

1 **The *Helioseismic and Magnetic Imager* (HMI) Vector Magnetic Field**
2 **Pipeline: SHARPs – Space-weather HMI Active Region Patches**

3 M.G. Bobra¹ · X. Sun¹ · J.T. Hoeksema¹ · M. Turmon² ·
4 Y. Liu¹ · K. Hayashi¹ · G. Barnes³ · K.D. Leka³

5 © Springer ●●●

6 **Abstract** A new data product from the *Helioseismic and Magnetic Imager* (HMI) onboard the
7 *Solar Dynamics Observatory* (SDO) called Space-weather HMI Active Region Patches (SHARPs)
8 is now available. SDO/HMI is the first space-based instrument to map the full-disk photospheric
9 vector magnetic field with high cadence and continuity. The SHARP data series provide maps in
10 patches that encompass automatically tracked magnetic concentrations for their entire lifetime;
11 map quantities include the photospheric vector magnetic field and its uncertainty, along with
12 Doppler velocity, continuum intensity, and line-of-sight magnetic field. Furthermore, keywords in
13 the SHARP data series provide several parameters that concisely characterize the magnetic-field
14 distribution and its deviation from a potential-field configuration. These indices may be useful for
15 active-region event forecasting and for identifying regions of interest. The indices are calculated
16 per patch and are available on a twelve-minute cadence. Quick-look data are available within
17 approximately three hours of observation; definitive science products are produced approximately
18 five weeks later. SHARP data are available at jsoc.stanford.edu and maps are available in either of
19 two different coordinate systems. This article describes the SHARP data products and presents
20 examples of SHARP data and parameters.

21 **Keywords:** Active Regions, Magnetic Fields; Flares, Relation to Magnetic Field; Instrumentation
22 and Data Management

23 **1. Introduction**

24 This article describes a data product from the *Solar Dynamics Observatory's Helioseismic and*
25 *Magnetic Imager* (SDO/HMI) called Space-weather HMI Active Region Patches (SHARPs).
26 SHARPs follow each significant patch of solar magnetic field from before the time it appears
27 until after it disappears. The SHARP data series presently include 16 indices computed from
28 the vector magnetic field in active-region patches. These parameters, many of which have been
29 associated with enhanced flare productivity, are automatically calculated for each solar active
30 region using HMI vector magnetic-field data with a 12-minute cadence. The indices and other
31 keywords can be used to select regions and time intervals for further study. The active-region
32 patches are automatically identified and tracked for their entire lifetime (Turmon *et al.*, 2014).
33 In addition to the indices, the four SHARP data series include the photospheric vector magnetic-
34 field data for the patches, as well as co-registered maps of Doppler velocity, continuum intensity,
35 line-of-sight magnetic field, and other quantities.

36 Measurements of the photospheric magnetic field provide insight into understanding and
37 possibly predicting eruptive phenomena in the solar atmosphere, such as flares and coronal
38 mass ejections. For example, it is generally accepted that large, complex, and rapidly evolving
39 photospheric active regions are the most likely to produce eruptive events (Zirin, 1988; Priest,
40 1984). As such, it is an active area of research to seek a correlation (or its rejection) between
41 eruptive events and quantitative parameterizations of the photospheric magnetic field. Many
42 studies have found a relationship between solar-flare productivity and various indices: magnetic
43 helicity (*e.g.* Tian *et al.*, 2005; Török and Kliem, 2005; LaBonte, Georgoulis, and Rust, 2007), free
44 energy proxies (*e.g.* Moore, Falconer, and Sterling, 2012), magnetic shear angle (*e.g.* Hagyard

¹ W.W. Hansen Experimental Physics Laboratory, Stanford University,
Stanford, CA, USA

email: jthoeksema@sun.stanford.edu

² Jet Propulsion Laboratory, Pasadena, CA, USA

³ Northwest Research Associates, Inc., Boulder, CO, USA

45 *et al.*, 1984; Leka and Barnes, 2003a, 2003b, 2007), magnetic topology (*e.g.* Cui *et al.*, 2006;
 46 Barnes and Leka, 2006, Georgoulis and Rust, 2007), or the properties of active-region polarity
 47 inversion lines (*e.g.* Mason and Hoeksema, 2010; Falconer, Moore, and Gary, 2008; Schrijver,
 48 2007). However, when Leka and Barnes (2003a) conducted a discriminant analysis of over a
 49 hundred parameters calculated from vector magnetic-field measurements of seven active regions,
 50 they could identify “no single, or even small number of, physical properties of an active region
 51 that is sufficient and necessary to produce a flare.” Larger statistical samples show correlations
 52 between some vector-field non-potentiality parameters and overall flare productivity (Leka and
 53 Barnes, 2007; Yang *et al.*, 2012), as well as correlations between the parameters themselves. Still,
 54 characteristics have yet to be identified that uniquely distinguish imminent flaring in an active
 55 region.

56 The SHARP data series will provide a complete record of all visible solar active regions
 57 since 1 May 2010. SHARP data are stored in a database and readily accessible at the Joint
 58 Science Operations Center (JSOC). JSOC data products from SDO, as well as source code for
 59 the modules, can be found at jsoc.stanford.edu. Continuously updated plots of near-real-time
 60 parameters are available online (see Table 2 for URLs). We describe how the SHARP series are
 61 created and show results for two representative active regions. We also present examples of four
 62 active-region parameters for 12 X-, M-, and C-class flaring active regions.

63 2. Methodology: SHARP Data and Active Region Parameters

64 Data taken onboard SDO/HMI are downlinked to the ground, automatically processed through
 65 the HMI data pipeline, and made available at jsoc.stanford.edu organized in data series (Schou
 66 *et al.*, 2012a; Scherrer *et al.*, 2012). Conceptually, a JSOC data series consists of a sequence
 67 of *records*, each of which includes i) a table of keywords and ii) associated data arrays, called
 68 *segments*. A record exists for each time step or unique set of prime keyword(s). Keywords and
 69 data array segments are merged by the JSOC into FITS files in response to a user’s request to
 70 download (or export) the data series. SHARP data for export can be selected by time, given in
 71 the keyword T_REC, and the region number, HARNUM; additionally, requests for data from the
 72 JSOC can also take advantage of simple SQL database queries on keywords to select data of
 73 interest. A complete overview of JSOC Data Series is available on the JSOC wiki (see Table 2).

74 Certain HMI data series are processed on two time scales: near real time (NRT) and definitive.
 75 NRT data are processed quickly, ordinarily within three hours of the observation time, but with
 76 preliminary calibrations. Section 7 describes the differences between definitive and NRT HARPs.
 77 Although most NRT data series are not archived and go offline after approximately three months,
 78 the NRT SHARP data since 14 September 2012 are archived. NRT data are primarily intended for
 79 quick-look monitoring or as a forecasting tool. This section briefly describes the elements of the
 80 HMI data pipeline necessary to create the definitive SHARP data. A more detailed explanation
 81 of the HMI vector magnetic-field pipeline processing is given by Hoeksema *et al.* (2014) and
 82 references therein.

- 83 • In each 135-second interval HMI samples six points across the Fe I 6173.3 Å spectral line
 84 and measures six polarization states: $I \pm Q$, $I \pm U$, and $I \pm V$, generating $36 \times 4096 \times 4096$
 85 full-disk filtergrams.
- 86 • To reduce noise and minimize the effects of solar oscillations, a tapered temporal average
 87 is performed every 720 seconds using 360 filtergrams collected over a 1350-second interval
 88 to produce 36 corrected, filtered, co-registered images (Couvidat *et al.*, 2012).
- 89 • A polarization calibration is applied and the four Stokes polarization states, $[IQUV]$, are
 90 determined at each wavelength, giving a total of 24 images at each time step (Schou *et al.*,
 91 2012b) that are available in the data series `hmi.S_720s`.
- 92 • Active-region patches are automatically detected and tracked in the photospheric line-
 93 of-sight magnetograms (Turmon *et al.*, 2014). The detection algorithm identifies both a
 94 rectangular bounding box on the CCD image that encompasses the entire region and,
 95 within this box, creates a bitmap that both encodes membership in the coherent magnetic
 96 structure and indicates strong-field pixels. Specifically, the bitmap array assigns a value to

Table 1. Listed below are URLs relevant for finding the SHARP data, codes, documentation, and data visualizations. These URLs will be maintained for at least the duration of the SDO mission.

Uniform Resource Locator	Description
jsoc.stanford.edu/data/hmi/sharp/dataviewer	Continuously updated plots of near-real-time SHARP parameters.
jsoc.stanford.edu/doc/data/hmi/sharp/sharp.htm	Description of the SHARP data product.
jsoc.stanford.edu/jsocwiki/DataSeries	A complete overview of the Joint Science and Operations Center (JSOC) Data Series.
jsoc.stanford.edu/jsocwiki/PipelineCode	Guide to HMI pipeline code and processing notes.
jsoc.stanford.edu/jsocwiki/Lev1qualBits	Description of bits in QUALITY keyword.
jsoc.stanford.edu/cvs/JSOC/proj/sharp/apps/sharp.c ../cvs/JSOC/proj/sharp/apps/sw_functions.c	The SHARP data are created via this publicly available C module (<code>sharp.c</code>) that includes a library of active region parameter calculations (<code>sw_functions.c</code>).
jsoc.stanford.edu/jsocwiki/sharp_coord	A technical note on SHARP coordinate systems, mapping, and vector transformations (Sun, 2013).
jsoc.stanford.edu/jsocwiki/HARPDaDataSeries	Description of the HARP data series (Turmon <i>et al.</i> , 2014)
hmi.stanford.edu/magnetic	Portal to HMI magnetic field data, image catalogs, coverage maps, and documentation
www.lmsal.com/sdouserguide.html	Comprehensive guide to SDO Data Analysis

- 97 each pixel inside the bounding box, depending on whether it i) resides inside or outside the
 98 active region, and ii) corresponds to weak or strong line-of-sight magnetic field. This coding
 99 scheme permits non-contiguous active-region patches.
- 100 • The tracking module numbers each HMI Active Region Patch (HARP) and generates a
 101 time series of bitmaps large enough to contain the maximum known heliographic extent of
 102 the region. Each numbered HARP (keyword HARPNUM) corresponds to one active region or
 103 AR complex (see Figure 1). The HARP database generally captures more patches of solar
 104 magnetic activity than the NOAA active region database because coherent regions that are
 105 small in extent or have no associated photometric sunspot are detected and tracked by our
 106 code; such faint HARPs often have no NOAA correspondence. A HARP may include zero,
 107 one, or multiple NOAA active regions (for example, see HARP 2360 in Figure 1); about
 108 one-third of HARPs correspond to a single NOAA region. The bitmap array described
 109 above is in the BITMAP segment of the data series `hmi.Mharp_720s`.
 - 110 • The full-disk Stokes data are inverted using the Very Fast Inversion of the Stokes Vector
 111 (VFISV) code, which assumes a Milne–Eddington model of the solar atmosphere, to yield
 112 vector magnetic field data (Borrero *et al.*, 2011; Centeno *et al.*, 2014). Inverted data are
 113 available in the data series `hmi.ME_720s_fd10`. Full disk inversions are being computed for
 114 all HMI data since 1 May 2010. An improvement made to the inversion code in May 2013
 115 (Hoeksema *et al.*, 2014) to use time-dependent information about the HMI filter profiles
 116 introduces measurable systematic differences in inversion results. Data in the interval from
 117 1 August 2012–24 May 2013 were processed before the improvement. Some care must be
 118 taken when comparing data computed with different versions of the analysis code (see the
 119 entry under PipelineCode referenced in Table 2).
 - 120 • The azimuthal component of the vector magnetic field is disambiguated using the Minimum
 121 Energy Code (ME0) to resolve the 180° ambiguity (Metcalfe, 1994; Leka *et al.*, 2009).
 122 Through 14 January 2014 SHARP regions have been disambiguated individually using `fd10`
 123 data inside a rectangle that extends beyond the HARP bounding box by the number of
 124 pixels given in the AMBNPAD keyword. Disambiguation results for each HARPNUM at each
 125 time step are stored in the DISAMBIG segment of the `hmi.Bharp_720s` data series. All pixels
 126 inside the rectangular bounding box are annealed in the patchwise SHARP disambiguation;
 127 however, pixels below a noise threshold are also smoothed (Barnes *et al.*, 2014; Hoeksema

SDO/HMI Tracked AR (HARP)
2013/01/13
00:48

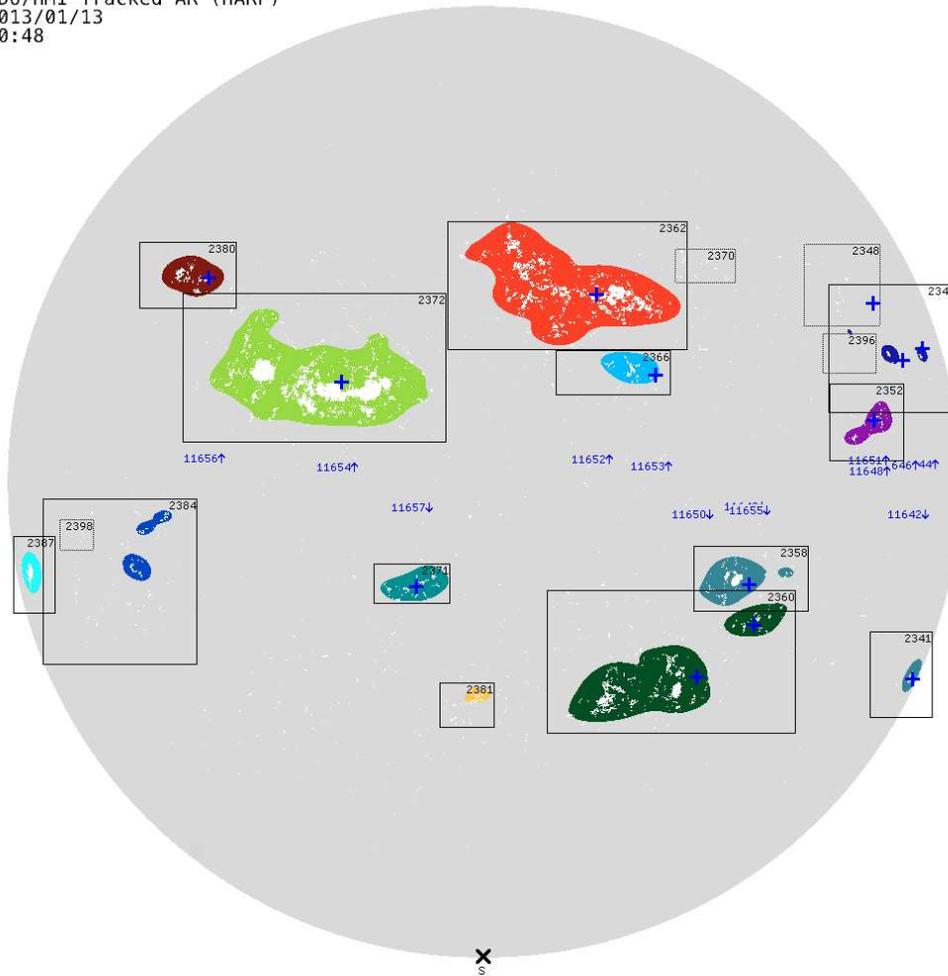


Figure 1. The results of the active-region automatic detection algorithm applied to the data on 13 January 2013 at 00:48 TAI. NOAA active-region numbers are labeled in blue near the Equator, next to arrows indicating the hemisphere; the HARP number is indicated inside the rectangular bounding box at the upper right. Note that HARP 2360 (lower right, in green) includes two NOAA active regions, 11650 and 11655. The colored patches show coherent magnetic structures that comprise the HARP. White pixels have a line-of-sight field strength above a line-of-sight magnetic-field threshold (Turmon *et al.*, 2014). Blue “+” symbols indicate coordinates that correspond to the reported center of a NOAA active region. The temporal life of a definitive HARP starts when it rotates onto the visible disk or two days before the magnetic feature is first identified in the photosphere. As such, empty boxes, *e.g.* HARP 2398 (on the left), represent patches of photosphere that will contain a coherent magnetic structure at a future time.

