

A repository of precision flatfields for high resolution MDI continuum data

H.E. Potts¹ · D.A. Diver¹

© Springer ●●●

Abstract We describe an archive of high-precision MDI flat fields, that can be used to refine most MDI high resolution continuum data. The archive consists of many flat field images representing different time ranges over the full operating period of SOHO. The residual flatfield error on the standard level 1.5/1.8 calibrated continuum images represents a significant proportion (25-100%) of the true data variation data on the quiet sun. Using the flat fields in this archive will reduce that error by a factor of 10–30, greatly increasing the accuracy of any tracking or photometric operations. The access, use and accuracy of these flatfields is described in this paper.

1. Introduction

SOHO has been a spectacularly successful solar imaging satellite, producing high resolution images of the Sun in many wavelengths nearly continuously for over 12 years (Scherrer *et al.*, 1995). Space is however a harsh environment, and the satellite has suffered due to radiation and thermal damage over that period, and as a consequence the calibration of instruments has drifted. Many instrumental parameters such as dark currents and focus positions can be successfully recalibrated in space, but recalibration of the flat-field – a measure of the gain of individual pixels on a CCD – is hard. The variation in sensitivity of MDI pixels on lev1.5/1.8 calibrated data is around 2% rms across the full field-of-view of the instrument, and around 0.4% across a 10×10 pixel subimage. In comparison the intensity variation in quiet-sun conditions due to the granulation and *p*-mode signals is only around 2% rms, so it can be seen that the flat field variation has a very significant effect on the images. This can result in large systematic errors in the results from any algorithms that depend on correlations between frames, and cause problems with other automated feature recognition methods, and any photometric measurements.

We have developed a way to retrospectively produce secondary CCD flatfields for use with level 1.5/1.8 high resolution MDI continuum images from existing archived data (Potts and Diver, 2008). We have made a comprehensive set of

¹ University of Glasgow, Glasgow, G12 8QQ, UK email: hugh@astro.gla.ac.uk

these flatfields available for download which cover the full 12 year operating period of the satellite. In this paper we describe how to access and use these flatfields, their accuracy, and also describe other useful functions that can be performed with them. The precision of these flat-fields is much higher than even the original MDI calibration, reducing the residual flatfield variation by around an order of magnitude.

2. Flatfield production

The method for producing the flat-fields is outlined briefly here, and is described in detail in Potts and Diver (2008). The method uses the consistent and well-understood statistical properties of the quiet solar photosphere. If a large number of continuum images from the quiet sun are averaged together, the time variations due to p -modes and granulation rapidly reduce, just leaving a measure of the CCD and optics response of the telescope and the limb darkening of the sun. Long lived deviations from this average brightness such as those created by sunspots, faculae and pores need to be masked out from the data before the frame averaging is done. This is achieved using a combination of the continuum intensity and co-spatial MDI magnetograms. Cosmic ray hits are removed by comparisons across consecutive frames. The result of this process is a flatfield image, with residual noise approximately proportional to \sqrt{n} where n is the number of frames used to make the flatfield image.

3. Data store

The flat field images are hosted on the MDI webserver at Stanford University:

http://soi.stanford.edu/sssc/MDI_continuum_hr_flatfields/flatfields.html (US)

an alternative repository in Europe may be found at:

http://www.astro.gla.ac.uk/users/hugh/MDI_flatfields/flatfields.html (UK)

A preview of the data table contained may be seen in Figure 1. From these pages the full range of flatfields can be browsed, with links to full size images, and previews of the flatfield data and errors. The individual flatfields can be downloaded as FITS files, and also a MATLAB file containing the full set of flatfields, along with other pieces of pertinent metadata, such as focus states, errors distributions, exposure times and basic statistical properties. Also shown in the data table are the step changes in the flat fields caused by the changes in exposure, and other instrumental effects, which are explained in more detail in Section 4.

