COMPARISON OF THREE SUNSPOT AREA DATABASES

T. Baranyi[1], A. Ludmány[1], L. Győri[2], H.E. Coffey[3]

[1]Heliophysical Observatory of the Hungarian Academy of Sciences, Debrecen, P.O.Box 30, H-4010, Hungary
Phone/Fax: 36 52 311015, e-mail: baranyi@tigris.klte.hu, ludmany@tigris.klte.hu
[2]Heliophysical Observatory, Gyula Observing Station, 5701 Gyula, P.O.Box 93, Hungary
Phone: 36 66 361553, e-mail: ls@mail.gyula.hungary.net
[3]NOAA/NESDIS/NGDC E/GC2 Solar-Terrestrial Physics Division, 325 Broadway, Boulder, CO 80303
Phone: 1 303 4976223, Fax: 1 303 4976513, e-mail: hcoffey@ngdc.noaa.gov

ABSTRACT

Sunspot area measurements play an important role in the studies of flux emergence in spots, development and decay of sunspot groups and variations in solar irradiance. These studies require a long homogeneous time-series of sunspot areas but after ceasing of Greenwich Photoheliographic Results there is no widely accepted standard for these data. That is why there is a need to compare the area databases provided by different observatories. In this study we statistically investigated the differences between the area data by Debrecen and Yunnan Observatories as well as those published in Solnechnie Dannie for the year 1986 and 1987.

Key words: sunspot area; flux; irradiance.

1. INTRODUCTION

Sunspot areas are important parameters for lots of studies in various fields of solar physics. Some examples are listed here which are related with the most important fields: growth and decay of spots (Lustig & Wöhl 1995), emergence of fluxes in spots (Baranyi & Ludmány 1992), evolution of sunspot groups and interaction between them (van Driel-Gesztelyi et al. 1993), axial tilt and rotation rate of sunspot groups (Howard 1991, 1992), periodicities in solar activity (Oliver & Ballester 1995), fragmentation of flux tubes (Ludmány et al. 1999) and irradiance variability (Fröhlich et al. 1991).

Because of its importance the area of sunspot is measured in several observatories. However, if someone wants to determine the sunspot area, he or she has to cope with many difficulties which result in random and systematic errors (Győri 1998). Furthermore, after ceasing of Greenwich Photoheliographic Results (GPR) there is no widely accepted standard for these data. That is why there is a need to compare the area databases provided by different observatories, so that we could make a relative calibration of them at first and an absolute calibration at the end. This work has been started and several comparative studies are published (Sivaraman et al. 1993; Lustig & Wöhl 1995; Fligge & Solanki 1997; Hathaway 1999).

The Debrecen Heliophysical Observatory has an extended history of detailed observation of sunspot areas and positions. The team of Debrecen Photoheliographic Results (DPR) (Dezső et al. 1987) makes a very great effort to provide an extraordinary detailed compilation of reduced measurements of the sunspot active regions starting with 1977. More recently, a related project was begun to reduce the data more quickly, but restricted to the basic data. The goal of this team is only to produce a research-quality data set of the sunspot areas and positions. The starting year of Debrecen Photoheliographic Data (DPD) is 1986 and it is also available for the year 1987. On the basis of these recently published catalogue we can compare the data measured in Debrecen with the ones of other observatories.

2. OBSERVATIONAL DATA

The DPD data are published in several kinds of form (printed, CD-ROM journal and electronic form) (Győri et al. 1996, 1998) and it contains the data for the whole group and each spot in it. The Debrecen data are measured on white-light full-disk photographic plates as it happened in Greenwich. A similar data base is also published in the Solnechnie Dannie (SD 1986-1987) based on measurements of white-light photographic plates but they are measured with a different method. This catalogue contains the areas of the whole group as well as that of the largest spot of the group. The third data base is provided by the Yunnan Observatory (CSGD 1986-1987; Coffey & Hanchett 1998; www.ngdc.noaa.gov/stp) and the areas are measured by using drawings here. The area of the largest spot is also available in Yunnan. These three data bases were statistically investigated in this study and we could compare the DPD with another photographic catalogue and a data base measured with a more widely used method.

