

bradscholars

A multi-wavelength analysis of active regions and sunspots by comparison of automatic detection algorithms

Item Type	Article
Authors	Verbeeck, C.;Higgins, P.A.;Colak, Tufan;Watson, F.T.;Delouille, V.;Mampaey, B.;Qahwaji, Rami
Citation	Verbeeck C, Higgins PA, Colak T et al (2013) A multi-wavelength analysis of active regions and sunspots by comparison of automatic detection algorithms. Solar Physics. 283(1): 67-95.
DOI	https://doi.org/10.1007/s11207-011-9859-6
Rights	© 2013 Springer Netherlands. Full-text reproduced in accordance with the publisher's self-archiving policy.
Download date	2026-05-18 02:49:42
Link to Item	https://bradscholars.brad.ac.uk/handle/10454/7682.2



The University of Bradford Institutional Repository

<http://bradscholars.brad.ac.uk>

This work is made available online in accordance with publisher policies. Please refer to the repository record for this item and our Policy Document available from the repository home page for further information.

To see the final version of this work please visit the publisher's website. Access to the published online version may require a subscription.

Link to original published version: <http://dx.doi.org/10.1007/s11207-011-9859-6>

Citation: Verbeeck C, Higgins PA, Colak T, Watson FT, Delouille V, Mampaey B and Qahwaji RSR (2013) A multi-wavelength analysis of active regions and sunspots by comparison of automatic detection algorithms. *Solar Physics*. 283(1): 67-95.

Copyright statement: © 2013 Springer Netherlands. Full-text reproduced in accordance with the publisher's self-archiving policy.

Editorial Manager(tm) for Solar Physics
Manuscript Draft

Manuscript Number:

Title: A multi-wavelength analysis of active regions and sunspots by comparison of automatic detection algorithms

Article Type: TI: Image Processing in the Petabyte Era

Keywords: "Active Regions"; "Magnetic Fields"; "Coronal Structures"; "Sunspots"

Corresponding Author: Fraser Thomas Watson

Corresponding Author's Institution:

First Author: Cis Verbeeck

Order of Authors: Cis Verbeeck;Paul Anthony Higgins;Tufan Colak;Fraser Thomas Watson;Veronique Delouille;Benjamin Mampaey;Rami Qahwaji

Abstract: The launch of the Solar Dynamics Observatory (SDO) in early 2010 has provided the solar physics community with the most detailed view of the Sun to date. However, this presents new challenges for the analysis of solar data. Currently, SDO sends over 1 terabyte of data per day back to Earth and methods for fast and reliable analysis are more important than ever. This article details four algorithms developed separately at the Universities of Bradford and Glasgow, the Royal Observatory of Belgium and Trinity College Dublin for the purposes of automated detection of solar active regions (ARs) and sunspots at different levels of the solar atmosphere.

The algorithms involved in this study are as follows:

1. The Solar Monitor Active Region Tracker (SMART) extracts, characterises, and tracks the evolution of active regions across the solar disk using line-of-sight magnetograms and a combination of image processing techniques.
2. The Automated Solar Activity Prediction code (ASAP) converts continuum images from heliocentric coordinates to Carrington heliographic coordinates, detects and tracks sunspots using thresholding and morphological methods.
3. The Sunspot Tracking And Recognition Algorithm (STARA) is used to detect and track sunspots from continuum images using a technique known as the top-hat transform.
4. The Spatial Possibilistic Clustering Algorithm (SPoCA) is a multi-channel unsupervised spatially-constrained fuzzy clustering method that automatically segments solar EUV images into active regions, coronal holes and quiet Sun. In the present paper, it is used to detect, characterise and track coronal active regions.

We describe the fundamental properties of each algorithm along with a detailed comparison of outputs obtained from the analysis of about one month of data from the SOHO-MDI and SOHO-EIT instruments during 12 May - 23 June, 2003. We track two active regions over time to study their properties in detail, and exploit the entire dataset to

investigate correlations between physical properties determined by the algorithms. This study allows us to prepare the algorithms in the best possible way for robust analysis of the large SDO data-stream.

The detection rates of the algorithms are compared with findings of the National Oceanic and Atmospheric Administration (NOAA) and the Solar Influences Data Analysis Centre (SIDC). By performing an inter-comparison of the algorithms, the physical properties of the solar features detected are measured at different heights of the solar atmosphere.

A multi-wavelength analysis of active regions and sunspots by comparison of automatic detection algorithms

C. Verbeeck¹ · P. A. Higgins² · T. Colak³ · F. T. Watson⁴ · V. Delouille¹ · B. Mampaey¹ · R. Qahwaji³

© Springer ●●●●

Abstract The launch of the Solar Dynamics Observatory (SDO) in early 2010 has provided the solar physics community with the most detailed view of the Sun to date. However, this presents new challenges for the analysis of solar data. Currently, SDO sends over 1 terabyte of data per day back to Earth and methods for fast and reliable analysis are more important than ever. This article details four algorithms developed separately at the Universities of Bradford and Glasgow, the Royal Observatory of Belgium and Trinity College Dublin for the purposes of automated detection of solar active regions (ARs) and sunspots at different levels of the solar atmosphere.

The algorithms involved in this study are as follows:

1. The Solar Monitor Active Region Tracker (SMART) extracts, characterises, and tracks the evolution of active regions across the solar disk using line-of-sight magnetograms and a combination of image processing techniques.
2. The Automated Solar Activity Prediction code (ASAP) converts continuum images from heliocentric coordinates to Carrington heliographic coordinates, detects and tracks sunspots using thresholding and morphological methods.
3. The Sunspot Tracking And Recognition Algorithm (STARA) is used to detect and track sunspots from continuum images using a technique known as the top-hat transform.
4. The Spatial Possibilistic Clustering Algorithm (SPoCA) is a multi-channel unsupervised spatially-constrained fuzzy clustering method that automatically segments solar EUV images into active regions, coronal holes and quiet Sun. In the present paper, it is used to detect, characterise and track coronal active regions.

We describe the fundamental properties of each algorithm along with a detailed comparison of outputs obtained from the analysis of about one month

¹ Royal Observatory of Belgium, Belgium email: cis.verbeeck@oma.be

² Trinity College Dublin, Ireland email: pohuigin@gmail.com

³ University of Bradford, UK email: t.colak@bradford.ac.uk

⁴ University of Glasgow, UK email: f.watson@astro.gla.ac.uk

of data from the *SOHO*-MDI and *SOHO*-EIT instruments during 12 May – 23 June, 2003. We track two active regions over time to study their properties in detail, and exploit the entire dataset to investigate correlations between physical properties determined by the algorithms. This study allows us to prepare the algorithms in the best possible way for robust analysis of the large SDO data-stream.

The detection rates of the algorithms are compared with findings of the National Oceanic and Atmospheric Administration (NOAA) and the Solar Influences Data Analysis Centre (SIDC). By performing an inter-comparison of the algorithms, the physical properties of the solar features detected are measured at different heights of the solar atmosphere.

Keywords: Active Regions; Magnetic Fields; Coronal Structures; Sunspots

1. Introduction

Sunspot observations are the first form of solar data and were analysed by an observer carefully drawing what they could see. Famous observers include Galileo, who made some of the first sunspot drawings in the 17th century and Carrington who observed in the 19th century. In fact, this method of solar observation is still in use at the Mount Wilson observatory where daily sunspot drawings have been taken since 1917. In the early to mid-1900s, observers such as Hale *et al.* (1919) used photographic plates to record their observations.

Solar data analysis has been revolutionised over the last 50 years with the advent of charge-coupled devices and space-born instruments. In the 1960s computer technology became more accessible to universities and academics and paved the way for a dramatic change in the way solar data was accessed and analysed. At the time, NASA launched the spacecraft Pioneer 6, 7, 8 and 9. These satellites were tasked with observing the solar wind and the solar magnetic fields, and formed the first space solar weather network. The maximum data recording rate of these spacecraft was 512 bits per second. For comparison, the recently launched Solar Dynamics Observatory (SDO) is currently relaying data back to Earth at a rate of 150 000 000 bits per second.

With SDO returning the equivalent of an image with 4096 by 4096 pixels every second, human analysis of the data would require a large team of people working 24 hours a day. The technological advances that have allowed the increased flow of data, such as improving communication bandwidths and on-board processing power, allows us to record data with a much greater cadence and spatial resolution than ever before. However, there are problems with such a large flow of data which can be separated into three categories — storage, transfer and analysis. SDO generates around 1.5 TB of data per day which is unprecedented in solar physics. Moving this volume of data around as well as storing it in convenient places for analysis is essential to make good use of the data. Here we compare new methods of automated analysis for such a large and constant flow of data. These fast and robust solar feature detection tools allow all of the data to be exploited, adding consistency to large scale statistical studies and reducing human biases in the detection of features in solar data.

The research on automated detection and identification of solar features has increased dramatically in the recent years following the increase in the volume of data available. An overview of the fundamental image processing techniques used in automated feature detection algorithms is given in Aschwanden (2010). These techniques are used to detect many features in various types of observations at different heights in the solar atmosphere (Pérez-Suárez *et al.*, 2011). The literature includes descriptions of a diverse collection of algorithms, designed for sensitivity to a specific feature. In Curto, Blanca, and Martínez (2008), full disk white light images were used to automatically detect and cluster sunspots into groups. In Zharkov *et al.* (2004), an automated system for the detection of sunspots on the Ca K1 and *SOHO*-MDI white light images was presented. The authors of Nguyen, Nguyen, and Nguyen (2005) used image processing, and clustering methods on *SOHO*-MDI white light images for the recognition and classification of sunspots according to the modified Zurich class of the McIntosh system. Also Colak and Qahwaji (2008) used *SOHO*-MDI white light and magnetogram images to detect and classify sunspots and magnetic regions. In Gao, Wang, and Zhou (2002), local thresholding and region-growing methods were used to detect filament disappearances. In Benkhalil (2003), active regions (ARs) were detected based on region growing. Filaments were detected in H-alpha images in Shih and Kowalski (2003) using morphological closing operations with multi-directional linear structuring elements to extract elongated shapes. The Singular Spectrum Analysis of signals was used to detect ARs on the solar disk, in Lefebvre and Rozelot (2004). Morphological operations and neural networks, were used in Qahwaji and Colak (2006), for the detection and verification of solar features such as filaments and ARs.

The purpose of this article is to compare four algorithms that have been created independently with the objective of extracting and characterising interesting features on solar images. The tools that we consider are the Solar Monitor Active Region Tracker (SMART) (Higgins *et al.*, 2010), the Automated Solar Activity Prediction code (ASAP) (Colak and Qahwaji, 2009), the Sunspot Tracking And Recognition Algorithm (STARA) (Watson *et al.*, 2009) and the Spatial Possibilistic Clustering Algorithm (SPoCA) (Barra *et al.*, 2009). More detail on how these algorithms operate is provided in Section 3.

This article focuses on the comparison of these algorithms. The performance of the algorithms is compared, as well as the properties of their associated detections in the different data types. The correlations found between the properties is investigated. Many scientific applications are possible from the output of a comparative study such as this. Studies that examine coronal heating as a result of large scale magnetic fields (Schrijver (1987) and Fisher *et al.* (1998)), coupling between the photosphere and corona (Handy and Schrijver, 2001), sources of coronal mass ejections (Subramanian and Dere, 2001), flux emergence and distribution (Liu and Kurokawa (2004) and Abramenko and Longcope (2005)) and flare forecasting (Gallagher, Moon, and Wang, 2002) can all be repeated with these automated detection methods. Using these methods allows a far greater number of features to be analysed and reduces human bias in the detection of features in the solar data.

Here, the evolution of two ARs is studied in detail, including their emergence in each of the atmospheric layers. Lites *et al.* (1995) present a similar multi-layered analysis of the emergence of an AR. As non-potentiality increases in an AR, it may begin to exhibit enhanced coronal activity. This effect has been studied in many papers resulting from dynamic behaviours such as helicity injection (Morita and McIntosh, 2005), turbulent cascades (Hewett *et al.*, 2008; Conlon *et al.*, 2008, 2010), enhanced polarity separation line gradient (Falconer, Moore, and Gary, 2008), and changes in magnetic connectivity (Georgoulis and Rust, 2007; Ahmed *et al.*, 2010). In this paper we simultaneously study multiple behaviours in the same AR using magnetic properties determinations. Finally, photospheric and coronal decay is compared. To our knowledge, this is the first time that automated feature extraction algorithms have been used to study the temporal evolution using properties of magnetic non-potentiality, sunspot characteristics, and coronal activity of ARs simultaneously.

The remainder of this paper details the investigation described above. In Section 2 the observations and solar features of interest are described, followed by Section 3 in which each of the algorithms used in this article are explained. Our results are presented in Section 4, including a correlation study of the complete sample of active regions and a detailed case study of two different active regions. We finish with a discussion of the results and concluding remarks in Section 5.

