**School of Science**

# Solar Variability and Automated Analysis for Space Weather

**Owen David Giersch**

**This thesis is presented for the Degree of**
**Doctor of Philosophy**
**of**
**Curtin University**

**August 2016**

## Acknowledgements

I would like to to give special thanks to all of the staff at Global Oscillation Network Group (GONG) without which this work would not be possible. Particular thanks go to Frank Hill, Jack Harvey and Tim Purdy who answered all of my questions about the GONG systems.

I would also like to thank all of my friends and colleagues at Learmonth Solar Observatory, with whom I worked collaboratively with over the last 10 years.

I would like to thank my supervisors, Dr John Kennewell and Prof Mervyn Lynch, who gave unfailing support to me.

## Abstract

In the modern era, particularly from 1980 onward, researchers have questioned the reliability of various long term solar index data sets such as sunspot number, sunspot area, radio flux and H-alpha flares. Have these measures changed over time, particularly when observations have ceased at one site or organisation and been taken up by another? Researchers such as Hathaway suggest that they have, and others cite more recent anecdotal evidence that sunspot area measurements, in particular, have changed since the United States Air Force took over this task in 1980. This thesis explores whether various solar measurements have changed in the period 1945 to 2015. It is found that a particular solar index may vary significantly, depending on the site of observation, era of observation, techniques or equipment used. These variations may be, at least in part, due to intrinsic variation in the Sun.

In order to minimise future problems, algorithms have been developed to measure sunspot area and sunspot number automatically with no human input. These programs used Global Oscillation Network Group (GONG) intensity and magnetogram images to extract the sunspot number and sunspot area from each of the 6 GONG sites every day from 2002 to 2014. These were subsequently averaged to produce a GONG daily Sunspot Number and Sunspot Area. These were then averaged over a year and compared with the International Sunspot Number (ISN) and Space Weather Prediction Center (SWPC) sunspot area reports. The results showed that GONG produced very good correlation with the International Sunspot Numbers, with greater variations between cycles when compared with the American Association of Variable Star Observers' (AAVSO) sunspot number. Using these analysis techniques it can be shown that a significant proportion of the variation in the sunspot number to sunspot area ratio between sunspot cycles is a feature of the Sun, and only partly contributed by current observing techniques. Therefore, this ratio should not be used to determine long term variations in measurement of sunspot number or sunspot area.

A careful re-measurement of a selection of sunspot drawings was made to test the claims that the USAF has been significantly underestimating sunspot areas. It was found that total disk sunspot areas were being under reported by up to 12% due to current measurement techniques. This error can easily be removed by more careful measurement of the area limb correction factors than are currently used by the USAF.

In addition, to improve consistency of sunspot classifications, algorithms were developed that produced the McIntosh classification for each sunspot group and also a Mt Wilson Magnetic Classification. This required determining specific physical measures for the penumbral shape, compactness of the region, and inter-mixing of magnetic field parameters. Good results were produced given the resolution limits on GONG.

For a long time it has been known that reporting of Hα flares varied from site to site. In 2010 GONG deployed a H-alpha system at each of its sites. Automated flare analysis software was developed and trialled on this data and inter-site comparisons were made. Whilst both the flare areas and flare brightness varied from site-to-site, they varied in a more consistent way than does the USAF Solar Optical Observing Network (SOON) observations. However, there are still significant differences between individual GONG sites. An analysis was also undertaken comparing Learmonth SOON flares with Learmonth GONG flare analyses that showed a significant difference between the two. This implies that variations between sites are not due to atmospheric effects that may influence seeing conditions, but rather to variations in equipment.

A set of parameters from the sunspot and magnetic field analyses were examined to establish if these observations could be used as predictors for M class or higher X-ray flares. Magnetic field strength, inversion line complexity, magnetic field gradient, sunspot area growth rates and sunspot orientation were all examined to generate a set of predictors. It was found that the magnetic inversion line analysis provided the most consistent results when examining flare rates from individual regions.

Finally, GONG region analysis was used to try determine if morphological features were able to predict the onset of solar particle events (SPEs). Unfortunately, no distinctive features could be found that were able to be associated with SPEs for events greater than 10 pfu at the 10 MeV energy level. Similarly, no features could be found that predicted Ground Level Enhancements. This is probably due to the primary acceleration method being Fermi acceleration some distance away from the sun, making morphological predictions more difficult.

# Table of Contents

# List of Figures

# List of Tables

## List of Abbreviations

| | |
|---|---|
| AAVSO | American Association of Variable Star Observers |
| AFWA | U.S. Air Force Weather Agency (now 557$^{th}$ Weather Wing) |
| ASSC | Australian Space Science Conference |
| BB | Big Bear |
| CME | Coronal Mass Ejection |
| CMP | Central Meridian Passage |
| CT | Cerra Tololo |
| FITS | Flexible Image Transport System |
| GLE | Ground Level Enhancement |
| GNSS | Global Navigation Satellite Systems |
| GONG | Global Oscillation Network Group |
| GPS | Global Positioning System |
| GSN | Group Sunspot Number |
| ISN | International Sunspot Number |
| LAT | Latitude |
| LEAR | Learmonth |
| LE | Learmonth |
| LON | Longitude |
| LSO | Learmonth Solar Observatory |
| ML | Mauna Loa |
| NGDC | U.S. National Geophysical Data Center |
| PALE | Palehua |
| PDF | Probability Density Function |
| PFU | Particle Flux Unit |
| RFI | Radio Frequency Interference |
| RGO | Royal Greenwich Observatory (U.K.) |
| RIMS | Radio Interference Measuring Set |
| RSTN | U.S. Air Force Radio Solar Telescope Network |
| SGMR | Sagamore Hill |
| SIDC | Solar Influences Data Center (Belgium) |
| SOON | U.S. Air Force Solar Observing Optical Network |
| SSN | Smoothed Sunspot Number |
| SPAN | Solar Particle Alert Network |

| | |
|---|---|
| SPE | Solar Particle Event |
| SRS | Solar Region Summary |
| SRS | Solar Radio Spectrograph |
| SFU | Solar Flux Unit = $10^{-22}$ W/m$^2$/Hz = -220 dBW/m$^2$/Hz |
| SVTO | San Vito |
| SWPC | U.S. Space Weather Prediction Center |
| TD | Tiede |
| UD | Udaipur |
| USAF | United States Air Force |
| WBMOD | WideBand MODel (of ionospheric scintillation) |

# 1. Introduction

## 1.1 Background

From a scientific perspective, the Sun has been studied for over 400 years. Over this time several significant and historical solar indices have been developed as long term measures of solar activity. These include sunspot number, sunspot area, sunspot classification, magnetic field classification, 10.7 cm flux and flare rates across a range of wavelengths from radio to x-ray.

In the last two decades various researchers have questioned the historical reliability of some of these parameters, in particular sunspot number and sunspot area. Discrepancies in these indices seemed to appear after sunspot analysis moved from Zurich to Brussels and the sunspot area analysis was taken over by the United States Air Force (USAF) after Royal Greenwich Observatory sunspot measurements ceased. Of course there were other groups monitoring these parameters (for example the American Association of Variable Star Observers (AAVSO) for sunspot number), but Zurich and Greenwich were the longest running.

On shorter time scales, from the 1960s it was noted that differing sites were producing different areas and brightness for the same Hα flares. Certainly this sort of effect had been noticed early on for sunspots, with different sites producing different sunspot numbers. In the case of sunspots, k-factors were applied to each observer and observatory to try to correct this variation, which was largely due to the large range of telescopes used.

To add complexity to this problem, at different times different measurement techniques were used across time periods as technology has advanced. The one exception to this is probably radio observations. The equipment from the 1960s to present has largely remained unchanged, only with advances in the computer interface used to collect the data from the receivers and how the data is transmitted to various interested parties.

Since sunspot classification is often used by space forecast centres to predict potentially geo-effective solar events, reliable and consistent classifications are needed to be produced on a daily basis. Over the last century there have been three major classification schemes introduced based on white light observations: Cortie's scheme (Cortie, 1901), the Zurich scheme (Carrasco et al., 2015) and the McIntosh scheme (McIntosh 1990).

There is also the Mount Wilson magnetic classification scheme (Hale et al., 1919). The Mount Wilson scheme has slowly been added to and updated. In all cases, different telescopes and observers will produce different classifications for the same sunspot groups. Reconciling this data into a consistent and coherent set is a challenge for both researchers and forecast centres.

In July 2015, the Solar Influences Data Center (SIDC) in Belgium revised the entire sunspot data set. Part of the reason for this revision was to remove the scaling factor of 0.6 from this index (see section 2.2). The other reason was to remove perceived inaccuracies in the data. Due to the timing of this revision and publication of this thesis, the old data set is used throughout this thesis.

### 1.2 Research Questions

The primary questions that this thesis attempts to answer are:

1. Are variations in sunspot number, sunspot area and 10 cm flux across sunspot cycles a result of intrinsic changes in the sun or changes in measurement?

2. How do different USAF Solar Observing Optical Network (SOON) sites and Radio Solar Telescope Network (RSTN) sites compare with each other for measuring Hα flares and radio bursts respectively? Is the recently installed Global Oscillation Network Group (GONG) system a more consistent system for flare observations?

3. Can software be produced that accurately and consistently measures various solar parameters such as sunspot area, sunspot number, sunspot classifications, and magnetic field parameters? If so, how do these automatic techniques compare with manually reduced parameters?

4. Can parameters from automatic optical observations be derived that predict X-ray flares or solar particle events (SPEs)?

### 1.3 Observing Sites

Much of the data used in this thesis is derived from the USAF SOON and RSTN sites. The automated techniques were used on GONG data. Figure 1.1 shows the site locations

for the USAF and table 1.1 their latitude and longitude. Figure 1.2 shows the sites for GONG and table 1.2 their latitude and longitude.

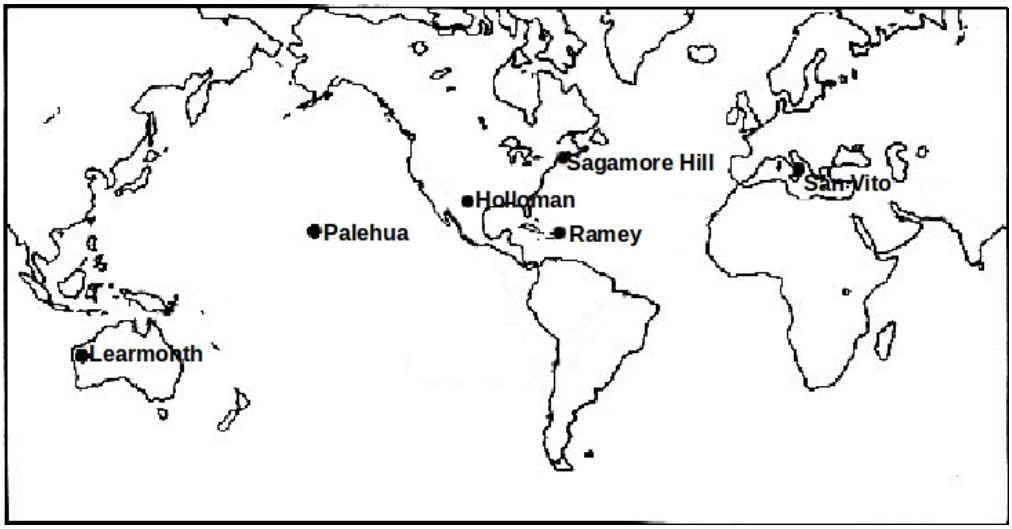



**Figure 1.1**: USAF SOON and RSTN Sites. Learmonth and San Vito have both optical (SOON) and radio (RSTN). Sagamore Hill only has Radio and Holloman only has optical. Palehua had both optical and radio until the SOON portion was closed in 1996. Ramey was optical only and closed in 2002.

| Site Name | Abbreviation | Latitude | Longitude |
|---|---|---|---|
| Learmonth | LEAR | S 22º 13’ | E 114º 6’ |
| San Vito | SVTO | N 40º 24’ | E 17º 26’ |
| Ramey | RAMY | N 18º 31’ | W 67º 6’ |
| Sagamore Hill | SGMR | N 42º 38’ | W 70º 49’ |
| Holloman | HOLL | N 32º 50’ | W 106º 6’ |
| Palehua | PALE | N 21º 23’ | W 158º 7’ |

**Table 1.1**: USAF SOON and RSTN Observatory Locations

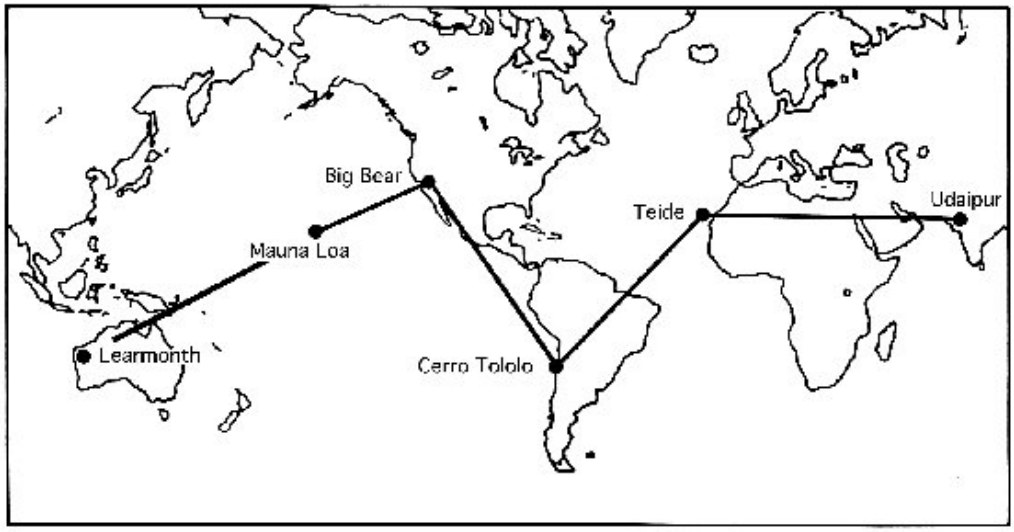



**Figure 1.2**: GONG Site Map

| Site Name | Abbreviation | Latitude | Longitude |
|---|---|---|---|
| Learmonth | LE | S 22º 13’ | E 114º 6’ |
| Udaipur | UD | N 24º 37’ | E 73º 40’ |
| Teide | TD | N 28º 18’ | W 16º 31’ |
| Cerra Tololo | CT | S 30º 10’ | W 70º 48’ |
| Big Bear | BB | N 34º 16’ | W 116º 55’ |
| Mauna Loa | ML | N 19º 32’ | W 155º 35’ |

**Table 1.2**: GONG Observatory Locations

The site locations are important, as each will have different seeing conditions that may affect observations. In addition sunspot features can change and evolve over the course of a few hours, meaning that sites at opposite sides of the globe may report the same region's parameters differently.

### 1.4 Methodology

Initially a review of each of the parameters is undertaken. The specific parameters studied are: sunspot number, sunspot area, sunspot classification, magnetic classification, 10.7 cm radio flux, solar radio bursts, Hα flares, x-ray flares and solar particle events.

These parameters are compared across sunspot cycles, to determine how they vary from cycle to cycle. In addition, measurements of these parameters are compared between different observing sites and groups.

Automated techniques are developed to measure sunspot and magnetic field parameters of the Sun using GONG intensity and magnetogram images. Sunspot numbers and sunspot areas generated by this process are compared with manually reduced data from other agencies, in an attempt to determine accuracy of currently used techniques, and the reliability of the automated process.

Automated techniques are also developed to measure Hα flares from GONG Hα images. As with sunspots, the GONG flares were compared with other reports to determine their reliability.

The sunspot parameters are also compared with x-ray flare rates. This is to determine if any particular parameter or group of parameters is a good predictor for flares. A similar approach is used for solar particle events.

Most of the analysis is of a statistical nature. That is determining if a particular parameter is significantly different between sunspot cycles or sites. For comparisons between two data sets, typically either linear regression analysis is used or $\chi^2$ goodness of fit is used. Goodness of fit is also used for determining how well predictive techniques work.

**1.5 Thesis Layout**

Chapter 2 provides an introduction to sunspots and the various sunspot parameters that are currently routinely measured. It provides a historical perspective and discusses the various observation techniques.

Chapter 3 gives an overview of radio observations of the sun over a wide range of wavelengths. Variations in calibration of the RSTN system is examined. An analysis of burst rates at various frequencies and across sunspot cycles is examined. Burst comparisons between sites are studied in order to determine how consistently the RSTN systems perform. Radio interference to GPS is reviewed as a specific example of how the Sun can affect technological systems.

Chapter 4 considers the relationships between sunspot area, sunspot number and 10.7 cm radio flux. Initially these parameters are examined individually for differences between observing sites and between sunspot cycles. These parameters are compared between each other to determine if the relationships between them change over sunspot cycles.

Chapter 5 discusses both past and current Hα flare observing techniques. Site discrepancies in the SOON systems are examined in some detail.

Chapter 6 examines the pre-processing of GONG images prior to being examined. Specifically the image normalisation, rotation and translation techniques are discussed. Finally, limb darkening correction techniques are examined.

Chapter 7 discusses the automated region analysis program in depth. It examines sunspot identification, region identification and analysis using GONG imagery. The results from this analysis are then compared with existing data sets in order to determine the reliability of the automated processes.

Chapter 8 discusses sunspot parameters associated with x-ray flares. Statistical analysis is used in an attempt to correlate various sunspot and magnetic features to x-ray flare rates. From this, the probability of a flare occurring in an active region can be determined. This analysis is performed for both archived data (where appropriate parameters are available) and for the automated GONG parameters.

Chapter 9 discusses the mechanisms for solar particle acceleration. Sunspot parameters are examined in order to determine if there are any significant sunspot features that can be correlated with solar particle events.

Chapter 10 discusses the automated Hα flare analysis algorithm. The results from this algorithm are compared with SOON flare reports and also between GONG sites in order to determine if GONG with automated processing performs better than the SOON system.

Finally, Chapter 11 sums up the thesis, discussing conclusions in regard to automated analysis. Also discussed is future work and potential for other automated techniques that were not covered in this thesis.

### 1.6 Note on Units

In this thesis, SI units are used on most occasions. There are, however, some notable exceptions. Firstly, the solar physics community generally still uses Gauss as the preferred unit for the magnetic field. The GONG magnetograms were scaled in Gauss and in this thesis the same convention is used. The conversion is given by 1 G = $10^{-4}$ T.

The second non-standard unit is solar area. The area of flares, sunspots and other solar features is measured in millionths of the solar hemisphere. There is no symbol for this unit, so the same notation as Fokaul (2014) of the micro-hemisphere (μh) where 1 μh = $3.04 \times 10^6$ km$^2$ has been adopted.

Thirdly, the Angstrom is still in frequent use for X-ray wavelength and it is used on occasion here where 1 Å = 0.1 nm.

Throughout this thesis two main statistical tests are used. Firstly slopes and gradients of lines are compared with each other, and secondly, $\chi^2$ tests are used to compare discrete data sets or data to a model. These will be discussed briefly in turn.

### 1.7 Linear Regression Tests

Throughout this thesis slopes and gradients of lines are compared with each other.

A good reference for obtaining the line of best fit to a set of data is Kirkup (1994). The regression parameters are found by equations 1.1 to 1.5

$$m = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \quad , \qquad \text{Eqn 1.1}$$

where $m$ is the gradient, $n$ is the number of data points in the plot, $x_i$ is the $i$th x data point and $y_i$ is the $i$th y data point.

$$b=\frac{\sum_{i=1}^{n} x_i^2 y_i-\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2} \quad , \qquad \text{Eqn 1.2}$$

where $b$ is the y-axis intercept.

$$\sigma=\left[\frac{1}{n-2}\sum_{i=1}^{n}\left(y_i-m x_i-b\right)^2\right]^{\frac{1}{2}} \quad , \qquad \text{Eqn 1.3}$$

where $\sigma$ is the uncertainty in the y values.

$$s_m=\left[\frac{\sigma^2 n}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}\right]^{\frac{1}{2}} \quad , \qquad \text{Eqn 1.4}$$

where $s_m$ is the standard deviation in the gradient.

$$s_b=\left[\frac{\sigma^2\left(\sum_{i=1}^{n} x_i\right)^2}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}\right]^{\frac{1}{2}} \quad , \qquad \text{Eqn 1.5}$$

where $s_b$ is the standard deviation in the gradient.

A standard statistical test, similar to testing differences between means, is used to determine differences between means. Throughout this thesis a confidence interval of 2 standard deviations is used. That is, if 2 gradients, $m_1$ and $m_2$, with standard deviations, $s_{m1}$ and $s_{m2}$, are to be considered statistically similar then the following condition must be met:

$$\left|m_2-m_1\right|<2\left[s_{m2}+s_{m1}\right] \quad . \qquad \text{Eqn 1.6}$$

A similar equation to 1.6 is used to determine if the intercepts are similar.

In chapter 8, the $\chi^2$ test is also used, and this will be described in that section.

### 1.8 Heliographic Coordinates

A review of the heliographic coodinate system can be found in Çakmak (2014). An overview of heliographic coordinates is given below. Figure 1.3 shows the coordinate system.

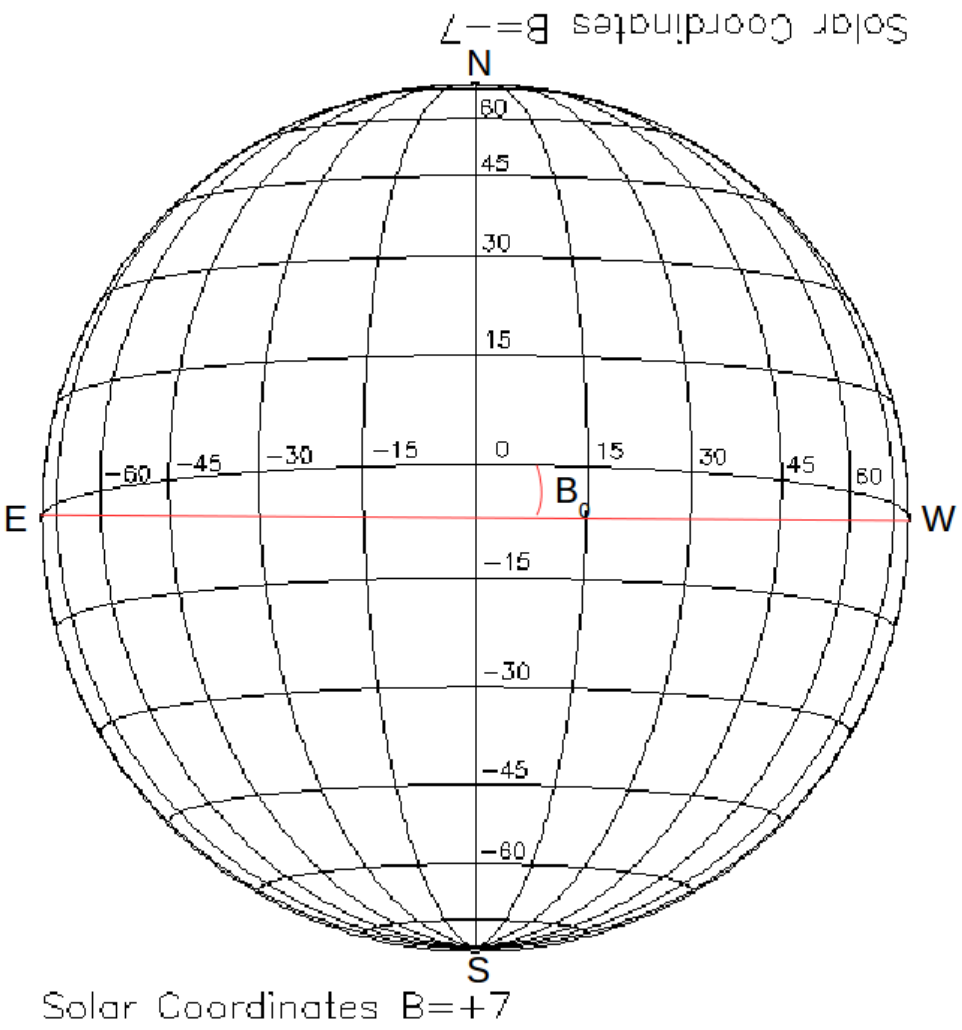



**Figure 1.3**: Heliographic Coordinate System. The red line indicates the centre of the disk. $B_0$ is the B-angle, the tilt towards or away from the Earth.

Heliographic co-ordinates are assigned as follows: negative east, positive west, negative

south and positive north.

The Sun also tilts clockwise or anti-clockwise. This angle is referred to as the P-angle. The P-angle is not used in this thesis as it is removed by rotating the image.

When computing distance between features or extent of features, the B-angle is not incorporated into these calculations. The B-angle is incorporated into the final positions of features for easy comparison between other databases.

# 2. Review of Sunspots

## 2.1 Introduction

Sunspots occur where magnetic fields emerge through the photosphere and magnetically cool the plasma in the region (Zirin, 1988). It should be noted, though, that it is possible to have magnetic field in a region, but no spots.

Sunspots are dynamic and grow and decay over time. Small spots maybe short lived (a few hours) while some sunspot regions may last for months. Sunspots occur in groups. This is due to the nature of the magnetic fields, that there must be complimentary positive and magnetic field in the same region. The more complex the magnetic field configuration, the more complex the sunspot group.

A detailed description of sunspots can be found in Bray and Loughhead (1964). A more recent overview can be found in Solanki (2003). A summary in point form is provided below.

- The number of sunspots visible on the solar disk goes through a cycle of minimum to maximum and back to minimum numbers approximately every 11 years. The length of the cycle, and maximum and minimum sunspot numbers vary from cycle to cycle.
- Sunspots are located at solar latitudes between 40º south and 40º north. At the beginning of a sunspot cycle, emerging from a minimum, the spots are at higher latitudes and appear closer to the equator towards the end of the cycle, entering the next minimum.
- Sunspots typically have a dark core (umbra) and a halo around it (penumbra). Small spots often have either rudimentary or no penumbra, and those with no penumbra are sometimes referred to as pores.
- Sunspots occur in areas of magnetic field. In essence they are where magnetic flux tubes emerge through the photosphere.
- Sunspots often occur in groups with spots on each magnetic polarity. Single spots can occur, but there is always a region of opposite polarity in the vicinity of the single spot. A sunspot group is therefore determined based on proximity of individual spots and their magnetic polarities.
- During the life time of a spot group, individual spots can grow, merge, split apart or shrink.

- A large sunspot can have multiple umbra within it. Thus when counting sunspots usually what is meant is counting umbra. In this thesis when counting spots, it is the individual number of umbra that are used.
- The darkness of spots, relative to the photosphere, is due to them being cooler and therefore radiating less. The umbra of spots can be up to 1900 K cooler than the photosphere of the quiet Sun.
- A sunspot life time can be less than a day to several months and is correlated with its size, that is larger spots survive longer than smaller ones.

The major index for monitoring sunspots is the International Sunspot Number (ISN). There is also the McIntosh sunspot group classification system and the Mount Wilson magnetic classification system. Other significant sunspot parameters that have been measured are sunspot area and the umbra area to total area ratio. Each of these will be described in turn.

**2.2 Sunspot Number**

The most basic sunspot parameter is the ISN, also known as the Wolf Sunspot Number (Izenman et al., 1983), named after the creator of the index. Each observer computes a sunspot index by equation 2.1:

$$R = k(10g + n) \quad , \qquad \text{Eqn 2.1}$$

where $R$ is the observer's sunspot number, $g$ is the number of sunspot groups and $n$ is the number of sunspot umbra visible on the disk. $k$ is a correction factor to adjust for different observers and equipment.

This is then averaged for all of the observers that make up the index to produce the ISN. There are typically between 10 and 30 individual observations per day that go into computing the final sunspot number (Clete et al., 2014).

The ISN, up until recently, was assigned a value of k=0.6. This was because Wolf used small telescopes for his observations (Izenman et al., 1983).

The sunspot number is periodic. This was first identified by Schwabe who estimated with limited data a period of 10 years. It was later refined to an average value of 11.1 years by

Wolf (Izenman et al., 1983). The actual sunspot number cycle length varies from cycle-to-cycle (see chapter 4 for details).

Individual daily sunspot reports, even from the same observers, can vary significantly due to changing seeing conditions. The constant *k* in equation 2.1 is therefore an average for an observer, but may be changed over time if it drifts.

Reliable records of this index date back to the 1700s making it the longest index for a solar feature or activity. However, in the 1940s the American Association of Variable Star Observers (AAVSO) started to report this index as well, which here will be referred to as the AAVSO sunspot number. This value was scaled to bring it into line with the then Zurich Sunspot Number (Shapley, 1949),

Other individual observatories also record sunspot counts, some of which are available on the National Geophysical Data Center (NGDC) website.

The ISN has gone through some changes, the most notable being the change of co-ordinating agency from Zurich to Belgium in 1980. Prior to 1980 the ISN was known as the Zurich Sunspot Number or Wolf Sunspot Number. More details on the various changes to the International Sunspot Number will be given in Chapter 4.

The recommended method of determining this number is for an for an observer to draw the sunspots from a projection of the Sun onto a whitelight board. The analyst then groups and counts the sunspots to obtain the sunspot number for that site. This may then be transmitted to the Solar Influences Data Center (SIDC) in Belgium. The appropriate *k* correction, for that observer and time, is applied and then all of the sites sunspot numbers are averaged to produce the official daily sunspot number.

The Group Sunspot Number (GSN) is a more recent introduction and is simply defined as 12.08 times the number of sunspot groups (Hoyt and Schatten, 1998). The number of individual sunspots is not considered. Formally the GSN is defined as:

$$R_G = \frac{12.08}{N} \sum k_i' g_i \quad , \qquad \text{Eqn 2.2}$$

where $R_G$ is the GSN, $N$ is the number of observers, $k_i'$ is the ith observer's correction factor and $g_i$ is the ith observers number of sunspot groups.

The value of 12.08 in equation 2.2 was chosen to make the mean GSN obtained from Royal Greenwich Observatory (RGO) the same as the Wolf Sunspot Number from 1874 to 1976.

Finally, a note on smoothing. There is a large day-to-day variation in sunspot number values. In order to overcome this a monthly value is often used as shown in equation 2.3.

$$R_m = \frac{\sum R_i}{N_m} \quad , \qquad \text{Eqn 2.3}$$

where $R_m$ is the average monthly sunspot number for month $m$, $R_i$ is the daily sunspot number for a day in a given month and $N_m$ is the number of days in that month.

Even this index can have rapid oscillation and so a further smoothing scheme is used (equation 2.4). This is called the smoothed monthly sunspot number or sometimes just the smoothed sunspot number (SSN).

$$R_{ms} = \frac{R_{m-6} + R_{m+6} + 2\sum_{j=m-5}^{m+5} R_j}{24} \quad , \qquad \text{Eqn 2.4}$$

where $R_{ms}$ is the smoothed monthly sunspot number for month $m$, and $R_j$ is the monthly sunspot number for month $j$.

An alternate index, less often used, is the annual sunspot number. This number is simply the average of the daily sunspot numbers over a calendar year. This index is not discussed any further in this thesis.

**2.3 Sunspot Area**

Daily observations of sunspot area began at Royal Greenwich Observatory in 1875 (Sarychev and Roshchina, 2006). Greenwich sunspot area measurements ceased in 1976. It was around this time that the USAF SOON (United States Air Force Solar Observing

Optical Network) sites started operating, although there is a period when neither Greenwich nor SOON were observing. This has led to some discussion as to whether the current sunspot areas are being under reported (for example Foukal, 2014).

Sunspot areas are currently measured from the sunspot drawings performed at the 3 SOON observatories. Three transparent overlays are used that match the size of the projected solar disk. The first has a series of ellipses of different size and shape. Each sunspot is measured by finding the ellipse that best matches a spot in area. For irregular spots, area is either added or subtracted to account for the irregular shape. The second overlay has correction for limb area foreshortening. This is a correction factor applied to the sunspot for regions not at the centre of the disk. The corrected area of each sunspot in a region is then added to give a group area. The third overlay is a Stonyhurst overlay marked with latitude and longitude so that the analyst can accurately identify the location of the spot region. Figures 4.15 to 4.17 in chapter 4 show examples of these overlays.

The typical unit used when measuring sunspot areas (or other areas of features on the solar disk) are millionths of the solar hemisphere or micro-hemispheres (μh), where $1\ \mu h = 3.04 \times 10^6\ km^2$.

There have been several analyses of sunspot area, although these are usually combined with other parameters such as sunspot number, and may include a discussion of how observations of sunspot features have changed over the historical record (for example: Wilson and Hathaway, 2005; Wilson and Hathaway, 2006; Baranyi et al., 2013). These will be discussed in chapter 4.

**2.4 Sunspot Classification**

Sunspot regions have been classified primarily as predictors for solar flare activity. Sunspots may be used as proxies for the magnetic field in a region (McIntosh, 1990).

The first attempt to classify sunspots was by Cortie (1901). His classification scheme is described in table 2.1. Under this scheme sunspot groups are assigned a type (I through V) with these categories having subtypes designated by a letter.

| Type | Description |
|---|---|
| **I** | **A group of 1 or more scattered spots** |
| **II** | **The two spot formation:** |
| IIa | In which the leader is the principal spot. |
| IIb | In which the following spot is the principal spot. |
| IIc | In which both spots are more or less equal. |
| **III** | **A train of spots:** |
| IIIa | With well defined principal spots. |
| IIIb | Without well defined principal spots, but consisting mostly of penumbral patches patches with scattered irregular umbra. |
| **IV** | **Single spots with:** |
| IVa | A single spot of round and regular outline. |
| IVb | A single spot of round and regular outline with smaller companions. |
| IVc | A single spot of irregular outline. |
| IVd | A single spot of irregular outline with a train of smaller companions. |
| IVe | A single spot of irregular outline with smaller companions not in a train. |
| **V** | **An irregular group of larger spots** |

**Table 2.1**: Reproduction of Cortie's Sunspot Classification System (Cortie, 1901)

Generally, sunspots evolve from Type I through to Type IV or V. This has been a theme of most classification systems which try to incorporate the sunspot evolution into their schemes. Cortie goes into somewhat more detail and suggests that a typical sunspot evolution may be: Type I, IIb, IIa, IIIa, IIa, IVd, IVa, I. It is interesting to note that most of the parameters that later schemes used are in Cortie's classification system, although often not in as much detail. There is a 'shape' for the sunspots (regular or irregular), some indication of distribution of sunspots in a region and presence of umbra.

In the 1930s, observers in Zurich developed an alternate scheme to Cortie's earlier work. They used nine categories to describe the evolution of sunspot groups (Bray and Loughhead, 1964). Table 2.2 summarises the Zurich classification system.

The Zurich system has little information about the distribution of spots or the complexity of regions, but does describe penumbra and size of the groups better. It is also a better description of the evolution of sunspot groups.

| Class | Description |
|---|---|
| A | One or more tiny spots that do not demonstrate bi-polarity or exhibit penumbra. |
| B | Two or more tiny spots that demonstrate bi-polarity but do not exhibit penumbra. |
| C | Two or more spots that demonstrate bi-polarity and either the lead spot or trailing spot has a penumbra. |
| D | Two or more spots that demonstrate bi-polarity and the lead spot and trailing spots display a penumbra. The "D" Type will occupy 10 degrees or less of Solar longitude. |
| E | This group type is similar to the "D" type but spreads between 10 and 15 degrees of Solar longitude. |
| F | Largest and most extensive of groups, similar to "E" type but will cover in excess of 15 degrees of Solar longitude. |
| G | The decayed remnant of "D", "E", and "F" groups. Demonstrates a bi-polar group with penumbras |
| H | The decayed remnant of "C", "D", "E", and "F" groups. A single spot group with penumbra. Must be larger than 2.5º in diameter. The "H" type occasionally is accompanied by a few small spots. |
| J | The same as the "H" type but has a diameter less than 2.5º. |

**Table 2.2**: Reproduction of the Zurich Sunspot Classification System from the American Association of Variable Star Observers Website: http://www.aavso.org/zurich-classification-system-sunspot-groups

In the 1930s regular Hα images of the solar disk began, with the aim of observing flares, and it was realised that there were some correlations between the structure of sunspot groups and flares. Giovanelli (1939) was the first person to attempt to use sunspots as a predictor for solar flares. Specific details of this paper will be discussed further in chapters 5 and 10. It is noted here though that flare prediction played a significant role in classification schemes.

In the 1960s McIntosh modified the Zurich system (McIntosh, 1990). He removed the G and J class and added 2 other parameters to the scheme. This was called the Modified Zurich Classification, but has now largely become known as the McIntosh Sunspot Classification System. This is the system currently in use. Tables 2.3 to 2.5 are derived from McIntosh's 1990 paper.

| Type | Description |
|---|---|
| A | A single spot or spots with extent less than 3º of one polarity without penumbra. |
| B | Bi-polar group with no penumbra. |
| C | Bi-polar group with penumbra on spots of one polarity. |
| D | Bi-polar group with penumbra on spots of both polarities and extent of less than 10º. |
| E | Bi-polar group with penumbra on spots of both polarities and extent between 10º and 15º. |
| F | Bi-polar group with penumbra on spots of both polarities and extent of more than 15º. |
| H | A single spot (possibly surrounded by small spots) with penumbra of single polarity. |

**Table 2.3**: Z Parameter – Zurich Classification

| Type | Description |
|---|---|
| X | No penumbra. |
| R | Rudimentary penumbra. |
| S | Small symmetric penumbra (less than 2.5º in extent). |
| A | Small asymmetric penumbra (less than 2.5º in extent). |
| H | Large symmetric penumbra (more than 2.5º in extent). |
| K | Large asymmetric penumbra (more than 2.5º in extent). |

**Table 2.4**: P Parameter – Penumbra Classification (largest spot in group)

| Type | Description |
|---|---|
| X | Not applicable. |
| O | Open: few if any spots between the leader and trailing spots. |
| I | Intermediate: Some spots between the positive and negative polarities |
| C | Compact: The entire region is heavily populated with spots, often with little delineation between spots of opposite polarity. In extreme cases the region is just one large penumbra filled with many umbra. |

**Table 2.5**: C Parameter– Compactness Classification

Obviously this system has some combinations of parameters that are not physically feasible. For example, a group would not be allowed to have a ZPC combination from tables 2.3 to 2.5 of ASO. This is because a Z classification of A has no penumbra. The

allowed classes are summarised below in table 2.6, a reproduction of McIntosh's Table II (McIntosh, 1990).

| **Zurich** | **Penumbra** | **Compactness** | **# of Types** |
|---|---|---|---|
| A | X | X | 1 |
| B | X | O, I | 2 |
| C | R, S, A, H, K | O, I | 10 |
| D, E, F | R | O, I | 6 |
| D, E, F | S, A, H, K | O, I, C | 36 |
| H | R, S, A, H, K | X | 5 |
| **Total Allowed Types** | | | **60** |

**Table 2.6**: Allowed McIntosh Classifications

Figure 2.1 shows a sample of regions with their McIntosh classifications.

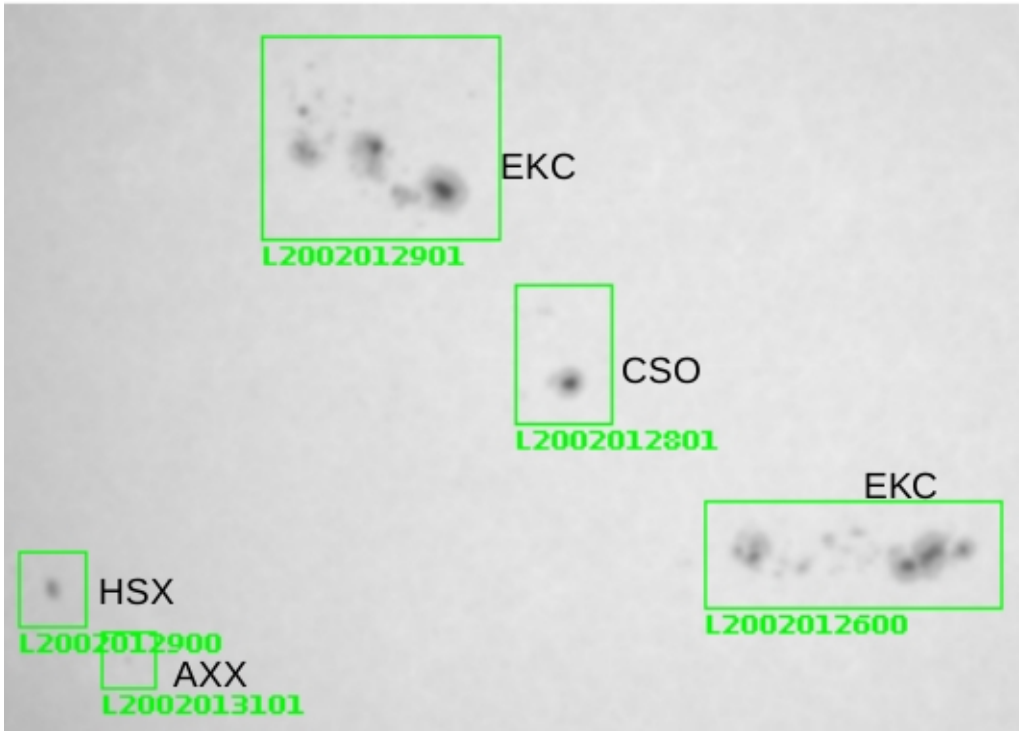



**Figure 2.1**: Sample McIntosh Classifications. This image was taken from Learmonth GONG on 2002/01/31 00:01:16 UT

The McIntosh Classification has been adopted by the USAF SOON sites which perform

daily sunspot classifications for each sunspot region.

Sunspot magnetic parameters have been observed at Mt Wilson since about 1920. It was Hale et al. (1919) who created the magnetic classification system. This classification system is still in use today with some extra categories added to it. Hale's original scheme only had three categories: α, β and γ. The description for these categories remain unchanged. An extra category has been added subsequently, the βγ class. The Mount Wilson Magnetic classification system is shown in table 2.7.

| Type | Description |
|---|---|
| α | A unipolar region (a significant flux of only one polarity in the region) |
| β | A bipolar region with a clear delineation between the polarities (flux of both polarities) |
| βγ | A bipolar region with some intermingling between polarities |
| γ | A bipolar region in which the polarities are completely intermingled (no deliniation between polarities) |

**Table 2.7**: Magnetic Classification System

In addition, a δ can be appended to β, βγ or γ classifications if a sunspot has umbra of both polarities within the same penumbra. Thus the magnetic classification has in total seven categories.

### 2.5 Umbra to Sunspot Area Ratio

Another parameter often discussed is umbral-to-penumbral area ratio or umbral to total sunspot area ratio. In this discussion the latter is used. Spot umbral-to-area ratios are often used in irradiance models (Brandt and Schmidt, 1990). This parameter is not used or measured as often as it was. The USAF SOON ceased measurement of this parameter in the early 1980s as the USAF considered this an additional task for the analysts with no significant benefit (Personal Communication, John Kennewell). Nonetheless significant work has been done analysing this particular parameter. Bray and Loughhead (1964) report a umbral-to-spot area of 0.17, although express some caution as this value varies significantly from sunspot-to-sunspot and this ratio is smaller near sunspot maximum.

Vaquero et al. (2005) have a table summarising some measurements made by different observers from about 1850 onwards. Values of umbral/spot area range from 0.17 to 0.24.

Brandt and Schimdt suggest different ratios for different sized spots. They report a ratio of 24% for large spots and 19% for small spots (Brandt and Schmidt, 1990).

### 2.6 Summary

Sunspots have been studied in some detail from the 17$^{th}$ century on. Various parameters are now used to describe and quantify sunspots. The most common of these are:

- International Sunspot Number (ISN), formerly the Wolf or Zurich Sunspot Number. The recording of this index is undertaken by the Solar Influences Data Center in Belgium.
- Sunspot areas: Formerly recorded by Royal Greenwich Observatory and now undertaken by the USAF SOON network.
- Sunspot classifications: the currently used schemes are the McIntosh scheme for sunspot morphology and the Mount Wilson scheme for sunspot magnetic fields.
- Umbral-to-Sunspot area ratios: These are unfortunately no longer recorded on a regular basis, but used to be recorded by RGO.

Other observatories and groups have undertaken these observations at different times, but the ones listed above are the most continuous data sets.

# 3. Solar Radio Bursts

## 3.1 Introduction

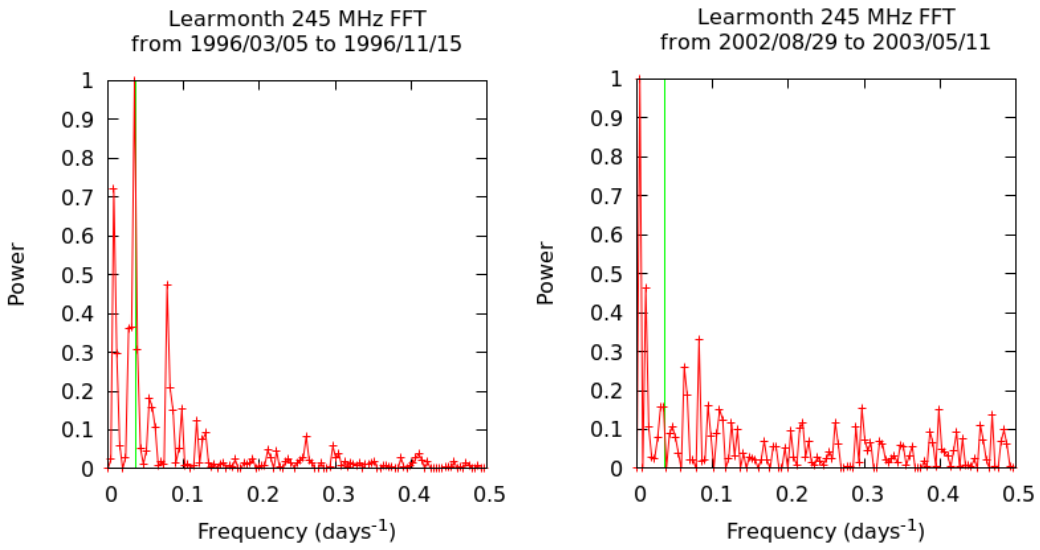
True solar radio astronomy began in the early 1940s when radar operators in England detected interference from the Sun on their equipment. Within months scientists began observing the Sun at radio wavelengths (Smith, 1967). Prior to this radio operators had noted interference, but had failed to identify the source correctly as radio emission from the Sun (Smith, 1967). In the current era, solar radio interference has become more important, as satellite communications, GPS, wireless devices and other technologies have become prevalent that may be impacted by solar activity.

The unit used throughout this chapter for radio flux is the Solar Flux Unit, where 1 SFU = $10^{-22}$ $Wm^{-2}Hz^{-1}$.

The Sun emits a continuous background radio energy. This is due to thermal emission. Figure 3.1 shows a typical quiet sun radio spectrum near solar minimum. The frequency ranges from 245MHz at high-VHF, through the UHF range of 300 MHz to 3 GHz and up to 15.4 GHz into the SHF range (3-30GHz).

This curve does not follow a black body radiation curve, as emission at different wavelengths are associated with different heights in the solar corona, and these vary significantly in temperature (Smith, 1967).

The slowly varying component is associated with plage and is most easily seen in the 3 GHz range (Smith, 1967) where there is an inflection in the spectrum in figure 3.1. This component has two distinct periods: one of 27 days and is associated with the rotation rate of the sun and one of 11 years associated with the sunspot cycle, or rather the plage areas (Smith, 1967). The 27 day periodicity can disappear often for months at a time, particularly near sunspot maximum. This is shown in figure 3.1 (a and b). These figures show the power spectrum of FFTs performed on 256 days of data near solar minimum and solar maximum respectively. The green line shows the 27 day sunspot cycle. Figure 3.1a (near sunspot minimum) shows a distinct peak at a frequency of 0.037 $days^{-1}$ (27 day period). Figure 3.1b (near sunspot maximum) is lacking such a distinct peak. This effect is due to active regions being distributed more evenly over the sun at solar maximum than at solar minimum.




(a) (b)

**Figure 3.1**: FFT of Solar Radio Flux Near (a) Solar Minimum and (b) Solar Maximum. Data from Learmonth RSTN obtain from NGDC. The FFT was done over daily flux values averaged between 05:55 UT and 06:05 UT for 256 consecutive days.

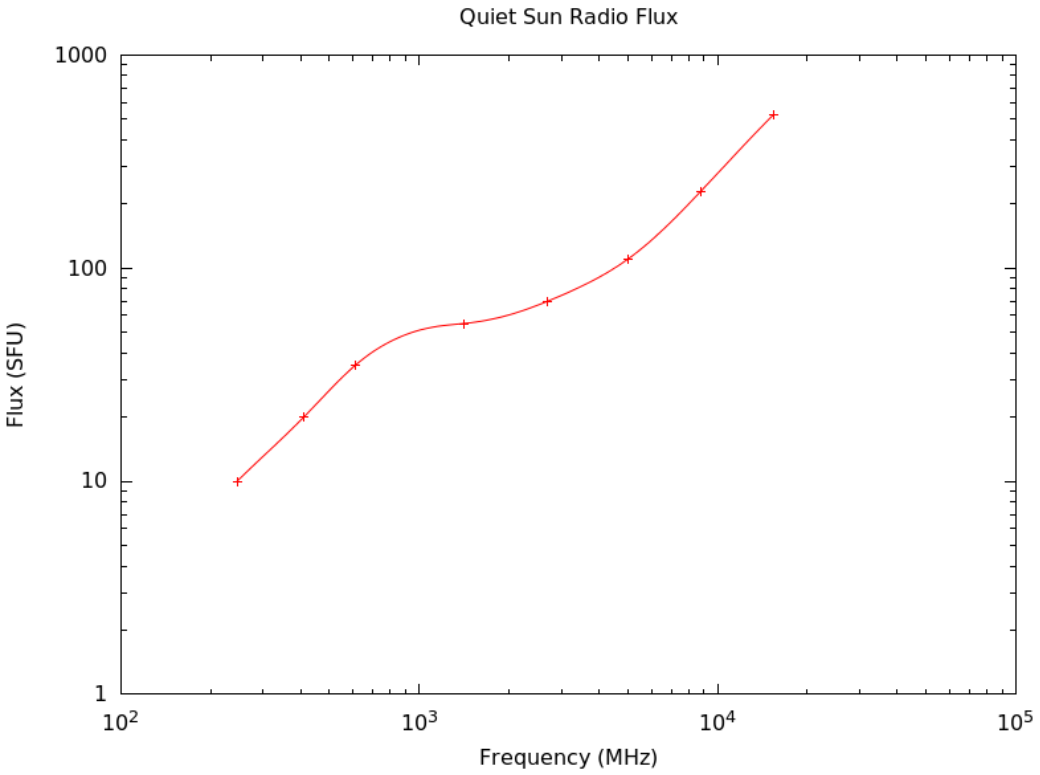



**Figure 3.2**: Quiet Sun Radio Flux

Finally there is the burst component. This originates in a rapid release of energy from a region. At metric wavelengths (300 MHz near the lower end of the spectrum in figure 3.1) the solar output can increase by more than five orders of magnitude above its background level, often in the space of minutes.

Mathematically then, the radio output of the sun at a particular frequency, *S*, at a particular time *t* can be written as:

$$S(t) = S_{back} + S_{SVC}(t) + S_{burst}(t) \quad , \qquad \text{Eqn 3.1}$$

where $S_{back}$ is the background component, $S_{SVC}$ is the slowly varying component and $S_{burst}$ is the burst component.

Unfortunately none of the terms in equation 3.1 are well known, they vary from cycle-to-cycle (see section 3.3 and 4.5). Understanding the behaviour of the burst component in particular is one of the more important aspects of space weather prediction. In addition each of the above terms is frequency dependent.

### 3.2 Burst Mechanisms

There are three main mechanisms thought to be responsible for solar radio bursts: (i) synchrotron emission, (ii) bremsstrahlung (free – free electron emission) and (iii) plasma emission. A fourth mechanism, cyclotron emission, is responsible for noise storms which are long duration bursts typically observed below 1 GHz. Details of these mechanisms can be found in the review paper by Bastian et al. (1998). A summary for each mechanism is provided below.

#### 3.2.1 Synchrotron and Cyclotron Emission

Both cyclotron and synchrotron emission is caused by charged particles moving in a magnetic field. The gyrofrequency, $\nu$, for an electron is given by:

$$\nu = 2.8 B / \gamma \quad , \qquad \text{Eqn 3.2}$$

where $\nu$ is the frequency in MHz, $B$ is is the magnetic field in Gauss and $\gamma$ is the Lorentz factor (Serway et al., 1997).

As an electron's speed increases (g approaches 0) more harmonics appear until, at increasingly higher energies, the emission can appear as a continuum. The frequency of maximum power is given by:

$$\nu = 1.2 B \gamma \quad , \qquad \text{Eqn 3.3}$$

(Lyne and Graham-Smith, 2006).

Further complicating matters is synchrotron self-absorption, where ions and electrons in the plasma absorb the emitted radiation to maintain equilibrium. This causes an upper intensity limit of solar radio bursts which appears to be around 50 kSFU for frequencies of about 3 GHz or higher.

### 3.2.2 Bremsstrahlung Emission

Bremsstrahlung (German for breaking radiation) emission occurs when a charged particle interacts with another charged particle (Kruger, 1979). In this case, it is the interaction of an electron with an ion that is of interest. The electron is accelerated by the electric field of the ion, causing it to emit radiation, conserving overall energy in the process. In a plasma such as the solar corona, the electron's path wiggles as a result of the interactions with the ion, causing broadband emission (Gary, lecture notes: https://web.njit.edu/~gary/728/Lecture2.html).

### 3.3.3 Plasma Emission

Plasma emission is caused by electron energy being converted into a plasma (Langmuir) wave (Kruger 1979). This wave can then be converted into electromagnetic energy when it interacts with another plasma wave or an electromagnetic wave.

The frequency of emission is at the plasma frequency, or its harmonic, given in equation 3.3

$$f_p \approx 9\sqrt{n_e} \quad , \qquad \text{Eqn 3.3}$$

where $f_p$ is the plasma frequency in Hz and $n_e$ is the electron number density in $m^{-3}$.

### 3.3 Site Variations

Kintner et al. (2009) claim that there are inconsistencies between the USAF RSTN sites, at least for large bursts. Their analysis was performed on several large bursts observed on 1415 MHz that interfered with GPS signals. The first thing to consider is the saturation of the RSTN receivers. These values are shown in table 3.1. A proposal was submitted to the USAF after the December 2006 bursts to have the 1415 MHz amplifiers adjusted to saturate at 500 kSFU, but this was rejected (Kennewell, Personal Communication).

| Frequency (MHz) | Maximum Measurable Burst (kSFU) |
|---|---|
| 245 | 500 |
| 410 | 500 |
| 610 | 500 |
| 1415 | 100 |
| 2695 | 50 |
| 4995 | 50 |
| 8800 | 50 |
| 15400 | 50 |

**Table 3.1**: Saturation of RSTN Radio Amplifiers

It should be noted that the RSTN specifications allow a 20% variation from these values, so saturation may occur at different points on different days, depending on the results of daily calibrations.

The previous RSTN software (which was replaced in late 2015) included a glitch. When bursts saturated the amplifiers the computer locked-up, requiring a restart of the computer. This occurred with the December 2006 bursts and may account for some of the discrepancies across the network reported by Kintner et al. (2009).

Kintner et al. (2009) show that, during these bursts, Sagamore Hill reported much lower values than Palehua. The exact cause is unknown but may have been because the Sagamore Hill computer locked up and data was lost, thus the highest value they observed was the one that was reported.

So the question that follows is, how consistent is the Radio Solar Telescope Network across sites? Fortunately, at least a partial answer to that question can be obtained using

noon flux values that are taken each day. After the midday calibration the antenna is kept stationary and the Sun is allowed to drift through the antenna beam. The peak flux value during this procedure is recorded at all eight frequencies. This is then reported to the Space Weather Prediction Center (SWPC) and archived at NGDC. Noon flux data is available from all sites from 1988 to 2011.

The U.S. Air Force Weather Agency (AFWA) have required, at various times, for each RSTN site to agree to within 10% of either (a) the network average or (b) a reference site. Unfortunately it is unknown during which periods the network average or the reference site was used, or which reference sites. In this analysis the network average is used.

The other limitation of this analysis is that the flux values are not obtained at the same time at all sites. The requirement is that the measurements be done within 1 hour of Central Meridian Passage (CMP) when the sun is directly overhead. Observations can be made up to ±3 hours of CMP, but those outside the 1 hour criteria must be qualified.

Bursts can be qualified for weather (which also covers attenuation for being outside 1 hour of CMP range), interference, equipment problems or bursts (usually noise storms). Unfortunately these qualifiers are not given in the archive.

The daily network average ($\overline{S_j}$) for each frequency *j*, was computed for each day and frequency where 3 or 4 sites had noon flux values reported. The percentage difference from the average was calculated for each day using equation 3.4:

$$d_{ij} = \frac{100\left(\overline{S_j} - S_{ij}\right)}{\overline{S_j}} \quad , \qquad \text{Eqn 3.4}$$

where $d_{ij}$ is the percentage difference for the *i*th site and and *j*th frequency, $\overline{S_j}$ is the network average flux for the *j*th frequency, and $S_{ij}$ is the flux for the *i*th site and *j*th frequency.

The noon flux values on 410 MHz and 610 MHz from San Vito were omitted from this analysis, and subsequent analysis of bursts. These two frequencies at San Vito are prone to RFI, so even though they are sometimes reported, the times when they were not used due to interference are unavailable.

The average difference and the standard deviation of the differences were found for solar cycles 22 to 24. These results are shown in tables 3.2 a-d.

| Frequency (MHz) | | Cycle 22 | Cycle 23 | Cycle 24 |
|---|---|---|---|---|
| 245 | Average difference (%) | 1.25 | -3.21 | 1.54 |
| | Std Dev of Differences (%) | 23.33 | 34.77 | 15.99 |
| 410 | Average difference (%) | 0.96 | 2.79 | 0.16 |
| | Std Dev of Differences (%) | 12.44 | 16.57 | 7.02 |
| 610 | Average difference (%) | -0.57 | -1.89 | 0.84 |
| | Std Dev of Differences (%) | 11.06 | 12.57 | 7.39 |
| 1415 | Average difference (%) | -0.35 | -2.64 | 0.98 |
| | Std Dev of Differences (%) | 8.78 | 13.75 | 4.02 |
| 2695 | Average difference (%) | -0.21 | 0.83 | -0.22 |
| | Std Dev of Differences (%) | 7.38 | 7.18 | 9.29 |
| 4995 | Average difference (%) | 0.62 | 0.73 | -2.17 |
| | Std Dev of Differences (%) | 6.21 | 5.84 | 7.39 |
| 8800 | Average difference (%) | -0.41 | 0.83 | 0.67 |
| | Std Dev of Differences (%) | 6.36 | 6.70 | 8.45 |
| 15400 | Average difference (%) | -0.54 | -2.92 | -0.82 |
| | Std Dev of Differences (%) | 4.98 | 6.11 | 8.20 |

**Table 3.2a**: Learmonth Noon Flux Deviation from Site Average (1988 to 2011)

| Frequency (MHz) | | Cycle 22 | Cycle 23 | Cycle 24 |
|---|---|---|---|---|
| 245 | Average difference (%) | -0.64 | 4.50 | -1.66 |
| | Std Dev of Differences (%) | 24.52 | 29.21 | 15.77 |
| 1415 | Average difference (%) | -0.16 | 1.92 | 2.71 |
| | Std Dev of Differences (%) | 7.23 | 8.67 | 2.66 |
| 2695 | Average difference (%) | -1.95 | -1.30 | 0.77 |
| | Std Dev of Differences (%) | 8.26 | 6.76 | 4.10 |
| 4995 | Average difference (%) | -1.08 | 0.68 | 4.23 |
| | Std Dev of Differences (%) | 6.39 | 6.09 | 5.46 |
| 8800 | Average difference (%) | 0.88 | -1.04 | 0.52 |
| | Std Dev of Differences (%) | 6.19 | 7.77 | 9.41 |
| 15400 | Average difference (%) | 0.67 | 0.88 | 1.31 |
| | Std Dev of Differences (%) | 4.95 | 4.45 | 6.59 |

**Table 3.2b**: San Vito Noon Flux Deviation from Site Average

| Frequency (MHz) | | Cycle 22 | Cycle 23 | Cycle 24 |
|---|---|---|---|---|
| 245 | Average difference (%) | 1.58 | 1.20 | -1.49 |
| | Std Dev of Differences (%) | 23.51 | 30.80 | 12.27 |
| 410 | Average difference (%) | -1.33 | -2.05 | 3.09 |
| | Std Dev of Differences (%) | 12.40 | 13.60 | 6.62 |
| 610 | Average difference (%) | 0.71 | 1.19 | 2.16 |
| | Std Dev of Differences (%) | 9.45 | 9.03 | 7.62 |
| 1415 | Average difference (%) | 0.16 | 0.89 | -1.33 |
| | Std Dev of Differences (%) | 6.36 | 9.37 | 2.25 |
| 2695 | Average difference (%) | 0.99 | 0.92 | 0.26 |
| | Std Dev of Differences (%) | 7.75 | 5.88 | 4.63 |
| 4995 | Average difference (%) | -0.09 | -0.91 | -0.30 |
| | Std Dev of Differences (%) | 6.43 | 6.12 | 4.50 |
| 8800 | Average difference (%) | -0.36 | 0.35 | -1.78 |
| | Std Dev of Differences (%) | 5.38 | 7.16 | 7.21 |
| 15400 | Average difference (%) | -0.09 | 0.24 | -0.29 |
| | Std Dev of Differences (%) | 5.61 | 6.15 | 8.63 |

**Table 3.2c**: Sagamore Hill Noon Flux Deviation from Site Average

| Frequency (MHz) | | Cycle 22 | Cycle 23 | Cycle 24 |
|---|---|---|---|---|
| 245 | Average difference (%) | -2.18 | -2.58 | 1.67 |
| | Std Dev of Differences (%) | 28.95 | 36.13 | 17.68 |
| 410 | Average difference (%) | 0.36 | -0.75 | -3.25 |
| | Std Dev of Differences (%) | 13.41 | 14.88 | 7.21 |
| 610 | Average difference (%) | -0.14 | 0.70 | -3.01 |
| | Std Dev of Differences (%) | 10.23 | 9.39 | 6.40 |
| 1415 | Average difference (%) | 0.32 | -0.22 | -2.32 |
| | Std Dev of Differences (%) | 6.64 | 10.49 | 3.40 |
| 2695 | Average difference (%) | 1.19 | -0.43 | -0.82 |
| | Std Dev of Differences (%) | 6.80 | 7.28 | 4.40 |
| 4995 | Average difference (%) | 0.61 | -0.49 | -1.84 |
| | Std Dev of Differences (%) | 7.81 | 6.20 | 4.26 |
| 8800 | Average difference (%) | -0.12 | -0.13 | 0.62 |
| | Std Dev of Differences (%) | 6.97 | 5.94 | 4.41 |
| 15400 | Average difference (%) | -0.04 | 1.75 | -0.23 |
| | Std Dev of Differences (%) | 5.03 | 6.79 | 4.55 |

**Table 3.2d**: Palehua Deviation Noon Flux from Site Average

All of the average differences in tables 3.2 are well within the standard deviation of differences for the same site and frequency, suggesting that there is no significant baseline offset at any particular site.

The lower three frequencies are not in very good agreement, with a wide spread of noon flux differences (over 35% in the case of Palehua in cycle 23). This is not surprising as noon flux measurements are taken even if there is a noise storm in progress (but not if there is any other type of burst). Noise storms predominantly occur at frequencies less than 1 GHz (Wild et al., 1963). As noise storms can last for days and vary dramatically even on a scale of minutes, the flux values observed will be significantly different from site-to-site on the frequencies susceptible to noise storms.

For 1415 MHz and above the standard deviations are typically between 5% and 10%, showing that about 65% of the time RSTN meets its stated criteria for each site to agree within 10% of the daily network average. However this implies that up to 35% of the time they are not meeting their own criteria. Of course some cycles, sites and frequencies

are better than others. For example 1415 MHz from Palehua in cycle 24 shows a standard deviation of 3.4%, implying that 95% of the time this frequency was within 7% of the network average, well within RSTN guidelines.

Consistency between sites could be improved by upgraded equipment, better training of maintenance and analyst personnel or more stringent procedures, or combinations of all of these. However this analysis does not invalidate the RSTN data set. but researchers need to be aware of the range of variation across the network. Indeed, since the noon flux values are available, at least from 1988, researchers should be able to take into account site variations in an analysis of bursts.

Whilst the above analysis gives a general overview of potential offsets and variations, a better measure is a comparison between the burst reports from each site. To this end, the RSTN burst reports archived on the NGDC website were examined. Noise storms were excluded from the analysis as were the 410 MHz and 610 MHz from San Vito. Only bursts greater than 100 SFU were considered. Over the years 1980 to 2015, RSTN has increased the burst threshold from 10 SFU to 100 SFU currently. For consistency, the higher value is used. The data set used ends at the end of 2010. Unfortunately there are some significant errors, mainly an incorrect end time (for example small non-noise storm bursts with durations of hours).

When comparing one site against another, not only did the burst start and end times for each site need to overlap, but the peak times also needed to be within 5 minutes of each other. This also had the effect of removing bursts where the peak was either before sunset or after sunset at one of the sites.

Figure 3.3 shows an example of such a comparison for Learmonth and Palehua at 245 MHz. The other plots can be found in appendix A. Tables 3.3 (a) to (e) summarise these comparisons where n is the number of points in the plot, m is the gradient $s_m$ is the standard deviation in the gradient, b is the y-axis intercept (SFU), $s_b$ is the standard deviation of the intercept (SFU) and r is the correlation coefficient. In all cases the fit parameters were performed on the log-log plots.

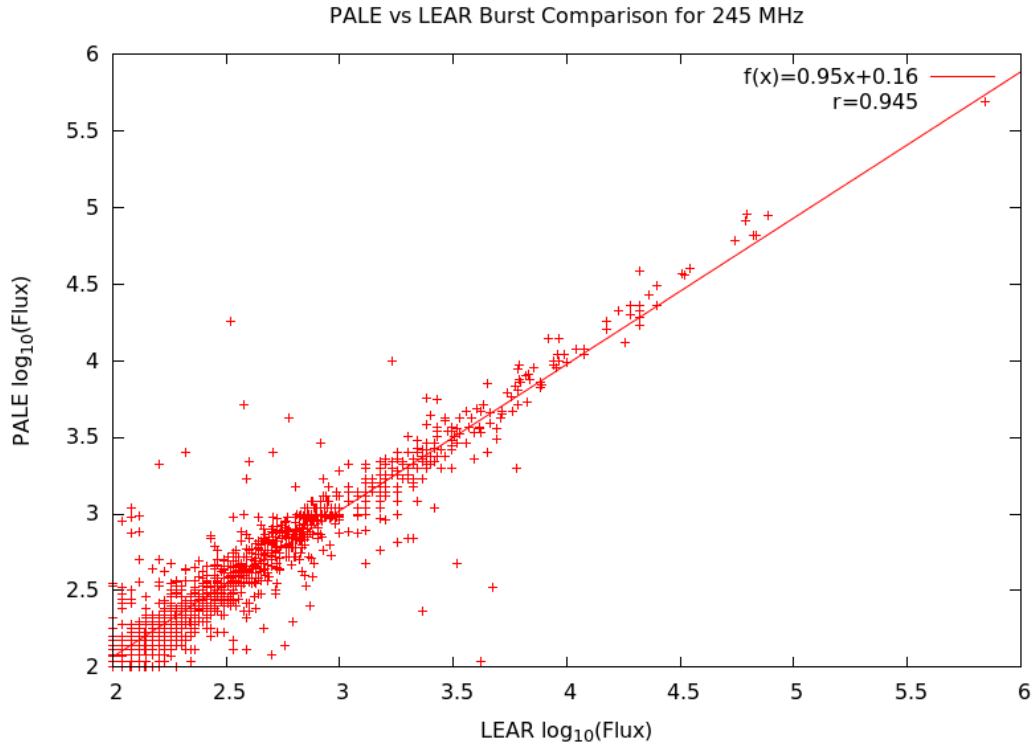



**Figure 3.3**: Palehua and Learmonth Burst Comparison for 245 MHz (1981 to 2010)

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1296 | 0.954 | 0.009 | 0.161 | 0.024 | 0.945 |
| 410 | 523 | 0.944 | 0.017 | 0.207 | 0.046 | 0.922 |
| 610 | 360 | 0.974 | 0.014 | 0.038 | 0.038 | 0.965 |
| 1415 | 187 | 1.000 | 0.012 | 0.022 | 0.032 | 0.987 |
| 2695 | 214 | 0.952 | 0.017 | 0.101 | 0.043 | 0.970 |
| 4995 | 259 | 0.979 | 0.011 | 0.063 | 0.029 | 0.983 |
| 8800 | 284 | 0.957 | 0.014 | 0.115 | 0.037 | 0.970 |
| 15400 | 215 | 1.004 | 0.012 | -0.024 | 0.034 | 0.984 |

**Table 3.3a**: Palehua vs Learmonth Burst Comparison Regression Analysis

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1394 | 0.982 | 0.007 | 0.041 | 0.018 | 0.967 |
| 1415 | 192 | 0.934 | 0.017 | 0.146 | 0.044 | 0.971 |
| 2695 | 214 | 0.956 | 0.011 | 0.109 | 0.027 | 0.987 |
| 4995 | 263 | 0.981 | 0.011 | 0.072 | 0.029 | 0.983 |
| 8800 | 297 | 0.930 | 0.015 | 0.191 | 0.038 | 0.966 |
| 15400 | 233 | 0.961 | 0.011 | 0.075 | 0.030 | 0.985 |

**Table 3.3b**: San Vito vs Learmonth Burst Comparison Regression Analysis

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1353 | 0.954 | 0.008 | 0.151 | 0.021 | 0.955 |
| 1415 | 199 | 0.932 | 0.026 | 0.169 | 0.068 | 0.931 |
| 2695 | 191 | 0.965 | 0.018 | 0.075 | 0.045 | 0.969 |
| 4995 | 252 | 0.963 | 0.010 | 0.080 | 0.026 | 0.986 |
| 8800 | 276 | 0.948 | 0.015 | 0.136 | 0.039 | 0.967 |
| 15400 | 237 | 1.000 | 0.012 | 0.008 | 0.030 | 0.984 |

**Table 3.3c**: Sagamore Hill vs San Vito Burst Comparison Regression Analysis

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 137 | 0.947 | 0.026 | 0.164 | 0.069 | 0.953 |
| 1415 | 29 | 0.652 | 0.070 | 0.910 | 0.200 | 0.874 |
| 2695 | 28 | 1.012 | 0.033 | -0.059 | 0.087 | 0.986 |
| 4995 | 33 | 0.942 | 0.060 | 0.157 | 0.153 | 0.943 |
| 8800 | 38 | 0.925 | 0.052 | 0.182 | 0.134 | 0.948 |
| 15400 | 25 | 0.969 | 0.041 | 0.108 | 0.113 | 0.980 |

**Table 3.3d**: Palehua vs San Vito Burst Comparison Regression Analysis

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1295 | 0.950 | 0.010 | 0.164 | 0.027 | 0.936 |
| 410 | 542 | 0.915 | 0.019 | 0.274 | 0.049 | 0.905 |
| 610 | 418 | 0.916 | 0.016 | 0.210 | 0.042 | 0.944 |
| 1415 | 229 | 0.952 | 0.014 | 0.126 | 0.036 | 0.976 |
| 2695 | 225 | 1.013 | 0.010 | -0.028 | 0.027 | 0.988 |
| 4995 | 293 | 1.001 | 0.012 | -0.017 | 0.030 | 0.981 |
| 8800 | 311 | 0.966 | 0.014 | 0.064 | 0.037 | 0.970 |
| 15400 | 250 | 0.972 | 0.014 | 0.079 | 0.038 | 0.976 |

**Table 3.3e**: Palehua vs Sagamore Burst Comparison Regression Analysis

These tables show a high correlation coefficient (>0.9) for most comparisons. The stand out exception is for 1415 MHz Palehua vs San Vito which is 0.874 (table 3.3d). There are some significant outliers here associated with the December 2006 bursts as discussed by Kintner et. al. (2009).

Most of the comparisons had a gradient (m) more than 2 standard deviations from a value of 1, demonstrating that there are relative differences between amplifiers at different sites, because the burst peak fluxes should be measured the same at all sites. Few of the comparisons have an intercept (b) close to zero. In most cases there is a significant offset. However the uncertainties ($s_b$) in the intercept are all quite large, making it difficult to say if there is a significant offset from site-to-site.

One factor that may affect higher frequencies is attenuation due to low sun angle. This normally affects frequencies greater than 4 GHz, although, if there is thick cloud or rain, attenuation can affect lower frequencies (Zubair et al., 2011). When a burst occurs, analysts are supposed to look at the pre-burst quiet sun value and compare that with the most recent noon flux value that was measured. If there is a difference of more than 10% they are supposed to correct the burst value by the relative difference. As bursts are not corrected if the difference is less than 10% then this may account for some of the variation.

### 3.4 Statistics of Radio Bursts

In order to determine the risk of solar radio interference to various systems it would be useful to know the rate at which radio bursts occur. It is acknowledged at the outset that

radio bursts roughly follow the sunspot cycle, thus making the distribution in time non-uniform. However, it will be shown that the largest bursts, the ones most likely to cause interference problems, occur infrequently and on a time scale significantly larger than a solar cycle. The following work was initiated in Giersch and Kennewell (2013) but only included an analysis of bursts on 1415 MHz. This work encompasses all eight RSTN frequencies. Also re-analysis was done to further refine the data set to remove errors that were unfortunately not taken into account in the original paper. Most of these error were due to a format change in 1982 for the peak flux in the NGDC database.

A similar criterion was used as in section 3.3 in order to combine together burst data from multiple sites.

The first step was determination when the various frequencies at different sites first started collecting data. Table 3.3 shows the first occurrences of bursts at each site. Therefore the data spans from 1966 to 2010, but different sites and frequencies started observing at different times. There was not complete 24 hour coverage of the Sun until San Vito began operation in 1986.

Table 3.4 gives the locations of each site. Using both of the data in these tables, an estimated patrol time for each site and frequency was generated. It is noted that the RSTN sites start and end patrol 15 minutes after sunrise and 15 minutes prior to sunset to reduce multipath interference from the ground at low sun angles. An example of where this effect was used was the sea interferometer (Thompson et al., 2001). These patrol time values assume no downtime for weather, maintenance or equipment problems, which is unrealistic, but shouldn't affect the patrol times more that a few percent, due to the overlapping times when multiple sites were observing the sun.

Table 3.5 shows a summary of the burst statistics. Of note is the 'Burst Rate' column, showing that lower frequencies have significantly more bursts than the higher ones. This is potentially significant for radio communications designers, in that there are fewer bursts in the 3 GHz range, and they are, on average, lower in intensity. The burst rate is simply given by:

$$R_{f0} = N_f / T_f \quad , \qquad \text{Eqn 3.5}$$

where $R_{f0}$ is the burst rate for frequency *f*, in bursts/year, $N_f$ is the number of bursts for frequency *f* and $T_f$ is the estimated observational patrol time in years for frequency *f*.

$N_f$ is the site-combined number of bursts. That is, if two sites reported the same burst then that burst is simply counted as one.

Figure 3.5 shows the median values for bursts. The median flux values are significantly lower than the mean values, due to the larger number of lower than higher flux events and the median dividing equal numbers of events, rather than averaging their fluxes. There is also a much smaller absolute variation in flux values across frequencies (that is, the curve is flatter) than for the mean. These variations in shape between mean and median occur due to the power law distribution of bursts. This is illustrated in figures 3.6 and 3.7.

Figure 3.4 shows the average peak burst flux. Of note is the significant minima at about 3 GHz. This was thought to be significant for Solar Particle Event (SPE) prediction, but is now no longer considered so. This will be discussed further in chapter 9.

| Frequency | LEAR | SVTO | SGMR | PALE |
|---|---|---|---|---|
| 245 | 1979/08/14 | 1986/09/08 | 1969/02/07 | 1979/08/11 |
| 410 | 1979/08/12 | ---------- | 1971/04/19 | 1979/08/09 |
| 610 | 1979/08/12 | ---------- | 1966/01/17 | 1979/08/09 |
| 1415 | 1979/08/15 | 1987/10/16 | 1966/03/16 | 1974/07/01 |
| 2695 | 1979/08/16 | 1987/07/24 | 1966/03/15 | 1979/08/26 |
| 4995 | 1979/08/16 | 1987/04/16 | 1966/08/28 | 1979/09/02 |
| 8800 | 1979/08/16 | 1986/10/24 | 1966/03/16 | 1974/07/01 |
| 15400 | 1979/08/20 | 1987/04/15 | 1968/01/04 | 1979/09/14 |

**Table 3.3**: Date (in YYYY/MM/DD format) of First Observed Burst by Site and Frequency

| Site | Longitude | Latitude |
|---|---|---|
| Learmonth | 114.08 E | 22.23 S |
| San Vito | 17.43 E | 40.40 N |
| Sagamore Hill | 70.82 W | 42.63 N |
| Palehua | 158.11 W | 21.38 N |

**Table 3.4**: RSTN Site Locations

| Frequency (MHz) | Number of Bursts Observed | Average Burst Peak Flux (SFU) | Median Burst peak Flux (SFU) | Estimated Patrol Time (Days) | Burst Rate (Bursts/Year) |
|---|---|---|---|---|---|
| 245 | 15285 | 1142 | 230 | 13236 | 421.79 |
| 410 | 5066 | 1270 | 230 | 12538 | 147.58 |
| 610 | 3457 | 1841 | 230 | 13901 | 90.83 |
| 1415 | 1945 | 1491 | 220 | 13759 | 51.63 |
| 2695 | 1872 | 803 | 210 | 13759 | 49.69 |
| 4995 | 2399 | 831 | 220 | 13759 | 63.68 |
| 8800 | 2820 | 1004 | 230 | 13759 | 74.86 |
| 15400 | 2226 | 1328 | 240 | 13863 | 58.65 |

**Table 3.5**: Burst Statistics for the Eight Discrete RSTN Frequencies

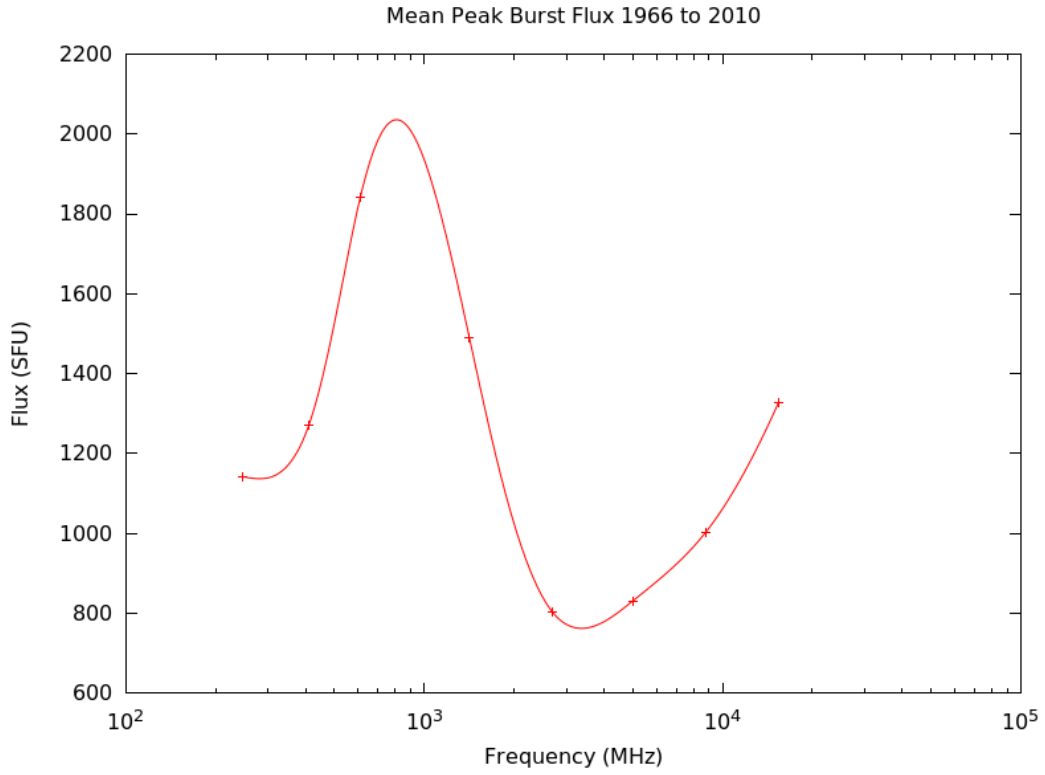



**Figure 3.4**: Mean Peak Burst Flux

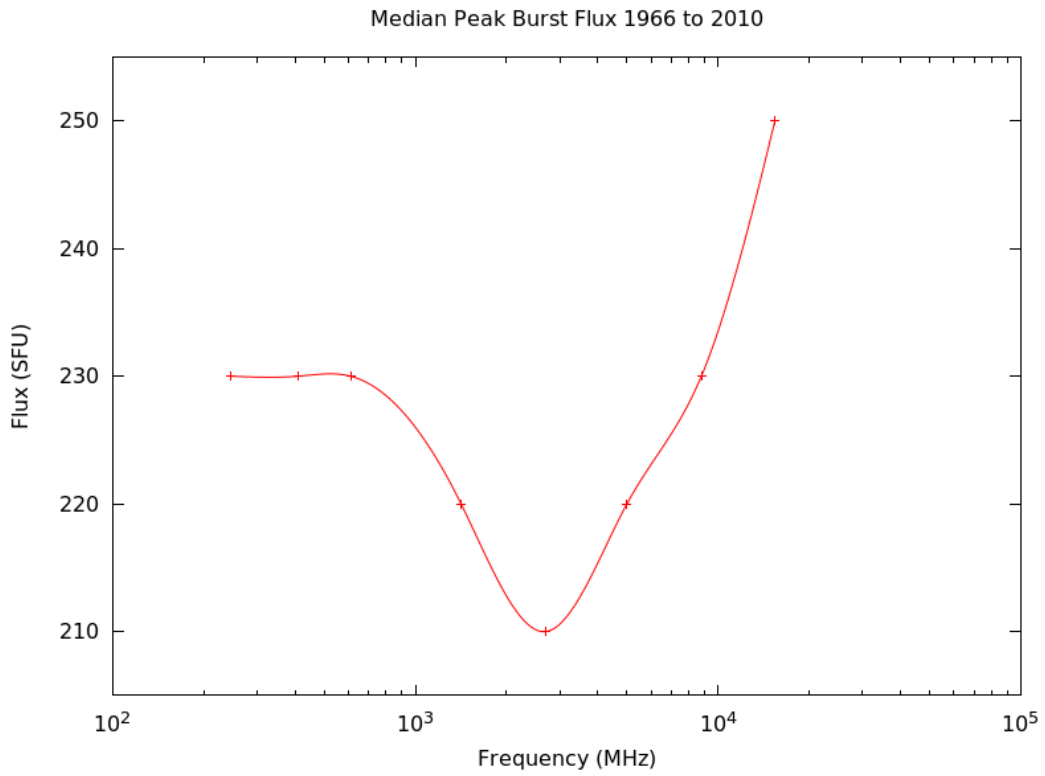
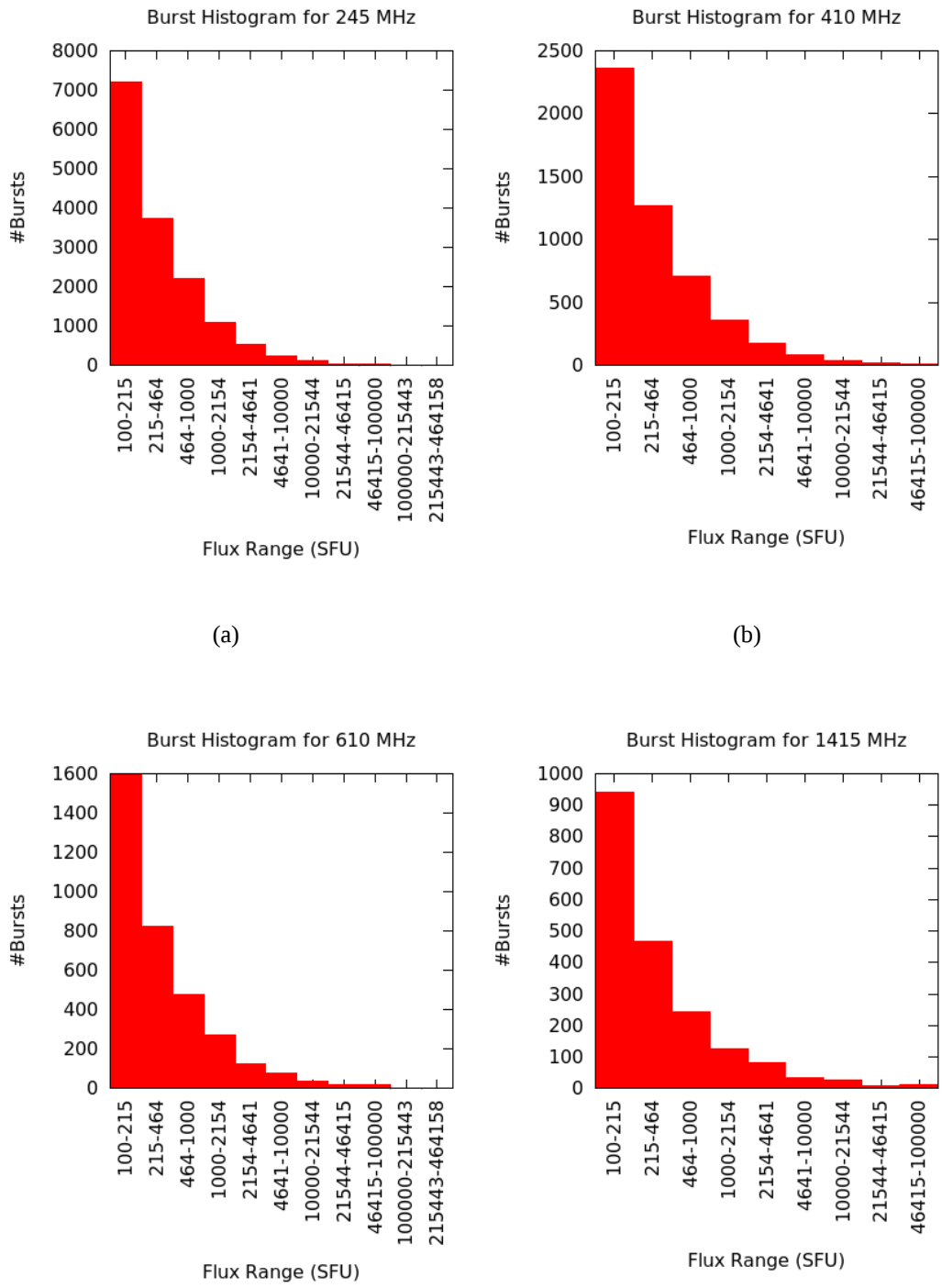



**Figure 3.5**: Median Peak Burst Flux

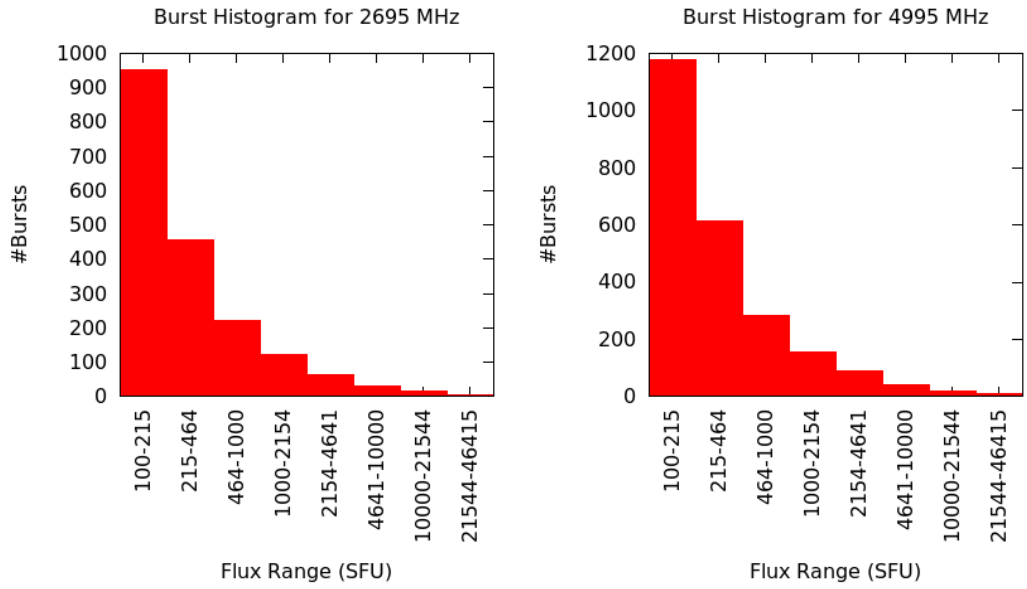
The burst distributions at each frequency are shown in tables 3.6 (a) to (g). In all cases the bursts are clustered at the low flux end. Categories with no bursts have been omitted. Because the flux of bursts is not known when an amplifier saturates, bursts in this range were omitted from these plots.




(a) (b)

(c) (d)

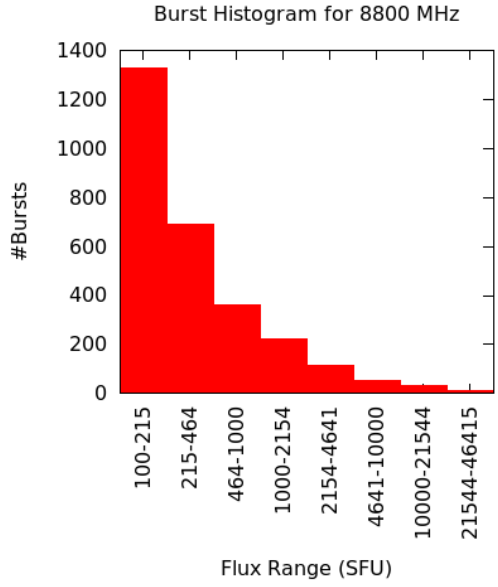
**Figure 3.6 (a-d)**: Histograms of Burst Distributions for RSTN Frequencies 245 MHz to 1415 MHz



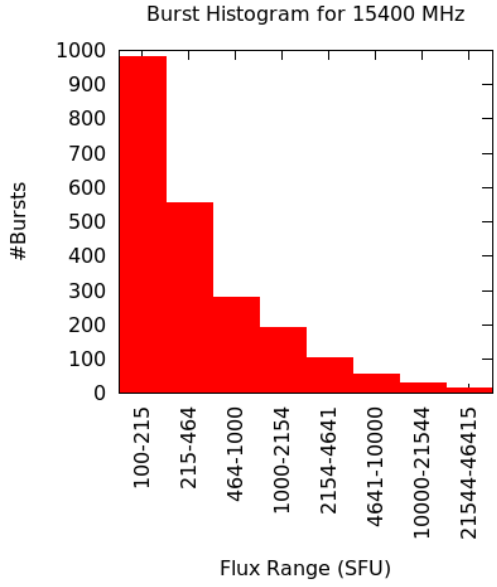

(e)




(f)




(g)




(h)

**Figure 3.6 (e-h)**: Histograms of Burst Distributions for RSTN Frequencies 1415MHz to 15400 MHz

These plots suggest either an exponential or power law fit, and indeed it is a power law that produces the best fits. This is not surprising as other types of solar flares such as

gamma ray flares follow a power law (Newman, 2005). The power law plots are shown for each frequency in Figure 3.7 a – h.

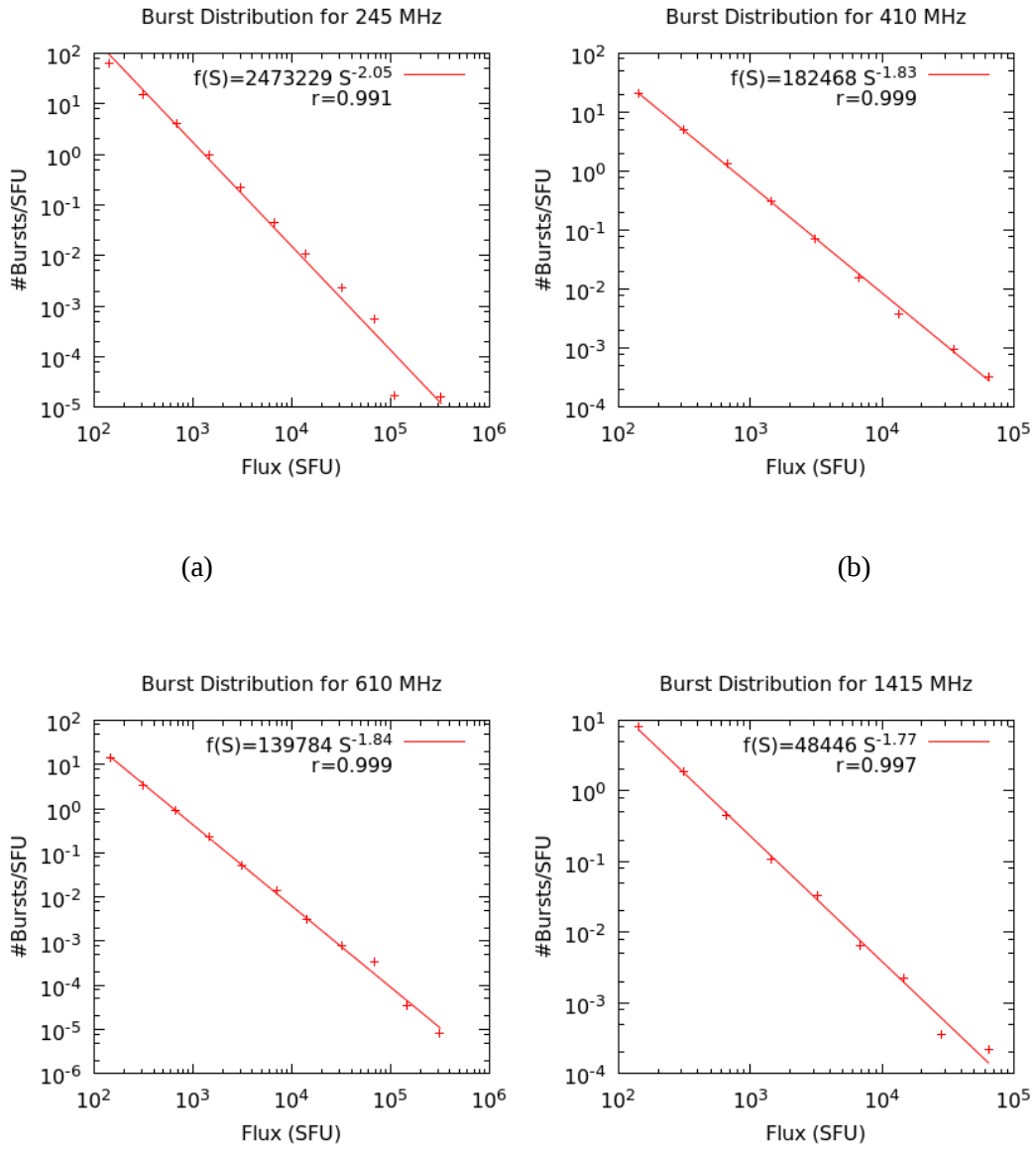
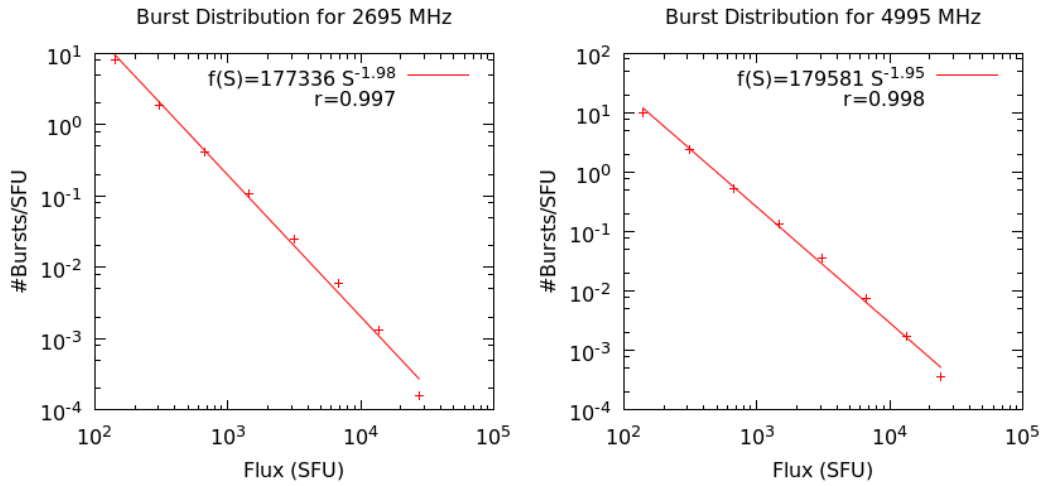
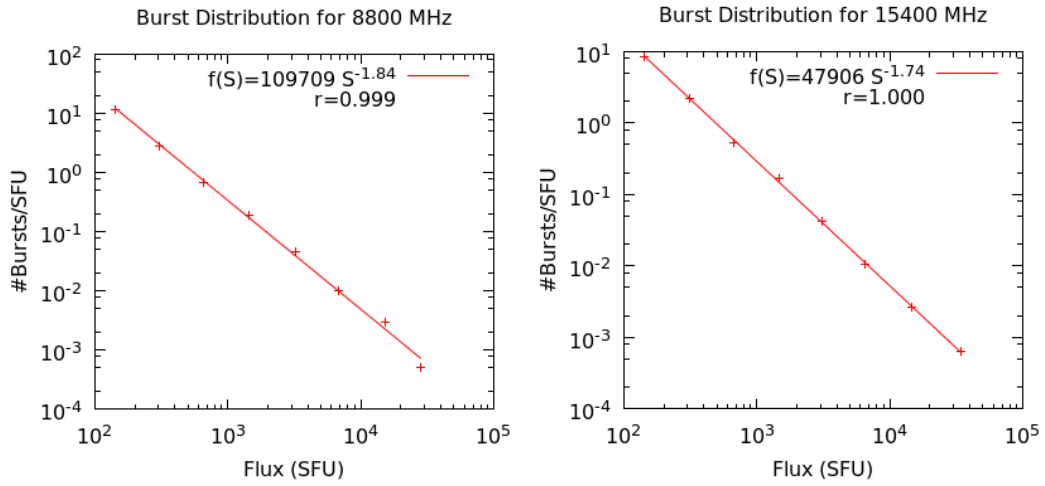



**Figure 3.7 (a-d)**: Burst Distribution Plots for RSTN Frequencies 245 MHz to 1415 MHz




(e) (f)




(g) (h)

**Figure 3.7 (e-h)**: Burst Distribution Plots for RSTN Frequencies 245 MHz to 1415 MHz

These distributions were then normalised to create probability density functions. This procedure, along with a description of power law distributions, can be found in Newman (2005). These are shown in figure 3.8 together with the lines of best fit. In practice these distributions cut-off for frequencies of 2 GHz or higher at around $5\times10^4$ SFU due to synchrotron self absorption. Details of synchrotron self absorption can be found in Kruger (1979). For the lower frequencies this upper cut-off is not entirely known. Two massive bursts observed on 1415 MHz in December 2006 saturated the RSTN receivers. Estimates put these bursts at around $10^6$ SFU (Kintner et al., 2009). However the area

under the curves past these estimated cut-offs is less than 1%, so they do not significantly affect any analysis.

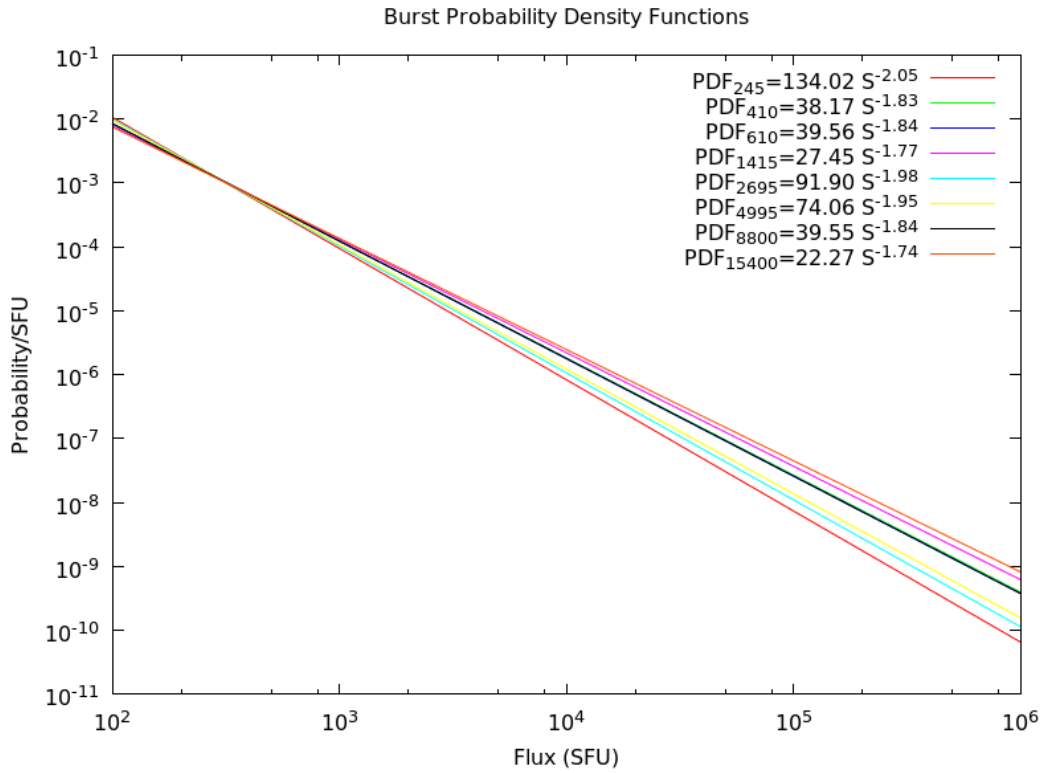



**Figure 3.8**: Burst Probability Density Functions

For any frequency the rate at which bursts will occur in a specific flux range can be calculated by equation 3.6:

$$R = R_0 \int_{S_1}^{S_2} P(S)\,dS \quad , \qquad \text{Eqn 3.6}$$

where $R$ is the burst rate for a burst falling in the flux range $S_1$ to $S_2$, $R_0$ is the burst rate from table 3.5 and $P(S)$ is the appropriate probability density function from figure 3.8.

Since the PDF's all have the form $P(S)=kS^{-a}$, and it is of interest if a burst is greater than a particular value, the expected rate for the flux to exceed a particular value, $S_1$, will be given by:

$$R(S>S_1)=\frac{R_0 k S_1^{1-a}}{a-1} \quad , \qquad \text{Eqn 3.7}$$

Equation 3.7 will be used in the next section with a specific example on interference to GPS.

Some care does need to be taken in using these expressions if the gain pattern of an antenna beam becomes significant (Kraus and Marhefka, 2002). If an antenna is particularly directional, the chances of interference will be significantly smaller than if the antenna is rarely pointing at the Sun. An example of this is satellite TV, where for a month or so each year, it is susceptible to solar interference due to orientation of the satellite TV antenna, geostationary satellites and the sun, but for the remainder of the year less so.

Table 3.6 gives the estimated burst occurrences in bursts per year for various flux values. Burst rates have been omitted for 8800 MHz and 15400 MHz at the $10^5$ SFU level, as it is expected that synchrotron self absorption (see section 3.2.1) will restrict the the peak flux to less than $5\times10^5$ SFU. 2695 MHz and 4995 MHz may also be restricted for the same reason.

| | Burst Rate (Bursts/Decade) | | |
|---|---|---|---|
| Freq (MHz) | S>$10^3$ SFU | S>$10^4$ SFU | S>$10^5$ SFU |
| 245 | 381.13 | 33.97 | 3.03 |
| 410 | 219.62 | 32.48 | 4.80 |
| 610 | 129.18 | 18.67 | 2.70 |
| 1415 | 90.15 | 15.31 | 2.60 |
| 2695 | 53.50 | 5.60 | 0.59 |
| 4995 | 70.12 | 7.87 | 0.88 |
| 8800 | 106.44 | 15.39 | -- |
| 15400 | 106.35 | 19.35 | -- |

**Table 3.6**: Estimated Burst Rates for Solar Radio Flux exceeding $10^3$, $10^4$ and $10^5$ SFU

How consistent are burst rates over solar cycles? To answer this, bursts of greater than 1000 SFU were examined over sunspot cycles 20 to 23 from 1966 to 2010, from the NGDC burst reports. Table 3.7 shows this analysis. From table 3.7, it seems that the burst

rate was significantly lower in cycles 20 and 21 (by more than half in some cases) than cycles 22 and 23. Cycles 22 and 23 are largely comparable.

As will be shown in later chapters, other solar indices were higher in cycles 20 and 21 than in cycles 22 and 23. What is the reason for this significant discrepancy between cycles?