3.1. Fits keywords

The FITS files containing the flat field images have keywords that describe both the properties of the flatfield itself, and the pointing of MDI during its capture. In Table 1 the new and redefined keywords are described, all other keys have

| Date | Frames | Data rejected | | Flatfield | Flatfield error (%) | Flatfield error | | Other params: | | |
|-----------------|--------|---------------|-------|-----------------|---------------------|-----------------|-------|---------------|-------------------|--|
| | | mean % | max % | | | mean % | max % | exposure (ms) | focus | |
| 2004-06-03 | 12215 | 2.7% | 10.4% | 500 : 0 | | | 0.12% | 0.22% | 990ms focus 4 | |
| EXPOSURE CHANGE | | | | | | | | | | |
| 2004-10-25 | 3239 | 8.8% | 64.2% | 700 : 200 | | | 0.14% | 0.29% | 1080ms focus 5 | |
| 2004-10-30 | 4574 | 4.5% | 30.0% | 505 : 5 | | | 0.12% | 0.25% | 1080ms focus 5 | |
| 2005-05-01 | 2164 | 15.6% | 62.9% | 500 : 0 | | | 0.19% | 0.68% | 1080ms focus 5 | |

Figure 1. Small section of the data table from the flat-field repository webpage.

their standard SOI meaning¹ and are taken directly from the CENTFITS file with is the fits file observed at the median time of all the flatfield frames, T_OBS. The keywords that describe the pointing of MDI allow Solarsoft routines such as fits2map to work with the flatfields.

4. Time variation of the flat-fields

Over the 12 years of operation of SOHO the flat fields have changed continuously. There are three main components to the time variation: a gradual global loss of sensitivity, mainly due to the darkening of the front window of MDI², variations in the performance of the interferometer, giving large scale effects, and individual pixel degradation. The darkening of the front window has resulted in a decrease in sensitivity of the instrument of nearly a factor of two over the life of the instrument, but this effect has been somewhat mitigated by increasing the exposure times several times³. The effect of this is a non-continuous decrease in the average intensity of MDI images, which is shown by the upper blue curve in Figure 2, which shows the mean value of rows 381–500 of the CCD from the flat-field images over the full time period, along with the exposure times used. If

¹SOI keywords: <http://soi.stanford.edu/cgi-bin/keyword.pl>

²MDI operations page: MDI Front Window Transmission History: http://mdisas.nascom.nasa.gov/transmission/index_hr.html

³MDI calibration page: events list: <http://mdisas.nascom.nasa.gov/events/events.html> MDI Exposure Changes

Table 1. Description of the new fits keywords

| Keyname | Description |
|----------|--|
| T_OBS | A representative time for the flatfield, taken from the median time from all the frames. This frame is used for all the pointing information contained in the standard MDI keywords. |
| T_FIRST | Time of the first fits file used in the observation |
| T_LAST | Time of the last fits file used in the observation |
| FRSTFITS | Name of the first fits file used in the flatfield |
| CENTFITS | Name of the median time fits file with observation time T_OBS |
| LASTFITS | Name of the last fits file used in the flatfield |
| REJ_MEAN | Mean proportion of data removed due to sunspots and magnetically active regions |
| REJ_MAX | Maximum amount of data rejected due to sunspots and magnetically active regions |
| ERR_MEAN | rms error of flatfield in dn |
| ERR_MAX | Maximum error of flatfield in dn |
| EXPOSURE | MDI exposure time in ms for the frames |
| FOCUS | Position of focus wheel |

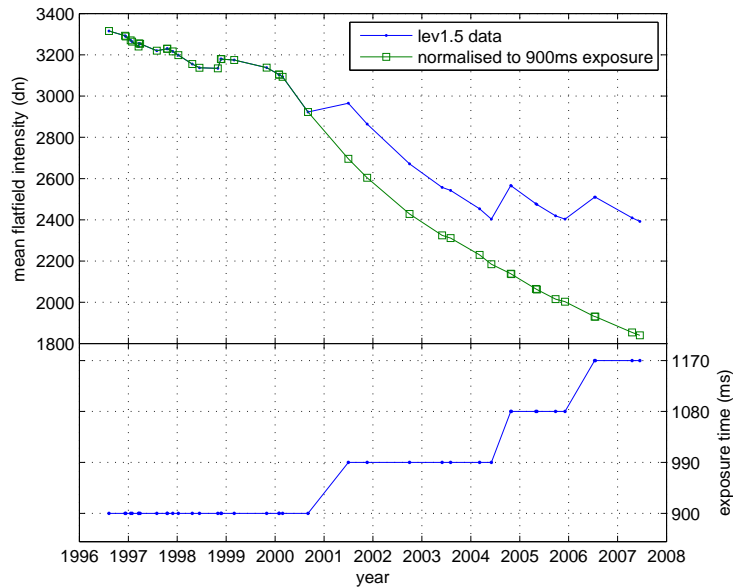


Figure 2. Reduction in sensitivity of MDI over 12 years of operation, measured from the flatfield intensity, and the change in exposure time over the same period.

the data are all normalised to the initial 900 ms exposure a continuous reduction in sensitivity can be observed from 2000–2008, also shown in the figure.