2. Observations

In this study we analyse data from the interval 12 May – 23 June, 2003. Individual NOAA ARs 10377 and 10365 are studied in detail in Section 4.3, while the detections obtained from each algorithm from the entire data set are studied as a whole in Section 4.2. This particular dataset was selected for the diversity of solar features present. *SOHO*-MDI magnetograms are used for magnetic region detection by SMART, while *SOHO*-MDI continuum images are used for sunspot detection by ASAP and STARA, and *SOHO*-EIT images are employed for active region detection by SPoCA.

The MDI instrument on SOHO provides almost continuous observations of the Sun in the white light continuum, in the vicinity of the Ni I 676.78 nm photospheric absorption line. White light pictures show how the Sun appears to the naked eye and MDI intensitygram images are primarily used for sunspot observations. The MDI data is available in several processed “levels”. The MDI images used in this research are level 2 images, which are smoothed, filtered and rotated (Scherrer *et al.*, 1995). SOHO provides two to four MDI intensitygram images per day with continuous coverage since 1995.

Using the same instrument level 1.8 line-of-sight (LOS) MDI magnetograms are recorded with a nominal cadence of 96 minutes. Magnetograms are recorded in order to measure the magnetic field strengths in the Sun’s photosphere. The magnetograms show the magnetic fields of the solar photosphere, with negative (represented as black) and positive (as white) areas indicating opposite LOS magnetic field orientations.

EIT, the Extreme Ultraviolet telescope aboard SOHO (Delaboudinière *et al.*, 1995), delivers synoptic observations consisting of 1024^2 images of the solar corona recorded in four different wavelengths every six hours. Every SPoCA segmentation in this paper was based on a pair of 17.1 and 19.5 nm EIT images. All images used have been preprocessed using the standard *eit-prep* procedure of the *solar software* library. A fixed-center segmentation with six classes was performed on the logarithms of the image pixel values. The AR center values are 401.74 and 324.25 DN/s in 17.1 and 19.5 nm, respectively. These values were derived from a cumulative run of SPoCA on a dataset of monthly EIT image pairs from February 1997 till April 2005, see Barra *et al.* (2009).

In the case studies presented in Section 4.3, we compare observations of NOAA 10365 and 10377 with flares characterised by the Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) team and distributed in the RHESSI flare list¹.

3. Methods

SMART, ASAP and STARA were created to examine photospheric features such as active regions and sunspots while SPoCA uses images in the extreme ultraviolet to observe coronal holes and active regions at the coronal level. This difference in atmospheric heights allows one to compare different feature detections directly to see how dynamics in an active region in the photosphere are propagated through the solar atmosphere into the corona. In this section, we will describe each of the feature detection algorithms used (Sections 3.1, 3.2, 3.3, and 3.4). The outputs from each of the algorithms are associated using the method explained in Section 3.5.

3.1. The SMART algorithm

The Solar Monitor Active Region Tracker (SMART) is an algorithm which automatically extracts, characterises, and tracks active regions using magnetograms over multiple solar rotations — from first emergence to decay. This allows one to study the complete life-cycle of ARs. A detailed description of SMART is presented in Higgins *et al.* (2010). The algorithm uses a combination of image processing techniques to determine the boundary of an AR. Two consecutive line-of-sight magnetograms are smoothed and thresholded to identify potential features. The two detections are overlaid to identify and remove features which are not present in both magnetograms. The remaining detection boundaries are then dilated to create the final one. Dilation is performed to include nearby decaying and plage fragments which may have separated from the main AR. This is intended to help conserve the measured polarity balance of the AR as it evolves. An example set of SMART detections is shown in Figure 1. In this paper, *SOHO*-MDI LOS magnetograms are used for detection, but recently the algorithm has been adapted for use with *SDO*-HMI magnetograms.

¹http://sprg.ssl.berkeley.edu/~jim/hessi/hsi_flare_list.html

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

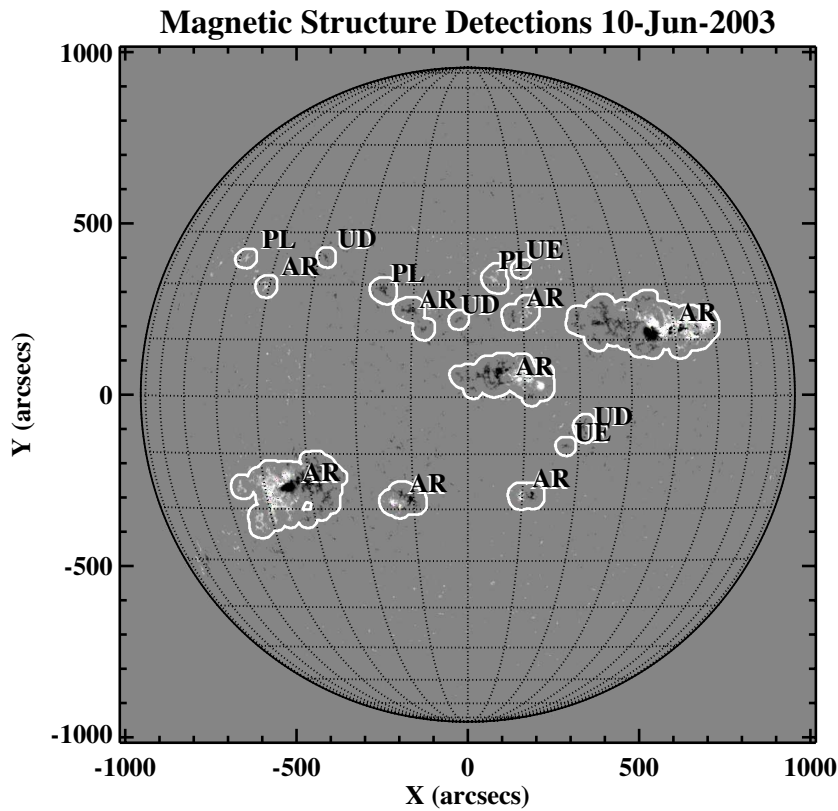


Figure 1. An example set of SMART detections from 10 June 2003. PL, UD, and UE identify three classes of unipolar feature, while AR denotes multipolar features.

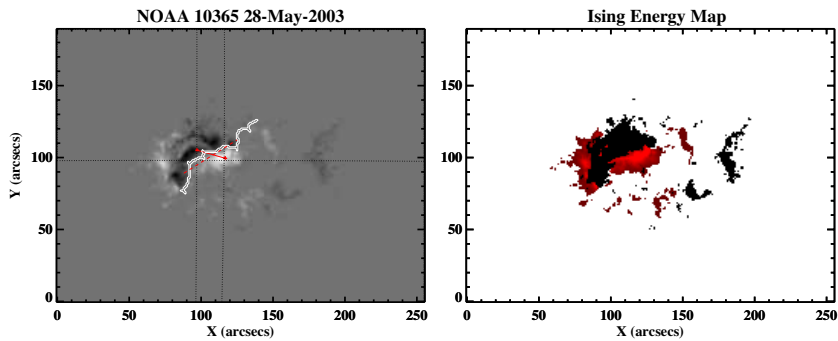


Figure 2. *Left:* NOAA 10365 highlighting the PSL (white contour), bipolar connection line (solid red line), and heliographic latitude reference lines (dotted black lines). *Right:* An Ising energy map of the same region. Red represents the magnitude of energy for each pixel from highest (light) to lowest (dark). Since the connection between pixels of opposite polarity is being represented, the energy map is only shown for one of the polarities.

Several new physical property modules have been added to SMART for this study. Rotational information about an AR is obtained by measuring the angle between the line connecting the centroids of its two largest unipolar regions (of opposite polarity) and the heliographic latitude line passing through the centroid of the AR. The length of the line connecting the centroids is also measured. This bipolar connecting line (BCL) length gives an idea of the compactness of an AR when compared to the overall AR size (left panel of Figure 2). Additionally, the angle detected between the best-fit line to the locus of pixels forming the main polarity separation line (PSL) and the nearest heliographic latitude line is measured. The time derivative of this angle is shown to be a useful proxy for the occurrence of helicity injection in an AR (Morita and McIntosh, 2005), which may be an important flare predictor (LaBonte, Georgoulis, and Rust, 2007). These properties are less informative when studying AR complexes or non-bipolar ARs, since often no main axes can be discerned, making a description of the AR orientation meaningless.

Finally, a calculation of Ising energy which is a proxy for magnetic connectivity within an AR is included (right panel of Figure 2). ARs exhibiting many opposite polarity regions in close proximity have a larger Ising energy than those that are simple with few unipolar regions. The model and details of this module are given in Ahmed *et al.* (2010).

3.2. The ASAP algorithm

ASAP (Automated Solar Activity Prediction) is a tool that is composed of a set of algorithms for sunspot and active region detection (Colak and Qahwaji, 2008) and solar flare prediction (Colak and Qahwaji, 2009). In this paper the newly developed sunspot detection algorithm for ASAP is used. This new sunspot detection algorithm works with continuum images and is described in Colak *et al.* (2010b). In the first step of the algorithm, continuum images in heliocentric coordinates are converted to Carrington heliographic coordinates and enhanced. After that, morphological imaging and intensity thresholding methods are applied to the newly created images for the final detection of sunspots.

ASAP is the collective name for a set of algorithms to process solar images. It can be used for sunspot, faculae and active region detections. Unlike other algorithms described in this paper ASAP uses quick look (in GIF or JPEG format) images for its calculations. Pre-processing of images such as solar disk detection, limb darkening removal and heliographic coordinate transformations are done using internal algorithms. Although heliographic conversion can be computationally expensive, detections and calculations are more accurate compared to the ones done on heliocentric coordinates. Also there is no need for longitudinal corrections in calculations. A tracking algorithm is added to ASAP for this study. This tracking algorithm is made easy by the use of heliographic coordinates. The algorithm finds the intersections between objects (e.g. sunspots) on two consecutive heliographic images since the differential rotation is very small in these coordinates. Only ASAP's sunspot detection is evaluated in this study. ASAP tends to detect small sunspots which can be classified as pores. This is useful especially when grouping and classifying sunspots but the amount

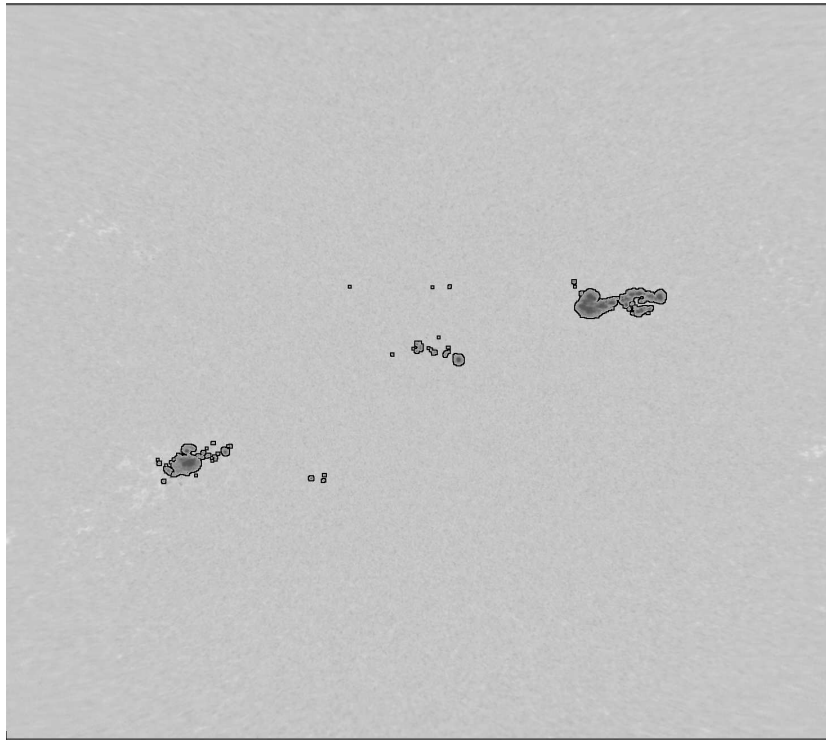


Figure 3. An example set of ASAP detections from 10 June 2003.

of tracked sunspots increases because they are usually visible for a few hours on the solar disk.

3.3. The STARA algorithm

The STARA (Sunspot Tracking And Recognition Algorithm) code was written in 2008 in order to perform consistent long term observations of sunspots over solar cycle 23, see Watson *et al.* (2009). It was originally developed for use with MDI data but has since been extended for use with data from the Solar Dynamics Observatory as well as data from a number of ground based instruments. A simple detection method was required to speed up processing when large datasets were used and suitable techniques were found in the field of morphological image processing. The methods used had added benefits in that the variation of intensity across the solar disk did not need to be taken into account, speeding up the algorithm further.