Nita et al. (2002) estimated that 50% of bursts were being 'missed' by the various solar radio observatories. They used the expected time between bursts to derive this figure. They offered no explanation as to why there may be missed bursts.

The answer probably lies in the way the RSTN observatories operate. These observatories are real-time patrol observatories mainly interested in real-time effects on both civilian and military communications and other systems. As a result it is not USAF policy to archive this data. Burst reports are supposed to be sent to NGDC for archiving but this has not always happened.

Up until the mid 1990s, the individual burst reports were mailed from the individual observatories to NGDC. On some occasions, observatories did not mail the results (personal communication, John Kennewell). Commencing in the 1990s data was transmitted automatically to NGDC electronically, probably around 1996 when the Solar Region Summaries started to become available.

John Kennewell, who was the Principal Physicist at LSO from 1981 to 2006, estimates that during the period 1980 to 1982 (cycle 22) there was at least 1 burst per week at 1415 MHz that exceeded 1000 SFU (personal communication). If this estimate is correct, this would give a burst rate of ~50 bursts/year. This was near sunspot maximum, so it would probably be expected, over the entire cycle, that the burst rate would be about half this. This would lead to a burst rate of about double that observed in cycle 23 on 1415 MHz.

Therefore, the bursts were not so much missed by observatories, as they were not archived at NGDC. This does not help current researchers, however, and needs to be taken into consideration when using burst statistics.

| Freq (MHz) | | Cycle 20 | Cycle 21 | Cycle 22 | Cycle 23 |
|---|---|---|---|---|---|
| **245** | Patrol Time (Years) | 3.51 | 8.44 | 9.66 | 12.50 |
| | #Bursts | 108 | 303 | 771 | 915 |
| | Burst Rate (Bursts/Year) | 30.77 | 35.90 | 79.81 | 73.20 |
| | σ | 2.96 | 2.06 | 2.87 | 2.42 |
| **410** | Patrol Time (Years) | 2.45 | 8.44 | 9.32 | 12.08 |
| | #Bursts | 34 | 93 | 256 | 328 |
| | Burst Rate (Bursts/Year) | 13.88 | 11.02 | 27.47 | 27.15 |
| | σ | 2.38 | 1.14 | 1.72 | 1.50 |
| **610** | Patrol Time (Years) | 5.40 | 9.21 | 9.32 | 12.08 |
| | #Bursts | 91 | 76 | 197 | 187 |
| | Burst Rate (Bursts/Year) | 16.85 | 8.25 | 21.14 | 15.48 |
| | σ | 1.77 | 0.95 | 1.51 | 1.13 |
| **1415** | Patrol Time (Years) | 4.94 | 8.44 | 9.66 | 12.50 |
| | #Bursts | 41 | 45 | 76 | 126 |
| | Burst Rate (Bursts/Year) | 8.30 | 5.33 | 7.87 | 10.08 |
| | σ | 1.30 | 0.79 | 0.90 | 0.90 |
| **2695** | Patrol Time (Years) | 4.94 | 8.44 | 9.66 | 12.50 |
| | Patrol Time (Years) | 19 | 36 | 73 | 113 |
| | Burst Rate (Bursts/Year) | 3.85 | 4.27 | 7.56 | 9.04 |
| | σ | 0.88 | 0.71 | 0.88 | 0.85 |
| **4995** | Patrol Time (Years) | 4.94 | 8.44 | 9.66 | 12.50 |
| | #Bursts | 36 | 42 | 95 | 149 |
| | Burst Rate (Bursts/Year) | 7.29 | 4.98 | 9.83 | 11.92 |
| | σ | 1.21 | 0.77 | 1.01 | 0.98 |
| **8800** | Patrol Time (Years) | 4.94 | 8.44 | 9.66 | 12.50 |
| | #Bursts | 53 | 57 | 131 | 199 |
| | Burst Rate (Bursts/Year) | 10.73 | 6.75 | 13.56 | 15.92 |
| | σ | 1.47 | 0.89 | 1.18 | 1.13 |
| **15400** | Patrol Time (Years) | 4.45 | 9.22 | 9.66 | 12.50 |
| | #Bursts | 36 | 55 | 127 | 188 |
| | Burst Rate (Bursts/Year) | 8.09 | 5.97 | 13.15 | 15.04 |
| | σ | 1.35 | 0.80 | 1.17 | 1.10 |

**Table 3.7**: Burst Rates for Bursts>1000 SFU by Sunspot Cycle. σ is the uncertainty in the burst rate (65% confidence).

This should not affect overall burst statistics, provided it can be assumed that the bursts listed in the NGDC database are suitably random and representative of all bursts. It will effect the estimated burst rates however. Unfortunately, there is no real way to correct for these deficiencies.

As a final note, Nita et al. (2002), give power law indices of between -1.7 and -1.9 for burst distributions for burst data between 1960 and 1999). This compares well with the power laws here of between -1.74 and -2.05.

### 3.4 Interference to GPS

#### 3.4.1 Background

This work was first presented at the Australian Space Science Conference (ASSC) meeting in 2012 (Giersch and Kennewell, 2013). The method used to examine the burst reports was slightly revised and thus the estimated frequency of bursts that affect GPS is higher than those reported in that paper.

Ionospheric scintillation is a known problem for Global Navigation Satellite Systems (GNSS) in equatorial regions (Wanninger, 1993). Figure 3.9 shows the scintillation intensity at 2300 local time, when scintillation is likely to be highest (produced from the WBMOD Ionospheric Scintillation Model (North West Research Associates). WBMOD is the WideBand MODel of ionospheric scintillation). The colours represent the S4 scintillation index. The orange and red regions are areas of high scintillation and primarily occur at equatorial latitudes or the poles. At mid-latitudes, scintillation is not significant (the white regions on figure 3.9). The S4 scintillation index is the standard deviation of the received power divided by the mean power of a radio signal (Aarons 1982).

However, in the daylight sector at all latitudes, solar radio bursts have affected the Global Positioning System (GPS) in the past (Kintner et al., 2009). In order to describe these effects, a brief review of GPS is required (Logston, 1992).

GPS uses spread-spectrum modulation which spreads the signal over a wide frequency band (Goldsmith, 2005). At any one frequency the signal lies well below the external noise level. Demodulation of the signal in the receiver is accomplished by correlating the incoming signal with a copy of the digital pseudo-random noise code that was used to

modulate the signal. This increases the signal to noise ratio (S/N) by a factor of around 30 dB, effectively by concentrating the signal in one thousandth of the bandwidth, whilst the noise in the receiver is uncorrelated with the psuedo-code and is not concentrated.

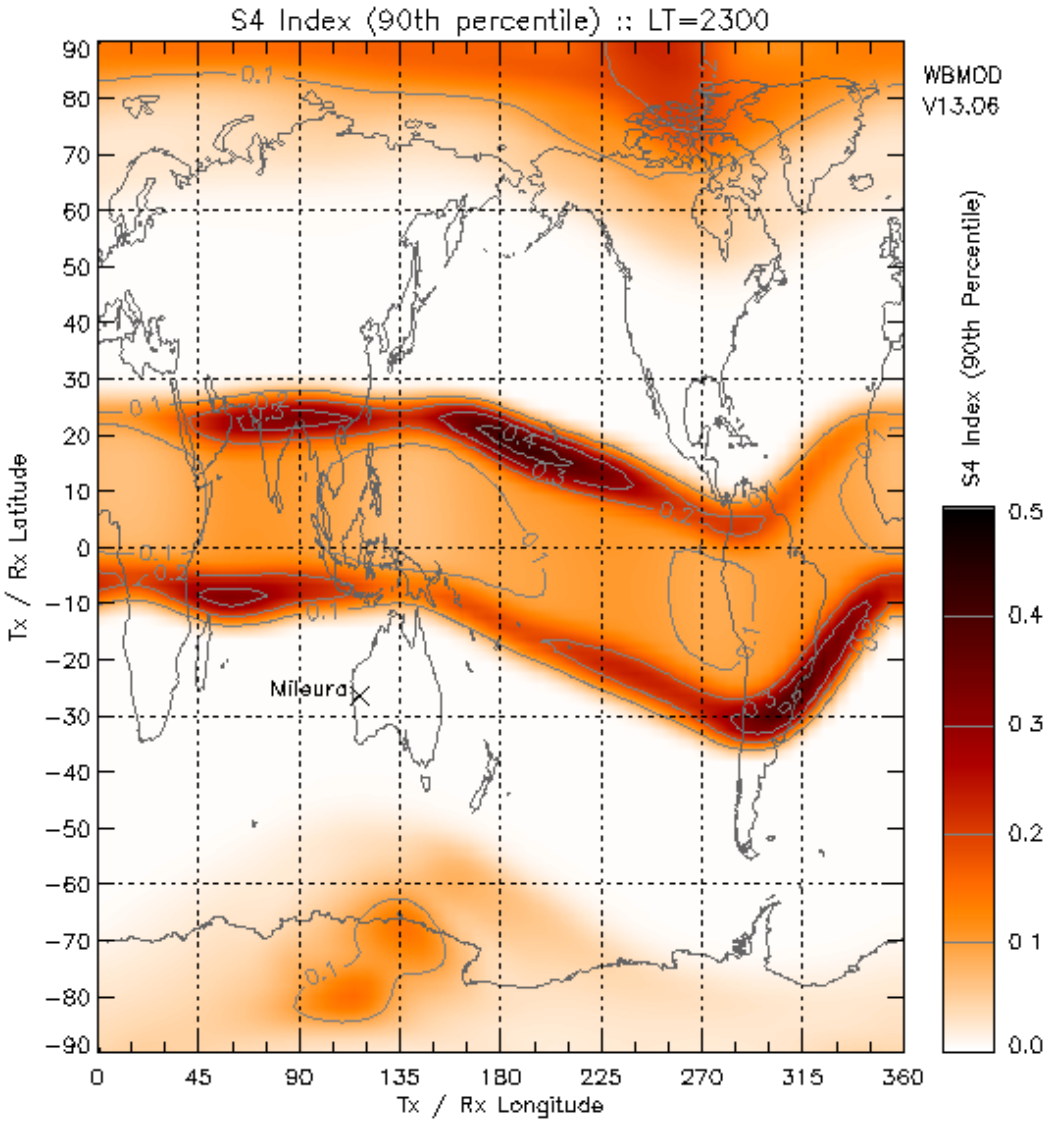



**Figure 3.9**: Global Scintillation for a Frequency of 1575 MHz (GPS L1), a SSN=150 and Kp=1 (an index of geomagnetic activity). The day of the year was 91 with a local time (globally) of 2300 hours. The plot shows maximum equatorial ionisation.

Figure 3.10 shows a sample GPS signal (http://www.insidegnss.com/node/2138). On the right vertical axis, solar radio flux in terms of Solar Flux Units has been added. It is interesting to note that the quiet sun (when no bursts are in progress) around 1400 MHz

near sunspot maximum is about 100 SFU. This is the peak level of the GPS spread signal. Receiver demodulation effectively raises this signal by 30 dB (a thousand-fold increase), giving an equivalent solar flux density of $10^5$ SFU. It would be expected that a solar burst exceeding this level to cause problems to the GPS signal.

It is noted that the quiet Sun flux at 1175 MHz will be slightly lower, and the flux at 1575 MHz slightly higher, than that at 1415 MHz. However, the quiet Sun flux at GPS frequencies will be within 10% of the flux at 1415 MHz, making 1415 MHz a good frequency to examine for effects on GPS.

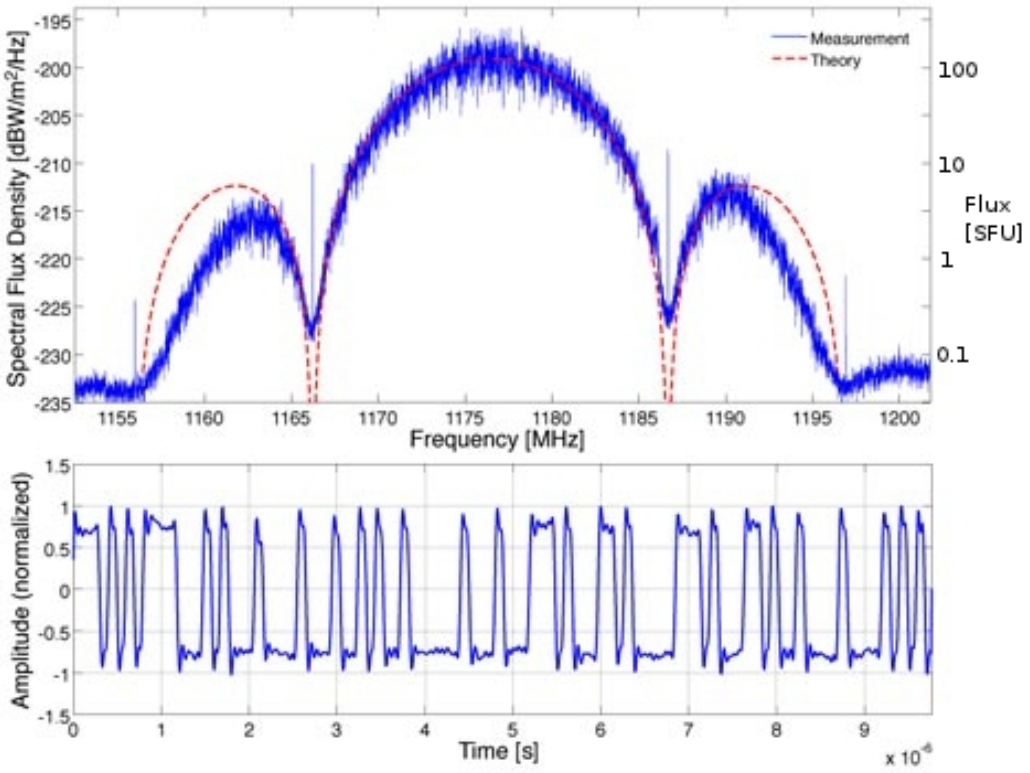



**Figure 3.10**: Sample GPS Frequency Spectrum. The spectral flux density (SFD = 10 $\log_{10}$(S) , where S is the flux in W/m$^2$/Hz) is on the left vertical axis and Solar Radio Flux (SFU) is on the right vertical axis. The lower panel is an example of the psuedo-code used for correlation.

Some more formal estimates of when GPS might suffer degradation have been made. Chen et al. (2005) estimated GPS degradation to occur at solar burst flux densities from 4 to 10 kSFU for uncorrelated GPS. Klobuchar et al. (1999) estimated 40 kSFU and Kennewell (http://www.sws.bom.gov.au/Educational/1/3/10) estimated 100 kSFU.

### 3.4.2 Solar Radio Interference to GNSS

On 6 December 2006 and again on 13 December 2006 two large bursts that possibly exceeded 1 MSFU were observed (Kintner et al., 2009). The RSTN 1415 MHz radiometers were saturated by these bursts (figure 3.11), as were most other solar radiometers around the world (Kintner et al., 2009). The GNSS became unusable for up to 10 minutes in some locations and GPS lock became impossible. Kintner et al. (2009). discusses these bursts in some detail. Solar radio bursts of this magnitude should not have come as a complete surprise. On 29 April 1973 a radio burst of 169 kSFU was observed at Sagamore Hill on 1415 MHz. This was found in the NGDC solar radio burst listings. It has possibly been overlooked as there was no GNSS to suffer interference in 1973. However, bursts that exceed 100 kSFU on 1415 MHz should be considered as rare and extreme events. Figure 3.11 shows the radio burst recorded at LSO on 13 December 2006. This lasted for about 2 hours. The flat tops on the peaks are where the radiometer saturated. The total time the burst exceeded 100 kSFU is quite small, about 10 minutes. The data for this plot was obtained from NGDC.

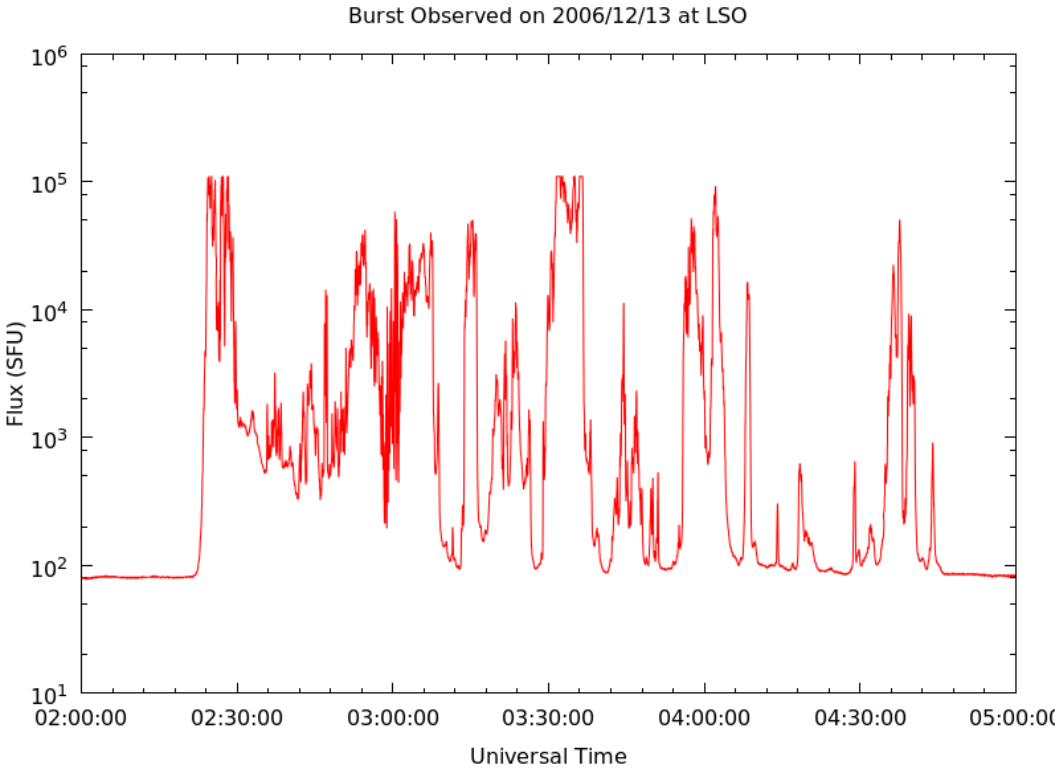



**Figure 3.11**: 1415 MHz Solar Radio Burst Recorded at Learmonth on 13 December 2006

### 3.4.3 1415 MHz Burst Statistics

As described in section 3.3, the PDF for bursts on 1415 MHz was found. This is shown in equation 3.8:

$$P(S) = 27.45\, S^{-1.77} \quad , \qquad \text{Eqn 3.8}$$

Substituting the appropriate values for 1415MHz into equation 3.7, the expression for the burst rate for bursts greater than a flux value of $S_1$ can be obtained:

$$R(S > S_1) = 1840.58\, S_1^{-0.77} \quad , \qquad \text{Eqn 3.9}$$

Using equation 3.9 the rate at which bursts will exceed a specified threshold can be obtained. This has been done for $S_1$ = 100 kSFU (SNR = 0, when GPS lock may not be possible) and $S_1$ = 1000 kSFU (SNR = -10, when GPS may become unusable). The estimated burst frequency is found to be 2.6 bursts/decade and 0.44 bursts/decade respectively. These numbers may be on the low side because of assumed interference thresholds at the high end of published speculation. Some GNSS systems may well be affected by lower levels of solar radio emission.

## 3.5 Summary and Conclusions

The main mechanisms for solar radio bursts are synchrotron, bremsstrahlung, plasma and cyclotron emission.

The USAF manages the Solar Radio Telescope Network that operates at 8 frequencies between 245 MHz and 15400 MHz. The consistency of measurement varies somewhat across the network.

For quiet Sun (noon flux) values, at lower frequencies (less than 1 GHz) the inter-site variation is primarily due to noise storms, which are long duration rapidly varying phenomena. At higher frequencies inter-site variations are still present, frequently outside the AFWA guidelines. These variations are probably due to calibration, equipment or environmental factors (temperature and humidity).

For bursts, variations between sites can be significant. On average, most sites did not

measure the same bursts as having the same values (within 2 standard deviations). This is significant when trying to compile the data into a single burst list.

Solar radio bursts at a given frequency follow a power law distribution. The coefficients are different for each frequency. Since observing period for each frequency is known, rates can be estimated for bursts of a particular magnitude. However, up to half the bursts prior to 1996 are not listed in the NGDC archives. Hence, burst rates computed using these data will be underestimates of the true burst rate.

Bursts of sufficient intensity to reach a GPS systems interference threshold are estimated to occur approximately 2.6 times per decade, and those to make GPS unusable approximately 0.44 times per decade (that is about once every 20 years).

# 4. Solar Variation

## 4.1 Introduction

There are many different solar cycles, both in terms of solar properties as well as duration. The one most commonly referred to is the approximately 11 year sunspot number cycle (Zirin, 1989). This is often, somewhat misleadingly, called the solar cycle.

In this chapter variations of sunspot number, sunspot area and 10.7 cm radio flux are examined in order to determine if these variations are caused by intrinsic changes in the sun, or variations in analysis over the decades that these parameters have been observed. The most commonly cited reasons for discrepancies over long time periods is change of directors of observatories or change of agency monitoring the particular parameters (for example: Wilson and Hathaway, 2005; Wilson and Hathaway, 2006; Foukal 2014).

The specific data sources used are:

- International Sunspot Number (ISN): http://sidc.oma.be/silso/versionarchive
- American Association of Variable Star Observers (AAVSO) sunspot number: ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-indices/sunspot-numbers/american/lists/list_aavso-arssn_monthly.txt
- Solar Region Summary sunspot number and area (SRS): ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_REGIONS/USAF_MWL/
- Greenwich sunspot area: ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_REGIONS/Greenwich/
- Rome sunspot area: ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_REGIONS/Rome/
- Debrecen sunspot area: ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_REGIONS/Debrecen/

All of the following USAF data was obtained from ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SUNSPOT_REGIONS/USAF_MWL/ .

- US-AFWA sunspot number and area
- Learmonth (LEAR) sunspot number and area
- San Vito (SVTO) sunspot number and area
- Holloman (HOLL) sunspot number and area

- Ramey (RAMY) sunspot number and area
- Palehua (PALE) sunspot number and area

The AFWA values were calculated by averaging the USAF sites (Learmonth, San Vito, Holloman, Ramey and Palehua). The Solar Region Summary data was taken from the daily reports produced by SWPC and 12 month smoothed means generated.

Finally the Penticton 10.7 cm radio flux was obtained from ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton_averaged/penticton_solflux-monthly-average.txt

Unless otherwise stated, all of the parameters discussed in this chapter are 12 month smoothed means (derived as shown in chapter 2).

Some of the data (primarily from USAF-AFWA sites, Rome and Debrecen) were missing or omitted. If a month had less than 5 days of observations from a particular site then it was excluded. If a month of data was missing in the values used in the 12 month running mean, then no running mean was calculated. Thus there are gaps of 12 months or more for some sites.

In the case of the USAF AFWA site data, there are no reports if there were no spots on the Sun, or no observations could be made due to weather or equipment problems. Thus for USAF data there are gaps around solar minimum during low sunspot counts as it was impossible to determine if a site was missing data or there were no spots.

**4.2 Sunspot Number Variations**

In July 2015, the International Sunspot Number data set for the last 400 years was revised (Clete et al., 2014). This new set was derived because of perceived problems with the way the old data was computed. These recalculations were done after four Sunspot Number meetings between 2010 and 2014. Some of the reasons Clete et al. (2014) listed for this recalculation are:

- The Locarno station weighting of sunspots based on their size and development of penumbra.
- A significant discontinuity in the data when Max Waldemier became the Zurich

Observatory Director in 1945.

- A significant discontinuity in the data when sunspot measurements moved from Zurich to Brussels.
- Removal of the 0.6 constant applied to the International Sunspot Number to scale the data to Wolf's original observations (Wolf used a small telescope for his observations).

In this section, relationships are analysed between various sites' sunspot number reports in an attempt to see how sunspot number observations have changed over time from site to site. For the international sunspot number, the old data set (pre 2015) is used, as most of the analysis was done prior to the new data set being generated.

Figure 4.1 shows the sunspot numbers calculated from various sites from 1945 to 2014. AAVSO acknowledge that the data prior to 1950 maybe suspect due to rescaling around this time.

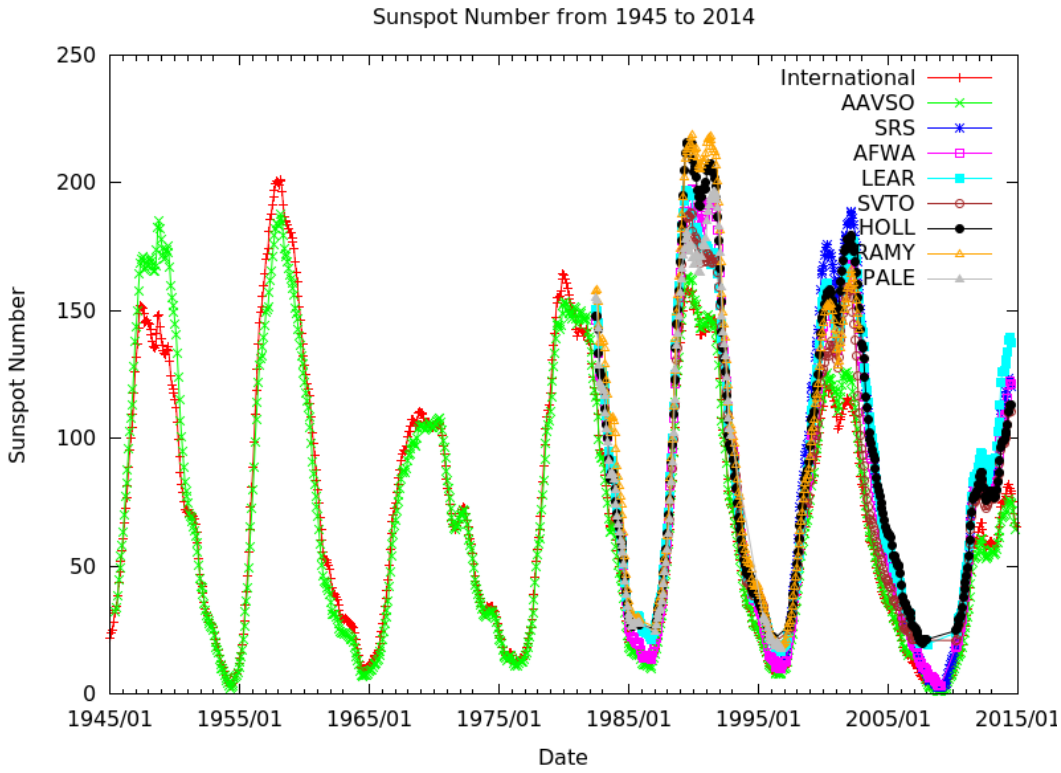



**Figure 4.1**: Smoothed Monthly Mean Sunspot Numbers from Various Sites and Institutions from 1945 to 2014, Spanning Cycles 18 to 24.

Figure 4.1 shows that the AAVSO sunspot number tracks with the International Sunspot Number very well, although there are some variations. The individual USAF-AFWA sites sunspot number are higher than the ISN as no scaling constant was applied for them, also causing the sites to differ from each other. These differences are seen more clearly in Figure 4.2 which spans the shorter period from 1982 to 2014.

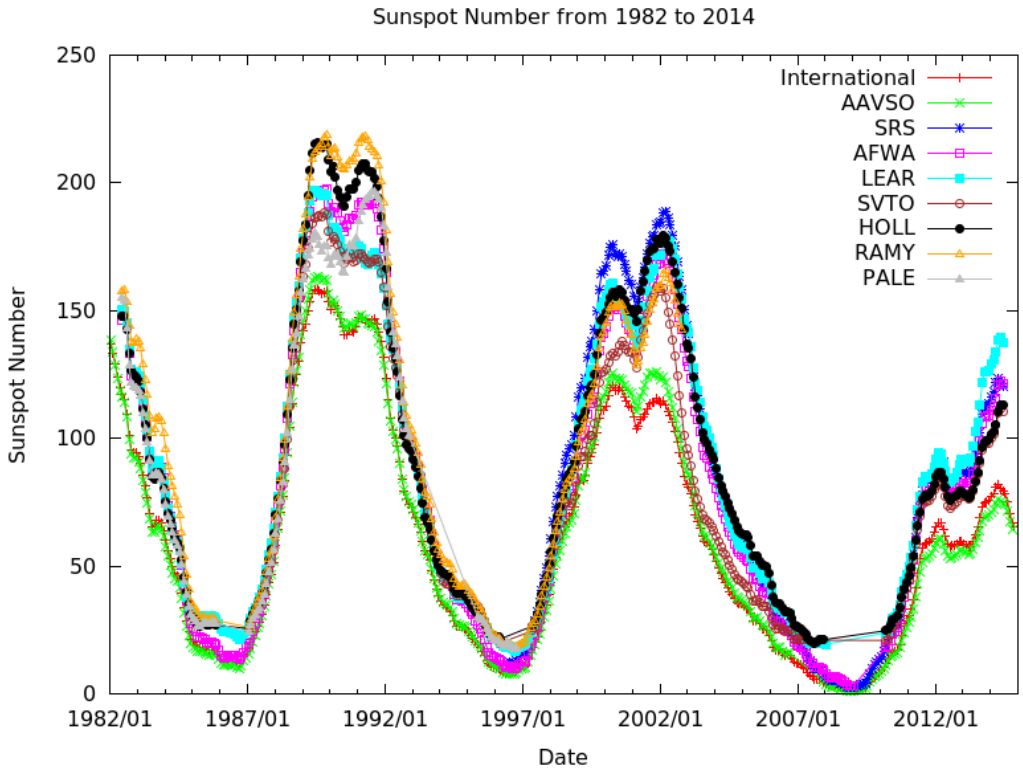



**Figure 4.2**: Smoothed Monthly Mean Sunspot Numbers from Various Sites and Institutions from 1982 to 2014, Spanning Cycles 21 to 24

The correlations between the different measurements shown in figures 4.3 to 4.7 where the relationship between the International, AAVSO, AFWA and SRS sunspot numbers are compared by sunspot cycle. Tables 4.1 to 4.5 below each figure show the regression statistics for each fit, where m is the gradient, $s_m$ is the standard deviation in the gradient, b is the intercept, $s_b$ is the standard deviation of the intercept and r is the correlation coefficient.

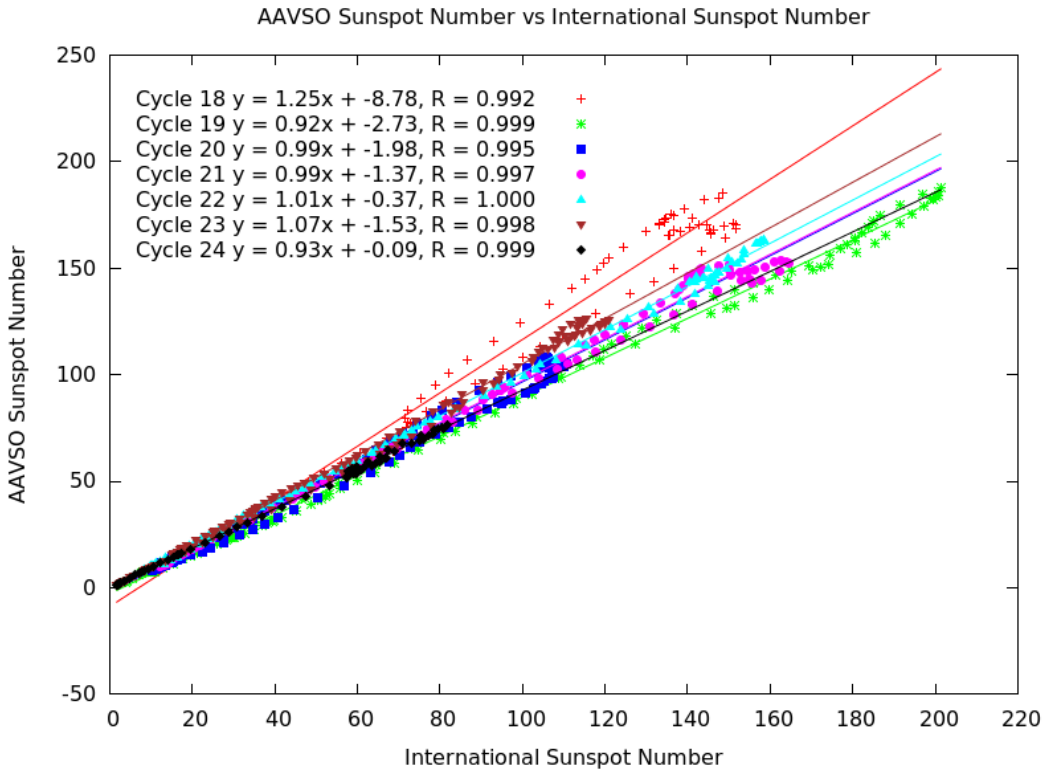



**Figure 4.3**: AAVSO Sunspot Number vs International Sunspot Number vs AAVSO for Cycles 18 to 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 18 | 1.253 | 0.016 | -8.778 | 1.503 | 0.992 |
| 19 | 0.922 | 0.004 | -2.733 | 0.423 | 0.999 |
| 20 | 0.986 | 0.008 | -1.984 | 0.571 | 0.995 |
| 21 | 0.986 | 0.007 | -1.366 | 0.630 | 0.997 |
| 22 | 1.013 | 0.003 | -0.367 | 0.289 | 1.000 |
| 23 | 1.066 | 0.005 | -1.527 | 0.368 | 0.998 |
| 24 | 0.928 | 0.004 | -0.090 | 0.207 | 0.999 |

**Table 4.1**: AAVSO vs International Sunspot Number Regression Parameters

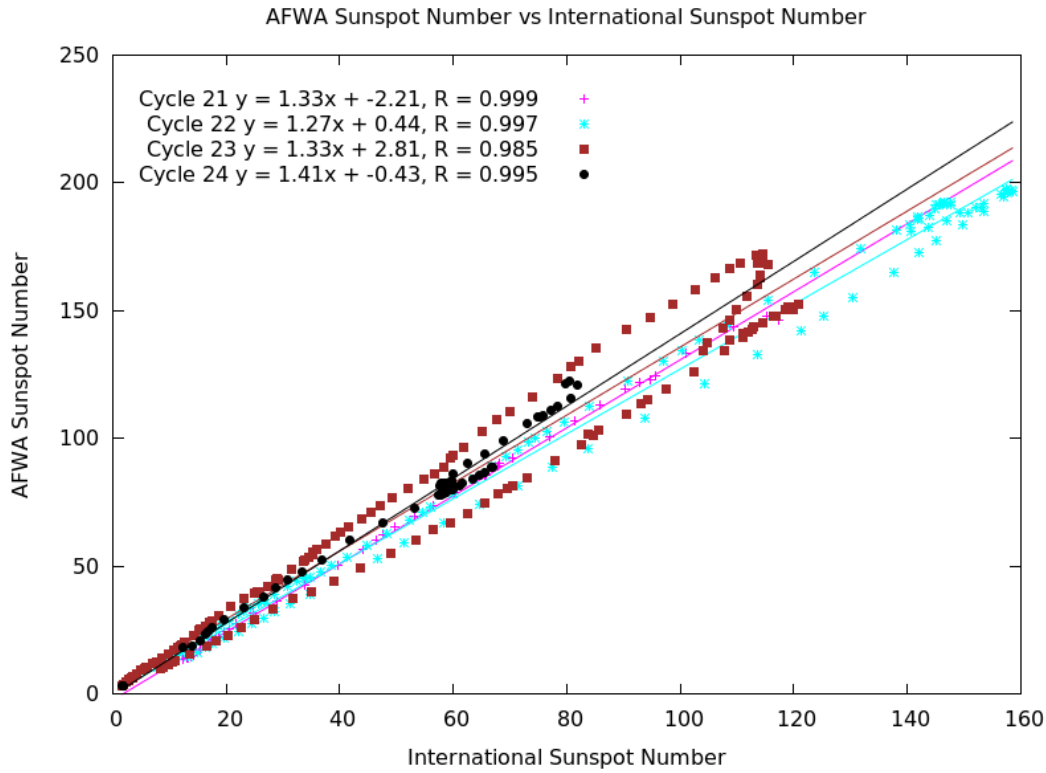



**Figure 4.4**: AFWA Sunspot Number vs International Sunspot Number for Cycles 21 to 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 21 | 1.330 | 0.006 | -2.207 | 0.357 | 0.999 |
| 22 | 1.267 | 0.009 | 0.442 | 0.887 | 0.997 |
| 23 | 1.330 | 0.019 | 2.814 | 1.318 | 0.985 |
| 24 | 1.415 | 0.020 | -0.431 | 1.149 | 0.995 |

**Table 4.2**: AFWA vs International Sunspot Number Regression Parameters

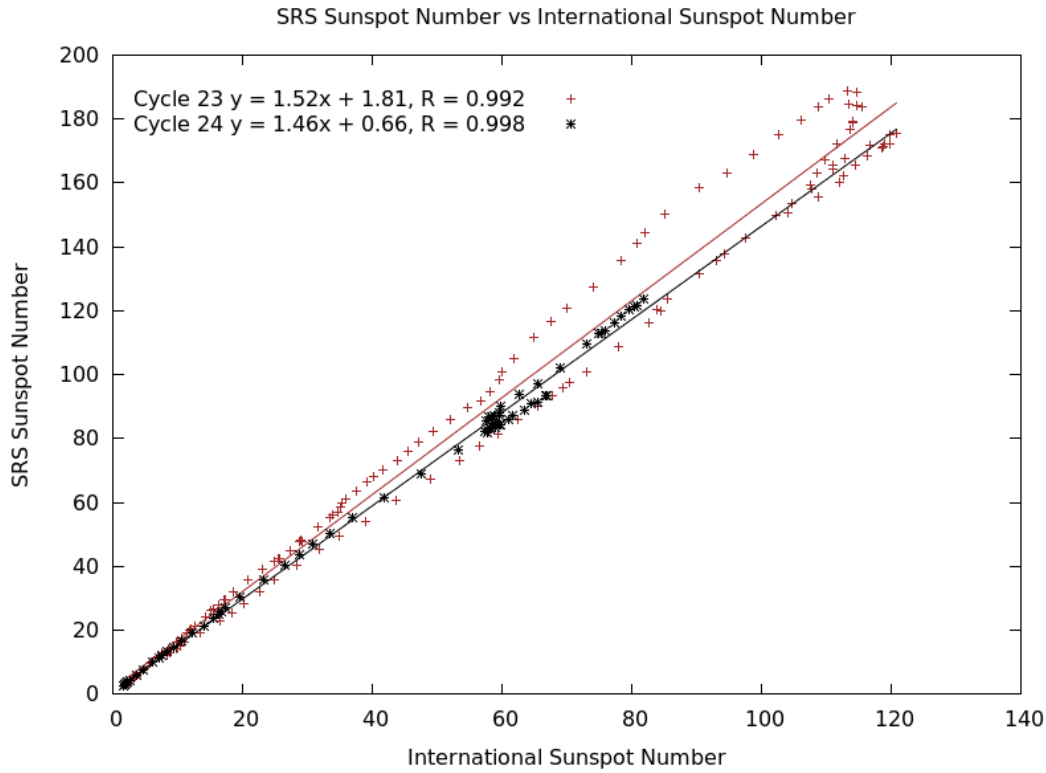



**Figure 4.5**: Solar Region Summary Sunspot Number vs International Sunspot Number for Cycles 23 and 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 23 | 1.516 | 0.016 | 1.808 | 1.073 | 0.992 |
| 24 | 1.458 | 0.010 | 0.664 | 0.505 | 0.998 |

**Table 4.3**: Solar Region Summary vs International Sunspot Number Regression Parameters

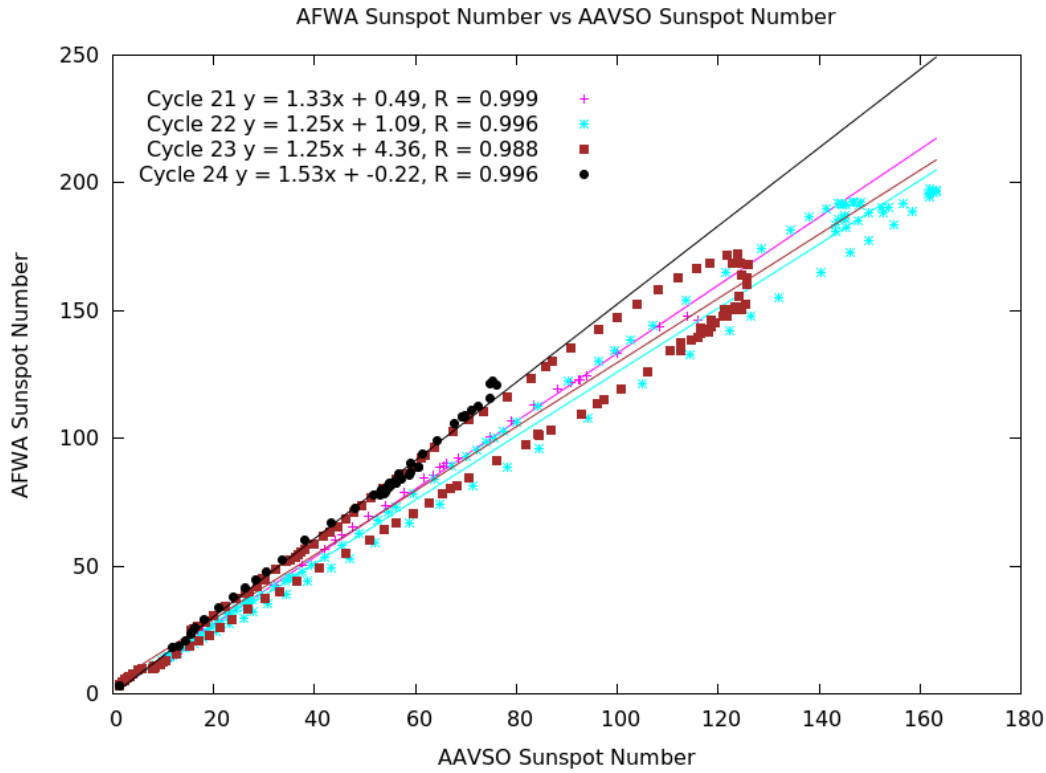



**Figure 4.6**: AFWA Sunspot Number vs AAVSO Sunspot Number for Cycles 21 to 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 21 | 1.329 | 0.007 | 0.491 | 0.415 | 0.999 |
| 22 | 1.249 | 0.011 | 1.088 | 1.065 | 0.996 |
| 23 | 1.254 | 0.017 | 4.360 | 1.251 | 0.988 |
| 24 | 1.528 | 0.018 | -0.217 | 0.944 | 0.996 |

**Table 4.4**: AFWA vs AAVSO Sunspot Number Regression Parameters

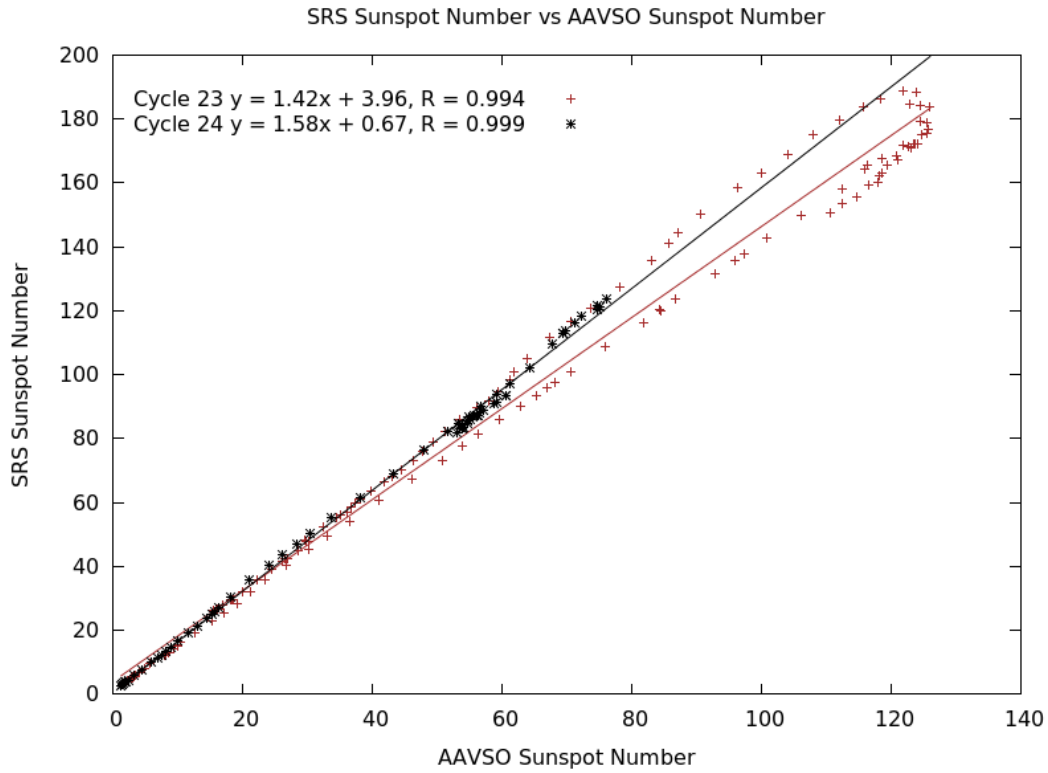



**Figure 4.7**: Solar Region Summary Sunspot Number vs AAVSO Sunspot Number for Cycles 23 and 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 23 | 1.424 | 0.014 | 3.962 | 1.024 | 0.994 |
| 24 | 1.578 | 0.008 | 0.674 | 0.384 | 0.999 |

**Table 4.5**: Solar Region Summary vs AAVSO Sunspot Number Regression Parameters

Two things of importance are revealed in the above graphs. The relationships between sites shown in figures 4.3 to 4.7 (i) changes between cycles and (ii) the relationship changes at solar maximum. That is the the relationships between two given sites show hysteresis between the declining phase of a cycle and the rising phase of a cycle. This variation is is seen in all observations where complete cycles have been monitored.

The variation in the relationships between cycles is better seen in the fit parameters in tables 4.1 to 4.5. There are only 4 cases where the gradients of the site comparisons across sunspot cycles are within 2 standard deviations of each other (as described in

section 1.7.1). These are shown in table 4.6.

| Sites | Cycles | Intercept Similar |
|---|---|---|
| AFWA vs ISN (Table 4.2) | 21 and 23 | No |
| AFWA vs AAVSO (Table 4.4) | 22 and 23 | Yes |
| AAVSO vs ISN (Table 4.1) | 19 and 24 | No |
| AAVSO vs ISN (Table 4.1) | 20 and 21 | Yes |

**Table 4.6**: List of Sunspot Number Site Comparisons with Statistically Similar Gradients (within 2 standard deviations). The third column states whether there the intercepts were statistically similar.

Therefore most of the sites relationships vary over time scales of less than a sunspot cycle. Why should this be the case? After all, the sites are measuring the same features. Certainly sunspots change in the course of a 24 hour period, but as most of these agencies incorporate more than one site over a wide range of longitudes, these differences should be averaged out. The *k* scaling factors for observers and sites are intended to overcome the differences in equipment, conditions and observers. It is possible that *k* values are not being updated regularly enough.

One possible solution, expensive though it would be, is for each agency to deploy identical equipment to different observatories, minimising the effects of different optics and recording techniques. The USAF already does this with their SOON sites, but between these sites there are differences as seen in figure 4.2. It is noted here that the USAF have rapid staff turnover at most of their sites (the exception being San Vito, which is run by contractors). An analyst is typically assigned to a solar observatory for between 15 and 24 months. However, rotation of personnel is constant and on average should not affect techniques used if constant procedures are adhered to.

In chapter seven another approach is taken, that of removing human observers and automating the analysis. This approach is recommended to improve consistency and will be discussed further in that chapter.

### 4.3 Sunspot Area Variations

It was recognised by Giovanelli (1939) that different sites were measuring the same sunspots but with different areas. This was a problem as Giovanelli was looking to use sunspot area as a predictor for Hα flares. Table 4.7 shows the area correction factors that were used to correct back to Greenwich (RGO) values.

| Observatory | Correction Factor |
|---|---|
| U.S. Naval (1935 - 1936) | 1.37 |
| U.S. Naval (1937) | 0.875 |
| Mount Wilson | 1.10 |
| Harvard | 1.01 |

**Table 4.7**: Giovanelli Sunspot Area Correction Factors

Of note is that the US Naval Observatory seems to have adjusted their measurements between 1936 and 1937 for unknown reasons. So, sunspot area variations between sites are not a new problem. It is generally assumed that the Greenwich (RGO) values are the best, as many authors seem to use it as a standard (for example, Wilson and Hathaway, 2005). This may just be due to it being the longest running data set.

Wilson and Hathaway (2005) examined the sunspot number to sunspot area ratio for various sites in order to determine if the USAF have been underestimating the area from 1980 onward, around the time when Greenwich (RGO) ceased sunspot area measurements. Wilson and Hathaway (2005) showed that the Greenwich observatory sunspot areas were 12% higher than Rome values, and that USAF areas were 12.7% lower than Rome's, therefore USAF were measuring areas that were approximately 22% lower than Greenwich (Wilson and Hathaway, 2005). An underlying assumption is that each of the observatories sunspot measurements are consistent over time.

Comparisons between Debrecen's Heliophysical Observatory and Greenwich's areas have also been made (Baranyi et al., 2013). The relationship is shown in equation 4.1.

$$A_G = 1.08(\pm 0.11) A_D \quad , \qquad \text{Eqn 4.1}$$

where $A_G$ is the Greenwich area and $A_D$ is the Debrecen area.

Foukal (2014) suggested reasons for the reported underestimate of USAF sunspot areas, including small spots were not counted in USAF's areas. He makes the statement:

> *"This range of g implies that our explanation is reasonable if the SOON program essentially neglected the areas of spots smaller than approximately 10 μh."*

Foukal's claim in not strictly correct, as small spots measured by USAF are assigned a value of 2 μh. This may indicate spots in the range 2 to 10 μh may be assigned either value. This will be analyst dependent.

The remainder of this section will investigate the relationships between area measurements at various sites. Again, the smoothed monthly means are used to remove daily variations in observations and minimise analyst bias.

Figure 4.8 shows sunspot areas measured from 1945 to 2014 over solar cycles 18 to 24. Figure 4.9 shows sunspot areas from 1982 to 2014 over cycles 21 to 24. From a visual inspection of these sites it can be seen that in different sunspot cycles the relationships between the sites change. To quantify this variation, the relationships between sites was plotted for individual cycles. These plots are shown in figures 4.10 to 4.14. The regression parameters are shown below each plot in tables 4.8 to 4.12.

Of particular note are figures 4.11 and 4.12. The cycle 23 relationship becomes distinctly non-linear. As this occurs for both Debrecen and AFWA data, the problem seems to be with Rome data. This illustrates that even within a given cycle observations can be inconsistent.

However despite some weaker correlations, there is a trend evident. AFWA and SRS (which is derived in part from AFWA data) values are consistently lower than other sites. AFWA derived sunspot area values are between 17 % and 63 % lower than the other observing institutions, depending on site and cycle.

In the sunspot area plots there is less hysteresis, that is the sunspot area measurements are more similar between the rise and fall of the cycle than in sunspot numbers. As can be seen from the regression tables (tables 4.8 to 4.12), the gradients are generally different between cycles, the one exception being Debrecen vs Rome for cycles 22 and 23.

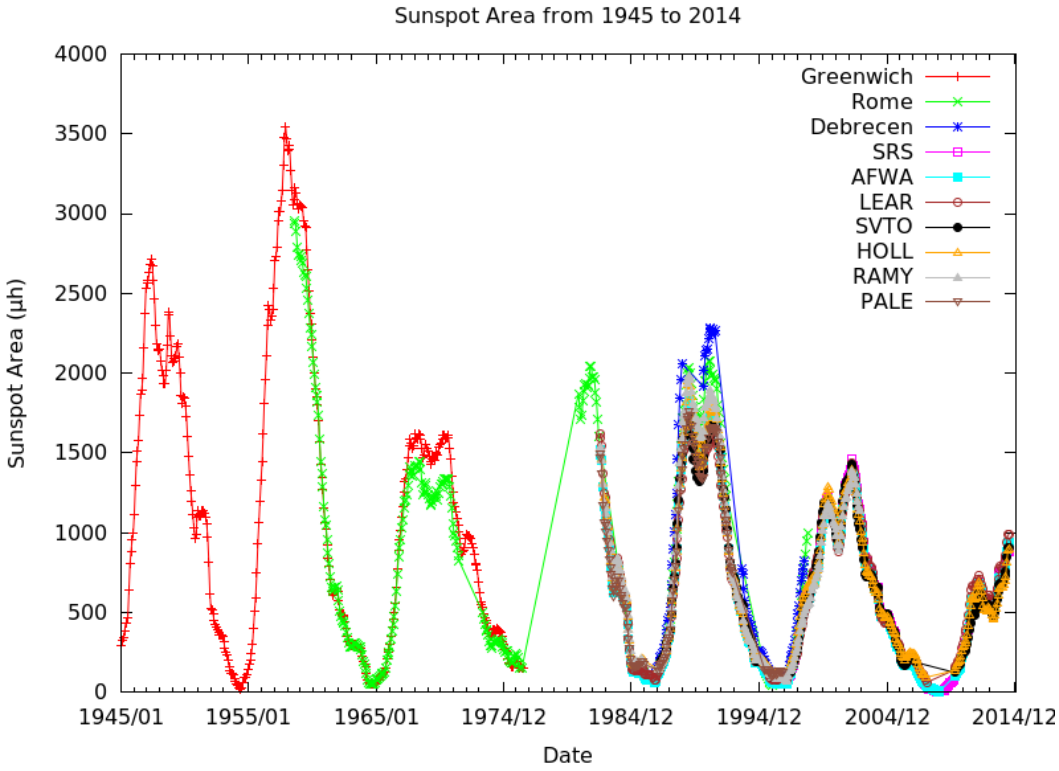



**Figure 4.8**: Smoothed Monthly Mean Sunspot Areas from Various Sites and Institutions from 1945 to 2014

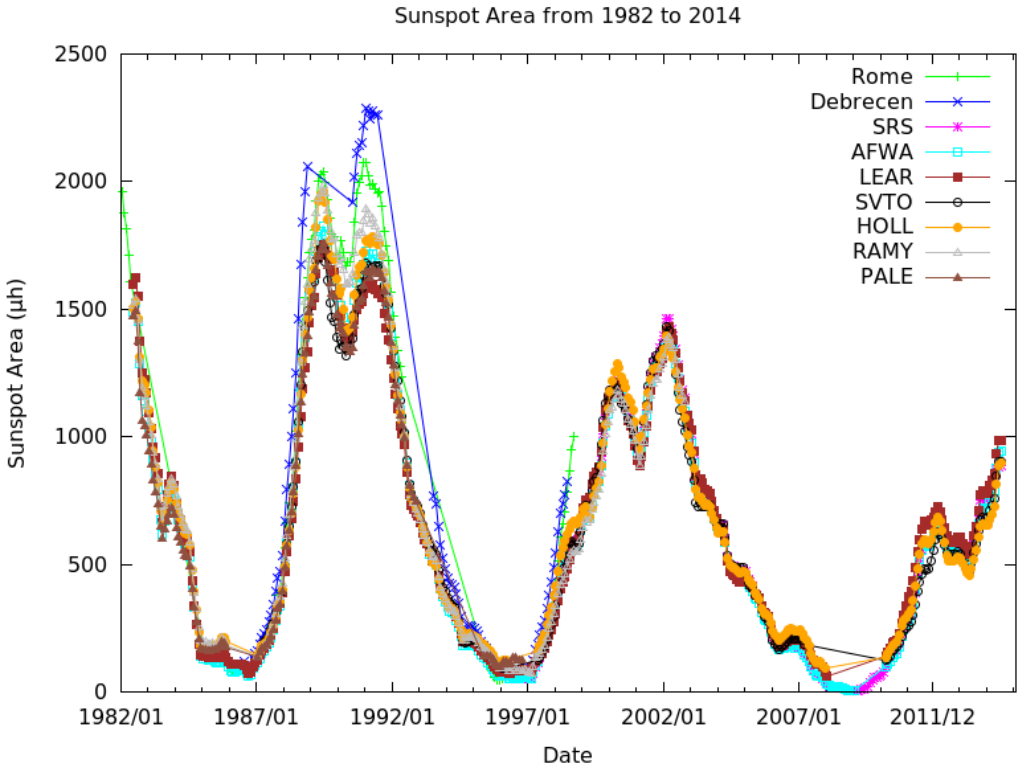



**Figure 4.9**: Smoothed Monthly Mean Sunspot Areas from Various Sites and Institutions from 1982 to 2014

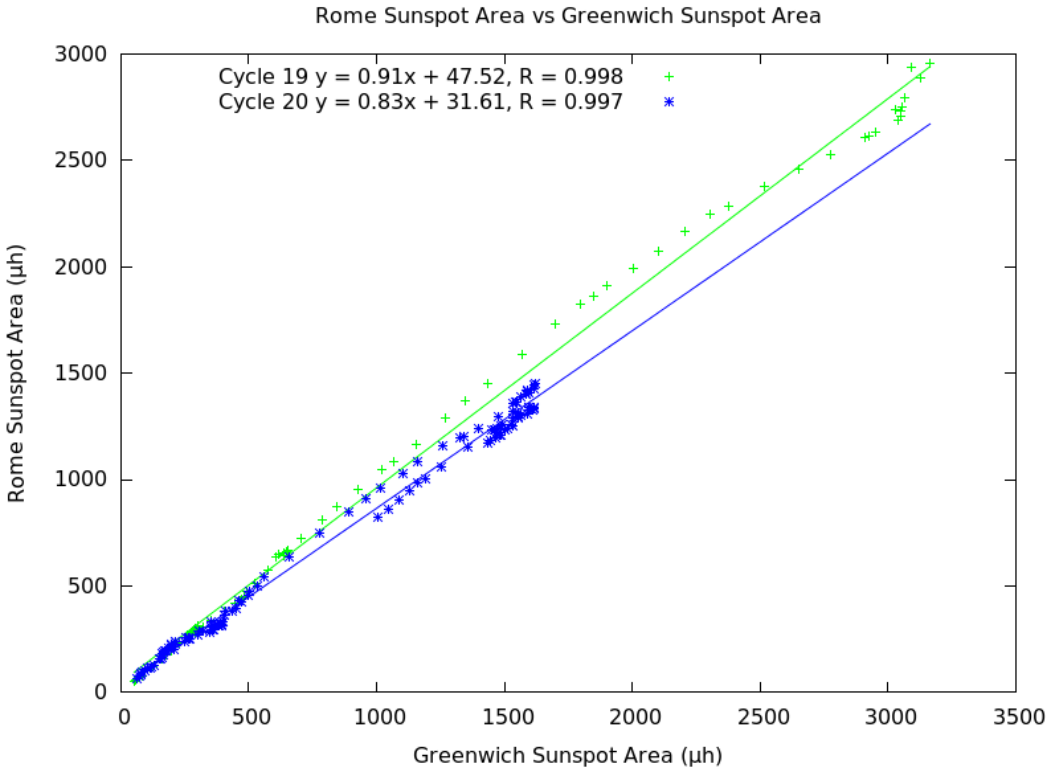



**Figure 4.10**: Rome Sunspot Area vs Greenwich Sunspot Area for Cycles 19 and 20

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 19 | 0.914 | 0.007 | 47.521 | 11.045 | 0.998 |
| 20 | 0.834 | 0.006 | 31.612 | 6.186 | 0.997 |

**Table 4.8**: Rome vs Greenwich Sunspot Area Regression Parameters

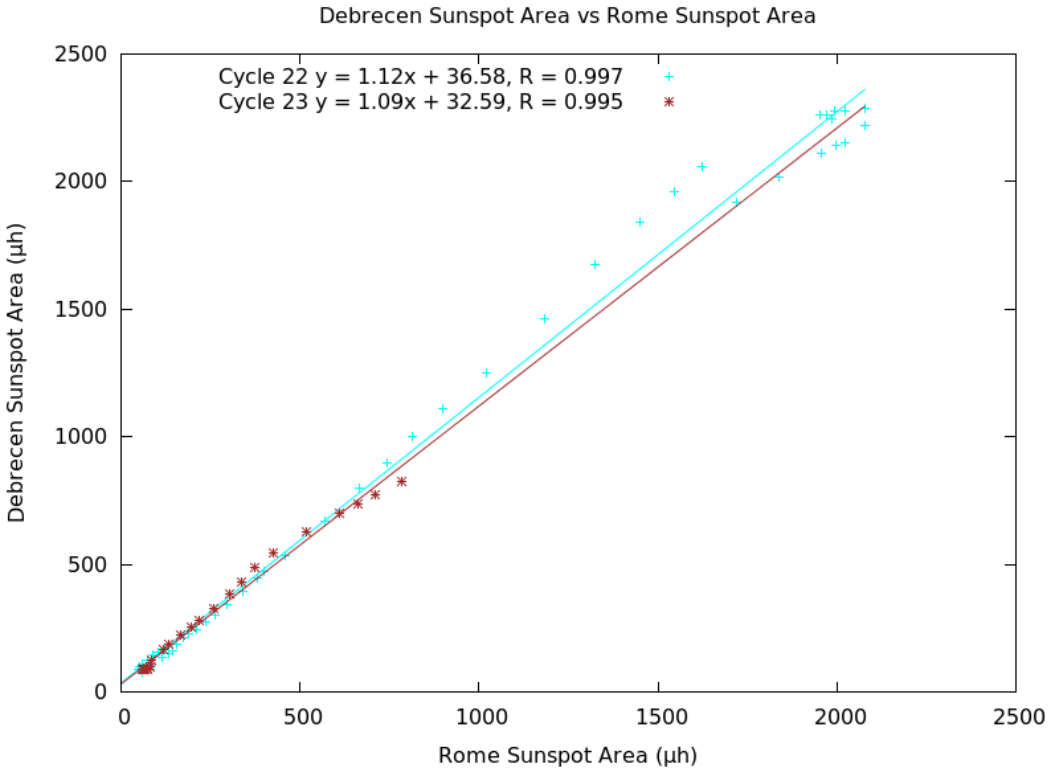



**Figure 4.11**: Debrecen Sunspot Area vs Rome Sunspot Area for Cycles 22 and 23

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 22 | 1.118 | 0.013 | 36.580 | 14.980 | 0.997 |
| 23 | 1.089 | 0.024 | 32.593 | 8.173 | 0.995 |

**Table 4.9**: Debrecen vs Rome Sunspot Area Regression Parameters

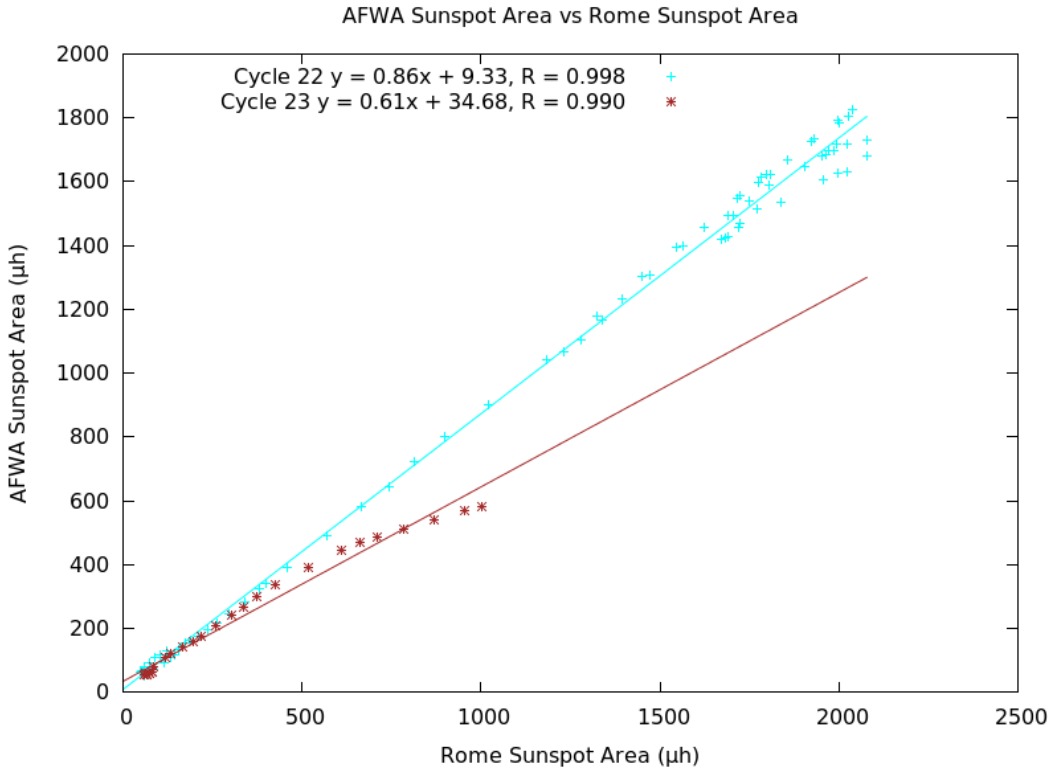



**Figure 4.12**: AFWA Sunspot Area vs Rome Sunspot Area for Cycles 22 and 23

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 22 | 0.864 | 0.006 | 9.334 | 8.177 | 0.998 |
| 23 | 0.609 | 0.017 | 34.677 | 7.796 | 0.990 |

**Table 4.10**: AFWA vs Rome Sunspot Area Regression Parameters

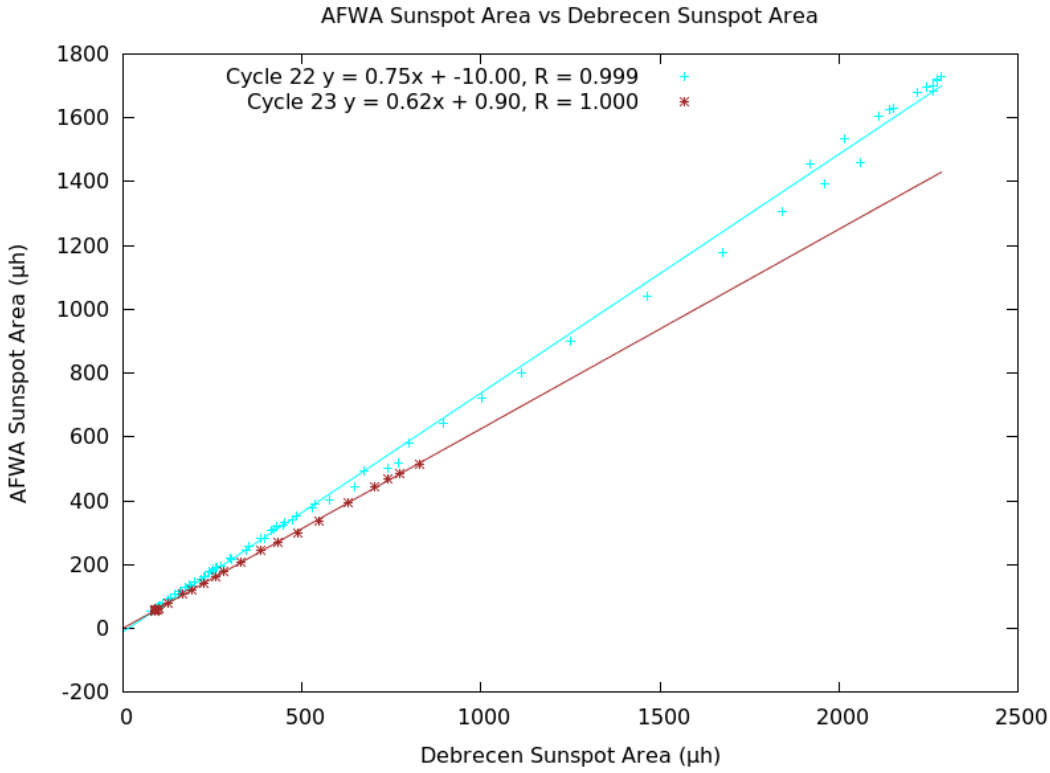



**Figure 4.13**: AFWA Sunspot Area vs Debrecen Sunspot Area for Cycles 22 and 23

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 22 | 0.748 | 0.003 | -10.001 | 3.741 | 0.999 |
| 23 | 0.625 | 0.002 | 0.903 | 0.977 | 1.000 |

**Table 4.11**: AFWA vs Debrecen Sunspot Area Regression Parameters

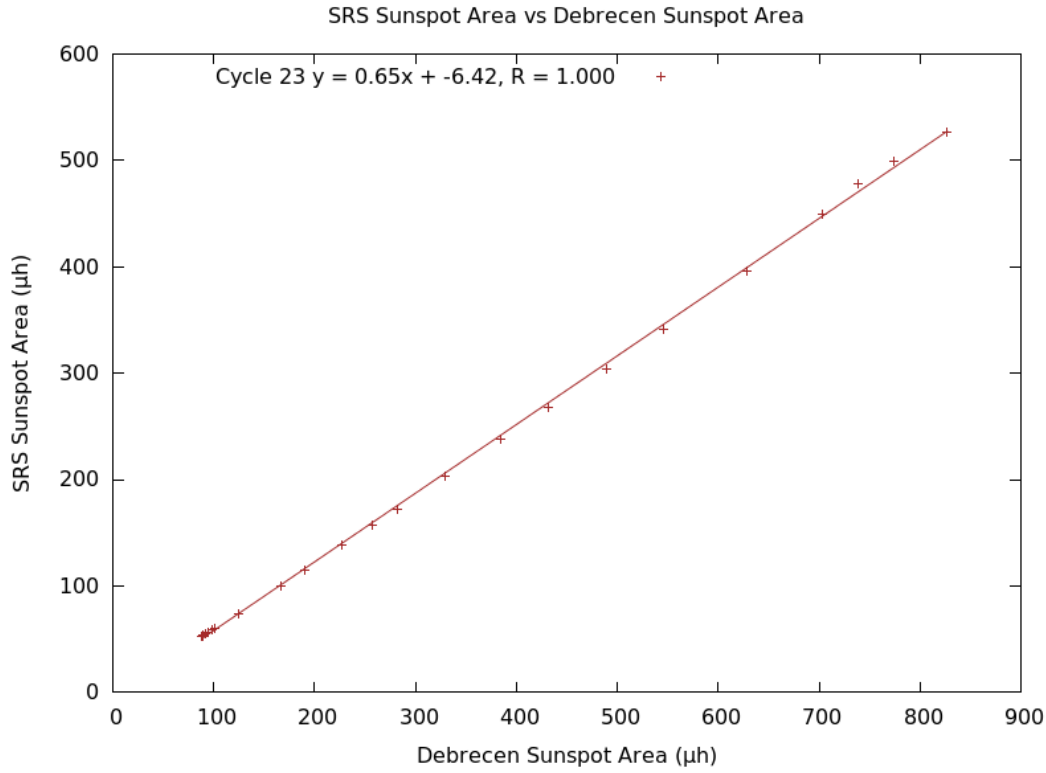



**Figure 4.14**: Solar Region Summary Sunspot Area vs Debrecen Sunspot Area vs for Cycle 23

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 23 | 0.645 | 0.003 | -6.424 | 1.210 | 1.000 |

**Table 4.12**: Solar Region Summary vs Debrecen Sunspot Area Regression Comparison

The USAF AFWA give their analysts instructions as to how sunspot drawings are to be performed and measured (USAF AFWAMAN 15-1, 2013). Of specific interest is the limb foreshortening corrections, as detailed in section 4.6.5 of this manual:

> *"4.6.5. Determine the total corrected area of each sunspot group using the limb foreshortening overlay.*
>
> *4.6.5.1. Center the limb foreshortening overlay on the sunspot drawing and rotate it so that the N/S line of the overlay corresponds with the N/S line of the form and runs from the sun center through the geometric center of the group. The 'hash mark' across the group's center represents a limb foreshortening correction factor.*
> *4.6.5.1.1. If the group's center is between two hash marks use the smaller correction factor (i.e. the hash mark closer to the disk center) no matter which*

*hash mark is closest to the group's center.*

*4.6.5.1.2. If the group's center lies below the 1.1 hash mark, give it a correction factor of 1.*

*4.6.5.1.3. If the group's center lies beyond the 3.0 hash mark, give it a correction factor of 3.*

*4.6.5.2. To calculate a group's corrected spot area (in millionths of the solar hemisphere) multiply the total uncorrected spot area by the foreshortening correction factor. Report each group's area in whole increments of 10 millionths of the solar hemisphere rounded up or down as appropriate.*"

Figure 4.15 shows the overlay referred to above. There is a fair bit of area between the end of the scaling factors and edge of the disk. It should therefore be possible to measure spots with area correction factors greater than 3, but this is not done.

The distance to the hash marks were remeasured to confirm that they were assigned the correct limb correction factor. This is indeed the case which means that the USAF sunspot measurements have an inbuilt underestimate of the area due to the rounding down to the lower hash mark, if a sunspot lies between 2 hash marks

In addition the USAF uses a series of ellipses to measure the size of spots rather than a grid, as observers at Greenwich used. This USAF overlay is shown in figure 4.16. The analyst would find the best fitting ellipse to a sunspot and use that for the area.

Finally the USAF has a series of Stoneyhurst overlays which allow an analyst to determine the position of the sunspots. An example is shown in figure 4.17.

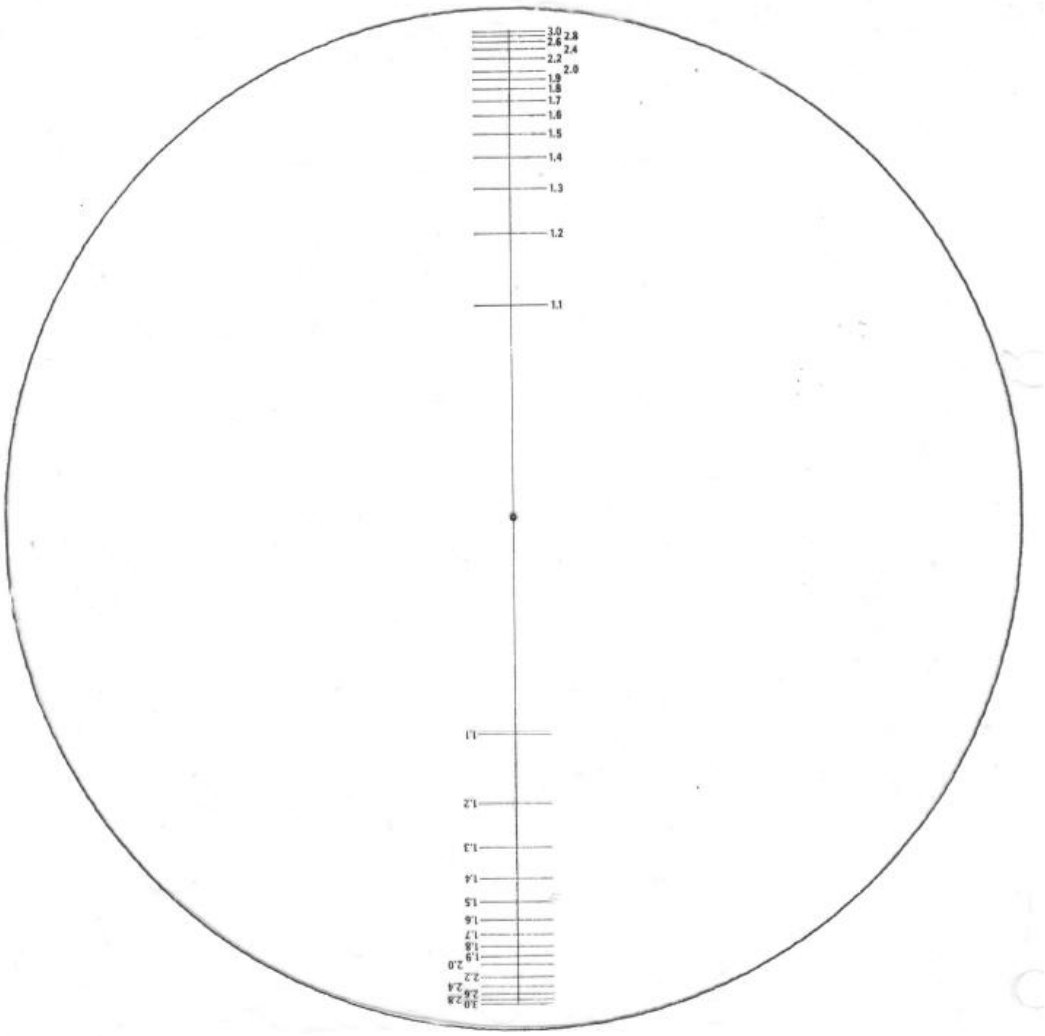



**Figure 4.15**: USAF Limb Foreshortening Overlay

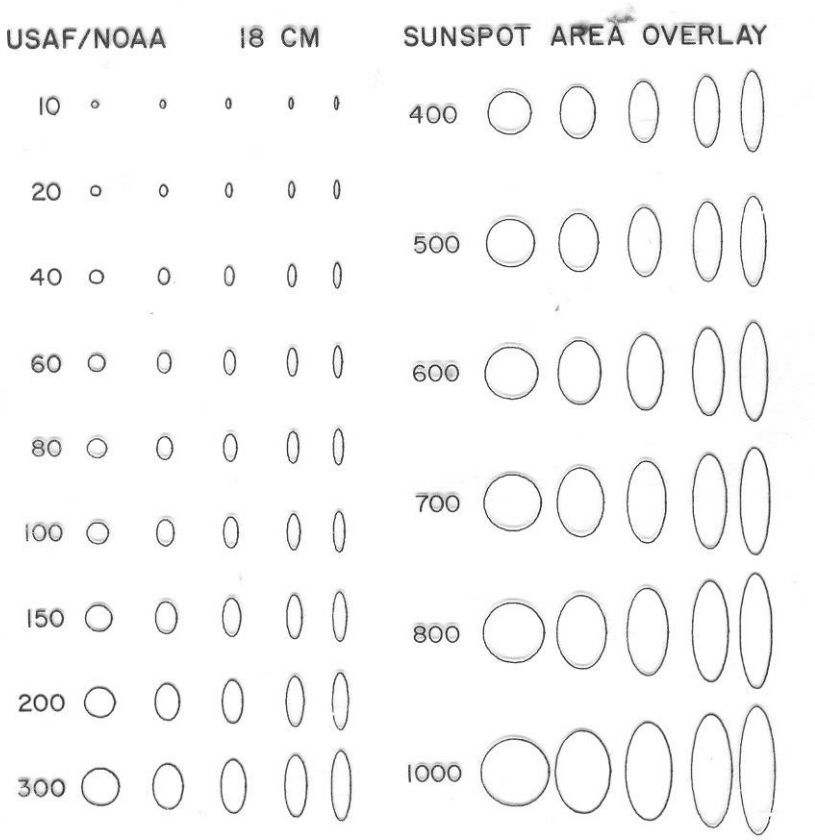



**Figure 4.16**: USAF Sunspot Area Overlay

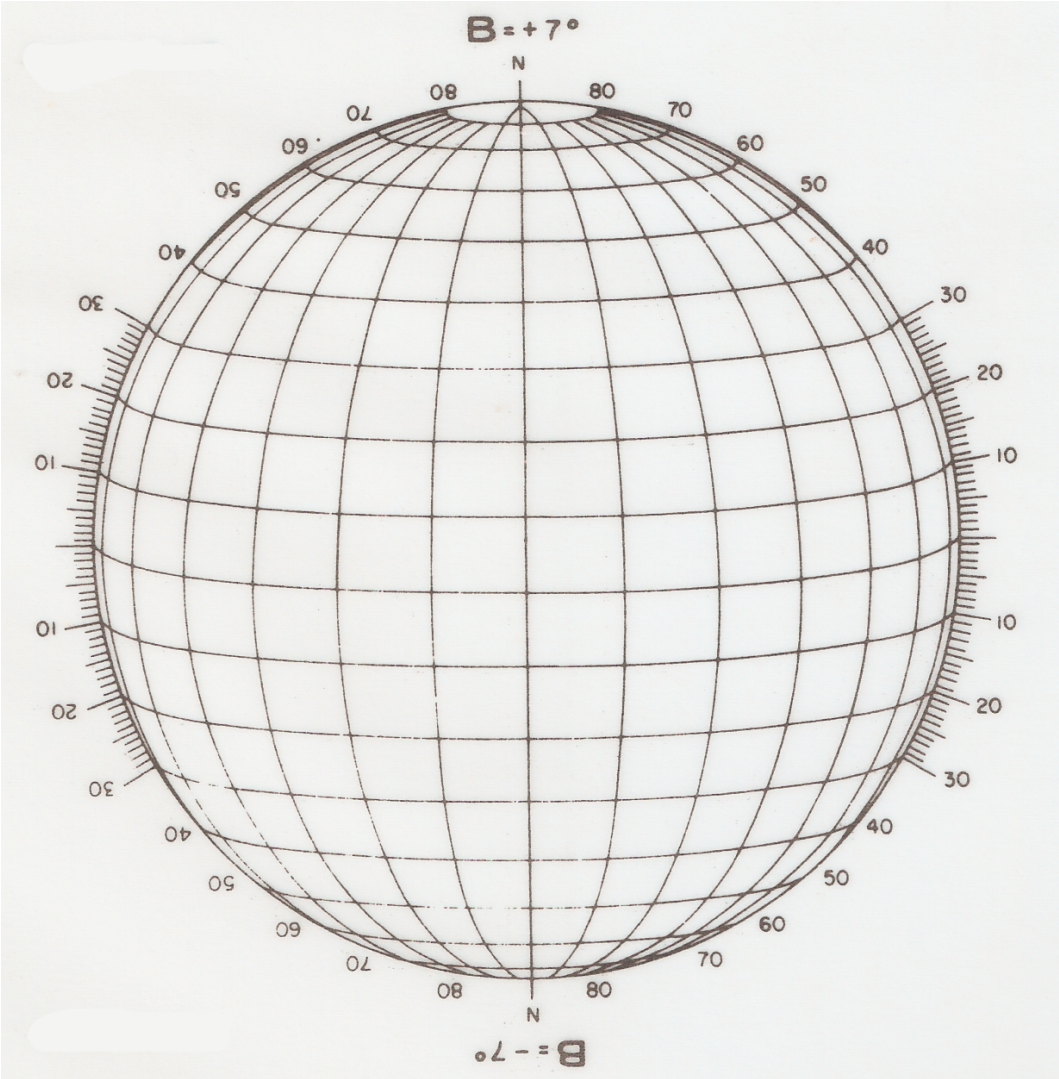



**Figure 4.17**: Example Stoneyhurst Overlay. There are 8 of these overlays for different solar B-angles.

In order to test whether this technique, particularly rounding of area correction factors, was adding a bias, two entire month's of sunspot drawings, from Learmonth for January 1990 and January 2002 were remeasured. Two different cycles were chosen to determine if there were long term differences between cycles or observers. It is noted that even within a month there are different observers reporting sunspots.

The sunspot images were printed from the scans archived at NGDC. Rather than using the overlays above, a 1 mm grid was used to measure sunspot areas. The distance to each group was also measured using this same grid, as was the radius of the disk. In this way, a more precise measure of the area correction factor could be obtained. Area correction

factors of up to 6 could be used using this technique. The area correction factor is found by equation 4.2.

$$c_f = \frac{1}{\sqrt{1 - r_v^2}} \quad , \qquad \text{Eqn 4.2}$$

where $c_f$ is the area correction factor and $r_v$ is the fractional radial distance from the centre of the disk to the sunspot group.

Figure 4.18 shows the comparison between the original USAF area and the remeasured areas for individual regions. There is little change between the 1990 and 2002 drawings, and from the regression equations there is a consistent 12% increase in area for the rescaled values over the USAF values. There are several possibilities to account for this difference. As discussed earlier, the rounding down of the scaling factor is going to be a contributing factor. It may also be that the ellipses used, particularly for smaller spots, aren't quite the right size (scaling errors in printing the overlays for example). Does the area of the ellipse include the area occupied by the line? The specific overlays used by each of the sites would need to be carefully measured to determine this.

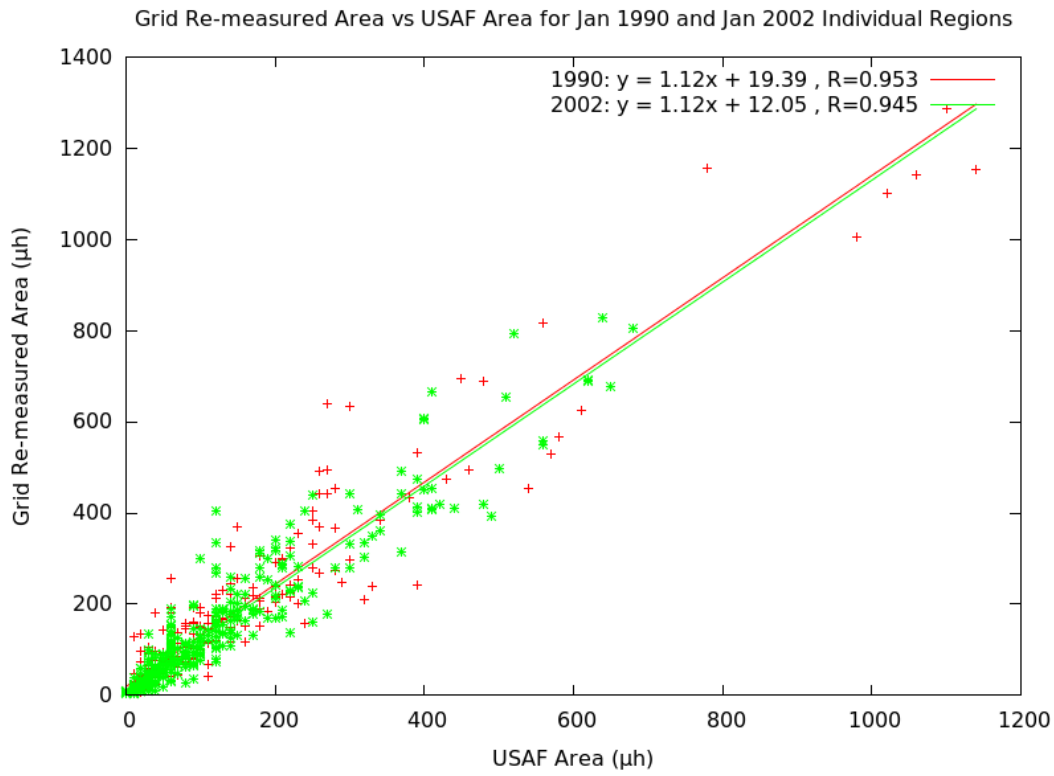



**Figure 4.18**: Grid Re-measured Area vs USAF Area for January 1990 and January 2002 for Individual Regions

The approach in determining if the USAF underestimate of area was due to the rounding of correction factors was to examine regions with an area correction factor of less than 1.5, where the effects of rounding would be less pronounced. This is shown in Figure 4.19. This plot shows that the remeasured areas are now closer to the USAF areas, although 1990 showed a better result than 2002. However, the result does suggest that limb correction may be at least part of the problem of USAF AFWA underestimating sunspot area.

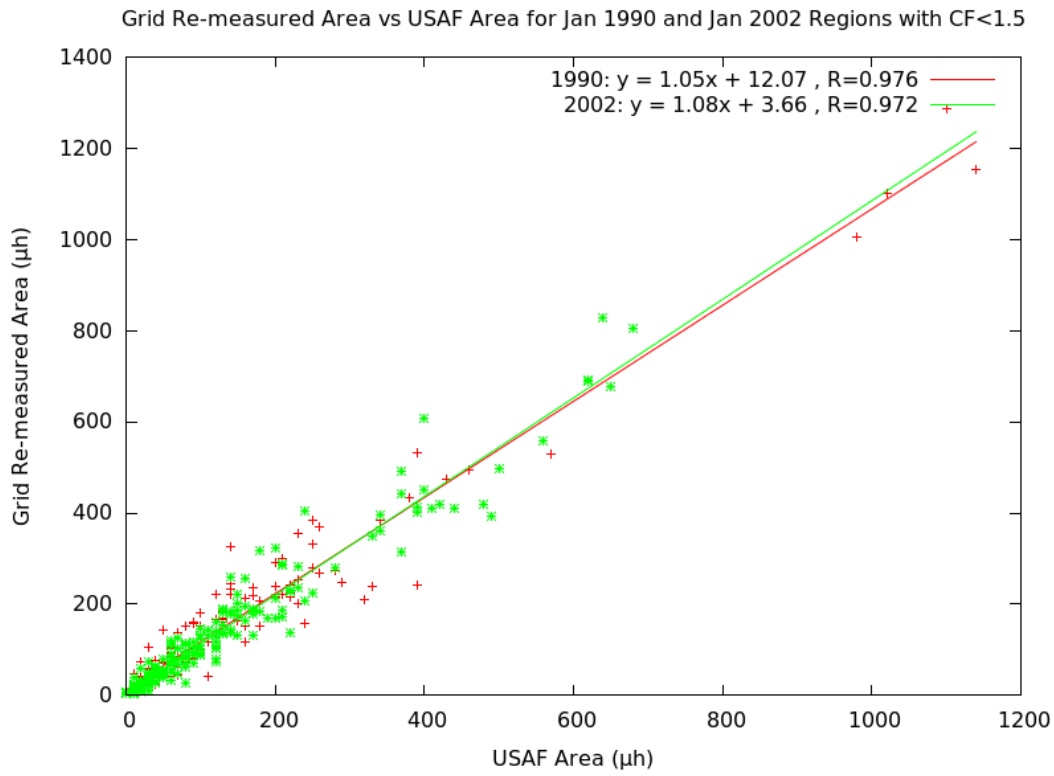



**Figure 4.19**: Grid Re-measured Area vs USAF Area for January 1990 and January 2002 for Individual Regions with Area Correction Factors <1.5

If the USAF were underestimating sunspot area, is it possible to estimate the average size of the error and potentially correct for it in the data set? The sunspot distribution by latitude is required to answer this. USAF sunspot data from 1985 to 2014 was used to extract the positions and area of each sunspot group. The positions were adjusted to remove the B-angle correction (tilt of the sun with respect to the ecliptic towards or away from the earth) and the distance from the centre of the disk to each spot group calculated. The correction factor was calculated for each spot region. However, as the USAF analysts round down to the lower intervals in th Limb Foreshortening Overlay (figure 4.15), the USAF correction factor was found by using the following equations:

$$c_u := \begin{cases} \lfloor 10\,c_f \rfloor / 10 : & c_f < 2 \\ \lfloor 5\,c_f \rfloor / 5 : & 2 \leqslant c_f < 3 \\ 3 : & c_f \geqslant 3 \end{cases} , \qquad \text{Eqn 4.3}$$

where $c_u$ is the USAF correction factor and cf the area correction factor from equation

4.2. The $\lfloor x \rfloor$ function is the floor function and returns the largest integer less than or equal to x.

The uncorrected area of a region is then found by Equation 4.4 :

$$A_{uc} = A_u / c_u \quad , \qquad \text{Eqn 4.4}$$

where $A_{uc}$ is the uncorrected area, and $A_u$ is the area reported by the USAF analyst.

Figure 4.20 shows the histogram of estimated uncorrected sunspot area by latitude from 1985 to 2014, normalised so that the total area is 100%.

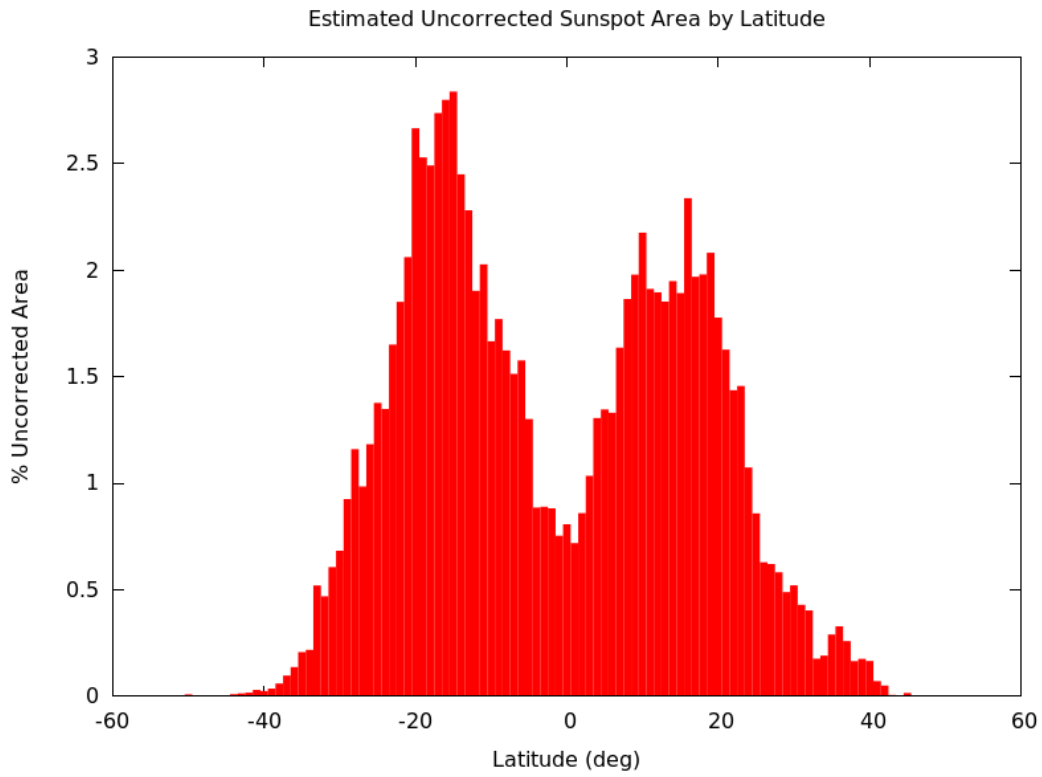



**Figure 4.20**: Normalised Estimated Distribution of Uncorrected Sunspot Area by Latitude

A brief aside on figure 4.20. The sunspot distribution is asymmetrical between the northern and southern hemispheres. This asymmetry is well recognised and can change within a cycle or between cycles (Bray and Loughhead, 1964).

If a uniform distribution of sunspot area by longitude is assumed then a relative measure of sunspot area is simply:

$$A_R = \int_{-90}^{90} \int_{-90}^{90} P_A(\lambda)\, c(\lambda, \varphi)\, d\lambda\, d\varphi \quad , \qquad \text{Eqn 4.5}$$

where $A_R$ is the relative area, $P_A$ is the distribution in figure 4.20, and $c$ is the correction factor as a function of longitude $\lambda$ and latitude $\phi$.

Equation 4.5 was solved numerically for both the USAF correction factor and the non-rounded (true) correction factor. The ratio between these two values ($A_c/A_u$) was 1.09. That is, the USAF values underestimate the sunspot area by 9% just due to the rounding of correction factors. For individual regions on the limb ($c_f$~6), areas may be under-measured by up to 100%.

This was further confirmed by computing smoothed monthly averages (to remove noise) of the USAF original area and recalculating the area without rounding the correction factors. This was done for each day and site by recalculating the areas for each region and then adding these to obtain the daily recalculated areas for each site. The daily values for each site were then averaged to obtain a daily area. Monthly and smoothed monthly averages were then computed. The results are shown in figure 4.21 and the gradient of 1.09 confirms the result from equation 4.5.

In summary it can be confirmed just from analysis of the USAF sunspot area analysis procedures that there is a 9% deficiency in their calculation of total total sunspot area. This result refutes claims (e.g. Wilson and Hathaway [2005] and Foukal [2014]) of up to a 40% sunspot area deficiency.

A simple correction to the way the USAF perform their sunspot analysis, by more accurate scaling factors, would remove this error in sunspot areas. The database could be adjusted by simply applying the corrections detailed above to each region, although a more rigorous way would be to reanalyse all of the sunspot drawings. This however would be an exceedingly time consuming task.

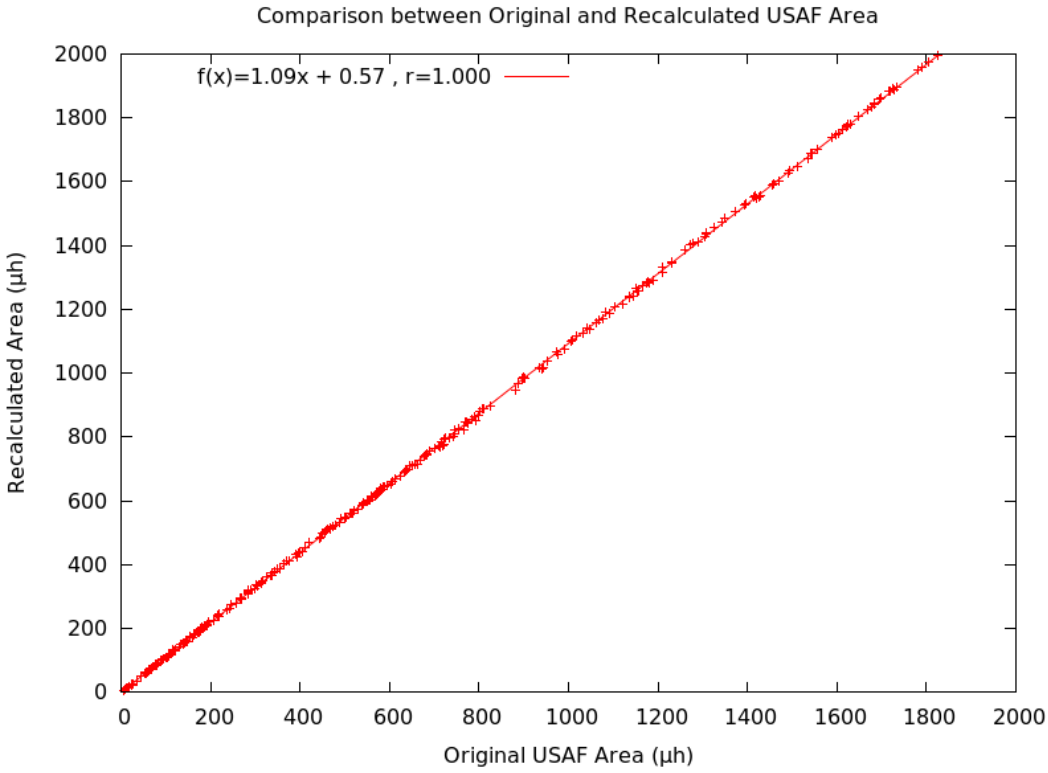



**Figure 4.21**: Comparison between Original and Recalculated USAF Area

### 4.4 10.7 cm Radio Flux

There are few direct observations of the 10.7 cm radio flux other than the Penticton data (Tapping and Morton, 2013) set, so comparisons are difficult. However there is some anecdotal evidence to suggest that the flux values at various times may have suffered variations in measurements at this wavelength. The USAF RSTN network checks their 2695 MHz flux values from the radio against Penticton. However, during the 1980s and 1990s it seems that Penticton was also checking their values against the USAF (Personal communication, John Kennewell). Therefore during this period some values may be questionable, if the two institutions were making corrections based on these comparisons, in a mutually reinforcing feedback loop. Currently Penticton uses horn antennas which can be more easily calibrated to check their flux values (Personal Communication, John Kennewell).

### 4.5 Comparison of Indices

Various researchers have compared solar indices discussed in the previous sections against each other. Within the first decade of operation of Penticton (then located at

Algonquin), Dodson et al. (1954) performed a study comparing properties of Hα flares to 2800 MHz radio bursts. This was not so much a comparison of long term variation as more an individual flare-burst comparison. None-the-less, there was an acknowledgement that different types of solar activity are somewhat related.

Dodson et al. (1974) also compared properties of sunspots in cycles 18, 19 and 20 and noted that the duration of cycles, the time of the peak of cycles and the maximum and values measured in the cycles varied depending on which parameter was being measured.

As mentioned in section 4.3, Wilson and Hathaway (2006) have used sunspot area to sunspot number ratios to show discrepancies in the both the sunspot number and area measured by different groups. However, this approach is problematic in that it is unclear that these ratios remain constant across cycles, or indeed within cycles.

Indeed, Özgüç and Ataç (2001) showed that the relationships between sunspot area, magnetic field and coronal index show some hysteresis between the rise and fall of cycles 21 and 22. Whether this is due to actual solar variation or problems in measurement as discussed earlier was not determined by Özgüç and Ataç (2001).

In this section, data from 1945 to 2014 is reviewed to compare cycle properties. It is noted at the outset all of the limitations detailed in the previous sections; although some significant conclusions can still be drawn.

Figure 4.21 shows the sunspot number, sunspot area and 10.7 cm flux from 1945 to 2014, across cycles 18 to 24. The values used are all smoothed monthly means, to remove short term variations from the data. The sunspot number and 10.7 cm flux are the International Sunspot Number and Penticton radio flux values. As stated in chapter 2, the International Sunspot Number site moved from Zurich to Brussels in 1980 and the 10.7 cm flux instrumentation moved from Ottawa to Penticton in 1991.

The sunspot area observations are more problematic than sunspot numbers. For the period up to 1976, Greenwich values were used. From 1977 to 2014 the value for any month was chosen depending on availability of data from a site. The value was chosen in order of highest to lowest preference as follows: Solar Region Summary, USAF/AFWA, Rome, Debrecen. This means there may be discontinuities in the area observations. Preference

was given to agencies with the fewest missing number of months.

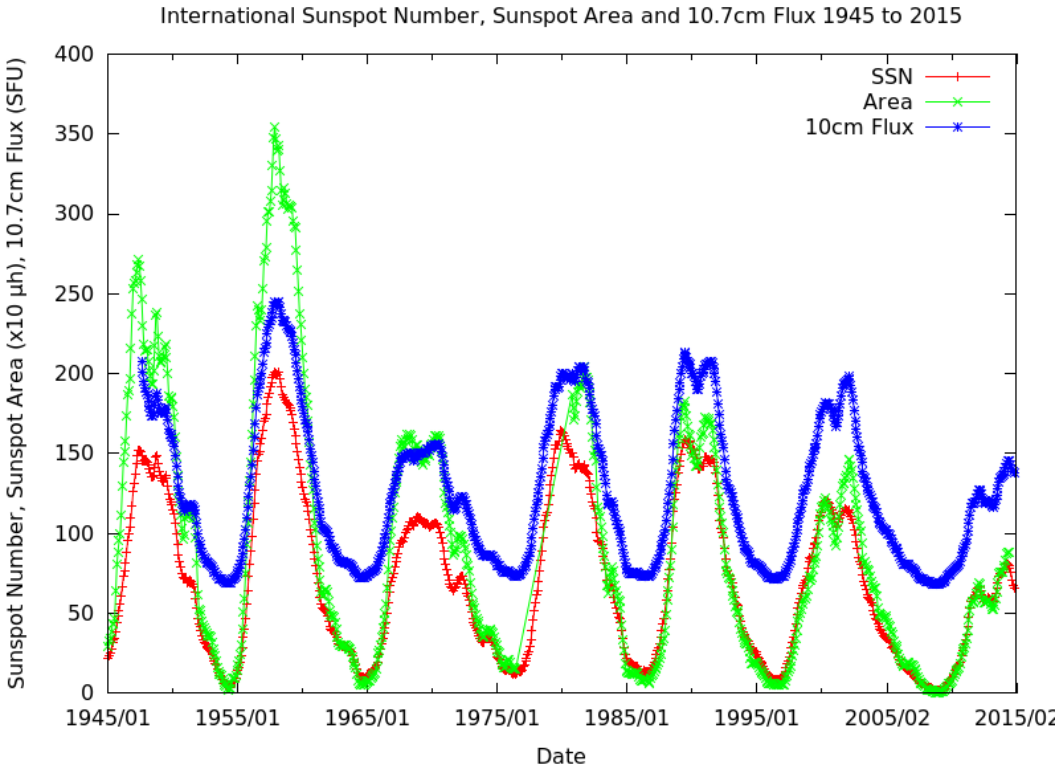



**Figure 4.22**: International Sunspot Number, Sunspot Area and 10.7cm Flux from 1945 to 2014, across cycles 18 to 24.

Figure 4.22 shows that all of these indices vary from cycle to cycle. Cycle 19 (April 1954 to November 1968) is the largest for all indices. The figure also shows that these indices change relationships between each other from cycle-to-cycle, discussed in detail below

Tables 4.13 and 4.14 give the dates and values for minimum and maximum of the three parameters in figure 4.22 across solar cycles shown. It can be seen that the minima for the different parameters often occur within a few months of each other, but maxima can often differ by years. Also, the start and end of cycles may differ. Therefore researchers and analysts should clearly specify which cycle they are talking about and which dates they are using for the start and end of cycles.

| | Sunspot Number | | Sunspot Area (μh) | | 10.7cm Flux (SFU) | |
|---|---|---|---|---|---|---|
| Cycle | Date | Value | Date | Value | Date | Value |
| 19 | 1954/04 | 3.4 | 1954/04 | 23 | 1954/04 | 69.0 |
| 20 | 1964/10 | 9.6 | 1964/10 | 51 | 1964/06 | 72.3 |
| 21 | 1976/05[1] | 12.2 | 1976/03[2] | 152 | 1976/06 | 73.3 |
| 22 | 1986/09 | 12.3 | 1986/09 | 63 | 1986/09 | 72.9 |
| 23 | 1996/05 | 8.0 | 1996/07[3] | 53 | 1996/05 | 71.5 |
| 24 | 2008/12 | 1.7 | 2008/11[4] | 5 | 2008/10 | 68.1 |

**Table 4.13**: Cycle Minima for Sunspot Number, Sunspot Area and 10cm Flux

Explanatory notes:

1. Cycle 21 sunspot number minimum value of 12.2 occurred in 1976/03 and 1976/06. 1976/05 was chosen as the midpoint of these values.