128 *et al.*, 2014). Since 19 December 2013 we have disambiguated the entire disk and use that
129 data from the consistently derived DISAMBIG segment of the hmi.B_720s data series for
130 definitive SHARPs observed from 15 January 2014 onward.

- 131 • Finally, to complete the SHARP data series the analysis pipeline collects maps of HMI
132 observables and computes a set of active region summary parameters using a publicly
133 available module (see Table 2 and Section 4).

134 3. SHARP Coordinates: CCD Cutouts and Cylindrical Equal Area Maps

135 HMI data series use standard World Coordinate System (WCS) for solar images (Thompson,
136 2006). SHARP data series are available in either of two coordinate systems: one is effectively cut
137 out directly from corrected full-disk images, which are in helio-projective Cartesian CCD image
138 coordinates, and the other is remapped from CCD coordinates to a heliographic Cylindrical

Table 2. Listed below are four series that contain SHARP data. SHARP active region parameters are stored as keywords for these series. For a list of parameters, see Table 3.

Data Series Name	Description
hmi.sharp_720s	Definitive data with 31 map segments in CCD coordinates wherein the vector \mathbf{B} is comprised of azimuth, inclination, and field strength.
hmi.sharp_cea_720s	Definitive data with 11 segments wherein all quantities have been remapped to a heliographic Cylindrical Equal-Area coordinate system centered on the patch and the vector \mathbf{B} has been transformed into the components B_r , B_θ , and B_ϕ .
hmi.sharp_720s_nrt	Near-real-time data; otherwise same as hmi.sharp_720s.
hmi.sharp_cea_720s_nrt	Near-real-time data; otherwise same as hmi.sharp_cea_720s.

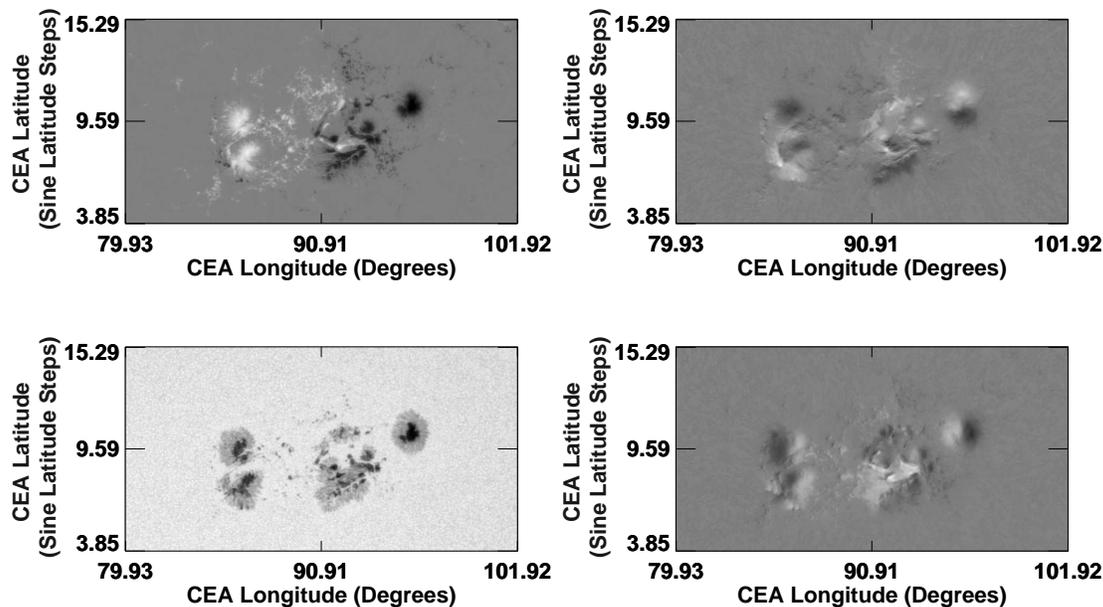


Figure 2. The first three panels, clockwise from upper left, show the inverted and disambiguated data wherein the vector \mathbf{B} has been remapped to a Cylindrical Equal-Area projection and decomposed into B_r , B_θ , and B_ϕ , respectively, for HARP 401 (NOAA AR 11166) on 9 March 2011 at 23:24:00 TAI. The color table is scaled between ± 2500 Gauss for all three magnetic-field arrays. The lower-left panel shows the computed continuum intensity for the same region at the same time. The patch is centered at longitude 90.91° , latitude 9.59° in Carrington Rotation 2107. CEA longitude and latitude are described in the text.

139 Equal-Area (CEA) projection centered on the patch. Table 2 lists the four available SHARP
140 data series.

141 For standard CCD-cutout SHARPs, the pipeline module collects 31 maps, including many of
142 the primary HMI observable data segments (line-of-sight magnetogram, Dopplergram, continuum
143 intensity, and vector magnetogram), other inversion and disambiguation quantities, uncertainty
144 arrays, and the HARP bitmap. Using the HARP bounding box as a stencil, the module extracts
145 the corresponding arrays of observable data. The first six tables in Appendix A give a description
146 of each of the cut-out SHARP series segment maps.

147 Additional processing is applied to the CEA versions of the SHARPs to convert selected
148 segments from CCD pixels in plane-of-the-sky coordinates to a heliographic coordinate system
149 in the photosphere. Table A.7 in Appendix A lists the 11 segment maps that are available in
150 CEA coordinates.

151 The expression relating the final CEA map coordinate $[x, y]$ to the heliographic longitude
152 and latitude $[\phi, \lambda]$ follows Equations (79) and (80) of Calabretta and Greisen (2002), compliant
153 with the World Coordinate System (WCS) standard (*e.g.* Thompson, 2006). The remapping
154 uses the patch center as reference point, thus effectively de-rotating the patch center to $\phi = 0$,

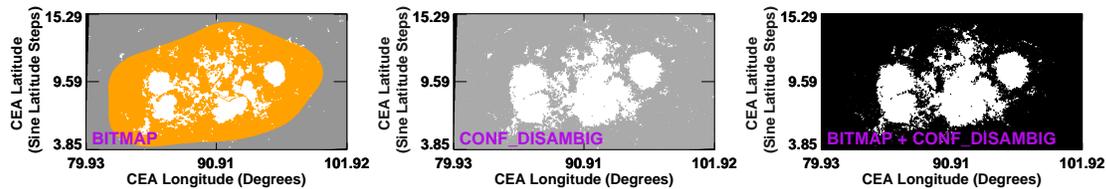


Figure 3. Only pixels that are both within the HARP (shaded orange in map segment BITMAP, left) and above the high-confidence disambiguation threshold (shown in white in the middle panel where segment CONF_DISAMBIG = 90) contribute to the active region parameters (represented in the rightmost panel). This example from hmi.sharp_720s_cea shows HARP 401 (NOAA AR 11166) on 9 March 2011 at 23:24:00 TAI, where the quantities have been remapped to a Cylindrical Equal-Area coordinate system. Black areas at the edge of the BITMAP and CONF_DISAMBIG images fall outside the maximal CCD HARP bounding box; therefore, the azimuthal ambiguity resolution has not been applied to these areas. As in Figure 2, the axes are labeled in CEA coordinates, as described in the text.

155 $\lambda = 0$ before CEA projection in order to minimize distortion (see Section 2.5 of Calabretta and
 156 Greisen, 2002). As a consequence, the correspondence between what are labeled CEA degrees
 157 and the familiar Carrington latitude and longitude is complex. The Carrington coordinates of
 158 the patch center are indicated in the keywords CRVAL1 and CRVAL2. The SHARP CEA pixels
 159 have a linear dimension in the x -direction of 0.03 heliographic degrees in the rotated coordinate
 160 system and an area on the photosphere of $1.33 \times 10^5 \text{ km}^2$. The size in the y -direction is defined
 161 by the CEA requirement that the area of each pixel be the same, so the pixels are equally spaced
 162 in the sine of the angular distance from the great circle that defines the x -axis and the step size
 163 is fixed such that the pixel dimension is equal to 0.03 degrees at patch center. In Figures 2 and
 164 3 the axes are labeled in CEA degrees with the center point having the Carrington longitude
 165 and latitude values. In our remapping process the CEA grid is oversampled by interpolating the
 166 nearby CCD values and then smoothed with a Gaussian filter to the final sampling. Details are
 167 provided by Sun (2013).

168 The remapping of the uncertainty images, as well as the BITMAP and CONF_DISAMBIG maps,
 169 is done a little differently. For these the center of each pixel in the remapped CEA coordinate
 170 system is first located in the original CCD image; then the nearest neighboring pixel in the
 171 original image is identified and the value for that nearest original CCD pixel is reported.

172 For the CEA version, the native three-component vector magnetic-field output from the
 173 inversion – expressed as field strength $[B]$, inclination $[\gamma]$, and azimuth $[\psi]$ in the image plane –
 174 is transformed to the components B_r , B_θ , and B_ϕ in standard heliographic spherical coordinates
 175 $[\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi]$ following Equation (1) of Gary and Hagyard (1990). Figure 2 shows the three
 176 components of the vector magnetic field and the computed continuum intensity for HARP 401
 177 on 9 March 2011 at 23:24 TAI in CEA coordinates. We note that because $(\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi)$ is a
 178 spherical coordinate system with the rotation axis at the pole $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$ is a planar cylindrical
 179 equal-area coordinate system centered on the patch, the unit vectors $(\hat{e}_\theta, \hat{e}_\phi)$ do not precisely
 180 align with (\hat{e}_x, \hat{e}_y) except at the center of the patch. In general, only along the y -axis passing
 181 through patch center do \hat{e}_ϕ and \hat{e}_y align. See Figure 2 of Calabretta and Greisen (2002) for an
 182 illustrative example. For more information on SHARP coordinate systems, mapping, and vector
 183 transformations, see Sun (2013).

184 4. SHARP Summary Parameters

185 The SHARP module calculates summary parameters every twelve minutes on the inverted and
 186 disambiguated data using the vector field and other γ quantities in the CEA projection. The
 187 SHARP series presently contain sixteen summary parameters, as detailed in Table 3. This initial
 188 list parametrizes some of the features of solar active regions that have been associated with
 189 enhanced flare productivity (*e.g.* Leka and Barnes, 2003a, 2007, and references therein) and
 190 includes different kinds of indices such as the total magnetic flux, the spatial gradients of the
 191 field, the characteristics of the vertical current density, current helicity, and a proxy for the

Table 3. Active-region parameters are stored as keywords in each SHARP series. This table lists each active-region parameter keyword with a brief description and formula. The keyword for the error associated with each parameter is given in the last column. Each parameter represents either a mean, sum, or integral of the distribution in the high-confidence part of the HARP; this is indicated in the Statistic column. The active-region parameters were generally adapted from Leka and Barnes (2003b) except as noted in the text. WCS-standard keywords such as CDELTA1, RSUN_OBS, and RSUN_REF, as well as fundamental constants, were used to convert to the units specified in the eponymous column. Calculations are performed on the CMASK high-confidence pixels in the CEA SHARP. Derivations of the errors can be found at the SHARP web page (see Table 2). Further description of the parameters can be found in Section 5.

Keyword	Description	Unit ¹	Formula ²	Statistic	Error Keyword
USFLUX	Total unsigned flux	Mx	$\Phi = \sum B_z dA$	Integral	ERRVF
MEANGAM	Mean angle of field from radial	Degree	$\bar{\gamma} = \frac{1}{N} \sum \arctan\left(\frac{B_h}{B_z}\right)$	Mean	ERRGAM
MEANGBT	Horizontal gradient of total field	G Mm ⁻¹	$ \nabla B_{tot} = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2}$	Mean	ERRBT
MEANGBZ	Horizontal gradient of vertical field	G Mm ⁻¹	$ \nabla B_z = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B_z}{\partial x}\right)^2 + \left(\frac{\partial B_z}{\partial y}\right)^2}$	Mean	ERRBZ
MEANGBH	Horizontal gradient of horizontal field	G Mm ⁻¹	$ \nabla B_h = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B_h}{\partial x}\right)^2 + \left(\frac{\partial B_h}{\partial y}\right)^2}$	Mean	ERRBH
MEANJZD	Vertical current density	mA m ⁻²	$\bar{J}_z \propto \frac{1}{N} \sum \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)$	Mean	ERRJZ
TOTUSJZ	Total unsigned vertical current	A	$J_{z_{total}} = \sum J_z dA$	Integral	ERRUSI
MEANALP	Characteristic twist parameter, α	Mm ⁻¹	$\alpha_{total} \propto \frac{\sum J_z \cdot B_z}{\sum B_z^2}$	Mean	ERRALP
MEANJZH	Current helicity (B_z contribution)	G ² m ⁻¹	$\bar{H}_c \propto \frac{1}{N} \sum B_z \cdot J_z$	Mean	ERRMIH
TOTUSJH	Total unsigned current helicity	G ² m ⁻¹	$H_{c_{total}} \propto \sum B_z \cdot J_z $	Sum	ERRTUI
ABSNJZH	Absolute value of the net current helicity	G ² m ⁻¹	$H_{c_{abs}} \propto \left \sum B_z \cdot J_z \right $	Sum	ERRTAI
SAVNCPP	Sum of the modulus of the net current per polarity	A	$J_{sum} \propto \left \sum_{B_z^+} J_z dA \right + \left \sum_{B_z^-} J_z dA \right $	Integral	ERRJHT
MEANPOT	Proxy for mean photospheric excess magnetic energy density	erg cm ⁻³	$\bar{\rho} \propto \frac{1}{N} \sum \left(\vec{B}^{Obs} - \vec{B}^{Pot} \right)^2$	Mean	ERRMPOT
TOTPOT	Proxy for total photospheric magnetic free energy density	erg cm ⁻¹	$\rho_{tot} \propto \sum \left(\vec{B}^{Obs} - \vec{B}^{Pot} \right)^2 dA$	Integral	ERRTPOT
MEANSHR	Shear angle	Degree	$\bar{\Gamma} = \frac{1}{N} \sum \arccos\left(\frac{\vec{B}^{Obs} \cdot \vec{B}^{Pot}}{ \vec{B}^{Obs} \vec{B}^{Pot} }\right)$	Mean	ERRMSHA
SHRGT45	Fractional of Area with Shear > 45°		Area with Shear > 45° / HARP Area	Fraction	

¹The HMI vector-magnetogram data are in units of Mx cm⁻², whereas the active-region parameters use units of Gauss. Currently, the filling factor is set to unity, so the two units have the same meaning.

