This is a repository copy of *Active longitude and CME occurrences*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/113087/

Article:

Gyenge, N., Singh, T., Kiss, T.S. et al. (2 more authors) Active longitude and CME occurrences. (Unpublished)

ACTIVE LONGITUDES AND OCCURRENCE OF CORONAL MASS EJECTIONS

N. Gyenge[1,2*] , T. Singh[3], T. S. Kiss[1,4], A. K. Srivastava[3], R. Erdélyi[1]

[1]Solar Physics and Space Plasmas Research Centre (SP2RC), School of Mathematics and Statistics, University of Sheffield
Hounsfield Road, Hicks Building, Sheffield S3 7RH, UK
[2]Debrecen Heliophysical Observatory (DHO), Konkoly Observatory, Research Centre for Astronomy and Earth Sciences
Hungarian Academy of Sciences, Debrecen, P.O.Box 30, H-4010, Hungary
[3] Department of Physics, Indian Institute of Technology (Banaras Hindu University), Varanasi, India
[4] Department of Physics, University of Debrecen, Egyetem tér 1, Debrecen, H-4010, Hungary

Draft version February 23, 2017

ABSTRACT

The spatial inhomogeneity of the distribution of coronal mass ejection (CME) occurrences in the solar atmosphere could provide a tool to estimate the longitudinal position of the most probable CME-capable active regions in the Sun. The anomaly in the longitudinal distribution of active regions themselves is often referred to as active longitude (AL). In order to reveal the connection between the AL and CME spatial occurrences, here we investigate the morphological properties of active regions. The first morphological property studied is the separateness parameter, which is able to characterise the probability of the occurrence of an energetic event, such as solar flare or CME. The second morphological property is the sunspot tilt angle. The tilt angle of sunspot groups allows us to estimate the helicity of active regions. The increased helicity leads to a more complex built-up of the magnetic structure and also can cause CME eruption. We found that the most complex active regions appear near to the AL and the AL itself is associated with the most tilted active regions. Therefore, the number of CME occurrences is higher within the AL. The origin of the fast CMEs is also found to be associated with this region. We concluded that the source of the most probably CME-capable active regions is at the AL. By applying this method we can potentially forecast a flare and/or CME source several Carrington rotations in advance. This finding also provides new information for solar dynamo modelling.

1. SOLAR NON-AXISYMMETRIC ACTIVITY

Since the beginning of the last century the non-homogeneous spatial properties of solar activity have been studied extensively (Carrington 1863; Chidambara 1932; Maunder 1905; Losh 1939; Bumba 1965; Bumba & Obridko 1969). These early investigations conjectured initially that the longitudinal distribution of sunspot groups or sunspot numbers shows non-homogeneous behaviour. These analyses concluded that there are preferred longitudes, where solar activity concentrates.

Later, different approximations and assumptions were applied to understand the essence of this phenomenon. The topic soon became controversial (see e.g. Pelt et al. 2006; Henney & Durney 2005). Overall, three approaches can be distinguished. The first approach is the quasi-rigid structure model by Warwick (1966). This model describes a constantly rotating frame which carries the persistent domains of activity (Ivanov 2007). Bogart (1982) applied an autocorrelation statistical method based on long-term sunspot number data. Balthasar & Schüssler (1983, 1984) applied period-analysis to the Greenwich Photoheliographics Results (GPR). These studies concluded that the angular velocity of the quasi-rigid rotating frame varies. The angular velocity depends on the solar cycle, but during one cycle the angular velocity is constant.

The second approach, promoted by e.g. Becker (1955); Castenmiller et al. (1986) and Brouwer & Zwaan (1990)

*e-mail: n.g.gyenge@sheffield.ac.uk

discovered the 'active nest' and defined it as a small and isolated area on the solar surface. Here, the enhanced longitudinal activities are considered as individual entities. These isolated entities can be absent for several rotations.

The third group of models assumes a migrating activity in the Carrington coordinate system. Berdyugina & Usoskin (2003); Berdyugina (2004, 2005) found persistent ALs under the influence of the differential rotation. Usoskin et al. (2005) and Berdyugina et al. (2006) introduced a 'dynamic reference frame'. This frame describes the longitudinal migration of active longitude in Carrington coordinate system and the frame has a similar dynamics to differential rotation. Usoskin et al. (2007) concluded that the migration of the enhanced activity is just apparent. In the 'dynamic reference frame', the rotation of the active longitude remains constant and the active longitude itself is a persistent quasi-rigidly rotating phenomenon. Usoskin et al. (2007) proposed that a seemingly migrating AL may occur as a result of interaction between the equatorward propagating dynamo wave and a quasi-rigidly rotating non-axisymmetric active zone. The role of differential rotation is also controversy. Various studies concluded that the differential rotation is not the reason of the migration of the AL (Balthasar 2007; Juckett 2006, 2007; Gyenge et al. 2014).

In our previous work (i.e. Gyenge et al. (2016), hereafter GY16), we found evidence supporting the third group of models. Moreover, the migration of ALs does not appear to correspond very well to the form of the 11-year cycle as suggested by previous studies (Usoskin

et al. 2005; Berdyugina et al. 2006). The half-width of the active longitudinal belt is fairly narrow during moderate activity but is wider at maximum activity. We also found that the AL is not always identifiable.

Several studies (Warwick 1965; Bai 2003a,b) suggested that the spatial distributions of eruptive solar phenomena also show non-axisymmetric properties. Bai (1987, 1988) analysed the coordinates of energetic solar flares based of 5 years of time period and concluded that longitudinal spatial distribution is non-homogeneous. Zhang et al. (2007, 2008, 2015) concluded that the dominant and co-dominant AL contains 80% of C- and X-flares. In GY16, we conducted a similar study based on four solar cycles and we did not find significant co-dominant activity; instead, we found that only the dominant AL contains 60% of the solar flares.

The flares and CMEs could occur independently of each other. Numerous CMEs have associated flares but several non-flaring filament lift-offs also lead to CME (Gosling *et al.* 1976; Harrison 1995). Furthermore, in the case of the 'stealth' CMEs there are no easily identifiable signatures to locate the source of the eruption on the solar surface (Howard 2013). Hence, separate flare - CME spatial distribution investigations are justified.

2. CME AND SUNSPOT DATA

We used the SOHO/LASCO HALO CME catalog by the CDAW data centre[1]. The CME catalog spans over 20 years, i.e. between 1996 and 2016. This is the most extensive catalog that contains the source location of CMEs. Only halo CMEs are reported, i.e. their angular width is 360 degrees in C3 coronagraph Field of View. Gopalswamy *et al.* (2009) describes the catalog in great details. The CME source is defined as the centre of the associated active region. The source of the CME is identified using SOHO EIT running difference images. Later, the ability of the STEREO mission to observe the backside of Sun was used to identify the source of back-sided halo CMEs (Gopalswamy *et al.* 2015). This catalog also provides the space speed of CMEs which is the actual speed with which the CME propagates in the interplanetary space. The Plane of Sky speed obtained from the single SOHO viewpoint, converted into space speed using the cone model (Xie *et al.* 2004).

The Debrecen Photoheliographic Data[2] (DPD) sunspot catalogue is used for estimating the longitudinal position of the AL. The catalog provides information about the date of observation, position and area for each sunspot. The precision of the position is 0.1 heliographic degrees and the estimated accuracy of the area measurement is ∼10 percent.

