Article:
Korsos, M.B., Georgoulis, M.K., Gyenge, N. et al. (2020) Solar flare prediction using magnetic field diagnostics above the photosphere. Astrophysical Journal, 896 (2). 119. ISSN: 0004-637X

https://doi.org/10.3847/1538-4357/ab8fa2

The Astrophysical Journal, 896:119 (12pp), 2020 June 20 https://doi.org/10.3847/1538-4357/ab8fa2

Solar Flare Prediction Using Magnetic Field Diagnostics above the Photosphere

M. B. Korsós[1,2], M. K. Georgoulis[4,3], N. Gyenge[5], S. K. Bisoi[6], S. Yu[6,7], S. Poedts[8,9], C. J. Nelson[10], J. Liu[11], Y. Yan[6], and R. Erdélyi[2,11]

[1] Department of Physics, Aberystwyth University, Ceredigion, Cymru, SY23 3BZ, UK
[2] Department of Astronomy, Eötvös Loránd University, Pázmány Péter sétány 1/A, H-1112 Budapest, Hungary; robertus@sheffield.ac.uk
[3] Department of Physics & Astronomy, Georgia State University, Atlanta, GA 30303, USA
[4] RCAAM of the Academy of Athens, 11527 Athens, Greece
[5] Research Computing Support, IT services, The University of Sheffield, 10-12 Brunswick Street, Sheffield, S10 2FN, UK
[6] Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China
[7] New Jersey Institute of Technology, Center for Solar-Terrestrial Research, 323 Martin Luther King Blvd., Newark, NJ 07102, USA
[8] Dept. Mathematics/Centre for Mathematical Plasma Astrophysics, KU Leuven, Celestijnenlaan 200B, B-3001 Leuven, Belgium
[9] Institute of Physics, University of Maria Curie-Skłodowska, Lublin, Poland
[10] Astrophysics Research Centre (ARC), School of Mathematics and Physics, Queen's University, Belfast, BT7 1NN, NI, UK
[11] Solar Physics & Space Plasma Research Center (SP2RC), School of Mathematics and Statistics, University of Sheffield, Hounsfield Road, Sheffield, S3 7RH, UK
Received 2020 March 21; revised 2020 April 25; accepted 2020 April 30; published 2020 June 18

Abstract

In this article, we present the application of the weighted horizontal gradient of magnetic field (WG_M) flare prediction method to three-dimensional (3D) extrapolated magnetic configurations of 13 flaring solar active regions (ARs). The main aim is to identify an optimal height range, if any, in the interface region between the photosphere and lower corona, where the flare onset time prediction capability of WG_M is best exploited. The optimal height is where flare prediction, by means of the WG_M method, is achieved earlier than at the photospheric level. 3D magnetic structures, based on potential and nonlinear force-free field extrapolations, are constructed to study a vertical range from the photosphere up to the low corona with a 45 km step size. The WG_M method is applied as a function of height to all 13 flaring AR cases that are subject to certain selection criteria. We found that applying the WG_M method between 1000 and 1800 km above the solar surface would improve the prediction of the flare onset time by around 2–8 hr. Certain caveats and an outlook for future work along these lines are also discussed.

Unified Astronomy Thesaurus concepts: The Sun (1693); Space weather (2037); Solar corona (1483); Solar flares (1496); Solar activity (1475); Solar photosphere (1518); Solar chromosphere (1479); Solar magnetic fields (1984); Solar active regions (1974); Sunspots (1653); Delta sunspots (1979); Solar active region magnetic fields (1975)

1. Introduction

The short-term (i.e., hours to days) interaction of solar activity manifestations with geospace occurs through a complex series of events, commonly referred to as space weather (SW). Solar activity contributing to SW generally falls under one of four major components: solar flares, coronal mass ejections (CMEs), high-speed solar wind, and solar energetic particles. From these occurrences, two most prominent ones are, arguably, solar flares and CME eruptions (Schwenn 2006). Earth is always impacted by Earth-facing solar flares, with impacts increasing with flare size. Major flares can generate long-lasting radiation storms in the Earth's upper atmosphere causing serious radio or data communication blackouts, among other damaging effects. Flares of larger GOES classes are more frequently associated with CMEs (see, for example, Yashiro et al. 2005).

CMEs, however, can be even more hazardous than flares. They are large clouds of magnetized plasma that may plow right through the Sun–Earth interplanetary space at high speeds. The impact of CMEs on Earth's magnetosphere can influence or even damage a number of socioeconomically vital ground-based (e.g., long-distance oil or gas pipelines, and electric power networks) and space-borne (satellites for communication, navigation (GPS, ISS, etc.) infrastructures (Eastwood et al. 2017). Many of these societal assets and services are key to the global economy, security, and wellbeing. Considerable infrastructure failures by CMEs have indeed happened in the past (e.g., the 1989 March electrical power blackout in Quebec, Canada). The largest known and potentially most dangerous solar eruption in recent history avoided Earth by only ∼30 degrees in 2012 (Temmer & Nitta 2015).

The frequency of occurrence of these most energetic eruptions in the entire solar system follows the 11 year solar cycle. At the peak of the cycle, intense flares and powerful CMEs occur frequently (i.e., around 2–3 daily). It is widely accepted that major solar eruptions (i.e., flares and CMEs, or eruptive flares) originate mostly from magnetically complex, highly twisted and sheared elements of an active region (AR), typically around sunspot groups with mixed magnetic polarities (called δ-sunspots; e.g., Georgoulis et al. 2019; Toriumi & Wang 2019). A key direction of research in solar eruptive activity aims to understand the dynamics of δ-sunspots preceding flare and CME eruptions in order to predict these eruptions within practical timescales, enabling protection of our high-tech facilities and, of course, ourselves. Predicting reliably and accurately these solar eruptions is a major scientific endeavor on its own. The question is not whether but when a potentially devastating flare (or CME) may happen, with adverse effects on our technosphere.

There are a number of methods available in the literature that rely on a range of predictive parameters of solar eruptions (see, e.g., Barnes et al. 2016; Leka et al. 2019, and references therein). Most flare and CME forecast methods apply photospheric magnetic and Doppler data of ARs for forecasting. Some recent, pioneering approaches with various degrees of

success attempt to incorporate solar atmospheric extreme ultraviolet (EUV) data and/or use machine learning in order to improve forecasting accuracy (see, e.g., Qahwaji & Colak 2007; Bobra & Couvidat 2015; Florios et al. 2018; Campi et al. 2019; Kim et al. 2019; Wang et al. 2019). Detailed information on measuring, and the consequent modeling, of the 3D magnetic field structure of an AR would be important to obtain more accurate insight into the preflare evolution locally in the solar atmosphere. However, direct routine observations of the 3D magnetic field in the lower solar atmosphere, above the photosphere up into the top of the chromosphere, are currently not available, with an overwhelming majority of observations referring now to either the line-of-sight (LOS) component or the full magnetic field vector in the photosphere. Nowadays, approximate methods for modeling the local magnetic field vector in the solar atmosphere include its construction using current free (potential, PF) or nonlinear force-free field (NLFFF) extrapolation techniques. In practice, however, to construct an accurate and reliable 3D magnetic field structure of an AR from photospheric measurements is still a challenging task with a number of caveats, see, e.g., Wiegelmann & Sakurai (2012).

Another potentially insightful approach may be the numerical simulation of AR from the subphotosphere to their emergence and evolution in the lower solar atmosphere. With the aim of testing flare prediction with simulated data, a flaring AR with δ-sunspots was modeled by Korsós et al. (2018a). They introduced and applied two flare precursors, part of the WG_M method (Korsós et al. 2019): one is related to the inverted V-shape feature of the WG_M proxy and the other is obtained from the U-shape of the so-called distance parameter prior to each investigated flare at a certain height range in the solar atmosphere. Korsós et al. (2018a) further conjectured the existence of the so-called *o*ptimal height, where the U-shape manifests itself earlier and reaches its minimum value earlier than in the photosphere. In their modeling study it was also shown that these optimal heights agreed reasonably well with the heights of flare occurrence identified by an analysis of thermal and ohmic heating signatures enabled by the magnetohydrodynamic (MHD) simulations of Korsós et al. (2018a). Next, for NOAA AR 11429, Korsós et al. (2018b) used PF extrapolation to construct the 3D magnetic field above the photosphere and studied the preflare evolution of this AR prior to two M-class flares. There, it was found again, that the earliest onset time estimation was enabled at a distinct and specific height range, i.e., at an optimal height, when compared to patterns derived from data in the photosphere or other atmospheric heights.

The two abovementioned studies prompt us to attempt to further establish the details of the conjectured wide applicability and benefits of 3D preflare analyses using a larger sample of flaring ARs and seeking the relevant optimal height (s) with better statistical significance. This work is organized as follows: Section 2 describes the adopted tools for the preflare analysis of a given 3D solar magnetic structure. Section 3 introduces and describes the application of the WG_M to the lower solar atmosphere before the flare occurrences. Section 4 discusses our results in detail. We summarize our key findings and draw our conclusions in Section 5. The Appendix contains more cases that support the analysis presented in the main body of the work.

