

# Thesis Title

by

E. Ross



A thesis submitted to the  
University of Birmingham  
for the degree of  
DOCTOR OF PHILOSOPHY

Solar and Stellar Physics Group (SASP)  
School of Physics and Astronomy  
University of Birmingham  
Birmingham, B15 2TT  
Month 20XX

# Abstract

You're awesome. Make sure the examiners know it

# Acknowledgements

Who do you hate least?

# Contents

<b>Abstract</b>	i
<b>Acknowledgments</b>	ii
<b>List of Figures</b>	xii
<b>List of Tables</b>	xiv
<b>List of Abbreviations</b>	xvi
<b>1 HiSPARC as a Space Weather Detector</b>	1
1.1 Introduction . . . . .	1
1.1.1 Space Weather Effects . . . . .	2
1.1.2 HiSPARC Project . . . . .	2
1.1.3 HiSPARC Detector and Station Configuration . . . . .	2
1.2 Aims . . . . .	7
1.3 HiSPARC Properties . . . . .	8
1.4 HiSPARC Observations . . . . .	11
1.4.1 HiSPARC Observations of Ground Level Enhancements . . . . .	12
1.4.2 HiSPARC Observations of Forbush Decreases . . . . .	16
1.5 Air Shower Simulations . . . . .	19
1.5.1 Air Shower Footprints . . . . .	19
1.5.2 Muon Flux . . . . .	20
1.6 Standardisation of HiSPARC Data . . . . .	20
1.6.1 Motivation . . . . .	20
1.6.2 Barometric Correction . . . . .	21
1.6.3 Temperature Correction . . . . .	24
1.7 HiSPARC Observations After Pressure Corrections . . . . .	26
1.7.1 Pressure Corrected Observations of Ground Level Enhancements	26
1.7.2 Pressure Corrected Observations of Forbush Decreases . . . . .	27
1.8 Discussion . . . . .	27
1.9 Conclusion . . . . .	27
<b>2 HiSPARC Station 14008</b>	29
2.1 Introduction . . . . .	29
2.2 Aims . . . . .	30
2.3 HiSPARC 14008 Detector Set-up . . . . .	30

2.3.1	Configuration . . . . .	30
2.3.2	Calibration . . . . .	31
2.3.3	Monitoring Temperature . . . . .	32
2.4	Observations . . . . .	33
2.5	Conclusions . . . . .	33
<b>3</b>	<b>Galactic Cosmic Ray Behaviour During Solar Cycle 24</b>	<b>34</b>
3.1	Introduction . . . . .	34
3.2	Data . . . . .	39
3.3	Time-Lag Analysis . . . . .	41
3.3.1	Method . . . . .	41
3.3.2	Results . . . . .	41
3.4	Hysteresis Effect Analysis . . . . .	47
3.4.1	Method . . . . .	47
3.4.2	Results . . . . .	48
3.5	Conclusions . . . . .	52
3.6	Update: End of Cycle 24 . . . . .	53
3.7	Comparison with HiSPARC . . . . .	53
<b>4</b>	<b>A Frequency Domain Investigation on the Morphology of the Solar Mean Magnetic Field</b>	<b>57</b>
4.1	Introduction . . . . .	57
4.2	Aims . . . . .	61
4.3	Data . . . . .	61
4.3.1	Summary of the Data Set . . . . .	61
4.3.2	Obtaining the SMMF from BiSON . . . . .	62
4.3.3	Comparison between WSO and BiSON . . . . .	67
4.4	Methodology . . . . .	70
4.4.1	Identifying Features in the SMMF Power Spectrum . . . . .	70
4.4.2	Parameterisation of the SMMF Power Spectrum . . . . .	72
4.4.3	Comparison with the WSO SMMF . . . . .	74
4.5	Results . . . . .	74
4.5.1	Investigation of the Window Function . . . . .	74
4.5.2	Modelling the BiSON Power Spectrum . . . . .	80
4.5.3	Comparison to the WSO Power Spectrum . . . . .	83
4.6	Discussion . . . . .	86
4.6.1	Asymmetry in the Power Spectrum . . . . .	86
4.6.2	Testing the Effects of Differential Rotation and Active Region Migration . . . . .	91
4.6.3	Further Morphology of the SMMF using SDO/HMI Data . . . . .	95
4.7	Conclusion . . . . .	96
<b>5</b>	<b>Rossby Modes in the Solar Mean Magnetic Field</b>	<b>99</b>
5.1	Introduction . . . . .	99
5.2	Aims . . . . .	101
5.3	Theory . . . . .	102
5.4	Methodology . . . . .	103

5.4.1	Testing the Residual Spectrum . . . . .	103
5.4.2	Modelling <i>r</i> mode Profiles . . . . .	105
5.5	Results . . . . .	106
5.5.1	Testing the Residual Spectrum . . . . .	106
5.5.2	Modelling <i>r</i> mode Profiles . . . . .	107
5.6	Discussion . . . . .	110
5.6.1	Manifestation of Rossby Waves in the Power Spectrum . . . . .	110
5.6.2	Rossby Modes in Other Sources of SMMF Data . . . . .	114
5.7	Conclusion . . . . .	115
<b>A</b>	<b>Simulations of Cosmic Ray Air Showers using CORSIKA</b>	<b>118</b>
<b>B</b>	<b>Simulations of the Artificial SMMF</b>	<b>122</b>
B.1	Model . . . . .	122
B.2	Configuration of the Simulations . . . . .	124
B.3	Outputs . . . . .	127
<b>Bibliography</b>		<b>130</b>

# List of Figures

1.1	Schematic diagram of the HiSPARC scintillation detector. (A): PMT; (B): light-guide adaptor; (C): light-guide; (D): scintillator. . . . .	3
1.2	Typical formations of two-detector and four-detector stations. In each, the grey circle denotes a GPS antenna which is located in between the detectors to provide a precise timestamp for each signal. . . . .	5
1.3	Azimuthal and zenith angle variations in the allowed and forbidden rigidity trajectories for HiSPARC station 501. . . . . . . . . . .	10
1.4	The vertical asymptotic viewing directions of 5 HiSPARC stations. The rigidity range of the simulations were from $1.0 \text{ GV} < R < 20.0 \text{ GV}$ , and the results are plotted in geographic coordinates on January 20th 2005. The diamonds correspond to the HS ground location and the circles correspond to the AVD for a specific rigidity value. . . . . . . . . . .	11
1.5	GLEs observed by the NM stations based at Oulu. Top panel: GLE 70; middle panel: GLE 71, bottom panel: GLE 72. The solid-black line shows the 2-minute-averaged, pressure corrected data and the vertical, dashed-red lines show the epochs of each GLE onset. . . . .	13
1.6	FDs observed by the NM stations based at Oulu. Top left panel: FDs during March 2012; top right panel: FD during July 2012, bottom left panel: FD during December 2014; bottom right panel: FD during September 2017. The solid-black line shows the 2-minute-averaged, pressure corrected data and the vertical, dashed-blue lines show the epochs of each FD onset. . . . . . . . . . .	13
1.7	HiSPARC data for stations 501 and 3001 around the epoch of GLE 70. The plot shows the minute-averaged and 5-minute-averaged trigger events between detectors within the station. The vertical red, dashed line depicts the approximate onset time of the GLE. The units of time on the x-axis are, MM-DD HH. . . . . . . . . . .	14
1.8	HiSPARC data for stations 8001 and 3001 around the epoch of GLE 71. The plot shows the minute-averaged and 5-minute-averaged trigger events between detectors within the station. The vertical red, dashed line depicts the approximate onset time of the GLE. The units of time on the x-axis are, MM-DD HH. . . . . . . . . . .	14

1.9	HiSPARC data for 4 stations around the epoch of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical red, dashed line depicts the approximate onset time of the GLE. The units of time on the x-axis are, MM-DD HH. . . . .	15
1.10	HiSPARC data for stations 501 and 8001 around the epoch of the FDs in March 2012. The plot shows the minute-averaged and hourly-averaged trigger events between detectors within the station. The vertical blue-dashed lines show the approximate onset-time of the FDs. . . . .	16
1.11	HiSPARC data for stations 501 and 8001 around the epoch of the FD in July 2012. The plot shows the minute-averaged and hourly-averaged trigger events between detectors within the station. The vertical blue-dashed line shows the approximate onset-time of the FD. . . . .	16
1.12	HiSPARC data for 4 stations around the epoch of the FD in December 2014. The plot shows the minute-averaged and hourly-averaged trigger events between detectors within the station. The vertical blue-dashed line shows the approximate onset-time of the FD. . . . .	17
1.13	HiSPARC data for [n] stations around the epoch in which there were several FDs close to the onset of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical blue-dashed lines show the approximate onset-times of the two FDs observed around this epoch and the red-dashed line depicts the approximate onset time of the GLE. . . . .	18
1.14	Mean muon density footprints for (a) proton-initiated air showers and (b) $\alpha$ -particle-initiated air showers with initial PCR trajectories with zenith angles $\theta = 0^\circ$ and various PCR energies. The error bars given represent $1\sigma$ . . . . .	20
1.15	Mean number of muons produced at ground level by the PCR for (a) proton-initiated air showers and (b) $\alpha$ -particle-initiated air showers, for various PCR energy. The uncertainty bars given represent $1\sigma$ . . . . .	21
1.16	The anti-correlation between CR count rates and the atmospheric pressure. (a) shows the CR and the local atmospheric pressure measured at a NM in the South Pole; (b) shows the CR and pressure measured by HiSPARC station 501. . . . .	22
1.17	The barometric coefficient calculation: (a) during November 2017 for the South Pole (SOPO) NM station, (b) during November 2019 for HiSPARC station 501 at Nikhef. . . . .	23
1.18	A comparison between the monthly barometric coefficient computed in this work and using the online barometric coefficient tool throughout the year 2017 for the SOPO NM station. . . . .	23

1.19	Pressure corrected HiSPARC data for 2 stations around the epoch of GLE 71. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical red, dashed line depicts the approximate onset time of the GLE. . . . .	26
1.20	Pressure corrected HiSPARC data for 2 stations around the epoch of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical red, dashed line depicts the approximate onset time of the GLE. . . . .	27
1.21	.....	27
1.22	.....	28
1.23	Pressure corrected HiSPARC data for [n] stations around the epoch in which there were several FDs close to the onset of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical blue-dashed lines show the approximate onset-times of the two FDs observed around this epoch and the red-dashed line depicts the approximate onset time of the GLE. . . . .	28
2.1	Schematic diagram of the HiSPARC station 14008 detector set-up. . . . .	30
2.2	HiSPARC 14008 assembly and configuration. (a) shows the stacked arrangement of the scintillators on the lab work bench, between layers of protective foam. (b) shows the complete detector inside the ski-box on the University of Birmingham campus. . . . .	31
2.3	Voltage calibration curve for the PMTs of station 14008. The upper, red-dashed line indicates the upper limit for the low threshold singles rate (400 Hz), and the lower 2, black-dashed lines indicate the upper and lower bounds for the high threshold singles rate (100–130 Hz). . . . .	32
2.4	Schematic diagram of the DS18B20 temperature sensor circuit, whereby the voltage, ground, and GPIO interfaces connect directly into pins of the Raspberry Pi board. . . . .	33
3.1	SSN (top), with vertical lines showing the beginning of each solar cycle. Cosmic ray intensity recorded by NM (bottom), with vertical lines showing the approximate epochs of solar magnetic field polarity reversals. (MCMD = McMurdo, NEWK = Newark, SOPO = South Pole, THUL = Thule). . . . .	36
3.2	Hysteresis plots between yearly averaged SSN and yearly averaged GCR intensity for each of the 4 main NM stations over cycles 19-24. . . . .	38
3.3	Variation in the correlation coefficient with time-lag NM station GCR intensity and SSN during solar cycles 20-23. . . . .	43
3.4	Variation in the correlation coefficient with time-lag between NM GCR intensity and SSN during solar cycle 24. . . . .	44

3.5	Variation in the correlation coefficient with time-lag between NM GCR intensity and SSN between 2000-2012. . . . .	45
3.6	Variation in time-lag plotted against NM station rigidity cut-off for the 16 NM stations detailed in Table 3.1. . . . .	46
3.7	The hysteresis plots for even solar cycles 20 and 22 and the linear regression fit to the data. . . . .	49
3.8	The hysteresis plots for odd solar cycles 21 and 23 and the linear regression fit to the data. . . . .	50
3.9	The hysteresis plots for odd solar cycles 21 and 23 and the ellipse fit to the data. . . . .	50
3.10	The hysteresis plot for solar cycle 24, and the linear regression fit to the data (left) and ellipse fit to the data (right). . . . .	52
3.11	The monthly-mean CR-induced muon count rate recorded by HiS-PARC (blue), and the SSN (black) between 2008 – 2019. . . . .	54
3.12	. . . . .	55
4.1	An example of the BiSON ratios data over a 30-minute period. The separation between the two ratios is due to the solar mean magnetic field. Other excursions in the individual ratios are due to the other effects measured by the RSS. . . . .	64
4.2	(a) 40-second cadence observations of the SMMF from the Sutherland BiSON station between 1992 and 2012. The sense of the field was chosen to match the Chaplin et al. (2003) and the WSO observations, where positive is for a field pointing outwards from the Sun. (b) Power spectrum of the SMMF on a 40-second cadence truncated to $10\mu\text{Hz}$ , however, the nyquist frequency is $12.5\text{ mHz}$ . . . . .	66
4.3	Daily averaged SMMF measured by WSO (blue) and by the Sutherland BiSON station (black). . . . .	67
4.4	Comparisons between the BiSON SMMF data and the WSO SMMF data. (a) shows the correlation between daily averaged BiSON SMMF and the WSO SMMF. The solid-green line provides the fit to the data, while the dashed-red line shows a 1:1 relation for comparison. (b) shows a comparison between the power spectra of BiSON and SMMF. . . . .	68
4.5	Formation heights of emission lines over the wavelength range used for WSO observations (top) and for BiSON observations (bottom). The vertical lines show the wavelengths of the emission lines used for observations in air (dashed) and in vacuum (solid). This plot was provided in private communication courtesy of Rinat Tagirov using the NESSY code. . . . .	69
4.6	Full power spectrum of the BiSON SMMF on a logarithmic scale up to the Nyquist frequency. . . . .	70
4.7	Locations of aliased power in side-band peaks. The orange, dotted-lines show the locations of frequencies at multiples of 1/day. The green, dashed-lines show the location of the side-band peaks – harmonic frequencies reflected around multiples of 1/day. The inset shows a zoom of one set of side-band peaks around 1/day. . . . .	75

4.8	(a) The effect on the power spectrum due to periodic gaps in the data. (b) The effect of the convolution between the the window function and the input power spectrum used to make the fake data, highlighting the need to adjust the power for the duty cycle. . . . .	77
4.9	Full, modelled power spectrum of the BiSON SMMF on logarithmic axes. The data is displayed in black and the convolved model using asymmetric Lorentzian peaks is shown in green. . . . .	82
4.10	Full, modelled power spectrum of the WSO SMMF on logarithmic axes. The data is displayed in black and the convolved model using asymmetric Lorentzian peaks is shown in blue. . . . .	85
4.11	(a) shows the KDE of the monthly averaged sunspot number used to draw samples of source seed times; (b) histogram of the rotation frequency of sources sampled from the KDE, after using the model for the migration and differential rotation. . . . .	87
4.12	(a) shows the latitudinal migration model as a function of time for each model; (b)-(d) shows the histogram of the rotation frequencies of sources sampled from the KDE for linear, quadratic, and exponential migration models, respectively. . . . .	88
4.13	The median value of the fitted global asymmetry parameter for several simulations of artificial data. The colour represents the migration model used: linear:blue; quadratic:orange; exponential: green. The marker represents the source model used: cosine:circle; sign change:cross; combinations of the cosine and sign change models in ratios of 5:95:downwards triangle; 10:90:plus; 20:80:upwards triangle.	89
4.14	(a) Shows the peak distribution before and after the time-averaged broadening, and the fit to the broadened peak. (b) Shows the peak distribution before and after the analytical broadening, and the fit to the broadened peak. In both plots the broadened peaks have been shifted by the relevant frequency to overlay them on top of the true $\nu_0$ for comparison. . . . .	94
4.15	Investigations of timescales in the SDO/HMI magnetograms over 2011 and 2014. Both plots show in the top panel, the hemispheric Mean Magnetic Field (MMF) and full-disk Solar Mean Magnetic Field (SMMF) from the magnetograms. The lower panel of each plot displays a comparison between the hemispheric and full-disk mean of the synoptic charts, compared to the box-car smoothed MF from the magnetograms. N: Northern hemisphere; S: Southern hemisphere. Full HMI: considers the full solar disk; 99 HMI: considers only the inner 99% of the solar disk, by radius. . . . .	96
5.1	Full, modelled power spectrum of the BiSON SMMF on logarithmic axes. The data is displayed in black and the convolved model using asymmetric Lorentzian peaks is shown in green. . . . .	101

5.2	Residual power spectrum of the BiSON SMMF. Over plotted in the green curve is the model of the main SMMF signal. Also over plotted as vertical solid lines are the expected locations of the 4 lowest-frequency sectoral $r$ modes and the dashed lines, the locations of the $B_0$ variation frequency splitting. . . . .	106
5.3	Realisations of the statistics tests of BiSON data for different re-binning factors ( $n$ ). The panels of each sub figure are: (top) the full PSD and fit, (second panel) the full and re-binned residuals, (third panel) the probability of statistical noise in each bin, (bottom) distribution of the residuals compared to a $\chi^2$ 2n-DOF. . . . .	108
5.4	Model fit to the $r$ mode in the BiSON PSD residuals using the median from the parameter posterior distributions. . . . .	109
5.5	Mode displacement schematic for an $l = m = 2$ $r$ mode (Strohmayer & Mahmoodifar, 2014) . . . . .	111
5.6	Time series of the velocity and amplitude modulation toy model simulations. The blue curve shows the velocity modulation, i.e. modulating using a cosine with period of 1 year, whereas the orange curve shows the amplitude modulation, i.e. modulating using a rectified cosine with period 1 year. . . . .	112
5.7	Power spectra for the two modulation methods, showing the difference in the way the modulation has changed the frequency of the observed mode. . . . .	112
5.8	SDO/HMI SMMF split into hemispheres and compared to other SMMF sources during (a) 2011, and (b) 2014. The top panel in each figure shows the north (N), south (S), and total disk-averaged mean magnetic field, for both the full solar disk and from pixels within 99% of solar radius. The bottom panels show a comparison between the SMMF, as observed BiSON, WSO, and SDO/HMI (full disk and the 99% disk). . . . .	113
5.9	Comparison of the power spectra for BiSON, WSO, and SDO/HMI. In both figures, the top panel shows the BiSON PSD and the bottom panel shows either the WSO or HMI PSD. The dashed, black line shows the location of theoretical $l = 2 = m$ $r$ mode frequency, and the red, solid line shows the location of the peak fit in the BiSON PSD residuals. . . . .	115
B.1	Schematic representations of the two models of the artificial SMMF. (a) shows the cosine model with a single source of constant polarity transiting the visible disk. (b) shown the sign change model whereby there are 2 regions of opposite polarity transiting the disk, and their contribution to the SMMF changes during the transit. . . . .	123
B.2	A single realisation of (a) the cosine model, and (b) the sign change model. . . . .	124
B.3	Flowchart showing the step-by-step processes in the generation of the artificial SMMF time series data. . . . .	125

B.4	(a) the burn-in period required to ensure that the interference of sources stabilises. The dashed-vertical line marks one year, showing the sources have stabilised. (b) histogram of the number of sources which remain in the simulation (out of the 20000 input), after the burn-in is removed. (c) the number of simulated sources on the visible disk over the BiSON observational epoch. (d) the daily-averaged sunspot number (blue, and the monthly-averaged sunspot number (orange)). . . . .	126
B.5	Artificial butterfly diagram generated from the simulations, allowing for the parameters to be drawn from distributions, to add stochasticity into the simulation. . . . .	127
B.6	Limit spectrum from 100 realisations of the cosine model (a) and the sign change model (b) using a single source in each model. . . . .	128
B.7	Time series and power spectra for realisations of the simulations using the cosine and sign change models. . . . .	129

# List of Tables

1.1	Properties of some of the HiSPARC stations: geographic longitude ( $\lambda$ ), geographic latitude ( $\phi$ ), altitude ( $h$ ), and the geomagnetic vertical cut-off rigidity ( $R_C$ ) calculated from the PLANETOCOSMICS simulations. . . . .	10
1.2	Space weather events investigated within the HiSPARC data. The percentage change column provides a reference of how much the CR counts observed by the NM station at Oulu ( $R_c=0.81$ GV) increased or decreased by, due to the space weather event. More precise times for the event onset can be found at NMDB (2018) (for GLEs) and Lingri et al. (2016) (for FDs). . . . .	12
3.1	Neutron monitor stations used in this study and their vertical geomagnetic cut-off rigidity ( $R_c$ ), longitude, latitude, and altitude acquired from NEST. The first four stations have been used for all of the analysis while the lower 12 stations have been used exclusively for the investigation into the dependence of $R_c$ on the time-lag. . . . .	40
3.2	Time-lags and the corresponding cross-correlation coefficient between NM CR count and SSN for solar cycles 20-23. . . . .	42
3.3	Time-lags and the corresponding cross-correlation coefficient between NM GCR intensity and SSN for solar cycle 24. . . . .	44
3.4	Time-lags and the corresponding cross-correlation coefficient between NM GCR intensity and SSN during 2000-2012. . . . .	45
3.5	Correlation coefficients of the linear regression and ellipse modeling of the hysteresis plots for solar cycles 20-23. . . . .	48
3.6	Correlation coefficients of the linear regression and ellipse modeling of the hysteresis plots for solar cycle 24. . . . .	51
4.1	Model parameter values for the generation of artificial data, and the median posterior values for the fit to the power spectra generated with and without the gaps in the data. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median. . . . .	79

4.2	Median values of the marginalised posterior distributions for each model parameter in the fit to the BiSON power spectrum, adjusted for the duty cycle factor ( $\sim 0.156$ ) in the convolution process. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median. . . . .	81
4.3	Median values of the marginalised posterior distributions for each model parameter in the fit to the WSO power spectrum, adjusted for the duty cycle factor ( $\sim 0.778$ ) in the convolution process. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median. . . . .	84
4.4	Input line-width and the median posterior values of the Lorentzian model each simulation. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median. . . . .	93
5.1	Predicted <sup>+</sup> and observed <sup>◦</sup> $r$ mode frequencies for combinations of $l$ and $m$ . Predicted frequencies and conversions of observations to different frames of reference use equation (5.1), equation (5.2), and equation (5.3), with $\Omega = 453.1$ nHz. The predicted splitting for the $B_0$ angle variation is also provided. The key for the source column is: LPT for Löptien et al. (2018), LNG for Liang et al. (2019), and LZA for Lanza et al. (2019). . . . .	103
5.2	Median posterior values of the Lorentzian model for the $r$ mode peak in the BiSON SMMF PSD residuals. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median. . . . .	107
A.1	Details of the proton-initiated air showers simulated using CORSIKA. . . . .	120
A.2	Details of the $\alpha$ -particle-initiated air showers simulated using CORSIKA. . . . .	121

# List of Abbreviations

**AR** Active Region.

**AVD** Asymptotic Viewing Direction.

**BiSON** Birmingham Solar Oscillations Network.

**BMR** Bipolar Magnetic Region.

**CORSIKA** Cosmic Ray Simulations for Kascade.

**CR** Cosmic Ray.

**EAS** Extensive Air Shower.

**EEMD** Empirical Mode Decomposition.

**FA** False Alarm.

**FD** Forbush Decrease.

**GLE** Ground Level Enhancement.

**GMF** General Magnetic Field.

**HMC** Hamiltonian Monte Carlo.

**IGRF** International Geomagnetic Reference Field.

**KDE** Kernel Density Estimate.

**LOS** Line Of Sight.

**MCMC** Markov Chain Monte Carlo.

**MD** Muon Detector.

**MFC** Magnetic Flux Concentration.

**MIP** Minimum Ionising Particle.

**MMF** Mean Magnetic Field.

**NLTE** Non-Local Thermal Equilibrium.

**NM** Neutron Monitor.

**NUTS** No U-Turn Sampler.

**PCR** Primary Cosmic Ray.

**PMMA** Polymethylmethacrylate.

**PMT** Photo Multiplier Tube.

**PSD** Power Spectral Density.

**RM** Rotationally Modulated.

**RSS** Resonant Scattering Spectrometer.

**SB** Stochastic Background.

**SCR** Solar Cosmics Ray.

**SDO/AIA** Solar Dynamic Observatory Atmospheric Imaging Assembly.

**SDO/HMI** Solar Dynamic Observatory Helioseismic and Magnetic Imager.

**SEP** Solar Energetic Particle.

**SMMF** Solar Mean Magnetic Field.

**SOHO/MDI** Solar and Heliospheric Observatory Michelson Doppler Imager.

**SSN** Sun Spot Number.

**UHECR** Ultra-High Energy Cosmic Ray.

**WSO** Wilcox Solar Observatory.

# 1 HiSPARC as a Space Weather Detector

## 1.1 Introduction

... [on daily variations (DV)] Dr. Rolf Butikofer (in a reply from Danislav Sapundzjiev, dasapund@meteo.be) said:

*"The daily cosmic ray variation near Earth is caused by the anisotropy of the cosmic ray intensity in the interplanetary space. Cosmic ray particles follow the field lines of the interplanetary magnetic field when they travel towards the interior of the heliosphere. Because of the rotation of the Earth, the angle between the asymptotic cone of acceptance of various energies at the location of ground-based cosmic ray detectors (neutron monitors) and the direction of the interplanetary magnetic field varies with a time period of 24 hours. As a consequence cosmic ray detectors look in different directions in the course of a day and observe therefore a diurnal variation. The daily variations of neutron monitors is mainly seen by high latitude stations which have asymptotic directions at low energies (rigidities) near the equator."*

... [on cosmic ray electron (CRE) losses and lower than protons] Tinivella (<http://arxiv.org/abs/1610.03672>) said:

*"The first term describes ionization losses in ISM and is dominant for energies up to a few tens of MeV. The second term is due to bremsstrahlung, adiabatic losses and pair production in electron-gamma interaction, while the last term represents losses by synchrotron emission and Inverse Compton scattering (IC) under the Thomson approximation, that holds very well for electrons up to a few TeV of energy."*

### 1.1.1 Space Weather Effects

Put something in here about the type of effects that have been observed, and how/why to refer back to with our observations...

### 1.1.2 HiSPARC Project

HiSPARC stands for *High School Project on Astrophysics and Research with Cosmics*, and it is a scientific outreach project that was initiated in the Netherlands in 2002 (Bartels, 2012). The HiSPARC project has two main goals: the study of Ultra-High Energy Cosmic Ray (UHECR) for astroparticle physics research, and to serve as a resource to expose high school students to scientific research (Bartels, 2012).

HiSPARC is a global network of muon detectors spread across the Netherlands, Denmark, the UK, and Namibia. The detectors at each station record muon counts and may be used for many scientific experiments, such as: reconstruction of the direction of a cosmic ray induced air shower, reconstruction of the energy of the air shower's primary particle, investigation between the atmospheric conditions and the number of cosmics rays observed, etc.

Data recorded by the HiSPARC stations are stored and are available publicly at <http://www.hisparc.nl>, where the Cosmic Ray (CR) counts, atmospheric data, station metadata, and more can be found.

### 1.1.3 HiSPARC Detector and Station Configuration

The detection philosophy of HiSPARC is to sample the footprints of Extensive Air Showers (EASs) using coincident triggers between scintillation detectors. As HiSPARC was set up as an outreach programme for high schools, this impacted detector design. Resources are limited in schools and the detectors are usually financed by the participating high schools, colleges, and universities. In addition, students (accompanied by their teachers and local node support staff) are responsible for assembly and installation their detectors, which are typically installed on the roofs of schools.

Due to this, the detectors needed to be cheap, robust, and easily maintainable, therefore the scintillation detector was selected for the HiSPARC network.

Scintillators consist of materials that emit light when charged particles pass through them with sufficient energy to ionise the scintillator material. The total light produced is proportional to the number of charged particles, and can be collected by a Photo Multiplier Tube (PMT). Each HiSPARC detector utilises a plastic scintillator of dimensions 1000 mm x 500 mm x 20 mm, providing a detection area of 0.5 m<sup>2</sup>. A vertically incident Minimum Ionising Particle (MIP) has a most probable energy loss in 2 cm of the scintillation material of 3.51 MeV ( $\equiv$  1 MIP) (van Dam et al., 2020).

The scintillator is glued to a triangular/‘fish-tailed’ light-guide (dimensions, base: 500 mm; top: 25 mm; height: 675 mm), and a light-guide adapter provides the optical interface between the square end of the light-guide and the cylindrical aperture of the PMT. The configuration of a single HiSPARC detector is shown in Figure 1.1.

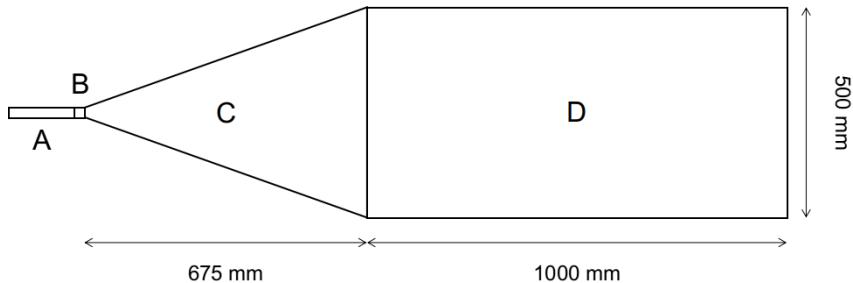


Figure 1.1: Schematic diagram of the HiSPARC scintillation detector. (A): PMT; (B): light-guide adaptor; (C): light-guide; (D): scintillator.

The scintillator is made of a material consisting of polyvinyltoluene as the base, with anthracene as the fluor, and the emission spectrum peaks at a wavelength of 425 nm (Fokkema, 2012; Bartels, 2012). The light-guide is made from Polymethylmethacrylate (PMMA) and has a comparable refractive index to the scintillator (1.58 and 1.49, respectively), reducing refraction effects between the two materials (van Dam et al., 2020).

The PMT used is an ETEEnterprises 9125B PMT, with a 25 mm aperture, blue-

green sensitive bialkali photocathode, and 11 high-gain dynodes (Bartels, 2012; ETEEnterprises, 2020). The quantum efficiency of the PMT used in the HiSPARC detectors peaks at around 375 nm at 28%, and at 425 nm the quantum efficiency is 25% (Fokkema, 2012).

Each detector is wrapped in aluminium foil (thickness 30  $\mu\text{m}$ ) and a black, vinyl material (thickness 0.45 mm), which is usually used as a pond liner, to ensure light-tight detectors and to reduce the noise level from stray photons (van Dam et al., 2020). In addition, each detector is placed inside of its own a plastic ski-box to again ensure that it is light-tight, and to also ensure that it is weather-proof, as the detectors are usually located on the roofs of schools, colleges, and universities.

A HiSPARC station combines either 2 or 4 detectors, to observe coincident muons ('events'), and typical configurations of each are shown in Figure 1.2. The separation between detectors varies from station-to-station. In addition some stations have the capability to measure the local atmospheric properties, such as temperature, pressure, relative humidity etc. Moreover, some stations also record the 'singles' rates, i.e. the frequency at which an individual detector is triggered, independently of the other detectors in the station. The singles rates are important when investigating non-EAS events.

The PMTs of the detector in a station are connected to HiSPARC electronics boxes by cables with a standard length of 30 m, to minimise any timing offsets between detectors (Fokkema, 2012; van Dam et al., 2020). The electronics boxes are capable of controlling and reading two PMTs, therefore a four-detector station requires two electronics boxes: a master and a slave.

The HiSPARC experiment is set up in such a way as to ensure that each station across the HiSPARC network reads a similar count rate of muons, in order to aid the direct comparison between the different stations in the network. When configuring the station, a trigger threshold must be applied for the PMT signals. This is standardised across the HiSPARC network. There are two thresholds, low:



(a) Two-detector station configuration



(b) Four-detector station configuration (triangle arrangement)



(c) Four-detector station configuration (diamond arrangement)

Figure 1.2: Typical formations of two-detector and four-detector stations. In each, the grey circle denotes a GPS antenna which is located in between the detectors to provide a precise timestamp for each signal.

30 mV, which represents 0.2 of a MIP; high: 70 mV, which represents 0.5 of a MIP.

The thresholds were chosen to increase the sensitivity of the stations for observing gamma rays and low energy electrons, but this has the effect of making it more

difficult to determine whether an individual detection is from a muon, or another MIP. This is why the HiSPARC network usually relies on detecting ‘events’, from coincident muons. It is requested that each detector in the network is set up such that the high threshold allows a mean count rate on the order 100 counts per second and the low threshold allows a mean count rate of the order 400 counts per second; these can be tuned by adjusting the PMT voltage. It could be argued that in setting up the detectors in this way, there is an immediate bias in the data to reject lower energy CRs.

[show image of a typical PMT trace to help explain the thresholds used for the signal triggers / include also a plot of the pulseheight diagram]

The pulseheight spectrum (see Figure ...) is composed of two main regions: the left side which falls off rather steeply and the main part of the spectrum which features a peak and a long tail. The left side of the spectrum is understood to be from high energy photons produced in air showers (Fokkema, 2012). These high energy photons may undergo pair production when interacting with the scintillator which may produce ionising electron and positron pairs. The trigger thresholds are placed to reject these noise signals from the data.

The main, asymmetric distribution which features a peak and a tail is from charged particles (muons and electrons). The mean energy loss of particles in a material is described by the Blethe-Bloch formula; however this does not account for fluctuations in energy loss (Fokkema, 2012). A Landau distribution in fact describes the fluctuations in energy loss of particles. Due to the resolution of the HiSPARC detectors the distribution in Figure ... is best described by the convolution of the Landau distribution with a normal distribution which describes the resolution of the detector (Fokkema, 2012). The peak of the distribution, the most probable values (MPV), is the most likely energy lost by a particle in the detector, i.e. the 3.51 MeV MIP (van Dam et al., 2020). It has been shown that the location of the MPV can vary due to the effects of atmospheric temperature (Bartels, 2012; van Dam et al.,

2020).

The default trigger conditions for detecting an air shower event between multiple PMTs within a station differ for a two/four-detector station. In a two-detector station, an event is recorded if the PMT signals from both detectors exceed the low threshold within the coincidence time window ( $1.5\ \mu\text{s}$ ). In a four-detector station, there are two conditions: (i) at least two detectors exceed the high threshold within the coincidence time window; (ii) at least three detectors exceed the low threshold within the coincidence time window. These are the default conditions, but there are other, user configurable ways of triggering the station.

The scientific goals that can be achieved also vary between the two/four-detector stations. When at least three detectors in a four-detector station observe particles of an EAS, the direction of the EAS (and thus the direction of the Primary Cosmic Ray (PCR)) can be acquired using triangulation calculations. When only two detectors in a station observe particles of an EAS, i.e. the limit for a two-detector station, it is only possible to reconstruct the arrival direction along the axis that connects the centres of those two detectors (thus it is not possible to reconstruct the direction of the PCR).

## 1.2 Aims

The HiSPARC project was set up with the detection philosophy of observing EAS, which are typically associated with PCRs with energy of  $\sim 10^{14}$  eV and above, that produce large footprints observable with many HiSPARC stations simultaneously. For PCRs with energy below  $\sim 10^{14}$  eV the air shower is small, with almost no observable footprint, and for PCRs with energy below  $\sim 10^{11}$  eV, there is typically fewer than one or two muons that reach the ground, making their observation difficult.

The HiSPARC detectors are capable of observing any muons that reach them, therefore the project was motivated by the existing network of Muon Detector (MD)

which may have the capability of observing the CRs associated with space weather events.

The principle aim of the project was to determine whether the existing HiSPARC network is capable of observing space weather events. To do this, we investigated the properties of the HiSPARC detectors, to learn about what typical PCRs we observe. This was initially achieved by investigating the data during periods of space weather activity to search for the associated signatures. We searched through some of the most reliable HiSPARC stations to determine whether these events were observed in the data. This was done to determine whether, without much effort, we could get a binary answer on whether these events were observed by HiSPARC.

Following this, we performed simulations of air showers initiated by CRs to understand the expected muon flux and dispersion at ground level. This helped us to understand how likely it is to observe the PCRs associated with space weather with the HiSPARC detectors, observing muons.

Finally, ground-based observations of muons from air showers are susceptible to the conditions in the atmosphere; therefore, where possible, we corrected for atmospheric effects and again reviewed the corrected data to determine whether the space weather events were observed.

### 1.3 HiSPARC Properties

To understand the PCR spectrum that the HiSPARC stations are capable of observing, PCR transport simulations were performed using the PLANETOCOSMICS software. PLANETOCOSMICS performs Geant4 Monte Carlo simulations of charged particle transport through Earth's magnetosphere based on Størmers transport equation for charged particles (Desorgher et al., 2006). PLANETOCOSMICS simulates backward trajectories of charged particles from a given location (latitude, longitude, and altitude) out to the magnetopause for a set of PCR rigidities.

For each trajectory there are two possible outcomes: (i) the particles trace out to

the magnetopause where they escape Earth's magnetosphere, an allowed trajectory; (ii) the particles are sufficiently bent by the effect of the Earth's magnetosphere that they do not reach the magnetopause and cannot escape the Earth's magnetosphere, a forbidden trajectory (Desorgher et al., 2006). The coordinates of the asymptotic direction at the magnetosphere are provided as an output to the simulations projected back down to the Earth's surface. In this work PLANETOCOSMICS was configured with the Tsyganenko-89 model for the external magnetospheric magnetic field and the International Geomagnetic Reference Field (IGRF) internal field model.

For each rigidity simulated, whether it was an allowed or forbidden trajectory was stored, which was used to provide an insight into the rigidity spectrum for a given station. From the allowed trajectories the effective cut-off rigidity ( $R_C$ ) for the stations was computed using equation (1.1), where  $R_U$  is the upper rigidity (the last allowed trajectory before the first forbidden trajectory);  $R_L$  is the lower rigidity (the last allowed trajectory before which all other trajectories with a lower rigidity are forbidden);  $\Delta R$  is the rigidity step size in the simulation (Desorgher et al., 2006; Herbst et al., 2013).

$$R_C = R_U - \sum_{i=R_L}^{R_U} \Delta R_i \quad (1.1)$$

The rigidity spectrum for each of the HiSPARC stations were investigated to determine  $R_C$  for each station. The cut-off rigidity calculated for the six HiSPARC stations for a vertical incidence upon the atmosphere (i.e.  $0^\circ$  zenith angle) are shown in Table 1.1 which show that there is little variation in  $R_C$  between the HiSPARC stations and that they observe protons with rigidities in excess of  $\sim 3$  GV. This analysis was initially carried out for the vertical direction (i.e. azimuth =  $0^\circ$ , zenith =  $0^\circ$ ); however further trajectories were simulated for different azimuth and zenith angles to determine the dependence of the rigidity spectrum on the detector acceptance angle. The analysis for the azimuthal dependence was carried out at a

Table 1.1: Properties of some of the HiSPARC stations: geographic longitude ( $\lambda$ ), geographic latitude ( $\phi$ ), altitude ( $h$ ), and the geomagnetic vertical cut-off rigidity ( $R_C$ ) calculated from the PLANETOCOSMICS simulations.

Station Name/ID	$R_C$ [GV]	$\lambda$ [deg]	$\phi$ [deg]	$h$ [m]	No. Detectors
Nikhef/501	3.19	4.95 E	52.36 N	56.18	4
College Hageveld/203	3.18	4.63 E	52.35 N	53.71	2
Leiden/3001	3.23	4.45 E	52.17 N	54.08	2
Eindhoven/8001	3.44	5.49 E	51.45 N	70.12	2
Birmingham University/14001	3.06	1.93 W	52.45 N	204.14	4

zenith angle of  $20^\circ$  as this is around the most probable angle for HiSPARC events, and the analysis of the zenith dependence was carried out at an azimuth angle of  $0^\circ$ .

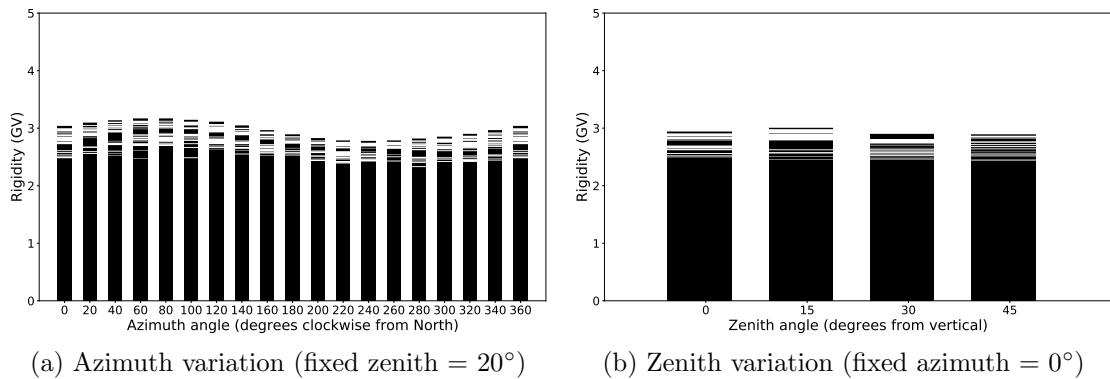


Figure 1.3: Azimuthal and zenith angle variations in the allowed and forbidden rigidity trajectories for HiSPARC station 501.

The small variation between HiSPARC stations is due to their close proximity in geographic latitude and longitude. The values of  $R_C$  calculated for the HiSPARC stations suggest that they should be able to observe higher energy Solar Cosmic Ray (SCR), but may not be as susceptible as the higher latitude Neutron Monitor (NM) where the effects of Ground Level Enhancements (GLEs) are highly observable.

As a result of the PLANETOCOSMICS simulations it was possible to understand the trajectories of particles that enter the Earth's magnetosphere prior to arrival at the atmosphere. It can be seen from Figure 1.4 that the Asymptotic Viewing Directions (AVDs) for each of the HiSPARC stations investigated are rather similar,

and that they mostly straddle the equator for low rigidity PCRs.

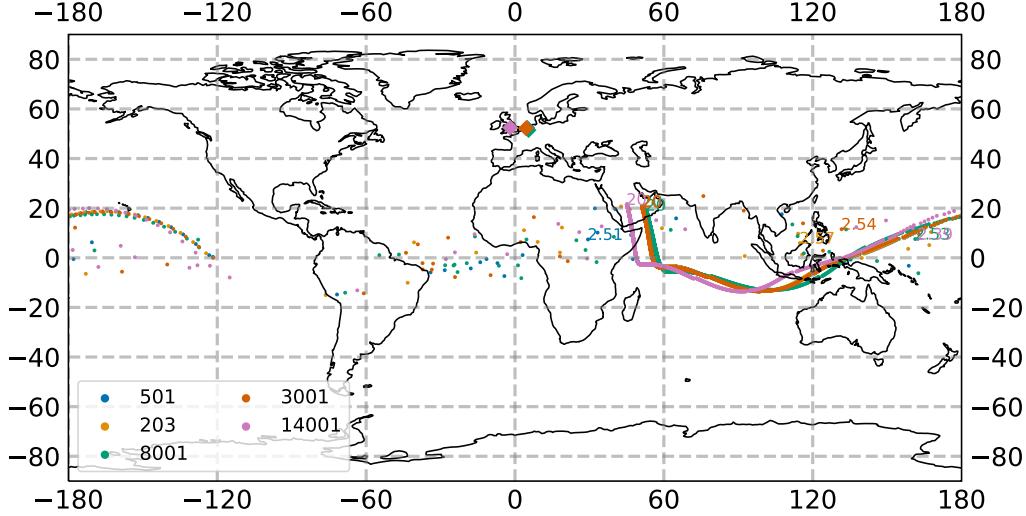


Figure 1.4: The vertical asymptotic viewing directions of 5 HiSPARC stations. The rigidity range of the simulations were from  $1.0 \text{ GV} < R < 20.0 \text{ GV}$ , and the results are plotted in geographic coordinates on January 20th 2005. The diamonds correspond to the HS ground location and the circles correspond to the AVD for a specific rigidity value.