²Constant terms are not shown.

192 integrated free magnetic energy. Until now, indices based on vector-field values have not been
 193 available with the coverage, cadence, and continuity afforded by HMI. With previously available
 194 data, none of the parameters were found to be necessary or sufficient to forecast a flaring event
 195 (Leka and Barnes, 2007). As of this writing, the SHARP indices focus on low-order statistical
 196 moments of observables and readily derived quantities. As the SHARP database develops further,
 197 new quantities will be added, including ones that characterize the magnetic inversion lines, the
 198 relevant fractal indices, and models of the coronal field (see Section 9 for further discussion).

199 The pixels that contribute to any given index calculation are selected by examining two data
 200 segment maps: BITMAP and CONF_DISAMBIG. The BITMAP segment, an example of which is
 201 shown in the left panel of Figure 3, identifies pixels within the HARP ($\text{BITMAP} \geq 33$). Pixels
 202 with strong line-of-sight magnetic field strength are shown in white, whether inside or outside the
 203 orange HARP area. The CONF_DISAMBIG segment has a high value for clusters of pixels above the
 204 spatially and temporally dependent disambiguation noise threshold (≈ 150 G, CONF_DISAMBIG
 205 $= 90$; see Table A.5 and Hoeksema *et al.*, 2014). Only data that are both within the HARP
 206 and above the high-confidence threshold contribute to the SHARP parameter calculation; the
 207 number of contributing CEA pixels is given in the keyword CMASK. The rightmost panel of
 208 Figure 3 shows the pixels that contribute to the active region parameters for HARP 401 (NOAA
 209 AR 11166) on 9 March 2011 at 23:24:00 TAI. The indices in all four SHARP series are computed
 210 from the CEA data.

211 5. SHARP Parameters for an Illustrative Region: HARP 401

212 The SHARP indices are common active-region parameters described in the literature, as discussed
 213 in the previous section, and the formulae are given in Table 3. Figures 4 and 5 show the SHARP
 214 indices for HARP 401 from the time it first rotated onto the disk on 2 March 2011 through
 215 its final disappearance on 15 March. Computed quantities from Table 3 are plotted with error
 216 bars, except those that are areas or pixel counts. In most cases the error bars are smaller than
 217 the size of the dots because formal errors are small and systematic errors are not reflected. We
 218 have excluded data points with poor status bits set in the QUALITY keyword, which provides
 219 information about data reliability (see Table A.8 and Lev1qualBits referenced in Table 2 for more
 220 information about QUALITY).

221 The photospheric AREA (Figure 4 Panel A1, top left) is determined by the HARP module
 222 using the HMI line-of-sight magnetic field measurements. The AREA includes everything inside
 223 the orange patch in the left panel of Figure 3. This established active region rotates onto the
 224 disk on 2 March and grows steadily as it crosses the disk. The patch reaches a maximum area
 225 of ≈ 7500 microhemispheres on 11 March before it starts to decrease as it rotates off the disk.
 226 The panel below (Figure 4 Panel A2) shows the total number of high-confidence pixels that
 227 contribute to the SHARP index calculation, CMASK, *i.e.* the pixels in white in the right panel of
 228 Figure 3. Once the region is on the disk, the number of CMASK pixels increases from about 40 000
 229 to nearly 80 000. The number of contributing pixels changes with the size of the region and also
 230 depends on the noise threshold that varies with location on the disk and velocity of SDO relative
 231 to the Sun (see Section 7.1 of Hoeksema *et al.*, 2014). A histogram of the total-field noise level
 232 (not shown) increases and broadens near 60° from central meridian, consequently increasing the
 233 number of pixels above the noise threshold relative to disk center.

234 For comparison, Figure 4 Panel A3 shows the area of the strong active pixels determined from
 235 the line-of-sight field during the initial identification of the HARP region. This area, AREA_ACR,
 236 associated with the white pixels inside the orange patch on the left of Figure 3, is smaller than
 237 the area associated with the high-confidence pixels in the center panel of that figure. The area
 238 of strong field shows a steady 40% increase during the new flux emergence on 7 – 8 March.
 239 The total unsigned flux [USFLUX] computed from the radial component of the vector magnetic
 240 field appears in Figure 4 Panel A4, at the bottom of the left column. The total flux, initially
 241 about 3×10^{22} Mx, decreases by 20% on 6 March, recovers by a similar amount late on 7 March,
 242 and then gradually builds to about 5×10^{22} Mx on 13 March. Variations in USFLUX in this
 243 time interval do not exactly track changes in the area of the region, the number of pixels in

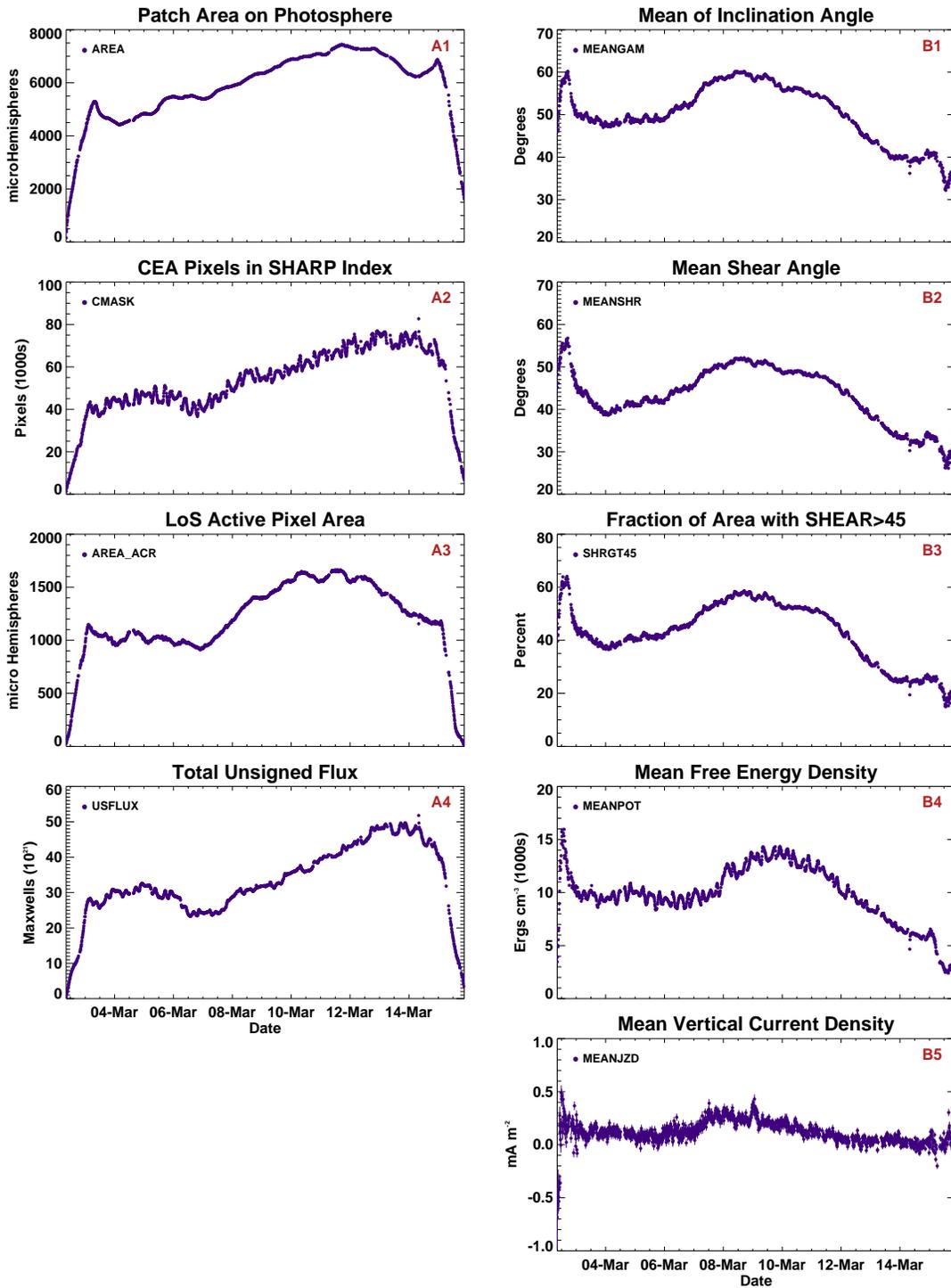


Figure 4. SHARP Active-Region Parameters for HARP 401, 2–15 March 2011. Column A on the left shows four quantities: Panel A1, AREA; A2, CMASK; A3, AREA_ACR; and A4, USFLUX; Column B on the right shows five quantities: Panel B1, MEANGAM; B2, MEANSHR; B3, SHRGT45; B4, MEANPOT; and B5, MEANJZD.

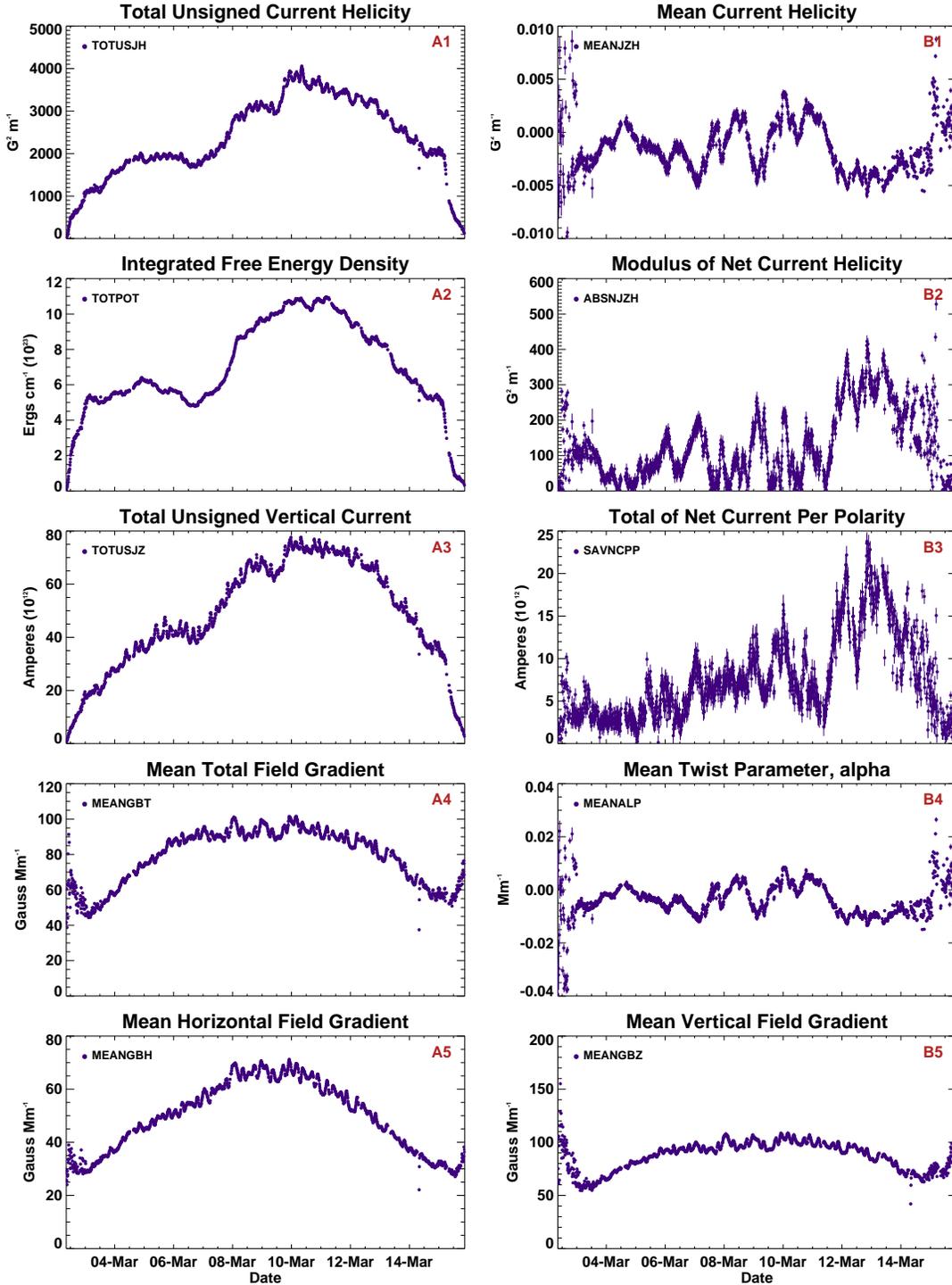


Figure 5. Additional SHARP Active-Region Parameters for HARP 401, 2–15 March 2011. Column A on the left shows five quantities: Panel A1, TOTUSJH; A2, TOTPOT; A3, TOTUSJZ; A4, MEANGBT; and A5, MEANGBH. Column B on the right shows five quantities: Panel B1, MEANJZH; B2, ABSNJZH; B3, SAVNCP; B4, MEANALP; and B5, MEANGBZ.