3. COMPARISON OF DATA BASES

By using the position data we identified the sunspot groups. There were 715 cases when the same group was measured in all three observatories on the same day and the relative distance (r) was smaller than 0.98. The independent variable was the area (D) published in the DPD and the dependent variables were the areas measured in Yunnan (Y) and Solnechnie Dannie (S). We made a linear regression analysis and a curve estimation in the form: $dependent = a + b * independent$.

Proc. 9th European Meeting on Solar Physics, 'Magnetic Fields and Solar Processes', Florence, Italy, 12-18 September 1999 (ESA SP-448, December 1999)

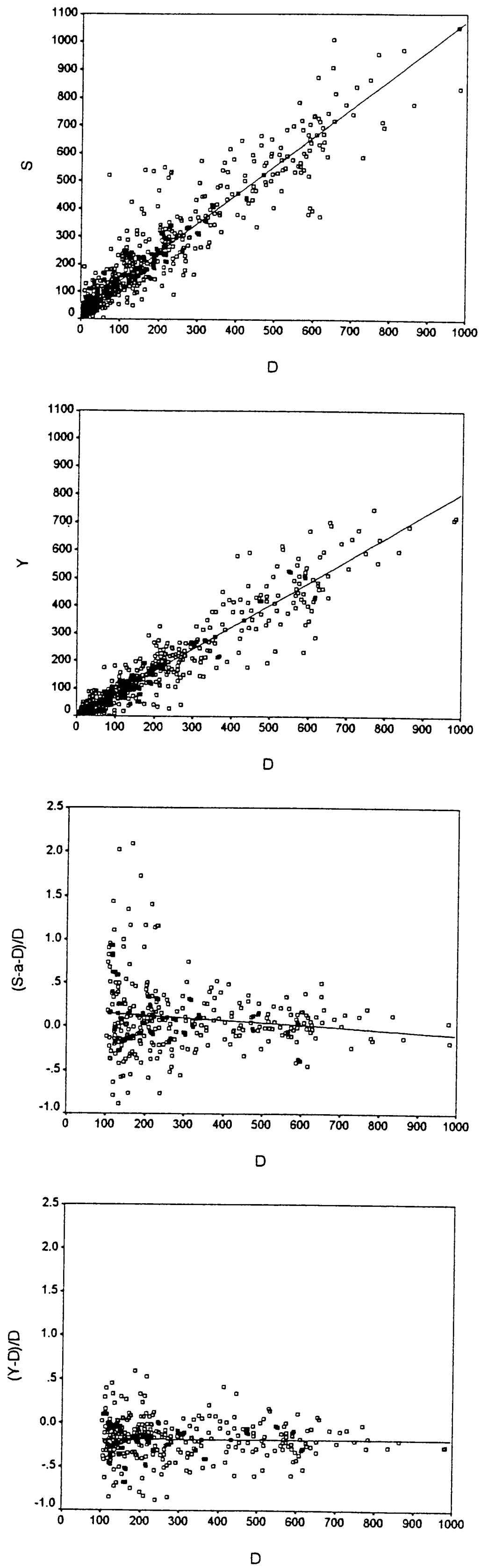

*Figure 1. a.) Area of sunspot groups published in SD versus area of sunspot groups measured in Debrecen (S=34.09+1.04*D). b.) Area of sunspot groups measured in Yunnan versus area of sunspot groups measured in Debrecen (Y=0.38+0.81*D).*
*c.) Relative error of area of sunspot groups published in SD versus areas of sunspot groups measured in Debrecen ((S-a-D)/D=0.1741-0.0003*D).*
*d.) Relative error of area of sunspot groups measured in Yunnan versus area of sunspot groups measured in Debrecen ((Y-D)/D=-0.1876-0.000008*D).*

3.1. Area of the whole group

For the S data the a was 34.09 and the b was 1.039 while for the Y the a was 0.38 and $b = 0.806$ (Fig.1.a and b). One can see that the S is larger with a relatively big constant but neglecting this fact the average areas are equal to the DPD data. Concerning the Yunnan data there is no need to correct them with a constant but the Yunnan areas are smaller with $19 - 20\%$ than that of DPD. It is also remarkable in these figures that there are large deviations from the linear curve mainly for S. The standard error for S is 67.27 and for Y is 46.41. This fact is worth some further examination. The precision of the area measurement depends on the area as the possible error is smaller in case of a small spot than a large spot. In case of a large group the standard error means only a small percentage of the whole area. For small groups the area variations caused by the evolution (times of observations differ by a half day) or by measurement problems may be smaller than the standard error but the relative differences may be large comparing with their small areas.

