Solar Phys (2011) 268: 389–406
DOI 10.1007/s11207-010-9552-1

HELIOSEISMOLOGY

# Subsurface Vorticity of Flaring *versus* Flare-Quiet Active Regions

**R. Komm · R. Ferguson · F. Hill · G. Barnes · K.D. Leka**

Received: 8 January 2010 / Accepted: 29 March 2010 / Published online: 30 April 2010


**Abstract** We apply discriminant analysis to 1023 active regions and their subsurface-flow parameters, such as vorticity and kinetic helicity density, with the goal of distinguishing between flaring and non-flaring active regions. We derive synoptic subsurface flows by analyzing GONG high-resolution Doppler data with ring-diagram analysis. We include magnetic-flux values in the discriminant analysis derived from NSO Kitt Peak and SOLIS synoptic maps binned to the same spatial scale as the helioseismic analysis. For each active region, we determine the flare information from GOES and include all flares within 60° central meridian distance to match the coverage of the ring-diagram analysis. The subsurface-flow characteristics improve the ability to distinguish between flaring and non-flaring active regions. For the C- and M-class flare category, the most important subsurface parameter is the so-called structure vorticity, which estimates the horizontal gradient of the horizontal-vorticity components. The no-event skill score, which measures the improvement over predicting that no events occur, reaches 0.48 for C-class flares and 0.32 for M-class flares, when the structure vorticity at three depths combined with total magnetic flux are used. The contributions come mainly from shallow layers within about 2 Mm of the surface and layers deeper than about 7 Mm.

**Keywords** Active regions, velocity field · Helioseismology, observations · Interior, convective zone · Velocity fields, interior · Flares

Helioseismology
Guest Editors: G. Houdek, H. Shibahashi, and J. Zhao

R. Komm (✉) · R. Ferguson · F. Hill
National Solar Observatory, Tucson, AZ 85719, USA
e-mail: komm@nso.edu

R. Ferguson
Michigan State University, East Lansing, MI 48824, USA

G. Barnes · K.D. Leka
NorthWest Research Associates, Inc., CoRA, Boulder, CO 80301, USA

## 1. Introduction

Previous studies have shown that the flare activity of active regions is associated with the vorticity of subsurface flows on temporal and spatial scales comparable to the size and lifetime of active regions (Ferguson *et al.*, 2009; Komm and Hill, 2009; Mason *et al.*, 2006). In this study, we further explore the relation between surface magnetic flux of the Sun and solar subsurface-flow vorticity for flaring and non-flaring active regions. We analyze synoptic maps focusing on the average behavior of active regions. Highly non-potential magnetic fields are very probably responsible for strong eruptive phenomena such as flares and coronal mass ejections (CME). Many new parameters based on surface magnetic-field measurements are considered as flare-forecasting criteria (Abramenko, 2005; McAteer, Gallagher, and Ireland, 2005; Cui *et al.*, 2006; Georgoulis and Rust, 2007; Leka and Barnes, 2007; Schrijver, 2007). Barnes and Leka (2008) made a quantitative comparison of these results by applying the published techniques to a common data set. They found that these parameters show only modest improvements over climatological or uniform forecasts when comparing skill scores to account for the low rate of flare events. Falconer *et al.* (2009) found that flare-productive active regions are located on a straight line in a phase-space diagram of the total free energy of the magnetic field and the total magnetic flux, and they concluded that the input of free energy via convection in and below the photosphere balances the loss of free energy via CMEs and flares.

Flows below the solar surface are accessible with the techniques of local helioseismology, which are currently the only way to measure anything related to active regions below the solar surface (see Gizon and Birch, 2005, for a review). The twist of subsurface flows might serve as a proxy for the twist of magnetic-flux tubes below the solar surface because either the flows have to respond to the presence of magnetic fields or the fields are being pushed and twisted by subsurface turbulent flows. A plausible physical connection between subsurface flows and flare activity is that the non-potential nature of flux tubes results from the kinetic helicity of turbulent subsurface flows, as studied in numerical models by Longcope, Fisher, and Pevtsov (1998).

The question remains whether the measured vorticity of subsurface flows contains sufficient additional information that can be used to identify potentially dangerous active regions. In this study, we apply statistical tests based on linear discriminant analysis (Kendall, Stuart, and Ord, 1983) to several subsurface-flow parameters with the goal of differentiating between flaring and non-flaring active regions. Some preliminary results of this analysis have been reported by Ferguson *et al.* (2009). Other simple tests have been performed on the same data set (Komm and Hill, 2009). The vorticity of subsurface flows has been studied before (Braun, Birch, and Lindsey, 2004; Komm *et al.*, 2004; Zhao and Kosovichev, 2004), but Mason *et al.* (2006) performed the first large statistical survey of flare-producing active regions and their subsurface plasma-flow complexity. The data set now consists of 1023 active regions and their subsurface flow measurements derived from Global Oscillation Network Group (GONG) data using ring-diagram analysis. This data set includes flaring as well as non-flaring active regions. The flux values were derived from NSO Kitt Peak and SOLIS synoptic maps and the level of flare activity is characterized by individual X-ray flare classes reported from the *Geostationary Operational Environmental Satellite* (GOES). In this study, we focus on synoptic flow parameters and thus on the average behavior of active regions. We investigate how well single subsurface parameters work for classifying flaring and non-flaring active regions, and which subsurface layers work best, with a goal of determining an optimum set of subsurface parameters.

## 2. Data and Analysis

### 2.1. Subsurface Flow Measurements

We analyze observations obtained during 81 Carrington Rotations CR 1979 – 2059 (27 July 2001 – 13 August 2007) for which we have high-resolution full-disk Doppler data from the GONG network (http://gong.nso.edu/data). We determine the horizontal components of solar subsurface flows with a ring-diagram analysis using the dense-pack technique (Haber *et al.*, 2002) adapted to GONG++ data (Corbard *et al.*, 2003). The full-disk Doppler images are divided into 189 overlapping regions with centers spaced by 7.5° ranging over ±52.5° in latitude and central meridian distance (CMD). Each region is apodized with a circular function reducing the effective diameter to 15° before calculating three-dimensional power spectra. We derive daily flow maps of horizontal velocities from the 189 dense-pack patches. For this study, we combine them to form synoptic flow maps at 16 depths from 0.6 to 16 Mm. We also estimate the vertical-velocity component from the divergence of horizontal flows using mass conservation (Komm *et al.*, 2004; Komm, 2007). To focus on the spatial variation of the flows, we remove the large-scale trends in latitude of the differential rotation and the meridional flows and calculate residual synoptic flow maps by subtracting a low-order polynomial fit in latitude of the longitudinal average of the flows for each Carrington rotation (Komm *et al.*, 2004, 2005).