The STARA detection method works as follows. The data is read in and the image is inverted so that the sunspots appear as bright areas on a dark disk. A morphological top-hat transform is then performed which records the intensity fluctuations on the solar disk and removes the background areas leaving only the sunspot detections. More detail on this step and the top hat transform is given in Watson *et al.* (2009) and Dougherty and Lotufo (2003). A size filter

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

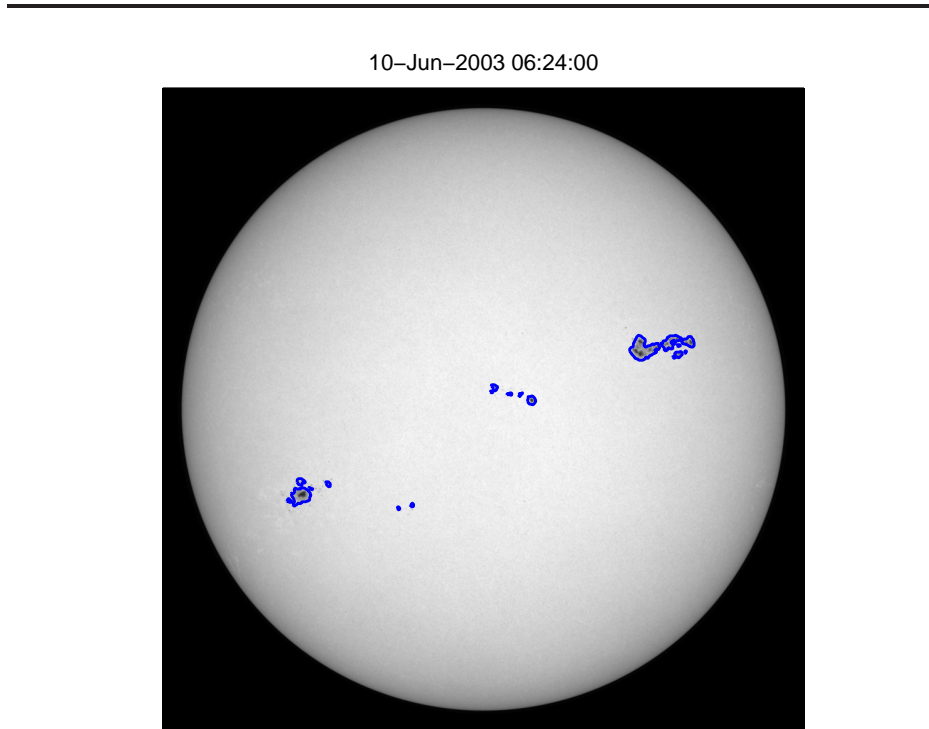
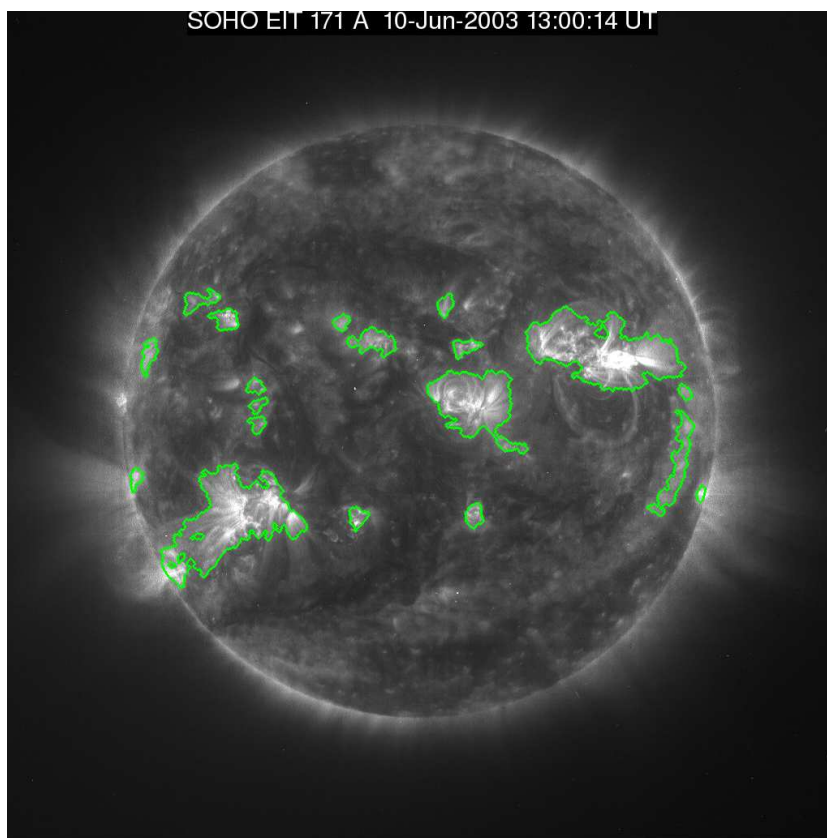


Figure 4. An example set of STARA detections from 10 June 2003.

is then applied which removes areas too small to be a sunspot. The remaining areas are recorded along with their locations as well as the number of umbral regions detected within the sunspot boundary. This is repeated for a number of consecutive images and using the solar rotation model of Howard, Harvey, and Forgach (1990) the sunspots can be tracked throughout the set of data. This allows the evolution of individual sunspots to be followed as well as the overall properties of the sunspot population as a whole. An example of a typical set of STARA sunspot detections is given in Figure 4.

3.4. The SPoCA algorithm

In a previous work (Barra *et al.*, 2009), SPoCA (Spatial Possibilistic Clustering Algorithm) was presented. This is a multi-channel unsupervised spatially-constrained fuzzy clustering method that automatically segments solar EUV images into its regions of interest. It separates in an optimal way active regions (AR), quiet Sun (QS) and coronal holes (CH), even though the boundaries of these regions are not always well defined. The description of the segmentation process in terms of fuzzy logic was motivated by the facts that information provided by a solar EUV image is noisy (corruption by Poisson and readout noise as well as by cosmic ray hits) and subject to both observational biases (line



35 **Figure 5.** An example set of SPoCA detections from 10 June 2003.

36
37 of sight integration of a transparent volume) and interpretation (the apparent
38 boundary between regions is a matter of convention).

39 This technique attributes to every pixel a probability of belonging to a par-
40 ticular class (AR, QS, CH). By assigning each pixel to the class for which it has
41 the largest probability of belonging, an image is segmented. In order to cope
42 successfully with intensity outlier pixels such as those affected by cosmic rays
43 and proton storms, a spatial regularization term was integrated in the clustering
44 algorithm. Since the solar corona is optically thin, and since the intensity in
45 EUV images is obtained through an integration along the line of sight, there is
46 a limb brightening effect in those images which may hinder the segmentation
47 process. Therefore, we first process the EUV images so as to lower the enhanced
48 brightness near the limb. The initial SPoCA class contours are automatically
49 postprocessed using a conditional morphological opening with a circular isotropic
50 element of size unity.
51

52 SPoCA is the only algorithm presented here which detects ARs in the solar
53 corona. The method is generic enough to allow the introduction of other channels
54 or data. It has been applied to *SOHO-EIT*, *SDO-AIA*, *PROBA2-SWAP* and
55 *STEREO-EUVI* images, and could potentially be used on other multi-channel
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7 maps such as Differential Emission maps. In this paper we focus on ARs, but
8 QS and CH can also be detected and tracked.

9 10 3.5. Association of features detected by the algorithms

11
12 The SMART tracking module, Multiple Disk Passage (MuDPie) is used to asso-
13 ciate individual SMART detections of the same physical feature by comparing
14 the centroids of all detections in consecutive magnetograms. A description of this
15 algorithm is provided in Higgins *et al.* (2010). The tracked SMART detections are
16 then associated with the best matched detections in each of the other algorithms
17 as described in the following paragraphs.

18 In order to analyse the relation between the features detected by different
19 algorithms a program has been created in Python. This program associates data
20 from different algorithm outputs in two ways. First outputs from ASAP, STARA
21 and SPoCA are associated to output from SMART based on time and location
22 information. Second, individual association outputs (ASAP vs SMART, STARA
23 vs SMART, SPoCA vs SMART) from the first step are combined using SMART
24 IDs and timing information.

25 For associations, SMART is chosen as base algorithm because SMART de-
26 tections are larger and usually encircle the corresponding ASAP and STARA
27 detections. They are more stable over time than SPoCA detections, due to
28 splitting and merging in SPoCA regions. Also, SMART detects magnetic regions
29 from MDI images which are more frequently available than the continuum and
30 EIT images that the other algorithms are working on. The association rules are
31 described below.

32 33 First Step: Individual associations (ASAP, STARA, SPoCA vs SMART)

- 34
35
- 36 • The time difference between the solar detections under consideration (i.e.,
37 sunspots from ASAP and STARA, active regions from SPoCA versus mag-
38 netic regions from SMART) is calculated.
 - 39 • If the time difference between a magnetic region detected by SMART and
40 a solar region detected by another algorithm is less than 0.2 Julian day
41 and their heliographic bounding boxes intersect, then these detections are
42 associated. Since SPoCA does not deliver heliographic bounding boxes, a
43 bounding box of 5 degrees in longitude and latitude is assumed.
 - 44 • If the same solar detection is associated to more than one SMART region,
45 only the closest (in terms of time and distance between centers) SMART
46 region is selected as associated.
 - 47 • Associations are saved in separate files (3 files; ASAP vs SMART, STARA
48 vs SMART, SPoCA vs SMART) including the selected characteristics from
49 each algorithm output that is going to be analysed.
- 50

51 52 Second Step: Combining all of the associations

- 53
- 54 • The SMART algorithm uses an ID for each magnetic region detected and in
55 this second step, the association data saved in the 3 separate files from the
56
- 57
58
59
60
61
62
63
64
65

first step are combined using this ID and time information. The association data with same SMART ID and closest timing are combined together. Timing information still has to be used due to the difference between the image times.

- The final combined data is saved in one file.

SMART provided 9356 detections (207 magnetic region features), ASAP 3039 detections (952 sunspot features), STARA 1329 detections (433 sunspot features) and SPoCA 1222 detections (190 coronal active region features) within the considered timeframe. In the first step, 714 SMART detections were associated to 2889 ASAP detections, 550 SMART detections were associated to 1315 STARA detections and 1089 SMART detections were associated to 1117 SPoCA detections. In the second step when all these data were combined, 350 detection (33 feature) associations were created for SMART, ASAP, STARA and SPoCA. The daily averages of some of the outputs such as average daily sunspot numbers, active region numbers and average areas are compared to the NOAA active region catalog in Section 4.1. In the considered period, NOAA recorded 217 detections (37 features).

The dilation used in the SPoCA algorithm sometimes causes an AR to merge with a neighboring AR, and to split again later on. In such cases, the association procedure described above does not relate the new SPoCA detection to the corresponding SMART detection. This happened a few times in the case studies in Section 4.3. For these cases, we applied a manual association of SPoCA detections to SMART detections.

4. Results

The feature detections from each algorithm are compared in the following sections. First, in Section 4.1, the detection performance of the individual algorithms is presented, and compared to the corresponding NOAA detections and the daily international sunspot number. Next, a Principal Component Analysis is performed on the full set of detections to probe the overall structure of the physical properties (variables) found by the algorithms in Section 4.2. Finally, in Section 4.3 the evolution and activity of NOAA 10377 and 10365 is analysed in depth, using physical properties determined by each algorithm.

4.1. Algorithm Performance

The performance of the algorithms is measured by comparing the daily total and average values of some of the solar feature properties to each other, to values reported by NOAA² and to the international daily sunspot numbers (SIDC-Team, 2003) between 12 May, 2003 and 23 June.

