Recurrence Large Sunspot Structures and 27-Day Component of Solar Activity as Proxies to Axis-Nonsymmetry

Alexander Shapoval 1,2,* and Mikhail Shnirman 3

1 Department of Mathematics and Computer Science, University of Lodz, 90-238 Lodz, Poland
2 Cybersecurity Center, Universidad Bernardo O’Higgins, Santiago 8370993, Chile
3 Institute of Earthquake Prediction Theory and Mathematical Geophysics RAS, 117997 Moscow, Russia;
shnir@mitp.ru
* Correspondence: abshapoval@gmail.com

Abstract: The purpose of this paper is to design tools that quantify the structure of the nonaxisymmetric component of the solar magnetic field. With the Fourier transform and the machine learning identification of recurrent objects, we define the 27-day component of solar proxies and recurrent large sunspot structures (ReLaSS), respectively. These two closely related characteristics are established to represent different components of the asymmetry of the solar magnetic field. We derive that the 27-day component and ReLaSS have anticorrelated since 1970 after dozens of years of a strong correlation. The persistence of the correlation sign during few solar cycles reflects yet unknown regularities of solar activity. The contribution of both proxies to the nonsymmetry of solar activity is shown to be lower in 1990–2010 than ∼100 years earlier. This property may be the trace of the asymmetry at the scales that are longer than the centennial Gleissberg cycle.

Keywords: axis-nonsymmetry; active longitudes; Gleissberg cycle; correlation

1. Introduction

Modern predictions of the solar cycle involve a persistent longitudinal asymmetry exposed by solar activity [1]. The origin of this asymmetry can be explained by the existence of the nonaxisymmetric enhancement of the underlying magnetic field [2]. The longitudinal asymmetry is related to the concentration of a strong magnetic field at specific locations for weeks, months, years, and even longer [3–7]. The asymmetry can be detected with the Fourier transform and wavelet analysis [8–11]. The specification of the time, spatial characteristics, and the methods of identification leads to somewhat different quantities including sunspot nests [12] nestlets [13], active regions [14], active longitudes [15], and activity complexes [16]. Direct observations of large sunspots, which are likely to be caused by the concentrated strong magnetic field, can be followed by the identification of recurrent objects rotating with the Sun and returning to the solar disk [13,17,18].

A direct and rather successful method of solar dynamo modeling applies the mean field approximation to the magnetohydrodynamic equations and implements additional effects related to the break of symmetries including the sustainable transformations between the poloidal and toroidal magnetic fields of the Sun [19–25]. While the modeling is performed in two dimensions, the calibration of 3D-spatial effects related to the symmetry breaking are impeded. As a positive example, we mention the paper [26] that constructed a model generating a large magnetic field rigidly rotating with the Sun. Clustering in time as a model output is induced through assumptions regarding the sunspot emergence and the diffusion rate [27,28]. Three-dimensional surface flux transport models with the parameters responsible for clustering of fluxes in space and time are calibrated [29,30] and used for the prediction of the solar cycle [31].

Researchers turn from findings of specific features exhibited by the nonsymmetric component of solar activity to the formulation of regularities extended to the decadal
Typically, the nonsymmetric component is dominated by symmetric one [33], but a temporal large growth in the longitudinal asymmetry is detected [2]. This asymmetry is primarily associated with large sunspots [34,35]. The papers [8,9,36] focused on the Fourier frequencies corresponding to the \( \sim 27 \) day rotation period of the Sun as it is observed from the Earth to analyze the component of the solar activity just related to the solar rotation. Le Mouël et al. [37] combined these frequencies into a packet of waves and, investigating the dynamics of the spectrum part defined with this packet, found that its growth had preceded the rise of the whole activity which occurred in the 1930s–1940s by 1–2 solar cycles. Identifying recurrent sunspot groups, Henwood et al. [13] exposed that the lifetime of large sunspot groups increased in the 1920s–1940s in \( \sim 1.4 \) times as was earlier found implicitly in [38]. The 80–100-year trend of the lifetime of large sunspots conjectured in [13] stays in agreement with the wave of the Gleissberg cycle that, as known [32,39], exhibits secular variations of solar activity and modulates the solar indices. The traces of the modulations that can be related to the Gleissberg cycle are seen in other quantities exposing active regions [40,41].

The purpose of this paper is to compare the dynamics of two quantities describing the longitudinal asymmetry: (i) the packet of waves being introduced by [37] and exhibiting a 27-day spectrum component of the underlying time series and (ii) a composite derived from the recurrent large sunspot structures (ReLaSS). ReLaSS are obtained with the algorithm, closely related to those proposed by Henwood et al. [13] and Nagovitsyn et al. [18], who identified mostly recurrent sunspot groups. The comparison is performed on the decadal-to-centennial scale. We expect to establish a general correlation between the two quantities, which agrees with the unity of mechanisms staying behind their formation. The break of correlation would expose new features of the longitudinal asymmetry of solar activity.