2. Greenwich data ended in 1976 and the USAF did not take over until 1980. During this period, Rome data was sporadic and unreliable smoothed monthly averages were able to be calculated. The date given may not be true minimum for sunspot area in cycle 21.

3. Cycle 23 sunspot area minimum value of 53 μh was recorded in 1996/05, 1996/08 and 1996/09. 1996/07 was chosen as the minimum as it lies in the middle of these dates.

4. Cycle 24 sunspot area minimum value of 5 μh was recorded in 2008/11 and 2008/12. 2008/11 was chosen as the minimum.

| | Sunspot Number | | Sunspot Area (μh) | | 10.7cm Flux (SFU) | |
|---|---|---|---|---|---|---|
| Cycle | Date | Value | Date | Value | Date | Value |
| 18 | 1947/05 | 151.8 | 1947/04 | 2677 | 1947/08[3] | 207.5 |
| 19 | 1958/03 | 201.3 | 1957/11 | 3547 | 1958/03 | 245.2 |
| 20 | 1968/11 | 110.6 | 1968/04 | 1619 | 1970/07 | 156.3 |
| 21 | 1979/12 | 164.5 | 1981/10[1] | 2046 | 1981/05 | 204.6 |
| 22 | 1989/07 | 158.5 | 1989/06 | 1825 | 1989/06 | 213.0 |
| 23 | 2000/04 | 120.8 | 2002/02[2] | 1462 | 2002/02 | 198.3 |

**Table 4.14:** Cycle Maxima for Sunspot Number, Sunspot Area and 10cm Flux

Explanatory Notes:

1. Greenwich sunspot area data ended in 1976 and the USAF did not take over until 1980. During this period, Rome data was sporadic and unreliable smoothed monthly averages were able to be calculated. The date given may not be true maximum for sunspot area in cycle 21.

2. Cycle 23 sunspot area maximum value of 1462μh was recorded in 2002/02 and 2002/03. 2002/02 was chosen as the maximum.

3. The cycle 18 10.7 cm flux maximum may have occurred before this value. This was the first date for which smoothed values were obtained.

The explanatory notes for the Minima and Maxima of Sunspot Number Cycles from NGDC (ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-indices/sunspot-numbers/cycle-data/table_cycle-dates_maximum-minimum.txt) state:

> "**When observations permit, a date selected as either a cycle minimum or maximum is based in part on an average of the times extremes are reached in the monthly mean sunspot number, in the smoothed monthly mean sunspot number, and in the monthly mean number of spot groups alone. Two more measures are used at time of sunspot minimum: the number of spotless days and the frequency of occurrence of 'old' and 'new' cycle spot groups.*
>
> ***The smoothed monthly mean sunspot number is defined here as the arithmetic average of two sequential 12 month running means of monthly mean numbers."*

Therefore, the official sunspot number cycle dates do not necessarily coincide with the dates of smoothed monthly sunspot number maxima or minima, and further averaging is performed on the sunspot numbers to generate these dates. Therefore, it cannot be assumed that the NGDC reported sunspot number cycle dates coincide exactly with the dates of smoothed monthly sunspot number maxima or minima. This is often due to the confusion between the sunspot number and a so-called and undefined solar activity cycle.

In the discussion to follow, the cycle dates are taken from table 4.13. For purposes of comparison, the smoothed monthly sunspot number cycle dates are used for all indices.

Figure 4.23 shows Sunspot Area vs Sunspot Number for cycles 18 to 24. Table 4.15 shows the regression parameters for these fits. The only two cycles that are statically similar (by the test given in equation 1.6) when comparing sunspot area and sunspot number are cycles 23 and 24, although cycle 24 contains only partial data as it was not a complete cycle at the time of analysis.

In addition the only intercepts that are statistically similar are also cycle 23 and 24 (by the test given in equation 1.6). However most intercepts do not go through the origin. One might think that they should, because if there is zero sunspots then there is zero sunspot area. However the smallest allowable sunspot number aside from zero is 11 (1 group with 1 spot). Sunspot area, at least up until 1977) could have values less than 10 μh (when USAF took over, area is rounded to the nearest 10 μh). This discontinuity in the sunspot number may be biasing the intercept value.

There does appear to be a splitting of the cycles, where cycles 18, 19 and 20 all have gradients of greater than 1.5, where cycles 22, 23 and 24 have gradients of approximately 1. As most of the cycles are statistically different in terms of $m$ and $b$, it is difficult to determine if the change in gradient is naturally occurring or if it has to do with the USAF taking over area measurements from Greenwich on the rise of cycle 21.

Hysteresis is also seen in these plots, showing that the sunspot area to sunspot number ratio changes after the maximum of the cycle, again complicating any corrections that may be based on these ratios.

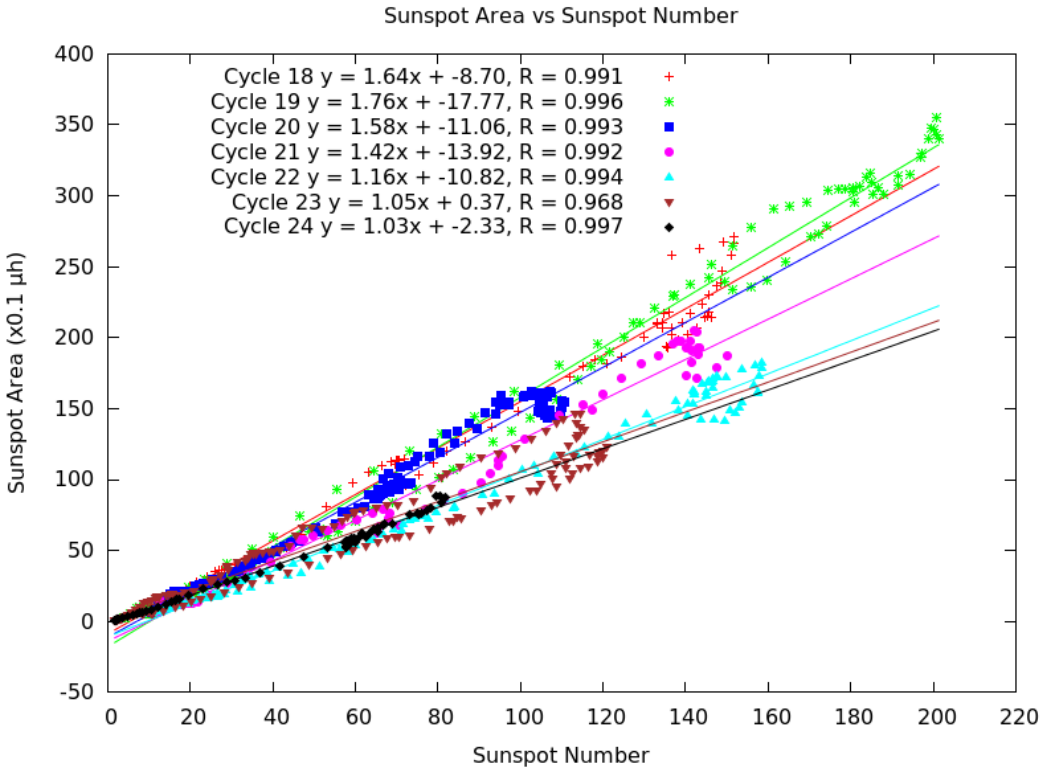



**Figure 4.23**: Sunspot Area vs Sunspot Number for Cycles 18 to 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 18 | 1.636 | 0.024 | -8.697 | 2.382 | 0.991 |
| 19 | 1.758 | 0.014 | -17.766 | 1.627 | 0.996 |
| 20 | 1.585 | 0.017 | -11.064 | 1.163 | 0.993 |
| 21 | 1.419 | 0.021 | -13.918 | 1.795 | 0.992 |
| 22 | 1.159 | 0.012 | -10.817 | 1.130 | 0.994 |
| 23 | 1.052 | 0.022 | 0.368 | 1.514 | 0.968 |
| 24 | 1.035 | 0.010 | -2.334 | 0.511 | 0.997 |

**Table 4.15**: Sunspot Area vs Sunspot Area Regression Parameters

Figure 4.24 shows 10.7cm Flux vs Sunspot Number and table 4.16 the regression parameters. Here the relationships in terms of gradient have less of a range (~15% variation) compared to the Sunspot Area – Sunspot Number plots (~60% variation). This would suggest that either there is simply a greater variation of sunspot area, or that sunspot area is not being measured as consistently as sunspot number.

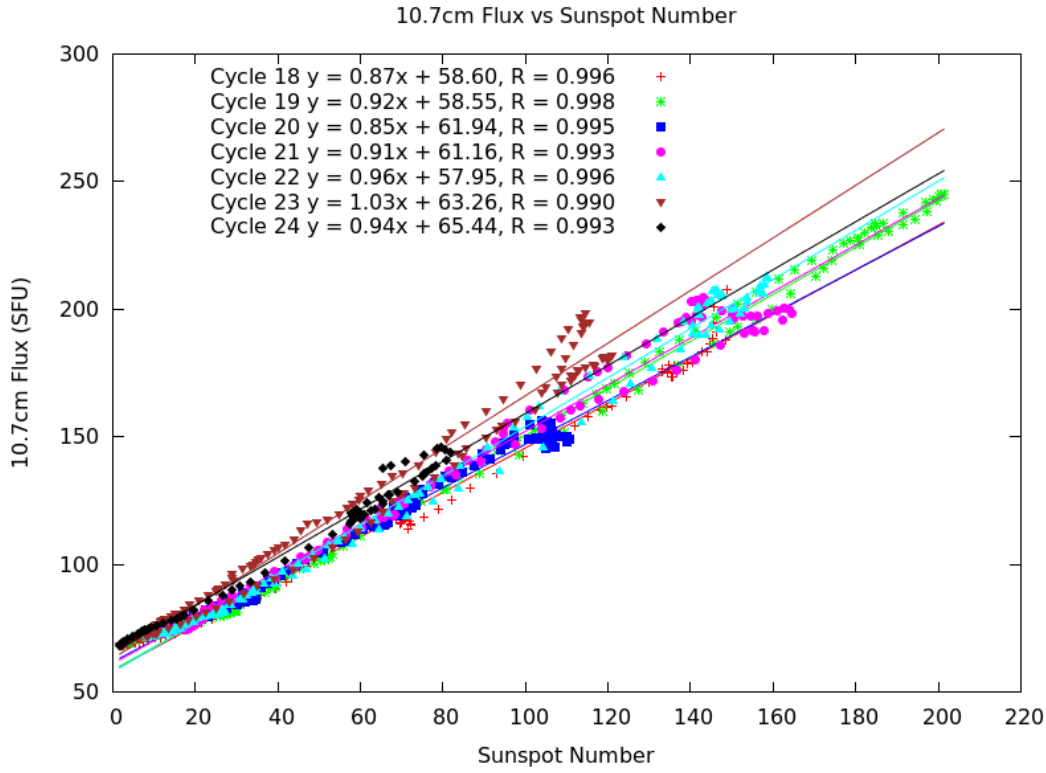



**Figure 4.24**: 10.7 cm Flux vs Sunspot Number for Cycles 18 to 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 18 | 0.872 | 0.009 | 58.599 | 0.870 | 0.996 |
| 19 | 0.921 | 0.005 | 58.553 | 0.557 | 0.998 |
| 20 | 0.853 | 0.007 | 61.941 | 0.508 | 0.995 |
| 21 | 0.911 | 0.010 | 61.162 | 0.933 | 0.993 |
| 22 | 0.961 | 0.008 | 57.947 | 0.764 | 0.996 |
| 23 | 1.029 | 0.012 | 63.256 | 0.799 | 0.990 |
| 24 | 0.938 | 0.013 | 65.440 | 0.705 | 0.993 |

**Table 4.16**: 10.7cm Flux vs Sunspot Number Regression Parameters

There are some cycles in the 10.7cm flux vs sunspot number plots that have statistically similar (within two standard deviations) gradients: cycles 18 and 20, 22 and 23, 21 and 24. There is not a difference in gradient as significant as that seen in the sunspot area – sunspot number comparison (figure 4.22).

The significance of the intercept (*b*) in figure 2.4 and table 4.16 is that this will be the

10.7 cm flux when there are no spots on the solar disk facing the Earth. The values here are about 10 SFU lower than the minimum values actually observed (table 4.13). This should not be surprising as there is no sunspot number minimum with a smoothed sunspot number of zero (although cycle 24 gets close). Also, as already mentioned the minimum of these indices do not necessarily occur simultaneously.

Finally, figure 4.25 shows the relationship between 10.7cm flux and sunspot area with table 4.17 giving the regression parameters. In this case it can be seen that the gradients for cycles 18, 19 and 20 are statistically similar, as are the intercepts for cycles 19 and 20. The gradients for cycles 22, 23 and 24 are approximately 35% higher than those for cycles 18 to 20, further suggesting that cycles 18 to 20 are different from cycles 22 to 24.

The intercepts here are higher than those for 10.7 cm flux vs sunspot number, but still lower than the 10.7 cm flux minimums that were observed for these cycles.

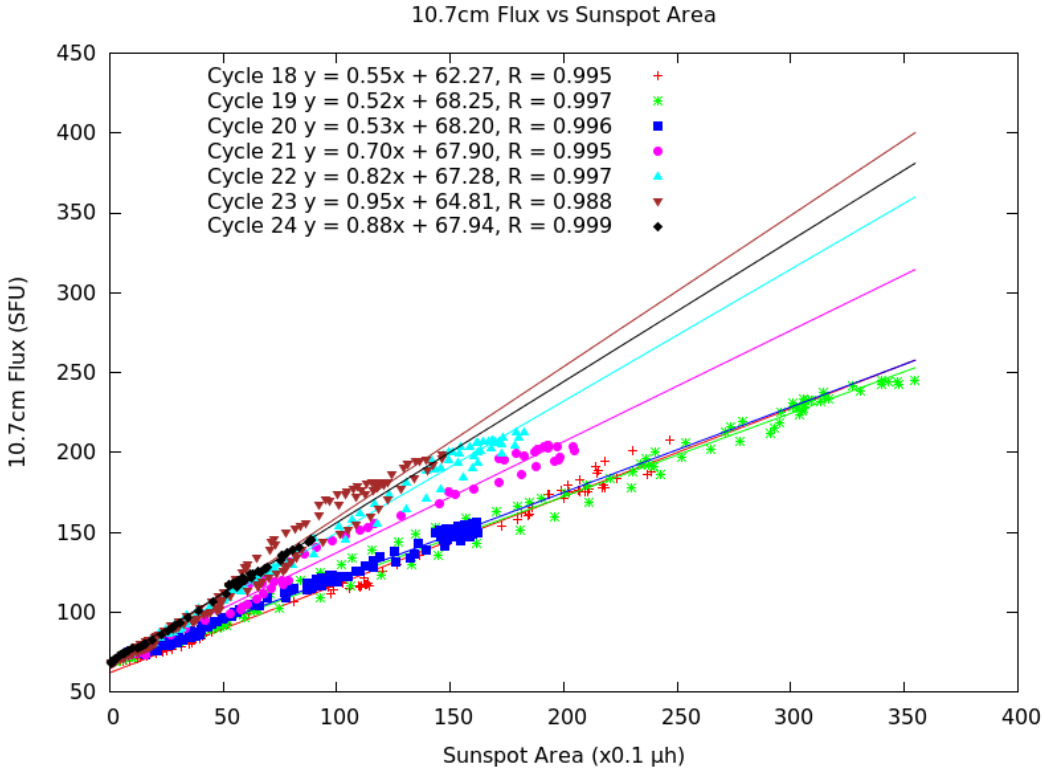



**Figure 4.25**: 10.7 cm Flux vs Sunspot Area for Cycles 18 to 24

| Cycle | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| 18 | 0.551 | 0.006 | 62.271 | 0.905 | 0.995 |
| 19 | 0.521 | 0.004 | 68.246 | 0.652 | 0.997 |
| 20 | 0.534 | 0.004 | 68.204 | 0.430 | 0.996 |
| 21 | 0.695 | 0.008 | 67.901 | 0.896 | 0.995 |
| 22 | 0.825 | 0.006 | 67.284 | 0.668 | 0.997 |
| 23 | 0.945 | 0.012 | 64.806 | 0.875 | 0.988 |
| 24 | 0.883 | 0.006 | 67.941 | 0.290 | 0.999 |

**Table 4.17**: 10.7cm Flux vs Sunspot Area Regression Parameters

Finally, the sunspot area to sunspot number ratio is revisited. Wilson and Hathaway (2006) attempted to use this ratio to demonstrate that sunspot area was being incorrectly measured. Figure 4.26 showing this ratio from 1945 to 2015 over cycles 18 to 24, demonstrates why this is not a good method to correct perceived sunspot area errors. This ratio varies significantly on time scales less than a sunspot cycle. Some of these variations may be due to measurement error, but attempting to remove intrinsic solar variation using this index to be left only with measurement errors seem to be an intractable task.

In addition, this index seem to be decreasing. The green line is a simple regression line. That is, the sunspot area to sunspot number ratio is, on average, decreasing at a rate of 0.00987 μh/year. However, no predictions are made as to whether this trend will continue into future cycles.

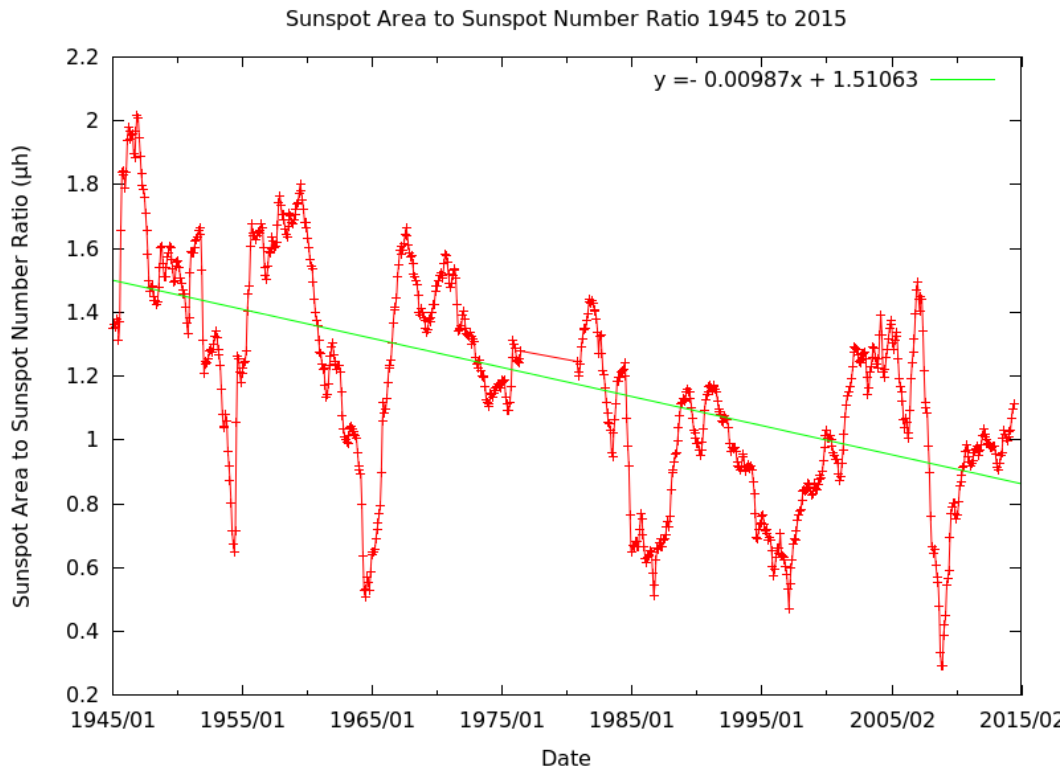



**Figure 4.26**: Susnpot Area to Sunspot Number Ratio from 1945 to 2014

### 4.6 Summary and Conclusions

Sunspot area, sunspot number and 10.7 cm flux were examined across 7 decades (cycles 18 to 24). The cycles for each of these parameters vary in length and size. Part of this variability results from alterations in observation techniques and systematic error. However, there also appears to be intrinsic solar variation between and within solar cycles, making corrections difficult.

Of particular note is the reported discrepancies between USAF and Greenwich sunspot area measurements. By carefully remeasuring a selection of sunspot drawings, it was found that the USAF analysts are under reporting total disk areas by about 12%, just based on measurement technique. This error is less than the 22% that Wilson and Hathaway (2005) claim and 40% to 50% that Foukal (2014) claims.

Furthermore, 9% of the under reporting can be attributed to the rounding down of limb correction factors that the USAF perform. If this practice is modified to using the non rounded values, the average areas increase by 9%.

Comparison of these indices on a cycle-by-cycle basis show significant variation in the relationships between indices across sunspot cycles. There is suggestive but not conclusive evidence, particularly in the 10.7 cm flux versus sunspot area comparisons, that sunspot areas in cycles 18 to 20 were different from those in cycles 22 to 24. Since cycle 21 was when Greenwich Observatory ceased sunspot area measurements and the USAF started them, this shift of observing agency cannot be excluded as being a contributing factor to these variations in sunspot area relationships. However the 10.7 cm flux vs sunspot number plots still show variation, suggesting that there are underlying solar mechanisms contributing to these variations.

# 5. Hα Flares

## 5.1 Introduction

The Hα line has a wavelength of 653.6 nm and is useful for observing flares as the line is a chromospheric line, one that is particularly intense and associated with the chromosphere (Bhatnagar and Livingston, 2005).

Hα images of the Sun have been observed since the early 1900s. The first devices used were spectrohelioscopes which comprised a grating and 2 slits. One of the slits was scanned across the image of the Sun to produce a 2 dimensional image. This process was detailed by Hale (1924). This device had the advantage that it could be 'tuned' to a particular frequency by selecting the appropriate wavelength from the grating. They were often designed to maximise brightness in a particular part of the spectrum. The disadvantage was the disk had to be scanned, which could take time. This was fine for observing plage and filaments, or other mostly static features, but problematic when observing flares. Another difficulty was that the slits had to be moved synchronously to acquire useable images.

In the late 1920s Lyot developed a filter that was specific to a given wavelength. These types of filters consist of several plates, each half the thickness of the previous plate. Each plate is separated by a polaroid. Different starting thickness of the plate select different wavelengths. This whole system needs to be temperature stabilised. Variations in temperature will mean that the filter goes 'off-band' and the desired wavelength will no longer be selected. The filters are generally referred to as Lyot or birefringent filters. For full details see Paul (1974).

Filters typically have a band pass of less than 0.1 nm. The USAF SOON filter has a 0.05 nm band pass (although for calibration it can be reduced to half that). The GONG Hα filter has a band pass of about 0.04 nm. If the band pass is too broad then the Hα line saturates and the images have poor contrast. This narrow band pass has some disadvantages. During a flare the Hα line is thermally broadened, smearing the spectral line. This smearing can be as large as 2 nm, far exceeding the filter bandpass (Reid, 1963). The other drawback is that, if flaring material is moving, the flare will be Doppler shifted. Hα emission from material achieving escape velocity will be Doppler shifted by 1.35 nm, or 27 times the bandwidth of the filter. The SOON filter can make

measurements off-band by up to 0.1 nm, whereas there is no facility to measure off-band with GONG.

In the remainder of this chapter, Hα flare classification techniques will be discussed. It will be shown that in the past 60 years, little improvement has been made in obtaining consistent flare observations.

### 5.2 Flare Classification

Hα flares are classified by their area, A, (Importance class) and intensity, (Brightness Class) (Bhatnagar and Livingston, 2005).

As with sunspots, the area of a flare is measured in millionths of the solar hemisphere (μh). The Importance class has 5 categories from 0 to 4 and these are related to the area of a flare as shown in table 5.1:

| Importance | Area (μh) |
|---|---|
| 0 | 10≤A<100 |
| 1 | 100≤A<250 |
| 2 | 250≤A<600 |
| 3 | 600≤A<1200 |
| 4 | A>1200 |

**Table 5.1**: Hα Flare Importance Classification

Prior to 1975, Importance 0 flares were referred to as 'sub-flares'.

The intensity of a flare is measured relative to the background continuum of the sun, and is expressed as a percentage relative to this background. The background value is typically found by choosing the most commonly occurring intensity value in a region (that is the mode of intensities in a region). This method assumes that most of the pixels in a region are non-flaring. The Brightness class has three categories and are related to the the intensity as shown in table 5.2.

| Brightness | Intensity (% Background) |
|---|---|
| Faint (F) | 150%<I≤250% |
| Normal (N) | 250%<I≤350% |
| Brilliant (B) | I>350% |

**Table 5.2**: Hα Flare Brightness Classification

This method provides a two parameter flare classification system with the lowest being 0F and highest being 4B. The minimum requirements for a flare is that it must have at least 10 μh of area brighter than 150% of the background.

In addition, typically, the following criteria should be met:

- For a flare to begin it must be above the minimum 0F criteria for at least 2 consecutive minutes
- For a flare to end it must be below the 0F criteria for at least 2 consecutive minutes
- For a flare to considered to be in a particular brightness category it must have a minimum 10 μh of area above the minimum intensity for that brightness category.

The USAF SOON apply these criteria for flares. These procedures and criteria can be found in section 6 of AFWAMAN 15-1.

These criteria have been modified over the decades (for example: Report of the Working Committee on the Improvement in Assignment of Hα Flares importance by Commission 10 (on Solar Activity) of the XII General Assembly of the International Astronomical Union, Hamburg 1964). Prior to computerisation, the analysis was assessed by eyepiece, and was an observer's 'best guess' as to the Importance and Brightness. Some analysis was performed by recording films of a flare, thus giving more information. It was not until the late 1960s when videometers were developed that could supposedly produce consistent analyses between sites (Reid and Vorhaben, 1971). This at least removed some of the subjective elements out of flare analysis.

Unfortunately, in flare reports, usually only the largest peak is reported (although some

reports contain significant secondary peaks). From about 1975 to 1983 the entire flare profile was archived on what was known as MOSAV tapes. The flare profile was stored in a binary format. Unfortunately, NGDC never archived this data, and presumably the tapes are now lost. Thus, there are no complete flare profiles stored from the SOON system. However, in 2010, GONG installed Hα filters and cameras with full data archives. Figure 5.1 shows a flare observed from the Big Bear GONG.

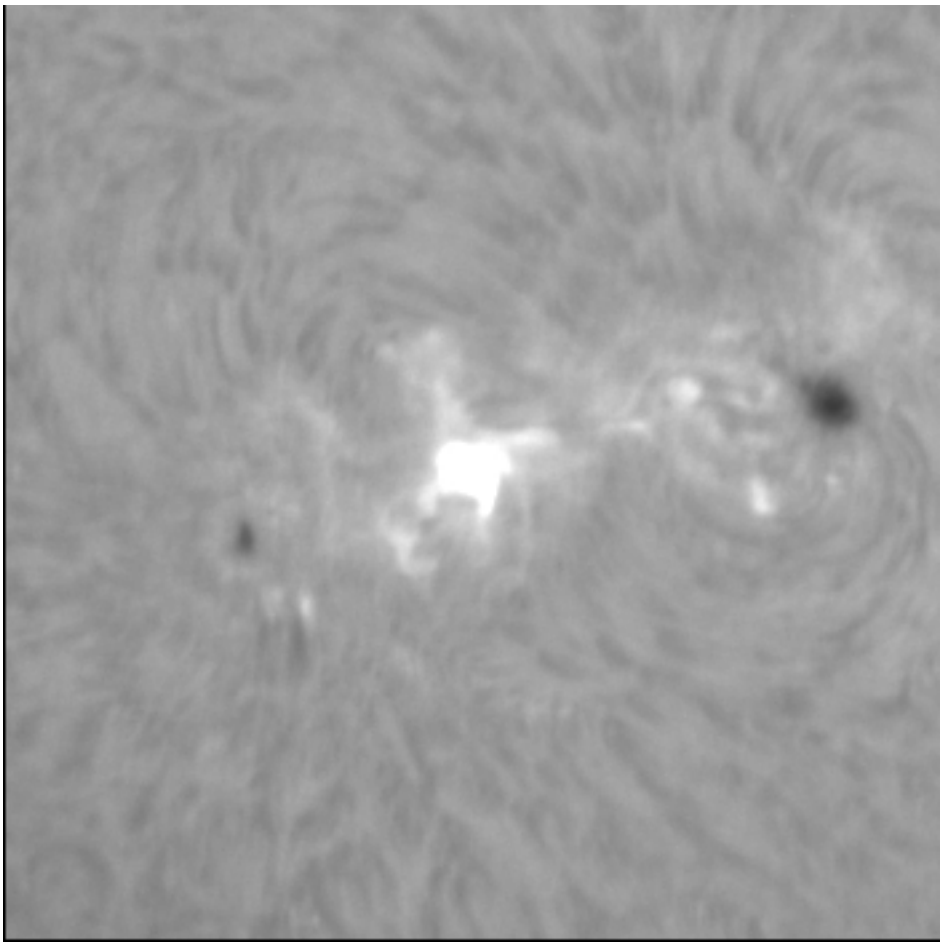


**Figure 5.1**: Flare Observed from Big Bear GONG on 2012/07/30

Figure 5.2 shows the intensity histogram of this flare at peak brightness. This histogram is used to determine both the flare Brightness and Importance. The plot is deliberately scaled so that the intensity bin with the highest area is 100%. In this case, the maximum intensity was 373%, but this bin only contained 0.22 μh of area. Counting back from this bin, it is found that 10 μh of area was contained in bins 313 or higher, making 313 the flare intensity, falling into the Normal (N) category.

The area is simply the sum of all bins of 151% or greater, in this case 119 μh, giving it an importance of 1. Therefore the flare class for this flare was 1N.

Figures 5.3 and 5.4 show the progression of intensity and area respectively.

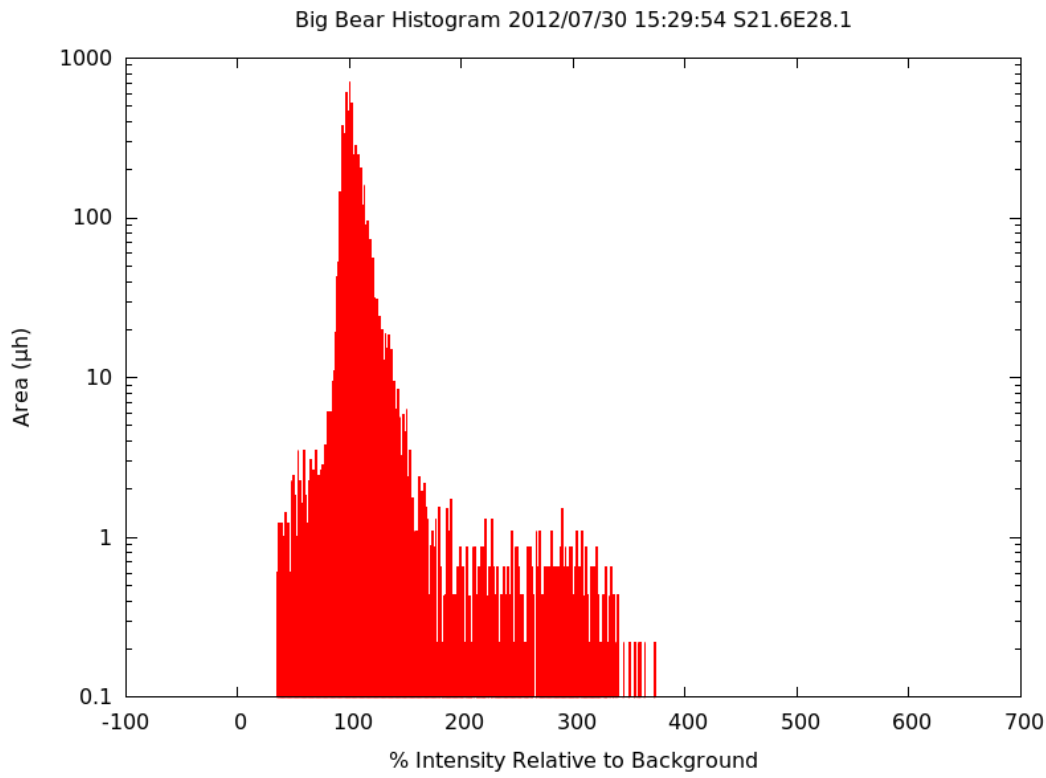



**Figure 5.2**: Peak Intensity Histogram for Flare Observed at Big Bear in Figure 5.1

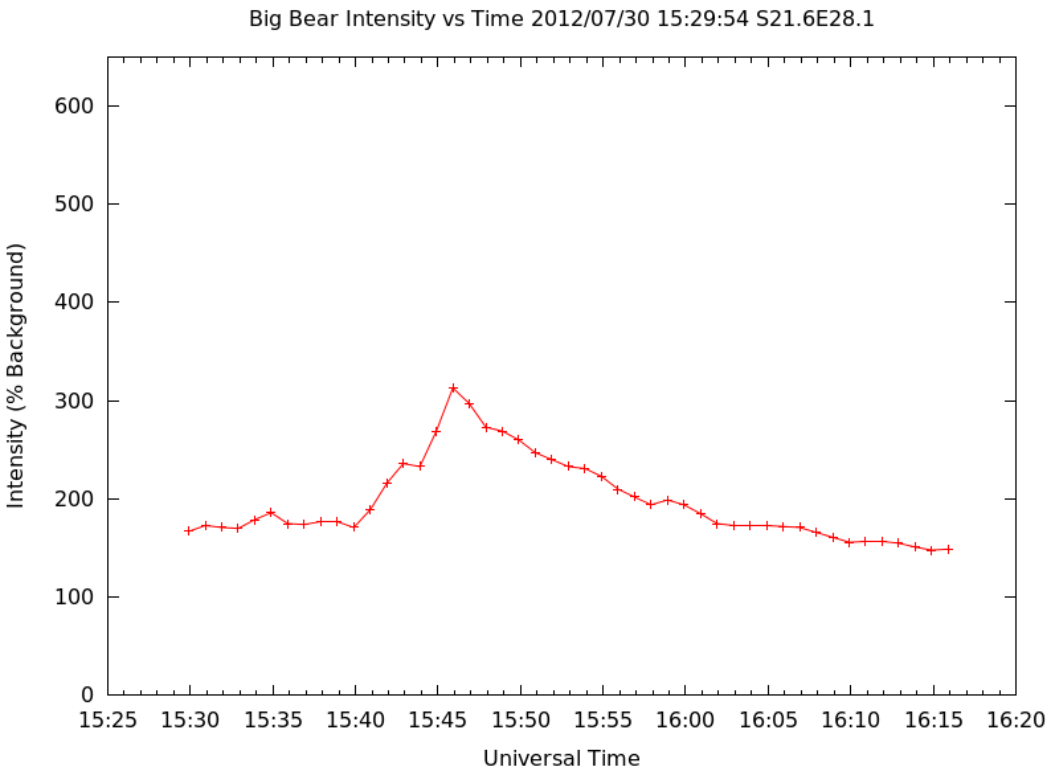



**Figure 5.3**: Intensity vs Time for  Flare Observed at Big Bear in Figure 5.1

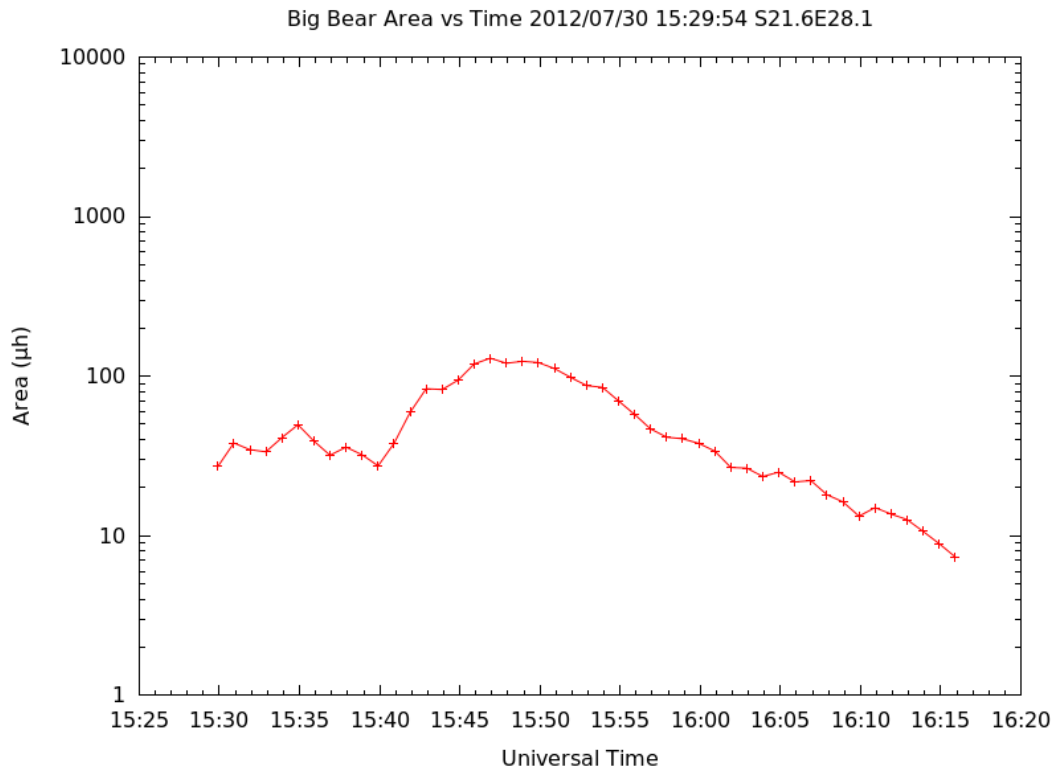



**Figure 5.4**: Area vs Time for  Flare Observed at Big Bear in Figure 5.1

The area peaked a minute after the intensity but, by convention, the area is reported at the time of peak brightness. This is not unusual as flares tend to spread out the longer they progress.

The USAF SOON system only has 64 intensity bins. This was because it used a 6 bit digital system when performing flare analysis, but this should not affect analysis greatly. The general technique described above still applies, except the background Sun is scaled to have a value of 10 instead of 100.

Further details on GONG flare analysis will be given in Chapter 10.

Other parameters that have been studied are flare duration and rise and fall times of flares. These parameters all vary with both area and intensity of the flare, as well as time in the sunspot cycle (Temmer et al., 2001). However such studies are somewhat restricted as they do not have access to the full flare profiles.

### 5.3 Site Discrepancy of Flare Measurements

It has long been recognised that there have been discrepancies between flare reports from different sites (for example Warwick, 1963; Reid, 1963). The USAF SOON provides an opportunity to examine the accuracy and precision of Hα flare measurements. Between 1976 and 1988 five USAF observatories became operational that comprised the Solar Optical Observing Network: Holloman, Palehua, Ramey, Learmonth, and San Vito. Unfortunately, Palehua was closed in the mid 1990s and Ramey in 2002. All of the equipment at these sites is the same, or as close to similar as can be produced, with virtually no modification in four decades of operation.

The NGDC contains an archive of the flare reports from each of these sites from 1982 to 2010. Flares between sites were matched by region and by time. Each flare was assigned a numeric flare category value of 0, 1 or 2 corresponding to the flare Brightness category of F, N or B respectively. A site-by-site comparison was made for each flare by subtracting one site's numeric flare category from another, giving 5 brightness difference categories. A histogram of flare difference was then created for each site pairing. If only one site recorded a flare, this flare was omitted from the analysis. Note that NGDC does not archive the actual percentage intensity of each flare, which is why only flare brightness categories can be analysed.

Figure 5.5 shows each of the site – site comparisons. There were no flares that could be matched between Ramey and Learmonth, and so few flares matched between San Vito and Palehua that these were both omitted. In these plots the zero bin represents the same flare Brightness category being measured at each site. Increasing (both positive and negative) values indicate that the flare was measured differently at the 2 sites.

At first glance it seems that there is a good match between sites, with each pairing giving typically a 70 to 80% agreement. However over half of the flares were faint. If the analysis is restricted to flares that had normal or brilliant brightness, a somewhat different pattern emerges.

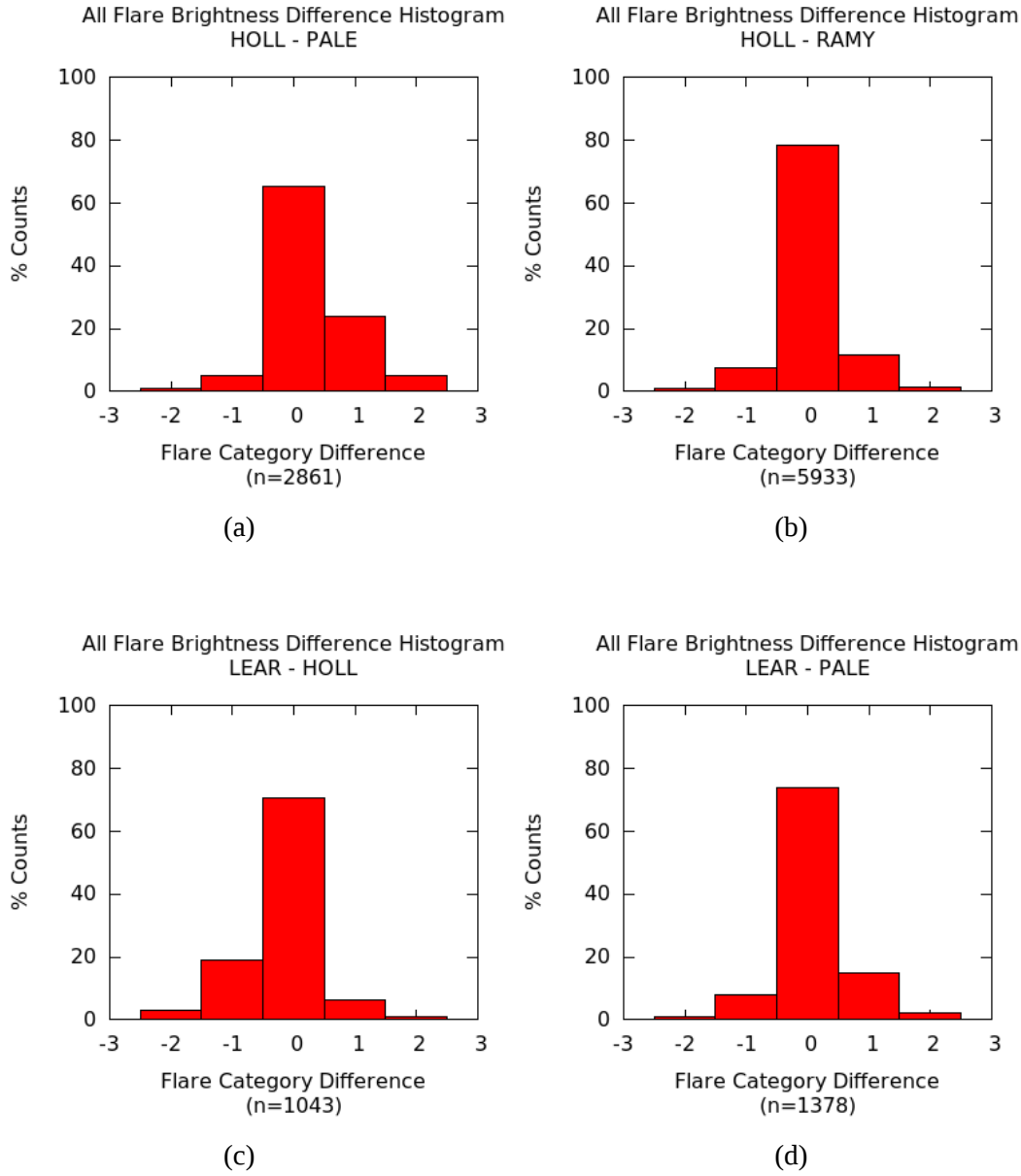



**Figure 5.5 (a-d)**: Comparison of Flare Brightness Between USAF Sites for Flares of all Brightness Categories

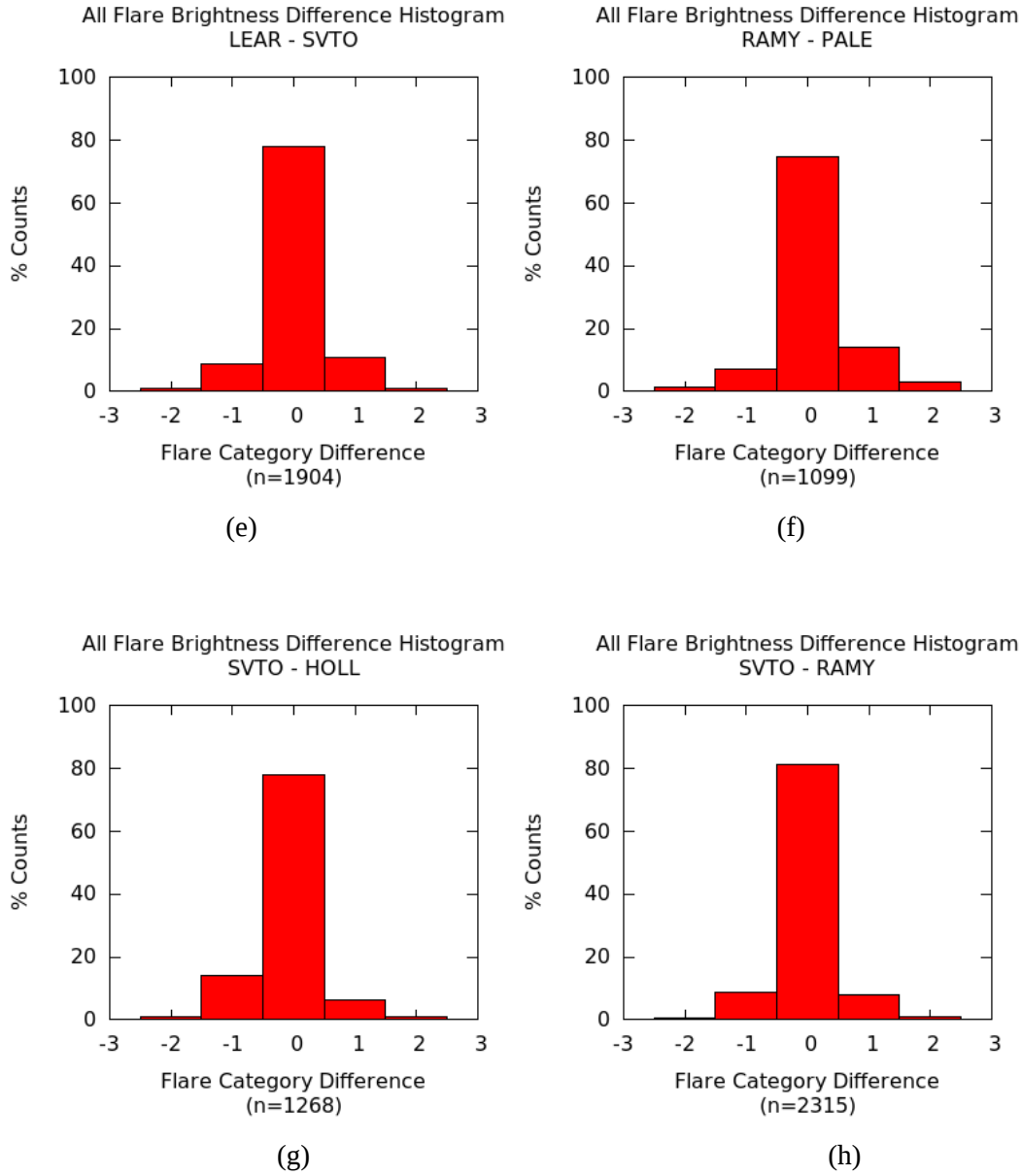



**Figure 5.5 (e-h)**: Comparison of Flare Brightness Between USAF Sites for Flares of all Brightness Categories

Figure 5.6 shows the plots with faint flares excluded. This means that at least one of the two sites in the comparison had to report the flare as normal or brilliant (the other site may have reported it as faint).

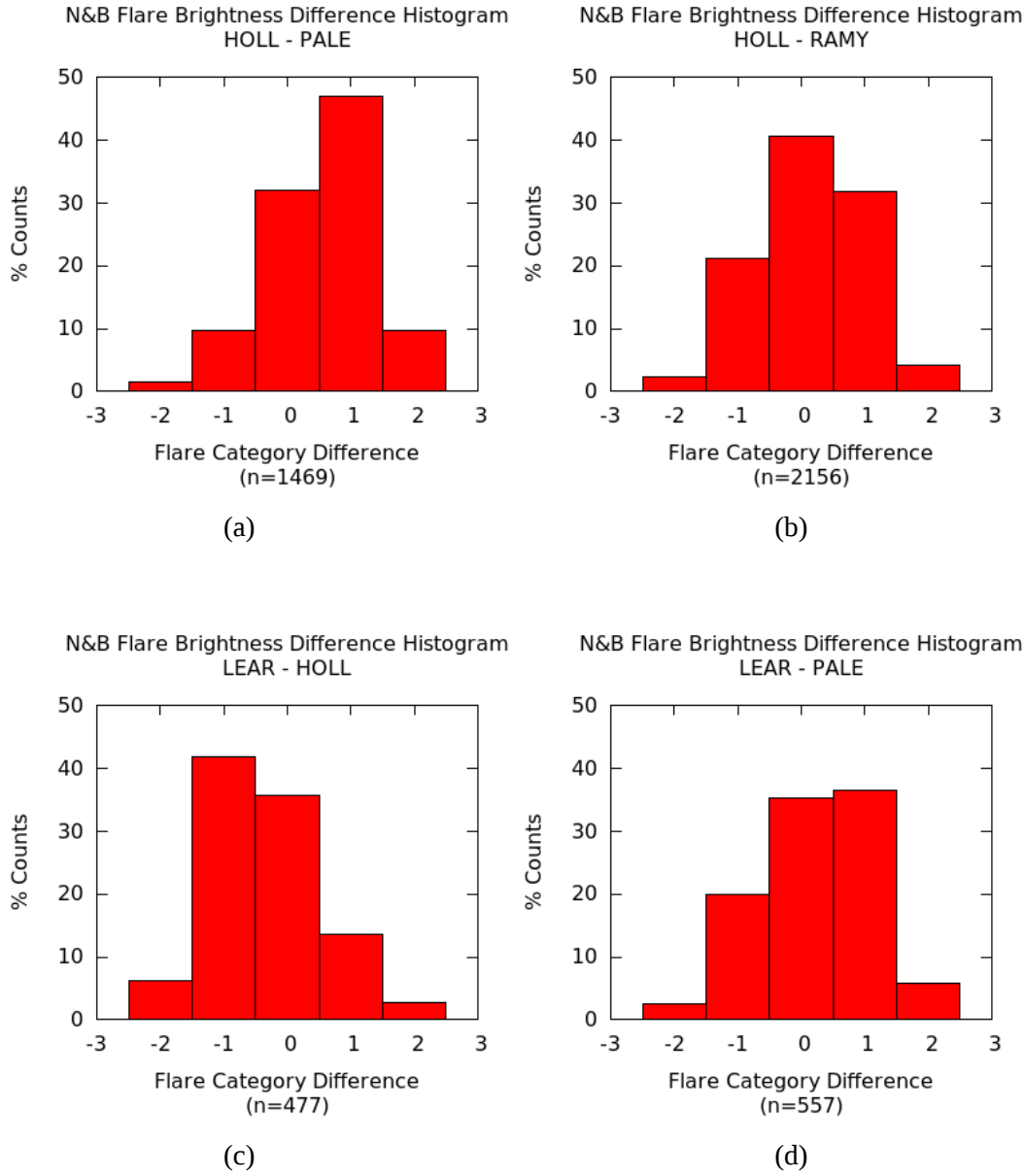



**Figure 5.6 (a-d)**: Comparison of Flare Brightness Categories Between USAF Sites for Normal and Brilliant Flares

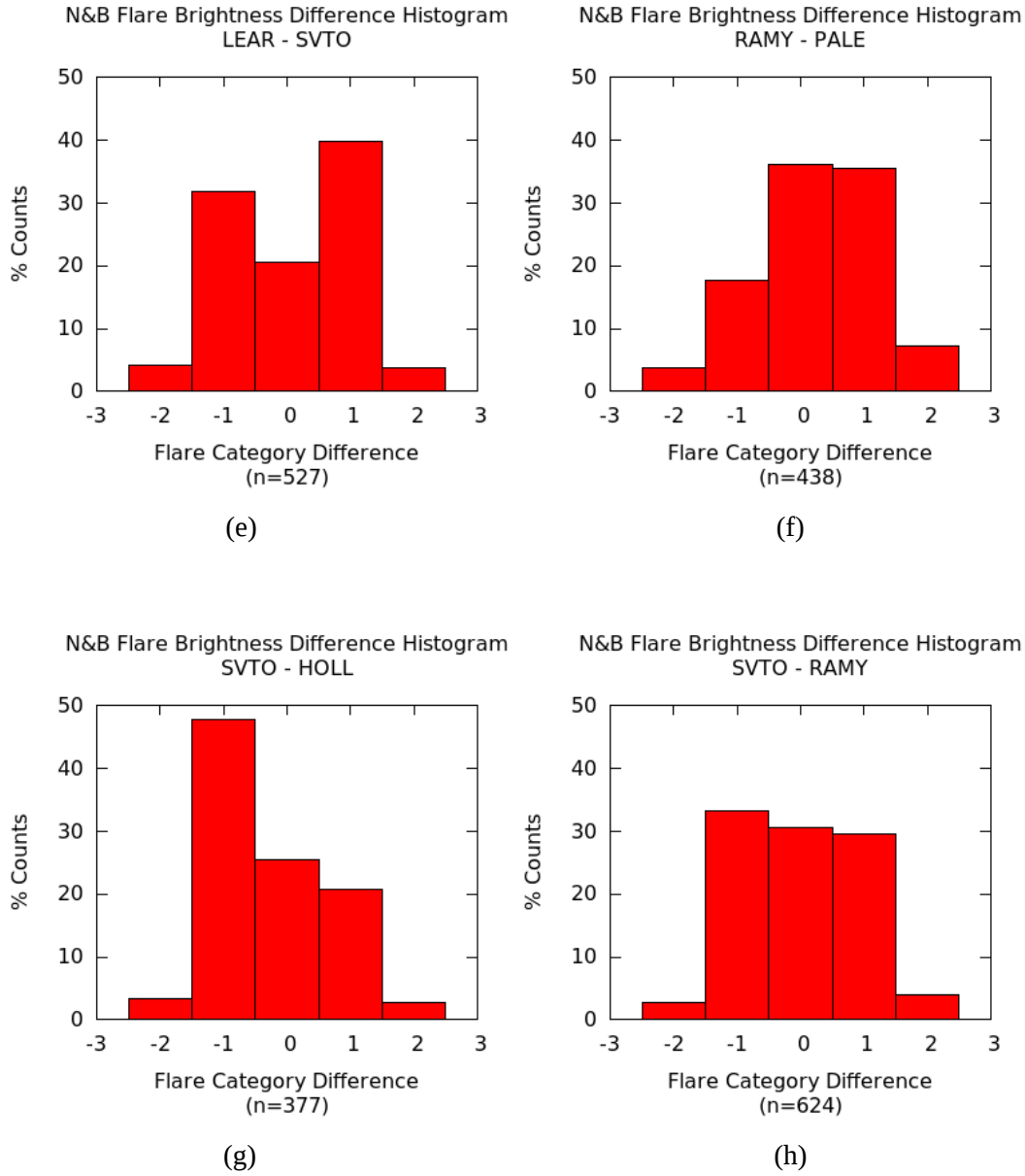



**Figure 5.6 (e-h)**: Comparison of Flare Brightness Categories Between USAF Sites for Normal and Brilliant Flares

For all cases, normal and brilliant flares were assigned the same flare Brightness category less than 50% of the time.. In most cases, the zero difference category (same flare at both sites) was not even the most commonly occurring. The only sites where the same flare category was the most commonly occurring was Holloman – Ramey (figure 5.6b) and Ramey – Palehua (figure 5.6f).

A similar approach is used to compare the importance of flares. For each site pairing, the

flare importance was subtracted, in this case leading to 9 difference categories.

Figure 5.7 shows these results. There again seems to be a good match between sites with typically greater than 80% agreement on Importance class. As with Brightness only the larger area flares are considered. In this case, only sites where at least one flare of Importance 2 or greater was reported were compared. Figure 5.8 shows these results. The first thing to note is that less than 10% of flares fall into these three categories, as compared with about 40% of flares falling into N and B categories for Brightness. Again it is seen that there is less than 50% agreement between sites. In this case, the only sites where the Importance difference class is in a bin other than zero are San Vito – Holloman (figure 5.8g) and San Vito – Ramey (figure 5.8h).

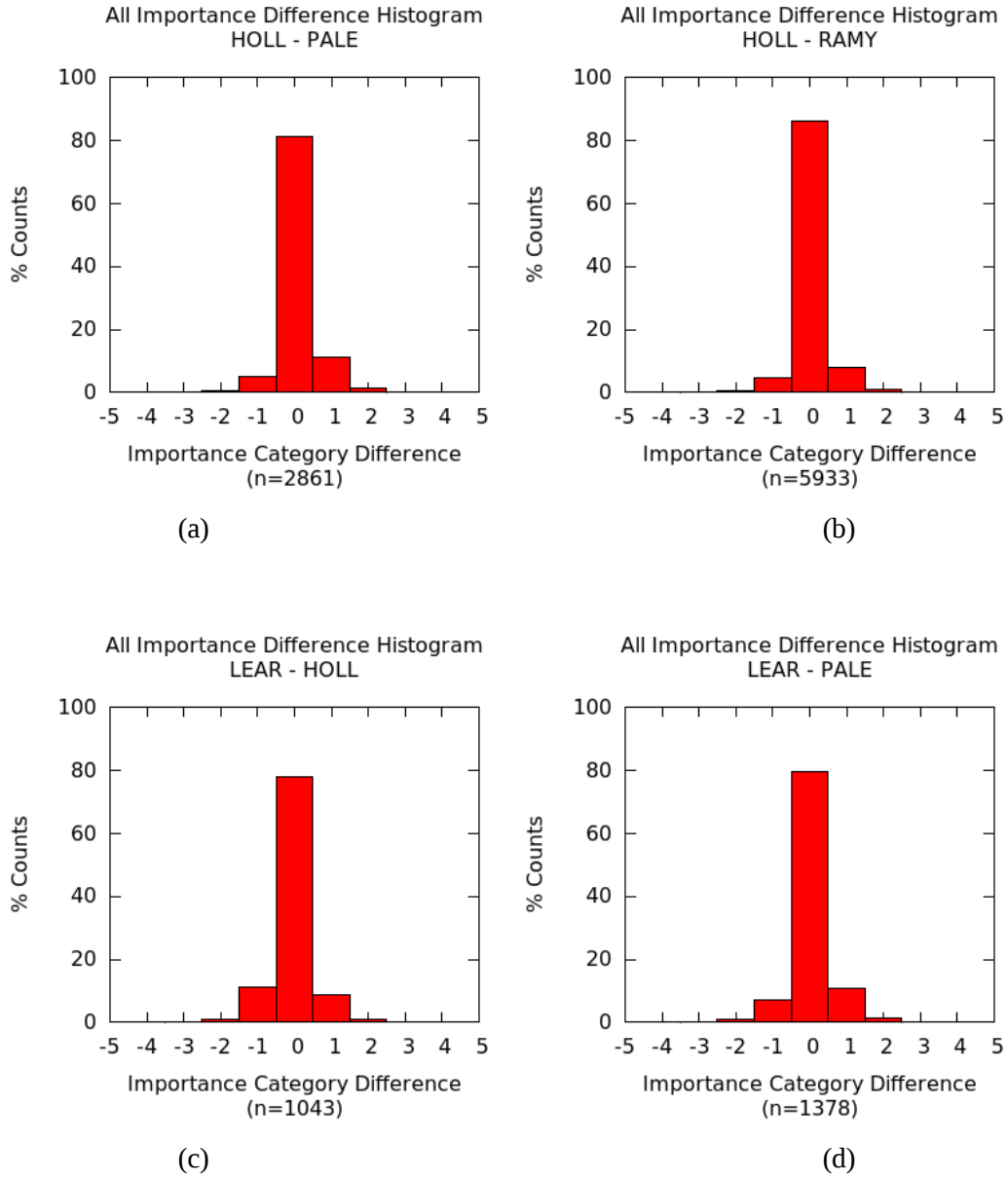
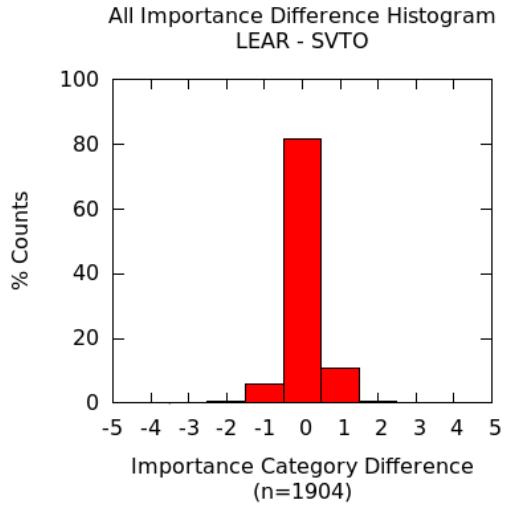
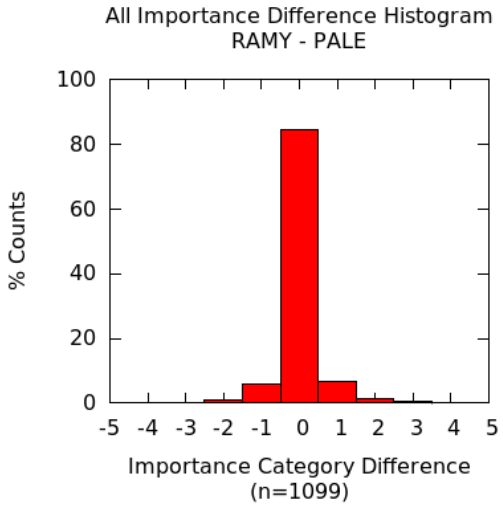


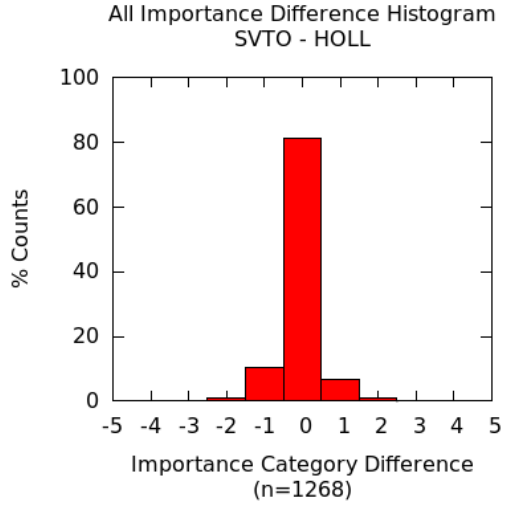

**Figure 5.7 (a-d)**: Comparison of Flare Importance Between USAF Sites for Flares of all Importance Categories



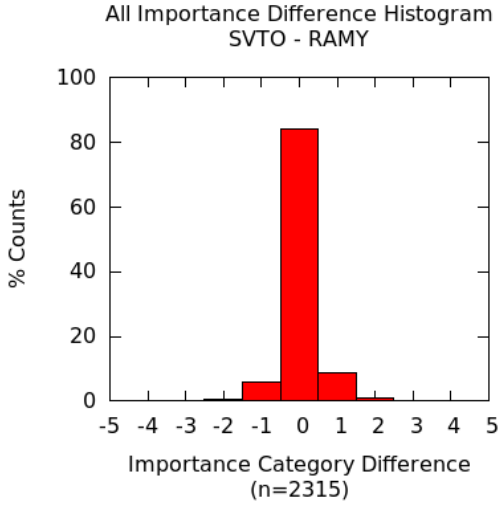

(e)




(f)




(g)




(h)

**Figure 5.7 (e-h)**: Comparison of Flare Importance Between USAF Sites for Flares of all Importance Categories

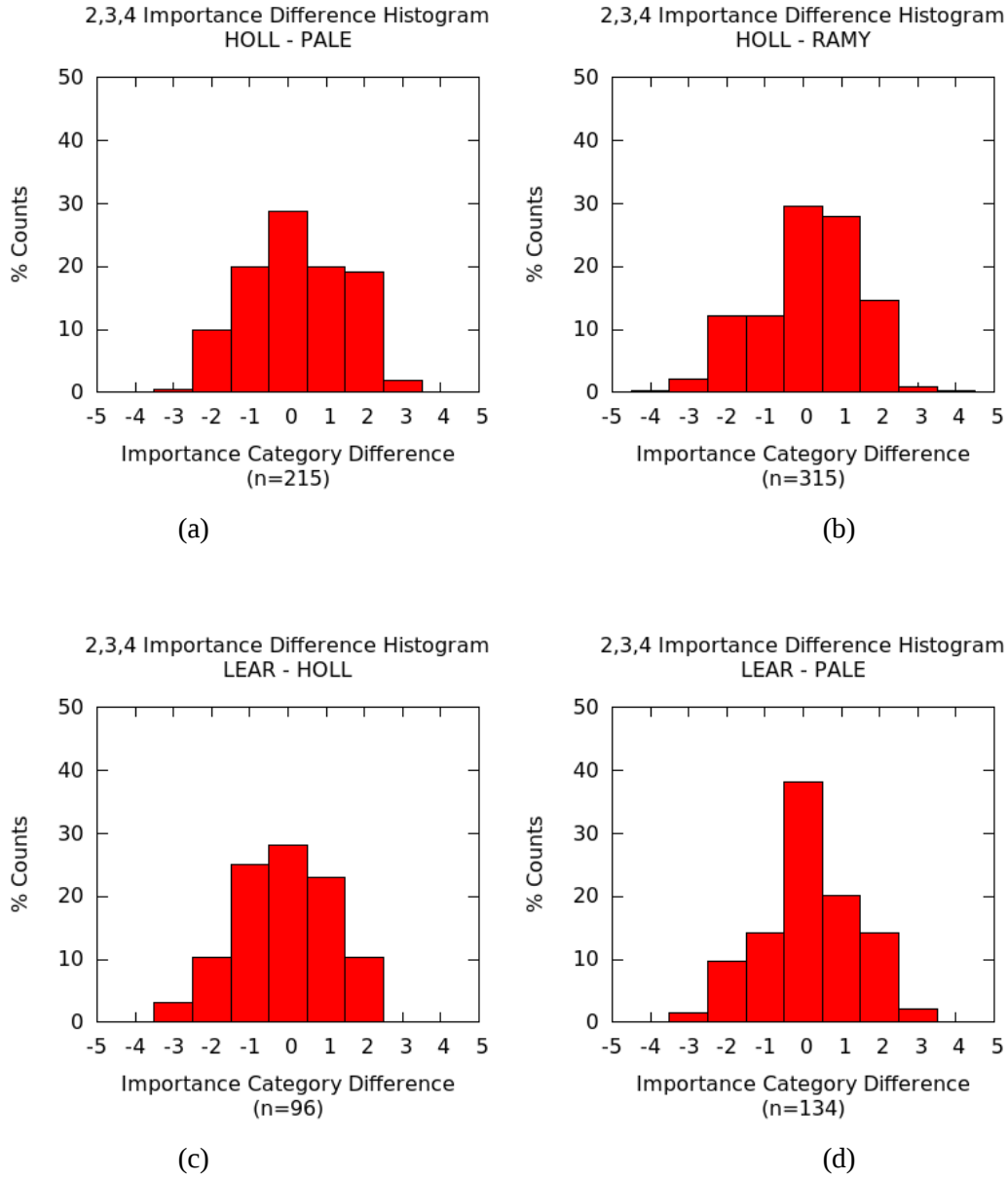



**Figure 5.8 (a-d)**: Comparison of Flare Importance Between USAF Sites for Importance 2, 3 and 4 Flares

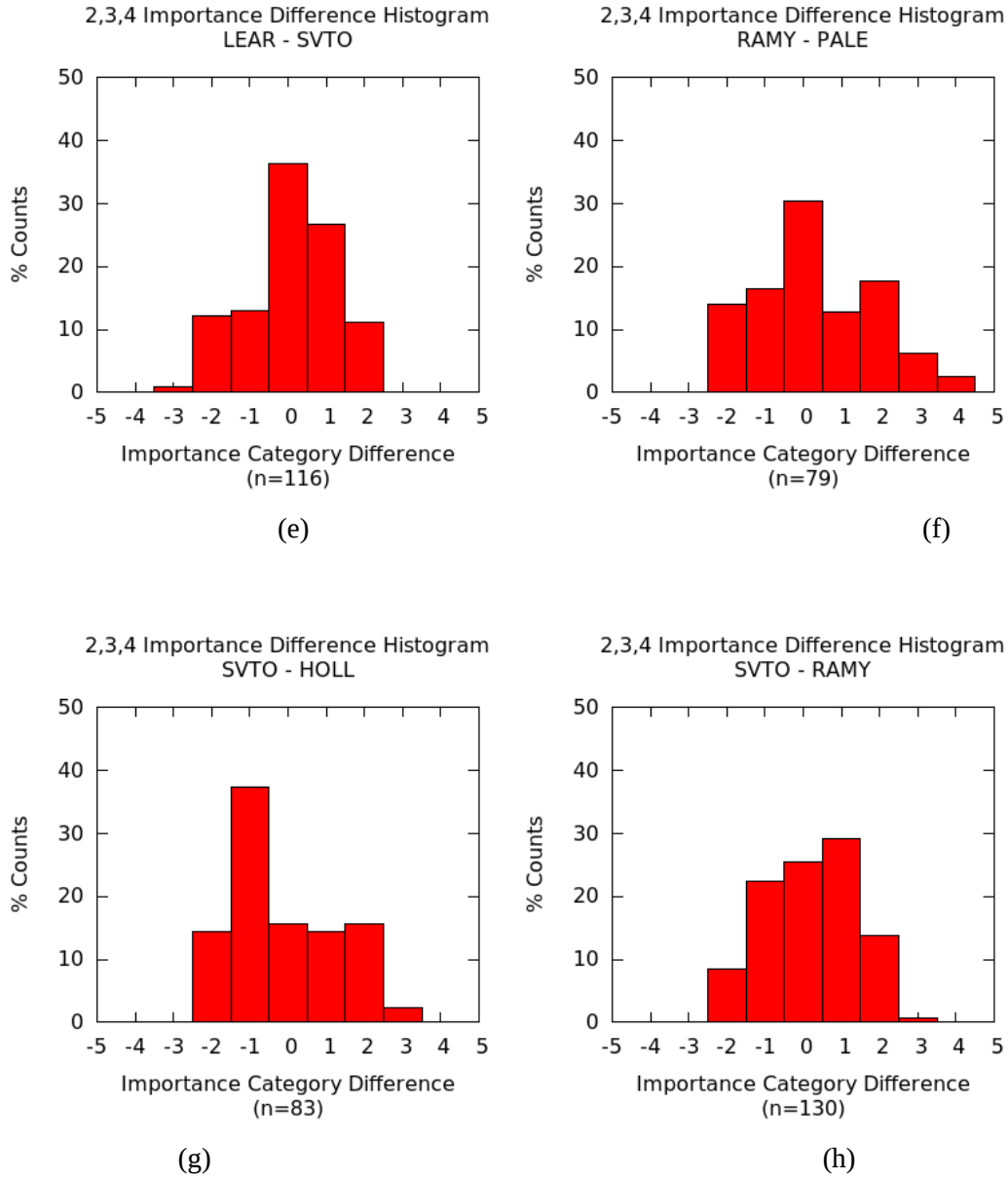



**Figure 5.8 (e-h)**: Comparison of Flare Importance Between USAF Sites for Importance 2, 3 and 4 Flares

What is the explanation for the discrepancies of flare parameters between sites? There may be several, and each is difficult to quantify.

1. Variation in the filters and calibration. All of the sites are supposed to be calibrated the same and weekly checks on the system are made. So, while slight variations maybe expected, it is unlikely to be significantly responsible.

2. Errors in the the limb area correction factors. An analyst, at the time of the flare occurrence, is supposed to select the flare centre. If the analyst does not, then the region centre is chosen as the flare centre by default. If a flare is near the limb and analysts at different sites choose different flare centres, then a different limb area correction factor will be used at each site, as the computer uses the flare centre for correcting the area of the entire flare.

3. ‘Seeing’ effects through the atmosphere may play a role in how well a flare's parameters can be measured. Of note is figure 5.6b , Holloman – Ramey Brightness difference for N & B flares. This was one of the only site pairings where the 0 difference bin was the highest. These two sites were geographically closest, ~28 degrees apart in longitude. Thus Sun-angle difference between these two sites is small when compared to other site pairs. This contrasts with Holloman and Learmonth for example, where the Sun is setting at one site but rising at the other site, so that both sites are looking at large zenith angles through long path lengths of atmosphere and turbulence. In addition, altitudes of sites vary across the SOON network, potentially giving varying seeing conditions.

4. There are potential errors when there are low light levels. The USAF try to maintain solar patrol even when conditions may be substandard. In normal conditions, the flare analysis is done automatically by the computer. However when light levels are too low, the computer ceases patrol and the analyst must manually process a flare. This is done using overlays on a monitor to compute area and by analysing the red and blue shift of a flare. This process is called ‘semi-automatic patrol’. This process takes some time so the end result may not be the flare brightness and importance at peak brightness. Typically semi-automatic patrol occurs within about 30 minutes of sunrise or sunset, when there is high cloud or occasionally dust in the atmosphere. As a result sites separated by large longitudes may both be in semi-automatic patrol simultaneously. Alternately one site may be in automatic patrol while another in semi-automatic, thus the measuring techniques will differ at those times and places.

5. As the actual intensity values or area values for the flare are not available it is not possible to know how many flares were on the edge of a category. For example one site may report a flare as having an intensity 240% above the continuum, where as another site may report an intensity of 260%. Even if there are a significant number of flares like

this, it cannot account for flares that are reported more than one category apart.

6. The SOON system will not end a flare if there is a break of 45 minutes or less in observations. If something occurs that would interrupt observations (clouds, equipment problems), and observations are resumed within 45 minutes, and a flare at the same location is progress, then the SOON system assumes that it is the same flare. Therefore, the peak of a flare may not actually be observed by a particular site under those circumstances.

The type of patrol (semi-automatic or automatic) is stated in the flare reports that the USAF observatories submit. There is also a field in the reports for seeing conditions. Both of these parameters are not archived by NGDC, so it is difficult to determine how much effect these have on inter-site discrepancies.

There are different types of calibration that are performed on the SOON system. The first is the real time image calibration. This is performed once every 30 seconds. A region is set (called MAGR) that is near the centre of the disk that is to have no plage, spots or filaments in it. Finding such a region can be challenging during times of high activity with many such regions on the Sun. This is used to calibrate other regions brightness and adjust the gain of the system. The dark level is set by a strip of black tape down the side of the camera. This is just electrical tape, and unfortunately, there are no guarantees that this tape is the same across all sites (or eras of observation).

Secondly, there are weekly and monthly checks of the system to ensure equipment is within specified USAF guidelines. During these calibrations, various pieces of equipment (for example amplifiers, cameras or filters) may be adjusted.

On rare occasions a full calibration is done using an epidiascope. Unfortunately, the details of this procedure are not available to the general public (personal communication, Major Haley Homan). It is unclear whether these are the same across sites. In 30 years this type of calibration has only been performed twice at Learmonth Solar Observatory (personal communication, John Kennewell).

### 5.4 Summary and Conclusions

Hα flares have been observed since the early 1900's. Flares are categorised by their area

(importance) and intensity (brightness). It was quickly realised that different sites were reporting different classifications for the same flares. There may be different reasons for these discrepancies including seeing conditions, different quality filters, area correction factor errors and the way different observatories manage interrupted observations.

The USAF has operated its SOON system for 3 decades at 5 different sites around the world with close to identical equipment. Even this comprehensive network, with the brighter flares, cannot get agreement on either area or brightness of flares.

The GONG installation of a Hα system in 2010 gives observers another global network to compare against. In addition, unlike the USAF SOON system which still requires some manual analysis, the GONG system can be fully automated, which will hopefully provide more consistent flare results. This is examined in detail in chapter 10.

# 6. Image Preprocessing

## 6.1 Introduction

Digital image preprocessing of solar imagery is required to assure the images of a particular type (Hα, magnetogram, intensity) are all consistent over a long time span. Pre-processing also minimises errors in subsequent analysis of the images.

In general the disk centre and disk radius must be determined, the disk brightness must be brought to a consistent level (image normalisation), and limb darkening correction applied to produce consistent brightness across the disk. Translation of the disk may be required in some images for future processing. Care must be taken throughout this process not to distort or eliminate features that may be of interest.

These pre-processing steps are described in this section.

## 6.2 File Formats

All of the images used are in standard Flexible Image Transport System (FITS) formats (http://fits.gsfc.nasa.gov/fits_primer.html). The Hα images are 2048 by 2048 pixels in size and the intensity continuum images (usually just referred to as intensity images) and magnetogram images are 860 by 860 pixels in size. The pixel values in the magnetogram and intensity images are signed 16 bit integers, where as for Hα they are 16 bit unsigned integers.

In summary, the FITS file contains a header portion and a data portion. Each individual header field is 80 characters long (padded with spaces if necessary) and the full header is a multiple of 2880 bytes (36 fields). The data portion then follows, the size defined by the BITPIX, NAXIS, NAXIS1, NAXIS2 keywords (note there can be up to 999 axes for multi-dimensional data sets).

The image data was obtained from the GONG servers. The specific locations are:

| | |
|---|---|
| Intensity Images: | ftp://gong2.nso.edu/vmb/archive/izi/ |
| Magnetogram Images: | ftp://gong2.nso.edu/vmb/archive/bzi/ |
| Hα Images: | ftp://gong2.nso.edu/HA/haf/ |

### 6.3 Disk Centre and Radius Algorithms

Fortunately, the solar disk centre and radius are already specified in the FITS headers for the intensity and magnetogram images. There is minimal processing on the Hα images, and the disk radius and centre needs to be found. The algorithm below was suggested by Jack Harvey of GONG (personal communication).

The average pixel brightness at the centre of the disk in the raw Hα image was originally set at approximately 2400 by GONG. After 2011, this value was changed to about 4000 and stored in the FITS file header. The value at the centre of the disk will be referred to as the disk brightness ($B_d$). Even with limb darkening, disk pixel values are usually brighter than 1/3 $B_d$, and prominences located off the disk are typically darker than this value.

To find the disk centre, first the pixel positions of all of the pixels brighter than 1/3 $B_d$ are summed as follows:

$$X_s = \begin{cases} X_s + x & if\ p(x,y) > B_d/3 \\ X_s & if\ p(x,y) \leqslant B_d/3 \end{cases}, \qquad \text{Eqn 6.1}$$

and

$$Y_s = \begin{cases} Y_s + y & if\ p(x,y) > B_d/3 \\ Y_s & if\ p(x,y) \leqslant B_d/3 \end{cases}, \qquad \text{Eqn 6.2}$$

where $X_s$ and $Y_s$ are the sums of the positions on the disk, x and y is the position of the current pixel and p(x,y) is the brightness of the current pixel.

In addition, each time the above criteria are met, the value $n_{pix}$ is incremented, giving the total number of pixels on the disk. $n_{pix}$ is also the area of the disk in square pixels.

Now the centre of the disk is found simply by:

$$X_c = \frac{X_s}{n_{pix}}, \qquad \text{Eqn 6.3}$$

$$Y_c = \frac{Y_s}{n_{pix}}. \qquad \text{Enq 6.4}$$

Since $n_{pix}$ is the area in pixels of the disk, the radius of the disk is found as follows:

$$r = \sqrt{\frac{n_{pix}}{\pi}} \quad . \qquad \text{Eqn 6.5}$$

These values can be used as an initial check to see if the image is of good quality. If $X_c$, $Y_c$, and r are within a certain predetermined range then the image can be considered good, otherwise it is considered bad. Initially the range is set so that the centre of the disk is within the bounds of where a perfectly centred solar disk would be, that is the range is dependent on the disk radius for each image.

### 6.4 Image Translation

In the case of intensity and magnetogram images, it is desirable to have the disk centres aligned. This simplifies later processing. Because these images are acquired at the same time, the disk radius for both images will be the same. But due to different optics for the magnetogram whereby the light goes through an additional set of polaroids, the image may be offset from the white light image.

To align the two images, each pixel in the magnetogram and intensity image is moved horizontally and vertically so that the centre of the disk is at the centre of the image (in the case of intensity and magnetogram images this would be at location [430,430]). If a pixel ends up being off the image then it is discarded. As a result of this translation, there will be blank pixels on one or two edges of the image, which are filled with zeros.

### 6.5 Image Normalisation

Before any analysis can be performed the intensity images must be normalised. The GONG FITS files have individual pixel values that are signed 16 bit integers. These are converted into unsigned integers by subtracting the minimum pixel value from all the pixels. A histogram is then generated, shown in figure 6.1. This histogram is only of the solar disk. The dark sky pixels dominate the image and have been removed as they are not needed for further processing. The range here is potentially from 0 to 65535 (16 bit). Because dark sky pixels have been omitted, there are no pixels with brightness less than 15000.

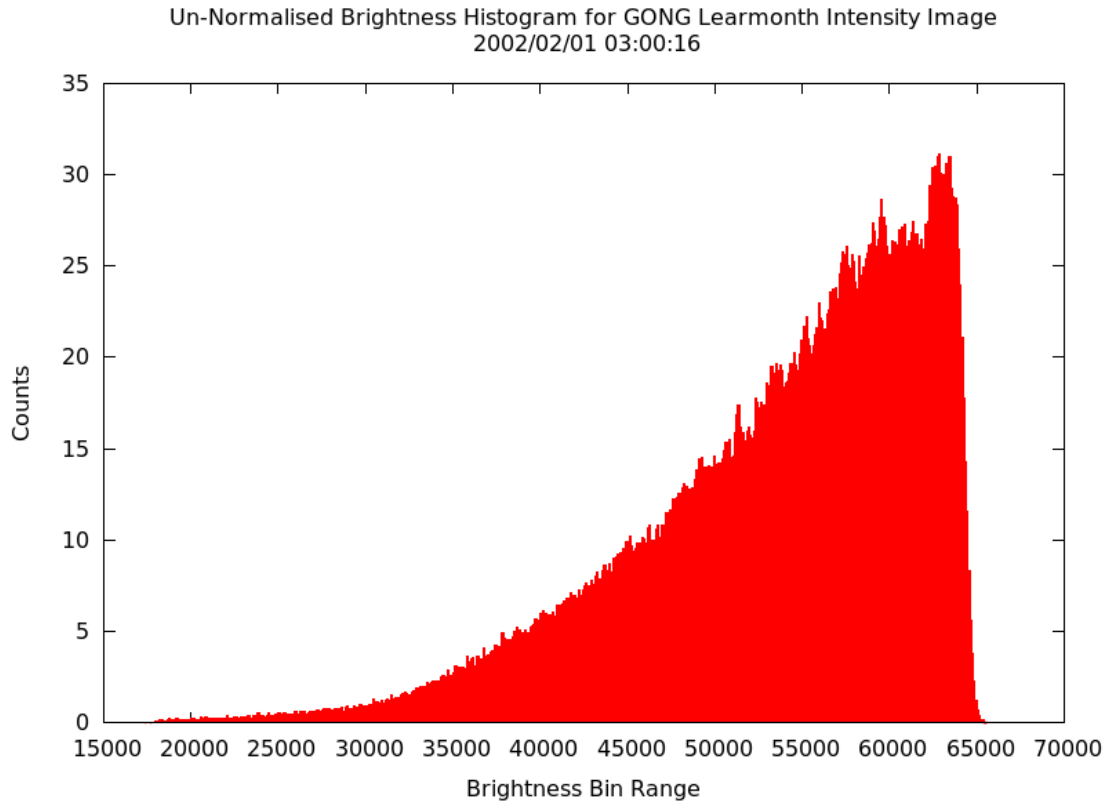



**Figure 6.1**: Sample Un-normalised Brightness Histogram of a GONG Intensity Image

To normalise the image, the median brightness bin of the histogram is found. Each pixel is then scaled using the following equation:

$$p_n = p_o \times \frac{200}{m} \quad , \qquad \text{Eqn 6.6}$$

where $p_n$ is the normalised pixel intensity value, $p_o$ is the original pixel intensity image and $m$ is the median of the disk histogram. The factor of 200 is used so that it is rare that even the brightest pixel in the image will exceed 255. If any pixel, once adjusted in this manner, exceeds a value of 255, then it is set to 255. This is to avoid saturation when plotting a grey-scale image. An example normalised image is shown in figure 6.2.

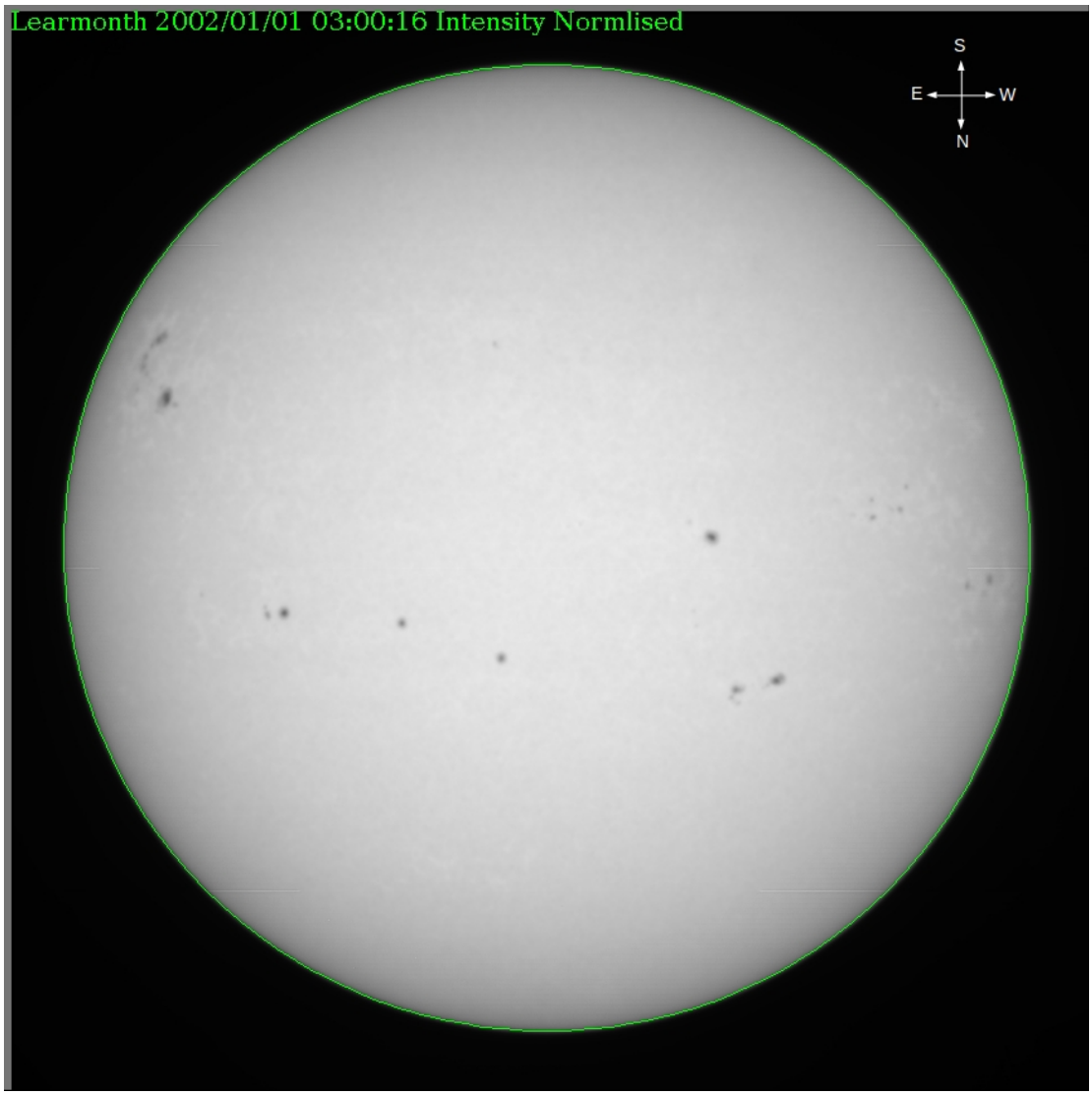



**Figure 6.2**: Sample Normalised GONG intensity Image. The green circle shows the edge of the solar disk.

Figure 6.2 clearly shows sunspots scattered across the disk, but due to limb darkening this image still cannot be used for sunspot analysis.

### 6.6 Limb Darkening Correction

Limb darkening occurs due to the geometry of an emitting sphere. Most of the radiation tends to travel out at right angles from the surface, but other factors such as the solar atmosphere also affect the brightness at a point on the disk (Bhatnagar and Livingston, 2005). Thus, rather than a simple trigonometric function, a polynomial is used to fit a limb darkening curve to the disk. It is assumed that the brightness of the disk is radially

symmetrical. This assumption is because the GONG images are flat fielded and dark framed, thus removing most of effects of the camera. In analysing images, there were no cases where the images weren't radially symmetrical.

The first stage is to find the brightness at a given solar radius (referred to as Rv). For a given Rv, the brightness of 360 pixels in one degree steps are taken, and the median of these points is calculated. This procedure minimises the effect of a pixel being selected on a sunspot. This process is performed for radial positions from 0 Rv to 0.95 Rv in 0.05 increments. The final position is at 0.98 Rv, giving 20 points in total.

Using this radius – brightness data, a $5^{th}$ order polynomial is fitted to produce a limb darkening function. This function is different for each image and needs to be determined every time an image is acquired. A sample plot is shown in figure 6.3 with the limb darkening function. This plot was generated from the normalised image shown in figure 6.2.

This method is adapted from the work of Joshi et al. (2010).

Once the limb darkening function is found, it is applied to each pixel out to 0.98 Rv. For an individual normalised pixel at radius $r$, $p_n$, the limb corrected value for that pixel, $p_l$, is given by equation 6.7:

$$p_l = \frac{p_n}{f(r)} \qquad \text{Eq 6.7}$$

Figure 6.4 shows the result of this procedure and shows the brightening of the limb in comparison with figure 6.2.

The limb of the disk starts to brighten up again. This is due to the limb correction algorithm not working perfectly at the extreme edge of the Sun. For this reason sunspot analysis is only performed out to 0.95 Rv.

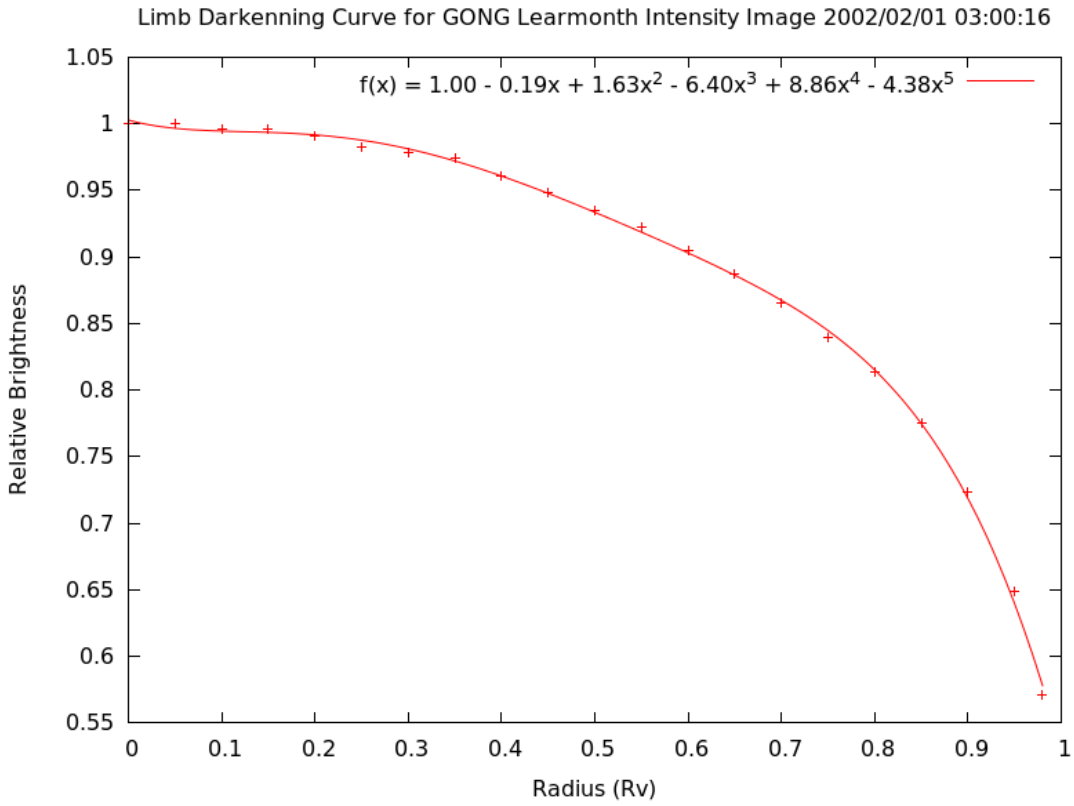



**Figure 6.3**: Sample Limb Darkening Curve from GONG Intensity Image in Figure 6.2

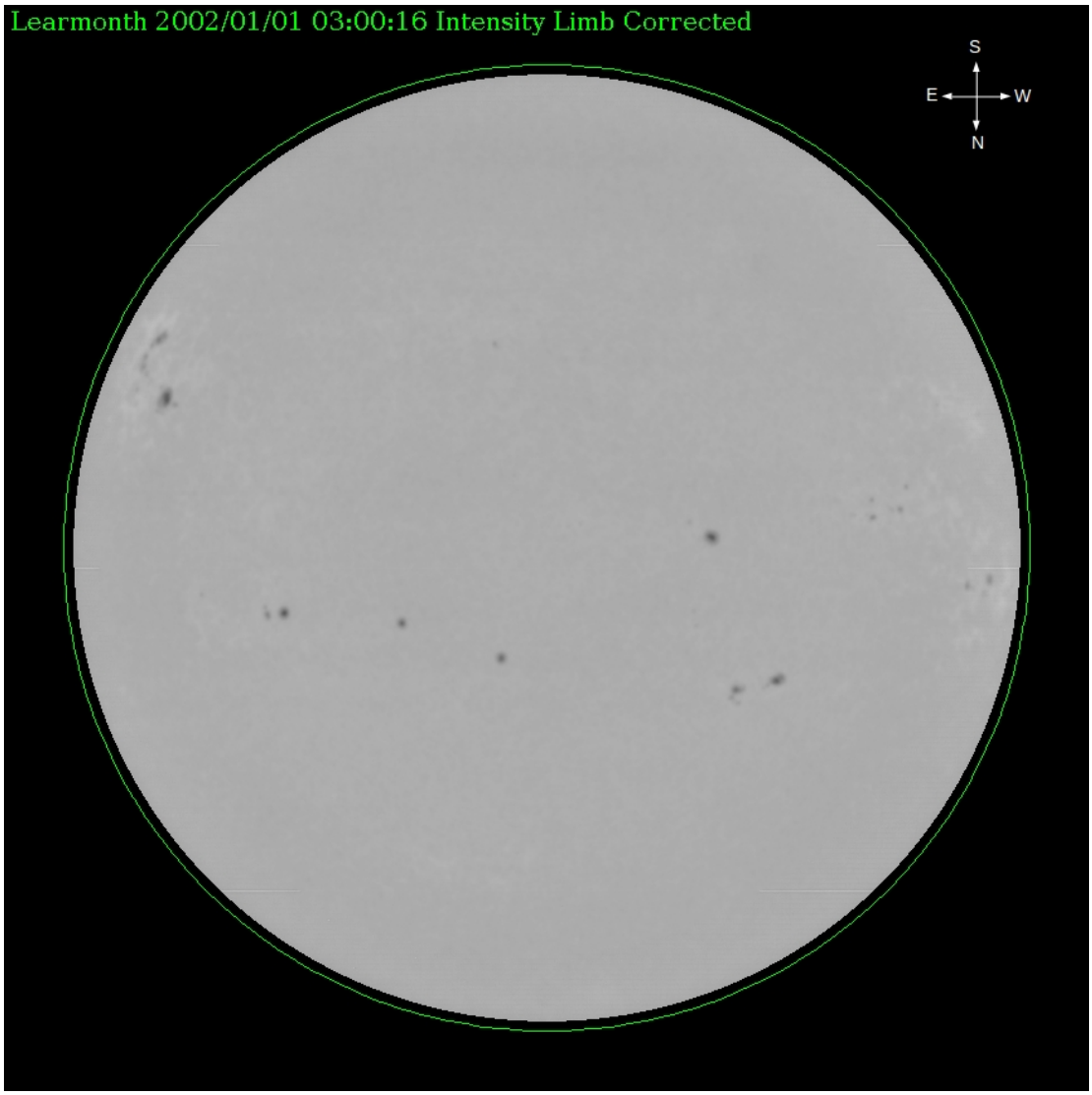



**Figure 6.4**: Sample Limb Corrected GONG Intensity Image from that in Figure 6.2. The green circle shows where the actual limb of the sun would be.

### 6.7 Summary

This section has detailed the procedures required to preprocess GONG magnetogram, intensity and Hα images. The critical components are the normalisation and limb darkening correction. Limb normalisation is performed by setting the brightness of the centre of the disk to a value of 200 and scaling the image brightness appropriately. Limb darkening correction is performed by fitting a $5^{th}$ order polynomial brightness vs radius curve to the solar disk (the assumption being that the disk is radially symmetrical). This polynomial is used to correct the limb darkening.

These algorithms have produced consistent bright images with little distortion for the 13 years of data that were analysed.

# 7. Automated Sunspot Region Detection and Analysis

## 7.1 Introduction

The aim of this project is to simultaneously produce sunspot number, sunspot areas, sunspot classification and magnetic classification in one package using simple, well defined physical parameters with straightforward and well tested algorithms. In this way it is hoped that a set of consistent observations over diverse temporal scales can be made. Whilst GONG data were used for this analysis, these techniques may be used on any well calibrated data sets. Of course, different results will be obtained due to difference in equipment (for example, higher resolution cameras will be able to resolve smaller sunspots).

As discussed in chapter 4, there do seem to be some errors in measurement of sunspots. Certainly the ratios between parameters from different observing sites seem to change their relationships across sunspot cycles. By using an automated procedure, both accuracy and consistency in observations can be improved. Note, the aim is not to replicate the exact results of manual analysis, but produce results of the same standard and accuracy across sunspot cycles and observing sites.

Solar analysis has been at least partially performed with automation since the early 1970s (Reid and Vorhaben, 1971). The SOON Hα system used some computerised procedures for flare analysis, and the RSTN system used automated techniques to detect and classify bursts. But even here there was significant user input to verify that the computer was producing results correctly. It wasn't until the 1990's that a concerted effort was started to try to analyse sunspots automatically.

Pettauer and Brandt (1997) developed two techniques for measuring sunspot areas automatically. The first was to look at a cumulative histogram of the sunspot region to isolate dark pixels, and the second was to use brightness gradients in a region to identify rapidly changing brightness and isolate sunspots as a result. However these early techniques still required an analyst to initially isolate a region. At around the same time, Győri (1997) developed a method for finding sunspot areas that used a combination of intensity thresholding and gradient changes to isolate sunspots.

Curto et al. (2008) used a series of morphological image operators in order to locate sunspots. They used various image operators in a specified order over the entire disk in

order to isolate sunspot regions.

Colak and Qahwaji (2008) used both magnetogram and intensity images to identify sunspot regions and then a neural network to classify sunspot regions.

In the remainder of this chapter, some of the techniques described in the above papers are used and adapted to produce a set of algorithms that can output parameters that describe individual sunspot regions. These are then compared with the records archived at NGDC and SIDC. Chapter 8 will examine the use of automatically determined parameters for predicting x-ray flares.

## 7.2 Algorithm

A flow chart of the sunspot region analysis algorithm is shown in figure 7.1. This algorithm is run once for each [intensity image – magnetogram] pair by a script. This script passes the appropriate image files names to the algorithm.

### 7.2.1 Image Normalisation and Limb Darkening Correction

The intensity images were preprocessed as described in section 6. This is to ensure consistency across the entire range of the data set.

### 7.2.2 Magnetogram Thresholding

As sunspots are always associated with areas of significant magnetic field, the magnetogram is used to remove potential noise from the intensity image. Before this can be achieved the magnetogram must be analysed to find areas of positive or magnetic field. This procedure is similar to work done by Colak and Qahwaji (2008).

Figure 7.2 shows a GONG magnetogram. This image is then processed in the following way: if a pixel has a magnetic field of greater than 35 Gauss then it assigned a value of +1; if a magnetic field less than -35 Gauss then it is a assigned a value of -1, and all other pixels assigned a value of 0.

Once a pixel is determined as being positive or negative a search is made of the limb corrected intensity image to find pixels that are part of sunspots.

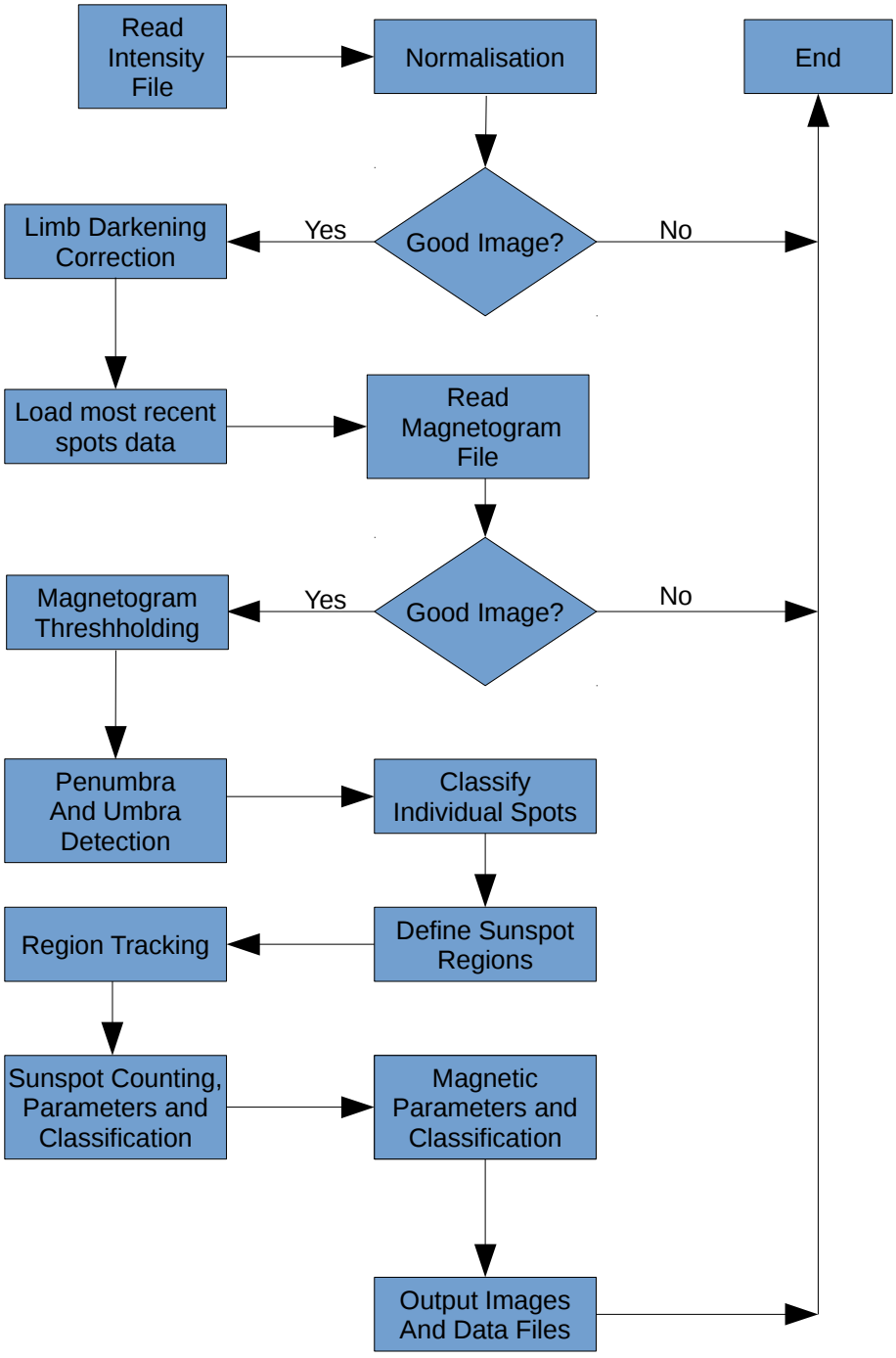



**Figure 7.1**: Schematic of Sunspot Region Analysis Algorithm

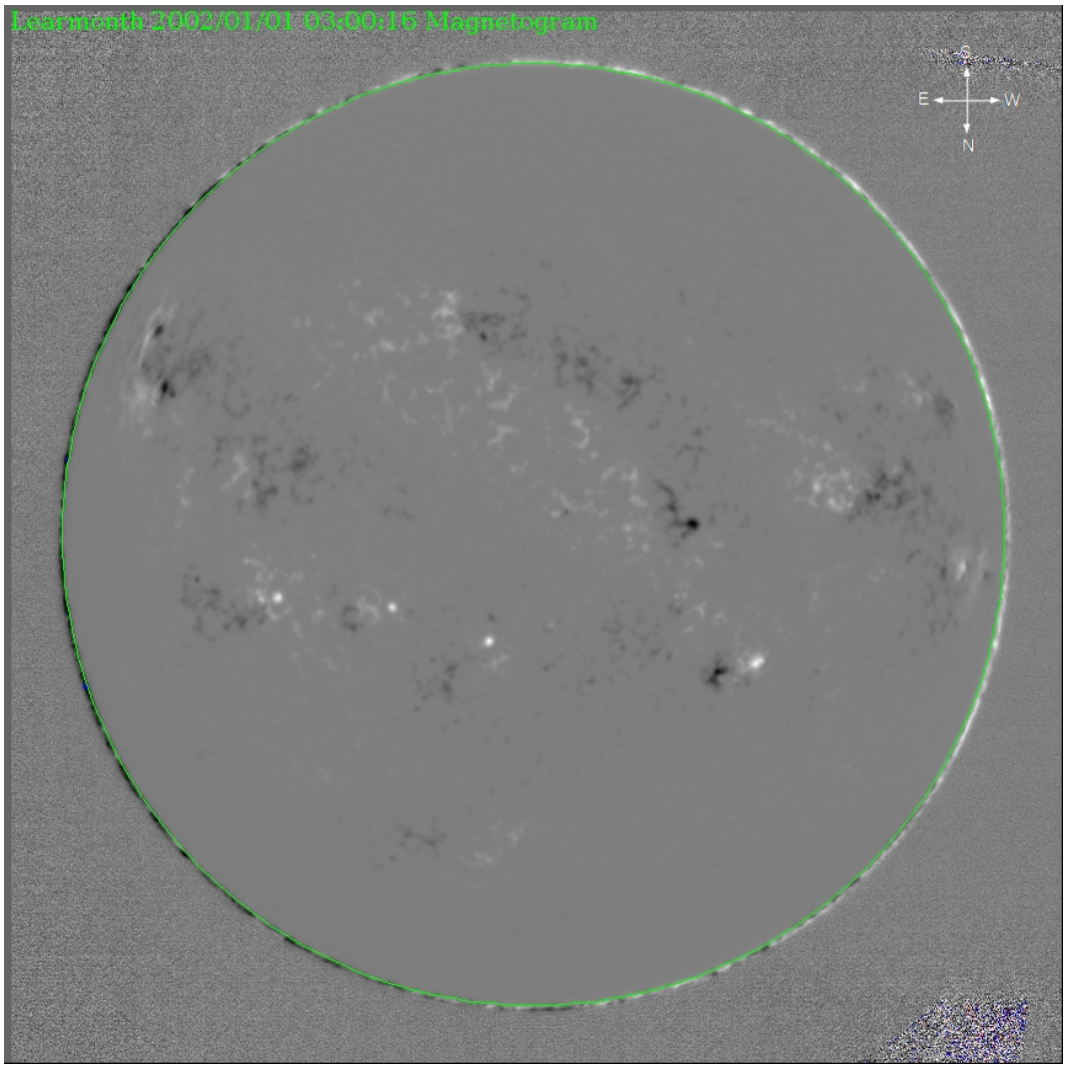



**Figure 7.2**: Sample GONG Magnetogram. Here, the black areas represent negative magnetic field (into the Sun) and white positive magnetic field (out of the Sun). The green circle shows the limb of the disk in the intensity image.

