An analysis of sources and predictability of geomagnetic storms

A thesis submitted in fulfilment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

of

RHODES UNIVERSITY

by

Jean Uwamahoro

November 2011

Abstract

Solar transient eruptions are the main cause of interplanetary-magnetospheric disturbances leading to the phenomena known as geomagnetic storms. Eruptive solar events such as coronal mass ejections (CMEs) are currently considered the main cause of geomagnetic storms (GMS). GMS are strong perturbations of the Earth's magnetic field that can affect space-borne and ground-based technological systems. The solar-terrestrial impact on modern technological systems is commonly known as Space Weather. Part of the research study described in this thesis was to investigate and establish a relationship between GMS (periods with $Dst \leq -50$ nT) and their associated solar and interplanetary (IP) properties during solar cycle (SC) 23. Solar and IP geoeffective properties associated with or without CMEs were investigated and used to qualitatively characterise both intense and moderate storms. The results of this analysis specifically provide an estimate of the main sources of GMS during an average 11-year solar activity period. This study indicates that during SC 23, the majority of intense GMS (83%) were associated with CMEs, while the non-associated CME storms were dominant among moderate storms. GMS phenomena are the result of a complex and non-linear chaotic system involving the Sun, the IP medium, the magnetosphere and ionosphere, which make the prediction of these phenomena challenging. This thesis also explored the predictability of both the occurrence and strength of GMS. Due to their non-linear driving mechanisms, the prediction of GMS was attempted by the use of neural network (NN) techniques, known for their non-linear modelling capabilities. To predict the occurrence of storms, a combination of solar and IP parameters were used as inputs in the NN model that proved to predict the occurrence of GMS with a probability of 87%. Using the solar wind (SW) and IP magnetic field (IMF) parameters, a separate NN-based model was developed to predict the storm-time strength as measured by the global Dst and ap geomagnetic indices, as well as by the locally measured K-index. The performance of the models was tested on data sets which were not part of the NN training process. The results obtained indicate that NN models provide a reliable alternative method for empirically predicting the occurrence and strength of GMS on the basis of solar and IP parameters. The demonstrated ability to predict the geoeffectiveness of solar and IP transient events is a key step in the goal towards improving space weather modelling and prediction.

Acknowledgements

Many people have contributed to the successful completion of this thesis. In particular, I am most grateful to my supervisor Dr Lee-Anne McKinnell for critically reading and commenting on the full thesis. Her useful suggestions and constructive comments contributed to the improvement of my work. I am indebted to her for her patience, encouragement and meticulous guidance during whole period of my Ph.D studies. I would like to express my sincere gratitude to Dr John Bosco Habarulema for his invaluable contribution by reading and commenting on my thesis, as well as for his technical assistance in Matlab programming and use of the linux operating system.

I am grateful to the National Astrophysics and Space Science Programme (NASSP), University of Cape Town, for funding my postgraduate studies. I wish to acknowledge the Rwandan government's Student Financing Agency (SFAR) for providing partial financial support for the completion of my studies. I am grateful to the South African National Space Agency [SANSA] Space Science, in Hermanus, South Africa, for providing financial support, a good working environment and other facilities which allowed me to complete my research project. Many thanks to the SANSA Space Science management, researchers and other staff members who assisted me in one way or another. Throughout my Ph.D studies, I had opportunity for international research collaboration. I am grateful to the Rhodes University, the International Center for Theoretical Physics (ICTP) and the SANSA Space Science for the funds provided to participate in conference, school and research visit respectively in Bremen, Germany, Trieste, Italy and Prague, Czech Republic.

I would like to express my sincere gratitude to all my colleagues, the SANSA Space Science students. I benefited from the fruitful discussions and socialisation. Special thanks are addressed to my classmates Mr Stefan and Mr Chigo as well as Dr Patrick, Mr Emmanuel, Mr Daniel and Mrs Mpho. I also appreciate encouragement and advices received from my friends and colleagues, particularly Dr Pheneas and Professor Yadav of Kigali Institute of Education (KIE), as well as Dr Schadrack of Rhodes University.

Last, but not least, my heartiest thanks are addressed to my wife Mrs Innocente, and sons, Regis and Aurel. I am most grateful to all members of my family for their moral support, in particular my Mother Astérie and young brother Leonard. Thanks, oh Lord, for your blessings!

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Transient solar phenomena occurring on the Sun are often the sources of interplanetary-magnetospheric disturbances known as geomagnetic storms (GMS), which can affect human technological systems in various ways. The research study described in this thesis focuses on the analysis of sources and predictability of GMS.

1.1 Solar-Terrestrial interaction

The Sun is the most important and nearest star to Earth, sustaining life on our planet through its irradiation. However, the Sun's dynamic processes can sometimes expel huge amounts of magnetised plasma, high-energy particles and harmful radiation producing disturbances in the Earth's space environment. In recent years, research on the solar-terrestrial environment has increased dramatically following advancements in satellite observation from the near-Earth environment. The growing interest in the study of the solar-terrestrial environment is driven by modern life, which relies heavily on ground-based and space-borne technological systems, susceptible to space weather effects (Moldwin, 2008).

Space weather is defined as the physical conditions on the Sun, which drive disturbances in the solar wind, magnetosphere, ionosphere and thermosphere that can influence the performance of space-borne and ground-based technological systems or even directly affect human health. The origin of space weather is the Sun, in which a combination of complex phenomena occur, including the dynamo mechanism, convection and differential rotation. These phenomena produce plasma fluid motions in a turbulent way extending to the outer layers of the Sun and into interplanetary space (Chian and Kamide, 2007).

Historically, the connection between solar activity and disturbances in the Earth's magnetic field was first suspected by Richard Carrington (a British amateur astronomer) after observing a powerful flare on 1 September 1859, which was followed by a severe magnetic storm about one day later. Hence, in the early days of solar forecasting, it was assumed that the occurrence of a solar flare would be followed by a geomagnetic disturbance. However, it was later discovered that CMEs were the main causes of GMS (Gosling et al., 1991; Cargill and Harra, 2007). CMEs-like phenomena and their terrestrial impact were first suggested by Lindermann (1919). In the entire solar system, CMEs and often associated solar flares are observed as the most powerful explosions during which the emitted magnetised plasma and energetic particles may strongly interact with the Earth's magnetosphere. On 9 March 1989, a CME erupted on the Sun and arrived on Earth four days later, producing a severe magnetic storm on 13 March 1989. This storm caused the collapse of Canadian Hydro-Qubec power grid for nine hours and led to a significant economic distress (Boteler, 2003). Other similar events and their effect on the Swedish technical systems can be found in the paper by Wik et al. (2009).

CMEs are transient expulsions of plasma and magnetic field from the Sun which are often responsible for strong interplanetary (IP) disturbances and subsequent non-recurrent and recurrent GMS (Sheeley et al., 1985; Crooker and McAllister, 1997). GMS are strong perturbations of the Earth's magnetosphere that can affect our modern technological society in various ways. The main solar sources of GMS are believed to be the CME eruptions from the Sun and the corotating interaction regions (CIRs).

1.2 Research framework

The new era of space-based instruments has allowed great advances in observation and understanding of storm events occurring on the Sun. The Solar and Heliospheric Observatory / Large Angle Spectrometric Coronagraph (SOHO/LASCO) (Brueckner et al., 1995) has been detecting the occurrence of CMEs on the Sun for almost two decades. However, although CMEs are considered major drivers of GMS, there is no one-to-one relationship between the CME eruptions and the occurrence of GMS events.

Solar observations show that CMEs that appear to surround the occulting disk of the observing coronagraphs, known as *halo* CMEs, have the highest probability of impacting the Earth's magnetosphere (Webb et al., 2000) when they originate from the visible solar disc and are Earth-directed. However, not all halo CMEs are associated with GMS, and non-halo CMEs can also cause intense GMS events if they arrive at Earth with an enhanced southward component of the Earth's magnetic field with high speed (Gopalswamy et al., 2007). A number of GMS events are often identified without any link to frontside halo CMEs (Schwenn et al., 2005). Hence, studies towards exploring the ability to estimate the geoeffectiveness of CMEs are of practical importance in the domain of space weather modelling and prediction.

The ability of CMEs to produce GMS does not only depend on geoeffective properties of CMEs when launched from the Sun, but is also governed by the way in which the expelled Sun's magnetised plasma connects with the Earth's magnetic field. A quantitative estimate of solar and interplanetary (IP) parameters associated with halo CMEs is, therefore, an important step towards achieving an accurate prediction of GMS.

Despite clear advancement in CME monitoring, it is still difficult to predict the onset and evolution of solar transient events (i.e., CMEs, solar flares) due to the complex nature of the solar activity which generates them. In the last decades, advances towards developping theoretical models of the solar-terrestrial environment have been achieved. However, being able to apply those models for space weather prediction is still subject of intense research. The current prediction of space weather (i.e., GMS) is dominated by empirical methods, relying mostly on the observable storm precursors of the Sun and in the IP medium (Fox and Murdin, 2001).

The study described in this thesis focuses on two major topics, namely: (a) An investigation and characterisation of the probable solar and IP properties associated with GMS during the first 11 years (1996 - 2006) of SC 23, and (b) the development of neural network based models for the prediction of the occurrence and strength of GMS. The reasons for choosing the period 1996-2006 for the investigation are the following:

- This period corresponds to the period during which spacecrafts (e.g., SOHO and the Advanced Composite Explorer [ACE]) consistently monitored transient phenomena occurring on the Sun as well as in situ measurements of the associated IP phenomena.
- The 11-year period (1996-2006) represents an average mean of a solar activity cycle during which solar storm events and correlated magnetic storms can be analysed. In fact, even though SC 23 had a prolonged solar minimum lasting up to 2009, the three year period after 2006 was quiet in terms of solar storms and magnetic activity and hence of not much interest for this analysis.

1.3 Research motivation

More than four years after experiencing an unusually long period of solar activity minimum, the Sun has entered a new cycle of activity expected to reach its maximum during the next two years (2012-2013) and during which large solar storms and subsequent GMS events are expected. Indeed, on 15 February 2011, a powerful solar flare (class $X2.0$) of the new SC 24 was observed by NASA's Solar Dynamic Observatory (SDO). It was the first major X-ray flare event since December 2006 (http://www.spaceweather.com). Therefore, the upcoming period of solar activity maximum offers an opportunity to study the solar-terrestrial interactions by investigating the short-term as well as long-term evolution of solar storms and their relationship to GMS events.

GMS represent an important link in the chain of solar-terrestrial relations (Prölss, 2004). A wider interest in magnetic storms in recent years is due to their severe impact on technological systems on which modern society relies. Therefore, the predictability of the occurrence and magnitude of GMS, as well as a quantitative estimate of the associated solar and IP properties is of practical importance for space weather prediction with the purpose of minimising their effects. It is anticipated that this research will contribute globally to the advancement of space weather predictions. Locally, the Space Science Directorate of the South African National Space Agency [SANSA], based in Hermanus, South Africa, was given a mandate by the International Space Environmental Service (ISES) to become the Regional Warning Centre for Africa (RWC). The results presented in this thesis would be among the products and services to be provided to the relevant scientific community in general, and to the users of the SANSA-RWC in particular.

1.4 Research objectives

The following are the main objectives addressed in this thesis:

- Identification of GMS events in SC 23 (1996 - 2006).
- Investigation, characterisation and statistical analysis of the probable solar sources (i.e. halo CMEs) and IP properties associated with GMS events in SC 23.
- The development of NN-based models to predict the occurrence and strength of GMS using both solar and IP input parameters.

1.5 Organisation of the content

This thesis comprises 7 chapters. Chapter 1 is an introductory chapter, describing the domain of research, the research problem, the motivation and objectives. In Chapter 2, basic concepts of the solar-terrestrial environment are provided. The Sun's magnetic activity is the key to all solar activity. An overview of the Sun's magnetism is given, outlining the basics of magnetohydrodynamics (MHD) as pertaining to solar magnetism. Since solar dynamics are complex and varying, the description of phenomena linked to solar activity is limited to the description of CME eruptions and their evolution in the IP medium. This chapter also describes the basics of the coupling between the solar wind and the magnetosphere, leading to GMS phenomena and related space weather effects. Chapter 3 provides an introduction to the basic principles of artificial neural network (ANN) prediction techniques and their application to space weather prediction.

Chapter 4 gives a detailed quantitative analysis of the main sources of GMS during SC 23. Chapter 5 describes the developed NN model which was used to predict the probability occurrence of GMS using a combination of solar and IP parameters as inputs. Chapter 6 describes another NN-based model for the prediction of the magnitude of a storm by using solar wind parameters as inputs. Concluding remarks including conclusions and suggestions for future work are provided in Chapter 7.

Chapter 2

The solar-terrestrial environment: basic concepts

The Sun's dynamic processes, including prominences, flares, and CMEs, produce energetic particles and electromagnetic fields that, by interaction with the Earth's atmosphere and magnetic field, may lead to GMS and subsequent adverse space weather effects. The aim of this chapter is to provide an overview of some of the fundamental concepts and principles of solar-terrestrial interaction. The Sun interacts with the Earth's environment via the solar wind (SW). The structure of SW and its constituents undergo disturbances following interaction with CMEs, leading to large scale ionosphere-magnetospheric disturbances. In this chapter, all phenomena leading to GMS events are briefly described.

2.1 The Sun

The Sun is an ordinary star, the nearest to us and the source of heat which sustains life on Earth and controls both terrestrial and space weather. The following are the main physical characteristics of the Sun, as adapted from Kivelson and Russell (1995) and Lang (2001).

- Age = 4.5×10^9 years
- Mass, $M_\odot = 1.99 \times 10^{30} kg$ (332 946 times Earth's mass)
- Principal chemical constituents = hydrogen (92.1%), helium (7.8%).
- Volume = 1.412×10^{27} m^3 (or 1.3 million Earths)
- Radius, $R_\odot$ = 696 000 km (109 Earth radii)
- Mean distance from Earth ($1AU = 1.5 \times 10^8$ km)
- Emitted radiation (luminosity)= 3.86×10^{26} W ($3.86 \times 10^{33} ergs^{-1}$)

- Equatorial rotational period = 26.8 days (30.8 days at 60^0 latitude).
- Surface temperature (photosphere) = 5785 K (= 1.56×10^7 K in Sun's centre and about 2×10^6 K in the corona)
- Density (centre)= 1.513×10^5 kg m^{-3}
- Pressure (centre)= 2.334×10^{16} Pa
- Magnetic field (sunspots)= $0.1 - 0.4T = 1 \times 10^3 - 4 \times 10^3$ G

The Sun is a giant mass of incandescent gas. Starting from the interior, the Sun's atmosphere consists of three layers: the photosphere, the chromosphere and the corona. The photosphere is the lowest and densest level of the solar atmosphere and is the only part of the Sun that is visible to the naked eye. However the apparent surface of the Sun is actually an illusion caused by a gas of extremely high opacity; in reality, the Sun does not have a solid surface. The entire solar atmosphere contains magnetic fields which are generated in the solar interior, while the tachocline ($\sim 0.7R_\odot$ at the base of the convection zone) plays an important role in the observed dynamic behaviour (Miesch, 2005).

The Sun's magnetic field is due to the movement of its plasma. As the solar plasma moves, any magnetic field line is pulled along with it, i.e. the magnetic field lines are *frozen* into the solar material. A complete understanding of solar dynamics requires an understanding of solar magnetism. In fact, it is well known that all solar activity is a consequence of the existence of the magnetic field on the Sun (Stix, 1989).

2.1.1 Sunspots and solar activity cycle

Detailed observations indicate that the photosphere is often pitted with dark spots called *sunspots*, the largest being much bigger than the size of the Earth. *Sunspots* are the common indicator of the Sun's activity level and correspond to the solar region of intensified magnetic field. Magnetic fields in sunspots were first measured in 1908 by the American astronomer George Ellery Hale, who suggested a cyclical sunspot period of 22 years, covering two polar reversals of the solar magnetic dipole. The magnetic field strength in sunspots is about 0.3 T, which is thousands of times stronger than the Earth's magnetic field (3×10^{-5} T) at the equator.

According to Hathaway et al. (1999), the cyclic magnetic behaviour observed through sunspots can be explained by the Sun's differential rotation, meridional circulation, and large-scale convective motions. A qualitative model to explain the dynamics of solar magnetism characterised by sunspot activity was first proposed by Babcok (1961):

- The 22-year cycle begins with a well-established dipole field component aligned along the solar rotational axis.
- Due to the solar differential rotation (the solar rotation at the equator is 20 percent faster than it is at the poles), the magnetic field lines are wrapped.
- After many rotations, the field lines are highly twisted and bundled resulting in the increase of the magnetic field intensity. The resulting buoyancy lifts the bundle to the solar surface and forms a bipolar field that appears as two spots.

During one 11-year sunspot cycle, the magnetic polarity of all the leading (westernmost) spots of the bipolar field in the northen hemisphere is the same, and is opposite to that of leading spots in the southern hemisphere. The magnetic polarity of the leading spots reverses in each hemisphere at the beginning of the next 11-year sunspot cycle as does the dipolar magnetic field at the solar poles. For the next 11 years in the new cycle, all polarities will be exchanged such that a full magnetic solar cycle of the Sun takes an average of 22 years (Lang, 2001).

The positions of sunspots and their associated active regions vary during an 11-year cycle. The first spots of each cycle appear at a latitude of about $30^o - 35^o$ in both hemispheres. As the cycle advances, the zone of sunspot occurrence migrates towards lower latitudes, and the last spots of a cycle are normally within $\pm 10^o$ of the solar equator. More details on the basics of solar cyclic magnetism and related sunspot dynamics can be found in various textbooks including Stix (1989), Kivelson and Russell (1995) and Lang (2001). As with sunspots, other forms of solar activity such as solar flares and CMEs also occur with a frequency that varies in step with the 11-year sunspot cycle.

Near the maximum of the 11-year SC (when there are more spots on the Sun's surface), solar storms such as powerful flares and fast and energetic CMEs are most frequent, producing disturbances of the Earth's magnetosphere. When geomagnetic measurements are averaged on a global scale, it is observed that the largest storms vary in step with the 11-year sunspot cycle. In the 1930s, Chapman and Bartels (1940) showed the existence of a certain relationship between geomagnetic disturbances and solar activity. Figure 2.1 shows this relationship for the last hundred years using yearly averaged sunspot numbers (SSN) and the geomagnetic *aa* index.

2.1.2 Solar magnetism

Solar activity is a consequence of the existence of the magnetic field on the Sun (Stix, 1989). The Sun's magnetic field has active effects on the plasma. It may exert a force on the plasma

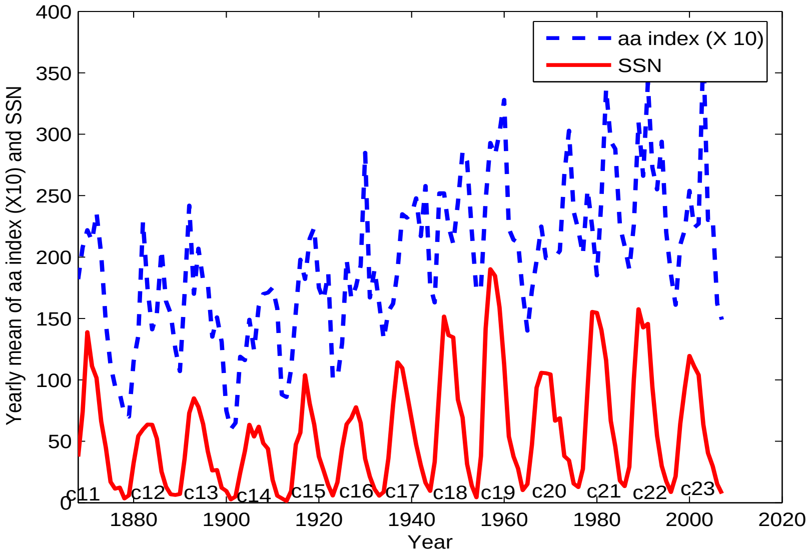


Figure 2.1: Illustration of the relationship between the solar activity cycle and geomagnetic activity using the *aa* index. This figure clearly shows that geomagnetic disturbances and solar activity correlate.

and creates structures such as sunspots. It may also store energy for a while and suddenly release it (Kivelson and Russell, 1995). Hence, an understanding of solar dynamics requires an understanding of solar magnetism. MHD principles are used to model the interaction of plasma and the magnetic field, in which the plasma is treated as a continuous medium.