In this study, we examine vorticity [$\boldsymbol{\omega}$] which is a vector quantity and kinetic helicity density [$h$] which is a scalar that correspond to changing orientation in space of fluid particles and are thus quantities associated with mixing and turbulence (see Lesieur, 1987, for example). The vorticity vector is defined as the curl of the velocity vector [$\boldsymbol{v}$]:

$$\boldsymbol{\omega} = \nabla \times \boldsymbol{v} \tag{1}$$

and the kinetic helicity density is defined as the scalar product of the velocity and vorticity vector:

$$h = \boldsymbol{\omega} \cdot \boldsymbol{v}. \tag{2}$$

Komm (2007) discusses these topics in more detail. In maps of flow vorticity, most active regions show a "dipolar" pattern in the zonal and meridional vorticity components (Komm and Hill, 2009). These patterns are most pronounced in strong active regions that produce flares. We quantify this dipolar pattern for each active region and derive a so-called structure component of vorticity [$\omega_s$] (Mason *et al.*, 2006). For each active region, a swath covering two dense-pack regions is used along the anticipated vorticity dipole for the zonal and meridional components. For the zonal-vorticity component, the minimum value from the southern half of the swath is subtracted from the maximum value in the northern half. For the meridional-vorticity component, the minimum value from the eastern half is subtracted from the maximum value in the western. The resulting differences in $\omega_x$ and $\omega_y$ are added to obtain $\omega_s$. If the flows within an active region show a dipolar structure of this orientation, the value of $\omega_s$ will be of the same order as the sum of $\omega_x$ and $\omega_y$. Negative values indicate that the direction of the dipole is reversed and values close to zero indicate that there is no such pattern. This quantity can be interpreted as the magnitude of the gradient of the horizontal-vorticity components across an active region on the dense-pack grid. The vertical contribution to the helicity scalar [$h_z$] is also of interest since it might be related to the footpoint motion of magnetic-flux tubes. We include $h_z$ separately from the helicity because the helicity is dominated by the other two contributions when derived on dense-pack scales (Komm, 2007).

As a measure of solar activity, we use the NSO Kitt Peak synoptic maps (http://nsokp.nso.edu/dataarch.html) and the NSO SOLIS synoptic maps (http://solis.nso.edu/solis_data.html). The SOLIS synoptic maps have been calibrated to ensure compatibility with Kitt Peak synoptic maps (Jones *et al.*, 2004; Henney and SOLIS Team, 2007). We convert the magnetogram data to absolute values and bin them into circular areas with 15° diameter centered on a grid with 7.5° spacing in latitude and longitude to match the dense-pack mosaic. As in Mason *et al.* (2006), active regions are identified in magnetograms using the NOAA active region numbers listed in the Active Region Monitor at NASA Goddard Space Flight Center's Solar Data Analysis Center (SDAC) (http://www.solarmonitor.org). Circular masks are drawn manually on the magnetic synoptic maps to fully enclose each active region. Applying these masks to the synoptic maps of flux, vorticity, and kinetic helicity density, we determine the total and the maximum values of each quantity using the values on all dense-pack grid points within each mask. The maximum values might not be robust due to fluctuations produced by noise. However, averaging used to create synoptic maps should reduce the influence of noise on the results. The X-ray flare activity of each active region during its disk passage is determined using data from the *Geostationary Operational Environmental Satellite* (GOES) (http://www.ngdc.noaa.gov/stp/SOLAR/ftpsolarflares.html#xray). We include all flares within 60° CMD to match the coverage of the dense-pack analysis.

Our data sets consist of a total of 1023 active regions with 521 regions (50.9%) producing at least C-class flares. More energetic flares are rare by comparison; 177 regions (17.3%) produce M-class or stronger flares and only 30 regions (2.9%) produce X-class flares. A given active region might produce flares in all three categories. For each active region, we analyze the maximum and total values of the three components of the vorticity vector $[\omega_x, \omega_y, \omega_z]$, the size of the vorticity, as represented by the enstrophy $[\omega^2]$, and the structure vorticity $[\omega_s]$. In addition, we use the helicity scalar $[h]$ and the vertical contribution to helicity $[h_z]$.

### 2.2. Discriminant Analysis

The goal of discriminant analysis is to classify a new object as belonging to one of (at least) two mutually exclusive populations in order to maximize the rate of correct classifications. It can consider multiple variables simultaneously and help determine which factors are the most important (Kendall, Stuart, and Ord, 1983). In our case, the two populations are active regions that either do or do not produce a flare during their disk passage. Parameters used are the depth-independent total and maximum values of magnetic flux, and the depth-dependent total and maximum values of all vorticity components and kinetic helicity density yielding 226 parameters.

Discriminant analysis uses an estimate of the probability-density function for each population to determine the population to which a new measurement is most likely to belong. The discriminant boundary occurs where there is equal probability of belonging to each population; the boundary thus indicates a 50% probability forecast. In its simplest form, the probability-density function of each variable is assumed to have a Gaussian distribution, and the mean and standard deviation are estimated from the data; when the covariance matrices for the two populations are also assumed to be equal, the discriminant function is linear in all variables. Leka and Barnes (2007) found that the results of discriminant analysis for flare prediction from photospheric magnetic-field observations tend not to be very sensitive to assumptions about the functional form of the probability density. Further, using a nonparametric estimate for the probability density is likely to improve the classifications if the distributions are non-Gaussian. Thus the results presented here are all based on linear discriminant analysis, but may be improved with more sophisticated analysis. Figure 1

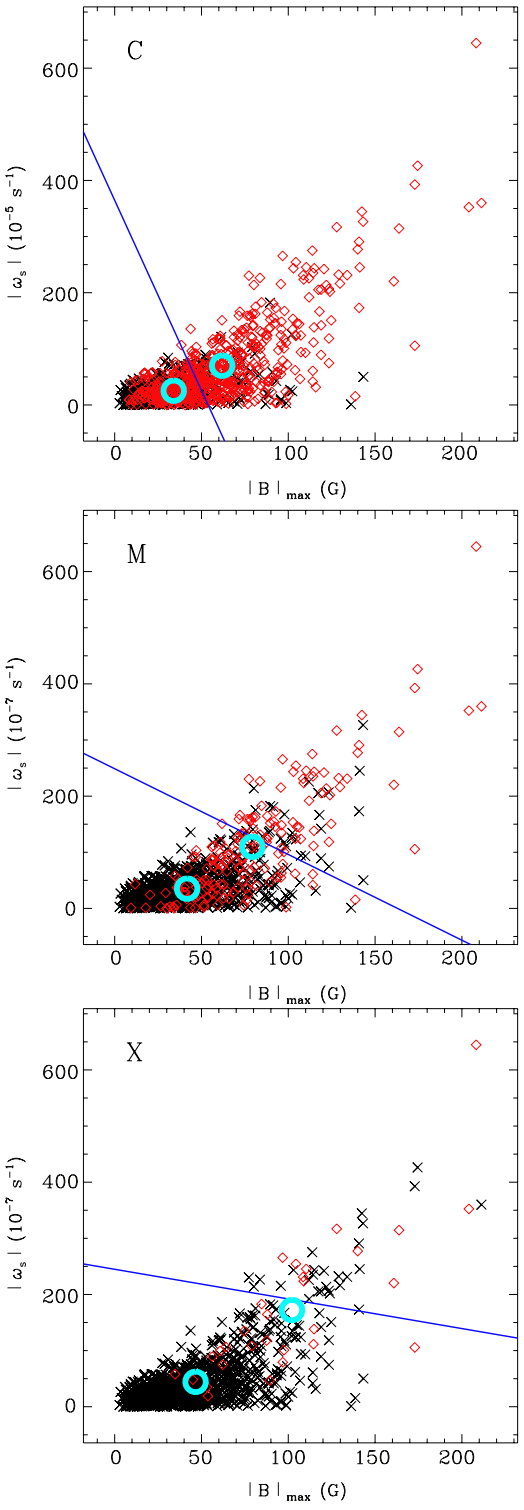


**Figure 1** Discriminant analysis of maximum surface flux and structure vorticity at a depth of 11 Mm, used to distinguish between flare-producing active regions (red diamonds) and non-flaring regions (black crosses) for regions that produce at least C-class (top), M-class (middle), and X-class flares (bottom). The large (cyan) circles indicate the mean values of the two populations. Active regions above and to the right of the discriminant line are predicted to flare, regions below are not.

shows an example of using two variables to discriminate on flare productivity for C-, M-, and X-class flares.

