Three-Dimensional Solar Flare Forecasting

a PhD thesis,

prepared under a cotutelle agreement between

University of Sheffield

and

University of Ioannina

authored by

Shreeyesh Biswal

June 27, 2026

Supervisory & Advisory Panel

The project "Three-Dimensional Solar Flare Forecasting" was jointly administered by (i) the University of Sheffield - serving as the primary host institution in the United Kingdom and (ii) the University of Ioannina - serving as the secondary host institution in Greece.

Prof. Robertus von Fay-Siebenbürgen of the School of Mathematical and Physical Sciences, University of Sheffield served as the project supervisor. The thesis advisory panel consisted of:

1. Prof. Michael Ruderman, University of Sheffield (01/11/2021 - 14/04/2025)
2. Dr. Sam Dolan, University of Sheffield (15/04/2025 - 13/02/2026)

The 3-member supervisory panel at the Department of Physics, University of Ioannina, consisted of:

1. Prof. Alexander Nindos, University of Ioannina
2. Dr. Manolis Georgoulis, Johns Hopkins University Applied Physics Laboratory
3. Prof. Robertus von Fay-Siebenbürgen, University of Sheffield

Examination Committee

The first draft of this thesis was jointly submitted to the University of Sheffield and the University of Ioannina on February 13, 2026. This thesis was successfully defended at the University of Sheffield on April 01, 2026 and at the University of Ioannina on June 03, 2026. The current draft of the thesis has taken into account, the minor corrections suggested by examination committees for both the Universities.

The 3-member examination committee at the School of Mathematical and Physical Sciences, University of Sheffield, consisted of:

1. Prof. Vassilios Archontis, University of Ioannina
2. Prof. Jasmina Magdalenic, KU Leuven
3. Dr. Ashley Willis, University of Sheffield

The 7-member examination committee at the Department of Physics, University of Ioannina, consisted of:

1. Prof. Alexander Nindos, University of Ioannina
2. Dr. Manolis Georgoulis, Johns Hopkins University Applied Physics Laboratory
3. Prof. Robertus von Fay-Siebenbürgen, University of Sheffield
4. Prof. Spyridon Patsourakos, University of Ioannina
5. Prof. Vassilios Archontis, University of Ioannina
6. Prof. Constantinos Gontikakis, University of Ioannina
7. Prof. Nikolaos Bakas, University of Ioannina

Declaration

All sentences or passages quoted in this document from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure.

Name: Shreeyesh Biswal

Signature: 

Date: 13/02/2026

Acknowledgments

First and foremost, I would like to thank all of my supervisors and mentors (*gurus*) - Prof. Robertus, Dr. Korsos, Prof. Nindos, Prof Spiros and Dr. Manolis who have guided me over the last 4.5 years.

At the outset of this project, when I arrived in Sheffield in January 2022, I felt a degree of apprehension about the journey ahead, having heard accounts of how strained supervisor–student relationships can pose significant challenges to research and the careers of young students. However, my first in-person meeting with Prof. Robertus dispelled all of my fears. Despite a daily schedule weighed down by an extraordinary workload, he manages to send emails at 4:00 A.M., sustains himself on muesli bars (I think), and quite remarkably replies to my emails within five minutes on most occasions. Prof. Robertus's office in H16 is a myriad of fascinating collections where time and thought seem to fold into one another. Venerable scientific textbooks from another era gaze down upon students like silent sentinels, clocks run backwards and tick in defiance of time, and tender saplings grow side by side in quiet pursuit of a promised "Garden of Eden". If Prof. Robertus taught me how to swim in lake Balaton, Dr. Korsós taught me how to run a 5k on early mornings, during the time I spent with them in Hungary. She has been my go-to person before every poster and conference presentations. Always cheerful, Dr. Korsós continues to amaze me through her creative methods of visualising data. I thank her for her guidance, mentorship and unwavering support throughout the PhD project.

While Sheffield initially appeared as a quiet contrast to the bustle of London, over time it revealed itself through its valley views, and abundant greenery woven through its streets, each season adding a new layer of beauty. The city, truly a jewel in the heart of Britain, became closely tied to everyday moments, from quiet walks framed by

sweeping landscapes to pauses at sunset points overlooking the skyline and surrounding hills. A deep appreciation remains for these views, especially J. M. W. Turner's Viewpoint in Meersbrook Park, a place I frequently visited.

In English, the idiom "It's all Greek to me" suggests that something is entirely incomprehensible. Yet, thanks to my three experienced Greek supervisors, no concept in solar physics remained beyond my fundamental understanding. Prof. Nindos embodies the calm focus of a monk. Composed and deliberate, he moves thoughtfully from one step to the next, noticing every detail in my writing. I extend my sincere appreciation to him for his rigorous standards, incisive insight and for helping me refine my scientific approach. Prof. Spiros is a naturally cheerful presence, embracing optimism even when the rest of the team is skeptical or contemplating a certain topic. On occasions when I fail to make a good presentation in our fortnightly meetings, Prof. Spiros is often the first to grasp the key details. I am deeply grateful to him for his constant encouragement, approachability, and generous support throughout this journey. Last, but not the least, if there was one word to describe Dr. Manolis, it would probably be precision. I am deeply appreciative of Dr. Manolis's perfectionism, which ensured that every aspect of scientific deliverables and publications met the highest standards. His insistence on rigour has left a lasting impact on my approach to scientific writing and I am indebted to him for his guidance.

I am sincerely grateful to Prof. Vassilios Archontis and Prof. Jasmina Magdalenic for their valuable feedback and suggestions, which greatly contributed to the improvement of this thesis. I would also like to express my appreciation to Dr. Ashley Willis for coordinating and facilitating the external review process.

I would like to thank my SWATNET (Space Weather Awareness Training and Network) friends Augustin (a.k.a. *Agastya*), Shifana, Guilherme (a.k.a. *Gilly*), Mayank, Andreas, Lidiya, Edin, Ronish, Gregoire, Simone (a.k.a. *Suman*) and Slava for the

truly wonderful times we spent together at summer schools, conferences and clubs. Among these friends of the SWATNET family, Simone, Augustin, Ronish, Slava and Gilly were my "Sheffield-cousins". I learnt a lot about fitness and training-discipline from Simone, who was truly like a brother to me. We have walked countless miles together, be it in the streets of Manchester or Rome. Augustin - "the friendly giant", was my partner in crime during my stay in Ioannina and Budapest. Ronish, apart from being a solar physics and an astrophotography enthusiast, shared a common hobby with me - numismatics. Andreas, the master of the Rubik's cube, was also my go-to dining partner and is quite a connoisseur when it comes to Indian cuisine. During the SWATNET conference dinners, I truly enjoyed discussing geopolitics with brother Edin. Slava and Gregoire, my French SWATNET friends were always up for an additional glass of beer. Lastly, the presence of my country-folks Mayank, Shifana and Lidiya never made me miss India when I was at SWATNET gatherings. In fact, throughout my time in Sheffield, my family and my close friends - Utpal and Debi, were always in touch with me over phone and so I never really missed being at home.

This thesis is dedicated to my father, Prof. Bhabendra Narayan Biswal who instilled scientific thinking in me and motivated me to pursue a research focused education at IISER Pune for my undergraduate studies, where I was actively supported at every step by my supervisors and mentors - Prof. Shyam Rai, Prof. Utsav Mannu, Prof. Prasad Subramanian and Prof. Dibyendu Chakrabarty. Further, I would like to express my gratitude to my three pillars of support - my mother Mrs. Sasmita Biswal, my sister Esha and my wife Disha.

This research project 'Three-Dimensional Solar Flare Forecasting' is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955620. I am thankful to the University of Sheffield for supporting my research through the Publication Scholarship Programme.

Abstract

Solar flares are intense bursts of electromagnetic radiation spanning across the electromagnetic spectrum and are key drivers of space weather phenomena. They are closely linked to other solar energetic phenomena such as coronal mass ejections and energetic particle events which have the potential to cause damage to both satellite-based or ground-based critical electronic infrastructure. Hence prediction or forecasting of such events is of key value to space weather research. Traditional techniques related to solar flare forecasting have focused on studying these parameters on the photosphere, where the magnetic field measurements are routinely available. However, recent studies on magnetic helicity and proxy parameters of the horizontal gradient of the vertical magnetic field, have shown that extrapolating the vertical magnetic field data obtained at the photosphere to higher altitudes could potentially improve flare prediction time. In this connection, we decide to extend such an approach to (i) polarity inversion lines - lines that demarcate regions of oppositely directed vertical field regions on the solar surface and (ii) R-value, a filtered version of magnetic flux, computed in close vicinity of polarity inversion lines. In this thesis, these features are studied based on their magnitudes and temporal evolution, spanning across multiple heights in the low solar atmosphere (up to $\sim$ 3 Mm).

Project Outcomes

Peer-reviewed publications:

1. 'Case Studies on Pre-eruptive X-class Flares using R-value in the Lower Solar Atmosphere'; S. Biswal, M. B. Korsós, M. K. Georgoulis, A. Nindos, S. Patsourakos, and R. Erdélyi; The Astrophysical Journal, (2024); ApJ 974 259; DOI: https://doi.org/10.1051/swsc/2024004

2. The Space Weather Awareness Training Network; T. Barata, D. Del Moro, R. Erdélyi, J. Fernandes, R. Gafeira, M. K. Georgoulis, K. Murawski, A. Nindos, O. Oliveira, S. Patsourakos, K. Petrovay, S. Poedts, R. Vainio, A. Afanasiev, M. B. Korsós, A. Boiko, L. Annie John, S. Biswal, S. Bourgeois, S. Chierichini, G. Francisco, A. André-Hoffmann, E. Husidic, S. Koya, M. Kumar, R. Mugatwala, G. A. L. Nogueira, A. Wagner, E. Kilpua, Journal of Space Weather and Space Climate (accepted)

3. 'Distinguishing between flaring and non-flaring solar active regions with height-dependent polarity inversion line lengths'; S. Biswal, M. B. Korsós, M. K. Georgoulis, A. Nindos, S. Patsourakos, and R. Erdélyi (accepted to ApJ, subject to minor corrections as of June 24 2026)

Open-source repositories:

This thesis adheres to the FAIR (Findable, Accessible, Interoperable, and Reusable) principles for scientific data management and sharing. All relevant codes, datasets, and plots have been made publicly available through the following GitHub repositories to promote transparency, reproducibility, and traceability of any inadvertent errors.

1. https://github.com/shreeyesh-biswal/Rvalue_3D

2. https://github.com/shreeyesh-biswal/PILs_3D

Contents

Chapter 1

Prologue

1.1 Historical Preface

The Sun, our closest star, has captivated the interest of humankind since time immemorial. Early human civilisations worshipped the Sun through dedicated deities. For example, the ancient Greek civilisation considered *Helios* to be the Sun God. Similarly, *Ra* was the Egyptian deity for the Sun in ancient times. Interestingly, the emblem on the back of the golden chair of the pharaoh Tutankhamun is reminiscent of the shape of the solar corona (Priest 2014)! Even today, people belonging to the Hindu faith worship *Surya* (also known as *Aditya*), the deity for the Sun. The Sun temple in Konark, India, designated as a UNESCO heritage site, serves as a testament to such rich tradition. To say that the interest of ancient human civilisations in the Sun was solely restricted to cultural beliefs, myths and legends would be an understatement. In fact, solar phenomena like the solar eclipse had also captivated the scientific interest of ancient civilisations. For example, there are records of astronomical observations of a solar eclipse in ancient China (around 2000 BC). The Sun also inspired technological inventions in ancient civilisations. Historically, humans have observed the motion of the Sun through the sky for keeping track of time and in this regard, the earliest known device for tracking time is the 'sundial', dating as far back in time as 1500 BC. It was

first used by the ancient Egyptians who would trace the motion of the shadow cast onto the dial to measure time. The dial used to have 12 sections for daytime and the motion of the shadow from one section to another would constitute an 'hour' as we understand in the modern context. The night-time was also considered to be equivalent to one cycle of motion of the shadow through the 12 sections. Thus, a day came to be known as consisting of 24 hours.

Ancient civilizations also made an effort to understand features on the surface of the Sun. This statement is supported by the records of sunspot observation made with the naked eye in China (around 960 – 1279) in the chronicles of the Sòng dynasty (Hayakawa et al. 2015). However, it was not until the early 1600s, that the first 'telescopic' observations of the Sun were made in the middle ages. During this period, sunspot observations in Europe were pioneered by Thomas Harriot, Galileo Galilei (see samples in Figure 1.1), Christoph Scheiner and Johannes Fabricius.

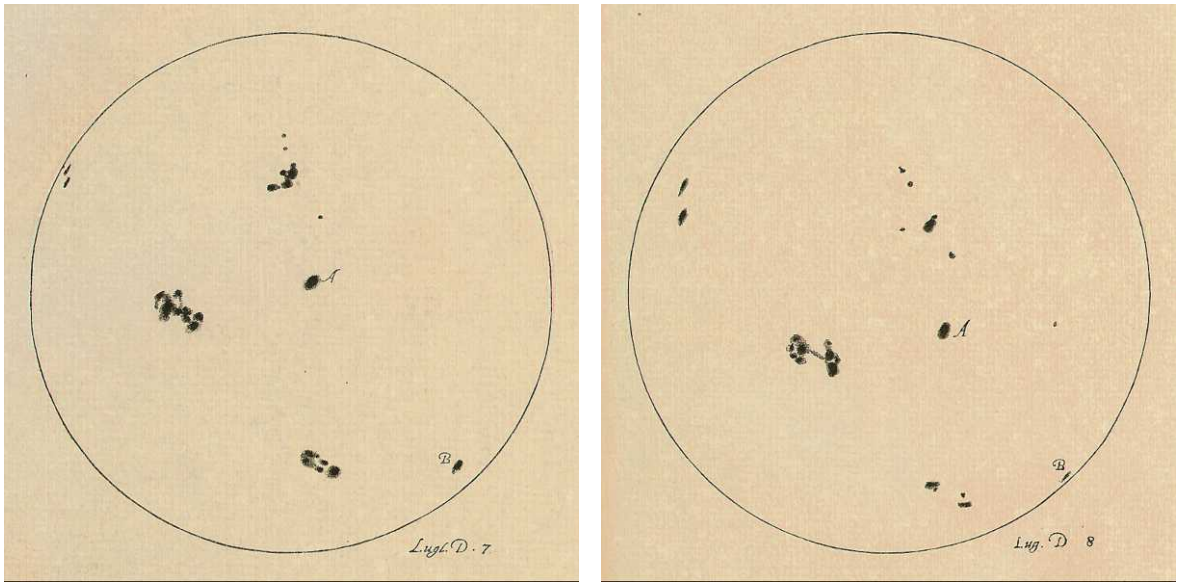

Figure 1.1: (Hand-drawn pictures of sunspots by Galileo on Jul 7, 1613 (left) and Jul 8, 1613 (right), courtesy: https://galileo.library.rice.edu/images/things/sunspot_drawings)

The observations of sunspots have continued steadily since then until now, with the exception of the time periods accounting for the 'Maunder Minimum' (1645 - 1715) and to some extent, the 'Dalton Minimum' (1800 - 1824). While the underlying causes of low solar activity in connection to the decreased frequency of sunspot occurrence during the Maunder and Dalton minima are not clearly understood, it is worth noting that Europe experienced lower than average temperatures during these periods, something that may have been caused due to low levels of solar activity. But since the cooling period began before the Maunder Minimum, climate models suggest that high levels of volcanic activity may have contributed to the cooling period (Owens et al. 2017).

Continuous study of sunspots for over two centuries yielded an interesting discovery when Schwabe (1843) concluded that the sunspot frequency has an 11-year cycle. Today, we know more precisely that the solar activity (which can not only be judged from the count of sunspots, but also from proxies like the magnetic field, the total solar irradiance, 10.7 cm radio flux, galactic cosmic ray flux and radioisotope studies inside tree rings and ice-cores) undergoes a decadal cycle with a mean period of 11 years and a standard deviation of 14 months (Hathaway 2010).

1.2 Life of the Sun

Out of the several hypotheses that explain the formation and evolution of the solar system, the most widely accepted hypothesis is the modern nebular hypothesis (Woolfson 1993). As per the nebular hypothesis, the formation of our solar system was triggered by the gravitational collapse of a molecular cloud, called the solar nebula. Modern research shows that stars are not born in isolation, but in clusters, although the Sun is thought to have formed in a relatively isolated or weakly interacting environment compared to typical clustered star formation (Pudritz 2002; Clark & Bonnell 2005). Giant molecular clouds over millions of years, undergo fragmentation and these smaller fragments of gas clouds collapse into stars (Pudritz 2002; Clark & Bonnell 2005). In any stellar nebula, if the contracting gravitational forces outweigh the expansive internal pressure forces (gas, thermal and magnetic pressures), the cloud contracts and forms a stellar configuration in its nascent stages, called a protostar. During the process of collapse, a bit of gravitational shearing imparts some amount of rotation. As the cloud contracts, the cloud gains further spin so as to conserve angular momentum. After some time in the stellar evolution timeline, the protostar is surrounded by a flat spinning disk called a protoplanetary disk (Pringle 1981), and this disk in turn slows down the rotation of the protostar. In case of the Sun, the collapse of the solar nebula occurred about 4.6 billion years ago and it was followed by the formation of the protostar in about 100,000 years (Montmerle et al. 2006). Most of the nebula's mass contracted towards the center to produce the Sun and thus, the Sun today accounts for about 99.9% of our solar system's mass. The formation of the protostar was succeeded by the formation of outer planets within 10 million years (Lineweaver 2001) and subsequently, the formation of inner planets within a time-frame of 100 million years (Gomes et al. 2005).

On the Hertzsprung-Russell diagram (Hertzsprung 1905; Russell 1919; Nielsen 1964), which is a scatter plot between stellar luminosity and stellar temperature for stars, the Sun falls within the cluster of main-sequence stars (see Figure 1.2). Main

sequence stars have a mass ranging between $0.08M_S$ - $120M_S$ (LeBlanc 2010) and produce helium by the nuclear fusion of deuterium. While upper main-sequence stars of mass greater than $1.5M_S$, produce helium via CNO cycle, lower main-sequence stars of mass less than $1.5M_S$, produce helium via p-p chain process (LeBlanc 2010). The energy released in the process of nuclear fusion causes the main sequence stars to remain in equilibrium by overcoming the inward gravitational forces. At some point when the hydrogen in the core finishes, the core continues to become hotter and contracts further. However, the hydrogen present in the layer surrounding the core, fuelled by the energy supplied from the hot core, continues to burn and causes the outer layer to expand. As the outer layer expands, it cools down after a certain point and the star appears red in colour, and thus the main sequence star transitions into a red-giant.

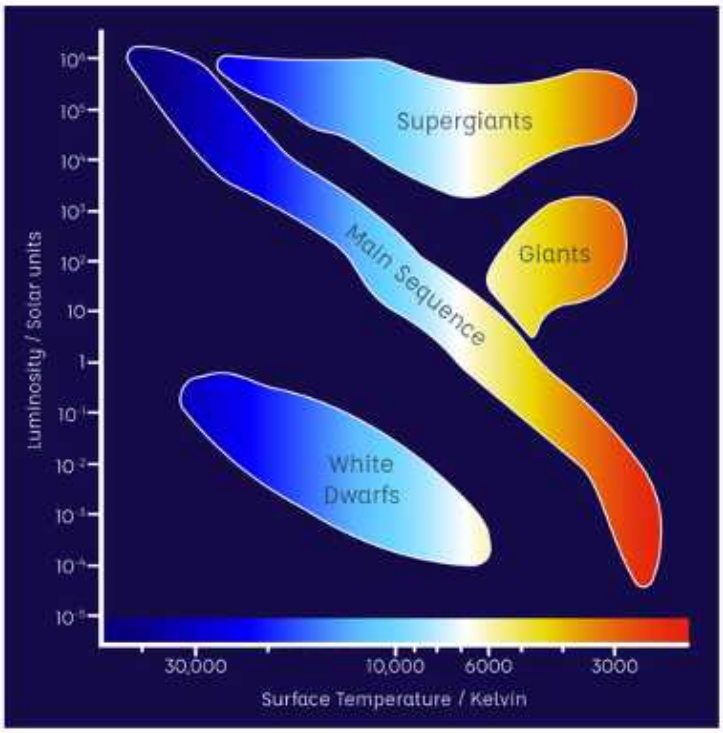

Figure 1.2: The Hertzsprung Russell diagram showing the different branches of stars on a luminosity-temperature map; courtesy: A. Hopkinson, https://lco.global/spacebook/stars/h-r-diagram/

Red-giant stars have an extremely hot core, hot enough to cause further fusion

of helium into heavier elements such as carbon and oxygen and this causes the core of the star to contract and become blue again until the helium content gets depleted (LeBlanc 2010). This process continues for about 100 million years and in the end, the core of heavier elements is essentially surrounded by thin shells of helium and hydrogen successively (LeBlanc 2010). While the core continues to collapse under gravity, helium burning in the outer shell supports the expansion of the star once again, and the star becomes a giant star again in about 10 million years, but only temporarily (LeBlanc 2010). The expanding outer layer eventually escapes into the interstellar space, leaving behind a dense core. This dense core, although hot in general terms, is not hot enough to sustain further thermonuclear fusion of heavier elements such as carbon and oxygen and thus remains as it is structurally and cools with time. This is the final evolutionary state of main-sequence stars which have masses similar to that of the Sun and the star is called a white dwarf at this stage. Massive stars however, follow an entirely different path that leads to a supernova explosion. The Sun is about 4.6 billion years old currently and has a composition of 73% hydrogen and 25% helium. It is expected to become a red-giant in the next 5 billion years and collapse into a white dwarf after that (LeBlanc 2010).

1.3 Space Weather

As mentioned in the previous section, the solar energy which we receive on the Earth today through radiation is effectively the outcome of p-p chain thermonuclear fusion of hydrogen into helium in the Sun's core. Since the Sun is extremely hot, practically all matter present is ionized and exists in plasma state. The motion of the charged plasma particles inside the Sun gives rise to magnetic fields, which serve as a conduit to transfer the energy from the Sun's core to its atmosphere, from where it is subsequently released into the heliosphere. Magnetic fields on the Sun are dynamic and evolve all the time. Sometimes when the magnetic field configuration in the solar atmosphere becomes unstable, it results in the occurrence of energetic processes, which are accompanied by a sudden release of energy. The most common energetic processes are solar flares and Coronal Mass Ejections (CMEs). While, a solar flare is observed as an intense brightening in the solar atmosphere on a scale of minutes, leading to a significant increase in the emission of radiation across the electromagnetic spectrum (Benz 2008), a CME is essentially an ejection of magnetic fields and plasma from the coronal region in the solar atmosphere into the heliosphere (Webb & Howard 2012). Sometimes eruptions on the Sun can lead to Solar Energetic Particle (SEP) events which refer to the transient injection of energetic particles, such as protons, electrons and heavier ions with kinetic energy ranging between several keVs to GeVs (corresponding to relativistic velocities), into the heliosphere (Whitman et al. 2023).

Just like the atmospheric conditions on the Earth at any given location defines the weather at that location, the term 'space weather' collectively refers to the highly dynamic conditions on the Sun's atmosphere, interplanetary medium and the Earth's geospace boundaries, more precisely within the magnetosphere (Baker 1998). Solar energetic phenomena such as flares, CMEs and SEP events are of technological interest because these phenomena are key drivers of space weather (Temmer 2021; Zhang et al. 2021; Georgoulis et al. 2024).

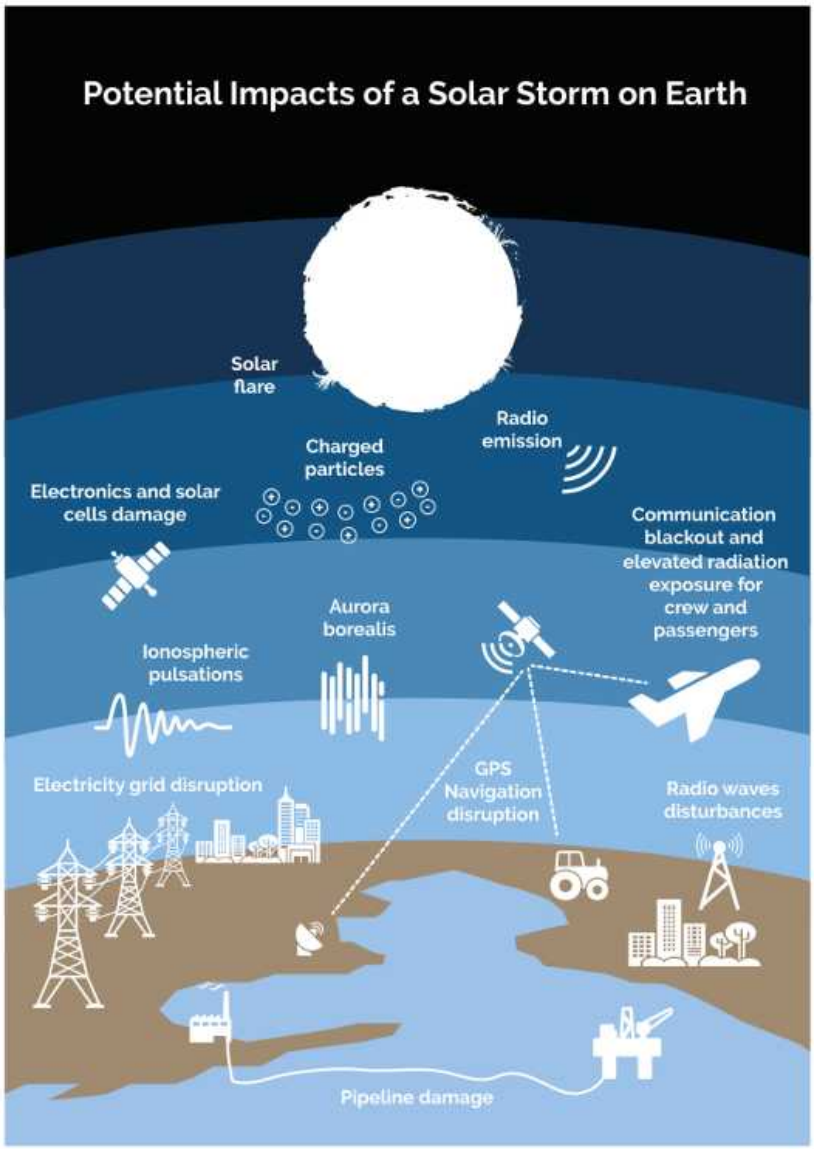

Figure 1.3: A cartoon depicting the influence of solar activity on the technosphere; adapted from Georgoulis et al. (2021); credit: University of Applied Sciences and Arts Northwestern Switzerland

Solar flare events are associated with a spike in the emitted x-ray and EUV (Extreme Ultra-Violet) fluxes, which leads to an increase in the ionization of the ionosphere, and as a consequence, radio systems, satellite communications systems and navigation services which rely on signals propagating through the ionosphere, may get disrupted (see Figure 1.3; Tsurutani et al. 2009; Ishii et al. 2024). Solar flare and CME events may

also trigger SEP events and pose a risk to the health of astronauts (Cucinotta 2015) and to the safety of satellite based electronic equipment in space (Iucci et al. 2005). The magnetic fields released into the interplanetary medium due to CME events may interact with the Earth's magnetic field and induce geomagnetic storms (see Section 5 in Gopalswamy 2022 and references therein). Geomagnetic storms, the effects of which were first observed by the German polymath Alexander von Humboldt in Berlin in 1806 when he saw disturbances in a compass needle at the time of an aurora (Lakhina & Tsurutani 2016), pose a huge threat to the safety of electrical infrastructure on the Earth because they induce huge currents in power transmission lines and electrical grids (Ngwira & Pulkkinen 2019).

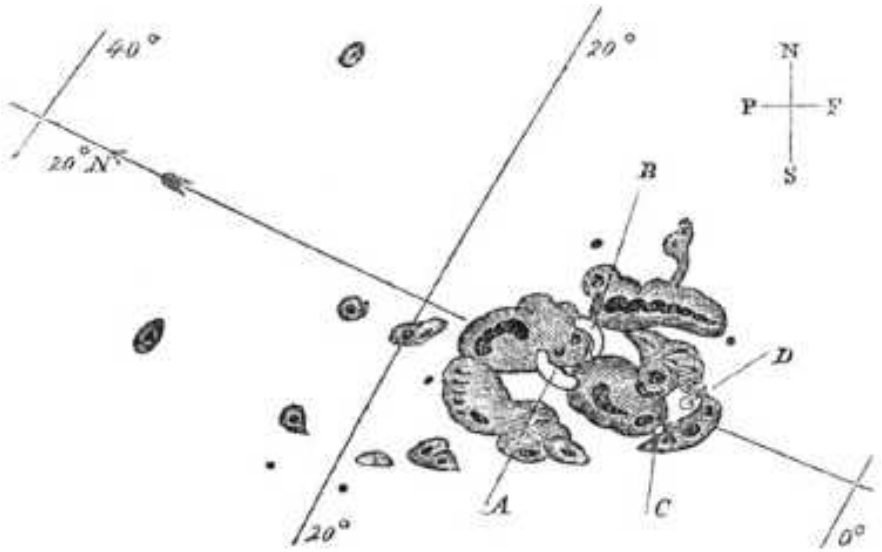

Figure 1.4: Carrington's sketch of the observed brightenings over the sunspot marked by A and B, courtesy: https://www.americanscientist.org/article/sunspotting

The most intense geomagnetic storm ever recorded in human history is the Carrington event of Sep 1-2, 1859. The geomagnetic storm caused fires at several telegraph stations worldwide and incidentally, is the first astronomical observation of a solar flare. Richard Carrington, a British astronomer based in Colaba, Mumbai at the time observed two bright patches of white light (Lakhina & Tsurutani 2016), alluding to the solar flares (see Figure 1.4) connected to the CME which induced the geomagnetic

storm. In 1989, the 9 hr Quebec power blackout arising out of a series of geomagnetic substorms (Boteler 2019) caused a total damage of 13.2 million CAD (Bolduc 2002). Several studies have attempted to estimate the economic damage arising out of extreme space weather events (see Eastwood et al. 2017 and references therein; Oughton et al. 2017). If a Carrington-level storm were to occur today (the probability of which occurring is 12% within a decade and 62% within a 50 year period (Cannon et al. 2013), a crude estimate is that it may inflict a damage of 0.6-2.6 trillion USD, in the United States (Lloyd's 2013). Oughton et al. (2017) estimate that the global damage due to an extreme magnetic storm could be between 7-42 billion USD per day.

To summarise, forecasting solar flares (and other solar energetic phenomena) is important because they can significantly affect modern technological infrastructure and human activities in space and on Earth. Accurate forecasting helps mitigate the impact of space weather on satellites, communication systems, navigation networks, aviation operations, and electrical power grids. It also supports astronaut safety during space missions by providing advance warning of enhanced radiation events. As humanity becomes more dependent on satellites and expands toward sustained lunar and interplanetary missions, reliable solar flare prediction becomes increasingly critical. Further it also helps in assessing the potential habitability of exoplanets orbiting magnetically active stars, as stellar flares, coronal mass ejections, and energetic particle events collectively drive atmospheric erosion and surface radiation environments that may critically constrain the emergence and sustainment of life (Airapetian et al. 2020; Samara et al. 2021). In addition to its practical applications, flare forecasting contributes to a better understanding of solar magnetic activity and energetic processes.

1.4 Thesis Outline

The damage estimates quoted in Section 1.3, arising out of solar energetic phenomena highlight the need for effective space weather monitoring and forecasting systems in order to secure critical ground-based electrical infrastructure and space-based satellite infrastructure. In this regard, the traditional approach of forecasting/predicting solar eruptive phenomena has primarily relied on tracing the evolution of photospheric magnetic precursors. This is primarily because flares and CMEs are magnetic phenomena, i.e. they are powered by stored budgets of free magnetic energy and the photosphere (see Section 2.1.1) remains the only place on the Sun where the morphology of solar magnetic fields can be studied routinely with the help of ground-based and space-based magnetographs, at present. However, studies in the past decade have established that flare prediction may be improved by studying the magnetic field configuration at a certain height range above the photosphere (depending on the flare-prediction parameter) called the OHR (Optimal Height Range) (Korsós et al. 2020; Korsós et al. 2022). This approach may be justified by the fact that the earliest signs of solar flares and CMEs are identified from the chromosphere and low solar corona and not exactly the photosphere. In this context, the (i) WG_M morphological parameter - a proxy for the horizontal gradient of the line-of-sight component of the magnetic field (Korsós et al. 2015, Korsós et al. 2018) and the (ii) emergence component of the magnetic helicity (Soós et al. 2022) have shown promise in improving flare prediction lead time by 2-8 hr (compared to the photosphere) when studied in height ranges of 1-1.8 Mm and 0.36-1.5 Mm above the photosphere, respectively (Korsós et al. 2020; Korsós et al. 2022).

Motivated by such promising results this thesis is primarily centered around the prediction of major solar flares and attempts to investigate if other critical parameters which are closely linked to WG_M morphological parameter and magnetic helicity, such as length of Polarity Inversion Lines (PILs) (see Section 2.4.4), R-value (see Section 2.4.5) and the unsigned magnetic flux around PILs (see Section 2.4.5) could also help

improve flare prediction when studied above the photosphere. Crucially, PILs - particularly strong, flux-massive and highly sheared ones - occupy a central role in the origin of solar eruptions, as eruptive active regions harbouring such PILs store non-potential magnetic energy and helicity far more efficiently than their non-eruptive counterparts, engaging in an irreversible path toward eruption once critical thresholds of free energy and helicity are crossed (Georgoulis et al. 2019). This thesis also aims to assess if these parameters could potentially help develop operational probabilistic flare prediction/forecasting models. In this context, these parameters (along with similar proxies to quantify them) are defined and described in detail in Chapter 2. Chapter 2 also contains a (i) brief description of the solar atmosphere, (ii) mechanism of occurrence of solar energetic phenomena and (iii) other mathematical concepts relevant to this thesis. The methodology and algorithms used to compute the flare prediction parameters of focus in this thesis are described in detail in Chapter 4. Detailed analysis of each parameter in the low solar atmosphere and further discussions regarding their relevance to flare prediction and forecasting are presented in Chapters 5, 6, 7. Finally, the major findings of this thesis and conclusive remarks are presented in Chapter 8.

While analysing different ARs over the course of this PhD project, sufficient time was dedicated to visualising solar magnetic field lines. This effort culminated with the preparation of a manual centered around plotting solar magnetic field lines using *ParaView*. This manual has been shared in Appendix A. Further, under the SWATNET program, ESRs (Early Stage Researchers) were mandatorily required to complete a three-month long industrial training internship. For project 3, i.e. 'three-dimensional solar flare forecasting', the industrial training internship was carried out in Hungary in close collaboration with mentors and collaborators from (i) Astrotech KFT, Baja, (ii) GSO (Gyula Solar Observatory), (iii) Eotvos Lorand University, Budapest, (iv) University of Durham and (v) University of Ioannina between Jul 15 - Oct 15, 2024. The technical report submitted at the conclusion of this internship has been shared in Appendix B.

Chapter 2

Introduction

2.1 The Solar Atmosphere

The Sun's magnetic field is believed to be generated in its interior through dynamo processes, arising from the interplay between differential rotation and convective plasma motions (Charbonneau 2010). This mechanism continuously regenerates and amplifies magnetic fields, which then emerge through the solar surface into the atmosphere. The solar magnetic field which serves as a link between the interior of the Sun and the solar atmosphere, also drives phenomena like flares and CMEs due to configurational instabilities (Wiegelmann et al. 2014). From a scientific standpoint - if developing predictive models to forecast solar flares is the objective, it is important to understand the physical mechanism of solar energetic phenomena. This, once again requires a preliminary understanding of the solar atmosphere, and thus, an overview of the different layers of the solar atmosphere - (i) the photosphere, (ii) the chromosphere, (iii) the transition region, (iv) the solar corona, along with the important magnetic field-driven features and physical phenomena found in these layers is provided in this chapter.

2.1.1 The Photosphere

The photosphere is the lowest visible layer in the solar atmosphere that stretches up to a few 100 km in altitudinal thickness. It is the layer of the Sun from where most of the visible light is emitted (Priest 2014). Upon observation, the photospheric surface appears granulated, the granules (∼1 Mm in size) being the top surfaces of the convection currents of hot plasma, rising from beneath the photosphere (Priest 2014). These currents of plasma play a key role in transporting energy. While granules (see Figure 2.1) are like small-scale convective cells, supergranules, which are also seen on the photosphere, are much larger convective cells that organise plasma motions and magnetic flux on scales of 20–70 Mm (Priest 2014). The plasma flow pattern in supergranules is largely along the horizontal direction. The emergence of magnetic flux through the photosphere can happen: (i) on a small spatial scale through tiny granular magnetic loops (flux as low as $\sim 10^{16}$ Mx), or (ii) on an intermediate scale through ephemeral regions (flux $\sim 10^{19}$ Mx), embedded in supergranular flows (Hagenaar 2001), or (iii) on a large scale (flux up to $\sim 10^{23}$ Mx) in active regions (ARs) (see details in Section 2.2; Priest 2014). ARs (i) span all four major layers of the solar atmosphere, (ii) are characterised by concentrated magnetic fields that are much more intense than the surrounding solar surface, and (iii) are the source of most solar flares, CMEs and solar radiation storms. (Priest 2014). When strong coherent magnetic flux associated with ARs emerges in the photosphere, it suppresses the convective motion of plasma at granular scales and forms darker and cooler regions (relative to their surroundings) called sunspots (Solanki 2003). In simpler words, sunspots are like 'photospheric footprints' of ARs, where the typical field strength is about 3000 G (Rezaei et al. 2012; Livingston & Watson 2015; Okamoto & Sakurai 2018), significantly higher compared to the ambient photospheric magnetic field which is about 10 G, ordinarily. The smallest sunspots are about 3.5 Mm in diameter (Bray & Loughhead 1964), while the largest ones have diameters exceeding 60 Mm (Priest 2014). The temperature in the photosphere ranges between 4000 - 6000 K and the upper layers are cooler than the bottom layers. The

photospheric plasma is weakly ionized, and while convective transport dominates in its deeper layers, radiative energy transport becomes dominant near the upper photospheric boundary where the plasma becomes optically thin (Vögler et al. 2005). The evolution of photospheric magnetic fields provides the boundary conditions for magnetic energy buildup in the chromosphere and corona, thus linking the photosphere to solar flares and coronal mass ejections (Forbes et al. 2006; Wiegelmann et al. 2014).

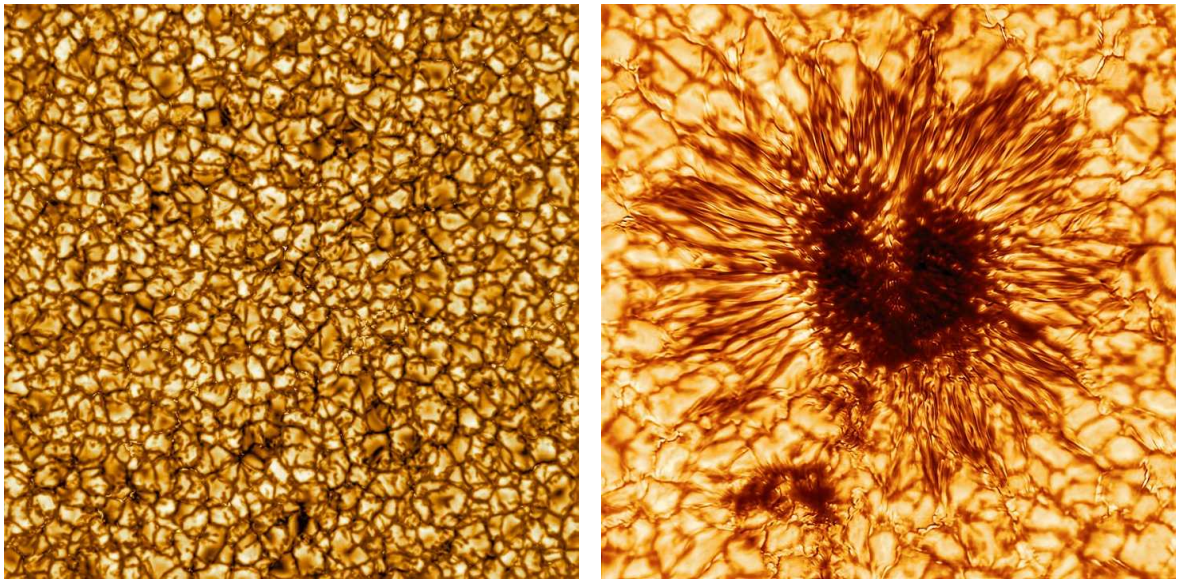

Figure 2.1: The solar surface as imaged by the Daniel K. Inouye Solar Telescope; granules (left); a sunspot (right); the umbra of the sunspots is usually tens of granules in size

2.1.2 The Chromosphere

The chromosphere is the middle layer of the solar atmosphere which is not normally visible because its weak line emission is overwhelmed by the intense continuum radiation of the photosphere. However, during a total solar eclipse, when the photospheric light is blocked by the Moon, the chromosphere becomes briefly visible as a thin coloured layer above the lunar limb (Carlsson et al. 2019). The emission corresponds to a wavelength of 656 nm (red colour) emitted by hydrogen - the Hα spectral line in the Balmer series (Carlsson et al. 2019). The chromosphere ranges from about 500 km to about 2500 km in altitude, relative to the base of the photosphere. In the chromosphere the temperature increases with height from 4000 K at the bottom to 20000 K at the top.

The chromosphere is where magnetically guided structures such as fibrils, mottles and spicules are observed. These are chromospheric plasma structures guided by magnetic fields, appearing differently because of the viewing angle and the magnetic environment. They play an important role in chromospheric heating, plasma transport and wave transport (De Pontieu et al. 2004; Zaqarashvili & Skhirtladze 2008; Zaqarashvili & Erdélyi 2009; van Ballegooijen et al. 2011).

Spicules (see example in Figure 2.2) are ejected from the boundaries of supergranules and are usually found in ARs and less commonly in quiet-Sun regions and coronal holes (see Section 2.1.4). They reach speeds of 10 to 50 kms^{-1} and about 5 Mm in height are persisting for about 3-10 min before falling back into the Sun (Beckers 1972; Sterling 2000; De Pontieu et al. 2007; Priest 2014).

Unlike spicules, which are primarily observed at the solar limb due to line-of-sight and visibility effects, fibrils and mottles are typically identified on the solar disc, although these structures are thought to represent similar magnetic field-guided phenomena viewed under different geometrical conditions (Hansteen et al. 2006; Rouppe

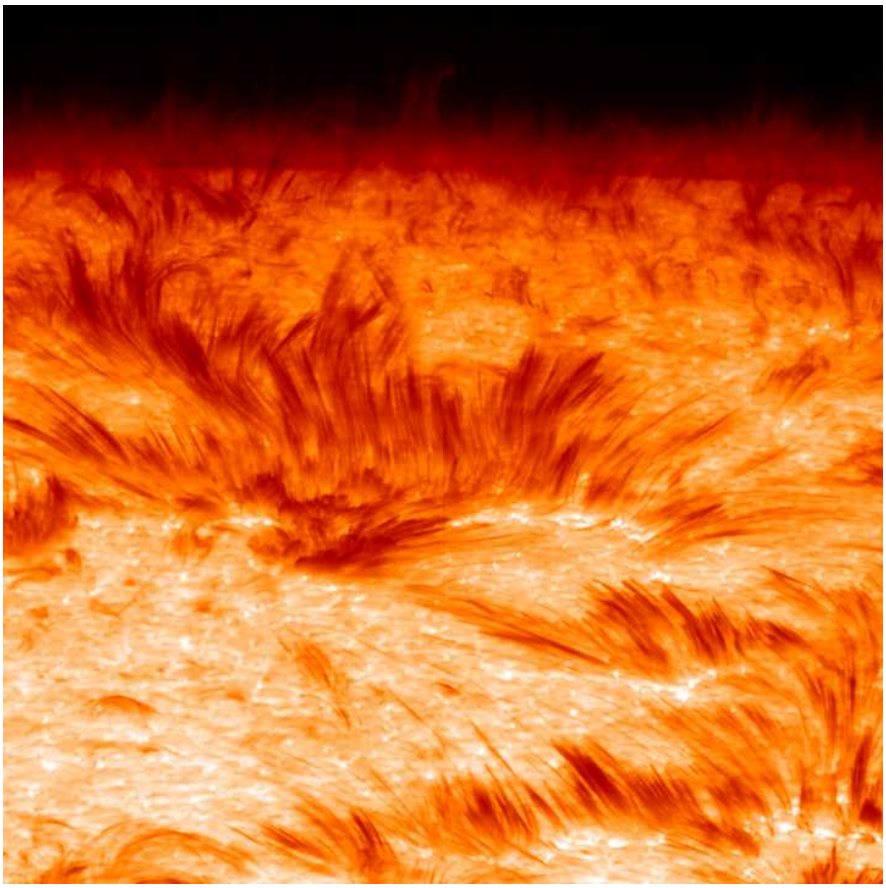

Figure 2.2: Image of spicules, arranged in the periphery of supergranular cells, observed by the Solar Optical Universal Polarimeter (SOUP) instrument on the Swedish Solar Telescope (SST) at La Palma

van der Voort et al. 2007; Kuridze et al. 2012; Rutten et al. 2019). Fibrils are usually found in ARs and may or may not exhibit a jet-like behaviour. The ones that don't exhibit a jet-like behaviour are static fibrils. These are heavily inclined loop-like structures that trace magnetic field lines connecting regions of opposite polarity (Jing et al. 2011; Leenaarts et al. 2012) and have lifespans on the scale of hours. Dynamic fibrils on the other hand, are shock driven jet-like features that are about 1-4 Mm in length and are shorter than their static counterparts. Mottles are like quiet-Sun counterparts of spicules, in general. They appear as dark, elongated features (2-10 Mm in length) in the Hα disc images, lasting between 2-15 min and supporting plasma motion of speeds 10-30 kms^{-1} (Suematsu et al. 1995).

Energetic phenomena in the chromosphere include jets, UV bursts and Ellerman bombs. While UV bursts are small-scale, short-lived brightenings observed in ultraviolet wavelengths in the upper chromosphere - typically linked to magnetic reconnection, Ellerman bombs (see Figure 2.3) are moderately energetic upper photospheric or chromospheric phenomena (energy dissipation $\sim 10^{27} - 10^{28}$ ergs), observed as transient, small-scale brightenings in the lower chromosphere (Georgoulis et al. 2002). Ellerman bombs are possibly explained by the stochastic magnetic reconnection (Nelson et al. 2013), arising out of turbulent evolution of the magnetic fields and are thus interpreted as signatures of magnetic reconnection in partially ionized plasma (Georgoulis et al. 2002). The temperature enhancement due to Ellerman bombs in the radiating volume is about 2000 K (Georgoulis et al. 2002).

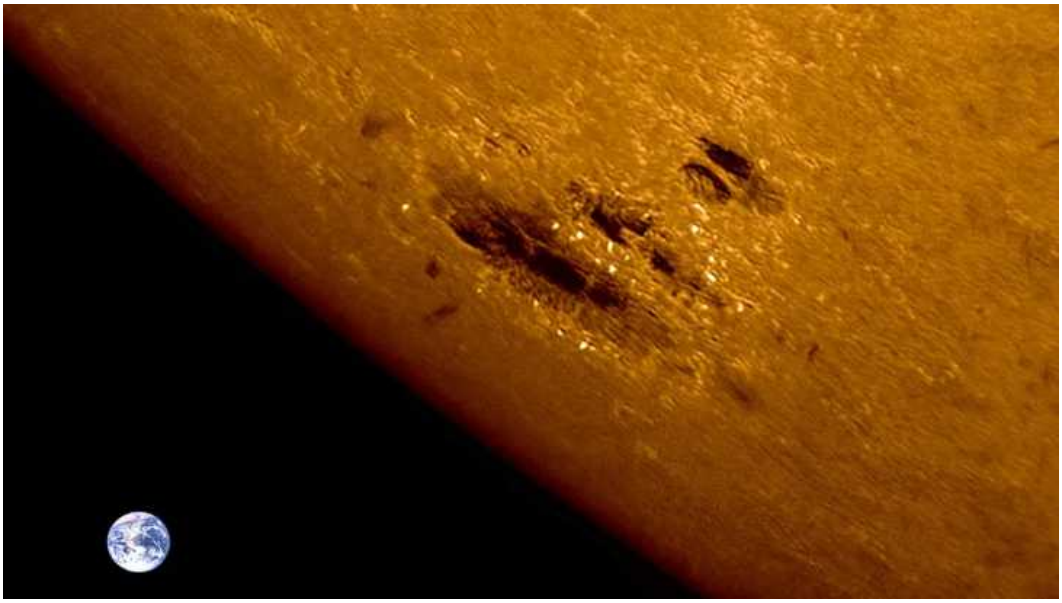

Figure 2.3: An image of Ellerman bombs (marked by bright spots) surrounding a sunspot (Image credit: Philippe Tosi)

2.1.3 The Transition Region

The thin layer (thickness $\sim$ a few hundred km) between the chromosphere and the solar corona is marked by an abrupt increase in temperature and an abrupt decrease of density. This layer is called the transition region (see Figure 2.4). The rapid rise in temperature is primarily due to downward thermal conduction from the excessively hot corona ($\sim 10^6$ K). Because thermal conductivity of a fully ionized plasma increases strongly with temperature, a large conductive heat flux flows downward and must be balanced by intense radiative losses in the lower atmosphere. This balance, together with the decreasing density required to maintain approximate pressure equilibrium, results in the very steep temperature gradient observed in the transition region (Priest 2014). The temperature increase and density decrease counterbalance each other to maintain a uniform spatial gas pressure (Lang 2000). In upper chromospheric heights the Lorentz force starts to dominate over the plasma pressure and gravity forces and above the transition region, the plasma is magnetically dominated, giving rise to different simplifications of MHD equations which are useful to model the solar atmosphere (Priest 2014; Wiegelmann et al. 2014; Wiegelmann & Sakurai 2021).

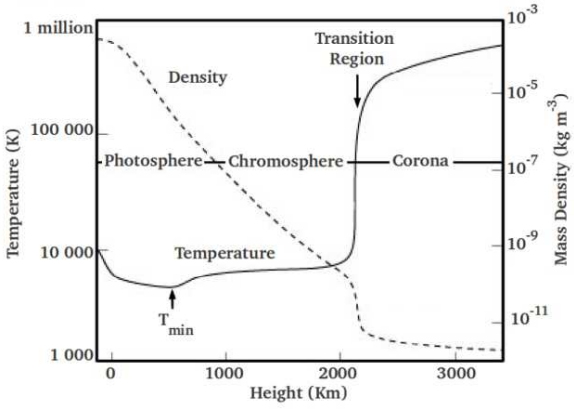

Figure 2.4: Variation of plasma density and temperature as a function of height above the photosphere (courtesy - Lang 2000); notice the sharp spike in temperature and the sharp dip in density at the transition region

2.1.4 The Solar Corona

The corona is the outermost layer of the solar atmosphere. It has no sharp lower boundary, but typically becomes dominant above the transition region at heights of a few thousand kilometres, depending on local magnetic and thermodynamic conditions. The corona is extremely hot with temperatures about 1-2$\times 10^6$ K (Edlén 1943; Priest 2014). Much like the chromosphere, the solar corona can only be seen during a total solar eclipse. The corona is studied in white light with the help of a coronagraph (Lyot 1932), which is designed to artificially replicate a total solar eclipse, allowing scientists to study the corona at any time, rather than only during naturally occurring total solar eclipses. The solar corona exhibits a wide range of structures and dynamical phenomena, many of which persist independently of major energetic events. In this section, attention is restricted to non-eruptive coronal features, while other major phenomena such as solar flares, coronal mass ejections, filaments, and prominences are discussed separately in Section 2.3. In this section, we mainly focus on three major types of coronal features: (i) quasi-static magnetic structures, (ii) small-scale heating and transient brightenings and (iii) waves oscillations and mass transport phenomena.

The morphology of the solar corona is shaped mainly by relatively stable magnetic structures (quasi static structures) that change gradually over periods ranging from hours to days. The most prominent of these features are coronal loops (see Figure 2.5). Coronal loops are closed magnetic field structures that are filled with plasma (temperature $\sim 10^6$K) and have lengths in the range 1 - 1000 Mm (Reale 2010; Priest 2014). Direct feature tracking observations and doppler shift studies show that coronal loops are associated with plasma upflows from the chromosphere, because of which it is thought that these loops originate below the chromosphere (Brekke et al. 1997; Peter & Judge 1999; Schrijver et al. 1999; Aschwanden 2001). In the quiet Sun, coronal loops are low-lying, short and faint, whereas they appear as dense, extended arcades in ARs when viewed in EUV and soft X-ray wavelengths (Webb & Zirin 1981; Reale 2010;

Priest 2014). Closed magnetic fields may form fan-like structures and helmet streamers (see Figure 2.6) that extend into the outer corona. The boundaries of helmet streamers produce slow solar wind (Priest 2014). The low plasma-β environment ensures that the observed coronal structures closely follow the magnetic field configuration, making such features valuable descriptors of the magnetic configuration in the corona (Priest 2014; Decraemer et al. 2019).

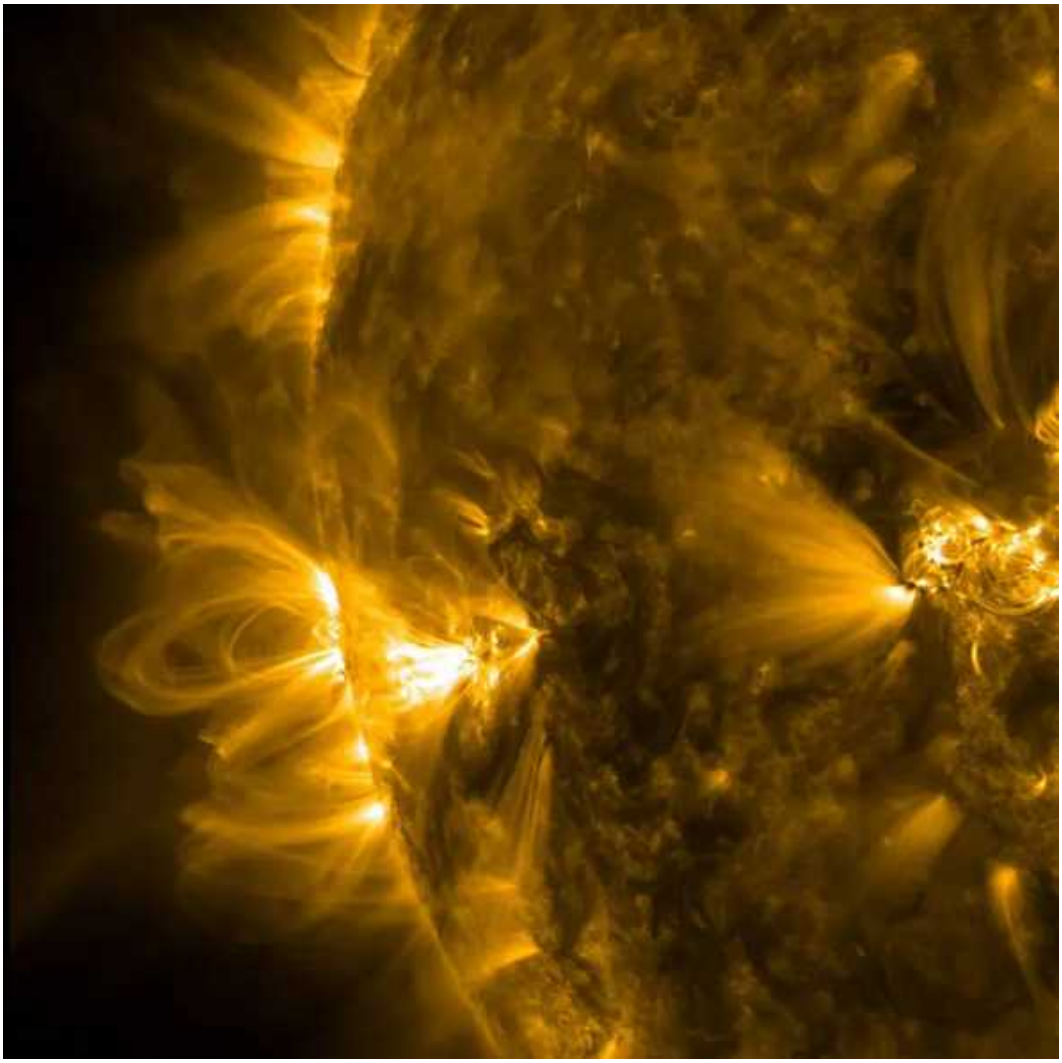

Figure 2.5: Image of a coronal loop as observed in the 17.1 nm channel by SDO/AIA in Feb 2014; courtesy: https://www.nasa.gov/

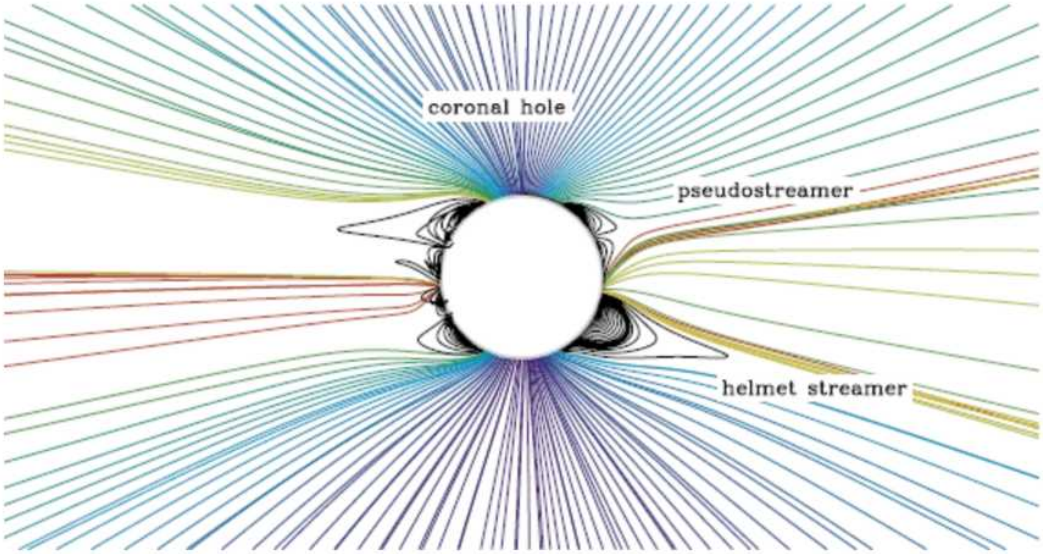

Figure 2.6: A diagrammatic magnetic field line sketch of different coronal features - coronal holes, streamers and pseudo-streamers; courtesy: Cranmer et al. (2017)

In contrast to closed-field systems, open-field regions of the corona are characterised by magnetic field lines that extend into interplanetary space. The most prominent of the open field quasi static magnetic structures are coronal holes (see Figure 2.6). Coronal holes appear as dark regions in EUV and X-ray images due to their lower density and temperature, relative to the ambient corona (Zirker 1977; Cranmer 2009) and are well established as the source of the fast solar wind (Neugebauer & Snyder 1966; Krieger et al. 1973 Cranmer 2009). Jets linked to coronal holes are magnetically driven collimated plasma ejections which may contribute to coronal heating and solar wind acceleration (Raouafi et al. 2016). Within coronal holes, plume and interplume structures are frequently observed, with plumes appearing as narrow, bright density enhancements aligned with open magnetic fields. Singular plumes usually last about a day (Priest 2014). Pseudostreamers (see Figure 2.6) represent large-scale open-field structures that separate coronal holes of the same magnetic polarity (unlike helmet streamers, which separate coronal holes of opposite polarity) and usually contribute to the slow/intermediate solar wind (Tokumaru & Fujiki 2024). These open-field structures play a central role in coupling the solar corona to the heliosphere and in controlling large-scale solar

wind properties (Tokumaru & Fujiki 2024).

At smaller spatial and temporal scales, localised heating events and transient brightenings are found in the solar corona. Such transient brightenings could be an outcome of nano/micro-flares (Parker 1988), wave dissipation signatures (De Moortel & Browning 2015) or jet-related phenomena (Shibata & Magara 2011). Coronal bright points are ubiquitous in the quiet Sun and coronal hole regions (Madjarska 2019) and are typically associated with small magnetic bipoles undergoing flux emergence or cancellation (Mou et al. 2018). Transient brightenings associated with nano/micro-flares (Parker 1988) represent impulsive energy release events that may occur continuously throughout the corona (Klimchuk 2006). Although individually these phenomena are weak, their ubiquitous occurrence makes it quite plausible that they contribute to coronal heating, especially in the quiet Sun where major flares are never seen (Parker 1988; Krucker 2002; Patsourakos & Klimchuk 2005; Brooks et al. 2008).

The corona also supports a variety of wave phenomena and mass transport processes that redistribute energy without necessarily leading to eruptions. These waves, such as MHD waves, Alfvén waves and magnetoacoustic waves provide an important mechanism for energy transport (Erdélyi & Ballai 2007; Taroyan & Erdélyi 2009) and form the basis of solar-magneto seismology techniques used to infer plasma and magnetic properties (Taroyan & Erdélyi 2009; Nakariakov et al. 2024). Mass motions such as coronal rain arise from thermal instabilities within closed loops, leading to the condensation and subsequent downward flow of cool plasma along magnetic field lines (Schrijver et al. 1999; Antolin & Rouppe van der Voort 2012; Vashalomidze et al. 2015; Antolin 2019).

2.2 Active Regions

Active regions (ARs) are localised concentrations of strong (see Figure 2.7) and complex magnetic fields on the photosphere and the solar atmosphere above it, serving as the sites of solar activity such as solar flares and CMEs (Toriumi & Wang 2019). They originate as buoyant magnetic flux tubes, rising as a consequence of the solar dynamo operating at the tachocline, the base of the Sun's convection zone (Charbonneau 2010). Above the photosphere, ARs extend into the chromosphere and corona, where magnetic field lines form loops, arcades, and more complex configurations (van Driel-Gesztelyi & Green 2015). The evolution of ARs follows a life cycle of emergence, growth, peak activity, and decay (van Driel-Gesztelyi & Green 2015). The number and complexity of ARs (and sunspots - their footprints on the photosphere) follows the 11-year solar cycle. During periods of high solar activity, ARs are more numerous, larger, and magnetically complex, increasing the likelihood of eruptive events (Guo et al. 2010).

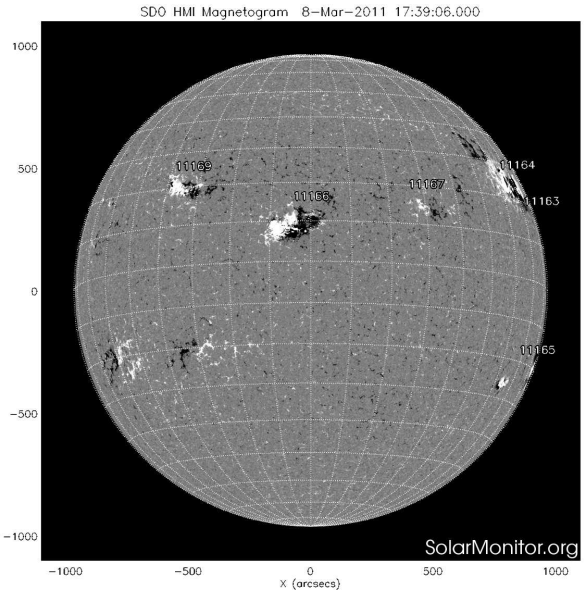

Figure 2.7: ARs as observed by SDO/HMI on Mar 08, 2011 - 17:39 UTC (courtesy: https://www.solarmonitor.org/)

2.3 Solar Energetic Phenomena

The Sun is a highly dynamic star whose magnetic activity drives a wide range of energetic phenomena in the solar atmosphere. In this section, we discuss the different solar energetic phenomena - (i) solar flares, (ii) CMEs, (iii) erupting filaments and prominences, and associated (iv) SEP (solar energetic particle) events. These phenomena are manifestations of the rapid release of magnetic energy stored in the solar corona. Despite their distinct observational signatures, these phenomena are often interconnected, with magnetic reconnection and instabilities in the coronal magnetic field serving as the fundamental triggers. Thus, understanding the mechanisms, frequency, and impacts of these eruptions is crucial, not only for solar physics but also for operational space weather forecasting, as these events can significantly affect the heliosphere, planetary environments, and technological systems in near-Earth space.

2.3.1 Solar Flares

A solar flare is an intense, transient release of energy in the solar atmosphere caused by magnetic reconnection, manifested as a burst of electromagnetic radiation (see Figure 2.8) spanning the spectrum from gamma-rays to radio waves, along with plasma heating, particle acceleration, and the generation of shock waves, releasing up to 10^{33} ergs on time scales of 10–100 seconds (Kopp & Pneuman 1976; Benz 2008; Gordovskyy et al. 2010; Korsos 2018). In magnetic reconnection, oppositely directed magnetic field lines in a plasma medium break and rejoin thereby causing the stored magnetic field energy to be dissipated in the form of plasma kinetic energy and thermal energy (Parker 1957; Sweet 1958; Petschek 1964; Zweibel & Yamada 2009). Due to reconnection, electrons and ions in the solar atmosphere are accelerated over a broad energy spectrum, ranging from thermal energies $\sim$ KeV) to non-thermal relativistic energies extending to several MeV for electrons and up to GeV for ions (Gordovskyy et al. 2010), emitting electromagnetic radiation as a consequence of their acceleration. Solar

flares are classified into A, B, C, M or X classes depending upon the peak value of the soft X-ray flux measured by GOES satellite in the 0.1-0.8 nm band (Garcia 1994). The peak X-ray flux for each class is ten times greater than the preceding one. X-class flares are the strongest with the peak value of the soft X-ray flux exceeding $10^{-4}\mathrm{Wm}^{-2}$. The duration of a solar flare may range from minutes to hours and is measured by the FWHM of the soft X-ray flux curve (Reep & Knizhnik 2019).

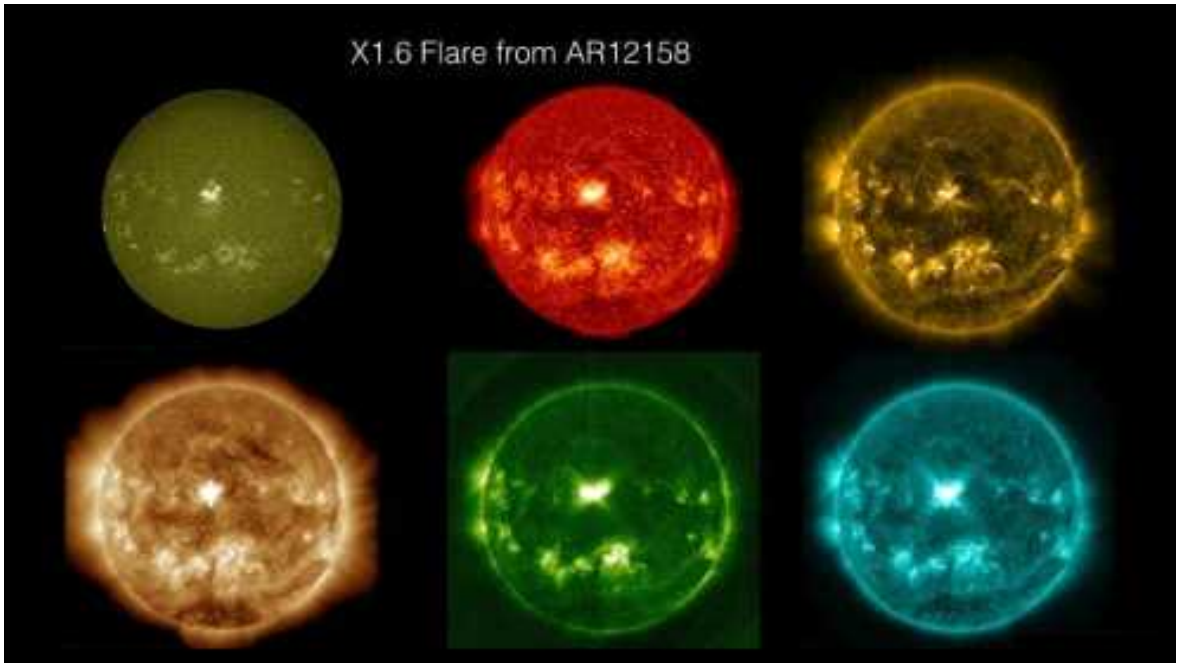

Figure 2.8: Observation of a solar flare by SDO/AIA (Atmospheric Imaging Assembly) at 17:45 UTC on Sep 10, 2014 in different wavelengths, each mapping to a specific layer in the solar atmosphere: (a) 160 nm - upper photosphere, (b) 30.4 nm (He II line) - chromosphere and transition region, (c) 17.1 nm (Fe IX line) - quiet corona, (d) 19.3 nm (Fe XII, Fe XIV lines) - AR corona, (e) 9.4 nm (Fe XVIII line) - hot flare plasma, (f) 13.1 nm (Fe VIII, Fe XXI lines) - hot flare core; notice the intense brightening at the centre which is consistent across observations in different wavelengths; courtesy: https://www.thesuntoday.org/

Flares are classified as eruptive or confined depending on whether the magnetic configuration in the solar corona is significantly altered (Woods et al. 2011; Priest 2014). While eruptive flares are associated with a significant change in the magnetic

configuration and may be accompanied with CMEs (Aschwanden et al. 2009), confined flares do not alter the magnetic configuration significantly (Švestka 1989).

Early solar flare reconnection modelling studies were predominantly two-dimensional, with the CSHKP model serving as the canonical example (Carmichael 1964; Sturrock 1966; Hirayama 1974; Kopp & Pneuman 1976). As per the CSHKP model, prior to the onset of a flare, the coronal magnetic field appears like arched loops (see Figure 2.9a). In the center of these arched loops, a prominence - cool dense plasma matter surrounded by magnetic field lines may be observed. Sometimes coronal magnetic fields are stressed (highly sheared and twisted) and have field line dips (concave-up segments of stressed magnetic fields that serve as gravitational traps - magnetic tension forces balancing the solar gravitational force) that allow dense prominence plasma to accumulate and remain suspended in the corona. Thus, prominences are one of the many symptoms of a stressed coronal magnetic field. These plasma particles may be supplied into the coronal through thermal instabilities or chromospheric evaporation (Shibata & Magara 2011). Sometimes, disturbances in the magnetic field could cause the prominence to rise upwards (see Figure 2.9b). In this phase, a vertical current sheet forms below the rising prominence, but above the flare loops, across which in the horizontal direction, the magnetic field changes direction in a relatively short distance. As the prominence rises, it stretches the magnetic field below it, pulling field lines into a narrow region, which serves as the reconnection site within the current sheet. As reconnection starts at these sites, thermal energy is released. This energy is transferred downwards causing the heated surface plasma to evaporate upward into a newly formed loop. As the reconnection sustains, its continues to create more hot loops. Meanwhile, the loops that had formed previously cool down and drain (see Figure 2.9c). The overall summary of the flare's main phase is that (i) the bright ribbons in the Hα lines (which serve as the footpoints of the newly reconnected magnetic field lines) separate, and (ii) the soft X-ray producing hot loops keep on developing and appear to rise with time, encapsulating the previously formed loops (Priest 2014).

