Solar thematic maps for space weather operations

E. Joshua Rigler,[1,2] Steven M. Hill,[3] Alysha A. Reinard,[2] and Robert A. Steenburgh[3]

[1] Thematic maps are arrays of labels, or "themes," associated with discrete locations in space and time. Borrowing heavily from the terrestrial remote sensing discipline, a numerical technique based on Bayes' theorem captures operational expertise in the form of trained theme statistics, then uses this to automatically assign labels to solar image pixels. Ultimately, regular thematic maps of the solar corona will be generated from high-cadence, high-resolution SUVI images, the solar ultraviolet imager slated to fly on NOAA's next-generation GOES-R series of satellites starting ~2016. These thematic maps will not only provide quicker, more consistent synoptic views of the sun for space weather forecasters, but digital thematic pixel masks (e.g., coronal hole, active region, flare, etc.), necessary for a new generation of operational solar data products, will be generated. This paper presents the mathematical underpinnings of our thematic mapper, as well as some practical algorithmic considerations. Then, using images from the Solar Dynamics Observatory (SDO) Advanced Imaging Array (AIA) as test data, it presents results from validation experiments designed to ascertain the robustness of the technique with respect to differing expert opinions and changing solar conditions.

Citation: Rigler, E. J., S. M. Hill, A. A. Reinard, and R. A. Steenburgh (2012), Solar thematic maps for space weather operations, *Space Weather*, 10, S08009, doi:10.1029/2012SW000780.

1. Introduction

[2] Forecasters at the NOAA Space Weather Prediction Center (SWPC) examine solar images and identify solar features by hand once a day, every day. The resulting hand-drawn full sun drawings are used to identify and track features that have potential space weather effects; e.g., coronal holes, which translate to high speed streams in the solar wind, and active regions, which have the potential to produce coronal mass ejections or flares (R. A. Steenburgh, SWPC Forecast Office, personal communication, 2011). This constitutes a form of thematic mapping, a process that involves assigning a label or "theme" consistent with some application of interest to an image pixel, or more generally, some spatial coordinate that can be associated with available data. This label constitutes one form of information that helps to bridge the gap between raw measurements and the scientific understanding of an observable scene (i.e., data→information→understanding, as described by *Richards* [2005]).

[3] Representative samples of SWPC synoptic drawings are shown in Figure 1. These are a fairly labor-intensive form of thematic map, whereby forecasters examine various input data before assigning labels to solar regions based on their expert judgment. Thematic maps of terrestrial features were once constructed in a similar manner, requiring many man-hours to collect in situ measurements from the field and register these onto a geographic coordinate system. Today however, most terrestrial data is collected remotely using satellites with high-resolution multispectral imagers, and the thematic maps are generated using automatic multichannel pixel classification algorithms on powerful computers. Sometimes first-principle physics are used to convert raw measurements into thematic map labels, but more often than not it is assumed that different terrestrial features possess statistically separable spectral signatures provided that 1) there exist enough well-calibrated image channels, and 2) a reasonably representative database of in situ so-called ground truth data that can be used to train and test a statistical classifier.

[4] "Ground truth" data for remotely sensed images of the solar corona are not readily available, but there exists a tremendous amount of scientific research and operational experience interpreting solar imagery to draw on instead. We contend that this combination of scientific knowledge and operational experience, combined with careful experimental design, can substitute for the ground truth data normally required to train and test automated statistical pixel classifiers. Such tools will become essential to space weather forecasting as NOAA's next generation GOES-R

[1]Geomagnetism Program, U.S. Geological Survey, Denver, Colorado, USA.
[2]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, USA.
[3]Space Weather Prediction Center, NOAA, Boulder, Colorado, USA.

Corresponding author: E. J. Rigler, Geomagnetism Program, U.S. Geological Survey, Box 25046, MS 966, Denver, CO 80225, USA. (erigler@usgs.gov)
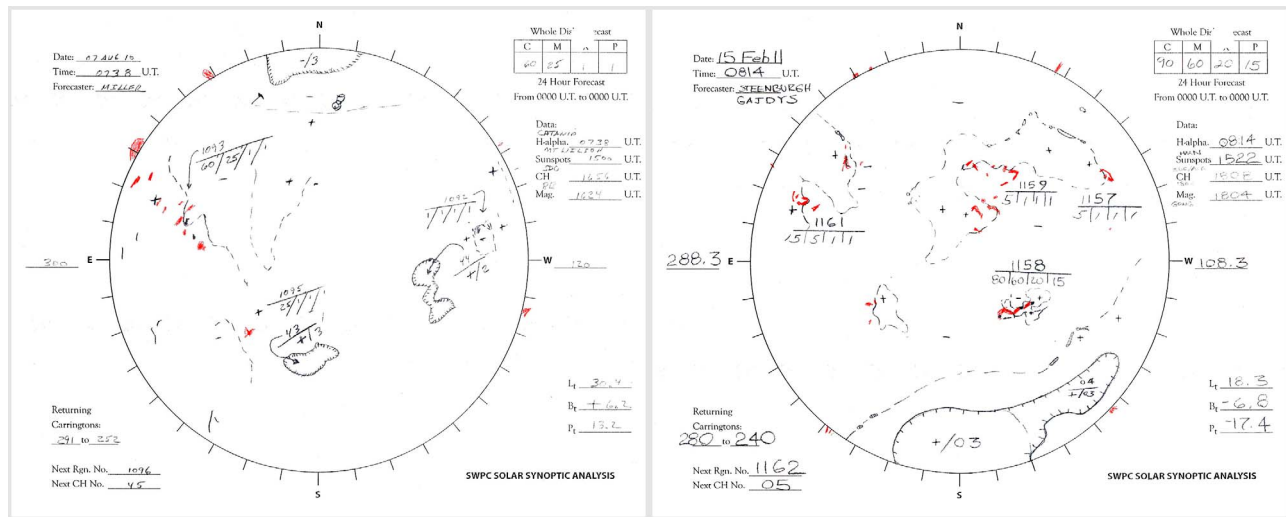
Figure 1. Sample SWPC full-disk synoptic drawings for 7 August 2010 and 15 February 2011.

satellite comes online within the next few years, and begins to return high-resolution solar images at rates never encountered by SWPC forecasters previously.

[5] As emphasized by *Richards* [2005], much of the research into automated statistical pixel labeling techniques for terrestrial remote sensing draws on pioneering work at Purdue's Laboratory for Applications of Remote Sensing (LARS), beginning with a simply named technical report "On Pattern Recognition" [*Cardillo and Landgrebe*, 1966]. Theoretical research into new classification techniques continues, but the field has become largely an applied discipline, with various textbooks [e.g., *Tso and Mather*, 2009], and even commercial software packages like ENVI [*Research Systems Inc.*, 2004] that collect the best methodologies and present them in a digestible manner for users to apply to their own thematic mapping problems. Image pixel classification has in fact come to be regarded as a "fundamental process in [terrestrial] remote sensing, which lies at the heart of the transformation from satellite image to usable geographic product" [*Wilkinson*, 2005, p. 433]. This includes the mapping of land surfaces [e.g., *Friedl et al.*, 2002; *Homer et al.*, 2004; *Bannari et al.*, 2006], sea and ice surfaces [e.g., *Ainsworth and Jones*, 1999; *Andréfouët et al.*, 2003; *Karvonen*, 2004], cloud cover [e.g., *Nair et al.*, 1999; *Platnick et al.*, 2003; *Lee et al.*, 2004], and combinations of all the above [*Simpson and Keller*, 1995].

[6] While much of the inspiration for the present article was drawn from the long history of satellite-driven thematic mapping in terrestrial applications, it should be acknowledged that automated image pixel classification is by no means a novel concept in the field of solar physics. Single-channel intensity threshold-based segmentation algorithms have been used since some of the earliest images of the sun became available [e.g., *McIntosh*, 1972; *Vaiana et al.*, 1973; *Timothy et al.*, 1975]. These have grown in

robustness and flexibility by applying histogram clustering, region growing techniques, morphological considerations, and allowing selective use of multisource images [e.g., *Worden et al.*, 1998; *Harvey and White*, 1999; *Preminger et al.*, 2001; *Qu et al.*, 2005; *Barra et al.*, 2009; *Krista and Gallagher*, 2009]. Mostly within the last half-decade, considerable progress has been made in statistical spatial feature recognition [e.g., *LaBonte et al.*, 2003; *Inhester et al.*, 2008], multispectral feature recognition [e.g., *de Wit et al.*, 2006], and combinations of the two [*Turmon et al.*, 2002, 2010]. Many of these research efforts were presented at one of a popular ongoing series of bi-annual Solar Image Processing Workshops (i.e., SIPWork I-V), and ultimately published in corresponding Solar Physics Journal topical issues [*Gallagher et al.*, 2005; *Young and Ireland*, 2008; *Ireland and Young*, 2010]. *Aschwanden* [2010] and *Martens et al.* [2012] both do a thorough job reviewing and summarizing recent work in solar image analysis.

[7] Despite this large body of solar image processing research, thematic mapping has not yet become a "fundamental process" in solar physics, or space weather research and operations for that matter. This is at least in part because it has been almost entirely motivated by, and optimized for, specific scientific inquiry. Sophisticated numerical tools have been designed to answer focused scientific questions, or if expanded to address broader problems, they still require that users be intimately familiar with the scientific questions being addressed, and/or the data being analyzed. They do not allow for the possibility that a knowledgeable non-specialist, like a space weather forecaster, might be able to derive utility from the tool, and maybe even add to its capability. Perhaps the most limiting characteristic of such predominantly science-driven solar image classifiers is that they assume a single correct answer. They generally do not allow for a more democratic,

or ensemble, solution that accounts for a broader range of expert analyses. This is a fundamental aspect of operational space weather forecasting.

