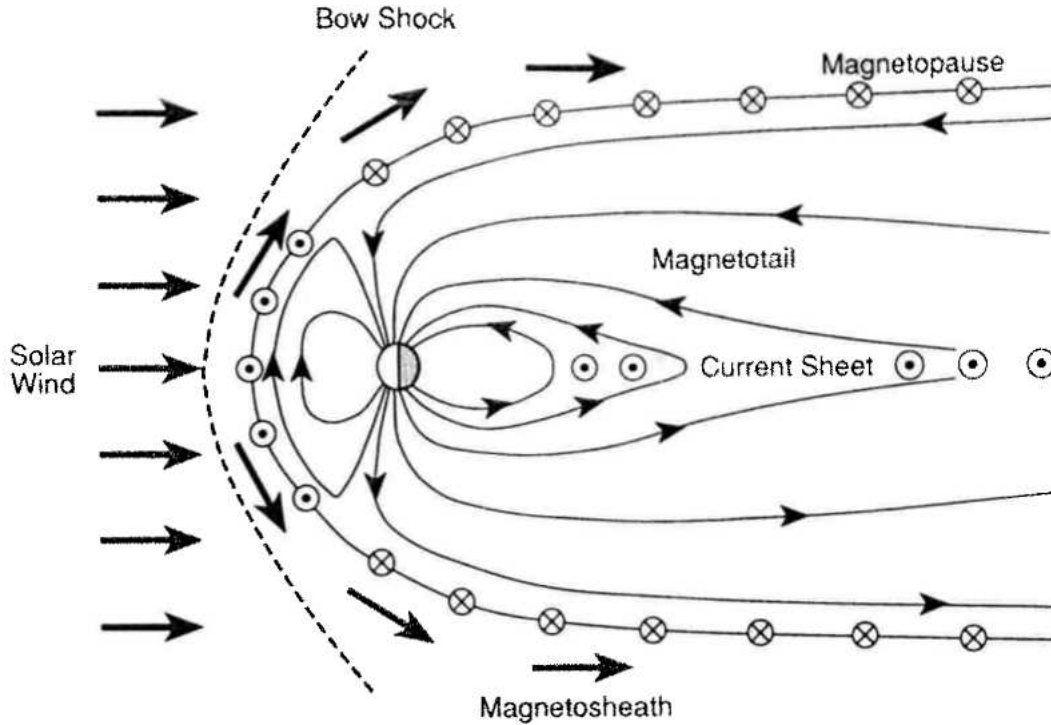


Figure 2.4: The outer magnetosphere of the Earth and magnetospheric regions seen from the noon-midnight meridian. The incident solar wind forms a bow shock as it interacts with the magnetosphere of a planet with a magnetic field. A balance between solar wind and magnetic field pressures occurs at the magnetopause. The Chapman-Ferraro current flows in this region of pressure balance (current coming out of the page on the sunward side of the magnetosphere). The magnetosheath is a region of shocked plasma between the bow shock and magnetopause. The magnetotail is caused by the solar wind stretching out the magnetosphere downstream. Magnetospheric currents are indicated by circles filled with dots or Xs and are discussed in more detail in the text. Figure 9.1 from *Kivelson and Russell (1995)*.

2.3.2 Magnetospheric Regions

Since the solar wind is supersonic, upstream of the magnetopause a bow shock is produced when the solar wind transitions from supersonic flow to subsonic flow (*Kivelson and Russell, 1995*). The location of the bow shock is illustrated in Figure 2.4. Between the magnetopause and the bow shock is a region of shocked plasma called the magnetosheath. The locations of the magnetospheric currents induced in the magnetopause and magnetotail are also illustrated in Figure 2.4.

The magnetosphere is composed of several regions. Figure 2.5 shows several of these

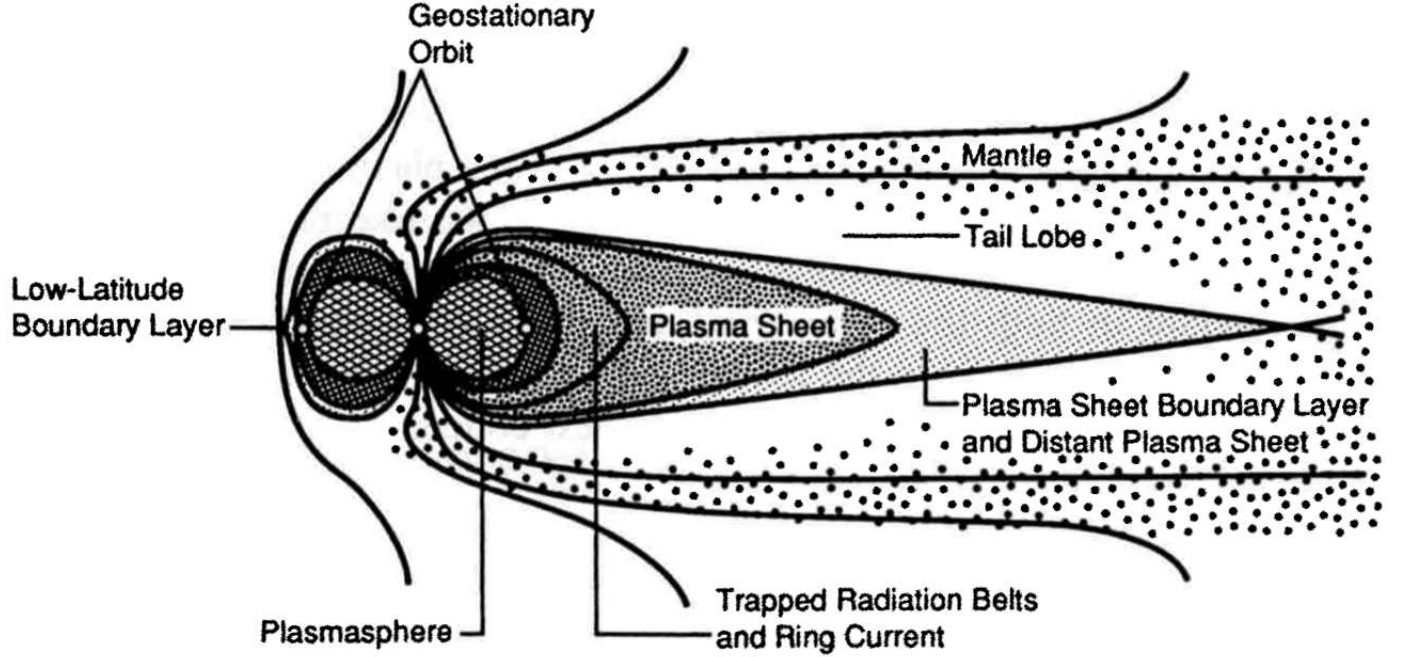


Figure 2.5: The locations of magnetospheric plasma layers and magnetospheric currents of the Earth are shown as viewed from the noon-midnight meridian plane. The Chapman-Ferraro current is shown in Figure 2.4 and lies sunward of the LLBL in this figure. Layer characteristics are discussed in the text. Figure 10.4 from *Kivelson and Russell (1995)*.

regions, where the magnetosheath and solar wind are “upwind” to the left in the figure. Different layers are characterized by both different particle energies and different particle sources. Table 2.2 summarizes typical values of density, temperature, and magnetic field strength observed in different regions of the magnetosphere. The magnetospheric region closest to the Earth is called the plasmasphere, which extends from $3\text{--}5 R_E$ (*Carpenter, 1963*). To first order, the plasmasphere co-rotates with the Earth, according to Ferraro’s Theorem, and is populated by particles from the ionosphere of the Earth (*Brekke, 2013*). Measurements by the IMAGE satellite showed that plasma originating from high latitudes in the plasmasphere lags the corotation by approximately 10–15% due to having lower angular momentum than plasma originating from lower latitudes (*Burch et al., 2004*). The Plasma Sheet (PS) extends from $5 R_E$ to $8 R_E$ and contains charged particles with energies of several keV that mostly originate from the ionosphere of the Earth. On the dayside of the magnetosphere, the Low-Latitude Boundary Layer (LLBL) is a region between the PS and

Parameter	Magneto-	Mantle/	Plasma		
	sheath	Tail Lobe	LLBL	Sheet	Plasmasphere
Density (cm^{-3})	8	0.01	0.1	0.3	10^3
Ion Temperature (eV)	150	300	1,000	4,200	1
Electron Temperature (eV)	25	50	150	600	1
Magnetic Field (nT)	2.5	3×10^{-3}	10^{-1}	6	$\sim 100 - 1000$

Table 2.2: A summary of typical plasma and magnetic field parameter values for regions of the magnetosphere shown in Figure 2.5 (*Kivelson and Russell, 1995*). The magnetic field strength in the plasmasphere can be approximated using the magnitude of Equation 2.9, and is therefore altitude dependent.

the mantle (*Kivelson and Russell, 1995; Brekke, 2013*). The LLBL is populated by plasma from both the magnetosheath and the magnetosphere and is located poleward of the PS. Poleward of the LLBL is the cusp, a region associated with particle precipitation of low energy ($\lesssim 1$ keV) originating in the solar wind (*Kivelson and Russell, 1995*). The mantle is poleward of the cusp and it contains plasma with density, temperature, and flow speed decreasing as a function of the distance from the cusp. Across the polar region and into the magnetotail, the mantle forms the tail lobe. When enough lobe lines “pile up” in the tail, the magnetic pressure increases until reconnection closes them to form the Plasma Sheet Boundary Layer in the tail.

2.4 The Ionosphere

The ionosphere of the Earth is a region of the upper atmosphere extending from approximately 60 – 1000 km (*Kivelson and Russell, 1995*). Ionospheres result from a combination of two ingredients, a neutral atmosphere and a source of ionization. The atmospheric density profile varies with altitude according to the exponential relationship given by

$$n(z) = n_0 \exp\left(-\frac{z - z_0}{H}\right), \quad (2.10)$$

where z is the altitude, n_0 is the a reference density at reference altitude z_0 , and H is the scale-height or the increase in altitude where the density decreases by a factor of e (scale-

height is also called the e -folding height) (Brekke, 2013). The scale height is $H = k_B T / mg$, where k_B is the Boltzmann constant, T is the neutral atmosphere temperature, m is the mass of a gas particle, and g is the acceleration due to gravity. Figure 2.6 shows the density profiles for the neutral gases Ar, He, O₂, N₂, and O as a function of altitude in the ionosphere. The ionosphere of the Earth is produced by ionizing these neutral atmosphere gases.

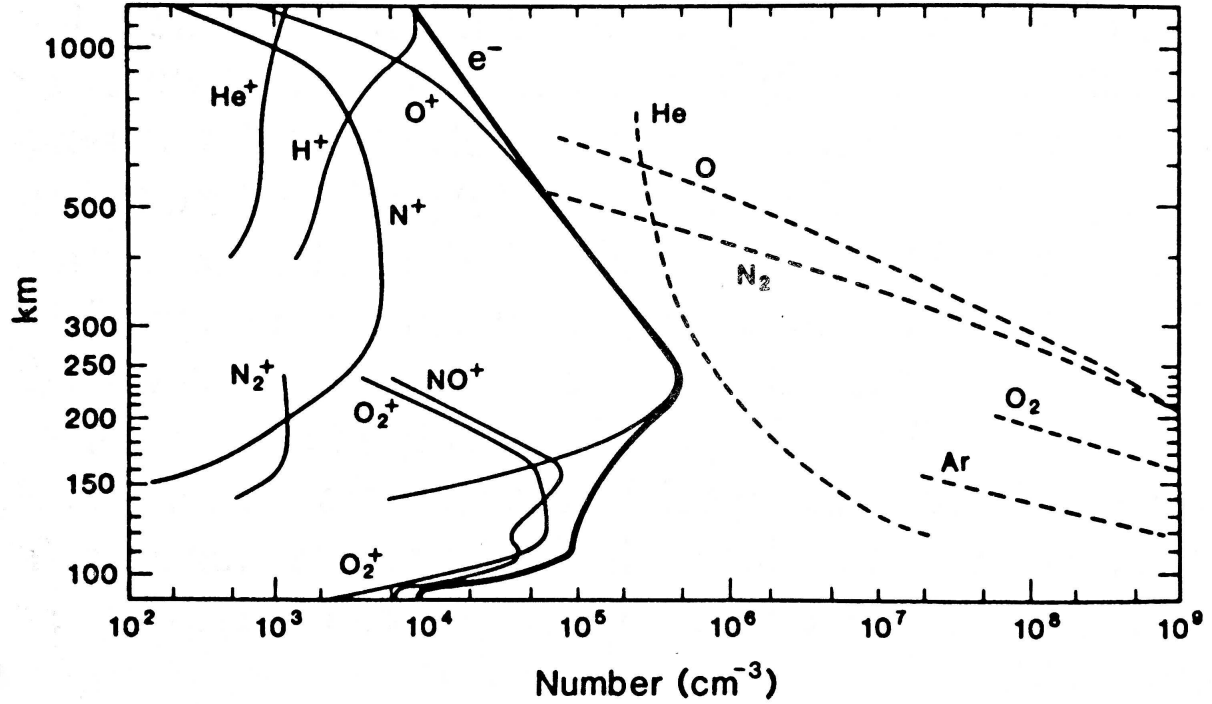


Figure 2.6: The altitude profile for a variety of atoms and molecules in both the neutral atmosphere and ionosphere of the Earth at ionospheric altitudes. Note that in the E region, NO⁺ and O₂⁺ molecular ions dominate and in the F region, O⁺ atomic ions are the major ionic species (Kivelson and Russell, 1995).

Both photoionization and particle impact ionization (Kivelson and Russell, 1995) produce the ionosphere. Photoionization is produced from gamma rays, x-rays, and ultraviolet rays and impact ionization is produced from energetic particle precipitation where charged particles (ions and electrons) precipitate down magnetic field lines into the atmosphere. During the day photoionization is the dominant plasma production process. Different atmospheric constituents have different photoionization cross-sections and the amount of ionization produced is a function of cross-section, initial intensity, zenith angle, and density profile. The ionization profile can be calculated using the Chapman production function

$$q = q_m \exp \left[1 - \frac{z - z_m}{H} - \sec \chi \exp \left(\frac{z - z_m}{H} \right) \right], \quad (2.11)$$

where z is altitude, q_m is the peak production which occurs at an altitude of z_m , χ is the solar zenith angle, and H is the scale-height of the neutral atmosphere (*Brekke, 2013*). The Chapman production function can also be used to model the impact ionization caused by energetic particle precipitation. Higher energy particles penetrate deeper into the atmosphere than lower energy particles. Figure 2.6 shows the density profile for several ions in the ionosphere: O_2^+ , NO^+ , N_2^+ , N^+ , O^+ , H^+ , and He^+ . Notice that different altitude regions are dominated by different ions and neutral particles. To first-order, the O^+ profile has the same “nose” shape as the Chapman production function for O^+ , since the dominant mechanism for O^+ production is photoionization. The profiles for NO^+ and O_2^+ do not have the same nose-like shape due to plasma transport, recombination, and chemical reaction based production processes dominating the ion production (*Brekke, 2013*).

At night, the ionospheric density decreases significantly and the dominant mode of plasma production is impact ionization from energetic particle precipitation. *Fang et al. (2010)* describes parametrization methods of calculating the expected ionization profiles based on particle precipitation (for mono-energetic particles, the profile resembles the Chapman production function for a single atmospheric species). Precipitating electrons may produce x-rays through bremsstrahlung due to Coulomb collisions. These x-rays may produce ionization at altitudes as low as 25 km and the expected ionization profile may be calculated as discussed in *Fang et al. (2010)* and references therein.

Differing neutral atmosphere composition at different altitudes produces regions in the ionosphere with differing chemistry. An illustration of these regions can be seen in the ion and electron density profile of Figure 2.6. The D region (not shown), extends from around 60 km to 90 km and contains the most complex chemistry, with plasma consisting of electrons, positive ions, and negative ions. The neutral density in the D region produces a large electron-neutral collision frequency. The E region extends from 90 km to 130 km (*Kivelson and Russell, 1995*), 140 km (*Gombosi, 1998*), or 150 km (*Kelley, 2009*) and contains only positive ions and electrons since the neutral atmosphere density and electron-neutral collision frequency have decreased too much to produce negative ions above 90 km. The F region is

the uppermost layer of the ionosphere and extends upwards from 150 km. Throughout each of the layers, there exists a background neutral density that interacts with the plasma mainly through ion-neutral collisions. Figure 2.6 shows that the E region plasma is dominated by the molecular ions NO^+ and O_2^+ , whereas the F region is dominated by the atomic ion O^+ .

2.4.1 Plasma Waves

Understanding the waves that exist in ionospheric plasma allows for its diagnosis and study. This principle, coupled with models, allows radars, satellites, heaters, and other instruments to probe and/or measure the parameters of space plasmas. Ionospheric plasma is host to many natural wave modes due to the collective behaviour of charged particles. The collective behaviour of electrons and ions in the plasma is mediated via the Coulomb force and the Lorentz force. This means that electrons and ions are not completely free in the sense that their motions are correlated and influence one another. Here we will briefly present and discuss the linear waves that exist in magnetized and unmagnetized plasmas. The dispersion equations for these waves are derived from the fluid momentum equations for a cold plasma and assuming quasi-neutrality (the density of the ions and electrons is approximately equal) of the plasma. These derivations can be found in many introductory plasma physics textbooks (e.g. *Chen*, 1984; *Kivelson and Russell*, 1995; *Baumjohann and Treumann*, 1997).

In an unmagnetized plasma or along the magnetic field lines in a magnetized plasma (like the ionosphere of the Earth), electrons undergo an electrostatic oscillate at the plasma frequency given by

$$\omega_p^2 = \frac{n_e e^2}{\epsilon_0 m_e}, \quad (2.12)$$

where n_e is the electron density, e is the elementary charge, ϵ_0 is the vacuum permittivity, and m_e is the electron mass. This mode is restorative in that it is produced when electrons are displaced which produces electric fields that attempt to correct the displacement. Due to the large difference between electron and ion mass, the plasma frequency is much faster than the time scales that the ions can respond on. The random thermal motion of the electrons cause these oscillations to propagate as electron plasma waves

$$\omega_{ep}^2 = \omega_p^2 + \frac{3}{2} k^2 v_{th}^2, \quad (2.13)$$

where $k = 2\pi/\lambda$ is the wavenumber of the plasma waves, λ is the wavelength of the waves, and v_{th}^2 is the thermal velocity of the electrons. Electron plasma waves are dispersive; their velocity, given by ω/k , is a function of k . This means that the propagation velocity of the plasma waves depends on the wavelength of the waves, where waves with smaller wavelengths (larger k) propagate faster than waves with larger wavelengths (smaller k).

Another natural mode of an unmagnetized plasma is the ion-acoustic wave. These waves are akin to sound waves in a neutral gas, except in a plasma the Coulomb repulsion between ions adds to the pressure force. The frequency of ion-acoustic waves is given by

$$\omega_{ia}^2 = \left(\frac{k_B T_e + 3k_B T_i}{m_i} \right) k^2, \quad (2.14)$$

where k_B is the Boltzmann constant, T_e and T_i are the electron and ion temperatures, and m_i is the ion mass. Since electrons respond much faster to perturbations in the plasma, the effect of electron inertia on ion-acoustic waves is neglected in deriving Equation 2.14 (*Baumjohann and Treumann, 1997*). In a more general form of Equation 2.14, the numerator is replaced by $\gamma_e k_B T_e + \gamma_i k_B T_i$, where $\gamma_{e,i}$ is the specific heat capacity ratio of the electrons and ions. In the most commonly used form of the equation, Equation 2.14, the electrons are isothermal ($\gamma_e = 1$) and the ions are adiabatic with one-dimensional motion ($\gamma_i = 3$). In contrast to electron plasma waves, the velocity of ion-acoustic waves is non-dispersive; it is not a function of k .

In a magnetized plasma, both ions and electrons gyrate about the magnetic field with a frequency given by

$$\Omega_q = \frac{qB}{m_q}, \quad (2.15)$$

where B is the magnetic field strength, q is the charge of the particle, and m_q is the mass of the particle. For waves that propagate at oblique angles to the magnetic field, two modes for each wave exist simultaneously and the expressions get much more complicated.

2.4.2 Currents

Unlike the solar wind and the magnetosphere, the conductivity of the plasma in the ionosphere is finite, allowing for currents to flow. The currents in the ionosphere can be classified

into three types: magnetic field-aligned current, Hall current, and Pederson current. Satisfying Ohm's law, each current is a function of its associated conductance. For a single ion species plasma, the current density taken in the reference frame of the neutral atmosphere motion may be written as (*Brekke, 2013*)

$$\mathbf{j} = \sigma_P \mathbf{E}_\perp - \sigma_H \frac{\mathbf{E} \times \mathbf{B}}{B} + \sigma_\parallel \mathbf{E}_\parallel, \quad (2.16)$$

where the terms on the right hand side of the equation give the Pederson current, the Hall current, and the field-aligned current. \mathbf{E}_\perp and \mathbf{E}_\parallel denote the components of the electric field that are perpendicular and parallel to the magnetic field \mathbf{B} , respectively. The conductivities are given by

$$\sigma_P = \frac{n_e e}{B} \left(\frac{k_e}{1 + k_e^2} + \frac{k_i}{1 + k_i^2} \right), \quad (2.17)$$

$$\sigma_H = \frac{n_e e}{B} \left(\frac{k_e}{1 + k_e^2} - \frac{k_i}{1 + k_i^2} \right), \quad (2.18)$$

$$\sigma_\parallel = \frac{n_e e}{B} (k_e + k_i), \quad (2.19)$$

respectively, where $k_e = \Omega_e / \nu_{en}$, $k_i = \Omega_i / \nu_{in}$, ν_{en} is the electron-neutral collision frequency, and ν_{in} is the ion-neutral collision frequency. As given by Equation 2.16, Pedersen currents flow in the direction of the electric field perpendicular to the magnetic field, Hall currents flow in the $\mathbf{E} \times \mathbf{B}$ direction and parallel currents flow along magnetic field lines.

In the E region and the F region the electron-neutral collision frequency is low and the electron motion is dominated by $\mathbf{E} \times \mathbf{B}$ drift. Below approximately 125 km, the ion-neutral collision frequency is large enough that the ions predominantly drift in the \mathbf{E} direction. Above 125 km the ion-neutral collision frequency is small enough that the ions $\mathbf{E} \times \mathbf{B}$ drift with the electrons. This results in the Hall current being larger in the E region than the F region and the Pedersen current being the dominant current in the F region. Hall currents are carried mostly by electrons in the E region due to the large collision frequency for ions, compared to a small collision frequency for the electrons. Field-aligned currents can be carried by electrons or ions, either by particle precipitation or up-flow events. Field aligned currents are sourced from the magnetosphere and participate in magnetosphere-ionosphere coupling processes.

2.5 Magnetosphere-Ionosphere Coupling

The magnetosphere and ionosphere are coupled by currents that flow along magnetic field lines into and out of the ionosphere. Currents that are sourced by magnetospheric generators and flow into the ionosphere are closed by Pedersen currents in the ionosphere before flowing up magnetic field lines and returning to the magnetosphere. Two important currents linking the ionosphere and magnetosphere are the region 1 (R1) and region 2 (R2) currents, known as the Birkeland² currents that flow from the magnetosphere into the polar ionosphere (*Kelley, 2009*).

The region above the Earth near the magnetic poles containing the open geomagnetic flux in the mantle region of the magnetosphere is called the polar cap. In general, the size of the polar cap can increase and decrease depending on the reconnection rates on the dayside and nightside of the magnetosphere. An expansion of the polar cap (increase in amount of open magnetic flux in the mantle) occurs when the dayside reconnection rate is greater than the nightside reconnection rate and contraction of the polar cap (decrease in amount of open magnetic flux in the mantle) occurs when the nightside reconnection rate is larger than the dayside reconnection rate (*Milan et al., 2012*). Figure 2.7 shows a diagram of the polar region ionosphere in the Northern hemisphere as viewed from above. In the figure, the polar cap is the region bounded by the inner circle. The R1 current sheet flows into the ionosphere on the dawn side of the polar cap and out of the ionosphere on the dusk side of the polar cap. The opposite is true of the R2 current sheets (not shown) which lies equatorward of the R1 currents.

Figure 2.7 also shows a dawn–dusk electric field, E_{DD} , that exists in the polar cap. As the solar wind propagates past the Earth, an electric field of

$$\mathbf{E}_{DD} = -\mathbf{v} \times \mathbf{B} \quad (2.20)$$

is induced in the frame of reference of the magnetosphere, where v is the solar wind velocity relative to the magnetosphere and B is the IMF. Since the conductivity along magnetic field

²In 1909, with the terrella experiment, Kristian Birkeland first experimentally demonstrated how the magnetic field of the Earth may direct charged particles in the solar wind to produce patterns similar to the aurora.

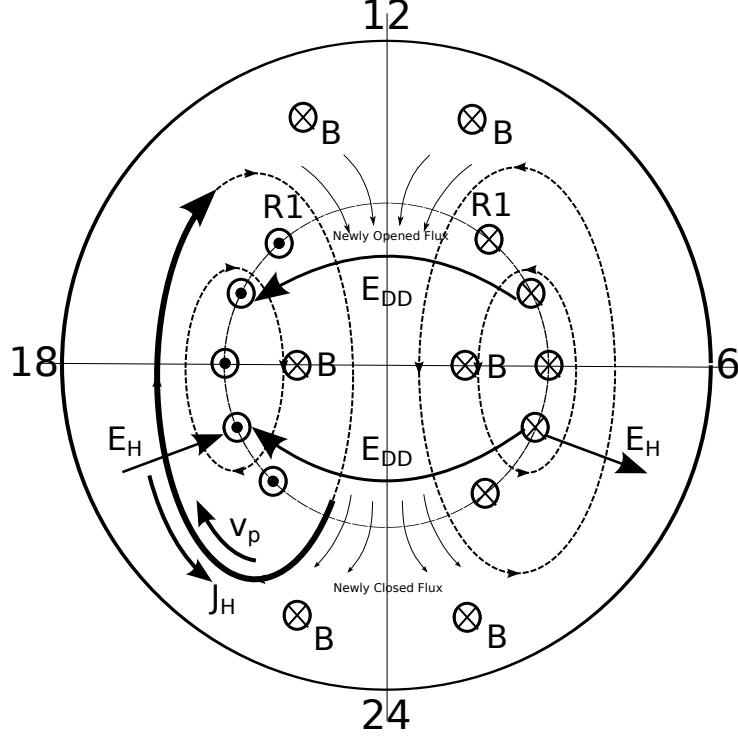


Figure 2.7: Figure illustrating magnetic convection as seen from above the northern hemisphere. The oval dashed curves illustrate the flow pattern of magnetic field lines from the dayside (12) to nightside (24) and back around to the dayside that results when the IMF is dominated by $B_z < 0$ (more information found in Section 2.3). E_{DD} is the Dawn-Dusk electric field generated by magnetic convection from 12 (noon) to 24 (midnight) over the polar cap (small circle inside largest circle). The region 1 current (R1), the polarization electric field (E_H), the eastward-electrojet (J_H) and plasma flow velocity (v_p) in the F region of the ionosphere are also indicated. The bold arrowed arc following the convection flow pattern in the dusk sector (18) indicates the eastward electrojet.

lines is nearly infinite, the dawn–dusk electric field induced by the solar wind maps along magnetic field lines down into the polar cap. This electric field then drives plasma flows in the polar cap ionosphere. In general, this electric field drives an $\mathbf{E} \times \mathbf{B}$ drift,

$$\mathbf{v}_{E \times B} = \frac{\mathbf{E} \times \mathbf{B}}{|\mathbf{B}|^2}, \quad (2.21)$$

in the F region of the polar cap ionosphere as illustrated in Figure 2.7 (recall the frozen-in discussion in Section 2.2, equations 2.1 to 2.8). Plasma is driven from dayside to nightside by the $\mathbf{E} \times \mathbf{B}$ drift producing a two cell plasma convection pattern. In the auroral region of the magnetosphere, a polarization electric field, \mathbf{E}_H , is generated between the R1 and R2 currents. This electric field points in the opposite direction to the dawn-dusk electric field

on the dusk side of the polar cap (towards the R1 current sheet as shown in Figure 2.7). On the dawn side of the polar cap, the polarization electric field points away from the R1 current sheet. The polarization electric field drives an additional $\mathbf{E} \times \mathbf{B}$ drift that produces the return flow from nightside to dayside by travelling around both the duskside or dawnside of the polar ionosphere.

The shape of the two cell convection pattern changes depending on the orientation of the IMF (*Cowley and Lockwood, 1992*). The IMF may take on a variety of orientations and is commonly described by a B_x , B_y , and B_z component in the geocentric solar magnetospheric coordinate system with origin at the center of the Earth, the Sun-Earth line as the x-axis (positive toward the Sun) and the y-axis as the cross product between the x-axis and the magnetic dipole z-axis. For $B_z < 0$ magnetic reconnection occurs on the dayside and a two-cell convection pattern is developed as illustrated in Figure 2.7. In what is known as the Svalgaard-Mansurov (*Kivelson and Russell, 1995*) effect, IMF $B_y \neq 0$ alters the dayside shape of the convection pattern. The convection pattern can even develop into a three- or more-cell pattern depending on IMF conditions. For example, when IMF $B_z > 0$ the two-cell convection pattern transforms into a four-cell pattern with two cells in the noon sector and two cells in the midnight sector. As will be presented in Chapter 3, SuperDARN is used to measure the plasma convection pattern in both the northern and southern hemispheres.

2.6 Ionospheric Irregularities

In order for SuperDARN to measure the plasma convection pattern, SuperDARN radars need a target from which radio waves can be scattered, that is inhomogeneities or irregularities embedded in the bulk plasma in F region ionosphere. SuperDARN radars receive scattered signals from decameter-scale irregularities. In general, the source of ionospheric irregularities are plasma waves and instabilities. Instabilities are created when conditions cause waves to grow and become unstable, like ocean waves growing until they become unstable and crash on a beach. The primary plasma process believed to produce ionospheric irregularities from which SuperDARN receives scattered signals is the gradient drift instability (GDI).

The GDI is caused by a combination of the local density gradient, ∇n_0 , electric field,

\mathbf{E}_0 , and magnetic field, \mathbf{B}_0 , in the plasma. For a density gradient with an $\mathbf{E} \times \mathbf{B}$ drift, the trailing edge of the density gradient is unstable to the GDI. Figure 2.8 shows the geometry that leads to the GDI, with a sinusoidal perturbation in the trailing edge of the density gradient. The background electric field produces a separation of charge in the perturbed region, which produces a polarization electric field, \mathbf{E}' . The polarization electric field adds to the background field in some regions and reduces it in other regions, which causes the perturbation to grow.

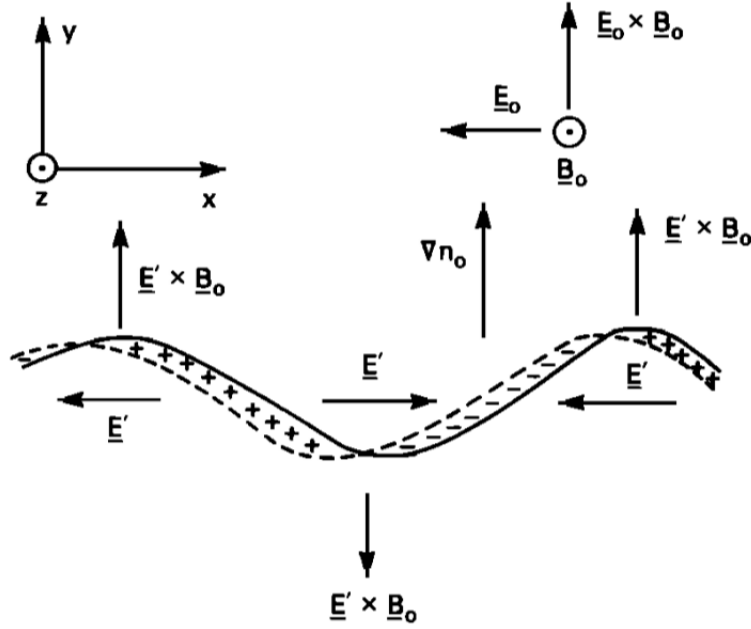


Figure 2.8: An illustration of the required conditions for GDI growth ($\gamma > 0$). Adapted from Figure 2 in *Keskinen and Ossakow* (1983).

The linear growth rate, γ , of the GDI in the F region of the ionosphere is given by

$$\gamma = \left[-\frac{\nu_{ei}}{\Omega_e} \frac{1}{L} \left(\frac{\nu_{in}}{\Omega_i} \frac{cE_0}{B} - \theta V_d \right) \right] / \left(\theta^2 + \frac{\nu_{in}}{\Omega_i} \frac{\nu_{ei}}{\Omega_e} \right) - D_{\perp} k_x^2 - D_{\parallel} k_y^2, \quad (2.22)$$

where ν_{ei} and ν_{in} are the electron-ion and ion-neutral collision frequencies, respectively, V_d is the reative electron and ion speeds along the magnetic field line, $\theta = k_z/k_x$ is the alignment (or aspect) angle of the GDI with respect to the magnetic field, $L^{-1} = (1/n_0)(\delta n_0/\delta y)$ where L is the scale length of the density gradient, and D_{\perp} and D_{\parallel} are the perpendicular and parallel plasma diffusion coefficients, respectively (*Keskinen and Ossakow*, 1983). The GDI growth rate reaches a maximum value for an alignment angle θ_m such that $\delta\gamma/\delta\theta = 0$,

$$\theta_m = \frac{\nu_{in}}{\Omega_i} \frac{cE_{0x}}{BV_d} \pm \sqrt{\left(\frac{\nu_{in}}{\Omega_i}\right)^2 \left(\frac{cE_{0x}}{BV_d}\right)^2 + \frac{\nu_{ei}\nu_{in}}{\Omega_e\Omega_i}}. \quad (2.23)$$

Using typical F region plasma parameters, *Keskinen and Ossakow* (1983) showed that $\theta_m \approx 0$ in the F region, meaning that the GDI growth rate is largest perpendicular to magnetic field lines. In the F region, the growth times for the GDI are $\gamma^{-1} \approx 10^2$ s (*Keskinen and Ossakow*, 1983). SuperDARN receives signals scattered from these irregularities in the F region produced by the GDI. Additionally, Equation 2.22 shows that the GDI is a convective instability. If $\mathbf{E}_0 = 0$, then there is no $\mathbf{E} \times \mathbf{B}$ drift and consequently no GDI. The faster the $\mathbf{E} \times \mathbf{B}$ drift, the larger the GDI growth rate.

2.7 Summary

The atmospheres of the Sun and the Earth are magnetically coupled by the IMF. The dynamics of solar activity drive dynamic activity in the magnetosphere of the Earth producing a variety of phenomena, including the Aurora Borealis and Aurora Australis. The ionizing radiation from the Sun produces the ionosphere of the Earth at an altitude above 85 km. As the solar wind propagates past the Earth it induces an electric field on the magnetosphere, driving plasma convection in both the magnetosphere and the ionosphere of the Earth. This electric field also drives field aligned currents which couple the magnetosphere and the ionosphere. Ionospheric irregularities, generated by instabilities like the GDI, are embedded in the F region plasma and propagate along with the bulk plasma flow. SuperDARN radars scatter radio waves off of the ionospheric irregularities to measure the bulk plasma flow. This allows SuperDARN to study the dynamic plasma convection in the magnetosphere by measuring the ionospheric plasma convection. In order for SuperDARN to measure the plasma convection it needs ionospheric targets from which to receive scattered signals. Ionospheric magnetic field-aligned irregularities produced by the GDI are the target from which SuperDARN receives scattered signals. In the next Chapter, the purpose and scientific role of SuperDARN in the context of geospace physics research will be discussed along with the radiophysics of radio wave propagation and scattering in the ionosphere. Some general radar principles and hardware used to operate SuperDARN radars will also be discussed. Concluding the chapter

will be a discussion of the standard data products produced by SuperDARN.

CHAPTER 3

THE SUPER DUAL AURORAL RADAR NETWORK

3.1 Introduction

The Super Dual Aurora Radar Network (SuperDARN) is an international network of ionospheric pulse-Doppler radars used to measure plasma flow in the F region ionosphere (*Greenwald et al.*, 1995; *Chisham et al.*, 2007). Specifically, the radars are used to measure the bulk $E \times B$ drift velocity of field-aligned plasma density irregularities in the F region. As discussed in section 2.5, F region plasma in the polar ionosphere is driven by the dawn-dusk electric field in the magnetosphere, therefore SuperDARN radars can measure the ionospheric footprint of magnetospheric plasma circulation. Figure 3.1 shows plots of the fields-of-view (FOV) of all the SuperDARN radars currently operating in the Northern and Southern hemispheres. As can be seen in the figure, SuperDARN radars have large FOVs that extend to at least 3,500 km range and approximately 52 degrees in azimuth, covering an area spanning several million square kilometers.

SuperDARN consists of 35 radars operating in both the Northern and Southern hemispheres, which is accomplished by an international collaboration consisting of 17 institutions in 10 countries. Tables 3.1 and 3.2 list every past and present SuperDARN radar located in the Northern and Southern hemispheres respectively. The institutes responsible for each radar are also listed in the tables. Two radars that are no longer operational, the radar at Schefferville in the Northern hemisphere and the radar at the Falkland Islands¹ in the Southern hemisphere, are included in the tables. In data analysis software, it is common to refer to each radar by a unique three letter code, which is also listed in the tables. Future

¹Which may become operational again.

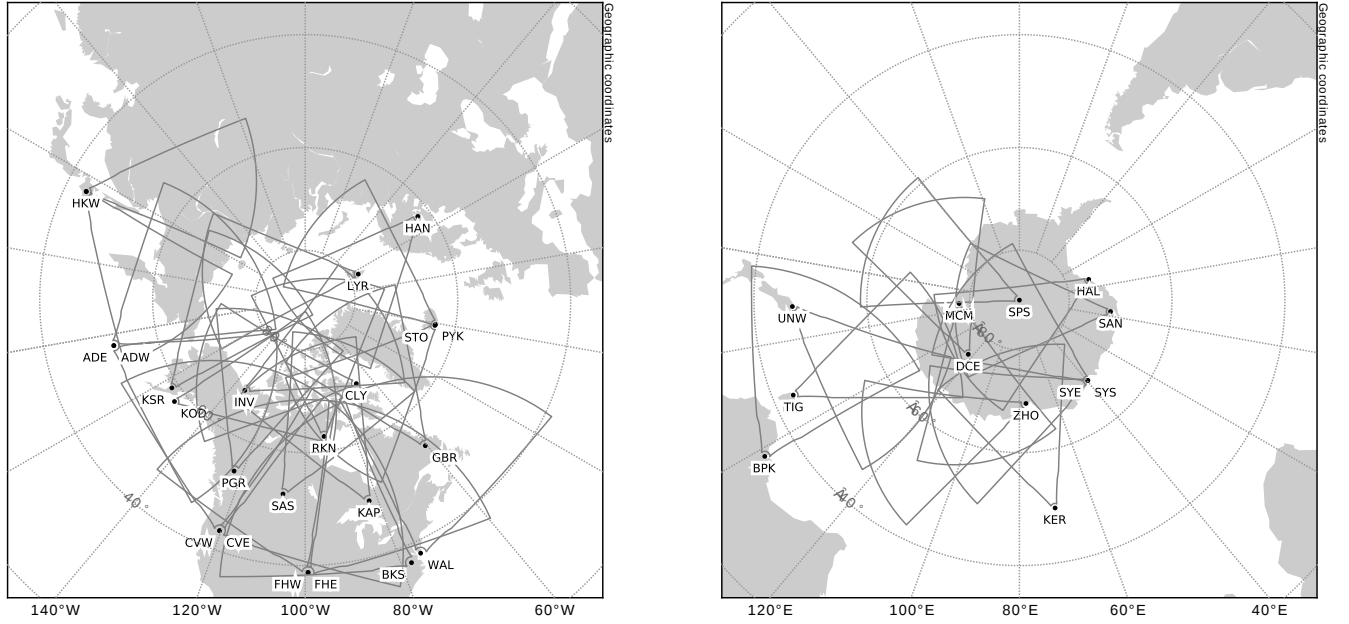


Figure 3.1: The FOVs of all active and inactive SuperDARN radars in the Northern (left) and Southern (right) hemispheres. The FOVs here were modelled starting at 180 km from the radar and using 75 45 km range gates.

radars are not listed, but at the time of this writing, there is a new SuperDARN radar being constructed in Southern France/Western Europe².

From the beginning, SuperDARN was designed to be an instrument that could be used to study the magnetosphere [George Sofko, private communication, 2012]. Previous ionospheric measurements by other radar experiments, such as the Scandinavian Twin Auroral Radar Experiment (STARE) (*Greenwald et al.*, 1978), observed high-latitude F region ionospheric plasma flow consistent with that expected due to magnetospheric electrodynamics (e.g. *Ecklund et al.*, 1975; *Greenwald et al.*, 1978, 1983). Compared with other instruments, such as satellites, SuperDARN is able to continuously sample large areas of the ionosphere, sampling the electrodynamics of huge volumes of the magnetosphere. In contrast, satellite measurements are small-scale and susceptible to spatial-temporal ambiguities. SuperDARN radars are also inexpensive to build and operate (less than one million CAD each). This means that SuperDARN is well-suited to provide a complementary global-scale data set to compare

²The precise location of this radar is still being determined.

with satellite measurements. Indeed, SuperDARN collaborates with several ongoing satellite missions, such as Time History of Events and Macroscale Interactions during Substorms (THEMIS) mission (*Angelopoulos, 2008*), the Radiation Belt Storm Probes (RBSP) mission (*Stratton and Fox, 2012*), and the Exploration of energization and Radiation in Geospace (ERG) mission (*Miyoshi et al., 2013*). SuperDARN radars are also used to study ionospheric processes and magnetosphere-ionosphere coupling phenomena, with similar FOV and cost advantages compared to other ionospheric instruments.

A wide variety of ionospheric and magnetospheric physics has been and continues to be studied using SuperDARN. A short list of topics follows: Magnetospheric reconnection rates and evolution of the Dungey-Cycle (*Hubert et al., 2006; Milan et al., 2007, 2012*), determination of the open-closed field-line boundary (*Wild et al., 2004*), the location of the auroral oval (*Imber et al., 2013*), flux transfer events (*Fear et al., 2009*), sub-auroral polarization streams (*Koustov et al., 2008; Clausen et al., 2012*), ultra-low frequency magnetohydrodynamic waves in the ionosphere (*Ponomarenko et al., 2003, 2005; Yeoman et al., 2012; Teramoto et al., 2016*), meteor echoes and mesospheric winds (*Hall et al., 1997; André et al., 1998; Jenkins and Jarvis, 1999; Hussey et al., 2000; Chisham and Freeman, 2013*), polar mesospheric summer echoes (*Hosokawa et al., 2005; Ogawa et al., 2013*), mapping field-aligned currents (*Sofko et al., 1995; Cousins et al., 2015*), ion-frictional heating (*Bjoland et al., 2015*), and studies of polar cap patches (*Oksavik et al., 2006; Dahlgren et al., 2012*). SuperDARN has proven to be an invaluable ionospheric and magnetospheric research instrument.

3.2 SuperDARN Radar Principles

SuperDARN radars are pulse-Doppler radar systems that determine the range and Doppler characteristics of a target by transmitting a pulse-sequence and recording the signal that returns. SuperDARN radars receive echoes from an illuminated volume of the ionosphere, referred to as a range cell, which generally extends 45 km in range with a 3 dB beam width of 3.24 degrees at 15 MHz. SuperDARN radars are monostatic phased-array radars with a main array and interferometer array. Figure 3.2 is a picture of the Saskatoon SuperDARN

Radar Name	Institution	Latitude	Longitude	Code
Adak Island East	University of Alaska, Fairbanks	51.88	-176.62	ade
Adak Island West	University of Alaska, Fairbanks	51.88	-176.62	adw
Blackstone	Virginia Tech	37.10	-77.95	bks
Christmas Valley East	Dartmouth College	43.27	-120.36	cve
Christmas Valley West	Dartmouth College	43.27	-120.36	cvw
Clyde River	University of Saskatchewan	70.49	-68.50	cly
Fort Hays East	Virginia Tech	38.86	-99.39	fhe
Fort Hays West	Virginia Tech	38.86	-99.39	fhw
Goose Bay	Virginia Tech	53.32	-60.46	gbr
Hankasalmi	University of Leicester	62.32	26.61	han
Hokkaido East	Nagoya University	43.53	143.61	hok
Hokkaido West	Nagoya University	43.54	143.60	
Inuvik	University of Saskatchewan	68.414	-133.772	inv
Kapuskasing	Virginia Tech	49.39	-82.32	kap
King Salmon	National Institute of Information and Communications Technology	58.68	-156.65	ksr
Kodiak	University of Alaska, Fairbanks	57.62	-152.19	kod
Longyearbyen	University Centre in Svalbard	78.153	16.074	lyr
Pykkvibaer	University of Leicester	63.77	-20.54	pyk
Prince George	University of Saskatchewan	53.98	-122.59	pgr
Rankin Inlet	University of Saskatchewan	62.82	-93.11	rkn
Saskatoon	University of Saskatchewan	52.16	-106.53	sas
Schefferville	CNRS/LPCE	54.80	-66.80	sch
Stokkseyri	Lancaster University	63.86	-22.02	sto
Wallops Island	JHU Applied Physics Laboratory	37.93	-75.47	wal

Table 3.1: Northern Hemisphere SuperDARN radars as of September 2017. Schefferville is no longer operational.

Radar Name	Institution	Latitude	Longitude	Code
Buckland Park	La Trobe University	-34.620	138.460	bpk
Dome C	Institute for Space Astrophysics and Planetology	-75.090	123.350	dce
Falkland Islands	British Antarctic Survey	-51.83	-58.98	fir
Halley	British Antarctic Survey	-75.52	-26.63	hal
Kerguelen	CNRS/LPCE	-49.22	70.14	ker
McMurdo	University of Alaska, Fairbanks	-77.88	166.73	mcm
SANAE	South African National Space Agency	-71.68	-2.85	san
South Pole Station	University of Alaska, Fairbanks	-89.995	118.291	sps
Syowa East	National Institute of Polar Re- search	-69.00	39.58	sye
Syowa South	National Institute of Polar Re- search	-69.00	39.58	sys
Tiger	La Trobe University	-43.40	147.20	tig
Unwin	La Trobe University	-46.51	168.38	unw
Zhongshan	Polar Research Institute of China	-69.38	76.38	zho

Table 3.2: Southern Hemisphere SuperDARN radars as of September 2017. Falkland Islands is no longer operational.

radar during the 20 December 2016 geomagnetic storm³. In the picture, both the main (background) and interferometer (foreground) arrays of the radar are visible, along with the control hardware shed (between the arrays). In the photograph, only 10 of the 16 main array antennas are visible.



Figure 3.2: The Saskatoon SuperDARN radar during the 20 December 2016 geomagnetic storm. All 4 antennas of the interferometer array can be seen in the foreground and 10 of the main array antennas can be seen in the background, along with the shed housing the transmitters, receiver, and control systems.

Each of the 16 antennas in the main array are fed by a transmitter consisting of a direct digital synthesis (DDS)–based signal generator system and a transmitter box. The DDS system generates the signal to be transmitted, including the frequency, phase offsets needed for beam forming, and pulse shape. The transmitter box contains RF power amplifiers and both the high and low power transmit/receive switches necessary to isolate the receive chain from the higher power output while transmitting. Each transmitter box amplifies the

³Not pictured, the author, who took this photo, digging out of a snow bank that he got stuck in when he parked his car before taking this picture.

milliwatt DDS signal up to 600 watts. The receiver chain uses DDS to generate the phasing necessary for beam-forming and before the signals from each antenna are combined, filtered, and digitized. The resulting digital signal is a stream of complex voltage samples that are ready for further signal processing. All of the RF electronics and computers necessary for operating the radar are located inside the control hardware shed.

3.2.1 Radar Principles

This section will discuss some general radar principles used by SuperDARN radars. The most well-known equation in radar is the radar equation. In general, the power received (P_r) from a radar target can be described by the radar equation, given by

$$P_r = \frac{P_t G}{4\pi R_t^2} \frac{\sigma A_e}{4\pi R_r^2}, \quad (3.1)$$

where P_t is the transmitted power emitted by the antenna, G is the gain of the transmitting antenna, σ is the radar cross section of (a measure of the power scattered by) the target, and A_e is the effective area of the receiving antenna (e.g. *Skolnik*, 1980; *Ulabiy et al.*, 1982; *Levanon and Mozeson*, 1988). The factors of $4\pi R_{r,t}^2$ are known as the spreading losses caused by the spreading of the transmitted and scattered (received) powers as they propagate isotropically in a spherical pattern. In the general case, the distance from the transmitter to the target R_t and the distance from the target to the receiver R_r are not the same. When the positions of the transmitter and receiver are not co-located the radar system is described as bistatic. When the positions of the transmitter and receiver are co-located the radar system is described as monostatic. In the general case of multiple transmitters and receivers, the radar system is multistatic. Even though SuperDARN radars are monostatic, they often share overlapping FOVs with other SuperDARN radars, so it may be possible to perform bistatic or even multistatic measurements with SuperDARN.

The simplest example of a monostatic pulse-Doppler radar system involves transmitting a pulse and listening for an echo, where an echo is a received signal resulting from the scattering or reflection of the transmitted pulse by a target. Using the time delay between transmission and reception the range of the target can be determined. If the pulse propagates at the speed of light c then the range R of the target can be determined from the time delay between

transmission and reception of the signal T using the relationship

$$R = \frac{cT}{2}. \quad (3.2)$$

The simplest way to obtain velocity information about the target is to transmit a series of uniformly time-spaced pulses and then record the echoes from each pulse. The Doppler shift of the target is determined by comparing the frequency of the received signal from each pulse with the frequency of the signal that was transmitted (e.g. *Levanon and Mozeson*, 1988; *Ulabiy et al.*, 1982). For monostatic radars, the Doppler shift, f_D , due to a target moving relative to the radar with a speed v_{los} toward or away from the radar is given by

$$f_D = \frac{2v_{los}}{\lambda_T}, \quad (3.3)$$

where λ_T is the wavelength of the transmitted pulses. Essentially, the Doppler frequency is the rate of change of the phase of the received signal produced by the change of distance in time (speed) between the radar and the target. Note that in general, any change of phase with respect to time in the received signal will produce an effective Doppler shift (this will be discussed in more detail in section 3.3). Finally, an additional Doppler shift is produced if the target is accelerating toward or away from the radar. For SuperDARN radars, it is assumed that targets are moving with uniform velocity, so the Doppler shift due to target acceleration is not considered.

3.2.2 Pulse Sequence

SuperDARN radars can receive backscatter from ranges 180 km⁴ up to 4500 km and measure ionospheric plasma flowing at speeds on the order of 1 km/s. Simultaneously measuring the range and Doppler of such signals using traditional radar methods results in range–Doppler ambiguities (*Levanon and Mozeson*, 2004; *Ulabiy et al.*, 1982). For example, a standard technique is to transmit a series of pulses with uniform spacing, however, for fast moving targets at long ranges this technique produces ambiguities for reasons that will be discussed in more detail in Chapter 4. Essentially, each pulse produces a noise-like interference called

⁴Actually, the radars can measure at distances as close as a pulse width away, but 180 km is standard.

self-clutter that makes it impossible to determine which ranges the received signals from other pulses originated. To overcome the range-Doppler ambiguities, SuperDARN radars transmit a non-uniform pulse-sequence where pulses are non-redundantly spaced according to the multiple-pulse technique described in *Farley* (1972). The multiple-pulse technique reduces the range-Doppler ambiguities, but even so self-clutter is present, albeit reduced.

The top plot of Figure 3.3 is a range-time diagram of an example 3 pulse multiple-pulse sequence. Pulse spacings are integer multiples of a fundamental pulse repetition time (PRT), more commonly referred to as the lag time (τ) in the SuperDARN community. Receiver voltage samples that are spaced the same as the pulses in the pulse sequence are used to estimate the spectral properties of the scattered signal. Signal processing techniques are discussed in more detail in Chapter 4.

A multiple-pulse sequence is characterized by four parameters: the pulse length (t_p), the lag time (τ), the number of pulses (N_p), and the pulse table (ptab). The pulse table is an array of integers describing the pulse spacing in multiples of τ . The example multiple-pulse sequence in the top plot of Figure 3.3 has parameters $\tau = 2.4$ ms, $t_p = 300$ μ s, $N_p = 3$, and ptab = [0,1,3]. The grey filled rectangles indicate transmitted pulses and the rectangles with blue borders indicate receiver samples that can be used to estimate the spectrum of the backscattered signal. The red diamonds indicate signal backscattered from the range of interest, and the black diamonds indicate signals originating from unwanted ranges. The bottom plot of Figure 3.3 shows the 8 pulse “katscan” multiple-pulse sequence that is typically used by SuperDARN radars. The katscan pulse sequence has parameters $\tau = 1.8$ ms, $t_p = 300$ μ s, $N_p = 8$, and ptab = [0,14,22,24,27,31,42,43].

While a monostatic radar is transmitting, it cannot receive signals (data). This means that for some ranges there are “bad” voltage samples that correspond to times when the receiver was blanked while the transmitter was on. In practice, usually 2 or 3 voltage samples are blanked due to the finite turn on/off time for the transmitters and the high/low power transmit/receive switches.

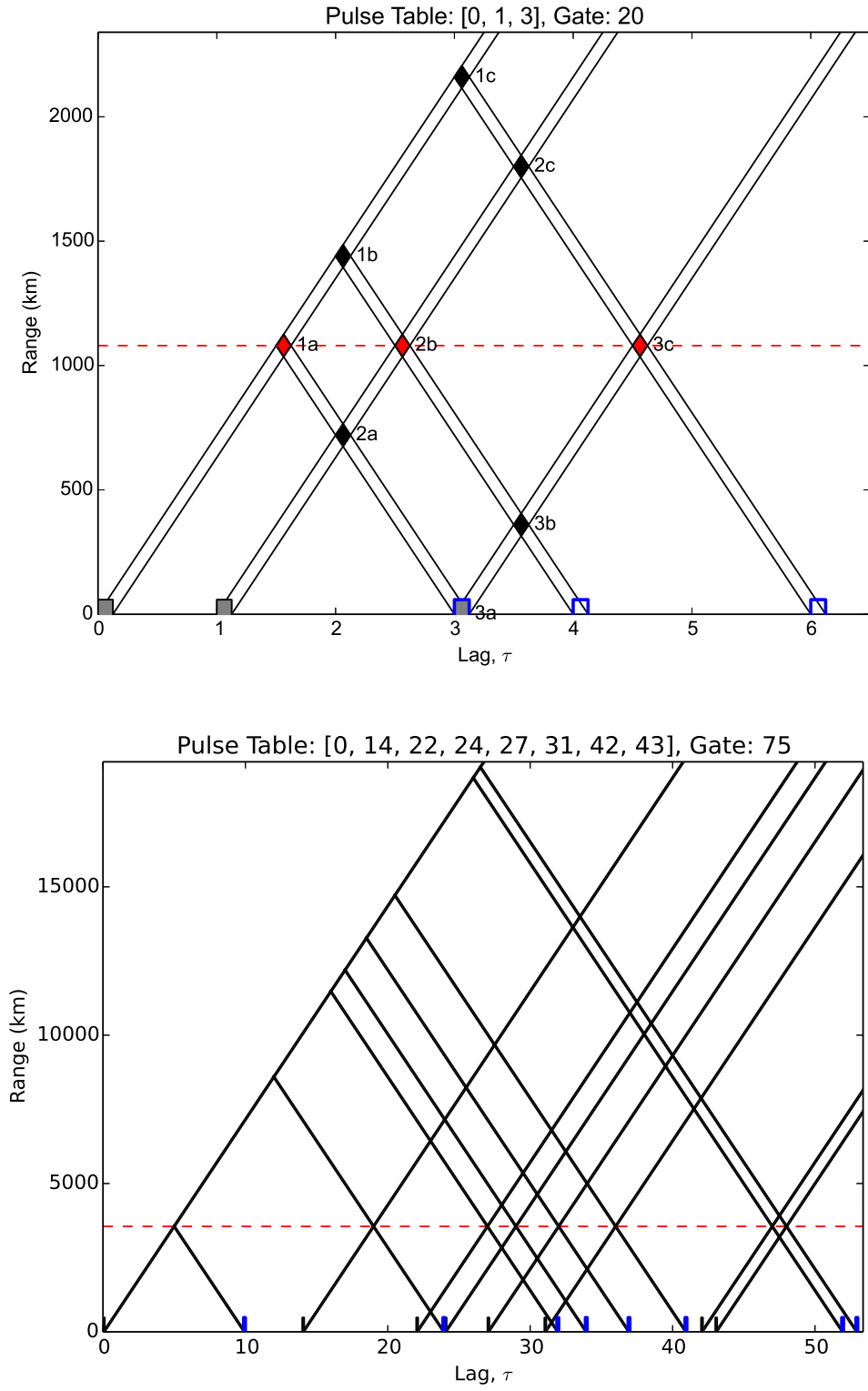


Figure 3.3: An example 3 pulse multiple-pulse sequence (top panel) and the 8 pulse katscan multiple-pulse sequence used by SuperDARN radars (bottom panel). In the bottom panel, the horizontal dashed line indicates the location of range gate 75 at a distance of 3500 km.

3.2.3 Doppler Shift From Quadrature

SuperDARN radars use a quadrature receiver to obtain complex voltage samples. This receiver performs a quadrature demodulation on the received signals, which involves mixing the received signal with sine and cosine waves and then low pass filtering the resulting signals. The signal resulting from the cosine mixing is called the “in-phase” component and the signal resulting from the sine mixing is called the “quadrature” component. Mathematically, a quadrature receiver operates on a signal received by the radar, $s(t)$, which can be represented by

$$s(t) = A(t) \cos([\omega + \omega_D]t), \quad (3.4)$$

where $A(t)$ is a time dependent amplitude of the signal, ω is the carrier frequency, and ω_D is the Doppler shift. To obtain the in-phase component, $s_I(t)$, $s(t)$ is mixed with a cosine signal, $\cos(\omega t)$, which is equivalent to a multiplication

$$s_I(t) = s(t) \cos(\omega t) = A(t) \cos([\omega + \omega_D]t) \cos(\omega t), \quad (3.5)$$

which is equivalent to⁵

$$s_I(t) = \frac{1}{2}A(t)[\cos([2\omega + \omega_D]t) + \cos(\omega_D t)]. \quad (3.6)$$

After low pass filtering to remove the cosine with a frequency of 2ω , the in-phase component is recovered:

$$s_I(t) = \frac{1}{2}A(t) \cos(\omega_D t). \quad (3.7)$$

Similarly, mixing a sine wave⁶, $\sin(\omega t)$, with $s(t)$ and then low pass filtering yields the quadrature component, $s_Q(t)$,

$$s_Q(t) = \frac{1}{2}A(t) \sin(\omega_D t). \quad (3.8)$$

Finally, the received complex signal

$$s_R(t) = s_I(t) + js_Q(t) = \frac{1}{2}A(t)[\cos(\omega_D t) + j \sin(\omega_D t)] \quad (3.9)$$

⁵Use the product trigonometric identity: $\cos \alpha \cos \beta = \frac{1}{2}[\cos(\alpha + \beta) + \cos(\alpha - \beta)]$.

⁶Use the product trigonometric identity: $\cos \alpha \sin \beta = \frac{1}{2}[\sin(\alpha + \beta) + \sin(\beta - \alpha)]$.

can be represented using Euler’s Identity to obtain

$$s_R(t) = \frac{1}{2}A(t)e^{j\omega_D t}, \quad (3.10)$$

which reveals that all of the Doppler shift information is contained in the phase of the received signals.

3.2.4 Phased Antenna Arrays

Antennas are transducers used to convert electrical signals into radio waves and vice versa. Antenna theory and design is a vast and complicated topic, beyond the scope of this thesis, but some basic concepts will be discussed here, as they are relevant to understanding how SuperDARN radars operate. Only the far-field descriptions of the radiation patterns of antennas will be presented, where the distance from the antenna R is much larger than the radar wavelength ($R \gg \lambda$).

There are several types of antennas, all with a variety of advantages and applications, but one of the primary objectives of these designs is to direct the radiated (and received) energy in some particular desirable direction or directions while suppressing it in other directions (Skolnik, 1980). The directive gain of an antenna is a measure of this enhancement and suppression and is defined as the ratio of the radiation intensity $I(\theta, \phi)$ in a particular azimuth ϕ , and elevation θ , to its mean radiation intensity \bar{I} expressed by

$$D(\theta, \phi) = \frac{I(\theta, \phi)}{\bar{I}}. \quad (3.11)$$

The directive gain is typically referred to as the “radiation pattern” of an antenna. An antenna is typically characterized by the maximum of its radiation pattern, where the directivity given by Equation 3.11 is largest. For radar, it is ideal that the radiation pattern would have maximum directivity in a beam pointed towards a target of interest and zero directivity in every other direction. In reality, however, antennas always have some directivity in unwanted directions.

SuperDARN radars use directive antennas in order to more efficiently radiate power in a preferred direction in the ionosphere. For nearly two decades, SuperDARN radars were constructed with log-periodic antennas, but as of 2010 newly constructed radars began using

the twin-terminated folded-dipole (TTFD) antenna design described in *Sterne et al.* (2011). The change was made because of the reduced cost and simplified construction process when building a SuperDARN radar using TTFD antennas compared to one using log-periodic antennas.

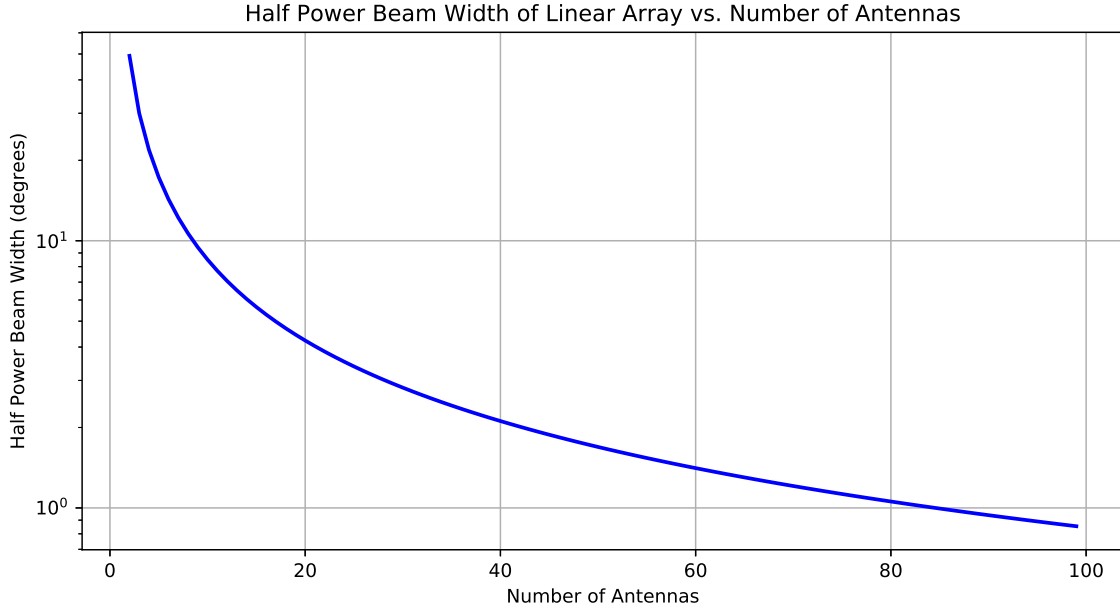


Figure 3.4: The half power beam width of a linear antenna array consisting of N equally spaced antennas.