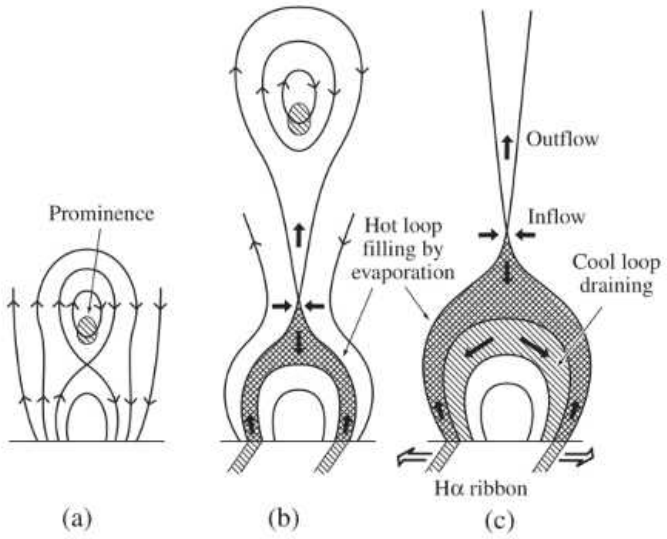

Figure 2.9: Different stages of a solar flare as per the CSHKP model; (a) prior to flare onset; (b) formation of the first flare loop; (c) formation of successive hot flare loops and draining of cooler loops; adapted from Priest (2014)

Shibata & Magara (2011) have provided a synoptic chronological overview of the developments made to the CSHKP model after the first version provided by Kopp & Pneuman (1976). Kopp & Pneuman (1976) proposed that after the open field lines reconnect to form a closed loop, the solar wind upflows along these reconnected lines collide, producing a gas dynamic shock within the loop that facilitates the heating of coronal plasma (see Figure 2.10a). While acknowledging the merits of the CSHKP model, Cargill & Priest (1982) further suggested that the heating in the flaring loops may be attributed to the slow MHD shock produced as an outcome of Petschek-type reconnection (see Figure 2.10b). The model of Cargill & Priest (1982) however did not explain how reconnection outflows interact with closed flare loops. This deficiency was later addressed by Forbes & Priest (1984), who showed that magnetic reconnection in the solar corona produces bi-directional reconnection outflows or jets (in the vertical direction of the CSHKP model). When the downward reconnection jet interacts with the closed flare loops below, it decelerates abruptly, leading to the formation of a fast-mode MHD shock (see Figure 2.10c). The fast MHD shock acts as a termination shock

for the reconnection outflow and causes a sudden localised increase in temperature and density, which serves as a source of thermal X-rays (Forbes & Priest 1984). The earlier models (Cargill & Priest 1982; Forbes & Priest 1984) largely explained the occurrence of flares in an ideal-MHD setting and did not account for thermal conduction in the solar corona. In this context, Forbes & Malherbe (1986) made a major conceptual advance by integrating thermodynamics into the MHD-based model. They showed that classical slow-mode MHD shocks cannot remain intact because of thermal conduction and must disintegrate into (i) an isothermal slow shock, and (ii) a thermal conduction front extending ahead of the shock (see Figure 2.10d). They demonstrated that a large fraction of the energy released at reconnection shocks is conducted downward which drives chromospheric heating and evaporation of plasma material into flare loops.

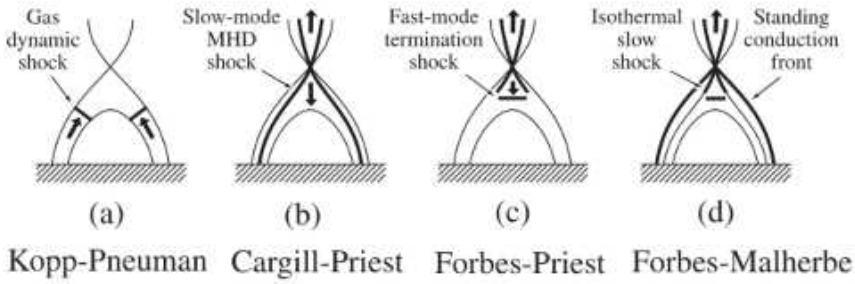

Figure 2.10: Development of different models of flare loop formation; adapted from Priest (2014)

The energy dissipation in solar flares happens in two major phases - (i) an impulsive phase and (ii) a gradual phase, each describing the release of radiation from a specific wavelength range and thus, having its own characteristic X-ray signature (Hoyng et al. 1981; Hudson 2011; Woods et al. 2011; Woods 2014; Pandey et al. 2023). The impulsive phase is characterised by non-thermal hard X-ray bremsstrahlung and microwave gyrosynchrotron radiation and can also be seen in the transition region He II line (30.4 nm) (Hudson 2011). By non-thermal, we imply that the radiation is produced by electrons whose energy distribution is not thermal (not Maxwellian), but instead consists

of accelerated, high-energy electrons interacting with ions in the solar plasma. The gradual phase of a solar flare is characterised by thermal soft X-ray emission from the coronal plasma, EUV emission and emissions from Fe XVI ($\sim 3 \times 10^6$ K, 33.5 nm), Fe XX ($\sim 10^7$ K, 12.2 nm) and Fe XXII ($\sim 1.3 \times 10^7$ K, 13.6 nm) coronal lines, all of which diagnose and trace a range of corona temperatures (Del Zanna & Woods 2013; Pandey et al. 2023). Solar flares are classified as impulsive or gradual depending on whether their temporal evolution exhibits a steep, rapid rise phase or a more gradual and extended increase in emission. A detailed diagram based on the CSHKP model for both these phases is shown in Figure 2.11.

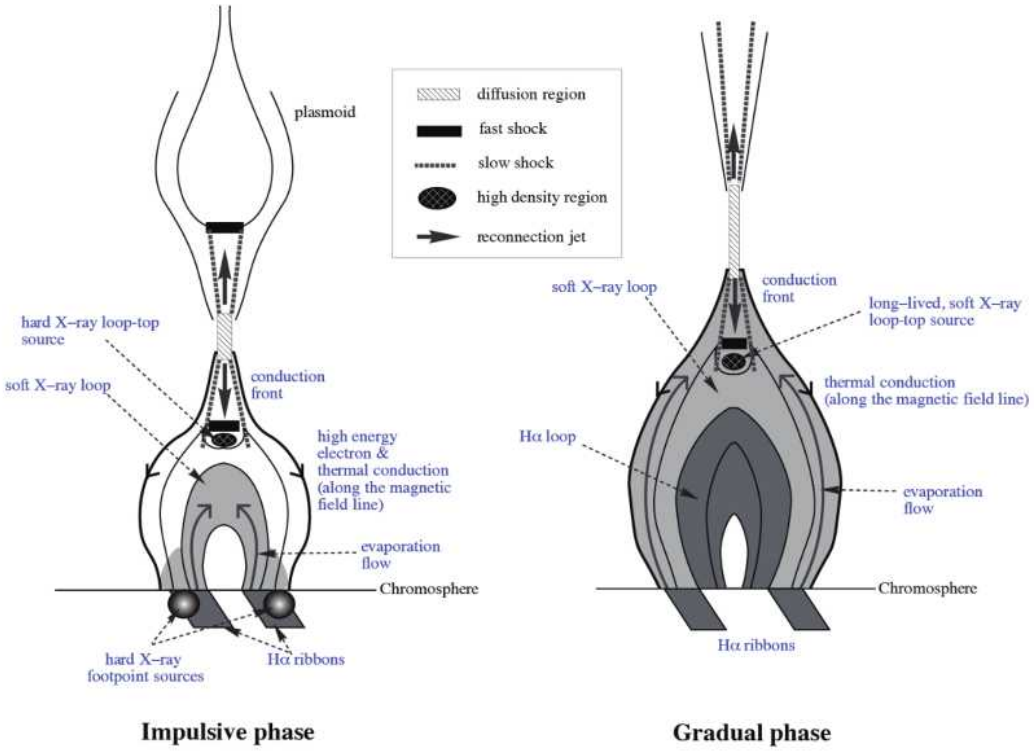

Figure 2.11: Schematic comparison of solar flare morphology (in two dimensions) during the impulsive and gradual phases; the impulsive phase exhibits compact, sharply localised loop-top and chromospheric footpoint structures associated with narrow flare loops, whereas the gradual phase is characterised by expanded, filled loop arcades with diffuse, long-lived loop-top emission and widely separated chromospheric ribbons (adapted from Magara et al. 1996 and Shibata & Magara 2011)

Although fully 3D MHD simulations of flares (or magnetic reconnection) existed earlier (see examples such as Mikic & Linker 1994; Amari et al. 1996; Fan & Gibson 2003; Roussev et al. 2003; Roussev et al. 2004), it was not until the early 2010s that synthesis and refinement of existing ideas led to the development of the 'the standard flare model in three dimensions', a 3D analogue of the 2D CSHKP model (Aulanier et al. 2012; Aulanier et al. 2013; Janvier et al. 2013; see Figure **??** as an example). In the first paper in a series of three papers on the 3D standard flare model, Aulanier et al. (2012) using the Observationally-driven High-order scheme Magnetohydrodynamic (OHM) code (Aulanier et al. 2005; Aulanier et al. 2009) and addressed how 3D reconnection dynamics drive the evolution of shear from strong to weak in post-flare loops - an observation that could not be explained in the 2D CSHKP model. The discussion in the second paper by Aulanier et al. (2013) was solely confined to probing the energetics of flares using the 3D MHD model. Aulanier et al. (2013) suggested that the maximum energy that may be released in a flare is $\sim 6 \times 10^{33}$ erg. The final paper in this sequel by Janvier et al. (2013) showed that slip-running reconnection in quasi-separatrix layers (regions where connectivity of magnetic field lines changes very rapidly) governs flux rope and flare loop evolution, thereby completing the standard flare model by providing its essential 3D physical framework. A comprehensive review of the standard 3D flare model and other recent developments in this regard can be found in Pontin & Priest (2022).

2.3.2 Coronal Mass Ejections

A CME is characterised by a large release of magnetic energy and plasma particles, often thought to be in the form of a helical magnetic flux rope (a bundle of twisted magnetic field lines containing hot plasma), from the solar corona into the interplanetary medium (Low 1994; Dere et al. 1999; Webb & Howard 2012; Patsourakos et al. 2020). The existence of CMEs was firmly established in the early 1970s through white-light coronagraph observations aboard the Skylab mission (Hansen et al. 1971; Tousey 1973; Gopalswamy 2016). Instruments such as the Large Angle and Spectrometric Coronagraph (LASCO) on board the Solar and Heliospheric Observatory (SOHO) and the Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI) on board the Solar TErrestrial RElations Observatory (STEREO) have provided continuous observations of CMEs propagating through the corona (Gopalswamy et al. 2009; Byrne et al. 2009; Webb & Howard 2012). From a time series of coronographic sequences, basic attributes of a CME such as mass, speed and width can be determined.

A significant fraction of CMEs ($\geq$40%) are launched as magnetic flux ropes which expand radially outward as they erupt (Vourlidas et al. 2013; Patsourakos et al. 2013). Traditional description of a CME as an expanding flux rope has led to the conceptualisation of the 'three-part morphology' - an observational paradigm to describe its shape (see Figure 2.12). As per this model, a CME comprises of (i) a bright leading edge - interpreted as a compressed plasma region, (ii) a surrounding dark cavity - interpreted as flux rope cross section or sometimes a low density region, and (iii) a bright core - interpreted as an embedded dense material such as an erupting prominence, respectively (Illing & Hundhausen 1985; Webb & Hundhausen 1987; Song et al. 2022). However, not all CMEs conform to the three-part morphological description. If a CME has a bright, filamentary loop but lacks a cavity or a core, then it is considered as a loop CME (Vourlidas et al. 2013). Some CMEs appear to extend outward from the Sun in nearly straight lines and thus are called jet CMEs (Vourlidas et al. 2013). There are

also events that are wider than jets, without a clear loop front or cavity, and are called outflow CMEs (Vourlidas et al. 2013). Thus, loop CMEs, jet CMEs, outflow CMEs and other irregularly shaped CMEs are some of the prominent non-three-part morphology-type CMEs (Vourlidas et al. 2013). As CMEs propagate outward from the Sun, due to excessive internal magnetic pressure, they expand and along radially outward magnetic field lines, naturally forming a cone-shaped volume. Thus the the angular width of the CME is studied to quantify the spread arising out of radial expansion. The angular width of CMEs may vary between as little as a few degrees to more than 120°, with an average value ∼46° (Gopalswamy 2006). Based on their angular width (θ), CMEs may be classified as narrow ($\theta < 30°$) or wide ($\theta > 60°$). Subject to viewing geometry, while some narrow CMEs may be jet-like in appearance (Bronarska et al. 2018), some wide CMEs may appear as partial halo CMEs ($120° < \theta < 360°$) or halo CMEs ($\theta \sim 360°$) - a special category of CMEs which appear to surround the occulting coronograph disk and are either Earth or anti-Earth directed (see Figure 2.12).

A typical CME carries a mass between 10^{13} - 10^{16} g (Gopalswamy 2006) and can reach speeds ranging from a few kms^{-1} to about 3000 kms^{-1}, with the average value $\sim$ 483 kms^{-1} (Yashiro et al. 2004; Gopalswamy 2006). From a statistical study, Yurchyshyn et al. (2005) found that the speed of CMEs is log-normally distributed. From a comprehensive analysis of CMEs of SC 23, Vourlidas et al. (2010) showed that the mass and energy properties of CME may fluctuate at the start but reach constant levels eventually and must be measured only above $\sim 10 R_S$. Vourlidas et al. (2010) found that pseudo CMEs (apparent CME-like features seen in coronagraph images that are not true mass ejection from the Sun) usually reach a peak mass at about $7 R_S$ in the solar corona and then suddenly disappear altogether. Just like the speed, the CME mass and energy distribution also follow a log-normal distribution when measured at about 10-15R_S (Vourlidas et al. 2010). From the measurements of mass and speed, it is estimated that the kinetic energy of a CME is between 10^{27}-10^{32} erg and the average kinetic energy is $\sim 5 \times 10^{29}$ erg (Gopalswamy 2006).

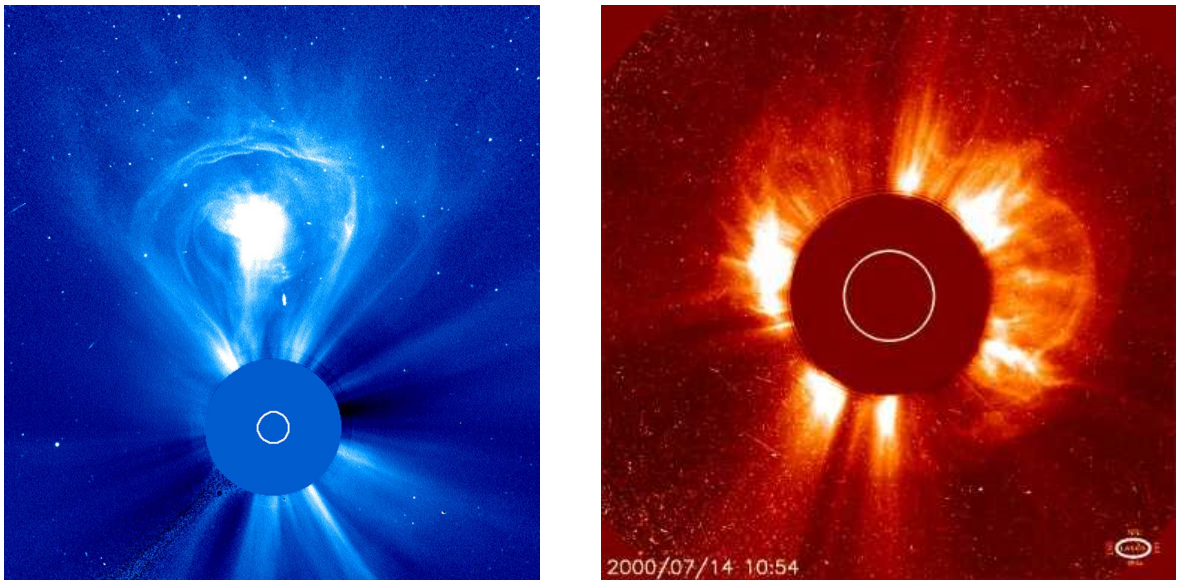

Figure 2.12: A lightbulb-shaped CME (left) showing the three-part morphology: a leading edge of compressed plasma, a void (magnetic flux rope - 'the bulb' here), and a bright core, taken on Feb 27, 2000 by LASCO; a halo CME (right) as recorded by LASCO on Jul 14, 2000 (courtesy: https://soho.nascom.nasa.gov/)

The frequency of CME occurrence is well correlated with solar activity. While during a solar minima one would expect 1 CME per day, one would expect about 5 CMEs per day during a solar maxima (St. Cyr et al. 2000; Webb & Howard 2012). Although CMEs may occur throughout different latitudes during a solar maximum, they are usually confined to lower latitudes during a solar minima (Webb & Howard 2012). Hundhausen (1993) who first pointed this out, noted that the latitude variation of CMEs follows the locations of streamers and prominences more closely than that of ARs or sunspots. Based on their origin, CMEs may be classified as AR CMEs, if they originate from within ARs or filament/prominence-eruptions linked to ARs (Subramanian & Dere 2001; Majumdar et al. 2020; Maričić et al. 2020; Pant et al. 2021). CMEs which are linked to ARs and flares are impulsively accelerated and are faster (velocities $\geq$ 750 kms^{-1}) compared to those originating from weak-field prominence eruptions which undergo gradual acceleration (velocities between 400-600 kms^{-1}) (MacQueen & Fisher 1983; Sheeley Jr. et al. 1999; Filippov 2019).

While, multi-point observations from the STEREO spacecraft have enabled three-dimensional reconstruction of CME morphology and shaped our understanding of CME kinematics, complementary observations in EUV and X-ray wavelengths reveal coronal and low coronal signatures associated with CME initiation, including observational features such as coronal dimmings, filament eruptions, and post-eruption arcades.

Zuccarello et al. (2013) have presented an overview of different mechanisms for CME initiation and have broadly classified the initiation mechanism into two types - (i) those requiring magnetic reconnection, and (ii) those not requiring magnetic reconnection. The models that require magnetic reconnection, such as the flux cancellation model (van Ballegooijen & Martens 1989; Forbes & Priest 1995) and breakout model (Antiochos et al. 1999; Green et al. 2011; Chen 2011), reconnection plays a fundamental role in the evolution of the coronal magnetic field. These mechanisms can remove the stabilising magnetic flux and contribute to the formation or strengthening of a flux rope, which may eventually erupt if it becomes unstable, leading to the rapid release of stored magnetic energy. In such models, reconnection enables the loss of equilibrium and provides efficient acceleration, often leading to fast, flare-associated CMEs. However, for the models that do not require magnetic reconnection, such as the mass loading model (Low 1996; Wolfson & Dlamini 1997), the kink instability model - caused by excessive twisting of flux rope (Roussev et al. 2003; Török & Kliem 2003; Török et al. 2004; Török & Kliem 2005), and the torus instability model - cause by rapid weakening of overlying magnetic confinement fields (Kliem & Török 2006; Török & Kliem 2007), the underlying concept is that the existing magnetic structure gradually loses stability. In these models, the eruption occurs when internal forces within the magnetic structure can no longer be balanced by gravity or the surrounding coronal field, and magnetic reconnection, if present, plays only a secondary role and is not required to drive the eruption. In this context, it is further known that the kink instability is not quite efficient in producing CMEs compared to the torus instability (Liu 2008; Jing et al. 2018).

The causal physical mechanisms behind CMEs have been discussed elaborately in Green et al. (2018) and these processes have been classified as 'triggers' or 'drivers'. Green et al. (2018) define drivers as mechanisms that may account for the rapid acceleration and huge expansion of plasma and magnetic field observed during eruptions, i.e., those mechanisms that may produce a CME. Examples include flare reconnection and torus instability. On the other hand, those mechanisms that initiate eruptions, but cannot independently produce a CME are considered triggers. Examples of triggers are much more numerous and include sunspot rotation, shearing of arcade, magnetic breakout, flux cancellation, flux dispersion, flux emergence (Archontis & Syntelis 2019), kink instability etc (see Table 2 of Green et al. (2018) for more examples and references).

Once initiated, what sustains the motion of CMEs has been a topic of scientific enquiry over the last couple of decades. In the low to mid corona (2-20R_S), CMEs are propelled by Lorentz forces. In this regime, it is the free magnetic energy (defined as the difference between the total magnetic energy and the energy content of the potential field configuration) that acts as a viable source of energy to propel the CME outside the solar atmosphere (Subramanian & Vourlidas 2007). However once the CME enters mid to outer corona ($\geq 15R_S$), as per the widely accepted drag-based model, it is the ambient solar wind that helps in the propagation of CMEs by exerting a drag force arising out of collision-free transfer of momentum and energy between the CME and the ambient solar wind (Cargill 2004; Vršnak & Žic 2007; Vršnak et al. 2013; Žic, Vršnak & Temmer 2015; Dumbović et al. 2018; Napoletano, Gianluca et al. 2018; Chierichini, Simone et al. 2024; Mugatwala et al. 2024). Based on their speeds, CMEs are classified as fast (speed $\geq$ 450 kms^{-1}), slow (speed $\leq$ 250 kms^{-1}) or intermediate (speed between 250-450 kms^{-1}) (Yashiro et al. 2004). Yashiro et al. (2004) found that most of the slow CMEs undergo acceleration and most of the fast CMEs undergo deceleration because of the drag force. This is because the drag force depends on the ambient solar wind speed (400 kms^{-1}) and causes CMEs to decelerate if they exceed the solar wind speed and accelerate if they are lesser than the solar wind speed.

The geo-effectiveness of CMEs, i.e. the ability of CMEs to produce a significant geomagnetic disturbance, is not determined solely by their launch properties, but also by their subsequent evolution during heliospheric propagation. In particular, CME deflection and rotation can substantially modify the orientation of the embedded flux rope relative to Earth, thereby influencing the southward magnetic-field component that governs geomagnetic coupling. Studies have shown that significant deflection and rotation may occur both below 30 R_S and during propagation to 1 AU (Astronomical Unit), with interactions involving the background solar wind, Parker-spiralled magnetic field, and heliospheric current sheet playing important roles in shaping CME trajectories and orientations (Isavnin et al. 2014). Furthermore, magnetic erosion caused by reconnection between the CME flux rope and the ambient heliospheric magnetic field can progressively remove magnetic flux and outer-shell mass from magnetic clouds during transit (Pal et al. 2020). Such erosion may reduce CME geo-effectiveness and alter propagation dynamics, including arrival time and speed at 1 AU, with modelling studies suggesting delays of up to three hours for strongly eroded events (Stamkos et al. 2023). Consequently, accurate assessment of CME geo-effectiveness requires consideration not only of eruption initiation, but also of the complex evolutionary processes acting throughout CME propagation in the heliosphere, which remain major challenges for space-weather forecasting (Vourlidas et al. 2019).

Even though (i) flares - which are primarily emissions of electromagnetic radiation and (ii) CMEs - which are primarily ejections of plasma and magnetic fields from the Sun, appear as distinct phenomena, they are actually considered as different manifestations of the same underlying mechanism (see Figure 2.13) of magnetic energy release (Gosling 1990; Harrison 1995; Gopalswamy 2016; Gou et al. 2020). A basic observational test on CME-flare pairs showing that they share a common formation mechanism are the soft/hard X-ray light-curves and CME acceleration timings (Zhang et al. 2004; Temmer et al. 2010). Despite decades of observations and modelling, several fundamental challenges remain in CME research. Predicting the onset time and location of

CMEs remains difficult due to incomplete knowledge of coronal magnetic fields. Further, the precise dynamics of CMEs interaction with the Earth's magnetosphere is also not fully understood. Since it is established that strong flares (particularly X-class flares) show a strong association with CMEs (Yashiro et al. 2005), the importance of forecasting strong flares to secure critical infrastructure in the technosphere, stands reinforced once again. In this context, future missions and advanced data-driven MHD models are expected to improve our understanding of CME initiation, evolution, and impact.

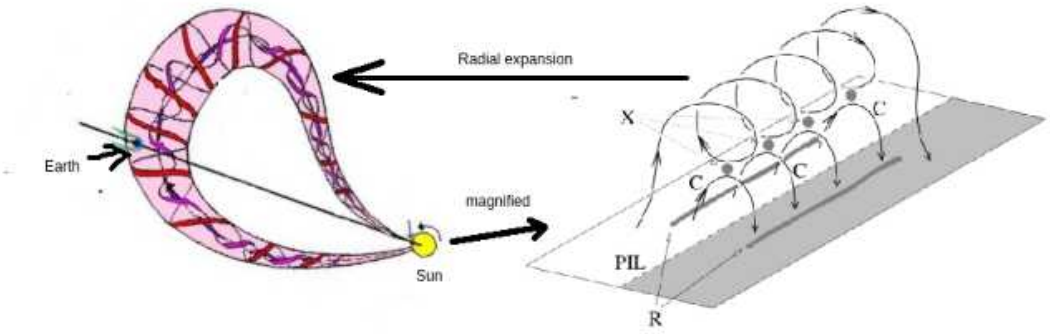

Figure 2.13: A representative image (not to scale) showing how the flux rope in the model by (Longcope et al. 2007) eventually expands into a CME (Marubashi 1997); adapted from Dr. Nat Gopalswamy's public talk and presentation on CMEs at the ISWI-SCOSTEP School at the Indian Institute of Geomagnetism, Mumbai (Jan 2026)

2.3.3 Erupting Filaments and Prominences

Filaments and prominences (see Figure 2.14) are essentially the same physical structures. They are clouds of cooler, denser plasma suspended above the Sun's surface by magnetic fields (Parenti 2014). When seen against the solar disk, they appear dark and are called filaments and when seen projected beyond the solar limb, they appear bright against the blackness of space and are called prominences (Parenti 2014). Filaments and prominences are always found above PILs and can exist for days or weeks as stable structures (Parenti 2014). However, when they become unstable and are ejected into space, they are involved in eruptive solar phenomena like coronal mass ejections (CMEs). So, a filament or prominence can be part of a solar eruption, but not every filament/prominence is eruptive. An eruption happens when the magnetic field slowly evolves due to motions on the solar surface and eventually crosses a stability threshold. The loss of equilibrium can happen through torus instability (Kliem & Török 2006; Parenti 2014).

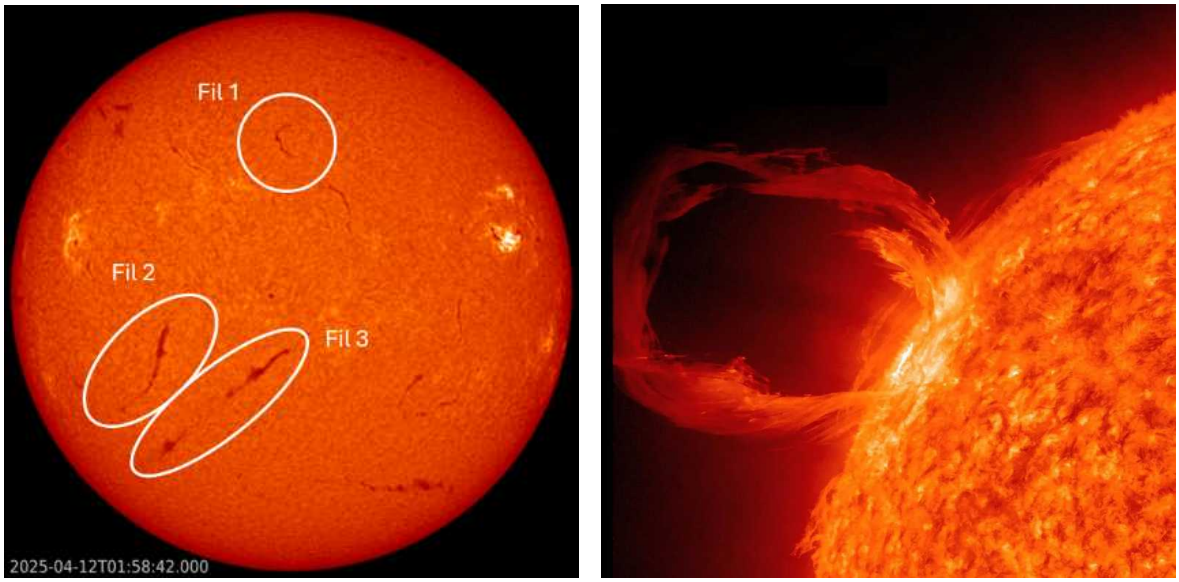

Figure 2.14: An image of filaments on the solar disk (left, courtesy: Royal Observatory of Belgium, https://www.sidc.be/article/filament-eruptions) and a prominence on the solar limb (right, courtesy: NASA, https://www.nasa.gov/image-article/what-solar-prominence/)

2.3.4 Solar Energetic Particle Events

Solar energetic particle (SEP) events are episodes in which ions and electrons are accelerated to high energies and propagate through the heliosphere, posing hazards to spacecraft, astronauts, and technological systems (Desai & Giacalone 2016). SEP events are commonly categorized into two idealised classes - (i) impulsive and (ii) gradual, based on their observational characteristics, source regions, and acceleration mechanisms, although many events exhibit mixed properties (Reames 1999; Desai & Giacalone 2016).

Impulsive SEP events are typically short-lived, electron-rich, and are associated with solar flares. These events are thought to originate from compact reconnection sites in the low corona, where particles are rapidly accelerated and injected onto open magnetic field lines (Miller et al. 1997; Reames 1999). They are characterised by strong enhancements of He^3 and heavy ions (Desai & Giacalone 2016).

Gradual SEP events, by contrast, are long-lasting and are dominated by protons, often extending over days and broad longitudinal ranges (Desai & Giacalone 2016). They are closely associated with fast coronal mass ejections (CMEs) that drive shock waves through the corona and interplanetary medium. Diffusive shock acceleration at shocks driven by CMEs, is widely accepted as the primary mechanism responsible for accelerating particles to energies exceeding hundreds of MeV (Desai & Giacalone 2016).

2.4 Flare Predictors

Predictors are measurable quantities used to estimate the likelihood, timing, or magnitude of a future event. Thus in this context, a flare predictor would indicate any measurable quantity that helps estimate the likelihood, timing or magnitude (GOES class) of a potential solar flare. In solar flare forecasting research, efforts have been made to identify parameters that could help distinguish between flaring and non-flaring regions on the Sun. For example, the FLARECAST project (http://flarecast.eu/) of the European Union which was focused on developing a quantitative solar flare forecasting system, identified a total of 209 solar flare predictors (see Georgoulis et al. 2021 and references therein). In this section, we shall (i) briefly discuss different predictors which have been used to predict the occurrence of solar flares in the past, (ii) classify them into different groups based on their attributes/functionalities and (iii) delve deeper into some specific predictors which are of significance for this thesis.

2.4.1 Flare Predictors in Literature

Solar flares and CMEs are fundamentally magnetic energy release phenomena. Thus to predict the occurrence of solar flares (and CMEs), it would be quite intuitive to trace the evolution of the solar magnetic field and identify any specific patterns in the magnetic field in form of predictors which are characteristic to flares. At the same time, we do note that the solar magnetic field does not evolve in isolation. There could be other parameters that may quantify non-magnetic features such as plasma properties (density, temperature), plasma dynamics (flow, velocity), coronal emission measures or historical behaviour to indirectly signal flare readiness. While, it should be noted that nearly all processes in the solar atmosphere are ultimately influenced by magnetic fields, for simplicity and ease of classification, the predictors are classified as magnetic or non-magnetic, depending on whether they directly depend on the magnetic field (see Table 2.1).

Class	**Sub-class**	**FLARECAST Predictors**
Magnetic	Morphological	sunspot structure (Hale/McIntosh classification)
		sunspot counts & sizes, PIL shape & count
		fractal dimension related parameters
Magnetic	Configurational	PIL-related metrics, unsigned flux, R-value
		effective connected field, ising energy
		average shear angle, WG_M parameter
		total/mean free energy
		vertical current density, unsigned vertical current
		twist, current helicity density, unsigned helicity
		Fourier and wavelet power spectral indices
Magnetic	Topological	null points, decay index, magnetic helicity
		relative magnetic helicity, helicity injection rate
Non-magnetic	Radiation-based	GOES class related (X-ray), EUV imaging
		line-of-sight doppler velocities
Non-magnetic	Helioseismic	solar oscillations, vorticity enhancements
		divergent flows, acoustic wave signatures

Table 2.1: Classification of FLARECAST predictors into different classes and sub-classes; the classification introduced here is intended for simplicity and does not represent a standard taxonomy in the literature

2.4.2 Magnetic Field-based Predictors

Magnetic field-based predictors are directly based on (or inferred from) the magnetic field and are further divided into three different sub-classes, namely - (i) morphological, (ii) configurational and (iii) topological.

1. Magnetic morphological predictors

Magnetic morphological predictors describe the large-scale structure and appearance of the photospheric magnetic field and help quantify how the magnetic polarities are arranged in general. Key examples of the magnetic morphological predictors are schemes to group sunspots based on their morphological complexity (see Section 2.4.6) and other sunspot structure metrics such as counts and sizes. Polarity Inversion Lines (PILs) are regions that demarcate areas of opposite polarities in the solar atmosphere (see Section 2.4.4). The qualitative PIL shape when converted into quantifiable measures, the number of PILs (and even the presence or absence of PILs altogether), may also serve as morphological predictors (Ji et al. 2023).

One major question in solar flare forecasting is whether, in flare-productive ARs, the magnetic field is smoothly organised at large spatial scales or is instead fragmented into a wide range of finer spatial structures. Observational studies have shown that ARs with higher flare productivity exhibit stronger intermittency and broader multi-fractal behaviour at small spatial scales, indicative of an increased degree of small-scale magnetic fragmentation (Abramenko & Yurchyshyn 2010). In this context, some classic multi-scale predictors such as fractal dimension, multi-fractal structure functions, scaling exponents and intermittency index also capture the morphology of the solar atmosphere (Abramenko et al. 2003; Abramenko 2005; McAteer et al. 2005; Conlon et al. 2008).

Traditionally, many of the configurational predictors mentioned in the next segment such as unsigned magnetic flux (Schrijver 2007), R-value (Schrijver 2007), flux weighted PIL length (Falconer et al. 2002; Falconer et al. 2003; Mason & Hoeksema 2010), WG_M parameter (Korsós et al. 2015; Korsós et al. 2020), effective connected magnetic field strength B_{eff} (Georgoulis & Rust 2007), total unsigned electric current (Georgoulis et al. 2012), also implicitly describe the morphology of the solar surface and may be considered as morphological parameters (Kontogiannis et al. 2018). Even though these parameters describe the overall appearance of the magnetic field, the algorithms used to estimate these parameters are heavily dependent on the magnetic configuration, in quantitative terms and thus we categorise them separately.

2. Magnetic configurational predictors

Magnetic configurational predictors describe and quantify the spatial distribution of the magnetic field in the form of geometrical metrics, orientations in space, gradients, and relationships between opposite polarity regions. These predictors are computed from the magnetic field distribution, measured at the photosphere and often are direct indicators of physical processes on the Sun. In this context, the structure of PILs and related metrics such as the length of PILs (Falconer et al. 2002; Falconer et al. 2003; Mason & Hoeksema 2010), the unsigned flux near PILs (Schrijver 2007; Hazra et al. 2020) and the R-value parameter (Schrijver 2007; Hazra et al. 2020) are examples of magnetic configurational predictors (more details in Section 2.4.5). B_{eff}, the effective connected magnetic field, developed by Georgoulis & Rust (2007), quantifies the strength and compactness of magnetic connections between opposite polarities in a solar AR by weighting magnetic flux by separation distance. Similar to B_{eff}, the Ising energy (Kontogiannis et al. 2018) is a complexity measure that also quantifies the interaction between opposite polarity magnetic elements in a solar AR, with stronger contributions from closely spaced opposite polarities such as those found along PILs.

The Sun's magnetic field is considered potential if it is current-free and in its minimum-energy state (see details in Section 2.6). In such a state, the solar magnetic field lines point radially outwards or inwards. However when the solar magnetic fields become stressed i.e. either tilted, sheared or twisted, the system deviates from its potential configuration and free energy begins to accumulate. This stored excess energy gets released in the form of solar eruptive phenomena when there is loss of equilibrium. The total and mean photospheric magnetic free energy densities are also configurational predictors that quantify free energy build-up available for release during a flare (Hazra et al. 2020).

In this connection, several SHARP (Space-weather HMI Active Region Patch) predictors identified by Bobra et al. (2014) such as the spatial gradients of the magnetic field (see application in Wang et al. 2006), average shear angle (Leka & Barnes 2003; Tiwari et al. 2010) and the fraction of area with high shear angle ($> 45°$) (similar description in Falconer et al. 2002), the G_M parameter (Korsós et al. 2014) and the WG_M parameter (Korsós et al. 2018; Korsós et al. 2020) measure the shear in the magnetic field (effectively serving as measures of non-potentiality) and thus, fall under the category of magnetic configurational predictors. Similarly the inclination angle from the radial field measures the tilt in the solar magnetic fields and is a descriptor of non-potentiality in the system (Bobra et al. 2014).

As noted earlier, the presence of currents in the system is an indicator of non-potentiality. While predictors such as the vertical current density (see kurtosis calculations in (Leka & Barnes 2003; De Vita, Gaetano et al. 2015; Sinha et al. 2022) and the total unsigned vertical current (Hazra et al. 2020) primarily quantify the strength of electric currents associated with magnetic non-potentiality and shear in the solar atmosphere, there are other current-based predictors that capture the helical and twisted nature of magnetic fields. These include the current helicity density (Zhang 2002; Leka

& Barnes 2003), total unsigned current helicity (Sinha et al. 2022; Garland et al. 2023), absolute value of the net current helicity (Garland et al. 2023), and twist-related characteristic parameters (Leka & Barnes 2003; Sinha et al. 2022). Helicity-based quantities such as helicity flux (Soós et al. 2022) are commonly interpreted as proxies for magnetic twist (Bobra et al. 2014).

Similar to the morphological multi-scale predictors configurational multi-scale predictors such as Fourier and wavelet power spectral indices (Hewett et al. 2008) quantify how magnetic energy is distributed across spatial scales and provide a means of characterising the multi-scale organisation of the magnetic field.

3. Magnetic topological predictors

Topological predictors quantify the connectivity, linkage, and the global organisation of the magnetic field (which cannot be removed by smooth deformations), rather than just its local morphology. They describe how magnetic field lines are interlinked in space, which directly relates to magnetic energy storage and stability. While configurational predictors such as current helicity density and unsigned current helicity measure the local or averaged field twisting, topological predictors such as magnetic helicity (Nindos 2008; Hazra et al. 2020) and relative magnetic helicity (Pariat et al. 2017) measure the globally conserved linkage of magnetic field lines. Similarly magnetic helicity injection rate is a topological predictor that measures how fast helicity is transported across the photosphere into the corona and studies show that flare-productive ARs often exhibit sustained, high helicity injection before major events (Yokoyama et al. 2003; Moon et al. 2003).

While in 2D, null points represent true topological discontinuities, in 3D quasi-separatrix layers are quasi-topological structures characterised by strong gradients in

field-line mapping. These features are indicators of likely reconnection sites. The number and quantifiable properties of 2D-null points (Haynes & Parnell 2007), 3D-quasi-separatrix layers (Priest & Démoulin 1995; Démoulin et al. 1996; Vemareddy 2021; quasi separatrix layers identified by high squashing factor - see Dudík et al. 2025) are topological predictors of flare occurrence.

Predictor remarks	Equation
Unsigned magnetic flux is obtained by summing the absolute vertical magnetic field value B_z, multiplied with pixel area i.e. a	$\phi_{unsigned} = a \sum_{ij} \lvert B_z \rvert_{ij}$ (2.1)
R-value is obtained by taking a weighted sum of unsigned flux elements; the weight element W_{ij} is computed from a strength threshold bitmap M (say $\lvert B_z \rvert > 150$ G), convolved with an area-normalized Gaussian with FWHM of D_{sep} (say 15 Mm)	$R\text{-value} = a \sum_{ij} \lvert B_z \rvert_{ij} \cdot W_{ij}$ (2.2) $W = M * FWHM(D_{sep})$
WG_M parameter measures the horizontal gradient of the vertical magnetic field, weighted by the magnetic flux of opposite-polarity sunspot umbrae; d^{+-} is the distance between the he area-weighted centroids of the opposite-polarity umbrae	$WG_M = \left\lvert \frac{\sum_i \phi_i^+ - \sum_j \phi_j^-}{d^{+-}} \right\rvert$ (2.3)
B_{eff} is computed through partitions; N^+, N^- are the number of positive and negative magnetic partitions, ϕ_{ij} is the magnetic flux connecting the i-th positive partition to the j-th negative partition and L_{ij} is separation distance between the flux-weighted centroids of the connected partitions.	$B_{eff} = \sum_{i=1}^{N^+} \sum_{j=1}^{N^-} \frac{\phi_{ij}}{L_{ij}^2}$ (2.4)

Table 2.2: Mathematical definitions for different flux-based predictors

Predictor remarks	Equation
L_{SS} is obtained by summing over PIL elements where horizontal magnetic field $\|B_h\| > 150$ G and shear angle $\theta_{sh} > 45°$	$L_{SS}(\|B_h\|, \theta_{sh}) = \int_{PIL} dl$ (2.5)
... where the horizontal magnetic field B_h is given as:	$B_h = \sqrt{B_x^2 + B_y^2}$
Magnetic shear angle θ_{sh} is the angle between the observed magnetic field ($\vec{B_{obs}}$) and the potential field ($\vec{B_{pot}}$); obtained using the dot product formula	$\theta_{sh} = cos^{-1}\left(\frac{\vec{B_{obs}}.\vec{B_{pot}}}{\|\vec{B_{obs}}\|\|\vec{B_{pot}}\|}\right)$ (2.6)
L_{SG} is obtained by summing over PIL elements where horizontal magnetic field $B_h > 150$ G and horizontal gradient of vertical magnetic field $\|\nabla_h B_z\| > 50$ G/Mm ...	$L_{SG}(B_h, \|\nabla_h B_z\|) = \int_{PIL} dl$ (2.7)
... where the horizontal gradient of the vertical magnetic field is:	$\|\nabla_h B_z\| = \sqrt{\left(\frac{\partial B_z}{\partial B_x}\right)^2 + \left(\frac{\partial B_z}{\partial B_y}\right)^2}$
$GWILL$ is obtained by taking the line integral of the horizontal gradient $\|\nabla_h B_z\|$ of the vertical magnetic field B_z across the PIL	$GWILL = \int_{PIL} \|\nabla_h B_z\| dz$ (2.8)

Table 2.3: Mathematical definitions for different PIL-related predictors

Predictor remarks	Equation
Ising energy measures the packing compactness of opposite polarity fluxes; the distance between polarities d_{ij} is a weighing factor.	$E_{\text{Ising}} = -\sum_{i,j} \frac{B_i B_j}{d_{ij}^2}$ (2.9)
Total photospheric free energy measures the excess energy in the system compared to a potential-field configuration.	$U = \frac{1}{2\mu_0} \sum_{i=1}^{N} (\vec{B_{obs}} - \vec{B_{pot}})_i^2$ (2.10)
Mean photospheric free energy measures the excess energy density in the system compared to a potential-field configuration; N is the number of pixels in the map	$\bar{U} = \frac{1}{2N\mu_0} \sum_{i=1}^{N} (\vec{B_{obs}} - \vec{B_{pot}})_i^2$ (2.11)
Vertical current density $\bar{j_z}$ measures the amount of vertical current in the system, averaged over the magnetogram.	$\bar{j_z} = \frac{1}{N\mu_0} \sum_{i=1}^{N} \left(\frac{\partial B_y}{\partial B_x} - \frac{\partial B_x}{\partial B_y} \right)_i$ (2.12)
Total unsigned vertical current j_z measures the amount of vertical current in the system.	$j_z = \frac{a}{\mu_0} \sum_{i=1}^{N} \left\| \left(\frac{\partial B_y}{\partial B_x} - \frac{\partial B_x}{\partial B_y} \right)_i \right\|$ (2.13)

Table 2.4: Mathematical definitions for energy and vertical current based different predictors

Predictor remarks	Equation
Magnetic helicity density is by definition the dot product of magnetic field and a vector potential $\vec{A}$ for which $\vec{\nabla} \times \vec{A} = \vec{B}$.	$h_B = \vec{A} \cdot \vec{B}$ (2.14)
Magnetic helicity is by definition the volume integral of magnetic helicity density.	$H_B = \int_V \vec{A} \cdot \vec{B}\, dV$ (2.15)
Relative magnetic helicity is often measured with reference to a potential field. Here $\vec{A} = \vec{\nabla} \times \vec{B}$ and $\vec{A_{pot}} = \vec{\nabla} \times \vec{B_{pot}}$.	$H_B^{rel} = \int_V (\vec{A} + \vec{A_{pot}}) \cdot (\vec{B} - \vec{B_{pot}})\, dV$ (2.16)
Current helicity density is by definition the dot product of magnetic field and current density. Here $\vec{\nabla} \times (\vec{B}/\mu_0) = \vec{j}$	$h_C = \frac{1}{\mu_0} \vec{B} \cdot \vec{j}$ (2.17)
Current helicity is by definition the volume integral of current helicity density	$H_C = \frac{1}{\mu_0} \int_V \vec{B} \cdot \vec{j}\, dV$ (2.18)
Total unsigned current helicity (vertical component) measures the overall amount of twist in current regardless of sign	$H_C^{unsigned} = a \sum_{i=1}^{N} \lvert B_z \cdot j_z \rvert$ (2.19)
Net current helicity (vertical component) measures the overall amount of twist	$H_C^{tot} = a \sum_{i=1}^{N} B_z \cdot j_z$ (2.20)

Table 2.5: Mathematical definitions for different magnetic and current helicity-based predictors

Predictor remarks	Equation
Given 2D magnetogram map $B_z(x,y)$ and scale s, $W(x,y,s)$ denotes the wavelet transform using a wavelet function $\psi_s(x,y)$. The wavelet power spectrum $P(s) \propto s^{\delta}$, over an inertial range of scales, where δ is the Wavelet power spectral index.	$W(x,y,s) = B(x,y) * \psi_s(x,y)$ $P(s) = \lvert W(x,y,s)\rvert^2$ $P(s) \propto s^{\delta}$ (2.21)
$\tilde{B}_z(x,y)$ is the Fourier transform of the 2D magnetogram data $B_z(x,y)$. $P(k_x,k_y)$ is the power spectrum of $\tilde{B}_z(x,y)$, reduced to 1-D spectrum $P(k)$ through azimuthal averaging; $k = \sqrt{k_x^2 + k_y^2}$. Over an inertial range of scales $P(k) \propto k^{-\gamma}$, where γ is the Fourier power spectral index.	$\tilde{B}_z(x,y) = \int\int B(x,y) e^{-i(k_x.x + k_y.y)} dxdy$ $P(k_x,k_y) = \lvert \tilde{B}_z(x,y)\rvert^2$ $P(k) \propto k^{-\gamma}$ (2.22)
Fractal Dimension D_f is a function of spatial scale s. Higher D_f indicates more complexity and fragmentation. $N(s)$ indicates the number of boxes that contain the feature/predictor above a given threshold.	$D_f = lim_{s \to 0} \frac{logN(s)}{log(1/s)}$ (2.23)
The multi-fractal structure function of order ξ measures how much the magnetic field changes between two points separated by a distance l after the change is raised to the power ξ and spatially averaged.	$S_{\xi}(l) = \langle \lvert B_z(x+l_x, y+l_y) - B_z(x,y)\rvert^{\xi} \rangle$ (2.24)

Table 2.6: Mathematical definitions for some of the most commonly used multiscale predictors

2.4.3 Predictors not directly linked to the Magnetic Field

We group the predictors not directly based on the magnetic field into two different categories, namely - (i) radiation-based and (ii) helioseismic predictors.

1. Radiation-based predictors

These predictors are based on the observed coronal emission across the electromagnetic spectrum, primarily from GOES X-ray flux measurements. Such predictors are important because flares exhibit persistence i.e past flare behaviour is a strong predictor of future flares (Zirin & Marquette 1991; Park et al. 2020) and since the GOES data provides the temporal flare history for ARs, this may drop hints on the occurrence of future flares. Key examples of radiation-based predictors include (i) previous flare magnitude i.e. GOES maximum class (X/M/C) of prior flares in the AR, (ii) flare duration of prior flares i.e. time interval of flare emission above a threshold, (iii) flare rise and decay times of prior flares indicating energy release rate and efficiency and (iv) cumulative X-ray flux i.e. the total energy output over a period, often used in forecasting models as a memory of prior activity.

Coronal structures observed in EUV and Hα channels, such as coronal loops, filaments, and prominences, provide quantitative non-magnetic predictors of flare activity through metrics like loop twist, filament height, and activation signatures (Schmieder et al. 2009, Régnier et al. 2011). Pre-flare phenomena including coronal dimming and EUV waves, measurable via intensity changes and propagation patterns, serve as additional radiation-based indicators of magnetic restructuring (Reinard & Biesecker 2008, Warmuth 2015). Together, these imaging-based features complement the X-ray flux-based information in forecasting flares and CMEs. Further, EUV observations in high-temperature passbands have revealed hot pre-eruptive magnetic flux ropes in the hours preceding CME onset; Nindos et al. (2020) found that in two-thirds of 30 events

the flux rope was present from 50 minutes to over eight hours before eruption, providing statistical support for CME initiation models involving pre-existing flux ropes.

Spectroscopic flare predictors derived from EUV and soft X-ray observations include line-of-sight Doppler velocities, which trace chromospheric and coronal upflows or downflows preceding flare onset (Brosius & Holman 2007). Non-thermal line broadening is interpreted as evidence of unresolved turbulent motions, wave activity, and magnetic energy release (Milligan 2011), while temperature and density sensitive emission line ratios help identify stressed coronal regions. Observations from Hinode/EIS (Extreme-Ultraviolet Imaging Spectrometer) and IRIS (Interface Region Imaging Spectrograph) have further revealed enhanced non-thermal velocities, precursor brightenings, and non-Gaussian line profiles in flare-producing regions, suggesting complex plasma dynamics and turbulence during the pre-eruptive and impulsive phases of solar flares (Nishizuka et al. 2017, Jeffrey et al. 2018, Panos et al. 2018, 2021). In this context, To et al. (2025) showed that non-thermal velocities at flare footpoints increase 4–25 minutes before GOES soft X-ray onset across C and M-class flares, with CME-associated events exhibiting earlier and more uniform precursor onsets than non-eruptive flares, suggesting a link between extended pre-flare non-thermal broadening and successful eruptions.

2. Helioseismic predictors

Helioseismology is a branch of solar physics that studies the solar oscillations to probe the Sun's interior structure and dynamics. Helioseismic flare predictors use measurements of solar oscillations to detect subsurface flow anomalies and wave-speed perturbations in ARs (Komm 2007). Features such as enhanced subsurface vorticity, divergent flows, or acoustic wave signatures can indicate emerging or stressed magnetic flux prior to flares. The NHGV (Normalised Helicity Gradient Variance) parameter derived from

the GONG (Global Oscillation Network Group) helioseismic data statistically distinguishes flaring from non-flaring ARs and correlates with flare magnitude (Reinard et al. 2010). The NHGV parameter peaks 2–3 days prior to flaring, reflecting an initially large but systematically shrinking spread of subsurface kinetic helicity with depth that converges toward zero at flare onset (Reinard et al. 2010). This suggests that subsurface vortical flows and helicity build-up may play a role in flare initiation and forecasting. These non-magnetic indicators provide early-warning information complementary to photospheric and coronal flare predictors.

2.4.4 Polarity Inversion Lines

Given any AR on the Sun there would be regions of positive magnetic polarity (outward vertical magnetic flux) and negative magnetic polarity (inward magnetic vertical flux). If we start off from the region of positive polarity and move towards the region of negative polarity, at some point on the path, the vertical flux (or the vertical magnetic field) would be zero, before an inversion of polarity occurs. Hence, the locus of neutral points separating the patches of opposite polarities is called a Polarity Inversion Line.

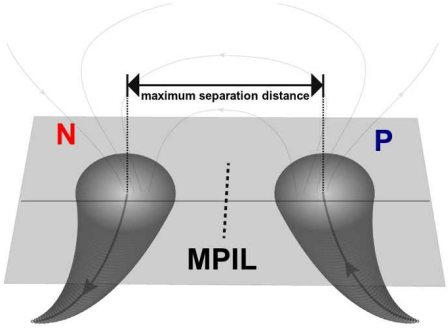

Figure 2.15: Diagrammatic sketch of a PIL, courtesy: Chintzoglou et al. (2019)

PILs were previously referred to as neutral lines and the first explicit mention of neutral lines in the modern solar physics literature appears in the 1960s once routine magnetograms observations were available (Severny 1964). Soon, neutral lines were shown to be closely associated with sunspots, filaments and solar flares (Sturrock & Coppi 1966). From a study of flares in August 1972, Zirin & Tanaka (1973) noted that the flares arose from to strong magnetic shears and high gradients across the neutral lines produced by sunspot motions. As magnetogram resolution and coverage improved in the 1980s and 1990s, studies increasingly emphasized the role of strong magnetic gradients and shear along these lines in driving solar activity and eruptions (Hagyard et al. 1984; Wang et al. 1994; Park et al. 2018). Finally in the early 2000s, PILs were formally quantified and incorporated into statistical and predictive frameworks, where strong-field PIL properties were shown to be effective indicators of flare and coronal mass ejection productivity (Falconer et al. 2002; Falconer et al. 2003; Schrijver 2007;

Falconer et al. 2008; Mason & Hoeksema 2010). After studying the central PIL in several ARs, Falconer et al. (2002) and Falconer et al. (2003) concluded that (i) the strong shear length parameter L_{SS} (with PIL segments where $B_h > 150$ G and shear angle $\theta_{sh} > 45°$) and (ii) the strong gradient length parameter L_{SG} (with PIL segments where $B_h > 150$ G and horizontal gradient of the vertical field $B_{los} > 50$ G/Mm), respectively, are well correlated to CME productivity with a $\pm$ 48 hr time window. However, the thresholds that Falconer et al. (2002, 2003) arrived at, breaching which CMEs are to be expected, were different for each case (62 Mm for L_{SS} and 36 Mm for L_{SG}). Similar to L_{SG}, Mason & Hoeksema (2010) introduced the $GWILL$, a PIL-based metric as a predictor for solar flares, also governed by the horizontal gradient of the vertical field. This reinforces the idea that the control parameters, which effectively capture configurational aspects of the magnetic field distribution around, determine the PIL-based metric/predictor.

Thus, how PILs are determined (based on algorithms) and defined (based on control parameters) is central to the idea of developing flare predictors. In this regard, a number of algorithms are available in the scientific literature. One of the most widely used algorithms is to use magnetic field threshold-based PIL masks to highlight strong magnetic flux patches of opposite polarities and then subsequently convolution-based filtering techniques in order to detect PIL(s) separating them (Schrijver 2007; Wang et al. 2019; Sun et al. 2021; Ji et al. 2023). Mason & Hoeksema (2010) followed a different approach where the input magnetogram (radial field map) is smoothed and a vector potential field is computed (Alissandrakis 1981). Using the vector potential field configuration, neutral regions are identified as preliminary guesses and then these guesses are filtered out, subject to control parameters such as thresholds on the horizontal gradient of the vertical field to identify high-shear regions. Sadykov & Kosovichev (2017) implemented a computed vision based image segmentation algorithm to detect PILs and this is a good example of how machine learning concepts are becoming increasingly relevant for PIL detection.

2.4.5 Unsigned Magnetic Flux & R-value Farameter

By definition, magnetic flux is defined as the integral of the magnetic field over a surface. However, rather than the (signed) magnetic flux, the unsigned magnetic flux is more relevant as a predictor in solar physics because both the positive/negative polarity regions contribute to the stored magnetic energy. Thus, magnetic energy storage is not sensitive to polarity balance. The unsigned magnetic flux has been a fundamental magnetic measure of the magnetic field since routine magnetogram observations became available in the late 20th century. By the early 2000s, the unsigned magnetic flux was routinely included in statistical flare/CME forecasting frameworks alongside other parameters such as neutral line length or free energy proxies (Georgoulis & Rust 2007; Mason & Hoeksema 2010; Hazra et al. 2020).

Compared to the unsigned magnetic flux across the full magnetogram, the unsigned magnetic flux measured near PILs is generally a better predictor of solar flares and eruptions because it preferentially samples regions of strong magnetic gradients, high shear and high free energy (e.g. Georgoulis et al. 2021 have computed unsigned magnetic flux within 2 Mm distance from PILs). Further restricting the computation of unsigned flux to high gradient PILs (i.e. high value of $\nabla_h B_z$) could yield better results for solar flare forecasting (Schrijver 2007; Mason & Hoeksema 2010). In this context, Schrijver (2007) devised the R-value parameter, which assures a 100% probability of major flare occurrence within 24 hr once it reaches a value close to 2×10^{21} Mx. For the calculation of R-value, PILs are first identified where opposite polarities lie close together above a threshold field strength ($|B_z| > 150$ G), and subsequently the unsigned flux is then summed within a fixed separation distance ($D_{sep} > 15$ Mm) of these strong gradient PILs. Essentially, the R-value is a filtered version of the unsigned magnetic flux, with its own unique algorithm.

2.4.6 Sunspots

Sunspots are the earliest studied morphological predictors. It is well established that sunspot number is a useful tracer of the solar magnetic activity. While Schwabe's law (Schwabe 1843) describes the approximately 11-year solar cycle of sunspot number (and average solar area fraction), Hale's law (Hale et al. 1919; Hale & Nicholson 1925) states that sunspots emerge in bipolar pairs with a consistent orientation of magnetic polarity that reverses between successive solar cycles (see Figures 2.16 & 2.17).

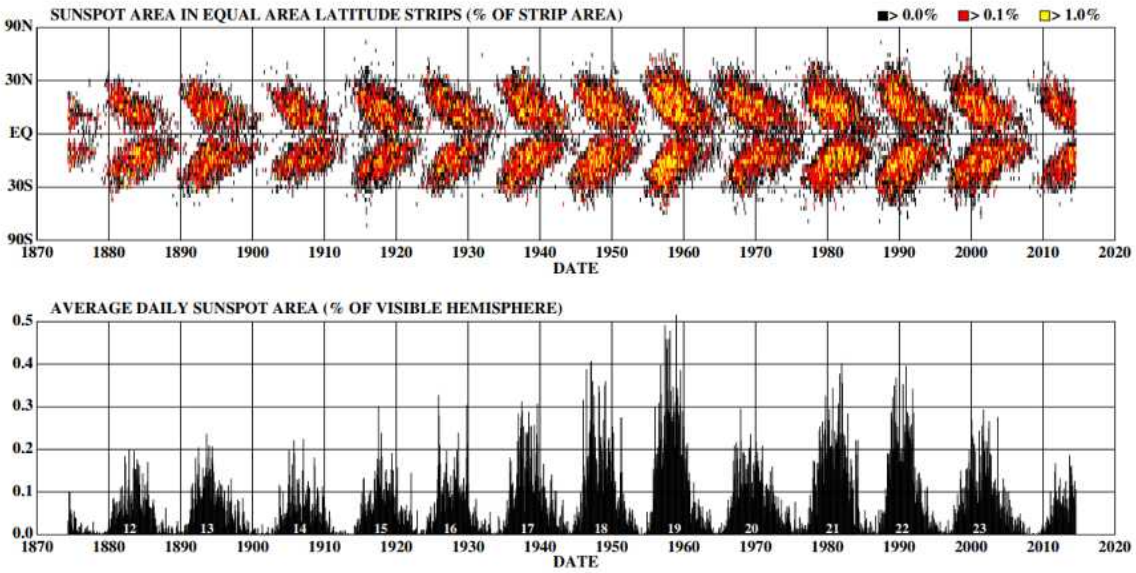

Figure 2.16: The butterfly diagram adapted from Hathaway (2015) shows how sunspots appear at mid-latitudes at the start of each solar cycle and migrate towards the equator over time; this behaviour is periodic and follows an 11 year cycle, consistent with the periodicity of the average sunspot area as a fraction of the hemisphere.

It is observed that large, complex sunspot groups were more frequently associated with eruptive solar flares (Künzel 1959; Sammis et al. 2000). In this regard, there exist sunspot classification schemes such as McIntosh classification (McIntosh 1990) and Mount Wilson classification (Hale et al. 1919; Künzel 1965) based on complexity. These classification schemes link flare productivity to properties such as sunspot area,

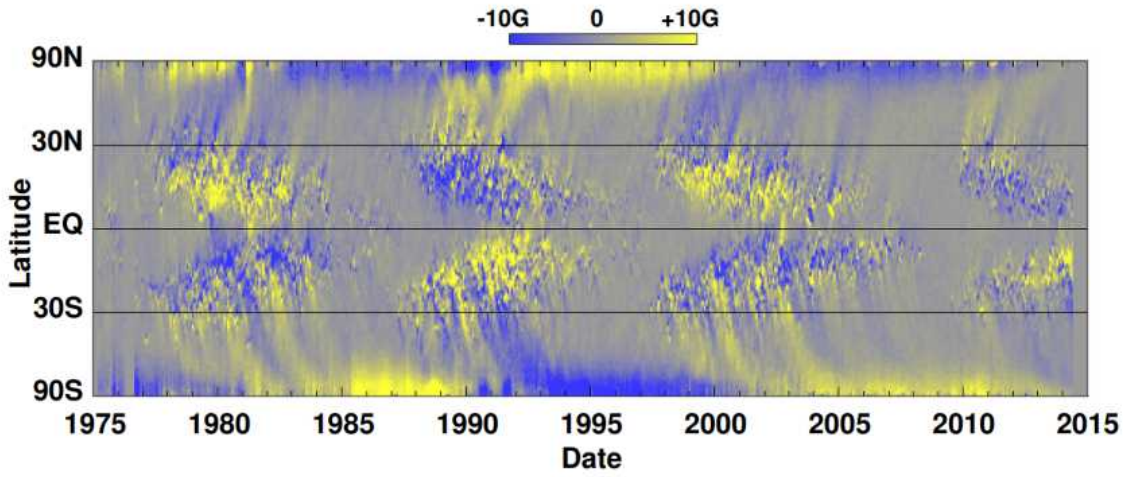

Figure 2.17: The magnetic butterfly diagram adapted from Hathaway (2015) shows the reversal of magnetic polarity every solar cycle

magnetic complexity, and the presence of closely spaced opposite polarities. Although sunspot-based predictors are inherently subjective and limited in their ability to capture magnetic non-potentiality, they laid the foundation for quantitative approaches and are still relevant today as simple, physically intuitive indicators of solar activity (Korsos 2018).

Mount Wilson Classification Scheme

The Mount Wilson sunspot classification scheme was developed at the the Mount Wilson Observatory in the early 20th century by George Hale and colleagues and is sometimes also called the Hale sunspot classification scheme in the research literature. It is a magnetic classification system that categorises solar ARs according to the photospheric magnetic polarity configuration of their sunspots. This scheme classifies ARs based on the number of magnetic polarities present, their arrangement, and the degree to which opposite polarities are intermixed. The underlying assumption is that as the complexity of polarity configurations increases, the sunspot becomes associated with stronger electric currents, greater non-potentiality, and a higher likelihood of magnetic

reconnection, all of which contribute to flare productivity (Sammis et al. 2000; Toriumi & Wang 2019). The basic Mount Wilson classes are:

1. α-type: This class represents unipolar sunspots with a single magnetic polarity (positive or negative). These regions are magnetically simple and generally associated with low flare productivity.

2. β-type: This class represents sunspot groups with a bipolar morphology i.e. clearly separated positive and negative polarities. This is typical of many emerging ARs.

3. γ-type: A γ-type sunspot-group is a complex multi-polar region in which the magnetic polarities are irregularly distributed and cannot be separated into a simple bipolar configuration, indicating significant magnetic complexity.

4. δ-type: This is a special designation applied when umbrae of opposite magnetic polarities are contained within a single penumbra. δ-type sunspot configurations indicate strong polarity mixing on small spatial scales and are strongly associated with major solar flares (GOES X/M-class flares). This sunspot group if present, indicates the presence of high magnetic shear and strong gradients present in ARs (Sammis et al. 2000).

In practice, these classes are often combined to describe intermediate or evolving configurations, such as $\beta\gamma$, $\beta\delta$, or $\beta\gamma\delta$, reflecting increasing degrees of complexity. ARs commonly evolve through multiple Mount Wilson classes over their lifetimes. For example, a newly emerging bipolar region may initially be classified as β, later developing additional flux emergence or shear that leads to a $\beta\gamma$ or $\beta\gamma\delta$ configuration. Conversely, decaying ARs may become simpler over time as magnetic flux disperses. This temporal variability underscores the fact that the Mount Wilson classification scheme is a snapshot description rather than a fixed property of an AR.

McIntosh Classification Scheme

The McIntosh sunspot classification scheme was introduced by Patrick McIntosh in 1990 as an extension of the original Zurich classification, aimed at providing a compact yet detailed description of sunspot group structure and its evolutionary state. Owing to its operational usefulness, the McIntosh classification has been widely adopted in space weather forecasting and statistical studies of solar activity (Gallagher et al. 2002; Bloomfield et al. 2012; McCloskey et al. 2018). The McIntosh classification builds on the Zurich system of sunspot classification (Kiepenheuer 1953) by recognising that sunspot group complexity cannot be fully characterised or explained by a single evolutionary label. Instead, it decomposes group morphology into three independent components (each represented by a character) that describe three different attributes of an AR: (i) the Zurich (Z) component, describing the overall evolutionary class of the group, (ii) the Penumbral (p) component, describing the penumbral development of the largest sunspot, and (iii) the Compactness (c) component, describing the spatial distribution of sunspots within the group. Further details of different types in each McIntosh component can be found in Table 2.7.

Component	Type	Description
Zurich (Z)	A	very small or unipolar group with no penumbra and no clear bipolar structure
	B	bipolar group with opposite polarities but no penumbrae on either spot
	C	bipolar group with a penumbra on one of the principal spots
	D	mature bipolar group with penumbrae on both main spots and moderate longitudinal separation
	E	large bipolar group with penumbrae on both main spots and large longitudinal separation
	F	very large developed bipolar group with penumbrae on both main spots and large longitudinal extent
	H	unipolar group consisting of a single mature sunspot with a penumbra, typically in decay
Penumbral (p)	x	no penumbra present
	r	rudimentary or incomplete penumbra
	s	small, symmetric penumbra
	a	small, asymmetric penumbra
	h	large, symmetric penumbra
	k	large, asymmetric penumbra
Compactness (c)	x	no interior spots
	o	open layout, few/no spots between main polarities
	i	intermediate layout, some spots between main polarities
	c	compact layout, many spots between main polarities

Table 2.7: Details of McIntosh sunspot classification scheme with classes and sub-classes

2.5 Solar Magnetic Field Measurements

The photospheric magnetic field data has been obtained routinely from spectropolarimetric observations obtained by space-based instruments such as SOHO-MDI (Scherrer et al. 1995) or SDO-HMI (Scherrer et al. 2012). These measurements rely on the Zeeman effect, in which the presence of a magnetic field causes a spectral line of rest wavelength λ_0 to split into its polarised components (Stenflo 2013). Under a weak field approximation (< 1000 G), given the effective Landé factor g_{eff}, λ_0 (in Å), and the magnetic field strength B (in Gauss), the wavelength separation between the Zeeman components i.e $\Delta\lambda_B$ (in Å) is (see Stenflo 2013):

$$\Delta\lambda_B = 4.67 \times 10^{-13} g_{eff} B \lambda_0^2 \tag{2.25}$$

The Zeeman effect is the primary physical basis for magnetic field diagnostics in photospheric spectropolarimetry - it induces polarisation signals in spectral lines that can be inverted to produce magnetograms (Solanki 1993; Del Toro Iniesta & Ruiz Cobo 1996). Although the wavelength separation given by Equation 2.25 provides the fundamental physical scale of the Zeeman effect, in most photospheric observations, the splitting is much smaller than the intrinsic width of the spectral line and cannot be resolved directly. Instead, the presence of a magnetic field manifests itself primarily through changes in the polarisation state of the emergent radiation. It is this polarisation, rather than a measurable line splitting, that is detected by modern spectropolarimeters (Solanki 1993; Del Toro Iniesta & Ruiz Cobo 1996; Reiners 2012; Stenflo 2013). Consequently, magnetic field information must be inferred by analysing the polarised components of the spectral line profile, linking the underlying Zeeman splitting to observable polarisation signals (Solanki 1993).

The polarisation state of the emitted or absorbed radiation is fully described by the four Stokes parameters (I, Q, U, V). The Stokes parameter I represents the to-

tal intensity (Solanki 1993). While Q and U measure the linear polarisation, the Stokes parameter V measures the circular polarisation (Solanki 1993). In the weak-field approximation ($\Delta\lambda_B << \Delta\lambda_D$, where $\Delta\lambda_D$ is the thermal Doppler width - the natural non-magnetic broadening of a spectral line caused by the motion of atoms or molecules along the line of sight, usually due to thermal motion or turbulence in the gas), commonly applicable to photospheric observations, the circular polarisation signal is directly related to the LOS magnetic field component as (see Stenflo 2013):

$$V(\lambda) \simeq \left[-4.67 \times 10^{-13} g_{eff} \lambda_0^2 \left(\frac{dI(\lambda)}{d\lambda}\right)\right] B_{LOS} \tag{2.26}$$

The factor $dI(\lambda)/d\lambda$ in Equation 2.26 represents the wavelength derivative of the emergent intensity profile, which is itself the solution to the polarised radiative transfer equation through the magnetised solar atmosphere. In general, the emergent Stokes profiles depend on the temperature, pressure, velocity, and magnetic field stratification along the line of sight, and their accurate computation requires solving the full set of polarised radiative transfer equations. A full treatment of polarised radiative transfer (via. spectral line inversion codes) is beyond the scope of this thesis (refer to Ruiz Cobo & del Toro Iniesta (1992, 1994) and Degl'innocenti & Landi (2004) for a comprehensive account).

The transverse component of the magnetic field is obtained from the linear polarisation Stokes parameters Q and U, with the field azimuth θ_{stokes} in the plane of the sky given by:

$$\theta_{stokes} = \frac{1}{2} \arctan\left(\frac{U}{Q}\right) \tag{2.27}$$

and the transverse field strength $B_\perp$ is proportional to $\sqrt{Q^2 + U^2}$.

2.6 Extrapolation Models

Although we can only routinely measure the solar magnetic field at the photosphere, most of the physical processes concerning flares and CMEs take place in the corona. As a result, to study coronal phenomena, one needs to employ mathematical models to extrapolate the photospheric magnetic configuration to coronal heights. This procedure essentially converts a two-dimensional magnetic field configuration into a three-dimensional magnetic field data-grid. In this thesis, we obtain a three dimensional data-grid for several ARs using extrapolation models to study (i) PILs, (ii) the unsigned magnetic flux and (iii) the R-value parameter, even though these predictors are traditionally studied on the photosphere. In this section, we specifically discuss the assumptions/approximations behind three different types of extrapolation models: (i) NLFFF (Non-Linear Force-Free Field) Extrapolation, (ii) LFFF (Linear Force-Free Field) Extrapolation and (iii) PF (Potential Field) Extrapolation.