[8] The remainder of this paper describes an expert-supervised solar image pixel classification scheme that, while relying on input from a variety of experts familiar with solar imagery to train the classifier up front, ultimately generates solar thematic maps that are more consistent with forecasters' operational expectations, and provides useful information to a broad range of space weather applications. Section 2 provides a practical description of the theory and algorithm used to implement this thematic mapper in an operational setting. Section 3 describes data that is expected to serve as input to the thematic mapper. Section 4 describes a validation experiment for the thematic mapper algorithm, and the results of this experiment. Finally, we conclude with comments on the limitations of this particular thematic mapper, possible improvements, and the broader implications of such a tool for space weather.

2. Statistical Theory and Algorithm Implementation

[9] Most of the following draws on long-established statistical theory, obtainable from any number of statistics, optimization, or image analysis textbooks [e.g., *Tso and Mather*, 2009], and is presented here as a convenience for the reader who might be interested in practical details of our numerical implementation. For our purposes, the most basic solar measurement is a pixel value, x, that can be assigned a unique spatial location. Pixel values obtained from the kth of ρ spectral channels, all assumed to be contemporaneous, may be combined into a multichannel pixel $\mathbf{x}_i$, where i uniquely locates the pixel on a 2D Cartesian grid (i.e., an image array), and

$$\mathbf{x} \equiv \begin{bmatrix} x_1 & x_2 & \cdots & x_{k=\rho} \end{bmatrix}^\mathsf{T}. \tag{1}$$

[10] The objective of the thematic mapper then is to assign a theme corresponding to one of a finite set of r pixel labels, w_j, to each pixel location i given its spectral, and possibly spatial, context. We deliberately chose an expert-supervised classification scheme which, while requiring input during its initial training phase, should lead to the automatic generation of thematic maps that are more consistent with the experts' expectations. It is similar to techniques promoted by *Turmon et al.* [2002], *de Wit* [2006], or *Turmon et al.* [2010], in that the label assigned to each multichannel pixel is that which maximizes the posterior probability in a Bayesian sense. In that spirit, let us start with a brief review of Bayesian statistics.

2.1. Bayes' Theorem

[11] In its most general form, Bayes' Theorem is

$$P(w_j|\mathbf{x}_i) = \frac{P(\mathbf{x}_i|w_j)P(w_j)}{P(\mathbf{x}_i)}. \tag{2}$$

[12] In words, this states that the posterior probability that multichannel pixel $\mathbf{x}_i$ should be assigned theme label w_j is equal to the probability of drawing pixel vector $\mathbf{x}_i$ from its multivariate distribution when w_j is assumed (i.e., the theme-conditioned probability), scaled by the ratio of the prior probability of being labeled w_j to the prior probability of drawing pixel vector $\mathbf{x}_i$ when no assumptions are made about its label. In multispectral pixel classification, $P(\mathbf{x}_i)$ is often ignored because one is only interested in the relative probabilities of theme membership, and $P(\mathbf{x}_i)$ is not a function of theme w_j. This leaves

$$P(w_j|\mathbf{x}_i) \propto P(\mathbf{x}_i|w_j)P(w_j). \tag{3}$$

[13] Finally, if there is no a priori reason to believe a pixel should be assigned label w_j, $P(w_i)$ can be assumed to be drawn from a uniform distribution, leaving the conditional probability

$$P(w_j|\mathbf{x}_i) \propto P(\mathbf{x}_i|w_j). \tag{4}$$

[14] Clearly the label w_j maximizing the RHS of (4) also maximizes the LHS, so it is assigned to pixel i. We refer to this as the maximum likelihood (ML) solution from here forward.

2.2. Supervised Classifier Training

[15] Before calculating theme-conditioned probabilities, it is necessary to know what the probability density function (PDF) for each pixel theme looks like. We assume these to be multivariate normal, allowing each distribution to be parameterized by a theme-dependent mean vector $\boldsymbol{\mu}_j$ and covariance matrix $\mathbf{C}_j$. These statistics are determined from representative subsets of multichannel pixels that have been assigned a theme by an expert in an offline training phase. Once training data have been collected, each mean vector is calculated as

$$\boldsymbol{\mu}_j = \frac{\sum_{i=1}^{n_j} \mathbf{x}_{j,i}}{n_j}, \tag{5}$$

where n_j is the number of multichannel pixels labeled w_j in a training data set. The mean vector for each theme looks much like the multichannel pixel vector.

$$\boldsymbol{\mu}_j \equiv \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_{k=\rho} \end{bmatrix}_j^\mathsf{T} \tag{6}$$

Each covariance matrix is a ρ-dimensional generalization of the variance, or 2nd statistical moment of a distribution of multichannel pixel values. Using the training pixels chosen for each pixel class, the covariance is calculated as

$$\mathbf{C}_j = \frac{\sum_{i=1}^{n_j} \left[\mathbf{x}_{j,i} - \boldsymbol{\mu}_j\right] \times \left[\mathbf{x}_{j,i} - \boldsymbol{\mu}_j\right]^\mathsf{T}}{n_j}. \tag{7}$$

[16] Note that (7) is not an unbiased estimate of the sample covariance, since the denominator is not adjusted to account for the reduced degrees of freedom that arise from calculating the sample mean μ_j. However, if n_j is sufficiently large, the difference is negligible, and (7) allows sample covariances to be easily and consistently merged using standard mixture reduction (SMR) techniques [*Reece and Roberts*, 2010].

[17] Given that $\mathbf{x}_{j,i}$ and μ_j are column vectors of length ρ, matrix multiplication rules turn the term inside the summation of (7) into a symmetric square matrix of dimension ρ. The diagonal elements of $\mathbf{C}_j$ hold the theme-dependent variance of each channel, while the off-diagonal elements hold the theme-dependent covariances between each channel. A covariance matrix for the *j*th class will look like

$$\mathbf{C}_j \equiv \begin{bmatrix} V_1 & C_{1,2} & \cdots & C_{1,\rho} \\ C_{2,1} & V_2 & \cdots & C_{2,\rho} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\rho,1} & C_{\rho,2} & \cdots & V_\rho \end{bmatrix}. \tag{8}$$

[18] Equation (7) guarantees that (8) will be symmetric, and that the variances will always be positive. However, a valid covariance matrix must also be positive-definite. Positive definiteness simply means that all eigenvalues for matrix $\mathbf{C}_j$ are positive, which is guaranteed if n_j is at least equal to $\rho + 1$. However, to avoid issues related to noisy measurements, n_j should be much larger than $\rho + 1$. It may be desirable to manually adjust a covariance matrix, but doing so generally removes positive definiteness. In such cases, a "nearest" positive definite matrix can be found [e.g., *Higham*, 2002], but this can be computationally prohibitive in time-critical operations.

2.3. Conditional Probabilities

[19] Given the definitions and constraints above, the probability of obtaining multichannel pixel $\mathbf{x}_i$ while assuming class label w_j is

$$P(\mathbf{x}_i|w_j) = \frac{1}{\sqrt{2\pi}^\rho \sqrt{|\mathbf{C}_j|}} \times \exp\left(-\frac{1}{2} \times \left(\mathbf{x}_i - \mu_j\right)^T \times \mathbf{C}_j^{-1} \times \left(\mathbf{x}_i - \mu_j\right)\right). \tag{9}$$

[20] Here, $|\mathbf{C}_j|$ is the scalar-valued determinant of the covariance matrix, $\mathbf{C}_j^{-1}$ is the inverse of the covariance matrix, and all other terms have been previously defined. This can be

expanded to fully represent the necessary matrix operations in (10).

[21] The large matrix operation inside the exponent reduces to a scalar, so $P(\mathbf{x}_i|w_j)$ is just a scalar that, given (4), is proportional to the probability that pixel i should be assigned label w_j. The label that maximizes this probability is assigned to pixel i, thus providing the ML thematic map solution.

2.4. Smoothness Prior

[22] ML thematic maps are prone to noise in the form of spatially isolated pixels that are mistakenly classified as one solar pixel type, even though they may be surrounded by a homogenous field of a different solar pixel type. While this may be a perfectly legitimate phenomenon, it is more likely that a solar pixel will be surrounded by pixels similar to itself. To formalize this prior assumption of smoothness, we turn to Markov Random Field (MRF) theory. *Li* [2009] provides a detailed description of how one specifies the joint probability of an MRF by maximizing the conditional probability of each image pixel using the labels of each pixel's nearest neighbors. We simply restate the end result here: given pixel i and its eight nearest neighbors N_i, and assuming that the influence of these neighboring pixels on pixel i is isotropic, the probability that pixel i should be assigned label w_j is

$$P(w_j|N_i) = \frac{\exp(\alpha_j + \beta n_i)}{\sum_{j=1}^{r} \exp(\alpha_j + \beta n_i)}. \tag{11}$$

[23] Here, r is the number of possible pixel labels, n_i is the number of pixels in neighborhood N_i that are labeled w_j, α_j is a relative weight for w_j, and β is an isotropic smoothness parameter. The parameters α_j and β are defined by the user. Roughly speaking, if α_j is bigger than other weights, theme w_j is more likely to be assigned to a given pixel, and the larger the value of β, the less likely a pixel is to be spatially isolated, resulting in a smoother thematic map. It turns out that this smoothing is not extremely sensitive to changes in β [*Besag*, 1986]. *Owen* [1989] reviews a number of studies arguing for different constraints on values for β, ultimately showing that the utility of $\beta < 0.7$ or $\beta > 4$ is negligible.