This means that sometimes the absolute error (standard error received in this linear regression) is not as characteristic as the relative error is. By taking into account the a constant difference we define the relative error by the following expression: $(dependent - a - independent)/independent$.

If we want to estimate the accuracy of the measurements, we have to study this parameter in cases of larger groups. Let us choose the $D > 100$ value, which is bigger than the standard errors. Fig.1.c and d shows the relative error in case of $D > 100$ for S and Y respectively. For the S data the a was 0.1741 and the b was -0.0003.0 These data mean that the S data are larger than D with an average difference of 14% at $D = 100$ and the difference decreases with the increase of D. The standard error is 0.39, which means that the standard error in the scatter of data is 39%. However the relative error is not constant. It may be $100 - 150\%$ at about $D = 200$ and it is decreasing with increasing D.

Concerning the relative errors for Y data the $a = -0.1876$, b is practically zero and the standard error is 0.22. These mean that in this case the average accuracy is much better. There is a stable systematic difference between D and Y, which is equal with the previously determined value of 19%. The scatter is much smaller and it is slightly decreasing with increasing D. After taking into account the systematic difference one can say that the D and Y data are equal within about $\pm$ 22% random error.

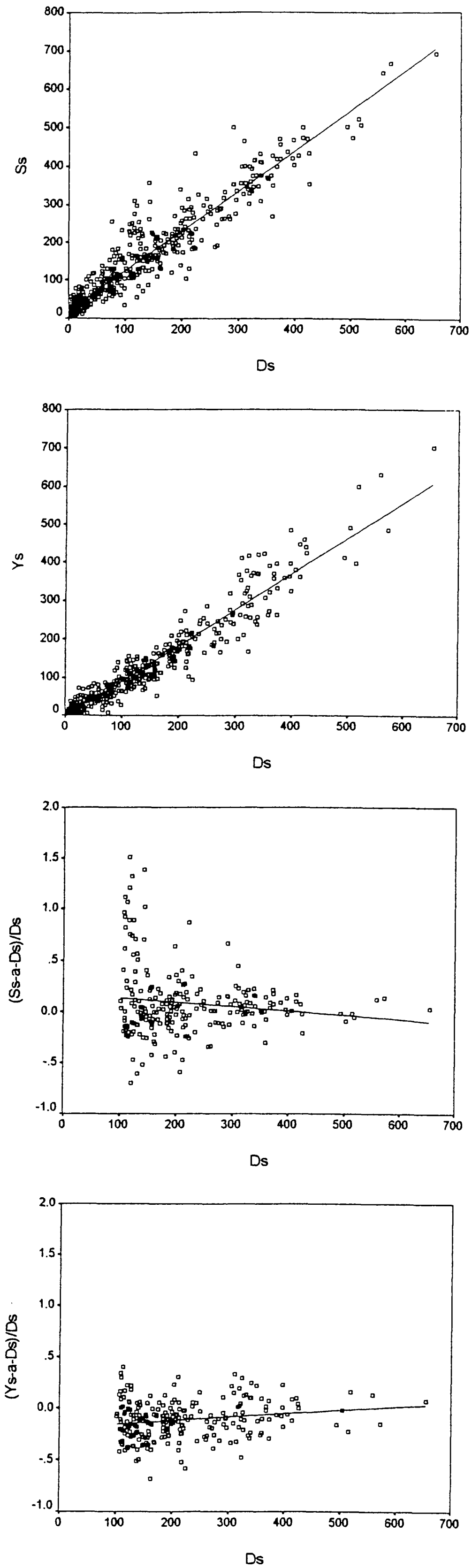

*Figure 2. a.) Area of the largest spots published in SD versus area of the largest spots measured in Debrecen (Ss=18.88+1.05*Ds). b.) Area of the largest spots measured in Yunnan versus area of the largest spots measured in Debrecen (Ys=-4.58+0.94*Ds).*
*c.) Relative error of area of largest sunspots published in SD versus area of largest spots measured in Debrecen ((Ss-a-Ds)/Ds=0.1765-0.0004*Ds).*
*d.) Relative error of area of largest spots measured in Yunnan versus area of largest spots measured in Debrecen ((Ys-a-Ds)/Ds=-0.1500+0.0003*Ds).*

