A DIGITAL ANALYSIS OF SUNSPOT AREAS

G. A. CHAPMAN and G. GROISMAN

San Fernando Observatory, Department of Physics and Astronomy, California State University, Northridge, Calif., U.S.A.

(Received 9 January, 1984)

Abstract. Full-disk white light images of the Sun have been digitized, calibrated, and examined to determine objective sunspot areas for the early part of the operation of the Solar Maximum Mission satellite. We find that published sunspot areas determined from synoptic programs compare favorably with our digital areas. The mean residual between published areas and our digital areas is approximately 80 millionths of a hemisphere. The largest residual found is 642 millionths on April 1980 for Hale No. 16752.

1. Introduction

The discovery that the solar irradiance is affected by sunspots was clearly made by the Active Cavity Radiometer Irradiance Monitor (ACRIM) flown aboard the Solar Maximum Mission (SMM) satellite (Willson et al., 1981; Willson and Hudson, 1981). There have been several attempts to match the ACRIM results using simple models of sunspot deficits based on sunspot areas (Willson et al., 1981; Oster et al., 1982; Eddy et al., 1982). These authors have discussed the inaccuracies in the sunspot data as being a major problem in fitting model sunspot deficits to the ACRIM variations. A similar but more severe problem probably exists for modeling facular brightness excesses from facular areas.

We have completed the major part of the development and testing of a computer program to obtain the area of sunspots from a digitized full disk photograph of the Sun such as is commonly obtained at various observatories as part of their routine or 'synoptic' data collection programs. In this paper we compare these digital sunspot areas with two commonly available compilations of sunspot areas. These sunspot areas are among those used in several studies of the ACRIM data. All of the full-disk photographs analyzed here were from the Sacramento Peak Observatory (SPO) and were sent to us by Mr Lou Gilliam.

2. Description of the Digitization

We have scanned original 35 mm film negatives using an Optronics P-1000 film scanner at a spacing and spot size of 50 µm. The machine outputs an 8-bit binary number, on tape, that is proportional to film density. The entire disk and some surrounding sky are scanned for each day. For the SPO images, one pixel corresponds to approximately 7 arc sec on the Sun.

The software package is called the Full Disk Analysis Program (FDAP) and is configured to run on the CSUN Cyber 750 computer. The operations carried out by this


© Kluwer Academic Publishers • Provided by the NASA Astrophysics Data System
program include determining the center and radius of the image, setting up a calibration table, finding the intensity contrast of each pixel, and locating and counting negative contrast pixels. The criterion for a pixel to be identified as a sunspot pixel is a negative contrast of $\geq 15\%$. The program produces a printer plot of pixels identified as sunspot

Fig. 1. Printer-plot of digitized photograph of the full-disk for 7 April, 1980. The image is distorted in this presentation because of the character spacing of the printer and the geometry of the copy camera.
pixels to aid in the identification of sunspot groups. As a test of the quality of the calibration, printer plots are occasionally made of the input image and the final image of contrast. Figure 1 shows a photograph of a three-part printer plot of the input image. The original is a Sac Peak patrol photograph of 7 April, 1980. Figure 2 is a copy of the contrast map produced by FDAP from the image shown in Figure 1.

The most obvious aspect of Figure 1 is the limb darkening and the two sunspot groups Hale No. 16747 and 16752. These two groups were responsible for the large decrease in solar irradiance seen in the ACRIM signal in April 1980 (Willson et al., 1981). Figure 2 shows the contrast relative to an average of a north–south and east–west (terrestrial) slice of quiet Sun.

Fig. 2. Printer-plot of contrast for the image of Figure 1. The northern part of the solar disk is about 3% brighter than the local average for the quiet Sun.
Clearly evident is a rather smooth, large scale variation in contrast that shows neither north–south or east–west symmetry nor does it show rotational or azimuthal symmetry. We believe the lack of these qualities suggests that the variation is in the original image. We can also see a three-part 'paw print' in the center of each image that suggests the presence of an internal reflection in the telescope system. The maximum variation, from a symmetric quiet Sun, in this particular image is approximately 3%. This variation from circular symmetry prevents any search for facular regions.