[24] Equation (11) provides us with a prior probability that pixel i should be assigned label w_j. This can be

$$P(\mathbf{x}_i|w_j) = \frac{1}{\sqrt{2\pi}^\rho \sqrt{|\mathbf{C}_j|}} \times \exp\left(-\frac{1}{2} \times \begin{bmatrix} \left[\left(x_{i,1} - \mu_{j,1}\right)\left(x_{i,2} - \mu_{j,2}\right) \cdots \left(x_{i,\rho} - \mu_{j,\rho}\right)\right]_i \times \\ \begin{bmatrix} V_1 & C_{1,2} & \cdots & C_{1,\rho} \\ C_{2,1} & V_2 & \cdots & C_{2,\rho} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\rho,1} & C_{\rho,2} & \cdots & V_\rho \end{bmatrix}_j^{-1} \times \\ \left[\left(x_{i,1} - \mu_{j,1}\right)\left(x_{i,2} - \mu_{j,2}\right) \cdots \left(x_{i,\rho} - \mu_{j,\rho}\right)\right]_i^\mathsf{T} \end{bmatrix}\right) \tag{10}$$

combined with the conditional probability determined from (9) or (10) to provide a posterior probability that pixel i belongs to a particular theme. The theme that results in the maximum a posteriori probability (MAP) is then assigned to pixel i. This MAP thematic map is generally smoother than the ML thematic map, but if the data-conditioned probabilities are strongly suggestive that an isolated pixel belongs to a theme that differs from its neighbors, it will retain that label.

2.5. Iterated Conditional Modes

[25] While pixel i relies on its nearest neighbors to provide the spatial context necessary to generate prior probabilities, it also provides context to its neighbors. Therefore, if a pixel label is altered to impose smoothness constraints, it becomes necessary to recalculate smoothness priors that account for the changed thematic map. This constitutes a nonlinear estimation problem, and requires some form of iterative solver. *Tso and Mather* [2009] propose three possible optimization techniques: simulated annealing (SA); maximizing posterior marginals (MPM); and iterated conditional modes (ICM). We chose the latter for its simplicity and speed. The algorithm is summarized as follows for generating a multichannel MAP thematic map.

Begin ICM
1. Calculate class-conditional probabilities for each pixel according to (5)–(10), then use these as initial posterior probabilities (i.e., LHS of (3));
2. Generate a MAP thematic map by assigning each pixel a label corresponding to the theme that maximizes posterior probabilities;
3. Calculate prior probabilities for each MAP thematic map pixel using its 8 nearest neighbors according to (11);
4. Combine class-conditional and prior probabilities into updated posterior probabilities according to (3);
5. Repeat steps 2–4 until convergence is achieved.

End ICM

[26] Convergence is guaranteed for ICM if pixels are processed and updated one-at-a-time. We find such sequential processing to be unusably slow for operational utility, so batch processing is applied (i.e., whole-image updates), thus allowing optimized matrix processing routines to be used. This leads to a common phenomenon whereby a proportionally small number of pixels ($\ll 1\%$) switch between two labels from one iteration to the next. This might be acceptable if a 2nd order convergence criterion was imposed, and experience suggests that $\sim$10–20 iterations would be sufficient to reach such a state of equilibrium. It is still possible, however unlikely, that pixels oscillate between more than two labels. We therefore prefer to simply set a maximum number of iterations for the ICM in lieu of defining higher-order convergence criteria. Setting this number of iterations to zero results in the ML thematic map solution.

[27] It should be noted that ICM only converges to a local maximum in posterior probability space [*Besag*, 1986],

but given our inclination to trust data over prior assumptions, we do not find the added computational complexity necessary to obtain a globally optimal solution worth the trouble of implementing a more sophisticated solver. Recent work in the optimization community employing a family of graph cut algorithms that are currently on the research frontier promises closer-to-optimal results for difficult problems [*Boykov and Kolmogorov*, 2004], and may prove useful in a future version of our algorithm.

3. Inputs and Experimental Setup

[28] The next generation of GOES weather satellites will fly with the Solar Ultraviolet Imager (SUVI). SUVI will return full dynamic range megapixel images of the sun in six relatively narrow extreme ultraviolet (EUV) spectral channels at least once every minute. SWPC forecasters and scientific modelers will be called upon to translate this flood of information into operationally meaningful reports, warnings, and predictions. This entails, among other things, maintaining consistency in solar image pixel classification over time, and perhaps more importantly, between on-duty forecasters, as well as between forecasters and scientists. Our solar thematic mapper addresses this consistency requirement. The experiments described below, while serving to validate our thematic map algorithm, should also provide guidance to forecasters and scientists who may be tasked with training, maintaining, and periodically re-evaluating the thematic mapper as it necessarily co-evolves with a changing operational and scientific understanding of the data returned from SUVI and other solar imagers.

3.1. Solar Images

[29] The SUVI instrument is not yet operational, however its design is very similar to the Solar Dynamics Observatory (SDO) Atmospheric Imaging Array (AIA). Both are normal incidence generalized Cassegrain telescopes that use a combination of filters and multilayer mirrors optimized for similar narrow-band EUV channels to focus solar photons onto back-illuminated CCDs [*Lemen et al.*, 2012; *Martinez-Galarce et al.*, 2010]. However, there are three important differences: (1) SDO-AIA's mirror area per spectral channel is $\sim$4$\times$ SUVI's mirror area per spectral channel; (2) SDO-AIA's CCDs have 4096 $\times$ 4096 pixels, while SUVI's has 1024 $\times$ 1024 pixels; (3) SDO-AIA has four telescopes, while SUVI has just one.

[30] Assuming a similar field of view and exposure time, the larger mirror area and smaller pixel size of SDO-AIA means its per-pixel signal-to-noise ratios (SNRs) will be $\sim$1/4 SUVI's. However, four telescopes mean SDO-AIA can have exposures 4$\times$ longer than SUVI and still meet the same per-channel cadence requirements, thus bringing the two instruments' SNRs roughly in line. Binning/averaging SDO-AIA pixels to a similar resolution as SUVI then results in an SNR that is $\sim$4$\times$ better.

[31] Finally, the thematic mapper assumes that all input images are of the same solar scene, just measuring different physical quantities organized by channels. That is, each
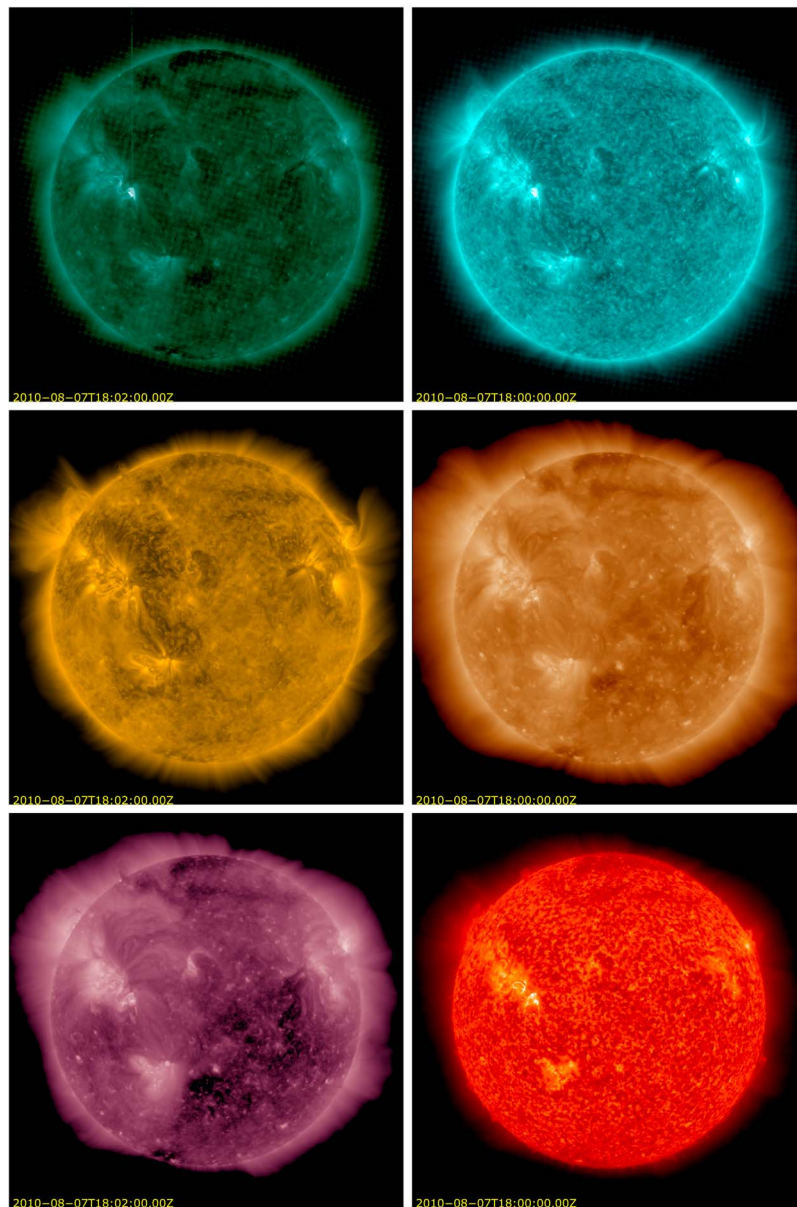
Figure 2. A multichannel SUVI proxy image is generated from comparable SDO-AIA EUV channels.

pixel across the different channels corresponds to the same location in space and time. Each SDO-AIA image contains metadata that describe its orientation and position with respect to a fixed solar coordinate system, as well as an accurate time stamp. This information is used to center, rotate, and rescale each image to a common point of view, as well as correct for solar rotation when appropriate. A representative set of these proxy images is presented in Figure 2.