To determine how successful a discriminant function is in classifying new active regions, we calculate a $2 \times 2$ classification table. With $N$ total data points, the first row lists the number of flare-producing active regions that are predicted to flare, $n_{11}$ (number of hits), and that are predicted not to flare, $n_{12}$ (number of missed events). The second row lists the number of non-flaring active regions which are predicted to flare, $n_{21}$ (number of false alarms), and that are predicted not to flare, $n_{22}$ (number of correct nulls). This allows us to calculate the success rate as the number of correct predictions divided by the total number of active regions:

$$r = \frac{n_{11} + n_{22}}{n_{11} + n_{12} + n_{21} + n_{22}} = \frac{n_{11} + n_{22}}{N} \tag{3}$$

which can range between zero and one. In the case of rare events, such as M- and X-class flares, it is easy to get a high success rate by simply assuming that no event will occur. Even a method that would produce a perfect forecast will improve on the success rate of this "prediction" by a very small amount. For this reason, we also calculate general skill scores, which measure the relative performance with respect to a reference forecast:

$$s = \frac{(n_{11} + n_{22}) - N_{\mathrm{ref}}}{N - N_{\mathrm{ref}}}, \tag{4}$$

where $N_{\mathrm{ref}}$ is the value of the reference forecast. Two commonly used skill scores are the one relative to uniform forecasts of always predicting that no event will occur and to forecasts that are correct by chance. With $N_{\mathrm{ref}} = n_{21} + n_{22}$, the no-event skill score is defined as the difference between the number of hits and the number of false alarms divided by the number of events:

$$s_{\mathrm{no}} = \frac{(n_{11} + n_{22}) - (n_{21} + n_{22})}{N - (n_{21} + n_{22})} = \frac{n_{11} - n_{21}}{n_{11} + n_{12}}. \tag{5}$$

Positive values indicate an improvement over the reference forecasts with a skill score of one for a perfect classification, while a negative value indicates that the no-event prediction is better. We mainly use the no-event skill score for the binary classifications in this study. A similar skill score can be constructed for probability forecasts to measure improvements over climatology, which always predicts the same probability ($[n_{11} + n_{12}]/N$) of an event occurring (*e.g.* Balch, 2008).

We expect that a certain number of classifications would be correct by chance. The number of correct forecasts by chance is:

$$N_{\mathrm{chance}} = \frac{1}{N}\big[(n_{11} + n_{12})(n_{11} + n_{21}) \; + \; (n_{12} + n_{22})(n_{21} + n_{22})\big], \tag{6}$$

where the first term describes the number of correct forecasts of events by chance and the second term describes the number of correct no-event predictions by chance (see Balch, 2008, for a review). The Heidke skill score, or the skill score relative to correct forecasts by chance, is then defined as:

$$s_{\mathrm{H}} = \frac{(n_{11} + n_{22}) - N_{\mathrm{chance}}}{N - N_{\mathrm{chance}}}. \tag{7}$$

Positive values indicate an improvement with a skill score of one for a perfect classification. A value of zero implies that the classifications are no better than chance. In this case, the

reference model depends on the method under consideration (Equation (6)). The Heidke skill score provides complementary information to the no-event skill score. It determines if a classification result is better than could be expected by chance, while the no-event skill score determines whether a given parameter is useful as a classification tool.

To construct an unbiased classification table, we use a leave-one-out technique removing one object from the samples and classifying the excluded point by a discriminant analysis of the remaining objects (Hills, 1966). This procedure is then repeated excluding each object in turn. More details about the linear discriminant analysis, as used here, can be found in Leka and Barnes (2007).

## 3. Results

### 3.1. Single Subsurface Parameter

First, we examine how well vorticity and kinetic helicity perform at different depths and compare the results with that of maximum and total magnetic flux, to see where the greatest improvements occur. For the C-class flare threshold (Figure 2), we show, as an example, the rate of correct classification and the no-event and the Heidke skill score for the structure vorticity as a function of depth. In this case, the structure vorticity is better at classifying flaring and non-flaring regions than a no-event prediction ($r = 0.491$) and the skill scores track each other, which is to be expected since the two sample sizes are approximately the same size. The skill-score values are greater than approximately 0.1 at all depths, but even the highest no-event skill scores from applying linear discriminant analysis to the structure vorticity measured between 7 and 10 Mm depth are smaller than those of the magnetic measures.

For the M-class flare threshold (Figure 3), the structure vorticity at all depths leads to positive skill scores, indicating that it is better at classifying flaring and non-flaring regions than a no-event prediction. The Heidke skill score is generally much larger than the no-event

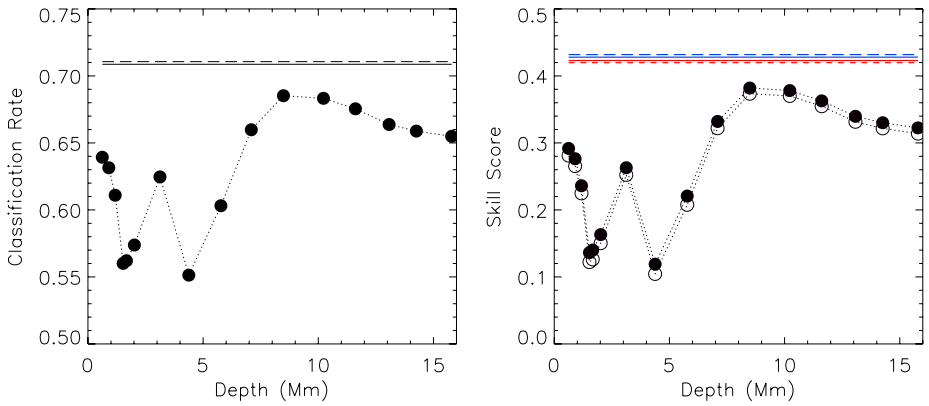


**Figure 2** The rate of correct classification of the C-class threshold category (left) using maximum structure vorticity (filled circles) compared to using total flux (solid) or maximum flux (long-dashed). The corresponding skill score (right) measures the performance of vorticity compared to a no-event prediction (filled circles) compared to the skill score of total flux (blue solid) or maximum flux (blue long-dashed). For comparison, we include the Heidke skill score for structure vorticity (open circle) and total flux (red dashed) or maximum flux (red dot-dashed). The multiple horizontal lines are almost indistinguishable.

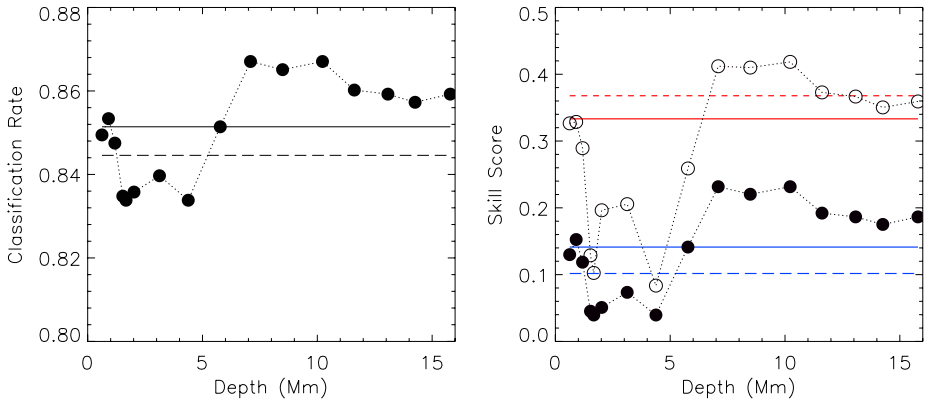


**Figure 3** The rate of correct classification of the M-class threshold category (left) using maximum structure vorticity (filled circles) compared to using total flux (solid) or maximum flux (long-dashed). The corresponding skill score (right) measures the performance compared to a no-event prediction (filled circles) compared to the skill score of total flux (blue solid) or maximum flux (blue long-dashed). For comparison, we include the Heidke skill score for structure vorticity (open circle) and total flux (red dashed) or maximum flux (red dot-dashed).

skill score at each depth, which is a result of the much smaller sample of flaring regions. At depths of 7 Mm and deeper, linear discriminant analysis applied to the structure vorticity leads to higher classification rates, and a higher no-event skill score than either magnetic measure. The best depths for the structure vorticity are 7 to 10 Mm, with the maximum no-event skill score being $s_{\mathrm{no}} = 0.232$, compared to $s_{\mathrm{no}} = 0.141$ for the total flux.