The simulations were only performed up to a rigidity of 20 GV; however, at higher rigidities, we would see the AVDs spiral in towards the geographic location of the station, and the PCR would enter the magnetosphere and atmosphere almost vertically above the detector. This map of the AVDs also informs us that we should expect to be able to observe some lower energy PCRs when the zenith of the detector is not facing the asymptotic direction of the PCR.

## 1.4 HiSPARC Observations

[...!...end with discussion on unknown PCRs observable and the effect of atmospheric weather conditions that need to be accounted for...]

The effects of space weather on CRs has been outlined in [REF intro]. It was highlighted during private communication with the UK Met Office that observations of GLEs are of more interest and importance to space weather forecasts and nowcasts. Forbush Decreases (FDs) are of lower interest and importance, we still searched for FDs within the HiSPARC data. Table 1.2 outlines the specific space

Table 1.2: Space weather events investigated within the HiSPARC data. The percentage change column provides a reference of how much the CR counts observed by the NM station at Oulu ( $R_c=0.81$  GV) increased or decreased by, due to the space weather event. More precise times for the event onset can be found at NMDB (2018) (for GLEs) and Lingri et al. (2016) (for FDs).

<b>GLE Onset</b>	<b>GLE</b>	<b>% Change (Oulu)</b>	<b>FD Onset</b>	<b>% Change (Oulu)</b>
13/12/2006	70	~ 100%	08/03/2012	~ 10%
17/05/2012	71	~ 15%	12/03/2012	~ 3 – 5%
10/09/2017	72	~ 5%	14/07/2012	~ 3%
			21/12/2014	~ 5%
			06/09/2017	~ 2%
			07/09/2017	~ 8%

weather driven GLEs and FDs that we searched for within the HiSPARC data.

The specific events in Table 1.2 were selected as: (i) for the GLEs, they are the only three that fall in the HiSPARC operational period; (ii) for the FDs, they are the only individual, or set of, FDs that result in a count variation in excess of ~ 5% and the largest FDs are likely to be the most promising candidates for observation with HiSPARC.

For comparison with the HiSPARC results shown below, we show the GLEs, as observed by the Oulu NM station, in Figure 1.5. It is clear from Figure 1.5 that the relative increase of the GLEs was large for GLE 70 and 71, but much more subtle for GLEs 72. We expect that if we are to observe any of the GLEs, we shall have the best chance of observing GLEs 70.

Similarly, we show a comparison plot for the FDs, as observed by the Oulu NM station, in Figure 1.6.

As there are only a few space weather events that we were particularly interested in, and only a few HiSPARC stations that we felt were reliable for our investigation, we conducted the search for these GLEs and FDs by-eye in the data .

#### 1.4.1 HiSPARC Observations of Ground Level Enhancements

The search for evidence of GLEs within the HiSPARC data was conducted for GLE 70, 71, and 72, as they are the only GLEs that span the operational epoch of the

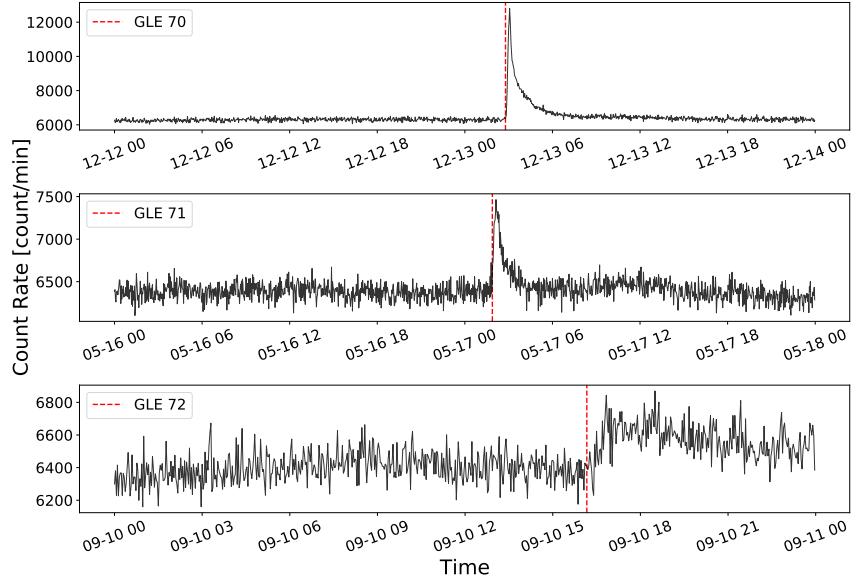


Figure 1.5: GLEs observed by the NM stations based at Oulu. Top panel: GLE 70; middle panel: GLE 71, bottom panel: GLE 72. The solid-black line shows the 2-minute-averaged, pressure corrected data and the vertical, dashed-red lines show the epochs of each GLE onset.

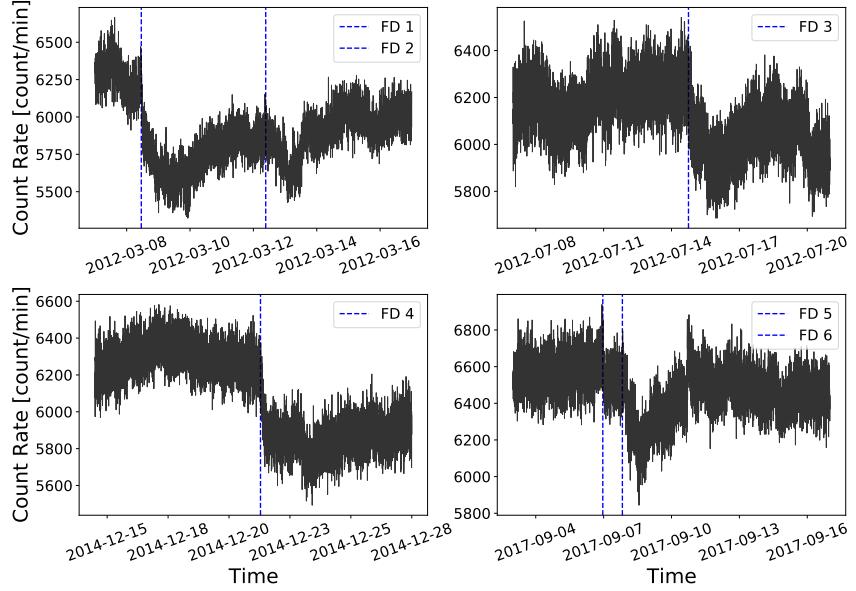


Figure 1.6: FDs observed by the NM stations based at Oulu. Top left panel: FDs during March 2012; top right panel: FD during July 2012, bottom left panel: FD during December 2014; bottom right panel: FD during September 2017. The solid-black line shows the 2-minute-averaged, pressure corrected data and the vertical, dashed-blue lines show the epochs of each FD onset.

HiSPARC network. Figure 1.7, Figure 1.8, and Figure 1.9 shows the HiSPARC observations around the epochs of GLE 70, 71, and 72, respectively.

Most of the observations show only the HiSPARC events data (i.e. coincidences between the detectors of a station); however, where possible, we also show the singles rates from each of the individual detectors in a station when the singles rate data is available.

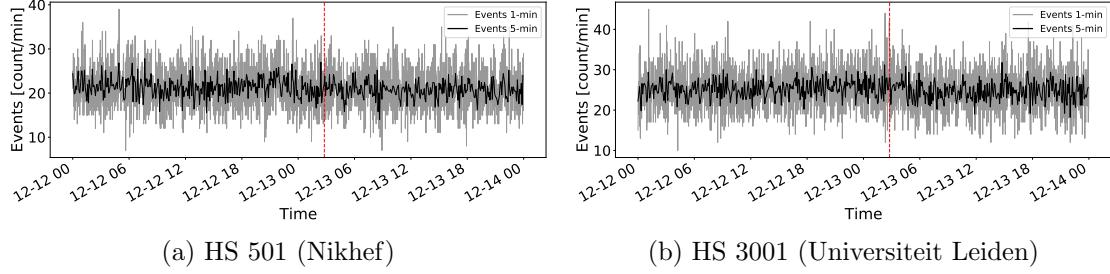


Figure 1.7: HiSPARC data for stations 501 and 3001 around the epoch of GLE 70. The plot shows the minute-averaged and 5-minute-averaged trigger events between detectors within the station. The vertical red, dashed line depicts the approximate onset time of the GLE. The units of time on the x-axis are, MM-DD HH.

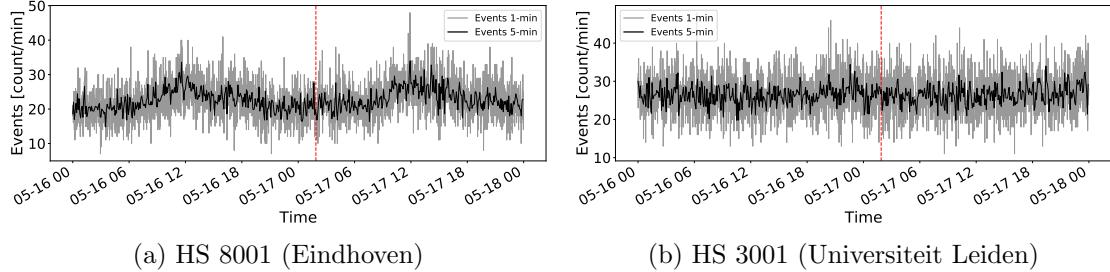


Figure 1.8: HiSPARC data for stations 8001 and 3001 around the epoch of GLE 71. The plot shows the minute-averaged and 5-minute-averaged trigger events between detectors within the station. The vertical red, dashed line depicts the approximate onset time of the GLE. The units of time on the x-axis are, MM-DD HH.

We can see from Figures 1.7, 1.8, and 1.9 that there are no clear and obvious signs of the GLE signals in the HiSPARC observations. This is the case for both the events data and the singles data.

There are some excursions from the mean count rate, this is significantly more prominent in the singles rates which are shown in the GLE 72 plots for stations 501, 203, and 8001. It is believed that these excursions are the effect of atmospheric pressure on the muon count rates; in Section 1.6 this is discussed further and is accounted for.

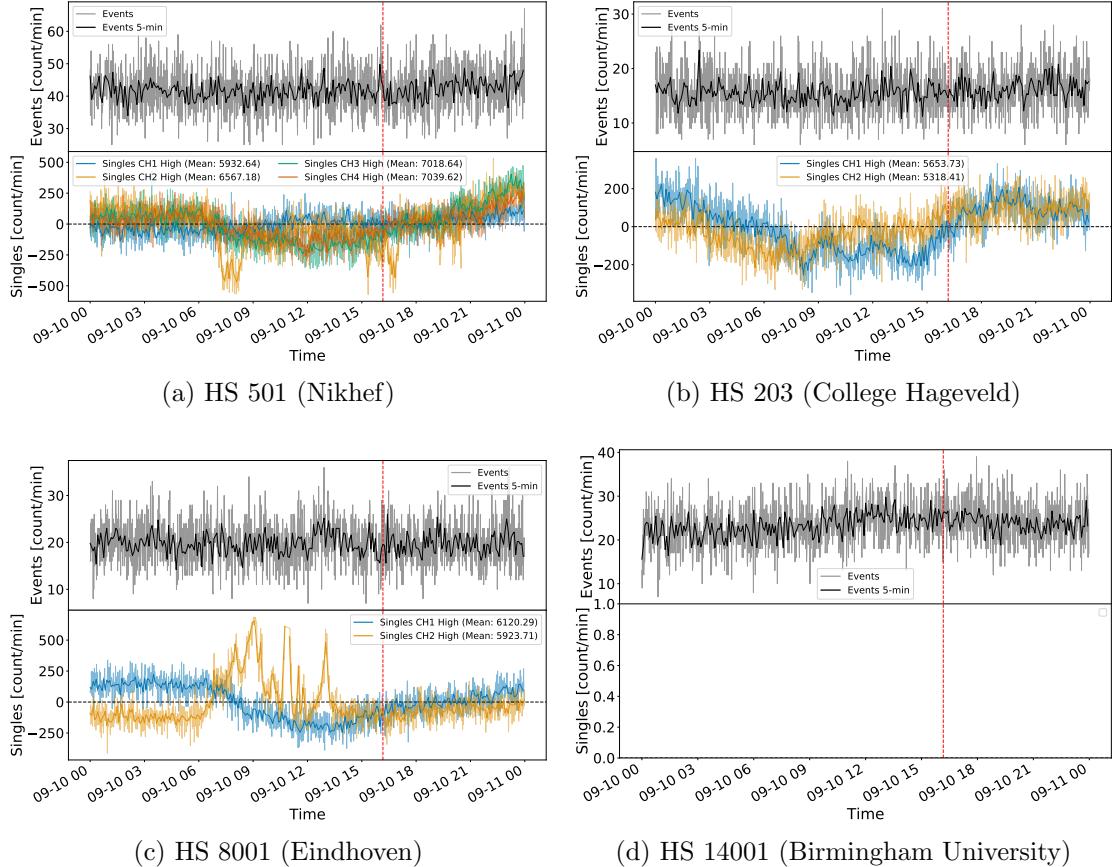


Figure 1.9: HiSPARC data for 4 stations around the epoch of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical red, dashed line depicts the approximate onset time of the GLE. The units of time on the x-axis are, MM-DD HH.

No clear GLEs have been observed in the HiSPARC data. We believe this is due to the rigidity cut-off of the HiSPARC stations, as GLEs are caused by Solar Energetic Particles (SEPs) with a lower energy. Typically GLEs are observed by NMIs, and only the most energetic have been observed by MDs [...(...cite to <https://doi.org/10.1088/0004-637X/761/2/101> and maybe also to <https://doi.org/10.1093/pasj/psv1>]. In Section ?? we investigated the CR spectrum to infer our ability to measure GLEs with the HiSPARC stations. We do also note that the atmospheric effects in the raw do not help our ability to observe the space weather and these effects were later removed (see Section 1.6.2).

### 1.4.2 HiSPARC Observations of Forbush Decreases

The search for evidence of FDs within the HiSPARC data was conducted for the FDs highlighted in Table 1.2. Figure 1.10, Figure 1.11, and Figure 1.12 show the HiSPARC observations around the epochs of the first four FDs listed in Table 1.2. Each of the plots shows only observations using the HiSPARC events data (i.e. coincidences between the detectors of a station).

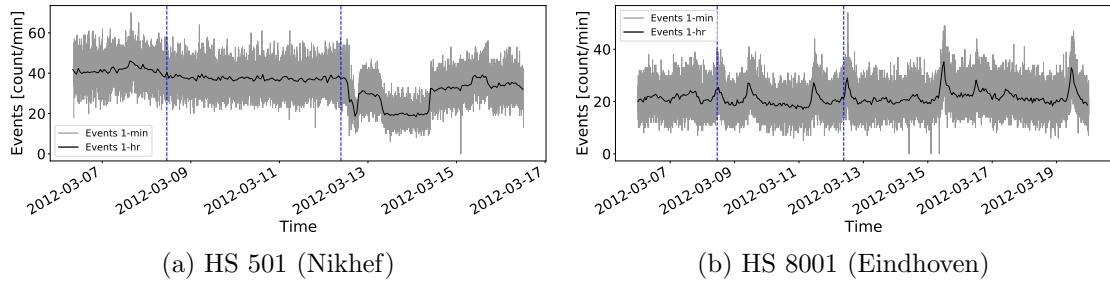


Figure 1.10: HiSPARC data for stations 501 and 8001 around the epoch of the FDs in March 2012. The plot shows the minute-averaged and hourly-averaged trigger events between detectors within the station. The vertical blue-dashed lines show the approximate onset-time of the FDs.

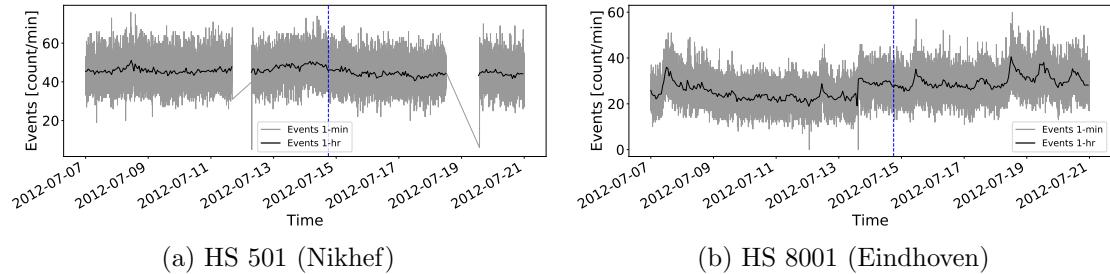


Figure 1.11: HiSPARC data for stations 501 and 8001 around the epoch of the FD in July 2012. The plot shows the minute-averaged and hourly-averaged trigger events between detectors within the station. The vertical blue-dashed line shows the approximate onset-time of the FD.

We can see from the plots that there are no clear signs of the anticipated FD signals in the HiSPARC observations. We observed a set of significant decreases in the muon count rate in station 501 after the second FD in March 2012 (see Figure 1.10a); however, it is unclear whether this is a consequence of the FD or other, hardware reasons, as the FD was not observed in the other HiSPARC station. The

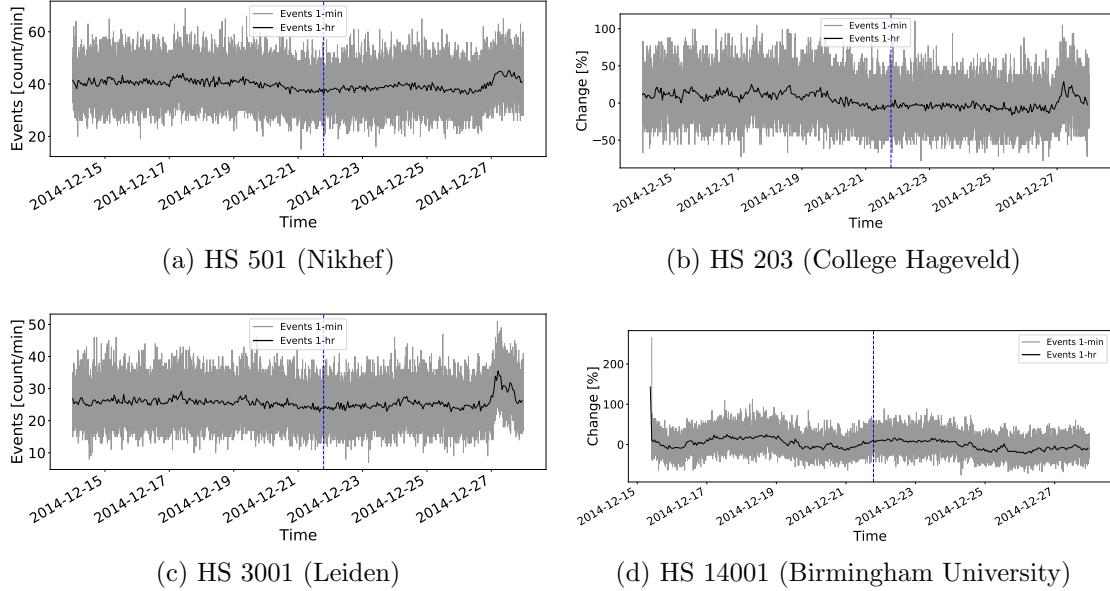


Figure 1.12: HiSPARC data for 4 stations around the epoch of the FD in December 2014. The plot shows the minute-averaged and hourly-averaged trigger events between detectors within the station. The vertical blue-dashed line shows the approximate onset-time of the FD.

shape of the FD in the NM data shows a sudden decrease and a smooth recovery within two days, but the shape of the HiSPARC data shows a more complicated effect, which suggests that the cause is not the FD, but rather a result of hardware.

In the other station we also observe some variations in the count rate which vary over longer time scales, but this is due to variations in the atmospheric pressure. Note that this needs accounting for and comes later...

It is quite clear from Figure 1.10b and Figure 1.11b that stations 8001 (Eindhoven) displays a semi-persistent diurnal variation in the count rate...

For the final two FDs listed in Table 1.2, the plot of the HiSPARC observations is shown in Figure 1.13. Plotted are the HiSPARC events data, and where possible, we also show the singles rates from each of the individual detectors in a station when the singles rate data is available. Furthermore, as these FDs were precursory to GLE72, we also marked on the epoch of the GLE for completeness.

As with the other FD epochs, we again do not observe any clear signs of the FD signals in either the events or singles data. In each of the three stations for which

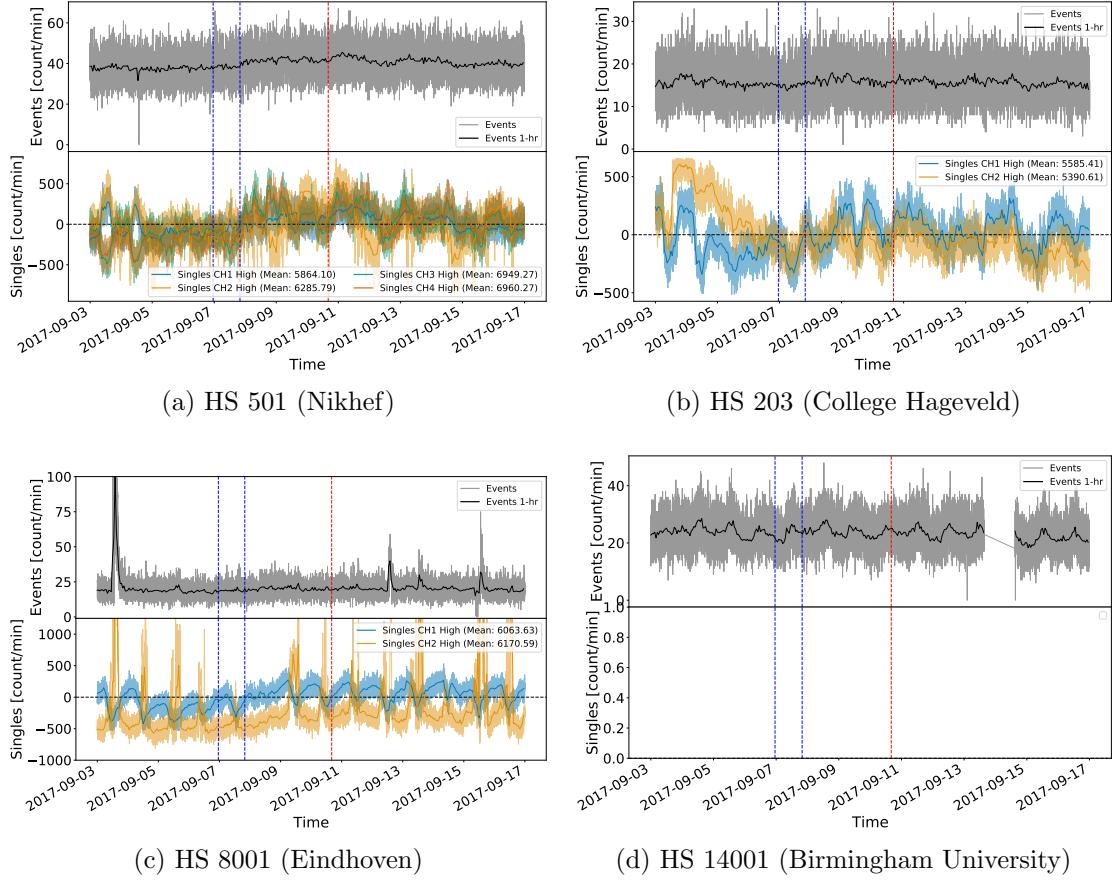


Figure 1.13: HiSPARC data for [n] stations around the epoch in which there were several FDs close to the onset of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical blue-dashed lines show the approximate onset-times of the two FDs observed around this epoch and the red-dashed line depicts the approximate onset time of the GLE.

there was singles data, we observed the semi-persistent diurnal signal, and again in the events for stations 14001. Furthermore, we also observed a similar slower variation in the count rate which is due to atmospheric pressure and needs to be accounted for.

No clear signal of FDs has been observed in the raw HiSPARC data. We again believe this could be due to the rigidity cut-off of the HiSPARC stations. We also note that the atmospheric effects in the raw do not help our ability to observe the space weather and these effects were later removed (see Section 1.6.2).

## 1.5 Air Shower Simulations

In order to understand the footprint of air showers produced by PCRs, simulations of air showers were performed for a range of PCRs energies for both primary protons and  $\alpha$ -particles.

To simulate the CR air shower development, the software Cosmic Ray Simulations for Kascade (CORSIKA) was employed: a Monte Carlo programme providing detailed simulations of the evolution of air showers initiated by PCRs through the atmosphere (Heck & Pierog, 2017). The particles in the CORSIKA simulations are tracked through the atmosphere until they undergo interactions with atmospheric nuclei, decay due to their instability, or reach the ground level defined as the simulation terminator.

Proton and  $\alpha$ -particle initiated air showers were generated with energies ranging from  $10^9$  to  $10^{20}$  eV, and  $4 \times 10^9$  to  $10^{20}$  eV, respectively. In total  $\sim 230000$  proton-initiated showers were simulated and  $\sim 180000$   $\alpha$ -particle-initiated air showers were simulated. The lists detailing the breakdown of PCR energies and number of simulations is provided in Appendix A, along with a brief discussion of the settings chosen within the simulations.

### 1.5.1 Air Shower Footprints

The average footprint of muons at ground level was acquired from the output CORSIKA simulations, by taking the distribution of the muons at ground level at the end of the simulation as a function of their distance from the shower core for each individual simulation realisation. For a given PCR energy, the average footprint distribution is calculated by combining all of the individual simulation realisations. Figure 1.14 shows the distributions for air showers induced by vertically incident protons and  $\alpha$ -particles.

The interpretation of Figure 1.14 provides an understanding of the minimum energy PCRs observable by the different stations within the HiSPARC network.

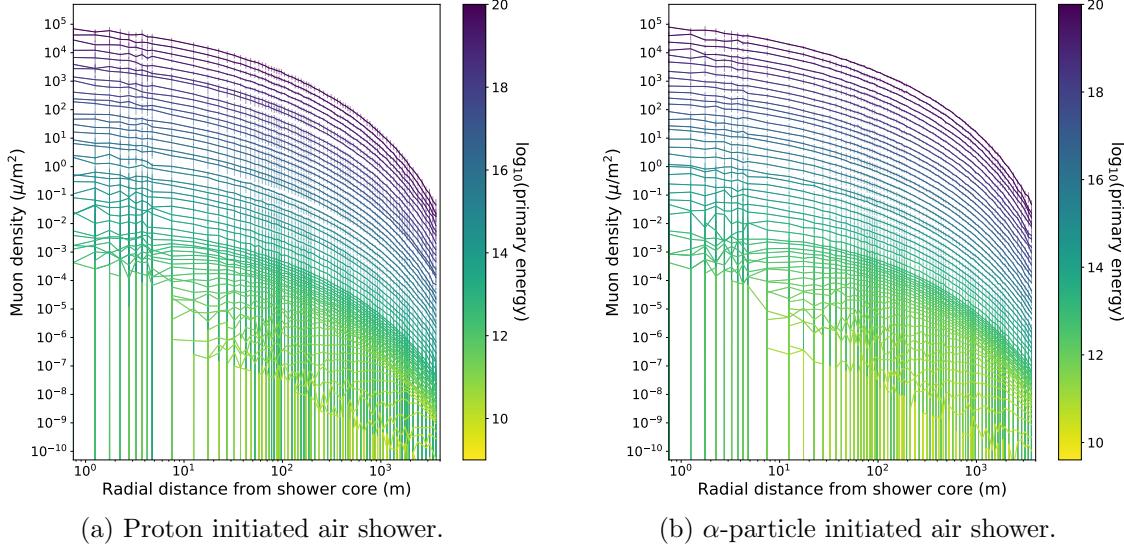


Figure 1.14: Mean muon density footprints for (a) proton-initiated air showers and (b)  $\alpha$ -particle-initiated air showers with initial PCR trajectories with zenith angles  $\theta = 0^\circ$  and various PCR energies. The error bars given represent  $1\sigma$ .

The typical separation between the detectors in a HiSPARC station is  $\sim 10$  m; however, the separation between detectors varies from station-to-station and can be up to as much as some 20 m or as low as just a couple of metres.

This analysis of the air shower footprint shows the variation in PCR energy sampled varies marginally over this range of distances and suggests that HiSPARC stations will typically observe PCR with an energy of  $\sim 10^{14} - 10^{15}$  eV to meet the required trigger conditions.

### 1.5.2 Muon Flux

From the air shower simulations it is possible to gain an estimate of the muon flux at ground-level based on the number of

## 1.6 Standardisation of HiSPARC Data

### 1.6.1 Motivation

- HiSPARC stations are individually managed and guidelines aren't stringent - Variability between stations exists and also apparently between detectors within a station

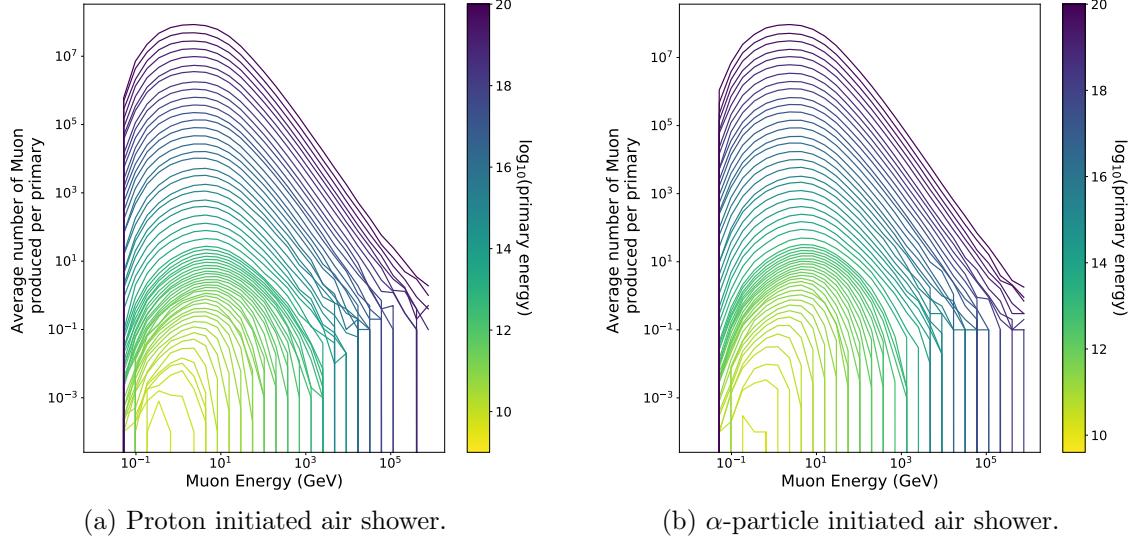


Figure 1.15: Mean number of muons produced at ground level by the PCR for (a) proton-initiated air showers and (b)  $\alpha$ -particle-initiated air showers, for various PCR energy. The uncertainty bars given represent  $1\sigma$ .

(i.e. see singles during GLE 72)

### 1.6.2 Barometric Correction

It is understood that observations made by ground-based CR detectors are susceptible to atmospheric conditions. Atmospheric pressure effects the CR travel path due to the expansion and contraction of the atmosphere with varying pressure; hence the CR counts are observed to be negatively correlated to atmospheric pressure as shown for both NMs and MDs in Figure 1.16. A correction for this barometric effect is routinely applied as part of the data calibration for all NM stations within the NMDB NEST, but there is no such process in the HiSPARC network.

The method of correcting for the barometric effect is discussed widely in the literature regarding NMs and is shown to depend on the barometric coefficient. Assuming the cosmic ray flux variation, absent of the atmospheric effects, is reasonably stable, then a simple corrected can be made. The CR variations ( $N$ ) that depend on the local atmospheric pressure are described by equation (1.2), where  $\Delta N$  is the change in count rate,  $\beta$  is the barometric coefficient, and  $\Delta P = P - P_0$  is the deviation in pressure from the average ( $P_0$ ) in the given time-period (Paschalis et al.,

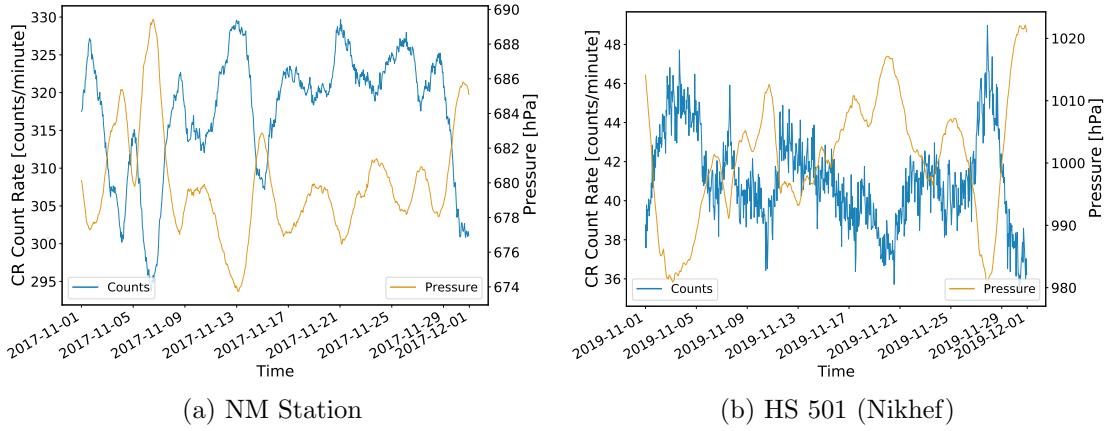


Figure 1.16: The anti-correlation between CR count rates and the atmospheric pressure. (a) shows the CR and the local atmospheric pressure measured at a NM in the South Pole; (b) shows the CR and pressure measured by HiSPARC station 501.

2013):

$$\Delta N = -\beta N \Delta P \quad (1.2)$$

Through the integration of equation (1.2), the solution shows the dependence of cosmic ray intensity on pressure as given in equation (1.3).

$$N = N_0 e^{-\beta \Delta P} \quad (1.3)$$

Therefore by taking the logarithm of equation (1.3), one can obtain the barometric coefficient by fitting the straight line given by equation (1.4) to the observed data, where  $N_0$  may be assumed as the mean count rate over the given time-period of observations considered.

$$\ln \left( \frac{N}{N_0} \right) = -\beta \Delta P \quad (1.4)$$

A demonstration of the barometric correction method of fitting a straight line to the data described by equation (1.4) is shown for both a NM and a HiSPARC station in Figure 1.17.

An online barometric coefficient tool is available which allows user to perform the

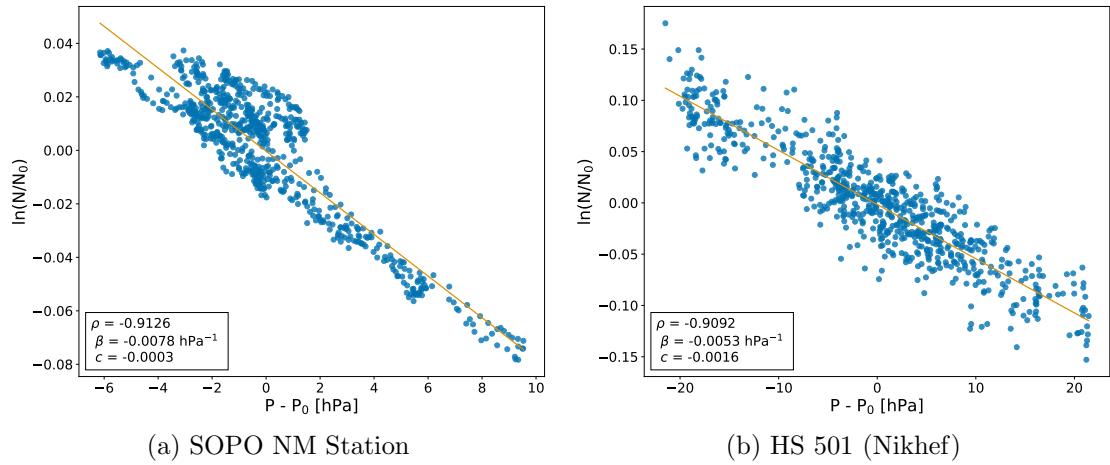


Figure 1.17: The barometric coefficient calculation: (a) during November 2017 for the South Pole (SOPO) NM station, (b) during November 2019 for HiSPARC station 501 at Nikhef.

barometric correction for a given station over a user-defined epoch (<http://cosray.phys.uoa.gr/index.php/data/nm-barometric-coefficient>). Using this tool, it was possible to provide a comparison between the method used in this work to that of the online NM barometric correction tool which is used for the correction of the NMDB stations. This is provided in Figure 1.18 for monthly corrections throughout 2017 for the NM station at the South Pole (SOPO).

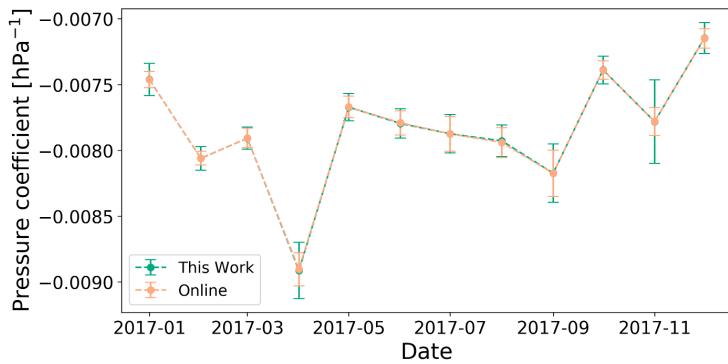


Figure 1.18: A comparison between the monthly barometric coefficient computed in this work and using the online barometric coefficient tool throughout the year 2017 for the SOPO NM station.

Figure 1.18 shows a close agreement between the barometric coefficient calculated in this work and those acquired from the online tool for the SOPO NM. This was also true for other stations tested (APTY and ROME), thus providing confidence

that the method used in this work was suitable for application on the HiSPARC data. The barometric correction was performed on the stations where sufficient pressure data and count rates exist, and were re-investigated to determine whether the space weather events were observed in the HiSPARC data. These results are provided in Section 1.7.

### 1.6.3 Temperature Correction

It has been discussed in the literature that the effect of atmospheric temperature on muon intensity has to be treated differently to the pressure effect (Berkova et al., 2011), as the temperature influences both the creation and disintegration processes for muons, such that there is a positive effect and a negative effect on muon intensity as a consequence of temperature variations (Mendonça et al., 2016).

The positive effect is related to pion decay and its dependence on temperature variation. The higher the temperature, the lower the atmospheric pion absorption, which implies a higher generation rate of muons (Mendonça et al., 2016).

The negative effect corresponds to the decrease of muon intensity at ground level as the muon average path length varies with temperature. Due to the heating and the expansion of the atmosphere during summer periods muons are produced higher in the atmosphere; hence the muon propagation path increases meaning more atmosphere for muons to traverse before reaching the ground, and an increased decay probability and ionisation losses (Savi et al., 2015; Mendonça et al., 2016).

Due to the difference in decay probability, the negative effects dominate for low energy muons (i.e. those detected by ground-level MDs), and the positive effect dominates for high energy muons (i.e. those detected by underground MDs) (Berkova et al., 2011); therefore it is expected that the negative effect should dominate for the HiSPARC network. Temperature effects are also observed by NMs; however the effect is less significant than for MDs hence temperature corrections are not widely applied for NMs (Mendonça et al., 2016).

This is in contradiction with the observations of diurnal variation with the HiSPARC detector, as one can quite clearly see that the HiSPARC stations register higher count rates during local noon.

Several methods of correcting for the negative temperature effect are summarised by Berkova et al. (2011) which utilise different measures of atmospheric temperature when performing the temperature correction. Mendon{\c{c}}a et al. (2016) provides a comparative summary of these methods applied to correct for atmospheric temperature variations observed by GMDN detectors. The methods discussed here however are typically applied over long timescales of years with low temporal resolution rather than to account for short timescale variations with periods of less than a day; hence the suitability of these methods is uncertain.

Mendon{\c{c}}a et al. (2016) concludes that correcting for temperature using the atmospheric mass weighted temperature is one of the most suitable methods for the GMDN as it allows for the highest correlation between long-term CR variations and temperature. The mass weighted method is an approximation for integrating over the vertical atmospheric temperature as is given in Eq. (1.5):

$$\left( \frac{\Delta N}{N} \right)_T = \bar{\alpha} \int_0^{h_0} \delta T(h) dh = \sum_{i=0}^n \frac{x(h_i) - x(h_{i+1})}{x(h_0)} T(h_i) = \alpha_{\text{MSS}} \delta T_{\text{MSS}} \quad (1.5)$$

where  $h_0$  is the closest to ground altitude;  $\delta T_{\text{MSS}}$  is the deviation of the mass weighted atmospheric temperature;  $T(h_i)$  is the temperature in degrees kelvin observed at the altitude  $h_i$ ;  $x(h_i)$  is the atmospheric depth at the altitude  $h_i$  which is given by Eq. 1.6:

$$x(h) = \int_h^\infty \rho(h) dh \quad \rho(h) = \frac{P(h)}{T(h)} \frac{M_{\text{mol}}}{R} \quad (1.6)$$

where  $P(h)$  is the atmospheric pressure profile as a function of depth;  $T(h)$  is the atmospheric temperature;  $\rho(h)$  is the air density at a given altitude  $h$ ;  $M_{\text{mol}}$  is

the molar mass of air;  $R$  is the universal gas constant.

The temperature correction is therefore used in a formalism the same as Eq. (1.4), replacing replacing  $\beta$  for  $\alpha$  and  $\Delta P$  for  $\Delta T$ .

In addition it is discussed by Berkova et al. (2011) and Mendon{\c{c}}a et al. (2016) that the effective generation level temperature is a suitable assumption for this purpose. This method is based on the assumption that muons are mostly generated at a certain isobaric level, taken as 100 mbar, and therefore the temperature at 100 mbar in the atmospheric pressure profile is used,  $T_{100\text{ mbar}}$ .

As discussed above, Mendon{\c{c}}a et al. (2016) provide this as a method for correcting for the long-term variation in atmospheric temperature which varies seasonally rather than to correct for diurnal variations; therefore it is unsure how relevant this method of atmospheric temperature correction will be to the diurnal variations observed in the HiSPARC data.