Sunspot area is determined by converting each pixel into area as a fraction of a hemisphere, corrected by \( \mu = \cos \theta \) for foreshortening.

3. Results

The data processed is from 27 February, 1980 until 15 May, 1980. This period covers the early operation of the SMM satellite. The two sets of sunspot data are that of the weekly forecast made available by the Space Environmental Services Center (SESC) and that of the World Data Center’s Solar Geophysical Data (SGD). We have carried out a linear regression analysis of these two sets versus our digital sunspot areas, labeled FDAP. Since we believe the noise to be smaller in our digital analysis we have chosen the FDAP areas as the independent variable and either SESC or SGD areas as the dependent variable. The regression analysis follows Bevington (1969).

![Graph](image.png)

Fig. 3. Linear regression of SESC sunspot area versus digitized area from FDAP using Sacramento Peak full-disk photographs. The solid line is the least squares fit and the dashed line is a slope of 1.0.

© Kluwer Academic Publishers • Provided by the NASA Astrophysics Data System
For the SESC data we have $\text{SESC} = A + B \times \text{FDAP}$ where $A = (30 \pm 8) \times 10^{-6}$ and $B = 0.84 \pm 0.02$, where $A$ is in units of a hemisphere.

The linear correlation coefficient is 0.96 and the number of pairs of data points is 171. This regression is shown graphically in Figure 3, where the dashed line represents a slope of unity. The root mean square residual is $83 \times 10^{-6}$ of a hemisphere.

The regression of SGD data versus FDAP data, $\text{SGD} = A + B \times \text{FDAP}$, gives $A = (4 \pm 13) \times 10^{-6}$ and $B = 0.99 \pm 0.03$, where again $A$ is in fractions of a hemisphere. The linear correlation coefficient is 0.97 for 78 data points. The root mean square residual is $81 \times 10^{-6}$. This regression is shown graphically in Figure 4.

It appears that the SESC sunspot areas are more complete than the SGD areas but that they differ significantly from our digital areas. An example of some interest is Hale region numbers 16747 and 16752 of mid April 1980. On April 9, the SGD and FDAP areas differ by 227 and $-42$, respectively, in millionths of the solar hemisphere in the sense of SGD–FDAP. For the same day, the differences, SESC–FDAP, are $-103$ and $-113$, respectively. Differences can be larger as on April 7 when this difference was $-607$ and $-642$, respectively.

These particular regions 16747 and 16752 partly account for the SESC–FDAP regression having a slope less than one. On other days these spot groups have smaller residuals from the regression line.

Fig. 4. Linear regression of SGD sunspot area versus digitized area from FDAP using SPO photographs. Solid and dashed lines are least squares solution and unity slope, respectively.
We have split both data sets into two parts and carried out linear regressions on each part separately. It appears that for both data sets there is a slight non-linearity in the sense that the published estimates overestimate the area of small sunspots. In addition, small sunspots are often unreported but detected by the FDAP.

If we assume that typically there might be about 4 major spot groups on any day and that the errors were approximately gaussian, then the error in the total spot area might be the order of \(80/\sqrt{4} \approx 40\) millionths. Since the bolometric factor is roughly \(\frac{1}{3}\) and projection effects will reduce the influence of some spots, we estimate that a typical daily uncertainty in comparisons with ACRIM data are 10–20 millionths. This uncertainty is in the range of the residuals for daily means of the ACRIM (Willson, 1982). However, because some days may have larger spot residuals, we believe it is important to improve the photographic patrols and digitize spot areas. This statement holds especially for the spin stabilized mode of the SMM satellite.

We are planning to carry out tests of the FDAP with data digitized at twice the resolution of that reported here. We are also digitizing selected San Fernando Observatory (SFO) white light images which are 15\% larger than SPO images. These tests will help determine the accuracy of the FDAP. Comparisons of FDAP spot areas with those from our Reticon photoelectric data (Chapman et al., 1983) will also be made.

Acknowledgements

We wish to acknowledge the efforts of Ms M. E. Brennan of the Aerospace Corporation for her help in the early development and testing of this program. The generous support of the CSUN computer center is appreciated. This work has also received some support under NSF grant No. AST–8121863. The assistance of the Solar Patrol of the San Fernando Observatory is also acknowledged.

References