A comparison of these values is provided in Figure 6. The graph on the upper left side of Figure 6 compares the daily number of sunspots detected

²<http://www.swpc.noaa.gov/ftplib/forecasts/SRS/README>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

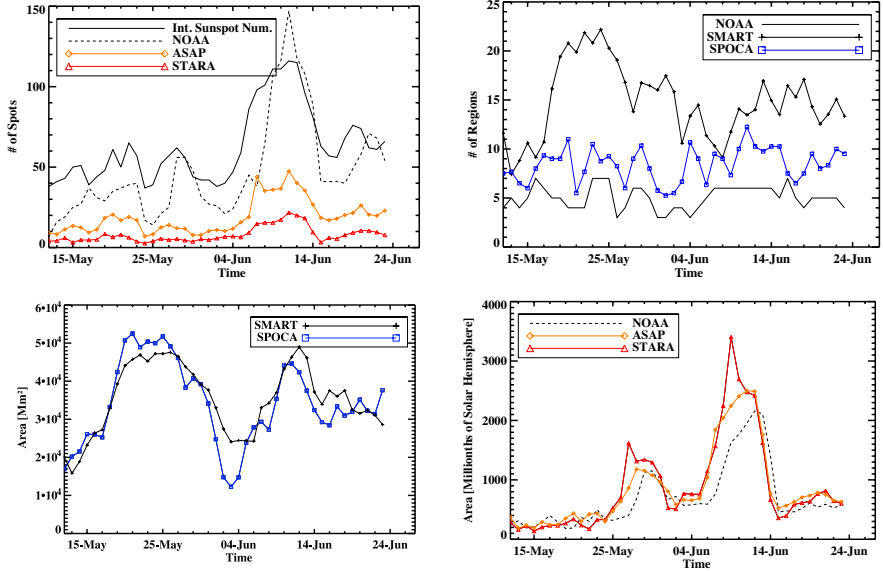


Figure 6. Comparison of average detection results of algorithms to reported NOAA and international sunspot numbers. Up-left: Comparison of number of sunspots detected by ASAP and STARA and reported by NOAA and recorded international sunspot numbers. Up-right: Number of regions detected by SMART and SPoCA compared with the ones reported by NOAA. Down-left: Comparison of average daily region areas detected by SMART and SPoCA. Down-right: Comparison of average daily sunspot areas detected by ASAP and STARA versus NOAA.

by ASAP and STARA to the total number of spots within NOAA regions and to the international sunspot number R_i . Generally, peaks and valleys in all of the series follow each other but the international sunspot numbers and the sunspot numbers for NOAA are usually higher than the sunspot numbers for ASAP and STARA. When sunspots are detected manually, each umbra within a penumbra is counted as one sunspot, whereas the automated algorithms discussed here count each penumbra as one sunspot although it could have more than one umbra within. Therefore the difference in sunspot numbers increases when the number of complex sunspot regions increases. Also, the number of sunspots detected by ASAP is always higher than the ones detected by STARA. This is because ASAP tends to detect very small sunspots (sometimes pores) while STARA has a threshold for the size of sunspot candidates.

The graph on the upper right side of Figure 6 compares the daily number of regions detected by SMART and SPoCA to the total number of NOAA regions. SMART and SPoCA detect more regions than NOAA because the NOAA number is given to a region only if it has one or more sunspots, while SMART and SPoCA regions do not depend on the existence of sunspots within. Because of the projection effects of large coronal loops, two close but distinct regions in the photosphere will often be detected by SPoCA as one region. This explains why SMART has a higher tally of daily regions than SPoCA. This effect is most important near the solar limb.

Comparison of the areas of AR and sunspots as detected by the four algorithms and in the NOAA catalog are presented in the lower part of Figure 6. SMART, ASAP and STARA areas were corrected for the projection effect that compresses the observed area as the region moves away from the central meridian. Since the projection of coronal loops onto the solar disk enlarges the observed area while the abovementioned effect decreases the observed extend of the base area, no net effect is expected for the observed SPoCA area. Hence, the raw SPoCA area is presented. Areas are given here in millionths of solar hemisphere, so as to be consistent with the units of the NOAA catalog.

The graph on the lower left side of Figure 6 shows the comparison between active region areas detected daily by SMART and by SPoCA. Considering we are dealing here with photospheric versus coronal areas, a pretty good correspondence between the areas is obtained. Both SMART and SPoCA areas vary smoothly. Moreover, they are large enough to include all group sunspots if these are present, and to see in a consistent way changes in topology or complexity.

The graph on the lower right side of Figure 6 compares the sunspot areas detected daily by ASAP and STARA to the NOAA sunspot areas. These three series follow each other consistently with values very close to each other, but it looks like there is a one day shift in NOAA sunspot areas. This could be caused by the averaging of the results detected by the algorithms in order to find daily values as mentioned earlier. NOAA region areas are not daily averages and are calculated once every day.

In all the graphs, series tends to follow each other and have similar changes within time and although there are differences in the values these differences are most of the time stable or change with the complexity of the regions or sunspots.

4.2. Principal Component Analysis

In this section, Principal Component Analysis (Jolliffe, 2002) is employed to improve our understanding of the structure of the system of variables observed by the algorithms described above. The Principal Components are the directions in n -dimensional variable space in which the dataset exhibits most variance. The following SMART variables were considered in this analysis: Schrijver R value (Schrijver, 2007), length of the strong gradient line, magnetic flux, maximum B field, area, length of the bipole connecting line, Ising energy, and Ising energy per pixel (Ising E ppx). For ASAP, the sunspot area and number of sunspots were taken into account. Since ASAP and STARA data are very similar, we did not include any STARA variables, as this would provide too much weight to the sunspot numbers and areas. Finally, the SPoCA variables retained for this survey are the raw AR area as well as the maximum, variance, kurtosis and skewness of the EUV intensity over the AR pixels. We excluded data points corresponding to regions whose center was more than 60 degrees from the central meridian, as projection errors involved become too large. Table 1 lists the percentage and cumulative percentage of the variance explained by the Principal Components. As 84% of the variance is contained in the first four components, we plot the correlations of the observed variables with components 1 and 2, and 3 and 4, respectively, in Figures 7 and 8.

Table 1. Percentage and cumulative percentage of the variance of the 15-dimensional variable space described above, that can be explained by the consecutive Principal Components. Note that the first four Principal Components comprise 84% of the variance.

Component	% variance	Cumulative % variance
1	52.86	52.86
2	14.61	67.47
3	10.30	77.77
4	6.55	84.32
5	5.17	89.49
6	3.61	93.11
7	2.82	95.93
8	1.61	97.54
9	0.93	98.47
10	0.44	98.92
11	0.39	99.30
12	0.29	99.59
13	0.22	99.81
14	0.12	99.93
15	0.07	100.00

The projection of every variable upon the first and second Principal Components is plotted in Figure 7. This shows the extent to which these variables are correlated with the first and second Principal Components. The Schrijver R value, the length of the strong gradient line Lsg, the Ising energy (also per pixel), the ASAP sunspot area, and, to a lesser extent, the flux, clearly form a cluster in this plot (cluster 1). This means that these variables are closely related, suggesting they basically measure the same underlying physical entity. This related behaviour of PSL length, R value, Ising energy and ASAP area is apparent in Figures 10, 12, 14, 16, 17 and 19.

These variables are linked to a measure of complexity of the AR and its capability to produce a flare, see Colak *et al.* (2010a).

The maximum, variance, skewness and kurtosis of the SPoCA EUV intensity seem to form another cluster (cluster 2). It is hardly a surprise that those coronal EUV parameters are closely related. For instance, a high value of maximum EUV flux means that there will be a long upper trail in the distribution of flux values, implying a high skewness. Time series of the variables in this cluster behave in a similar fashion (results not shown here).

A third tentative cluster comprises the maximum magnetic field B max, the length of the Bipole Connecting Line and the SPoCA area. We do not, however, discern very related behaviour of B max, BCL length and SPoCA area in Figures 10, 11, 12, 14, 15, 16, 17, 18 and 19.

In Figure 8, the projections of the variables upon the third and fourth Principal Components are shown. We see that most of the variables in the first cluster still form a cluster in these two other dimensions (R , Lsg, Ising energy, flux,

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

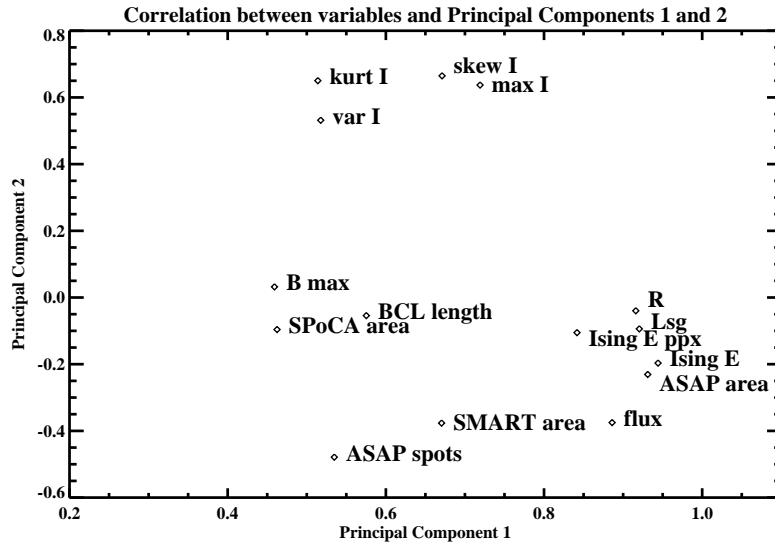


Figure 7. The projections of the algorithm variables upon the first and second Principal Components are plotted. They provide a measure of the extent to which these variables are correlated with the first and second Principal Components.

ASAP sunspot area). This confirms a strong relationship between the variables involved. Though the SPoCA maximum intensity does not belong to this 4-dimensional cluster, its projection upon components 3 and 4 falls within the projection of this cluster. The other two clusters found in Figure 7 do not form clusters in this 4-dimensional space: though their projections upon the first two Principal Components form a cluster, their projections upon components 3 and 4 are scattered. This means the variables within these 2-dimensional clusters are related, but not as closely as the ones within cluster 1.

Note that all variables are highly correlated with Principal Component 1, especially the ones in cluster 1. For Principal Component 2, it's mainly cluster 2 which is highly correlated. Cluster 1 is not correlated with components 3 nor 4, meaning its contribution is mainly present in component 1, and to a lesser extent in component 2.

4.3. Case Studies

In the following section we analyse the time evolution of the ARs which emerge as NOAA 10377 and 10365. Of special interest is how activity in the corona results from changes in the photosphere. Drawing this connection is essential for flare prediction, since the photosphere is more easily physically characterised than the corona, where flares actually occur. This connection is not well understood, see e.g., the work of Leka and Barnes (2007) and of Handy and Schrijver (2001), with the references therein.

As mentioned before, we compare observations of NOAA 10365 and 10377 with flares characterised by the Reuven Ramaty High Energy Solar Spectroscopic

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

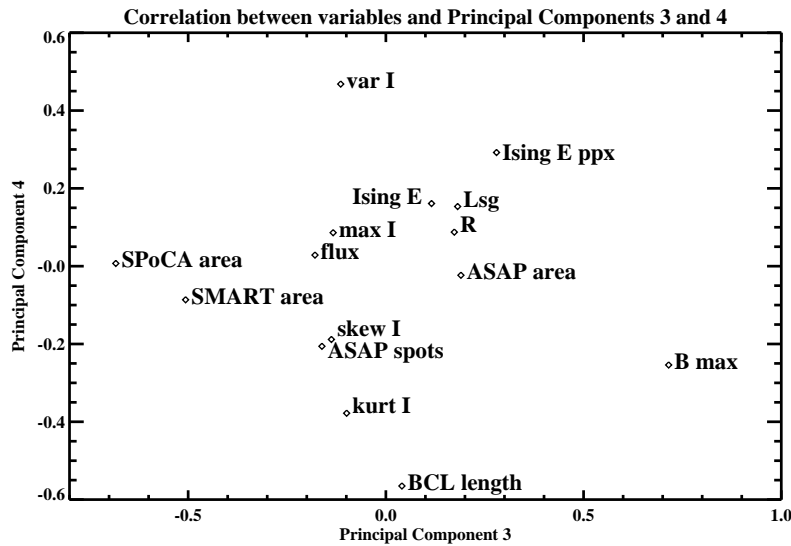


Figure 8. The projections of the algorithm variables upon the third and fourth Principal Components are plotted. They provide a measure of how much these variables are correlated with the third and fourth Principal Components.

Imager (RHESSI) team and distributed in the RHESSI flare list³. The flares, which have been associated with the individual ARs by the RHESSI team, are represented in plots (Section 4.3) as downward pointing arrows, whose size is logarithmically proportional to their peak count rate.

4.3.1. NOAA 10377

NOAA 10377 first emerges just before rotating onto the visible disk on 4 June 2003. It continues to gradually develop as it progresses across the disk producing very little activity (only one B9.1 event is listed in the NOAA events catalog⁴). Figure 9 shows the SMART detection of 10377 in red, while other features are outlined in black. The extended dashed blue contours are SPoCA AR detections and the small symbols and contours are sunspot detections from ASAP and STARA, respectively. It is clear from Figure 9 that positions of the SMART, ASAP, STARA and SPoCA detections agree very well. Whereas the sunspots detected by ASAP and STARA are well confined within the SMART magnetic region boundary, the SPoCA region most often contains most of the SMART detection. In case a coronal loop is formed to a nearby AR, the SMART detection will not lie near the center of the AR detected by SPoCA. This is especially apparent near the solar limb.