3.1.1. Solar EUV Channels

[32] SUVI's six spectral channels are 94Å, 131Å, 171Å, 195Å, 284Å, and 304Å. The shortest wavelength channel was chosen primarily to capture flare location and morphology; the next three wavelengths capture different aspects of quiet to active corona; the 284Å channel maximizes contrast between open-field coronal holes and the surrounding closed-field corona; and the 304Å HeII channel peers through most of the corona into the transition region and chromosphere. The SDO-AIA proxy channels selected for this study are 94Å, 131Å, 171Å, 193Å, 211Å, and 304Å. Five of these channels match SUVI's channels very closely or exactly. The 211Å channel was chosen to substitute for the SUVI 284Å channel based on a
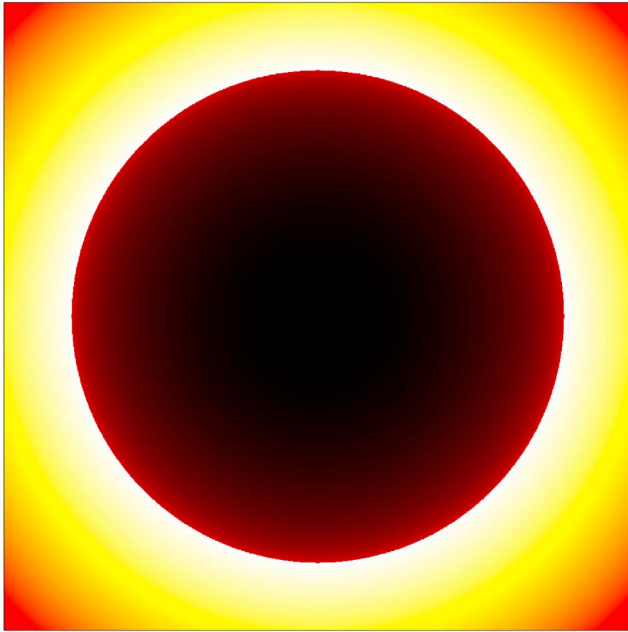
Figure 3. The path length ($\log_{10}$ km) through a simulated corona with a uniform plasma density, uniform brightness temperature, and a constant height of 1 R_S above the solar surface, serves as a pseudo-channel that helps alleviate certain geometric artifacts in pixel classifications.

limited set of visual comparisons of SDO-AIA and SOHO-EIT images at these respective wavelengths, not for any known similarities in plasmas that generate these spectral lines.

[33] While each of the individual spectral channels was chosen carefully to meet well-defined operational and scientific objectives, in principle they all help improve one's ability to isolate and classify all of the solar features just described, and more. For example, it is expected that the longest wavelength channel will help disambiguate dim coronal holes and filaments, the on-disk manifestation of solar prominences. So long as pixels from different channels can be considered co-located, every channel should help refine the classification accuracy of interesting solar features, even if the channel wasn't specifically designed to do so. In reality the solar corona is a three-dimensional structure with dominant plasma distributions that vary by height above the photosphere, each with different effective temperatures to which each spectral channel is more or less sensitive. This means that pixels from different channels are not necessarily co-located, and it is possible that a poorly chosen channel might reduce classification accuracy by effectively masking a feature below it, or even out-shining a dimmer feature above it. This problem warrants further study that is beyond the scope of this article. For the purpose of this study, we simply assume a common planar photon emitting surface for all channels.

3.1.2. Auxiliary Pseudo-Channels

[34] Occasionally it is useful to incorporate a priori information about a given scene into a so-called pseudo-channel. In terrestrial remote sensing, a common example is an elevation map that might help distinguish between spectrally similar land-cover types (e.g., an ice sheet near sea level, and an alpine glacier). In solar imagery it is useful to distinguish between off-disk and on-disk, so a pixel array that consisted of zeros for off-disk, and ones for on-disk might be included as an auxiliary image. We actually take this a step further by assuming an extremely simple slab-like model of the solar corona (i.e., a constant height above the solar surface), and calculating the path length through this volume for the look angle corresponding to each pixel. The resulting pseudo-channel possesses a well-defined solar disk, but also helps correct for limb effects that arise due to the fact that detected photons are the integration of photon emission through an optically thin plasma. This pseudo-channel is presented in Figure 3. For completeness we note that an empirical functional fit, similar to a technique used by *Harvey and White* [1999], was considered in order to remove limb effects from each channel. We decided that, in addition to the operationally prohibitive computational expense of implementing this in an operational setting, the technique would not work well with EUV images, where on-disk dynamic range can vary by many orders of magnitude, leading to sub-optimal fits.

3.2. Training and Test Data Selection

[35] The pixel classifier described here relies on expert judgment to assign the jth of a set of r predetermined theme labels to a subset of image pixels. This is considered a supervised classification scheme. We enlisted three solar imaging experts: one an image processing and statistics expert; one a solar physicist; and one a space weather forecaster. First they were consulted to help define a set of solar pixel labels that are considered both scientifically and operationally useful, and subsequently tasked to analyze two different solar events captured in the proxy EUV channels and assign thematic labels to subsets of pixels. The eight candidate pixel labels chosen for this study were: (1) Outer Space, (2) Coronal Hole (on-disk), (3) Coronal Hole (off-disk), (4) Quiet Corona (on-disk), (5) Quiet Corona (off-disk), (6) Active Region, (7) Prominence, and (8) Flare.

[36] Outer Space pixels are off-disk pixels that do not fit into one of the other categories. The on-disk/off-disk designation for Coronal Hole and Quiet Corona pixel labels was invented because initial studies demonstrated that more accurate classifications ensued when this a priori differentiation was implemented. On the other hand, the same initial studies suggested that this differentiation did not add much to the classification accuracy of Active Region, Prominence, or Flare pixels. The application of an automatic spectral clustering algorithm may well have eliminated the need for this design iteration [e.g., *Cheeseman and Stutz*, 1996], but since our tool is designed for space weather
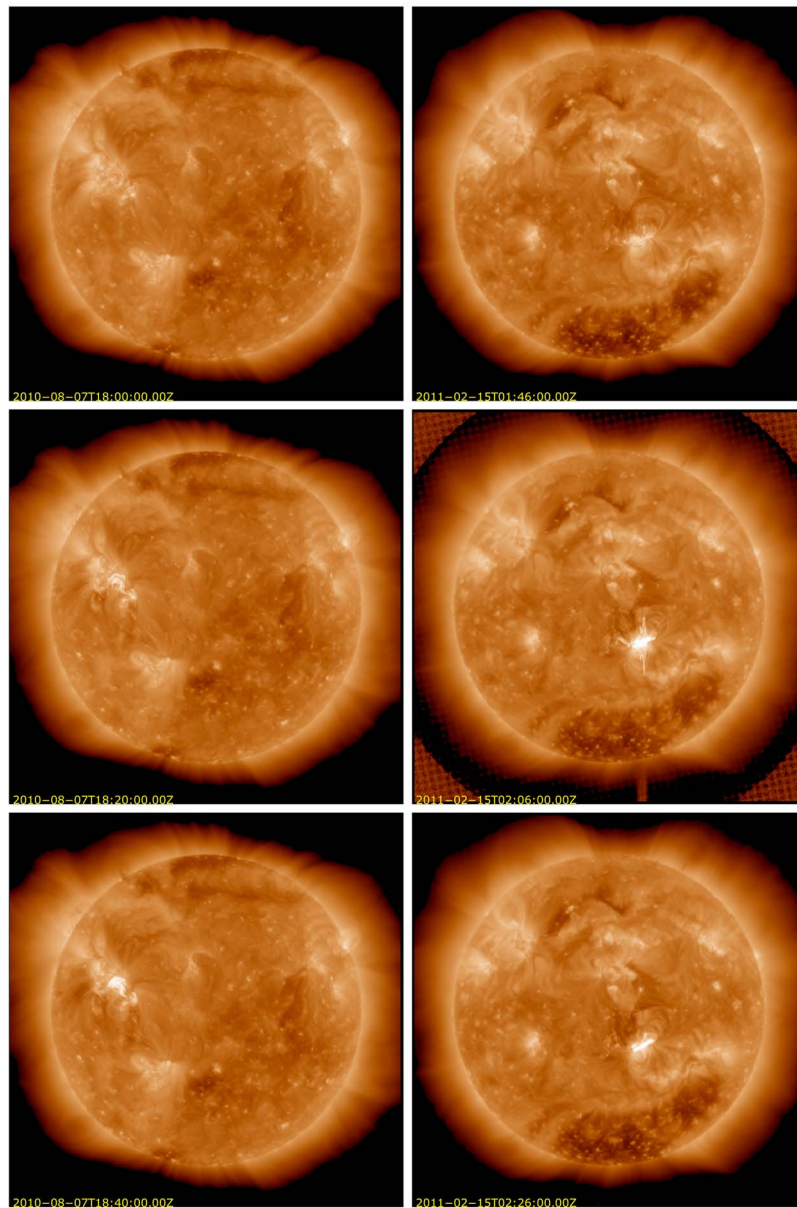
Figure 4. The SUVI proxy 193Å channel is shown for events on (left) 7 August 2010 and (right) 15 February 2011. Events are defined as (top) pre-, (middle) peak-, and (bottom) post-flare.

operations, it was deemed sensible to let categories be defined by those with appropriate operational expertise.

[37] Each expert was provided with this list of labels, a computerized graphical selection tool to assign labels to pixels, some basic explanatory material, and identical data sets that contained solar coronal images in the six EUV channels for two flaring events selected from the SWPC Edited Events Tables: an M-class flare on 7 August 2010; and an X-class flare on 15 February 2011. There were three nearly simultaneous sets of six images for each event: an

immediately pre-flare scene; a peak-flare scene; and a post-peak-flare scene (these designations are relative to X-ray event times listed in SWPC's Edited Events Tables for the dates in question, and may differ from EUV flare times). For brevity, these are shown in Figure 4 for the 193Å channel only. These scenes are presented for the other channels in the auxiliary material.[1]

---

[1]Auxiliary materials are available in the HTML. doi:10.1029/2012SW000780.

