Automated detection of solar eruptions

N. Hurlburt*

Lockheed Martin Solar and Astrophysics Laboratory, Lockheed Martin Advanced Technology Center, 3251 Hanover Street, Palo Alto, CA, USA

*Corresponding author: hurlburt@lmsal.com

Received 21 November 2015 / Accepted 1 December 2015

ABSTRACT

Observation of the solar atmosphere reveals a wide range of motions, from small scale jets and spicules to global-scale coronal mass ejections (CMEs). Identifying and characterizing these motions are essential to advancing our understanding of the drivers of space weather. Both automated and visual identifications are currently used in identifying Coronal Mass Ejections. To date, eruptions near the solar surface, which may be precursors to CMEs, have been identified primarily by visual inspection. Here we report on Eruption Patrol (EP): a software module that is designed to automatically identify eruptions from data collected by the Atmospheric Imaging Assembly on the Solar Dynamics Observatory (SDO/AIA). We describe the method underlying the module and compare its results to previous identifications found in the Heliophysics Event Knowledgebase. EP identifies eruptions that are consistent with those found by human annotations, but in a significantly more consistent and quantitative manner. Eruptions are found to be distributed within 15 Mm of the solar surface. They possess peak speeds ranging from 4 to 100 km/s and display a power-law probability distribution over that range. These characteristics are consistent with previous observations of prominences.

Key words. Sun – Eruptions – Solar image processing – Data mining

1. Introduction

Eruptions in the low solar atmosphere are key elements in generating space weather. Large eruptions can evolve into coronal mass ejections (CMEs) that can plow through the solar wind and ultimately impact the earth’s magnetosphere (Munro et al. 1979; Gosling 1993). The source of many of these CMEs has been associated with prominence eruptions (e.g., Gopalswamy et al. 2003). Yan et al. (2011) find that approximately half of active region filament eruptions are associated with CMEs and over 90% are associated with flares. Smaller eruptions may provide the ultimate source for the solar wind (Tian et al. 2014). Regardless of their magnitude, eruptions play a significant role in the structure and dynamics of the solar atmosphere.

Identifying eruptions occurring near the solar surface is complicated by the presence of a wide variety of features and scales. Active regions, coronal holes, and filaments persist for long periods, while the short-lived eruptions pass through among them. Flares and other sudden changes in intensity add distractions that can mask or mimic motions that would otherwise be visible. As a result, automated detection of these eruptions has been challenging.

Previous studies have developed automated methods to detect and track filaments primarily in H-alpha images (Gao et al. 2002; Bernasconi et al. 2005; Qu et al. 2005; Zharkova & Schetinin 2005; Foullon & Verwichte 2006; Aboudarham et al. 2008; Scholl & Habbal 2008; Joshi et al. 2010; Labrosse et al. 2010; Yuan et al. 2011; Hao et al. 2013; Schuh et al. 2014). These methods typically use image-based feature detection followed by a tracking step comparing the results of sequential detections. Measured velocities for prominence eruptions in these studies tend to lie in the range of 10–100 km/s while quiescent prominences show velocities of around 4 km/s or less.

We take a different approach by first extracting velocities from a sequence of images and then identifying features within the resulting velocity fields. Here we use an optical flow method to identify regions of significant motion. The use of these velocity fields to define regions of interest, rather than working directly from the images, removes many of the distracting features and permits us to identify and characterize the flows in sufficient detail for further analysis. We note that Gissot et al. (2008) have previously applied an optical flow method to an erupting filament using simultaneous pairs of 30.4 nm images acquired by the EUVI instruments on the two STEREO spacecraft (Kaiser et al. 2008). However their objective was to reconstruct the three-dimensional structure of a filament that had been identified by visual inspection rather than to find the eruption autonomously.

In the following sections we present the underlying method used by Eruption Patrol (EP), assess its performance, survey the statistical properties of the resulting detections and summarize our findings in Section 4.