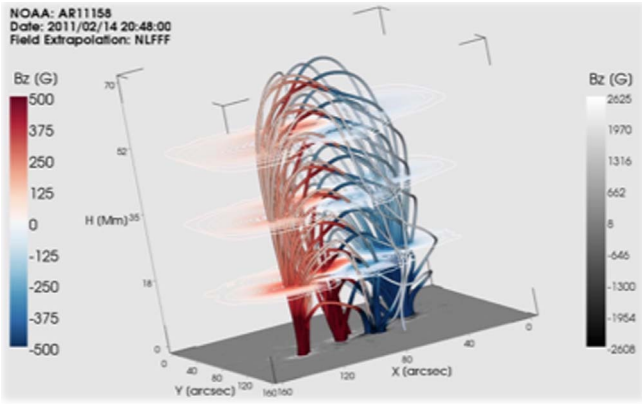

Figure 1. Three-dimensional NLFFF extrapolation of AR 11158 at 20:48:00 on 2011 February 14. The red–blue color bar refers to the positive and negative polarity magnetic field values at different heights in the solar atmosphere. The gray color bar represents the photospheric vertical magnetic field component, B_z. The three horizontal slices represent the identified sunspot at various heights in the lower solar atmosphere.

2. Methodology

Motivated by the case studies of Korsós et al. (2018a, 2018b), we now extend the application of the WG_M method to more ARs by constructing a data catalog of sunspots using 3D PF and NLFFF extrapolations. An extrapolation example from the collected data catalog is shown in Figure 1.

With extrapolations accomplished, the WG_M method is applied to the 3D PF and NLFFF extrapolated data. The WG_M method is then applied to both data sets and the results are compared.

It is clear that each of the two extrapolation types, PF and NLFFF, has its strengths and weaknesses: the NLFFF is most likely a much better reconstruction method of the magnetic field in the lower solar atmosphere of an AR than the PF, it is considerably more expensive computationally. Hence, if the advantage in using NLFFF against PF toward improving the lead time for flare onset prediction is trivial, one might opt to save computing time by using the PF approximation instead of the NLFFF one.

2.1. Selection of Studied ARs

Before we initiate the 3D analysis of the preflare dynamics of ARs with the WG_M method, a consistent data set of ARs is required. The data processing and the PF/NLFFF magnetic field extrapolations of ARs were carried out with the extensive use of SolarSoft,[12] with purpose- and instrument-specific routines. For an AR to be included in the analysis, the following four selection criteria are set to be satisfied:

1. The studied AR is located between $-60°$ and $+60°$ in central meridian distance during the examined period of time.
2. The AR hosted a GOES *X*-class flare. This is dictated by practical, computational reasons and can be revisited when sufficient resources are available.
3. The easternmost central meridian distance of X-flaring locations is not more than $\sim -40°$.
4. The AR had at least one δ-spot(s).

[12] http://www.mssl.ucl.ac.uk/surf/sswdoc/solarsoft/ssw_install_howto.html

Over Solar Cycle 24, 13 ARs were found to satisfy the above four selection criteria.

2.2. 3D Lower Atmospheric Magnetic Field of ARs

Both the PF and the NLFFF extrapolations require photospheric boundary conditions. We employed the Solar Dynamics Observatory Helioseismic and Magnetic Imager (HMI, Scherrer et al. 2012) LOS magnetograms as a boundary condition for the PF extrapolation. For the NLFFF, the vector magnetic field measurements of HMI Active Region Patches (HARP; Bobra et al. 2014) are used as a boundary condition. In this work, the magnetogram data were studied every hour and were resized by a factor of 8 thus giving rise to photospheric magnetograms with a pixel size of 4 arcsec.

2.2.1. PF Extrapolation

To determine the magnetic field above the photosphere with the PF extrapolation method, we employed the linear force-free field (LFFF) IDL extrapolation code (see www.heliodocs.com), that is based on Gary (1989), where we set the force-free parameter (α) equal to zero.

In brief, the PF is the simplest possible assumption for the solar atmospheric magnetic field. The LOS magnetograph is used as a boundary condition to solve Laplace's equation,

$$\nabla^2\phi = 0, \tag{1}$$

where ϕ is the associated scalar potential for the PF.

2.2.2. NLFFF Extrapolation

We apply the direct boundary integral formulation for nonlinear force-free magnetic field extrapolation as outlined by Yan & Li (2006). The predecessor of the NLFFF extrapolation used here is the so-called boundary integral method, first developed by Yan & Sakurai (2000), and recently implemented with GPU-acceleration by Wang et al. (2013).

The method uses the Green's function to reformulate the NLFFF problem. The obtained nonlinear integral equations allow the independent calculation of the vector magnetic field at any location of the extrapolation volume. The method considers the half-space above the lower boundary with vanishing magnetic field at infinity. The solution at a given point i inside the volume V for the boundary magnetic field values ($\boldsymbol{B}_0$) on $\Gamma = \partial$V, is given by:

$$c_i \boldsymbol{B}_i = \oint_\Gamma \left(\boldsymbol{Y} \frac{\partial \boldsymbol{B}}{\partial n} - \frac{\partial \boldsymbol{Y}}{\partial n} \boldsymbol{B}_0 \right) d\Gamma, \tag{2}$$

with $c_i = 1$ for points in the volume and $c_i = 1/2$ for boundary points. $\boldsymbol{Y}$, in Equation (2), is a kernel function that depends on $\boldsymbol{B}$ (for more details, see Equation (19) in Yan & Sakurai 2000).

2.3. Catalog

We tracked sunspots above their photospheric altitudes using the Yet Another Feature Tracking Algorithm (YAFTA; Welsch & Longcope 2003; DeForest et al. 2007). YAFTA is accessible from the Solarsoft IDL library. The detection algorithm is based on the so-called clumping method, which enrolls together all contiguous-like-polarity pixels with absolute flux densities above a specified threshold, and marks them as unique elements. In our study, YAFTA groups pixels into an element (such as a sunspot) when the given criteria are satisfied: (i) minimum number of pixels is 30 and (ii) local vertical magnetic fields exceed a |150| G threshold.

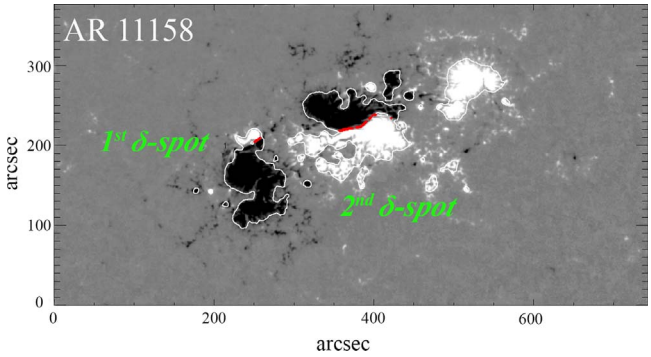

Figure 2. Magnetogram snapshot showing two δ-spots in NOAA AR11158 on 2011 February 14 at 20:48:00. The red dotted lines are the automatically identified PILs of the AR. White contours define areas that enclose the identified strong flux elements.

The magnetic field strength, area, and cross-sectional diameter of all identified sunspots of ARs are then recorded for each relevant frame and saved in the sunspot data catalog. The generated 3D catalog includes the area, mean magnetic field, and location (Carrington coordinates, L and B) of identified sunspots using a 45 km step above the photosphere toward the lower corona. The 45 km step size was chosen as the highest vertical resolution, i.e., smallest grid size, implemented in the NLFFF extrapolations.

2.4. Identification of δ-spot of an AR

Before we begin to apply the WG_M method, we need to identify the δ-spot(s) of the selected ARs. Here, we adopt and employ the automatic PIL recognition algorithm developed by Cui et al. (2006). The program first computes the horizontal component of the PF. Next, the pixels are selected, based on whether the strength of the deduced transverse component of the magnetic field is higher than |150| G. Also, the pixels are identified where the horizontal gradient of the longitudinal component of the magnetic field is larger than |50| G/Mm. In the example given in Figure 2, the contoured area with PIL(s) corresponds to the δ-spots in the region, where we apply the WG_M method.

3. Analysis of Preflare Behavior Based on Extrapolation Data

Through the case of AR 11158 with an X2.2 flare at 01:56 on 2011 February 15, let us now demonstrate the application of the WG_M method as a function of height, applying it to PF and NLFFF extrapolations. AR 11158 has two δ-spots, labeled as the first and second δ-spots, see Figure 2.

For both the PF and NLFFF extrapolation results, let us now track the evolution of (i) the WG_M proxy, (ii) the D_{pn} distance between the area-weighted barycenters of opposite polarities, and (iii) the unsigned magnetic flux Φ in the first and second δ-spots at consecutive 45 km steps in height (z). To identify the inverted V- and U-shape preflare features, we use the maximum and minimum values of the best nth degree polynomial fit to the WG_M and D_{pn} data, respectively.