### 7.2.3 Sunspot Detection

In section 7.2.2 thresholding on the magnetogram image was performed. In order to find sunspots dark pixels on the intensity image need to be found. The sunspot detection portion of the algorithm is shown in figure 7.3. This analysis is performed on the limb corrected image as described in chapter 6.

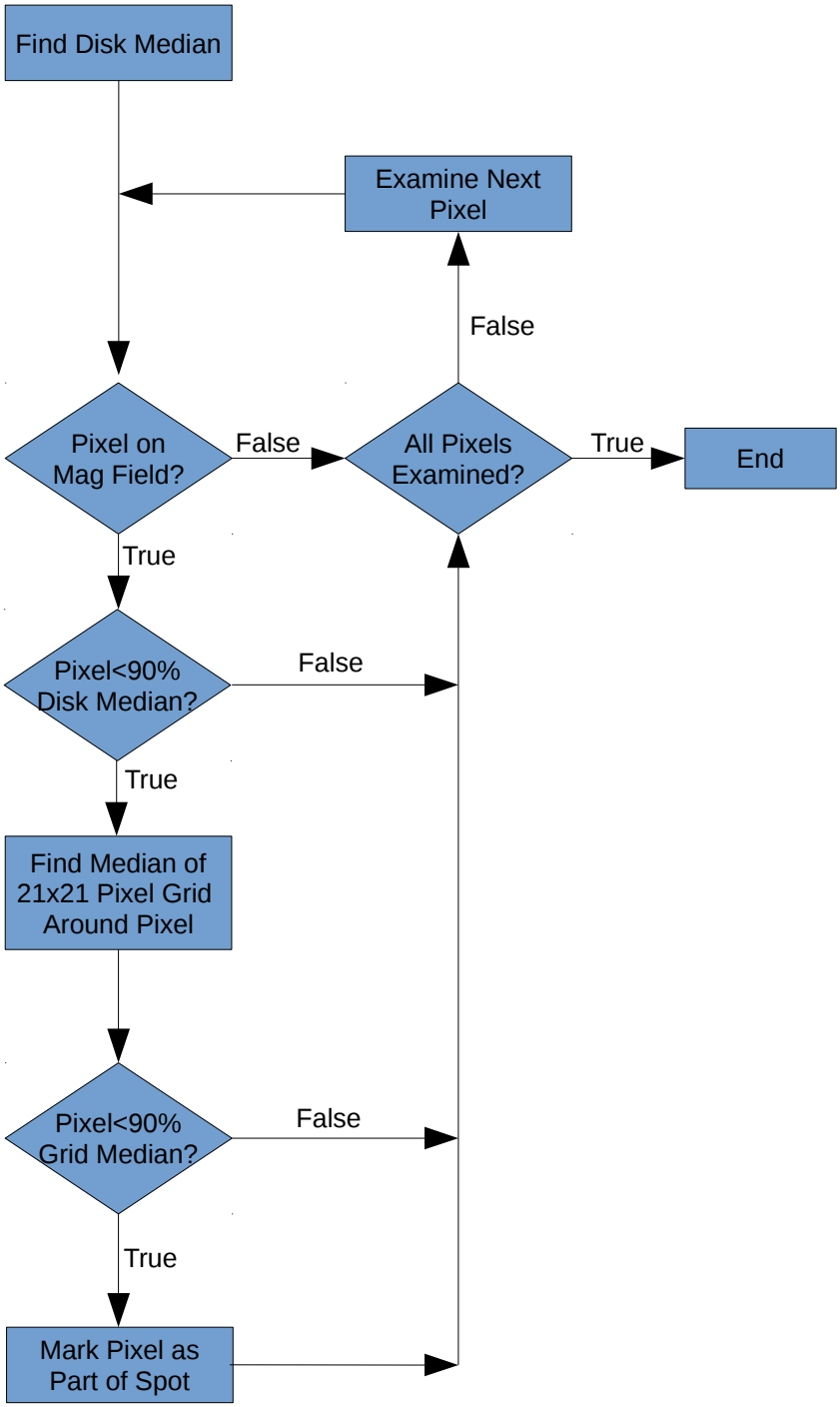



**Figure 7.3**: Sunspot Detection Algorithm

In order to determine if a limb corrected image pixel is part of a sunspot, some thresholds are used to determine how dark the pixel is. The first test is if the pixel is darker than 90% of the entire disk median in the limb corrected image as shown in equation 7.1.

$$p_l < 0.9 \times m_l \quad , \qquad \text{Eqn 7.1}$$

where $p_l$ is a limb corrected pixel brightness and $m_l$ is the median brightness of the limb corrected disk.

If equation 7.1 is true then a 21 by 21 pixel grid with the pixel of interest at the centre is constructed. If the pixel brightness is less than 95% of the median brightness of this 21 by 21 pixel grid then its considered to be part of a sunspot.

The procedure is performed only if a pixel is determined to be part of significant magnetic field as determined in section 7.2.2. If a sunspot pixel in an area of magnetic field is found then the area of magnetic field is kept, otherwise it is rejected. Two images are produced by the combination of step in sections 7.2.2 and 7.2.3: a magnetic map shown in figure 7.4 and sunspot map shown in figure 7.5. In practice, the two algorithms run simultaneously and reject noise in both images as only the magnetic field that occurs with sunspots is of interest.

Magnetic thresholding, and therefore sunspot detection, is restricted to within 0.95 $R_v$ due to the limb correction algorithm not working particularly well at extreme $R_v$.

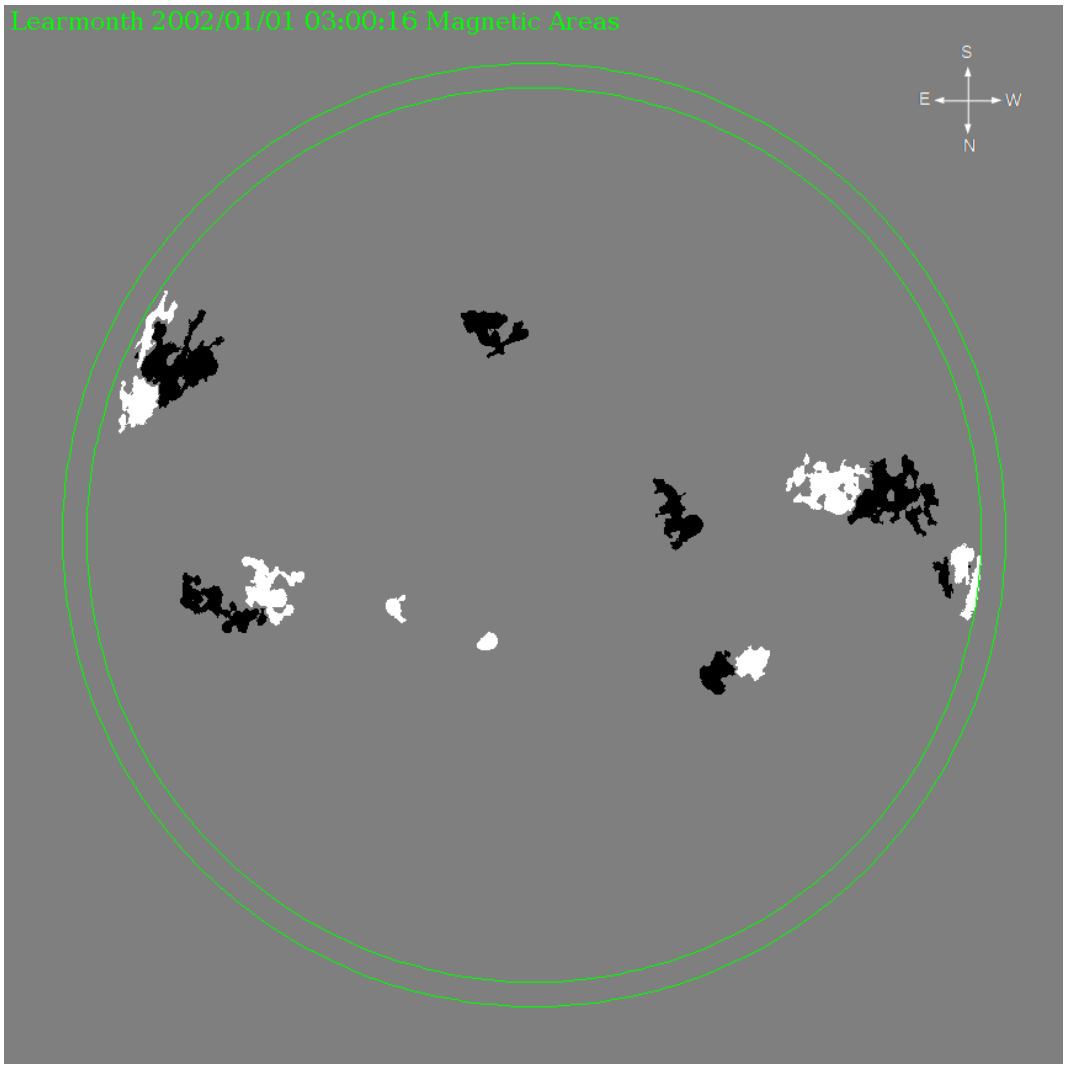



**Figure 7.4**: Sample GONG Magnetic Area Map. Black regions correspond to the magnetic field <-35 Gauss, white regions correspond to the magnetic field > +35 Gauss. The inner green ring is at 0.95 Rv (where analysis stops) and the outer ring is the disk limb.

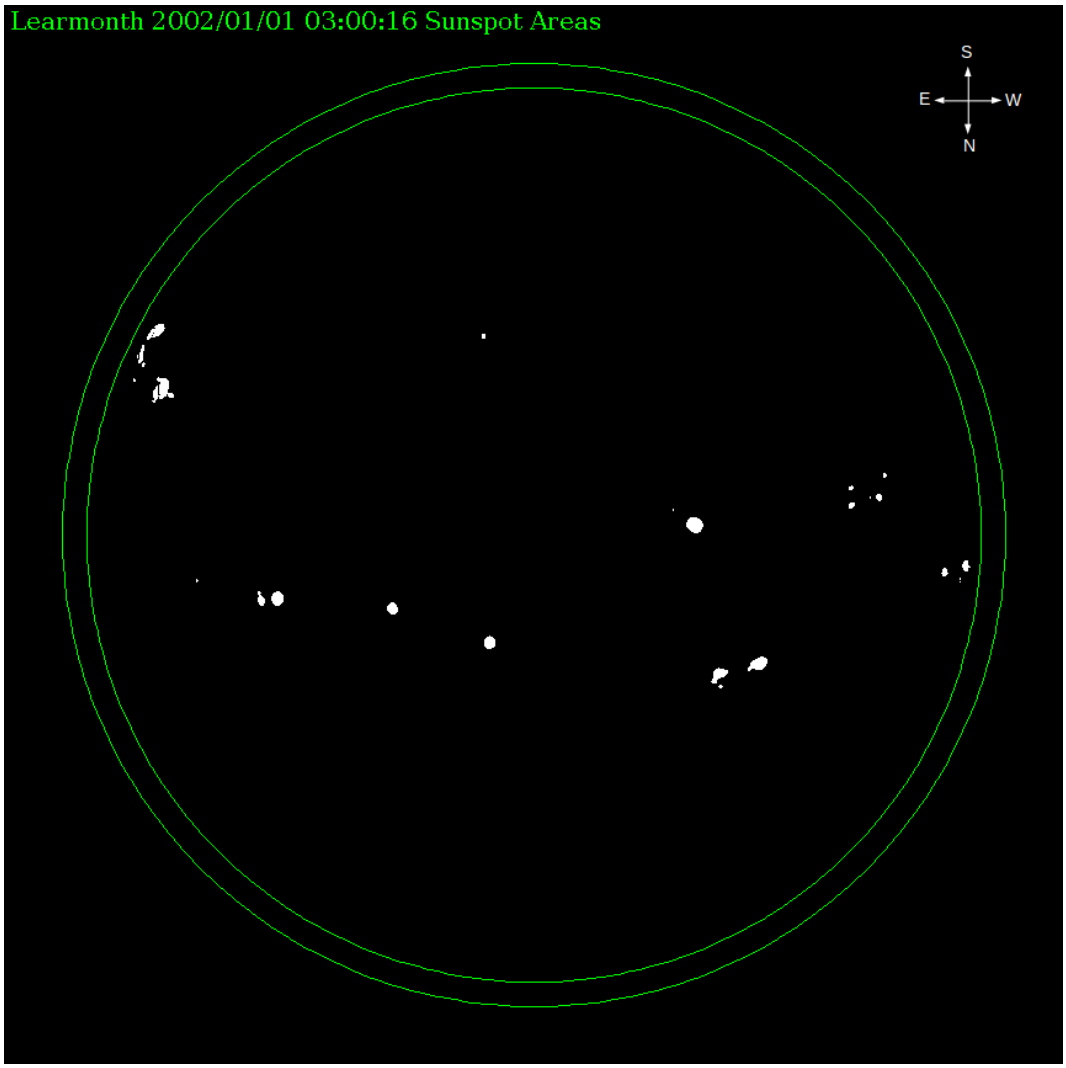



**Figure 7.5**: GONG Sample Sunspot Area Map. The white regions are sunspots detected by the algorithm in figure 7.3.

### 7.2.4 Penumbra and Umbra Detection

For a description on the structure of umbra and penumbra, see chapter 2.

Whilst figure 7.5 is suitable for finding the total sunspot area, it is not suitable for finding the sunspot number or sunspot classification, as the sunspot number requires the number of umbra. The presence of penumbra are required for the Z and P parameters in the McIntosh classification.

Umbra in an individual sunspot are found using the following procedure:

- If there are less than 10 pixels in the sunspot, then there are not enough pixels to begin umbral searching and the procedure stops here, setting all of the pixels in the sunspot to umbra, otherwise it continues.
- The maximum and minimum pixel brightness in a sunspot is found. The image depth of the sunspot is the maximum pixel brightness minus the minimum pixel brightness. If the depth is less than 25 then all of the pixels are set to umbra and the procedure ends here.
- The brightness of pixels two pixels away from the pixel of interest is found. If all of these pixels have greater brightness than the pixel of interest then it is at the minimum brightness of the sunspot (there may be more than one minima if there is more than one umbra in a sunspot). Figure 7.6 shows how this algorithm works:

| | | | | |
|---|---|---|---|---|
| E | E | E | E | E |
| E | NE | NE | NE | E |
| E | NE | PoI | NE | E |
| E | NE | NE | NE | E |
| E | E | E | E | E |

**Figure 7.6**: Sample Pixel Array for Umbral Algorithm. PoI is the Pixel of Interest. The pixels immediately surrounding the PoI are not examined (NE). Those 2 pixels out are Examined (E) to establish if they are darker than the PoI.

- Once a minima in the sunspot is identified, a recursive algorithm is used to identify dark pixels surrounding the PoI. A brightness threshold is defined by equation 7.2. If a pixel immediately surrounding the PoI has a value less than this threshold then this pixel is marked as being part of the umbra and becomes the new Pixel of Interest. This whole step is repeated until no more umbral pixels are found.

$$T = U + (p_{max} - U) \times 0.2 \quad , \qquad \text{Eqn 7.2}$$

where $T$ is the threshold brightness, $U$ is the brightness of the first umbral pixel found in an umbra and $p_{max}$ is the maximum pixel brightness in the sunspot

- The entire procedure is repeated until all the pixels in a sunspot are searched.

The coefficients for equations 7.1 and 7.2 were determined by trial and error such that the values of 0.9 and 0.2 in these expressions were adjusted until the region produced a reasonable sunspot as compared with the raw intensity image. These coefficients can greatly affect the results. For equation 7.1, choosing a value of less than 0.9 tends to miss significant portions of the penumbra and selecting a higher value tends to select large numbers of pixels not associated with the spot. For equation 7.2, choosing a value less than 0.2 tends to miss small umbra and underestimate the umbra size for larger spots, while choosing a larger value selects most of the spot as being umbra.

### 7.2.5 Individual Sunspot Parameters

It is at this stage that parameters for individual sunspots are found. Aside from gaining the penumbral parameter (P) for the individual sunspots, this also aids in defining the sunspot regions.

Firstly, if no dark pixels (umbral pixels) were found, then the entire sunspot is set to being umbra and the penumbral class set to X, (no penumbra).

The sunspot area is calculated by simply adding the individual area of each pixel in the sunspot. This area is calculated by equation 7.3.

$$A_p = \frac{500000}{\pi R_s^2 \sqrt{1 - R_v^2}} \quad , \qquad \text{Eqn 7.3}$$

where $A_p$ is the area of an individual pixel, $R_s$ is the radius of the solar disk in pixels, and $R_v$ is the distance from the centre of the disk to the pixel as a fraction of the solar radius. The factor $500000/\pi$ is to convert from steradian to millionths of the solar hemisphere.

The umbral area is computed in the same way, except only for pixels in the umbra.

Sunspot magnetic polarity is determined by one of two ways depending on the particular sunspot. For a sunspot without penumbra a simple count is made of the number of positive and negative pixels (from the mag-map, figure 7.4). If there are more positive

than negative pixels then the sunspot is positive or negative otherwise.

Alternatively, if a sunspot has penumbra then the polarity of the umbra is found and used as the polarity of the sunspot. If the sunspot has more than one umbra then it is assigned a value of zero if there is umbra of both polarities present.

Therefore the polarity of any given sunspot can have a value of -1 (negative), 0 (mixed polarity) or 1 (positive).

The sunspot shape is determined by looking at the north-south and east-west extent of the sunspot. These extents are all computed in heliographic extent (the program uses radians for all heliographic co-ordinates of pixels). The east-most and west-most pixels in a given sunspot are found and a value, EW, is found by subtracting the east-most position from the west-most position. The same procedure is done in the north-south direction to produce the NS value. The value $P_{ratio}$ is simply the ratio between these values: EW/NS. If the $P_{ratio}$ is in the range 0.85 to 1.15 (axes with 15% of each other) the sunspot is considered symmetric, otherwise it is asymmetric.

In addition the maximum and minimum magnetic fields are found within the sunspot, and also their heliographic co-ordinates.
The penumbral class is then determined by table 7.1. Note that the automated procedure does not try to detect rudimentary penumbra, as the GONG images have insufficient resolution to do it well (see chapter 2 for a definition of rudimentary penumbra).

| Area | Size | $P_{ratio}$ | Shape | Class |
|---|---|---|---|---|
| <250 μh | Small | 0.85 to 1.15 | Symmetric | S |
| | | <0.85 or >1.15 | Asymmetric | A |
| ≥250 μh | Large | 0.85 to 1.15 | Symmetric | H |
| | | <0.85 or >1.15 | Asymmetric | K |

**Table 7.1**: Automated Penumbral Class Determination

### 7.2.6 Defining Sunspot Regions

The sunspot classifications are performed over a whole region, so in order classify them, each individual sunspot must be grouped into a region. A sunspot can be in a group on its

own (unipolar) or be part of a multi-spot region (typically bipolar). This grouping procedure is implemented based on a sunspot's proximity to other sunspots and polarities of the sunspots. The grouping is done in the following order, with distances in heliographic degree:

- any sunspots that are within 3º of each other are grouped together.
- any unipolar regions that are of opposite polarity and within 7º of each other are grouped together.
- any bipolar region that is within 5º of a unipolar region is grouped together.
- if the magnetic field associated with a unipolar sunspot region is within 1º of magnetic field associated with another unipolar sunspot region, then these are grouped together.

Here, the region's distance is defined as the closest distance between two sunspots in different regions. The magnetic field distance is that of the closest significant (<-35 Gauss or >+35 Gauss) magnetic field pixel in one region to the closest significant magnetic field in another region.

The simple procedure defined above produces very good correlation with SWPC regions as will be shown in Section 7.3.3.

### 7.2.7 Region Tracking

SWPC assigns regions an Active Region Number for its entire life across the disk. The same procedure is used here, although a different numbering system is used, as the automated procedure sometimes assigns regions differently than SWPC.

Each region is assigned a designation as follows: LYYYYMMDDNN where L is the site designator, YYYY, MM and DD respectively are the year, month and day when the region was first identified, and NN is a two digit number in case more than one region is identified on a given day. NN typically does not exceed 05, unless there was a significant break (days) in observations.

A given region needs to be tracked across the Sun. This is performed simply by finding the closest previous day's region (corrected for solar rotation) that is within 5º of a current

region. If no such region can be found then it assigned a new number. If a match is found, then the region is given the previous number.

Figure 7.7 shows the sunspots with umbra marked and regional groupings.

It is noted here that the purpose is to assign one region number to each region, regardless of which site identified the region, which is used by all of the GONG sites. However, for ease of testing the overall algorithm's performance, the software is currently set up (by a script) so that each site assigns its own unique number. In a real time environment this can easily and simply be changed.

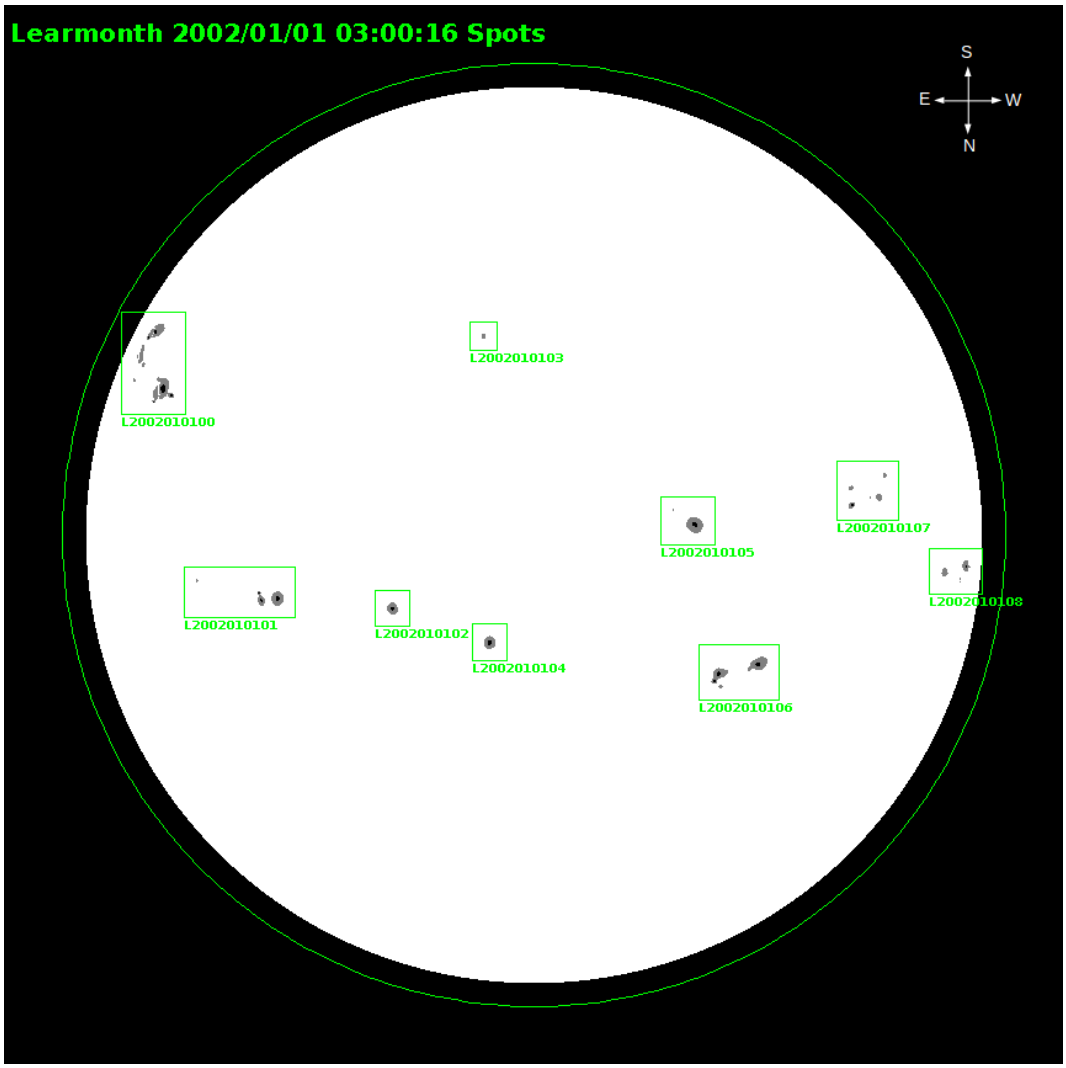



**Figure 7.7**: Sample Automated Sunspot Map with Regions. The green boxes show the individual regions, the grey areas show penumbra and black umbra. If a sunspot has no penumbra, it will just be marked grey.

### 7.2.8 Sunspot Counting, Parameters and Classification

Sunspot parameters that are extracted are: sunspot number, sunspot area for both the region and largest sunspot, umbral area for both the region and largest sunspot, extent of the sunspot group and compactness. The full sunspot classification is also found.

The number of sunspots in each region is found by simply counting the number of umbra in each sunspot in the region.

The sunspot area is found by adding all of the sunspot areas in a region together, and this

is also implemented for all the umbral area of the region, to give the umbral area of the region.

The extent of the region (in degrees longitude) is found by subtracting the east-most longitude from the west-most longitude of sunspots in the region. The extent is required for the Zurich parameter in the McIntosh classification and in addition, is used to determine the Compactness parameter for the region.

The position of the region is just the centre of the region in heliographic co-ordinates.

The largest negative sunspot and largest positive sunspot areas are found in millionths of the solar hemisphere (if the largest sunspot is mixed polarity, then the positive and negative sunspot will be the same). However if the largest sunspot of given polarity does not have penumbra, and there is a sunspot with penumbra that is smaller, the sunspot with penumbra is used (this scenario happens less than 1% of the time, and usually with regions on the limb).

From the extent, polarity and penumbral class (found in section 7.2.4) of the largest sunspot, the Zurich parameter can be found using Table 7.2

| Polarity | Penumbra | Extent | Class |
|---|---|---|---|
| Unipolar | No | N/A | A |
| | Yes | N/A | H |
| Bipolar | No | N/A | B |
| | On One Polarity Only | N/A | C |
| | On Both Polarities | <10° | D |
| | | 10° to 15° | E |
| | | >15° | F |

**Table 7.2**: Automated Zurich Parameter Classification

The penumbral class is chosen from the largest spot in the region, and has already been discussed in section 7.2.5

The Compactness class is found by computing the compactness ratio ($C_{ratio}$). This ratio is

simply the sunspot area of the region divided by the extent of the region. The area here is expressed in the unit millionths of the solar disk and extent is in degree. This parameter is a measure of the sunspot density of the region, the larger it is the more densely packed with spots. Table 7.3 shows how the Compactness class is found.

| Zurich Class | $C_{ratio}$ | Compactness Class |
|---|---|---|
| A,H | N/A | X |
| B, C | <20 | O |
| | ≥20 | I |
| D, E, F | <20 | O |
| | 20 to 50 | I |
| | ≥50 | C |

**Table 7.3**: Automated Compactness Parameter Classification

This gives the full McIntosh classification (with Rudimentary Penumbral class excluded) with no forbidden classes being assigned (see chapter 2).

Table 7.3 shows a summary of the parameters that the program produces for sunspot analysis for figure 7.7.

| Site | Rgn # | Location | #Spots | Spot Area | Umbra Area | L-Spot Area | L-Spot Umbra | Ex | $P_{ratio}$ | $C_{ratio}$ | ZPC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| L | L2002010100 | S19.3E60.3 | 6 | 777.2 | 211.3 | 367.5 | 42.7 | 12 | 1.52 | 63.215 | CKI |
| L | L2002010101 | N04.7E39.1 | 3 | 191.6 | 82.8 | 121.9 | 13.1 | 13 | 1.01 | 14.206 | CSO |
| L | L2002010102 | N06.1E17.6 | 1 | 78.4 | 8.1 | 78.4 | 8.1 | 01 | 0.94 | 60.901 | HSX |
| L | L2002010103 | S27.9E07.0 | 1 | 14.5 | 14.5 | 14.5 | 14.5 | 00 | 1.00 | 40.205 | AXX |
| L | L2002010104 | N10.1E05.5 | 1 | 87.3 | 12.3 | 87.3 | 12.3 | 01 | 0.82 | 69.721 | HAX |
| L | L2002010105 | S05.4W19.1 | 2 | 160.4 | 17.3 | 158.2 | 15.0 | 03 | 1.15 | 42.430 | HAX |
| L | L2002010106 | N13.9W26.7 | 3 | 289.3 | 34.3 | 158.3 | 13.0 | 07 | 1.51 | 38.335 | DAI |
| L | L2002010107 | S08.6W45.7 | 5 | 106.8 | 106.8 | 38.5 | 38.5 | 06 | 1.00 | 15.667 | BXO |
| L | L2002010108 | N02.7W63.8 | 4 | 177.5 | 177.5 | 109.4 | 109.4 | 07 | 1.00 | 22.864 | BXI |

**Table 7.4**: Sample Sunspot Parameters Determined from Figure 7.7. 'Site' is the site designator, 'Rgn #' is the unique identifier for a region, 'Location' is the position of the centre of the region in heliographic co-ordinates, '#Spots' is the number of umbra in the region, 'Spot Area' is the total sunspot area (μh) in the region, 'Umbra Area' is the total umbral area (μh) in the region, 'L-Spot Area' is the area of the largest sunspot (μh) in the region, 'L-Spot Umbra' is the umbral area (μh) of the largest spot, 'Ex' is the east-west extent of the region in heliographic degree, '$P_{ratio}$' is the penumbral ratio, '$C_{ratio}$' is the compactness ratio and 'ZPC' is the McIntosh sunspot classification of the region.

### 7.2.9 Magnetic Parameters and Classification

Once the regions have been found, the sunspot parameters and simplified magnetic map (figure 7.4) are examined to extract magnetic information.

For each sunspot the magnetic field is searched to find the strongest positive and negative fields in the region. From this the magnetic field gradient between the fields is calculated in Gauss per megametre (1 heliographic degree = 12.15 Mm = 12150 km).

In addition, the orientations of the magnetic fields are found. Normally sunspots (and the magnetic fields) are oriented in an east-west direction. The field orientation is assigned a value from 0 to 90, where 0 is a north-south orientation and 90 an east-west orientation.

If a region is unipolar then its magnetic classification is automatically set as α. Further processing is required for bipolar regions.

In order to determine if a region should be classified as β, βγ or γ, the region is broken by

longitude in half degree steps. The number of positive and negative pixels within each half degree step are counted. If a longitudinal step contains both positive and negative fields, then the amount of field overlap is incremented by 0.5. This overlap is then divided by the total magnetic field extent of the region. This ratio is called the $\gamma_{ratio}$. Figure 7.8 demonstrates this process. Table 7.5 shows the magnetic classification by $\gamma_{ratio}$.

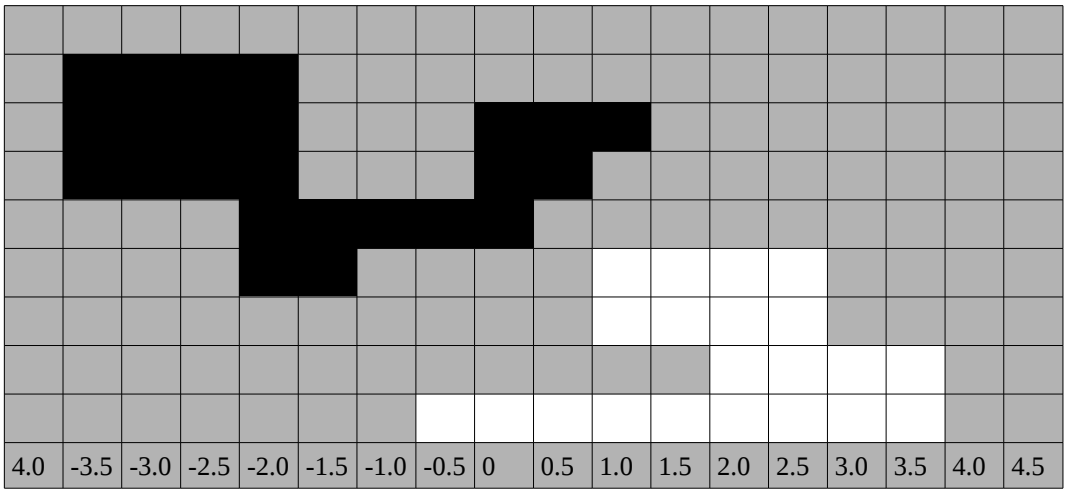



**Figure 7.8**: Idealised Magnetic Region Example to Demonstrate $\gamma_{ratio}$ Determination. The black boxes are negative polarity pixels and the white are positive polarity pixels, with grey neutral. Each column represents a step of 0.5º longitude along the region from the region centre. This region has a field extent of 7º (-3.5 to 3.5). There are 4 columns that have both positive and negative field (-0.5 to 1.0) giving the amount of overlap of positive and negative field of 2º. Thus in this example the $\gamma_{ratio}$ is 0.28. From table 7.5 below this would make it a βγ region.

| $\gamma_{ratio}$ | Magnetic Class |
|---|---|
| <0.25 | β |
| 0.25 to 0.4 | βγ |
| ≥0.4 | γ |

**Table 7.5**: Magnetic Classification by $\gamma_{ratio.}$

The relative field positions are inspected to determine if the sunspot region is reversed polarity. If the leading (west-most) sunspot polarity is opposite to what would be expected for the current sunspot cycle (Bray and Loughhead, 1964) then the sunspot region is classified reversed.

Finally the magnetic inversion line needs to be found so that the inversion line length and

complexity can be found. Figure 7.9 shows an idealised magnetic region with the inversion line marked on it.

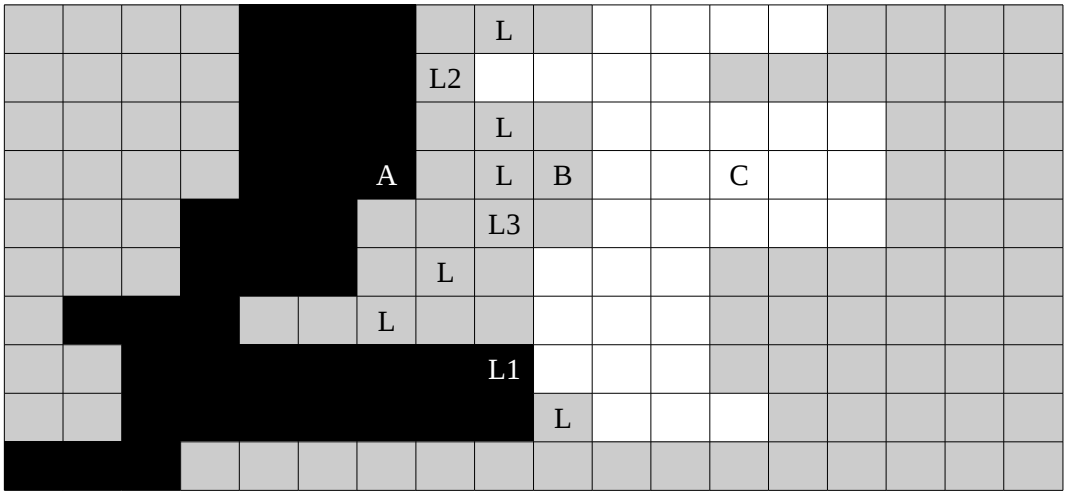



**Figure 7.9**: Idealised Magnetic Region with Neutral Line Marked with L. The methodology of how this diagram is used is given below.

The neutral line is determined as follows with reference to figure 7.9:

- If a pixel in the region is not neutral (either polarity) then adjacent pixels are examined to see if they are of the opposite polarity. If an opposite polarity pixel is found then this pixel is on the inversion line. An example of this is the pixel marked as L1 in figure 7.9.
- If the pixel is neutral, immediately adjacent pixels on opposite sides are examined. If these pixels have non-zero opposite polarity then the neutral pixel between them is on the inversion line. The pixel L2 in figure 7.9 is an example of this where the pixels immediately adjacent to L2 are opposite polarity.
- If a non-neutral pixel is found within 2 degrees of a neutral pixel and the pixel directly opposite is of opposite polarity and not completely surrounded by non-neutral pixels then the neutral pixel is on the inversion line. In figure 7.9, the pixel L3 satisfies this condition,with there being a positive pixel directly above 3 pixels and a negative pixel below 3 pixels. The pixel B does have opposite polarities 3 pixels to the left and right of it (at pixels A and C), but C is surrounded by other positive pixels, and therefore B is not on the neutral line.

As inversion line pixels are found, the length of each pixel in degree is calculated and

added to the inversion line length.

The north-south extent of the inversion line is found (the usual orientation of the inversion line). The inversion line complexity is just the inversion line length divided by the north south extent of the inversion line. The inversion line complexity typically ranges from 1 to 5. Values less than 1 are possible if the inversion line is broken up and not well defined.

Figure 7.10 shows a sample processed magnetic map with regions and inversion lines marked.

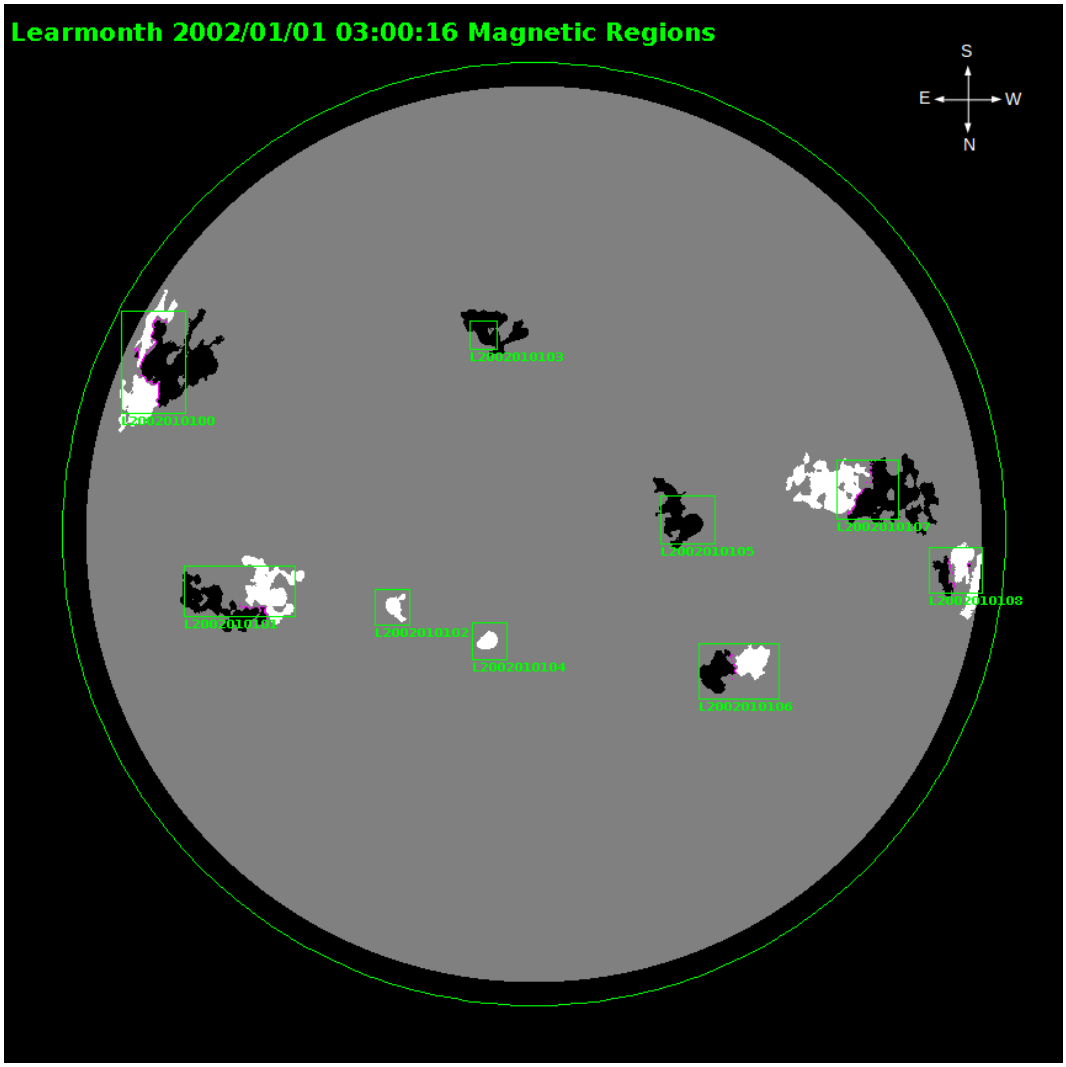



**Figure 7.10**: Sample Processed Magnetic Map. The black areas are negative field (<-35 G), white areas positive field (> +35 G) and the pink lines shows the inversion lines. The regions are marked

by the green boxes.

Table 7.6 show the sample magnetic field parameters that the program generated from figure 7.10.

| Site | Rgn # | -ve Fld | -ve Location | +ve Fld | +ve Location | Grad | O | ILC | ILL | $\gamma_{ratio}$ | M |
|---|---|---|---|---|---|---|---|---|---|---|---|
| L | L2002010100 | -0790 | S19.8E56.4 | +0319 | S26.0E65.8 | 8.6 | 56 | 39 | 3.381 | 0.41 | G |
| L | L2002010101 | -0274 | N03.4E45.8 | +0757 | N05.1E32.9 | 6.5 | 82 | 6 | 2.794 | 0.20 | B |
| L | L2002010102 | ///// | ////////// | +0696 | N06.0E17.6 | ////// | // | /// | ///// | 0.00 | A |
| L | L2002010103 | -0443 | S27.9E07.0 | ///// | ////////// | ////// | // | /// | ///// | 0.00 | A |
| L | L2002010104 | ///// | ////////// | +0886 | N10.1E05.5 | ////// | // | /// | ///// | 0.00 | A |
| L | L2002010105 | -0873 | S04.2W19.9 | ///// | ////////// | ////// | // | /// | ///// | 0.00 | A |
| L | L2002010106 | -0811 | N14.2W23.8 | +0855 | N13.3W29.0 | 26.6 | 79 | 3 | 1.168 | 0.05 | B |
| L | L2002010107 | -0446 | S06.7W47.4 | +0297 | S07.9W42.6 | 12.6 | 75 | 9 | 1.596 | 0.17 | B |
| L | L2002010108 | -0276 | N03.0W60.5 | +0457 | N02.7W65.8 | 11.4 | 86 | 10 | 2.801 | 0.18 | B |

**Table 7.6**: Sample Magnetic Parameters Determined from Figure 7.10. 'Site' is the site designator, 'Rgn #' is the unique identifier for a region, '-ve Fld' is the most negative magnetic field strength in Gauss, '-ve Location' is the position of the most negative field in heliographic co-ordinates, '+ve Fld' is the most positive magnetic field strength in Gauss, '+ve Location' is the position of the most positive field in heliographic co-ordinates, 'Grad' is the magnetic field gradient in G/Mm between the most positive and most negative field, 'O' is orientation from north in degrees, 'ILL' is the inversion line length in heliographic degrees, 'ILC' is the inversion line complexity, '$\gamma_{ratio}$' is the amount of mixing of polarities in the region and 'M' is the Mount Wilson magnetic classification.

### 7.2.10 Programming Language and Performance

The entire program was written in Java version 7. Java was chosen due to its easy portability across platforms. Testing was done using an 8 core PC that was running Ubuntu version 14.04. The program accepts as input the filename of the intensity and magnetogram FITS files. Running the program once will process one set of images. It was designed this way in order for a daily Cron to be run, processing one data set per day per site.

However, since archived data was being used, a script was written that simply looped through the entire data set for each GONG site. The data set comprised, where possible, one magnetogram and intensity FITS file for each day from 2002/01/01 to approximately

2015/08/31 (the end date varied with site). A total of 23,079 intensity – magnetogram file pairs were processed. Processing one site on each core resulted in the entire dataset being processed in about 18 hours, or about 15 seconds per file pair.

## 7.3 Results

### 7.3.1 Sunspot Number

A daily sunspot number for each site was calculated in the usual manner (Equation 2.1) with a *k* value of 1. From this an average daily GONG sunspot number was computed simply by averaging all of the daily GONG individual site sunspot numbers. A site was rejected from the average if it had (i) a sunspot number greater than 75 and the sunspot number exceeded the average by more than 50% or (ii) if the sunspot area was greater than 500 μh and the sunspot area exceeded the average area by more than 50%. Out of the 13 years of data, only on 66 occasions was a site rejected.

Figure 7.11 shows the GONG daily sunspot number for the individual sites and the daily average derived as described in the previous sections.

These daily averages are then averaged further in the same way as the ISN (monthly averages are calculated and these are then smoothed over a year).

Figure 7.12 shows the GONG smoothed sunspot number calculated as described above along with the International Sunspot Number and AAVSO Sunspot Number. This plot shows that there are no significant deviations from either the International or AAVSO sunspot numbers. These relationships are plotted in figures 7.13 and 7.14.

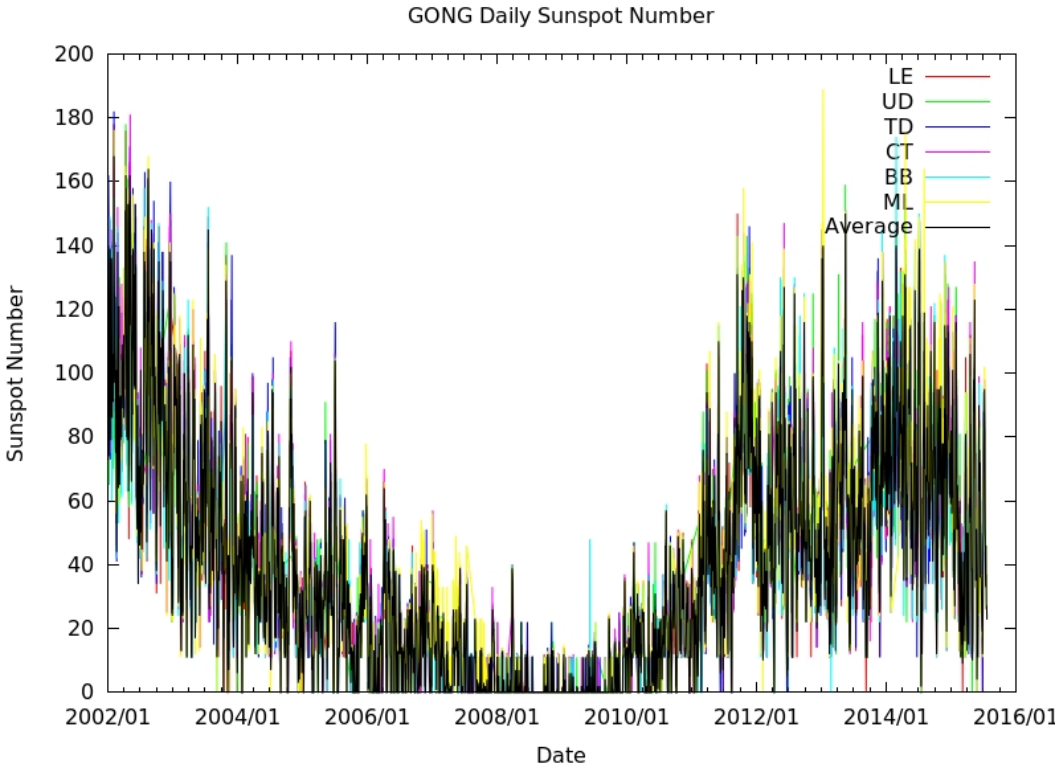



**Figure 7.11**: GONG Daily Sunspot Number Derived using the Methodolgy in the Text. This plot has had outliers removed. The black line is the average of all the sites that had an observation on a given day.

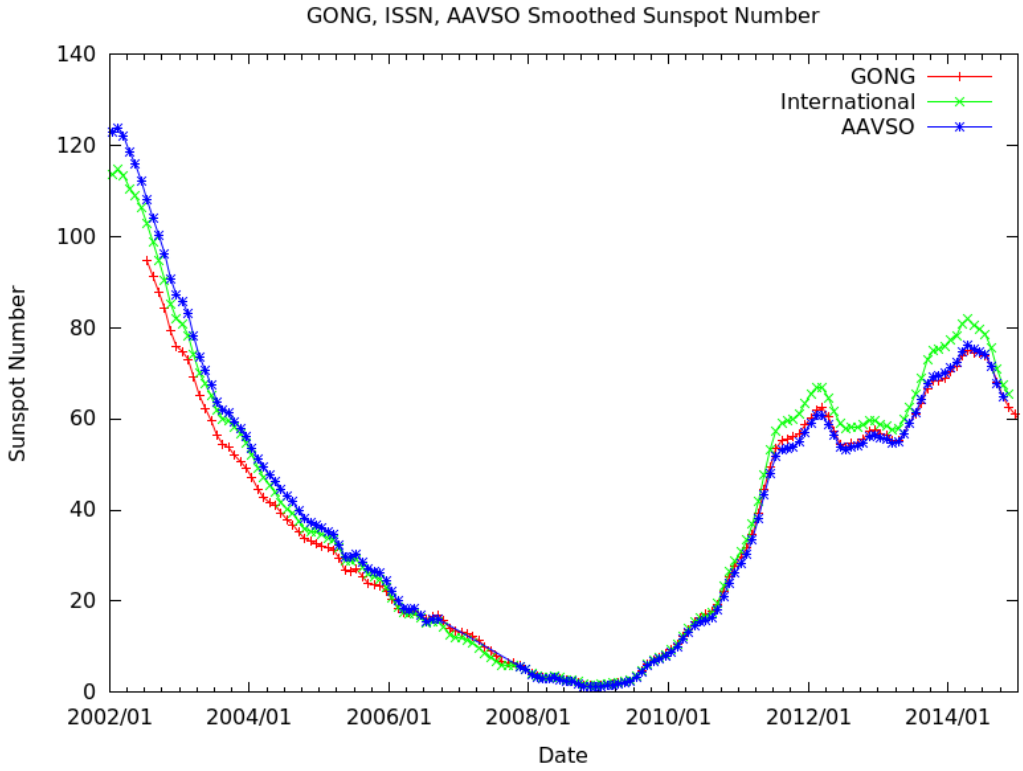



**Figure 7.12**: GONG, International and AAVSO Smoothed Monthly Sunspot Numbers from 2002 to 2015

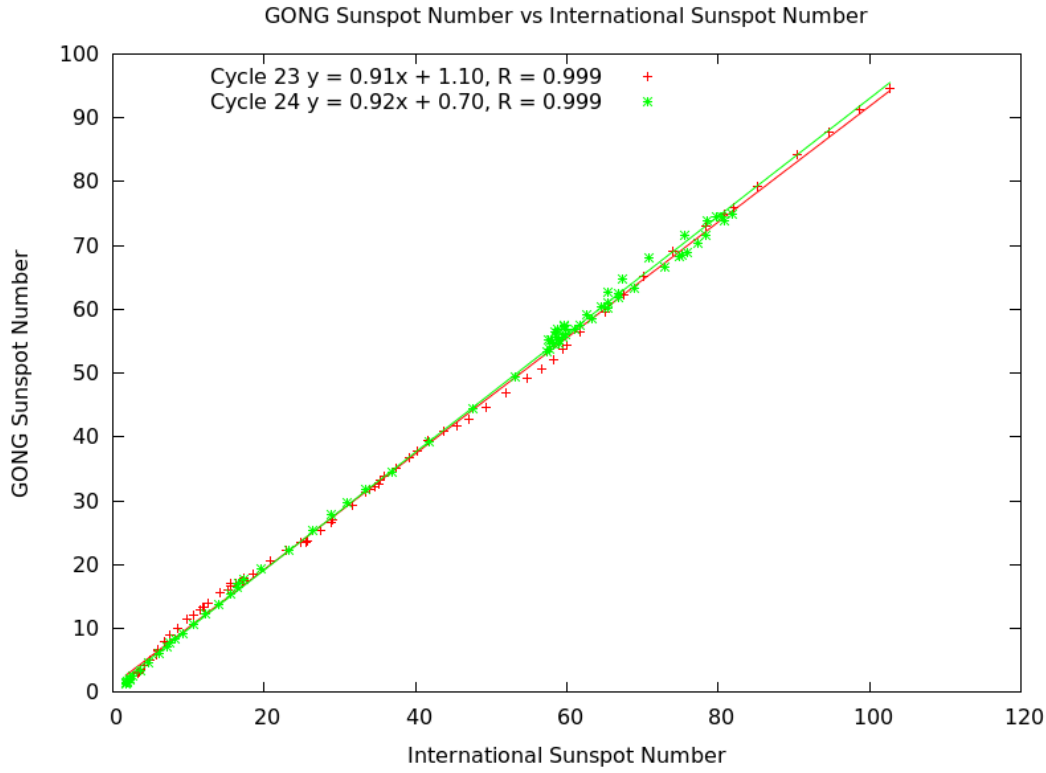



**Figure 7.13**: GONG Sunspot Number from the Present Research vs International Sunspot Number for Cycles 23 and 24

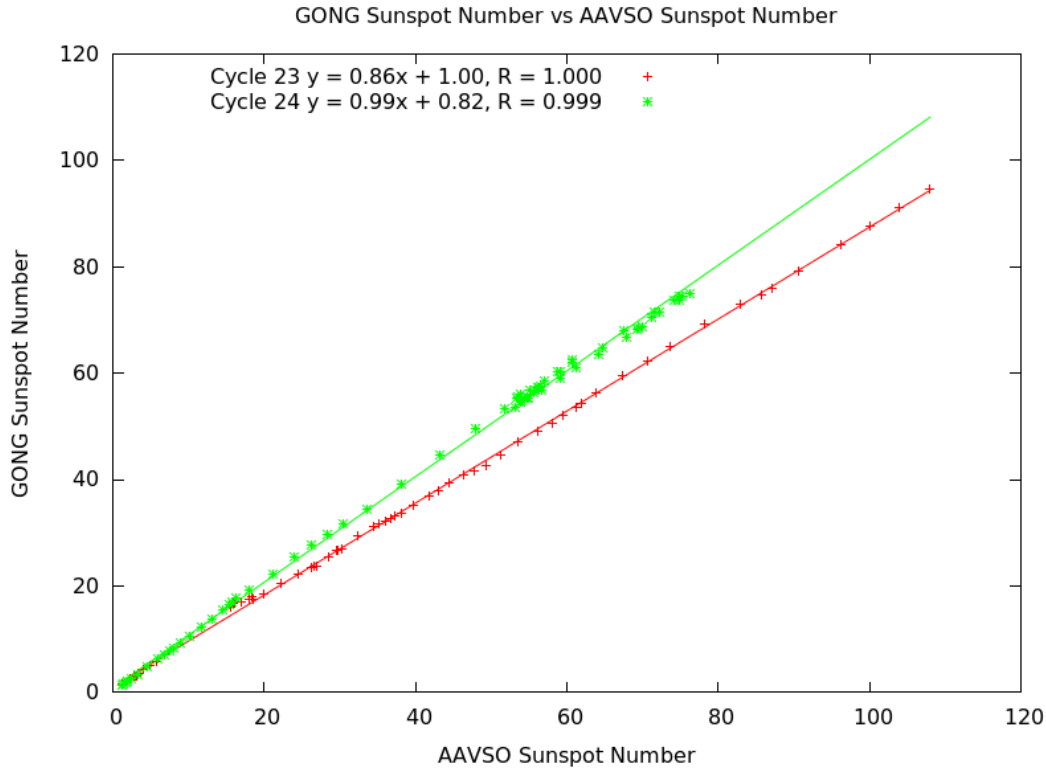



**Figure 7.14**: GONG Sunspot Number from the Present Research vs AAVSO Sunspot Number for Cycles 23 and 24

In all cases, the correlation coefficients are close to 1. Furthermore, from figure 7.13, there does not appear to be a significant change in relationship between cycles 23 and 24 for the International Sunspot Number, but cycle 24 has a different slope than cycle 23 for the GONG and AAVSO comparison in figure 7.14. This would seem to suggest that measurement of the AAVSO sunspot number has changed between these cycles.

Further investigation was done on the slopes of the lines to determine if there was any statistical difference between cycles. Table 7.7 shows the fit parameters for figures 7.13 and 7.14.

The difference between the slopes in cycle 23 and 24 for the GONG SSN vs International SSN plots is greater than 2 standard deviations. Therefore these slopes are still statistically different. A similar process was used for the GONG SSN vs AAVSO SSN. The cycle variation in this case is even greater.

| X-Axis | Y-Axis | Cycle | m | $s_m$ | b | $s_b$ |
|---|---|---|---|---|---|---|
| International SSN | GONG SSN | 23 | 0.9069 | 0.0040 | 1.10 | 0.17 |
| | | 24 | 0.9231 | 0.0041 | 0.70 | 0.21 |
| AAVSO SSN | GONG SSN | 23 | 0.8644 | 0.0025 | 1.00 | 0.12 |
| | | 24 | 0.9933 | 0.0040 | 0.82 | 0.19 |

**Table 7.7** Fit Parameters for GONG Sunspot Number Comparison

The explanation for the statistical difference between GONG SSN vs International SSN differences is simply that the fits are so good and the standard deviations so small that even a small variation in slope comes up as statistically significant. However it is still argued here that the GONG algorithm is far more consistent than most manual reductions, at least over the time period studied. It is too early to tell yet whether these slight variations are random between cycles or if they suggest something more systematic. Furthermore, it cannot deduced which data set is varying (however slightly) between cycles, whether it is GONG or the International sunspot numbers. Hopefully this will be able to be determined once a complete sunspot cycle's worth of data is collected.

### 7.3.2 Sunspot Area

A daily sunspot area for each site was calculated. As with sunspot number, there were some outliers and these were rejected using the same criteria as with sunspot number. Each site's daily sunspot area was averaged to produce a daily GONG sunspot area. These results are shown in figure 7.15.

As with sunspot number a smoothed monthly average area was calculated using GONG data processed in this project. This is plotted with the SRS smoothed monthly area in figure 7.16

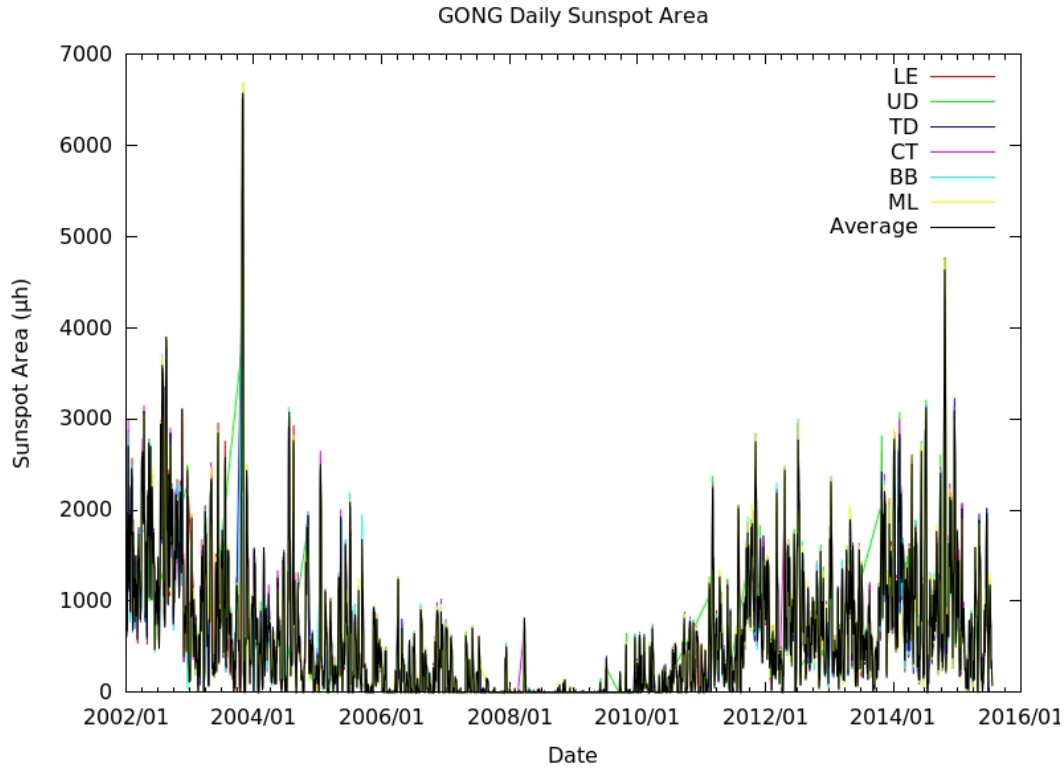



**Figure 7.15**: GONG Daily Sunspot Area Derived in this Project. This plot has had outliers removed. The black line is the average of all the sites that had an observation on a given day.

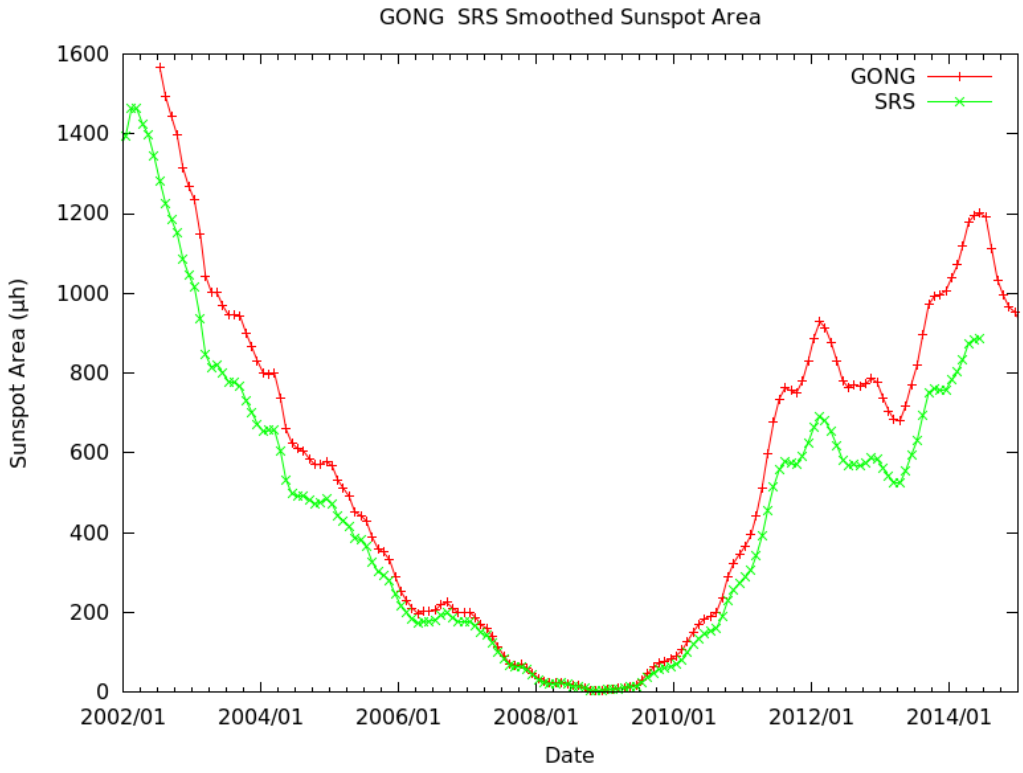



**Figure 7.16**: GONG and SRS Smoothed Monthly Sunspot Areas from 2002 to 2015

The GONG procedure, as shown in figure 7.16, is measuring consistently higher sunspot areas than the SRS are reporting. This is further illustrated in figure 7.17 which plots GONG Area vs SRS Area. Table 7.6 shows the fit parameters for this plot, showing that there is again a statistically significant difference between the two cycles.

It is still consistent that GONG was providing a significantly higher sunspot number than SRS, but both cycles have a correlation coefficient of one. The correlation coefficient indicates how well the variations correlate with each other, independent of offset, and Fig 7.16 shows they correlate well, hence the unity coefficient. However, the non-unity intercepts reflect some kind of offset in number that the correlation coefficient can't represent, and Fig 7.16 indicates the offset is solar cycle phase dependent.

It is hard to believe that an automated procedure working on similar images could produce this level of variation. Therefore, it is suspected that it is the manual measurement of sunspot areas that is varying across cycles. Also, these results add credence to the claim that current area measurements are under-estimating sunspot areas

by 25% to 50%. However it would be difficult to correct the data set as the the measurements change on a cycle-to-cycle basis.

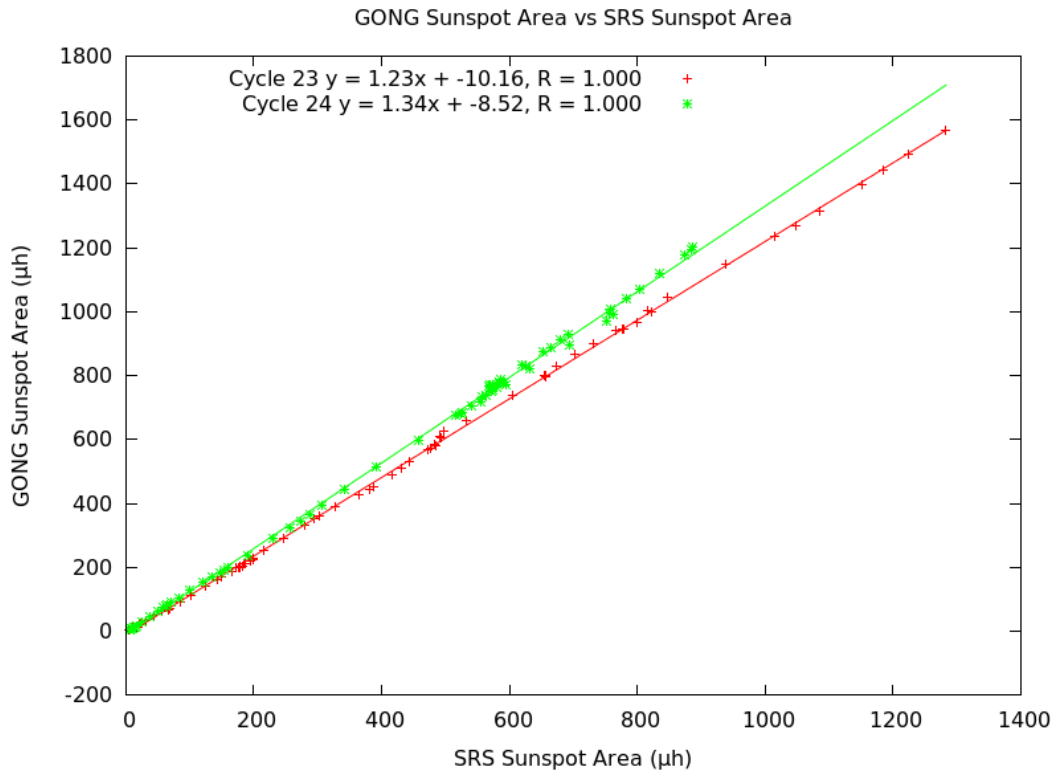



**Figure 7.17**: GONG Sunspot Area vs Solar Region Summary Sunspot Area for Cycles 23 and 24

| Cycle | m | $s_m$ | b | $s_b$ |
|---|---|---|---|---|
| 23 | 1.2299 | 0.0025 | -10.16 | 1.35 |
| 24 | 1.3389 | 0.0045 | -8.52 | 2.28 |

**Table 7.8**: Fit Parameters for GONG – SRS Sunspot Area Comparison

Table 7.8 shows the regression parameters for the GONG – SRS sunspot area comparisons. The standard deviation in the slope can not account for the variation between the two cycles.

The umbra area to sunspot area ratio was calculated for both cycle 23 and 24 in order to compare with other published results as a test for the algorithm. This was performed for both the entire region and for the largest spot in the region. Only regions that had a

penumbra on at least one spot were used in these statistics. Tables 7.9 to 7.11 summarise these results.

The first thing to note here is the the standard deviations are quite high, making it difficult to draw significant conclusions. Tables 7.9 and 7.10 show an increase in the umbra to spot area ratio for spots with more than 1000 μh of area, indicating very large regions have on average larger umbra than smaller regions. There does not appear to be a significant change in this ratio across sunspot cycles, however, its noted here again, that only the declining part of cycle 23 and increasing part of cycle 24 were measured.

For the largest spot (tables 7.11 and 7.12) this ratio is similar, although slightly lower than that reported by Bray and Loughhead (1964). Similarly, there does appear to be some difference between the ratio for spots with less than 100 μh of area and those with more than 100 μh, but this is not a statistically significant difference. In fact, the difference here is larger than that reported by Hathaway (2005).

| Area (μh) | | LE | UD | TD | CT | BB | ML | Combined |
|---|---|---|---|---|---|---|---|---|
| A<100 | Mean | 0.202 | 0.204 | 0.213 | 0.208 | 0.204 | 0.214 | 0.208 |
| | SD | 0.142 | 0.151 | 0.149 | 0.145 | 0.147 | 0.154 | 0.148 |
| 100<=A<250 | Mean | 0.196 | 0.213 | 0.205 | 0.202 | 0.204 | 0.202 | 0.203 |
| | SD | 0.151 | 0.164 | 0.155 | 0.155 | 0.151 | 0.150 | 0.154 |
| 250<=A<=600 | Mean | 0.209 | 0.201 | 0.197 | 0.198 | 0.205 | 0.206 | 0.203 |
| | SD | 0.136 | 0.135 | 0.127 | 0.130 | 0.137 | 0.138 | 0.134 |
| 600<=A<1000 | Mean | 0.196 | 0.190 | 0.207 | 0.209 | 0.193 | 0.195 | 0.199 |
| | SD | 0.096 | 0.087 | 0.110 | 0.107 | 0.106 | 0.101 | 0.103 |
| A>=1000 | Mean | 0.244 | 0.275 | 0.262 | 0.266 | 0.257 | 0.271 | 0.261 |
| | SD | 0.117 | 0.156 | 0.124 | 0.120 | 0.127 | 0.125 | 0.126 |
| All | Mean | 0.203 | 0.208 | 0.207 | 0.205 | 0.205 | 0.208 | 0.206 |
| | SD | 0.140 | 0.150 | 0.143 | 0.142 | 0.142 | 0.144 | 0.143 |
| A>=100 | Mean | 0.203 | 0.209 | 0.205 | 0.204 | 0.206 | 0.207 | 0.205 |
| | SD | 0.140 | 0.149 | 0.141 | 0.141 | 0.141 | 0.141 | 0.142 |

**Table 7.9**: Umbra to Spot Area Ratio for Entire Regions Cycle 23. Mean is the average sunspot ratio and SD is the standard deviation of the mean.

| Area (μh) | | LE | UD | TD | CT | BB | ML | Combined |
|---|---|---|---|---|---|---|---|---|
| A<100 | Mean | 0.201 | 0.216 | 0.191 | 0.223 | 0.212 | 0.215 | 0.211 |
| | SD | 0.135 | 0.153 | 0.131 | 0.163 | 0.147 | 0.146 | 0.148 |
| 100<=A<250 | Mean | 0.197 | 0.202 | 0.200 | 0.200 | 0.199 | 0.198 | 0.199 |
| | SD | 0.154 | 0.157 | 0.157 | 0.156 | 0.156 | 0.153 | 0.155 |
| 250<=A<=600 | Mean | 0.205 | 0.206 | 0.220 | 0.194 | 0.205 | 0.197 | 0.204 |
| | SD | 0.137 | 0.131 | 0.145 | 0.125 | 0.135 | 0.127 | 0.134 |
| 600<=A<1000 | Mean | 0.207 | 0.187 | 0.194 | 0.182 | 0.182 | 0.189 | 0.190 |
| | SD | 0.117 | 0.100 | 0.100 | 0.090 | 0.091 | 0.103 | 0.101 |
| A>=1000 | Mean | 0.248 | 0.243 | 0.251 | 0.247 | 0.241 | 0.230 | 0.243 |
| | SD | 0.123 | 0.130 | 0.126 | 0.124 | 0.126 | 0.110 | 0.124 |
| All | Mean | 0.202 | 0.206 | 0.205 | 0.204 | 0.204 | 0.202 | 0.204 |
| | SD | 0.143 | 0.146 | 0.147 | 0.146 | 0.145 | 0.142 | 0.145 |
| A>=100 | Mean | 0.202 | 0.204 | 0.208 | 0.198 | 0.202 | 0.198 | 0.202 |
| | SD | 0.145 | 0.144 | 0.149 | 0.141 | 0.144 | 0.140 | 0.144 |

**Table 7.10**: Umbra to Spot Area Ratio for Entire Regions Cycle 24

| Area (μh) | | LE | UD | TD | CT | BB | ML | Average |
|---|---|---|---|---|---|---|---|---|
| A<100 | Mean | 0.140 | 0.141 | 0.146 | 0.139 | 0.146 | 0.141 | 0.142 |
| | SD | 0.116 | 0.131 | 0.114 | 0.104 | 0.134 | 0.119 | 0.119 |
| 100<=A<250 | Mean | 0.119 | 0.124 | 0.114 | 0.116 | 0.119 | 0.115 | 0.118 |
| | SD | 0.122 | 0.132 | 0.104 | 0.113 | 0.114 | 0.106 | 0.115 |
| 250<=A<=600 | Mean | 0.111 | 0.107 | 0.110 | 0.111 | 0.110 | 0.113 | 0.110 |
| | SD | 0.094 | 0.094 | 0.090 | 0.099 | 0.097 | 0.106 | 0.097 |
| 600<=A<1000 | Mean | 0.129 | 0.127 | 0.147 | 0.145 | 0.129 | 0.129 | 0.135 |
| | SD | 0.077 | 0.087 | 0.104 | 0.103 | 0.092 | 0.090 | 0.094 |
| A>=1000 | Mean | 0.210 | 0.256 | 0.226 | 0.229 | 0.218 | 0.244 | 0.228 |
| | SD | 0.143 | 0.181 | 0.164 | 0.154 | 0.158 | 0.162 | 0.159 |
| All | Mean | 0.125 | 0.127 | 0.128 | 0.126 | 0.126 | 0.127 | 0.126 |
| | SD | 0.113 | 0.124 | 0.109 | 0.110 | 0.116 | 0.115 | 0.114 |
| A>=100 | Mean | 0.122 | 0.123 | 0.122 | 0.122 | 0.122 | 0.123 | 0.122 |
| | SD | 0.112 | 0.122 | 0.107 | 0.112 | 0.111 | 0.113 | 0.112 |

**Table 7.11**: Umbra to Spot Area Ratio for Largest Spot in a Region Cycle 23

| Area ( μh) | | LE | UD | TD | CT | BB | ML | Average |
|---|---|---|---|---|---|---|---|---|
| A<100 | Mean | 0.131 | 0.141 | 0.126 | 0.151 | 0.137 | 0.135 | 0.138 |
| | SD | 0.091 | 0.117 | 0.062 | 0.148 | 0.094 | 0.089 | 0.107 |
| 100<=A<250 | Mean | 0.113 | 0.115 | 0.116 | 0.113 | 0.118 | 0.117 | 0.115 |
| | SD | 0.103 | 0.111 | 0.121 | 0.106 | 0.116 | 0.118 | 0.113 |
| 250<=A<=600 | Mean | 0.110 | 0.111 | 0.109 | 0.111 | 0.115 | 0.114 | 0.112 |
| | SD | 0.098 | 0.096 | 0.099 | 0.107 | 0.118 | 0.107 | 0.105 |
| 600<=A<1000 | Mean | 0.128 | 0.127 | 0.130 | 0.122 | 0.126 | 0.135 | 0.128 |
| | SD | 0.100 | 0.105 | 0.092 | 0.080 | 0.100 | 0.115 | 0.099 |
| A>=1000 | Mean | 0.221 | 0.207 | 0.206 | 0.223 | 0.222 | 0.191 | 0.213 |
| | SD | 0.160 | 0.165 | 0.151 | 0.157 | 0.161 | 0.131 | 0.156 |
| All | Mean | 0.119 | 0.123 | 0.119 | 0.124 | 0.124 | 0.122 | 0.122 |
| | SD | 0.103 | 0.112 | 0.107 | 0.120 | 0.115 | 0.110 | 0.111 |
| A>=100 | Mean | 0.117 | 0.118 | 0.117 | 0.117 | 0.121 | 0.119 | 0.118 |
| | SD | 0.105 | 0.110 | 0.114 | 0.109 | 0.119 | 0.115 | 0.112 |

**Table 7.12**: Umbra to Spot Area Ratio for Largest Spot in a Region Cycle 24

### 7.3.3 Sunspot Area to Sunspot Number Ratios

The automated procedures detailed in section 7.2 should produce consistent sunspot numbers and sunspot areas over long time periods. Therefore, it should be possible to see how the sunspot area to sunspot number ratio changes between cycles. In this case however, analysis is somewhat limited due to only having GONG observations for the declining half of cycle 23 and rising half of cycle 24. As detailed in section 4.4, it seems that this ratio changes after solar maximum in a cycle.

Figure 7.18 shows that this ratio has changed between the two parts of the cycles that were measured. The fit parameters for these lines are shown in Table 7.13. Once again the gradients here are statistically different (outside 2 standard deviations from each other). As stated previously, as these are in different parts of the cycle, some care needs to be taken in interpreting this result. However this difference does show that using this ratio to correct sunspot areas is difficult as there is variation between cycles.

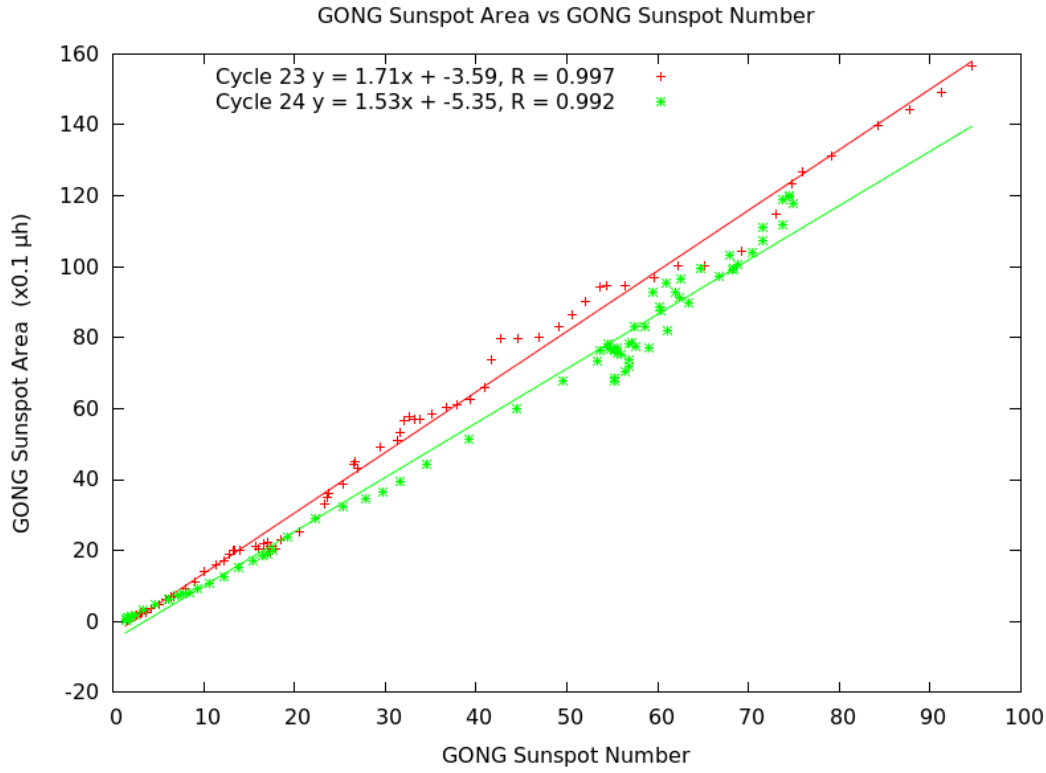



**Figure 7.18**: GONG Sunspot Area vs GONG Sunspot Number

| Cycle | m | $s_m$ | b | $s_b$ |
|---|---|---|---|---|
| 23 | 1.707 | 0.016 | -3.59 | 0.65 |
| 24 | 1.532 | 0.023 | -5.35 | 1.11 |

**Table 7.13**: Fit Parameters for GONG Sunspot Area vs GONG Sunspot Number

### 7.3.4 Sunspot Classification

In order to determine how the automated region classification system compares with manual analysis, firstly it is necessary to determine if the region locations are the same as the solar region summaries report. This is achieved simply by finding the closest SRS region to a GONG region and establishing if it is within 5º (after being corrected for solar rotation). If it is, then its location is considered a match. Table 7.14 shows how well the region locations match for each GONG site.

| GONG Site | Number of GONG Regions | Number of SRS Regions | Number of Matched Regions | % GONG Matched | % SRS Matched |
|---|---|---|---|---|---|
| LE | 11694 | 14640 | 10687 | 91.39 | 73.00 |
| UD | 8466 | 10177 | 7614 | 89.94 | 74.82 |
| TD | 10993 | 13641 | 10014 | 91.09 | 73.41 |
| CT | 11966 | 14360 | 10740 | 89.75 | 74.79 |
| BB | 11241 | 13892 | 9985 | 88.83 | 71.88 |
| ML | 11435 | 13345 | 9715 | 84.96 | 72.80 |

**Table 7.14**: GONG - Solar Region Summary Region Location Summary

The first thing to note is that there are more SRS regions than there are GONG regions. This is due to the smaller telescope size and moderate size (~780 by 780 pixels) of the GONG solar disk. Small spots are not resolved on GONG reducing the number of regions. In addition, since the GONG analysis only extends out to 0.95 $R_v$, regions on the extreme edge of the limb are not being recorded by GONG.

The '% GONG Matched' column is simply the 'Number of Matched Regions' divided by the 'Number of GONG Regions', and the '% SRS Matched' is the 'Number of Matched Regions' divided by the 'Number of SRS Regions'. Thus as there are more SRS regions, the '% SRS Matched' is smaller.

In most cases around 90% GONG regions were matched with an SRS region. This is considered to be a good result. Reasons for the 10% unmatched regions will be detailed later in this section.

This result should not be surprising due to the approximate 90% relationship between the International and GONG sunspot number as seen in figure 7.13. If there were a lot more or less regions being found then this sunspot number relationship would be different.

The McIntosh Classification scheme (see chapter 2 for a description) is examined first. Initially, consistency between GONG sites is examined. This is a simple process of matching each region from a specific GONG site with the same region from every other GONG site. The full McIntosh class is used (ZPC). These results are shown in table 7.15.

| | LE | UD | TD | CT | BB | ML |
|---|---|---|---|---|---|---|
| LE | 100.00% | 64.09% | 55.25% | 49.21% | 45.64% | 43.38% |
| UD | 63.93% | 100.00% | 56.56% | 52.52% | 46.32% | 43.94% |
| TD | 55.05% | 56.41% | 100.00% | 60.55% | 54.84% | 50.85% |
| CT | 49.18% | 52.52% | 60.67% | 100.00% | 63.89% | 59.35% |
| BB | 45.45% | 46.18% | 54.90% | 63.64% | 100.00% | 65.52% |
| ML | 43.14% | 43.78% | 50.81% | 59.14% | 65.46% | 100.00% |

**Table 7.15**: Inter-Site GONG McIntosh Class Comparison

Table 7.15 has the sites arranged in order of time, that is that Learmonth observations are on average the first available on a UT day, while Mauna Loa observations are at the end of the UT day. This is significant as sites that are closest in time have the highest match. This is primarily due to sunspot evolution over a 24 hour period. Sunspots may change significantly in a few hours.

How does the variation seen in Table 7.15 compare with manual reduction techniques? The Solar Region Summaries put out by SWPC are, in part, compiled from the sunspot drawings from the USAF solar observatories. These reports were used to produce table 7.16.

| | LEAR | SVTO | HOLL | RAMY | PALE |
|---|---|---|---|---|---|
| LEAR | 99.98% | 30.54% | 27.21% | 29.03% | 27.90% |
| SVTO | 30.53% | 99.99% | 26.77% | 31.99% | 29.76% |
| HOLL | 27.20% | 26.75% | 99.98% | 32.36% | 42.09% |
| RAMY | 28.98% | 32.02% | 32.44% | 99.98% | 32.41% |
| PALE | 27.89% | 29.67% | 41.94% | 32.39% | 99.98% |

**Table 7.16**: Inter-Site USAF McIntosh Class Comparison

Table 7.16 shows that for the USAF the site comparisons are much worse, with Holloman and Palehua having the best comparisons. It should be noted that there appears to be a few errors in the region reports shown on the NGDC database, namely the positions seem to have been misreported. This is why Table 7.13 does not have 100% match between the same sites. These errors appear to be rare and do not significantly affect any conclusions. Also, the GONG comparisons did not have a rudimentary penumbra class, so the USAF

have one extra class category to match.

Figures 7.19 and 7.20 show histograms of the GONG and SRS McIntosh classifications respectively. These graphs, in part, explain why the GONG matches are better than the SRS matches.

GONG has nearly 30% of regions with a class of AXX, while the SRS AXX regions occur less than 10% of the time. With 3 times more regions in one class, it is not surprising that the inter-site matches are better with GONG than with the USAF system. Figure 7.19 also shows that the number of regions in each class is remarkably consistent across the GONG sites.

McIntosh Class comparisons between the GONG and SRS classifications were made. Initially the full 3 parameter classifications were compared. These results are shown in Table 7.17

Table 7.17 shows that the class match between GONG and SRS is poor. Part of this is explained by the number of AXX regions GONG identifies as stated earlier. The other point to note is that there is no match for a region with an SRS penumbral class of R. If all of the regions with penumbral classes of R in figure 7.20 are added together, these regions only comprise about 7% of all regions, so this cannot account for the low matching rate.

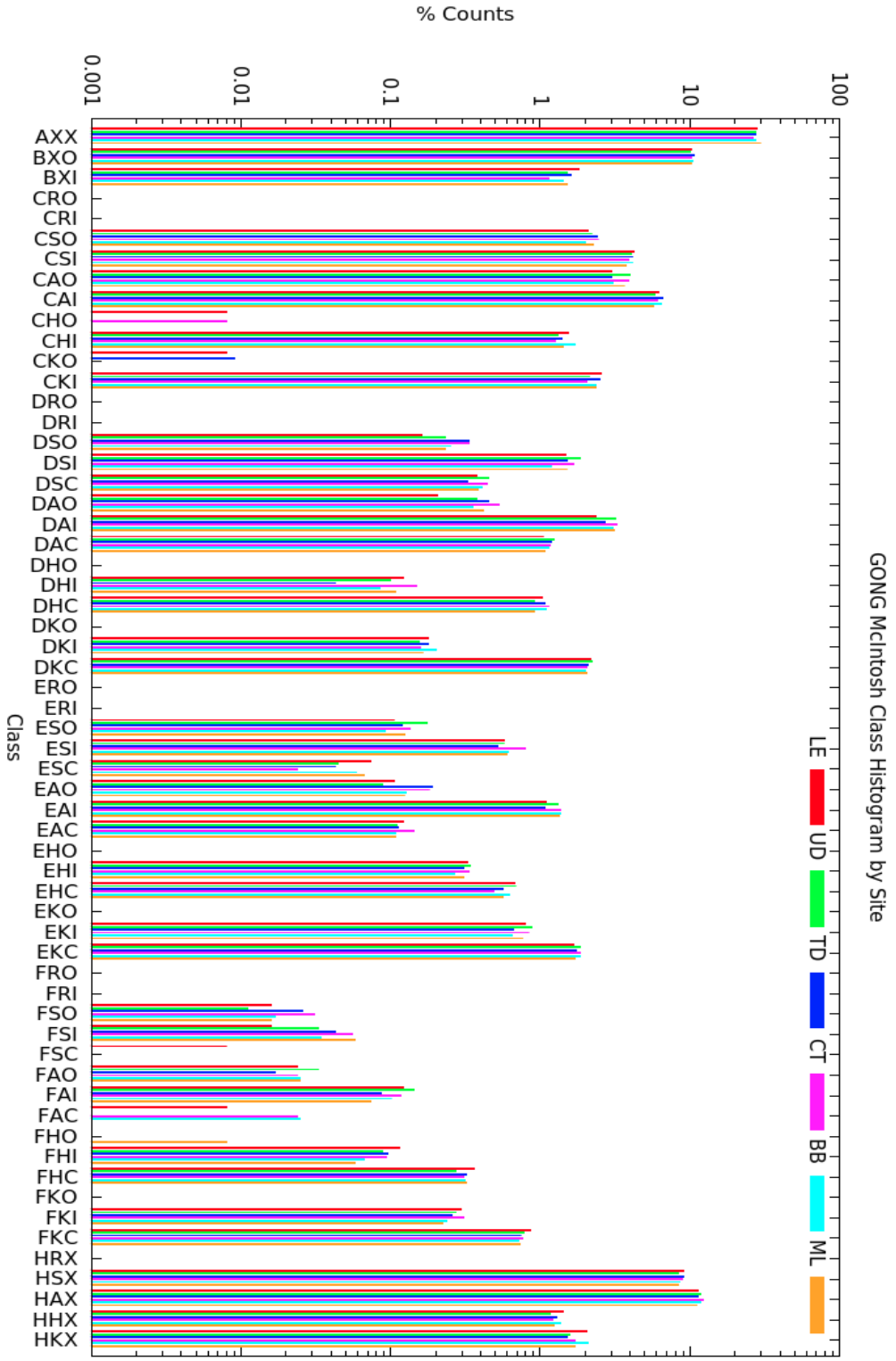



**Figure 7.19**: GONG McIntosh Class Histogram by Site

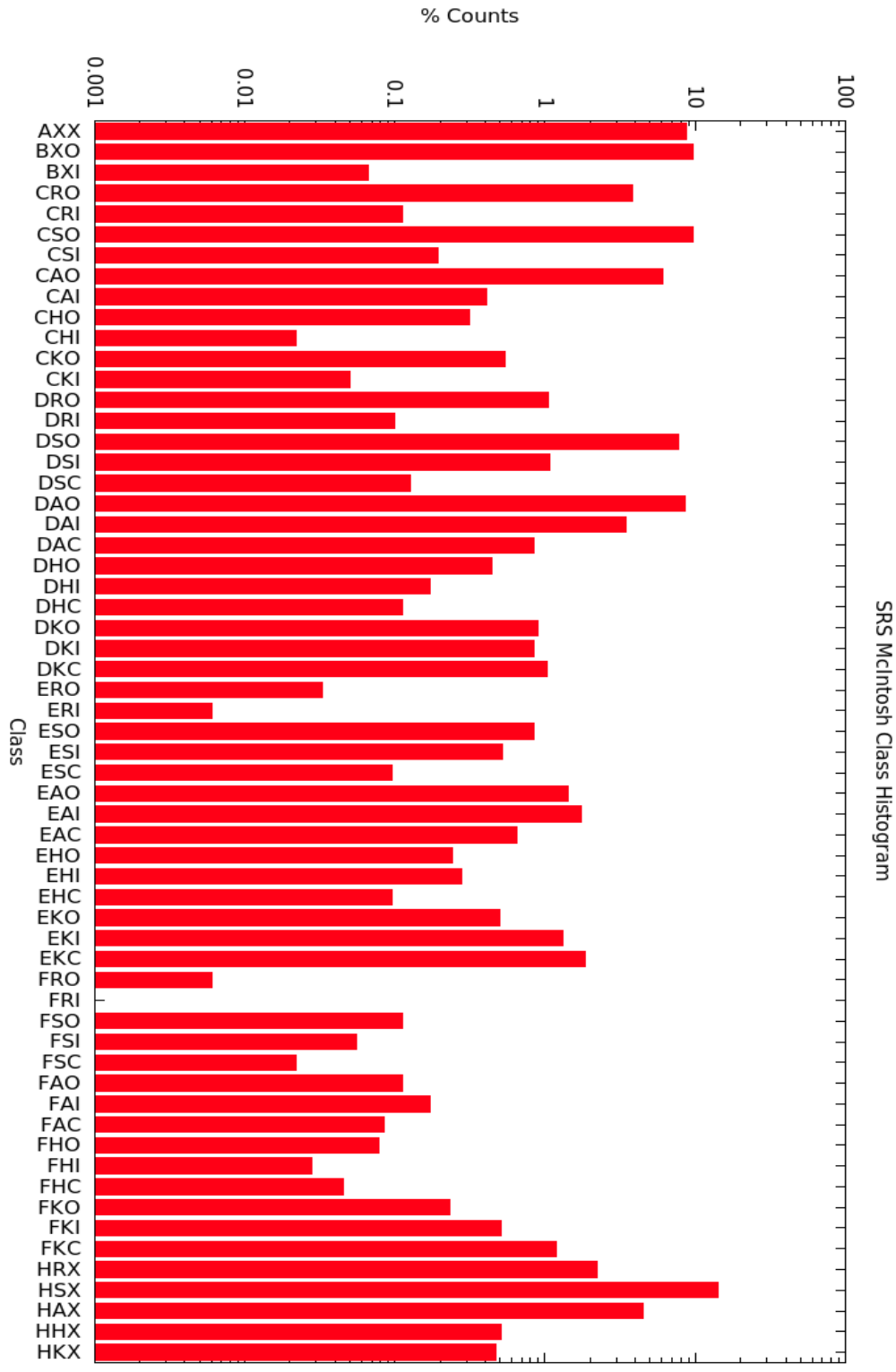



**Figure 7.20**: SRS McIntosh Class Histogram

| Site | %SRS Matched |
|---|---|
| LE | 16.141 |
| UD | 16.877 |
| TD | 15.019 |
| CT | 16.639 |
| BB | 14.502 |
| ML | 15.564 |

**Table 7.17**: GONG - SRS Full McIntosh Class Comparison

To investigate further, the individual parameters of the McIntosh classification were compared. This method too has its problems as the individual parameters in the M parameters are not independent. Tables 7.18 to 7.20 show the comparison for Learmonth. The complete set of tables can be found in Appendix B.

These tables require some explanations. The 'Matches' row for each parameter is simply the number of GONG regions of a particular parameter that were matched with an SRS region with another parameter. The '%GONG' rows are the matches for a given class divided by the total number of GONG regions of that class. Similarly the '%SRS' rows are the matches for a given class divided by the total number of SRS regions of that class. For example, if the Z parameter (table 7.18) is considered and compare the GONG B parameter ($4^{th}$ row down) and the SRS D parameter ($6^{th}$ column across), then the number of GONG B regions that were matched with SRS D regions were 624. There were a total of 1347 GONG B regions and 3279 SRS D regions, giving %GONG of 46.33% and %SRS of 19.03%

These tables show that the automated procedure with GONG imagery has difficulty finding small spots, or resolving the penumbra on small spots. Nearly 60% of Z=A regions for GONG were classified as B, C or D regions in SRS reports . A significant proportion of H regions were also Classified as A. A large proportion of GONG C regions were classified as D regions in SRS reports.

This is further clarified by table 7.19 where a large number of P=X regions for GONG were classified as R, S or A regions in SRS reports.

Appendix C gives comparisons for specific examples of where the GONG and SRS McIntosh classifications differ.

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | A | B | C | D | E | F | H | Total GONG |
| A | Matches | 459 | 458 | 764 | 339 | 39 | 9 | 624 | 2692 |
| | %GONG | 17.05 | 17.01 | 28.38 | 12.59 | 1.45 | 0.33 | 23.18 | |
| | %SRS | 92.91 | 70.35 | 31.40 | 10.34 | 3.09 | 2.71 | 27.89 | |
| B | Matches | 17 | 174 | 413 | 624 | 77 | 18 | 24 | 1347 |
| | %GONG | 1.26 | 12.92 | 30.66 | 46.33 | 5.72 | 1.34 | 1.78 | |
| | %SRS | 3.44 | 26.73 | 16.97 | 19.03 | 6.11 | 5.42 | 1.07 | |
| C | Matches | 4 | 13 | 478 | 1119 | 394 | 60 | 143 | 2211 |
| | %GONG | 0.18 | 0.59 | 21.62 | 50.61 | 17.82 | 2.71 | 6.47 | |
| | %SRS | 0.81 | 2.00 | 19.65 | 34.13 | 31.25 | 18.07 | 6.39 | |
| D | Matches | 2 | 0 | 41 | 754 | 137 | 20 | 86 | 1040 |
| | %GONG | 0.19 | 0.00 | 3.94 | 72.50 | 13.17 | 1.92 | 8.27 | |
| | %SRS | 0.40 | 0.00 | 1.69 | 22.99 | 10.86 | 6.02 | 3.84 | |
| E | Matches | 1 | 0 | 5 | 89 | 495 | 52 | 4 | 646 |
| | %GONG | 0.15 | 0.00 | 0.77 | 13.78 | 76.63 | 8.05 | 0.62 | |
| | %SRS | 0.20 | 0.00 | 0.21 | 2.71 | 39.25 | 15.66 | 0.18 | |
| F | Matches | 0 | 0 | 1 | 8 | 40 | 150 | 0 | 199 |
| | %GONG | 0.00 | 0.00 | 0.50 | 4.02 | 20.10 | 75.38 | 0.00 | |
| | %SRS | 0.00 | 0.00 | 0.04 | 0.24 | 3.17 | 45.18 | 0.00 | |
| H | Matches | 11 | 6 | 731 | 346 | 79 | 23 | 1356 | 2552 |
| | %GONG | 0.43 | 0.24 | 28.64 | 13.56 | 3.10 | 0.90 | 53.13 | |
| | %SRS | 2.23 | 0.92 | 30.05 | 10.55 | 6.26 | 6.93 | 60.62 | |
| | Total SRS | 494 | 651 | 2433 | 3279 | 1261 | 332 | 2237 | |

**Table 7.18**: GONG Learmonth - SRS Z Parameter Comparison

| | | SRS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GONG | | X | R | S | A | H | K | Total GONG |
| X | Matches | 1108 | 671 | 1278 | 935 | 6 | 41 | 4039 |
| | %GONG | 27.43 | 16.61 | 31.64 | 23.15 | 0.15 | 1.02 | |
| | %SRS | 96.77 | 86.69 | 32.30 | 28.43 | 2.04 | 3.34 | |
| S | Matches | 16 | 50 | 1137 | 745 | 25 | 54 | 2027 |
| | %GONG | 0.79 | 2.47 | 56.09 | 36.75 | 1.23 | 2.66 | |
| | %SRS | 1.40 | 6.46 | 28.73 | 22.65 | 8.50 | 4.40 | |
| A | Matches | 20 | 51 | 1330 | 1299 | 41 | 128 | 2869 |
| | %GONG | 0.70 | 1.78 | 46.36 | 45.28 | 1.43 | 4.46 | |
| | %SRS | 1.75 | 6.59 | 33.61 | 39.50 | 13.95 | 10.42 | |
| H | Matches | 0 | 0 | 81 | 96 | 95 | 339 | 611 |
| | %GONG | 0.00 | 0.00 | 13.26 | 15.71 | 15.55 | 55.48 | |
| | %SRS | 0.00 | 0.00 | 2.05 | 2.92 | 32.31 | 27.61 | |
| K | Matches | 1 | 2 | 131 | 214 | 127 | 666 | 1141 |
| | %GONG | 0.09 | 0.18 | 11.48 | 18.76 | 11.13 | 58.37 | |
| | %SRS | 0.09 | 0.26 | 3.31 | 6.51 | 43.20 | 54.23 | |
| | Total SRS | 1145 | 774 | 3957 | 3289 | 294 | 1228 | |

**Table 7.19**: GONG Learmonth - SRS P Parameter Comparison

| | | SRS | | | | |
|---|---|---|---|---|---|---|
| GONG | | X | O | I | C | Total GONG |
| X | Matches | 2449 | 2594 | 147 | 54 | 5244 |
| | %GONG | 46.70 | 49.47 | 2.80 | 1.03 | |
| | %SRS | 89.71 | 46.34 | 9.83 | 6.26 | |
| O | Matches | 55 | 1478 | 252 | 25 | 1810 |
| | %GONG | 3.04 | 81.66 | 13.92 | 1.38 | |
| | %SRS | 2.01 | 26.40 | 16.84 | 2.90 | |
| I | Matches | 152 | 1364 | 837 | 341 | 2694 |
| | %GONG | 5.64 | 50.63 | 31.07 | 12.66 | |
| | %SRS | 5.57 | 24.37 | 55.95 | 39.51 | |
| C | Matches | 74 | 162 | 260 | 443 | 939 |
| | %GONG | 7.88 | 17.25 | 27.69 | 47.18 | |
| | %SRS | 2.71 | 2.89 | 17.38 | 51.33 | |
| | Total SRS | 2730 | 5598 | 1496 | 863 | |

**Table 7.20**: GONG Learmonth - SRS C Parameter Comparison

A similar process, as above, was performed with the magnetic class (description of these classes can be found in chapter 2). This analysis is somewhat simpler as there is only the one parameter with 7 classes. Table 7.21 shows the comparison between GONG sites. Here, the matches are better than with the McIntosh classification, but again the same trend is seen, that the further away in time the observations are, the poorer the match.

| | LE | UD | TD | CT | BB | ML |
|---|---|---|---|---|---|---|
| LE | 100.00% | 77.31% | 71.23% | 67.44% | 65.56% | 63.31% |
| UD | 77.11% | 100.00% | 73.55% | 69.32% | 65.31% | 63.76% |
| TD | 70.93% | 73.39% | 100.00% | 75.16% | 71.46% | 69.02% |
| CT | 67.30% | 69.25% | 75.24% | 100.00% | 76.99% | 74.42% |
| BB | 65.32% | 65.17% | 71.49% | 76.75% | 100.00% | 78.78% |
| ML | 62.92% | 63.53% | 68.78% | 74.17% | 78.71% | 100.00% |

**Table 7.21**: Inter-Site GONG Magnetic Class Comparison

Table 7.22 shows the USAF inter-site comparison for magnetic class. It can be seen that the magnetic classification is marginally better than for GONG and is also a little more consistent.

| | LEAR | SVTO | HOLL | RAMY | PALE |
|---|---|---|---|---|---|
| LEAR | 99.99% | 74.71% | 70.80% | 73.05% | 70.81% |
| SVTO | 74.72% | 99.99% | 74.00% | 75.61% | 72.96% |
| HOLL | 70.82% | 73.96% | 99.99% | 75.61% | 76.97% |
| RAMY | 73.04% | 75.64% | 75.72% | 99.99% | 74.33% |
| PALE | 70.90% | 72.93% | 76.83% | 74.24% | 99.99% |

**Table 7.22**: Inter-Site USAF Magnetic Class Comparison

Figures 7.21 and 7.22 show the distribution of magnetic classes for GONG and SRS. Of note is the difference with the δ configurations. While approximately the same proportion of regions in SRS and GONG have δ configurations, SRS finds far fewer γδ regions than does GONG. Additionally, there are far fewer γ regions found by SRS. This may mean that the criteria for γ region classifications in the automated algorithm are too generous.