2. Method

Our approach to identifying solar eruptions is to extract velocity fields from sequences of solar images using the opflow3d method described in Hurlburt & Jaffey (2015) as applied to images obtained by the Atmospheric Imaging Assembly on the Solar Dynamics Observatory (SDO/AIA, Lemen et al. 2012). The Interactive Data Language (IDL) routine, opflow3d, creates an estimate for an unknown,
steady, velocity field \( \mathbf{v}(x, y) \) from a time sequence of images \( I(x, y, t) \) with the assumption that it satisfies the advection equation

\[
\frac{\partial I}{\partial t} + \mathbf{v} \cdot \nabla I = 0. \tag{1}
\]

The best-fit velocity field \( \mathbf{v}_f \) can then be found using least squares applied over a set of images \( I(x, y, t) \) over some time interval, \( \Delta t \). \texttt{opflow3d} expands the fit velocity \( \mathbf{v}_f \) in a truncated Fourier series

\[
\mathbf{v}_f = \sum_{i=-N_x}^{N_x-1} \sum_{j=-N_y}^{N_y-1} (z_i \hat{x} + \beta_{ij} \hat{y}) e^{-2\pi i (iX+jY)}, \tag{2}
\]

where \( J = \sqrt{-1} \), and \( X \) and \( Y \) are the spatial dimensions of the image set. It then determines the optimal values of the complex amplitudes \( z_i \) and \( \beta_{ij} \), where \( \hat{x} \) and \( \hat{y} \) are unit vectors, and \( N_x \) and \( N_y \) are the numbers of Fourier modes retained in the expansion.

The primary goal of EP is to identify regions of high velocity while operating within the near-real time AIA data pipeline. Hence we choose parameters for EP that can keep up with the data flow. By setting \( \Delta t = 2 \) min, \( N_x = N_y = 8 \), and \( X, Y = 4096 \) pixels, \texttt{opflow3d} can extract a velocity fit from the full-resolution AIA images in approximately one minute of computation on a 2013-vintage Apple iMac. The effective resolution element of this fit is one quarter of the shortest wavelength captured in equation (2), which is 0.25\( X/N_x = 128 \) pixels, or about 80 arcsec. Systematic errors are generally negligible given the quality and consistency of AIA data, but there may be some ringing due to Gibbs phenomena when strong motions are present at the edge of the field of view. Additional sources of error are discussed below.

A two-minute sequence of ten, full-resolution Level 1 images taken in the He II line at 30.4 nm is processed to create one velocity map. These images have had dark current and flat-field corrections and spikes caused by bad pixels and radiation hits removed. The intensity in AIA images is expected to scale as the square of the density of the emitting plasma multiplied by a function temperature (Gaetz & Salpeter 1983). Over the 2 min used in determining the fit, we expect that changes in intensity are mostly due to changes in density since temperature changes are damped relatively quickly (Priest 1982) except in actively flaring regions. Hence, the square-root of these images is a proxy for the density, which we expect to approximately obey equation (1). This is consistent with the visual impression given by viewing the image sequences. We therefore apply a square-root to the images before passing them into \texttt{opflow3d}.

Previous studies found velocities exceeding 100 km/s, which translates to about 17 arcsec over the 2-minute sampling period used for a velocity fit. This is well within our effective resolution so systematic error due to smearing should be small; it also suggests that the maximum reliable velocity estimate we can expect for our sampling choice is about 500 km/s, i.e., a displacement of 80 arcsec over the interval \( \Delta t = 2 \) min. Fitting for higher speeds would require either smaller \( \Delta t \) or larger spatial windows. For instance, we could detect speeds approaching 5 Mm/s over 80 arcsec by setting \( \Delta t = 12 \) s, the maximum AIA cadence. The minimum speed of the \texttt{opflow3d} method itself is only limited by roundoff errors. Note that if we were to apply this method to the coronal images collected by AIA, say 19.3 nm where velocities are expected to reach these ranges, we would need to adjust our parameters accordingly. In that case, the more complex variations in intensity may require an explicit treatment, such as the multiscale optical flow method described by Gissot & Hochedez (2007), which was used by Gissot et al. (2008).

The derived velocity fields are composed of multiple components, some of which are sources of error for our application: these include solar rotation, super-granulation, and other quasi-static motions. Most of these motions are small and reasonably isotropic. The solar rotation profile is neither, with a peak value of about 2.2 km/s. This can introduce a bias when using a thresholding technique to identify eruption sites. Hence, EP subtracts a background velocity corresponding to that of solid-body rotation. We do not attempt to remove smaller effects such as differential rotation and meridional circulations.

An example of the resulting flow field is displayed in Figure 1. The spatial resolution of our velocity fit was doubled to 40 arcsec here to better define the regions. The flow associated with the large filament eruption near the northern pole is clearly captured, as well as a few smaller-scale flows around the limb.

EP samples velocities every 20 min and records the time, location, and velocity at the point of maximum speed within each sample. As described above, this velocity corresponds to the best-fit over a region of about 80 arcsec in a 2 min interval. Hence the precise position of the peak is only known to that resolution. Figure 2 displays the raw output of the patrol over a seven week period starting on March 29, 2014. The effect of the rotation removal can be seen as a drop in the floor of the velocity measurements to values consistent with those expected from super-granulation (e.g., Shine et al. 2000) and other sources. Peaks corresponding to eruptions and spacecraft motions are also clearly visible. The latter are excluded in the production version of the method.