We studied the effect of the fact that close to the limb the measurement may be much difficult. It is much hard to find the border of spots or/and the correction for foreshortening may enlarge the small errors. Two subsets of data bases were separated by the relative distances of groups from the disc center: $r >= 0.7$ and $r < 0.7$.

Comparing the "center" and "limb" subsets the linear regression gives similar results for Y and D, while for S and D the limb regions will substantially differ from each other, the a constant deviation, the value of b as well as the standard error are larger if $r >= 0.7$. This means that close to the limb the errors in S are increasing because of the method of measurements.

3.2. Area of the largest spot

We exploited the fact that all three data bases contain the area of the largest spot of each group. We chose those cases where it seemed to be unambiguous that the same spot was measured at all three observatories (519 cases). By using this subset we can estimate the accuracy of the measurements of single spots.

Figure 2.a. shows the the result of linear regression between the areas of spots of DPD (Ds) and that of Solnechnie Dannie (Ss). For Ss $a = 18.88$, $b = 1.051$ and the standard error is 40.18. Comparing these parameters with that of recieved in the previous subsection it can be seen that the a and the standard error of the estimate decreased to a large extent but the b has practically the same value. This means that areas of individual spots are less burdened with errors but the Ss data are systematically larger with a constant about 19 and a percentage of 5%.

Figure 2.c. shows the result for the relative error in case of $Ds > 100$. The parameters of the linear estimation are $a = 0.1765$, $b = -0.0004$ and the std. error is 33%. This means that the 5% difference between Ss and Ds is an average value over the interval of Ds. The relative error depends on the Ds it is about 14% at $Ds = 100$ and it is decreasing with the encrease of Ds. These are in good agreement with the results received in the study of group and they can characterize the relative differences between the DPD and SD data bases.

We can make a further investigation for the systematic differences. When there is no observation in the Debrecen Observatory then the main source of the plates for the DPD is the Kislovodsk Observing Station which is also the main source of data for SD. If we separated the data of Kislovodsk in DPD (117 data) we can conclude that in lots of cases the same plates were measured for DPD (in Debrecen) and SD (in Kislovodsk). Considering only the data of Kislovodsk plates from DPD and the data for the same day from SD we received the following values comparing the Ds and Ss data $a = 19.05$, $b = 1.002$, std. error $= 63.71$. The value of a is almost the same as that for all Ds and Ss data and we can conclude that the constant difference between Ds and Ss is systematic and it is likely resulted by the measuring method. Taking into account this constant one can say that there is no systematic difference (0.2%) between Ds and Ss measured on the same plates although the random deviations are large.

Figure 2.b. shows the result of linear estimation between Ds and Yunnan's areas of spots (Ys). For Ys $a = -4.58$, $b = 0.937$ and the standard error is 30.31. As it was expected the std. error is smaller than it was in case of the area of the whole group so the random errors are smaller. However there is a surprising result here: there is only 7% difference between the Ys and Ds areas, which is much less than the 19% found in case of the whole group.

For detailed study of this difference we displayed the relative error again (Fig.2.d.). The results for the linear regression are: $a = -0.1500$, $b = 0.0003$, std. error $= 0.178$. This means that Ys is smaller than Ds with 12% at $Ds = 100$ but their difference is decreasing with the increase of Ds. In contrast with the constant difference between D and Y the systematic difference between Ds and Ys depends on Ds and it is smaller if Ds larger.

4. DISCUSSION

It is well-known that the differences between area measurements can occur because of several reasons. Repeated measurements of a given observer show a few percentage random differences. Comparing two or more observers we can find smaller systematic and larger random errors (Gerlei 1977). Area measurements of the same spot at the same time from different observatories depend on, among other things, the differences in seeing (Fligge & Solanki 1997 and references therein). However on longer timescale when data are gathered from several observers in case of different seeing conditions there are only two things which mainly influence the systematic differences: the observing and the measuring techniques.