## 1.7 HiSPARC Observations After Pressure Corrections

### 1.7.1 Pressure Corrected Observations of Ground Level Enhancements

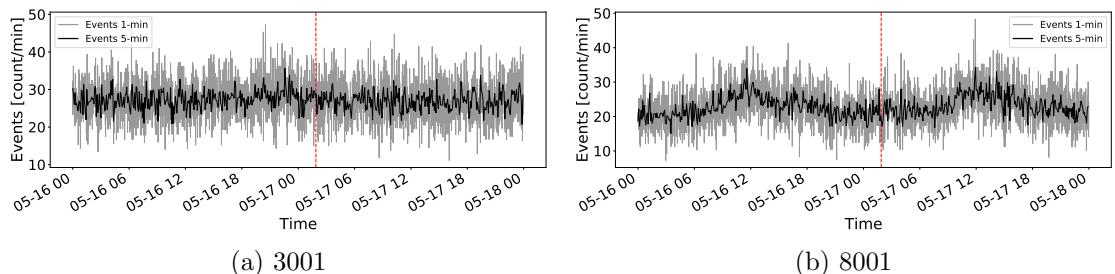


Figure 1.19: Pressure corrected HiSPARC data for 2 stations around the epoch of GLE 71. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical red, dashed line depicts the approximate onset time of the GLE.

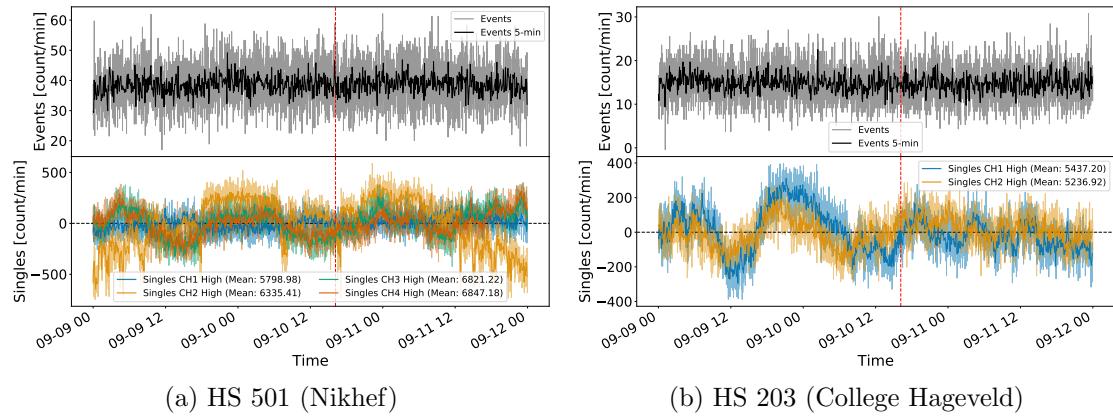


Figure 1.20: Pressure corrected HiSPARC data for 2 stations around the epoch of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical red, dashed line depicts the approximate onset time of the GLE.

## 1.7.2 Pressure Corrected Observations of Forbush Decreases

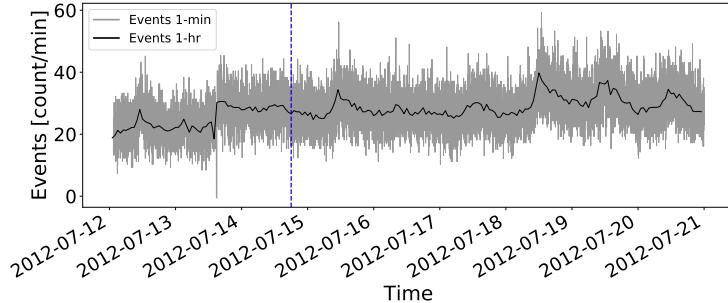


Figure 1.21: ...

## 1.8 Discussion

Throughout this chapter the feasibility of using the HiSPARC network of muon detectors has been analysed. This has involved performing cosmic ray air shower simulations using CORSIKA and performing backwards

## 1.9 Conclusion

• • •

We leave the reader with the following points:

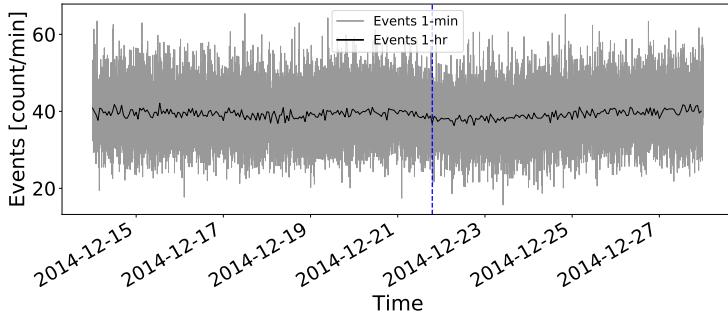


Figure 1.22: ...

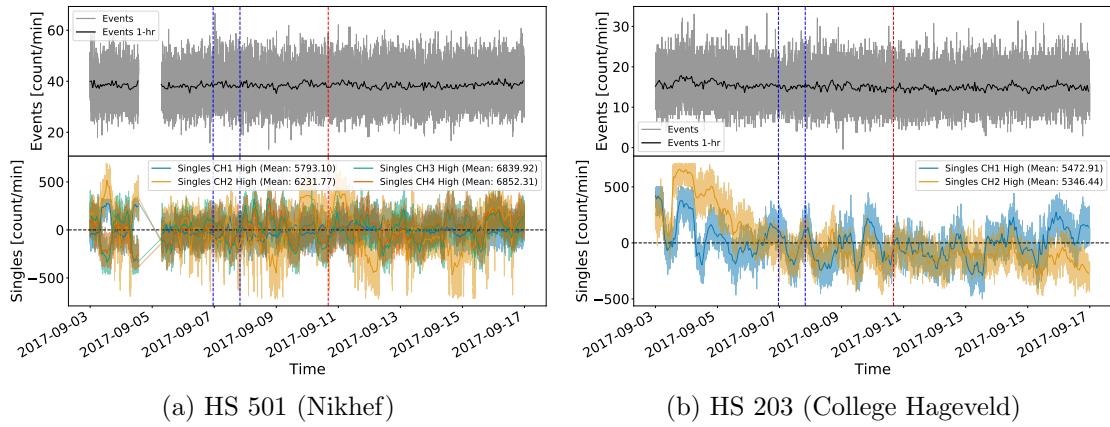


Figure 1.23: Pressure corrected HiSPARC data for [n] stations around the epoch in which there were several FDs close to the onset of GLE 72. The top panel of each subplot shows the minute-averaged trigger events between detectors within the station, while the bottom panel shows the mean-shifted, minute-averaged counts by each individual detector in the station. The vertical blue-dashed lines show the approximate onset-times of the two FDs observed around this epoch and the red-dashed line depicts the approximate onset time of the GLE.

1. ...

2. ...

3. ...

## 2 HiSPARC Station 14008

### 2.1 Introduction

It was clear from the work covered in Chapter 1, using data from the HiSPARC network, that the HiSPARC detectors were not clearly capable of observing space weather events and this is also hindered as they are rather sensitive to variations in the terrestrial conditions.

To some extent, it was possible to eliminate the variation in CRs due to terrestrial variation from the HiSPARC data; however it was shown to be not always so simple, as different detectors in the HiSPARC network showed different responses to pressure and temperature variation. The non-linear relationship between temperature and CR count means the correction of the count rate due to thermal fluctuations is non-trivial, unlike the counterpart correction for pressure.

It is believed that the atmospheric thermal fluctuations induce thermal noise in the Photo Multiplier Tubes (PMTs), and although the temperature inside the HiSPARC ski-boxes have not been measured, it is suspected that the PMTs can get quite hot, in particular when the sky boxes are in direct sunlight.

An instance of thermal noise in a single PMT will be random, and uncorrelated with an instance of thermal noise in another PMT. It is therefore possible to hypothesise that it is unlikely that within the coincidence window of  $\sim 1.5 \mu\text{s}$ , that a coincidence between 2 PMTs would be due to random thermal noise induced in the PMTs.

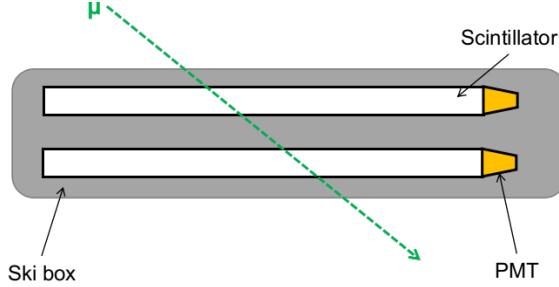


Figure 2.1: Schematic diagram of the HiSPARC station 14008 detector set-up.

To exploit this, it is possible to stack 2 detectors on top of each other to measure a single muon which traverses both scintillators, hence inducing signals in both PMTs.

## 2.2 Aims

The aim of creating a new HiSPARC station was to test whether an alternative configuration of HiSPARC station could minimise atmospheric deviations in the data and allow for the observation of space weather events...

## 2.3 HiSPARC 14008 Detector Set-up

### 2.3.1 Configuration

The configuration of HiSPARC station 14008 is shown in Figure 2.1; the station is composed of two scintillators stacked on top of each other, inside one ski-box.

Each of the plastic scintillators has a thickness of  $\Delta x = 0.5$  cm, and density,  $\rho = TBC \text{ gcm}^{-3}$ . We know that the stopping power of the scintillator for a minimum ionising particle is  $\sim 2 \text{ MeV/cm}^{-2}$  [REF]... [discuss this w/ Angela as it seems to be derivative of Blethe-Bloch, but cannot find link]...

$$E = \Delta x S \rho \cos(\theta) \quad (2.1)$$

Bartels (2012) state that typical energy loss of a muon in a single scintillator is 3.38 MeV, hence in this configuration, as a muon traverses two scintillators, the

lower limit on the energy loss by muons in the detector is  $\sim 6.76$  MeV.

To protect the scintillators and PMTs within the ski boxes, we sandwiched the scintillators between layers of foam, as can be seen on the lab work bench in Figure 2.2a. Upon complete assembly of the detector, the scintillators and PMTs are placed within the ski-box on the roof of the Poynting Physics building on the campus of the University of Birmingham, as shown in Figure 2.2b.

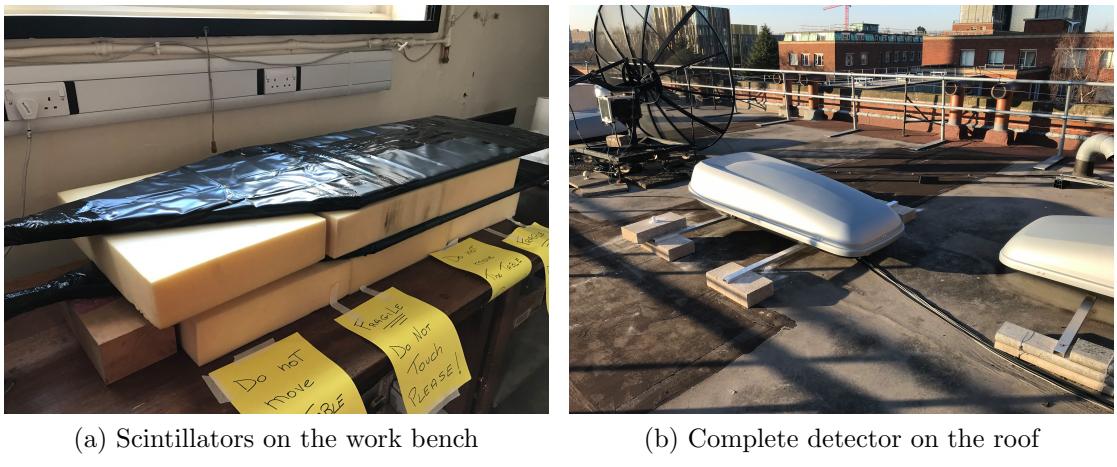


Figure 2.2: HiSPARC 14008 assembly and configuration. (a) shows the stacked arrangement of the scintillators on the lab work bench, between layers of protective foam. (b) shows the complete detector inside the ski-box on the University of Birmingham campus.

### 2.3.2 Calibration

When setting up the HiSPARC station, it was required to set several operating parameters for the detectors and the HiSPARC electronics box. One such setting was the PMT operating voltage. Each of the detector PMTs needs to be powered with a high enough operating voltage such to provide an amplified signal, but not too high such as to over-amplify the noise.

In general, the PMTs has an advised operating voltage of around 700 V (Fokkema, 2019); however, best practise is to operate the PMT at the plateau region, whereby the counts/voltage no longer increases. As can be seen from Figure 2.3, neither of the PMTs have clear plateau regions, hence there was no obvious PMT set point.

The HiSPARC installation manual does, however, suggest to tune the PMT

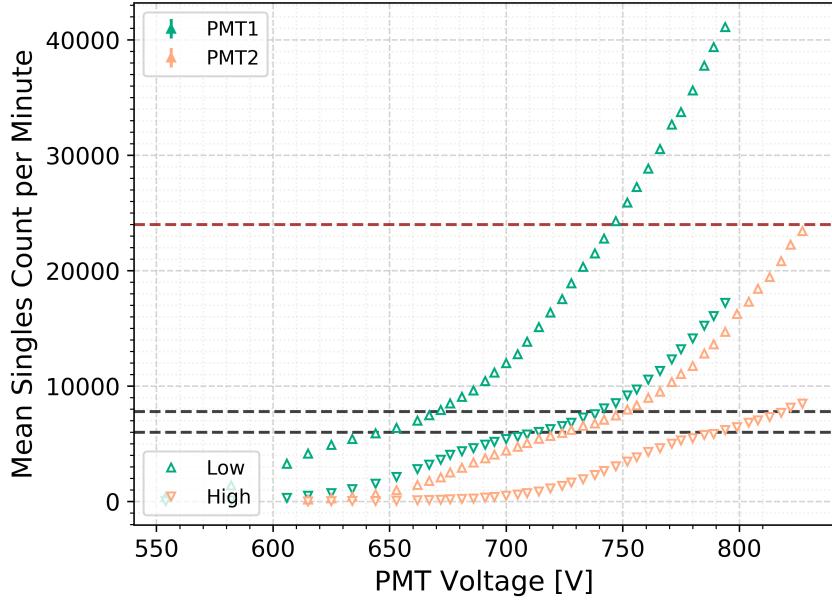


Figure 2.3: Voltage calibration curve for the PMTs of station 14008. The upper, red-dashed line indicates the upper limit for the low threshold singles rate (400 Hz), and the lower 2, black-dashed lines indicate the upper and lower bounds for the high threshold singles rate (100–130 Hz).

voltages such that the singles rates for each detector meet the following criteria: singles rate of 100–130 Hz for signal above the high trigger threshold, and singles rate of <400 Hz for signal above the low trigger threshold (Fokkema, 2019).

In order to calibrate the PMTs to the correct level, we measured the singles rates above the high and low thresholds as a function of PMT operating voltage, as is shown in Figure 2.3 [UPDATE THIS PLOT...!!!!]. The voltage calibration plot shows drastically the different performances one can get from different PMTs, therefore it is necessary to treat each PMT individually when calibrating.

### 2.3.3 Monitoring Temperature

It was suspected in Chapter 1 that the singles count rates (and thus event count rates also) are affected by the temperature of the PMT within the HiSPARC ski-boxes.

Some of the existing HiSPARC stations monitor local temperature however none measure the temperature of the PMT within the ski box; therefore the temperature of the PMT itself is unknown, and thus we cannot account for the thermal noise.

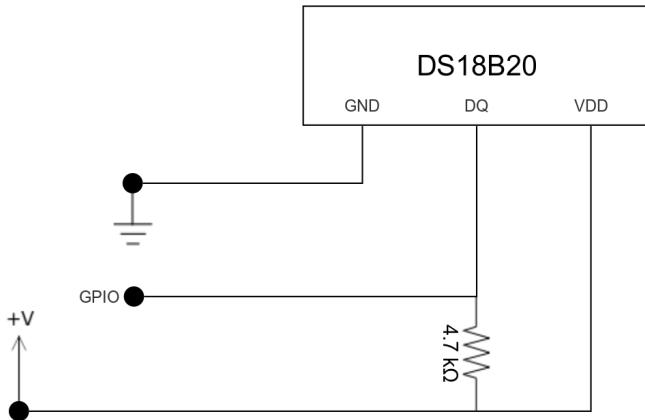


Figure 2.4: Schematic diagram of the DS18B20 temperature sensor circuit, whereby the voltage, ground, and GPIO interfaces connect directly into pins of the Raspberry Pi board.

When building this new HiSPARC station, a temperature sensor was placed into the ski box which allowed us to monitor the temperature.

Figure 2.4 shows the schematic for the temperature sensor. We used the DS18B20 temperature sensor with the one-wire telemetry protocol, which used a single wire to transmit the temperature readings to the microcontroller; the microcontroller used was a Raspberry Pi 4. Three wires were used for the operation of the DS18B20: constant current voltage, ground, and data.

We use the Raspberry Pi to control the data acquisition by running a Python program to output the temperature readings from the sensor to a local file. The temperature is read on a 10-second cadence and is recorded in degrees Celsius.

## 2.4 Observations

## 2.5 Conclusions

# 3 Galactic Cosmic Ray Behaviour During Solar Cycle 24

*The majority of the text in this chapter is taken from Ross & Chaplin (2019). I was first author on this journal article and conducted the work of the investigation. Section [CITE] has been written since the publication of Ross & Chaplin (2019) to document the continuation of this analysis to the end of Solar Cycle 24.*

## 3.1 Introduction

Galactic cosmic rays (GCRs) are charged particles and atomic nuclei with energies spanning the range from a few MeV up to approximately  $10^{21}$  eV, that encroach upon the Earth from all directions (Giacalone, 2010). They mainly originate outside the solar system, within the Milky Way; however they are also expected to originate from other galaxies (Aab et al., 2017). GCRs at the top of the atmosphere are mostly composed of protons ( $\sim 87\%$ ) and  $\alpha$ -particles ( $\sim 12\%$ ), with a smaller contribution ( $\sim 1\%$ ) from heavier nuclei (Dunai, 2010).

When cosmic rays (CRs) enter the atmosphere, they interact with atmospheric atoms and produce cascades of secondary particles, which at ground level are primarily neutrons and muons. Neutron monitors (NMs) and muon detectors (MDs) located at different locations on Earth have been used since the 1950s to observe GCRs. Information on GCRs prior to the modern epoch of NMs and MDs, and the space age, rely on the studies of cosmogenic isotope records from ice cores and tree

rings (Owens & Forsyth, 2013).

It has long been established that there exists an anti-correlation between GCR intensity and the level of solar activity, over a cyclic 11-year period, with perhaps some time-lag (Forbush, 1958; Parker, 1965; Usoskin et al., 1998; Van Allen, 2000). Figure 3.1 shows clearly the anti-correlation between GCRs and sunspot number (SSN).

It is well known that the 11-year solar activity cycle is in fact a 22-year cycle - the Hale cycle - which describes the alternating polarity of the large-scale solar magnetic field (Thomas et al., 2014a). The interchanging peaked and flat-topped shape of GCR intensity in Figure 3.1 is a manifestation of this effect in addition to other CR transport processes (Aslam & Badruddin, 2012).

The polarity of the solar field,  $A$ , is taken to be negative when the field axis is aligned with the axis of rotation, and positive when the opposite is true (Thomas et al., 2014a). The solar field polarity conventionally is described in combination with particle charge,  $q$ , due to the effect of curvature and gradient drift on charged particles; thus it is customary to define the solar polarity as  $qA$ . Vertical lines showing the approximate epochs at which the polarity reverses are plotted in Figure 3.1 (Janardhan et al., 2018; Thomas et al., 2014a).

Particle drifts differ during different  $qA$  cycles, with positive CRs (i.e. protons) predominantly arriving into the heliosphere from the heliospheric poles and outwards to Earth during periods when  $qA > 0$ , whereas when  $qA < 0$ , positive CRs predominantly arrive at Earth inwards along the heliospheric current sheet (HCS) (Belov, 2000; Thomas et al., 2014b). As the solar magnetic dipole axis is tilted to the solar rotation axis, so is the HCS; the HCS tilt varies with solar cycle and is typically smaller during solar minimum and larger during solar maximum (Owens & Forsyth, 2013). The tilt angle of the HCS has also been shown to be strongly correlated to the GCR intensity and the GCR lag behind the solar activity (Belov, 2000; Mavromichalaki et al., 2007).

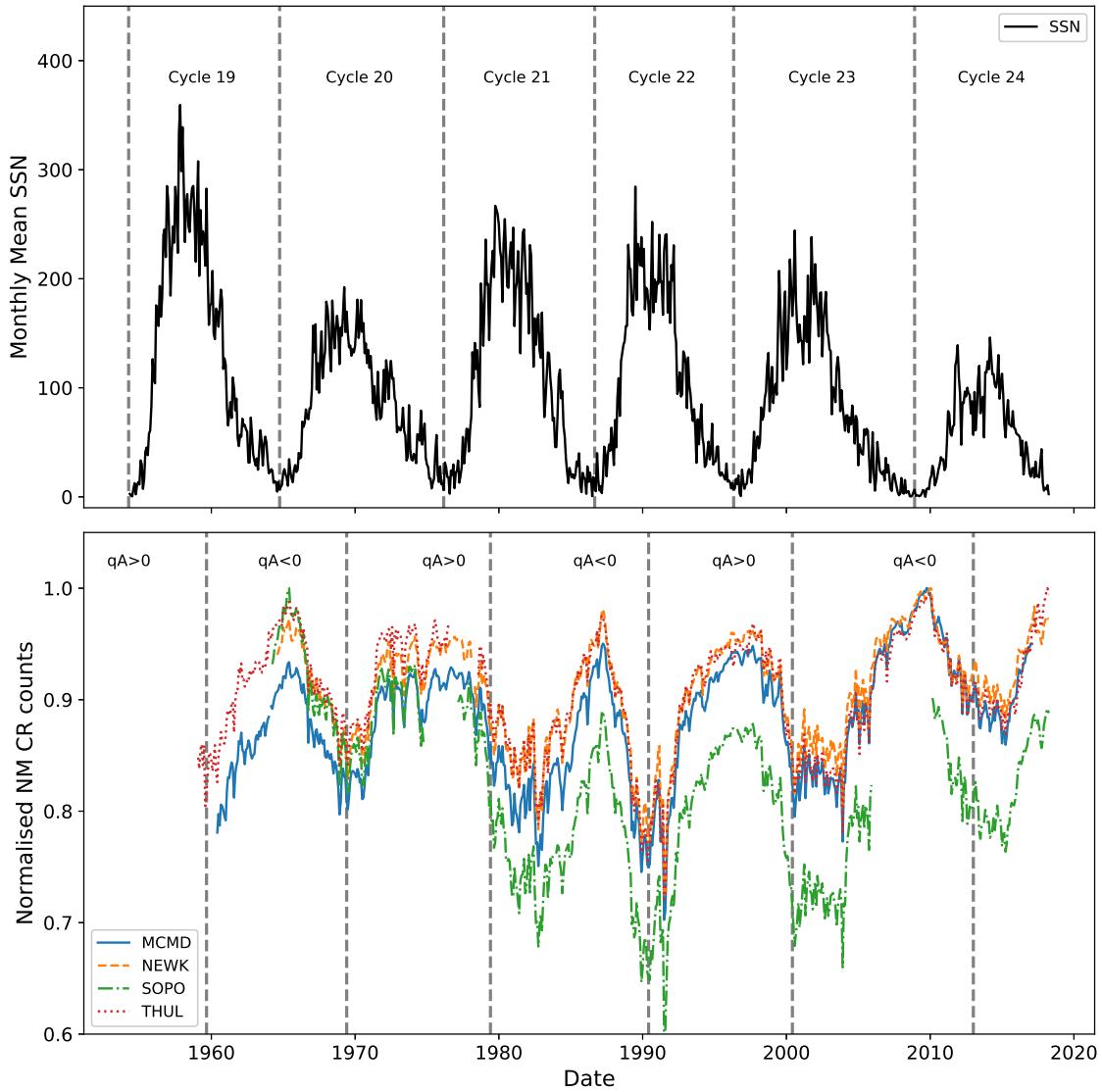


Figure 3.1: SSN (top), with vertical lines showing the beginning of each solar cycle. Cosmic ray intensity recorded by NM (bottom), with vertical lines showing the approximate epochs of solar magnetic field polarity reversals. (MCMD = McMurdo, NEWK = Newark, SOPO = South Pole, THUL = Thule).

Aslam & Badruddin (2012, 2015) found that the different processes of CR transport have varying levels of importance throughout the solar activity cycle, but around solar maximum it is likely that drifts play less of a role and disturbances in the solar wind (and hence HCS) are the predominant factor of CR modulation. Even cycles encounter  $qA < 0$  polarity during their onset phase and  $qA > 0$  during their declining phase, thus experiencing a faster GCR recovery after solar maximum as the GCRs predominantly enter the heliosphere from the heliospheric poles and experience an outwards drift towards Earth. Odd cycles encounter  $qA > 0$  polarity during their onset phase and  $qA < 0$  during their declining phase and so experience a slower recovery after solar maximum, as the GCRs predominantly enter the heliosphere along the HCS. When the HCS is tilted and disturbed during the declining activity phase, the path length that GCRs must travel to Earth increases; hence resulting in an increased time-lag.

Several studies have demonstrated the lag between GCR and solar activity proxies is approximately zero (i.e. no lag) during even solar cycles, and that there exists a lag of around a year or more during odd solar cycles (Usoskin et al., 1998; Mavromichalaki et al., 2007; Singh et al., 2008).

Van Allen (2000) showed through cross-plotting the annual mean intensity of GCRs against sunspot number (SSN) between 1953 and 1999 (covering solar cycles 19-22), that there is a distinct difference in the plot shapes between the different solar cycles, with 19 and 21 producing broad ovals, and 20 and 22 as approximately flat lines. The striking difference between odd-numbered and even-numbered cycles is shown for cycles 19-24 in Figure 3.2. It is believed that this hysteresis effect is caused by the combination of the heliospheric magnetic field (HMF), solar magnetic field polarity and thus the particle drift, and the tilt of the HCS leading to a slow recovery of GCR intensity after maxima in odd cycles and a fast recovery after maxima in even cycles (Van Allen, 2000; Belov, 2000; Thomas et al., 2014a).

An extension of this work has since been carried out by Inceoglu et al. (2014)

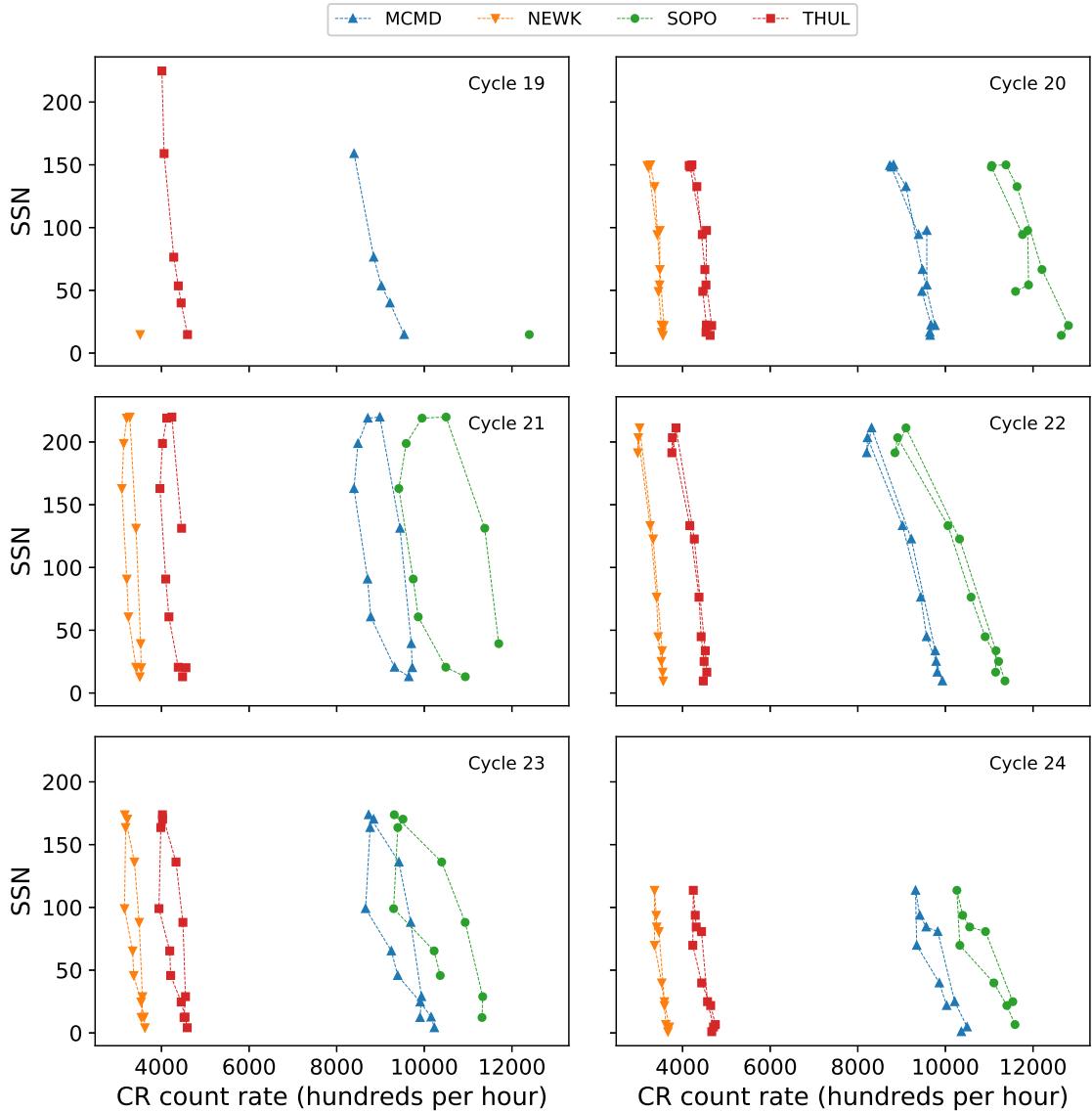


Figure 3.2: Hysteresis plots between yearly averaged SSN and yearly averaged GCR intensity for each of the 4 main NM stations over cycles 19-24.

showing that the even numbered solar activity cycles can be best modeled using a linear fit due to the narrow shape of the hysteresis loops; whereas odd-numbered solar activity cycles are better represented by ellipses due to their broader shape.

There has been speculation in the literature on the behaviour of cycle 24 compared to recent odd and even cycles. It has been suggested that there exists a lag between SSN maxima and GCR intensity minima in excess of 10 months (Kane, 2014; Mishra & Mishra, 2016) which does not follow the previous even cycles having a near-zero lag and in fact suggesting that cycle 24 behaved similarly to previous odd cycles; however these studies do not make use of a complete cycle of data and thus may draw inaccurate conclusions about the behaviour of the whole cycle because of the unusually extended nature of the declining phase of cycle 23 and the low amplitude of cycle 24 maximum (Broomhall, 2017). Mishra & Mishra (2016) make use of a more complete data set for cycle 24, yet still incomplete, and conclude that it is also likely that a 4 month lag could exist between GCRs and SSN.

This work aims to provide a timely update on the statistical relationship between GCR intensity and solar activity during solar cycle 24, since the cycle has now almost declined to a minimum. These aims have been achieved through a time-lag analysis and hysteresis effect analysis between SSN and GCR intensity.

In Section 3.2 we provide a brief description of the data that was used throughout this study for both CRs and SSN.

We show in Section 3.3 through a correlative time-lag analysis that there exists a small time-lag between the SSN and GCR intensity over solar cycle 24, which is slightly longer than preceding even-numbered cycles but not as high as observed in previous odd-numbered cycles. We also discuss whether the time-lag between SSN and GCR shows a dependence on the rigidity cut-off of the observing station.

In Section 3.4 we model the shapes of hysteresis plots between GCR intensity and SSN. We show that the behaviour of the hysteresis loops for cycle 24 follow the preceding even-numbered solar activity cycles and is better represented by a straight

line fit rather than an elliptical model.

## 3.2 Data

For the majority of the work in this study we have considered the pressure corrected count rates measured by four NM monitor stations as acquired from the NM data base (NMDB, 2018) event search tool (NEST) (<http://nmdb.eu/nest/>). The four stations are McMurdo (MCMD), Newark (NEWK), South Pole (SOPO), and Thule (THUL), i.e. the same NM stations used in the study by Inceoglu et al. (2014) to provide a comparison to existing literature. Table 3.1 details the basic characteristics of the NM stations used in this study.

We have investigated the long-term GCR modulation in the heliosphere from 1964-2018, spanning solar cycles 20-24, for the cycle epochs: 20: (10/1964 - 03/1976); 21: (03/1976 - 09/1986); 22: (09/1986 - 05/1996); 23: (05/1996 - 12/2008); 24: (12/2008 - 03/2018). Early predictions on solar cycle 25 suggest that solar cycle 24 is unlikely to reach a minimum earlier than the middle of 2019 up to as far as early 2021 (see Howe et al. (2018); Upton & Hathaway (2018); Pesnell & Schatten (2018)). The data used in this study are therefore of an incomplete cycle 24; however we believe this to now have a minimal effect on the results as cycle 24 draws to a minimum. Cycle 19 was omitted from this study due to the incomplete data set for this period (see Figure 3.1 and Figure 3.2).

During the time-lag correlation analysis, as our results suggested there may be a rigidity dependence on the time lag, we introduced a further 12 NM stations with data acquired from NEST spanning cycles 20-24 to increase the rigidity spectrum utilised in this study; these stations and their basic characteristics are also detailed in Table 3.1. These stations are not included in the rest of the results however as the results from these stations do not change the conclusions of this study.

Furthermore, we have also used monthly/yearly averaged SSN, as collected by WDC-SILSO (<http://sidc.be/silso/>), for the time-lag analysis/hysteresis anal-

Table 3.1: Neutron monitor stations used in this study and their vertical geomagnetic cut-off rigidity ( $R_c$ ), longitude, latitude, and altitude acquired from NEST. The first four stations have been used for all of the analysis while the lower 12 stations have been used exclusively for the investigation into the dependence of  $R_c$  on the time-lag.

	Station	$R_c$ [GV]	Long. [deg]	Lat. [deg]	$h$ [m]
Time-Lag & Hysteresis	McMurdo (MCMD)	0.30	166.6 E	77.9 S	48
	Newark (NEWK)	2.40	75.8 W	39.7 N	50
	South Pole (SOPO)	0.10	0.0 E	90.0 S	2820
	Thule (THUL)	0.30	68.7 W	76.5 N	26
$R_c$ dependence of time-lag	Oulu (OULU)	0.81	25.5 E	65.1 N	15
	Kerguelen (KERG)	1.14	70.3 E	49.4 S	33
	Magadan (MGDN)	2.10	151.1 E	60.0 N	220
	Climax (CLMX)	3.00	106.2 W	39.4 N	3400
	Dourbes (DRBS)	3.18	4.6 E	50.1 N	225
	IGY Jungfraujoch (JUNG)	4.49	7.98 E	46.6 N	3570
	Hermanus (HRMS)	4.58	19.2 E	34.4 S	26
	Alma-Ata B (AATB)	6.69	76.9 E	43.0 N	3340
	Potchefstroom (PTFM)	6.98	27.1 E	26.7 S	1351
	Mexico (MXCO)	8.28	99.2 W	19.8 N	2274
	Tsumeb (TSMB)	9.15	17.6 E	19.2 S	1240
	Huancayo (HUAN)	12.92	75.3 W	12.0 S	3400

ysis respectively as our chosen proxy of solar activity.

### 3.3 Time-Lag Analysis

#### 3.3.1 Method

To investigate the time delay between the modulation of GCRs compared to the solar activity, a time-lag cross-correlation analysis was performed between monthly mean GCR intensity and monthly mean SSN for each station, following the approach of Usoskin et al. (1998). We used a time window of width  $T$  centred on a time  $t$ , i.e. shifting within the interval  $t - T/2$  to  $t + T/2$ . Here we used  $T = 50$  months.

The window was shifted in steps  $\Delta t = 1$  month within this interval and for each step the Spearman's rank correlation coefficient ( $\rho$ ) between GCR intensity and SSN was calculated. The lag between GCR and SSN was then estimated by finding the peak correlation coefficient within the time interval  $T$ .

The results from the four main NM stations used suggested that there may be a relationship between the rigidity cut-off ( $R_c$ ) of a NM station and the time-lag for GCRs; hence four additional NM stations were introduced to determine whether this was so, as detailed above.

#### 3.3.2 Results

The correlation ( $\rho$ ) between monthly averaged GCR counts and SSN for different time-lags was calculated for cycles 20-23. The variation in  $\rho$  is presented in Figure 3.3, showing that for each cycle there is a time-lag corresponding to peak anti-correlation between GCR intensity and SSN. Table 3.2 summarises the time-lag with the highest correlation and the corresponding correlation coefficient for all stations in each individual solar cycle.

As previously reported in the literature, we see here that all of the NM stations clearly exhibit almost no lag during even solar cycles, and a longer lag varying between 11-21 months during odd solar cycles. There is a strong agreement between

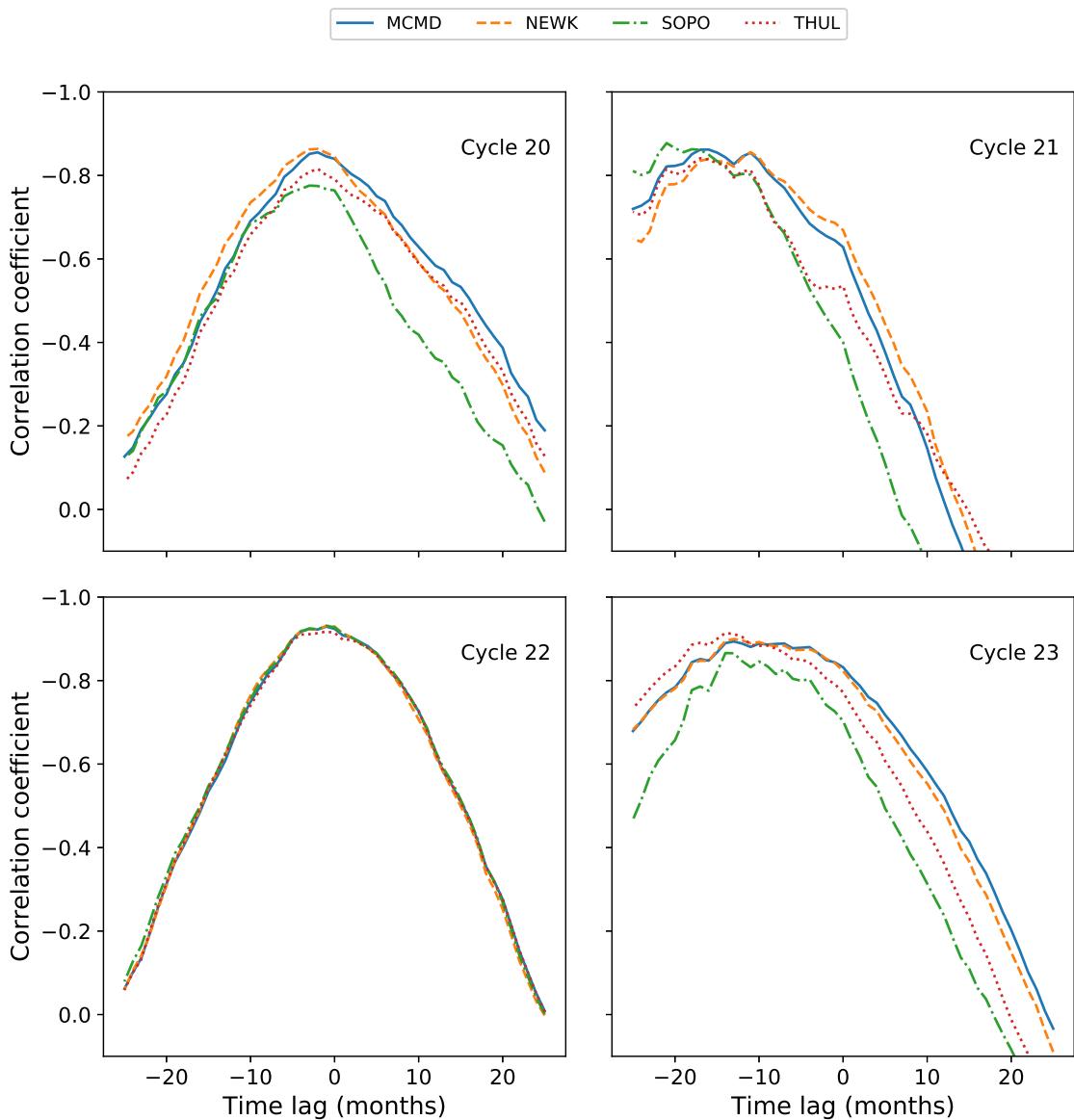


Figure 3.3: Variation in the correlation coefficient with time-lag NM station GCR intensity and SSN during solar cycles 20-23.

Table 3.2: Time-lags and the corresponding cross-correlation coefficient between NM CR count and SSN for solar cycles 20-23.

	Cycle 20		Cycle 21	
	Lag [months]	$\rho$	Lag [months]	$\rho$
McMurdo	2	-0.855	16	-0.862
Newark	2	-0.863	11	-0.856
South Pole	3	-0.776	21	-0.877
Thule	2	-0.816	17	-0.841
Cycle 22		Cycle 23		
	Lag [months]	$\rho$	Lag [months]	$\rho$
McMurdo	1	-0.929	13	-0.894
Newark	1	-0.931	13	-0.900
South Pole	1	-0.931	14	-0.866
Thule	1	-0.917	14	-0.914

the results presented in Table 3.2 and those of Mavromichalaki et al. (2007), Kane (2014), and Paouris et al. (2015), thus providing further evidence on the distinction between odd and even solar cycles due to particle transport in the heliosphere. The agreement with existing literature provides evidence of a suitable methodology in this study.

The same cross-correlation technique was then applied to cycle 24 between the dates 12/2008 - 03/2018 and the results are presented in Figure 3.4 and Table 3.3.

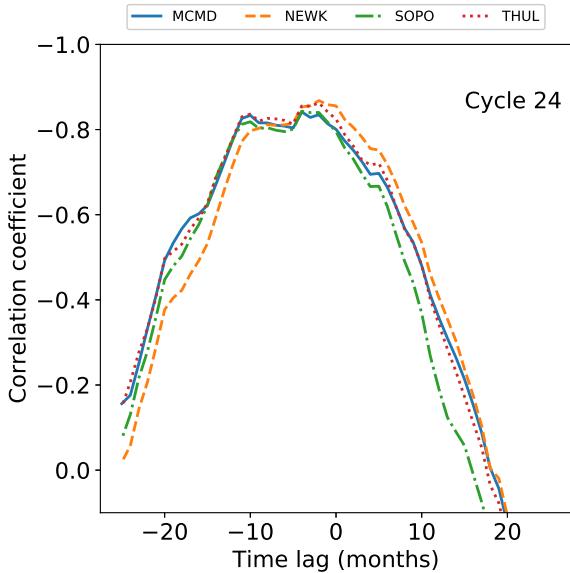


Figure 3.4: Variation in the correlation coefficient with time-lag between NM GCR intensity and SSN during solar cycle 24.

Table 3.3: Time-lags and the corresponding cross-correlation coefficient between NM GCR intensity and SSN for solar cycle 24.