With the large number of variables being considered, one must be aware of the possibility of obtaining a non-zero skill score through chance alone. Therefore, a numerical experiment was performed, in which pairs of samples were randomly drawn from the same Gaussian distribution. Discriminant analysis was applied to these samples, and the distribution of the resulting skill scores was analyzed. For the same sample sizes as the C-class flare threshold, it was found that the probability of producing a skill score of 0.1 or larger was less than 1%. Thus it is likely that the maximum structure vorticity measured at any depth contains some information about whether an active region will produce C-class or larger flares. For the same sample sizes as the M-class flare threshold, it was found that the probability of producing any non-zero skill score was less than 1%. Thus it is likely that the maximum structure vorticity measured at any depth also contains some information about whether an active region will produce M-class or larger flares.

For the X-class threshold (Figure 4), the structure vorticity measured at depths from 1 – 6 Mm leads to higher classification rates and no-event skill scores than the magnetic measures. But, the highest no-event skill scores are still negative. While the large Heidke skill scores imply that these classifications are not due to chance, the consistently negative skill scores result from the assumption of a Gaussian distribution not being met. In particular, the distribution of the structure vorticity for no-event regions has a relatively long positive tail which is not well represented by a Gaussian. For classification of X-threshold events to result in positive skill scores, it is necessary to accurately represent this tail.

We repeat the same analysis for all parameters and determine the depth with the largest no-event skill score for each parameter. Figure 5 shows the results for the maximum and the total values of single subsurface parameters. For the C-class threshold, the parameters reach no-event skill scores greater than zero at most depths (crosses). Exceptions are the maximum

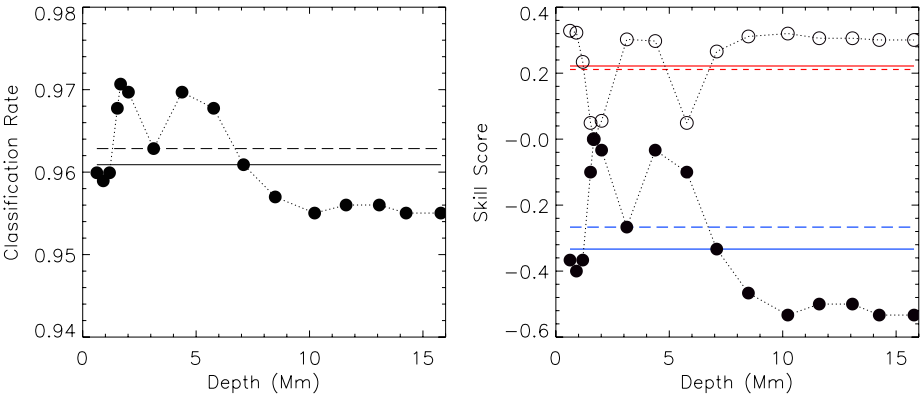


**Figure 4** The rate of correct classification of the X-class category (left) using maximum-structure vorticity (filled circles) compared to using total flux (solid) or maximum flux (long-dashed). The corresponding skill score (right) measures the performance compared to a no-event prediction (filled circles) compared to the skill score of total flux (blue solid) or maximum flux (blue long-dashed). For comparison, we include the Heidke skill score for structure vorticity (open circle) and total flux (red dashed) or maximum flux (red dot-dashed).

and total structure vorticity at 2 and 4 Mm and the maximum values of the vertical-vorticity and helicity contributions, $\omega_z$ and $h_z$ at depths between 2 and 6 Mm (no symbols). All skill scores of the subsurface parameters are smaller than the ones of total or maximum flux, as in the example of Figure 2. The largest skill scores occur for parameters measured near the surface or at greater depth (blue or red squares).

For the M-class flare category, the parameters reach no-event skill scores greater than zero at most depths. Exceptions are again the vertical-vorticity and helicity contributions, as in the case of the C-class category. When using maximum values (left), three parameters, the zonal-vorticity component [$\omega_x$], the enstrophy [$\omega^2$], and the structure vorticity [$\omega_s$], are better at separating flaring from non-flaring regions than both magnetic-flux measures at several depths (filled circles). When using total values (right), the helicity is the fourth parameter in addition to the three other ones that produce higher no-event skill scores than the magnetic measures. The largest skill scores occur mainly from parameters measured at depths of 7 Mm and below. The highest skill score of $s_{no} = 0.232$ is reached by using the structure vorticity measured at depths of 8 and 10 Mm (see Figure 3). With 177 active regions producing M-class flares, this skill score implies that 16 additional regions are correctly classified compared to a skill score of $s_{no} = 0.141$ for total flux and that 41 regions are correctly classified compared to a no-event forecast.

For the X-class flare category, we find that there is no single subsurface parameter that results in a positive no-event skill score. Most subsurface parameters have some depths where they will be better at classifying the data set than does the total or maximum magnetic flux. However, even the best-performing parameters (enstrophy) lead to negative no-event skill scores as in Figure 4. This confirms the earlier statement that the assumptions of linear discriminant analysis are violated; the distributions are highly non-Gaussian. For the rest of this study, we will focus on C- and M-class flare activity and use only the no-event skill score, since the Heidke skill score tracks it quite well.

### 3.2. Combinations of Subsurface and Magnetic Parameters

Next, we determine whether combining a magnetic with a subsurface parameter improves the skill. We perform the discriminant analysis for all subsurface parameters and determine