2. Data

We work with the daily ISSN index (from 1870 onwards) of solar activity provided by WDC-SILSO, Royal Observatory of Belgium, Brussels [42], referred to further as \( W(t) \), where \( t \) indicates the day of observations.

The daily areas of sunspot groups constitute another well-known index of solar activity. The Royal Greenwich Observatory (RGO) provided the data from 1874 until 1976. The 1977–2020 data were compiled by US Air Force (USAF) from the Solar Optical Observing Network (SOON) with the support of the National Oceanic and Atmospheric Administration (NOAA) We used RGO/USAF/NOAA daily data available at solarcyclescience.com/activeregions.html (accessed on 26 May 2023) [43]. Following Hathaway [44], we multiplied the 1977–2020 data by factor 1.4 to reduce the non-homogeneity related to the change in the observatory in 1976–1977. The sum of the areas of sunspot groups observed at the day \( t \) is denoted by \( D(t) \).

The RGO/USAF/NOAA records contain the location of the sunspot groups. This allows us (as well as other authors [13,18]) to derive large recurrent sunspot groups, i.e., those that are likely to return to the solar disk at least once and characterized by a sufficiently large maximal area. Our identification algorithm is fully described in [45] and sketched in the Appendix A. The result of the identification is referred to as recurrent large sunspot structures (ReLaSS). The sum of the areas of sunspot groups observed on the day \( t \) and included into ReLaSS is denoted by \( L(t) \). As the ReLaSS are derived from the sunspot groups, \( L(t) \leq D(t) \). We note that ReLaSS significantly contributes to the areas of all sunspots. For example, the sum of daily \( L(t) \)-values found with 22-year windows constitutes the 0.15–0.30 part of that of \( D(t) \).

3. Methods

3.1. 27-Day Signal

We analyze several daily proxies of solar activity deriving from them a component related to the \( \sim 27 \)–day solar rotation. This component, called a packet of waves after [37], is computed with the Fourier transform within sliding windows of \( N = 8109 \) days (approx-
imately 2 solar cycles or slightly more than 22 years). Let \( \Omega(t) \) be a solar proxy of interest (\( W(t), D(t), \) or \( L(t) \) described in Section 2) and \( \{\hat{\Omega}(k, t)\}_{k=1}^N \) be the discrete Fourier transform of \( \{\Omega(t')\}_{t'=-N/2+1/2}^{t'=t+N/2-1/2} \) computed within the window \( [t-N/2+1/2, t+N/2-1/2] \). Then, for each sliding time window \( \{t'\}_{t'\in[t-N/2+1/2, t+N/2-1/2]} \), the packet of waves \( E_\Omega^{(27)}(t) \) is obtained through the selection of the Fourier orthogonal harmonics related to the periods from the 25.7–28.5 range, the summation of the squares of the corresponding magnitudes, and the assignment of the normalized sum to the center \( t \) of the window:

\[
E_\Omega^{(27)}(t) = \frac{1}{N^2} \sum_{N/k\in[25.7,28.5]} |\hat{\Omega}(k, t)|^2.
\]

The computation scheme including the range of the periods, which corresponds to the solar differential rotation, remains in line with [37]. The usage of the 22-year sliding windows makes the 11-year variations subtle and allows us to focus on longer trends.

The packet of waves is compared with the whole energy represented by

\[
E_{\Omega}(t) = \frac{1}{N} \sum_{j=-(N-1)/2}^{(N-1)/2} (\Omega(t+j))^2, \quad \Omega(t) \in \{D(t), L(t), W(t)\},
\]

and computed within the same sliding windows. The contribution of the 27-day component is estimated with the ratio \( r_\Omega(t) \) of the packet of waves to the whole energy:

\[
r_\Omega(t) = \frac{E_\Omega^{(27)}(t)}{E_{\Omega}(t)} \quad \Omega = W \text{ or } \Omega = D,
\]

### 3.2. A Few Proxies to Axis-Nonsymmetry

The daily \( r_W(t) \) and \( r_D(t) \) ratios computed with the sunspot and sunspot group indices, \( \Omega = W \) and \( \Omega = D \), represent the dynamics of the axial asymmetry exhibited by the solar magnetic field. Indeed, these quantities are affected by magnetic fluxes appearing at specific locations for few \( \sim 27 \)-day solar rotations (as our Fourier transform-based computation measures the presence of objects which appear and disappear once per 27 days). We note that the composite \( L(t) \) found with the areas of ReLaSS also describes the axial asymmetry of the solar magnetic field. According to Figure 1, the composites \( E_D, E_D^{(27)}, \) and \( E_l \) have much in common, including a general two-peak pattern and the fall at the 20th cycle. Yet, with this scale of observations, a stronger high-frequency component is seen in the dynamics of the total energy \( E_D \) (red curve).