³http://sprg.ssl.berkeley.edu/~jimm/hessi/hsi_flare_list.html

⁴<http://www.swpc.noaa.gov/ftplib/indices/events/README>

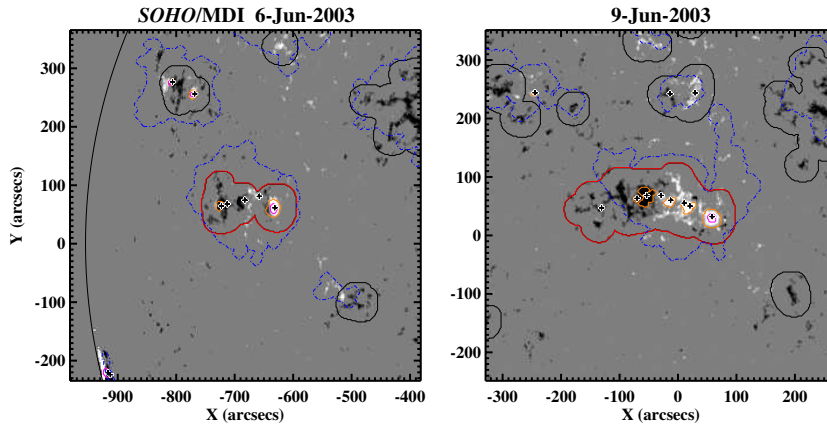


Figure 9. A comparison of NOAA AR 10377 detections. ASAP sunspots are represented by black crosses. The contours represent SMART in black (with NOAA 10377 outlined in red) for the magnetic features, SPoCA in dashed blue for coronal features, and STARA in orange for sunspot penumbrae and magenta for umbrae.

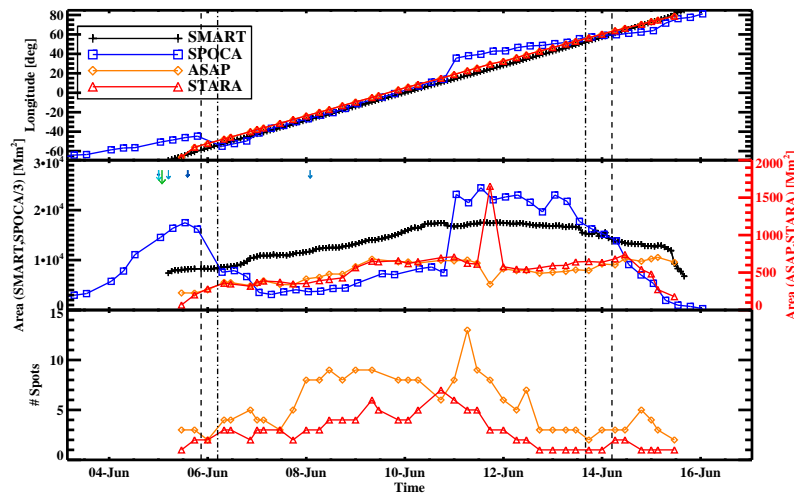


Figure 10. Time series of position, area, and sunspot information characterising the evolution of NOAA AR 10377. The legend indicates symbols and colors for each of the detection algorithms. The axes of the area plot are split between left (SPOCA and SMART) and right (ASAP and STARA). The SPOCA data have been divided by 3 for display.

Figures 10–12 show the evolution of NOAA 10377 as it progresses across the disk. In the top panel of Figure 10 the Stonyhurst longitudes of the region centroids from each algorithm are shown.

The vertical dotted lines indicate where the AR magnetic bounding box edges (dashed-dotted) and centroid (dashed) cross -60 and 60 degrees longitude. The cosine correction used for magnetic field parameters is not reliable outside of this

1
2
3
4
5
6
7 range. At 60 degrees, sunspot visibility is $\sim \frac{1}{3}$ of that at disk center (Watson
8 *et al.*, 2009).

9 The centroid of the magnetic footpoints (SMART) follows behind the sunspot
10 centroids (ASAP and STARA). Since the longitudinal speed of the white light
11 and magnetic detections are the same, this implies that the eastern polarity
12 of the AR extends beyond its embedded sunspots, while the western polarity
13 remains compact.

14 The centroid of the coronal structure detection (SPoCA) is at times cospatial
15 with SMART, but exhibits discontinuities as loops are formed with nearby mag-
16 netic regions ahead of the AR. As the coronal NOAA region 10377 is close to
17 10375, this last region interferes with the SPoCA detections. From 3 till 6 June,
18 SPoCA detects both NOAA regions combined as one region. When this region
19 splits into two parts on June 6, the SPoCA longitude and area curves take a jump
20 downwards, and can now be better compared to the photospheric structures.
21 This changes when the two NOAA regions merge again into one SPoCA region
22 on 11 June. Whenever the region detected by SPoCA corresponds to the region
23 detected by the other algorithms, all four longitudes agree fairly well.

24 The areas in this section are given in Mm^2 (projected upon the photosphere).
25 The total sunspot area determined by ASAP and STARA (Figure 10, middle) is
26 very similar except for one data point near 12 June 2003. This is due to the MDI
27 image on 11 June 2003 at 17^h36^m UT being distorted. Most of the distortion is
28 visible on the south limb of the image where this area is darker than the rest
29 of the solar disk. Because ASAP detects the solar disc directly from the image,
30 while STARA uses FITS keywords, the determination of the solar disc by these
31 two methods is different. This explains why on this image ASAP sunspot areas
32 are much smaller than the STARA area: Whereas the distorted area is detected
33 by STARA as a big sunspot, it is completely discarded by ASAP. The SMART
34 and SPoCA areas of photospheric magnetic regions and coronal active regions
35 obey the same general trend as the sunspot areas, though the absolute scales
36 are different. For SPoCA, this is only true between 6 and 11 June, when the
37 coronal detection corresponds to NOAA region 10377 only. Before and after this
38 period, the SPoCA area increases hugely since the coronal detection of 10377
39 then connects to that of 10375.
40

41 While the area measurements are stable, the total number of sunspots is much
42 less so. The total area is dominated by the largest sunspots, while the total
43 number of spots is affected by small transient spots which ASAP is especially
44 sensitive to.
45

46 In the top panel of Figure 11, the emergence of the magnetic structure of
47 10377 is clearly seen in measurements of its total flux. The AR is stable until
48 ~ 8 June 2003 when a phase of rapid emergence begins, lasting until ~ 11 June
49 2003 when the total magnetic flux has more than doubled.

50 The same general smooth trend is observed in the SPoCA total EUV intensity
51 between 6 and 11 June (i.e., when the SPoCA detection corresponds to NOAA
52 10377 only). On June 11, the coronal detection of NOAA 10377 merges with
53 that of NOAA 10375. From that moment on, we see a clear decay of the total
54 EUV intensity of this combined coronal region. Note that both SMART flux and
55 SPoCA total EUV intensity behave quite similar to the region area time series.
56
57
58
59
60
61
62
63
64
65

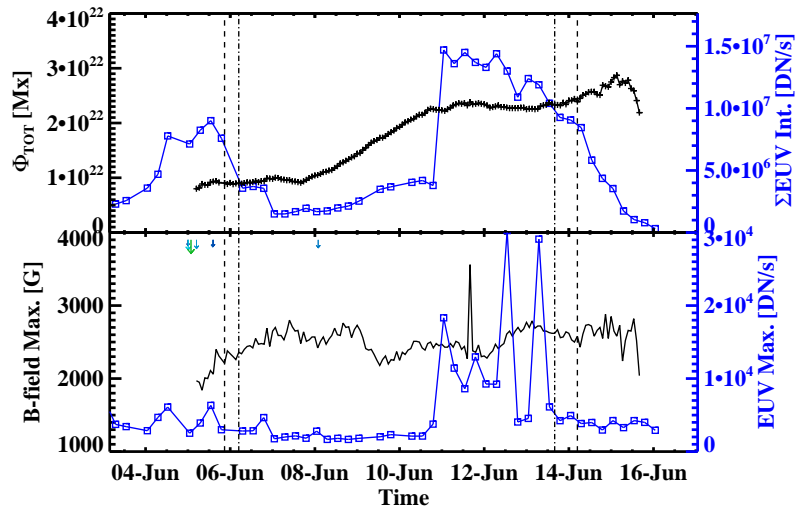


Figure 11. Time series of (top) total magnetic flux, total EUV intensity, (bottom) maximum magnetic field, and maximum EUV intensity for NOAA AR 10377. The axes of the plots are split between left (magnetic field properties, black crosses) and right (coronal properties, blue squares). RHESSI flares associated with the AR are indicated by downward arrows.

In the bottom panel the maximum magnetic field is much less stable than the flux, but also does not produce any dramatic changes.

The maximal SPoCA EUV intensity is rather constant between 6 and 11 June for NOAA region 10377, but exhibits 3 clear peaks afterwards. The first peak, on 11 June, must be attributed to SPoCA suddenly joining NOAA 10375 to its detection. The peak on 12 June, near 13^h, is probably related to the M1.0 flare in 10375 around 13^h58^m, whereas the 13 June, 7^h peak probably involves the M1.8 (6^h28^m) or C6.1 (7^h10^m) flares from 10375. This confirms our expectations that SPoCA maximum intensity is a good indicator of solar eruptions.

Magnetic properties related to polarity mixing and complexity are shown in Figure 12. In the top panel, the angle between the bipole connection line and PSL is presented. Because this AR has only a PSL of several megameters, as seen in the middle-top panel, this angle is very uncertain, as a small growth in the PSL detection in any direction can cause the angle to change dramatically, in this case. In the middle-bottom panel the total flux near the PSL, R , is very small until it begins to increase as a false PSL is detected due to the near-horizontal fields of the large leading polarity sunspot approaching the west limb.

A proxy for magnetic connectivity, Ising energy, is shown in the bottom panel. This property increases during the main magnetic emergence phase (~ 8 to 10 June 2003) since it is dependent on the magnetic field strength and inversely dependent on the distance between individual magnetic elements. Comparing Figure 10 and Figure 11, we see that the total magnetic flux increases faster than the magnetic area, implying that the AR magnetic fields increase in intensity relatively faster than they diffuse. The Ising energy per pixel (dashed line)

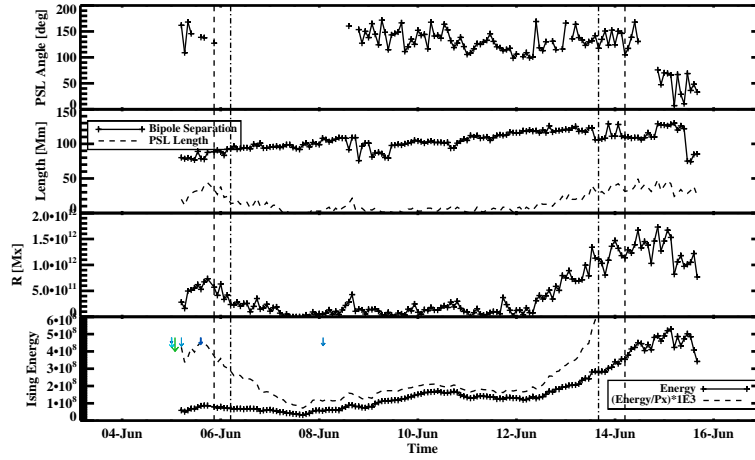


Figure 12. Time series of (top) PSL orientation with respect to the bipole separation line, (middle-top) bipole separation line length (crosses) and PSL length (dashed), (middle-bottom) R , and (bottom) Ising energy (crosses) and Ising energy per pixel (dashed; multiplied by 1000 for display) for NOAA AR 10377.

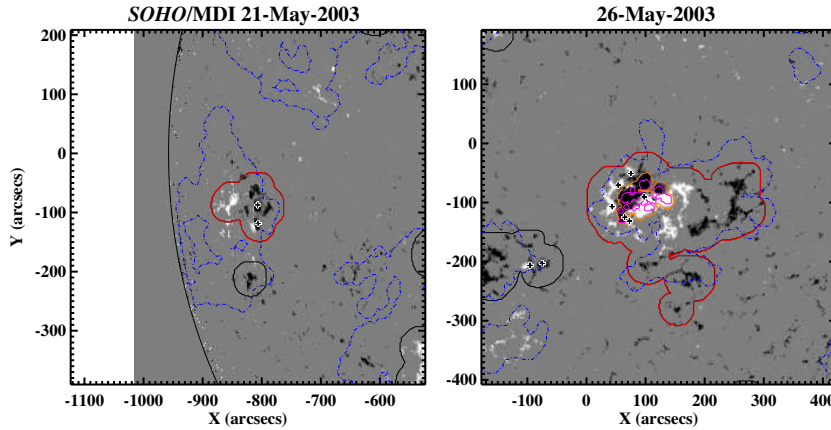


Figure 13. A comparison of detection contours for NOAA AR 10365 using the same color scheme as Figure 9.

appears to be more susceptible to geometrical effects, as the large decrease near the west limb and increase near the east limb occur at the same time as false PSLs form in the leading sunspot. It should be noted that this quantity was calculated without remapping the data to disk-center as done in Ahmed *et al.* (2010), giving it a large viewing angle dependence.