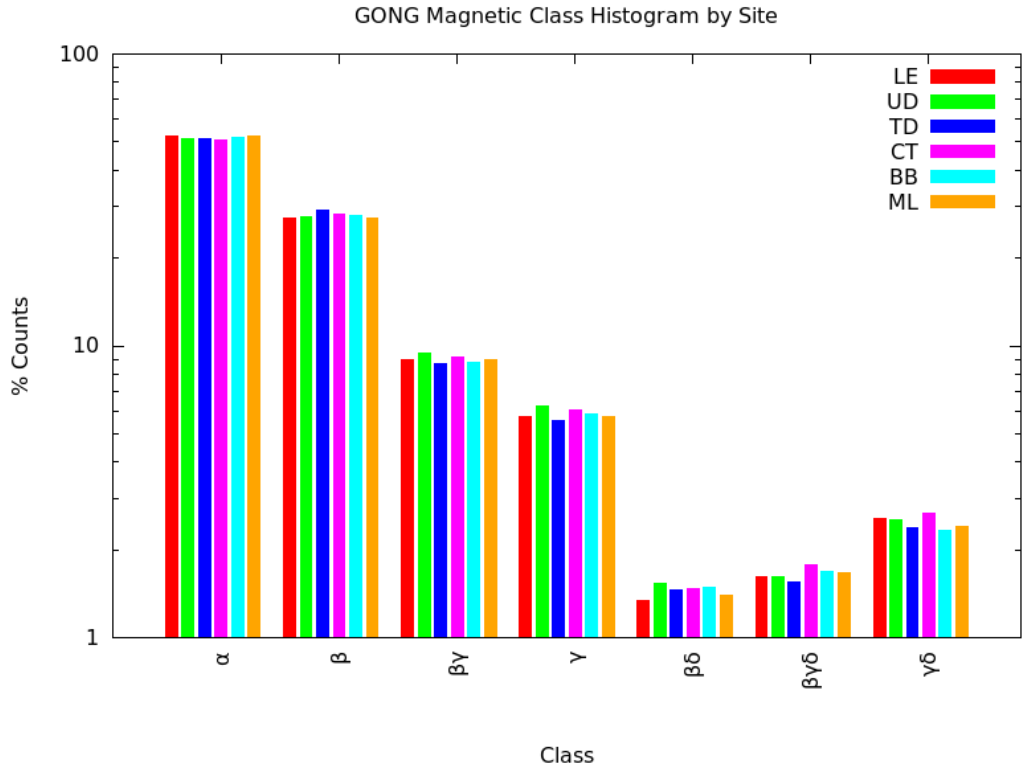



**Figure 7.21**: GONG Magnetic Class Histogram

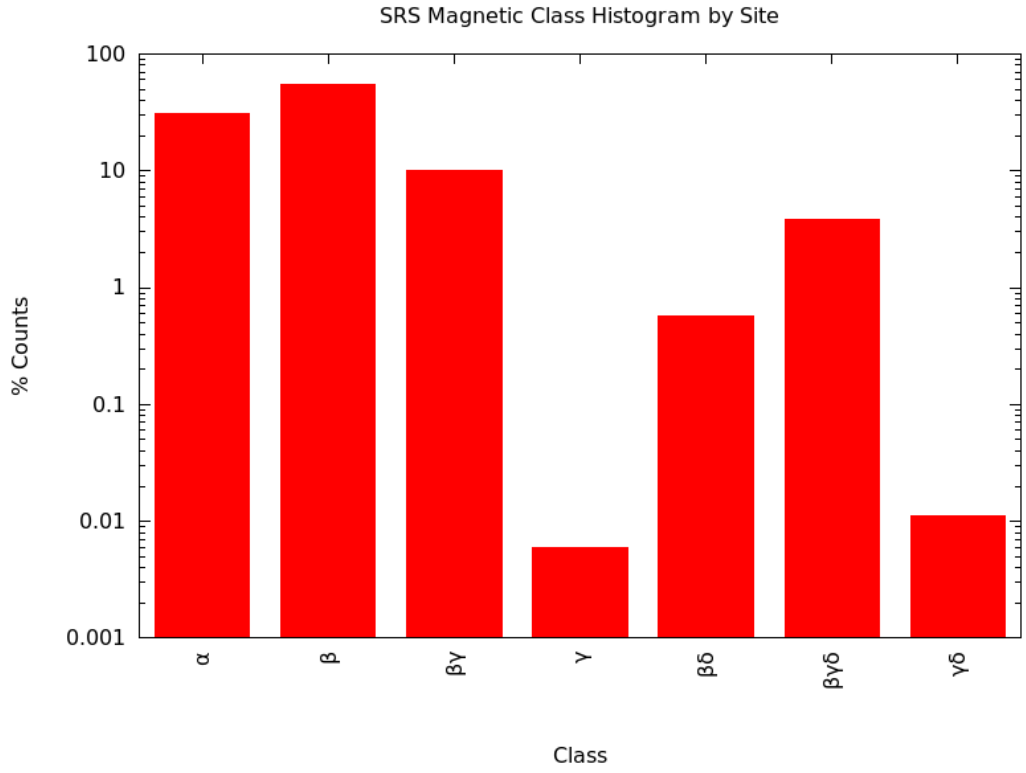



**Figure 7.22**: SRS Magnetic Class Histogram

Table 7.23 shows shows the comparison between GONG and SRS magnetic class for Learmonth. For example, consider the βγ row for GONG (3rd column down). There were 9 matches with βδ SRS regions (5th column across). There were a total of 1013 βγ GONG regions and 68 SRS βδ regions, giving 0.89% GONG matches between these classes and 13.24% SRS matches.

Of interest here is that no SRS γδ regions matched GONG γδ Regions. GONG is also finding more α regions than SRS. This is for the same reason mentioned previously, the GONG images have lower resolution. The tables for other sites can be found in Appendix D. Appendix E shows examples of when the magnetic classifications did not agree.

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | α | β | βγ | γ | βδ | βγδ | γδ | Total GONG |
| α | Matches | 2455 | 2604 | 178 | 0 | 11 | 52 | 0 | 5300 |
| | %GONG | 46.32 | 49.13 | 3.36 | 0.00 | 0.21 | 0.98 | 0.00 | |
| | %SRS | 89.76 | 43.63 | 12.95 | 0.00 | 16.18 | 9.67 | 0.00 | |
| β | Matches | 145 | 2364 | 524 | 1 | 10 | 61 | 0 | 3105 |
| | %GONG | 4.67 | 76.14 | 16.88 | 0.03 | 0.32 | 1.96 | 0.00 | |
| | %SRS | 5.30 | 39.61 | 38.14 | 100.0 | 14.71 | 11.34 | 0.00 | |
| βγ | Matches | 34 | 580 | 323 | 0 | 9 | 67 | 0 | 1013 |
| | %GONG | 3.36 | 57.26 | 31.89 | 0.00 | 0.89 | 6.61 | 0.00 | |
| | %SRS | 1.24 | 9.72 | 23.51 | 0.00 | 13.24 | 12.45 | 0.00 | |
| γ | Matches | 32 | 329 | 227 | 0 | 8 | 67 | 2 | 665 |
| | %GONG | 4.81 | 49.47 | 34.14 | 0.00 | 1.20 | 10.08 | 0.30 | |
| | %SRS | 1.17 | 5.51 | 16.52 | 0.00 | 11.76 | 12.45 | 100.0 | |
| βδ | Matches | 44 | 40 | 18 | 0 | 8 | 26 | 0 | 136 |
| | %GONG | 32.35 | 29.41 | 13.24 | 0.00 | 5.88 | 19.12 | 0.00 | |
| | %SRS | 1.61 | 0.67 | 1.31 | 0.00 | 11.76 | 4.83 | 0.00 | |
| βγδ | Matches | 14 | 32 | 39 | 0 | 8 | 85 | 0 | 178 |
| | %GONG | 7.87 | 17.98 | 21.91 | 0.00 | 4.49 | 47.75 | 0.00 | |
| | %SRS | 0.51 | 0.54 | 2.84 | 0.00 | 11.76 | 15.80 | 0.00 | |
| γδ | Matches | 11 | 19 | 65 | 0 | 14 | 180 | 0 | 289 |
| | %GONG | 3.81 | 6.57 | 22.49 | 0.00 | 4.84 | 62.28 | 0.00 | |
| | %SRS | 0.40 | 0.32 | 4.73 | 0.00 | 20.59 | 33.46 | 0.00 | |
| | Total SRS | 2735 | 5968 | 1374 | 1 | 68 | 538 | 2 | |

**Table 7.23**: GONG Learmonth - SRS Magnetic Class Comparison

### 7.4 Summary and Conclusions

A series of automated algorithms were developed to extract sunspot and magnetic field information from GONG intensity and magnetogram images. The 12 month smoothed sunspot numbers were compared with the international and AAVSO 12 month smoothed sunspot numbers for cycles 23 and 24. These results show a very good match to the International Sunspot Number and a poorer match to the AAVSO sunspot numbers. The GONG sunspot area was compared to the USAF sunspot area for 12 month smoothed averages. These fits were not as good with a significant difference between cycles 23 and 24.

The umbra to sunspot area ratio was calculated and compared with published literature. No significant differences were found in this ratio between the parts of the cycles 23 and 24 that were measured. For the entire region an umbra to sunspot area ratio was 20.6% and 20.4% respectively for each cycle. For individual spots they were found to be 12.6% and 12.4% for the largest spot in the region. These are slightly lower than the published values discussed in Section 2.3. There is some variation based on sunspot size, but this variation, strictly speaking, is not statistically significant. None the less, it is certainly possible that there is a real variation in this ratio based on sunspot size.

It is suspected that the reason for the lower ratio values is the algorithm not finding the entirety of the umbra. The main concern in algorithm development was detection and number of umbra for sunspot classification, rather than its accurate area measurement.

The agreement between GONG region locations and the Solar Region Summary locations produced by SWPC is around 90%, however, GONG finds significantly fewer regions. Thus SRS regions agree only around 73% of the time. The GONG images are lower resolution than the USAF SOON white light board produces, so that smaller regions, in particular, are not found.

Additionally, the GONG images are cropped at 95% of the solar radius to remove the annulus of the disk where the limb correction algorithm does not work at the extreme edge. Whilst this may seem like a small area of the disk to remove, this annulus comprises 31% of the disk area. However, sunspots generally only occur at latitudes between -40º and +40º. The omitted area in this range is 22%. A rough estimate then would be that up to 22% of regions are being missed by the GONG automated procedure which would easily take into account the difference between GONG and SRS reports. Large regions are less likely to be missed though, as some part of them may still be visible on the uncropped portion. In addition, smaller spots may disappear entirely due to limb foreshortenning. By visual inspection of a selection of GONG-produced images, it is clear that some small regions (normally with a Z parameter of A or B) are being missed entirely.

The Sunspot Area to Sunspot Number relationship between cycles 23 and 24 was found to vary in a statistically significant way. However, care needs to be taken interpreting this result as only the declining phase of cycle 23 and rising phase of cycle 24 were measured.

Thus, this difference may simply be variations in the sunspot area to sunspot number ratio at different times in the sunspot cycle. However, even if it is just variations in a cycle, care needs to be taken using this ratio as a correction for historical sunspot area, as there is at least variations on time scales of a sunspot cycle. This ratio is one of the reasons why an automated technique is useful, as the same procedure is applied on all of the data, and becomes even more powerful when used on data sets from identical (or as close as one can get to identical) equipment over long time scales.

For the same reasons detailed above, the GONG automated McIntosh classification system tends to miss small sunspots and has difficulty resolving penumbra. Small regions with a D parameter, for example, are sometimes classified as B or C. For this reason the full GONG McIntosh classification is only in agreement about 16% of the time with the SRS classification. Of course, there are 3 parameters to match and, individually, each parameter performs better than this. Visual inspection of a selection of images shows that it is unusual for GONG to give an incorrect classification based on the quality of the images used.

The consistency between individual sites for GONG classification varies between 40% and 65%, largely based on the time difference between observations. For example, Learmonth (observations typically taken at the start of the UT day and no later than 0900 UT) has a poor match with Mauna Loa (observations taken between 1900 and 2300 UT) of only 43%. The matches between individual USAF SOON sites is worse; between 25 and 40%.

One of the reasons for the better match is that nearly 3 times more GONG regions than SRS regions are classified as AXX. Thus, it is easier to correctly match for these simple regions.

The magnetic classification matches are somewhat better than those of the McIntosh class, simply because there is only one parameter to match. GONG has about a 70% consistency between its own sites as does the USAF sites. Generally there is a fair match between the SRS and GONG with again GONG sites over reporting unipolar groups. The worst are the δ regions. While the number of USAF and GONG delta regions are similar, there is clearly a significant difference in how the delta parameter is being applied. This maybe due to the lower resolution of the GONG images, which are unable to either

resolve small umbra or small portions of opposite magnetic field.

In summary, the automated GONG analysis produces excellent results for sunspot number and sunspot area. The sunspot classification is limited somewhat by the spatial resolution of the images and also by the inability for the procedure to resolve rudimentary penumbra. Given these constraints, the GONG sunspot classification system does work well.

# 8. Sunspot Parameters Associated with X-ray Flares

## 8.1 Introduction

When flares are referred to, it is usually X-ray flares that are being considered, rather than Hα flares. X-ray flares are classified by their intensity and assigned a class (Phillips, 1992). This can be seen in the X-ray plot shown in Figure 8.1. There are two different frequency ranges that the GOES satellites observe: 0.5 to 4.0 Å and 1.0 to 8.0 Å. Normally it is the latter spectral interval that is used when classifying X-ray flares. One of the reasons for having two channels is that the flare temperature can be derived if the sensor properties are sufficiently well characterised (Garcia, 1994).

X-ray flares are specified by the observed irradiance as seen in figure 8.1. Table 8.1 shows a summary of the classification system. Thus an M5.4 flare would have an X-ray power of $5.4\times10^{-4}$ $Wm^{-2}$. The GOES system saturates at about X17, and extrapolation techniques need to be used to measure flares that saturate the detectors.

It is noted here that, in classifying X-ray flares, the background emission is not subtracted. For significant flares, this does not pose a problem as the background will be typically less than ten percent of the peak flux. Also, the analysis performed in this chapter is concerned with predicting X-ray flares which are classified as M class or higher.

Like other parameters discussed earlier, the flare rate varies in different sunspot cycles. This is shown in figure 8.2 for cycles 21 to 22. The flare rate in cycles 21 and 23 lagged the sunspot number peak. It can also be seen in cycles 23 and 24 that the flare rates are lower than in cycles 21 and 22. This may be at least in part due to active regions being smaller in these cycles (as shown in chapter 4). Fig 8.2 also shows that the twin peak due to hemispheric differences can be even more pronounced in flare rate than sunspot number, particularly for cycle 22. This lack of consistency across cycles makes flare prediction somewhat more difficult.

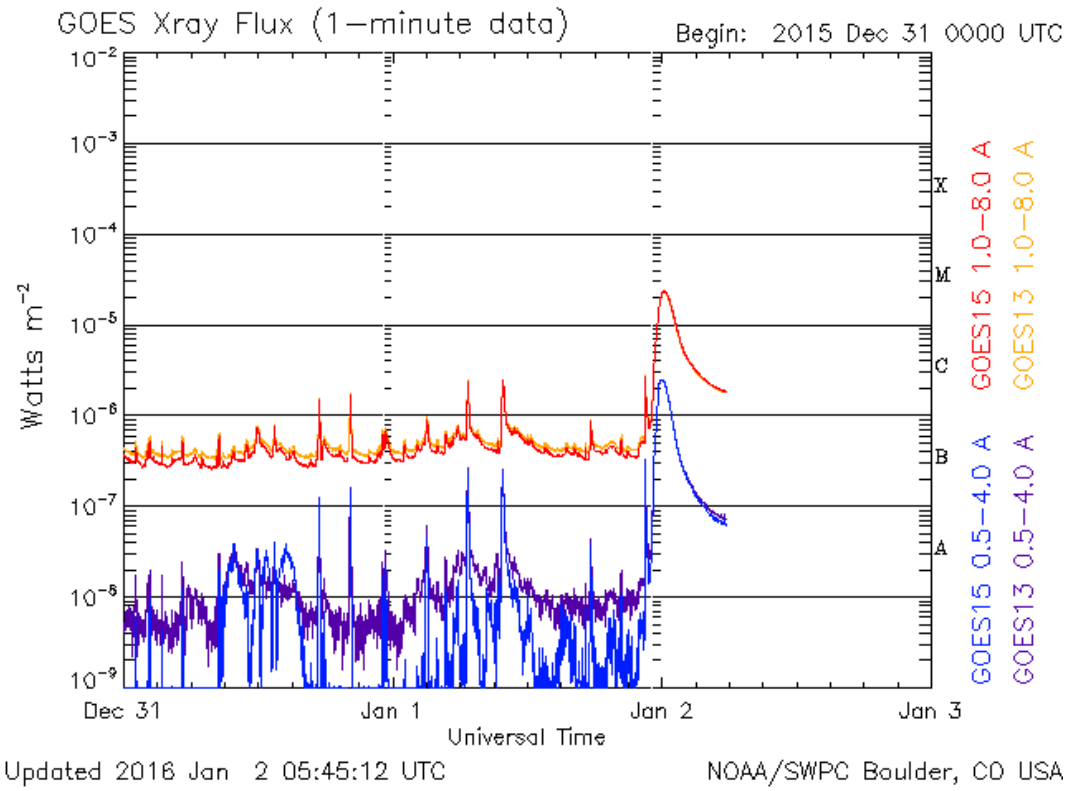



**Figure 8.1**: An M2 flare observed on the 2nd Jan 2016

(image from http://www.swpc.noaa.gov/products/goes-x-ray-flux)

| X-Ray Class | Irradiance ($Wm^{-2}$) |
|---|---|
| A | $10^{-8}$ |
| B | $10^{-7}$ |
| C | $10^{-6}$ |
| M | $10^{-5}$ |
| X | $10^{-4}$ |
| X10 (Y) | $10^{-3}$ |

**Table 8.1**: X-Ray Flare Classification System

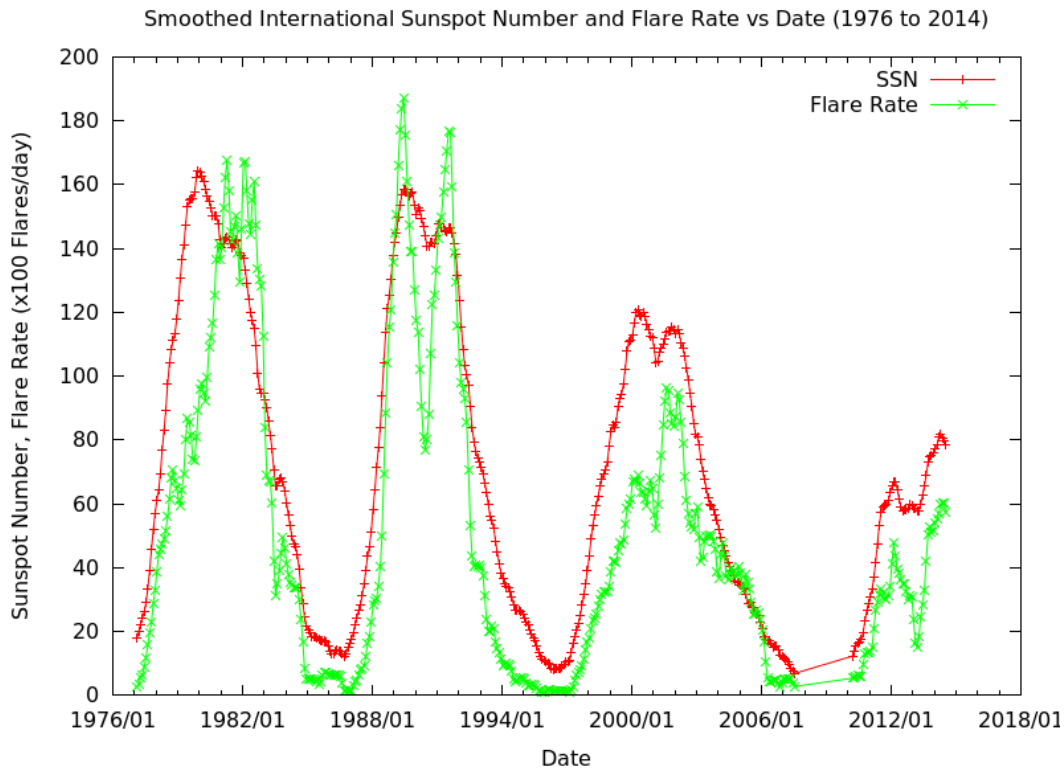



**Figure 8.2**: Smoothed Sunspot Number and Smoothed Flare Rate from 1976 to 2014

Since Hα observations started, observers have realised that there is at least some correlation between a region's sunspot classification and how flare productive it will be. The first person to attempt to correlate sunspot features with Hα flares was Giovanelli (1939). He considered the sunspot area, sunspot area growth, and magnetic classification for sunspot regions. He derived Equation 8.1 which links these parameters (Giovanelli, 1939).

$$R = 2.6\,k\,.\,a\,.\,f(i) \quad , \quad \text{Eqn 8.1}$$

where $R$ is the flare rate, $k$ is a weighting factor for the magnetic classification of the region, $a$ is weighting factor for the sunspot area and $f(i)$ is a function for the growth rate of the sunspot area. 2.6 was a scaling factor introduced to account for the lack at the time of continuous observations (an average of 9 hours per day).

Giovanelli (1939) did look at the magnetic field strength, but the flare rates showed no correlation with magnetic field strength. However, this was probably due to limitations of

the spectrographs at the time, rather than there being no correlations (the lowest value reported was 1000 Gauss). Subsequent research has shown a link between magnetic fields and flares (for example Schrijver, 2007).

All of these parameters are still used in some form in modern flare predictions.

Both the sunspot class and magnetic class are used for flare prediction. One of the reasons McIntosh modified the Zurich scheme was because he found it a poor predictor for flares (McIntosh, 1990).

The Space Environment Services Center (now the Space Weather Prediction Center) used weights of various parameters to determine a flare rate (Sawyer and Warwick, 1986). These weighted parameters were quite extensive, some of which do not seem to be available. For example inversion line complexity which measures the number of kinks in the neutral line of a region. SOON analysts reported many sunspot region parameters, including number of kinks in the neutral line, for a few years in the latter half of the 1980s (personal communication, John Kennewell).

One method of flare prediction is simply to determine if a specific region has produced a flare in the past (i.e. persistence). If it has, then it is likely to produce another. This process was formalised using Bayesian statistics (Wheatland, 2005). The idea of persistence does seem rather obvious, as if the conditions are suitable for one flare to occur, it is likely that, as long as those conditions remain, another will occur. Whilst not expressed in this form, the SWPC has been using flare history of a region for flare prediction since at least 1980. In addition, this approach does not not help elucidate which physical properties of an active region lead to flare generation.

More typically, sunspot class is used often in conjunction with other parameters. Sunspot growth for example, is thought to be significant. A study showed that, for the most productive regions, an increase in area corresponded to an increase in flare rate (Lee, et al., 2012). However this study only examined 11 of the McIntosh classes and did not separate these into cycles. In addition Lee et al. (2012) only investigated if the sunspot area of the region increased, decreased or stayed the same size; they did not categorise the spot growth by an actual quantified rate. Thus the study by Lee et al. (2012) was somewhat limited.

In the following discussion, various solar features were examined in order to determine which were good predictors of solar flare rates. If a Poisson distribution is assumed then the probability of at least one flare occurring will be given by equation 8.2,

$$P(t) = 1 - e^{-Rt} \quad , \qquad \text{Eqn 8.2}$$

where $P$ is the probability of at least one flare occurring, $R$ is the flare rate in flares per day and $t$ is the time interval.

The time period of interested in is simply one day.

In the analysis to follow, some of the properties of interest are discrete classes (sunspot and magnetic class, orientation class), while others are continuous distributions (sunspot area and area growth, magnetic field, inversion line complexity). In order to compute flare rates for these continuous parameters, the following procedure was used.

- A particular continuous parameter, $p$, is binned into several ranges. The number of regions in each bin was determined. For example, the first five sunspot area bins were in μh: 0 to 25, 25 to 50, 50 to 75, 75 to 100 and 100 to 150.
- Flares were matched with the appropriate region, and the number of flares occurring in each of the parameter bins determined.
- A flare rate was then determined using equation 8.3.

$$R = \frac{N_p}{N_f} \quad , \qquad \text{Eqn 8.3}$$

where $R$ is the flare rate, $N_p$ is the number of regions within a given bin and $N_f$ the number of flares that were associated with a region falling into the appropriate bin.

The flare rate was then plotted against the average value in each bin. Curves of best fit were then found. Thus a flare rate for a given parameter could just be read off the graph.

$\chi^2$ values for these plots were found by comparing the number of observed flares for each bin with the expected flares in each bin (scaled appropriately for the number of regions). Thus comparisons across cycles with various models could be performed. Details of the

$\chi^2$ test can be found in Freund (1988).

For classes obtained from GONG and SWPC prior to 1996 (because solar region summaries were not available electronically prior to this date), the individual reports from the observatories were used. From these reports a single sunspot and magnetic class had to be obtained. SWPC has an analyst that takes all of these reports, as well as satellite imagery, and manually analyses each region to produce a sunspot class (personal communication, SWPC Customer Support). Manual analysis was not practical for the data prior to 1996 due to the large volume. On each day for each region the class with the highest number of reports was chosen. The site which reported this class closest to the centre of the disk was chosen as the best site. This was to minimise limb fore-shortening errors for the areas.

## 8.2 Sunspot Area and its Growth as a Predictor for X-Ray Flares

### 8.2.1 Sunspot Area

Figure 8.3 shows the Flare Rate vs Sunspot Area plot for cycles 21 to 24 using SWPC sunspot areas. It should be noted that cycles 21 and 24 are only partial cycles.

The linear fits were forced through the origin. This approach was taken, as often the line of regression, at low areas, produced negative flare rates, so forcing the fit through the origin is physically sensible.

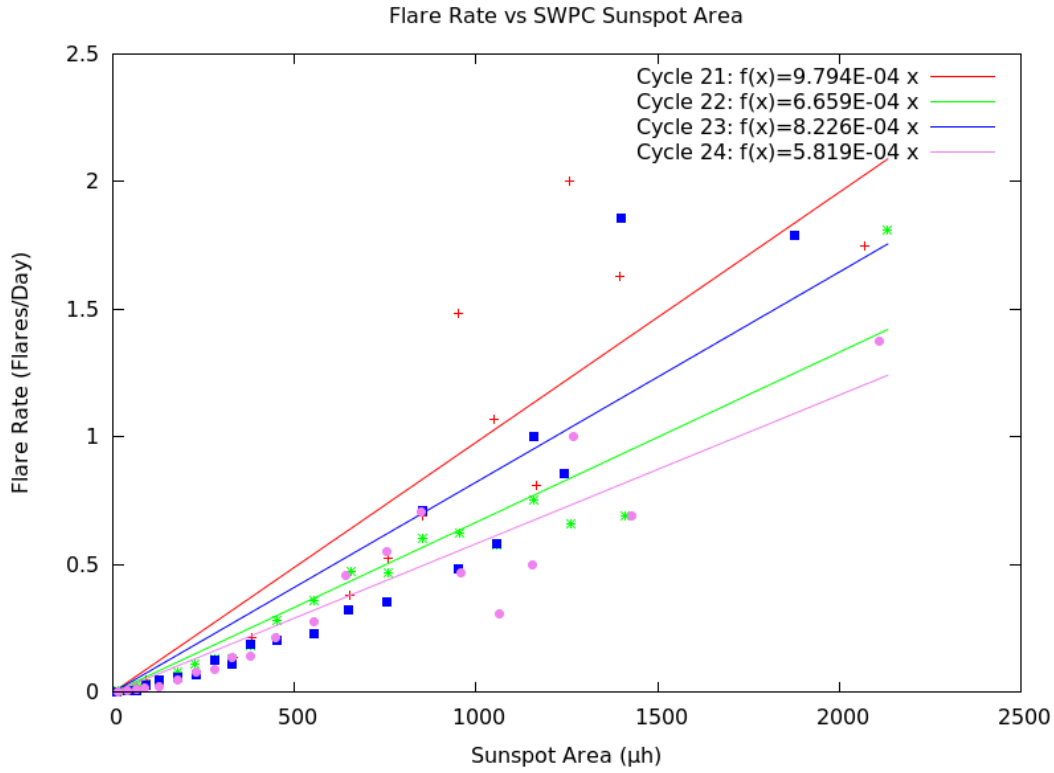



**Figure 8.3**: Flare Rate vs SWPC Area for Cycles 21 to 24

The first thing to note is that the individual cycles have very different flare rates. Secondly, the scatter in the points is quite broad.

A comparison was done between the model (linear fit) and observed flares. This analysis is performed by calculating the expected number of flares for each area as shown in equation 8.4:

$$F_e = N_p R_m \quad , \qquad \text{Eqn 8.4}$$

where $F_e$ is the expected number of flares for a given sunspot area, $N_p$ is the number of regions found for that area and $R_m$ is the model flare rate determined by the equations shown in figure 8.3.

This was then compared to the number of observed flares to generate the values used in the calculation of the $\chi^2$ statistic. The same process is used for other continuous parameters.

Table 8.2 shows that the model fits are a poor representation of the observed data as the $\chi^2$ values are far higher than the confidence intervals. From this result the only real statement that can be made is that as a general rule, the flare rate increases with area, but not in any statistically predictable way.

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 21 | 20 | 276.4 |
| 22 | 20 | 172.0 |
| 23 | 20 | 580.2 |
| 24 | 19 | 124.8 |

**Table 8.2**: SWPC Observed Flares vs Model Flares by Sunspot Area $\chi^2$ Values. The confidence interval for 20 degrees of freedom (DoF) is 31.4 and for 19 DoF is 30.1.

A similar comparison of flare rates versus sunspot area was performed using the processed GONG areas. Figure 8.4 shows the flare rates vs area for GONG. Here, it is seen that the scatter is similar to that of the SWPC data. Table 8.3 shows the $\chi^2$ values for the lines of regression compared to observed data. These are comparable with SWPC data. The poor correlations between flare rate and sunspot areas do not appear to be caused by area measurement uncertainties, but rather something intrinsic to the physical flare mechanism.

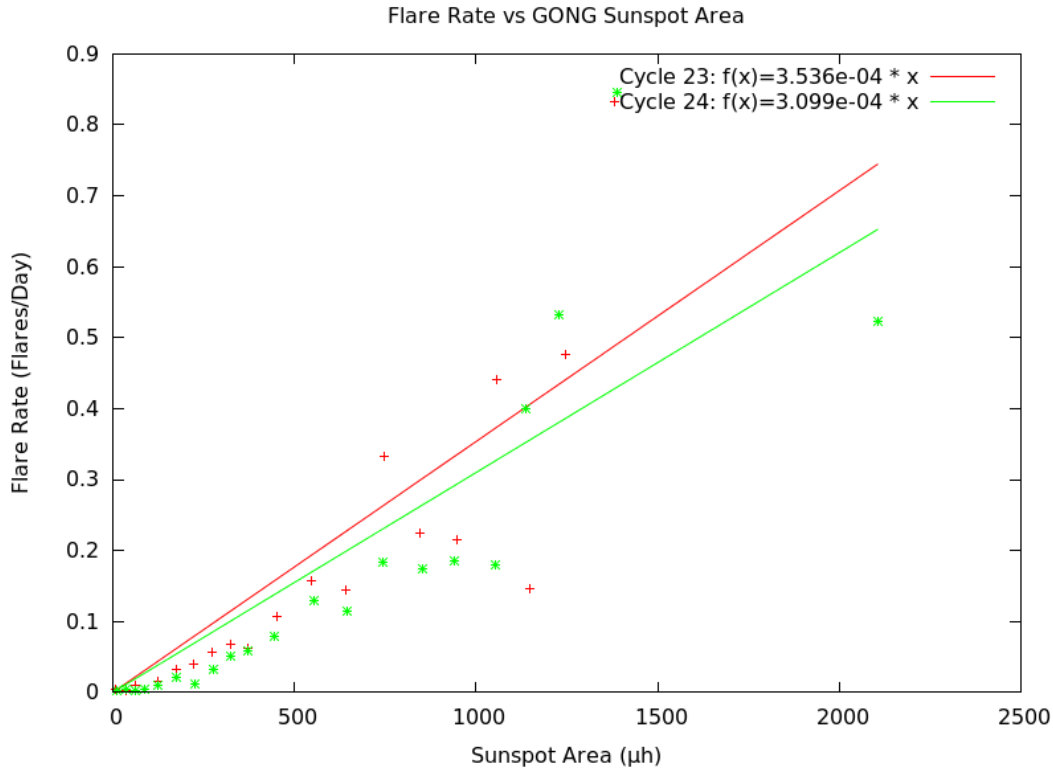



**Figure 8.4**: Flare Rate vs GONG Sunspot Area

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 23 | 20 | 179.6 |
| 24 | 20 | 138.7 |

**Table 8.3**: GONG Observed Flares – Model Flares by Sunspot Area $\chi^2$ Values. The critical value for 20 degrees of freedom is 31.4.

### 8.2.2 Absolute Sunspot Area Growth

The rate of absolute sunspot area growth was examined using a similar approach. The absolute sunspot growth for each day was calculated by equation 8.5:

$$G_A = \frac{A_1 - A_0}{\delta t} \quad , \qquad \text{(Eqn 8.5)}$$

where $G_A$ is the growth rate of a sunspot group in μh/day, $A_0$ is the initial area of the region, $A_1$ is the area of the region on the following day and $\delta t$ is the time in days between

the observations.

Once the area growths were found, they were binned and flare rates computed in the same manner as sunspot area. Figure 8.5 shows these results for SWPC data. A piecewise function was used in an attempt to model the flare rates with both the negative and positive growth, both functions being linear.

It is immediately obvious that, on average, a change of sunspot area leads to an increase in flare rate, regardless as to whether the change is positive or negative. However the fits to linear correlations are very poor, and the scatter worse than for the sunspot area vs flare rate plots in figures 8.3 and 8.4. Also the fits vary greatly between sunspot cycles. This seems to show that absolute area growth is a poorer variable for use in flare prediction than the area of the sunspots.

The $\chi^2$ values between observed flare rates and model flare rates (table 8.4) are considerably larger than the critical values for the degrees of freedom (DoF) used, showing that the linear models do not represent the observed data well. The area growth model performs as poorly as using the area only model, and significantly worse in cycle 22.

As seen in Figure 8.6, the absolute area growth rate using GONG data is a very poor parameter for predicting flares. No reasonable lines could be fitted to this data.

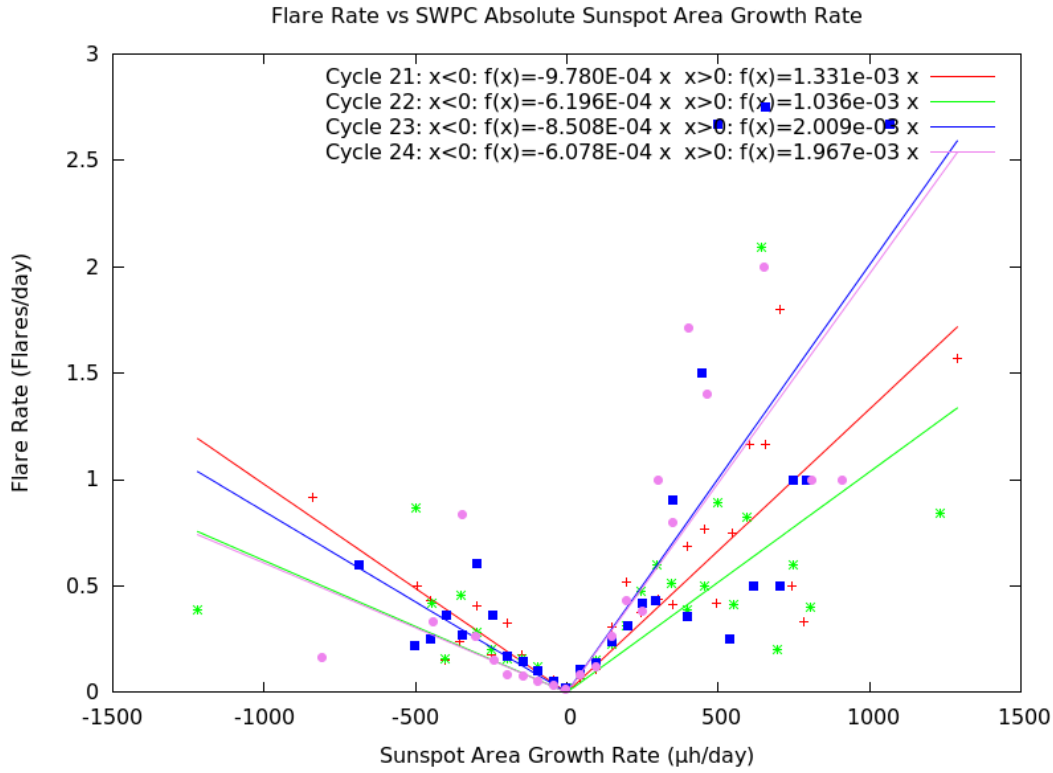



**Figure 8.5**: Flare Rate vs SWPC Absolute Sunspot Area Growth Rate

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 21 | 27 | 108.6 |
| 22 | 27 | 720.7 |
| 23 | 26 | 240.3 |
| 24 | 24 | 278.5 |

**Table 8.4**: SWPC Observed Flares vs Model Flares by Absolute Sunspot Area Growth $\chi^2$ Values. The critical values for 27, 26 and 24 DoF are 40.1, 38.9 and 36.4 respectively.

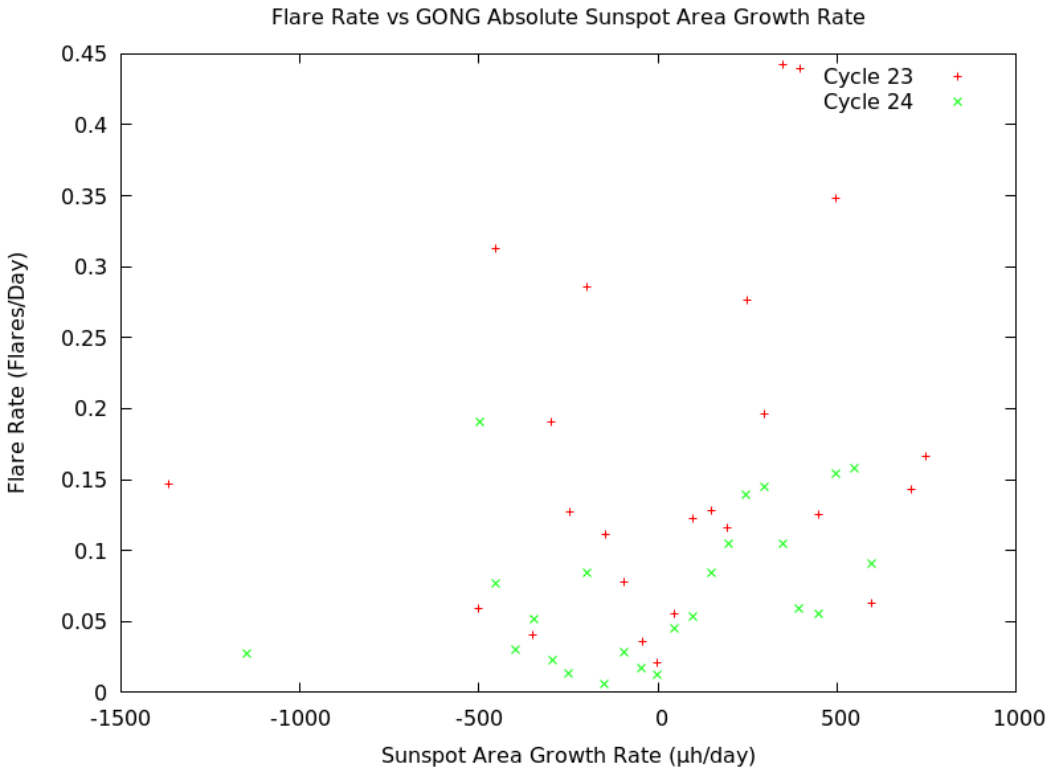



**Figure 8.6**: Flare Rate vs GONG Absolute Sunspot Area Growth Rate

### 8.2.3 Relative Sunspot Area Growth

Relative sunspot area growth was considered next. It is a proportional growth rate as shown in equation 8.6:

$$G_r = \frac{2}{\delta t}\left(\frac{A_1 - A_0}{A_1 + A_0}\right) \quad , \qquad \text{Eqn 8.6}$$

where $G_r$ is the relative growth rate and all of the other variables are the same as equation 8.4.

Figure 8.7 shows a scatter plot of flare rate versus relative sunspot area, measured by SWPC.

No decent linear or polynomial fits could be found for these data, at least not consistently over the 4 cycles. Relative sunspot area growth is therefore poorest predictor examined for flare rate, in terms of its statistical dependence as a dependent variable.

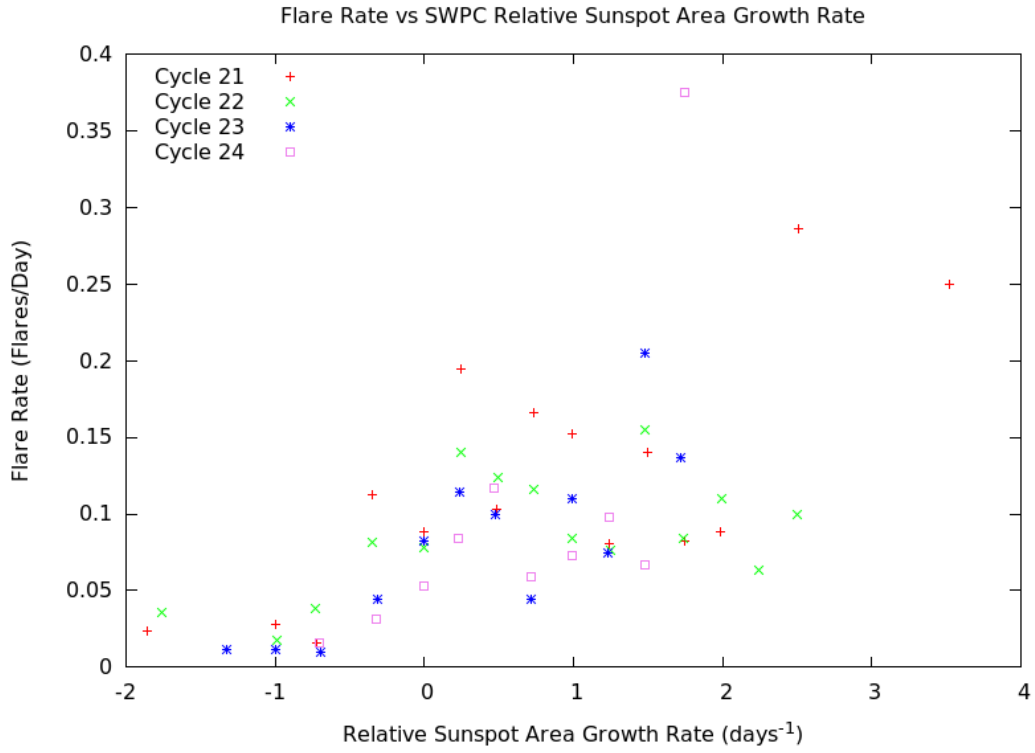



**Figure 8.7**: Flare Rate vs SWPC Relative Sunspot Area Growth Rate

Figure 8.8 shows the flare rate vs relative sunspot area growth for GONG. No good linear or polynomial fits could be found for these data also. Again, the conclusion is drawn that relative sunspot area is a very poor predictor of flares.

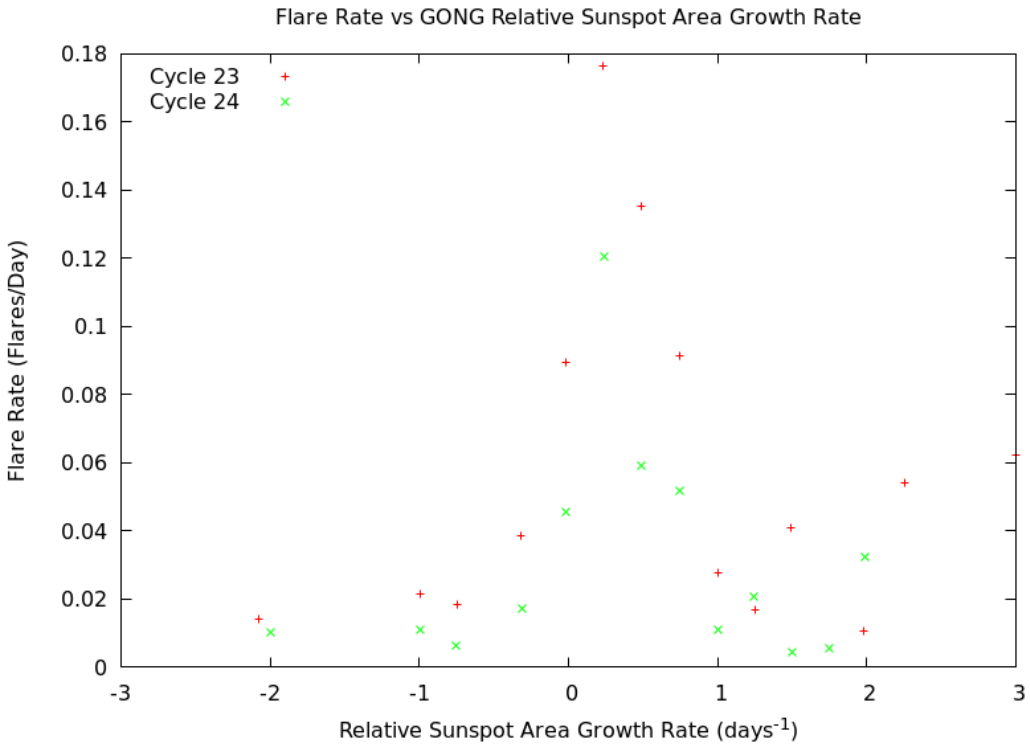



**Figure 8.8**: Flare Rate vs GONG Relative Sunspot Area Growth Rate

Of all of the types of sunspot area measurement, or its changes, the sunspot area itself is the best quantitative predictor of flares. However the relationship changes from one cycle to another and so needs to be recalculated in each cycle once sufficient statistics are collected.

### 8.3 Sunspot Class as a Predictor for X-ray Flares

An analysis was performed comparing the flare rates categorised by McIntosh sunspot class between sunspot cycles. As with the previous analysis, the number of each region of a given class was found, as was the number of flares for each class. From this a daily flare rate was computed. A $\chi^2$ contingency table was constructed comparing each cycle with every other cycle. As usual, any class that produced an expected number of flares less than five was combined with another class so that the combined class had the number of expected flares greater than 5. The individual flare rates for each class can be found in appendix F.

Table 8.5 shows the relationship between solar cycles, using a contingency table

with critical values (the CV rows) for each cycle. It should be noted that the cycle 20 flare rates were compiled by Kildahl (1980). Only the flare rates were available and as a result cycle 20 could not be used as an observed cycle.

| | | Expected Cycle | | | | |
|---|---|---|---|---|---|---|
| Observed Cycle | | 20 | 21 | 22 | 23 | 24 |
| 21 | $\chi^2$ | **489.0** | **0.0** | **210.2** | **809.5** | **1426.8** |
| | DoF | 32 | 24 | 23 | 20 | 17 |
| | CV | 46.2 | 36.4 | 35.2 | 31.4 | 27.6 |
| 22 | $\chi^2$ | **1747.5** | **572.8** | **0.0** | **560.7** | **1564.7** |
| | DoF | 38 | 30 | 34 | 27 | 23 |
| | CV | 53.4 | 43.8 | 48.6 | 40.1 | 35.2 |
| 23 | $\chi^2$ | **2240.1** | **1224.8** | **366.8** | **0.0** | **317.4** |
| | DoF | 38 | 28 | 30 | 33 | 26 |
| | CV | 53.4 | 41.3 | 43.8 | 47.4 | 38.9 |
| 24 | $\chi^2$ | **1143.3** | **960.7** | **296.8** | **114.8** | **0.0** |
| | DoF | 36 | 21 | 24 | 23 | 20 |
| | CV | 51.0 | 32.7 | 36.4 | 35.2 | 41.4 |

**Table 8.5**: SWPC $\chi^2$ Flare Rate Contingency Table for Mcintosh Sunspot Classes

Firstly, the $\chi^2$ values are significantly higher than the corresponding critical values, indicating that flare rates based on McIntosh sunspot class differ significantly between cycles. Secondly the $\chi^2$ values increase the further apart the sunspot cycles are. That is the flare rates get increasingly different the further apart in time they are. Therefore, if sunspot classes are going to be used as a predictor for flares, it needs to be clearly stated for which cycle the flare statistics were computed.

The same analysis was conducted using the GONG sunspot classifications. For this analysis, both cycles 23 and 24 are incomplete. Appendix G contains the full flare rate data for the GONG sunspot class flare rates. The $\chi^2$ value for cycle 23 as the observed cycle and cycle 24 as the expected value was 179.6 with 19 degrees of freedom. This compares with the critical value of 30.14. The GONG automated classification performs similarly to the SWPC manual reduction for the purposes of flare prediction.

Here, a brief note on the Bornmann and Shaw (1994) flare prediction model.

Bornmann and Shaw (1994) produced two different models, linear and multiplicative, in an attempt to predict flares. In these models each individual parameter in the sunspot class is treated as independent. Each parameter is assigned a flare weight. In the linear model these weights are then added to give the expected flare rate. In the multiplicative model the weights are multiplied.

Bornmann and Shaw set as the origin of the system A,B,H for the Zurich parameter and both of the X classes for the penumbra and compactness classes. These are combined in one parameter for the purposes of this model. Thus in total there are 12 weights that need to be found. To create a flare rate from the weights, equation 8.7 is used for the linear model,

$$y(x_i) = a_0 + \sum_{j=1}^{m} a_j x_{ij} \quad , \qquad \text{Eqn 8.7}$$

where $y(x_i)$ is the flare rate for the $i$th valid McIntosh parameter, $a_0$ is the origin of the system, $a_j$ is the weight for the $j$th parameter and $x_{ij}$ is a delta function which is unity if the parameter is in the $x_i$ class and zero otherwise.

Similarly the multiplicative scheme can be used, but in this case the weights are $b_j$'s, as shown in equation 8.8:

$$y(x_i) = b_0 \prod_{j=1}^{m} b_j^{x_{ij}} \quad , \qquad \text{Eqn 8.8}$$

Figure 8.9 shows Bornmann and Shaw's results for cycle 20, exhibiting large scatter. Whilst an interesting intellectual exercise, one would be better off using the normal sunspot class to obtain the flare rate, rather than this scheme. In addition the multiplicative model does not produce a 1:1 ratio for observed and expected flares. That is, the model underestimates the number of flares. For this reason , this model was not pursued any further.

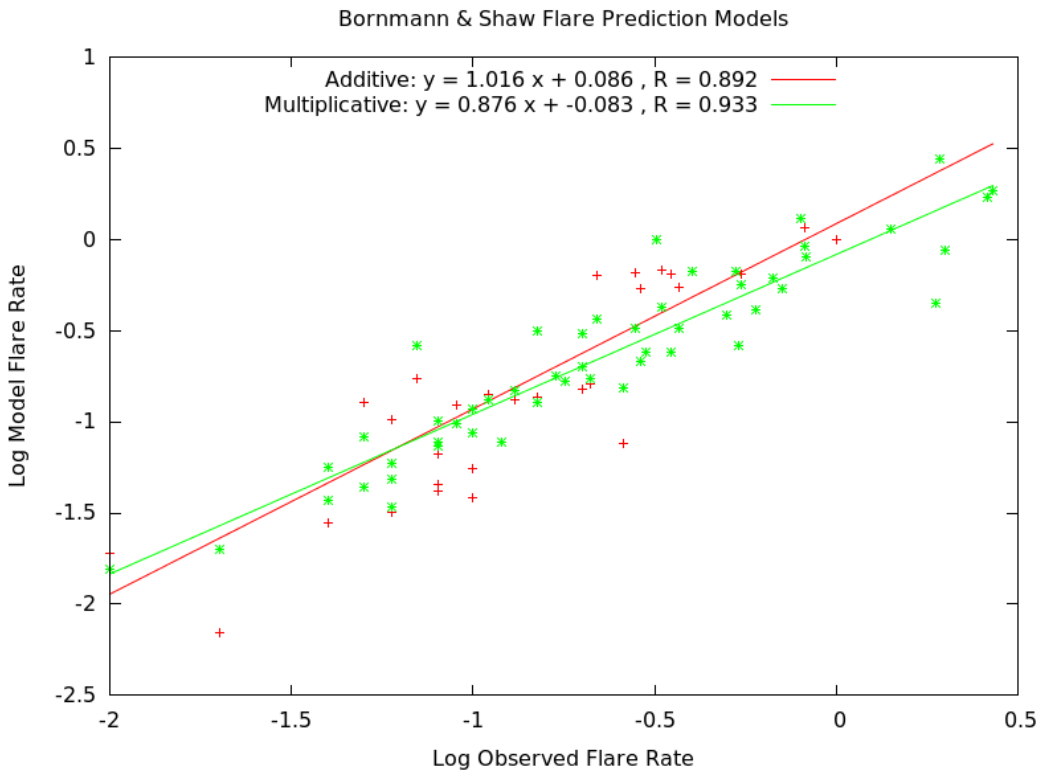



**Figure 8.9**: Bornmann and Shaw Flare Prediction Models for Cycle 20

### 8.4 Magnetic Features as a Predictor for X-ray Flares

In this section, four properties of magnetic fields are examined as flare predictors: the magnetic class, the orientation of the magnetic fields, the magnetic field strength, and the inversion line complexity. The only parameter of these that is available from SWPC is magnetic class. All of the other data was computed using the GONG magnetic field reduction techniques described in chapter 7.

#### 8.4.1 Magnetic Classification

The same approach to magnetic class flare rates was used as for sunspot class in section 8.3. Table 8.6 shows the flare rates for cycles 21 to 24 found using SWPC data. Table 8.6 shows that δ regions have higher flare rates than the non-δ regions, although the low statistics for γ and γδ do not especially assist in this analysis. The flare rates, generally speaking tend to decrease from cycle 22 to 24. Cycle 21 has some classes with higher flare rates than cycle 22 and others are lower. Of particular note is that flare rates from βγδ and γδ regions decreased from cycle 21.

| | Cycle 21 | | | Cycle 22 | | | Cycle 23 | | | Cycle 24 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Class | #Rgn | #Flrs | Rate | #Rgn | #Flrs | Rate | #Rgn | #Flrs | Rate | #Rgn | #Flrs | Rate |
| α | 2144 | 33 | 0.015 | 6659 | 68 | 0.010 | 6413 | 44 | 0.007 | 2779 | 12 | 0.004 |
| β | 4377 | 164 | 0.037 | 14373 | 818 | 0.057 | 12673 | 368 | 0.029 | 4849 | 98 | 0.020 |
| βγ | 364 | 106 | 0.291 | 866 | 271 | 0.313 | 1731 | 294 | 0.170 | 936 | 111 | 0.119 |
| γ | 6 | 2 | 0.333 | 2 | 0 | 0.000 | 4 | 2 | 0.500 | -- | -- | -- |
| βδ | 147 | 75 | 0.510 | 231 | 135 | 0.584 | 87 | 33 | 0.379 | 53 | 11 | 0.208 |
| βγδ | 214 | 267 | 1.248 | 368 | 401 | 1.090 | 603 | 426 | 0.706 | 322 | 165 | 0.512 |
| γδ | 15 | 26 | 1.733 | 5 | 3 | 0.600 | 10 | 5 | 0.500 | -- | -- | -- |

**Table 8.6**: SWPC Flare Rates by Magnetic Class. Flares are analysed by number per class (#Flrs). #Rgns is the number of regions in a particular class.

The $\chi^2$ values comparing the magnetic class between cycles were calculated and shown on table 8.7. As with table 8.5 it can be seen that the magnetic class flare rates vary significantly between the cycles. Magnetic classification does not seem to perform any better as a flare predictor than sunspot classification across cycles. The general trend however, from a qualitative perspective, is the more complex a region, the more likely it is to produce a flare. This should not be surprising as in more complex regions there is more magnetic energy available to produce flares.

| | | Expected Cycle | | | |
|---|---|---|---|---|---|
| Observed Cycle | | 21 | 22 | 23 | 24 |
| 21 | $\chi^2$ | **0.0** | **73.7** | **205.0** | **507.5** |
| | DoF | 5 | 5 | 5 | 4 |
| 22 | $\chi^2$ | **171.4** | **0.0** | **600.4** | **1684.7** |
| | DoF | 5 | 4 | 4 | 4 |
| 23 | $\chi^2$ | **294.8** | **381.0** | **0.0** | **153.5** |
| | DoF | 5 | 5 | 5 | 4 |
| 24 | $\chi^2$ | **305.5** | **348.5** | **51.4** | **0.0** |
| | DoF | 4 | 4 | 4 | 4 |

**Table 8.7**: $\chi^2$ Flare Rate by Magnetic Class Contingency Table. The critical values for 4 and 5 DoF are 9.49 and 11.07 respectively.

Table 8.8 shows the GONG magnetic class flare rates. It can be seen that in every class

except βγδ the flare rates decrease from cycle 23 to 24. δ regions still produce more flares than non-δ regions. Again the GONG data shows that as a general rule the more complex the region the more likely a flare is to occur.

| | Cycle 23 | | | Cycle24 | | |
|---|---|---|---|---|---|---|
| Class | # Regions | # Flares | Rate | # Regions | # Flares | Rate |
| α | 4864 | 68 | 0.014 | 5553 | 56 | 0.010 |
| β | 2163 | 37 | 0.017 | 2355 | 37 | 0.016 |
| βγ | 643 | 62 | 0.096 | 770 | 43 | 0.056 |
| γ | 477 | 54 | 0.113 | 492 | 43 | 0.087 |
| βδ | 153 | 30 | 0.196 | 142 | 10 | 0.070 |
| βγδ | 147 | 45 | 0.306 | 130 | 48 | 0.369 |
| γδ | 250 | 184 | 0.736 | 178 | 81 | 0.455 |

**Table 8.8**: GONG Flare Rates by Magnetic Class

The $\chi^2$ value for the GONG magnetic class comparison between cycle 23 and 24 was 109.4 with 6 degrees of freedom and a critical value of 12.59. Once again it is seen that the flare rates change between sunspot cycles based on magnetic classification. The GONG magnetic classification is no better than the SWPC classification for the purposes of flare prediction.

While the magnetic class gives a measure of the complexity of the region, the sunspot area is also important. That is to say, it would be expected that two regions of the same magnetic class but with different sunspot areas to have different flare rates. Can a simple method be used to incorporate the area of a region into the magnetic class? In order to answer this question the first step is to find the areas in each magnetic class.

Table 8.9 shows the area for each class using SWPC data. The areas for each class in different sunspot cycles vary, especially with the more complex regions. This may help explain the variation across cycles in the flare rates, given the dependence of flare rate on area shown in section 8.2.1. The standard deviations are quite large for each class showing a significant range of areas even within a class.

| | | Cycle | | | |
|---|---|---|---|---|---|
| Class | | 21 | 22 | 23 | 24 |
| α | Mean Area | 61.7 | 48.8 | 49.6 | 56.4 |
| | Std Dev | 77.3 | 70.8 | 60.6 | 63.5 |
| β | Mean Area | 108.8 | 123.0 | 114.1 | 103.9 |
| | Std Dev | 153.1 | 173.2 | 113.5 | 95.1 |
| βγ | Mean Area | 418.1 | 452.3 | 366.6 | 252.9 |
| | Std Dev | 304.7 | 342.4 | 232.3 | 179.4 |
| γ | Mean Area | 1035.0 | 415.0 | 597.5 | 0.0 |
| | Std Dev | 304.3 | 35.0 | 67.2 | 0.0 |
| βδ | Mean Area | 620.0 | 742.3 | 556.2 | 263.2 |
| | Std Dev | 456.6 | 512.8 | 352.9 | 150.3 |
| βγδ | Mean Area | 993.9 | 1103.0 | 704.8 | 615.7 |
| | Std Dev | 645.0 | 744.3 | 460.3 | 496.5 |
| γδ | Mean Area | 1518.7 | 498.0 | 501.0 | 0.0 |
| | Std Dev | 732.4 | 465.9 | 277.4 | 0.0 |

**Table 8.9**: SWPC Areas for Magnetic Classes

Next a simple technique was used to try to scale the flare rates to take into account the area. Cycle 23 was used as the 'standard cycle'. This was chosen because there was complete data for this cycle and all classes produced flares in this cycle. Equation 8.9 shows the relationship that related the number of flares across cycles to area for each class.

$$N_e = r_{23} R_c \frac{A_c}{A_{23}} \quad , \qquad \text{Eqn 8.9}$$

where $N_e$ is the number of expected flares in cycle c, $r_{23}$ is flare rate in cycle 23, $R_c$ is the number of regions in cycle $c$, $A_c$ is the average sunspot area in cycle $c$, $A_{23}$ is the average sunspot area in cycle 23.

The observed and expected flares for each cycle are shown in table 8.10.

The $\chi^2$ values between the observed and expected flares are shown in table 8.11. These values are improved over just using the magnetic class (Table 8.8) but they are still not

statistically significant compared with the critical values. Thus, there probably is some dependency of flare rates on sunspot area for magnetic class, but this simple method is only a marginal improvement over flare prediction using magnetic classes.

| | | Cycle | | | |
|---|---|---|---|---|---|
| Class | | 21 | 22 | 23 | 24 |
| α | Observed | 33 | 68 | 44 | 12 |
| | Expected | 18 | 44 | 44 | 21 |
| β | Observed | 164 | 818 | 356 | 91 |
| | Expected | 146 | 543 | 356 | 122 |
| βγ | Observed | 106 | 271 | 294 | 111 |
| | Expected | 70 | 181 | 294 | 109 |
| γ | Observed | 2 | 0 | 2 | 0 |
| | Expected | 5 | 0 | 2 | 0 |
| βδ | Observed | 75 | 135 | 33 | 11 |
| | Expected | 62 | 116 | 33 | 9 |
| βγδ | Observed | 267 | 401 | 426 | 165 |
| | Expected | 213 | 407 | 426 | 199 |
| γδ | Observed | 26 | 3 | 5 | 0 |
| | Expected | 22 | 2 | 5 | 0 |

**Table 8.10**: SWPC Magnetic Class Observed and Expected Values when Scaled for Area

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 21 | 6 | 49.98 |
| 22 | 4 | 196.62 |
| 23 | 5 | 0.00 |
| 24 | 4 | 18.68 |

**Table 8.11**: $\chi^2$ Values Between Observed Flare Rates and Flare Rates Corrected for Area. The Critical values for 4, 5 and 6 DoF are 9.5, 11.1 and 12.6.

This same process was used for GONG data. The GONG areas by sunspot class are shown in table 8.12.

Table 8.13 shows the number of observed and expected flares for cycles 23 and 24, using

cycle 23 as a reference cycle. The $\chi^2$ value for cycle 24 was 33.68 with 6 degrees of freedom and a critical value of 12.6. This is a significant improvement over using just the class ($\chi^2$ = 109.4 with 6 degrees of freedom), but still not a statistically significant result.

| | | Cycle | |
|---|---|---|---|
| Class | | 23 | 24 |
| α | Area | 75.9 | 73.9 |
| | Std Dev | 139.4 | 117.7 |
| β | Area | 186.8 | 173.8 |
| | Std Dev | 185.6 | 160.9 |
| βγ | Area | 319.2 | 282.4 |
| | Std Dev | 270.0 | 230.5 |
| γ | Area | 371.2 | 305.0 |
| | Std Dev | 323.1 | 270.2 |
| βδ | Area | 416.3 | 406.1 |
| | Std Dev | 307.8 | 296.9 |
| βγδ | Area | 646.6 | 744.4 |
| | Std Dev | 354.5 | 784.7 |
| γδ | Area | 1024.9 | 893.7 |
| | Std Dev | 678.2 | 544.5 |

**Table 8.12**: GONG Areas for Magnetic Classes

| | | Cycle | |
|---|---|---|---|
| Class | | 23 | 24 |
| α | Observed | 68 | 56 |
| | Expected | 68 | 75 |
| β | Observed | 37 | 37 |
| | Expected | 37 | 37 |
| βγ | Observed | 62 | 43 |
| | Expected | 61 | 65 |
| γ | Observed | 54 | 43 |
| | Expected | 54 | 45 |
| βδ | Observed | 30 | 10 |
| | Expected | 30 | 27 |
| βγδ | Observed | 45 | 48 |
| | Expected | 45 | 45 |
| γδ | Observed | 184 | 81 |
| | Expected | 184 | 114 |

**Table 8.13**: GONG Magnetic Class Observed and Expected Values when Scaled for Area

### 8.4.2 Magnetic Field Orientation

The region orientation, magnetic field strength and inversion line complexity are not reported by SWPC, so GONG analysis is the only data set available.

For orientation a simple four class classification scheme was developed. This is described in table 8.14.

| Orientation Class | Angle from East-West Line | Leader Spot Polarity |
|---|---|---|
| N0 | <45 Degrees | Expected |
| N1 | >45 Degrees | Expected |
| R1 | >45 Degrees | Opposite |
| R0 | <45 Degrees | Opposite |

**Table 8.14**: Orientation Class Definition. The leader spot polarity denotes whether the leader spot is reversed from its usual polarity for the cycle or if it is the expected polarity.

It would be expected the that the order of increasing flare rate would be: N0, N1, R1, R0.

As a region increases in rotation from the east-west line and dominant cycle polarity (N0) to eventually be fully reversed (R0), there will be increasing magnetic complexity and tension resulting in greater flare likelihood (Tian and Alexander, 2005).

Table 8.15 shows the flare rates by orientation class. In both cycles 23 and 24 the flare rate increases with orientation, through the escalating classes N0, N1, R1 and R0.

| | | Cycle 23 | | | Cycle24 | |
|---|---|---|---|---|---|---|
| Class | # Regions | # Flares | Rate | # Regions | # Flares | Rate |
| N0 | 5166 | 125 | 0.0242 | 5834 | 73 | 0.0125 |
| N1 | 2778 | 242 | 0.0871 | 3279 | 202 | 0.0616 |
| R1 | 534 | 67 | 0.1255 | 371 | 29 | 0.0782 |
| R0 | 220 | 46 | 0.2091 | 136 | 14 | 0.1029 |

**Table 8.15**: GONG Flare Rates by Orientation Class

The $\chi^2$ value for the orientation class comparison between cycles was 125.1 with 3 degrees of freedom and a critical value of 7.82. Therefore once again, the flare rates vary significantly between cycles for this measure.

### 8.4.3 Magnetic Field

The absolute value of the magnetic field was examined. In this case the maximum absolute value of the magnetic field of the region was used. For example, if a region had maximum field of 500 Gauss and minimum of -800 Gauss, the value of 800 would be used.

Figure 8.10 shows the relationship between magnetic field and flare rates. The lines show different slopes in cycles 23 and 24. Of note is the lower cut-off value. No flares in either cycle were observed when a region had a maximum absolute magnetic field strength of less than 200 Gauss. Any region with a field strength more than this has at least some chance of producing a flare.

In practice, this cut-off maybe higher, as GONG only looks at line of sight magnetic field. Regions near the limb probably have fields stronger than GONG reports.

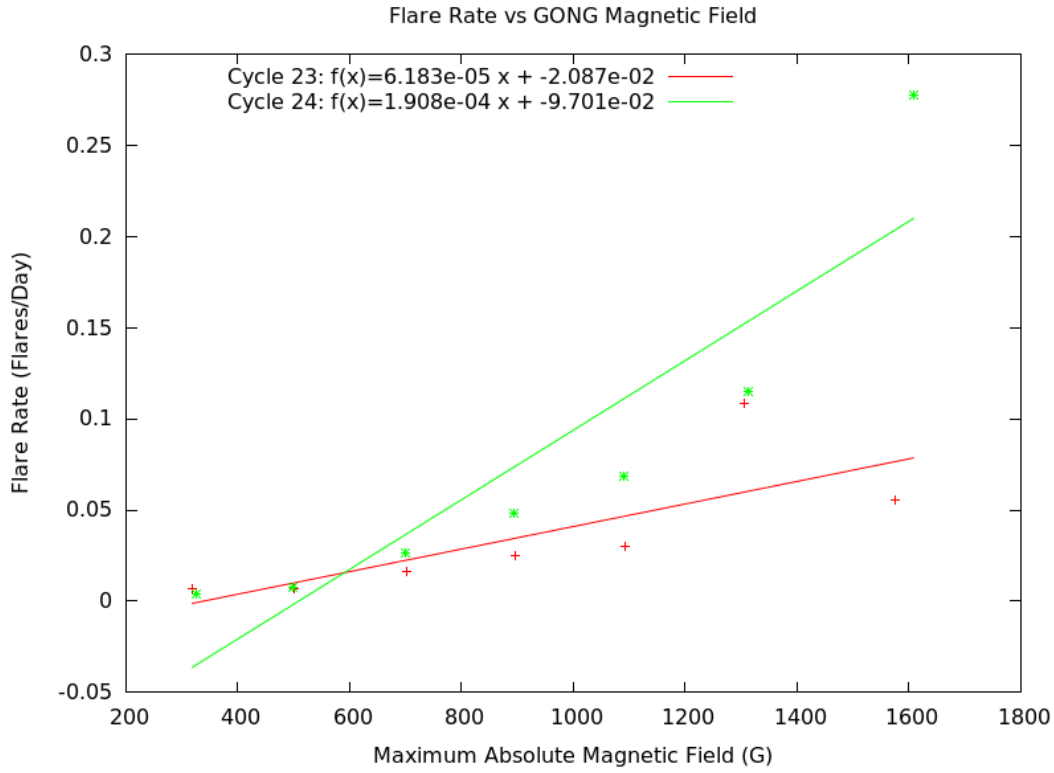



**Figure 8.10**: Flare Rate vs GONG Absolute Magnetic Field Strength

Table 8.16 shows the model fit contingency table for the magnetic field. The linear model does not work for cycle 24. Indeed, it appears that the function changes between cycle 23 and cycle 24.

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 23 | 6 | 96.6 |
| 24 | 3 | 1762.3 |

**Table 8.16**: GONG Observed Flares vs Model Flares by Absolute Magnetic Field $\chi^2$ Values. The critical value for 6 and 3 DoF are 12.6 and 7.8 respectively.

### 8.4.4 Magnetic Flux

The magnetic flux was examined as this is intuitively a better estimate than other sunspot parameters of the total energy available to produce flares. The flux computed here is simply the absolute maximum magnetic field of a region multiplied by the sunspot area. Therefore, this measure is an over-estimate of the true flux value.

Figure 8.11 shows magnetic flux vs flare rate. A log-log scale is used so the curve fits do not appear to be as close as shown in table 8.17. The best fit for these curves was a power law fit, with the exponent for both cycles being similar but the coefficient differing significantly. Again there was a distinct cut-off, although this was different in each cycle. For cycle 23 this was $10^4$ G.μh, and for cycle 24 it was $2\times10^4$ G.μh.

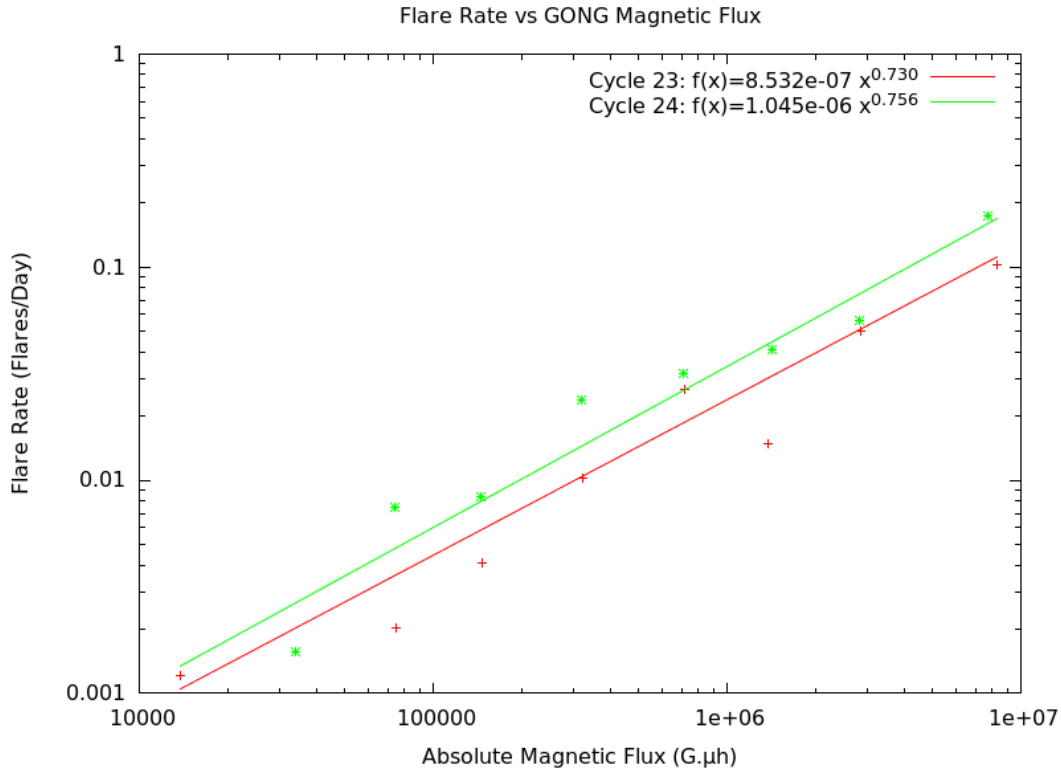



**Figure 8.11**: Flare Rate vs GONG Magnetic Flux

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 23 | 7 | 2490.9 |
| 24 | 6 | 1114.7 |

**Table 8.17**: GONG Observed Flares – Model Flares by Magnetic Flux $\chi^2$ Values. The critical value for 6 and 7 DoF are 12.6 and 14.1 respectively.

### 8.4.5 Inversion Line Complexity

Finally, dependence of flare rate on inversion line complexity was examined. Complexity is the inversion line length divided by the north-south extent, with values typically in the

range 1 to 5. The top two bins in cycle 23 and top bin in cycle 24 had to be rejected due to them being outliers, with significantly low flare rates. The relationships are shown in figure 8.12.

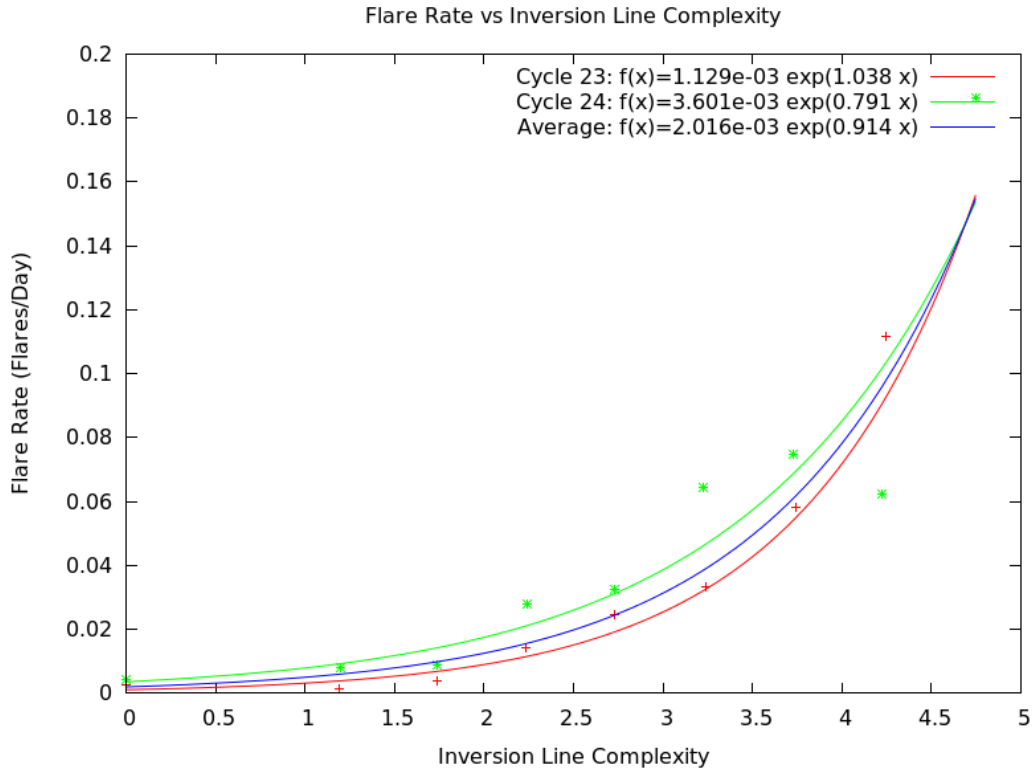



**Figure 8.12**: Flare Rate vs GONG Inversion Line Complexity

For inversion line complexity, it was found that an exponential fit produced the best fit as shown in Table 8.18. This is the first parameter for which a model in one cycle that reasonably fits the data (Cycle 24) as $\chi^2$ of 9.4 is less than the critical value of 14.1. The blue line in figure 8.12 was a fit with cycle 23 and 24 data combined to produce an average curve. Table 8.19 shows the $\chi^2$ values between the average model and data. This blue curve produced a better fit for cycle 23 with a $\chi^2$ of 9.2 less than a critical value of 11.1, but a worse fit for cycle 24.

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 23 | 5 | 16.3 |
| 24 | 7 | 9.4 |

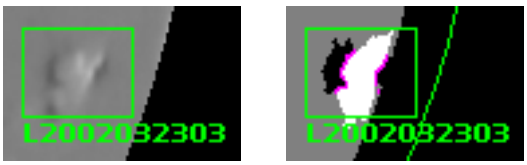
**Table 8.18**: GONG Flare Rates by Inversion Line Complexity Model vs Measured $\chi^2$ Values. The critical values for 5 and 7 DoF are 11.1 and 14.1.

| Cycle | DoF | $\chi^2$ |
|---|---|---|
| 23 | 5 | 9.2 |
| 24 | 7 | 31.4 |

**Table 8.19**: GONG Flare Rates by Average Inversion Line Complexity Model vs Measured $\chi^2$ Values. The critical values for 6 and 8 DoF are 11.1 and 14.1.

The outliers rejected in figure 8.12 and subsequent analysis all had inversion line complexities greater than 5. A brief investigation as to why this may be the case was undertaken. Many of the regions with high complexities had correct values. However, the following two examples illustrate where the automated algorithm produced erroneous complexities.

Figure 8.13 shows a region near the limb, which produced two inversion lines. The algorithm correctly identified the lines, but this is probably a case where the magnetic maps near the limb of the sun become unreliable, due to most of the field being tangential to the line of sight.

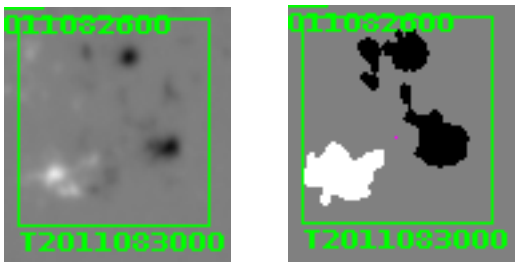



(a) (b)

**Figure 8.13**: Magnetic Region Near the Limb Producing Erroneous Inversion Line Complexity. (a) is the magnetogram and (b) is the magnetic area image. Image taken from GONG Learmonth on 2002/03/23 01:00:16 UT and was assigned an inversion line complexity of 7.3.

Figure 8.14 shows a magnetic region near the centre of the solar disk. In this case the

inversion line is so small and poorly defined that the algorithm calculated a large inversion line complexity. This kind of error could be overcome by requiring a minimum inversion line length for it to be included in the statistics.




(a) (b)

**Figure 8.14**: Magnetic Region with Poorly Defined Inversion Line Producing Erroneous Inversion Line Complexity. (a) is the magnetogram and (b) is the magnetic area image. Image taken from GONG Teide on 2011/09/01 08:00:16 UT and was assigned an inversion line complexity of 7.3.

### 8.5 Summary and Conclusions

A variety of solar parameters were examined in order to determine which were good predictors of solar X-ray flares. The flare rates for regions with specific features were compared. It was found that the flare rate for M or X class flares changed significantly between sunspot cycles such that, for a given parameter of a sunspot region, the number of flares that regions with the same parameter value produced, was different between cycles.

It was found that sunspot area was a better predictor for flares than sunspot area growth. The sunspot class can be used within a single sunspot cycle, but flare rates need to be recalculated for each sunspot cycle as they too vary significantly. Magnetic class was no better a predictor than area, area growth or sunspot class. While a general statement that a δ region produces more flares than non-δ regions may be made, the flare rates again vary significantly between cycles.

Applying a correction factor to each magnetic class to adjust for variations in area slightly improved the consistency between cycles, but not in a significant manner. Thus area does seem to play a part but not to a significant extent.

Regions in cycles 23 and 24 with an absolute magnetic field of less than 200 Gauss

produced no flares. A general trend of larger field strengths producing higher flare rates was found. Magnetic flux also had a flare rate cut-off below $10^4$ G.μh.

The inversion line complexity was the only parameter where a simple function was able to be fitted to the flare rates and produce a statistically significant fit, although this was only in one cycle. However outliers in each cycle had to be rejected in order to achieve this result. These results may be able to be improved with more rigorous criteria in examining the inversion line.

From these results, it certainly seems that inversion line complexity is the most reliable parameter for predicting flares. In addition, a significant number of smaller regions can automatically assigned a zero flare probability just based on the magnetic field strength being below the measured threshold of 200 Gauss, if future cycles behave similarly to cycles 23 and 24.

In practice, flare rates for any parameter will need to be recalculated for each sunspot cycle. The flare rates for SWPC sunspot classification can be found in appendix F and those for GONG sunspot classification can be found in appendix G. These are displayed by sunspot cycle.

# 9. Prediction of Solar Particle Events

## 9.1 Introduction

Solar Particle Events, SPEs (Solar Energetic Particles, SEPs), are relativistic particles, mainly protons, that are accelerated either from the Sun during a flare or in the interplanetary medium from shocks (see section 9.2).

There is a constant stream of low energy particles at low flux being emitted by the Sun. SWPC define an SPE as a flux greater than 10 particle flux units (pfu) of energy greater than 10 MeV, where 1 pfu = 1 particle $cm^{-2}s^{-1}sr^{-1}$ (http://www.swpc.noaa.gov/phenomena/solar-radiation-storm). Figure 9.1 shows an example solar particle event measured at the GOES-13 spacecraft in geostationary orbit. The red line shows the flux for >10 MeV particles exceeding the 10 pfu level, at the dotted line, for at least 3 days.

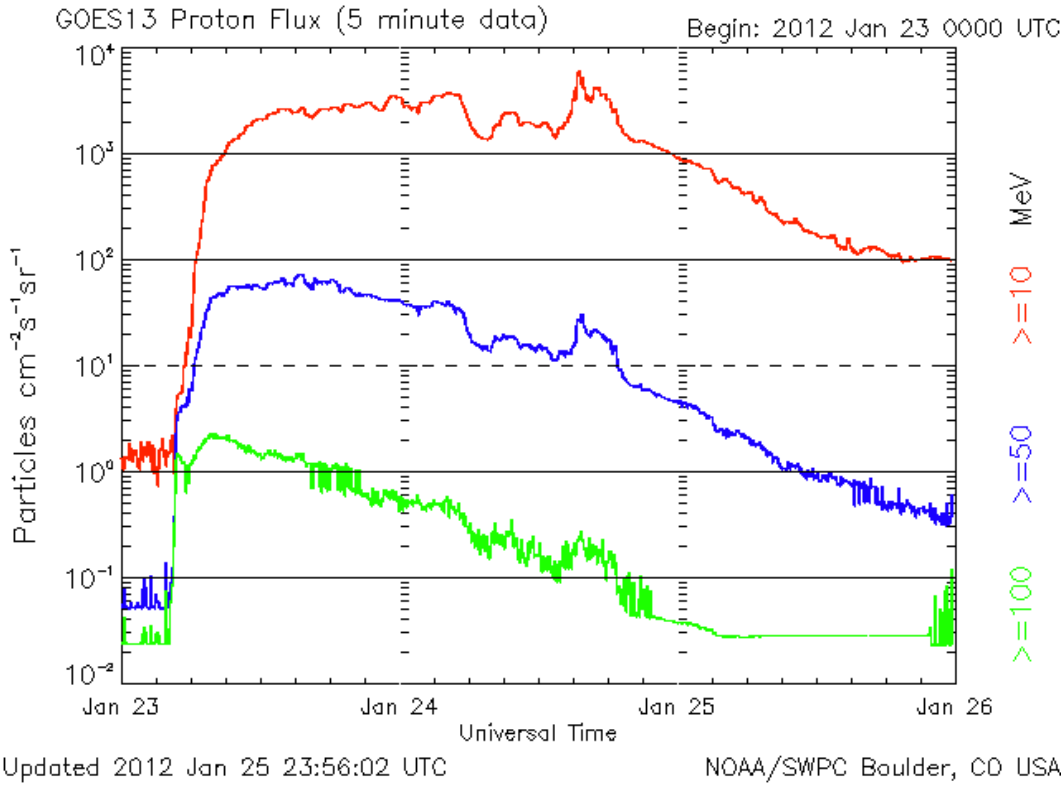



**Figure 9.1** Solar Particle Event beginning January 23 2012 (image obtained from http://www.swpc.noaa.gov/products/goes-proton-flux). The red, blue and green lines show the flux for particle energies greater than 10 MeV, 50 MeV and 100 MeV respectively.

Rarely do SPEs reach the surface of the Earth. Most are stopped by the Earth's

atmosphere. When an SPE is detected at the Earths surface it is often referred to as a Ground Level Enhancement (GLE).

There are some difficulties in detecting GLEs at the surface of the Earth. Neutron monitors are used to detect secondary particles. Detection depends upon the cut-off rigidity for a given site (Nitta et al., 2012). The atmosphere alone has a cut-off rigidity of about 450 MV and is higher near the equator due to the effect of the Earth's magnetic field (Shea and Smart, 2012). Therefore a GLE that is detected at the poles may not be detected at lower latitudes.

The importance of SPEs from a space weather perspective is for the airline and space industry. Airline crews suffer a higher radiation dose as aircraft fly above most of the atmosphere, and for economic reasons, airline operators are wanting to fly to higher altitudes. The space industry is concerned with protecting satellites. Energetic particles can damage solar cells, flip bits in the satellite instrumentation and are a hazard to astronauts (Jiggens et al. 2014). With increasing interest in long duration manned Mars missions, the space radiation environment is becoming even more important.

Figure 9.2 shows the comparison of SPEs with sunspot number. It is seen that around solar maximum there are significantly more SPEs than at solar minimum. That is not to say that these events cannot occur at solar minimum; just that they are less likely. It is also seen that there appears to be little correlation between the sunspot number of a cycle and the flux of the SPEs in a sunspot cycle.

Of the four sunspot minima seen on figure 9.2, all show SPEs occuring, with an SPE almost simultaneous with the 1976 minimum.

Therefore, figure 9.2 gives some indication of how difficult devising a scheme of SPE prediction based on solar features may be to achieve.

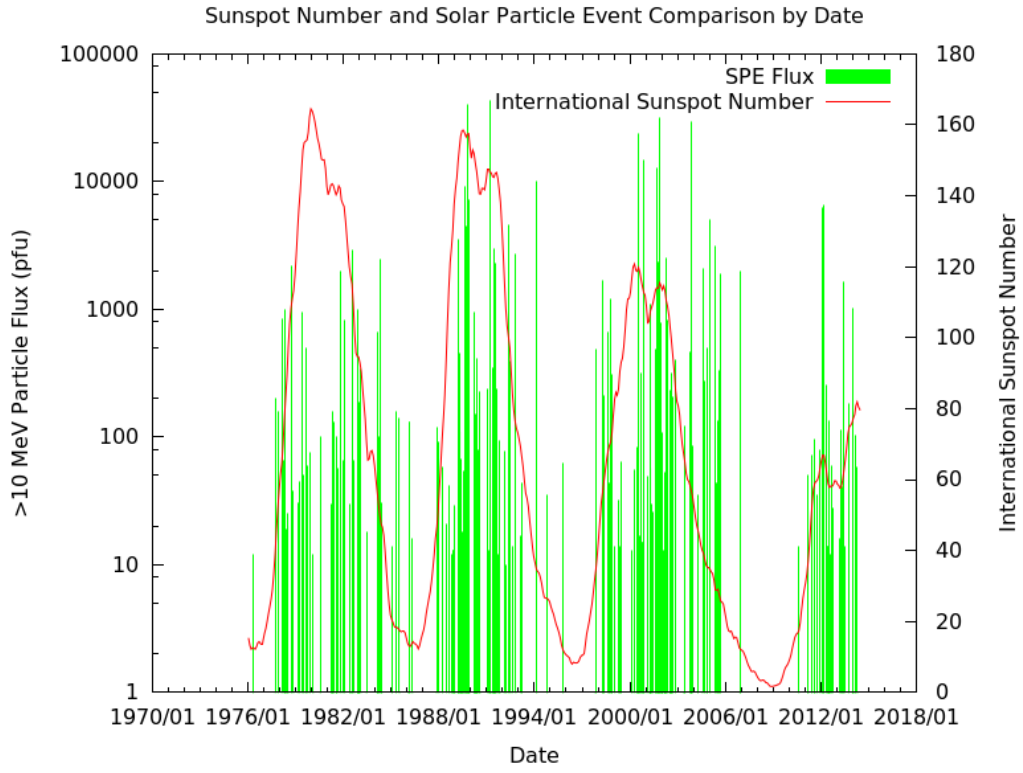



**Figure 9.2**: Sunspot Number and Solar Particle Event Comparison by Date Over 4 Solar Cycles

There are three mechanisms for accelerating particles (discussed in section 9.2): direct acceleration associated with flares (usually referred to as flare accelerated), second order Fermi acceleration and first order Fermi Acceleration.

## 9.2 Particle Acceleration Mechanisms

### 9.2.1 Direct or Flare Associated Acceleration

One of the most commonly used flare acceleration models is the CSHKP model. A good description of this model is found in Masson et al. (2013). In this model a magnetic flux rope can lift off the Sun at the time of a flare. At the magnetic reconnection site, particles are able to be accelerated as they are trapped in the field lines wrapped around the flux rope. The difficulty with this model is how particles then escape into the interplanetary medium. Figure 9.3 shows this model and is found in Masson et. al. (2013)

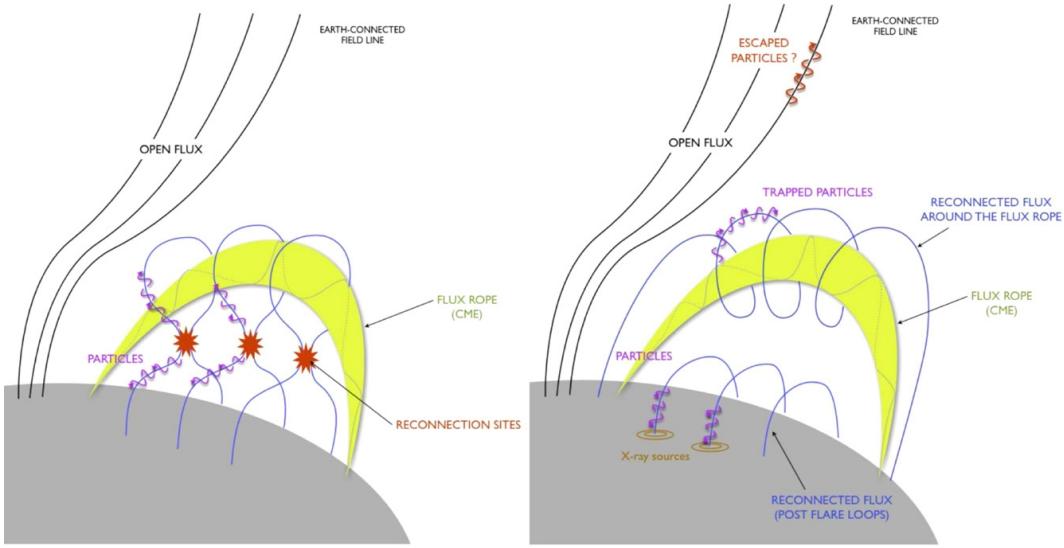



**Figure 9.3**: Standard Model for Eruptive Flares and its Implications on Particle Trapping. "The left panel shows the rising CME plasmoid represented by the green arch, along with the flare reconnection below the arch. The right panel shows that all the field lines resulting from the reconnection remain closed and, hence, do not allow any flare-accelerated particles to escape." (Masson et al., 2013).

### 9.2.2 Second Order Fermi Acceleration

This method was first proposed to explain the energies of galactic cosmic rays by Fermi (1949). Here a particle gets accelerated and reflected by a travelling magnetic field. If the collision is head-on with the magnetic field, then the particle will gain energy, if it is tail-on then the particle will lose energy. As tail-on collisions are less likely than head-on ones, on average the particles gain energy. A representation of this process is shown in figure 9.4.

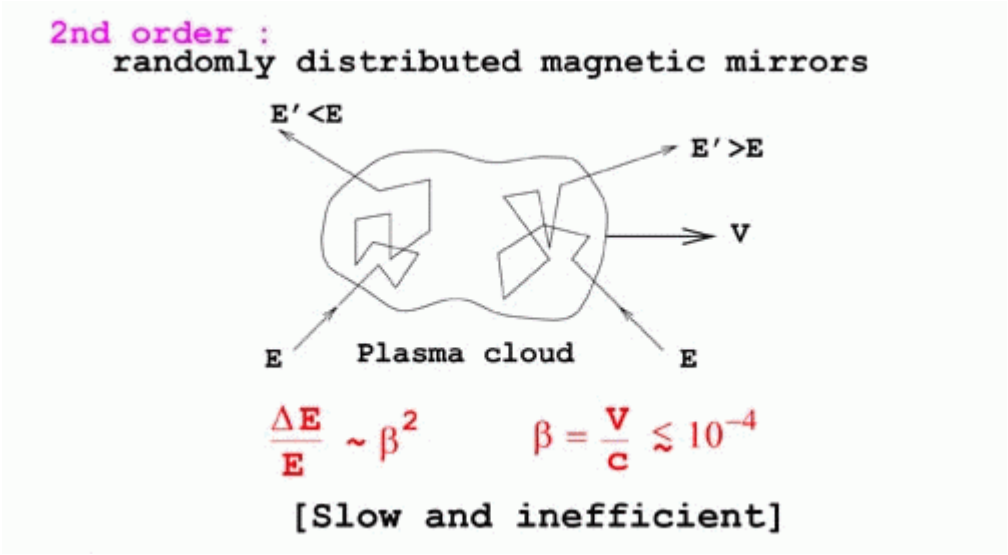



**Figure 9.4**: Diagram of Second Order Fermi Acceleration (obtained from http://www.cosmic-ray.org/reading/uhecr.html). The plasma cloud is moving to the right. Particles entering the cloud from the left will lose energy, and those entering from the right will gain energy.

### 9.2.3 First Order Fermi Acceleration

In first order acceleration, the particles gain energy due to inhomogeneities in a shock. That is, there may be different magnetic fields preceding and following the shock. A particle can be accelerated and reflected by these inhomogeneities (Li et al., 2003). Figure 9.5 shows particles crossing a shock front.

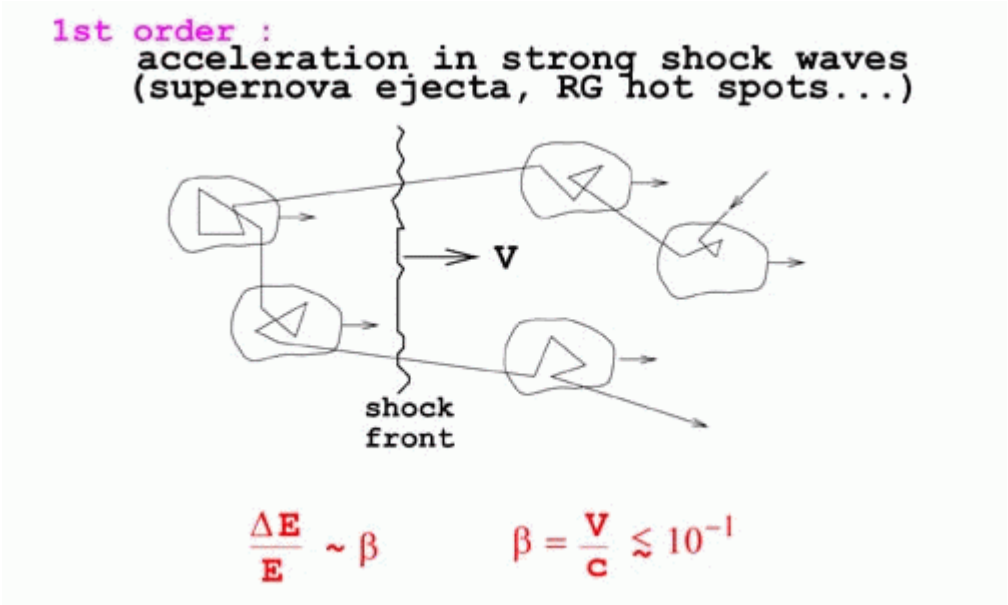



**Figure 9.5**: Diagram of First Order Fermi Acceleration (obtained from http://www.cosmic-ray.org/reading/uhecr.html). The plasma cloud is moving to the right. Inhomogeneities in the plasma cloud cause particle to be accelerated when they cross a shock boundary.

### 9.3 Radio Burst Features Associated with SPEs

Radio signatures have long been thought to be a good indicator of SPEs. In the 1960s and 1970s a type of radio signature called the Castelli-U was thought to be a good predictor of SPEs. Then, Type-2 and Type 4 radio bursts were considered an improvement over the Castelli-U for SPE prediction. Finally, type 3L emission (a subset of type 3 emission) was recognised as improved predictor of SPEs. These will each be discussed in turn.

#### 9.3.1 Castelli-U

In the 1960s, the Solar Particle Alert Network (SPAN) was set up by NASA. This was to support the Apollo moon missions. In particular, it was an attempt to predict the onset of solar particle events that may have negative impacts on the missions. It consisted of Hα telescopes, radio telescopes in the range of 1415 MHz to 4995 MHz, with some sites having 245 MHz to 610 Mhz, 15.4 GHz and 30 GHz. In addition, it had neutron monitors. For full details see Robbins and Reid (1969).

Castelli and Guidice did a significant amount of work using radio data as a precursor to SPEs. They found that a spectral profile matching all of the following conditions was a

good predictor of SPEs:

- A burst of 1000 SFU or greater in the 245 MHz to 610 MHz range.
- A burst of 1000 SFU or greater in the 4995 MHz to 15.4 GHz range.
- An integrated flux from the start of the burst to the end of the burst on 8800 MHz of 100 kSFU×s,
- A pronounced dip in the flux in the 1415 MHz to 2695 MHz range.

All of the above criteria must be met by a multi-spectral burst for it to be considered a Castelli-U event (Castelli and Guidace, 1976). Furthermore, RSTN added an additional requirement that all of the peaks that make up a Castelli-U event occur within 2 minutes of each other (http://static.e-publishing.af.mil/production/1/afwa/publication/afwaman15-2/afwaman15-2.pdf).

Figure 9.6 shows an example Castelli-U event that was observed on 29 April 1973 at Sagamore Hill. Of note is the similarity of the general shape to the average burst profile shown in chapter 3 and reproduced in figure 9.7; that is, the average burst profile actually meets the requirements for a Castelli-U.

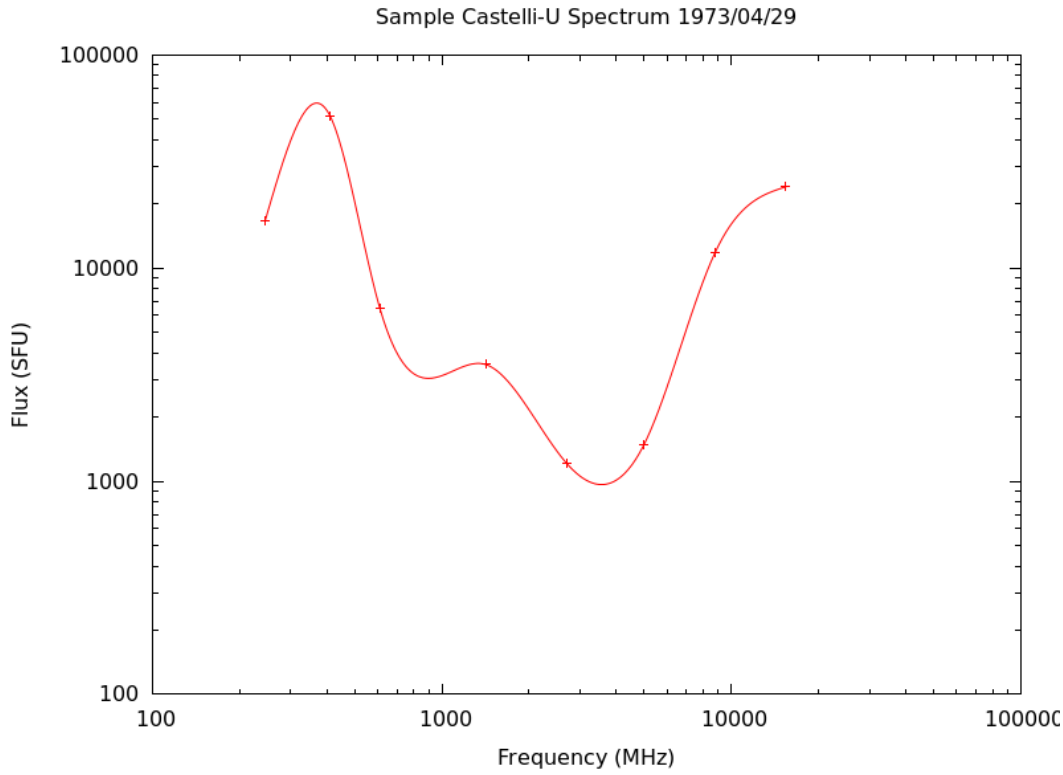


**Figure 9.6**: Castelli-U Event Observed on 29 April 1973 at Sagamore Hill

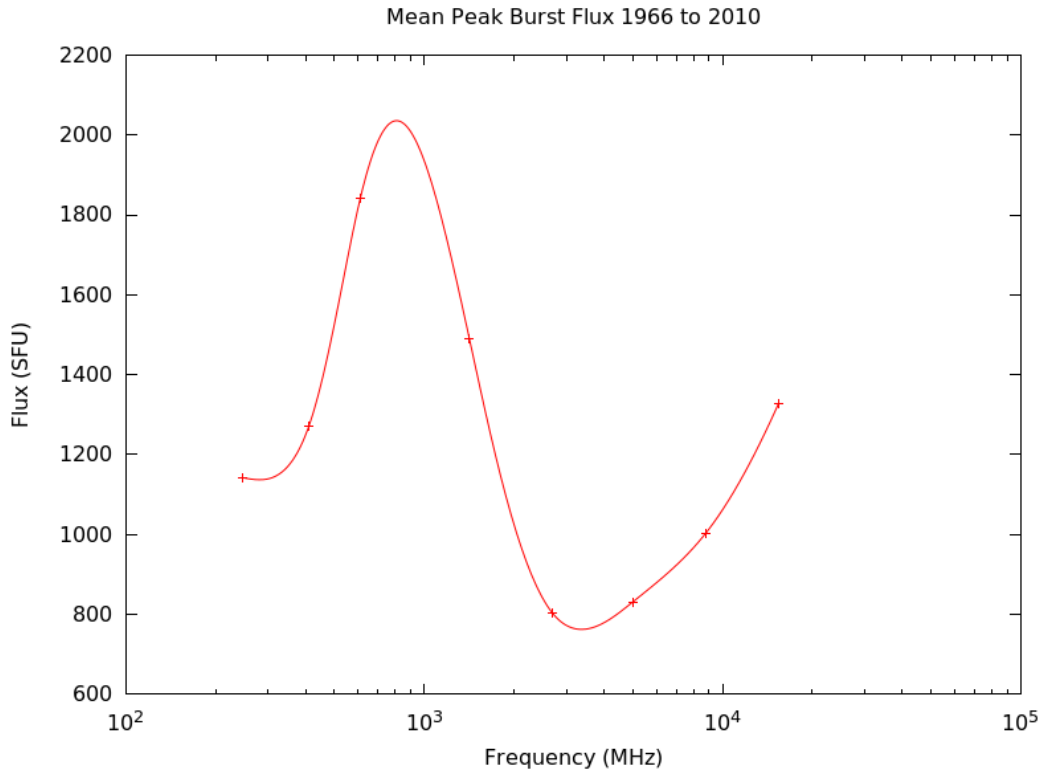



**Figure 9.7**: Average Peak Burst Flux

If the average burst spectrum profile is similar to that of a Castelli-U event, how good a predictor is the Castelli-U for SPEs? An estimate of the reliability for prediction will be obtained in this work. Unfortunately, the NGDC does not contain a list of the occurrences of Castelli-U spectral bursts. But it does contain some of the information needed in individual burst messages. Unfortunately none of the integrated fluxes are listed, so these are estimated using equation 9.1:

$$I_S = \frac{1}{3} S(t_e - t_s) \quad , \qquad \text{Eqn 9.1}$$

where $I_S$ is the estimated integrated flux, $S$ is the peak flux, $t_e$ is the end time for the burst and $t_s$ the start time for the burst. Both $t_e$ and $t_s$ are in seconds.