3. LONGITUDINAL SUNSPOT DISTRIBUTION

The first step of our identification procedure is to divide the solar surface by 18 equally sliced longitudinal belts. Hence, one bin equals to a zone with $20°$ width. We take into account all sunspot groups from the moment when they reach their maximum area. This filtering criteria is chosen for the following reasons: Firstly, if we select all sunspot groups at every moment of time then

the statistics will be biased by the long-lived sunspot groups. Secondly, the maximum area of the sunspot groups is a well-defined and easily identifiable moment. This is different from the used practice of considering only the first appearance of each sunspot group. Let us define the matrix W by:

$$W_{\lambda,\mathrm{CR}} = \frac{A_{i,\mathrm{CR}}}{\sum_{i=1}^{n} A_{i,\mathrm{CR}}}. \tag{1}$$

Here, the area of all sunspot groups ($A_{i,\mathrm{CR}}$) is summed up in each longitudinal bin for each CR. $A_{i,\mathrm{CR}}$ is divided by the summarised area over the entire solar surface ($\sum_{i=1}^{n} A_{i,\mathrm{CR}}$). The range of W must be always between 0 and 1, depending on the local appearance of the activity. In the case of $W_{\lambda,\mathrm{CR}} = 1$, all of the flux emergence takes place in one single longitudinal strip. A 3σ significance level threshold is applied to filter the noise. Moving average is also applied for data smoothing with a time-window of 3 CRs.

We standardised the matrix $W_{\lambda,\mathrm{CR}}$ (defined in Eq 1) by removing the mean of the data and scaling to unit variance. Then, cluster analysis was performed for grouping the obtained significant peaks. Here, the DBSCAN clustering algorithm was chosen which is a density-based algorithm. The method groups together points that are relatively closely packed together in a high-density region and it marks outlier points that stand alone in low-density regions (for details see Ester *et al.* (1996)). The parameter epsilon ($= 0.2$) defines the maximum distance between two points to be considered to be in the same group. The parameter m ($= 3$) specifies the desired minimum cluster size. Clusters, containing less than three points, were omitted. The longitudinal location of the clusters ($\lambda_{cluster,\mathrm{CR}}$) represent the position of the AL.

Panel A of Figure 1 shows an example of the initial identification steps outlined above. The sample time period covers 6 years, and it corresponds to 80 CR between 1920 CR and 2000 CR. The quantity W is represented by the shades with blue colour. The dark blue regions denote the significance (3σ) presence of activity. The brighter shades stand for a weaker manifestation of activity (2σ and 1σ). The grey squares shows the sunspot group clusters.

In Panel B of Figure 1, the Carrington longitudes of the most significant cluster is transformed into Carrington Phase for each CR:

$$\psi_{\mathrm{Al, CR}} = \frac{\lambda_{\mathrm{cluster, CR}}}{360}. \tag{2}$$

The range of the quantity ψ_{AL} must be between 0 and 1 where $\psi_{\mathrm{AL}} = 1$ represents the entire circumference of the Sun. In Panel B of Figure 1, Panel A is repeated three times so that we are able to track the migration of the activity thought the phases. To the data Panels A and B, we applied a polynomial least squares fitting based on multiple models. Linear, quadratic, cubic and higher-order polynomial models were tested. The quadratic or parabolic regression ($Ax^2 + Bx - C$) shows the best goodness of fit, hence this model is chosen. Table 1 shows the coefficients and uncertainties.

The shape of migration clearly follows parabolic-shaped path as found in several earlier studies (Berdyugina et al. 2006; Berdyugina & Usoskin 2003; Zhang et al.
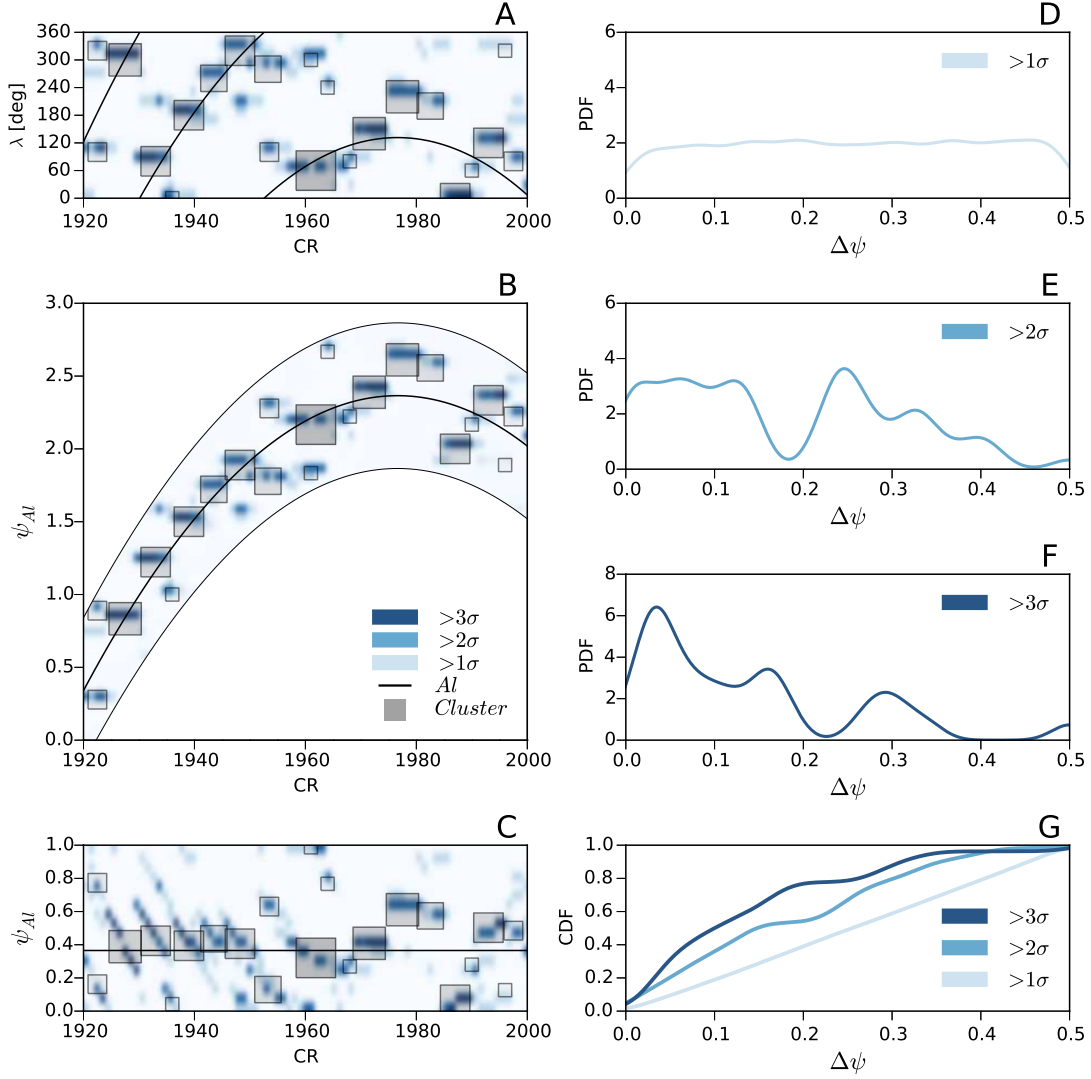
FIG. 1.— The Panels A, B and C show an example of the migration of the active longitude between 1920 CR and 2000 CR (01/03/1997-20/02/2003) based on data of the solar northern hemisphere. The shades of blue indicates the significance of sunspot group activity. The grey squares show high-density areas, i.e. the detected clusters. The solid black line represents the migration patch, fitted by the least-square-method and considering only the most significant (above 3σ) clusters. Panel A shows the observed longitudinal distribution of sunspots in Carrington coordinate system. Panel B, similar to Panel A, depicts solar circumference (Carrington phase) repeated three times. Panel C is the phase-corrected migration path. The parabolic migration pattern is now transformed to a constant line, and provides an insight into the non-homogeneous spatial property. The panels A and C use the same colour scale as defined in Panel B. The colour scale is displayed in the lower right corner of the Panel B. Panels D, E and F show the sunspot distribution around the active longitude (corresponding to $\Delta\psi = 0$) with different significance levels taken for the entire time period and for both hemispheres. The horizontal axes ($\Delta\psi$ - Carrington Phase Difference) represent the shortest distance between the migration of enhanced longitudinal activity and given sunspot groups. Panel G is the cumulative distribution of the above three PDF.