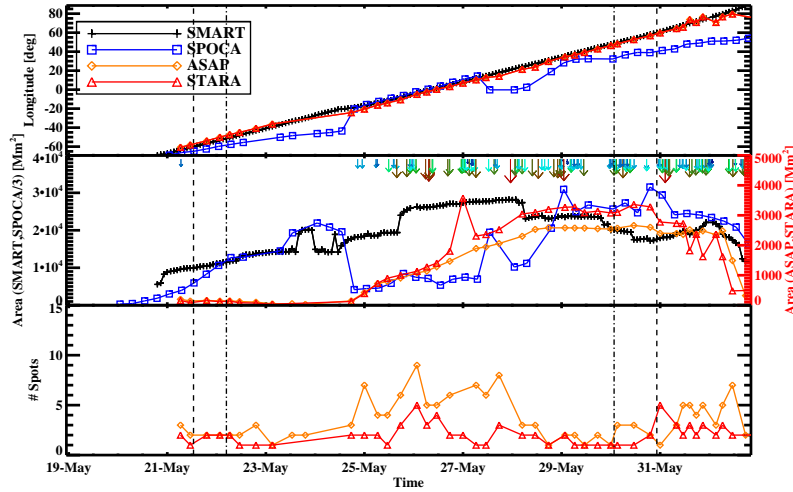


Figure 14. Time series of position, area, and sunspot information characterising the evolution of NOAA AR 10365. The labelling is the same as Figure 10.

4.3.2. NOAA 10365

Active region NOAA 10365 rotates onto the visible solar disk on 19 May 2003 at heliographic latitude -5° . At this point 10365 is mature and decaying, having emerged and evolved on the far side of the Sun. On 24 May 2003, a new bipolar structure rapidly emerges in the extended plage of the trailing (positive) polarity. NOAA switches the 10365 designation to this newly emerged bipole several days later. As the bipole evolves it develops a strong double PSL by merging with the old magnetic flux. It produces many C- and M-class flares and several X-class flares. The AR progresses around the visible disk, eventually returning as NOAA 10386. The onset of decay occurs as C- and M-class flares are produced with decreasing frequency.

Figure 13 shows a comparison of the heliographic positions and sizes of two sets of SMART, ASAP, STARA and SPoCA detections of NOAA 10365. We can see that positions of the SMART, ASAP, STARA and SPoCA detections agree well. The SPoCA detection, however, includes many more coronal loops than just those of NOAA 10365.

Before 24 May, SPoCA adds NOAA region 10367 to its detection of 10365. From 24 May till 27 May it only detects 10365, on 27 May at 13^h there is a single data point where these regions are merged by SPoCA, and from 29 May at 1^h onwards, SPoCA merges them for good. The longitudes of all detections within 24–27 May agree well. Afterwards, we see the SPoCA longitude drifting, reflecting the changes in the combined shape and hence centroid position of both coronal regions involved.

Unlike 10377, the magnetic centroid of 10365 at first trails behind the sunspot centroid but then precedes it, as evidenced by the top panel in Figure 14. This is

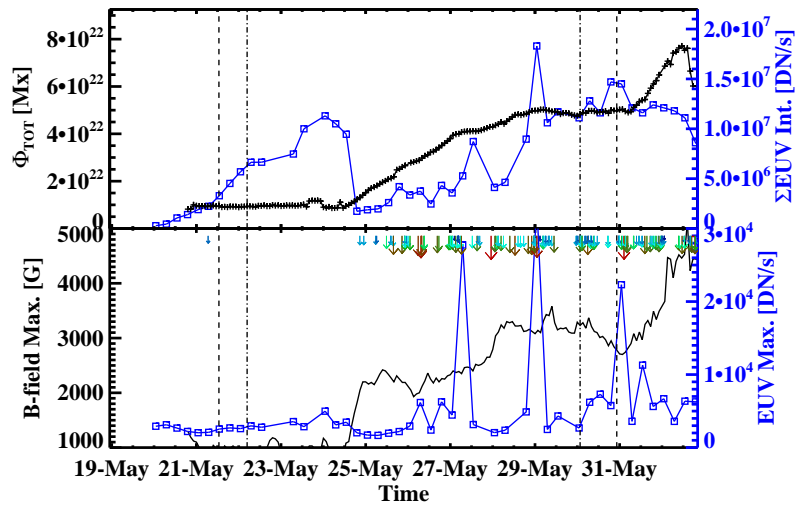


Figure 15. Time series of (top) total magnetic flux, total EUV intensity, (bottom) maximum magnetic field, and maximum EUV intensity for NOAA AR 10365. The labelling is the same as Figure 11.

because the new bipole, which develops many spots, emerges behind the existing weakly spotted bipole. The new emergence is clear in the plot of total sunspot area (middle panel), and is unclear in the magnetic and EUV area plots since the new bipole emerges partially within the boundary of the old one. Note that all areas for NOAA 10365 are much larger than those for simpler region 10377. For SMART, area is very sensitive to weak magnetic plage, however. This can be seen in the sudden jumps around May 25 and 28, which is due to nearby plage temporarily merging with the AR.

When we examine the number of spots detected by the ASAP and STARA algorithms, we find that ASAP is almost always detecting more spots. This shows us the effects of the size threshold applied in STARA as opposed to ASAP where many small features, including pores, are included in the analysis. The jump in STARA area on 27 May can be attributed to a bad datafile (note that there is no ASAP data point at that time).

From 25 May onwards, the total magnetic flux increases gradually to over four-fold the initial value during development and levels off around 29 May (see Figure 15). The maximum magnetic field increases abruptly on 25 May and also increases over time, albeit in a less smooth fashion than the magnetic flux. Note that the maximum magnetic field did not show a particular increasing or decreasing trend in the case of simpler NOAA region 10377.

The M1.9 flare in 10365 at 5^h34^m on 26 May most probably causes the small peak in SPoCA total intensity around that time. One day later, we observe a large peak in SPoCA total intensity, which can be attributed to the M1.6 flare in 10365 at 05^h06^m on 27 May. It is strange that there is no sign of the X1.3 flare at 22^h56^m on 27 May in the SPoCA maximum intensity data, even though

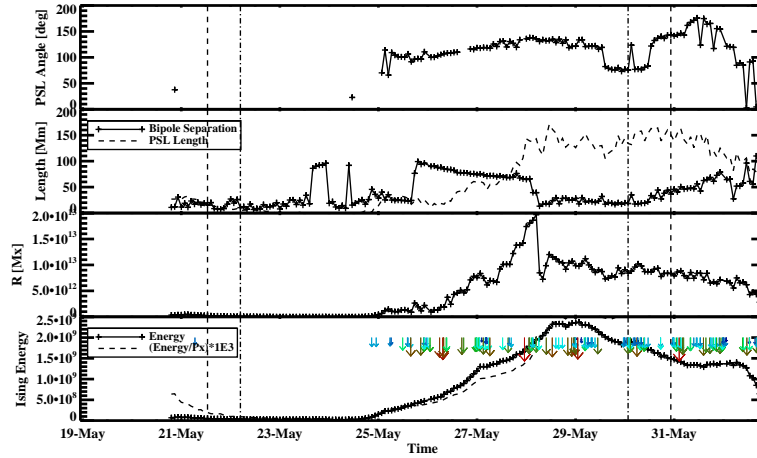


Figure 16. Time series showing proxies for the complexity and polarity mixing in 10365. The plots and labelling are the same as for Figure 12.

this flare occurred in the same NOAA region 10365. The X1.2 flare from this region at 00^h51^m on 29 May, on the other hand, generates a very high SPoCA maximum intensity value. Also the M9.3 flare in NOAA 10365 of 02^h13^m on 31 May is registered quite well.

The last two flares are even visible in the total SPoCA intensity, which shows more or less a gradual increase over time, but less smooth than in the case of the simpler NOAA region 10377.

Signatures in the evolution of the magnetic topology of NOAA 10365 precede its intense coronal activity, indicated by the associated RHESSI flares represented by downward arrows in Figure 16. Just before 25 May 2003 the new emergence causes a jump in the main bipole separation line length. As the emergence continues and strong PSLs develop, this length decreases, while the total PSL length increases, as shown in the middle-top panel of Figure 16. Also, there are signs of gradual helicity injection as the angle between the main bipole connection line and the main PSL grows from near perpendicular to around 120 degrees (top panel). The flux near PSLs (R) grows during this time, as does the Ising energy (middle-bottom and bottom panels, respectively). A bump in R just before 26 May is followed by an intense RHESSI flare. Intense flaring begins again around the second bump in R on 28 June. Examining the development of Ising energy, it appears that the most intense flaring begins around a day after the energy plateaus.

Comparing the evolution of the PSL angle in 10377 and 10365, we see that it varies more smoothly in the more complex region 10365. In 10365, the PSL length builds up to more than 3 times the values it reaches in 10377, testimony to the higher magnetic energy involved in region 10365. Both the R and Ising

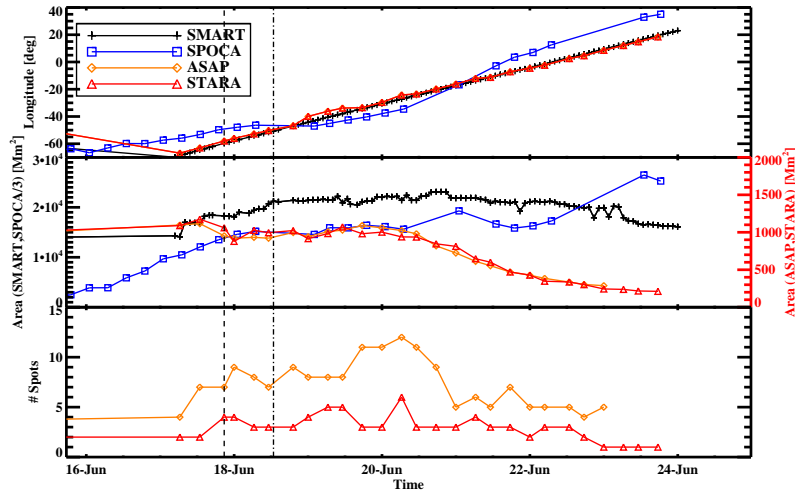


Figure 17. Time series of position, area, and sunspot information characterising the decay phase of NOAA AR 10365 (renamed 10386) during its second disk passage. The labelling is the same as Figure 10.

energy time series for 10365 are smoother and reach up to five to ten times the maximum values for 10377.

NOAA 10365 returns for a second disk passage, renamed 10386. We are able to observe its decay phase, as shown in Figure 17. As no RHESSI data on flares is available for this period, no flare arrows were added to Figures 17–19.

Whereas the longitude of the SMART magnetic centroid increases linearly with time, the ASAP and STARA sunspot centroids show small departures from this line between 19 and 21 June, preceding the magnetic centroid. SPoCA has a hard time tracking just this particular NOAA 10386, and merges with regions 10388 and 10389, rendering interpretations of the SPoCA curves for NOAA 10386 rather tempting.

While the magnetic area remains rather constant, the total sunspot area clearly decreases (middle panel), and has already decreased substantially since the previous passage as NOAA 10365.

The total magnetic flux clearly decreases (top panel, Figure 18) as its magnetic fields diffuse and weaken. Comparing the values to Figure 15, we notice that the flux had already decreased quite a lot since the previous solar rotation.

The total EUV intensity does not change substantially, regardless of the weakening magnetic footpoints, but it too has decreased since the previous solar rotation. Its rise near the end is an artifact due to SPoCA’s detection merging into a huge region near the limb.

The maximum magnetic field value shows a gradual, though not very smooth decrease, and it too has decreased since the previous passage. The maximum EUV intensity shows no particular trend, and though a few flares are still detected, the intensity levels are nothing like those accompanying flares in the

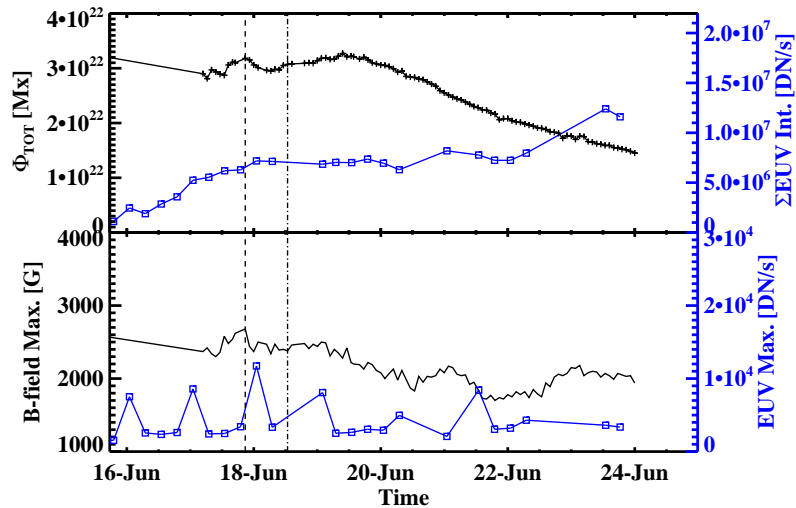


Figure 18. Time series of (top) total magnetic flux, total EUV intensity, (bottom) maximum magnetic field, and maximum EUV intensity for NOAA AR 10365 on its second disk passage as 10386. The labelling is the same as Figure 11.

previous passage. The peak on 18 June at 1^h UT, for instance, can probably be associated to the M6.8 flare produced by region 10386 at 22^h27^m on 17 June.