Cycle 24		
	Lag [months]	$\rho$
MCMD	4	-0.841
NEWK	2	-0.868
SOPO	4	-0.843
THUL	2	-0.862

Cycle 24 is seen here to follow the pattern of almost no lag for even cycles; however cycle 24 does display a lag that is larger than the previous two even-numbered cycles, despite not being as large as the two previous odd-numbered cycles. The cause for the increased time-lag in cycle 24, as compared to the previous two even-numbered cycles, is likely due to the combined effects of the unusually deep and extended minimum between solar cycles 23 and 24, which delayed the decline in GCR intensity and caused record-breaking high GCR intensities (Pacini & Usoskin, 2015), and the small amplitude of the cycle 24 maximum.

The results presented in this study, using data for a near-complete cycle 24, show that the results of Kane (2014), and Mishra & Mishra (2016), were likely unduly influenced by the unusually deep and extended declining phase of cycle 23 given that they had a limited data set. Mishra & Mishra (2016) used data for just over half of cycle 24 and resulted in a time-lag of 4 months which agree with the results of this study.

As a further note on time-lag, Tomassetti et al. (2017) showed that through the introduction of time-lag as a parameter in the CR transport calculations of CR spectra that there exists a time-lag of  $8.1 \pm 1.2$  months during the period 2000-2012 spanning across cycle 23 and 24. We performed the time-lag analysis for the first 4 NM stations detailed in Table 3.1 for the period between 2000-2012 to investigate whether these results can be reproduced. The results of this analysis are presented in Figure 3.5 and Table 3.4.

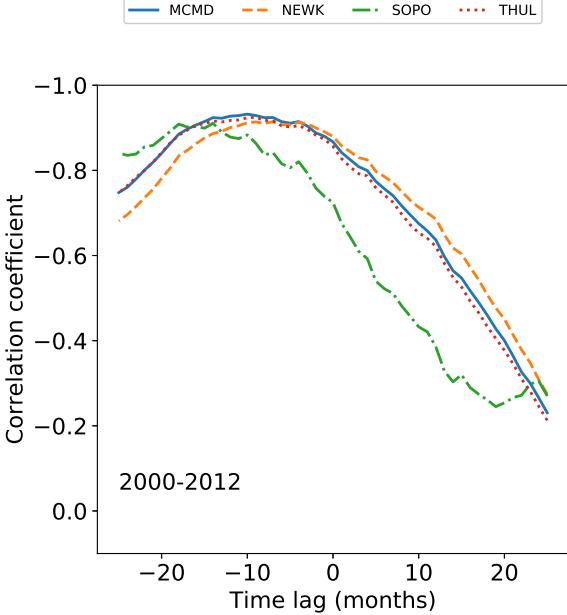


Figure 3.5: Variation in the correlation coefficient with time-lag between NM GCR intensity and SSN between 2000-2012.

From the time-lag analysis of the 4 stations presented there is a mean lag of  $10.00 \pm 1.47$  months, which is in good agreement with the results of Tomassetti et al. (2017).

Finally, allowing for the odd/even cycle dependence, we see in Figure 3.3, Figure 3.4, and Figure 3.5 that the time-lag appears to depend on the rigidity of the NM station used for observation. Such a dependence may impact the conclusions depending on the choice of NM station. We expect that if a dependence exists, a station with a higher rigidity cut-off ( $R_c$ ) would have a shorter lag as this station observes higher energy CRs which are affected less by solar modulation and thus able to recover faster from solar maximum. Whereas a station with a lower cut-off rigidity observing lower energy CRs, which are more influenced by solar modulation, would recover more slowly from solar modulation and therefore experience a longer time-lag. This is supported by Figure 3.3, Figure 3.4, and Figure 3.5, but in order to provide more conclusive evidence of such a relationship we introduced the additional NM stations detailed in Table 3.1 to provide a rigidity range spanning 0-13 GV. We present in Figure 3.6 a plot of the time-lag versus station  $R_c$  for all

Table 3.4: Time-lags and the corresponding cross-correlation coefficient between NM GCR intensity and SSN during 2000-2012.

2000-2012		
	Lag	$\rho$
	[months]	
MCMD	10	-0.932
NEWK	7	-0.914
SOPO	14	-0.911
THUL	9	-0.925

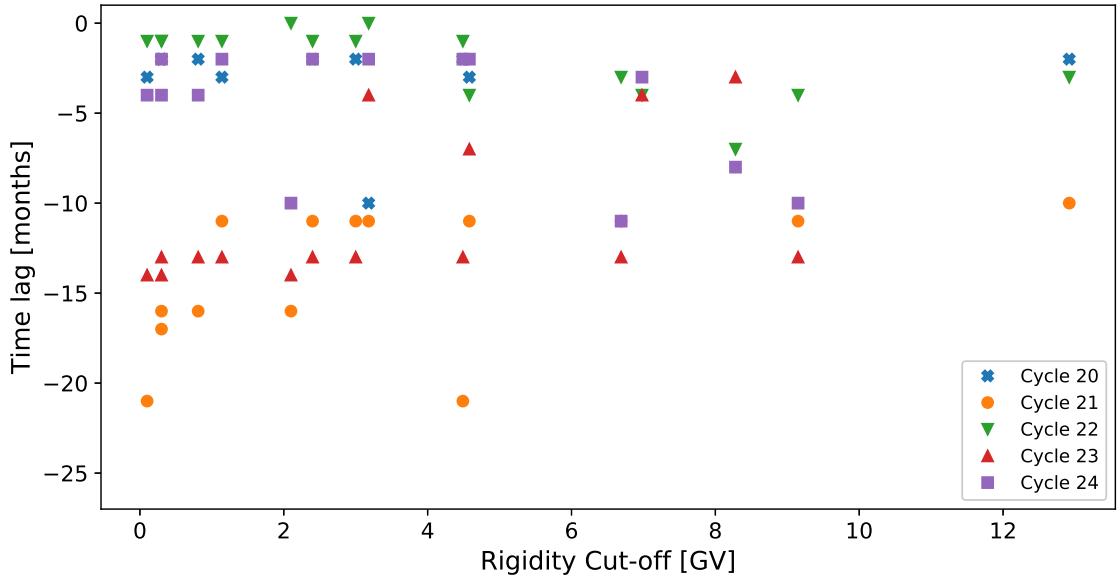


Figure 3.6: Variation in time-lag plotted against NM station rigidity cut-off for the 16 NM stations detailed in Table 3.1.

16 stations over cycles 20-24. To acquire uncertainties on the time-lag we ran 1000 Monte Carlo simulations of the time-lag analysis, sampling from the uncertainty distributions for each of the monthly averaged SSN and GCR counts; however the uncertainties in the data propagated in the Monte Carlo simulations produced no scatter in the overall results.

The results of this analysis do not show a clear rigidity dependence on the time-lag between SSN and GCR intensity; the sampling of higher  $R_c$  is too low, due to the availability of high  $R_c$  stations, to reasonably conclude on such a dependence at high rigidities, despite cycle 21 suggesting the existence of a dependence for low  $R_c$  stations as per our expectations. For low  $R_c$  stations there appears to be a more pronounced distinction between the time-lag observed between odd-numbered and even-numbered cycles than for higher  $R_c$  stations, however again this is not definitive due to the low sampling at higher  $R_c$ . We therefore conclude that there will be no significant dependence of the time-lag analysis on the  $R_c$  of the observing station.

## 3.4 Hysteresis Effect Analysis

### 3.4.1 Method

To investigate the hysteresis effect, we have adopted the approaches of Van Allen (2000), Singh et al. (2008), and Inceoglu et al. (2014). Plots of the annual mean SSN versus the annual mean GCR intensity were generated for cycle 20-24 and analysed by fitting different models to the data.

As highlighted in Inceoglu et al. (2014), even-numbered solar cycles can be suitably modeled by a linear fit due to their narrow hysteresis shape, and odd-numbered solar cycles were shown to be better modeled by ellipses due to their broadened hysteresis shape. Here, we first repeated for solar cycles 20-23 the linear and ellipse fitting to confirm the method reproduces the results reported in Inceoglu et al. (2014) before applying the method to cycle 24.

For even cycles which display narrow hysteresis loops, an unweighted least squares linear regression was used to reconstruct estimates of the GCR intensity from SSN. As odd-numbered solar cycles display a broader hysteresis loop, they were separately modeled using unweighted linear regression and ellipse fitting to determine the model providing the better fit.

The equation of the ellipse fitting took the form:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + R(\phi) \begin{bmatrix} a \cos \theta \\ b \sin \theta \end{bmatrix} \quad (3.1)$$

where  $x$  is GCR intensity;  $y$  is SSN;  $(x_0, y_0)$  are the centroid coordinates of the fitted ellipse,  $R(\phi)$  is the rotation matrix;  $\phi$  is the ellipse tilt angle;  $a$  and  $b$  are the semi-major and semi-minor axes respectively; and  $0 \leq \theta \leq 2\pi$  is the polar angle measured anti-clockwise from the semi-major axis.

In order to regain the GCR intensity from the model, where linear regression was used to model the data, GCR was acquired directly from the SSN for each year.

For the ellipse model the GCR was acquired from the model as a function of time from  $\theta$ , where the time-lag calculated from the analysis in Section 3.3 was used to correctly phase the ellipse allowing  $\theta$  to be calculated using standard equations of ellipses.

GCR intensity predicted by the linear regression and ellipse models were compared to the measured GCR intensity using Spearman's rank correlation as per Inceoglu et al. (2014).

### 3.4.2 Results

The hysteresis loops between yearly averaged SSN and GCR intensity for each station were first modeled with a linear regression for both odd and even solar activity cycles, then the odd cycles were separately re-modeled by ellipse fitting to show that this provides a more representative fit as suggested in Inceoglu et al. (2014). The correlation between measured CR intensities and modeled CR intensities for cycles 20-23 are presented in Table 3.5.

Table 3.5: Correlation coefficients of the linear regression and ellipse modeling of the hysteresis plots for solar cycles 20-23.

	Cycle 20		Cycle 21		Cycle 22		Cycle 23	
	Linear	Ellipse	Linear	Ellipse	Linear	Ellipse	Linear	Ellipse
McMurdo	0.867	-	0.664	0.946	0.964	-	0.846	0.852
Newark	0.888	-	0.700	0.964	0.955	-	0.857	0.874
South Pole	0.746	-	0.358	0.939	0.936	-	0.733	0.855
Thule	0.783	-	0.912	0.964	0.900	-	0.813	0.929

There is a consistent and good agreement between the measured and modeled CR intensities for even solar cycles modeled through linear regression, because the hysteresis loops are quite narrow as shown in Figure 3.7. These results support the findings of Inceoglu et al. (2014). Note that discrepancies in the correlation coefficients between this study and Inceoglu et al. (2014) are likely due to a number of reasons: Inceoglu et al. (2014) used data smoothing processes where in this study raw data are used; Inceoglu et al. (2014) made use of monthly mean data, whereas

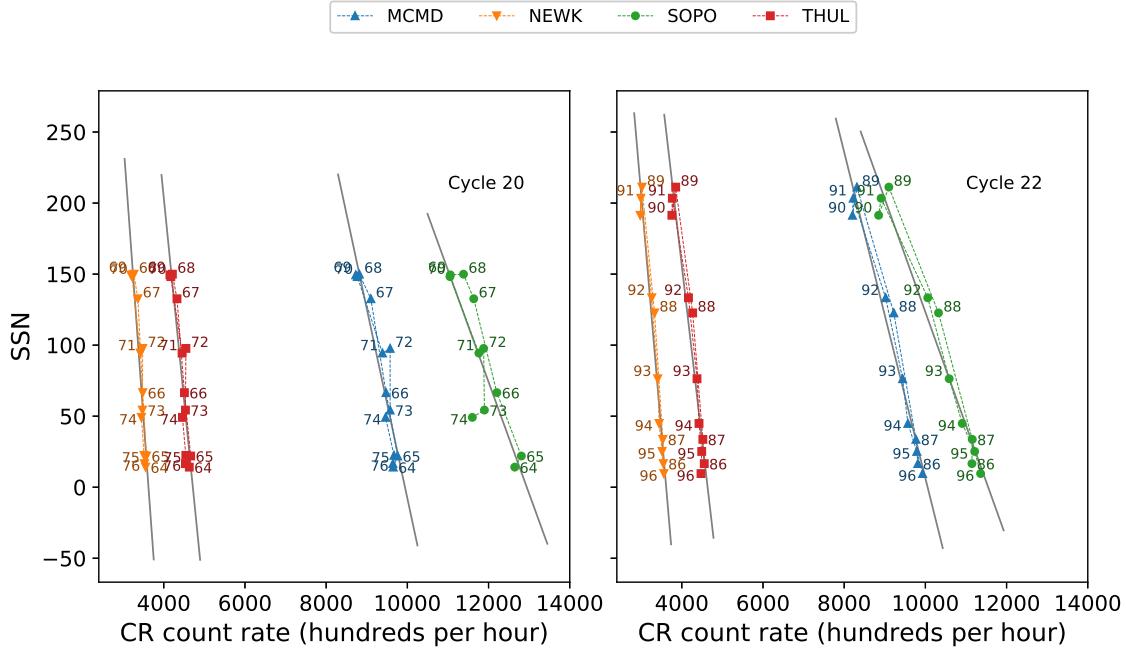


Figure 3.7: The hysteresis plots for even solar cycles 20 and 22 and the linear regression fit to the data.

annual mean data was used in this study; Inceoglu et al. (2014) interpolated missing data, whereas gaps have been left untreated in this study.

The linear relations for odd solar cycles are less consistent in their agreement with observed CR intensities, with the correlation during cycle 21 as low as 0.34 for South Pole and as high as 0.91 for Thule. Across both of the odd cycles considered in this study linear regression is not as good a representation of the data as for even cycles. Figure 3.8 shows the wider hysteresis loops which is a characteristic of odd solar cycles and shows visually that a linear fit does not provide a good representation of the data.

In agreement with the results of Inceoglu et al. (2014), it can be seen from the results in Table 3.5 that the ellipse models provide estimates of the CR intensity which are in good agreement with the measured intensities due to the increased correlation coefficient for each station during cycle 21 and cycle 23; for SOPO during cycle 21 the increase in  $\rho$  is seen to be 0.58 proving the benefit of the ellipse model.

If cycle 24 follows the pattern between odd and even cycles, it is expected that

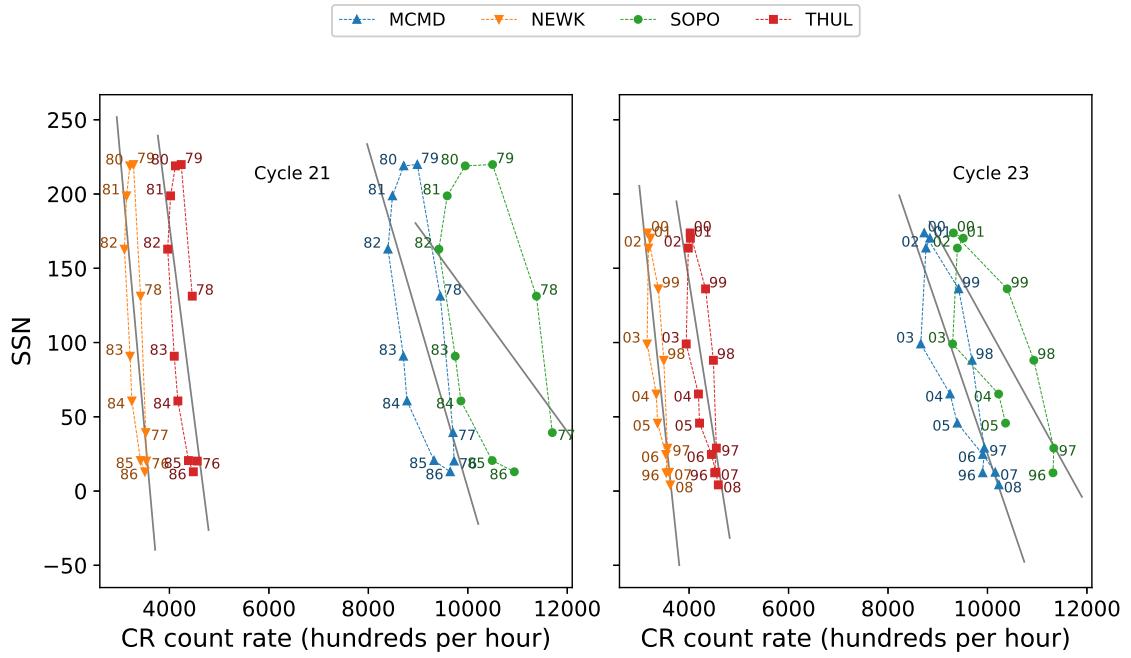


Figure 3.8: The hysteresis plots for odd solar cycles 21 and 23 and the linear regression fit to the data.

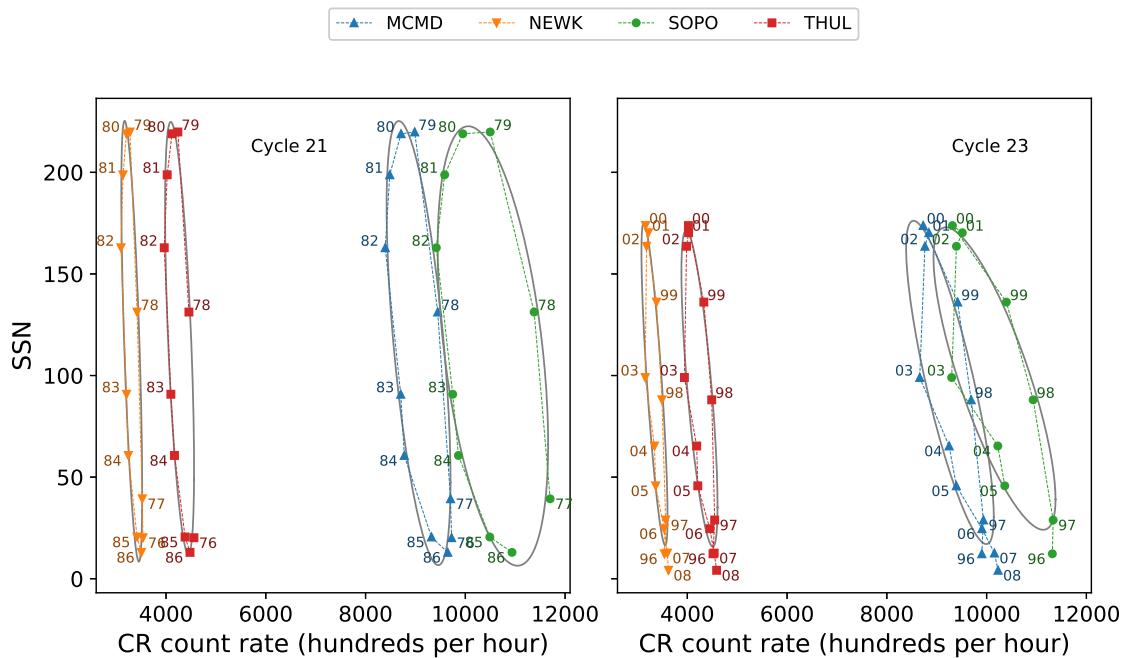


Figure 3.9: The hysteresis plots for odd solar cycles 21 and 23 and the ellipse fit to the data.

the best fit will be provided by the linear model; however it can be seen in Figure 3.2 that cycle 24 appears to display a wider hysteresis loop than the two preceding even-numbered cycles. Both the linear model and ellipse model were applied to the hysteresis plots for solar cycle 24 to determine which model would provide the better fit; the correlation between measured CR intensities and modeled CR intensities are presented in table 3.6.

Table 3.6: Correlation coefficients of the linear regression and ellipse modeling of the hysteresis plots for solar cycle 24.

	Cycle 24	
	Linear	Ellipse
McMurdo	0.903	0.927
Newark	0.936	-
South Pole	0.883	-
Thule	0.873	0.936

The linear model for cycle 24 shows a good correlation between the observed and modeled CR intensities providing evidence to suggest that cycle 24 follows the two preceding even-numbered cycles. The ellipse model does however improve the relation between the observed and modeled CR intensities for 2 out of the 4 stations: McMurdo and Thule. For South Pole and Newark the ellipse model was not able to provide a fit at all, which is believed to be due to the Newark data points crossing where the semi-major axis would be defined for the ellipse model causing the calculation of the semi-major and semi-minor axes to return as not a number, and the South Pole has data missing at the beginning of cycle 24.

The results for solar cycle 24 do not provide a conclusive answer as to whether cycle 24 behaves like past even solar cycles or odd cycles from this data set alone; however the ellipse model does not provide as significant an improvement for the two modeled NM stations as for odd cycles. The small improvement in the ellipse fit is likely due again to the effects of the extended declining phase of cycle 23 and the unusually low activity of cycle 24.

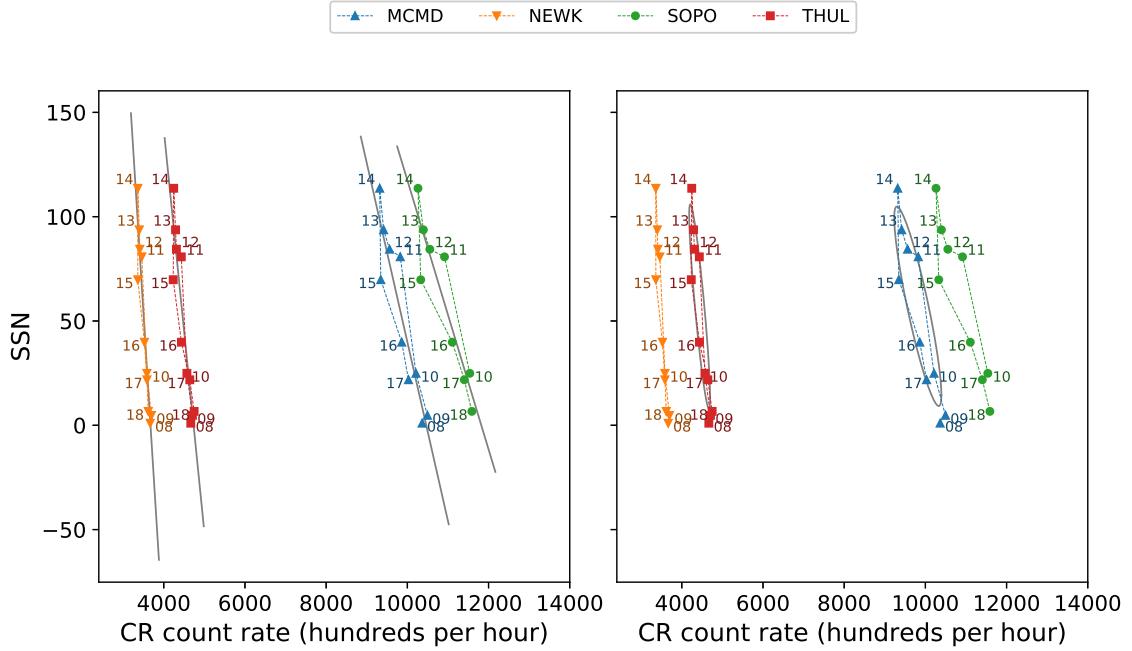


Figure 3.10: The hysteresis plot for solar cycle 24, and the linear regression fit to the data (left) and ellipse fit to the data (right).

We repeated the analysis for the additional 4 NM stations that featured in this study. Again, the linear model provided a good fit to the data however the ellipse model was not able to provide a fit; favouring the conclusion that cycle 24 is best represented by a simple linear model as was true for the preceding even-numbered cycles.

Despite cycle 24 having not yet declined to a minimum, it is clear from the observations shown in the hysteresis plots that further data in cycle 24 is unlikely to broaden the loop any further. The hysteresis loop begins to tighten up after 2016 following the broadening between 2014-2016; hence it appears unlikely that by the end of cycle 24 further observations will support the ellipse model and instead will favour the linear fit to the data.

### 3.5 Conclusions

As cosmic rays are modulated by the heliosphere during the 11-year solar activity cycle, and this effect has been studied for previous solar cycles, the principal aim of

this study was to investigate the nature of GCRs during the current activity cycle 24 as it draws to a minimum.

In this study we presented a time-lag analysis between GCR intensity and SSN which showed that cycle 24 has a longer lag (2-4 months) than the preceding even-numbered solar activity cycles (typically 0-1 months); however its lag is not as large as preceding odd-numbered cycles, and cycle 24 follows the trend of a short or near-zero lag for even-numbered cycles. We suggest here that the cause of the extended lag in cycle 24 compared to previous even-numbered cycles is likely due to the deep, extended minimum between cycle 23 and 24, and the low maximum activity of cycle 24 (Broomhall, 2017).

It has been previously shown in the literature that there is a striking difference in the shape of the plot of SSN and GCR intensity between odd-numbered and even-numbered solar cycles. Due to the difference in the shape of the hysteresis plots for odd-numbered and even-numbered cycles, we have modeled the hysteresis plots using both a simple linear model and an ellipse model. The results of this study tend to support that cycle 24 follows the same trend as preceding even-numbered cycles and is best represented by a straight line rather than an ellipse, such is the case for odd-numbered activity cycles.

We emphasise that although cycle 24 has not yet ended, the shape of the hysteresis plots suggest that we are now past the main broadening region and the inclusion of further data for cycle 24 will very likely only support the linear model. This study will continue to follow the evolution of the cycle 24 until the onset of cycle 25, in around 2019-2021 (Howe et al., 2018; Upton & Hathaway, 2018; Pesnell & Schatten, 2018), when an update on the final results of cycle 24 should be provided.

## 3.6 Update: End of Cycle 24

## 3.7 Comparison with HiSPARC

The analysis was repeated using data from HiSPARC station 501, over the period between 2008 – 2019, to investigate whether the HiSPARC (HS) network was capable of monitoring the solar cycle variation in the muon count rate. HiSPARC station 501 is the best candidate for this analysis due to its long operational lifetime, in comparison to other HS stations, which spans most of cycle 24, and due to the fact that this station is used as the *gold standard* HS station.

This was further motivated by the observations of Fan & Velthuis (2018), in which they stated a confirmation of the accepted anti-correlation between solar activity and CRs after analysing the count rate of the HS 501 station, comparing the count rate with the number of solar flares and the SSN during the rising phase of cycle 24.

We also wanted to determine whether the results of the analysis for the behaviour of CRs during cycle 24, using HS data, agreed with those acquired above, using NM data.

A plot of the HS 501 data over this period is given in Figure 3.11, which shows a comparison between the HS count rate and SSN. By-eye, no clear anti-correlation is displayed between the HS count rate and the SSN.

The spearman correlation coefficient between the HS count rate and the SSN was calculated between the two data sets, without performing the time-lag analysis, to be,  $\rho = 0.078$ . It was quantitatively concluded therefore that there exists no correlation between the SSN and the HS muon count rate, confirming what was suspected by-eye.

This lack of anti-correlation is, unfortunately, symptomatic of the set-up of the HS stations. There exists a stand-out reason why the set-up of the HS stations limits the capability of HS to observe the solar cycle modulation of CR intensity,

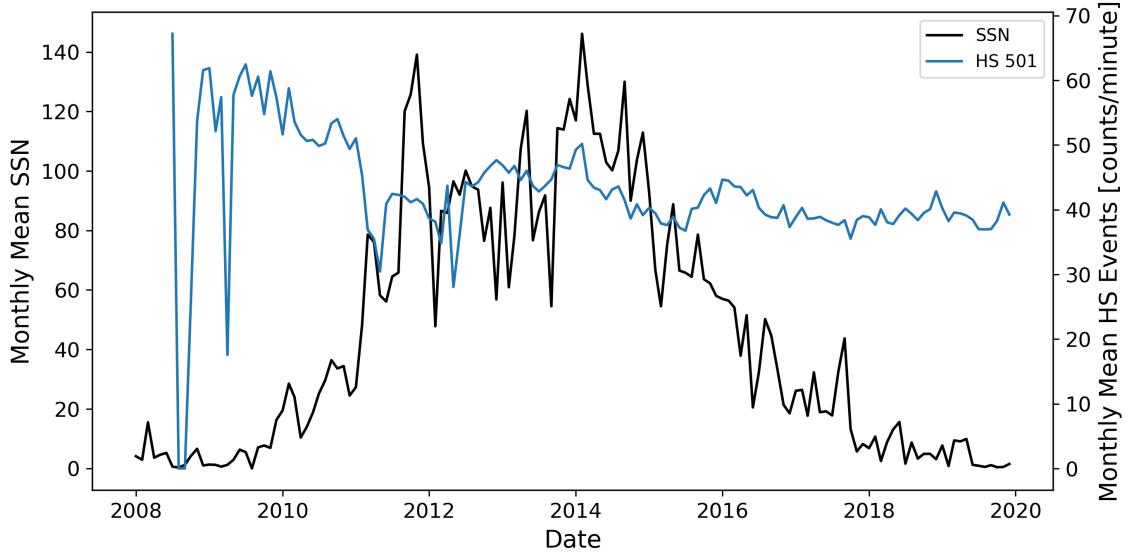


Figure 3.11: The monthly-mean CR-induced muon count rate recorded by HiSPARC (blue), and the SSN (black) between 2008 – 2019.

tweaking of the detector threshold setting.

The HiSPARC stations are typically set up to achieve a singles rate of around 100 Hz above the higher threshold limit, resulting in an events rate of around 1 Hz for the station. The PMT is often adjusted when the count rate strays away from the nominal value, and this mitigates the ability of HiSPARC to reliably observe the solar cycle variation.

Modifications to the PMT voltage of each detector in HS 501 is illustrated in Figure 3.12, and shows the scale of the adjustments, with the voltages varying by as much as  $\sim 80$  V. The effect of varying the PMT voltage is dependent on the individual PMT, but a change of 50–100 V will have an effect on the count rate, therefore it is likely to be a root cause for not observing the solar cycle modulation.

The claim that the solar cycle variation was observed in the count rate of HS 501 data by Fan & Velthuis (2018) was highly likely based on their limited data set, spanning not enough of solar cycle 24. One could be convinced from the count rate during the rising phase of cycle 24 that the cycle variation is observed...

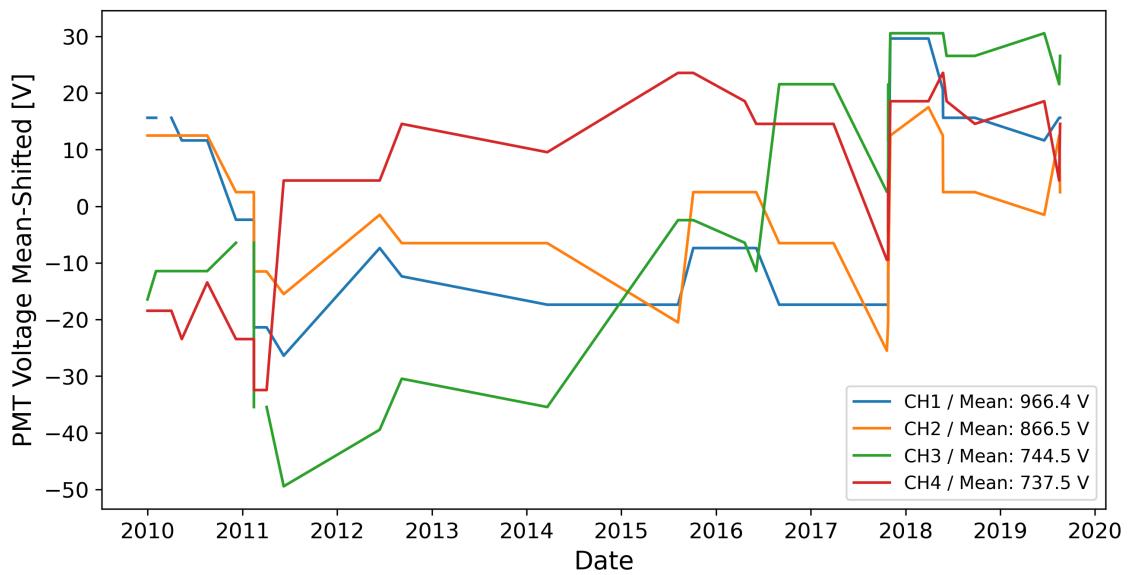


Figure 3.12

# 4 A Frequency Domain Investigation on the Morphology of the Solar Mean Magnetic Field

## 4.1 Introduction

The Sun has a complicated magnetic field structure; many features of the Sun and proxies for the solar activity are related to the evolution of the Sun's magnetic field (Wu et al., 2018).

The Solar Mean Magnetic Field (SMMF) is a surprising, non-zero measurement of the imbalance of opposite magnetic flux polarities observed on the full, visible disk of the Sun (Svalgaard et al., 1975), and is defined as the mean Line Of Sight (LOS) magnetic field when observing the Sun-as-a-star (Scherrer et al., 1977a,b; Garca et al., 1999). In the literature the SMMF is also commonly referred to as the General Magnetic Field (GMF) (Severny, 1971) or the Mean Magnetic Field (MMF) (Kotov, 2008) of the Sun.

Observations of the SMMF have typically been made by measuring the Zeeman splitting of spectral lines using a ground-based Babcock-type magnetograph (Scherrer et al., 1977a), although more recently the SMMF has been calculated from full-disk LOS magnetograms taken from space-borne telescopes such as the Solar Dynamic Observatory Helioseismic and Magnetic Imager (SDO/HMI), in order to

understand the morphology of the SMMF (Kutsenko et al., 2017; Bose & Nagaraju, 2018). It is understood that the strength of the SMMF may vary depending on the spectral line used to measure the SMMF (Kotov, 2008, 2012), however, it is generally accepted in the literature that the SMMF varies slowly with the solar activity cycle, with a amplitude of up to around  $\pm 2$  G during solar maximum and up to around  $\pm 0.2$  G during solar minimum (Plachinda et al., 2011). In addition, the SMMF displays a strong, quasi-coherent rotational signal (Chaplin et al., 2003; Xie et al., 2017), which we assume arises from inhomogeneities on the solar disk with lifetimes of a few rotations.

Despite a wide-ranging existing literature on SMMF observations, spanning several decades, ultimately, our understanding is limited and the SMMF origin remains a crucial, open question in solar physics. The principle component of the SMMF is commonly assumed to be weak, large-scale magnetic flux, distributed over vast areas over the entire solar disk, rather than from more concentrated regions such as Active Regions (ARs) or sunspots (Severny, 1971; Scherrer et al., 1977a; Xiang & Qu, 2016). However, conversely, Scherrer et al. (1972) found that the SMMF was most highly correlated with only the inner-most one quarter, by area, of the solar disk, which is more sensitive to active latitudes.

In recent literature, Bose & Nagaraju (2018) provided an interesting and novel approach to understanding the SMMF whereby they decomposed the SMMF through feature identification and pixel-by-pixel analysis of full-disk magnetograms. Bose & Nagaraju (2018) concluded that: (i) the observed variability in the SMMF lies in the polarity imbalance of large-scale magnetic field structures on the visible surface of the Sun, (ii) the correlation between the flux from sunspots and the SMMF is statistically insignificant, (iii) and more critically that the background flux dominates the SMMF, accounting for around 89% of the variation in the SMMF. There still remained a strong component due to rotation in the background magnetic field presented by Bose & Nagaraju (2018), which is indicative of inhomogeneous mag-

netic features with lifetimes on the order of several solar rotations rather than the shorter-lived, weaker fields usually associated with the large-scale background.

In order to identify the contours of specific features Bose & Nagaraju (2018) used an adaptive thresholding technique on various Solar Dynamic Observatory Atmospheric Imaging Assembly (SDO/AIA) images of the solar disk to create binary masks for different types of features. These masks were then applied to scaled SDO/HMI magnetograms in order to segment which features contributed to the SMMF. Upon a closer inspection of the example magnetogram in Figure 2 of the paper, with over-plotted contours of identified features from SDO/AIA images, there are clearly regions of strong Magnetic Flux Concentrations (MFCs) in the local vicinity of, and connected to, the identified features that are outside of the contour lines and are therefore allocated to the background magnetic field rather than attributed to the specific features. It seems an obvious statement to suggest that SDO/AIA optical counterparts of the magnetograms will not exactly align with the observed magnetic flux in the magnetograms, as Sun's the magnetic field is a holistic feature. We would be expected that the magnetic field will manifest itself differently in the optical observations and the magnetograms, which leads one to believe that the background component in this study could mistakenly contain flux from some of the identified features.

One particular note; there was a careful treatment of plages in this work from additional chromospheric observations, but a separate, specific handling of faculae in the photosphere was absent. Focussing on this could have contributed to the completeness of the study. Furthermore, a decomposition of the identified background component into regimes of strong and weak field would have provided more clarity on the exact morphology of the SMMF, and would have likely provided evidence to conclude whether flux from AR features were incorporated into the background.

Despite these findings, it is known that the strength of the SMMF is weaker during solar minimum, when there are fewer ARs, and stronger during solar maximum,

when there are more ARs. This is suggestive that the processes which underpin the evolution of ARs affect the SMMF.

There is another view in the literature which suggests AR flux dominates the SMMF. It was shown earlier by Kutsenko et al. (2017) that a large component of the SMMF may be explained by strong and intermediate flux regions, that are associated with ARs. Again using a thresholding technique, they showed between 65% to 95% of the SMMF could be attributed to strong and intermediate flux, while the fraction of the occupied area varied between 2% to 6% of the disk area, depending on the chosen threshold for separating weak and strong flux. This finding suggests that strong, long-lived, inhomogeneous MFCs produce the strong rotation signal in the SMMF. Potential sources could be sunspots, plages, faculae, etc. and Kutsenko et al. (2017) discussed that there is an entanglement of strong flux (typically associated with ARs) and intermediate flux (typically associated with network fields and remains of decayed ARs). Disentangling the flux would have provided a more accurate analysis of the SMMF owing to a clearer picture of the main contributor to the SMMF.

The Sun's dynamo and hence magnetic field is directly coupled to the solar rotation. The Sun exhibits latitude-dependent and depth-dependent differential rotation with a sidereal, equatorial period of around 25 days (Howe, 2009). To Earth-based observers, the synodic rotation of the Sun is observed at around 27 days, and the SMMF displays a dominant periodicity of around 27 days due to the solar rotation (Chaplin et al., 2003; Xie et al., 2017; Bose & Nagaraju, 2018). It was also reported by Xie et al. (2017) that the differential solar rotation was observed in the SMMF with measured rotational periods of  $28.28 \pm 0.67$  days and  $27.32 \pm 0.64$  days for the rising and declining phases, respectively, of all of the solar cycles in their considered time-frame.

On the other hand, Xiang & Qu (2016) utilised ensemble Empirical Mode Decomposition (EEMD) analysis to extract rotational modes in the SMMF and found

two rotation periods which are derived from different strengths of magnetic flux elements. They found that a rotation period of 26.6 days was related to a weaker magnetic flux element within the SMMF, while for stronger magnetic flux elements in the SMMF, the measured rotation period was 28.5 days.

Ultimately, to date, our understanding of the SMMF and its origin remains rather limited.

## 4.2 Aims

In this work an investigation of high-cadence (sub-minute) observations of the SMMF, made by Birmingham Solar Oscillations Network (BiSON) (Chaplin et al., 1996, 2005; Hale et al., 2016), was performed. The aim of the investigation was to understand the morphology of the SMMF.

This work provides a frequency domain analysis of the SMMF data, where a model was built up and fit to the power spectrum of the SMMF which allowed us to understand the characteristics of its possible source(s).

The rotational modulation signal in the SMMF was clearly observed as several low-frequency peaks in the power spectrum. In addition, the use of the high-cadence data was especially crucial for inferences on components of the SMMF with periods of less than a day at higher frequencies in the power spectrum, with the intention to determine whether the background magnetic field exhibited a stochastically excited component, which evolved on short timescales.

After fitting a model of the power spectrum, artificial data was simulated and comparisons were made to other sources of SMMF data to aid the clarification of the inferences against the observations.

## 4.3 Data

### 4.3.1 Summary of the Data Set

Chaplin et al. (2003) provided the first examination of the SMMF using data from

BiSON, and the work presented in this paper is a continuation of that study.

BiSON is a six-station, ground-based, global network of telescopes attempting to continuously monitor the Sun, which principally makes precise measurements of the LOS velocity of the photosphere due to solar *p* mode oscillations. Through the use of polarising optics and additional electronics, the BiSON spectrometers can measure both the disk-averaged LOS velocity and magnetic field in the photosphere (Chaplin et al., 2003), however, not all BiSON sites measure the SMMF.

In this study we focus on the data collected by the Sutherland node, in South Africa, which was also used by Chaplin et al. (2003). Data are sampled on a 40-second cadence, and the SMMF data collected by the Sutherland station pertains the epochs from 01/1992 – 12/2012. Over this period, the duty cycle of solar observations is low because we are using only a single site, and averages to be  $\sim 15.6\%$  of the total epoch.

As a comparison to the BiSON data, SMMF observations were also acquired from the Wilcox Solar Observatory (WSO) (<http://wso.stanford.edu/>) (Scherrer et al., 1977b). The WSO SMMF data are sampled daily from 16/05/1975 – present day, but for comparison with the BiSON SMMF, we used data over the same temporal range.

The WSO also measures the LOS SMMF using a Babcock-type magnetograph, which allows the measurement of the amount and sense of circular polarisation in the wings in an absorption line (Scherrer et al., 1977a). WSO uses two absorption lines for the measurement of the magnetic field: the Fe I at 5250 Å( $\lambda$ 5250) is used for measurement of the field, and Fe I at 5124 Å( $\lambda$ 5124) is used to determine the instrument's zero offset, as this line is magnetically insensitive. Scherrer et al. (1977a) describes that a single, complete observation takes  $\sim 20$  minutes, consisting of four 3-minute integrations. The data provided by WSO is a daily weighted mean of the observations, where the weighting used is the statistical uncertainty and the magnitude of the zero offset measured with the  $\lambda$ 5124 line.

### 4.3.2 Obtaining the SMMF from BiSON

There is no catalogued BiSON SMMF data-set, so it was necessary to compute the SMMF from the other, available BiSON data. To acquire the SMMF from BiSON data, the method as described by Chaplin et al. (2003) was adopted; here we discuss the key aspects.

Each BiSON site employs a Resonant Scattering Spectrometer (RSS) to measure the Doppler shift of the  $^2S_{1/2} \rightarrow ^2P_{1/2}$  line (D1 line) of potassium, at  $\sim 770$  nm (Brookes et al., 1978). A potassium vapour cell placed within a longitudinal magnetic field Zeeman splits the laboratory line into the two allowed D1 transitions (Lund et al., 2017). The intensity of the longer wavelength (red;  $I_R$ ) and shorter wavelength (blue;  $I_B$ ) components of the line may be measured by the RSS almost simultaneously, by using polarising optics to switch between the red and blue wings of the line, to form the ratio given by equation (4.1), which is used as a proxy for the Doppler shift from the LOS velocity of the photosphere (see Brookes et al. (1976, 1978); Elsworth et al. (1995a); Chaplin et al. (2003); Lund et al. (2017)).

$$\mathcal{R} = \frac{I_B - I_R}{I_B + I_R} \quad (4.1)$$

There are known effects which occur when making observations of the entire solar disk at one time, such as LOS Doppler-imaging and limb-darkening (Davies et al., 2014b). Some BiSON stations, that do not measure the SMMF, use optics to spatially scramble incoming sunlight to remove Doppler-imaging effects, to ensure a more accurate measure of the disk-averaged Sun-as-a-star. Sutherland, however, is not free from these effects, but they are assumed to be small in the analysis.

Photospheric magnetic fields Zeeman split the Fraunhofer line and the Zeeman-split components have opposite senses of circular polarization (Chaplin et al., 2003). Additional polarising optics are used in the RSS to manipulate the sense of circular polarization (either + or -) that is passed through the instrument. The ratio  $\mathcal{R}_+$

or  $\mathcal{R}_-$  is formed, and the ratios  $\mathcal{R}_\pm$  would be equal if there was no magnetic field present.