![Figure 1](image-url)
The previous discussion implies that the ratio

\[ r_L(t) = \frac{E_L(t)}{E_D(t)}. \]  

(2)

is expected to exhibit some similarities with \( r_D \) defined in (1). The distinctions between the two ratios would expose differences between the impact of magnetic fluxes, estimated with the Fourier transform, and large objects observed in approximately the same location during few solar rotations. Targeting these distinctions, the ratio

\[ r^* = \frac{r_D}{r_L} = \frac{E_D^{(27)}}{E_L}. \]  

(3)

is defined.

3.3. Verification of Significance

The significance of the values of \( r^* \) is estimated with the following stochastic model that destroys the day-to-day dependence between the data but keeps the average values of the time series within yearly intervals. The value of the artificial time series at day \( t \) is generated in two steps: first, the number of groups with the Poisson random variable \( \zeta(t) \) and, second, the area of each group. The mean of \( \zeta(t) \) is set to the average daily number of the groups observed on \( I_t = [t - 364/2, t + 364/2] \). The area of each group assigned to day \( t \) is randomly drawn from the sample of the areas from the RGO/USA/NOAA or the ReLaSS catalogue. We define \( r^* \) by the analogue of (3), where the numerator is computed with the artificial time series, whereas the denominator \( E_L \) remains unchanged, still being computed with the real catalog. Small values of \( r^* \) with respect to \( r^* \) would support the significance of the \( r^* \)-values. The elimination of dependencies between the elements of time series is frequently used to validate the statistical significance of empirical procedures; we mention paper [46], where the authors assess the earthquake clustering, as another example.

4. Results

The contribution of the 27-day component to solar activity until the end of the 1970s was first exposed in paper [37]; see their Figure 8 and our Figure 2a,b. The authors found a growth of the 27-day component 1–2 solar cycle prior to a general rise of solar activity. Extending their graph with the modern data, we observe a slow steady decay of \( r_W(t) \) during the last 3 cycles, which indicates a gradually decreasing role of the 27-day component (gray curve in Figure 2a). On the one hand, the level of the \( r_W \) ratio attained during cycles 21–23 is just slightly below that observed in 1885–1905. This feature agrees with the existence of 70–100-year oscillations of the solar activity called the Gleissberg cycle. On the other hand, both ratios \( r_W \) and \( r_D \) at the right points of Figure 2a are lower than at the left points, which may be related to regularities at the scales that are (significantly?) longer than the Gleissberg cycle.

The contribution of the 27-day component estimated with the RGO data follows a similar pattern to that of \( r_W \) since 1910. The most remarkable difference is in the absence of small values of \( r_D(t) \) in 1885–1900 (red curve in Figure 2). We also indicate a larger variability of \( r_D(t) \) than \( r_W(t) \) on the month-to-year scale. In particular, cycles 21–23 are not characterized by a steady decay of \( r_D(t) \).

Figure 3 exhibits the similarity and dissimilarity of the contribution of the 27-day component and the derived ReLaSS to solar activity by displaying the dynamics of the ratios \( r_L(t) \) and \( r_D(t) \). Splitting the time interval into three parts: \( t \in [0, 14818] \) corresponding to 1885–1926, \( t \in [14818, 17740] \) corresponding to 1934–1967, \( t > 17740 \) corresponding to 1934–2020—we report the coefficients of correlation equaled to 0.47, 0.95, and −0.64, respectively. The inclusion of cycle 16 with a short episode of anti-correlation to the first interval reduces the correlation. Thus, the correlation of \( r_L \) and \( r_D \) is almost exact on cycles
13–19 except a couple of years related to cycle 16. A larger variability of $r_D$, including a drop in cycle 18, underlies a slight distortion between the dynamics of $r_L$ and $r_D$. However, a turn to the epoch of a strong anti-correlation occurred in cycle 20.

![Figure 2](image)

**Figure 2.** (a) Ratios $r_\Omega = E_{\Omega}(27)/E_\Omega$, $\Omega = W$ and $D$, of the energy of the packet of waves to that of the total energy (in gray) computed with ISSN, $\Omega = W$ (in gray), and RGO/USAF/NOAA, $\Omega = D$ (in red). (b) Evolution of $E_{\Omega}(27)(t)$ and $E_\Omega(t)$ separately.