The PSL length has decreased since the previous solar rotation, and shows a further gradual decrease in Figure 19. The same is true for both R and the Ising energy.

5. Discussion and Conclusion

In this paper, we demonstrated the capability to perform a comparative study of the properties of photospheric active regions, sunspots, and coronal active regions, using independent and fully automated feature detection algorithms. The algorithms have a similar performance as measured in number or area of sunspots or active regions, and provide an excellent agreement with the NOAA numbers and areas as well as the daily international sunspot number (Section 4.1). Moreover, the algorithms corroborate each other by the good correspondence of their positions and areas (Section 4.3). Together, they allow us to directly compare photospheric properties such as number of sunspots or magnetic flux to coronal properties such as maximum EUV intensity.

The algorithms sketched above are also complementary to some extent: whereas the SMART algorithm provides many magnetic properties, ASAP and STARA each provide their own approach to sunspots, with ASAP taking into account even pores and STARA setting a firm size limit on the sunspots. SPoCA, on the other hand, allows us to have a look at the coronal loops and their properties.

The present paper is just a first exploration based on a month's worth of SOHO data, and paves the way for a more systematic analysis of the full SOHO

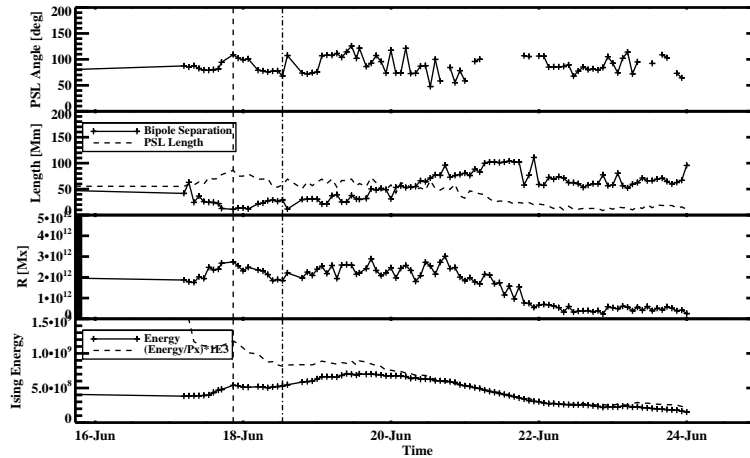


Figure 19. Time series showing proxies for the complexity and polarity mixing in NOAA 10386. The plots and labelling are the same as for Figure 12.

archive, as well as in detail studies of combined photospheric and coronal SDO data sets.

The solar feature tracking method used in this paper is described in Section 3.5 and is a simple algorithm associating the magnetic flux-weighted centroids of detected features to associate them between frames. Its advantage over other algorithms is that it allows features to be tracked between multiple disk passages by comparing newly detected features to where they would have appeared near the end of their previous disk passage. Multiple disk passage tracking is often essential for analysing the complete life-cycle of an AR, as exemplified in the analysis of NOAA 10365 (Section 4.3.2).

Note that the algorithms presented above are automatic, independent, and unsupervised (they do not rely on a training set). A more sophisticated AR tracking algorithm is Yet Another Feature Tracking Algorithm (Welsch and Longcope, 2003, YAFTA) which utilises the spatial extent of features. This yields a more robust knowledge of merging and splitting. Future work on AR evolution should combine both morphological information and multiple disk passages for tracking.

Several physical property modules were added to SMART for this work: magnetic feature orientation and rotation, polarity separation line orientation and rotation, main polarity separation distance, and the number of unipolar regions within the feature. These modules will be used for future large-scale AR studies and will be added to the pipe-line versions of SMART running at the Heliophysics Events Knowledgebase⁵ (Hurlburt *et al.*, 2010) and included in the

⁵<http://www.lmsal.com/hek/index.html>

Heliophysics Integrated Observatory⁶ (Bentley *et al.*, 2009). In the future, other property modules will be added to calculate a physically motivated magnetic connectivity measurement (Georgoulis and Rust, 2007), multi-scale energy spectrum slope (Hewett *et al.*, 2008), and multi-fractal spectrum properties (Conlon *et al.*, 2008).

Although there were not any changes in the sunspot detection part of ASAP since we started this work and no bugs have been discovered, some computational issues have been noticed during the course of these studies. An important one is the application of algorithm to SDO/HMI images larger than 1024x1024, this shows that heliographic conversion algorithms are slow and have to be updated to tackle larger images. Also experiments in this manuscript show how automated detection algorithms are vulnerable to corrupted or missing data. None of the new or existing data are 100% correct or complete and even slight errors in the images can affect the algorithms' capabilities to work. One interesting contribution to the solar imaging community in the future would be to create an automated algorithm to detect structural or visible errors on solar images or data. This can also improve automated systems that are using solar images for prediction or modelling.

STARA has undergone very few changes over the course of this work as the code was well established beforehand. Nevertheless, some subtle problems have been discovered in the process. As sunspots approach the limb (at longitudes greater than 75 degrees) the sunspot position returned by the code quickly loses accuracy. This is a common problem with feature detection methods as the geometrical foreshortening effects test the limits of automated systems. There are also potential problems present with bad data. Obviously the best remedy is to remove bad data altogether but with MDI it is possible to have data with only half of the solar disk present or large artifacts. Both of these situations can have large effects on the detected global properties and cause problems with analysis.

Since the publication of Barra *et al.* (2009), the SPoCA algorithm was optimized and extended in several ways:

- In order to gain more consistent results, we introduced a more compact fuzzy membership function and some constraints on the η_i parameters.
- The limb correction is now applied in a continuously increasing way towards the limb instead of introducing it abruptly from some point onwards.
- For individual AR detection, first the Bright Points are removed (size threshold = 1500 square arcseconds) and then a spherical dilation is employed to group the remaining bright blobs into individual active regions.
- Individual AR are tracked through time by comparing the masks of regions in two consecutive time frames, taking into account differential rotation.

SPoCA has been running in near real time on AIA data since August 2010 as part of the SDO Feature Finding Project, a suite of software pipeline modules for automated feature recognition and analysis for the imagery from SDO (Martens

⁶<http://www.helio-vo.eu/index.php>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

et al., 2011). The resulting AR events are automatically ingested by the Helio-
physics Events Knowledgebase (Hurlburt *et al.*, 2010).

Whereas SPoCA provides a consistent output of tracked ARs, two points could
benefit from further improvement: different but close ARs near the limb tend to
be detected as one huge AR, resulting in AR regions which are unrealistically
large and long-lived. Further, in a complex data set such as the one presented
here, SPoCA occasionally splits an AR in two parts and later rejoins them where
a human observer would interpret them as one and the same region throughout
its lifetime.

Comparing the two sunspot properties measured by ASAP and STARA,
sunspot area is more interesting than total number, since the latter fluctuates
wildly as transient spots flick in and out of view. Also the visibility curve is
important here since the number of small spots detected is highly impacted by
the observer's viewing angle (Dalla, Fletcher, and Walton, 2008; Watson *et al.*,
2009). The area of large spots, however, is less impacted by this effect, and
dominates the total area measurement. In general, time series of the sunspot
and AR areas detected by the four algorithms show similar trends.

The total magnetic flux is much more telling of an AR evolution than the area
or maximum magnetic field value, which is often affected by the MDI saturation
problem (Liu, Norton, and Scherrer, 2007). Saturation affects the reliability of
magnetic field measurements above 1000 Gauss due to on-board data processing.
Also, the maximum magnetic field value is unstable since different parts of the
AR will over-take each other in field magnitude as they develop causing the
location that the value is sampled from to vary wildly.

Due to the fact that SPoCA regularly merges its detected region with a
neighboring one, its time series feature a few discontinuities. Whenever SPoCA
detects a single region as with the other algorithms, however, its time series are
stable. SPoCA's total EUV intensity has a smooth behaviour over time and is
closely linked to the SPoCA area. Its maximum EUV intensity, on the other
hand, is more variable, and peaks when the active region emits a large flare.

Through the time-series analysis of two AR case studies (Section 4.3), we have
observed three physical processes evident in their evolution: emergence of the
bipolar magnetic structure, sunspots, and EUV loops (Section 4.3.1); increase
and peak in non-potentiality, followed by the onset of flaring (Section 4.3.2);
decay and weakening of magnetic footpoints (Section 4.3.2).

Emergence is observed in the simple region NOAA 10377, as evidenced by
its time series of areas, magnetic flux, total EUV intensity, R value and Ising
energy.

We follow the complex NOAA region 10365 from the moment when it is
already well established. The region shows a steady increase in size (for all al-
gorithms), magnetic flux, magnetic field value, total EUV intensity, PSL length,
 R value and Ising energy, and releases several major flares, which are clearly
visible in the maximum EUV intensity.

In our analysis of AR 10365, we find that the main PSL rotates with respect
to the bipole connection line, which is a sign of helicity injection. This may
cause the following decrease in connection line length and increase in polarity
mixing. The increase in PSL length, R , and Ising energy are likely a result of

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

this mixing. Helicity injection has been established as a method of increasing the non-potentiality in ARs. The cause of the injection may result from the emergence of subsurface twisted flux ropes, as seen in Dun *et al.* (2007).

As NOAA 10365 returns after one solar rotation, decay is seen in its magnetic footprint in total magnetic flux and total sunspot area, as well as in the maximum magnetic field, the PSL length, the R value and the Ising energy. The magnetic area and EUV signatures are not seen to decrease, however. In the photosphere, supergranular diffusion, which occurs during decay, likely keeps the area at a constant, if not increasing, value.

Coronal structures do not appear to decay as readily as their magnetic footpoints. This result agrees with Lites *et al.* (1995), in which it is reasoned that if the coronal magnetic structure is closed, it may be in a state of quasi-static equilibrium, whereby the magnetic buoyancy of the loops is cancelled by the weight of plasma trapped at the bottom of the closed structure.

Acknowledgements Funding of CV by the Belgian Federal Science Policy Office (BEL-SPO) through the ESA/PRODEX SIDC Data Exploitation program, as well as by the Solar-Terrestrial Center of Excellence/ROB is hereby appreciatively acknowledged. FTW acknowledges Ph. D. funding from the Science and Technology Facilities Council and the guidance of his supervisor, Lyndsay Fletcher. We acknowledge support from ISSI through funding for the International Team on “Mining and exploiting SDO data in Europe”. ASAP is supported by an EPSRC Grant (EP/F022948/1), which is entitled ‘Image Processing, Machine Learning and Geometric Modelling for the 3D Representation of Solar Features’. PAH acknowledges support from ESA/PRODEX and a grant from the EC Framework Programme 7 (HELIO) and the guidance of his supervisor, Peter T. Gallagher. We would like to thank the SOHO team for making both their data and analysis software publicly available and Omar W. Ahmed for allowing us the use of the Ising Energy software he was involved in developing.

References

- Abramenko, V.I., Longcope, D.W.: 2005, Distribution of the Magnetic Flux in Elements of the Magnetic Field in Active Regions. *Astrophys. J.* **619**, 1160–1166. doi:10.1086/426710.
- Ahmed, O., Qahwaji, R., Colak, T., DudokDeWit, T., Ipson, S.: 2010, A new technique for the calculation and 3d visualisation of magnetic complexities on solar satellite images. *The Visual Computer* **26**, 385–395. 10.1007/s00371-010-0418-1. <http://dx.doi.org/10.1007/s00371-010-0418-1>.
- Aschwanden, M.J.: 2010, Image Processing Techniques and Feature Recognition in Solar Physics. *Solar Phys.* **262**, 235–275. doi:10.1007/s11207-009-9474-y.
- Barra, V., Delouille, V., Kretzschmar, M., Hochedez, J.: 2009, Fast and robust segmentation of solar EUV images: algorithm and results for solar cycle 23. *Astron. Astrophys.* **505**, 361–371. doi:10.1051/0004-6361/200811416.