One method for improving the directivity of an antenna is to combine several antennas together in an antenna array. This is the configuration that SuperDARN radars use, where 16 antennas are used to form a linear antenna array. To understand how this works, consider a linear array of N antennas with isotropic directivity and a uniform spacing of d between antennas. The radiation pattern of the array D_{array} as a function of azimuth angle ϕ is given by

$$D_{array} = \frac{\sin(N\gamma/2)}{\sin(\gamma/2)}, \quad \gamma = \frac{2\pi d}{\lambda} (\sin(\phi) - \sin(\alpha)) \quad (3.12)$$

where α is the desired angle towards which the main beam of the pattern is to be steered (*Skolnik*, 1980). As the number of antennas increases, the beam width decreases and the

directivity of the radiation pattern of the array improves. This is demonstrated in Figure 3.4, where the half power beam width (HPBW) for an array of N antennas with $d/\lambda = 0.6$ is plotted. *Skolnik* (1980) provides an approximate analytic formula for the beam width, given by

$$\theta_{HPBW} = \frac{0.886\lambda}{Nd}. \quad (3.13)$$

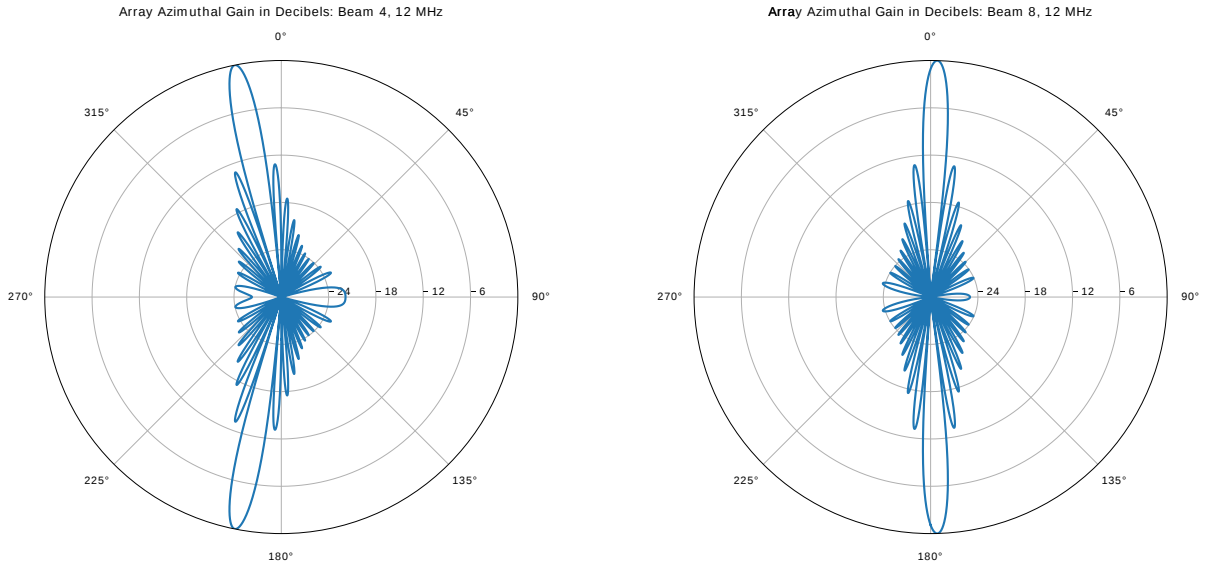


Figure 3.5: The azimuthal gain pattern for a phased antenna array consisting of 16 antennas spaced 15 meters apart for two different SuperDARN beam directions. See text for details.

For radar or communication systems using large parabolic or spherical dish antennas, the entire antenna must be moved to steer the beam of an antenna in a desired direction. A SuperDARN radar array uses 16 antennas spaced by approximately 15 meters, resulting in an array on the order of 215 meters across. Rotating such a large array to steer the beam would be at the very least, impractical. A significant advantage of the antenna array design is that the beam can be electronically steered, without any physical movement. This is accomplished by applying a linearly increasing phase offset of $2\pi d \sin(\alpha)/\lambda$ to the signal sent to each antenna across the array. Note that the required phase offset is a function of the wavelength and the antenna spacing. This means that the phase offsets required to steer the radar beam for one frequency are not the same as for another frequency. Another advantage

of this design is that the beam steering is inertia-less and effectively instantaneous, practically limited by the speed of the electronics used.

Figure 3.5 shows plots of the steered beam radiation pattern similar to that which the Saskatoon SuperDARN radar produces. The figure was made using Equation 3.12 with an antenna spacing of 15 meters and a radar wavelength of 24.98 meters (a radar frequency of 12 MHz). In the left plot of Figure 3.5 a phase offset with a slope of -0.741 radians was used to steer the beam counter clockwise by 11.34 degrees and in the right plot of Figure 3.5 a phase offset with a slope of 0.107 radians was used to steer the beam clockwise by 1.62 degrees. Note that this pattern is the array radiation pattern assuming isotropic antennas and that the array pattern has both a forward and backward facing beam. In practice, the antennas used in a SuperDARN array are directional, producing a predominately forward facing radiation pattern such that the back facing beam in the array pattern is significantly attenuated. For a more accurate representation of the full radiation pattern of a SuperDARN radar, one needs to simulate it (for both the log-periodic and TTFD designs), as performed by *Sterne* (2010) and *Sterne et al.* (2011).

3.2.5 Interferometry

SuperDARN radars consist of two linear arrays of antennas, a main array of 16 antennas and an interferometer array of 4 antennas⁷. The characteristics and advantages of the main array have already been discussed in section 3.2.4. The interferometer array is used, along with the main array, to measure the angle of arrival of signals received by a SuperDARN radar using interferometry. Essentially, this is accomplished using the difference between the time of arrival of a signal at the main array and time of arrival of a signal at the interferometer array (e.g. *McDonald et al.*, 2013; *Burrell et al.*, 2016). Elevation angle measurements are an important tool for SuperDARN users since such measurements can be used to geolocate received signals and to characterize from which part of the ionosphere the received signals originated (e.g. *Yeoman et al.*, 2008; *Burrell et al.*, 2015, 2016).

For a radar with a main array and an interferometer array, where the arrays are at the

⁷The Goose Bay radar is unique in that it has 16 interferometer antennas.

same altitude, centered, and separated by some distance D , the elevation angle Δ of a received signal is given by

$$\Delta = \arcsin \sqrt{\cos^2 \alpha - \left(\frac{\Phi \lambda}{2\pi D} \right)^2}, \quad (3.14)$$

where α is the angle between the radar beam and normal to the main array, λ is the wavelength of the received signal, and Φ is a phase offset produced by the received signal arriving to the main and interferometer arrays at different times (*Burrell et al.*, 2016). The phase offset Φ is measured by the SuperDARN radar and then the elevation angle Δ is calculated using Equation 3.14.

It is important to note that not all elevation angles can be unambiguously measured using the two array design that SuperDARN radars use. This is due to elevation angle aliasing that occurs due to the spacing between the main and interferometer arrays relative to the wavelength of the received signal. These ambiguities can be resolved by the addition of another array, which has been discussed in detail in *McDonald et al.* (2013).

3.3 Radiophysics

SuperDARN radars operate in the high-frequency (HF) band and are designed to transmit and receive on frequencies from 8 to 20 MHz, which corresponds to vacuum wavelengths of 34.47 to 14.99 meters (e.g. *Greenwald et al.*, 1995). Due to interactions with the ionosphere, properties of the HF radio waves are modified as they propagate through the ionosphere. Some of these interactions result in scattering, refraction, dispersion, and absorption of the propagating waves.

3.3.1 Coherent Scatter

SuperDARN radars receive signals that have been scattered back to the radar by plasma density irregularity structures in the ionosphere. These structures may be generated by plasma instabilities (recall section 2.6), which produce magnetic field-aligned quasi-periodic electron density structures such as the GDI. These structures persist throughout an extended volume of ionosphere, filling the volume illuminated by the radar beam (*Schlegel*, 1996). Scatter

from these field-aligned irregularities is aspect sensitive, such that a probing radio wave is scattered most strongly only where the k-vector (wave vector) of the wave is orthogonal to the magnetic field (e.g. *Krall and Trivelpiece*, 1973; *Schlegel*, 1996; *Ulabay et al.*, 1982). An important criteria for SuperDARN scatter is the Bragg condition

$$\lambda_{rad} = 2\lambda_{irr} \sin\left(\frac{\theta}{2}\right), \quad (3.15)$$

where λ_{rad} is the radar transmitted wavelength in the medium, λ_{irr} is the wavelength of the irregularities, and θ is the scattering angle relative to the incident radio wave in the plane perpendicular to the magnetic field. When this condition is satisfied, a significant gain of scattered signal is observed that is reminiscent of resonance phenomena, which is why the Bragg condition is sometimes referred to as Bragg Resonance (*Ulabay et al.*, 1982). As the radio waves scattered by each field-aligned density structure add coherently when the Bragg condition is satisfied, this type of scatter is referred to as coherent scatter. For monostatic radars like SuperDARN, direct backscatter results (e.g. $\theta = 180^\circ$), so the Bragg condition simplifies to

$$\lambda_{irr} = \frac{\lambda_{rad}}{2}, \quad (3.16)$$

which states that for backscatter, coherent scatter radars observe electron density irregularity structures with a wavelength that is half the radar wavelength (approximately 7.5 to 15 meters for SuperDARN radars).

SuperDARN radars are unable to make ionospheric measurements without ionospheric irregularities. It is important to remember that density gradients and strong electric fields are required for irregularities to be generated (see Section 2.6). Even when large electric fields exist, if the conductance in the E region is large relative to the F region conductance, the E region may dissipate the electric fields resulting in a reduction of F region irregularities (*Milan et al.*, 1999). The effect can be significant enough that no F region signals are observed by the radar. It is also possible for the plasma density to be reduced by enhanced recombination rates, such as when large electric fields frictionally heat the ions in the plasma (e.g. *St.-Maurice and Hanson*, 1982; *Davies et al.*, 1995). This means that in addition to satisfying

the aspect conditions of equation 3.15, SuperDARN radars require ionospheric conditions that favour irregularity generation for the coherent scattering process to occur.

3.3.2 Refraction

As radio waves propagate through the ionosphere they interact with the magnetized plasma and are modified according to magneto-ionic theory (*Budden, 1961; Hargreaves, 1995a*). This interaction can be succinctly described using the Appleton-Hartree equation, which describes the complex index of refraction of a magnetized plasma n given by

$$n^2 = 1 - \frac{X}{1 - jZ - \left(\frac{Y_T^2}{2(1-X-jZ)} \right) \pm \left(\frac{Y_T^4}{4(1-X-jZ)^2} + Y_L^2 \right)^{\frac{1}{2}}}, \quad (3.17)$$

where

$$X = \omega_p^2/\omega^2, \quad Y = \omega_B/\omega, \quad Z = \nu/\omega, \quad (3.18)$$

$$Y_L = \omega_B \cos \theta/\omega, \quad Y_T = \omega_B \sin \theta/\omega, \quad (3.19)$$

where ω is the angular frequency of the radio wave, ω_p is the electron plasma frequency, ω_B is the electron gyrofrequency, ν is the electron-neutral collision frequency and θ is the angle between the radio propagation direction and the magnetic field. The Appleton-Hartree equation is a general dispersion relation that describes the frequency dependence of propagation and absorption of radio waves.

At HF in both the E and F regions, $Z \approx 0$ since the electron-neutral collision frequency is small compared to the radio wave frequency. In ionospheric plasma, HF radio waves propagate in two regimes described by two special cases of Equation 3.17: quasi-longitudinal propagation (propagation approximately along the magnetic field, $Y_L > Y_T$) and quasi-transverse propagation (propagation nearly perpendicular to the magnetic field, $Y_T > Y_L$). For SuperDARN frequencies, quasi-longitudinal propagation dominates for $\theta \lesssim 75^\circ$ and $\theta \gtrsim 105^\circ$ (assuming the electron density is $5.0 \times 10^{11} m^{-3}$). The index of refraction for quasi-longitudinal propagation is

$$n^2 = 1 - \frac{X}{1 \pm Y_L}, \quad (3.20)$$

and for quasi-transverse propagation

$$n^2 = 1 - \frac{X(1-X)}{1-X-(Y_T^2 \pm Y_T^2)/2}. \quad (3.21)$$

For SuperDARN radars, the most important insight provided by the Appleton-Hartree equation is that the propagation of radio waves in the ionosphere is a dispersive phenomenon, depending on both the plasma density and the frequency of the radio wave. At HF, the index of refraction is less than one, meaning HF radio waves are refracted by the ionosphere as they propagate through it. As the HF waves refract, the propagation angle relative to the magnetic field changes such that radio waves launched from the ground refract to become perpendicular to the magnetic field at F region altitudes, satisfying the magnetic aspect angle requirement for coherent backscatter (equation 3.15).

Using a known electron density profile and the Appleton-Hartree equation, it is possible to model the propagation of SuperDARN radio waves. Figure 3.6 shows a plot of the results of ray tracing a 9 MHz radio wave from the Saskatoon SuperDARN radar through a model of the ionosphere produced by the 2011 International Reference Ionosphere model (*Bilitza et al.*, 2011). The ray tracing was performed using the “RayDARN” ray tracing software (*de Larquier et al.*, 2013) that is available on the Virginia Tech SuperDARN website (<http://vt.superdarn.org>). The ground range is plotted along the curved x-axis of the figure and the altitude above the ground is plotted along the y-axis. The background colour denotes the plasma density and the magenta nearly-vertical lines show the magnetic field. The grey lines are the ray paths of radio waves originating from the bottom left of the plot. White lines intersecting with the ray paths indicate constant range. The short black horizontal lines in the center of the figure indicate regions where the radio waves are within one degree of perpendicular to the magnetic field (regions where coherent backscatter is possible).

Figure 3.6 did not include an E region layer, but the effect of E region plasma density on the propagation path of HF radio waves has been previously studied (e.g. *Hussey et al.*, 1999; *Bahcivan et al.*, 2013). For example, when the plasma density in the E region is large enough, it may refract all incident HF waves that would otherwise reach the F region for coherent backscatter (*Milan et al.*, 1999). Also not shown in the figure is the possibility of HF

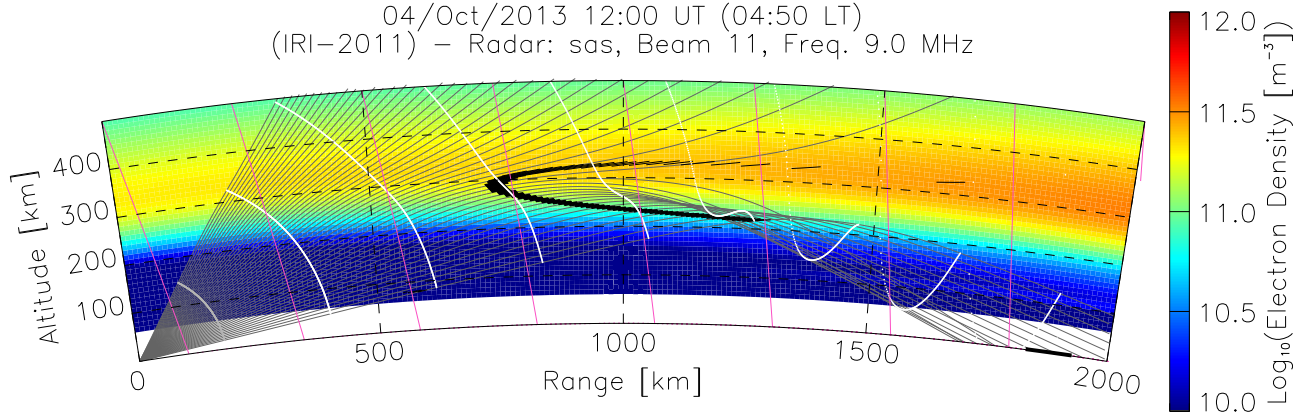


Figure 3.6: A simulation of the expected ray paths for the Saskatoon SuperDARN radar on beam 11 with a transmit frequency of 9 MHz. The grey lines indicate ray paths with white lines crossing the rays indicating the along ray distance in 200 km increments. The magenta lines indicate magnetic field lines. The thick black lines indicate regions where the rays are within one degree of aspect conditions required for coherent backscatter. The background colour indicates the electron density. The ray trace was performed using the Virginia Tech SuperDARN online Ray-Tracing Tool.

radio waves to be refracted towards the ground where they are reflected back upwards and continue towards the ionosphere where they may coherently backscatter. Using ionospheric refraction, it is possible for HF waves to achieve multiple “hops” and produce backscatter at large ranges from the radar.

Additionally, using Equation 3.17 one can see that changes to the electron density (which changes the plasma frequency) in the ionosphere results in changes to the index of refraction. If the density changes occur fast enough in time, as it does during auroral particle precipitation events, the changes to the index of refraction may produce induced Doppler shifts measured by SuperDARN (e.g. *Scoular et al.*, 2013). This is because the Doppler shift is measured by observing the change in phase with time of a received signal and changes to the index of refraction along the propagation path of a signal produce time varying changes to the path length phase. This results in an additional apparent Doppler shift that is indistinguishable from the Doppler shift due to motion of the target.

Absorption

When the plasma density in the D region becomes enhanced due to proton storms, solar flares, or high energy electron precipitation events, the increased electron density results in enhanced absorption of HF radio waves propagating through the D region. When this occurs, Z at HF frequencies is no longer negligible and the index of refraction becomes complex.

Figure 3.7 illustrates the effects of D region absorption during an X1.3 solar flare on 25 April, 2014. In the bottom panel of each plot, the X-ray power flux observed is plotted. The top panels show plots of the backscatter seen by the five SuperDARN Canada radars, where the colour scale denotes the signal to noise ratio of the backscatter measured in dB (blue is small, red is large). The black region at the right in each panel shows the day-night terminator. The left (right) panel of the figure shows the radar scatter seen before (during) the solar flare event. Note that before the solar flare, the radars were observing significantly more scatter and significantly stronger (higher signal to noise) scatter than was observed during the solar flare. The latter is due to the increase in X-ray flux producing enhanced D region electron density (where the electron-neutral collision frequency is large compared to the radio wave frequency, $Z > 0$), resulting in absorption of the HF radio waves transmitted by SuperDARN.

3.4 Standard Data Products

The main data product of SuperDARN is the global maps of F region ionospheric plasma convection in both the Northern and Southern hemispheres. Figure 3.8 shows a convection map made by combining data from 13 individual SuperDARN radars using a spherical harmonic regression procedure (*Ruohoniemi and Baker, 1998a*). Alternative data assimilation techniques to the SuperDARN FIT technique have also been proposed for generating convection maps (e.g. *Fiori et al., 2010a*; *Cousins et al., 2013*). Furthermore, SuperDARN data may also be ingested into assimilative models, as in *Cousins et al. (2015)* where data from SuperDARN and data from the Active Magnetosphere and Planetary Electrodynamics Response Experiment (AMPERE) were combined to produce maps of both the electrostatic and magnetic vector potentials.

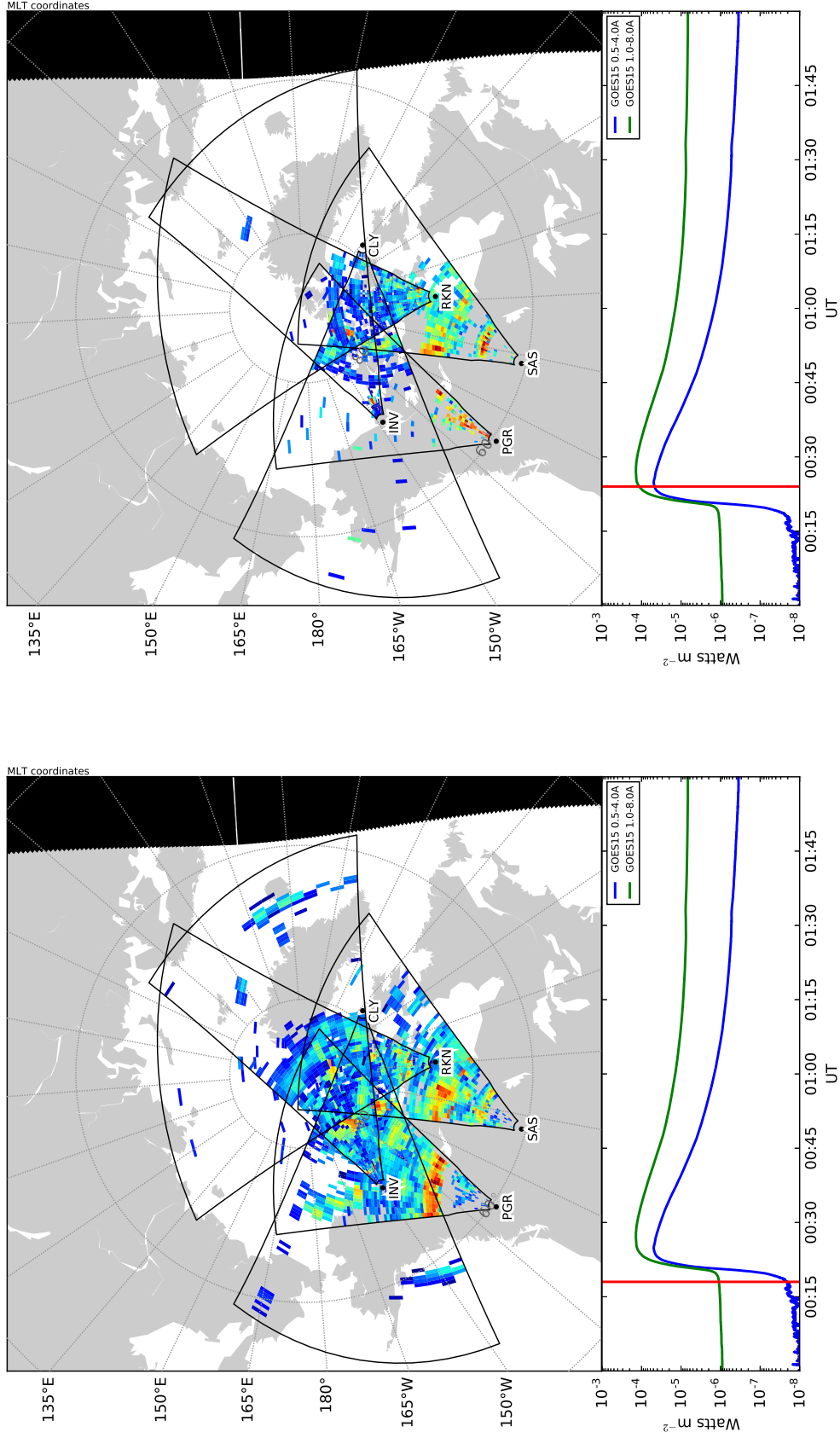


Figure 3.7: Data from the Canadian SuperDARN radars collected on 25 April 2014, both before (left) and during (right) an X1.3 solar flare from sunspot region 12035. The top of each plot shows the Canadian SuperDARN radars FOVs and backscatter, while the bottom of each plot indicates the X-ray flux incident on the Earth as measured by the GOES 15 satellite.

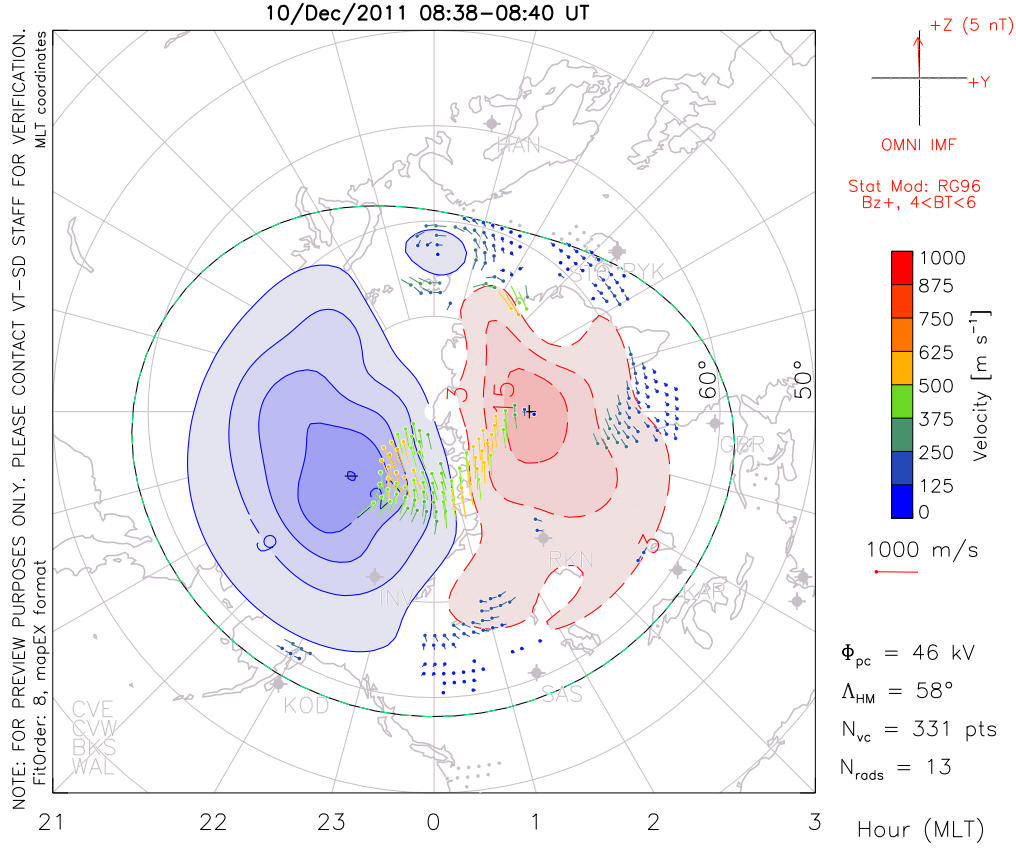


Figure 3.8: A SuperDARN convection map for the Northern hemisphere on 10 December 2011 from 8:38–8:40 UT. The contours indicate lines of equal electric potential. The vectors indicate the F region plasma velocity. This map was generated using Virginia Tech SuperDARN online Interactive Map Potential Plotting.

The primary data product of individual SuperDARN radars is the so-called line-of-sight (LOS) velocity. The LOS velocity is the component of the velocity vector of the plasma irregularities projected along the look direction of the radar beam. SuperDARN radars also measure the signal-to-noise ratio (SNR) and the spectral width of the received signals.

Due to the multiple hop propagation that is possible at HF, SuperDARN radars receive backscattered signals from the ground. Ground echoes typically have both low velocity and low spectral width, so a first-order ground-scatter identification procedure used by SuperDARN is to flag echoes with speeds v less than 30 m/s and spectral widths w less than 35 m/s. A more sophisticated probabilistic method developed by *Blanchard et al. (2009)* found

that the ground scatter identification criteria

$$v < 33.1\text{m/s} + 0.139w - (0.00133\text{s/m})w^2 \quad (3.22)$$

more effectively identifies ground scatter. Additionally, in the mid-latitude regions, ionospheric echoes frequently have both low velocity and low spectral width meaning that they are falsely identified as ground scatter. During geomagnetically quiet periods, mid-latitude SuperDARN scatter should be characterized as ground or ionospheric in origin using the algorithm developed by *Ribeiro et al.* (2011), which improves the number of correctly identified ionospheric scatter echoes by 50% compared to the first-order method.

3.4.1 A New Density Data Product

When an instrument-transmitted frequency is the same as the plasma frequency, the signal is reflected. This principle is used by ionosondes, which are scientific instruments used to measure the vertical plasma density profile of the ionosphere by sweeping through a range of transmission frequencies and recording that time delay between transmission and reception (e.g. *Chen et al.*, 2012). A SuperDARN radar can be thought of as a sort of oblique ionosonde (except coherent scattering occurs instead of reflection) such that using the principles of the Appleton-Hartree equation and multiple transmit frequencies, the electron density can be inferred (*Gillies et al.*, 2012). For example, in the scattering region, the index of refraction n_s for a SuperDARN transmitted signal is given by

$$n_s^2 = 1 - \frac{f_p^2}{f^2}, \quad (3.23)$$

where f_p is the plasma frequency in the scattering region and f is the radar transmitted frequency. The Doppler velocity measured by the radar v_m is influenced by the index of refraction such that

$$v_m = n_s v_s, \quad (3.24)$$

where v_s is the true Doppler velocity in the scattering region. Since the index of refraction in the ionosphere is generally less than one at SuperDARN frequencies, this means that in general, SuperDARN radars measure an underestimated velocity in the ionosphere. By

combining Equations 3.23 and 3.24 for two different transmitted frequencies, *Gillies et al.* (2012) determined that the plasma frequency in the scattering region could be determined using

$$f_p^2 = \frac{f_1^2(1 - v_1^2/v_2^2)}{(1 - (v_1^2 f_1^2)/(v_2^2 f_2^2))}, \quad (3.25)$$

where v_1 and v_2 are the velocities measured by the radar using frequencies f_1 and f_2 respectively. Using a superposed epoch analysis of transmit frequency shifting events, *Gillies et al.* (2012) was able to correct SuperDARN velocity measurements. This frequency shifting analysis utilized 14×10^6 events and was a statistical analysis.

From May 2012 to April 2013, a frequency shifting experiment utilizing the Kodiak SuperDARN radar in Alaska was designed and performed. I analyzed the data in an attempt to retrieve plasma density measurements using the SuperDARN radars. This was to be one of the original goals of this thesis, but it was not successful. Following a significant amount of work, it was unclear whether retrieving plasma density measurements from individual SuperDARN measurements would be possible. It was determined that the failure of the experiment was due to the experiment design, which assumed that the uncertainty in the measured SuperDARN velocities was more accurate than it was and, more significantly, no reliable error estimates were available. Additionally, investigating the SuperDARN signal processing code, several fundamental potential improvements to the SuperDARN signal processing methodology were identified. It became clear that in order to apply the frequency shifting technique to shorter temporal scales, a major reworking of how SuperDARN parameters are obtained and, more importantly, how they are constrained using accurate error estimates or error bars, was required.

The result of this initial research work was a refocusing of my PhD research to improve SuperDARN parameter determination and quantify and develop robust parameter error estimates. Due to the importance of SuperDARN observations to the space physics community (not just the frequency shifting technique) this research is critical. Global convection maps and global modelling efforts (e.g. *Gao, 2012*) are vital to the community who benefit significantly from improvements to SuperDARN data.

The next chapter provides a general introduction and overview of the signal processing techniques and statistical methods that were used to develop a methodology for processing

SuperDARN data. In subsequent chapters, more detailed discussions of the signal processing techniques that are used are provided. In Chapter 5 a new self-clutter estimation technique is presented, which is required to accurately account for the self-clutter contributed to SuperDARN measurements due to using the multiple-pulse technique. In Chapter 6 the statistical properties of SuperDARN measurements are presented, including how to use self-clutter estimates to accurately obtain error estimates required to accurately fit SuperDARN data. Chapter 7 presents a new fitting methodology that requires the use of the techniques developed in Chapters 5 and 6.

CHAPTER 4

RADAR SIGNAL PROCESSING: GENERAL CONCEPTS AND TECHNIQUES

To better quantify errors for measured SuperDARN data parameters, improvements to SuperDARN signal processing techniques were required. Therefore, the majority of the work performed for this dissertation was focused on signal processing. SuperDARN signal processing begins with the complex voltage samples produced by the quadrature receivers and ends with fitted data parameters, such as SNR, spectral width, and velocity, that researchers typically utilize while studying the geospace environment. To extract the fitted parameters, the voltage samples are processed using statistical methods, signal processing techniques, and models of the scattering process.

Probability theory is central to processing SuperDARN signals, which are received from volume scatter where a large number of randomly distributed scatterers contribute to the total received signal. Probability theory provides knowledge of the probability distribution or probability density function (PDF) of the received voltage samples. Using the PDF of the voltage samples, statistical estimates of the autocorrelation function (ACF) and its variance can be obtained. Finally, using signal processing techniques and a model of the ACF, fitted parameters and their associated errors can be obtained. This chapter will discuss general topics in probability theory and radar signal processing that are fundamental to the SuperDARN signal processing techniques and methods that will be discussed in greater detail in Chapters 5, 6, 7.

4.1 Probability and Random Variables

Probability theory provides a set of mathematical techniques for characterizing the uncertainty in radar measurements. In general, probability theory provides understanding for all types of scientific measurements and the uncertainty in those measurements. Statistical measurements and uncertainties are characterized as quantities called random variables, which have statistical properties such as a mean, variance, and PDF. Using Estimator theory, the statistical properties of quantities estimated using random variables can be determined. The mathematical techniques used to estimate these properties are called estimators. The concepts of probability seem to permeate much of radar signal processing with concepts such as ambiguity and information playing central roles in developing optimal measurement techniques and determining the range and Doppler resolution of the radar. In this section, some general topics related to random variables and estimators used in SuperDARN radar signal processing will be presented and discussed.

4.1.1 Probability Density and Statistics

For events with discrete outcomes, such as a coin toss, the probability of an event occurring can be thought of as a fraction of equally likely possibilities¹, half the time resulting in heads and half the time resulting in tails (e.g. *Woodward*, 1960). Such outcomes are said to be uniformly distributed. In the coin tossing example each individual outcome is a measure of a random variable. In general, probability is not uniform and some events are more likely to occur than others, in other words, the random variable will tend towards some outcomes more often than others. In this case, it is useful to describe the probability using a probability density function (PDF). Using continuous notation, a random variable x has a PDF $p(x)$ (e.g. *Westfall and Henning*, 2013). The probability of event A occurring can be written as

$$P(A) = \int_A p(x)dx, \quad (4.1)$$

¹Only ideally, of course. In reality, coins can land on their edge (*Murray and Teare*, 1993) and probability of heads or tails depends on how the coin is tossed (*Diaconis et al.*, 2007).

where A is the range of values of x that correspond to event A . The total area of a PDF is given by

$$\int_{-\infty}^{\infty} p(x)dx = 1, \quad (4.2)$$

which means that the probability of all events must be equal to 1. A PDF can be characterized by a number of important statistics, such as the mean and variance of the random variable.

The mean, μ , is given by

$$\mu = \int_{-\infty}^{\infty} xp(x)dx \quad (4.3)$$

and the variance, σ^2 , is given by

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 p(x)dx. \quad (4.4)$$

In general, the expectation value $E[f(x)]$ of a function $f(x)$ operating on a random variable x is given by

$$E[f(x)] = \int_{-\infty}^{\infty} f(x)p(x)dx. \quad (4.5)$$

This equation is sometimes referred to as the law of the unconscious statistician, since one may obtain the expectation value of $f(x)$ without knowledge of the PDF of $f(x)$ (*Westfall and Henning, 2013*). If $f(x) = x$, then Equation 4.5 reduces to Equation 4.3. Unless specified otherwise, the expectation value will refer to the mean of a random variable. In the context of radar experiments, the expectation value is the value of the random variable for which convergence is expected after taking a sufficiently large number of measurements. For example, for an ideal coin toss, a small number of coin tosses may result in more heads than tails, but the expectation value of obtaining “heads” after a large number of coin tosses is 50%.

To describe the probability of some events occurring may require two or more random variables. For example, the probability of drawing a red face card from a deck of cards depends on the probability of the colour of the card being red and the probability of the card being a Jack, Queen, or King. For these events, a multivariate PDF of random variables is used. For example, for random variables x and y , the PDF is written as $p(x, y)$. The marginal distribution can be useful to examine the PDF of one random variable of a multivariate PDF. For the multivariate distribution $p(x, y)$, the marginal distribution of x can be obtained by

integrating over the domain of y , such that

$$p(x) = \int_{-\infty}^{\infty} p(x, y) dy. \quad (4.6)$$

Conditional Probability

Conditional probability describes the probability of one event A given information that another event B has occurred. If the information about B does not modify the probability, or the probability of A does not depend on the probability of B , then the events are called independent. However, if the addition of information about one event changes the probability of another, then the probability of A given B , is given by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad (4.7)$$

where $P(A \cap B)$ is the probability of A occurring without B and $P(B)$ is the probability of B occurring (*Woodward, 1960*). It is important to note that in general $P(A|B)$ is not equal to $P(B|A)$. For example, an airport radar having a 90% chance of detecting a plane is the probability that a plane is detected given that there is a plane to detect. Here, event B is “there is a plane” and event A is “detecting a plane”, such that $P(A|B) = 90\%$. It may be that the radar falsely detects a plane even when there is in fact no plane there, so the probability of there being a plane when a plane is detected $P(B|A)$ is larger than 10%. Conditional probabilities can be properly reversed using Bayes’ theorem, given by

$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}. \quad (4.8)$$

4.1.2 Gaussian Random Variables

The type of random variable most fundamental to SuperDARN radar measurements is the Gaussian random variable. Gaussian random variables are characterized by the Gaussian (or normal) distribution. If x is a Gaussian random variable, with mean \bar{x} , and variance σ^2 , it has a PDF given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \bar{x})^2}{2\sigma^2}\right), \quad (4.9)$$

where $p(x)$ is usually denoted as $\mathcal{N}(\bar{x}, \sigma^2)$. As shown in Chapter 6, SuperDARN voltage samples are Gaussian random variables due to the fact that targets for SuperDARN are volume filling such that a large number of independent randomly distributed irregularities scatter the signal back to the radar. This is due to the Central Limit Theorem, which states that the PDF of a random variable that is the sum of a “large” number of independent and identically distributed random variables is given by a Gaussian distribution (e.g. *Ulaby et al.*, 1982; *Levanon and Mozeson*, 1988; *Westfall and Henning*, 2013). What is “large” enough depends on the PDF of the random variables being summed, but the resulting random variable becomes increasingly more Gaussian as the number of variables contributing to the sum increases. For SuperDARN voltage samples to be Gaussian distributed, the illuminated volume is required to be a “large” number of wavelengths in extent. Additionally, the summed phase contributions of the individual scatterers is uniformly distributed in $[0, 2\pi]$ because they are randomly spatially distributed (*Ulaby et al.*, 1982).

4.1.3 Statistical Estimation

Discriminating between relevant and irrelevant information in the signals received by SuperDARN radars is important. The voltage samples produced by the radar are a random sample of the plasma population. The characteristics of the plasma population has an unknown probability distribution that is a function of some plasma parameters. These parameters characterize the population and therefore the parameters also characterize the sample data. Using statistical estimation techniques and more specifically using estimator theory, an attempt to estimate the parameters of the plasma population from the sample of data can be made (e.g. *Fisher*, 1958; *Sampath*, 2005).

An estimator² is a mathematical function that is used to estimate a parameter from a sample of measurements (*Sampath*, 2005). The estimated parameter, or estimand, is an estimate of the population parameter using the sample (*Westfall and Henning*, 2013).

An estimator itself is a random variable with a PDF, mean, and variance. If an estimator

²Interestingly, *Fisher* (1958) appears to dislike the term “estimator”, as they claim “It is difficult in any particular case to know whether by “estimator” is meant a method of estimation, or the algebraic specification of the estimate reached by that method, or the particular value in a single instance.” This may be why terms such as “Estimand” are used by later authors such as *Westfall and Henning* (2013).

is unbiased and the PDF of the estimator narrows (the variance decreases) about the value of the population parameter as the numbers of samples increases, then the estimator is called a consistent estimator. The expectation value of an estimator is the expected value of the parameter being estimated, which is the value of the parameter for the population that was sampled. Bias is the difference between the expectation value of an estimator and the parameter being estimated. If the expected value of the estimator is equal to its population parameter, then the estimator is unbiased. The mean squared error (MSE) of an estimator provides a quantitative measure of the quality of the estimator. The MSE of an estimator \hat{E} with expectation value E is given by

$$\text{MSE} = \langle (\hat{E} - E)^2 \rangle = \text{Var}(\hat{E}) - \text{Bias}(\hat{E}), \quad (4.10)$$

where $\text{Var}(\hat{E})$ and $\text{Bias}(\hat{E})$ is the variance and bias of the estimator, respectively. The notation $\langle \dots \rangle$ will be used to denote the expected value throughout this dissertation. If the bias is 0, then the MSE of an estimator is equal to the variance of the estimator.

Efficiency of Estimators

The problem of estimation involves designing the best estimator for the task at hand. The best estimator is one that is said to be most efficient, that is, the best estimator is one which has the least MSE for a given number of samples (e.g. *Fisher*, 1958). Inefficient estimators produce larger variances than efficient estimators and require a larger number of samples to obtain the same MSE as a more efficient estimator. Inefficient estimators can be useful in that they might be computationally faster, while still accurate enough to produce a meaningful estimate.

For example, SuperDARN radars are used to determine the power of signals received as a function of range, or the echo power range-profile, of the ionospheric targets. The estimated echo power is produced using the measured voltage samples and a mathematical function. The estimated power is an estimate of the average power that would be scattered back from the plasma if an infinite number of measurements could be made. If there is a problem with the radar hardware such that a voltage offset was present in every measurement, or the estimator was poorly constructed, the estimated echo power will be biased; that is, it will

have an offset compared to what is expected. For this example, a more efficient estimator would be one that includes terms to remove the hardware offset.

The power and autocorrelation function estimators used to process SuperDARN radar data are presented and discussed in detail in Chapters 5 and 6. The notation used in these chapters and throughout this dissertation is that a tilde is used to denote a variable as a measurement (e.g. \tilde{E}) and a hat is used to denote a variable as an estimator (e.g. \hat{E}).

Finally, it must be mentioned that SuperDARN ACF estimates presented in this dissertation assume that the voltage samples used are independent and identically distributed Gaussian random variables. This concept is more thoroughly explored in Chapter 6, but briefly, the lags of the ACF are estimated by “sharing” voltage samples that were used to estimate other lags. As such, the ACF lag estimates have an additional error term that has not been addressed in this dissertation. As a result, the ACF estimator used by SuperDARN (see Equation 4.46) is not the most efficient due to additional variance from the error correlations (e.g. *Farley*, 1969). This has an important implication for Chapter 7 since the least-squares regression estimator has a variance that is biased “downwards” making the least squares estimates of the SuperDARN fitted parameters (power, velocity, and spectral width) seem to have less fitting error than they actually do (e.g. *Haines*, 1978). Although this is a second order effect and is not expected to produce a significant underestimation of the fitted errors, it should be addressed in future work.

4.2 Radar Waveforms and Measurements

Most modern radar applications focus on obtaining both range and Doppler information about a target. In communication systems, such as WiFi, information is encoded and transmitted and then received and decoded. In contrast, in radar, the transmitted signal provides a carrier onto which the target encodes information. A receiver must then decode the information contained in a returned signal by accounting for both the transmitted signal and any modifications to the transmitted signal that may occur as the signal propagates. The optimal signal that should be transmitted by the radar depends on several things, including:

1. the range-Doppler characteristics of the target,

2. the desired accuracy of the measurement,
3. the physics of the radio wave propagation and interaction, and
4. the capabilities of the hardware and software of the radar.

Points 3 and 4 have already been discussed in Chapter 3 for SuperDARN radars. This section will address points 1 and 2 and introduce radar concepts that influence the accuracy of a measurement, including general concepts about radar waveform design.

Throughout this section and this dissertation, it is assumed that the targets being measured by SuperDARN radars exhibit weak-sense stationarity with respect to the measurement time-scales. This means that it is assumed that the mean and autocorrelation of the ionospheric irregularities does not appreciably change in the time it takes to measure the irregularities. This is a common ionospheric radar constraint³, which places fundamental upper limits on appropriate measurement time-scales that can be used (e.g. *Hysell et al.*, 2008; *Nikoukar et al.*, 2012). Strictly speaking, for a weak-sense stationary process, the mean of the voltage samples $m(t)$ does not change in measurement time t_m , namely

$$m(t) = m(t + t_m), \quad (4.11)$$

and the autocorrelation R does not depend on the times when the measurements were taken, t_1 or t_2 , but on the time difference between measurements

$$R(t_1, t_2) = R(t_1 - t_2). \quad (4.12)$$

Weak-sense stationarity is a fundamental requirement for the multiple-pulse technique to be valid, especially the condition described by Equation 4.12. The multiple-pulse technique does not measure all lags of SuperDARN ACF simultaneously, nor does it generally measure them in order (*Farley*, 1972). Proper treatment of non-stationary plasma irregularities and the resulting consequences for SuperDARN are beyond the scope of this dissertation. So long as a process is weak-sense stationary, by the Wiener-Khinchin theorem, all of the spectral information about the process is specified by the ACF. This means that as long as ionospheric irregularities are weak-sense stationary, all of the desired information received by the radar is contained in the SuperDARN ACF.

³There does not seem to be any literature that explores the concept of stationarity for SuperDARN radars.

4.2.1 Signal, Noise, and Clutter

Some of the most fundamental concepts in radar are signal, noise, and clutter. Signals carry the desired information about a radar target. Noise is unwanted interference that limits sensitivity of radar measurements and may contaminate the information carried by the signal, depending on the signal-to-noise ratio (SNR). Clutter is unwanted signal-like interference, such as echoes from targets that are not desired targets (satellites and ground clutter are typical clutter sources for ionospheric radar) (e.g. *Skolnik*, 1980; *Wehner*, 1987; *Turunen et al.*, 2000; *Blanchard et al.*, 2009). The main objective in radar is to gain information from what is measured, which is always a sum of signal, noise, and clutter, and to be efficient in extracting all the information that exists in the measurement (e.g. *Woodward*, 1960). The challenge is to eliminate the unwanted information (noise and clutter), while retaining what is wanted (signal).

Noise is produced by many sources, including the hardware and electronics that are used to transmit and receive signals. Thermal noise constitutes the lower limit of the noise that can be expected. The thermal noise power in a received signal N_T is given by

$$N_T = k_B T B_n, \quad (4.13)$$

where k_B is the Boltzmann constant and T is the temperature of the receiver hardware. The receiver bandwidth B_n is the integral over the range of frequencies that the receiver is sensitive to. The bandwidth is given by

$$B_n = \frac{\int_{-\infty}^{\infty} |H(f)|^2 df}{|H(f_0)|^2}, \quad (4.14)$$

where $H(f)$ is the frequency response of the receiver filter, and f_0 is the frequency with the largest response. Typically, B_n is approximately equal to the half-power bandwidth of the filter as long as the filter is well-designed (*Skolnik*, 1980). For SuperDARN radars, the dominate contribution to noise is from atmospheric and man-made sources (*Landee et al.*, 1957). Noise can be mitigated from measurements by measuring the noise and subtracting it from measurements that contain both signal and noise.

Chapter 5 discusses noise and noise processing for SuperDARN radars in greater detail. The chapter also discusses the dominant source of clutter to SuperDARN measurements and

derives novel estimation and mitigation techniques. Without the proper identification of signal and without properly accounting for and/or mitigating the noise and clutter contributions to radar measurements, it is impossible to accurately quantify the uncertainty in the measurements and fitted parameters (see Chapters 6 and 7). The relative contributions of signal, noise, and clutter also affect the fundamental range and Doppler resolution of the radar measurements.

4.2.2 Matched Filtering

The receiver stage in a radar includes filtering to limit the contribution of noise and interference from frequencies that are not of interest. Matched filtering is a type of filtering that maximized the SNR of the received signal. The impulse response of a filter that maximizes the SNR can be derived in the following manner (e.g. *Levanon and Mozeson, 2004; Richards et al., 2010*). In the frequency domain, the SNR of a received signal that has been filtered is given by

$$\left(\frac{S}{N}\right) = \frac{|\int_{-\infty}^{\infty} H(\omega)S(\omega) \exp(j\omega t) d\omega|^2}{\pi N_0 \int_{-\infty}^{\infty} |H(\omega)|^2 d\omega} \quad (4.15)$$

where $H(\omega)$ and $S(\omega)$ are the frequency responses of the filter and the signal, respectively, and N_0 is the noise power. Using the Schwarz Inequality

$$|\int_{-\infty}^{\infty} H(\omega)S(\omega) d\omega|^2 \leq \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega \int_{-\infty}^{\infty} |H(\omega)|^2 d\omega, \quad (4.16)$$

equation 4.15 can be simplified to

$$\left(\frac{S}{N}\right) \leq \frac{1}{\pi N_0} \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega, \quad (4.17)$$

which implies that

$$H(\omega) = KS^*(\omega) \exp(-j\omega t), \quad (4.18)$$

where K is an arbitrary constant. This result shows that a matched filter is the filter that is a time-reversed and complex-conjugate copy of the signal.

Originally, matched filtering was approximated by designing analogue filtering stages that approximated the shape of the waveform being transmitted. For coded waveforms, complex analogue circuits were required to provide matched filtering (*Skolnik, 1980*). Modern

digital radar systems directly digitize both the transmitted and received signals and can use computers to matched filter the signals using samples of the transmitted waveform.

4.2.3 Waveform Design

The accuracy of a radar measurement is dependent on how the measurement was made, or more specifically, the waveform that was transmitted. Fundamentally, waveform design is a global optimization problem where competing constraints need to be balanced. As will be discussed, the range and Doppler resolutions, maximum unambiguous range and Doppler, propagation conditions, and hardware and software complexities need to be considered. The optimal transmitted waveform therefore depends on what an experiment is trying to accomplish.

Assuming that radar hardware and software are not the limiting factors, the transmitted waveform determines the range resolution. The range resolution σ_τ for a small hard target is determined by the bandwidth of the transmitted waveform f_B using

$$\sigma_\tau = \frac{c}{2f_B}, \quad (4.19)$$

where c is the speed of light (e.g. *Woodward*, 1960; *Levanon and Mozeson*, 1988; *Richards et al.*, 2010). Equation 4.19 is sometimes referred to as the Rayleigh resolution, since the equation is derived using the Rayleigh criterion (*Richards et al.*, 2010).

The simplest waveform for pulse-Doppler radars is a single pulse. The bandwidth of a single pulse is the inverse of the pulse length. Using Equation 4.19, the range resolution of a single pulse increases as the pulse length decreases. It is important to note that the SNR of the received signal is also a function of the pulse length, so shortening the pulse length t_p without increasing the transmitted power will result in a decrease in SNR. This decrease may result in the target no longer being distinguishable from the noise. Increasing the bandwidth of a pulse by using some kind of modulation scheme, such as phase coding, also improves the range resolution.

Similar to range resolution, the Doppler resolution of a radar is dependent on the transmitted waveform. In general, the Doppler resolution is given by the half-width of the Doppler-ambiguity function, which will be discussed below, but first a few analytic expressions will

be presented. The Doppler resolution Δf_D of a single pulse is given by

$$\Delta f_D = \frac{1}{t_p}, \quad (4.20)$$

where t_p is length of the pulse (e.g. *Wehner*, 1987; *Richards et al.*, 2010). Essentially, the Doppler resolution increases as the time observing the target increases. For example, using Equation 4.20, to obtain a Doppler resolution of 10 m/s at 10 MHz, a pulse width of 3 seconds is required. This is only practical for bistatic radars systems and poses significant challenges for range resolution without any pulse coding.

A more practical technique that can be used by monostatic radar and allows both a useful range resolution and useful Doppler resolution involves transmitting a coherent train of pulses. Here, coherent means that both the amplitude and phase of the received signals are recorded. For a coherent pulse train, the Doppler resolution is given by

$$\Delta f_D \simeq \frac{1}{nT_R}, \quad (4.21)$$

where n is the number of pulses and T_R is pulse spacing, also known as the interpulse-period (e.g. *Ulaby et al.*, 1982). Using Equation 4.21, the Doppler resolution can be increased by increasing the number of pulses in the train. Note that the product $nT_R = 3$ s is needed to achieve the same 10 m/s Doppler resolution at 10 MHz as for Equation 4.20.

If a radar target is at a range greater than $cT_R/2$, transmitting several pulses in a coherent train results in a range ambiguity. This ambiguity is fundamentally caused by not being able to discern which echoes being received by the radar are resulting from which pulses. Additionally, there is an upper limit on the largest Doppler shift that can be measured using a sequence of pulses. Therefore, proper operation of pulsed-Doppler radars requires an understanding of two additional concepts: maximum unambiguous range (MUR) and maximum unambiguous Doppler (MUD). The MUR is the largest distance that can be measured between transmission pulses without range aliasing. Similarly, the MUD is the largest Doppler shift that can be measured according to the Nyquist frequency (*Ulaby et al.*, 1982; *Skolnik*, 1980; *Levanon and Mozeson*, 2004; *Richards et al.*, 2010, e.g.). For a train of transmitted pulses with equal time-spacing T_R ,

$$\text{MUR} = \frac{cT_R}{2}, \quad (4.22)$$

$$\text{MUD} = \pm \frac{\lambda}{4T_R}, \quad (4.23)$$

where c is the speed of light and λ is the transmitted wavelength. Increasing T_R causes an increase in the MUR, but it also causes a decrease in the MUD. It is clear that if a radar target requires both a very large MUR and a very large MUD, the experiment designer needs to resolve a paradox, which is most commonly known as the Doppler Dilemma. For example, SuperDARN radars must be able to measure plasma velocities >1 km/s at ranges >3000 km at a transmission frequency of, say, 10 MHz ($\lambda=30$ m). Using these values, Equation 4.23 yields $T_R \approx 7.5$ ms and Equation 4.22 the $\text{MUR} \approx 1123$ km, meaning that it is not possible to unambiguously resolve targets moving faster than 1 km/s at ranges greater than 1123 km by transmitting a coherent pulse train. The multiple-pulse technique, introduced in Chapter 3 resolves this paradox at the expense of SNR.

Routinely, SuperDARN radars transmit $300 \mu\text{s}$ pulses, resulting in a range resolution, as determined by Equation 4.19, of $\sigma_r = 45$ km. Generally, SuperDARN radars improve range resolution via shortening the pulse length only, with the shortest pulse lengths being $100 \mu\text{s}$ corresponding to 15 km range resolution. Experiments have been performed by the Alaskan SuperDARN group using Barker-coded pulses and these have attained a range resolution of better than 20 km, however these experiments have not been widely repeated due to radio licensing concerns regarding the increased transmitter bandwidth required (Bill Bristow, private communication, 2012).

So far, range resolution for SuperDARN radars has been presented in the context of “hard” point-like targets. However, as was discussed in Chapter 3, ionospheric irregularities are “soft” beam-filling targets. For ionospheric radars, Equation 4.19 provides a fundamental lower limit on the maximum attainable range resolution. The ionosphere is significantly larger than the transmitted pulse, so for ionospheric volume filling targets, this received pulse is effectively lengthened. For example, for a rectangular pulse shape, this means that the received power comes from a trapezoidal area that is larger than the pulse that was transmitted (*Ulaby et al.*, 1982). SuperDARN radars typically record voltage samples with gate sizes equal to the transmitted pulse-length. Typically SuperDARN range gates are 45 km and 15 km for the $300 \mu\text{s}$ and $100 \mu\text{s}$ pulse lengths respectively, but the range resolution for the same pulse lengths is 90 km and 30 km respectively. Therefore, it is important to remember

that the *range gate size* (given by $ct_s/2$ where t_s is the sampling time) is not the same thing as the *range resolution*. This is especially important to remember when interpreting SuperDARN data, since the data for a particular range gate is a weighted sum of the signals from the range gate itself and the adjacent range gates. A good approximation of the range resolution for ionospheric radars, such as SuperDARN radars, is to multiply Equation 4.19 by 2. This means that for routine operations, SuperDARN radars have a range resolution of ≈ 90 km, but can attain a range resolution of ≈ 30 km by shortening the pulse length to $100 \mu\text{s}$. This approximation works best when pulses are approximately rectangular in shape. To obtain the true range and Doppler resolutions for any pulse or pulse sequence, one must calculate the half-power width of the range-Doppler ambiguity function.

Ambiguity Functions

An ambiguity function is a mathematical function used in pulsed radar signal processing that provides a quantitative measure of the quality of a transmitted waveform (Woodward, 1960). Generally, a range-Doppler ambiguity function, sometimes referred to as a delay-Doppler ambiguity function, describes what combinations of target ranges and Doppler shifts produce indistinguishable signals when received by the radar (e.g. Woodward, 1960; Levanon and Mozeson, 2004; Richards *et al.*, 2010). The delay-Doppler ambiguity function $\chi(\tau, \nu)$ is calculated from the magnitude of the convolution of the transmitted waveform $u(t)$ with its Doppler-shifted matched-filter response $u^*(t + \tau) \exp(j2\pi\nu t)$, using

$$|\chi(\tau, \nu)| = \left| \int_{-\infty}^{\infty} u(t) u^*(t + \tau) \exp(j2\pi\nu t) dt \right|, \quad (4.24)$$

where $u(t)$ is the complex-valued function describing the transmitted waveform, τ is the time-delay to a target at range R (e.g. $\tau = 2R/c$), and ν is Doppler frequency. The range-Doppler ambiguity function has a maximum at $\tau = 0$ and $\nu = 0$ and is symmetric about the origin. It also has unity volume under the surface, such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\chi(\tau, \nu)|^2 d\tau d\nu = 1. \quad (4.25)$$

Proofs of these properties and derivations are widely available in the literature (e.g. Woodward, 1960; Levanon and Mozeson, 2004; Richards *et al.*, 2010). The most important property

of the ambiguity function is constant unity volume, meaning that waveform design is fundamentally about making compromises. If an undesirable peak in the ambiguity function is removed by changing the transmitted waveform $u(t)$, then that the volume that was under that peak needs to be accounted for elsewhere.

A single rectangular pulse, defined by the **rect** function

$$u(t) = \frac{1}{\sqrt{t_p}} \text{rect} \frac{t}{t_p}, \quad (4.26)$$

where t_p is the length of the pulse, has a range-Doppler ambiguity function (*Levanon and Mozeson, 2004*) given by

$$|\chi(\tau, \nu)| = \left| \left(1 - \frac{|\tau|}{t_p} \right) \frac{\sin[\pi t_p (1 - |\tau|/t_p) \nu]}{\pi t_p (1 - |\tau|/t_p) \nu} \right|, \quad |\tau| \leq t_p. \quad (4.27)$$

Setting $\nu = 0$ yields the delay ambiguity function of a rectangular pulse

$$|\chi(\tau, 0)| = 1 - \frac{|\tau|}{t_p}, \quad |\tau| \leq t_p, \quad (4.28)$$

which is simply a triangular shaped function. The half-power width of Equation 4.28 yields the range resolution of the pulse. The half-power width is t_p , meaning that the range resolution of a rectangular pulse for ionospheric radar is equal to the pulse length, as previously discussed. Setting $\tau = 0$ yields the Doppler ambiguity function of a rectangular pulse

$$|\chi(0, \nu)| = \left| \frac{\sin \pi t_p \nu}{\pi t_p \nu} \right|, \quad (4.29)$$

which is the well known **sinc** function. The Doppler resolution is given by the half-power width of Equation 4.29, $\approx 1.2/t_p$, which is similar to the Doppler resolution given by Equation 4.20.

The ideal ambiguity function is a “thumbtack”, or delta function, where the ambiguity function is 0 when $|\tau|, |\nu| > 0$, but this would require an infinitely long white-noise waveform (*Levanon and Mozeson, 2004*). For finite transmitted waveforms, at best an increased “noise” floor of the ambiguity is achieved with a narrow range-Doppler peak centered at $\tau = \nu = 0$. The most promising transmitter waveforms for SuperDARN radars that can achieve a thumbtack-like ambiguity function are those produced using long pseudo-random phase codes (*Richards et al., 2010*), since the code approximates white noise. However, long waveforms

are challenging to implement in monostatic radar configurations due to the large dynamic range required to measure both the transmitted and received signals simultaneously. If the transmitting and receiving elements can be sufficiently isolated, such as by using a co-located pair of antennas, long pseudo-random waveforms may become more practical. In Chapter 5, the ambiguity function is used to analyse the self-clutter produced by the multiple-pulse technique used by SuperDARN radars.

4.3 Regression Techniques

In the previous section, a brief background in how radars use different waveforms to achieve measurements with desired range and Doppler resolution was presented. This section will describe how radars extract desired information from the measured voltage samples using regression techniques, with a specific focus on the least-squares regression techniques used to obtain the SuperDARN fitted parameters, power, spectral width, and Doppler velocity.

Regression techniques use probability theory to determine the likelihood of a physical model being consistent with the measured data. In general, some kind of mathematical function that quantifies the “closeness” of the model to the data as a function of the physical parameters is minimized until the physical parameters are found that produce the “closest” fit of the model to the data. Confidence intervals, which describe the likelihood of the physical parameters having other values, are then constructed. The confidence intervals can be interpreted as error bars for the fit. The confidence intervals and the quality of the regression are a function of the quality and quantity of the data. More and better data reduces the uncertainty in the parameters being fitted for. Essentially, data provides information that constrains the set of possible values for the parameters (*Westfall and Henning, 2013*).

To perform a regression, first a forward model is developed that predicts what is expected to be measured depending on the physical parameters. For ionospheric radars, the forward model predicts what is expected to be measured by the radar given ionospheric parameters. Regression techniques are then used to “fit” the forward model to the data (measured signals) by changing the values of the physical parameters until the best fit of the model to the data is achieved. In SuperDARN radar signal processing, one of the most commonly used forward

models is a model of the ACF $R(\tau_i)$ given by

$$R(\tau_i) = P e^{-2\pi w_d \tau_i / \lambda} e^{j4\pi v_d \tau_i / \lambda}, \quad (4.30)$$

where P , w_d , and v_d are the lag-zero signal power, spectral width, and Doppler velocity, respectively, λ is the wavelength of the transmitted radar pulses, and τ_i is the lag time (e.g. *Ribeiro et al.*, 2013a). Comparing the model with an ACF measured by a SuperDARN radar, the values of P , w_d , and v_d are changed until values of the parameters are found that “best match” the model with the data. The “best match” is a quantitative measure that is dependent on the regression technique used.

4.3.1 Least-Squares Regression

To extract parameters from received signals, SuperDARN signal processing utilizes least-squares regression techniques. The metric used to determine the “best match” for least-squares is the chi-squared, or χ^2 , sum given by

$$\chi^2 = \sum_i^N \left(\frac{d_i - m_i}{\sigma_i} \right)^2, \quad (4.31)$$

which determines the squared difference between N measurements (data) d_i and N model values m_i all normalized by the standard deviation of the measurements σ_i (e.g. *Haines*, 1978; *Press et al.*, 1986; *Westfall and Henning*, 2013). If $\sigma_i = 1$, then the technique is called ordinary least-squares regression; otherwise, it is known as weighted least-squares regression. The term in the brackets is called the residual and is a random variable with a PDF, mean, and variance. If the residual is a Gaussian random variable, then χ^2 is a random variable characterized by a chi-squared distribution, hence the notation (e.g. *Simon*, 2007).

As the name suggests, the goal of least-squares regression is to minimize the χ^2 sum. The model values m_i are dependant on the physical parameters and the physical parameters are modified until values are found that produce the minimum possible χ^2 . Confidence intervals for the “best-fit” parameter values are then calculated. Least-squares regression is fundamentally a global-minimum search problem, which can either be simple, if the model is linear and the number of parameters is low, or complex, if the model is non-linear and the number of parameters is large.

4.3.2 Linear Least-Squares Regression

If the model used to produce values m_i in Equation 4.31 is linear, analytic solutions for the best-fit parameters can be derived, which significantly simplifies the regression process (e.g. *Press et al.*, 1986). To obtain the analytic solutions, first start with the χ^2 sum

$$\chi^2 = \sum_i^N \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2, \quad (4.32)$$

where y_i are the dependent variables (data) obtained at independent variable values x_i , and a and b are the parameter values of the linear model, specifically the y-intercept and slope, respectively. To minimize χ^2 , the derivative of χ^2 with respect to each parameter is taken and set to zero, obtaining

$$\frac{\delta\chi^2}{\delta a} = 2 \sum_{i=1}^N \frac{y_i - a - bx_i}{\sigma_i^2} = 0, \quad (4.33)$$

$$\frac{\delta\chi^2}{\delta b} = 2 \sum_{i=1}^N \frac{x_i(y_i - a - bx_i)}{\sigma_i^2} = 0. \quad (4.34)$$

These equations can be rewritten as a system of linear equations with 2 equations and 2 unknowns

$$aS + bS_x = S_y, \quad (4.35)$$

$$aS_x + bS_{xx} = S_{xy}, \quad (4.36)$$

where the sums S , S_x , S_y , S_{xx} , and S_{xy} are given by

$$S = \sum_{i=1}^N \frac{1}{\sigma_i^2}, \quad S_x = \sum_{i=1}^N \frac{x_i}{\sigma_i^2}, \quad S_y = \sum_{i=1}^N \frac{y_i}{\sigma_i^2}, \quad (4.37)$$

$$S_{xx} = \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2}, \quad S_{xy} = \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2}. \quad (4.38)$$

Solving the system of linear equations yield analytic expressions for a and b

$$a = \frac{S_{xx}S_y - S_xS_{xy}}{\Delta} \quad (4.39)$$

$$b = \frac{SS_{xy} - S_xS_y}{\Delta} \quad (4.40)$$

where Δ is the determinant given by

$$\Delta = SS_{xx} - (S_x)^2. \quad (4.41)$$

In this way, the best-fit values of a and b are analytically determined using computationally simple operations. A more detailed discussion about the usage of linear regression in fitting SuperDARN measured ACFs is found in Chapter 7.

The variance and covariance of the fitted parameters are given by

$$\sigma_a^2 = \frac{S_{xx}}{\Delta}, \quad \sigma_b^2 = \frac{S}{\Delta}, \quad (4.42)$$

and

$$\sigma_{ab}^2 = \frac{-S_x}{\Delta}, \quad (4.43)$$

respectively. The variance and covariance are used to construct confidence intervals for the fitted parameters.

4.3.3 Non-Linear Least-Squares Regression

If the model is non-linear, such as 4.30 (which is typically used by SuperDARN), there is no generally applicable analytic solution when solving for the best fit parameters. One possible technique is to perform a brute force global search for the parameter values that produce a minimum χ^2 , however if the number of parameters in the model is M , the search space scales exponentially with M , so this brute force method is computationally expensive and slow. For example, to find the minimum for a 2 parameter model using a grid of 1000 values of each parameter requires calculating 10^6 χ^2 values, and a model depending on 3 parameters would require calculating 10^9 χ^2 values.