2011a; Usoskin et al. 2007, 2005; Gyenge et al. 2014).

In Panel C of Figure 1, let us now plot the parabola-shaped migration path in a coordinate system that follows the parabolic shape of Panel B. Hence, the active longitude is easily recognisable without repeating Carrington Phases. Several new features are noticeable in Panel C; parallel aligned and tilted lanes between CR 1930 and CR 1950. These shapes are artificial, created by the coordinate system transformation. In the new dynamic coordinate system the AL stands still as represented by the black regression line.

Panels D, E and F of Figure 1 show the kernel density estimation (KDE) of the spatial difference between the AL ($\psi_{AL,CR}$) and the longitudinal position of individ-

ual sunspot groups (ψ_{CR}) in Carrington phase difference ($\Delta\psi$, see Eq. 3) applied to the entire period investigated (i.e between 06/01/1997 and 30/12/2015) and for both hemispheres.

$$\Delta\psi^* = |\psi_{Al,\ CR} - \psi_{CR}|,$$
$$\Delta\psi = \begin{cases} \Delta\psi^*, & \text{if } \Delta\psi^* \leq 0.5, \\ 1 - \Delta\psi^*, & \text{if } \Delta\psi^* > 0.5. \end{cases} \quad (3)$$

The KDE is a non-parametric method to estimate the probability density function (Connolly 2000). We used a Gaussian kernel function. The optimal value of the Gaussian kernel bandwidth is 0.02. If $\Delta\psi = 0$ the sunspot group is located the AL. In case of $\Delta\psi = 0.5$, the loca-

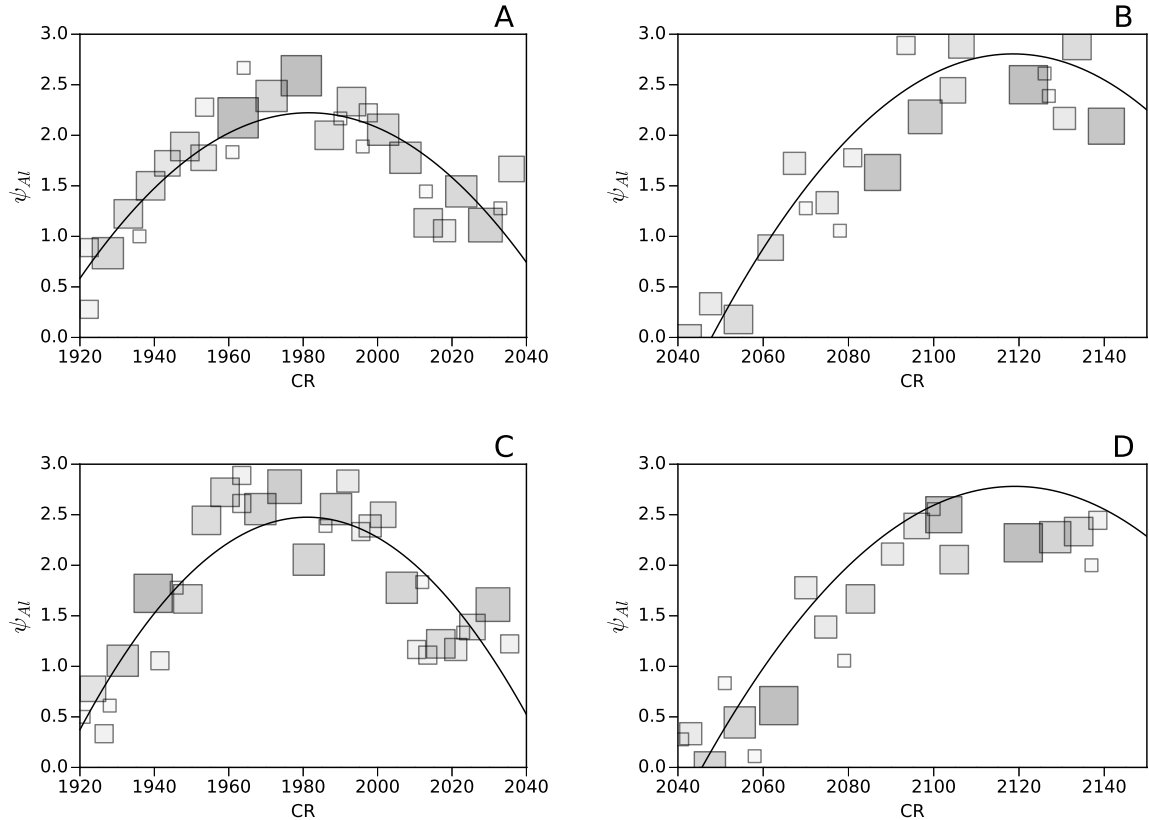| Solar Cycle	Hemisphere	Coefficient A	Coefficient B	Coefficient C
23	North	$-4.32 \times 10^{-4} \pm 5.44 \times 10^{-5}$	$1.71 \pm 2.15 \times 10^{-1}$	$1.69 \times 10^3 \pm 2.13 \times 10^2$
24	North	$-5.61 \times 10^{-4} \pm 1.78 \times 10^{-4}$	$2.37 \pm 7.47 \times 10^{-1}$	$2.51 \times 10^3 \pm 7.83 \times 10^2$
23	South	$-5.62 \times 10^{-4} \pm 6.62 \times 10^{-5}$	$2.23 \pm 2.62 \times 10^{-1}$	$2.20 \times 10^3 \pm 2.59 \times 10^2$
24	South	$-5.16 \times 10^{-4} \pm 1.85 \times 10^{-4}$	$2.18 \pm 7.77 \times 10^{-1}$	$2.31 \times 10^3 \pm 8.12 \times 10^2$

Fig. 2.— The four panels demonstrate the phase-corrected migration path based on the entire analysed period. The Panels A and C show the AL in Solar Cycle 23 and the Panels B and D in Solar Cycle 24. Panels A and B demonstrate the northern hemisphere data and Panels C and D show the southern hemisphere. The grey squares visualise the result of the DBSCAN algorithm, i.e. the detected and significant sunspot group clusters. The quadratic model function is applied. The coefficients and uncertainties are given in Table 1.

tion of the sunspot groups are shifted by 180 degrees from the position of the AL. This shifted position is marking the area of the co-dominant AL. The three KDEs are obtained from different sunspot group datasets with different threshold levels. Panel F of Figure 1 shows the most significant regions (3σ significance threshold was applied) which tend to be formed near the active longitude. The less significant (2σ) sunspot groups show a more disperse distribution. Below the 1σ threshold, the KDE does not show obvious peaks. Panel G of Figure 1 depicts cumulative distribution functions (CDF) of the three KDEs. Note that 70% of the most significant sunspot groups appear closer than $\Delta\psi = 0.17$. This value corresponds to a longitudinal zone with ±60 degrees of width around AL. The Figure 2 show the migration path of the AL based on the entire analysed period. The coefficients and uncertainties of the employed quadratic model functions are displayed in Table 1.