The observed ratio ( $\mathcal{R}_\pm$ ) may be decomposed as given by equation (4.2); where  $\mathcal{R}_{\text{orb}}$  is the radial component of the Earth's orbital velocity around the Sun,  $\mathcal{R}_{\text{spin}}$  is the component towards the Sun of the Earth's diurnal rotation about its spin axis as a function of latitude and time,  $\mathcal{R}_{\text{grs}}$  is from the gravitational red-shift of the solar line due to the Sun's mass (Elsworth et al., 1995b; Dumbill, 1999). The LOS velocity due to  $p$  mode oscillations are given by  $\delta r_{\text{osc}}(t)$ , and  $\delta r_B(t)$  is due to the magnetic field (+/- from the polarity) (Dumbill, 1999). The effect of the magnetic field on the ratio is shown in Fig. 4.1, and it is clear to see from equation (4.3) that the difference between the opposite magnetic field ratios is twice the magnetic ratio residual.

$$\mathcal{R}_\pm = \mathcal{R}_{\text{orb}} + \mathcal{R}_{\text{spin}} + \mathcal{R}_{\text{grs}} + \delta r_{\text{osc}}(t) \pm \delta r_B(t) \quad (4.2)$$

$$\mathcal{R}_+ - \mathcal{R}_- = 2 \delta r_B(t) \quad (4.3)$$

In fact, the BiSON RSS is measuring the velocity variation on the solar disk, and therefore a calibration from the ratio to a velocity is necessary. One method of calibration is achieved by first fitting the observed ratio, averaged over both magnetic polarities, to a 2nd- or 3rd-order polynomial as a function of velocity, as discussed by Elsworth et al. (1995b). Here we chose to fit the ratio in terms of velocity,  $\mathcal{R}_{\text{calc}}(u)$ , see equation (4.4):

$$\mathcal{R}_{\text{calc}}(u) = \sum_n \mathcal{R}_n u^n \quad (4.4)$$

where:

$$u = v_{\text{orb}} + v_{\text{spin}} \quad (4.5)$$

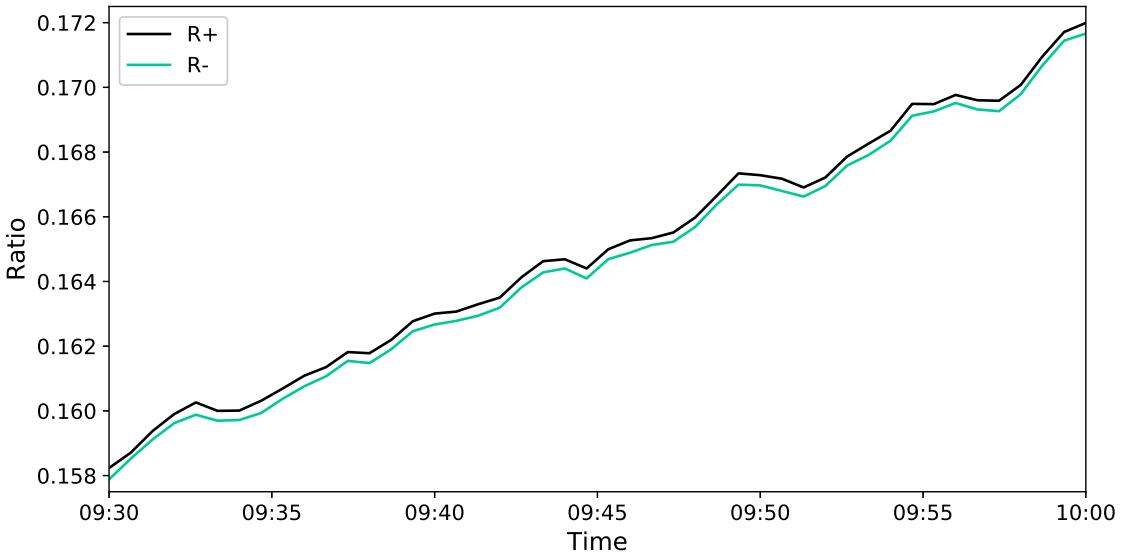


Figure 4.1: An example of the BiSON ratios data over a 30-minute period. The separation between the two ratios is due to the solar mean magnetic field. Other excursions in the individual ratios are due to the other effects measured by the RSS.

and  $v_{\text{orb}}$  is the velocity component related to the ratio,  $\mathcal{R}_{\text{orb}}$ ;  $v_{\text{spin}}$  is related to the ratio,  $\mathcal{R}_{\text{spin}}$ ; and  $n$  is the polynomial order.

It is possible to see that through the removal of  $\mathcal{R}_{\text{calc}}(u)$  from the observed ratios, one is left with the ratio residuals of the  $p$  mode oscillations and the magnetic field (see equation (4.6)). Conversion from ratio residuals into velocity residuals uses the calibration given by equation (4.7).

$$\mathcal{R}_{\pm} - \mathcal{R}_{\text{calc}}(u) = \delta r_{\text{osc}}(t) \pm \delta r_{\text{B}}(t) \quad (4.6)$$

$$\delta v(t) = \left( \frac{d\mathcal{R}_{\text{calc}}}{dV} \right)^{-1} \delta r(t) \quad (4.7)$$

In order to finally obtain the SMMF in units of magnetic field, one must combine equation (4.3) and equation (4.7) with the conversion factor in equation (4.8) (Dumbill, 1999), where  $\mu_B$  is the Bohr magneton,  $h$  is Planck's constant,  $c$  is the speed of light, and  $\nu$  is the frequency of the photons, and the entire procedure can be simplified into equation (4.9).

$$K_B = \frac{8}{3} \frac{\mu_B}{h} \frac{c}{\nu} \approx 2.89 \dots \text{ms}^{-1} \text{G}^{-1} \quad (4.8)$$

$$B(t) = \frac{1}{2} \left( \frac{d\mathcal{R}_{calc}}{dV} \right)^{-1} (\mathcal{R}_+ - \mathcal{R}_-) / K_B \quad (4.9)$$

Through the application of this methodology, one acquires the SMMF as shown in Fig. (4.2a). The power spectrum of the SMMF is shown in Fig. (4.2b), and it shows a strong rotational signal at a period of  $\sim 27$  days.

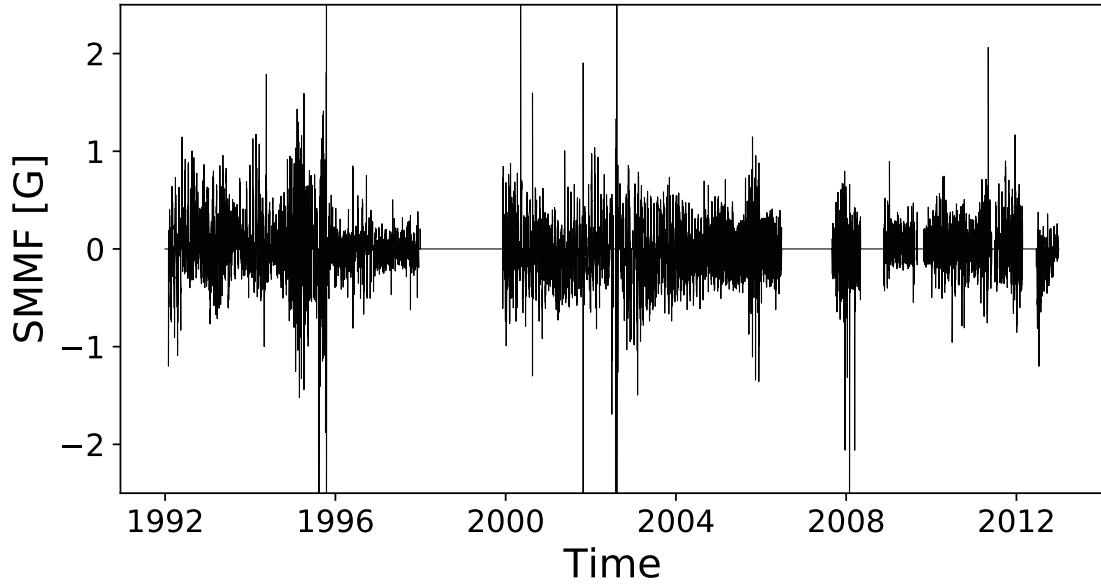
The power spectrum shows a clear set of strong peaks at low frequency, which are due to the persistent rotation signal in the SMMF. The largest peak is the fundamental rotation frequency, and the following peaks are its harmonics.

### 4.3.3 Comparison between WSO and BiSON

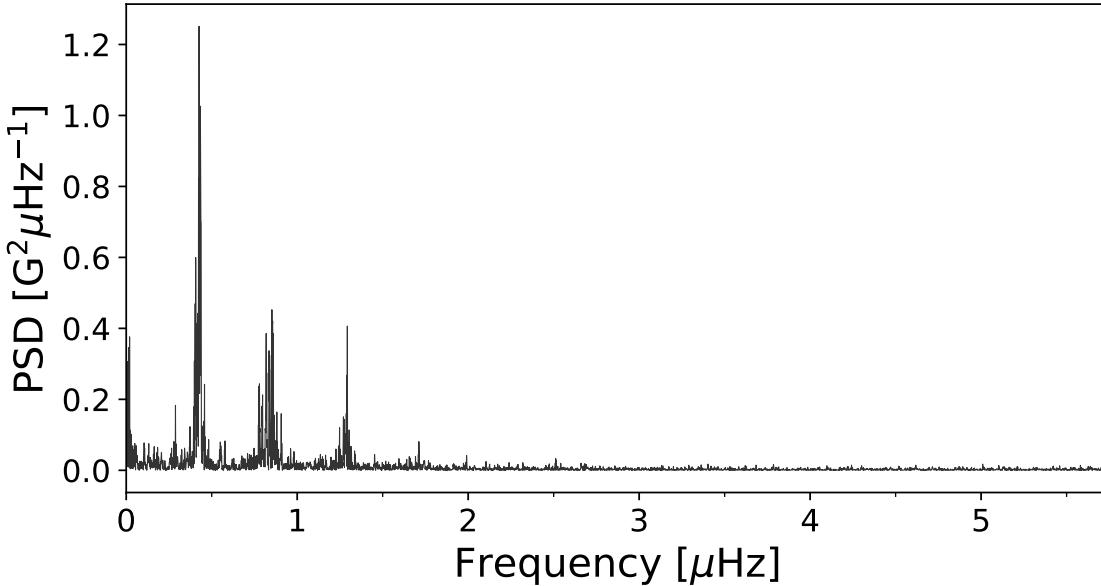
The relationship between the SMMF measured by the Sutherland BiSON station and the SMMF measured by Stanford WSO was examined. The daily SMMF measured by the Stanford WSO and Sutherland BiSON station are plotted in Fig 4.3.

Fig. 4.4 shows the correlation between the two data sets, and also a comparison of their power spectra at low frequencies. The gradient of the line in Fig. 4.4a is  $0.4999 \pm 0.0001$  and informs us that the BiSON SMMF is half of the magnitude of that observed by WSO. Interestingly Chaplin et al. (2003) stated that the SMMF observed by the Sutherland BiSON station was roughly twice that of WSO, which means there is a self-consistency problem with the calibration of the BiSON SMMF and suggests that there is a factor of 4 difference between the measurements of the SMMF by Chaplin et al. (2003) and in this work.

A possible reason for the differences between the WSO observations and the BiSON observations may be explained by the formation heights of the lines. Assuming at each level of the solar atmosphere, regions threaded by magnetic field are in pressure equilibrium with those without any field, then the pressure balance can be expressed by equation (4.10), where  $P_{\text{with}}$  is the gas pressure in regions with



(a) BiSON SMMF 40-second cadence time series



(b) Power spectral density of the BiSON SMMF

Figure 4.2: (a) 40-second cadence observations of the SMMF from the Sutherland BiSON station between 1992 and 2012. The sense of the field was chosen to match the Chaplin et al. (2003) and the WSO observations, where positive is for a field pointing outwards from the Sun. (b) Power spectrum of the SMMF on a 40-second cadence truncated to  $10\mu\text{Hz}$ , however, the nyquist frequency is  $12.5\text{ mHz}$ .

magnetic field,  $P_{\text{without}}$  in those without, and  $B^2/2\mu_0$  is the magnetic pressure ( $\mu_0 =$  magnetic permeability of free space). This implies that the observed magnetic field scales like  $\sqrt{P}$ .

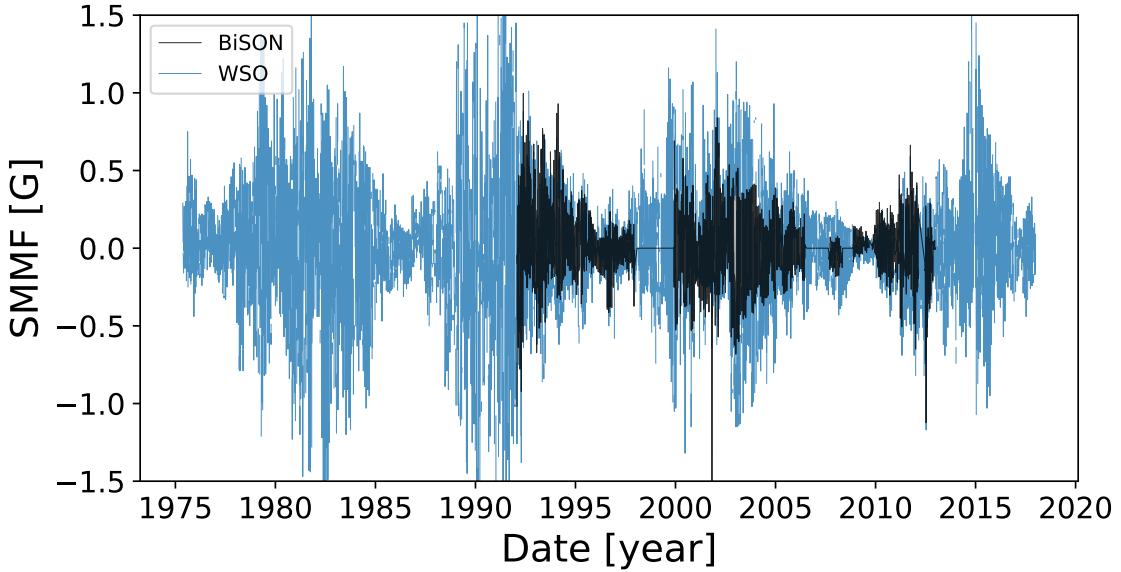


Figure 4.3: Daily averaged SMMF measured by WSO (blue) and by the Sutherland BiSON station (black).

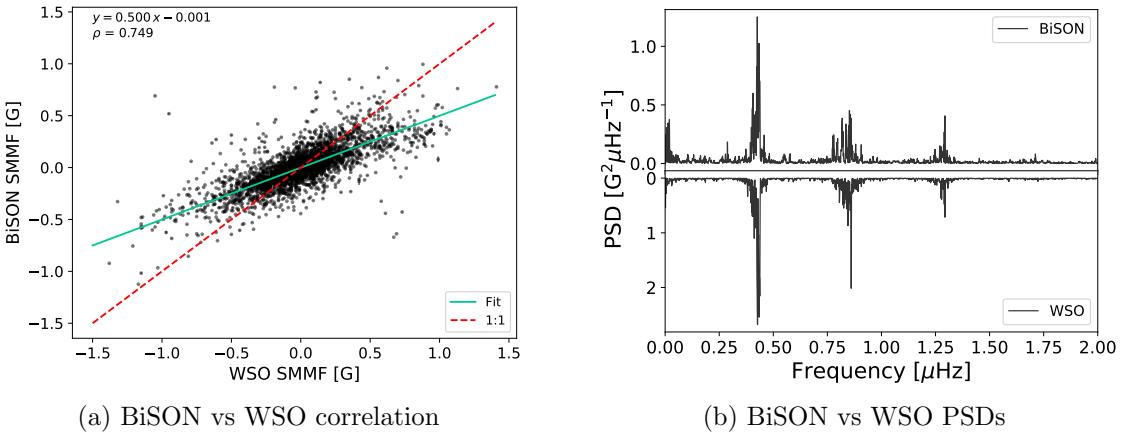


Figure 4.4: Comparisons between the BiSON SMMF data and the WSO SMMF data. (a) shows the correlation between daily averaged BiSON SMMF and the WSO SMMF. The solid-green line provides the fit to the data, while the dashed-red line shows a 1:1 relation for comparison. (b) shows a comparison between the power spectra of BiSON and SMMF.

$$P_{\text{with}} + \frac{B^2}{2\mu_0} = P_{\text{without}} \quad (4.10)$$

Pressure changes exponentially with the scale height,  $H_P$  ( $\sim 100 - 150$  km for the photosphere (Christensen-Dalsgaard et al., 1996)), as described by equation (4.11). Therefore combining both equation (4.10) and equation (4.11), we arrive at equation (4.12).

$$P_2 = P_1 e^{-(h_2 - h_1)/H_P} \quad (4.11)$$

$$B_2 = B_1 e^{-(h_2 - h_1)/(2H_P)} \quad (4.12)$$

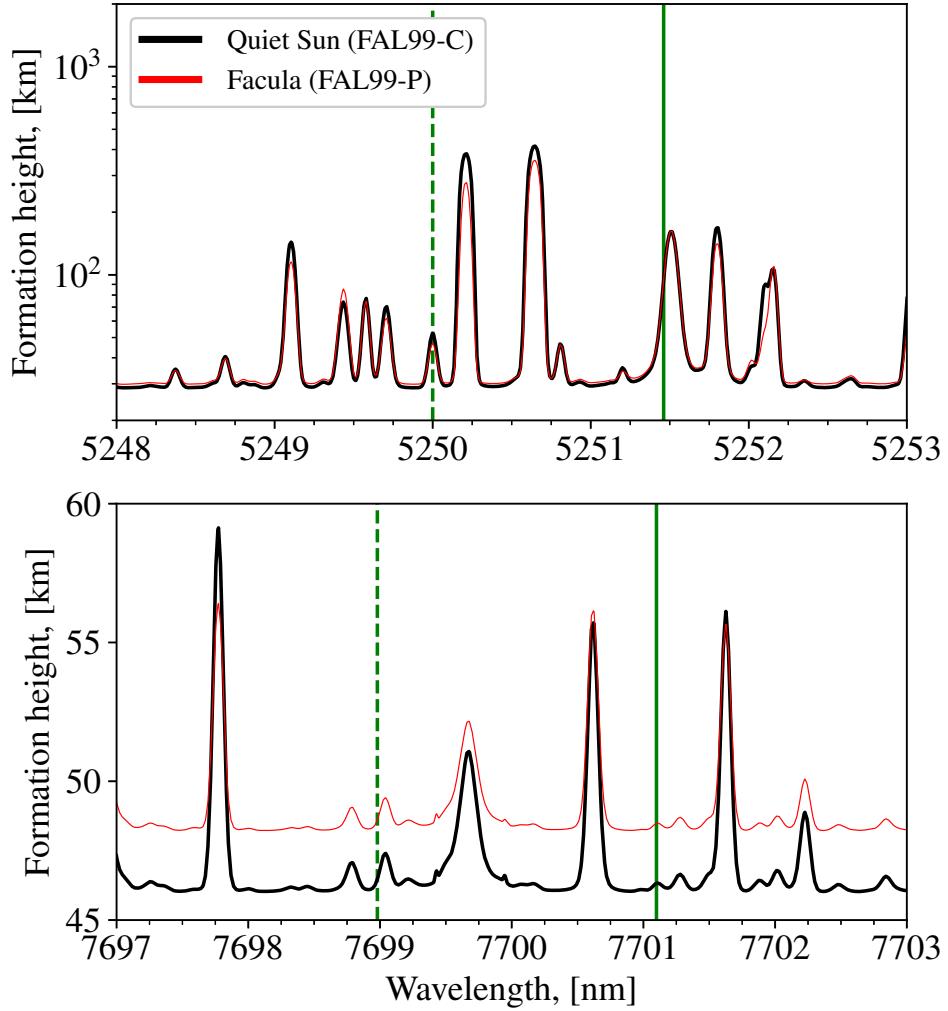


Figure 4.5: Formation heights of emission lines over the wavelength range used for WSO observations (top) and for BiSON observations (bottom). The vertical lines show the wavelengths of the emission lines used for observations in air (dashed) and in vacuum (solid). This plot was provided in private communication courtesy of Rinat Tagirov using the NESSY code.

Using a Non-Local Thermal Equilibrium (NLTE) simulation to model the spectrum of solar magnetic features (see Figure 4.5, and assuming for some slight calibration issues and Doppler shifting, we might suppose that the formation height of

the BiSON line (in vacuum) is around 50 km, and for the WSO line (in vacuum), around 150 km. This would mean that  $B_{\text{WSO}} \approx 0.6 B_{\text{BiSON}} - 0.7 B_{\text{BiSON}}$  which disagrees with our observations and the comparison in Figure 4.4, but agrees more favourably with Chaplin et al. (2003).

[Need to decide still how to neatly handle the discrepancy between old paper, this work, and WSO...]

Finally, we can see quite clearly that the two power spectra align very closely, as shown in Fig. 4.4b, and that both spectra appear to display an asymmetric peak shape, with a negative asymmetry. The WSO power spectrum does also appear to display a lower degree of noise compared to BiSON.

## 4.4 Methodology

### 4.4.1 Identifying Features in the SMMF Power Spectrum

The full power spectrum of the 40-second cadence SMMF is shown in Figure 4.6, covering a frequency range up to the Nyquist frequency of 12.5 mHz, with a resolution of 1.516 nHz.

There are a number of features in the power spectrum. First, the peaks between  $0.1 - 2.0 \mu\text{Hz}$  are a manifestation of a persistent rotational signal the SMMF. The distinct set of peaks indicates the existence of a long-lived, inhomogeneous, Rotationally Modulated (RM) source. The SMMF signal exhibits a quasi-coherent behaviour in the time domain, and based on the comparatively short timescales for the emergence of magnetic features compared to their slow decay (Zwaan, 1981; Harvey & Zwaan, 1993; Hathaway & Choudhary, 2008), we assume the evolution of the RM component with time is a sudden appearance and a long, exponential decay.

As we have 40-second cadence observations of the SMMF, we were able to investigate the power spectrum up to a high Nyquist frequency. This was critical in uncovering a red-noise-like component in the power spectrum. This component could arise from continuously evolving, short-lived regions of magnetic field linked

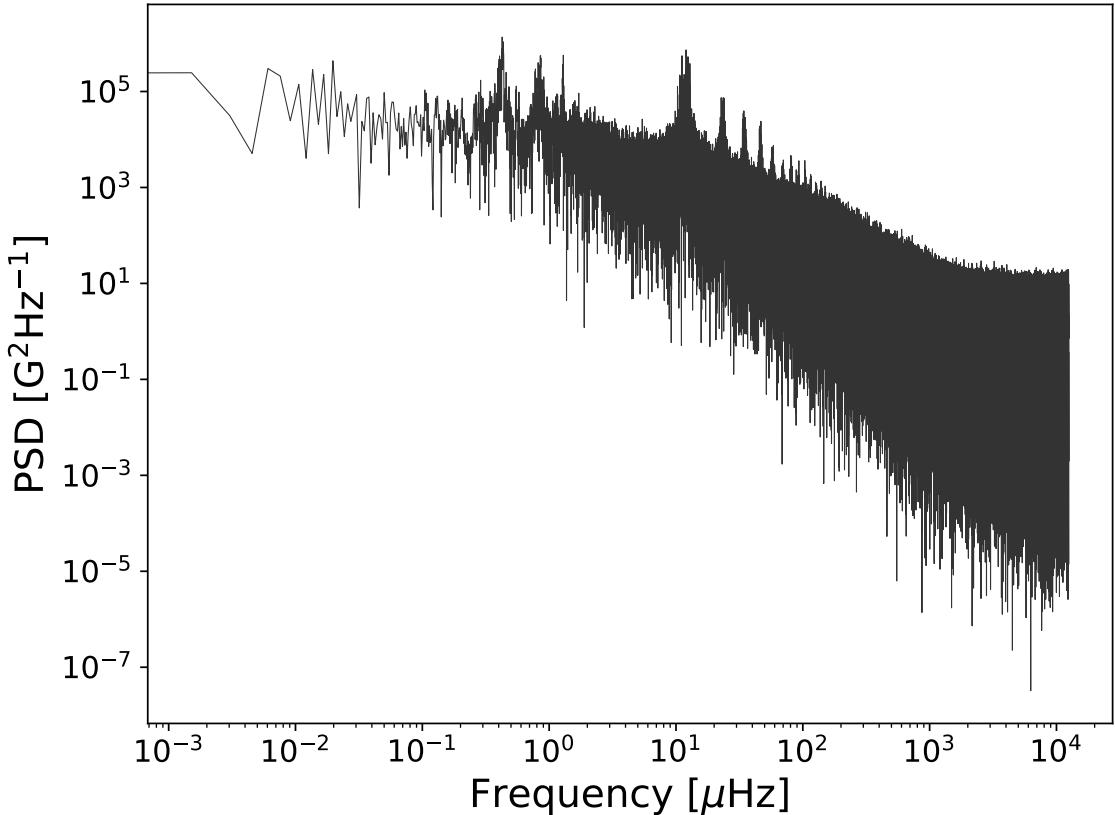


Figure 4.6: Full power spectrum of the BiSON SMMF on a logarithmic scale up to the Nyquist frequency.

to magneto-convection, akin to a random walk, which we will dub the Stochastic Background (SB) component. Analogous to the SB, is the granulation signal observed in the Doppler-velocity measurements of the solar surface (Basu & Chaplin, 2017).

In addition, at low-frequency there is power associated with instrumental noise and solar activity, and at very-high frequency shot-noise is captured which sets the lower limit in power in the spectrum.

There are also side-band features in the power spectrum at multiples of 1/day  $\sim 11.57\mu\text{Hz}$ . The side-bands are a well-known phenomena in ground-based helioseismology. They arise from gaps in the data which are a consequence of making single-site, ground-based observations of the Sun.

The duty cycle of the BiSON observations is very low, at around  $\sim 15\%$ , therefore

it was important to take into consideration the effect that gaps in the data have on the power spectrum. Gaps in the data cause an aliasing of power from actual signal frequencies spread to other frequencies in the spectrum, and the nature of the aliasing depends on the properties of the window function of the observations. It is also possible that the SB component is an effect of power aliasing due to the low duty cycle of the data. Hence, before modelling the power spectrum, the window function was well-characterised.

Through understanding how the duty cycle of the observations affected the power spectrum informed the way we finally parametrised the full model of the power spectrum.

#### 4.4.2 Parameterisation of the SMMF Power Spectrum

In the frequency domain, each of the RM peaks models well as a Lorentzian distribution, similar to peak-bagging modes of solar oscillation (Handberg & Campante, 2011; Davies et al., 2014a), which is due to the quasi-coherent nature of the source. The exponential decay of the RM SMMF source gives width to the peaks in the power spectrum, which we can measure to infer their lifetime.

A single, symmetric Lorentzian peak can be modelled by equation (4.13), where  $\nu$  is frequency,  $A_n$  is the mode amplitude of the RM component,  $\Gamma$  is the mode line-width, and  $\nu_n$  is the frequency of the mode.

$$L_n(\nu; \Gamma, A_n, \nu_n) = \frac{2A_n^2}{\pi\Gamma} \left(1 + \left(\frac{\nu - \nu_n}{\Gamma/2}\right)^2\right)^{-1} \quad (4.13)$$

Upon closer inspection of the power spectrum it is possible to see that the peaks appear to exhibit an asymmetric shape, see Figure 4.2. Taking inspiration from (Howe et al., 2020), it is possible to allow for asymmetry in the Lorentzian peak, which is controlled by the asymmetry parameter,  $\alpha$ , in equation (4.14):

$$L_n(\nu; \Gamma, A_n, \nu_n) = \frac{2A_n^2}{\pi\Gamma(\nu)} \left(1 + (2X(\nu))^2\right)^{-1} \quad (4.14)$$

where

$$X(\nu) = (\nu - \nu_n)/\Gamma(\nu) \quad (4.15)$$

$$\Gamma(\nu) = 2\Gamma/[1 + \exp^{-\alpha(\nu - \nu_n)}]. \quad (4.16)$$

This parameterisation of the asymmetric Lorentzian stops any of the power from becoming negative. In the limit where  $\alpha \rightarrow 0$ , we see that  $\Gamma(\nu) \rightarrow \Gamma$ , thus the asymmetric expression equates to the symmetric expression.

The model function used to describe the RM signal in the power spectrum is given by equation (4.17); the sum of  $N$  Lorentzian-peaks. The subscript,  $n$ , describes a single peak in the power spectrum; in implementing the model we constrain the mode frequencies such that they must be integer values of  $\nu_0$ :  $\nu_n = n\nu_0$ . This means that we define a single rotation frequency only, and subsequent peaks are harmonics. It is worth noting explicitly that this function assumes the line-width of each Lorentzian peak is the same; only their amplitudes and central frequency differ.

$$P(\nu) = \sum_{n=1}^N L_n(\nu; \Gamma, A_n, \nu_n) \quad (4.17)$$

When modelling the power spectrum we attempted with both the symmetric and asymmetric Lorentzian expressions, independently, to determine whether there is a necessity for the extra asymmetry parameter.

Through this formulation we can measure the lifetime of the RM component ( $L$ ), as it is related to the line-width of the peak by equation (4.18).

$$\Gamma = (\pi L)^{-1} \quad (4.18)$$

The low-frequency power can be incorporated into the model via the inclusion of

a zero-frequency centred Lorentzian, i.e. Harvey-function, given by equation (4.19); where  $\sigma$  is the characteristic amplitude of the low frequency signal, and  $\tau$  describes the characteristic timescale of the excursions around zero.

$$H(\nu; \sigma, \tau) = \frac{4\sigma^2\tau}{1 + (2\pi\nu\tau)^2} \quad (4.19)$$

The SB component can also be modelled using the Harvey-function, where  $\sigma$  is the characteristic amplitude of the red-noise signal and  $\tau$  is its characteristic timescale.

Finally, the high frequency power is accounted for by the inclusion of a constant offset due to shot-noise,  $c$ .

When modelling the power spectrum we used the affine-invariant Markov Chain Monte Carlo (MCMC) sampler `emcee` (Foreman-Mackey et al., 2013) to explore the posterior parameter space. The chains are not independent within `emcee`, therefore convergence was interrogated using the integrated autocorrelation time, to quantify the effects of sampling error on the results and to ensure we used a sufficient number of effective samples.

#### 4.4.3 Comparison with the WSO SMMF

To provide comparative results on the inferences from the BiSON SMMF, we repeated the analysis on the power spectrum of the WSO SMMF. The WSO data are only provided on a daily cadence, hence the Nyquist frequency is lower than for BiSON, at  $\sim 5.79 \mu\text{Hz}$ , and it was not possible to observe the SB component.

The same parametrisation as outlined above was relevant to the modelling of the features in the WSO Power Spectral Density (PSD), and the RM peak were fit using a model with symmetric Lorentzian peaks and separately with asymmetric Lorentzian peaks.

## 4.5 Results

### 4.5.1 Investigation of the Window Function

Daily gaps in the data cause some power from the low-frequency RM component in the power spectrum to be aliased to higher frequencies, specifically to harmonics of the frequency of the gaps in the data. In this case there are daily gaps, hence power is aliased to a frequency of  $1/\text{day} \sim 11.57 \mu\text{Hz}$  and its harmonics.

The mode frequency (and harmonics) of the RM component are located near zero ( $\nu_0 \sim 0.4 \mu\text{Hz}$ ). We are usually only interested in the real, positive frequencies but due to their close proximity to zero, they are reflected back as a product of the aliasing and hence there are negative and positive side-bands in the complete power spectrum. When considering the aliased power, both the positive and negative side-bands must be taken into account. The aliased power is located at frequencies defined in equation (4.20), where  $i$  denotes the side-band number, and  $n$  denotes the harmonic of the mode. The locations of side-bands are shown clearly to obey equation (4.20) in the SMMF power spectrum show in Fig. 4.7.

$$\nu_{n,i} = i \left( \frac{1}{\text{day}} \pm \nu_n \right) \quad (4.20)$$

It is clear that we could therefore have used the predicted locations of the aliased power and incorporated them into the model for the full power spectrum. This would, however, have required us to explicitly model some  $\sim 1100$  groups of side-bands in order to cover this effect over the entire frequency range, and each group would have required a unique parameter to control the fraction of power that was contributed to the full PSD. It would have become computationally expensive to model each aliased peak and there would certainly have been room for degeneracy issues to occur.

An alternative approach was to utilise the power spectrum of the window function

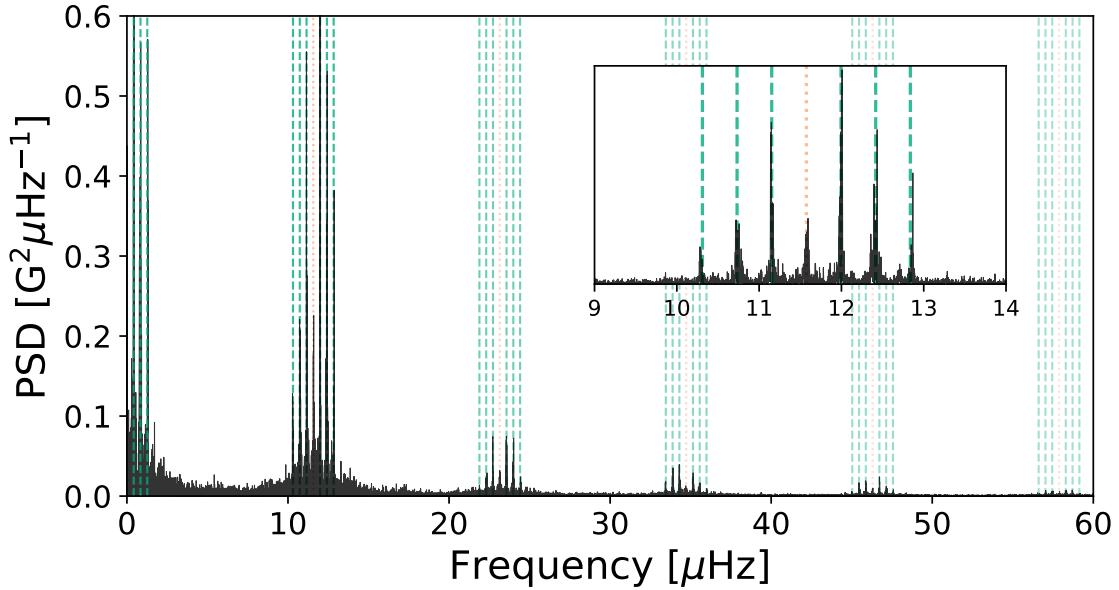


Figure 4.7: Locations of aliased power in side-band peaks. The orange, dotted-lines show the locations of frequencies at multiples of 1/day. The green, dashed-lines show the location of the side-band peaks – harmonic frequencies reflected around multiples of 1/day. The inset shows a zoom of one set of side-band peaks around 1/day.

itself. To do this the Fourier transform of the window function describing the duty cycle of observations was computed (i.e.  $|\mathcal{F}[g(t)]|^2$ ), where the duty cycle function,  $g(t)$ , is given by equation (4.21).

$$g(t) = \begin{cases} 1 & \text{for } |B(t)| > 0 \\ 0 & \text{for } |B(t)| = 0 \end{cases} \quad (4.21)$$

To demonstrate the effect of the window function on the power spectrum, an artificial power spectrum was simulated with a single Lorentzian peak which followed equation (4.17). By computing the inverse Fourier transform, an artificial time-series was generated over the same epoch as the BiSON SMMF observations. We were then able to examine the effects of injecting gaps into the data which were concurrent with the BiSON SMMF gaps.

In Figure 4.8 the power spectrum of the window function is shown, as well as the noiseless peak used to generate the fake data, and the output power spectra of the artificial data with and without the injected gaps. The power spectrum of the

BiSON SMMF data is also plotted for comparison.

It is strikingly clear from Figure 4.8 that the shape of the spectrum of the window function has a remarkable resemblance to the BiSON SMMF spectrum and the output of the artificial spectrum with gaps injected. This demonstrates that the periodic window function, with such a low duty cycle, has a dominating effect on the power spectrum of the input signal which not only produces the diurnal sidebands, but also a broadband spread of the power.

Due to the broadband shape of the window function power spectrum compared to the BiSON SMMF, it appears that there is actually no red-noise-like component in the SMMF and it is instead a manifestation of the gaps on the data. There is still a necessity for a low-frequency Harvey function to describe the power close to zero.

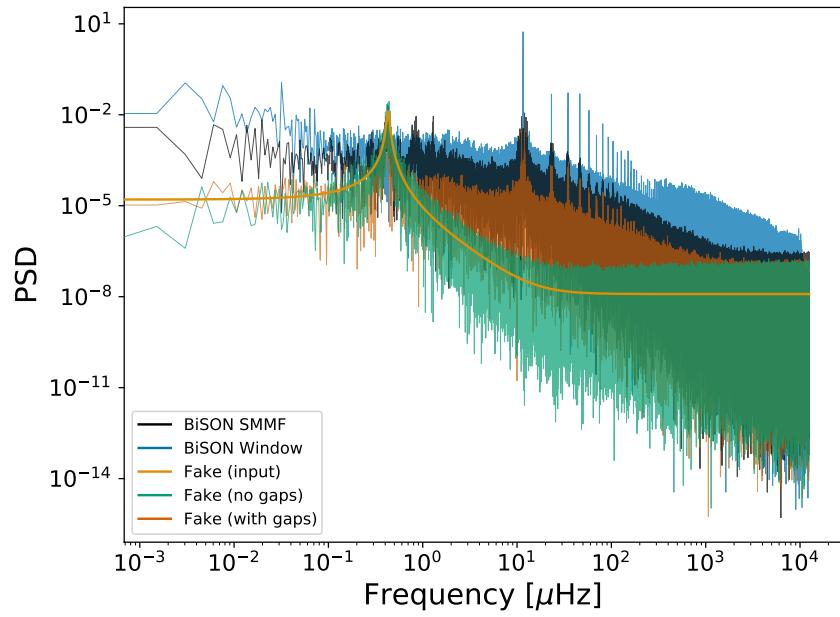
To further understand this effect, we can express the time series data ( $y(t)$ ) as a multiplication of the signal ( $f(t)$ ) with the window function ( $g(t)$ ), as given by equation (4.22).

$$y(t) = f(t) g(t) \quad (4.22)$$

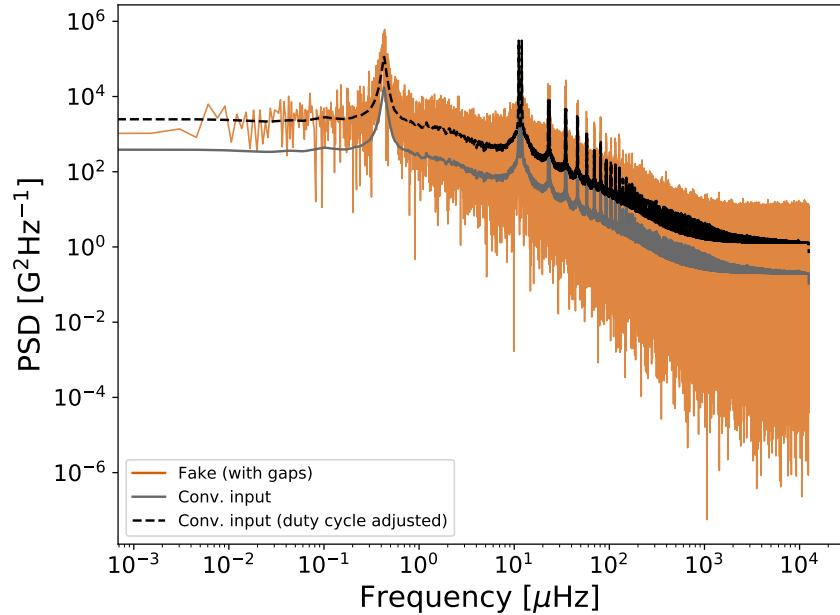
In the frequency domain, as the Fourier transform of a product becomes the convolution of the transformed components, it is possible to express the observed power spectrum in terms of the window function and the gap-free, model power spectrum, given in equation (4.23)

$$P'(\nu; \mathbf{a}) = P(\nu; \mathbf{a}) * |\mathcal{F}[g(t)]|^2 \quad (4.23)$$

Therefore to model the observed power spectrum in a robust manner, which takes into account the intricacies caused by gaps in the data, we used a model which was formed of a model power spectrum,  $P(\nu; \mathbf{a})$ , convolved with the Fourier transform of the window function describing the duty cycle of observations ( $|\mathcal{F}[g(t)]|^2$ ), i.e. a



(a) Black line: BiSON SMMF PSD; blue line: power spectrum of the window function; green and dark-orange lines: the power spectrum of the artificial data without and with gaps, respectively; light orange line: the input peak used to generate the artificial data over-plotted. The lines have been offset for clarity.



(b) Dark-orange line: the power spectrum of the artificial data with gaps; solid-grey line: input PSD convolved with the window function; dashed-black line: input PSD convolved with the power spectrum and adjusted for the duty cycle.

Figure 4.8: (a) The effect on the power spectrum due to periodic gaps in the data. (b) The effect of the convolution between the the window function and the input power spectrum used to make the fake data, highlighting the need to adjust the power for the duty cycle.

model described by equation (4.23), where

$$P(\nu; \mathbf{a}) = \sum_{n=1}^N L_n(\nu; \Gamma, A_n, \nu_n) + H(\nu; \sigma, \tau) + c. \quad (4.24)$$

Care was taken to ensure Parseval’s theorem was obeyed, and no power was lost or gained from the convolution operation. The likelihood of the resulting model from the convolution,  $P'(\nu; \mathbf{a})$ , was then maximised to give the best fitting parameters.

We demonstrated the convolution process by performing the convolution between the noiseless peak used to generate the artificial data and the window function. The result of this is shown in Figure 4.8. Due to method adopted for the convolution, the overall power was reduced as can be seen by the solid-grey line in Figure 4.8. It was therefore necessary to adjust the resultant power spectrum by the duty cycle factor, i.e.  $\sim 15\%$  for BiSON 40-s cadence observations, which produced the dashed-black line in Figure 4.8, which aligns with the power spectrum of the artificial data with gaps.

To further demonstrate the effects of the convolution process during the modelling, we fit a model of a single Lorentzian peak, plus a shot-noise background, to the gap-free fake PSD (without the convolution) and the fake PSD using the gaps (requiring the convolution). The modelling was pefromed using the affine-invariant MCMC sampler `emcee` (Foreman-Mackey et al., 2013) to explore the posterior parameter space, using 6000 iterations on 50 chains. The results of this fit are summarised in Table 4.1.

Table 4.1: Model parameter values for the generation of artificial data, and the median posterior values for the fit to the power spectra generated with and without the gaps in the data. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median.