244 the computation, or the strong-pixel area, indicating that the strength of the field in the region
 245 is also changing. Correlated daily variations in USFLUX and CMASK are associated with SDO's
 246 geosynchronous orbital velocity. The episode of flux emergence during 7 and 8 March is reflected
 247 in a number of the quantities. The largest flare produced by HARP 401, an X 1.5 flare, peaked
 248 at 23:23 TAI on 9 March, about the time that the active-pixel area first reaches a maximum.
 249 Numerous C-class and M-class flares occurred during the lifetime of the region.

250 The systematic change in the transverse-field noise level is reflected in the trend of the mean
 251 value of the inclination angle (MEANGAM) shown in Panel B1 at the top right of Figure 4.
 252 The plot shows both the evolution of the region and a position-dependent trend that results
 253 from the different strengths and noise levels in the circular and linear polarization signals. (See
 254 Borrero and Kobel (2011) for a relevant discussion of the effects of noise on the interpretation of
 255 vector field measurements.) At disk center, the vertical magnetic-field component [B_z] is closest
 256 to the lower-noise line-of-sight direction that depends on the stronger Stokes- V ; the horizontal
 257 component [B_h] reflects the sensitivity to noise in Stokes- Q and $-U$. In weak-field pixels this
 258 tends to bias the inclination angle away from 0° . The relative contributions of noise to the
 259 vertical and horizontal field components change with center-to-limb angle [μ]. As a consequence
 260 the ratio B_z/B_h in the weak-field pixels increases, decreasing the horizontal bias in the reported
 261 inclination. MEANGAM reaches a maximum of $\approx 60^\circ$ from radial near disk center and shows two
 262 broad minima at 45° and 40° when the region is near the east and west limbs, respectively, where
 263 the noise contributions to the vertical and horizontal field components are roughly the same.

264 The mean shear angle [MEANSHR] in Figure 4 Panel B2 shows a similar variation across the
 265 disk, with a maximum a little over 50° near central meridian passage and broad minima below
 266 40° and 35° in the East and West, respectively. The shear angle is calculated by determining the
 267 angle between the observed field [B^{Obs}] and a potential field [B^{Pot}]. To compute the parameters
 268 that require a potential-field model, we use the discretized Green's function based on Equation
 269 (2.14) of Sakurai (1982), which is the potential due to a submerged monopole at a depth of
 270 $\Delta/\sqrt{2\pi}$. In that case, Δ is the size of a pixel, which preserves the total flux of B_z . However,
 271 using that depth yields a B_z map that is blurry compared to the original observational data,
 272 which, in turn, yields blurry calculated B_x and B_y maps. Therefore, we choose a smaller Δ that
 273 corresponds to 0.001 pixels. Since this yields a sharper B_z map, with a resolution similar to the
 274 original observational data, the calculated B_x and B_y maps are of a higher resolution as well.
 275 We preserve the original observational data for the z -component of the potential magnetic field.
 276 Figure 4 Panel B3, the fraction of CMASK pixels with shear greater than 45° [SHEARGT45] shows
 277 a pattern very similar to the mean shear and mean inclination angle. Trends in the large-scale
 278 averages are affected by what is happening in the weak and intermediate field strength pixels
 279 near the noise level and the systematic change in reported field direction from center to limb.
 280 There is a few percent decrease in the fraction of strong-shear pixels over the course of 9 March,
 281 prior to the X-class flare, which may or may not be significant.

282 Figure 4 Panel B4 presents the mean value of the free energy density averaged over the
 283 patch, MEANPOT. MEANPOT shares evolutionary characteristics of the shear and inclination angle.
 284 Figure 4 Panel B5 (bottom right) shows the evolution of the mean vertical current density
 285 [MEANJZD]. The point-to-point scatter and the uncertainties in this quantity are relatively larger
 286 than for most of the other SHARP parameters. The mean vertical-current density more than
 287 doubles from about 0.1 to 0.25 mA m $^{-2}$ on 7 March when new flux began to rapidly emerge.
 288 The vertical current is computed using derivatives of the horizontal magnetic-field components.
 289 To compute any of the parameters that require a computational derivative, we use a second-order
 290 finite-difference method with a nine-point stencil centered on each of the CMASK pixels.

291 We now consider Figure 5, which shows additional SHARP parameters for the same HARP
 292 401. Figure 5 Panels A1 and A2 on the upper left show the total unsigned current helicity
 293 [TOTUSJH] and a proxy for the integrated total free-energy density [TOTPOT]. Both quantities
 294 show a sustained increase on 7 March when new flux was emerging. The total current helicity
 295 showed a sharp increase from 3100 to 3900 G 2 m $^{-1}$ on 9 March leading up to the X-class flare.
 296 The integrated free-energy density is the difference between the observed and potential magnetic-
 297 field energy integrated over the region. TOTPOT nearly doubles from 5×10^{23} to 9×10^{23} erg cm $^{-1}$
 298 on 7 March; however, no obvious signal associated with the flare or its immediate aftermath is

299 reflected in the free energy density plot. In fact TOTPOT continues to increase gradually until 11
300 March.

301 The total unsigned vertical current (TOTUSJZ in Figure 5 Panel A3) changes dramatically
302 during the life of HARP 401. Like the current helicity and integrated free energy density, it
303 reaches a plateau on 5 March and then increases rapidly on 7 and 8 March from 4×10^{13} to
304 7×10^{13} A. A dip and rapid rise occur on 9 March before the X-class flare, after which the
305 current stabilizes for several days.

306 Figure 5 Panels A4, A5 (bottom left), and B5 (bottom right) show the temporal dependence
307 of the horizontal gradients of the field. Each index is the mean value of the gradient computed
308 at the CMASK pixels in the patch. Figure 5 Panel A4 shows the mean horizontal gradient of
309 the total field magnitude [MEANGBT]. There is a fairly clear daily periodicity associated with
310 the spacecraft velocity and the number of pixels in CMASK. The daily variation is superposed
311 on a broad peak near central meridian at about 100 G Mm^{-1} . The same shape is evident in
312 Figure 5 Panel A5, which shows the horizontal gradient of the horizontal component of the field
313 [MEANGBH]. The peak is a little sharper, ranging from $\approx 20\text{--}65 \text{ G Mm}^{-1}$ during the disk passage
314 of the region. Figure 5 Panel B5 (on the lower right), shows that the horizontal gradient of the
315 vertical component of the field [MEANGBZ] is less sharply peaked near central meridian and has
316 a more pronounced daily variation. Consideration of other regions (see the discussion of HARP
317 2920 and Figure 6) suggests that the broad shape tends to follow that of CMASK and AREA; so,
318 perhaps the mean gradient of the vertical field is more heavily influenced by the contributions
319 of the variable number of weak-field pixels than are the means of the total or horizontal field
320 gradient.

321 Figure 5 Panel B1 (upper right) shows the mean of the contribution to the current helicity
322 from the vertical components of the magnetic field and the current density [MEANJZH]. We cannot
323 calculate the other terms that contribute to the total helicity because HMI cannot determine
324 the field gradient in the vertical direction. The mean current helicity is generally negative for
325 this region through much of its lifetime and shows relatively strong variability while the region
326 is evolving rapidly from 6–11 March. Starting 12 March the helicity was relatively large in
327 magnitude, at $-0.004 \text{ G}^2 \text{ m}^{-1}$, but stable. Indices plotted in the next three panels, B2, B3,
328 and B4, are related to physical quantities associated with helicity, and thus all share a similar
329 temporal profile. The sum of the absolute values of the net current helicity [ABSNJZH] is shown
330 in Figure 5 Panel B2; the sum of absolute values of the net current determined separately in
331 the positive and negative B_z regions [SAVNCP] appears in Figure 5 Panel B3; and the mean of
332 the magnetic field twist, α , of the region [MEANALP] is in Figure 5 Panel B4. All exhibit some
333 degree of daily variation. Periodic variations are particularly strong on 6, 7, 9, and 11 March.
334 All experience a steep increase in magnitude on 11–12 March, after which the indices remain
335 fairly stable. The sum of the net currents in the two polarity regions [SAVNCP] peaks above
336 2×10^{13} A on 13 March.

337 The average twist parameter [MEANALP] posed a challenge. The simple definition of twist,
338 $\alpha = J_z/B_z$, is noisy for individual pixels when the field is low and near the noise level (*cf.* Leka
339 and Skumanich, 1999). Simply averaging the computed α in the high-confidence SHARP region
340 pixels results in a meaningless scatter of points from one time step to the next, suggesting that a
341 higher threshold may be more appropriate. Instead we calculate a parameter intended to reflect
342 the mean twist of the field in the entire active region. A variety of methods have been proposed
343 (Pevtsov, Canfield, and Metcalf, 1995; Leka *et al.*, 1996; Leka and Skumanich, 1999; Falconer,
344 Moore, and Gary, 2002) based on fits to differences from a linear force-free field, moments of
345 the distribution of α , and taking ratios of spatial averages determined in parts of the active
346 region. None of the methods is clearly superior. For the SHARP index MEANALP we adopt the
347 B_z^2 -weighted α method proposed by Hagino and Sakurai (2004) in which one simply computes
348 the sum of the product of $J_z B_z$ at the CMASK pixels and divides by the sum of B_z^2 .

349 6. Selected Parameters for a Second Region, HARP 2920

350 Considering a single active-region complex does not provide sufficient context to understand how
351 regions differ from each other or how much of the variation in a quantity depends on disk position

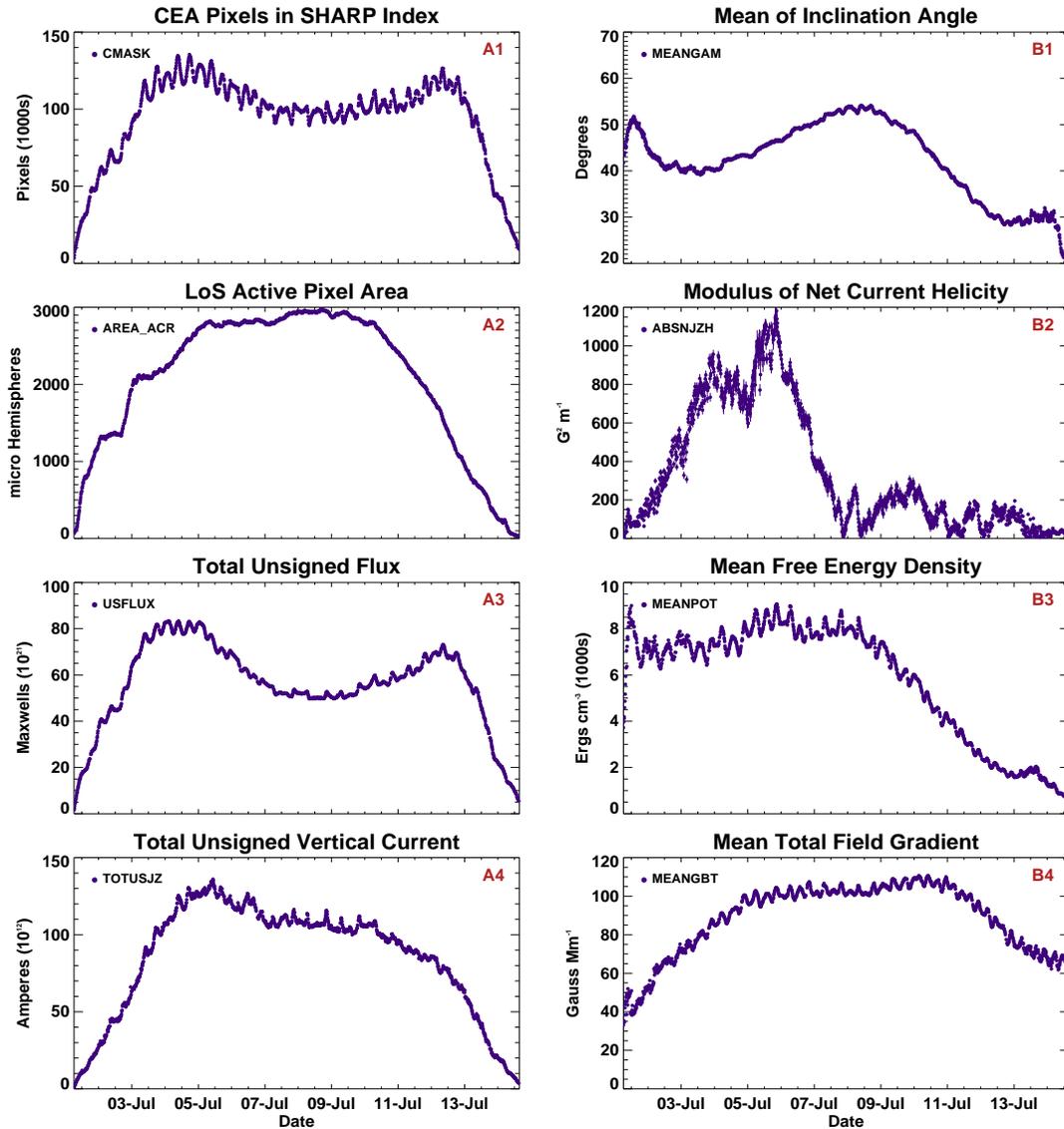


Figure 6. SHARP Active Region Parameters for HARP 2920, 1–14 July, 2013. Column A on the left shows four quantities: Panel A1, CMASK; A2, ACR_AREA; A3, USFLUX; and A4, TOTUSJZ. Column B on the right shows four quantities: Panel B1, MEANGAM; B2, ABSNJZH; B3, MEANPOT; and B4, MEANGBT.

352 or other typical evolutionary characteristics. To illustrate the differences between regions, Figure
 353 6 shows selected SHARP indices for HARP 2920 from the time that it first rotated onto the disk
 354 on 1 July 2013 through its final disappearance on 14 July. HARP 401 was energetic and large, but
 355 had reasonably simple large-scale topology. HARP 2920 was larger and more complex, ultimately
 356 including three NOAA regions: 11785, 11787, and 11788. HARP 2920 produced numerous C-class
 357 flares; the largest, class M 1.5, occurred at 07:18 UT on 3 July while the region was still near
 358 the east limb. Figure 6 Panel A1 (CMASK, upper left) shows the number of high-confidence CEA
 359 pixels that contribute to the indices. Panel A2 shows the area associated with strong pixels,
 360 AREA_ACR. The region grows as it rotates onto the disk and then on 3 and 4 July its size nearly
 361 doubles from about 1400 microhemispheres on 2 July to 2100 on 3 July as a second activity
 362 complex (AR11787) rotates over the limb and then to 2800 by the end of 4 July as new flux
 363 emerges. In the NRT HARP this appearance and nearby emergence results in the merger of two
 364 regions. The size of the region remains fairly stable as it continues to rotate across the disk. The
 365 active pixel area [AREA_ACR] starts to diminish on 10 March, but the size of the high-confidence

366 pixel area [CMASK] only begins to decrease rapidly starting on 12 July as the HARP rotates off
 367 the limb. Contrast this with the strong emergence of new flux within the existing flux system
 368 seen in HARP 401 on 8–9 July.