[38] Once the experts labeled a statistically significant number of pixels to generate theme-dependent means and covariances, a series of numerical experiments were conducted using different subsets of the data as training data and test data. The first experiment combined all the experts' data into a single training set, generated thematic maps, and finally compared the pixel classifications to the same data used to train the classifier. This "in-sample" comparison serves as a kind of upper bound on expected accuracy, and is primarily designed to highlight limitations related to algorithmic assumptions like distribution shape (i.e., are the pixel value distributions really Gaussian?), and statistical separability (i.e., are the pre-determined pixel classes actually distinguishable given the data provided). The next experiment generated pixel classifications using data from one expert for training, and data obtained from another expert for testing (pixel class statistics from the different events were merged prior to classification). The training and testing data were used to generate multivariate metrics for subsequent analysis. This comparison provides insight into the algorithm's sensitivity to variations in experts' opinions. The final experiment generated pixel classifications using one event's data for training, and the other event's data for testing (pixel class statistics from different experts were merged prior to classification). This comparison provides insight into the sensitivity of the algorithm to variations in solar conditions and/or instrument performance. Figure 5 shows color-coded thematic masks that correspond to the labels assigned by all experts after analyzing the training data scenes shown in Figure 4. Separate thematic masks generated by each expert are provided in the auxiliary material, but even these composite thematic masks illustrate a certain lack of consistency in how pixel labels are assigned by experts, a shortcoming that, as stated earlier, our supervised classification scheme is designed to help overcome.

4. Results and Discussion

4.1. Solar Thematic Maps

[39] The MAP thematic map solutions shown in Figure 6 were generated using input parameters $\alpha_j = 0$, and $\beta = 1$. Again for the sake of brevity, we only present maps generated from pixel theme statistics accumulated from all training data. This is because the thematic maps generated from different sets of training data only differ in rather small details that are best appreciated by looking at a more detailed error analysis derived from direct comparisons with test data (next subsection).

[40] Starting with the M-class flare event from 2010 (Figure 6, left), it is immediately apparent that the classifier does a good job reproducing the experts' differentiation between non-coronal (i.e., "outer space") and coronal pixels, off-disk open and closed field regions, and between on-disk and off-disk pixels in general. In fact one might argue that the latter is overly segmented, since near-limb

active region pixels seem to transition directly into prominence pixels in both the northeast and northwest sectors, when one might expect that the active region pixels would extend above the limb due to the 3D nature of the coronal loops that tend to define active regions in EUV wavelengths. Upon closer inspection of the original images however, it is clear in the 304Å channel that there really are significant prominences above the limb in these locations. At this point it is worth reminding ourselves that this does not mean that these pixels do not represent an active region, but rather given the constraint that only one class may be assigned to any given pixel, the algorithm has decided that a prominence is more likely above the limb.

[41] Moving onto the solar disk, it is clear that the algorithm has difficulty differentiating between coronal hole and filament (i.e., on-disk prominence) pixels. This is not surprising since even human observers struggle to differentiate between these two features in EUV solar imagery. Closer inspection of training data (not shown here) indicates a clear bias toward assigning dimmer low-latitude pixels the prominence/filament label by at least one expert. The blending of statistics from the different training data sets results in a classification ambiguity along the boundaries between coronal holes and quiet corona. It may prove valuable for the different experts to discuss their disagreement over these pixel classifications, and reconsider their training data selection for future thematic maps so that unnecessary confusion is avoided.

[42] Finally, we note that the M-class flaring region determined by the classifier covers more solar disk area than any expert might reasonably assign. This is a consequence of the algorithm's imperfect assumption that pixel value distributions are Gaussian. As a result, the algorithm struggles to place the decision boundary between active region and flaring pixels in a place that is intuitively satisfying to the experts. There are two ways to address this problem: 1) use a different algorithm that does not rely on such rigid assumptions about pixel value distributions; or 2) invent a spectrum of additional flaring pixel labels, each of which cover a more limited range of the brightness distribution. The former is not an option for our classifier at this time. The latter is an option, and is in fact quite similar, although somewhat more ad hoc, to *Turmon et al.* [2002, 2010].

[43] Similar statements may be made about the thematic maps for the X-class flare event (Figure 6, right) as were made about the M-class event. Additionally, the X-class flare event thematic map suffers from saturation effects in several of the EUV channels used as classifier inputs during the peak of the flare. This is apparent in the less well-defined boundaries of major regions of contiguous pixel labels, and a higher number of small isolated pixel label regions (mostly active region) randomly scattered throughout the quiet corona. While not shown graphically here, it turns out that higher pixel value variances that arose in certain X-class flare event channels tended to
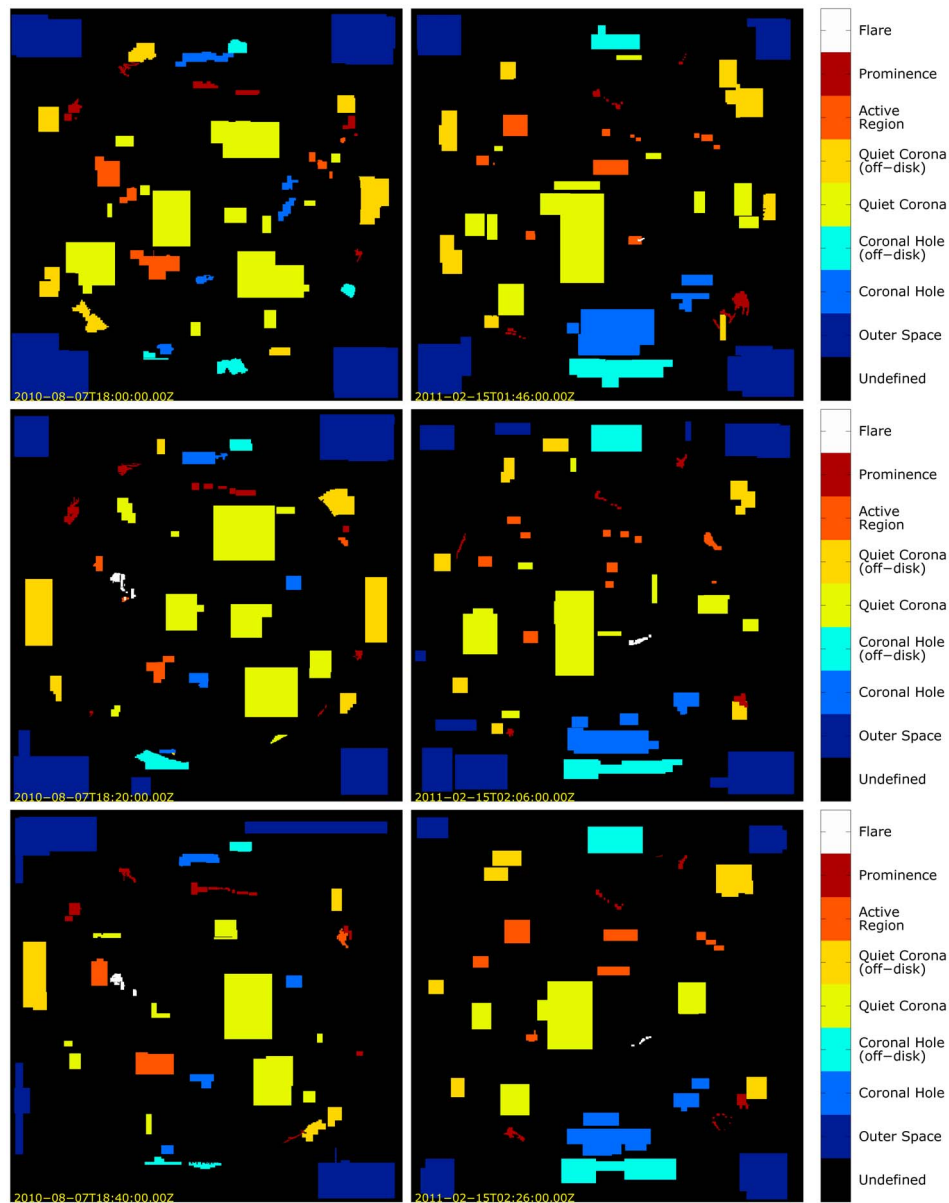
Figure 5. Manual thematic classifications for subsets of solar pixels are presented as discrete color labels. These classifications were determined by expert opinion using training data for events on (left) 7 August 2010 and (right) 15 February 2011. Events are defined as (top) pre-, (middle) peak-, and (bottom) post-flare.

cause substantial out-of-sample classification error, to the point that the resulting thematic maps looked rather non-physical. We discuss this result in more detail below.