At present there is no overlap between GPR and DPD so we cannot compare them directly. However there is a comparison between GPR and DPR (Gerlei 1987). As the method of measurement is the same for DPR and DPD it is assumed that results of this comparison are valid for DPD, too. In this case we can say that the Greenwich areas are larger with about 8% than that of DPD and this fact arises from the different method of measurement.

The SD data are also larger than DPD areas. There is a constant systematic deviation between them (34 for the whole group and 19 for the spots) and the cause of it is in the method of measurement of SD. Besides this constant systematic deviation there is also a size-dependent systematic deviation between them $(4 - 5\%)$. Concerning only the areas of the largest spots of Kislovodsk plates we concluded that there is no further systematic difference between the SD and DPD data, both of the methods provide the same average result (with random errors). As the result is the same for the same plates we have to seek for the cause of the difference in the fact that in most cases the DPD uses different kind of plates as the SD does. The only important difference between the Debrecen and Kislovodsk plates is the gamma (slope of the linear portion of density-vs-intensity (H&D) curve) of the photographic material.

The variation of the intensity across a spot to the umbra is the following (see Fig.1. in the Beck and Chapman 1993): approaching the spot there is a thin zone where the intensity decreases from the level of photosphere to the level of penumbra and coming through the penumbra there is a similar zone where the intensity decreases from the level of penumbra to that of umbra. On the photographic plates the points of maximum slopes in these zones are usually defined as the boundaries between photosphere and penumbra as well as penumbra and umbra. When we take a photo we make a non-linear intensity transformation by the H&D curve. This non-linear transformation result in that the points of maximum slope can be closer to the penumbra and umbra than they are in fact. The higher the gamma of the film is, the closer these points are. (By all means, we cannot say that the lower gamma is better than the higher one as in case of that it is likely that the large umbrae are underexposed in a large extent.)

The gamma is higher for the Debrecen plates than for SD plates so this fact can explain that the DPD areas are systematically smaller with $4 - 5\%$. This can explain the linear decrease of the relative error, too (Fig.2.c.). The thickness of the zone between the photosphere and penumbra does not change linearly with the extension of the spot. So the larger the spot is, the smaller the difference between the DPD and SD areas caused by the gamma of film is.

By accounting these systematic differences caused by the method of measurement or the used photographic plates we can conclude that the GPR and DPD and SD areas are in good agreement although they can be burdened with random errors.

Concerning the Yunnan data the most surprising thing was the fact that the Yunnan areas of the whole spots were smaller than that of DPD with uniformly 19% but the deviation for the single spots was substantially smaller. For the largest spots of Yunnan the relative deviation depends on the area of DPD in a similar way as we found at the SD data but in opposite sense. This means that at $Ds = 100$ the Ss data are larger with 14% while Ys smaller with 12% but their difference is decreasing if Ds is increasing. The relative error between the three observations is practically zero at about $Ds = 500$. On the basis of this similarity we can conclude that the root of the deviation of Ys is the same as that of Ss. This means that making drawings can be modeled with usage of the photographic plate of the highest gamma. In this case the position of the point of the maximum slope is practically at the edge of the penumbra. This result and the fact that the Yunnan areas of the whole spots are smaller uniformly with about 20% is seems to be explainable with characteristics of sensitivity of eyes.

In case of measuring of photographic plates as well as making drawings the main measuring "facilities" are the human eye and brain. We can suppose that the cause of the systematic difference between these two groups of data is likely to be hidden in the different conditions for the determination of the borders of spots. It is well known that the viewer's eye (and brain) estimates the brightness on a given place of the retina on the base of its deviation from the mean of the intensity of the surround of that place. This means that we make relative measurements of brightness. For example this fact results in the well-known effect that we can see the same grey square to be brighter if its background is dark than it is white (simultaneous contrast, see e.g. Kaiser 1999). In case of making a drawing there is a bright white background in which we have to determine the dark border of spots. However the negative film gives a gray background in which we have to find the border of bright spots. Comparing the usage of films with making drawings we can find other differences which can influence our perception: the interval between the intensity of brightest and the darkest points is less as it suffered a logarithmic transformation by the H&D curve; we can change the illumination to the optimal level by changing the voltage of the lamp; the average brightness is smaller and the rods can play larger role in the perception than in case of drawing.