![Figure 3](image)

**Figure 3.** Ratio $r_D$ (red curve) of the packet of waves to the averaged squared area of all sunspot groups and ratio $r_L$ (blue curve) of averaged squared area of large recurrent groups to the averaged squared area.

We perform another comparison between the 27-day component of the composite based on the sunspot group areas and ReLaSS by displaying their ratio $r^*$ defined by Equation (3) in Figure 4. This ratio exhibits two ranges of values: 0.07–0.11 prior to 1926 and 0.04–0.07 after 1926. The transition was fast around 1926. Remarkably, anomalous cycle 20 is characterized by a quick growth of $r^*$ to the first range of values followed by a decay. Thus, relatively large values of $r^* = E_{D}(27)/E_L$ are attained during the episodes of a weaker activity.

We formulate and solve the problem addressing to what extent the values of $r^*$ exhibit factual effects. In this way, the values of the investigated time series $D(t)$ are randomized to keep 1-year averages but eliminate day-to-day correlations. The ratio $r^*$ constructed with this artificial time series (see Section 3) in the same method as $r^*$ is constructed with $D(t)$ is displayed in Figure 4 in gray. The values of $r^*$ go 1 order of magnitude below than that of $r^*$. Therefore, the changes in several hundredths observed in $r^*$ are significant.

The last available data used in the construction of $r^*$ correspond to the current activity, which is close to the minimum of the Gleissberg cycle. Therefore, one may expect an uprising trend of $r^*$ towards the 0.07–0.11 interval in the nearest future.
Figure 4. Ratio $r^* = \frac{E_{D}^{(27)}}{E_L}$ of the packet of waves to the averaged squared area of derived recurrent objects (ReLaSS, in black); $r^*$ (in gray) is the same but constructed with the model series, which lack the day-to-day dependence between the data, to exhibit the significance.

5. Discussion

The components of solar proxies related to the longitudinal asymmetry, in general, follow the trends exhibited by the proxies themselves after 22-year smoothing, which eliminates the 11 and 22-year cycles (Figure 1). This is in line with a rather good agreement between different composites derived from the daily observations of the Sun [47]. The source of the disagreement between the 27 day component of the RGO and ISSN series in 1885–1904 seen with the normalized quantities (Figure 2) is unclear. One may also assume its counterfactual nature. A rise in the 27-day component after 1915, which preceded a general growth of solar activity, (the drop after cycle 17 in Figure 3) was discovered in [38].

We highlight an almost flat part of this component observed in cycles 21–23, whereas the whole activity and its components related to large sunspots evolved to the minimum of the Gleissberg cycles detected in, e.g., paper [39].

We compared the measurement of the longitudinal asymmetry exploring the ratio of $E_{D}^{(27)}$ to $E_L$ obtained with the Fourier transform and recurrent large sunspot structures (ReLaSS), respectively (Figure 4). The variations of this ratio are typically small, which signals that $E_{D}^{(27)}$ and $E_L$ measure the longitudinal asymmetry in a similar way. Nevertheless, the episodes of low solar activity are characterized by larger values of the ratio. A somewhat reduced rate of ReLaSS during a weak activity can be explained by the choice of the absolute lower border of the large sunspot groups. It is worth adjusting and applying the definition of large sunspot groups relative to the whole activity. One can also test to what extent the phenomenon of active longitudes is understood as the appearance of (possibly remote) persistent magnetic fluxes along longitudes disagrees with large objects rigidly rotating with the Sun. Our finding may manifest that the solar magnetic field is prone to causing the existence of persistent remote fluxes along the longitudes. According to Vernova [35], 82% of the longitudinal asymmetry is generated by the sunspots with the area from 100 to 2000 MSH. However, during weak activity with the shortage of large sunspots, the asymmetry is probably transformed into just active longitudes. One may expect an upward trend of the $r^*$-ratio in Figure 4 as the Gleissberg cycle is progressing. Our method of analysis cannot estimate the time frame of this upward trend; the next 11 years of observations are desired to verify this hypothesis. The discussion regarding the Gleissberg wave is justified to some extent by not only the normalization of the series with factor 1.4 (as Hathaway [48] suggested) but also the fact that the numerator and the denominator of all our ratios were affected after 1976 similarly.