There are other techniques for solving non-linear least squares (such as particle swarm optimization (e.g. *Liao et al.*, 2011)), but the most commonly used non-linear least-squares technique is the Levenburg-Marquardt algorithm (e.g. *Levenburg*, 1944; *Marquardt*, 1963; *Press et al.*, 1986; *Westfall and Henning*, 2013). The Levenburg-Marquardt algorithm requires an initial guess of the best-fit parameters and then iteratively determines the best way to perturb the model parameters in the “direction” that decreases the χ^2 , using finite-difference approximations of the derivatives of the χ^2 with respect to each of the parameters. The algorithm then updates the initial guess of the best-fit parameters in the direction of decreasing χ^2 while also modifying the size of the step in that direction such that χ^2 always decreases. The iterative process continues until the change in χ^2 is smaller than some

user-defined threshold value, at which point the final value of the parameters is the best-fit value.

A caveat of the Levenburg-Marquardt algorithm is that it finds a minimum in χ^2 , but that minimum may not be the global minimum. The minimum χ^2 value found is highly sensitive to the shape of the χ^2 surface, which may have many local minima that change location depending on the data being fitted, and it is highly sensitive to the initial guess of the best-fit parameters. Chapter 7 contains a detailed discussion about the application of the Levenburg-Marquardt algorithm as applied to fitting SuperDARN ACFs and a new fitting methodology.

4.3.4 Confidence intervals

Confidence intervals are essentially errors bars for fitted parameters. More strictly speaking, for least-squares regression, a confidence interval quantifies the probability of obtaining values of a fitted parameter that would lie within the confidence interval if the experiment were repeated. For example, for a fitted parameter μ with a confidence interval $24.39 < \mu < 29.23$, according to *Westfall and Henning* (2013), the most rigorous and accurate description of the confidence interval for μ is the following: “Since μ will lie within the upper and lower limits of similarly constructed intervals for 95% of the repeated samples, my sample is likely to be one of those samples where μ is within the upper and lower limits, and I am therefore 95% confident that μ is between 24.39 and 29.23”. The confidence interval also quantifies how the value of the fitted parameter changes with a change in χ^2 .

When the residuals of the χ^2 sum are “small”, they are Gaussian distributed random variables and a confidence interval for a fitted parameter can be constructed using a change in the chi-squared value, a $\Delta\chi^2$ (e.g. *Press et al.*, 1986). The $\Delta\chi^2$ and standard deviation of the fitted parameter $\sigma_{\hat{\mu}}$ can be used to construct the confidence interval from the standard deviation of the fitted parameter using

$$\delta\hat{\mu} = \pm\sqrt{\Delta\chi^2}\sigma_{\hat{\mu}}. \quad (4.44)$$

For a single parameter fit, a 95% confidence interval can be constructed using a $\Delta\chi^2 = 4$,

resulting in

$$\delta\hat{\mu} = \pm 2\sigma_{\hat{\mu}}. \quad (4.45)$$

When the residuals are not small, they are generally not Gaussian distributed and thus a confidence interval for the fitted parameters cannot be constructed using the above method. Instead, confidence intervals must be constructed by tracing along the chi-squared curve, moving up by $\Delta\chi^2$ from the minimum best-fit χ_{min}^2 . The values of the fitted parameters at $\chi_{min}^2 + \Delta\chi^2$ bound the confidence interval for the fitted parameters. This method is known as the “constant chi-square boundary as confidence limits” and will be discussed in more detail in Chapter 7.

4.4 SuperDARN Signal Processing

SuperDARN signal processing specifically requires the concepts discussed in this chapter, beginning with the statistical properties of the voltage samples recorded by the radar receiver and ending with the least-squares regression used to obtain the fitted parameters, power, spectral width, and line of sight Doppler velocity. This section will provide a broad overview of SuperDARN signal processing, including a discussion about the general workflow required to process the voltage samples into ACFs and the ACFs into the fitted data products. The specific details regarding signal reception, filtering, and digitization, including the hardware used to perform these tasks, are radar dependent. A proper treatment of the receive chain for SuperDARN radars would require a dedicated chapter and is beyond the scope of this thesis. For the purposes of this thesis, it is sufficient to assume we have been provided complex voltage samples from a quadrature receiver and proceed from there.

The volume of ionosphere sampled by SuperDARN radars is much larger than the wavelength of HF radio waves. As discussed in Chapter 3, the ionospheric irregularities that contribute to scatter must be half the radar wavelength in scale. This means that it can be reasonably assumed that the volume contains a large number of randomly distributed scatterers, such that any HF signal scattered by this volume could be characterized as the sum of a large number of individually scattered signals. As will be discussed in Chapter 6, the consequence of this is that the voltage samples recorded by SuperDARN radars are Gaussian

random variables. Since all later stages of the signal processing workflow operate on the voltage samples or quantities derived from the voltage samples, it is of fundamental importance that the statistical properties of the voltage samples are known. This is especially important when performing error analysis.

4.4.1 Current SuperDARN Signal Processing

The following summarizes how SuperDARN presently⁴ processes measured voltage samples into the SuperDARN ionospheric parameters power, spectral width, and velocity. The first step in SuperDARN signal processing is to obtain gated (binned) complex voltage samples from the quadrature receiver at the radar. This is performed at the radar site and the voltage samples are generally not saved for later usage. However there is an option to save the voltage samples and they are stored in an `iqdat` file. The format of this file is given in Appendix D.1. A set of voltage samples is collected for each multiple-pulse sequence that is transmitted. Several of these voltage sets are then used to obtain a time-domain estimate of the complex radar auto-correlation function (ACF). The multiple-pulse sequence samples the ionospheric irregularities at different lag times such that using the received voltage samples $\tilde{V}(t)$, a mean value of a lag of the ACF can be estimated from K pulse sequences using

$$\hat{R}(\tau) = \frac{1}{K} \sum_{k=1}^K \tilde{V}_k(t) \tilde{V}_k^*(t + \tau), \quad (4.46)$$

where $\hat{R}(\tau)$ is an estimate of the mean value of the ACF at lag time τ . This estimate is obtained from random variables (the voltage samples) and is therefore a random variable with a probability distribution function. The statistical properties of ACF estimates are discussed in detail in Chapter 6. The mean ACF is recorded in a `rawacf` file so it can be processed at a later time. The desired ionospheric parameters (power, spectral width, and velocity) are then obtained by processing the `rawacf` files and fitting the ACFs using least-squares regression.

Weighted linear least-squares regression is applied to the magnitude and phase of the ACFs to obtain the ionospheric parameters. More details of how this fitting is performed

⁴At the time of this writing, Autumn 2017.

are presented in Chapter 7, but briefly, the power and spectral width are obtained by fitting the magnitude of the ACF and the velocity is obtained by fitting the phase. Finally, these fitted parameters are saved in `fitacf` files and are ready for any space physics analysis. The `fitacf` files from all SuperDARN radars are also combined to create global convection maps.

The original fitting software and signal processing workflow was developed in the late 1980s when computational resources were more expensive than they are today. Additionally, the SNR of SuperDARN signals are large, routinely 10 dB or greater. Due to these constraints, at the time it was desirable to use computationally inexpensive linear fitting algorithms over non-linear algorithms. Moreover, since the SuperDARN SNR is usually relatively large, errors for the fitted parameters are expected to be small, and as such, less effort was spent on error analysis than would have been required had the SuperDARN SNR been much lower. Using ad hoc error estimates and linear fitting techniques provided a simple-to-understand and inexpensive data processing methodology that aided in the growth and maintenance of SuperDARN as the network grew to the present day 35 radars.

With the modern availability of inexpensive and powerful computational resources, it is possible to reconsider the SuperDARN signal processing methodology⁵. As discussed at the end of Chapter 3, this entire thesis was motivated by the failed attempt at designing an experiment that enabled SuperDARN derived electron density measurements to be routinely available. The review identified some challenges with the current signal processing methodology and provided a path towards solutions, which are described in Chapters 5 and 6, 7.

⁵Sufficient computational resources have been available since the late 1990s or at the very latest the early 2000s.

CHAPTER 5

SELF-CLUTTER ESTIMATION

5.1 Background

The objective of this dissertation is to develop an improved SuperDARN signal processing methodology such that accurate fitted ionospheric parameter errors can be obtained. In order to achieve this objective, an investigation of the current SuperDARN signal processing methodology first needed to be performed. The investigation primarily involved reading through the source code for the SuperDARN Radar Software Toolkit to reverse engineer the signal processing workflow. Several publications were also consulted, including *Villain et al.* (1987), *Ponomarenko and Waters* (2006), and *Ribeiro et al.* (2013a), however, there is no single publication that describes the signal processing methodology in detail. The software and literature investigation revealed several challenges, summarized as follows:

1. Error estimates for the magnitude and phase of the SuperDARN ACFs are inaccurate and incorrect,
2. ACF lags containing self-clutter are discarded from linear regression (fitting) instead of being weighted with appropriately larger errors, and
3. Reliable phase unwrapping in the presence of self-clutter is not possible, especially when the SNR is low.

As is discussed in detail in Chapter 7, the error estimates used for the magnitude and phase of SuperDARN ACFs are ad hoc expressions that were crafted based on reasonable arguments about the expected behaviour of the errors in the magnitude and phase as a function of SNR. In weighted least-squares fitting, using inaccurate error estimates leads to inaccurate

fitted parameter errors, thus explaining the reasons why the fitted velocity errors provided inaccurate expectations in designing the plasma density experiment. Additionally, the phase of the ACF is a periodic or circular quantity, such that it “wraps” where 0 and 2π are equivalent. This means that in order to perform a linear fit to the phase, the phase needs to first be “unwrapped”. In the presence of noise and self-clutter, it is not always possible to unambiguously unwrap the phase to perform a linear fit.

Since fitting depends on the ACF lag errors, the first task that was identified was to determine the true expressions for the variance of ACF lag estimates. Once the variance was determined, challenges 2 and 3, regarding fitting the ACF lags, was addressed. Research into ionospheric radar signal processing, specifically an incoherent scatter radar publication by *Farley* (1969), revealed that a first-order error estimate σ for the lags of the SuperDARN ACF may be given by

$$\sigma = \frac{S + N + C}{\sqrt{K}}, \quad (5.1)$$

where K is the number of pulse sequences transmitted and S , N , and C are the signal, noise, and clutter powers, respectively. Equation 5.1 shows that the uncertainty in an estimate of a SuperDARN ACF lag is increased by contributions due to noise and clutter. This means that in order to obtain estimates of the ACF errors, one needs to quantify the signal, noise, and clutter powers. For SuperDARN, it was determined that the dominant form of clutter contributing to measurements was self-clutter caused by using the multiple-pulse technique as the transmitted radar waveform (*Farley*, 1972).

Reviewing the SuperDARN software and literature determined that while SuperDARN signal processing does include techniques for estimating noise, it does not include any techniques for estimating self-clutter. Instead, SuperDARN ACFs are discarded using a self-clutter filtering criteria developed in *Ponomarenko and Waters* (2006). This means that instead of weighting lags that contain self-clutter with appropriate larger errors, the data was being discarded. Further investigation of the SuperDARN software found that this self-clutter criteria would sometimes discard data, which was revealed to be valid data when the ACFs were visually inspected. All of this research resulted in the realization that an estimate of the self-clutter was needed in order to provide an accurate estimate of error in SuperDARN ACF lags.

In summary, the following improvements were identified as solutions to the challenges and inaccuracies inherent to the current SuperDARN signal processing workflow:

1. Development of a method to estimate self-clutter (Chapter 5),
2. Rigorous analysis of the statistical properties of SuperDARN ACF estimates, including exact analytic formulas for the error (Chapter 6), and
3. Development of a new fitting methodology that utilizes both the self-clutter estimates, ACF variance (error) estimates, and non-linear fitting techniques for fitting the real and imaginary components of SuperDARN ACFs instead of fitting the magnitude and phase as is done with current SuperDARN fitting techniques (Chapter 7).

Improvements 1 and 2 were needed to address challenge 1 and improvement 3 was needed to address challenges 2 and 3.

This chapter addresses improvement 1 and contains a detailed analysis of the self-clutter produced by the multiple-pulse technique and how it affects SuperDARN measurements. Several estimators of self-clutter were derived and tested.

Chapter 6 considers improvement 2 by deriving and discussing the statistical properties of SuperDARN ACF lag estimates, as well as deriving analytic expressions for the probability distribution functions of the ACF lags. Full expressions for the variance of ACF lag estimates are also presented, validated, and discussed.

Chapter 7 discusses current SuperDARN fitting methods and presents an improved fitting methodology that incorporates the power-based self-clutter estimator, which developed and presented in the current chapter, with the exact expressions of the variance of SuperDARN ACF lag estimates developed in Chapter 6.

5.1.1 Published Manuscript

Authors: *Ashton S. Reimer and Glenn C. Hussey*

Reference: Reimer, A. S., and G. C. Hussey (2015), Estimating self-clutter of the multiple-pulse technique, *Radio Sci.*, 50, 698–711, doi:10.1002/2015RS005706.

This chapter contains the first published work towards developing an improved SuperDARN data processing methodology, which provides improved fitted parameter errors, by developing three novel self-clutter estimators. Most of the research developing the self-clutter estimators involved reviewing and understanding all of the SuperDARN signal processing code written in C¹, drawing countless Farley diagrams (explained in the publication below), and producing stacks of paper containing attempted self-clutter estimator derivations. Understanding the current treatment of self-clutter by SuperDARN algorithms was primarily obtained reading the SuperDARN processing C code and reading *Ponomarenko and Waters* (2006). *Farley* (1972) was studied in detail to understand the origins of self-clutter for the multiple-pulse sequence. *Farley* (1969) provided a theoretical foundation from the perspective of estimator theory applied to Gaussian random variables such that several self-clutter estimators and their associated variance could be derived (the derivations are provided in Appendix B). Additionally, a familiarity with incoherent scatter radar measurement techniques and data processing methods provided a foundation from which the derivations could be based upon.

Extensive verification of the derived self-clutter estimators was performed using a significantly modified (by this author) version of the realistic SuperDARN data simulator (*Ribeiro et al.*, 2013b). The simulator was originally developed to produce simulated voltage samples that could be used to calculate realistic simulated SuperDARN ACFs. However, there were a number of significant improvements to the code that were required by this research in order to accurately simulate SuperDARN voltage samples.

Initially, an attempt to estimate the self-clutter exactly was made, such that the estimated self-clutter could be subtracted from the ACF lag estimates in the same way that noise is estimated and subtracted from lag-zero of the ACF. Using the voltage sample simulator, this author was able to explore what ionospheric backscatter conditions allowed self-clutter subtraction to work (the conditions are discussed below). This research led to the development of a *voltage-based self-clutter estimator*, however this estimator could only be used for a narrow subset of ionospheric conditions. It became clear from this research that a more

¹At the time (2013), this was thousands of lines of undocumented code in the Radar Software Toolkit (RST). Since then, RST has been significantly simplified and documented by others.

general estimator was needed. Therefore, a *maximal voltage-based self-clutter estimator* was developed giving an upper limit estimate of the self-clutter for any ionospheric conditions. The only problem with this estimator was that the voltage sample data (`iqdat` files) are not regularly produced by SuperDARN radars, so any voltage-based estimator could not be used to reprocess the entire SuperDARN database. As such, a *maximal power-based self-clutter estimator* that can be used to process all past SuperDARN data was developed. The *maximal power-based self-clutter estimator* is an important and novel contribution to SuperDARN research because it can be used to determine the variance of ACF lag estimates (discussed in Chapter 6). Using this estimator also means that ACF lags with self-clutter do not need to be discarded when fitting SuperDARN ACFs (discussed in Chapter 7). Without this quantification of the self-clutter contributing to SuperDARN measurements, it would not be possible to accurately estimate the ACF lag variance.

What follows is the published self-clutter manuscript with the same content as was published in *Radio Science*. I am licensed (see Appendix E for details) to reproduce this manuscript within this thesis. The figure numbers, section headings and general formatting have been altered to ensure consistency with the rest of this thesis; however, no other alterations have been made to the manuscript.

5.2 Abstract

Autocorrelation function (ACF) estimates from voltage data measured by high frequency (HF) ionospheric radar systems that utilize the multiple-pulse technique of *Farley* (1972) are susceptible to interference from self-clutter. Self-clutter is caused by simultaneous returns from multiple transmitted pulses echoing from unwanted, or ambiguous ranges. Without accurate estimates of self-clutter it is impossible to account for all the uncertainty in estimates of the radar ACF. Voltage- and power-based self-clutter estimators are presented and evaluated using a modified version of the radar data simulator of *Ribeiro et al.* (2013b) and data from the Super Dual Auroral Radar Network (SuperDARN) (*Greenwald et al.*, 1995). It is shown that self-clutter caused by ambiguous ranges filled with ground scatter can be accurately estimated using a voltage-based self-clutter estimator, but that for ionospheric origin self-clutter a maximal estimator must be used. Two maximal self-clutter estimators are discussed and verified using the radar data simulator. A discussion of the application of the self-clutter estimator as it is applied to ACFs obtained with Saskatoon SuperDARN radar is also presented.

5.3 Introduction

Decameter scale ionospheric irregularities can be probed using radars in the high frequency (HF) Super Dual Auroral Radar Network (SuperDARN) (*Greenwald et al.*, 1995; *Chisham et al.*, 2007). A quadrature receiver is used to recover the gated (binned) complex voltage signal, called voltage samples, received from these irregularities (e.g. *Villain et al.*, 1987). Lags (time steps) of the complex radar auto-correlation function (ACF) for each range are derived from the average of pairs of correlated lagged voltage samples (e.g. *Farley*, 1969). Desired parameters (Doppler shift, echo power, etc.) are then extracted from the radar ACF using fitting techniques (e.g. *Ribeiro et al.*, 2013a). The radar ACF is composed of contributions of signal (which contains the useful information), noise, and clutter. In general, the noise and clutter increase uncertainty associated with extracting useful information from the radar ACF (*Farley*, 1969; *Ulaby et al.*, 1982). For example, *Farley* (1972) showed that

to first order, the variance in estimating a lag of the radar ACF (\hat{R}) is given by,

$$\langle(\delta\hat{R})^2\rangle = \frac{1}{K} (S + N + C)^2, \quad (5.2)$$

where S , N , and C are the signal, noise, and clutter powers, respectively, with K pairs of voltage samples used to produce the lag estimate.

Noise interference in the HF band is primarily due to both atmospheric and man-made sources (*Landee et al.*, 1957). Noise may be readily removed by subtracting estimates of the average noise power (*Farley*, 1969; *Ulaby et al.*, 1982). In contrast, it is generally more difficult to account for clutter. Clutter is defined as unwanted signal-like interference (*Skolnik*, 1980) and for ionospheric radars typical sources include ground scatter (e.g. *Turunen et al.*, 2000; *Blanchard et al.*, 2009), echoes from satellites or planes (e.g. *Porteous et al.*, 2003), and ambiguous range echoes (e.g. *Zrnic et al.*, 1999; *Grydeland and Gustavsson*, 2011). Removal or suppression of clutter usually requires complex signal processing algorithms and/or modifications to transmitted radar waveforms (e.g. *Turunen et al.*, 2000). Ambiguous range echoes are a function of the transmitted radar waveform and are sometimes referred to as cross-range interference (e.g. *Ponomarenko and Waters*, 2006), but are more commonly referred to as self-clutter (e.g. *Rihaczek*, 1965; *Zrnic et al.*, 1999; *Kolawole*, 2002; *Grydeland and Gustavsson*, 2011; *Cook*, 2012). As stated in Equation 5.2, without accurate quantification of signal, noise, and clutter, it is impossible to determine the error in estimating lags of the radar ACF.

Some ionospheric radars such as the high-frequency (HF) Super Dual Auroral Radar Network (SuperDARN) radars (*Greenwald et al.*, 1995; *Chisham et al.*, 2007) utilize the multiple-pulse technique of *Farley* (1972) to obtain estimates of the radar ACF. The multiple-pulse technique utilizes the transmission of non-redundantly spaced pulses to suppress range ambiguities. Such spacings enable improved range and Doppler performance compared to a uniformly spaced coherent pulse train (*Farley*, 1972); however, with reduced radar duty cycle and consequently reduced average signal-to-noise ratio (SNR). Reduced SNR is not typically a problem for SuperDARN radars, which routinely detect ionospheric targets with SNR of 20 dB or larger. The multiple-pulse technique suppresses ambiguous range echoes and extends the maximum unambiguous range of the radar, but as will be shown, some ambiguity is still

present. The delays between transmitted pulses are integer multiples of a fundamental pulse-repetition time (PRT), τ , a parameter that determines the maximum unambiguous Doppler shift that can be resolved by the radar waveform. The maximum unambiguous range is given by the sequence repetition time (SRT), the time delay between pulse sequences. Three other important quantities of a multiple-pulse sequence are the pulse length, t_p , the number of pulses transmitted, N_p , and the pulse table, PTAB, which specifies at what time pulses are to be transmitted. Often, measurements from K pulse sequences are averaged together to estimate target parameters. The multiple-pulse technique is especially useful for studying overspread radar targets, such as those probed by HF over-the-horizon ionospheric radar systems (targets at long range: > 1000 km and with high velocity: ~ 1 km/s).

Despite the superior range ambiguity suppression of the multiple-pulse technique, in general range ambiguities are still present and under certain conditions significantly “clutter” the received signal. While techniques for estimating clutter power due to ground and aircraft echoes are known, for the multiple-pulse technique it has remained an open problem to estimate the self-clutter present in estimates of the radar ACF. As Equation 5.2 indicates, without quantification of *all* types of clutter power it is impossible to obtain accurate error estimates.

The present study discusses a generalized signal-processing algorithm for quantitative estimation of self-clutter contained in ACF estimates by utilizing voltage samples obtained from echo power (lag-zero) measurements. A maximal estimate of the self-clutter power is also discussed. The self-clutter estimators derived in this study enable the estimation of signal-derived errors for the radar ACFs obtained using the multiple-pulse technique, which will allow the use of error-weighted fitting algorithms to extract ionospheric parameters from the radar ACFs.

5.3.1 Power and ACF Estimators

Ionospheric radars are often able to determine the echo power as a function of range, or the power range-profile, of ionospheric targets. Echo power is estimated from measured voltage samples using a power estimator. The voltage samples are the result of gating the continuous received voltage signal into bins of time. Range gates are determined by time of

flight or transit time (i.e. voltage samples of 300 μs correspond to range gates of 45 km). In general, an estimator or statistic, is a function used to estimate a parameter from a sample of measurements (*Sampath, 2005*). To estimate the parameter power, P , in range gate r , from K independent voltage samples from the same range, \tilde{V}_{rk} , one may use the estimator

$$\hat{P}_r = \frac{1}{K} \sum_{k=1}^K \tilde{V}_{rk} \tilde{V}_{rk}^*, \quad (5.3)$$

where \hat{P}_r is the estimator for parameter P at range r , the asterisk denotes the complex conjugate, and the tilde indicates a measurement. The indices r and k indicate the k th sample at range r . The expected value of the estimator is $\langle \hat{P}_r \rangle = P_r$, which is to say that when a large number of measurements are used ($K \rightarrow \infty$) the estimator \hat{P}_r approaches the true value P_r . The notation using tilde (measurement), hat (estimator), asterisk (complex conjugate), and $\langle \dots \rangle$ (expected value) will be used throughout this paper.

The uncertainty in an estimator is given by the mean-squared error (MSE), defined as the expected value of the average of the square of the difference between the statistic and the parameter being estimated. For example, the MSE ($\langle (\hat{P}_r - P_r)^2 \rangle$) of the power estimator defined in Equation 5.3 is

$$\text{MSE}(\hat{P}_r) = \frac{P_r^2}{K}. \quad (5.4)$$

In the above discussion, and for the remainder of this study, the real and imaginary components of the complex voltage samples, \tilde{V}_{rk} , are assumed to be zero-mean Gaussian random variables (*Farley, 1969; Ulaby et al., 1982*). Also, Equations 5.3 and 5.4 were derived without considering the effects of noise and clutter. Noise contributions to estimates of echo power can be accounted for and removed (*Ulaby et al., 1982*). For radar systems like SuperDARN, contributions of self-clutter to estimates of power are negligible for properly designed transmitted waveforms (as will be discussed in detail later). In general, the presence of noise and clutter will increase the MSE of the power estimator in a manner given by the RHS of Equation 5.2 (e.g. *Farley, 1969*). To simplify the discussion that will follow as much as possible effects due to noise will not be included in derivations. It is a simple exercise to add contributions of noise to the final results that will be presented.

Radar ACFs combined with fitting techniques (e.g. *Ribeiro et al.*, 2013a) are used to extract ionospheric parameters from radar measurements. The individual data points of the estimated radar ACF are referred to as “lag-products” or simply “lags”. To estimate a lag of the ACF at time t ($R(t)$), the average of the correlation between K pairs of complex voltage samples, with each voltage sample in each pair separated in time by t , is calculated using

$$\hat{R}(t) = \frac{1}{K} \sum_{k=1}^K \tilde{V}_{1k} \tilde{V}_{2k}^*, \quad (5.5)$$

where $\hat{R}(t)$ is the estimator for $R(t)$ and the subscripts 1 and 2 each refer to one of the two complex voltage samples separated by a time of t . The MSE of the ACF estimator is given by

$$\text{MSE}(\hat{R}(t)) = \frac{P_r^2}{K}. \quad (5.6)$$

Figure 5.1 shows a typical ACF estimated from SuperDARN voltage measurements using the 8 pulse multiple-pulse sequence described in Figure 5.4. It is important to note that the estimate of the radar ACF at lag $t = 0\tau$, called lag-zero, is equivalent to an estimate of the echo power. Both are often referred to as the “lag-zero power” and so will be referred to as lag-zero power in this study. As a final comment, lag-zero power measurements may be made with a multiple-pulse sequence by adding a large time delay between the first and second pulse of the pulse sequence. Ideally, the time delay should be large enough for the signal to propagate out to and return from the largest range to be measured. For the sequence plotted in Figure 5.4, $\tau = 1.8$ ms is sufficient for lag-zero power measurements out to a range of 3780 km. Detailed discussion of Figures 5.1 and 5.4 follow in the next section.

5.4 Self-Clutter Estimation

The origins of self-clutter can be more readily understood through examination of an example multiple-pulse sequence shown in the Farley diagram plotted in Figure 5.2, which was adapted from Figure 1 of *Farley* (1972). Figure 5.2 was made using $\tau = 2.4$ ms, $t_p = 300$ μ s, $N_p = 3$, and PTAB = $[0\tau, 1\tau, 3\tau]$. Elapsed time since the start of the first transmitted pulse is plotted

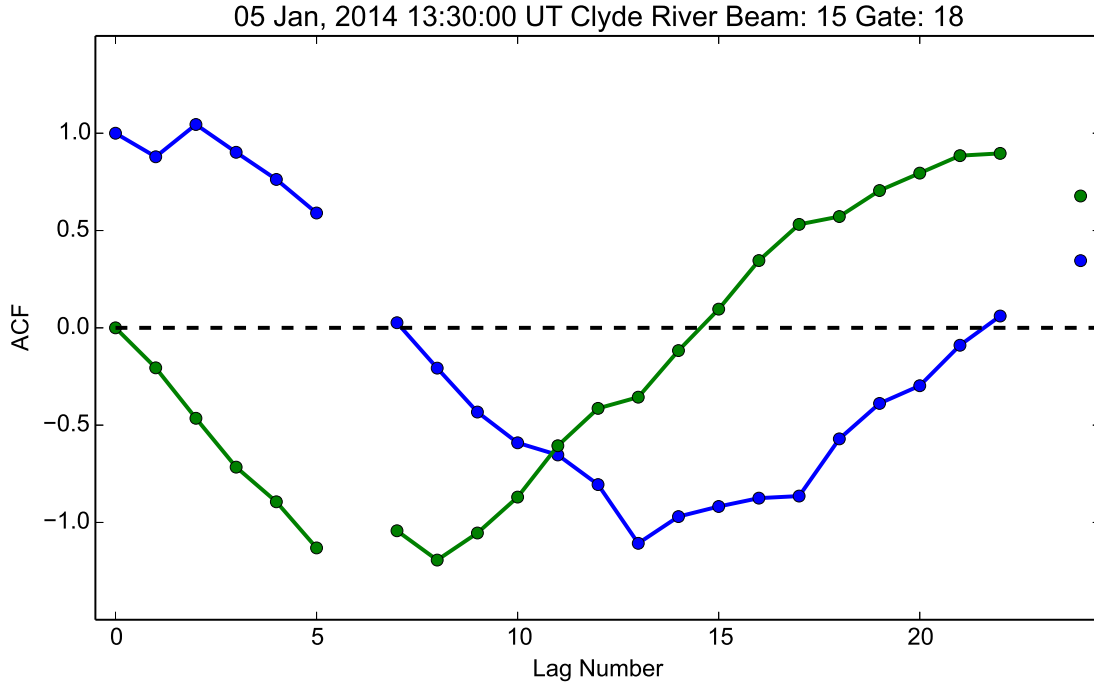


Figure 5.1: An estimated ACF derived from voltage samples measured by the Clyde River SuperDARN radar using the 8 pulse multiple-pulse sequence illustrated in the Farley Diagram in Figure 5.4. The blue curve and the green curve indicate the real and imaginary parts of the estimated ACF, respectively. The missing points are caused by missing lags in the multiple-pulse sequence.

along the x-axis and range from the radar is plotted along the y-axis (see caption for more details). Notice in Figure 5.2 that each received voltage sample (outlined in blue) contains contributions from both the range of interest (red diamonds) and unwanted interfering ranges (self-clutter, marked by black diamonds). In Figure 5.2, if a power range-profile yielded a narrow band of scatter centered around the ranges near the black diamond 1b, increasing (decreasing) τ would move 1b further (closer) in range and out of the region where it would contribute to self-clutter. From this example, it is evident that self-clutter varies with the lag-zero power range-profile of the ionosphere and the value of τ for the multiple-pulse sequence being used.

The zero-Doppler (or delay) ambiguity function (*Woodward, 1960*) for the same 3 pulse pulse-sequence is plotted in Figure 5.3. The delay ambiguity function is the modulus of the auto-correlation function of the transmitted waveform in Figure 5.2, equivalent to collapsing the Farley diagram in Figure 5.2 along the time axis. In contrast the the Farley diagram,

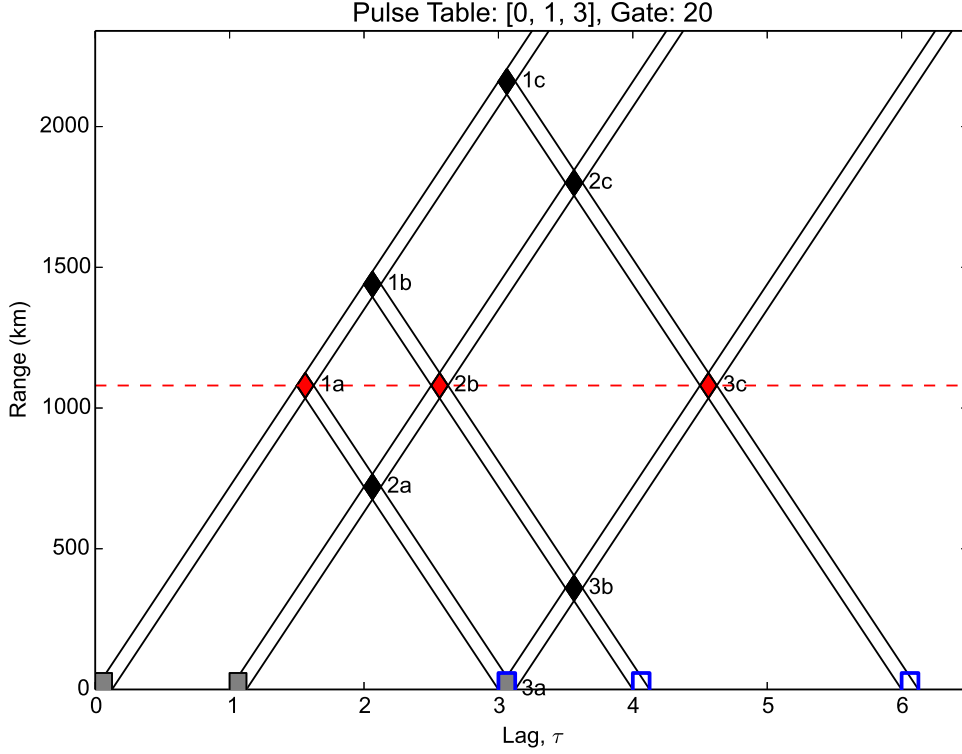


Figure 5.2: A Farley diagram (c.f. Figure 1 in *Farley* (1972)) showing an example of a multiple-pulse sequence using 3 pulses with $\tau = 2.4$ ms, $t_p = 300$ μ s, $N_p = 3$, and PTAB = $[0\tau, 1\tau, 3\tau]$. The grey filled rectangles indicate transmission pulses and the rectangles with blue outlines indicate received voltage samples (from scattered signals returning to the radar). The red diamonds indicate signal backscattered from a repeated range of interest, in contrast to the black diamonds indicating signal backscattered from non-repeated unwanted ranges.

which explicitly details the ambiguity in each voltage sample, the delay ambiguity function summarizes the ambiguity in the 3 pulse pulse-sequence. The small peaks in Figure 5.3 (1c, 2c, 1b, 2a, 3b, 3a) indicate delays (ranges) that may contribute range ambiguity to a measurement made with the 3 pulse pulse-sequence. Varying τ changes the separation in delay (range) between range ambiguities in the same way as discussed above. Figure 5.3 reiterates the fact that if echoes exist at an ambiguous range, as can be determined by measuring the power range-profile, it may be possible to vary τ in such a manner as to move the ambiguous range into a region with weaker or non-existent echoes, thereby limiting or

removing the self-clutter contribution to the radar ACF. The exact variation of τ would need to be balanced with maximum unambiguous Doppler shift requirements, since decreasing (increasing) τ will increase (decrease) the maximum unambiguous Doppler shift of the pulse-sequence; an effect that can be understood by studying the Doppler (zero-delay) ambiguity function (not shown).

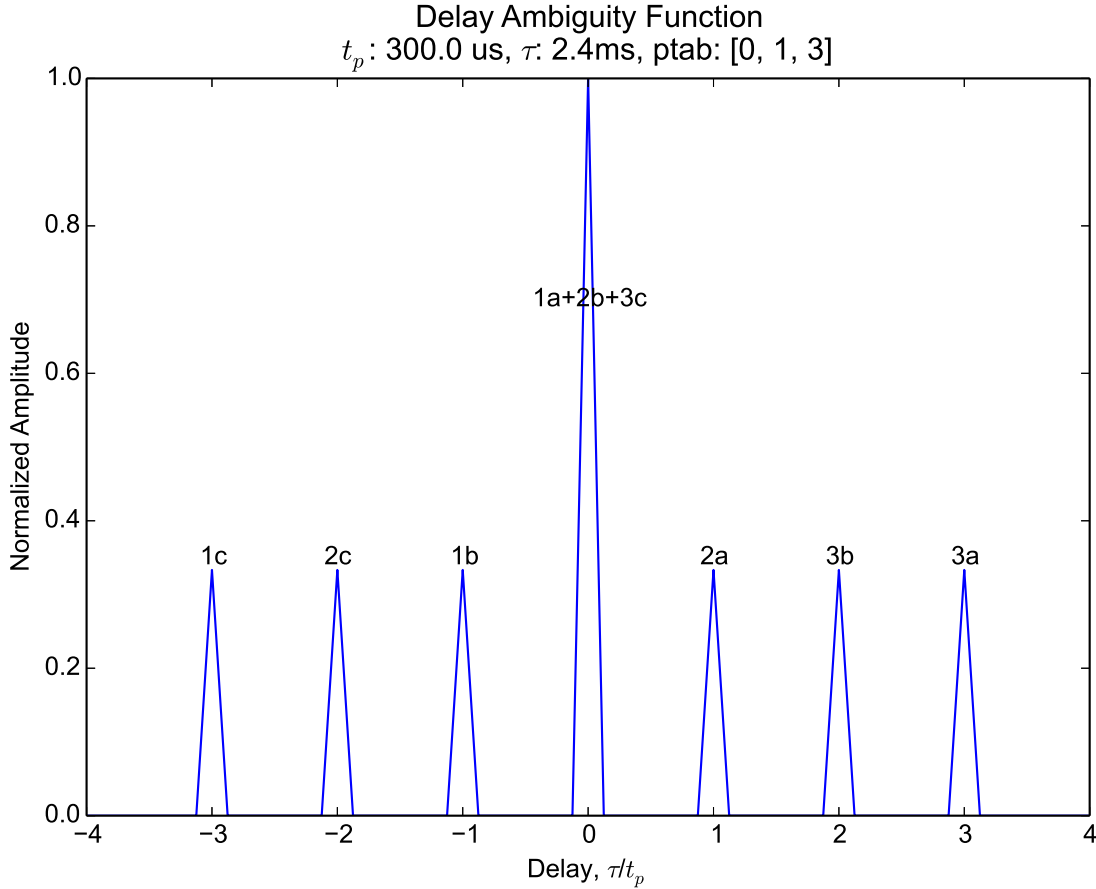


Figure 5.3: The delay ambiguity function for the pulse-sequence in Figure 5.2.

While techniques for suppressing or removing self-clutter are discussed in the literature for other radar measurement techniques (e.g. *Zrnic et al.*, 1999; *Rihaczek*, 1965), the authors are not aware of any techniques for *removing* self-clutter for the multiple-pulse technique. *Grydeland and Gustavsson* (2011) discusses a method for suppressing self-clutter by selectively transmitting pulses at orthogonal polarizations. This technique is only possible on radar systems capable of polarized transmission and reception, but is able to decrease measurement errors by a factor of 4. *Ponomarenko and Waters* (2006) discussed a method for

identifying the lags of a radar ACF that are influenced by self-clutter in an effort to remove lags that contained “too much” self-clutter. The detection criteria was based on individual comparisons of received lag-zero power between interfering ranges and the range of interest. At each range, for each lag, if the lag-zero power of any interfering range in the lag was found to be greater than the lag-zero power of the range of interest, then that lag was discarded. After implementing this detection criteria, *Ponomarenko and Waters* (2006) noted that a more realistic value of the spectral width was obtained with the current SuperDARN fitting technique. However, rather than discarding lags of the estimated ACF, it is more desirable to estimate the self-clutter contained in a lag such that the signal-derived error in the lag estimate can be obtained (recall Equation 5.2). Then it is possible to use an error-weighted fitting technique to extract ionospheric parameters from ACF estimates in a self-consistent manner (without requiring externally imposed criteria). In addition, as is done for noise, accurate estimates of self-clutter may be subtracted from the radar ACF to extract the desired signal ACF.

5.4.1 Self-Clutter Power in a Voltage Sample

For a pulse sequence with N_p pulses, *Farley* (1972) stated that when operating at incoherent scatter radar frequencies (VHF and UHF) and with short measurement time scales (the entire pulse sequence is transmitted in < 1 ms), an estimate of the self-clutter power (C) contained in any single voltage sample is

$$C \approx (N_p - 1)S, \quad (5.7)$$

assuming that the scattering volume over which the pulse sequence extends returns an average signal power S for a single transmitted pulse and that all N_p pulses have been transmitted before the voltage samples are measured. It is also assumed that interfering ranges return statistically independent voltages such that on average, products of voltages from different range gates are zero. Examining the voltage sample at lag 6 in the Farley diagram in Figure 5.2, with $N_p = 3$ pulses transmitted, it is observed that the voltage sample contains $(N_p - 1) = 2$ interfering ranges. Given a uniform lag-zero power range-profile, where each

range returns a power of S , the self-clutter in the voltage sample is therefore $(N_p - 1)S$, as given by Equation 5.7.

Some assumptions made in *Farley* (1972) do not hold for HF radars. Due to the multiple-hop over-the-horizon propagation at HF frequencies, the lag-zero power range-profile for HF radars is never uniform. Only interfering ranges that return non-zero lag-zero power will contribute to the self-clutter. With these considerations, Equation 5.7 can be generalized to

$$C = \sum_{n=1}^N S_n, \quad (5.8)$$

where N is the number of interfering ranges and S_n is the lag-zero power at interfering range n . For a voltage sample received after all pulses have been transmitted $N = N_p - 1$ and with a uniform power range-profile $S_n = S$, we recover Equation 5.7 from Equation 5.8. It is apparent that for the generalized result the signal-to-clutter ratio (SCR) will vary with both the multiple-pulse sequence used and the lag-zero power range-profile in contrast to Equation 5.7, which states that SCR is solely a function of the pulse sequence.

5.4.2 Constructing the Self-Clutter Estimator

The generalization of Equation 5.7, combined with the process of how the radar ACF is estimated from complex voltage samples, can be used to obtain an estimate of the self-clutter contribution to the lags of the radar ACF. With multiple transmitted pulses, each voltage sample measured by a radar contains a voltage contribution from the range of interest but also from interfering ranges. Referring to Figure 5.2, recall that in general the radar ACF at time t is the correlation between two voltage samples separated in time by t . For instance, the radar ACF at $t = 2\tau$ is the product of the voltage sample at $t = 4\tau$ (\tilde{V}_1) and the complex conjugate of the voltage sample at $t = 6\tau$ (\tilde{V}_2). As verified by Figure 5.2, the voltage samples \tilde{V}_1 and \tilde{V}_2 are the sums of individual voltage contributions from different ranges: 1 from the range of interest (2b for \tilde{V}_1 and 3c for \tilde{V}_2), and 2 from interfering ranges (1b and 3b for \tilde{V}_1 and 1c and 2c for \tilde{V}_2). In general, for an arbitrary number of pulses transmitted, \tilde{V}_1 and \tilde{V}_2 can be written as

$$\tilde{V}_1 = \tilde{V}_{R1} + \sum_{n=1}^N \tilde{V}_n, \quad \tilde{V}_2 = \tilde{V}_{R2} + \sum_{m=1}^M \tilde{V}_m \quad (5.9)$$

where \tilde{V}_n and \tilde{V}_m correspond to voltage contributions from N and M interfering ranges contained in voltage samples \tilde{V}_1 and \tilde{V}_2 , respectively and \tilde{V}_{R1} and \tilde{V}_{R2} represent the voltage contribution from the range of interest in voltage samples \tilde{V}_1 and \tilde{V}_2 , respectively. Correlating \tilde{V}_1 and \tilde{V}_2 to obtain the radar ACF at lag 2 will result in 9 different correlations: 1 from the range of interest (referring to Figure 5.2 and representing the correlation with the notation [2b,3c]), 4 between the interfering ranges and the range of interest ([2b,2c], [2b,1c], [1b,3c], [3b,3c]), and 4 between the interfering ranges ([3b,2c], [3b,1c], [1b,2c], [1b,1c]) which can be written algebraically as,

$$\tilde{V}_1 \tilde{V}_2^* = \tilde{V}_{R1} \tilde{V}_{R2}^* + \sum_{n=1}^N \tilde{V}_n \tilde{V}_{R2}^* + \sum_{m=1}^M \tilde{V}_{R1} \tilde{V}_m^* + \sum_{n=1}^N \sum_{m=1}^M \tilde{V}_n \tilde{V}_m^*. \quad (5.10)$$

Examining Equation 5.10, it is clear that the first term of the RHS corresponds to the signal ACF and the remaining terms on the RHS must necessarily be the self-clutter. For a multiple-pulse sequence with N_p transmitted pulses, the self-clutter contribution may yield up to as many as $N_p^2 - 1$ correlations.

Figure 5.4 shows a Farley diagram of the 8-pulse multiple-pulse sequence, known as *katscan*, routinely used by SuperDARN radars. Notice that different voltage samples contain a different number of interfering range contributions (cf. the first and last voltage samples). Even though *katscan* is a more complicated pulse-sequence with a different interfering range structure, Equation 5.10 remains valid. For a given range, selecting two voltage samples for \tilde{V}_1 and \tilde{V}_2 and identifying the N and M interfering range contributions for each, respectively, allows for an estimate of the self-clutter.

For Equation 5.10, it has been assumed that all of the pulses are transmitted with equal amplitude (such that no variations in the received voltages are due to variations in the transmitted pulse power) which is a reasonable assumption for the SuperDARN radars. It is also assumed that the plasma processes responsible for backscatter are stationary with respect to measurement timescales. For identifying the correlations due to self-clutter, it is not necessary to make assumptions about the uniformity of the scattering volume; the

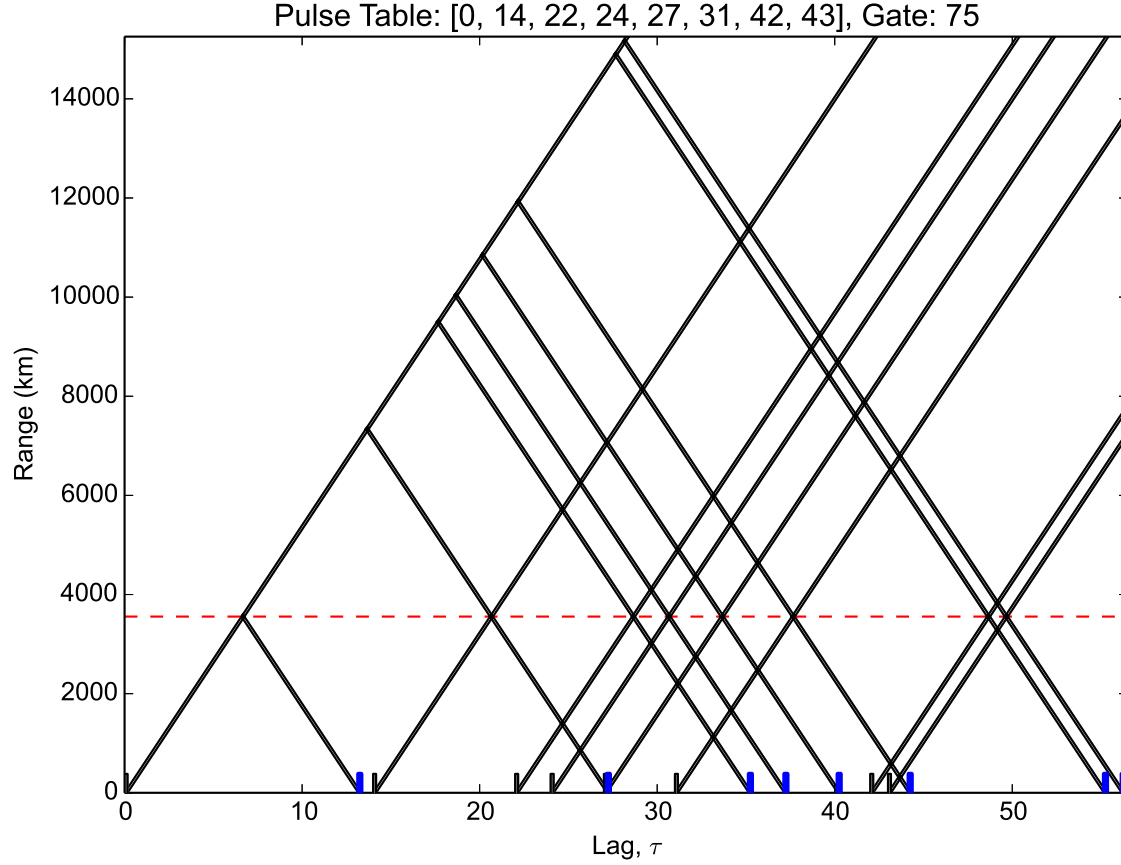


Figure 5.4: The 8-pulse multiple-pulse sequence *katscan* with $\tau = 1.8$ ms, $t_p = 300$ μ s, $N_p = 8$. The dashed line indicates the location of the range of interest. Near the x-axis, the black vertical lines at the base of positive sloped lines indicate transmission pulses and the blue vertical lines at the base of negative sloped lines indicate received samples. Intersections of the slanted black lines that do not lie on the dashed line indicate where signal is received from interfering ranges.

lag-zero power range-profile need not be, on average, constant in range. Finally, it is not necessary to assume that all N_p pulses have been transmitted before receiving any returns as was required for Equation 5.7. For pulse sequence transmission times longer than the time required for a pulse to return from the range of interest (c.f. Figure 5.4), the number of pulses “in the air” will be a function of the measurement time of the voltage sample and the pulse spacings in the pulse sequence.

5.4.3 Generalized Self-Clutter Estimator

Given the preceding subsection, we now explicitly write a generalized estimator for the self-clutter contained in a lag-product of the radar ACF. From K transmitted pulse-sequences, the self-clutter terms in Equation 5.10 are averaged to form a self-clutter estimator given by

$$\hat{C} = \frac{1}{K} \sum_{k=1}^K \left(\sum_{n=1}^N \tilde{V}_{nk} \tilde{V}_{R2k}^* + \sum_{m=1}^M \tilde{V}_{R1k} \tilde{V}_{mk}^* + \sum_{n=1}^N \sum_{m=1}^M \tilde{V}_{nk} \tilde{V}_{mk}^* \right). \quad (5.11)$$

However only the \tilde{V}_1 and \tilde{V}_2 voltage samples are known (measured). Notice that evaluation of Equation 5.11 requires knowledge of the \tilde{V}_{R1} and \tilde{V}_{R2} voltage contributions, quantities that are impossible to measure directly (hence the motivation of this study). The \tilde{V}_n and \tilde{V}_m contributions are also unknown. \tilde{V}_{R1} and \tilde{V}_{R2} can be removed by rearranging the expressions in Equation 5.9 and substituting the results into Equation 5.11 to obtain the voltage-based self-clutter estimate (VSE)

$$\hat{C}_{\text{VSE}} = \frac{1}{K} \sum_{k=1}^K \left(\sum_{n=1}^N \tilde{V}_{nk} \tilde{V}_{2k}^* + \sum_{m=1}^M \tilde{V}_{1k} \tilde{V}_{mk}^* - \sum_{n=1}^N \sum_{m=1}^M \tilde{V}_{nk} \tilde{V}_{mk}^* \right), \quad (5.12)$$

where \tilde{V}_n and \tilde{V}_m are approximated using voltage samples from lag-zero power measurements. Generally, this approximation will introduce extra uncertainty into the VSE, unless the correlations involving \tilde{V}_n and \tilde{V}_m have a long lifetime (large constant coherence) and a large period (constant phase), as is the case for ground scatter or slow moving long-lived ionospheric irregularities. If either or both the lifetime and period of the correlations are short, the approximation will result in increased errors in the VSE. Equivalently, the condition can be written as $t_3 \gg |t_1 - t_2|$ for the approximation to be accurate, where t_1 is the lag time of the actual correlation being approximated, t_2 is the lag time between the voltage samples used to create the approximation of the correlation, and t_3 is the time scale of the rate of change of the correlation between the two ranges causing the self-clutter. A further consequence of this approximation is that the VSE is unable to account for self-clutter originating from ranges that return echoes when said ranges are not included in lag-zero power measurements. Recall that the first pulse in a multiple-pulse sequence is usually used to provide lag-zero power estimates, but only until the second pulse is transmitted. For HF radar systems like

SuperDARN, it is unlikely that beyond 5000 km echoes will be received due to propagation and signal attenuation effects. The true range will depend on radar specifications and is something that may be experimentally determined for each individual radar system.

Up to N_p of the approximated correlations share a lag-time that is identical to the lag-time in the actual correlations. For example, Figure 5.2 shows that for the voltage samples at $t = 4\tau$ and $t = 6\tau$, N_p of the correlations, specifically [3b,3c], [2b,2c], and [1b,1c] (recall the discussion surrounding Equation 5.10), share a lag-time that is common with the voltages samples from lag-zero power measurements that are used to approximate the correlations. Assuming stationarity, regardless of assumptions about coherence and Doppler characteristics, up to N_p of the approximated correlations will be accurate within statistical sampling uncertainty. However, the remaining approximated correlations have a lag-time that differs from the actual lag time, meaning that unless the long coherence and constant phase assumptions hold, such approximations will be inaccurate (larger uncertainty than statistical sampling uncertainty).

The MSE of the VSE in Equation 5.12 is given by

$$\text{MSE}(\hat{C}_{\text{VSE}}) = \frac{C^2}{K}, \quad (5.13)$$

where C is the self-clutter power. To simplify the derivation of this estimate, \tilde{V}_n and \tilde{V}_m were assumed to be known. This means that Equation 5.13 will generally underestimate the error in Equation 5.12 due to correlations that do not satisfy the long coherence and constant phase assumptions made above. In deriving Equation 5.13 it was also assumed that all ranges that contribute to self-clutter have been included in \tilde{V}_n and \tilde{V}_m via the lag-zero power voltage samples.

Despite the limitations on the applicability of the VSE and its MSE, as will be shown, the estimator is able to accurately account for ground scatter origin self-clutter. This capability allows for removal of ground scatter self-clutter from the radar ACF obtained by SuperDARN radars that often obtain measurements of ionospheric irregularities contaminated by ground scatter self-clutter. In these situations, first order error estimates on estimates of the radar ACF are now possible.

5.4.4 Maximal Self-Clutter Estimators

We shall now discuss two maximal self-clutter estimators for use when the VSE cannot be accurately implemented. The maximal estimators estimate the upper limit or “worst-case” self-clutter contained in an estimate of the radar ACF and enable first order error estimates for radar ACF estimates (using Equation 5.13) to be self-consistently obtained, even if the VSE cannot be used.

First a maximal voltage-based self-clutter estimator (MVSE) was derived by replacing each voltage correlation in Equation 5.11 with their respective magnitudes (e.g. $|\tilde{V}_n \tilde{V}_m|$) and assuming complete correlation. Equivalently, each voltage correlation is replaced with the product of the magnitudes of the contributing voltage samples, which can be written in terms of a product of voltage magnitudes ($|\tilde{V}_n| |\tilde{V}_m|$) and normalized coherence function (ρ) with the added assumption of complete coherence ($\rho = 1$). This substitution enables identification of the upper limit on the magnitude of (worst case scenario) self-clutter. Algebraically the MVSE is written,

$$\hat{C}_{\text{MVSE}} = \frac{1}{K} \frac{4}{\pi} \sum_{k=1}^K \left(\sum_{n=1}^N |\tilde{V}_{nk}| |\tilde{V}_{Rk}| + \sum_{m=1}^M |\tilde{V}_{Rk}| |\tilde{V}_{mk}| + \sum_{n=1}^N \sum_{m=1}^M |\tilde{V}_{nk}| |\tilde{V}_{mk}| \right). \quad (5.14)$$

The magnitude of each complex voltage sample follows a Rayleigh distribution of order 2 and has an expected value of $(\sqrt{\pi}/2)|V|$ (*Simon, 2007*). The factor of $4/\pi$ in Equation 5.14 is required to normalize the factor of $\pi/4$ that results when calculating the expected value of \hat{C}_{MVSE} (see below). In contrast to the VSE, the MVSE is real valued and is only an estimate of the upper limit of self-clutter *power*; it cannot be used for estimating the phase of the self-clutter.

Historically, due to limitations on data storage it was not always possible to store complex voltage measurements. This is currently the case for most SuperDARN radars where the typical mode of operation only records estimates of radar ACFs using Equation 5.5, and estimates of lag-zero power, using Equation 5.3. In this case, a power-based maximal self-clutter estimator (MPSE) is therefore desired if error estimates are to be calculated. The MPSE can be derived from the MVSE using the Cauchy Inequality (*Abramowitz and Stegun, 2012*) and the power estimator given by Equation 5.3 to obtain,

$$\hat{C}_{\text{MPSE}} = \sum_{n=1}^N \sqrt{\hat{P}_n \hat{P}_R} + \sum_{m=1}^M \sqrt{\hat{P}_R \hat{P}_m} + \sum_{n=1}^N \sum_{m=1}^M \sqrt{\hat{P}_n \hat{P}_m}, \quad (5.15)$$

for the MPSE, where the symbol \hat{P}_i indicates the average lag-zero power estimator (a quadratic average of voltage samples) from a range i . In the absence of lag-zero power estimates, lag-zero estimates may be inferred from the radar ACF using fitting techniques (e.g. *Ribeiro et al.*, 2013a), enabling implementation of the MPSE.

The expected value of the maximal self-clutter estimators is given by

$$C_{\text{MAX}} = \sum_{n=1}^N \sqrt{P_n P_R} + \sum_{m=1}^M \sqrt{P_R P_m} + \sum_{n=1}^N \sum_{m=1}^M \sqrt{P_n P_m}, \quad (5.16)$$

where the symbol P_i indicates the expected value of the lag-zero power from range i . The notation R , n , and m is as before. This is an intuitive result that arises from the fact that $\sqrt{P} = |V|$ assuming unit resistance.

The MSE of the MxSE estimator (where x is either V or P for either voltage-based or power-based, respectively), following the same notation with additional indicies a and b, is given by,

$$\begin{aligned} \text{MSE}(\hat{C}_{\text{MxSE}}) &= \sum_{n=1}^N \sum_{a=1}^N \alpha_{na} P_R \sqrt{P_n P_a} + \sum_{m=1}^M \sum_{b=1}^M \alpha_{mb} P_R \sqrt{P_m P_b} + 2 \sum_{n=1}^N \sum_{m=1}^M \alpha_{n0} P_R \sqrt{P_n P_m} \\ &+ 2 \sum_{n=1}^N \sum_{a=1}^N \sum_{m=1}^M \beta_{na} \sqrt{P_n P_a P_m P_R} + 2 \sum_{n=1}^N \sum_{m=1}^M \sum_{b=1}^M \beta_{mb} \sqrt{P_n P_m P_b P_R} + \sum_{n=1}^N \sum_{a=1}^N \sum_{m=1}^M \sum_{b=1}^M \gamma_{namb} \sqrt{P_n P_a P_m P_b}. \end{aligned} \quad (5.17)$$

To obtain the MSE of the voltage-based maximal self-clutter estimator, the following coefficients must be used with Equation 5.17,

$$\alpha_{na} = \begin{cases} (1/K) ((4/\pi)^2 - 1) & n = a \\ (1/K) ((4/\pi) - 1) & n \neq a \end{cases}$$

$$\beta_{na} = \begin{cases} (1/K) ((4/\pi) - 1) & n = a \\ 0 & n \neq a \end{cases}$$

$$\gamma_{nmab} = \begin{cases} (1/K) ((4/\pi)^2 - 1) & n = a, m = b \\ (1/K) ((4/\pi) - 1) & n \neq a, m = b \text{ or } n = a, m \neq b \\ 0 & n \neq a, m \neq b \end{cases}, \quad (5.18)$$

and to obtain the MSE of the power-based maximal self-clutter estimator, the following coefficients must be used with Equation 5.17,

$$\begin{aligned} \alpha_{na} &= \begin{cases} 1/K & n = a \\ \sqrt{1 + 1/K} - 1 & n \neq a \end{cases} \\ \beta_{na} &= \begin{cases} \sqrt{1 + 1/K} - 1 & n = a \\ 0 & n \neq a \end{cases} \\ \gamma_{nmab} &= \begin{cases} 1/K & n = a, m = b \\ \sqrt{1 + 1/K} - 1 & n \neq a, m = b \text{ or } n = a, m \neq b \\ 0 & n \neq a, m \neq b \end{cases}. \end{aligned} \quad (5.19)$$

Similar to the MSE for the VSE, the MSE for the MVSE and MPSE will underestimate the error in any maximal self-clutter estimate where interfering ranges that contribute are not included in lag-zero power voltage measurements. This is unlikely to be a problem for typical SuperDARN operation modes that measure out to 3500 or 4000 km, beyond which back-scattered signal strength has been degraded too much by cumulative signal propagation and attenuation effects. The largest practical range can easily be investigated for each radar utilizing an operating mode designed to check for the largest ranges that echos are returned from.

5.5 Testing and Discussion

The radar data simulator of *Ribeiro et al.* (2013b) was modified by the authors of the current study so it could be used to evaluate the self-clutter estimators presented in the previous section. Using the simulator it was possible to compare the simulated actual self-clutter with estimates of self-clutter derived from simulated complex voltage samples.

The assumptions made in deriving Equation 5.12 were tested with the results presented in Figure 5.5. To test the assumptions, two cases were simulated: one with self-clutter of

ground scatter origin, and the other with self-clutter of ionospheric origin. The simulator was run using the katscan multiple-pulse sequence with $\tau = 1.8$ ms and $K = 18$ (similar to parameters in common SuperDARN operations) and the simulated scattered signal strength from each range gate were all set to be equal.

The top set of panels in Figure 5.5 show that for self-clutter caused by interfering ranges filled with ground scatter the VSE is able to accurately extract the signal ACF from the radar ACF. Note that the top left panel shows the ACF with no self-clutter mitigation (current technique), the top middle panel shows the same ACF after the new method (subtracting the VSE), and the top right panel shows the ideal (true) ACF that should result if the technique were perfect. The extracted signal ACF (top middle) agrees with the simulated actual signal ACF (top right) within error. The difference between the actual ACF and the extracted signal ACF (\hat{R}_S) is quantitatively compared with the relative error in Table 5.1. The error is given by $\sqrt{\langle(\delta\hat{R}_S)^2\rangle}$ where

$$\langle(\delta\hat{R}_S)^2\rangle = \frac{1}{K} [(S + C)^2 + C^2] \quad (5.20)$$

with the relative error given by $\sqrt{\langle(\delta\hat{R}_S)^2\rangle}/S$. Comparing the differences with the expected error shows that there is agreement between the extracted signal ACF and the simulated actual signal ACF within the expected relative error. The bottom panels in Figure 5.5 show that the VSE is unable to accurately extract the signal ACF when self-clutter originates from interfering ranges filled with ionospheric scatter.

These results echo the previously discussed consequences of the assumptions made in deriving the VSE, which require that self-clutter originate from ground echoes if the VSE is to be used accurately. The ground scatter in the simulation had zero velocity and zero spectral width, but it must be noted that in practice, ground scatter signals propagate through the ionosphere which may impart small non-zero Doppler shifts due to particle precipitation and/or vertical oscillations in the ionosphere (e.g. *Scoular et al.*, 2013) and/or small non-zero spectral width due to wave front decorrelation caused by random density variations in the ionospheric plasma (e.g. *Vallières et al.*, 2004). Nevertheless, as shown below, such small non-zero enhancements to the Doppler shift and spectral width of ground scatter do not prevent the VSE from removing self-clutter from measurement-based radar ACFs.

Lag Number	0	1	2	3	4	5	7	8	9	10	11	12
Diff. (Re)	0.00	0.37	0.02	0.15	0.21	0.02	0.03	0.02	0.06	0.06	0.05	0.02
Diff. (Im)	0.00	0.05	0.12	0.19	0.10	0.06	0.02	0.11	0.08	0.17	0.20	0.14
Error	0.24	0.37	0.44	0.83	0.55	0.59	0.59	0.25	0.43	0.26	0.61	0.34
Lag Number	13	14	15	16	17	18	19	20	21	22	24	
Diff. (Re)	0.04	0.00	0.15	0.13	0.00	0.27	0.27	0.00	0.03	0.01	0.00	
Diff. (Im)	0.01	0.00	0.08	0.05	0.11	0.01	0.00	0.13	0.14	0.04	0.19	
Error	0.25	0.24	0.82	0.52	0.30	0.58	0.34	0.47	0.44	0.27	0.29	

Table 5.1: Extracted Signal ACF Compared with Actual Signal ACF: A quantitative comparison of the absolute value of the difference between the real and imaginary components for all lags of the extracted signal ACF and the simulated actual ACF depicted in Figure 5.5 (the top middle and top right panels, respectively). The differences are compared with the relative error defined by Equation 5.20.

A second test of the VSE was performed using measured complex voltage samples recorded by the Saskatoon SuperDARN radar on 14 July, 2014 at 6:01:07 UT. Figure 5.6 presents the use of the VSE on measured voltage samples from range gate 25 on beam 2. The figure shows the magnitude (bottom row) and phase (top row) of a typical complex radar ACF obtained for an ionospheric target with self-clutter originating from ranges filled with ground scatter, both with (left column) and without (right column) the VSE subtracted from the radar ACF. The usage of the triangular markers in the plots will be discussed shortly.