4. COMPLEXITY PROPERTIES OF AL SUNSPOTS

4.1. Separateness parameter

The relationship between flare occurrences and the morphological properties of sunspot groups is widely accepted and is reported in numerous studies (e.g. Schrijver et al. 2007; Cui et al. 2006; Mason & Hoeksema 2010; Korsós et al. 2014, 2015b). The first classification scheme (Hale et al. 1919) is published by Waldmeier (1938). The scheme examines the role of the size and morphology of sunspot groups in relation of determining the capacity of their flare-productivity. The system was modified by Waldmeier (1947) and is known today as the modified Zürich classification system (Kiepenheuer 1953). The classification was further developed by McIntosh (1990) and its version is still in use widely today. Later the classification was automated by (Colak & Qahwaji 2008). However, the classification of the Zürich or McIntosh system is still subjective; further there are just a limited
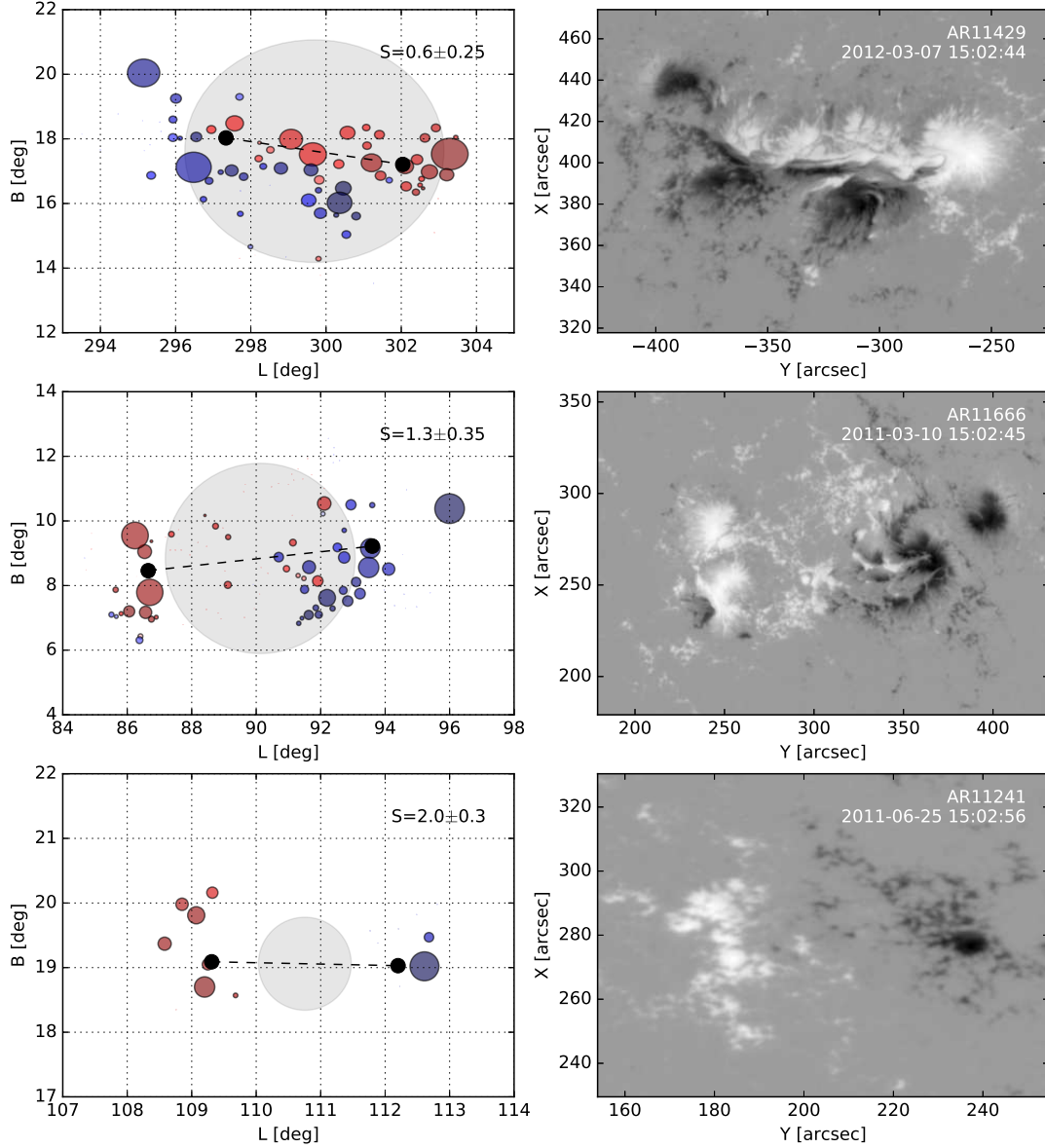
FIG. 3.— Representation of the reconstructed sunspot groups (left-hand side) using the Debrecen Sunspot Data (DPD) catalogue applied to active regions NOAA 11429 (07/03/2012 15:02:44), NOAA 11666 (0/03/2011 15:02:45), NOAA 11241 (25/06/2011 15:02:26). The polarities of the spots are distinguished by the different colours (red and blue). The grey colour is the hypothetical circle, having summarized area of the group. The dashed line is the distance between the following and leading subgroups. The area is measured in MSH (millionths of solar hemisphere). The sunspots are corrected for foreshortening. On the right-hand side, the magnetograms of the example sunspot groups are displayed.

number of classes (Bornmann & Shaw 1994).

The separateness parameter (S) is used to reveal the morphological properties of the sunspot groups near and far from the AL. This parameter was introduced by Korsós et al. (2015b) and its investigation and application to flares showed that the separateness parameter can be a numerical indicator/precursor besides the traditional (Zürich, McIntosh and Mount Wilson) classifications of sunspot groups. The parameter S is considered as an indicator for the potential flaring outbreak and CME capability of active regions.

The separateness parameter is determined by the angular distance between the area-weighted centres of the leading and following subgroups divided by the angu-

lar diameter of a hypothetical circle whose area is equal to the total area of all umbrae constituting the sunspot group.

The angular distance is the shortest distance between two points on the surface of a sphere. The distance (in degrees) between the leading and following subgroups is provided by the spherical law of cosines:

$$\Delta\theta = 2\arcsin\left[\sin^2(\frac{|B_l - B_f|}{2})+ \right.$$
$$\left. + \cos(B_l)\cos(B_f)\sin^2(\frac{|L_l - L_f|}{2})\right]^{\frac{1}{2}}. \qquad (4)$$

Here, B and L refer to the heliographical latitude and longitude of l leading and f following subgroups. If the absolute difference is greater than 180 ($|L_l - L_f| > 180$), then the absolute difference is $360 - |L_l - L_f|$.

The corrected area of individual sunspots (A^*) in millions of solar hemisphere is converted to Mm^2, using:

$$A = \frac{1}{2}(4\pi R_{Sun}^2)10^{-7}A^*, \qquad (5)$$

where R_{Sun} is in Mm. The total sunspot group area (T) is calculated. The number of sunspots in a certain sunspot group is represented by the quantity n. The total area (Mm^2) means the summed up area of the individual sunspots:

$$T = \sum_{i=1}^{n} A_i. \qquad (6)$$

The diameter of an individual sunspot group, ($\Delta\Omega$), is estimated by:

$$\Delta\Omega = \frac{2\sqrt{T/\pi}}{2(R_{Sun}\cos(\frac{1}{2}(B_l + B_f))\pi)}360^o. \qquad (7)$$

The numerator is the diameter (Mm) of a hypothetical circle whose area is equal to the T total area. The denominator represents the circumference of small circle (Mm), which connects all locations with a given latitude. The fraction is multiplied by 360 degrees, equals to the angular distance between the endpoints of the sunspot group diameter in degrees.