4.2. Error Analysis

[44] Visual analysis of thematic maps is valuable, especially when ascertaining whether the maps match basic operational expectations for pixel classification. However, there is far too much information at the pixel-level for the human eye/mind to fully process and comprehend. Therefore we invoke a tool common to terrestrial image classification: the so-called confusion matrix [e.g., *Tso and Mather*, 2009, chapter 2]. This is typically a square matrix where the jth row or column corresponds to the jth pixel class. Thematic map pixel labels are compared to test data (i.e., a set of pixel classifications for the scene obtained independently from the classifier itself). If a thematic map and test pixel both match the jth pixel theme, the jth
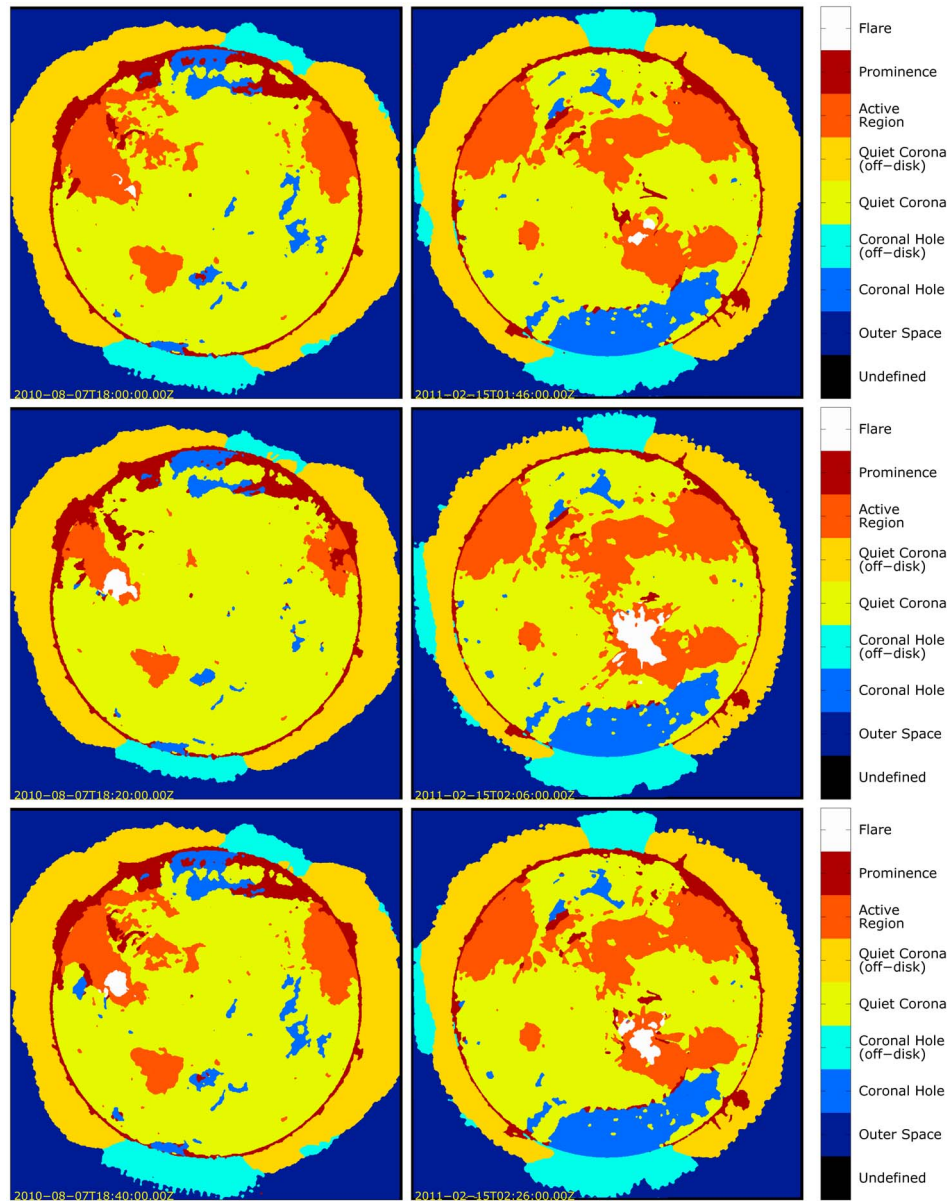
Figure 6. MAP thematic maps are presented using discrete color labels. These thematic maps were generated using multivariate statistics obtained from the manual thematic classifications collected from all experts for all events on (left) 7 August 2010 and (right) 15 February 2011. Events are defined as (top) pre-, (middle) peak-, and (bottom) post-flare.

diagonal element of the confusion matrix is incremented. If the thematic map pixel does not match the test pixel, the confusion matrix cell at the row corresponding to the test pixel label, and the column corresponding to the mis-classified thematic map pixel label, is incremented. This might also be thought of as a multioutput truth or contingency table.

[45] Once complete, diagonal elements of the confusion matrix contain the total number of true-positive classifications. The off-diagonal elements along a row represent the count of thematic map pixels that were labeled according to that row, but should have been labeled according to the column. This is a form of commission, or type I, error (i.e., false-positive). The off-diagonal elements of a column represent the count of thematic map pixels that should have been labeled according to that column, but were instead labeled according to the row. This is omission, or type II, error (i.e., false-negative). The

Table 1. Confusion Matrix: All Versus All (in Sample)[a]

Pixel Label	Outer Space	Coronal Hole	Coronal Hole (Off-Disk)	Quiet Corona	Quiet Corona (Off-Disk)	Active Region	Prominence	Flare
Outer space	**100**	0	0	0	0	0	0	0
Coronal hole	0	**91**	0	**1**	0	0	**25**	0
Coronal hole (off-disk)	0	0	**97**	0	0	0	0	0
Quiet corona	0	**7**	0	**98**	0	**2**	**1**	0
Quiet corona (off-disk)	0	0	**1**	0	**98**	0	**7**	0
Active region	0	0	0	**1**	0	**95**	0	**7**
Prominence	0	**2**	**1**	0	**2**	**1**	**67**	0
Flare	0	0	0	0	0	**2**	0	**93**

[a]Classification accuracies (diagonal), commission errors (rows), and omission errors (columns), are all specified as percentages of the number of pixels manually labeled by experts. The multiclass accuracy κ is still calculated from raw counts, not percentages. "In sample" implies classifier output was tested against its own training data. $\kappa = 0.960$. Bold indicates non-zero values.

sum of each column is the total number of corresponding pixel labels in the test data; the sum of each row is the total number of labels from the thematic map at the test data pixel positions. Independently generated confusion matrices may be merged by simple element-by-element summation of pixel labels.

[46] Finally, while per-theme classification statistics and inter-theme comparisons are important, it is also valuable to ascertain the overall accuracy of the classifier, accounting for type I and II errors across multiple possible themes. A common metric for this purpose is Cohen's κ-coefficient (sometimes cited phonetically as "kappa-coefficient") [e.g., *Jones et al.*, 2008]:

$$\kappa = \frac{N \sum_{j=1}^{r} x_{jj} - \sum_{j=1}^{r} \left(x_{j+} \cdot x_{+j} \right)}{N^2 - \sum_{j=1}^{r} \left(x_{j+} \cdot x_{+j} \right)}. \qquad (12)$$

[47] Here, r is the number of columns in the confusion matrix, x_{jj} is the number of correct classifications for theme j, and x_{j+} and x_{+j} are the row and column totals for each theme, respectively. N is the total number of pixels being compared. If all automatic classifications match the test data $\kappa = 1$, while $\kappa = 0$ implies that the automatic classification was no better than a random assignment of labels to pixels. The confusion matrices in the following tables display classification accuracy as percentages instead of pixel

counts. This facilitates inter-theme comparisons, but due to round-off error, these percentages do not always sum to 100. However, the κ-coefficient presented below each confusion matrix is still calculated from actual pixel counts, not percentages.

[48] Table 1 summarizes the accuracy of the classifier when all available data is used for training and testing. Not surprisingly for in-sample validation (comparison of classifier output with its own training data), most theme-specific accuracies are quite high, as is κ. For the most part, theme-specific accuracies deviate from 100% only at transitional boundaries between similar pixel types (e.g., quiet corona to active region, or active region to flare). This suggests that our assumptions of normally distributed pixel values were appropriate, and for the training data at least, that these theme-dependent distributions were separable. A noticeable exception to this is the classifier's apparent inability to classify prominences (accuracy only ~67%). In fact, it appears that the classifier struggles to distinguish prominences from on-disk coronal holes, a solar feature that has no known physical reason to be spatially associated with prominences.

[49] One might be tempted to simply attribute this higher degree of misclassification to the fact that both these solar features tend to manifest as dimmer on-disk regions in the EUV, and that this was nothing more than an example of poor statistical separability. However, a quick glance at Tables 2 and 3 shows us that the average in-sample accuracy for prominence classification is

Table 2. Confusion Matrix: Expert Versus Expert (in Sample)[a]

Pixel Label	Outer Space	Coronal Hole	Coronal Hole (Off-Disk)	Quiet Corona	Quiet Corona (Off-Disk)	Active Region	Prominence	Flare
Outer space	**100**	0	0	0	0	0	0	0
Coronal hole	0	**93**	0	**1**	0	0	**3**	0
Coronal hole (off-disk)	0	0	**98**	0	0	0	0	0
Quiet corona	0	**5**	0	**98**	0	**1**	**1**	0
Quiet corona (off-disk)	0	0	**1**	0	**98**	0	**5**	0
Active region	0	0	0	**1**	0	**96**	0	**6**
Prominence	0	**2**	0	0	**2**	0	**92**	0
Flare	0	0	0	0	0	**2**	0	**94**

[a]Classification accuracies (diagonal), commission errors (rows), and omission errors (columns), are all specified as percentages of the number of pixels manually labeled by experts. The multiclass accuracy κ is still calculated from raw counts, not percentages. "In sample" implies classifier output was tested against its own training data. $\kappa = 0.978$. Bold indicates non-zero values.