369 The evolution of the total unsigned flux [USFLUX] appears in Figure 6 Panel A3. The change
 370 in CMASK pixel number creates broad peaks near 60° from central meridian on 4 July and 12
 371 July in the USFLUX. The variations of CMASK and USFLUX were also correlated for HARP 401,
 372 but the evolution across the disk was much different. The trend also seems to be reflected in
 373 an inverse fashion in the mean inclination angle [MEANGAM] plotted in Figure 6 Panel B1 (top
 374 right). A similar inverted trend appears, with a broad peak near central meridian on 8–9 July,
 375 in the measures of shear angle and the mean vertical current density (not shown). The similarity
 376 of the MEANGAM profile for 401 and 2920 confirms that significant effects due to the relative
 377 noise levels in Stokes QUV are important.

378 Figure 6 Panel B2 shows the modulus of the net current helicity [ABSNJZH]. There is a strong
 379 rise on 2–4 July and again on 5 July followed by a sharp decline on 6 and 7 July. The mean-
 380 current-helicity, net-current-per-polarity, and mean-twist parameters (not shown) have a similar
 381 profile. Contrast this with the weaker and relatively less volatile behavior of HARP 401 (note
 382 the difference in plot scale) even though 401 was emerging much more new flux. The mean free
 383 energy density [MEANPOT, Figure 6, Panel B3] remains fairly stable at $7000 \text{ ergs cm}^{-3}$ from the
 384 time the region appeared until a steady decrease begins on 9 July. The mean free energy density
 385 of HARP 401 was significantly greater and increased by $\approx 30\%$ during its disk passage before
 386 beginning a similar decline. The variations of the total unsigned vertical current [TOTUSJZ, Figure
 387 6 Panel A4] are representative of the total unsigned current helicity and integrated free energy
 388 density proxy. Unlike HARP 401, these quantities in HARP 2920 do not follow the evolution of
 389 the unsigned flux or the area. There is an interesting small excursion in the vertical current on
 390 6 July just after the helicity measures reach their peak and begin their rapid decline. No similar
 391 relationship is seen in HARP 401.

392 Finally, Figure 6 Panel B4 plots the mean of the horizontal gradient of the total field strength
 393 [MEANGBT] which is indicative of the evolution of the mean gradients of the other field compo-
 394 nents. The broad hump on the MEANGBT curve that occurs on 9–10 July is not apparent in
 395 any of the indices unrelated to field strength gradients. Otherwise the evolution is very smooth,
 396 much smoother than for HARP 401. All gradient indices exhibit a short-term (12-hour) variation
 397 that is related to the sensitivity of the vector-field measurement to the orbital velocity of the
 398 spacecraft (Hoeksema *et al.*, 2014). The general profile of the mean gradient of the horizontal-
 399 field component (not shown) for HARP 2920 has a broad peak near central meridian passage, as
 400 does the area of the strong-field elements. The mean gradients of the total and vertical field (not
 401 shown) follow more closely the flatter shape of the total area, with additional broad increases
 402 appearing near 60° from central meridian associated with the increase in the number of weak
 403 and intermediate strength pixels, though both start to fall off steadily on 10 July.

404 7. Definitive and Near-Real-Time (NRT) SHARPs

405 The definitive HARP processing module groups and tailors the identified regions according to
 406 their complete life history. The definitive HARP geometry is determined only after an active-
 407 region patch has crossed the face of the disk. At each time step the rectangular bounding box
 408 of a definitive HARP on the CCD encloses the fixed heliographic region that encompasses the
 409 greatest geometric extent attained by the patch during its entire lifetime. The temporal life of
 410 a definitive HARP starts when it rotates onto the visible disk or two days before an emerging
 411 magnetic feature is first identified in the photosphere. The HARP expires two days after the
 412 feature decays or when it rotates completely off the disk. The center of the HARP at central
 413 meridian passage is uniformly tracked at the differential rotation rate appropriate for its latitude,
 414 given in keyword OMEGA_DT. There is necessarily a delay of about five weeks before definitive
 415 SHARPs can be created.

416 Operational space-weather forecasting requires more timely data and would need to rely on
 417 the HMI NRT data stream. We outline below three primary differences between the NRT data

418 and definitive SHARP data. Note that the HARNUM for a particular region will be different for
 419 the definitive and NRT SHARP series. The NRT SHARPs are offered “as is”, *i.e.* there is no plan
 420 to necessarily correct the NRT data series when updates are made to the definitive SHARPs.
 421 The NRT SHARP archive begins 14 September 2012, but because of the inferior quality of the
 422 NRT data, we strongly recommend against use of the NRT data except for forecasting and
 423 development of forecasting tools.

- 424 i) The NRT and definitive observables input data differ in completeness and calibration. Roughly
 425 4 % of the data are delayed more than one hour; delays tend to be more clustered than random.
 426 Calibrations and corrections to the NRT data rely on predicted conditions or on calibration
 427 information that may be increasingly out of date as the day progresses. Effects of cosmic rays
 428 are not corrected. The differences are generally minor or localized. For a detailed summary
 429 of calibration procedures and the differences between the NRT and definitive input data, see
 430 Hoeksema *et al.* (2014).
- 431 ii) NRT HARP geometry is determined as soon as possible, before the full life cycle of the region
 432 is known. For that reason the photospheric area enclosed by the box bounding the active
 433 region can grow (but will never shrink) with time. In addition, the heliographic center of the
 434 NRT HARP bounding box may shift in time as a region evolves. In general the size and shape
 435 of the patch itself is the same in NRT and definitive HARPs. It is important to note that
 436 NRT HARPs may merge, resulting in the termination of one HARP and the continuation of
 437 another HARP, but augmented by the content of the terminated HARP. This will typically
 438 cause a major discontinuity in the NRT SHARP indices at that time step. The H_MERGE
 439 keyword is set when such a merge occurs, so that merging can be taken into account when
 440 the discontinuities are observed. The H_MERGE keyword is also carried over into the definitive
 441 HARPs, but in this case the region configuration is consistent before and after the merge (the
 442 entire future of all regions is available), so for definitive HARPs, the relic H_MERGE keyword
 443 is not particularly significant. At least one merger occurred during the lifetime of 494 of the
 444 first 3213 HARPs. Note again that the NRT and definitive HARNUM will not be the same.
- 445 iii) For NRT processing the annealing parameters for the disambiguation code are adjusted to
 446 enable faster computation (Barnes *et al.*, 2014) and a smaller buffer outside the HARP is
 447 used to compute the potential-field starting point. The keyword AMBNPAD gives the size of
 448 the buffer and is reduced to 50 currently for NRT SHARPs from the 500 used for definitive
 449 processing. To investigate how these input parameters affect the active-region indices, we
 450 disambiguated a five-day cube of inverted data for HARP 401 using the two different sets of
 451 disambiguation parameters. The resulting active-region indices generally differ by less than
 452 a percent. For example, the typical difference in the total field gradient was less than 0.05 %
 453 with a maximum difference of 0.3 %. Starting on 15 January 2014, the definitive SHARPs rely
 454 on full-disk rather than patch-wise disambiguation.

455 Hoeksema *et al.* (2014) present for HARP 2920 a detailed comparison between the definitive
 456 and quick-look total unsigned flux parameter and find that the typical difference is about 1 % (see
 457 their Figure 5). The differences have some systematic periodic components, likely attributable
 458 to differences in calibration. The differences increase to a few percent when SHARPs are near
 459 the limb. By far the largest difference ($\approx 30\%$) is due to a merger.

460 8. Sources of Uncertainty

461 The vector-magnetogram data used in this study have uncertainties and limitations that are
 462 discussed at length by Hoeksema *et al.* (2014). Many of these issues are more significant in weak-
 463 field regions, which do not contribute directly to the computation of active-region parameters,
 464 except that in intermediate field-strength regions near the noise threshold the number of pixels
 465 can change appreciably. Systematic errors remain, the largest associated with the daily variation
 466 of the radial velocity of the spacecraft inherent to the geosynchronous orbit (*e.g.* small periodic
 467 variations in Figures 4–7). For each index we characterize the formal random error in the
 468 computed active-region parameter. The inversion code provides estimates of uncertainties at

Table 4. The following active regions that produced X-, M-, and C-class flares were used in our sample data. In the table, we list the time and position of the active region during the GOES X-Ray flux peak; however, we analyzed a five-day time series of data per active region. The latitude and longitude are given in Stonyhurst coordinates and correspond to the latitude and longitude of the flux-weighted center of active pixels at the time of the GOES X-Ray flux peak. These correspond to keywords LAT_FWT and LON_FWT.

Flare Peak [TAI]	Class	HARP	NOAA AR	(Lat., Lon.) in Degrees
2011.02.15.01:56:00	X2.2	377	11158	(-20.20, 12.77)
2011.03.09.23:23:00	X1.5	401	11166	(8.86, 10.30)
2011.09.06.22:20:00	X2.1	833	11283	(15.13, 14.19)
2012.03.07.00:24:00	X5.4	1449	11429	(17.72, -25.90)
2012.11.21.15:30:00	M3.5	2220	11618	(7.88, -5.19)
2012.11.27.21:26:00	M1.0	2227	11620	(-13.40, 41.18)
2013.01.13.00:50:00	M1.0	2362	11652	(19.49, 12.28)
2013.02.17.15:50:00	M1.9	2491	11675	(12.43, -22.75)
2012.12.25.06:43:00	C1.8	2314	11635	(11.07, 6.60)
2013.01.01.09:06:00	C1.2	2337	11640	(27.21, -0.38)
2013.01.31.04:34:00	C1.1	2420	11663	(-10.96, 9.63)
2013.02.03.18:01:00	C1.5	2433	11665	(10.66, -2.94)

469 each pixel, including χ^2 , the computed standard deviations, and certain correlation coefficients
 470 of the errors in the derived parameters. They effectively provide a way to estimate a lower
 471 limit on the uncertainties. We use the uncertainty determined for each component of the vector
 472 magnetic field and formally propagate these error estimates per pixel per unit time per quantity
 473 for each SHARP index. The uncertainty keyword is listed in the last column of Table 3. To test
 474 the results, we verified our formal error propagation via a Monte Carlo analysis in which we
 475 varied the input Stokes parameters according to the error estimates, a relatively early stage in
 476 the vector field pipeline. The variability found in the final SHARP indices is consistent with the
 477 formal error propagation results.

478 9. Sample Data and Discussion

479 For illustrative purposes, Figure 7 shows the evolution of a few SHARP parameters for selected
 480 active regions associated with X-, M-, and C-class flares (Table 4). A more complete analysis
 481 with comprehensive statistics is left for a future publication. Region selection was based on the
 482 following criteria. i) To minimize the effects of the increased noise in limb-ward data, we require
 483 that (a) the active region must be within 45 degrees of central meridian during the GOES X-
 484 Ray flux peak, and (b) for active regions that produced multiple flares, we chose the flare that
 485 occurred while the region was closest to disk center. ii) In some cases the identification and
 486 extraction algorithm (Turmon *et al.*, 2014) identifies as one coherent magnetic structure – *i.e.*,
 487 one HARP – a region associated with multiple NOAA active regions. For simplicity such HARPs
 488 were excluded from this sample. iii) We selected the largest flare class associated with that active
 489 region (*e.g.* a multi-flaring active region chosen for a C-class flare would not be associated with
 490 an M- or X-class flare). From that list we then arbitrarily selected four regions of each flare class
 491 to show as a demonstration of the presently available SHARP parameters.

492 Figure 7 shows temporal profiles for each active region, color-coded by flare class, for the
 493 unsigned flux, the absolute value of the net current helicity, the mean of the absolute value of