Finally, let us define the dimensionless separateness parameter:

$$S = \frac{\Delta\theta}{\Delta\Omega}. \qquad (8)$$

In Figure 3, typical active regions NOAA 11429, 11666 and 11241 (from top to bottom) are selected to demonstrate the usefulness of the parameter S. The visualisation of the active regions is plotted on the left-hand side. The blue and red colours distinguish the different magnetic polarities, the radius of a circle represents the area of the spot. The black dots indicate the weighted average position of the leading and following subgroups. The black dashed line between the black dots is the calculated angular distance ($\Delta\theta$), described by the Eq. 4. The grey circle around the spots is the hypothetical circle whose area is equal to the total area of all umbrae constituting the sunspot group ($\Delta\Omega$), defined by Eq. 7. Panels on the right-hand side are the snapshots (HMI magnetogramm by SDO) of the active regions.

The upper two panels of the Figure 3 are visualising a complex sunspot group (namely, NOAA 11429). This sunspot group is a beta-gamma-delta magnetic configuration according to the Mount Wilson classification. The AR is extremely complex, having umbrae of opposite polarity within the same penumbra. The calculated separateness parameter is 0.6±0.25 (Eq. 8). The middle panels of Figure 3 show NOAA Active Region 11666, which is a moderate complex sunspot group ($S = 1.3 \pm 0.35$). The bottom panels display NOAA Active Region 11241, a less complex bipolar sunspot group ($S = 2.0 \pm 0.30$)

Based on the study by Korsós et al. (2015b), there is a hight risk of X-class flare and/or fast CME occurrence(s) if $S < 1$. There is a moderate risk of flaring (M-class) and/or CME occurrence(s) if $S > 1$ and $S < 2$. In case of $S > 3$, only bipolar sunspot groups appear with relatively simple morphological properties. These ARs have a rather low probability of a significant flaring (above GOES C-class) or CME activities.

4.2. Separateness parameter within AL

In this section, we investigate the longitudinal spatial distribution of the parameter S (Eq. 8) and the area of the investigated sunspot groups. The spatial distance of sunspot groups ($\Delta\psi$) is defined by Eq. 3.

The raw area measures are converted to standard scores (standardized statistics). The standard score is a dimensionless quantity calculated by subtracting the average of sunspot group area ($\overline{T}$) from a given group area T_i and dividing by the sample-corrected standard deviation ($\sigma(T)$) of the area data:

$$Z_i = \frac{T_i - \overline{T}}{\sigma(T)}. \qquad (9)$$

Spatial multi-variable (linear) interpolation is applied to $f(S, \Delta\psi, Z)$. The method results in a regular matrix (M) form unstructured 3-dimensional data. The range of $\log S$ is $[-1, 0.5]$, and is divided by 1500 equal bins (n). The range of $\Delta\psi$ $[0, 0.5]$ is divided by 500 bins (m), i.e.

$$M = \begin{bmatrix} Z_{1,1} & Z_{1,2} & ... & ,Z_{1,n} \\ Z_{2,1} & Z_{2,2} & ... & ,Z_{2,n} \\ ... & ... & ... & ,... \\ Z_{m,1} & Z_{m,2} & ... & ,Z_{m,n} \end{bmatrix}. \qquad (10)$$

Figure 4 shows the results of the statistics. The panels A, B and C are obtained for data of the northern hemisphere (red-coloured plots) and panels D, E and F are that of the southern hemisphere (blue-coloured figures). In panels A (northern hemisphere) and F (southern hemisphere), the matrix M (Eq. 10) is visualised. The horizontal axis is the distance from AL. The vertical axis is the logarithm of parameter S; $\log S < 0$ stands for a high-risk flare or CME occurrence. The colour code is the standard score of sunspot group area. The red or blue shades indicate positive standard score ($T_i > \overline{T}$). The white color stands for negative standard score ($T_i < \overline{T}$ or no data). Panels B and E are the row averages of matrix M. Panels C and D are the column averages of matrix M.

In Panel A, significant islands are visible between $0 > \log S > -0.5$ at $\Delta\psi < 0.1$. There is one more obviously visible island around $\Delta\psi = 0.2$ at $\log S < -0.4$. However, above $\Delta\psi > 0.2$ there is no remarkable island. Panel C of Figure 4 also reveals a peak below $\Delta\psi < 0.1$. The statistics suggests, that the most complex and largest sunspot groups appear near the AL.

Analysis of the data of the southern hemisphere data show similar results. There are easily notable islands below $\Delta\psi < 0.2$ at $\log S = -0.6$. Panels D and E of Figure 4 clearly reveal that significant sunspot groups appear only near the AL.
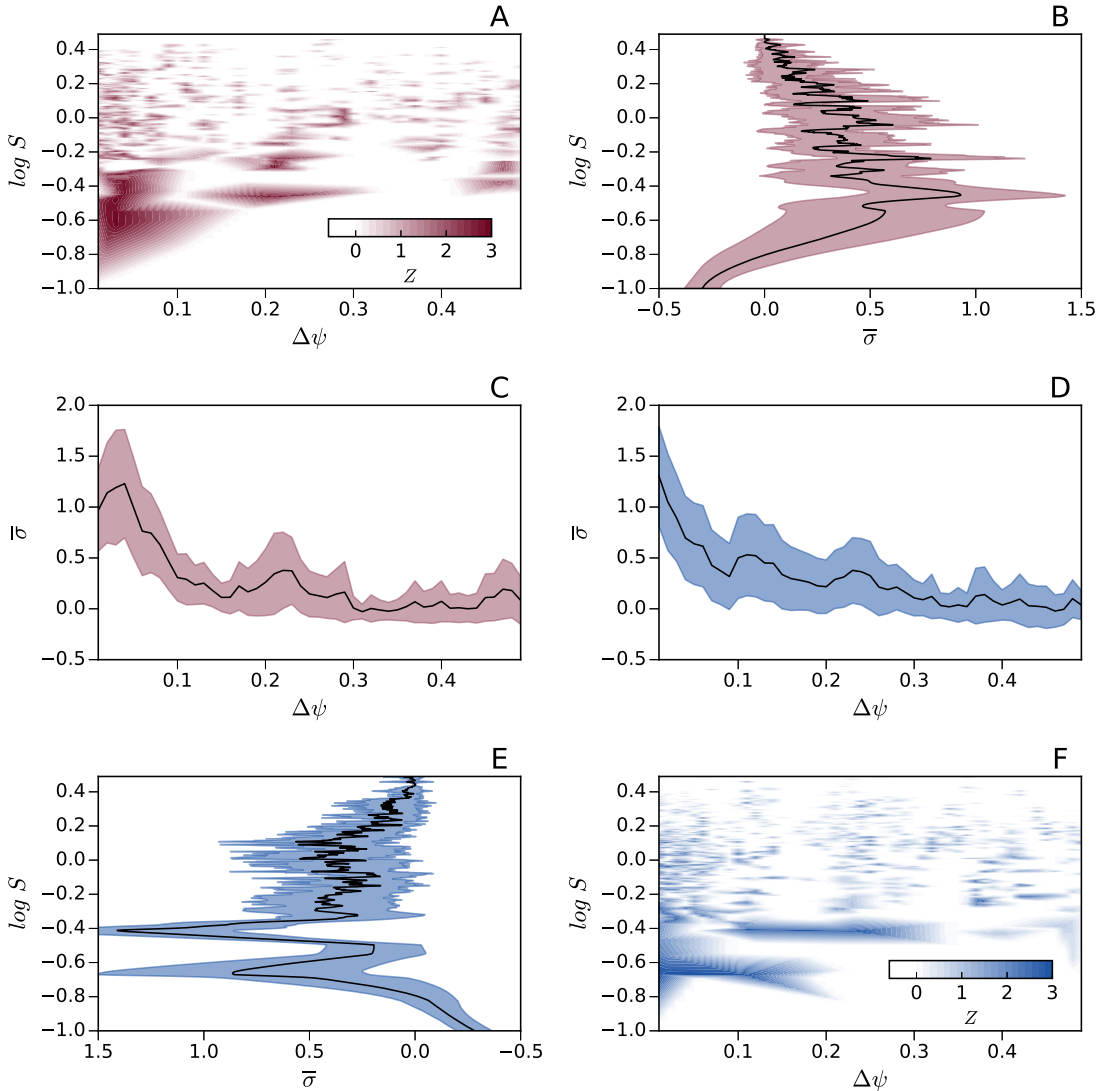
FIG. 4.— The separateness - phase - standard score statistic based on the northern and souther hemisphere, indicated by blue and red colours. The panels A and E shows the separateness of sunspot groups versus the longitudinal location of the active region. The shade of red and blue colours indicate the positive standard score. The Panels B, C, D and E are the cross-section PDF, produced by the KDE method.