Unfortunately only the primary peak is reported. The Castelli-U event can use a secondary peak if it falls within the time criteria specified by the USAF-RSTN. For this reason the time criteria has been ignored in this analysis, thus potentially allowing more

bursts to be classified as a Castelli-U than the current definition allows.
The analysis of SPE to Castelli-U event correlations only begins from 1980, as this was when three of the four RSTN sites were operational. Prior to 1979, only Sagamore Hill was monitoring all 8 frequencies; thus, a significant number of bursts may have been missed.

The bursts from 1980 to 2011 were grouped together if they were in progress at the same time. Each group was then examined to see if it met the Castelli-U burst criteria. Once this was done, the Castelli-U bursts were compared against the Solar Particle Event list to establish if an event occurred with 24 hours of a Castelli-U. Table 9.1 summarises these results.

| | |
|---|---|
| Number of Castelli-U Events | 182 |
| Number of SPEs | 206 |
| Number of Castelli-U Events followed by an SPE | 38 |

**Table 9.1**: Summary of Castelli-U Events as a Predictor for SPEs

Castelli claimed a near 100% success rate with few false alarms using this criteria (Castelli and Guidice, 1976). In fact it was considered so reliable that prediction schemes were developed from the Castelli-U that predicted onset time and the energy spectrum of the associated proton event (Smart and Shea, 1979).

However, as shown, the Castelli-U is a poor indicator of when an SPE will occur. 20.9% of Castelli-U bursts were followed by an SPE, and 18.4% of the time an SPE was preceded by a Castelli-U.

In the mid 1980s it was recognised that the Castelli-U event was probably not a good indicator for SPEs. Cliver et al. (1985) claimed that a coincident type 2 and type 4 emission was a better predictor and that the correlation of Castelli-U bursts and SPEs were simply due to them both occuring during a large flare. The poor Castelli-U – SPE correlation has since continued.

The RSTN sites still report a Castelli-U event when it occurs, and some space weather agencies still use it in predictive tools.

### 9.3.2 Type 2/Type 4 Emission

This section is derived from a paper presented at the Australian Institute of Physics 18th National Congress (Giersch and Kennewell, 2008). Type 2 emissions were described in chapter 3.

Type 2 emission results from a shock moving through the solar corona (Smith, 1967). The frequency drift allows a calculation of the shock speed if an electron density model is assumed for the corona. If the type 2 extends to low frequencies (below 10 MHz) it is assumed to be a shock preceding a coronal mass ejection (CME), whereas a low frequency cut-off above 20 MHz may imply a coronal shock produced by a flare blast wave (Leblanc et al., 2001). The former class is often referred to as an interplanetary (IP) type 2, and is usually not seen on ground-based equipment. Figure 9.8 is an example of a type 2 seen on a metric-wave terrestrial spectrograph.

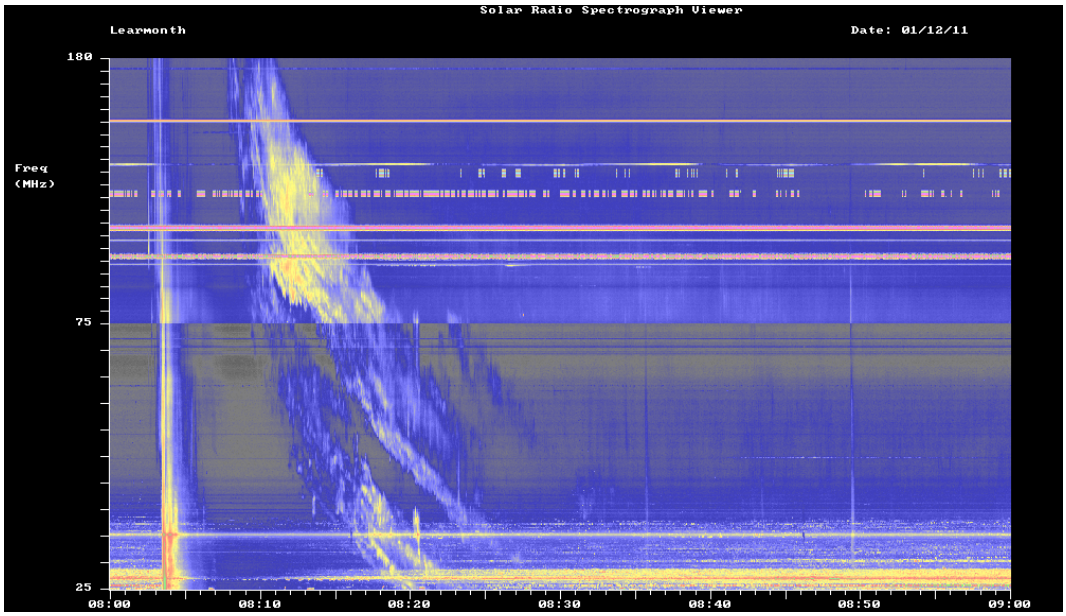



**Figure 9.8**: Example of Type 2 Solar Radio Emission. The type two begins at 08:08 UT. It is preceded by a conventional type 3 emission and followed by a weak type 4 emission beginning at about 08:29 UT.

Most type 2 emissions are not followed by an SPE. It is the type 2/type 4 complex which is considered a more significant SPE indicator. Type 4 emission is synchrotron radiation, but the source is variable, and some classification schemes (Kruger, 1979) show many different type 4 subclasses. The so-called ‘moving type 4s’ are probably associated with

CME's and thus provide a better indicator of SPE acceleration. Unfortunately, a radioheliograph is needed to distinguish the type 4 source region, and these are few and far between. Cliver et al. (1985) used statistics from sunspot cycle 20 to show that metre - wavelength type 2/type 4 emission was a better predictor of SPEs than the Castelli-U signature.

An analysis by Giersch and Kennewell (2008) of 8 years of solar radio spectrograph data for type 2/type 4 emission and possible SPE correlation these events to SPEs, found that 36% of type 2/4 events were followed by an SPE, and 24% of SPE's were preceded by a Type 2/4 event.

These results show that the Type 2/4 emission complex is a better predictor for SPEs than the Castelli-U event, but really not acceptable for space weather forecasts.

### 9.3.3 Type 3Lm Emission

As with section 9.3.2, this section is derived from a paper presented at the Australian Institute of Physics 18th National Congress (Giersch and Kennewell, 2008).

Attempts to relate type 3 (fast drift) emission, produced by relativistic electron streams, to solar energetic particles date back to the early 1970's. Graedel and Lanzerotti (1970) analysed data between July 28 to August 4 1967. They found that 61% of 222 proton events in the 1 MeV range had type 3 association. This is a very low energy threshold, and an SPE is now defined by a threshold flux above 10 MeV.

Measurement of low frequencies (<18 MHz) solar radio emission is difficult due to the terrestrial ionosphere. Not only is the ionosphere opaque to solar radio waves at these low frequencies, but it also reflects man-made terrestrial signals creating a challenging Radio Frequency Interference (RFI) environment for solar radio astronomers working at these frequencies. Although space based instrumentation can access lower frequencies, some attempts have been made to observe the Sun at these low frequencies with ground-based equipment. This is particularly important as spacecraft may fail, and timely replacements are not always possible.

Type 3L is a subset of type 3 emission that (1) lasts a longer time, (2) extends to lower frequency and (3) starts later relative to the start of the flare. Type $3L_m$ is a further subset

of type 3L emission at metric wavelengths that could be observed on ground based spectrographs. This distinguishes it from other type 3L which extend to much lower wavelengths.

For example the type 3 as seen in figure 9.8 lasts about one minute, whereas the type $3L_m$ emission in figure 9.9 lasts 20 minutes at 25 MHz. These are seen clearly in space-based WIND spacecraft data (Cane et al., 2002). Of concern is how many type 3Ls are observed on ground based solar radio spectrographs. The Bruny Island Radio Spectrograph (BIRS) measured frequencies down to 4 MHz and had software which looked for gaps in the high frequency (HF) radio spectrum, which are free from RFI (Erickson, 1997). However most radio spectrographs do not go as low as 4 MHz due to the difficult observing environment. The Culgoora radiospectrograph extends down to 18 MHz and the Learmonth instrument to 25 MHz. Space-based 3L detection shows a high association with SPE's (Cane et al., 2002, MacDowall et al., 2003). This investigation considers metre-wave type 3L – SPE associations.

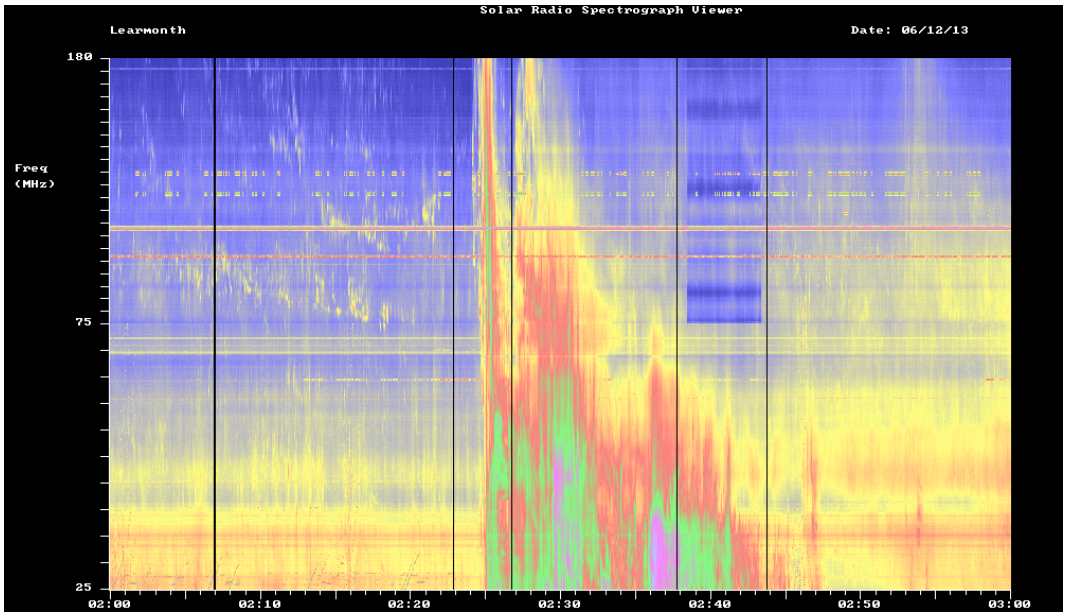



**Figure 9.9**: Strong type $3L_m$ burst observed by the Learmonth Solar Radio Spectrograph on 13 December 2006

In examining spectrograph data from 2000 to 2008 from RSTN sites, it was found that up to 37% of Type $3L_m$s were were followed by SPE's, but up to 62% of SPE's were preceded by Type $3L_m$s (Giersch and Kennewell, 2008). There is uncertainty, as in many

cases it was difficult to determine if an event was a type $3L_m$ or not. In fact Cane et. al. (2002) stated that often analysts misidentified type 3L events, considering them to be type 2 instead. Type $3L_m$ events usually occur in conjunction with type 2 or type 4 events, and may even dominate, washing out the lower intensity emission. Figure 9.8 is an example of where the emission was misidentified as type 2. The results are summarised in table 9.2.

| Event | # Samples | SPE Prediction Rate | False Alarm Rate |
|---|---|---|---|
| Type 2/4 | 116 | 36% | 64% |
| Type $3L_m$ (C) | 14 | 16% | 84% |
| Type $3L_m$ (C+U) | 32 | 37% | 63% |
| Type 3L | 6[1] | >80% | <20% |

**Table 9.2**: Association of SPEs with Type 2/4 complexes and Type 3Ls

Explanatory Notes: 3Lm events are those detected only at metre wavelengths, 3L events are those detected down to hectometric wavelengths, (C) refers to certain events, (U+C) to combined certain and uncertain events. 1 refers to the Cane et al. study (2002).

These results show that ground based detection of type 3Lm emission is still a poor predictor of SPEs, due to a low SPE prediction rate compared with a high false alarm rate.

### 9.4 Sunspot Features Associated with Solar Particle Events

The aim here is to determine what, if any, features of sunspot regions maybe an indication of solar particle events. This is somewhat problematic. Flares, for example, can usually be associated reasonably reliably with an active region. This cannot be said for an SPE, because in many cases the area of acceleration may be at a shock front many solar radii from the photosphere of the Sun.

Attempts have been made though to associate SPEs with active regions. Nitta et al. (2012) examined GLEs that occurred during sunspot cycle 23. They could not find a correlation between active region features or active region magnetic fields and GLEs. The question addressed in this section is “Does the automated feature analysis yield any better results for this correlation than did the manual region analysis?”

An analysis was performed of regions that were on the disk 24 hours prior to the

occurrence of an SPE. The region parameters used were the results of the GONG automated analysis discussed in chapter 7. For each day prior to an SPE, the most complex region was found. Of this list, the minimum for each parameter was found in order to determine the potential minimum requirements for an SPE. Regions with magnetic class of α were omitted. These results are summarised in table 9.3.

| | |
|---|---|
| Field (G) | 198 |
| Flux (G.μh) | 3.58E+03 |
| Inversion Line Complexity | 0.783 |
| Inversion Line Length (degrees) | 0 |
| Gradient (G/Mm) | 5.7 |
| Orientation (degrees) | 89 |

**Table 9.3**: Minimum Parameters Associated with SPEs

Unfortunately, nearly all β or greater complexity regions meet the above criteria. Due to the low number of SPEs that occur, it makes rate analysis like that applied to the analysis of flares (chapter 8) difficult.

To illustrate, a histogram of inversion line complexity for the most complex region on the Sun prior to an SPE was generated. This is shown in figure 9.10. For Inversion Line Complexities greater than 2.0, this is a nearly flat distribution. That is, there is no dependence on SPE occurrence when the inversion line complexity exceeds 2.0.

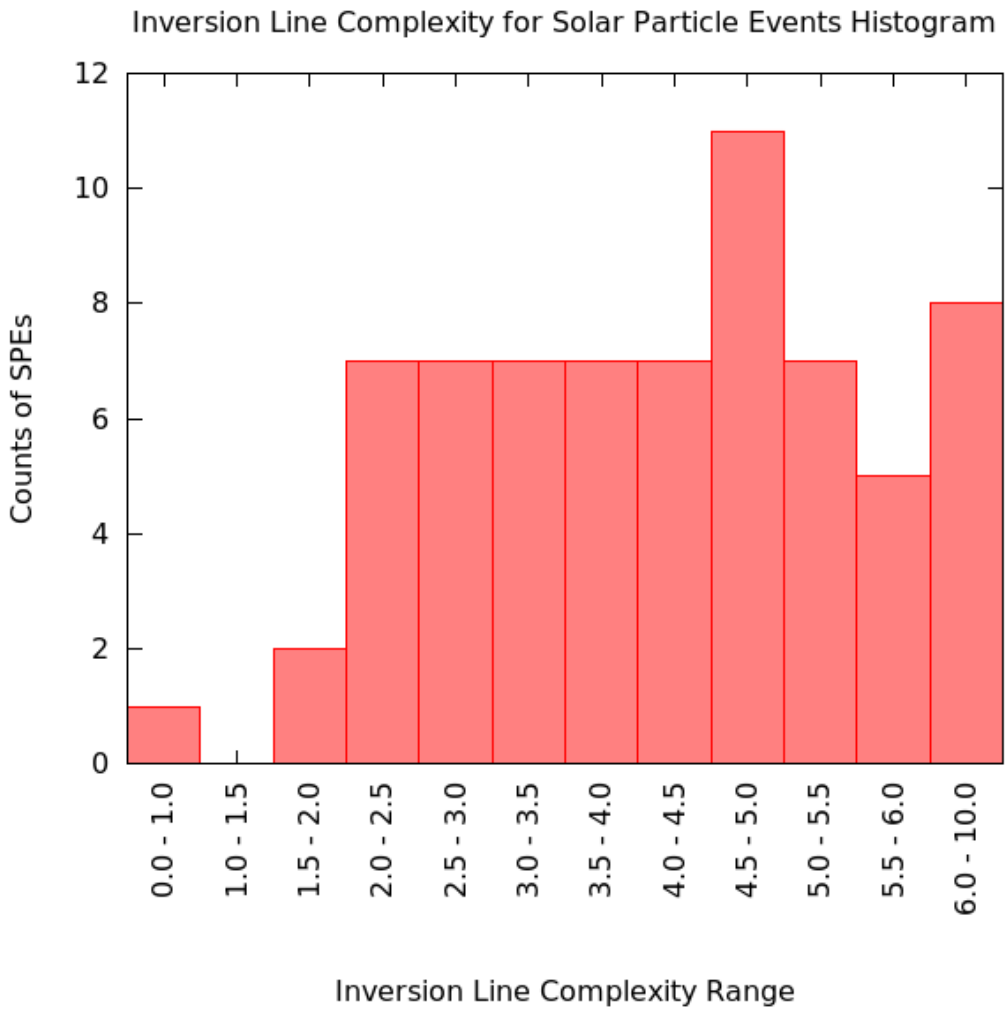



**Figure 9.10**: Histogram of Inversion Line Complexities for Regions that were Associated with SPEs

Possibly of more significance is the percentage of regions in each Inversion Line Complexity category that appears associated with an SPE. That is, a histogram of the most complex regions on each day was created. Each category in figure 9.10 was divided by the total number of regions in that category, producing figure 9.11.

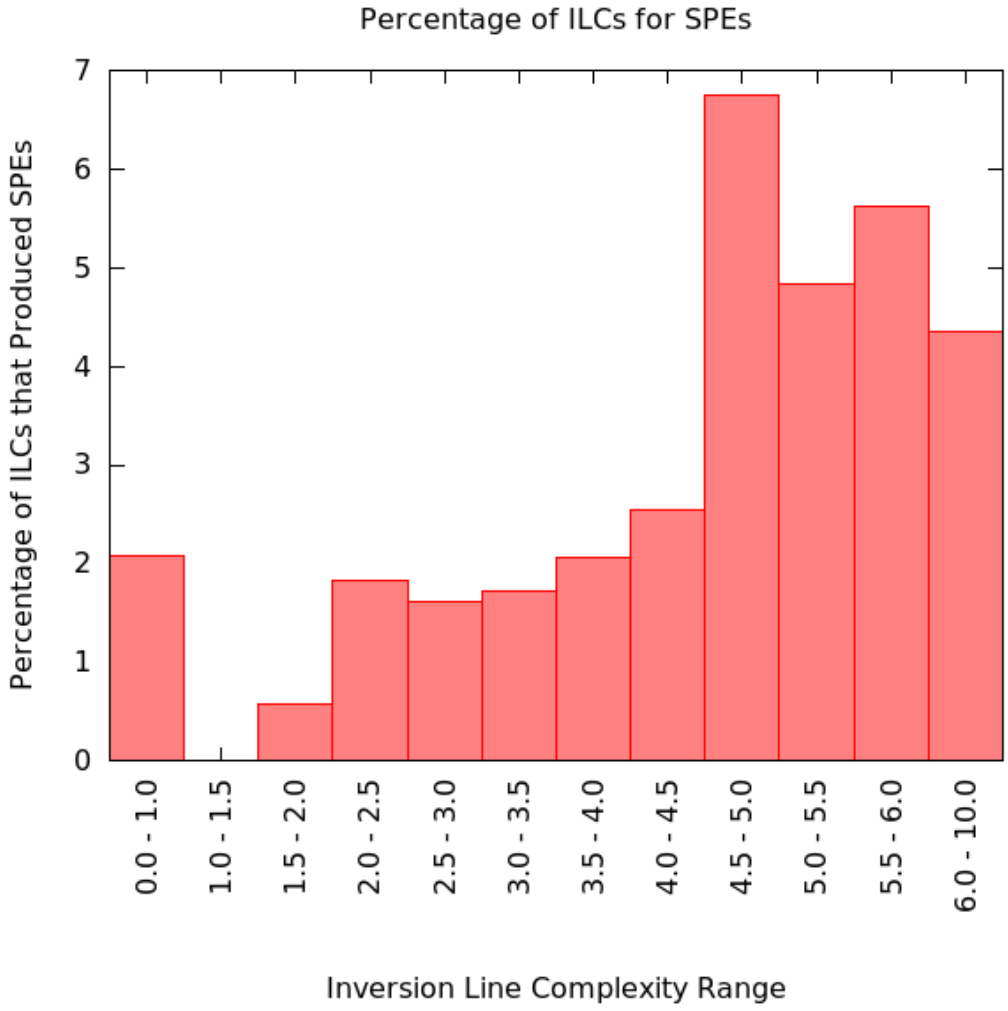



**Figure 9.11**: Percentage of Inversion Line Complexities in Each Category that were Associated with SPEs

Figure 9.11 shows that at best a region with an Inversion Line Complexity of 4.5 to 5.0 has a 6.8% chance of producing an SPE. Alternatively, there is nearly a 93% false alarm rate for this category. The other categories are worse.

Figures 9.10 and 9.11 incorporate all events from 2002 to 2014 that had a particle flux greater than 10 MeV. It is sensible to conder the most energetic events, namely those that make it to the ground as GLEs. NGDC contains a list of GLEs from 1942 to 2012. The events examined in this work are listed in table 9.4, along with the SPE counterpart.

| GLE Event # | Date of Enhancement | SPE Start Date | SPE Flux (pfu) |
|---|---|---|---|
| 064 | 24 August 2002 | 24 August 2002 | 317 |
| 065 | 28 October 2003 | 28 October 2003 | 29500 |
| 066 | 29 October 2003 | 28 October 2003 | 29500 |
| 067 | 02 November 2003 | 2 November 2003 | 1570 |
| 068 | 17 January 2005 | 16 January 2005 | 5040 |
| 069 | 20 January 2005 | - | - |
| 070 | 13 December 2006 | 13 Dec 2006 | 698 |
| 071 | 17 May 2012 | 17 May 2012 | 255 |

**Table 9.4**: List of Ground Level Enhancements

It is seen that the 20 January 2005 enhancement had no correlating SPE, but this may have simply been an extension of the 16 January event. Also worthy of noting is that the 24 August 2002 and 17 May 2012 enhancements had SPE fluxes of less than 500 pfu, hardly a large flux. Finally events 65 to 67 can probably all be considered part of the same long duration events. However for the purposes of this analysis they have been treated separately.

Table 9.5 shows the region parameters associated with GLEs, with the parameters explained in table 7.6. Again there no parameters that stand out as being unique to these type of events.

The unfortunate conclusion is that morphological analysis does not yield good prediction results for SPEs. This is probably due to the particles themselves not usually being accelerated from the region, but rather being accelerated by one of the Fermi methods, several solar radii or further out from the Sun.

| Date | Field (G) | Flux (G.μh) | ILL | ILC | Gradient (G/Mm) | Orientation (deg) | Delta Region |
|---|---|---|---|---|---|---|---|
| 2002/08/23 | 1416 | 970951 | 16 | 2.465 | 46.7 | 32 | N |
| 2002/08/23 | 776 | 136654 | 10 | 4.580 | 24.9 | 66 | N |
| 2003/10/27 | 1465 | 2383848 | 51 | 6.358 | 38.5 | 63 | Y |
| 2003/10/27 | 1306 | 2789094 | 62 | 5.385 | 32.2 | 73 | Y |
| 2003/11/01 | 294 | 10172 | 5 | 2.632 | 5.1 | 74 | N |
| 2005/01/15 | 1251 | 2524518 | 23 | 4.008 | 76.4 | 03 | Y |
| 2005/01/20 | - | - | - | - | - | - | - |
| 2006/12/12 | 1402 | 1074633 | 20 | 3.448 | 87.6 | -26 | Y |
| 2012/05/17 | 677 | 296458 | 16 | 3.800 | 59.6 | -81 | Y |

**Table 9.5**: Region Parameters Associated with Ground Level Enhancements

### 9.5 Summary and Conclusions

Solar particles are protons and occasionally heavier ions that are accelerated from the Sun. The threshold for a particle event is a flux greater than 10 pfu for particles with more that 10 MeV energy. These events can occur at any time during the sunspot cycle although are more common near sunspot maximum.

The three main mechanisms for accelerating these particle are flare associated acceleration, first order Fermi acceleration and second order Fermi acceleration.

Solar radio signatures were examined in order to determine which may provide the best short-term (within minutes to hours) prediction of SPEs. From literature, the Type 3L emission seems to be the best with greater than 80% success. However type 3Lm (metric 3L) has a up to 69% success rate. The worst was the Castielli-U event with 21% success.

The results from GONG region analysis were used to establish if there were any significant sunspot features that were distinctive prior to SPEs. Unfortunately none were found that could be used as predictor for SPEs.

Events associated with Ground Level Enhancements were also analysed and correlated with SPEs. Again, the results were poor, only showing that usually large magnetic fluxes were required, but most regions were not associated with GLEs.

The most likely cause for morphological analysis being a poor guide to SPE prediction is that the particles are most often accelerated away from the Sun at or between CME shocks (Gargate et al., 2014). The conditions in the shocks need to be suitable to accelerate particles, which is why the majority of CMEs are not associated with SPEs (Park et al., 2012).

# 10. Automated Hα Flare Analysis

## 10.1 Introduction

As discussed in chapter 5, current Hα flare analysis techniques often yield inconsistent results between sites. The GONG network offers another network for inter-site comparison. To be fair, SOON flare analysis is largely automated, but it still requires human input at various times. For example, the flare threshold is sometimes adjusted if a region has a long duration and low intensity flare (plage fluctuations). The SOON system requires an analyst to set the regions to be analysed for flares, and once a flare occurs, to select the flare centre. These regions are 'boxed in', and any flare outside a boxed region will not be analysed. As a result Hyder flares (also known as two ribbon flares), flares outside an active region associated with disappearing filaments (Hyder, 1967), may be missed, or an inattentive analyst may not have boxed in an active region.

Most of these limitations can be overcome with full automation and analysis of GONG Hα images.

The GONG Hα system splits off light from the main light path and redirects it through a mica birefringent filter with a bandwidth of approximately 0.04 nm. The camera gives a 2048 by 2048 pixel image. The final images are FITS files with a 16 bit unsigned integer for each pixel, giving 8 MB sized images.

## 10.2 Algorithm

The general outline of the flare analysis algorithm is shown in figure 10.1. Described below are each of the steps in the algorithm.

### 10.2.1 Normalisation and Limb Darkening Correction

Each Hα image is pre-processed as described in chapter 6. This is to ensure consistency across the entire range of the data set.

### 10.2.2 Find Flaring Pixels

The disk median pixel brightness value is found for the limb corrected image. Each pixel in the disk is examined to see if it has 150% of the median disk brightness. If a pixel does then it is flagged as being a flaring pixel.

### 10.2.3: Total Area of Flaring Pixels

The total area of flaring pixels is found. If the total flaring area is large, it means the limb darkening correction did not work. This usually occurs if the Sun is partially obscured by clouds, causing some pixels to be much brighter and others much darker. If the total area of flaring pixels is greater than 10000 μh, then the limb correction has failed and the image is unsuitable for further processing.

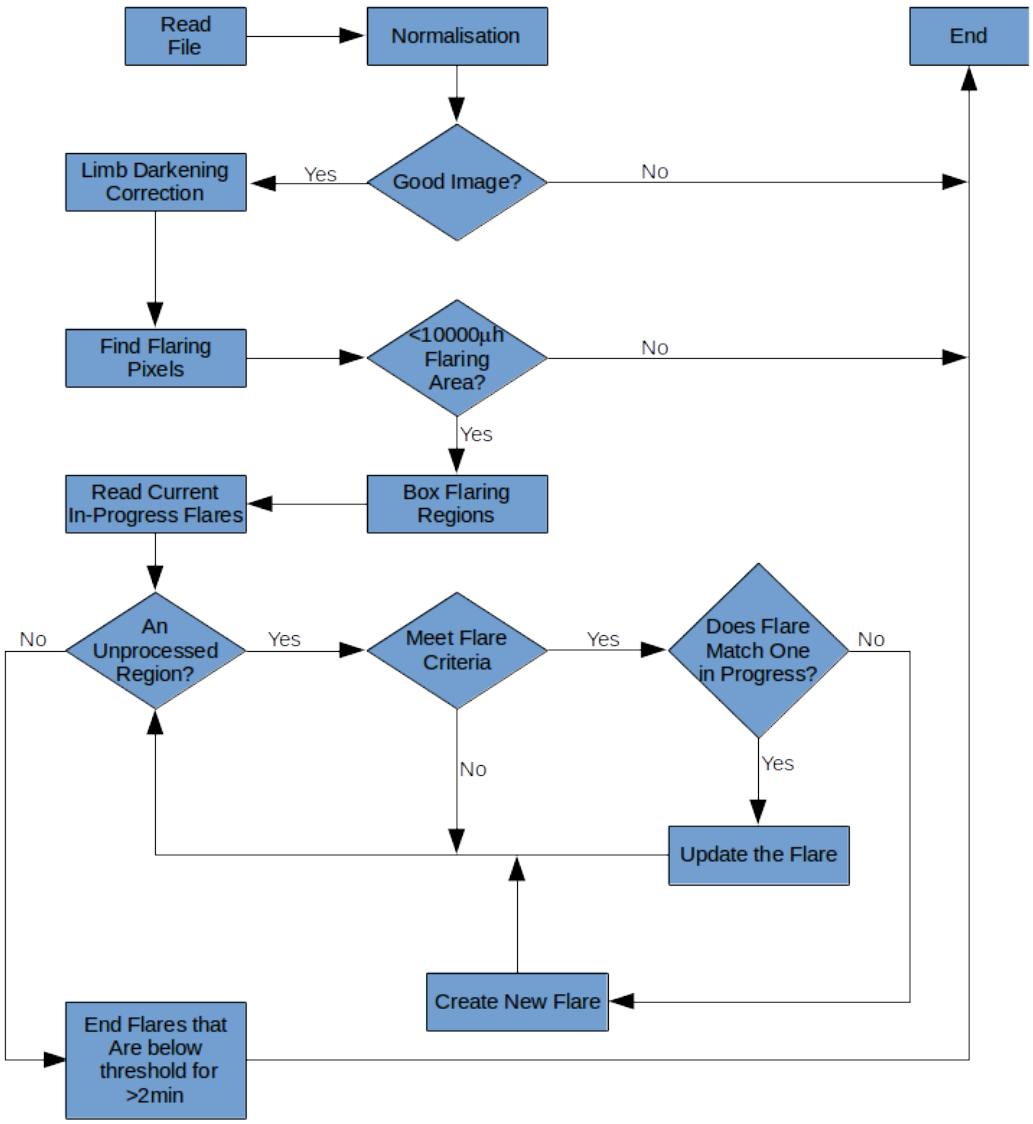



**Figure 10.1**: Hα Flare Analysis Flow Chart

### 10.2.4 Box Flaring Regions

The flaring pixels are grouped together by proximity. Flaring pixels within 5 degrees of each other are grouped together. Initially a region is boxed that extends 30 pixels out from the lower, upper, left and right extreme flaring pixels in a region. If the number of flaring pixels is greater than half the number of pixels in the box then each dimension of the box is expanded by a factor of 1.5 times the extent of the flaring pixels. This ensures that more than half of the box contains non-flaring pixels. This amount of 'quiet Sun' is required so that the brightness of the background Sun pixels can be determined (the mode of the pixel brightness in a region is taken as the background Sun).

This means that for a region, the box size is dynamic. The advantage of this system is that it is unlikely that a flaring pixel will be outside the box. However the drawback is that the same pixels aren't necessarily being sampled from minute-to-minute. This should not be a problem provided the number of non-flaring pixels is sufficiently large.

The other potential effect that this has is that flares with multiple centres maybe split. For very large regions, flare centres can be more than 5 degrees apart. However, if these other flaring centres meet flare criteria then they will be recorded as separate flares and, if necessary, can be combined later.

The advantages of a global search algorithm with adjustable boxes outweighs the inconvenience of some extra post-processing.

### 10.2.5 Read Current Flares in Progress

Flares from previous images that have not met the end-of-flare criteria are loaded from file. Each ongoing flare's parameters are read as is the last peak histogram and the brightness and area of the flare for each minute that the flare has been in progress.

### 10.2.6 Loop through unprocessed regions

Each region that has been found is analysed and processed individually.

### 10.2.7 Meeting Flare Criteria

Each region has a histogram generated. The histogram is created from pixels found within

the box created in 10.2.4. If a pixel is at an Rv of more than 0.95, it is omitted from the analysis because the limb darkening correction does not work well beyond this range. In addition it excludes pixels that are not on the disk from the analysis. The entire box is re-analysed, even though 'flaring' pixels have been found. This is because the pixels found in 10.2.2 were merely to identify regions, not for formal flare analysis. The new histogram is rescaled so that the mode occurs at an intensity value of 100%. The range of the histogram is 0 to 650%, with pixels brighter than this being assigned a value of 650.

### 10.2.8 Flare Matching

Each newly found flare is compared against flares found in the previous image processed. If a new flare centre is within 5º of a previous flare centre then the older flare is updated with the new flare parameters (typically just the flare end time). The peak times, intensity, area and centre are only updated if the last analysis produced a higher flare intensity than the previous peak intensity.

### 10.2.9 Creating a New Flare.

A new flare is added to the list of flaring regions if a flare is found that could not be matched to a current in-progress flare.

### 10.2.10 Ending Flares

In-progress flares are ended if flare criteria are not met in a region for 2 successive images (2 minutes). In addition, if there have been no observations for 5 minutes (due to cloud, equipment problems, maintenance or sunset) then flares are also ended. This criteria is somewhat arbitrary, but a short time span was chosen to be sure that two flares were not treated as one. It is relatively straightforward in post-processing to merge data from 2 flares together based on information from other sites.

### 10.2.11 Implementation

This algorithm was implemented in Java and run under Linux. It is run within a script that passes the program files to be processed in sequential order. In a real time environment the files are pulled off a server and run once a minute using a scheduling script. The program was tested in this manner at Learmonth Solar Observatory. Archive data is processed in a similar manner, with a script passing the archived files to be processed to

the program.

Each iteration of the program takes between 10 and 15 seconds. The fundamental requirement for a real-time system is that it take less than one minute per iteration, as this is the acquisition rate for imagery at each site. When processing archived data, an 8 core computer was used, using 1 core for each GONG site. This produced little degradation in the time of processing, still only requiring 15 seconds per image per site, therefore 6 images per 15 seconds (each from a different site) can be analysed.

### 10.2.12 Output files

Several files are output after a flare has ended. The first is a text file that describes the flare. An example is shown in figure 10.2. From this data the history of the flare is known without needing any reanalysis of the original imagery.

From this data three graphs are generated (at the time of analysis of the flare). Firstly, figure 10.3 shows the peak histogram of a flare observed at Learmonth Solar Observatory. This is simply data plotted from lines 5 to 656 in the flare text file (figure 10.3).

Secondly, figure 10.4 shows the area vs time plot for this flare. This is a plot of the time and area columns from figure 10.2.

Finally, figure 10.5 shows the brightness vs time plot for this flare. This is a plot of the FLB (Flare Brightness) and time columns in figure 10.2.

In addition to these plots, actual images of the flares are archived. A 512 by 512 pixel image centred on the flare centre is recorded for each minute that the flare is in progress. These can be compiled later into an animated GIF or other movie format. Figure 10.6 shows a sample of these images for this example flare.

```
S    DATE    STARTTIME  PKTIME   LOCATION   AREA  FLB MXB IB ENDTIME
L 2012/07/11 08:27:34  08:28:34 S22.5E04.8 0195.4 276 301 1N 08:47:34

Peak Histogram
BIN AREA
000 0000.00
001 0000.00
…
098 0496.25
099 0594.62
100 0907.78
101 0542.91
102 0473.04
…
648 0000.00
649 0000.00
650 0000.00

FLARE PROGRESSION
DATE       TIME     AREA    FLB
2012/07/11 08:27:34 0163.01 237
2012/07/11 08:28:34 0195.44 276
2012/07/11 08:29:34 0178.67 265
2012/07/11 08:30:34 0180.10 260
2012/07/11 08:31:34 0178.62 247
2012/07/11 08:32:34 0180.69 252
2012/07/11 08:33:34 0174.56 231
2012/07/11 08:34:34 0176.88 232
2012/07/11 08:35:34 0167.48 220
2012/07/11 08:36:34 0147.54 211
2012/07/11 08:37:34 0128.75 199
2012/07/11 08:38:34 0105.94 198
2012/07/11 08:40:34 0086.79 192
2012/07/11 08:41:34 0077.19 189
2012/07/11 08:42:34 0068.30 182
2012/07/11 08:43:34 0074.48 175
2012/07/11 08:44:34 0081.88 178
2012/07/11 08:45:34 0076.31 174
2012/07/11 08:46:34 0074.52 168
2012/07/11 08:47:34 0057.20 167
```

**Figure 10.2**: Sample Flare Text Output. Line 2 provides the basic flare parameter summary. Lines 5 to 656 is the peak histogram. Lines 660 onwards show the area and brightness for the entire duration of the flare.

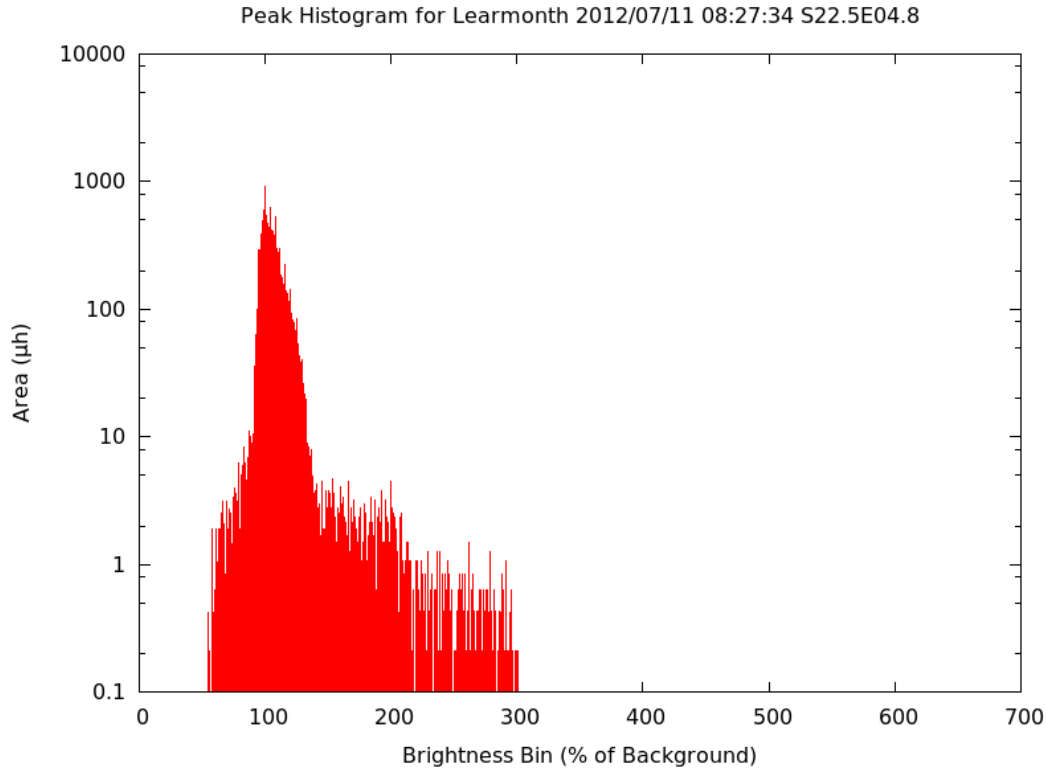



**Figure 10.3**: Example Peak Histogram from GONG Learmonth

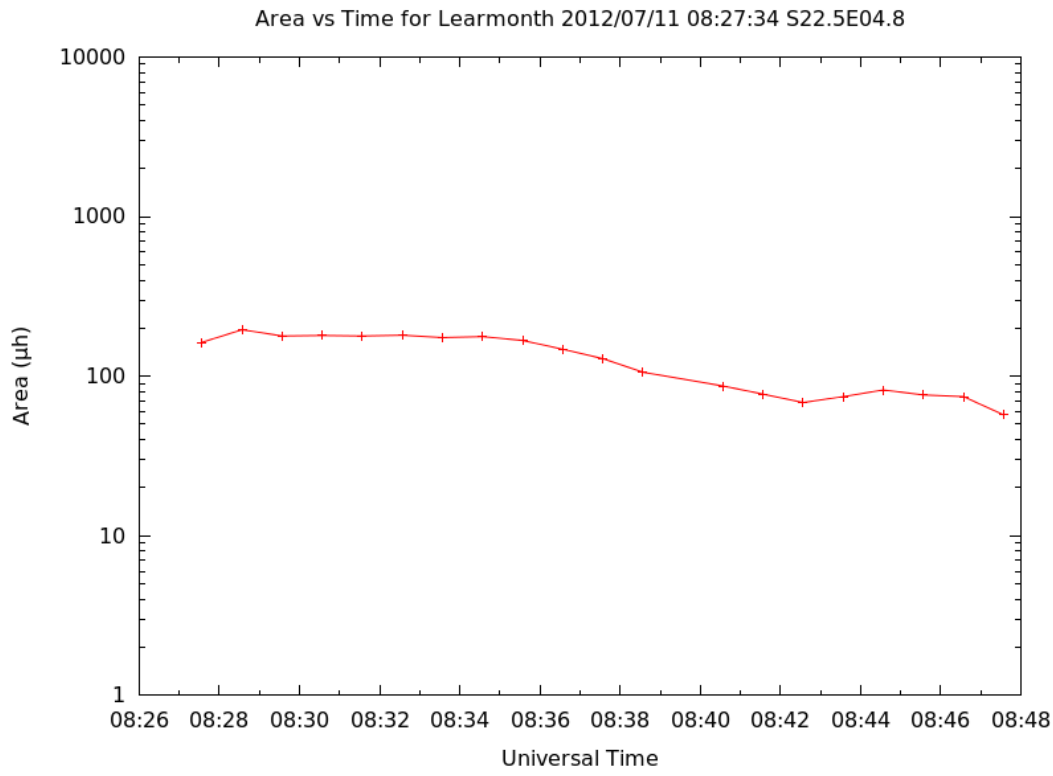



**Figure 10.4**: Example Area vs Time Plot for GONG Learmonth

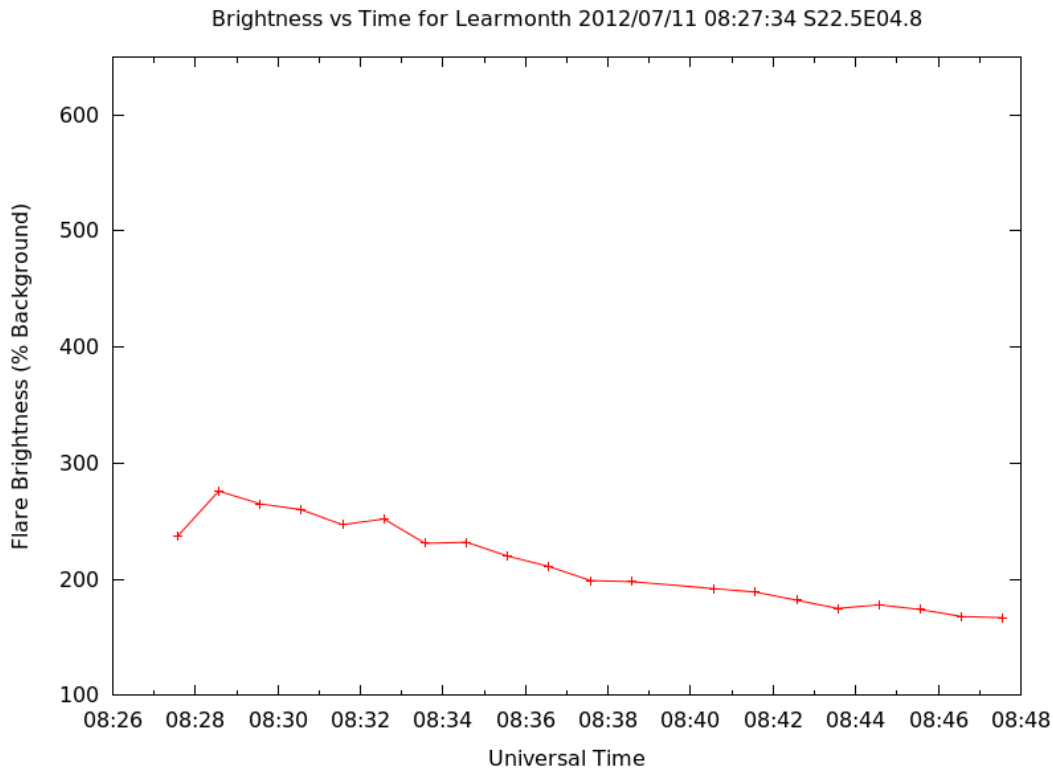
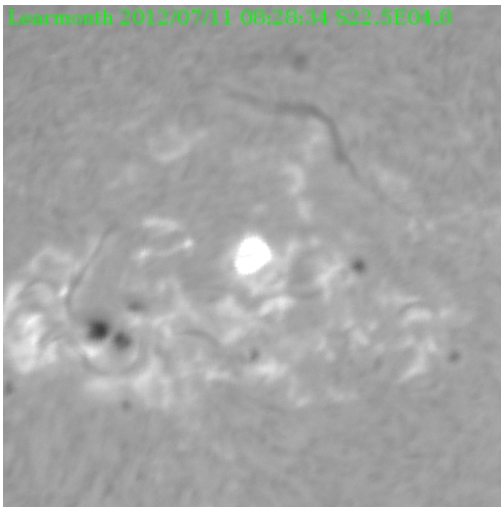
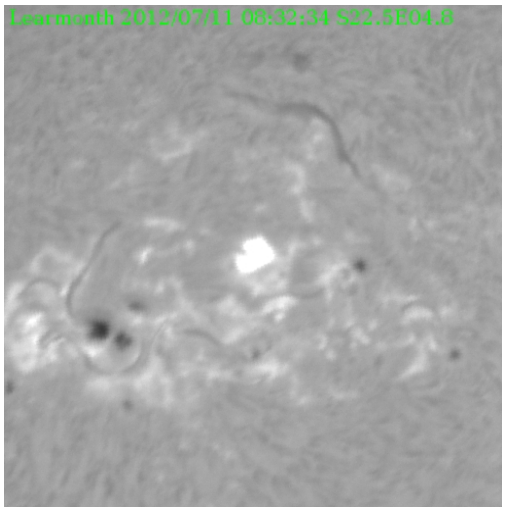



**Figure 10.5**: Example Brightness vs Time Plot for GONG Learmonth

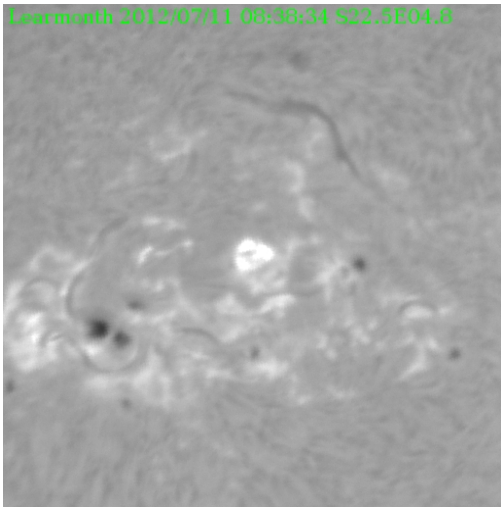
These images explain the shape of the histogram in figure 10.3. As noted earlier, the histogram is deliberately scaled so the largest brightness bin is centred on the value 100. The tail to the right of the peak is due to the flare, in addition to the plage in the region which is brighter than the background but not flare-bright. The truncated tail to the left of the peak is due to the filaments and spots that can be seen in the image.



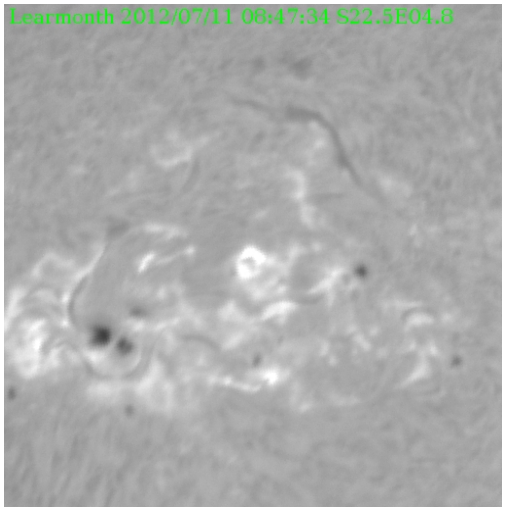

(a) 08:28:34 – Peak brightness




(b) 08:32:34




(c) 08:38:34




(d) 08:47:34 - Flare End

**Figure 10.6**: Images of the Example Flare From GONG Learmonth

These files and data show that, at least at first glance, the algorithm is producing sensible results for individual sites.

### 10.3 GONG Flare Comparisons

As discussed in chapter 5, getting sites to agree on flare brightness and areas is difficult. Does the GONG Hα system perform any better than the USAF SOON system? If it doesn't, can correction factors be applied to bring consistency to this data? The advantage

with GONG is that all of the data is available , including actual values for brightness and area, rather than just the classes.

### 10.3.1 Data

The GONG system operates at each site while the Sun has an elevation greater than 10 degrees. Lower elevations mean the images are distorted due to poor seeing from looking through significant atmosphere. However, as there are 6 sites, normally two are observing the Sun at any given time.

The data set chosen was initially the entire month of June of 2012. There were a particularly high number of flares at this time, which meant a good sample of flares could be examined without processing a large number of images. However, this is also when Udaipur closes down due to the monsoon season in India. Therefore, images from May to July 2012 were also examined to enlarge the data set for all sites.

This provides a lot of images to analyse, so in order to reduce the task, the list of flare reports from SOON sites for this time was extracted from NGDC. GONG images one hour either side of the start and end time of a flare reported by USAF were examined. This means the sample is slightly biased by what the USAF reported, but will be convenient for comparing GONG flares with USAF flares.

### 10.3.2 GONG Inter-site Comparisons

As with the SOON flare data, comparisons were made between sites for the flare brightness and areas to determine if they performed any better than the USAF SOON network.

As the actual area and brightness values are known, trend lines were fitted to the data rather than using the flare classifications.

In some cases, two or more flares from one site occurred during the entire duration of a flare at another site. These multiple flares were treated as one. Such interruptions could have been caused by cloud or the flare dropping below thresholds at one site but not the other. Also analysis was only performed on site pairs when there were more than 5 flares observed at both sites.

Figure 10.7 shows the brightness comparisons between sites.

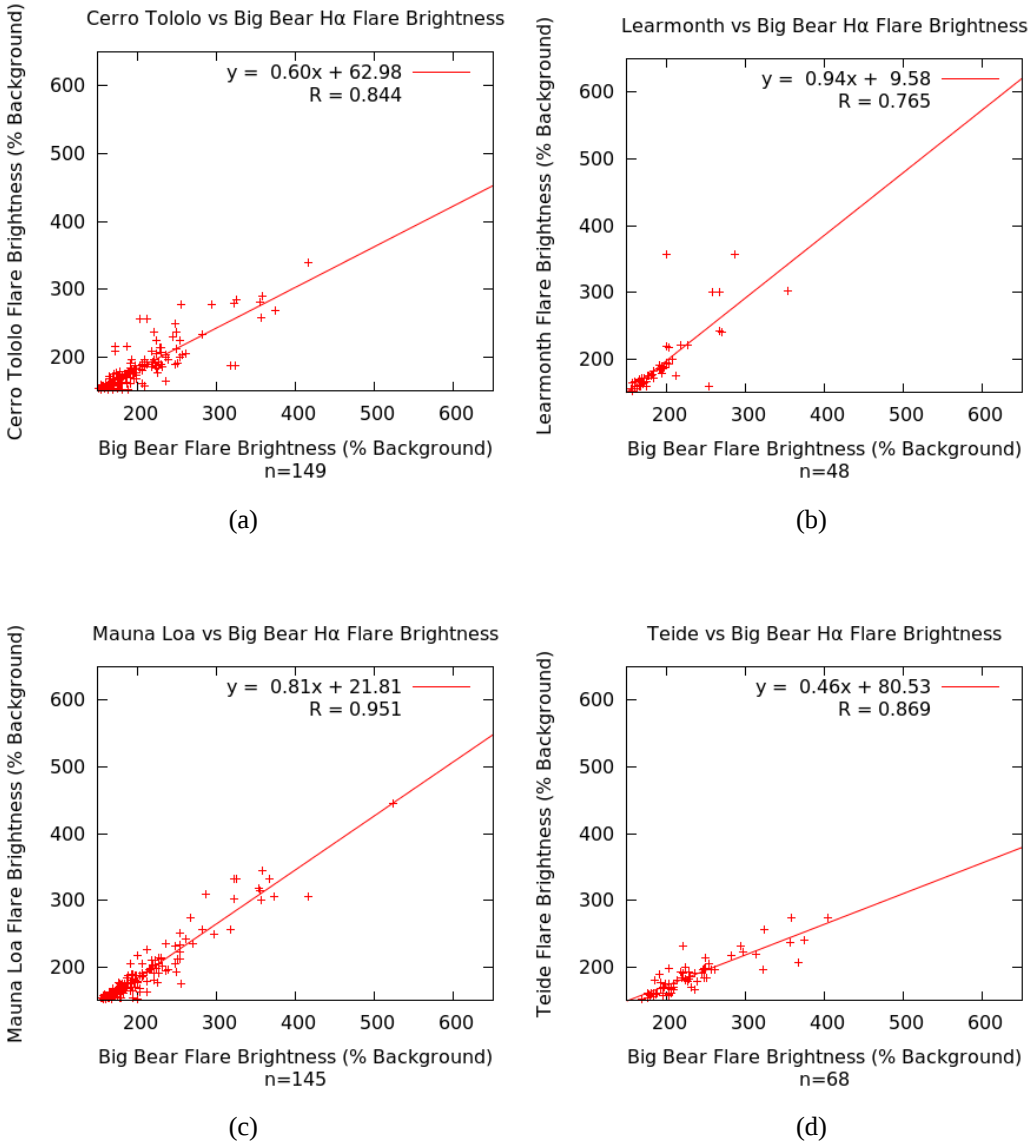
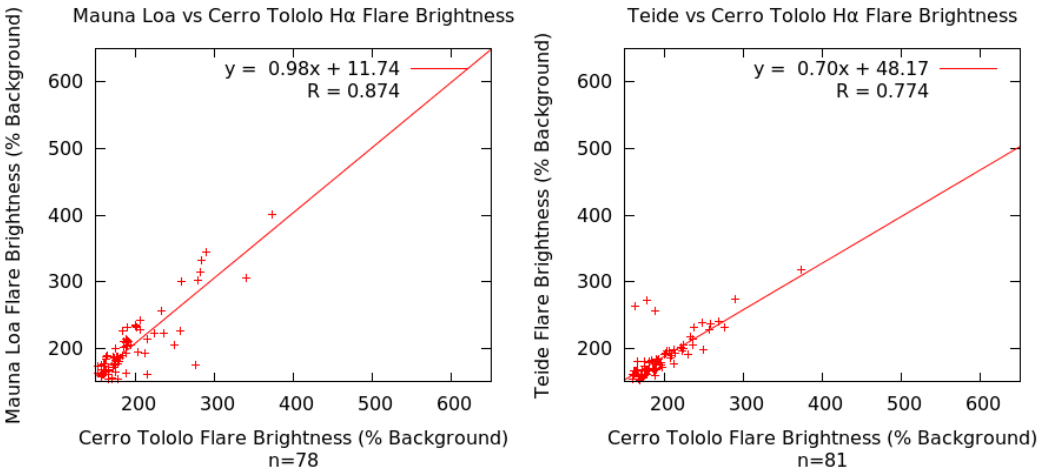
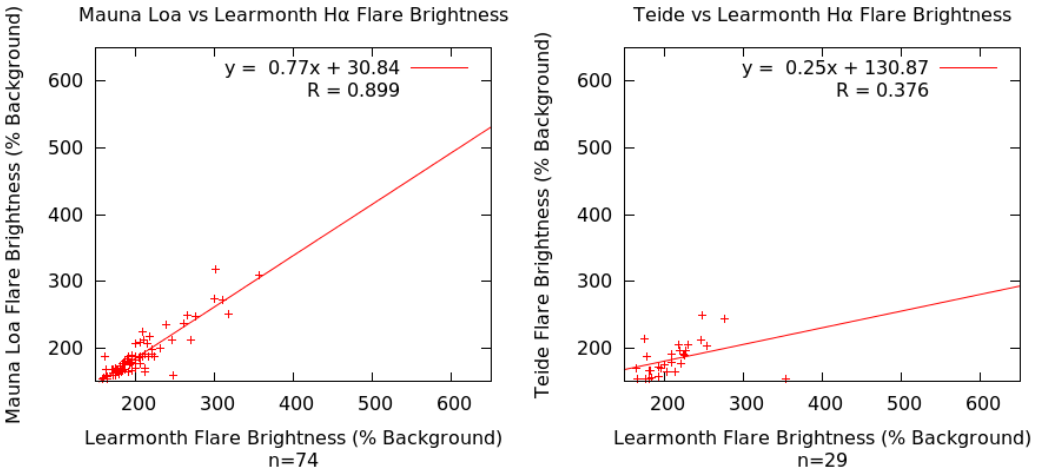



**Figure 10.7 (a-d)**: Plots of Relationships Between GONG sites for Hα Flare Brightness

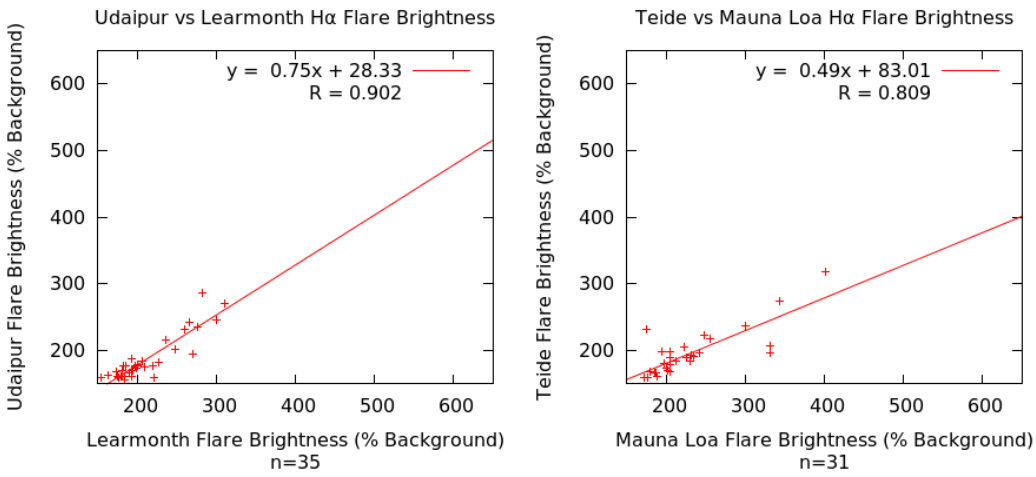



(e) (f)




(g) (h)

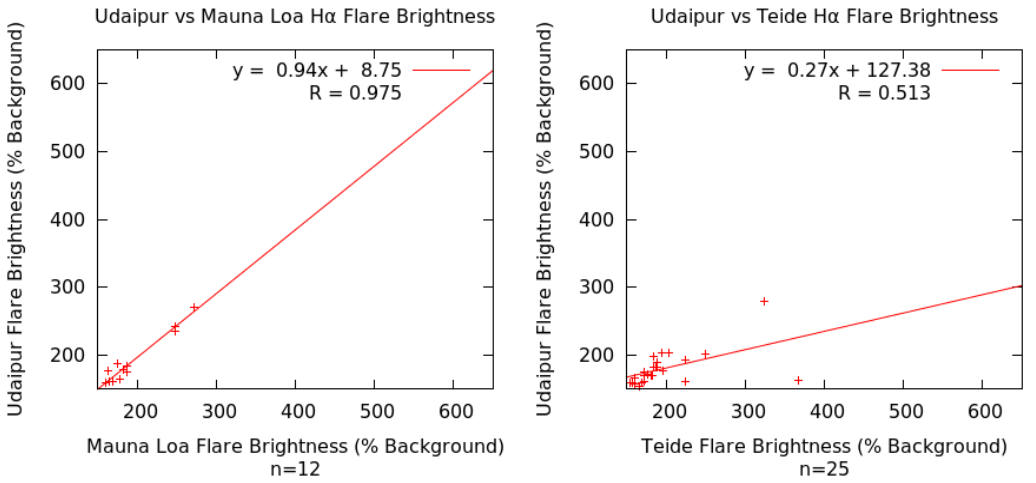
**Figure 10.7 (e-h)**: Plots of Relationships Between GONG sites for Hα Flare Brightness




(i)




(j)




(k)




(l)

**Figure 10.7 (i-l)**: Plots of Relationships Between GONG sites for Hα Flare Brightness

These plots show wide ranges of gradients, intercepts and correlation coefficients, suggesting initially that, as with the SOON sites, there is a lack of consistency across the network. The parameters of these fitted lines are summarised in table 10.1

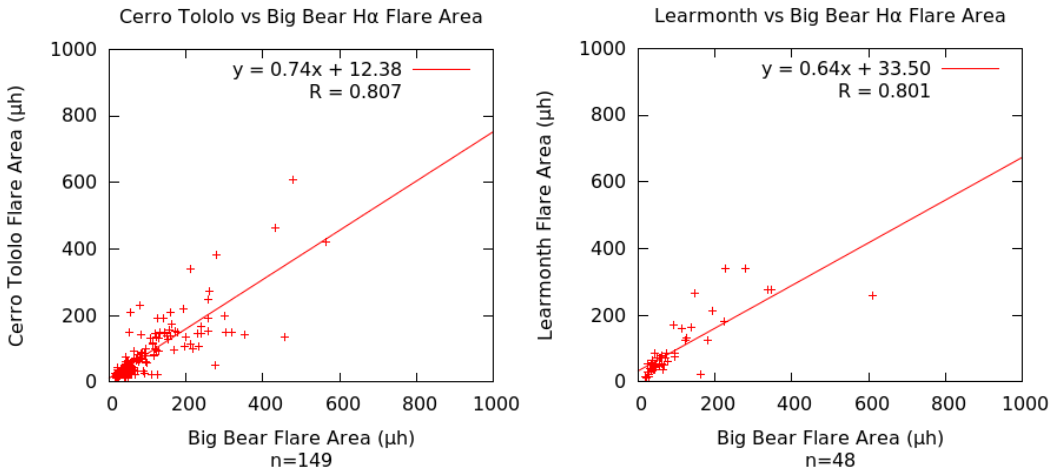
| Sites | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| Cerro Tololo vs Big Bear | 0.599 | 0.031 | 62.984 | 6.584 | 0.844 |
| Learmonth vs Big Bear | 0.944 | 0.117 | 9.576 | 23.492 | 0.765 |
| Mauna Loa vs Big Bear | 0.811 | 0.022 | 21.810 | 4.812 | 0.951 |
| Teide vs Big Bear | 0.457 | 0.032 | 80.529 | 7.666 | 0.869 |
| Mauna Loa vs Cerro Tololo | 0.976 | 0.062 | 11.738 | 12.337 | 0.874 |
| Teide vs Cerro Tololo | 0.702 | 0.065 | 48.167 | 12.989 | 0.774 |
| Mauna Loa vs Learmonth | 0.769 | 0.044 | 30.845 | 9.276 | 0.899 |
| Teide vs Learmonth | 0.249 | 0.118 | 130.868 | 25.487 | 0.376 |
| Udaipur vs Learmonth | 0.748 | 0.062 | 28.331 | 13.426 | 0.902 |
| Teide vs Mauna Loa | 0.492 | 0.066 | 83.007 | 15.622 | 0.809 |
| Udaipur vs Mauna Loa | 0.938 | 0.068 | 8.750 | 13.461 | 0.975 |
| Udaipur vs Teide | 0.265 | 0.092 | 127.376 | 18.841 | 0.513 |

**Table 10.1**: Fit Parameters for GONG Site Flare Brightness Comparison

Of note is that all but two of the correlation coefficients (r) are greater than 0.7, suggesting that for most pairs of sites there is a reasonable correlation, possibly enabling adjustment of flare brightness in some cases. However, the range in gradients suggest that the response of filters varies across the network.

Also of interest is the intercepts (b) and in particular sites compared with Teide have large offsets. Teide vs Learmonth (with the worst intercept) can probably be explained by being far apart in longitude and therefore at opposite ends of the day when seeing conditions are likely to be poor. However, there are only three comparisons (Udaipur vs Mauna Loa, Mauna Loa vs Cerro Tololo, and Learmonth vs Big Bear) where the intercepts are within 2 standard deviations ($2s_b$) of the origin.

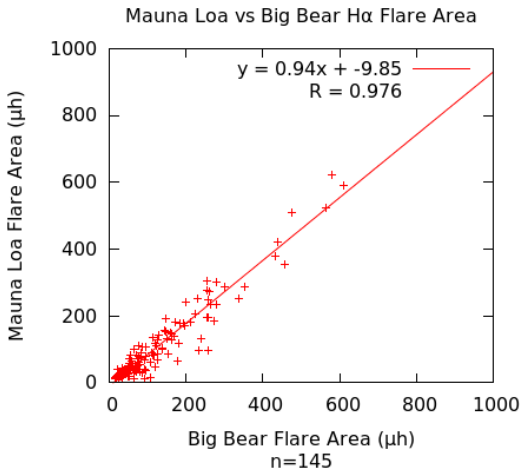
Figure 10.8 shows the flare area comparison between sites and table 10.2 the summary of the corresponding lines of regression. As for brightness most of the correlation coefficients are above 0.7 suggesting reasonable correlation between sites, but the range of gradients suggest inconsistencies across the network.



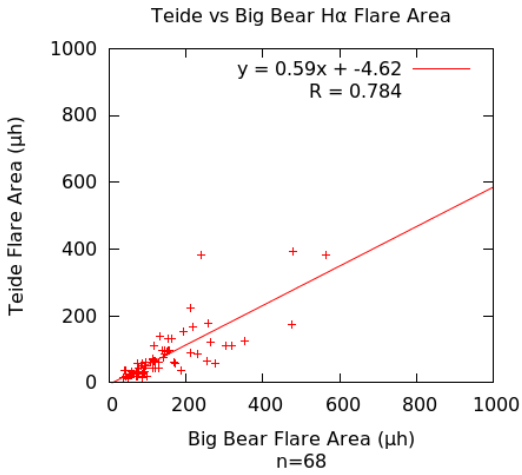

(a)




(b)

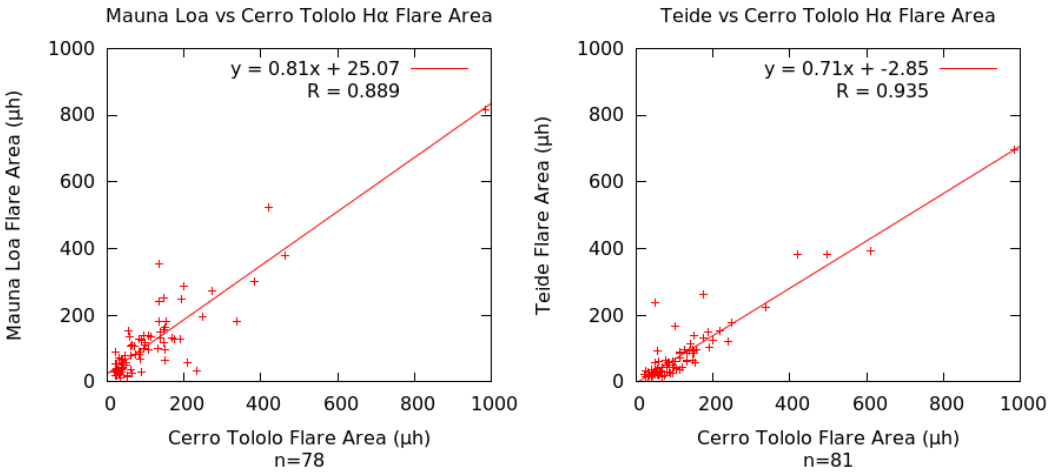



(c)

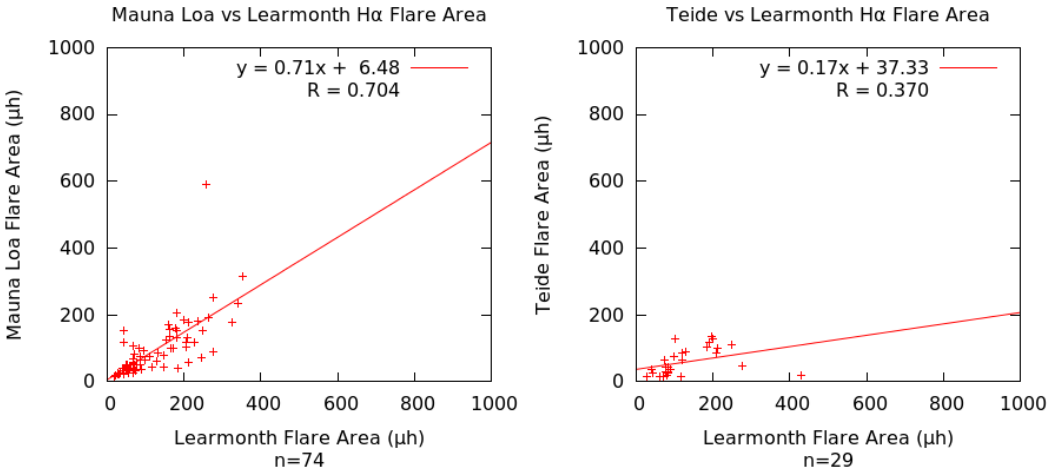



(d)

**Figure 10.8 (a-d)**: Plots of Relationships Between GONG sites for Hα Flare Area

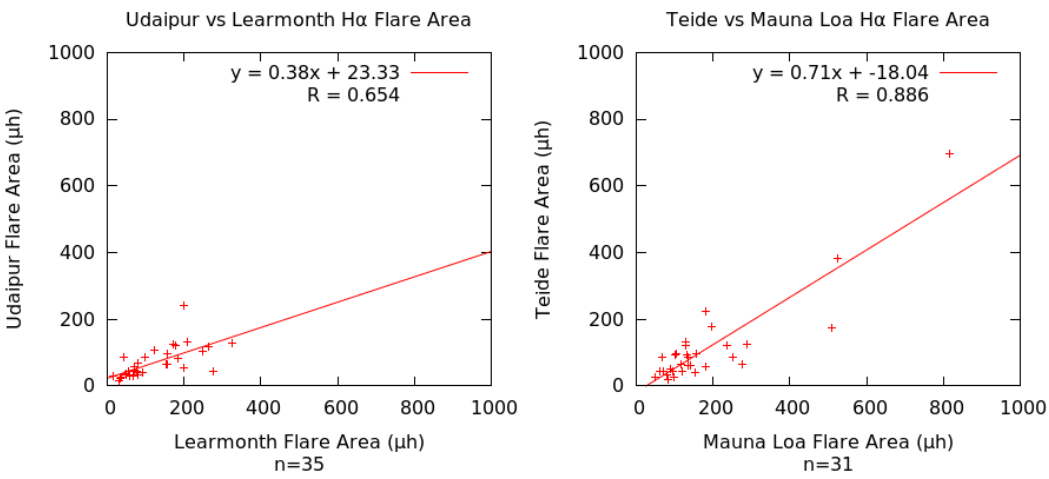



(e) (f)




(g) (h)

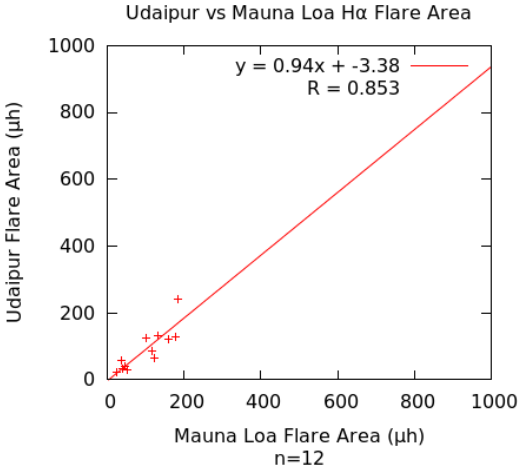
**Figure 10.8 (e-h)**: Plots of Relationships Between GONG sites for Hα Flare Area



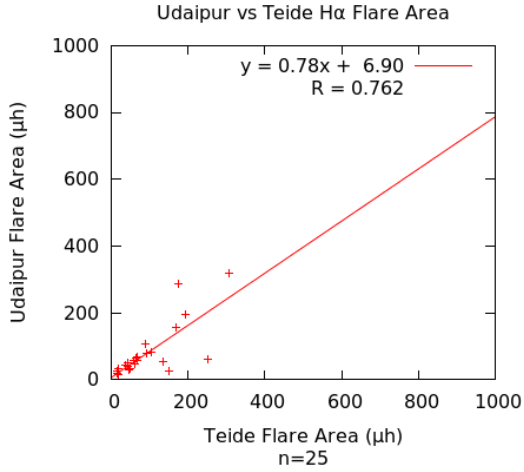

(i)




(j)




(k)




(l)

**Figure 10.8 (i-l)**: Plots of Relationships Between GONG sites for Hα Flare Area

| Sites | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|
| Cerro Tololo vs Big Bear | 0.736 | 0.044 | 12.378 | 6.364 | 0.807 |
| Learmonth vs Big Bear | 0.643 | 0.071 | 33.499 | 10.529 | 0.801 |
| Mauna Loa vs Big Bear | 0.940 | 0.018 | -9.850 | 3.558 | 0.976 |
| Teide vs Big Bear | 0.585 | 0.057 | -4.622 | 10.414 | 0.784 |
| Mauna Loa vs Cerro Tololo | 0.805 | 0.048 | 25.073 | 8.519 | 0.889 |
| Teide vs Cerro Tololo | 0.712 | 0.030 | -2.855 | 5.659 | 0.935 |
| Mauna Loa vs Learmonth | 0.709 | 0.084 | 6.480 | 12.769 | 0.704 |
| Teide vs Learmonth | 0.171 | 0.083 | 37.332 | 12.973 | 0.370 |
| Udaipur vs Learmonth | 0.376 | 0.076 | 23.330 | 10.788 | 0.654 |
| Teide vs Mauna Loa | 0.708 | 0.069 | -18.045 | 16.736 | 0.886 |
| Udaipur vs Mauna Loa | 0.937 | 0.182 | -3.379 | 20.739 | 0.853 |
| Udaipur vs Teide | 0.781 | 0.139 | 6.901 | 16.597 | 0.762 |

**Table 10.2**: Fit Parameters for GONG Site Flare Area Comparison

Table 10.2 shows that, similar to brightness, Teide – Learmonth has the worst correlation for area of flares. However most of the site comparison's intercepts fall within two standard deviations of the origin. There is a significant variation in gradients, which largely cannot be accounted for in statistical variation.

At this stage it is difficult to comment if these differences are due to variations in the filters or if they are specific with the sites themselves (eg weather or seeing conditions). Another contributing factor may be that the sites do not acquire images at the same time. They are staggered 20 seconds apart. This was done so that 3 images would be acquired across the network with 20 second time resolution. However this technique only works if the Hα system response at each site is identical. SOON acquires an image every 30 seconds at each site, even though they only report the times to the nearest minute. Flares can increase in brightness rapidly, therefore, a 1 minute sampling interval may not be sufficient to capture the peak brightness causing discrepancies between sites. One might expect, however, that this would average out (sometimes the flare peak may be before an observation at a site and at another site at a time after it).

### 10.3.3 GONG – SOON Comparisons

The co-location of GONG and USAF SOON equipment at Learmonth Solar Observatory

provides a unique opportunity to compare two different sets of Hα equipment. In this case the effects of seeing and weather should be consistent for both instruments.

It was stated in chapter 5 that NGDC does not contain the brightness bins or areas for flares, only the classes. However, up until mid-2014, the original messages transmitted from the USAF observatories were stored on the individual computers at each site. These messages were lost whenever a hard disk failed, a computer was replaced or upgrades performed. The new system installed mid-2014 only stores messages for 45 days before they are deleted.

The original messages for June to August 2012 were obtained from the USAF computers on-site. From these messages the brightness and area values for flares observed at Learmonth were extracted and compared against those measure from GONG.

Figures 10.7 and 10.8 show brightness and area comparisons for flares observed on both instruments. These graphs show very poor correlation, suggesting that most of the variance between sites is due to instrumentation and not site location.

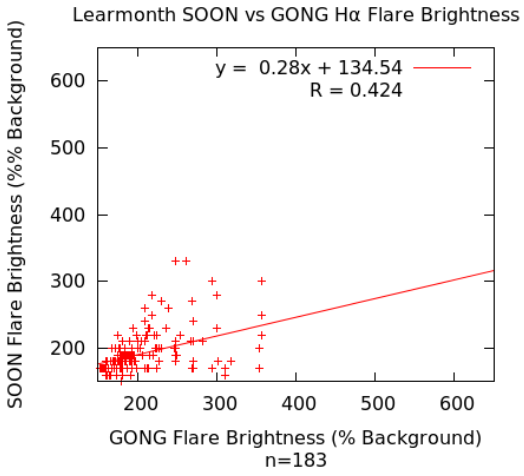



**Figure 10.7**: SOON vs GONG Flare Brightness for Learmonth

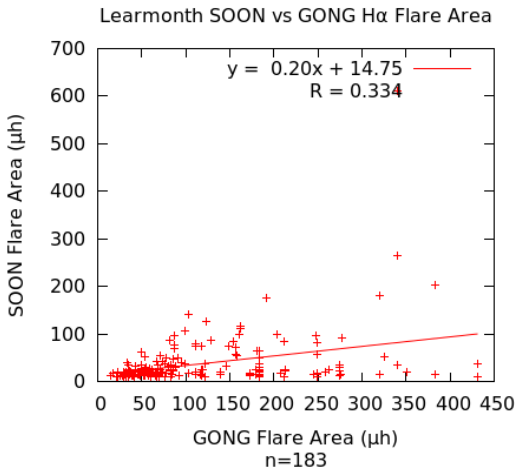



**Figure 10.8**: SOON vs GONG Flare Area for Learmonth

### 10.4 Summary and Conclusions

An algorithm was developed that automatically analyses GONG Hα imagery for flares, measures and quantifies them. Flares from the six sites were then compared to see if the sites produced the same results for the same flares.

In general there was a greater than 0.7 correlation coefficient (50% correlation) between the sites, but both the area and brightness gradients in the fits varied wildly. For some cases the measured values at one site averaged double the flare brightness and area at another (for example Teide and Big Bear). Thus, while results are probably more consistent across the network, there are great variations in the values measured.

The Learmonth GONG and SOON flares were compared as these are co-located at Learmonth Solar Observatory. Analysis of this data produced a very poor correlation, less than 0.5 for both area and intensity. Therefore it would seem that variations are more dependent on equipment than individual seeing conditions.

After 8 decades of regular Hα observations, little improvement has been seen in consistency of measurement of flares. It is suspected that it is the filter that is the primary cause for these discrepancies, due to the tolerances required in manufacturing and operation. However all of these telescopes have multiple mirrors and lenses in the light path which could potentially contribute to variations. Also different responses of the camera, particularly for bright flares if the camera becomes non-linear, may be an

additional contributing factor.

# 11. Conclusions and Recommendations

## 11.1 Conclusions

In this thesis it has been shown that solar indices are not constant. Specifically, sunspot number, sunspot area and 10.7 cm flux vary significantly between sunspot cycles. Part of this variation for some indices is due to measurement error, but some is also due to intrinsic solar variation.

The Solar Influences Data Center in recent years has attempted to 'correct' the historical sunspot number. In July 2015 the new corrected sunspot number series was released. Due to the timing of this release, the analysis was performed on the old series.

In order to show the difficulties involved in measuring these indices, the sunspot numbers and areas from different sites or organisations were compared. There appears to be a change in the measurement of these indices at around each sunspot minimum and maximum. To add further complexity, the time of minima and maxima for each index can be different, implying that there is not a single solar cycle, but that different indices have their own cycles.

For sunspot areas, USAF SOON sites took over measurement from Royal Greenwich Observatory in 1980. An analysis was performed on sunspot drawings from 1990 and 2002. These drawings were remeasured using a grid rather than the ellipses that the USAF use. It was found that there was up to a 12% difference between the reported and remeasured areas. It was found that 9% of this error was due to the way the SOON analysts round the limb area correction factors down. A simple change of technique will reduce this error. This compares with Wilson and Hathaway (2005) and Foukal's (2014) claims of a 20% to 40% sunspot area deficiency in sunspot area.

In addition, site comparisons were made of solar radio bursts from the USAF RSTN sites. Few of the sites and frequencies showed statistically significant agreement in the peak flux reports. That is to say when the peak burst fluxes of two sites were plotted against each other, the slopes differed significantly (at 95% confidence) from an expected value of 1. This is probably due to the difficulties of calibration the radio receivers identically over the entirety of RSTN.

For Hα flares, there was good agreement between USAF SOON sites for faint flares.

However, for brighter flares sites varied significantly. The same was true for flare area, with larger area flares often disagreeing by more than one category class.

It is fair to say then, that solar observers still have work to do in improving these observations. The question then becomes, “is manual reduction part of the problem and can automated techniques be used to mitigate these errors?”

To answer this question, a set of programs was developed to analyse GONG intensity, magnetogram and Hα images.

In the case of intensity and magnetogram images, the software located sunspot regions, counted sunspots, computed sunspot areas and performed sunspot and magnetic field classifications. One intensity-magnetogram image pair per day was examined from each GONG site. The results of this analysis was compared to the Solar Region Summary reports produced each day by SWPC.

The relationship between the smoothed GONG sunspot number, $R_G$, and old International Sunspot Number, $R_I$, for cycle 23 was found to be:

$$R_G = (0.9069 \pm 0.0040) R_i + (1.10 \pm 0.17) \quad , \qquad \text{Eqn 11.1}$$

and for cycle 24:

$$R_G = (0.9231 \pm 0.0041) R_i + (0.70 \pm 0.21) \quad . \qquad \text{Eqn 11.2}$$

This is a much improved result given other site comparisons, but these cycles are still statistically different even though they only vary by about 2%.

A similar approach was taken for the area, in this case the comparison between GONG sunspot area $A_G$ and the SRS sunspot area $A_S$. For cycle 23 the relationship was:

$$A_G = (1.2299 \pm 0.0025) A_S - (10.16 \pm 1.35) \quad , \qquad \text{Eqn 11.3}$$

and for cycle 24:

$$A_G = (1.3389 \pm 0.0045) A_S - (8.52 \pm 2.28) \quad . \qquad \text{Eqn 11.4}$$

These relationships are statistically different from each other for the two cycles, more-so than in the case for sunspot number. Thus, it is suspected that the SRS sunspot area measurements are changing between cycles.

Comparisons of the McIntosh classification generated by GONG and those reported in the SRS reports were made. There were significant differences in the classifications, most likely due to lower spatial resolution of GONG. That is, GONG generally reported a lower complexity for a region than did SRS. Similarly, for magnetic regions, more mixing of polarity in regions ($\beta\gamma$ and $\gamma$ regions) was reported. This was most likely due to the criteria used in the GONG algorithm.

In the case of H$\alpha$ flare analysis, flare reports were examined from SOON sites to obtain the times of flares. The GONG H$\alpha$ images were processed in a time frame of 1 hour prior to the reported start of the flare to 1 hour after the reported end of the flare. When comparisons between the GONG sites were made, there were significant differences between both the flare areas and flare brightness. When comparisons were made between the Learmonth GONG and SOON reports, the correlation coefficient for both flare area and brightness were less than 0.5, showing that it is probably differences in equipment that causes variation, rather than seeing conditions at the sites.

An attempt was made to perform X-ray flare prediction using morphological parameters of regions reported both by SOON and GONG. In all cases there were significant differences in flare rates between cycles regardless of which region parameters were used. The exception was GONG magnetic inversion line complexity. Here the flare rate for M class or higher X-ray flares of a region was found to be reasonably approximated across solar cycles by the following relationship:

$$R_F = 2.016 \times 10^{-3} e^{0.914C} \quad , \qquad \text{Eqn 11.5}$$

where $R_F$ is the flare rate in flares per day and $C$ is the inversion line complexity.

This relationship worked best in cycle 23 and was poorer in cycle 24. Furthermore, in order to obtain this fit, regions with complexity greater than 5 were omitted. It may be

possible to improve this result by incorporating inversion line length or magnetic field strength to eliminate those regions where the inversion line is incorrectly assigned a large complexity. Inversion line complexity is not currently reported by SWPC.

Solar radio signatures were examined in order to attempt to find a correlation with Solar Particle Events (SPEs). It was found that for ground based observation metric type 3L events (type 3Lm) produced a 69% successful prediction rate. The worst predictor was the Castelli-U event with a 21% prediction rate. The Type 2/4 complex produced a 36% prediction rate.

Finally, an attempt to relate morphological features to Solar Particle Events (SPEs) was attempted. Unfortunately, no parameters could be found that could be used as a predictor, even for those events that produced Ground Level Enhancements (GLEs).

In summary, automated techniques for sunspot number and sunspot area produce a more reliable and consistent data set than current manual reduction techniques. Of course if instrumentation or software changes then these values will need to be recalculated or adjusted.

Whether automated techniques for region classification are an improvement over manual techniques is probably still a subject for discussion. These results seem to suggest that the results are at least consistent, but are possibly limited by the lower spatial resolution of the images used.

The limiting factor for Hα flare analysis seems to be the optical systems themselves, most likely to be variations in the filters from site-to-site.

**11.2 Recommendations**

1. Where possible, solar observing networks should have as close to identical equipment as possible. This is already the case for GONG, SOON and RSTN. Admittedly this is unlikely to be possible for organisations such as the American Association for Variable Star Observers (AAVSO) as they are often amateur observers.

2. Correction of the historical record due to errors in sunspot area and sunspot number is a difficult proposition due to innate variations that occur between sunspot cycles. It may

be preferable to keep existing records, as it is impossible to tell if natural solar variation is being corrected along with perceived errors.

3. The USAF SOON observers should modify their limb area correction techniques to allow correction factors greater than 3 for sunspot areas. The ellipse overlays should be replaced with a grid, which enables more consistent measurement of sunspot area.

4. Automated techniques should be employed in the observation and measurement of sunspot area and sunspot number. For GONG this was simpler due to the acquisition of both intensity and magnetogram images simultaneously, making it easier to determine which spots belong to each region. It should be possible however, to modify the techniques used here for a whitelight-only observing system.

5. Inversion line analysis should be given a greater priority than it currently has for prediction of solar x-ray flares. Of all the parameters analysed here, it was the best predictor.

### 11.3 Future Work

It is hoped to publish the automated GONG sunspot number and sunspot area records on the Australian Space Academy (ASA) website (www.spaceacademy.net.au) updated on a monthly basis. There is the possibility of whitelight and Hα telescopes being deployed at ASA and software will be developed to analyse this imagery based on the work presented in this thesis.

There is an occasional glitch in the Automated Sunspot Analysis Program that classifies sunspots as not having penumbra when they do. It is unclear at this stage where the error lies, but future work will be able to rectify this problem.

An attempt has been made to automatically analyse RSTN radio data, but this is not yet complete. The difficulty here is removing un-flagged calibration data from the data set. This should not be an insurmountable problem however, and will hopefully result in less errors in burst reports. In addition a simple 1415 MHz monitoring system has been proposed that uses Software Defined Radio (SDR), and a simple dipole. The purpose for this system is to monitor large radio bursts that would saturate the RSTN system in order to support the GNSS.

The sunspot analysis performed here will be repeated with the revised International Sunspot Number. The GONG automated analysis should be able to determine if the revised numbers are more or less consistent over the years for which both GONG and the old sunspot number series cover.

Finally, implementation of a similar method to Joshi et al. (2010) for eruptive filaments using GONG images will be attempted. Some preliminary results have been obtained but is not yet robust enough for reliable detection.

## 12. References

"Inside GNSS News", 17 June 2010, http://www.insidegnss.com/node/2138

Aarons, J. (1982). Global Morphology of Ionospheric Scintillations. IEEE Proceedings, 70, pp. 360-378.

Baranyi, T., Kiraly, S., Coffey, H. E. (2013). Indirect Comparison of Debrecen and Greenwich Daily Sums of Sunspot Area. *Monthly Notices of the Royal Astronomical Society*, 434(2), pp. 1713-1720.

Bastian, T. S., Benz, A. O. and Gary, D. E. (1998). Radio Emission from Solar Flares. *Annual Review of Astronomy and Astrophysics*, 36, pp. 131-188.

Bhatnagar, A. and Livingston, W. (2005). *Fundamentals of Solar Astronomy*. Singapore: World Scientific Publishing Company, pp. 94-97, 253-256, 401.

Bornmann, P. L. and Shaw, D. (1994). Flare Rate and the McIntosh Active-Region Classifications. *Solar Physics*, 150, pp. 127-146.

Brandt, P. N. and Schmidt, W. (1990). On the Umbra-Penumbra Area Ratio of Sunspots. *Solar Physics*, 129, pp. 191-194.

Bray, R. J. and Loughhead, R. E. (1964). *Sunspots.* Bath: Pitman Press, pp. 78-79, 227-228, 242-245.

Cane, H.V., Erickson, W. C. and Prestage, N. P. (2002). Solar flares, type III radio bursts, coronal mass ejections and energetic particles. *Journal of Geophysical Research*, 107 (A10), pp. 1315-1333.

Carrasco, V. M. s>, Lefevre, L. Vaquero, J. M. and Gallego, M. C. (2015). Equivalence Relations Between the Cortie and Zürich Sunspot Group Morphological Classifications. *Solar Physics*, 290 (5), pp. 1445-1455.

Castelli, J. P. and Guidice, D. A. (1976). Impact of Current Solar Radio Patrol Observations. *Vistas in Astronomy*, 19, pp. 355-383.

Chen, Z., Gao, Y. and Liu, Z. (2005). Evaluation of solar radio bursts' effect on GPS receiver signal tracking within International GPS Service network. *Radio Science*, 40.

Clete, F., Svalgaard, L., Vaquero, J. M. and Cliver, E. W. (2014). Revisiting the Sunspot Number: A 400-Year Perspective on the Solar Cycle. *Space Science Reviews*, 186 (1-4), pp. 35-103.

Cliver, E. W., McNamara, L. F. and Gentile, L. C. (1985). Peak Flux Density Spectra of Large Solar Radio Bursts and Proton Emisison from Flares. *Journal of Geophysical Research*, 90 (A7), pp. 6251 – 6266.

Colak, T. and Qahwaji, R. (2008). Automated McIntosh-Based Classification of Sunspot Groups Using MDI Images. *Solar Physics*, 248, pp. 277-296.

Cortie, A. L. (1901). On the Types of Sun-Spot Disturbances. *Astrophysical Journal*, 13, pp. 260-263.

Curto, J. J., Blanca, M. and Martínez E. (2008). Automatic Sunspots Detection on Full-Disk Solar Images Using Mathematical Morphology. *Solar Physics*, 250, pp. 411-429.

Dodson, H. W., Hedeman, E. R. and Covington. A. E. (1954). Solar Flares and Associated 2800MC/SEC (10.7 cm) Radiation. *Astrophysical Journal*, 119, pp. 541-563.

Dodson, H. W., Hedeman, E. R. and Mohler, O. C. (1974). Comparison of Activity in Solar Cycles 18, 19 and 20. *Reviews of Geophysics and Space Physics*, 12 (3), pp. 329-341.

Erickson, W. C. (1997). The Bruny Island Radio Spectrometer. *Publication of the Astronomical Society of Australia*, 14, pp 278-282.

Fermi, E. (1949). On the Origin of the Cosmic Radiation. *Physical Review*, 75(8), pp. 1169-1174.

Foukal, P. (2014). An Explanation of the Differences Between the Sunspot Area Scales of

the Royal Greenwich and Mt. Wilson Observatories, and the SOON Program. *Solar Physics*, 289, pp. 1517-1529.

Freund, J. E., (1988). *Modern Elementary Statistics*. 7th ed. New Jersey: Prentice-Hall, pp. 367-372.

Garcia, H. A. (1994). Temperature and Emission Measure from GOES Soft X-Ray Measurements. *Solar Physics*, 154, pp. 275-308.

Gargate, L., Fonsec, R. A., Silva, L. O., Bamford, R. A. and Bingham, R. (2014). SEP Acceleration in CME Driven Shocks Using a Hybrid Code. *The Astrophysical Journal*, 792 (1).

Giersch, O., Kennewell, J. (2008). Type 3L Radio Emission as a Predictor for Solar Particle Events. In: *Australian Institute of Physics 18th National Congress: Adelaide*. Australian Institute of Physics, pp. 235-238. Available at: http://www.aip.org.au/info/sites/default/files/Congress2008/AIPC2008/PDF/AUTHOR/AP081160.PDF [Accessed 9 Jun. 2016].

Giersch, O. and Kennewell, J. (2013). Solar Radio Interference to the GNSS. In: *Proceedings of the 12th Australian Space Science Conference: Melbourne.* pp. 141-150. Available at www.nssa.com.au/12assc/12assc-proceeding-web.pdf [Accessed 7 Dec 2016].

Giovanelli, R. G. (1939). The Relations Between Eruptions and Sunspots, *The Astrophysical Journal.* 89, pp. 555-567.

Goldsmith, A. (2005). *Wireless Communications*. Cambridge: Cambridge University Press (2005), pp. 403-451.

Graedel, T. E. and Lanzerotti, L. J. (1970). Interplanetary-Particle Associations with Type III Solar Bursts. *Journal of Geophysical Research,* 76(28), pp. 6932 – 6938.

Győri, L. (1997). Automation of Area Measurements of Sunspots. *Solar Physics*, 180, pp. 109-130.

Hale, G. E., Ellerman, F., Nicholson, S. B. and Joy, A. H. (1919). The Magnetic Polarity of Sun-Spots. *Astrophysical Journal*, 49, pp. 153-178.

Hale, G. E. (1924). The Spectrohelioscope. *Proceedings of the National Academy of Sciences of the United States of America*, 10 (8), pp. 361-363.

Hoyt, D. V. and Schatten, K. H. (1998). Group Sunspot Numbers: A New Solar Activity Reconstruction. *Solar Physics*, 181, pp. 491-512.

Hyder, C. L. (1967). A Phenomenological Model for Disparitions Brusques followed by Flarelike Chromospheric Brightenings, I: The Model, its Consequences, and Observations in Quiet Solar Regions. *Solar Physics*, 2(1), pp. 49-74.

Izenman, A. J. (1983). J. R. Wolf and H. A. Wolfer: An Historical Note on the Zurich Sunspot Relative Numbers. *Journal of the Royal Statistical Society*, 46 (3), pp. 311-318.

Jiggens, P., Chavy-Macdonald, M., Santin, G., Menicucci, A., Evans, H. and Hilgers, A. (2014). The Magnitude and Effects of Extreme Solar Particle Events. *Journal of Space Weather and Space Climate*, 4.

Joshi, A. D., Srivastava, N. and Mathew, S. K. (2010). Automated Detection of Filaments and Their Disappearance Using Full-Disc Hα Images. *Solar Physics*, 262(2), pp. 425-436.

Kildahl, K. J. N. (1980). Frequency of Class M and X flares by Sunspot Class (1969-1976). In: R. F. Donnelly, ed., *Solar Terrestrial Predictions Proceedings Vol 3: Solar Activity Predictions.*, 1st ed. Washington, D.C.: U.S. Government Printing Office, pp. 166-172

Kintner, P. M. Jr., O'Hanlon, B., Gary, D. E. and Kintner, P. M. S. (2009). Global Positioning System and Solar Radio Burst Forensics. *Radio Science*, 44.

Kirkup, L., (1994). *Experimental Methods*. Milton: John Wiley and Sons Australia, pp. 106-115.

Klobuchar, J. A., Kunches, J. M. and Van Dierendonck, A. J. (1999). Eye on the ionosphere: Potential solar radio burst effects on GPS signal to noise. *GPS Solutions*, 3(2), pp. 69-71.

Kraus, J. D. and Marhefka, R. J. (2002). *Antennas for all Applications*. 3rd Ed. New York: Mcgraw-Hill, pp. 23-27.

Kruger, A. (1979). *Introduction to Solar Radio Astronomy and Radio Physics.* Dordrecht: D. Reidel Publishing Company, pp. 113-116, 192-197, 215-216, 224-229.

Leblanc Y., Dulk, G. A., Vourlidas, A. and Bougeret, J. L. (2001). Tracing Shockwaves from the Corona to 1 AU: Type II Radio Emission and Relationship with CMEs. *Journal of Geophysical Research*, 106(A11), pp. 25301-25312.

Lee, K., Moon, Y. J., Lee, J., Lee, K. and Na, H. (2012). Solar Flare Occurence Rate and Probability in Terms of The Sunspot Classification Supplemented with Sunspot Area and Its Changes. *Solar Physics*, 281, pp. 639-650.

Li, G., Zank, G.P. and Rice, W. K. M. (2003). Acceleration and Transport of Energetic Particles at CME-Driven Shocks. *Advanced Space Research*, 32(12), pp. 2597-2602.

Lyne, A. and Graham-Smith, F. (2006). *Pulsar Astronomy*, Cambridge: Cambridge University Press, pp. 226-227.

MacDowall, R. J., Lara, A., Manoharan, P. K., Nitta, N. V., Rosas, A. M. and Bougeret, J. L. (2003). Long-Duration Hectometric Type III Radio Bursts and their Association with Solar Energetic Particle (SEP) Events. *Geophysical Research Letters*, 30 (12), pp. 8018-8021.

Masson, S., Antiochos, S. K. and DeVore, C. R. (2013). A Model for the Escape of Solar-Flare-Accelerated Particles. *Astrophysical Journal*, 771 (2), pp. 82-96.

McIntosh, P. S. (1990). The Classification of Susnpot Groups. *Solar Physics*, 125, pp. 251-267.

Newman, M. E. J. (2005). Power Laws, Pareto Distributions and Zipf's Law. *Contemporary Physics*, 46 (5), pp. 323-351

Nita, G. M., Gary, D. E., Lanzerotti, L. J. and Thomson, D. J. (2002). The Peak Flux Distribution of Solar Radio Bursts. *The Astrophysical Journal*, 570, pp. 423-438.

Nitta, N. V., Liu, Y., DeRosa, M. L. and Nightingale, R. W. (2012). What are Special About Ground Level Events? Flares, CME's Active Regions and Magnetic Field Connection. *Space Science Reviews*, 171, pp. 61-83.

Park, J., Moon Y.-J., Gopalswamy, N. (2012). Dependence of Solar Proton Events On Their Associated Activities: Coronal Mass Ejection Parameters. *Journal of Geophysical Research*, 117.

Paul, H. E., (1974). Building a Birefringent Polarizing Monochromator for Solar Prominences. In: A. G. Ingalls, ed., *Amateur Telescope Making Volume 3*, 1st ed. Kingsport: Kingsport Press, pp. 376-412.

Pettauer, T., and Brandt, P. N. (1997). On Novel Methods to Determine Area of Sunspots from Photoheliograms. *Solar Physics*, 175, pp. 197-203.

Phillips, K. J. H. (1992), *Guide to the Sun.* Cambridge: Cambridge University Press, p. 203.

Reid, J. H. (1963). Classification of Solar Flares. *The Irish Astronomical Journal*, 6 (2), pp. 45-51.

Reid, J. H. and Vorhaben, K. H. (1971). The Videometer – An Instrument for Quantitative Measures on Solar Flares. In: *Automation in Optical Astrophysics.* Edinburgh: Proceedings of IAU Colloq. 11, pp. 168-172.

Robbins, D. E. and Reid, J. H. (1969). Solar Physics at the NASA Manned Spacecraft Center. *Solar Physics*, 10, pp. 502-510.

Sarychev, A. P. and Roshchina, E. M. (2006). Total Sunspot Area as a Solar Activity

Index. *Solar System Research*, 40 (6), pp. 521-526.

Sawyer, C. and Warwick, J. W. (1986). Solar Flare Prediction. Colorado: Colorado University Press, pp 115-117.

Serway, R. A., Moses, C. and Moyer C. A. (1997). *Modern Physics*. 2nd ed. Fort Worth: Saunders College Publishing, p. 16.

Schrijver, C. J. (2007). A Characteristic Magnetic Field Pattern Associated with all Major Solar Flares and Its Use in Flare Forecasting. *Astrophysical Journal*, 655, pp L117-L120.

Shapley, A.H. (1949). Reduction of Sunspot Number Observations. *Publications of the Astronomical Society of the Pacific*, 61 (358), pp. 13-21.

Smart, D.F. and Shea, M.A. (1979). PPS76 – A Computerized 'Event Mode' Solar Proton Forecasting Technique. In: NOAA *Solar-Terrestrial Predictions Proceedings 1*. Boulder, NOAA, pp. 406-427.

Shea, M. A. and Smart, D. F. (2012). Space Weather and Ground Level Solar Proton Events of the 23rd Solar Cycle. *Space Science Reviews*, 171, pp. 161-188.

Smith, A. J. (1967). Radio Exploration of the Sun. New Jersey: D. Van Nostrand Company Inc., pp 9-12, 62-74, 86-87, 99-101.

Solanki, S. K. (2003). Sunspots: An Overview. *The Astronomy and Astrophysics Review*, 11**,** pp. 153-286

Tapping, K. F. and Morton, D. C. (2013). The Next Generation of Canadian Solar Flux Monitoring. *Journal of Physics: Conference Series*, 440.

Temmer, M. Veronig, A., Hanslmeier, A., Otruba, W. and Messerotti, M. (2001). Statistical Analysis of Solar Ha Flares. *Astronomy and Astrophysics*, 375, pp. 1049-1061

Thompson, A. R., Moran, J. M., and Swenson, G. W. Jr. (2001). *Interferometry and Synthesis in Radio Astronomy*. New York: John Wiley and Sons, Inc., pp. 18-19.

Tian, L. and Alexander, D. (2006). Role of Sunspot and Sunspot Group-Rotation in Driving Sigmoidal Active Region Eruptions. Solar Physics. 233, pp 29.-43.

USAF AFWAMAN 15-1. Available at: http://static.e-publishing.af.mil/production/1/afwa/publication/afwaman15-1/afwaman15-1.pdf [Accessed 19 Dec. 2016].

USAF AFWAMAN 15-1. Available at: http://static.e-publishing.af.mil/production/1/afwa/publication/afwaman15-1/afwaman15-1.pdf [Accessed 23 Dec. 2016].

Vaquero, J. M. Gordillo, A., Gallego, M. C., Sanchez-Bajo, F. and Garcia, J. A. (2005) The umbra-penumbra area ratio of sunspots from the de la Rue data", *The Observatory*, 125, pp. 152-125.

Wanninger, L. (1993). Effects of Equatorial Ionosphere on GPS. *GPS World*, 4, pp. 48-54.

Warwick, C. S. (1963). Flare Importance Ratings – Some Hope For Improvement. The Physics of Solar Flares. In *Proceedings of the AAS/NASA Symposium*, Goddard Space Flight Center, pp. 27-28.

Wheatland, M. S. (2005). Initial Test of a Bayesian Approach to Solar Flare Prediction. *Publications of the Astronomical Society of Australia*, 22(2), pp. 153-156.

Wild, J. P., Smerd, S. F. and Weiss, A. A. (1963). Solar Bursts. *Annual Review of Astronomy and Astrophysics*, 1, pp. 291-366.

Wilson, R.M. and Hathaway, D.H. (2005). A Comparison of Rome Observatory Sunspot Area and Sunspot Number Determinations With International Measures, 1958-1998. *NASA Technical Paper*, TP-2005-214191.

Wilson, R.M. and Hathaway, D.H. (2006). On the Relation Between Sunspot Area and Sunspot Number, *NASA Technical Paper*, TP-2006-214324.

Zirin, H. (1988). *Astrophysics of the Sun*. Cambridge: Cambridge University Press, pp. 303, 327-228.

Zubair, M., Haider, Z., Khan, S. A., and Nasir, J. (2011). Atmospheric Influences on Satellite Communications. *Prseglad Elektrotechniczny (Electrical review)*, R87(5), pp. 261-264.