The decision to only sample 10% of the images, i.e., 2 out of 20 min, risks missing short-lived events. However, eruptions with lifetimes shorter than 20 min probably have little impact on their surroundings. Our goal here is to identify eruptions that may have significant impacts, so the computational cost savings outweigh the loss of information. Our intent is to return to periods of significant motion for a more detailed analysis in the future.

The results of this first pass are then processed to identify time periods where velocity exceeds the threshold of 3.6 km/s. This value was settled upon by the need to exclude the background motions seen in Figure 2 while generating a moderate detection rate. It also corresponds to the level of motion found in quiescent prominences. These periods, along with the largest velocity and its position, are then recorded to the Heliophysics Events Knowledgebase (HEK, Hurlburt et al. 2012) as preliminary reports of eruptions. Our intent is to analyze these more carefully in a second "characterization" routine and then replace or update these entries with more details.

### 3. Results

#### 3.1. Comparison with manual selection

We assess the performance of our method by comparing it to eruptions recorded manually to the HEK. These entries are primarily provided by members of the SDO/AIA science team who monitor data as they arrive at the AIA Validation Center.
The volunteer annotators regularly sign up for three-day shifts. Thus all the datasets used by EP, along with other AIA channels, have also been reviewed by this team. Over the interval from April 18, 2014 to July 17, 2014, a total of 43 filament eruptions and 44 eruptions were recorded by the team. For this case we consider an eruption

Fig. 1. The velocity field obtained by applying our method to 2 min of SDO/AIA 30.4 nm data starting at 2010-Aug-01T21:20. The arrows are aligned with the local velocity with areas proportional to the speed. The corresponding image is shown in the color background. Two regions are seen to be erupting: a long filament on disk is ascending into the corona; and another region in the upper left that may be part of the same eruption, or a sympathetic response. Only velocities over 1 km/s are displayed and solar rotation is not removed. The peak velocity here is 3.8 km/s.

Fig. 2. The peak speeds reported by EP in 3000 samples taken between March 29, 2014 to May 16, 2014. The preliminary results for April did not remove the effects of solar rotation, and hence have a floor of about 2 km/s. Data in May have that correction applied, and the remaining floor is a combination of differential rotation, meridional flow, super-granulation, and other ubiquitous sources. Spacecraft maneuvers and other calibrations are also present in the April results, most notably during April 23.

(see Hurlburt et al. 2012 for details).
to be any of two classes accepted by the HEK: eruptions and filament eruptions. The first is a catch-all category that may or may not be associated with a filament; the latter is associated with a filament that the observer considered to have ejected material into the corona.

As a first test, we queried the HEK for both classes of eruptions using iSolSearch\(^1\) to select the events and then exporting them into SolarSoft (Freeland et al. 2000) and using the IDL routine \(\text{hek}\_\text{match}\_\text{events}\). For this study we considered events that overlapped within an hour in time. The results are displayed in Table 1. Of the 43 filament eruptions reported by humans, 37 (79%) matched times reported by EP. The success drops to 24 (44%) when we also require a separation of less than 120 arcsec. Human reports of eruptions displayed a similar behavior, which is partly due to some observers selecting both when they generate their reports. We will discuss this further below.

As a second test, we selected the 29 events with speeds exceeding 30 km/s from EP over the same interval and compared them to entries reported by human annotators. Nine (31%) match the human annotations in time, while the remaining 20 did not. Only 7 of those 9 also overlapped spatially. All of these missed events were reviewed visually using the daily movies posted by the AIA team\(^2\) and were found to be associated with significant eruptions.

These patterns persist over the entire AIA dataset, as can be seen in Table 2. Overall EP finds about 70% of all time periods manually reported as erupting in either category. This moderate drop for \(\text{Hit}(t)\) in Table 2, as compared to Table 1, may be due to the fact that humans are particularly good at catching slow, long-duration eruptions. Comparisons of EP detections exceeding 30 km/s to manual ones show a similar drop from over 50% to 25% in Table 2. Overall EP finds about 70% of all time periods exceeding 30 km/s to manual ones show a similar drop from 31% in Table 1 to 24% in Table 2. This might indicate that the converse is also true: fast, short-duration eruptions may be overlooked by human annotators. This is consistent with what we found in the smaller sample, but we have not manually reviewed the 289 events to confirm this.