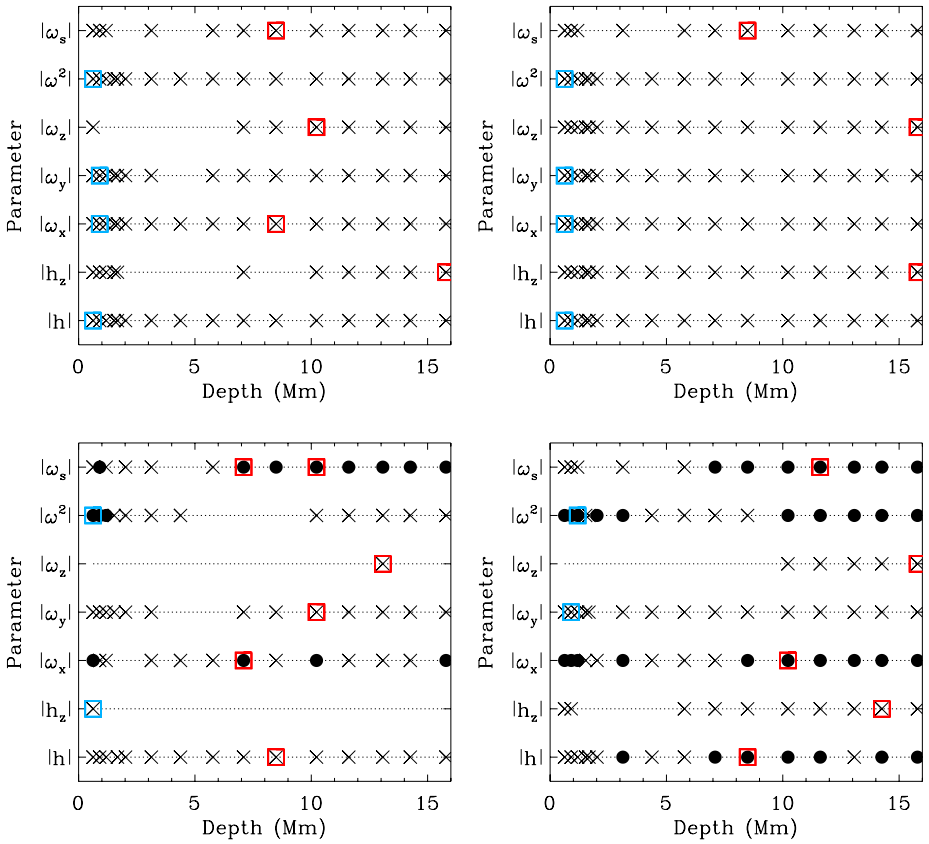


**Figure 5** The no-event skill score as a function of depth for maximum (left) and total value (right) of subsurface parameters for C-class (top) and M-class category (bottom) indicating positive skill scores smaller than the ones of both magnetic flux measures (crosses: good), and positive skill scores greater than the ones of both flux measures (filled circles: best). No symbols are plotted for negative skill scores (bad). Square symbols indicate the depth of the maximum skill score; color indicates depth ranges (blue: 0.6 – 2.0 Mm, black: 3.1 – 5.8 Mm, red: 7.1 – 15.8 Mm).

the depth with the largest no-event skill score for each parameter. Figure 6 shows the results for the maximum values of the parameters. For the C-class category, the two-parameter combinations lead to higher classification rates and skill scores than the subsurface parameters alone. However, the combined use of a subsurface parameter with magnetic flux is barely better than a magnetic parameter alone. The combination of total flux with structure vorticity leads to the highest skill score of $s_{\mathrm{no}} = 0.478$ compared to combining total and maximum flux with $s_{\mathrm{no}} = 0.464$. The largest skill scores occur for subsurface parameters measured near the surface or at greater depth (blue or red symbols); see Section 4.

For the M-class category, the two-parameter combinations with total flux lead to higher classification rates and skill scores than either the magnetic or subsurface parameters alone. The maximum-flux combinations lead to better values than the subsurface parameters alone but they are not always better than magnetic measures alone. The highest skill score of $s_{\mathrm{no}} = 0.260$ is reached by combining structure vorticity with total flux (see Figures 3 and 5). The number of incorrectly classified active regions decreases to 131 compared to 136 when

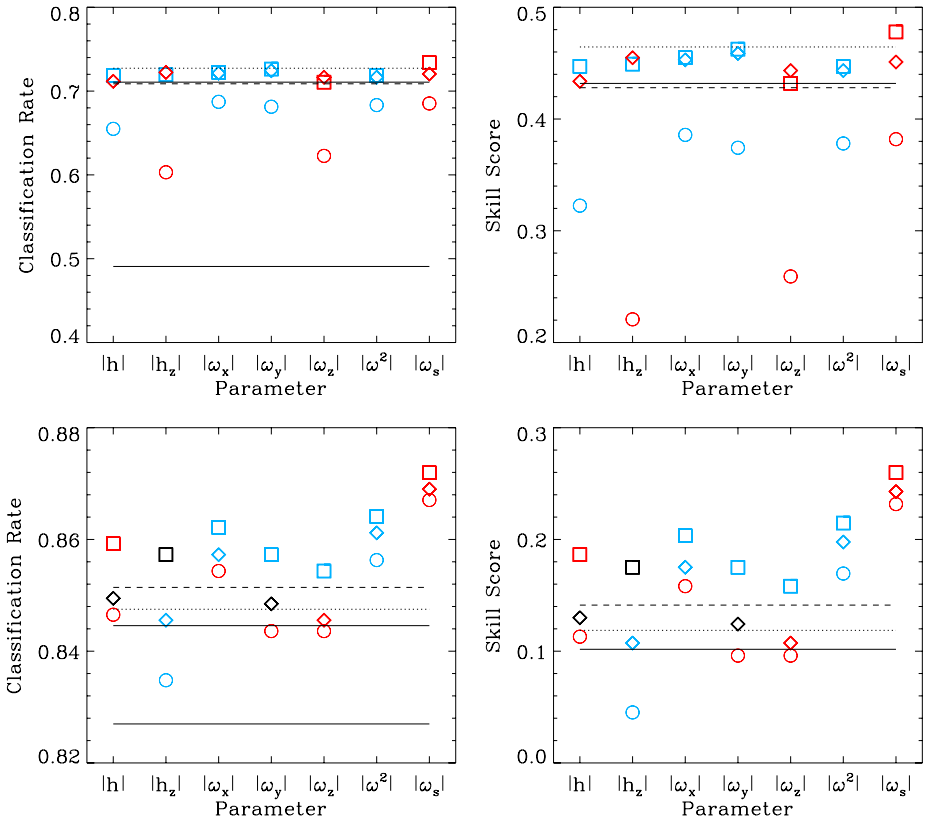


**Figure 6** The highest correct classification rate (left) and the highest no-event skill score (right) for each subsurface flow parameter using maximum values combined with total flux (squares) or maximum flux (diamonds) for C-class (top) and M-class categories (bottom). For comparison, we include the values of single subsurface parameters (circles), total flux (dashed), maximum flux (dot-dashed), and the combination of both magnetic measures (dotted) as well as the rate of predicting that no events occur (solid). The colors indicate depth ranges (blue: 0.6 – 2.0 Mm, black: 3.1 – 5.8 Mm, red: 7.1 – 15.8 Mm).

using structure vorticity alone. However, the other two-parameter combinations fail to improve upon the best single parameters.

Figure 7 shows the same as Figure 6 for the total subsurface parameters. This leads to similar results as for the maximum values. For the C-class category, the best skill scores are reached by combining maximum magnetic flux with total zonal or meridional vorticity components [$\omega_x$ and $\omega_y$] but they are slightly lower than the results of combining maximum structure vorticity and total flux (Figure 6). For the M-class category, the largest skill scores are reached by combining total flux with total enstrophy for a skill score of $s_{\rm no} = 0.237$. This is smaller than the skill score reached by combining maximum structure vorticity with total flux (Figure 6) but larger than every other two-parameter combination. We confirm that including maximum structure vorticity enhances the ability to discriminate between the flaring and non-flaring populations for the C- and M-class thresholds.

Note that the results of combining the magnetic parameters for the M-class flare category results in a lower success rate and skill score than the total magnetic flux considered alone (*e.g.* Figure 6). As in the X-flare threshold case, this suggests that the assumption

