Solar Physics (2026) 301:34
https://doi.org/10.1007/s11207-026-02632-2

RESEARCH

Quantifying Sunspot Group Nesting with Density-Based Unsupervised Clustering

Nurdan Karapınar[1] · **Emre Işık**[2] · **Natalie A. Krivova**[2] · **Hakan V. Şenavcı**[3]

Received: 10 December 2025 / Accepted: 13 February 2026
© The Author(s) 2026

Abstract

Sunspot groups often emerge in spatial–temporal clusters, known as nests or complexes of activity. Quantifying how frequently such nesting occurs is important for understanding the organisation and recurrence of solar magnetic fields. We introduce an automated approach based on kernel density estimation and DBSCAN clustering to identify nests in the longitude–time domain and to measure the fraction of sunspot groups that belong to them. The method combines a smooth representation of emergence patterns with a density-based clustering procedure, validated using synthetic solar-like cycles and corrected for variations in data density.

We apply this method to 151 years of sunspot-group observations from the Royal Greenwich Observatory Photoheliographic Results (RGO, 1874 – 1976) and Kislovodsk Mountain Astronomical Station (KMAS, 1955 – 2025) catalogues. Across all cycles and latitude bands, the mean nesting degree is $\langle D \rangle = 0.61 \pm 0.12$, implying that about 60 percent all sunspot groups emerge within nests. Nesting is strongest at mid-latitudes ($10^\circ - 20^\circ$), and results from the two independent datasets agree in the period of overlap. The nesting degree significantly correlates with the solar activity level, with the correlation strengthening when small groups are excluded. The characteristic inter-nest spacing contracts from $\sim 200 - 500$ Mm at low activity to $\sim 60 - 100$ Mm at solar maximum, approaching typical sunspot-group dimensions. The identified nests range from compact clusters to long-lived, drifting structures, offering new quantitative constraints on the persistence and organisation of solar magnetic activity.

✉ E. Işık
isik@mps.mpg.de

N. Karapınar
nurdankarapinar@gmail.com

1 Department of Astronomy and Space Sciences, Graduate School of Natural and Applied Sciences, Ankara University, Ankara, Türkiye

2 Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany

3 Department of Astronomy and Space Sciences, Faculty of Science, Ankara University, Ankara, Türkiye

Published online: 27 February 2026

1. Introduction

Sunspots appear in the solar photosphere as visible signatures of magnetic-flux bundles emerging from the solar interior. These bundles typically form structures with positive and negative polarities, which outline the footpoints of rising magnetic-flux loops (van Driel-Gesztelyi and Green 2015; Weber et al. 2023). The emergence of sunspot groups (SGs) is one of the primary indicators of solar magnetic activity at photospheric levels.

It has long been recognised that sunspot groups do not emerge randomly in longitude. Instead, they tend to cluster in space and time, forming so-called nests or complexes of activity. Early statistical work by Becker (1955) and later analyses of synoptic magnetograms by Gaizauskas et al. (1983) highlighted the recurrent nature of these emergence patterns (Becker 1955 called them *Fleckenherde*, or 'spot herds'). These clusters, which may persist for several solar rotations, were found to exhibit characteristic longitudinal motions that deviate from the local (differential) rotation profile in both directions and occasionally show diverging and converging behaviour over time.

A systematic and quantitative treatment of nesting was initiated by Castenmiller, Zwaan, and van der Zalm (1986), who applied cluster analysis to the Greenwich sunspot group data. Their analysis revealed statistically significant nests typically spanning less than 30° in longitude and persisting up to 15 solar rotations. Brouwer and Zwaan (1990) extended this work using three-dimensional (time, longitude, and latitude) clustering analysis of Greenwich data, showing that at least one-third of all sunspot appearances belong to intrinsically physical clusters, with characteristic longitudinal and latitudinal spreads of $\sim 2°$ and $\sim 1°$, respectively, and lifetimes of 1 – 7 months. The patterns showed hierarchical organisation that were called 'nested nests' with component separations below 25° longitude.

Benevolenskaya (2005) examined activity nests within the broader context of complexes of solar activity, linking their formation and evolution to the topology of the non-axisymmetrical component of the solar magnetic field. Her analysis connected nesting patterns to large-scale coronal structures and suggested that activity complexes may have roots deep in the tachocline (see also Raphaldini, Dikpati, and McIntosh 2023).

SG nesting was also sometimes connected to the so-called active longitudes, which exhibit persistent activity patterns across multiple rotations with 180° separation between hemispheric activity herds and flip-flop cycles of 3.7 years (Berdyugina and Usoskin 2003; Usoskin, Berdyugina, and Poutanen 2005; Usoskin et al. 2007).

Despite these efforts, the quantitative characterisation of nesting remains challenging (Pojoga and Cudnik 2002). Earlier analyses often relied on subjective criteria or methods that assumed specific cluster shapes or rotation profiles. Moreover, the increasing availability of long-term, homogeneous sunspot catalogues motivates the development of automated and reproducible techniques to quantify clustering behaviour. Modern data-driven machine-learning techniques, including non-parametric density estimation and unsupervised pattern-recognition methods, offer new opportunities to revisit these questions (Asensio Ramos et al. 2023).

In this work, we present a systematic approach to identify sunspot-group nests in the longitude–time domain and to quantify the degree of nesting across solar cycles. We combine a smooth representation of the emergence density with a density-based clustering approach that detects clusters of arbitrary shape, and we validate the method using synthetic solar-like data. Applying this approach to 151 years of the Royal Greenwich Observatory Photoheliographic Results (RGO, 1874 – 1976) and Kislovodsk Mountain Astronomical Station (KMAS, 1955 – 2025) observations, we investigate the statistical properties of nesting as a function of latitude and solar cycle, assess their significance, and discuss the implications for the organisation of solar magnetic fields.

2. Data and Methodology

2.1. Data

We analyse sunspot group (SG) data from two complementary observational programs. Our primary dataset comprises the Royal Greenwich Observatory Photoheliographic Results (RGO, 1874 – 1976), providing systematic observations spanning Cycles 12 – 19 (Willis et al. 2013). To extend coverage through the modern era, we use data from the Kislovodsk Mountain Astronomical Station (KMAS, 1955 – 2025), covering Cycles 19 – 25 (Nagovitsyn, Makarova, and Nagovitsyna 2007).

Both datasets provide heliographic coordinates (latitude, longitude), observed times, and maximum area measurements for each sunspot group, making them suitable for spatial-temporal clustering analysis. The RGO catalog is the established reference standard for sunspot observations during the photographic era, with well-documented observing protocols and systematic daily coverage. KMAS maintains comparable observational standards using similar measurement techniques.

The datasets overlap during 1954 – 1976 (Cycles 19 – 20), enabling direct comparison and cross-calibration. Cross-calibration studies show that KMAS individual group areas are systematically smaller than RGO measurements by approximately 3%, with a scaling factor $b_{\mathrm{RGO-KMAS}} = 1.031 \pm 0.056$ (Mandal et al. 2020). For our analysis, we apply this correction to bring KMAS areas to the RGO scale, ensuring homogeneity across the combined dataset. The overlap period also enables validation of our nesting methodology across independent observational programs (Section 3.2).

Together, these datasets provide 151 years of sunspot group observations (1874 – 2025) covering 13 complete solar cycles plus the ongoing Cycle 25, with RGO contributing Cycles 12 – 19 and KMAS extending through Cycles 19 – 25.

2.2. Methods

In the present work, we analyse sunspot-group statistics in the longitude-time plane, focusing on emergence locations and their phase speeds in the Carrington reference frame, defined to be co-rotating with about $\pm 16^\circ$ latitudes with a sidereal period of about 25.4 days.

To evaluate SG-nesting quantitatively, we employ Kernel Density Estimation (KDE) for visualisation of emergence density patterns (Scott 1979), combined with the DBSCAN (Density Based Spatial Clustering of Applications with Noise) clustering algorithm (Ester et al. 1996) for discrete nest identification. The KDE provides intuitive visual guidance, while DBSCAN assigns each sunspot group to either an identified nest or classifies it as an isolated emergence. As the main quantity that quantifies nesting, we define the *degree of nesting* as the fraction of SGs in nests, derived from each latitude slice's time window, which is a full solar cycle. It is expressed as

$$D_{ij} := \frac{N_{\mathrm{nest}}}{N}, \tag{1}$$

where N_{nest} is the number of SGs in nests and N the total number of SGs in a given time-latitude range, represented by the ith cycle and jth latitude slice.

2.2.1. Kernel Density Estimation

Kernel Density Estimation (KDE) is a non-parametric technique for estimating the probability density function of a dataset by placing a selected kernel function K_H at each data point

and summing their contributions (Scott 1979). For a dataset of n SGs with emergence longitudes and times (ϕ_i, t_i) and maximum observed areas A_i, we define area-weighted weights as $w_{A,i} = \ln(A_i)/\ln(A_{\max})$, where $A_{\max} = \max_j(A_j)$ is the largest SG area in the analysis window. The two-dimensional area-weighted kernel density estimator in the longitude-time space is

$$\hat{f}(\phi, t) = \frac{1}{\sum_{i=1}^{n} w_{A,i}} \sum_{i=1}^{n} w_{A,i} \, K_H \left(\frac{\phi - \phi_i}{h_\phi}, \frac{t - t_i}{h_t} \right), \tag{2}$$

where K_H is a bivariate Gaussian kernel with bandwidth parameters h_ϕ and h_t determined using Scott's rule (Scott 1979). Prior to KDE, both longitude and time coordinates are standardised using z-score normalisation. The area weighting ensures that larger SGs contribute more significantly to the density estimation, reflecting their enhanced magnetic flux content.

In our analysis, KDE serves primarily as a *visualisation tool* to display the underlying density structure of sunspot emergence patterns. The discrete identification of nests is performed by the DBSCAN algorithm described below.

2.2.2. Clustering with DBSCAN

To identify discrete nesting patterns in SG emergence, we employ the Density Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester et al. 1996). Unlike traditional clustering methods such as k-means, DBSCAN does not require specifying the number of clusters beforehand and can identify clusters of arbitrary shapes, making it well-suited for detecting the elongated or irregular nesting patterns expected in solar active region emergence.

DBSCAN operates using two hyperparameters: the neighbourhood radius ε and the minimum number of points m_p required to form a dense region. The algorithm classifies data points (each point being a sunspot group) into *core points* (forming cluster interiors), *border points* (on cluster peripheries), and *noise points* (isolated outliers). In our context, noise points correspond to SGs emerging in isolation, while clustered points represent nest members.

2.2.3. Parameter Selection and Validation

The DBSCAN parameters ε (neighbourhood radius) and m_p (minimum number of points) determine the clustering sensitivity and must be chosen to reflect the intrinsic spatial–temporal structure of sunspot-group emergence. We explored two complementary approaches: (i) an adaptive selection informed by the KDE bandwidth, and (ii) a fixed-parameter approach calibrated using synthetic data. The KDE-informed approach assumes that the smoothing scales appropriate for visualising the emergence density should also approximate the characteristic scales of potential clusters. However, validation tests showed that KDE bandwidths respond strongly to the overall spatial–temporal extent of the data rather than to the intrinsic cohesive scales of nests. As a result, this approach systematically overestimates the clustering strength when the true nesting degree is low ($D \lesssim 0.5$).

We therefore adopted a fixed pair of parameters, $\varepsilon = 0.11$ and $m_p = 3$, chosen after extensive injection-recovery testing (Section 2.3). These values were selected to minimise the mean absolute bias between injected and recovered nesting degrees across the solar-relevant range $0.4 \lesssim D \lesssim 0.7$. The choice $m_p = 3$ reflects the minimum plausible number of

independently observed sunspot groups required to constitute a physically meaningful nest, consistent with earlier definitions based on longitudinal clustering in active longitudes.

Although a fixed ε performs well for typical data densities, sparse data windows, particularly high-latitude bands during low-activity periods, can lead DBSCAN to artificially fragment genuine nests. To account for this, varying data density across different temporal windows and latitude bands, we introduced a sparsity correction that adjusts the effective radius ε:

$$\varepsilon_{\text{eff}} = \begin{cases} \varepsilon_0 \times \min\left(1.5, \sqrt{N_{\text{ref}}/N}\right), & N < N_{\text{ref}} \\ \varepsilon_0, & N \geq N_{\text{ref}} \end{cases} \tag{3}$$

where $\varepsilon_0 = 0.11$ is the baseline parameter, N is the number of SGs in the analysis window, and $N_{\text{ref}} = 350$ is a reference sample size. The reference value was determined empirically: synthetic tests showed that windows with $N \gtrsim 350$ yield stable recovery of nests without excessive fragmentation, so the correction compensates only in sparser cases. This correction increases ε by up to 50% for sparse datasets, preventing excessive fragmentation of genuine nesting structures during low-activity periods. For dense data slices with $N > N_{\text{ref}}$, no correction is applied ($\varepsilon_{\text{eff}} = \varepsilon_0$).

This hybrid procedure – fixed ε with a bounded sparsity correction – produces robust and nearly unbiased recovery of the nesting degree in both synthetic and observational data.

2.2.4. Analysis Windows and Boundary Handling

We process SG emergence data in non-overlapping temporal windows corresponding to full solar cycles (minimum to minimum), subdivided into 5° latitude bands from $-35°$ to $+35°$. This latitude slicing is motivated by the following observations: (i) SG nesting occurs in both longitude and latitude. (ii) SG nests can exhibit various rotation rates in the Carrington frame and in many cases intrinsic phase motion, but no systematic motion in latitude was reported.

A critical consideration is the proper handling of the longitude discontinuity at $0°/360°$. Since DBSCAN identifies spatially coherent structures, nests that traverse the discontinuity would be artificially split if the domain were treated naively. To avoid this, we implement a wraparound correction by replicating sunspot groups near the boundary: groups with longitude $\phi > 275°$ are duplicated at $\phi - 360°$, and groups with longitude $\phi < 85°$ are duplicated beyond $\phi + 360°$. The clustering is then performed on the propagated dataset, but only groups within the original $[0°, 360°)$ range are retained when computing the nesting degree. This ensures that nests are identified consistently irrespective of their location relative to the coordinate singularity.

2.2.5. Statistical Significance

The primary output of our analysis is the nesting degree, defined by Equation 1. To validate the statistical significance of identified nests and avoid spurious clustering from random alignments in sparse data, we test each identified nest against the null hypothesis of uniform random emergence. Under this hypothesis, sunspot groups are distributed uniformly across the longitude-time domain, and the number of groups falling within any subregion follows a Poisson distribution. For each cluster occupying a bounding box of area $A_c = \Delta\phi \times \Delta t$, the expected number of groups under uniform distribution is

$$N_{\text{expected}} = \frac{N_{\text{total}}}{360° T_{\text{window}}} A_c, \tag{4}$$

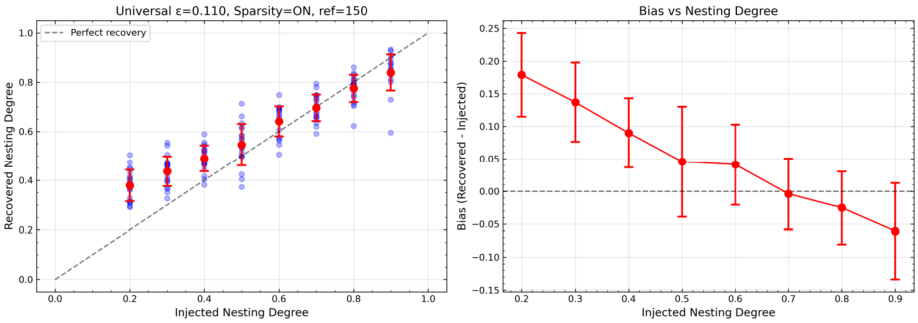


Figure 1 Injection-recovery validation of DBSCAN clustering for synthetic SG-emergence data of solar-like cycles, showing recovered vs. injected nesting degree (left) and the corresponding bias as a function of the injected degree (right). Red symbols show mean values with error bars indicating standard deviation over 160 independent realisations.

where N_{total} is the total number of SGs in the analysis window and T_{window} is its temporal extent. We then calculate the probability of observing n_c or more groups within this area by chance yields

$$p(X \geq n_c | N_{\text{expected}}) = 1 - F_{\text{Poisson}}(n_c - 1; N_{\text{expected}}), \tag{5}$$

where F_{Poisson} is the Poisson cumulative distribution function. Clusters with $p < 0.05$ are classified as statistically significant, indicating that their overdensity is unlikely to arise from random fluctuations.

2.3. Validation with Synthetic Data

To validate our methodology and characterise any systematic biases, we performed injection-recovery tests using synthetic sunspot data with prescribed nesting degrees. The synthetic data generator produces realistic SG distributions including: (i) latitude drift following the butterfly diagram pattern, (ii) area distributions matching observed statistics, and (iii) controlled nesting with specified probability D_{inj} that each new group emerges within an existing nest rather than at a random location.

Figure 1 shows the results of injection-recovery tests for nesting degrees D_{inj} ranging from 0.2 to 0.9. For each injected value, we generated 20 independent realisations and applied DBSCAN with $\varepsilon = 0.11$ and sparsity correction (Equation 3).

The recovered nesting degree shows good agreement with injected values across the tested range, with mean bias $\langle D_{\text{rec}} - D_{\text{inj}} \rangle = 0.062$ and RMS scatter of 0.10. At low nesting degrees ($D_{\text{inj}} < 0.4$), slight underestimation occurs. The overall scatter corresponds to 16relative error for typical $D \sim 0.6$. In the solar-relevant range $D = 0.5–0.7$, the scatter reduces to $\sigma = 0.07$ (11% for $D \sim 0.6$), indicating good precision for typical active region nesting degrees. This validation demonstrates that our fixed-parameter approach with sparsity correction provides robust nesting degree estimates across the range of values expected in solar data.

3. Results

We applied the validated DBSCAN method described in Section 2.2 to RGO and KMAS SG datasets spanning 1878 – 2025, covering 14 solar cycles (Cycles 12 – 25). The analysis

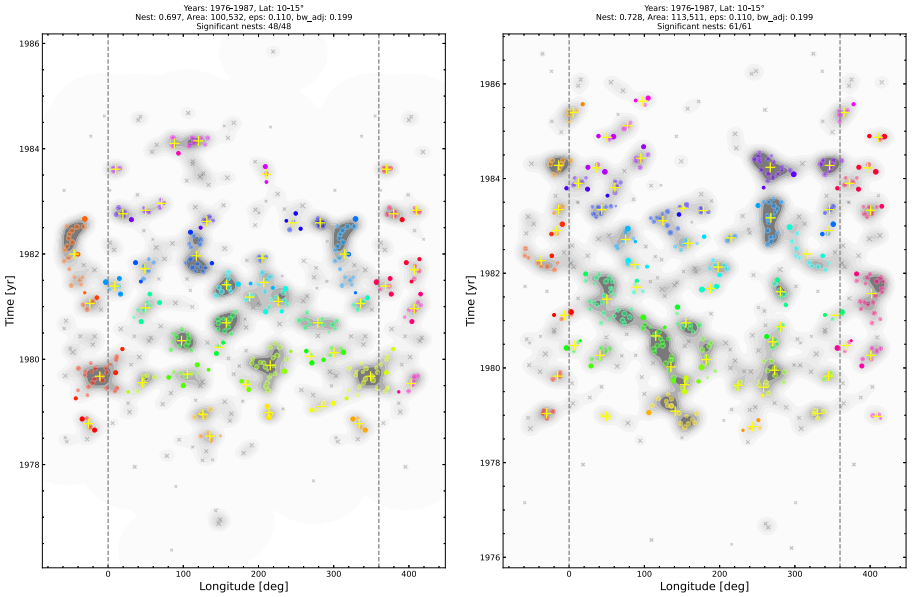


Figure 2 Carrington longitudes and times of SG emergence from KMAS data, during solar cycle 21 across the latitude band $10° - 15°$ of the northern (left panel) and southern (right panel) hemisphere. Grey-scale background shows the SG-area-weighted KDE density distribution where the areas were taken in their peak value. The coloured circles denote nest members and grey crosses non-members according to the area-weighted DBSCAN. Yellow plus signs show the area-weighted centres of nests. The nesting degree, total SG area, adopted ε, and the adjusted KDE bandwidth (bw_adjust), and the number of significant vs. total nests are given on plot titles.

was performed using full solar cycle temporal windows, with each cycle analysed separately over 5° latitude bands from $-35°$ to $+35°$.

3.1. Identification of Sunspot Group Nests

Figure 2 presents representative examples of identified SG nests on the Carrington longitude-time plane. The background contours show the area-weighted KDE density estimation, providing visual context for the underlying emergence patterns. Coloured points represent SGs assigned to individual nests by the DBSCAN algorithm, with cluster centres marked by yellow crosses. Black crosses indicate isolated groups not participating in nesting events.

These examples demonstrate the capability of our method to detect clustering patterns of varying spatial and temporal extent. Some nests are tightly confined (compact clusters), while others show extended temporal persistence spanning multiple years, reflecting the long-lived nature of subsurface magnetic structures that give rise to repeated emergence in preferred longitudes. These preferred longitudes often drift in time, shifting the nest centre at various rates and directions,

3.2. Cross-Validation Between Datasets

To assess the robustness of our methodology and quantify systematic differences between observational programmes, we applied our nesting analysis independently to both RGO

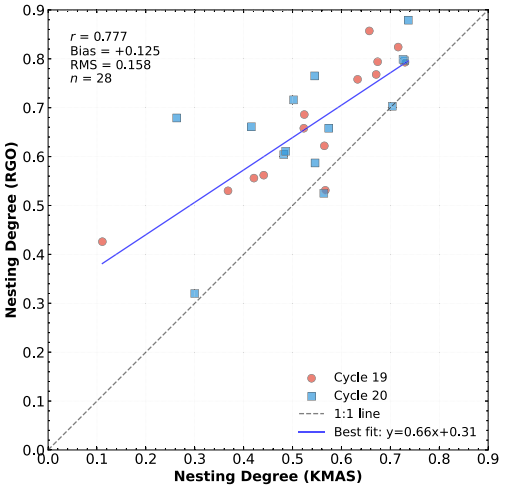


Figure 3 Comparison of nesting degrees from RGO and KMAS during overlap period (Cycles 19 – 20, 1954 – 1976). Red circles: Cycle 19 ($n = 14$); blue squares: Cycle 20 ($n = 14$). Strong correlation ($r = 0.777$) with systematic offset following $D_{\mathrm{RGO}} = 0.66\, D_{\mathrm{KMAS}} + 0.31$ (RMS = 0.158).

and KMAS data during their overlap period (1954 – 1976, covering Cycles 19 – 20). Both datasets were processed using identical analysis windows and latitude bands (5° intervals), with KMAS areas scaled to the RGO reference frame using $b_{\mathrm{RGO-KMAS}} = 1.031$ (Mandal et al. 2020).

Figure 3 shows the comparison of recovered nesting degrees across 28 matched time-latitude windows. Despite the area calibration, RGO systematically recovers higher nesting degrees. The datasets are strongly correlated ($r = 0.777$, $p < 10^{-6}$), demonstrating that both capture the same underlying nesting patterns. The relationship follows:

$$D_{\mathrm{RGO}} = 0.66\, D_{\mathrm{KMAS}} + 0.31 \tag{6}$$

with RMS scatter of 0.158. The positive intercept implies that RGO systematically yields higher nesting degrees than KMAS over most of the observed range. The slope below unity indicates that the magnitude of this offset decreases with increasing nesting degree (~ 0.28 at $D_{\mathrm{KMAS}} = 0.1$ and ~ 0.07 at $D_{\mathrm{KMAS}} = 0.7$). This pattern is consistent for both cycles (Cycle 19: $r = 0.891$; Cycle 20: $r = 0.657$) and both hemispheres (N: mean bias +0.140; S: mean bias +0.106).

The weaker correlation in Cycle 20 (r = 0.657) compared to Cycle 19 (r = 0.891) suggests that systematic differences between the datasets might depend on the average activity level. During lower-activity periods (Cycle 20 as compared to 19), RGO's superior small-group detection may become more important for identifying weak or dispersed nests. Alternatively, this difference may partly reflect changes in KMAS observing procedures during its early operational phase following its start in 1955.

This completeness effect remains important even at high nesting degrees, where numerous small groups cluster around large active regions. Figure 4 illustrates this effect for a representative window (Cycle 19, 15° – 20° latitude): RGO identifies substantially more groups, resulting in higher recovered nesting degrees in both hemispheres ($D_{\mathrm{RGO}} = 0.768$ vs. $D_{\mathrm{KMAS}} = 0.663$ for the northern band).

The strong correlation validates that our fixed-parameter DBSCAN methodology captures reproducible nesting patterns across independent observational programmes. For the combined analysis spanning 1874 – 2025, we use RGO data for Cycles 12 – 20 (1874 – 1976)

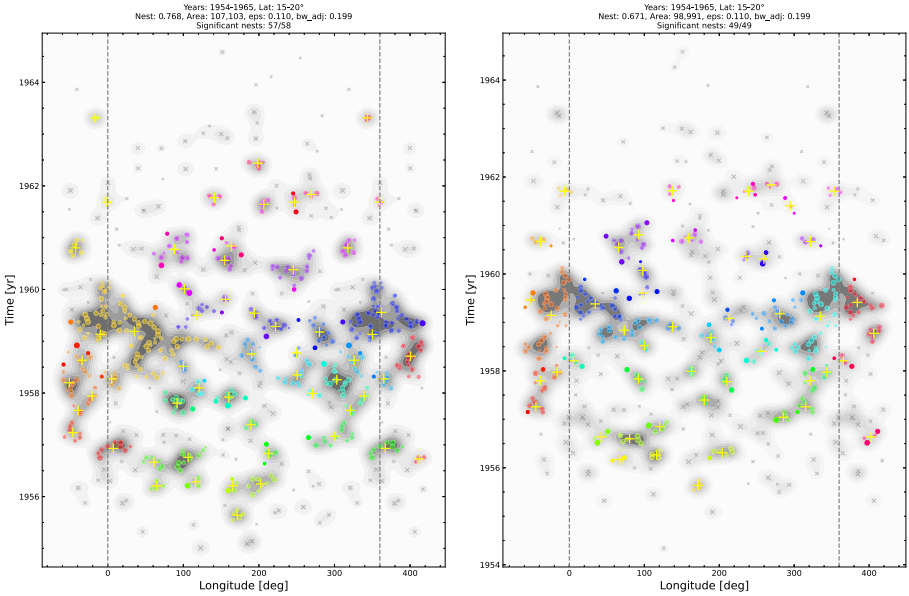


Figure 4 Example comparison for Cycle 19, latitude band $15° - 20°$. Left: RGO detects substantially more groups ($D = 0.768$). Right: KMAS shows sparser coverage ($D = 0.671$) for the identical temporal and spatial window. Both datasets identify similar large-scale nesting structures, but RGO's superior completeness captures additional small groups that increase the recovered nesting degree.

and KMAS data for Cycles 19 – 25 (1954 – 2025). When comparing results across the transition, Equation 6 provides the empirical transformation between the two scales.

3.3. Spatio-Temporal Organisation of Nesting

We have shown that nesting persists consistently across 14 solar cycles. Next, we examine how nesting patterns vary with heliographic position. The combined cycle-averaged statistics for the two datasets is presented in Table 1. Spatial analysis reveals organisation in both systematic and apparently random configurations, persisting throughout the observational record.

3.3.1. Latitude Dependence

Table 2 quantifies the strong latitude dependence of the nesting degree observed across both datasets. Peak nesting occurs in the $10° - 15°$ latitude band with $\langle D \rangle = 0.690 \pm 0.074$ in the combined RGO-KMAS dataset (after applying the calibration correction from Equation 6). The adjacent $15° - 20°$ band shows similarly high values with $\langle D \rangle = 0.688 \pm 0.060$. This mid-latitude enhancement spans the region where toroidal flux emergence rates are highest throughout the solar cycle.

Figure 5 illustrates the distribution of nesting degrees across latitude bands using the combined RGO–KMAS dataset. The box plots demonstrate remarkably consistent behaviour at mid-latitudes, with the $10° - 20°$ bands showing tight distributions ($\sigma_D \approx 0.07$) that reflect stable clustering patterns across all 14 analysed cycles. The median values (horizontal lines within boxes) closely track the mean values (red diamonds), indicating approximately symmetric distributions at these latitudes. In contrast, the $30° - 35°$ band exhibits

Table 1 Nesting statistics by solar cycle. RGO data (Cycles 12 – 19) and KMAS data (Cycles 18 – 25) with calibration correction applied (Equation 6). For overlap cycles (18 – 19), both datasets are shown. All values represent statistically significant clusters ($p < 0.05$).

Cycle	Years	Dataset	N_{win}	$\langle N_{\text{nest}} \rangle$	σ_N	$\langle D \rangle$	σ_D
12	1878 – 1890	RGO	6	42.2	24.6	0.552	0.128
13	1890 – 1902	RGO	7	54.6	39.7	0.539	0.151
14	1902 – 1913	RGO	6	52.5	34.8	0.560	0.095
15	1913 – 1923	RGO	7	64.4	45.1	0.593	0.178
16	1923 – 1933	RGO	7	62.6	43.5	0.610	0.081
17	1933 – 1944	RGO	7	68.3	44.3	0.587	0.158
18	1944 – 1954	RGO	7	68.4	46.3	0.590	0.181
19	1954 – 1964	RGO	7	79.9	37.3	0.643	0.092
		KMAS	7	64.9	37.6	0.684	0.070
20	1964 – 1976	RGO	7	70.7	33.1	0.640	0.072
		KMAS	7	60.6	37.3	0.661	0.087
21	1976 – 1986	KMAS	7	61.4	39.4	0.671	0.103
22	1986 – 1996	KMAS	7	52.7	33.5	0.642	0.072
23	1996 – 2008	KMAS	7	51.3	27.7	0.652	0.091
24	2008 – 2019	KMAS	6	49.3	30.0	0.664	0.063
25	2019 – 2025	KMAS	7	45.3	33.3	0.627	0.101

N_{win}: number of analysis windows (latitude bands); $\langle N_{\text{nest}} \rangle$: mean number of significant nests per window; $\langle D \rangle$: mean nesting degree; σ: standard deviation across windows.

Table 2 Nesting statistics by latitude band. Combined RGO and KMAS datasets with calibration applied. Values represent averages across all solar cycles covered by each dataset.

Latitude Band (°)	Dataset	N_{win}	$\langle N_{\text{nest}} \rangle$	$\langle D \rangle$	σ_D
0 – 5	RGO	9	30.7	0.575	0.064
	KMAS	7	22.7	0.609	0.066
5 – 10	RGO	9	87.9	0.649	0.067
	KMAS	7	69.0	0.709	0.033
10 – 15	RGO	9	107.4	0.671	0.084
	KMAS	7	92.4	0.728	0.033
15 – 20	RGO	9	97.7	0.677	0.070
	KMAS	7	93.7	0.717	0.031
20 – 25	RGO	9	65.9	0.628	0.053
	KMAS	7	61.0	0.682	0.029
25 – 30	RGO	9	29.8	0.518	0.103
	KMAS	7	31.1	0.628	0.039
30 – 35	RGO	7	10.9	0.374	0.172
	KMAS	6	9.8	0.505	0.050

Combined statistics from RGO (Cycles 12 – 19) and KMAS (Cycles 18 – 25).

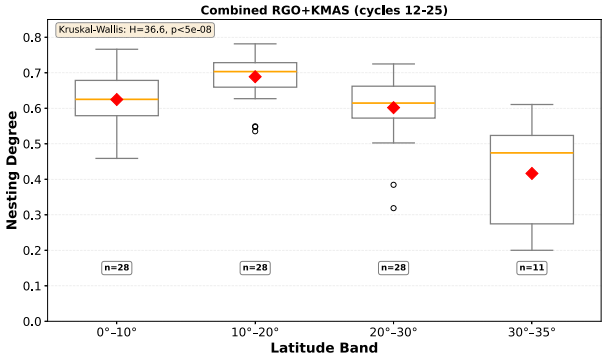


Figure 5 Distribution of nesting degrees across latitude bands from the combined RGO (Cycles 12 – 20) and KMAS (Cycles 21 – 25, calibrated) dataset. Box plots show distributions across 95 independent full-cycle analysis windows in 10°-wide latitude bands. Box boundaries mark the first and third quartiles (25th and 75th percentiles), with the median indicated by the horizontal orange line and the mean by red diamonds. Whiskers extend to 1.5 times the interquartile range, with outliers shown as circles. Sample sizes (n) indicate the number of analysis windows contributing to each band. The Kruskal–Wallis test ($H = 36.6$, $p = 5.5 \times 10^{-8}$) confirms highly significant latitude dependence.

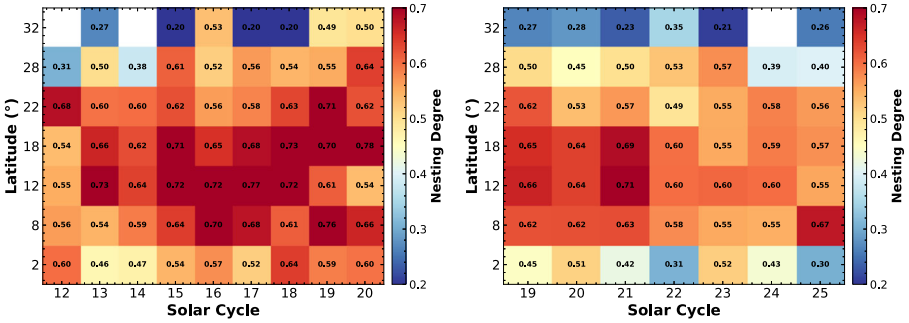


Figure 6 Estimated nesting degree for different cycles in the 5° latitude bands considered, for RGO (left panel) and calibrated KMAS (right panel) datasets. Values are annotated in each cell.

substantially larger variance, consistent with the more sporadic emergence at high latitudes, particularly near cycle minima.

Nesting degree decreases systematically toward both lower and higher latitudes. At the equatorial edge (0° – 5°), the combined dataset yields $\langle D \rangle = 0.582 \pm 0.066$, while at high latitudes (30° – 35°), nesting drops to $\langle D \rangle = 0.416 \pm 0.148$. Statistical testing confirms the high significance of this latitude dependence: a Kruskal–Wallis test on the combined dataset yields $H = 47.6$ with $p = 1.4 \times 10^{-8}$. Furthermore, mid-latitude regions (10° – 20°) show significantly higher nesting than high-latitude regions ($\geq 25°$) with a Mann–Whitney U test giving $p = 5.8 \times 10^{-8}$.

Figure 6 shows the cycle-to-cycle evolution of this latitude pattern for both datasets separately. The persistent mid-latitude enhancement appears consistently across all 14 cycles spanning 147 years. Both RGO (left panel, Cycles 12 – 20) and KMAS (right panel, cycles 19 – 25) independently reproduce the characteristic pattern, with darker colours (higher D values) concentrated in the 10° – 20° bands. Cycle 19 exhibits particularly strong nesting ($D > 0.68$) over the 10° – 25° range in both datasets, while weaker cycles such as 14 and

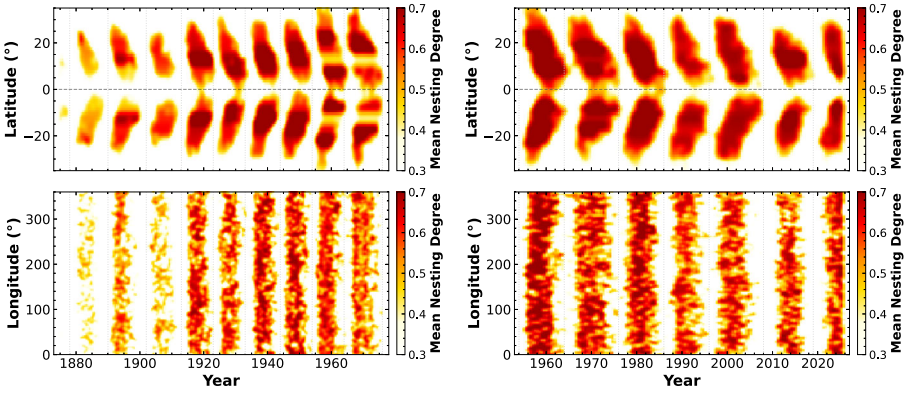


Figure 7 D-weighted KDE showing latitude-time (top) and longitude-time (bottom) projections for RGO (left) and calibrated KMAS (right).

24 maintain the mid-latitude peak despite lower overall activity levels. The agreement between the two datasets after calibration correction demonstrates that the latitude dependence represents a fundamental feature of solar magnetic emergence rather than an observational artefact.

The physical interpretation of this pattern relates directly to toroidal field structure of the solar dynamo. The $10^\circ - 20^\circ$ bands mark the latitude range where:

i) Toroidal magnetic flux reaches its maximum values during most of each cycle.
ii) Emergence rates are highest, providing numerous opportunities for repeated flux emergence in clustered patterns.
iii) The subsurface magnetic field remains coherent over extended longitudinal ranges, enabling the formation of large-scale active region complexes.

At higher latitudes ($> 25^\circ$), both emergence frequency and toroidal field coherence decrease, reducing opportunities for nest formation. Near the equator ($< 10^\circ$), while emergence rates remain substantial, the nesting degree shows greater cycle-to-cycle variation, possibly reflecting the transition between cycles.

3.3.2. Longitude Distribution

In contrast to the strong latitude dependence, sunspot group nesting shows remarkable uniformity in Carrington longitude. Figure 7 shows D-weighted kernel density estimates on both the latitude-time (top panels) and longitude-time (bottom panels) planes for RGO and KMAS datasets separately.

The latitude-time projections (top panels of Figure 7) reproduce the classic butterfly diagram pattern, with high nesting degrees (red tones, $D > 0.65$) concentrated in mid-latitude bands that migrate equatorward through each cycle. This visualisation confirms that the latitude dependence documented in Table 2 and Figure 6 persists throughout the temporal evolution of each cycle rather than reflecting a time-averaged artefact.

The longitude-time projections (bottom panels) reveal strikingly different behaviour. Unlike the strongly structured latitude-time patterns, the longitude-time plane shows diffuse, relatively homogeneous nesting patterns distributed across all Carrington longitudes. While individual nests exhibit clear longitudinal localisation (as shown in Figure 2), the aggregate distribution indicates no preferred emergence longitudes when averaged over full solar

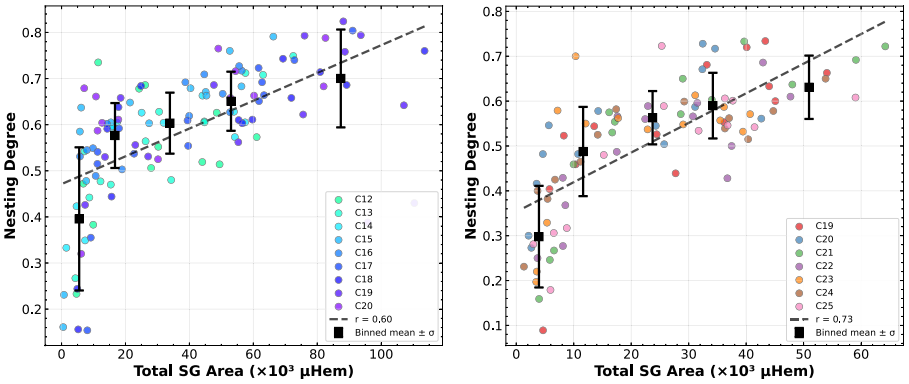


Figure 8 Nesting degree change with solar activity expressed in terms of total SG areas within analysed windows for RGO (left) and KMAS (right).

cycles. This longitudinal uniformity occurs despite substantial non-uniformity in the temporal distribution of nesting, with some periods showing concentrated emergence and others showing sparse activity.

The homogeneous longitude distribution of nests has important implications:

i) Nests forming at different latitudes experience different rotation rates in the Carrington frame. Over the 10 – 12 year duration of each analysis window, this differential rotation effectively smears any initial longitude preferences throughout the full 360° range.
ii) While individual nests can persist for multiple rotations at fixed Carrington longitudes (or exhibit systematic drift patterns), such features appear and disappear on timescales shorter than full solar cycles. The cycle-averaged view thus captures a succession of distinct nesting events rather than persistent active longitude structures.
iii) Many identified nests exhibit systematic longitude drift independent of the local differential rotation rate (Figure 2). Such drifting nests further contribute to longitudinal homogenisation when integrated over multi-year periods.

This longitude independency distinguishes our full-cycle analysis from shorter timescale studies of active longitudes. On rotational timescales ($\lesssim 1$ year), pronounced longitude preferences are commonly observed (Berdyugina and Usoskin 2003; Usoskin et al. 2007; Korsós et al. 2024). However, these structures evidently lack the multi-year persistence necessary to generate cycle-averaged longitude asymmetries in nesting patterns.

3.4. Correlation with Solar Activity

Figure 8 shows the relationship between the nesting degree and the overall solar activity level, quantified by the total SG area within each analysis window. We find a clear positive correlation in both datasets, with Pearson coefficients of $r = 0.60$ for RGO and $r = 0.73$ for KMAS. This indicates that stronger cycles, characterised by larger total emergent flux, tend to exhibit higher nesting degrees.

The trend is not strictly linear and there is considerable cycle-to-cycle scatter, especially in the KMAS dataset where small-number statistics affect low- and high-activity edges. Nevertheless, the overall behaviour suggests that more coherent toroidal flux systems during strong cycles give rise to more frequent or more pronounced longitudinal clustering. This

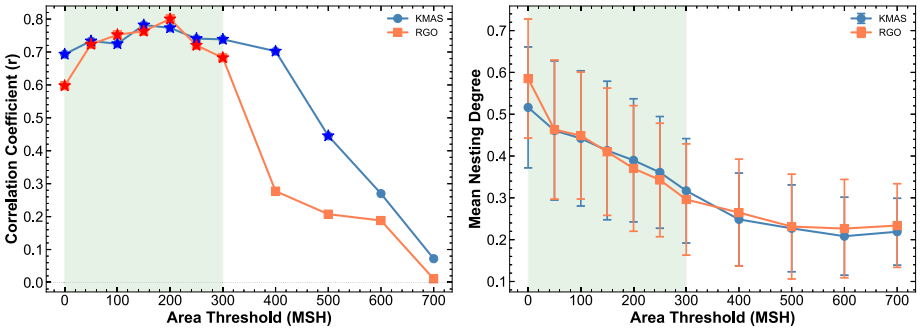


Figure 9 Pearson correlation coefficients between the nesting degree and the activity level (left) and the nesting degree (right) as a function of the minimum area threshold, for KMAS (blue) and RGO (red) datasets. Star symbols denote correlations with $p < 0.01$. The green region shows the threshold range where the nesting-activity correlation is rather stable. The error bars show 1σ intervals.

finding is consistent with the enhanced large-scale magnetic organisation expected during periods of elevated solar activity.

For each threshold, we recomputed nesting degrees for all time-latitude windows and recalculated the correlation with activity level.

3.4.1. Correlation Stability

One key question is whether the correlation between nesting degree and activity level persists when the analysis is restricted to larger groups, or whether it arises primarily because high-activity periods include more groups available to form apparent clusters. To test this, we repeated the analysis for progressively increasing minimum area thresholds.

Figure 9 (left panel) shows the correlation coefficient between nesting degree and the activity level (total area across time-latitude windows) as a function of the area threshold. For thresholds up to 200 MSH (KMAS) and 300 MSH (RGO), both datasets maintain strong, highly significant correlations (KMAS: $r = 0.70 - 0.79$; RGO: $r = 0.60 - 0.78$, all $p < 10^{-12}$). Notably, in RGO the correlation *increases* from $r = 0.60$ at zero threshold to $r = 0.78$ at about 200 MSH, indicating that the nesting–activity relationship is strongest for medium-to-large groups. This behaviour argues against a statistical counting artefact: if the correlation were driven mainly by abundant small groups during high-activity periods, excluding them would *weaken* the correlation. Instead, the correlation strengthens, suggesting that the organisation of medium-to-large active regions ($\sim 100 - 300$ MSH) plays a dominant role. At thresholds above 300 MSH, the correlation decreases as sample sizes become small.

3.4.2. Nesting Hierarchy

Beyond validating the nesting-activity correlation, the threshold analysis in Section 3.4.1 reveals hierarchical organisation in SG emergence. Figure 9 (right panel) shows mean nesting degree as a function of the area threshold for both datasets. Both KMAS (1955 – 2025, Cycles 19 – 25) and RGO (1874 – 1976, Cycles 11 – 20) exhibit consistent patterns: mean nesting degree decreases from $\sim 0.50 - 0.60$ for all groups to $\sim 0.23 - 0.27$ for groups exceeding 400 – 450 MSH. This decline indicates that small SGs are preferentially associated with clustered emergence, whereas large regions alone exhibit weaker clustering.

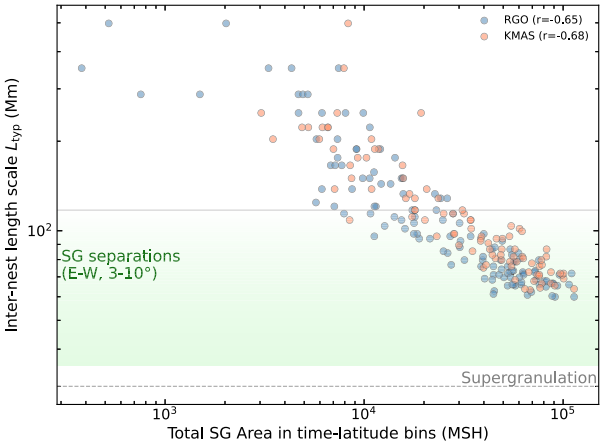


Figure 10 Inter-nest length scale L_{typ} as a function of the overall activity level in terms of the total SG area in each window considered. The range of SG longitudinal (E-W) separations is marked with a green band, where the gradient qualitatively indicates the size-dependent emergence likelihood.

The non-zero asymptotic nesting degree of ~ 0.25 at high thresholds indicates that even large active regions retain some tendency to form nests. The substantially higher nesting degree at low thresholds (~ 0.50) reflects 'parasitic' flux emergence, where small groups preferentially emerge near existing large active regions, possibly led by fragmentation of large-scale toroidal flux loops into smaller parts, clustering around the nest centre.

Overall, these results are consistent with a hierarchical emergence pattern: (1) small, parasitic groups (< 200 MSH) preferentially emerge near large active regions, producing strong clustering, and (2) large active regions themselves (> 300 MSH) emerge with reduced but persistent longitudinal organisation (see also Jiang et al. 2011, Section 3.3).

The error bars in Figure 9 represent 1σ standard deviation across time-latitude windows, reflecting variation with solar activity level and cycle phase. The large error bars ($\sim 0.10-0.20$) predominantly in the low-threshold regime, demonstrate substantial window-to-window variation, which is the source of the nesting-activity correlation presented in Section 3.4.1.

3.5. Inter-Nest Length Scale

To characterise the spatial organisation of nesting, we compute an inter-nest length scale for each analysis window (5° latitude band over a full solar cycle),

$$L_{\text{typ}} = (A/N_{\text{sig}})^{1/2}, \tag{7}$$

where $A = 360° \times 5°$ is the total surface area of the latitude band, and N_{sig} is the number of statistically significant nests identified in that window. This metric represents the mean separation obtained by uniformly distributing N_{sig} nests across the available longitude range. We convert this angular separation to physical distance (Mm) using a reference latitude of 15° ($\cos(15°) \simeq 0.966$), representative of the mid-latitude active region belt where most nesting occurs.

Figure 10 shows the inter-nest length scale as a function of total SG area (a proxy for solar activity level) for both RGO (1874–1976) and KMAS (1954–2025) latitude-time windows. Both datasets exhibit a strong negative correlation between L_{typ} and activity level, with Pearson coefficients -0.647 ($p \sim 10^{-15}$) for RGO and -0.682 ($p \sim 10^{-13}$) for KMAS. At low activity levels (sparse emergence), L_{typ} reaches 200–500 Mm, while at high activity

levels it converges to $\sim 60-100$ Mm, comparable to typical SG elongations in the E-W direction ($3-10°$, corresponding to $\sim 35-117$ Mm at 15° latitude). The mean value of $L_{\rm typ}$ is approximately 125 Mm for both datasets.

However, $L_{\rm typ}$ represents a phase–space density measure that integrates over the full solar cycle duration, including periods of minimal or no emergence. Since nests preferentially cluster in active longitude zones during emergence (e.g., Gaizauskas et al. 1983), the actual local spacing between nests within these zones is likely substantially smaller than $L_{\rm typ}$. Nevertheless, the strong anti-correlation with activity level demonstrates that nests pack more compactly in phase space during solar maximum, with the convergence to SG elongation scales suggesting dense spatial organisation during peak activity.

4. Discussion

Our analysis provides the first fully automated, cycle-by-cycle quantification of sunspot-group nesting using density-based clustering on a century-scale dataset. Several qualitative results long recognised in earlier studies naturally emerge from our DBSCAN-based approach.

First, the typical scale and temporal persistence of nests identified here align well with the classical findings of Castenmiller, Zwaan, and van der Zalm (1986) and Brouwer and Zwaan (1990). Nests commonly persist for one to several solar rotations and can extend across tens of degrees in Carrington longitude, consistent with the recurrent emergence of magnetic flux from long-lived subsurface concentrations. The observed displacement of nest centres in longitude, often independent of the local differential rotation, matches earlier descriptions of active-region complexes and reflects underlying magnetic structures that rotate with their own characteristic phase speeds.

Second, the latitude dependence of the nesting degree provides an important constraint on models of flux emergence. We find the highest nesting degree ($D \approx 0.6$) within $10° - 20°$ latitude, where the majority of active regions appear during the maximum and declining phases of cycles. Lower values toward higher latitudes primarily reflect reduced emergence rates rather than a weakening of the clustering tendency itself. This behaviour mirrors the butterfly diagram and the decreasing coherence of toroidal flux systems near cycle onset.

Recent studies using density-based clustering identified persistent active longitudes ($\sim 40°$ wide, 15 – 25 Carrington rotations) by analysing all latitudes together (Korsós et al. 2024; Császár et al. 2026). Our latitude-sliced approach quantifies nesting in both longitude and latitude over full solar cycles, revealing strong mid-latitude enhancement but no persistent longitudinal preference (Figure 7). This cycle-scale longitudinal homogeneity reflects differential rotation dispersing shorter-lived active longitude structures across all longitudes over decadal periods.

The third result comes from analysing the RGO and KMAS catalogues independently over their overlap period. RGO consistently yields slightly higher nesting degrees, which we attribute to its more complete sampling of small sunspot groups. Despite this offset, both datasets display the same latitude dependence, similar ranges of D, and comparable cycle-to-cycle variability. This agreement indicates that the clustering patterns recovered by our method represent intrinsic properties of solar emergence rather than artefacts of a particular catalogue.

Fourth, the synthetic-data validation demonstrates that the chosen clustering parameters recover the true nesting degree with $\lesssim 10\%$ scatter. Importantly, the behaviour of DBSCAN differs from classical active-longitude analyses, which typically rely on sinusoidal fits or

phase-folding of emergence patterns. DBSCAN captures clustering of arbitrary shape, enabling us to detect both compact nests and extended, drifting structures that would be difficult to characterise using traditional methods.

A further result is the moderate but clear positive correlation between nesting degree and overall cycle activity (Section 3.4). Stronger cycles, characterised by larger emergent flux, tend to show higher clustering. This behaviour is consistent with the idea that more coherent toroidal flux systems during strong cycles feed repeated localised flux emergence, whereas weaker cycles exhibit reduced magnetic organisation. It is also consistent with a similar possibility proposed for solar-type stars, where an increase in the nesting degree with the activity level could explain the observed trends in photometric variability amplitudes (Işık et al. 2020).

Finally, these results have implications for Sun-as-a-star variability studies. If more than half of all sunspot groups emerge within nests, as our analysis indicates, then the rotational modulation of solar irradiance is strongly influenced by nest persistence and longitudinal asymmetry. This connects directly to stellar observations, where active longitudes and long-lived spot complexes are commonly inferred in photometric time series (Özavcı et al. 2018; Breton, Lanza, and Messina 2024). The quantitative nesting degree ($D \approx 0.61$) provides a benchmark for comparison with stellar active-region nesting (Işık et al. 2020). If similar nesting fractions occur on other Sun-like stars, photometric variability and rotational modulation amplitudes should reflect both the instantaneous distribution of active regions and their tendency to cluster in spatial-temporal complexes.

A number of general limitations should be noted. DBSCAN relies on a single spatial–temporal scale parameter, which cannot capture variations in nest size within an analysis window. Our approach also treats longitude and time symmetrically after standardisation, whereas the physical processes governing flux emergence may introduce anisotropy. Furthermore, the longitude–time projection does not account for potential clustering in latitude. Future work could incorporate full 3D clustering or probabilistic models that allow continuous tracking of nest evolution.

The inter-nest length scale L_{typ} (Section 3.5) represents a phase-space density measure integrating over full cycle durations and uniform spatial distribution across $360^\circ \times 5^\circ$ bands. Consequently, L_{typ} does not reflect actual local spacing between neighbouring nests within active zones, where emergence concentrates preferentially (Gaizauskas et al. 1983). Nevertheless, the robust anti-correlation with activity ($r \approx -0.65$, $p < 10^{-13}$) reveals a significant pattern: during solar maximum, nests pack into denser configurations, approaching typical SG elongations. This convergence may reflect hierarchical magnetic organisation, with compact nests clustering into larger composite structures spanning up to 55° longitude (Brouwer and Zwaan 1990). Refined analysis focusing on active emergence periods rather than cycle-integrated densities would better characterise local inter-nest spacing but lies beyond the scope of the present study.

Taken together, our results show that nests constitute real, temporally coherent structures on timescales of several rotations, yet their collective contributions do not produce long-lived active longitudes when averaged over full solar cycles. Differential rotation and intrinsic longitudinal drift disperse nests on timescales shorter than a cycle, preventing persistent non-axisymmetric patterns.

5. Conclusions

We presented an automated, data-driven approach to identify sunspot-group nests in the longitude–time domain and applied it independently to the RGO (1874 – 1976) and KMAS

(1955 – 2025) catalogues. The method reliably detects coherent emergence patterns and is supported by tests on synthetic data.

Our main findings are as follows.

- Nests are a persistent feature of solar activity: across all cycles and latitude bands, the mean nesting degree is $\langle D \rangle = 0.613 \pm 0.118$, indicating that roughly 61% of all sunspot groups emerge within clusters.
- Nesting exhibits a clear latitude dependence, peaking ($D \approx 0.69$) between 10° and 15°, consistent with densest phase of the butterfly diagram and the organisation of toroidal flux in the activity belts.
- Nesting degree correlates positively with overall cycle strength, suggesting that more coherent toroidal flux bundles sustain repeated flux emergence. Importantly, this correlation strengthens when small groups are progressively excluded, indicating that nesting represents genuine clustering in flux emergence rather than statistical artefacts from abundant small groups.
- Although individual nests can remain coherent for several rotations, differential rotation and longitudinal drift disperse them on timescales short compared to a solar cycle, preventing long-lived active longitudes in cycle-averaged emergence.
- Nesting exhibits hierarchical structure: small sunspot groups cluster strongly around large active regions (nesting degree $\sim 0.5 – 0.6$), while large groups emerge with greater spatial independence (asymptotic nesting degree ~ 0.25).
- Characteristic inter-nest spacing shows strong anti-correlation with activity, contracting from $\sim 200 – 500$ Mm during low activity to $\sim 60 – 100$ Mm during solar maximum, comparable to typical sunspot group elongations (35 – 117 Mm) and indicating dense spatial packing of emergence sites.
- Despite quantitative differences, RGO and KMAS yield qualitatively consistent nesting picture, suggesting the robustness of our approach. Quantitative differences are at least partially due to non-exhaustive observational coverage.

Our findings demonstrate that unsupervised machine-learning–based clustering offers a powerful unified approach for quantifying nesting and for linking solar flux-emergence patterns to long-term magnetic activity. Extensions of this approach to full 3D clustering and to stellar photometric data may provide new constraints on the magnetic activity of Sun-like stars.

Acknowledgments We thank the anonymous reviewer for their comments and suggestions. We also thank Sudip Mandal for useful discussions about sunspot group datasets.

Author Contributions N.K. carried out data compilation, preprocessing, nesting algorithm validation, hyperparameter search, correlations, wrote the main manuscript text; E.I. developed the sunspot-group nesting identification methods and analysis tools, solar cycle generator for model validation, and interpreted results; N.A.K. supervised on data selection and analysis, contributed to interpretation of results; H.V.S. reviewed intermediate steps of the study and contributed to the final manuscript. All authors contributed to the writing of the text and reviewed the manuscript.

Funding Information Open Access funding enabled and organized by Projekt DEAL.

Data Availability The nesting statistics and longitude-time diagrams (Section 3) spanning the full time range are available in the Edmond repository of the Max Planck Society at https://doi.org/10.17617/3.MJFOKL.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Asensio Ramos, A., Cheung, M.C.M., Chifu, I., Gafeira, R.: 2023, Machine learning in solar physics. *Living Rev. Sol. Phys.* **20**, 4. DOI. ADS.

Becker, U.: 1955, Untersuchungen über die Herdbildung der Sonnenflecken. Mit 13 Textabbildungen. *Z. Astrophys.* **37**, 47. ADS.

Benevolenskaya, E.E.: 2005, The formation and evolution of complexes of activity, activity nests and the large-scale connectivity in the solar corona. In: Sankarasubramanian, K., Penn, M., Pevtsov, A. (eds.) *Large-Scale Structures and Their Role in Solar Activity*, *Astronomical Society of the Pacific Conference Series* **346**, 129. ADS.

Berdyugina, S.V., Usoskin, I.G.: 2003, Active longitudes in sunspot activity: century scale persistence. *Astron. Astrophys.* **405**, 1121. DOI. ADS.

Breton, S.N., Lanza, A.F., Messina, S.: 2024, Tracking active nests in solar-type pulsators: ensemble starspot modelling of Kepler asteroseismic targets. *Astron. Astrophys.* **682**, A67. DOI. ADS.

Brouwer, M.P., Zwaan, C.: 1990, Sunspot nests as traced by a cluster analysis. *Sol. Phys.* **129**, 221. DOI. ADS.

Castenmiller, M.J.M., Zwaan, C., van der Zalm, E.B.J.: 1986, Sunspot nests - manifestations of sequences in magnetic activity. *Sol. Phys.* **105**, 237. DOI. ADS.

Császár, K., Korsós, M.B., Soós, S., Erdélyi, R.: 2026, Exploring spatial and temporal patterns across solar cycles: focusing on active longitudes. *Astron. Astrophys.* **707**, A39. DOI.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X.: 1996, A density-based algorithm for discovering clusters in large spatial databases with noise. In: Pfitzner, D.W., Salmon, J.K. (eds.) *Second International Conference on Knowledge Discovery and Data Mining (KDD'96), Proceedings of a Conference Held August 2 – 4*, 226. ADS.

Gaizauskas, V., Harvey, K.L., Harvey, J.W., Zwaan, C.: 1983, Large-scale patterns formed by solar active regions during the ascending phase of cycle 21. *Astrophys. J.* **265**, 1056. DOI. ADS.

Işık, E., Shapiro, A.I., Solanki, S.K., Krivova, N.A.: 2020, Amplification of brightness variability by active-region nesting in solar-like stars. *Astrophys. J. Lett.* **901**, L12. DOI. ADS.

Jiang, J., Cameron, R.H., Schmitt, D., Schüssler, M.: 2011, The solar magnetic field since 1700. I. Characteristics of sunspot group emergence and reconstruction of the butterfly diagram. *Astron. Astrophys.* **528**, A82. DOI. ADS.

Korsós, M.B., Elek, A., Zuccarello, F., Erdélyi, R.: 2024, Exploring spatial and temporal patterns in the Debrecen solar faculae database: part II. *Astrophys. J.* **975**, 248. DOI. ADS.

Mandal, S., Krivova, N.A., Solanki, S.K., Sinha, N., Banerjee, D.: 2020, Sunspot area catalog revisited: daily cross-calibrated areas since 1874. *Astron. Astrophys.* **640**, A78. DOI. ADS.

Nagovitsyn, Y.A., Makarova, V.V., Nagovitsyna, E.Y.: 2007, Series of classical solar activity indices: Kislovodsk data. *Sol. Syst. Res.* **41**, 81. DOI. ADS.

Özavcı, I., Şenavcı, H.V., Işık, E., Hussain, G.A.J., O'Neal, D., Yılmaz, M., Selam, S.O.: 2018, Recurrent star-spot activity and differential rotation in KIC 11560447. *Mon. Not. R. Astron. Soc.* **474**, 5534. DOI. ADS.

Pojoga, S., Cudnik, B.: 2002, The clustering properties of active regions during the first part of solar cycle 23. *Sol. Phys.* **208**, 17. DOI. ADS.

Raphaldini, B., Dikpati, M., McIntosh, S.W.: 2023, Information-theoretic analysis of longitude distribution of photospheric magnetic fields from MDI/HMI synoptic maps: evidence for Rossby waves. *Astrophys. J.* **953**, 156. DOI. ADS.

Scott, D.W.: 1979, On optimal and data-based histograms. *Biometrika* **66**, 605. DOI.

Usoskin, I.G., Berdyugina, S.V., Poutanen, J.: 2005, Preferred sunspot longitudes: non-axisymmetry and differential rotation. *Astron. Astrophys.* **441**, 347. DOI. ADS.

Usoskin, I.G., Berdyugina, S.V., Moss, D., Sokoloff, D.D.: 2007, Long-term persistence of solar active longitudes and its implications for the solar dynamo theory. *Adv. Space Res.* **40**, 951. DOI. ADS.

van Driel-Gesztelyi, L., Green, L.M.: 2015, Evolution of active regions. *Living Rev. Sol. Phys.* **12**, 1. DOI. ADS.

Weber, M.A., Schunker, H., Jouve, L., Işık, E.: 2023, Understanding active region origins and emergence on the sun and other cool stars. *Space Sci. Rev.* **219**, 63. DOI. ADS.

Willis, D.M., Coffey, H.E., Henwood, R., Erwin, E.H., Hoyt, D.V., Wild, M.N., Denig, W.F.: 2013, The Greenwich photo-heliographic results (1874 – 1976): summary of the observations, applications, datasets, definitions and errors. *Sol. Phys.* **288**, 117. DOI. ADS.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.