The success at matching both the time and location of filament eruptions drops to 27% in the larger sample. This may again be due to large-scale filament eruptions. In that case, the position selected by annotators will tend to be the geometric center of the filament, while the position reported by EP will be the fastest moving element, or perhaps that of a separate fast, short-lived eruption. The two eruptions seen in Figure 1 provide an example of this situation.

Table 1. A bi-directional comparison between filament eruptions (FE) and generic eruptions (ER) reported manually and EP reports over the time span from April 18, 2014 to July 17, 2014. The first two rows show the success of EP in matching times recorded by FE and ER. The third row shows the success of the manual reports in matching times reported by EP. The columns show the total number of reports in each set, the percentage of hits and misses for the reported times, and hits and misses for both time and position.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Count</th>
<th>Hit (t)</th>
<th>Miss (t)</th>
<th>Hit (x, t)</th>
<th>Miss (x, t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP from FE</td>
<td>43</td>
<td>79%</td>
<td>21%</td>
<td>44%</td>
<td>56%</td>
</tr>
<tr>
<td>EP from ER</td>
<td>44</td>
<td>84%</td>
<td>16%</td>
<td>55%</td>
<td>45%</td>
</tr>
<tr>
<td>FE + ER from EP</td>
<td>29</td>
<td>31%</td>
<td>69%</td>
<td>24%</td>
<td>86%</td>
</tr>
</tbody>
</table>

The matches to generic eruption continue to be slightly over 50%. In this larger time span, the two sets (filament eruptions and generic eruptions) are reasonably independent, with the latter catching a larger range of behavior which is left to the discretion of the annotator. The fact that there is not higher success rate here is to be expected, since EP reports only a single location for a given time sample. If multiple eruptions are occurring, all but one of them will be missed. This is another issue that will be addressed in the characterization module.

The overall accuracy of manual entries, finding less than a quarter of significant time periods identified by EP and with only 11% overlapping spatially, is noteworthy. Their accuracy appears to be uncorrelated with the eruption’s duration or magnitude. There may be some relation to spatial extent, but in reviewing a smaller sample with members of the annotation team, the main source of discrepancy is most likely a lack of attention or interest. Each annotator has particular interests related to their individual research, and there is a clear correlation between their success at this task and their interests.

3.2. Statistical properties

The EP module described above was run over the entire SDO mission up to July 12, 2014, thus spanning just over 4 years. Here we give an overview of the statistical properties of this sample. The left panel in Figure 3 displays the histogram of peak speeds detected for each recorded eruption. The distribution appears to have power-law dependence on the speed, with the largest event having a speed of 96 km/s. These are consistent with previous studies, such as Gopalswamy et al. (2003) and Wang et al. (2010). The right panel in Figure 3 displays the distribution of velocities as a polar plot. There does not appear to be a significant directional bias in the sample.

Figure 4 displays the spatial distribution of all events over this period. Eruptions are detected almost everywhere on the disk, as seen in the left panel, but are clearly clustered near the activity belts and the limb. This distribution appears to be independent of the magnitude of the events (as indicated by the color of the dots). There is a clear lack of eruptions reported near the poles which may be due to relatively slow-moving polar crown filament eruptions being masked by more dynamic regions as described in the last section.

The apparent clustering near the limb is examined in the right panel, where the histogram as a function of radius \(r\) for these events is displayed. The distribution rises from zero near disk center \((r = 0)\) until the active region bands begin to contribute at \(r \approx 0.4 R_{\text{sun}}\), where \(R_{\text{sun}}\) is the solar radius. The distribution remains relatively constant between that point until near the limb \((r \approx R_{\text{sun}})\), where the number of eruptions climbs rapidly before falling back to zero. This distribution is consistent with that expected in the case of a shallow,
optically-thin, formation layer (between 3 and 15 Mm) containing a uniformly random distribution of eruptions. Gopalswamy et al. (2003) report that (relatively large) eruptive prominences have heights between 1.1 and 1.5 $R_{\odot}$. This suggests that there may be some scale dependence on the distribution that is neglected in our simple model, which might also explain the deviations from our model for $r > R_{\odot}$. Some level of scale dependence is expected in the structuring of the solar atmosphere by magnetic fields, as in the magnetic carpet model of Title & Schrijver (1998).