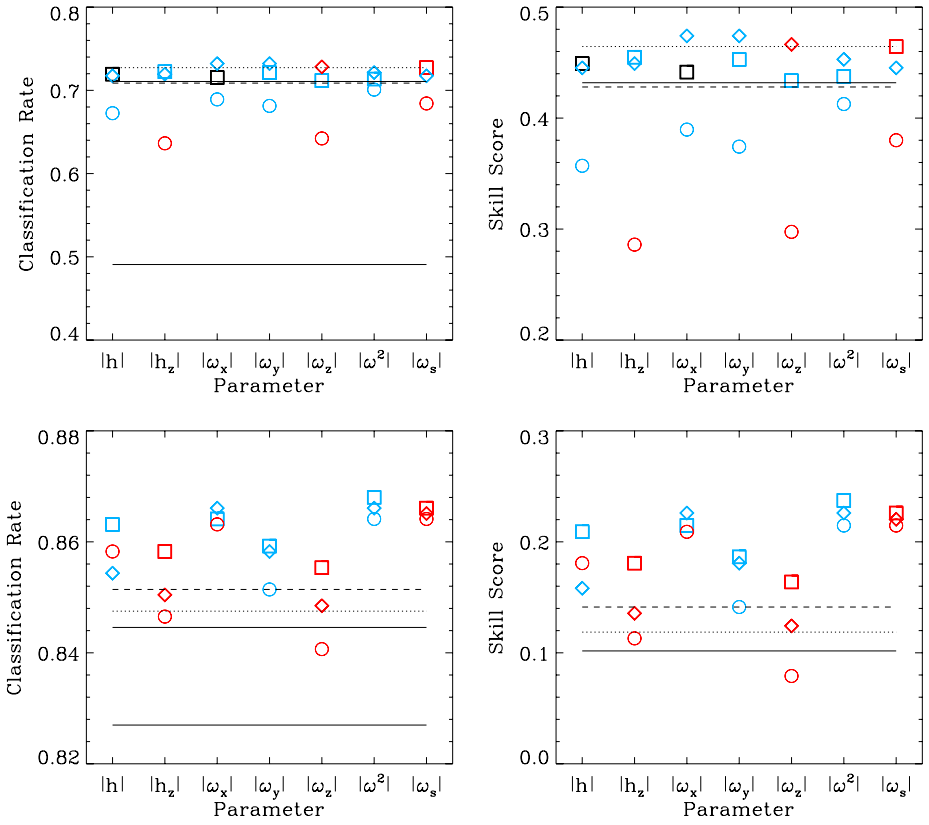


**Figure 7** The highest correct classification rate (left) and the highest no-event skill score (right) for each subsurface flow parameter using total values combined with total flux (squares) or maximum flux (diamonds) for C-class (top) and M-class categories (bottom). For comparison, we include the values of single subsurface parameters (circles), total flux (dashed), maximum flux (dot-dashed), and the combination of both magnetic measures (dotted) as well as the rate of predicting that no events occur (solid). The colors indicate depth ranges (blue: 0.6 – 2.0 Mm, black: 3.1 – 5.8 Mm, red: 7.1 – 15.8 Mm).

of Gaussian distribution for the variables does not hold, as combining variables should not lead to less information than the variables considered alone, provided that the assumptions are correct.

### 3.3. Multiple Subsurface Parameters

Does increasing the number of subsurface parameters improve the classification of flaring and non-flaring active regions? We calculate all combinations of two subsurface plus total magnetic-flux $(2 + 1)$ parameters for the 14 subsurface parameters (maximum and total value) at 16 depths. We calculate the skill scores as a function of depth for each of the $14 \times 14$ combinations and find that for the C-class category the highest skill score is $s_{\mathrm{no}} = 0.476$ when combining maximum structure vorticity at 1.2 and 7.1 Mm. For the M-class category, the best-performing $2 + 1$ parameter combination with a skill score of $s_{\mathrm{no}} = 0.294$ is the combination of maximum structure vorticity at 10.2 and 15.8 Mm. This skill score is larger than the best pair of a single subsurface and magnetic parameter which means that adding

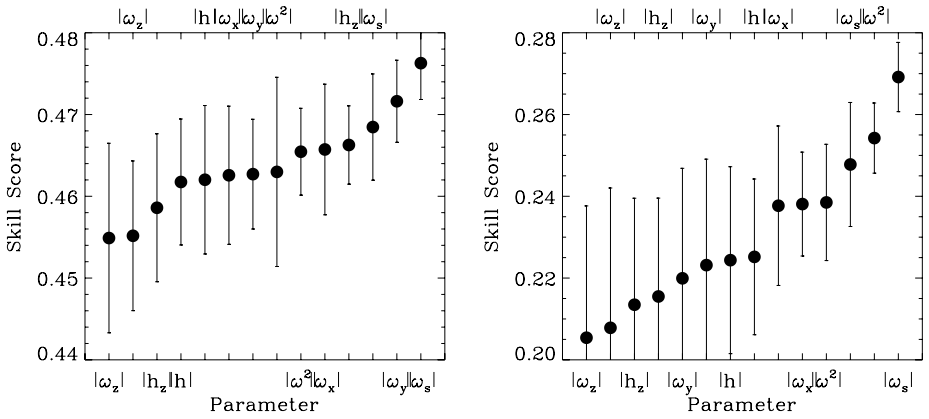


**Figure 8** The average no-event skill score calculated over the 14 maximum skill scores of all $2+1$ combinations that include the indicated parameter, for the C-class (left) and M-class (right) thresholds. The top (bottom) label indicates the total (maximum) value of the parameter shown. The error bars indicate the standard deviation.

a second subsurface parameter can improve the classification. For both flare thresholds, the well-performing parameters lead to high skill scores with almost all other parameters, while low-skill parameters lead to low values with most parameters except the well-performing ones. As a consequence, we can rank the subsurface parameters by determining the average skill scores that are calculated over the 14 maximum skill scores of all $2+1$ combinations that include a given parameter (Figure 8). The maximum structure vorticity is on average the best-performing parameter for both flare categories. The standard deviation is small indicating that this quantity consistently leads to high skill scores. The vertical-vorticity component leads to the smallest skill scores with large standard deviations.

Now, we use multiple depths of the best-performing parameter, maximum structure vorticity, to check whether adding more depths improves the classification. In Table 1, we include the total enstrophy as an example of a well-performing parameter. For the C-class category, the combination of three depths of structure vorticity with total flux leads to the highest skill score of $s_{\mathrm{no}} = 0.480$. The three depths cover near-surface and deeper layers. The result is marginally better than the $2+1$ parameter combination discussed above. For the M-class category, the combination of three depths of structure vorticity with total flux leads to the highest skill score of $s_{\mathrm{no}} = 0.316$ which is better than any $2+1$ combination discussed above. Again, the three depths cover shallow and deeper layers. For total enstrophy, the maximum skill score is reached for the combination of two or three depths with a magnetic measure. However, the largest skill scores achieved by using total enstrophy are smaller than those using structure vorticity. These results suggest that the variation with depth of the subsurface parameters contains additional information and that a small number of parameters leads to the best results in linear discriminant analysis. The information comes mainly from deep and shallow layers and less from intermediate depths.

Finally, we see whether increasing the number of subsurface parameters beyond four improves the classification of flaring and non-flaring active regions. Since the number of subsurface parameters is large [224], it is challenging to calculate all combinations for large numbers of parameters and depths. As an alternate search strategy, we use a "step-up" procedure (Klecka, 1980), starting with the best-performing combination of a subsurface and a magnetic parameter for each flare class. We then add a third parameter and perform the

**Table 1** The no-event skill score [the number of correctly classified flaring active regions (number of hits)] and the depths as a function of the number of subsurface parameters used. Top: total enstrophy [$\omega^2$] and maximum structure vorticity [$\omega_s$] for the C-class category (521 flaring active regions), Bottom: the same for the M-class category (177 flaring active regions). A magnetic measure included in the analysis is indicated by +1 in the first column (total flux was used except for enstrophy in the C-class category). The skill score [$s_{\rm no}$] together with the number of hits [$n_{11}$], the number of events, and the number of non-events, describes the information contained in the $2 \times 2$ classification tables.