Table 3. Confusion Matrix: Expert Versus Expert (out of Sample)[a]

Pixel Label	Outer Space	Coronal Hole	Coronal Hole (Off-Disk)	Quiet Corona	Quiet Corona (Off-Disk)	Active Region	Prominence	Flare
Outer space	**100**	0	**3**	0	**7**	0	0	0
Coronal hole	0	**80**	0	0	0	0	**23**	0
Coronal hole (off-disk)	0	0	**95**	0	**1**	0	0	0
Quiet corona	0	**14**	0	**97**	0	**5**	**5**	0
Quiet corona (off-disk)	0	0	**1**	0	**91**	0	**31**	0
Active region	0	0	0	**2**	0	**92**	**1**	**9**
Prominence	0	**6**	**1**	0	**2**	0	**41**	0
Flare	0	0	0	0	0	**2**	0	**91**

[a]Classification accuracies (diagonal), commission errors (rows), and omission errors (columns), are all specified as percentages of the number of pixels manually labeled by experts. The multiclass accuracy κ is still calculated from raw counts, not percentages. "Out of sample" implies classifier output was tested against independent data. $\kappa = 0.925$. Bold indicates non-zero values.

actually fairly high (∼92%), while the out-of-sample accuracy (accuracy when classifier output is compared to independent test data) is extremely low (∼41%). This is a tell-tale sign of major discrepancies between each expert's training data. Upon closer inspection of the individual experts' training/testing data it was ascertained that one expert tended to label certain dim on-disk pixels as filaments (i.e., on-disk prominences), while others simply left these pixels unlabeled. It turns out that such selection bias shows up in the confusion matrices as an imbalance in classification accuracy between two labels (e.g., 91% correct coronal hole classification versus 67% correct prominence classification), rather than both labels exhibiting similarly reduced accuracies. Upon further examination of Tables 2 and 3, as well as closer inspection of each expert's training data labels, a similar difference of opinion appears to exist regarding prominences and off-disk quiet corona. This discrepancy was not so easily discerned from the all-data, in-sample comparisons shown in Table 1, so it would seem that separate and independent analysis of the experts' pixel classifications was justified.

[50] Tables 4 and 5 describe the classifier's performance when the different experts' opinions were merged, but two independent space weather events were analyzed. The in-sample analysis looks similar to the in-sample analysis for all events in Table 1. This indicates that whatever disagreements between the experts may exist, they are

consistent from event to event. More interesting from an operational perspective are the out-of-sample comparisons. The multitheme accuracy κ drops considerably; much more than it did when different experts were compared out-of-sample. This implies that, while the experts are generally in good agreement about solar pixel labels, something about the theme-dependent pixel value distributions changes when different events are analyzed, leading to higher misclassification rates.

[51] The most significant change from the all-data in-sample comparison is in the flare classification. A large fraction of flare pixels identified by experts for one event tend to get misclassified as active region pixels when classified using training data from the other event. In fact, if individual out-of-sample comparisons are studied (not shown here), nearly 100% of the pixels that should have been identified as flare pixels in the 2010 event were misclassified as active region. This misclassification is not very symmetric (i.e., active region pixels do not get misclassified as flares very often), so again, this cannot be dismissed as a lack of statistical separability. A closer inspection of the individual events, and their respective training data, shows that the flare-dependent mean values are much higher for the 2011 event. This should not be surprising since this was an X-class flare, compared to the weaker M-class flare in the 2010 event. This extreme misclassification rate offers justification for including

Table 4. Confusion Matrix: Event Versus Event (in Sample)[a]

Pixel Label	Outer Space	Coronal Hole	Coronal Hole (Off-Disk)	Quiet Corona	Quiet Corona (Off-Disk)	Active Region	Prominence	Flare
Outer space	**100**	0	0	0	0	0	0	0
Coronal hole	0	**93**	0	0	0	0	**28**	0
Coronal hole (off-disk)	0	0	**97**	0	0	0	0	0
Quiet corona	0	**6**	0	**98**	0	**1**	**1**	0
Quiet corona (off-disk)	0	0	**1**	0	**98**	0	**6**	0
Active region	0	0	0	**1**	0	**96**	0	**3**
Prominence	0	**1**	**1**	0	**2**	**1**	**64**	0
Flare	0	0	0	0	0	**2**	0	**97**

[a]Classification accuracies (diagonal), commission errors (rows), and omission errors (columns), are all specified as percentages of the number of pixels manually labeled by experts. The multiclass accuracy κ is still calculated from raw counts, not percentages. "In sample" implies classifier output was tested against its own training data. $\kappa = 0.965$. Bold indicates non-zero values.

Table 5. Confusion Matrix: Event Versus Event (out of Sample)[a]

Pixel Label	Outer Space	Coronal Hole	Coronal Hole (Off-Disk)	Quiet Corona	Quiet Corona (Off-Disk)	Active Region	Prominence	Flare
Outer space	**95**	0	**14**	0	**18**	0	**1**	0
Coronal hole	0	**89**	0	**17**	0	0	**26**	0
Coronal hole (off-disk)	0	0	**84**	0	**6**	0	**1**	0
Quiet corona	0	**6**	0	**71**	0	**1**	**1**	0
Quiet corona (off-disk)	**1**	0	**1**	0	**75**	**1**	**26**	0
Active region	**4**	0	0	**9**	0	**98**	0	**75**
Prominence	0	**5**	**1**	**1**	**2**	0	**45**	0
Flare	0	0	0	**1**	0	0	0	**25**

[a]Classification accuracies (diagonal), commission errors (rows), and omission errors (columns), are all specified as percentages of the number of pixels manually labeled by experts. The multiclass accuracy κ is still calculated from raw counts, not percentages. "Out of sample" implies classifier output was tested against independent data. $\kappa = 0.783$. Bold indicates non-zero values.

additional labels that each correspond to a different class of solar flare event.

[52] Another point of interest is the apparent confusion between outer space pixels and several other classes where no such confusion existed previously. This was difficult to explain until a careful examination of the input images showed that there was a tremendous amount of photon counting noise in a single frame of the 94Å channel for the 2011 event. This does reduce the statistical separability of outer space pixels with several other classes.

5. Summary and Conclusions

[53] In summary, a Bayesian solar pixel classifier designed to emulate a range of expert opinions was described, validated using independent test data, and while its general performance already meets or exceeds many of our initial operational requirements, several shortcomings were identified that can be easily fixed without changing the underlying classification algorithm. Once previously noted discrepancies between the experts' opinions on certain pixel types are reconciled, perhaps the improvement to consider next is adding more expert opinions and events for training data. This should improve out-of-sample comparisons significantly given the limited set of training data considered for this study.

[54] At some point it becomes unwieldy to maintain a database of every training pixel, from every expert, for every event. It is more efficient to characterize each training set by its respective theme-dependent statistics (means, covariances, and pixel counts for each theme), then blend these in some optimal fashion. In fact we did this using standard mixture reduction (SMR), with theme-dependent pixel counts as weights. One drawback to this approach is that it gives equal credibility to each expert, or worse, allows a potentially unreliable expert to unduly influence the theme-dependent statistics by simply selecting more training pixels. Expert credibility might be described and applied using relative weights, but this generally leads to inconsistent theme statistics when SMR is used. *Reece and Roberts* [2010] propose a solution they call generalized

covariance union (GCU) which promises to reliably generate consistent statistics.

[55] While the event versus event comparison focused mainly on the two themes with the largest out-of-sample performance decreases, there is clearly an overall decrease in out-of-sample performance that would seem to suggest a time-varying aspect of the theme-dependent statistics. This may be due to genuine nonstationarity in the solar scene being observed, or possibly even changes in instrument performance. With this in mind, it might prove beneficial to develop a recursive algorithm to blend training data that employs some sort of forgetting factor that weighed recent training results more than old training results. This could be accomplished in an ad hoc fashion by simply re-scaling the weights used in SMR or GCU by some function of time-since-last-update. However, the well-known Kalman Filter, or one of its many variants, is specifically designed to do this in an optimal manner, and is generally considered engineering "best practice" for time-varying estimation problems.

[56] Also, a lack of time-stationarity is often a symptom of an insufficient number of independent inputs. *Rigler et al.* [2007] demonstrated this concept by revisiting the problem of radiation belt predictions using multi-input linear filters. For the problem at hand, namely solar image pixel classification, we are clearly not considering all the relevant inputs. Part of this is a practical consideration: NOAA requires that operational data products rely only on operational NOAA assets, like SUVI. However, even if non-operational data cannot be exploited to improve our pixel classifications, we can transform existing operational data to extract more information. One promising possibility is to use difference images in addition to the snapshots used in this study.

[57] Finally, while every pixel analyzed in the above results was assigned one of the preselected labels, it may be desirable that some pixels be considered unclassifiable. This might be for as simple a reason as a single input channel's pixel being flagged "bad," in which case the probability of membership to all classes should be set to zero. More useful still would be the ability to specify a

minimum acceptable probability for each pixel theme, preferably one based on the training data and robust statistics. The square root of the scalar product of the matrix multiplication in (9) or (10), or the Mahalanobis distance (MD), follows a chi-square distribution for ρ degrees of freedom. A threshold MD could be calculated given a specified critical value (e.g., 0.95, 0.99, etc.). If this threshold is not met, the conditional probability is simply set to zero, and if all conditional probabilities equal zero, the thematic map pixel is labeled unclassifiable. This approach may suffer from a well-known "masking effect," whereby training data outliers tend to bias or inflate theme-dependent statistics such that otherwise obvious outliers do not exceed the threshold MD. *Rousseeuw and van Zomeren* [1990] propose techniques for generating robust theme-dependent statistics, and thereby a robust distance (RD) that should largely mitigate this problem.

[58] These improvements will serve to standardize solar pixel classifications, increase thematic map accuracy, and ensure compatibility with hand-drawn synoptic maps, which play an essential role in SWPC's daily forecasts by integrating a variety of information into a coherent picture of the solar environment. Future research will exploit these improvements to investigate solar cycle effects, test better ways to categorize flares and other solar features that may not fit the statistical assumptions used here, and maybe even incorporate additional non-EUV channels to better differentiate certain solar features like filaments. This will all be done in close collaboration with SWPC so that the eventual real time implementation of this solar thematic mapper can be tailored to their operational needs. Perhaps more importantly, SWPC forecasters will grow to understand more fully the strengths and limitations of this tool as a key component of their daily forecasts, and ultimately extend and enhance its capabilities by contributing their own expertise to the refinement of its theme-dependent statistics.