Notice that subtraction of the VSE has reduced the fluctuations in the ACF phase for the first few lags. Also notice that subtraction of the VSE significantly suppresses the large peaks in lag power caused by self-clutter at lags 11, 12, 15, 16, 18, 20, and 21. It is clear that the magnitude and phase of the ACF have improved due to subtraction of the VSE, a result that is consistent with the ideal case simulated in Figure 5.5. The ACF lag power is more greatly affected by removal of the VSE than the ACF phase since ground scatter is characterized by nearly zero phase (small Doppler shift) and large correlation times (small spectral width). This means that ground scatter dominated contributions of self-clutter on the estimates of the radar ACF have little effect on the estimated ACF *phase*, as confirmed in Figure 5.6. The self-clutter at large lags, although significantly improved is not completely

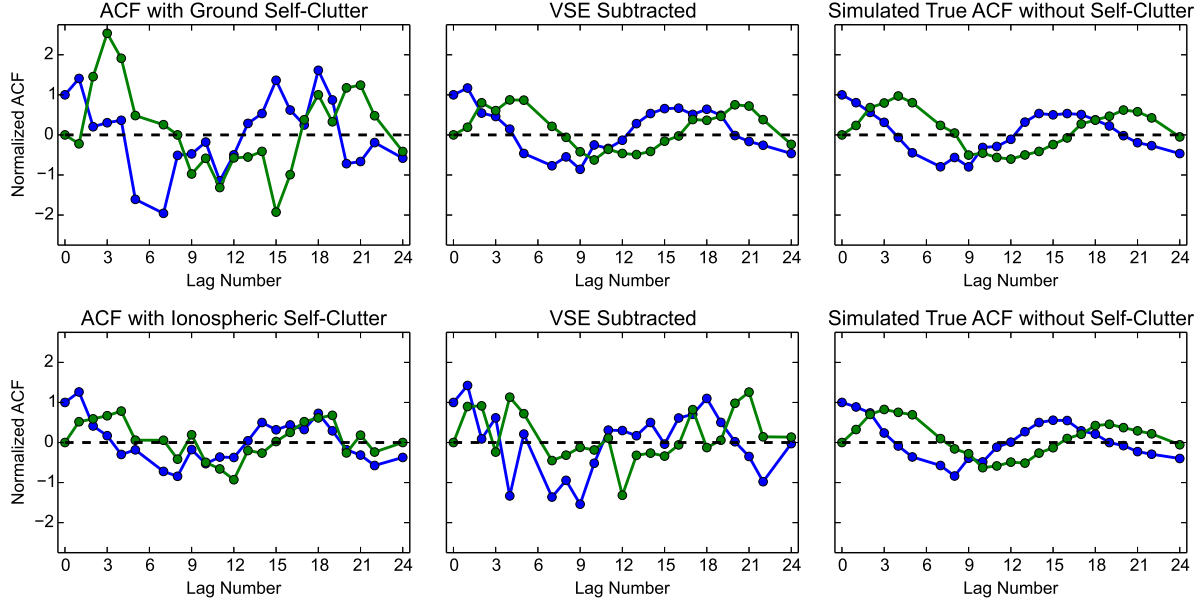


Figure 5.5: Plots comparing the use of the VSE on simulated radar measurements for estimating self-clutter originating from ground echoes (top panels) and for self-clutter originating from ionospheric echoes (bottom panels). For the top panels, all interfering ranges were filled with ground echoes (targets with a velocity of 0 m/s and a spectral width of 0 m/s) and for the bottom panels, all interfering ranges were filled with ionospheric echoes (targets with a velocity of 500 m/s and a spectral width of 100 m/s).

removed, as it was for the simulation in Figure 5.5. This result is expected as it is likely caused by enhancements of the ground scatter spectral width due to wave front decorrelation of the propagating radio waves (*Vallières et al.*, 2004).

Finally, the effect of the VSE on current SuperDARN fitted radar data (extracted ionospheric parameters) was examined. The VSE was applied to 4 days of complex voltage samples collected by the Saskatoon SuperDARN radar (from 10 July, 2014 at 10:00 UT until 14 July, 2014 at 10:00 UT). Ionospheric parameters were then extracted using the standard SuperDARN fitacf software (*Ribeiro et al.*, 2013a) both with and without subtraction of the VSE (for examples of fits of SuperDARN radar data, see *Ribeiro et al.* (2013a)).

Distributions of the differences in fitted data products between the fitted radar data processed with and without VSE subtraction were plotted (see Figure 5.7) and the full-width half-maximum (FWHM) of each distribution was determined. The FWHM of the difference distribution for ionospheric velocities was 6.2 m/s, for ground velocities was 9.1 m/s, for

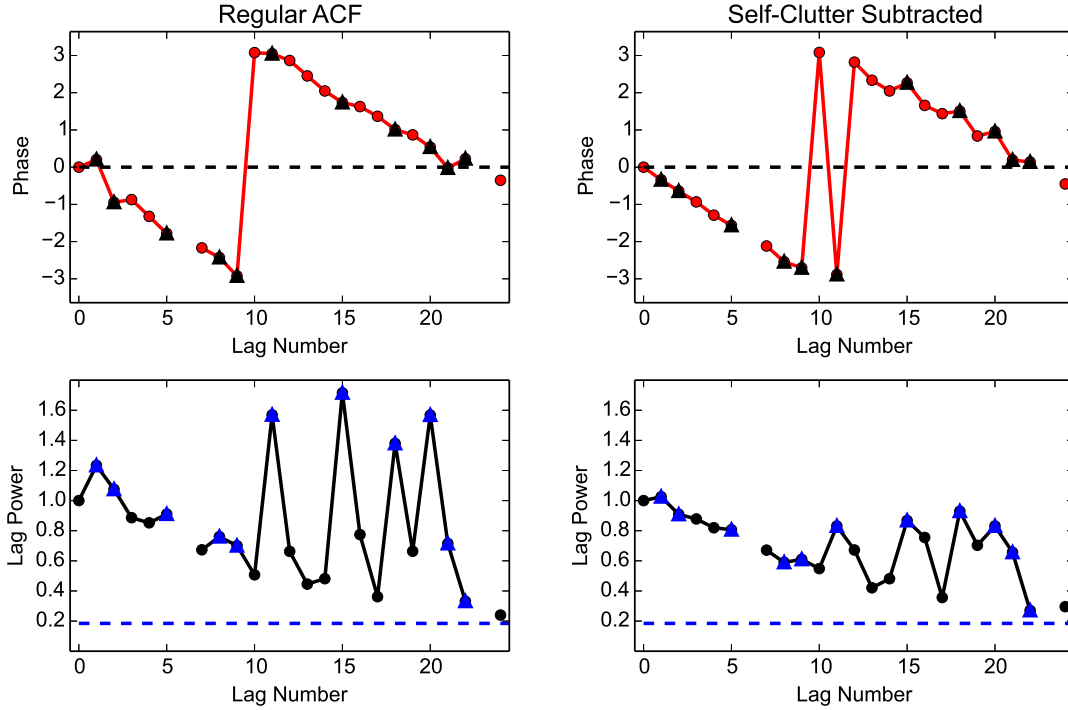


Figure 5.6: A comparison between measured SuperDARN radar ACF data both without (left column) and with (right column) the self-clutter estimate of Equation 5.12 removed. The ACF was recorded on 14 July, 2014 at 6:01:07 UT in gate 25 and beam 2. The top row of panels show ACF phase and the bottom row show ACF power. The triangular data points indicate data that is currently discarded by the self-clutter identification algorithm of *Ponomarenko and Waters (2006)* (the current technique used by SuperDARN to remove data contaminated by self-clutter).

ionospheric spectral width was 7.0 m/s, and for ground spectral width was 12.3 m/s. The distributions of differences are narrower for the velocity and spectral width of ionospheric echoes compared with the same distributions for ground scatter echoes. This is caused, in part, by the assumption that voltage samples at interfering ranges can be approximated with voltage samples from lag-zero power measurements. Due to the nature of HF radio propagation, ground echoes and ionospheric echoes return from different ranges in a manner such that self-clutter affecting ground scatter usually originates from interfering ranges containing ionospheric echoes and self-clutter affecting ionospheric scatter usually originates from interfering ranges containing ground scatter. Self-clutter caused by ground scatter signals satisfies the assumptions made in deriving Equation 5.12 better than self-clutter caused by ionospheric signals, leading to larger errors in the estimation of the self-clutter that affects

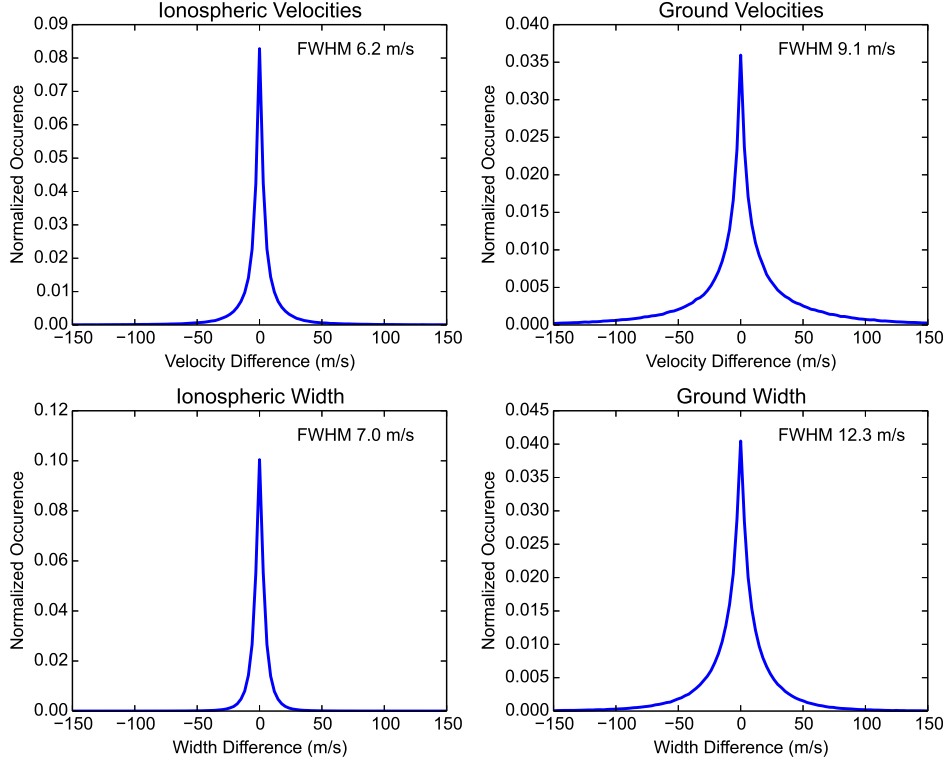


Figure 5.7: Distributions of the difference in fitted velocities and fitted spectral widths between fitted SuperDARN data and VSE subtracted fitted SuperDARN data from the Saskatoon SuperDARN radar from 10 July, 2014 at 10:00 UT until 14 July, 2014 at 10:00 UT. The full-width half-maximum (FWHM) is included in the top right of each panel.

ground scatter ACF measurements and consequently broadening the FWHM of the difference distributions for ground scatter.

5.5.1 Error Estimates

The mean-square error estimates for the VSE, MVSE, and MPSE were evaluated using the modified radar data simulator. The simulator was primed with 2 hours of SuperDARN data from the Saskatoon radar on 14 July, 2014 starting at 6:01 UT to ensure realistic simulations of both complex voltage samples and exact simulated values of the self-clutter. The simulated actual self-clutter was compared to estimates of self-clutter calculated from the simulated

voltage data.

The average relative difference between the VSE and the simulated actual self-clutter was compared with the square root of the relative MSE of the VSE for the two hours of simulated radar data. The average relative difference was found to be 1.02 whereas it was expected to be $1/\sqrt{32} = 0.18$, meaning that on average for realistic SuperDARN data, Equation 5.13 produces an underestimation of the error in the VSE. As was stated earlier this result was expected since Equation 5.13 was derived assuming that the interfering voltages, \tilde{V}_n and \tilde{V}_m , were known, while in practice, these voltage samples were approximated using lag-zero power measurements. However, as was quantified in Table 5.1, when the self-clutter is caused by ground scatter, and therefore the assumptions necessary for accurate usage of the VSE are satisfied, the error estimate is accurate.

The maximal clutter estimates given by Equations 5.14 and 5.15 were also compared with the simulated actual self-clutter. Both the MVSE and the MPSE were found to be maximal within the error estimates given by Equation 5.17 for all simulated data (3,312,000 data points). These results validate both the MVSE and MPSE as robust maximal estimates of the self-clutter power contained in radar ACF estimates.

5.6 Conclusions

The multiple-pulse technique of *Farley* (1972) is widely used by HF ionospheric radar systems like SuperDARN, but, before the present work, it was not possible to quantitatively account for the effects of self-clutter in the errors in autocorrelation function (ACF) estimates. As shown in *Farley* (1969), estimates of self-clutter (transmitted waveform induced range ambiguities) are needed to properly account for errors in estimating the signal ACF. A voltage-based self-clutter estimator (VSE), a maximal voltage-based self-clutter estimator (MVSE), and a maximal power-based self-clutter estimator (MPSE) for use with the multiple-pulse technique have been derived and tested in the present work. The mean-square-error (MSE) for each self-clutter estimator has also been presented and tested.

The VSE given by Equation 5.12 relies on assumptions that are required for approximating interfering voltage samples and may be successfully used to account for self-clutter on a case

by case basis as shown in Figure 5.6. Further study of the VSE is required to develop an estimator that may be more generally used. We suggest development of a more accurate method for approximating interfering voltage samples that would involve using extracted parameters from fitted radar data in an iterative process to produce higher order corrections to the \tilde{V}_n and \tilde{V}_m approximations.

Each maximal estimator has been independently verified utilizing a version of the radar data simulator of *Ribeiro et al.* (2013b) as discussed in Section 5.5. As will be demonstrated in a subsequent study, the maximal self-clutter estimators can be used to self-consistently estimate the measurement error in radar ACF estimates. Radar systems using the multiple-pulse technique will benefit from these self-clutter estimators. For example, for SuperDARN data, the maximal self-clutter estimators can be used to produce error estimates for use in yet to be implemented error-weighted fitting algorithms for improved ionospheric parameter extraction.

Further modifications beyond error-weighted fitting algorithms will need to be made to parameter extraction algorithms for SuperDARN to benefit from the results of this study. The SuperDARN fitacf process currently utilizes the *Ponomarenko and Waters* (2006) cross-range interference detection algorithm meaning that in general significant differences between voltage sample data processed with and without subtraction of the VSE are not expected. For instance, the triangles plotted in Figure 5.6 indicate data points that were flagged by the *Ponomarenko and Waters* (2006) algorithm as “bad data” to be discarded, yet it is clear that some of the lags that have been flagged are in fact not contaminated. Also, despite improvements to the radar ACF made by the VSE the same lags are still flagged in the current implementation.

Modifications to the SuperDARN fitting software and implementation of a new error-weighted data fitting algorithm that utilizes information provided by the maximal self-clutter estimates is currently under development and will be the topic of a future study. Nonetheless, utilizing the MPSE it will be possible to retroactively reprocess *all* SuperDARN data using an error-weighted fitting algorithm, allowing for extracted parameters (Doppler shift, etc.) with signal derived confidence bounds (error bars). Such an algorithm may potentially also enable improved accuracy in the extraction of parameters.

5.7 Acknowledgments

The radar data used in this work was collected with the support of the Canadian Space Agency's Geospace Observatory (GO Canada) continuation initiative. This work was supported by funding provided by Natural Sciences and Engineering Research Council of Canada. The authors acknowledge Gareth Perry and Matthew Wessel for helpful discussions and George Sofko for insight and never-ending encouragement.

CHAPTER 6

SUPERDARN AUTO-CORRELATION FUNCTION STATISTICS

6.1 Background

With estimators for the self-clutter in SuperDARN ACF lag estimates developed in Chapter 5 (in particular Equations 5.12, 5.14, and 5.15), effort was focused on remaining work addressing challenge 1 (Section 5.1, p. 80). More specifically, the current chapter details the research required to develop improvement 2 (Section 5.1, p. 82), which involves determining the statistical properties of SuperDARN ACF lag estimates. From *Farley* (1969), it is clear that in order to accurately estimate the variance of SuperDARN ACF lag estimates, the statistical properties of the SuperDARN voltage samples must be considered and understood. Once the statistical properties of the estimates of the magnitude, phase, real, and imaginary components of ACF lag estimates are known, the variance of these estimates can be used in weighted fitting techniques to extract SuperDARN ionospheric parameters. The current chapter focuses on SuperDARN ACF statistics, while current and new fitting methodologies are discussed in Chapter 7.

6.2 Published Manuscript

Authors: *Ashton S. Reimer, Glenn C. Hussey, and Seth R. Dueck*

Reference: Reimer, A. S., G. C. Hussey, and S. R. Dueck (2016), On the statistics of SuperDARN autocorrelation function estimates, *Radio Sci.*, 51, 690–703, doi:10.1002/2016RS005975.

This chapter contains the second published work towards developing an improved SuperDARN data processing methodology that provides improved fitted parameter errors. This chapter derives and discusses the statistical properties of SuperDARN ACF lag estimates. *Farley* (1969) discussed the statistical properties of *normalized* ACF lag estimates used in incoherent scatter radar signal processing, but SuperDARN signal processing utilizes *unnormalized* ACF lag estimates. Nevertheless, the techniques discussed in *Farley* (1969) are applicable here and were used as a foundation to derive the statistical properties of SuperDARN ACFs.

As will be discussed in detail below, SuperDARN ACF lag estimates are estimated from complex (real and imaginary) voltage samples, which are random variables with statistical properties. This means that the lag estimates are also random variables with statistical properties; however, these properties depend on the statistical properties of the voltage samples. Initially, the PDFs of the lag estimates were determined numerically using Monte Carlo methods, but ultimately the PDFs were analytically derived. This included PDFs for magnitude, phase, real, and imaginary components of the lag estimates. Equipped with the PDFs, the variance of the lag estimates could finally be determined.

The current chapter begins with an analysis and discussion of the statistical properties of SuperDARN voltage samples, including both theoretically and empirically derived PDFs for the real and imaginary components of the voltage samples. Next the methodology used to derive both the PDFs of the ACF lag estimates and the lag estimate variances is presented. Finally, the obtained analytic expressions are validated using both Monte Carlo methods and comparisons to previous work.

It was found that there are no exact analytic expressions for the variance of the magnitude nor the variance of the phase of the ACF lag estimates, however an algorithm for obtaining the variances numerically was developed. In contrast, exact and computationally simple analytic expressions for the variance of the real and variance of the imaginary components of ACF lag estimates were derived and validated. Combining the variance expressions for the real and imaginary components of the ACF lag estimates derived in this chapter with the maximal power-based self-clutter estimator developed in Chapter 5 addresses challenge 1 (Section 5.1, p. 80), enabling an improved data fitting methodology to be developed in

Chapter 7.

What follows is the published ACF statistics manuscript with the same content as was published in *Radio Science*. I am licensed (see Appendix E for details) to reproduce this manuscript within this thesis. The figure numbers, section headings and general formatting have been altered to ensure consistency with the rest of this thesis; however, no other alterations have been made to the manuscript.

6.3 Abstract

Time-domain signal processing techniques are employed by the Super Dual Auroral Radar Network (SuperDARN) to obtain bulk measurements of the velocity and spectral width of F region ionospheric plasma irregularities. The measurements are obtained by fitting estimates of the mean autocorrelation function (ACF) of the radar target. To accurately and consistently extract target parameters from the mean unnormalized ACF, it is necessary to utilize error-weighted fitting algorithms with a weight given by the variance of the ACF. Currently implemented weights are ad hoc and a detailed description of the statistical characterization of SuperDARN ACFs is needed. Following the discussions in *Farley* (1969) and *Woodman and Hagfors* (1969), which describe the variance for the mean normalized ACF used with Incoherent Scatter Radars, we present analytic expressions for obtaining the variance of the real and imaginary components of the mean unnormalized SuperDARN ACF. These expressions are based on models by *André et al.* (1999) and *Moorcroft* (2004) of the voltage signal received by SuperDARN radars, but may be used for other soft target radar systems. An algorithm for obtaining the variance of both the magnitude and phase of the mean ACF is also presented. The results of this study may be directly integrated into existing SuperDARN data analysis software and other pulse-Doppler radar systems that utilize estimates of the mean unnormalized ACF.

6.4 Introduction

The Super Dual Auroral Radar Network (SuperDARN) is an international network of ionospheric pulse-Doppler radars used to measure the bulk $E \times B$ drift velocity in the F region ionosphere (*Greenwald et al.*, 1995; *Chisham et al.*, 2007). The radars operate in the high-frequency (HF) band from 8 to 20 MHz and diagnose decameter-scale field-aligned plasma density irregularities. Pulse-Doppler Radar systems diagnose the range and Doppler characteristics of a target by transmitting a pulse-sequence and recording the signal that is returned. Plasma parameters may be obtained via a frequency-domain analysis of the power spectrum of the signal, or via time-domain auto-covariance techniques if the conditions of the Wiener-

Khinchin Theorem hold (*Chatfield, 2013*). SuperDARN uses a time-domain analysis to obtain an estimate of the auto-correlation function (ACF) of the plasma irregularities. Estimates of the *mean* ACF are obtained by averaging several complex products of pairs of time-lagged voltage samples. Written algebraically, the population mean ACF at lag time t ($R(t)$), may be estimated using K pairs of complex voltage samples (where the voltage samples in each pair are separated, or lagged, in time by t), is given by

$$\hat{R}(t) = \frac{1}{K} \sum_{k=1}^K \tilde{V}_{1k} \tilde{V}_{2k}^*, \quad (6.1)$$

where $\hat{R}(t)$ is the estimator for the sample mean of $R(t)$ and the subscripts denote each one of the two complex voltage samples (V_{1k} , V_{2k}) separated by a time of t (e.g. *Reimer and Hussey, 2015*). Equation 6.1 yields a complex number describing the mean “lag” product of the mean ACF at time t . Equation 6.1 is applied to calculate each of the available lags to obtain the time-dependent behaviour of the mean ACF. The number of unique lags of the mean ACF obtained depends on the pulse-sequence transmitted by the radar (e.g. *Reimer and Hussey, 2015*). The multiple pulse technique used by SuperDARN radars involves transmitting several pulses with non-redundant delays between pulses. This reduces range ambiguities in comparison to uniform pulse spacing techniques, but unwanted echoes from multiple ranges returning to the radar at the same time cause interference in the ACF known as self-clutter (*Farley, 1972; Barthes et al., 1998; Reimer and Hussey, 2015*)

An example of the components of a typical mean ACF measured by a SuperDARN radar are plotted in Figure 6.1. The real and imaginary components of the ACF, as obtained using Equation 6.1, are plotted in the top left and top right panels of the figure, respectively. The magnitude and phase of the mean ACF, plotted in the bottom left and bottom right panels respectively, are obtained using

$$\langle |\hat{R}| \rangle = \sqrt{\left(\mathcal{R} \left[\langle \hat{R}(t) \rangle \right] \right)^2 + \left(\mathcal{I} \left[\langle \hat{R}(t) \rangle \right] \right)^2} \quad (6.2)$$

and

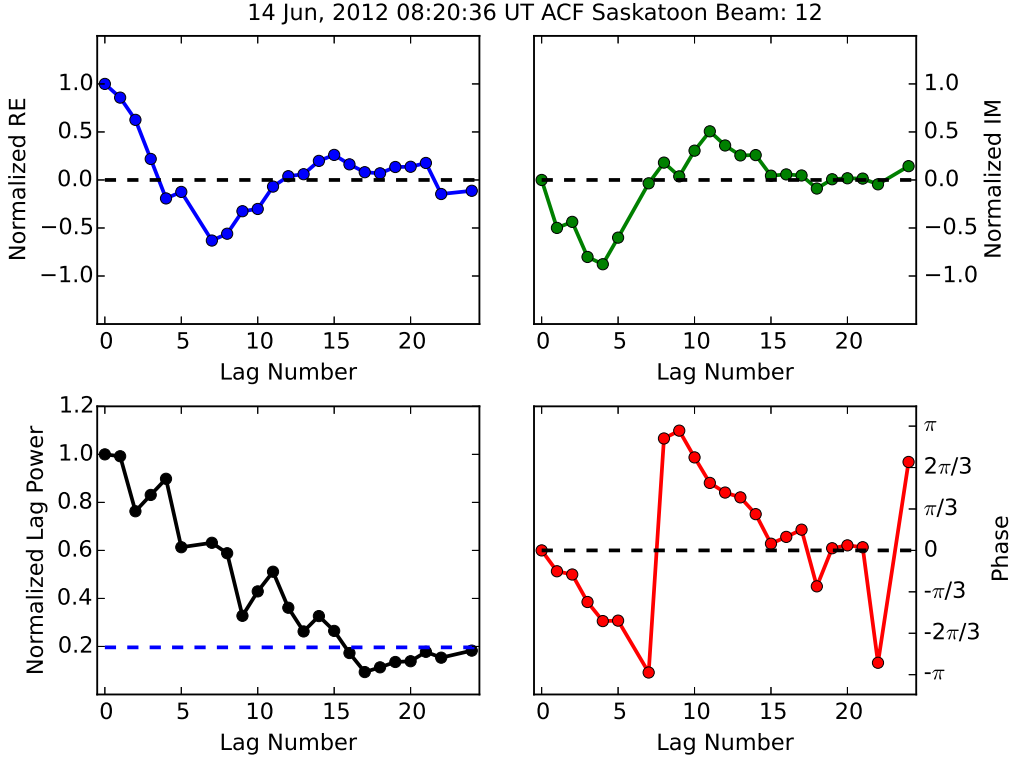


Figure 6.1: Plots of the real (top left), imaginary (top right), magnitude (bottom left), and phase (bottom right) of a mean ACF estimated from measurements by the Saskatoon SuperDARN radar, for beam 0 range gate 20, on 14 July, 2014 at 7:42 UT.

$$\angle \hat{R}(t) = \tan^{-1} \left(\frac{\mathcal{I} [\langle \hat{R}(t) \rangle]}{\mathcal{R} [\langle \hat{R}(t) \rangle]} \right). \quad (6.3)$$

In Equations 6.2 and 6.3, \mathcal{R} and \mathcal{I} denote the real and imaginary components of $\hat{R}(t)$, respectively. The time varying behaviour of the mean ACF for a single Doppler-peaked scatterer can be modeled by a decaying complex exponential, given by

$$R(t) = P e^{-2\pi w_d t / \lambda} e^{j 2\pi f_d t} \quad (6.4)$$

where $P = R(0)$ is the echo power at time $t = 0$ (the so-called lag-zero power), w_d is the spectral width, f_d is the Doppler shift, and λ is the transmitted signal wavelength (e.g. *Ribeiro et al.*, 2013a; *Spaleta et al.*, 2015). Shortly, we will discuss $R(t)$ in terms of its real

and imaginary components with the notation $R(t) = R_r + jR_i$. The mean ACF may also be described in terms of a correlation coefficient, ρ , at time t such that

$$R(t) = R(0)\rho, \quad \rho = \rho_r + j\rho_i. \quad (6.5)$$

Here we note that from Equations 6.4 and 6.5, $\rho_r, \rho_i \in [-1, 1]$ and $|\rho| \in [0, 1]$ for $t > 0$ and that $\rho_r = \rho \cos(\phi)$ and $\rho_i = \rho \sin(\phi)$, where ϕ is the population mean phase.

As discussed by *Villain et al.* (1996), the functional form of the magnitude of Equation 6.4 changes for different plasma conditions. Similarly, it is important to note that Equation 6.4 will change in the presence of mixed scatter (e.g. *Ponomarenko et al.*, 2008). However, the results of this study do not depend on the time varying behaviour of the ACF.

In the present SuperDARN data analysis procedures, the velocity and spectral width of the plasma are obtained by fitting analytic models to the mean ACF using weighted least-squares techniques (e.g. *Ribeiro et al.*, 2013a). Usually, the velocity is obtained by fitting the phase of Equation 6.4 to the phase of the mean ACF and the spectral width is obtained by fitting the magnitude of Equation 6.4 to the magnitude of the mean ACF. Currently, the weighted least-squares fitting algorithms use ad hoc weights. These weights are described in a technical report by Kile Baker (White Paper on FITACF, 2003) and are based on the assumption that the variance of the mean ACF estimates is inversely proportional to the lag power ($|R(t)|$). As will become evident, these ad hoc weights are only qualitatively correct for the phase of the mean ACF and only for lags with large signal-to-noise ratio (SNR). For error-weighted least squares techniques, inaccurate weights lead to inaccurate estimates of the fitted parameter and the fitted parameter error. Quantitatively accurate weights are required to obtain accurate fitted parameter errors (*Press et al.*, 1986).

Previously, *Farley* (1969) studied the statistics and estimation errors of the *normalized* mean ACF used in Incoherent Scatter Radar (ISR) applications. They presented the statistics of the real and imaginary components of the normalized mean ACF for three different normalization factors, including derivations of the analytic expressions for the variance of the estimates both with and without the presence of noise. Small estimation errors require small normalization factor estimation errors, which can only be obtained with a large number of samples (i.e. > 1000) of the normalization factor. For SuperDARN, typically only ≈ 30

samples are obtained, meaning that a normalized ACF estimator cannot be used, therefore SuperDARN utilizes the *unnormalized* ACF estimator. The literature does not appear to discuss the statistics and estimation errors of an *unnormalized* mean ACF.

Here we present and discuss the statistics of the *unnormalized* mean ACF with applications to SuperDARN data analysis, with a focus on the variance of the ACF estimates (i.e. the variance of Equation 6.1). An immediate application of these results to SuperDARN radars, as will be presented in a manuscript under preparation, is replacing the ad hoc weights with the variance of the unnormalized mean ACF estimates will lead to more reliable and self-consistent fitted parameter errors. For example, accurate errors are particularly useful when comparing datasets between different instruments, or when assimilating SuperDARN data into predictive models (e.g. *Ridley et al.*, 2006). Accurate velocity errors will enable accurate error estimates of SuperDARN derived data products (e.g. convection maps, cross polar cap potential, etc.) through error propagation methods.

Section 6.5.1 discusses the statistics of the complex voltage samples measured by SuperDARN and their relationship to the mean ACF. In Section 6.5.2, the analytic form of the distribution of the mean ACF is derived and discussed. The distributions for the magnitude and phase of the mean ACF are also presented. Section 6.6 is dedicated to numerical verification of the mean ACF distributions presented in Section 6.5.2 and comparison of the results to experimentally obtained mean ACFs. An efficient way to numerically compute the variance of these distributions is also discussed. As will be shown in a subsequent publication, the variance of these distributions can be used in error-weighted fitting techniques to obtain improvements to SuperDARN data products, including statistically consistent estimates of plasma parameter errors.

6.5 Mean ACF Estimate Statistics

The statistics of the mean ACF estimator, given by Equation 6.1, follow from the statistical distribution of the complex voltage samples that are received by SuperDARN radars. Using the statistics of the voltage samples allows one to derive analytic expressions for the variance of the estimators for the real and imaginary components of the mean ACF. Applying the

Central Limit Theorem, analytic probability distribution functions for the mean ACF and for the magnitude and phase of the mean ACF may be obtained.

6.5.1 Voltage Sample Statistics

SuperDARN radars receive echoes from a large illuminated volume of the ionosphere, referred to as a range cell, which generally extends 45 km in range with 3 dB beam-width of 3.24 degrees. At the largest transmitted SuperDARN wavelength, 37.5 m (8 MHz), the volume is more than a thousand wavelengths in each dimension. To explain the shape of the power spectra measured by the SHERPA HF radar, *Hanuise et al.* (1993) assumed that received voltage signals are given by the volume integral of the contributions from multiple field aligned irregularities inside a range cell. To explain the behaviour of SuperDARN ACFs observed in the cusp, *André et al.* (1999) modeled the scattered signal as a sum of the scattered signals from many individual irregularities. *Moorcroft* (2004) derived a backscatter model capable of explaining the observed shape of auroral backscatter spectra, for radar frequencies from 10 to 933 MHz, by assuming that a large number (L) of individual scatters, $s_\ell(t)$, each contribute to the total scattered signal, $V(t)$, such that

$$V(t) = \sum_{\ell=1}^L s_\ell(t). \quad (6.6)$$

The total received signal is a sum of the signals provided by each individual scatterer. A radar data simulator based on the work of *Moorcroft* (2004) was shown by *Ribeiro et al.* (2013b) to realistically model experimentally measured SuperDARN ACFs.

Following *Moorcroft* (2004), if the field aligned scatterers are randomly distributed in the volume illuminated by the radar, then by the Central Limit Theorem, the distribution of the sum of the random scattered signals is given by a Gaussian distribution (*Westfall and Henning*, 2013). This result is valid regardless of the form of the probability distribution function that characterizes the individual scatterers. To test the validity of this model, a typical example of the distributions of the real and imaginary components of the voltage signal obtained from the Saskatoon SuperDARN radar is shown in Figure 6.2. The plots were produced using 256 signal dominated voltage samples, from 32 pulse-sequences with 8 pulses each obtained in ≈ 3 seconds, with SNR > 20 dB. These were then binned into 10 bins

(shown by the red dots). The grey curves were generated by running 500 simulations of the voltage data by sampling Equation 6.7 (presented below). As expected, due to the finite number of voltage samples used, there is some deviation of the data from the population mean (black dashed line); however, as indicated by the grey curves, this deviation is within the variance given by the model.

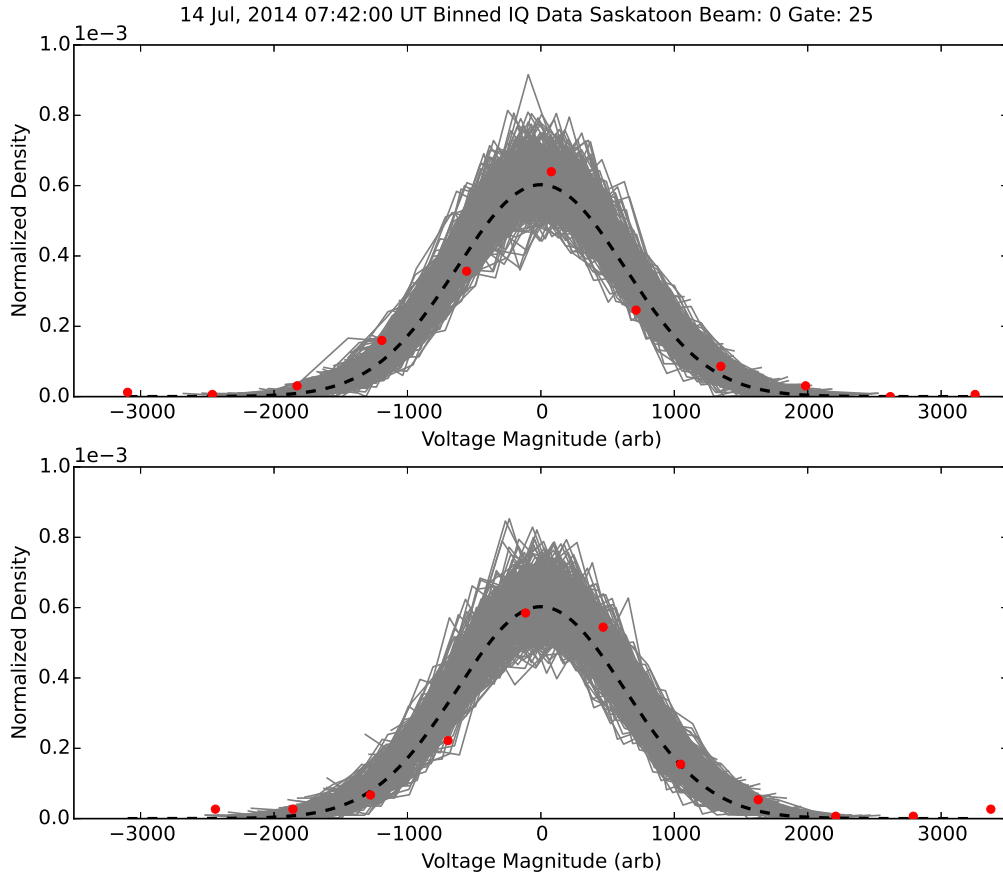


Figure 6.2: The distributions of the real (top) and imaginary (bottom) components of complex voltage samples measured using the Saskatoon SuperDARN radar (red dots), for beam 0 range gate 25, on 14 July, 2014 at 7:42 UT. The grey lines indicate distributions derived from Monte Carlo simulations of the data using Equation 6.7.

As assumed theoretically and just demonstrated experimentally, the voltage samples measured by SuperDARN radars follow a zero-mean Gaussian distribution. Specifically, the real and imaginary components of two voltage samples (i.e. $V_1 = x_1 + ix_2$ and $V_2 = x_3 + ix_4$) can

be described by a 4-dimensional joint Gaussian distribution (e.g. *Simon*, 2007)

$$p(x_1, x_2, x_3, x_4) = \frac{1}{(2\pi)^2 |\mathbf{C}|^{1/2}} \exp \left(-\frac{1}{2|\mathbf{C}|} \sum_{i,j=1}^4 |C_{ij}| x_i x_j \right) \quad (6.7)$$

with a covariance matrix \mathbf{C} . Following the example given by Donald Farley and Tor Hagfors in an unpublished textbook (AGF-304, Arctic Geophysics, The University Courses on Svalbard, Textbook manuscript, 1999), the covariance matrix of Equation 6.7 is written as

$$\mathbf{C} = \sigma^2 \begin{bmatrix} 1 & 0 & \rho_r & -\rho_i \\ 0 & 1 & \rho_i & \rho_r \\ \rho_r & \rho_i & 1 & 0 \\ -\rho_i & \rho_r & 0 & 1 \end{bmatrix}. \quad (6.8)$$

The covariance matrix \mathbf{C} describes the correlation between the real and imaginary components of 2 voltage samples received at 2 different times by a quadrature receiver (such as is used by SuperDARN radars). The real and imaginary components of a single voltage sample are not correlated, but the real and imaginary component between two voltage samples separated in time will be correlated as described by the population mean coefficients of correlation, ρ_r and ρ_i . The variance, σ^2 , is proportional to the echo power or lag-zero power ($R(0) = 2\sigma^2$) received by the radar. As detailed in Section 6.6, Equation 6.8 may be used in a Monte Carlo simulation to numerically verify the analytically derived distribution and variance of the mean ACF and to test for when the assumption that the Central Limit Theorem can be applied breaks down.

6.5.2 Mean ACF Component Distributions

To derive the probability distribution function for the mean ACF, an approach similar to that taken by *Woodman and Hagfors* (1969) was used. We have not considered contributions of noise as it is a simple exercise to add this later (e.g. *Farley*, 1969). The noise in SuperDARN voltage samples is Gaussian (see Figure 8 in *Goh et al.* (2014)), so the noise contribution to the voltage samples is accounted for by adding the variance of the noise, σ_N^2 , to the main diagonal of \mathbf{C} . Equation 6.1 shows that estimates of the mean ACF are given by the sum of random variables (the voltage samples, V_1 and V_2 , are random variables and therefore their

products are random variables). The sample mean estimators for the real and imaginary components of the ACF are given by

$$\hat{R}_r = \frac{1}{K} \sum_{k=1}^K (x_{1k}x_{3k} + x_{2k}x_{4k}), \quad \hat{R}_i = \frac{1}{K} \sum_{k=1}^K (x_{2k}x_{3k} - x_{1k}x_{4k}) \quad (6.9)$$

with expected values $R_r = P\rho_r$ and $R_i = P\rho_i$, where $P = R(0)$. The variance for each of these estimators is the difference of the expectation value of the square of the estimator, $\langle \hat{R}^2 \rangle$, and the square of the expectation value, $\langle \hat{R} \rangle^2$, of the estimator (e.g. *Farley*, 1969; *Reimer and Hussey*, 2015). For voltage samples characterized by Equation 6.7, the variance of the real (σ_r) and imaginary (σ_i) components of the population mean ACF, derived from Equation 6.9, are given by

$$\sigma_r^2 = P^2 \left(\frac{1 - |\rho|^2}{2K} + \frac{\rho_r^2}{K} \right) \quad (6.10)$$

and

$$\sigma_i^2 = P^2 \left(\frac{1 - |\rho|^2}{2K} + \frac{\rho_i^2}{K} \right) \quad (6.11)$$

where again P is the signal power at lag-zero. Note that this result is independent of the functional form of the ACF and therefore does not depend on Equation 6.4 nor does it depend on the probability distribution function of the mean ACF, but rather on the probability distribution function of the voltage samples. For voltage samples of a different distribution, Equations 6.10 and 6.11 need to be re-derived following the discussion presented here and in *Farley* (1969). To include contributions of Gaussian noise, N , and clutter (including self-clutter), C :

$$P \rightarrow P + N + C, \quad \rho \rightarrow \frac{R(t) + R_N(t) + R_C(t)}{P + N + C}, \quad (6.12)$$

where R_N and R_C are the noise and clutter correlation functions (*Farley*, 1969). In current practice for SuperDARN, $R_N(0) = N$ with $R_N = 0$ elsewhere and the clutter term may be estimated using one of the estimators derived in *Reimer and Hussey* (2015).

The probability distribution function of the real and imaginary components can be described by a 2-dimensional joint Gaussian distribution. Here the Central Limit Theorem is invoked by assuming “large” values of K . How large K must be depends on how “close” one wishes the distribution to be to Gaussian, where only $K \rightarrow \infty$ produces a distribution

identical to a Gaussian (*Westfall and Henning, 2013*). *Westfall and Henning (2013)* suggests a “dirty rule of thumb” of $K = 30$ is sufficient and as will be shown in Section 6.6, $K = 30$ (the typical number of sample obtained in 3 seconds of integration on SuperDARN radars) produces acceptable results. Assuming large K , the real and imaginary components of the mean ACF are given by (e.g. *Simon, 2007*)

$$p(x, y) = \frac{1}{2\pi\sigma_r\sigma_i\sqrt{1-\rho_{ri}^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma_r^2(1-\rho_{ri}^2)} - \frac{(y-\bar{y})^2}{2\sigma_i^2(1-\rho_{ri}^2)} + \frac{\rho_{ri}(x-\bar{x})(y-\bar{y})}{\sigma_r\sigma_i(1-\rho_{ri}^2)}}, \quad \rho_{ri} = 2\rho_r\rho_i \quad (6.13)$$

where x and y correspond to samples of the real and imaginary components of the mean ACF. Here the bar notation is used to denote expected values (population mean), with $\bar{x} = R_r$ and $\bar{y} = R_i$.

It is a straightforward matter to derive the distribution functions of both the magnitude and phase of the mean ACF from Equation 6.13 (e.g. *Beckmann, 1962*; *Aalo et al., 2007*; *Dharmawansa et al., 2009*). Equation 6.13 is first converted to polar coordinates and then integrated with respect to either r or θ . The distribution of the phase of the mean ACF, given by Equation C.3 in the Appendix, is found by integrating Equation C.1 (Appendix) with respect to r from 0 to ∞ , yielding a result similar to *Aalo et al. (2007)* but differing from *Woodman and Hagfors (1969)*. The distribution of the magnitude of the mean ACF (see Equation C.13 in the Appendix) is found by integrating Equation C.1 with respect to θ from 0 to 2π . This is straightforward in principle, but as detailed in the Appendix the calculation is rather involved. It is difficult if not impossible to calculate an analytic expression for the variance of both Equations C.3 and C.13 (Appendix), thus we defer this to numeric methods as discussed in Section 6.6. Equations C.3 and C.13 are also computationally slow and therefore it is impractical to use them to determine the variance of the phase and magnitude distributions for error-weighted fitting in place of other methods, but they were included for the sake of completeness. Here again, we will defer to numeric methods in Section 6.6 where we present a more computationally efficient numeric method for determining the variance of magnitude and phase of the mean ACF estimator given in Equation 6.1.

6.6 Analysis and Discussion

As a test of the analytic expressions derived in Section 6.5.2, they were compared with distributions that were numerically derived from Equation 6.7 with a Monte Carlo simulation. While similar analytic results were obtained by *Beckmann* (1962) and *Aalo et al.* (2007), their expressions were not verified numerically. Using the Monte Carlo method, we were able to determine the value of K at which the validity of invoking the Central Limit Theorem to obtain Equation 6.13 from Equation 6.7 breaks down. Experimentally obtained SuperDARN data was also compared with Equation 6.13.

6.6.1 Comparison with Monte Carlo

To test the Central Limit Theorem assumption, K voltage samples were drawn from Equation 6.7 and using Equation 6.1 a sample of the mean ACF was numerically obtained; this process was repeated several times until M samples were obtained and a distribution of the mean ACF was numerically produced. Values for σ , ρ_r , and ρ_i in Equation 6.8 were determined using Equation 6.4 with realistic values for echo power, P , spectral width, w_d , and Doppler shift, f_d , derived from real SuperDARN data. Various permutations of velocities, spectral widths, and number of samples, K , were explored.

Figure 6.3 compares the numerically and analytically derived mean ACFs for the real and imaginary (equation 6.13), magnitude (equation C.13), and phase (equation C.3) distributions. The top panels in the figure show the real and imaginary marginal distributions and the bottom panels compare the numeric magnitude and phase distributions with the analytic results given by Equations C.3 and C.13. The relative difference between the Monte Carlo and analytic distributions was studied and found to be less than the expected statistical sampling error (i.e. $< 1/\sqrt{M}$ with M as the number of samples of the Monte Carlo derived mean ACF distribution).

The good agreement observed in the Figure 6.3 is an expected result, but it confirms that the expressions derived here and by previous authors (e.g. *Aalo et al.*, 2007) are correct. This agreement was observed for all lags, regardless of the values of velocity and spectral width selected. Specifically testing various values of K (not shown), we observed that the

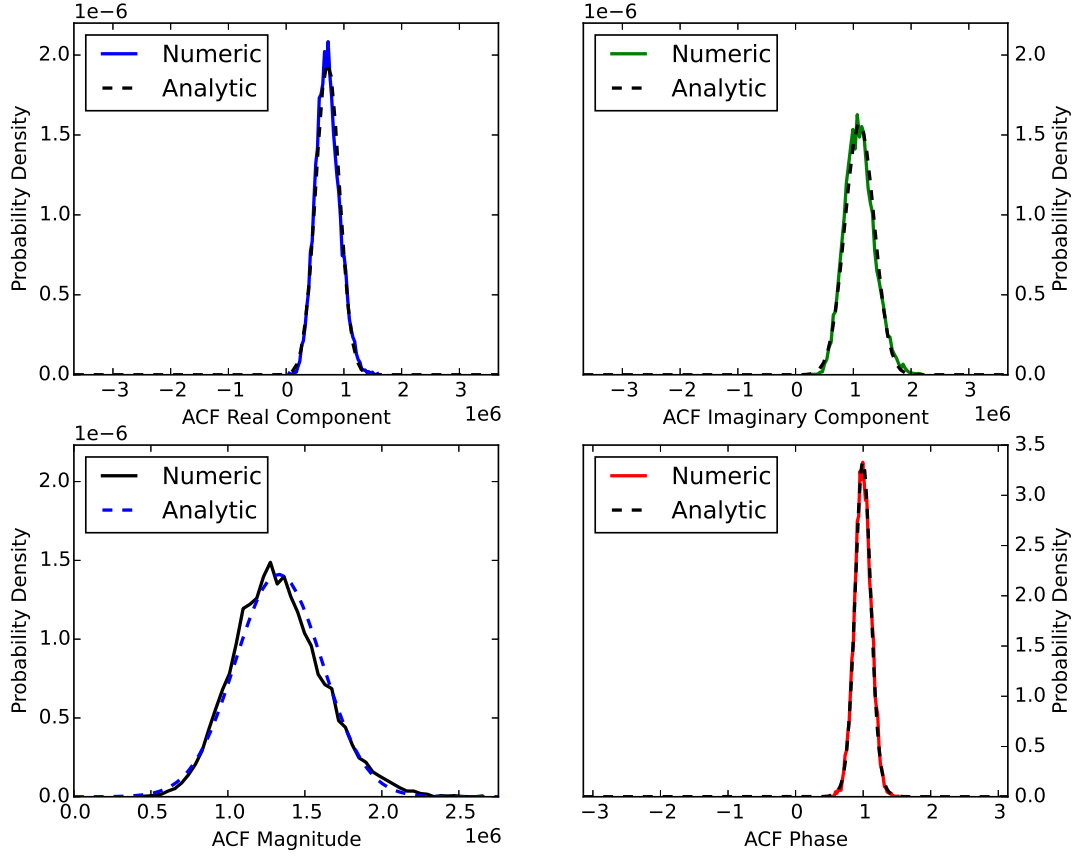


Figure 6.3: A comparison of the Monte Carlo generated marginal distributions for lag 5 of the mean ACF with the respective analytic expressions for the distributions of the components of the mean ACF given by Equations 6.13, C.3, and C.13. For a velocity of 300 m/s and a spectral width of 200 m/s.

assumption that the Central Limit Theorem can be applied to derive Equation 6.13 is invalid when $K < 10$, indicating that Equation 6.13 may be used so long as $K \geq 10$ (note that SuperDARN radars typically operate with $K > 20$). So for $K < 10$, Equation 6.7 must be used to numerically obtain the probability distribution function of the mean ACF. As will be discussed, using Equation 6.7 is more computationally intensive by a factor of K .

The analytic expressions for the variance of the mean ACF (Equations 6.10 and 6.11) were also directly compared with Monte Carlo-derived results. This was done to both visualize the variance for typical SuperDARN mean ACFs and to demonstrate that Equations 6.10 and 6.11 have been properly derived. Figure 6.4 shows the variances of the real (left column) and imaginary (right column) components of the mean ACF. A velocity of 300 m/s was used for all plots and spectral widths of 10 m/s, 200 m/s, and 1000 m/s, were used for the top, middle, and bottom rows of plots respectively. These values were chosen to explore the typical range of spectral widths observed by SuperDARN radars.

In contrast to the real and imaginary components of the ACF, analytic expressions for the variance of the magnitude and phase distributions are not easily, if at all, calculable, and therefore must be obtained numerically. One could use the distribution for the voltages (equation 6.7) in combination with a Monte Carlo method to obtain the distributions of the magnitude and phase of the mean ACF and then numerically determine the variance for each, but this process is slower (it requires $K \times M$ samples of Equation 6.7). Instead, using Equation 6.13 in a Monte Carlo simulation enables the variance of the magnitude and phase to be obtained with K fewer computations. For some SuperDARN experiments with $K < 10$, such as *Greenwald et al.* (2008) where $K = 2$ and *Theurer and Bristow* (2012) where $K = 9$, Equation 6.13 cannot be used, and the computationally slower method of using Equation 6.7 must be performed. To obtain the variance of the magnitude and phase distributions, one must obtain M samples of x and y from Equation 6.13, calculate vectors of the magnitudes r and phases θ from the x and y samples, and then calculate the variance of the magnitudes and phases. This method is also faster at obtaining the variance of the magnitude and phase than a numeric algorithm that obtains samples of the magnitude and phase using Equation C.3 or C.13. For directional data (such as the phase of the mean ACF), it is important to note that the variance must be calculated using the square of the mean

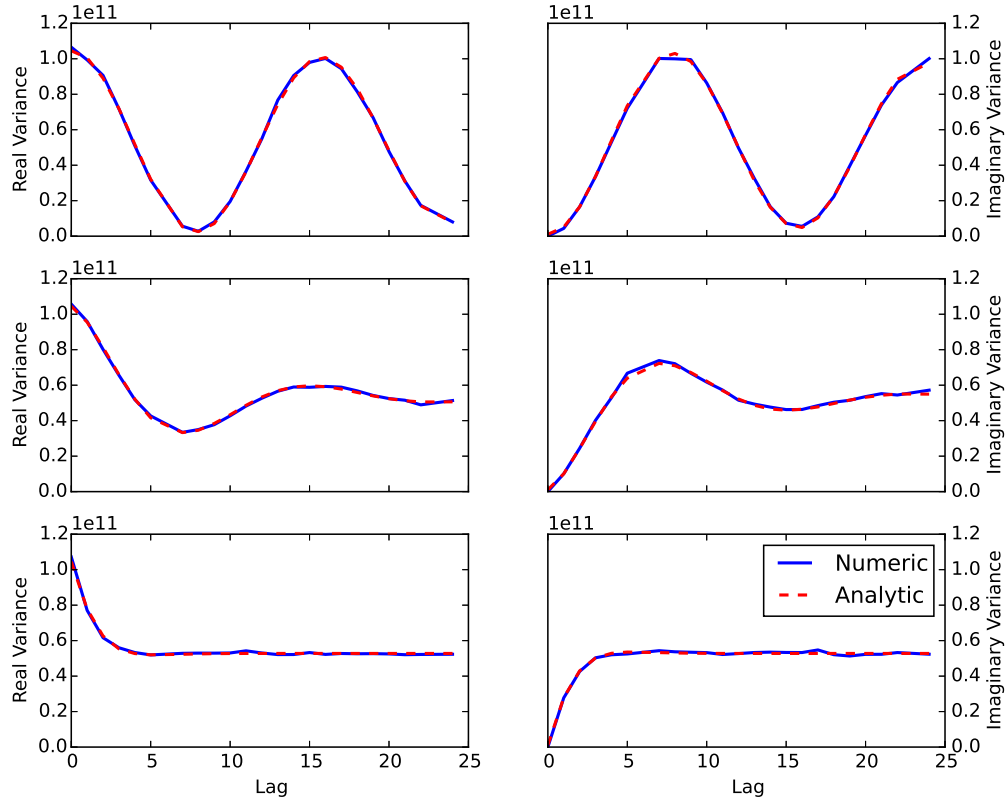


Figure 6.4: A comparison of the variance of the analytic (red dashed curves) and Monte Carlo simulated (solid blue curves) real (left column) and imaginary (right column) component distributions. A power of 1830000 (arb), a velocity of 300 m/s, and spectral widths of 10 m/s (top), 200 m/s (middle), and 1000 m/s (bottom) were used.

deviation, d_0 , given by *Mardia* (1972) as

$$d_0 = \pi - \frac{1}{N} \sum |\pi - |\theta_s - \bar{\theta}|| \quad (6.14)$$

where N is the number of samples of the phase distribution (θ_s) and $\bar{\theta}$ is the sample circular mean.

Figure 6.5 compares the Monte Carlo derived and analytic variances for the magnitude (left column) and phase (right column) of a mean ACF. The figure was generated using the same velocity and spectral widths as in Figure 6.4. Note that the phase variance in the top row is essentially 0 being barely visible along the x-axis.

The behaviour of the real, imaginary, magnitude, and phase variance of the mean ACF

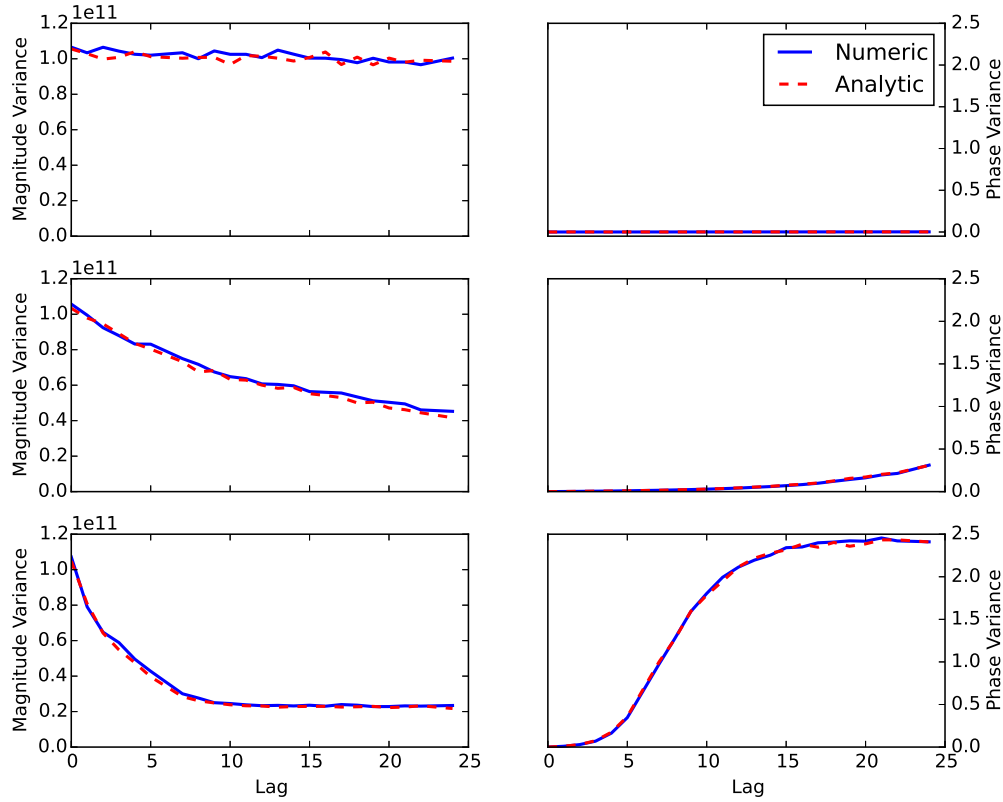


Figure 6.5: A comparison of the variance of the analytic (red dashed curves) and Monte Carlo simulated (solid blue curves) magnitude (left column) and phase (right column) component marginal distributions. A power of 1830000 (arb), a velocity of 300 m/s, and spectral widths of 10 m/s (top), 200 m/s (middle), and 1000 m/s (bottom) were used.

components differs for various values of input parameters ρ and K . Figures 6.4 and 6.5 were produced with spectral width increasing from the top row down. As the spectral width increases, the correlation coefficient, ρ , decreases more quickly as a function of increasing lag number (not shown). In general, as described by Equation 6.4, as the lag number increases (t increases) the correlation coefficient will decrease.

Exploring the asymptotic behaviour of the variances, we observe different behaviours for the different components. Note that to generate the figures, $P = 1830000$ and $K = 32$ were used. For narrow spectral width (top row Figure 6.4), the variance of the real component (c.f.

Equation 6.10) is P^2/K at $t = 0$ but then oscillates between P^2/K and $P^2/(2K)$ for larger lags. Similarly, the variance of the imaginary component (c.f. Equation 6.11) is 0 at $t = 0$, but then oscillates between P^2/K and $P^2/(2K)$ for larger lags. This behaviour is observed for both the Monte Carlo (numeric) and analytic variances of the real and imaginary components. As the spectral width increases, the variance of the real and imaginary components both converge to a value of $P^2/(2K)$ for increasing lag number.

Observing the asymptotic behaviour seen in Figure 6.5, the variance of the magnitude distribution at lag-zero (variance of $|\hat{R}(t)|$ for $t = 0$) is P^2/K for small lags but approaches $\frac{4-\pi}{2} \frac{P^2}{2K}$ for large lags when the spectral width is large (bottom row). This behaviour at large lags results from the distribution of the magnitude of the mean ACF, given by Equation C.13, becoming a Rayleigh distribution as $\rho \rightarrow 0$ (e.g. *Beckmann*, 1962). The variance of the phase distribution is 0 at lag-zero, but increases as ρ decreases (as the lag number increases). For $\rho \rightarrow 0$ the variance approaches $(\pi/2)^2 \approx 2.47$ as observed in the bottom right panel of Figure 6.5, as the phase distribution approaches a circular uniform distribution.

Here it is important to note that the ad hoc weights for fitting the magnitude and phase as reported by Kile Baker (White Paper on FITACF, 2003) are inconsistent with the behaviour of the variance seen in Figure 6.5. Whereas the ad hoc weights discussed by Baker produce an increasing variance as lag power decreases ($Z/|R(t)|^2$ where Z is a constant), the variance of the magnitude and phase shown in Figure 6.5 clearly do not follow a simple inverse-squared relation.

6.6.2 Comparison with Experimental ACFs

The probability distribution function of the mean ACF, given by Equation 6.13, was compared with experimentally obtained SuperDARN ACFs from the Saskatoon SuperDARN radar. For each comparison, it was determined that the model was able to reproduce the corresponding experimentally obtained mean ACF. To perform the comparison, first the power, velocity, and spectral width of the experimental ACF was determined using a least-squares fit to Equation 6.4 and then the fitted parameters were substituted into Equations 6.4, 6.10, 6.11, and 6.13. Figure 6.6 compares 100 Monte Carlo simulated ACFs (grey lines) obtained using Equation 6.13 with an experimentally measured ACF (red dots). The average of the

simulated mean ACFs is indicated by the black line, which corresponds to Equation 6.4 with $K = 26$, $P = 1141899$, $v_d = 307$ m/s and $w_d = 188$ m/s. It is clear from the figure that the model given by Equation 6.13 is able to accurately reproduce the experimental data since all of the data is within the variability observed in the simulated mean ACFs.

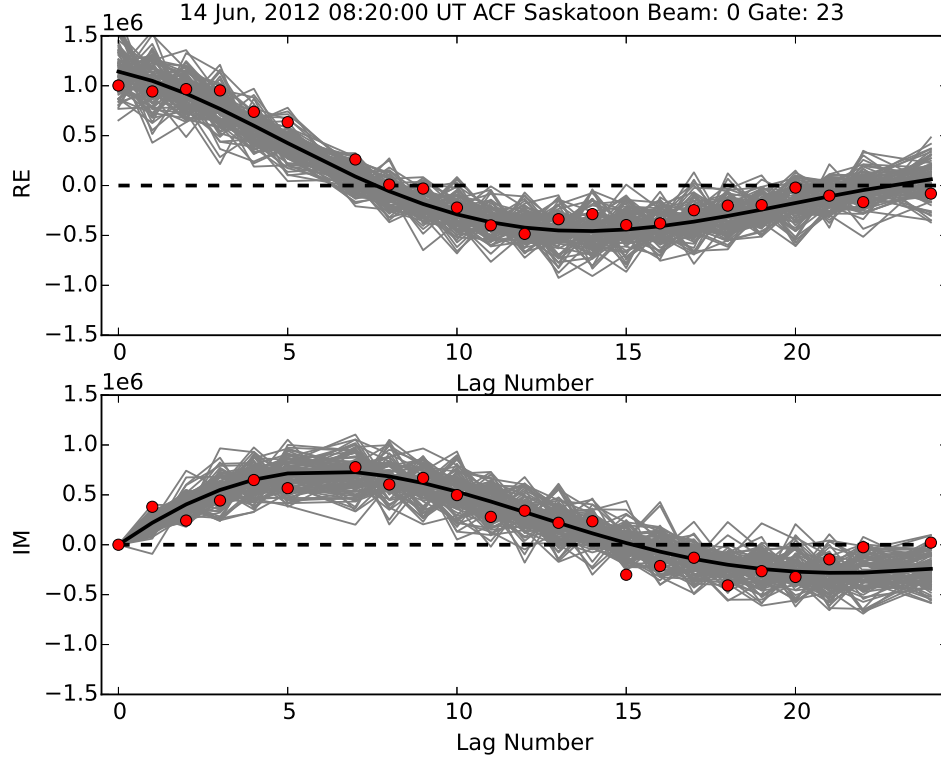


Figure 6.6: A comparison of the ACF model given by Equation 6.13 with an experimentally measured SuperDARN mean ACF. See text for details.

6.6.3 Comparison to Previous Work

Woodman and Hagfors (1969) derived an expression for the distribution of the phase of the mean ACF and an approximate analytic form for the variance of the phase, given by

$$\langle (\delta\hat{\phi})^2 \rangle = \frac{1}{2K} \frac{1 - |\rho|^2}{|\rho|^2}, \quad (6.15)$$

valid only where

$$\frac{K\rho^2}{1 + \rho^2} > 1. \quad (6.16)$$

The distribution and variance expressions were derived for use with ISRs, which use the normalized mean ACF, but the results are also applicable to SuperDARN radars that use the unnormalized mean ACF. This is because any normalization constant applied to the mean ACF is cancelled while calculating the phase using Equation 6.3.

Figures 6.7 and 6.8 compare the results derived by *Woodman and Hagfors* (1969) with those presented in the current study. Figure 6.7 shows a comparison between the phase distribution of *Woodman and Hagfors* (1969) (solid blue curve) with Equation C.3 (red dashed curve); the plot was made using $\rho = 0.5$, $K = 30$, and $\phi = \pi/4$. Notice that the distribution of *Woodman and Hagfors* (1969) is bimodal and encloses a total area greater than 1 compared to the curve produced by Equation C.3, which is unimodal and encloses a total area equal to 1. Comparing the largest peaks, we see that both curves follow an approximately Gaussian shape and have equivalent full-width-at-half-maximums. Here we remark that the phase distribution given by Equation C.3 is only unimodal due to the major axis of the mean ACF distribution being restricted to lie on the radial axis (in polar coordinates) by the assumed mean ACF model given by Equation 6.4. For other assumed mean ACF models, such as the mixed scatter model given by Equation 1 in *Barthes et al.* (1998), Equation C.3 may become bimodal; however, it is important to note this bimodality does not explain the bimodality observed in *Woodman and Hagfors* (1969).

Figure 6.8 compares the approximate expression for the phase variance of *Woodman and Hagfors* (1969) (blue solid curve) with that derived from Equation 6.13 (red dashed curve). The curves were made using $K = 10$, and $\phi = \pi/4$. Good agreement between the two curves is observed in the region to the right of the black dashed line ($\rho > 0.33$), where the condition given by Equation 6.16 is satisfied. To the left side of the black dashed line Equation 6.15 diverges to infinity, whereas the red dashed curve converges to $(\pi/2)^2$ as previously discussed (c.f. bottom right panel of Figure 6.5).