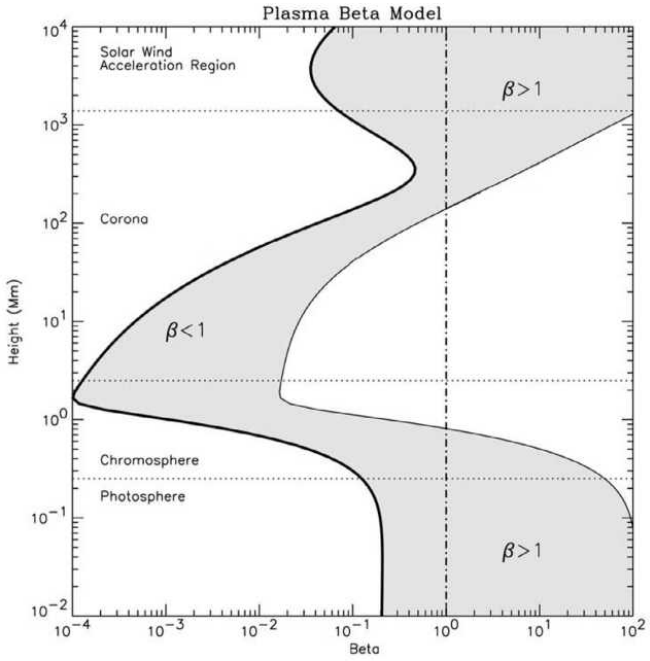

Figure 2.18: Height-wise variation of plasma parameter β in ARs, adapted from Gary (2001)

In the solar ARs, the plasma pressure and the overall magnetic field strength vary as a function of height and thus, the plasma parameter β (ratio between plasma pressure and magnetic pressure) also varies with height above the photosphere (see Figure 2.18). While in the quiet Sun, the height above the base of the photosphere at which magnetic forces start to dominate over other non-magnetic forces (i.e. $\beta << 1$) is found to vary between 0.8-1.6 Mm (Rosenthal et al. 2002; Wiegelmann et al. 2014), however, in case of ARs, it has been determined that the minimum height for the condition $\beta << 1$ is about 0.4 Mm (Metcalf et al. 1995; Wiegelmann et al. 2014). In this thesis, we explore the magnetic field configuration up to a height of 3.24 Mm (see Section 2.6.3) above the photosphere (lower solar atmosphere) using a step size of 0.36 Mm (0.5 arcsec) and thus the $\beta << 1$ approximation shall be valid for a large part of the data. This approximation is quite helpful to simplify the governing MHD equations.

2.6.1 Framework for NLFFF Extrapolation

In the MHD momentum equation for solar plasmas, given mass density ρ, magnetic field $\vec{B}$ and current density $\vec{j}$, the balance (or imbalance) between magnetic pressure gradient force ($\vec{\nabla} p$), Lorentz force ($\vec{j} \times \vec{B}$) and gravitational force ($\rho \vec{g}$) is written as follows:

$$\rho \left(\frac{d\vec{v}}{dt} \right) = -\vec{\nabla} p + \vec{j} \times \vec{B} + \rho \vec{g}. \tag{2.28}$$

Since $\beta << 1$, the plasma pressure gradient force can be neglected. Also particularly in ARs, the gravitational force is negligible compared to the Lorentz force (Priest 2014). Hence the MHD momentum equation reduces to:

$$\rho \left(\frac{d\vec{v}}{dt} \right) = \vec{j} \times \vec{B}. \tag{2.29}$$

For a static configuration, this equation further reduces to something called the Force-Free Field (FFF) approximation' - $\vec{j} \times \vec{B} = 0$. The FFF approximation implies that the current must flow parallel to the magnetic field. Thus the current density must be related to the magnetic field by a scalar factor $\alpha(\vec{r})$ varying in space.

$$\vec{j} = \alpha(\vec{r}).\vec{B} \tag{2.30}$$

As per Ampere's circuital law, the curl of the magnetic field is proportional to the current density.

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{j}. \tag{2.31}$$

Substituting the value of $\vec{j}$ in Equation 2.31, we have:

$$\vec{\nabla} \times \vec{B} = \mu_0 \alpha(\vec{r}).\vec{B} \tag{2.32}$$

Using the vector identity $\vec{\nabla} \cdot (\vec{\nabla} \times \vec{B}) = 0$ in Equation 2.32, we obtain:

$$\vec{\nabla} \cdot (\alpha(\vec{r}).\vec{B}) = 0 \tag{2.33}$$

Expanding the left-hand side gives:

$$\vec{\nabla} \cdot (\alpha(\vec{r}).\vec{B}) = \vec{B} \cdot \vec{\nabla}\alpha(\vec{r}) + \alpha(\vec{r})(\vec{\nabla} \cdot \vec{B}) \tag{2.34}$$

Another important equation is classical electrodynamic theory is the Gauss's law for magnetism which states that the net magnetic flux through any closed surface is zero. This can be expressed in a differential equation as $\vec{\nabla}.\vec{B} = 0$. Therefore:

$$\vec{B} \cdot \vec{\nabla}\alpha(\vec{r}) = 0 \tag{2.35}$$

Equation 2.30 describes the Non-Linear Force-Free Field (NLFFF) model for extrapolation. The non-linearity arises because the force-free parameter α varies spatially and is constrained to be constant along magnetic field lines (see Equation 2.35), making it an implicit function of the magnetic field itself. In case of a NLFFF extrapolation model, $\alpha(\vec{r})$ has to be computed iteratively while the extrapolation is in progress. This makes NLFFF extrapolation more computationally expensive than LFFF extrapolation (Korsós et al. 2024). Wiegelmann & Sakurai (2021) give a summary of the different approaches/procedures to compute the solar magnetic fields by using LFFF/NLFFF models. There exist analytical equilibrium models such as Low and Lou's equilibrium (Low & Lou 1990) and Titov-Démoulin equilibrium (Titov & Démoulin 1999), which are used for testing and benchmarking NLFFF extrapolation techniques. As an example, we briefly discuss a solution proposed by Low & Lou (1990) which uses simplifying assumptions such as axisymmetry. They introduced a flux function $A(r, \theta)$ that labels magnetic field lines and further added a twist function $Q(A)$ that tells how much the

field twists around the axis. The magnetic field $\vec{B}$, can be expressed in terms of flux function A and $Q(A)$ (ϕ component of the field B) as:

$$\vec{B} = \frac{1}{rsin\theta}\left(\frac{1}{r}\frac{\partial A}{\partial \theta}\hat{r} - \frac{\partial A}{\partial r}\hat{\theta} + Q\hat{\phi}\right) \tag{2.36}$$

The benefit of expressing $\vec{B}$ as shown above is that the solenoidal condition is automatically satisfied. In this case, Equation 2.32 reduces to the Grad-Shafranov equation, in spherical coordinates:

$$\frac{\partial^2 A}{\partial r^2} + \left(\frac{1-\eta^2}{r^2}\right)\frac{\partial^2 A}{\partial \eta^2} + Q\frac{dQ}{dA} = 0; \qquad \eta = cos\theta \tag{2.37}$$

The solution (A, Q) obtained by solving Equation 2.37 can be plugged into Equation 2.36 to obtain $\vec{B}$. In the solution (A, Q) shown below, n and λ are constants and serve as control parameters for radial decay and twist of the magnetic field, respectively.

$$Q(A) = \lambda A^{1+\frac{1}{n}}; \qquad A(r,\theta) = \frac{P(\eta)}{r^n} \tag{2.38}$$

2.6.2 Frameowrk for LFFF Extrapolation

The LFFF extrapolation model can be considered as a special case of the NLFFF extrapolation model in which the coefficient $\alpha(\vec{r})$ is a constant (α) and does not vary with space. Usually α is determined from observations before actual extrapolation is carried out (Wiegelmann & Sakurai 2021). The Helmholtz equation, which is a linear partial differential equation, is critical to obtain the LFFF extrapolation solution. To derive it, first, we take the curl of the force-free equation:

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{B}) = \vec{\nabla} \times (\alpha \vec{B}) = \alpha \vec{\nabla} \times \vec{B} = \alpha^2 \vec{B} \tag{2.39}$$

Using the vector identity $\vec{\nabla} \times (\vec{\nabla} \times \vec{B}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{B}) - \nabla^2 \vec{B}$ and applying the solenoidal condition, we obtain:

$$\nabla^2 \vec{B} + \alpha^2 \vec{B} = 0. \tag{2.40}$$

Seehafer (1978) derived analytical solutions for LFFF in Cartesian geometry by solving the Helmholtz equation using a Green's function approach, and line-tied boundary conditions at the photosphere. The magnetic field components were expressed as Fourier series that satisfy both the solenoidal condition and the force-free constraint, with an exponential decay of field strength with height. Building on this formulation, Alissandrakis (1981) implemented an efficient computational method based on fast Fourier transforms to perform magnetic field extrapolations from observed photospheric magnetograms. In the solution, the magnetic field $\vec{B} = (B_x, B_y, B_z)$ is expressed in Cartesian coordinates (x, y, z), within a rectangular photospheric domain of size (L_x, L_y) and z denotes the height above the photosphere (see Equation 2.41). The rectangular photospheric domain relates to the magnetogram data, which is usually the line-of-sight component or the radial component of the photospheric vector magnetic field.

$$B_x = \sum_{m,n=1}^{\infty} \gamma_{mn} \left[\alpha \frac{n\pi}{L_y} sin\left(\frac{m\pi x}{L_x}\right) cos\left(\frac{n\pi y}{L_y}\right) - r_{mn} \frac{m\pi}{L_x} cos\left(\frac{m\pi x}{L_x}\right) sin\left(\frac{n\pi y}{L_y}\right) \right]$$

$$B_y = - \sum_{m,n=1}^{\infty} \gamma_{mn} \left[\alpha \frac{m\pi}{L_x} cos\left(\frac{m\pi x}{L_x}\right) sin\left(\frac{n\pi y}{L_y}\right) + r_{mn} \frac{n\pi}{L_y} sin\left(\frac{m\pi x}{L_x}\right) cos\left(\frac{n\pi y}{L_y}\right) \right]$$

$$B_z = \sum_{m,n=1}^{\infty} \gamma_{mn} \lambda_{mn} . sin\left(\frac{m\pi x}{L_x}\right) sin\left(\frac{n\pi y}{L_y}\right)$$

$$\gamma_{mn} = \frac{C_{mn}}{\lambda_{mn}} e^{-r_{mn} z}; \quad \lambda_{mn} = \pi^2 \left(\frac{m^2}{{L_x}^2} + \frac{n^2}{{L_y}^2} \right); \quad r_{mn} = \sqrt{\lambda_{mn} - \alpha^2} \tag{2.41}$$

The parameter α is the constant force-free parameter. The integers (m,n) label the Fourier modes in the horizontal directions. The eigenvalue λ_{mn} represents the squared horizontal wavenumber and the coefficients C_{mn} are obtained from the Fourier decomposition of the photospheric vertical magnetic field. While r_{mn} controls the exponential decay of each mode with height, γ_{mn} incorporates both the boundary amplitudes and their vertical dependence.

2.6.3 Framework for PF Extrapolation

The PF extrapolation model is a special case of the LFFF extrapolation model where the force free parameter $\alpha = 0$, effectively so because in a PF, the system has no current ($\vec{j} = 0$). This implies that $\vec{\nabla} \times \vec{B} = 0$. Hence we can express $\vec{B}$ as the gradient of a scalar potential Φ, given by $\vec{B} = -\vec{\nabla}\Phi$. Using the solenoidal condition for the magnetic field we obtain:

$$\nabla^2 \Phi = 0. \tag{2.42}$$

Thus, we arrive at the classic Laplace's equation as a solution for the magnetic field. Usually the line-of-sight/radial-field magnetogram data serve as boundary conditions to solve the Laplace's equation described above. An analytical solution to the Laplace equation is provided by the Green's function approach of Schmidt (1964), who showed that the scalar potential satisfying the current-free condition is uniquely determined by the vertical magnetic field component on the photospheric boundary alone. A homogenised form of Schmidt's analytical expressions is presented in Georgoulis & LaBonte (2007). With today's computational power these integral formulae are entirely tractable. However the FFT approach of (Alissandrakis 1981) remains comparatively faster.

The main advantage of the PF extrapolation method over LFFF/NLFFF is that it is computationally cheap and serves as the lowest-order approximation for modelling the coronal magnetic field. The research work shared in this thesis has been carried out using the PF extrapolation (Fast Fourier Method - Alissandrakis (1981)) up to a height of 3.24 Mm to analyse the magnetic field configuration in ARs. The upper boundary of 3.24 Mm was selected to focus on the low-lying magnetic structures associated with the photosphere and the chromosphere, where the extrapolated magnetic field is expected to reliably represent the local magnetic topology.

Chapter 3

Mathematical Techniques

This chapter is dedicated to the mathematical techniques used in the research presented in this thesis. There are two major mathematical areas that we address in this chapter.

1. An autoregressive model is a time-series model where the current value of a variable is expressed as a linear combination of its own past values, plus noise (Hamilton 1994; Lutkepohl 2005). Autoregressive models are widely used to model systems that have a certain memory dependence and its application spans economics, signal processing, geophysics, astrophysics, and machine learning. Chapter 6 focuses on the time series analysis of PIL lengths using autoregressive models and so, we discuss the mathematical background behind autoregressive models in the current chapter.

2. In Chapter 7, there is substantial discussion on using PIL length based thresholds to separate out ARs producing strong/weak/no flares. In this connection, we have used the one-sided Chebyshev inequality to make inferences on the distribution of strongly/weakly/non-flaring AR populations. Hence, the derivation (and assisting theory) of the Chebyshev inequality is provided in this chapter.

3.1 Autoregressive Models

3.1.1 Definition

Let us consider an autoregressive model $A(p)$, where p is the lag order. The current term Y_t can be expressed as a linear combination of its previous values Y_{t-i}, weighted by coefficients ψ_i, and a white noise error term ϵ_t, where $\epsilon_t \sim \text{Gaussian}(0, \sigma^2)$.

$$Y_t = \sum_{i=1}^{p} \psi_i Y_{t-i} + \epsilon_t \tag{3.1}$$

Let Ω denote the back-shift (lag) operator, defined by

$$\Omega Y_t = Y_{t-1}, \quad \Omega^2 Y_t = Y_{t-2}, \quad \dots \tag{3.2}$$

Using the back-shift operator, the autoregressive model in Equation (3.1) may be rewritten solely in terms of Y_t as

$$Y_t = \psi_1 \Omega Y_t + \psi_2 \Omega^2 Y_t + \cdots + \psi_p \Omega^p Y_t + \epsilon_t. \tag{3.3}$$

Rearranging terms yields the lag-operator form

$$\Psi(\Omega) Y_t = \epsilon_t, \tag{3.4}$$

where the characteristic lag operator polynomial is given by

$$\Psi(\Omega) = 1 - \psi_1 \Omega - \psi_2 \Omega^2 - \cdots - \psi_p \Omega^p. \tag{3.5}$$

3.1.2 Condition for Stationarity

To analyse stationarity, the corresponding characteristic equation is obtained by replacing the operator Ω with a complex variable ω:

$$\Psi(\omega) = 1 - \psi_1\omega - \psi_2\omega^2 - \cdots - \psi_p\omega^p = 0. \tag{3.6}$$

The autoregressive process is (weakly) stationary if and only if all p roots ω_i of the characteristic equation satisfy $|\omega_i| > 1$. If at least one root satisfies $|\omega_i| \leq 1$, the process is non-stationary. In the special case where $\omega = 1$ is a root (a unit root), the characteristic polynomial satisfies

$$\Psi(1) = 1 - \psi_1 - \psi_2 - \cdots - \psi_p = 0. \tag{3.7}$$

To illustrate why roots lying on or inside the unit circle imply non-stationarity, consider the simplest case of an autoregressive model of order one:

$$Y_t = \psi Y_{t-1} + \epsilon_t. \tag{3.8}$$

The characteristic equation associated with this model is $1 - \psi\omega = 0$, which has the root $\omega = 1/\psi$. By recursive substitution, Y_t may be expressed in terms of the initial value Y_0 as

$$Y_t = \psi^t Y_0 + \sum_{i=1}^{t} \psi^{t-i}\epsilon_i. \tag{3.9}$$

Taking expectations and using $E[\epsilon_i] = 0$, we obtain

$$E[Y_t] = \psi^t E[Y_0]. \tag{3.10}$$

The variance of Y_t is given by

$$\mathrm{Var}(Y_t) = \psi^{2t}\mathrm{Var}(Y_0) + \sigma^2 \sum_{i=0}^{t-1} \psi^{2i}. \tag{3.11}$$

Three cases arise:

Case 1: $|\psi| < 1$**.** As $t \to \infty$,

$$\lim_{t\to\infty} E[Y_t] = 0, \quad \lim_{t\to\infty} \mathrm{Var}(Y_t) = \frac{\sigma^2}{1-\psi^2}.$$

Both the mean and variance converge to finite constants, and the process is stationary.

Case 2: $|\psi| > 1$**.** Both the mean and variance grow exponentially with time, and the process is explosively non-stationary.

Case 3: $|\psi| = 1$**.** When $\psi = 1$, the process reduces to

$$Y_t = Y_{t-1} + \epsilon_t, \tag{3.12}$$

which describes a pure random walk. In this case,

$$E[Y_t] = E[Y_0], \quad \mathrm{Var}(Y_t) = \mathrm{Var}(Y_0) + t\sigma^2,$$

so the variance grows linearly over time and the process is non-stationary.

Thus, for an autoregressive process to be stationary, all roots ω_i of the characteristic equation must lie strictly outside the unit circle.

3.1.3 The ADF Regression Equation

To check for the presence of a unit root (or non-stationarity, in a stochastic sense) for any given time series, we use the ADF (Augmented Dickey Fuller) test (Dickey & Fuller 1979, Said & Dickey 1984, Elliott et al. 1996; see Giannattasio et al. 2019, Giannattasio, F. et al. 2022 for applications). The ADF test is a regression involving the lagged difference term ΔY_t with the previous term Y_{t-1} and previous lagged differences ΔY_{t-i}. The ADF regression equation is a rearrangement of the equation corresponding to $A(p)$ (see example provided).

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta Y_{t-i} + \epsilon_t \tag{3.13}$$

Here δ_i are the coefficients of the lagged differences and γ is the coefficient of previous term and it controls how much the previous term influences the current lagged difference. The intuition behind this is that when $\gamma < 0$, the series is mean reverting. Originally, the Dickey Fuller (DF) test was devised for an $A(1)$ process, equivalently in the differenced form: $\Delta Y_t = \alpha + \gamma Y_{t-1} + \epsilon_t$. The augmentation of the additional difference terms with coefficients δ_1, δ_2, ... δ_m makes it model more generalised and complex autoregressive processes.

Example

Let us consider an autoregressive model of lag order 2. The equation for Y_t can be simplified as:

$$Y_t = \sum_{i=1}^{p} \psi_i Y_{t-i} + \epsilon_t = \psi_1 Y_{t-1} + \psi_2 Y_{t-2} + \epsilon_t \tag{3.14}$$

Subtracting Y_{t-1} from both sides we have:

$$\Delta Y_t = Y_t - Y_{t-1} = (\psi_1 - 1)Y_{t-1} + \psi_2 Y_{t-2} + \epsilon_t \tag{3.15}$$

On the RHS of the equation, adding and subtracting $\psi_2 Y_{t-1}$ we have:

$$\Delta Y_t = (\psi_1+\psi_2-1)Y_{t-1}+(\psi_2 Y_{t-2}-\psi_2 Y_{t-1})+\epsilon_t = (\psi_1+\psi_2-1)Y_{t-1}-\psi_2\Delta Y_{t-1}+\epsilon_t \tag{3.16}$$

The generalised version of the regression equation reduces to the following for $p = 2$:

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta Y_{t-i} + \epsilon_t = \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \epsilon_t \tag{3.17}$$

Upon comparing the previous Equations 3.16 and 3.17 we find equivalence when:

$$\gamma = \psi_1 + \psi_2 - 1 \quad \& \quad \delta_1 = -\psi_2 \tag{3.18}$$

3.1.4 Implementing the ADF Test

In real-world time series data, the autoregressive component may be present along with other functional trends $f(t)$ (i.e. linear, quadratic, logarithmic, exponential, power-law and sinusoidal). Thus, a more accurate description of the time series would be:

$$Y_t = f(t) + \sum_{i=1}^{p} \psi_i Y_{t-i} + \epsilon_t \tag{3.19}$$

Similarly, the ADF regression equation is also modified accordingly.

$$\Delta Y_t = f(t) + \gamma Y_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta Y_{t-i} + \epsilon_t \tag{3.20}$$

Hence, prior to the application of the ADF test, the data should be detrended using curve fitting techniques if visual inspection suggests the presence of a definite functional trend $f(t)$. This also ensures that ADF test does not model the functional trend as a memory-dependent process. Before carrying out the ADF test an initial guess for the maximum number of lags n_{max} may be manually selected and then optimisation parameters such as the AIC - Akaike Information Criterion (Akaike 1974) and the BIC - Bayesian Information Criterion (Schwarz 1978) may be used to select the best lag value between $\{0,\ 1,\ 2\ ...\ n_{max}\}$. In the ADF test, the null hypothesis states that the series has a unit root (i.e., it is non-stationary), while the alternative hypothesis suggests that the series is stationary. If the test statistic τ_{ADF} is more negative than the critical value (1%, 5% or 10% S_l significance levels), the null hypothesis is rejected and the time series is considered as stationary.

$$\tau_{ADF} = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \tag{3.21}$$

Here $\hat{\gamma}$ is the regression estimate of γ and $SE(\hat{\gamma})$ refers to the standard error in the estimate.

3.2 Chebyshev's Inequality (one-sided)

3.2.1 Statement

In probability theory, as per the one-sided Chebyshev's inequality (or Cantelli's inequality), if X is a random variable (there is no restriction on the distribution of X) with mean μ and standard deviation σ, then for $\lambda > 0$, the following are true (see Lin & Bai 2011):

$$P(X \geq \mu + \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2} \quad \equiv \quad P(X \leq \mu + \lambda) \geq \frac{\lambda^2}{\sigma^2 + \lambda^2} \tag{3.22}$$

$$P(X \leq \mu - \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2} \quad \equiv \quad P(X \geq \mu - \lambda) \geq \frac{\lambda^2}{\sigma^2 + \lambda^2} \tag{3.23}$$

Substituting λ with $\{\sigma, 2\sigma, 3\sigma, 10\sigma\}$, we get the following relations for example:

$$P(X \geq \mu + \sigma), P(X \leq \mu - \sigma) \leq 50\% \quad \equiv \quad P(X \leq \mu + \sigma), P(X \geq \mu - \sigma) \geq 50\%$$

$$P(X \geq \mu + 2\sigma), P(X \leq \mu - 2\sigma) \leq 20\% \quad \equiv \quad P(X \leq \mu + 2\sigma), P(X \geq \mu - 2\sigma) \geq 80\%$$

$$P(X \geq \mu + 3\sigma), P(X \leq \mu - 3\sigma) \leq 10\% \quad \equiv \quad P(X \leq \mu + 3\sigma), P(X \geq \mu - 3\sigma) \geq 90\%$$

$$P(X \geq \mu + 10\sigma), P(X \leq \mu - 10\sigma) \leq 1\% \quad \equiv \quad P(X \leq \mu + 10\sigma), P(X \geq \mu - 10\sigma) \geq 99\%$$

3.2.2 Derivation

Let $Y = X - \mu$ be a random variable. Then $E[Y] = 0$, and $E[Y^2] = \sigma^2$. For any $c > 0$, the random variable $(Y + c)^2$ is non-negative. On the event $\{Y \geq \lambda\}$, where $\lambda > 0$ then $(Y + c)^2 \geq (\lambda + c)^2$. This implies:

$$\frac{(Y + c)^2}{(\lambda + c)^2} \geq 1 = P(Y \geq \lambda)$$

Taking expectations on both sides,

$$P(Y \geq \lambda) \leq \frac{E[(Y + c)^2]}{(\lambda + c)^2}.$$

Expanding the numerator,

$$E[(Y + c)^2] = E[Y^2] + 2cE[Y] + c^2 = \sigma^2 + c^2$$

Thus,

$$P(Y \geq \lambda) \leq \frac{\sigma^2 + c^2}{(\lambda + c)^2}.$$

Minimizing (the second derivative can be shown to be > 0) the right-hand side with respect to $c > 0$, (differentiating w.r.t. c) we have:

$$\begin{aligned}
(\sigma^2 + c^2)'(\lambda + c)^2 &= (\sigma^2 + c^2)[(\lambda + c)^2]' \\
\implies 2c(\lambda + c)^2 &= (\sigma^2 + c^2)[2(\lambda + c)] \\
\implies c(\lambda + c) &= (\sigma^2 + c^2)
\end{aligned}$$

$$\implies c\lambda = \sigma^2$$
$$\implies c = \frac{\sigma^2}{\lambda}.$$

Substituting this value,

$$P(Y \geq \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

Returning to X,

$$P(X \geq \mu + \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

Taking complements i.e. subtracting each side from 1, we have:

$$P(X \leq \mu + \lambda) \geq \frac{\lambda^2}{\sigma^2 + \lambda^2}.$$

Applying the result to $-X$,

$$P(-X \geq -\mu + \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

which is equivalent to

$$P(X \leq \mu - \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

Taking complements again,

$$P(X \geq \mu - \lambda) \geq \frac{\lambda^2}{\sigma^2 + \lambda^2}.$$

Chapter 4

Data Processing and Computational Methods

This chapter is centered around the methods we used to compute the four major predictors of focus in this thesis: (i) L_{tot} - total PIL length, (ii) L_{max} - the maximum PIL length, (iii) ϕ - unsigned flux around PILs and (iv) R-value. The algorithm for each predictor is described in a systematically in a sequential step-wise format.

4.1 Observational Data Sources

For all the research work connected to different ARs in this thesis, we have worked with the HMI-derived Lambert CEA (Cylindrical Equal Area) projection (i.e. a map projection that preserves surface areas by projecting a spherical surface onto a cylinder) of the radial component of the vector magnetogram (B_r), bearing the extension `hmi.sharp_cea_720s`, available at the official data-link website of JSOC (Joint Science Operations Center): https://jsoc.stanford.edu. JSOC, is a core data center for NASA's Solar Dynamics Observatory (SDO) at Stanford University in the United States. The SDO is a NASA mission launched in 2010 to provide continu-

ous, high-cadence observations of the Sun and its magnetic activity. It operates in a geosynchronous orbit and carries multiple instruments designed to investigate the solar atmosphere, magnetic field evolution, and drivers of space weather. In this work, observations from the Helioseismic and Magnetic Imager (HMI) are employed. The HMI observes the full solar disk in the Fe I 6173 Å spectral line and provides measurements of the photospheric magnetic field, Doppler velocity, and continuum intensity. The instrument records full-disk images with a plate scale of approximately 0.5 arcsec/pixel corresponding to a spatial sampling of about 360 km per pixel at the solar disk centre, and achieves an effective spatial resolution of roughly 1 arcsec. By providing the SHARP number corresponding to the desired AR, we obtained the radial magnetogram data at 1 hour cadence within 60° EW of the central meridian of the Sun as the data is known to suffer from projection effects (or geometric distortions) when obtained close to the limb regions (Bobra et al. 2014).

4.2 PIL Detection and Length Computation

There are two PIL detection and length computation methods discussed in this section - (i) the FLARECAST code (Georgoulis et al. 2021) adapted from Mason & Hoeksema (2010) and Park et al. (2018) and (ii) the detection algorithm of Cai et al. (2020), used to generate the publicly available PIL length dataset (Ji et al. 2023).

4.2.1 The FLARECAST Code

The code for the identification of PILs is available in the FLARECAST project repository: https://dev.flarecast.eu/stash/projects/FE/repos/mpil-idl/browse. The main program file is `mpil.pro` and there are six other assisting dependencies: `alex.pro`, `edpt.pro`, `cropcont.pro`, `find_nl_n.pro`, `get_mpil_region.pro`, `isth.pro`.

Control Parameter	Physical Value	Scale	Function
`pbt_thresh1` (Bh_{th1})	120 G	- - -	to set a threshold over interpolated horizontal field Bh to filter PIL candidates at different levels
`pbt_thresh2` (Bh_{th2})	100 G	- - -	
`thresh1` (Bzg_{th1})	20 G	- - -	to set a threshold over interpolated spatially (x,y) averaged vertical field Bzg to filter PIL candidates
`thresh2` (Bzg_{th2})	8 G	- - -	
`minsize1` (ms_1)	$\sim$ 5.5 Mm	15 pixel	to select PIL candidates greater greater than a specific length to minimise noisy features / artefacts
`minsize2` (ms_2)	$\sim$ 2 Mm	5 pixel	
`isth_cut1` (ic_1)	$\sim$ 2 Mm	5 pixel	to set a length threshold to connect separate filament like structures within a single PIL
`isth_cut2` (ic_2)	$\sim$ 1 Mm	3 pixel	
`grad_width` (gw)	$\sim$ 3.5 Mm	10 pixel	spatial scale for smoothing
`sep` (s)	$\sim$ 11 Mm	30 pixel	spatial scale to group PILs
`xy_pixel` (w)	$\sim$ 2 Mm	5 pixel	width around PILs to compute flux

Table 4.1: Table listing the important parameters used in the detection of PILs and the flux surrounding PILs using `find_nl_n.pro`; `xy_pixel` is the only flux specific control parameter, not defined in `find_nl_n.pro` but is defined `get_mpil_region.pro` to detect flux boundaries

The FLARECAST algorithm detects PILs using two major control parameters - (i) the horizontal magnetic field value B_h and (ii) the spatially averaged value of B_z in the horizontal plane of the magnetogram i.e. B_{zg}. B_{zg} may be also understood as a heuristic proxy for the horizontal gradient of the vertical field because it is computed with a fixed horizontal averaging scale (or box size) i.e. gw (see Table 4.1). The PIL detection and length computation procedure consists of seven steps.

1. Pre-processing:

The input parameters are defined (see Table 4.1) and the input magnetogram map of the vertical magnetic field (B_r or B_{los} - see Figure 4.1) is smoothed using a smoothing routine in IDL to remove noise.

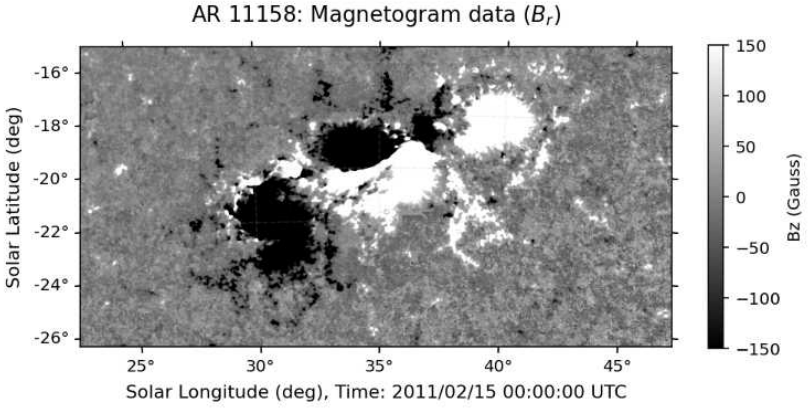

Figure 4.1: The original magnetogram map (B_r) without any pre-processing for AR 11158 at 00:00 UTC on 15 Feb, 2011

2. Magnetic field computation:

A PF is computed from the input magnetogram using the routine `alex.pro`, which extrapolates the observed vertical magnetic field to obtain the three components of the magnetic field vector. Although this extrapolation provides the full vector field, only the photospheric values are used in the PIL identification procedure described here.

From the input magnetogram, two additional maps are constructed: (i) a smoothed positive-polarity map and (ii) a smoothed negative-polarity map. These are obtained by first separating the magnetogram into its positive and negative components and then applying spatial smoothing over a scale set by the parameter $gw = 3.5$ Mm (see Table 4.1). The resulting maps represent spatially averaged measures of nearby positive and negative magnetic flux density (in Gauss). These smoothed polarity maps (see Figure 4.2) quantify the presence of strong, coherent opposite polarities in the neighbourhood of a candidate PIL. Their role is to suppress PIL segments that arise from weak-field regions or small-scale noise, ensuring that detected PILs are associated with substantial magnetic flux concentrations.

3. Zero Gauss contour detection:

All possible neutral line candidates are identified by using IDL's built-in `contour` routine (see Figure 4.3).

4. Filtering and structuring:

For each candidate PIL identified from the zero-Gauss contour, the algorithm interpolates three quantities at each contour point: (i) the smoothed positive-polarity magnetic field (B^+_{zg}), (ii) the smoothed negative-polarity magnetic field (B^-_{zg}), and (iii) the horizontal magnetic field strength (B_h) derived from the PF (see Figure 4.4). All three interpolated quantities are expressed in Gauss. The thresholds Bh_{th} and Bzg_{th} are applied directly to the smoothed magnetic field amplitudes across the neutral line. A contour point is retained as part of a valid PIL candidate only if it satisfies all of the following conditions (within gw of the contour point): (i) $B^+_{zg} > Bzg_{th}$, (ii) $B^-_{zg} < -Bzg_{th}$ and (iii) $B_h > Bh_{th}$ (`pbt_thresh`).

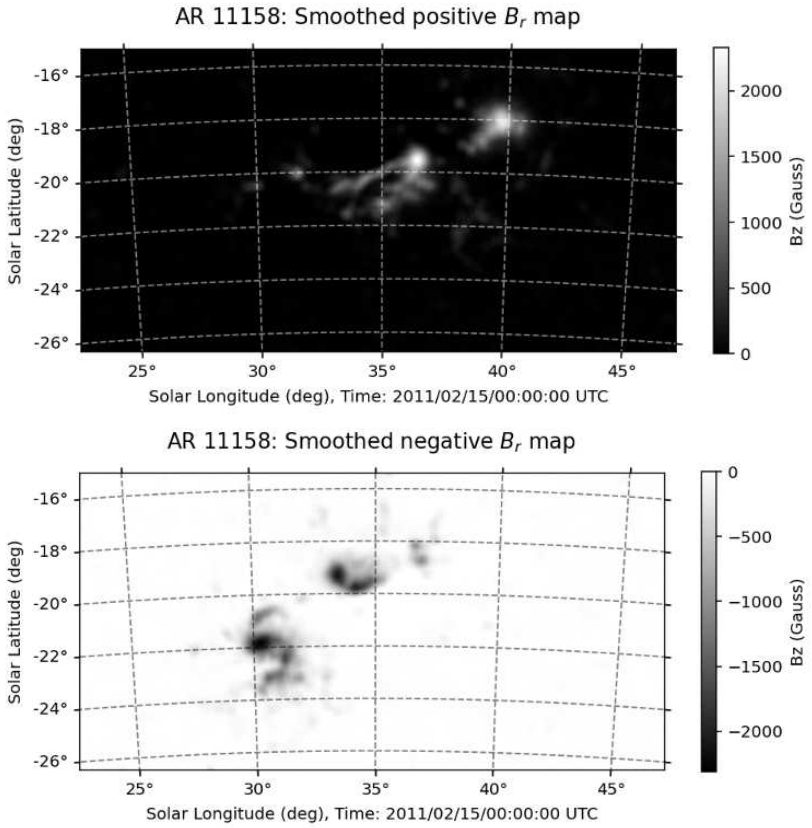

Figure 4.2: Smoothed magnetic field maps for the input magnetogram data of AR 11158 as shown in Figure 4.1; smoothed positive B_r map (left) and smoothed negative B_r map (right).

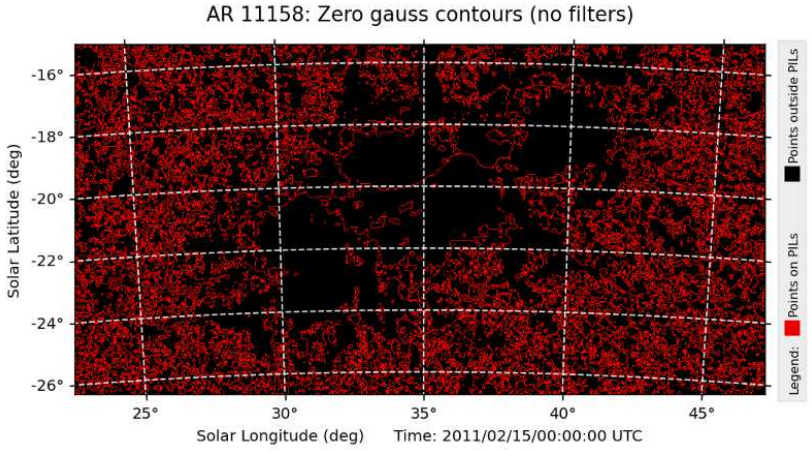

Figure 4.3: All candidate PILs computed by the IDL `contour` routine without any filtering

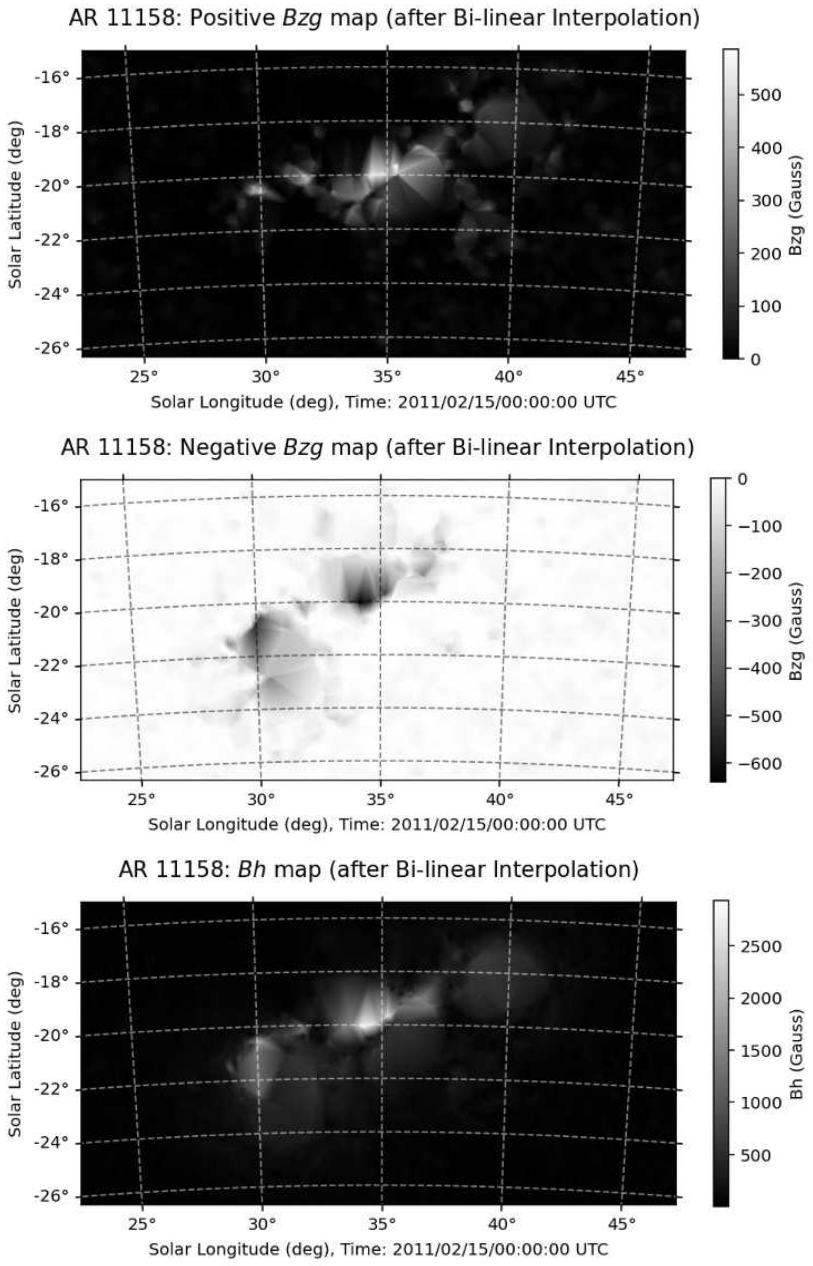

Figure 4.4: Smoothed magnetic field maps after PF computation and subsequent interpolation, for (i) the smoothed positive polarity map (top), (ii) the smoothed negative polarity map (middle) and (iii) the horizontal magnetic field (bottom); the values B_{zg}^{+}, B_{zg}^{-} and B_h are only computed on the detected candidate PILs as shown in Figure 4.3; to enable visualisation, we have used Bilinear interpolation to generate a 2D map

Although the FLARECAST code description in `mpil.pro` refers to the input control parameter `thresh` as a "horizontal gradient of the vertical field", B_{zg} which is computed after interpolation and compared to `thresh` is not a true spatial derivative of B_z. Instead, the code evaluates the spatial dominance of opposite magnetic polarities within a smoothing window of width gw, effectively measuring the consistency of vertical field over a horizontal spatial scale. However, when normalised by the smoothing length, this quantity may be interpreted as a proxy for the horizontal gradient (in G/Mm unit), explaining both the numerical scale of the adopted thresholds and their sensitivity to the choice of smoothing width. The combined thresholding using Bh_{th} and Bzg_{th} enforces the requirement that a PIL must separate strong, spatially coherent opposite polarities and be associated with a sufficiently strong horizontal field. After this filtering step, the surviving contour points are grouped into line segments and stored in internal data structures for subsequent length filtering, topological cleanup, and optional comparative smoothing.

5. Post-processing:

In this step, extremely short lines, lower than the user-defined threshold ms and thin connections (called isthmuses), lower than the user-defined threshold ic are removed. Following this secondary filtering step, structured data is passed as output.

6. Comparative smoothing:

Steps 1-5 show the overall design of the IDL routine `find_nl_n.pro`. In step 4, there is an option of enforcing comparative smoothing. However, this does not get executed if `find_nl_n.pro` were to be executed on a standalone basis. Hence, the FLARECAST project has another IDL routine `get_mpil_region.pro`, where `find_nl_n.pro` is called at two different levels. At each level, the input parameters vary. The first call is a

more conservative one and it leads to a coarser identification of PILs. The second call however, leads to finer PIL candidates being identified. The second call is only activated when the initial detection succeeds and it is used to refine the results of the first call (see sample detection in Figure 4.5).

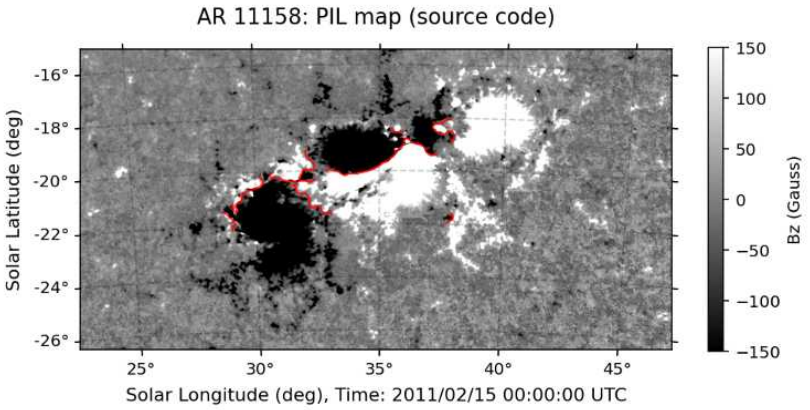

Figure 4.5: PILs detected using the FLARECAST source code with application of comparative smoothing; In the first stage of PIL detection (Bh_{th1}, Bzg_{th1}) = (120 G, 20 G); In the second stage of PIL detection (Bh_{th2}, Bzg_{th2}) = (100 G, 8 G); this is the result corresponding to the input magnetogram shown in Figure 4.1

7. Computation of PIL parameters:

After PIL detection, the main program file `mpil.pro` is used to numerically compute the PIL parameters (i) L_{tot} i.e. the total length of PILs, (ii) L_{max} i.e. the length of the longest PIL and (iii) ϕ i.e. the unsigned flux in the vicinity of PILs (within 2 Mm). The details related to ϕ are shared in Section 4.3.

4.2.2 Modifications of the FLARECAST Code

Case 1: PIL detection without two-step comparative smoothing

In the first step of PIL detection (Bh_{th1}, Bzg_{th1}) = (120 G, 20 G) are the coarser set of thresholds prescribed by FLARECAST. Similarly, in the second step of PIL detection (Bh_{th2}, Bzg_{th2}) = (100 G, 8 G) are the finer set of thresholds. If we do not execute `get_mpil_region.pro`, and slightly modify `find_nl_n.pro` and execute it on a standalone basis, then the difference between the two set of thresholds may be observed (see Figure 4.6).

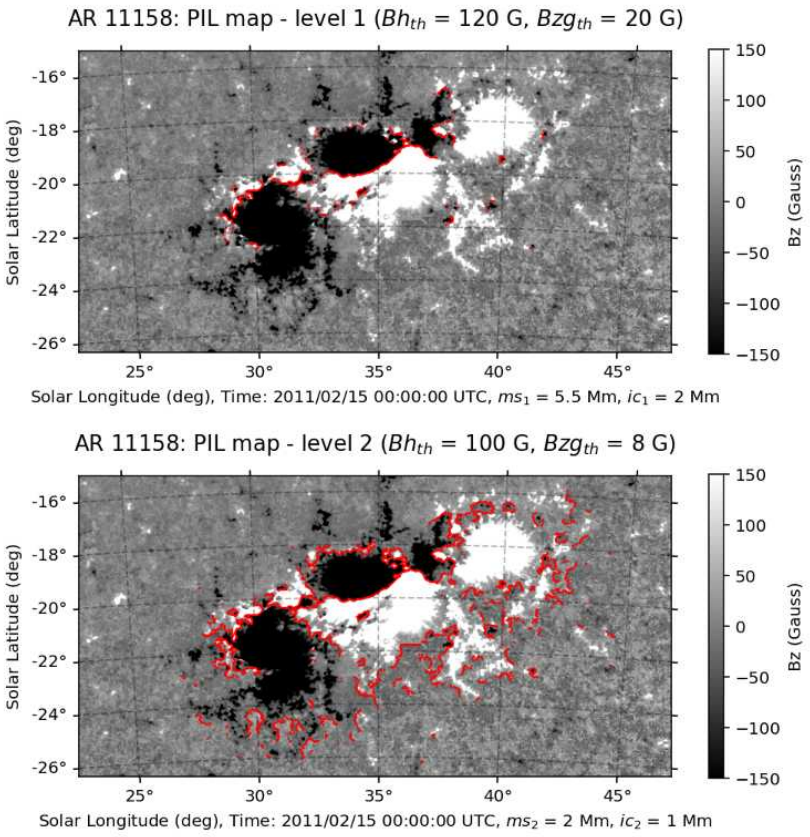

Figure 4.6: Preliminary detection of PILs without the two-step threshold before using any length based filter; using coarser thresholds at level 1 (top) and finer thresholds at level 2 (bottom); notice how using finer thresholds helps identify more fragments of PILs compared to the coarser thresholds

Case 2: Switching off the horizontal gradient of the vertical field threshold

In this case, we switched off Bzg_{th}, the control parameter for B_{zg}. In the first instance, we used $Bh_{th} = 120$ G (see Figure 4.7), and the result included many noisy PIL fragments being detected. However, when we increased the Bh_{th} to 500 G, the noisy PIL fragments were largely removed, leaving mostly the central PIL to be detected. These results appear to be consistent with the findings of Sun et al. (2012), who reported B_h values of 1200-1500 G, around the X2.2 flare time at 01:58 UTC. However, the PILs detected were still not very continuous as we find in the two-step smoothing process (see Figure 4.5).

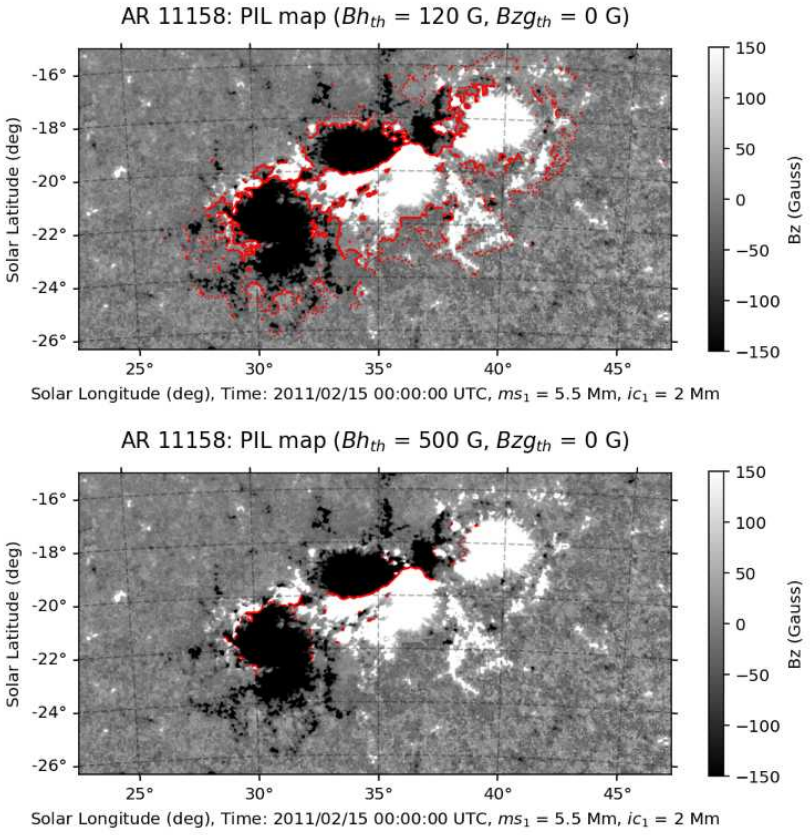

Figure 4.7: PILs detected without any Bzg_{th}, the threshold for the spatially averaged vertical magnetic field B_{zg}; here the control parameter is only B_h with $Bh_{th} = 120$ G (top) and $Bh_{th} = 500$ G (bottom); notice how the central PIL has an extremely strong B_h with values exceeding 500 G

Case 3: Switching off the horizontal field threshold

In this case, we switched off Bh_{th}, the control parameter for B_h. In the first instance, we used $Bzg_{th} = 20$ G (see Figure 4.8). At this threshold, we were able to demarcate positive and negative polarity B_z regions. However, the PILs detected were largely fragmented and not quite smooth or continuous as in Figure 4.5. Upon increasing Bzg_{th} to 100 G, despite most of the noisy PIL fragments being removed, a major section of the central PIL still survived, albeit in fragments. Overall, from these case studies, we find that using both B_h and B_{zg} as control parameters, in conjunction with the two-step smoothing procedure may be a more robust PIL detection approach.

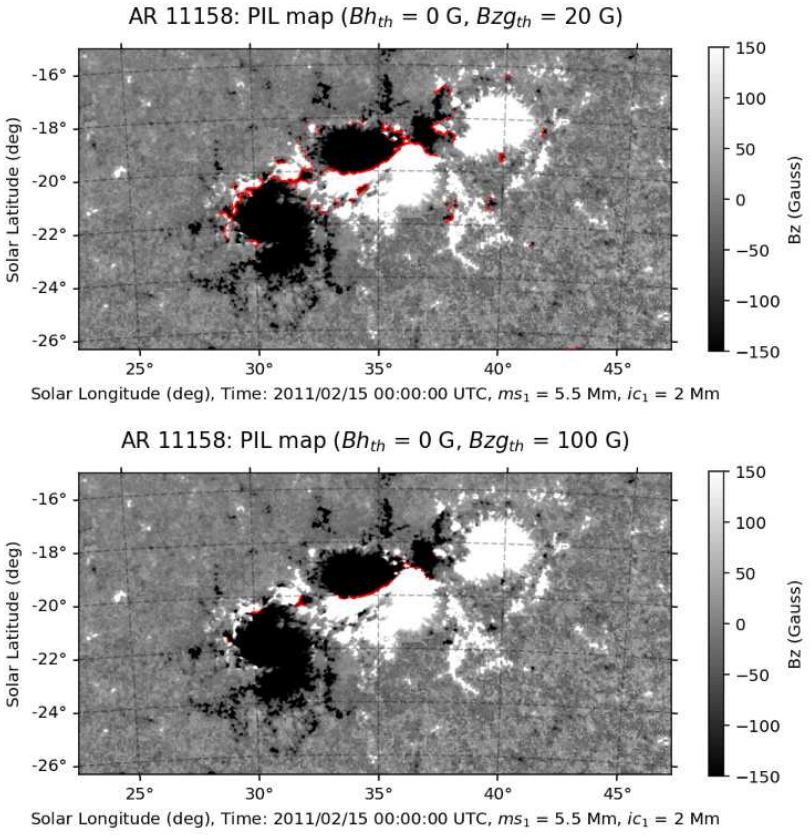

Figure 4.8: PILs detected without any Bh_{th}, the threshold for the horizontal field B_h; here the control parameter is only B_{zg} with $Bzg_{th} = 20$ G (top) and $Bzg_{th} = 100$ G (bottom); notice how the central PIL has an extremely strong B_{zg} with values exceeding 100 G

4.2.3 Detection Method of Cai et al. 2020

Cai et al. (2020) developed a method for detecting PILs from line-of-sight longitudinal magnetograms (B_{los}) or from the radial component of vector magnetograms (B_r). Based on this PIL detection mechanism, Ji et al. (2023) published an extensive open-source dataset which contained information on PILs detected across 4090 HARP series between May 2010 and March 2019, for both B_{los} and B_r data as input. This dataset contains details of several important parameters such as (i) the number of PILs detected, (ii) the total PIL length, (iii) region of Polarity Inversion (RoPI), which is then used to estimate the (iv) unsigned magnetic flux by summing the absolute magnetic field values for all pixels within the RoPI and the (v) fractal value, quantifying the self similarity in PILs. The detection method may be simplified into six steps.

1. Polarity identification:

The first step is to identify the positive and negative polarity regions, by setting a certain threshold value for the vertical magnetic field (Ji et al. (2023) chose a threshold of 100 G for the vertical magnetic field B_z). Information regarding the positive and negative polarity regions are stored in the form of positive and negative bitmaps, respectively.

2. Edge detection:

Using the Canny edge detection mechanism (Ji et al. 2023), the bitmaps are used in the identification of the interfaces of regions with strong polarity. As a result, in this step two binary masks are produced that contain the edges of the strong polarity regions, one each for the positive and negative bitmaps.

3. Morphological dilation:

RoPI (Region of Polarity Inversion) denotes the regions where the positive and negative polarity regions are adjacent. This is done by expanding the detected edges in the two binary masks using a technique called morphological dilation (Ji et al. 2023), and then subsequently selecting the region where these dilated edges intersect.

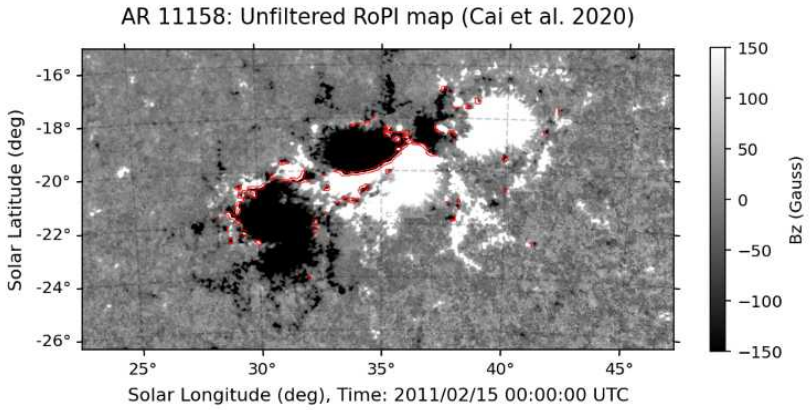

Figure 4.9: RoPI map after morphological dilation

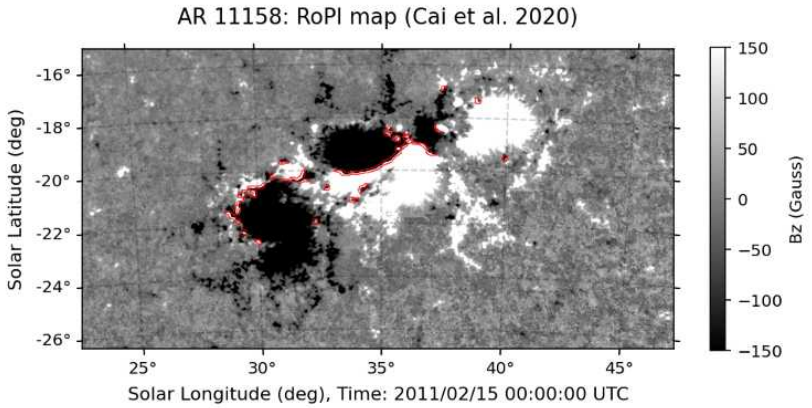

Figure 4.10: Final RoPI map (top 95%) after filtering out insignificant regions

4. Filtering RoPIs:

The RoPI identified at the previous step usually contains several isolated fragments of polarity inversion, each fragment corresponding to some associated unsigned magnetic flux. All isolated fragments are ranked in descending order, based on their unsigned fluxes. Then a user-defined threshold percentage (say $x = 95\%$) is to be used to identify the top n isolated fragments which account for $x\%$ of the total unsigned flux for the RoPI computed in the previous step. This step essentially filters off isolated regions that do not contribute significantly to the unsigned flux, even though they indicate the presence of strong polarity in step 2.

5. Morphological Thinning:

A technique called morphological thinning (Ji et al. 2023) is then applied to the RoPI detected in step 4 and this produces a pixel chain representation of the candidate PILs. This usually produces several disconnected strands of candidate PILs (see Figure 4.11).

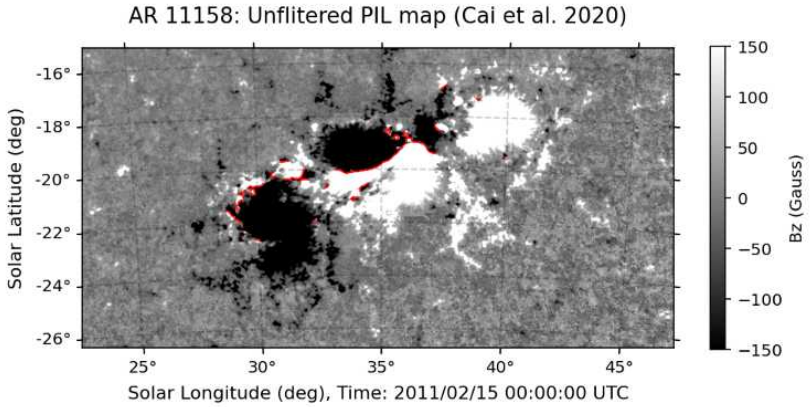

Figure 4.11: Unfiltered PIL map after morphological thinning of the RoPI map

6. Filtering PILs:

The candidate PILs identified in step 5 may include some very small isolated PILs. For a coarse detection, only the PILs longer than a minimum length threshold are retained and the insignificant candidate PILs are filtered off (see Figure 4.12).

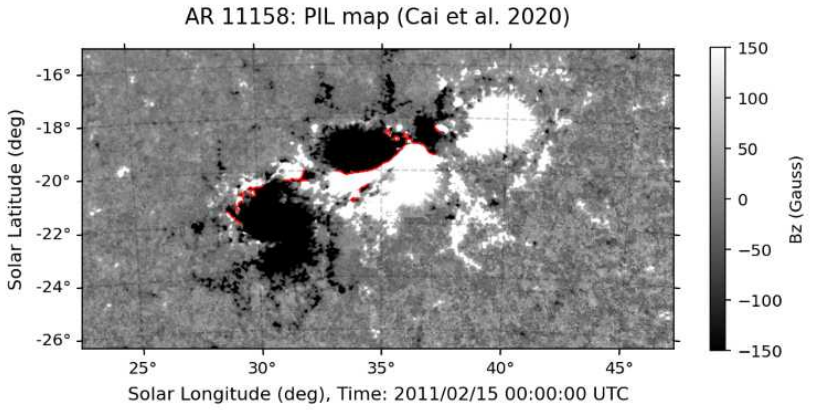

Figure 4.12: Final PIL map, after filtering out PILs shorter than 5 Mm

4.2.4 A Comparison of the PIL Detection Methods

The procedure of detecting PILs proposed by Cai et al. (2020) is similar to the one described in the FLARECAST code (Georgoulis et al. 2021). These methods are similar in the sense that both these methods employ a priori thresholds (i) to identify the interfaces between regions with strong opposite polarities and (ii) reject extremely short PILs. However, key differences between the two algorithms exist and these differences may lead to PILs being detected differently.

While the FLARECAST code sets a minimum threshold on the (i) spatial average of the vertical magnetic field (horizontally along (x, y)) and the (ii) horizontal magnetic field across PILs, the approach by Cai et al. (2020) sets a threshold only on the vertical magnetic field to identify PILs. Thus, unlike the FLARECAST code, where both the horizontal and vertical components of the magnetic field dictate the identification of PILs, the algorithm published by Cai et al. (2020) does not have any threshold/metric to filter PILs based on the horizontal magnetic field pattern. Further, the FLARECAST code has a provision to 'connect' segregated filaments if they happen to be sufficiently close. However, no such provision exists in the algorithm of Cai et al. (2020). Another advantage of the FLARECAST code is that it also leads to a more robust detection of PILs as it enables a comparative smoothing process (in its default version). Owing to these factors, it can also be said that the FLARECAST code is more complex and computationally expensive compared to the approach suggested by Cai et al. (2020) and the latter may be useful for a much quicker detection of PILs and computation of useful features as a first hand estimate.

4.3 Computation of Unsigned Magnetic Flux

The method of computing the unsigned magnetic flux ϕ is similar for both the FLARECAST method and that by Cai et al. (2020). In the FLARECAST codes, the parameter `xy_pixel` (w) creates a width around the PIL (default value is 2 Mm) within which the $|B_z|$ values are summed and multiplied with pixel area i.e. 1.33×10^{15} cm^2, to compute the unsigned magnetic flux (see Figure 4.13).

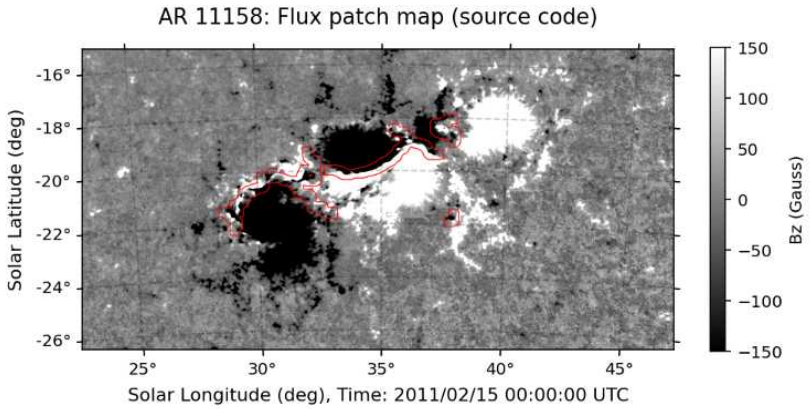

Figure 4.13: Flux patch overlaid in red over the magnetogram data; the closed contour region encloses a radial distance of 2 Mm from the PIL

In the PIL detection approach of Cai et al. (2020), one of the outputs produced by Ji et al. (2023) in the open source dataset is the sum of $|B_{los}|$ within the RoPI (in Gauss). When this is multiplied with the pixel area i.e. 1.33×10^{15} cm^2, the unsigned flux may be obtained in Mx.

4.4 Computation of the R-value Parameter

The algorithm for the computation of R-value in the FLARECAST codes requires only two SSW-IDL program files: `r_value.pro` and `smear.pro`. Both the program files can be accessed at the FLARECAST's project repository: https://dev.flarecast.eu/stash/projects/FE/repos/rvalue-idl/browse. The algorithm to compute the R-value parameter is essentially a four step process:

1. Creating the filter

First, a positive polarity bitmap and a negative polarity bitmap are calculated, starting from the input magnetogram data (see Figure 4.1). These maps have binary values (0/1) depending on (i) whether the field intensity is greater than the threshold Bz_{th} in the case of the positive bitmap and (ii) whether the field intensity is lesser than the threshold $-Bz_{th}$ in the case of the negative bitmap. Although, the threshold Bz_{th} can be set to any reasonable value as per the user, Schrijver (2007) after a statistical analysis concluded that on the photosphere, it is optimal to have $Bz_{th} = 150$ G. Both these positive and negative polarity maps are smoothed over a kernel of size 3 pixel $\times$ 3 pixel. The resulting maps M_{pos} and M_{neg} are shown in Figure 4.14. In these maps, every pixel contains information on the occupancy fraction i.e. the number of pixels in the 9-pixel smoothing box that satisfy $B_z > Bz_{th}$ in M_{pos} and $B_z < -Bz_{th}$ in M_{neg}.

2. Detection of PIL skeleton

In the previous step, we detected the positive and negative polarity bitmaps. If we take their product, then the result would be zero as at every point both $B_z > Bz_{th}$ and $B_z < -Bz_{th}$ cannot be satisfied. However, since we carry out smoothing of both the bitmaps and obtain M_{pos} and M_{neg}, the edges of the high polarity regions represented in these maps get dilated. Thus when we multiply M_{pos} and M_{neg}, the corresponding occupancy fractions get multiplied and we obtain a PIL map.

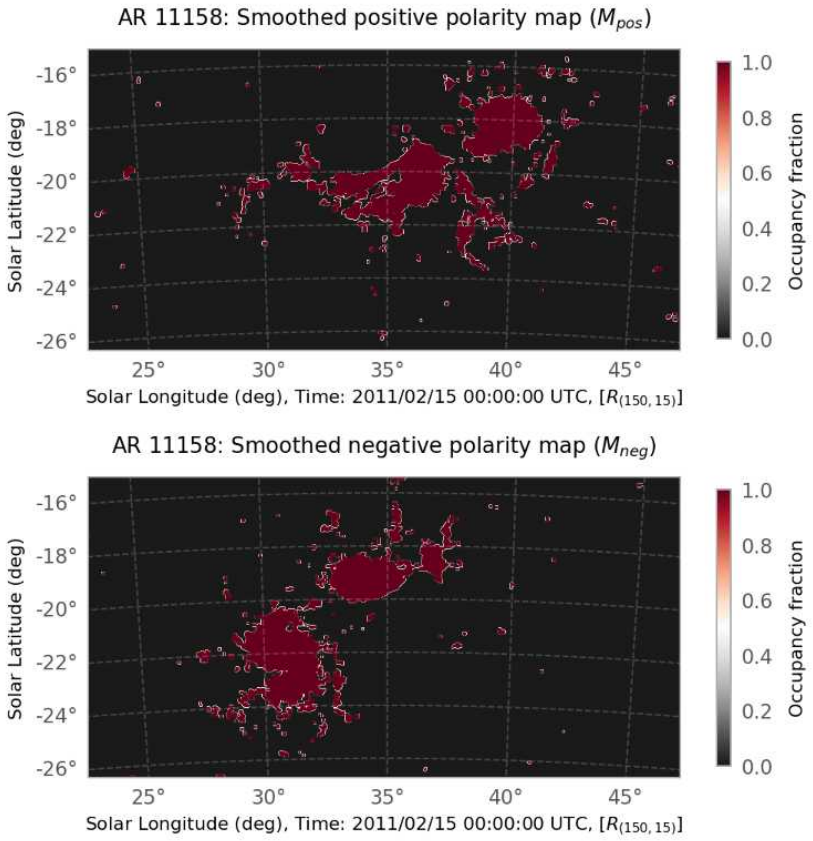

Figure 4.14: Smoothed polarity maps M_{pos} (top) and M_{neg} (bottom); the occupancy is the number of pixels in the 9-pixel smoothing box that satisfy $B_z > Bz_{th}$ in M_{pos} and $B_z < -Bz_{th}$ in M_{neg}

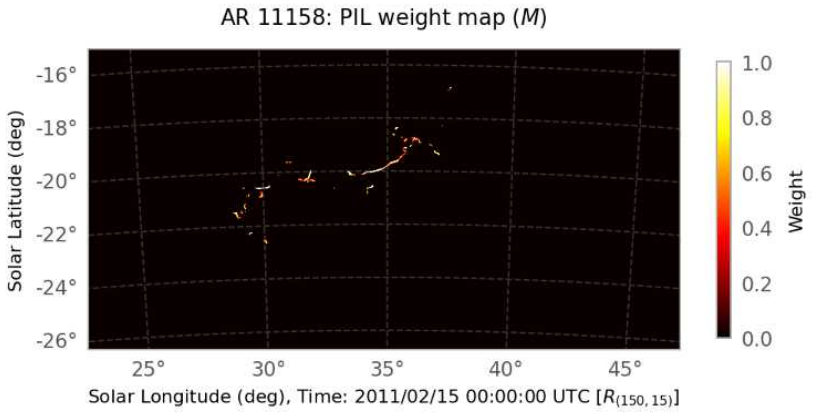

Figure 4.15: PIL skeleton map (M) detected by the multiplication of M_{pos} and M_{neg}

3. Convolution of PIL skeleton

The PIL skeleton identified is too thin to measure any unsigned flux contained on/around it meaningfully. Hence it is expanded into a finite-width neighbourhood using Gaussian convolution, defining the spatial region over which magnetic flux is summed to compute the R-value. This is done by applying the `smear` function onto the PIL weight map M which carries out Gaussian convolution over FWHM $= D_{sep}$. Schrijver (2007) used $D_{sep} = 15$ Mm for R-value calculations because about two thirds of the data had a separation of $D < 15$ Mm between PILs and the brightest point in the TRACE (Transition Region and Coronal Explorer) EUV image. The post-convolution weight map W corresponding to Figure 4.15 is shown in Figure 4.16. Each pixel in this map contains a weight arising from the convolution procedure

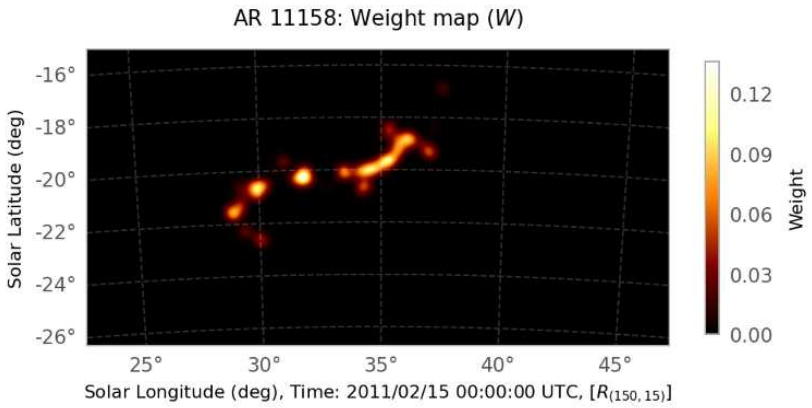

Figure 4.16: The weight map W, obtained by convolving the PIL skeleton map with a 2D Gaussian filter of FWHM $= D_{sep}$

4. Computation of R-value

In the weight map W created in the previous step, the weight information contained in each pixel is multiplied with the original input magnetogram to obtain the normalised magnetic field that shows the region around PILs that is meaningful for R-value calculations (see Figure 4.17-R_1). If we (i) multiply the weight map W element-wise with $|B_z|$

of the original magnetogram data (see Figure 4.17-R_2), (ii) sum over all the magnetic field values, and (iii) multiply the result with the pixel area, then we obtain the R-value.