Both statistics suggest that the most complex active regions tend to cluster near the AL The co-dominant AL around $\Delta\psi = 0.5$ does not have a significant activity. This statistical investigation also highlights a non-equivalent AL and co-dominant AL activity.

4.3. *Tilt angle of investigated Active Regions within AL*

The last investigated morphological property of ARs, here, is the sunspot group tilt angle. The definition of the tilt angle γ^* is given by Howard (1991);

$$\gamma^* = \mathrm{atan}\left(\frac{(B_f - B_l)/(L_f - L_l)}{\mathrm{sign}(|B_f| - |B_l|)\cos(B)}\right). \qquad (11)$$

Parameters B and L are the Carrington latitude and longitude. The following and leading subgroups have subscripts f and l, respectively.

Figure 5 demonstrates the relationship found between

the separateness parameter S and scaled tilt angle $\gamma = |\gamma^*/90|$. Only most significant ARs are taken into account; the area of sunspot groups has to be at least 2σ greater than average. The results applicable to the two hemispheres are distinguished by the colours of dots.

Linear regression cannot be used because the data is associated with a considerable uncertainty both in the X and Y directions (Isobe et al. 1990). For that reason, Principal Component Analysis (PCA) is used to fit the dataset. The PCA method is a linear dimensionality reduction keeping only the most significant singular vectors to project the data to a lower dimensional space (Einbeck *et al.* 2007). The eigenvector of the first component $e_1 = [-0.8387, 0.5445]$ shows the direction of the maximum variance of the data ($\sigma^2 = 4.1227$), i.e where the data is most spread out.

Based on the result of the PCA we performed Principal Component Regression (PCR). In Figure 5, the solid black line is the regression fit and the grey halo repre-
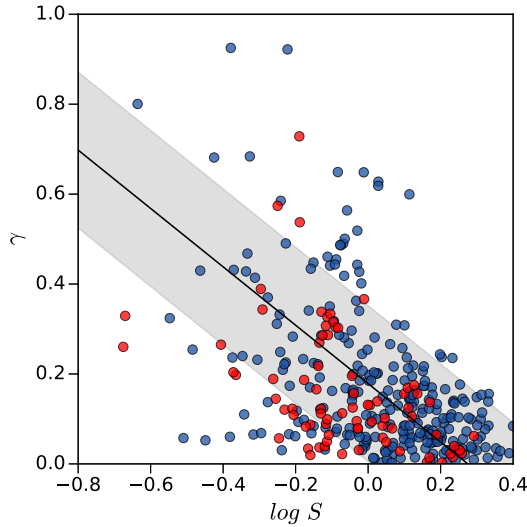
FIG. 5.— The tilt angle versus separateness parameter of sunspot groups for each hemispheres. Data from northern and souther hemispheres, respectively, are distinguished by blue and red dots. The yellow arrows visualise the result of the PCA method.
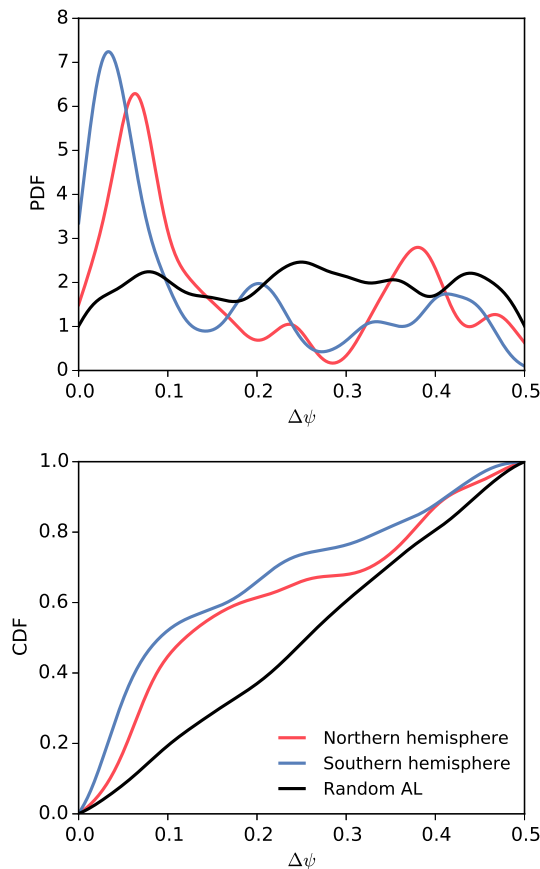

FIG. 6.— The upper/lower panel shows the PDF/CDF for CME occurrence. In the upper panel, the blue and red coloured lines are the PDF, obtained by KDE method, distinguished by the northern and southern hemispheres, respectively. The black line is from determining the PDF based on random-generated AL. In the lower panel, the CDFs of the above defined functions are shown.

sents a 1σ standard deviation of the sample along the regression line. The difference between the samples of two hemispheres is statistically insignificant, therefore the regression was applied to the data of both hemispheres. The obtained statistics suggests that there is a clear relationship between the tilt angle and the separateness parameter of the sunspot groups.

5. ENHANCED LONGITUDINAL BEHAVIOUR OF CME EVENTS

5.1. Spatial probably CME occurrences

In this section, the connection between the CME occurrence and AL is revealed. Panel A of Figure 6 shows the kernel probability density function of the longitudinal distribution of CME occurrence. This statistics is based on data from both the northern and southern hemispheres. There is only one significant peak visible around $\Delta\psi = 0.05$. Besides this remarkable peak, there is a long plateau with some insignificant local peaks. Only one more peak, above the significance level at $\Delta\psi = 0.4$, is present, but with a relatively weak activity when compered to the first peak.

A random-generated control group is also used in this statistic. The longitudinal position of AL now is a random position. This test was inspired by Pelt et al. (2005) who expressed a critical view on the identification method of Berdyugina & Usoskin (2003) employed for AL. In our study, we applied the methodology introduced by Pelt et al. (2005), who reconstructed the distribution of AL with random sunspot longitude data. The KDE plot of the control group does not show any peaks. This homogeneous distribution means that AL identification does not cause false significant peaks, which would affect the results.

Panel B of Figure 6 shows the cumulative distribution of the above-defined spatial distributions. The blue and red lines have a steep increasing phase between values of 0 and 0.1 followed by a less steep increasing trend. These results allow us to estimate that most of CMEs (around 60%) occur in a $\pm 36°$ belt around the position of AL. Hence, the width of the longitudinal belt of CME occurrences is equal to the width of the longitudinal belt of solar flare occurrences (GY16). The black line is the cumulative distribution obtained from the analysis applied to random longitudinal positions. This distribution would only contain 20% of CMEs. This latter finding means that AL plays a significant role in the spatial distribution of CME occurrences.