Parameter	Input	Fit (no gaps)	Fit (gaps)	Unit
$\nu_0$	0.42867	$0.4277^{(+18)}_{(-18)}$	$0.4261^{(+03)}_{(-03)}$	$\mu\text{Hz}$
$\Gamma$	0.030	$0.0279^{(+37)}_{(-36)}$	$0.0340^{(+06)}_{(-06)}$	$\mu\text{Hz}$
$A$	100.0	$101.2^{+7.3}_{-6.0}$	$282.56 \pm 0.25$	mG
$c$	0.20	$0.1872^{(+03)}_{(-03)}$	$1.2009^{(+08)}_{(-08)}$	$\text{G}^2\text{Hz}^{-1}$

There were several points uncovered about the convolution process in this fitting procedure that we had to take into consideration. Firstly, we found agreement with the effect on the total power observed in Figure 4.8. The median posterior values of the amplitude and noise parameters are affected by the duty cycle, compensating for the effect on the power, which resulted in the fitted amplitude being too large by a factor of the square-root of the duty cycle ( $\sim \sqrt{0.15}$ ) and the noise being too large by a factor of the duty cycle ( $\sim 0.15$ ). This again highlights the necessity to incorporate the duty cycle into the convolution to ensure Parseval’s theorem is obeyed.

In addition, we found that the width of the posterior distributions for the parameters were generally smaller as a result of the convolution process. We therefore observed that the likelihood evaluation, when using the convolved model, does not fully take into account the correlated noise on the power spectrum. To perform the full likelihood evaluation with the correlated noise requires large-data computational linear algebra (i.e. the inversion of an N-by-N diagonal-constant/Toeplitz matrix, where N is the size of data). Unfortunately the process of fully accounting for the correlated noise in this scenario is too computationally expensive, due to the large data set with  $> 10^6$  data points. We highlight this, but cannot account for it.

From this exercise we learned that during the modelling of the BiSON (and WSO) power spectra, we needed to account for the duty cycle to ensure no loss of power from the convolution process. We have also learnt that the likelihood evaluation does not properly take into consideration the effect of the correlated noise in the power spectrum, but it is too computationally expensive to account for.

### 4.5.2 Modelling the BiSON Power Spectrum

As there were many data points in the power spectrum, each likelihood calculation was computationally expensive. In order to reduce the required computation, the BiSON power spectrum was cut at a frequency of  $7000 \mu\text{Hz}$ , as at very high fre-

quency, the spectrum purely represents the noise in the SMMF, and it was deemed as a sufficient limit to still fully converge on the shot noise parameter.

We have shown the likelihood evaluation doesn't fully take into account the correlated noise due to the convolution, and doing so would have been too expensive to compute; hence, the pragmatic approach was to ignore the effect of the correlated noise in the convolution in the likelihood calculation.

The BiSON power spectrum was modelled against equation (4.23) (which used equation (4.24) with  $N = 4$  peaks) using the affine-invariant MCMC sampler `emcee` (Foreman-Mackey et al., 2013) to explore the posterior parameter space, using 7000 iterations on 50 chains.

$$\nu_0 \sim \mathcal{U}(0.38, 0.50) \mu\text{Hz}$$

$$\Gamma \sim \mathcal{U}(0.00, 0.11) \mu\text{Hz}$$

$$A_1 \sim \mathcal{U}(300, 900) \text{ mG}$$

$$A_2 \sim \mathcal{U}(100, 500) \text{ mG}$$

$$A_3 \sim \mathcal{U}(50, 350) \text{ mG}$$

$$A_4 \sim \mathcal{U}(20, 150) \text{ mG}$$

$$\sigma \sim \mathcal{U}(0.10, 1000) \text{ mG}$$

$$\tau \sim \mathcal{U}(0.10, 200) 10^6 \text{ s}$$

$$c \sim \mathcal{U}(10^{-2}, 10^2) \text{ G}^2 \text{ Hz}^{-1}$$

$$\alpha \sim \mathcal{U}(-500, 0)$$

In Table 4.2 the median values of marginalised posterior distributions for each of the model parameters are displayed. Reported uncertainties on the parameters correspond to the  $1\sigma$  (68%) credible intervals either side of the median.

Table 4.2: Median values of the marginalised posterior distributions for each model parameter in the fit to the BiSON power spectrum, adjusted for the duty cycle factor ( $\sim 0.156$ ) in the convolution process. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median.

Parameter	40-s symm.	40-s asymm.	Unit
$\nu_0$	$0.42699^{(+13)}_{(-13)}$	$0.42778^{(+14)}_{(-14)}$	$\mu\text{Hz}$
$\Gamma$	$0.02639^{(+49)}_{(-48)}$	$0.03163^{(+71)}_{(-71)}$	$\mu\text{Hz}$
$A_1$	$166.0 \pm 2.0$	$178.9 \pm 1.2$	mG
$A_2$	$115.9 \pm 2.3$	$129.0 \pm 1.2$	mG
$A_3$	$83.2 \pm 2.7$	$93.5 \pm 1.2$	mG
$A_4$	$32.6 \pm 4.3$	$38.9 \pm 1.9$	mG
$\tau$	$51.8^{+3.9}_{-3.7}$	$62.7^{+4.9}_{-4.6}$	days
$\sigma$	$83.4 \pm 3.5$	$79.1 \pm 1.4$	mG
$c$	$0.2103^{(+09)}_{(-09)}$	$0.2102^{(+01)}_{(-01)}$	$\text{G}^2\text{Hz}^{-1}$
$\alpha$	—	$-119.8^{+8.0}_{-7.9}$	—

As the asymmetry parameter converged reasonably within the prior bounds we therefore deduced that the extra parameter was necessary, and the model utilising asymmetric Lorentzian peaks was a better fit to the data than the model with symmetric Lorentzian peaks. The convolved model of the data, using asymmetric Lorentzian peaks is shown in Figure 4.9 over-plotted on top the BiSON SMMF power spectrum.

The central frequency of this model,  $\nu_0$  implies a rotation period of  $27.06^{+0.01}_{-0.01}$  days, and accounting for sidereal rotation,  $25.19^{+0.01}_{-0.01}$  days. The rotation period measured is in agreement with other literature values for the rotation signal in the SMMF (Chaplin et al., 2003; Xie et al., 2017).

According to the model for differential rotation given by Snodgrass (1983) and Brown et al. (1989), the measured rotation period implies the RM component of the SMMF is sensitive to a time-averaged latitude of around  $11^\circ$ . This latitude is consistent with the latitudes spanned by sunspots and ARs over the solar activity cycle (Maunder, 1904; McIntosh et al., 2014; Thomas et al., 2019), and particularly during the declining phase of the solar cycle. This strongly implies that the origin of the RM component of the SMMF is linked to ARs and MFCs.

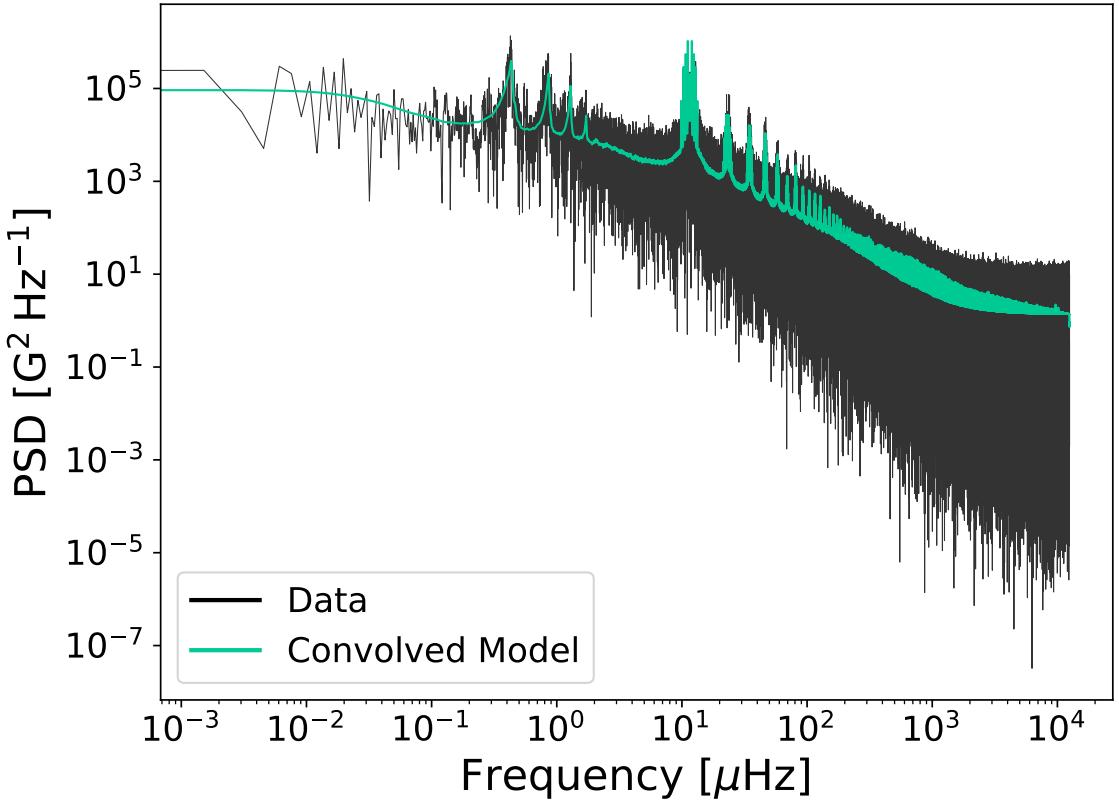


Figure 4.9: Full, modelled power spectrum of the BiSON SMMF on logarithmic axes. The data is displayed in black and the convolved model using asymmetric Lorentzian peaks is shown in green.

Furthermore, from the measured line-width of the Lorentzian peaks, we have calculated the lifetime of the RM component using equation (4.18). The line-width suggests a lifetime of  $116.5 \pm 2.6$  days, which is in the region of  $\sim 16.5$  weeks. The typical lifetime of ARs and sunspots is usually on the order of weeks to months, dependent on their size (Zwaan, 1981; Schrijver & Harvey, 1994; Howard, 2001; Hathaway & Choudhary, 2008; van Driel-Gesztelyi & Green, 2015), therefore we have measured a lifetime of the RM component which is consistent with the lifetime of ARs and sunspots. This again suggests that the source of the signal is linked to active regions of magnetic field.

Taking into account the work performed by Bose & Nagaraju (2018), which showed compelling evidence that the sunspots did not contribute to the SMMF, we should rule out sunspots as the strong RM source of the SMMF. The method

of identifying ARs or strong MFCs in magnetograms by Bose & Nagaraju (2018) potentially mis-identified regions of magnetic flux associated with ARs and MFCs as background flux however, hence it is possible that these do, in fact, contribute to the SMMF. By comparison with the work carried out by Kutsenko et al. (2017), the work performed here agrees that the SMMF is dominated by features with properties in-line with ARs and MFCs, i.e. long-lived and make up a fraction of the solar disk. With all this considered, we state that our results indicate the SMMF has its origin in the vicinity of ARs and other concentrations of strong flux that are long-lived on the solar disk and exist in active latitudes.

#### 4.5.3 Comparison to the WSO Power Spectrum

When modelling the WSO power spectrum, although the duty cycle of the WSO observations over the same epoch as BiSON observations was  $\sim 78\%$ , therefore there was less of an effect from the window function on the observations, the power spectrum was modelled using the convolution with the window function, as per the BiSON power spectrum.

The BiSON power spectrum was modelled against equation (4.23) (which used equation (4.24) with  $N = 3$  peaks) using the affine-invariant MCMC sampler `emcee` (Foreman-Mackey et al., 2013) to explore the posterior parameter space, using 10000 iterations on 50 chains. The chains are not independent within `emcee`, therefore convergence was interrogated using the integrated autocorrelation time, to quantify the effects of sampling error on the results and to ensure a sufficient number of effective samples were used.

In Table 4.3 the median values of marginalised posterior distributions for each of the model parameters are displayed. Reported uncertainties on the parameters correspond to the  $1\sigma$  (68%) credible intervals either side of the median. The convolved model of the data, using asymmetric Lorentzian peaks is shown in Figure 4.10 over-plotted on top the BiSON SMMF power spectrum.

Table 4.3: Median values of the marginalised posterior distributions for each model parameter in the fit to the WSO power spectrum, adjusted for the duty cycle factor ( $\sim 0.778$ ) in the convolution process. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median.

Parameter	24-hr symm.	24-hr asymm.	Unit
$\nu_0$	$0.4273^{(+05)}_{(-05)}$	$0.4290^{(+04)}_{(-05)}$	$\mu\text{Hz}$
$\Gamma$	$0.0209^{(+16)}_{(-15)}$	$0.0210^{(+16)}_{(-15)}$	mG
$A_1$	$134.4^{+3.2}_{-3.3}$	$137.0^{+3.4}_{-3.5}$	mG
$A_2$	$120.7 \pm 3.4$	$124.8^{+3.5}_{-3.6}$	mG
$A_3$	$73.7 \pm 3.5$	$77.4 \pm 3.8$	mG
$\tau$	$22.5^{+10.1}_{-6.9}$	$23.7^{+11.1}_{-7.3}$	$10^6\text{s}$
$\sigma$	$52.1^{+4.7}_{-4.4}$	$51.5^{+4.6}_{-4.5}$	mG
$c$	$11.4^{+17.8}_{-8.5}$	$12.0^{+18.6}_{-8.9}$	$\text{G}^2\text{Hz}^{-1}$
$\alpha$	—	$-61.7^{+13.3}_{-15.3}$	—

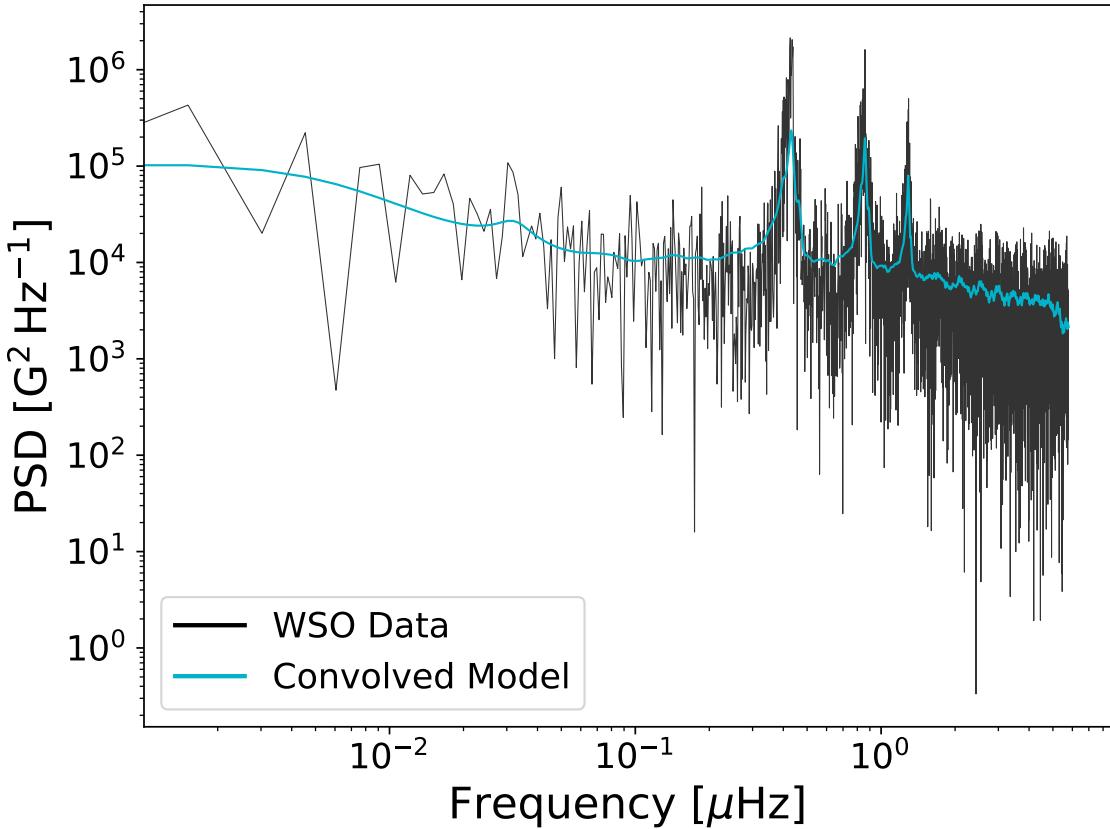


Figure 4.10: Full, modelled power spectrum of the WSO SMMF on logarithmic axes. The data is displayed in black and the convolved model using asymmetric Lorentzian peaks is shown in blue.

The fit to the WSO power spectrum using asymmetric Lorentzian profiles produces results that are close to those measured with the BiSON power spectrum. The

rotation period is in agreement with that measured using BiSON data to within  $2\sigma$ , and this period implies a cycle-averaged latitude of around  $8^\circ$ . This agrees with the conclusions drawn from our inferences of the BiSON data that the RM source is linked to ARs.

The line-width suggests a RM lifetime of  $175 \pm 13$  days, which is in the region of  $\sim 25$  weeks or half a year. This lifetime is inconsistent with that measured using the BiSON data, however; the lifetime measured in the BiSON power spectrum is around 50% larger than that measured with the WSO data. A possible explanation for this could be the resolution differences of the two instruments, or more likely it is a by-product of the different regions within the photosphere they probe, as discussed above. The stronger field measured by WSO implies the sources would have a longer lifetime. These limits are still consistent however with the lifetime of large, strong ARs (Schrijver & Harvey, 1994; van Driel-Gesztelyi & Green, 2015).

## 4.6 Discussion

### 4.6.1 Asymmetry in the Power Spectrum

We have shown that the BiSON and WSO SMMF power spectra are best modelled with asymmetric Lorentzian profiles. In this section, we investigate the cause of this asymmetry.

An initial hypothesis on the origin of the asymmetry was the migration of active regions towards the equator with the progression of the solar cycle. This is illustrated clearly by the asymmetry in the rotation frequencies of sources at latitudes sampled from a Kernel Density Estimate (KDE) of the SSN, in combination with models for the differential rotation (Snodgrass, 1983) and equatorial migration (Li et al., 2001a), shown in Figure 4.11. The migration of ARs towards the equator has an asymmetry which is skewed towards lower frequencies, in accordance with the asymmetry of the Lorentzian peaks in the BiSON and WSO power spectra. If the total power spectrum is the sum of the power spectra for each contributing source, which have

central frequencies following the distribution shown in Figure 4.11, then we may expect to see this same asymmetry manifested in the power spectrum.

To investigate migration as a source of the asymmetry, we used artificial data which was created to simulate the differential rotation and migration of active regions during the solar cycle. The methodology involved with generating the simulated data is discussed in Appendix B, and here we discuss the outcomes.

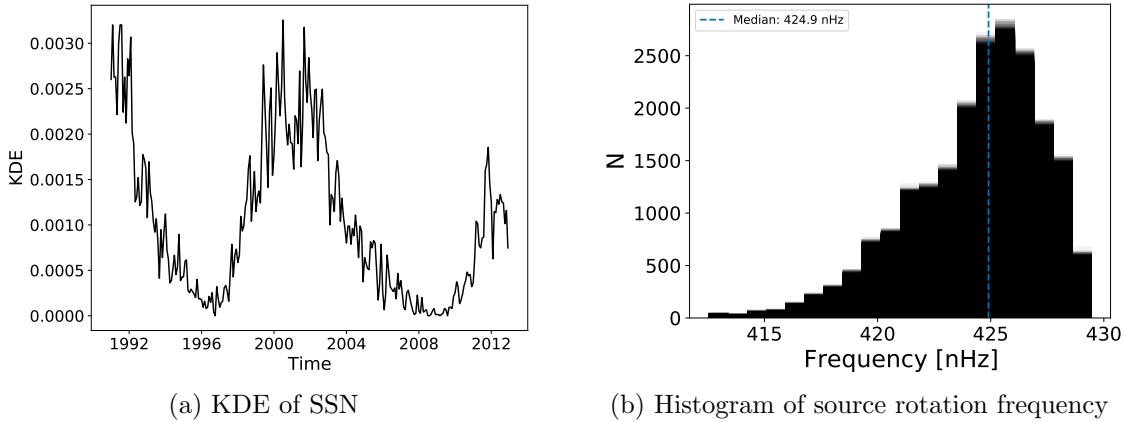


Figure 4.11: (a) shows the KDE of the monthly averaged sunspot number used to draw samples of source seed times; (b) histogram of the rotation frequency of sources sampled from the KDE, after using the model for the migration and differential rotation.

Three separate models for the migration of ARs towards the equator were used in the simulations and they are shown in Figure 4.12. The quadratic model was taken from Li et al. (2001a), which represents a ‘typical’ migration of ARs, and the linear and exponential models were parametrised to provided opposite extremes for a slower and faster migration towards the equator, respectively. The aim of this was to investigate whether varying the migration rate affected the asymmetry. Each migration model is given in equation (4.25), (4.26), and (4.27), where  $t$  is the time since the start of the cycle, in years.

$$\lambda_q = 0.0893t^2 - 2.8t + 27.24 \quad (4.25)$$

$$\lambda_l = -1.84t + 27.24 \quad (4.26)$$

$$\lambda_e = 20.24e^{-t/2} + 7 \quad (4.27)$$

In Figure 4.12 we also note that for each model of the migration, we see a different median rotation frequency in the distribution, which is lower for the slower migration, and higher for the faster migration. This is expected, due to the differential rotation, but it shows that the measured central rotation frequency of the SMMF power spectrum is sensitive to the migration and tells us something about the source. This also suggests that the RM component on the SMMF power spectrum is consistent with a faster migration model.

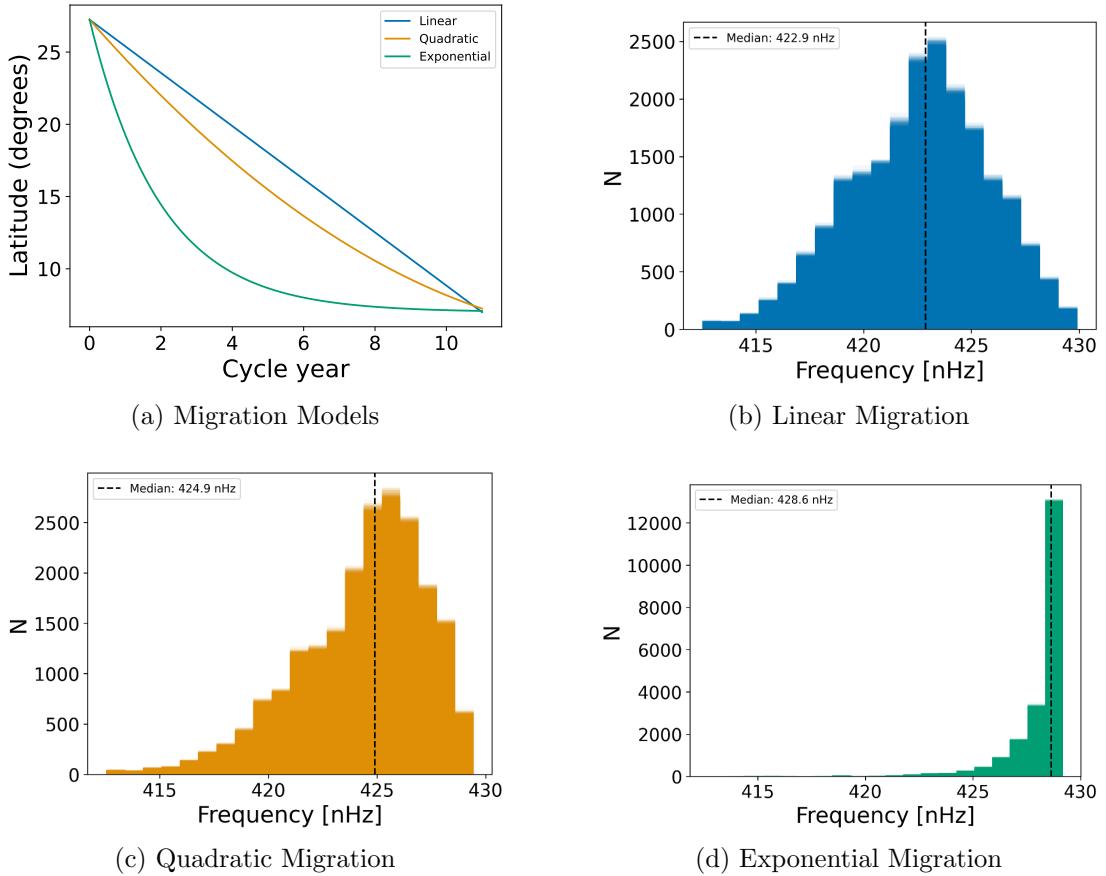


Figure 4.12: (a) shows the latitudinal migration model as a function of time for each model; (b)-(d) shows the histogram of the rotation frequencies of sources sampled from the KDE for linear, quadratic, and exponential migration models, respectively.

The simulations were run with 15 different configurations which used the three migration models combined with five models for the source transits across the visible

disk (see Appendix B: cosine model, sign change model, and combinations of the cosine and sign change models in ratios of 5:95, 10:90, and 20:80). In each case, 250 simulations were performed in order to produce a limit spectrum. Each limit spectrum was modelled with a series of Lorentzian peaks sharing a global asymmetry parameter, and the resultant asymmetry values are shown in Figure 4.13.

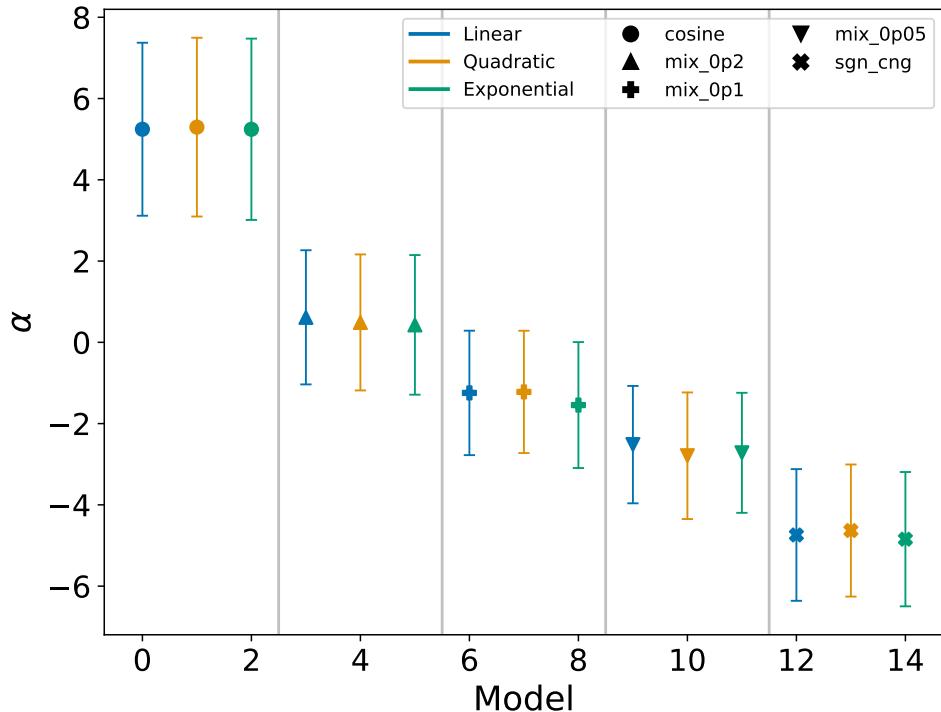


Figure 4.13: The median value of the fitted global asymmetry parameter for several simulations of artificial data. The colour represents the migration model used: linear:blue; quadratic:orange; exponential: green. The marker represents the source model used: cosine:circle; sign change:cross; combinations of the cosine and sign change models in ratios of 5:95:downwards triangle; 10:90:plus; 20:80:upwards triangle.

It is immediately clear that the choice of migration model has no significant effect on the asymmetry of the Lorentzian peaks, hence we are able to rule out AR migration as the origin of the asymmetry observed. A possible reason for this explanation is that it is not correct to suggest that the total power spectrum is the sum of power spectra for each individual source, because this neglects the interference between sources in the time domain. However, this figure does show an obvious difference between the way the different models of the sources have asymmetry manifested in

their spectra.

This links to a still-open debate in helioseismology, on the source of asymmetry of the modes in the power spectra, and the asymmetry reversal between the observations in intensity and Doppler velocity. Asymmetry is regularly observed in  $p$  modes of oscillation in helioseismic data and we observe a difference in the sign of the asymmetry term which is negative for Doppler velocity observations (i.e. more power in the low-frequency wing of the mode) and positive for measurements made in intensity (i.e. more power in the high-frequency wing of the mode) (Duvall et al., 1993; Chaplin & Appourchaux, 1999; Howe et al., 2015; Basu & Chaplin, 2017).

There are believed to be two main causes of asymmetry in acoustic modes:

1. The spatially localised nature of the excitation source of acoustic modes in the near-surface layers of the outer convection zone and the interference between multiple waves that have accumulated a phase difference since their emission. This source of asymmetry is believed to dominate in Doppler velocity observations.
2. The correlation between the convective granulation (i.e. correlated noise) and the signal from the modes themselves, or the correlation between the excitations of one mode to another. This source of asymmetry is believed to dominate in intensity observations due to the lower signal-to-noise ratio.

We believe that some combination of phase difference between sources, and noise arising the randomness of the simulations is responsible for the asymmetry difference observed between the cosine and sign change models, as they replicate the difference in the ways we observe the ARs, either by intensity (cosine model) or Doppler velocity (sign change model).

In the case of the SMMF observations, we believe the asymmetry is explained by the first of the above explanations. Active magnetic regions are typically localised in the near-surface layers of the Sun and we see clear interference between

different regions of magnetic flux which form the observed more prominent active regions, which we believe contribute strongly to the SMMF. The interference between small regions of flux will have accumulated a phase difference and therefore could contribute to the observed asymmetry.

On the second point, the SMMF is measured by the Zeeman splitting of an emission line due to the magnetic field, which is inherently a Doppler velocity method of observing. There is no bleed through of  $p$  mode power in the spectrum and we have shown that there is no red-noise-like signal or stochastic background with a short timescale resembling granulation in the power spectrum, therefore it is unlikely that we see a strong correlated noise contribution to the asymmetry in the SMMF RM component peaks.

#### 4.6.2 Testing the Effects of Differential Rotation and Active Region Migration

We know the rotation period of ARs varies throughout the solar cycle as a result of solar differential rotation and latitudinal migration. As we have inferred that the RM component of the SMMF is likely linked to ARs and MFCs, we may therefore assume that the RM component is also sensitive to these effects. Here we analyse the effect of migration and differential rotation on our ability to make inferences on the lifetime of the RM component.

Several studies have modelled the the solar differential rotation, and its variation with latitude and radius of the Sun (see Beck (2000) and Howe (2009) for an in depth review of the literature on solar differential rotation). Magnetic features have been shown to be sensitive to rotation deeper than the photosphere; therefore, in general, magnetic features can be seen to rotate with a shorter period than the surface plasma (Howe, 2009).

Chaplin et al. (2008) analysed the effects of differential rotation on the shape of asteroseismic low- $l$   $p$  modes of oscillation, and showed that the consequence of

differential rotation is to broaden the observed line-width of a mode peak. The authors provide a model of the resultant profile of a  $p$  mode whose frequency is shifted in time to be a time-average of several instantaneous Lorentzian profiles with central frequency  $\nu(t)$ , given by equation (4.28). The angled brackets indicate an average over time.  $H$  and  $\Gamma$  are the mode height (maximum power spectral density) and line-width, respectively. The full period of observation is given by  $T$ .

$$\langle P(\nu) \rangle = \frac{1}{T} \int_0^T H \left( 1 + \left( \frac{\nu - \nu(t)}{\Gamma/2} \right)^2 \right)^{-1} dt \quad (4.28)$$

Chaplin et al. (2008) also show that by assuming a simple, linear variation of the unperturbed frequency,  $\nu_0$ , from the start to the end of the time-series by a total frequency shift  $\Delta\nu$  (see equation (4.29)),

$$\nu(t) = \nu_0 + \Delta\nu \frac{t}{T} \quad (4.29)$$

the resultant profile of a  $p$  mode can analytically be modelled by equation (4.30):

$$\langle P(\nu) \rangle = \frac{H}{2\epsilon} \arctan \left( \frac{2\epsilon}{1 - \epsilon^2 + X^2} \right) \quad (4.30)$$

where  $\epsilon$  and  $X$  are defined in equation 4.31 and equation 4.32.

$$\epsilon = \frac{\Delta\nu}{\Gamma} \quad (4.31)$$

$$X = \frac{\nu - [\nu_0 + (\Delta\nu/2)]}{\Gamma/2} \quad (4.32)$$

As the mode line-widths are broadened by this effect, we evaluate whether our ability to revolve the true line-width of the RM component, and hence its lifetime, is affected. To evaluate this we computed the broadened profiles given by both equation (4.28) and equation (4.30), and fit the model for a single Lorentzian peak, to determine whether the line-width is recovered.

In the first instance, we computed the broadened peak using equation (4.28). Over the duration of the observations, we computed the daily instantaneous profile,  $P(\nu(t))$ . The time-averaged profile,  $\langle P(\nu) \rangle$ , is a weighted average of each instantaneous profile, where the weights are given by the squared, daily-averaged SMMF, in order to allow a larger broadening contribution at times when the SMMF amplitude is large.

In the second instance, we computed the broadened peak using equation (4.30). Over the duration of the observations the daily frequency shift is computed,  $\Delta\nu$ . The time-averaged shift,  $\Delta\nu$ , is a weighted average, where again the weightings are given by the squared, daily-averaged SMMF.

To determine the shift in the rotation rate with migration, we used the model of the solar differential rotation as traced by magnetic features ( $\Omega_m$ ) given by equation (4.33), where  $\mu = \cos\theta$  and  $\theta$  is the co-latitude (Snodgrass, 1983; Brown et al., 1989). Finally, the time-dependence on the latitude of the active regions used the best-fitting quadratic model by Li et al. (2001b).

$$\frac{\Omega_m}{2\pi} = 462 - 74\mu^2 - 53\mu^4 \text{ nHz} \quad (4.33)$$

In both instances, the broadened peak was modelled as a single Lorentzian peak using equation (4.13), with a width equivalent to that which was inferred from modelling the BiSON power spectrum. We use `emcee` (Foreman-Mackey et al., 2013) to explore the posterior parameter space with priors similar to the fit to the full power spectrum.

Over the entire duration of the SMMF observations, the time-averaged profile was calculated, using equation (4.28), and this is shown in Fig. 4.14a. The broadened mode used the input parameters outlined in Table 4.2, however, with the background parameter set to zero.

By eye the broadened profile does not appear to have a significantly larger line-width. The fit to the time-averaged broadened peak produced a line-width

of  $0.0315 \pm 0.0041 \mu\text{Hz}$ . The line-width of the broadened peak under this method is rather unchanged from that of the true peak, and both line-widths are within uncertainties of each other.

Table 4.4: Input line-width and the median posterior values of the Lorentzian model each simulation. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median.

Input Value	Weighted Fit	Analytic Fit	Unit
$0.0316(^{+07})_{-07}$	$0.0315 \pm 0.0041$	$0.0315 \pm 0.0041$	$\mu\text{Hz}$

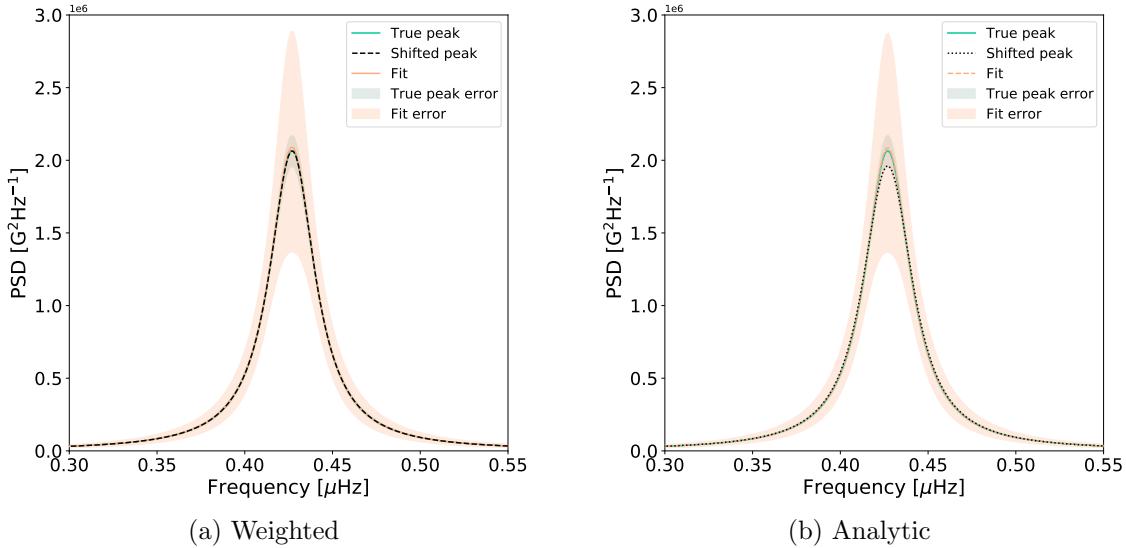


Figure 4.14: (a) Shows the peak distribution before and after the time-averaged broadening, and the fit to the broadened peak. (b) Shows the peak distribution before and after the analytical broadening, and the fit to the broadened peak. In both plots the broadened peaks have been shifted by the relevant frequency to overlay them on top of the true  $\nu_0$  for comparison.

The time-averaged frequency shift due to differential rotation was calculated, much in the same way as equation (4.28), to be  $\Delta\nu = 0.01285 \mu\text{Hz}$ . This shift was used to generate the broadened profile using equation (4.30). The broadened mode distribution also used the input parameters outlined in Table 4.2, however, with the background parameter set to zero.

Similarly to the numerically broadened peak, by eye, the analytically broadened profile does not appear to have a significantly larger line-width and the resultant

median posterior values of the fitted Lorentzian profile are shown in Table 4.4. The line-width of the analytically broadened peak from the fit is  $0.0315 \pm 0.0041 \mu\text{Hz}$ , which is within the uncertainties of the line-width of the input peak.

These results show that both numerically and analytically, the mode broadening effect of differential rotation and latitudinal migration does not affect our ability to resolve the line-width of the peaks. Both broadening methods applied have been shown to have a negligible effect on the measured line-width. This result provides confidence that the line-width in Table 4.2 is the true line-width of the RM peaks, thus providing the correct lifetime for RM component, unaffected by migration and differential rotation.

#### 4.6.3 Further Morphology of the SMMF using SDO/HMI Data

In Chapter 5 we acquired SDO/HMI full-disk magnetograms, using the SunPy python module (Barnes et al., 2020), to support our investigations into Rossby waves.

Owing to having the SDO/HMI magnetograms, which provided the capability to separately analyse the Northern and Southern Hemispheres' MMF contribution to the SMMF during the rising phase of Cycle 24 in 2011 and during solar maximum in 2014, we also investigated whether there were hemispheric differences in the data, which resulted from the opposite polarities at high latitudes and towards the poles. This served as a further analysis into other timescales which may exist in the SMMF. In particular, we investigated if the SMMF exhibited an anti-correlation between the two hemispheres due to the oppositely polarised field near the polar regions, such as that which is found in synoptic charts, on a time-scale of the solar cycle.

To support this investigation, we acquired the synoptic charts from SDO/HMI. It was possible to also average the signal over the Northern and Southern Hemispheres of the synoptic charts, as well as the full solar surface, thus providing a comparison

to the hemispheric MMF and the full disk SMMF.

To compare the magnetogram data to the synoptic charts, we smoothed the separately averaged Northern and Southern Hemispheres' MMF and the full-disk SMMF signals using a box-car filter with a window width of a Carrington period, i.e.  $\sim 27$  days. The resultant time series have been plotted along with the hemispheric mean of the synoptics charts from SDO/HMI. Figure 4.15 shows the resultant smoothed hemispheric MMF and full-disk SMMF along with the average of the synoptic charts.

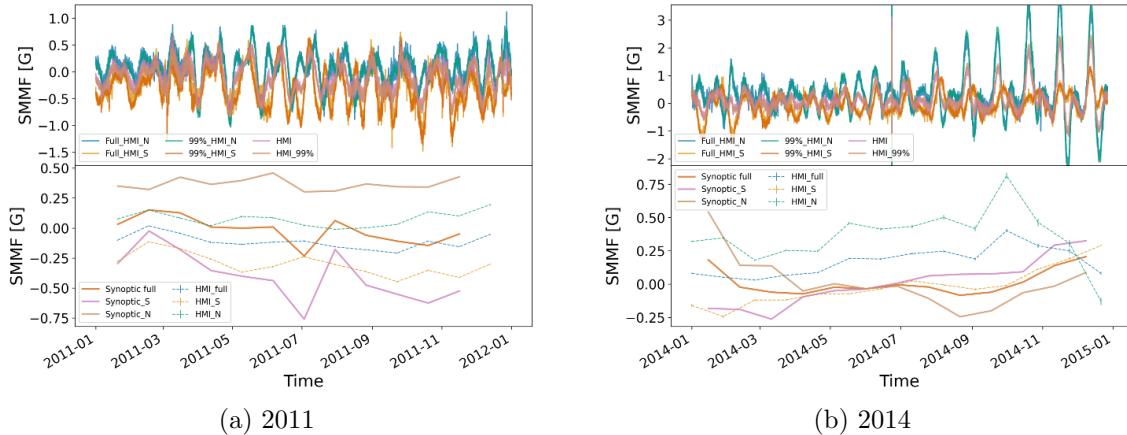


Figure 4.15: Investigations of timescales in the SDO/HMI magnetograms over 2011 and 2014. Both plots show in the top panel, the hemispheric MMF and full-disk SMMF from the magnetograms. The lower panel of each plot displays a comparison between the hemispheric and full-disk mean of the synoptic charts, compared to the box-car smoothed MF from the magnetograms. N: Northern hemisphere; S: Southern hemisphere. Full HMI: considers the full solar disk; 99 HMI: considers only the inner 99% of the solar disk, by radius.

We can see from Figure 4.15 that there does exist a longer timescale in the hemispheric MMF when we average out the effects of the RM component. This is visible from the reversal of the field polarity in 2014. This longer timescale component resembles the average of the synoptic charts, and shows the secular variation which is contributed from the solar dipole at high latitudes. This timescale is the solar activity period, and it can be seen particularly in Figure 4.15b whereby the dipole magnetic field undergoes a the beginning of a reversal at around solar maximum, with the onset of the rush of magnetic field to the poles.

Interestingly, the field reversal is located during different epochs when comparing the synoptic chart data to the visible disk, hemispheric MMF data, and is delayed in the hemispheric MMF by around 7 Carrington rotations.

Naturally, when the full-disk averaged SMMF is smoothed using the box-bar filter, the RM component is averaged out and the result is a near-flat line. This is however expected as it's the average of the opposite hemisphere, and is consistent with the synoptic charts.

## 4.7 Conclusion

Observations of the SMMF have been computed from the Zeeman split D1 line of Potassium at  $\sim 770$  nm, as measured by the Sutherland node of BiSON. The observations covered a period from 1992 – 2012 with a cadence of 40 seconds. A frequency-domain analysis of the SMMF was performed; the short cadence and long baseline of observations gave a fine frequency resolution in the power spectrum up to a high Nyquist frequency, allowing us to probe the elements that underpin the observed SMMF.

The duty cycle for the 40-second cadence observations was very low, hence the effect of the low fill on the power spectrum of the SMMF was investigated to help inform how to best model the full power spectrum. We highlighted that although there appeared to exist a red-noise-like, stochastic background component in the power spectrum, this was a feature added into the power spectrum due to the low duty cycle of the observations.