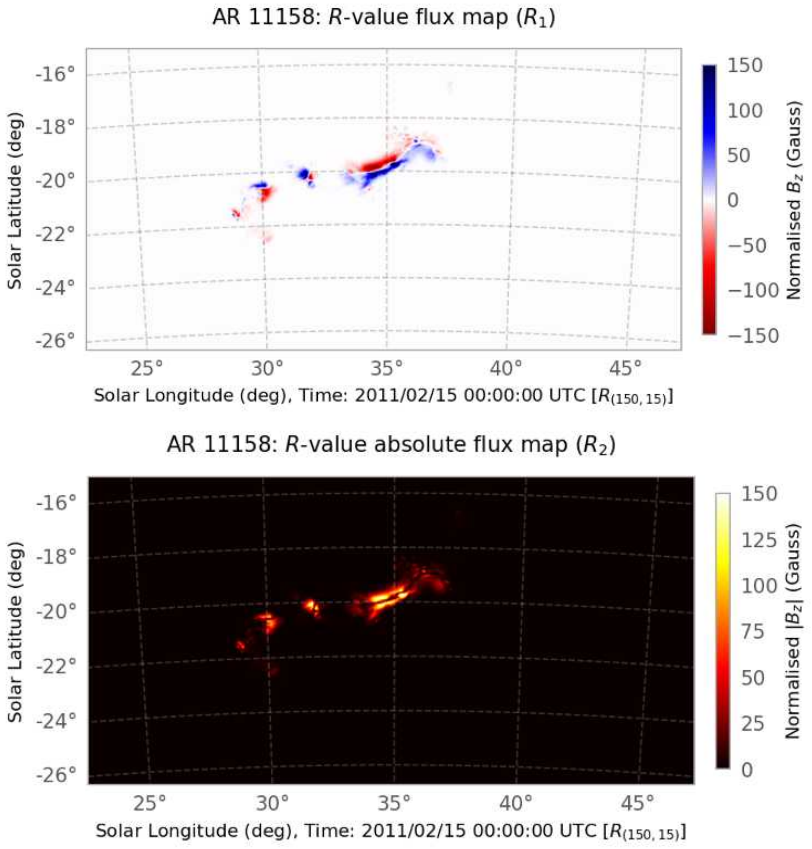

Figure 4.17: R-value maps identified by multiplying W with magnetogram B_z (top) and magnetogram $|B_z|$ (bottom)

Chapter 5

Analysis of R-value

5.1 Objective

In this chapter we discuss if the evolution of flux based parameters i.e. the unsigned magnetic flux and more importantly the R-value (as it is specifically designed to be a proxy for the unsigned magnetic flux in high B_z polarity regions) may help predict major solar flares. Motivated by the 3-dimensional approach of Korsós et al. (2020) and Korsós et al. (2022), we examine if the R-value may provide hints of an imminent major solar flare, when studied in the lower solar atmosphere above the photosphere. We also check if an OHR may be determined for the R-value which may help improve flare prediction when the R-value is studied within it, compared to a photospheric analysis (see section 1.4). The approach followed in this chapter is to extrapolate the magnetogram data (using PF approach, see section 2.6) for ARs specified in the dataset (see section 5.2) at step sizes of 0.5" (0.36 Mm) up to a height of 3.24 Mm. The extrapolation and data analysis is carried out with a cadence of 1 hr in the pre-flaring environments of ARs (mainly linked to X-class flares) and non-flaring ARs (ARs not producing any flare stronger than GOES C1.0 class). Some M-class flare producing AR cases and a C-class flare producing AR are studied as intermediate cases. The results shared in this chapter are published and peer-reviewed (Biswal et al. 2024).

5.2 Dataset

We selected 8 ARs belonging to the solar cycle 24 that produced X-class flares as their strongest flare (see Table 5.1). The selection criteria for these ARs is as follows:

1. The SDO/HMI SHARP magnetogram data corresponding to the AR is continuously available for at least 120 hr, suggesting that the AR has sufficient longevity.

2. The AR must be within 60° EW of the central meridian during the full duration (at least 120 hr), so that the magnetic field can be reliably used to construct the PF extrapolated magnetic field data-grid (see section 4.1)

3. The minimum time available for analysis before the onset of the first flare is at least 48 hr, so that there is sufficient time to analyse the pre-flare environment.

4. The AR is linked to a δ-type sunspot group, signifying a complex arrangement of the photospheric magnetic field (B_z).

Since flares exhibit persistence i.e. ARs that have flared in the past are likely to flare again (Park et al. 2020), we restrict our analysis of the R-value to the first X-class flare produced by the AR. For example, consider the case of AR 12673. It is considered as an X-class flaring AR as the strongest flare was an X9.3 flare (2017/09/06 11:53 UTC). However, because the first X-class flare for AR 12673 was an X2.2 flare (2017/09/06 08:57 UTC), we study the R-value evolution in the 48 hr interval immediately preceding the first (X2.2) flare.

We also analysed 4 M-class flaring ARs, the details of which are shared in Table 5.2 and a C-class flaring AR (using the same conditions written above).

Observed Data (X-class flaring ARs)						
No.	AR	Class	T_{start}	T_{end}	$T_{flare\ onset}$	ΔT (hrs)
01	11158	X2.2	2011/02/11 00:00	2011/02/17 00:00	2011/02/15 01:44	97.73
02	11166	X1.5	2011/03/05 00:00	2011/03/10 00:00	2011/03/09 23:13	119.22
03	11283	X2.1	2011/09/01 00:00	2011/09/09 00:00	2011/09/06 22:12	142.2
04	11520	X1.4	2012/07/08 00:00	2012/07/15 00:00	2012/07/12 15:37	111.62
05	12017	X1.0	2014/03/23 00:00	2014/04/01 00:00	2014/03/29 17:35	161.58
06	12158	X1.6	2014/09/07 00:00	2014/09/14 00:00	2014/09/10 17:21	89.35
07	12297	X2.1	2015/03/09 00:00	2015/03/15 00:00	2015/03/11 16:11	64.18
08	12673	X2.2	2017/09/03 00:00	2017/09/08 00:00	2017/09/06 08:57	80.95

Table 5.1: Table listing the details of the studied X-class flaring ARs, GOES flare classes, observation intervals (between T_{start} and T_{end}), flare onset time ($T_{flare\ onset}$), time difference between flare onset and start of the observing interval ΔT (in hours)

Observed Data - (M-class flaring ARs)						
No.	AR	Class	T_{start}	T_{end}	$T_{flare\ onset}$	ΔT (hrs)
01	11620	M2.2	2012/11/24 00:00	2012/11/30 00:00	2012/11/28 21:20	117.33
02	11719	M6.5	2013/04/08 00:00	2013/04/13 00:00	2013/04/11 06:55	78.92
03	11818	M3.3	2013/08/12 00:00	2013/08/19 00:00	2013/08/17 18:16	138.27
04	12497	M1.0	2016/02/08 00:00	2016/02/15 00:00	2016/02/12 10:36	106.60

Table 5.2: Table listing the same details as in table 5.1 but for M-class flaring ARs

Observed Data - non flaring ARs			
No.	AR	T_{start}	T_{end}
01	11143	2011/01/07 00:00	2011/01/13 00:00
02	11710	2013/03/30 00:00	2013/04/05 00:00
03	12240	2013/08/12 00:00	2013/08/19 00:00

Table 5.3: Table listing the details of studied non-flaring ARs and their corresponding time windows

5.3 Results and Discussion

The results and discussion are split into six sections.

The first section (5.3.1) discusses the dependence of R-value on the apriori input parameters that control flux calculations. Based on a series of numerical tests, we determine which R-value models are really useful for analysis of magnetogram data.

Sections 5.3.2 and 5.3.3 discuss the results for X-class flaring ARs and M-class flaring ARs, respectively. Section 5.3.4 explores whether the R-value could be used as a tool to discriminate between flaring and non-flaring ARs.

Section 5.3.5 is dedicated to the analysis of AR 12353, a rare case, where an AR with a δ-type sunspot could only produce C-class flares. This serves as an intermediate case between strongly flaring ARs and non-flaring ARs.

Finally in the last section (5.3.6), the discussion is dedicated to how we may envision the R-value to be useful in a real time flare prediction scenario. In this section a number of important parameters such as lead times and optimal times are defined, analysed and explained with the help of an example.

5.3.1 Input Parameter Sensitivity Analysis

As discussed previously in section 4.4, the computed R-value depends upon the B_z field threshold Bz_{th} and the convolution width D_{sep}. Schrijver (2007) suggested that it may be optimal to have (Bz_{th}, D_{sep}) as (150 G, 15 Mm). However, the calculations of Schrijver (2007) were based on the photospheric B_{los} data as input and for flaring ARs of solar cycle 23. Since we are analysing ARs of solar cycle 24 with B_r data, and also within a height range of 0-3.24 Mm above the photosphere, it may be possible that the optimal values of (Bz_{th}, D_{sep}) are quite different for our case. Thus we start off with the values of Schrijver (2007) as an initial estimate (or as a reference model) and then modify the input parameter to check how sensitive the R-value is to each of the input parameters. We use a total of 5 different models of $R_{(Bz_{th},D_{sep})}$ to test input parameter sensitivity: (i) the reference model - $R_{(150,15)}$, (ii) $R_{(100,15)}$, (iii) $R_{(50,15)}$, (iv) $R_{(150,10)}$, (v) $R_{(50,10)}$ (justification provided in further discussion). We have presented an example of AR 11166 for which the evolution of R-value is presented height-wise in the form of a stack-plot (see Figure 5.1). Further, as an example, at a specific timestamp for AR 11166 (at 18:00 UTC, on Mar 8, 2011), we have presented the numerical values of different R-value models to discuss the sensitivity to change in input parameters.

1. Testing sensitivity to Bz_{th}

Since, we shall analyse magnetic field data at extrapolated heights, B_z shall become weaker with increase in height. Using a higher cutoff for B_z may not be optimal as the high polarity regions which usually get detected at the photosphere, may not get detected at higher altitudes. Thus it is better for us to reduce Bz_{th} in an alternate model, while testing parameter sensitivity. To check the impact of Bz_{th} on R-value calculations we keep D_{sep} fixed at 15 Mm and reduce Bz_{th} from 150 G to 100 G and 50 G (see Table 5.4). We don't reduce Bz_{th} to less than 50 G because the noise level on the photosphere in an AR is about 10 G (Liu et al. 2012).

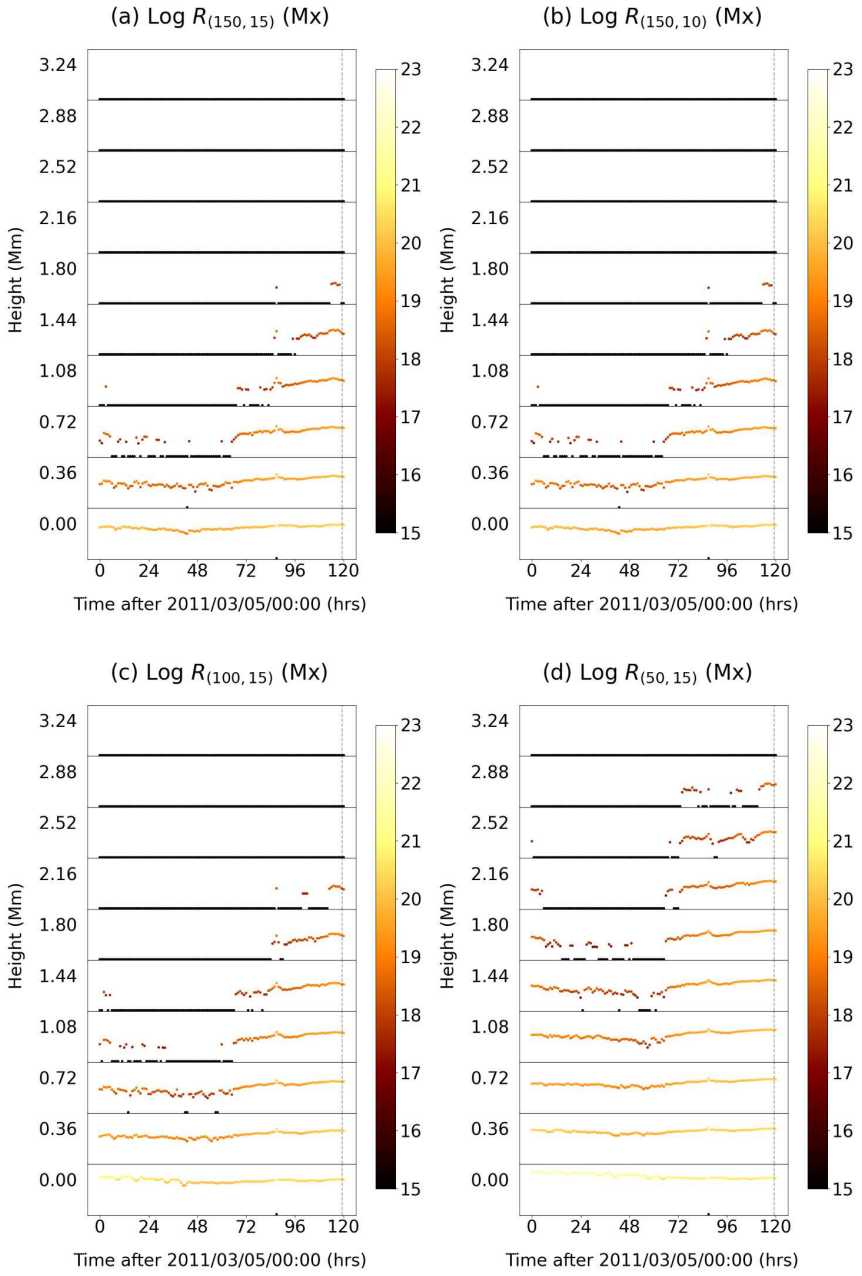

Figure 5.1: R-value stack-plots at a height range of 0-3.24 Mm above the photosphere for AR 11166 (Mar 05-10, 2011); notice how the reducing D_{sep} from 15 Mm to 10 Mm does not impact the overall evolution pattern as significantly (compare plots a & b) as reducing Bz_{th} from 150 G to 100 G and 50 G (compare plots a with b & a with c); flat black lines represent instances where the R-value code output is a null output - indicating that at no point in the magnetogram map B_z exceeds Bz_{th}

Overall, when we reduce Bz_{th}, we find that the R-value increases and black flat lines (representing null output in R-value) begin to get replaced by finite values (see Figure 5.1a, c & d). We find that $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$ are quite distinct in their evolution pattern and the minimum height where a single non-null output in R-value is seen, keeps increasing with a decrease in Bz_{th} (1.80 Mm at 150 G, 2.16 Mm at 100 G and 2.88 Mm at 50 G). Reduction of Bz_{th} enables the identification of more high-polarity regions and these regions in the magnetogram data contribute to the flux calculation (see Figure 5.2). For example, a point on the magnetogram which has $B_z = 125$ G, may get filtered out using a Bz_{th} of 150 G but would be detected using filters of $Bz_{th} = 100$ G and 50 G.

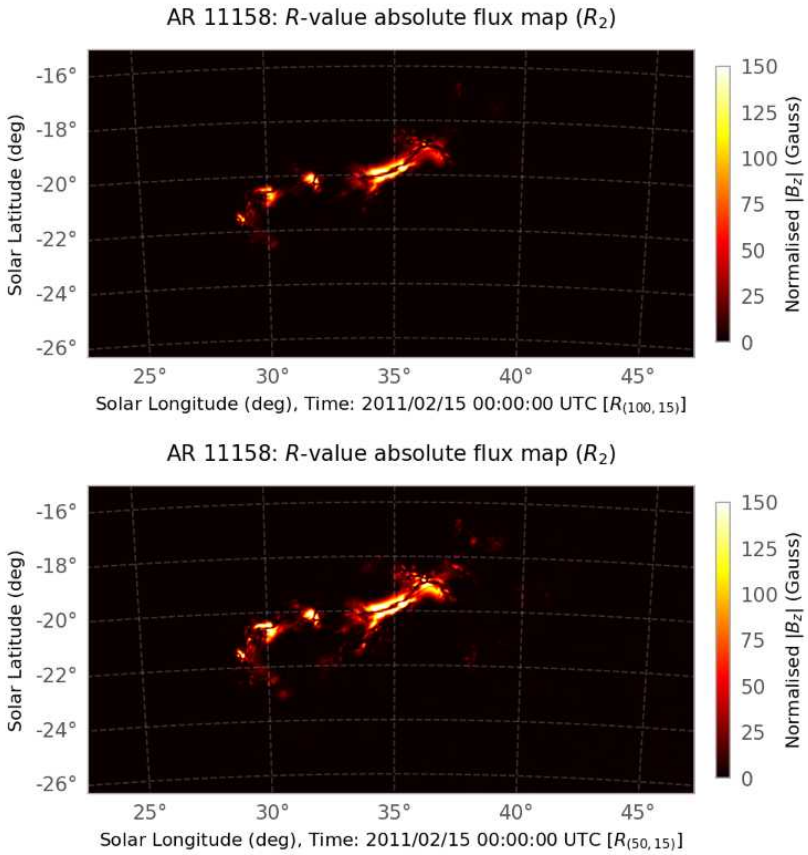

Figure 5.2: A comparison of different R-value flux maps for AR 11158 on 15 Feb 2011 (00:00 UTC) computed using $Bz_{th} = 100$ G (top) and 50 G (bottom); see Figure 4.17 for $Bz_{th} = 150$ G

If we take the example of the single time-stamp mentioned before, we find that reducing Bz_{th} from 150 G to 100 G, keeping D_{sep} fixed at 15 Mm, increases the R-value by 64% on the photosphere and by a factor of ~ 10 at $z = 1.08$ Mm (see Table 5.4). Similarly if we reduce Bz_{th} from 150 G to 50 G, keeping D_{sep} fixed at 15 Mm, the R-value increases by a factor of ~ 3 on the photosphere and by a factor of ~ 50 at $z = 1.08$ Mm (see Table 5.4). A similar pattern is also seen when we fix D_{sep} at 10 Mm and reduce Bz_{th} from 150 G to 50 G (see Table 5.4). Further, we also note that as we reduce Bz_{th} (make the threshold weaker), the number of heights where R-value is a non-null output increases.

Unsigned flux (ϕ) and R-values (in 10^{20} Mx) - testing sensitivity to Bz_{th}									
z (Mm)	ϕ	$R_{(150,15)}$	$R_{(100,15)}$	$R_{(50,15)}$	r_1	r_2	$R_{(150,10)}$	$R_{(50,10)}$	r_3
0.00	23.54	2.1758	3.5746	7.4204	1.64	3.41	2.0698	6.7488	3.26
0.36	15.21	0.5619	1.1095	2.5417	1.97	4.52	0.5240	2.2674	4.33
0.72	11.06	0.1382	0.3935	1.1530	2.85	8.34	0.1228	0.9991	8.13
1.08	5.39	0.0114	0.1291	0.5310	11.36	46.74	0.0103	0.4435	43.18
1.44	3.99	N.V.	0.0218	0.2230	N.V.	N.V.	N.V.	0.1812	N.V.
1.80	0.90	N.V.	N.V.	0.0626	N.V.	N.V.	N.V.	0.0529	N.V.
2.16	1.61	N.V.	N.V.	0.0196	N.V.	N.V.	N.V.	0.0157	N.V.
2.52	0.48	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.
2.88	0.18	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.
3.24	0.00	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.

Table 5.4: A comparison of unsigned flux and different models of R-value for AR 11166 at 18:00 UTC, Mar 08, 2011, to explore the sensitivity to B_{th} (refer to columns 6, 7 and 10); $r_1 = R_{(100,15)}/R_{(150,15)}$; $r_2 = R_{(50,15)}/R_{(150,15)}$; $r_3 = R_{(50,10)}/R_{(150,10)}$; N.V.: No value, indicates null output; column 1 lists the height z above the photosphere; column 2 specifies the unsigned flux; in columns 3, 4 and 5, D_{sep} is fixed at 15 Mm but B_{th} is 150, 100 and 50 G, respectively. In columns 8 and 9, D_{sep} is fixed at 10 Mm but Bz_{th} is 150 and 50 G, respectively.

2. Testing sensitivity to D_{sep}

To check the impact of D_{sep}, we reduced it from 15 Mm to 10 Mm, primarily because Schrijver (2007) suggested that the distribution of D (see section 4.4) is heavily peaked at lower values (67 % of D values were less than 15 Mm). If we check Figure 5.1 (plots a & b), we find that the evolution pattern of R-value is not impacted as significantly as we saw with Bz_{th}. From the example of AR 11166 (single time-stamp), it is seen that a reduction of D_{sep} from 15 Mm to 10 Mm, only amounts to (i) a 5-10 % reduction in R-value, and (ii) a 10-20 % reduction in R-value keeping Bz_{th} fixed at 150 G and 50 G, respectively. The slight decrease in R-value could be explained by the idea that there may be very few important flux patches that may lie between 10-15 Mm from PILs.

Unsigned flux (ϕ) and R-values (in 10^{20} Mx) - testing sensitivity to D_{sep}							
z (Mm)	ϕ	$R_{(150,15)}$	$R_{(150,10)}$	r_4	$R_{(50,15)}$	$R_{(50,10)}$	r_5
0.00	23.54	2.1758	2.0698	0.95	7.4204	6.7488	0.91
0.36	15.21	0.5619	0.5240	0.93	2.5417	2.2674	0.89
0.72	11.06	0.1382	0.1228	0.89	1.1530	0.9991	0.87
1.08	5.39	0.0114	0.0103	0.90	0.5310	0.4435	0.84
1.44	3.99	N.V.	N.V.	N.V.	0.2230	0.1812	0.81
1.80	0.90	N.V.	N.V.	N.V.	0.0626	0.0529	0.85
2.16	1.61	N.V.	N.V.	N.V.	0.0196	0.0157	0.80
2.52	0.48	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.
2.88	0.18	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.
3.24	0.00	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.

Table 5.5: A comparison of unsigned flux and different models of R-value for AR 11166 at 18:00 UTC, Mar 08, 2011 to explore the sensitivity to D_{sep} (refer to columns 5 and 8); $r_4 = R_{(150,10)}/R_{(150,15)}$; $r_5 = R_{(50,10)}/R_{(50,15)}$; columns 1 and 2 are the same as in Table 5.4; in columns 3 and 4, B_{th} is fixed at 150 G but D_{sep} is 15 and 10 Mm, respectively; in columns 6 and 7, B_{th} is fixed at 50 G but D_{sep} is 15 and 10 Mm, respectively; N.V.: No value, indicates null output.

3. Comparison of input parameter sensitivities

While reducing Bz_{th} by about 33% from 150 G to 100 G may increase the R-value by an order at $z = 1.08$ Mm, reducing D_{sep} by the same percentage from 15 Mm to 10 Mm, impacts the R-value significantly lesser. This shows that changes in Bz_{th} tend to impact the R-value more significantly compared to D_{sep}. This trend is true for all heights we analysed, even though the scales may differ.

We know that R-value is a filtered form of ϕ (Schrijver 2007) and since the D_{sep} directly influences the flux that may be summed for R-value calculations, we also checked if the reduction of R-value was significant relative to ϕ. We find that at the photosphere, where the R-value as a fraction of ϕ is the maximum compared to higher altitudes, both $R_{(150,15)}$ and $R_{(150,10)}$ were only about 9% of ϕ (see Table 5.5), suggesting that the reduction in R-value was also not quite significant to ϕ. This fraction may largely be justified from the fact that the B_z values that are effectively used for R-value calculations are normalised by the kernel-based averaging and smoothing procedures described in section 4.4.

Hence, for further analysis of the R-value for different flaring and non-flaring ARs in this thesis, we keep D_{sep} fixed at 15 Mm and only check three different R-value models: (i) $R_{(150,15)}$, (ii) $R_{(100,15)}$ and $R_{(50.15)}$.

5.3.2 Analysis of X-class Flaring ARs

To simplify the discussion of X-class flaring ARs, we classified the eight cases listed in Table 5.1 into two broad categories, based on the evolution of the unsigned flux ϕ (on the photosphere only) in the 48 hr duration immediately before the occurrence of the first X-class flare - (i) ARs exhibiting an increase in ϕ, the examples being ARs 11158, 11166, 12017 and 12673 and (ii) ARs exhibiting a decrease in ϕ, the examples being ARs 11520, 12158 and 12297. Only AR 11283 defied categorisation as ϕ was more or less stable and thus, it is discussed separately.

1. Cases with increase in unsigned flux

It is known that flux emergence near PILs is associated with X-class flares (Toriumi 2022). In this context, we try to obtain a quantitative estimate of the increase in the photospheric ϕ for ARs 11158, 11166, 12017 and 12673. We introduce two parameters: (i) T_{fe} - the time of flux emergence, representing the time when ϕ begins to rise sharply and (ii) T_{fo} - the timestamp in our dataset immediately before flare onset (see Table 5.6). T_{fe} may be determined from the visual inspection of the evolution of ϕ (see Figure 5.3).

Photospheric unsigned Flux ϕ near PILs and $R_{(150,15)}$ (in 10^{20} Mx) at T_{fe} and T_{fo}						
AR	T_{fe}	ϕ (T_{fe})	$R_{(150,15)}$	T_{fo}	ϕ (T_{fo})	$R_{(150,15)}$
11158	2011/02/13 01:00	5.63	0.24	2011/02/15 01:00	25.54	2.31
11166	2011/03/06 16:00	5.93	0.18	2011/03/09 23:00	30.76	2.40
12017	2014/03/28 00:00	2.94	0.05	2014/03/29 17:00	8.30	0.76
12673	2017/09/03 18:00	12.70	0.87	2017/09/06 06:00	61.17	7.35

Table 5.6: Table comparing changes in $R_{(150,15)}$ and ϕ for flux increasing ARs on the photosphere

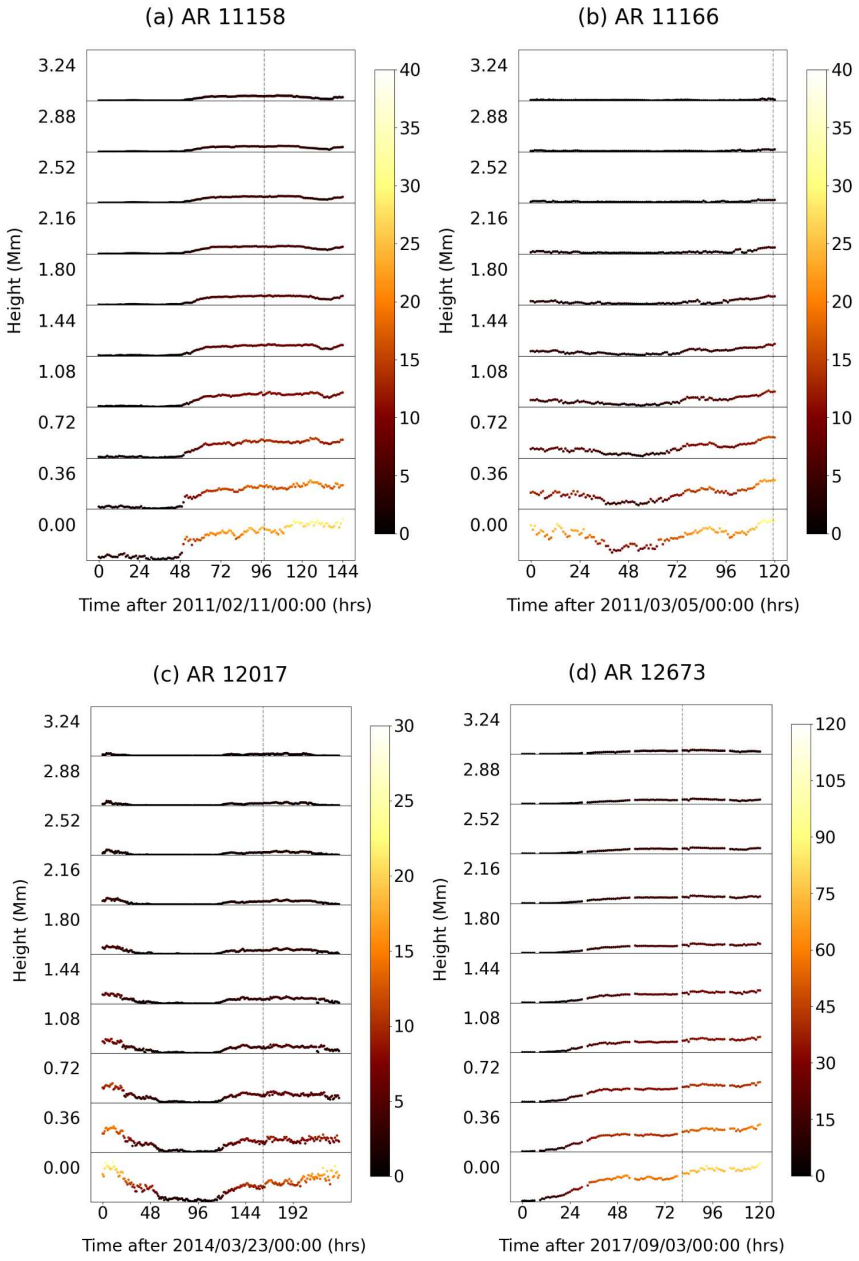

Figure 5.3: Stack-plots showing the height-wise variation of ϕ (in 10^{20} Mx, see colorbar for colour-coding) between 0-3.24 Mm for (a) AR 11158, (b) AR 11166, (c) AR 12017 and (d) AR 12673; the vertical dashed line denotes the time of occurrence of X-class flares in each plot

It is seen that the ϕ (T_{fo}) is about 5 times of ϕ (T_{fe}). In each of these cases, $R_{(150,15)}$ increased by an order between T_{fe} and T_{fo} (see Table 5.6). Further, the ratio between $R_{(150,15)}$ and ϕ is larger at T_{fe} compared to T_{fo}, suggesting that flux close to high-gradient PIL (i.e. PILs across which B_z changes sharply) contributes more to ϕ at a time when the strong flare is about to occur, compared to a time when flux is starting to emerge. Further, the time at which ϕ emerges is similar (or in very close proximity) to the time at which R-value emerges T_{rve}, i.e. when R-value begins to show a sudden jump (see Figure 5.4).

T_{fe} and T_{rve} are quite closely spaced in time. This can be shown visually with the help of stack-plots, and also numerically using an example case of AR 11158, which hosted the first X-class flare (X2.2) of solar cycle 24 on Feb 15, 2011 (01:44 UTC) and is one of the most widely studied ARs. AR 11158 showed a gradual increase in sunspot group complexity from $\beta\gamma$-type on Feb 11 to a $\beta\gamma\delta$-type, a day after the occurrence of the first X-class solar flare. If we observe Figure 5.4, we find that in each of the stack-plots corresponding to a specific R-value model (differing by the choice of Bz_{th}), a continuous flat black line precedes the non-null output in R-value, suggesting that the R-value which remains fairly dormant for a long time suddenly jumps at some point, as we approach the time of occurrence of the first X-class flare. To distinguish a jump in the R-value from noise (i.e. frequent fluctuations between null and finite outputs in R-value), we impose the condition in retrospect that the R-value must have a non-null value continuously for minimum duration of 6 hr, after the last null output value and should not descend back to null outputs again. In this context, the first time-stamp when R-value is a non-null value is called T_{rve}. Table 5.7 lists the T_{rve} for each model of R-value, height-wise. It is seen that T_{rve} tends to occur later in time (i) at higher altitudes for a given Bz_{th} and (ii) at higher values of Bz_{th} for the same height (compare T_{rve} for the three different models at $z = 1.08$ Mm in Table 5.7).

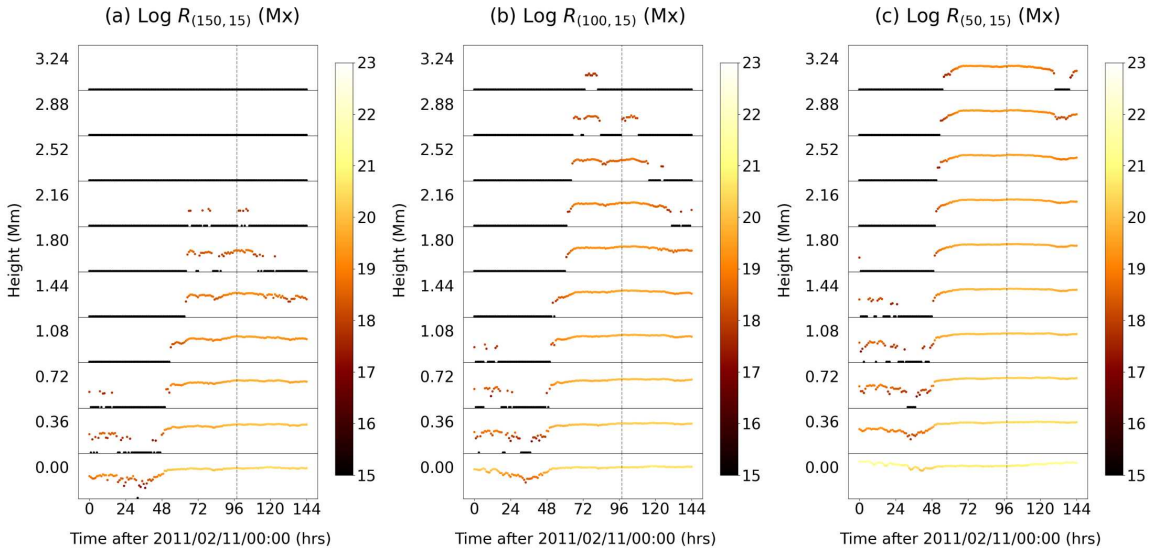

Figure 5.4: Different R-value stack-plots for AR 11158 after 11 Feb, 2011 (00:00 UTC); (a) log $R_{(150,15)}$; (b) log $R_{(100,15)}$ and (c) log $R_{(50,15)}$; the vertical dashed line indicates the time of occurrence of X-class flare; the colourbar gives the logarithm (base 10) of the R-value

R-value emergence times T_{rve} for different R-value models (AR 11158)			
Height (in Mm)	$T_{rve}[R_{(150,15)}]$	$T_{rve}[R_{(100,15)}]$	$T_{rve}[R_{(50,15)}]$
0.00	- - -	- - -	- - -
0.36	2011/02/13 00:00	2011/02/12 14:00	- - -
0.72	2011/02/13 03:00	2011/02/13 00:00	- - -
1.08	2011/02/13 06:00	2011/02/13 03:00	2011/02/12 23:00
1.44	2011/02/13 16:00	2011/02/13 07:00	2011/02/13 01:00
1.80	- - -	2011/02/13 13:00	2011/02/13 02:00
2.16	- - -	2011/02/13 14:00	2011/02/13 03:00
2.52	- - -	2011/02/13 17:00	2011/02/13 04:00
2.88	- - -	- - -	2011/02/13 06:00
3.24	- - -	- - -	2011/02/13 08:00

Table 5.7: The time at which a jump in R-value is observed, listed both model-wise and height-wise, T_{rve} indicates the first time-stamp where following a continuous period of null values, a non-null finite R-value output is observed

With regards to R-value emergence, observed as a jump in R-value in the height-wise stack-plots, an OHR may be defined as 'a collection of heights where a clear and sustained jump in R-value is observed' (Biswal et al. 2024). For AR 11158, the OHRs for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$ are 0.36 - 1.44 Mm, 0.36 - 2.52 Mm and 1.08 - 3.24 Mm, respectively (see Figure 5.4).

Similar to AR 11158, the sunspot-group complexity increased for ARs 11166, 12017 and 12673. For AR 11166, the complexity of the sunspot group evolved from a β-type sunspot group on Mar 04, 2011 to a $\beta\gamma$-type sunspot group on Mar 06, 2011. Eventually the sunspot group attained the maximum complexity - a $\beta\gamma\delta$-type, on Mar 08, 2011. We observe that increasing sunspot complexity, together with temporal proximity to the X-class flare, is associated with persistently non-zero R-values at greater heights. For instance, on 05 Mar the highest altitude at which $R_{(150,15)}$ remains continuously non-zero is approximately 0.36 Mm, whereas by 09 Mar non-zero $R_{(150,15)}$ values extend up to 1.44 Mm (see Figure 5.1a). The OHRs for AR 11166 were identified as 0.72 - 1.44 Mm, 1.08 - 1.80 Mm and 1.44 - 2.52 Mm for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$, respectively (see Figure 5.1). This suggests a possible relationship between the temporal changes in the height-dependent R-value and the evolving complexity of the sunspot group.

For AR 12017, the magnetic flux is initially high (see Figure 5.3c during 23-24 Mar 2014), followed by a decline to near-zero levels before rising sharply once again. The sunspot group complexity mirrored the trend in unsigned flux. AR 12017 was a $\beta/\beta\gamma$-type sunspot group initially on 23-24 Mar and it reduced to a β-type sunspot group between 26-28 Mar. However the sunspot group suddenly increased its complexit, transitioning into a $\beta\delta$-type on 29 Mar and eventually to a $\beta\gamma\delta$-type on 30 Mar. Interestingly, the temporal evolution of the R-value closely follows the trends shown by the unsigned flux and sunspot-group complexity. An examination of the photospheric R-value reveals a distinct V-shaped temporal pattern, indicating that the R-value underwent a decay before re-emergence. This feature becomes increasingly prominent

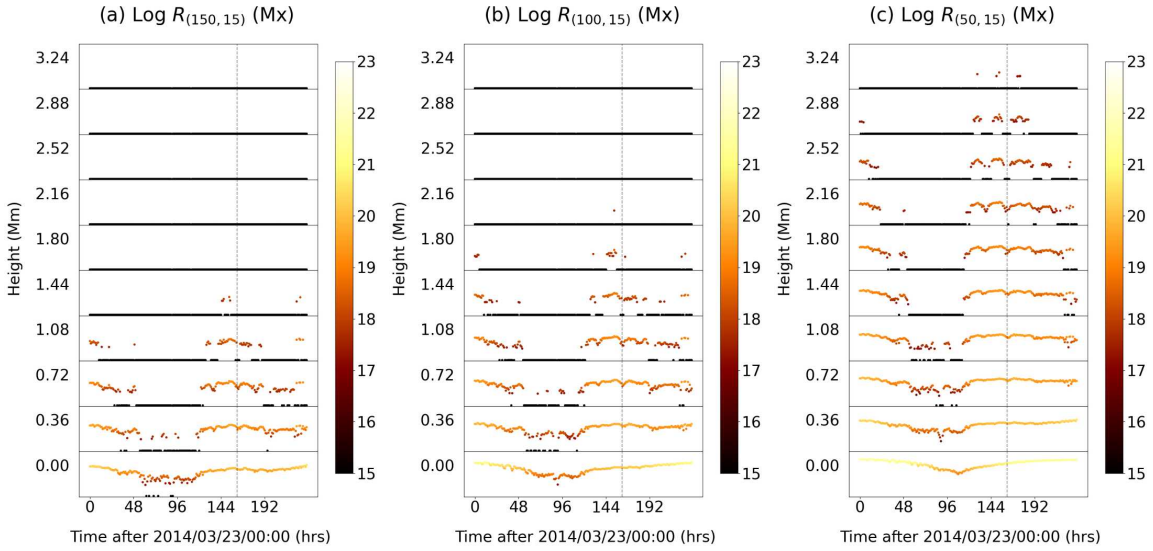

Figure 5.5: *R*-value stack-plots (same format as in Figure 5.4) for AR 12017 after 23 Mar 2014 (00:00 UTC); the OHRs are (a) 0.36 - 1.08 Mm, (b) 0.72 - 1.44 Mm, and (c) 1.08 - 2.16 Mm.

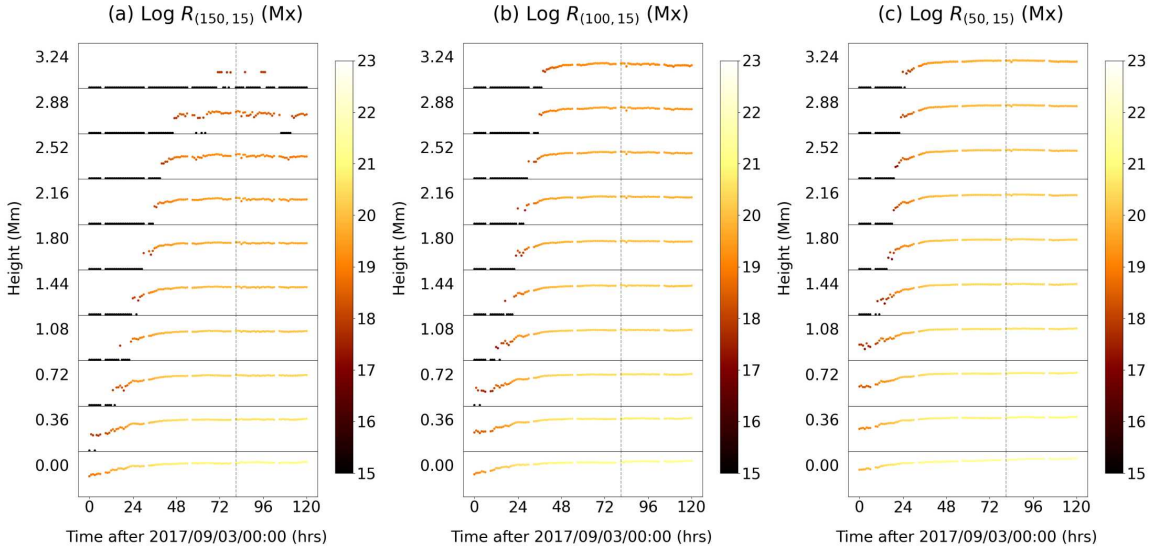

Figure 5.6: *R*-value stack-plots (same format as in Figure 5.4) for AR 12673 after 3 Sep 2017 (00:00 UTC); the OHRs are (a) 0.72 - 2.88 Mm, (b) 1.08 - 3.24 Mm, and (c) 1.44 - 3.24 Mm.

with height, as indicated by the black lines where the *R*-value transitions to zero, such as at 0.36 Mm for $R_{(150,15)}$, at 1.08 Mm for $R_{(100,15)}$ and at 1.44 Mm for $R_{(50,15)}$ (see Figure 5.5). The OHRs for AR 12017 were identified as 0.36 - 1.08 Mm, 0.72 - 1.44 Mm, and 1.08 - 2.16 Mm for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$, respectively (see Figure 5.5).

In the case of AR 12673, one of the most flare-productive ARs of solar cycle 24, 4 X-class flare events were observed, 3 of which occurred within the 60° EW of the disk centre. Prior to the first X-class flare on Sep 6 2017, 12 M-class flares occurred within a 72 hr window. On Sep 3, AR 12673 had a β-type sunspot group. It evolved into a $\beta\gamma$-type sunspot group on Sep 4 and eventually into a $\beta\gamma\delta$-type sunspot group by Sep 6. The unsigned flux increase is quite gradual with T_{fe} being in close temporal proximity to the time of emergence of bipoles (Liu et al. 2019). The R-value once again closely follows the trends shown by the unsigned flux and sunspot-group complexity (see Figure 5.6). The OHRs for AR 12673 were identified as 0.72 - 2.88 Mm, 1.08 - 3.24 Mm, and 1.44 - 3.24 Mm for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$, respectively (see Figure 5.6).

2. Cases with decrease in unsigned flux

It is known that major eruptive solar flares at times could be driven by magnetic flux cancellation (Zhang et al. 2001; Burtseva & Petrie 2013) and it might be possible that the X-class flares related to ARs 11520, 12158 and 12297 which show a decrease in ϕ, (see Figures 5.8a-c) were related to magnetic flux cancellation (see Table 5.8). The decrease in ϕ is not as substantial as the increase in ϕ as seen for flux-emerging ARs. All the three ARs showed a constant $\beta\gamma\delta$-type sunspot group for the 48 hr interval prior to the first flare suggesting a saturation in magnetic complexity and no OHRs could be identified in any of these cases.

Photospheric unsigned Flux ϕ near PILs and $R_{(150,15)}$ (in 10^{20} Mx) at T_{ref} and T_{fo}						
AR	$T_{ref} = T_{fo} - 48$ hr	$\phi\ (T_{ref})$	$R_{(150,15)}$	T_{fo}	$\phi\ (T_{fo})$	$R_{(150,15)}$
11520	2012/07/10 15:00	30.90	1.89	2012/07/12 15:00	25.80	1.47
12158	2014/09/08 17:00	18.94	0.63	2014/09/10 17:00	11.41	0.58
12297	2015/03/09 16:00	40.33	3.33	2015/03/11 16:00	35.19	2.78

Table 5.8: Table comparing changes in $R_{(150,15)}$ and ϕ for flux decreasing ARs on the photosphere; a reference time T_{ref}, is used 48 hr prior to flare onset time T_{fo} to keep a track of ϕ and $R_{(150,15)}$; it may be noted here that the decrease in photospheric ϕ and $R_{(150,15)}$ does not undergo a significant decrease as compared to the increases in these parameters in flux emerging cases

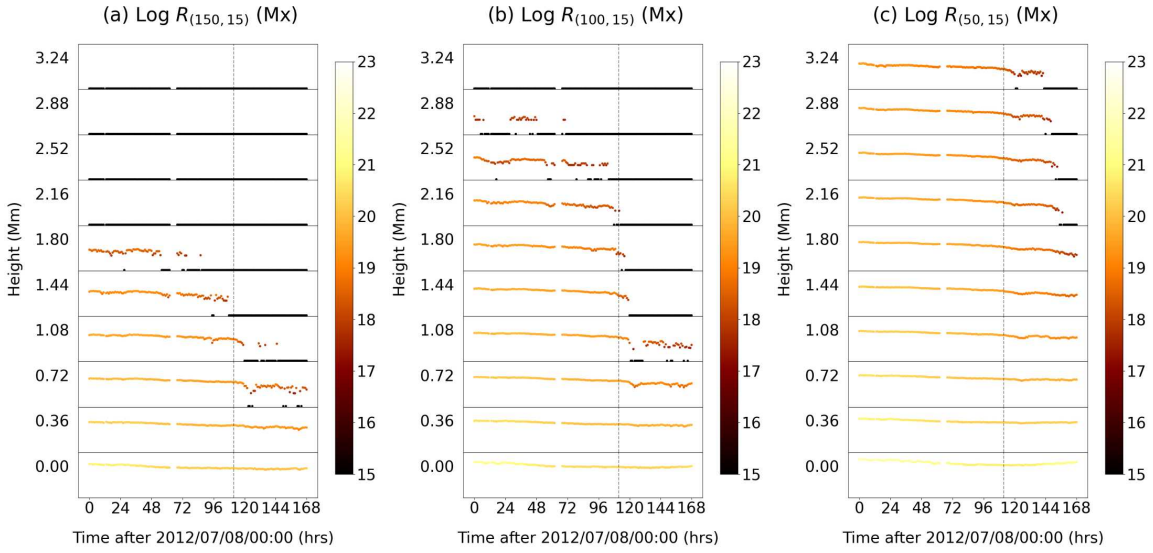

Figure 5.7: R-value stack-plots (same format as in Figure 5.4) for AR 11520 after 8 Jul 2012 (00:00 UTC); no OHRs could be identified in any R-value model

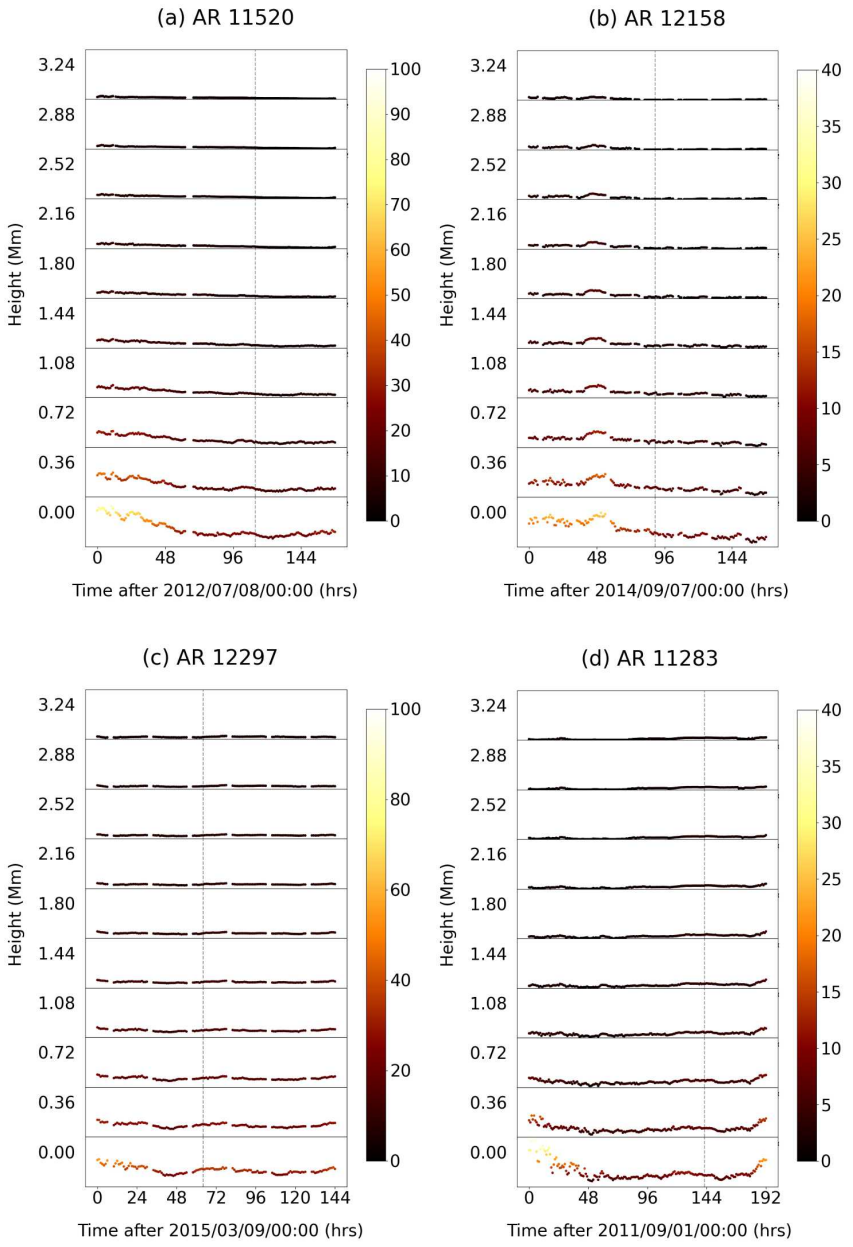

Figure 5.8: Stack-plots showing the height-wise variation of ϕ (in 10^{20} Mx, see colorbar for colour-coding) between 0-3.24 Mm for (a) AR 11520, (b) AR 12158, (c) AR 12297 and (d) AR 11283

3. AR 11283

AR 11283 could not be classified into any of the previous two categories because ϕ remained nearly constant with time (see Figure 5.8d) and the R-value pattern was not very consistent across different values of Bz_{th} and with heights for $R_{(150,15)}$, even though the sunspot-group showed an increase in complexity from β-type between Sep 1-5 2011 to a $\beta\gamma\delta$-type on Sep 7, making it technically a flux-emerging AR. $R_{(150,15)}$ and $R_{(100,15)}$ fluctuate between null and non-null values at $z = 0.72$ Mm and $z = 1.08$ - 1.44 Mm respectively (see Figure 5.9). In case of $R_{(150,15)}$, the height of $z = 1.08$ Mm remains the only height where we can identify a jump in R-value and beyond this we have flat black lines up to 3.24 Mm. In case of $R_{(100,15)}$, if we ignore the instances of R-value fluctuations before Sep 6, 2011 at $z = 1.44$ - 1.80 Mm, these are flat black lines, showing a clear jump around Sep 6, 2011 (00:00 UTC). However, the results for $R_{(50,15)}$ are different from the other models. If we remove the noisy data fluctuations in the height range of 2.16 - 3.24 Mm, the jumps in R-value can be clearly observed. Overall, the OHR for $R_{(50,15)}$ shows a jump in R-value much before $R_{(150,15)}$ and $R_{(100,15)}$.

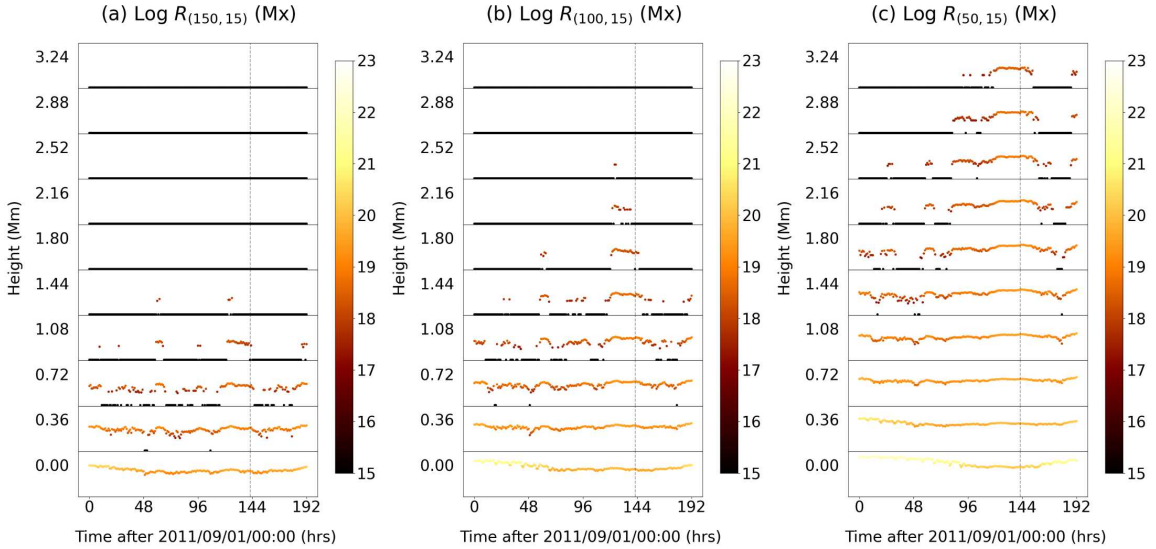

Figure 5.9: R-value stack-plots (same format as in Figure 5.4) for AR 11283 after 1 Sep 2011 (00:00 UTC); an OHR was identified at (a) 1.08 Mm (b) 1.44 - 1.80 Mm and (c) 2.16 - 3.24 Mm for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$, respectively

5.3.3 Analysis of M-class Flaring ARs

Similar to the analysis of X-class flaring ARs, we analysed 4 different M-class flaring ARs - 11620, 11719, 11818 and 12497 (see Table 5.2 for details). Similar to the X-class cases, we analysed the pre-flaring (prior to the first M-class flare) environment in the lower solar atmosphere for these M-class cases with the condition that the ARs hosted a δ-type sunspot at some point.

Of these, ARs 11620, 11818 and 12497 showed an increase in ϕ, indicating flux emergence, while AR 11719 showed a partial decrease in ϕ suggesting possible flux cancellation. These changes in ϕ were also well correlated to the changes R-value (see an example of increase in R-value in Figure 5.10 and an example of decrease in R-value in Figure 5.11). ARs 11620 and 12497 showed an increase in sunspot-group complexity (from a β-type at the start of the observation window to a $\beta\gamma\delta$-type at the time of the first M-class flare) which was closely in line with the R-value and ϕ patterns. AR 11818 showed an initial decrease in sunspot-group complexity, followed by an increase, the same kind of pattern that led to a V-shape R-value pattern in AR 12017. In case of AR 11719, the sunspot-group type remained $\beta\gamma$ for several days before the flare until it was a $\beta\gamma\delta$-type sunspot-group temporarily at the time of flare occurrence.

The overall impression here is that studying the R-value in conjunction with the concept of OHR may not be sufficient in distinguishing between an imminent X-class flare and a M-class flare at a qualitative level directly from the visual inspection of the R-value stack-plots.

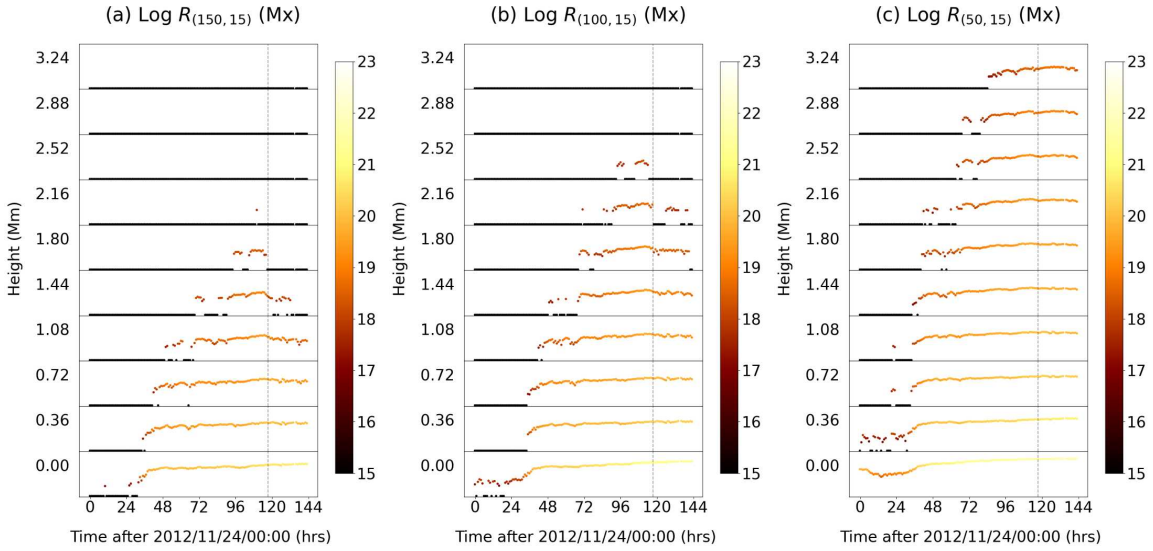

Figure 5.10: *R*-value stack-plots (same format as in Figure 5.4) for AR 11620 after 24 Nov 2012 (00:00 UTC); the vertical dashed line denotes the time of occurrence of the M2.2 flare; an OHR was identified at (a) 0.00 - 1.80 Mm (b) 0.36 - 2.52 Mm and (c) 0.72 - 3.24 Mm for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$, respectively

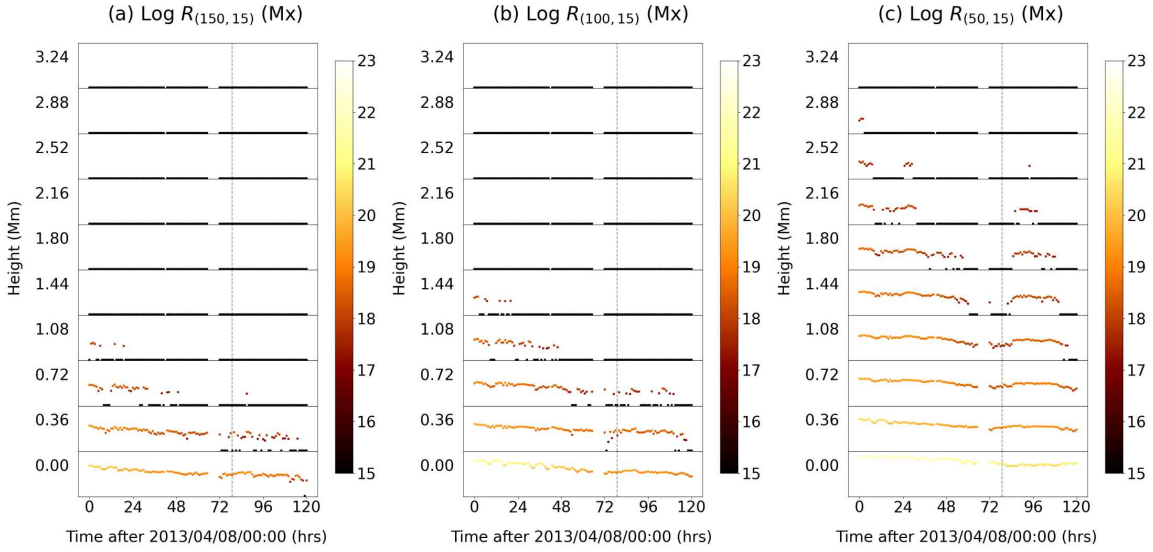

Figure 5.11: *R*-value stack-plots (same format as in Figure 5.4) for AR 11719 after 8 Apr 2013 (00:00 UTC); the vertical dashed line denotes the time of occurrence of the M6.5 flare; an OHR could not be identified for this AR from any of these models

5.3.4 Analysis of Non-flaring Cases

In the previous sections we observed qualitatively that there is a jump in R-value prior to the occurrence of major flares (M/X-class) in flux emerging ARs. In ARs that show slightly decreasing/stable ϕ, the R-value also appears to sustain finite non-null values before dropping down to null values (see Figures 5.7 and 5.11 as examples - no OHR identified for these cases). In this context, we explore if the R-value behaves similarly or differently for non-flaring ARs. Here the connotation for a non-flaring AR is that the AR does not host any flare stronger than C1.0. We studied 3 non-flaring ARs, all of which hosted a β sunspot during the time-window in which they were studied (refer Table 5.3). Similar to X/M-class ARs, we also obtained the height-wise time series data for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$ (see Figures 5.12 and 5.13).

We find that $R_{(150,15)}$ does not consistently sustain non-zero values at heights as low as 0.36 Mm for AR 11143 or barely does so for ARs 11710 and 12064 (see Figure 5.12). This is clearly a distinct behaviour from what we observed for M/X-class flaring ARs where the maximum height at which $R_{(150,15)}$ sustains non-zero values may be as low as 0.72 Mm or as high as 3.24 Mm, or possibly beyond that, given that we have restricted our analysis to 3.24 Mm. This leads us to conclude that the vertical variation of $R_{(150,15)}$ may be used to distinguish between flaring and non-flaring ARs.

Other R-value models such as $R_{(100,15)}$ and $R_{(50,15)}$ yield stack-plots which are consistent with the variation in ϕ (see example of AR 11710 in 5.13) and also are in general, weaker than their counterparts for X/M-class flaring ARs. However, when we discriminate between flaring ARs from non-flaring ARs, the presence or absence of a few high-polarity regions ($B_z > 150$ G) appears to be more of a distinguishing factor rather than presence or absence of a number of relatively moderate polarity regions ($B_z > 50$ G, for example). Hence $R_{(150,15)}$ appears to be more useful as a tool over $R_{(100,15)}$ or $R_{(50,15)}$ as it can more quickly filter out high-polarity regions with height.

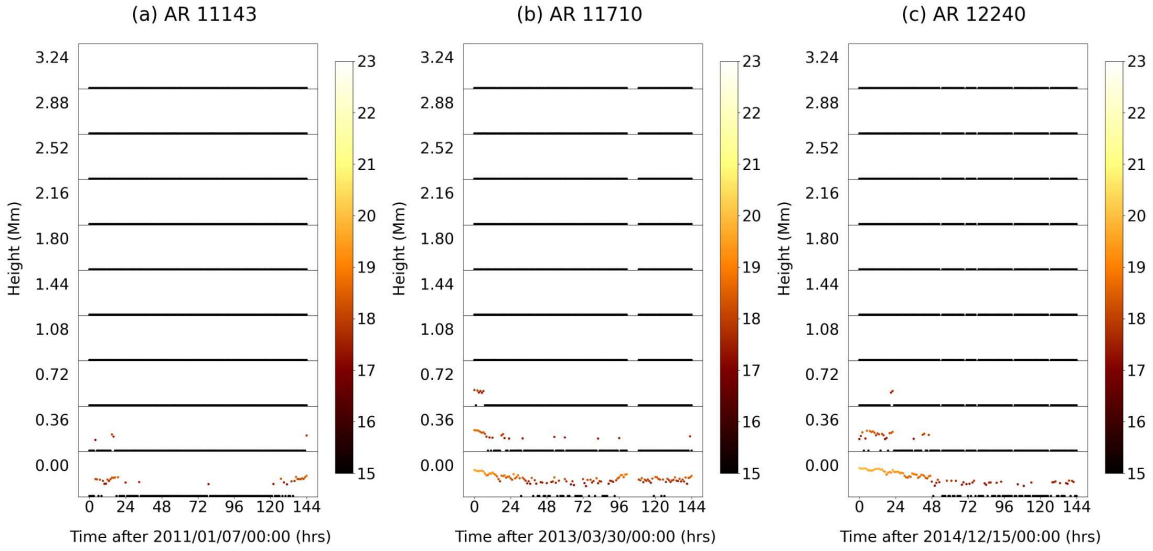

Figure 5.12: $R_{(150,15)}$ stack-plots (same format as in Figure 5.4) for non-flaring ARs

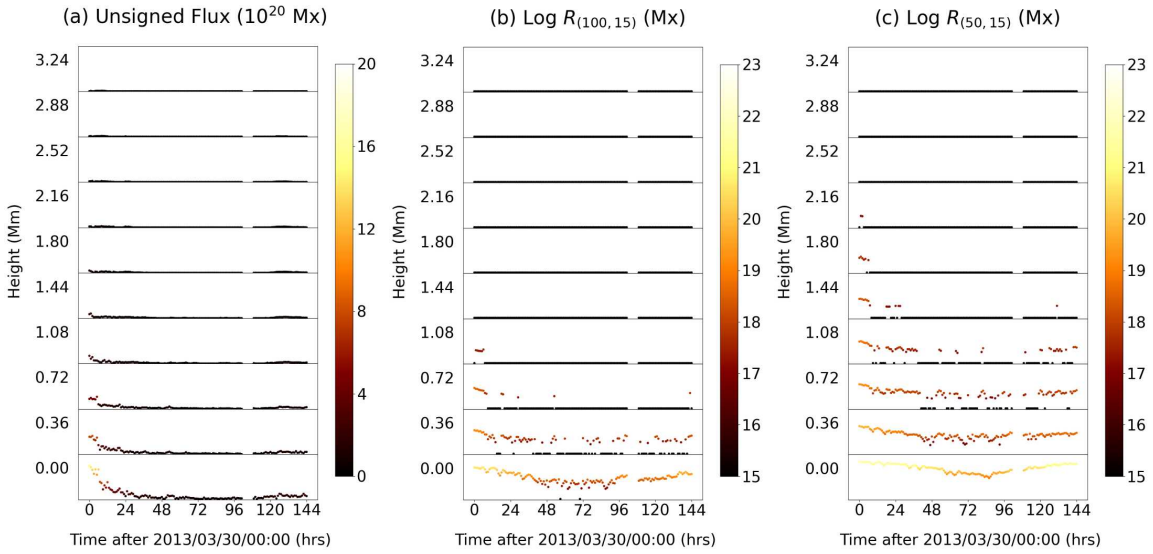

Figure 5.13: ϕ stack-plot and R-value stack-plots (b, c) for AR 11710

5.3.5 Analysis of a C-class Flaring Case: AR 12353

It is quite uncommon for ARs associated with a δ-type sunspot-group to not produce X/M-class flares and just produce C-class flares. Further it is quite hard to find cases that satisfy the AR selection criteria in Section 5.2. As a result, in this section, we discuss only one example - AR 12353, which satisfied the criteria listed in Section 5.2. It produced 3 C-class flares on May 23, 2015; C1.0 (03:27 hrs); C1.1 (07:18 hrs) and C2.3 (17:30 hrs) and hosted a δ-type sunspot group on the same day (Wiśniewska et al. 2024).

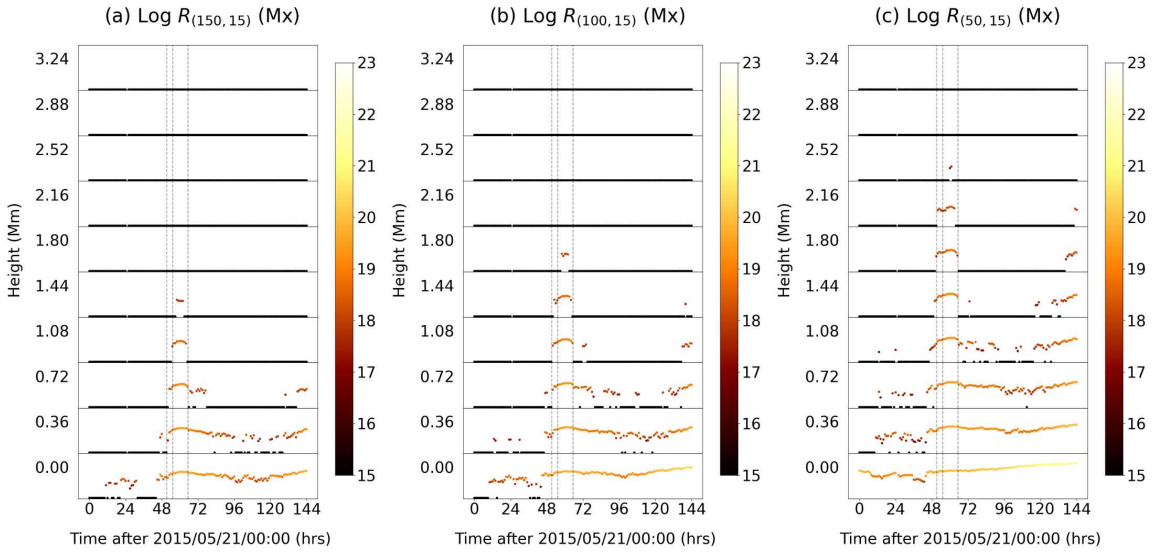

Figure 5.14: R-value stack-plots (same format as in Figure 5.4) for AR 12353 after 21 May 2015

From the stack-plots for AR 12353, an interesting observation is that a spike in R-value at higher altitudes completely envelopes the time duration over which the three C-class flares occurred. In this context, one may specifically look at a height ranges of (i) $z = 0.72$ Mm for $R_{(150,15)}$, (ii) $z = 1.08$ - 1.44 Mm for $R_{(100,15)}$ and (iii) 1.44 - 2.16 Mm for $R_{(50,15)}$ (see Figure 5.14). This observation may be explained by the idea that the high polarity regions that are associated with flares are better detected at the aforementioned height-ranges compared to the photosphere because other high-polarity

regions detected on the photosphere which do not contribute to flares get filtered off at these heights. Once the configuration that probably supported the production of C-class flares got disturbed after the C2.3 flare, the high polarity regions did not get detected anymore and the R-value may have come down to the null-output state once again, as a direct consequence.

5.3.6 Implications on Flare Predictability

So far, we were able to identify an OHR qualitatively for (i) X-class flaring ARs linked to flux emergence, i.e. ARs 11158, 11166, 12017 and 12673 and the special case of AR 11283 and (ii) M-class flaring ARs linked to flux emergence i.e ARs 11620, 11818 and 12497. The details of the identified OHRs in each case are tabulated in Table 5.9. However, in this section, we move forward to a more quantitative estimate of how much time an identified OHR leaves us to respond to a flare. In this connection we define the lead time T_{diff} i.e. the time difference between (i) T_{fo} i.e. the latest timestamp preceding flare onset in our dataset and (ii) T_{rve} i.e. the time of jump in R-value (or R-value emergence). For every AR studied, each height and R-value model may have a distinct T_{rve}. Thus T_{diff} is a function of height within the OHR (see details in Table 5.10). We plotted T_{diff} with height (see Figure 5.15) and colour-coded the lead time with flare strength (X-ray flux in 10^{-4} Wm^{-2}) to check if any patterns exist that may connect flare strength with lead time.

No.	AR	OHR $[R_{(150,15)}]$	OHR $[R_{(100,15)}]$	OHR $[R_{(50,15)}]$
01	11158	0.36 - 1.44 Mm	0.36 - 2.52 Mm	1.08 - 3.24 Mm
02	11166	0.72 - 1.44 Mm	1.08 - 1.80 Mm	1.44 - 2.52 Mm
03	11283	1.08 Mm	1.44 - 1.80 Mm	2.16 - 3.24 Mm
04	11520	- - -	- - -	- - -
05	12017	0.36 - 1.08 Mm	0.72 - 1.44 Mm	1.08 - 2.16 Mm
06	12158	- - -	- - -	- - -
07	12297	- - -	- - -	- - -
08	12673	0.72 - 2.88 Mm	1.08 - 3.24 Mm	1.44 - 3.24 Mm
09	11620	0.00 - 1.80 Mm	0.36 - 2.52 Mm	0.72 - 3.24 Mm
10	11719	- - -	- - -	- - -
11	11818	0.72 - 1.80 Mm	1.08 - 2.52 Mm	1.44 - 3.24 Mm
12	12497	0.72 - 1.08 Mm	1.08 - 2.16 Mm	1.80 - 3.24 Mm

Table 5.9: List of OHRs for different R-value models for X-class (1-8) and M-class (9-12) flare cases

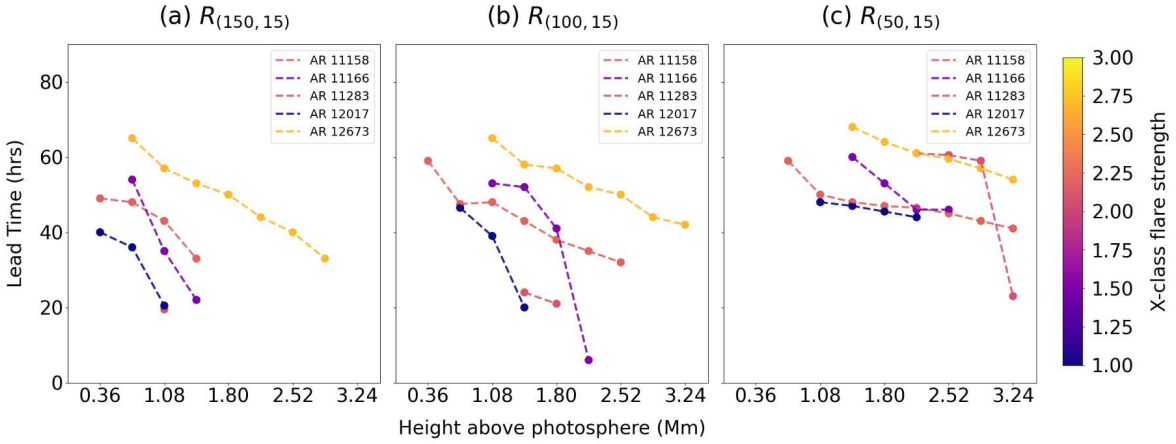

Figure 5.15: Lead time T_{diff} vs height (z) plots for the five X-class flaring ARs with OHR; (a) $R_{(150,15)}$, (b) $R_{(100,15)}$ and (c) $R_{(50,15)}$

From the T_{diff} vs z plots, we find that lead time decreases with increase in height above the photosphere. However, the nature or rate of decrease may vary from case to case. The data suggests that flare strength may not be well correlated with T_{diff}, although this requires reviewing with a much larger statistical dataset. For example, consider ARs 11158 and 11166 in Figure 5.15a which intersect at heights between 0.72 Mm and 1.08 Mm. While at $z = 0.72$ Mm the weaker X1.5 flare belonging to AR 11166, is associated with a higher lead time, at $z = 1.08$ Mm the stronger X2.2 flare belonging to AR 11158 is associated with a higher lead time.

Now that we have a quantitative estimate of T_{diff} as a function of height z, we take the maximum value of T_{diff} across different heights (denoted as T^*_{diff}) and tabulate it for different R-value models, varying as a function of Bz_{th}. In simpler terms, every AR shall now have three different T^*_{diff}, each corresponding to a R-value model (see Table 5.11). The maximum value of T^*_{diff} gives T_{opt}, i.e. the optimal time (optimised in terms of height and Bz_{th}) available to respond to a flare, and in some sense a theoretical maximum available time to respond to a flare. However these calculations for T_{opt} are done purely in retrospect, and in real time flare prediction, when one encounters a jump in R-value, it may not always signal a major flare and there is a chance that it could be

noise. This is why when we defined the OHR for R-value before, we emphasised that the jump in R-value should be a "sustained" jump.

To distinguish sustained jumps from noisy ones in a real time flare prediction scenario, we enforce a subjective condition - the R-value must sustain non-null outputs for at least 24 hr (we call this a confidence interval or C_{int}) after an increase so that a flare warning may be issued with a certain degree of confidence. It is within this confidence interval one might cross-validate the inputs received from the optimal R-value model with the other ones. If for some AR, the C_{int} validates the presence of a jump in R-value, a flare prediction alert may be issued immediately. Hence the actual time agencies would have in hand to respond to a flare alert would be $T_{res} = T_{opt} - C_{int}$.

Example

Let us consider the case of AR 11166. Prior to the X1.5 flare on Mar 9, 2011 (23:00 UTC), $R_{(50,15)}$ showed the earliest spike, as early as 60 hr (T_{opt}) before the flare i.e. at 11:00 UTC on Mar 7, 2011. The jump in $R_{(50,15)}$ is monitored for 24 hr (C_{int}) until 11:00 UTC on Mar 8, 2011. Within this time-frame it is seen that $R_{(100,15)}$ and $R_{(150,15)}$ also undergo a sustained jump for 17 and 18 hr respectively. If in real time, a flare warning were to be issued at the end of the confidence interval at 11:00 UTC on Mar 7, 2011, that would leave $T_{res} = 36$ hr.

From the analysis of T_{opt} (see Table 5.11), we find that $R_{(50,15)}$ is more likely to maximise T_{opt}. Due to a weaker threshold (less Bz_{th} compared to other models), $R_{(50,15)}$ at the initiation time of flux emergence may be detecting the emergence of weaker B_z values before stronger B_z values begin to dominate flux emergence. We find that T_{opt} may vary between 48 - 68 hr for X-class flaring and flux emerging ARs, leaving a T_{res} of 24 - 44 hr. In case of M-class flaring and flux ARs, T_{opt} varies between 18 - 97 hr.

Approximate time between jump in $R_{(150,15)}$ and flare occurrence time											
No.	AR	0.00	0.36	0.72	1.08	1.44	1.80	2.16	2.52	2.88	3.24
01	11158	- - -	**49 hr**	48 hr	43 hr	33 hr	- - -	- - -	- - -	- - -	- - -
02	11166	- - -	- - -	**54 hr**	35 hr	22 hr	- - -	- - -	- - -	- - -	- - -
03	11283	- - -	- - -	- - -	**20 hr**	- - -	- - -	- - -	- - -	- - -	- - -
04	12017	- - -	**40 hr**	36 hr	20 hr	- - -	- - -	- - -	- - -	- - -	- - -
05	12673	- - -	- - -	**65 hr**	57 hr	53 hr	50 hr	44 hr	40 hr	33 hr	- - -
06	11620	**85 hr**	80 hr	71 hr	48 hr	26 hr	22 hr	- - -	- - -	- - -	- - -
07	11818	- - -	- - -	**54 hr**	52 hr	47 hr	40 hr	- - -	- - -	- - -	- - -
08	12497	- - -	- - -	**18 hr**	3 hr	- - -	- - -	- - -	- - -	- - -	- - -

Approximate time between jump in $R_{(100,15)}$ and flare occurrence time											
No.	AR	0.00	0.36	0.72	1.08	1.44	1.80	2.16	2.52	2.88	3.24
01	11158	- - -	**59 hr**	47 hr	48 hr	43 hr	38 hr	35 hr	32 hr	- - -	- - -
02	11166	- - -	- - -	- - -	**53 hr**	52 hr	41 hr	6 hr	- - -	- - -	- - -
03	11283	- - -	- - -	- - -	- - -	**24 hr**	21 hr	- - -	- - -	- - -	- - -
04	12017	- - -	- - -	**47 hr**	39 hr	20 hr	- - -	- - -	- - -	- - -	- - -
05	12673	- - -	- - -	- - -	**65 hr**	58 hr	57 hr	52 hr	50 hr	44 hr	42 hr
06	11620	- - -	**82 hr**	82 hr	72 hr	49 hr	48 hr	26 hr	23 hr	- - -	- - -
07	11818	- - -	- - -	- - -	**54 hr**	52 hr	48 hr	47 hr	43 hr	- - -	- - -
08	12497	- - -	- - -	- - -	**7 hr**	3 hr	1 hr	0 hr	- - -	- - -	- - -

Approximate time between jump in $R_{(50,15)}$ and flare occurrence time											
No.	AR	0.00	0.36	0.72	1.08	1.44	1.80	2.16	2.52	2.88	3.24
01	11158	- - -	- - -	- - -	**50 hr**	48 hr	47 hr	46 hr	45 hr	43 hr	41 hr
02	11166	- - -	- - -	- - -	- - -	**60 hr**	53 hr	46 hr	46 hr	- - -	- - -
03	11283	- - -	- - -	- - -	- - -	- - -	- - -	**61 hr**	60 hr	59 hr	23 hr
04	12017	- - -	- - -	- - -	**48 hr**	47 hr	46 hr	44 hr	- - -	- - -	- - -
05	12673	- - -	- - -	- - -	- - -	**68 hr**	64 hr	61 hr	60 hr	57 hr	54 hr
06	11620	- - -	- - -	**83 hr**	82 hr	78 hr	76 hr	53 hr	49 hr	37 hr	32 hr
07	11818	- - -	- - -	- - -	- - -	**97 hr**	90 hr	67 hr	52 hr	51 hr	49 hr
08	12497	- - -	- - -	- - -	- - -	- - -	**18 hr**	7 hr	7 hr	4 hr	3 hr

Table 5.10: Lead times T_{diff} listed at all heights (in Mm) for different ARs computed using $R_{(150,15)}$, $R_{(100,15)}$, and $R_{(50,15)}$; bold text indicates T^*_{diff}, the maximum lead time across heights for the AR.

No.	AR	Flare	T_{fo}	$T^*_{diff}(150)$	$T^*_{diff}(100)$	$T^*_{diff}(50)$	T_{opt}	T_{res}
01	11158	X2.2	2011/02/15 01:00	49 hr	**59 hr**	50 hr	59 hr	35 hr
02	11166	X1.5	2011/03/09 23:00	54 hr	53 hr	**60 hr**	60 hr	36 hr
03	11283	X2.1	2011/09/06 22:00	20 hr	24 hr	**61 hr**	61 hr	37 hr
04	11520	X1.4	2012/07/12 15:00	- - -	- - -	- - -	- - -	- - -
05	12017	X1.0	2014/03/29 17:00	40 hr	47 hr	**48 hr**	48 hr	24 hr
06	12158	X1.6	2014/09/10 17:00	- - -	- - -	- - -	- - -	- - -
07	12297	X2.1	2015/03/11 16:00	- - -	- - -	- - -	- - -	- - -
08	12673	X2.2	2017/09/06 08:00	65 hr	65 hr	**68 hr**	68 hr	44 hr
09	11620	M2.2	2012/11/28 21:00	**85 hr**	82 hr	83 hr	85 hr	61 hr
10	11719	M6.5	2013/04/11 06:00	- - -	- - -	- - -	- - -	- - -
11	11818	M3.3	2013/08/17 18:00	54 hr	54 hr	**97 hr**	97 hr	73 hr
12	12497	M1.0	2016/02/12 10:00	**18 hr**	7 hr	**18 hr**	18 hr	-6 hr

Table 5.11: A list of optimal times (indicated in bold) and response times for each OHR for X-class flare cases (1-8) and M-class flare cases (9-12); column 4 lists the latest time-stamp preceding the flare T_{fo}; Columns 5-7 list the maximum lead time across all heights T^*_{diff} for $R_{(150,15)}$, $R_{(100,15)}$ and $R_{(50,15)}$ respectively. The maximum T^*_{diff} across multiple R-value models is T_{opt} and is listed in column 8. Column 9 lists the response time T_{res} after subtracting the confidence interval C_{int} from T_{opt}. For exact information on T_{diff} and T^*_{diff}, refer to Tables 5.10; the negative value for T_{res} in the case of AR 12497 is because the jumps in R-value are seen less than 24 hr before T_{fo}

5.4 Summary

The key finding of this chapter is that R-value can be used as a discriminant between flaring and non-flaring ARs. It is seen that the $R_{(150,15)}$ attenuates to null-outputs faster in height for non-flaring ARs compared to flaring ARs.

From the case studies discussed in this chapter, in case of X-class flaring ARs, a jump in the R-value may been as early as 48 - 68 hr before the first X-class of emerging ARs and the height range of 0.36 - 3.24 Mm behaves as the OHR to study R-value. Similarly, a jump in the R-value may be seen as early as 18 - 97 hr before the first M-class flare. In this context, we note that the lead time prior to a flare decreases with height for a given AR.

Following this approach, a statistical study of X/M-class flaring ARs using inputs from $R_{(50,15)}$, $R_{(100,15)}$ and $R_{(150,15)}$, could be useful to refine our understanding of the range/spread of the lead times.

Chapter 6

Time Series Analysis of PIL length

6.1 Objective

In the previous chapter, we saw that the evolution of R-value has the potential to detect the first major flare of an emerging AR. Since R-value and ϕ are closely connected to PILs, our objective for this chapter remains to assess if the analysis of the time series data for PILs could also help detect major flares. In this regard, the simplest (and the most intuitive) parameters to analyse are those representing PIL length and thus for every AR (PF-extrapolated between 0 - 3.24 Mm), we analyse (i) L_{tot} and (ii) L_{max}. While the detection of major flares through R-value time series data is aided by the presence of discontinuities in the data, the data for PIL length (obtained from FLARECAST (Georgoulis et al. 2021) or the publicly available dataset of (Ji et al. 2023)) is relatively smoother, especially for flaring ARs (see Figure 6.1 and stack-plot samples in Section 6.3.8). We use autoregressive models to probe the memory structure in data and see if the previous points in the time series also influence the current point in the dataset. Similar to the qualitative analysis made in case of R-value case studies, we also check if the evolution of L_{tot} or L_{max} may help distinguish between flaring or non-flaring ARs.

6.2 Dataset

Similar to our analysis of the R-value, we selected the same 12 flaring ARs (8 X-class and 4 M-class) as listed in Section 5.2. The selection criteria for ARs, regarding the (i) AR hosting a δ-type sunspot group and (ii) being within 60° EW of the solar central meridian, are the same as discussed in Chapter 5. Further, due to the property of persistence, exhibited by flares, we only analyse the time series data for the 48 hr prior to the first major flare of the AR (X/M-class depending upon case). Thus, the only major point of difference between the datasets in Chapters 5 and 6 relates to the observation interval of the studied AR (see details in Table 6.1).

In addition to the flaring ARs, we also studied nine 48-hr intervals of 3 non-flaring ARs within 60° EW, which were not associated with δ-type sunspots (see details in Table 6.1). We also checked if the PIL length estimates obtained using FLARECAST codes (dataset A or d_A) were consistent with another PIL detection algorithm. We chose to compare our data with a readily available open-source data at 12 min cadence (dataset B or d_B) published by Ji et al. (2023), based on the algorithm of Cai et al. (2020). The fundamental differences in the PIL detection algorithms corresponding to the datasets were discussed elaborately previously in Section 4.2.4. Since d_B is not subject to feature extraction at extrapolated heights and does not contain information on L_{max}, the comparison of the datasets is restricted only to the photospheric values of L_{tot}. The complete dataset and results can be accessed at the GitHub Project Repository.[2]

Observed Data						
No.	AR No.	Class	T_{start}	T_{end}	$T_{flare\ onset}$	PCC (d_A, d_B)
01	11158	X2.2	2011/02/13 01:00	2011/02/15 01:00	2011/02/15 01:44	0.52
02	11166	X1.5	2011/03/07 23:00	2011/03/09 23:00	2011/03/09 23:13	0.57
03	11283	X2.1	2011/09/04 22:00	2011/09/06 22:00	2011/09/06 22:12	0.30
04	11520	X1.4	2012/07/10 15:00	2012/07/12 15:00	2012/07/12 15:37	0.20
05	12017	X1.0	2014/03/27 17:00	2014/03/29 17:00	2014/03/29 17:35	0.77
06	12158	X1.6	2014/09/08 17:00	2014/09/10 17:00	2014/09/10 17:21	0.84
07	12297	X2.1	2015/03/09 16:00	2015/03/11 16:00	2015/03/11 16:11	0.78
08	12673	X2.2	2017/09/04 09:00	2017/09/06 09:00	2017/09/06 08:57	0.20
09	11620	M2.2	2012/11/26 21:00	2012/11/28 21:00	2012/11/28 21:20	0.82
10	11719	M6.5	2013/04/09 06:00	2013/04/11 06:00	2013/04/11 06:55	0.75
11	11818	M3.3	2013/08/15 18:00	2013/08/17 18:00	2013/08/17 18:16	0.26
12	12497	M1.0	2016/02/10 10:00	2016/02/12 10:00	2016/02/12 10:36	0.41
13	11143-1	- - -	2011/01/07 00:00	2011/01/09 00:00	- - -	- - -
14	11143-2	- - -	2011/01/09 00:00	2011/01/11 00:00	- - -	- - -
15	11143-3	- - -	2011/01/11 00:00	2011/01/13 00:00	- - -	- - -
16	11710-1	- - -	2013/03/30 00:00	2013/04/01 00:00	- - -	- - -
17	11710-2	- - -	2013/04/01 00:00	2013/04/03 00:00	- - -	- - -
18	11710-3	- - -	2013/04/03 00:00	2013/04/05 00:00	- - -	- - -
19	12240-1	- - -	2014/12/15 00:00	2014/12/17 00:00	- - -	- - -
20	12240-2	- - -	2014/12/17 00:00	2014/12/19 00:00	- - -	- - -
21	12240-3	- - -	2014/12/19 00:00	2014/12/21 00:00	- - -	- - -

Table 6.1: Table listing the details of the studied flaring (1-12) and non-flaring (13-21) ARs, GOES flare classes, observation time window and the time of flare occurrence; the final column shows the PCC between d_A and d_B, calculated for L_{tot} at the photosphere; non-flaring ARs were studied in three distinct 48-hr intervals each; full data and results are available at: [2] https://github.com/shreeyesh-biswal/PILs_3D

Comparison of Datasets

For the flaring ARs, the PCC (Pearson Correlation Coefficient) ranges between 0.20-0.84, suggesting that the correlation between d_A and d_B may be strong or negligible, varying from case to case. We excluded non-flaring ARs from the PCC calculations because at several instances of time, no PILs were detected and this could have led to erroneous results. In Figure 6.1, we show the comparison between datasets A and B using 2 examples each from X-class and M-class flaring cases. For example, if we observe AR 11818 (Figure 6.1c), we find that both the datasets are probably separated by a fixed factor because of which the PCC is really good but the datasets still have a considerable difference. The PCC is a linear correlation coefficient and it identifies trends but does not recognise differences by a fixed linear factor. Hence, it is important to have a look in the actual values and differences while checking whether two datasets agree on the estimate. Further if we look at AR 11158 (Figure 6.1a), it appears to us that the data are separated by a fixed difference even though some of the trends in time are similar. ARs 12673 and 11818 are examples (see Figures 6.1b & 6.1d) of poor correlation and highlight the fact that changing the control parameters to detect PILs eventually do not always guarantee the same outcome. The fact that the algorithm of Cai et al. (2020) does not take into account the horizontal magnetic field distribution (or any metric based on it) as a control parameter may be a reason why it consistently underestimates PIL lengths compared to the FLARECAST codes (refer to the discussion in 4.2.4). Moving forward, we only use the FLARECAST codes for time series analysis they are available for both L_{tot} and L_{max}. These parameters are further computed in the height range of 0 - 3.24 Mm after PF extrapolation.

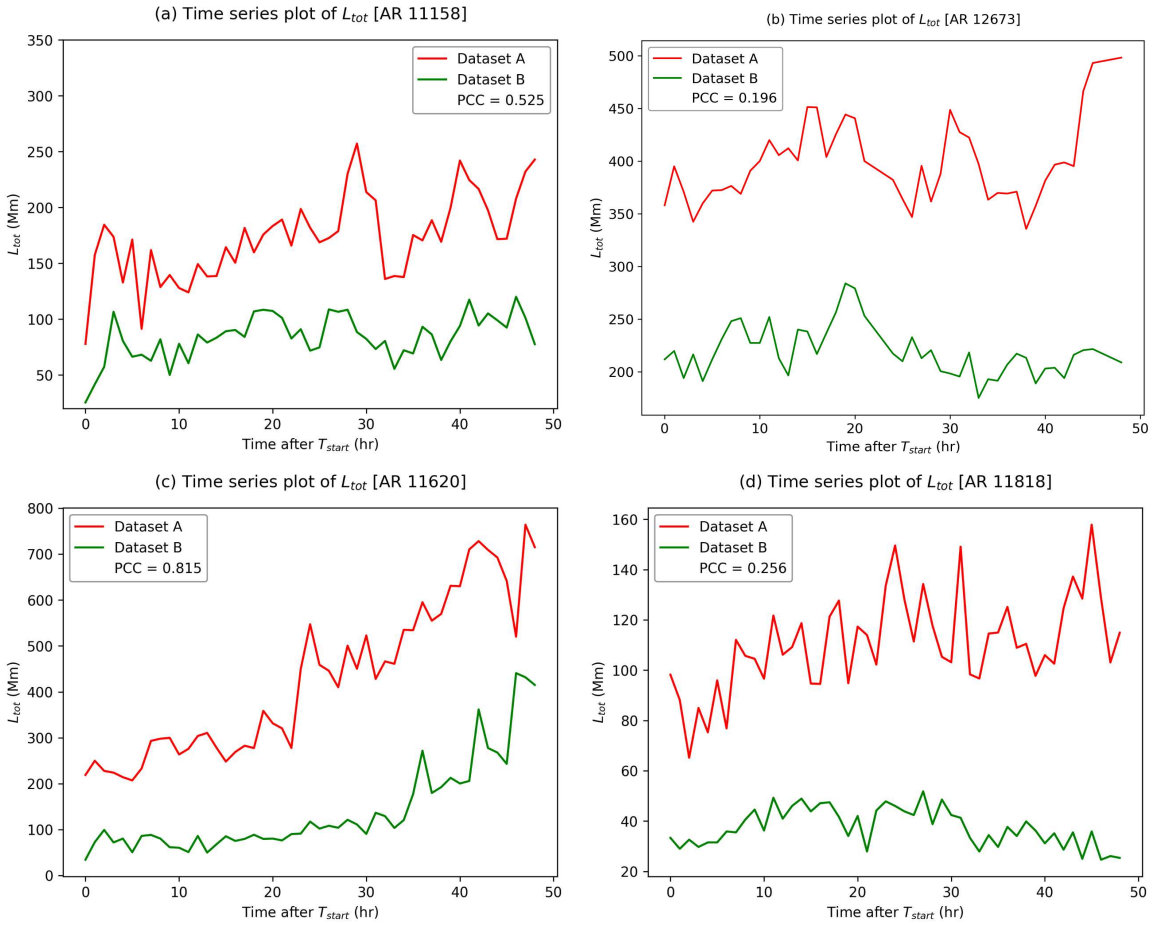

Figure 6.1: Simultaneous time series plots for L_{tot} on the photosphere obtained at a cadence of 1 hr from dataset A (red) and dataset B (green); PCC values are shown in the box

6.3 Results and Discussion

An outline of the step-wise mathematical modelling of the time series data (L_{tot}/L_{max}) is presented with the help of an example - AR 12158, in Section 6.3.1. The following Sections between 6.3.2 - 6.3.5 present the full results for all the analysed ARs and check for the presence of any trends/patterns in time which may help distinguish flaring from non-flaring ARs.

Normally while studying a stochastic dataset using autoregressive models, the dataset is detrended if any underlying functional time trend is suspected to be present. While this has been the approach adopted for Sections 6.3.1 - 6.3.5, Section 6.3.6 is more of a curiosity exercise where we have deliberately ignored the presence of functional time trends and treated the raw data entirely as an autoregressive model.

Further discussion on the implications of this study for flare predictability in the presence of magnitude-based thresholds is discussed in Section 6.3.7.

Finally, different time series data are visualised in the form of stack-plots and a very small-scale magnitude analysis of L_{tot} variation with height for flaring and non-flaring ARs is presented along with it in Section 6.3.7.

6.3.1 An Example of Time Series Modelling: AR 12158

1. Overview

Prior to the mathematical time series modelling of ARs, we ensured that all instances of missing data were removed with the help of linear interpolation. We assume that the time series of L (denoting L_{tot} or L_{max}) can be split into two parts, (i) a time dependent deterministic trend $f(t)$ - to account for the dependency of PIL parameters with time and (ii) a partially stochastic part $g(t)$ that is modelled as an autoregressive process (see Section 3.1). This decomposition of time series data is formulated as:

$$L = f(t) + g(t); \quad g(t) = Y_t = \sum_{i=1}^{p} \psi_i Y_{t-i} + \epsilon_t \tag{6.1}$$

$g(t)$ can be further decomposed into a purely deterministic autoregression process (see the summation term in Equation 6.1) and a purely stochastic white noise ϵ_t, to account for the random changes in the PIL parameters, essentially the purely stochastic white-noise component in the data. In this data modelling process, we are essentially isolating the purely stochastic ϵ_t from the deterministic part of the model.

2. Determining the functional time trend

To find out the time varying deterministic part i.e. $f(t)$, the time series data on the photosphere are fitted with six types of curve fitting approaches - (i) linear: $y = a_0 + a_1 t$, (ii) quadratic: $y = a_0 + a_1 t + a_2 t^2$, (iii) power law: $y = at^b + c$, (iv) logarithmic: $y = a\ ln(t) + c$, (v) exponential: $y = ae^{bt} + c$ and (vi) sinusoidal: $y = a\ sin(\omega t + b) + c$. Among these curve fits, only the sinusoidal curve fitting approach looks for seasonality or periodicity in the data, an example of which is shown in Figure 6.2h. For AR 12158, we found the sinusoidal fit to be the best possible curve fit $f(t)$. While finding the best $f(t)$, we take into account several practical considerations, enumerated below:

1. The curve fit must not yield a negative value estimate for the PIL parameter as PIL length is always positive or zero.

2. The coefficient of determination R^2, of the best model fit must be greater than 0.35 i.e. the fitted equation must be able to account for more than 35% of the variance of the data (see Equation 6.2: y_i, $\hat{y}_i$, $\bar{y}$ are the data point, estimate value and mean value, respectively).

3. The best fit should not be an approximation of other types of curve fits. For example, one may be able to represent a straight line as an approximation of an exponential curve (see Figure 6.2e - a linear fit would be a better choice compared to the exponential curve), or a quadratic curve may be expressed as an approximation of a sinusoidal curve.

4. While fitting exponential and power law curves, we checked if the addition of a constant improves the fit (see Figures 6.2c-f). In certain cases such as exponential decay, where visual inspection suggests that the PIL length may decay to zero as time approaches infinity, the constant may be set to zero as a natural constraint.

5. In order to confirm the presence of seasonality in the data we looked for the presence of at least one (or nearly one) complete cycle of the sinusoidal curve.

6. In case of multiple competing models, we made use of the reduced coefficient of determination R^2_{red} to ascertain the best fit for $f(t)$ as R^2_{red} imposes a penalty for extra terms in the curve fitting equation (see Equation 6.3: N and P are the sample size and number of independent variables, respectively).

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \tag{6.2}$$

$$R_{adj} = 1 - (1 - R^2) \left[\frac{N - 1}{N - P - 1} \right] \tag{6.3}$$

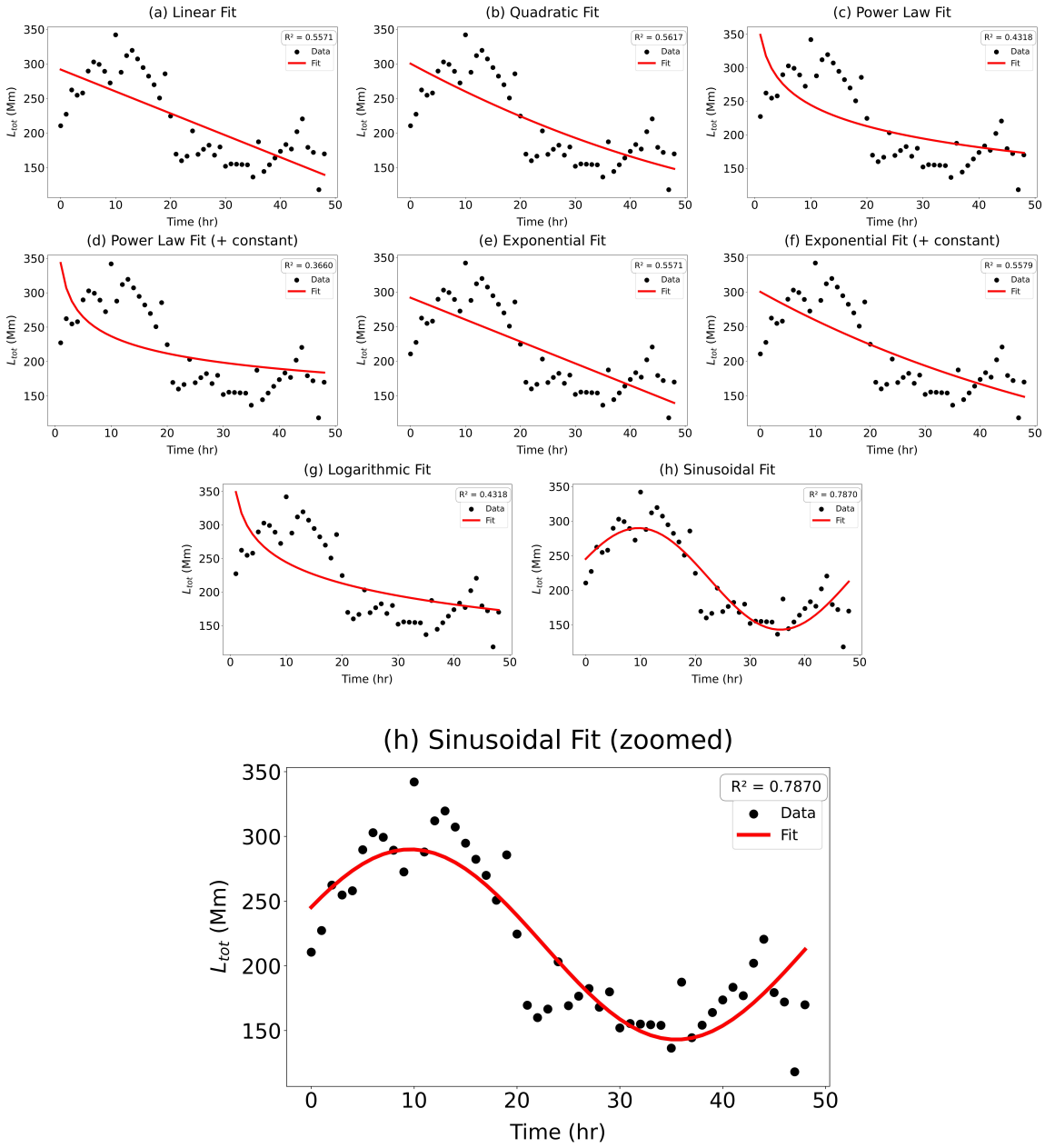

Figure 6.2: Different type of curve fits for the functional trend $f(t)$ (in red); notice that the sinusoidal fit (plot h) is the best possible curve fit for $f(t)$ in case of AR 12158

3. Determining the autoregressive model component

Once $f(t)$ is determined, $g(t)$ can be obtained by subtracting the best curve fit $f(t)$ from the data L (L_{max} or L_{tot}). In case, no time-varying deterministic trend $f(t)$ is found, the entire data L is treated as $g(t)$. Before selecting an autoregressive model to explain $g(t)$, it is important to first determine if autoregressive models are relevant in the first place. To do so we begin by carrying out an Auto Correlation Function (ACF) test, which is a tool to measure the correlation between a series and its past values at different lags. If the ACF of $g(t)$ does not show any significant spikes, then it simply means that $g(t)$ lacks an autocorrelation structure and is essentially white noise. In such a case, the input data can be thought of as a curve fit plus white noise i.e. $L = f(t) + \epsilon_t$, in mathematical terms. However, if significant spikes are observed, then $g(t)$ has autocorrelation that needs to be accounted for by an autoregressive model. For example, in Figure 6.3c, it may be observed that the ACF of $g(t)[L_{tot}]$ for AR 12158 has a significant spike (beyond the shaded blue region) at $n = 1$. This means that an autoregressive model is necessary to model AR 12158. In Sections 3.1.3 and 3.1.4, we saw how AR models can be re-parametrised and represented in the ADF regression equation form. If $g(t)$ has autocorrelation, we express the $g(t)$ in the form of an ADF regression equation. Doing so, also helps obtain the following, additionally: (i) getting the order p, for an autoregressive process $A(p)$, where $p = m + 1$ and with (ii) the detection of stationarity with m lags using the Augmented Dickey Fuller (ADF) test (Dickey & Fuller 1979, Said & Dickey 1984, Elliott et al. 1996; see Giannattasio et al. 2019, Giannattasio, F. et al. 2022 for applications).

4. Determining the lag number in the ADF regression equation

If the ACF test of $g(t)$ indicates the presence of significantly correlated lags, the first step is to find out the most suitable number of lags m to model the differenced equation for $g(t)$. In theory, a differenced regression equation with m lags consists of $m+2$

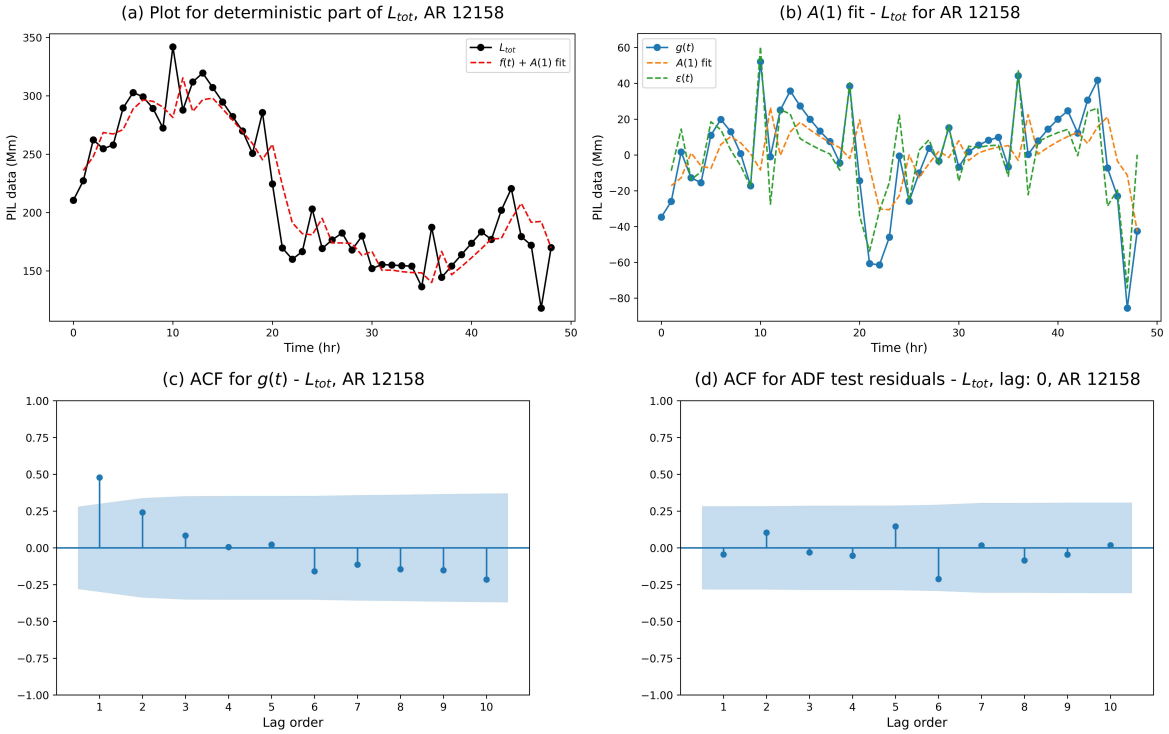

Figure 6.3: Plots describing L_{tot} evolution for AR 12158; (a) a plot of the deterministic part comprising of a sinusoidal curve fit and an $A(1)$ model; (b) a plot of correlated residuals (in blue) after fitting the sinusoidal curve, along with its decomposition into an $A(1)$ model (in orange) and purely stochastic noise component (in green); (c) ACF plot of the correlated residuals $g(t)$; the plot shows one significant spike and which is an evidence of serial correlation in $g(t)$; (d) ACF plot of ADF test residuals (also equivalent to ϵ_t); the plot shows no significant spikes implying that the $A(1)$ model fit is appropriate and the stationary output of the ADF test is reliable

independent variables. Since we have $N = 49$ data points in each time series, we would expect the number of independent variables to not exceed $N/10 \sim 5$ as a thumb rule for small samples. Using more terms (say a $A(8)$ process) to model the data increases the risk of over-fitting and multicollinearity. Thus we expect the number of lags m to be either 0, 1, 2 or 3, corresponding to an $A(1)$, $A(2)$, $A(3)$ or $A(4)$ process, respectively. Normally, m is determined by minimising the optimisation parameters such as the AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion) over a range of values (Akaike 1974; Schwarz 1978). We use the BIC for our calculations because it tends to select simpler $A(p)$ models by definition, compared to AIC. However, we do

not merely proceed with the best lag suggestion m_B from the BIC test, rather, we take it as an initial guess. We also check if residuals after fitting are uncorrelated. If the residuals are uncorrelated, then m_B is an ideal suggestion for the lag order. If not, m_B is increased until the residuals show no correlation at lag m. Uncorrelated residuals are an important assumption of the ADF test, the verification of which is carried out by carrying out an ACF test once again.

In case of AR 12158, it was found from BIC testing that among $m = \{0, 1, 2, 3\}$, $m_B = 0$ minimises the BIC. Thus we took $m = 0$ as our initial guess and carried out ADF regression with the model $\Delta Y_t = \gamma Y_{t-1} + \epsilon_t$, corresponding to an $A(1)$ process. The residuals in this case were subjected to an ACF test once again and showed no more autocorrelation (see Figure 6.3d). This validates the assumption that the residuals are uncorrelated and assures that $m = 0$ (or $p = 1$) is suitable to model the data. With this information in hand, we generated plots a & b in Figure 6.3 to illustrate the mathematical modelling of L_{tot} and $g(t)$ corresponding to L_{tot}, respectively.

5. Carrying out the ADF test

From the regression estimate $\hat{\gamma}$, the ADF statistic τ_{ADF} may be computed manually. However, we carried out the test using the *adfuller* routine on Python, using a 'constant-only' keyword assumption since we have removed the functional trend $f(t)$ before. The $g(t)$ data was found to be stationary at a 1% significance level.

6.3.2 Model Fitting Results

The mathematical time series modelling approach in this chapter decomposes the PIL length into a functional time trend $f(t)$ and a stochastic component $g(t)$, which consists of a deterministic autoregressive model $A(p)$ and a purely stochastic white noise component ϵ_t. The coefficient of determination R^2 described before in the earlier section quantifies what fraction of data (L) variation is explained by $f(t)$. We numerically computed the variance fractions $V_{g(t)}$, $V_{A(p)}$ and V_{ϵ_t}, pertaining to the (i) stochastic component of the data, (ii) the purely autoregressive component of the data and (iii) the purely stochastic component of the data, respectively to understand numerically the weight of each type of function in the data. This was done by taking the variance of the time series of $g(t)$, $A(p)$ and ϵ_t after removing their respective mean values and then dividing by the variance of the mean-removed original data (L_{tot} or L_{max}). The results for the photospheric PIL data (see Tables 6.2 and 6.3) indicate that L_{tot} is a much more reliable parameter to study PILs as compared to L_{max}. This is reflected in the number of cases in which we could find a useful fit for $f(t)$. Whereas in case of L_{tot}, a good fit for $f(t)$ could be determined for all the flaring cases, in case of L_{max}, we could not find a suitable time varying deterministic trend in five out of eight cases for ARs linked to X-class flares and three out of four cases for ARs linked to M-class flares. For non-flaring ARs, the results for L_{tot} and L_{max} are comparable. If we take a look at the variance fraction of the purely stochastic uncorrelated white noise component, in general, it is less for L_{tot} when compared to L_{max} for all the flaring cases (see Tables 6.2 and 6.3). However the variance fractions for L_{max} and L_{tot} are equivalent for non-flaring ARs.

The differences in results for L_{tot} and L_{max}, i.e. the evolution of L_{max} being more stochastic than L_{tot} could be explained by the idea that L_{max} may not always be tracking the same PIL over 48 hours. It could be the case that longest PIL for a flaring AR at two different points in time may be different. In case of non-flaring ARs, the

number of PILs are usually quite low and L_{max} may be specifically tracking a certain PIL for most part of the data, yielding reliable results. For example, for AR 11143-1, when we computed the ratio between L_{max} and L_{tot}, we obtained 1 at several instances, implying that only one PIL was identified. However, since L_{tot} takes into consideration all the PILs in the data, it is more suitable to study PIL lengths and the analysis of variance fractions of the white noise component shows that it is much less stochastic compared to L_{max}. Hence for further analysis, we only proceed with results for L_{tot}.

Results of model fitting - photospheric L_{tot}									
AR	$f(t)$ fit	ACF $[g(t)]$	m_B	ADF $[g(t)]$ (m)	$A(p)$	R^2	$V_{g(t)}$ %	$V_{A(p)}$ %	V_{ϵ_t} %
11158	Linear	Correlated	0	Stationary (0)	$A(1)$	0.40	60.32 %	10.36 %	54.14 %
11166	Sinusoidal	Uncorrelated	0	Stationary (0)	- - -	0.63	36.70 %	0.00 %	36.70 %
11283	Sinusoidal	Uncorrelated	0	Stationary (0)	- - -	0.49	51.23 %	0.00 %	51.23 %
11520	Quadratic	Correlated	0	Stationary (0)	$A(1)$	0.64	35.94 %	4.86 %	31.41 %
12017	Quadratic	Correlated	0	Stationary (0)	$A(1)$	0.78	21.98 %	2.13 %	19.39 %
12158	Sinusoidal	Correlated	0	Stationary (0)	$A(1)$	0.79	21.30 %	5.11 %	15.50 %
12297	Quadratic	Correlated	0	Stationary (2)	$A(3)$	0.82	17.36 %	6.37 %	8.08 %
12673	Sinusoidal	Correlated	0	Non-stationary (0)	$A(1)$	0.40	59.66 %	25.29 %	35.41 %
11620	Exponential (c=0)	Correlated	0	Stationary (0)	$A(1)$	0.90	10.31 %	1.08 %	9.58 %
11719	Exponential (c=0)	Uncorrelated	0	Stationary (0)	- - -	0.81	18.96 %	0.00 %	18.96 %
11818	Exponential (c)	Uncorrelated	0	Stationary (0)	- - -	0.35	64.96 %	0.00 %	64.96 %
12497	Quadratic	Correlated	0	Stationary (0)	$A(1)$	0.36	64.09 %	22.74 %	48.02 %
11143-1	Exponential (c=0)	Correlated	0	Stationary (0)	$A(1)$	0.58	42.23 %	6.63 %	28.48 %
11143-2	None	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -
11143-3	Exponential (c=0)	Correlated	3	Non-stationary (3)	$A(4)$	0.70	30.48 %	9.59 %	22.13 %
11710-1	Exponential (c=0)	Correlated	2	Stationary (2)	$A(3)$	0.93	7.45 %	1.82 %	9.23 %
11710-2	None	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -
11710-3	Exponential (c)	Uncorrelated	0	Stationary (0)	- - -	0.48	51.23 %	0.00 %	51.23 %
12240-1	Exponential (c=0)	Correlated	3	Stationary (3)	$A(4)$	0.83	16.60 %	9.68 %	13.28 %
12240-2	Exponential (c)	Correlated	2	Stationary (2)	$A(3)$	0.61	39.42 %	15.51 %	30.90 %
12240-3	None	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -

Table 6.2: Details of the curve fitting for all the 12 flaring and 9 non-flaring ARs for L_{tot}; $V_{g(t)}, V_{A(p)}$ and V_{ϵ_t} quantify the percentage of variance explained by the residual after time trend fitting, autoregressive model and white noise, respectively; stationarity is reported at 1% S_l

Results of model fitting - photospheric L_{max}									
AR	$f(t)$ fit	ACF $[g(t)]$	m_B	ADF $[g(t)]$ (m)	$A(p)$	R^2	$V_{g(t)}$ %	$V_{A(p)}$ %	V_{ϵ_t} %
11158	None	Correlated	0	Non-stationary (0)	$A(1)$	- - -	100 %	25.02 %	74.98 %
11166	None	Correlated	0	Non-stationary (0)	$A(1)$	- - -	100 %	38.74 %	61.26 %
11283	None	Correlated	0	Stationary (3)	$A(4)$	- - -	100 %	22.09 %	78.91 %
11520	None	Uncorrelated	0	Stationary (0)	- - -	- - -	100 %	0.00 %	100 %
12017	Quadratic	Uncorrelated	0	Stationary (0)	- - -	0.47	52.59 %	0.00 %	52.59 %
12158	Quadratic	Uncorrelated	0	Stationary (0)	- - -	0.59	41.44 %	0.00 %	41.44 %
12297	Sinusoidal	Correlated	0	Stationary (0)	$A(1)$	0.53	46.86 %	4.26 %	42.64 %
12673	None	Uncorrelated	0	Stationary (0)	- - -	- - -	100 %	0.00 %	100 %
11620	None	Correlated	0	Stationary (0)	$A(1)$	- - -	100 %	31.78 %	68.22 %
11719	Exponential (c=0)	Correlated	1	Stationary (2)	$A(3)$	0.60	39.64 %	17.60 %	17.21 %
11818	None	Correlated	0	Stationary (0)	$A(1)$	- - -	100 %	11.96 %	88.04 %
12497	None	Correlated	0	Stationary (0)	$A(1)$	- - -	100 %	26.30 %	73.70 %
11143-1	Exponential (c=0)	Correlated	0	Stationary (0)	$A(1)$	0.60	39.72 %	5.90 %	27.16 %
11143-2	None	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -
11143-3	Exponential (c=0)	Uncorrelated	0	Stationary (0)	- - -	0.72	27.54 %	0.00 %	27.54 %
11710-1	Exponential (c=0)	Uncorrelated	- - -	Stationary (0)	- - -	0.81	18.92 %	0.00 %	18.92 %
11710-2	None	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -
11710-3	Exponential (c)	Correlated	0	Stationary (0)	$A(1)$	0.40	60.46 %	15.96 %	44.87 %
12240-1	Exponential (c=0)	Correlated	0	Stationary (3)	$A(4)$	0.58	42.26 %	5.32 %	39.60 %
12240-2	Exponential (c)	Correlated	2	Stationary (2)	$A(3)$	0.54	45.66 %	16.95 %	34.85 %
12240-3	None	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -

Table 6.3: Details of the curve fitting for all the 12 flaring and 9 non-flaring ARs for L_{max}; $V_{g(t)}, V_{A(p)}$ and V_{ϵ_t} quantify the percentage of variance explained by the residual after time trend fitting, autoregressive model and white noise, respectively; for cases where a suitable $f(t)$ is not found, L_{max} is directly treated as $g(t)$ modelled using an autoregressive model; V_{ϵ_t} is 100% for ARs 11520 and 12673 suggesting that the L_{max} data is purely random noise; stationarity is reported at 1% S_l

6.3.3 Time Trend Curve Fitting

From these photospheric case studies, we find that in case of ARs linked to X-class flares, there is a greater probability of finding seasonal patterns or a quadratic pattern (increase in L_{tot} followed by a decrease or vice-versa) as compared to ARs linked to M-class cases or non-flaring ARs. These conclusions are weak at this stage and require further validation from a statistical study. The seasonal pattern was most pronounced for AR 12158, as seen from the R^2 value of 0.79. The sinusoidal curve also fit well for ARs 11520, 12017 and 12297 hinting at the presence of a seasonal pattern in these cases. However since we did not observe one complete cycle, the presence of seasonality is not conclusive and the quadratic curve was considered as the best fit curve for these cases. For ARs linked to M-class flares and non flaring ARs, we find that the exponential curve fit of the form $f(t) = ae^{bt} + c$ may be a reasonable model to explain the PIL data. Some good examples for the exponential fit are ARs 11620 and 11719 for M-class cases and ARs 11710-1 and 12240-1 for non-flaring cases. This may possibly indicate that PIL growth or decay may be proportional to the PIL length and may follow a first order differential equation of the form (neglecting the autoregressive component):

$$\frac{dL_{tot}}{dt} = bL_{tot} \tag{6.4}$$

It may be the case that for ARs associated with X-class flares, PIL lengths may have more or less saturated or may be influenced by other physical processes which are not concerned with PIL growth. Thus these cases may not be exhibiting exponential growth as we find in case of non-flaring ARs or those whose strongest flares belong to weaker classes. Once again, this aspect requires further verification from a statistical study.

One of the objectives of the study was to check that in the absence of a defined PIL length threshold that correlates to flare occurrence, if the pattern in which PIL length evolves, provides hints regarding a potential flare. We found that the time evolution of flaring ARs can be functionally similar to non-flaring ARs, making it tough to distinguish between flaring and non-flaring ARs. As an example, we present and compare the case of ARs 11719 and AR 11710-1. While AR 11719 hosted a δ-type sunspot-group and an M class flare, AR 11710-1 hosted a β-type sunspot-group and was not associated with any flare stronger than C1.0. However, both the cases showed an exponential decay in L_{tot} at similar length scales - from about 300 Mm at the start to less than 50 Mm within 48 hr (see Figure 6.4).

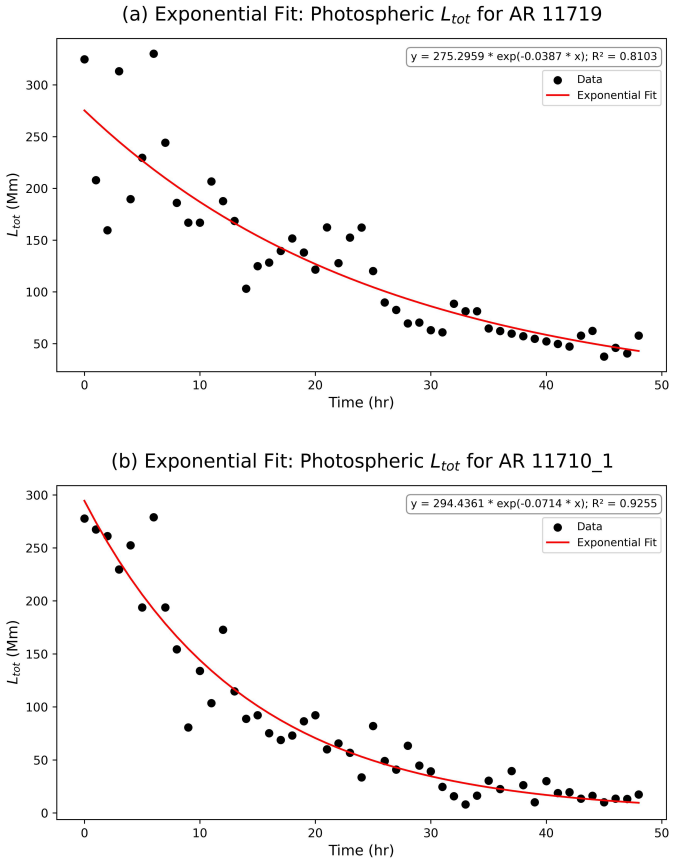

Figure 6.4: Figure showing exponential decay of L_{tot} for (a) a flaring AR 11719 and (b) a non-flaring AR 11710; the start time of the data for each of the cases is listed in Table 6.1

6.3.4 Autoregressive Model Fitting

In two instances, ARs 11166 and 11283, the residual $g(t)$ was effectively white noise after fitting a sinusoidal curve and thus autoregressive models were not necessary to model L_{tot}. However, in cases where the ACF of $g(t)$ showed that the residuals were correlated, using an autoregressive model makes the system more deterministic as autoregressive models may explain up to an additional 25% of the variance in the data after curve fitting. In some cases, such as AR 12017 where the $A(1)$ model accounts for only about 2% of the total data variance or about 10% of the variance not explained by $f(t)$, autoregressive models may not be useful to improve the data modelling significantly and may only be used as a tool to remove serial correlation from the model fit residual $g(t)$. For such cases, we may say that the impact of previous points in the data is minimal. In general for most flaring cases, we find that the $A(1)$ model, is suitable for data modelling suggesting that the dependence of the L_{tot} series on the past is short memory, and only the recent-most past value matters. Even in case of AR 12297 where we used an $A(3)$ model to explain the data, the initial guess from BIC optimisation suggested an $A(1)$ model. However, the residuals after fitting the $A(1)$ model were correlated and thus we had to manually increase the model order to 3 until it yielded uncorrelated residuals. For non-flaring ARs, the impact of memory in the data was slightly stronger and as a result we found quite a few cases where more complex autoregressive models such as $A(3)$ or $A(4)$ were needed to eliminate the serial correlation in $g(t)$ after curve fitting.

6.3.5 ADF Test Results

We carried out the ADF test at the optimal lag m for each of the ARs, and ensured that the ACF of the ADF residuals at the optimal lag did not yield any significant correlations as a means of cross checking the assumption behind the ADF test. In 16 out of 18 cases (see Table 6.2), we found a stationary output suggesting that the curve fit was reliable. Only two cases produced non-stationary output - ARs 12673 and 11143-3. AR 12673 produced the biggest flare of solar cycle 24 - an X9.3 flare at 2017/09/06, 12:00 UTC, approximately three hours after the first X2.2 flare at about 09:00 UTC on the same day. The non stationary output may be either because the curve fitting is not accurate or it may possibly indicate the presence of stochastic trends i.e. a random walk like behaviour. This may also indicate long-term persistence (see Figure 6.5a) - which in this context means that future values in this time series are highly dependent on the cumulative effect of previous shocks to the system. In case of AR 11143-3, the non-stationary output may indicate that the model fitting does not explain the data quite well. A single spike at lag 3 in $g(t)$ but not at other values (see Figure 6.5b) shows that long term persistence is not the cause of non-stationary output.

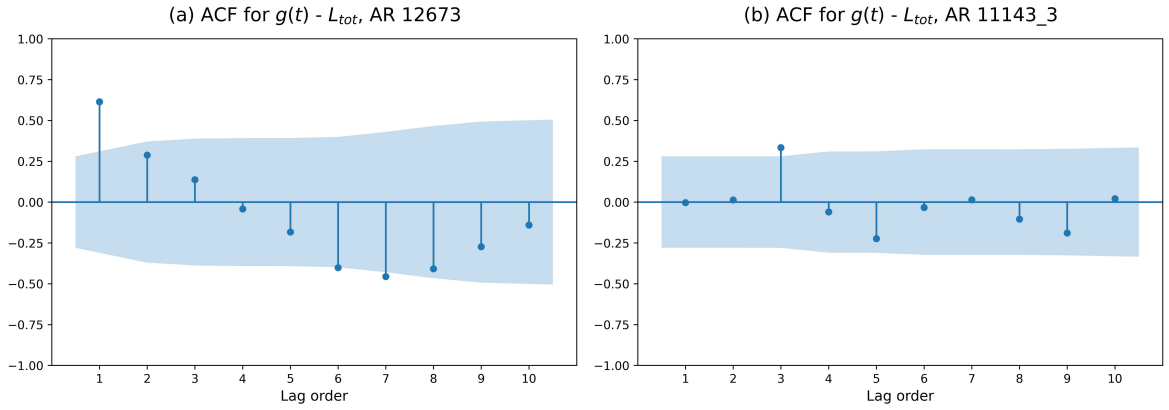

Figure 6.5: ACF plots of L_{tot} for (a) AR 12673 and (b) 11143-3; the ACF of $g(t)$ for AR 12673, shows persistence as the spikes don't cut off sharply and provides evidence of non-stationary behaviour; the ACF of $g(t)$ for AR 11143-3 however does not shows a slow decay of correlation with lag value but rather an isolated spike at lag 3

6.3.6 Direct Application of Autoregressive Models

In the previous sections, the time series modelling approach took into account the functional time trend $f(t)$. However, in this section we model the data directly using autoregressive models without finding $f(t)$. While this is not a mathematically sound technique to analyse stationarity in the data, the underlying idea (out of curiosity) is to check if there is any clear pattern in terms of the presence of unit root to distinguish flaring ARs from non-flaring ones without assuming any deterministic time trend $f(t)$ at all heights between 0 - 3.24 Mm in our data. This also implies that we are modelling even the deterministic time trends, if present, as a memory model. In this light, any result obtained from the ADF test is purely in terms of the presence of a unit root i.e. whether shocks (random changes) have permanent impact on the data (non-stationary output - unit root present) or if they fade over time (stationary output - unit root absent). The results from the ADF test may not be used to conclude constancy in mean, variance or autocorrelation as we are not accounting for the deterministic time trend $f(t)$, unless the data appears to not have a deterministic time trend in the first place.

We used two different approaches of carrying out the ADF test - (i) by employing an automatic lag selection method that optimises the BIC, (ii) by uniformly using a lag 1 calculation, corresponding to using a $A(2)$ model, justified by a 91% accuracy (tests carried out separately - see project repository) of uncorrelated residuals for the set of 120 flaring time series (12 ARs x 10). Each of these approaches has its own limitations. While sometimes the BIC estimate for the lag m_B may not necessarily ensure uncorrelated residuals, using a $A(2)$ model may sometimes over-fit or under-fit the data. For each of these cases, we generated a heat-map of the p-value (MacKinnon 2010), denoted as p_{ADF}, for all ARs and heights (see Figure 6.6). p_{ADF} is the probability of obtaining results as extreme as the observed ones, given that the null hypothesis - a unit root is present, is true.

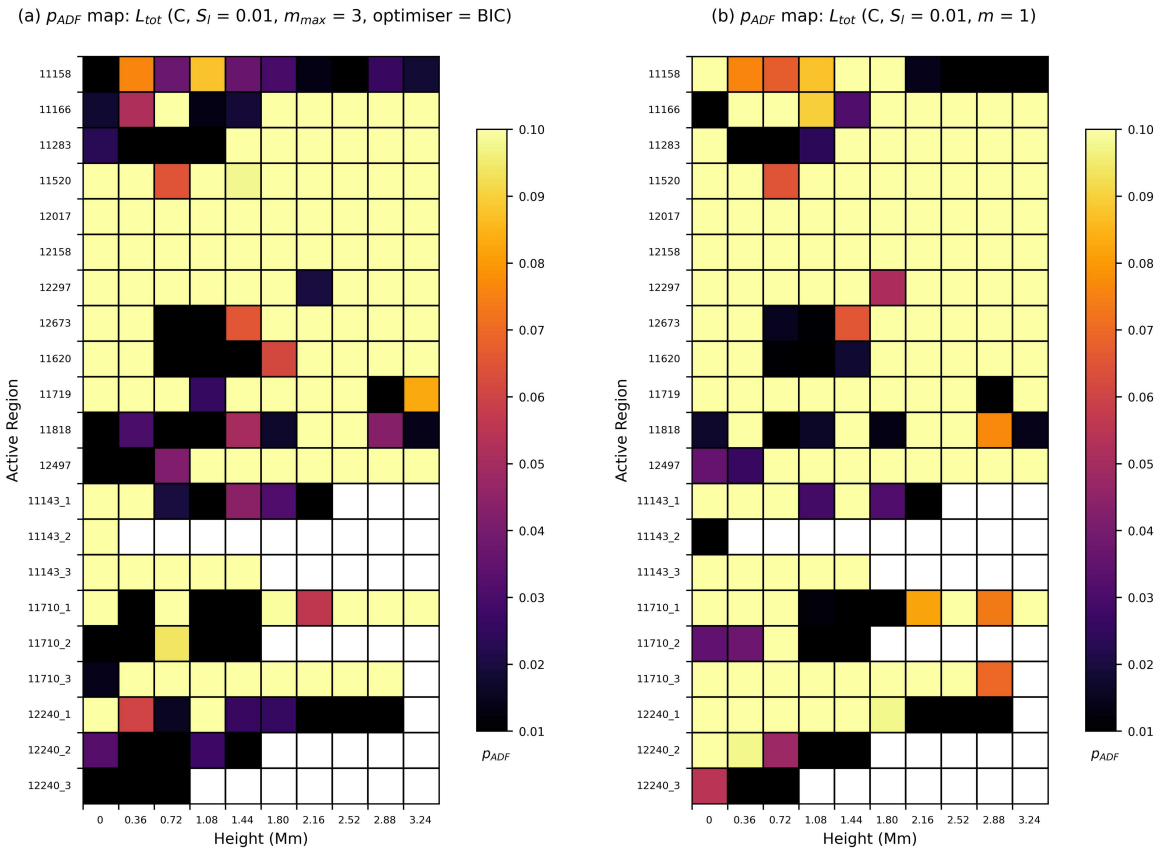

Figure 6.6: Heatmaps of the p_{ADF} - the probability that the null hypothesis of assuming non-stationarity is valid; (a) for an automated lag selection method using BIC with maximum lag set to 3 and (b) for a uniform lag set to 1; the colorbar denotes the value of p_{ADF}, the white boxes for the non-flaring ARs represent the cases where no PILs were detected in the extrapolated B_z data and thus the ADF test could not be performed

We found that there were no clear patterns to distinguish flaring ARs from non flaring ARs. For e.g. if we consider AR 11158 and AR 12158, we find that in case of AR 12158, a unit root is present at all heights for the 48 hr L_{tot} data, implying that stochastic shocks (ϵ_t) have a permanent effect, while this is not the case for AR 11158. Further, among the flaring cases, there was no clear difference in terms of the presence of a unit root, between the two classes - M/X or between the types of AR - flux decreasing (ARs 11520, 12158, 12297, 11818) or flux increasing (ARs 11158, 11166, 12017, 12673, 11620, 11818, 12497).

6.3.7 Implications on Flare Predictability and Forecasting

Even though it is known that in general, PILs are connected to solar eruptive phenomena, at present, there are no statistical studies to suggest that there is a relationship between a defined threshold on PIL length (or a related length based parameter) and the class-wise probability of flare occurrence. However, there are studies that link CME productivity to PIL length (Falconer et al. 2002; Falconer et al. 2003). In this context, assuming that no thresholds on PIL lengths exist to forecast/predict flares, we find that the time series analysis approach using autoregressive models does not satisfactorily distinguish flaring from non-flaring ARs. However, despite the negative results, one positive aspect of this study was that we learned that L_{tot} is a much less stochastic parameter to analyse compared to L_{max}.

Using a mathematical modelling approach for L_{tot} that involves deterministic trend fitting followed by residual modelling with an autoregressive model (as implemented before in this chapter) shows that the stochastic white noise component can vary between 8% - 54% for X-class flares and between 10% - 65% for M-class flares. A high white noise component in the data, especially above $\sim 50\%$, suggests that the data has more noise than information and is not ideal for forecasting using the method we have described in Section 6.3.1. If we consider a 50% white noise level cut-off, then 75% of the flaring ARs and 83% of the non flaring cases listed are suitable for forecasting (see Table 6.2). Only 1 out of 6 non-flaring cases - AR 11710-3 and 3 out of the 12 flaring cases - ARs 11158, 11283 and 11818 are not suitable for forecasting.

Assuming that in future, a relation between a threshold in L_{tot} (or a length based PIL parameter) and class-wise flare occurrence probability is found, then only for the cases which are suitable for forecasting (i.e. cases with low white noise percentage), the previous 48 hr data (or 24 hr data, the time duration T is subjective) may be thought of as a model training time and a forecast of the PIL parameter may be made

up to $T/4$ hr, so as to maintain a balance between training data and test data in a 4:1 ratio. In this framework, if the forecast exceeds the defined threshold after T_{resp} hr ($T_{resp} < T/4$), then T_{resp} can be thought of as a response time, i.e. the minimum time prior to flares which may be used for issuing a flare warning and for safeguarding assets which may be impacted by a flare. This hypothesised method of flare prediction through PIL parameter forecasting assumes that (i) PIL parameter does not breach the threshold during the model training time and (ii) once a PIL parameter breaches a certain threshold, flares are to be expected with a certain known time window, similar to what has been found for CMEs by Falconer et al. (2003). This approach is most likely to work well for flux emerging ARs where the PIL length is continuously increasing. One of the best examples in our dataset is AR 11620 which showed flux emergence and an exponential growth in L_{tot} with less than 10% white noise component for 48 hr prior to flares.

For some of the ARs, the non-stationary output of the ADF test on $g(t)$ suggests that the current mathematical model may not completely explain the evolution of L_{tot}. For such cases, differencing the data may be necessary to forecast the data. For such cases, ARIMA (Autoregressive Integrated Moving Average) models are necessary for forecasting the data (Hamilton 1994; Lutkepohl 2005). ARIMA models can also be applied directly to the data without looking for deterministic time trends. This is because differencing eliminates time trends and on certain occasions seasonality in the data. The only major limitation in using an ARIMA model is that once differencing is applied, the underlying physics which may be controlling the time trend such as type of growth/decay, seasonality etc cannot be explained.

6.3.8 PIL Length Stack-plots

In this section, we present a few stack-plots of L_{tot} (in the same format as we have seen before for R-value) between 0 - 3.24 Mm and analyse the variation of PIL length with height. We have chosen ARs 11158, 12497 and 11143 as sample cases from X-class flaring ARs, M-class flaring ARs and non-flaring ARs, respectively (see Figure 6.7). From a visual inspection of the stack-plots in Figure 6.7, it appears that looking for the presence of PILs at heights above the photosphere, could potentially work as a discriminator between flaring and non-flaring ARs.

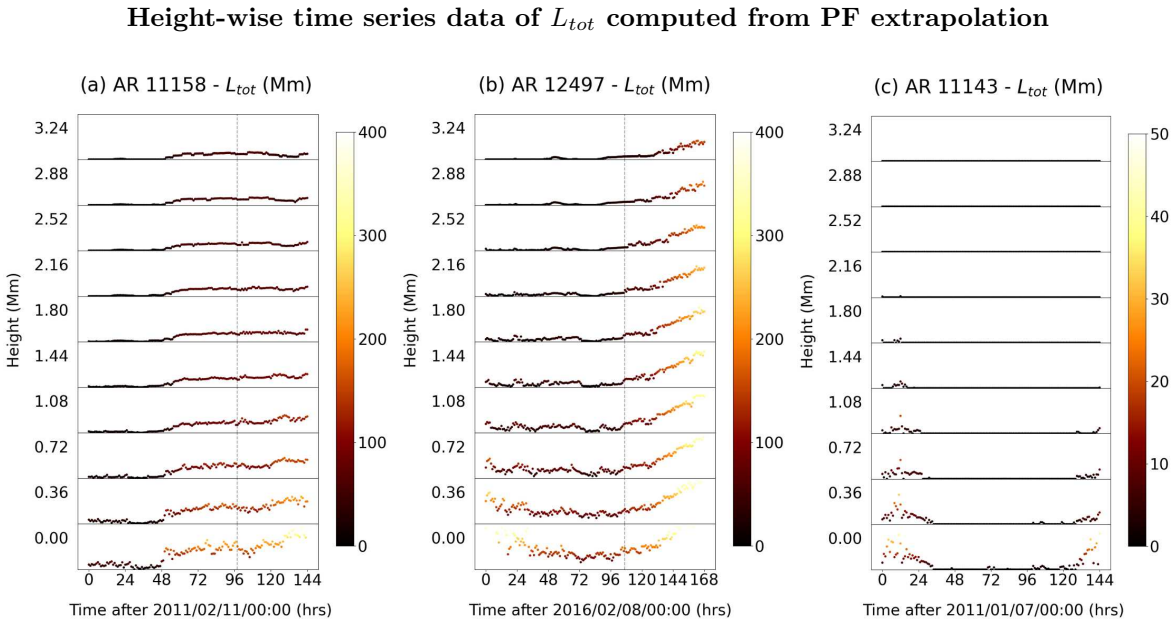

Figure 6.7: Multiple-height stack plots of L_{tot} for (0-3.24 Mm) for (a) AR 11158 (X-class flare case), (b) AR 11710 (M-class flare case) and (c) AR 11158 (non-flaring case) as a function of time; the colorbar in each case represents the magnitude of the L_{tot} (in Mm); the colorbar limit is set to 400 Mm for flaring cases - ARs 11158 and 12497 and 50 Mm for the non-flaring case - AR 11143; the dashed line in (a) and (b) represents the time of occurrence of the X-class flare; the height-wise evolution of L_{tot} is starkly different for flaring and non-flaring cases; no PILs found at 2.52 Mm or beyond for the non-flaring case - AR 11143

For the non-flaring cases in our dataset, while the highest altitude at which a single PIL was observed varied from case to case, in eight out of the nine ($\sim 89\%$) cases, no PILs were observed at a height of 3.24 Mm (see Table 6.4). Hence, based on the limited AR sample studied in this work, it appears that non-flaring ARs may not consistently show the presence of PILs over a 48 hr period at about 3 Mm above the photosphere and such behaviour is significantly different from pre-flaring ARs (prior to the first strong flares, such as M and X classes). In this connection we define a new parameter L^*_{tot} which denotes the maximum value of L_{tot} in a given time series. At $z = 3.24$ Mm, L^*_{tot} varied between 23 - 115 Mm for X-class cases and 18 - 64 Mm for M-class cases. For 9 non-flaring cases studied, only 1 non-zero L^*_{tot} value was observed. Given such promising results, we conclude that all heights up to 3.24 Mm require a statistical investigation to check if some L^*_{tot} value could be found at each height which may help distinguish between flaring and non-flaring ARs. Such an analysis may also provide information on the most optimal height within the 0 - 3.24 Mm height range, which when analysed gives the best distinction between flaring and non-flaring ARs.

L^*_{tot} - Maximum L_{tot} at 3.24 Mm height (in Mm)											
X-class flares				M-class flares		Non-flaring cases					
AR	L^*_{tot}	AR	L^*_{tot}	AR	L^*_{tot}	AR	L^*_{tot}	AR	L^*_{tot}	AR	L^*_{tot}
11158	56.57	12017	22.66	11620	63.95	11143-1	0.00	11710-1	4.87	12240-1	0.00
11166	41.81	12158	61.53	11719	18.30	11143-2	0.00	11710-2	0.00	12240-2	0.00
11283	31.55	12297	78.48	11818	40.69	11143-3	0.00	11710-3	0.00	12240-3	0.00
11520	53.43	12673	115.21	12497	26.85						

Table 6.4: A tabulation of the maximum value of L_{tot} at 3.24 Mm for different ARs

6.4 Summary

It may be argued that PIL evolution is a somewhat deterministic process if the approach is to use appropriate MHD models and boundary conditions to study the LSA. However, if we rely on observational data of the solar vector magnetic field, then PIL evolution can be considered to be a stochastic process. As evidence to support this, our sample of 12 ARs showed that the white noise component can vary between 8% - 54% for X-class flares and between 10% - 65% for M-class flares. This also implies that some cases, which have a very high white noise component in the L_{tot} time series data cannot be forecasted accurately, unless a better model is found that captures the deterministic time trend and the memory structure well. Further, L_{tot} appears to be a less stochastic parameter compared to L_{max} based on the results of curve fitting.

Using the examples of AR 11719 and AR 11710-1 (where both the cases lacked a significant stochastic component), we showed that the evolution of L_{tot} may be similar for flaring and non-flaring ARs and may not provide a clear picture to distinguish flaring from non-flaring ARs. While using autoregressive models may not be suitable to study each and every AR, it may be useful to understand the memory structure of specific AR cases (ARs 12497 and 12673 for example).

Unlike the R-value predictor, where temporal trends indicated a jump in the data for emerging ARs prior to the first M/X-class flare, there are no patterns in the evolution of L_{tot} (disregarding the magnitude aspect here) that provide hints of an imminent major flare.

From the stack-plot analysis of PILs, it appears that the decay of L_{tot} with height may serve as a discriminant between flaring and non-flaring ARs. Hence, a more consolidated study in this direction, with specific focus to L_{tot} magnitudes is required (see Chapter 7).

Chapter 7

Analysis of PIL length data

7.1 Objective

Our main objective in this chapter is to investigate if the PIL length (L_{tot}) derived from the extrapolated magnetic field could (i) help distinguish between flaring and non-flaring ARs, especially with the help of L_{tot}-based thresholds and (ii) assist in the prediction of major flares. In this connection, we explore how the magnetic field configuration differs between strongly flaring/non-flaring ARs, and how it is for intermediate cases - weakly flaring ARs (categories defined in Section 7.2). Similar to Chapters 5 & 6, the extrapolation data is obtained between 0 - 3.24 Mm with a 1 hr cadence and the stack-plots of L_{tot} are analysed, with specific focus to L_{tot} magnitudes, observed at different extrapolated heights.

7.2 Dataset

Unlike Chapters 5 & 6 which were largely case studies, the dataset discussed in this chapter may be considered as a small statistical dataset. Our dataset consists of 40 ARs belonging to solar cycle 24, and a total of 92 different 48-hr L_{tot} time series data are analysed. Among these 40 ARs, 24 are flaring ARs (indicating that the strongest flare strength $\geq$ C1.0) and the remaining 16 are non-flaring ARs *[1] (non-flaring active regions), indicating that the strongest flare $<$ C1.0. Further, among these 24 flaring ARs, we have 8 X-class flaring ARs, 8 M-class flaring ARs and 8 C-class flaring ARs, indicating that the strongest flare observed was a GOES X-class, M-class and a C-class flare, respectively (see Table 7.1). For simplicity (see further discussion in Section 7.3) we call the 16 ARs (with X/M-class flares) as SFARs* (strongly flaring active regions) and the 8 CFARs* (C-class flaring active regions) may be considered as weakly flaring cases i.e. cases intermediate to strongly flaring and non-flaring cases.

For all the flaring ARs, the L_{tot} data in the 48-hr period immediately before the first flare of focus is analysed. For example, AR 11620 is an M-class flaring AR as the strongest flare was an M2.2 flare, which occurred on Nov 28, 2012 (21:20 UTC). Hence our observation interval is between 26-28 Nov, 2012 (21:00 UTC). However, non-flaring ARs do not have any flares stronger than C1.0 by definition. Thus, we sample the non-flaring AR data into distinct and independent time-series data. For example, the L_{tot} data for AR 11143 is obtained for a 144-hr duration between Jan 7-13, 2011 (00:00 UTC). Thus it is independently sampled into three 48-hr distinct intervals - Jan 7-9, Jan 9-11 and Jan 11-13 (00:00 UTC), denoted by AR 11143-1, AR 11143-2 and AR 11143-3, respectively. In total we have 45 different time series in our dataset for non-flaring ARs (see Table 7.2).

[1] Use of acronyms - SFAR, NFAR, CFAR and ISCFAR is limited to figures and tables, and has been avoided in the text to improve readability.

We note here that the definition of non-flaring active regions (as defined in the previous paragraphs) is a subjective one. If our goal would be to solely focus on the occurrence of X/M-class flares, whether C-class flares occur or not would be immaterial. In this context, one may also consider weakly flaring cases such as C-class flaring ARs as non-flaring. Hence, to explore this consideration further, we have also sampled the C-class flaring ARs into distinct and independent 48-hr L_{tot} time-series data. This gives us the data for ISCFARs*[2] (independently sampled C-class flaring active regions). The benefit of using such an approach is that we have an idea of whether the behaviour of C-class flaring ARs is more strongly inclined towards strongly flaring ARs or non-flaring ARs. For example, AR 12353 is a C-class flaring AR with the first C-class flare (C1.0) occurring on May 23, 2015 (03:27 UTC). Thus we analyse the pre-flare data between May 21-23, 2015 (03:00 UTC). The magnetogram data was available for this AR between May 21 - 27 within 60° EW of the solar central meridian and thus we split this 144-hr duration into three independent and distinct 48-hr intervals, similar to that of non-flaring ARs. In total, we have 23 different time series for independently sampled C-class flaring ARs in our dataset (see Table 7.3).

Our analysis of the different flaring categories also takes into account the sunspot-group information. In this light, we define three terms for an AR/time-series: (i) ISC (Interval Sunspot Complexity), referring to the evolution of sunspot-group complexity of the AR in a given interval, (ii) MISC (Maximum Interval Sunspot Complexity), referring to the sunspot-group type of the most complex sunspot configuration in a given interval and (iii) MSC (Maximum Sunspot Complexity), referring to maximum sunspot complexity throughout the evolution of the AR, not restricted to the observation interval in our study (see Tables 7.1, 7.2 & 7.3).

[2] Use of acronyms - SFAR, NFAR, CFAR and ISCFAR is limited to figures and tables, and has been avoided in the text to improve readability.

For example, between Sep 4-6, 2011 (22:00 UTC), the sunspot-group corresponding to AR 11283 increased in complexity from β-type to $\beta\gamma$-type. In this interval, the ISC is specified as $\beta/\beta\gamma$ highlighting the evolution pattern and the MISC is $\beta\gamma$. However, the most complex sunspot-group morphology for AR 11283 was a $\beta\gamma\delta$-type morphology between Sep 7-9, 2011 and thus the MSC for AR 11283 is $\beta\gamma\delta$. It is important to note here that for the terminology we defined, ISC and MISC are interval specific definitions, while MSC is an AR specific definition.

While selecting strongly flaring ARs, we selected those cases that had a δ-type sunspot as the MSC, similar to what we have done earlier in Chapters 5 and 6. While selecting C-class flaring ARs, we prioritised cases with a δ-type sunspot as the MSC (though such cases are rare), as we wanted to see if strongly flaring ARs may be distinguished from their C-class flaring counterparts even when the samples in our dataset are the most complex ones from the C-class flaring AR population. While selecting non-flaring ARs, we did not find any case that had a δ-type sunspot as the MSC. Thus, we selected only those cases which had a β-type or a γ-type sunspot-group as the MSC.

Dataset for flaring ARs (X/M/C-class)								
No.	AR	SHARP	T_{start}	$T_{flare\ onset}$	Class	ISC	MISC	MSC
01	11158	377	2011/02/13 01:00	2011/02/15 01:44	X2.2	$\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
02	11166	401	2011/03/07 23:00	2011/03/09 23:13	X1.5	$\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
03	11283	833	2011/09/04 22:00	2011/09/06 22:12	X2.1	$\beta/\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
04	11520	1834	2012/07/10 15:00	2012/07/12 15:37	X1.4	$\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
05	12017	3894	2014/03/27 17:00	2014/03/29 17:35	X1.0	$\beta/\beta\delta$	$\beta\delta$	$\beta\gamma\delta$
06	12158	4536	2014/09/08 17:00	2014/09/10 17:21	X1.6	$\beta\delta/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
07	12297	5298	2015/03/09 16:00	2015/03/11 16:11	X2.1	$\beta\gamma\delta/\beta\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
08	12673	7115	2017/09/04 09:00	2017/09/06 08:57	X2.2	$\beta\gamma/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
09	11620	2227	2012/11/26 21:00	2012/11/28 21:20	M2.2	$\beta\gamma/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
10	11719	2635	2013/04/09 06:00	2013/04/11 06:55	M6.5	$\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
11	11731	2693	2013/04/30 04:00	2013/05/02 04:58	M1.1	$\beta\gamma/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
12	11818	3056	2013/08/15 18:00	2013/08/17 18:16	M3.3	$\beta/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
13	12253	5011	2015/01/01 09:00	2015/01/03 09:40	M1.1	$\beta\gamma/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
14	12403	5885	2015/08/19 01:00	2015/08/21 01:56	M1.2	$\beta/\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
15	12422	5983	2015/09/25 20:00	2015/09/27 20:54	M1.2	$\beta\gamma/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
16	12497	6327	2016/02/10 10:00	2016/02/12 10:36	M1.0	$\beta\gamma/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
17	11440	1484	2012/03/19 12:00	2012/03/21 12:38	C2.9	$\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
18	11553	1959	2012/08/27 23:00	2012/08/29 23:40	C1.1	β	β	β
19	11640	2337	2012/12/30 08:00	2013/01/01 08:47	C1.2	β	β	$\beta\gamma\delta$
20	11730	2691	2013/04/27 20:00	2013/04/29 20:22	C5.9	β	β	$\beta\gamma\delta$
21	12255	5022	2015/01/09 14:00	2015/01/11 14:42	C1.6	$\alpha/\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
22	12293	5249	2015/03/02 13:00	2015/03/04 13:35	C1.1	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\delta$
23	12353	5596	2015/05/21 03:00	2015/05/23 03:27	C1.0	$\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
24	12552	6599	2016/06/07 12:00	2016/06/09 12:28	C1.4	β	β	$\beta\delta$

Table 7.1: Dataset for flaring AR (X/M/C-class) with the sunspot group information in columns 7-9; columns 2-3 give the NOAA AR number and the SHARP number; columns 4-5 give the observation window details and column 6 contains the strength of the flare in focus

Dataset for NFARs (non-flaring ARs)							
No.	AR	SHARP	T_{start}	T_{end}	ISC	MISC	MSC
1	11143-1	335	2011/01/07 00:00	2011/01/09 00:00	β	β	β
2	11143-2	335	2011/01/09 00:00	2011/01/11 00:00	β	β	β
3	11143-3	335	2011/01/11 00:00	2011/01/13 00:00	β	β	β
4	11416-1	1389	2012/02/09 00:00	2012/02/11 00:00	β	β	$\beta\gamma$
5	11416-2	1389	2012/02/11 00:00	2012/02/13 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma$
6	11416-3	1389	2012/02/13 00:00	2012/02/15 00:00	β	β	$\beta\gamma$
7	11673-1	2489	2013/02/17 00:00	2013/02/19 00:00	β	β	$\beta\gamma$
8	11673-2	2489	2013/02/19 00:00	2013/02/21 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma$
9	11673-3	2489	2013/02/21 00:00	2013/02/23 00:00	β	β	$\beta\gamma$
10	11682-1	2501	2013/02/25 00:00	2013/02/27 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma$
11	11682-2	2501	2013/02/27 00:00	2013/03/01 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma$
12	11682-3	2501	2013/03/01 00:00	2013/03/03 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma$
13	11710-1	2599	2013/03/30 00:00	2013/04/01 00:00	β	β	β
14	11710-2	2599	2013/04/01 00:00	2013/04/03 00:00	β	β	β
15	11710-3	2599	2013/04/03 00:00	2013/04/05 00:00	β	β	β
16	11813-1	3049	2013/08/08 00:00	2013/08/10 00:00	$\beta/\beta\gamma$	$\beta\gamma$	$\beta\gamma$
17	11813-2	3049	2013/08/10 00:00	2013/08/12 00:00	β	β	$\beta\gamma$
18	11886-1	3326	2013/10/29 00:00	2013/10/31 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma$
19	11886-2	3326	2013/10/31 00:00	2013/11/02 00:00	β	β	$\beta\gamma$
20	11922-1	3490	2013/12/10 00:00	2013/12/12 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma$
21	11922-2	3490	2013/12/12 00:00	2013/12/14 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma$
22	11957-1	3648	2014/01/18 00:00	2014/01/20 00:00	β	β	$\beta\gamma$
23	11957-2	3648	2014/01/20 00:00	2014/01/22 00:00	β	β	$\beta\gamma$
24	11957-3	3648	2014/01/22 00:00	2014/01/24 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma$
25	11957-4	3648	2014/01/24 00:00	2014/01/26 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma$

Table continues on the following page ...

Dataset for NFARs (non-flaring ARs)							
No.	AR	SHARP	T_{start}	T_{end}	ISC	MISC	MSC
26	12064-1	4135	2014/05/15 00:00	2014/05/17 00:00	β	β	β
27	12064-2	4135	2014/05/17 00:00	2014/05/19 00:00	β	β	β
28	12064-3	4135	2014/05/19 00:00	2014/05/21 00:00	β	β	β
29	12240-1	4943	2014/12/15 00:00	2014/12/17 00:00	α	α	β
30	12240-2	4943	2014/12/17 00:00	2014/12/19 00:00	β	β	β
31	12240-3	4943	2014/12/19 00:00	2014/12/21 00:00	β	β	β
32	12400-1	5865	2015/08/11 00:00	2015/08/13 00:00	β	β	β
33	12400-2	5865	2015/08/13 00:00	2015/08/15 00:00	β	β	β
34	12400-3	5865	2015/08/15 00:00	2015/08/17 00:00	β	β	β
35	12528-1	6464	2016/04/04 00:00	2016/04/06 00:00	β	β	$\beta\gamma$
36	12528-2	6464	2016/04/06 00:00	2016/04/08 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma$
37	12663-1	7050	2017/06/16 00:00	2017/06/18 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma$
38	12663-2	7050	2017/06/18 00:00	2017/06/20 00:00	β	β	$\beta\gamma$
39	12664-1	7058	2017/06/22 00:00	2017/06/24 00:00	α	α	$\beta\gamma$
40	12664-2	7058	2017/06/24 00:00	2017/06/26 00:00	α/β	β	$\beta\gamma$
41	12664-3	7058	2017/06/26 00:00	2017/06/28 00:00	$\beta/\beta\gamma$	$\beta\gamma$	$\beta\gamma$
42	12664-4	7058	2017/06/28 00:00	2017/06/30 00:00	β	β	$\beta\gamma$
43	12719-1	7299	2018/08/20 00:00	2018/08/22 00:00	β	β	$\beta\gamma$
44	12719-2	7299	2018/08/22 00:00	2018/08/24 00:00	β/α	β	$\beta\gamma$
45	12719-3	7299	2018/08/24 00:00	2018/08/26 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma$

Table 7.2: Dataset for non-flaring ARs in the same format as Table 7.1, except for the flare class column; the NFAR time-series data was independently sampled into 45 distinct 48-hr intervals

Dataset for ISCFARs (independently sampled C-class flaring active regions)							
No.	AR	SHARP	T_{start}	T_{end}	ISC	MISC	MSC
1	11440-1	1484	2012/03/19 00:00	2012/03/21 00:00	β	β	$\beta\gamma\delta$
2	11440-2	1484	2012/03/21 00:00	2012/03/23 00:00	$\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
3	11440-3	1484	2012/03/23 00:00	2012/03/25 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma\delta$
4	11553-1	1959	2012/08/24 00:00	2012/08/26 00:00	α	α	β
5	11553-2	1959	2012/08/26 00:00	2012/02/28 00:00	β	β	β
6	11553-3	1959	2012/08/28 00:00	2012/08/30 00:00	β	β	β
7	11553-4	1959	2012/08/30 00:00	2012/09/01 00:00	β	β	β
8	11640-1	2337	2012/12/30 00:00	2013/01/01 00:00	β	β	$\beta\gamma\delta$
9	11640-2	2337	2013/01/01 00:00	2013/01/03 00:00	β	β	$\beta\gamma\delta$
10	11640-3	2337	2013/01/03 00:00	2013/01/05 00:00	$\beta/\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
11	11730-1	2691	2013/04/26 00:00	2013/04/28 00:00	β	β	$\beta\gamma\delta$
12	11730-2	2691	2013/04/28 00:00	2013/04/30 00:00	β	β	$\beta\gamma\delta$
13	11730-3	2691	2013/04/30 00:00	2013/05/02 00:00	$\beta\gamma\delta$	$\beta\gamma\delta$	$\beta\gamma\delta$
14	12255-1	5022	2015/01/06 00:00	2015/01/08 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma\delta$
15	12255-2	5022	2015/01/08 00:00	2015/01/10 00:00	β/α	β	$\beta\gamma\delta$
16	12255-3	5022	2015/01/10 00:00	2015/01/12 00:00	$\alpha/\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
17	12293-1	5249	2015/02/27 00:00	2015/03/01 00:00	$\alpha/\beta\delta$	$\beta\delta$	$\beta\delta$
18	12293-2	5249	2015/03/01 00:00	2015/03/03 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\delta$
19	12293-3	5249	2015/03/03 00:00	2015/03/05 00:00	β	β	$\beta\delta$
20	12353-1	5596	2015/05/21 00:00	2015/05/23 00:00	$\beta\gamma$	$\beta\gamma$	$\beta\gamma\delta$
21	12353-2	5596	2015/05/23 00:00	2015/05/25 00:00	$\beta\gamma/\beta$	$\beta\gamma$	$\beta\gamma\delta$
22	12353-3	5596	2015/05/25 00:00	2015/05/27 00:00	β	β	$\beta\gamma\delta$
23	12552-1	6599	2016/06/08 00:00	2016/08/10 00:00	β	β	$\beta\delta$

Table 7.3: Dataset for independently sampled C-class flaring active regions in the same format as 7.1, except for the flare class column; the CFAR time-series data was independently sampled into 23 distinct 48-hr intervals

7.3 Results and Discussion

The results and discussion section is divided into four major sections:

1. The first Section (7.3.1) presents a visualisation of how PILs look when they are detected at extrapolated altitudes. This also gives a basic picture of how changes to the PIL configuration with time appear at different altitudes at/above the photosphere.

2. Section 7.3.2 is heavily focused on the descriptive statistics of L^*_{tot} (defined previously in Chapter 6) and solely focused on the current data sample we have. The statistical descriptors of L^*_{tot} for different categories of ARs are compared and we check whether L^*_{tot} may be able to distinguish strongly flaring from non-flaring ARs. Here we try to check whether an optimal height (or an OHR) can be ascertained, optimal in the sense of bringing about a distinction between strongly flaring and non-flaring categories.

3. Section 7.3.3 is inference-based and tries to infer details about the overall population of the different AR categories discussed in the chapter from the limited data sample we have. In this context, we also check if an optimal height (or an OHR) may be ascertained which may separate out population clusters of different ARs.

4. In the final Section 7.3.4, we discuss the implications of our results and explain how a flare prediction model may be envisioned from this study taking practical considerations into account.

7.3.1 Visualising PIL Detection

In this section, we present a few examples of the height-wise detection of PILs for (i) a couple of X-class flaring ARs including AR 11158 - a flux emerging case and AR 12158, which is not a flux emerging case and (ii) AR 11143, an example of a non-flaring AR (see Figure 7.1). The examples of PIL detection are provided at three different heights: $z = 0$ Mm (the photosphere), $z = 1.80$ Mm and $z = 3.24$ Mm. For the flaring cases while Figures 7.1 a-c and Figures 7.1 g-i correspond to a timestamp 2 days prior to the occurrence of the first X-class flare, Figures 7.1 d-f and Figures 7.1 j-l correspond to the time immediately before the time of occurrence of the first X-class flare. Figures 7.1 m-o show the PIL detection for AR 11143 at a timestamp when L_{tot} had its maximum value on the photosphere between Jan 7-13, 2011.

It is seen that the two X-class flaring ARs which differ in the evolution of ϕ also differ in terms of PIL detection. In case of AR 11158, the regions of opposite polarities moved towards each other between Feb 13-15, 2011 (01:00 UTC) in addition to flux emergence which took place in the background. Both these factors may have contributed to the strengthening of B_h and B_{zg} across PILs leading to PIL growth over time (see Figures 7.1a-f). The results at the extrapolated altitudes validate our observations of the photosphere.

The opposite happens in the case of AR 12158. The opposite polarity regions move away from each other, producing a clear visible gap. As a result, the interfaces where the regions of opposite polarities meet become restricted and that causes a reduction in PIL length with time (see Figures 7.1 g-l). Once again, the results at the extrapolated altitudes validate our observations of the photosphere.

As we move up in height the magnetogram data becomes smoother compared to the photosphere due to PF extrapolation. While smoothing may make the data less

noisy and make it easier to trace the motion of opposite polarity regions as observed at a height of $z = 1.80$ Mm, excessive smoothing may make fields too weak compared to the thresholds for PILs to be detected and may suppress the detection of smaller isolated fragments of PILs, as observed at a height of $z = 3.24$ Mm.

For the non-flaring AR, even for the most compact configuration within the 144-hr interval between Jan 7-13, 2011, a clear gap exists between the major opposite polarity regions. Only a small amount of positive flux region borders the negative flux region and vice versa. Upon PF smoothing these small scale local features get dominated by the larger opposite polarity regions in their immediate neighbourhood and thus we, find only a ~ 1 Mm long PIL at $z = 1.80$ Mm and no PILs at $z = 3.24$ Mm, above the photosphere.

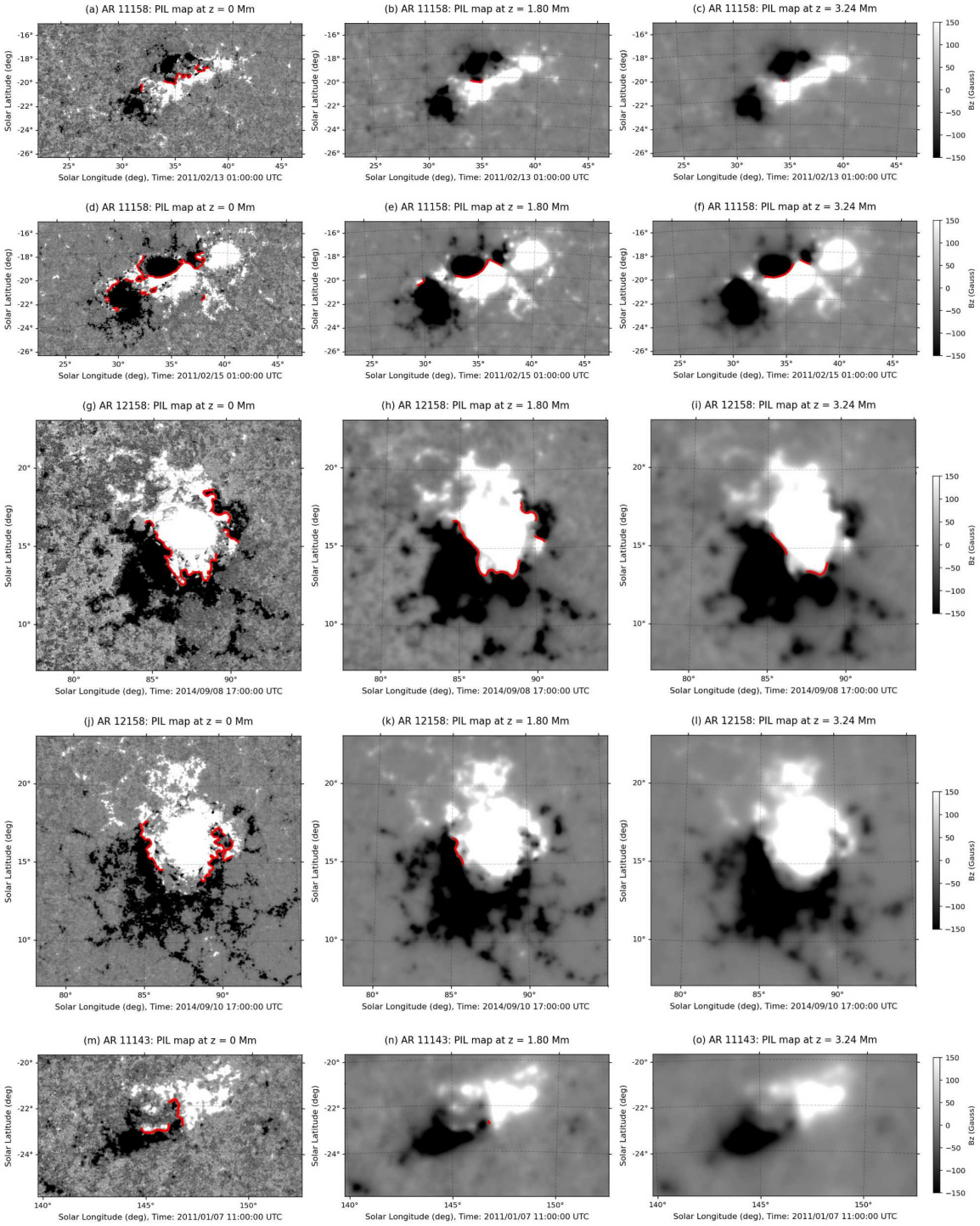

Figure 7.1: Plots showing detected PILs overlaid in red on the extrapolated magnetic field for XFARs - 11158 (a-f), 12158 (g-l) and NFAR 11143 (m-o) at a height of 0 Mm (col 1), 1.80 Mm (col 2) and 3.24 Mm (col 3); the colorbar thresholds are saturated at $\pm$ 150 G

7.3.2 Descriptive Statistics

1. L^*_{tot} decrease profiles with height

In Chapter 6, we had defined L^*_{tot} as the maximum value of L_{tot} in a given time series data. In this chapter, we analyse this parameter extensively with height. We computed the L^*_{tot} value with height for all the 92 different time series in our dataset. Then we computed the mean μ and standard deviation σ of L^*_{tot} for different flaring categories as shown in Table 7.4. Statistical details of 10 different categories are presented in Table 7.4, out of which 4 types consider sunspot information. These 4 types are actually sub-classifications of independently sampled C-class flaring ARs and non-flaring ARs into two types each. While the C-class flaring ARs (independently sampled) are classified into two types - (i) having an MISC of either an α or β-type and (ii) having an MISC of either a γ or δ-type, non-flaring ARs are classified into two types - (i) having an MISC of α or β-type and (ii) having an MISC of γ-type. The 10 different categories (see first column in Table 7.4) which are analysed further (see comparisons in Figure 7.2) are enumerated below:

1. $\mu(X)$ & $\sigma(X)$ - X-class flaring ARs
2. $\mu(M)$ & $\sigma(M)$ - M-class flaring ARs
3. $\mu(X, M)$ & $\sigma(X, M)$ - Strongly flaring ARs
4. $\mu(C)$ & $\sigma(C)$ - C-class flaring ARs
5. $\mu(ISC)$ and $\sigma(ISC)$ - C-class flaring ARs (independently sampled)
6. $\mu(ISC_{\alpha,\beta})$ & $\sigma(ISC_{\alpha,\beta})$ - C-class flaring ARs (independently sampled) with MISC α or β
7. $\mu(ISC_{\gamma,\delta})$ & $\sigma(ISC_{\gamma,\delta})$ - C-class flaring ARs (independently sampled) with MISC γ or δ

8. $\mu(NF)$ & $\sigma(NF)$ - Non-flaring ARs

9. $\mu(NF_{\alpha,\beta})$ & $\sigma(NF_{\alpha,\beta})$ - Non-flaring ARs with MISC α or β

10. $\mu(NF_{\gamma})$ & $\sigma(NF_{\gamma})$ - Non-flaring ARs with MISC γ

Statistical details (μ, σ) for different categories of data										
AR No.	h0	h1	h2	h3	h4	h5	h6	h7	h8	h9
$\mu(X)$	341.99	242.92	191.87	153.60	128.35	114.16	94.52	78.91	68.27	57.90
$\sigma(X)$	168.38	98.17	82.13	69.50	60.03	53.51	48.31	41.80	36.13	29.12
$\mu(M)$	443.77	258.32	181.28	142.76	109.68	96.50	74.86	60.13	52.29	47.04
$\sigma(M)$	257.74	112.06	57.26	45.73	45.56	44.29	42.30	36.07	32.48	31.03
$\mu(X,M)$	392.88	250.62	186.57	148.18	119.02	105.33	84.69	69.52	60.28	52.47
$\sigma(X,M)$	216.78	102.08	68.61	57.11	52.37	48.32	45.02	38.94	34.19	29.61
$\mu(C)$	129.24	82.13	60.31	42.15	34.93	30.53	25.76	23.23	18.59	16.11
$\sigma(C)$	107.53	30.78	19.66	18.72	19.42	18.59	18.05	18.12	18.74	17.40
$\mu(ISC)$	237.03	114.72	73.62	55.22	42.24	36.20	26.59	21.67	17.04	15.03
$\sigma(ISC)$	226.98	90.50	63.19	47.65	39.49	36.20	27.82	24.33	21.96	20.27
$\mu(ISC_{\alpha,\beta})$	154.64	85.78	61.9	45.56	36.46	30.98	25.34	20.65	15.08	13.24
$\sigma(ISC_{\alpha,\beta})$	132.65	53.86	39.35	30.08	28.89	25.85	24.93	23.40	20.38	20.23
$\mu(ISC_{\gamma,\delta})$	300.61	139.75	88.49	67.36	49.22	42.80	27.56	22.12	18.60	16.33
$\sigma(ISC_{\gamma,\delta})$	277.58	116.7	86.20	64.44	51.88	47.96	33.54	27.73	25.61	22.14
$\mu(NF)$	128.38	63.09	41.74	30.58	23.68	18.15	13.13	9.85	7.64	4.6
$\sigma(NF)$	135.59	57.78	35.04	26.55	22.42	18.82	14.6	11.6	10.36	6.51
$\mu(NF_{\alpha,\beta})$	123.82	58.54	37.69	27.51	21.45	16.23	11.08	8.20	6.51	3.92
$\sigma(NF_{\alpha,\beta})$	138.56	59.49	34.73	24.69	20.64	14.24	10.43	8.51	7.54	5.43
$\mu(NF_{\gamma})$	137.49	72.20	49.83	36.72	28.13	22.00	17.24	13.16	9.89	5.95
$\sigma(NF_{\gamma})$	133.73	55.04	35.42	29.87	25.81	25.86	20.43	15.99	14.53	8.32

Table 7.4: A tabulation of statistical details (μ, σ) for different categories of data

We find that the X-class and M-class flaring ARs are quite hard to distinguish (see Figure 7.2g) as there is a significant overlap between the $\mu \pm \sigma$ boundaries of L^*_{tot} for both these categories. From this observation, it may be deduced that the longest PIL configuration in the 48-hr duration immediately before the first flare for an AR may be hitting a saturation point as the GOES flare class reaches M-class levels for the first flare. Thus, we consider these two categories under a single set of strongly flaring ARs.

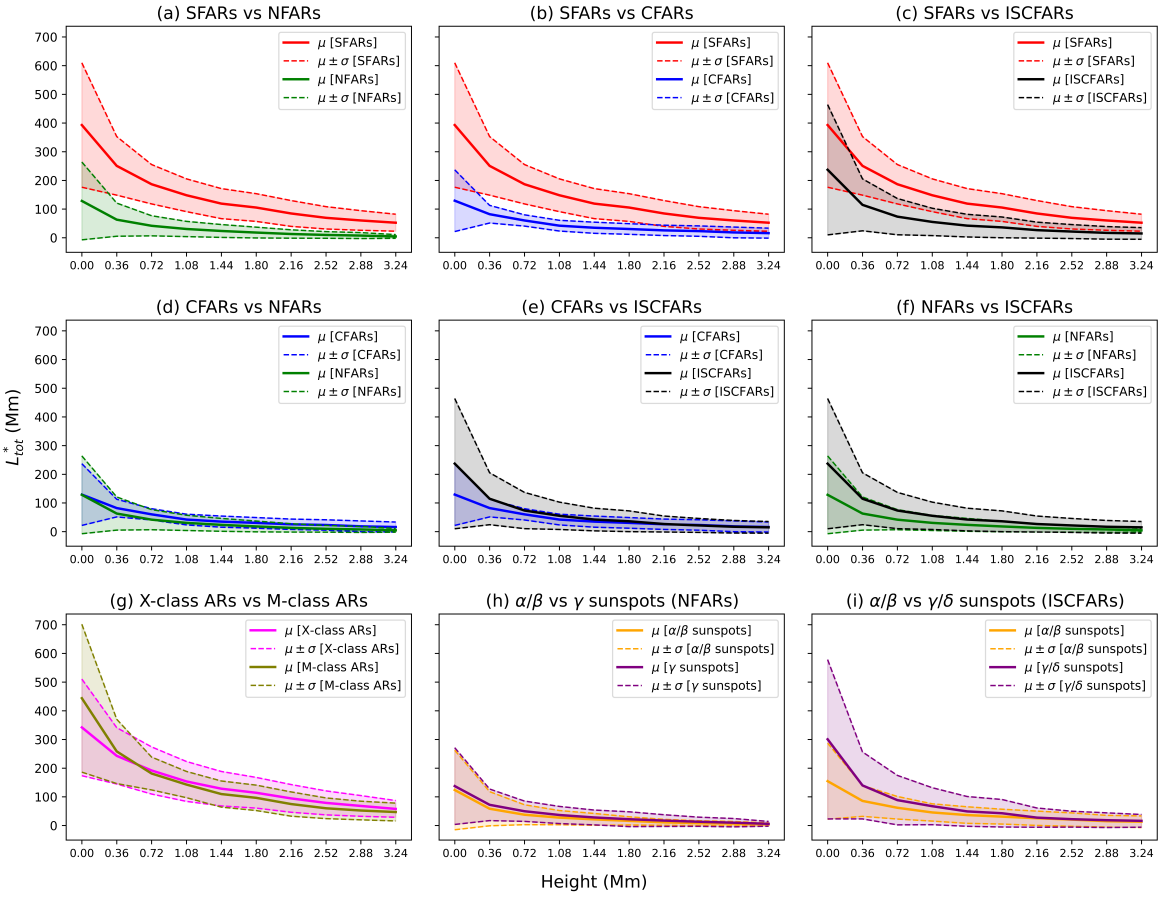

Figure 7.2: A comparison of L^*_{tot} decrease profiles (as a function of height) for different categories

Comparing strongly flaring and non-flaring ARs (see Figure 7.2a), we find a clear gap between the $\mu + \sigma$ level for non-flaring ARs and the $\mu - \sigma$ level for strongly-flaring ARs, suggesting that heights ≥ 0.36 Mm may serve as an OHR to distinguish between these two categories.

Comparing strongly flaring ARs with C-class flaring ARs, we find that the 48-hr duration prior to the first X/M-class flare is quite distinct from that of the first C-class flare and thus it suggests that L^*_{tot} does not saturate at C-class levels. Once again, there exists a clear gap between strongly flaring ARs and C-class flaring ARs at a height range of 0.36 Mm - 1.80 Mm, where $\mu + \sigma$ level of C-class flaring ARs is less than the $\mu - \sigma$ level of strongly flaring ARs (see Figure 7.2b). This height range may serve as an OHR to distinguish strongly flaring ARs from C-class flaring ARs.

In both the comparisons: (i) strongly flaring ARs vs non-flaring ARs and (ii) strongly flaring ARs vs C-class flaring ARs, there still is partial overlap between the analysed categories on the photosphere. This highlights the importance of studying PILs in the extrapolated configuration of ARs - as higher altitudes may separate these categories more cleanly compared to the photosphere.

However, when we compare strongly flaring ARs with independently sampled C-class flaring ARs, even though we find that μ of the former category is greater than μ of the latter category (see Figure 7.2c), there is still an overlap between the $\mu \pm \sigma$ zones for both these categories. In principle, ARs that produce C-class flares, considered within a specific constrained pre-flare time window (CFARs in our dataset) form a subset of independently sampled ARs that produce C-class flares (ISCFARs), because the independently sampled set is not limited to the 48-hour interval preceding the first C-class flare of an AR. Instead, it may also include 48-hour intervals following the occurrence of the first C-class flare. As a result, this dataset can capture periods during which the PIL configuration becomes more complex, which may in turn lead to an increase in the total length scale parameter L_{tot}. This is why we find that the $\mu \pm \sigma$ range of L^*_{tot} for C-class flaring ARs is sort of enveloped by the $\mu \pm \sigma$ range of L^*_{tot} of their independently sampled counterparts (see Figure 7.2e).

The L^*_{tot} $\mu \pm \sigma$ profiles for C-class and non-flaring ARs are almost similar (see Figure 7.2d), suggesting that the 48-hr interval before the first C-class flare for an AR effectively behaves like a non-flaring AR. As a result, similar to C-class flaring ARs, $\mu \pm \sigma$ zone for non-flaring ARs is enveloped by the $\mu \pm \sigma$ zone for independently sampled C-class flaring ARs (see Figure 7.2f).

Within the 45 non-flaring AR time series, in 15 instances the MISC is γ-type. In the remaining 30 cases, the MISC is α or β-type. The $\mu \pm \sigma$ profiles of L^*_{tot} for both these categories are identical (see Figure 7.2h), suggesting that the sunspot-group complexity may not be quite useful to distinguish these categories. Within the 23 independently sampled C-class flaring AR time series, in 10 instances the MISC is either γ or δ-type, while for the remaining 13 cases, the MISC is either α or β-type. Here both the categories of sunspot-group types are not distinguishable from $\mu \pm \sigma$ zones for L^*_{tot}. However, we find that the α/β-type category is sort of enveloped by the γ/δ-type category, which has a higher data spread (see Figure 7.2i).

2. Optimal height for distinguishing SFARs from NFARs or ISCFARs

The decrease profiles of L^*_{tot} as a function of height showed that it may be possible to distinguish strongly flaring ARs (i) from C-class flaring ARs within a height range of 0.36 - 1.80 Mm and (ii) from non-flaring ARs at heights $\geq$ 0.36 Mm. While these OHR estimates are made from statistical descriptors μ and σ, these descriptors do not tell us how the data is distributed. Hence we made a comparative box-plot analysis of (i) strongly flaring, (ii) C-class flaring (independently sampled), and (iii) non-flaring ARs to get insights into the distribution of the data (see Figure 7.3). We did not include CFARs in the box-plot analysis due to the lack of a sufficient number of samples.

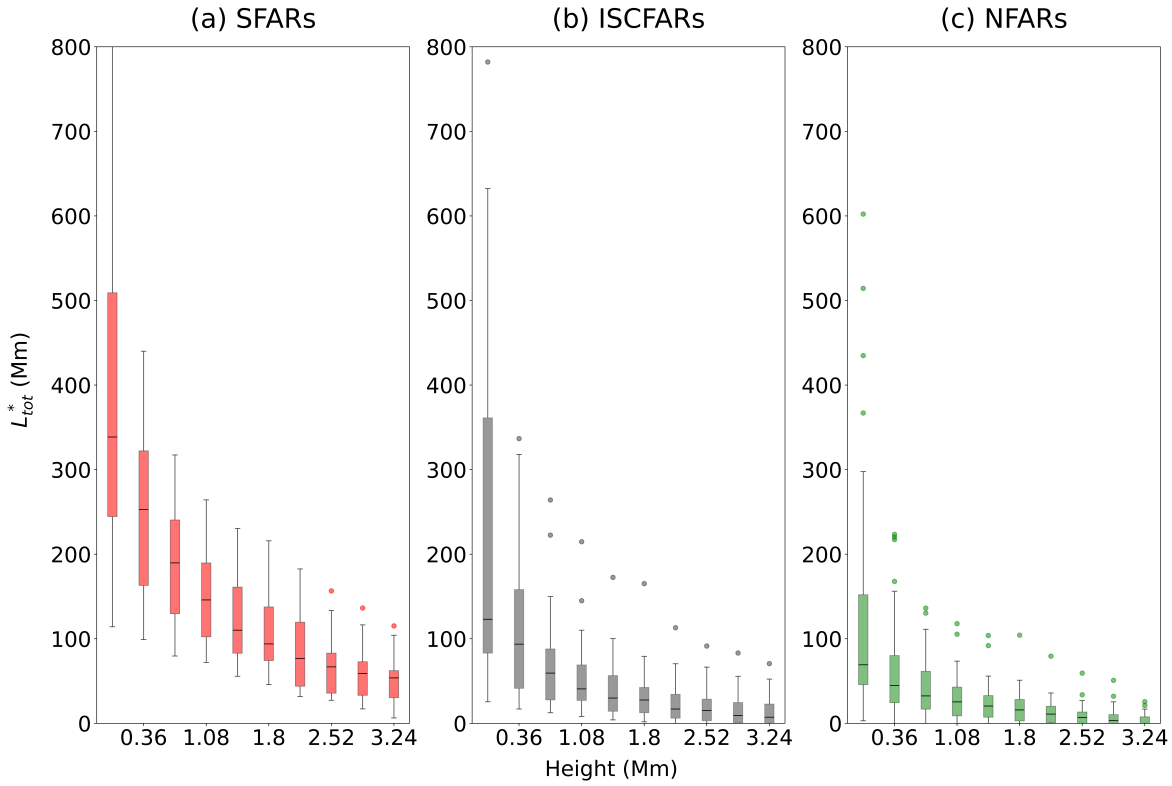

Figure 7.3: Height-wise (0 - 3.24 Mm) boxplots of L^*_{tot} for (a) SFARs, (b) ISCFARs and (c) NFARs; the lower and upper boundaries of each box represent the 25th (Q_1) and 75th (Q_3) percentiles; the median is denoted by a black line inside each box; the whisker lines denote the levels at 1.5 $\times$ IQR (inter-quartile range = $Q_3 - Q_1$) from Q_1 and Q_3; the coloured dots denote outliers

From the box-plot distribution (see Figure 7.3) we found that the upper whisker is longer than the bottom whisker for all the three categories, indicating that the data is skewed and not normally distributed. Further the IQRs or the middle 50% of the data for non-flaring ARs is clearly below the middle 50% of the data for strongly-flaring ARs. The same holds true for C-class flaring ARs (sampled independently) when compared to strongly flaring ARs at all heights except for the photosphere. In other words, the 75th percentile value of L^*_{tot} for non-flaring ARs is below the 25th percentile value of L^*_{tot} for strongly flaring ARs at all heights. Similarly the 75th percentile value of L^*_{tot} for independently sampled C-class flaring ARs is below the 25th percentile value of L^*_{tot} for strongly flaring ARs at all heights $\geq$ 0.36 Mm.

Considering all the three aforementioned categories together, we find that the box-plots in Figure 7.3 indicate that the heights 1.80 Mm and 2.16 Mm contain the least number of outliers, suggesting that these heights may be quite suitable for bringing about a distinction. In this regard we go one step further and compare how the minimum value of L^*_{tot} for strongly flaring ARs (i.e. L_{th}) at all heights compares with L^*_{tot} values of other categories (see Table 7.5). It is seen that the 1.80 Mm height is the most optimal height within the OHR to separate strongly flaring ARs from independently sampled C-class flaring ARs or from non-flaring ARs. More precisely, $\sim 96\%$ of time series of non-flaring ARs and $\sim 78\%$ time series of independently sampled C-class flaring ARs have L^*_{tot} below $L_{th} \sim 46$ Mm.

Height (Mm)	L_{th} (Mm)	NFAR	$(L^*_{tot} < L_{th})\%$	ISCFAR	$(L^*_{tot} < L_{th})\%$
0.00	114.24	28/45	62.22%	11/23	47.83%
0.36	99.03	37/45	82.22%	12/23	52.17%
0.72	79.68	38/45	84.44%	16/23	69.57%
1.08	72.07	42/45	93.33%	17/23	73.91%
1.44	55.67	42/45	93.33%	17/23	73.91%
1.80	**45.99**	**43/45**	**95.56%**	**18/23**	**78.26%**
2.16	31.62	43/45	95.56%	16/23	69.57%
2.52	27.38	43/45	95.56%	16/23	69.57%
2.88	17.12	39/45	86.67%	15/23	65.22%
3.24	6.45	31/45	68.89%	11/23	47.83%

Table 7.5: Variation of L_{th} i.e. the minimum L^*_{tot} value observed for SFARs with height (column 2) and details of the percentage of NFARs and ISCFARs which satisfied the condition: $L^*_{tot} < L_{th}$ at each height (columns 3–6). Optimal distinguishing power is achieved for a height of 1.8 Mm (in bold).

7.3.3 Inferential Statistics: Identifying Population Clusters

In Figure 7.2a, the $\mu + \sigma$ line of L^*_{tot} for non-flaring ARs is below the $\mu - \sigma$ line of L^*_{tot} for strongly flaring ARs at all heights ≥ 0.36 Mm. If we multiply σ by a fixed factor $k > 1$, then the gap between the two shaded regions can be minimised. Upon solving the linear equation: $\mu(SFAR)$ - $k \cdot \sigma(SFAR) = \mu(NFAR) + k \cdot \sigma(NFAR)$, we note that this gap can be reduced to zero for a unique value of k. Thus, k is given by:

$$k = \frac{\mu(SFAR) - \mu(NFAR)}{\sigma(SFAR) + \sigma(NFAR)} \tag{7.1}$$

For the nine heights in ascending order between 0.36 - 3.24 Mm, it is determined that $k(z) = \{1.17, 1.39, 1.40, 1.27, 1.30, 1.20, 1.18, 1.18, 1.33\}$. The maximum value k occurs at $z = 1.08$ Mm. We plotted the $\mu \pm 1.4\sigma$ profiles for non-flaring and strongly flaring ARs and at $z = 1.08$ Mm (see Figure 7.4), the boundary level was found to be at $L^*_{tot} \simeq 68$ Mm. Using Chebyshev's (one sided) inequality (discussed in Section 3.2), we obtain that in at least 66% of non-flaring ARs L^*_{tot} is guaranteed to be less than 68 Mm and in at least 66% of strongly flaring ARs, L^*_{tot} is guaranteed to be greater than 68 Mm. The figures we have quoted here are strictly inferential and describe the populations of strongly flaring and non-flaring ARs.

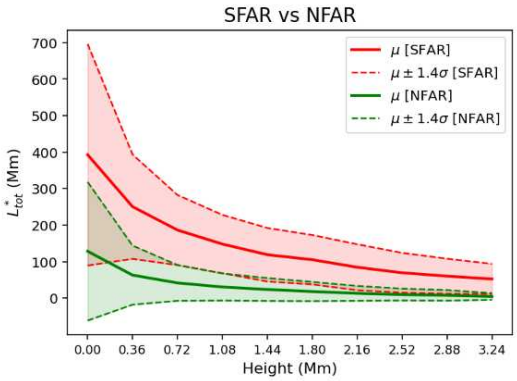

Figure 7.4: $\mu \pm 1.4\sigma$ profiles of L^*_{tot} for SFARs and NFARs; notice how at heights 0.72 - 1.08 Mm above the photosphere, the boundaries of SFARs and NFARs nearly touch each other

7.3.4 Implications on Flare Predictability

Preliminary analysis of L^*_{tot} decrease with height using standard statistical descriptors such as μ and σ showed that the OHR to distinguish strongly flaring ARs from non-flaring ARs was $\geq$ 0.36 Mm. It was also seen that even though strongly flaring ARs could not be separated from C-class flaring ARs when independently sampled, an OHR of 0.36 - 1.80 Mm could separate strongly flaring ARs from C-class flaring ARs when we sample such ARs 48 hr before the first C-class flare. Motivated by the presence of an OHR, we used two different techniques to determine the most optimal height within the extrapolated height ranges, to bring about the distinction. The two approaches are enumerated below.

1. Using descriptive statistics (which describes only the current data sample), we found that at a height of 1.80 Mm above the photosphere, the most optimal distinction between strongly flaring and non-flaring ARs may be obtained. About 96% non-flaring ARs had L^*_{tot} value less than the minimum L^*_{tot} value for strongly flaring ARs. Additionally it was seen that $\sim$78% of C-class flaring ARs (independently sampled) had L^*_{tot} value less than the minimum L^*_{tot} value for strongly flaring ARs.

2. The second approach was largely inferential and inferences were made about the distribution of L^*_{tot} for strongly flaring ARs and non-flaring ARs at a population level. Here, $z = 1.08$ Mm is the most optimal height to segregate strongly flaring AR population from non-flaring AR population. It is guaranteed that at least 66% of the data of strongly flaring ARs have $L^*_{tot} > 68$ Mm and at least 66% of the data of non-flaring ARs have $L^*_{tot} < 68$ Mm. It is to be noted that the percentage figure quoted here is a lower bound and not an expectation figure.

While these results described in the approach using descriptive statistics appear promising at this stage to develop a binary (TRUE/FALSE) model to detect the first X/M-class flare for an AR, we acknowledge that our current data sample has been sta-

tistically small and working on a larger statistical sample shall help refine the existing results.

The main benefit of using a larger statistical dataset is that we can obtain information regarding the the relative natural occurrence frequencies of (i) strongly flaring ARs (f_1), (ii) C-class flaring ARs (f_2) and (iii) non-flaring ARs (f_3). The number of samples of each of these categories in our dataset most likely do not accurately represent how frequently these categories of ARs naturally occur on the Sun. From a statistical dataset if we determine L_{th} (a threshold based on some relevant statistical metric e.g. minimum value, 5%ile value etc.) pertaining to the L^*_{tot} distribution for strongly flaring ARs, then the probabilities of $L^*_{tot} \geq L_{th}$ for strongly flaring ARs (p_1), C-class flaring ARs (p_2) and non-flaring ARs (p_3), may provide a much more accurate description of the effectiveness of L_{th} as a threshold.

Once we know the natural occurrence frequencies (f_1, f_2, f_3) and the threshold breaching probabilities (p_1, p_2, p_3), the problem of forecasting the first M/X-class flare for an AR over a 48-hr window (in a real-time scenario) may be translated into a simpler problem of Bayesian inference. If we find $L^*_{tot} \geq L_{th}$ at the optimal height, then the probability that it would produce an M/X-class flare within 48 hr shall be:

$$p = \frac{f_1 p_1}{f_1 p_1 + f_2 p_2 + f_3 p_3} \tag{7.2}$$

7.4 Summary

Overall, we find L^*_{tot} - the maximum value of L_{tot} in a 48-hr interval to be a quite promising flare prediction parameter. The computation of L_{tot} is quite similar to L_{SG} in terms of control parameters. While an association between $L_{SG} = 36$ Mm as a threshold and CME eruptivity within $\pm$ 48 hr, has been drawn previously by Falconer et al. (2003), L_{tot} may be seen as a counterpart of L_{SG}, suitable for studying the lower solar atmosphere between 0 - 3.24 Mm. In this regard, a statistical study of L_{tot} may be helpful to link L^*_{tot}-based thresholds to flare occurrence. In a best case scenario, this would be of significant value to space weather forecasting in general.

Both the (i) magnitude-based analysis in terms of μ and σ (using L^*_{tot} decrease profile with height) and the (ii) distribution or percentile-based analysis (using box-plots), suggest that strongly flaring and non-flaring ARs are better distinguished above the photosphere than at the photosphere. The same holds for independently sampled C-class flaring ARs. This highlights the importance of extrapolation-based studies as traditionally solar flare forecasting studies have been restricted to the photosphere.

From the magnitude analysis of L_{tot}, we found that around 96% of non-flaring ARs and 78% independently sampled C-class flaring ARs have $L^*_{tot} < 46$ Mm at a height of 1.80 Mm. We acknowledge this as a preliminary understanding of the populations of the aforementioned categories and hope that a statistical study shall refine our existing thresholds and present a more accurate understanding of PILs at extrapolated altitudes.

Chapter 8

Conclusions

This project set out to investigate whether key magnetic field–based predictors of solar flare activity, specifically the R-value parameter and the length of PILs can provide improved insight into flare predictability when examined beyond the photosphere and across multiple heights in the lower solar atmosphere. The work presented in Chapters 5, 6 and 7 collectively advances this objective by combining parameter-based analysis, time-series modelling, and statistical inference into a unified framework for understanding pre-eruptive magnetic evolution in ARs. Overall, the approach of analysis in this thesis is phenomenological. We start from experimental observations (i.e. the occurrence of flares) and then introduce effective parameters (or predictors) that are determined from data to analyse the data.

8.1 Conclusions on the R-value Parameter

The results shared by Schrijver (2007) with regards to the R-value, amounted to a huge leap in space weather forecasting as it was known that threshold in R-value (2×10^{21} Mx) could guarantee the occurrence of strong flares in under 24 hr. The findings in this thesis provide critical complementary information. If we had no information about the magnitude of R-value, by simply visually inspecting the height-wise stack-plots we could say that for flux emerging ARs, a jump in R-value may be expected 48 - 68 hr before an AR may produce its first X-class flare. Also the height-wise decay of R-value observing patterns such as the rate of decay with height and whether R-value gives finite/null outputs, could be used to differentiate between flaring and non-flaring ARs. A key outcome of Chapter 5 is the identification of systematic differences in the temporal evolution of R-values between flaring and non-flaring regions. X-class and M-class flaring ARs consistently exhibit: (i) higher magnitudes of R-values, (ii) more pronounced temporal build-up prior to flare onset and (iii) clear pre-eruptive trends distinguishable from background variability. In contrast, non-flaring regions lack such coherent R-value build-up. Further, the analysis of the R-value parameter across different classes of active regions demonstrates that magnetic flux concentration near PILs remains a robust diagnostic of flare productivity, but its predictive capability is significantly enhanced when examined in a multi-height and multi-model (i.e. with different apriori input values for the R-value parameter) framework.

8.2 Conclusions on the PIL Length Parameters

While the R-value captures magnetic flux concentration near PILs, the PIL length captures geometric complexity and spatial extent of polarity boundaries. In this connection we found that the evolution (time series) of PIL length parameters L_{tot} (total length of PILs) and L_{max} (length of longest PIL) are quite different from the evolution of R-value. Unlike the R-value which may show a clear jump prior to the first M/X-class flare for emergent ARs, these parameters do not show any such jumps. In Chapter 6 we tried to decompose the PIL length time series (L_{tot}/L_{max}) using various types of curve fits and autoregressive models. We found that the evolution of L_{max} is more stochastic compared to L_{tot} as the variance fraction from the white noise component was greater for L_{max} in almost all ARs we studied. This makes L_{tot} a much better parameter to trace PIL lengths compared to L_{max}. This could be because when there are multiple PILs in any AR, the longest PIL may differ from time to time as the AR continues to evolve. Unlike the R-value parameter, the evolution of PIL length parameters L_{tot} and L_{max} shows no specific patterns that may distinguish between flaring and non-flaring ARs.

However, in Chapter 7 we showed that by studying the height-wise evolution of the PIL length parameter L_{tot}, we may also be able to tell whether it is likely that a given AR is flaring or non-flaring. For example, we found from our dataset that about 78% of independently sampled CFARs and 96% of non-flaring ARs have the maximum value of L_{tot} within a 48 hr interval less than 46 Mm at a height of about 1.8 Mm. Given that there are PIL based metrics in the literature that correlate to CME-eruptivity, having a L^*_{tot}-based threshold (as a function of height above the photosphere) that correlates to flare occurrence would be key result in space weather forecasting.

We find that the time series of L_{tot} and its maximum value L^*_{tot} are useful tools to distinguish between different categories of flares. The extrapolation-based approach

used for computing PILs is justified because comparisons between (i) strongly flaring ARs (sampled in the 48 hr duration immediately before the first M/X-class flare) and non-flaring ARs and between (ii) strongly flaring ARs and C-class flaring ARs (prior to first C-class flare) led us to the finding that the $\mu \pm \sigma$ boundaries of the PIL parameter L^*_{tot} are much more clearly separated above the photosphere than at the photosphere for each of the two sets of aforementioned categories. In this context, we also found that as C-class flaring ARs continue to evolve beyond the occurrence of the first C-class flare, the $\mu \pm \sigma$ boundaries of L^*_{tot} begins to overlap with corresponding boundaries of strongly flaring ARs. Hence, we would recommend studying L_{tot} and L^*_{tot} within a height range of 1-2 Mm above the photosphere to distinguish between flaring and non-flaring regions. As we study more and more ARs in these height ranges, we shall have a more refined understanding of thresholds necessary to distinguish strong flaring ARs from weakly flaring or non-flaring ones. In this context, Chapter 7 effectively treats the problem of identifying strongly flaring ARs (by distinguishing from non-flaring or weakly flaring ARs) based on L^*_{tot} values as a conditional probability problem.

8.3 Future Work

Overall, this thesis has contributed to solar flare forecasting in the following ways:

1. Extending classical flare predictors into the lower solar atmosphere using PF extrapolation

2. Demonstrating of the importance of temporal evolution in magnetic parameters, with regards to flare prediction and forecasting

3. Evidencing and supporting the existence of optimal height ranges for prediction of flares

4. Identifying patterns with height and time to discriminate between strongly flaring and non-flaring ARs

5. Laying out a framework to forecast flares using magnetic field-based predictors in a real time scenario

A major limitation of this thesis has been that our dataset has been limited. Moving forward, we seek to repeat these numerical calculations for a much larger statistical dataset so that (i) a clearer picture regarding the distribution of optimal times (in terms of R-value) prior to flux emerging ARs may be obtained and (ii) a more refined understanding of the L^*_{tot}-based thresholds as a function of height may be obtained which may then be used to distinguish strongly flaring ARs from non-flaring ones. Throughout the thesis, we have used PF extrapolation to compute the magnetic field and subsequently the magnetic field-based predictors in three dimensions. Moving forward, we repeat our calculations using the NLFFF extrapolation approach to cross-check and validate our current findings. We are also looking extend this study to more predictors which are indicators of non-potentiality in an AR. The broader idea here is that once, a sufficient number of predictors are successfully tested using the current three dimensional approach, an efficient neural network model maybe developed using a select few of them, to address the problem of predicting and forecasting major flares.

Bibliography

Abramenko, V. I. (2005), 'Relationship between magnetic power spectrum and flare productivity in solar active regions', *The Astrophysical Journal* **629**(2), 1141.
URL: *https://doi.org/10.1086/431732*

Abramenko, V. I., Yurchyshyn, V. B., Wang, H., Spirock, T. J. & Goode, P. R. (2003), 'Signature of an avalanche in solar flares as measured by photospheric magnetic fields', *The Astrophysical Journal* **597**(2), 1135.
URL: *https://doi.org/10.1086/378492*

Abramenko, V. & Yurchyshyn, V. (2010), 'Intermittency and multifractality spectra of the magnetic field in solar active regions', *The Astrophysical Journal* **722**(1), 122.
URL: *https://doi.org/10.1088/0004-637X/722/1/122*

Airapetian, V. S., Barnes, R., Cohen, O., Collinson, G. A., Danchi, W. C., Dong, C. F., Del Genio, A. D., France, K., Garcia-Sage, K., Glocer, A., Gopalswamy, N., Grenfell, J. L., Gronoff, G., Güdel, M., Herbst, K., Henning, W. G., Jackman, C. H., Jin, M., Johnstone, C. P., Kaltenegger, L., Kay, C. D., Kobayashi, K., Kuang, W., Li, G., Lynch, B. J., Lüftinger, T., Luhmann, J. G., Maehara, H., Mlynczak, M. G., Notsu, Y., Osten, R. A., Ramirez, R. M., Rugheimer, S., Scheucher, M., Schlieder, J. E., Shibata, K., Sousa-Silva, C., Stamenković, V., Strangeway, R. J., Usmanov, A. V., Vergados, P., Verkhoglyadova, O. P., Vidotto, A. A., Voytek, M., Way, M. J., Zank, G. P. & Yamashiki, Y. (2020), 'Impact of space weather on climate and habitability of terrestrial-type exoplanets', *International Journal of Astrobiology* **19**(2), 136–194.

Akaike, H. (1974), 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control* **19**(6), 716–723.

Alissandrakis, C. E. (1981), 'On the computation of constant alpha force-free magnetic field', *aap* **100**(1), 197–200.

Amari, T., Luciani, J. F., Aly, J. J. & Tagger, M. (1996), 'Very Fast Opening of a Three-dimensional Twisted Magnetic Flux Tube', *apjl* **466**, L39.

Antiochos, S. K., DeVore, C. R. & Klimchuk, J. A. (1999), 'A model for solar coronal mass ejections', *The Astrophysical Journal* **510**(1), 485.
URL: *https://doi.org/10.1086/306563*

Antolin, P. (2019), 'Thermal instability and non-equilibrium in solar coronal loops: from coronal rain to long-period intensity pulsations', *Plasma Physics and Controlled Fusion* **62**(1), 014016.
URL: *https://doi.org/10.1088/1361-6587/ab5406*

Antolin, P. & Rouppe van der Voort, L. (2012), 'Observing the Fine Structure of Loops through High-resolution Spectroscopic Observations of Coronal Rain with the CRISP Instrument at the Swedish Solar Telescope', *apj* **745**(2), 152.

Archontis, V. & Syntelis, P. (2019), 'The emergence of magnetic flux and its role on the onset of solar dynamic events', *Philosophical Transactions of the Royal Society of London Series A* **377**(2148), 20180387.

Aschwanden, M. J. (2001), 'An evaluation of coronal heating models for active regions based on yohkoh, soho, and trace observations', *The Astrophysical Journal* **560**(2), 1035.
URL: *https://doi.org/10.1086/323064*

Aschwanden, M. J., Nitta, N. V., Wuelser, J.-P., Lemen, J. R., Sandman, A., Vourlidas, A. & Colaninno, R. C. (2009), 'First measurements of the mass of coronal mass

ejections from the euv dimming observed with stereo euvi a+b spacecraft', *The Astrophysical Journal* **706**(1), 376.
URL: *https://doi.org/10.1088/0004-637X/706/1/376*

Aulanier, G., Démoulin, P. & Grappin, R. (2005), 'Equilibrium and observational properties of line-tied twisted flux tubes', *aap* **430**, 1067–1087.

Aulanier, G., Démoulin, P., Schrijver, C. J., Janvier, M., Pariat, E. & Schmieder, B. (2013), 'The standard flare model in three dimensions. II. Upper limit on solar flare energy', *aap* **549**, A66.

Aulanier, G., Janvier, M. & Schmieder, B. (2012), 'The standard flare model in three dimensions. I. Strong-to-weak shear transition in post-flare loops', *aap* **543**, A110.

Aulanier, G., Török, T., Démoulin, P. & DeLuca, E. E. (2009), 'Formation of torus-unstable flux ropes and electric currents in erupting sigmoids', *The Astrophysical Journal* **708**(1), 314.

Ayachit, U. (2015), *The paraview guide: a parallel visualization application*, Kitware, Inc.

Baker, D. (1998), 'What is space weather?', *Advances in Space Research* **22**(1), 7–16. Solar-Terrestrial Relations: Predicting the Effects on the Near-Earth Environment.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117797010958*

Beckers, J. M. (1972), 'Solar Spicules', *araa* **10**, 73.

Benz, A. O. (2008), 'Flare observations', *Living Reviews in Solar Physics* **5**(1), 1.
URL: *https://doi.org/10.12942/lrsp-2008-1*

Biswal, S., Korsós, M. B., Georgoulis, M. K., Nindos, A., Patsourakos, S. & Erdélyi, R. (2024), 'Case studies on pre-eruptive x-class flares using r-value in the lower solar atmosphere', *The Astrophysical Journal* **974**(2), 259.
URL: *https://doi.org/10.3847/1538-4357/ad6c33*

Bloomfield, D. S., Higgins, P. A., McAteer, R. T. J. & Gallagher, P. T. (2012), 'Toward Reliable Benchmarking of Solar Flare Forecasting Methods', *apjl* **747**(2), L41.

Bobra, M. G., Sun, X., Hoeksema, J. T., Turmon, M., Liu, Y., Hayashi, K., Barnes, G. & Leka, K. D. (2014), 'The Helioseismic and Magnetic Imager (HMI) Vector Magnetic Field Pipeline: SHARPs – Space-Weather HMI Active Region Patches', *Solar Physics* **289**(9), 3549–3578.
URL: *http://link.springer.com/10.1007/s11207-014-0529-3*

Bolduc, L. (2002), 'Gic observations and studies in the hydro-quebec system', *Journal of Atmospheric and Solar-Terrestrial Physics* **64**, 1793–1802.

Boteler, D. H. (2019), 'A 21st century view of the march 1989 magnetic storm', *Space Weather* **17**(10), 1427–1441.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019SW002278*

Bray, R. J. & Loughhead, R. E. (1964), *Sunspots*.

Brekke, P., Kjeldseth-Moe, O. & Harrison, R. A. (1997), 'High-velocity flows in an active region loop system observed with the coronal diagnostic spectrometer (cds) on soho', *Solar Physics* **175**(2), 511–521.
URL: *https://doi.org/10.1023/A:1004950330900*

Bronarska, Wheatland, M. S., Gopalswamy, N. & Michalek, G. (2018), 'Very narrow coronal mass ejections producing solar energetic particles', *A&A* **619**, A34.
URL: *https://doi.org/10.1051/0004-6361/201833237*

Brooks, D. H., Ugarte-Urra, I. & Warren, H. P. (2008), 'The role of transient brightenings in heating the solar corona', *The Astrophysical Journal* **689**(1), L77.
URL: *https://doi.org/10.1086/595745*

Brosius, J. W. & Holman, G. D. (2007), 'Chromospheric Evaporation in a Remote Solar Flare-like Transient Observed at High Time Resolution with SOHO's CDS and RHESSI', *apjl* **659**(1), L73–L76.

Burtseva, O. & Petrie, G. (2013), 'Magnetic flux changes and cancellation associated with x-class and m-class flares', *Solar Physics* **283**(2), 429–452.
URL: *https://doi.org/10.1007/s11207-013-0241-8*

Byrne, Gallagher, P. T., McAteer, R. T. J. & Young, C. A. (2009), 'The kinematics of coronal mass ejections using multiscale methods', *A&A* **495**(1), 325–334.
URL: *https://doi.org/10.1051/0004-6361:200809811*

Cai, X., Aydin, B., Ji, A., Georgoulis, M. K. & Angryk, R. (2020), A framework for detecting polarity inversion lines from longitudinal magnetograms, *in* '2020 IEEE International Conference on Big Data (Big Data)', pp. 4175–4183.

Cannon, P., Angling, M., Barclay, L., Curry, C., Dyer, C., Edwards, R., Greene, G., Hapgood, M., Horne, R., Jackson, D., Mitchell, C., Owen, J., Richards, A., Rogers, C., Ryden, K., Saunders, S., Sweeting, M., Tanner, R., Thomson, A. & Underwood, C. (2013), 'Extreme space weather: impacts on engineered systems and infrastructure'.

Cargill, P. J. (2004), 'On the aerodynamic drag force acting on interplanetary coronal mass ejections', *Solar Physics* **221**(1), 135–149.
URL: *https://doi.org/10.1023/B:SOLA.0000033366.10725.a2*

Cargill, P. J. & Priest, E. R. (1982), 'Slow-shock heating and the kopp-pneuman model for 'post'-flare loops', *Solar Physics* **76**(2), 357–375.
URL: *https://doi.org/10.1007/BF00170991*

Carlsson, M., De Pontieu, B. & Hansteen, V. H. (2019), 'New view of the solar chromosphere', *Annual Review of Astronomy and Astrophysics* **57**(Volume 57, 2019), 189–226.
URL: *https://www.annualreviews.org/content/journals/10.1146/annurev-astro-081817-052044*

Carmichael, H. (1964), A Process for Flares, *in* 'NASA Special Publication', Vol. 50, p. 451.

Charbonneau, P. (2010), 'Dynamo models of the solar cycle', *Living Reviews in Solar Physics* **7**(1), 3.
URL: *https://doi.org/10.12942/lrsp-2010-3*

Chen, P. F. (2011), 'Coronal mass ejections: Models and their observational basis', *Living Reviews in Solar Physics* **8**(1), 1.
URL: *https://doi.org/10.12942/lrsp-2011-1*

Chierichini, Simone, Francisco, Gregoire, Mugatwala, Ronish, Foldes, Raffaello, Camporeale, Enrico, De Gasperis, Giancarlo, Giovannelli, Luca, Napoletano, Gianluca, Del Moro, Dario & Erdelyi, Robertus (2024), 'A bayesian approach to the drag-based modelling of icmes', *J. Space Weather Space Clim.* **14**, 1.
URL: *https://doi.org/10.1051/swsc/2023032*

Chintzoglou, G., Zhang, J., Cheung, M. C. & Kazachenko, M. (2019), 'The origin of major solar activity: Collisional shearing between nonconjugated polarities of multiple bipoles emerging within active regions', *The Astrophysical Journal* **871**(1), 67.

Clark, P. C. & Bonnell, I. A. (2005), 'The onset of collapse in turbulently supported molecular clouds', *Monthly Notices of the Royal Astronomical Society* **361**(1), 2–16.
URL: *https://doi.org/10.1111/j.1365-2966.2005.09105.x*

Conlon, P. A., Gallagher, P. T., McAteer, R. T. J., Ireland, J., Young, C. A., Kestener, P., Hewett, R. J. & Maguire, K. (2008), 'Multifractal Properties of Evolving Active Regions', *solphys* **248**(2), 297–309.

Cranmer, S. R. (2009), 'Coronal holes', *Living Reviews in Solar Physics* **6**(1), 3.
URL: *https://doi.org/10.12942/lrsp-2009-3*

Cranmer, S. R., Gibson, S. E. & Riley, P. (2017), 'Origins of the ambient solar wind: Implications for space weather', *Space Science Reviews* **212**(3), 1345–1384.
URL: *https://doi.org/10.1007/s11214-017-0416-y*

Cucinotta, F. A. (2015), 'Review of nasa approach to space radiation risk assessments for mars exploration', *Health Physics* **108**(2).
URL: *https://journals.lww.com/health-physics/fulltext/2015/02000/*

De Moortel, I. & Browning, P. (2015), 'Recent advances in coronal heating', *Philosophical Transactions of the Royal Society of London Series A* **373**(2042), 20140269–20140269.

De Pontieu, B., Erdélyi, R. & James, S. P. (2004), 'Solar chromospheric spicules from the leakage of photospheric oscillations and flows', *Nature* **430**(6999), 536–539.
URL: *https://doi.org/10.1038/nature02749*

De Pontieu, B., Hansteen, V. H., Rouppe van der Voort, L., van Noort, M. & Carlsson, M. (2007), 'High-resolution observations and modeling of dynamic fibrils', *The Astrophysical Journal* **655**(1), 624.
URL: *https://doi.org/10.1086/509070*

De Vita, Gaetano, Vecchio, Antonio, Sorriso-Valvo, Luca, Briand, Carine, Primavera, Leonardo, Servidio, Sergio, Lepreti, Fabio & Carbone, Vincenzo (2015), 'Cancellation analysis of current density in solar active region noaa10019', *J. Space Weather Space Clim.* **5**, A28.
URL: *https://doi.org/10.1051/swsc/2015029*

Decraemer, B., Zhukov, A. N. & Van Doorsselaere, T. (2019), 'Three-dimensional density structure of a solar coronal streamer observed by soho/lasco and stereo/cor2 in quadrature', *The Astrophysical Journal* **883**(2), 152.
URL: *https://doi.org/10.3847/1538-4357/ab3b58*

Degl'innocenti & Landi, E. (2004), *Radiative Transfer for Polarized Radiation*, Springer Netherlands, Dordrecht, pp. 349–373.
URL: *https://doi.org/10.1007/1-4020-2415-0_8*

Del Toro Iniesta, J. C. & Ruiz Cobo, B. (1996), 'Stokes Profiles Inversion Techniques', *solphys* **164**(1-2), 169–182.

Del Zanna & Woods (2013), 'Spectral diagnostics with the sdo eve flare lines', *A&A* **555**, A59.
URL: *https://doi.org/10.1051/0004-6361/201220988*

Dere, K. P., Brueckner, G. E., Howard, R. A., Michels, D. J. & Delaboudiniere, J. P. (1999), 'LASCO and EIT Observations of Helical Structure in Coronal Mass Ejections', *apj* **516**(1), 465–474.

Desai, M. & Giacalone, J. (2016), 'Large gradual solar energetic particle events', *Living Reviews in Solar Physics* **13**(1), 3.
URL: *https://doi.org/10.1007/s41116-016-0002-5*

Dickey, D. A. & Fuller, W. A. (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**(366a), 427–431.
URL: *https://doi.org/10.1080/01621459.1979.10482531*

Dudík, J., Aulanier, G., Lörinčík, J. & Zemanová, A. (2025), 'Quasi-separatrix layers and three-dimensional magnetic reconnection: Theory and observations of solar flares (invited review)', *Solar Physics* **300**(10), 139.
URL: *https://doi.org/10.1007/s11207-025-02549-2*

Dumbović, M., Čalogović, J., Vršnak, B., Temmer, M., Mays, M. L., Veronig, A. & Piantschitsch, I. (2018), 'The drag-based ensemble model (dbem) for coronal mass ejection propagation', *The Astrophysical Journal* **854**(2), 180.
URL: *https://doi.org/10.3847/1538-4357/aaaa66*

Démoulin, P., Priest, E. R. & Lonie, D. P. (1996), 'Three-dimensional magnetic reconnection without null points: 2. application to twisted flux tubes', *Journal of Geophysical Research: Space Physics* **101**(A4), 7631–7646.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95JA03558*

Eastwood, J. P., Biffis, E., Hapgood, M. A., Green, L., Bisi, M. M., Bentley, R. D., Wicks, R., McKinnell, L.-A., Gibbs, M. & Burnett, C. (2017), 'The economic impact of space weather: Where do we stand?', *Risk Analysis* **37**(2), 206–218.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.12765*

Edlén, B. (1943), 'Die Deutung der Emissionslinien im Spektrum der Sonnenkorona. Mit 6 Abbildungen.', *zap* **22**, 30.

Elliott, G., Rothenberg, T. J. & Stock, J. H. (1996), 'Efficient tests for an autoregressive unit root', *Econometrica* **64**(4), 813–836. Full publication date: Jul., 1996.
URL: *https://doi.org/10.2307/2171846*

Erdélyi, Korsós, Marianna B., Huang, Xin, Yang, Yong, Pizzey, Danielle, Wrathmall, Steven A., Hughes, Ifan G., Dyer, Martin J., Dhillon, Vikram S., Belucz, Bernadett, Brajsa, Roman, Chatterjee, Piyali, Cheng, Xuewu, Deng, Yuanyong, Domínguez, Santiago Vargas, Joya, Raúl, Gömöry, Peter, Gyenge, Norbert G., Hanslmeier, Arnold, Kucera, Ales, Kuridze, David, Li, Faquan, Liu, Zhong, Xu, Long, Mathioudakis, Mihalis, Matthews, Sarah, McAteer, James R.T., Pevtsov, Alexei A., Pötzi, Werner, Romano, Paolo, Shen, Jinhua, Temesváry, János, Tlatov, Andrey G., Triana, Charles, Utz, Dominik, Veronig, Astrid M., Wang, Yuming, Yan, Yihua, Zaqarashvili, Teimuraz & Zuccarello, Francesca (2022), 'The solar activity monitor network - samnet', *J. Space Weather Space Clim.* **12**, 2.
URL: *https://doi.org/10.1051/swsc/2021025*

Erdélyi, R. & Ballai, I. (2007), 'Heating of the solar and stellar coronae: a review', *Astronomische Nachrichten* **328**(8), 726–733.

Falconer, D. A., Moore, R. L. & Gary, G. A. (2002), 'Correlation of the coronal mass ejection productivity of solar active regions with measures of their global nonpotentiality from vector magnetograms: Baseline results', *The Astrophysical Journal* **569**(2), 1016.
URL: *https://dx.doi.org/10.1086/339161*

Falconer, D. A., Moore, R. L. & Gary, G. A. (2003), 'A measure from line-of-sight magnetograms for prediction of coronal mass ejections', *Journal of Geophysical Research: Space Physics* **108**(A10).
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JA010030*

Falconer, D. A., Moore, R. L. & Gary, G. A. (2008), 'Magnetogram measures of total nonpotentiality for prediction of solar coronal mass ejections from active regions of any degree of magnetic complexity', *The Astrophysical Journal* **689**(2), 1433.
URL: *https://doi.org/10.1086/591045*

Fan, Y. & Gibson, S. E. (2003), 'The emergence of a twisted magnetic flux tube into a preexisting coronal arcade', *The Astrophysical Journal* **589**(2), L105.
URL: *https://doi.org/10.1086/375834*

Filippov, B. (2019), 'Difference of source regions between fast and slow coronal mass ejections', *Publications of the Astronomical Society of Australia* **36**, e022.

Forbes, T. G., Linker, J. A., Chen, J., Cid, C., Kóta, J., Lee, M. A., Mann, G., Mikić, Z., Potgieter, M. S., Schmidt, J. M., Siscoe, G. L., Vainio, R., Antiochos, S. K. & Riley, P. (2006), 'Cme theory and models', *Space Science Reviews* **123**(1), 251–302.
URL: *https://doi.org/10.1007/s11214-006-9019-8*

Forbes, T. G. & Malherbe, J. M. (1986), 'A Shock Condensation Mechanism for Loop Prominences', *apjl* **302**, L67.

Forbes, T. G. & Priest, E. R. (1984), 'Numerical simulation of reconnection in an

emerging magnetic flux region', *Solar Physics* **94**(2), 315–340.
URL: *https://doi.org/10.1007/BF00151321*

Forbes, T. G. & Priest, E. R. (1995), 'Photospheric Magnetic Field Evolution and Eruptive Flares', *apj* **446**, 377.

Gallagher, P. T., Moon, Y.-J. & Wang, H. (2002), 'Active-region monitoring and flare forecasting – i. data processing and first results', *Solar Physics* **209**(1), 171–183.
URL: *https://doi.org/10.1023/A:1020950221179*

Garcia, H. A. (1994), 'Temperature and emission measure from goes soft x-ray measurements', *Solar Physics* **154**(2), 275–308.

Garland, S. H., Yurchyshyn, V. B., Loper, R. D., Akers, B. F. & Emmons, D. J. (2023), 'Evolution of coronal magnetic field parameters during x5.4 solar flare', *Frontiers in Astronomy and Space Sciences* **Volume 10 - 2023**.
URL: *https://www.frontiersin.org/journals/astronomy-and-space-sciences/articles/10.3389/fspas.2023.1148293*

Gary, G. A. (2001), 'Plasma beta above a solar active region: rethinking the paradigm', *Solar Physics* **203**(1), 71–86.

Georgoulis, M. K., Bloomfield, D. S., Piana, M., Massone, A. M., Soldati, M., Gallagher, P. T., Pariat, E., Vilmer, N., Buchlin, E., Baudin, F. et al. (2021), 'The flare likelihood and region eruption forecasting (flarecast) project: flare forecasting in the big data & machine learning era', *Journal of Space Weather and Space Climate* **11**, 39.

Georgoulis, M. K. & LaBonte, B. J. (2007), 'Magnetic energy and helicity budgets in the active region solar corona. i. linear force-free approximation', *The Astrophysical Journal* **671**(1), 1034.
URL: *https://doi.org/10.1086/521417*

Georgoulis, M. K., Nindos, A. & Zhang, H. (2019), 'The source and engine of coronal mass ejections', *Philosophical Transactions of the Royal Society of London Series A* **377**(2148), 20180094.

Georgoulis, M. K. & Rust, D. M. (2007), 'Quantitative forecasting of major solar flares', *The Astrophysical Journal* **661**(1), L109.
URL: *https://doi.org/10.1086/518718*

Georgoulis, M. K., Rust, D. M., Bernasconi, P. N. & Schmieder, B. (2002), 'Statistics, morphology, and energetics of ellerman bombs', *The Astrophysical Journal* **575**(1), 506.
URL: *https://doi.org/10.1086/341195*

Georgoulis, M. K., Titov, V. S. & Mikić, Z. (2012), 'Non-neutralized electric current patterns in solar active regions: Origin of the shear-generating lorentz force', *The Astrophysical Journal* **761**(1), 61.
URL: *https://doi.org/10.1088/0004-637X/761/1/61*

Georgoulis, M. K., Yardley, S., Guerra, J., Murray, S., Ahmadzadeh, A., Anastasiadis, A., Angryk, R. A., Aydin, B., Banerjee, D., Barnes, G., Bemporad, A., Benvenuto, F., Bloomfield, S., Bobra, M., Campi, C., Camporeale, E., DeForest, C., Emslie, A. G., Falconer, D. A., Feng, L., Gan, W., Green, L., Guastavino, S., Hapgood, M., Kempton, D. J., Kitiashvili, I., Kontogiannis, I., Korsos, M., Leka, K. D., Massa, P., Massone, A. M., Nandi, D., Nindos, A., Papaioannou, A., Park, S.-H., Patsourakos, S., Piana, M., Rawafi, N., Sadykov, V., Toriumi, S., Vourlidas, A., Wang, H., Wang, J., Whitman, K., Yan, Y. & Zhukov, A. (2024), Prediction of Solar Energetic Events Impacting Space Weather Conditions: A Roadmap of Findings and Recommendations, *in* '45th COSPAR Scientific Assembly', Vol. 45, p. 2903.

Giannattasio, F., Consolini, G., Berrilli, F. & De Michelis, P. (2022), 'Scaling properties of magnetic field fluctuations in the quiet sun', *A&A* **659**, A180.
URL: *https://doi.org/10.1051/0004-6361/202142940*

Giannattasio, F., Consolini, G., Berrilli, F. & Del Moro, D. (2019), 'The Complex Nature of Magnetic Element Transport in the Quiet Sun: The Lévy-walk Character', *apj* **878**(1), 33.

Gomes, R., Levison, H. F., Tsiganis, K. & Morbidelli, A. (2005), 'Origin of the cataclysmic late heavy bombardment period of the terrestrial planets', *Nature* **435**(7041), 466–469.
URL: *https://doi.org/10.1038/nature03676*

Gopalswamy, N. (2006), 'Coronal mass ejections of solar cycle 23', *Journal of Astrophysics and Astronomy* **27**(2), 243–254.
URL: *https://doi.org/10.1007/BF02702527*

Gopalswamy, N. (2016), 'History and development of coronal mass ejections as a key player in solar terrestrial relationship', *Geoscience Letters* **3**(1), 8.
URL: *https://doi.org/10.1186/s40562-016-0039-2*

Gopalswamy, N. (2022), 'The sun and space weather', *Atmosphere* **13**(11).
URL: *https://www.mdpi.com/2073-4433/13/11/1781*

Gopalswamy, S., Y., G., M., G., S., A., V., S., F. & R., H. (2009), 'The soho/lasco cme catalog', *Earth, Moon, and Planets* **104**(1-4), 295–313.
URL: *https://cir.nii.ac.jp/crid/1362825894469231488*

Gordovskyy, Browning, P. K. & Vekstein, G. E. (2010), 'Particle acceleration in a transient magnetic reconnection event', *A&A* **519**, A21.
URL: *https://doi.org/10.1051/0004-6361/200913569*

Gosling, J. T. (1990), *Coronal Mass Ejections and Magnetic Flux Ropes in Interplanetary Space*, American Geophysical Union (AGU), pp. 343–364.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GM058p0343*

Gou, T., Veronig, A. M., Liu, R., Zhuang, B., Dumbović, M., Podladchikova, T., Reid, H. A. S., Temmer, M., Dissauer, K., Vršnak, B. & Wang, Y. (2020), 'Solar flare–cme

coupling throughout two acceleration phases of a fast cme', *The Astrophysical Journal Letters* **897**(2), L36.
URL: *https://doi.org/10.3847/2041-8213/ab9ec5*

Green, Kliem & Wallace, A. J. (2011), 'Photospheric flux cancellation and associated flux rope formation and eruption', *A&A* **526**, A2.
URL: *https://doi.org/10.1051/0004-6361/201015146*

Green, L. M., Török, T., Vršnak, B., Manchester, W. & Veronig, A. (2018), 'The Origin, Early Evolution and Predictability of Solar Eruptions', *ssr* **214**(1), 46.

Guo, J., Zhang, H. Q., Chumak, O. V. & Lin, J. B. (2010), 'The evolution of the quantified magnetic complexity of active regions in the 23rd solar cycle', *Monthly Notices of the Royal Astronomical Society* **405**(1), 111–114.
URL: *https://doi.org/10.1111/j.1365-2966.2010.16465.x*

Hagenaar, H. J. (2001), 'Ephemeral regions on a sequence of full-disk michelson doppler imager magnetograms', *The Astrophysical Journal* **555**(1), 448.
URL: *https://doi.org/10.1086/321448*

Hagyard, M. J., Smith, J. B., Teuber, D. & West, E. A. (1984), 'A quantitative study relating observed shear in photospheric magnetic fields to repeated flaring', *Solar Physics* **91**(1), 115–126.
URL: *https://doi.org/10.1007/BF00213618*

Hale, G. E., Ellerman, F., Nicholson, S. B. & Joy, A. H. (1919), 'The Magnetic Polarity of Sun-Spots', *apj* **49**, 153.

Hale, G. E. & Nicholson, S. B. (1925), 'The Law of Sun-Spot Polarity', *apj* **62**, 270.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press, Princeton.
URL: *https://doi.org/10.1515/9780691218632*

Hansen, R. T., Garcia, C. J., Grognard, R. J.-M. & Sheridan, K. V. (1971), 'A coronal disturbance observed simultaneously with a white-light corona-meter and the 80 MHz Culgoora radioheliograph', *pasa* **2**, 57.

Hansteen, V. H., De Pontieu, B., Rouppe van der Voort, L., van Noort, M. & Carlsson, M. (2006), 'Dynamic fibrils are driven by magnetoacoustic shocks', *The Astrophysical Journal* **647**(1), L73.
URL: *https://doi.org/10.1086/507452*

Harrison, R. A. (1995), 'The nature of solar flares associated with coronal mass ejection.', *aap* **304**, 585.

Hathaway, D. H. (2010), 'The solar cycle', *Living Reviews in Solar Physics* **7**(1), 1.
URL: *https://doi.org/10.12942/lrsp-2010-1*

Hathaway, D. H. (2015), 'The Solar Cycle', *Living Reviews in Solar Physics* **12**(1), 4.

Hayakawa, H., Tamazawa, H., Kawamura, A. D. & Isobe, H. (2015), 'Records of sunspot and aurora during ce 960–1279 in the chinese chronicle of the sòng dynasty', *Earth, Planets and Space* **67**(1), 82.
URL: *https://doi.org/10.1186/s40623-015-0250-y*

Haynes, A. L. & Parnell, C. E. (2007), 'A trilinear method for finding null points in a three-dimensional vector space', *Physics of Plasmas* **14**(8), 082107.
URL: *https://doi.org/10.1063/1.2756751*

Hazra, S., Sardar, G. & Chowdhury, P. (2020), 'Distinguishing between flaring and nonflaring active regions', *aap* **639**, A44.

Hertzsprung, E. (1905), 'Zur strahlung der sterne, teil 1', *Zeitschrift für wissenschaftliche Photographie, Photochemie und Photophysik* **3**, 429–442.

Hewett, R. J., Gallagher, P. T., McAteer, R. T. J., Young, C. A., Ireland, J., Conlon, P. A. & Maguire, K. (2008), 'Multiscale analysis of active region evolution', *Solar*

Physics **248**(2), 311–322.
URL: *https://doi.org/10.1007/s11207-007-9028-0*

Hirayama, T. (1974), 'Theoretical Model of Flares and Prominences. I: Evaporating Flare Model', *solphys* **34**(2), 323–338.

Hoyng, P., Duijveman, A., Machado, M. E., Rust, D. M., Svestka, Z., Boelee, A., de Jager, C., Frost, K. T., Lafleur, H., Simnett, G. M., van Beek, H. F. & Woodgate, B. E. (1981), 'Origin and Location of the Hard X-Ray Emission in a Two-Ribbon Flare', *apjl* **246**, L155.

Hudson, H. S. (2011), 'Global properties of solar flares', *Space Science Reviews* **158**(1), 5–41.
URL: *https://doi.org/10.1007/s11214-010-9721-4*

Hundhausen, A. J. (1993), 'Sizes and locations of coronal mass ejections: SMM observations from 1980 and 1984-1989', *jgr* **98**(A8), 13177–13200.

Illing, R. M. E. & Hundhausen, A. J. (1985), 'Observation of a coronal transient from 1.2 to 6 solar radii', *Journal of Geophysical Research: Space Physics* **90**(A1), 275–282.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JA090iA01p00275*

Inoue, S., Magara, T., Pandey, V. S., Shiota, D., Kusano, K., Choe, G. S. & Kim, K. S. (2014), 'Non-linear force free extrapolation of the coronal magnetic field based on the magnetohydrodynamic relaxation method', *The Astrophysical Journal* **780**(1), 101.

Isavnin, A., Vourlidas, A. & Kilpua, E. K. J. (2014), 'Three-Dimensional Evolution of Flux-Rope CMEs and Its Relation to the Local Orientation of the Heliospheric Current Sheet', *solphys* **289**(6), 2141–2156.

Ishii, M., Berdermann, J., Forte, B., Hapgood, M., Bisi, M. M. & Romano, V. (2024), 'Space weather impact on radio communication and navigation', *Advances in Space*

Research .
URL: *https://www.sciencedirect.com/science/article/pii/S0273117724000863*

Iucci, N., Levitin, A. E., Belov, A. V., Eroshenko, E. A., Ptitsyna, N. G., Villoresi, G., Chizhenkov, G. V., Dorman, L. I., Gromova, L. I., Parisi, M., Tyasto, M. I. & Yanke, V. G. (2005), 'Space weather conditions and spacecraft anomalies in different orbits', *Space Weather* **3**(1).
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003SW000056*

Janvier, M., Aulanier, G., Pariat, E. & Démoulin, P. (2013), 'The standard flare model in three dimensions. III. Slip-running reconnection properties', *aap* **555**, A77.

Jeffrey, N. L. S., Fletcher, L., Labrosse, N. & Simões, P. J. A. (2018), 'The development of lower-atmosphere turbulence early in a solar flare', *Science Advances* **4**(12), eaav2794.
URL: *https://www.science.org/doi/abs/10.1126/sciadv.aav2794*

Ji, A., Cai, X., Khasayeva, N., Georgoulis, M. K., Martens, P. C., Angryk, R. A. & Aydin, B. (2023), 'A systematic magnetic polarity inversion line data set from sdo/hmi magnetograms', *The Astrophysical Journal Supplement Series* **265**(1), 28.
URL: *https://dx.doi.org/10.3847/1538-4365/acb43a*

Jing, J., Liu, C., Lee, J., Ji, H., Liu, N., Xu, Y. & Wang, H. (2018), 'Statistical analysis of torus and kink instabilities in solar eruptions', *Astrophysical Journal* **864**(2). Publisher Copyright: © 2018. The American Astronomical Society. All rights reserved.

Jing, J., Yuan, Y., Reardon, K., Wiegelmann, T., Xu, Y. & Wang, H. (2011), 'Non-potentiality of chromospheric fibrils in noaa active regions 11092 and 9661', *The Astrophysical Journal* **739**(2), 67.
URL: *https://doi.org/10.1088/0004-637X/739/2/67*

Kiepenheuer, K. O. (1953), Solar Activity, *in* G. P. Kuiper, ed., 'The Sun', p. 322.

Kliem, B. & Török, T. (2006), 'Torus Instability', *prl* **96**(25), 255002.

Klimchuk, J. A. (2006), 'On Solving the Coronal Heating Problem', *solphys* **234**(1), 41–77.

Komm, R. W. (2007), 'Value-added maps: fluid-dynamics descriptors from ring diagrams', *Astronomische Nachrichten* **328**(3), 269.

Kontogiannis, I., Georgoulis, M. K., Park, S.-H. & Guerra, J. A. (2018), 'Testing and improving a set of morphological predictors of flaring activity', *Solar Physics* **293**(6), 1–18.

Kopp, R. A. & Pneuman, G. W. (1976), 'Magnetic reconnection in the corona and the loop prominence phenomenon.', *solphys* **50**(1), 85–98.

Korsos, M. (2018), Leap forward in space weather forecast: Novel prediction of flares. Unpublished.
URL: *https://etheses.whiterose.ac.uk/id/eprint/22683/*

Korsós, M. B., Gyenge, N., Baranyi, T. & Ludmány, A. (2015), 'Dynamic precursors of flares in active region noaa 10486', *Journal of Astrophysics and Astronomy* **36**(1), 111–121.
URL: *https://doi.org/10.1007/s12036-015-9329-x*

Korsós, M., Chatterjee, P. & Erdélyi, R. (2018), 'Applying the weighted horizontal magnetic gradient method to a simulated flaring active region', *The Astrophysical Journal* **857**(2), 103.

Korsós, M., Georgoulis, M., Gyenge, N., Bisoi, S., Yu, S., Poedts, S., Nelson, C., Liu, J., Yan, Y. & Erdélyi, R. (2020), 'Solar flare prediction using magnetic field diagnostics above the photosphere', *The Astrophysical Journal* **896**(2), 119.

Korsós, M. B., Baranyi, T. & Ludmány, A. (2014), 'Pre-flare dynamics of sunspot groups', *The Astrophysical Journal* **789**(2), 107.
URL: *https://doi.org/10.1088/0004-637X/789/2/107*

Korsós, M. B., Erdélyi, R., Huang, X. & Morgan, H. (2022), 'Magnetic helicity flux oscillations in the atmospheres of flaring and nonflaring active regions', *The Astrophysical Journal* **933**(1), 66.
URL: *https://doi.org/10.3847/1538-4357/ac7469*

Korsós, M. B., Jarolim, R., Erdélyi, R., Veronig, A. M., Morgan, H. & Zuccarello, F. (2024), 'First insights into the applicability and importance of different 3d magnetic field extrapolation approaches for studying the preeruptive conditions of solar active regions', *The Astrophysical Journal* **962**(2), 171.
URL: *https://dx.doi.org/10.3847/1538-4357/ad18bd*

Krieger, A. S., Timothy, A. F. & Roelof, E. C. (1973), 'A coronal hole and its identification as the source of a high velocity solar wind stream', *Solar Physics* **29**(2), 505–525.
URL: *https://doi.org/10.1007/BF00150828*

Krucker, S. (2002), 'Nanoflare heating in the quiet corona: evidence and problems', *Advances in Space Research* **30**(3), 493.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117702003253*

Künzel, H. (1965), 'Zur Klassifikation von Sonnenfleckengruppen', *Astronomische Nachrichten* **288**, 177.

Kuridze, D., Morton, R. J., Erdélyi, R., Dorrian, G. D., Mathioudakis, M., Jess, D. B. & Keenan, F. P. (2012), 'Transverse oscillations in chromospheric mottles', *The Astrophysical Journal* **750**(1), 51.
URL: *https://doi.org/10.1088/0004-637X/750/1/51*

Kusano., K., Iijima, H., Kaneko, T., Masuda, S., IJu, T. & Inoue, S. (2021), 'Isee database for nonlinear force-free field of solar active regions'.

Kusano., K., Iju, T., Bamba, Y. & Inoue, S. (2020), 'A physics-based method that can predict imminent large solar flares', *Science* **369**(6503), 587–591.

Künzel, H. (1959), 'Die flare-häufigkeit in fleckengruppen unterschiedlicher klasse und magnetischer struktur (mitteilungen des astrophysikalischen observatoriums potsdam nr. 87)', *Astronomische Nachrichten* **285**(5-6), 271–273.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/asna.19592850516*

Lakhina, G. S. & Tsurutani, B. T. (2016), 'Geomagnetic storms: historical perspective to modern view', *Geoscience Letters* **3**(1), 5.
URL: *https://doi.org/10.1186/s40562-016-0037-4*

Lang, K. R., ed. (2000), *The sun from space*.

LeBlanc, F. (2010), *An Introduction to Stellar Astrophysics*.

Leenaarts, J., Carlsson, M. & Rouppe van der Voort, L. (2012), 'The formation of the hα line in the solar chromosphere', *The Astrophysical Journal* **749**(2), 136.
URL: *https://doi.org/10.1088/0004-637X/749/2/136*

Leka, K. D. & Barnes, G. (2003), 'Photospheric magnetic field properties of flaring versus flare-quiet active regions. i. data, general approach, and sample results', *The Astrophysical Journal* **595**(2), 1277.
URL: *https://doi.org/10.1086/377511*

Lin, Z. & Bai, Z. (2011), *Bounds of Probabilities in Terms of Moments*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 51–66.
URL: *https://doi.org/10.1007/978-3-642-05261-3_6"*

Lineweaver, C. H. (2001), 'An Estimate of the Age Distribution of Terrestrial Planets in the Universe: Quantifying Metallicity as a Selection Effect', *icarus* **151**(2), 307–313.

Liu, L., Cheng, X., Wang, Y. & Zhou, Z. (2019), 'Formation of a magnetic flux rope in the early emergence phase of noaa active region 12673', *The Astrophysical Journal*

884(1), 45.
URL: *https://dx.doi.org/10.3847/1538-4357/ab3c6c*

Liu, Y. (2008), 'Magnetic field overlying solar eruption regions and kink and torus instabilities', *The Astrophysical Journal* **679**(2), L151.
URL: *https://doi.org/10.1086/589282*

Liu, Y., Hoeksema, J. T., Scherrer, P. H., Schou, J., Couvidat, S., Bush, R. I., Duvall, T. L., Hayashi, K., Sun, X. & Zhao, X. (2012), 'Comparison of line-of-sight magnetograms taken by the solar dynamics observatory/helioseismic and magnetic imager and solar and heliospheric observatory/michelson doppler imager', *Solar Physics* **279**(1), 295–316.
URL: *https://doi.org/10.1007/s11207-012-9976-x*

Livingston, W. & Watson, F. (2015), 'A new solar signal: Average maximum sunspot magnetic fields independent of activity cycle', *grl* **42**(21), 9185–9189.

Lloyd's (2013), 'Solar storm risk to the north american electric gridrep'.
URL: *https://https://assets.lloyds.com/assets/pdf-solar-storm-risk-to-the-north-american-electric-grid/1/pdf-Solar-Storm-Risk-to-the-North-American-Electric-Grid.pdf*

Longcope, D., Beveridge, C., Qiu, J., Ravindra, B., Barnes, G. & Dasso, S. (2007), 'Modeling and measuring the flux reconnected and ejected by the two-ribbon flare/cme event on 7 november 2004', *Solar Physics* **244**(1), 45–73.
URL: *https://doi.org/10.1007/s11207-007-0330-7*

Low, B. C. (1994), Coronal mass ejections and magnetic helicity, *in* J. J. Hunt, ed., 'Solar Dynamic Phenomena and Solar Wind Consequences, the Third SOHO Workshop', Vol. 373 of *ESA Special Publication*, p. 123.

Low, B. C. (1996), 'Solar Activity and the Corona', *solphys* **167**(1-2), 217–265.

Low, B. & Lou, Y. (1990), 'Modeling solar force-free magnetic fields', *The Astrophysical Journal* **352**, 343–352.

Lutkepohl, H. (2005), Var order selection and checking the model adequacy, *in* 'New Introduction to Multiple Time Series Analysis', Springer Books, Springer, chapter 4, pp. 135–192.

Lyot, B. (1932), 'Étude de la couronne solaire en dehors des éclipses. Avec 16 figures dans le texte.', *zap* **5**, 73.

MacKinnon, J. G. (2010), Critical values for cointegration tests, Working Paper 1227, Economics Department, Queen's University.
URL: *https://ideas.repec.org/p/qed/wpaper/1227.html*

MacQueen, R. M. & Fisher, R. R. (1983), 'The Kinematics of Solar Inner Coronal Transients', *solphys* **89**(1), 89–102.

Madjarska, M. S. (2019), 'Coronal bright points', *Living Reviews in Solar Physics* **16**(1), 2.

Magara, T., Mineshige, S., Yokoyama, T. & Shibata, K. (1996), 'Numerical Simulation of Magnetic Reconnection in Eruptive Flares', *apj* **466**, 1054.

Majumdar, S., Pant, V., Patel, R. & Banerjee, D. (2020), 'Connecting 3d evolution of coronal mass ejections to their source regions', *The Astrophysical Journal* **899**(1), 6.
URL: *https://doi.org/10.3847/1538-4357/aba1f2*

Maričić, D., Vršnak, B., Veronig, A. M., Dumbović, M., Šterc, F., Roša, D., Karlica, M., Hržina, D. & Romštajn, I. (2020), 'Sun-to-earth observations and characteristics of isolated earth-impacting interplanetary coronal mass ejections during 2008 – 2014', *Solar Physics* **295**(7), 91.
URL: *https://doi.org/10.1007/s11207-020-01658-4*

Marubashi, K. (1997), *Interplanetary Magnetic Flux Ropes and Solar Filaments*, American Geophysical Union (AGU), pp. 147–156.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GM099p0147*

Mason, J. & Hoeksema, T. (2010), 'Testing automated solar flare forecasting with 13 years of michelson doppler imager magnetograms', *The Astrophysical Journal* **723**, 634.

McAteer, R. T. J., Gallagher, P. T. & Ireland, J. (2005), 'Statistics of active region complexity: A large-scale fractal dimension survey', *The Astrophysical Journal* **631**(1), 628.
URL: *https://doi.org/10.1086/432412*

McCloskey, Gallagher, Peter T. & Bloomfield, D. Shaun (2018), 'Flare forecasting using the evolution of mcintosh sunspot classifications', *J. Space Weather Space Clim.* **8**, A34.
URL: *https://doi.org/10.1051/swsc/2018022*

McIntosh, P. S. (1990), 'The Classification of Sunspot Groups', *solphys* **125**(2), 251–267.

Metcalf, T. R., Jiao, L., McClymont, A. N., Canfield, R. C. & Uitenbroek, H. (1995), 'Is the solar chromospheric magnetic field force-free?', *The Astrophysical Journal* **439**, 474–481.

Mikic, Z. & Linker, J. A. (1994), 'Disruption of Coronal Magnetic Field Arcades', *apj* **430**, 898.

Miller, J. A., Cargill, P. J., Emslie, A. G., Holman, G. D., Dennis, B. R., LaRosa, T. N., Winglee, R. M., Benka, S. G. & Tsuneta, S. (1997), 'Critical issues for understanding particle acceleration in impulsive solar flares', *Journal of Geophysical Research: Space Physics* **102**(A7), 14631–14659.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JA00976*

Milligan, R. O. (2011), 'Spatially Resolved Nonthermal Line Broadening during the Impulsive Phase of a Solar Flare', *apj* **740**(2), 70.

Montmerle, T., Augereau, J.-C., Chaussidon, M., Gounelle, M., Marty, B. & Morbidelli, A. (2006), '3. solar system formation and early evolution: the first 100 million years', *Earth, Moon, and Planets* **98**(1), 39–95.
URL: *https://doi.org/10.1007/s11038-006-9087-5*

Moon, Y.-J., Chae, J., Wang, H. & Park, Y. (2003), 'Magnetic helicity change rate associated with three x-class eruptive flares', *Advances in Space Research* **32**(10), 1953–1958.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117703906326*

Mou, C., Madjarska, M. S., Galsgaard, K. & Xia, L. (2018), 'Eruptions from quiet Sun coronal bright points. I. Observations', *aap* **619**, A55.

Mugatwala, R., Chierichini, S., Francisco, G., Napoletano, G., Foldes, R., Giovannelli, L., De Gasperis, G., Camporeale, E., Erdélyi, R. & Del Moro, D. (2024), 'A catalogue of observed geo-effective CME/ICME characteristics', *Journal of Space Weather and Space Climate* **14**, 6.

Nakariakov, V. M., Zhong, S., Kolotkov, D. Y., Meadowcroft, R. L., Zhong, Y. & Yuan, D. (2024), 'Diagnostics of the solar coronal plasmas by magnetohydrodynamic waves: magnetohydrodynamic seismology', *Reviews of Modern Plasma Physics* **8**(1), 19.
URL: *https://doi.org/10.1007/s41614-024-00160-9*

Napoletano, Gianluca, Forte, Roberta, Moro, Dario Del, Pietropaolo, Ermanno, Giovannelli, Luca & Berrilli, Francesco (2018), 'A probabilistic approach to the drag-based model', *J. Space Weather Space Clim.* **8**, A11.
URL: *https://doi.org/10.1051/swsc/2018003*

Nelson, C. J., Shelyag, S., Mathioudakis, M., Doyle, J. G., Madjarska, M. S., Uitenbroek, H. & Erdélyi, R. (2013), 'Ellerman bombs—evidence for magnetic reconnec-

tion in the lower solar atmosphere', *The Astrophysical Journal* **779**(2), 125.
URL: *https://doi.org/10.1088/0004-637X/779/2/125*

Neugebauer, M. & Snyder, C. W. (1966), 'Mariner 2 observations of the solar wind: 1. average properties', *Journal of Geophysical Research (1896-1977)* **71**(19), 4469–4484.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ071i019p04469*

Ngwira, C. M. & Pulkkinen, A. A. (2019), *An Introduction to Geomagnetically Induced Currents*, American Geophysical Union (AGU), chapter 1, pp. 1–13.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/9781119434412.ch1*

Nielsen, A. V. (1964), 'Contributions to the history of the hertzsprung-russell diagram', *Centaurus* **9**(4), 219–253.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0498.1964.tb00285.x*

Nindos, A. (2008), 'Magnetic helicity of solar active regions', *Proceedings of the International Astronomical Union* **4**(S257), 133–143.

Nishizuka, N., Sugiura, K., Kubo, Y., Den, M., Watari, S. & Ishii, M. (2017), 'Solar flare prediction model with three machine-learning algorithms using ultraviolet brightening and vector magnetograms', *The Astrophysical Journal* **835**(2), 156.
URL: *https://doi.org/10.3847/1538-4357/835/2/156*

Okamoto, T. J. & Sakurai, T. (2018), 'Super-strong magnetic field in sunspots', *The Astrophysical Journal Letters* **852**(1), L16.
URL: *https://doi.org/10.3847/2041-8213/aaa3d8*

Oughton, E. J., Skelton, A., Horne, R. B., Thomson, A. W. P. & Gaunt, C. T. (2017), 'Quantifying the daily economic impact of extreme space weather due to failure in electricity transmission infrastructure', *Space Weather* **15**(1), 65–83.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016SW001491*

Owens, M. J., Lockwood, M., Hawkins, E., Usoskin, I., Jones, G. S., Barnard, L., Schurer, A. & Fasullo, J. (2017), 'The maunder minimum and the little ice age: an update from recent reconstructions and climate simulations', *J. Space Weather Space Clim.* **7**.

Pal, S., Dash, S. & Nandy, D. (2020), 'Flux Erosion of Magnetic Clouds by Reconnection With the Sun's Open Flux', *grl* **47**(8), e86372.

Pandey, K., Chakrabarty, D., Kumar, A., Bhardwaj, A., Biswal, S., Hussey, G. & Yadav, A. (2023), 'Characteristics of x-class flares of solar cycles 23 and 24 in x-ray and euv bands', *Advances in Space Research* **71**(12), 5438–5452.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117723001412*

Panos, B., Kleint, L., Huwyler, C., Krucker, S., Melchior, M., Ullmann, D. & Voloshynovskiy, S. (2018), 'Identifying typical mg ii flare spectra using machine learning', *The Astrophysical Journal* **861**(1), 62.
URL: *https://doi.org/10.3847/1538-4357/aac779*

Panos, B., Kleint, L. & Voloshynovskiy, S. (2021), 'Exploring mutual information between iris spectral lines. i. correlations between spectral lines during solar flares and within the quiet sun.', *The Astrophysical Journal* **912**(2), 121.
URL: *https://doi.org/10.3847/1538-4357/abf11b*

Pant, V., Majumdar, S., Patel, R., Chauhan, A., Banerjee, D. & Gopalswamy, N. (2021), 'Investigating width distribution of slow and fast cmes in solar cycles 23 and 24', *Frontiers in Astronomy and Space Sciences* **Volume 8 - 2021**.
URL: *https://www.frontiersin.org/journals/astronomy-and-space-sciences/articles/10.3389/fspas.2021.634358*

Parenti, S. (2014), 'Solar prominences: Observations', *Living Reviews in Solar Physics* **11**(1), 1.
URL: *https://doi.org/10.12942/lrsp-2014-1*

Pariat, Leake, J. E., Valori, G., Linton, M. G., Zuccarello, F. P. & Dalmasse, K. (2017), 'Relative magnetic helicity as a diagnostic of solar eruptivity', *A&A* **601**, A125.
URL: *https://doi.org/10.1051/0004-6361/201630043*

Park, S.-H., Guerra, J. A., Gallagher, P. T., Georgoulis, M. K. & Bloomfield, D. S. (2018), 'Photospheric shear flows in solar active regions and their relation to flare occurrence', *Solar Physics* **293**(8), 114.
URL: *https://doi.org/10.1007/s11207-018-1336-z*

Park, S.-H., Leka, K. D., Kusano, K., Andries, J., Barnes, G., Bingham, S., Bloomfield, D. S., McCloskey, A. E., Delouille, V., Falconer, D., Gallagher, P. T., Georgoulis, M. K., Kubo, Y., Lee, K., Lee, S., Lobzin, V., Mun, J., Murray, S. A., Hamad Nageem, T. A. M., Qahwaji, R., Sharpe, M., Steenburgh, R. A., Steward, G. & Terkildsen, M. (2020), 'A comparison of flare forecasting methods. iv. evaluating consecutive-day forecasting patterns', *The Astrophysical Journal* **890**(2), 124.
URL: *https://doi.org/10.3847/1538-4357/ab65f0*

Parker, E. N. (1957), 'Sweet's mechanism for merging magnetic fields in conducting fluids', *Journal of Geophysical Research (1896-1977)* **62**(4), 509–520.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ062i004p00509*

Parker, E. N. (1988), 'Nanoflares and the Solar X-Ray Corona', *apj* **330**, 474.

Patsourakos, S. & Klimchuk, J. A. (2005), 'Coronal loop heating by nanoflares: The impact of the field-aligned distribution of the heating on loop observations', *The Astrophysical Journal* **628**(2), 1023.
URL: *https://doi.org/10.1086/430662*

Patsourakos, S., Vourlidas, A. & Stenborg, G. (2013), 'Direct evidence for a fast coronal mass ejection driven by the prior formation and subsequent destabilization of a magnetic flux rope', *The Astrophysical Journal* **764**(2), 125.
URL: *https://doi.org/10.1088/0004-637X/764/2/125*

Patsourakos, S., Vourlidas, A., Török, T., Kliem, B., Antiochos, S. K., Archontis, V., Aulanier, G., Cheng, X., Chintzoglou, G., Georgoulis, M. K., Green, L. M., Leake, J. E., Moore, R., Nindos, A., Syntelis, P., Yardley, S. L., Yurchyshyn, V. & Zhang, J. (2020), 'Decoding the Pre-Eruptive Magnetic Field Configurations of Coronal Mass Ejections', *ssr* **216**(8), 131.

Peter, H. & Judge, P. G. (1999), 'On the doppler shifts of solar ultraviolet emission lines', *The Astrophysical Journal* **522**(2), 1148.
URL: *https://doi.org/10.1086/307672*

Petschek, H. E. (1964), Magnetic Field Annihilation, *in* W. N. Hess, ed., 'NASA Special Publication', Vol. 50, p. 425.

Pontin, D. I. & Priest, E. R. (2022), 'Magnetic reconnection: Mhd theory and modelling', *Living Reviews in Solar Physics* **19**(1), 1.
URL: *https://doi.org/10.1007/s41116-022-00032-9*

Priest, E. (2014), *Magnetohydrodynamics of the Sun*, Cambridge University Press.

Priest, E. R. & Démoulin, P. (1995), 'Three-dimensional magnetic reconnection without null points: 1. basic theory of magnetic flipping', *Journal of Geophysical Research: Space Physics* **100**(A12), 23443–23463.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95JA02740*

Pringle, J. E. (1981), 'Accretion discs in astrophysics', *Annual Review of Astronomy and Astrophysics* **19**(Volume 19, 1981), 137–160.
URL: *https://www.annualreviews.org/content/journals/10.1146/annurev.aa.19.090181.001033*

Pudritz, R. E. (2002), 'Clustered star formation and the origin of stellar masses', *Science* **295**(5552), 68–76.
URL: *https://www.science.org/doi/abs/10.1126/science.1068298*

Raouafi, N. E., Patsourakos, S., Pariat, E., Young, P. R., Sterling, A. C., Savcheva, A., Shimojo, M., Moreno-Insertis, F., DeVore, C. R., Archontis, V., Török, T., Mason,

H., Curdt, W., Meyer, K., Dalmasse, K. & Matsui, Y. (2016), 'Solar coronal jets: Observations, theory, and modeling', *Space Science Reviews* **201**(1), 1–53.
URL: *https://doi.org/10.1007/s11214-016-0260-5*

Reale, F. (2010), 'Coronal loops: Observations and modeling of confined plasma', *Living Reviews in Solar Physics* **7**(1), 5.
URL: *https://doi.org/10.12942/lrsp-2010-5*

Reames, D. V. (1999), 'Particle acceleration at the Sun and in the heliosphere', *ssr* **90**, 413–491.

Reep, J. W. & Knizhnik, K. J. (2019), 'What determines the x-ray intensity and duration of a solar flare?', *The Astrophysical Journal* **874**(2), 157.

Reinard, A. A. & Biesecker, D. A. (2008), 'Coronal Mass Ejection-Associated Coronal Dimmings', *apj* **674**(1), 576–585.

Reinard, A. A., Henthorn, J., Komm, R. & Hill, F. (2010), 'Evidence that temporal changes in solar subsurface helicity precede active region flaring', *The Astrophysical Journal Letters* **710**(2), L121.
URL: *https://doi.org/10.1088/2041-8205/710/2/L121*

Reiners, A. (2012), 'Observations of cool-star magnetic fields', *Living Reviews in Solar Physics* **9**(1), 1.
URL: *https://doi.org/10.12942/lrsp-2012-1*

Rezaei, R., Beck, C. & Schmidt, W. (2012), 'Variation in sunspot properties between 1999 and 2011 as observed with the Tenerife Infrared Polarimeter', *aap* **541**, A60.

Rosenthal, C., Bogdan, T., Carlsson, M., Dorch, S., Hansteen, V., McIntosh, S., McMurry, A., Nordlund, Å. & Stein, R. (2002), 'Waves in the magnetized solar atmosphere. i. basic processes and internetwork oscillations', *The Astrophysical Journal* **564**(1), 508.

Rouppe van der Voort, L. H. M., De Pontieu, B., Hansteen, V. H., Carlsson, M. & van Noort, M. (2007), 'Magnetoacoustic shocks as a driver of quiet-sun mottles', *The Astrophysical Journal* **660**(2), L169.
URL: *https://doi.org/10.1086/518246*

Roussev, I. I., Forbes, T. G., Gombosi, T. I., Sokolov, I. V., DeZeeuw, D. L. & Birn, J. (2003), 'A three-dimensional flux rope model for coronal mass ejections based on a loss of equilibrium', *The Astrophysical Journal* **588**(1), L45.
URL: *https://doi.org/10.1086/375442*

Roussev, I. I., Sokolov, I. V., Forbes, T. G., Gombosi, T. I., Lee, M. A. & Sakai, J. I. (2004), 'A numerical model of a coronal mass ejection: Shock development with implications for the acceleration of gev protons', *The Astrophysical Journal* **605**(1), L73.
URL: *https://doi.org/10.1086/392504*

Ruiz Cobo, B. & del Toro Iniesta, J. C. (1992), 'Inversion of Stokes Profiles', *apj* **398**, 375.

Ruiz Cobo, B. & del Toro Iniesta, J. C. (1994), 'On the sensitivity of Stokes profiles to physical quantities.', *aap* **283**, 129–143.

Russell, H. N. (1919), 'Some problems of sidereal astronomy*', *Proceedings of the National Academy of Sciences* **5**(10), 391–416.
URL: *https://www.pnas.org/doi/abs/10.1073/pnas.5.10.391*

Rutten, Rouppe van der Voort, Luc H. M. & De Pontieu, Bart (2019), 'Solar htures with hot onsets - iv. network fibrils', *A&A* **632**, A96.
URL: *https://doi.org/10.1051/0004-6361/201936113*

Régnier, Walsh, R. W. & Alexander, C. E. (2011), 'A new look at a polar crown cavity as observed by sdo/aia - structure and dynamics', *A&A* **533**, L1.
URL: *https://doi.org/10.1051/0004-6361/201117381*

Sadykov, V. M. & Kosovichev, A. G. (2017), 'Characteristics of the polarity inversion line and solar flare forecasts', *AAS/Solar Physics Division Abstracts# 48* **48**, 116–01.

Said, S. E. & Dickey, D. A. (1984), 'Testing for unit roots in autoregressive-moving average models of unknown order', *Biometrika* **71**(3), 599–607.
URL: *https://doi.org/10.1093/biomet/71.3.599*

Samara, E., Patsourakos, S. & Georgoulis, M. K. (2021), 'A Readily Implemented Atmosphere Sustainability Constraint for Terrestrial Exoplanets Orbiting Magnetically Active Stars', *apjl* **909**(1), L12.

Sammis, I., Tang, F. & Zirin, H. (2000), 'The dependence of large flare occurrence on the magnetic structure of sunspots', *The Astrophysical Journal* **540**(1), 583.
URL: *https://dx.doi.org/10.1086/309303*

Scherrer, P. H., Bogart, R. S., Bush, R., Hoeksema, J.-a., Kosovichev, A., Schou, J., Rosenberg, W., Springer, L., Tarbell, T., Wolfson, C. et al. (1995), The solar oscillations investigation—michelson doppler imager, *in* 'The SOHO Mission', Springer, pp. 129–188.

Scherrer, P. H., Schou, J., Bush, R., Kosovichev, A., Bogart, R., Hoeksema, J., Liu, Y., Duvall, T., Zhao, J., Schrijver, C. et al. (2012), 'The helioseismic and magnetic imager (hmi) investigation for the solar dynamics observatory (sdo)', *Solar Physics* **275**(1), 207–227.

Schmidt, H. U. (1964), On the Observable Effects of Magnetic Energy Storage and Release Connected With Solar Flares, *in* W. N. Hess, ed., 'NASA Special Publication', Vol. 50, p. 107.

Schmieder, B., Aulanier, G., Démoulin, P. & Pariat, E. (2009), 'Coronal loops, flare ribbons and aurora during slip-running', *Earth, Planets and Space* **61**(5), 565–568.
URL: *https://doi.org/10.1186/BF03352923*

Schrijver, C. J. (2007), 'A characteristic magnetic field pattern associated with all major solar flares and its use in flare forecasting', *The Astrophysical Journal* **655**(2), L117.

Schrijver, C. J., Title, A. M., Berger, T. E., Fletcher, L., Hurlburt, N. E., Nightingale, R. W., Shine, R. A., Tarbell, T. D., Wolfson, J., Golub, L., Bookbinder, J. A., DeLuca, E. E., McMullen, R. A., Warren, H. P., Kankelborg, C. C., Handy, B. N. & De Pontieu, B. (1999), 'A new view of the solar outer atmosphere by the Transition Region and Coronal Explorer', *solphys* **187**(2), 261–302.

Schwabe, S. H. (1843), 'Die sonne', *Astronomische Nachrichten* **20**(17), 283–286.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/asna.18430201706*

Schwarz, G. (1978), 'Estimating the dimension of a model', *Annals of Statistics* **6**, 461–464.
URL: *https://api.semanticscholar.org/CorpusID:123722079*

Seehafer, N. (1978), 'Determination of constant α force-free solar magnetic fields from magnetograph data', *Solar Physics* **58**(2), 215–223.
URL: *https://doi.org/10.1007/BF00157267*

Severny, A. (1964), 'Solar Magnetic Fields', *ssr* **3**(4), 451–486.

Sheeley Jr., N. R., Walters, J. H., Wang, Y.-M. & Howard, R. A. (1999), 'Continuous tracking of coronal outflows: Two kinds of coronal mass ejections', *Journal of Geophysical Research: Space Physics* **104**(A11), 24739–24767.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999JA900308*

Shibata, K. & Magara, T. (2011), 'Solar flares: Magnetohydrodynamic processes', *Living Reviews in Solar Physics* **8**(1), 6.
URL: *https://doi.org/10.12942/lrsp-2011-6*

Sinha, S., Gupta, O., Singh, V., Lekshmi, B., Nandy, D., Mitra, D., Chatterjee, S., Bhattacharya, S., Chatterjee, S., Srivastava, N., Brandenburg, A. & Pal, S. (2022),

'A comparative analysis of machine-learning models for solar flare forecasting: Identifying high-performing active region flare indicators', *The Astrophysical Journal* **935**(1), 45.
URL: *https://doi.org/10.3847/1538-4357/ac7955*

Solanki, S. K. (1993), 'Smallscale Solar Magnetic Fields - an Overview', *ssr* **63**(1-2), 1–188.

Solanki, S. K. (2003), 'Sunspots: An overview', *The Astronomy and Astrophysics Review* **11**(2), 153–286.
URL: *https://doi.org/10.1007/s00159-003-0018-4*

Song, H., Li, L. & Chen, Y. (2022), 'Toward a unified explanation for the three-part structure of solar coronal mass ejections', *The Astrophysical Journal* **933**(1), 68.
URL: *https://doi.org/10.3847/1538-4357/ac7239*

Soós, S., Korsós, M., Morgan, H. & Erdélyi, R. (2022), 'On the differences in the periodic behavior of magnetic helicity flux in flaring active regions with and without x-class events', *The Astrophysical Journal* **925**(2), 129.

St. Cyr, O. C., Howard, R. A., Sheeley Jr., N. R., Plunkett, S. P., Michels, D. J., Paswaters, S. E., Koomen, M. J., Simnett, G. M., Thompson, B. J., Gurman, J. B., Schwenn, R., Webb, D. F., Hildner, E. & Lamy, P. L. (2000), 'Properties of coronal mass ejections: Soho lasco observations from january 1996 to june 1998', *Journal of Geophysical Research: Space Physics* **105**(A8), 18169–18185.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999JA000381*

Stamkos, S., Patsourakos, S., Vourlidas, A. & Daglis, I. A. (2023), 'How Magnetic Erosion Affects the Drag-Based Kinematics of Fast Coronal Mass Ejections', *solphys* **298**(7), 88.

Stenflo, J. O. (2013), 'Solar magnetic fields as revealed by stokes polarimetry', *The*

Astronomy and Astrophysics Review **21**(1), 66.
URL: *https://doi.org/10.1007/s00159-013-0066-3*

Sterling, A. C. (2000), 'Solar Spicules: A Review of Recent Models and Targets for Future Observations - (Invited Review)', *solphys* **196**(1), 79–111.

Sturrock, P. A. (1966), 'Model of the High-Energy Phase of Solar Flares', *nat* **211**(5050), 695–697.

Sturrock, P. A. & Coppi, B. (1966), 'A New Model of Solar Flares', *apj* **143**, 3.

Subramanian, P. & Dere, K. P. (2001), 'Source regions of coronal mass ejections', *The Astrophysical Journal* **561**(1), 372.
URL: *https://doi.org/10.1086/323213*

Subramanian & Vourlidas (2007), 'Energetics of solar coronal mass ejections', *A&A* **467**(2), 685–693.
URL: *https://doi.org/10.1051/0004-6361:20066770*

Suematsu, Y., Wang, H. & Zirin, H. (1995), 'High-Resolution Observation of Disk Spicules. I. Evolution and Kinematics of Spicules in the Enhanced Network', *apj* **450**, 411.

Sun, H., Manchester IV, W. & Chen, Y. (2021), 'Improved and interpretable solar flare predictions with spatial and topological features of the polarity inversion line masked magnetograms', *Space Weather* **19**(12), e2021SW002837. e2021SW002837 2021SW002837.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021SW002837*

Sun, X., Hoeksema, J. T., Liu, Y., Wiegelmann, T., Hayashi, K., Chen, Q. & Thalmann, J. (2012), 'Evolution of Magnetic Field and Energy in a Major Eruptive Active Region Based on SDO/HMI Observation', *apj* **748**(2), 77.

Švestka, Z. (1989), 'Solar flares: The gradual phase', *Solar Physics* **121**(1), 399–417.
URL: *https://doi.org/10.1007/BF00161709*

Sweet, P. A. (1958), The Neutral Point Theory of Solar Flares, *in* B. Lehnert, ed., 'Electromagnetic Phenomena in Cosmical Physics', Vol. 6 of *IAU Symposium*, p. 123.

Taroyan, Y. & Erdélyi, R. (2009), 'Heating Diagnostics with MHD Waves', *ssr* **149**(1-4), 229–254.

Temmer, M. (2021), 'Space weather: the solar perspective', *Living Reviews in Solar Physics* **18**(1), 4.
URL: *https://doi.org/10.1007/s41116-021-00030-3*

Temmer, M., Veronig, A. M., Kontar, E. P., Krucker, S. & Vršnak, B. (2010), 'Combined STEREO/RHESSI Study of Coronal Mass Ejection Acceleration and Particle Acceleration in Solar Flares', *apj* **712**(2), 1410–1420.

Titov, V. & Démoulin, P. (1999), 'Basic topology of twisted magnetic configurations in solar flares', *Astronomy and Astrophysics* **351**, 707–720.

Tiwari, S. K., Venkatakrishnan, P. & Gosain, S. (2010), 'Magnetic non-potentiality of solar active regions and peak x-ray flux of the associated flares', *The Astrophysical Journal* **721**(1), 622.
URL: *https://doi.org/10.1088/0004-637X/721/1/622*

To, A. S. H., Burden, A., Baker, D., Eklund, H., Brooks, D. H., Hayes, L. A., Martínez-Sykora, J., Testa, P., Reep, J., Janvier, M., Imada, S., Hernandez Camero, J., Long, D. M., Mihailescu, T. & Weberg, M. J. (2025), 'Systematic Nonthermal Velocity Increase Preceding Soft X-Ray Flare Onset: A Large-scale Hinode/EIS Study', *apj* **993**(1), 102.

Tokumaru, M. & Fujiki, K. (2024), 'Coronal magnetic-field configuration associated with pseudostreamer and slow solar wind', *Solar Physics* **299**(11), 160.
URL: *https://doi.org/10.1007/s11207-024-02398-5*

Toriumi, S. (2022), 'Flux emergence and generation of flare-productive active regions', *Advances in Space Research* **70**(6), 1549–1561. Magnetic Flux Ropes in Solar Environments.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117721004373*

Toriumi, S. & Wang, H. (2019), 'Flare-productive active regions', *Living Reviews in Solar Physics* **16**(1), 3.

Török, T. & Kliem, B. (2003), 'The evolution of twisting coronal magnetic flux tubes', *aap* **406**, 1043–1059.

Tousey, R. (1973), The solar corona., *in* M. J. Rycroft & S. K. Runcorn, eds, 'Space Research Conference', Vol. 2, pp. 713–730.

Tsurutani, B. T., Verkhoglyadova, O. P., Mannucci, A. J., Lakhina, G. S., Li, G. & Zank, G. P. (2009), 'A brief review of "solar flare effects" on the ionosphere', *Radio Science* **44**(1).
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008RS004029*

Török, Kliem, B. & Titov, V. S. (2004), 'Ideal kink instability of a magnetic loop equilibrium', *A&A* **413**(3), L27–L30.
URL: *https://doi.org/10.1051/0004-6361:20031691*

Török, T. & Kliem, B. (2005), 'Confined and ejective eruptions of kink-unstable flux ropes', *The Astrophysical Journal* **630**(1), L97.
URL: *https://doi.org/10.1086/462412*

Török, T. & Kliem, B. (2007), 'Numerical simulations of fast and slow coronal mass ejections', *Astronomische Nachrichten* **328**(8), 743–746.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/asna.200710795*

van Ballegooijen, A. A., Asgari-Targhi, M., Cranmer, S. R. & DeLuca, E. E. (2011), 'Heating of the Solar Chromosphere and Corona by Alfven Wave Turbulence', *apj* **736**(1), 3.

van Ballegooijen, A. A. & Martens, P. C. H. (1989), 'Formation and Eruption of Solar Prominences', *apj* **343**, 971.

van Driel-Gesztelyi, L. & Green, L. M. (2015), 'Evolution of active regions', *Living Reviews in Solar Physics* **12**(1), 1.
URL: *https://doi.org/10.1007/lrsp-2015-1*

Vashalomidze, Z., Kukhianidze, V., Zaqarashvili, T. V., Oliver, R., Shergelashvili, B., Ramishvili, G., Poedts, S. & De Causmaecker, P. (2015), 'Formation and evolution of coronal rain observed by SDO/AIA on February 22, 2012', *aap* **577**, A136.

Vemareddy, P. (2021), 'Magnetic structure in successively erupting active regions: Comparison of flare-ribbons with quasi-separatrix layers', *Frontiers in Physics* **Volume 9 - 2021**.
URL: *https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2021.749479*

Vourlidas, A., Howard, R. A., Esfandiari, E., Patsourakos, S., Yashiro, S. & Michalek, G. (2010), 'Comprehensive analysis of coronal mass ejection mass and energy properties over a full solar cycle', *The Astrophysical Journal* **722**(2), 1522.
URL: *https://doi.org/10.1088/0004-637X/722/2/1522*

Vourlidas, A., Lynch, B. J., Howard, R. A. & Li, Y. (2013), 'How many cmes have flux ropes? deciphering the signatures of shocks, flux ropes, and prominences in coronagraph observations of cmes', *Solar Physics* **284**(1), 179–201.
URL: *https://doi.org/10.1007/s11207-012-0084-8*

Vourlidas, A., Patsourakos, S. & Savani, N. P. (2019), 'Predicting the geoeffective properties of coronal mass ejections: current status, open issues and path forward', *Philosophical Transactions of the Royal Society of London Series A* **377**(2148), 20180096.

Vršnak, B., Žic, T., Vrbanec, D., Temmer, M., Rollett, T., Möstl, C., Veronig, A., Čalogović, J., Dumbović, M., Lulić, S., Moon, Y.-J. & Shanmugaraju, A. (2013), 'Propagation of interplanetary coronal mass ejections: The drag-based model', *Solar*

Physics **285**(1), 295–315.
URL: *https://doi.org/10.1007/s11207-012-0035-4*

Vršnak & Žic (2007), 'Transit times of interplanetary coronal mass ejections and the solar wind speed', *A&A* **472**(3), 937–943.
URL: *https://doi.org/10.1051/0004-6361:20077499*

Vögler, Shelyag, S., Schüssler, M., Cattaneo, F., Emonet, T. & Linde, T. (2005), 'Simulations of magneto-convection in the solar photosphere* - equations, methods, and results of the muram code', *A&A* **429**(1), 335–351.
URL: *https://doi.org/10.1051/0004-6361:20041507*

Wang, H., Ewell, Jr., M. W., Zirin, H. & Ai, G. (1994), 'Vector Magnetic Field Changes Associated with X-Class Flares', *apj* **424**, 436.

Wang, H.-M., Song, H., Jing, J., Yurchyshyn, V., Deng, Y.-Y., Zhang, H.-Q., Falconer, D. & Li, J. (2006), 'The relationship between magnetic gradient and magnetic shear in five super active regions producing great flares', *Chinese Journal of Astronomy and Astrophysics* **6**(4), 477.
URL: *https://doi.org/10.1088/1009-9271/6/4/11*

Wang, J., Liu, S., Ao, X., Zhang, Y., Wang, T. & Liu, Y. (2019), 'Parameters derived from the sdo/hmi vector magnetic field data: Potential to improve machine-learning-based solar flare prediction models', *The Astrophysical Journal* **884**(2), 175.
URL: *https://dx.doi.org/10.3847/1538-4357/ab441b*

Warmuth, A. (2015), 'Large-scale globally propagating coronal waves', *Living Reviews in Solar Physics* **12**(1), 3.
URL: *https://doi.org/10.1007/lrsp-2015-3*

Webb, D. F. & Howard, T. A. (2012), 'Coronal mass ejections: Observations', *Living Reviews in Solar Physics* **9**(1), 3.
URL: *https://doi.org/10.12942/lrsp-2012-3*

Webb, D. F. & Hundhausen, A. J. (1987), 'Activity associated with the solar origin of coronal mass ejections', *Solar Physics* **108**(2), 383–401.
URL: *https://doi.org/10.1007/BF00214170*

Webb, D. F. & Zirin, H. (1981), 'Coronal loops and active region structure', *Solar Physics* **69**(1).
URL: *https://doi.org/10.1007/bf00151258*

Whitman, K., Egeland, R., Richardson, I. G., Allison, C., Quinn, P., Barzilla, J., Kitiashvili, I., Sadykov, V., Bain, H. M., Dierckxsens, M., Mays, M. L., Tadesse, T., Lee, K. T., Semones, E., Luhmann, J. G., Núñez, M., White, S. M., Kahler, S. W., Ling, A. G., Smart, D. F., Shea, M. A., Tenishev, V., Boubrahimi, S. F., Aydin, B., Martens, P., Angryk, R., Marsh, M. S., Dalla, S., Crosby, N., Schwadron, N. A., Kozarev, K., Gorby, M., Young, M. A., Laurenza, M., Cliver, E. W., Alberti, T., Stumpo, M., Benella, S., Papaioannou, A., Anastasiadis, A., Sandberg, I., Georgoulis, M. K., Ji, A., Kempton, D., Pandey, C., Li, G., Hu, J., Zank, G. P., Lavasa, E., Giannopoulos, G., Falconer, D., Kadadi, Y., Fernandes, I., Dayeh, M. A., Muñoz-Jaramillo, A., Chatterjee, S., Moreland, K. D., Sokolov, I. V., Roussev, I. I., Taktakishvili, A., Effenberger, F., Gombosi, T., Huang, Z., Zhao, L., Wijsen, N., Aran, A., Poedts, S., Kouloumvakos, A., Paassilta, M., Vainio, R., Belov, A., Eroshenko, E. A., Abunina, M. A., Abunin, A. A., Balch, C. C., Malandraki, O., Karavolos, M., Heber, B., Labrenz, J., Kühl, P., Kosovichev, A. G., Oria, V., Nita, G. M., Illarionov, E., O'Keefe, P. M., Jiang, Y., Fereira, S. H., Ali, A., Paouris, E., Aminalragia-Giamini, S., Jiggens, P., Jin, M., Lee, C. O., Palmerio, E., Bruno, A., Kasapis, S., Wang, X., Chen, Y., Sanahuja, B., Lario, D., Jacobs, C., Strauss, D. T., Steyn, R., van den Berg, J., Swalwell, B., Waterfall, C., Nedal, M., Miteva, R., Dechev, M., Zucca, P., Engell, A., Maze, B., Farmer, H., Kerber, T., Barnett, B., Loomis, J., Grey, N., Thompson, B. J., Linker, J. A., Caplan, R. M., Downs, C., Török, T., Lionello, R., Titov, V., Zhang, M. & Hosseinzadeh, P. (2023), 'Review of solar energetic particle prediction models', *Advances in Space Research* **72**(12), 5161–5242. COSPAR Space

Weather Roadmap 2022-2024: Scientific Research and Applications.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117722007244*

Wiegelmann, T. & Sakurai, T. (2021), 'Solar force-free magnetic fields', *Living Reviews in Solar Physics* **18**(1), 1.
URL: *https://doi.org/10.1007/s41116-020-00027-4*

Wiegelmann, T., Thalmann, J. K. & Solanki, S. K. (2014), 'The magnetic field in the solar atmosphere', *The Astronomy and Astrophysics Review* **22**(1), 1–106.

Wiśniewska, Korsós, M. B., Kontogiannis, I., Soós, Sz. & Erdélyi, R. (2024), 'Long-period oscillations in the lower solar atmosphere prior to flare events', *A&A* **686**, A224.
URL: *https://doi.org/10.1051/0004-6361/202348606*

Wolfson, R. & Dlamini, B. (1997), 'Cross-field currents: An energy source for coronal mass ejections?', *The Astrophysical Journal* **483**(2), 961.
URL: *https://doi.org/10.1086/304263*

Woods, T. N. (2014), 'Extreme ultraviolet late-phase flares: Before and during the solar dynamics observatory mission', *Solar Physics* **289**(9), 3391–3401.
URL: *https://doi.org/10.1007/s11207-014-0483-0*

Woods, T. N., Hock, R., Eparvier, F., Jones, A. R., Chamberlin, P. C., Klimchuk, J. A., Didkovsky, L., Judge, D., Mariska, J., Warren, H., Schrijver, C. J., Webb, D. F., Bailey, S. & Tobiska, W. K. (2011), 'New solar extreme-ultraviolet irradiance observations during flares', *The Astrophysical Journal* **739**(2), 59.
URL: *https://doi.org/10.1088/0004-637X/739/2/59*

Woolfson, M. M. (1993), 'The Solar - Origin and Evolution', *qjras* **34**, 1–20.

Yashiro, S., Gopalswamy, N., Akiyama, S., Michalek, G. & Howard, R. A. (2005), 'Visibility of coronal mass ejections as a function of flare location and intensity',

Journal of Geophysical Research: Space Physics **110**(A12).
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JA011151*

Yashiro, S., Gopalswamy, N., Michalek, G., St. Cyr, O. C., Plunkett, S. P., Rich, N. B. & Howard, R. A. (2004), 'A catalog of white light coronal mass ejections observed by the soho spacecraft', *Journal of Geophysical Research: Space Physics* **109**(A7).
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JA010282*

Yokoyama, T., Kusano, K., Maeshiro, T. & Sakurai, T. (2003), 'Relation between magnetic helicity injection and flare activities in active region noaa 8100', *Advances in Space Research* **32**(10), 1949–1952.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117703906314*

Yurchyshyn, V., Yashiro, S., Abramenko, V., Wang, H. & Gopalswamy, N. (2005), 'Statistical distributions of speeds of coronal mass ejections', *The Astrophysical Journal* **619**(1), 599.
URL: *https://doi.org/10.1086/426129*

Zaqarashvili, T. V. & Erdélyi, R. (2009), 'Oscillations and waves in solar spicules', *Space Science Reviews* **149**(1), 355–388.
URL: *https://doi.org/10.1007/s11214-009-9549-y*

Zaqarashvili, T. V. & Skhirtladze, N. (2008), 'Helical motion of magnetic flux tubes in the solar atmosphere', *The Astrophysical Journal* **683**(1), L91.
URL: *https://doi.org/10.1086/591524*

Zhang, H. (2002), 'Magnetic field, helicity and the 2000 july 14 flare in solar active region 9077', *Monthly Notices of the Royal Astronomical Society* **332**(2), 500–512.
URL: *https://doi.org/10.1046/j.1365-8711.2002.05345.x*

Zhang, J., Dere, K. P., Howard, R. A. & Vourlidas, A. (2004), 'A Study of the Kinematic Evolution of Coronal Mass Ejections', *apj* **604**(1), 420–432.

Zhang, J., Temmer, M., Gopalswamy, N., Malandraki, O., Nitta, N. V., Patsourakos, S., Shen, F., Vršnak, B., Wang, Y., Webb, D., Desai, M. I., Dissauer, K., Dresing, N., Dumbović, M., Feng, X., Heinemann, S. G., Laurenza, M., Lugaz, N. & Zhuang, B. (2021), 'Earth-affecting solar transients: a review of progresses in solar cycle 24', *Progress in Earth and Planetary Science* **8**(1), 56.

Zhang, J., Wang, J., Deng, Y. & Wu, D. (2001), 'Magnetic flux cancellation associated with the major solar event on 2000 july 14', *The Astrophysical Journal* **548**(1), L99.
URL: *https://dx.doi.org/10.1086/318934*

Zirin, H. & Marquette, W. (1991), 'BEARALERTS: A successful flare prediction system', *solphys* **131**(1), 149–164.

Zirin, H. & Tanaka, K. (1973), 'The flares of august 1972', *Solar Physics* **32**(1), 173–207.
URL: *https://doi.org/10.1007/BF00152736*

Zirker, J. B. (1977), 'Coronal holes and high-speed wind streams', *Reviews of Geophysics* **15**(3), 257–269.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG015i003p00257*

Zuccarello, Balmaceda, Laura, Cessateur, Gael, Cremades, Hebe, Guglielmino, Salvatore L., Lilensten, Jean, Dudok de Wit, Thierry, Kretzschmar, Matthieu, Lopez, Fernando M., Mierla, Marilena, Parenti, Susanna, Pomoell, Jens, Romano, Paolo, Rodriguez, Luciano, Srivastava, Nandita, Vainio, Rami, West, Matt & Zuccarello, Francesco P. (2013), 'Solar activity and its evolution across the corona: recent advances', *J. Space Weather Space Clim.* **3**, A18.
URL: *https://doi.org/10.1051/swsc/2013039*

Zweibel, E. G. & Yamada, M. (2009), 'Magnetic Reconnection in Astrophysical and Laboratory Plasmas', *araa* **47**(1), 291–332.

Žic, T., Vršnak, B. & Temmer, M. (2015), 'Heliospheric propagation of coronal mass ejections: Drag-based model fitting', *The Astrophysical Journal Supplement Series* **218**(2), 32.

URL: *https://doi.org/10.1088/0067-0049/218/2/32*

Appendices

Appendix A

Visualising solar magnetic fields in ParaView

Introduction

ParaView VTK (Visualisation Toolkit) is a software application intended to help researchers visualise fluid data (Ayachit 2015). *ParaView* is of use to solar physicists because it can be used to visually represent the streamlines of magnetic field, thereby providing insights into the solar atmosphere. This manual contains a step by step approach to generate visualisations of solar magnetic field using *ParaView 5.7.0*. We have shared two methods of plotting solar magnetic fields in this manual. In the first method, described in section A.1, we use manually devised *SSW-IDL* codes to perform PF extrapolation of vector magnetogram data up to 3.24 Mm, subsequently using the 3D data grid as input. In the second method, described in section A.2, the approach relies on directly importing a 3D-grid of PF/NLFFF data from an open-source ISEE Database directly as input (Kusano. et al. 2020, 2021). This extrapolated magnetic field data is based on the MHD relaxation described in Inoue et al. (2014). Dependencies necessary to replicate the steps/results shown in this manual may be accessed from the following GitHub repository: https://github.com/shreeyesh-biswal/Paraview_SMF.

A.1 Method 1: Manual extrapolation

A.1.1 Input Data Assembly

The first step is to download the Lambert cylindrical equal-area projection data corresponding to the photospheric vector magnetic field (B_r) of any desired AR from the Stanford-JSOC website: http://jsoc.stanford.edu/. The requisite data extension is: `hmi.sharp_cea_720s`. A sample data file appears as:

```
hmi.sharp_cea_720s.377.20110215_000000_TAI.Br.fits
```

The *IDL* program file `PF_Sharp_sav.pro` is then compiled along with a dependency file `poten_sheffield.pro` to produce the PF extrapolation (α is set to 0 in the source code) of the 2D-photospheric input data B_r. The result is a Datacube of 10 planes, each spaced 0.36 Mm apart with each plane containing (B_x, B_y, B_z) information. For every time-stamp, there shall be 30 output FITS files, 3 at each height. In addition to these 30 files, there shall be a `.sav` file containing the magnetic field information in the Datacube. The files should be easy to identify. For example, the file corresponding to B_y at 2.88 Mm height (plane no. 9) is the following:

```
hmi.sharp_cea_720s.377.20110215_000000_0.00000_8.00000_pby.fits
```

A.1.2 Input Data Visualisation

The FITS files generated in the previous step may be visualised using *Python-Astropy* but there exists a much quicker way to visualise it in *IDL* CLI (Command Line Interface). In order to do that, we pass the following commands in sequence in the same directory where the FITS file is located after opening the *SSW-IDL*:

```
>> f = 'hmi.sharp_cea_720s.377.20110215_000000_0.00000_8.00000_pby.fits'
>> afile = findfile(f)
>> read_sdo, afile, aindex, adata
>> plot_image, adata
```

The dimension of the 3D Datacube is 733 x 382 x 10. This means that the length scales along x, y and z axes are 732 x 381 x 9 units where 1 unit = 0.36 Mm. In terms of physical units, the figures are 263.52 Mm, 137.16 Mm and 3.24 Mm. If we view from the top (which we do in case of satellite based observations, we shall see an area of 263.52 x 137.16 = 36144.40 Mm^2. The radius of the Sun is $R_s = 696.34$ Mm. Hence the 'visible' (or projected) surface area is $\pi {R_s}^2 = 1523324.96$ Mm^2. Hence, the area of the Sun that we are plotting the magnetic field lines for is equal to 2.37% of the total projected area.

A.1.3 Generating ParaView Readable Files

The next step is to convert the data inside a `.sav` file into a `.vtk` file because *ParaView* does not support visualisation using `.sav` file. The *IDL* program file `sav2vtk.pro` is designed for this conversion and is executed at this stage.

A.1.4 ParaView Visualisation

1. Orientation

Now we open the `.vtk` file with *ParaView*. For proper orientation of the axes, adjust camera view (5th button below the Layout bar in *ParaView 5.7.0*) by loading the file `camera.pvcc`. This file normally saves the orientation layout, which is of interest to the user. *ParaView* also supports quick access to `XY, YZ, ZX` views (top/bottom) under `Standard Viewpoints`.

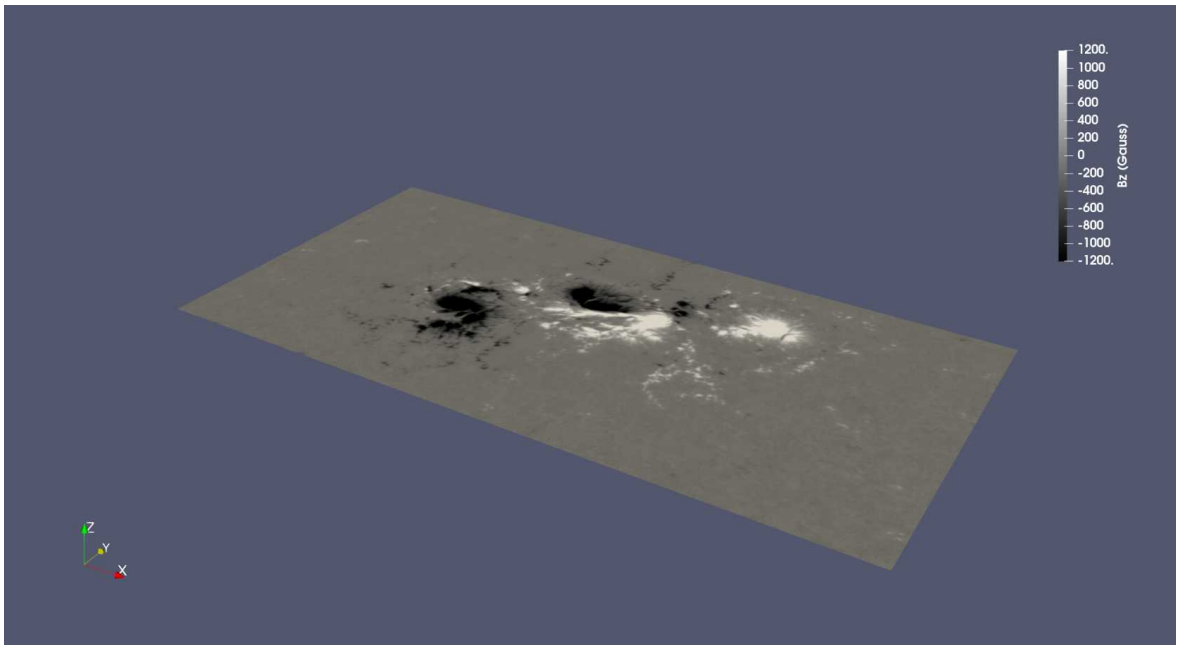

Figure A.1: A slice of the photospheric B_z for AR 11158 on 15 Feb 2011, 00:00 UTC

2. Surface Plot and Slice Filter

In order to visualise the B_z data in 3D, go to the `Display Bar` and change `solid colour` to `bfield`, change `magnitude` to `z`, and select `surface` instead of `outline`. The `surface` option coupled with the `Slice` filter allows the user to view the (B_x, B_y, B_z) field intensities layer by layer. To visualise the photospheric magnetic field in just like magnetogram data, apply the slice filter, set the `normal` to (0, 0, 1) and the `origin` as (0, 0, 0.01). The next step is to go into the `coloring` section for the `slice` and use the `Grayscale` preset. To enhance contrast in the visualisation, the user might want to click on `rescale to custom range` and set the limits between $\pm$ 1200 G. Disable automatic rescaling again. Click on `edit color legend properties` button, set the window location to `any location` and set the desired position. We recommend using (0.9, 0.6) as the location. We also suggest setting the title text as 'Bz (Gauss)', leaving the `component title` blank. Use bold font and set the `font size` to 24. Upon disabling the data axes grid and axes grid options, one may arrive at Figure A.1.

3. Importing Seed Points with Table to Points Filter

Before applying the `Stream Tracer`, it is possible to manually specify the location of the seed points that serve as end points or boundary conditions between streamlines. It is preferable to specify the seed points on the photosphere. To do so, open the CSV file containing seed points where $B_z > 20$ G: `data_selected_origin_centered_20G.csv`. This CSV file is created using the *Python* file `FITS_read_origin_centered_20G.py` starting with photospheric B_z magnetogram data. A similar CSV file but for 100 G can also be used if needed. Once the CSV file is opened in *ParaView*, apply the `Table to Points` filter and specify the `X`, `Y` and `Z` columns appropriately in the filter properties. The seed points shall now appear at the desired locations. A sample image showing the seed locations is presented in Figure A.2.

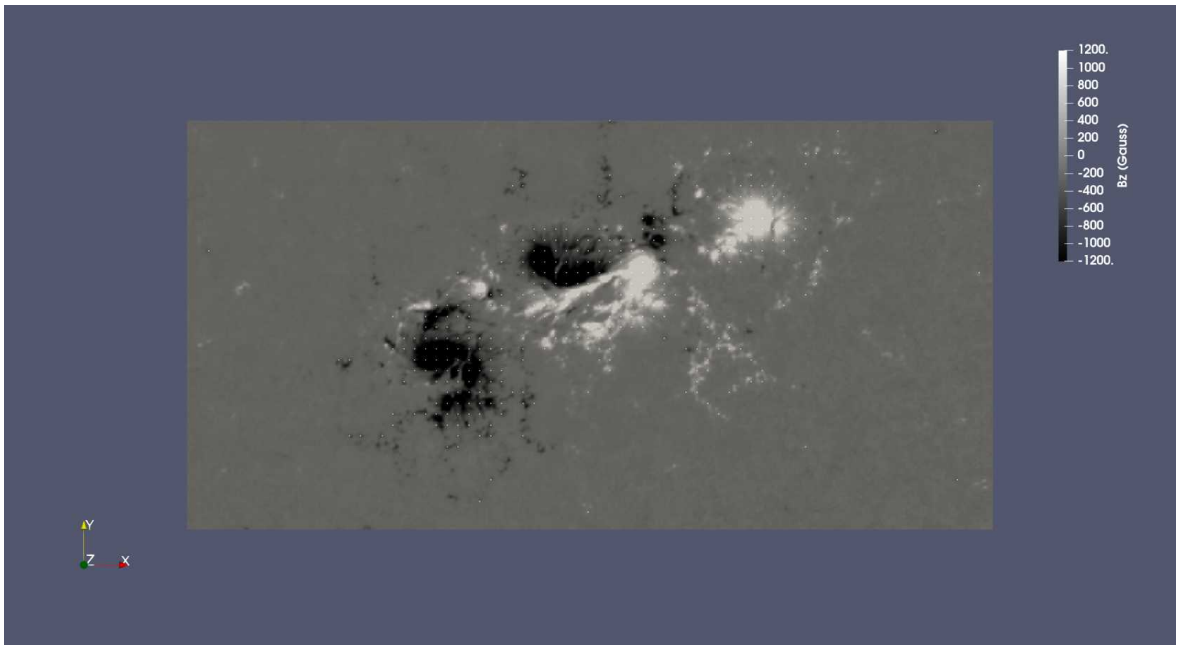

Figure A.2: A vertical view of the photospheric Bz with seed locations (look closely for white dots) for AR 11158 on 15 Feb 2011, 00:00 UTC using a threshold of 100 G

4. Application of Stream Tracer

This is the main step in generating the 3D magnetic field streamlines. The `Stream Tracer` filter generates the streamlines by first generating a select number of seed points and then takes the field values as boundary condition. *ParaView* offers the option of generating seed points on a straight line or within a sphere of pre-specified volume. However, since we have already specified the seed locations manually, we shall proceed with selecting the `.vtk` file in the pipeline browser and applying the `Stream Tracer with Custom Source` filter. In the dialog box that pops up immediately, select the `.vtk` file as `Input` and the `TableToPoints1` as `Seed Source`. Ensure that the `Runge Kutta 4-5 Intergation` is carried out in `both directions`. The streamlines shall now appear within the integration volume. Keeping the colorbar at (0.9, 0.1) and saturating the B_z colorbar thresholds to $\pm$ 400 G yields the output as shown in Figure A.3. After appropriate text annotation is added (see subsection A.1.4), the photospheric `slice` and field lines can be visualised simultaneously in one frame (see Figure A.4). Sometimes in switching between different filters/objects in the pipeline browser, the `colorbar` location/attributes may change due to default settings in *ParaView*. Please select `colorbar` and re-specify the location/attributes if needed.

5. Text Annotation

Go to `sources` in the menu bar and click on `text` within `annotation`. For the `title text`, use `font size` 8, `bold font` and `text format`: `AR ***** - yyyy/mm/dd/hr:mn:sc`. To show the `title text` at `center-top`, use switch number 2 under `text position` (6 switches in all). For the scale, repeat the process using another text box, preferably with switch number 4 (bottom-left location). Use `font size` 6, `bold font` to add any scale or dimension related information.

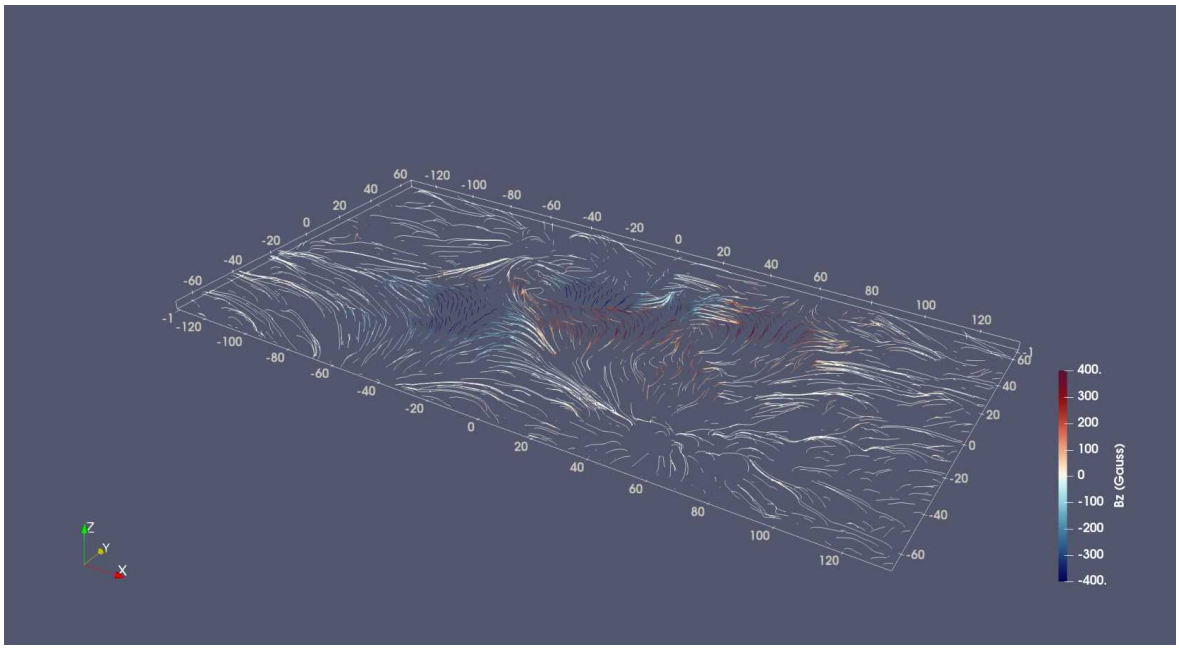

Figure A.3: Streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using a threshold of 20 G

6. Saving Data

Click on file and `save screenshot` in the desired location. We recommend saving the results in form of a `PVSM` state file for a quick retrieval later on, when needed.

7. Axis Exaggeration with Transform Filter

It is possible to exaggerate the vertical axis to get a better view. To do so, apply the `Transform` filter to the `StreamTracerWithCustomSource1` object in the pipeline browser after setting the scale as $1 : 1 : 20$ and unchecking the `Transform All Input Vectors` option. Uncheck the `Show Box` and reset the colorbar location if needed (see Figure A.5).

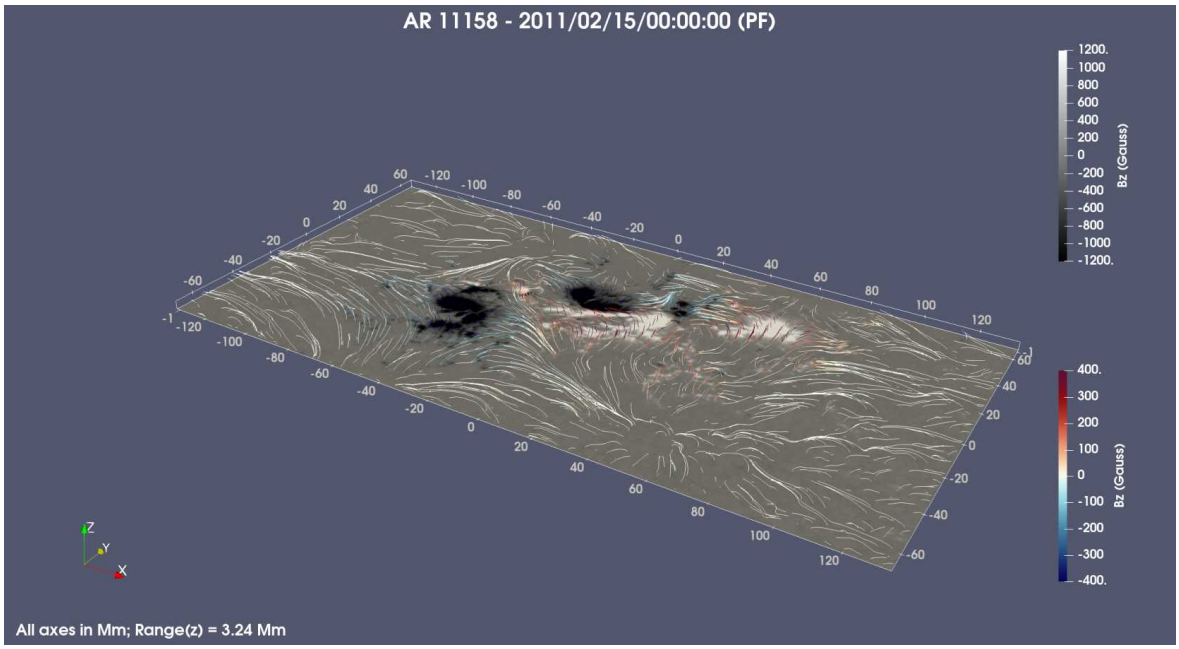

Figure A.4: Streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using a threshold of 20 G along with photospheric Bz, visualised along with the input magnetogram data

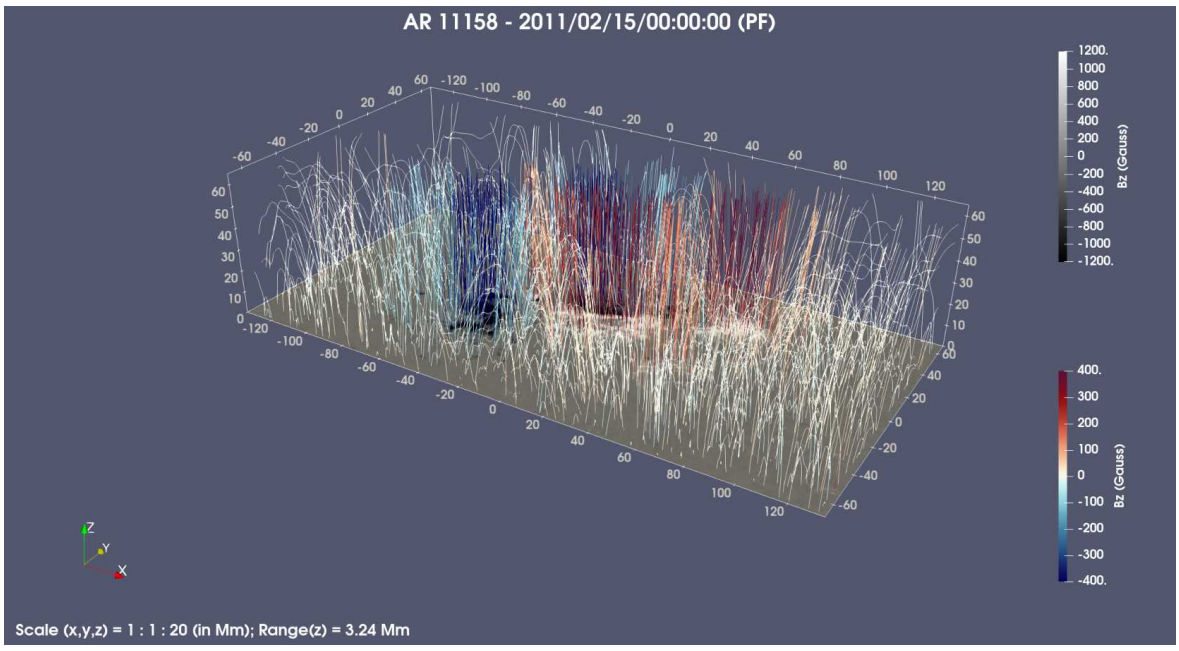

Figure A.5: An exaggerated view of Bz streamlines for AR 11158 on 15 Feb 2011, 00:00 UTC using a threshold of 20 G; the vertical axis has been exaggerated by a factor of 20

A.2 Method 2: ISEE data of Kusano et al. 2021

A.2.1 Input Data Assembly

The first step is to download the CDF file (e.g. `11158_20110215_000000.nc` for the same example as before) from the ISEE open-source website (https://hinode.isee.nagoya-u.ac.jp/nlfff_database) This file can be opened with *ParaView*, using the *NetCDF reader*. This CDF file contains the magnetic field information computed using PF and NLFFF methods. While the `(Bx, By, Bz)` vector information corresponds to the NLFFF case, `(Bx_pot, By_pot, Bz_pot)` corresponds to the PF case.

A.2.2 Visualisation with ParaView

1. Orientation and Surface Plot

Once the data is loaded, we first change the `attribute` from `outline` to `surface` and then select `Bz` if we wish to visualise NLFFF data and `Bz_pot` if we wish to visualise the PF data. For proper orientation of the axes, `adjust camera view` (5th button below the `Layout bar` in *ParaView 5.7.0*) by loading the file `standard_camera_137Mm.pvcc`. Unlike the data we handled in the previous section where the vertical extent was 3.24 Mm, it may be observed that the vertical extent of the data here is 137 Mm.

2. Application of Slice Filter

Apply the `Slice` filter as shown before in subsection A.1.4 with (0, 0, 1) as the normal vector and origin at (0, 0, 0.01). It is suggested that the `Grayscale` preset with a threshold ± 1200 G may be optimal for visualisation.

3. Importing Seed Points with Table to Points Filter

The seed points specified in the CSV file as chosed per the desired threshold (20 G/ 100 G). They are imported into *ParaView* in the same manner as before (refer Subsection A.1.4) using the `Table to Points` filter.

4. Applying the Calculator Filter

In order to obtain the vector information, apply the `Calculator` filter to the CDF input file. This step is absent in the method discussed previously as we had directly imported the vector data into the `.vtk` file. Below the `result array name` box, type:

```
>> Bx_pot*iHat+By_pot*jHat+Bz_pot*kHat (to calculate the PF vector) OR
>> Bx*iHat+By*jHat+Bz*kHat (to calculate the NLFFF vector)
```

5. Application of Stream Tracer

Select the `Calculator1` object (`input`) in the pipeline browser and apply the `Stream Tracer` filter as described before in subsection A.1.4, following which we shall have the complete visualisation of streamlines for the PF/NLFFF data upto 137 Mm (see Figures A.6 - A.9).

6. Application of Tube Filter

Select the `StreamTracerWithCustomSource1` object in the pipeline browser and apply the `Tube` filter with a `Capping Radius` of 0.2. Use a separate `colorbar`, preferably '`Cool to Warm`' with a threshold of ± 400 G. We suggest placing this `colorbar` at (0.9, 0.1).

7. Annotation and Saving Data

The title may be in `AR ***** - yyyy/mm/dd/hr:mn:sc (PF/NLFFF, 20/100 G)` format. The state files may be saved as `Kusano_(POT/NLF)_137Mm_20G.pvsm`. Additionally we have a camera file `vertical_camera_137Mm.pvcc` to generate the vertical view.

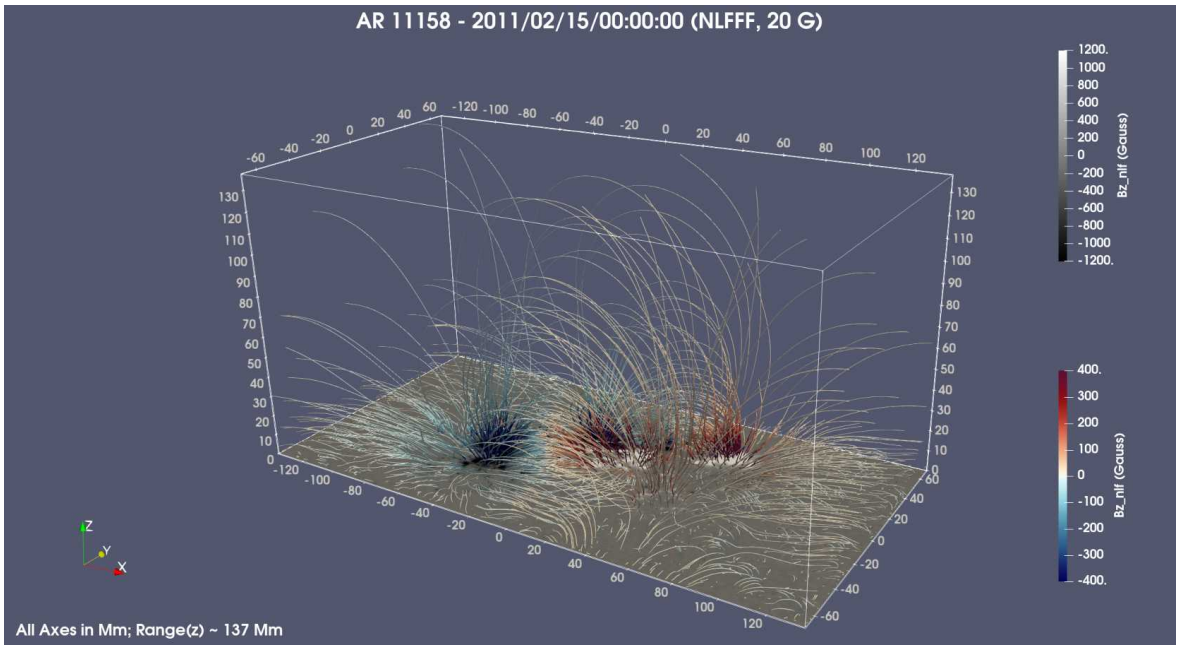

Figure A.6: Streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using ISEE NLFFF database with a threshold of 20 G

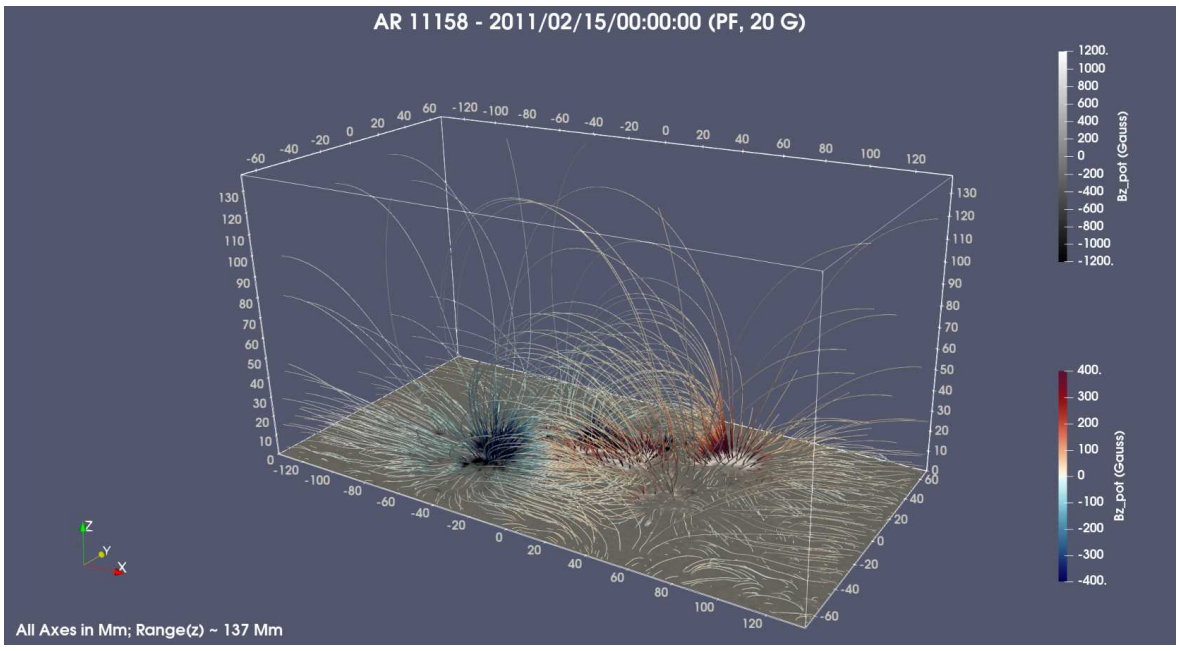

Figure A.7: Streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using ISEE PF database with a threshold of 20 G

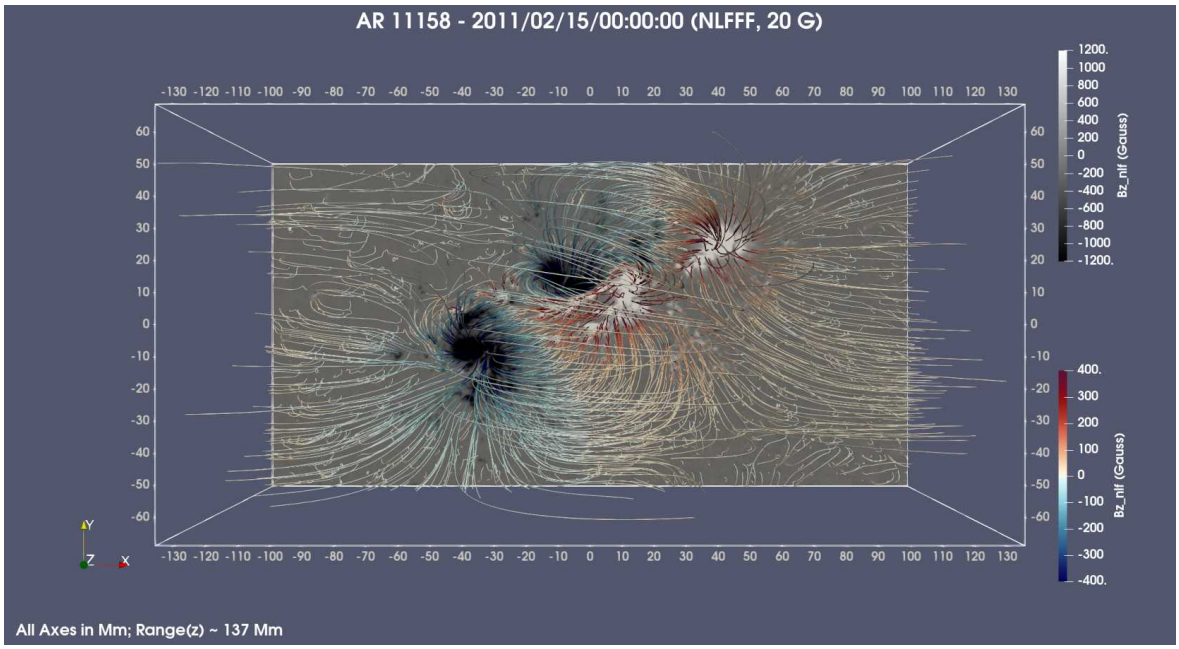

Figure A.8: Vertical view of streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using ISEE NLFFF database with a threshold of 20 G

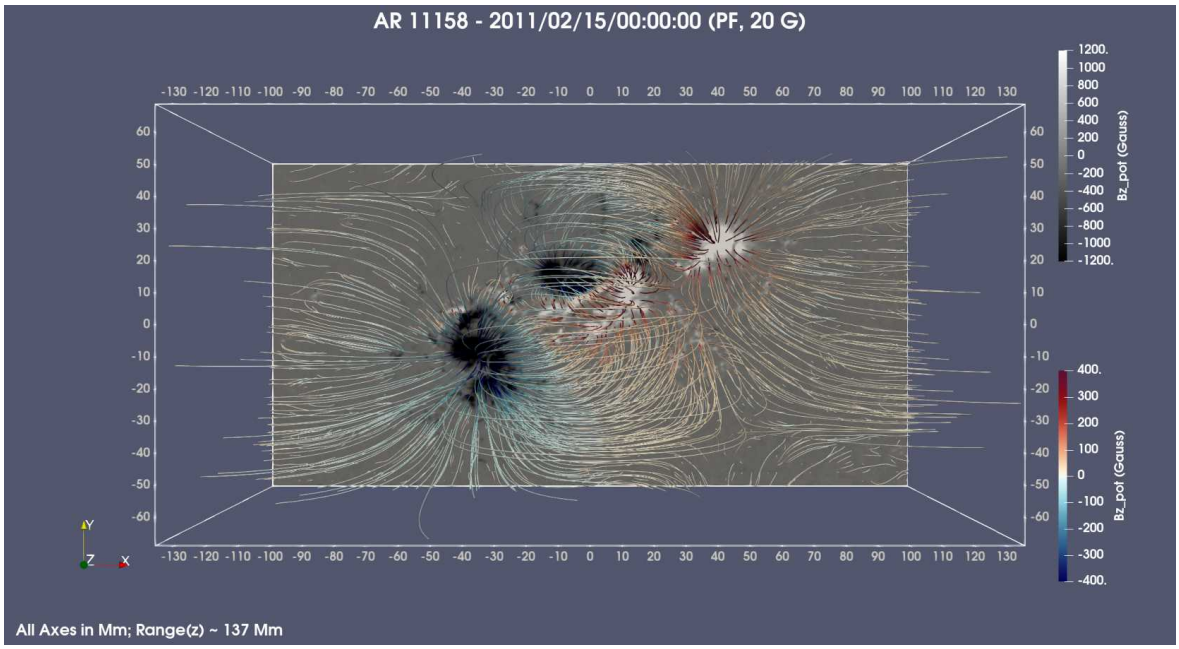

Figure A.9: Vertical view of streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using ISEE PF database with a threshold of 20 G

8. Using the Clip Filter

Now, that we have visualised the fields in the entire vertical extent of 137 Mm, it might be useful to know how to extract the visualisation of a subspace of the entire volume. This can be achieved using the `Clip` filter. For example, if the user is interested to access the B_z field in the 0 - 50 Mm height range for NLFFF case with a threshold of 20 G, then it best to load the pre-saved state file `Kusano_NLF_137Mm_20G.pvsm` into *ParaView*. Apply the `Clip` filter on to `Tube1` object in the pipeline browser. Set the `Clip Type` as `Box` and the z-coordinate of the `length` option to 50. To exaggerate the vertical axis, follow the steps as described before in subsection A.1.4 and use a suitable scaling factor. A sample of the result with a scaling factor of 2.5 is shown in Figure A.10.

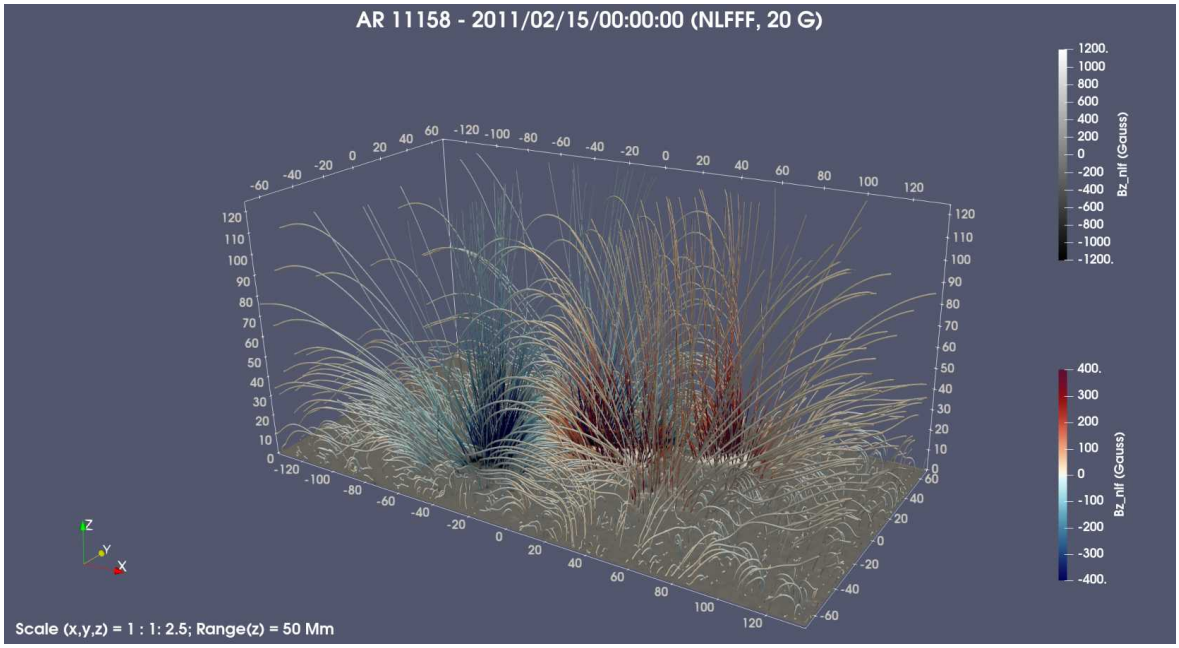

Figure A.10: Streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using ISEE NLFFF database with a threshold of 20 G upto a vertical extent of 50 Mm

A.3 Important notes

Overall, the PF and NLFFF magnetic configurations for AR 11158 closely resemble one another, especially when a higher threshold such as 100 G is chosen. At a lower threshold such as 20 G, the PF extrapolation appears to be smoother compared to the NLFFF extrapolation (for reference see top left section of Figures A.8 and A.9).

One key mistake that a user might commit while plotting magnetic field lines in *ParaView* is to exaggerate/scale the axes before performing the stream-tracing. By doing so we make the space non-homogeneous that is not suitable for the application of Runge-Kutta routines. This may lead to development of non-potential planes, indicating swirls in magnetic field lines, an example of which is shown in Figure A.11.

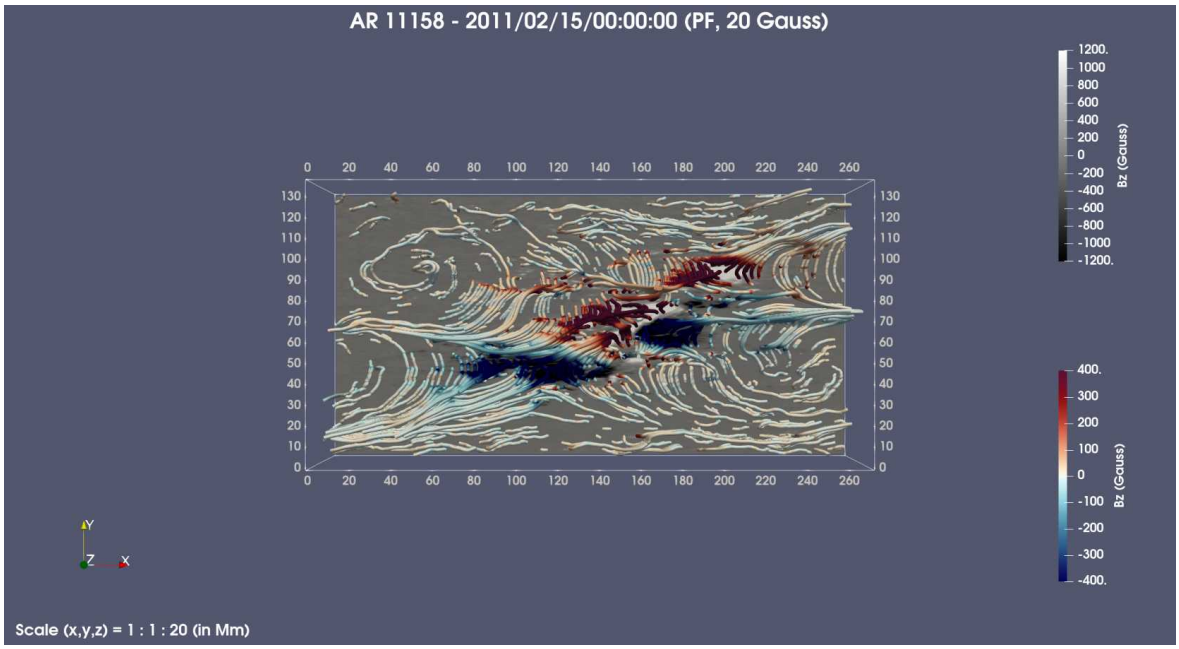

Figure A.11: In this figure the vertical extent has been exaggerated prior to streamline tracing

A.4 Additional Figures

This section contains additional visualisation samples (see figure captions for details).

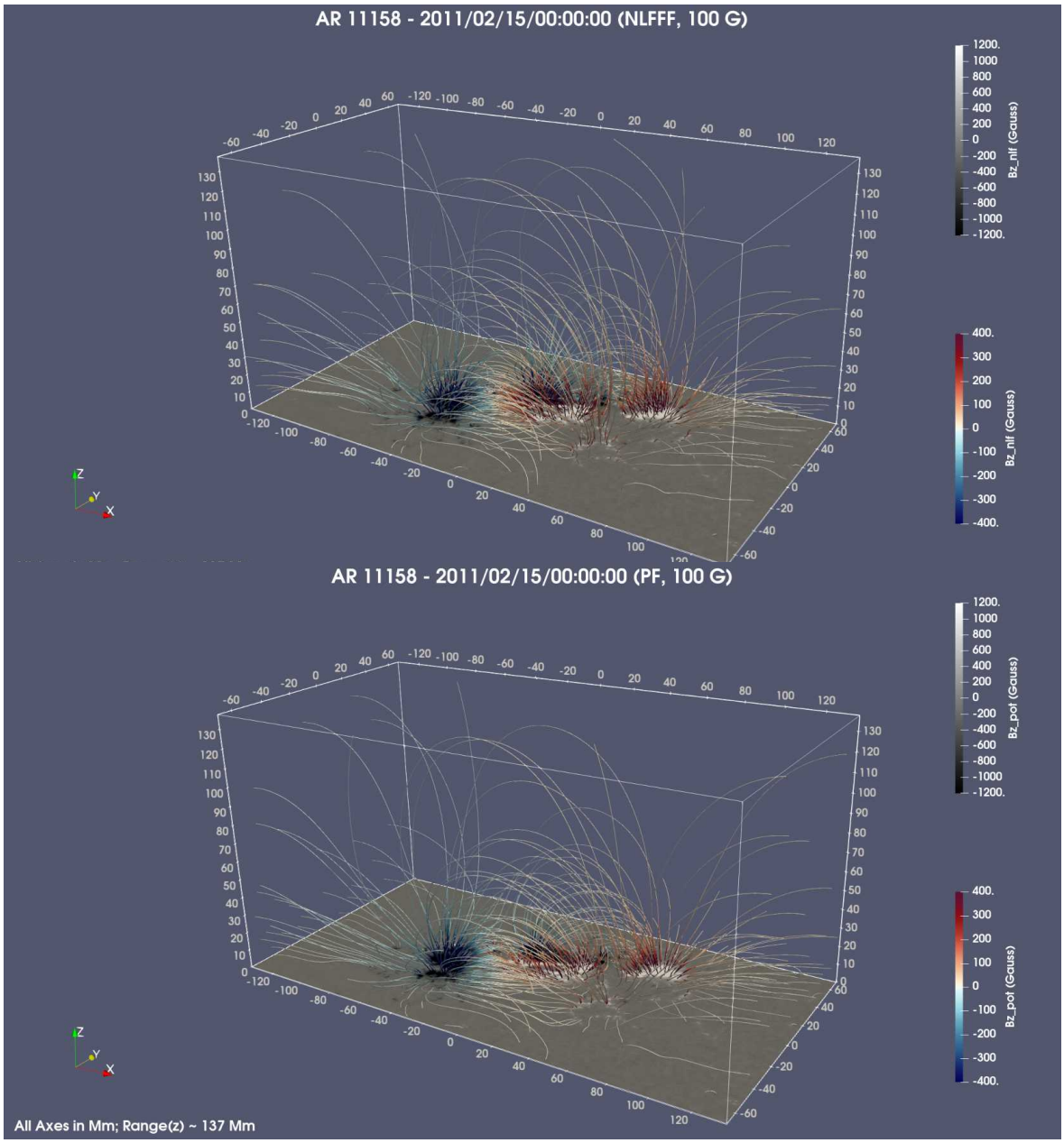

Figure A.12: Comparison of streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using NLFFF (top) and PF (bottom) ISEE databases with a threshold of 100 G

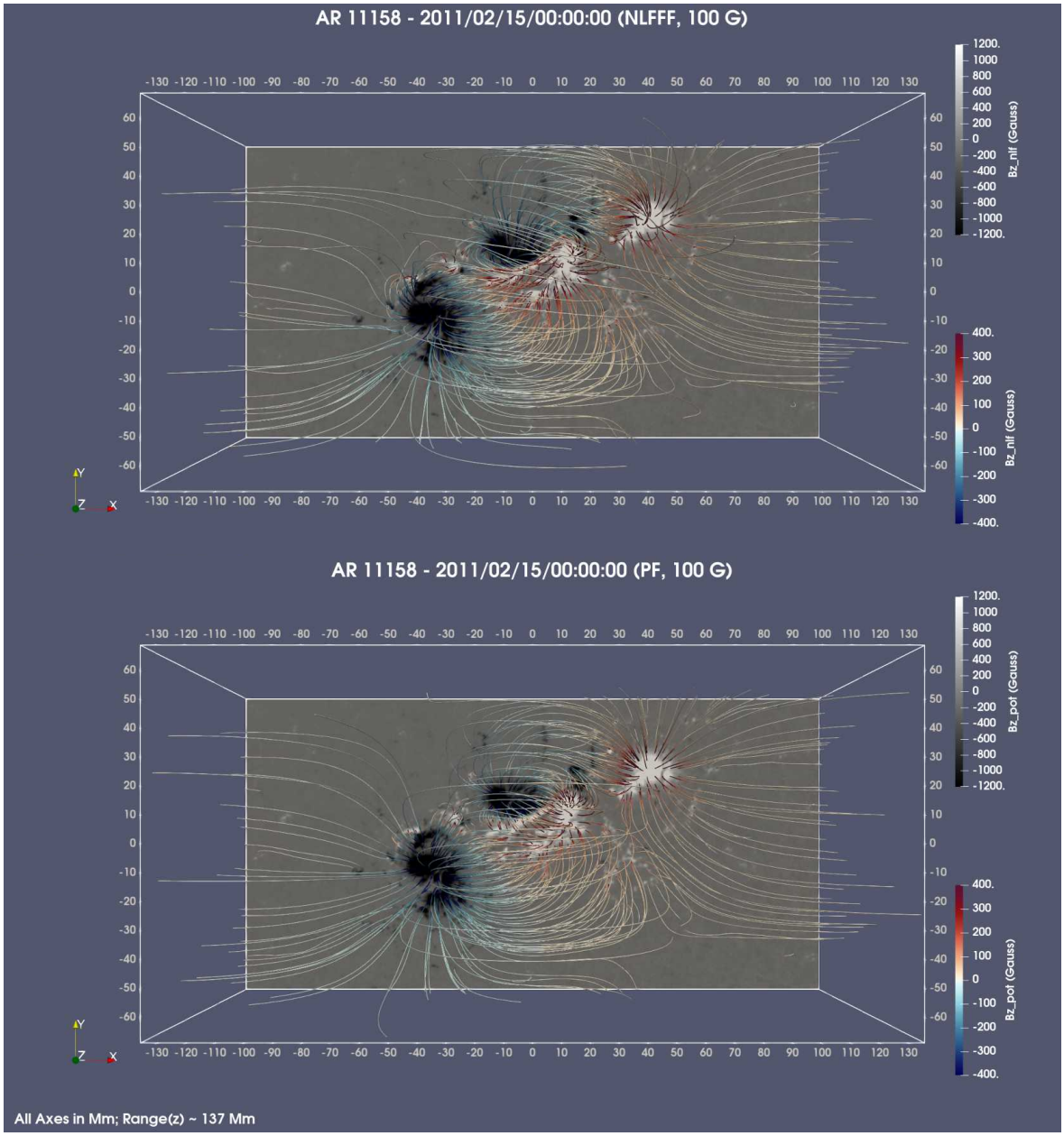

Figure A.13: Comparison of vertical views of streamlines computed for AR 11158 on 15 Feb 2011, 00:00 UTC using NLFFF (top) and PF (bottom) ISEE databases with a threshold of 100 G.

Appendix B

Industrial Training: SAMNET GUI Design

B.1 Introduction

The solar telescope at the GSO is a part of SAMNET (Solar Activity Monitoring Network - an international network of ground-based solar observatories) and it images the solar atmosphere at K-I spectral line which is at a wavelength of 770 nm (Erdélyi et al. 2022). In this internship, the preliminary work included preparing and assembling a freshly procured NUC (Next Unit of Computing) device, designed for high performance in space constrained environments. In this report, an overview of the overall software and hardware architecture at GSO are presented. Towards, the end of the internship, considerable effort was directed towards understanding GUI design using *PyQt* and *Qt Designer* and then preparing the layout of an additional tab called the `Fast Tab` to improve the overall GUI (Graphical User Interface) for SAMNET.

B.2 Installation of Ubuntu on the NUC device

The name of the NUC device is NUC-B for all future reference.

B.2.1 Preparation of a bootable USB drive

The pre-requisites for the installation of Ubuntu on the NUC device are (i) a Linux based PC and (ii) a bootable USB stick having a minimum storage capacity of 8.0 GB.

1. Download the *Linux-Ubuntu 20.04.6 LTS Focal Fossa** version file named, '`ubuntu-20.04.6-desktop-amd64.iso`' (size $\sim$ 4.4 GB) on the PC from the following website: https://releases.ubuntu.com/focal/

2. (i) To make the USB stick bootable using Launch Startup Disk Creator, follow the guidelines mentioned here: https://ubuntu.com/tutorials Extension ... tutorials/create-a-usb-stick-on-ubuntu#4-iso-and-usb-selection

   OR

   (ii) Alternatively, the bootable USB stick may be prepared from the terminal. To identify the USB stick name, type:

   ```
   >> lsblk
   ```

   Look for the storage space identifiers. For example, the USB stick with size 8GB may be identified as '`/dev/sdb`' if the size i.e. 8GB appears next to '`/dev/sdb`'. Once the USB stick name is known, type the following in the terminal:

   ```
   >> sudo dd bs = 4M if = /filepath/ubuntu-20.04.6-desktop-amd64.iso
   of = <stickname>
   ```

3. Go to `Disks` application and ensure that partition type reads `0x00 (Bootable)`. The bootable flash drive should be ready for use now.

4. Insert the flash drive in the NUC device and start the normal booting process. If the NUC device is unable to detect the USB stick (message: `"no bootable device detected"`), it may be possible that the NUC device only supports UEFI booting.

* Installation related problems were encountered with Ubuntu 20.04 LTS version (file name: `ubuntu-20.04-desktop-amd64.iso`, size $\sim$ 2.7 GB).

B.2.2 Booting process from the UEFI terminal

The Unified Extensible Firmware Interface (UEFI, successor of the EFI) is an interface between operating systems and firmware. It provides a standard environment for booting an operating system and running pre-boot applications. It is distinct from the MBR boot code method that was used by legacy BIOS systems.

1. Open BIOS settings (Ctrl + Alt + Del, followed by pressing F2 repeatedly). Ensure that `Secure Boot` is disabled. Go to `Boot Priority` and ensure that in addition to UEFI Boot, the following options are ticked: (i) `Boot USB Devices First`, (ii) `Internal UEFI Shell` and (iii) `USB`. Boot Option #1 should read '`UEFI: Built-in EFI Shell`'. Save and exit (press F10).

2. Restart the NUC device and select `Normal Boot` (Esc). After a while, the UEFI Shell screen should appear. Type `>> mount` in the UEFI shell. This should display a mapping table with a set of identifiers. Try to guess the USB stick identifier from the set (for ex. the correct identifier could be `blk0`).

3. Type the following to complete the booting process (retry if initial guess is incorrect):

`>> mount blk0`

`>> fs0:`

`>> cd efi\boot\` (type `ls` & to see `bootx64.efi`, `grubx64.efi`, `mmx64.efi`)

`>> grubx64.efi` (installation process should be pretty simple after this point)

B.2.3 Installation checklist

Make sure to check the `install third party drivers` option during the installation procedure. Also, allow the installation of basic application and utilities during the installation process. Do not allow any driver/software updates during the installation process. Finish the installation process by setting the time to UTC. Type the following in the terminal:

```
>> sudo timedatectl set-timezone "UTC"
```

B.3 Installation of Software Packages

B.3.1 Packages installed directly from terminal

Commands for the installation of requisite software packages/dependencies (25 in total) from the terminal window are provided below:

```
>> sudo apt install openssh-server
>> sudo apt install python-is-python3
>> sudo apt install python3-serial
>> sudo apt install python3-astropy
>> sudo apt install python3-matplotlib
>> sudo apt install python3-dev
>> sudo apt install python3-netcdf4
>> sudo apt install python3-regions
>> sudo apt install python3-pyro4
>> sudo apt install python3-pyqt5
>> sudo apt install python3-ephem
>> sudo apt install python3-pip
>> sudo apt install htop
>> sudo apt install curl
>> sudo apt install saods9
>> sudo apt install libusb-dev
>> sudo apt install libcfitsio-dev
>> sudo apt install build-essential
>> sudo apt install libgsl-dev
>> sudo apt install emacs-gtk
>> sudo apt install screen
>> sudo apt install net-tools
```

```
>> sudo apt install lm-sensors
```

```
>> sudo apt install hddtemp
```
(not need for NUCs without external HDD)

```
>> pip install rpyc
```
(can only be installed after pip is installed)

B.3.2 Packages requiring manual installation

It is preferable to transfer the downloaded file to a new directory called `Installations` to keep a track of downloaded softwares. In this folder store the downloaded installation files for *ASI Studio* and *Atom* for future reference. The *ASI Studio* software package (version 1.12) for Linux systems may be obtained from the following website: https://www.zwoastro.com/software/. To install ASI Studio, type the following in the terminal at the location of the downloaded file `ASIStudio_V1.12.run`:

```
>> chmod +x ASIStudio_V1.12.run
>> ./ASIStudio_V1.12.run
```

The *Atom* software package (version 1.60.0) for linux systems may be downloaded from the following GitHub repository: https://github.com/atom/atom/releases/tag/v1.60.0. Download the file named `atom-amd64.deb` from the GitHub repository and move it to the `Installations` directory (preferably). Type the following at the location of the debian file:

```
>> sudo apt install ./atom-amd64.deb
```

B.3.3 Other non-essential packages

For testing and analysis of the control software and data pipeline, it is preferable to have *conda* and *Spyder* to interact with and manage python environments. *Mini-*

conda3 is a light-weight variant of *conda* and the source bash file for installation may be accessed from the following website: https://docs.anaconda.com/miniconda/. Download the file of the name `Miniconda3-latest-Linux-x86_64.sh` and execute the bash script from the terminal window at its location by typing in the following:

```
>> bash Miniconda3-latest-Linux-x86_64.sh
```

During the installation procedure, respond 'yes' to the permission request for automatic editing of the source shell file. After the installation is complete, exit the terminal and enter back again into the terminal. You may now find that *conda* `base` environment opens by default. Since it is not preferable to open *conda* by default, type the following in the terminal:

```
>> conda config --set auto_activate_base false
```

To install *Spyder*, type the following in the terminal:

```
>> sudo apt install spyder
```

Some really basic but important conda commands are listed below:

`>> conda activate` (activates the base conda environment)
`>> conda deactivate` (deactivates the current conda environment)
`>> conda create -n <environmentname>` (creates a new conda environment)
`>> conda activate <environmentname>` (activates a specific conda environment)
`>> conda list` (lists all packages in the current conda environment)
`>> conda install <packagename>` (installs package in current conda environment)
`>> spyder` (opens the Spyder GUI to manage, compile and execute python files)

B.4 Establishing Remote Access (at ELTE)

B.4.1 Within a LAN

To access the NUC device via LAN and ssh, it is important to have ssh installed on a second PC on the same LAN. To install ssh on the local PC, type the following in the terminal after obtaining the username* and the IP address** of the NUC device.

```
>> sudo apt install openssh-server
>> sudo systemctl status ssh (to verify if ssh is actively running)
>> ssh <username>@<ipaddress>
```

* To obtain the username of the NUC device, check for the name appearing in the terminal window of the form `username@machinename`.

** To obtain the IP address of the NUC device, type `>> ifconfig` in the terminal of the NUC device. The IP address should be of the form `192.168.x.y`, where x and y are integers.

In case of the NUC device that was assembled at ELTE, the username and IP address on the LAN were `gsonuc2@gsonuc2-NUC12WSHi7` and `192.168.3.102`, respectively. The NUC device was connected to the LAN via. ethernet.

B.4.2 Using Anydesk

```
>> wget http://ftp.de.debian.org/debian/pool/main/g/gtkglext/
libgtkglext1_1.2.0-11_amd64.deb
>> sudo dpkg -i libgtkglext1_1.2.0-11_amd64.deb
>> sudo dpkg -i anydesk_6.3.2-1_amd64.deb
```

B.5 Current hardware architecture at GSO

B.5.1 Original layout

The original layout of the hardware/network architecture at GSO is presented in Figure B.2. The workstation is connected to 6 major devices: The dome control box, the weather station, cameras A and B, a 4 TB HDD for data storage, a router and hub. The hub is connected to the temperature control monitors for MOF and WS, cooling control systems for cameras A and B, finder camera (F), mount control and the focuser. The router is further connected to the four IP cameras. The workstation draws internet from the `192.168.100` network and can be connected to any office PC on this network via SSH.

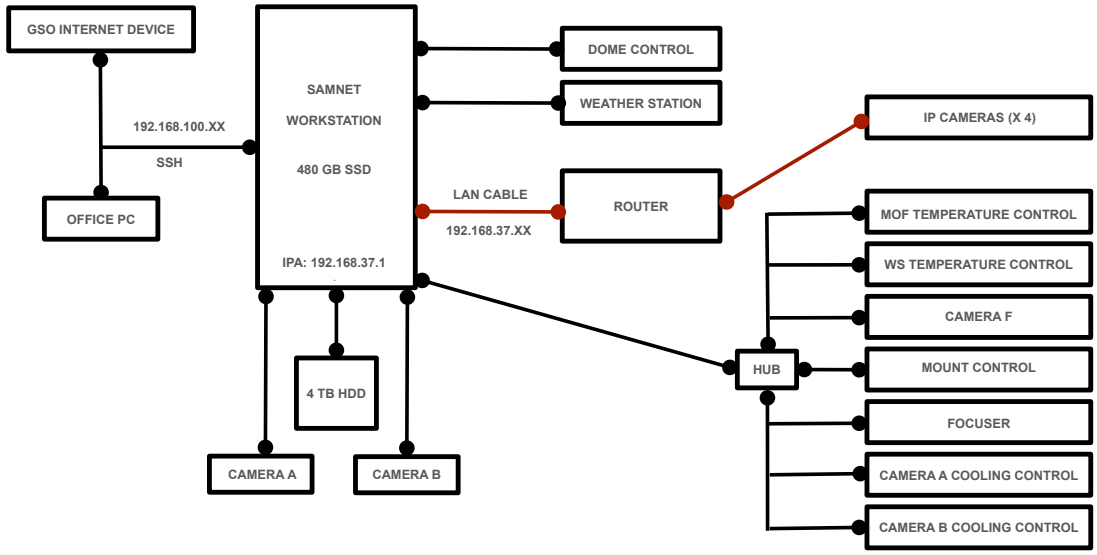

Figure B.1: Original layout of the network architecture at GSO; the cables in red are LAN cables

B.5.2 Updated layout for testing purposes (Oct 03-04, 2024)

The updated layout of the hardware/network architecture at GSO for testing purposes is presented in Figure B.2. In this setup, the workstation was not connected to the cameras A and B for testing purposes. The router was connected to NUC device in addition to the workstationa and the IP cameras. Since NUC-B had a very limited number of USB ports, the the dome control, hub and weather station were switched with cameras A and B at the USB ports for occasional testing.

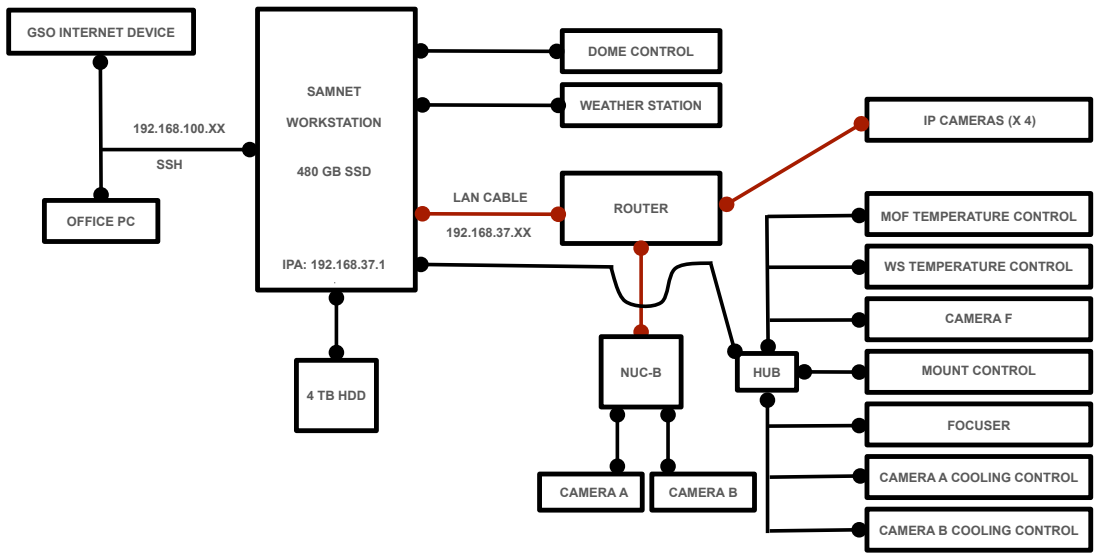

Figure B.2: Updated layout of the network architecture at GSO; the cables in red are LAN cables

B.5.3 Connecting NUC-B and Telescope, Oct 03, 2024

The central network at the observatory that hosts the wifi is identifiable by the IP-address: `192.168.100.156`. A network branch hosting the telescope equipment such as the scientific cameras is connected to this central network. The IP-address of the network branch is `192.168.37.1`. The IP address details of the devices on SAMNET network are:

- `192.168.37.41` - ipcam1
- `192.168.37.42` - ipcam2
- `192.168.37.43` - ipcam3
- `192.168.37.44` - ipcam4
- `192.168.37.51` - power1
- `192.168.37.52` - power2

The following lines are desirable for IPv6 capable hosts:

- ::1 ip6-localhost ip6-loopback
- fe00::0 ip6-localnet
- ff00::0 ip6-mcastprefix
- ff02::1 ip6-allnodes
- ff02::2 ip6-allrouters

These details may be verified by typing the following command in the terminal:

```
>> cat /etc/hosts
```

Via. Wifi

To connect to this IP-address via ssh, access the central network user first and then connect to the branch network on the same user. To do that, type following in the terminal, sequentially:

```
>> ssh -Y samnet@192.168.100.156
```

`>> ssh -Y samnet@192.168.37.1` (user switches to `samnet@samnet-workstation`)

Via. Ethernet

Alternatively NUC-B may be connected directly to the router via. an ethernet cable. For this the default ethernet settings need to be edited on the NUC-B. To edit the ethernet settings, select the IPv4 option in the 'Wired' settings and change the Ipv4 method to 'Manual' instead of 'Automatic (DHCP)', which is the default setting for use at ELTE. The IP Address, Netmask and Gateway should be set to '`192.168.37.2`', '`255.255.255.0`' and '`192.168.37.1`', respectively. The DNS (Domain Name System) must be set to '`8.8.8.8, 8.8.4.4`'. Following this, the NUC would be assigned an IP address '`192.168.37.2`' and it can be connected to the equipment network via. SSH.

```
>> ssh -Y samnet@192.168.37.1
```

B.5.4 Additional information regarding running samnetGUI

Kindly ensure that the following checks are amde before the control software is run from samnetGUI.

- The user for running the control software should be added to the dialout group from the main user which has `sudo` access. The command is:

  `>> sudo usermod -a -G dialout samnet`

- In the samnet directory of the source code, hostname '`samnet-workstation`' files should be changed to '`gsonuc2-NUC12WSHi7`'.

  (a) More particularly, in the parameters file `samnetParams.py`,

  `if socket.gethostname() in ['samnet-workstation']:` is modified to

  `if socket.gethostname() in ['samnet-workstation', 'gsonuc2-NUC12WSHi7']:`

  (b) Similarly in `updateWebPage.py`,

  `assert socket.gethostname() == 'samnet-workstation'` is modified to

  `assert socket.gethostname() == ['samnet-workstation', 'gsonuc2-NUC12WSHi7']`

- In the samnet directory, the path must be appropriately specified in the main script file `samnetgui.py`.

- To run the GUI from the terminal, the command is `samnetGUI`. This user on NUC-B used to run this command should be '`samnet`' (and not `gsonuc2`).

- The relevant hostnames must be added to `cat/etc/hosts` (see the hostnames after the space in the hosts file in the root directory)

B.6 Control Software Architecture and Modules

The control software is written in Python and C. The main file that launches a GUI to interactively control the telescope is `samnetGUI.py`. The main file imports an UI (User Interface) file called `samnetGUI.ui` which concerns the overall design and layout of the GUI and contains the descriptions of the control software application components that the user may modify. Before running the control software make sure that the path is correctly set in `samnetGUI.py`. For example, for the NUC device at ELTE, the correct path is:

```
if os.path.isfile('/home/samnet/samnet/samnetGUI.ui'):
        qtcreator_file = "/home/samnet/samnet/samnetGUI.ui"
```

Here some information on important dependencies concerning the design of the GUI are shared are shared:

B.6.1 About PyQt

PyQt is a Python binding of the cross-platform GUI toolkit *Qt*, implemented as a Python plug-in. *PyQt5* contains the following Python modules: *QtQml, QtQtuick, QtCore, QtGui, QtPrintSupport, QtWidgets, QGLContext, QGLFormat, QGLWidget, QtWebKit* and *QtWebKitWidgets*. *PyQt* helps the user to build web applications and GUIs easily.

For development of the 'Fast Tab' (described later), we made of some Qt C++ classes like `QObject`, `QLabel` and `QPushButton`. For offcial documentation of these classes, please visit https://doc.qt.io/qt-6/classes.html.

B.6.2 About Pyro

Pyro is the abbreviation of *PYthon Remote Objects*. It is a library that enables building applications in which objects can communicate to each other over the network, with minimal programming effort. *Pyro* is used for the devices to communicate with each other on a network and for this project, it is used to make the buttons in the Fast Tab functional (for Fast Tab, see discussion in next section).

B.7 New Layout Template for GUI

B.7.1 Limitations of the original GUI

1. The GUI in its basic version is not very efficient in the sense that there are multiple buttons to be pushed in sequence to activate power supply to the telescope components. Ideally, we would like the `Power control` button to cascade the turning on of of the following components in sequence (a-h) without turning each one on individually:

   (a) `MOF`

   (b) `WS`

   (c) `Mount`

   (d) `spare(4)`

   (e) `Rotation Stage`

   (f) `Focus Stage`

   (g) `CamCool1`

   (h) `CamCool2`.

2. Similarly, we would like a button to cascade the turning on of all the components in the `Process Control`.

3. During the process of taking images from camera A and camera B, a considerable amount of time is being spent currently in switching between cameras A and B and in switching between 0 and 45 degrees of rotation for the half-wave plate. Ideally we would like to have a fast single button that (a) stops auto-guiding of science cameras, (b) obtain the set of 4 images, (c) saves them with the help of a unique file name and (d) restarts auto-guiding, in sequence.

B.7.2 Objectives for the new GUI template

The main objective is to improve the GUI layout and minimise time-loss during its operation. This may be done by addressing each of the limitations mentioned in the previous segment. We would like to achieve all the desired changes in a new tab named `Fast Tab`. For convenience, the `Fast Tab` must have:

1. a button to cascade power control
2. a button to cascade process control
3. buttons to operate quick measurements and save the observations in FITS file format.
4. space to display the images from the weather camera (to check the presence of clouds along the line-of-sight), finder camera, science cameras A and B
5. details of the half-wave rotation and focus adjustment
6. a table checking the temperature stability of `MOF`, `WS` and the science cameras

B.7.3 Installation of Qt5 Designer

The *Qt5 Designer* helps create a layout of the GUI template and generates the `.ui` file, subsequently. The .ui file may then be imported into the main python file for launching the GUI. Currently in the main python file i.e. `samnetGUI.py`, the following patch of code at the end imports the `samnetGUI.ui` file:

```
if __name__ == "__main__":
app = QtWidgets.QApplication(sys.argv)
if 'sim' in sys.argv:
window = MyWindow(sim=True)
else:
```

```
window = MyWindow()
window.show()
sys.exit(app.exec_())
```

To install *Qt5 Designe*r, type the following in the terminal:

```
>> sudo apt install qtcreator pyqt5-dev-tools
```

B.7.4 Details of the Fast Tab

The following steps were used to generate the Fast Tab (see design in Figure **??**):

1. In the `Object` tab of the `ObjectInspector`, the `tabWidget` option was right-clicked and `Insert Page` was selected. A new tab is added in the existing layout.
2. The `currentTabText` in the `QTabWidget` tab was selected and the tab was renamed as `Fast Tab`.
3. A `Push Button` was dropped into the `Fast Tab` and the name was saved as `Fast Observations` in the text option under `QAbstractButton`.

Several components of the `Fast Tab` are descried below:

1. Fast Observations Section

The following steps were used to generate the `Fast Observations Section` under the `Fast Tab`:

1. A grid layout of `QLabels` was created to show the status of cascade process i.e. `Fast Cascade Status`. The default text in the status shows as '- - -' when the app is opened or when the refresh cascade switch is turned on but when the

individual processes (`Stop Autoguiding`, Camera A 0 degree - 100 images ... `Restart Autoguiding`) are completed, the text shows as 'complete'.

2. A `QPushButton` was dropped below the `Fast Cascade Status` grid and the named as `Refresh Cascade`. The purpose of the `Refresh Cascade` button is to prepare the GUI for the next set of observations.

2. Imaging Section

To generate the layout for the `Imaging Section`, a horizontal grid layout (4 x 1) was created to show the images obtained from `Weather Camera`, `Finder Camera`, `Camera A` and `Camera B` using 4 `QObjects`. The title `QLabels` are centered using the `AlignHCenter` option.

3. Mechanisms Section

This layout for this section for `Half-wave plate rotation` and `Focus adjustment` was copied from the `Mechanisms Tab` and slightly adapted to match new dimensions.

4. Temperature Monitor Section

The layout for this section includes a 5 x 2 grid layout of `QLabels` to show the overall temperature status and temperature status of `MOF`, `WS`, `Camera A` and `Camera B`. In addition, `QPushButtons` were included for turning ON/OFF the `Set Cooling` mode.

5. Control Sections

This section mainly has two major `QPushButtons`, namely `Cascade Power Control` and `Cascade Process Control` to turn on the `Power Control` and `Process Control` in sequence.

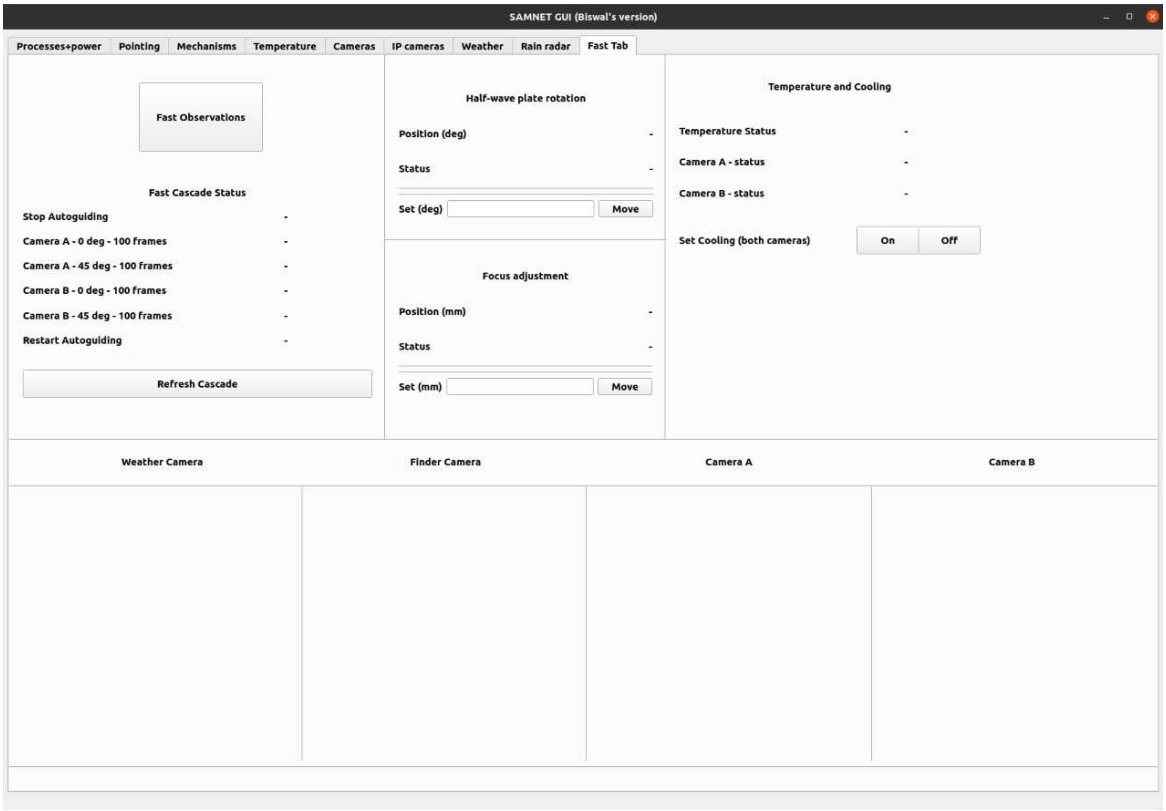

Figure B.3: The Fast Tab design created using *Qt Designer*

Thanks ...

An explanation of "the aerodynamics of paper planes" at the SWATNET mini-MBA course, University of Sheffield (March 2023)

Lunch at John's Van, University of Sheffield during the SWATNET Annual Meeting (March 2023)