5.2. CME dynamics

Let us now consider the apparent and space velocities (V_a and V_s) of CME events. Two-dimensional kernel density estimations are applied with an axis-aligned bivariate Gaussian kernel, evaluated on a square grid of the $\Delta\psi - V_a$ and $\Delta\psi - V_s$ space. Figure 7 shows the result, based on data from both hemispheres. The significance levels $1 - 3\sigma$ are indicated by coloured contour lines.

In both panels of Figure 7, there are four islands above the 1σ significance level. The statistics shows that the source of fast CMEs (speeds between 1500 km/s and 3000 km/s) is indeed an active region, located within AL. However, slow (i.e. speed less than 1500 km/s) CMEs can occur outside of AL.
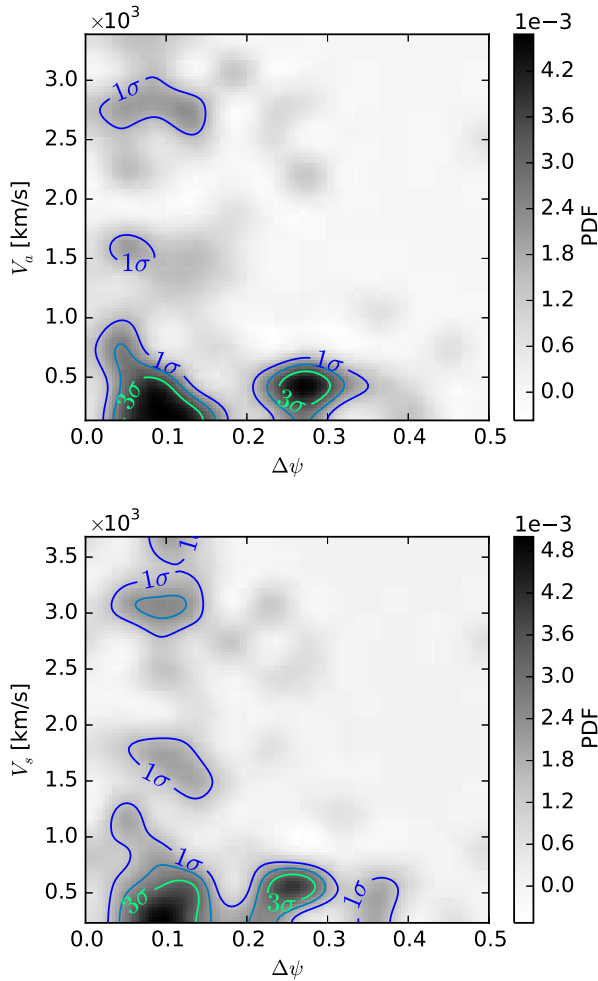
FIG. 7.— The result of the two-dimensional KDE, using parameters $\Delta\psi$ and V_a (apparent velocity of CME, upper panel), V_s (space velocity of CME, lower panel). The shade of the grey colour represent the probability density. The significant islands are indicated by blue (1σ), dark green (2σ) and bright green (3σ) colours. The northern and souther hemispheres are not distinguished from each other.

Above the significance level of 3σ, there are only two islands. These are only slow CMEs inside and outside of AL. Analysis of this statistics also indicates that the probability of a slow CME is two standard deviation units higher than the probability that of a fast CME.

6. DISCUSSION

The AL identification method presented here reveals new spatial properties of the longitudinal distribution of the sunspot groups (Panels D, E, F of the Fig 1). The spatial distribution of smaller sunspot groups (less then 1σ) is homogeneous. There is no enhanced longitudinal belt identifiable based on small sunspots; small groups appear everywhere as function of longitude. Moderate sunspot groups (between $1 - 2\sigma$) show already inhomogeneous properties. However, these results still have to be treated with caution. Only sunspot groups above the 3σ significance level have signatures of obvious and remarkable inhomogeneous spatial distribution.

The idea of two, almost equally significant longitudinal

zones is widely accepted by numerous studies, see e.g. Berdyugina & Usoskin (2003). The dominant and co-dominant active longitude is separated by 180 degrees (Zhang et al. 2011b; Bumba et al. 2000). However, we do not find such equally strong ALs (neither here nor in GY16). In our investigation, the co-dominant AL plays a less important role.

The spatial distribution of the separateness parameter (defined by Eq. 8) shows that complex active regions with a high CME capability appear near mostly the AL (Figure 4). The appearance of moderate and simple complex configurations are everywhere on the solar surface. These groups are also able to have CMEs with a significantly lower probability.

We also found that, the most tilted sunspot groups have a complex configuration (Figure 5). Simple bipolar sunspot groups show relatively small tilt angle. Sakurai (2003) and Canfield (2012) concluded, that there is positive correlation between magnetic helicity and sunspot tilt angle. The sunspot rotation could play important role in helicity transport across the photosphere. Sunspot rotation may increase helicity in the corona leading to flares and CMEs (Pevtsov 2012). This property may also have the consequence of a more complex built-up of the underlying magnetic structure, and, the well-studied magnetic arches of the upper solar layer could be oriented at a large angle to the equator (Grigorev 2012). The more complex active regions are the more flares and they will be associated with CMEs (Huang et al. 2013; Jetsu et al. 1997; Kitchatinov & Olemskoi 2005). Hence, we conclude that the above physical process can take place within AL and anywhere else but only with low probability there.

Several studies have investigated east-west asymmetry of CMEs occurrence. Skirgiello (2007) found asymmetry using data provided by the SOHO-LASCO $1996 - 2004$. The asymmetric behaviour could be a consequence of AL. Our result obtained here shows (see e.g. Fig 7) that the number of CME occurrence is marginally higher within AL. The mean of the apparent and space velocity of CME occurrences is around 500 km/s considering the entire surface of the Sun. This mean velocity is known an 'slow' CMEs, found in a number of earlier studies (Ying et al. 2016). However, the mean velocity is significantly higher if only the AL itself is considered. Within AL the average (or space) velocity is around 1000 km/s (see e.g. Michalek et al. 2009). There is no fast CME occurrence found outside of AL. Therefore, interestingly and notably, the fast HALO CMEs are also AL CMEs.

7. CONCLUSION

Our new findings (together with the results of GY16) could provide novel aspects both for space weather forecast and for solar dynamo theory. Usually, the flare and/or CME prediction tools are based on only the behaviour of active regions, such as complexity of magnetic fields or other morphological properties. However, the spatial distribution of active regions can also assist in forecasting as suggested by e.g Zhang et al. (2008). We conclude, that the main source of CME and solar flare (GY16) occurrences is the AL. Hence, the detection of this enhanced longitudinal belt may allow us to find the most flare- and CME-capable regions of the Sun preceding the appearance of an active region. This potential

flare and CME source is predictable even several solar rotations in advance.

The observed properties of the non-axisymmetric solar activity need to be taken into account in developing and verifying suitable dynamo theory: the observations analysed here show that there is only one significant AL with a relatively wide ($\pm 20 - 30$ degrees) belt. Furthermore, the tilt angle of the active regions is also an important observed constraint for dynamo theory: the tilt angle of sunspot groups shows non-axisymmetric behaviour, which is a completely new (and surprising) finding.