This work utilises data from the National Solar Observatory Integrated Synoptic Program, which is operated by the Association of Universities for Research in Astronomy, under a cooperative agreement with the National Science Foundation and with additional financial support from the National Oceanic and Atmospheric Administration, the National Aeronautics and Space Administration, and the United States Air Force. The GONG network of instruments is hosted by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias, and Cerro Tololo Interamerican Observatory.

# Appendix A: RSTN Site Radio Burst Comparison Plots

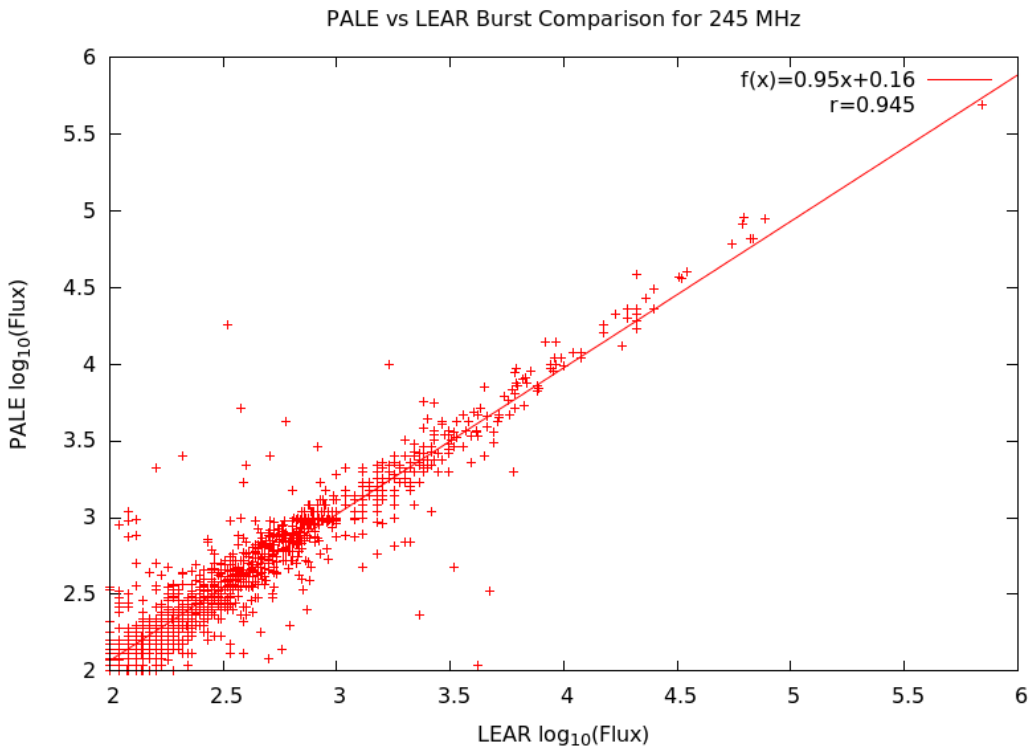



**Figure A.1**: Palehua and Learmonth Burst Peak Flux Comparison for 245 MHz

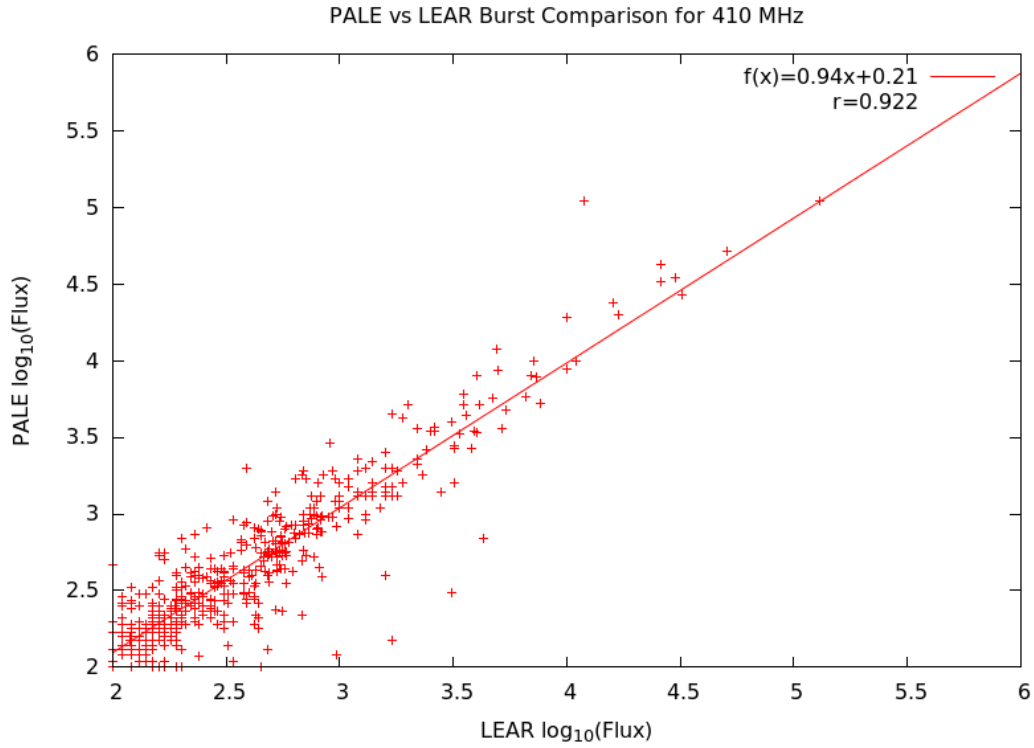



**Figure A.2**: Palehua and Learmonth Burst Peak Flux Comparison for 410 MHz

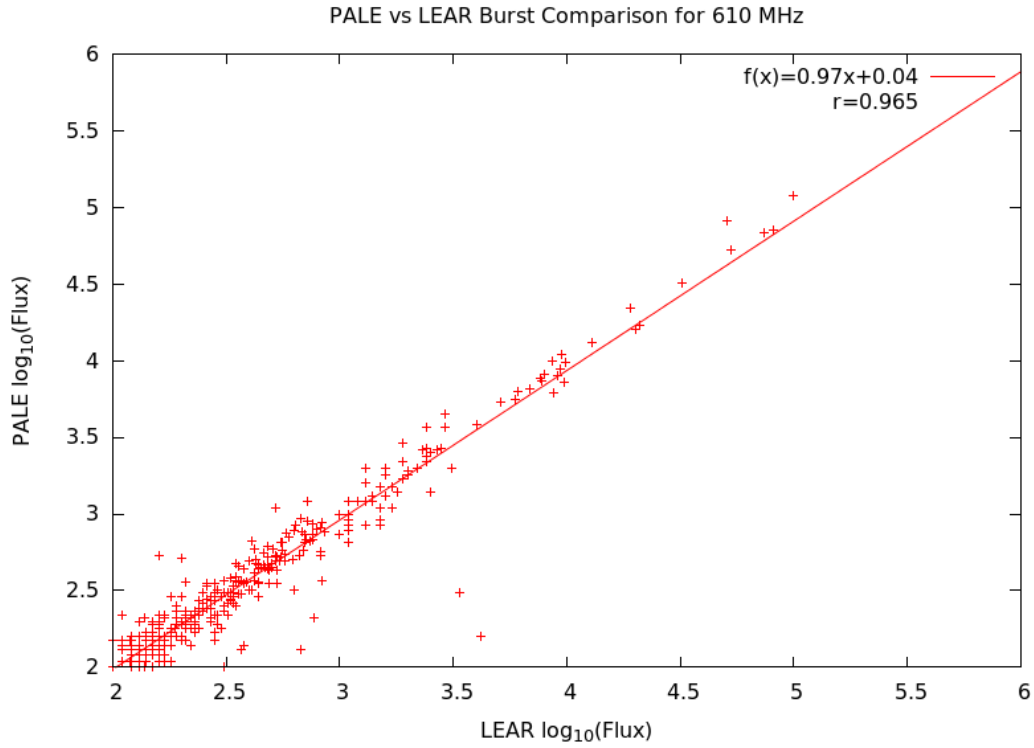



**Figure A.3**: Palehua and Learmonth Burst Peak Flux Comparison for 610 MHz

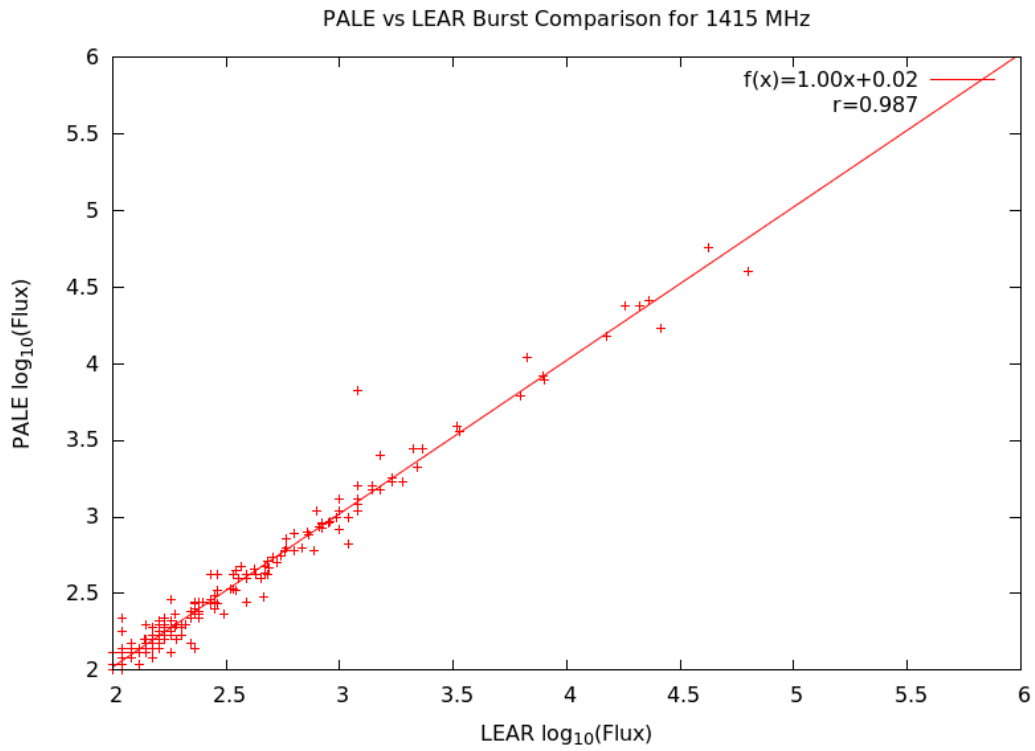



**Figure A.4**: Palehua and Learmonth Burst Peak Flux Comparison for 1415 MHz

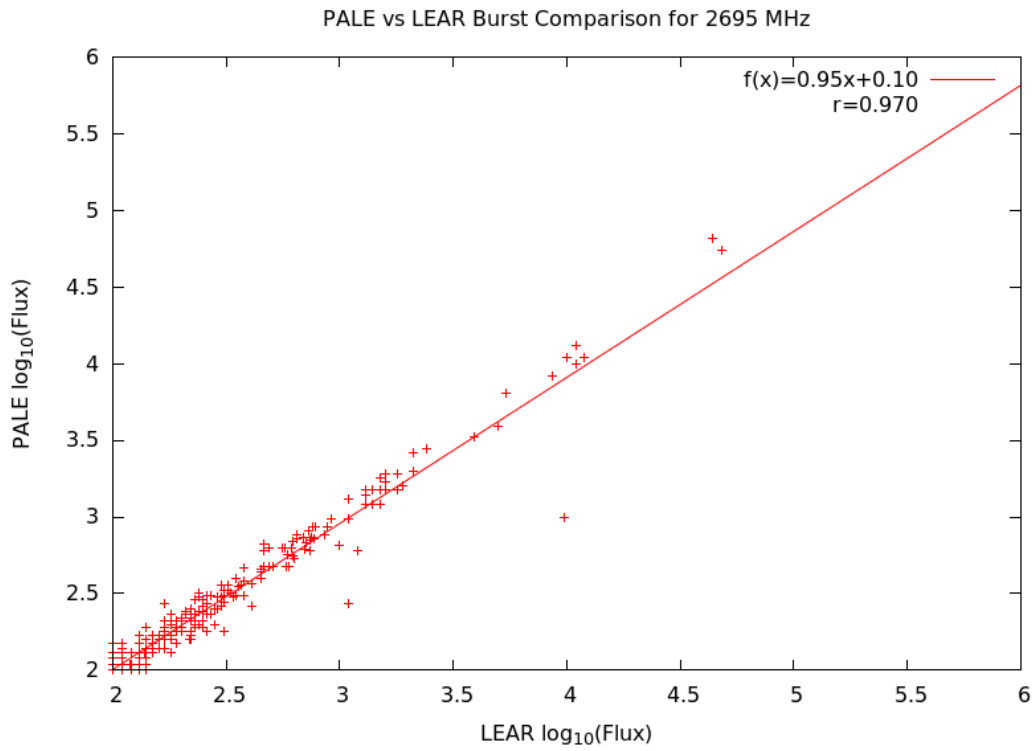



**Figure A.5**: Palehua and Learmonth Burst Peak Flux Comparison for 2695 MHz

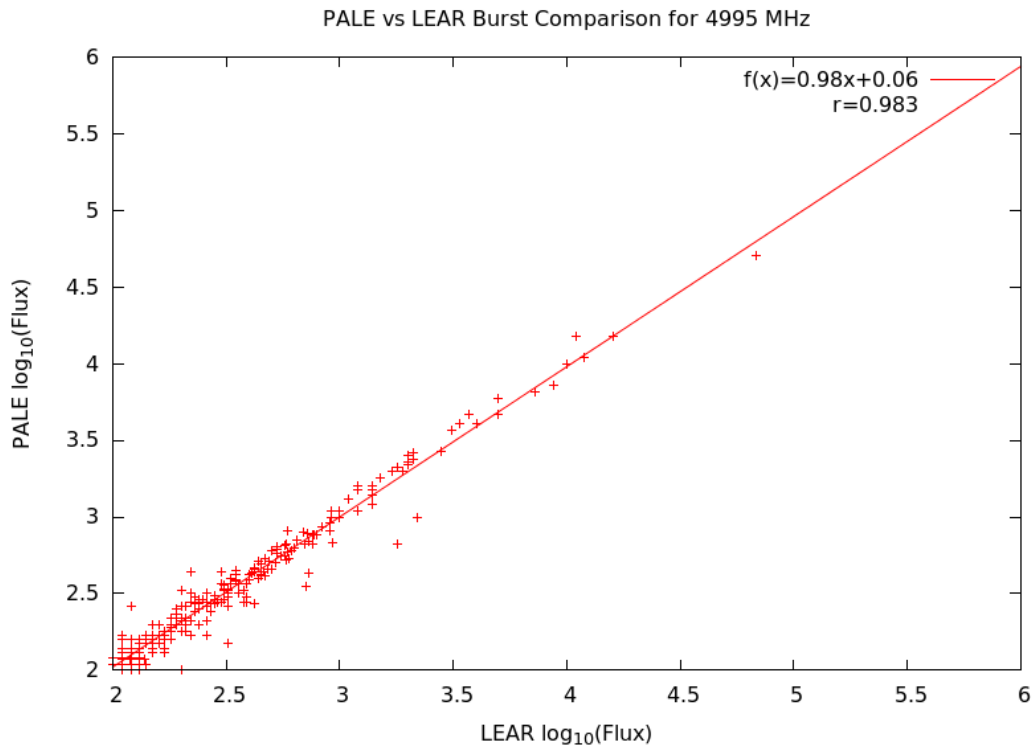



**Figure A.6**: Palehua and Learmonth Burst Peak Flux Comparison for 4995 MHz

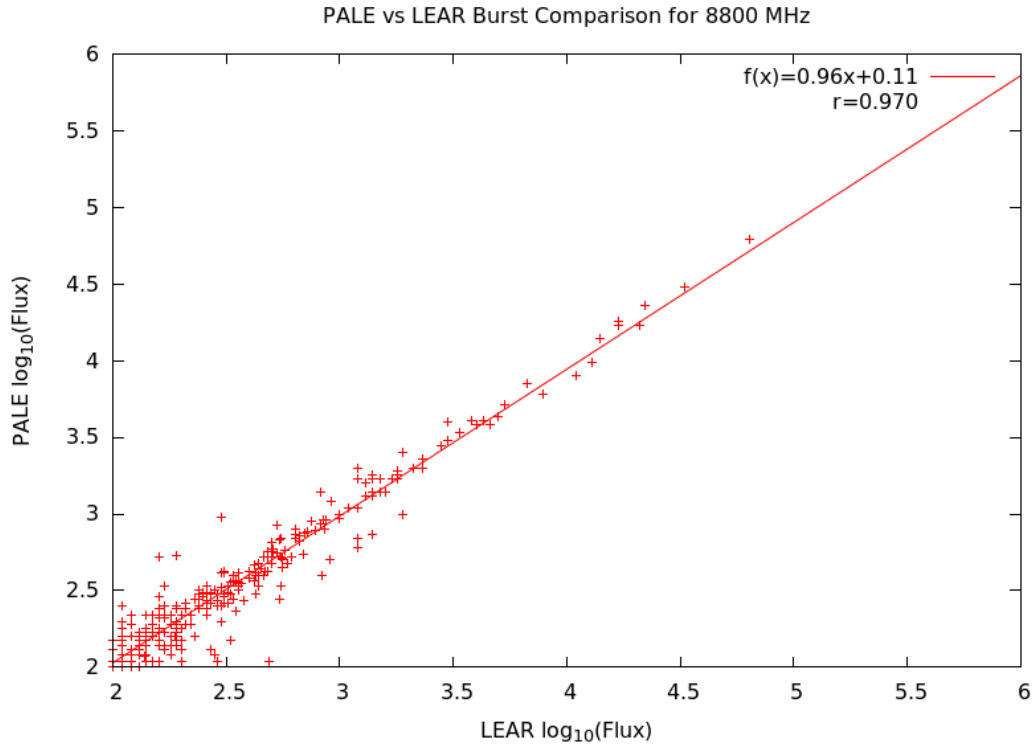



**Figure A.7**: Palehua and Learmonth Burst Peak Flux Comparison for 8800 MHz

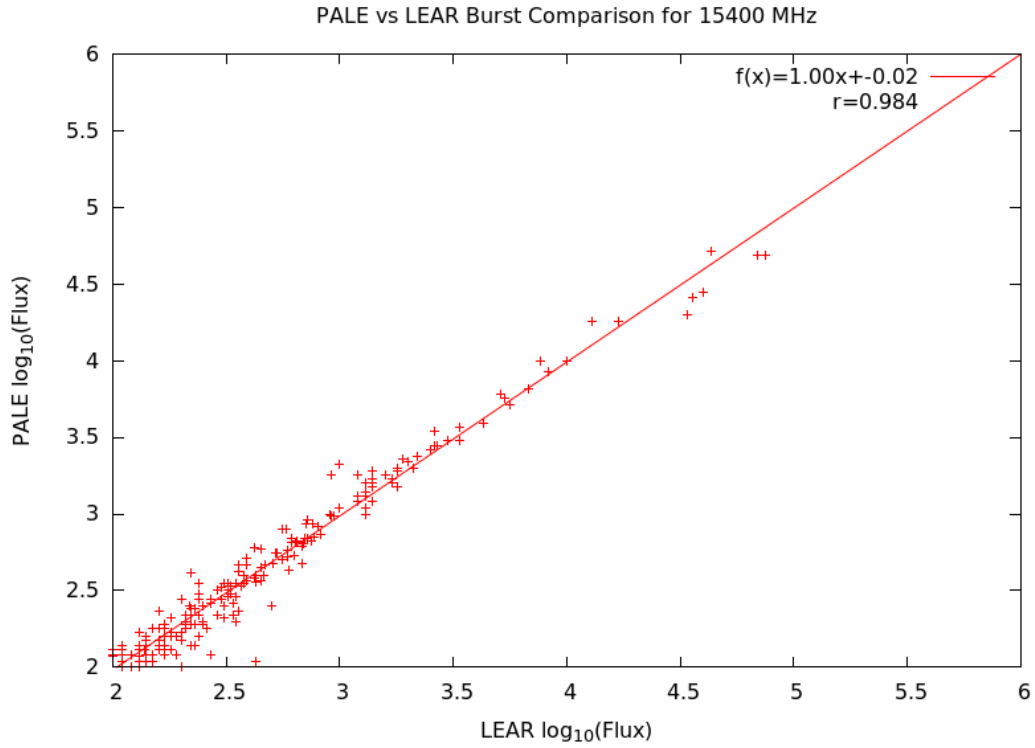



**Figure A.8**: Palehua and Learmonth Burst Peak Flux Comparison for 15400 MHz

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1296 | 0.954 | 0.009 | 0.161 | 0.024 | 0.945 |
| 410 | 523 | 0.944 | 0.017 | 0.207 | 0.046 | 0.922 |
| 610 | 360 | 0.974 | 0.014 | 0.038 | 0.038 | 0.965 |
| 1415 | 187 | 1.000 | 0.012 | 0.022 | 0.032 | 0.987 |
| 2695 | 214 | 0.952 | 0.017 | 0.101 | 0.043 | 0.970 |
| 4995 | 259 | 0.979 | 0.011 | 0.063 | 0.029 | 0.983 |
| 8800 | 284 | 0.957 | 0.014 | 0.115 | 0.037 | 0.970 |
| 15400 | 215 | 1.004 | 0.012 | -0.024 | 0.034 | 0.984 |

**Table A.1**: Palehua vs Learmonth Burst Peak Flux Comparison Regression Analysis

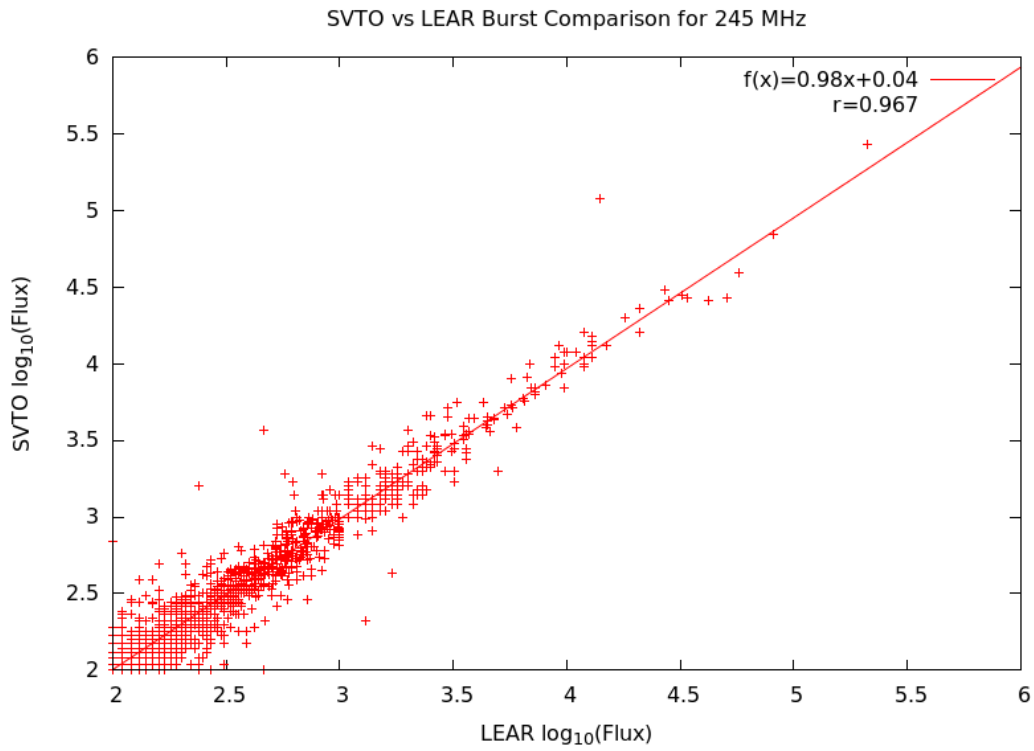



**Figure A.9**: San Vito and Learmonth Burst Peak Flux Comparison for 245 MHz

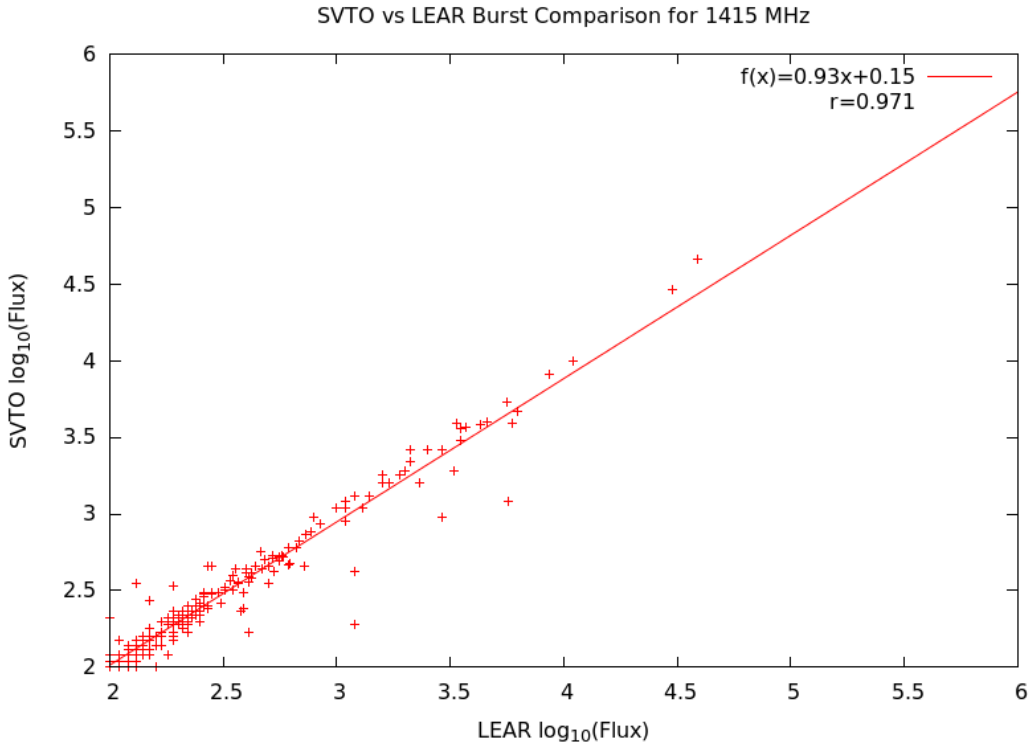



**Figure A.10**: San Vito and Learmonth Burst Peak Flux Comparison for 1415 MHz

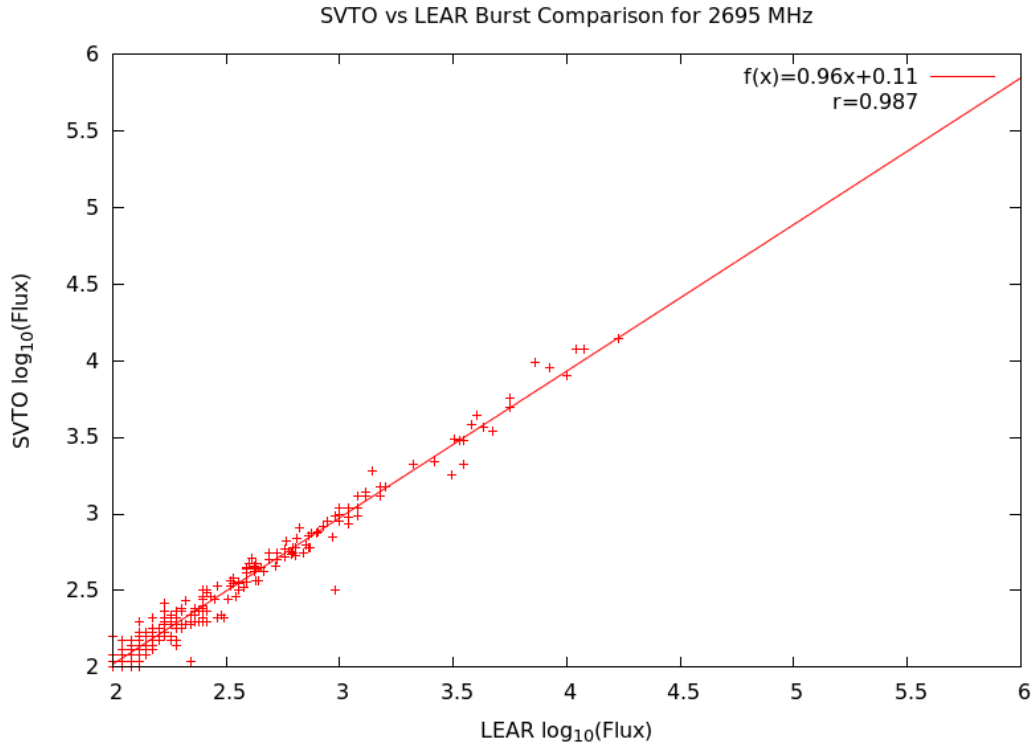



**Figure A.11**: San Vito and Learmonth Burst Peak Flux Comparison for 2695 MHz

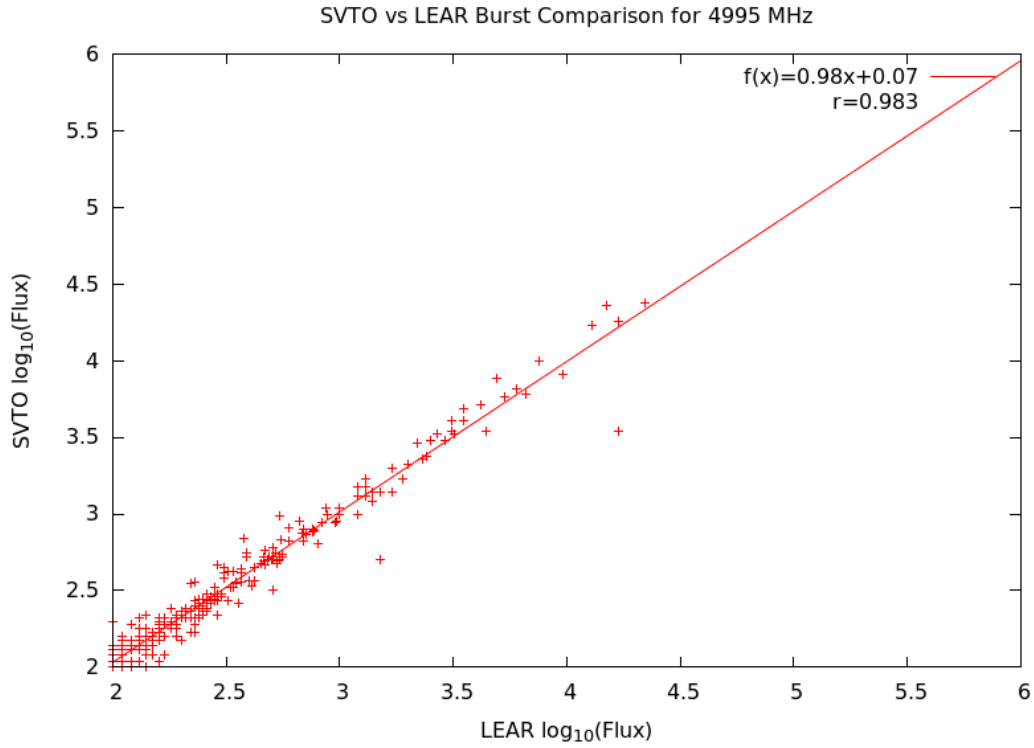



**Figure A.12**: San Vito and Learmonth Burst Peak Flux Comparison for 4995 MHz

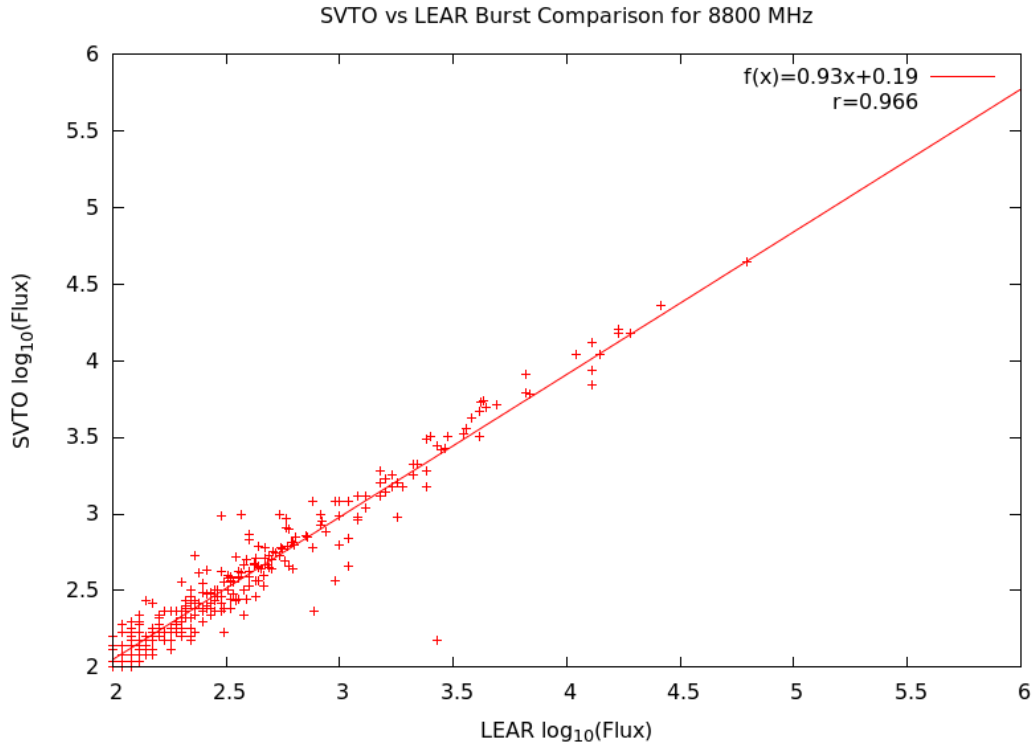



**Figure A.13**: San Vito and Learmonth Burst Peak Flux Comparison for 8800 MHz

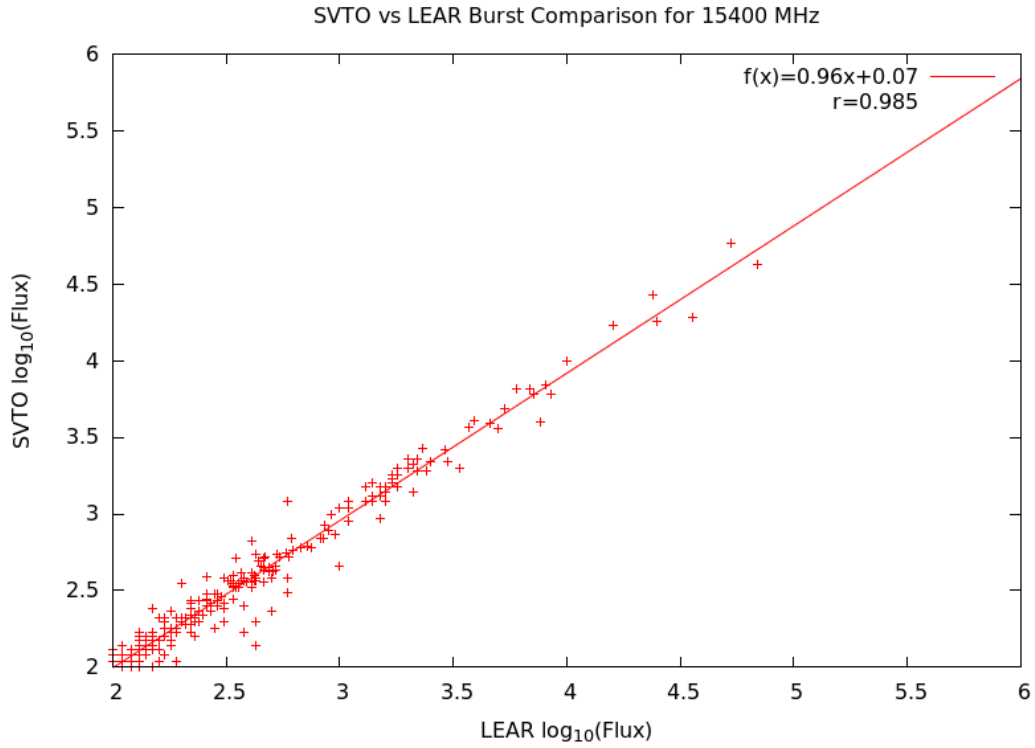



**Figure A.14**: San Vito and Learmonth Burst Peak Flux Comparison for 15400 MHz

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1394 | 0.982 | 0.007 | 0.041 | 0.018 | 0.967 |
| 1415 | 192 | 0.934 | 0.017 | 0.146 | 0.044 | 0.971 |
| 2695 | 214 | 0.956 | 0.011 | 0.109 | 0.027 | 0.987 |
| 4995 | 263 | 0.981 | 0.011 | 0.072 | 0.029 | 0.983 |
| 8800 | 297 | 0.930 | 0.015 | 0.191 | 0.038 | 0.966 |
| 15400 | 233 | 0.961 | 0.011 | 0.075 | 0.030 | 0.985 |

**Table A.2**: San Vito vs Learmonth Burst Comparison Regression Analysis

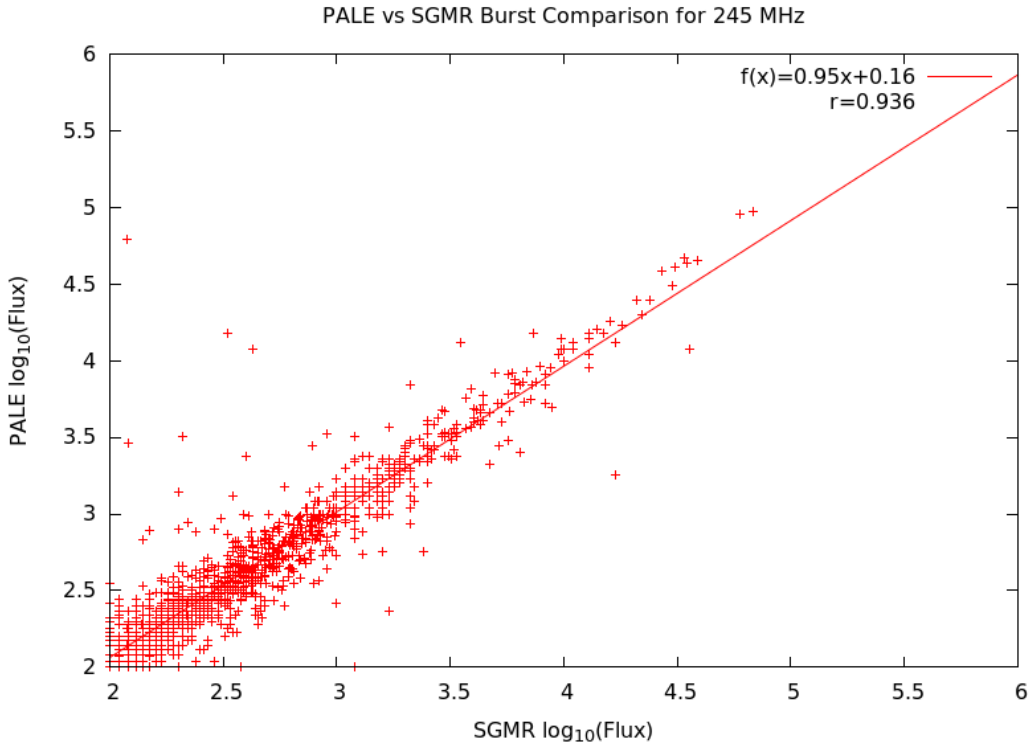



**Figure A.15**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 245 MHz

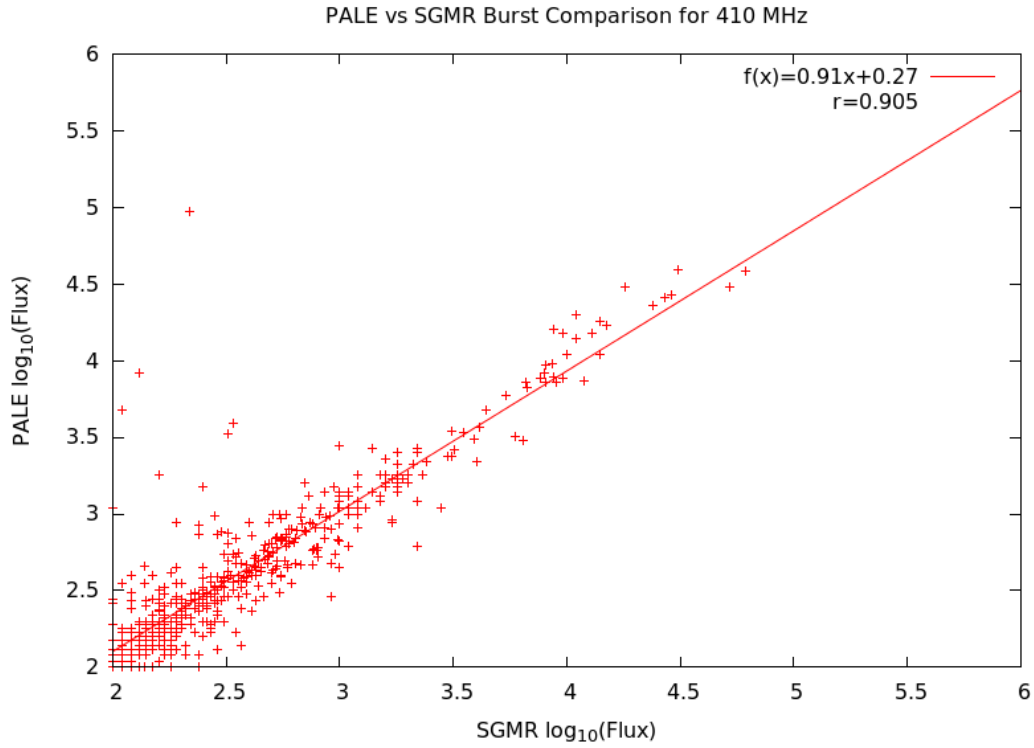



**Figure A.16**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 410 MHz

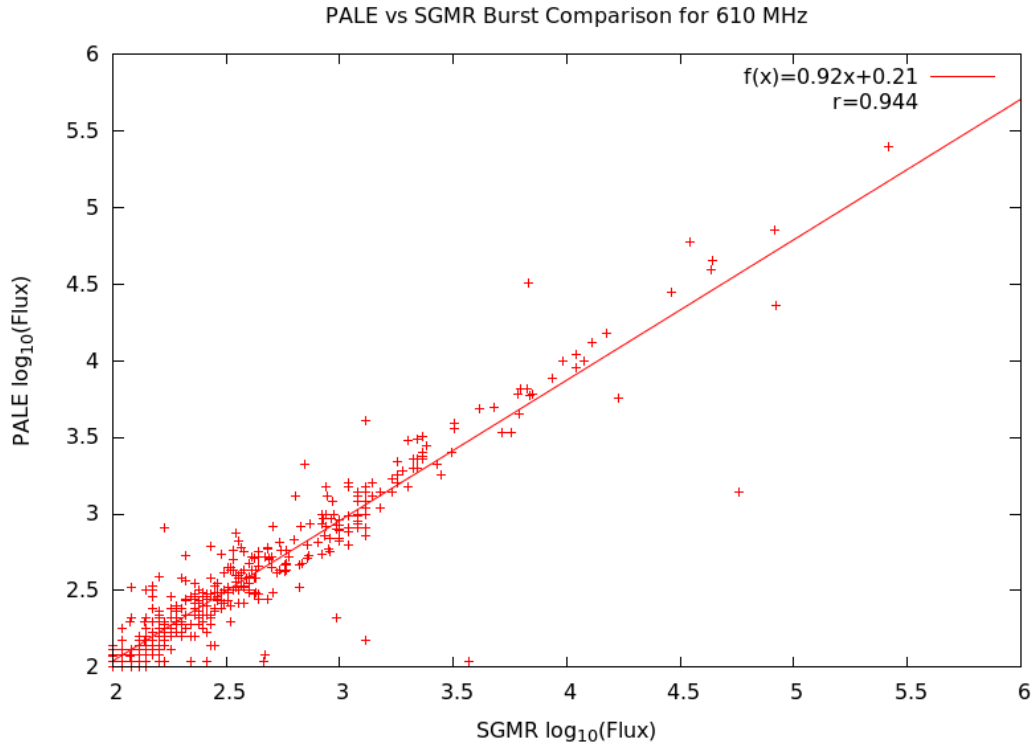



**Figure A.17**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 610 MHz

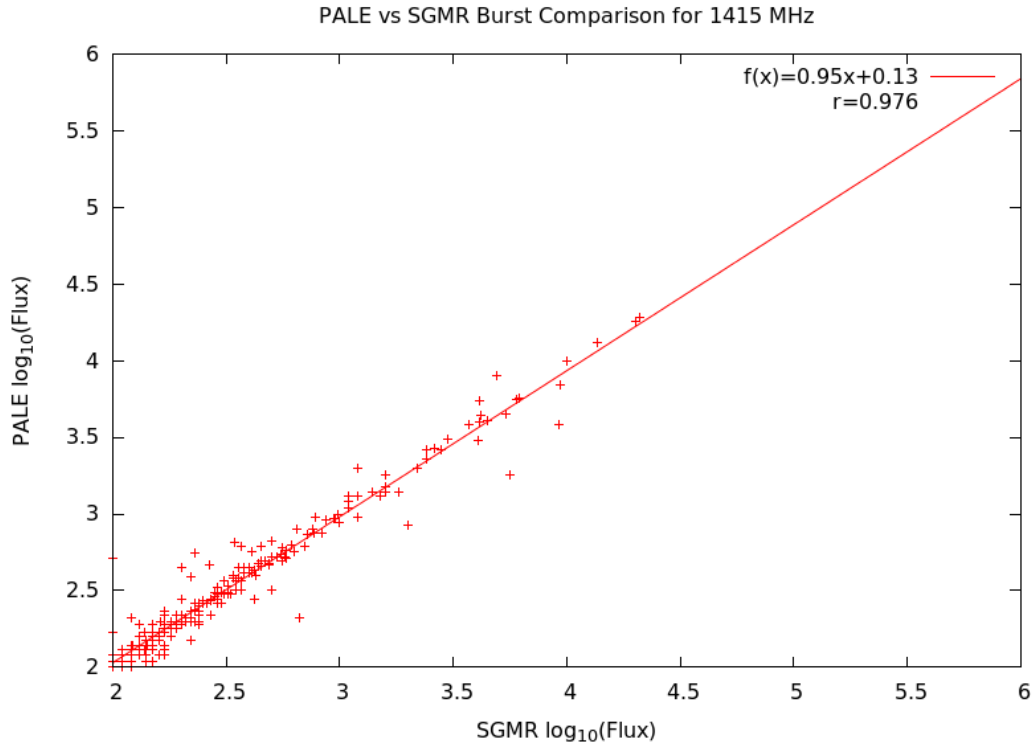



**Figure A.18**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 1415 MHz

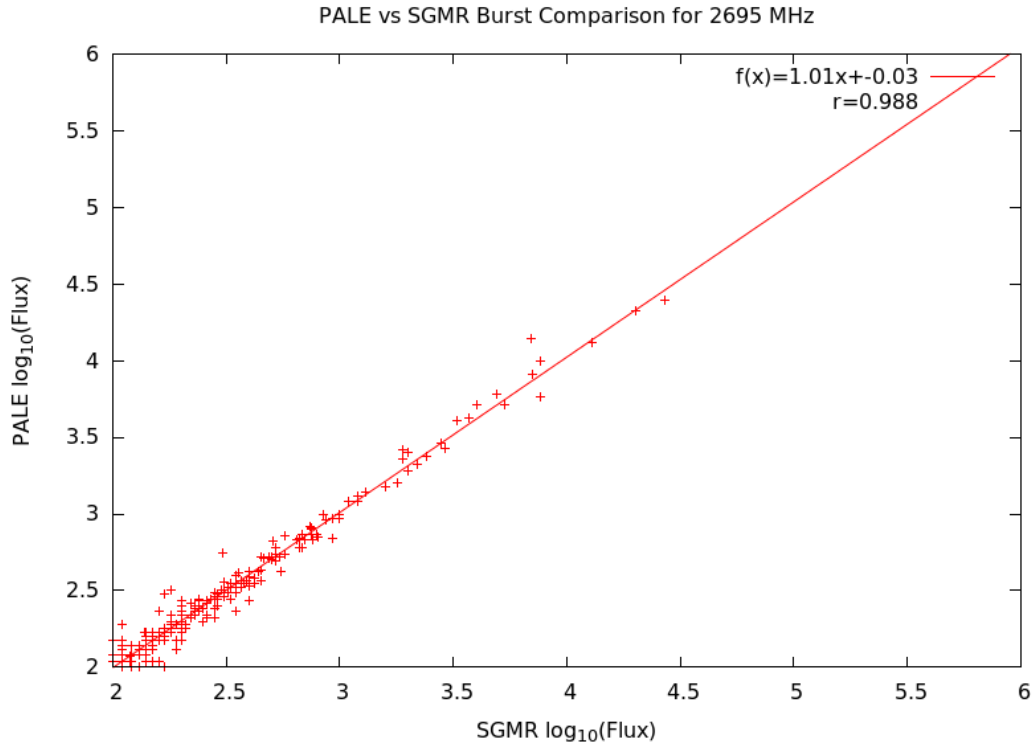



**Figure A.19**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 2695 MHz

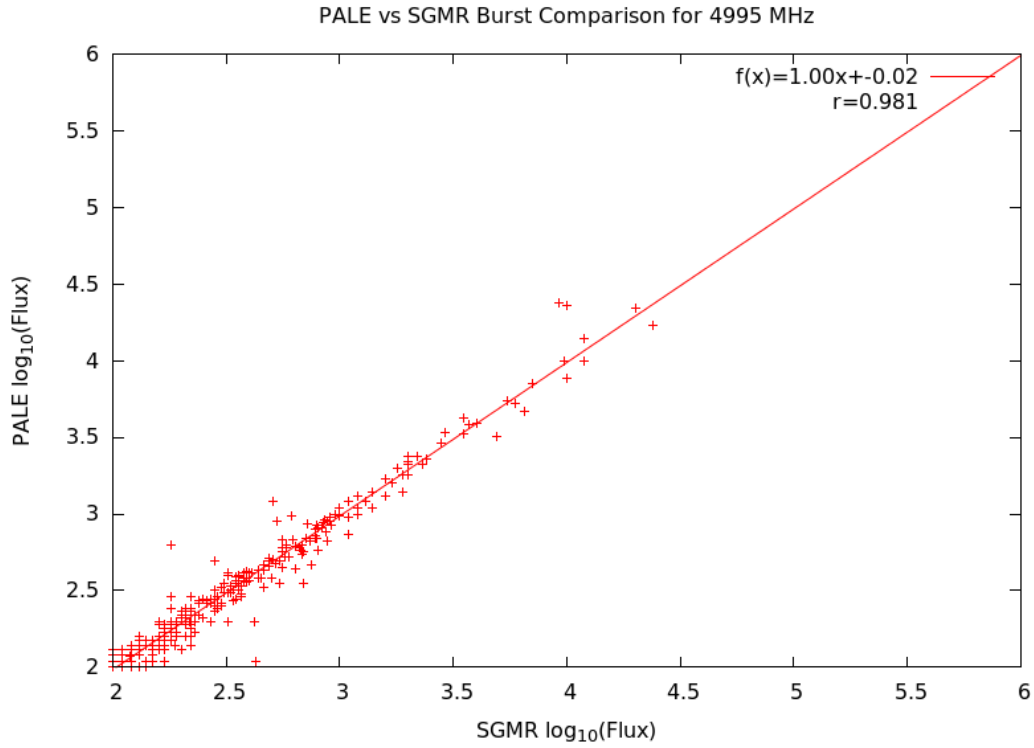



**Figure A.20**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 4995 MHz

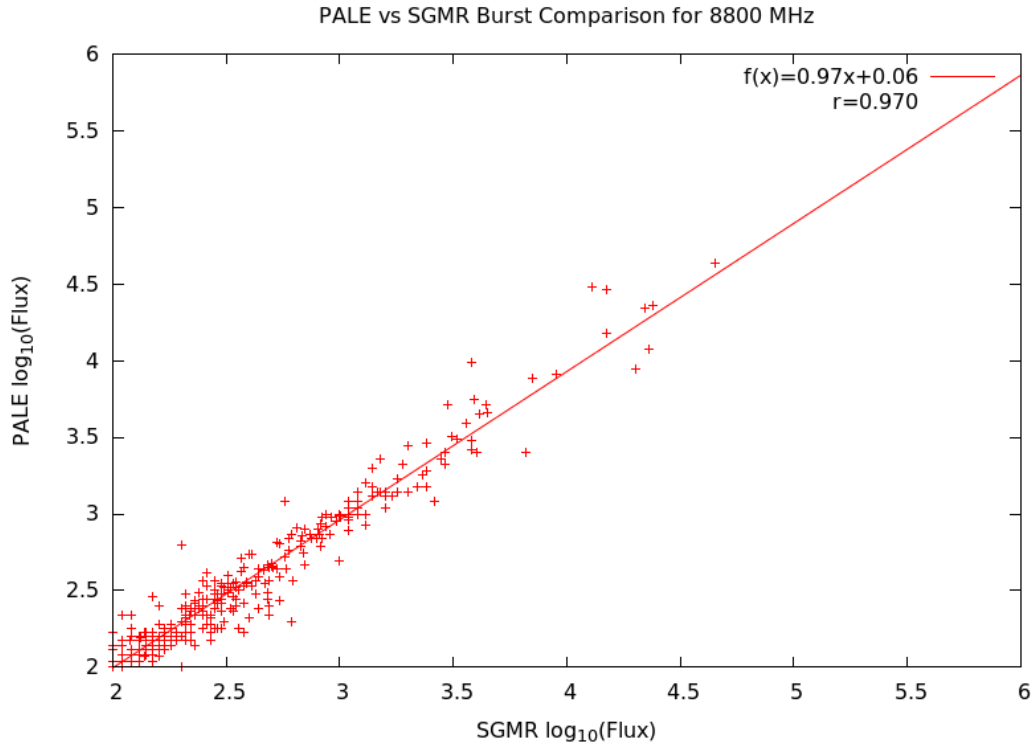



**Figure A.21**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 8800 MHz

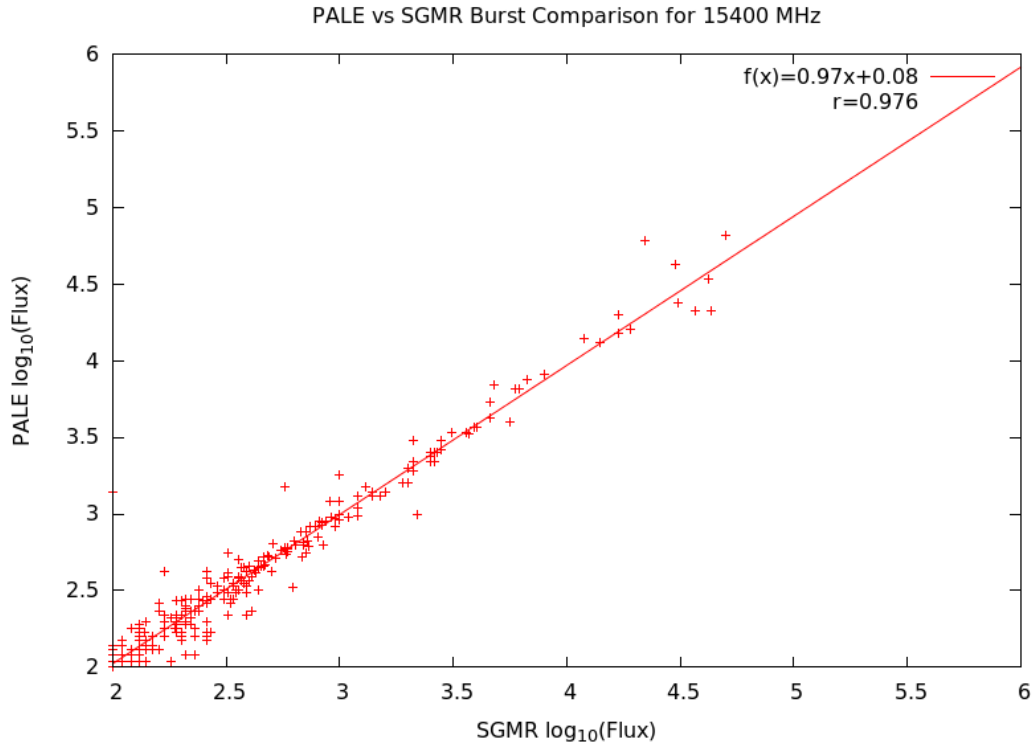



**Figure A.22**: Palehua and Sagamore Hill Burst Peak Flux Comparison for 15400 MHz

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1353 | 0.954 | 0.008 | 0.151 | 0.021 | 0.955 |
| 1415 | 199 | 0.932 | 0.026 | 0.169 | 0.068 | 0.931 |
| 2695 | 191 | 0.965 | 0.018 | 0.075 | 0.045 | 0.969 |
| 4995 | 252 | 0.963 | 0.010 | 0.080 | 0.026 | 0.986 |
| 8800 | 276 | 0.948 | 0.015 | 0.136 | 0.039 | 0.967 |
| 15400 | 237 | 1.000 | 0.012 | 0.008 | 0.030 | 0.984 |

**Table A.3**: Sagamore Hill vs San Vito Burst Comparison Regression Analysis

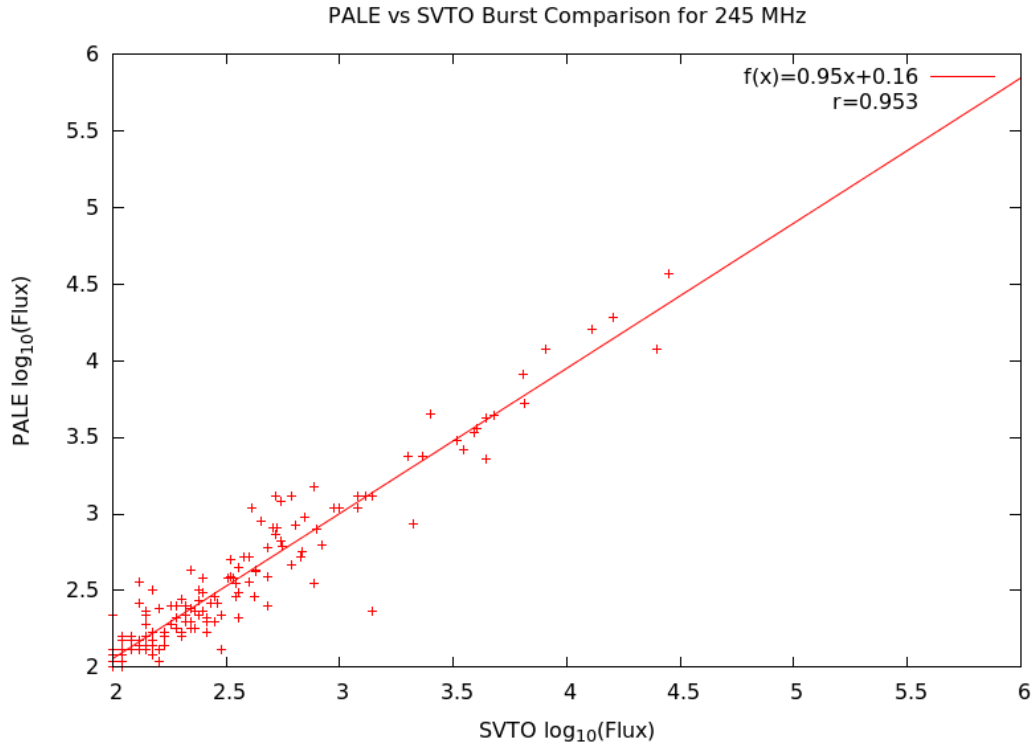



**Figure A.23**: Palehua and San Vito Burst Peak Flux Comparison for 245 MHz

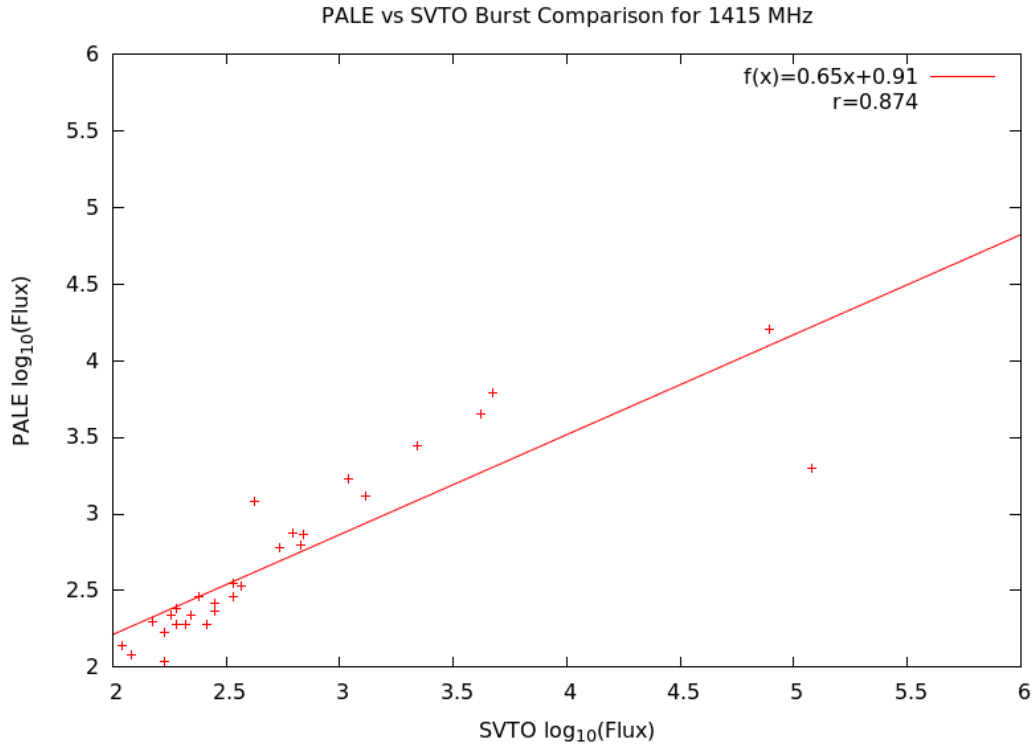



**Figure A.24**: Palehua and San Vito Burst Peak Flux Comparison for 1415 MHz

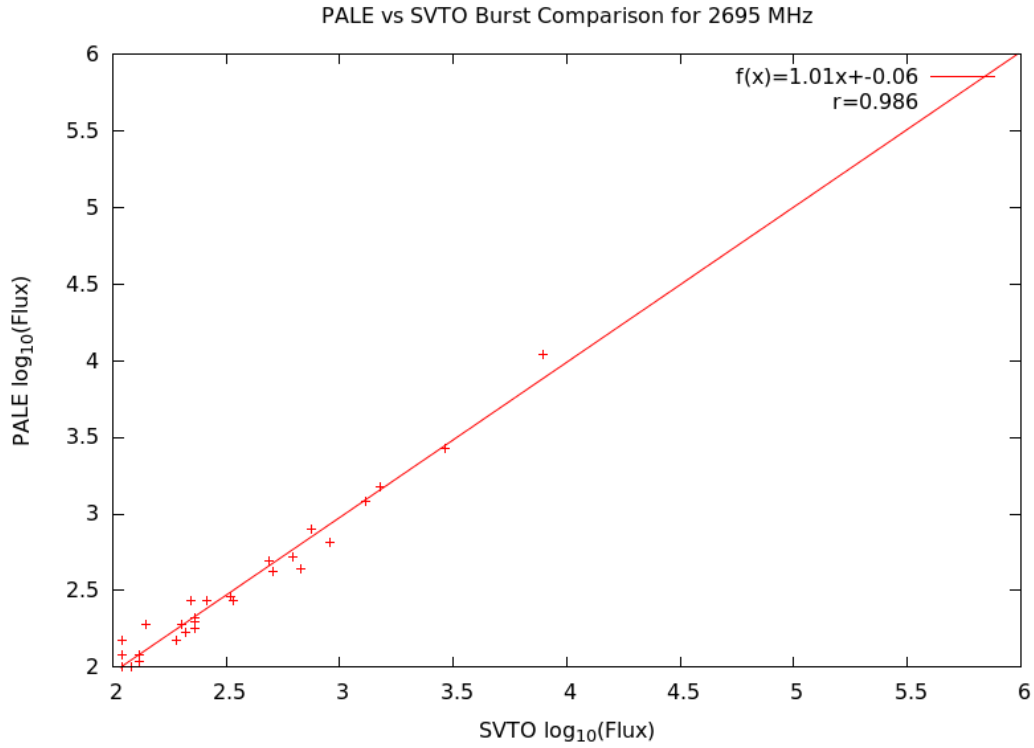



**Figure A.25**: Palehua and San Vito Burst Peak Flux Comparison for 2695 MHz

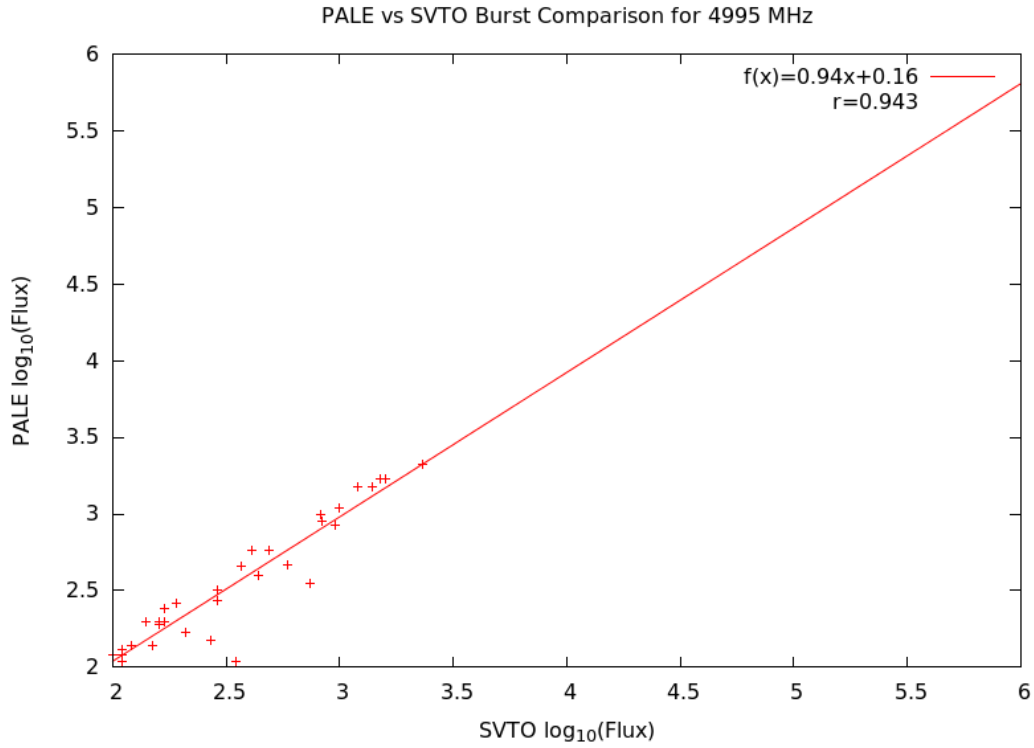



**Figure A.26**: Palehua and San Vito Burst Peak Flux Comparison for 4995 MHz

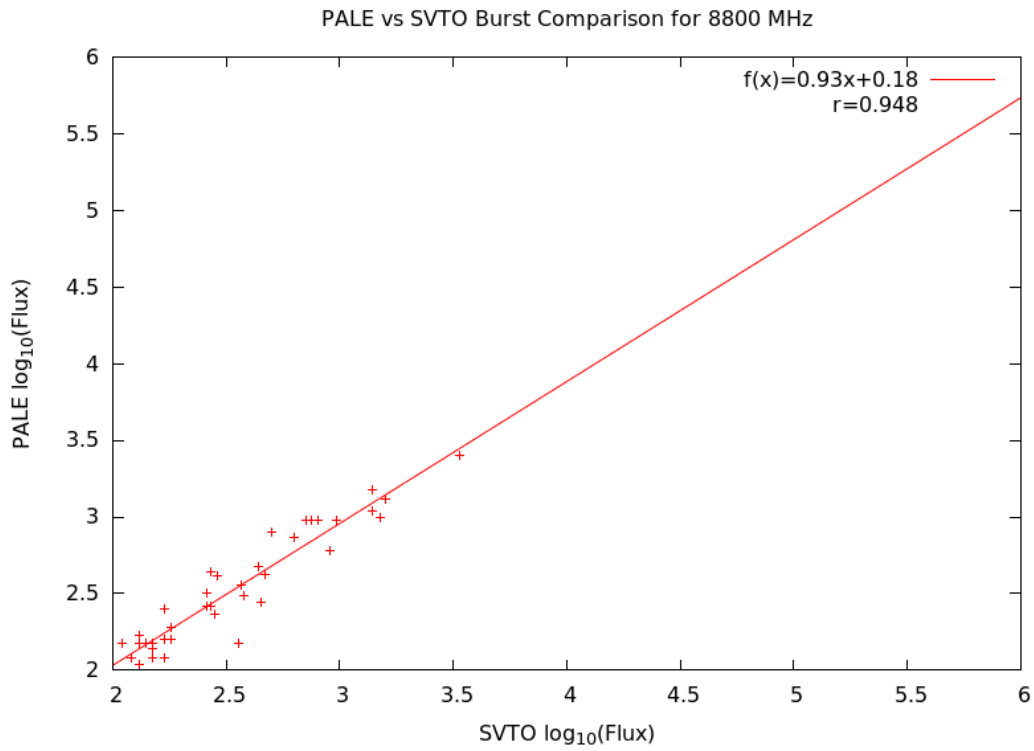



**Figure A.27**: Palehua and San Vito Burst Peak Flux Comparison for 8800 MHz

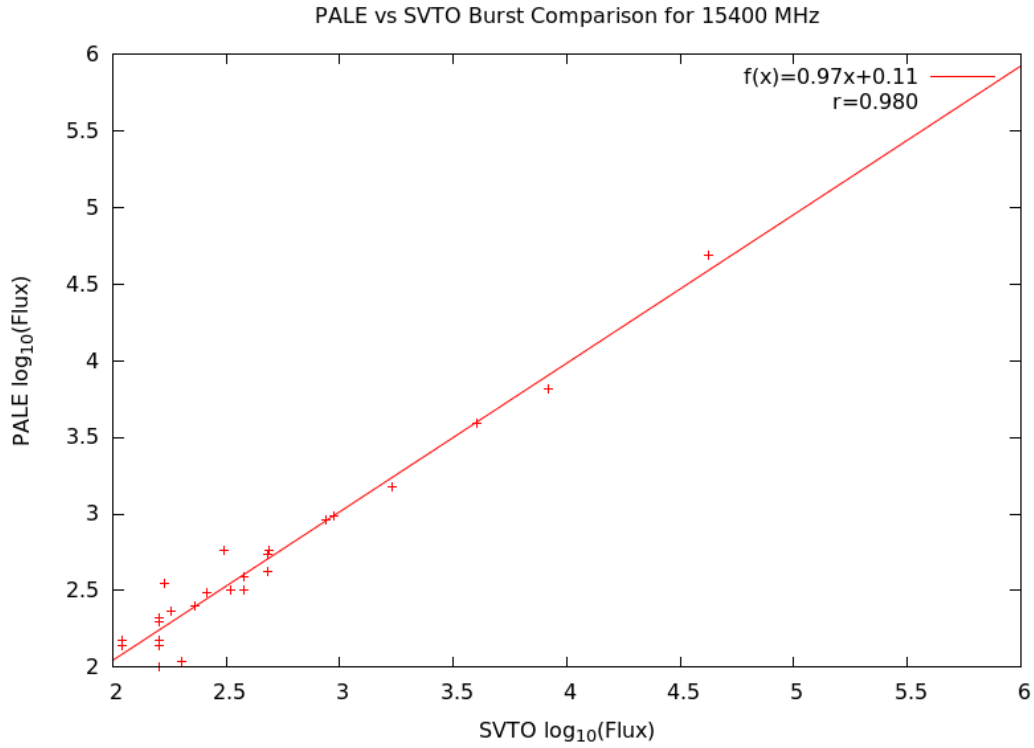



**Figure A.28**: Palehua and San Vito Burst Peak Flux Comparison for 15400 MHz

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 137 | 0.947 | 0.026 | 0.164 | 0.069 | 0.953 |
| 1415 | 29 | 0.652[1] | 0.070 | 0.910 | 0.200 | 0.874 |
| 2695 | 28 | 1.012 | 0.033 | -0.059 | 0.087 | 0.986 |
| 4995 | 33 | 0.942 | 0.060 | 0.157 | 0.153 | 0.943 |
| 8800 | 38 | 0.925 | 0.052 | 0.182 | 0.134 | 0.948 |
| 15400 | 25 | 0.969 | 0.041 | 0.108 | 0.113 | 0.980 |

**Table A.4**: Palehua vs San Vito Burst Comparison Regression Analysis

1. The gradient for 1415 MHz deviates significantly from 1, more-so than any other site relationship. There is however an outlier, and due to the small number of points due to the time zones that Palehua and San Vito are in, this is distorting the relationship.

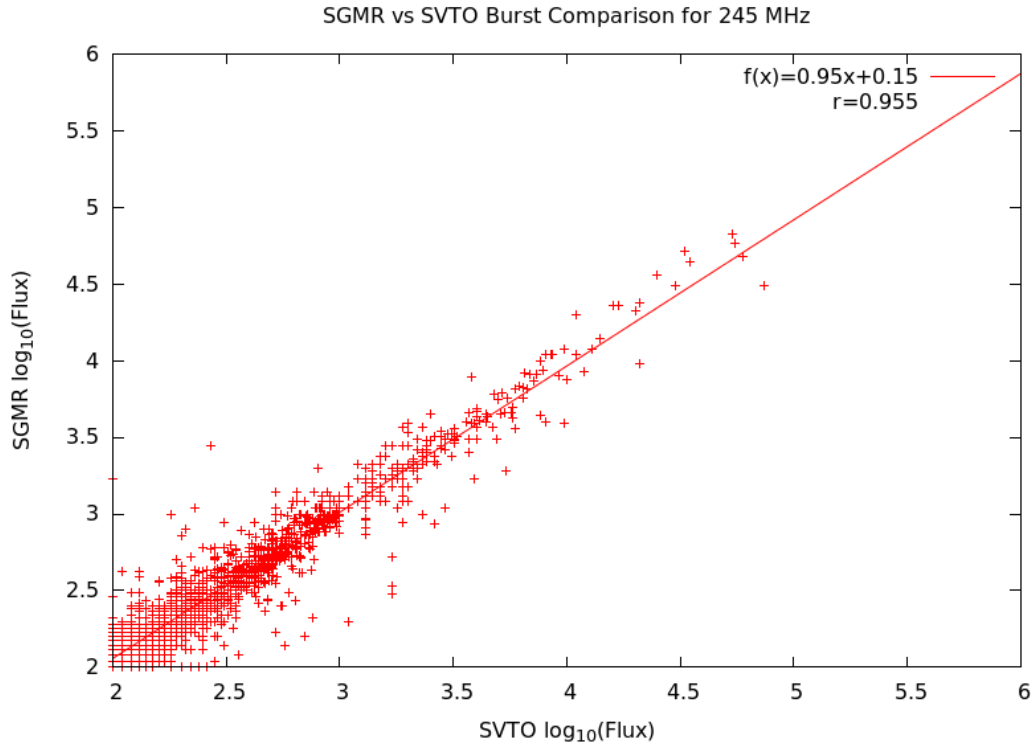



**Figure A.29**: Sagamore Hill and San Vito Burst Peak Flux Comparison for 245 MHz

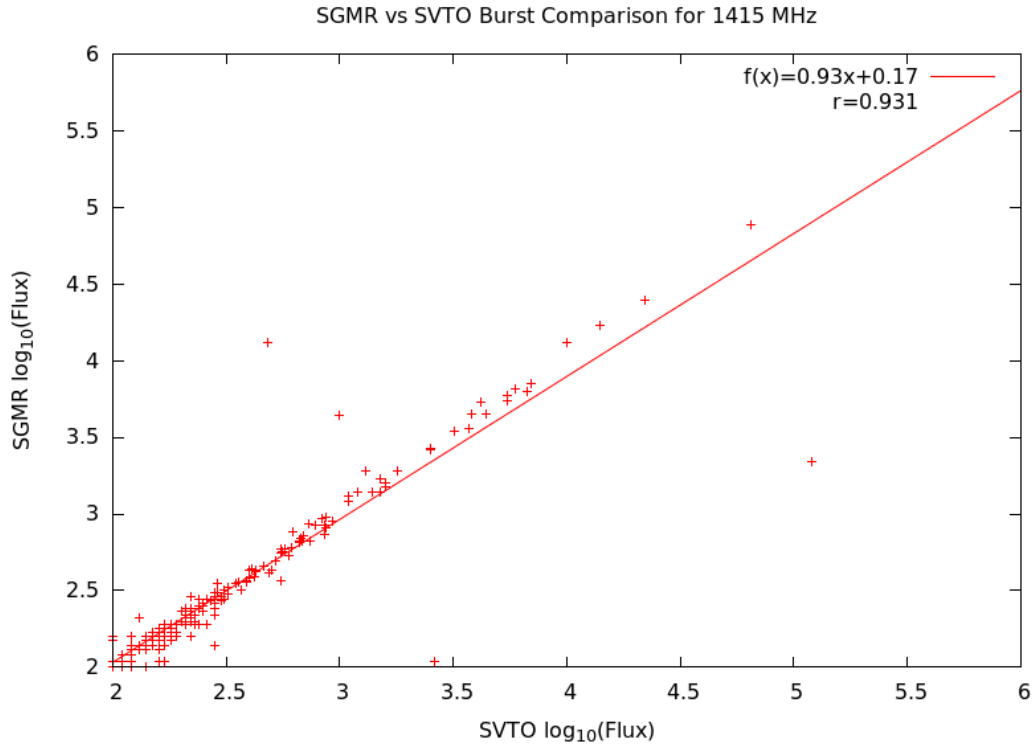



**Figure A.30**: Sagamore Hill and San Vito Burst Peak Flux Comparison for 1415 MHz

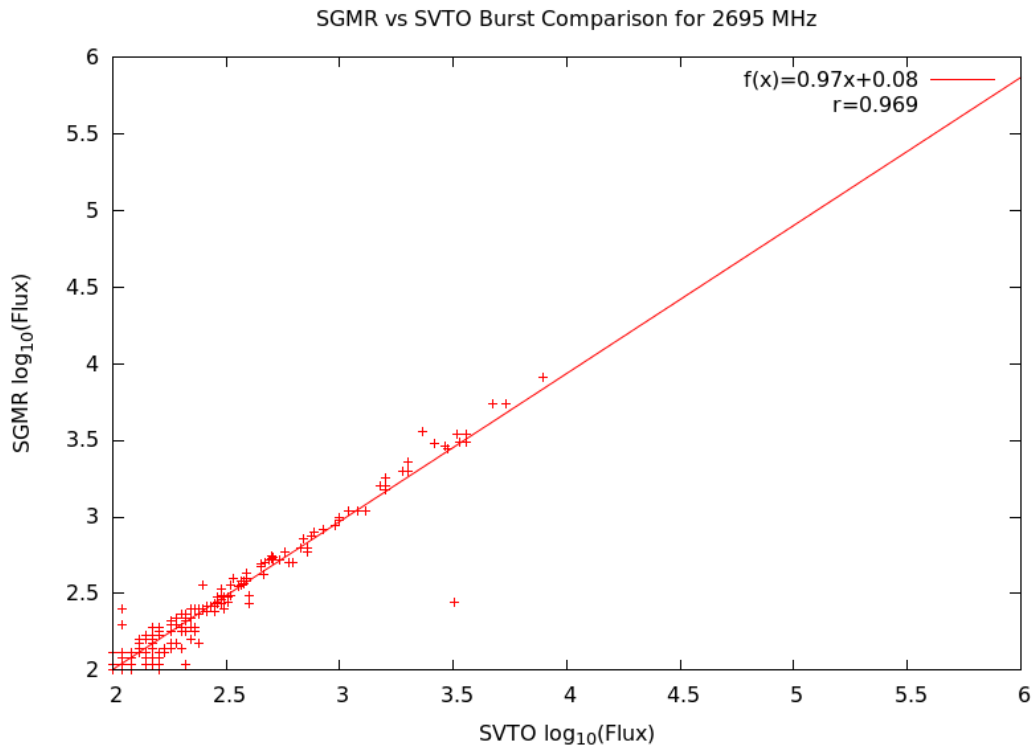



**Figure A.31**: Sagamore Hill and San Vito Burst Peak Flux Comparison for 2695 MHz

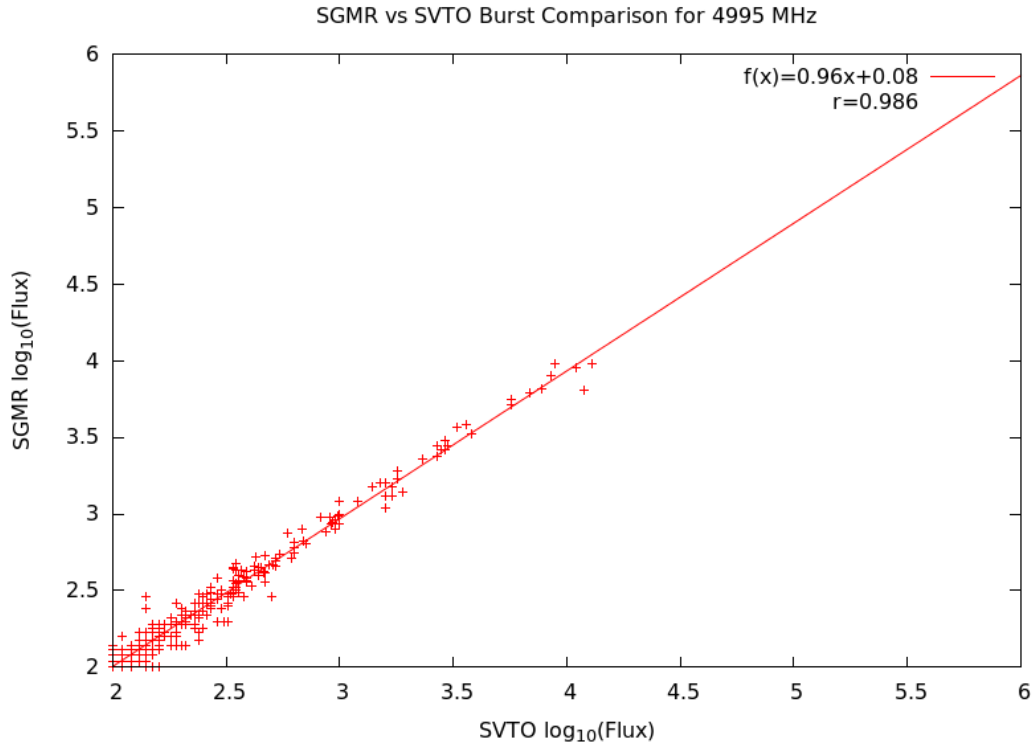



**Figure A.32**: Sagamore Hill and San Vito Burst Peak Flux Comparison for 4995 MHz

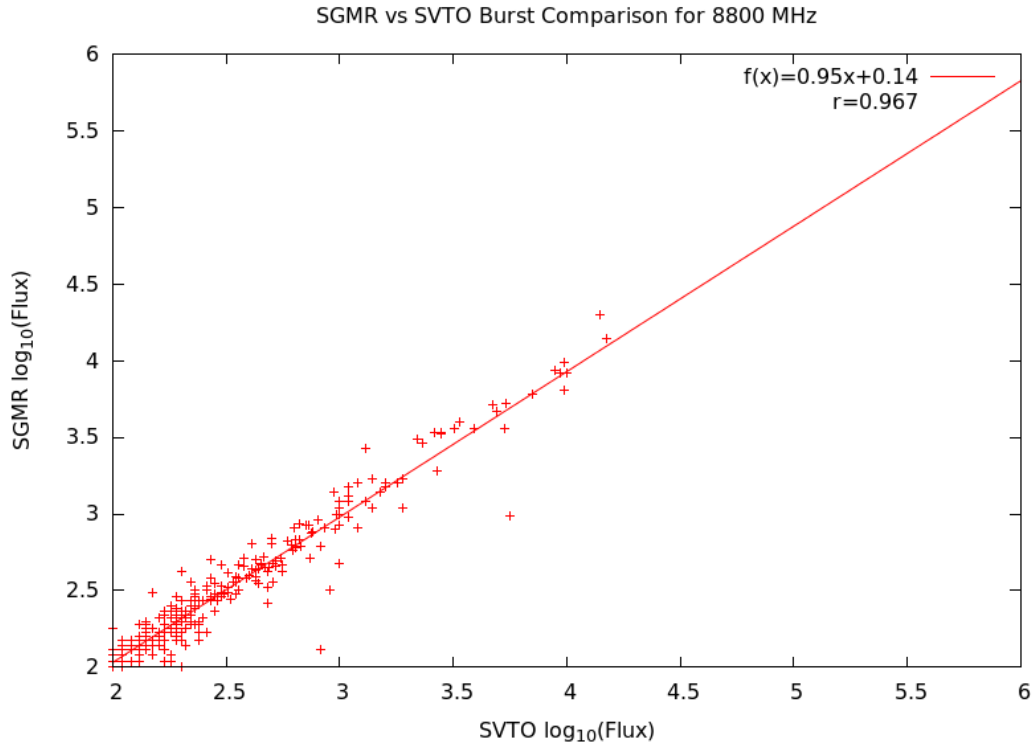



**Figure A.33**: Sagamore Hill and San Vito Burst Peak Flux Comparison for 8800 MHz

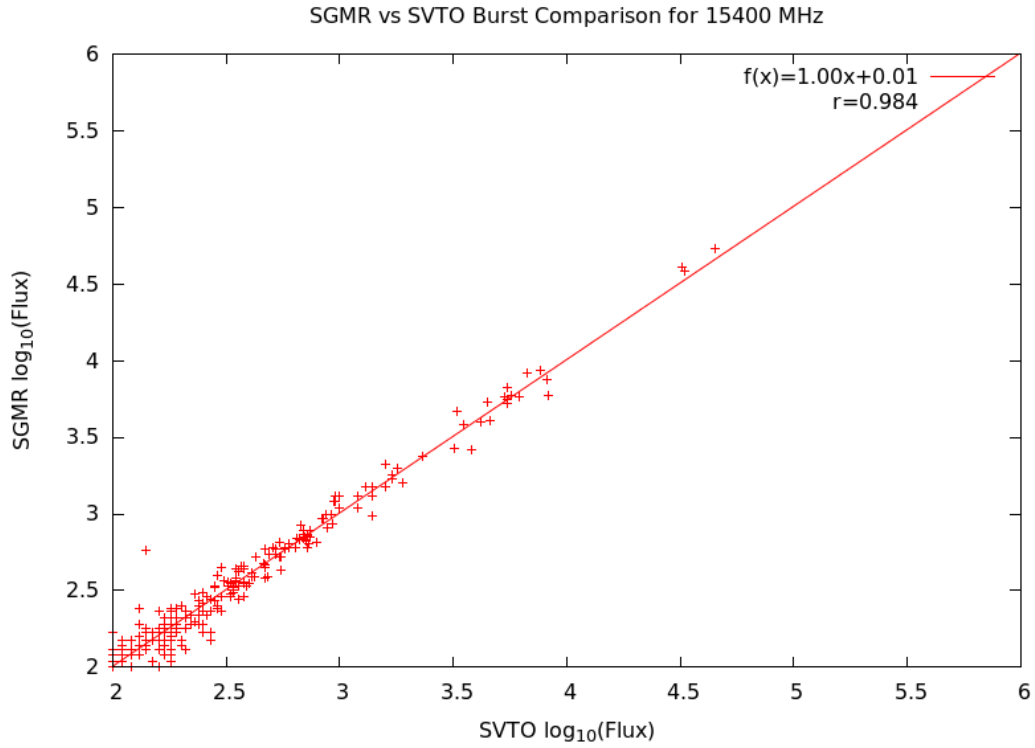



**Figure A.34**: Sagamore Hill and San Vito Burst Peak Flux Comparison for 15400 MHz

| Freq (Mhz) | n | m | $s_m$ | b | $s_b$ | r |
|---|---|---|---|---|---|---|
| 245 | 1295 | 0.950 | 0.010 | 0.164 | 0.027 | 0.936 |
| 410 | 542 | 0.915 | 0.019 | 0.274 | 0.049 | 0.905 |
| 610 | 418 | 0.916 | 0.016 | 0.210 | 0.042 | 0.944 |
| 1415 | 229 | 0.952 | 0.014 | 0.126 | 0.036 | 0.976 |
| 2695 | 225 | 1.013 | 0.010 | -0.028 | 0.027 | 0.988 |
| 4995 | 293 | 1.001 | 0.012 | -0.017 | 0.030 | 0.981 |
| 8800 | 311 | 0.966 | 0.014 | 0.064 | 0.037 | 0.970 |
| 15400 | 250 | 0.972 | 0.014 | 0.079 | 0.038 | 0.976 |

**Table A.5**: Palehua vs Sagamore Burst Comparison Regression Analysis

## Appendix B: GONG – Solar Region Summary McIntosh Classification Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | A | B | C | D | E | F | H | Total GONG |
| A | Matches | 459 | 458 | 764 | 339 | 39 | 9 | 624 | 2692 |
| | %GONG | 17.05 | 17.01 | 28.38 | 12.59 | 1.45 | 0.33 | 23.18 | |
| | %SRS | 92.91 | 70.35 | 31.40 | 10.34 | 3.09 | 2.71 | 27.89 | |
| B | Matches | 17 | 174 | 413 | 624 | 77 | 18 | 24 | 1347 |
| | %GONG | 1.26 | 12.92 | 30.66 | 46.33 | 5.72 | 1.34 | 1.78 | |
| | %SRS | 3.44 | 26.73 | 16.97 | 19.03 | 6.11 | 5.42 | 1.07 | |
| C | Matches | 4 | 13 | 478 | 1119 | 394 | 60 | 143 | 2211 |
| | %GONG | 0.18 | 0.59 | 21.62 | 50.61 | 17.82 | 2.71 | 6.47 | |
| | %SRS | 0.81 | 2.00 | 19.65 | 34.13 | 31.25 | 18.07 | 6.39 | |
| D | Matches | 2 | 0 | 41 | 754 | 137 | 20 | 86 | 1040 |
| | %GONG | 0.19 | 0.00 | 3.94 | 72.50 | 13.17 | 1.92 | 8.27 | |
| | %SRS | 0.40 | 0.00 | 1.69 | 22.99 | 10.86 | 6.02 | 3.84 | |
| E | Matches | 1 | 0 | 5 | 89 | 495 | 52 | 4 | 646 |
| | %GONG | 0.15 | 0.00 | 0.77 | 13.78 | 76.63 | 8.05 | 0.62 | |
| | %SRS | 0.20 | 0.00 | 0.21 | 2.71 | 39.25 | 15.66 | 0.18 | |
| F | Matches | 0 | 0 | 1 | 8 | 40 | 150 | 0 | 199 |
| | %GONG | 0.00 | 0.00 | 0.50 | 4.02 | 20.10 | 75.38 | 0.00 | |
| | %SRS | 0.00 | 0.00 | 0.04 | 0.24 | 3.17 | 45.18 | 0.00 | |
| H | Matches | 11 | 6 | 731 | 346 | 79 | 23 | 1356 | 2552 |
| | %GONG | 0.43 | 0.24 | 28.64 | 13.56 | 3.10 | 0.90 | 53.13 | |
| | %SRS | 2.23 | 0.92 | 30.05 | 10.55 | 6.26 | 6.93 | 60.62 | |
| | Total SRS | 494 | 651 | 2433 | 3279 | 1261 | 332 | 2237 | |

**Table B.1**: GONG Learmonth – Solar Region Summary Z Parameter Comparison

| | | SRS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GONG | | X | R | S | A | H | K | Total GONG |
| X | Matches | 1108 | 671 | 1278 | 935 | 6 | 41 | 4039 |
| | %GONG | 27.43 | 16.61 | 31.64 | 23.15 | 0.15 | 1.02 | |
| | %SRS | 96.77 | 86.69 | 32.30 | 28.43 | 2.04 | 3.34 | |
| S | Matches | 16 | 50 | 1137 | 745 | 25 | 54 | 2027 |
| | %GONG | 0.79 | 2.47 | 56.09 | 36.75 | 1.23 | 2.66 | |
| | %SRS | 1.40 | 6.46 | 28.73 | 22.65 | 8.50 | 4.40 | |
| A | Matches | 20 | 51 | 1330 | 1299 | 41 | 128 | 2869 |
| | %GONG | 0.70 | 1.78 | 46.36 | 45.28 | 1.43 | 4.46 | |
| | %SRS | 1.75 | 6.59 | 33.61 | 39.50 | 13.95 | 10.42 | |
| H | Matches | 0 | 0 | 81 | 96 | 95 | 339 | 611 |
| | %GONG | 0.00 | 0.00 | 13.26 | 15.71 | 15.55 | 55.48 | |
| | %SRS | 0.00 | 0.00 | 2.05 | 2.92 | 32.31 | 27.61 | |
| K | Matches | 1 | 2 | 131 | 214 | 127 | 666 | 1141 |
| | %GONG | 0.09 | 0.18 | 11.48 | 18.76 | 11.13 | 58.37 | |
| | %SRS | 0.09 | 0.26 | 3.31 | 6.51 | 43.20 | 54.23 | |
| | Total SRS | 1145 | 774 | 3957 | 3289 | 294 | 1228 | |

**Table B.2:** GONG Learmonth – Solar Region Summary P Parameter Comparison

| | | SRS | | | | |
|---|---|---|---|---|---|---|
| GONG | | X | O | I | C | Total GONG |
| X | Matches | 2449 | 2594 | 147 | 54 | 5244 |
| | %GONG | 46.70 | 49.47 | 2.80 | 1.03 | |
| | %SRS | 89.71 | 46.34 | 9.83 | 6.26 | |
| O | Matches | 55 | 1478 | 252 | 25 | 1810 |
| | %GONG | 3.04 | 81.66 | 13.92 | 1.38 | |
| | %SRS | 2.01 | 26.40 | 16.84 | 2.90 | |
| I | Matches | 152 | 1364 | 837 | 341 | 2694 |
| | %GONG | 5.64 | 50.63 | 31.07 | 12.66 | |
| | %SRS | 5.57 | 24.37 | 55.95 | 39.51 | |
| C | Matches | 74 | 162 | 260 | 443 | 939 |
| | %GONG | 7.88 | 17.25 | 27.69 | 47.18 | |
| | %SRS | 2.71 | 2.89 | 17.38 | 51.33 | |
| | Total SRS | 2730 | 5598 | 1496 | 863 | |

**Table B.3**: GONG Learmonth – Solar Region Summary C Parameter Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | A | B | C | D | E | F | H | Total GONG |
| A | Matches | 332 | 344 | 478 | 189 | 41 | 2 | 413 | 1799 |
| | %GONG | 18.45 | 19.12 | 26.57 | 10.51 | 2.28 | 0.11 | 22.96 | |
| | %SRS | 90.46 | 67.45 | 27.49 | 7.93 | 4.53 | 1.08 | 27.10 | |
| B | Matches | 8 | 150 | 294 | 405 | 33 | 5 | 19 | 914 |
| | %GONG | 0.88 | 16.41 | 32.17 | 44.31 | 3.61 | 0.55 | 2.08 | |
| | %SRS | 2.18 | 29.41 | 16.91 | 17.00 | 3.64 | 2.69 | 1.25 | |
| C | Matches | 3 | 9 | 416 | 786 | 258 | 34 | 88 | 1594 |
| | %GONG | 0.19 | 0.56 | 26.10 | 49.31 | 16.19 | 2.13 | 5.52 | |
| | %SRS | 0.82 | 1.76 | 23.92 | 33.00 | 28.48 | 18.28 | 5.77 | |
| D | Matches | 0 | 2 | 47 | 658 | 106 | 13 | 56 | 882 |
| | %GONG | 0.00 | 0.23 | 5.33 | 74.60 | 12.02 | 1.47 | 6.35 | |
| | %SRS | 0.00 | 0.39 | 2.70 | 27.62 | 11.70 | 6.99 | 3.67 | |
| E | Matches | 1 | 0 | 8 | 93 | 385 | 24 | 3 | 514 |
| | %GONG | 0.19 | 0.00 | 1.56 | 18.09 | 74.90 | 4.67 | 0.58 | |
| | %SRS | 0.27 | 0.00 | 0.46 | 3.90 | 42.49 | 12.90 | 0.20 | |
| F | Matches | 0 | 0 | 0 | 2 | 31 | 91 | 0 | 124 |
| | %GONG | 0.00 | 0.00 | 0.00 | 1.61 | 25.00 | 73.39 | 0.00 | |
| | %SRS | 0.00 | 0.00 | 0.00 | 0.08 | 3.42 | 48.92 | 0.00 | |
| H | Matches | 23 | 5 | 496 | 249 | 52 | 17 | 945 | 1787 |
| | %GONG | 1.29 | 0.28 | 27.76 | 13.93 | 2.91 | 0.95 | 52.88 | |
| | %SRS | 6.27 | 0.98 | 28.52 | 10.45 | 5.74 | 9.14 | 62.01 | |
| | Total SRS | 367 | 510 | 1739 | 2382 | 906 | 186 | 1524 | |

**Table B.4**: GONG Udaipur – Solar Region Summary Z Parameter Comparison

| | | SRS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GONG | | X | R | S | A | H | K | Total GONG |
| X | Matches | 835 | 426 | 834 | 589 | 9 | 20 | 2713 |
| | %GONG | 30.78 | 15.70 | 30.74 | 21.71 | 0.33 | 0.74 | |
| | %SRS | 95.10 | 80.08 | 30.67 | 24.11 | 4.55 | 2.37 | |
| S | Matches | 21 | 43 | 756 | 575 | 13 | 50 | 1458 |
| | %GONG | 1.44 | 2.95 | 51.85 | 39.44 | 0.89 | 3.43 | |
| | %SRS | 2.39 | 8.08 | 27.80 | 23.54 | 6.57 | 5.92 | |
| A | Matches | 20 | 61 | 999 | 1064 | 32 | 102 | 2278 |
| | %GONG | 0.88 | 2.68 | 43.85 | 46.71 | 1.40 | 4.48 | |
| | %SRS | 2.28 | 11.47 | 36.74 | 43.55 | 16.16 | 12.09 | |
| H | Matches | 0 | 0 | 44 | 67 | 61 | 210 | 382 |
| | %GONG | 0.00 | 0.00 | 11.52 | 17.54 | 15.97 | 54.97 | |
| | %SRS | 0.00 | 0.00 | 1.62 | 2.74 | 30.81 | 24.88 | |
| K | Matches | 2 | 2 | 86 | 148 | 83 | 462 | 783 |
| | %GONG | 0.26 | 0.26 | 10.98 | 18.90 | 10.60 | 59.00 | |
| | %SRS | 0.23 | 0.38 | 3.16 | 6.06 | 41.92 | 54.74 | |
| | Total SRS | 878 | 532 | 2719 | 2443 | 198 | 844 | |

**Table B.5**: GONG Udaipur – Solar Region Summary P Parameter Comparison

| | | | | SRS | | |
|---|---|---|---|---|---|---|
| GONG | | X | O | I | C | Total GONG |
| X | Matches | 1713 | 1751 | 93 | 29 | 3586 |
| | %GONG | 47.77 | 48.83 | 2.59 | 0.81 | |
| | %SRS | 90.59 | 43.85 | 8.49 | 4.57 | |
| O | Matches | 35 | 1131 | 204 | 25 | 1395 |
| | %GONG | 2.51 | 81.08 | 14.62 | 1.79 | |
| | %SRS | 1.85 | 28.32 | 18.61 | 3.94 | |
| I | Matches | 95 | 975 | 613 | 261 | 1944 |
| | %GONG | 4.89 | 50.15 | 31.53 | 13.43 | |
| | %SRS | 5.02 | 24.42 | 55.93 | 41.17 | |
| C | Matches | 48 | 136 | 186 | 319 | 689 |
| | %GONG | 6.97 | 19.74 | 27.00 | 46.30 | |
| | %SRS | 2.54 | 3.41 | 16.97 | 50.32 | |
| | Total SRS | 1891 | 3993 | 1096 | 634 | |

**Table B.6**: GONG Udaipur – Solar Region Summary C Parameter Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | A | B | C | D | E | F | H | Total GONG |
| A | Matches | 368 | 419 | 732 | 263 | 47 | 9 | 593 | 2431 |
| | %GONG | 15.14 | 17.24 | 30.11 | 10.82 | 1.93 | 0.37 | 24.39 | |
| | %SRS | 92.70 | 70.78 | 31.16 | 8.23 | 4.08 | 3.56 | 28.61 | |
| B | Matches | 11 | 149 | 420 | 610 | 65 | 7 | 23 | 1285 |
| | %GONG | 0.86 | 11.60 | 32.68 | 47.47 | 5.06 | 0.54 | 1.79 | |
| | %SRS | 2.77 | 25.17 | 17.88 | 19.08 | 5.64 | 2.77 | 1.11 | |
| C | Matches | 2 | 15 | 440 | 1121 | 341 | 48 | 121 | 2088 |
| | %GONG | 0.10 | 0.72 | 21.07 | 53.69 | 16.33 | 2.30 | 5.80 | |
| | %SRS | 0.50 | 2.53 | 18.73 | 35.06 | 29.58 | 18.97 | 5.84 | |
| D | Matches | 3 | 3 | 48 | 795 | 126 | 24 | 83 | 1082 |
| | %GONG | 0.28 | 0.28 | 4.44 | 73.48 | 11.65 | 2.22 | 7.67 | |
| | %SRS | 0.76 | 0.51 | 2.04 | 24.87 | 10.93 | 9.49 | 4.00 | |
| E | Matches | 2 | 0 | 11 | 84 | 447 | 39 | 6 | 589 |
| | %GONG | 0.34 | 0.00 | 1.87 | 14.26 | 75.89 | 6.62 | 1.02 | |
| | %SRS | 0.50 | 0.00 | 0.47 | 2.63 | 38.77 | 15.42 | 0.29 | |
| F | Matches | 0 | 0 | 0 | 6 | 44 | 114 | 0 | 164 |
| | %GONG | 0.00 | 0.00 | 0.00 | 3.66 | 26.83 | 69.51 | 0.00 | |
| | %SRS | 0.00 | 0.00 | 0.00 | 0.19 | 3.82 | 45.06 | 0.00 | |
| H | Matches | 11 | 6 | 698 | 318 | 83 | 12 | 1247 | 2375 |
| | %GONG | 0.46 | 0.25 | 29.39 | 13.39 | 3.49 | 0.51 | 52.51 | |
| | %SRS | 2.77 | 1.01 | 29.71 | 9.95 | 7.20 | 4.74 | 60.15 | |
| | Total SRS | 397 | 592 | 2349 | 3197 | 1153 | 253 | 2073 | |

**Table B.7**: GONG Teide – Solar Region Summary Z Parameter Comparison

| | | SRS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GONG | | X | R | S | A | H | K | Total GONG |
| X | Matches | 947 | 675 | 1188 | 873 | 8 | 25 | 3716 |
| | %GONG | 25.48 | 18.16 | 31.97 | 23.49 | 0.22 | 0.67 | |
| | %SRS | 95.75 | 87.21 | 31.45 | 28.17 | 3.02 | 2.25 | |
| S | Matches | 23 | 43 | 1081 | 688 | 25 | 60 | 1920 |
| | %GONG | 1.20 | 2.24 | 56.30 | 35.83 | 1.30 | 3.13 | |
| | %SRS | 2.33 | 5.56 | 28.62 | 22.20 | 9.43 | 5.41 | |
| A | Matches | 16 | 55 | 1348 | 1291 | 37 | 127 | 2874 |
| | %GONG | 0.56 | 1.91 | 46.90 | 44.92 | 1.29 | 4.42 | |
| | %SRS | 1.62 | 7.11 | 35.69 | 41.66 | 13.96 | 11.44 | |
| H | Matches | 1 | 0 | 60 | 67 | 94 | 311 | 533 |
| | %GONG | 0.19 | 0.00 | 11.26 | 12.57 | 17.64 | 58.35 | |
| | %SRS | 0.10 | 0.00 | 1.59 | 2.16 | 35.47 | 28.02 | |
| K | Matches | 2 | 1 | 100 | 180 | 101 | 587 | 971 |
| | %GONG | 0.21 | 0.10 | 10.30 | 18.54 | 10.40 | 60.45 | |
| | %SRS | 0.20 | 0.13 | 2.65 | 5.81 | 38.11 | 52.88 | |
| | Total SRS | 989 | 774 | 3777 | 3099 | 265 | 1110 | |