2.1.3 Solar magnetohydrodynamics: an overview

MHD equations are derived from Maxwell's equations of electromagnetism and the equations of fluid mechanics (hydrodynamics). Maxwell's equations describe how current and charge density affect the magnetic and electric fields. MHD equations, therefore, incorporate familiar mechanical laws, but they also account for electromagnetic properties. The plasma is treated as an electrically neutral fluid made of two species: ions and electrons. Maxwell's equations incorporate Ampere's law (equation 2.1) which relates the magnetic field to the net current $\mathbf{j}$:

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{j} + \varepsilon_0 \frac{\partial E}{\partial t} \right), \tag{2.1}$$

the vanishing of the divergence of the magnetic field,

$$\nabla \cdot \mathbf{B} = 0,$$

Faraday's law (equation 2.2)

$$\frac{\partial B}{\partial t} = -\nabla \times \mathbf{E}, \tag{2.2}$$

and Poisson's equation (equation 2.3)

$$\nabla \cdot E = \rho q / \varepsilon_0. \tag{2.3}$$

On the other hand, Ohm's law defines the relationship between the net current and the magnetic field

$$\mathbf{j} = \sigma \left(\mathbf{E} + \mathbf{u} \times \mathbf{B} \right), \tag{2.4}$$

where σ represents the electrical conductivity and $\mathbf{u}$ the plasma velocity.
For a fluid made of electrons and protons, fluid mechanics equations are the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0, \tag{2.5}$$

where ρ is the total mass density and $\mathbf{u}$ the centre of mass velocity. Considering that $m_e << m_p$, the momentum equation can be written as

$$\rho \left(\frac{\partial u}{\partial t} + u \cdot \nabla u \right) = -\nabla p + \mathbf{j} \times \mathbf{B} + \rho F_g / m_p. \tag{2.6}$$

In the case where $n_e = n_p = n$ and both electrons and protons have the same temperature, the pressure in equation 2.6 is related to the temperature by the ideal-gas law

$$P = 2nkT. \tag{2.7}$$

For a plasma with speed u much slower than the speed of light, the displacement current, $\varepsilon_0 \partial E / \partial t$ is neglected and hence, Ampere's law becomes

$$\mathbf{j} = \nabla \times \frac{\mathbf{B}}{\mu_0},$$

With the magnitude $\mathbf{j} \sim \frac{\mathbf{B}}{\mu_0 L}$ where L is the scale length for magnetic variation Ohm's law can be written as

$$\mathbf{E} = -\mathbf{u} \times B + \frac{\mathbf{j}}{\sigma}.$$

By taking the curl of the above equation and using equation 2.2, the first reduced MHD equation, the *induction equation*, is obtained:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B}, \tag{2.8}$$

where $\eta = 1/(\mu_0\sigma)$ is the magnetic diffusibility, which is assumed to be uniform in this case. The ratio of the first term to the second term on the right of equation 2.8 gives the magnetic Reynolds number:

$$R_m = \frac{uL}{\eta} = \mu_0\sigma uL,$$

which is enormous ($10^6 - 10^{12}$) for solar phenomena. Here L is the characteristic scale length for changes in the field and the flow. The magnetic field is thus frozen to the plasma, and the electric field does not drive the current, but is simply $\mathbf{E} = -\mathbf{u} \times \mathbf{B}$.

The second main MHD equation is the *momentum equation*:

$$\rho\frac{d\mathbf{u}}{dt} = -\nabla p + \mathbf{j} \times \mathbf{B} + \rho\mathbf{g}, \tag{2.9}$$

where the first two terms on the right-hand side of equation 2.9 represent the effects of thermal pressure and curvature. In the case where the plasma beta $\beta = \dfrac{2\mu p}{B^2}$ is small, the magnetic forces usually dominate the pressure forces like in active regions of the solar outer atmosphere. Equilibria of solar structures such as sunspots and prominences are described by the force balance

$$\mathbf{j} \times \mathbf{B} - \nabla p + \rho\mathbf{g} = 0. \tag{2.10}$$

There is no contribution of the magnetic force along the magnetic field and there is a hydrostatic equilibrium balance between pressure gradients and gravity. In the Sun's active regions, the magnetic field dominates and equation 2.10 reduces to

$$\mathbf{j} \times \mathbf{B} = 0. \tag{2.11}$$

There is no force to compensate the magnetic force and the fields are said to be *force-free.* Here, $\mathbf{j} = \nabla \times \frac{\mathbf{B}}{\mu}$ and $\nabla \cdot \mathbf{B} = 0$. Electric currents are parallel to the magnetic field and so $\nabla \times \mathbf{B} = \alpha\mathbf{B}$ where α is a constant, a scalar function of position. It is believed that solar transient phenomena such as solar flares result from eruptive MHD instability. To describe how flares may start, it is suggested to solve the MHD equations:

$$\mathbf{j} \times \mathbf{B} = 0,$$

and

$$\frac{\partial\mathbf{B}}{\partial t} = \nabla \times (u \times \mathbf{B}).$$

for an evolution through a series of force-free equilibria due to photospheric foot point motions (Priest, 2007).

As indicated by Manchester IV et al. (2004), solutions to MHD equations are the only self-consistent mathematical descriptions of the Sun-Earth space environment. The above overview of the basic MHD concepts as applied to solar phenomena were retrieved from Kivelson and Russell (1995), Priest (1995) and Priest (2007) in which more details can be found. It must, however, be mentioned that the Sun remains intrinsically an object with such a rich variety of MHD phenomena yet still to be understood. The current models and theories describing solar explosions are described in the next section with special focus on CMEs.

2.2 Explosive solar phenomena

Solar flares and CMEs are the most explosive phenomena in the solar system. During these explosions, the emitted radiation and particles can lead to disturbances in the Earth-space environment. Both flares and CMEs are believed to be the result of an explosive release of energy from the active regions in the solar outer atmosphere.

A solar flare is defined as the transient brightening on the solar surface observed in the H_α emission line ($\lambda = 656.3$ nm). This explosion of energy radiates electromagnetic emissions from γ and X-rays to radio waves. In the emission line H_α, flares normally appear as two ribbon-like bright areas, known as a two-ribbon flare. X-ray flares are classified according to the order of magnitude of the Geostationary Operational Environmental Satellite (GOES) X-ray (0.1- 0.8 nm) peak burst intensity, I (Wm^{-2}), measured on the Earth. The following is an X-ray flares classification with corresponding energy range in Wm^{-2} (http://www.spaceweather.com/glossary/flareclasses.html):

- Class B with $I < 10^{-6}$
- Class C with $10^{-6} \leq I \leq 10^{-5}$
- Class M with $10^{-5} \leq I \leq 10^{-4}$
- Class X with $I \geq 10^{-4}$

CMEs are defined as large-scale expulsions of plasma and magnetic field from the lower corona into the IP medium, events during which about $10^{15} - 10^{16}$ g are ejected into the IP space with kinetic energy of the order of 10^{31} ergs (Manchester IV et al., 2004). There is a close link between CMEs and solar prominences. Prominences arise as arches extending into the coronal darker region and appearing as long dark filaments. As indicated by Forbes (2000), CMEs, prominence eruptions and large flares are closely related and may be different manifestations of a single physical process.

2.2.1 Flares and CMEs relationships

The relationship between CMEs and solar flares is an issue that has been a point of contention among scientists. For sometime, it was thought that CMEs were a simple visualisation of disturbances produced by solar flares (Ondoh, 2001), but new observational techniques later revealed that CMEs were not necessarily connected to flare phenomena. While CMEs occur at a wider range of solar latitudes, solar flares tend to be restricted to lower latitudes (Wallace, 1997).

There exists a physical connection between CMEs and solar flares, especially between fast CMEs and major flares. Dynamic flares occur as a consequence of CME eruptions initially driven by ideal MHD processes. According to Vršnak (2008), CME dynamics is closely related to the energy release in the associated flares. Messerotti et al. (2009) suggest that the violent launch of a CME causes magnetic fields of opposite polarity to reconnect, and this quickly leads to sporadic electromagnetic radiation in the form of flares.

The next section provides a detailed description of CMEs. The practical interest in CMEs follows large disturbances they produce in the SW that are the primary causes of GMS (Gosling, 1993).

2.2.2 CME characteristics

CMEs were first discovered on 14 December 1971 and later on 8 February 1972 using the white-light coronagraph aboard NASA's seventh Orbiting Solar Observatory (OSO 7) (Lang, 2001). However, it was in 1973 that the Skylab clearly identified a CME. Since then various missions, including the Solar Maximum Mission (SMM), the Yohkoh, the Solar and Heliospheric Observatory (SOHO) as well as the recent Transition Region and Coronal Explorer (TRACE) and the Solar Terrestrial Environment Observatory (STEREO), have allowed reliable observations and provided more information and knowledge about the morphology and properties of CMEs (Mikić and Lee, 2006). The SOHO spacecraft is a NASA/ESA joint project. The LASCO instrument aboard SOHO has 3 coronagraphs (C_1, C_2 and C_3) that have operated since 1996 and detected more than 10 000 CMEs during SC 23 (Gopalswamy, 2009a).

CMEs are detected in visible-light observations of the solar corona from spacecraft. The coronagraphs image the CMEs in Thomson-scattered photospheric light by blocking the direct sunlight with an occulting disk (Gopalswamy, 2009a). Mass ejections are seen as bright moving loop-like features in the corona, blasting out from the edge of the occulted photosphere as illustrated in Figure 2.2. CMEs originate from closed magnetic field regions (e.g active regions) on the Sun which may or may not correspond to sunspots. It is believed that CME eruptions are associated with a large-scale reconfiguration of the coronal magnetic field that contributes to the magnetic polarity reversal over the SC (Low, 2001).

A typical CME is characterised by a three-part structure, namely: the leading bright edge (or frontal rim), the void (or dark cavity) and the bright core also known as a prominence (Illing and Hundhausen, 1986). This prominence-corona structure hints that the CME morphology has its roots in the pre-eruption magnetic field configuration. It is believed that the frontal rim indicates the leading edge of the erupting arcade whereas the cavity was a large magnetic flux rope with low density plasma, hence appearing as dark regions in the coronagraphs (Low, 1996). In situ measurements of interplanetary CMEs (ICMEs) have confirmed the flux-rope characteristics of CMEs in the IP medium known as magnetic clouds (Lepping et al., 1990).

The CME images in Figure 2.2 were adapted from Gopalswamy (2009a) and illustrate the CME morphology using two CMEs. The left images are direct images with SOHO's Extreme-Ultraviolet Imaging Telescope (EIT) superimposed on the LASCO images. The top panels of this figure relate to the CME of 20 December 2001 which originated close to the limb, while the bottom panels relate to the CME of 18 November 2003 which originated close to the disk center. The CME of 18 November 2003 (bottom panels) shows a single structure only. The right images are difference images where the previous frames have been subtracted in order to see the changes taking place in the corona. Difference images show the perturbations around the CME, where dark regions correspond to material depletion, which is an indication of the displacement of a structure between the two frames used.

The CME occurrence rate is SC phase dependent and is roughly about 0.5 per day at solar minimum and can be > 6 per day during a solar maximum period (Gopalswamy, 2009a). During solar maximum CMEs occur over a wide range of latitudes, but are commonly found near the equator at solar minimum (Cyr et al., 2000). Other basic properties of CMEs are their speed, apparent width and their acceleration. The speed measured in the sky plane varies from ~ 20 km/s to > 3000 km/s. The highest CME speed (~ 3387 km/s) in SC 23 was recorded on November 10, 2004 (Gopalswamy, 2009a). The CME width (W) ranges from $< 5^0$ to 360^0. CMEs that appear to surround the occulting disk of the observing coronagraphs are known as halo CMEs (Howard et al., 1982) and often signal a future terrestrial hit.

However, a halo appearance of CMEs itself does not directly indicate whether the CME is directed towards the Earth or moving away from it. Halo CMEs are more energetic than other CMEs with an average speed of ~ 1000 km/s, and are very important in space weather research (Gopalswamy, 2009a). Full halo CMEs have an apparent width of $W = 360^0$, while partial halo CMEs have a width of $120^0 < W < 360^0$. It should be mentioned that the true width of halo CMEs is not well known. Table 2.1 provides a summary of CME properties as adapted from Gopalswamy (2009a).

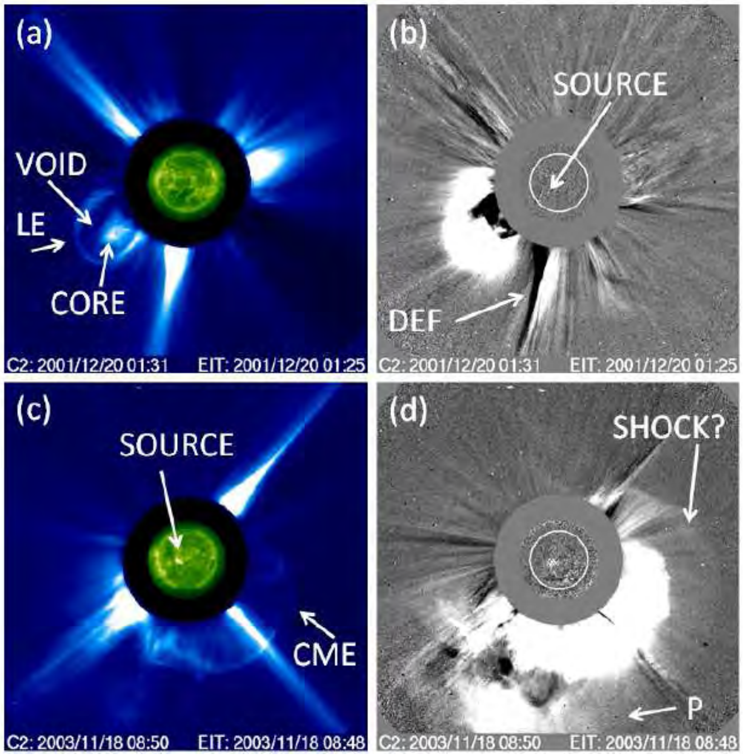


Figure 2.2: The compact bright features of the EIT images (SOURCE) indicate the location of the eruptions. The top images, (a) and (b), clearly show the CME three-part structure: the CME leading edge (LE) or frontal structure is the outermost feature followed by a void (cavity) and the prominence core. The streamer deflection (DEF) can be seen (b). Fig. (d) shows a shock and associated sheath surrounding the CME. A previous CME (P) is in progress when the second CME was ejected. Images and description adapted from Gopalswamy (2009a).

Table 2.1: A summary of the main properties of CMEs

Property	range	average
Speed	~ 20 km/s to > 3000 km/s	~ 470 km/s
Mass	$\sim 10^{12}$ g to $> 10^{16}$ g	$\sim 4 \times 10^{14}$ g
Kinetic energy	$\sim 10^{27}$ erg to 10^{33} erg	$\sim 5 \times 10^{29}$ erg
Angular width	$< 5^0$ to 360^0	$\sim 54^0$
Daily occurrence rate	< 0.5 to > 6 CMEs	Solar min - solar max

2.2.3 CME generation: an overview of theories and models

A number of recent studies (e.g Forbes, 2000; Forbes et al., 2006; Mikić and Lee, 2006; Vršnak, 2008) provide an overview of the current understanding of CME generation. The common understanding is that the coronal magnetic field plays a dominant role in CME dynamics and most of the theoretical CME concepts are based on the MHD equations and their simplifications (e.g $\beta = 0$ and force-free). As indicated by Forbes et al. (2006), much of the CME modelling effort has been dominated by numerical methods due to the complexity of the MHD equations governing the CME dynamics. In general, theorists have constructed their models with a focus on force-free models of the corona in which all forces other than magnetic forces are neglected (Mikić and Lee, 2006). In such models, the equilibrium force-balance condition is simplified to equation 2.11 (which is itself a difficult nonlinear problem (Mikić and Lee, 2006)), where $\mathbf{j} = c\nabla \times \mathbf{B}/4\pi$ is the electric current density and $\mathbf{B}$, the magnetic field intensity. In equilibrium, neglecting gravity and in the presence of a strong magnetic field, the momentum equation expresses a balance between the tension in the magnetic field lines that are line-tied in the photosphere and magnetic and thermal pressure:

$$\mathbf{B} \cdot \nabla \mathbf{B} = \nabla(4\pi p + B^2/2). \tag{2.12}$$

Eruption involves forcing the system to evolve into a state in which this delicate balance can no longer be maintained (Mikić and Lee, 2006).

Although the most generally accepted explanation for the cause of CMEs is that they are produced by a loss of stability or equilibrium of the coronal magnetic field (Low, 1996), there is still however no consensus as to what mechanism causes the loss of equilibrium (Forbes, 2000; Low, 2001). Forbes (2000) suggests a model whereby a continual emergency of new flux from the Sun's convection zone, combined with the random motions of the footpoints of closed coronal field lines, cause stresses to build up in the coronal field. Eventually, these stresses exceed a threshold beyond which a stable equilibrium cannot be maintained and the field erupts. This mechanism drives the release of the magnetic energy stored in the fields associated with coronal currents, and hence, models based on this mechanism are described as *storage and release models* (Forbes, 2000).

According to Forbes (2000), field lines in the photosphere (which is an excellent conductor) are frozen into plasma, hence a sudden injection of flux from the convection zone into the corona must necessarily move the photospheric plasma. In the corona the magnetic energy density is much larger than the thermal and gravitational energy density, hence, the currents associated with the magnetic energy stored there must either be force-free (i.e. the current flow along the magnetic field direction) [see Figure 2.3 (a)] or confined to the current sheets

[Figure 2.3 (b)]. Storage models for flares and CMEs are therefore generally divided into those based on force-free currents and based on current sheets.