The large scale variation in the sensitivity of the images across the frame is dominated by the performance of the Michelson interferometer (Scherrer *et al.*, 1995). This has changed gradually over the full time period, with one exception: between June and October 1998, when MDI was lost in space, where there was a step change, probably due to thermal stress. This change is clearly visible in the flatfield previews on the web pages. At the smallest scales, the sensitivity of the individual pixels drifts with time. This process is gradual, but there a few times where more rapid changes occur.

All the major discontinuities in exposure and large and small scale variations are indicated by coloured bars across the data table.

5. Using the flat-field images

5.1. Choosing an appropriate flat field

If a flat field is available that is made containing the data set you are using, this should always be used, and will give the best accuracy. If this is not the case it is best to choose the closest flat field in time that covers the region of the CCD you are interested in, although there are a few caveats:

- The flat field does not change continuously: there are occasional step changes at small and large scales. These are indicated by horizontal bars in the data table. If possible choose a flat field where the observation time is within the same block as your data.
- If the data of the observations is between two flat fields of similar residual noise within the same data block, then an interpolation between the two, weighted by the number of contributing frames will generally give better results. The improvement over the single flat field can be checked using the method described below.

If it is not clear which flatfield is the best to use, then the best flat field to use is the one one that has the best correlation with the data that you are using, at the scales you are interested in. The best procedure to test for this is as follows:

1. Make a time average of a few frames of the data to be corrected. A minimum of 10 well-spaced frames is recommended, but the more the better. If the data contains an active region, mask out sunspots and pores using the method described in section 7.2
2. If any spatial filtering of the data is required, for example for the removal of p -modes, apply the filter to both the averaged frame, and the possible flat field candidates. This is to ensure that any errors are not dominated by changes at a different spatial scale.
3. Calculate the correlation between the filtered average image and each of the flat field candidates. Choose the flat-field with the best correlation. If there are two similar results try a weighted average, and see if this gives a better result.

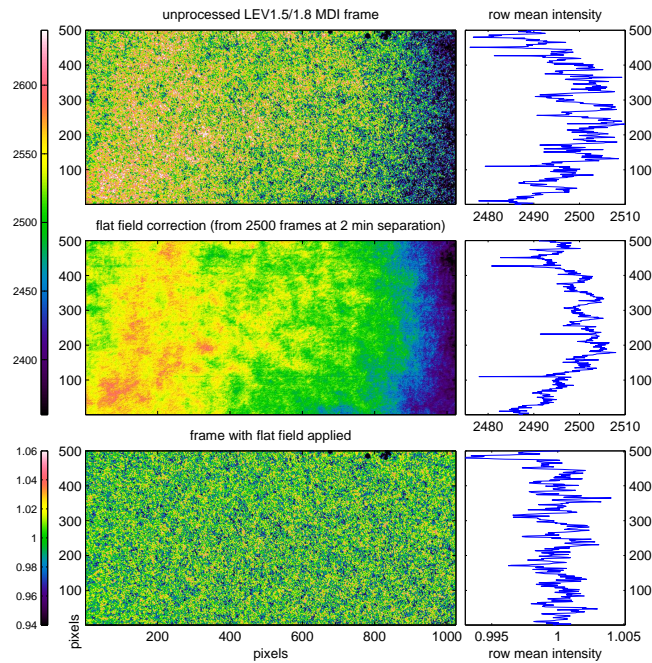


Figure 3. An example of using a flat field, showing a typical frame (top row), the flat field(middle), and the corrected image (bottom)

If there is a requirement for the best possible accuracy, and a large dataset is available (> 2000 frames), contact the authors of this paper, and we will endeavour to add an appropriate flat field to the archive.