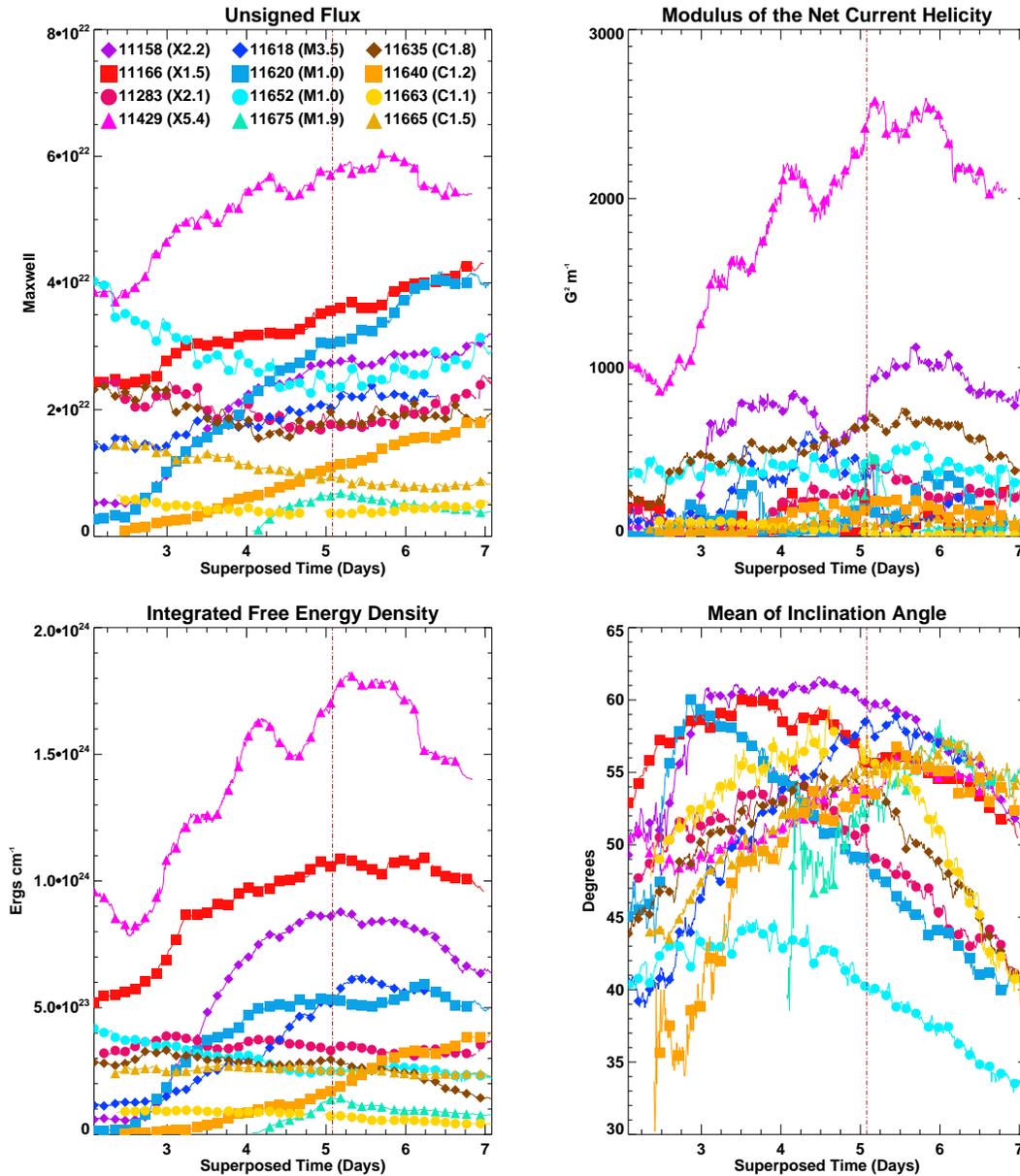


Figure 7. Clockwise from top left, temporal profiles of the total unsigned flux [USFLUX], the modulus of net current helicity [ABSNJZH], the mean value of the inclination angle [MEANGAM], and the integrated total free-energy density per active region [TOTPOT]. The entire sample is color coded: Active regions associated with X-class flares are represented with red-purples, M-class by blue-greens, and C-class by yellow-browns. For clarity a larger symbol is plotted every three hours, *i.e.* every 15th point. The legend is in the top-left panel. The time profiles are adjusted to align the flare peaks a little after the start of Day 5, as denoted by the red dotted-dashed line. Error bars are plotted for all points; however, in most cases, they are smaller than the point size. Scatter in the active-region parameters for NOAA AR 11429 for a few points following the flare peak is due to poor data quality following an eclipse: thermal changes in the HMI front window affect the focus. Periodicities in some of the parameters, most prominently in some temporal profiles of unsigned flux, are systematic effects due to the daily variation of the radial velocity of the spacecraft inherent to the geosynchronous orbit.

494 the inclination angle, and a proxy for the total free-energy density. These and other active region
 495 parameters appear as keywords in the SHARP data series and so can be displayed, retrieved, or
 496 used in a query with the JSOC data-handling tools without having to retrieve the image data.
 497 A link to examples that can be used interactively with the JSOC lookdata program can be found
 498 at the magnetic field portal (see Table 2). The temporal profiles are adjusted to align the flare
 499 occurrence time to a little after the start of Day 5, as indicated by the red dotted-dashed line.
 500 The SHARP data can be used to create temporal profiles of the parameters for any active region
 501 since 1 May 2010. Note that at the time of writing, the HMI analysis pipeline is running as fast
 502 as practical to close the remaining gap in SHARP coverage by mid-2014.

503 We chose the four parameters in Figure 7 to suggest possible uses of SHARP indices for quickly
 504 and easily comparing regions of interest. Magnetic flux has been well correlated with flaring
 505 activity (*e.g.* Barnes and Leka, 2008; Komm *et al.*, 2011; Welsch, Christe, and McTiernan, 2011;
 506 and Georgoulis, 2012), although the line-of-sight magnetic field data are known to suffer from
 507 bias. Region 11429 was much greater in both total unsigned flux (upper left panel of Figure 7)
 508 and in flare magnitude (Class X5.4). Small flux regions showed little flare activity. It is easy
 509 to track the growth rate of total flux, *e.g.* region 11620 grows rapidly during its disk transit.
 510 Statistical studies of flare-related magnetic field configurations, including the best determinations
 511 of the true total magnetic flux, have been performed with vector magnetic data (*e.g.* Leka and
 512 Barnes, 2007, Barnes and Leka, 2008, Barnes *et al.*, 2007), albeit with the recognized limitations
 513 of ground-based data sources, many of which are now ameliorated with the SDO/HMI SHARP
 514 series. Several studies use line-of-sight magnetogram data to show that the photospheric magnetic
 515 field can store up to 50 % of the total magnetic energy (*e.g.* Priest and Forbes, 2002 and references
 516 therein); however, this percentage may change when considering the transverse component of the
 517 vector magnetic field. The integrated free-energy density TOTPOT, shown in the lower-left panel,
 518 seems to increase significantly for most, but not all, of the large-flare regions; the exception was
 519 region 11283. Fan (2009) and Fang *et al.* (2012) suggest that some eruptive flares result in an
 520 imbalance of magnetic torque at the photosphere; this may have implications for the photospheric
 521 current helicity. Two of the largest regions, 11429 and 11158, had a large net current helicity
 522 and showed abrupt changes at the time of their X-class flares (upper right panel). C1.8-class
 523 region 11631 also had reasonably high net current helicity. A more comprehensive analysis is
 524 required to see if a significant relationship exists. Hudson, Fisher, and Welsch (2008) noted
 525 that explosive events should decrease coronal magnetic energy and thus lead the coronal field
 526 to contract, increasing the inclination angle or the angle between the vertical and horizontal
 527 photospheric field. Indeed, several studies (Liu *et al.*, 2005; Petrie, 2012, 2013; Sun *et al.*, 2012;
 528 Wang, Liu, and Wang, 2012) show that the horizontal component of the magnetic field changes
 529 within select areas of an active region – in particular, near the polarity inversion line. However,
 530 the mean inclination angles shown in the lower-right panel give no indication of an obvious
 531 systematic relationship to flare size or timing. Such field changes may not be detectable in the
 532 large-scale SHARP averages shown in Figure 7.

533 We have implemented an interface to automatically submit SHARP parameters, as well as
 534 HARP geometry and location keywords, to the Heliophysics Events Knowledgebase (HEK;
 535 Hurlburt *et al.*, 2012). The HEK is a web-based tool designed to aid researchers in finding
 536 features and events of interest. Various features extracted or extrapolated from HMI data, such
 537 as the location of sunspots, polarity-inversion lines, and non-linear force-free numerical models,
 538 are already available in the HEK (see Sections 13–15 of Martens *et al.*, 2012).

539 The list of active-region parameters in the SHARP data series is by no means exhaustive. We
 540 plan to include additional parameters, including those that characterize polarity-inversion lines
 541 and field morphologies of varying complexity. Several studies show a relationship between flaring
 542 activity and properties of the polarity-inversion line. For example, Schrijver (2007) defined a
 543 parameter $[R]$ that measures the flux contribution surrounding polarity-inversion lines. After
 544 determining R for 289 active regions using line-of-sight magnetograms from the *Solar and He-*
 545 *liospheric Observatory’s Michelson Doppler Imager (SOHO/MDI)*, he found that “large flares,
 546 without exception, are associated with pronounced high-gradient polarity-separation lines.” Ma-
 547 son and Hoeksema (2010) developed a similar parameter, called the Gradient-Weighted Inversion
 548 Line Length (GWILL), applied it to 71 000 MDI line-of-sight magnetograms of 1075 active

549 regions, and found that GWILL shows a 35% increase during the 40 hours prior to an X-class
 550 flare. Falconer, Moore, and Gary (2008) devised a similar parameter $[WL_{sg}]$ and computed it for
 551 56 vector magnetic field measurements of active regions. Using WL_{sg} , they could predict CMEs
 552 with a 75% success rate.

553 Two additional approaches have been widely used to characterize active regions in the context
 554 of energetic-event productivity. One is to model the coronal magnetic field from the observed
 555 photospheric boundary and parametrize the results in order to gauge the coronal magnetic field
 556 complexity and morphology. Examples of relevant parameterizations include descriptions of the
 557 magnetic connectivity (*e.g.* ϕ_{ij} from Barnes and Leka, 2006, and B_{eff} from Georgoulis and Rust,
 558 2007), and topological descriptions (Barnes and Leka, 2006; Barnes, 2007; Ugarte-Urra, Warren,
 559 and Winebarger, 2007; Cook, Mackay, and Nandy, 2009). The results are fairly convincing that
 560 parameters based on models of the coronal magnetic field can add unique information to what
 561 is otherwise available from characterizing the photosphere. Secondly, the fractal spectrum and
 562 related parameterizations of the photospheric field provide additional measures of the magnetic
 563 complexity, although the event-predictive capabilities of such measures require additional re-
 564 search. While McAteer, Gallagher, and Ireland (2005) and Abramenko and Yurchyshyn (2010)
 565 found a relation between fractal dimension and the range of multifractality spectra and flare
 566 productivity, respectively, Georgoulis (2012) found that “both flaring and non-flaring active
 567 regions exhibit significant fractality, multifractality, and non-Kolmogorov turbulence, but none
 568 of the three tested parameters manages to distinguish active regions with major flares from the
 569 flare-quiet ones.” More study is required using these analysis approaches. As the database of
 570 SHARP active-region parameters grows, it will include parameters derived from these and other
 571 relevant studies.

572 10. Summary

573 The four SHARP data series provide a systematic active-region database of patches of photo-
 574 spheric vector magnetic field, Doppler velocity, continuum intensity, and line-of-sight magnetic
 575 field extracted and tracked to mitigate cumbersome handling of full-disk data. At each 12-minute
 576 time step the SHARP pipeline module automatically calculates sixteen indices that characterize
 577 active regions. The parameters have been chosen because they are representative examples of
 578 the types of quantities linked to active-region flare productivity in the the literature. These and
 579 other keywords can be used to identify and select regions of interest. Definitive data are available
 580 a few weeks after regions complete their passage across the disk; quick-look data for forecasting
 581 purposes are available within a few hours of being observed. We compare temporal profiles of
 582 four SHARP indices for 16 selected regions at the times of flares of various classes. We expect to
 583 add several more parameters to the database. The SHARP database can enable a more thorough
 584 investigation of these parameters as statistics accumulate.

585 Acknowledgments

586 We thank the many team members who have contributed to the success of the SDO mission and
 587 particularly to the HMI instrument. This work was supported by NASA Contract NAS5-02139
 588 (HMI) to Stanford University. Some of the research described here was carried out by staff of the
 589 Jet Propulsion Laboratory, California Institute of Technology. Efforts at NWRRA were also sup-
 590 ported through NASA Contracts NNH09CF22C and NNH12CG10C and by NNG12PP28D/C#
 591 GS-23F-0197P from NASA/Goddard Space Flight Center. The authors thank Huned Botee for
 592 development of the SHARP Data Viewer.

593 Appendix

594 A. SHARP Data Segment Descriptions

595 The `hmi.sharp_720s` and `hmi.sharp.720s_nrt` data series, which are in CCD coordinates, include
 596 31 data arrays, or segments. Table A.1 describes the segments associated with maps of the line-

Table A.1. Line-of-Sight Observables

Segment Name	Unit	Description
MAGNETOGRAM	Mx cm^{-2}	The MAGNETOGRAM segment contains HARP-sized line-of-sight magnetic field strength data from the series <code>hmi.M_720s</code> .
DOPPLERGRAM	m s^{-1}	The DOPPLERGRAM segment contains HARP-sized line-of-sight velocity data from the series <code>hmi.V_720s</code> .
CONTINUUM	DN s^{-1}	The CONTINUUM segment contains HARP-sized computed continuum intensity data from the series <code>hmi.lc_720s</code> .

Table A.2. HARP BITMAP Information

Name	Description
BITMAP	The dimensionless BITMAP segment defines the bounding box and identifies which pixels are located within the HARP, and which are above the HARP noise threshold, by labeling each pixel with the following: <ul style="list-style-type: none"> 0 Off-disk. 1 Weak field, outside the HARP. 2 Strong field, outside the HARP. 33 Weak field, inside the HARP. 34 Strong field, inside the HARP.

597 of-sight HMI observables. Table A.2 describes the BITMAP segment associated with geometry
598 of the HARP data series. Table A.3 describes the segments associated with the vector-field
599 inversion, including the vector magnetic-field data as inclination, disambiguated azimuth, and
600 field strength. Segments that provide estimates of the uncertainties are listed separately in Table
601 A.4. Table A.5 describes two data segments associated specifically with the disambiguation
602 module. Table A.6 describes three segments that contains bits set by either the inversion or
603 disambiguation module.

604 Table A.7 describes the eleven map segments associated with the `hmi.sharp_cea_720s` and
605 `hmi.sharp_cea_720s_nrt` data series. The CEA data-series segments have been remapped to helio-
606 graphic Cylindrical Equal-Area coordinates centered on the patch. Several of the segments, such
607 as plasma parameters from the inversion module, are not included in the CEA data series.

608 The keywords for the SHARP-computed active region quantities and their associated uncer-
609 tainties are described in Table 3.

610 Each of the SHARP data series includes more than 300 keywords that provide information
611 about HARP geometry, disk position, upstream processing, data statistics, *etc.* A few are
612 described in Table A.8. Additional documentation can be found on the JSOC wiki (see Table 2).

Table A.3. In order to solve the inverse problem of inferring a vector magnetic field from polarization profiles, the Very Fast Inversion of the Stokes Vector (VFISV) module solves a set of differential equations that fit the parameters below.

Segment Name	Unit	Description
INCLINATION	Degree	The INCLINATION segment contains the magnetic field inclination with respect to the line-of-sight.
AZIMUTH	Degree	The AZIMUTH segment contains the magnetic field azimuth. Zero corresponds to the up direction of a column of pixels on the HMI CCD; values increase counter-clockwise. The AZIMUTH in the SHARP series has been disambiguated. Keyword CROTA2 give the angle between up on the CCD and North on the Sun.
FIELD	Mx cm^{-2}	The FIELD segment contains the magnetic flux density. Currently, the filling factor is set equal to unity, so this quantity is also representative of the average magnetic field strength. The uncertainty (see FIELD_ERR in Table A.4) accounts for noise in both the line-of-sight and transverse field components. Values of $\approx 220 \text{ Mx cm}^{-2}$ or less (2σ) are generally considered to be noise.
VLOS_MAG	cm s^{-1}	The VLOS_MAG segment contains the velocity of the plasma along the line-of-sight from the VFISV inversion. Positive means redshift. [Note: These data are in cm s^{-1} , whereas the Dopplergram data are in m s^{-1} .]
DOP_WIDTH	$\text{m}\text{\AA}$	The DOP_WIDTH segment contains the Doppler width of the spectral line, computed as if it were assumed to be a Gaussian.
ETA_0		The ETA_0 segment contains the center-to-continuum absorption coefficient.
DAMPING	$\text{m}\text{\AA}$	The DAMPING segment contains the electron dipole oscillation approximated as a simple harmonic oscillator. In the current version of the VFISV code, this parameter is constant and set to 0.5.
SRC_CONTINUUM	DN s^{-1}	The SRC_CONTINUUM segment contains the source function at the base of the photosphere. In the Milne–Eddington approximation, the source function varies linearly with optical depth.
SRC_GRAD	DN s^{-1}	The SRC_GRAD segment contains gradient of the source function with optical depth. By definition, $\text{SRC_CONTINUUM} + \text{SRC_GRAD} = \text{observed continuum intensity}$.
ALPHA_MAG		The segment ALPHA_MAG is defined as the portion of the resolution element that is filled with magnetized plasma. In the current version of the VFISV code, this parameter is constant and set to unity.