Another source of systematic error may result from projection effects. If all eruptions were predominantly radial, we would expect a radial dependence in the magnitude and direction of the reported velocities. Gopalswamy et al. (2003) found this to be the case overall, but with many eruptions possessing tangential motions of the order of 10 km/s. In contrast, the motions we detect are randomly oriented and show no significant projection effects. We may resolve this discrepancy by noting that the former study was effectively tracking the centroid of a prominence while our method is measuring local velocities, which includes twisting, writhing, and streaming motions that frequently accompany eruptions. Hence, we can identify times and location of eruptions based on these associated motions regardless of where they appear.

Fig. 3. (A) Histogram of speeds (black dots) of all eruptions from May 15, 2010 to July 12, 2014. Eruptions appear to possess a power-law distribution near that of an inverse square (solid line). (B) A polar plot of the velocity vectors shows them to be reasonably isotropic with a maximum speed of 96.2 km/s. For clarity, only speeds over 10 km/s are displayed.

Fig. 4. (A) The distribution of all eruptions from May 15, 2010 to July 12, 2014 show clustering near activity belts and the limb. Black, blue, and red circles display weak (< 10 km/s), medium (< 30 km/s) and strong (> 30 km/s) eruptions respectively. (B) Histogram of radial positions. Dots represent the number of events in 10 arcsec bins. The solid line is a theoretical distribution assuming the formation layer is uniform 15 Mm thick, while the dashed is the same for a 3.5 Mm layer. Both lines are scaled to match the maximum. The distribution is compatible with a layer thickness somewhere within the range of 3–15 Mm.
on the disk. A more complete reckoning of these velocity contributions falls to the future characterization module.

4. Conclusion
We have developed an automated method for finding eruptions in the lower solar atmosphere and have deployed it within the SDO/AIA Event Detection System which operates on the data as they arrive. The method has been found to measure velocities with statistical properties consistent with previous studies. The reported eruptions also appear to be consistent with those reported by human reviewers. The automated detections are less prone to lapses in attention or skewed by personal interests, but may miss slow, long-duration eruptions. They also provide a more complete characterization of eruptions by reporting both the location and plane-of-sky velocity. The reported events are found to be distributed in a layer near the solar surface and possess a power-law distribution in peak speed. Details of these events, including summary movies, can be found using a variety of tools including Helioviewer\(^3\), iSolSearch\(^4\), and SolarSoft. As part of the HEK, they are automatically cross-referenced with solar datasets obtained by the Hinode (Kosugi et al. 2007) and Interface Region Imaging Spectrograph (IRIS, De Pontieu et al. 2014) missions.

Subsequent papers will explore how these eruptions compare with those found with other automated processes recording in the HEK and will describe a characterization module that confirms and extracts more detailed information on the eruptions reported by the EP.

Acknowledgements. Thanks to Dr. Paul Higgins for his helpful discussions and to Ryan Timmons and Sam Freeland for help in implementing the module. This work has been supported by NASA under Contract No. NNG04EA00C and Lockheed Martin Internal Informatics on the disk. A more complete reckoning of these velocity contributions falls to the future characterization module.

4. Conclusion
We have developed an automated method for finding eruptions in the lower solar atmosphere and have deployed it within the SDO/AIA Event Detection System which operates on the data as they arrive. The method has been found to measure velocities with statistical properties consistent with previous studies. The reported eruptions also appear to be consistent with those reported by human reviewers. The automated detections are less prone to lapses in attention or skewed by personal interests, but may miss slow, long-duration eruptions. They also provide a more complete characterization of eruptions by reporting both the location and plane-of-sky velocity. The reported events are found to be distributed in a layer near the solar surface and possess a power-law distribution in peak speed. Details of these events, including summary movies, can be found using a variety of tools including Helioviewer\(^3\), iSolSearch\(^4\), and SolarSoft. As part of the HEK, they are automatically cross-referenced with solar datasets obtained by the Hinode (Kosugi et al. 2007) and Interface Region Imaging Spectrograph (IRIS, De Pontieu et al. 2014) missions.

Subsequent papers will explore how these eruptions compare with those found with other automated processes recording in the HEK and will describe a characterization module that confirms and extracts more detailed information on the eruptions reported by the EP.

Acknowledgements. Thanks to Dr. Paul Higgins for his helpful discussions and to Ryan Timmons and Sam Freeland for help in implementing the module. This work has been supported by NASA under Contract No. NNG04EA00C and Lockheed Martin Internal Research Funds. The editor thanks Jean-François Hochedez and two anonymous referees for their assistance in evaluating this paper.

References

3 http://helioviewer.org
4 http://www.lmsal.com/isolsearch


Cite this article as: Hurlburt N. Automated detection of solar eruptions. *J. Space Weather Space Clim.*, 5, A39, 2015, DOI: 10.1051/swsc/2015042.