It is clear that despite the bimodal nature of phase distribution derived by *Woodman and Hagfors* (1969), Equation 6.15 is valid so long as the condition in Equation 6.16 is satisfied. When Equation 6.16 is satisfied, one may utilize Equation 6.15 to calculate the variance in the phase of the mean ACF, but when it is not satisfied, the method discussed earlier to numerically obtain the variance in the phase from Equation 6.13 should be used, which is

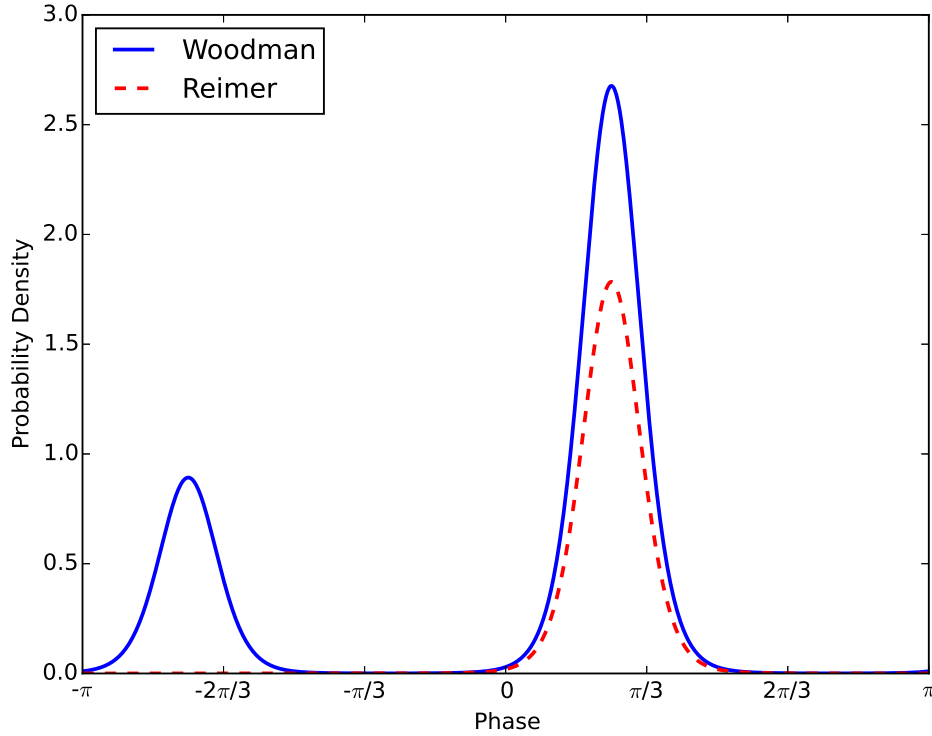


Figure 6.7: A plot of the probability distribution function for the phase of the mean ACF given by *Woodman and Hagfors* (1969) (blue solid curve) and Equation C.3 (red dashed curve). The distributions were made using $\rho = 0.5$, $K = 30$, and $\phi = \pi/4$.

computationally slower than using Equation 6.15.

6.7 Conclusions

The present study has derived and numerically verified analytic expressions for the mean ACF based on the statistics of the voltage samples measured by SuperDARN radars. The study was motivated by a need to know the variance of the mean ACF estimator (equation 6.1) and the variance of its magnitude and phase (Equations 6.2 and 6.3) for use in the error-weighted fitting data analysis techniques used by SuperDARN. The results presented here are broadly applicable to any pulse-Doppler radar systems that utilize components of the mean *unnormalized* ACF to diagnose radar targets. First, it was shown that pairs of

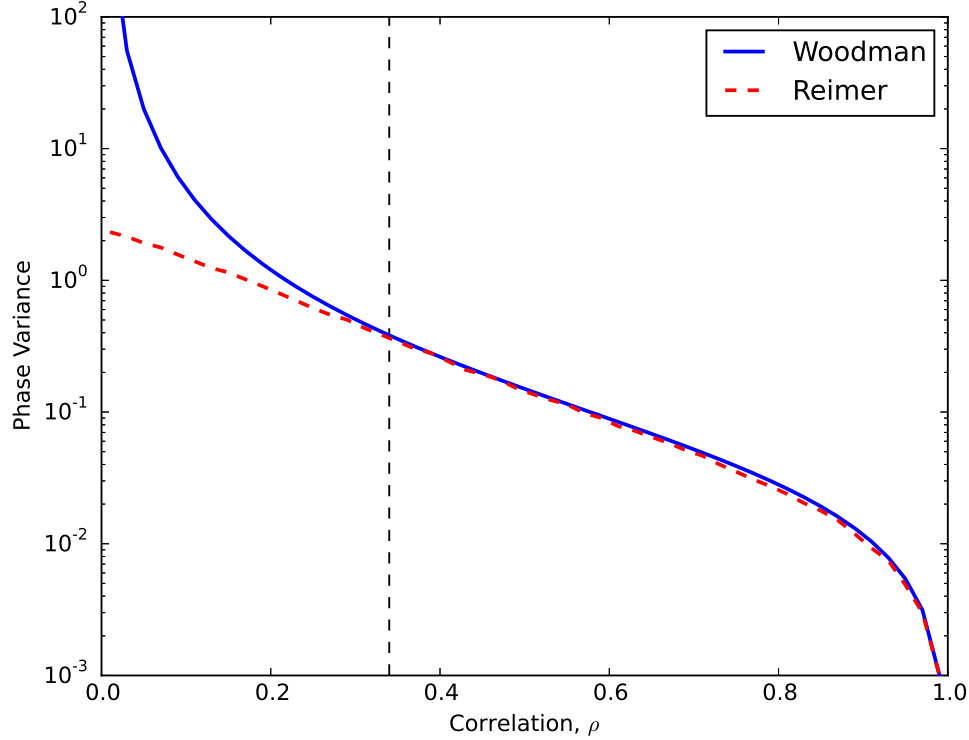


Figure 6.8: A comparison of the variance of the phase distribution given by Equation 6.15 (blue solid curve) and that derived from 6.13 (red dashed curve). The plots were made using $K = 10$, and $\phi = \pi/4$.

voltage samples obtained by the SuperDARN radars can be characterized by a 4-dimensional Gaussian distribution by modeling the received voltage signal from field-aligned ionospheric plasma irregularities as the sum of a large number of individual scattered signals (*André et al.*, 1999; *Moorcroft*, 2004; *Ribeiro et al.*, 2013b). This result was confirmed with experimentally obtained voltage data (for example Figure 6.2). Next, using the Central Limit Theorem it was shown that the distribution of the mean ACF may be accurately characterized by the 2-dimensional Gaussian distribution described by Equation 6.13, with variances given exactly by Equations 6.10 and 6.11. Analytic expressions for the distributions of the phase and magnitude of the mean ACF, Equations C.3 and C.13 (see Appendix), were derived, numerically verified, and studied with a Monte Carlo method. It was determined that both the Central Limit Theorem assumption and Equation 6.13 are valid so long as $K \geq 10$.

Equation 6.13 was also compared with experimentally obtained SuperDARN mean ACFs and found to accurately reproduce measured ACFs (see Figure 6.6).

The variance of the distributions given by Equations 6.13, C.3, and C.13 were also studied and numerically verified. The variance of the real component of the mean ACF was found to be bounded by $\sigma_r^2 \in [P^2/(2K), P^2/K]$, decreasing as the correlation approaches 0 ($\rho \rightarrow 0$). The variance of the imaginary component of the mean ACF was found to be bounded by $\sigma_i^2 \in [0, P^2/K]$, starting at 0 for complete correlation ($\rho = 1$) and then approaching $P^2/(2K)$ as the correlation approaches 0. The variance of the magnitude distribution was found to be bounded by $\sigma_m^2 \in [P^2(4 - \pi)/(4K), P^2/K]$, decreasing as the correlation approaches 0. The variance of the phase distribution was found to be bounded by $\sigma_\theta^2 \in [0, (\pi/2)^2]$, approaching 0 as the correlation approaches 1 and approaching $(\pi/2)^2$ as the correlation approaches 0. It is important to note that the real and imaginary variances were derived directly from Equation 6.9 and are exact results that only depend on the statistics of the voltage samples and not do not depend on the assumption that usage of the Central Limit Theorem to obtain Equation 6.13 is valid.

The variance results presented in this study can be used in error-weighted fitting algorithms that extract plasma parameters from the mean ACF. For the real and imaginary components of the mean ACF, analytic expressions for variance (Equations 6.10 and 6.11) were derived. An computationally efficient algorithm for numerically calculating the variance of the magnitude and phase from Equation 6.13 was discussed, although Equation 6.15 (from earlier work by *Woodman and Hagfors* (1969)) for the variance of the phase is more efficient and can be used as long as the condition in Equation 6.16 is satisfied. The results of this study may be incorporated into existing radar signal processing routines (as discussed in *Ribeiro et al.* (2013a)) currently utilized by the SuperDARN collaboration or new fitting techniques, such as the Bayesian Inference method recently described by *Spaleta et al.* (2015). Other pulse-Doppler radar systems that extract target characteristics from estimates of the unnormalized mean ACF of soft targets would also benefit from this investigation. Future work on approximate analytic expressions for the variance of the magnitude and phase based on spline fitting would also be beneficial.

6.8 Acknowledgments

This study was supported by funding from the Canada Foundation for Innovation (CFI) Major Science Initiatives Program, the Canadian Space Agency (CSA) Geospace Observatory (GO Canada) contribution agreement, and Natural Sciences and Engineering Research Council (NSERC) grants CGSD3-425247-2012 and 414292-00014756. The authors acknowledge the use of SuperDARN data. SuperDARN is a collection of radars funded by national scientific funding agencies of Australia, Canada, China, France, Japan, South Africa, United Kingdom, and United States of America. SuperDARN data used in this study is available from the corresponding author upon request. AR acknowledges Ella Coulter for advice during the writing process.

CHAPTER 7

SELF-CONSISTENT FITTED PARAMETER ERRORS FOR SUPERDARN

7.1 Background

As was discussed in Section 5.1, during the investigation into current SuperDARN signal processing methods, several improvements were identified to address numerous challenges. Improvements 1 and 2 (Section 5.1, p. 82) were developed and discussed in Chapters 5 and 6. In Chapter 5 a novel maximal power-based self-clutter estimator (Equation 5.15) was developed that can be used to provide an upper-limit estimate of the self-clutter that contributes to each lag estimate of the SuperDARN ACF. The estimator accomplishes the objective of improvement 1 (Section 5.1, p. 82) since it can be used in all past, present, and future SuperDARN data processing. In Chapter 6, the statistical properties of the SuperDARN ACF lag estimates were determined, validated, and discussed. This research accomplished the objective of improvement 2 by deriving techniques and expressions to determine the variance of the magnitude, phase, real and imaginary components of the lag estimates. Using both the power-based self-clutter estimator and the ACF lag variance expressions, specifically the lag variance of the real and imaginary components (Equations 6.10 and 6.11), challenge 1 has been addressed, with challenges 2 and 3 remaining (Section 5.1, p. 80).

The current chapter presents an improved SuperDARN signal processing methodology that addresses challenges 2 and 3 using a weighted non-linear least squares fit to the real and imaginary components of the ACF lag estimates, where the weights are determined using both the *maximal power-based self-clutter estimator* developed in Chapter 5 and the exact real and imaginary ACF lag variance expressions developed in Chapter 6. This is in

contrast to the current SuperDARN fitting methodology, which performs linear least-squares fits to the magnitude and phase of the ACF and uses ad hoc expressions for the variances. Performing fits to the real and imaginary components avoids the problems associated with unwrapping the phase.

7.2 Published Manuscript

Authors: *Ashton S. Reimer, Glenn C. Hussey, and Kathryn A. McWilliams*

Reference: Reimer, A. S., G. C. Hussey, and K. A. McWilliams (2018), Statistically self-consistent and accurate errors for SuperDARN data, *Radio Sci.*, 53, 93–111.

doi:10.1002/2017RS006450

This chapter contains the final research towards developing an improved SuperDARN data processing methodology, with a focus on providing improved and accurate fitted parameter errors. The chapter presents and discusses a first-principles fitting methodology (FPFM) that avoids the challenges associated with fitting the magnitude and phase of SuperDARN ACF lag estimates by instead performing weighted non-linear least-squares regression on the real and imaginary components of the ACF lag estimates. The exact analytic variance expressions for the real and imaginary components of the ACF, which were developed in Chapter 6, are used to weight the least-square fitting process. The variance includes estimates of the self-clutter contained in each ACF lag using the maximal power-based self-clutter estimator developed in Chapter 5.

Initially, research in to developing a new fitting methodology focused on a “brute-force” technique, where the minimum χ^2 value was determined by exploring the range of valid physical values for the fitted parameters. The advantage of this method was that it enabled a thorough exploration of the χ^2 space, and accurate confidence intervals, to be easily developed. The disadvantage was that the method was computationally very slow for non-linear models involving transcendental functions such as sine, cosine, and exponential. Eventually, the Levenburg-Marquardt algorithm was chosen to perform non-linear regression. Fitting the magnitude and phase of the ACF lag estimates was also attempted. While attempting

to fit the magnitude, it was found that the exponential model used in current SuperDARN fitting was fundamentally incompatible with the statistical nature of the magnitude of the ACF lag estimates. This is thoroughly discussed below. While attempting to fit the phase, an error-weighted phase unwrapping algorithm was attempted; however, this was abandoned after it was realized that a phase unwrapping algorithm that can unambiguously unwrap the phase is fundamentally incompatible with the statistical properties of the phase of the ACF lag estimates. This research was performed in conjunction with the research presented in Chapter 6, which informed the development of the FPFM.

Significant testing of the FPFM was performed using synthetic SuperDARN data generated with the same simulator used to validate the self-clutter estimators in Chapter 5. These simulation tests proved that the confidence intervals (errors bars) of the fitted parameters varied as expected with signal, noise, and self-clutter, while also verifying that the confidence intervals were consistent with the true fitting error. Finally, comparisons between the FPFM and the current SuperDARN signal processing methodology were performed with real SuperDARN observations in a statistical study spanning eight years of data collected by the Saskatoon SuperDARN radar. All of the testing and comparisons determined that the FPFM is an improvement compared to the current SuperDARN signal processing methodology, especially in regards to the fitted parameter errors. Significantly more data is made available, all with accurate errors, and all of the SuperDARN data that has ever been collected may be reprocessed for improved research data analysis and insight.

What follows is the fitting methodology manuscript with the same content as was published in *Radio Science*. I am licensed (see Appendix E for details) to reproduce this manuscript within this thesis. The figure numbers, section headings and general formatting have been altered to ensure consistency with the rest of this thesis; however, no other alterations have been made to the manuscript.

7.3 Abstract

The Super Dual Auroral Radar Network (SuperDARN) fitted data products (e.g. spectral width and velocity) are produced using weighted least-squares fitting. We present a new First-Principles Fitting Methodology (FPFM) that utilizes the first-principles approach of *Reimer et al.* (2016) to estimate the variance of the real and imaginary components of the mean autocorrelation functions (ACF) lags. SuperDARN ACFs fitted by the FPFM do not use *ad hoc* variance or data filtering criteria. Currently, the weighting used to fit the ACF lags are derived from *ad hoc* estimates of the ACF lag variance. Additionally, a lag filtering criterion is used that is over-cautious and sometimes discards data that contains useful information. In low signal-to-noise (SNR) and/or low signal-to-clutter regimes these *ad hoc* criteria lead to underestimated errors for the fitted parameter because the relative contributions of signal, noise, and clutter to the ACF variance is not taken into consideration. The FPFM variance expressions include contributions of signal, noise, and clutter. The clutter is estimated using the maximal power-based self-clutter estimator derived by *Reimer and Hussey* (2015). The FPFM was successfully implemented and tested using synthetic ACFs generated with the radar data simulator of *Ribeiro et al.* (2013b). The fitted parameters and the fitted-parameter errors produced by the FPFM are compared with the current SuperDARN fitting software, FITACF. Using self-consistent statistical analysis, the FPFM produces reliable or trustworthy quantitative measures of the errors of the fitted parameters. For an SNR in excess of 3 dB and velocity error below 100 m/s, the FPFM produces 52% more data points than FITACF.

7.4 Introduction

Super Dual Auroral Radar Network (SuperDARN) radars can be used to study a wide variety of magnetospheric and ionospheric processes. A search of the Scopus database (<https://www.scopus.com>) on 22 July, 2017 revealed more than 600 published journal articles using SuperDARN data. For example: vorticity observed in F region SuperDARN convection maps can be used to estimate the Birkeland currents (e.g. *Sofko et al.*, 1995); individual radar data are used in studies of polar cap patches (e.g. *Dahlgren et al.*, 2012); SuperDARN

data have been used to infer magnetospheric reconnection rates (e.g. *Hubert et al.*, 2006). Furthermore, SuperDARN data may also be ingested into assimilative models, as in *Cousins et al.* (2015) where data from SuperDARN and data from the Active Magnetosphere and Planetary Electrodynamics Response Experiment (AMPERE) were combined to produce maps of both the electrostatic and magnetic vector potentials.

It is clear that SuperDARN contributes an invaluable dataset to the space physics research community. Fitted-parameter errors that can be trusted as quantitative measures of the error in fitted parameters will improve SuperDARN data products and are therefore needed. Historically, due to limitations of available computer power in the 1980s and 1990s, when SuperDARN fitting software was originally developed, *ad hoc* conditions and estimates were implemented in the software. In an effort to eliminate self-clutter, SuperDARN fitting software uses *ad hoc* conditions to filter what are referred to as “bad” lags. The software also uses *ad hoc* variance expressions to weight the fitting of the remaining “good” ACF lags. This results in non-optimal estimates of the fitted-parameter errors (*Fisher*, 1958), because the *ad hoc* variance does not weight each individual lag estimate with the appropriate relative contributions of signal, noise, and clutter. The resulting fitted-parameter errors therefore do not vary as a function of signal-to-noise (SNR) and signal-to-clutter (SCR) as one would expect, with some good data being discarded and some bad data being erroneously fitted.

In this study we present a First-Principles Fitting Methodology (FPFM) based on previous work by *Reimer and Hussey* (2015) and *Reimer et al.* (2016). By employing physics based first-principles, FPFM does not rely on *ad hoc* criteria to deal with noise and self-clutter. The FPFM is implemented in a new SuperDARN fitting algorithm called LMFIT2 and has been tested and compared with existing SuperDARN software. It will be shown that the fitted-parameter errors from FPFM are accurate measures of the uncertainty in the fitted parameters. In addition, the amount of fitted data obtained with FPFM is increased compared to the amount of fitted data obtained using the current SuperDARN FITACF software. For clarity, we will refer to FPFM and its implementation in software, LMFIT2, as FPFM-LMFIT2. If FPFM-LMFIT2 is used to process all SuperDARN data, researchers will be able, for the first time, to quantitatively select which data from the SuperDARN dataset to use based on meaningful fitted-parameter errors. Since the FPFM-LMFIT2 is applied in

post-processing, the entire existing SuperDARN can be processed using FPFM-LMFIT2.

7.4.1 SuperDARN Measurements

The pulsed Doppler radars of SuperDARN measure the F region $\mathbf{E} \times \mathbf{B}$ plasma drift with minute-scale temporal resolution (*Greenwald et al.*, 1995; *Chisham et al.*, 2007). Operating in the high-frequency (HF) band, between 8 and 20 MHz, the radars record signals that are scattered by decameter-scale magnetic field-aligned ionospheric plasma density irregularities. To overcome the range-Doppler ambiguities associated with the long range (in excess of 1000 km range from the radar) and high velocity (of the order of 1 km/s) ionospheric targets, SuperDARN radars use the multiple-pulse technique of *Farley* (1972), which involves transmitting a sequence of pulses with non-redundant spacings at a variety of integer multiples of a base inter-pulse period. Complex voltage samples are decoded from scattered signals with a quadrature receiver. Time-lagged voltage samples from each pulse-sequence are used to estimate the time-dependent auto-correlation function (ACF) of the received signals to reveal the spectral properties of ionospheric irregularities (*Villain et al.*, 1996). To reduce both noise and statistical fluctuations, ACFs from several pulse-sequences (typically 30 or so) are combined to produce a mean ACF. Using K pulse sequences, the i -th lag of the mean ACF, $\hat{R}(\tau_i)$, is estimated from the voltage samples using

$$\hat{R}(\tau_i) = \frac{1}{K} \sum_{k=1}^K V_k(t + \tau_i) V_k^*(t), \quad (7.1)$$

where $V_k(t + \tau_i)$ and $V_k(t)$ are two voltage samples separated by a lag time of τ_i and the asterisk denotes the complex conjugate (e.g. *Villain et al.*, 1987; *Reimer et al.*, 2016). For SuperDARN, the lag times, τ_i , are integer multiples of the base inter-pulse period of the transmitted multiple-pulse sequence (typically 1800 or 2400 μ s). For simplicity, we will refer to the mean ACF as the ACF. Estimating all lags of the ACF using Equation 7.1 yields complex numbers describing the mean lag-products of the ACF at lag times τ_i . For a pulse sequence with L pulses, the multiple-pulse technique enables $L(L - 1)/2$ lags of the ACF to be estimated, which allows the time-dependent behaviour of the ACF to be observed (e.g. *Farley*, 1972; *Villain et al.*, 1996; *Reimer and Hussey*, 2015). From the ACF the echo power,

spectral width, and line-of-sight Doppler velocity are extracted using error-weighted least-squares fitting techniques (*Ribeiro et al.*, 2013a). Figure 7.1 is an example of a measured ACF (blue dots) from the SuperDARN radar in Clyde River, Nunavut. The top left and top right panels in the figure show the real and imaginary components of the ACF, respectively. The magnitude ($|\hat{R}(\tau_i)|$) and phase ($\angle \hat{R}(\tau_i)$) of the ACF are plotted in the bottom left and bottom right panels, respectively, and were obtained using

$$|\hat{R}(\tau_i)| = \sqrt{\left(\mathcal{R}[\hat{R}(\tau_i)]\right)^2 + \left(\mathcal{I}[\hat{R}(\tau_i)]\right)^2} \quad \angle \hat{R}(\tau_i) = \tan^{-1} \left(\frac{\mathcal{I}[\hat{R}(\tau_i)]}{\mathcal{R}[\hat{R}(\tau_i)]} \right), \quad (7.2)$$

where $\mathcal{R}[\hat{R}(\tau_i)]$ and $\mathcal{I}[\hat{R}(\tau_i)]$ denote the real and imaginary components of $\hat{R}(\tau_i)$, respectively. Note that the magnitude and phase given by Equation 7.2 are not the mean magnitude and mean phase of the ACF, but rather the magnitude and phase of the mean ACF. This distinction is important since the mean magnitude/phase and magnitude/phase of the ACF do not have the same statistical properties. The real and imaginary components plotted in Figure 7.1 exhibit a decaying complex sinusoidal behaviour. The magnitude decays in a similar manner, while the phase exhibits a linear wrapping behaviour.

7.4.2 Noise and Clutter

In general, the ACF measured by a SuperDARN radar contains contributions of signal, noise, and clutter. Due to the statistical nature of the scattering process, the signal contribution to the ACF has a sampling uncertainty. The presence of noise and clutter results additional uncertainty in extracting information about the signal. This uncertainty is quantitatively described by the variance of the ACF with both noise and clutter increasing the variance. It is therefore important to quantitatively account for noise and clutter in SuperDARN measurements in order to accurately extract signal information from the ACF.

Noise is generally only significant in zero-th lag of the ACF. This is because the first lag after lag-zero occurs at a time that is much longer than the impulse response of the final stage of filtering used for SuperDARN radars. For example, the last filter in the receiver chain is typically approximately 300 μs wide where as the shortest lag time is typically 1800 μs . If

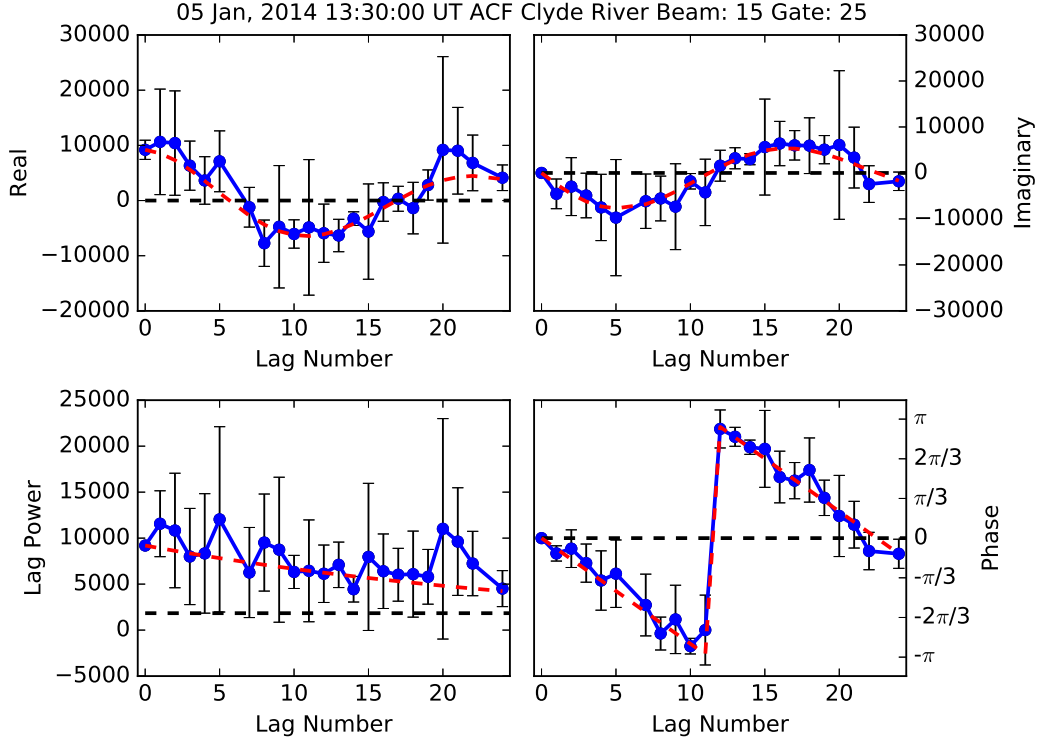


Figure 7.1: The real, imaginary, magnitude, and phase components of a typical ACF estimated from voltage data obtained with the Clyde River SuperDARN radar. The error bars were calculated using the methodology described in *Reimer et al. (2016)*. The blue solid line and dots show the ACF data and the red dashed line indicates the curve of the best least-squares error-weighted fit. The red dashed curves are plotted using Equation 7.3 with best fit parameters of $P = 9194.75$ arb, $w_d = 95.8$ m/s, and $v_d = -413.9$ m/s.

this were not the case, white noise passing through the filter would contribute a non-zero real component to more lags of the ACF than just the zero-th lag. Estimates of the noise power in a beam are currently made by averaging the lag-zero power from the 10 range gates with the lowest lag-zero power.

Self-clutter is a type of signal-like noise that is a by-product of the waveform that is transmitted by a radar. For SuperDARN radars, self-clutter is the dominant form of clutter, and it is produced by the multiple-pulse technique (*Farley, 1972; Barthes et al., 1998; Ponomarenko and Waters, 2006; Reimer and Hussey, 2015*). When more than one pulse is transmitted, echoes from multiple ranges can arrive simultaneously at the receiver. The

received signal is the sum of the signals that arrive together. Signals received at different times have different sets of interfering range gates contributing to self-clutter. The variance of each signal will be different than for any other signals and thus the self-clutter produces a heteroscedastic contribution to the ACF lag variance.

When processing multiple-pulse radar data, it is assumed that the physical processes that scatter signals in disjoint regions of the ionosphere are uncorrelated. Under this assumption, the correlation between signals from different locations randomly fluctuate such that they cancel out in the mean ACF, but the variance of the mean ACF is still affected by self-clutter. If this assumption is valid, then for K transmitted pulse-sequences the contribution of self-clutter, C , to a lag of the mean ACF is reduced to C/\sqrt{K} . SuperDARN radars routinely estimate ACFs from $K \approx 30$ pulse-sequences, which according to the assumption should attenuate the self-clutter contribution by ≈ 7.4 dB. SuperDARN radars routinely receive echoes with a SNR greater than 20 dB, so self-clutter originating from such strong sources may not be sufficiently attenuated. Additionally, *Reimer and Hussey (2015)* recently demonstrated that self-clutter caused by signals with long correlation time and zero-Doppler characteristics, such as ground scatter or long-lived plasma irregularities with velocity perpendicular to the look direction of the radar, do not adhere satisfy this assumption. All of this means that self-clutter may contribute significantly to the ACF and must be properly accounted for when processing SuperDARN data.

In the current SuperDARN fitting software, the contribution to the received signal by self-clutter from other range gates is not estimated (*Ribeiro et al., 2013a*). Instead, the current software discards lags where signals from interfering ranges are believed to dominate the signal from the discarded range. More details regarding the treatment of self-clutter in current SuperDARN processing are presented in Section 7.6.

7.4.3 Fitting for Ionospheric Parameters

Estimates of the lag-zero signal power, P , spectral width, w_d , and Doppler velocity, v_d , are obtained from least-squares fits of either the magnitude and phase, or the real and imaginary component of the ACF (e.g. *Ribeiro et al., 2013a*). Estimates of the ACF lags are heteroscedastic (e.g. *Reimer et al., 2016*), so SuperDARN fitting software must utilize

weighted least-squares fitting techniques since ordinary (non-weighted) least-squares fits do not provide an efficient estimate of the fitted parameters for heteroscedastic data (*Haines, 1978*). Usually, a single Doppler-peaked scatterer model is used to fit to the ACF, such as

$$R(\tau_i) = P e^{-2\pi w_d \tau_i / \lambda} e^{j4\pi v_d \tau_i / \lambda} \quad (7.3)$$

with λ being the wavelength of the transmitted radar pulses. The red dashed curves in Figure 7.1 show the model given by Equation 7.3 fitted to ACF data. Since $R(\tau_i)$ is complex valued, one may write it in terms of real and imaginary components, $R(\tau_i) = R_r(\tau_i) + jR_i(\tau_i)$, in terms of the lag-zero power and normalized ACF, $R(\tau_i) = P\rho(\tau_i)$, or in terms of the lag-zero power and the real and imaginary components of the normalized ACF, $R(\tau_i) = P(\rho_R(\tau_i) + j\rho_I(\tau_i))$. The functional form of the ACF model may vary depending on how the plasma irregularities decay (e.g. *Villain et al., 1996*) or depending on the presence of mixed scatter (e.g. *Ponomarenko et al., 2008*). Equation 7.3 is the model for the exponentially decaying ACF that is typically used in SuperDARN fitting software. In this study we will use Equation 7.3 to validate the FPFM-LMFIT2 technique, but the results obtained do not depend on the particular ACF model that is used.

7.5 Variance of ACF Lag Estimates

The fitted parameters and the fitted-parameter errors are dependent on the variance of the lags of the ACF (referred to as the ACF variance). For example, as the ACF variance increases so do the errors in the fitted parameters. Here we discuss several methods for estimating the ACF variance. The dependence of fitted-parameter error on ACF variance will be explored in more detail in Section 7.7.

Each lag of the ACF is composed of signal, noise, and clutter (e.g. *Ulaby et al., 1982*). The variance of a lag increases/decreases as the noise and clutter increases/decreases (e.g. *Farley, 1969*). It is therefore paramount to accurately estimate the noise and clutter in order to accurately estimate the ACF variance. *Reimer et al. (2016)* showed that in the absence of noise and clutter, the ACF lags are heteroscedastic, meaning that the variance of each lag differs. Since SuperDARN radars use the multiple pulse technique, an additional

heteroscedasticity due to self-clutter is contributed to the ACF lags. Self-clutter is discussed in greater detail in Section 7.6.

7.5.1 Ad Hoc Variance Estimates

Currently, the SuperDARN FITACF software uses *ad hoc* variance expressions for weighting the least-squares fit of the lags of the ACF. The software performs linear weighted fits to the logarithm of the ACF magnitude and to the unwrapped phase of the ACF. In a technical report by Kile Baker (White Paper on FITACF, 2003), it is stated that the *ad hoc* variance was implemented due to a lack of *a priori* knowledge of the ACF variance. It was reasonably assumed that the variance of the magnitude and phase of lag i should be inversely proportional to the magnitude (or power) of lag i , which we will denote P_i . The *ad hoc* variance for lag i , $\sigma_{adhoc,i}^2$, suggested by Baker is

$$\sigma_{adhoc,i}^2 = \frac{\bar{P}^2}{P_i^2} \sigma_{all}^2, \quad \sigma_{all}^2 = \frac{n}{n-1} \frac{\sum P_i^2 (d_i - m_i)^2}{\sum P_i^2} \quad (7.4)$$

where \bar{P} is the average magnitude of all lags of the ACF. The i -th measured ACF lag is d_i , the value of the best-fit model for the ACF lag is m_i , and the magnitude in each lag is P_i . The magnitude-weighted average difference between the ACF and the best-fit model is given by σ . In order to determine $\sigma_{adhoc,i}$, one must first fit the phase of the ACF to determine the best fit model values, m_i , so that σ_{all} can be calculated.

The *ad hoc* variance given in Equation 7.4 is used to fit both the magnitude and the phase of SuperDARN ACFs, but Equation 7.4 does not accurately represent the variance of the magnitude nor the variance of the phase. This can be seen by comparing the inverse-square behaviour of the ad-hoc variance expression, $1/P_i^2$, given by Equation 7.4 with the more complicated behaviour of the variance of the magnitude and phase derived by and presented in *Reimer et al. (2016)*. Equation 7.4 also does not accurately account for variance contributions from noise and clutter. This is why the *ad hoc* filtering conditions were implemented as an attempt to remove lags contaminated with self-clutter.

7.5.2 Variance Estimates from First Principles

Following *Farley* (1969) and *Woodman and Hagfors* (1969), *Reimer et al.* (2016) presented the statistical characteristics of the lag estimates of the SuperDARN ACF. Using a model that characterizes the voltage samples received by SuperDARN radars as correlated Gaussian random variables, *Reimer et al.* (2016) derived the exact analytic expressions for the variances of the real and imaginary components of the SuperDARN ACF. Expressions for the variance of the magnitude and variance of the phase were also derived. The variance of the real (σ_R^2) and imaginary (σ_I^2) components of an ACF are given by

$$\sigma_R^2 = P^2 \left(\frac{1 - |\rho|^2}{2K} + \frac{\rho_R^2}{K} \right) \quad (7.5)$$

and

$$\sigma_I^2 = P^2 \left(\frac{1 - |\rho|^2}{2K} + \frac{\rho_I^2}{K} \right) \quad (7.6)$$

where P is the lag-zero power, $\rho = \rho_R + j\rho_I$ is the normalized ACF (e.g. $\rho(\tau_i) = R(\tau_i)/P$ and *Farley* (e.g. 1969)), and K is the number of pulse sequences averaged together to produce the ACF. Here we have dropped the τ_i notation that was used in Equation 7.3 for simplicity of presentation. These expressions are valid for all values of $K \geq 1$. Equations 7.5 and 7.6 describe the true heteroscedastic nature of the ACF. They can be used to perform a weighted least-squares fit of the real and imaginary components of the SuperDARN ACF (not the magnitude and phase of the ACF as is currently done). There are several advantages to fitting the real and imaginary components of the SuperDARN ACF, compared to fitting the magnitude and phase, which are discussed in Section 7.7, so we will not discuss further the variance of the magnitude and phase.

Notice that Equations 7.5 and 7.6 depend on the normalized ACF, which depends on the best-fit values of w_d and v_d from the fitting procedure. This means that one needs *a priori* knowledge of fitted parameters before one may fit for said parameters. A solution to this apparent dilemma is to perform an iterative fitting procedure, where the fitted parameters from one fitting step are used to estimate the variance for the next fitting step. The initial variance estimates may be provided by an approximate variance estimate.

A first order estimate for the variance of the real and imaginary components of the

SuperDARN ACF is readily available in the literature (e.g. *Farley, 1969; Ulaby et al., 1982; Skolnik, 1980*). The first-order estimate of the variance of real and imaginary components (σ_{FO}^2) is given by

$$\sigma_{FO}^2 = \frac{P^2}{K}. \quad (7.7)$$

Equations 7.5, 7.6, and 7.7, provide a solid foundation built on first-principles for performing a properly weighted fit of the real and imaginary components of the SuperDARN ACF.

Including Noise and Self-Clutter

As with all radars, the signals that SuperDARN measures include contributions due to noise and clutter. The noise and clutter increase the variance of the lags of the SuperDARN ACF. To include the contribution of noise and clutter in Equations 7.5, 7.6, and 7.7 a normalized ACF in the presence of noise and clutter $\rho'(\tau_i)$ *Farley* (e.g 1969) and total power P' must be substituted for $\rho(\tau_i)$ and P (e.g. *Farley, 1969; Reimer et al., 2016*). We will refer to $\rho'(\tau_i)$ as the reduced normalize ACF, which is given by

$$\rho'(\tau_i) = \frac{P\rho(\tau_i)}{P + N + C}. \quad (7.8)$$

The total power must be used since the magnitude of the lags of the ACFs measured by SuperDARN radars include contributions of signal, noise, and clutter power. The expression for the total received power is given by

$$P' = P + N + C. \quad (7.9)$$

As discusses in subsection 7.4.2, noise only contributes to the zeroth-lag so $N = 0$ is used in Equation 7.9 for all other lags.

7.6 Estimating Self-Clutter

Since Equation 7.4 does not account for self-clutter, the current SuperDARN fitting software attempts to identify lag estimates that are dominated by self-clutter so that those lags can be eliminated from the fitting process. Elimination of the self-cluttered lags using an *ad hoc* filtering criterion was found to improve estimates of spectral width (*Ponomarenko and*

Waters, 2006). To filter the self-clutter contaminated lags, *Ponomarenko and Waters* (2006) suggested that the lag-zero power, P_n , of the n potentially interfering ranges, be compared with the lag-zero power, P_R , of the range being fitted. If the power in any of the n interfering ranges is larger than P_R , the lag is discarded. The *ad hoc* criterion used to reject a lag due to strong self-clutter can be written as

$$P_R < P_n. \quad (7.10)$$

Figure 7.2 shows an example of a SuperDARN ACF with discarded lags denoted by black triangles. The discarded lags are not necessarily of any poorer quality or noisier in appearance than the lags that have not been discarded. It is clear that the filtering criterion imposes a false dichotomy of “good” and “bad” lags, when they are in reality somewhere in between, depending on the SCR.

In addition, it can be argued that Equation 7.10 should have been written as

$$P_R < \frac{P_n}{\sqrt{K}} \quad (7.11)$$

since the multiple-pulse technique reduces self-clutter by a factor of $1/\sqrt{K}$. Regardless of whether this would improve the *ad hoc* filtering or not, such criterion will always be subjective. Additionally, even a lag heavily contaminated by self-clutter contains some useful information that is discarded by a filtering criterion. For these reasons, *ad hoc* criteria are undesirable. Finally, an *ad hoc* criterion does not help determine the self-clutter term, C , which is required to use Equations 7.8 and 7.9.

Motivated by the need to determine the contribution of self-clutter power to the variance of ACF lag estimates, a maximal self-clutter power estimator was derived and validated using a first-principles statistical methodology (*Reimer and Hussey*, 2015). The estimator includes contributions from the interfering ranges that contribute to each of the voltage samples ($V_k(t + \tau)$ and $V_k(t)$ in Equation 7.1, respectively) that are used to estimate a lag of the ACF (*Reimer and Hussey*, 2015). The maximal power-based self-clutter estimator (MPSE) is given by

$$\hat{C}_{\text{MPSE}}(\tau_i) = \sum_{n=1}^N \sqrt{\hat{P}_n \hat{P}} + \sum_{m=1}^M \sqrt{\hat{P} \hat{P}_m} + \sum_{n=1}^N \sum_{m=1}^M \sqrt{\hat{P}_n \hat{P}_m}, \quad (7.12)$$

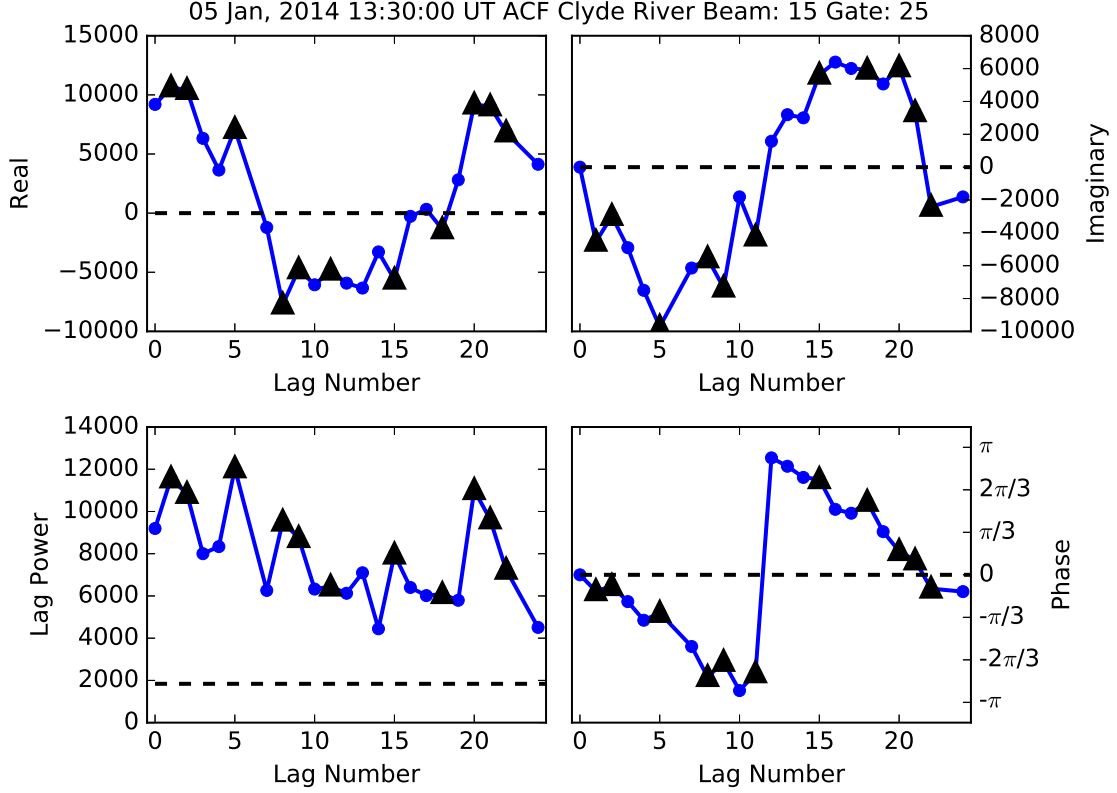


Figure 7.2: An illustration of the self-clutter filtering criterion of *Ponomarenko and Waters* (2006). Black triangles indicate lags that satisfy Equation 7.10.

where \hat{P}_n and \hat{P}_m are the estimates of lag-zero power at the n -th and m -th interfering ranges, respectively. Voltage samples $V_k(t + \tau)$ and $V_k(t)$ will have up to N and M interfering range contributions (*Reimer and Hussey* (see 2015)). The MPSE provides an upper limit estimate of the self-clutter contained in a lag of the ACF. The MPSE can be used with a measured lag-zero power or a fitted lag-zero power. It is important to note that the lag-zero power is a parameter that is saved for all SuperDARN radars, so the MPSE can be used to reprocess the entire historical SuperDARN dataset.

When used in conjunction with Equations 7.5, 7.6, 7.8, and 7.9, the MPSE enables the ACF variance to be estimated from first-principles for use in error weighted fitting. This means that it is possible to perform error weighted fitting of SuperDARN ACFs without using any *ad hoc* criteria (like Equation 7.10) or any *ad hoc* variance expressions (like Equation 7.4).

7.7 Current SuperDARN Fitting Software

Presently, three fitting routines are used to process SuperDARN ACFs to obtain the fitted parameters, power, P , velocity, v_d , and spectral width, w_d . These routines are called FITACF, FITEX2, and LMFIT (*Ribeiro et al.*, 2013a). The fitting routines rely on least squares regression techniques to fit SuperDARN ACFs. Least squares fitting involves minimization of the sum of the square of residuals, or minimization of the χ^2 sum

$$\chi^2 = \sum_i^N \left(\frac{d_i - m_i}{\sigma_i} \right)^2, \quad (7.13)$$

where d_i and m_i are the i -th data and model values, respectively, and σ_i is the square root of the variance of the i -th data point. The i -th residual is the weighted difference $(d_i - m_i)/\sigma_i$. Fitting the lags of an ACF (the d_i values) to some model (such as Equation 7.3, therefore the model m_i values) is accomplished by varying the parameters of the model (P , w_d , and v_d) until a global minimum value for χ^2 is found. The parameters that minimize χ^2 are known as the “best-fit” parameters.

If the model is linear, analytic solutions for the best-fit parameters are readily available, which simplifies the fitting process (e.g. *Press et al.*, 1986). FITACF performs a linear least-squares fit to the logarithm of the ACF magnitude to obtain the fitted lag-zero power and spectral width. The fitted velocity is obtained from a linear fit to the unwrapped phase. Historically, implementing analytic solutions for linear least-squares fitting of SuperDARN was favoured because linear fitting is much less computationally expensive compared with non-linear alternatives. However, with the availability of modern computational resources, non-linear least-squares techniques are now more attractive and more easily implemented. FITEX2 performs a linear least squares fit to obtain the power and spectral width in the same way as FITACF, but FITEX2 performs a non-linear least squares fit of the phase to obtain the velocity. FITEX2 generates candidate wrapped-phase models for several velocities and searches for the model, and therefore the best-fit velocity, that produces the minimum χ^2 . If the model is non-linear (as in Equation 7.3), one can use the Levenburg-Marquardt algorithm to minimize Equation 7.13 (*Levenburg*, 1944; *Marquardt*, 1963; *Press et al.*, 1986). It is important to note that the Levenburg-Marquardt algorithm requires an initial guess of

best fit parameters and will only converge on a *local* minimum in χ^2 . The local minimum is not necessarily the *global* minimum. This means that if the χ^2 surface contains multiple minima, the choice of initial guess will determine whether a local or global minimum is found. LMFIT performs a non-linear least squares fit to an ACF model (similar to Equation 7.3) using the Levenburg-Marquardt algorithm and FITEX2 is used to provide an initial guess of the best fit parameters. LMFIT assumes a lag variance of 1 for all lags.

7.7.1 Challenges of Fitting the Magnitude and Phase

In Section 7.10.1 it will be shown that the fitted data produced by FITACF is of good quality, especially when the SNR is large, despite the challenges discussed here. While the FITACF performs well with the *ad hoc* self-clutter filtering criterion enabled, some case studies are presented to demonstrate that incorrect fitted velocities can result from FITACF due to the *ad hoc* criteria and *ad hoc* variance it uses.

Out of the FITACF, FITEX2, and LMFIT algorithms, FITACF is the most commonly used SuperDARN fitting routine. It uses an error-weighted linear least-squares method to fit the magnitude and phase of the ACF (*Ribeiro et al.*, 2013a). The signal power, P , and spectral width, w_d , are obtained by fitting the logarithm of the magnitude of the ACF to the logarithm of the magnitude of Equation 7.3. The Doppler velocity, v_d , is obtained by fitting the phase of the ACF to the phase of Equation 7.3. FITACF uses the self-clutter filtering criterion discussed by *Ponomarenko and Waters* (2006) to discard lags dominated by self-clutter. FITACF uses an *ad hoc* variance (Equation 7.4) to perform the error-weighted linear fits. There are two difficulties associated with fitting the magnitude and phase of the ACF, but these challenges are not present when fitting the real and imaginary components of the ACF.

One of the challenges is related to the fundamental statistical characteristics of the magnitude of the ACF when the value of the ACF lags approach zero. *Reimer et al.* (2016) demonstrated that as the correlation in the ACF approaches zero ($\rho \rightarrow 0$) the value of the magnitude of the ACF lags does not approach zero. This is because the magnitude of the ACF lags are Rayleigh distributed at zero correlation and the mean of a Rayleigh distribution is generally non-zero. FITACF uses the logarithm of the magnitude of Equation 7.3 to fit the

logarithm of the magnitude of the ACF, but the magnitude of Equation 7.3 is an exponential function, which approaches 0 for large negative values of the argument. Using the magnitude of Equation 7.3 as a model for the magnitude of the ACF is therefore not valid when the correlation is weak. The model given by Equation 7.3 is only valid for fitting to the real and imaginary components of the ACF. It is incompatible with fitting the magnitude of the ACF. In FITACF, the approach to dealing with this problem has been to use an additional *ad hoc* criterion to discard lags that have a lag power below P/\sqrt{K} .

A second difficulty encountered in FITACF involves fitting the phase of the ACF. Since the phase may be wrapped, due to the 2π ambiguity when estimating the phase from the real and imaginary components of the ACF, the phase must be unwrapped before a linear fit may be performed. For example, see the phase wrapping between lags 11 and 12 in Figure 7.1. Unwrapping the phase is inherently challenging due to the presence self-clutter in the ACF. To clearly demonstrate this point, 75,000 ACFs were synthesized using the SuperDARN radar data simulator developed by *Ribeiro et al.* (2013b). We simulated 1000 ACFs at 75 different uniformly distributed velocities between 0 m/s and 2000 m/s, both with and without strong self-clutter. All ACFs were simulated with $P = 10^4$, $w_d = 50$ m/s, $K = 18$, and a noise power equal to 104.5. Power is in arbitrary units since SuperDARN hardware is not calibrated to measure power relative to a calibrated source. Using a modified version of FITACF that had the *ad hoc* self-clutter filtering removed, the simulated ACFs were then using FITACF. The average fitted velocity of the 1000 fitted ACFs at each of the 75 simulated velocities was compared with the expected velocity. This comparison is shown in Figure 7.3 where a line with a slope of one corresponds to a fitted velocity that is exactly the same as the simulated (expected) velocity. Without self-clutter (top left plot), FITACF velocities are nearly identical to the simulated velocities. In the presence of self-clutter (green dots, top right plot), FITACF without the self-clutter filtering was unable to unwrap the phase, resulting in incorrect fitted velocities (note that below ~ 250 m/s the fits are ok since no phase jumps occur below this speed). This clearly demonstrates that the phase of the ACF cannot be unwrapped by the FITACF algorithm without the use of some criterion to eliminated the self-clutter. The orange dots in the top right plot of Figure 7.3 show the results of processing the same simulated ACFs using FITACF with the self-clutter

filtering turned on. The orange dots show that, even with the *ad hoc* criterion to discard cluttered lags, for velocities larger than approximately 800 m/s, FITACF is much less capable at unwrapping the phase. *Ribeiro et al.* (2013a) also found that FITACF has difficulty with fitting ACFs with velocities greater than 1000 m/s (see Figure 7 of *Ribeiro et al.* (2013a) and note the large RMS velocity errors at high velocity). Finally, we attempted to perform phase unwrapping using an error-weighted unwrapping algorithm where the phase variance used was determined using the methodology of *Reimer et al.* (2016), but this was also found to be unable to consistently unwrap the phase (not shown).

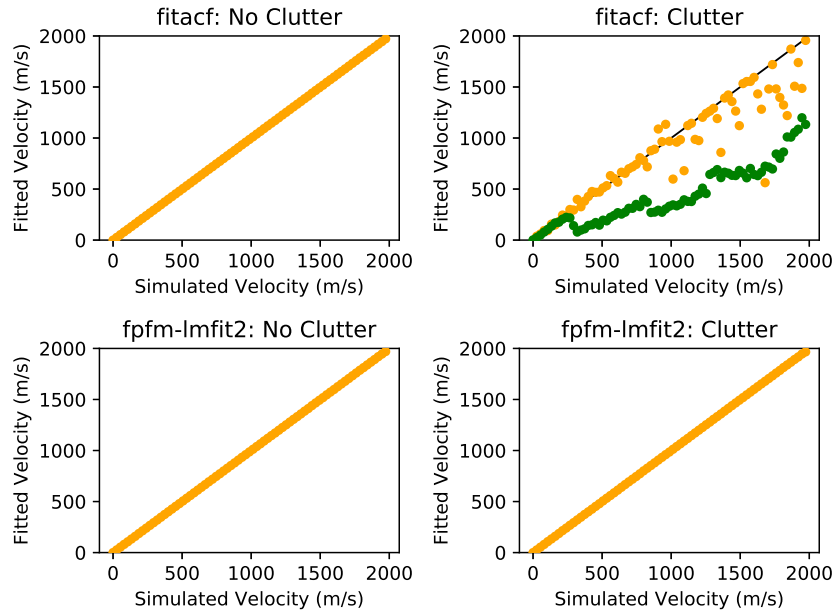


Figure 7.3: This figure illustrates the difficulties associated with unwrapping the phase of ACF data in FITACF (top panels) versus fitting to the real and imaginary components of the ACF data in FPFM-LMFIT2 (bottom panels), both with and without strong self-clutter in the ACF data. The black line in each panel indicates the expected one-to-one trend of fitted velocities versus simulated velocities. Each dot represents the mean fitted velocity obtained from 1,000 simulated ACFs. In green dots in the top right panel, show fitted velocities from a version of FITACF without the self-clutter criterion, which illustrates that the phase unwrapping is significantly more challenging without using the self-clutter criterion. In the top panels, the orange dots represent mean fitted velocities obtained using the standard unmodified version of FITACF. The bottom panels show that fitting the real and imaginary components of the ACF, as performed by FPFM-LMFIT2, results in correctly fitted velocities, even with strong self-clutter present.

7.8 A First-Principles Fitting Methodology

To address the issues identified in Section 7.7, a First-Principles Fitting Methodology (FPFM) was developed. The FPFM uses Equations 7.5, 7.6, 7.7, 7.8, 7.9 and 7.12 to perform an error-weighted fit to the real and imaginary components of the ACF. FPFM involves an iterative fitting process where the best-fit parameters from the previous fit are used to provide initial estimates of the ACF variance for the next fit. Using Equations 7.5 and 7.6 to determine the variance of the real and imaginary components of the ACF requires knowledge of the fitted power, spectral width, and velocity, so the first fit is performed using the first-order error given by Equation 7.7. Subsequent fits are performed using Equations 7.5 and 7.6. One could perform many iterations of fitting like this until the fitted parameters converged within a user-specified tolerance, but in practice we observed that the process converged within two iterations.

The software implementation of the FPFM was developed and called LMFIT2. The name LMFIT2, for Levenburg-Marquardt Fitter version 2, was chosen to follow the naming convention started by another fitting software, LMFIT. LMFIT was developed by *Ribeiro et al.* (2013a) to perform an error weighted non-linear least-squares fit of the real and imaginary components of the ACF. LMFIT uses the Levenburg-Marquardt algorithm to simultaneously fit for the power, spectral width, and velocity. However, LMFIT employs the *ad hoc* self-clutter filtering criterion and the *ad hoc* variance expression (with $\sigma^2 = 1$ in Equation 7.4).

The new software, LMFIT2, also uses the Levenburg-Marquardt algorithm to perform an error-weighted fit of the ACF, but it performs the ACF fits using the FPFM. A non-linear fit of the ACF using the model given by Equation 7.3 results in a χ^2 surface with multiple minima in the velocity dimension, as shown in Figure 7.4. The choice of the initial guess for the model will determine whether a local minimum or global minimum is converged upon. For example, in Figure 7.4, choosing an initial guess for velocity of 0 m/s would result in the Levenburg-Marquardt algorithm finding the local minimum near 300 m/s instead of the global minimum at -413.9 m/s. In LMFIT this issue was addressed by using the fitted velocity from FITEX2 as an initial guess for LMFIT. LMFIT2 utilizes an alternate approach. The maximum unambiguous velocity, usually called the Nyquist frequency, that can be measured

by the multiple-pulse sequence is given by

$$v_{nyquist} = \frac{\lambda}{4\tau_B} \quad (7.14)$$

where λ is the radar transmitted wavelength and τ_B is the base inter-pulse period of the transmitted multiple-pulse sequence. The domain of measured Doppler velocity is bounded by $\pm v_{nyquist}/2$. Since the range of possible velocities is bounded, and the number of minima on the χ^2 surface is finite, a local versus global minimum determination can be made from the output of the Levenburg-Marquardt algorithm by using multiple initial guesses that are uniformly distributed between $-v_{nyquist}/2$ and $+v_{nyquist}/2$. The χ^2 for the fitted velocities resulting from each initial guess can then be compared to find the fitted velocity that produces the global minimum χ^2 . The number of initial guesses needs to be larger than the number of minima in the chi-squared surface. The technique of using multiple uniformly distributed initial guesses also enables better estimation of the confidence interval for the fitted velocity, as will be discussed later.

Including the multiple initial velocity guesses technique, FPFM-LMFIT2, can be summarized as follows:

For each SuperDARN ACF:

1. Estimate self-clutter for each lag of the ACF using the MPSE (Equation 7.12).
2. Estimate the first-order errors for the ACF using Equations 7.7 and 7.9.
3. Perform an error-weighted fit of the ACF using the Levenburg-Marquardt algorithm using multiple initial guesses for the velocity
4. Determine which fitted parameters result in the global minimum χ^2
5. Use best fit parameters to estimate real and imaginary variances using Equations 7.5, 7.6, 7.8, and 7.9
6. Perform an error-weighted fit of the ACF using the Levenburg-Marquardt algorithm using multiple initial guesses for the velocity.
7. Determine which fitted parameters results in the global minimum χ^2 .

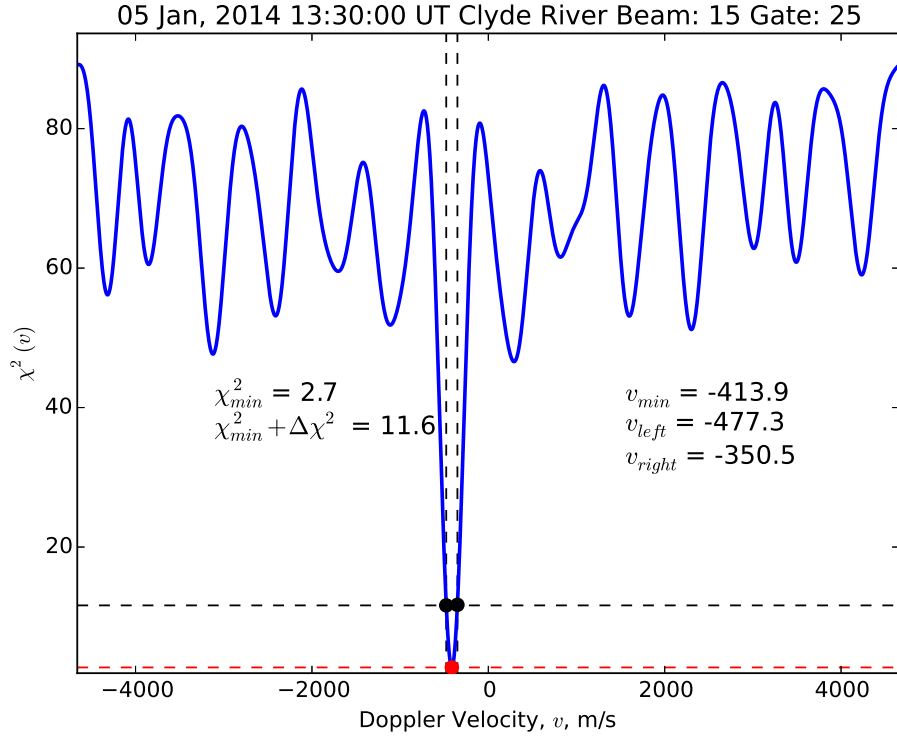


Figure 7.4: An example curve of least-squares χ^2 as a function of velocity (v_d , a fitted parameter). Note that the curve reaches a minimum value at $\chi^2 = 2.7$ corresponding to $v_d = -413.9$ m/s (red dot). The CCBCL method (see text for details) is also illustrated for a $\Delta\chi^2 \approx 9$. Tracing the curve from χ^2_{min} to $\chi^2_{min} + \Delta\chi^2$ results in bounds for a 3σ confidence limit (black dots intersected by dashed lines). On the left side of the fitted value, the bound extends to $\Delta v_{left} = 63.4$ m/s ($v_{left} = -477.3$ m/s) and on the right side it extends to $\Delta v_{right} = 63.4$ m/s ($v_{right} = -350.5$ m/s), resulting in the best fit $v_d = -413.9$ m/s ± 63.4 m/s.

8. Repeat from step 5 until convergence criteria are met.
9. Determine the confidence intervals for the best-fit parameters.

It is important to note that because FPFM-LMFIT2 performs fits of the real and imaginary components of the ACF, it does not need to unwrap the phase. FPFM-LMFIT2 is therefore able to determine the Doppler-velocity even when self-clutter is present. This is clearly shown in Figure 7.3 when comparing the top plots (FITACF fitted velocities) and the bottom plots (LMFIT2 fitted velocities). FPFM-LMFIT2 is able to obtain accurate fitted velocities even when self-clutter causes inaccurate fitted velocities from FITACF.

7.8.1 Fitted-Parameter Errors

When the statistical characteristics of the residuals in Equation 7.13 can be described by a Gaussian distribution, the confidence interval of the best-fit parameters can be estimated using the covariance matrix of the best-fit parameters (e.g. *Press et al.*, 1986). The covariance matrix contains the squares of the so-called standard errors, from which a confidence interval for each parameter may be constructed. The fitted-parameter errors are therefore actually a measure of how certain one can be that the true value of the parameters lie within the confidence intervals. When the residuals for the best-fit parameters are not small, they may no longer be Gaussian, requiring alternative techniques to construct the confidence interval.

The “constant chi-square boundary as confidence limits” (CCBCL) method may be used to determine the error in the best-fit parameters (*Press et al.*, 1986). The CCBCL method involves adding some chosen $\Delta\chi^2$ to the global minimum chi-squared, χ_{min}^2 . The magnitude of the difference between the value of a fitted parameter at $\chi_{min}^2 + \Delta\chi^2$ and the value of the fitted parameter at χ_{min}^2 determines the bounds of the confidence interval for the fitted parameter. As depicted in Figure 7.4, the bounds are determined with the CCBCL method by “stepping up” from the minimum χ^2 by some $\Delta\chi^2$ value. For example, stepping up by $\Delta\chi^2 = 9$ from χ_{min}^2 on the right side yields a 3σ equivalent error bar of 63.4 m/s (the figure shows a $\Delta\chi^2$ of 8.9 due to error introduced by a discrete velocity spacing). Even if the residuals are not Gaussian distributed, a confidence limit for a fitted parameter can be obtained using the CCBCL method, but must be interpreted as “the confidence limit assuming Gaussian distributed residuals” (*Press et al.*, 1986).

In LMFIT2, both the CCBCL method and the covariance matrix method of obtaining confidence intervals have been implemented. First, fitted-parameter errors are determined using the covariance matrix method. The fitted-parameter errors are the product of the standard error and $\sqrt{\Delta\chi^2}$ (*Press et al.*, 1986). In LMFIT2, $\Delta\chi^2 = 4$ was used, which corresponds to a 2σ or 95% confidence interval; however, the value of $\Delta\chi^2$ is subjective and can be modified depending on the level of confidence desired.

After using the covariance matrix method, LMFIT2 also uses the CCBCL method. The χ^2 values at each minima of the χ^2 surface, χ_{local}^2 , which were obtained using the Levenburg-

Marquardt algorithm and the multiple uniformly distributed initial velocities, are compared to see if:

$$\chi_{local}^2 < \chi_{min}^2 + \Delta\chi^2. \quad (7.15)$$

If Equation 7.15 is satisfied, then the fitted-parameter error is given by the largest difference between the fitted parameter value at χ_{local}^2 and the fitted-parameter value at χ_{min}^2 , instead of the covariance matrix method. For example, referring to Figure 7.4, if $\Delta\chi^2$ of 50 were desired, corresponding to an excessive significance of $\sqrt{50}\sigma$, then the local minima at ~ -3100 m/s and ~ 300 m/s both satisfy Equation 7.15. Since the largest difference between the velocities at the local minima and the global minimum is 2686.1 m/s, the $\sqrt{50}\sigma$ confidence interval for the fitted velocity would be 2686.1 m/s. If Equation 7.15 is not satisfied, then the error bar for the fitted velocity is specified using the covariance matrix method.

7.9 Synthetic ACF Tests of FPFM-LMFIT2

To assess the accuracy of FPFM-LMFIT2 fitted parameters and the reliability of the fitted-parameter errors, a synthetic ACF dataset was generated using the SuperDARN radar data simulator (*Ribeiro et al.*, 2013b). The radar data simulator is available in the software package RSTLite available on Github (<https://github.com/vtsuperdarn/RSTLite>). Two datasets were generated: an ideal dataset without any self-clutter and a pathological dataset with a large amount of self-clutter. The ideal dataset represents “best-case scenario” SuperDARN ACFs, and the pathological dataset represents “worst-case scenario” SuperDARN ACFs. Any fitting methodology that is able to reliably fit both the ideal and the pathological datasets should be capable of reliably fitting real SuperDARN ACFs. The synthetic datasets were generated by simulating the 8 pulse “katscan” pulse sequence with typical values corresponding to an inter-pulse period of 1800 μs , a pulse-width of 300 μs , $K = 18$, and a delay from the radar to the first range of 1200 μs . A transmission frequency of 10.537 MHz was used. ACFs with a variety of powers, spectral widths, and velocities were simulated.