We establish that both components $E_D$ and $E_L$ related to longitudinal asymmetry, derived with the Fourier transform and ReLaSS, correlate almost everywhere in 1885–1967 (Figure 3). This is quite expected as the components exhibit the common phenomena. A large correlation in 1935–1960 is partly explained by a general growth of the sunspot areas, written in the denominator of the ratios $r_D$ and $r_L$. Nevertheless, the episode of
correlation is wider, including the time of a weak activity around 1900 and the data from the rather small cycle 20. We argue that the large correlation signals that the derivation of the components $E_L$ and $E_D$ of the solar activity from data is performed properly. However, the derived components anti-correlate in 1967–2009: a relative growth in the 27-day component is accompanied by a drop of the ReLaSS component and vice versa. We stress that the correlation and anti-correlation between the two quantities $r_D$ and $r_L$ are not directly related to either the level of these quantities or the ratios between them, or the level of whole activity (Figure 3. This feature is yet to be explained).

The characteristics of the longitudinal asymmetry found with the active regions and the Fourier formalism exhibit a certain persistence within a few solar cycles. The peculiarities of solar activity at this time scale are awaited to be included in the general modeling of solar dynamo represented by, e.g., papers [19,20,23,33].

6. Conclusions

We focus on the structure of the nonsymmetrical component of the solar magnetic field and expose the dynamics of the proxies to this structure at the centennial scale. The proxy are the 27-day component of the daily activity and recurrent large sunspot structures (ReLaSS). They exhibit similarities as being related to the same phenomenon and dissimilarities that probably correspond to the tendency of relatively moderate sunspots to specific longitudes exhibited during specific time intervals. The following results are obtained:

1. The solar activity and the contribution of ReLaSS to it was small in 1915–1935, whereas the 27-day component contributed disproportionately strongly at that time. The latter might have signaled the preparation of a forthcoming growth of the solar activity observed in the 1940s;
2. The 27-day component and ReLaSS have anticorrelated since 1970 after dozens of years of a strong correlation. The persistence of the correlation sign during few solar cycles reflects yet unknown regularities of solar activity;
3. The contribution of both proxies to the nonsymmetry of solar activity is lower nowadays than 100 years ago. We hypothesize that this property may be the trace of the asymmetry at scales that are longer than the centennial Gleissberg cycle;
4. The Fourier formalism applied to the sunspot indices and large recurrent sunspot groups represent somewhat different components of the asymmetry of the solar magnetic field. Each of these components and the relationship between them exhibits peculiar centennial waves.

The existence of these new regularities rises questions regarding axial asymmetry of the meridional circulation. The found regularities are yet to be explained with solar dynamo modeling and, potentially, used in the prediction of solar cycles.

Author Contributions: Methodology, A.S. and M.S.; Validation, A.S. and M.S.; Formal analysis, A.S.; Investigation, M.S.; Writing—original draft, A.S.; Writing—review & editing, A.S. and M.S.; Visualization, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No data were created in this study.

Acknowledgments: We are very thankful to V. Fufaev for numerous discussions of the manuscript.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

The identification of ReLaSS aims at the extraction of recurrent large sunspot groups by involving the following consideration. As a rule, sunspots drift only slightly in the Carrington coordinates related to the Sun. Therefore, recurrent solar groups leaving the solar disk on the east are expected to return on the west with approximately the same coordinates. Thus, a group emerged at the solar disk can be claimed as the complement of...
a previously observed group (in other words, identified as the same group) if it is observed inside an ellipse centered in the location of this previously observed group. The semi-axes of the ellipse constitute two parameters of the algorithms. Nagovitsyn et al. [18] chose these parameters based on the data regarding the migration of the sunspots across the Sun.

In contrast to [18], we identify only large recurrent groups (the area exceeds 950 MSH at least once). Their location can change significantly when a part of their sunspots appears or disappears. Accounting for this effect, we increase the size of the ellipse. This enlarges the number of the cases when more than a single group is detected inside the ellipse. The multiplicity problem is solved by the introduction of an additional rule that matches the area of the group and its complement. The semi-axes of the ellipse are adjusted on the sunspot groups that are observable on the solar disk at 7–12 days. In more detail, we scan the values of the semi-axes on these sunspot groups, choosing those that lead to the best identification in terms of the so-called F1-score. Narrowing the training set from the groups that are observing 12 consecutive days on the solar disk to those that are observed 11, . . . , 7 days, we establish the insensitivity of the chosen values to this reduction. This approach advocates for the application of the adjusted values of the semi-axes to the real identification that deals with the return of groups to the solar disk after ~14 days. The groups identified by the algorithm are called recurrent large structures in the text.

References

22. Maiernski, E.V.; Malova, H.V.; Popov, V.Y.; Zelenyi, L.M. Ulysses Flyby in the Heliosphere: Comparison of the Solar Wind Model with Observational Data. Universe 2022, 8, 324. [CrossRef]