REFERENCES

Bai, T., 1987, ApJ, 314, 795
Bai, T., 1988, ApJ, 328, 860
Bai, T., 2003a, ApJ, 585, 1114
Bai, T., 2003b, ApJ, 591, 406
Balthasar, H., 2007, A&A, 471, 281
Balthasar, H., Schüssler, M., 1983, Solar Phys., 87, 23
Balthasar, H., Schüssler, M., 1983, Solar Phys., 93, 177
Baranyi T. MNRAS, 2015
Becker, U., 1955, Z. Astrophys., 37, 47
Berdyugina, S.V., Usoskin, I.G., 2003, A&A, 405, 1121
Berdyugina, S.V., 2004, Solar Phys 224, 123-131
Berdyugina, S.V., 2005, ASP Conference Series, 346
Berdyugina, S.V., Moss, D., Sokoloff, D.D., Usoskin, I.G.. Astron. Astrophys. 445, 703714, 2006.
Brouwer, M. P., Zwaan, C., 1990, Solar Phys. 129, 221
Bumba, V. et al. 1965, The Astrophys. J., 141, 14921517.
Bogart, R.S.: 1982, Solar Phys. 76, 155 165.
Bornmann, P.L. & D. Shaw, V.I., 1994, Sol. Phys., 150, 127
Bumba, V., Obridko, V. N., 1969, Solar Phys., 6, 104
Bumba, V., Garcia, A., Klvaňa, M., 2000, Solar Phys., 196, 403
Carrington, R. C., 1863, Observations of the spots on the Sun from November 9, 1853, to March 24, 1861 made at Redhill, Williams and Norgate, London,, Section VI, p. 246.
Castenmiller, M. J. M., Zwaan, C., van der Zalm, E. B. J., 1986, Solar Phys. 105, 237
Canfield, R. C., & Pevtsov, A. A. 1998, Sol. Phys., 182, 145
Chidambara, A. 1932, Mon. Not. R. Astron. Soc., 93, 150152.
Cui, Y., Li, R., Zhang, L., He, Y., & Wang, H. 2006, SoPh, 237, 45
Colak, T. & Qahwaji R., 2008, Sol. Phys., 248, 277
Connolly, A. J., Genovese, C., Moore, A. W. Nichol, R. C.,Schneider, J. and Wasserman, L., 2000, ArXiv Astrophysics e-prints, http://adsabs.harvard.edu/abs/2000astro.ph..8187C
Gyenge, N., Baranyi, T., Ludmány, A., 2014, Solar Phys., 289, 579
Gyenge, N., Baranyi, T., Ludmány, A., 2016, ApJ, 818:127 (8pp)
Gopalswamy, N.; Yashiro, S.; Michalek, G.; Stenborg, G.; Vourlidas, A.; Freeland, S.; Howard, R., Earth, Moon, and Planets, Volume 104, Issue 1-4, pp. 295-313
N. Gopalswamy, H. Xie, S. Akiyama, P. Mkel, S. Yashiro, and G. Michalek ; The Astrophysical Journal Letters, 804:L23 (6pp), 2015 May 1 ; doi:10.1088/2041-8205/804/1/L23
Gosling, J.T., Hildner, E., MacQueen, R.M., Munro, R.H., Poland, A.I. & Ross, C.L., 1976, Sol. Phys., 48, 389 - 397
Grigorev, V., M., Ermakova, L., V. and Khlystova, A., I., 2012, Astronomy Reports, 56, 878
Győri, L., Baranyi, T., Ludmány, A. 2011, IAU Symp. 273, 403
Einbeck, J., Evers L., Bailer-Jones, C., Gorban, A., Kegl, B., Wunsch, D., Zinovyev A., Lecture Notes in Computational Science and Engineering, Springer, 2007, pp. 180–204
Hale, G. E., E., Ferdinand, Nicholson, S. B. & Joy, A. H. 1919, ApJ, 49, p.153
Harrison, R.A., 1995, Astron. Astrophys., 304, 585
Henney, C. J., Durney, B. R., 2005, ASPC, 346, 381
Howard R., 1991b, Solar Phys., 136, 251
Howard, T. A. & Harrison, R. A. 2013, Sol. Phys., 285, 1-2, 269-280
Huang, X., Zhang, L., Wang, H., Li, L., 2013, A&A, 549, 127

Isobe T., Feigelson E. D., Akritas M. G., Babu G. J., 1990, ApJ, 364, 105
Ivanov, E. V., 2007, Adv. Spa. Res. 40, 959
Jetsu, L., Pohjolainen, S., Pelt, J., Tuominen, I., 1997, A&A, 318, 293
Juckett, D. A., 2006, Solar Phys., 245, 37
Juckett, D. A., 2007, Solar Phys., 245, 37
Kitchatinov, L. L., Olemskoi, S. V., 2005, Astron.Lett., 31, 280
Kiepenheuer, K.O., 1953, The University of Chicago Press, p.322
Korsós, M., B., Ludmány, A.; Baranyi, T. 2015, Apj, 789, 107
Korsós, M., B., Ludmány, A.; Erdélyi, R.; Baranyi, T. 2015, Apj Lett, 802, L21
Korsós, M., B.; Erdélyi, R.; 2016, Apj
Losh, H. M., 1939, Publ.Obs.Michigan, 7, no. 5., 127-145
Mason, J. P., & Hoeksema, J. T. 2010, ApJ, 723, 634
Maunder, E. W., 1905, MNRAS, 65, 538
McIntosh, P. S., 1990, Sol. Phys., 125, 251
Michalek, G.; Gopalswamy, N.; Yashiro, S., Sol. Phys., Volume 260, Issue 2, pp.401-406
Mumford, S. *et al.* 2013, SunPy: Python for Solar Physicists, Proceedings of the 12th Python in Science Conference, 74-77.
Obridko, V N; Chertoprud, V E; Ivanov, E V. Solar Physics, 272.1 (Aug 2011): 59-71.
Pelt, J., Tuominen, I., Brooke, J., 2005, A&A, 429, 1093
Pelt, J., Brooke, J. M., Korpi, M. J., Tuominen, I., 2006, A&A, 460, 875
Pevtsov A., Astrophysics and Space Science Proceedings, Volume 30. ISBN 978-3-642-29416-7. Springer-Verlag Berlin Heidelberg, 2012, p. 83-91
Sakurai, T. and Hagino, M., 2003. Journal of The Korean Astronomical Society, 36, 7-12
Schrijver, C. 2007, ApJL, 655, L117
Skirgiello, M. 2005, Annales Geophysicae, 23, 31393147
Usoskin, I. G., Berdyugina, S. V., Poutanen, J., 2005. A&A, 441, 347
I.G. Usoskin, S.V. Berdyugina, D. Moss, D.D. Sokoloff ,2007, Advances in Space Research 40. 951958
Waldmeier, M., 1938, Zeitschrift für Astrophysik, 16, 276
Waldmeier, M., 1947, Publ. Zürich, Band IX, Heft 1.
Warwick, C. S., 1965, ApJ, 141, 500
Zhang, L.Y., Wang, H.N., Du, Z.L., Cui, Y. M., He, H., 2007, A&A, 471, 711
Zhang, L.Y., Wang, H.N., Du, Z.L., 2008, A&A, **484**, 523-527.
Zhang, L.Y., Mursula, K., Usoskin, I. G., Wang, H.N., 2011, A&A, 529, 23
Zhang, L.Y., Mursula, K., Usoskin, I. G., Wang, H.N., 2011, JASTP, 73. 258
Zhang, L.; Mursula, K.; Usoskin, I., 2015, A&A L., 575
Ester, M., H. P. Kriegel, J. Sander, and X. Xu, In Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, OR, AAAI Press, pp. 226231. 1996
Warwick, C. S, 1966, ApJ, 145, 215
Xie, H., Ofman, L., Lawrence, G., Cone model for halo CMEs: Application to space weather forecasting, J. Geophys. Res., vol. 109, 2004, A03109.
Ying et al., The Astrophysical Journal Supplement Series, 222, 2 , http://stacks.iop.org/0067-0049/222/i=2/a=23