Figure 7.5 presents the results of using FPFM-LMFIT2 to fit the ideal dataset. Each point on each line is an average result from 1000 fitted ACFs. In the left panel, all ACFs were simulated with $P = 10^4$ and a noise power of 104.5. 75 range gates were simulated with

different velocities, starting with 0 m/s for range gate 0 and linearly increasing to 1973 m/s at range gate 74. The blue lines indicate ACFs with $w_d = 50$ m/s, the red lines indicate ACFs with $w_d = 250$ m/s, and the green lines indicate ACFs with $w_d = 500$ m/s. The solid lines represent the magnitude of the difference between the fitted velocity and the true velocity (the true error). The dashed lines represent the fitted velocity error output by FPFM-LMFIT2, which in this case, the fitting error corresponds to the 95% confidence interval for the fit. The solid line and dashed line convention is used in all panels of the figure.

In the left panel of Figure 7.5, the fitting error and the true error (“error” and “diff”, respectively in the figure legends) follow the same trend and agree up to a multiplicative factor. The factor is approximately $\sqrt{2}$ since we chose to construct a 95% confidence interval $(2 - \sigma)$ and the true error is expected to be $1 - \sigma$ since we expect the fitting residuals to be small and Gaussian distributed with a variance of σ^2 . The fitting error is therefore an accurate estimate of the uncertainty in the fitted velocity. The fitting error and true error are not functions of velocity, but as the spectral width increases, the velocity error increases. Considering the variance of the real and imaginary components of the ACF (Equations 7.5, 7.6, 7.8, and 7.9), this is expected behaviour because the SNR, SCR, nor the correlation for any of the lags of the ACF are affected by the velocity, but vary with the spectral width (recall the model given in Equation 7.3). If FPFM-LMFIT2 had used a higher confidence interval, for example 3σ , the fitting error shown as the dashed lines in Figure 7.5 would have been larger. The opposite is true for a lower confidence interval.

The middle panel in Figure 7.5 is similar to the left panel, except that the fitting error and the true error are plotted as a function of spectral width. The ACFs were simulated with $P = 10^4$, a noise power of 104.5, and a spectral width linearly increasing from 0 m/s for range gate 0 to 1973 m/s for range gate 74. The blue lines indicate ACFs with $v_d = 50$ m/s, the red lines indicate ACFs with $v_d = 250$ m/s, and the green lines indicate ACFs with $v_d = 500$ m/s. Since the velocity error is not a function of the velocity (as seen in the left panel), the blue, red, and green lines all overlap. The spectral width of the ACFs determines the decorrelation time of the ACF (e.g. Equation 7.3), such that increasing the spectral width results in a larger number of lags that have a smaller SNR, smaller SCR, and smaller correlation. It follows that increasing the spectral width should increase the velocity error

and that is what is observed in the middle panel of Figure 7.5. In the right panel in Figure 7.5 the velocity error is plotted as a function of SNR. The ACFs were simulated using $w_d = 250$ m/s, $v_d = 250$ m/s, and linearly increasing SNR starting at 0 dB at range 0 and ending at 29.6 dB at range 74. As expected the velocity error is largest when the SNR is smallest and the velocity error is smallest when the SNR is largest. Simulations presented in Figure 7.5 demonstrate that FPFM-LMFIT2 is capable of accurately fitting ideal SuperDARN radar data, while also producing statistically self-consistent and accurate fitting errors.

Figure 7.6 shows the results of using FPFM-LMFIT2 to fit the synthetic pathological dataset, which contains an unrealistically (pathologically) large amount of self-clutter. The self-clutter was simulated by including contributions of signals from each of the 75 simulated range gates, according to the self-clutter characteristics of the katscan pulse-sequence (e.g. *Reimer and Hussey, 2015*). As the range increases, there are fewer self-clutter contributions since there are fewer ranges beyond the range of interest that could contribute self-clutter. The radar data simulator was not designed to simulate self-clutter contributions from ranges beyond the 75 range gates (*Ribeiro et al., 2013b*). The consequence of the decrease in the number of ranges contributing self-clutter with increasing range is shown most clearly in the right panel of Figure 7.6, where the velocity error is observed to decrease in a step-wise manner at SNR values of approximately 3 dB, 11 dB, 21 dB, 23 dB, and 25 dB, corresponding to ranges 9, 27, 51, 57, and 63. The steps occur starting with range 9 with additional steps occurring at later ranges that are a multiple of 6 further in range. The multiplicative factor is expected and is the ratio of the inter-pulse period and the pulse width (i.e. $1800 \mu\text{s}/300 \mu\text{s}$). This factor determines the spacing between transmitted pulses in units of range gate and therefore the ranges from which self-clutter originates (e.g. *Reimer and Hussey, 2015*).

Each panel in Figure 7.6 shows a similar behaviour as the corresponding panels in Figure 7.5, but with larger velocity errors. The confidence intervals were again constructed for 2σ . The increase in velocity errors from fitting the pathological dataset compared to the ideal dataset is due to the larger ACF variance caused by the addition of self-clutter to the ACFs. One would expect that the largest velocity error possible to be approximately the Nyquist velocity (± 1977 m/s for this synthetic data) and that is what is observed as the largest velocity error in each of the panels. In the left panel, the velocity error is independent of the velocity.

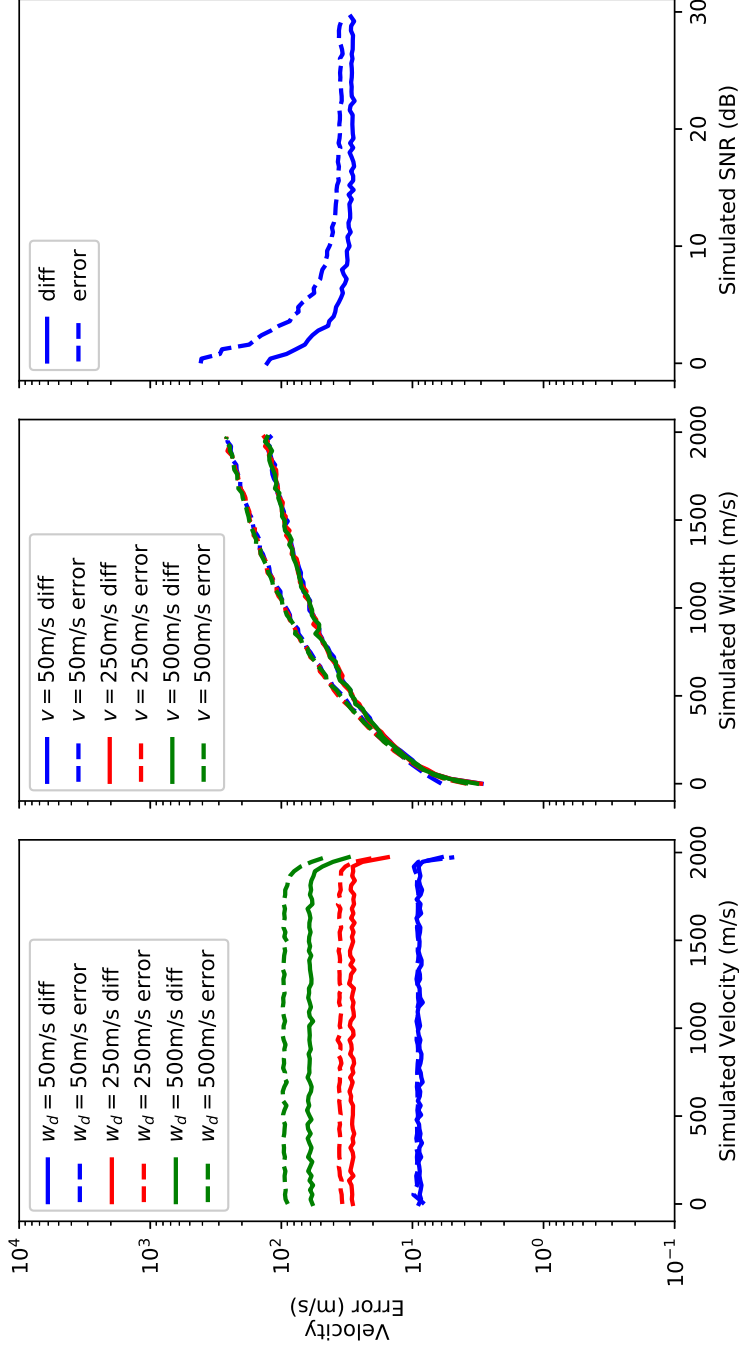


Figure 7.5: A comparison of the fitted velocity error (dashed lines) and true error (solid lines) for synthetic ACF data fitted by FPFM-LMFIT2. The synthetic data used to make these plots did not include self-clutter (the ideal dataset). The left panel shows fitted velocity error versus simulated velocity for synthetic ACF data with spectral widths of 50 m/s, 250 m/s, and 500 m/s corresponding to the blue, red, and green lines, respectively. The middle panel shows fitted velocity error versus simulated spectral width for synthetic ACF data with velocities of 50 m/s, 250 m/s, and 500 m/s corresponding to the blue, red, and green lines, respectively. The right panel shows fitted velocity error versus simulated signal to noise ratio for simulated ACF data with a velocity of 250 m/s and a spectral width of 250 m/s.

The slight decrease in error at the largest velocities is consistent with the decreased amount of self-clutter in the simulated ACFs as was noted above. In the middle panel, the velocity error is observed to increase with increasing spectral width as it did in Figure 7.5, plateauing at the Nyquist velocity. In the right panel, the velocity error is observed to decrease as the SNR increases, although it is clear that the SCR is the dominating factor influencing the velocity error when the SNR is small (compare the smooth error curves in the right panel of Figure 7.5 with the self-clutter induced step-wise error curves in Figure 7.6).

For the pathological dataset, Figure 7.6, the fitting error and true error follow similar trends and have similar magnitudes as the ideal fitted dataset, Figure 7.5, but the difference between the fitting error and true error is larger for the pathological dataset for the ideal dataset. This behaviour is expected. Since the MPSE is a maximal estimate of the self-clutter, the MSPE provides an estimate of the largest possible self-clutter contribution. In other words, the MPSE overestimates the self-clutter, which will then result in an over-estimation of the variance of the real and imaginary components of the ACF (*Reimer and Hussey, 2015*). This in turn produces an overestimate of the fitted-parameter error. Even so, Figure 7.6 clearly shows that, in the presence of a pathological amount of self-clutter, FPFM-LMFIT2 is still able to fit ACFs and produce statistically self-consistent and accurate fitting errors.

7.10 Radar Data Analysis

The testing of FPFM-LMFIT2 with synthesized data indicates that it is able to fit SuperDARN ACFs accurately while providing accurate error estimates for the fitted parameters. Fitted parameters (velocity and SNR) obtained using FPFM-LMFIT2 and FITACF from ACFs obtained from the Saskatoon SuperDARN radar are compared. For both the FITACF and the FPFM-LMFIT2 data ground scatter has not been highlighted nor removed.

Figure 7.7 includes range-time-intensity (RTI) plots of the fitted velocity obtained using FITACF (top panel) and FPFM-LMFIT2 (bottom panel) from the Saskatoon SuperDARN radar. The data were recorded on June 14, 2012. For both fitting methods, only data with SNR greater than 0 dB and velocity error less than 100 m/s are included. During the 8 hour

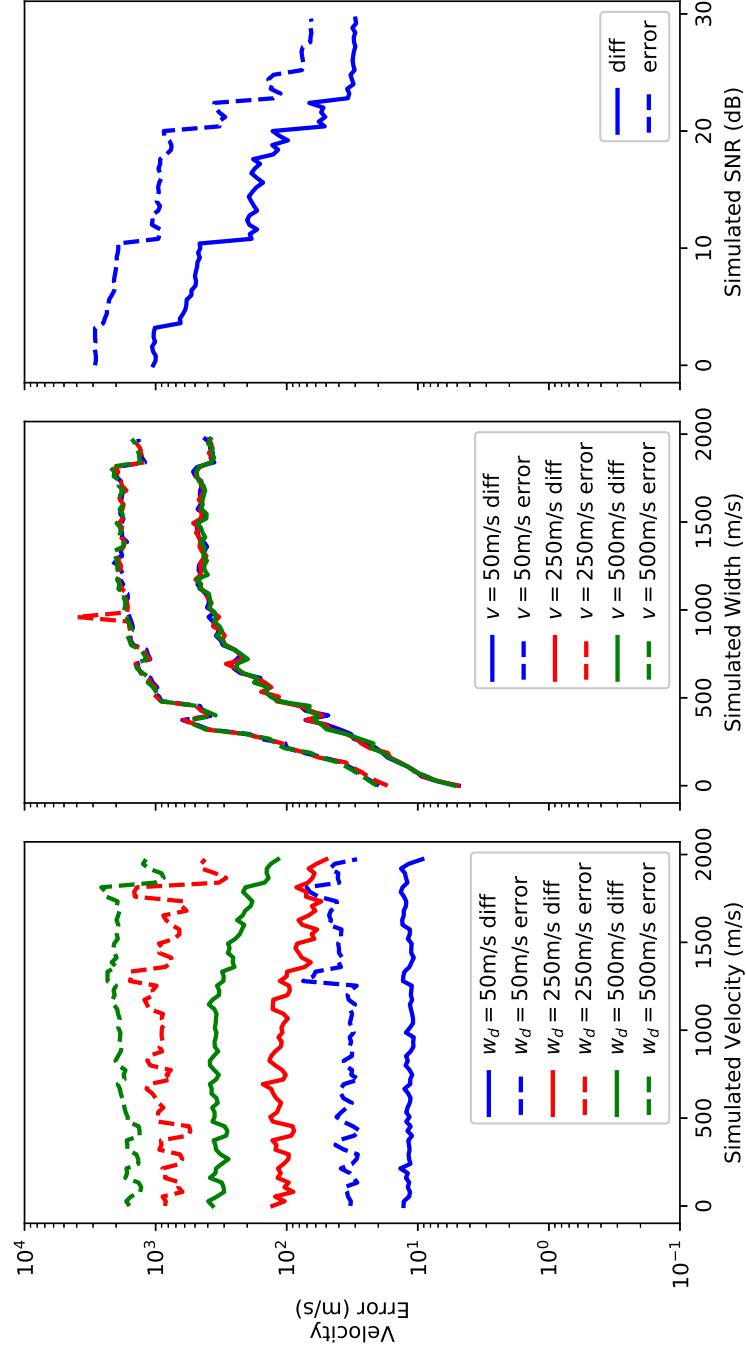


Figure 7.6: The same comparison as plotted in Figure 7.5, but using synthetic data with strong self-clutter (the pathological dataset).

interval, FPFM-LMFIT2 returns 79.8% more data than FITACF. Only a negligible number of FITACF data points were filtered out by the SNR and velocity error filters, so the observed increase in fitted data seen comparing FPFM-LMFIT2 to FITACF is not explained by this filtering. The larger amount of data obtained with FPFM-LMFIT2 is a direct result of the differences in how self-clutter treated and how the ACF lag variances are estimated by FPFM-LMFIT2 and FITACF. A significant amount of acceptable lags with signal are discarded by FITACF using the *ad hoc* self-clutter filtering criterion. FPFM-LMFIT2 does not use this criterion, which was demonstrated to erroneously discard “good” lags (e.g. Figure 7.2). If too many lags are discarded, there are not enough lags to perform a fit, and FITACF does not return any fitted parameters (a failed fit). Instead, FPFM-LMFIT2 addresses lags contaminated with self-clutter by using the MPSE (Equation 7.12), which identifies self-clutter as an increase in the variance (error) of the contaminated lags. Consequently, lags are not discarded, rather they are weighted appropriately, and FPFM-LMFIT2 is able to perform a weighted fit to the ACF to obtain the fitted parameters. As a result, more fitted data are produced by FPFM-LMFIT2 compared to FITACF. In Figure 7.7, only ranges with SNR greater than 0 dB and velocity error less than 100 m/s, but FPFM-LMFIT2 performs fits for all ranges. By changing the SNR and velocity error limits, more or fewer ranges can be included in an RTI.

In general, FPFM-LMFIT2 and FITACF show good agreement in the band of scatter closest in range to the radar. Comparing the SNR at ranges where FPFM-LMFIT2 and FITACF both have data (Figure 7.8), the two methods give similar results, especially at ranges less than 30 and where the SNR is greater than 20 dB. In all Figures 7.7, 7.8, and 7.9, in the FITACF panel before 5:30 UT the second band of scatter is visible, but after 5:30 UT the SNR of the first band increases and the second band almost completely disappears. In the FPFM-LMFIT2 panel in Figure 7.7, two bands of scatter are visible. The first band is approximately centered at range gate 20, and the second band is further from the radar than the first band starting approximately at range gate 30 and shifting to range gate 50 as time progresses. In the FITACF panel, the first band is visible, but the second band has a “speckled” appearance. This is an important distinction, since while interpreting the FITACF data one may erroneously conclude that the second band of data is due to noise

or perhaps the data are patchy due to poor propagation conditions. When interpreting the FPFM-LMFIT2 fitted data one would not be tempted to make such an interpretation. The second band found with FPFM-LMFIT2 extends over a much larger range and has very consistent fitted velocities and fitted SNR. This allows for more structure to be visible in the data than is visible in the FITACF panel.

The elimination of “bad” lags by FITACF is observed to favour the band of scatter at near ranges. The first band of scatter has a larger SNR and larger SCR resulting in fewer “bad” lags for ACFs in the first band of scatter compared to the ACFs in the second band of scatter at larger range with lower SNR and lower SCR. This is because the first band of scatter is producing self-clutter in the second band and the first band with a larger magnitude than the power in the second band. In the second band of scatter, the self-clutter filtering criterion in FITACF is discarding too many lags in the lower power range gates of the second band resulting in an almost a complete loss of the second band of data. The self-clutter criterion used by FITACF (Equation 7.10) is not satisfied when the power in the first band is larger than the power in the second band, so “bad” lags are discarded. It should be noted that the radars often record large power in a close range band of scatter and smaller power in a second larger range band of scatter in SuperDARN data. This situation is particularly susceptible to discarding too many lags when using the *ad hoc* self-clutter filtering criterion, resulting in a loss of otherwise meaningful data.

The fitted velocity errors from FITACF and FPFM-LMFIT2 are plotted in Figure 7.9. In the top panel, the FITACF velocity error is observed to be nearly zero for all data points, which is not consistent with expectations. To first order, the error in the lags of an ACF are a function of the SNR and SCR (recalling Equation 7.7), so to first order the fitted-parameter errors should also vary as a function of the SNR and SCR. Comparing the FITACF panels in Figures 7.8 and 7.9 we see that it would be reasonable for the velocity error to be low where the SNR is very large (> 20 dB), and for the velocity error to be larger in regions where the SNR is much lower (for example, at 10:00 UT), but the SNR trends in Figure 7.8 do not appear in the FITACF velocity errors in Figure 7.9. In contrast the velocity error for FPFM-LMFIT2 is generally lowest where the SNR is largest, and the velocity error increases as the SNR decreases. For example, compare the SNR and velocity error of the blob-like features

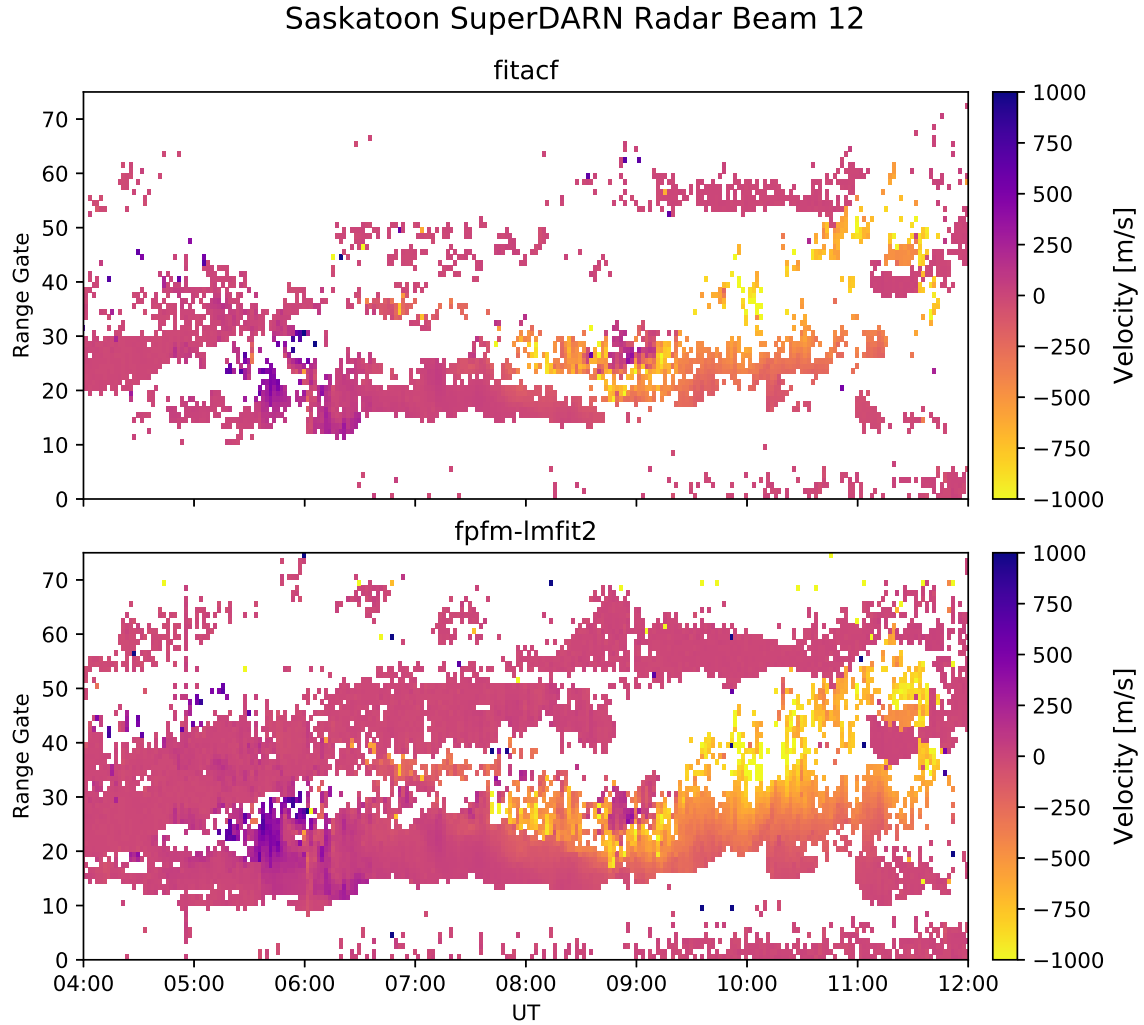


Figure 7.7: A RTI plot comparing fitted velocity data produced by FITACF (top panel) and by FPFM-LMFIT2 (bottom panel) for 8 hours using ACF data from the Saskatoon SuperDARN radar data. The velocity data are plotted with the range away from the radar is plotted along the y-axis and the UT time is plotted in x-axis. The colour represents the magnitude of the fitted velocity. Ground scatter has not been highlighted or removed and is just included in the plot.

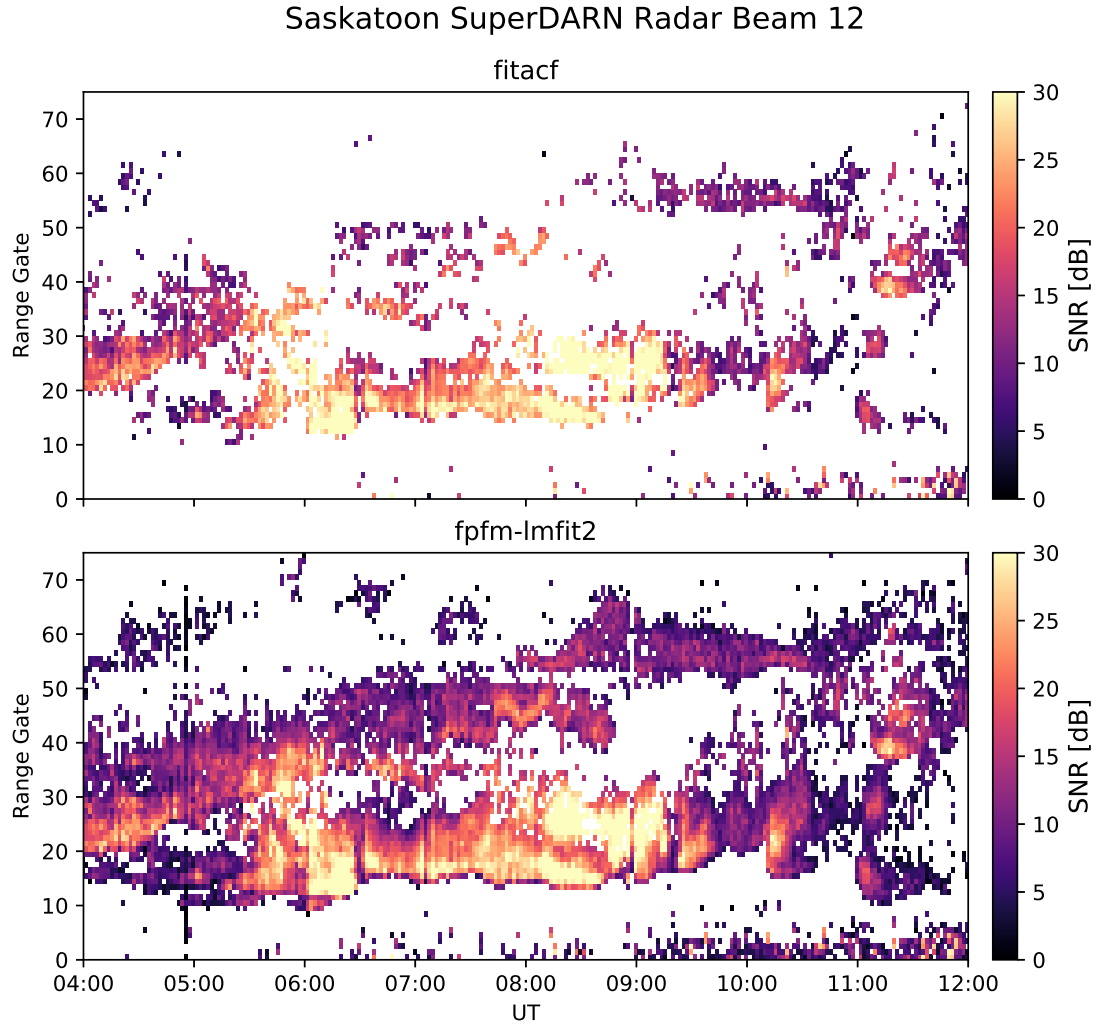


Figure 7.8: A RTI plot comparing fitted SNR data produced by FITACF (top panel) and by FPFM-LMFIT2 (bottom panel) for 8 hours using ACF data from the Saskatoon SuperDARN radar data. The SNR data are plotted with the range away from the radar is plotted along the y-axis and the UT time is plotted in x-axis. The colour represents the magnitude of the fitted SNR. Ground scatter has not been highlighted or removed and is just included in the plot.

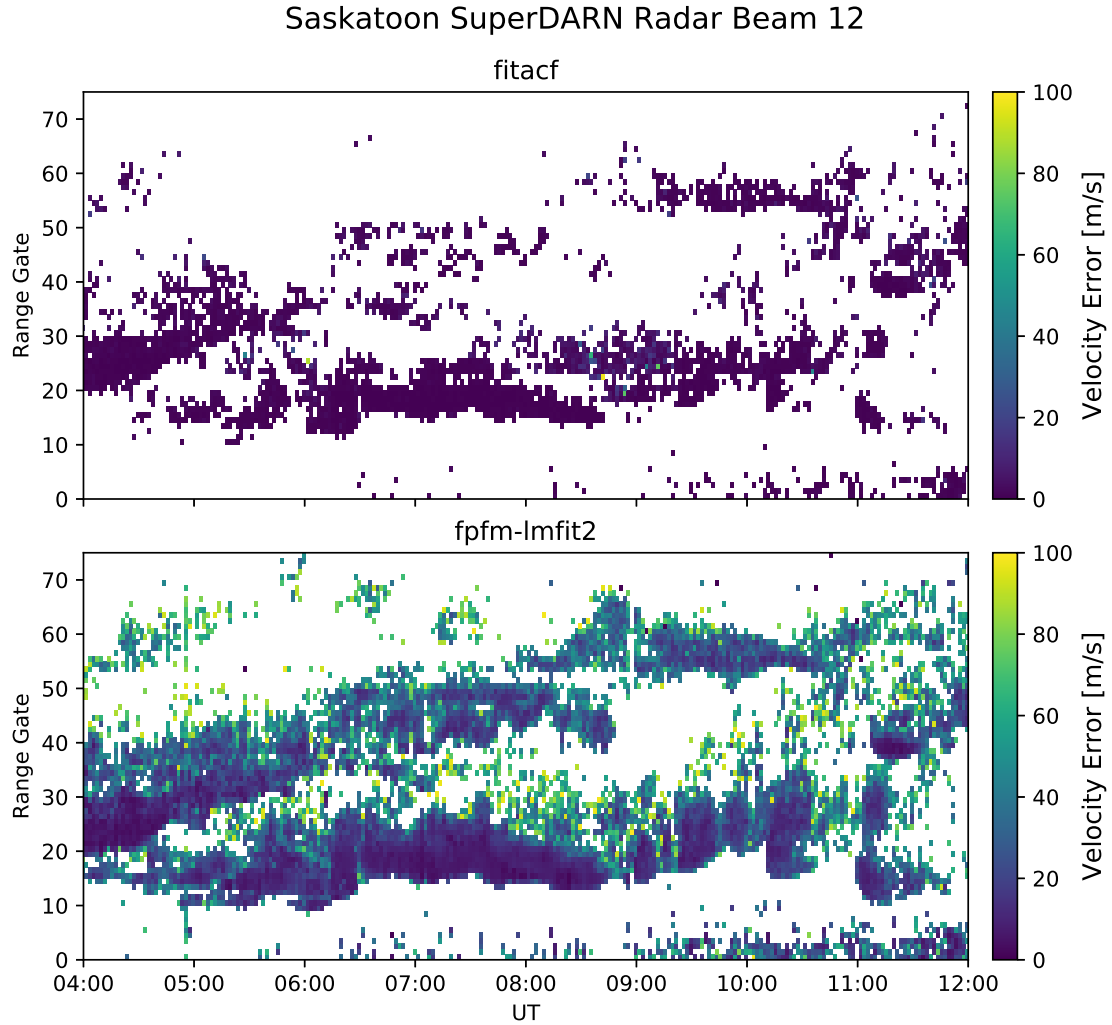


Figure 7.9: A RTI plot comparing fitted velocity errors produced by FITACF (top panel) and by FPFM-LMFIT2 (bottom panel) for 8 hours using ACF data from the Saskatoon SuperDARN radar data. The velocity errors are plotted with the range away from the radar is plotted along the y-axis and the UT time is plotted in x-axis. The colour represents the magnitude of the fitted velocity. Ground scatter has not been highlighted or removed and is just included in the plot.

between 11:00 UT and 11:30 UT. One can also see how the SCR affects the velocity error by comparing the SNR in Figure 7.8 and velocity error in Figure 7.9 for the FPFM-LMFIT2 data, specifically in the region with high SNR in the first band of scatter between ranges 20 and 30 and between 8:00 UT and 9:30 UT. Despite the SNR exceeding 30 dB in this region, the velocity error here is larger than it is in some ranges that have smaller SNR, like those in the first band of scatter at 10:00 UT. This is due to the contribution of self-clutter caused by the large SNR at interfering ranges.

7.10.1 Saskatoon SuperDARN Radar Analysis

For the eight hours of data plotted in Figures 7.7, 7.8, and 7.9, FPFM-LMFIT2 produced 79.8% more data with SNR greater than 0 dB and velocity error less than 100 m/s than FITACF. To determine whether this increase in data is a persistently observed improvement, eight years of Saskatoon SuperDARN radar data were analyzed.

SuperDARN data from July 2006 through August 2014, inclusive, was fitted using FPFM-LMFIT2 and FITACF and for each method, the fitted data was binned into three SNR bins: <3 dB, 3-20 dB, and >20 dB. The fitted data from FPFM-LMFIT2 was compared with fitted data from FITACF and the extra data from FPFM-LMFIT2 was counted. The purpose of this binning process was to determine how much additional data FPFM-LMFIT2 produces compared to FITACF as a function of SNR and velocity error. Figure 7.10 displays the results of statistical comparison where the amount of additional data obtained by FPFM-LMFIT2 compared to FITACF is plotted as a function of velocity error. The blue, green, and red lines represents the amount of additional data in the <3 dB, 3-20 dB, and >20 dB SNR bins, respectively.

From Figure 7.10, one can see that FPFM-LMFIT2 consistently produces more fitted data than FITACF and that a significant amount of the additional data have small velocity error. For example, note the amount of additional data with less than 100 m/s velocity error compared to the additional data with greater than 100 m/s velocity error in Figure 7.10. For the eight years of Saskatoon SuperDARN radar data, FPFM-LMFIT2 obtains more than 10^9 additional data points than FITACF, and of these 477×10^6 have less than 100 m/s velocity error. The statistical comparison finds that for velocity errors less than 100 m/s,

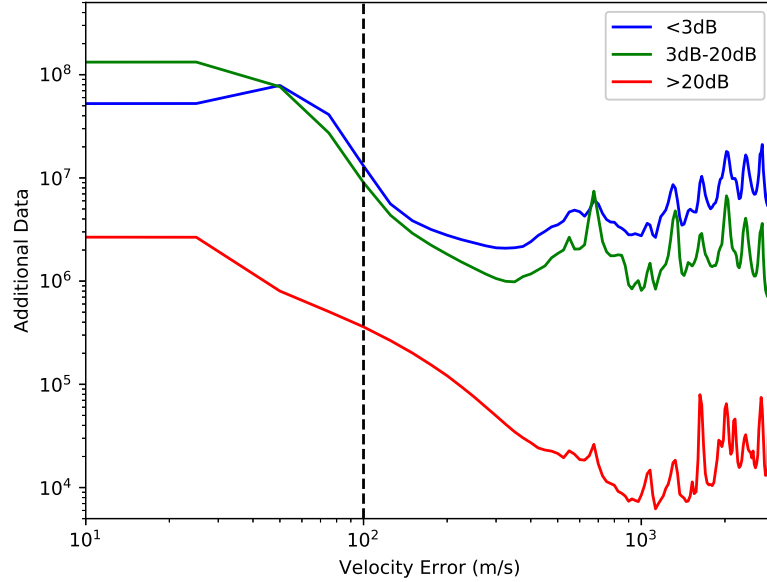


Figure 7.10: Distributions of the additional data produced by FPFM-LMFIT2 compared to FITACF for a statistical comparison of 8 years of Saskatoon SuperDARN data. The additional data is plotted as a function of velocity error and binned in to three SNR ranges. Each line plotted corresponds to a different range of SNR values. The blue, green, and red lines corresponds to a SNR values below 3 dB, from 3 to 20 dB, and above 20 dB.

FPFM-LMFIT2 produces 44.6% more data in the 3-20 dB SNR bin and 7.9% more fitted data in the >20 dB SNR bin than FITACF. Therefore, it may be inferred that for large SNR (>20 dB) FITACF data is of comparable quality to FPFM-LMFIT2 and has a velocity error of less than 100 m/s, but for SNR less than 20 dB, FITACF is discarding lags that contain meaningful information by using its self-clutter filtering criterion.

7.10.2 FITACF Velocity

Although FITACF performs well when fitting ACFs with large SNR, it does occasionally produce erroneous fitted parameters. This is illustrated in Figure 7.11, which compares FITACF (top panel) and FPFM-LMFIT2 (bottom panel) fitted velocities for 20 minutes of data from the Rankin Inlet SuperDARN radar. In several locations, including in proximity to the leading or trailing edges of a large band of scatter, FITACF returns a velocity that is

inconsistent with the adjacent large-scale velocity feature. For example, the FPFM-LMFIT2 velocities at range gate 20 from 8:05 UT to 8:10 UT, are approximately 600 m/s, which is consistent with the velocities in adjacent ranges, but FITACF returns velocities that are in the opposite direction with a magnitude of approximately 200 m/s compared to fitted velocities in adjacent ranges. By examining the ACF lags that are discarded by the self-clutter criterion (not shown), it can be determined that the incorrect velocity returned by FITACF was caused by discarding of lags that were crucial to performing the phase unwrapping. With the few remaining “good” lags, FITACF incorrectly unwrapped the phase, which resulted in a fitted velocity with the incorrect sign and an unrealistic magnitude. The FITACF fitted velocity errors for the incorrectly fitted ranges are small and do not provide any indication that these velocities are erroneous.

The issue of producing incorrect fitted velocities has a significant impact on the interpretation of SuperDARN velocity data. If one does not realize that these velocities are incorrect, one may erroneously interpret the FITACF data as having several interesting shears in the velocity flow. Since FITACF produces fitted velocity errors for the erroneously fitted velocities that are the same as for ranges that are correctly fitted, like those in Figure 7.9, one cannot use the FITACF velocity error to filter erroneously fitted velocities. Instead the ACF data needs to be examined to determine if the FITACF velocities are valid. For large datasets this is impractical. This demonstrates the importance of having accurate velocity errors, like those produced by FPFM-LMFIT2.

7.11 Summary

SuperDARN ACF fitting software obtains best-fit values for power, velocity, and spectral width by applying error-weighted least-squares regression techniques that require accurate knowledge of the statistical characteristics of the data. Current SuperDARN fitting routines rely on three *ad hoc* criteria: 1) the *ad hoc* variance expression given by Equation 7.4; 2) an invalid magnitude model to fit the magnitude of the ACF, requiring an *ad hoc* condition to discard lags of the ACF when the power in the lags has decreased below an *ad hoc* criterion; 3) a self-clutter lag discarding criterion. When the SNR in a range gate is large, FITACF

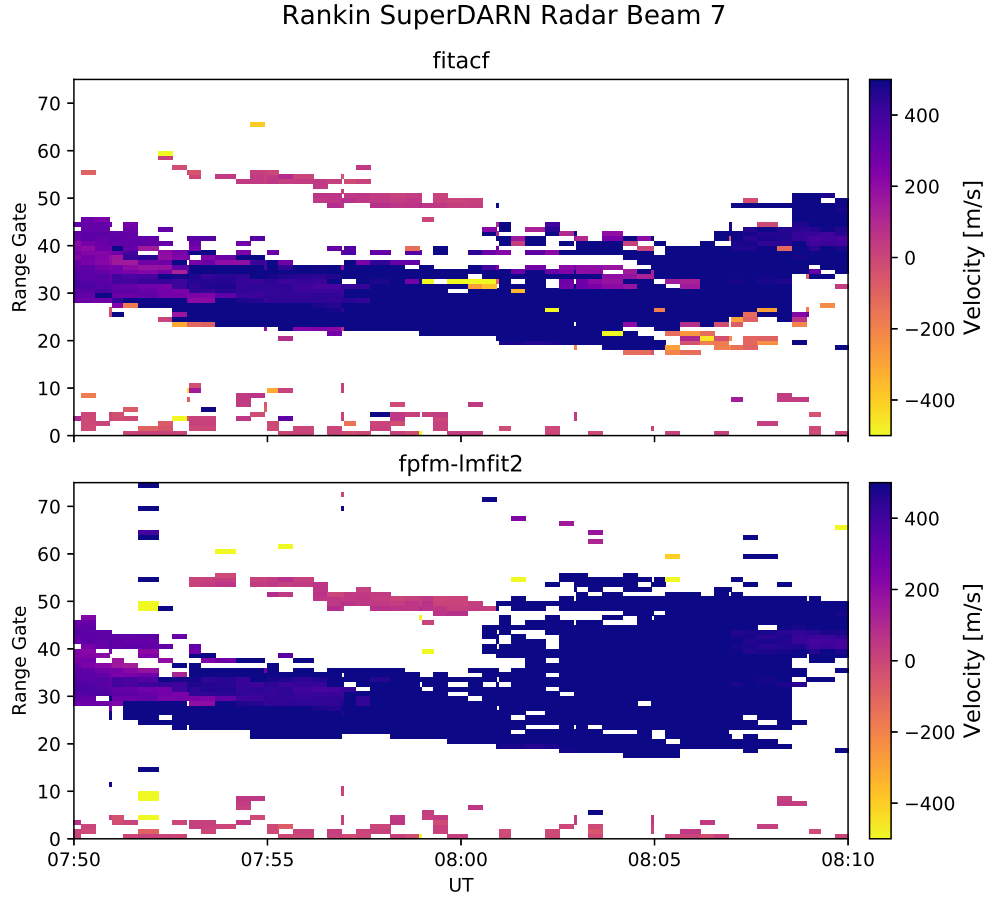


Figure 7.11: A RTI plot of the fitted velocity obtained using FITACF (top panel) and FPFM-LMFIT2 (bottom panel) for 20 minutes of ACF data from the Rankin Inlet SuperDARN radar, on February 22, 2014.

tends to produce realistic fitted values, but as shown in Figure 7.11, FITACF sometimes fails to properly fit SuperDARN data. Note that these *ad hoc* criteria were designed to be conservative, eliminating lags rather than risk self-clutter contamination during the fitting process. As discussed in Section 7.10.1, for the most part SuperDARN fitted data are good, but, with more widely available computational resources, modifications to the fitting procedures can now be made to improve the fitted data and provide accurate estimates of the fitting error.

A First-Principles Fitting Methodology (FPFM) was presented to address these issues and improve the fitting procedures. The FPFM was designed for fitting SuperDARN ACFs without the use of any *ad hoc* filtering criteria and without *ad hoc* variance estimates. Using

first-principles methods based on the statistical properties of the voltage signals received by SuperDARN radars, *Reimer et al.* (2016) derived the statistical characteristics of ACF lag estimates. The results of that study include exact analytic expressions for the variance of the real and imaginary components of the ACF lag estimates (Equations 7.5, 7.6, 7.8, and 7.9). To address the contribution of self-clutter to the ACF variance, *Reimer and Hussey* (2015) derived the MPSE (Equation 7.12), which can be used to estimate the upper limit of the self-clutter power in a lag of the ACF. These previous works have been combined with a fitting methodology to produce an improved operational ACF fitting methodology (the FPFM) for SuperDARN data processing. To avoid the challenges associated with fitting the magnitude and phase of the ACF (as detailed in Section 7.7.1), the FPFM performs a fit to the real and imaginary components of the ACF. The FPFM then uses Equations 7.5, 7.6, 7.8, 7.9, and 7.12 to obtain the variance of the real and imaginary components of the ACF and uses the Levenburg-Marquardt algorithm to perform a non-linear error-weighted least-squares fit of Equation 7.3 to the ACF. The FPFM has been implemented in C and in Python as the LMFIT2 fitting software. The LMFIT2 open source code is available on Github (<https://github.com/asreimer/lmfit2>).

The main result of this work is that by using the FPFM-LMFIT2, one is capable of producing statistically self-consistent and accurate measures of uncertainty for fitted parameters in all SuperDARN range gates. Tests of FPFM-LMFIT2 software using synthetic data showed that the errors in fitted data products from FPFM-LMFIT2 vary as a function of SNR, SCR, and ACF correlation time, as expected. Specifically, when the SNR, SCR, and correlation time are large the fitting errors are small. When the SNR, SCR, or the correlation time are small, the fitting errors are large. These results were illustrated in Figures 7.5 and 7.6. Additionally, all historical SuperDARN data can be reprocessed using FPFM-LMFIT2.

Another substantial advantage of FPFM-LMFIT2 is the significant increase in the amount of fitted data produced compared to the amount produced by the SuperDARN FITACF software. For example, in Figures 7.7, 7.8, and 7.9, FPFM-LMFIT2 produced 79.8% more data than FITACF with SNR greater than 0 dB and velocity errors less than 100 m/s. When comparing FPFM-LMFIT2 and FITACF statistically using eight years of Saskatoon SuperDARN radar data, FPFM-LMFIT2 produced more than 10^9 more fitted data than

FITACF. Of these additional FPFM-LMFIT2 fitted data points, 477×10^6 ranges had a velocity error less than 100 m/s, resulting in 52.5% additional fitted data points when the SNR is greater than 3 dB and 7.9% additional fitted data points when the SNR is greater than 20 dB. From this comparison it can be concluded that FITACF produces good data as long as the SNR is large (>20 dB).

FPFM-LMFIT2 may be used to process data from any other radar system that extracts fitted parameters from an ACF. FPFM-LMFIT2 may also be utilized with other fitting techniques. For example, instead of using least-squares regression, the Bayesian Inference method (*Spaleta et al.*, 2015) could be used in conjunction with Equations 7.5, 7.6, 7.8, 7.9, and 7.12.

It should be noted that FPFM-LMFIT2 does not include contributions to the ACF lag variance due to off-diagonal terms in the covariance matrix of the ACF. These terms are non-zero, since some voltage samples are reused to estimate more than one lag of the ACF (e.g. *Farley*, 1969). This means that some of the lags of the ACF are correlated with other lags, because the same voltage sample being used to estimate both of them. Future work is required to determine the significance of these off-diagonal terms and how to quantitatively estimate their contributions to the ACF lag variance.

FPFM-LMFIT2 is compatible with all SuperDARN data collected to date, so a re-analysis of all SuperDARN data products, such as velocity vectors in convection maps and the cross polar cap potential, could be performed using fitted data obtained using the FPFM-LMFIT2. Future studies could examine how the increase in the amount of data from FPFM-LMFIT2 affects SuperDARN-derived maps of F region ionospheric plasma convection. The fitted velocity error for each velocity measurement should be included in the global convection map fitting procedures when performing an error-weighted fitting of the statistical convection model to the input velocity measurements (e.g. *Ruohoniemi and Baker*, 1998b; *Fiori et al.*, 2010b; *Cousins et al.*, 2013). Such a study could produce convection maps that include confidence intervals for the output convection velocity vectors and the cross polar cap potential. In summary using the FPFM-LMFIT2 provides accurate confidence intervals to be determined for all SuperDARN data products.

7.12 Acknowledgments

This study was supported by funding from Canada Foundation for Innovation (CFI) Major Science Initiatives Program, the Canadian Space Agency's (CSA) Geospace Observatory (GO Canada) contribution agreement, Natural Sciences and Engineering Research Council (NSERC) grant CGSD3-425247-2012, and NSERC Discovery Grant 194324-2013. The authors acknowledge the use of SuperDARN data. SuperDARN is a collection of radars funded by national scientific funding agencies of Australia, Canada, China, France, Japan, South Africa, United Kingdom, and United States of America. SuperDARN data used in this study is available from the corresponding author upon request. SuperDARN Canada is supported by the Canadian Foundation for Innovation, the Canadian Space Agency, and the Province of Saskatchewan.

CHAPTER 8

CONCLUSIONS AND FUTURE APPLICATIONS

The main objective of this thesis was to develop accurate and statistically self-consistent error bars for SuperDARN fitted power, spectral width, and velocity. The objective has been accomplished by developing three improvements for the signal processing workflow: a novel method for estimating self-clutter, determining the statistical properties of SuperDARN ACF lag estimates, and a first-principles fitting methodology that improves upon existing SuperDARN fitting techniques. The fitting methodology relies on the variance of the ACF lag estimates, which was determined from the PDF of the lag estimates. The variance of each lag estimate requires quantification of the self-clutter contributing to that lag, which is provided using the novel self-clutter estimator that was developed. The first-principles fitting methodology provides accurate and statistically self-consistent errors bars for the fitted power, spectral width, and velocity obtained from fitting SuperDARN ACFs. Additionally, the methodology provides fitted data for every range such that the fitting errors can be used to filter undesirable data. In Section 8.1, summaries of the results and applications of Chapters 5, 6, and 7 are provided. The remainder of this chapter, Section 8.2, presents and discusses new research already in progress directly resulted from developing accurate SuperDARN fitted parameter errors.

8.1 Summary of Contributions

The three preceding chapters represent a significant contribution to the SuperDARN research community. No previous published work exists that attempts to quantify the self-clutter caused by the multiple-pulse technique. The contribution of self-clutter to the variance of ACF lag estimates had never been considered when processing SuperDARN data. Using the

power-based self-clutter estimator developed in Chapter 5, it is now possible to estimate the upper limit of the contribution of self-clutter for the entire archive of SuperDARN data. The new self-clutter estimator enables the reprocessing of all 25 years of SuperDARN data. Now the self-clutter is properly accounted for in the variance of the ACF lag estimates. Additionally, previous authors, such as *Farley* (1969), have explored the statistical properties of ACFs for ISRs, but the literature lacked any such explorations for CSRs, such as SuperDARN and its ACFs. The statistical properties of the SuperDARN ACF lag estimates were determined and presented for the first time in Chapter 6. This is a significant contribution because exact expressions for the variance of the lag estimates, which can be used in error-weighted fitting techniques, were also developed. In contrast, current SuperDARN processing uses ad hoc variance estimates that were found to result in inaccurate fitted parameter errors. Finally, a new fitting methodology was developed in Chapter 7, which was based on the first-principles research performed in Chapters 5 and 6. The new fitting methodology is also a significant contribution because it provides accurate and statistically self-consistent fitted parameter errors for SuperDARN data. Additionally, the fitting methodology enables much more data to be extracted from the SuperDARN ACFs than was previously possible.

What follows are detailed descriptions of the results of the research performed in Chapters 5, 6, and 7. A discussion of the applicability and the limitations of the results is also provided.

8.1.1 Self-Clutter Estimation

The research performed in Chapter 5 involved developing a technique to estimate the self-clutter that is present in SuperDARN ACFs caused by using the multiple-pulse technique of *Farley* (1972). As discussed in Section 5.1, self-clutter estimates are needed to accurately estimate the variance of ACF lag estimates. If the variance is not accurately estimated, the fitted parameter errors will be incorrect and not useful. Three estimators were developed for use in SuperDARN data processing: a voltage-based self-clutter estimator (VSE), a maximal voltage-based self-clutter estimator (MVSE), and a maximal power-based self-clutter estimator (MPSE). The mean-square-error of each of the self-clutter estimators was also derived. The estimators were tested and validated using a modified, by this author, version of the

radar data simulator of *Ribeiro et al.* (2013b).

The VSE, given by Equation 5.12, is only valid for a limited range of ionospheric parameters and cannot be used generally. However, it is possible to use the VSE to subtract self-clutter produced by ground scatter or slow moving long lived ionospheric irregularities from the SuperDARN ACF, as was demonstrated in Figure 5.6. The maximal estimators, MVSE and MPSE, which are given by Equations 5.14 and 5.15, are generally applicable, however they only provide an upper limit or “worst-case” estimate of the self-clutter. The MVSE is a more efficient estimator than the MPSE; however, it requires voltage samples to be used and voltage samples are not typically recorded and stored for later use by SuperDARN radars¹. Therefore, the MPSE, which uses the power estimates routinely recorded and stored by all SuperDARN radars is the most practical estimator to use with SuperDARN. Lastly, it should be pointed out that all three of the self-clutter estimators can be used with any radar system that uses the multiple-pulse technique.

Future improvement to the self-clutter estimators could involve developing a more accurate estimator for SuperDARN. It may be possible to develop an improved estimator that is based on the fitted parameters obtained at all ranges by utilizing an iterative fitting technique. Recall that the self-clutter contributions to a lag of the ACF depend on what signals are scattering from different ranges. Also recall that the voltage samples received from a range can be simulated using the fitted parameters obtained for that range. Using these two facts, it may be possible to obtain a more accurate estimate of self-clutter using fitted ionospheric parameters to simulate the self-clutter that is expected to result from ranges with those same ionospheric parameters. An iterative algorithm could take the form:

1. estimate the self-clutter for all lags at all ranges using the power-based self-clutter estimator,
2. fit all ranges using the First-Principles Fitting Methodology,
3. simulate the voltage samples that would produce ACFs consistent with the fitted pa-

¹At least, not at the time of this writing. Members of the SuperDARN community have argued that too much storage space is required to store the voltage samples. Although storage was an issue in the past, my investigations have revealed that iqdat files are often less than ten times larger than rawacf files and that several years of compressed iqdat files can fit on a 1TB hard drive.

rameters obtained in step 2,

4. use the simulated voltage samples to estimate the self-clutter for all lags at all ranges using the voltage-based self-clutter estimator, and
5. fit all ranges using the First-Principles Fitting Methodology and repeat the process from step 3 onward until a convergence in fitted parameters has been reached.

In step 4, the upper limit of the self-clutter is still being estimated, however this estimate will be in general better than that given by the MPSE since the simulated ionospheric-origin self-clutter will average away as discussed in Chapter 5. Such an algorithm could provide a more accurate estimate of self-clutter that would provide a more accurate estimate of the variance of ACF lags, which would ultimately improve the accuracy of the fitted ionospheric parameter errors.

8.1.2 SuperDARN Auto-Correlation Function Statistics

In Chapter 6, the statistical properties of ACF lag estimates were presented and discussed for the first time. These properties were derived by propagating the statistical properties of SuperDARN voltage samples through the ACF lag estimator given by Equation 6.1. Following the arguments presented by other authors, such as *André et al.* (1999); *Moorcroft* (2004); *Ribeiro et al.* (2013b), SuperDARN voltage signals were modelled as the sum of a large number of randomly spatially distributed ionospheric signals. It was then demonstrated that pairs of SuperDARN voltage samples are characterized by a 4-dimensional Gaussian distribution, a result that was experimentally verified (see Figure 6.2).

Next, the PDF of the real and imaginary components of the ACF lag estimates was derived from the voltage samples and found to be characterized by a 2-dimensional Gaussian distribution given by Equation 6.13. The PDFs of the magnitude and phase of the ACF lag estimates (Equations C.3 and C.13) were also derived and found to be significantly more complicated than the PDFs of the real and imaginary components. The variances of all components of ACF lag estimates were also derived. For the real and imaginary components, exact analytic expressions were obtained and are given by Equations 6.10 and 6.11. For the magnitude and phase, it was not possible to obtain analytic expressions for the variance;

however, an algorithm for calculating the variances numerically was developed. The PDFs and variances of the ACF lag estimates were verified numerically using Monte Carlo methods. It was determined that the PDF of the real and imaginary components of the ACF lag estimates is applicable so long as the number of pulse-sequences used to estimate the lags is ≥ 10 . A previously derived asymptotic expression for the variance of the phase of ACF lag estimates, derived by *Woodman and Hagfors* (1969), was compared and validated against the numerically derived phase variance. The variance of the real and imaginary components was also compared with experimentally obtained SuperDARN ACFs (see Figure 6.6) and found to be consistent.

Two statistical properties of the magnitude and phase relevant to current SuperDARN fitting procedures were determined. The first statistical property found was that the PDF of the magnitude of ACF lag estimates becomes a Rayleigh distribution when the correlation in the ACF approaches zero. This is a significant property since a Rayleigh distribution is identically equal to zero at zero power, meaning that the magnitude of the ACF lags does not approach zero when the correlation in the ACF approaches zero. Instead, it was found that the magnitude approaches $(\pi/2)^2$ as the correlation approaches zero. This is in direct conflict with the exponential model routinely used to describe and fit the magnitude of the ACF lags in the current SuperDARN fitting methodology, since the exponential model approaches zero as the ACF correlation approaches zero. The second statistical property found was related to the PDF of the phase of ACF lag estimates. As the correlation approaches zero, the PDF of the phase becomes a uniform distribution. This means that in low SNR conditions or strong self-clutter conditions, the statistical properties of the phase are fundamentally incompatible with unambiguous phase unwrapping. Unambiguous phase unwrapping is required to perform a linear fit to the phase and this cannot be reliably done in the currently used SuperDARN fitting methodology.

Finally, the results obtained in Chapter 6 are applicable to any pulse-Doppler radar system that estimates an ACF and then fits it to extract target parameters. It is also important to note that the real and imaginary variances were derived directly from Equation 6.9 and are therefore exact results that only depend on the PDF of the voltage samples being Gaussian and are not subject to the same criteria as the ACF lag PDF (Equation 6.13). Combining the

MPSE from Chapter 5 with the variance expressions for the real and imaginary components of the ACF lags from Chapter 6 enables statistically self-consistent error-weighted fitting of SuperDARN data to be performed for the first time.

The variance expressions developed in Chapter 6 do not include contributions to the variance from *error correlations* as discussed in *Farley* (1969). These error correlations are a consequence of reusing the same voltage samples to estimate more than one lag of the ACF. This results in correlations in the variance of the ACF lags and produces “smoother” looking ACFs than would otherwise occur. Future work should be performed to quantify these correlations as was done for ISRs by *Farley* (1969). Further future work could involve deriving approximate analytic expressions for the variance of the magnitude and phase based on spline fitting the numerically derived variances. The approximate expressions would be computationally faster, which would be beneficial to fitting algorithms.

8.1.3 Self-Consistent Fitted Parameter Errors for SuperDARN

Utilizing the results obtained in previous chapters, a new first-principles fitting methodology (FPFM) was developed in Chapter 7 that improves upon current SuperDARN fitting methodologies. The research performed in this chapter utilized the results from all previous chapters to achieve the objective of this thesis. Using the FPFM, it is now possible to process SuperDARN data and obtain accurate and statistically self-consistent fitted parameter errors. The FPFM was implemented in the C and Python programming languages as a software package called LMFIT2. The LMFIT2 source code is open source and is available on Github (<https://github.com/asreimer/lmfit2>).

The current standard SuperDARN fitting software uses error-weighted linear least squares algorithms to perform fits to the magnitude and phase of the ACF lag estimates to extract power, spectral width, and velocity of ionospheric irregularities. As mentioned in Section 5.1 and discussed in detail in Chapter 7, fitting the magnitude and phase requires overcoming significant challenges, some of which are fundamentally incompatible with the statistical properties of the magnitude and phase.

To avoid these challenges, the FPFM was instead developed to perform error-weighted non-linear least-squares fits to the real and imaginary components of the ACF lag estimates

using the Levenburg-Marquardt algorithm. To provide accurate weights, both the MPSE developed in Chapter 5 and the variance expressions for the real and imaginary components of the lag estimates developed in Chapter 6 are used. If either an incorrect model or incorrect variance is used to fit the data, the resulting fitted parameters and fitted parameter errors will be at best inaccurate, or in the worst case produce “good looking” data that is actually meaningless.

The FPFM was tested significantly using synthetic SuperDARN ACF data produced using the SuperDARN data simulator of *Ribeiro et al.* (2013b), where known ionospheric parameter values were used to seed the simulator and then the fitted parameters obtained using the FPFM were compared with the true (expected) parameter values. Two sets of synthetic ACF data were used, ideal and pathological, containing no self-clutter and the maximum possible self-clutter, respectively. The FPFM was found to be able to accurately fit both synthetic datasets while simultaneously producing accurate fitted parameter errors (see Figures 7.5 and 7.6).

The FPFM was also used to process real measured SuperDARN ACFs. SuperDARN data processed using FPFM was compared with data processed using the current SuperDARN fitting methodology, FITACF. By comparing the fitted data products, it was found that the FPFM produces a significant increase in the amount of fitted data when compared to the amount of fitted data produced by FITACF. In a statistical study comparing 8 years of Saskatoon SuperDARN data, the FPFM produced 10^9 more fitted data than FITACF. Of these additional FPFM data, 477×10^6 ranges had a velocity error less than 100 m/s, resulting in FPFM producing an average of 52.5% additional fitted data points compared to FITACF when the SNR > 3 dB. When SNR > 20 dB, an additional 7.9% fitted data points were obtained. In summary, not only does the FPFM produce accurate fitted parameter errors, but it also produces significantly more fitted data than the currently standard SuperDARN fitting methodology. Clearly, the objective of this thesis has been accomplished and exceeded.

The following is a list interesting future research topics:

- Reprocessing the entire historical SuperDARN archive so that all fitted SuperDARN data contains accurate fitted parameter errors.

- Investigating the effects of including fitted parameter errors in deriving convection maps. It is anticipated that the parameter errors can be used to filter and weight the data input to the convention mapping software such that more accurate convection maps can be derived.
- Investigate a potential systematic underestimation of fitted velocities from the FITACF fitting software. As was reported by *Ribeiro et al.* (2013a), when comparing FPFM with FITACF, it was found that the phase unwrapping algorithm in FITACF is unable to reliably unwrap the phase of the ACF in the presence of strong self-clutter for velocities $\gtrsim 1000$ m/s, resulting in systematically lower fitted velocities on average than expected (see orange dots in Figure 7.3). Such a systematic bias could have significant consequences for derived data products, such as convection maps and the cross polar cap potential.
- Elevation angle processing is not implemented in the LMFIT2 software, however I am working on implementing this as part of future research.

8.2 Progress on Ensuing Work

Several other related and important projects organically resulted from this thesis research. This dissertation will conclude with a brief description of three such projects², including some preliminary research results. All of the projects are fundamentally related to inferring electron density from SuperDARN measurements as a standard data product.

The first project describes a methodology for designing an experiment mode to run on a SuperDARN radar such that desired range and Doppler resolutions are achieved while simultaneously obtaining a desired fitted parameter error given an expected set of ionospheric

²I actually have more that I would like to include; but, for the sake of brevity, will just list them here: While working with the voltage sample data while developing the self-clutter estimators, I noticed that some sets of voltage samples sometimes contained strong interference from HAM radio operators. This interference should be filtered by discarding the “bad” voltage samples; however, in the current processing software, when one ACF (≈ 100 ms of data) is “poisoned” by strong intermittent interference the entire mean ACF (≈ 3 s of data) is discarded. Finally, an alternative fitting methodology that fits ACFs from each individual pulse sequence simultaneously instead of the mean ACF could be investigated. The averaging process destroys information, and fitting each individual ACF retains that information, resulting in a better fit.

conditions.

The second project describes a joint experiment between the Clyde River SuperDARN radar in Clyde River, Nunavut and the Resolute Bay Incoherent Scatter Radar (RISR-C) in Resolute Bay, Nunavut that was performed to compare the inferred SuperDARN electron density with the absolute electron density measurements obtained from RISR-C.

The third and final project describes a new and unique variation of the multiple-pulse sequence that suppresses self-clutter from correlated interfering signals such as ground scatter.

8.2.1 Optimal SuperDARN Pulse Sequence Investigations

The goal of this investigation was to design a pulse sequence that is optimal for standard operation on all SuperDARN radars. Using an optimal pulse-sequence ensures that the SuperDARN radars are collecting the highest quality data possible. For the purposes of this investigation, the pulse width will be assumed to be a fixed value. The conditions determining the optimal pulse sequence are fundamentally fixed by the characteristics of the ionospheric irregularities that are being measured. For example, the optimum pulse-sequence must be able to resolve the range of expected ionospheric velocities. The maximum unambiguous Doppler (MUD) determines the range of resolvable velocities and for the multiple-pulse technique the MUD is determined by the transmitted wavelength λ and the fundamental interpulse period (or pulse-repetition time (PRT)) τ , such that

$$\text{MUD} = \pm \frac{\lambda}{4\tau}. \quad (8.1)$$

Another metric for determining the optimal pulse-sequence is to determine which pulse-sequence produces the lowest fitted parameter errors. Fundamentally, the fitting errors depend on the variance of the ACF lag estimates (recall Chapter 6). To first order, the ACF standard deviation, which will be referred to as the ACF errors, depends on the SNR, the SCR, and the number of pulse-sequences transmitted (recall Equation 5.1, p. 81). This can be expressed as

$$dS/S = (1 + 1/\text{SNR} + 1/\text{SCR})/\sqrt{K}, \quad (8.2)$$

where dS/S is the first-order approximation of the variance of the ACF lag estimate divided by the signal power S . The SNR depends on propagation conditions and the presence of

ionospheric irregularities and does not depend on the pulse sequence transmitted (the pulse width controls the SNR too, but modifying it also modifies the range resolution). The SCR is the only quantity contributing to the ACF variance that depends on the pulse-sequence transmitted. Therefore, the self-clutter in SuperDARN ACF lag estimates depend on which pulse-sequence was transmitted. Recall from Chapter 5 that the ranges where self-clutter originates depend on pulse spacing, τ . The larger the spacing between the pulses, the further away the interfering ranges are from one another. Due to the multiple hopping nature of HF wave propagation in the ionosphere, it is common for SuperDARN radars to measure bands of scatter. This means that there are ranges between the bands from which no signal is typically received, so if the pulse spacings could be increased sufficiently to move the interfering ranges into the “gaps” between the bands of scatter, the SCR can be increased. The upper limit on pulse spacing is determined by the MUD. This means that an optimal pulse sequence is one where the MUD is set only as high as it is needed so that the pulse-spacing can be as large as possible to reduce the self-clutter and increase the SCR. Fundamentally, designing the optimal pulse sequence is an exercise in finding the optimal value for the PRT, τ .

SuperDARN radars are designed to operate in the HF band with frequencies between 8 and 20 MHz. Additionally, the PRTs used on SuperDARN radars typically range from 1.8 ms up to 2.4 ms. Figure 8.1 visualizes the relationship between the MUD, PRT, and transmitted frequency. The figure shows how, for a given MUD, the PRT needs to be decreased as the transmitted frequency increases. Also note that increasing the PRT or decreasing the transmitted frequency increases the MUD. The hatched area shows the range of values for the PRT and transmitted frequency for which SuperDARN radars typically operate.