These differences may contribute to the result that the spots seem to be smaller when we make a drawing compared to the case when we look at a photo. Concerning the area of the whole group these are likely to cause that making a drawing we may neglect a fraction of the group. In case of a large group there is a large dark feature on the bright disk. The larger the group is, the larger those areas can be where the intensity is closer to that of the photosphere than that of the large dark penumbras and umbras. Comparing the brightnesses we are likely to decide that those bright peripheral parts (over a brightness limit) do not belong to the group. So the larger the group is, the larger the omitted areas can be. A similar effect can be experienced if we look at positive prints of different hardness, brighter details may be missed on the harder photographic paper. From our results the rate of the omitted part may be about 20%. This assumption is in agreement with result of Fligge% Solanki (1997) who agree that there is no significant difference between the area data of those observatories which measure by using drawings (Rome, Yunnan, Catania, US Air Force) but they are systematically smaller than the Greenwich data with about $15 - 25\%$.

Summarizing the above findings we can conclude that there are two data groups of area measurements of sunspot groups. The first data group contains the areas measured on photographic plates and the second one contains the areas measured on drawings. Between these two data groups there is a systematic difference about 20%. The exact value will depend on time, which needs further studies.

ACKNOWLEDGMENTS

The work was supported by the grants of Hungarian National Foundation for Scientific Research Nos. OTKA-T025640 and F019829, as well as by U.S. -Hungarian Joint Found for Science and Technology under contract No. 95a-524.

REFERENCES

Baranyi, T., Ludmány, A. 1992, Sol. Phys. 139, 247

Coffey, H.E., Hanchett, C.D. 1998, ASP Conf. Ser. 140, 447

CSGD 1986-1987, Chinese Solar-Geophysical Data Montly Reports, Academia Sinica, Beijing

Dezső, L., Gerlei, O., Kovács, Á. 1987, Publ. Debrecen Obs. Heliogr. Ser. 1, 1

Fligge, M., Solanki, S.K. 1997, Sol. Phys. 173, 427

Fröhlich, C., Foukal, P.V., Hickey, J.R., Wilson, R.C. 1991, in: Sun in Time, eds.: C.P. Sonett et al., The University of Arizona Press, Tucson, 11

Gerlei, O. 1987, Publ. Debrecen Obs. Heliogr. Ser. 1, Appendix 2

Győri, L. 1998, Sol. Phys. 180, 109

Győri, L., Baranyi, T., Csepura, G., Gerlei, O. Ludmány, A. 1996, Publ. Debrecen Obs. Heliogr. Ser. 10, 1 Journ. of Astron. Data 4, 2 ftp://fenyi.sci.klte.hu/pub/DPD

Győri, L., Baranyi, T., Csepura, G., Gerlei, O. Ludmány, A. 1998, Publ. Debrecen Obs. Heliogr. Ser. 11, 1 Journ. of Astron. Data 6, 1 ftp://fenyi.sci.klte.hu/pub/DPD

Hathaway, D.H. 1999, www.ssl.msfc.nasa.gov /ssl/pad/solar/greenwch.htm

Howard, R.F. 1991, Sol. Phys. 136, 251

Howard, R.F. 1992, Sol. Phys. 142, 47

Kaiser, P.K. 1999, www.yorku.ca/eye/

Ludmány, A., Baranyi, T., Mező, G. 1999, JOSO Ann. Rep. 1998, 99

Lustig, G., Wöhl, H. 1995, Sol. Phys. 157, 389

Oliver, R., Ballester, L.J. 1995, Sol. Phys. 156, 145

Sivaraman, K.R., Gupta, S.S., Howard, R.F. 1993, Sol. Phys. 146, 27

Solnechnie Dannie Bulletin, 1986-1987, Montly Reports, Main Astrophys. Obs. Acad. USSR, Nauka, Leningrad

van Driel-Gesztelyi, L., Csepura, G., Nagy, I., Gerlei, O., Schmieder, B., Rayrole, J., Demoulin, P. 1993, Sol. Phys., 145, 77