References

Ainsworth, E. J., and I. S. F. Jones (1999), Radiance spectra classification from the ocean color and temperature scanner on ADEOS, *IEEE Trans. Geosci. Remote Sens., 37*(3), 1645–1656, doi:10.1109/36.763281.

Andréfouët, S., et al. (2003), Multi-site evaluation of IKONOS data for classification of tropical coral reef environments, *Remote Sens. Environ., 88*(1–2), 128–143, doi:10.1016/j.rse.2003.04.005.

Aschwanden, M. J. (2010), Image processing techniques and feature recognition in solar physics, *Sol. Phys., 262*, 235–275.

Bannari, A., et al. (2006), Estimating and mapping crop residues cover on agricultural lands using hyperspectral and IKONOS data, *Remote Sens. Environ., 104*(4), 447–459, doi:10.1016/j.rse.2006.05.018.

Barra, V., V. Delouille, M. Kretzschmar, and J.-F. Hochedez (2009), Fast and robust segmentation of solar EUV images: Algorithm and results for solar cycle 23, *Astron. Astrophys., 505*, 361–371, doi:10.1051/0004-6361/200811416.

Besag, J. (1986), On the statistical analysis of dirty pictures, *J. R. Stat. Soc., 48*(3), 259–302.

Boykov, Y., and V. Kolmogorov (2004), An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *IEEE Trans. Pattern Anal. Mach. Intell., 26*(9), 1124–1137, doi:10.1109/TPAMI.2004.60.

Cardillo, G., and D. Landgrebe (1966), On pattern recognition, *LARS Tech. Note, 101866*, Purdue Univ., West Lafayette, Indiana.

Cheeseman, P., and J. Stutz (1996), Bayesian classification (Auto-Class): Theory and results, in *Advances in Knowledge Discovery and Data Mining*, edited by U. M. Fayyad et al., pp. 153–180, Am. Assoc. for Artificial Intell., Menlo Park, Calif.

de Wit, T. D. (2006), Fast segmentation of solar extreme ultraviolet images, *Sol. Phys., 239*(1–2), 519–530, doi:10.1007/s11207-006-0140-3.

Friedl, M. A., et al. (2002), Global land cover mapping from MODIS: Algorithms and early results, *Remote Sens. Environ., 83*(1–2), 287–302, doi:10.1016/S0034-4257(02)00078-0.

Gallagher, P., D. Berghmans, and M. Aschwanden (Eds.) (2005), *Solar Image Processing, Sol. Phys., 228*(1–2).

Harvey, K. L., and O. R. White (1999), Magnetic and radiative variability of solar surface structures. I. Image decomposition and magnetic-intensity mapping, *Astrophys. J., 515*, 812–831, doi:10.1086/307035.

Higham, N. J. (2002), Computing the nearest correlation matrix—A problem from finance, *IMA J. Numer. Anal., 22*, 329–343, doi:10.1093/imanum/22.3.329.

Homer, C., et al. (2004), Development of a 2001 national land-cover database for the United States, *Photogramm. Eng. Remote Sens., 70*(7), 829–840.

Inhester, B., L. Feng, and T. Wiegelmann (2008), Segmentation of loops from coronal EUV images, *Sol. Phys., 248*, 379–393, doi:10.1007/s11207-007-9027-1.

Ireland, J., and C. A. Young (Eds.) (2010), *Solar Image Processing and Analysis, Sol. Phys., 262*(2).

Jones, H. P., G. A. Chapman, K. L. Harvey, J. M. Pap, D. G. Preminger, M. J. Turmon, and S. R. Walton (2008), A comparison of feature classification methods for modeling solar irradiance variation, *Sol. Phys., 248*(2), 323–337, doi:10.1007/s11207-007-9069-4.

Karvonen, J. A. (2004), Baltic Sea ice SAR segmentation and classification using modified pulse-coupled neural networks, *IEEE Trans. Geosci. Remote Sens., 42*(7), 1566–1574, doi:10.1109/TGRS.2004.828179.

Krista, L. D., and P. T. Gallagher (2009), Automated coronal hole detection using local intensity thresholding techniques, *Sol. Phys., 256*, 87–100.

LaBonte, B. J., D. M. Rust, and P. N. Bernasconi (2003), An automated system for detecting sigmoids in solar X-ray images, *Bull. Am. Astron. Soc., 35*, 814.

Lee, Y., G. Wahba, and S. A. Ackerman (2004), Cloud classification of satellite radiance data by multicategory support vector machines, *J. Atmos. Oceanic Technol., 21*(2), 159–169, doi:10.1175/1520-0426(2004)021<0159:CCOSRD>2.0.CO;2.

Lemen, J. R., et al. (2012), The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO), *Sol. Phys., 275*, 17–40, doi:10.1007/s11207-011-9776-8.

Li, S. Z. (2009), Mathematical MRF models, in *Markov Random Field Modeling in Image Analysis*, 3rd ed., pp. 21–48, Springer, London.

Martens, P. C. H., et al. (2012), Computer vision for the Solar Dynamics Observatory (SDO), *Sol. Phys., 275*, 79–113, doi:10.1007/s11207-010-9697-y.

Martinez-Galarce, D., et al. (2010), A novel forward-model technique for estimating EUV imaging performance: Design and analysis of the SUVI telescope, *Proc. SPIE, 7732*, 773237.

McIntosh, P. S. (1972), Solar magnetic fields derived from hydrogen alpha filtergrams, *Rev. Geophys., 10*(3), 837–846, doi:10.1029/RG010i003p00837.

Nair, U. S., et al. (1999), Detection of cumulus cloud fields in satellite imagery, *Proc. SPIE, 3750*, 345–355, doi:10.1117/12.363530.

Owen, A. (1989), Image segmentation via iterated conditional expectations, *SIMS Tech. Rep., 122*, Dep. of Stat., Stanford Univ., Stanford, Calif.

Platnick, S., et al. (2003), The MODIS cloud products: Algorithms and examples from Terra, *IEEE Trans. Geosci. Remote Sens.*, 41(2), 459–473, doi:10.1109/TGRS.2002.808301.

Preminger, D. G., S. R. Walton, and G. A. Chapman (2001), Solar feature identification using contrasts and contiguity, *Sol. Phys.*, 202(1), 53–62, doi:10.1023/A:1011896413891.

Qu, M., F. Shih, J. Jing, and H. Wang (2005), Automatic solar filament detection using image processing techniques, *Sol. Phys.*, 228, 119–135, doi:10.1007/s11207-005-5780-1.

Reece, S., and S. Roberts (2010), Generalized covariance union: A unified approach to hypothesis merging in tracking, *IEEE Trans. Aerosp. Electron. Syst.*, 46(1), 207–221, doi:10.1109/TAES.2010.5417157.

Research Systems Inc. (2004), *ENVI User's Guide, Version 4.1*, 1150 pp., Boulder, Colo.

Richards, J. A. (2005), Analysis of remotely sensed data: The formative decades and the future, *IEEE Trans. Geosci. Remote Sens.*, 43(3), 422–432, doi:10.1109/TGRS.2004.837326.

Rigler, E. J., M. Wiltberger, and D. N. Baker (2007), Radiation belt electrons respond to multiple solar wind inputs, *J. Geophys. Res.*, 112, A06208, doi:10.1029/2006JA012181.

Rousseeuw, P. J., and B. C. van Zomeren (1990), Unmasking multivariate outliers and leverage points, *J. Am. Stat. Assoc.*, 85(411), 633–639.

Simpson, J. J., and R. H. Keller (1995), An improved fuzzy logic segmentation of sea ice, clouds, and ocean in remotely sensed Arctic imagery, *Remote Sens. Environ.*, 54(3), 290–312, doi:10.1016/0034-4257(95)00175-1.

Timothy, A. F., A. S. Krieger, and G. S. Vaiana (1975), The structure and evolution of coronal holes, *Sol. Phys.*, 42, 135–156, doi:10.1007/BF00153291.

Tso, B., and P. M. Mather (2009), *Classification Methods for Remotely Sensed Data*, 2nd ed., CRC Press, Boca Raton, Fla., doi:10.1201/9781420090741.

Turmon, M., J. M. Pap, and S. Mukhtar (2002), Statistical pattern recognition for labeling solar active regions: Application to SOHO/MDI imagery, *Astrophys. J.*, 568(1), 396–407, doi:10.1086/338681.

Turmon, M., H. P. Jones, O. V. Malanushenko, and J. M. Pap (2010), Statistical feature recognition for multidimensional solar imagery, *Sol. Phys.*, 262(2), 277–298, doi:10.1007/s11207-009-9490-y.

Vaiana, G. S., A. S. Krieger, and A. F. Timothy (1973), Identification and analysis of structures in the corona from x-ray photography, *Sol. Phys.*, 32, 81–116, doi:10.1007/BF00152731.

Wilkinson, G. G. (2005), Results and implications of a study of fifteen years of satellite image classification experiments, *IEEE Trans. Geosci. Remote Sens.*, 43(3), 433–440, doi:10.1109/TGRS.2004.837325.

Worden, J. R., O. R. White, and T. N. Woods (1998), Evolution of chromospheric structures derived from Ca II K spectroheliograms: Implications for solar ultraviolet irradiance variability, *Astrophys. J.*, 496(2), 998–1014, doi:10.1086/305392.

Young, C. A., and J. Ireland (Eds.) (2008), *Solar Image Analysis and Visualization, Sol. Phys.*, 248(2).