5.2. Applying the flat-field to data

These flat fields are designed to be used with level 1.5/1.8 data. To use with level 0 data the flat-fields must be first combined with the relevant primary flat field. The simplest use of the flat fields is to divide the continuum fits data by the appropriate flat field. This will give a new image plane centred approximately on unity, with all the flatfield errors and limb darkening removed. Notice that the variation in instrument sensitivity across the imaging array is much larger than that due to the limb darkening, so this cannot be observed directly with MDI in its high resolution mode. An example of usage in IDL follows:

```
flatfield = readfits(<flatfield filename>)
data = readfits(<MDI continuum filename>)
data_corrected = data/flatfield
data_corrected_valid = data_corrected(3:1020,*)
```

Note the last line of the code - the first and last three columns of the MDI CCD are not valid image pixels, when normalised with the flatfield they will be set to approximately unity. An example showing the appearance of a continuum image

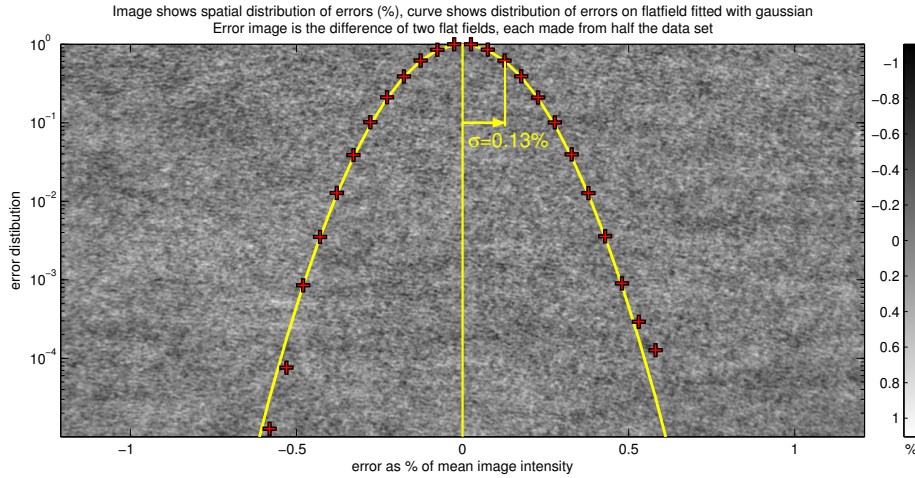


Figure 4. A typical error plot from a fairly quiet period (December 2002), made from 2990 frames. The background image shows the difference between the two flatfields, and the red crosses are a histogram of the distribution of these errors. The yellow curve is a gaussian fit to the error distribution. Note that the y axis for the histogram and fit is logarithmic.

of quiet sun before and after normalisation is shown in Figure 3. This process replaces the value of persistently stuck pixels with the mean intensity of the image. See Section 7.1 for more details.

6. Flatfield accuracy and errors

The errors on an individual flatfield depend on the number of frames that were used to construct it, and the level of activity of the sun over the period. In the web repository these errors are represented in a variety of ways, including the spatial and magnitude distributions of the errors, and their rms and maximum values. The error values were obtained by constructing two independent flat fields from each data set. The pairs of flat fields were made such that the number of contributing frames for a pixel on one frame (and hence its error) was the same as that for the corresponding pixel on the other flat field. The average pixel error, and spatial error distribution could then be obtained from the difference between these two flat fields. The final flat field is made from the mean of these two independent flat fields. Assuming that all the pixels have statistically similar noise, the rms pixel error σ_f on the combined flatfield is just:

$$\sigma_f = \frac{\sigma(f_1 - f_2)}{\sqrt{2}} \quad (1)$$

where sigma denotes the standard deviation function and $(f_1 - f_2)$ is the set of pixel-by-pixel differences between the flat fields. The distribution of these errors for quiet Sun cases is almost exactly gaussian. An example of an error plot from a fairly quiet sun period may be found in Figure 4. The background of this graph

is an image showing the difference between the two flatfields, expressed as a percentage of the mean flatfield. These graphs may be obtained for each flatfield by clicking on the thumbnail in the *Flatfield error* column. The thumbnail shows how even the spatial distribution of the errors is, and is obtained by taking the absolute value of the difference image and smoothing it. The amplitude of the error is roughly proportional to the square root of the number of frames that produce each pixel of the flat field. A consequence of this is that in highly active regions, where a large proportion of the data is rejected, the flat-field error will be much higher. This is particularly a problem during solar maximum, where extended active regions can reduce the frame count locally to less than half that of the source frames, resulting in more than $\sqrt{2}$ times the error in the surrounding quiet sun regions. This effect can clearly be seen on some of the the error images on the web pages, particularly for data sets between 2000 and 2005, around solar maximum. For cases where a good flatfield is required during a very active period the recently published method of Wachter and Schou (2009) may be better. This method has the advantage that it can use all the data, including that from highly active regions to generate a good quality flatfield for small scale features.