**Table B.8**: GONG Teide – Solar Region Summary P Parameter Comparison

| | | SRS | | | | |
|---|---|---|---|---|---|---|
| GONG | | X | O | I | C | Total GONG |
| X | Matches | 2219 | 2429 | 118 | 40 | 4806 |
| | %GONG | 46.17 | 50.54 | 2.46 | 0.83 | |
| | %SRS | 89.87 | 45.08 | 8.56 | 5.14 | |
| O | Matches | 40 | 1518 | 252 | 27 | 1837 |
| | %GONG | 2.18 | 82.63 | 13.72 | 1.47 | |
| | %SRS | 1.62 | 28.17 | 18.27 | 3.47 | |
| I | Matches | 136 | 1283 | 770 | 311 | 2500 |
| | %GONG | 5.44 | 51.32 | 30.80 | 12.44 | |
| | %SRS | 5.51 | 23.81 | 55.84 | 39.97 | |
| C | Matches | 74 | 158 | 239 | 400 | 871 |
| | %GONG | 8.50 | 18.14 | 27.44 | 45.92 | |
| | %SRS | 3.00 | 2.93 | 17.33 | 51.41 | |
| | Total SRS | 2469 | 5388 | 1379 | 778 | |

**Table B.9**: GONG Teide – Solar Region Summary C Parameter Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | A | B | C | D | E | F | H | Total GONG |
| A | Matches | 377 | 454 | 728 | 245 | 40 | 7 | 557 | 2408 |
| | %GONG | 15.66 | 18.85 | 30.23 | 10.17 | 1.66 | 0.29 | 23.13 | |
| | %SRS | 89.98 | 65.04 | 29.06 | 7.36 | 3.24 | 2.40 | 24.61 | |
| B | Matches | 12 | 207 | 401 | 530 | 59 | 10 | 27 | 1246 |
| | %GONG | 0.96 | 16.61 | 32.18 | 42.54 | 4.74 | 0.80 | 2.17 | |
| | %SRS | 2.86 | 29.66 | 16.01 | 15.92 | 4.79 | 3.42 | 1.19 | |
| C | Matches | 4 | 28 | 566 | 1107 | 334 | 58 | 127 | 2224 |
| | %GONG | 0.18 | 1.26 | 25.45 | 49.78 | 15.02 | 2.61 | 5.71 | |
| | %SRS | 0.95 | 4.01 | 22.59 | 33.24 | 27.09 | 19.86 | 5.61 | |
| D | Matches | 1 | 2 | 57 | 965 | 145 | 24 | 87 | 1281 |
| | %GONG | 0.08 | 0.16 | 4.45 | 75.33 | 11.32 | 1.87 | 6.79 | |
| | %SRS | 0.24 | 0.29 | 2.28 | 28.98 | 11.76 | 8.22 | 3.84 | |
| E | Matches | 1 | 0 | 9 | 141 | 528 | 40 | 7 | 726 |
| | %GONG | 0.14 | 0.00 | 1.24 | 19.42 | 72.73 | 5.51 | 0.96 | |
| | %SRS | 0.24 | 0.00 | 0.36 | 4.23 | 42.82 | 13.70 | 0.31 | |
| F | Matches | 0 | 0 | 1 | 7 | 46 | 138 | 0 | 192 |
| | %GONG | 0.00 | 0.00 | 0.52 | 3.65 | 23.96 | 71.88 | 0.00 | |
| | %SRS | 0.00 | 0.00 | 0.04 | 0.21 | 3.73 | 47.26 | 0.00 | |
| H | Matches | 24 | 7 | 743 | 335 | 81 | 15 | 1458 | 2663 |
| | %GONG | 0.90 | 0.26 | 27.90 | 12.58 | 3.04 | 0.56 | 54.75 | |
| | %SRS | 5.73 | 1.00 | 29.66 | 10.06 | 6.57 | 5.14 | 64.43 | |
| | Total SRS | 419 | 698 | 2505 | 3330 | 1233 | 292 | 2263 | |

**Table B.10**: GONG Cerra Tololo – Solar Region Summary Z Parameter Comparison

| | | SRS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GONG | | X | R | S | A | H | K | Total GONG |
| X | Matches | 1050 | 677 | 1113 | 775 | 7 | 32 | 3654 |
| | %GONG | 28.74 | 18.53 | 30.46 | 21.21 | 0.19 | 0.88 | |
| | %SRS | 94.00 | 82.36 | 28.38 | 22.69 | 2.38 | 2.74 | |
| S | Matches | 24 | 64 | 1146 | 784 | 36 | 63 | 2117 |
| | %GONG | 1.13 | 3.02 | 54.13 | 37.03 | 1.70 | 2.98 | |
| | %SRS | 2.15 | 7.79 | 29.22 | 22.95 | 12.24 | 5.39 | |
| A | Matches | 42 | 80 | 1476 | 1546 | 40 | 135 | 3319 |
| | %GONG | 1.27 | 2.41 | 44.47 | 46.58 | 1.21 | 4.07 | |
| | %SRS | 3.76 | 9.73 | 37.63 | 45.26 | 13.61 | 11.55 | |
| H | Matches | 0 | 0 | 63 | 95 | 95 | 313 | 566 |
| | %GONG | 0.00 | 0.00 | 11.13 | 16.78 | 16.78 | 55.30 | |
| | %SRS | 0.00 | 0.00 | 1.61 | 2.78 | 32.31 | 26.78 | |
| K | Matches | 1 | 1 | 124 | 216 | 116 | 626 | 1084 |
| | %GONG | 0.09 | 0.09 | 11.44 | 19.93 | 10.70 | 57.75 | |
| | %SRS | 0.09 | 0.12 | 3.16 | 6.32 | 39.46 | 53.55 | |
| | Total SRS | 1117 | 822 | 3922 | 3416 | 294 | 1169 | |

**Table B.11**: GONG Cerra Tololo – Solar Region Summary P Parameter Comparison

| | | SRS | | | | |
|---|---|---|---|---|---|---|
| GONG | | X | O | I | C | Total GONG |
| X | Matches | 2415 | 2498 | 120 | 38 | 5071 |
| | %GONG | 47.62 | 49.26 | 2.37 | 0.75 | |
| | %SRS | 90.08 | 43.41 | 8.19 | 4.52 | |
| O | Matches | 39 | 1670 | 254 | 25 | 1988 |
| | %GONG | 1.96 | 84.00 | 12.78 | 1.26 | |
| | %SRS | 1.45 | 29.02 | 17.34 | 2.98 | |
| I | Matches | 155 | 1398 | 830 | 333 | 2716 |
| | %GONG | 5.71 | 51.47 | 30.56 | 12.26 | |
| | %SRS | 5.78 | 24.30 | 56.66 | 39.64 | |
| C | Matches | 72 | 188 | 261 | 444 | 965 |
| | %GONG | 7.46 | 19.48 | 27.05 | 46.01 | |
| | %SRS | 2.69 | 3.27 | 17.82 | 52.86 | |
| | Total SRS | 2681 | 5754 | 1465 | 840 | |

**Table B.12**: GONG Cerra Tololo – Solar Region Summary C Parameter Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | A | B | C | D | E | F | H | Total GONG |
| A | Matches | 292 | 422 | 733 | 292 | 65 | 10 | 563 | 2377 |
| | %GONG | 12.28 | 17.75 | 30.84 | 12.28 | 2.73 | 0.42 | 23.69 | |
| | %SRS | 91.82 | 67.20 | 31.03 | 9.18 | 5.55 | 3.80 | 27.30 | |
| B | Matches | 10 | 170 | 383 | 543 | 71 | 9 | 24 | 1210 |
| | %GONG | 0.83 | 14.05 | 31.65 | 44.88 | 5.87 | 0.74 | 1.98 | |
| | %SRS | 3.14 | 27.07 | 16.22 | 17.08 | 6.06 | 3.42 | 1.16 | |
| C | Matches | 2 | 25 | 456 | 1088 | 346 | 54 | 123 | 2094 |
| | %GONG | 0.10 | 1.19 | 21.78 | 51.96 | 16.52 | 2.58 | 5.87 | |
| | %SRS | 0.63 | 3.98 | 19.31 | 34.21 | 29.52 | 20.53 | 5.97 | |
| D | Matches | 1 | 4 | 70 | 794 | 124 | 20 | 81 | 1094 |
| | %GONG | 0.09 | 0.37 | 6.40 | 72.58 | 11.33 | 1.83 | 7.40 | |
| | %SRS | 0.31 | 0.64 | 2.96 | 24.97 | 10.58 | 7.60 | 3.93 | |
| E | Matches | 1 | 0 | 5 | 133 | 451 | 44 | 6 | 640 |
| | %GONG | 0.16 | 0.00 | 0.78 | 20.78 | 70.47 | 6.88 | 0.94 | |
| | %SRS | 0.31 | 0.00 | 0.21 | 4.18 | 38.48 | 16.73 | 0.29 | |
| F | Matches | 0 | 0 | 0 | 7 | 43 | 112 | 0 | 162 |
| | %GONG | 0.00 | 0.00 | 0.00 | 4.32 | 26.54 | 69.14 | 0.00 | |
| | %SRS | 0.00 | 0.00 | 0.00 | 0.22 | 3.67 | 42.59 | 0.00 | |
| H | Matches | 12 | 7 | 715 | 323 | 72 | 14 | 1265 | 2408 |
| | %GONG | 0.50 | 0.29 | 29.69 | 13.41 | 2.99 | 0.58 | 52.53 | |
| | %SRS | 3.77 | 1.11 | 30.27 | 10.16 | 6.14 | 5.32 | 61.35 | |
| | Total SRS | 318 | 628 | 2362 | 3180 | 1172 | 263 | 2062 | |

**Table B.13**: GONG Big Bear – Solar Region Summary Z Parameter Comparison

| | | SRS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GONG | | X | R | S | A | H | K | Total GONG |
| X | Matches | 894 | 627 | 1161 | 857 | 12 | 36 | 3587 |
| | %GONG | 24.92 | 17.48 | 32.37 | 23.89 | 0.33 | 1.00 | |
| | %SRS | 94.50 | 84.27 | 31.18 | 27.36 | 4.11 | 3.14 | |
| S | Matches | 17 | 44 | 1010 | 670 | 25 | 58 | 1824 |
| | %GONG | 0.93 | 2.41 | 55.37 | 36.73 | 1.37 | 3.18 | |
| | %SRS | 1.80 | 5.91 | 27.13 | 21.39 | 8.56 | 5.05 | |
| A | Matches | 33 | 71 | 1351 | 1323 | 41 | 128 | 2947 |
| | %GONG | 1.12 | 2.41 | 45.84 | 44.89 | 1.39 | 4.34 | |
| | %SRS | 3.49 | 9.54 | 36.29 | 42.24 | 14.04 | 11.15 | |
| H | Matches | 0 | 0 | 73 | 85 | 99 | 325 | 582 |
| | %GONG | 0.00 | 0.00 | 12.54 | 14.60 | 17.01 | 55.84 | |
| | %SRS | 0.00 | 0.00 | 1.96 | 2.71 | 33.90 | 28.31 | |
| K | Matches | 2 | 2 | 128 | 197 | 115 | 601 | 1045 |
| | %GONG | 0.19 | 0.19 | 12.25 | 18.85 | 11.00 | 57.51 | |
| | %SRS | 0.21 | 0.27 | 3.44 | 6.29 | 39.38 | 52.35 | |
| | Total SRS | 946 | 744 | 3723 | 3132 | 292 | 1148 | |

**Table B.14**: GONG Big Bear – Solar Region Summary P Parameter Comparison

| | | SRS | | | | |
|---|---|---|---|---|---|---|
| GONG | | X | O | I | C | Total GONG |
| X | Matches | 2132 | 2472 | 149 | 32 | 4785 |
| | %GONG | 44.56 | 51.66 | 3.11 | 0.67 | |
| | %SRS | 89.58 | 46.03 | 10.30 | 4.06 | |
| O | Matches | 31 | 1406 | 236 | 28 | 1701 |
| | %GONG | 1.82 | 82.66 | 13.87 | 1.65 | |
| | %SRS | 1.30 | 26.18 | 16.31 | 3.55 | |
| I | Matches | 150 | 1310 | 817 | 318 | 2595 |
| | %GONG | 5.78 | 50.48 | 31.48 | 12.25 | |
| | %SRS | 6.30 | 24.39 | 56.46 | 40.36 | |
| C | Matches | 67 | 182 | 245 | 410 | 904 |
| | %GONG | 7.41 | 20.13 | 27.10 | 45.35 | |
| | %SRS | 2.82 | 3.39 | 16.93 | 52.03 | |
| | Total SRS | 2380 | 5370 | 1447 | 788 | |

**Table B.15**: GONG Big Bear – Solar Region Summary C Parameter Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | A | B | C | D | E | F | H | Total GONG |
| A | Matches | 307 | 374 | 659 | 279 | 55 | 6 | 533 | 2213 |
| | %GONG | 13.87 | 16.90 | 29.78 | 12.61 | 2.49 | 0.27 | 24.08 | |
| | %SRS | 88.47 | 59.74 | 29.19 | 8.96 | 4.94 | 2.32 | 26.66 | |
| B | Matches | 17 | 196 | 361 | 502 | 63 | 14 | 24 | 1177 |
| | %GONG | 1.44 | 16.65 | 30.67 | 42.65 | 5.35 | 1.19 | 2.04 | |
| | %SRS | 4.90 | 31.31 | 15.99 | 16.13 | 5.66 | 5.41 | 1.20 | |
| C | Matches | 3 | 41 | 481 | 1047 | 304 | 50 | 139 | 2065 |
| | %GONG | 0.15 | 1.99 | 23.29 | 50.70 | 14.72 | 2.42 | 6.73 | |
| | %SRS | 0.86 | 6.55 | 21.30 | 33.63 | 27.31 | 19.31 | 6.95 | |
| D | Matches | 0 | 7 | 73 | 832 | 114 | 25 | 69 | 1120 |
| | %GONG | 0.00 | 0.63 | 6.52 | 74.29 | 10.18 | 2.23 | 6.16 | |
| | %SRS | 0.00 | 1.12 | 3.23 | 26.73 | 10.24 | 9.65 | 3.45 | |
| E | Matches | 1 | 0 | 8 | 137 | 467 | 37 | 7 | 657 |
| | %GONG | 0.15 | 0.00 | 1.22 | 20.85 | 71.08 | 5.63 | 1.07 | |
| | %SRS | 0.29 | 0.00 | 0.35 | 4.40 | 41.96 | 14.29 | 0.35 | |
| F | Matches | 0 | 0 | 2 | 9 | 38 | 114 | 0 | 163 |
| | %GONG | 0.00 | 0.00 | 1.23 | 5.52 | 23.31 | 69.94 | 0.00 | |
| | %SRS | 0.00 | 0.00 | 0.09 | 0.29 | 3.41 | 44.02 | 0.00 | |
| H | Matches | 19 | 8 | 674 | 307 | 72 | 13 | 1227 | 2320 |
| | %GONG | 0.82 | 0.34 | 29.05 | 13.23 | 3.10 | 0.56 | 52.89 | |
| | %SRS | 5.48 | 1.28 | 29.85 | 9.86 | 6.47 | 5.02 | 61.38 | |
| | Total SRS | 347 | 626 | 2258 | 3113 | 1113 | 259 | 1999 | |

**Table B.16**: GONG Mauna Loa – Solar Region Summary Z Parameter Comparison

| | | SRS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| GONG | | X | R | S | A | H | K | Total GONG |
| X | Matches | 894 | 603 | 1078 | 773 | 11 | 31 | 3390 |
| | %GONG | 26.37 | 17.79 | 31.80 | 22.80 | 0.32 | 0.91 | |
| | %SRS | 91.88 | 79.24 | 29.86 | 25.97 | 3.90 | 2.79 | |
| S | Matches | 37 | 51 | 1023 | 662 | 23 | 62 | 1858 |
| | %GONG | 1.99 | 2.74 | 55.06 | 35.63 | 1.24 | 3.34 | |
| | %SRS | 3.80 | 6.70 | 28.34 | 22.24 | 8.16 | 5.58 | |
| A | Matches | 39 | 105 | 1317 | 1285 | 44 | 122 | 2912 |
| | %GONG | 1.34 | 3.61 | 45.23 | 44.13 | 1.51 | 4.19 | |
| | %SRS | 4.01 | 13.80 | 36.48 | 43.16 | 15.60 | 10.97 | |
| H | Matches | 0 | 0 | 64 | 72 | 98 | 286 | 520 |
| | %GONG | 0.00 | 0.00 | 12.31 | 13.85 | 18.85 | 55.00 | |
| | %SRS | 0.00 | 0.00 | 1.77 | 2.42 | 34.75 | 25.72 | |
| K | Matches | 3 | 2 | 128 | 185 | 106 | 611 | 1035 |
| | %GONG | 0.29 | 0.19 | 12.37 | 17.87 | 10.24 | 59.03 | |
| | %SRS | 0.31 | 0.26 | 3.55 | 6.21 | 37.59 | 54.95 | |
| | Total SRS | 973 | 761 | 3610 | 2977 | 282 | 1112 | |

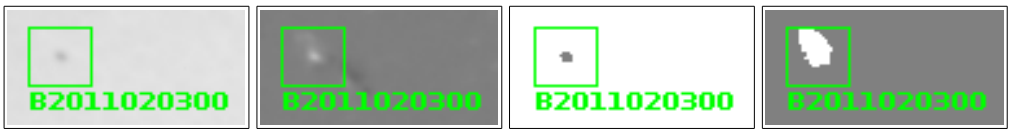
**Table B.17**: GONG Mauna Loa – Solar Region Summary P Parameter Comparison

| | | SRS | | | | |
|---|---|---|---|---|---|---|
| GONG | | X | O | I | C | Total GONG |
| X | Matches | 2085 | 2305 | 112 | 31 | 4533 |
| | %GONG | 46.00 | 50.85 | 2.47 | 0.68 | |
| | %SRS | 88.91 | 44.17 | 8.07 | 4.06 | |
| O | Matches | 42 | 1465 | 255 | 28 | 1790 |
| | %GONG | 2.35 | 81.84 | 14.25 | 1.56 | |
| | %SRS | 1.79 | 28.08 | 18.37 | 3.66 | |
| I | Matches | 163 | 1288 | 778 | 301 | 2530 |
| | %GONG | 6.44 | 50.91 | 30.75 | 11.90 | |
| | %SRS | 6.95 | 24.68 | 56.05 | 39.40 | |
| C | Matches | 55 | 160 | 243 | 404 | 862 |
| | %GONG | 6.38 | 18.56 | 28.19 | 46.87 | |
| | %SRS | 2.35 | 3.07 | 17.51 | 52.88 | |
| | Total SRS | 2345 | 5218 | 1388 | 764 | |

**Table B.18**: GONG Mauna Loa – Solar Region Summary C Parameter Comparison

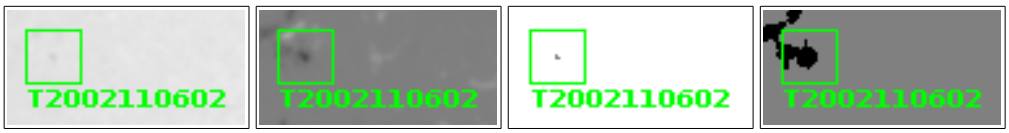
## Appendix C: Examples of McIntosh Classification Disagreement Between Automated GONG and the Solar Region Summary

In all of the following figures, (a) is the normalised intensity image, (b) is the magnetogram, (c) is the sunspot image, and (d) is the magnetic area image. In all cases for sunspots, grey is penumbra and black is umbra. For magnetic fields black is negative (into the sun) and white is positive (out of the sun). Heliographic north is down and heliographic west is to the right.



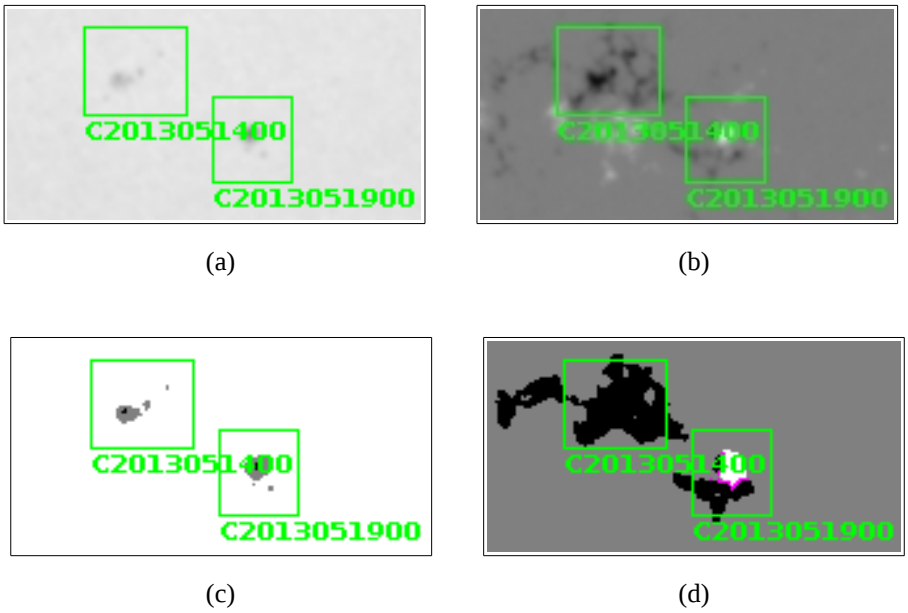

(a) (b) (c) (d)

**Figure C.1**: Example of GONG AXX classification assigned a BXO classification by SWPC. GONG processed images for Big Bear 2011/02/03 17:00:16 UT. This was SWPC region 1152 classified as BXO and GONG region B2011020300 classified as AXX. There is only positive field and no clear umbra visible on these images. Therefore, the GONG classification is correct using the imagery available. SOON analysts were able to resolve 6 sunspots using their equipment.




(a) (b) (c) (d)

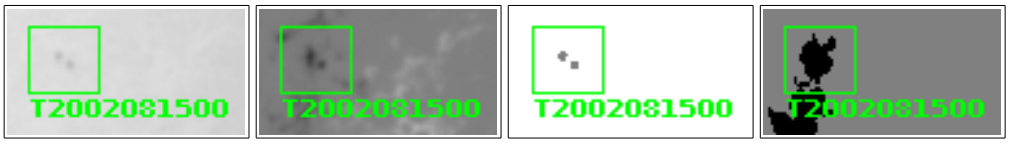
**Figure C.2**: Example of GONG AXX classification assigned a CSO classification by SWPC. GONG processed images for Tiede 2002/11/06 09:00:16 UT. This was SWPC region 0187 classified as CSO and GONG region T2002110602 classified as AXX. Some positive field is visible, but does not exceed the threshold of 35 Gauss. There is no clear umbra visible on these images. Therefore, the GONG classification is correct using the imagery available. SOON analysts were able to resolve 4 sunspots using their equipment.




(a) (b)

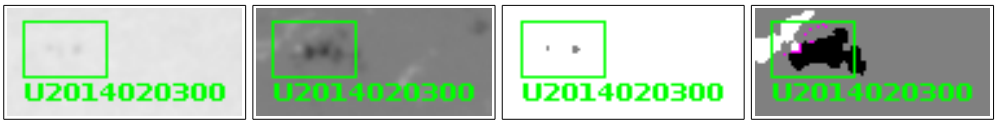
(c) (d)

**Figure C.3**: Example of GONG AXX classification assigned a FAO classification by SWPC. GONG processed images for Cerra Tololo 2013/05/19 15:00:16 UT. This was SWPC region 1748 classified as FAO and GONG region C2013110514 classified as AXX. This is an example where many things had gone wrong with the GONG classification. Firstly GONG has split the region into two, with upper left portion being unipolar. This is the primary reason for the discrepancy, Even if splitting the region can be considered correct, the program still should have classified the unipolar region as HAX, as there is an umbra and penumbra present. At this stage it is unclear where the error in the program is, or why some regions produce this error and others do not. (see section 11.3: Future Work).



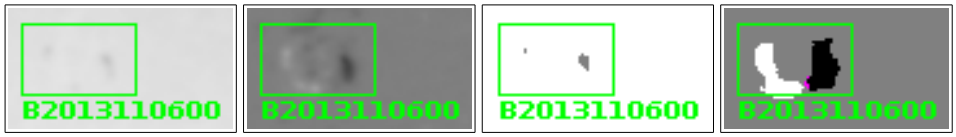

(a) (b) (c) (d)

**Figure C.4**: Example of GONG AXX classification assigned a HSX classification by SWPC. GONG processed images for Big Bear 2011/02/03 17:00:16 UT. This was SWPC region 0080 classified as HSX and GONG region T2002081500 classified as AXX. There are two sunspots visible on negative polarity, but no clear penumbra present. Therefore, the GONG classification is correct using the imagery available. SOON analysts were also able to resolve 2 sunspots using their equipment.



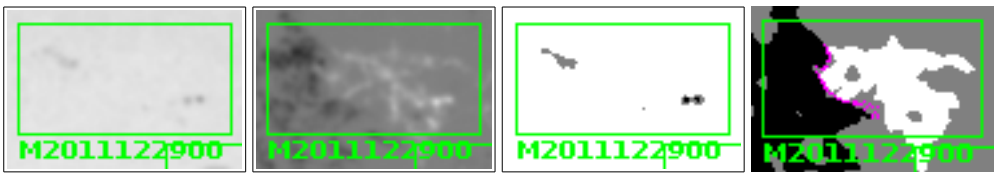

(a) (b) (c) (d)

**Figure C.5**: Example of GONG BXO classification assigned an AXX classification by SWPC. GONG processed images for Udaipur 2014/02/06 04:00:16 UT. This was SWPC region 1970 classified as AXX and GONG region U2014020300 classified as BXO. There are clearly two regions of magnetic field. Therefore, the GONG classification is correct using the imagery available. SOON analysts were also able to resolve 2 sunspots using their equipment, but not identify that the region was bipolar.



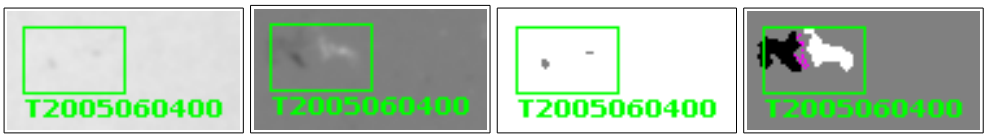

(a) (b) (c) (d)

**Figure C.6**: Example of GONG BXO classification assigned an CAO classification by SWPC. GONG processed images for Big Bear 2013/11/06 16:00:16 UT. This was SWPC region 1891 classified as CAO and GONG region B2013110600 classified as BXO. There is no visible penumbra on any of the sunspots. Therefore, the GONG classification is correct using the imagery available. SOON analysts were also able to resolve 2 sunspots using their equipment, but assigned one with penumbra.



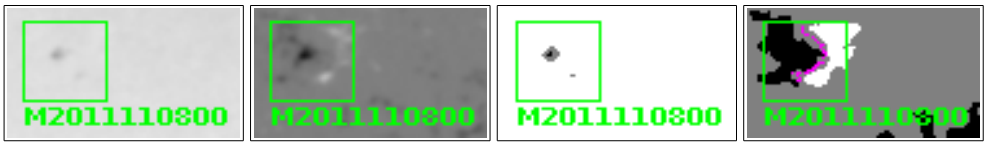

(a) (b) (c) (d)

**Figure C.7**: Example of GONG BXO classification assigned an ESO classification by SWPC. GONG processed images for Mauna Loa 2012/01/03 19:00:16 UT. This was SWPC region 1389 classified as ESO and GONG region M2011122900 classified as BXO. The GONG program has failed to assign this region a CSO classification based on this imagery as there are 2 sunspots with penumbra. SOON analysts were able to resolve 14 sunspots using their equipment. This is the same error that occurred in figure C.4.



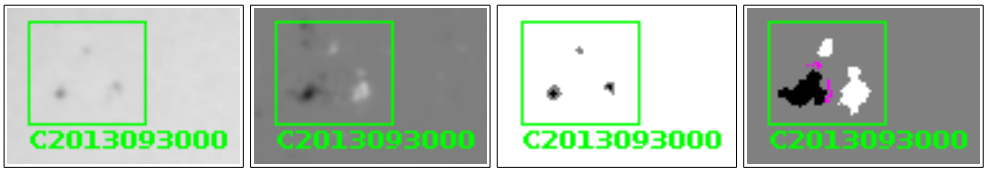

(a) (b) (c) (d)

**Figure C.8**: Example of GONG BXO classification assigned an HRX classification by SWPC. GONG processed images for Teide 2005/06/05 08:00:16 UT. This was SWPC region 0774 classified as HRX and GONG region T2005060400 classified as BXO. There are 2 sunspots present on opposite polarity field in this imagery. In addition this is an example of where rudimentary penumbra were unable to be resolved by GONG. SOON analysts were only able to resolve 1 sunspot using their equipment.



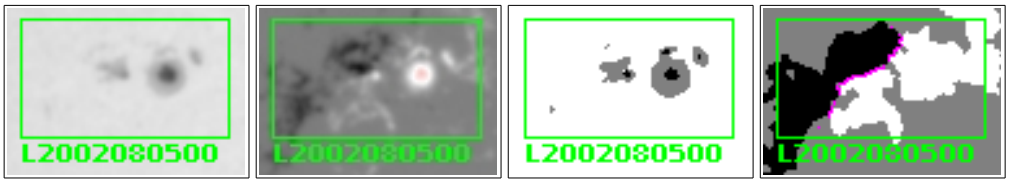

(a) (b) (c) (d)

**Figure C.9**: Example of GONG CAO classification assigned an AXX classification by SWPC. GONG processed images for Mauna Loa 2011/11/08 18:00:16 UT. This was SWPC region 1334 classified as AXX and GONG region M2011110800 classified as CAO. The GONG analysis is correct from the imagery available. SOON analysts were only able to resolve 1 sunspot using their equipment.



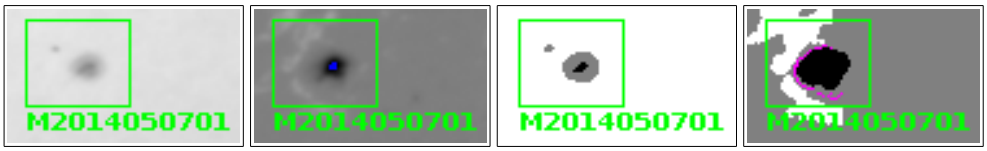

(a) (b) (c) (d)

**Figure C.10**: Example of GONG CSO classification assigned a BXO classification by SWPC. GONG processed images for Cerro Tololo 2013/09/30 13:00:16 UT. This was SWPC region 1855 classified as BXO and GONG region C2013093000 classified as CSO. Based on the imagery, GONG should have classified this region as DSO as there are sunspots of both polarities with penumbra. SOON analysts were only able to resolve 4 sunspots using their equipment, but did not assign any with penumbra. This is the same error that occurred in figures C.4 and C.7.



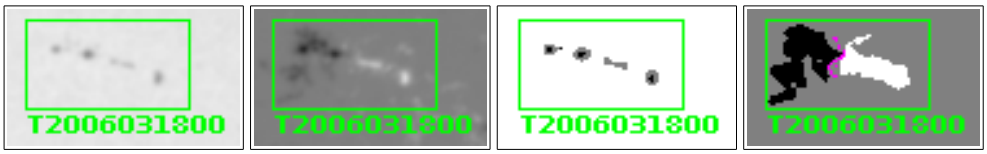

(a) (b) (c) (d)

**Figure C.11**: Example of GONG CKI classification assigned an EKC classification by SWPC. GONG processed images for Learmonth 2002/08/09 01:00:16 UT. This was SWPC region 0061 classified as EKC and GONG region L2002080500 classified as CKI. In this case, the umbra in the eastern spot is actually on a positive portion of field. Thus, the program considers this a C region, even though the spot also has negative field associated with it. SOON analysts were able to resolve 14 sunspots using their equipment.



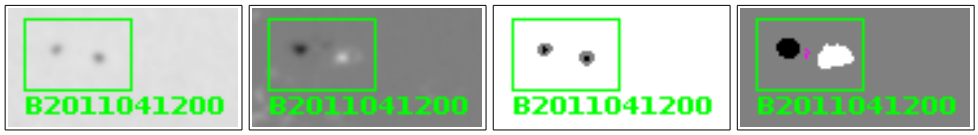

(a) (b) (c) (d)

**Figure C.12**: Example of GONG CAI classification assigned a HAX classification by SWPC. GONG processed images for Mauna Loa 2014/05/13 17:00:16 UT. This was SWPC region 2057 classified as HAX and GONG region M2014050701 classified as CAI. GONG has correctly classified this region based on the imagery available. SOON analysts were able to resolve 2 sunspots using their equipment, but did not identify that they were of opposite polarity.



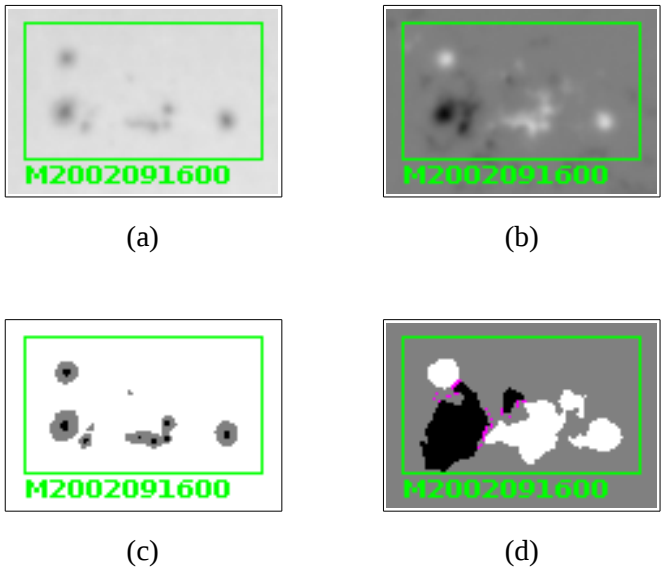

(a) (b) (c) (d)

**Figure C.13**: Example of GONG DAO classification assigned an AXX classification by SWPC. GONG processed images for Teide 2006/03/19 09:00:16 UT. This was SWPC region 0860 classified as AXX and GONG region T2006031800 classified as DAO. GONG has correctly classified this region based on the imagery available. SOON analysts were only able to resolve 2 sunspots using their equipment, but did not identify that they were of opposite polarity.




(a) (b) (c) (d)

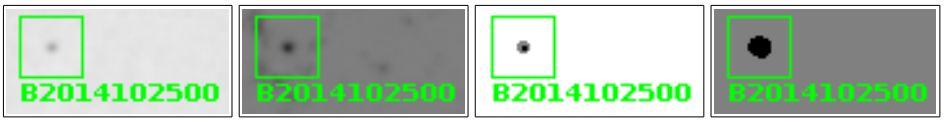
**Figure C.14**: Example of GONG DSO classification assigned a BXO classification by SWPC. GONG processed images for Big Bear 2011/04/12 15:00:16 UT. This was SWPC region 1192 classified as BXO and GONG region B2011041200 classified as DAO. GONG has correctly classified this region based on the imagery available. SOON analysts were able to resolve 3 sunspots using their equipment, but were unable to resolve penumbra.




(a) (b)

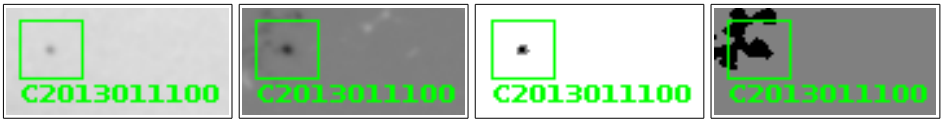
(c) (d)

**Figure C.15**: Example of GONG EAI classification assigned a HAX classification by SWPC. GONG processed images for Mauna Loa 2002/09/21 18:00:16 UT. This was SWPC region 0121 classified as HAX and GONG region M2002091600 classified as EAI. GONG has correctly classified this region based on the imagery available. SOON analysts were able to resolve 2 sunspots using their equipment, but were unable to resolve penumbra. They also possibly considered some of the larger spots to be individual H regions.



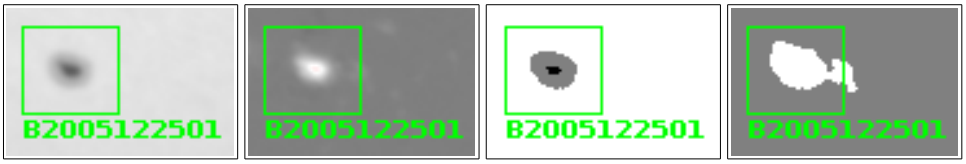

(a) (b) (c) (d)

**Figure C.16**: Example of GONG HSX classification assigned a AXX classification by SWPC. GONG processed images for Big Bear 2014/10/29 16:00:16 UT. This was SWPC region 2196 classified as AXX and GONG region B2014102500 classified as HSX. GONG has correctly classified this region based on the imagery available. SOON analysts were unable to resolve the penumbra.



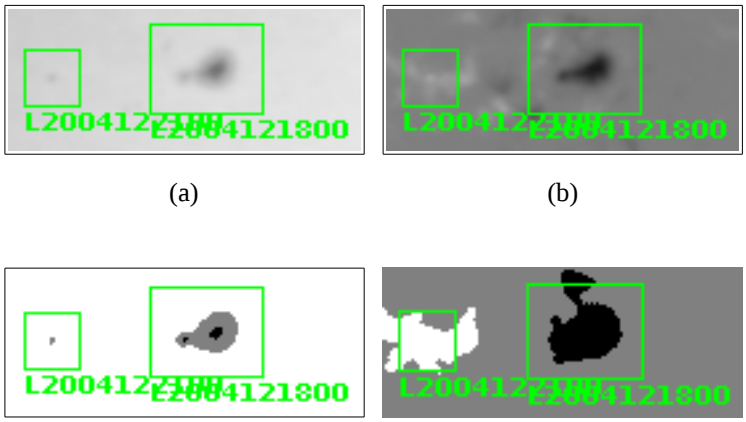

(a) (b) (c) (d)

**Figure C.17**: Example of GONG HAX classification assigned a BXO classification by SWPC. GONG processed images for Cerro Tololo 2013/01/14 13:00:16 UT. This was SWPC region 1656 classified as BXO and GONG region C2013011100 classified as HAX. GONG has correctly classified this region based on the imagery available. SOON analysts were able to resolve 7 sunspots , but were unable to resolve any penumbra.




(a) (b) (c) (d)

**Figure C.18**: Example of GONG HAX classification assigned a CAO classification by SWPC. GONG processed images for Big Bear 2005/12/29 18:00:16 UT. This was SWPC region 0841 classified as CAO and GONG region B2005122501 classified as HAX. GONG has correctly classified this region based on the imagery available. SOON analysts were able to resolve 4 sunspots. This is possibly a case of where the limited GONG spatial resolution was not able to resolve very small sunspots.



(a) (b)

(c) (d)

**Figure C.19**: Example of GONG HKX classification assigned a FAI classification by SWPC. GONG processed images for Learmonth 2004/12/22 00:00:16 UT. This was SWPC region 0713 classified as FAI and GONG region L2004121800 classified as HKX. GONG has correctly classified this region based on the imagery available. L2004121800 is the region on the right in the images above. SOON analysts had combined these to region together, and this accounts for the F class given to it by SWPC.

## Appendix D: GONG – Solar Region Summary Magnetic Classification Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | α | β | βγ | γ | βδ | βγδ | γδ | Total GONG |
| α | Matches | 2455 | 2604 | 178 | 0 | 11 | 52 | 0 | 5300 |
| | %GONG | 46.32 | 49.13 | 3.36 | 0.00 | 0.21 | 0.98 | 0.00 | |
| | %SRS | 89.76 | 43.63 | 12.95 | 0.00 | 16.18 | 9.67 | 0.00 | |
| β | Matches | 145 | 2364 | 524 | 1 | 10 | 61 | 0 | 3105 |
| | %GONG | 4.67 | 76.14 | 16.88 | 0.03 | 0.32 | 1.96 | 0.00 | |
| | %SRS | 5.30 | 39.61 | 38.14 | 100.00 | 14.71 | 11.34 | 0.00 | |
| βγ | Matches | 34 | 580 | 323 | 0 | 9 | 67 | 0 | 1013 |
| | %GONG | 3.36 | 57.26 | 31.89 | 0.00 | 0.89 | 6.61 | 0.00 | |
| | %SRS | 1.24 | 9.72 | 23.51 | 0.00 | 13.24 | 12.45 | 0.00 | |
| γ | Matches | 32 | 329 | 227 | 0 | 8 | 67 | 2 | 665 |
| | %GONG | 4.81 | 49.47 | 34.14 | 0.00 | 1.20 | 10.08 | 0.30 | |
| | %SRS | 1.17 | 5.51 | 16.52 | 0.00 | 11.76 | 12.45 | 100.00 | |
| βδ | Matches | 44 | 40 | 18 | 0 | 8 | 26 | 0 | 136 |
| | %GONG | 32.35 | 29.41 | 13.24 | 0.00 | 5.88 | 19.12 | 0.00 | |
| | %SRS | 1.61 | 0.67 | 1.31 | 0.00 | 11.76 | 4.83 | 0.00 | |
| βγδ | Matches | 14 | 32 | 39 | 0 | 8 | 85 | 0 | 178 |
| | %GONG | 7.87 | 17.98 | 21.91 | 0.00 | 4.49 | 47.75 | 0.00 | |
| | %SRS | 0.51 | 0.54 | 2.84 | 0.00 | 11.76 | 15.80 | 0.00 | |
| γδ | Matches | 11 | 19 | 65 | 0 | 14 | 180 | 0 | 289 |
| | %GONG | 3.81 | 6.57 | 22.49 | 0.00 | 4.84 | 62.28 | 0.00 | |
| | %SRS | 0.40 | 0.32 | 4.73 | 0.00 | 20.59 | 33.46 | 0.00 | |
| | Total SRS | 2735 | 5968 | 1374 | 1 | 68 | 538 | 2 | |

**Table D.1**: GONG Learmonth – Solar Region Summary Magnetic Class Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | α | β | βγ | γ | βδ | βγδ | γδ | Total GONG |
| α | Matches | 1711 | 1765 | 105 | 0 | 12 | 27 | 0 | 3620 |
| | %GONG | 47.27 | 48.76 | 2.90 | 0.00 | 0.33 | 0.75 | 0.00 | |
| | %SRS | 90.67 | 40.62 | 10.68 | 0.00 | 22.64 | 7.85 | 0.00 | |
| β | Matches | 92 | 1764 | 348 | 0 | 2 | 30 | 0 | 2236 |
| | %GONG | 4.11 | 78.89 | 15.56 | 0.00 | 0.09 | 1.34 | 0.00 | |
| | %SRS | 4.88 | 40.60 | 35.40 | 0.00 | 3.77 | 8.72 | 0.00 | |
| βγ | Matches | 25 | 449 | 254 | 1 | 5 | 46 | 0 | 780 |
| | %GONG | 3.21 | 57.56 | 32.56 | 0.13 | 0.64 | 5.90 | 0.00 | |
| | %SRS | 1.32 | 10.33 | 25.84 | 100.00 | 9.43 | 13.37 | 0 | |
| γ | Matches | 14 | 275 | 197 | 0 | 3 | 43 | 0 | 532 |
| | %GONG | 2.63 | 51.69 | 37.03 | 0.00 | 0.56 | 8.08 | 0.00 | |
| | %SRS | 0.74 | 6.33 | 20.04 | 0.00 | 5.66 | 12.50 | 0.00 | |
| βδ | Matches | 30 | 37 | 18 | 0 | 6 | 19 | 0 | 110 |
| | %GONG | 27.27 | 33.64 | 16.36 | 0.00 | 5.45 | 17.27 | 0.00 | |
| | %SRS | 1.59 | 0.85 | 1.83 | 0.00 | 11.32 | 5.52 | 0.00 | |
| βγδ | Matches | 7 | 31 | 33 | 0 | 7 | 51 | 0 | 129 |
| | %GONG | 5.43 | 24.03 | 25.58 | 0.00 | 5.43 | 39.53 | 0.00 | |
| | %SRS | 0.37 | 0.71 | 3.36 | 0.00 | 13.21 | 14.83 | 0.00 | |
| γδ | Matches | 8 | 24 | 28 | 0 | 18 | 128 | 0 | 206 |
| | %GONG | 3.88 | 11.65 | 13.59 | 0.00 | 8.74 | 62.14 | 0.00 | |
| | %SRS | 0.42 | 0.55 | 2.85 | 0.00 | 33.96 | 37.21 | 0.00 | |
| | Total SRS | 1887 | 4345 | 983 | 1 | 53 | 344 | 0 | |

**Table D.2**: GONG Udaipur – Solar Region Summary Magnetic Class Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | α | β | βγ | γ | βδ | βγδ | γδ | Total GONG |
| α | Matches | 2218 | 2435 | 151 | 0 | 18 | 30 | 0 | 4852 |
| | %GONG | 45.71 | 50.19 | 3.11 | 0.00 | 0.37 | 0.62 | 0.00 | |
| | %SRS | 89.91 | 42.08 | 12.41 | 0.00 | 22.22 | 6.52 | 0.00 | |
| β | Matches | 118 | 2393 | 498 | 0 | 7 | 47 | 0 | 3063 |
| | %GONG | 3.85 | 78.13 | 16.26 | 0.00 | 0.23 | 1.53 | 0.00 | |
| | %SRS | 4.78 | 41.36 | 40.92 | 0.00 | 8.64 | 10.22 | 0.00 | |
| βγ | Matches | 41 | 547 | 255 | 1 | 12 | 70 | 0 | 926 |
| | %GONG | 4.43 | 59.07 | 27.54 | 0.11 | 1.30 | 7.56 | 0.00 | |
| | %SRS | 1.66 | 9.45 | 20.95 | 100.00 | 14.81 | 15.22 | 0.00 | |
| γ | Matches | 22 | 316 | 195 | 0 | 7 | 67 | 1 | 608 |
| | %GONG | 3.62 | 51.97 | 32.07 | 0.00 | 1.15 | 11.02 | 0.16 | |
| | %SRS | 0.89 | 5.46 | 16.02 | 0.00 | 8.64 | 14.57 | 100.00 | |
| βδ | Matches | 46 | 34 | 29 | 0 | 12 | 21 | 0 | 142 |
| | %GONG | 32.39 | 23.94 | 20.42 | 0.00 | 8.45 | 14.79 | 0.00 | |
| | %SRS | 1.86 | 0.59 | 2.38 | 0.00 | 14.81 | 4.57 | 0.00 | |
| βγδ | Matches | 9 | 37 | 36 | 0 | 8 | 72 | 0 | 162 |
| | %GONG | 5.56 | 22.84 | 22.22 | 0.00 | 4.94 | 44.44 | 0.00 | |
| | %SRS | 0.36 | 0.64 | 2.96 | 0.00 | 9.88 | 15.65 | 0.00 | |
| γδ | Matches | 13 | 24 | 53 | 0 | 17 | 153 | 0 | 260 |
| | %GONG | 5.00 | 9.23 | 20.38 | 0.00 | 6.54 | 58.85 | 0.00 | |
| | %SRS | 0.53 | 0.41 | 4.35 | 0.00 | 20.99 | 33.26 | 0.00 | |
| | Total SRS | 2467 | 5786 | 1217 | 1 | 81 | 460 | 1 | |

**Table D.3**: GONG Teide – Solar Region Summary Magnetic Class Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | α | β | βγ | γ | βδ | βγδ | γδ | Total GONG |
| α | Matches | 2418 | 2519 | 146 | 0 | 14 | 32 | 0 | 5129 |
| | %GONG | 47.14 | 49.11 | 2.85 | 0.00 | 0.27 | 0.62 | 0.00 | |
| | %SRS | 90.26 | 40.86 | 10.99 | 0.00 | 17.72 | 6.58 | 0.00 | |
| β | Matches | 134 | 2494 | 501 | 1 | 9 | 45 | 0 | 3184 |
| | %GONG | 4.21 | 78.33 | 15.73 | 0.03 | 0.28 | 1.41 | 0.00 | |
| | %SRS | 5.00 | 40.45 | 37.73 | 100.00 | 11.39 | 9.26 | 0.00 | |
| βγ | Matches | 31 | 637 | 320 | 0 | 11 | 54 | 0 | 1053 |
| | %GONG | 2.94 | 60.49 | 30.39 | 0.00 | 1.04 | 5.13 | 0.00 | |
| | %SRS | 1.16 | 10.33 | 24.10 | 0.00 | 13.92 | 11.11 | 0.00 | |
| γ | Matches | 26 | 382 | 224 | 0 | 7 | 61 | 1 | 701 |
| | %GONG | 3.71 | 54.49 | 31.95 | 0.00 | 1.00 | 8.70 | 0.14 | |
| | %SRS | 0.97 | 6.20 | 16.87 | 0.00 | 8.86 | 12.55 | 100.00 | |
| βδ | Matches | 43 | 55 | 27 | 0 | 5 | 31 | 0 | 161 |
| | %GONG | 26.71 | 34.16 | 16.77 | 0.00 | 3.11 | 19.25 | 0.00 | |
| | %SRS | 1.61 | 0.89 | 2.03 | 0.00 | 6.33 | 6.38 | 0.00 | |
| βγδ | Matches | 16 | 47 | 55 | 0 | 13 | 74 | 0 | 205 |
| | %GONG | 7.80 | 22.93 | 26.83 | 0.00 | 6.34 | 36.10 | 0.00 | |
| | %SRS | 0.60 | 0.76 | 4.14 | 0.00 | 16.46 | 15.23 | 0.00 | |
| γδ | Matches | 11 | 31 | 55 | 0 | 20 | 189 | 0 | 306 |
| | %GONG | 3.59 | 10.13 | 17.97 | 0.00 | 6.54 | 61.76 | 0.00 | |
| | %SRS | 0.41 | 0.50 | 4.14 | 0.00 | 25.32 | 38.89 | 0.00 | |
| | Total SRS | 2679 | 6165 | 1328 | 1 | 79 | 486 | 1 | |

**Table D.4**: GONG Cerra Tololo– Solar Region Summary Magnetic Class Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | α | β | βγ | γ | βδ | βγδ | γδ | Total GONG |
| α | Matches | 2133 | 2507 | 152 | 0 | 15 | 40 | 0 | 4847 |
| | %GONG | 44.01 | 51.72 | 3.14 | 0.00 | 0.31 | 0.83 | 0.00 | |
| | %SRS | 89.89 | 43.15 | 12.11 | 0.00 | 18.75 | 8.62 | 0.00 | |
| β | Matches | 111 | 2281 | 511 | 1 | 9 | 51 | 0 | 2964 |
| | %GONG | 3.74 | 76.96 | 17.24 | 0.03 | 0.30 | 1.72 | 0.00 | |
| | %SRS | 4.68 | 39.26 | 40.72 | 100.00 | 11.25 | 10.99 | 0.00 | |
| βγ | Matches | 38 | 576 | 271 | 0 | 13 | 59 | 0 | 957 |
| | %GONG | 3.97 | 60.19 | 28.32 | 0.00 | 1.36 | 6.17 | 0.00 | |
| | %SRS | 1.60 | 9.91 | 21.59 | 0.00 | 16.25 | 12.72 | 0.00 | |
| γ | Matches | 26 | 328 | 197 | 0 | 7 | 67 | 2 | 627 |
| | %GONG | 4.15 | 52.31 | 31.42 | 0.00 | 1.12 | 10.69 | 0.32 | |
| | %SRS | 1.10 | 5.65 | 15.70 | 0.00 | 8.75 | 14.44 | 100.00 | |
| βδ | Matches | 46 | 55 | 25 | 0 | 11 | 20 | 0 | 157 |
| | %GONG | 29.30 | 35.03 | 15.92 | 0.00 | 7.01 | 12.74 | 0.00 | |
| | %SRS | 1.94 | 0.95 | 1.99 | 0.00 | 13.75 | 4.31 | 0.00 | |
| βγδ | Matches | 12 | 30 | 48 | 0 | 13 | 74 | 0 | 177 |
| | %GONG | 6.78 | 16.95 | 27.12 | 0.00 | 7.34 | 41.81 | 0.00 | |
| | %SRS | 0.51 | 0.52 | 3.82 | 0.00 | 16.25 | 15.95 | 0.00 | |
| γδ | Matches | 7 | 33 | 51 | 0 | 12 | 153 | 0 | 256 |
| | %GONG | 2.73 | 12.89 | 19.92 | 0.00 | 4.69 | 59.77 | 0.00 | |
| | %SRS | 0.29 | 0.57 | 4.06 | 0.00 | 15.00 | 32.97 | 0.00 | |
| | Total SRS | 2373 | 5810 | 1255 | 1 | 80 | 464 | 2 | |

**Table D.5**: GONG Big Bear – Solar Region Summary Magnetic Class Comparison

| | | SRS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| GONG | | α | β | βγ | γ | βδ | βγδ | γδ | Total GONG |
| α | Matches | 2087 | 2308 | 137 | 0 | 15 | 41 | 0 | 4588 |
| | %GONG | 45.49 | 50.31 | 2.99 | 0.00 | 0.33 | 0.89 | 0.00 | |
| | %SRS | 89.19 | 40.75 | 11.64 | 0.00 | 20.00 | 8.99 | 0.00 | |
| β | Matches | 122 | 2282 | 433 | 0 | 8 | 46 | 0 | 2891 |
| | %GONG | 4.22 | 78.93 | 14.98 | 0.00 | 0.28 | 1.59 | 0.00 | |
| | %SRS | 5.21 | 40.29 | 36.79 | 0.00 | 10.67 | 10.09 | 0.00 | |
| βγ | Matches | 48 | 626 | 270 | 1 | 10 | 57 | 0 | 1012 |
| | %GONG | 4.74 | 61.86 | 26.68 | 0.10 | 0.99 | 5.63 | 0.00 | |
| | %SRS | 2.05 | 11.05 | 22.94 | 100.00 | 13.33 | 12.50 | 0.00 | |
| γ | Matches | 29 | 332 | 197 | 0 | 8 | 55 | 1 | 622 |
| | %GONG | 4.66 | 53.38 | 31.67 | 0.00 | 1.29 | 8.84 | 0.16 | |
| | %SRS | 1.24 | 5.86 | 16.74 | 0.00 | 10.67 | 12.06 | 100.00 | |
| βδ | Matches | 33 | 49 | 27 | 0 | 11 | 22 | 0 | 142 |
| | %GONG | 23.24 | 34.51 | 19.01 | 0.00 | 7.75 | 15.49 | 0.00 | |
| | %SRS | 1.41 | 0.87 | 2.29 | 0.00 | 14.67 | 4.82 | 0.00 | |
| βγδ | Matches | 16 | 30 | 53 | 0 | 9 | 81 | 0 | 189 |
| | %GONG | 8.47 | 15.87 | 28.04 | 0.00 | 4.76 | 42.86 | 0.00 | |
| | %SRS | 0.68 | 0.53 | 4.50 | 0.00 | 12.00 | 17.76 | 0.00 | |
| γδ | Matches | 5 | 37 | 60 | 0 | 14 | 154 | 0 | 270 |
| | %GONG | 1.85 | 13.70 | 22.22 | 0.00 | 5.19 | 57.04 | 0.00 | |
| | %SRS | 0.21 | 0.65 | 5.10 | 0.00 | 18.67 | 33.77 | 0.00 | |
| | Total SRS | 2340 | 5664 | 1177 | 1 | 75 | 456 | 1 | |

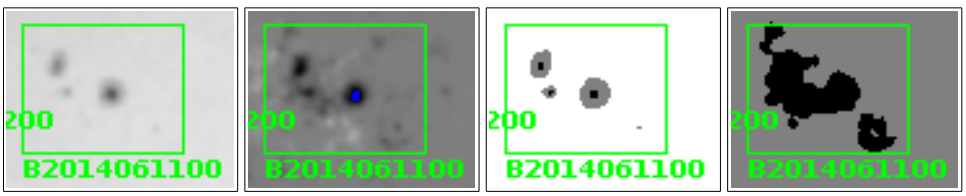
**Table D.6**: GONG Mauna Loa – Solar Region Summary Magnetic Class Comparison

## Appendix E: Examples of Magnetic Classification Disagreement Between Automated GONG and the Solar Region Summary

In all of the following figures, (a) is the normalised intensity image, (b) is the magnetogram, (c) is the sunspot image, and (d) is the magnetic area image. In all cases for sunspots, grey is penumbra and black is umbra. For magnetic fields black is negative (into the sun) and white is positive (out of the sun). Heliographic north is down and heliographic west is to the right.

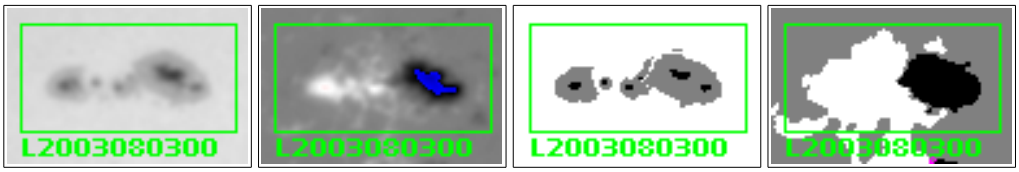

(a) (b) (c) (d)

**Figure E.1**: Example of GONG α classification assigned a β classification by SWPC. GONG processed images for Udaipur 2012/01/27 04:00:16 UT. This was SWPC region 1408 classified as β and GONG region U2012012200 classified as α. There is only negative field present. Based on the GONG imagery this classification is correct. SOON analysts were able to resolve 5 sunspots using their equipment and identified this region as bipolar.



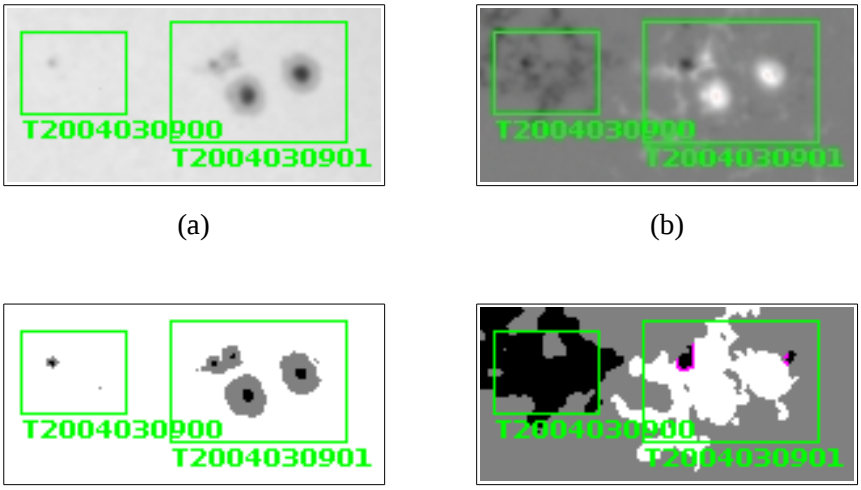

(a) (b) (c) (d)

**Figure E.2**: Example of GONG α classification assigned a βγ classification by SWPC. GONG processed images for Big Bear 2014/06/14 15:00:16 UT. This was SWPC region 2087 classified as α and GONG region B2014061100 classified as βγ. GONG was unable to resolve any spots on the positive field in the lower portion of the region, so classified this a unipolar region. SOON analysts were able to resolve 24 sunspots using their equipment and identified this region as bipolar.




(a) (b) (c) (d)

**Figure E.3**: Example of GONG α classification assigned a βδ classification by SWPC. GONG processed images for Learmonth 2003/08/07 01:00:16 UT. This was SWPC region 0424 classified as βδ and GONG region L2003080300 classified as α. This is an example of where the GONG classification system has completely failed. From the imagery this is clearly a βδ region, and suggests that the algorithm still needs some adjustment.



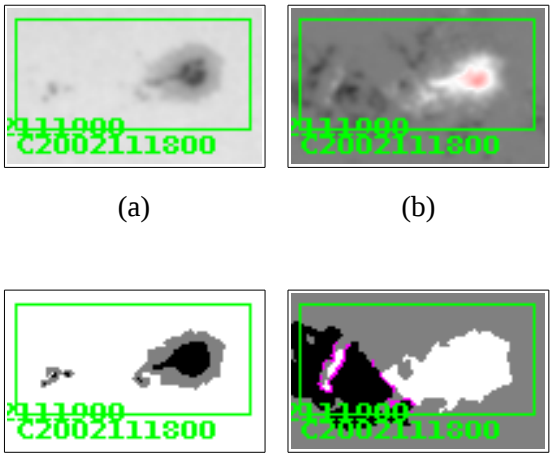

(a) (b)

(c) (d)

**Figure E.4**: Example of GONG α classification assigned a βγδ classification by SWPC. GONG processed images for Teide 2004/03/13 09:00:16 UT. This was SWPC region 0570 classified as βγδ and GONG region T2004030900 classified as α. This is an example of where the GONG algorithm split a region in 2. From the magmaps, the SOON analysts have probably classified this region correctly, keeping the two polarities together.


(a) (b) (c) (d)

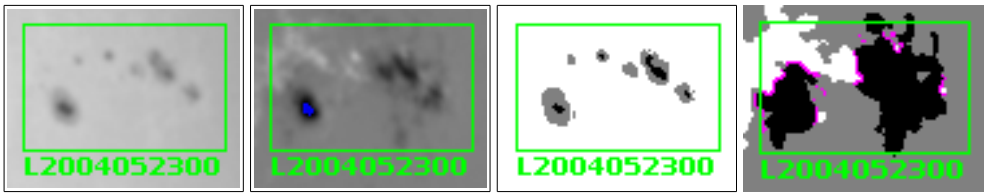
**Figure E.5**: Example of GONG β classification assigned a α classification by SWPC. GONG processed images for Learmonth 2006/09/11 01:00:16 UT. This was SWPC region 0908 classified as α and GONG region L2006090700 classified as β. There is a very small portion of negative field present which caused this region to be bipolar. SOON analysts were able to resolve 4 sunspots with their equipment, but were unable to determine the small area of negative field.




(a) (b)

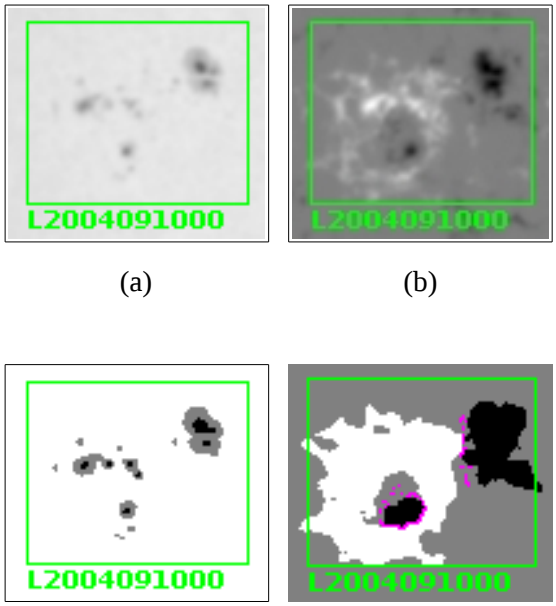
(c) (d)

**Figure E.6**: Example of GONG β classification assigned a βγ classification by SWPC. GONG processed images for Cerro Tololo 2002/11/21 13:00:16 UT. This was SWPC region 0198 classified as βγ and GONG region C2002111800 classified as β. This is clearly at least a β region. It is reasonable for SOON analysts to classify this as a βγ.




(a) (b) (c) (d)

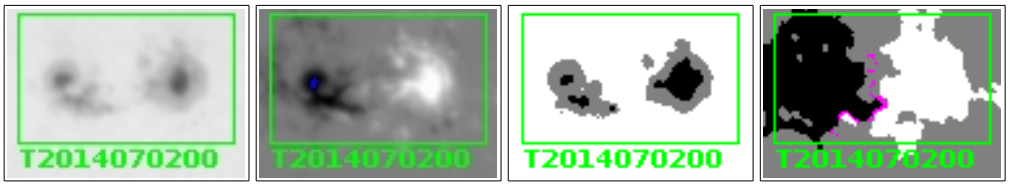
**Figure E.7**: Example of GONG β classification assigned a γ classification by SWPC. GONG processed images for Learmonth 2004/05/29 01:00:16 UT. This was SWPC region 0618 classified as β and GONG region L2004052300 classified as γ. This clearly at least a β region. It is reasonable for SOON analysts to classify this as a γ due to the distribution of polarities through out the region.




(a) (b)

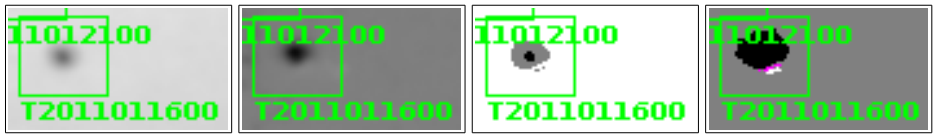
(c) (d)

**Figure E.8**: Example of GONG β classification assigned a βδ classification by SWPC. GONG processed images for Learmonth 2004/09/15 00:00:16 UT. This was SWPC region 0672 classified as βδ and GONG region L2004091000 classified as β. This clearly at least a β region. From the imagery there is no umbra of opposite polarity visible within the same penumbra. SOON analysts may have been able to resolve fainter penumbra particularly for the sunspots in the lower portion of the region.



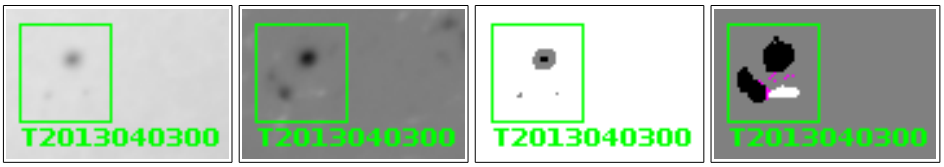

(a) (b) (c) (d)

**Figure E.9**: Example of GONG β classification assigned a βγδ classification by SWPC. GONG processed images for Teide 2014/07/07 08:00:16 UT. This was SWPC region 2108 classified as βγδ and GONG region T2014070200 classified as β. This is clearly at least a β region. From the imagery there is no umbra of opposite polarity visible within the same penumbra. There is no significant mixing of polarities present. Therefore, this is a reasonable classification by GONG.



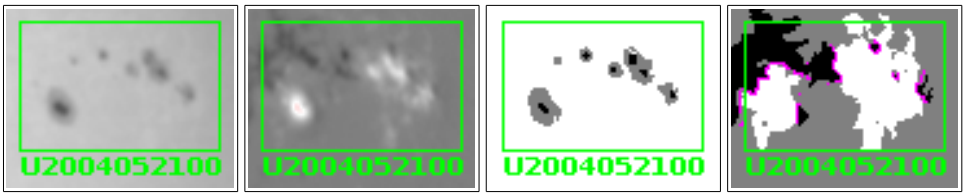

(a) (b) (c) (d)

**Figure E.10**: Example of GONG βγ classification assigned a α classification by SWPC. GONG processed images for Teide 2011/01/21 10:00:16 UT. This was SWPC region 1147 classified as α and GONG region T2011011600 classified as βγ. In this example there is a very small portion of magnetic field in the lower portion of the region. Because this is oriented north-south, the region was determined to have sufficient polarity mixing to be considered a βγ. The SOON analyst classification is probably correct in this case.




(a) (b) (c) (d)

**Figure E.11**: Example of GONG βγ classification assigned a β classification by SWPC. GONG processed images for Teide 2013/04/09 08:00:16 UT. This was SWPC region 1714 classified as β and GONG region T2013040300 classified as βγ. This clearly at least a β region. As with figure E.10, it is the orientation of the region that caused the GONG algorithm to classify this as βγ. The SOON analyst classification is probably correct in this case.

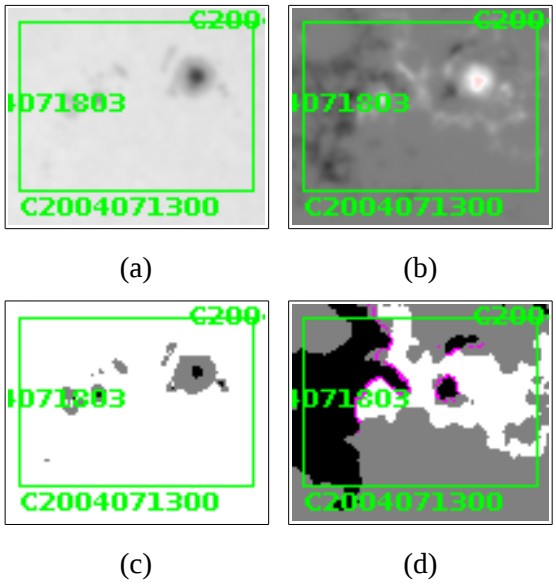



(a) (b) (c) (d)

**Figure E.12**: Example of GONG βγ classification assigned a γ classification by SWPC. GONG processed images for Udaipur 2004/05/29 03:00:16 UT. This was SWPC region 0618 classified as γ and GONG region U2004052100 classified as βγ. Either classification of βγ or γ could be considered correct for this region as there is significant mixing of polarities.


(a) (b) (c) (d)

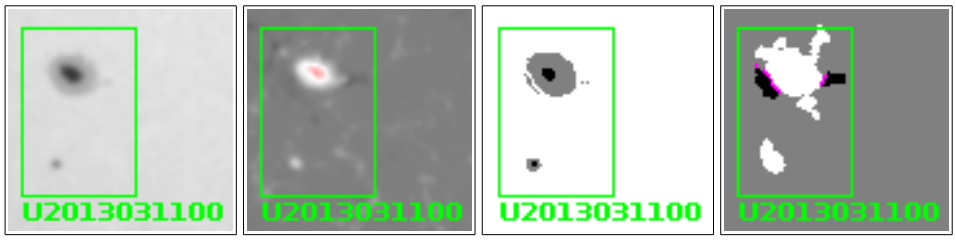
**Figure E.13**: Example of GONG βγ classification assigned a βδ classification by SWPC. GONG processed images for Udaipur 2013/11/07 05:00:16 UT. This was SWPC region 1891 classified as βδ and GONG region U2013110600 classified as βγ. In the GONG imagery no umbra of opposite polarity are within the same penumbra. Based on this, the GONG classification is correct.




(a) (b)

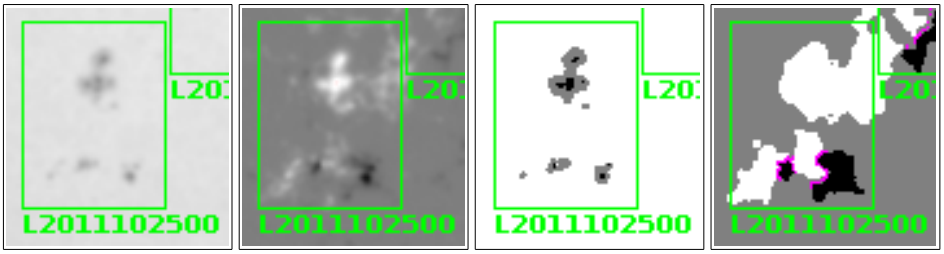
(c) (d)

**Figure E.14**: Example of GONG βγ classification assigned a βγδ classification by SWPC. GONG processed images for Cerro Tololo 2004/07/19 16:00:16 UT. This was SWPC region 0649 classified as βγδ and GONG region C2004071300 classified as βγ. As with figure E.13, there is no sunspot with a d configuration present, so based on this imagery the GONG classification is correct.



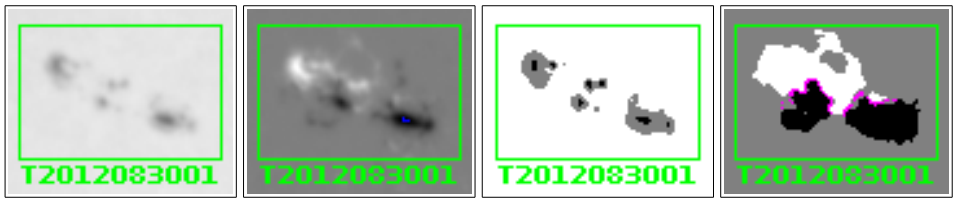

(a) (b) (c) (d)

**Figure E.15**: Example of GONG γ classification assigned a α classification by SWPC. GONG processed images for Udaipur 2013/03/14 03:00:16 UT. This was SWPC region 1692 classified as α and GONG region U2013031100 classified as γ. As with figure E.10, the γ classification comes about due to polarity distribution. In this case the region possibly should have been split into 2 α regions. The other issue here is that there is negative field directly surrounding the positive field towards the centre of the region, thus giving this region a bipolar classification.



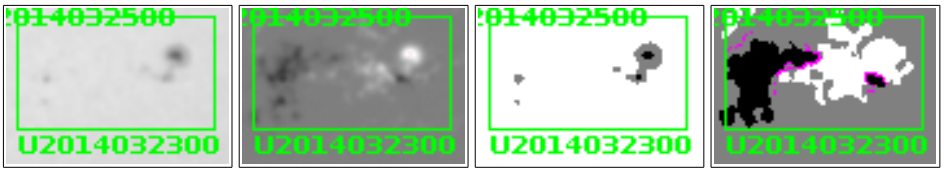

(a) (b) (c) (d)

**Figure E.16**: Example of GONG γ classification assigned a β classification by SWPC. GONG processed images for Learmonth 2011/10/28 00:00:16 UT. This was SWPC region 1333 classified as β and GONG region L2011102500 classified as γ. This is clearly a bipolar region, but as with earlier examples, negative field situated below the positive field causes the GONG algorithm to classify this as γ. The SOON analysts probably have the better classification of β.



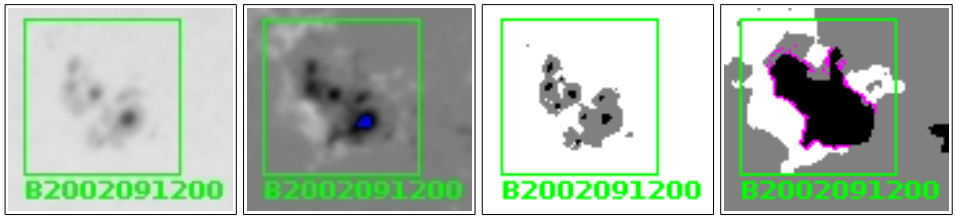

(a) (b) (c) (d)

**Figure E.17**: Example of GONG γ classification assigned a βγ classification by SWPC. GONG processed images for Teide 2012/09/01 08:00:16 UT. This was SWPC region 1560 classified as βγ and GONG region T2012083001 classified as γ. The orientation of the magnetic fields in the region support either a βγ or γ classification.



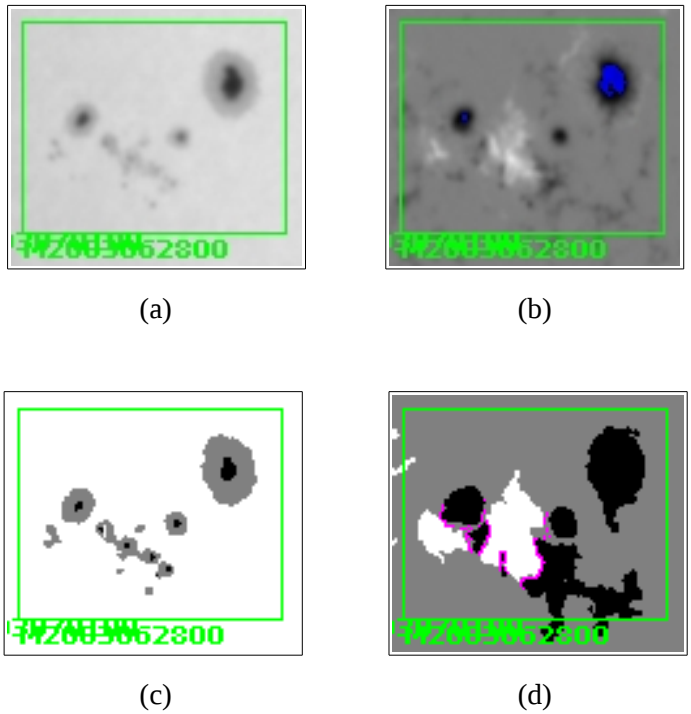

(a) (b) (c) (d)

**Figure E.18**: Example of GONG γ classification assigned a βδ classification by SWPC. GONG processed images for Udaipur 2014/03/28 03:00:16 UT. This was SWPC region 2017 classified as βδ and GONG region U2014032300 classified as γ. This region shows some mixing of polarities. However, from the imagery there is no δ configuration present in any of the sunspots.




(a) (b) (c) (d)

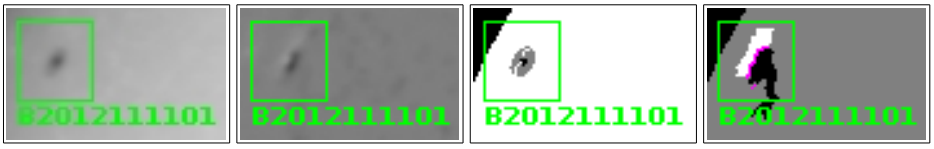
**Figure E.19**: Example of GONG γ classification assigned a βγδ classification by SWPC. GONG processed images for Big Bear 2002/09/14 15:00:16 UT. This was SWPC region 0114 classified as βγδ and GONG region B2002091200 classified as γ. This region shows some mixing of polarities, but again a δ configuration as reported by SOON analysts cannot be seen on this imagery.



(a) (b)

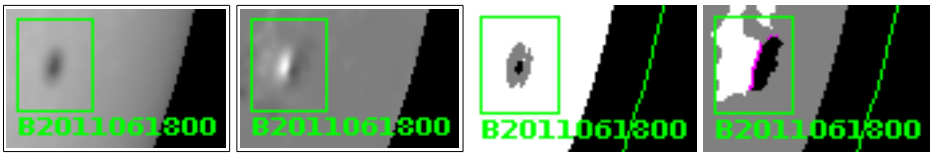
(c) (d)

**Figure E.20**: Example of GONG γ classification assigned a γδ classification by SWPC. GONG processed images for Mauna Loa 2003/07/01 18:00:16 UT. This was SWPC region 0397 classified as γδ and GONG region M2003062800 classified as γ. This is clearly a γ region. It is possible that GONG has not fully resolved penumbra in this case, thus not allocating a δ configuration to this region.



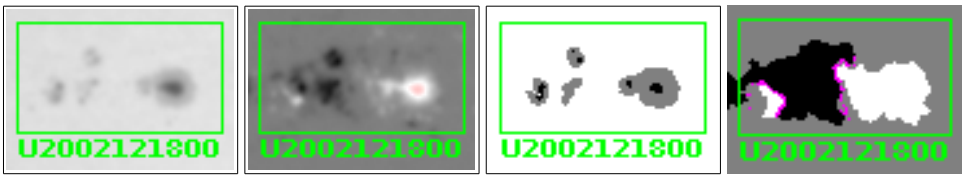

(a) (b) (c) (d)

**Figure E.21**: Example of GONG βδ classification assigned a α classification by SWPC. GONG processed images for Big Bear 2012/11/11 17:00:16 UT. This was SWPC region 1613 classified as α and GONG region B2012111101 classified as βδ. This is an example where a region close to the limb was probably mis-classified. Based on the imagery, this should probably be a β region. However as there is only one sunspot, the SOON analyst classified this correctly.



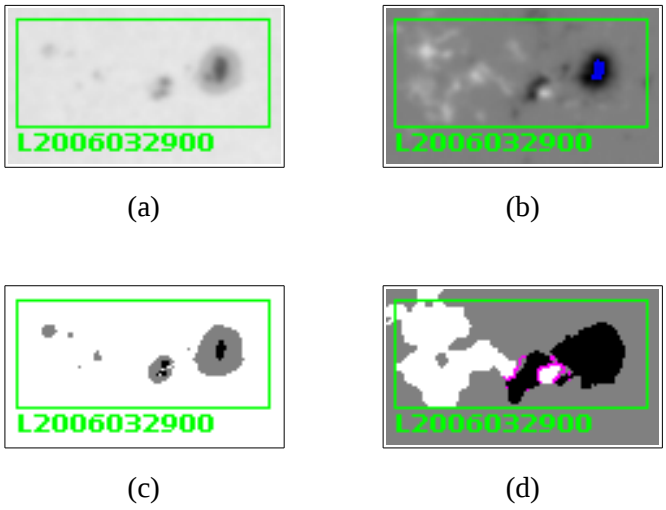

(a) (b) (c) (d)

**Figure E.22**: Example of GONG βδ classification assigned a β classification by SWPC. GONG processed images for Big Bear 2011/06/24 15:00:16 UT. This was SWPC region 1236 classified as β and GONG region B2011061800 classified as βδ. As with figure E.21, this is an example where a region close to the limb was probably mis-classified. Based on the imagery, this should probably be an α region as there is only one spot visible. The SOON analysts were able to resolve 4 sunspots of opposite polarity.




(a) (b) (c) (d)

**Figure E.23**: Example of GONG βδ classification assigned a βγ classification by SWPC. GONG processed images for Udaipur 2002/12/21 04:00:16 UT. This was SWPC region 0230 classified as βγ and GONG region U2002121800 classified as βδ. Here, the left-most penumbra has umbra of opposite polarity with it. GONG was correct on classifying this region using the above imagery.



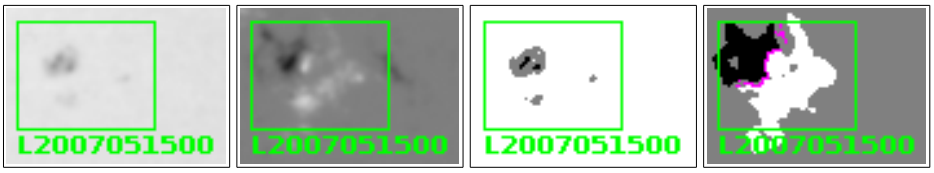

(a) (b)

(c) (d)

**Figure E.24**: Example of GONG βδ classification assigned a βγδ classification by SWPC. GONG processed images for Learmonth 2006/04/02 00:00:16 UT. This was SWPC region 0865 classified as βγδ and GONG region L2006032900 classified as βδ. Either classification is probably acceptable in this example as polarities are reasonably mixed and there is a sunspot with both polarities towards the centre of the region.

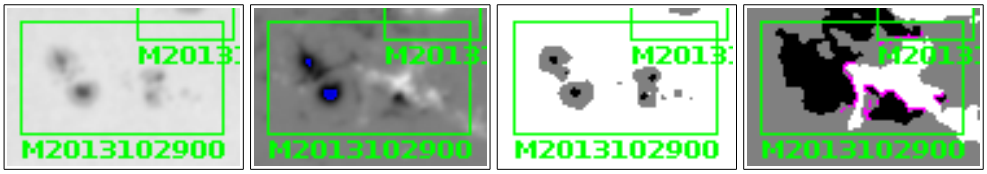

(a) (b) (c) (d)

**Figure E.25**: Example of GONG βγδ classification assigned a α classification by SWPC. GONG processed images for Cerro Tololo 2013/12/20 13:00:16 UT. This was SWPC region 1921 classified as α and GONG region C2013121000 classified as βγδ. This is an example of where the limb correction process has not worked well. Based on the imagery the region probably should be classified as β as there are two distinct sunspots. The SOON analysts were able to identify 2 sunspots with their equipment, but did not recognise they were of opposite polarities.



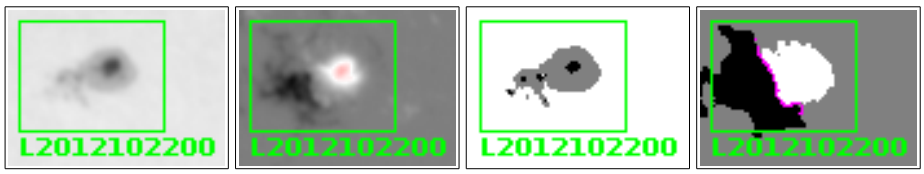

(a) (b) (c) (d)

**Figure E.26**: Example of GONG βγδ classification assigned a β classification by SWPC. GONG processed images for Learmonth 2007/05/20 04:00:16 UT. This was SWPC region 0956 classified as β and GONG region L2007051500 classified as βγδ. Based on the imagery, this classification seems reasonable. The SOON analyst did not identify the opposite polarity umbra.



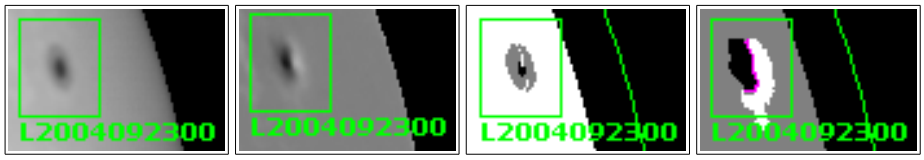

(a) (b) (c) (d)

**Figure E.27**: Example of GONG βγδ classification assigned a βγ classification by SWPC. GONG processed images for Mauna Loa 2013/11/01 19:00:16 UT. This was SWPC region 1884 classified as βγ and GONG region M2013102900 classified as βγδ. The SOON analysts failed to identify the bipolar umbra to the right of the region. Based on this imagery the GONG classification is correct.




(a) (b) (c) (d)

**Figure E.28**: Example of GONG βγδ classification assigned a βδ classification by SWPC. GONG processed images for Learmonth 2012/10/27 01:00:16 UT. This was SWPC region 1598 classified as βδ and GONG region L2012102200 classified as βγδ. The SOON analysts probably have the better classification as there is little mixing of polarities.

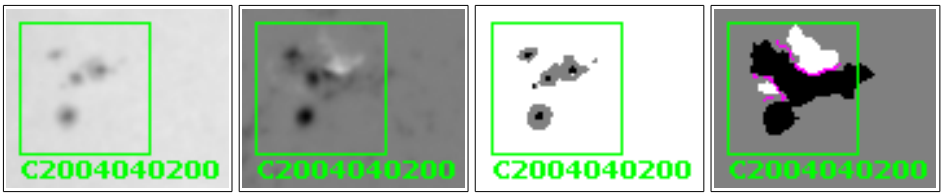



(a) (b) (c) (d)

**Figure E.29**: Example of GONG γδ classification assigned a α classification by SWPC. GONG processed images for Learmonth 2004/09/26 00:00:16 UT. This was SWPC region 0673 classified as α and GONG region L2004092300 classified as γδ. This region should be classified as α. This is an example where GONG analysis has difficulties near the limb of the disk. The SOON classification is correct.

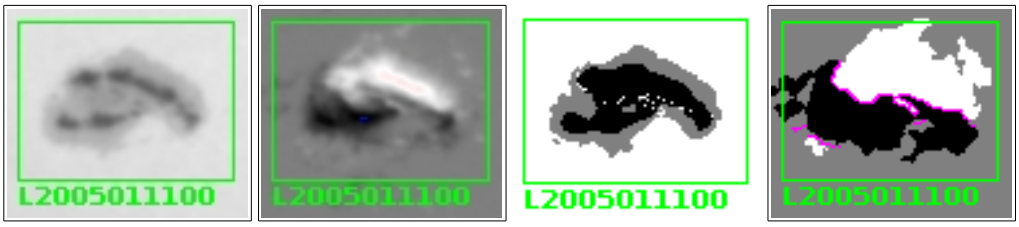

(a) (b) (c) (d)

**Figure E.30**: Example of GONG γδ classification assigned a β classification by SWPC. GONG processed images for Big Bear 2012/08/09 15:00:16 UT. This was SWPC region 1544 classified as β and GONG region B2012080900 classified as γδ. In this case there is minimal mixing and it appears that there are no sunspots of mixed polarity. Therefore the SOON analysts classified this region correctly.



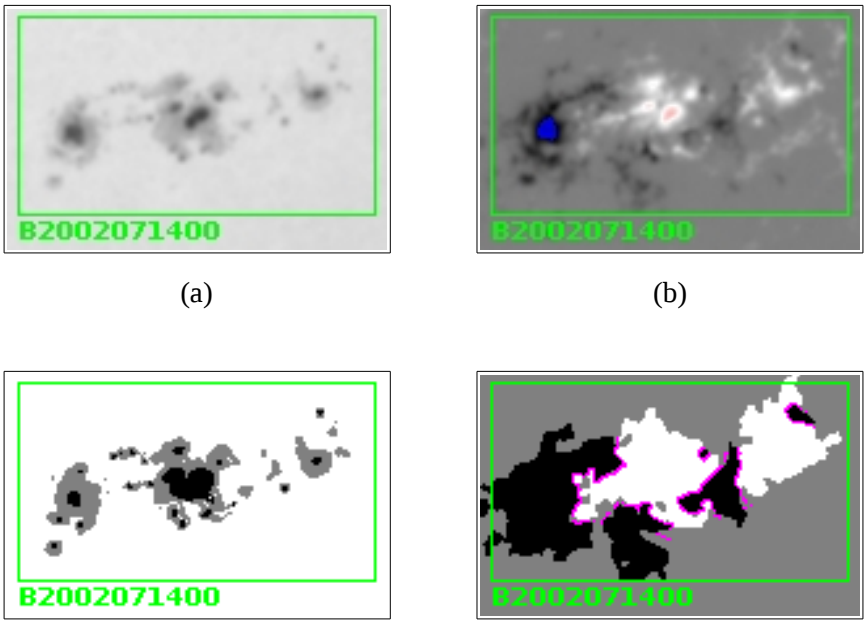

(a) (b) (c) (d)

**Figure E.31**: Example of GONG γδ classification assigned a βγ classification by SWPC. GONG processed images for Cerro Tololo 2004/04/05 13:00:16 UT. This was SWPC region 0588 classified as βγ and GONG region C2004040200 classified as γδ. From the intensity image, it is difficult to determine if the two umbra in the centre of the region are in the same penumbra. The GONG algorithm may have been in error in this case.




(a) (b) (c) (d)

**Figure E.32**: Example of GONG γδ classification assigned a βδ classification by SWPC. GONG processed images for Learmonth 2005/01/15 01:00:16 UT. This was SWPC region 0720 classified as βδ and GONG region L2005011100 classified as γδ. It is obvious that this region is a δ region. This was assigned a γδ classification by the GONG algorithm due to the north-south orientation of the region.




(a) (b)

(c) (d)

**Figure E.33**: Example of GONG γδ classification assigned a βγδ classification by SWPC. GONG processed images for Big Bear 2002/07/16 UT. This was SWPC region 1350 classified as βγδ and GONG region B2002071400 classified as γδ. In this example either classification could be considered correct. Due to the distribution of magnetic fields a γδ classification is reasonable.

## Appendix F: SWPC Flare Rates by Sunspot Region Classification

| | Cycle | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 20 | 21 | | | 22 | | | 23 | | | 24 | | |
| Class | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate |
| AXX | 0.010 | 3 | 843 | 0.004 | 7 | 3141 | 0.002 | 7 | 2140 | 0.003 | 1 | 818 | 0.001 |
| BX | 0.080 | 6 | 1390 | 0.004 | 56 | 4854 | 0.012 | -- | -- | -- | -- | -- | -- |
| HRX | 0.040 | 0 | 85 | 0.000 | 1 | 207 | 0.005 | 1 | 396 | 0.003 | 1 | 236 | 0.004 |
| HSX | 0.050 | 11 | 888 | 0.012 | 33 | 2409 | 0.014 | 21 | 2877 | 0.007 | 8 | 1312 | 0.006 |
| HAX | 0.060 | 6 | 166 | 0.036 | 9 | 675 | 0.013 | 10 | 835 | 0.012 | 1 | 351 | 0.003 |
| HHX | 0.120 | 13 | 116 | 0.112 | 2 | 130 | 0.015 | 2 | 92 | 0.022 | 0 | 36 | 0.000 |
| HKX | 0.180 | 0 | 44 | 0.000 | 17 | 102 | 0.167 | 3 | 96 | 0.031 | 1 | 25 | 0.040 |
| CRO | 0.060 | 4 | 323 | 0.012 | 5 | 589 | 0.008 | 5 | 704 | 0.007 | 5 | 466 | 0.011 |
| CRI | 0.050 | 0 | 9 | 0.000 | 0 | 15 | 0.000 | 0 | 2 | 0.000 | 0 | 20 | 0.000 |
| CSO | 0.040 | 16 | 636 | 0.025 | 34 | 1649 | 0.021 | 40 | 2325 | 0.017 | 14 | 740 | 0.019 |
| CSI | 0.090 | 0 | 9 | 0.000 | 5 | 29 | 0.172 | 0 | 7 | 0.000 | 2 | 29 | 0.069 |
| CAO | 0.080 | 7 | 332 | 0.021 | 41 | 1557 | 0.026 | 34 | 1214 | 0.028 | 9 | 560 | 0.016 |
| CAI | 0.110 | 0 | 7 | 0.000 | 3 | 46 | 0.065 | 2 | 23 | 0.087 | 2 | 56 | 0.036 |

**Table F.1a**: SWPC Flare Rates by Sunspot Region Classification

| | Cycle | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 20 | 21 | | | 22 | | | 23 | | | 24 | | |
| Class | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate |
| CHO | 0.080 | 0 | 76 | 0.000 | 3 | 146 | 0.021 | 2 | 75 | 0.027 | 0 | 27 | 0.000 |
| CHI | 0.210 | 0 | 5 | 0.000 | 0 | 3 | 0.000 | 0 | 3 | 0.000 | 0 | 2 | 0.000 |
| CKO | 0.290 | 4 | 80 | 0.050 | 38 | 217 | 0.175 | 10 | 134 | 0.075 | 1 | 35 | 0.029 |
| CKI | 0.220 | 3 | 13 | 0.231 | 4 | 20 | 0.200 | 1 | 7 | 0.143 | 0 | 4 | 0.000 |
| DRO | 0.080 | 1 | 69 | 0.014 | 1 | 71 | 0.014 | 1 | 150 | 0.007 | 1 | 115 | 0.009 |
| ERO | 0.000 | 0 | 8 | 0.000 | 0 | 3 | 0.000 | 0 | 3 | 0.000 | 0 | 4 | 0.000 |
| FRO | 0.000 | -- | -- | -- | -- | -- | -- | 0 | 1 | 0.000 | -- | -- | -- |
| DSO | 0.100 | 23 | 436 | 0.053 | 34 | 695 | 0.049 | 35 | 1652 | 0.021 | 24 | 554 | 0.043 |
| ESO | 0.170 | 2 | 83 | 0.024 | 10 | 170 | 0.059 | 8 | 255 | 0.031 | 4 | 80 | 0.050 |
| FSO | 0.540 | 0 | 5 | 0.000 | 0 | 26 | 0.000 | 5 | 49 | 0.102 | 0 | 16 | 0.000 |
| DAO | 0.100 | 27 | 445 | 0.061 | 135 | 1877 | 0.072 | 90 | 2053 | 0.044 | 14 | 584 | 0.024 |
| EAO | 0.300 | 13 | 94 | 0.138 | 70 | 534 | 0.131 | 50 | 582 | 0.086 | 3 | 66 | 0.045 |
| FAO | 0.000 | 3 | 14 | 0.214 | 16 | 82 | 0.195 | 10 | 98 | 0.102 | 1 | 6 | 0.167 |
| DHO | 0.260 | 4 | 38 | 0.105 | 5 | 66 | 0.076 | 3 | 49 | 0.061 | 1 | 50 | 0.020 |
| EHO | 0.150 | 2 | 40 | 0.050 | 4 | 52 | 0.077 | 2 | 29 | 0.069 | 1 | 31 | 0.032 |

**Table F.1b**: SWPC Flare Rates by Sunspot Region Classification Continued

| | Cycle | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 20 | 21 | | | 22 | | | 23 | | | 24 | | |
| Class | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate |
| FHO | 0.000 | 1 | 8 | 0.125 | 2 | 24 | 0.083 | 3 | 17 | 0.176 | 1 | 9 | 0.111 |
| DKO | 0.370 | 16 | 112 | 0.143 | 52 | 261 | 0.199 | 31 | 186 | 0.167 | 13 | 67 | 0.194 |
| EKO | 0.400 | 29 | 100 | 0.290 | 51 | 265 | 0.192 | 25 | 154 | 0.162 | 2 | 38 | 0.053 |
| FKO | 0.320 | 14 | 21 | 0.667 | 36 | 122 | 0.295 | 13 | 69 | 0.188 | 5 | 16 | 0.313 |
| DRI | 0.150 | 1 | 8 | 0.125 | 0 | 9 | 0.000 | 0 | 6 | 0.000 | 0 | 15 | 0.000 |
| ERI | 0.000 | 0 | 2 | 0.000 | -- | -- | -- | -- | -- | -- | 0 | 1 | 0.000 |
| FRI | 0.500 | -- | -- | -- | 0 | 1 | 0.000 | -- | -- | -- | -- | -- | -- |
| DSI | 0.130 | 7 | 47 | 0.149 | 8 | 81 | 0.099 | 7 | 86 | 0.081 | 3 | 141 | 0.021 |
| ESI | 0.200 | 1 | 14 | 0.071 | 6 | 46 | 0.130 | 4 | 63 | 0.063 | 1 | 67 | 0.015 |
| FSI | 1.880 | 0 | 5 | 0.000 | 1 | 9 | 0.111 | 2 | 18 | 0.111 | 3 | 5 | 0.600 |
| DAI | 0.200 | 23 | 99 | 0.232 | 70 | 438 | 0.160 | 47 | 448 | 0.105 | 15 | 357 | 0.042 |
| EAI | 0.600 | 15 | 58 | 0.259 | 61 | 338 | 0.180 | 61 | 446 | 0.137 | 8 | 113 | 0.071 |
| FAI | 0.670 | 2 | 9 | 0.222 | 46 | 104 | 0.442 | 24 | 142 | 0.169 | 0 | 2 | 0.000 |
| DHI | 0.070 | 11 | 21 | 0.524 | 3 | 19 | 0.158 | 0 | 9 | 0.000 | 2 | 23 | 0.087 |
| EHI | 0.710 | 1 | 21 | 0.048 | 6 | 45 | 0.133 | 7 | 27 | 0.259 | 0 | 33 | 0.000 |

**Table F.1c**: SWPC Flare Rates by Sunspot Region Classification Continued

| | Cycle | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 20 | 21 | | | 22 | | | 23 | | | 24 | | |
| Class | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate |
| FHI | 0.830 | 5 | 6 | 0.833 | 12 | 18 | 0.667 | 2 | 10 | 0.200 | 0 | 2 | 0.000 |
| DKI | 0.550 | 36 | 131 | 0.275 | 59 | 249 | 0.237 | 38 | 175 | 0.217 | 11 | 61 | 0.180 |
| EKI | 1.410 | 111 | 188 | 0.590 | 215 | 472 | 0.456 | 65 | 322 | 0.202 | 12 | 83 | 0.145 |
| FKI | 2.620 | 90 | 81 | 1.111 | 178 | 288 | 0.618 | 73 | 183 | 0.399 | 3 | 10 | 0.300 |
| DSC | 0.350 | 0 | 2 | 0.000 | 0 | 1 | 0.000 | 0 | 5 | 0.000 | 6 | 19 | 0.316 |
| ESC | 0.000 | -- | -- | -- | -- | -- | -- | 1 | 6 | 0.167 | 0 | 14 | 0.000 |
| FSC | 0.000 | -- | -- | -- | -- | -- | -- | 1 | 3 | 0.333 | 2 | 2 | 1.000 |
| DAC | 0.280 | 0 | 1 | 0.000 | 1 | 10 | 0.100 | 21 | 67 | 0.313 | 9 | 100 | 0.090 |
| EAC | 0.530 | 0 | 2 | 0.000 | 0 | 3 | 0.000 | 10 | 73 | 0.137 | 17 | 69 | 0.246 |
| FAC | 0.000 | 0 | 1 | 0.000 | -- | -- | -- | 8 | 31 | 0.258 | 2 | 7 | 0.286 |
| DHC | 0.330 | -- | -- | -- | 0 | 1 | 0.000 | 1 | 3 | 0.333 | 0 | 17 | 0.000 |
| EHC | 2.000 | -- | -- | -- | 0 | 2 | 0.000 | 0 | 4 | 0.000 | 3 | 14 | 0.214 |
| FHC | 0.800 | -- | -- | -- | -- | -- | -- | 10 | 10 | 1.000 | 3 | 4 | 0.750 |
| DKC | 0.820 | 12 | 29 | 0.414 | 40 | 70 | 0.571 | 50 | 98 | 0.510 | 21 | 109 | 0.193 |
| EKC | 2.700 | 64 | 16 | 4.000 | 130 | 110 | 1.182 | 120 | 230 | 0.522 | 71 | 204 | 0.348 |

**Table F.1d**: SWPC Flare Rates by Sunspot Region Classification Continued

| | Cycle | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 20 | 21 | | | 22 | | | 23 | | | 24 | | |
| Class | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate | #Flares | #Regions | Rate |
| FKC | 1.930 | 86 | 40 | 2.150 | 186 | 125 | 1.488 | 194 | 218 | 0.890 | 83 | 99 | 0.838 |

**Table F.1e**: SWPC Flare Rates by Sunspot Region Classification Continued

## Appendix G: Flare Rates by Automated GONG Sunspot Classification

| | Cycle | | | | | |
|---|---|---|---|---|---|---|
| | 23 | | | 24 | | |
| Class | # Regions | # Flares | Rate | # Regions | # Flares | Rate |
| AXX | 3013 | 19 | 0.006 | 3358 | 20 | 0.006 |
| BXO | 847 | 13 | 0.015 | 972 | 7 | 0.007 |
| BXI | 105 | 8 | 0.076 | 122 | 4 | 0.033 |
| HSX | 730 | 7 | 0.010 | 747 | 2 | 0.003 |
| HAX | 833 | 4 | 0.005 | 1129 | 18 | 0.016 |
| HHX | 114 | 3 | 0.026 | 105 | 1 | 0.010 |
| HKX | 122 | 6 | 0.049 | 182 | 5 | 0.027 |
| CSO | 167 | 0 | 0.000 | 172 | 1 | 0.006 |
| CSI | 312 | 7 | 0.022 | 315 | 7 | 0.022 |
| CAO | 206 | 3 | 0.015 | 295 | 7 | 0.024 |
| CAI | 437 | 31 | 0.071 | 541 | 18 | 0.033 |
| CHO | -- | -- | -- | -- | -- | -- |
| CHI | 177 | 10 | 0.056 | 82 | 0 | 0.000 |
| CKO | -- | -- | -- | -- | -- | -- |
| CKI | 194 | 11 | 0.057 | 198 | 10 | 0.051 |
| DSO | 15 | 0 | 0.000 | 18 | 0 | 0.000 |
| ESO | 9 | 0 | 0.000 | 11 | 0 | 0.000 |
| FSO | -- | -- | -- | 4 | 0 | 0.000 |
| DAO | 26 | 1 | 0.038 | 29 | 1 | 0.034 |
| EAO | 14 | 1 | 0.071 | 6 | 0 | 0.000 |
| FAO | 1 | 0 | 0.000 | 2 | 0 | 0.000 |
| DHO | -- | -- | -- | -- | -- | -- |
| EHO | -- | -- | -- | -- | -- | -- |
| FHO | -- | -- | -- | 1 | 0 | 0.000 |
| DKO | -- | -- | -- | -- | -- | -- |
| EKO | -- | -- | -- | -- | -- | -- |
| FKO | -- | -- | -- | -- | -- | -- |
| DSI | 119 | 6 | 0.050 | 107 | 5 | 0.047 |
| ESI | 38 | 1 | 0.026 | 46 | 6 | 0.130 |

**Table G.1a**: Flare Rates by Automated GONG Sunspot Classification

| | Cycle | | | | | |
|---|---|---|---|---|---|---|
| | 23 | | | 24 | | |
| Class | # Regions | # Flares | Rate | # Regions | # Flares | Rate |
| FSI | 4 | 0 | 0.000 | 4 | 0 | 0.000 |
| DAI | 231 | 16 | 0.069 | 276 | 10 | 0.036 |
| EAI | 94 | 12 | 0.128 | 115 | 18 | 0.157 |
| FAI | 5 | 0 | 0.000 | 11 | 7 | 0.636 |
| DHI | 10 | 2 | 0.200 | 8 | 1 | 0.125 |
| EHI | 24 | 2 | 0.083 | 22 | 6 | 0.273 |
| FHI | 10 | 2 | 0.200 | 3 | 0 | 0.000 |
| DKI | 7 | 1 | 0.143 | 21 | 4 | 0.190 |
| EKI | 61 | 15 | 0.246 | 58 | 7 | 0.121 |
| FKI | 31 | 11 | 0.355 | 11 | 1 | 0.091 |
| DSC | 28 | 2 | 0.071 | 36 | 0 | 0.000 |
| ESC | 4 | 0 | 0.000 | 4 | 1 | 0.250 |
| FSC | 0 | 0 | 0.000 | 0 | 0 | 0.000 |
| DAC | 104 | 16 | 0.154 | 101 | 4 | 0.040 |
| EAC | 5 | 0 | 0.000 | 8 | 0 | 0.000 |
| FAC | 0 | 0 | 0.000 | 2 | 0 | 0.000 |
| DHC | 112 | 22 | 0.196 | 68 | 9 | 0.132 |
| EHC | 55 | 12 | 0.218 | 34 | 11 | 0.324 |
| FHC | 26 | 17 | 0.654 | 23 | 18 | 0.783 |
| DKC | 185 | 64 | 0.346 | 178 | 33 | 0.185 |
| EKC | 152 | 75 | 0.493 | 142 | 39 | 0.275 |
| FKC | 71 | 80 | 1.127 | 53 | 37 | 0.698 |

**Table G.1b**: Flare Rates by Automated GONG Sunspot Classification Continued