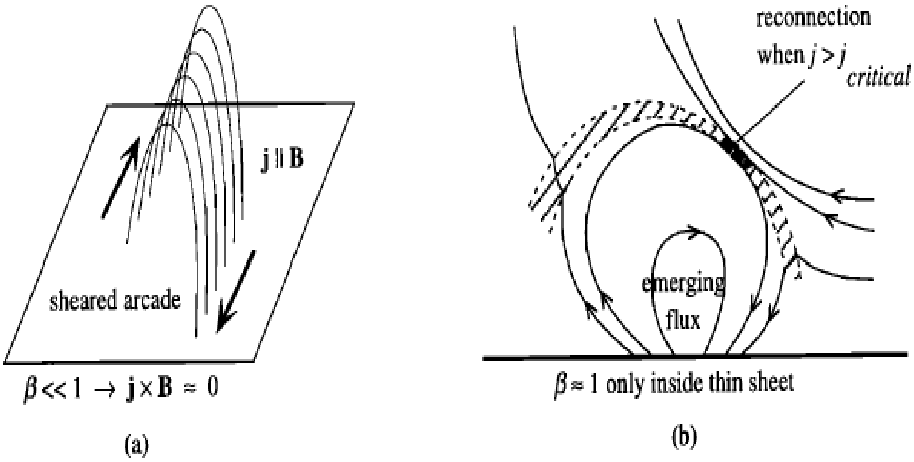


Figure 2.3: Schematic illustration of two types of models that use magnetic energy to power a flare or a CME: (a) Magnetic energy is stored in the corona in the form of field-aligned currents that eventually become unstable and (force-free currents), and (b) Magnetic energy is stored in the corona in the form of a thin current sheet that is suddenly dissipated when a micro-instability is triggered within the sheet. Sketch by Forbes (2000).

In the CME initiation models as presented by Mikić and Lee (2006), storage and release models were classified into *flux cancellation* and *breakout* models. The flux cancellation model refers to the mutual disappearance of magnetic fields of opposite polarity at the neutral line separating them. The breakout model by Antiochos et al. (1999) refers to magnetic reconnection, where the magnetic topology is due to multipolar flux distribution at the photosphere and contains at least one null point where reconnection can occur. Like the flux cancellation model, breakout requires strongly sheared fields near the neutral line, as observed in filament channels.

In summary, the CME generation phenomenon is the result of the interplay between the coupling of the solar differential rotation, convective motions and magnetic field MHD dynamo processes on different scales. The shearing and twisting motions induce electric currents and store free energy into current-carrying magnetic field structures. A part of this energy is transferred through the solar surface into the corona, where it is partly spent on coronal heating and partly released in eruptive processes, taking the form of CMEs and /or solar flares (Priest, 1982; Vršnak, 2008). Currently, the general consensus is that the CME mechanism involves the release of free magnetic energy associated with currents flowing in the corona. However, there is no consensus about the mechanism by which this energy is released (Forbes, 2000).

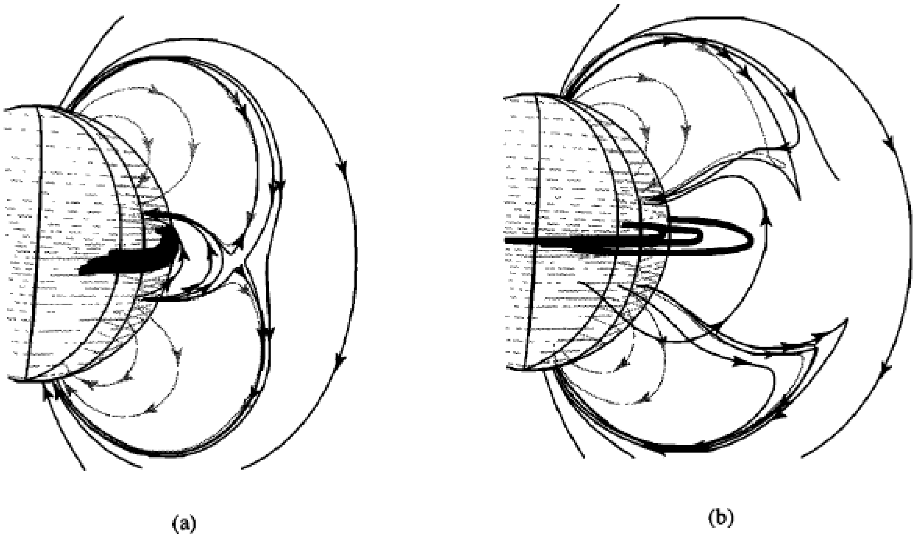


Figure 2.4: (a) Reconnection of the field lines and (b) opening up of the sheared field lines in the model by Antiochos et al. (1999).

After CME initiation, its dynamics involve several factors including acceleration, expansion, drag and distortion. While acceleration and expansion is an integral part of the initiation process, drag and distortion result from the interaction of the CME with the ambient SW, CIRs and other CMEs (Forbes et al., 2006). The next section discusses the behaviour of CMEs in the IP medium.

2.2.4 Interplanetary coronal mass ejection (ICMEs)

After a CME eruption on the Sun, its evolution is dominated mainly by its interaction with the SW and generally, it could take one to five days to reach the Earth's magnetosphere. CMEs carry into the heliosphere a large amount of coronal magnetic fields, forming the subset of CMEs in the IP medium known as the interplanetary counterpart of CMEs or ICMEs. These structures can be detected by remote sensing and in situ spacecraft observations. During the last two decades, the detection of CME signatures in the IP space has been achieved by a number of spacecraft, including NASA's WIND SW monitor and the interplanetary monitor (IMP) series. Launched in 1997, the NASA Advanced Composite Explorer (ACE) spacecraft has been at the Lagrangian point (L1), providing a variety of information about the SW status (Cargill and Harra, 2007). In addition, NASA's STEREO mission, which was launched in October 2006, has the ability to observe the Sun from both the front and the back. The combined views are currently providing a 3-D view of the Sun, tracking a CME as it erupts from the Sun and moves into the IP medium.

CMEs in the IP medium still carry magnetic field patterns from their parent solar filament. These are closed magnetic field structures with features which differ from the ambient medium. Ejecta, magnetic clouds (MCs) and flux ropes are the other names which are given to ICME structures near the Earth at about 1 AU (see e.g. Gopalswamy, 2009a). Figure 2.5 is a schematic illustration of an ICME. A variety of in situ signatures are interpreted to indicate the passage of CME material that passes a spacecraft. These include IP shocks, plasma density changes, depressed proton temperatures and magnetic field strength and topology. The structure and composition of a MC with respect to the normal SW wind provide a clue to the MCs solar origin. These structures are characterised by a long duration structure, with a magnetic field strength ranging between 20-30 nT and constitute about 40% of ICMEs. MCs are very well organised structures and are considered the residue of a large solar loop-like structure linked to the prominence cavity (see description of the three part structure of a CME, section 2.2.2, paragraph 3). Figure 2.6 is a schematic diagram of the MC. It resembles a bundle of twisted magnetic field lines. MCs are the main cause of geomagnetic disturbances, and their geoeffectiveness depends on whether the field at the leading edge has a strong northward or southward component (Kallenrode, 1998).

As a consequence of interaction between CMEs and the SW, slow CMEs may accelerate towards SW speed, while fast CMEs may decelerate toward the speed of the SW. Faster CMEs ($>$ 500 Km/s) generally drive shocks. CME-driven shocks are generated when the speed of the CME in the frame moving with the SW is faster than the local magnetosonic speed. CME-driven shocks are very important in space weather since they are the main accelerator of the solar energetic particles (SEPs) emanating from the flare reconnection regions (Gopalswamy, 2006a). CME-driven shocks are inferred from the type II radiobursts detected within $1R_{\odot}$ from the solar surface. Type II radiobursts at metre wavelengths are observed by ground-based radiotelescopes and constitute a primary means of tracking CME-driven shocks in the IP medium. IP shocks often indicate the presence of ICMEs although this is not always the case because the flanks of shocks may extend well beyond the associated ICMEs (Richardson and Cane, 1993). Therefore the ejecta (or ICME) may not hit the Earth while associated shocks may drive substantial GMS (Gosling et al., 1991). On the other hand, it should be noted that not all full or partial halo CMEs are followed by ICMEs near Earth nor are all observed ICMEs associated with halo CMEs. In their analysis, Cane and Richardson (2003), indicated that a significant fraction of ICMEs detected at Earth were without probable association with halo CME eruption at the Sun. The interest in the study of ICME structures arises from the fact that they are the main source of geomagnetic disturbances. As they approach the Earth, shocks and ICME structures couple to the magnetosphere to drive moderate to major storms (Webb, 2000).

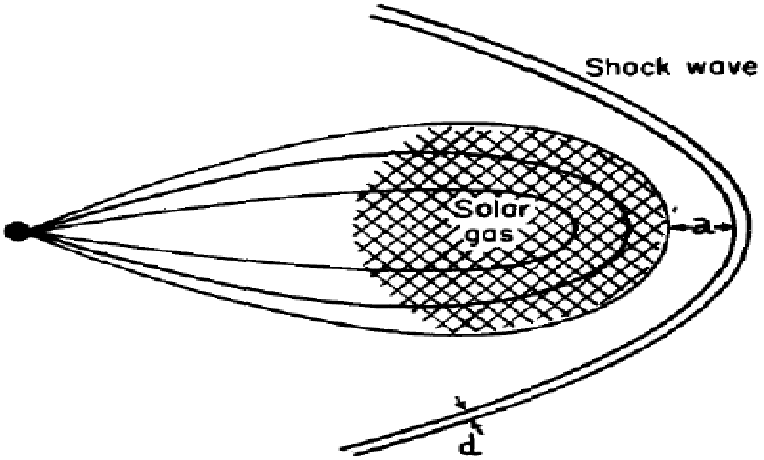


Figure 2.5: A schematic representation of the spatial structure of an ICME. The sketch shows an enormous size of the magnetised plasma cloud connected to the Sun and driving a shock ahead of it. Sketch by Gold (1962) and reprinted from Gopalswamy (2009a).

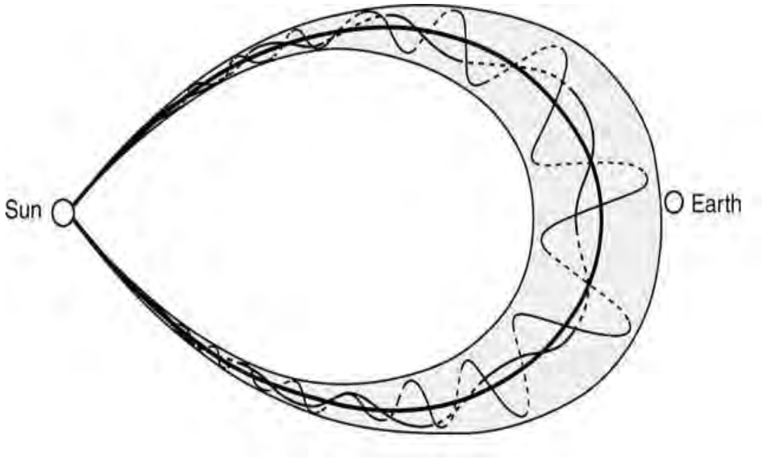


Figure 2.6: Proposed topology of a magnetic cloud in the interplanetary space. Sketch adapted from Burlaga (1991).

2.3 The solar wind

Rather than being an empty space, the IP medium is filled with particles and fields constituted mainly by the SW and the embedded interplanetary magnetic field (IMF). The first indications that the solar atmosphere in the IP space is a dynamic phenomena came from observations of comet tails in the 1940s and 1950s (Prölss, 2004). To explain the structures and topology of comet tails, Biermann (1951) postulated the existence of a plasma flow that is continually emitted from the Sun with variable flow velocity. However, it was Parker (1958) who first developed a successful SW model by predicting a continual high speed SW based upon hydrodynamic theory. Parker's SW model was confirmed first by *in situ* observations by the Mariner 2 spacecraft which collected the first SW data in 1962. Later the SW was probed from distances around 1 AU by Helios 1 and Helios 2 during the period 1974-1976 (Stix, 1989). In this thesis only the general basic properties of the SW are described. Details of the SW formation in the corona and its evolution can be found, among other literature in Hundhausen (1995).

The SW is a stream of charged particles resulting from the gradual outward acceleration of the solar corona to form the supersonic wind. Particles escape the Sun's gravity due to their high kinetic energy and the high temperature of the corona. Currently, the *in situ* observations by spacecraft provide us with reliable knowledge about the SW. The SW is composed essentially of protons, electrons and α particles (H^{++}) in low quantity. The SW is supersonic with a mean speed of about 400 km/s, travelling from the Sun to Earth within roughly 4 days. Table 2.2 gives a summary of the SW mean properties at the Earth's orbit, i.e. in the ecliptic plane at a heliocentric distance of 1 AU. While the SW velocity and density are highly fluctuating, the particle flux is relatively constant. It is possible to estimate the Sun's mass loss rate by knowing the particle flux using the following relation (Prölss, 2004):

Table 2.2: Mean properties of the SW, adapted from Prölss (2004)

Property		**$\simeq$ value at 1 AU**
Composition		$\simeq 96\% H^+$, $4\% \, (0-20\%) \, He^{++}$, e^-
Density	$(n_p \simeq n_e)$	$\simeq 6 \, (0.1 - 100) \, cm^{-3}$
Temperature	$(T_p \simeq T_e)$	$\simeq 10^5 \, (3500 - 5 \cdot 10^5) \, K$
Velocity	$(u_p \simeq u_e)$	$\simeq 470 \, (170 - 2000) \, km/s$
Proton flux	$(n_p u_3)$	$\simeq 3 \times 10^{12} m^{-2} s^{-1}$
Energy flux	$(n_p m_h u^3/2)$	$\simeq 0.5 mW/m^2$
Momentum flux	$(n_p m_h u^2)$	$\simeq 2 \times 10^{-9} N/m^2$

$$dMs/dt \simeq n_p u m_h 4\pi \, (1AU)^2 > 10^9 kg/s \tag{2.13}$$

This indicates that the Sun loses more than a million tons of mass each second via the SW. However, this loss is unnoticeable given the Sun's total mass ($M_\odot \simeq 2 \times 10^{30}$ kg).
Two types of SW plasma flow have been observed: the fast and the slow wind (Schwenn, 1990). Data from radio scintillation observations show that the two types of SW originate from different solar latitudes (Prölss, 2004; Lang, 2001). At mid-to high heliographic latitudes the SW flows with a very high speed (750-800 km/s). Fast SW is more stable with low density, compared to the slow ($\simeq$ 400 Km/s), highly variable, turbulent and denser SW originating from lower latitudes.

Near the minimum of the 11-year SC, the slow component of the SW is essentially confined to the low latitudes within the equatorial belt ($30 - 35^0$) or streamer belt, while the fast SW seems to originate from the higher heliographic latitudes. The source of fast SW is believed to be the coronal holes, the dark parts of the corona dominated by open field lines. At solar activity maximum the slow SW seems to originate from all over the Sun. The slow SW is characterised by a complex structure, often containing large scale-structures such as MCs and shocks (Kallenrode, 1998).

Figure 2.7 shows the latitudinal profile of the SW velocity recorded on the ULYSSES spacecraft along its polar orbit around the Sun (McComas et al., 1998). These velocity data obtained between 1991 and 1996 show that a first wind escapes from the polar regions where coronal holes are found, while a slow wind is associated with the Sun's equatorial region that contain coronal streamers. The Ulysses solar probe was launched by NASA in 1990 in an orbit that allows it to observe over the solar poles and completed its first solar orbit between 1992-1997 during the period of solar minimum.

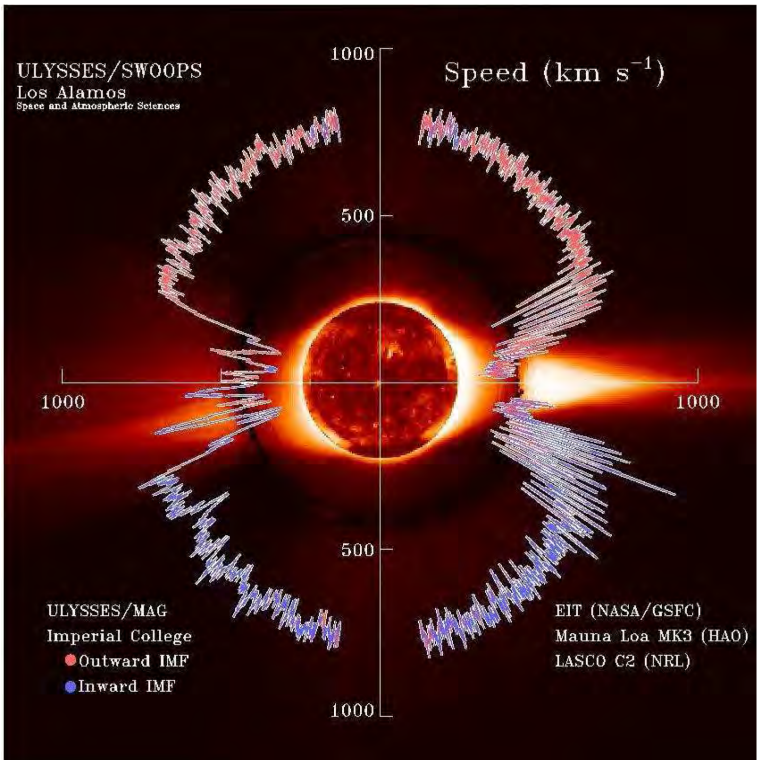


Figure 2.7: The image shows the SW speed measurements from its circling orbit around the Sun as a function of heliospheric latitude. Above 30^0 ($N \propto S$), the SW speed exceeds 750 km/s while at lower latitudes, the SW speed is $\simeq 450$ km/s. (Courtesy of the Ulysses mission, a project of international collaboration between ESA and NASA).

2.3.1 Corotating interactive regions (CIRs)

Solar coronal holes are polar areas of open magnetic fields, which migrate towards the solar equator during the descending phase of SC. M regions are boundaries between coronal holes and coronal streamers. Coronal holes are the sources of fast solar wind flows (750 - 800 km/s), while coronal streamers are (especially the streamer belt at the equatorial latitudes) are sources of the normal solar wind flow. The fast wind stream from coronal holes extends to the plane of the solar equator. When this fast wind overtakes the slow-speed, equatorial one, the two wind components interact. This interaction produces shock waves and intense magnetic fields that rotates with the Sun (Prölss, 2004).

The magnetic fields of the slower speed stream are more curved, while the fields of the higher speed stream are more radial due to their high speed. This interaction produce an interface (IF) which is the boundary between the slow stream and fast stream of plasma and fields. The front of the IF are the compressed and accelerated slower speed plasma and fields, while the back of the IF are the compressed and decelerated high speed stream plasma and fields (Tsurutani and Gonzalez, 1997). Due to the fact that both streams rotate with the Sun, the developed interaction region is called a corotating interaction region or CIR, a nomination given by Smith and Wolfe (1976), following observations of this IP structure by Pioneer 10 and 11. If the ambient magnetic field already possesses a negative IMF Bz-component, it can be amplified to the point where a GMS is triggered as a compression region passes the Earth. These CIRs are most responsible for the recurrent GMS, repeatedly occurring each 27 day interval (Burlaga and Lepping, 1977; Burlaga et al., 1978).

At a distance of about $> 1.5AU$, the CIR structures are well developed and bounded by fast forward (FS) and fast reverse (RS) shocks as illustrated by Figure 2.8. More details about CIR structures can be found in Tsurutani and Gonzalez (1997) and references therein.

2.4 Solar-terrestrial interactions

Solar-terrestrial relations are based on the interaction of the SW with the Earth's external magnetic field. The following sections outline the structure of the magnetosphere and its interaction with the SW leading to GMS events.