-
- 1
2
3
4
5
6
7 Benkhalil, A.*et al.*: 2003, Automatic Identification of Active Regions (Plages)
8 in the Full-Disk Solar Images Using Local Thresholding and Region Growing
9 Techniques. *AISB'03 Symposium on Biologically-inspired Machine Vision,*
10 *Theory and Application. 2003 University of Wales, Aberystwyth.*
- 11 Bentley, R.D., Abouadarham, J., Csillaghy, A., Jacquy, C., Hapgood, M.A.,
12 Messerotti, M., Gallagher, P., Bocchialini, K., Hurlburt, N.E., Roberts, D.,
13 Sanchez Duarte, L.: 2009, Addressing Science Use Cases with HELIO. *AGU*
14 *Fall Meeting Abstracts*, A6.
- 15
16 Colak, T., Ahmed, O.W., Qahwaji, R., Higgins, P.A.: 2010a, Automated So-
17 lar Flare Prediction: Is it a Myth? *Presentation in Seventh European Space*
18 *Weather Week.*
- 19
20 Colak, T., Qahwaji, R., Ipson, S., Ugail, H.: 2010b, Representation of Solar
21 Features in 3D for Creating Visual Solar Catalogues. *Advances in Space*
22 *Research.*
- 23
24 Colak, T., Qahwaji, R.: 2008, Automated McIntosh-Based Classification
25 of Sunspot Groups Using MDI Images. *Solar Phys.* **248**, 277–296.
26 doi:10.1007/s11207-007-9094-3.
- 27
28 Colak, T., Qahwaji, R.: 2009, Automated Solar Activity Prediction: A hybrid
29 computer platform using machine learning and solar imaging for automated
30 prediction of solar flares . *Space Weather*, Vol. 7, S06001.
- 31
32 Conlon, P.A., Gallagher, P.T., McAteer, R.T.J., Ireland, J., Young, C.A.,
33 Kestener, P., Hewett, R.J., Maguire, K.: 2008, Multifractal Properties of
34 Evolving Active Regions. *Solar Phys.* **248**, 297–309. doi:10.1007/s11207-007-
35 9074-7.
- 36
37 Conlon, P.A., McAteer, R.T.J., Gallagher, P.T., Fennell, L.: 2010, Quantifying
38 the Evolving Magnetic Structure of Active Regions. *Astrophys. J.* **722**, 577–
39 585. doi:10.1088/0004-637X/722/1/577.
- 40
41 Curto, J.J., Blanca, M., Martínez, E.: 2008, Automatic Sunspots Detection on
42 Full-Disk Solar Images using Mathematical Morphology. *Solar Phys.* **250**,
43 411–429. doi:10.1007/s11207-008-9224-6.
- 44
45 Dalla, S., Fletcher, L., Walton, N.A.: 2008, Invisible sunspots and rate of solar
46 magnetic flux emergence. *Astron. Astrophys.* **479**, L1–L4. doi:10.1051/0004-
47 6361:20078800.
- 48
49 Delaboudinière, J., Artzner, G.E., Brunaud, J., Gabriel, A.H., Hochedez, J.F.,
50 Millier, F., Song, X.Y., Au, B., Dere, K.P., Howard, R.A., Kreplin, R., Michels,
51 D.J., Moses, J.D., Defise, J.M., Jamar, C., Rochus, P., Chauvineau, J.P.,
52 Marioge, J.P., Catura, R.C., Lemen, J.R., Shing, L., Stern, R.A., Gurman,
53 J.B., Neupert, W.M., Maucherat, A., Clette, F., Cugnon, P., van Dessel, E.L.:
54 1995, EIT: Extreme-Ultraviolet Imaging Telescope for the SOHO Mission.
55 *Solar Phys.* **162**, 291–312. doi:10.1007/BF00733432.
- 56
57
58
59
60
61
62
63
64
65

-
- 1
2
3
4
5
6
7 Dougherty, E.R., Lotufo, R.A.: 2003, *Hands-on morphological image processing*,
8 SPIE Optical Engineering Press, Washington, 130.
- 9
10 Dun, J., Kurokawa, H., Ishii, T.T., Liu, Y., Zhang, H.: 2007, Evolution of
11 Magnetic Nonpotentiality in NOAA AR 10486. *Astrophys. J.* **657**, 577–591.
12 doi:10.1086/510373.
- 13
14 Falconer, D.A., Moore, R.L., Gary, G.A.: 2008, Magnetogram Measures of Total
15 Nonpotentiality for Prediction of Solar Coronal Mass Ejections from Active
16 Regions of Any Degree of Magnetic Complexity. *Astrophys. J.* **689**, 1433–
17 1442. doi:10.1086/591045.
- 18
19 Fisher, G.H., Longcope, D.W., Metcalf, T.R., Pevtsov, A.A.: 1998, Coronal Heat-
20 ing in Active Regions as a Function of Global Magnetic Variables. *Astrophys.*
21 *J.* **508**, 885–898. doi:10.1086/306435.
- 22
23 Gallagher, P.T., Moon, Y., Wang, H.: 2002, Active-Region Monitoring and Flare
24 Forecasting I. Data Processing and First Results. *Solar Phys.* **209**, 171–183.
25 doi:10.1023/A:1020950221179.
- 26
27 Gao, J., Wang, H., Zhou, M.: 2002, Development of an Automatic Filament
28 Disappearance Detection System. *Solar Phys.* **205**, 93–103.
- 29
30 Georgoulis, M.K., Rust, D.M.: 2007, Quantitative Forecasting of Major Solar
31 Flares. *Astrophys. J. Letters* **661**, L109–L112. doi:10.1086/518718.
- 32
33 Hale, G.E., Ellerman, F., Nicholson, S.B., Joy, A.H.: 1919, No. 165. The magnetic
34 polarity of Sun-spots. *Contributions from the Mount Wilson Observatory /*
35 *Carnegie Institution of Washington* **165**, 1–26.
- 36
37 Handy, B.N., Schrijver, C.J.: 2001, On the Evolution of the Solar Pho-
38 tospheric and Coronal Magnetic Field. *Astrophys. J.* **547**, 1100–1108.
39 doi:10.1086/318429.
- 40
41 Hewett, R.J., Gallagher, P.T., McAteer, R.T.J., Young, C.A., Ireland, J., Conlon,
42 P.A., Maguire, K.: 2008, Multiscale Analysis of Active Region Evolution. *Solar*
43 *Phys.* **248**, 311–322. doi:10.1007/s11207-007-9028-0.
- 44
45 Higgins, P.A., Gallagher, P.T., McAteer, R.T.J., Bloomfield, D.S.: 2010, Solar
46 Magnetic Feature Detection and Tracking for Space Weather Monitoring.
47 *ArXiv e-prints*.
- 48
49 Howard, R.F., Harvey, J.W., Forgach, S.: 1990, Solar surface velocity fields
50 determined from small magnetic features. *Solar Phys.* **130**, 295–311.
51 doi:10.1007/BF00156795.
- 52
53 Hurlburt, N., Cheung, M., Schrijver, C., Chang, L., Freeland, S., Green, S.,
54 Heck, C., Jaffey, A., Kobashi, A., Schiff, D., Serafin, J., Seguin, R., Slater,
55 G., Somani, A., Timmons, R.: 2010, Heliophysics Event Knowledgebase
56 for the Solar Dynamics Observatory (SDO) and Beyond. *Solar Phys.*, 195.
57 doi:10.1007/s11207-010-9624-2.
- 58
59
60
61
62
63
64
65

-
- 1
2
3
4
5
6
7 Jolliffe, I.T.: 2002, *Principal component analysis, 2nd edition*, Springer, New
8 York, 487.
- 9
10 LaBonte, B.J., Georgoulis, M.K., Rust, D.M.: 2007, Survey of Magnetic Helicity
11 Injection in Regions Producing X-Class Flares. *Astrophys. J.* **671**, 955–963.
12 doi:10.1086/522682.
- 13
14 Lefebvre, S., Rozelot, J.: 2004, A new method to detect active features at the so-
15 lar limb. *Solar Phys.* **219**, 25–37. doi:10.1023/B:SOLA.0000021818.97402.1e.
- 16
17 Leka, K.D., Barnes, G.: 2007, Photospheric Magnetic Field Properties of Flar-
18 ing versus Flare-quiet Active Regions. IV. A Statistically Significant Sample.
19 *Astrophys. J.* **656**, 1173–1186. doi:10.1086/510282.
- 20
21 Lites, B.W., Low, B.C., Martinez Pillet, V., Seagraves, P., Skumanich, A., Frank,
22 Z.A., Shine, R.A., Tsuneta, S.: 1995, The Possible Ascent of a Closed Magnetic
23 System through the Photosphere. *Astrophys. J.* **446**, 877. doi:10.1086/175845.
- 24
25 Liu, Y., Kurokawa, H.: 2004, On a Surge: Properties of an Emerging Flux Region.
26 *Astrophys. J.* **610**, 1136–1147. doi:10.1086/421715.
- 27
28 Liu, Y., Norton, A.A., Scherrer, P.H.: 2007, A Note on Saturation Seen in the
29 MDI/SOHO Magnetograms. *Solar Phys.* **241**, 185–193. doi:10.1007/s11207-
30 007-0296-5.
- 31
32 Martens, P.C.H., Attrill, G.D.R., Davey, A.R., Engell, A., Farid, S., Grigis,
33 P.C., Kasper, J., Korreck, K., Saar, S.H., Savcheva, A., Su, Y., Testa, P.,
34 Wills-Davey, M., Bernasconi, P.N., Raouafi, N., Delouille, V.A., Hochedez,
35 J.F., Cirtain, J.W., Deforest, C.E., Angryk, R.A., de Moortel, I., Wiegelm-
36 mann, T., Georgoulis, M.K., McAteer, R.T.J., Timmons, R.P.: 2011, Com-
37 puter Vision for the Solar Dynamics Observatory (SDO). *Solar Phys.*, 8.
38 doi:10.1007/s11207-010-9697-y.
- 39
40 Morita, S., McIntosh, S.W.: 2005, Genesis of AR NOAA10314. In: K. Sankara-
41 subramanian, M. Penn, & A. Pevtsov (ed.) *Large-scale Structures and their*
42 *Role in Solar Activity, Astronomical Society of the Pacific Conference Series*
43 **346**, 317.
- 44
45 Nguyen, S.H., Nguyen, T.T., Nguyen, H.S.: 2005, Rough set approach to sunspot
46 classification problem. In: Slezak, D., Yao, J., Peters, J.F., Ziarko, W., Hu, X.
47 (eds.) *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing Lecture*
48 *Notes in Computer Science* **3642**, Springer, Berlin / Heidelberg, 263–272.
- 49
50 Pérez-Suárez, D., Higgins, P.A., McAteer, R.T.J., Bloomfield, D.S., Gallagher,
51 P.T.: 2011, Solar Feature Detection. *Applied Signal and Image Processing:*
52 *Multidisciplinary Advancements*.
- 53
54 Qahwaji, R., Colak, T.: 2006, Hybrid imaging and neural networks techniques
55 for processing solar images. *I. J. Comput. Appl.* **13**(1), 9–16.
- 56
57
58
59
60
61
62
63
64
65

-
- 1
2
3
4
5
6
7 Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G.,
8 Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., Wolf-
9 son, C.J., Zayer, I., MDI Engineering Team: 1995, The Solar Oscilla-
10 tions Investigation - Michelson Doppler Imager. *Solar Phys.* **162**, 129–188.
11 doi:10.1007/BF00733429.
- 12 Schrijver, C.J.: 1987, Solar active regions - Radiative intensities and large-scale
13 parameters of the magnetic field. *Astron. Astrophys.* **180**, 241–252.
- 14
15 Schrijver, C.J.: 2007, A Characteristic Magnetic Field Pattern Associated with
16 All Major Solar Flares and Its Use in Flare Forecasting. *Astrophys. J. Letters*
17 **655**, L117–L120. doi:10.1086/511857.
- 18
19 Shih, F.Y., Kowalski, A.J.: 2003, Automatic Extraction of Filaments in H α Solar
20 Images. *Solar Phys.* **218**, 99–122. doi:10.1023/B:SOLA.0000013052.34180.58.
- 21
22 SIDC-Team: 2003, The International Sunspot Number. *Monthly Report on the*
23 *International Sunspot Number, online catalogue.*
- 24
25 Subramanian, P., Dere, K.P.: 2001, Source Regions of Coronal Mass Ejections.
26 *Astrophys. J.* **561**, 372–395. doi:10.1086/323213.
- 27
28 Watson, F., Fletcher, L., Dalla, S., Marshall, S.: 2009, Modelling the Longitu-
29 dinal Asymmetry in Sunspot Emergence: The Role of the Wilson Depression.
30 *Solar Phys.* **260**, 5–19. doi:10.1007/s11207-009-9420-z.
- 31
32 Welsch, B.T., Longcope, D.W.: 2003, Magnetic Helicity Injection by Horizontal
33 Flows in the Quiet Sun. I. Mutual-Helicity Flux. *Astrophys. J.* **588**, 620–629.
34 doi:10.1086/368408.
- 35
36 Zharkov, S., Zharkova, V., Ipson, S., Benkhalil, A.: 2004, Automated recognition
37 of sunspots on the soho/mdi white light solar images. In: *Knowledge-Based*
38 *Intelligent Information and Engineering Systems Lecture Notes in Computer*
39 *Science* **3215**, Springer, Berlin / Heidelberg, 446–452.
- 40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65