Table A.4. The following segments contain formal computed standard deviations and correlation coefficients of the uncertainties derived during the inversion that can be used to determine the statistical errors of the vector magnetic field. The standard deviations are the single-parameter quantities; the correlation coefficients are the double-parameter entries. The calculated uncertainties and covariances are only reliable if the VFISV solution is close to an absolute minimum.

Segment Name	Unit or Description
INCLINATION_ERR	Degree
AZIMUTH_ERR	Degree
FIELD_ERR	Mx cm^{-2}
VLOS_ERR	cm s^{-1}
ALPHA_ERR	Filling factor error, currently set to unity by VFISV.
FIELD_INCLINATION_ERR	Cross correlation of errors in field strength and inclination.
FIELD_AZ_ERR	Cross correlation of errors in field strength and azimuth.
INCLIN_AZIMUTH_ERR	Cross correlation of errors in inclination and azimuth.
FIELD_ALPHA_ERR	Cross correlation of errors in field strength and filling factor (set to unity).
INCLINATION_ALPHA_ERR	Cross correlation of errors in inclination angle and filling factor (set to unity).
AZIMUTH_ALPHA_ERR	Cross correlation of errors in azimuth and filling factor (set to unity).
CHISQ	A measure of how well the profiles are fit in the VFISV least squares iteration. CHISQ is not normalized.

Table A.5. Disambiguation Module Segments

Segment Name	Description
CONF_DISAMBIG	<p>The CONF_DISAMBIG segment identifies the confidence assigned to the final disambiguation solution for each pixel. The confidence value nominally ranges from 0–100 and depends on the field strength in the pixel compared to the estimated noise mask or proximity to strong-field areas. Currently only the values 90, 60, 50, and 0 are assigned. For patch-wise disambiguated SHARPs only three values are assigned: 90, 60, or 0.</p> <p>90 Highest Confidence: Clusters of pixels with transverse field strength that exceeds the disambiguation noise threshold by DOFFSET.</p> <p>60 Intermediate Confidence: Pixels adjacent to strong-field regions. For patch-wise disambiguated SHARPs (those based on hmi.Bharp_720s and all nrt HARPs), all pixels in the SHARP that do not exceed the noise threshold are considered intermediate. For full-disk disambiguation (SHARPs processed beginning 15 January 2014 that use hmi.B_720s), pixels within AMBNPAD=5 of a strong-field pixel.</p> <p>50 Lower Confidence: In full-disk disambiguation only, the weak-field pixels not within AMBNPAD=5 pixels of a strong-field pixel in either x or y.</p> <p>0 Not disambiguated, <i>e.g.</i> off-disk pixels.</p>
DISAMBIG	<p>The DISAMBIG segment encodes information about the results of the disambiguation calculation in three bits. Each bit represents a different disambiguation solution in weak and some intermediate confidence pixels, as described below. The three bits are identical for high-confidence pixels and for intermediate-confidence patch-wise disambiguations computed after August 2013. A bit is set when 180° needs to be added to the AZIMUTH returned by the VFISV fd10 inversion module. The SHARP module has added 180° to the reported azimuth value according to the rules described below.</p> <p>For all high-confidence pixels (CONF_DISAMBIG= 90) the HMI pipeline determines the azimuth disambiguation using the minimum energy method and records the result in Bit 0.</p> <p>For the intermediate confidence pixels (CONF_DISAMBIG= 60), the minimum energy disambiguation is determined and spatial smoothing is applied to the result and stored in Bit 0.</p> <p>The SHARPs use the results in Bit 0 to adjust the value in the AZIMUTH map segment for high and intermediate confidence pixels.</p> <p>For lower confidence pixels (CONF_DISAMBIG=50) the results of three solutions are provided.</p> <p>Bit 0 (lowest bit) gives the result of a potential field model solution. Bit 1 (middle bit) assigns a random disambiguation for the pixel. Bit 2 (higher bit) gives the radial-acute angle solution.</p> <p>The results for the radial-acute angle solution (Bit 2) are used in the SHARPs for lower confidence pixels.</p> <p>In some cases for intermediate-confidence patch-wise SHARPs disambiguated before August 2013 Bits 1 and 2 of DISAMBIG will include the results of the random or radial-acute angle solution. Use of these bits for intermediate-confidence pixels is deprecated.</p>

Table A.6. Per pixel information about the status/quality of inversion or disambiguation processing

Segment Name	Description
CONV_FLAG	<p>These values are set by the VFISV Code.</p> <ul style="list-style-type: none"> 0 Reached convergence criteria ($\chi_{old}^2 - \chi_{new}^2 < \epsilon$) 1 Continuum intensity not above required threshold. Pixel not inverted. 2 Reached maximum number of iterations before convergence. 3 Reached maximum number of iterations and finished with too many consecutive non-improving iterations (not used by the current <code>fd10</code> code). 4 Not-a-Number in the computation of χ^2. 5 Not-a-Number in Singular Value Decomposition of Hessian matrix.
CONFID_MAP	<p>The CONFID_MAP segment identifies the confidence index of the inversion output. The index value at each pixel will take the integer value from 0 (best) to 6 (worst), defined as the highest item number satisfying the following conditions:</p> <ul style="list-style-type: none"> 0 No issue found in the input Stokes. 1 Signals for the transverse field component in the input Stokes parameters [Q and U] were weak. 2 Signal for the line-of-sight field component in the input Stokes parameters [V] was weak. 3 Magnetic field signals of both LoS and transverse component were weak. 4 The ME-VFISV inversion did not converge within the iteration maximum of 200. 5 If the difference between the absolute value of the line-of-sight field strength derived from magnetogram algorithm and the absolute value of the LoS component from the VFISV inversion $B \cos(\text{inclination})$ is greater than 500 Gauss, we expect the inversion did not solve the problem correctly. 6 One (or more) of the 24 input Stokes arrays had NaN value.
INFO_MAP	<p>The dimensionless INFO_MAP segment identifies the quality index of the inversion output at each pixel. The 16 bits in the top 4 hex digits are set by the inversion module, while the 16 bottom bits are updated during the disambiguation step. The meaning of the bits is defined as follows (a star indicates an arbitrary number):</p> <p style="padding-left: 2em;">Set by Disambiguation</p> <ul style="list-style-type: none"> 0x****0000 Not disambiguated. 0x****0001 Weak field, not annealed (only for full disk, filled with potential field, radial acute, or random solution). 0x****0003 Weak field, annealed. 0x****0007 Strong field, annealed. <p style="padding-left: 2em;">Set by VFISV Inversion</p> <ul style="list-style-type: none"> 0x0000**** Pixel with no recorded inversion issue. 0x0{0-5}00**** Same as the three bits of the convergence index in CONV_FLAG. 0x0800**** Bad pixel, defined using the same criteria as 5 of CONFID_MAP. <p style="padding-left: 2em;">The following bits do not necessarily indicate errors:</p> <ul style="list-style-type: none"> 0x1000**** Low Q or U signal: $\sqrt{(Q_0 + \dots + Q_5)^2 + (U_0 + \dots + U_5)^2}$ was smaller than $0.206\sqrt{I_0 + \dots + I_5}$ (the nominal photon noise level). 0x2000**** Low V signal: $V_0 + V_1 + \dots + V_5$ was smaller than $0.206\sqrt{I_0 + \dots + I_5}$. 0x4000**** Low B_{LoS} value: B_{LoS} from magnetogram algorithm was smaller than 6.2 Gauss (the nominal noise level). 0x8000**** Missing data.

Table A.7. Map segments in the CEA SHARPs. The CEA magnetic-field values are represented differently, as spherical vector field components B_r , B_θ , and B_ϕ at each remapped grid point. Statistical uncertainties are given for each field component, but no cross-correlations are provided. The errors in B_r , B_θ , and B_ϕ at each remapped pixel are calculated from the variances of the inverted magnetic field $-B_{\text{Total}}$, inclination, and azimuth – and the covariances between them. The nearest-neighbor method is used to get the values of the variances and covariances at the original CCD pixel nearest the final remapped pixel. These values are then propagated to derive the errors for B_r , B_θ , and B_ϕ . If nothing is specified in the unit column, the quantity is dimensionless.

Segment Name	Unit	Description
BP	Mx cm^{-2}	ϕ (westward) component of the CEA vector magnetic field in the direction of solar rotation.
BT	Mx cm^{-2}	θ (southward) component of the CEA vector magnetic field.
BR	Mx cm^{-2}	Radial (out of photosphere) component of the CEA vector magnetic field.
BP_ERR	Mx cm^{-2}	Computed uncertainty (standard deviation) of the ϕ component of the CEA vector magnetic field. CEA uncertainties are determined at the nearest CCD pixel in the original computation.
BT_ERR	Mx cm^{-2}	Computed uncertainty (standard deviation) of the θ component of the CEA vector magnetic field.
BR_ERR	Mx cm^{-2}	Computed uncertainty (standard deviation) of the radial component of the CEA vector magnetic field.
MAGNETOGRAM	Mx cm^{-2}	The MAGNETOGRAM segment contains HARP-sized line-of-sight magnetogram data from the series <code>hmi.M_720s</code> . The field is remapped, but not transformed, <i>i.e.</i> it is still the line-of-sight component relative to HMI.
DOPPLERGRAM	m s^{-1}	The DOPPLERGRAM segment contains HARP-sized Dopplergram data from the series <code>hmi.V_720s</code> . The Doppler velocity is remapped, but not transformed, <i>i.e.</i> it is still the line-of-sight component relative to HMI.
CONTINUUM	DN s^{-1}	The CONTINUUM segment contains HARP-sized computed continuum intensity from the series <code>hmi.lc_720s</code> .
CONF_DISAMBIG		The CONF_DISAMBIG segment identifies the final disambiguation solution for each pixel with a value which maps to a confidence level in the result (roughly a probability). The CEA value is the same as the value of the nearest un-remapped CCD pixel.
BITMAP		The BITMAP segment identifies the pixels located within the HARP. The CEA value is the same as the value of the nearest un-remapped CCD pixel.

Table A.8. Definitions of selected SHARP keywords. See references in Table 2 for links to more information.

Keyword	Description
HARPNUM	The identifying number of the SHARP, one of two prime keywords.
T_REC	The center time of the observation, the other prime keyword.
QUALITY	A specific bit in QUALITY is set when a specific problem exists for this observation. <i>E.g.</i> bit 0x0100 is set during an eclipse. See the jsoc wiki entry <code>Lev1qualBits</code> referenced in Table 2 for details.
DATE	The time at which the SHARP module was run.
CODEVER7	The software version number of the SHARP code. Code version numbers are given for several modules in other keywords.
WCSNAME	World Coordinate System (WCS) coordinate system name. A number of keywords not listed in this table provide information about the coordinates (Thompson, 2006).
DSUN_OBS	The distance from HMI to the Sun center in meters.
OBS_VR	The radial velocity of HMI away from the Sun in m s^{-1} .
H_MERGE	Indicates if two NRT HARPs were merged at this time step.
OMEGA_DT	Rotation rate of the region in degrees per day.
NPIX	The number of CCD pixels in the patch.
CMASK	The number of pixels that contribute to the calculation of the SHARP indices.
AREA	The de-projected area of the patch in micro-hemispheres.
NACR	The number of strong LoS magnetic field pixels in the patch.
MTOT	The sum of the absolute values of the LoS magnetic field in the patch.
MNET	The sum of the LoS magnetic field in the patch.
T_FIRST	The first T_REC of this HARPNUM
T_LAST	The final T_REC of this HARPNUM
LON_FWT	The Stonyhurst longitude of the LoS flux-weighted center of the patch.
LAT_FWT	The Stonyhurst latitude of the LoS flux-weighted center of the patch.
NOAA_AR	The NOAA Active Region first associated with the patch, if any.
NOAA_NUM	The number of NOAA Active Regions associated with the patch.
NOAA_ARS	List of the NOAA Active Regions associated with the patch.