2.4.1 The Earth's magnetic field

From the 16^{th} century, it was known that our planet behaves like a great magnet. It is believed that the Earth's magnetic field is produced by electrically conducting currents in the Earth's molten core (Lang, 2001). As shown by Figure 2.9, the Earth's magnetic field is

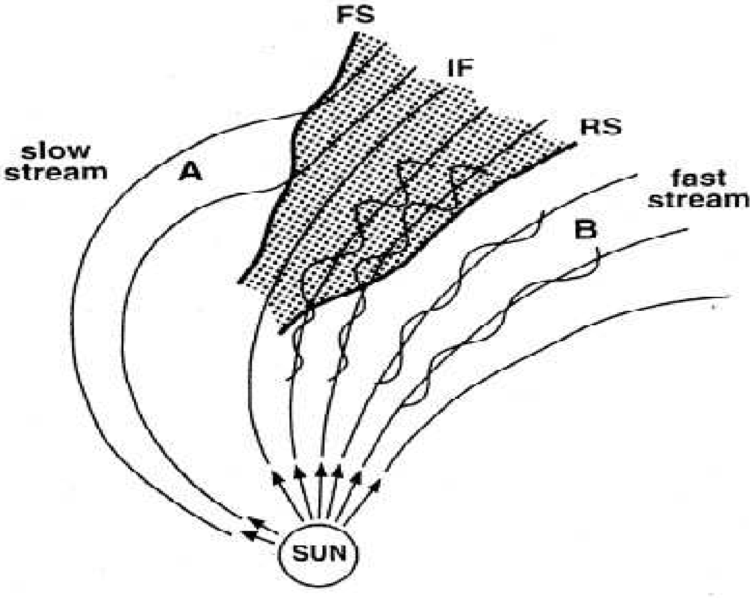


Figure 2.8: A schematic illustration of CIR formation by Kamide et al. (1998). Indicated are the interaction between a high-speed stream (B) and a slow speed stream (A) and the (shaded) CIR. FS represents the forward shock, IF the interface surface while RS represents the reverse shock.

characterised by a magnetic field direction parallel to the Earth's surface (i.e. horizontal) for regions at low geographic latitude and perpendicular to the Earth's surface (i.e. vertical) for regions of high geographic latitude. The Earth can therefore be approximated as a uniformly magnetised sphere along its dipole axis.

Although the positions of the magnetic inclination poles are subject to secular variations (Prölss, 2004), the Earth's dipole axis intersects the surface at two points: the austral (southern) pole (AP) at $79^0 S, 110^0 E$; and the boreal (northern) pole (BP) at $79^0 N, 290^0 E$. The AP is close to the Vostok station in Antarctica and the BP is close to Thule (Greenland). The two positions are about 800 km from the geographic poles. The magnetic dipole axis is inclined by 11.3^0 with respect to the Earth's axis of rotation.

The near-Earth structure of the terrestrial magnetic field corresponds to the familiar dipole field, thus obeying the magnetostatic theory. A magnetic field is described by the magnetic flux density **B**. This quantity determines the strength and direction of the magnetic force acting on a charge q and moving at the velocity **v**.

$$\mathbf{F}_B = q\mathbf{v} \times \mathbf{B} \tag{2.14}$$

where **F** is the Lorentz force. The unit of **B**, according to the international system of units

(SI), is the *Tesla* or T. The non-SI units of magnetic field used is the *gamma* or γ where 1γ is equal to 1 nanotesla (nT) and $1nT = 10^{-9}T$. The magnetic dipole moment of a dipolar field is a quantity that characterises the dipolar magnetic field and is expressed by the product of the pole strength and the distance between the poles, $\mathbf{M} = |P_B|\mathbf{d}$. Assuming a bar magnet at the centre of Earth, the approximated magnetic dipole moment of the Earth is $M_E \backsimeq 7.7 \times 10^{22} Am^2$.

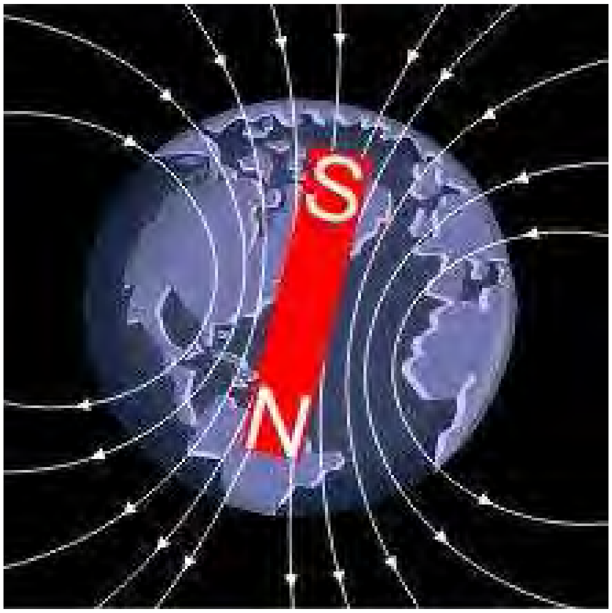


Figure 2.9: Representation of the orientation of the Earth's magnetic field (Figure from *physorg.com*).

Geomagnetic coordinates and components

In a spherical coordinate system oriented along the dipole axis, the position of a point P is described by its distance r from the centre of the dipole, and the angle θ from the dipole axis to the radius vector of P. Given the magnetic moment M_E, the components of the magnetic field can be written as (Kallenrode, 1998):

$$B_r = -2\frac{M_E}{r^3}cos\theta; B_\theta = -\frac{M_E}{r^3}sin\theta, \tag{2.15}$$

and the magnetic flux density is then expressed by:

$$B = \sqrt{B_r^2 + B_\theta^2} = \frac{M_E}{r^3}\sqrt{1 + 3cos^2\theta}. \tag{2.16}$$

The field strength falls oof with distance a $1/r^3$.

The geomagnetic coordinate system is oriented along the magnetic dipole axis and the equatorial plane intersects the dipole axis perpendicularly at the center of the Earth. The intersection of the equatorial plane with the Earth surface defines the geomagnetic equator. The geomagnetic longitude $\mathit{\Lambda}$ and latitude Φ are defined to the geographic longitude λ and latitude φ. With $\varphi_0 = 78.3^0$N and $\lambda_0 = 291^0$E as the latitude and longitude of the boreal magnetic pole, the magnetic and geographic coordinates are related by the transformations

$$sin\Phi = sin\varphi + sin\varphi_0 + cos\varphi cos\varphi_0 cos(\lambda - \lambda_0) \tag{2.17}$$

and

$$sin\Lambda = \frac{cos\varphi sin(\lambda - \lambda_0)}{cos\Phi} \tag{2.18}$$

The magnetic potential at a position r from the Earth's centre is expressed by

$$V = \frac{\mathbf{M}_E \cdot \mathbf{r}}{r^3} = -\frac{M_E sin\varphi}{r^2}. \tag{2.19}$$

The magnetic field strength can therefore be derived as $\mathbf{B} = -\nabla V$.

In a rectangular Cartesian system, the triple (X,Y,Z) gives the northward, eastward and vertical components, as shown on Figure 2.10. In a cylindrical system, the triple (D,H,Z) is used where Z and H represent the vertical and horizontal intensities respectively and D is the declination of the field. In a spherical system, where Z and X are the axes of reference, the field is described by the total intensity F, inclination I, as well as the declination D (see Figure 2.10). The dip equator or geomagnetic equator corresponds to the line with $I = 0^0$. The field components can be derived from equation 2.19. The radial or vertical component is

$$Z = -B_r = \frac{\partial V}{\partial r} = \frac{2M_E sin\phi}{r^3}, \tag{2.20}$$

and the horizontal component is

$$H = B\varphi = \frac{1}{r}\frac{\partial V}{\partial \phi} = -\frac{M_E cos\varphi}{r^3} \tag{2.21}$$

where ϕ is the geomagnetic latitude. At the pole, B equals Z while at the equator, B equals H. The magnetic inclination I is expressed by $I = \frac{Z}{H} = -2tan\phi$. The details of the above overview of the Earth's dipolar magnetic field coordinates and components can be found in Kallenrode (1998).

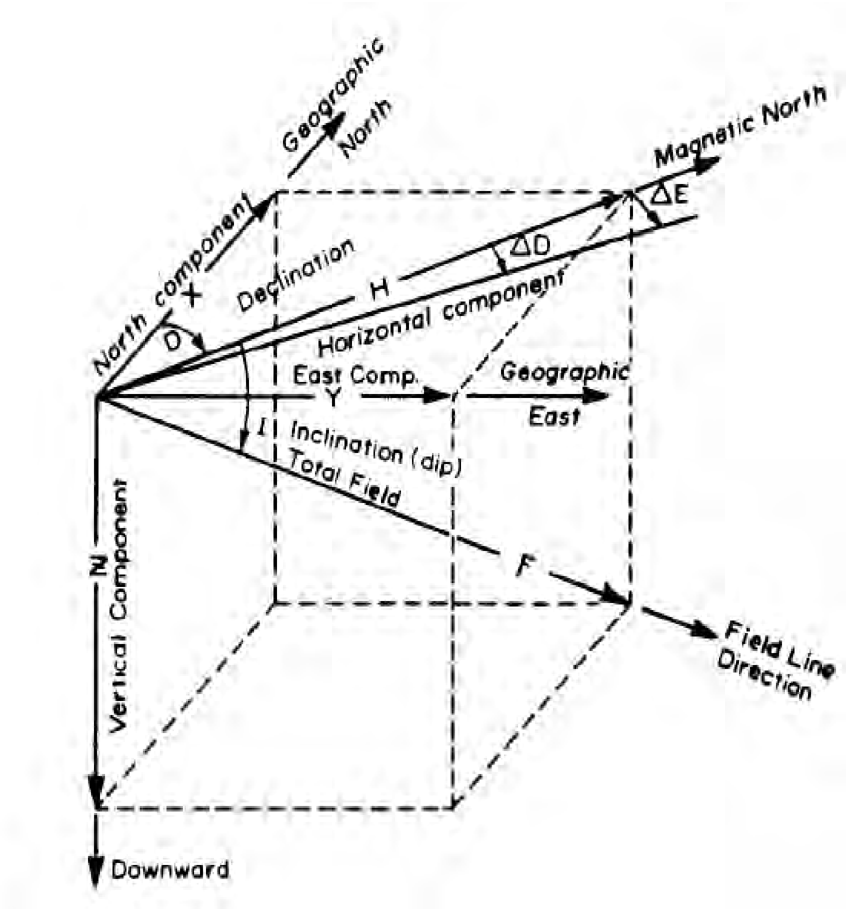


Figure 2.10: Spatial representation of the Earth's magnetic field components. Figure adapted from Wallace (2003)

2.4.2 The magnetosphere

At the surface, the Earth's magnetic field (near the equator) is about 3.1×10^{-5} T and the field decreases as the field lines extend to a greater distance from the Earth, but remain strong enough to shield the Earth from the SW force. Thus, the terrestrial magnetic field (magnetosphere) protects life on Earth from possibly lethal energetic particles (Lang, 2001). At a distance of about $10R_E$ on the dayside, the SW exerts pressure on the magnetosphere forming a boundary, called the magnetopause. Upstream of the magnetopause the supersonic SW is slowed down forming a bow shock.

As a result of its interaction with the IP medium, the Earth's magnetic field is confined into a finite volume called the magnetosphere, the *magnetopause* being the outer boundary of this volume. On the sunward side, the magnetosphere presents an ellipsoidal shape and the geocentric distance to the subsolar point of the magnetopause is about 10 Earth radii. (1 Earth radius or R_E is about 6 400 km). The nightside magnetosphere is greatly extended taking on a cylindrical shape. This region is named the *magnetotail* because of its similarity to the tail of a comet. The length of the magnetotail is variable and can extend beyond the Moon's orbit to about 60 R_E. Routine observations were even made beyond 200 R_E by Geotail.

To describe magnetospheric topology, the geocentric solar magnetospheric (GSM) coordinate system is frequently used. In this system the $x-$axis points to the Sun, along the Earth-Sun line. The y-coordinate measures the distance between the plane defined by the dipole axis and the Earth-Sun line. The z-axis completes the Cartesian system of coordinates and is given by, $\hat{z} = \hat{x} \times \hat{y}$.

2.4.3 The ring current

The terrestrial ring current is a toroidal-shaped electric current that flows westward around the Earth at geocentric distances between ~ 2 R_E and ~ 9 R_E (Daglis et al., 1999). GMS which are characterised by a decrease in the disturbance storm time (Dst) geomagnetic index are the consequence of the ring current enhancement, which produces the decrease in the H-component of the geomagnetic field on the Earth's surface. The ring current is carried mainly by energetic ions (in the range $10 - 200$ keV) that are trapped by the geomagnetic field lines and undergo an azimuthal drift (Daglis et al., 1999). Figure 2.11 is a schematic illustration of the terrestrial magnetosphere and the topology of the ring current. When energetic particles are injected into the inner magnetosphere on the nightside, they are influ-

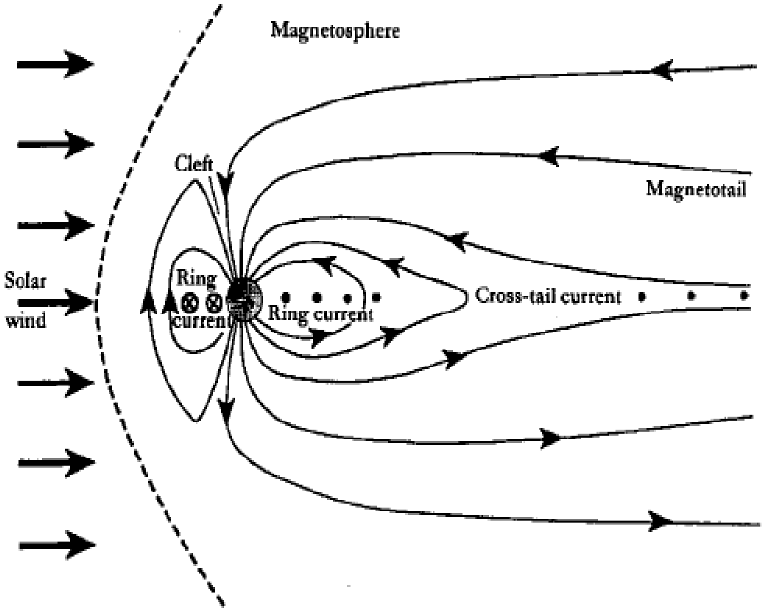


Figure 2.11: A schematic illustration of the topology of the ring current in the terrestrial magnetosphere. Adapted from Daglis et al. (1999).

enced by forces due to the curvature and gradient of the Earth's magnetic field. Because of these forces, protons drift westward from nightside toward dusk and electrons drift eastward from nightside towards dawn, and the resulting effect is the ring current encircling the Earth westward (Kamide and Maltsev, 2007).

Basic motions of particles in the presence of a magnetic field are the gradient drift motion, the curvature drift motion and the gyration (Baumjohann and Treumann, 1996). Parker (1957) established a hydromagnetic formalism which relates the magnetospheric ring currents to particle pressures both parallel ($P_{||}$) and perpendicular ($P_{\perp}$) to the magnetic field. The current due to particle drift driven by the magnetic field gradient is expressed as

$$\mathbf{J}_{\triangledown} = P_{\perp}\frac{\mathbf{B}\times\nabla B}{B^3}. \tag{2.22}$$

The current due to particle drift driven by the magnetic field curvature is expressed as

$$\mathbf{J}_C = \frac{P_{||}}{B^4}\mathbf{B}\times\left(\mathbf{B}\cdot\nabla\right)\mathbf{B}, \tag{2.23}$$

and the current due to gyration effects within the particle distribution is

$$\mathbf{J}_G = \frac{\mathbf{B}}{B^2}\times\left(\nabla P_{\perp} - \frac{P_{\perp}}{\mathbf{B}}\nabla B - \frac{P_{\perp}}{B^2}\left(\mathbf{B}\cdot\nabla\right)\mathbf{B}\right). \tag{2.24}$$

In equation 2.24 the three terms on the right-hand side represent currents due to the particle pressure gradient, the magnetic field gradient and the magnetic field line curvature. The total current which does not depend on the gradients of the magnetic field is then written as:

$$\mathbf{J} = \mathbf{J}_{\triangledown} + \mathbf{J}_C + \mathbf{J}_G = \frac{\mathbf{B}}{B^2}\times\left(\nabla P_{\perp} + \frac{P_{||}-P_{\perp}}{B^2}\left(\mathbf{B}\cdot\nabla\right)\mathbf{B}\right). \tag{2.25}$$

The quiet time average ring current density range is $\sim$ 1-4 nAm^{-2} while the storm time ring current can exceed 7 nAm^{-2}. During a strong magnetic storm, particles are injected from the plasma sheet in the magnetotail into the radiation belts, enhancing its density. The enhanced ring current then reduces the magnetic field measured on the ground. During disturbed conditions, particle density increase leads to an enhancement of electric current associated with the drift motion.

The magnetic disturbance level associated with the density enhancement of the ring current can be estimated (Prölss, 2004): Consider a current loop in the equatorial plane at a certain distance from the Earth. The equivalent magnetic field in the centre of a ring current with radius R is

$$B = \mu_0 I/2R, \tag{2.26}$$

where I is the ring current intensity and $\mu_0 = 4\pi$, the magnetic permeability. An approximated magnetic field disturbance at the Earth's surface is, therefore expressed as:

$$\triangle B_{rc}(R_E) \lesssim \triangle B_{rc}(r=0) = -\frac{\mu_0 \triangle I_{rc}}{2LR_E} \tag{2.27}$$

where r is the geocentric distance, $\triangle I_{rc}$ the disturbance-induced enhancement of the ring current intensity, while L is the shell parameter which is the geocentric distance of the field line in the geomagnetic equatorial plane measured in units of Earth radii. At $L = 5$, the intensity of the ring current would need to be increased by 5 MA to produce a magnetic field of $\triangle B_{rc} = -100$ nT (Prölss, 2004).

2.5 Geomagnetic storms and space weather

2.5.1 Geomagnetic storms (GMS)

GMS are complex phenomena that originate at the solar corona and occur in the SW, the magnetosphere, the ionosphere and thermosphere (Kamide and Maltsev, 2007). According to Kamide and Maltsev (2007), a magnetic storm is characterised by a depression in the H component of the magnetic field lasting from one to several days. This depression is a direct consequence of the increase in the ring current flowing westward in the magnetosphere and is generally measured by the Dst index which can be calculated as follows (Kamide and Maltsev, 2007):

$$Dst = \frac{1}{N}\sum_{n=1}^{N}\frac{H - H_q}{cos\phi} \tag{2.28}$$

where H represents the horizontal component of the magnetic field disturbance at a given station, H_q the same component over the quietest days; N is the total number of stations, and ϕ represents a particular station latitude. During a typical magnetic storm, the Dst index suddenly increases from zero level (the sudden storm commencement or SSC), drops to zero level and then decreases to a minimum (negative) value before recovering to the zero level. The minimum Dst value reached is an indication of the intensity of the storm. Generally, the minima of the Dst index less than -50 nT values are considered as magnetic storms (Kamide and Maltsev, 2007). A magnetic storm develops when the coupling of matter and energy from the SW with the magnetosphere is strong and long-lasting. The direction of the north-south component of the IMF regulates the growth of the ring current (Kamide et al., 1998) which begins with an injection of particles into the inner magnetosphere from the magnetotail.

In general, the storm begins with a sudden increase in the magnetic field. This is the SSC which may last a few hours. This initial phase is followed by a rapid and often disturbed decrease in *Dst* defining the storm's main phase. Subsequently, *Dst* initially recovers quickly, and then more gradually. The latter is the recovery phase which may last several days. Figure 2.14 is a graphical representation of the 15 May 2005 magnetic storm, indicating the three main phases of a typical storm.