| $N$ | skill score | | number of hits | | depth (Mm) | |
|---|---|---|---|---|---|---|
| | $\omega^2$ | $\lvert\omega_{\rm S}\rvert$ | $\omega^2$ | $\lvert\omega_{\rm S}\rvert$ | $\omega^2$ | $\lvert\omega_{\rm S}\rvert$ |
| C-class category | | | | | | |
| 1+0 | 0.413 | 0.382 | 281 | 290 | 0.6 | 8.5 |
| 2+0 | 0.422 | 0.403 | 285 | 282 | 0.6, 2.0 | 0.6, 8.5 |
| 1+1 | 0.453 | 0.478 | 327 | 319 | 0.9 | 8.5 |
| 2+1 | 0.464 | 0.476 | 340 | 319 | 0.9, 8.5 | 1.2, 7.1 |
| 3+1 | 0.470 | 0.480 | 343 | 318 | 0.9, 1.7, 8.5 | 1.5, 1.7, 8.5 |
| 4+1 | 0.476 | 0.480 | 340 | 331 | 0.9, 2.0, 3.1, 7.1 | 1.5, 5.8, 8.5 14.3 |
| 16+1 | 0.459 | 0.440 | 345 | 323 | all | all |
| M-class category | | | | | | |
| 1+0 | 0.215 | 0.232 | 50 | 64 | 1.2 | 7.1 (10.2) |
| 2+0 | 0.243 | 0.243 | 54 | 65 | 1.2, 3.1 | 0.9, 10.2 |
| 1+1 | 0.237 | 0.260 | 62 | 69 | 1.2 | 10.2 |
| 2+1 | 0.266 | 0.294 | 65 | 74 | 1.2, 3.1 | 10.2, 15.8 |
| 3+1 | 0.260 | 0.316 | 65 | 79 | 0.9, 3.1, 5.8 | 2.0, 10.2, 11.6 |
| 4+1 | 0.260 | 0.316 | 69 | 79 | 0.9, 1.7, 3.1, 5.8 | 2.0, 4.4, 10.2, 11.6 |
| 16+1 | 0.226 | 0.237 | 66 | 71 | all | all |

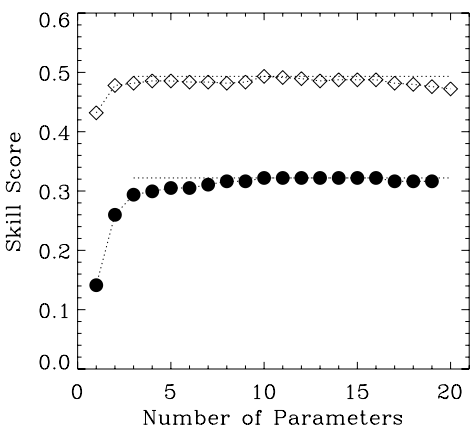


**Figure 9** The no-event skill score as a function of the number of parameters used for the C-class (diamonds) and M-class (filled circles) thresholds. The single parameter value is the magnetic-flux value. The dotted lines indicate the maximum skill score for each flare class.

discriminant analysis. We repeat this for all possible choices for the third parameter and determine the best-performing combination of three parameters. Next, we add a fourth parameter to the three-parameter combination and repeat the process. In this way, we add more and more parameters.

Figure 9 shows the resulting skill scores as a function of the number of parameters used. The number of parameters and the sample sizes result in non-unique combinations which

**Table 2** The no-event skill score and the number of correctly classified flaring active regions (number of hits) as a function of the number of parameters used in a discriminant analysis shown for the C-class category (521 flaring active regions) and the M-class category (177 flaring active regions). The result of the best magnetic measure (1 B) is included as baseline.

| $N$ | no-event skill score | | number of hits | |
|---|---|---|---|---|
| | C class | M class | C class | M class |
| 1 B | 0.432 | 0.141 | 328 | 61 |
| 1 | 0.413 | 0.232 | 281 | 62 |
| 2 | 0.478 | 0.260 | 319 | 69 |
| 5 | 0.486 | 0.305 | 324 | 74 |
| 10 | 0.493 | 0.322 | 326 | 77 |
| 20 | 0.480 | 0.316 | 319 | 77 |

return the same skill score. For the C-class flare category, the number of best choices at a given step remains small (four or fewer) and we can calculate all possibilities. The combination of four or five parameters leads to a skill score of $s_{\mathrm{no}} = 0.486$, while the ten-parameter combination results in a slightly larger value of $s_{\mathrm{no}} = 0.493$.

For the M-class category, there is gradual improvement up to five parameters and the choices are unique. Then, there are 38 parameters that lead to the same skill score when used as the sixth parameter. We select either the parameter at the greatest depth or the one closest to the surface. From the seventh to the tenth parameter, the choice of selecting the parameter at greater depth leads to consistently better skill scores than selecting the parameter at shallower depth. We have unsuccessfully tried a few other possible choices. However, we cannot rule out that we have missed the optimum combination. The combination of five parameters leads to a skill score of $s_{\mathrm{no}} = 0.305$, while the ten-parameter combination leads to a somewhat larger value of $s_{\mathrm{no}} = 0.322$.

For both flare classes using this search strategy, the maximum skill score is reached with the combination of ten parameters; additional parameters do not provide additional information. The resulting skill scores are slightly higher than the ones of the $3+1$ parameter combinations in Table 1. There is, of course, no guarantee that the applied search strategy leads to the best-performing parameter combination. However, the results gives us some confidence that the available information can be captured with a small number of parameters.

## 4. Summary and Conclusions

The subsurface-flow characteristics have the ability to distinguish between flaring and non-flaring active regions when analyzing synoptic data of 1023 active regions with linear discriminant analysis. This agrees with the fact that large values of inferred subsurface twisting motions occur mainly in active regions that produce energetic flares (Mason *et al.*, 2006; Komm and Hill, 2009). For C- and M-class flare thresholds, the most important subsurface parameter is the so-called structure vorticity, which can be interpreted as an estimate of the horizontal gradient of the horizontal vorticity components. We use the no-event skill score, which measures the improvement over predicting that no events occur, to determine the performance of each set of parameters. The highest no-event skill score is 0.48 for C-class flares and 0.32 for M-class flares, when the structure vorticity at three depths combined with the total magnetic flux are used. The subsurface vorticity contributions come mainly from shallow layers within about 2 Mm of the surface and layers deeper than about 7 Mm, which agrees with previous results indicating three depth regimes with distinct characteristics (Komm, 2007). The variation with depth of subsurface vorticity is important for the classification of flaring and non-flaring active regions.

We speculate that the flows at intermediate depths may produce a less clear signal with respect to flare productivity as a consequence of overlap of strong flows from different mechanisms simultaneously present below active regions. For example, some observational and theoretical work on emerging-flux systems suggests the presence of strong up-flows and possibly shearing flows in regions consistent with our deeper layers (Manchester, 2008; Komm, Howe, and Hill, 2009), and down-flows due to adiabatic expansion near the upper layers (Fan, 2001; Schüssler and Rempel, 2005). It can be easily pictured that, while emerging flux and shearing flows are likely properties of flare-productive active regions, the dominant signals of the associated mechanisms are prominent at depth and near the surface, but less distinct in the intermediate layers.

The greater discrimination resulting from structure vorticity compared to the vertical vorticity is due to noise levels and analysis regions. That is, since the vertical size of the inversion kernels is small compared to the horizontal size of the dense-pack patches, the zonal and meridional vorticity are probably better sampled than the vertical vorticity, calculated as the curl of the horizontal velocity components, and thus potentially less noisy. The presence of this dipolar pattern captured by the structure vorticity is consistent with the existence of gradients in the sub-surface plasma flow associated with flare-productive active regions. We plan to investigate the inferred flow characteristics using smaller patches to quantify the variation of the signal and the noise in cases using different analysis scales.

For the X-class threshold, the small sample size and non-Gaussian distributions of the parameters result in skill scores that are systematically negative. This result indicates that the assumptions needed to justify applying linear discriminant analysis are violated and no resulting skill score is credible.