Note for comparison that the error on the primary flat field (used for level 1.5/1.8 data), is given by the rms value of the secondary flat field, and is typically around 2% of the mean image intensity, compared to a value of around 0.1 – 0.2% for the secondary flat fields. The magnitude of the errors is described in detail in Potts and Diver (2008).

7. Other data processing operations

7.1. Identifying bad pixels

In the flat field image the intensity of the bad pixels is preserved. Correcting individual frames by dividing by the flat field will therefore set these pixels to unity, equal to the mean of the corrected image. If bad pixels need to be identified it is easy to do on the flat-field image, as this already has the small scale fluctuations due to granulation and *p*-modes removed. A good method is to pass the flat field through a high pass filter in order to remove the large scale variations, and then apply a threshold on the absolute value of the resultant image. A suitable way to filter the data is to subtract from the flatfield a copy of it that has been blurred by convolution with a 2d gaussian of width 5 pixels, and sum unity. Note that cosmic ray hits still need to be removed on a frame-by-frame basis.

7.2. Automated recognition of sunspots and pores

The use of these flat fields greatly simplifies the identification of sunspots and pores, as it effectively removes all the large scale variation on the data. An intensity threshold, ideally on the mean of several frames can then be used to simply identify any sunspots or pores. An example is shown in figure 5, where a

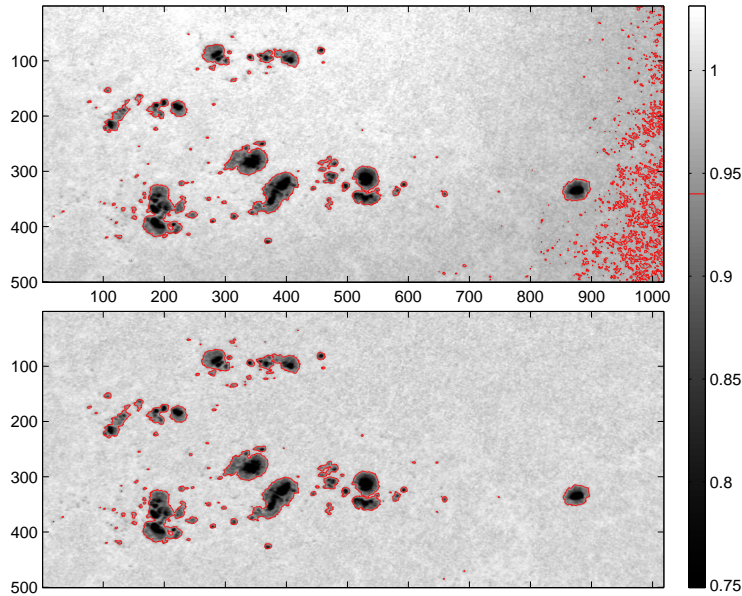


Figure 5. Demonstration of finding sunspots in a complex active region using a simple intensity threshold on uncorrected data (*top*) and flatfield corrected data (*bottom*). The image data is the mean of 20 high resolution continuum frames, taken on 25th June 1999

threshold of 0.95 on the flat-field normalised data was chosen (*bottom*), compared with an equivalent threshold on the original lev1.5 data (*top*). It can be seen that this simple method accurately picks up all the sunspots and pores on the flatfield corrected image, unlike on the original image where the large scale intensity variation causes small pores to be missed in bright areas, and false positives in dark areas. This technique is very useful in identifying areas to be masked from any granulation tracking procedures.

8. Summary

Use of the flat-fields in the repository described in this paper substantially increases the quality of all MDI high-resolution continuum data, effectively increasing the signal to noise ratio on the corrected data by a factor of 10–20 over almost all observation periods. These flatfields should be used whenever any photometric or motion tracking algorithms are being performed on high resolution continuum data. Accurate flat fielding of the data also allows simple identification of sunspots and pores and enables a detailed analysis of the degradation of the instrument over its years of service.

Acknowledgements This work was supported by UK STFC rolling grant number ST/F002149/1. Financial support by the European Commission through the SOLAIRE Network (MTRN-CT-2006-035484) is also gratefully acknowledged.

References

- Potts, H.E., Diver, D.A.: 2008, Post-hoc derivation of SOHO Michelson Doppler Imager flatfields. *Solar Phys.*
- Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., Wolfson, C.J., Zayer, I., MDI Engineering Team: 1995, The Solar Oscillations Investigation - Michelson Doppler Imager. *Solar Phys.* **162**, 188–188.
- Wachter, R., Schou, J.: 2009, Inferring Small-Scale Flatfields from Solar Rotation. *Solar Phys.*