[Note on the prominent RM signal in the SMMF power spectrum...]

Using a model comprising of a series of Lorentzian peaks to model the RM signal, a Harvey function to account for lower frequency drifts, and shot-noise limit, which was convolved with the Fourier transform of the window function to account for the low duty cycle artefacts, we modelled the full power spectrum and identified the properties of the RM signal.

...

[Don't forget to re-comment on the correlated noise issue...]

To further investigate the SMMF and our ability to infer the properties of the source, we used simulations to analyse the effects of differential rotation and AR migrations on our ability to measure the line-width, and the origin of the asymmetric peak shape.

We leave the reader with the following points:

1. We have shown that there does not exist a short time-scale component in the SMMF, and its emergence was due to the low duty cycle of our observations...
2. We have measured the central frequency of the RM component, allowing us to infer the sidereal period of the RM to be ... suggests an activity cycle average latitude of ...
3. The lifetime of the source of the RM component was inferred from the line-width of the Lorentzian peaks to be ...
4. As the properties of the RM component are consistent with ARs we have presented evidence suggesting them as the source of the dominant signal in the SMMF...
5. We have shown that our ability to determine the line-width and hence lifetime of the RM modes is unaffected by AR migration and differential rotation.
6. Something on asymmetry...

# 5 Rossby Modes in the Solar Mean Magnetic Field

## 5.1 Introduction

Rossby waves, as first derived by Rossby & Collaborators (1939), have recently been discovered in the Sun through observations of near-surface flows by the Solar Dynamic Observatory Helioseismic and Magnetic Imager (SDO/HMI) (Löptien et al., 2018; Liang et al., 2019).

Rossby waves, or  $r$  modes, are toroidal modes of oscillation of a rotating, fluid body for which the dominant restoring force against the pressure gradients is the Coriolis force (Lanza et al., 2019; Hathaway & Upton, 2020). Rossby wave are associated with an undulation of a flow resulting in a pattern of radial vorticity of alternating sign. They are understood to form in the high atmosphere on Earth, heavily influencing global weather. For the Sun we observe that Rossby waves propagate in the retrograde direction, in the Carrington reference frame.

Recently, Löptien et al. (2018) provided an unambiguous detection of sectoral solar  $r$  modes by tracking the horizontal flows of granules in the solar photosphere during a 6-year period, using observations by SDO/HMI. Following this study Liang et al. (2019) confirmed the detection of solar  $r$  modes with time-distance helioseismology to measure deeper, subsurface flows in the meridional direction along the solar equator using both Solar and Heliospheric Observatory Michelson Doppler Im-

ager (SOHO/MDI) and SDO/HMI data, covering 21 years. In addition, Hanasoge & Mandal (2019) were also able to show the detection of solar  $r$  modes using a normal-mode coupling technique on 2 years of SDO/HMI data. In both of the observation conducted by Löptien et al. (2018) and Liang et al. (2019), the average lifetime of the  $r$  modes were on the order of several months, and as long as a over a year for specific modes.

By monitoring the proper motions of solar supergranules using a local correlation tracking method Hathaway & Upton (2020) also report observing low latitude Rossby waves in full-disk Doppler images obtained by SDO/HMI, extending the measurements of Rossby waves to greater depths in the solar atmosphere, by an order of magnitude. The  $r$  modes observed using the supergranules have lifetimes which are only slightly longer than the Carrington rotation period, hence in slight disagreement with Löptien et al. (2018) and Liang et al. (2019), which Hathaway & Upton (2020) claim may be due the waves getting in and out of phase with each other as the low wave number waves propagate faster than the higher wave number waves. As these observations are at low latitudes, it is possible that they are linked to active magnetic regions and hence could be manifested in other sources of magnetic data.

Here we present the possible detection of the to-date unobserved, lowest frequency sectoral  $r$  mode using 21 years of Birmingham Solar Oscillations Network (BiSON) observations of the Line Of Sight (LOS) Solar Mean Magnetic Field (SMMF) (for details on the SMMF and BiSON, see Chaplin et al. (2003) and references within).

In an earlier study of the SMMF, the power spectrum of the BiSON observations was modelled using a combination of Lorentzian peaks and a convolution with the window function to account for the effect of gaps in the data. The full spectrum and the fit are shown in Figure 5.1.

We can divide through by the model, to achieve the residual spectrum, in which

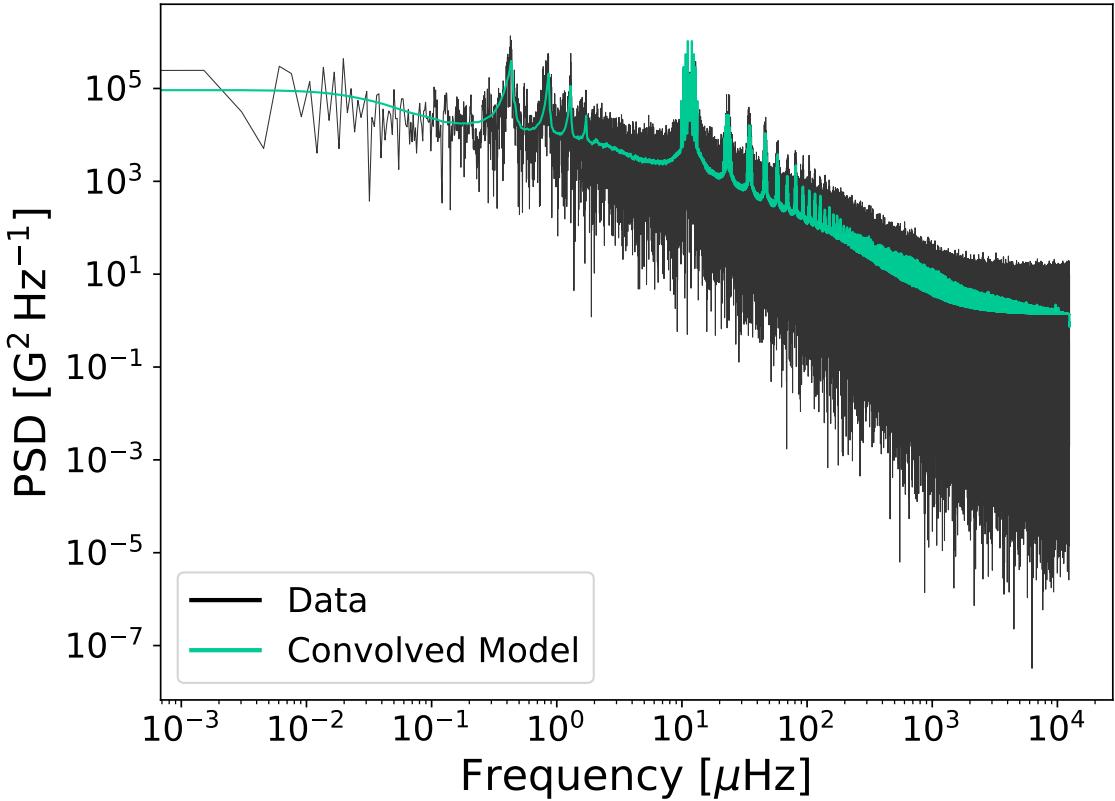


Figure 5.1: Full, modelled power spectrum of the BiSON SMMF on logarithmic axes. The data is displayed in black and the convolved model using asymmetric Lorentzian peaks is shown in green.

we have investigated and potentially found the signature of  $r$  modes. We find that although there exists a promising narrow-band region of significant power in the BiSON residuals spectrum, which could be the  $l = m = 2$  mode, the  $r$  mode does not exist in other sources of SMMF data, hence the observed peak is more likely explained as being a source of prominent noise in the BiSON observations.

## 5.2 Aims

In this chapter the residual power spectrum of the BiSON SMMF, after removing a model for the dominant signal, was investigated in order to search for the existence of  $r$  modes. Within the residual spectrum we test for the existence of  $r$  modes and where they appear to exist, a Lorentzian model is fit to the peak to understand the properties of the source.

Further analysis is provided using simulated data to better understand how an annual modulation of the  $r$  mode observations would affect their frequencies. In addition, SDO/HMI data is investigated to support the argument that  $r$  modes are not split in the power spectrum and instead we should observe the central mode frequency and lower amplitude side bands.

### 5.3 Theory

The detailed theory of the effect of  $r$  modes on observational data was re-visited recently by Lanza et al. (2019) in light of the solar observations, in an effort to determine the effect of the  $r$  modes on radial velocity detections of exoplanets.

Under the assumption of a slow, uniformly rotating sphere (with angular velocity,  $\Omega$ , where  $\Omega^2 \ll GMR^{-3}$ ), the frequencies of global  $r$  modes in the Carrington rotating frame is well approximated by equation (5.1), where  $l > 0$  is the angular degree and  $m$  is the azimuthal order (Löptien et al., 2018; Lanza et al., 2019).

$$\nu_{carr} = -\frac{2m\Omega}{l(l+1)} \quad (5.1)$$

In an inertial frame the observed  $r$  mode frequencies will be (Lanza et al., 2019):

$$\nu_{in}(l, m) \approx m\Omega - \frac{2m\Omega}{l(l+1)} = m\Omega \left(1 - \frac{2}{l(l+1)}\right) \quad (5.2)$$

where  $\Omega$  is the mean sidereal rotation rate, and  $l$  and  $m$  are the angular and azimuthal degree, respectively. Sectoral Rossby waves are obtained by setting  $l = m$  in this equation. A consequence is that they propagate with a retrograde phase velocity as  $\nu/m = -2\Omega/[m(m+1)] < 0$ .

An Earth-based observer, orbiting the sun, shall expect to observe frequencies adjusted by the orbital frequency,  $\nu_{\oplus} \approx 31.7\text{nHz}$ , given by equation (5.3).

$$\nu_{obs}(l, m) = \nu_{in}(l, m) - m\nu_{\oplus} \quad (5.3)$$

In addition, due to the tilt of the ecliptic with respect to the solar equatorial plane (the solar  $B_0$  angle), the visibility of the modes will vary on a timescale of 1 year, meaning we expect to actually observe split peaks at frequencies of  $\nu_{\text{obs}}(l, m) \pm \nu_{\oplus}$  (Lanza et al., 2019)).

Based on this theory, we can compare the observed sectoral  $r$  modes with those predicted. These frequencies are summarised in Table 5.1.

Table 5.1: Predicted<sup>+</sup> and observed<sup>◦</sup>  $r$  mode frequencies for combinations of  $l$  and  $m$ . Predicted frequencies and conversions of observations to different frames of reference use equation (5.1), equation (5.2), and equation (5.3), with  $\Omega = 453.1$  nHz. The predicted splitting for the  $B_0$  angle variation is also provided. The key for the source column is: LPT for Löptien et al. (2018), LNG for Liang et al. (2019), and LZA for Lanza et al. (2019).

Frequency	Source	$l = m = 2$	$l = m = 3$	$l = m = 4$	$l = m = 5$
$\nu_{\text{carr}} [nHz]$	LPT <sup>◦</sup>	–	-259	-194	-157
	LNG <sup>◦</sup>	–	-253	-198	-156
	LZA <sup>+</sup>	-302.1	-226.6	-181.2	-151.0
$\nu_{\text{in}} [nHz]$	LPT	–	1100	1618	2109
	LNG	–	1106	1614	2110
	LZA	604.1	1132.8	1631.2	2114.5
$\nu_{\text{obs}} [nHz]$	LPT	–	1005.2	1491.7	1950.1
	LNG	–	1011.2	1487.7	1951.1
	LZA	540.8	1037.7	1504.4	1956.0
$\nu_{\text{obs}} + \nu_{\oplus} [nHz]$	LPT	–	1036.9	1523.3	1981.8
	LNG	–	1042.9	1519.3	1982.8
	LZA	572.4	1069.4	1536.1	1987.7
$\nu_{\text{obs}} - \nu_{\oplus} [nHz]$	LPT	–	973.6	1460.0	1918.4
	LNG	–	979.6	1456.0	1919.4
	LZA	509.1	1006.0	1472.7	1924.3

## 5.4 Methodology

### 5.4.1 Testing the Residual Spectrum

In order to investigate the presence of Rossby wave modes in the power spectrum of the BiSON SMMF, statistical significance tests were employed using a false-alarm approach, to test the probability of finding prominent narrow-band power in the residual spectrum.

We assume negative exponential statistics (i.e.  $\chi^2$  2-degrees of freedom distribution), and that the bins in the power spectrum are uncorrelated (an assumption which may be incorrect due to the low fill-factor of the BiSON SMMF data). Then we can find the false alarm probability, or probability to observe power in a given frequency bin,  $\nu$ , that is greater than or equal to a given threshold.

The probability to observe power in a given frequency bin,  $\nu$ , that is greater than or equal to  $P(\nu)$  is:

$$p[P(\nu)] = \frac{1}{P_{lim}(\nu)} \exp\left(-\frac{P(\nu)}{P_{lim}(\nu)}\right) \text{ or; } p[P(\nu)] = \frac{1}{\langle P(\nu) \rangle} \exp\left(-\frac{P(\nu)}{\langle P(\nu) \rangle}\right), \quad (5.4)$$

where  $P_{lim}(\nu)$  is the limit spectrum or  $\langle P(\nu) \rangle$  is a well-fitting model/estimate to the spectrum. Considering a relative power approach (i.e. considering the power relative to the mean level or model of the Power Spectral Density (PSD)), equation (5.4) may be written more concisely as:

$$p(s_\nu) = e^{(-s_\nu)}, \quad (5.5)$$

where,

$$s_\nu = P(\nu)/\langle P(\nu) \rangle, \quad (5.6)$$

and  $\langle P(\nu) \rangle$  is reduced to 1 when we compare the power relative to the power spectrum residuals.

In reality, we use the  $\chi^2$  cumulative distribution function to compute the probability of power, which is given by equation (5.7), where  $k$  is the number of degrees of freedom,  $\gamma(s, t)$  is the lower incomplete gamma function and  $P(s, t)$  is the regularized gamma function.

$$F(x; k) = \frac{\gamma(\frac{k}{2}, \frac{x}{2})}{\Gamma(\frac{k}{2})} = P\left(\frac{k}{2}, \frac{x}{2}\right) \quad (5.7)$$

Using these expressions, we can rewrite the equation for  $P(s_\nu)$  as given by equation (5.8). This allows us to compute the probability of statistically significant peaks in the residuals for re-binned data.

$$p(s_\nu) = 1 - F(2s_\nu; k) = 1 - P\left(\frac{k}{2}, s_\nu\right) \quad (5.8)$$

The probability that a bin has power at or above the level  $s_\nu$  is therefore given by equation (5.5), or more generally by equation (5.8), hence the probability that we fail to find a bin with power at or above the level  $s_\nu$  is  $1 - p(s_\nu)$ ; thus the probability of failing to find a bin with power at or above  $s_\nu$  in  $N$ -bins in the spectrum is  $[1 - p(s_\nu)]^N$ . Therefore the probability to find at least one bin with power at or above  $s_\nu$  in  $N$ -bins in the spectrum is:

$$p_N = 1 - [1 - p(s_\nu)]^N \quad (5.9)$$

where a low value for  $p_N$  indicates that the spike in power in that bin is unlikely to be a statistical fluctuation, and therefore is considered a statistically significant spike.

This can be generalised using the cumulative binomial distribution. The probability of finding at least  $r$  spikes in  $N$ -bins at or above the relative power level  $s_\nu$  is given by equation (5.10), which is equal to equation (5.9) when  $r = 1$ .

$$p[r; p(s_\nu), N] = 1 - Pr(X < s_\nu) = \sum_{r=r}^N \binom{N}{r} p(s_\nu)^r [1 - p(s_\nu)]^{N-r} \quad (5.10)$$

By applying equation (5.10) to the residuals of in the power spectrum, we can test whether there are any significant peaks in the residual power spectrum. Again,

a low value for  $p[r; p(s_\nu), N]$  indicates that the power in that bin is unlikely to be a statistical fluctuation.

### 5.4.2 Modelling $r$ mode Profiles

In the location of any suspected  $r$  modes in the residual spectrum, we can model the profile of the peak by using a Lorentzian distribution, using the form expressed by equation (5.11), where  $A$  is the amplitude of the signal,  $\Gamma$  is the line-width of the distribution, and  $\nu_0$  is the central frequency of the distribution. This follows the methodology adopted by Löptien et al. (2018) and Liang et al. (2019).

$$P(\nu) = \frac{2A^2/(\pi\Gamma)}{1 + (2(\nu - \nu_0)/\Gamma)^2} \quad (5.11)$$

Modelling was performed using the `pymc3` No U-Turn Sampler (NUTS) extension to a Hamiltonian Monte Carlo (HMC) sampling algorithm (Salvatier et al., 2016).

## 5.5 Results

### 5.5.1 Testing the Residual Spectrum

The residual spectrum is shown in Figure 5.2, with the removed model and locations of  $r$  mode frequencies predicted by Lanza et al. (2019) over-plotted.

It is clear from Figure 5.2 that there appears to be a resolved peak of narrow-band power in both the location of the  $l = m = 2$   $r$  mode and perhaps also the upper  $B_0$ -variation-modulated  $l = m = 3$   $r$  mode.

The statistics tests were performed on the residual spectrum for various re-binning factors,  $n$ . The plots summarising the statistics tests are shown in Fig. 5.3 for re-binning factors of  $n = 1, 2, 5$ , and  $10$ .

These tests suggest that the feature around  $0.5 \mu\text{Hz}$  may be a signal related to  $r$  modes, as its location is roughly correct for the  $l = m = 2$  mode consistently has a low False Alarm (FA) probability. The other low FA probability features tend to

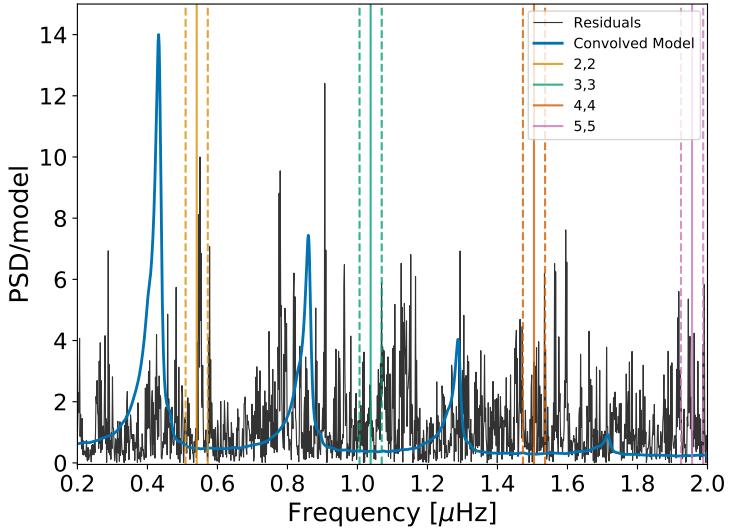


Figure 5.2: Residual power spectrum of the BiSON SMMF. Over plotted in the green curve is the model of the main SMMF signal. Also over plotted as vertical solid lines are the expected locations of the 4 lowest-frequency sectoral  $r$  modes and the dashed lines, the locations of the  $B_0$  variation frequency splitting.

be associated with the residuals from the main harmonics of the SMMF signal, and are the result of a slight failing of the main model of the power spectrum.

In particular in Figure 5.3b, there is compelling evidence to suggest that this peak is significant. In order to solidify this conjecture, we aimed to fit a model to the residuals around this peak in order to confirm whether the properties of the peak resembled those suggested by Löptien et al. (2018), Liang et al. (2019), and Lanza et al. (2019).

### 5.5.2 Modelling $r$ mode Profiles

Using the model for the Lorentzian peak (eq. 5.11), we modelled the residual spectrum around the location of the potential  $l = m = 2$  mode. The results of the fit are given in Table 5.2 and the fit to the residuals is shown in Figure 5.4.

The median of the posterior distribution for the width parameter suggests an e-folding lifetime ( $1/(\pi\Gamma)$ ) of around 650 days ( $\sim 1.8$  years), which follows the order magnitude of the lifetime suggested by low- $m$  Rossby modes. The lowest  $r$  mode observed,  $l = m = 3$ , was shown to exhibit a lifetime of over a year ( $\sim 1.4$  years) by

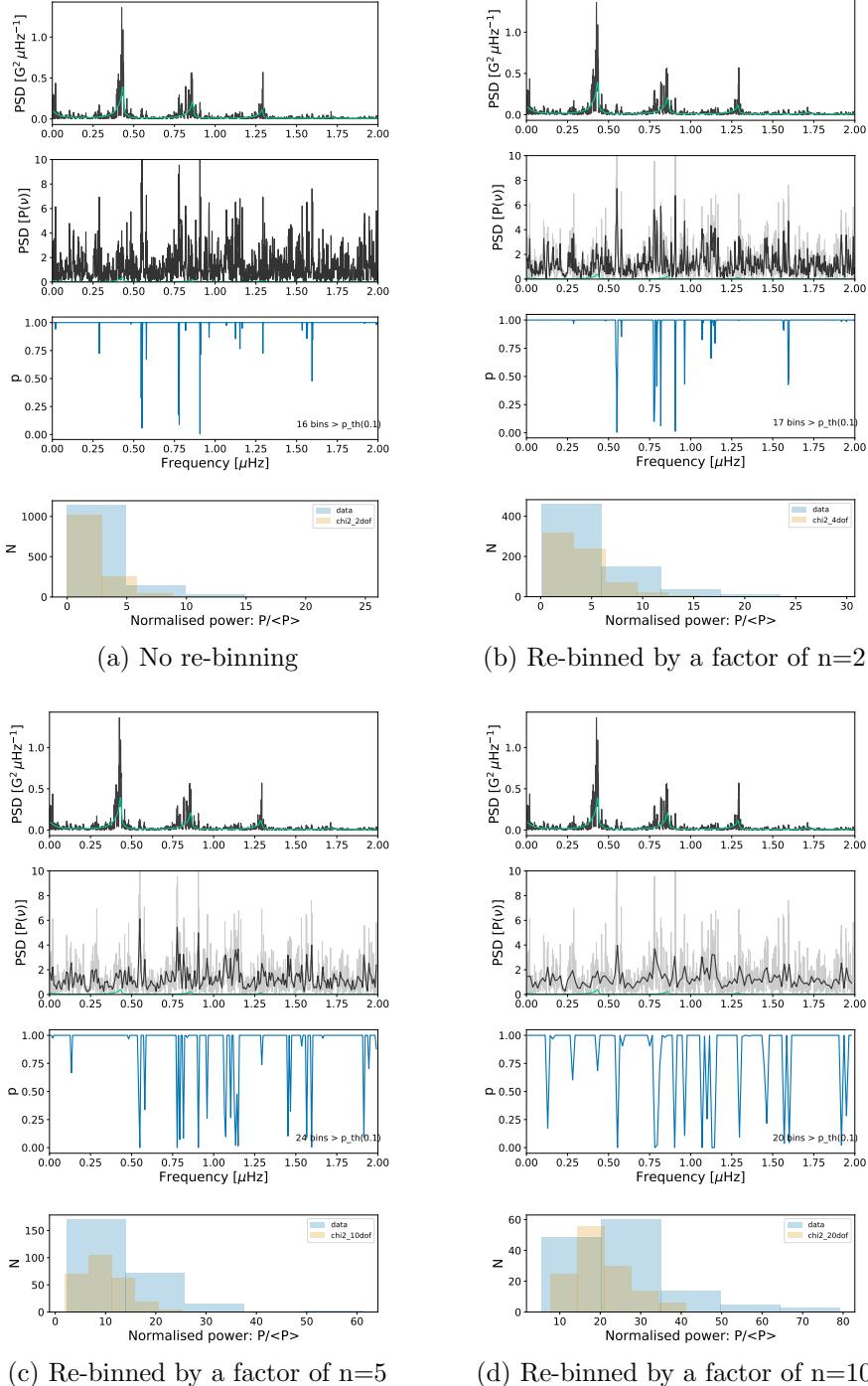


Figure 5.3: Realisations of the statistics tests of BiSON data for different re-binning factors ( $n$ ). The panels of each sub figure are: (top) the full PSD and fit, (second panel) the full and re-binned residuals, (third panel) the probability of statistical noise in each bin, (bottom) distribution of the residuals compared to a  $\chi^2$   $2n$ -DOF.

Liang et al. (2019). The  $l = m = 4$  were observed by both Löptien et al. (2018) and Liang et al. (2019) to have a lifetime of  $\sim 0.6$  years and  $\sim 0.3$  years, respectively.

Table 5.2: Median posterior values of the Lorentzian model for the  $r$  mode peak in the BiSON SMMF PSD residuals. Numbers in brackets denote uncertainties on the last 2 digits, and all uncertainties correspond to the 68% credible intervals either side of the median.

Parameter	Value	Unit
$\nu_0$	$0.5500^{(+16)}_{(-16)}$	$\mu\text{Hz}$
$\Gamma$	$0.0058^{(+45)}_{(-28)}$	$\mu\text{Hz}$
$A$	$0.261^{+0.063}_{-0.049}$	—
<i>bgnd</i>	$1.00 \pm 0.16$	—

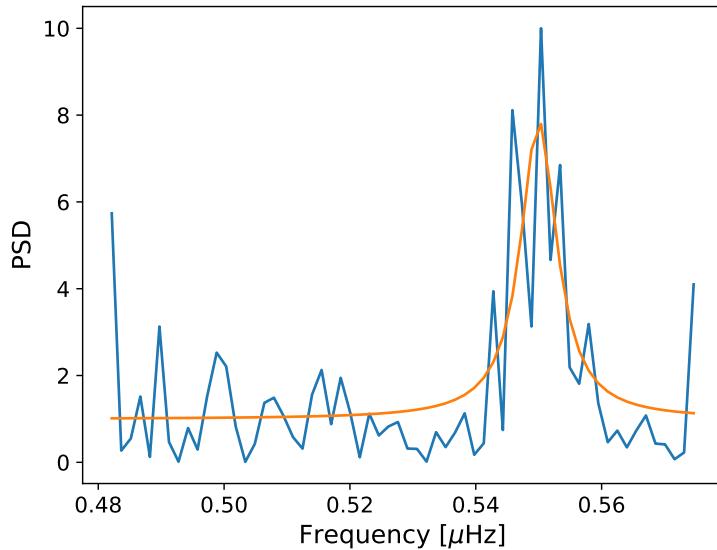


Figure 5.4: Model fit to the  $r$  mode in the BiSON PSD residuals using the median from the parameter posterior distributions.

There seems to be a slight increasing trend in the mode lifetimes observed by Löptien et al. (2018) and Liang et al. (2019) of longer lifetimes for lower  $m$ , therefore the long lifetime found here for the  $l = m = 2$  mode is entirely reasonable and in-line with current observations.

The model of the mode in the residuals spectrum has a magnetic amplitude of  $\sim 30$  mG, which equates to a radial velocity amplitude of  $\sim 8.7 \text{ cm}^{-1}$ . Lanza et al. (2019) state that the maximum RV amplitude of the  $l = m = 2$  mode is  $\sim 24.5 \text{ cm}^{-1}$ , meaning our observed peak is around a third of the maximum RV amplitude one might have expected to observe. This is however an upper limit given by Lanza et al. (2019), and therefore the lower amplitude in the model should not be concerning.

In particular, however, we can see that the background is  $\sim 1$ , which is expected for such a fit to residual PSD data. Due to the accuracy of the background, the agreement between the central frequency of the fit to the  $l = m = 2$  mode, the agreement in the order of magnitude of the e-folding lifetime, and the amplitude of the mode residing below the upper limit suggested by Lanza et al. (2019), we believe there is strong evidence to suggest that this peak, which has shown to be significant through the false-alarm statistics tests, is the  $l = m = 2$  sectoral  $r$  mode, observed in the BiSON SMMF data.

## 5.6 Discussion

Despite the compelling results presented in the previous section there remained an open question on the way the Rossby waves would manifest themselves in the power spectrum. The BiSON power spectrum was also compared to the power spectra of the Wilcox Solar Observatory (WSO) SMMF and the SDO/HMI SMMF to cross-reference the finding.

### 5.6.1 Manifestation of Rossby Waves in the Power Spectrum

It was suggested by Lanza et al. (2019) that the mode should be split into two frequencies due to the annual variation of the  $B_0$  angle, but the observed peak in the BiSON spectrum is located at approximately the location of the central frequency; therefore the observation is not split due to this modulation. We needed to determine if this was physically observable.

Figure 5.5 shows a schematic diagram of the flow of a  $l = 2 = m$  sectoral  $r$  mode. One can clearly see from the more visible purple region of the schematic, the Southern Hemisphere flow is oriented out of the page, whereas the Northern Hemisphere flow is oriented into the page. Due to the  $B_0$  modulation, a varying of the sign of the flow would be observed over this region, i.e the velocity of the flow.

In the more red-green regions the schematic, the flow is more transverse, hence this would contribute less to the effect of the  $B_0$  modulation.

We needed to therefore understand whether the SMMF observations have a hemispheric dependence that would lead to a change sign in the observations due to the  $B_0$  variation. In addition, this raised the question of how the mode was affected in the power spectrum by the  $B_0$  modulation; either split into separate peaks as suggested by Lanza et al. (2019) or instead was it possible that we could have a situation where the mode at the central frequency remained?

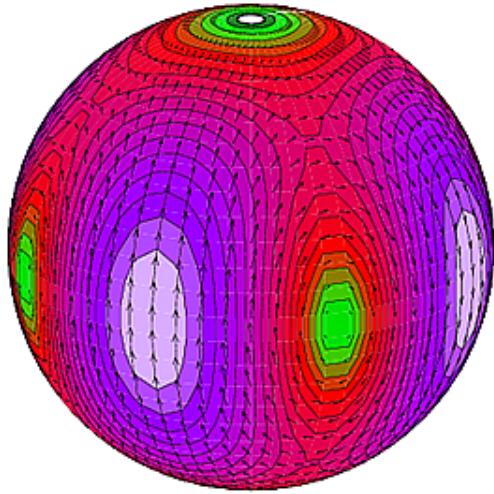


Figure 5.5: Mode displacement schematic for an  $l = m = 2$   $r$  mode (Strohmayer & Mahmoodifar, 2014)

To investigate the splitting of the mode, a simple model was produced whereby a sinusoidal function (with a period of  $\sim 25$  days) was modulated by either a cosine or rectified cosine function (with a period of 1 year). In the former, using the cosine modulation, this represents observing the sign of the flow varying with the  $B_0$  modulation. Conversely in the latter simulation, this instead represents a variation of the amplitude, and it does not change the sign. Figure 5.6 shows the time series of the two cases to more clearly show their difference. The power spectrum of each case was then computed and these are shown in Figure 5.7 to demonstrate the differences between the modes produced.

One can clearly see the difference in the power spectrum produced in each of the

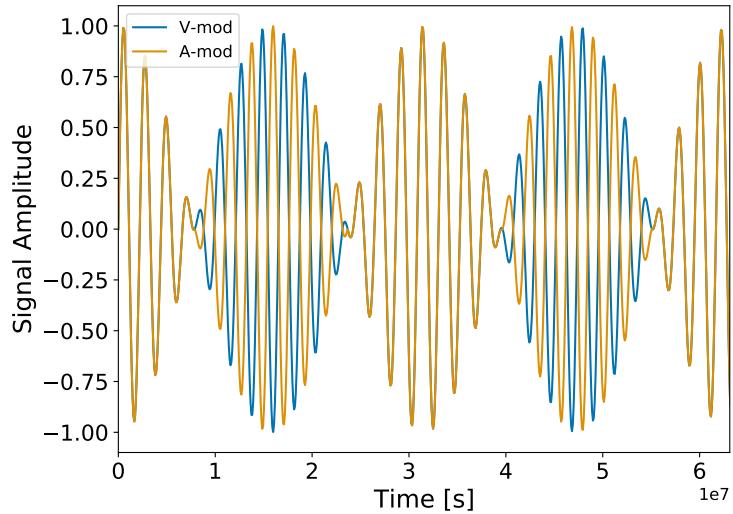


Figure 5.6: Time series of the velocity and amplitude modulation toy model simulations. The blue curve shows the velocity modulation, i.e. modulating using a cosine with period 1 year, whereas the orange curve shows the amplitude modulation, i.e. modulating using a rectified cosine with period 1 year.

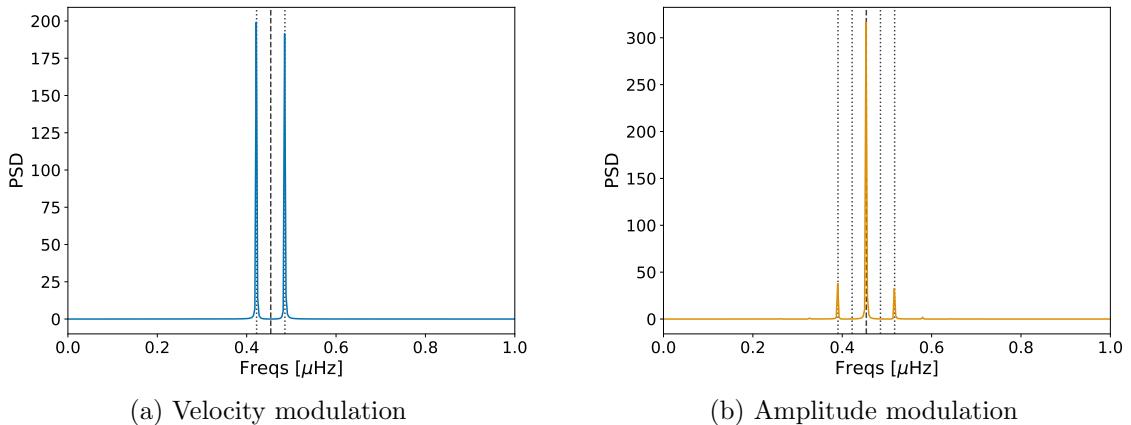


Figure 5.7: Power spectra for the two modulation methods, showing the difference in the way the modulation has changed the frequency of the observed mode.

cases. In the velocity modulation case, we see a splitting of the oscillation mode into two peaks split around the expected mode frequency by  $\pm\nu_{\oplus}$ . There exists no more power at the central mode frequency in this case and is in agreement with the scenario suggested by Lanza et al. (2019). In the amplitude modulation case, we see a production of sidebands at  $\pm 2\nu_{\oplus}$  however the expected mode frequency remains in this scenario and has a significantly higher peak height, a ratio of 90:10 in favour of the central peak.

We have shown that it is possible to retain the central frequency of the  $r$  mode in the power spectrum if the  $B_0$  modulates the amplitude of the observations and not the sign. With this known, it was then necessary to understand whether the two hemispheres of the Sun contribute signals that are more analogous with the velocity modulation or amplitude modulation. In the former, velocity modulation, we would expect to see a persistent anti-correlation between the two hemispheres. In the latter, amplitude modulation, we would expect to see the signals from each hemisphere that are correlated, which track each other and which can be both positive or negative.

We investigated how the two hemispheres of the Sun contribute to the SMMF through analysis of SDO/HMI data. To do this, we acquired 720s-cadence magnetograms from SDO/HMI using the SunPy python module (Barnes et al., 2020) for the rising phase of solar cycle 24 during 2011, and for the maximum of cycle 24 during 2014. It was possible to separately average the Northern and Southern Hemispheres' contributions to the total, disk-averaged SMMF. The plot of this data is shown in Figure 5.8.

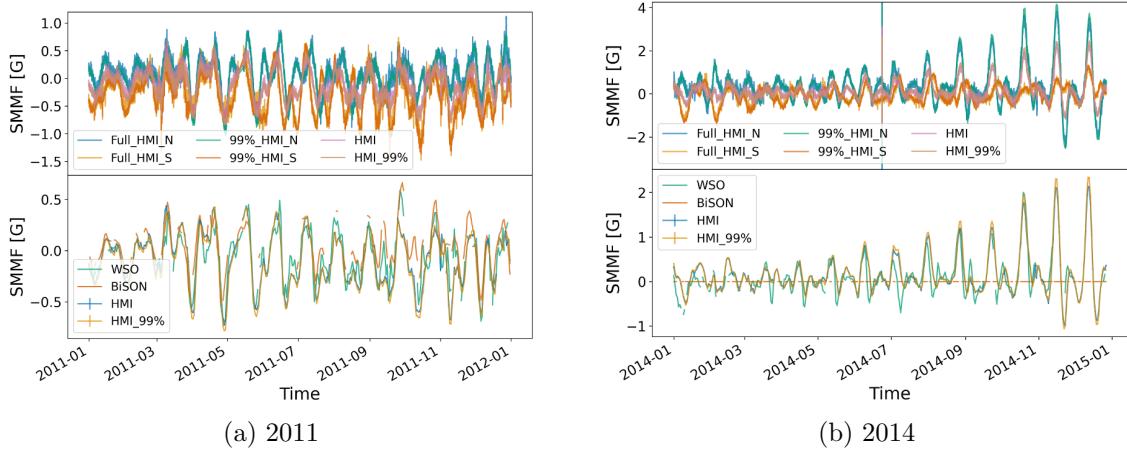


Figure 5.8: SDO/HMI SMMF split into hemispheres and compared to other SMMF sources during (a) 2011, and (b) 2014. The top panel in each figure shows the north (N), south (S), and total disk-averaged mean magnetic field, for both the full solar disk and from pixels within 99% of solar radius. The bottom panels show a comparison between the SMMF, as observed BiSON, WSO, and SDO/HMI (full disk and the 99% disk).

The hemispheric contributions to the total, disk-averaged, SMMF can be seen in Figure 5.8 to track each other during 2011, and both become positive or negative. By contrast, when observing the hemispheric contributions to the full SMMF in 2014, we see that there are frequent periods of strong anti-correlation between the Northern and Southern Hemispheres. There are also several periods in 2014 whereby the two hemispheres are correlated. This plot aids in our understanding of how the  $r$  modes would be manifested in the power spectrum due to the variation of the  $B_0$  angle.

As there are periods of both strong correlation and strong anti-correlation between the two hemispheres, this is a good indication that the  $r$  mode signal would result in a central frequency with sidebands due to the correlation between North and South. But due to the existence pf periods of anti-correlation between North and South, we also expect there to be some degree of frequency splitting in the power spectrum, but this is dependent on how prevalent the anti-correlation is over the entire solar cycle; in these short epochs however, we expect this to be minimal. It is therefore possible to conclude that we are confident we are observing the  $l = 2 = m$   $r$  mode.

As a further point, we see that there is both strong correlation and strong anti-correlation between the Northern and Southern Hemispheres, however this does not necessarily mean that the  $r$  mode signal would directly manifest itself in the same way. We can see from Figure 5.5 that if the  $r$  mode observations are constrained to active latitude bands, closer to the equator, then the effects of the  $B_0$  are less prominent.

### 5.6.2 Rossby Modes in Other Sources of SMMF Data

To further investigate whether the observation of the  $l = 2 = m$   $r$  mode is real, a comparison was made between the power spectrum of the BiSON observations of the SMMF and those from WSO and SDO/HMI, to determine whether the suspected

$r$  mode is visible. In the case of the WSO, the power spectrum was computed over the same observing epoch as the BiSON data (i.e. from 1992 – 2013); however SDO was not launched until 2010, so for HMI the power spectrum was computed on data from 2010 – 2020, hence at approximately half the frequency resolution of WSO and BiSON.

Figure 5.9 shows the comparison of the WSO and SDO/HMI power spectra reflected around the x-axis against the BiSON power spectrum. In both cases we see a good agreement between the different sets of data on the location of the rotational mode in the SMMF, but there does not appear to be a visible  $r$  mode candidate in the WSO or the SDO/HMI data.

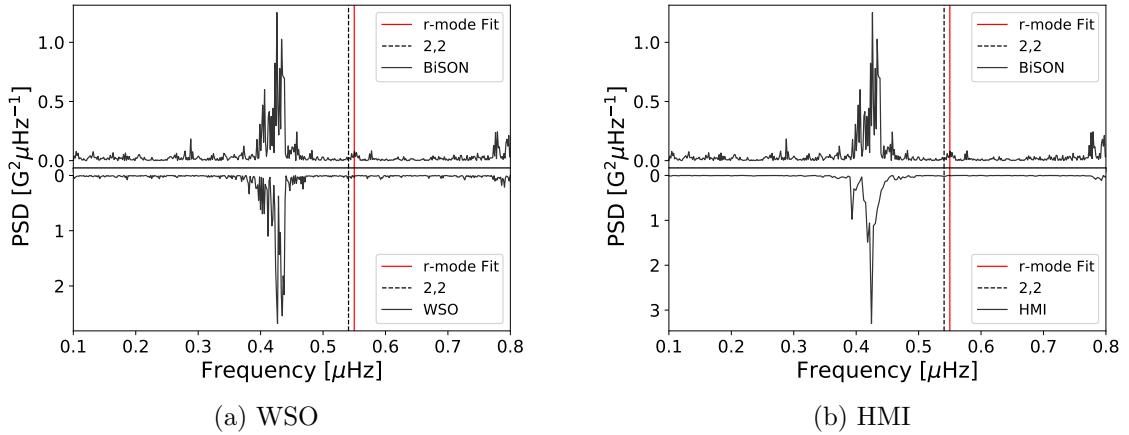


Figure 5.9: Comparison of the power spectra for BiSON, WSO, and SDO/HMI. In both figures, the top panel shows the BiSON PSD and the bottom panel shows either the WSO or HMI PSD. The dashed, black line shows the location of theoretical  $l = 2 = m$   $r$  mode frequency, and the red, solid line shows the location of the peak fit in the BiSON PSD residuals.

This provides concerning evidence that we may not, in fact, have observed the  $l = 2 = m$   $r$  mode in the BiSON power spectrum, and perhaps instead we have a persistent noise source in the BiSON SMMF.

However convincing the results are suggesting that we have observed the  $l = 2 = m$   $r$  mode in the BiSON power spectrum, we cannot ignore that it is not present in two other sources of SMMF observations, of which one of these telescopes is responsible for providing the recent observations of  $r$  modes. Owing to this, we

cannot conclude that we have observed the  $l = 2 = m$   $r$  mode in the BiSON power spectrum. Searching for  $r$  modes in the power spectra of the SMMF observations should be revisited in a few years, or say a solar cycle's time, when the frequency resolution in each spectra has increased significantly, to provide a follow-up study.

## 5.7 Conclusion

After removing the model for the BiSON SMMF power spectrum we investigated the residual spectrum to search for the existence of Rossby wave modes ( $r$  modes). Using a false-alarm approach we estimated the probability of finding narrow-band power. We identified the existence of a statistically significant peak which was near the  $l = 2 = m$   $r$  mode frequency calculated by Lanza et al. (2019).

Using a Lorentzian model for the peak, following the description of  $r$  mode given by Löptien et al. (2018) and Liang et al. (2019), we identified the properties of the peak, and compared them to the existing literature. The work by Lanza et al. (2019) stated that the  $l = 2 = m$   $r$  mode should be observed at a frequency of  $\sim 540.8$  nHz, with a maximum amplitude of  $24.5\text{ cm}^{-1}$ . The recent observations of low-degree sectoral  $r$  modes by Löptien et al. (2018) and Liang et al. (2019) claimed the line-width of the mode would be on the order of around 10 nHz, due to the e-folding lifetime of the mode being on the order of a year.

To further interrogate the inferences on the  $r$  mode in the BiSON SMMF power spectrum, we use simple simulations to determine how the  $r$  mode may manifest itself in the power spectrum, using the modulation of the signal due to the variation in the  $B_0$  angle. Furthermore, SDO/HMI hemispheric data was employed to verify these results.