Barnes and Leka (2008) analyze magnetic-field data and daily forecasts for an M-class threshold, comparing different derived parameters published by various authors in order to make a systematic comparison of different techniques. They find skill scores of $s_{\mathrm{no}} \leq 0.16$ and conclude that the parameters used show only modest improvements over climatological forecasts. Our results for the M-class threshold are not directly comparable since we use synoptic data to classify flaring and non-flaring active regions. Nevertheless, we find a comparable skill score of 0.14 using total flux, and find the higher skill scores that result from including subsurface parameters such as structure vorticity very encouraging.

We have tried to minimize the number of necessary subsurface parameters searching for an optimum parameter set. Our simple search strategy indicates that about ten parameters or less are the optimum number for our data sets. This is similar to the results of Leka and Barnes (2007) who found that about three parameters contain most of the available information from the magnetic-field data. The relatively small increase in skill when several parameters are used may indicate that parameters are correlated, or may indicate that there is no additional information contained. Since the inversion kernels used to determine subsurface flows as a function of depth extend in the radial direction with their widths getting larger in deeper layers (Birch *et al.*, 2007), flows measured at adjacent depths are not completely independent, leading to correlations among parameters. The best parameter combinations include information from shallow and deeper layers which are as uncorrelated as possible.

More work is clearly needed. One task is to better account for regions that are short-lived in the synoptic data or appear very close to larger active regions and thus are "mixed in" with the latter's signals.

A second task is to study the temporal behavior of subsurface flows associated with active regions on time scales of one day or shorter in order to measure the evolution of active regions and their flare activity. This will allow us to more directly compare our subsurface results with results derived with other methods using magnetic activity at the surface. A preliminary investigation of the timing of flares and changes in subsurface parameters is under

way (Reinard *et al.*, 2010). Initial results based on daily flow maps are encouraging and indicate that this method might be useful for flare forecasting.

We plan to investigate the practicality of producing flare forecasts using subsurface vorticity values. An important step will be to evaluate the quality of flow maps derived from shorter time series. For example, if a continuous time series of eight hours were sufficient, we could obtain flow maps from data obtained at a single ground-based site, such as one of the GONG stations. With continuous data available from the GONG program and the SDO spacecraft, we could then update the subsurface flow information about every four hours.

**Acknowledgements** We thank the referee for helpful comments. This work utilizes data obtained by the Global Oscillation Network Group (GONG) program, managed by the National Solar Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA), Inc. under a cooperative agreement with the National Science Foundation. The data were acquired by instruments operated by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias, and Cerro Tololo Interamerican Observatory. NSO/Kitt Peak data used here are produced cooperatively by NSF/NSO, NASA/GSFC, and NOAA/SEC. SOLIS VSM data used here are produced cooperatively by NSF/NSO and NASA/GSFC. This work was supported by NASA grant NNG08EI54I to the National Solar Observatory. This work is carried out through the NSO's REU site program, which is co-funded by the Department of Defense in partnership with the NSF REU Program. The discriminant analysis code was developed by Barnes and Leka with funding from AFOSR under contracts F49620-00-C-0004, F49620-03-C-0019, and from NASA under contract NNH07CD25C.

## References

Abramenko, V.I.: 2005, *Astrophys. J.* **629**, 1141.
Balch, C.C.: 2008, *Space Weather* **6**, 1001.
Barnes, G., Leka, K.D.: 2008, *Astrophys. J. Lett.* **688**, L107.
Birch, A.C., Gizon, L., Hindman, B.W., Haber, D.A.: 2007, *Astrophys. J.* **662**, 730.
Braun, D.C., Birch, A.C., Lindsey, C.: 2004, In: Danesy, D. (ed.) *Helio- and Asteroseismology: Towards a Golden Future* **SP-559**, ESA, Noordwijk, 337.
Corbard, T., Toner, C., Hill, F., Hanna, K.D., Haber, D.A., Hindman, B.W., Bogart, R.S.: 2003, In: Sawaya-Lacoste, H. (ed.) *Local and Global Helioseismology: The Present and Future* **SP-517**, ESA, Noordwijk, 255.
Cui, Y., Li, R., Zhang, L., He, Y., Wang, H.: 2006, *Solar Phys.* **237**, 45.
Falconer, D.A., Moore, R.L., Gary, A., Adams, M.: 2009, *Astrophys. J. Lett.* **700**, L166.
Fan, Y.: 2001, *Astrophys. J.* **554**, L111.
Ferguson, R., Komm, R., Hill, F., Barnes, G., Leka, K.D.: 2009, In: Dikpati, M., Hill, F., González Hernández, I., Arentoft, T. (eds.): *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology* **CS 416**, Astron. Soc. Pac., San Francisco, 127.
Georgoulis, M.K., Rust, D.M.: 2007, *Astrophys. J.* **661**, L109.
Gizon, L., Birch, A.C.: 2005, Local Helioseismology, *Living Rev. Solar Phys.* **2**(6). http://solarphysics.livingreviews.org/Articles/lrsp-2005-6.
Haber, D.A., Hindman, B.W., Toomre, J., Bogart, R.S., Larsen, R.M., Hill, F.: 2002, *Astrophys. J.* **570**, 885.
Henney, C.J., SOLIS Team: 2007, *AGU Fall Meeting Abstracts*, A1104.
Hills, M.: 1966, *J. Roy. Stat. Soc. B* **28**, 1.
Jones, H.P., Harvey, J.W., Henney, C.J., Keller, C.U., Malanushenko, O.M.: 2004, *Bull. Am. Astron. Soc.* **36**, 709.
Kendall, M.G., Stuart, A., Ord, J.K.: 1983, *Advanced Theory of Statistics* **3**, 4th edn., Macmillan, New York.
Klecka, W.R.: 1980, *Discriminant Analysis*, Sage Publications, Newbury Park.
Komm, R.: 2007, *Astron. Nachr.* **328**, 269.
Komm, R., Hill, F.: 2009, *J. Geophys. Res. A* **114**, 6105.
Komm, R., Howe, R., Hill, F.: 2009, *Solar Phys.* **258**, 13.
Komm, R., Howe, R., Hill, F., González Hernández, I., Toner, C., Corbard, T.: 2005, *Astrophys. J.* **631**, 636.
Komm, R.W., Corbard, T., Durney, B.R., González Hernández, I., Hill, F., Howe, R., Toner, C.: 2004, *Astrophys. J.* **605**, 554.
Leka, K.D., Barnes, G.: 2007, *Astrophys. J.* **656**, 1173.
Lesieur, M.: 1987, *Turbulence in Fluids*, Kluwer, Dordrecht.

Longcope, D.W., Fisher, G.H., Pevtsov, A.A.: 1998, *Astrophys. J.* **507**, 417.
Manchester, W. IV: 2008, In: Howe, R., Komm, R., Balasubramaniam, K.S., Petrie, G.J.D. (eds.) *Subsurface and Atmospheric Influences on Solar Activity* **CS 383**, Astron. Soc. Pac., San Francisco, 163.
Mason, D., Komm, R., Hill, F., Howe, R., Haber, D., Hindman, B.: 2006, *Astrophys. J.* **645**, 1543.
McAteer, R.T.J., Gallagher, P.T., Ireland, J.: 2005, *Astrophys. J.* **631**, 628.
Reinard, A.A., Henthorn, J., Komm, R., Hill, F.: 2010, *Astrophys. J. Lett.* **710**, L121.
Schrijver, C.J.: 2007, *Astrophys. J.* **655**, L117.
Schüssler, M., Rempel, M.: 2005, *Astron. Astrophys.* **441**, 337.
Zhao, J., Kosovichev, A.G.: 2004, *Astrophys. J.* **603**, 776.