The phases of a GMS can be simply interpreted as follows (Kallenrode, 1998):

- The SSC (increase in the magnetic field strength) at the beginning of the storm can be attributed to the compression of the magnetosphere when the magnetopause is pushed inward by increased solar wind speed.
- The decrease in the field strength during the main phase of the storm is due to an increase in the ring current, which creates a magnetic field opposite to the terrestrial one.
- The recovery phase of the storm corresponds to the ring current decays, primarily due to charge exchange.

Details describing storm manifestations can be found in reviews by (Gonzalez et al., 1994; Kamide et al., 1998; Kamide and Maltsev, 2007).

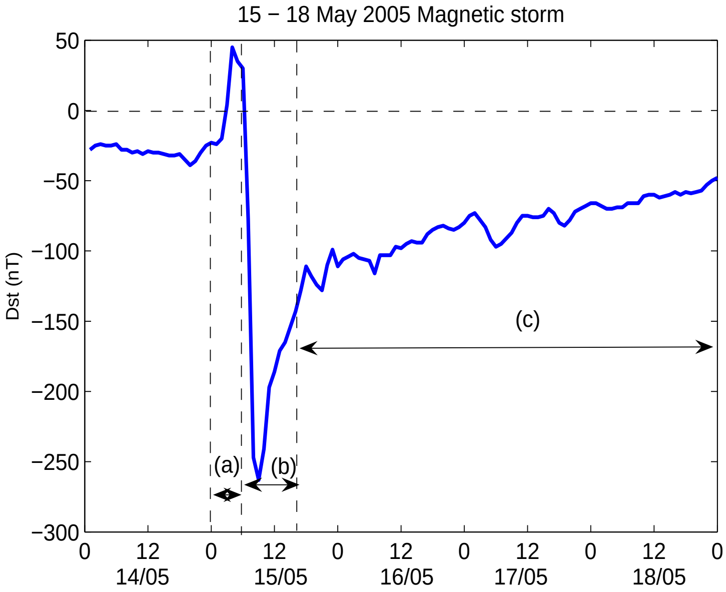


Figure 2.12: Graphical representation of a typical magnetic storm as measured by the *Dst* index. The three parts of a typical magnetic storm are (a) the SSC (b) the storm's main phase and (c) the recovery phase.

2.5.2 Solar and IP causes of geomagnetic storms

CMEs and IP structures such as ICMEs and CIRs produce SW disturbances leading to magnetic storms. The scale of a magnetic storm depends on the kind of disturbance in the SW. McPherron (1997) indicates that GMS events are a consequence of the SW dawn-to-dusk electric fields which are caused by a combination of SW velocity and the southward pointing IMF. The merging between the IMF and the Earth's field is most effective if they point in opposite directions. When the IMF has a southward component, the magnetosphere has an open configuration and plasma and energy are fed into it easily (Kallenrode, 1998). In other words, the energy flow into the magnetosphere is proportional to the rate at which the southward magnetic flux is carried to the magnetosphere by the SW (Burton et al., 1975).

As described by Tsurutani and Gonzalez (1997), there are two primary mechanisms that are the main sources of enhanced dawn-to-dusk electric fields in the IP medium. These are CMEs and CIRs as illustrated in Figure 2.13. During the maximum of a SC, CMEs are the dominant cause of GMS. Unlike the great sporadic storms, moderate GMS are not necessarily associated with solar transient phenomena (Lang, 2001). Moderate storm activity correlates with the 27-day solar rotation period at the equator. As described before, this recurrent storm activity is linked to CIR phenomena, commonly observed during the descending phase of an 11-year SC.

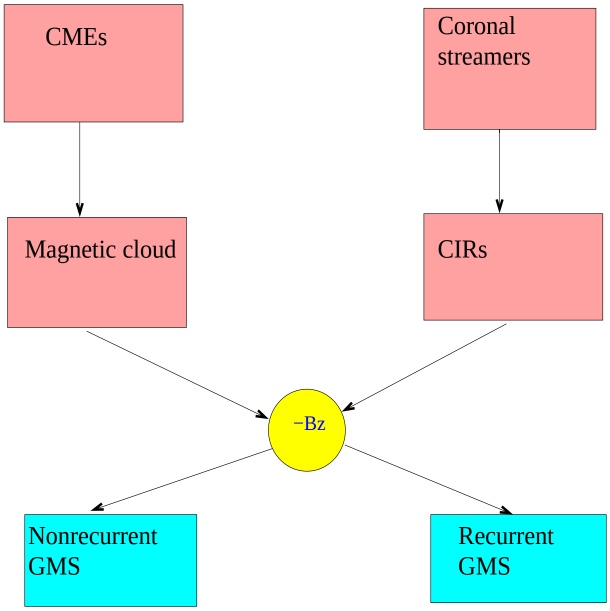


Figure 2.13: Some aspects of the solar and IP causes of non recurrent and recurrent GMS. Adapted from Prölss (2004)

2.5.3 Geomagnetic indices

In addition to the Dst index as a measure of geomagnetic disturbances, other geomagnetic indices are often used. Those include the AE (Auroral Electrojet), the ap, the K and the Kp indices. The Kp index is based primarily on data from magnetic observatories at middle and high northern latitudes and its values are generated with a time resolution of 3 hours. It therefore represents a quasi-logarithmic measure of the disturbance range, having values between 0 (very quiet) and 9 (very disturbed). One of the advantages of the Kp index is that it is a good measure of the general level of magnetic activity (Prölss, 2004). The linear ap index (in units of nT) is preferred for calculating sums and averages. The conversion of the Kp to ap index is shown in Table 2.3.

The K-index quantifies disturbances in the H-component of the Earth's magnetic field with an integer in the range 0-9; with 1 indicating calm conditions and 5 or more indicating a storm. The K-index is derived from the maximum fluctuations of the H-component observed on a magnetometer during a three-hour interval. The official Kp-index is derived by calculating a weighted average of the K-indices from a network of geomagnetic observatories. Ground stations (magnetometers) throughout the world monitor geomagnetic activity providing a local logarithmic K or linear a index of its strength during three-hour interval. Table 2.4 shows the relationship between the logarithmic K index and linear a index as a classification of geomagnetic disturbances.

Table 2.3: Conversion of the Kp to ap index, a table adapted from (Prölss, 2004)

Kp	0	1	2	3	4	5	6	7	8	9
ap (nT)	0	4	7	15	27	48	80	132	207	400

Table 2.4: Classification of geomagnetic disturbances using the K and a indices, by Lang (2001)

Description	K index	a index [nT]
Quiet	0 to 2	< 8
Unsettled	3	8 -15
Active	4	16-29
Minor storm	5	30-49
Major storm	6	50-99
Severe storm	7 to 9	100-400

The National Oceanic and Atmospheric Administration (NOAA) in United States provides a storm classification based on the Kp index (http://www.swpc.noaa.gov/NOAAscales/), where $G1$ (minor storm) corresponds to: $Kp = 5$; $G2$ (moderate storm): $Kp = 6$; $G3$ (strong storm): $Kp = 7$; $G4$ (severe storm): $Kp = 8$ and $G5$ (extreme storm): $Kp = 9$.

2.5.4 Solar and IP disturbance effects on ground and space-based technologies

Solar transient phenomena and the subsequent GMS are keenly monitored for their technological impacts. This section is an outline of some aspects of space weather, as adapted from Prölss (2004):

- GMS can produce geomagnetically induced currents (GICs). These are large fluctuations of the Earth's magnetic field strength during magnetic storms and can induce significant transient voltages and currents in the Earth and in extended conductors such as high voltage transmission cables and oil pipelines. In severe cases, the additional current can shift the operational range of high voltage transformers into saturation, leading to the overload and meltdown of the transformer windings. This is what occurred during the great magnetic storm of March 1989, when the power system in the Canadian province of Quebec was interrupted for approximately 9 hours.

- Magnetic storm events can be a threat to magnetic prospecting. This technique uses high precision airborne observations of the geomagnetic field to investigate anomalies that may provide evidence of valuable mineral deposits.

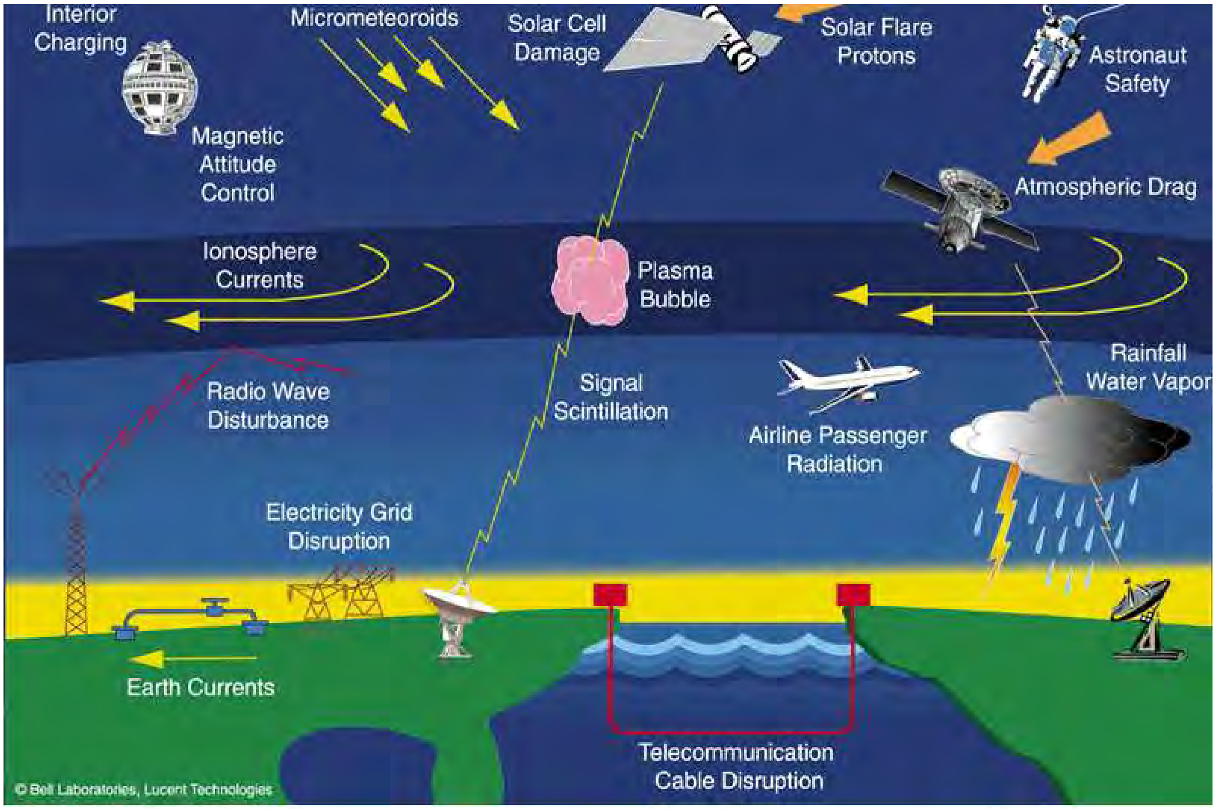
- Solar storms may be followed by thermospheric storms leading to increased air drag. As a consequence, there is a stronger and more erratic deceleration of low orbiting satellites which may necessitate a recalculation of their orbital elements. These conditions reduce satellite life expectancy due to the enhanced drag, forcing the satellite to re-enter the denser atmosphere prematurely.

- The preferred transmission frequency for ionospheric radio communication is set near the maximum plasma frequency of the F region. Disturbances can often produce negative *ionospheric storms* leading to the lowering of the maximum reflection frequency. Given that in this radio range the frequency bands are wholly allocated, lowering the transmission frequency is often impossible and the transmitted radio signal will be lost.

- Ionospheric storms also affect Global Positioning System (GPS) networks. When they pass through the dispersive ionosphere, GPS signals are slightly delayed. For single frequency transmission, this can result in a ranging error of about 50 m that would need correction. The magnitude of the error depends on the total electron content (TEC) of the ionosphere. During disturbed conditions, the TEC is highly variable, making error corrections difficult.

- Solar and IP storms are sources of *Energetic Particles* that can produce considerable damage to communication electronics on board a satellite, or can directly impair human health.
- Finally, it is important to note that SW magnetic fields and geomagnetic fields play crucial roles in space climate by modulating the precipitation of high energy galactic cosmic rays and solar particles into the Earth's atmosphere (Chian and Kamide, 2007).

Further details of various aspects of space weather can be found in other sources (e.g. Lanzerotti, 2007). Figures 2.14 and 2.15 illustrate various solar-terrestrial impact on ground-and space-based technology.



The effects of magnetic storms - what scientists call space weather - extend from the ground to geostationary orbit and beyond.

Figure 2.14: Illustration of various aspects of space weather [credit:www-istp.gsfc.nasa.gov]

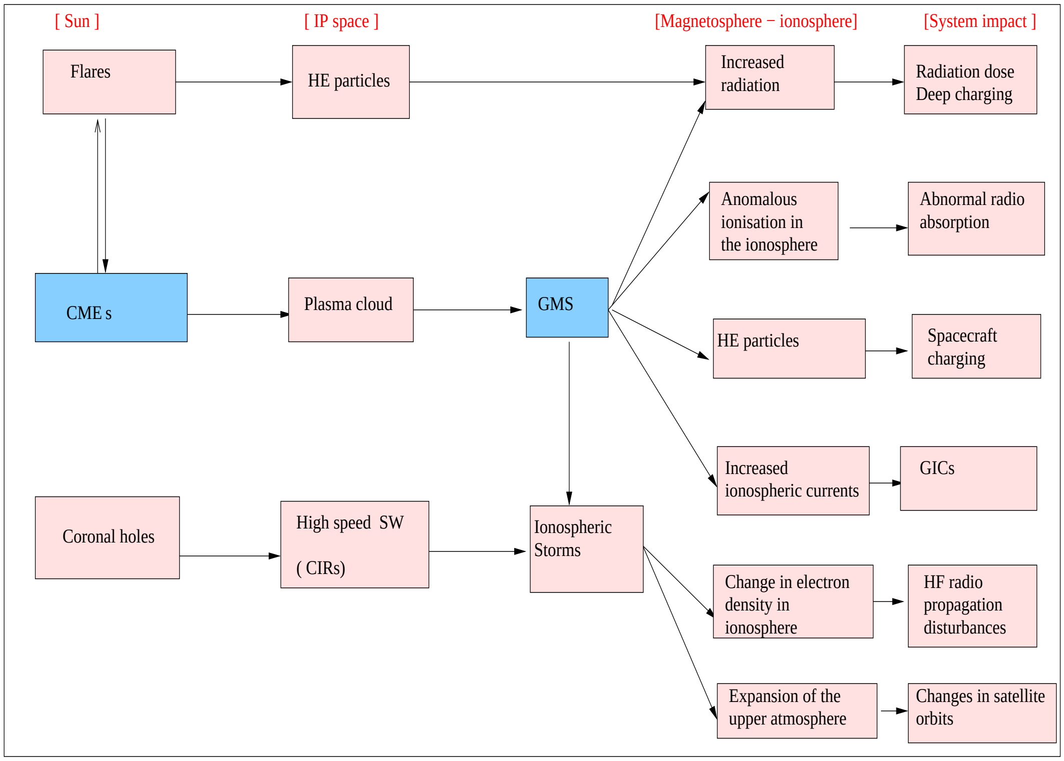


Figure 2.15: A simpified diagram representing some particular aspects of the solar-terrestrial impact on various technological systems, adapted from Ondoh (2001).

2.5.5 Space weather monitoring

During the past two decades, a number of spacecraft have been launched and their observations have provided new data that contribute to the improvement of space weather forecasting and prediction. In addition to the NASA/ESA SOHO spacecraft, two NASA STEREO spacecraft have been operating since they were launched in 2006. Currently, observations from the SECCHI (Sun Earth Connection Coronal and Heliospheric Investigation) on board STEREO spacecraft allows direct observation of CMEs, all along the Sun-Earth line from the low corona to the Earth's magnetosphere, providing new insight into the 3D structure of CMEs. On the other hand, the ACE spacecraft located upstream at L1, provides SW measurements in advance, with a view to warning of possible near-Earth disturbances.

With new observational techniques and modern technology, space scientists are now able to provide advance warning of probable threatening solar and IP events. For example, using the SOHO/LASCO coronal images, space scientists at the Solar Influences Data Analysis Centre (SIDC),in Brussels, Belgium, have developed a software package for the automatic detection and tracking of halo CMEs (see: http://sidc.oma.be/cactus/). The SIDC is part of a worldwide network of 13 RWCs, which are coordinated by the ISES to continuously monitor and provide space weather services. Other well known RWCs include among others, the NOAA Space Environment Centre in Boulder, USA and the Ionospheric Prediction Service (IPS) Radio and Space Services located in Sydney, Australia. In Africa, space weather services are provided by SANSA Space Science Directorate located in Hermanus, South Africa.

Chapter 3

Neural networks in space weather prediction

One of the aims in the domain of solar-terrestrial research is to develop models with ability to predict and forecast space weather and hence, minimise the cost of its impact. GMS represent a particular feature of space weather, and part of this thesis describe the development of models to predict GMS phenomena using Neural Networks (NNs). This chapter provides an outline of the current status of space weather prediction with focus on the prediction of GMS, and introduces the reader to the NN prediction techniques.

3.1 Space weather prediction challenges

Despite advances in ground-and space-based methods of monitoring solar and IP disturbances, accurate modelling and prediction of the solar-terrestrial environment is difficult to achieve. CME eruptions and subsequent GMS are governed by non-linear chaotic systems, thus complicating accurate modelling. At the moment, there is no reliable model that can accurately predict the onset and speed of a CME on the Sun (Messerotti et al., 2009), and most of the proposed models can only partially cope with the complexity of the solar-terrestrial environment system (Fox and Murdin, 2001). Currently, GMS prediction methods include among others, statistical, physics based as well as NN based methods.

Part of the study described in this thesis was devoted to developing NN-based models for prediction of the occurrence and strength of GMS. NN techniques are input-output models which have proved to be efficient in capturing the linear as well as the non-linear processes (Kamide et al., 1998). The main strength of the use of NNs in space weather prediction lies in their ability to model non-linear processes and this ability has been demonstrated in various studies (Lundstedt et al., 2005; Pallocchia et al., 2006; Woolley et al., 2010). As indicated by Messerotti et al. (2009), the use of NN techniques for predicting non-linear chaotic systems often perform better than other methods. The following sections introduce the reader to the basics of an Artificial NN (ANN) and its applications.

3.2 An introduction to the ANN prediction technique

An ANN is an information processing system consisting of a large number of simple processing elements called *neurons*, with reference to the neurons in the human brain. ANNs are characterised by (1) the pattern of connection between the neurons, (2) the method of determining the weights on the connections (training or learning algorithm) and (3) the activation function (Fausett, 1994).
For the NN models used for predictions, three types of neurons (or units) are defined:

1. input units, which are set to represent values within the time series
2. output units, which store the output values corresponding to a given set of input values and produce the results of the NN processing
3. hidden units, which keep the internal representation of the mapping

Units in layers are connected by weights which keep the knowledge of the network and govern the influence each input has on each output. Weights are adjusted by a learning process which involves the comparison of network calculations with input-output data for known cases. The process of adjusting weights is known as network training. During the training, weights are determined so that the network properly relates inputs to desired outputs. Hence, the network learns to predict outcomes from experience rather than from using causal laws (Macpherson et al., 1995).