As a final check, we compared the power spectrum of the SMMF observations from BiSON to those from WSO and SDO/HMI. There was no clear signature of the  $l = 2 = m$   $r$  mode in either of the other power spectra, which rules our observation of an  $r$  mode in the BiSON SMMF very suspicious, especially as the recent observations

of sectoral Rossby waves in the Sun have all used SDO/HMI data.

We leave the reader with the following points:

1. Through a series of false-alarm probability statistical tests, we have shown that there exists a statistically significant peak in the BiSON SMMF residuals spectrum which is located near the theoretical frequency of the  $l = 2 = m r$  mode.
2. By modelling the peak as a Lorentzian profile we find that the peak has a central frequency of  $550 \pm 16$  nHz (i.e. located  $\sim 9.2$  nHz from the theoretical frequency), a line-width of  $5.8^{+4.5}_{-2.8}$  nHz, and an amplitude of  $\sim 30$  mG. This profile is within the upper limit for the amplitude of the  $l = 2 = m r$  mode and the life time implied by the line-width is on the order of 1–2 years, which is in agreement with the observations by Löptien et al. (2018) and Liang et al. (2019).
3. Through the analysis of simulated data and hemispheric observations of the SMMF, we have shown that we should expect to see a prominent mode at the theoretical frequency, and not a split mode due to the effect of the  $B_0$  variation, which supported the findings of the  $r$  mode.
4. By comparing the power spectrum of the SMMF observed by BiSON, to those of WSO and SDO/HMI, we have shown that the  $r$  mode peak is not manifested in either the WSO or SDO/HMI spectra, therefore ruling it highly unlikely that the observed peak in the BiSON spectrum is the  $l = 2 = m r$  mode.

As we collect more observations of the SMMF using BiSON, the frequency resolution of the power spectrum increases. An obvious next step in this work is to collect more observations of the SMMF with BiSON, to further investigate if this suspected mode remains resolved, or whether it diminishes into the noise.

# A Simulations of Cosmic Ray Air Showers using CORSIKA

There are several user-definable settings within Cosmic Ray Simulations for Cascade (CORSIKA). These settings are explained in-depth in the CORSIKA user’s guide (Heck & Pierog, 2017). Here, the options selected when building and using CORSIKA are outlined.

For high energy, inelastic hadronic interactions within CORSIKA the QGSJET-II (Ostapchenko, 2006) model was selected. Interactions of hadrons with energies below 80 GeV are simulated using GHEISHA (Fesefeldt, 1985), which allowed for the simulation of PCRs in the regime of SCRs. In addition to these hadronic interactions, electromagnetic interactions within the CORSIKA simulations were described by the EGS4 (Nelson et al., 1985) model. Furthermore CORSIKA has a minimum muon energy limit that can be simulated of 10 MeV. This limit is sufficient, as the lowest energy muons that are observable with HiSPARC are those with energy of  $\sim 4$  MeV.

Simulation thinning was enable to reduce the computation time of the simulations and reduce the output file size. The observation level at which point the simulation cease was set at 100 m above sea level (compared to the  $\sim 50$  m typical of the stations; however, this difference is negligible for the air shower development). The pre-defined central European atmosphere in October was used for all simulations, and western-European magnetic field was used as calculated with the *Geomag*

programme [REF]:  $B_x = 18.799 \mu\text{T}$  and  $B_z = 44.980 \mu\text{T}$ .

Proton and  $\alpha$ -particle initiated air showers were generated with energies ranging from  $10^9$  to  $10^{20}$  eV, and  $4 \times 10^9$  to  $10^{20}$  eV, respectively. In total  $\sim 230000$  proton-initiated showers were simulated and  $\sim 180000$   $\alpha$ -particle-initiated air showers were simulated. The simulated air-shower-initiating PCRs are listed Table A.1 (proton) and Table A.2 ( $\alpha$ -particles).

Table A.1: Details of the proton-initiated air showers simulated using CORSIKA.

<b>E<sub>PCR</sub> (eV)</b>	<b>N<sub>sims</sub></b>	<b>E<sub>PCR</sub> (eV)</b>	<b>N<sub>sims</sub></b>
1.00E+09	10000	2.98E+12	1000
1.27E+09	10000	3.79E+12	1000
1.62E+09	10000	4.83E+12	1000
2.07E+09	10000	6.16E+12	1000
2.64E+09	10000	7.85E+12	1000
3.36E+09	10000	1.00E+13	1000
4.28E+09	10000	1.78E+13	100
5.46E+09	10000	3.16E+13	100
6.95E+09	10000	5.62E+13	100
8.86E+09	10000	1.00E+14	100
1.00E+10	10000	1.78E+14	50
1.13E+10	10000	3.16E+14	50
1.44E+10	10000	5.62E+14	50
1.83E+10	10000	1.00E+15	10
2.34E+10	10000	1.78E+15	10
2.98E+10	10000	3.16E+15	10
3.79E+10	10000	5.62E+15	10
4.83E+10	10000	1.00E+16	10
6.16E+10	10000	1.78E+16	10
7.85E+10	10000	3.16E+16	10
1.00E+11	10000	5.62E+16	10
1.27E+11	1000	1.00E+17	10
1.62E+11	1000	1.78E+17	10
2.07E+11	1000	3.16E+17	10
2.64E+11	1000	5.62E+17	10
3.36E+11	1000	1.00E+18	10
4.28E+11	1000	1.78E+18	10
5.46E+11	1000	3.16E+18	10
6.95E+11	1000	5.62E+18	10
8.86E+11	1000	1.00E+19	10
1.13E+12	1000	1.78E+19	10
1.44E+12	1000	3.16E+19	10
1.83E+12	1000	5.62E+19	10
2.34E+12	1000	1.00E+20	10

Table A.2: Details of the  $\alpha$ -particle-initiated air showers simulated using CORSIKA.

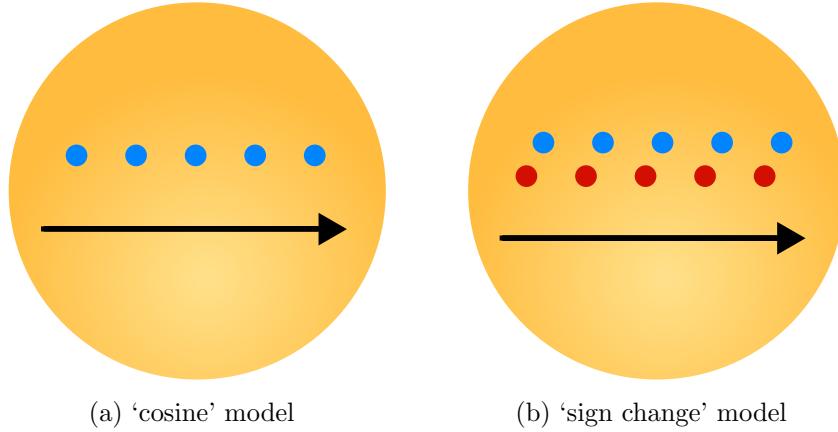
<b>E<sub>PCR</sub> (eV)</b>	<b>N<sub>sims</sub></b>	<b>E<sub>PCR</sub> (eV)</b>	<b>N<sub>sims</sub></b>
4.00E+09	10000	1.00E+13	1000
4.28E+09	10000	1.78E+13	100
5.46E+09	10000	3.16E+13	100
6.95E+09	10000	5.62E+13	100
8.86E+09	10000	1.00E+14	100
1.00E+10	10000	1.78E+14	50
1.13E+10	10000	3.16E+14	50
1.44E+10	10000	5.62E+14	50
1.83E+10	10000	1.00E+15	10
2.34E+10	10000	1.78E+15	10
2.98E+10	10000	3.16E+15	10
3.79E+10	10000	5.62E+15	10
4.83E+10	10000	1.00E+16	10
6.16E+10	10000	1.78E+16	10
7.85E+10	10000	3.16E+16	10
1.00E+11	10000	5.62E+16	10
1.27E+11	1000	1.00E+17	10
1.62E+11	1000	1.78E+17	10
2.07E+11	1000	3.16E+17	10
2.64E+11	1000	5.62E+17	10
3.36E+11	1000	1.00E+18	10
4.28E+11	1000	1.78E+18	10
5.46E+11	1000	3.16E+18	10
6.95E+11	1000	5.62E+18	10
8.86E+11	1000	1.00E+19	10
1.13E+12	1000	1.78E+19	10
1.44E+12	1000	3.16E+19	10
1.83E+12	1000	5.62E+19	10
2.34E+12	1000	1.00E+20	10
2.98E+12	1000		
3.79E+12	1000		
4.83E+12	1000		
6.16E+12	1000		
7.85E+12	1000		

## B Simulations of the Artificial SMMF

### B.1 Model

The artificial data used in the simulations of the Solar Mean Magnetic Field (SMMF) was created using 2 very simplistic models, either separately or in combination, which were physically motivated by sunspots and active regions on the solar surface:

1. **Cosine:** in this model a source was simulated as a single region which appears on the visible disk from one limb, traversing across the disk, before disappearing off the other limb (see Fig. B.1a). The physical motivation for this would be a single, concentrated source of imbalanced magnetic flux. This method induces no sign change of the simulated source; it remains the same polarity always. This was simulated as a rectified sinusoidal signal, with a 50% duty cycle (i.e. it is only visible during times on near-side of the disk).
2. **Sign change:** in this model a source was simulated as 2 regions of opposite polarity, such as sunspot pairs or a Bipolar Magnetic Region (BMR). The leading region contributes more at the start of the transit during ingress, and the trailing source contributes more at the end of the transit during egress; hence at the middle of the transit we assume there is a sign change in the overall signal, see Fig. B.1b. This was simulated as a rectified sinusoidal signal, and was multiplied by a cosinusoidal signal of the same period to provide the projection of the pair of regions. This operation results in a rectified sinusoidal signal with half the period.



(a) ‘cosine’ model

(b) ‘sign change’ model

Figure B.1: Schematic representations of the two models of the artificial SMMF. (a) shows the cosine model with a single source of constant polarity transiting the visible disk. (b) shown the sign change model whereby there are 2 regions of opposite polarity transiting the disk, and their contribution to the SMMF changes during the transit.

There were several variables that allowed us to change the physics of the simulations. These were:

- $N$ : Number of sources
- $t_0$ : Start time of source appearance
- $A$ : Amplitude of sources
- $\lambda$ : Latitude of sources
- $\tau$ : Decay time of sources
- $\phi$ : Additional phase of sources

Using these variables, the mathematical forms of the two model types for a single source are expressed by equation (B.1) and equation (B.2). In these equations  $t' = t - t_0$ ,  $\Pi_{P/2}(t)$  is a window function to define the transit period of the simulated source on the visible side of the solar disk,  $\text{III}_P$  is a Dirac comb of repetition period,  $P$ , and  $\Pi_T(t)$  is a window function defining the total observation period. The time series of a single modelled source for each model is shown in Figure B.1.

$$B_{\text{cosine}}(t) = \left[ A e^{-t'/\tau} \left( \cos\left(\frac{2\pi}{P}t' + \phi\right) \Pi_{P/2}(t) \right) * \text{III}_P \right] \times \Pi_T(t) \quad (\text{B.1})$$

$$B_{\text{sign-change}}(t) = \left[ A e^{-t'/\tau} \left( \cos\left(\frac{2\pi}{P}t' + \phi\right) \sin\left(\frac{2\pi}{P}t' + \phi\right) \Pi_{P/2}(t) \right) * \text{III}_P \right] * \Pi_T(t) \quad (\text{B.2})$$

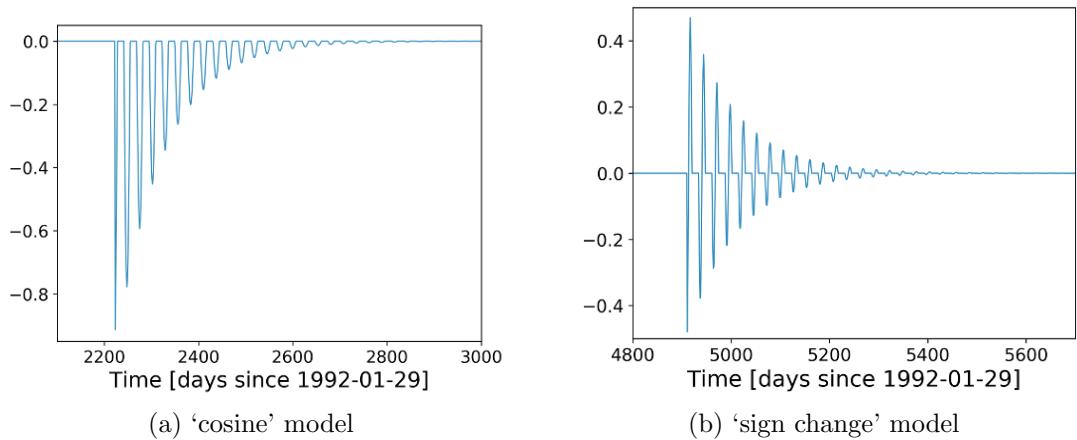


Figure B.2: A single realisation of (a) the cosine model, and (b) the sign change model.

## B.2 Configuration of the Simulations

A flowchart describing the steps in the simulations is shown in Figure B.3. The simulations require the user to select the number of sources to be modelled. Using this information, the user selects whether to draw  $N$  seed/start times ( $t_0$ ) from either a Kernel Density Estimate (KDE) of the Sun Spot Number (SSN), or from a uniform distribution between the start and end times. The former will give an output which is more physical, but the latter is useful for testing.

From the generated seed times, latitudes are computed using a model for the migration of spots during the Solar Cycle (Li et al., 2001b), and the differential rotation frequency at that latitude is taken from a model of the solar differential rotation (Snodgrass, 1983).

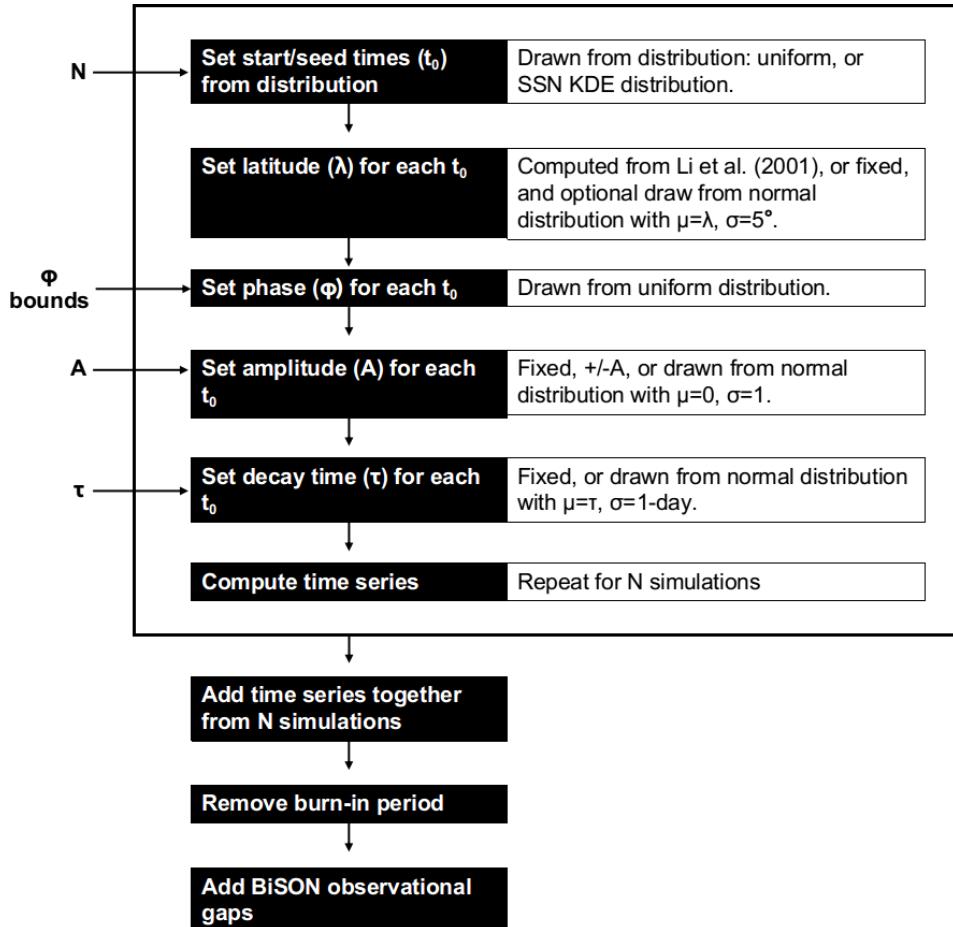


Figure B.3: Flowchart showing the step-by-step processes in the generation of the artificial SMMF time series data.

Values are then assigned for the phase, amplitude, and decay time of each source (the user can set these parameters, or it is also possible to randomly draw them from a distribution). Each individual source is simulated according to equation (B.1) or equation (B.2). Following the simulation of  $N$  sources, the full time-series is computed by adding all  $N$  source contributions together. Finally, a burn-in period is removed, which allows the artificial SMMF to settle prior to the start of ‘observations’, and then we can inject gaps into the artificial time series which are concurrent with the gaps in the Birmingham Solar Oscillations Network (BiSON) observations.

The aim of creating artificial data was to produce a representative power spectrum to that of the BiSON SMMF. We assumed that the sources of the SMMF are Active Region (AR)s and Magnetic Flux Concentration (MFC)s, therefore we aimed

to produce a time series that physically represented these sources, i.e. comparable to the sunspot number. To do this we produced an average number of sources on the visible disk during close to the sunspot number. At solar maximum during cycle 23, the number of daily spots on the disk is around  $150 - 200$ .

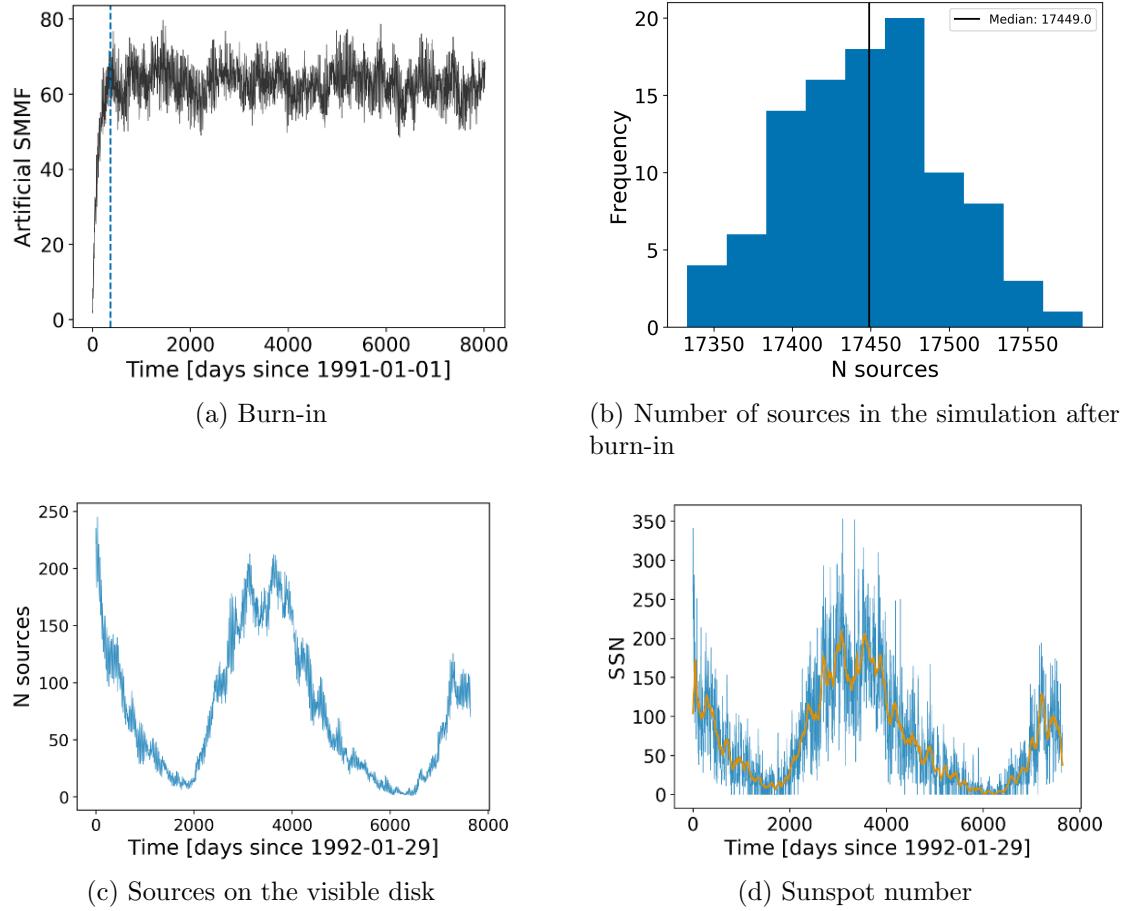


Figure B.4: (a) the burn-in period required to ensure that the interference of sources stabilises. The dashed-vertical line marks one year, showing the sources have stabilised. (b) histogram of the number of sources which remain in the simulation (out of the 20000 input), after the burn-in is removed. (c) the number of simulated sources on the visible disk over the BiSON observational epoch. (d) the daily-averaged sunspot number (blue, and the monthly-averaged sunspot number (orange)).

We included  $\sim 1$ -year for burn-in, such that the interferences between the sources stabilises (see Figure B.4a). After a year the number of sources on the visible disk have steadied, providing a stable SMMF, hence a burn-in period from 01/01/1991 – 29/01/1992 is sufficient. With this burn-in period, we chose to select  $N = 20000$  and  $\tau = 100$  days, which produces a median total of  $\sim 17450$  sources over the BiSON

observational epoch of 7633 days, with an average of around 150 – 200 sources per day on the visible disk during solar cycle 23 maximum, as shown in Figure B.4.

### B.3 Outputs

The simulation produces a time series output from the combination of all the sources with and without the inclusion of BiSON observational gaps, along with the Fourier transforms of both.

For quality assurance, a butterfly diagram is also generated per simulation to understand whether the simulation is representative of the Sun’s ARs. A butterfly diagram, from a realisation which allows for a more realistic, stochastic simulation whereby the values for parameters are drawn from a distribution with scatter, is shown in Figure B.5. This butterfly diagram shows a strong resemblance to a true, observational butterfly diagram of the Sun. This provides us with confidence that the simulations are well-representing ARs on the solar disk.

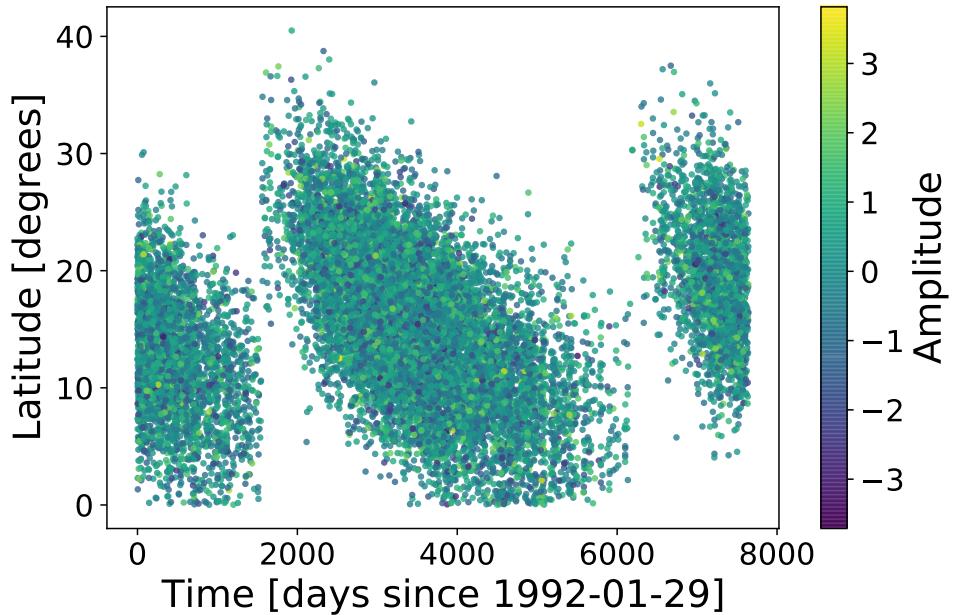


Figure B.5: Artificial butterfly diagram generated from the simulations, allowing for the parameters to be drawn from distributions, to add stochasticity into the simulation.

The resultant power spectra for the two different models ('cosine' and 'sign change') show different features, which can be seen in the limit spectra shown in

Figure B.6. These limit spectra were made by combining 100 realisations of the power spectrum. Each individual power spectrum was made using only a single source starting at  $t_0 = 0$ , and allowing only the phase to vary between the different realisations.

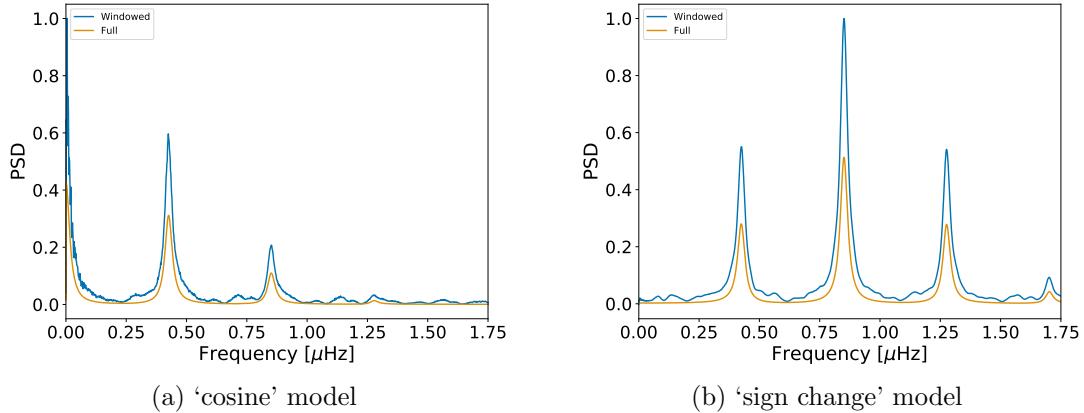


Figure B.6: Limit spectrum from 100 realisations of the cosine model (a) and the sign change model (b) using a single source in each model.

It is clear that the ‘cosine’ model produces a strong peak at the rotational period in the simulation, and a harmonic peak at twice that frequency. This model also produces a significant amount of power at low frequency due to the non-zero mean of the simulated time series.

On the other hand, the ‘sign change’ model has near-zero low frequency power due to the  $\sim$ zero mean of the time series. The ‘sign change’ model also produces a strong peak at half the period of rotation in the simulation. There is a smaller, yet strong peak at the rotational period and at a harmonic of the rotation of 3 times the period. In addition to the limit spectra, we show in Figure B.7 the time series and power spectra for individual realisations of the simulations.

These realisations show the simulated SMMF on a daily cadence. The plots show the stark difference in the time series produced by each model. There are many deviations from a zero mean in the ‘cosine’ model case, which clearly demonstrates the origin of the low frequency power. In the ‘sign change’ model realisation, the data shows a compelling near-zero mean, hence the origin of the near-zero low

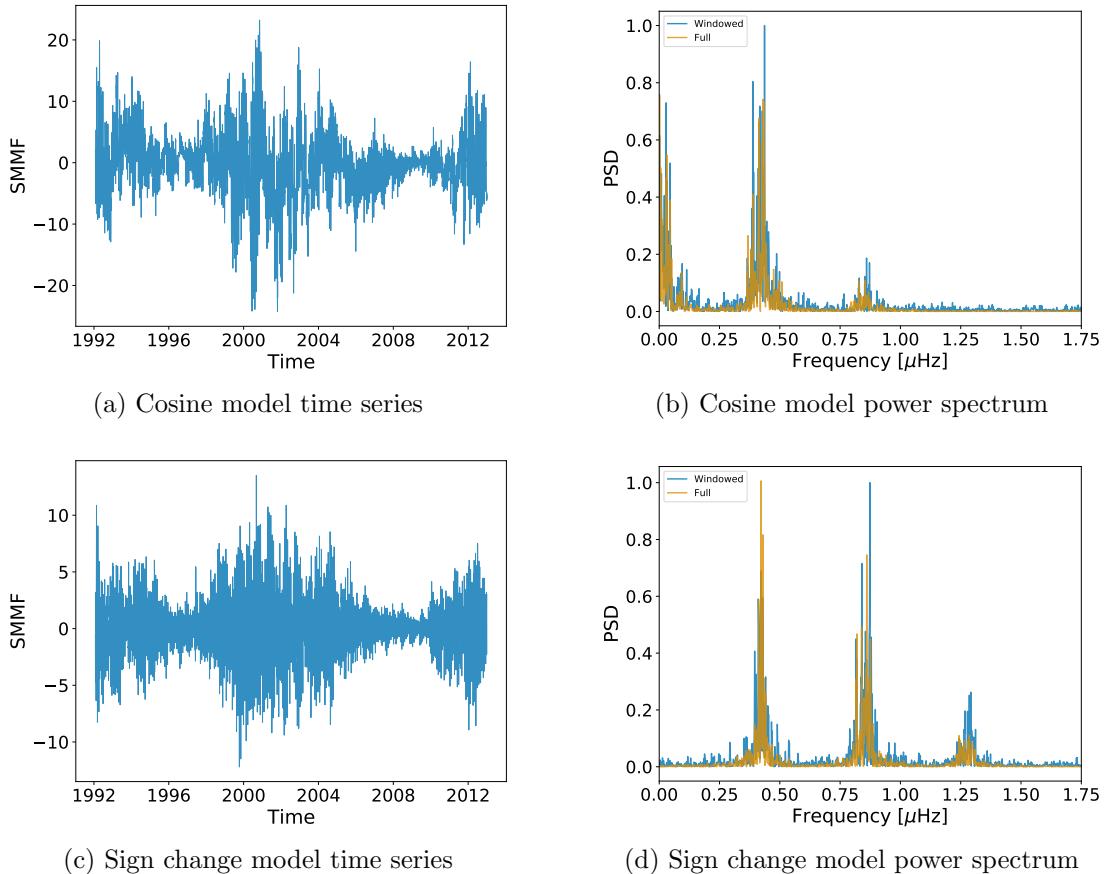


Figure B.7: Time series and power spectra for realisations of the simulations using the cosine and sign change models.

frequency power in the spectrum.

Both realisations of the simulations result in a time series which does resemble some of the features of the BiSON SMMF observations, and hence there is a justification for using both in combination to simulate a representative model of the SMMF. This is supported also by the power spectra for each realisation, and the limit spectra. The power spectrum of the BiSON SMMF shows a strong harmonic peak at three times the period, which is more prominent in the ‘sign change’ model than the ‘cosine’. The ‘sign change’ model power spectrum also shows a similar splitting around the half-period frequency, resembling that of the BiSON spectrum in Figure 4.4b.

# Bibliography

- Aab A., et al., 2017, Science, 357, 1266
- Aslam O. P. M., Badruddin 2012, Solar Physics, 279, 269
- Aslam O. P. M., Badruddin 2015, Sol Phys, 290, 2333
- Barnes a. W. T., et al., 2020, ApJ, 890, 68
- Bartels R. T., 2012, Technical report, The HiSPARC Experiment: An Analysis of the MPV and the Number of Events per Unit Time. University College Utrecht
- Basu S., Chaplin W. J., 2017, Asteroseismic Data Analysis: Foundations and Techniques
- Beck J. G., 2000, Solar Physics, 191, 47
- Belov A., 2000, Space Science Reviews, 93, 79
- Berkova M. D., Belov A. V., Eroshenko E. A., Yanke V. G., 2011, Bull. Russ. Acad. Sci. Phys., 75, 820
- Bose S., Nagaraju K., 2018, The Astrophysical Journal, 862, 35
- Brookes J. R., Isaak G. R., Raay H. B. v. d., 1976, Nature, 259, 92
- Brookes J. R., Isaak G. R., van der Raay H. B., 1978, Mon Not R Astron Soc, 185, 1
- Broomhall A.-M., 2017, Sol Phys, 292, 67
- Brown T. M., Christensen-Dalsgaard J., Dziembowski W. A., Goode P., Gough D. O., Morrow C. A., 1989, The Astrophysical Journal, 343, 526
- Chaplin W. J., Appourchaux T., 1999, Mon Not R Astron Soc, 309, 761
- Chaplin W. J., et al., 1996, Solar Physics, 168, 1
- Chaplin W. J., Dumbill A. M., Elsworth Y., Isaak G. R., McLeod C. P., Miller B. A., New R., Pintr B., 2003, Mon Not R Astron Soc, 343, 813
- Chaplin W. J., Elsworth Y., Isaak G. R., Miller B. A., New R., Pintr B., 2005, Mon Not R Astron Soc, 359, 607
- Chaplin W. J., Elsworth Y., New R., Toutain T., 2008, Monthly Notices of the Royal Astronomical Society, 384, 1668

- Christensen-Dalsgaard J., et al., 1996, Science, 272, 1286
- Davies G. R., Broomhall A. M., Chaplin W. J., Elsworth Y., Hale S. J., 2014a, Monthly Notices of the Royal Astronomical Society, 439, 2025
- Davies G. R., Chaplin W. J., Elsworth Y., Hale S. J., 2014b, Mon Not R Astron Soc, 441, 3009
- Desorgher L., Flckiger E. O., Gurtner M., 2006. p. 2361, <http://adsabs.harvard.edu/abs/2006cosp...36.2361D>
- Dumbill A. M., 1999, PhD thesis, School of Physics and Space Research, University of Birmingham
- Dunai T. J., 2010, in , Cosmogenic Nuclides: Principles, Concepts and Applications in the Earth Surface Sciences. Cambridge University Press, Cambridge, pp 1–24, doi:10.1017/CBO9780511804519.003, /core/books/cosmogenic-nuclides/cosmogenic-nuclides/F85F29AC333CD6F7C9CE2B9155B43B13
- Duvall Jr. T. L., Jefferies S. M., Harvey J. W., Osaki Y., Pomerantz M. A., 1993, The Astrophysical Journal, 410, 829
- ETEnterprises 2020, Technical Report 9125B, Data Sheet: 9125B Series, <http://et-enterprises.com/products/photomultipliers/product/p9125b-series>. ET Enterprises, <http://et-enterprises.com/products/photomultipliers/product/p9125b-series>
- Elsworth Y., Howe R., Isaak G. R., McLeod C. P., Miller B. A., van der Raay H. B., Wheeler S. J., New R., 1995a. p. 392, <http://adsabs.harvard.edu/abs/1995ASPC...76..392E>
- Elsworth Y., Howe R., Isaak G. R., McLeod C. P., Miller B. A., New R., Wheeler S. J., 1995b, Astronomy and Astrophysics Supplement Series, 113, 379
- Fan K.-z., Velthuis J., 2018
- Fesefeldt H., 1985, Technical Report PTHA-85/02, GHEISHA. The Simulation of Hadronic Showers: Physics and Applications. RWTH Aachen
- Fokkema D. B. R. A., 2012, PhD thesis, University of Twente
- Fokkema D. B. R. A., 2019, HiSPARC station software en hardware documentatie HiSPARC software 6.11 documentation, <https://docs.hisparc.nl/station-software/doc/index.html>
- Forbush S. E., 1958, Journal of Geophysical Research, 63, 651
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publications of the Astronomical Society of the Pacific, 125, 306
- Garca R. A., et al., 1999, Astronomy and Astrophysics, 346, 626

Giacalone J., 2010, in Schrijver C. J., Siscoe G. L., eds, , Heliophysics: Space Storms and Radiation: Causes and Effects. Cambridge University Press, Cambridge, pp 233–262, doi:10.1017/CBO9781139194532.010, /core/books/heliophysics-space-storms-and-radiation-causes-and-effects/energetic-particle-transport/FC3869AB4D484479501B015DCB07A976

Hale S. J., Howe R., Chaplin W. J., Davies G. R., Elsworth Y., 2016, Solar Physics, 291, 1

Hanasoge S., Mandal K., 2019, The Astrophysical Journal Letters, 871, L32

Handberg R., Campante T. L., 2011, Astronomy & Astrophysics, 527, A56

Harvey K. L., Zwaan C., 1993, Solar Physics, 148, 85

Hathaway D. H., Choudhary D. P., 2008, Sol Phys, 250, 269

Hathaway D. H., Upton L. A., 2020, arXiv e-prints, 2006, arXiv:2006.06084

Heck D., Pierog T., 2017, Technical Report 7.6400, Extensive Air Shower Simulation with CORSIKA: A Users Guide

Herbst K., Kopp A., Heber B., 2013, Annales Geophysicae, 31, 1637

Howard R. F., 2001, in , The Encyclopedia of Astronomy and Astrophysics. IOP Publishing Ltd, doi:10.1888/0333750888/2297, <http://eaa.crcpress.com/0333750888/2297>

Howe R., 2009, Living Reviews in Solar Physics, 6

Howe R., Davies G. R., Chaplin W. J., Elsworth Y. P., Hale S. J., 2015, Mon Not R Astron Soc, 454, 4120

Howe R., Hill F., Komm R., Chaplin W. J., Elsworth Y., Davies G. R., Schou J., Thompson M. J., 2018, ApJL, 862, L5

Howe R., et al., 2020, Mon Not R Astron Soc Lett, 493, L49

Inceoglu F., Knudsen M. F., Karoff C., Olsen J., 2014, Sol Phys, 289, 1387

Janardhan P., Fujiki K., Ingale M., Bisoi S. K., Rout D., 2018, Astronomy and Astrophysics, 618, A148

Kane R. P., 2014, Solar Physics, 289, 2727

Kotov V. A., 2008, Astron. Rep., 52, 419

Kotov V. A., 2012, Bull.Crim. Astrophys. Observ., 108, 20

Kutsenko A. S., Abramenko V. I., Yurchyshyn V. B., 2017, Sol Phys, 292, 121

Lanza A. F., Gizon L., Zaqrashvili T. V., Liang Z.-C., Rodenbeck K., 2019, A&A, 623, A50

Li K. J., Yun H. S., Gu X. M., 2001a, AJ, 122, 2115

Li K. J., Yun H. S., Gu X. M., 2001b, AJ, 122, 2115

- Liang Z.-C., Gizon L., Birch A. C., Duvall T. L., 2019, A&A, 626, A3
- Lingri D., Mavromichalaki H., Belov A., Eroshenko E., Yanke V., Abunin A., Abunina M., 2016, arXiv:1612.08900 [astro-ph]
- Löptien B., Gizon L., Birch A. C., Schou J., Proxauf B., Duvall T. L., Bogart R. S., Christensen U. R., 2018, Nature Astronomy, 2, 568
- Lund M. N., Chaplin W. J., Hale S. J., Davies G. R., Elsworth Y. P., Howe R., 2017, Mon Not R Astron Soc, 472, 3256
- Maunder E. W., 1904, Mon Not R Astron Soc, 64, 747
- Mavromichalaki H., Paouris E., Karalidi T., 2007, Sol Phys, 245, 369
- McIntosh S. W., et al., 2014, ApJ, 792, 12
- Mendon{\c{c}}a R. R. S. d., et al., 2016, ApJ, 830, 88
- Mishra V. K., Mishra A. P., 2016, Indian J Phys, 90, 1333
- NMDB 2018, NMDB Event Search Tool (NEST), <http://www.nmdb.eu/nest/>
- Nelson W. R., Hirayama H., Rogers D. W. O., 1985, Technical Report SLAC-265, EGS4 code system, <https://www.osti.gov/biblio/6137659-egs4-code-system>. Stanford Linear Accelerator Center, Menlo Park, CA (USA), <https://www.osti.gov/biblio/6137659-egs4-code-system>
- Ostapchenko S., 2006, Nuclear Physics B - Proceedings Supplements, 151, 143
- Owens M. J., Forsyth R. J., 2013, Living Rev. Sol. Phys., 10, 5
- Pacini A. A., Usoskin I. G., 2015, Sol Phys, 290, 943
- Paouris E., Mavromichalaki H., Belov A., Eroshenko E., Gushchina R., 2015, J. Phys.: Conf. Ser., 632, 012074
- Parker E. N., 1965, Planetary and Space Science, 13, 9
- Paschalis P., Mavromichalaki H., Yanke V., Belov A., Eroshenko E., Gerontidou M., Koutroumpi I., 2013, New Astronomy, 19, 10
- Pesnell W. D., Schatten K. H., 2018, Sol Phys, 293, 112
- Plachinda S., Pankov N., Baklanova D., 2011, Astronomische Nachrichten, 332, 918
- Ross E., Chaplin W. J., 2019, Sol Phys, 294, 8
- Rossby C.-G., Collaborators 1939, J Mar Res, 2, 38
- Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, PeerJ Comput. Sci., 2, e55
- Savi M., Maleti D., Jokovi D., Veselinovi N., Banjanac R., Udovii V., Dragi A., 2015, J. Phys.: Conf. Ser., 632, 012059
- Scherrer P. H., Wilcox J. M., Howard R., 1972, Sol Phys, 22, 418

- Scherrer P. H., Wilcox J. M., Kotov V., Severnyj A. B., Severny A. B., Howard R., 1977a, Solar Physics, 52, 3
- Scherrer P. H., Wilcox J. M., Svalgaard L., Duvall Jr. T. L., Dittmer P. H., Gustafson E. K., 1977b, Solar Physics, 54, 353
- Schrijver C. J., Harvey K. L., 1994, Solar Physics, 150, 1
- Severny A. B., 1971, Quarterly Journal of the Royal Astronomical Society, 12, 363
- Singh M., Singh Y. P., Badruddin 2008, Journal of Atmospheric and Solar-Terrestrial Physics, 70, 169
- Snodgrass H. B., 1983, The Astrophysical Journal, 270, 288
- Strohmayer T., Mahmoodifar S., 2014, A Neutron Star Oscillation Mode During a Superburst, <https://slideplayer.com/slide/10313230/>
- Svalgaard L., Wilcox J. M., Scherrer P. H., Howard R., 1975, Sol Phys, 45, 83
- Thomas S. R., Owens M. J., Lockwood M., 2014a, Sol Phys, 289, 407
- Thomas S. R., Owens M. J., Lockwood M., Scott C. J., 2014b, Solar Physics; Dordrecht, 289, 2653
- Thomas A. E. L., et al., 2019, Mon Not R Astron Soc, 485, 3857
- Tomassetti N., Orcinha M., Baro F., Bertucci B., 2017, ApJL, 849, L32
- Upton L. A., Hathaway D. H., 2018, Geophysical Research Letters
- Usoskin I. G., Kananen H., Mursula K., Tanskanen P., Kovaltsov G. A., 1998, Journal of Geophysical Research: Space Physics, 103, 9567
- Van Allen J. A., 2000, Geophysical Research Letters, 27, 2453
- Wu C. J., Usoskin I. G., Krivova N., Kovaltsov G. A., Baroni M., Bard E., Solanki S. K., 2018, A&A, 615, A93
- Xiang N. B., Qu Z. N., 2016, AJ, 151, 76
- Xie J. L., Shi X. J., Xu J. C., 2017, The Astronomical Journal, 153, 171
- Zwaan C., 1981, NASA Special Publication, 450
- van Dam K., van Eijk B., Fokkema D. B. R. A., van Holten J. W., de Laat A. P. L. S., Schultheiss N. G., Steijger J. J. M., Verkooijen J. C., 2020, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 959, 163577
- van Driel-Gesztelyi L., Green L. M., 2015, Living Rev. Sol. Phys., 12, 1