To determine the upper bound for the PRT, the distribution of velocities measured by every SuperDARN radar for 21 years (1992–2013, inclusive) were compiled and combined into groups according to the latitudinal region in which the radar is located. Figures 8.2, 8.3, and 8.4 show the resulting velocity distributions for the polar cap, high-latitude, and mid-latitude regions, respectively. The ground scatter identification method of *Blanchard et al.* (2009) was used to filter out ground scatter echoes from the velocity distributions, which results in the double peak structure observed in the velocity distributions centred at zero velocity.

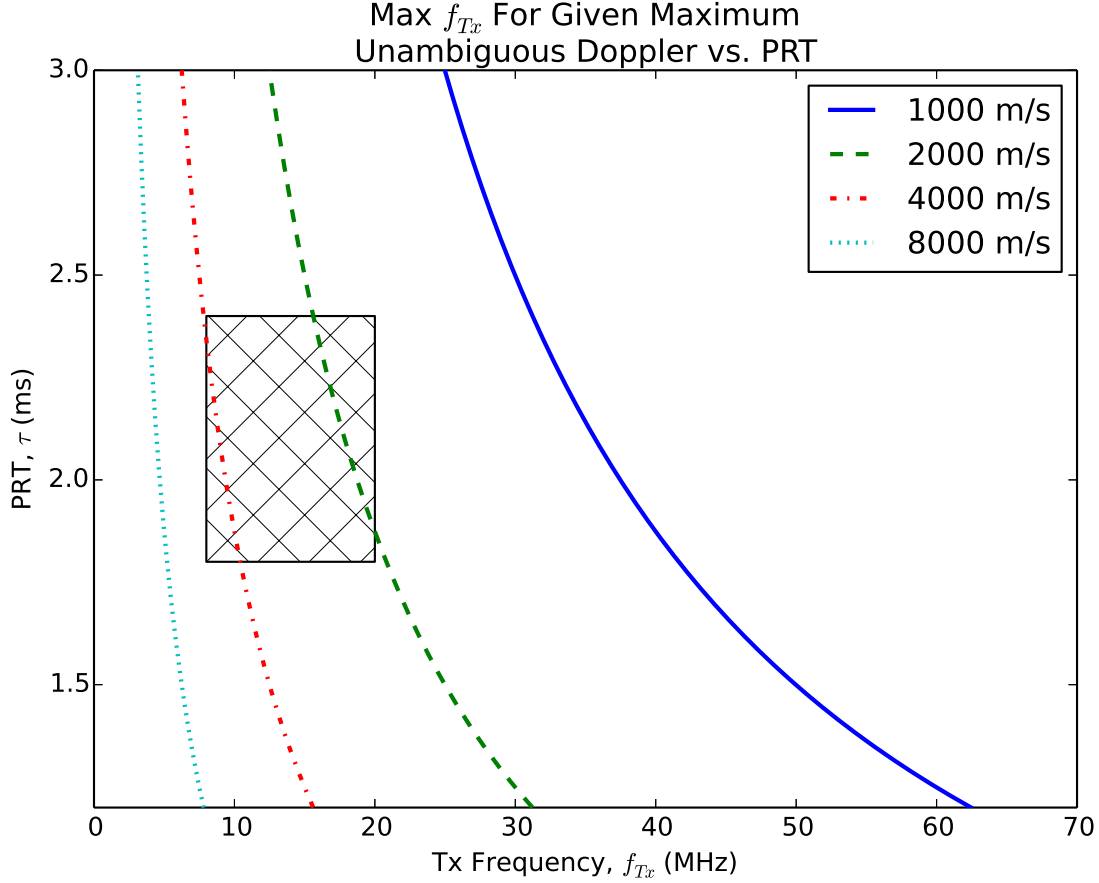


Figure 8.1: A contour plot of Equation 8.1. Choosing combinations of τ and transmission frequency (f_{Tx}) to the left side of a contour of MUD guarantee a $\text{MUD} \geq$ the MUD of the contour. The hatched area shows the region in which SuperDARN radars currently operate: $1.8 \text{ ms} \leq \tau \leq 2.4 \text{ ms}$ and $8 \text{ MHz} \leq f_{Tx} \leq 20 \text{ MHz}$.

Since the main objective of SuperDARN radars is to measure plasma velocity, one must be careful to ensure that the optimal choice of τ results in a MUD equal to or larger than the largest plasma velocity flows in the ionosphere. This means that determining the largest required MUD for SuperDARN is the same as determining the largest velocities that may be observed in the ionosphere by SuperDARN.

Both a Laplace distribution and a Gaussian distribution were fitted to each velocity distribution in an attempt to extract the standard deviation of the velocity data. Pearson's goodness of fit parameter for each fitted distribution is displayed as the χ^2 value in the legend for each plot. Notice that the polar cap SuperDARN radar velocity measurements (Figure 8.2) follow a Gaussian velocity distribution, whereas the high-latitude radar

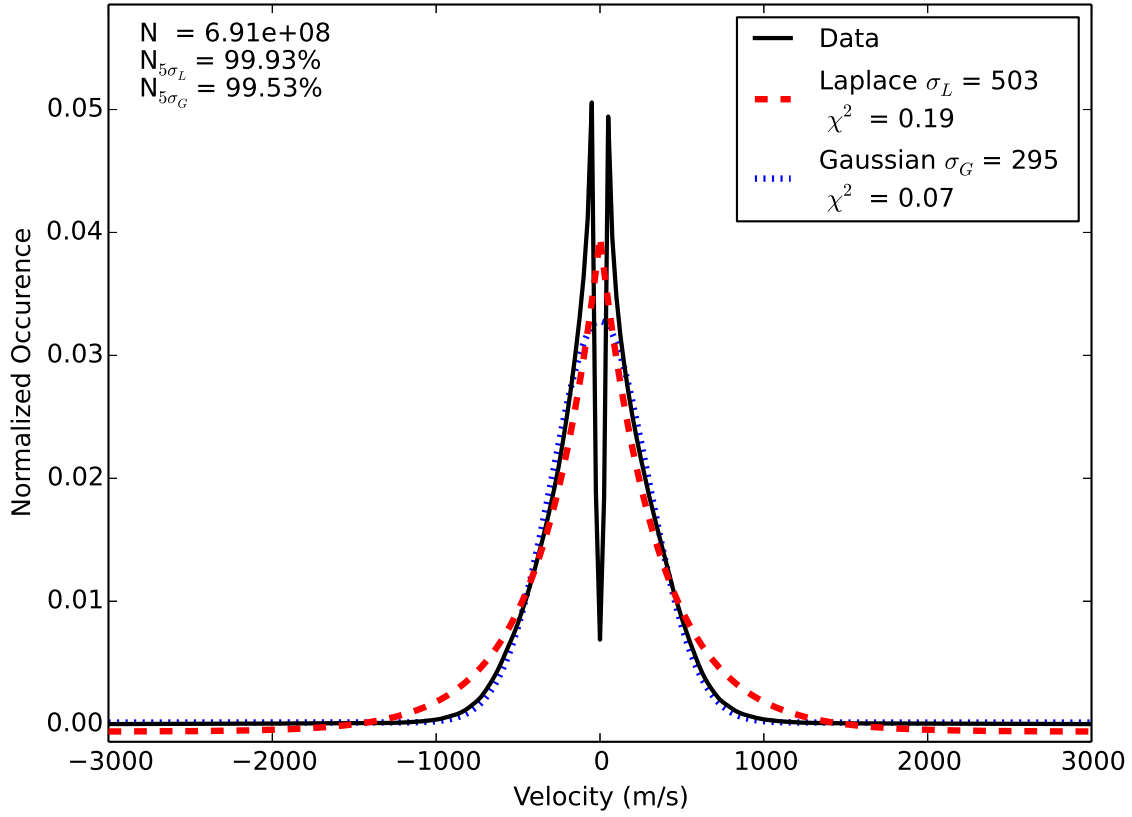


Figure 8.2: The normalized velocity distribution for SuperDARN radars in the polar cap. The velocity bins are 25 m/s in size. N is the total number of points in the plot. $N_{5\sigma_L}$ and $N_{5\sigma_G}$ indicate the amount of data points within 5 standard deviations of the mean for the Laplace fit and the Gaussian fit, respectively.

velocity measurements (Figure 8.3) follow a Laplace distributed velocity distribution. The mid-latitude radar velocity measurements (Figure 8.4) do not follow a velocity distribution that is accurately described by either a Laplace or a Gaussian distribution.

This statistical study resulted in the following: for the polar cap radars $5\sigma \approx 1500$ m/s, for the high-latitude radars $5\sigma \approx 2000$ m/s, and for mid-latitude radars $5\sigma \approx 1250$ m/s. Being conservative, setting $MUD = 2000$ m/s for polar cap (P) radars, $MUD = 2500$ m/s for high-latitude (HL) radars, and $MUD = 1750$ m/s for mid-latitude (ML) radars the following constraints for PRT are obtained from Equation 8.1,

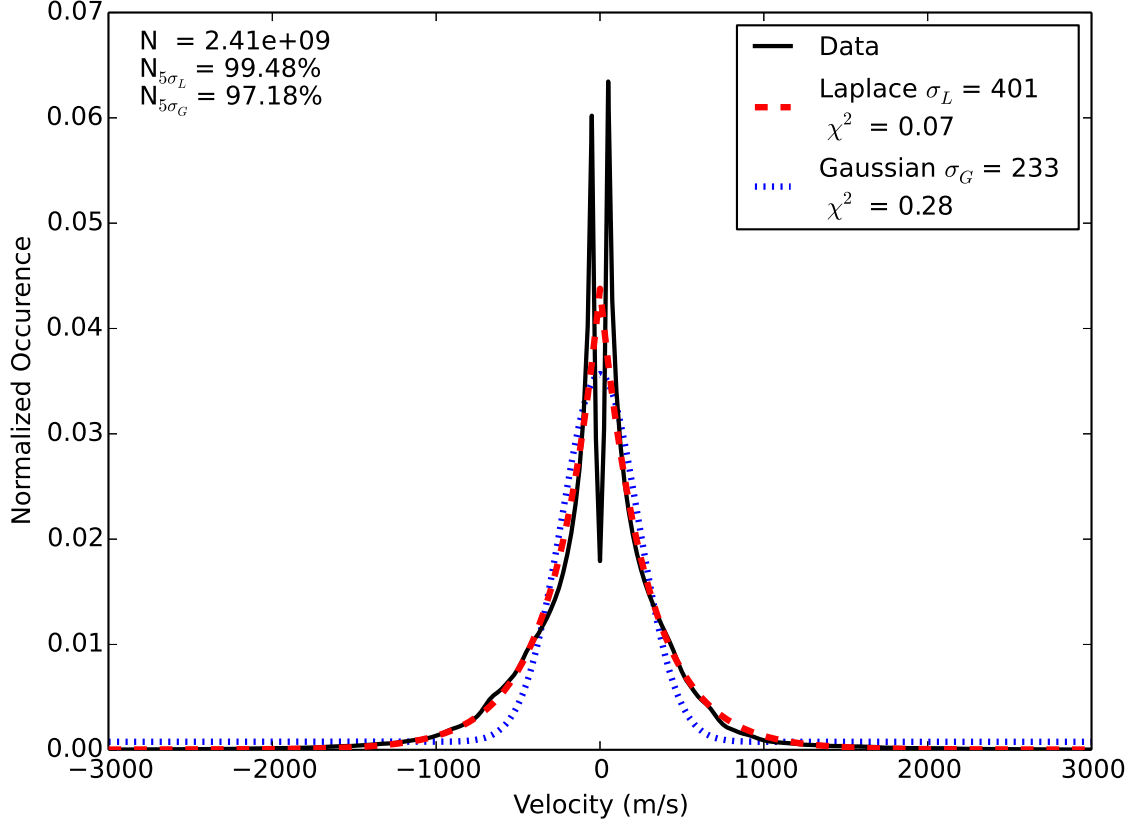


Figure 8.3: The normalized velocity distribution for SuperDARN radars in the high-latitudes. The velocity bins are 25 m/s in size. N is the total number of points in the plot. $N_{5\sigma_L}$ and $N_{5\sigma_G}$ indicate the amount of data points within 5 standard deviations of the mean for the Laplace fit and the Gaussian fit, respectively.

$$\tau_P \leq 75/f_{Tx}, \quad (8.3)$$

$$\tau_{HL} \leq 60/f_{Tx}, \quad (8.4)$$

$$\tau_{ML} \leq 86/f_{Tx}, \quad (8.5)$$

where f_{Tx} is the transmitted frequency in kHz.

With an upper limit for τ determined, how the ACF errors behave as a function of τ was briefly explored. Minimizing the ACF error will help to determine an optimal choice for τ . The average errors for 1 hour of data from the Saskatoon SuperDARN radar on 1 December

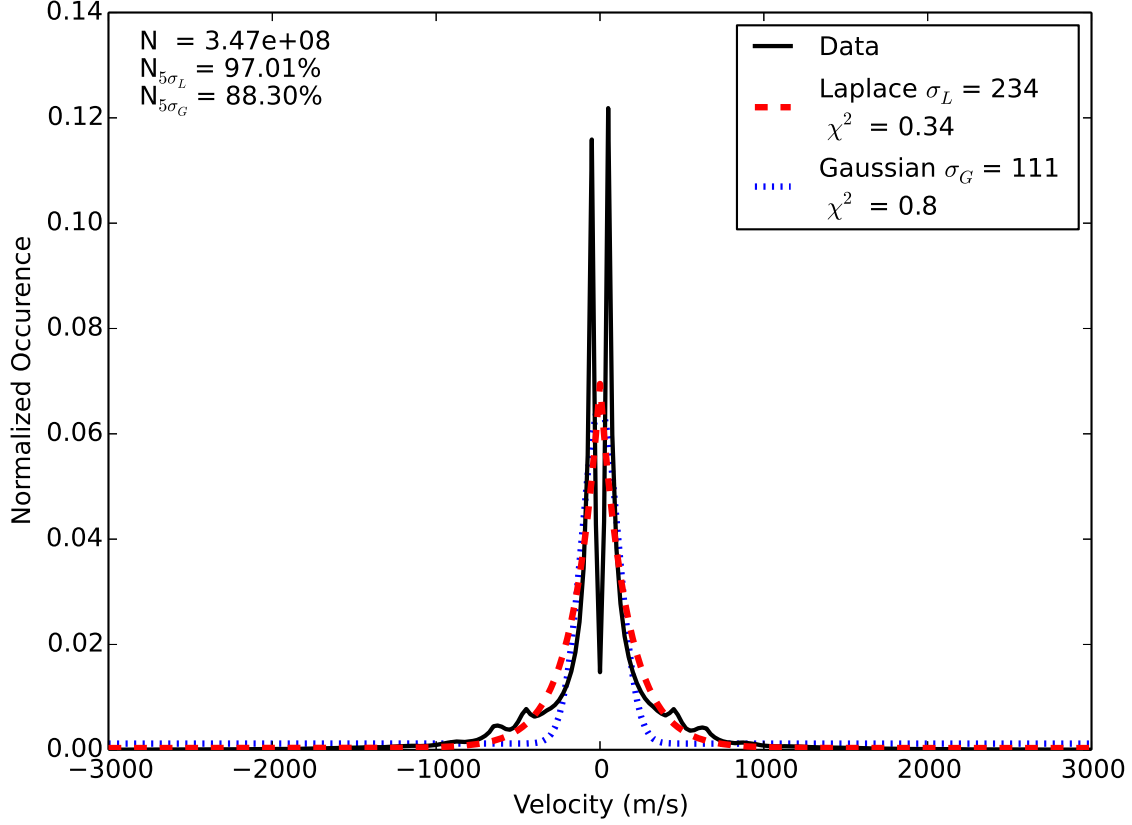


Figure 8.4: The normalized velocity distribution for SuperDARN radars in the mid-latitudes. The velocity bins are 25 m/s in size. N is the total number of points in the plot. $N_{5\sigma_L}$ and $N_{5\sigma_G}$ indicate the amount of data points within 5 standard deviations of the mean for the Laplace fit and the Gaussian fit, respectively.

2013 were computed using Equation 8.2 for the cases where $\tau = 1.8$ ms and $\tau = 2.4$ ms. The errors were then compared by plotting the difference between the errors for $\tau = 1.8$ ms and $\tau = 2.4$ ms in Figure 8.5.

Generally, there is improvement (decrease) in the ACF errors when increasing τ from 1.8 ms to 2.4 ms. This preliminary study suggests that a $\tau = 2.4$ ms is the optimum choice to balance both the MUD requirements and simultaneously minimize the ACF errors and therefore achieve the lowest fitted parameter errors. Using a value of 2.4 ms also satisfies Equations 8.3, 8.4, 8.5 for all SuperDARN operating frequencies. Future work on this topic should explore a wider range of SuperDARN data for determining improvements in ACF

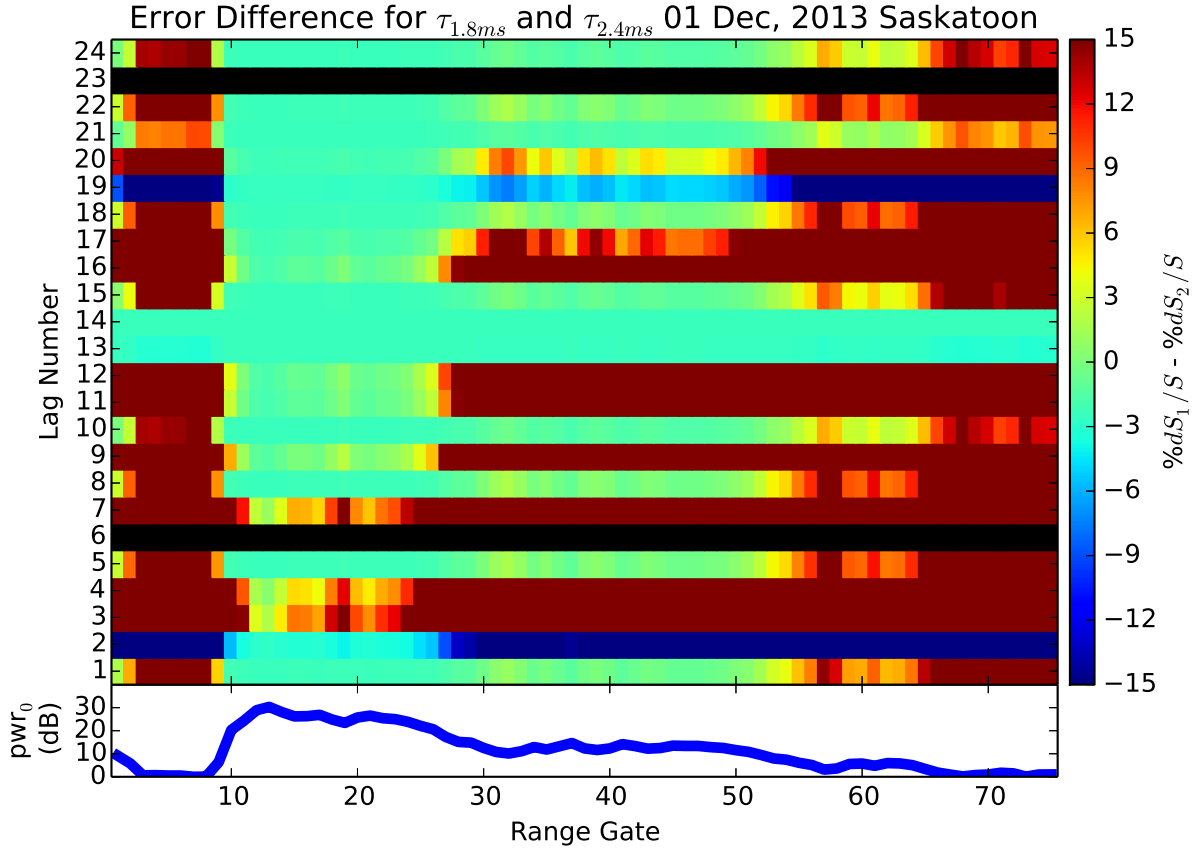


Figure 8.5: A plot of the average difference in ACF lag variance as calculated using Equation 8.2 between multiple-pulse sequences using $\tau = 1.8$ ms and $\tau = 2.4$ ms. Improvement in error is indicated by red and deterioration is indicated by blue.

errors beyond the single hour of SuperDARN data that was considered. Additionally, the range of theoretically observable velocities in the F region ionosphere should be considered in deciding what MUD to set for the SuperDARN radars in each region.

8.2.2 Inferred Electron Density as a Standard SuperDARN Data Product

SuperDARN radars are primarily used to measure the plasma circulation in the high-latitude ionosphere, but, as was discussed in Chapter 1, a recent statistical analysis using months of data by *Gillies et al. (2012)* showed that SuperDARN may also be used for electron den-

sity measurements. Using Bayesian inference and a multiple-frequency ACF model derived from the work of *Gillies et al. (2012)*, *Spaleta et al. (2015)* demonstrated that minute resolution electron density measurements may be inferred from multiple-frequency experiment modes running on SuperDARN radars. However, these minute resolution SuperDARN derived electron density measurements have yet to be quantitatively compared with or verified against established electron density measuring radar systems, such as incoherent scatter radars. Since SuperDARN radars are more numerous and have much larger fields-of-view than other ionospheric radar systems, verification of SuperDARN electron density measurements may provide a significant electron density dataset to the space physics community. The Canadian Resolute Incoherent Scatter Radar (RISR-C), located near Resolute, Nunavut, is a Canadian radar that has been performing absolute electron density measurements of the ionosphere since late 2015. The following presents the preliminary results of an experiment from 3 March, 2016 where the electron density measurements from RISR-C were compared with the electron density obtained using the Clyde River SuperDARN radar.

Following the previous work of *Spaleta et al. (2015)*, a Bayesian inference data fitting algorithm was built using the Python package `emcee`, which provides a Monte Carlo Markov Chain method of constructing the posterior distribution of model parameters (*Foreman-Mackey et al., 2012*). The Bayesian inference algorithm utilizes the FPFM developed in Chapter 7, but uses Bayesian inference in place of the non-linear least-squares fitting techniques. The resulting posterior distributions can be used to determine the most probable values of the parameters given the model and the data (e.g. *Spaleta et al., 2015*). Additionally, a joint experiment was performed where both the Clyde River SuperDARN radar and RISR-C were run at the same time. The objective of this experiment was to compare electron density measurements derived from SuperDARN radar data with calibrated absolute electron density measurements from RISR-C. Such a comparison would provide a way to validate the inferred SuperDARN electron densities.

In a statistical study using many months of SuperDARN data, *Gillies et al. (2012)* demonstrated that measurements taken using different SuperDARN transmission frequencies may be used to obtain estimates of the electron density. Subsequently, *Spaleta et al. (2015)* obtained minute resolution measurements of the electron density measurements with Super-

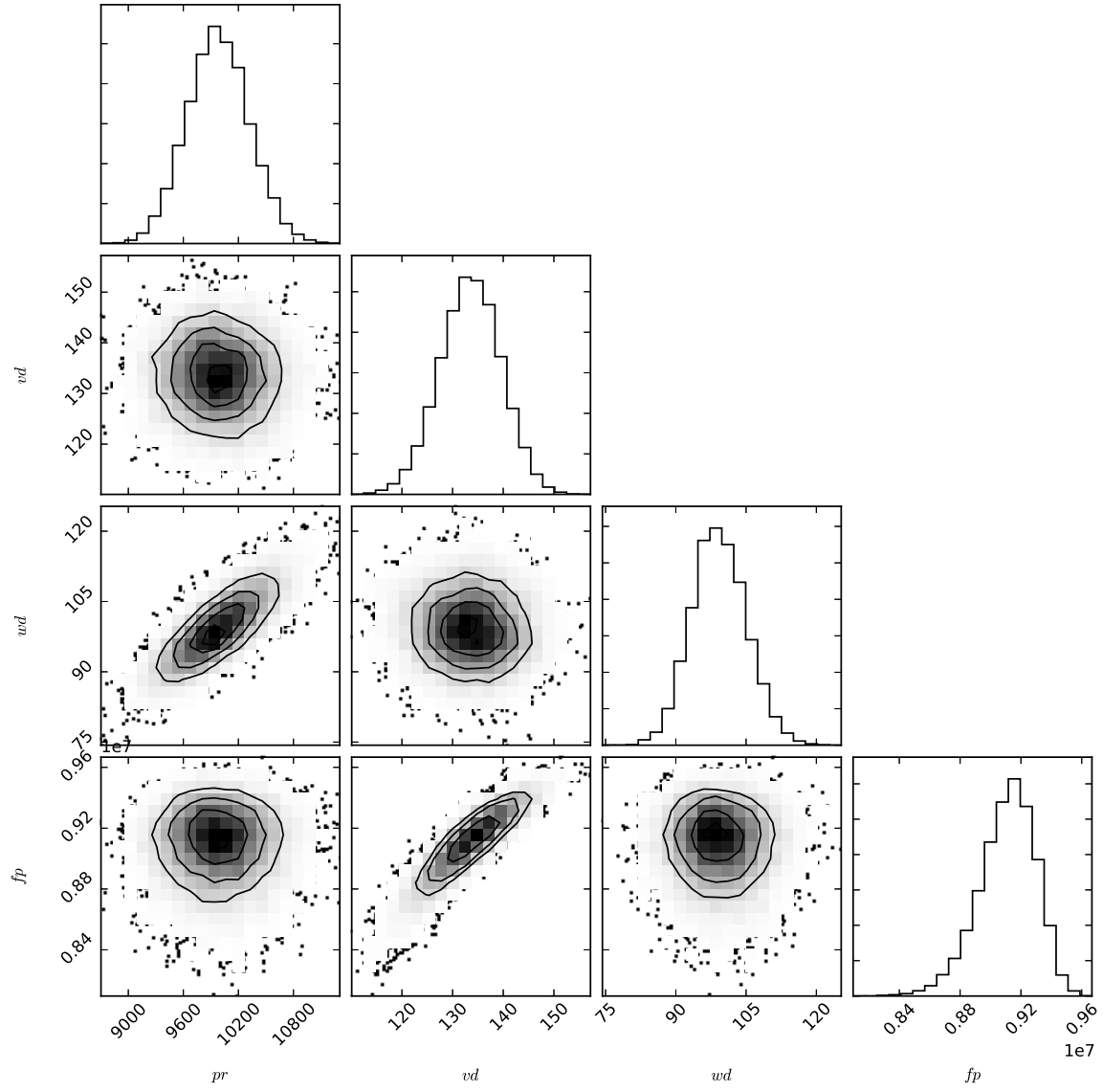


Figure 8.6: The posterior distribution functions for fitted power, velocity, spectral width, and plasma frequency resulting from a Bayesian fit of simulated data.

DARN using a model and a Bayesian fitting technique. To do this, SuperDARN ACFs were fitted to the model given by

$$R_i(t) = P_i e^{k_i w_d t} e^{j4\pi \sqrt{f_i^2 - f_p^2} v_d t / c}, \quad (8.6)$$

where i denotes data recorded from the f_i transmitted frequency, R_i is the autocorrelation function at time t , P_i is the echo power, k_i is the transmitted wavenumber, w_d is the spectral width, f_p is the plasma frequency, v_d is the Doppler velocity, and c is the speed of light. To obtain the electron density, parameters P_i , w_d , v_d , and f_p are fitted for and then simply

$$f_p^2 = \frac{n_e e^2}{m \epsilon_0}. \quad (8.7)$$

Fitting the data to the model performs best when the phase difference between the lags of ACFs measured from different frequencies is large. The difference in phase between the lags of different ACFs is increased by larger ionospheric velocities and larger frequency separations. Larger frequency separations also change the propagation paths of the two different frequencies meaning that it is less likely that scatter from both frequencies is received from the same volume of plasma. Figure 8.6 shows a plot of what the posterior distributions of the fitted parameters is expected to look like when the power pr , velocity vd , spectral width wd , and plasma frequency fp are all successfully resolved. As seen in the figure, if the parameters are well-resolved, then the posterior distributions are single peaked with a low variance (narrow). When a parameter is not well resolved, in general its posterior distribution will resemble a uniform distribution (wide).

A new 10 pulse multiple-pulse sequence, called *g10scan*, was designed and tested using a radar data simulation based on Equation 6.13 optimize the pulse-sequence parameters with the expected polar cap ionospheric conditions. Then a joint experiment between RISR-C and the Clyde River SuperDARN radar was performed on 3 March, 2016 from 16–20 UT. RISR-C operated in a 1-minute resolution, 51 beam mode. Clyde River operated in a 6 beam mode alternating transmitted frequencies scan-to-scan on 10 MHz and 12 MHz. The beams on Clyde River were chosen to overlap with RISR-C. Figure 8.7 shows the fields of view of both RISR-C and Clyde River during the experiment.

A comparison of the electron densities obtained by RISR-C and Clyde River are plotted in Figure 8.8. The electron densities from beam 42 of RISR-C, at 200 km, 250 km, and 300

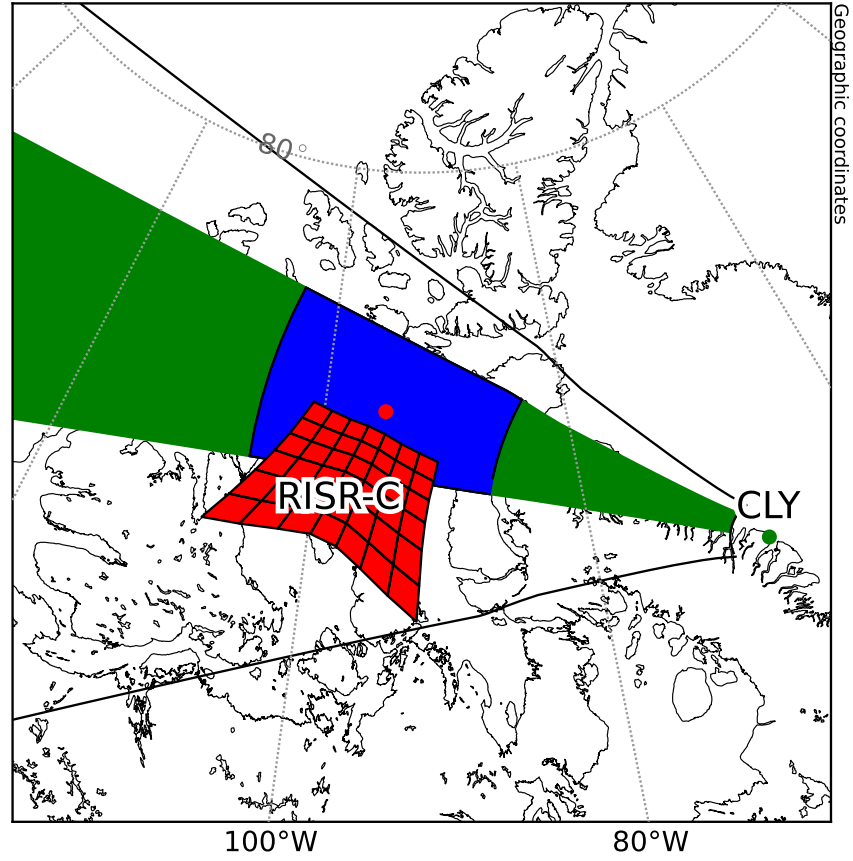


Figure 8.7: An illustration of the overlap of the fields of view of both RISR-C and the Clyde River SuperDARN radar during the 3 March, 2016 experiment.

km, are plotted as black, green, and red lines, respectively. The inferred Clyde River electron densities are plotted in blue. The comparison presented is preliminary and does not filter the data by considering propagation, nor does it attempt to use the Clyde River elevation angle to ray trace the expected altitude of the backscatter. Despite this, there does seem to be some agreement between the RISR-C and Clyde River electron densities near 16:15 UT and 17:15 UT. The anomalously large Clyde River densities (blue data points near 10^{12} m^{-3}) are unrealistic. These unrealistic densities are produced when the ACF measured using the lowest transmit frequency measures a larger velocity than the ACF measured using the

highest transmit frequency. When this occurs, it is physically impossible according to the magnetoionic theory (Appleton-Hartree equation), where the largest transmit frequency is expected to measure the largest velocity, unless the scattered signals from which the ACFs were made originated from two different places. Considering the frequency dependence of HF wave propagation in the ionosphere, this is a frustrating but not unexpected result.

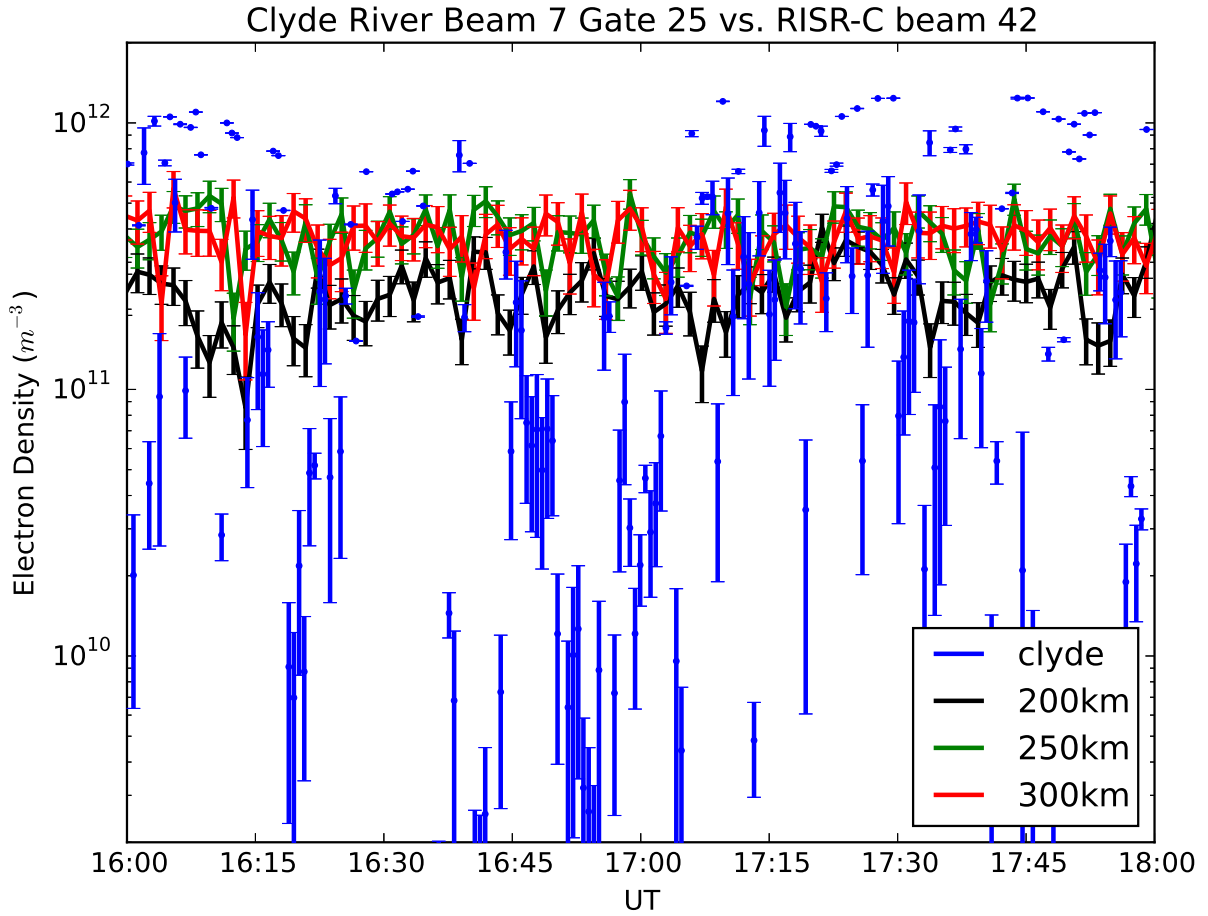


Figure 8.8: A plot comparing the electron density measured along beam 42 of RISR-C at different altitudes with the Clyde River inferred electron density.

Analysis of all the remaining coincident data is needed. Future experiments are also required such that a range of ionospheric conditions can be sampled and the optimal conditions for inferring electron densities using SuperDARN can be determined. A desirable improvement to the experiment mode would be to transmit on 2 or more frequencies simultaneously;

however, the Clyde River radar does not yet have this capability, but is expected to gain the ability in a year or two. To perform experiments using multiple simultaneous transmit and receive frequencies right now, the experiment could be performed using the Kodiak SuperDARN radar in Alaska in combination with the Poker Flat Incoherent Scatter Radar.

8.2.3 The Multiple-Incoherent-Pulse Technique

As discussed in detail in Chapter 5, the multiple-pulse technique of *Farley (1972)* solves the Doppler-dilemma for SuperDARN radars, but does so at the expense of contributing self-clutter to the SuperDARN ACF. When the interfering ranges that produce the self-clutter are uncorrelated, the self-clutter acts like an incoherent noise source and is attenuated by a factor of K , where K is the number of pulse sequences transmitted. However, when the interfering ranges are correlated, such as when ground scatter contributes, the self-clutter survives the averaging process, persisting and producing significant interference (recall the “spikes” in the bottom left panel of Figure 5.6). In this subsection, a novel variation on the multiple-pulse technique, called the Multiple-Incoherent-Pulse Technique (MIPT) will be introduced and discussed. As will be shown, MIPT causes all interfering signals to behave as incoherent noise, such that even ground scatter origin self-clutter is attenuated by a factor of K .

Previous efforts to reduce the self-clutter of the multiple-pulse technique have involved different variations of the multiple-pulse technique. A Multiple Pulse with Orthogonal Polarization (MPOP) technique, which involves alternating the transmitted pulses between left and right circular polarizations, was developed for use with the Jicamarca incoherence scatter radar (*Grydeland and Gustavsson, 2011*). For MPOP, the orthogonality of each pulse must be preserved during the entire transit of the pulses. Compared to the multiple-pulse technique, *Grydeland and Gustavsson (2011)* found that the MPOP reduces the variance of ACF lag estimates by up to a factor of 4. In order to measure the ionospheric electric field with a sample rate of 5 Hz, *Greenwald et al. (2008)* developed a novel pulse-sequence that can be transmitted on SuperDARN radars. The pulse-sequence involves transmitting a standard multiple-pulse sequence followed immediately by a time-reversed version of the same pulse sequence. The sequence used in *Greenwald et al. (2008)* produces approximately 50% more

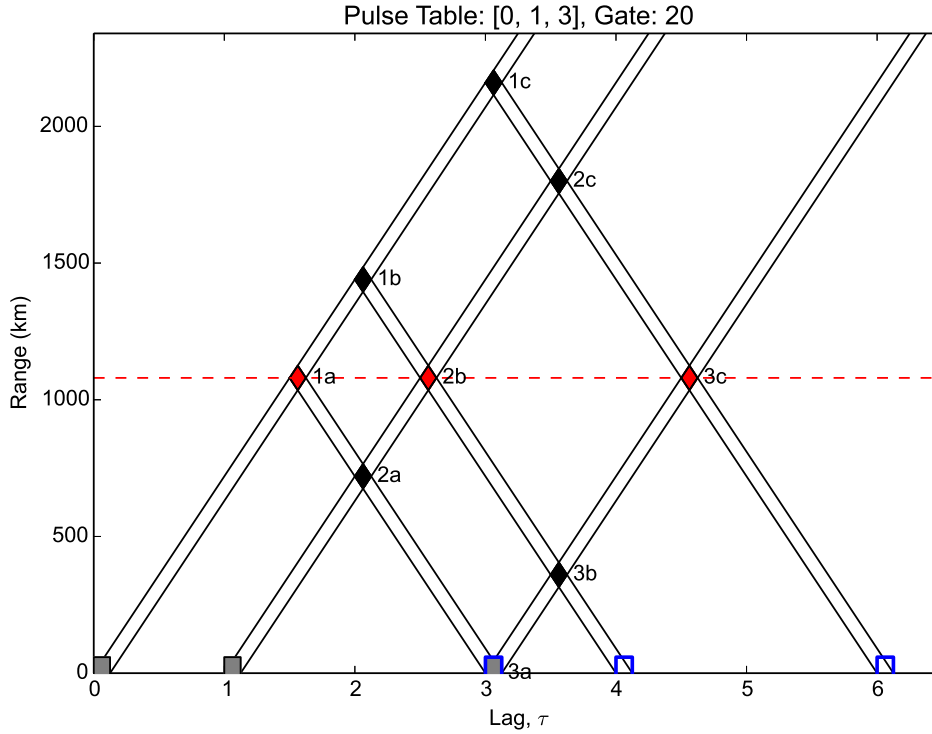


Figure 8.9: A Farley diagram of a multiple-pulse sequence using 3 pulses with $\tau = 2.4$ ms, $t_p = 300$ μ s, $N_p = 3$, and $\text{PTAB} = [0\tau, 1\tau, 3\tau]$. The grey filled rectangles indicate transmission pulses and the rectangles with blue outlines indicate received voltage samples (from scattered signals returning to the radar). The red diamonds indicate signal backscattered from a repeated range of interest, in contrast to the black diamonds indicating signal backscattered from non-repeated unwanted ranges.

lag estimates per second compared to the standard katscan pulse sequence typically used. An additional benefit of the *Greenwald et al.* (2008) technique is that it provides lag estimates for ranges where lag estimates are missing from the standard multiple-pulse sequence, due to receiver blanking when pulses are transmitted at the same time as signals are being received. The end result is that self-clutter is reduced by roughly a factor of two compared to the standard multiple-pulse technique and lag estimates are obtained for previously “bad” lags due to receiver blanking.

The approach taken by the MIPT technique is completely different than those previously attempted for SuperDARN, but is similar to the Simultaneous Multiple Pulse Repetition Frequency (SMPRF) technique proposed for usage with Doppler-weather radars (*Pirttilä et al.*,

2005). SuperDARN radars are not currently capable of transmitting on different polarizations. Additionally, the mirrored multiple-pulse technique does not address the problems caused by correlated self-clutter. Instead, the MIPT uses pseudo-random phase coding to transmit pulses that are incoherent with respect to one another. Due to the hardware used for Doppler-weather radars, transmitted pulses have a random phase with respect to one another and this was leveraged in the SMPRF technique to reduce self-clutter to an incoherent noise source. If the random phases used during transmission are recorded and then used to decode the received voltage samples, then on the average, the interfering ranges are not decoded by the pseudo-random coding. This means that regardless of the properties of the signals coming from the interfering ranges, they are “forced” to behave as an incoherent noise source due to the pseudo-random coding.

Consider the simple 3 pulse multiple-pulse sequence shown in Figure 8.9 and imagine that the pulses transmitted at times 0τ , 1τ , and 3τ are all coded with phase offsets given by $e^{j\phi_1}$, $e^{j\phi_2}$, and $e^{j\phi_3}$. The voltage samples received at 3τ , 4τ , and 6τ are then given by

$$V_{3\tau} = V_{1a}e^{j\phi_1} + V_{2a}e^{j\phi_2}, \quad (8.8)$$

$$V_{4\tau} = V_{1b}e^{j\phi_1} + V_{2b}e^{j\phi_2} + V_{3b}e^{j\phi_3}, \quad (8.9)$$

$$V_{6\tau} = V_{1c}e^{j\phi_1} + V_{2c}e^{j\phi_2} + V_{3c}e^{j\phi_3}. \quad (8.10)$$

In order to obtain the desired correlation of $V_{2b}V_{3c}^*$, we correlate $V_{4\tau}$ and $V_{6\tau}$ and obtain

$$(V_{4\tau}e^{-j\phi_2})(V_{6\tau}e^{-j\phi_3})^* = [(V_{1c}e^{j\phi_1} + V_{2c}e^{j\phi_2} + V_{3c}e^{j\phi_3})e^{-j\phi_2}] \quad (8.11)$$

$$\times [(V_{1b}e^{j\phi_1} + V_{2b}e^{j\phi_2} + V_{3b}e^{j\phi_3})e^{-j\phi_3}]^*, \quad (8.12)$$

noting that in order to obtain the desired correlation, the random phase offsets need to be removed (decoded). Simplifying yields

$$(V_{4\tau}e^{-j\phi_2})(V_{6\tau}e^{-j\phi_3})^* = V_{2c}V_{3b}^* + V_{2c}V_{1b}^*e^{-j(\phi_1-\phi_3)} + V_{2c}V_{2b}^*e^{-j(\phi_2-\phi_3)} \quad (8.13)$$

$$+ V_{1c}e^{j(\phi_1-\phi_2)}V_{3b}^* + V_{3c}e^{j(\phi_3-\phi_2)}V_{3b}^* \quad (8.14)$$

$$+ V_{1c}V_{1b}^*e^{j(\phi_3-\phi_2)} + V_{1c}V_{2b}^*e^{j(\phi_1-2\phi_2+\phi_3)} \quad (8.15)$$

$$+ V_{3c}V_{1b}^*e^{j(-\phi_1-\phi_2+2\phi_3)} + V_{3c}V_{2b}^*e^{j(-2\phi_2+2\phi_3)}. \quad (8.16)$$

Notice that all of the unwanted terms contain random phase offset. If several pulse-sequences are transmitted, each with a different set of random phase offsets, then even if the interfering ranges are correlated the random phase offset force the correlation of the interfering ranges to vary randomly from pulse-sequence to pulse sequence since

$$\langle V_{1i} V_{2i}^* e^{j\phi_i} \rangle = \langle V_{1i} V_{2i}^* \rangle \langle e^{j\phi_i} \rangle = 0, \quad (8.17)$$

because

$$\langle e^{j\phi_i} \rangle = 0, \quad (8.18)$$

if ϕ_i is a uniformly distributed random variable. So as long as the random phase offsets are generated by sampling a uniform distribution, then the self-clutter produced by the MIPT will be attenuated by a factor of K regardless of correlative properties of the interfering ranges.

The significance of this result is that the MIPT will suppress the strong self-clutter caused by ground scatter that is currently seen in SuperDARN data. Additionally, a new maximal estimate of the self-clutter, obtained by dividing Equation 5.15 by K , can be used in the FPFM to provide more accurate fitted parameter errors.

Currently, SuperDARN radar hardware is not capable of transmitting MIPT because the radars are not able to change the phase of the transmitter fast enough. New digital radar hardware that uses software defined radio technology is being developed and built by the Saskatoon SuperDARN group. The new fully digital radar will be capable of transmitting the MIPT. It is expected that the digital radar hardware will be operational at the Saskatoon SuperDARN radar site in 2018. In the meantime, testing of the MIPT can be performed using simulated SuperDARN radar data.

REFERENCES

- Aalo, V. A., G. P. Efthymoglou, and C. Chayawan (2007), On the envelope and phase distributions for correlated gaussian quadratures, *IEEE Communications Letters*, 11(12), 985–987, doi:10.1109/LCOMM.2007.071394.
- Abramowitz, M., and I. Stegun (2012), *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Dover Books on Mathematics, Dover Publications.
- Alfvén, H. (1942), On the existence of electromagnetic-hydrodynamic waves, *Ark. Mat. Astron. Fys.*, 29B.
- André, D., G. J. Sofko, K. Baker, and J. MacDougall (1998), Superdarn interferometry: Meteor echoes and electron densities from groundscatter, *Journal of Geophysical Research: Space Physics*, 103(A4), 7003–7015, doi:10.1029/97JA02923.
- André, R., M. Pinnock, and A. S. Rodger (1999), On the superdarn autocorrelation function observed in the ionospheric cusp, *Geophysical Research Letters*, 26(22), 3353–3356, doi:10.1029/1999GL003658.
- Angelopoulos, V. (2008), The THEMIS mission, *Space Science Reviews*, 141(1), 5, doi:10.1007/s11214-008-9336-1.
- Bahcivan, H., M. J. Nicolls, and G. Perry (2013), Comparison of SuperDARN irregularity drift measurements and F-region ion velocities from the resolute bay ISR, *Journal of Atmospheric and Solar-Terrestrial Physics*, 105(Supplement C), 325 – 331, doi:https://doi.org/10.1016/j.jastp.2013.02.002.
- Barabash, S., A. Fedorov, R. Lundin, and J.-A. Sauvaud (2007), Martian atmospheric erosion rates, *Science*, 315(5811), 501–503, doi:10.1126/science.1134358.
- Barthes, L., R. André, J.-C. Cerisier, and J.-P. Villain (1998), Separation of multiple echoes using a high-resolution spectral analysis for SuperDARN HF radars, *Radio Science*, 33(4), 1005–1017, doi:10.1029/98RS00714.
- Bauer, S., L. Brace, D. Hunten, D. Intriligator, W. Knudsen, A. Nagy, C. Russell, F. Scarf, and J. Wolfe (1977), The Venus ionosphere and solar wind interaction, *Space Science Reviews*, 20, 413–430, 10.1007/BF02186461.
- Baumjohann, W., and R. Treumann (1997), *Basic Space Plasma Physics*, Imperial College Press.

- Beckmann, P. (1962), Statistical distribution of the amplitude and phase of a multiply scattered field, *Journal of Research of the National Bureau of Standards*, 66D(3), 231–240.
- Bendat, J., and A. Piersol (2000), *Random data: analysis and measurement procedures*, Wiley series in probability and statistics: Texts and references section, Wiley.
- Bilitza, D., L.-A. McKinnell, B. Reinisch, and T. Fuller-Rowell (2011), The international reference ionosphere today and in the future, *Journal of Geodesy*, 85(12), 909–920, doi:10.1007/s00190-010-0427-x.
- Bjoland, L. M., X. Chen, Y. Jin, A. S. Reimer, Å. Skjæveland, M. R. Wessel, J. K. Burchill, L. B. N. Clausen, S. E. Haaland, and K. A. McWilliams (2015), Interplanetary magnetic field and solar cycle dependence of northern hemisphere F region Joule heating, *Journal of Geophysical Research: Space Physics*, doi:10.1002/2014JA020586.
- Blanchard, G. T., S. Sundeen, and K. B. Baker (2009), Probabilistic identification of high-frequency radar backscatter from the ground and ionosphere based on spectral characteristics, *Radio Science*, 44(5), doi:10.1029/2009RS004141.
- Brekke, A. (2013), *Physics of the Upper Polar Atmosphere*, Springer Atmospheric Sciences Series, Springer London, Limited.
- Brekke, P. (2012), *Our Explosive Sun*, Springer.
- Budden, K. (1961), *Radio Waves in the Ionosphere: The Mathematical Theory of the Reflection of Radio Waves from Stratified Ionised Layers*, University Press.
- Burbidge, E. M., G. R. Burbidge, W. A. Fowler, and F. Hoyle (1957), Synthesis of the elements in stars, *Rev. Mod. Phys.*, 29, 547–650, doi:10.1103/RevModPhys.29.547.
- Burch, J. L., J. Goldstein, and B. R. Sandel (2004), Cause of plasmasphere corotation lag, *Geophysical Research Letters*, 31(5), doi:10.1029/2003GL019164, 105802.
- Burch, J. L., R. B. Torbert, T. D. Phan, L.-J. Chen, T. E. Moore, R. E. Ergun, J. P. Eastwood, D. J. Gershman, P. A. Cassak, M. R. Argall, S. Wang, M. Hesse, C. J. Pollock, B. L. Giles, R. Nakamura, B. H. Mauk, S. A. Fuselier, C. T. Russell, R. J. Strangeway, J. F. Drake, M. A. Shay, Y. V. Khotyaintsev, P.-A. Lindqvist, G. Marklund, F. D. Wilder, D. T. Young, K. Torkar, J. Goldstein, J. C. Dorelli, L. A. Avanov, M. Oka, D. N. Baker, A. N. Jaynes, K. A. Goodrich, I. J. Cohen, D. L. Turner, J. F. Fennell, J. B. Blake, J. Clemmons, M. Goldman, D. Newman, S. M. Petriner, K. J. Trattner, B. Lavraud, P. H. Reiff, W. Baumjohann, W. Magnes, M. Steller, W. Lewis, Y. Saito, V. Coffey, and M. Chandler (2016), Electron-scale measurements of magnetic reconnection in space, *Science*, doi:10.1126/science.aaf2939.
- Burrell, A. G., S. E. Milan, G. W. Perry, T. K. Yeoman, and M. Lester (2015), Automatically determining the origin direction and propagation mode of high-frequency radar backscatter, *Radio Science*, 50(12), 1225–1245, doi:10.1002/2015RS005808, 2015RS005808.

- Burrell, A. G., T. K. Yeoman, S. E. Milan, and M. Lester (2016), Phase calibration of interferometer arrays at high-frequency radars, *Radio Science*, *51*(9), 1445–1456, doi:10.1002/2016RS006089, 2016RS006089.
- Carpenter, D. L. (1963), Whistler evidence of a ‘knee’ in the magnetospheric ionization density profile, *J. Geophys. Res.*, *68*(6), 1675–1682, doi:10.1029/JZ068i006p01675.
- Chatfield, C. (2013), *The Analysis of Time Series: Theory and Practice*, Monographs on Statistics and Applied Probability, Springer US.
- Chen, F. (1984), *Introduction to Plasma Physics and Controlled Fusion*, no. v. 1 in Introduction to Plasma Physics and Controlled Fusion, Springer.
- Chen, K., Z. Zhu, B. Ning, J. Lan, and F. Sun (2012), Developing a new mode for observation of ionospheric disturbances by digital ionosonde in ionospheric vertical sounding, *Radio Science*, *47*(3), doi:10.1029/2011RS004968, rS3009.
- Chisham, G., and M. P. Freeman (2013), A reassessment of SuperDARN meteor echoes from the upper mesosphere and lower thermosphere, *Journal of Atmospheric and Solar-Terrestrial Physics*, *102*(Supplement C), 207 – 221, doi:https://doi.org/10.1016/j.jastp.2013.05.018.
- Chisham, G., M. Lester, S. Milan, M. Freeman, W. Bristow, A. Grocott, K. McWilliams, J. Ruohoniemi, T. Yeoman, P. Dyson, R. Greenwald, T. Kikuchi, M. Pinnock, J. Rash, N. Sato, G. Sofko, J.-P. Villain, and A. Walker (2007), A decade of the Super Dual Auroral Radar Network (SuperDARN): scientific achievements, new techniques and future directions, *Surveys in Geophysics*, *28*, 33–109, doi:10.1007/s10712-007-9017-8.
- Choudhary, R. K., K. M. Ambili, S. Choudhury, M. B. Dhanya, and A. Bhardwaj (2016), On the origin of the ionosphere at the moon using results from chandrayaan-1 s-band radio occultation experiment and a photochemical model, *Geophysical Research Letters*, doi:10.1002/2016GL070612, 2016GL070612.
- Clausen, L. B. N., J. B. H. Baker, J. M. Ruohoniemi, R. A. Greenwald, E. G. Thomas, S. G. Shepherd, E. R. Talaat, W. A. Bristow, Y. Zheng, A. J. Coster, and S. Sazykin (2012), Large-scale observations of a subauroral polarization stream by midlatitude SuperDARN radars: Instantaneous longitudinal velocity variations, *Journal of Geophysical Research: Space Physics*, *117*(A5), doi:10.1029/2011JA017232, a05306.
- Cook, C. (2012), *Radar Signals: An Introduction to Theory and Application*, Electrical science series, Elsevier Science.
- Cousins, E. D. P., T. Matsuo, and A. D. Richmond (2013), Superdarn assimilative mapping, *Journal of Geophysical Research: Space Physics*, *118*(12), 7954–7962, doi:10.1002/2013JA019321.
- Cousins, E. D. P., T. Matsuo, and A. D. Richmond (2015), Mapping high-latitude ionospheric electrodynamics with SuperDARN and AMPERE, *Journal of Geophysical Research: Space Physics*, *120*(7), 5854–5870, doi:10.1002/2014JA020463, 2014JA020463.

- Cowley, S. W. H., and M. Lockwood (1992), Excitation and decay of solar wind-driven flows in the magnetosphere-ionosphere, *Annales Geophysicae Atmospheres Hydrospheres and Space Sciences*, *10*(1-2), 103–115.
- Dahlgren, H., G. W. Perry, J. L. Semeter, J.-P. St.-Maurice, K. Hosokawa, M. J. Nicolls, M. Greffen, K. Shiokawa, and C. Heinselman (2012), Space-time variability of polar cap patches: Direct evidence for internal plasma structuring, *Journal of Geophysical Research: Space Physics*, *117*(A9), doi:10.1029/2012JA017961, a09312.
- Davies, J. A., M. Lester, B. Jenkins, and R. J. Moffett (1995), Dayside ion frictional heating: EISCAT observations and comparison with model results, *J. Atmos. Terr. Phys.*, *57*(7), 775–793, doi:10.1016/0021-9169(94)00051-O.
- de Larquier, S., A. Ribeiro, N. A. Frissell, J. Spaleta, B. Kunduri, E. G. Thomas, J. Ruohoniemi, and J. B. Baker (2013), A new open-source Python-based Space Weather data access, visualization, and analysis toolkit, *AGU Fall Meeting Abstracts*.
- DeForest, C. E., W. H. Matthaeus, N. M. Viall, and S. R. Cranmer (2016), Fading coronal structure and the onset of turbulence in the young solar wind, *The Astrophysical Journal*, *828*(2), 66.
- Dharmawansa, P., N. Rajatheva, and C. Tellambura (2009), Envelope and phase distribution of two correlated gaussian variables, *IEEE Transactions on Communications*, *57*(4), 915–921, doi:10.1109/TCOMM.2009.04.070065.
- Diaconis, P., S. Holmes, and R. Montgomery (2007), Dynamical bias in the coin toss, *SIAM Review*, *49*(2), 211–235.
- Dungey, J. W. (1961), Interplanetary magnetic field and the auroral zones, *Phys. Rev. Lett.*, *6*, 47–48, doi:10.1103/PhysRevLett.6.47.
- Ecklund, W. L., B. B. Balsley, and R. A. Greenwald (1975), Crossed beam measurements of the diffuse radar aurora, *Journal of Geophysical Research*, *80*(13), 1805–1809, doi:10.1029/JA080i013p01805.
- Edberg, N. J. T., A. I. Eriksson, E. Odelstad, E. Vigren, D. J. Andrews, F. Johansson, J. L. Burch, C. M. Carr, E. Cupido, K.-H. Glassmeier, R. Goldstein, J. S. Halekas, P. Henri, C. Koenders, K. Mandt, P. Mokashi, Z. Nemeth, H. Nilsson, R. Ramstad, I. Richter, and G. S. Wieser (2016), Solar wind interaction with comet 67p: Impacts of corotating interaction regions, *Journal of Geophysical Research: Space Physics*, *121*(2), 949–965, doi:10.1002/2015JA022147, 2015JA022147.
- Eddington, A. S. (1920), The internal constitution of the stars, *Nature*, *106*, 14–20, doi:10.1038/106014a0.
- Enghoff, M. B., J. O. P. Pedersen, U. I. Uggerhøj, S. M. Paling, and H. Svensmark (2011), Aerosol nucleation induced by a high energy particle beam, *Geophysical Research Letters*, *38*(9), doi:10.1029/2011GL047036, l09805.

- Fang, X., C. E. Randall, D. Lummerzheim, W. Wang, G. Lu, S. C. Solomon, and R. A. Frahm (2010), Parametrization of monoenergetic electron impact ionization, *Geophys. Research Letters*, *37*(22).
- Farley, D. T. (1969), Incoherent scatter correlation function measurements, *Radio Science*, *4*(10), 935–953, doi:10.1029/RS004i010p00935.
- Farley, D. T. (1972), Multiple-pulse incoherent-scatter correlation function measurements, *Radio Science*, *7*(6), 661–666, doi:10.1029/RS007i006p00661.
- Farley, D. T., and T. Hagfors (2006), *AGF 304 Arctic Geophysics Textbook Manuscript (unpublished)*, The University Courses on Svalbard.
- Fear, R. C., S. E. Milan, A. N. Fazakerley, K.-H. Fornacon, C. M. Carr, and I. Dandouras (2009), Simultaneous observations of flux transfer events by THEMIS, Cluster, Double Star, and SuperDARN: Acceleration of FTEs, *Journal of Geophysical Research: Space Physics*, *114*(A10), doi:10.1029/2009JA014310, a10213.
- Fiori, R. A. D., D. H. Boteler, A. V. Koustov, G. V. Haines, and J. M. Ruohoniemi (2010a), Spherical cap harmonic analysis of Super Dual Auroral Radar Network (SuperDARN) observations for generating maps of ionospheric convection, *J. Geophys. Res.*, *115*, doi:10.1029/2009JA015055.
- Fiori, R. A. D., D. H. Boteler, A. V. Koustov, G. V. Haines, and J. M. Ruohoniemi (2010b), Spherical cap harmonic analysis of super dual auroral radar network (superdarn) observations for generating maps of ionospheric convection, *Journal of Geophysical Research: Space Physics*, *115*(A7), doi:10.1029/2009JA015055, a07307.
- Fisher, P. D. (1992), Improving on police radar, *IEEE Spectrum*, *29*(7), 38–43, doi:10.1109/6.144510.
- Fisher, R. (1958), *Statistical Methods for Research Workers*, Biological monographs and manuals, Hafner.
- Foreman-Mackey, D., D. W. Hogg, D. Lang, and J. Goodman (2012), emcee: The MCMC hammer, doi:10.1086/670067.
- Fortescue, P., G. Swinerd, and J. Stark (2011), *Spacecraft Systems Engineering*, Wiley.
- Fuselier, S. A., and I. H. Cairns (2015), Plasma properties at the voyager 1 crossing of the heliopause, *Journal of Physics: Conference Series*, *642*(1), 012,010.
- Gao, Y. (2012), Comparing the cross polar cap potentials measured by superdarn and amie during saturation intervals, *Journal of Geophysical Research: Space Physics*, *117*(A8), doi:10.1029/2012JA017690, a08325.
- Gillies, R., G. Hussey, G. Sofko, P. Ponomarenko, and K. McWilliams (2011), Improvement of HF coherent radar line-of-sight velocities by estimating the refractive index in the scattering volume using radar frequency shifting, *Journal of Geophysical Research A: Space Physics*, *116*(1).

- Gillies, R., G. Hussey, G. Sofko, and K. McWilliams (2012), A statistical analysis of SuperDARN scattering volume electron densities and velocity corrections using a radar frequency shifting technique, *Submitted to: Journal of Geophysical Research A: Space Physics*, In Press, Accepted 6 July 2012.
- Goh, C., J. C. Devlin, D. Deng, A. McDonald, and M. R. Kamarudin (2014), Uncorrelated weighted median filtering for noise removal in superdarn, in *Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on*, pp. 981–988, doi:10.1109/CSE.2014.196.
- Gombosi, T. (1998), *Physics of the Space Environment*, Cambridge Atmospheric and Space Science Series, Cambridge University Press.
- Greenwald, R. A., W. Weiss, E. Nielsen, and N. R. Thomson (1978), STARE: A new radar auroral backscatter experiment in northern Scandinavia, *Radio Science*, *13*(6), 1021–1039, doi:10.1029/RS013i006p01021.
- Greenwald, R. A., K. B. Baker, and J. P. Villain (1983), Initial studies of small-scale F region irregularities at very high latitudes, *Radio Science*, *18*(6), 1122–1132, doi:10.1029/RS018i006p01122.
- Greenwald, R. A., K. B. Baker, J. R. Dudeney, M. Pinnock, T. B. Jones, E. C. Thomas, J. P. Villain, J. C. Cerisier, C. Senior, C. Hanuise, R. D. Hunsucker, G. Sofko, J. Koehler, E. Nielsen, R. Pellinen, A. D. M. Walker, N. Sato, and H. Yamagishi (1995), DARN/SuperDARN, *Space Science Reviews*, *71*(1), 761–796, doi:10.1007/BF00751350.
- Greenwald, R. A., K. Oksavik, R. Barnes, J. M. Ruohoniemi, J. Baker, and E. R. Talaat (2008), First radar measurements of ionospheric electric fields at sub-second temporal resolution, *Geophysical Research Letters*, *35*(3), doi:10.1029/2007GL032164.
- Grydeland, T., and B. Gustavsson (2011), Orthogonal-polarization multipulse sequences, *Radio Science*, *46*(1), doi:10.1029/2010RS004425, rS1003.
- Haines, B. (1978), *An Introduction to Quantitative Economics*, Economics and society series, G. Allen & Unwin.
- Hall, G. E., J. W. MacDougall, D. R. Moorcroft, J.-P. St.-Maurice, A. H. Manson, and C. E. Meek (1997), Super dual auroral radar network observations of meteor echoes, *Journal of Geophysical Research: Space Physics*, *102*(A7), 14,603–14,614, doi:10.1029/97JA00517.
- Hanuise, C., J. P. Villain, D. Gresillon, B. Cabrit, R. A. Greenwald, and K. B. Baker (1993), Interpretation of HF radar ionospheric Doppler spectra by collective wave scattering theory, *Annales Geophysicae*, *11*, 29–39.
- Hargreaves, J. (1995a), *The Solar-Terrestrial Environment: An Introduction to Geospace - the Science of the Terrestrial Upper Atmosphere, Ionosphere, and Magnetosphere*, Cambridge Atmospheric and Space Science Series, Cambridge University Press.