A unique feature of NNs lies in their ability not only to learn the training data, but also to generalise by predicting unseen patterns within the boundaries given by the training set. In general, solving a non-linear problem with the NN technique requires (1) choosing a convenient network architecture, (2) selecting a large database of input-output pairs (patterns) that contains sufficient historical information about the time series, and (3) training the network to relate the inputs to the corresponding outputs. Several available algorithms have been proposed for ANN applications (Haykin, 1994; Fausett, 1994; Bishop, 1995). For the current study, two forms of ANNs were used, namely the feed-forward NN (FFNN) and the Elman Recurrent Network (ERN).

3.2.1 Feed-forward backpropagation NNs

The multilayer feed-forward error backpropagation algorithm (see Hertz et al., 1991, for details) belongs to the class of supervised networks, meaning that it learns from known answers. FFNN represents the simplest and most popular type of NN. In a FFNN arrangement, neurons (units) between layers are connected in a forward direction. Neurons in a given layer do not connect to each other and do not take inputs from subsequent layers. The input units which are set to the previous values of the time series, send the signals to the hidden units. These hidden units process the received information and pass the results to the output units which deliver the final response to the input signals. According to Cybenko (1989) networks with one hidden layer and a continuous sigmoidal transfer function are capable of approximating any continuous function with arbitrary accuracy. Figure 3.1 shows a simplified three-layer network configuration with input, hidden and output layers. For a network made by input k, hidden j and output i layers, a mathematical description of a FFNN is outlined by Kugblenu et al. (1999) as follows:

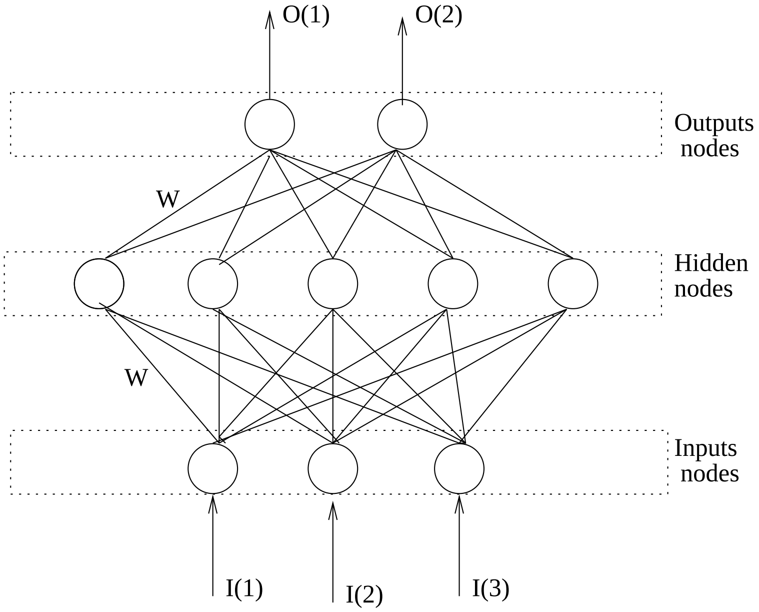


Figure 3.1: A simplified schematic of a single hidden layer FFNN. Units between layers are connected in a forward direction. Weight (W) connections between layers are adjusted during the training process.

For a given pattern μ, a neuron j in the hidden layer receives from a neuron k in the input layer, a net input described by

$$x_j^\mu = \sum_k w_{jk}\xi_k^\mu + b_k \tag{3.1}$$

where ξ_k^μ represents the input signal to neuron k in the input layer, w_{jk} represents the weight between neuron j in the hidden layer and neuron k in the input layer, and b_k is the bias connected to the input layer. The output V_j^μ from the hidden layer is written as:

$$V_j^\mu = g_h(x_j^\mu) = g_h\left(\sum_k w_{jk}\xi_k^\mu + b_k\right). \tag{3.2}$$

where $g_h(x)$ is the activation function. A neuron i in the output layer receives this signal from the neuron j in the hidden layer as input and produces an output, O_i^μ, which is written as:

$$O_i^\mu = g_o(x_i^\mu) = g_o\left(\sum_j W_{ij} V_j^\mu + b_j\right). \tag{3.3}$$

with $g_o(x)$ being the activation function which is generally assumed to be the same as $g_h(x)$. Omitting the biases b_k and b_j (i.e. they can be taken as extra input that are connected to all units in the network), the net output can hence be written as

$$O_i^\mu = g_o\left(\sum W_{ij} g_h\left(\sum w_{jk} \xi_k^\mu\right)\right). \tag{3.4}$$

The role of the activation function is to introduce non-linearity into the network and hence allow the network to represent non-linear functions. The most commonly chosen activation functions (for the backpropagation learning algorithms) are the sigmoid functions including the logistic, hyperbolic tangent and the Gaussian functions. The more general activation function, the binary sigmoid function, was used for the NN models developed for this research. This activation function has a range between 0 and 1, is differentiable and is expected to saturate (approach the maximum and minimum values asymptotically). The logistic sigmoid function is defined as (Fausett, 1994):

$$g(x) = g_h(x) = g_o(x) = \frac{1}{1 + exp(-bx)} \tag{3.5}$$

where b is the slope parameter which can modify the steepness of the logistic sigmoid function. The derivative of equation 3.5 is

$$g'(x) = bg(x)[1 - g(x)].$$

Note that the logistic sigmoid function is the default neuron activation function for the Stuttgart Neural Network Simulator (SNNS) software that was used in this study.

The training of a network by backpropagation consists of three main stages (Fausett, 1994), namely:

- The feed-forward of the input training vector to the input layer
- The computation and backpropagation of the associated error
- The adjustment of the weights

During the training process, the input ξ_k^μ is presented to the network together with its corresponding known output O_i^μ and the network system learns the relationship that exists between the two by adjusting the weights. The training proceeds until the network error (E) is minimised according to the relation

$$E = \frac{1}{2}\sum_{1}^{N}\left(T_i^{\mu} - O_i^{\mu}\right)^2 , \tag{3.6}$$

where T_i^{μ} is the target output with $i = 1, ..., N$, and O_i^{μ} the known output.

For the backpropagation of errors learning algorithm, the weights are updated by a small change in each time-step through the method of gradient descent, whereby each weight changes by an amount which is proportional to the gradient of the error E at the present location. The weight change by the gradient descent rule for the hidden to output connections is written as

$$W_{ij} = W_{ij} + \triangle\, W_{ij}, \tag{3.7}$$

where

$$\triangle\, W_{ij} = -\alpha \frac{\partial E}{\partial W_{ij}}, \tag{3.8}$$

with α the learning rate or learning parameter. Similarly, the weight adjustment for the input to hidden connections is given by

$$W_{jk} = W_{jk} + \triangle\, W_{jk}, \tag{3.9}$$

where

$$\triangle\, W_{jk} = -\alpha \frac{\partial E}{\partial w_{jk}} = \alpha \sum_{\mu} \frac{\partial E}{\partial V_j^{\mu}} \frac{\partial V_j^{\mu}}{\partial w_{jk}}. \tag{3.10}$$

3.2.2 The Elman neural network

The Elman NN (Elman, 1990) is a type of network that belongs to the class of recurrent NNs and hence, also commonly known as the Elman recurrent network (ERN). The ERN consists of an input layer, a hidden layer and an output layer. In this way, it ressembles a three-layer FFNN. However, it also has a context layer. The context neurons receive inputs from the hidden layers and also pass their output to the hidden layers. The context layer always stores the output from the hidden layer and relays this information in the next iteration. Context neurons therefore form a sort of short-term memory, very useful for improving prediction of sequences. Figure 3.2 is a simplified representation of the ERN configuration.

Using the same notation as by Cai et al. (2010), here follows a simplified mathematical description of the ERN. A two-part structure of the ERN input implies: the true external input $\mathbf{i} = I_1, I_2, I_3, ..., I_N$, and the feed-back input $\mathbf{C} = c_1, c_2, c_3, ...c_k$ as shown in Figure 3.2.

The feed-back input nodes are called context nodes or units. The value of context unit l at time step t is simply a copy of the hidden output at time step $t-1$, that is

$$C_l(t) = H_l(t-1) \qquad l = 1, 2, ..., S. \tag{3.11}$$

The j^{th} hidden unit output H_j is defined as

$$H_j = g\left(\sum_{i=1}^{M} w_{ji} I_i + \sum_{i=1}^{k} w_{jl} c_l + b_j\right). \tag{3.12}$$

In equation 3.12, g represents the sigmoid activation function. I_i is the value of the input neuron i, M is the number of input neurons, w_{ji} is the connecting weight between the input node i and the hidden neuron j, and b_j is the bias of the hidden neuron j. This means that the state of the whole network at a given time depends on an aggregate of the previous states, as well as on the current inputs (Pallocchia et al., 2006).

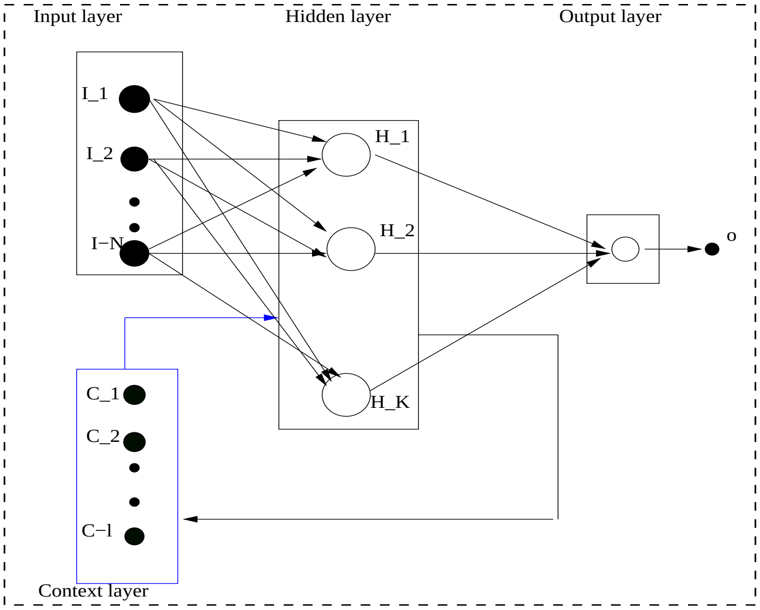


Figure 3.2: A simplified schematic illustration of the ERN with one input, one hidden and one output layers. The context neurons receive inputs from the hidden layer and also pass their output to the hidden layer.

3.2.3 The training procedure

Generally, the time series is split into two data sets: a training set and a test set. The training set is used to adjust the weights during training, while the test set is used to verify the prediction performance of the network (Fessant et al., 1996). The network is trained using past examples so that the NN has a minimal generalisation error in order for it to perform better on future examples. Both the training and testing data sets are split randomly in order to avoid training results to be biased towards a particular section of the database. In order to determine how well the NN has learned the behaviour in the input-output patterns, a data set known as a validation set (including data not seen by the network) is used.

The learning parameter α is generally small (range from 0 to 1) and is used to control the rate of the training process over hundreds or thousands of iterations. Training is repeated over a number of iterations until the network has attained the generalisation ability. Overtraining can lead to NN overfitting, where the NN learns the noise in the training set rather than learning to represent the system dynamics in general. This happens when the error on the testing set starts to increase, while the training set error starts decreasing. Just before this point is reached, the training is stopped, because then the network's generalisation ability has been optimised (Macpherson et al., 1995).

3.2.4 The Stuttgart Neural Network Simulator (SNNS)

SNNS is portable NN simulator software developed by researchers from the universities of Stuttgart and Tübingen, Germany, in collaboration with the European Particle Research Lab, CERN, Geneva in Switzerland (Zell et al., 1998; Reczko et al., 1998). The SNNS software has various built-in functions for optimisation and provides the relevant options for both the FFNN and ERN. Table 3.1 indicate the chosen functions for both ANNs. Further details on the use of the two NN approaches in a SNNS can be found in Zell et al. (1998) and Reczko et al. (1998).

Table 3.1: Selected functions for FFNN and ERN (Zell et al., 1998)

Function	Feed-forward network	Elman neural network
Update function	Topological Order	Jordan Elman (JE) Order
Learning function	Standard Backpropagation	JE Backpropagation
Initialisation function	Randomised Weights	JE Weights
Pattern remap function	None	None

3.3 Applications

NNs techniques have been successfully used for the prediction of various solar-terrestrial time series (e.g. Macpherson et al., 1995; Conway, 1998). Recently, this technique was also used to estimate the timing and amplitude of solar cycle 24 (Uwamahoro et al., 2009). With regards to the prediction of GMS, both FFNN and ERN-based algorithms have been applied and demonstrated good performance in predicting the geomagnetic *Dst* index during storm time (Wu and Lundstedt, 1997; Lundstedt et al., 2002). Currently, NNs techniques are widely used for real-time space weather forecasts:(e.g http://www.lund.irf.se/Helioshome/wg4nnhenrik.pdf). In this thesis, NN-based models were developed for predicting the occurrence and strength of GMS based on solar and IP parameters.

In summary, this chapter provided the basic principles of NN techniques as well as pointing out the difficulties inherent in space weather prediction. The complexity and non-linearity of the solar-terrestrial system justifies the preference given to NNs as a reliable tool for space weather prediction and forecasts. It must again be noted that the description of the NN techniques provided in this chapter is limited only to the basic principles. For more details about the FFNN and ERN techniques and their application, the reader is referred to references provided (Elman, 1990; Haykin, 1994; Fausett, 1994; Bishop, 1995).

Chapter 4

The main sources of geomagnetic storms in SC 23

Explosive events occurring on the Sun are primarily responsible for space weather which affects space-and ground-based technology, as well as life on Earth (e.g. Siscoe and Schwenn, 2006). This chapter provides a qualitative and quantitative analysis of the solar and associated IP causes of GMS during the 11-year period of SC 23: 1996 - 2006. The probable solar causes of 229 GMS ($Dst \leq -50$ nT) were investigated, with a focus on halo CMEs.

4.1 Geoeffectiveness of solar and IP phenomena

CMEs are transient expulsions of plasma and magnetic field from the Sun and are responsible for strong IP disturbances leading to GMS and related space weather phenomena, as described in Section 2.5.4, or in Lanzerotti (2007). GMS occur as a result of the energy transfer from the SW to the Earth's magnetosphere via magnetic reconnection. The main solar sources of GMS are: (a) the CME eruptions on the Sun (Gopalswamy et al., 2007), and (b) the corotating interaction regions (CIRs) that result from the interaction between the fast and slow SW originating from coronal holes (Zhang et al., 2007). The two phenomena evolve in geoeffective conditions in the SW, producing moderate to intense GMS when there is an enhanced and long-lasting IMF in the southward direction (Richardson et al., 2002; Richardson, 2006; Gonzalez et al., 1994). However, despite the prominent role played by CMEs in producing GMS, the prediction of GMS cannot be based only on CME observations. As noted by Wang et al. (2002), the properties of CMEs that give rise to magnetic storms are still not very clear. Therefore, improving the prediction of GMS requires the identification of the key solar and IP geoeffective parameters of CMEs (Srivastava, 2005).

Currently, space-based instruments have facilitated advanced observation and understanding of storms occurring on the Sun, i.e. detection of CME eruptions on the Sun since the onset of SC 23 (Brueckner et al., 1995). The CMEs that appear to surround the occulting disk of the observing coronagraphs are known as halo CMEs, among which those originating from the visible solar disc and are earth-directed have the highest probability of impacting the earth's magnetosphere (Webb et al., 2000). Currently, the STEREO spacecraft allows CME tracking from the Sun to Earth and the received data will contribute towards improving the predictability of geoeffective CMEs (Messerotti et al., 2009).

The geoeffectiveness of a CME refers to its ability to produce a magnetic storm with $Dst \leq -50$ nT. In their study, Webb et al. (2000) and Cyr et al. (2000) used 140^0 and 120^0 respectively as a threshold apparent angular width (AW) to define halo CMEs, while a study by Wang et al. (2002) considered a halo CME to be the one with apparent AW greater than 130^0. In this study halo CMEs are defined according to storm categorisation by Gopalswamy et al. (2007): full halo CMEs (F-type) with an apparent sky plane width of 360^0, and partial halos (P-type) with an apparent AW ranging between 120^0 and 360^0.

Over the 11-year period of SC 23 [from January 1996 to December 2006], 393 full halo CMEs were identified, representing 3.4% of the 11 683 CMEs recorded. During the same period the number of partial halo CMEs was 840. Therefore during the same period, LASCO observed 1 233 (10.5%) halo CMEs of which 393 (31.8%) were full halo CMEs and 840 (68.2%) partial halo CMEs. As illustrated in Figure 4.1, the frequency occurrence of full halo CMEs varies in step with the occurrence of GMS events for the 1996 - 2006 period. However, not all halo CMEs are geoeffective and the non halo CMEs can also cause GMS if they arrive at the Earth with an enhanced southward component of the IP magnetic field at high speed (e.g Gopalswamy et al., 2007). Indeed, and as indicated previously, both moderate and intense storms can also be caused by CIRs.

During the last few years, there has been increasing interest into research to estimate the geoeffectiveness of solar phenomena. A statistical study by Wang et al. (2002) investigated the geoeffectiveness of frontside CMEs during the rise of SC 23. The geoeffectiveness of both full and partial halo CMEs was discussed in detail by Gopalswamy et al. (2007) and Gopalswamy (2009b). Zhang et al. (2007) presented results of a study that investigated the solar and IP causes of major GMS events ($Dst \leq -100$ nT) during the period 1996 - 2005. In spite of all the research on this subject, persistent discrepencies between the various estimations of storm-causing effects of solar and IP events remain. According to Yermolaev et al. (2005), differing results may be due to the methods used for analysis of the data. On the other hand, a recent study by Gopalswamy (2009b) suggests that the range of reported geoeffectivity rates can be explained by the different definitions given to halo CMEs.

The main goal of this analysis is to contribute to the efforts towards estimating the GMS effectiveness of solar events. The focus is to investigate halo CMEs and associated solar and IP properties that were most likely causes of the 229 GMS events identified during the period 1996 - 2006. The present study is similar the analysis done by Zhang et al. (2007) which involved only 88 major storms ($Dst < -100$ nT). This study describes an additional quantitative estimate of the solar and IP properties that were associated with moderate GMS during SC 23. In addition, this study compares the geoeffectiveness of full and partial halo CMEs.