613 References

- 614 Abramenko, V., Yurchyshyn, V.: 2010, Intermittency and Multifractality Spectra of the Magnetic Field in Solar
615 Active Regions. *Astrophys. J.* **722**, 122. DOI. ADS.
- 616 Barnes, G.: 2007, On the Relationship between Coronal Magnetic Null Points and Solar Eruptive Events.
617 *Astrophys. J. Lett.* **670**, L53. DOI. ADS.
- 618 Barnes, G., Leka, K.D.: 2006, Photospheric Magnetic Field Properties of Flaring versus Flare-quiet Active Regions.
619 III. Magnetic Charge Topology Models. *Astrophys. J.* **646**, 1303. DOI. ADS.
- 620 Barnes, G., Leka, K.D.: 2008, Evaluating the Performance of Solar Flare Forecasting Methods. *Astrophys. J.*
621 *Lett.* **688**, L107. DOI. ADS.
- 622 Barnes, G., Leka, K.D., Schumer, E.A., Della-Rose, D.J.: 2007, Probabilistic forecasting of solar flares from vector
623 magnetogram data. *Space Weather* **5**, 9002. DOI. ADS.
- 624 Barnes, G., Leka, K.D., Crouch, A.D., Sun, X., Wagner, E.D., Schou, J.: 2014, The helioseismic and magnetic
625 imager (hmi) vector magnetic field pipeline: Disambiguation. *Solar Phys.*, in preparation.
- 626 Borrero, J.M., Kobel, P.: 2011, Inferring the magnetic field vector in the quiet Sun. I. Photon noise and selection
627 criteria. *Astron. Astrophys.* **527**, A29. DOI. ADS.
- 628 Borrero, J.M., Tomczyk, S., Kubo, M., Socas-Navarro, H., Schou, J., Couvidat, S., Bogart, R.: 2011, VFISV:
629 Very Fast Inversion of the Stokes Vector for the Helioseismic and Magnetic Imager. *Solar Phys.* **273**, 267.
630 DOI. ADS.
- 631 Calabretta, M.R., Greisen, E.W.: 2002, Representations of celestial coordinates in FITS. *Astron. Astrophys.*
632 **395**, 1077. DOI. ADS.
- 633 Centeno, R., Schou, J., Hayashi, K., Norton, A., Hoeksema, J.T., Liu, Y., Leka, K.D., Barnes, G.: 2014, The
634 Helioseismic and Magnetic Imager (HMI) Vector Magnetic Field Pipeline: Optimization of the Spectral Line
635 Inversion Code. *Solar Phys.*, in press. DOI.
- 636 Cook, G.R., Mackay, D.H., Nandy, D.: 2009, Solar Cycle Variations of Coronal Null Points: Implications for the
637 Magnetic Breakout Model of Coronal Mass Ejections. *Astrophys. J.* **704**, 1021. DOI. ADS.
- 638 Couvidat, S., Schou, J., Shine, R.A., Bush, R.I., Miles, J.W., Scherrer, P.H., Rairden, R.L.: 2012, Wavelength
639 Dependence of the Helioseismic and Magnetic Imager (HMI) Instrument onboard the Solar Dynamics
640 Observatory (SDO). *Solar Phys.* **275**, 285. DOI. ADS.
- 641 Cui, Y., Li, R., Zhang, L., He, Y., Wang, H.: 2006, Correlation Between Solar Flare Productivity and Photospheric
642 Magnetic Field Properties. 1. Maximum Horizontal Gradient, Length of Neutral Line, Number of Singular
643 Points. *Solar Phys.* **237**, 45. DOI. ADS.
- 644 Falconer, D.A., Moore, R.L., Gary, G.A.: 2002, Correlation of the Coronal Mass Ejection Productivity of Solar
645 Active Regions with Measures of Their Global Nonpotentiality from Vector Magnetograms: Baseline Results.
646 *Astrophys. J.* **569**, 1016. DOI. ADS.
- 647 Falconer, D.A., Moore, R.L., Gary, G.A.: 2008, Magnetogram Measures of Total Nonpotentiality for Prediction
648 of Solar Coronal Mass Ejections from Active Regions of Any Degree of Magnetic Complexity. *Astrophys. J.*
649 **689**, 1433. DOI. ADS.
- 650 Fan, Y.: 2009, The Emergence of a Twisted Flux Tube into the Solar Atmosphere: Sunspot Rotations and the
651 Formation of a Coronal Flux Rope. *Astrophys. J.* **697**, 1529. DOI. ADS.
- 652 Fang, F., Manchester, W. IV, Abbott, W.P., van der Holst, B.: 2012, Buildup of Magnetic Shear and Free Energy
653 during Flux Emergence and Cancellation. *Astrophys. J.* **754**, 15. DOI. ADS.
- 654 Gary, G.A., Hagyard, M.J.: 1990, Transformation of vector magnetograms and the problems associated with the
655 effects of perspective and the azimuthal ambiguity. *Solar Phys.* **126**, 21. ADS.
- 656 Georgoulis, M.K.: 2012, Are Solar Active Regions with Major Flares More Fractal, Multifractal, or Turbulent
657 Than Others? *Solar Phys.* **276**, 161. DOI. ADS.
- 658 Georgoulis, M.K., Rust, D.M.: 2007, Quantitative Forecasting of Major Solar Flares. *Astrophys. J. Lett.* **661**,
659 L109. DOI. ADS.
- 660 Hagino, M., Sakurai, T.: 2004, Latitude Variation of Helicity in Solar Active Regions. *Pub. Astron. Soc. Japan*
661 **56**, 831. ADS.
- 662 Hagyard, M.J., Teuber, D., West, E.A., Smith, J.B.: 1984, A quantitative study relating observed shear in
663 photospheric magnetic fields to repeated flaring. *Solar Phys.* **91**, 115. DOI. ADS.
- 664 Hoeksema, J.T., Liu, Y., Hayashi, K., Sun, X., Schou, J., Couvidat, S., Norton, A., Bobra, M.G., Centeno,
665 R., Leka, K.D., Barnes, G., Turmon, M.: 2014, The Helioseismic and Magnetic Imager (HMI) Solar Vector
666 Magnetic Field Pipeline: Overview and Performance. *Solar Phys.*, in press. DOI.
- 667 Hudson, H.S., Fisher, G.H., Welsch, B.T.: 2008, Flare Energy and Magnetic Field Variations. In: Howe, R., Komm,
668 R.W., Balasubramaniam, K.S., Petrie, G.J.D. (eds.) *Subsurface and Atmospheric Influences on Solar Activity*,
669 *Astron. Soc. Pac. Conf. Ser.* **383**, 221. ADS.
- 670 Hurlburt, N., Cheung, M., Schrijver, C., Chang, L., Freeland, S., Green, S., Heck, C., Jaffey, A., Kobashi, A., Schiff,
671 D., Serafin, J., Seguin, R., Slater, G., Somani, A., Timmons, R.: 2012, Heliophysics Event Knowledgebase for
672 the Solar Dynamics Observatory (SDO) and Beyond. *Solar Phys.* **275**, 67. DOI. ADS.
- 673 Komm, R., Ferguson, R., Hill, F., Barnes, G., Leka, K.D.: 2011, Subsurface Vorticity of Flaring versus Flare-quiet
674 Active Regions. *Solar Phys.* **268**, 389. DOI. ADS.
- 675 LaBonte, B.J., Georgoulis, M.K., Rust, D.M.: 2007, Survey of Magnetic Helicity Injection in Regions Producing
676 X-Class Flares. *Astrophys. J.* **671**, 955. DOI. ADS.
- 677 Leka, K.D., Barnes, G.: 2003a, Photospheric Magnetic Field Properties of Flaring versus Flare-quiet Active
678 Regions. I. Data, General Approach, and Sample Results. *Astrophys. J.* **595**, 1277. DOI. ADS.
- 679 Leka, K.D., Barnes, G.: 2003b, Photospheric Magnetic Field Properties of Flaring versus Flare-quiet Active
680 Regions. II. Discriminant Analysis. *Astrophys. J.* **595**, 1296. DOI. ADS.
- 681 Leka, K.D., Barnes, G.: 2007, Photospheric Magnetic Field Properties of Flaring versus Flare-quiet Active Regions.
682 IV. A Statistically Significant Sample. *Astrophys. J.* **656**, 1173. DOI. ADS.

- 683 Leka, K.D., Skumanich, A.: 1999, On the value of ‘ α AR’ from vector magnetograph data - I. Methods and
684 Caveats. *Solar Phys.* **188**, 3. DOI. ADS.
- 685 Leka, K.D., Canfield, R.C., McClymont, A.N., van Driel-Gesztelyi, L.: 1996, Evidence for Current-carrying
686 Emerging Flux. *Astrophys. J.* **462**, 547. DOI. ADS.
- 687 Leka, K.D., Barnes, G., Crouch, A.D., Metcalf, T.R., Gary, G.A., Jing, J., Liu, Y.: 2009, Resolving the 180°
688 Ambiguity in Solar Vector Magnetic Field Data: Evaluating the Effects of Noise, Spatial Resolution, and
689 Method Assumptions. *Solar Phys.* **260**, 83. DOI. ADS.
- 690 Liu, C., Deng, N., Liu, Y., Falconer, D., Goode, P.R., Denker, C., Wang, H.: 2005, Rapid Change of δ Spot
691 Structure Associated with Seven Major Flares. *Astrophys. J.* **622**, 722. DOI. ADS.
- 692 Martens, P.C.H., Attrill, G.D.R., Davey, A.R., Engell, A., Farid, S., Grigis, P.C., Kasper, J., Korreck, K., Saar,
693 S.H., Savcheva, A., Su, Y., Testa, P., Wills-Davey, M., Bernasconi, P.N., Raouafi, N.-E., Delouille, V.A.,
694 Hochedez, J.F., Cirtain, J.W., Deforest, C.E., Angryk, R.A., de Moortel, I., Wiegmann, T., Georgoulis,
695 M.K., McAteer, R.T.J., Timmons, R.P.: 2012, Computer Vision for the Solar Dynamics Observatory (SDO).
696 *Solar Phys.* **275**, 79. DOI. ADS.
- 697 Mason, J.P., Hoeksema, J.T.: 2010, Testing Automated Solar Flare Forecasting with 13 Years of Michelson
698 Doppler Imager Magnetograms. *Astrophys. J.* **723**, 634. DOI. ADS.
- 699 McAteer, R.T.J., Gallagher, P.T., Ireland, J.: 2005, Statistics of Active Region Complexity: A Large-Scale Fractal
700 Dimension Survey. *Astrophys. J.* **631**, 628. DOI. ADS.
- 701 Metcalf, T.R.: 1994, Resolving the 180-degree ambiguity in vector magnetic field measurements: The ‘minimum’
702 energy solution. *Solar Phys.* **155**, 235. DOI. ADS.
- 703 Moore, R.L., Falconer, D.A., Sterling, A.C.: 2012, The Limit of Magnetic-shear Energy in Solar Active Regions.
704 *Astrophys. J.* **750**, 24. DOI. ADS.
- 705 Petrie, G.J.D.: 2012, The Abrupt Changes in the Photospheric Magnetic and Lorentz Force Vectors during Six
706 Major Neutral-line Flares. *Astrophys. J.* **759**, 50. DOI. ADS.
- 707 Petrie, G.J.D.: 2013, A Spatio-temporal Description of the Abrupt Changes in the Photospheric Magnetic and
708 Lorentz-Force Vectors During the 15 February 2011 X2.2 Flare. *Solar Phys.* **287**, 415. DOI. ADS.
- 709 Pevtsov, A.A., Canfield, R.C., Metcalf, T.R.: 1995, Latitudinal variation of helicity of photospheric magnetic
710 fields. *Astrophys. J. Lett.* **440**, L109. DOI. ADS.
- 711 Priest, E.R.: 1984, *Solar Magnetohydrodynamics, Geophysics and Astrophysics Monographs*, Reidel, Dordrecht.
- 712 Priest, E.R., Forbes, T.G.: 2002, The magnetic nature of solar flares. *Astron. Astrophys. Rev.* **10**, 313. DOI.
713 ADS.
- 714 Sakurai, T.: 1982, Green’s Function Methods for Potential Magnetic Fields. *Solar Phys.* **76**, 301. DOI. ADS.
- 715 Scherrer, P.H., Schou, J., Bush, R.I., Kosovichev, A.G., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L., Zhao,
716 J., Title, A.M., Schrijver, C.J., Tarbell, T.D., Tomczyk, S.: 2012, The Helioseismic and Magnetic Imager (HMI)
717 Investigation for the Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 207. DOI. ADS.
- 718 Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., Bogart, R.S., Hoeksema,
719 J.T., Liu, Y., Duvall, T.L., Akin, D.J., Allard, B.A., Miles, J.W., Rairden, R., Shine, R.A., Tarbell, T.D.,
720 Title, A.M., Wolfson, C.J., Elmore, D.F., Norton, A.A., Tomczyk, S.: 2012a, Design and Ground Calibration
721 of the Helioseismic and Magnetic Imager (HMI) Instrument on the Solar Dynamics Observatory (SDO). *Solar*
722 *Phys.* **275**, 229. DOI. ADS.
- 723 Schou, J., Borrero, J.M., Norton, A.A., Tomczyk, S., Elmore, D., Card, G.L.: 2012b, Polarization Calibration of
724 the Helioseismic and Magnetic Imager (HMI) onboard the Solar Dynamics Observatory (SDO). *Solar Phys.*
725 **275**, 327. DOI. ADS.
- 726 Schrijver, C.J.: 2007, A Characteristic Magnetic Field Pattern Associated with All Major Solar Flares and Its
727 Use in Flare Forecasting. *Astrophys. J. Lett.* **655**, L117. DOI. ADS.
- 728 Sun, X.: 2013, On the Coordinate System of Space-Weather HMI Active Region Patches (SHARPs): A Technical
729 Note. *ArXiv e-prints*. ADS.
- 730 Sun, X., Hoeksema, J.T., Liu, Y., Wiegmann, T., Hayashi, K., Chen, Q., Thalmann, J.: 2012, Evolution of
731 Magnetic Field and Energy in a Major Eruptive Active Region Based on SDO/HMI Observation. *Astrophys.*
732 *J.* **748**, 77. DOI. ADS.
- 733 Thompson, W.T.: 2006, Coordinate systems for solar image data. *Astron. Astrophys.* **449**, 791. DOI. ADS.
- 734 Tian, L., Alexander, D., Liu, Y., Yang, J.: 2005, Magnetic Twist and Writhe of δ Active Regions. *Solar Phys.*
735 **229**, 63. DOI. ADS.
- 736 Török, T., Kliem, B.: 2005, Confined and Ejective Eruptions of Kink-unstable Flux Ropes. *Astrophys. J. Lett.*
737 **630**, L97. DOI. ADS.
- 738 Turmon, M., Hoeksema, J.T., Sun, X., Bobra, M., Sommers, J.: 2014, The Helioseismic and Magnetic Imager
739 (HMI) Magnetic Field Pipeline: HMI Active Region Patches. *Solar Phys.*, in preparation.
- 740 Ugarte-Urra, I., Warren, H.P., Winebarger, A.R.: 2007, The Magnetic Topology of Coronal Mass Ejection Sources.
741 *Astrophys. J.* **662**, 1293. DOI. ADS.
- 742 Wang, S., Liu, C., Wang, H.: 2012, The Relationship between the Sudden Change of the Lorentz Force and the
743 Magnitude of Associated Flares. *Astrophys. J. Lett.* **757**, L5. DOI. ADS.
- 744 Welsch, B.T., Christe, S., McTiernan, J.M.: 2011, Photospheric Magnetic Evolution in the WHI Active Regions.
745 *Solar Phys.* **274**, 131. DOI. ADS.
- 746 Yang, X., Zhang, H., Gao, Y., Guo, J., Lin, G.: 2012, A Statistical Study on Photospheric Magnetic Nonpoten-
747 tiality of Active Regions and Its Relationship with Flares During Solar Cycles 22 - 23. *Solar Phys.* **280**, 165.
748 DOI. ADS.
- 749 Zirin, H.: 1988, *Astrophysics of the Sun*, Cambridge University Press, Cambridge and New York. ADS.