We, hereafter, focus on those atmospheric heights, where the inverted V- and U-shape are identifiable prior to the X2.2 flare,

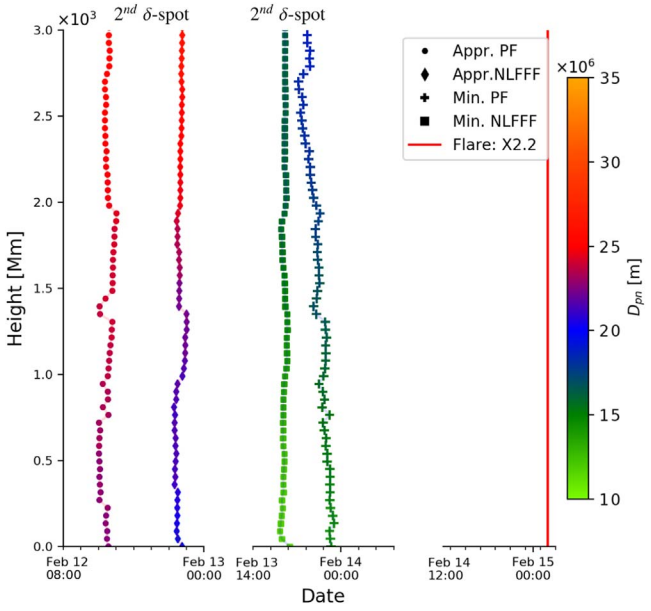

Figure 3. AR 11158: times associated with the start (point/diamond symbols) and closest (plus/square symbols) convergence of opposite-polarity area-weighted barycenters as a function of height. Results obtained using the PF and the NLFFF extrapolations are shown by point/plus and diamond/square symbols, respectively. The moments of starting and closest convergence times are deduced by the maximum and minimum values, respectively, of the best nth degree polynomial fit to the $D_{\rm pn}$ data. The color bar gives information about the actual value of $D_{\rm pn}$. Results from the second δ-spot in the AR are shown. The red vertical line marks the X2.2 flare occurrence time that occurred from the second δ-spot.

for both δ-spot cases. For NOAA AR 11158 we have found the following:

1. In the case of the first δ-spot, the inverted V-shape of the WG_M and the U-shape of the $D_{\rm pn}$ are not discernible in the two extrapolation cases (see, e.g., Appendix Figures A1(a)–(b)).
2. In the case of the second δ-spot, the inverted V- and U-shape each are observed prior to the X2.2 flare in both extrapolation cases (e.g., see Figures A2–A3(a)–(b)). These two preflare behaviors concurrently manifest, from the photosphere up to 3000 km in the low corona. Also, this δ-spot was the actual major source region of the X2.2 flare (e.g., Wang et al. 2012).

After, the relevant inverted V- and U-shapes are successfully identified at a given height, we carry on investigating their evolution as a function of height. In particular, based on Korsós et al. (2018a, 2018b), we concentrate on the starting (first orange dots in Figures A1–A3) and finishing (first blue dots in Figures A1–A3) moments of the converging phase of the U-shape illustrated at various heights/cases, because the elapsed time between these two moments provides information about the expected flare onset time (see for more details, see Figure 5 of Korsós et al. 2019).

Figure 3 shows the starting time ($D_{\rm pn}^{Max}$, point and diamond symbols) and finishing time ($D_{\rm pn}^{Min}$, plus and square symbols) of the converging phase at each 45 km step in the two δ-spot cases. The point/plus symbols represent the data derived from the PF extrapolation and the diamond/square symbols indicate the results of the NLFFF extrapolation for the constructed 3D lower atmospheric magnetic fields of AR 11158. The color code corresponds to the actual value of the $D_{\rm pn}$. Also, the red line marks the onset time of the X2.2 flare in AR 11158. We notice that the converging phase begins earlier and reaches its minimum distance also earlier at a certain height (referred to as the optimal height) than it does at the photosphere, as also found in Korsós et al. (2018a, 2018b).

Identifying the corresponding optimal heights, we estimate the expected largest flare intensity class ($S_{\rm flare}$) and onset time ($T_{\rm est}$), as in Korsós et al. (2019), to investigate the applicability of WG_M for the second δ-spot:

1. In the PF case, the optimal height is 1395 km because the converging phase (the max point of the fitted nth order polynomial) started 1.2 hr before and finished 2 hr earlier than in the photosphere, enabling the maximum lead time prior to flare onset as a function of solar atmospheric height.
2. In the NLFFF case, the converging phase began 0.9 hr before and finished 0.7 hr earlier at best, at 810 km.

Therefore, with the PF approximation one can estimate the expected flare onset time 1.3 hr earlier than in the case of NLFFF. It is also worth noting that $T_{\rm est}$ is very close to the actual values of T_{D+F} in the case of NLFFF. At these two optimal heights, the $S_{\rm flare}$ of the investigated flare are found to be fairly well estimated, i.e., the expected flare intensity is determined as an X-class flare (see Table A1).

The above obtained two estimates ($S_{\rm flare}$ and $T_{\rm est}$) are summarized in Table A1. Furthermore, Table A1 includes information about the time prior to the flare at the start ($T_{\rm lmp}^{C}$) and closest (T_{mlmp}^{M}) convergence at the optimal height, for both extrapolation approaches. T_{mlmp}^{M} can also be understood as the lead time at the corresponding optimal height, as we estimate the flare onset time from $D_{pn}^{\rm Min} - D_{pn}^{\rm Max}$. Actually, the values of $T_{\rm lmp}^{M}$ indicate how much time one could gain in the flare onset time estimation if one applies the WG_M method at an identified optimal height.

3.1. Additional Results of PF versus NLFFF

Here, we outline the results of investigating three additional ARs that demonstrate how to further improve the flare onset time prediction capability of the WG_M method by means of an optimal height analysis. For the detailed comparative analysis of ARs 11166, 11283, and 12192, see Table A1 in Appendix A. A visual summary of these results is given in Figure 4.

In Figure 4(a), the columns show the gained time at the end of the converging phase ($T_{\rm lmp}^{M}$) at the optimal height. In Figure 4(b), the columns represent the optimal height of the particular flare events. In Figures 4(a) and (b), the gray/line-crossed columns refer to PF/NLFFF extrapolations. The plotted values expressed in numbers are given in Table A1. The abscissas of Figures 4(a) and (b) are labeled with the name of the AR that hosted the flare in the same order as the names of the ARs listed in Table A1.

From Figure 4 and Table A1, we conclude that the optimal heights and the lead-time improvements are not identical for the four studied ARs and five major flare cases. We also note that, interestingly, the PF has better $T_{\rm lmp}^{M}$ improvement in four cases out of five. Based on this finding, one might be tempted to use PF in further studies, for computational efficiency. This might be changed in the future, of course, when computational advances allow for the routine application of NLFFF (or, indeed, even more sophisticated modeling) in much shorter times.

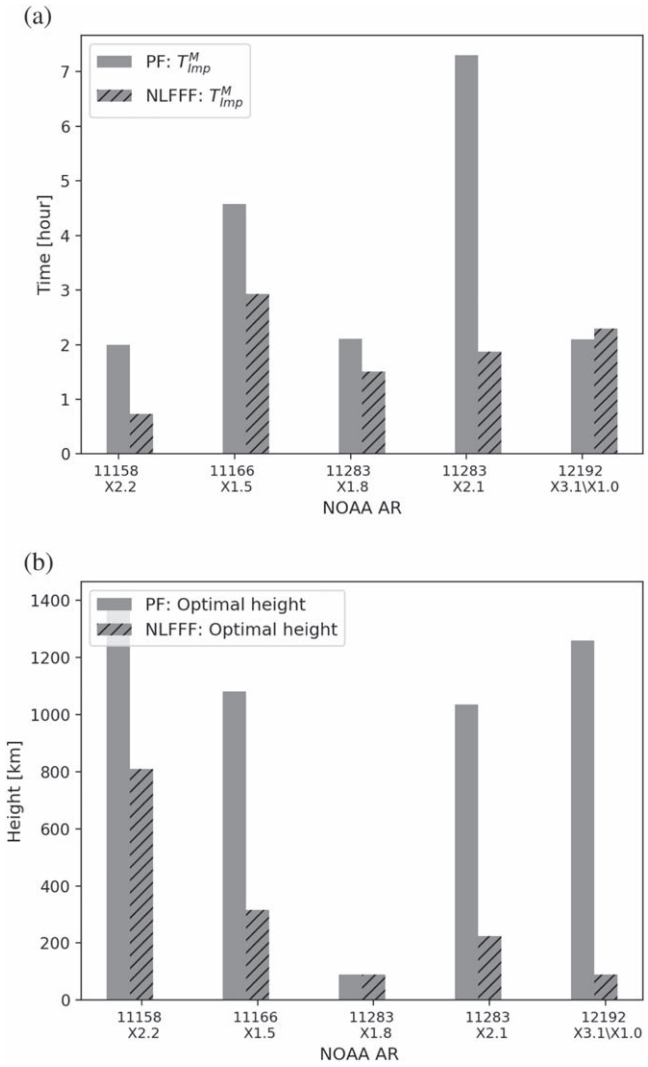

Figure 4. Summary of WG_M lead times (a) and optimal heights (b) for PF (gray columns) and NLFFF (line-crossed columns) extrapolations in case of four different, eruptive ARs and five major flares.

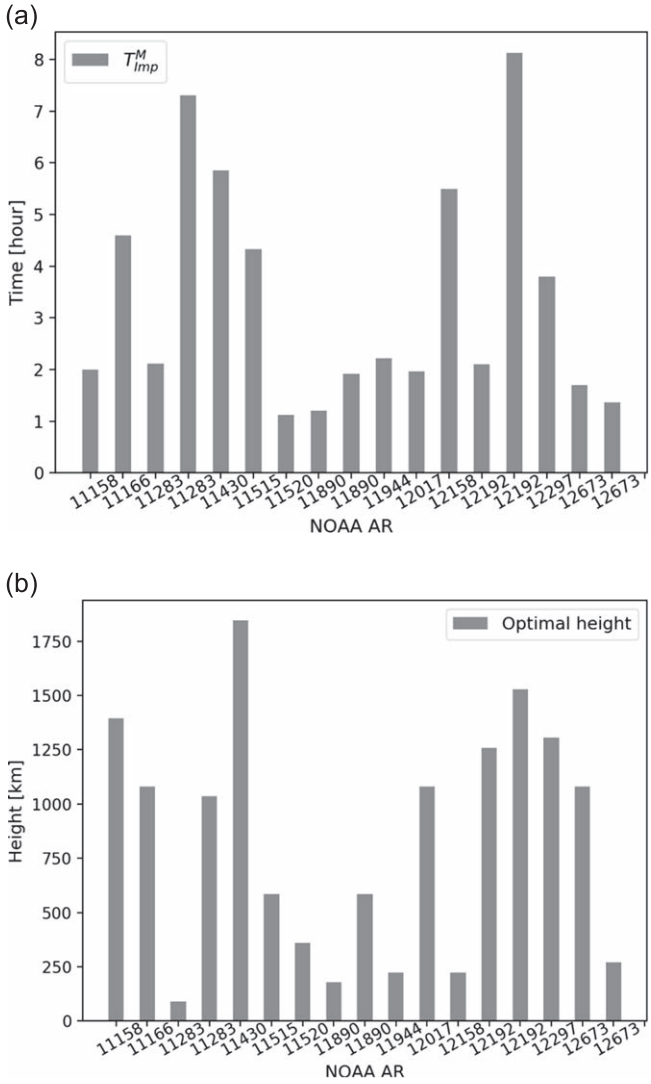

Figure 5. Summary of WG_M lead times (a) and optimal heights (b) for 13 AR cases under the PF extrapolation. For the actual values, see Tables A1 and B1.

4. Application of PF to More Active Regions

Let us now analyze 13 more flaring AR cases (see Table B1), which all satisfy the selection criteria given in Section 2.1.

First, we constructed the 3D PF extrapolations and identified δ-spots. Next, the WG_M method was applied to each δ-spot as a function of height with steps of 45 km. The analysis of Section 3 were carried out to identify the relevant inverted V- and U-shapes of the WG_M and D_{pn} parameters. Our findings are summarized in Figure 5 and Table B1.

Similar to the first four examples, we found again that the evolution of the three parameters, i.e., WG_M, D_{pn}, and Φ, vary as a function of height in all identified δ-spots. In all cases, the converging phase began earlier and reached its shortest distance also earlier at their respective optimal heights. Furthermore, we note here that we could not identify the concurrent precursor presence of the inverted V- or U-shape in the cases of nonflaring δ-spots (see two random examples of AR 11158 in Figure A1(a) and AR 12297 in Figure A6).

In regards to the 13 flaring ARs studied, we find that (i) the lead-time (T^M_{Imp}) values range between ~1 and 8 hr at the identified optimal heights, and (ii) optimal heights seem to fall under two distinct intervals, namely, ~90–600 km and ~1000–1800 km.

The estimated flare onset times (T_{est}) are much closer to the actual values of T_{D+F} at the optimal height of 1000–1800 km, even considering the 7.2 hr uncertainty. For the 90–600 km range, differences can be as large as 2 days between T_{est} and T_{D+F} (see, e.g., AR 11515, in Table B1).

Based on our findings above, there is a practically effective optimal height range of 1000–1800 km, for which the prediction capability of the WG_M method is considerably improved. In order to determine the potential lead time between 1000 and 1800 km, we use T^M_{Imp}. This is very important information because estimating the onset time of a flare seems to rely on the linear relationship between the converging and diverging motions of the opposite polarities (see Figure 5(a) of Korsós et al. 2019). Therefore, as a summary we conclude that, based on the values of T^M_{Imp} in Tables A1 and B1, we could estimate the flare onset times ~2–8 hr earlier with the WG_M method at an altitude range of ~1000–1800 km above the photosphere using the PF method, rather than working on the photospheric magnetic field.

5. Conclusions

In this work, we applied the WG_M method at different heights between the photosphere and the low corona to 3D magnetic field extrapolations of ARs containing δ-spots at their observed photospheric base, in order to identify an optimal height range where flare prediction could be achieved earlier than using only photospheric data. Our aims were realized by (i) implementing potential (PF) and nonlinear magnetic field (NLFFF) exploration techniques and (ii) creating a sample of 3D magnetic maps of sunspots in the lower solar atmosphere.

As in our previous works, we considered two unique preflare patterns of the WG_M method (namely, the inverted V-shape of the WG_M proxy and the U-shape of the distance (D_{pn}) parameter) as a function of height, instead of studying the otherwise popular quantity of the free energy of ARs. We still do not have a detailed physical explanation to capture the two preflare patterns. However, we put forward our conjecture in Section 4 of Korsós et al. (2019). Namely, that a current sheet develops during the convergence phase of the two opposite-polarity area-weighed barycenters, while magnetic reconnection takes place after the end of the divergence phase. The validation of this conjecture should be confirmed by 3D numerical simulations but that is beyond the scope of this work. An alternative suggestion was put forward by Tlatov et al. (2018), for more details see Figure 5 of their paper.

In this study, we compared the results obtained by applying the WG_M method to PF and NLFFF extrapolation data in four different flaring ARs. We discovered that, at a certain height, called the optimal height, the fitted U-shapes enabled us to estimate the expected flare onset time earlier than using only magnetic data at the photospheric level. This is a key finding of this work.

We also observed that the identified optimal heights and lead-time improvements for estimating the flare onset time vary with the applied extrapolation method. Namely, we found that sometimes the yielded lead time by PF is better by up to a few hours than using NLFFF extrapolation. This is another important practical aspect because the CPU running time differences between the PF and NLFFF extrapolations are substantial. It might be prohibiting to apply the WG_M method under the NLFFF extrapolation in near-realtime due to computational limitations, but the PF extrapolation may offer a viable alternative.

Next, we restricted to PF extrapolations only. We found that the potential lead-time improvement for estimating the flare onset time varies in the interval (2, 8) hr if we apply the WG_M method to an identified sunspot between $\sim$1000 and 1800 km above the photosphere.

In this study, we do not have a negative sample (i.e., nonflaring ARs with δ-spots), because ARs that form δ-spots tend to be flaring (e.g., Georgoulis et al. 2019; Toriumi & Wang 2019, and references therein). It seems that there are δ-spots in some of the ARs studied, here, which do not show the inverted V- and U-shapes though. In the future, we will also extend this work in at least two directions: (i) carry out magnetic field extrapolations and use the WG_M method to determine the evolution of the nonflaring and flaring δ-sunspots with flares of lower GOES class (e.g., M-, and even C-classes); and (ii) test the findings of this work, as well as the flare precursor capability of the WG_M method, with a larger statistical sample of ARs. We will also investigate, for near-realtime operational purposes, the time needed to determine whether a given D_{pn} is indeed the minimum and how this time relates to the estimated and actual onset times.

M.B.K. is grateful to the University of Sheffield for the support received while carrying out this research and STFC grant ST/S000518/1 to Aberystwyth University. M.B.K. also acknowledges the open research program of CAS Key Laboratory of Solar Activity, National Astronomical Observatories, No. KLSA201610. M.B.K. and R.E. acknowledges the CAS Key Laboratory of Solar Activity, National Astronomical Observatories Commission for Collaborating Research Program for support received to carry out part of this work. R.E. acknowledges the CAS Presidents International Fellowship Initiative, grant No. 2019VMA052. R.E. is also grateful to the Science and Technology Facilities Council (STFC, grant No. ST/M000826/1) UK and the Royal Society for enabling this research. S.K.B. acknowledges the support by the National Natural Science Foundation of China (NSFC grant No. 11750110422, 11433006, 11790301, and 11790305). C.J.N. thanks the Science and Technology Facilities Council (STFC) for the support received to conduct this research through grant No. ST/P000304/1. Y.Y. acknowledges NSFC supports with the grant Nos. 11790300 and 11790301.

Appendix A
Application of the WG_M-method at Different Atmospheric Heights

A.1. AR 11158

Figures A1–A3 show the evolution of WG_M, D_{pn}, and Φ before the X2.2 flare, which occurred at 01:56 on 2011 February 15 in AR 11158. Panels (a) and (b) of each figure reveal the evolution of the various preflare indicators, applied to the PF and NLFFF extrapolations. The upper panels in each figure are the temporal variations of WG_M. The preflare behavior of WG_M is fitted by an nth-order polynomial (red line), where the orange dot corresponds to the maximum of WG_M. The middle panels demonstrate the evolution of D_{pn}. The consecutive maximum–minimum–maximum (orange–blue–orange dots) locations of the fitted nth-degree polynomial denote the full converging–diverging phase uncovered by D_{pn}. The vertical blue stripes mark the flare peak time. The bottom panels show the evolution of the unsigned magnetic flux (Φ).

To find the best nth-order polynomial, we fit the data with a range of polynomial degrees and pick the degree that has the lowest rms error.

Figures A1–A2 show the evolution of WG_M, D_{pn}, and Φ at the photospheric level in the case of first and second δ-spots of AR 11158. Furthermore, Figures A3(a)–(b) correspond to findings obtained at the identified optimum heights in the case of PF an NLFFF analyses of the second δ-spot, respectively. The corresponding gain prediction times at the identified optimum heights are summarized in Table A1.

A.2. AR 12297

AR 12297 was the host of an X2.1 flare at 16:22 on 2015 March 11. This AR has two δ-spots determined by the method of Cui et al. (2006), labeled as the first and second δ-spots,

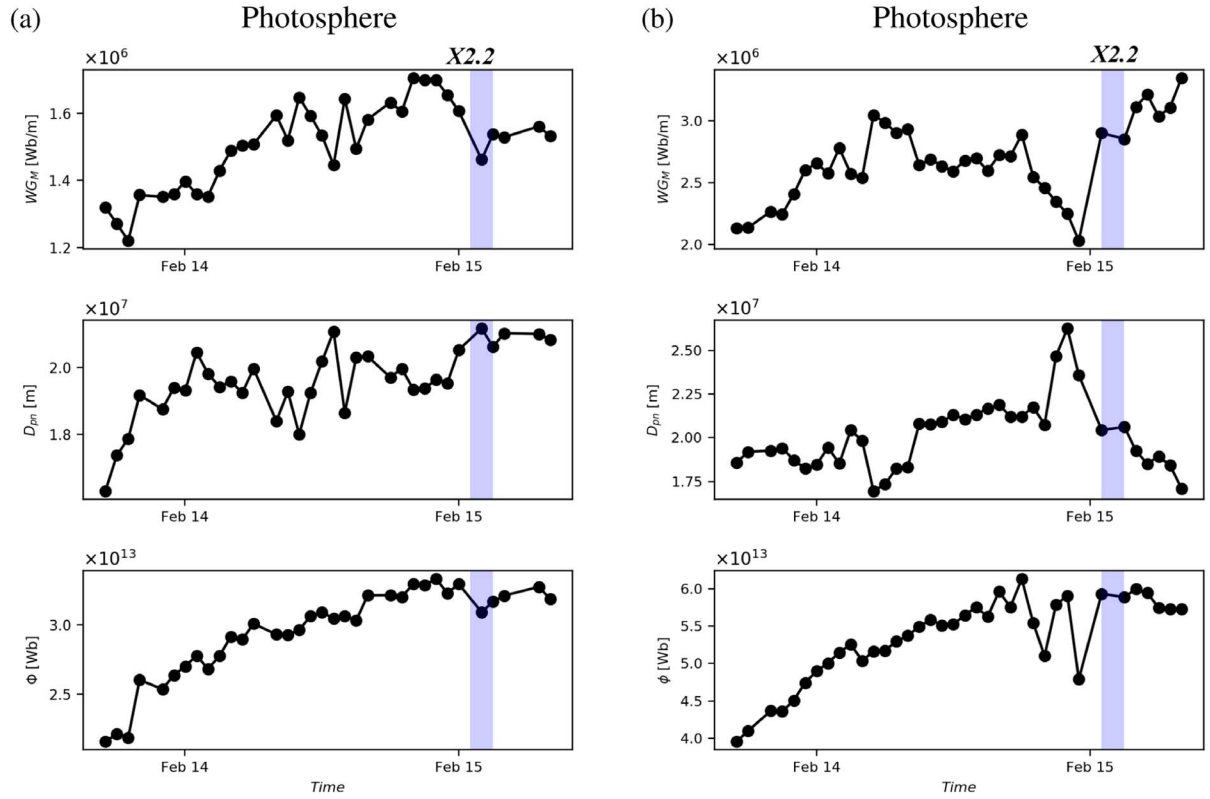

Figure A1. Columns (a) and (b) show the graphical visualization of the result of the WG_M analysis for the first δ-spot of AR 11158 (for the context image see Figure 2a) at the photosphere. Column (a) is the PF and (b) the NLFFF extrapolation case, respectively. The first δ-spot was not the cradle of the X2.2 flare.

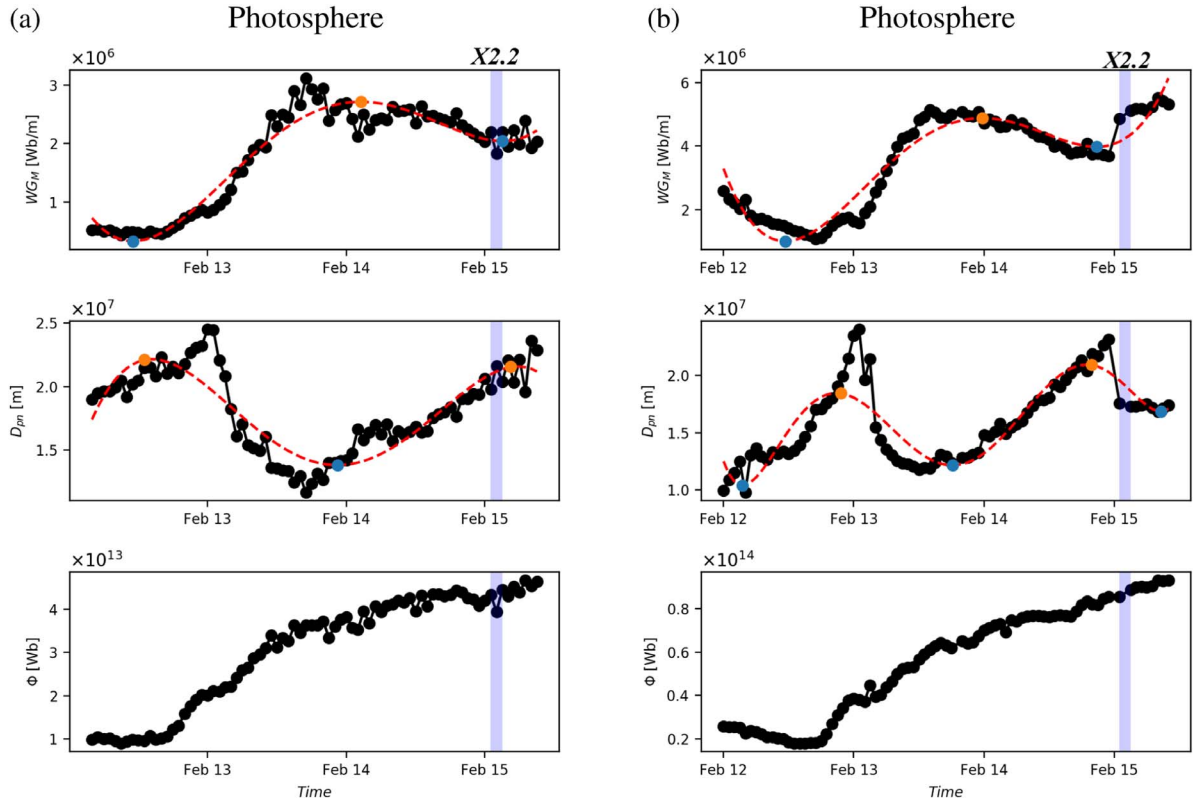

Figure A2. Same as Figure A1 but for the second δ-spot of AR 11158 at the photosphere. To identify the inverted V- and U-shape preflare features, we use the maximum (orange dots) and minimum (blue dots) values of the best nth degree polynomial fit (red dashed line) to the WG_M data and to the D_{pn} data. The blue shaded vertical line marks the X2.2 flare occurrence time. This second δ-spot was the host of the X2.2 flare.

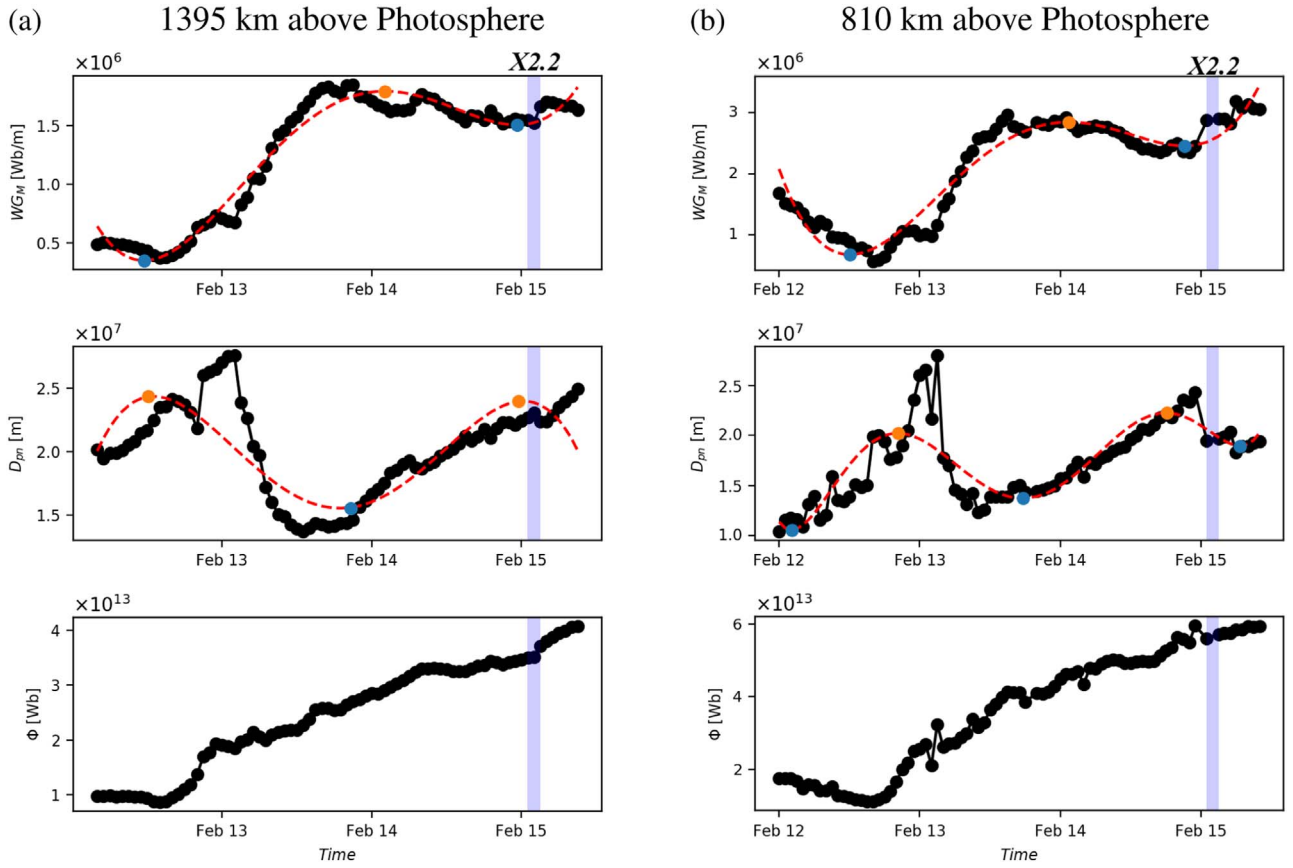

Figure A3. Same as Figure A2 for AR 11158 where panels (a) and (b) now illustrate the evolution of the flare precursor parameters at the optimal height. (a) The optimal height is 1395 km above the photosphere in the PF case. (b) The optimal height is 810 km from the photosphere in the NLFFF case.

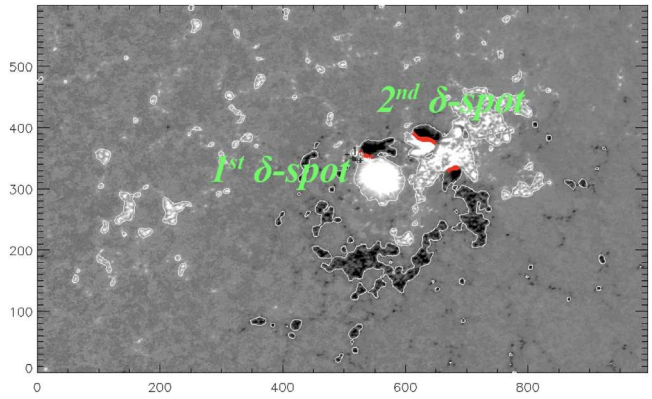

Figure A4. Magnetogram snapshots showing the analyzed two δ-spots of *AR 12297* on 10/03/2015 at 22:00:00. The red dotted lines are the automatically identified PILs of the AR. The white countered areas show the identification of strong flux elements.

respectively, in Figure A4. Figures A5–A6 show the temporal variation of WG_M, D_{pn}, and Φ at the photosphere (panel (a)); and 500 km above the solar surface (panel (b)), applied to the PF magnetic extrapolation data. Panels (a) and (b) of Figures A5 reveal the evolution of inverted V- and U-shapes of the WG_M and D_{pn} parameters before the X2.1 flare, in the case of the first δ-spot. Note that, in Figure A6, we cannot identify the inverted V- and U-shapes before the X2.1 in the case of the second δ-spot. The reason is because the first δ-spot was actually the source of the X2.1 flare (Lu et al. 2019).

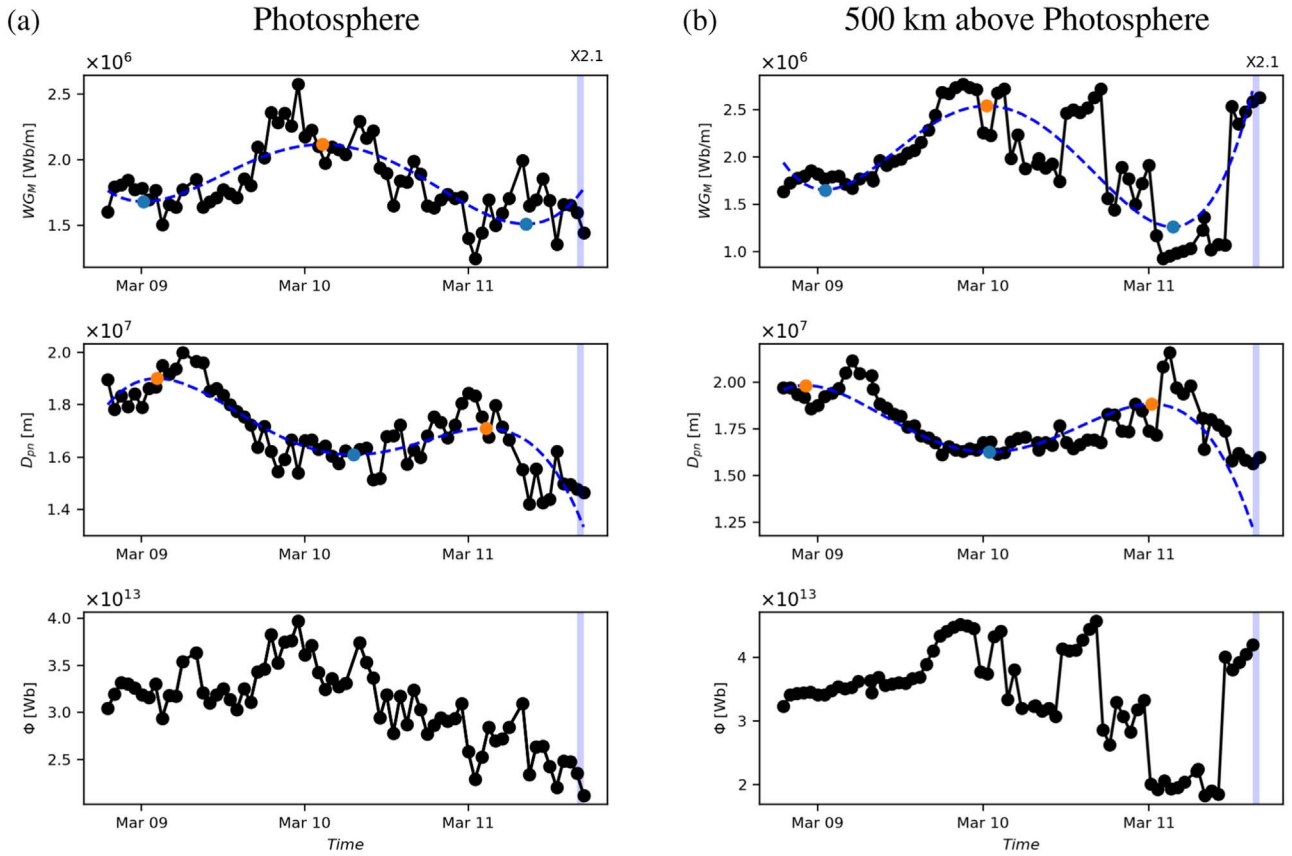

Figure A5. Columns (a) and (b) show the graphical visualization of the result of the WG_M analysis for the first δ-spot of *AR 12297* (for the context image, see Figure A4). Column (a) is for the photosphere and panel (b) the 500 km level case, respectively.

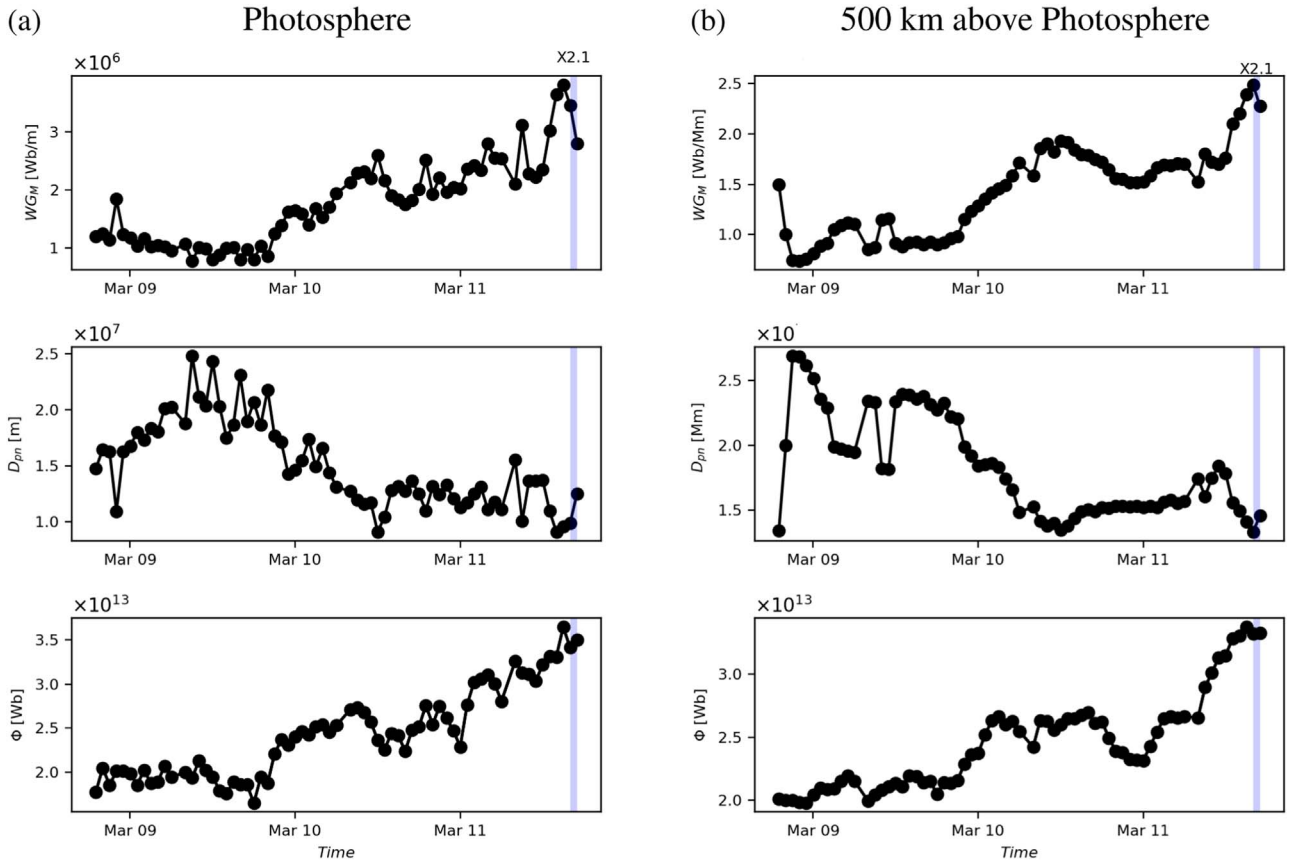

Figure A6. Same as Figure A5 but for the second δ-spot of *AR 12297*.

Table A1
Table to Compare the Results of Applying the WG_M Method, Obtained by Means of PF and NLFFF Extrapolations in Four Investigated ARs

		PF						NLFFF					
NOAA AR	Flare Intensity	$T^C_{\rm Imp}$ (h)	$T^M_{\rm Imp}$ (h)	Opt. Height (km)	$S_{\rm flare}$	$T_{\rm est}$ (h)	T_{D+F} (h)	$T^C_{\rm Imp}$ (h)	$T^M_{\rm Imp}$ (h)	Opt. Height (km)	$S_{\rm flare}$	$T_{\rm est}$ (h)	T_{D+F} (h)
11158	X2.2	1.2	2.0	1395	X	29.5	42	0.9	0.7	810	X	30	32.5
11166	X1.5	0.6	4.6	1080	>M5	39.1	46.8	0.5	2.9	315	X	39.5	36.2
11283	X1.8	1.6	2.1	90	X	22.2	53.6	1.2	1.5	90	X	21.8	59.6
	X2.1	6.7	7.3	1035	X	47.4	33.7	3.8	1.8	225	X	50.1	30.1
12192	X3.1/ X1.0	2.5	2.1	1260	X	21.9	43.7/63.7	1.7	2.3	90	X	41.3	32.9/52.9

Note. The table includes how many hours earlier the converging phase ($T^C_{\rm Imp}$) began and reached the minimum value ($T^M_{\rm Imp}$) at the optimal height (Opt. Height) when compared to the photosphere. $S_{\rm flare}$ is the estimated flare class. $T_{\rm est}$ is the estimated flare onset time in hours. T_{D+F} is the elapsed time from the moment of the closest location of the two opposite-polarity barycenters to flare onset in hours.

Appendix B
Brief Comparison Analyses of Preflare Behavior of the Further Three Active Regions Based on PF and NLFFF Extrapolations

B.1. AR 11166

The second example is AR 11166. Here, we investigate the preflare states before an X1.5 flare. This flare occurred in the single δ-spot of the AR at 23:23 on 2011 March 9 (Vemareddy & Wiegelmann 2014). We could identify the prominent and typical inverted V- and U-shapes prior to X1.5 in the vertical region from the photosphere up to 2000 km at each 45 km step. In the NLFFF extrapolation case, we further noticed that two consecutive precursors of the WG_M and the $D_{\rm pn}$ appear instead of one only above 500 km. Therefore, we can uniquely identify the optimal height only in the case of the first modeling approach. In the second case, we do not have photospheric reference data (see Figure B1). In the PF case, the optimal height is identified at 1080 km where the converging phase started ($T^C_{\rm Imp}$) 0.6 hr earlier and ended ($T^M_{\rm Imp}$) 4.6 hr earlier than in the photosphere. In the NLFFF case, $T^C_{\rm Imp}$ is 0.5 hr and $T^M_{\rm Imp}$ is 2.9 hr corresponding to an optimal height of 315 km. Here, we could estimate the flare onset time a couple of hours earlier using either of the extrapolations. Unfortunately, $T_{\rm est}$ seems to be rather overestimated, with 8 hr, in the PF extrapolation case. However, the $T_{\rm est}$ value is well in agreement with the T_{D+F} value when applying data from the corresponding NLFFF extrapolation. In the PF case, using the equation of Figure 5 of Korsós et al. (2019), we estimated the $S_{\rm flare}$ as an M-class that is an underestimate when compared to the measured X1.5 flare intensity. In the NLFFF case, $S_{\rm flare}$ is found to be correct.

B.2. AR 11283

The next example is *AR 11283* with a flare of X1.8 that occurred at 22:20 on 2011 September 6 and with another one of X2.1 at 22:38 on 2011 September 7. These two flares had their cradle in the same δ-spot of the AR (Liu et al. 2014). Here, the characteristic preflare behavior of the WG_M and the $D_{\rm pn}$ are evaluated, using the appropriate 3D constructed magnetic field structures, where both

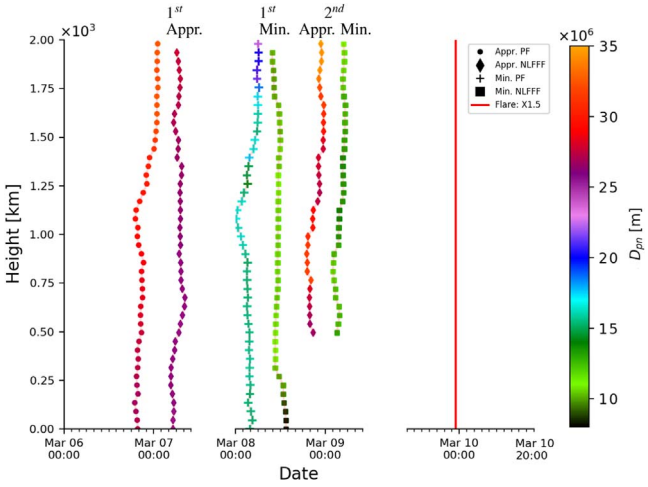

Figure B1. Same as Figure 3 but for AR 11166.

precursor patterns are identifiable prior to each of the two flares. We found that the two flare precursor behaviors of the X1.8 flare disappear from 1000 km upwards in both of the PF and NLFFF extrapolation modeling. For the X1.8 flare, the optimal height of the PF is found to be at 90 km and for the NLFFF it is also at 90 km, where, the value of $T^C_{\rm Imp}$ is 1.6/1.2 hr and the value of $T^M_{\rm Imp}$ is 2.1/1.5 hr in the PF/NLFFF case, respectively.

For the X2.1 flare, the optimal height is found to be 1035 km for PF magnetic field structures of AR 11283. The converging phase began 6.7 hr beforehand and finished 7.3 hr earlier at the optimum height 1035 km when compared to the result of analysis applied to the data in the photosphere. The 1035 km was chosen for the PF as the optimum height because the beginning and finishing moments of the converging phase started to shift continuously above this height until 3000 km (see Figure B2(b)). Extrapolation was carried out only up to 3000 km. In the NLFFF case, the optimum height is found to be at 225 km, where, the value of $T^C_{\rm Imp}$ is 3.8 hr and the value of $T^M_{\rm Imp}$ is 1.8 hr, respectively.

In summary, the overall situation with the estimates is similar to that of the AR 11166 in the PF case. Here, in both

Table B1
Same as Table A1 but for Several Active Regions and by Using PF Extrapolations Only

		PF					
NOAA AR	Flare Intensity	$T^{C}_{\rm Imp}$ (h)	$T^{M}_{\rm Imp}$ (h)	Opt. Height (km)	$S_{\rm flare}$	$T_{\rm est}$ (h)	T_{D+F} (h)
11430	X1.3	3.4	5.8	1845	<M5	13.4	8.6
11515	X1.1	8.7	4.3	585	X	89.8	35.3
11520	X1.4	16.5	1.1	360	X	65.8	67.7
11890	X1.1	2.4	1.2	180	X	52.5	17.5
	X1.1	2.5	1.9	585	>M5	27.6	21.4
11944	X1.2	4.7	2.2	225	X	25.7	43.4
12017	X1.0	7.5	1.9	1080	>M5	24.9	32.3
12158	X1.6	2.6	5.5	225	<M5	42.8	33.9
12192	X2.0	3.3	8.1	1530	X	41.6	33.4
12297	X2.1	12.7	3.8	1305	X	47.3	37.5
12673	X2.2/X9.3	6.2	1.7	1080	X	21.4	20.6
	X1.3	0.3	1.3	270	X	17.8	25.2

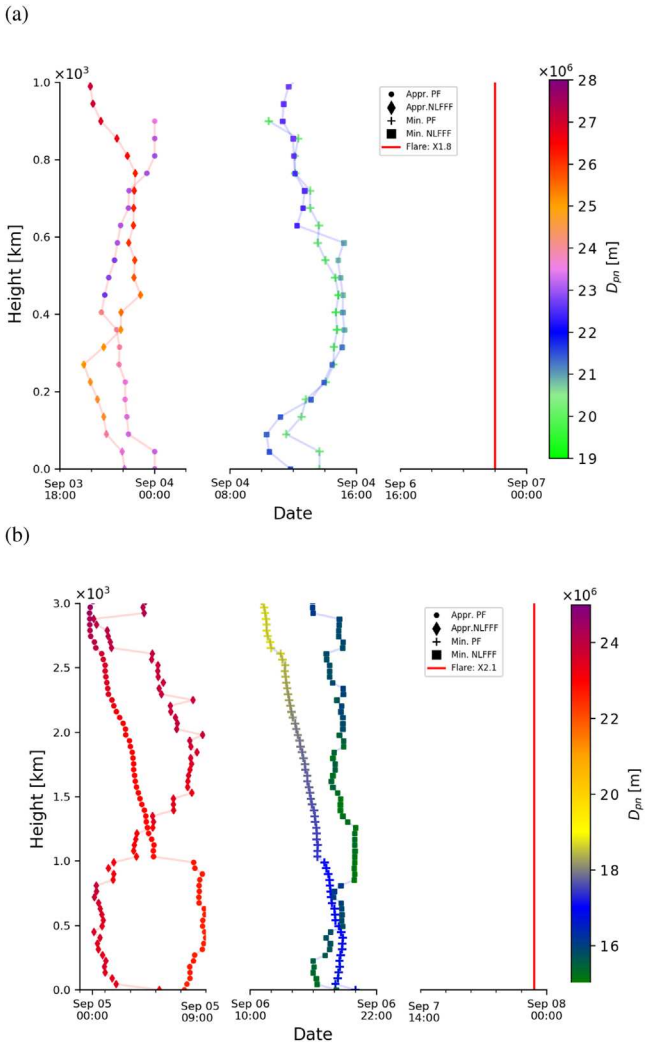

Figure B2. Plots (a) and (b) correspond to two *X*-class flares (X1.8 and X2.1 flares) of *AR 11283*.

extrapolation models, the $T_{\rm est}$ values are underestimated for the X1.8 flare, while $T_{\rm est}$ is overestimated for the X2.1 case, just like in the PF case for AR 11166. $S_{\rm flare}$ are fairly well estimated. In the case of AR 11283, we could estimate the onset time of the X1.8 flare 5.5 hr earlier with the PF data when compared to the counterpart obtained with the NLFFF extrapolation. For predicting the X2.1 flare, the PF and NLFFF extrapolations seem to be similarly beneficial.

B.3. AR 12192

The last randomly selected example to demonstrate the prediction capability of the $\rm WG_M$ method in a 3D lower solar atmosphere model is *AR 12192*, with a series of flares that occurred in the same δ-spot, according to Bamba et al. (2017). The first pair of characteristic preflare behaviors of the $\rm WG_M$ and $D_{\rm pn}$ are observed prior to the X1.6 flare, which occurred at 14:28 on 2014 October 22. However, the U-shape starts to form when the AR is on $\sim$67°, where the magnetic projection effects are not neglectable. Therefore, here, we did not investigate the case of the X1.6 flare.

Later, another pair of inverted V- and U-shapes are found and evaluated prior to the X3.1 (at 21:41 on 2014 October 24) and X1.0 (at 18:08 on 2014 October 25) flares, respectively. In the cases of the X3.1 and X1.0 flares, the optimal height of the PF approach is at 1260 km, and it is at 90 km for the NLFFF extrapolation (see Figure B3). The lead times are very similar in the case of NLFFF ($T^{M}_{\rm Imp} = 2.3$ hr) and PF approach ($T^{M}_{\rm Imp} = 2.1$ hr). Furthermore, in the case of NLFFF, the $T_{\rm est}$ values is well estimated for the first flare occurrence. The difference between the estimated and the actual occurrence times are close to the $\pm$7.2 hr uncertainty. We cannot say this about the case of PF extrapolation. Here, we also must mention that *AR 12192* produced a further X2.0 flare at 10:56 2014 October 26 (see Table B1). However, we can only observe the two typical preflare patterns when using the PF data.

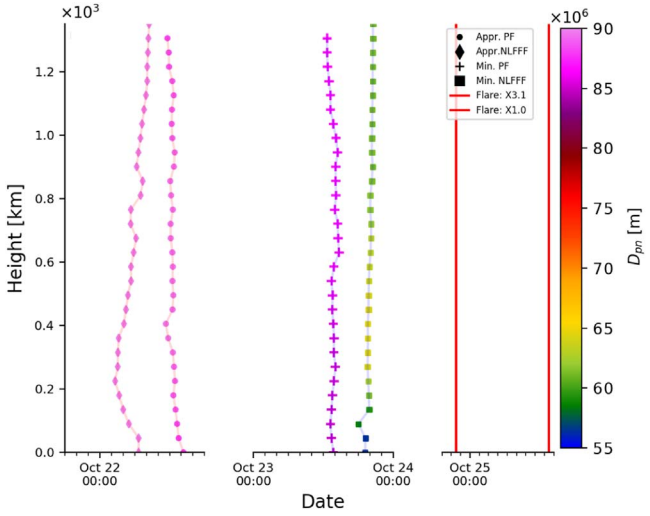

Figure B3. Plot corresponds to two *X*-class flares (X3.1 and X1.0 flares) of *AR 12192*.

ORCID iDs

M. B. Korsós https://orcid.org/0000-0002-0049-4798
M. K. Georgoulis https://orcid.org/0000-0001-6913-1330
N. Gyenge https://orcid.org/0000-0003-0464-1537
S. K. Bisoi https://orcid.org/0000-0002-9448-1794
S. Yu https://orcid.org/0000-0003-2872-2614
S. Poedts https://orcid.org/0000-0002-1743-0651
C. J. Nelson https://orcid.org/0000-0003-1400-8356
J. Liu https://orcid.org/0000-0003-2569-1840
Y. Yan https://orcid.org/0000-0002-7106-6029
R. Erdélyi https://orcid.org/0000-0003-3439-4127

References

Bamba, Y., Lee, K.-S., Imada, S., & Kusano, K. 2017, ApJ, 840, 116
Barnes, G., Leka, K. D., Schrijver, C. J., et al. 2016, ApJ, 829, 89
Bobra, M. G., & Couvidat, S. 2015, ApJ, 798, 135
Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, SoPh, 289, 3549
Campi, C., Benvenuto, F., Massone, A. M., et al. 2019, ApJ, 883, 150
Cui, Y., Li, R., Zhang, L., He, Y., & Wang, H. 2006, SoPh, 237, 45
DeForest, C. E., Hagenaar, H. J., Lamb, D. A., Parnell, C. E., & Welsch, B. T. 2007, ApJ, 666, 576
Eastwood, J. P., Biffis, E., Hapgood, M. A., et al. 2017, RiskA, 37, 206
Florios, K., Kontogiannis, I., Park, S.-H., et al. 2018, SoPh, 293, 28
Gary, G. A. 1989, ApJS, 69, 323
Georgoulis, M. K., Nindos, A., & Zhang, H. 2019, RSPTA, 377, 20180094
Kim, T., Park, E., Lee, H., et al. 2019, NatAs, 3, 397
Korsós, M. B., Chatterjee, P., & Erdélyi, R. 2018a, ApJ, 857, 103
Korsós, M. B., Poedts, S., Gyenge, N., et al. 2018b, in IAU Proc. 335, Space Weather of the Heliosphere: Processes and Forecasts, ed. C. Foullon & O. E. Malandraki (Cambridge: Cambridge Univ. Press), 294
Korsós, M. B., Yang, S., & Erdélyi, R. 2019, JSWSC, 9, A6
Leka, K. D., Park, S. H., Kusano, K., et al. 2019, ApJ, 243, 36
Liu, C., Deng, N., Lee, J., et al. 2014, ApJ, 795, 128
Lu, Z., Cao, W., Jin, G., et al. 2019, ApJ, 876, 133
Qahwaji, R., & Colak, T. 2007, SoPh, 241, 195
Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, SoPh, 275, 207
Schwenn, R. 2006, LRSP, 3, 2
Temmer, M., & Nitta, N. V. 2015, SoPh, 290, 919
Tlatov, A. G., Abramov-Maximov, V. E., Borovik, V. N., & Opeikina, L. V. 2018, Ge&Ae, 58, 1087
Toriumi, S., & Wang, H. 2019, LRSP, 16, 3
Vemareddy, P., & Wiegelmann, T. 2014, ApJ, 792, 40
Wang, R., Yan, Y., & Tan, B. 2013, SoPh, 288, 507
Wang, S., Liu, C., Liu, R., et al. 2012, ApJ, 745, L17
Wang, Y., Liu, J., Jiang, J., & Erdélyi, R. 2019, ApJ, 881, 15
Welsch, B. T., & Longcope, D. W. 2003, ApJ, 588, 620
Wiegelmann, T., & Sakurai, T. 2012, LRSP, 9, 5
Yan, Y., & Li, Z. 2006, ApJ, 638, 1162
Yan, Y., & Sakurai, T. 2000, SoPh, 195, 89
Yashiro, S., Gopalswamy, N., Akiyama, S., Michalek, G., & Howard, R. A. 2005, JGRA, 110, A12S05