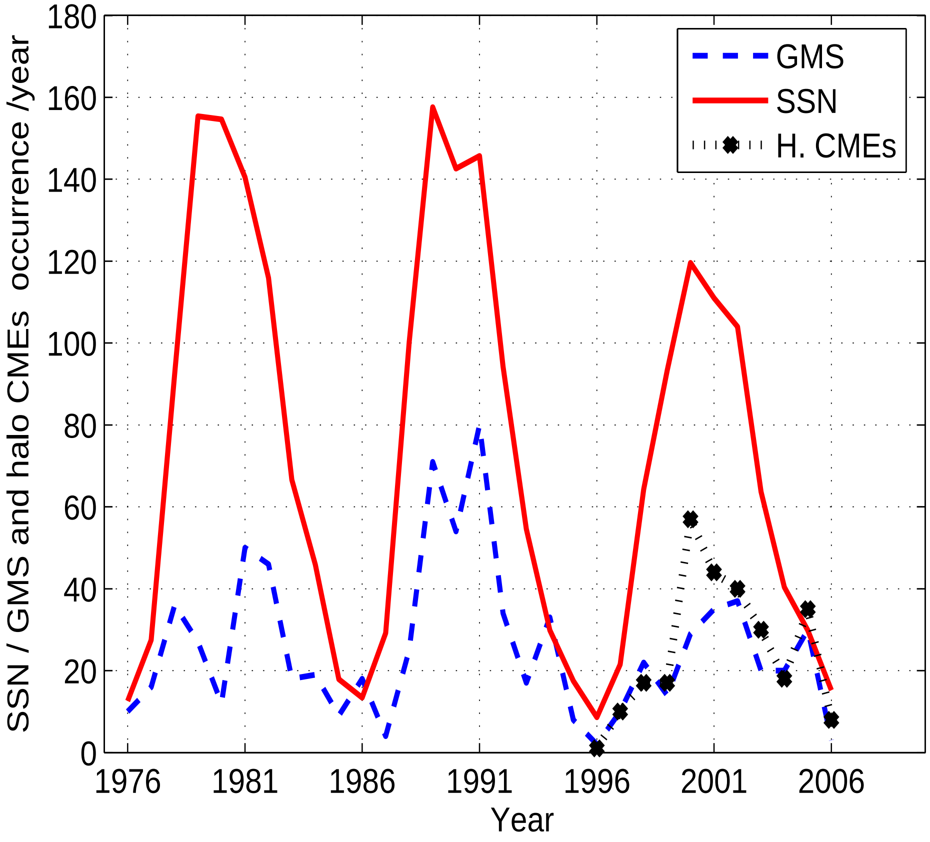


Figure 4.1: The correlation between solar activity cycle (in terms of sunspot number) and GMS occurrence during the last three solar cycles (SC 21, SC 22 and SC 23) is shown. For SC 23, the occurrence rate of GMS is shown for comparison with the occurrence rate of halo CMEs.

4.2 Data and methods

4.2.1 Selection of magnetic storm events

Despite the fact that phenomena leading to GMS always evolve in the Sun-Earth direction, it is common to proceed from Earth to the Sun when analysing the geoeffectiveness of solar events (Zhang et al., 2007). In addition, forecasting geomagnetic conditions on the basis of solar phenomena alone can lead to false alarms. Geomagnetic indices, such as the Dst and the Kp indices, are commonly used to quantify the level of geomagnetic response to solar transient phenomena.

For this study, only the Dst index was used for the selection of storm events. The Dst index is a measure of the horizontal component of the Earth's magnetic field at low to mid-latitudes. It is a proxy for the deviation of the horizontal component at the equator from a quiet day and is generally considered as a measure of the ring current. In the selection of storm events an effort was made to select clearly isolated storm periods, thus not considering any single event with $Dst \leq -50$ nT as a separate storm. This is because a given isolated storm can be characterised by multiple periods of $Dst \leq -50$ nT, especially during high magnetic activity periods.

Based on the minimum Dst value, Loewe and Prölss (1997) classify GMS as weak (in range -30 to -50 nT), moderate (-50 to -100 nT), intense (-100 to -200 nT), severe (-200 to -350 nT) and great (less than -350 nT). For simplicity, the storm classification by Gopalswamy et al. (2007) was followed in this analysis: moderate storms (-100 nT $\leq Dst <$ -50 nT) and intense storms $Dst \leq -100$ nT. Zhang et al. (2007) compiled a table indicating solar and IP sources of major GMS events ($Dst \leq -100$ nT) in SC 23 (1996 - 2005). Table 4.1 is an extension of the table by Zhang et al. (2007) to include moderate storms. It contains 244 storm events and probable associated halo CMEs during the period January 1996 to December 2006. Table 4.1 columns 2 and 3 indicate the storm time and the peak minimum value of the Dst index respectively. The Dst index data set was obtained from the National Geophysical Data Center (NGDC), available on the website: ftp://ftp.ngdc.noaa.gov/STP/GEOMAG/dst.html.

4.2.2 Near-Earth IP signatures of geoeffective solar events

Solar disturbances are transported from the Sun to near-Earth via the SW. The CME structures in the IP medium (at about 1 AU) are known as ICMEs. ICMEs often contain a strong and long-lasting southward component of the IMF in the driver plasma or in the plasma sheath which follows the CME-driven shock leading to intense GMS (Gosling et al., 1990). The geoeffectiveness of an ICME has been found to be correlated with the negative

Bz-component of the IMF, also written as B_s or $E_y \backsimeq VB_s$ (Richardson and Cane, 2011), where E_y is the y component of the SW convective electric field $E = -VB$. A catalogue of ICME properties as well as their geomagnetic effect can be found in Richardson and Cane (2010), and its updated version is available online:
(http://www.srl.caltech.edu/ACE/ASC/DATA/level3/icmetable2.htm.)
Therefore, Table 4.1 is a subset of the Richardson and Cane's table (RC) with two major differences:

- The RC table is based on ICME events and not on GMS events. Hence, Table 4.1 contains an additional 92 storm events which are not listed in the RC table. In the last column of Table 4.1, which indicates the start time of the ICME, GMS events which are missing in the RC table are marked as (m).
- Unlike the RC table, Table 4.1 only indicates moderate to intense storms. Hence all events with $Dst > -50$ nT do not appear in the table.

Figure 4.2 illustrates the perturbations induced in the SW and the IMF following the near-Earth passage of an ICME.

The magnitude of SW speed and the IMF negative Bz-component

A geoeffective SW is characterised by a prolonged and enhanced southward-directed magnetic field (Bs) that allows efficient SW energy transport into the Earth's magnetosphere. The intensity of a magnetic storm depends mainly on the magnitude of Bs and the speed V with which the CME impacts the Earth's magnetosphere.

To investigate SW geoeffective conditions associated with GMS events, the averaged hourly SW speed and the IMF Bs were considered. The peak values of the IMF B_s were selected within a one day window prior to the peak minimum Dst. For practical purposes, the SW velocity value considered in this analysis corresponds to the time of peak minimum Dst. In Table 4.1, the IMF Bs and the SW speed values are indicated in columns 4 and 5 respectively. SW and IMF Bs are in situ (measured at 1 AU from various spacecraft) measurements provided by the NSSDC and were obtained via its OMNIWEB: http://www.nssdc.gsfc.nasa/omniweb.html.

4.2.3 Solar properties associated with geoeffective halo CMEs

A large AW of CMEs is often an indication that they are directed along the Sun-Earth line, with a higher probability of impacting the Earth's magnetosphere (Howard et al., 1982). The AW and the first time of halo CME appearance in the coronagraphs is indicated in column 6 of Table 4.1. The AW of a CME is labelled FH for full halo (i.e $AW = 360^0$) and PH for a

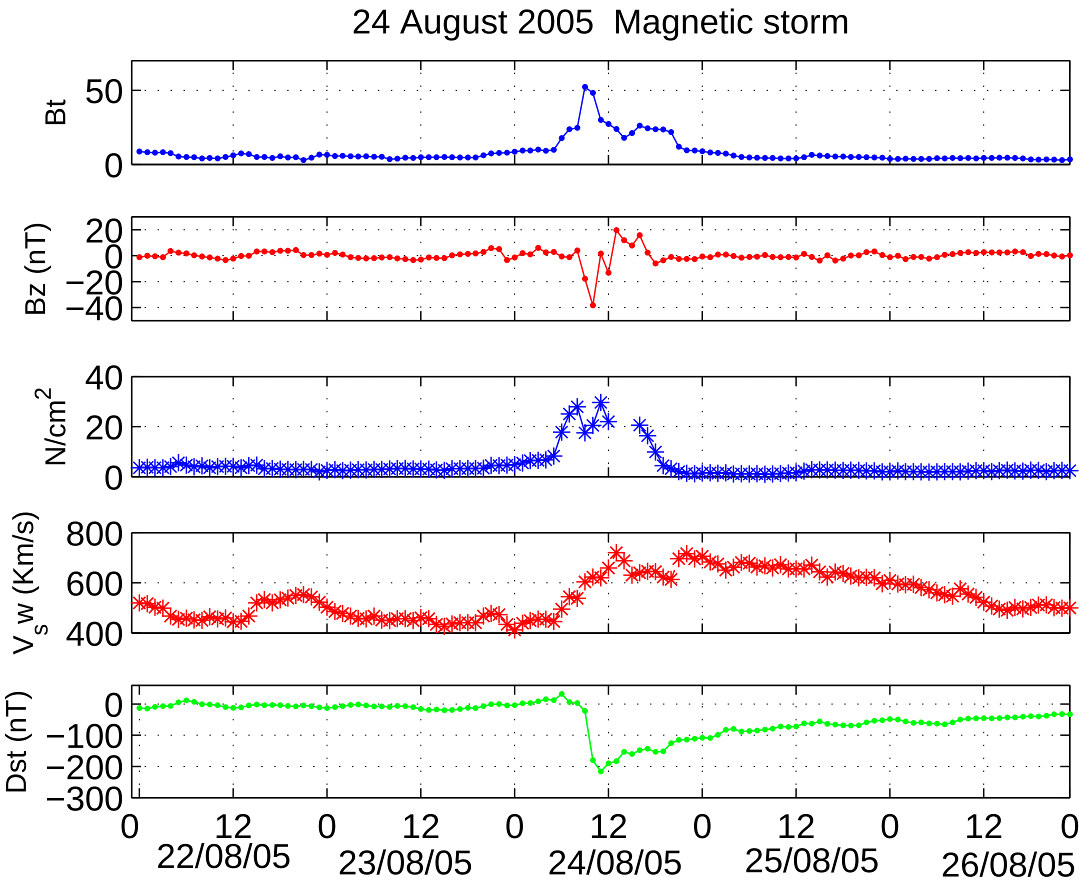


Figure 4.2: Illustration of the perturbations induced in the SW B_t, B_z, N and V parameters following the arrival of a geoeffective ICME. The bottom plot indicates the corresponding Dst index for the period around the GMS event of 24 August 2005.

partial halo ($120^0 \leq W \leq 360^0$). In addition to their AW appearance in the coronagraphs, other geoeffective CME properties include their speed (see column 7 of Table 4.1) and their association with X-ray flares. In his study, Kahler (1992) indicated that GMS events and their associated IP disturbances can be traced back directly to large flares. Indeed, a long-lasting flare can also be considered the source of the associated CME (Kahler, 1992). The X-ray flare class (see Chapter 2, Section 2.2 for flare classification) with the corresponding intensity level, as well as the surface coordinates are indicated in Table 4.1, columns 8 and 9 respectively.

Various observational techniques are available to identify the surface location of CME eruptions, e.g the SOHO's Extreme-Ultraviolet Imaging Telescope (EIT). The LASCO/EIT images provide a measure of heliographic source coordinates for halo CMEs associated with flare activity. Where available, the GOES data list also reports the source coordinates of flare eruptions which, in this study, have been assumed to be the source location of associated CMEs. Hence, the association of X-ray flares with halo CMEs can help to identify the solar surface location origin of CMEs (Gopalswamy, 2009b). In this analysis, the same criteria as used by Gopalswamy et al. (2007) was followed and a time lapse of ±0.5 hours was used to associate a halo CME eruption with a flare occurrence.

The above parameters were all considered for the analysis of GMS sources. The CME data used are from the LASCO/CME catalogue (Yashiro et al., 2004). The list is generated by NASA and the Catholic University of America in collaboration with the Naval Research Laboratory and is available online at http://cdaw.gsfc.nasa.gov/CME_list/. The solar flare data are provided by GOES via the Solar Geophysical Data Center and the data is available online on the website:
http://www.ngdc.noaa.gov/stp/solar/solarflares.html.

4.2.4 Methods of investigation

The average transit time of CMEs from the Sun to Earth is estimated at 80 hours (Brueckner et al., 1998). However, this transit time doesn't hold during solar maximum when very fast CMEs are observed. Cane and Richardson (2003) suggest that the transit time may generally vary in a large range from 1 to 5 days. For this analysis, a 5-day window prior to GMS occurrence was used to explore probable halo CMEs (and their properties) as causes of the subsequent GMS.

For every selected storm, the process consisted of examining the LASCO CME catalogue to track backward in the search window for the existence of a full (or partial) halo CME as the potential cause of the storm. Despite the fact that one isolated storm may be the result of more than one halo CME (Gopalswamy et al., 2007), all halo CMEs found within a 5-day time window were not considered as potential causes of the storm. Note that some halo CMEs may be backsided and therefore not likely to hit the Earth. Only frontside CMEs are able to reach the Earth and produce GMS. However, the possibility of backsided CMEs producing minor storms was reported in a study by Webb et al. (2000). In the current analysis, halo CME association with X-ray flares was the primary criterion for defining a potential cause of the corresponding storm, provided that the flare-CME association information was available (using EIT and GOES data). The importance of considering this criterion was explained in Section 4.2.3. The following serve as examples of GMS events and their corresponding solar and IP characteristics as listed in Table 4.1.

Event number 232 in Table 4.1 is the most recent among strong GMS in SC 23. The main phase of the storm occurred on 24 August 2005 with a peak minimum Dst of −216 nT. The IMF B_s peak reached a value of 38.3 nT while the SW speed at the time of peak minimum

Dst was 620 km/s. As indicated in Table 4.1, two high-speed ($V > 1000$ km/s) full halo CMEs were considered the probable sources of the storm. The two CMEs were associated with M-class flares and the corresponding surface coordinates are indicated. This is an example among many others where one storm is shared by more than one halo CME.

Another example is event number 187 in Table 4.1 which occurred on 20 November 2003 at 20:00 UT (peak minimum $Dst \leq -422$ nT). It was the greatest of all GMS events in SC 23. The source of this storm was a full halo and Earth-directed CME that occurred on 18 November 2003 at 08:50 UT. This very fast halo CME ($V = 1668$ km/s) was associated with an $M3.9$ class X-ray flare originating from the solar surface at $N0^0E18^0$. The storm was preceded by the presence of an ICME, first recorded on 20 November at 10:00 UT according to the RC catalogue.

Finally, Table 4.1 shows a good number of GMS events, including intense storms, without any association with halo CMEs. An example is storm event number 195 which occurred on 4 April 2004 with peak minimum *Dst* of -112 nT. Such storms may be caused by CIRs or their association with non-halo CMEs if they arrive near Earth with a southward pointing IMF B_z component (Gopalswamy et al., 2007). However, the catalogue of ICMEs at 1 AU by Richardson and Cane (2010) indicates that this storm event was associated with an ICME, starting on 3 April at 14:00 UT. This is one example among many other cases where an ICME was detected without any link to a halo CME in the time window. Such cases have also been reported by previous studies (e.g., Cane and Richardson, 2003). Hence, despite the fact that the near-Earth presence of an ICME is a potentially IP geoeffective property, there is no one-to-one association between halo CMEs and subsequent ICMEs (Richardson and Cane, 2008). Storm events without association with ICMEs are shown in column 10 of Table 4.1.

Table 4.1: GMS and associated probable solar sources and IP properties in SC 23

Event No.	**Storm time)**	**Dst min. (nT)**	**Bz (nT)**	**Vsw (Km/s)**	**CME /Time**	**V_cmes**	**XRFs**	**XRF_Loc.**	**ICME/start**
	Year 1996								
1	13/01/96 - 11:00	-90	-8.2	387	dg	-	-	-	—-
2	21/03/96 - 22:00	-66	-3.9	636	-	-	-	-	m
3	25/03/96 - 01:00	-60	-7.2	544	dg	-	-	-	—-
4	15/04/96 - 00:00	-56	-7.2	589	-	-	-	-	m
5	23/10/96 - 04:00	-105	-11	634	-	-	-	-	m
	Year 1997								
6	10/01/97 09:00	-78	-12.7	445	FH : 06/01 [15:10]	-	-	-	10/01 [04:00]
7	10/02/97- 11:00	-68	-7.6	456	FH: 07/02 [00:30]	490	-	-	10/02 [02:00]
8	27/02/97- 23:00	-86	-13.3	522	-	-	-	-	—-
9	11/04/97- 04:00	-82	-8.5	437	FH: 07/04 [14:27]	878	C6.8	S30E19	11/04 [06:00]
10	17/04/97 -05:00	-77	-9.3	499	-	-	-	-	m
11	21/04/97 - 23:00	-107	-9.6	409	-	-	-	-	21/04 [10:00]
12	15/05/97 - 12:00	-115	-23.9	447	FH: 12/05 [05:30]	464	C1.3	N21W08	15/05/ [09:00]
13	27/05/97 - 04:00	-73	-10.4	328	-	-	-	-	26/05/ [16:00]
14	09/06/97 -04:00	-84	-9.4	374	-	-	-	-	08/06 [18:00]
15	03/09/97 - 22:00	-98	-15.8	487	FH: 30/08 [01:30]	371	-	-	03/09 [13:00]
16	01/10/97 - 16:00	-98	-9.5	483	FH: 28/09 [01:08]	359	-	-	01/10 [16:00]
17	11/10/97- 03:00	-130	-10.3	422	PH:06/10 [15:28]	293	-	S27W05	10/10 [11:00]
18	25/10/97 - 03:00	-64	-9.1	428	FH: 21/10 [18:03]	523	C3.3	N16E07	27/10 [00:00]
19	07/11/97 - 04:00	-110	-12.5	456	FH: 04/11 [06:10]	785	X2.1	S14W33	07/11 [19:00]
20	23/11/97 - 06:00	-108	-12.8	500	FH: 19/11 [12:27]	150	C1.6	N24E05	22/11 [19:00]
21	11/12/97 -10:00	-60	-11.2	332	PH : 06/12 [22:06]	397	-	-	10/12 [18:00]
22	30/12/97 19:00	-77	-10.4	352	PH: 26/12 [02:31]	197	-	-	30/12 [10:00]
	Year 1998								
23	07/01/98 -04:00	-77	-9.9	405	FH : 02/01 [23:28]	-	-	-	07/01 [01:00]
24	28/01/98 12:00	-55	-7.4	374	FH : 25/01 [15:26]	693	C1.1	N21E25	29/01 [20:00]
25	18/02/98 - 00:00	-100	-15.1	409	PH: 14/02 [06:55]	123	-	-	17/02 [10:00]
26	10/03/98 - 20:00	-116	-14.8	529	-	-	-	-	m
27	21/03/98 - 15:00	-85	-12.4	429	PH: 18/03 [07:33]	413	-	-	m
28	25/03/98 - 16:00	-56	-7.2	401	-	-	-	-	25/03 [14:00]
29	24//04/98 - 07:00	-69	-8.9	435	FH: 23/04 [05:55]	1691	X1.2	-	m
30	02/05./98 17:00	-85	-11.6	596	FH : 29/04 [16:58]	1374	M6.8	S18E20	02/05 [05:00]
31	04/05/98 - 06:00	-205	-28.5	803	FH: 01/05 [23:40]	585	C2.8	N25E35	04/05 [10:00]
					FH: 02/05 [05:31]	542	C5.4	S20W07	
					FH : 02/05 [14:06]	938	X1.1	S15W15	
32	09/05/98 - 19:00	-63	-5.7	512	PH:06/05 [08:29]	1099	X2.7	S11W65	m
33	14/06/98 - 11:00	-55	-11.1	325	-	-	-	-	14/06 [04:00]
34	26/06/98 - 03:00	-101	-13	465	PH: 21/06 [05:35]	192	-	N15W30	26/06 [04:00]
35	16/07/98 - 15:00	-58	-10.6	578	dg	dg	-	-	
36	06/08/98 - 10:00	-138	-19.3	398	dg	dg	-	-	
37	07/08/98 05: 00	-108	-11.3		dg	dg	-	-	
38	20/08/98 - 20:00	-67	-10.2	342	dg	dg	-	-	
39	27/08/98 - 08:00	-155	-14.2	635	dg	dg	-	-	
40	25/09/98 - 07:00	-207	-17.9	797	dg	dg	-	-	
41	01/10/98 - 01:00	-58	-8.4	472	dg	dg	-	-	
42	7/10/98 - 22:00	-70	-10.5	580	dg	dg	-	-	
43	19/10/98 - 15:00	-112	-16.7	388	FH: 15/10 [10:04]	262	-	-	19/10 [04:00]
44	22/10/98 18:00	-53	-4	594	-	-	-	-	23/10 [15:00]
45	08/11/1998	-149	-19.9	623	FH:04/11 [07:54]	523	C1.6	N17W01	07/11 [22:00]
					FH:05/11 [20:44]	1118	M8.4	N22W18	
46	09/11/1998	-142	-15.2	436	FH:05/11 [20:44]	1118	M8.4	N22W18	09/11 [01:00]
					PH: 06/11 [02:18]	405	C4.4	N19W24	
47	13/11/98 - 18:00	-131	-17	360	PH:10/11 [06:18]	286	C3.3	N21W64	13/11 [02:00]
48	11/12/98 - 15:00	-69	-12.7	344	PH:10/12 [23:30]	591	B9.3	S22W51	m
49	28/12/98 11:00	-58	-6.9	404	dg	-	-	-	—-
	Year 1999								
50	13/01/99 - 23:00	-112	-14.6	418	dg	-	-	-	
51	23/01/99 22:00	-52	-4.9	541	dg	-	-	-	—-
52	18/02/99 - 09:00	-123	-21.8	635	PH:14/02 [11:08]	758	M1.0	N17E05	