- Hargreaves, J. K. (1995b), *The Solar-Terrestrial Environment: An Introduction to Geospace - the Science of the Terrestrial Upper Atmosphere, Ionosphere, and Magnetosphere*, Cambridge Press.
- Hosokawa, K., T. Ogawa, N. F. Arnold, M. Lester, N. Sato, and A. S. Yukimatu (2005), Extraction of polar mesosphere summer echoes from SuperDARN data, *Geophysical Research Letters*, *32*(12), doi:10.1029/2005GL022788, 112801.
- Hubert, B., S. E. Milan, A. Grocott, C. Blockx, S. W. H. Cowley, and J.-C. Gérard (2006), Dayside and nightside reconnection rates inferred from image fuv and super dual auroral radar network data, *Journal of Geophysical Research: Space Physics*, *111*(A3), doi:10.1029/2005JA011140, a03217.
- Hussey, G. C., C. Haldoupis, A. Bourdillon, and D. André (1999), Spatial occurrence of decameter midlatitude e region backscatter, *Journal of Geophysical Research: Space Physics*, *104*(A5), 10,071–10,080, doi:10.1029/1999JA900105.
- Hussey, G. C., C. E. Meek, D. André, A. H. Manson, G. J. Sofko, and C. M. Hall (2000), A comparison of northern hemisphere winds using SuperDARN meteor trail and mf radar wind measurements, *Journal of Geophysical Research: Atmospheres*, *105*(D14), 18,053–18,066, doi:10.1029/2000JD900272.
- Hysell, D. L., F. S. Rodrigues, J. L. Chau, and J. D. Huba (2008), Full profile incoherent scatter analysis at jicamarca, *Annales Geophysicae*, *26*(1), 59–75, doi:10.5194/angeo-26-59-2008.
- Imber, S. M., S. E. Milan, and M. Lester (2013), The Heppner-Maynard Boundary measured by SuperDARN as a proxy for the latitude of the auroral oval, *Journal of Geophysical Research: Space Physics*, *118*(2), 685–697, doi:10.1029/2012JA018222.
- Isserlis, L. (1918), On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, *Biometrika*, *12*(1/2), pp. 134–139.
- Jenkins, B., and M. J. Jarvis (1999), Mesospheric winds derived from superdarn hf radar meteor echoes at halley, antarctica, *Earth, Planets and Space*, *51*(7), 685–689, doi:10.1186/BF03353226.
- Kelley, M. (2009), *The Earth's Ionosphere: Plasma Physics and Electrodynamics*, International Geophysics Series, Academic Press.
- Keskinen, M. J., and S. L. Ossakow (1983), Theories of high-latitude ionospheric irregularities: A review, *Radio Science*, *18*(6), 1077–1091, doi:10.1029/RS018i006p01077.
- Kivelson, M., and C. Russell (1995), *Introduction to Space Physics*, Cambridge atmospheric and space science series, Cambridge University Press.
- Kolawole, M. (2002), *Radar Systems, Peak Detection and Tracking*, Communications engineering, Newnes.

- Koustov, A., N. Nishitani, Y. Ebihara, T. Kikuchi, M. R. Hairston, and D. Andre (2008), Subauroral polarization streams: observations with the hokkaido and king salmon SuperDARN radars and modeling, *Annales Geophysicae*, *26*(11), 3317–3327, doi:10.5194/angeo-26-3317-2008.
- Krall, N., and A. Trivelpiece (1973), *Principles of plasma physics*, no. v. 0-911351 in International series in pure and applied physics, McGraw-Hill.
- Landee, R., A. Albrecht, and D. Davis (1957), *Electronic Designers' Handbook*, McGraw-Hill Book Company.
- Laundal, K. M., I. Cnossen, S. E. Milan, S. E. Haaland, J. Coxon, N. M. Pedatella, M. Förster, and J. P. Reistad (2016), North–south asymmetries in earth’s magnetic field, *Space Science Reviews*, pp. 1–33, doi:10.1007/s11214-016-0273-0.
- Levanon, N., and E. Mozeson (1988), *Radar principles*, John Wiley and Sons, Inc.
- Levanon, N., and E. Mozeson (2004), *Radar signals*, John Wiley and Sons, Inc.
- Levenburg, K. (1944), A method for the solution of certain non-linear problems in least squares, *Quarterly of Applied Mathematics*, *2*(2), 164–168.
- Liao, R., H. Zheng, S. Grzybowski, and L. Yang (2011), Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers, *Electric Power Systems Research*, *81*(12), 2074 – 2080, doi:https://doi.org/10.1016/j.epsr.2011.07.020.
- Madden, J. L. (1982), Electrocoagulation in the treatment of cancer of the rectum., *Bulletin of the New York Academy of Medicine*, *58*(6), 552–558.
- Mardia, K. (1972), *Statistics of Directional Data*, Probability and Mathematical Statistics a Series of Monographs and Textbooks, Academic Press.
- Marquardt, D. W. (1963), An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society for Industrial and Applied Mathematics*, *11*(2), 431–441, doi:10.1137/0111030.
- McDonald, A. J., J. Whittington, S. de Larquier, E. Custovic, T. A. Kane, and J. C. Devlin (2013), Elevation angle-of-arrival determination for a standard and a modified superdarn hf radar layout, *Radio Science*, *48*(6), 709–721, doi:10.1002/2013RS005157.
- Milan, S. E., J. A. Davies, and M. Lester (1999), Coherent HF radar backscatter characteristics associated with auroral forms identified by incoherent radar techniques: A comparison of CUTLASS and EISCAT observations, *Journal of Geophysical Research: Space Physics*, *104*(A10), 22,591–22,604, doi:10.1029/1999JA900277.
- Milan, S. E., G. Provan, and B. Hubert (2007), Magnetic flux transport in the Dungey cycle: A survey of dayside and nightside reconnection rates, *Journal of Geophysical Research: Space Physics*, *112*(A1), doi:10.1029/2006JA011642, a01209.

- Milan, S. E., J. S. Gosling, and B. Hubert (2012), Relationship between interplanetary parameters and the magnetopause reconnection rate quantified from observations of the expanding polar cap, *Journal of Geophysical Research: Space Physics*, *117*(A3), doi: 10.1029/2011JA017082, a03226.
- Miyoshi, Y., T. Ono, T. Takashima, K. Asamura, M. Hirahara, Y. Kasaba, A. Matsuoka, H. Kojima, K. Shiokawa, K. Seki, M. Fujimoto, T. Nagatsuma, C. Cheng, Y. Kazama, S. Kasahara, T. Mitani, H. Matsumoto, N. Higashio, A. Kumamoto, S. Yagitani, Y. Kasahara, K. Ishisaka, L. Blomberg, A. Fujimoto, Y. Katoh, Y. Ebihara, Y. Omura, M. Nosé, T. Hori, Y. Miyashita, Y.-M. Tanaka, and T. T. Segawa (2013), *The Energization and Radiation in Geospace (ERG) Project*, pp. 103–116, American Geophysical Union, doi: 10.1029/2012GM001304.
- Moorcroft, D. R. (2004), The shape of auroral backscatter spectra, *Geophysical Research Letters*, *31*(9), doi:10.1029/2003GL019340, l09802.
- Mori, Y. (1981), Evidence of an 11-year periodicity in tree-ring series from formosa related to the sunspot cycle, *Journal of Climatology*, *1*(4), 345–353, doi:10.1002/joc.3370010407.
- Murray, D. B., and S. W. Teare (1993), Probability of a tossed coin landing on edge, *Phys. Rev. E*, *48*, 2547–2552, doi:10.1103/PhysRevE.48.2547.
- Mursula, K., and T. Hiltula (2003), Bashful ballerina: Southward shifted heliospheric current sheet, *Geophys. Res. Lett.*, *30*(22), doi:10.1029/2003GL018201.
- Musk, E. (2013), Tweet: Latest vertical landing rocket vid. 325M altitude, hold against wind, land w T/W₁, radar in loop, <https://twitter.com/elonmusk/status/353309492131278848>, tweet Retrieved: 29 September, 2016.
- Nikoukar, R., F. Kamalabadi, E. Kudeki, and M. Sulzer (2012), On resolution/error trade-offs in incoherent scatter radar measurements, *Radio Science*, *47*(1), doi:10.1029/2011RS004685, rS1008.
- Ogawa, T., N. Nishitani, S. Kawamura, and Y. Murayama (2013), Mesosphere summer echoes observed with the SuperDARN Hokkaido HF radar at Rikubetsu, Japan (43.5n), *Earth, Planets and Space*, *65*(12), 1593–1597, doi:10.5047/eps.2013.07.009.
- Oksavik, K., J. M. Ruohoniemi, R. A. Greenwald, J. B. H. Baker, J. Moen, H. C. Carlson, T. K. Yeoman, and M. Lester (2006), Observations of isolated polar cap patches by the european incoherent scatter (eiscat) svalbard and super dual auroral radar network (superdarn) finland radars, *Journal of Geophysical Research: Space Physics*, *111*(A5), doi: 10.1029/2005JA011400, a05310.
- Parker, E. N., and M. Krook (1956), Diffusion and Severing of Magnetic Lines of Force., *Astrophysical Journal*, *124*, 214, doi:10.1086/146216.
- Petschek, H. E. (1964), Magnetic Field Annihilation, *NASA Special Publication*, *50*, 425.

- Pirttilä, J., M. S. Lehtinen, A. Huuskonen, and M. Markkanen (2005), A proposed solution to the range–doppler dilemma of weather radar measurements by using the SMPRF codes, practical results, and a comparison with operational measurements, *Journal of Applied Meteorology*, *44*(9), 1375–1390, doi:10.1175/JAM2288.1.
- Ponomarenko, P. V., and C. L. Waters (2006), Spectral width of SuperDARN echoes: measurement, use and physical interpretation, *Annales Geophysicae*, *24*(1), 115–128, doi:10.5194/angeo-24-115-2006.
- Ponomarenko, P. V., F. W. Menk, and C. L. Waters (2003), Visualization of ulf waves in superdarn data, *Geophysical Research Letters*, *30*(18), doi:10.1029/2003GL017757, 1926.
- Ponomarenko, P. V., F. W. F. W. Menk, C. L. Waters, and M. D. Sciffer (2005), Pc3-4 ulf waves observed by the superdarn tiger radar, *Annales Geophysicae*, *23*(4), 1271–1280, doi:10.5194/angeo-23-1271-2005.
- Ponomarenko, P. V., C. L. Waters, and F. W. Menk (2008), Effects of mixed scatter on superdarn convection maps, *Annales Geophysicae*, *26*(6), 1517–1523, doi:10.5194/angeo-26-1517-2008.
- Porteous, J., A. M. Samson, K. A. Berrington, and I. W. McCrea (2003), Automated detection of satellite contamination in incoherent scatter radar spectra, *Annales Geophysicae*, *21*(5), 1177–1182, doi:10.5194/angeo-21-1177-2003.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1986), *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press.
- Reed, I. (1962), On a moment theorem for complex Gaussian processes, *Information Theory, IRE Transactions on*, *8*(3), 194–195, doi:10.1109/TIT.1962.1057719.
- Regan, H. M., M. Colyvan, and M. A. Burgman (2002), A taxonomy and treatment of uncertainty for ecology and conservation biology, *Ecological applications*, *12*(2), 618–628.
- Reimer, A. S., and G. C. Hussey (2015), Estimating self-clutter of the multiple-pulse technique, *Radio Science*, *50*(7), 698–711, doi:10.1002/2015RS005706, 2015RS005706.
- Reimer, A. S., G. C. Hussey, and S. R. Dueck (2016), On the statistics of superdarn autocorrelation function estimates, *Radio Science*, *51*(6), 690–703, doi:10.1002/2016RS005975, 2016RS005975.
- Ribeiro, A. J., J. M. Ruohoniemi, J. B. H. Baker, L. B. N. Clausen, S. de Larquier, and R. A. Greenwald (2011), A new approach for identifying ionospheric backscatter in midlatitude SuperDARN HF radar observations, *Radio Science*, *46*(4), doi:10.1029/2011RS004676, rS4011.
- Ribeiro, A. J., J. M. Ruohoniemi, P. V. Ponomarenko, L. B. N. Clausen, J. B. H. Baker, R. A. Greenwald, K. Oksavik, and S. de Larquier (2013a), A comparison of SuperDARN ACF fitting methods, *Radio Science*, *48*(3), 274–282, doi:10.1002/rds.20031.

- Ribeiro, A. J., P. V. Ponomarenko, J. M. Ruohoniemi, J. B. H. Baker, L. B. N. Clausen, R. A. Greenwald, and S. de Larquier (2013b), A realistic radar data simulator for the Super Dual Auroral Radar Network, *Radio Science*, *48*(3), 283–288, doi:10.1002/rds.20032.
- Richards, M., W. Holm, and J. Scheer (2010), *Principles of Modern Radar: Basic Principles*, Electromagnetics and Radar, Institution of Engineering and Technology.
- Ridley, A., Y. Deng, and G. Tóth (2006), The global ionosphere–thermosphere model, *Journal of Atmospheric and Solar-Terrestrial Physics*, *68*(8), 839 – 864, doi:http://dx.doi.org/10.1016/j.jastp.2006.01.008.
- Rihaczek, A. (1965), Radar signal design for target resolution, *Proceedings of the IEEE*, *53*(2), 116–128, doi:10.1109/PROC.1965.3572.
- Ruohoniemi, J. M., and K. B. Baker (1998a), Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations, *J. Geophys. Res.*, *103*(20), doi:10.1029/98JA01288.
- Ruohoniemi, J. M., and K. B. Baker (1998b), Large-scale imaging of high-latitude convection with super dual auroral radar network hf radar observations, *Journal of Geophysical Research: Space Physics*, *103*(A9), 20,797–20,811, doi:10.1029/98JA01288.
- Sampath, S. (2005), *Sampling Theory and Methods*, Alpha Science International.
- Schlegel, K. (1996), Coherent backscatter from ionospheric E-region plasma irregularities, *Journal of Atmospheric and Terrestrial Physics*, *58*, 933–941.
- Scott, C. J., R. G. Harrison, M. J. Owens, M. Lockwood, and L. Barnard (2014), Evidence for solar wind modulation of lightning, *Environmental Research Letters*, *9*(5), 055,004.
- Scoular, G., P. V. Ponomarenko, and J. St.-Maurice (2013), A new type of Doppler velocity fluctuations in HF ground scatter from the polar cap, *Geophysical Research Letters*, *40*(19), 4992–4997, doi:10.1002/grl.50960.
- Simon, M. (2007), *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*, International Series in Engineering and Computer Science, Springer US.
- Skolnik, M. I. (1980), *Introduction to Radar Systems*, McGraw-Hill.
- Sofko, G. J., R. Greenwald, and W. Bristow (1995), Direct determination of large-scale magnetospheric field-aligned currents with SuperDARN, *Geophysical Research Letters*, *22*(15), 2041–2044, doi:10.1029/95GL01317.
- Sojka, J., W. Raitt, and R. Schunk (1979), Effect of displaced geomagnetic and geographic poles on high-latitude plasma convection and ionospheric depletions, *Journal of Geophysical Research: Space Physics*, *84*(A10), 5943–5951, doi:10.1029/JA084iA10p05943.

- Spaleta, J., W. A. Bristow, and J. Klein (2015), Temporal and spatial resolved superdarn line of sight velocity measurements corrected for plasma index of refraction using bayesian inference, *Journal of Geophysical Research: Space Physics*, *120*(4), 3207–3225, doi:10.1002/2014JA020960, 2014JA020960.
- St.-Maurice, J.-P., and W. B. Hanson (1982), Ion frictional heating at high latitudes and its possible use for an in situ determination of neutral thermospheric winds and temperatures, *J. Geophys. Res.*, *87*(A9), 7580–7602, doi:10.1029/JA087iA09p07580.
- Sterne, K. T. (2010), Testing the re-designed SuperDARN HF radar and modeling of a twin terminated folded dipole array, Master’s thesis, Virginia Polytechnic Institute and State University.
- Sterne, K. T., R. A. Greenwald, J. B. H. Baker, and J. M. Ruohoniemi (2011), Modeling of a twin terminated folded dipole antenna for the Super Dual Auroral Radar Network (SuperDARN), in *2011 IEEE RadarCon (RADAR)*, pp. 934–938, doi:10.1109/RADAR.2011.5960673.
- Stix, M. (2003), On the time scale of energy transport in the sun, *Solar Physics*, *212*(1), 3–6.
- Stratton, J., and N. J. Fox (2012), Radiation belt storm probes (rbsp) mission overview, in *2012 IEEE Aerospace Conference*, pp. 1–10, doi:10.1109/AERO.2012.6187019.
- Sweet, P. A. (1958), The neutral point theory of solar flares, *Symposium - International Astronomical Union*, *6*, 123–134, doi:10.1017/S0074180900237704.
- SymPy Development Team (2014), *SymPy: Python library for symbolic mathematics*.
- Teramoto, M., N. Nishitani, Y. Nishimura, and T. Nagatsuma (2016), Latitudinal dependence on the frequency of pi2 pulsations near the plasmopause using themis satellites and asian-oceanian superdarn radars, *Earth, Planets and Space*, *68*(1), 22, doi:10.1186/s40623-016-0397-1.
- The Tesla Team (2016), Upgrading Autopilot: Seeing the World in Radar, <https://www.tesla.com/blog/upgrading-autopilot-seeing-world-radar>, accessed: 29 September, 2016.
- Theurer, T. E., and W. A. Bristow (2012), Observations and effects of artificial density layers on oblique high-frequency backscatter, *Radio Science*, *47*(2), doi:10.1029/2011RS004861, rS2010.
- Tokumaru, M., M. Kojima, and K. Fujiki (2010), Solar cycle evolution of the solar wind speed distribution from 1985 to 2008, *Journal of Geophysical Research: Space Physics*, *115*(A4), doi:10.1029/2009JA014628, a04102.
- Tsyganenko, N. A. (1995), Modeling the earth’s magnetospheric magnetic field confined within a realistic magnetopause, *Journal of Geophysical Research: Space Physics*, *100*(A4), 5599–5612, doi:10.1029/94JA03193.

- Turunen, T., J. Markkanen, and A. P. van Eyken (2000), Ground clutter cancellation in incoherent radars: solutions for eiscat svalbard radar, *Annales Geophysicae*, *18*(9), 1242–1247, doi:10.1007/s00585-000-1242-0.
- Ulaby, F. T., R. K. Moore, and A. K. Fung (1982), *Microwave Remote Sensing Active and Passive Volume 2*, Addison-Wesley Publishing Company.
- Vallières, X., J.-P. Villain, C. Hanuise, and R. André (2004), Ionospheric propagation effects on spectral widths measured by SuperDARN HF radars, *Annales Geophysicae*, *22*(6), 2023–2031, doi:10.5194/angeo-22-2023-2004.
- Villain, J. P., R. A. Greenwald, K. B. Baker, and J. M. Ruohoniemi (1987), Hf radar observations of e region plasma irregularities produced by oblique electron streaming, *Journal of Geophysical Research: Space Physics*, *92*(A11), 12,327–12,342, doi:10.1029/JA092iA11p12327.
- Villain, J. P., R. André, H. C., and D. Grésillon (1996), Observation of the high latitude ionosphere by {HF} radars: interpretation in terms of collective wave scattering and characterization of turbulence, *Journal of Atmospheric and Terrestrial Physics*, *58*(8-9), 943 – 958, doi:http://dx.doi.org/10.1016/0021-9169(95)00125-5.
- Wehner, D. R. (1987), *High Resolution Radar*, Artech House Inc.
- Westfall, P., and K. Henning (2013), *Understanding Advanced Statistical Methods*, Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis.
- Wild, J. A., S. E. Milan, C. J. Owen, J. M. Bosqued, M. Lester, D. M. Wright, H. Frey, C. W. Carlson, A. N. Fazakerley, and H. Rème (2004), The location of the open-closed magnetic field line boundary in the dawn sector auroral ionosphere, *Annales Geophysicae*, *22*(10), 3625–3639, doi:10.5194/angeo-22-3625-2004.
- Woodman, R. F., and T. Hagfors (1969), Methods for the measurement of vertical ionospheric motions near the magnetic equator by incoherent scattering, *Journal of Geophysical Research*, *74*(5), 1205–1212, doi:10.1029/JA074i005p01205.
- Woodward, P. (1960), *Probability and Information Theory: With Applications to Radar*, International series of monographs on electronics and instrumentation, Pergamon Press.
- Yeoman, T. K., G. Chisham, L. J. Baddeley, R. S. Dhillon, T. J. T. Karhunen, T. R. Robinson, A. Senior, and D. M. Wright (2008), Mapping ionospheric backscatter measured by the superdarn hf radars part 2: Assessing superdarn virtual height models, *Annales Geophysicae*, *26*(4), 843–852, doi:10.5194/angeo-26-843-2008.
- Yeoman, T. K., M. James, P. N. Mager, and D. Y. Klimushkin (2012), Superdarn observations of high-m ulf waves with curved phase fronts and their interpretation in terms of transverse resonator theory, *Journal of Geophysical Research: Space Physics*, *117*(A6), doi:10.1029/2012JA017668, a06231.

Zrnic, B., A. Zejak, and A. Petrovii (1999), Pulse compression radar: self-clutter suppression using modified rls algorithm, in *Telecommunications in Modern Satellite, Cable and Broadcasting Services, 1999. 4th International Conference on*, vol. 2, pp. 363–366 vol.2, doi:10.1109/TELSKS.1999.806231.

APPENDIX A

DISTRIBUTIONS DERIVED FROM GAUSSIAN RANDOM VARIABLES

If x is a Gaussian random variable, with mean \bar{x} , and variance σ^2 , it has a PDF given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \bar{x})^2}{2\sigma^2}\right). \quad (\text{A.1})$$

If x_1 and x_2 are zero mean ($\bar{x}_{1,2} = 0$) Gaussian random variables, then the sum of their squares, r ,

$$r = \sqrt{r_1^2 + r_2^2} \quad (\text{A.2})$$

is Rayleigh distributed. A Rayleigh distributed random variable, r , has a PDF given by

$$p(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0. \quad (\text{A.3})$$

The magnitude of complex voltage samples is Rayleigh distributed if the real and imaginary components are Gaussian random variables. The expectation value of r is $\sqrt{\frac{\pi}{2}}\sigma$.

The square of a Gaussian random variable, $y = x^2$, is a Chi-squared distributed random variable with a PDF given by

$$P(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y}{2\sigma^2}\right), \quad y \geq 0. \quad (\text{A.4})$$

A Chi-squared random variable has an expectation value of σ^2 . The Chi-squared distribution is central to the least-squares fitting techniques described in Chapter 7.

APPENDIX B

SELF-CLUTTER ESTIMATOR DERIVATIONS

B.1 Expected Maximum Magnitude of Self-Clutter

The expected value of any self-clutter estimator is given by

$$C = \frac{1}{K} \sum_{k=1}^K \left(\sum_{n=1}^N C_{n2} + \sum_{m=1}^M C_{1m} - \sum_{n=1}^N \sum_{m=1}^M C_{nm} \right), \quad (\text{B.1})$$

where C_{xy} denotes the expected value of the correlation between voltage samples \tilde{V}_x and \tilde{V}_y .

An estimate of the maximum magnitude of self-clutter contained in a lag estimate should have an expected value of

$$C_{\text{MAX}} = \sum_{n=1}^N |V_n| |V_R| + \sum_{m=1}^M |V_R| |V_m| + \sum_{n=1}^N \sum_{m=1}^M |V_n| |V_m|, \quad (\text{B.2})$$

or written in terms of expected power, assuming unit resistance,

$$C_{\text{MAX}} = \sum_{n=1}^N \sqrt{P_n P_R} + \sum_{m=1}^M \sqrt{P_R P_m} + \sum_{n=1}^N \sum_{m=1}^M \sqrt{P_n P_m} \quad (\text{B.3})$$

B.2 VSE

The voltage-based self-clutter estimator is written as

$$\hat{C}_{\text{VSE}} = \frac{1}{K} \sum_{k=1}^K \left(\sum_{n=1}^N \tilde{V}_{nk} \tilde{V}_{2k}^* + \sum_{m=1}^M \tilde{V}_{1k} \tilde{V}_{mk}^* - \sum_{n=1}^N \sum_{m=1}^M \tilde{V}_{nk} \tilde{V}_{mk}^* \right). \quad (\text{B.4})$$

with a MSE given by the

$$\text{MSE}(\hat{C}_{\text{VSE}}) = \frac{C^2}{K}. \quad (\text{B.5})$$

The MSE is equivalent to

$$\text{MSE}(\hat{C}_{\text{VSE}}) = \langle \hat{C}_{\text{VSE}}^2 \rangle - C^2. \quad (\text{B.6})$$

$\langle \hat{C}_{\text{VSE}}^2 \rangle$ requires a large but simple calculation. First we start with \hat{C}_{VSE}^2 and then we calculate the expected value of the result. As will be shown, the expected value calculation requires calculating 12 separate expected values.

$$\hat{C}_{\text{VSE}}^2 = \frac{1}{K} \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^K \left(\sum_{n=1}^N \tilde{V}_{nk} \tilde{V}_{2k}^* + \sum_{m=1}^M \tilde{V}_{1k} \tilde{V}_{mk}^* - \sum_{n=1}^N \sum_{m=1}^M \tilde{V}_{nk} \tilde{V}_{mk}^* \right)$$

$$\left(\sum_{n=1}^N \tilde{V}_{nl} \tilde{V}_{2l}^* + \sum_{m=1}^M \tilde{V}_{1l} \tilde{V}_{ml}^* - \sum_{n=1}^N \sum_{m=1}^M \tilde{V}_{nl} \tilde{V}_{ml}^* \right). \quad (\text{B.7})$$

$$\begin{aligned} \hat{C}_{\text{VSE}}^2 = & \frac{1}{K} \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^K \left(\sum_{n=1}^N \sum_{a=1}^N \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{2l}^* + \sum_{m=1}^M \sum_{b=1}^M \tilde{V}_{1k} \tilde{V}_{mk}^* \tilde{V}_{1l} \tilde{V}_{bl}^* \right. \\ & + 2 \sum_{n=1}^N \sum_{m=1}^M \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{1l} \tilde{V}_{ml}^* - 2 \sum_{n=1}^N \sum_{a=1}^N \sum_{m=1}^M \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{ml}^* \\ & \left. - 2 \sum_{n=1}^N \sum_{m=1}^M \sum_{b=1}^M \tilde{V}_{1l} \tilde{V}_{bl}^* \tilde{V}_{nk} \tilde{V}_{mk}^* + \sum_{n=1}^N \sum_{m=1}^M \sum_{a=1}^N \sum_{b=1}^M \tilde{V}_{nk} \tilde{V}_{mk}^* \tilde{V}_{al} \tilde{V}_{bl}^* \right) \end{aligned} \quad (\text{B.8})$$

The subscripts k and l select voltage samples from the the k and l rows of the \mathbf{V} matrix. Voltage samples in one row are statistically independent from voltage samples in another row. The expected value is equivalent to

$$\begin{aligned} \langle \hat{C}_{\text{VSE}}^2 \rangle = & \frac{1}{K} \left(\sum_{n=1}^N \sum_{a=1}^N \langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{2l}^* \rangle + \sum_{m=1}^M \sum_{b=1}^M \langle \tilde{V}_{1k} \tilde{V}_{mk}^* \tilde{V}_{1l} \tilde{V}_{bl}^* \rangle \right. \\ & + 2 \sum_{n=1}^N \sum_{m=1}^M \langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{1l} \tilde{V}_{ml}^* \rangle - 2 \sum_{n=1}^N \sum_{a=1}^N \sum_{m=1}^M \langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{ml}^* \rangle \\ & \left. - 2 \sum_{n=1}^N \sum_{m=1}^M \sum_{b=1}^M \langle \tilde{V}_{1l} \tilde{V}_{bl}^* \tilde{V}_{nk} \tilde{V}_{mk}^* \rangle + \sum_{n=1}^N \sum_{m=1}^M \sum_{a=1}^N \sum_{b=1}^M \langle \tilde{V}_{nk} \tilde{V}_{mk}^* \tilde{V}_{al} \tilde{V}_{bl}^* \rangle \right) \Big|_{k=l} \\ & + \frac{K-1}{K} \left(\sum_{n=1}^N \sum_{a=1}^N \langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{2l}^* \rangle + \sum_{m=1}^M \sum_{b=1}^M \langle \tilde{V}_{1k} \tilde{V}_{mk}^* \tilde{V}_{1l} \tilde{V}_{bl}^* \rangle \right. \\ & + 2 \sum_{n=1}^N \sum_{m=1}^M \langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{1l} \tilde{V}_{ml}^* \rangle - 2 \sum_{n=1}^N \sum_{a=1}^N \sum_{m=1}^M \langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{ml}^* \rangle \\ & \left. - 2 \sum_{n=1}^N \sum_{m=1}^M \sum_{b=1}^M \langle \tilde{V}_{1l} \tilde{V}_{bl}^* \tilde{V}_{nk} \tilde{V}_{mk}^* \rangle + \sum_{n=1}^N \sum_{m=1}^M \sum_{a=1}^N \sum_{b=1}^M \langle \tilde{V}_{nk} \tilde{V}_{mk}^* \tilde{V}_{al} \tilde{V}_{bl}^* \rangle \right) \Big|_{k \neq l} \end{aligned} \quad (\text{B.9})$$

The next step is to calculate the expected value of each of the 12 terms. The calculation is made simpler using a specific form of Isserlis' Theorem (*Isserlis*, 1918) given by *Reed* (1962) referred to as the fourth moment theorem,

$$\langle V_1 V_2 V_3 V_4 \rangle = \langle V_1 V_2 \rangle \langle V_3 V_4 \rangle + \langle V_1 V_3 \rangle \langle V_2 V_4 \rangle + \langle V_1 V_4 \rangle \langle V_2 V_3 \rangle. \quad (\text{B.10})$$

Then, assuming that correlations between voltage samples from differing multiple-pulse sequences are uncorrelated, we obtain

$$\langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{2l}^* \rangle = \begin{cases} 2C_{n2}C_{a2}, & k = l \\ C_{n2}C_{a2}, & k \neq l \end{cases} \quad (\text{B.11})$$

$$\langle \tilde{V}_{1k} \tilde{V}_{mk}^* \tilde{V}_{1l} \tilde{V}_{bl}^* \rangle = \begin{cases} 2C_{1m}C_{1b}, & k = l \\ C_{1m}C_{1b}, & k \neq l \end{cases} \quad (\text{B.12})$$

$$\langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{1l} \tilde{V}_{ml}^* \rangle = \begin{cases} C_{1m}C_{n2} + C_{nm}C_{12}, & k = l \\ C_{1m}C_{n2}, & k \neq l \end{cases} \quad (\text{B.13})$$

$$\langle \tilde{V}_{nk} \tilde{V}_{2k}^* \tilde{V}_{al} \tilde{V}_{ml}^* \rangle = \begin{cases} 2C_{n2}C_{am}, & k = l \\ C_{n2}C_{am}, & k \neq l \end{cases} \quad (\text{B.14})$$

$$\langle \tilde{V}_{1l} \tilde{V}_{bl}^* \tilde{V}_{nk} \tilde{V}_{mk}^* \rangle = \begin{cases} 2C_{nm}C_{1b}, & k = l \\ C_{nm}C_{1b}, & k \neq l \end{cases} \quad (\text{B.15})$$

$$\langle \tilde{V}_{nk} \tilde{V}_{mk}^* \tilde{V}_{al} \tilde{V}_{bl}^* \rangle = \begin{cases} 2C_{nm}C_{ab}, & k = l \\ C_{nm}C_{ab}, & k \neq l \end{cases} \quad (\text{B.16})$$

From the above relations we obtain

$$\langle \hat{C}_{\text{VSE}}^2 \rangle = \frac{2}{K}C^2 + \frac{K-1}{K}C^2 + \frac{2}{K} \sum_{n=1}^N \sum_{m=1}^M (C_{nm}C_{12} - C_{1m}C_{n2}) \quad (\text{B.17})$$

but $C_{nm}C_{12} - C_{1m}C_{n2} = 0$ since

$$\begin{aligned} C_{1m}C_{n2} &= V_1 V_m^* V_n V_2^* \\ &= V_1 V_2^* V_n V_m^* \\ &= C_{12}C_{nm} \end{aligned} \quad (\text{B.18})$$

which is to say that the correlations C_{nm} and C_{12} are each products of two complex numbers and that products of complex numbers are commutative. With this relation, we finally obtain

$$\text{MSE}(\hat{C}_{\text{VSE}}) = \frac{C^2}{K}. \quad (\text{B.19})$$

B.3 MVSE and MPSE

The maximum voltage-based self-clutter estimator is given by the relation

$$\hat{C}_{MVSE} = \frac{4}{\pi K} \sum_{k=1}^K \left(\sum_{n=1}^N |\tilde{V}_{nk}| |\tilde{V}_{Rk}| + \sum_{m=1}^M |\tilde{V}_{Rk}| |\tilde{V}_{mk}| + \sum_{n=1}^N \sum_{m=1}^M |\tilde{V}_{nk}| |\tilde{V}_{mk}| \right). \quad (\text{B.20})$$

The voltage samples are zero mean complex Gaussian random variables, which means that the magnitude of each voltage sample follows a Rayleigh distribution of order 2 (*Simon, 2007*). The expected value of the magnitude of a voltage sample is then given by

$$E[|\tilde{V}|] = \frac{\pi}{4}|V| \quad (\text{B.21})$$

where $|V|$ is the population mean value of the magnitude of the voltage sample. This relation shows why a constant factor of $4/\pi$ is used to ensure that the expected value of $\hat{C}_{MVSE} = C_{\text{MAX}}$ (e.g. *Sampath, 2005*).

The maximum power-based self-clutter is derived from the maximum voltage-based self-clutter estimator using Cauchy's inequality (*Abramowitz and Stegun*, 2012)

$$\left(\sum a_n b_n\right)^2 \leq \sum a_n^2 \sum b_n^2. \quad (\text{B.22})$$

Applying the inequality to the terms in the maximum voltage-based self-clutter estimator and taking the square root we obtain

$$\left(\sum_{k=1}^K |\tilde{V}_{nk}| |\tilde{V}_{mk}|\right) \leq \sqrt{\sum_{k=1}^K |\tilde{V}_{nk}|^2 \sum_{k=1}^K |\tilde{V}_{mk}|^2}. \quad (\text{B.23})$$

The terms inside the square root can be replaced with the power estimator in Chapter 5 (equation 5.3). The maximum power-based self-clutter estimator is then given by the relation

$$\hat{C}_{MPSE} = \sum_{n=1}^N \sqrt{\hat{P}_n \hat{P}_R} + \sum_{m=1}^M \sqrt{\hat{P}_R \hat{P}_m} + \sum_{n=1}^N \sum_{m=1}^M \sqrt{\hat{P}_n \hat{P}_m}, \quad (\text{B.24})$$

and relies on the power estimator. Calculation of the MSE in the MVSE and MPSE is accomplished in a similar manner as for the MSE in the VSE. For the MVSE the following relations were used,

$$\langle |\tilde{V}_{nk}| |\tilde{V}_{mk}| |\tilde{V}_{al}| |\tilde{V}_{bl}| \rangle = \begin{cases} \sqrt{P_n P_m P_a P_b}, & n = a, b = m, k = l \\ (\pi/4) \sqrt{P_n P_m P_a P_b}, & n = a, b \neq m, k = l \\ (\pi/4) \sqrt{P_n P_m P_a P_b}, & n \neq a, b = m, k = l \\ (\pi/4)^2 \sqrt{P_n P_m P_a P_b}, & n \neq a, b \neq m, k = l \\ (\pi/4)^2 \sqrt{P_n P_m P_a P_b}, & k \neq l \end{cases}. \quad (\text{B.25})$$

Here we have used the fact that for $n \neq a$ and $m \neq b$, voltage samples are statistically independent and

$$\langle |\tilde{V}_{nk}| |\tilde{V}_{mk}| |\tilde{V}_{al}| |\tilde{V}_{bl}| \rangle = \langle |\tilde{V}_{nk}| \rangle \langle |\tilde{V}_{mk}| \rangle \langle |\tilde{V}_{al}| \rangle \langle |\tilde{V}_{bl}| \rangle, \quad (\text{B.26})$$

and that for $n = a$ or $m = b$

$$\langle |\tilde{V}_n| |\tilde{V}_a| \rangle = \langle |\tilde{V}_n|^2 \rangle = P_n. \quad (\text{B.27})$$

Also, recall that voltage samples \tilde{V}_n and \tilde{V}_m are statistically independent.

Calculation of the MSE in the MPSE was based on the fact that

$$\left\langle \sqrt{\hat{P}_n \hat{P}_a} \right\rangle = \sqrt{\langle \hat{P}_n \hat{P}_a \rangle} \quad (\text{B.28})$$

with

$$\langle \hat{P}_n \hat{P}_a \rangle = \langle \hat{P}_n \rangle \langle \hat{P}_a \rangle \quad (\text{B.29})$$

for $n \neq a$ and that

$$\langle \hat{P}_n \hat{P}_a \rangle = \langle \hat{P}_n^2 \rangle = P_n^2 \left(1 + \frac{1}{K}\right). \quad (\text{B.30})$$

Finally, we can write a MSE for the MxPE (where x is either V or P for either voltage-based or power-based, respectively) as

$$\begin{aligned}
MSE(\hat{C}_{\text{MxSE}}) = & \left(\sum_{n=1}^N \sum_{a=1}^N \alpha_{na} P_R \sqrt{P_n P_a} + \sum_{m=1}^M \sum_{b=1}^M \alpha_{mb} P_R \sqrt{P_m P_b} \right. \\
& + 2 \sum_{n=1}^N \sum_{m=1}^M \alpha_{n0} P_R \sqrt{P_n P_m} + 2 \sum_{n=1}^N \sum_{a=1}^N \sum_{m=1}^M \beta_{na} \sqrt{P_n P_a P_m P_R} \\
& \left. + 2 \sum_{n=1}^N \sum_{m=1}^M \sum_{b=1}^M \beta_{mb} \sqrt{P_n P_m P_b P_R} \sum_{n=1}^N \sum_{a=1}^N \sum_{m=1}^M \sum_{b=1}^M \gamma_{namb} \sqrt{P_n P_a P_m P_b} \right). \quad (\text{B.31})
\end{aligned}$$

The coefficients for MSE of the voltage-based maximal self-clutter estimator are,

$$\begin{aligned}
\alpha_{na} &= \begin{cases} (1/K) ((4/\pi)^2 - 1) & n = a \\ (1/K) ((4/\pi) - 1) & n \neq a \end{cases} \\
\beta_{na} &= \begin{cases} (1/K) ((4/\pi) - 1) & n = a \\ 0 & n \neq a \end{cases} \\
\gamma_{nmab} &= \begin{cases} (1/K) ((4/\pi)^2 - 1) & n = a, m = b \\ (1/K) ((4/\pi) - 1) & n \neq a, m = b \\ 0 & n = a, m \neq b \\ 0 & n \neq a, m \neq b \end{cases}, \quad (\text{B.32})
\end{aligned}$$

and the coefficients for MSE of the power-based maximal self-clutter estimator are

$$\begin{aligned}
\alpha_{na} &= \begin{cases} 1/K & n = a \\ \sqrt{1 + 1/K} - 1 & n \neq a \end{cases} \\
\beta_{na} &= \begin{cases} \sqrt{1 + 1/K} - 1 & n = a \\ 0 & n \neq a \end{cases} \\
\gamma_{nmab} &= \begin{cases} 1/K & n = a, m = b \\ \sqrt{1 + 1/K} - 1 & n \neq a, m = b \\ 0 & n = a, m \neq b \\ 0 & n \neq a, m \neq b \end{cases}. \quad (\text{B.33})
\end{aligned}$$

APPENDIX C

MISCELLANEOUS DERIVATIONS

C.1 Derivation of Probability Distribution Functions of the Magnitude and Phase of the Mean Auto-Correlation Function

Below we present our derivation of the distributions of the magnitude (or envelope) and phase of Equation 6.13. It is worth noting that results in different forms have been obtained for the magnitude distribution by *Beckmann* (1962) (only for the uncorrelated case), *Aalo et al.* (2007), and *Dharmawansa et al.* (2009). The polar form of Equation 6.13 is given by

$$p(r, \theta) = \frac{r}{2\pi\sigma_r\sigma_i\sqrt{1-\rho_{ri}^2}} e^{-\frac{(r\cos(\theta)-\bar{x})^2}{2\sigma_r^2(1-\rho_{ri}^2)} - \frac{(r\sin(\theta)-\bar{y})^2}{2\sigma_i^2(1-\rho_{ri}^2)} + \frac{\rho_{ri}(r\cos(\theta)-\bar{x})(r\sin(\theta)-\bar{y})}{\sigma_r\sigma_i(1-\rho_{ri}^2)}}. \quad (\text{C.1})$$

To evaluate the integral for the phase distribution, as discussed in Section 6.5.2, we rearrange Equation C.1 to take advantage of the fact that

$$\int_0^\infty r e^{-ar^2+br+c} dr = \frac{e^c}{4\sqrt{a}} \left(\frac{b\sqrt{\pi}}{\sqrt{a}} \exp\left(\frac{b^2}{4a}\right) \left(1 + \operatorname{erf}\left(\frac{b}{2\sqrt{a}}\right)\right) + 2 \right). \quad (\text{C.2})$$

This result can be derived using the symbolic mathematics Python package SymPy (*SymPy Development Team*, 2014). Finally, only basic algebraic manipulations are required to obtain

$$p(\theta) = \frac{C}{\sqrt{A}} \left(\frac{B\sqrt{\pi}}{\sqrt{A}} \exp\left(\frac{B^2}{4A}\right) \left(1 + \operatorname{erf}\left(\frac{B}{2\sqrt{A}}\right)\right) + 2 \right) \quad (\text{C.3})$$

where A , B , and C are given by

$$A = \frac{\cos^2(\theta)}{2\sigma_r^2(1-\rho_{ri}^2)} + \frac{\sin^2(\theta)}{2\sigma_i^2(1-\rho_{ri}^2)} - \frac{\rho_{ri}\cos(\theta)\sin(\theta)}{\sigma_r\sigma_i(1-\rho_{ri}^2)} \quad (\text{C.4})$$

$$B = \frac{1}{\sigma_r(1-\rho_{ri}^2)} \left[\frac{\bar{x}}{\sigma_r} - \frac{\rho_{ri}\bar{y}}{\sigma_i} \right] \cos(\theta) + \frac{1}{\sigma_i(1-\rho_{ri}^2)} \left[\frac{\bar{y}}{\sigma_i} - \frac{\rho_{ri}\bar{x}}{\sigma_r} \right] \sin(\theta) \quad (\text{C.5})$$

$$C = \frac{\exp\left(\frac{\bar{x}^2}{2\sigma_r^2(1-\rho_{ri}^2)} + \frac{\bar{y}^2}{2\sigma_i^2(1-\rho_{ri}^2)} - \frac{\rho_{ri}\bar{x}\bar{y}}{\sigma_r\sigma_i(1-\rho_{ri}^2)}\right)}{4\pi\sigma_r\sigma_i(1-\rho_{ri}^2)}. \quad (\text{C.6})$$

which is a result similar in form to that presented in *Aalo et al.* (2007).

Evaluating the integral for the magnitude distribution is less trivial. First, we rearrange Equation C.1 into the following form (where we have applied double-angle trigonometric identities)

$$u(A, P, Q, R, H) = A e^{-\frac{P}{2}} \int_0^{2\pi} e^{-\frac{P}{2} \cos(2\theta) + Q \cos \theta + R \sin \theta + \frac{H}{2} \sin(2\theta)} d\theta. \quad (\text{C.7})$$

Next, the appropriate terms in the integral are substituted with the following trigonometric and Bessel function identities (from *Abramowitz and Stegun* (2012)),

$$a \cos(\omega t) + b \sin(\omega t) = \sqrt{a^2 + b^2} \cos\left(\omega t + \tan^{-1}\left(\frac{b}{a}\right)\right) \quad (\text{C.8})$$

$$e^{z \cos x} = I_0(z) + 2 \sum_{k=1}^{\infty} I_k(z) \cos(kx) \quad (\text{C.9})$$

$$e^{z \sin x} = I_0(z) + 2 \sum_{k=0}^{\infty} (-1)^k I_{2k+1}(z) \sin([2k+1]x) + 2 \sum_{k=1}^{\infty} (-1)^k I_{2k}(z) \cos(2kx), \quad (\text{C.10})$$

to obtain

$$\begin{aligned} u(A, P, Q, R, H) = A e^{-\frac{P}{2}} \int_0^{2\pi} \left[I_0\left(\frac{P}{2}\right) + 2 \sum_{k=1}^{\infty} I_k\left(\frac{P}{2}\right) \cos(2k\theta) \right] \\ \left[I_0\left(\frac{H}{2}\right) + 2 \sum_{m=0}^{\infty} (-1)^m I_{2m+1}\left(\frac{H}{2}\right) \sin([4m+2]\theta) + \right. \\ \left. 2 \sum_{m=1}^{\infty} (-1)^m I_{2m}\left(\frac{H}{2}\right) \cos(4m\theta) \right] e^{\sqrt{Q^2+R^2} \cos(\theta + \tan^{-1}(\frac{R}{Q}))} d\theta. \quad (\text{C.11}) \end{aligned}$$

Finally, by expanding and utilizing angle sum and difference trigonometric identities, and

$$\int_0^{2\pi} \cos(nx) e^{z \cos x} dx = 2\pi I_n(z) \quad (\text{C.12})$$

we obtain

$$\begin{aligned} p(r) = \frac{r e^{-P/2-D}}{\sigma_r \sigma_i \sqrt{1-\rho_{ri}^2}} \left[I_0\left(\frac{P}{2}\right) I_0\left(\frac{H}{2}\right) I_0\left(\sqrt{Q^2+R^2}\right) + \right. \\ 2 I_0\left(\frac{H}{2}\right) \sum_{k=1}^{\infty} (-1)^k I_k\left(\frac{P}{2}\right) I_{2k}\left(\sqrt{Q^2+R^2}\right) \cos(2k\phi) + \\ 2 I_0\left(\frac{P}{2}\right) \sum_{m=1}^{\infty} (-1)^m I_{2m}\left(\frac{H}{2}\right) I_{4m}\left(\sqrt{Q^2+R^2}\right) \cos(4m\phi) + \\ 2 \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} (-1)^k (-1)^m I_k\left(\frac{P}{2}\right) I_{2m}\left(\frac{H}{2}\right) \left[I_{2k-4m}\left(\sqrt{Q^2+R^2}\right) \cos((2k-4m)\phi) + \right. \\ \left. I_{2k+4m}\left(\sqrt{Q^2+R^2}\right) \cos((2k+4m)\phi) \right] \left. \right] \quad (\text{C.13}) \end{aligned}$$

where I_n is the modified Bessel function of the first kind and coefficients D, H, P, Q, and R are given by

$$D = \frac{1}{1-\rho_{ri}^2} \left(\frac{\bar{x}^2}{2\sigma_r^2} + \frac{\bar{y}^2}{2\sigma_i^2} + \frac{r^2}{2\sigma_i^2} - \frac{\rho_{ri}\bar{x}\bar{y}}{\sigma_r\sigma_i} \right) \quad (\text{C.14})$$

$$H = \frac{r^2}{1-\rho_{ri}^2} \left(\frac{\rho_{ri}}{\sigma_r\sigma_i} \right) \quad (\text{C.15})$$

$$P = \frac{r^2}{1-\rho_{ri}^2} \left(\frac{1}{2\sigma_r^2} - \frac{1}{2\sigma_i^2} \right) \quad (\text{C.16})$$

$$Q = \frac{1}{1-\rho_{ri}^2} \left(\frac{\bar{x}r}{\sigma_r^2} - \frac{\rho_{ri}\bar{y}r}{\sigma_r\sigma_i} \right) \quad (\text{C.17})$$

$$R = \frac{1}{1-\rho_{ri}^2} \left(\frac{\bar{y}r}{\sigma_i^2} - \frac{\rho_{ri}\bar{x}r}{\sigma_r\sigma_i} \right). \quad (\text{C.18})$$

C.2 Reduced Correlation in the Presence of Noise and Clutter

To obtain Equation 7.8 presented in Section 7.5, the total voltage received by SuperDARN radars, $V_T(t)$ is modelled as the sum of the received ionospheric signal voltage, $V_S(t)$, and a contribution to the voltage by noise, $V_N(t)$:

$$V_T(t) = V_S(t) + V_N(t). \quad (\text{C.19})$$

The noise and ionospheric signal are assumed to be statically independent. The ACF is estimated using Equation 7.1. Using equations 7.1 and C.19, $\hat{R}(\tau_i)$ can be written as

$$\hat{R}_T(\tau_i) = \hat{R}_S(\tau_i) + \hat{R}_N(\tau_i), \quad (\text{C.20})$$

where $\hat{R}_S(\tau_i)$ and $\hat{R}_N(\tau_i)$ are the ionospheric signal and noise ACFs, respectively. The received ionospheric signal voltage, $V_S(t)$, produced by the multiple-pulse technique contains both signal and self-clutter. Using Equation 8 from *Reimer and Hussey* (2015), then we can rewrite $V_S(t)$ as

$$V_S(t) = V_R(t) + \sum V_i(t) \quad (\text{C.21})$$

such that the ionospheric signal voltage is the sum of the signal from the range for which we wish to determine the ACF, V_R , and the sum of ionospheric signal voltages from interfering ranges, V_i . Rewriting $\hat{R}_S(t)$ as in Equation 9 in *Reimer and Hussey* (2015) gives

$$\hat{R}_S(t) = \hat{R}_R(t) + \hat{R}_C(t), \quad (\text{C.22})$$

where $\hat{R}_R(t)$ is the ACF of the ionosphere at the range of interest and $\hat{R}_C(t)$ is the self-clutter ACF. The estimate of the ACF from the measured voltage signal $V_S(t)$ is the sum of the ACF at the range of interest, $\hat{R}_R(t)$, and the ACF due to self-clutter, $\hat{R}_C(t)$.

Following the discussion in section 8.4 of *Bendat and Piersol* (2000), we obtain

$$\hat{\rho}_T = \frac{\hat{R}_S(t) + \hat{R}_N(t)}{(\hat{R}_S(0) + \hat{R}_N(0))}, \quad (\text{C.23})$$

which is the estimate of total normalized correlation coefficient, $\hat{\rho}_T$. Equation C.23 is similar to Equation 8.110 in *Bendat and Piersol* (2000) except that we have used the notation $\hat{R} = \hat{R}_{xx} = \hat{R}_{yy}$. Substituting Equation C.23 in to Equation C.22 and taking the population limit (number of samples becomes very large), we obtain

$$\rho_T = \frac{R_R(t) + R_N(t) + R_C(t)}{(R_R(0) + R_N(0) + R_C(0))} \quad (\text{C.24})$$

which is the total normalized correlation coefficient in the presence of noise and clutter. Since $R_R(\tau_i) = P\rho(\tau_i)$, $R_R(0) = P$, $R_N(0) = N$, and $R_C(0) = C$, the correlation coefficient of the *signal* at the range of interest, ρ' , is

$$\rho' = \frac{P\rho}{(P + N + C)}. \quad (\text{C.25})$$

APPENDIX D

REFERENCE MATERIALS

D.1 SuperDARN IQDAT File Contents

D.1.1 Scalars

Variable Name	Data Type	Description
radar.revision.major	char	Major version number of the radar operating system.
radar.revision.minor	char	Minor version number of the radar operating system.
origin.code	char	Code indicating origin of the data.
origin.time	string	ASCII representation of when the data was generated.
origin.command	string	The command line or control program used to generate the data.
cp	short	Control program identifier.
stid	short	Station identifier.
time.yr	short	Year.
time.mo	short	Month.
time.dy	short	Day.
time.hr	short	Hour.
time.mt	short	Minute.
time.sc	short	Second.
time.us	short	Micro-Second.
txpow	short	Transmitted power (kW).
nave	short	Number of pulse sequences transmitted.
atten	short	Attenuation level.
lagfr	short	Lag to first range (microseconds).
smsep	short	Sample separation (microseconds).
ercod	short	Error code.
stat.agc	short	AGC status word.
stat.lopwr	short	LOPWR status word.
noise.search	float	Calculated noise from clear frequency search.
noise.mean	float	Average noise across frequency band.
channel	short	Channel number for a stereo radar (zero for all others).
bmnum	short	Beam number.
bmazm	float	Beam azimuth.
scan	short	Scan flag.

offset	short	Offset between channels for a stereo radar (zero for all others).
rxrise	short	Receiver rise time (microseconds).
intt.sc	short	Whole number of seconds of integration time.
intt.us	short	Fractional number of microseconds of integration time.
txpl	short	Transmit pulse length (microseconds).
mpinc	short	Multi-pulse increment (microseconds).
mppul	short	Number of pulses in sequence.
mplgs	short	Number of lags in sequence.
nrang	short	Number of ranges.
frang	short	Distance to first range (kilometers).
rsep	short	Range separation (kilometers).
xcf	short	XCF flag.
tfreq	short	Transmitted frequency.
mxpwr	int	Maximum power (kHz).
lvmax	int	Maximum noise level allowed.
iqdata.revision.major	int	Major version number of the iqdata library.
iqdata.revision.minor	int	Minor version number of the iqdata library.
combf	string	Comment buffer.
seqnum	int	Number of pulse sequences transmitted.
chnum	int	Number of channels sampled (both I and Q quadrature samples).
smpnum	int	Number of samples taken per sequence.
skpnum	int	Number of samples to skip before the first valid sample.

D.1.2 Arrays

Variable Name	Data Type	Dimensions	Description
ptab[mppul]	short	mppul	Pulse table.
ltab[2][mplgs]	short	2,mplgs	Lag table.
tsc[seqnum]	int	seqnum	Seconds component of time past epoch of pulse sequence.
tus[seqnum]	int	seqnum	Microsecond component of time past epoch of pulse sequence.
tatten[seqnum]	short	seqnum	Attenuator setting for each pulse sequence.
tnoise[seqnum]	float	seqnum	Noise value for each pulse sequence.
toff[seqnum]	int	seqnum	Offset into the sample buffer for each pulse sequence.
tsze[seqnum]	int	seqnum	Number of words stored for this pulse sequence.
data[totnum]	int	totnum	Array of raw I and Q samples.

D.2 SuperDARN RAWACF File Contents

D.2.1 Scalars

Variable Name	Data Type	Description
radar.revision.major	char	Major version number of the radar operating system.
radar.revision.minor	char	Minor version number of the radar operating system.
origin.code	char	Code indicating origin of the data.
origin.time	string	ASCII representation of when the data was generated.
origin.command	string	The command line or control program used to generate the data.
cp	short	Control program identifier.
stid	short	Station identifier.
time.yr	short	Year.
time.mo	short	Month.
time.dy	short	Day.
time.hr	short	Hour.
time.mt	short	Minute.
time.sc	short	Second.
time.us	short	Micro-Second.
txpow	short	Transmitted power (kW).
nave	short	Number of pulse sequences transmitted.
atten	short	Attenuation level.
lagfr	short	Lag to first range (microseconds).
smsep	short	Sample separation (microseconds).
ercod	short	Error code.
stat.agc	short	AGC status word.
stat.lopwr	short	LOPWR status word.
noise.search	float	Calculated noise from clear frequency search.
noise.mean	float	Average noise across frequency band.
channel	short	Channel number for a stereo radar (zero for all others).
bmnum	short	Beam number.
bmazm	float	Beam azimuth.
scan	short	Scan flag.
offset	short	Offset between channels for a stereo radar (zero for all others).
rxrise	short	Receiver rise time (microseconds).
intt.sc	short	Whole number of seconds of integration time.
intt.us	short	Fractional number of microseconds of integration time.

txpl	short	Transmit pulse length (microseconds).
mpinc	short	Multi-pulse increment (microseconds).
mppul	short	Number of pulses in sequence.
mplgs	short	Number of lags in sequence.
nrang	short	Number of ranges.
frang	short	Distance to first range (kilometers).
rsep	short	Range separation (kilometers).
xcf	short	XCF flag.
tfreq	short	Transmitted frequency.
mxpwr	int	Maximum power (kHz).
lvmax	int	Maximum noise level allowed.
rawacf.revision.major	int	Major version number of the rawacf format.
rawacf.revision.minor	int	Minor version number of the rawacf format.
combf	string	Comment buffer.
thr	float	Thresholding factor.

D.2.2 Arrays

Variable Name	Data Type	Dimensions	Description
ptab[mppul]	short	mppul	Pulse table.
ltab[2][mplgs]	short	2,mplgs	Lag table.
pwr0[nrng]	float	nrng	Lag zero power.
slist[0-nrng]	short	0-nrng	List of stored ranges.
acfd[2][mplgs][0-nrng]	short	2,mplgs,0-nrng	Calculated ACFs.
xcfd[2][mplgs][0-nrng]	short	2,mplgs,0-nrng	Calculated XCFs.

D.3 SuperDARN FITACF File Contents

D.3.1 Scalars

Variable Name	Data Type	Description
radar.revision.major	char	Major version number of the radar operating system.
radar.revision.minor	char	Minor version number of the radar operating system.
origin.code	char	Code indicating origin of the data.
origin.time	string	ASCII representation of when the data was generated.
origin.command	string	The command line or control program used to generate the data.
cp	short	Control program identifier.
stid	short	Station identifier.

time.yr	short	Year.
time.mo	short	Month.
time.dy	short	Day.
time.hr	short	Hour.
time.mt	short	Minute.
time.sc	short	Second.
time.us	short	Micro-Second.
txpow	short	Transmitted power (kW).
nave	short	Number of pulse sequences transmitted.
atten	short	Attenuation level.
lagfr	short	Lag to first range (microseconds).
smsep	short	Sample separation (microseconds).
ercod	short	Error code.
stat.agc	short	AGC status word.
stat.lopwr	short	LOPWR status word.
noise.search	float	Calculated noise from clear frequency search.
noise.mean	float	Average noise across frequency band.
channel	short	Channel number for a stereo radar (zero for all others).
bmnum	short	Beam number.
bmazm	float	Beam azimuth.
scan	short	Scan flag.
offset	short	Offset between channels for a stereo radar (zero for all others).
rxrise	short	Receiver rise time (microseconds).
intt.sc	short	Whole number of seconds of integration time.
intt.us	short	Fractional number of microseconds of integration time.
txpl	short	Transmit pulse length (microseconds).
mpinc	short	Multi-pulse increment (microseconds).
mppul	short	Number of pulses in sequence.
mplgs	short	Number of lags in sequence.
nrang	short	Number of ranges.
frang	short	Distance to first range (kilometers).
rsep	short	Range separation (kilometers).
xcf	short	XCF flag.
tfreq	short	Transmitted frequency.
mxpwr	int	Maximum power (kHz).
lvmax	int	Maximum noise level allowed.
fitacf.revision.major	int	Major version number of the FitACF algorithm.
fitacf.revision.minor	int	Minor version number of the FitACF algorithm.
combf	string	Comment buffer.
noise.sky	float	Sky noise.
noise.lag0	float	Lag zero power of noise ACF.
noise.vel	float	Velocity from fitting the noise noise ACF.

D.3.2 Arrays

Variable Name	Data Type	Dimensions	Description
ptab[mppul]	short	mppul	Pulse table.
ltab[2][mplgs]	short	2,mplgs	Lag table.
pwr0[nrng]	float	nrng	Lag zero power.
slist[0-nrng]	short	0-nrng	List of stored ranges.
nlag[0-nrng]	short	0-nrng	Number of points in the fit.
qflg[0-nrng]	char	0-nrng	Quality of fit flag for ACF.
gflg[0-nrng]	char	0-nrng	Ground scatter flag for ACF.
p_l[0-nrng]	float	0-nrng	Power from lambda fit of ACF.
p_l_e[0-nrng]	float	0-nrng	Power error from lambda fit of ACF.
p_s[0-nrng]	float	0-nrng	Power from sigma fit of ACF..
p_s_e[0-nrng]	float	0-nrng	Powr error from sigma fit of ACF.
v[0-nrng]	float	0-nrng	Velocity from ACF.
v_e[0-nrng]	float	0-nrng	Velocity error from ACF.
w_l[0-nrng]	float	0-nrng	Spectral width from lambda fit of ACF.
w_l_e[0-nrng]	float	0-nrng	Spectral width error from lambda fit of ACF.
w_s[0-nrng]	float	0-nrng	Spectral width from sigma fit of ACF.
w_s_e[0-nrng]	float	0-nrng	Spectral width error from sigma fit of ACF.
sd_l[0-nrng]	float	0-nrng	Standard deviation of sigma fit.
sd_s[0-nrng]	float	0-nrng	Standard deviation of lambda fit.
sd_phi[0-nrng]	float	0-nrng	Standard deviation of phase fit of ACF.
x_qflg[0-nrng]	char	0-nrng	Quality of fit flag for XCF.
x_gflg[0-nrng]	char	0-nrng	Ground scatter flag for XCF.
x_p_l[0-nrng]	float	0-nrng	Power from lambda fit of XCF.
x_p_l_e[0-nrng]	float	0-nrng	Power error from lambda fit of XCF.
x_p_s[0-nrng]	float	0-nrng	Power from sigma fit of XCF.
x_p_s_e[0-nrng]	float	0-nrng	Power error from sigma fit of XCF.
x_v[0-nrng]	float	0-nrng	Velocity from XCF.
x_v_e[0-nrng]	float	0-nrng	Velocity error from XCF.
x_w_l[0-nrng]	float	0-nrng	Spectral width from lambda fit of XCF.
x_w_l_e[0-nrng]	float	0-nrng	Spectral width error from lambda fit of XCF.
x_w_s[0-nrng]	float	0-nrng	Spectral width from sigma fit of XCF.
x_w_s_e[0-nrng]	float	0-nrng	Spectral width error from sigma fit of XCF.

phi0[0-nrng]	float	0-nrng	Phase determination at lag zero of the ACF.
phi0_e[0-nrng]	float	0-nrng	Phase determination error at lag zero of the ACF.
elv[0-nrng]	float	0-nrng	Angle of arrival estimate.
elv_low[0-nrng]	float	0-nrng	Lowest estimate of angle of arrival.
elv_high[0-nrng]	float	0-nrng	Highest estimate of angle of arrival.
x_sd_l[0-nrng]	float	0-nrng	Standard deviation of lambda fit of XCF.
x_sd_s[0-nrng]	float	0-nrng	Standard deviation of sigma fit of XCF.
x_sd_phi[0-nrng]	float	0-nrng	Standard deviation of phase fit of XCF.

APPENDIX E

COPYRIGHT INFORMATION

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

Nov 26, 2017

This Agreement between Mr. Ashton Reimer ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4236351394851
License date	Nov 26, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Radio Science
Licensed Content Title	Estimating self-clutter of the multiple-pulse technique
Licensed Content Author	A. S. Reimer,G. C. Hussey
Licensed Content Date	Jul 24, 2015
Licensed Content Pages	14
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Improved SuperDARN radar signal processing: A first principles statistical approach for reliable measurement uncertainties and enhanced data products
Expected completion date	Feb 2018
Expected size (number of pages)	230
Requestor Location	Mr. Ashton Reimer 116 Science Place Saskatoon, SK S7N 5E2 Canada Attn: Mr. Ashton Reimer
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	Mr. Ashton Reimer 116 Science Place Saskatoon, SK S7N 5E2 Canada Attn: Mr. Ashton Reimer
Total	0.00 CAD
Terms and Conditions	

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

Nov 26, 2017

This Agreement between Mr. Ashton Reimer ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4236360133795
License date	Nov 26, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Radio Science
Licensed Content Title	On the statistics of SuperDARN autocorrelation function estimates
Licensed Content Author	A. S. Reimer,G. C. Hussey,S. R. Dueck
Licensed Content Date	Jun 18, 2016
Licensed Content Pages	14
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Improved SuperDARN radar signal processing: A first principles statistical approach for reliable measurement uncertainties and enhanced data products
Expected completion date	Feb 2018
Expected size (number of pages)	230
Requestor Location	Mr. Ashton Reimer 116 Science Place Saskatoon, SK S7N 5E2 Canada Attn: Mr. Ashton Reimer
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	Mr. Ashton Reimer 116 Science Place Saskatoon, SK S7N 5E2 Canada Attn: Mr. Ashton Reimer
Total	0.00 CAD
Terms and Conditions	

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

Feb 25, 2018

This Agreement between Mr. Ashton Reimer ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4295770218716
License date	Feb 25, 2018
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Radio Science
Licensed Content Title	Statistically Self-Consistent and Accurate Errors for SuperDARN Data
Licensed Content Author	A. S. Reimer,G. C. Hussey,K. A. McWilliams
Licensed Content Date	Jan 21, 2018
Licensed Content Pages	19
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Improved SuperDARN radar signal processing: A first principles statistical approach for reliable measurement uncertainties and enhanced data products
Expected completion date	Feb 2018
Expected size (number of pages)	230
Requestor Location	Mr. Ashton Reimer 116 Science Place Saskatoon, SK S7N 5E2 Canada Attn: Mr. Ashton Reimer
Publisher Tax ID	EU826007151
Total	0.00 CAD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms