AutoTAB: Automatic Tracking Algorithm for Bipolar Magnetic Regions

Anu Sreedevi[1] , Bibhuti Kumar Jha[2,3] , Bidya Binay Karak[1] , and Dipankar Banerjee[3,4,5]
[1] Department of Physics, Indian Institute of Technology (Banaras Hindu University), Varanasi 221005, India; anubsreedevi.rs.phy20@itbhu.ac.in,
karak.phy@iitbhu.ac.in
[2] Southwest Research Institute, Boulder, CO 80302, USA
[3] Aryabhatta Research Institute of Observational Sciences, Nainital 263002, Uttarakhand, India
[4] Indian Institute of Astrophysics, Koramangala, Bangalore 560034, India
[5] Center of Excellence in Space Sciences India, IISER Kolkata, Mohanpur 741246, West Bengal, India

Abstract

Bipolar magnetic regions (BMRs) provide crucial information about solar magnetism. They exhibit varying morphology and magnetic properties throughout their lifetime, and studying these properties can provide valuable insights into the workings of the solar dynamo. The majority of previous studies have counted every detected BMR as a new one and have not been able to study the full life history of each BMR. To address this issue, we have developed Automatic Tracking Algorithm for BMRs (AutoTAB) that tracks the BMRs for their entire lifetime or throughout their disk passage. AutoTAB uses the binary maps of detected BMRs and their overlapping criterion to automatically track the regions. In this first article of this project, we provide a detailed description of the working of the algorithm and evaluate its strengths and weaknesses by comparing it with existing algorithms. AutoTAB excels in tracking even for the small BMRs (with a flux of $\sim 10^{20}$ Mx), and it has successfully tracked 9152 BMRs over the last two solar cycles (1996–2020), providing a comprehensive data set that depicts the evolution of various properties for each BMR. The tracked BMRs exhibit the well-known butterfly diagram and 11 yr solar cycle variation, except for small BMRs, which appear at all phases of the solar cycle and show a weak latitudinal dependence. Finally, we discuss the possibility of adapting our algorithm to other data sets and expanding the technique to track other solar features in the future.

Unified Astronomy Thesaurus concepts: Solar magnetic fields (1503); Solar active regions (1974); Sunspots (1653); Bipolar sunspot groups (156); Solar cycle (1487); Solar magnetic flux emergence (2000); Sunspot groups (1651); Sunspot cycle (1650)

1. Introduction

Bipolar magnetic regions (BMRs) remain one of the most predominant signatures of solar magnetism as observed on the surface of the Sun. The number of such regions at a given time represents the solar activity that waxes and wanes cyclically over a period of 11 yr (Schwabe 1844; Hathaway 2015). These BMRs are generally the source regions of solar eruptive events (Schrijver 2009), and hence they are crucial for the understanding of space-weather conditions. BMRs are observed to be tilted by an angle with respect to the equator, which is found to increase statistically with latitude; this is known as Joy's law (Hale et al. 1919). It is observed that the decay and dispersal of tilted BMRs produce the poloidal field in the Sun through the Babcock–Leighton mechanism (Babcock 1961; Leighton 1964; Mordvinov et al. 2022).

Sunspot groups in white-light observations mimic the locations of large BMRs and have traditionally been used for the study of BMR properties (e.g., Howard 1991, 1996; Sivaraman et al. 2007; Dasi-Espuig et al. 2010) because of the unavailability of magnetogram data. Since the early 1970s, regular full-disk measurements of the Sun's magnetic field have been conducted, and now there exist vast archived data of solar magnetograms from the different ground-based (Synoptic Optical Long-term Investigations of the Sun, SOLIS; Keller et al. 2003; Global

Oscillation Network Group, GONG; Harvey et al. 1996) and space-based (Michelson Doppler Imager, MDI; Scherrer et al. 1995; Helioseismic and Magnetic Imager, HMI; Scherrer et al. 2012) observatories. To exploit such a large volume of data for the understanding of BMRs and solar magnetism, various automatic methods have been developed (Stenflo & Kosovichev 2012; Tlatov & Pevtsov 2014; Jha et al. 2020). In all these studies, each detection of BMRs has been treated as a new one (Stenflo & Kosovichev 2012; Jha et al. 2020). However, this approach may influence the analysis as bigger BMRs have a longer lifetime and thereby will have higher weightage in the analysis (van Driel-Gesztelyi & Green 2015). Furthermore, BMR properties, including morphology, magnetic field strength, and tilt angle, evolve significantly over its lifetime (Ugarte-Urra et al. 2015; McClintock & Norton 2016; Getling & Buchnev 2019; Schunker et al. 2019, 2020). Hence, to overcome these limitations, it is essential to track the BMR to get insight into the physics of the formation and evolution of a BMR. Moreover, the automatic detection and tracking of magnetic regions also become essential for monitoring solar activity and events (McAteer et al. 2005; LaBonte et al. 2007). Therefore, an enormous effort has been made to develop automatic algorithms to track the BMRs (Higgins et al. 2011; Muñoz-Jaramillo et al. 2016).

Solar monitor active region tracking (SMART; Higgins et al. 2011) was developed for the automatic detection and tracking of magnetic active regions in real time for solar eruptive event prediction. Although SMART does a decent job in identifying and extracting various magnetic features, it tends to miss some quiet-Sun magnetic regions. The bipolar active region detection

(BARD; Muñoz-Jaramillo et al. 2016) is another algorithm that was developed for the detection and tracking of BMRs. It uses similar techniques as SMART to detect BMRs, but uses the dual-maximum flux-weighted overlap method for feature association. The BARD uses human supervision to correct any errors in detection and tracking. Furthermore, Space-Weather MDI Active Region Patches (SMARPs; Bobra et al. 2021) and Space-Weather HMI Active Region Patches (SHARPs; Bobra et al. 2014) are the data sets derived from MDI and HMI magnetograms, respectively, which provide the tracked maps of active regions identified in the magnetic image of the Sun. We note that in SMARPs and SHARPs, active regions are not necessarily the BMRs, where a decent flux balance condition holds.

To the best of our knowledge, there is no existing catalog that can provide the properties of the individual BMRs throughout their lifetimes, or at least the course of their lifetimes. Therefore, we overcome the limitation of existing BMR tracking algorithms and develop a completely automatic method to track the BMR, which can be implemented in all sets of magnetogram data. In Section 2, we present a detailed description of our tracking algorithm and its comparison with the existing algorithm. In Section 3, we present some representative results based on our tracking algorithm, and finally, in Section 4, we conclude with our insight on this new algorithm.

2. Data and Method

In our study, we use line-of-sight (LOS) magnetic field observations of the Sun, usually referred to as magnetograms, for the period of 1996 September–2019 December (Cycles 23 and 24) from MDI (1996 to 2011; Scherrer et al. 1995) and HMI (2010–present; Scherrer et al. 2012) on board the Solar and Heliospheric Observatory (SOHO) and Solar Dynamics Observatory (SDO), respectively. Here, we utilize all the magnetogram data from MDI, which have a cadence of 96 minutes in 1024×1024 pixels with a spatial resolution of $4'' \times 4''$. Although HMI provides magnetograms with a cadence of 45 s in 4096×4096 resolution with a spatial resolution of $1'' \times 1''$, for the ease of computation, we only use one image every 96 minutes as available for MDI from the HMI series "hmi.M_720s."

The data sets used for the mentioned period contain 126,381 fits images (MDI: 56,384; HMI: 69,997). In addition to the quality keyword, the number of faulty pixels (identified as "Not a Number") on the solar disk region is also checked for every magnetogram. If the number of faulty pixels is found to be greater than 100, the magnetogram is considered defective and was not included in the analysis.

The data in the magnetograms correspond to the LOS component of the surface magnetic flux density, calculated across a spatial resolution window. These LOS magnetograms suffer from the projection effects, which are inversely proportional to the cosine of the heliocentric angular distance (μ).[6] We correct for the LOS component of the magnetic field by assuming that the magnetic field is normal to the solar surface. Therefore, we divide the surface magnetic flux density at each pixel by the cosine of μ. Furthermore, as we approach to limb, the projection effects become severe, and hence, initially, we restrict ourselves to less than $0.9\,R_\odot$.

2.1. BMR Detection Algorithm

The algorithm developed to detect BMRs from the LOS magnetograms follows the prescription given in Stenflo & Kosovichev (2012), which was also used in Jha et al. (2020) with slight modifications. A brief description of the detection algorithm is as follows. In the first step, an adaptive threshold, based on the average of the absolute magnetic field inside the $0.9\,R_\odot$, is applied to the projection-corrected magnetogram to identify the regions of strong magnetic fields. This step is followed by applying a threshold of 200 G to the identified regions and imposing a flux balance condition defined by $(F_+ + F_-)/(F_+ - F_-) < 0.4$ on them to isolate the BMRs. Here, F_+ (F_-) represents the sum of all the positive (negative) radial magnetic field values that exceed absolute 100 G for the identified regions (see the Appendix of Stenflo & Kosovichev 2012, for a detailed description of the detection method). Stenflo & Kosovichev (2012) only used this algorithm for the MDI data set, and it was extended for HMI in Jha et al. (2020). Jha et al. (2020) noted that the same method could be used for HMI data by multiplying the HMI magnetogram by a factor of 1.4, which comes from the comparison of MDI and HMI magnetograms in the overlapping period (Liu et al. 2012). To keep the consistency in the data set, for HMI magnetograms, we rebinned the magnetogram from 4096×4096 pixels to 1024×1024 pixels to further reduce the computational time. By comparing the detected BMRs from the original and the binned data, we find that binning tends to miss the detection of small BMRs with a flux of $\sim 10^{20}$ Mx, but the detection of other BMRs remains unaffected. One such example of identified regions from MDI is shown in Figure 1(a). The BMR from both data sets is stored in the form of a binary mask, which acts as the starting point of our tracking algorithm.

2.2. Preprocessing of BMR Masks

In Jha et al. (2020), the identified BMRs are stored as rectangular regions by binary masks (see Figure 1(a)), which include additional pixels that are not part of BMRs and may affect the tracking, particularly in the case of closely spaced BMRs during a high-activity period. Hence, to get rid of those extra pixels before we start the tracking, we go through the following preprocessing steps.

To get the exact morphology of the BMR in the rectangular regions of interest (ROI), we picked all the pixels with absolute magnetic field values greater than 100 G. This threshold not only leads to a few fragmented pixels in the ROI but also separates the poles of the BMR, as they do not always touch each other's boundaries (see Figure 1(b)). To eliminate the fragmented pixels and connect isolated parts, we use a technique called morphological closing. It works by filling in the gaps or holes in the image using a kernel with a certain shape and size.[7] Here, we use a kernel of a square matrix with values inside a set radius R equal to 1 and others 0. In our case, we start with a kernel that has an initial radius of 6 pixels, followed by an area threshold of 50 pixels for filled regions. The radius of the kernel is systematically increased up to 9 pixels at the step of one pixel until the number of connected regions matches the number of identified BMRs (i.e., the number of rectangular regions; see Figure 1(a)) in that magnetogram. Finally, the flux balance condition, as described

---

[6] Angular distance of the pixel in consideration from the disk center.

[7] We use morph_close.pro function available in IDL for this.
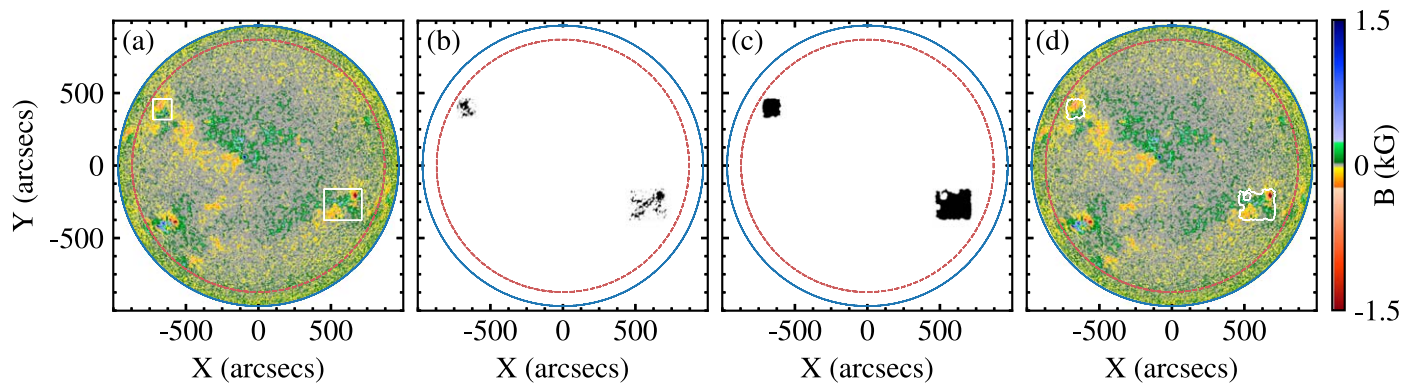
Figure 1. (a) Magnetogram from MDI for 2000 November 15 20:48 with detected BMRs represented by white rectangular boxes. (b) Binary mask of detected BMRs from panel (a), but after applying the threshold of 100 G to the detected regions. (c) Same as panel (b) after the preprocessing (described in Section 2.2). (d) Identified regions after the preprocessing depicted on the magnetogram. The solid blue and dashed red circles represent $1.0\,R_\odot$ and $0.9\,R_\odot$, respectively.
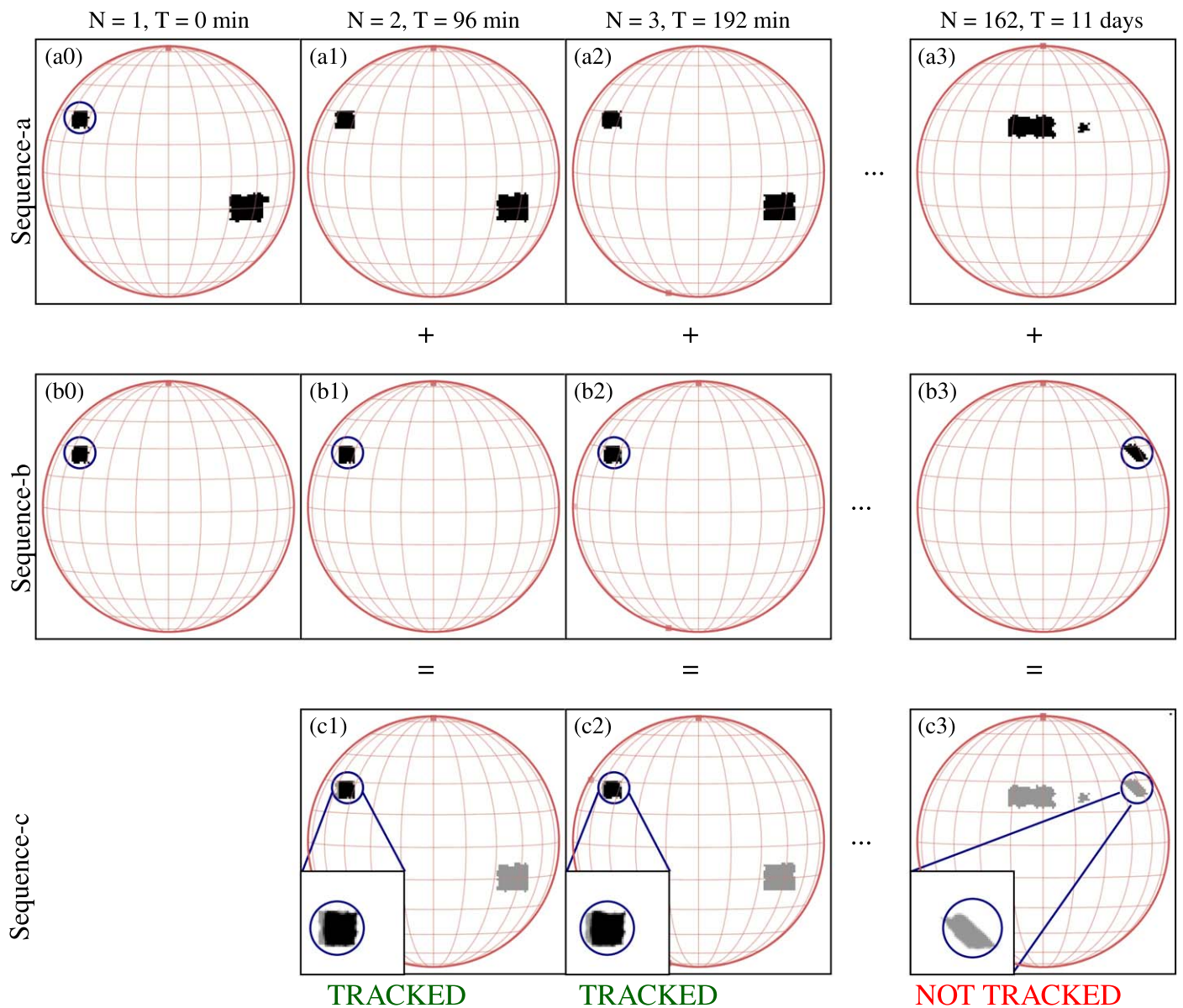
Figure 2. Representative example of the BMR tracking algorithm. The BMR intended to be tracked is marked by a circle. (a0), (a1), (a2), and (a3) represent the selected binary maps in Sequence-a. (b0), (b1), (b2), and (b3) represent Sequence-b, the isolated and differentially rotated binary masks of the BMR, corresponding to the time of observation in Sequence-a. (c1), (c2), and (c3) represent Sequence-c, obtained by adding Sequence-a and Sequence-b. The zoomed-in view of the overlapping region is shown in the inset.

previously, (Stenflo & Kosovichev 2012; Jha et al. 2020) is verified for each region to ensure that they are bipolar. In Figure 1(c), we show the BMR regions after the morphological closing operation. In Figure 1(d), white contours represent the BMRs after the preprocessing steps. The information of detected BMRs after preprocessing is also stored as binary masks so that it can be used for tracking.

2.3. BMR Tracking Algorithm

Now we come to the core of the work, which is to develop an automatic algorithm to track all of the identified BMRs. Our tracking algorithm employs binary masks obtained from the last step of preprocessing to follow the interested regions over their lifetimes/disk passage. The concept of the BMR tracking algorithm comes from the sunspot tracking algorithm developed in Jha et al. (2021), which is modified considerably to track the BMRs in the magnetograms. Now, we discuss the steps of our tracking algorithm as follows.

1. In the first step, a binary mask is selected from the archive (Figure 2(a1)). A BMR is selected from this mask and is checked for tracking history stored in Step 7. If the selected BMR is not tracked already (see the marked circle in Figure 2(a1)), it is isolated in a separate binary mask (Figure 2(b1)), and a unique BMR-ID (e.g., 10001) is assigned to it.

2. Now, we calculate the maximum tracking period (T_{max}), i.e., the time it takes to reach the west limb, based on the heliographic latitude θ_{BMR} and longitude ϕ_{BMR} of the BMR, as

$$T_{max} = \frac{90° - \phi_{BMR}}{\Omega(\theta_{BMR})}. \quad (1)$$

Here, $\Omega(\theta_{BMR})$ is the photospheric differential rotation rate (Howard et al. 1990) for a given θ_{BMR}.

3. In the next step, we take a sequence of binary masks (e.g., Sequence-a, as shown in Figures 2(a0)–(a3)) falling in the range of T_{max}.

4. The binary mask of an isolated BMR (Figure 2(b1)) is differentially rotated using the drot_map.pro routine in IDL[8] to the time of observation of Sequence-a. For example, b0 is differentially rotated to the time of observation of a1 to obtain b1; similarly, b0 is differentially rotated to the time of a2 to obtain b2, and so on. This is represented by Sequence-b and is shown in Figures 2(b0)–(b3).

5. In the following step, we add Sequence-a (Figures 2(a0)–(a3)) with the corresponding differentially rotated binary masks, Sequence-b (Figure 2(b0)–(b3)), e.g., c1 = a1 +b1, c2 = a2+b2 and so on. This is represented by Sequence-c and is shown in Figures 2(c1)–(c2).

6. Now, we go through all the images in Sequence-c (Figures 2(c1)–(c3)) and check for the overlapping pixels (pixels with value 2). If the overlap is more than 150 pixels (600 arcsec²), the BMR is marked as tracked (Figures 2(c1) and (c2)); otherwise, it is marked as not tracked (Figure 2(c3)). Furthermore, if the overlapping criteria are not met in the consecutive 30 binary masks (i.e., not enough overlap was detected for the next 48 hr), we increase the overlapping criteria to 250 pixels

(1000 arcsec²) for the next observation to ensure that the BMR tracked is the intended one and not a new one. Eventually, we only track BMRs with an area greater than 1000 arcsec².

7. The same BMR-ID, i.e., 10001, is assigned to all the tracked BMRs in Sequence-c. At this step, along with BMR-ID, we also store a few other parameters of the tracked BMR, such as B_{max}, B_{mean}, total unsigned flux, area, and heliographic coordinates. The time and unique region index, which is assigned to each connected region (pixels with value 2) in Sequence-c along with the BMR-ID, are appended in the tracking history file, which will be used in Step 1 to check for tracking history.

8. We repeat all the above steps for all the other detected BMRs within the chosen binary map from Step 1 (Figure 2(a0)).

9. Once all the BMRs in the first selected observation are tracked, we select the next observation and track only the new BMRs that are not tracked already. We do this by comparing the BMR-ID with the existing information in the tracked history file that was updated in the previous step.

3. Results

#### 3.1. Representative Tracking Results

To demonstrate the result of our tracking algorithm, in Figure 3, we show the evolution of NOAA region AR9232 (AutoTAB-ID 11436) and AR11390 (AutoTAB-ID 11201), which have been observed in MDI and HMI, respectively. AR9232 was first identified near the east limb (25.2° N, 49.3° E) and has been tracked throughout its disk passage from 2000 June 18 03:11 to 2000 June 24 22:23.[9] Although our algorithm tracks the AR9232 in the next 113 time steps with 96 minute cadence, in Figure 3(a), we only show one snapshot every day for 7 days. In Figure 3(c), we also show the evolution of B_{max} and the absolute total flux for AR9232, and we note a systematic and continuous decrease (except small fluctuation due to fragmentation of the negative polarity) in both these quantities. This continuous decrease suggests that Automatic Tracking Algorithm for BMRs (AutoTAB) has been able to track the BMR properly, and it has been picked during its decaying phase, which is also inferred from Figure 3. In contrast with our algorithm, the SMART algorithm (Higgins et al. 2011) tracks the same region from 2000 November 15 03:15 to 2000 November 23 20:48.

In Figure 3(b), we show the evolution of AR11390 observed in HMI in a similar way as for AR9232. This BMR was first identified near the east limb (8.3° N, 58.2° E), and AutoTAB tracked it for 8 days (in 122 observations) until it reached close to the west limb. The variation of total flux and B_{max} (Figure 3(d)) once again shows the continuous and systematic variation over the tracking period. Furthermore, in Figure 3(d), we see that B_{max} as well as the absolute total flux initially show a small variation and then they start to increase rapidly on day 5. Hence, we infer that this BMR has been tracked from its emergence to the time it crosses the west limb, when AutoTAB loses track of it.

Apart from these two BMRs, in the Appendix, we also show the evolution of another tracked BMR that was observed in both MDI and HMI (also see Figure 9). The comparison of

---

[8] It uses a differential rotation profile from Howard et al. (1990).

[9] All the times mentioned in this article are in UTC.
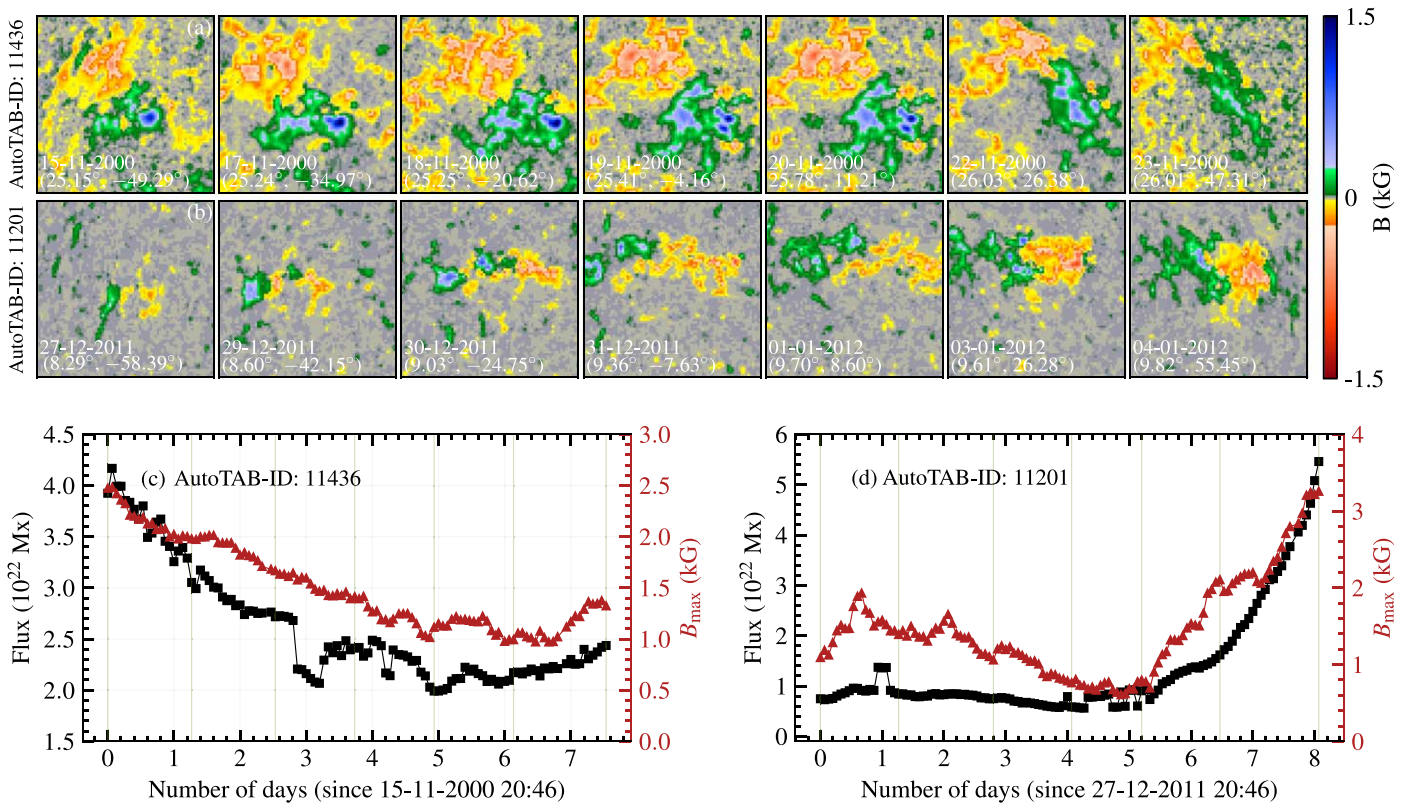
Figure 3. Panels (a) and (b) show the snapshots of the tracked BMRs with AR9232 (AutoTAB-ID 11436) and AR11390 (AutoTAB-ID 11201), respectively, corresponding to each day during the tracking. Panels (c) and (d) show the evolution of the absolute total flux and B_{max} for these BMRs. The vertical lines in panels (c) and (d) represent the times corresponding to the time of the snapshots shown in panels (a) and (b), respectively.
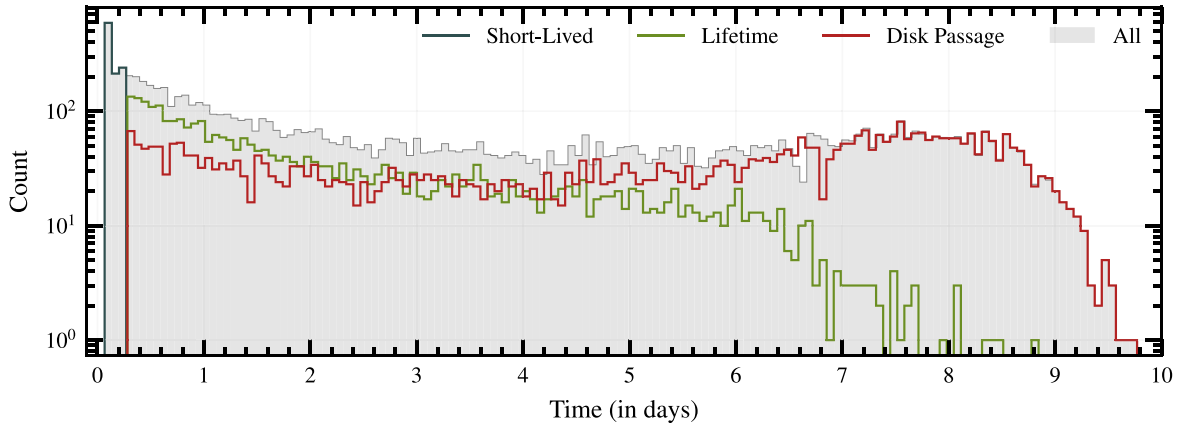
Figure 4. Distribution of the lifetimes of the tracked BMRs of different classes shown by different colors.

B_{max} and the total flux of this BMR obtained from MDI and HMI confirms that the AutoTAB is working in the same way for both data sets. Note that we have scaled up B_{max} measured in HMI by 1.4 to bring both MDI and HMI to the same level (Liu et al. 2012).

By selecting the BMRs randomly from the tracked data, we observe that similar types of variations are consistently present in all the tracked BMRs, with a few exceptions. Therefore, based on these findings, we say that AutoTAB is very efficient in tracking the BMRs during their appearance on the nearside of the Sun. Using our state-of-the-art algorithm, we have tracked 9232 BMRs in Cycles 23 and 24 from 1996 to 2022. In the following section, we discuss some of the statistical properties of these tracked BMRs.

3.2. Statistical Properties of BMRs

In Figure 4, we show the distribution of the lifetimes/disk-passage times for all the tracked BMRs. From Figure 4, we note that BMRs exhibit a broad range of lifetimes, varying from several hours to over a week. Although AutoTAB is not able to track a BMR that goes to the farside, it has been tracked only in its evolutionary stage. Other than that, we also encountered a few cases in which the BMR lived for a sufficiently long time, but they were detected for less than 5% of all times, in the number of expected frames[10] due to data gaps, corrupted data, or they were missed during detection.

---

[10] Number of expected frames = (time of last detection − time of first detection)/cadence of data.
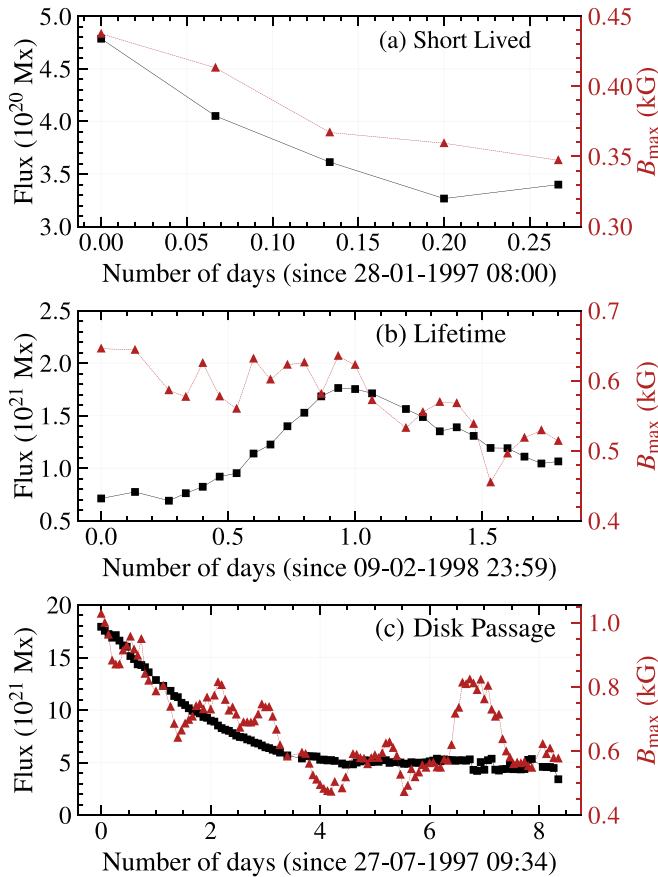
Figure 5. Evolution of the absolute total flux and B_{max} for a typical BMR from (a) SL, (b) LT, and (c) DP classes. AutoTAB-IDs of the BMRs shown in panels (a), (b), and (c) are 10076, 10399, and 10213, respectively. The NOAA number assigned only to the representative BMR of the DP class is AR08066.

Hence, after excluding such cases, based on how they were tracked, i.e., either for the whole life or only in the evolutionary stage, we classified the BMRs into three classes, which are as follows.

1. *Short-lived (SL).* The BMRs that emerge and decay on the nearside of the Sun and have a lifetime of 8 hr or shorter are classified as short-lived (SL) BMRs. We segregated this class as they mainly comprise small BMRs, which may include a few large ephemeral regions with a typical magnetic flux of $\approx 10^{20}$ Mx. The distribution of the SL class appears in the leftmost panel of Figure 4. In Figure 5(a), we show the evolution of the total flux and B_{max} of a typical BMR (AutoTAB-ID 10076) from this class. It is noted from this figure that the BMR shown here emerges with relatively low flux and B_{max} and decays within a few hours.

2. *Lifetime (LT).* The rest of the tracked BMRs that have a lifetime of more than 8 hr and have emerged and decayed in the nearside of the Sun, i.e., they were tracked throughout their lifetime, are classified as lifetime (LT) BMRs. The distribution of the LT class spans around a week, as shown in Figure 4. An example of the evolution of such a BMR (AutoTAB-ID 10213) is shown in Figure 5(b).

3. *Disk passage (DP).* This class of BMRs has not been tracked for their lifetime. Instead, they have only been tracked for a part of their life span on the visible solar

disk, and we classified them as disk passage (DP). This class includes (i) BMRs that appear near the east limb ($\leqslant 45°$ E) and disappear on the nearside; (ii) BMRs that appear on the nearside of the Sun, but cross the west limb (longitude $\geqslant 45°$ W); and (iii) BMRs that appear on the east limb ($\leqslant 45°$ E) and cross the west limb ($\geqslant 45°$ W). Here, we restrict ourselves to longitudes of $\pm 45°$, as the uncertainty in the magnetic field measurement increases toward the limb. The lifetime for this class is distributed all the way from a few hours to more than a week (Figure 4). The evolution of flux and B_{max} for one such BMR (AutoTAB-ID 10399; NOAA AR08066) is shown in Figure 5(c).

The snapshots of the evolution of all the three BMRs in Figure 5 can be found in the data repository[11] for AutoTAB.

The number of BMRs identified in each class is represented in Table 1 for the period of 1996–2020. Along with their numbers, we also calculated the mean of area, total flux, B_{max} and B_{mean} for all these classes, which are listed in Table 1. From this table, we note that all these quantities increase from the SL class to the LT and DP classes. Hence, we find that the larger BMRs have longer lifetimes, which is in agreement with the earlier findings from the sunspot (van Driel-Gesztelyi & Green 2015).

After discussing the various classes of BMRs and their time evolutions, we now discuss the collective behavior of the BMRs in the aforementioned three classes.

3.2.1. Solar Cycle Variation

The first property that we looked for is whether the number of newly emerging BMRs obeys the well-known solar cycle behavior. In contrast to previous studies (e.g., Stenflo & Kosovichev 2012; Jha et al. 2020), where the same BMR is counted multiple times during their appearance on the nearside, here we only count each tracked BMR once. Therefore, in Figure 6, we show the monthly number of newly emerging BMRs with the time for all three classes, SL (Figure 6(a)), LT (Figure 6(b)), and DP (Figure 6(c)). The number of BMRs in the LT and DP classes obediently follows the known solar cycle behavior based on the conventional sunspot number. Such behavior is not evident in the case of the SL class. A significant number of SL BMRs are also observed during the solar minima when the toroidal field of the Sun is weak and only a few large BMRs are produced.

3.2.2. Latitude–Time Distribution

Another crucial property of sunspots is their latitudinal distribution (Hathaway 2015; Jha et al. 2022), the so-called "butterfly diagram." In Figure 7 we show the latitude–time distribution of the identified BMRs for each class. Here, each point represents a unique BMR. The latitude and time of all the BMRs are chosen at a time when they attain their maximum total flux (represented by the colors of the points) during the tracking period. We note that for the DP class, the maximum flux may not be the actual maximum flux of the BMR, as AutoTAB only tracked them during a fraction of their life span. The interesting thing to note here is that although the BMR classes LT (Figure 7(b)) and DP (Figure 7(c)) follow the well-known butterfly diagram, the SL (Figure 7(a)) class is more

---

[11] https://github.com/sreedevi-anu/AutoTAB

Table 1
Some Key Parameters of Different Classes of Tracked BMRs

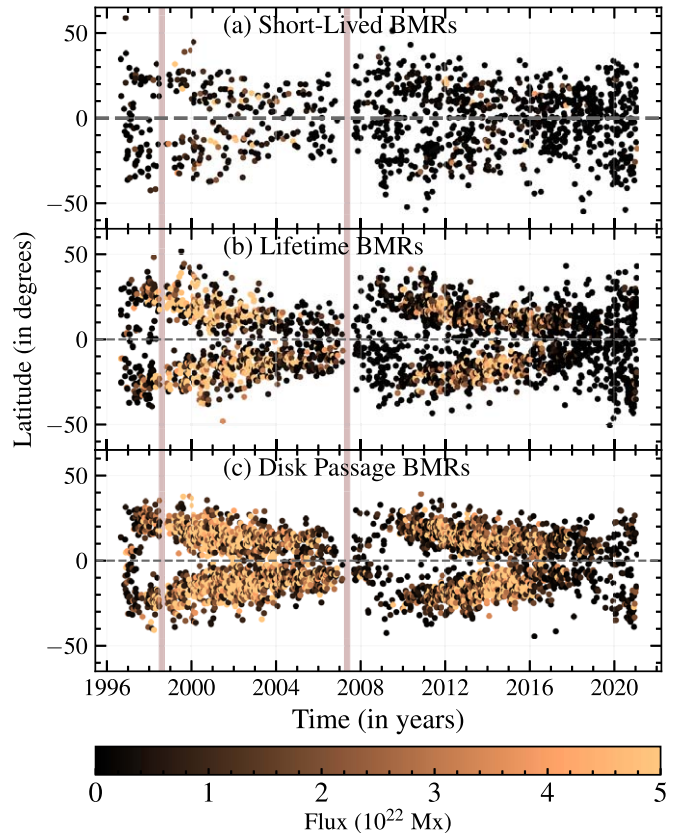| Classification	Number of BMRs	Area ± ΔArea (μHem)	Flux ± ΔFlux (10^{22} Mx)	$B_{max} \pm \Delta B_{max}$ (G)	$B_{mean} \pm \Delta B_{mean}$ (G)
Short-lived (SL)	1251	20.17 ± 0.71	0.26 ± 0.01	541.32 ± 5.38	197.46 ± 0.66
Lifetime (LT)	3191	88.65 ± 1.05	1.50 ± 0.02	949.20 ± 6.35	224.46 ± 0.81
Disk Passage (DP)	4710	116.87 ± 0.17	2.05 ± 0.01	1436.83 ± 7.02	281.07 ± 0.71
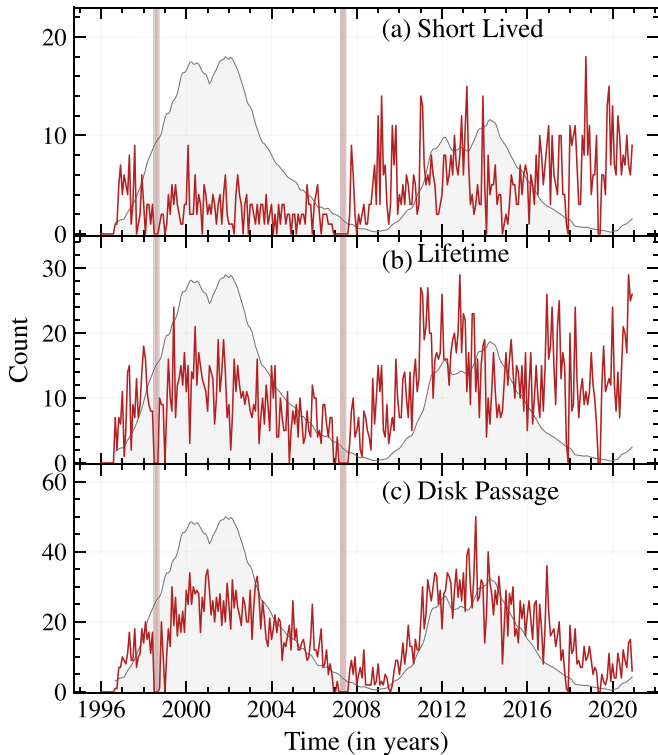
Figure 6. Time series of the monthly number of unique BMRs for (a) SL, (b) LT, and (c) DP classes. The gray shaded region represents the international smoothed sunspot number (scaled to BMR counts for better representation) from WDC-SILSO, Royal Observatory of Belgium, Brussels. Here, vertical strips represent the data gaps in the MDI magnetograms.

Figure 7. Latitude–time distribution (butterfly diagram) of the tracked BMRs of (a) SL, (b) LT, and (c) DP classes. The color of the points represents the total flux of the BMR. The vertical strips represent the data gaps in MDI magnetograms.

scattered over the latitude (a weak latitudinal dependence can still be seen) independent of the phase of the solar cycle. It could be that SL class BMRs that appear from the high latitude at the beginning of the cycle and continue until the end of the cycle at the solar minimum are part of the extended solar cycle (McIntosh et al. 2015), and they are formed through shredding and tangling of the large-scale magnetic field itself (Karak & Brandenburg 2016). However, Jha et al. (2020) have shown that the two classes of BMRs, namely the BMRs with a spot and without a spot in white-light images, follow the same latitude–time distribution based on their classification. Further study is required to dig deep into the detail, which will be done in the follow-up paper. Another point to note here is that although the total number of BMRs in Cycle 23 is less than in Cycle 24, the fraction of bigger BMRs that falls in the LT and DP classes is greater in Cycle 23.

3.3. Comparison with BARD

In this section, we evaluate the performance of AutoTAB with another existing BMR tracking algorithm, BARD (Muñoz-Jaramillo et al. 2016, 2021a, 2021b). To achieve this,

we compute the count of newly emergent BMRs detected by both algorithms for the overlapping period, which is represented in Figure 8. The number of newly emergent BMRs in BARD is obtained from the BARD tilt catalog (Muñoz-Jaramillo et al. 2016, 2021a, 2021b).

It can be seen from Figure 8(a) that AutoTAB detects more BMRs throughout the intersecting period, especially during Cycle 24. As BARD has limited the cadence of the observation to one magnetogram per day, any BMR detected and tracked by BARD cannot have a lifetime of less than a single day. Moreover, BARD employs different magnetic threshold values for MDI and HMI, which can bring about the disparity observed in the detected number of BMRs. Taking these points into consideration and in order to maintain consistency between the data sets, initially, we exclude all BMRs from the SL class and any BMRs that live for less than a day from the AutoTAB data set. Nevertheless, AutoTAB still detects 28% more new BMRs in Cycle 23 and 85% in Cycle 24. Additionally, removing the LT class BMRs from the AutoTAB data set and applying a 500 G magnetic threshold to the BMRs detected
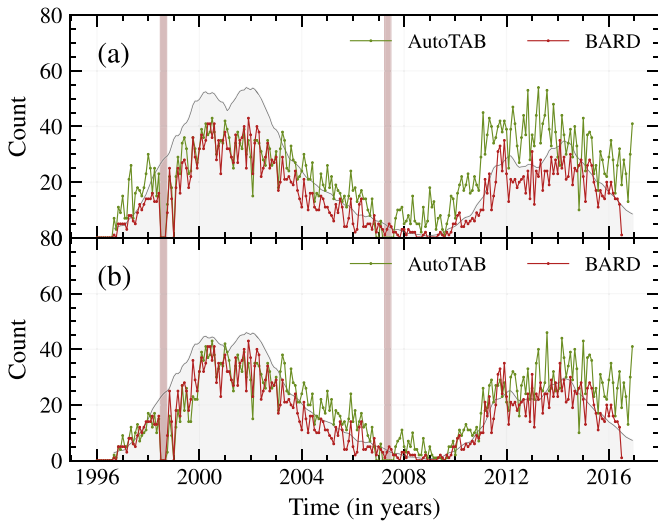
Figure 8. (a) Comparison of the monthly number of newly emergent BMRs obtained from AutoTAB (our algorithm) and BARD, along with the traditional monthly sunspot number (gray shaded area enclosed by line). The sunspot number has been multiplied by a factor of 0.6 to bring it to the scale of the BMR count. The vertical strips in the figure represent the data gaps in MDI magnetograms. (b) Same as panel (a), but only including the BMRs from DP class, which has a lifetime of greater than 1 day and the 500 G threshold applied to BMRs from MDI in the AutoTAB data set.

Table 2
Comparison of the Statistics of New Emergent BMRs between AutoTAB and BARD for Different Cases

	Cycle 23		Cycle 24	
	AutoTAB	BARD	AutoTAB	BARD
All BMRs	3861	2328	3900	1422
BMRs with lifetime $\geqslant$1 day	2989	2328	2742	1422
BMRs with lifetime $\geqslant$1 day and LT class removed (i.e., we are left with DP class)	2722	2328	2114	1422
All DP class BMRs plus 500 G threshold to Cycle 23	2649	2328	2114	1422

from MDI (Cycle 23) results in a close match of 13% of more new BMRs detected by AutoTAB in Cycle 23. However, we still observe that in Cycle 24, AutoTAB detects 42% of the newer BMRs; see Figure 8(b). This suggests that BARD competently detects BMRs with larger areas and higher flux and tends to miss smaller bipolar regions. The number of uniquely identified BMRs from both AutoTAB and BARD, over both cycles, are represented in Table 2.

3.4. Limitations of the Tracking Algorithm

As with any algorithm, AutoTAB also has some limitations, which are as follows.

(i) One of the greatest challenges that AutoTAB has to face is dealing with multiple BMRs lying closely together. This issue is mostly carried forward from the detection and also affects the tracking, and occasionally multiple BMRs have been tracked as a single one.

(ii) Once a BMR crosses the west limb, AutoTAB loses track of it, but in a few cases, particularly for long-living

BMRs, the same BMR could appear in the east limb after coming back from the farside. In this case, AutoTAB treats this BMR as a new one and gives a unique AutoTAB-ID to it even though they are the same.

(iii) Last but not least, the detection and preprocessing steps are strongly affected by the level of noise in the data while identifying the region of interest, and ultimately this also impacts the tracking.

3.5. Possible Application to Other Data

So far, we have only discussed and demonstrated the application of AutoTAB on MDI and HMI magnetogram data sets. Now the question is, will it be possible for AutoTAB to track the BMRs in other data sets? The answer is yes, but since our algorithm has three main parts, (i) detection, (ii) preprocessing, and (iii) tracking, all of them will not work in the same way with different data sets. Although the preprocessing and tracking are expected to work with various data sets, our detection algorithm, which is only optimized for MDI (Stenflo & Kosovichev 2012) and HMI (Jha et al. 2020) magnetograms, may not work with other data sets. Thus, for a given binary mask of the region of interest, the tracking algorithm can be used, along with preprocessing technique. Here, we would like to emphasize that even for preprocessing and tracking, a few constants or thresholds, such as morph closing kernel size or overlapping area threshold, need to be tuned according to the data. For example, in the case of HMI magnetogram data with an original resolution of 4096×4096, the initial appropriate radius of morph closing kernel size is 16 pixels. The radius is increased to 19 pixels in three different steps, as mentioned in Section 2.2. The area threshold mentioned in Section 2.2 can be scaled by the resolution of the data, in the case of HMI, by a factor of 4. Furthermore, the AutoTAB tracking algorithm can also be used to track the other features in the solar atmosphere by optimizing these parameters.

4. Conclusions

Observational study of the evolving properties of BMRs is crucial to understand the origin of the solar magnetic field and cycle. However, due to the unavailability of a completely automatic and efficient BMR tracking algorithm, the study of the temporal evolutions of various properties of BMRs is limited. We have developed an automatic algorithm, AutoTAB, for tracking the BMR from magnetograms. It works by taking the binary maps of the detected BMR as inputs. For detecting BMRs from magnetograms, we also built an algorithm following the idea of Stenflo & Kosovichev (2012) and Jha et al. (2020). Using our developed detection algorithm, we have produced the binary maps of the detected BMRs from HMI and MDI during Solar Cycles 23 and 24. By feeding these binary maps to AutoTAB, we have produced a homogeneous and comprehensive data set for 9152 tracked BMRs.

AutoTAB is successful in tracking the identified BMRs through their lives on the visible solar disk and in capturing their evolving magnetic and morphological properties. Representative examples of the data set have been presented by showing the snapshots of the evolution of the BMRs at various stages of their life, along with the progression of unsigned flux and $B_{\max}$. As AutoTAB works by checking for the overlap between consecutive binary maps, it remains independent of

the data set used and can work efficiently for tracking other solar features. A comparison between the number of new emergent BMRs in AutoTAB and the BARD catalog shows that AutoTAB consistently detects and tracks more BMRs through the past two solar cycles compared to BARD. As AutoTAB is fully automatic, multiple closely lying BMRs might have been identified as one, which remains one of the main challenges faced by AutoTAB that needs to be addressed in the future.

AutoTAB tracks the BMR in a wide range of its lifetime, from hours to days, which leaves us with a large volume of tracked information. Hence, the tracked BMRs have been classified into different groups, namely, short-lived (lives for less than 8 hr), lifetime (emerges and disperses in the visible surface), and disk passage (coming from and/or going to the farside of the Sun). The tracked BMRs show the usual signatures of solar cycle variation. We also observe that the LT and DP BMRs follow familiar latitudinal and temporal distributions, as seen by the sunspot butterfly diagram. This distribution is not observed in the case of SL BMRs as they are the small features that appear at all times on the Sun. In a follow-up publication, we will further exploit the results of tracked BMRs to compute various properties of BMR, which will help us to identify the origin of BMR formation. The tracked information of the BMRs used in this study, along with the code of the algorithm, will be publicly accessible at https://github.com/sreedevi-anu/AutoTAB, soon after our follow-up article.

Acknowledgments

Appendix
Comparison of the Tracking Results from MDI and HMI Magnetograms

Panels (a) and (b) in Figure 9 show the evolution of NOAA AR11068 from HMI and MDI, respectively. AutoTAB tracked this particular BMR for a total duration of 153 hr from 2010 May 5 19:10 to 2010 May 12 04:46, during its passage across the solar disk, with AutoTAB-ID 10349. As this BMR has been identified in the period when we have observations from both instruments (HMI and MDI), we track it in both data sets for comparison. This BMR was tracked for an extra 6 hr and more consistently by HMI (102 detections) than MDI, with only 87 detections. Fewer detections in MDI are expected because the noise in the data is higher than in the HMI data. Nevertheless, upon closer examination of the evolution of the total flux and B_{max} throughout the life span of the BMR (as shown in Figures 9(c)–(d)), we observe that the B_{max} and absolute total flux show excellent agreement in both data sets. For this particular BMR, the total flux consistently decreases over the lifetime of the BMR, while B_{max} shows a small increase in the first few hours and subsequently decreases. This small increase could be due to the uncertainty in the measurement of B_{max} near the limb.
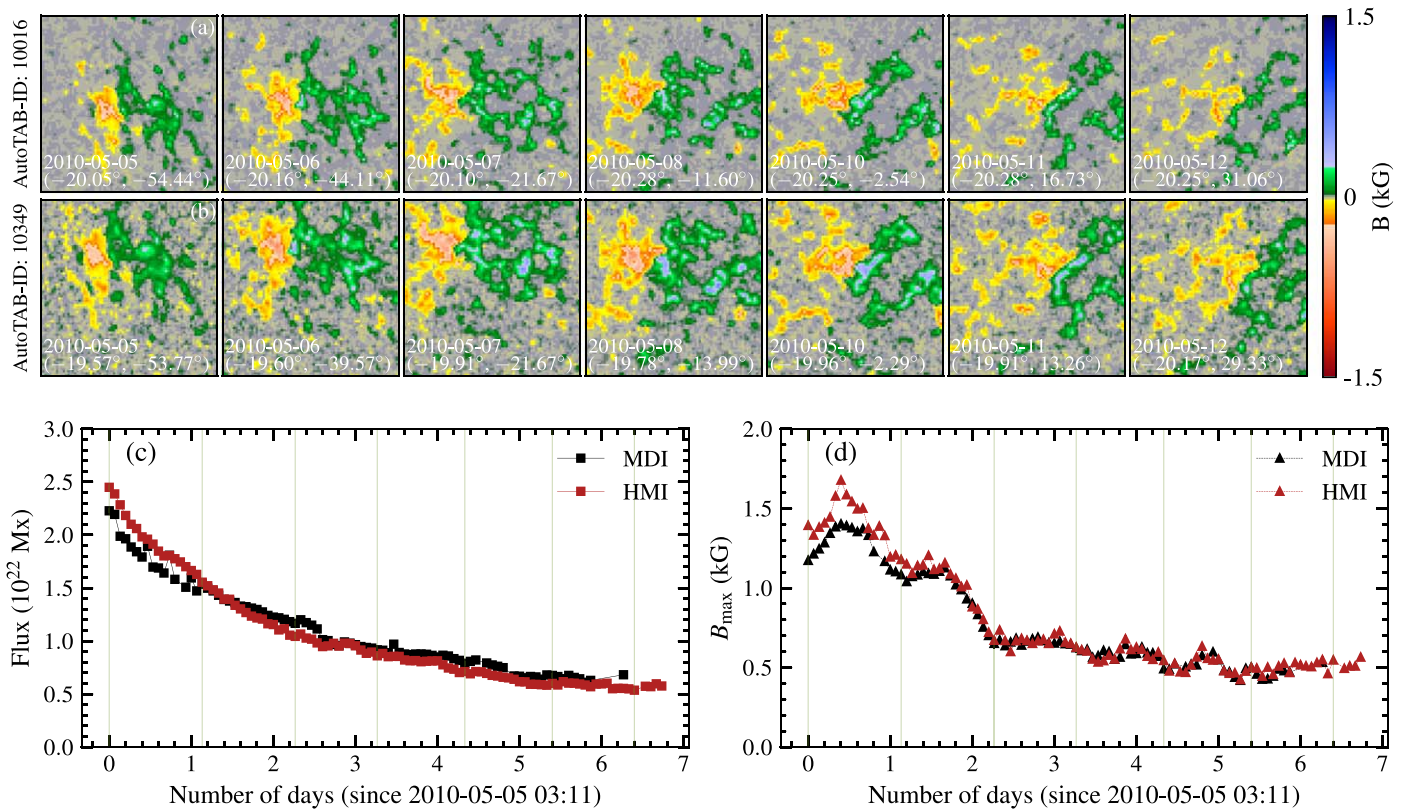
Figure 9. Snapshot of a tracked BMR with NOAA AR11068 during its passage on the observable disk as observed in (a) HMI and (b) MDI. Panels (c) and (d) show the variations of the total flux and B_{max}, respectively, obtained from both data sets. Vertical lines in panels (c) and (d) represent the times corresponding to the snapshots shown in panels (a) and (b), respectively.

ORCID iDs

Anu Sreedevi ⦿ https://orcid.org/0000-0001-7036-2902
Bibhuti Kumar Jha ⦿ https://orcid.org/0000-0003-3191-4625
Bidya Binay Karak ⦿ https://orcid.org/0000-0002-8883-3562
Dipankar Banerjee ⦿ https://orcid.org/0000-0003-4653-6823

References

Babcock, H. W. 1961, ApJ, 133, 572
Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, SoPh, 289, 3549
Bobra, M. G., Wright, P. J., Sun, X., & Turmon, M. J. 2021, ApJS, 256, 26
Dasi-Espuig, M., Solanki, S. K., Krivova, N. A., Cameron, R., & Peñuela, T. 2010, A&A, 518, A7
Getling, A. V., & Buchnev, A. A. 2019, ApJ, 871, 224
Hale, G. E., Ellerman, F., Nicholson, S. B., & Joy, A. H. 1919, ApJ, 49, 153
Harvey, J. W., Hill, F., Hubbard, R. P., et al. 1996, Sci, 272, 1284
Hathaway, D. 2015, LRSP, 12, 4
Higgins, P. A., Gallagher, P. T., McAteer, R. T. J., & Bloomfield, D. S. 2011, AdSpR, 47, 2105
Howard, R. F. 1991, SoPh, 136, 251
Howard, R. F. 1996, SoPh, 169, 293
Howard, R. F., Harvey, J. W., & Forgach, S. 1990, SoPh, 130, 295
Jha, B. K., Hegde, M., Priyadarshi, A., et al. 2022, FrASS, 9, 1019751
Jha, B. K., Karak, B. B., Mandal, S., & Banerjee, D. 2020, ApJL, 889, L19
Jha, B. K., Priyadarshi, A., Mandal, S., Chatterjee, S., & Banerjee, D. 2021, SoPh, 296, 25

Karak, B. B., & Brandenburg, A. 2016, ApJ, 816, 28
Keller, C. U., Harvey, J. W., & Giampapa, M. S. 2003, Proc. SPIE, 4853, 194
LaBonte, B. J., Georgoulis, M. K., & Rust, D. M. 2007, ApJ, 671, 955
Leighton, R. B. 1964, ApJ, 140, 1547
Liu, Y., Hoeksema, J. T., Scherrer, P. H., et al. 2012, SoPh, 279, 295
McAteer, R. T. J., Gallagher, P. T., Ireland, J., & Young, C. A. 2005, SoPh, 228, 55
McClintock, B. H., & Norton, A. A. 2016, ApJ, 818, 7
McIntosh, S. W., Leamon, R. J., Krista, L. D., et al. 2015, NatCo, 6, 6491
Mordvinov, A. V., Karak, B. B., Banerjee, D., et al. 2022, MNRAS, 510, 1331
Muñoz-Jaramillo, A., Navarrete, B., & Campusano, L. E. 2021a, ApJ, 920, 31
Muñoz-Jaramillo, A., Werginz, Z., Vargas, J. P., DeLuca, M., & Navarrete, B. 2021b, Bipolar Active Regions Detection (BARD) Tilt Catalog, v3, Harvard Dataverse, doi:10.7910/DVN/QEMSZ2
Muñoz-Jaramillo, A., Werginz, Z., Vargas-Acosta, J. P., et al. 2016, in 2016 IEEE International Conference on Big Data (Piscataway, NJ: IEEE), 3194
Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, SoPh, 162, 129
Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, SoPh, 275, 207
Schrijver, C. J. 2009, AdSpR, 43, 739
Schunker, H., Baumgartner, C., Birch, A. C., et al. 2020, A&A, 640, A116
Schunker, H., Birch, A. C., Cameron, R. H., et al. 2019, A&A, 625, A53
Schwabe, H. 1844, AN, 21, 233
Sivaraman, K. R., Gokhale, M. H., Sivaraman, H., Gupta, S. S., & Howard, R. F. 2007, ApJ, 657, 592
Stenflo, J. O., & Kosovichev, A. G. 2012, ApJ, 745, 129
Tlatov, A. G., & Pevtsov, A. A. 2014, SoPh, 289, 1143
Ugarte-Urra, I., Upton, L., Warren, H. P., & Hathaway, D. H. 2015, ApJ, 815, 90
van Driel-Gesztelyi, L., & Green, L. M. 2015, LRSP, 12, 1