Table 4.1: (continued)

E No.	Storm time (Dst min)	Dst (nT)	Bz (nT)	Vsw (Km/s)	CME /Time	V_cmes	XRFs	XRF_Loc.	ICME /start
53	01/03/99 19:00	-95	-15.1	456	-	-	-	-	m
54	10/03/99 08:00	-81	-9.5	429	-	-	-	-	—-
55	17/04/99 07:00	-91	-14	415	PH:13/04 [03:30]	291	-	-	16/04 [18:00]
56	31/07/99 01:00	-53	-10.6	625	FH:28/07 [09:06]	462	-	-	30/07 [20:00]
57	23/08/99 00:00	-66	-9.2	428	PH:20/08 [13:26]	209	M1.8	S28E76	20/08 [23:00]
58	13/09/99 04:00	-74	-7.4	566	PH:12/09 [00:54]	732	B9.9	-	m
59	16/09/99 08:00	-67	-6.1	579	PH:13/09 [09:30]	898	C4.9	N22E10	m
					PH: 13/09 [17:31]	444	C2.1	N17E75	
60	22/09/99 23:00	-173	-15.8	588	FH: 20/09 [06:06]	604	C2.8	S21W05	22/09 [19:00]
61	27/09/99 18:00	-64	-5.9	638	-	-	-	-	m
62	10/10/99 18:00	-67	-8.6	514	-	-	-	-	m
63	15/10/99 05:00	-67	-5.1	710	FH:14/10 [09:26]	1250	X1.8	N11E32	m
64	22/10/99 06:00	-237	-30.7	548	PH:18/10 [00:06]	144	-	-	21/10 [08:00]
					PH: 19/10 [05:30]	753	C2.9	-	
65	27/10/99 16:00	-66	-6.4	390	-	-	-	-	m
66	08/11/99 14:00	-73	-7.2	560	PH:05/11 [18:26]	378	M3.0	S19W45	m
67	13/11/99 22:00	-106	-11.5	443	dg	-	-	-	13/11 [20:00]
68	13/12/99 09:00	-85	-9.7	487	dg	-	-	-	12/12 [19:00]
	Year 2000								
69	11/01/00 21:00	-81	-15.1	524	-	-	-	-	m
70	22/01/00 03:00	-97	-15.7	388	FH :18/01 [17:04]	739	M3.4	S19E11	22/101 [17;00]
71	12/02/00 11:00	-133	-16.4	568	FH : 08/02 [09:30]	1079	M1.3	N25E26	12/02 [12:00]
					FH : 09/02 [19:54]	910	C7.4	S17W40	
72	31/03/00 -11:00	-60	-7.2	396	FH :29/03 [10:54]	949	C1.8	S17E75	29/03 [19:00]
73	05/04/00 01:00	-63	-7.6	376	-	-	-	-	m
74	07/04/00 00:00	-288	-27.3	571	FH:04/04 [16:32]	1188	C9.7	N16W66	07/04 [06:00]
75	16/04/00 11:00	-79	-8.2	393	-	-	-	-	m
76	24/04/00 14:00	-61	-11	480	FH: 23/04 [12:54]	1187	-	-	24/04 [04:00]
77	17/05/00 05:00	-92	-9.4	505	PH: 15/05 [08:50]	1549	M4.4	N21E75	16/05 [23:00]
					PH: 15/05 [16:26]	1212	M1.1	N21E73	
78	24/05/00 08:00	-147	-19.2	642	FH: 22/05 [01:50]	649	C6.3	N20W22	24/05 [12:00]
79	08/06/00 19:00	-90	-6.9	760	FH: 06/06 [15:54]	1119	X2.3	N20E18	08/06 [12:00]
80	26/06/00 -17:00	-76	-9.2	540	PH: 23/06 [14:54]	847	M3.0	N26W72	26/06 [12:00]
					PH: 25/06 [07:54]	1617	M1.9	N16W55	
81	16/07/00 - 00:00	-301	-49.4	1030	FH: 14/07 [10:54]	1674	X5.7	N22W07	15/07 [19:00]
82	20/07/00 - 09:00	-93	-7.9	563	-	-	-	-	20/07 [01:00]
83	23/07/00 - 22:00	-68	-10.2	379	PH: 22/07 [11:54]	1230	M3.7	N14W56	23/07 [15:00]
84	29/07/00 - 11:00	-71	-10	458	FH: 25/07 [03:30]	528	M8.0	N06W08	28/07 [12:00]
85	06/08/00 - 05:00	-56	-3.7	553	-	-	-	-	m
86	11/08/00- 06:00	-106	-13.2	426	PH: 08/08 [15:54]	867	C1.4	-	10/08 [19:00]
87	12/08/00 - 09:00	-235	-28.7	613	FH: 09/08 [16:30]	702	C2.3	N11W11	12/08 [05:00]
88	29/08/00 - 06:00	-60	-6	602	PH: 25/08 [14:54]	518	M1.4	S15E67	m
					PH: 26/08 [21:30]	326	C3.8	N26E02	
89	02/09/00 - 14:00	-57	-6.6	444	-	-	-	-	02/09 [19:00]
90	12/09/00 - 19:00	-73	-9.5	395	PH:09/09/ [08:56]	554	M1.6	N07W67	m
91	16/09/00 - 23:00	-68	-15.9	401	FH:12/09 [11:54]	1550	M1.0	S17W09	m
92	18/09/00 - 23:00	-201	-23.9	794	FH: 15/09 [21:50]	257	C7.4	N12E04	17/09 [21:00]
					FH: 16/09 [05:18]	1215	M5.9	N14W07	
93	26/09/00 - 02:00	-55	-5.6	568	-	-	-	-	m
94	30/09/00 - 14:00	-76	-11.5	408	FH: 25/09 [02:50]	587	M1.8	S11W58	m
95	05/10/00 - 13:00	-182	-20.2	523	FH: 02/10 [03:50]	525	-	-	03/10 [10:00]
					FH: 02/10 [20:26]	569	C8.4	N01E80	
96	14/10/00 - 14:00	-107	-11.5	413	FH: 09/10 [23:50]	798	C6.7	N01W14	13/10 [16:00]
97	29/10/00 - 03:00	-127	-17.1	403	FH: 25/10 [08:26]	770	C4.0	N06W60	28/10 [21:00]
98	06/11/2000-[21:00]	-159	-11.7	570	FH: 03/11 [18:26]	291	C3.2	N02W02	06/11 [17:00]
99	10/11/00 - 12:00	-96	-8	881	-	-	-	-	m
100	27/11/00 01:00	-80	-10.8	517	FH : 24/11 [15:30]	1245	X2.3	N22W07	27/11 [08:00]
					FH : 24/11 [22:06]	1005	X1.8	N21W14	
101	29/11/00 - 13:00	-119	-10.3	486	FH: 25/11 [09:30]	675	M8.2	N18W24	28/11 [11:00]
					FH: 25/11 [17:31]	671	M3.5	**N20W23**	
					FH: 26/11 [17:06]	980	X4.0	N18W38	

Table 4.1: (continued)

E No.	**Storm time (Dst min)**	**Dst (nT)**	**Bz (nT)**	**Vsw (Km/s)**	**CME /Time**	**V_cmes**	**XRFs**	**XRF_Loc.**	**ICME /start**
102	23/12/00 - 04:00	-62	-13.9	323	PH: 20/12 [10:30]	1076	-	-	23/12 [00:00]
					PH: 20/12 [21:30]	609	C3.5	S10E48	
	Year 2001								
103	24/01/01-18:00	-61	-6.8	436	FH:20/01 [19:31]	839	M1.2	S07E40	24/01 [09:00]
					FH:20/01 [21:30]	1507	M7.7	S07E46	
104	05/03/01 - 02:00	-73	-12.5	435	PH: 01/03 [18:26]	631	C1.2	S09W27	04/03 [04:00]
105	20/03/01 - 13:00	-149	-18.8	388	FH: 16/03 [03:50]	389	-	-	19/03 [17:00]
106	23/03/01- 16:00	-75	-7.5	413	PH: 20/03 [03:26]	478	-	-	m
107	28/03/01 - 15:00	-87	-6.8	606	FH: 24/03 [20:50]	906	M1.7	N15E22	27/03 [20:00]
					FH: 25/03 [17:06]	677	C9.0	N16E25	
108	31/03/01 - 08:00	-387	-44.7	644	FH:28/03 [12:50]	519	M4.3	N18E02	31/03 [05:00]
					FH: 29/03 [10:26]	942	X1.7	N20W19	
109	04/04/01 07:00	-50	-5.5	645	PH : 02/04 [22:06]	2505	X20.0	-	04/04 [18:00]
110	09/04/01 - 06:00	-63	-4.7	644	FH:05/04 [17:06]	1390	M5.1	S24E50	08/04 [14:00]
					FH:06/04 [19:30]	1270	X5.6	S21E31	
111	11/04/01 23:00	-271	-20.5	725	FH:09/04 [15:54]	1192	M7.9	S21W04	11/04 [22:00]
					FH:10/04 [05:30]	2411	X2.3	S23W09	
112	18/04/01- 06:00	-114	-19.6	512	PH:15/04 [14:06]	1199	X14.4	S20W85	18/04 [12:00]
113	22/04/01- 15:00	-102	-12.8	351	-	-	-	-	21/04 [23:00]
114	10/05/01 - 03:00	-76	-8.4	426	PH:07/05 [12:06]	1223	C3.9	N25W35	09/05 [12:00]
115	18/06/01- 08:00	-61	-11.6	363	FH:15/06 [15:56]	1701	C2.2	S16E18	m
116	17/08/01 21:00	-105	-18.1	519	FH:14/08 [16:01]	618	C2.3	N16W36	17/08 [20:00
117	13/09/01 07:00	-57	-7.2	395	FH : 11/09 [14:54]	791	C3.2	N13E35	13/09 [18:00]
118	23/09/01- 18:00	-73	-9	525	PH : 20/09 [19:31]	446	C3.2	S17E54	24/09 [00:00]
119	26/09/01 -01:00	-102	-6.4	627	FH:24/09 [10:30]	2402	X2.6	S16E23	m
120	01/10/01 - 08:00	-148	-12.7	481	FH: 28/09 [08:54]	846	M3.3	N10E182N	01/10 [08:00]
121	03/10/01 - 14:00	-166	-20.9	520	PH:09/29 [11:54]	509	M1.8	N13E03	02/10 [04:00]
122	12/10/01- 12:00	-71	-13.2	479	FH:09/10 [11:30]	973	M1.4	S28E08	12/10 [04;00]
123	19/10/01- 21:00	-57	-7.4	340	-	-	-	-	m
124	21/10/01- 21:00	-187	-16.4	608	FH: 19/10 [01:27]	558	X1.6	N16W18	21/10 [20:00]
					FH: 19/10 [16:50]	901	X1.6	N15W29	
125	28/10/01 - 11:00	-157	-14.5	474	FH: 25/10 [15:26]	1092	X1.3	S16W21	29/10 [22:00]
126	01/11/01 - 10:00	-106	-13	348	-	-	-	-	31/10 [20:00]
127	06/11/01 - 06:00	-292	-64	729	FH: 03/11 [19:20]	457	-	-	05/11 [19:00]
					FH: 04/11 [16:35]	1810	X1.0	N06W18	
128	24/11/01 - 16:00	-221	-27.8	1034	FH:21/11 [14:06]	518	C4.7	S14W19	24/11 [14:00]
					FH: 22/11 [20:30]	1443	M3.8	S25W67	
					FH: 22/11 [23:30]	1437	M9.9	S14W36	
129	21/12/01 - 22:00	-67	-9.1	423	-	-	-	-	m
130	30/12/01 05:00	-58	-9.7	382	PH : 26/12 [05:30]	1446	M7.1	N08W54	30/12 [00:00]
	Year 2002								
131	11/01/02 - 06:00	-72	-4.5	630	FH: 08/01 [17:54]	1794	C7.2	S18W42	m
132	02/02/02 - 09:00	-86	-12.7	370	-	-	-	-	m
133	05/02/02 20:00	-82	-7.7	523	PH:02/02 [15:54]	362	C2.2	S18W13	m
134	01/03/02 - 01:00	-71	-14.6	390	-	-	-	-	28/02 [17:00]
135	24/03/02 - 09:00	-100	-9.4	421	FH: 20/03 [17:54]	603	C4.0	S21W15	24/03 [12:00]
					PH: 19/03 [11:54]	860	M1	S10W58	
136	18/04/02 - 07:00	-127	-12.8	504	FH: 15/04 [03:50]	720	M1.2	S15W01	17/04 [16:00]
137	20/04/02 - 08:00	-149	-14.7	611	FH:17/04 [08:26]	1240	M2.6	S14W34	20/04 [00:00]
138	11/05/02 - 19:00	-110	-16.5	437	FH:08/05 [13:50]	614	C4.2	S12W07	11/05 [15:00]
139	15/05/02 - 00:00	-65	-6.8	412	-	-	-	-	m
140	23/05/02 - 17:00	-109	-14.1	871	FH: 22/05 [03:50]	1557	C5.0	S22W53	23/05 [20:00]
141	27/05/02 - 09:00	-64	-12	553	PH:25/05 [02:50]	880	C1.7	S18W13	m
142	02/08/02 - 05:00	-102	-12.5	524	PH:29/07 [12:07]	562	-	-	01/08 [09:00]
					PH:01/08 [04:06]	375	C2.5	S10E12	
143	21/08/02 - 02:00	-106	-9.2	434	FH:16/08 [12:30]	1585	M5.2	S14E20	19/08 [12:00]
144	04/09/02 - 05:00	-109	-18.3	514	-	-	-		m
145	08/09/02 - 00:00	-181	-21.5	520	FH: 05/09 [16:54]	1748	C5.2	N09E28	08/09 [04:00]
					FH: 06/09 [13:31]	909	C2.0	NA	
146	01/10/02 - 16:00	-176	-21.8	379	PH: 09/30 [01:31]	307	M2.1	N13E10	30/09 [20:00]
147	04/10/02 - 08:00	-146	-11.8	388	**PH: 02/10 [07:31]**	**903**	**B9.2**	**S18E20**	03/10 [01:00]
148	07/10/02 -07:00	-115	-8.4	371	PH: 04/10 [20:30]	743	C4.8	N12E46	—-

Table 4.1: (continued)

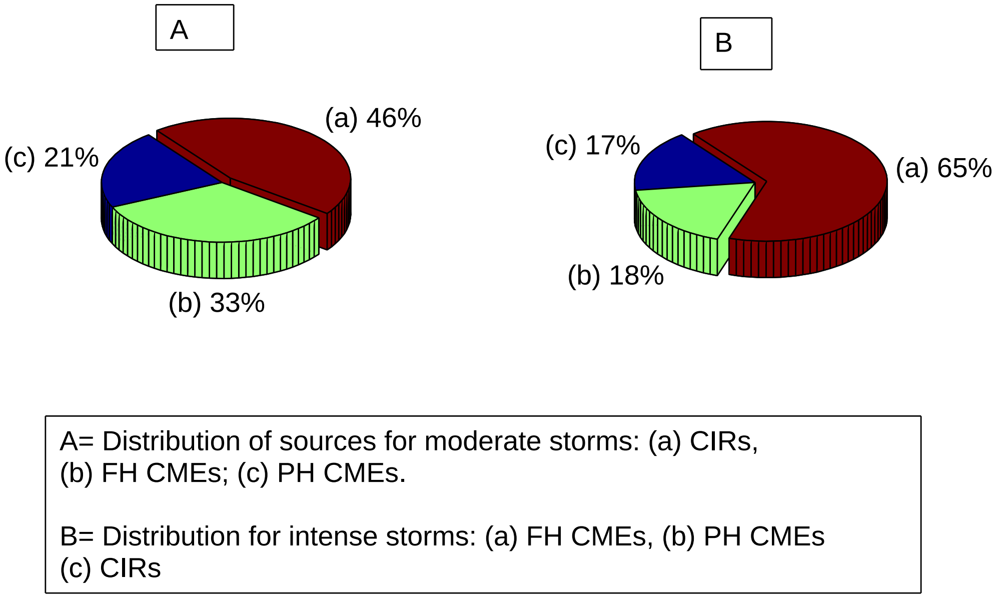
E No.	Storm time (Dst min)	Dst (nT)	Bz (nT)	Vsw (Km/s)	CME /Time	V_cmes	XRFs	XRF_Loc.	ICME /start
149	14/10/02 - 13:00	-100	-13.6	355	-		-	-	m
150	24/10/02 - 19:00	-94	-7.4	732	-	-	-	-	m
151	27/10/02 - 16:00	-65	-4.3	608	-	-	-	-	m
152	31/10/02 - 18:00	-52	-5.3	437	FH:27/10 [23:18]	2115	-	-	m
153	3/11/2002 - 06:00	-75	-6.3	480	-	-	-	-	m
154	18/11/02 - 22:00	-52	-7.7	379	FH: 16/11 [07:12]	1185	C1.9	S19W18	17/11 [10:00]
155	21/11/02 - 10:00	-128	-13.1	656	-	-	-	-	m
156	27/11/02 - 06:00	-64	-4.8	544	FH:24/11 [20:30]	1077	C6.4	S17W37	m
157	21/12/02 - 03:00	-75	-8.1	495	-	-			21/12 [23:00]
158	23/12/02 - 11:00	-67	-5.6	581	FH: 19/12 [22:06]	1092	M2.7	N15W08	m
159	27/12/02 - 04:00	-68	-8.1	638	-	-	-	-	m
	Year 2003								
160	30/01/03 - 00:00	-66	-8.8	438	PH:27/01 [22:23]	1053	C2.4	S17W23	m
161	04/02/03 - 09:00	-74	-8.2	593	PH : 30/01 [10:06]	325	B4.8	S11E52	01/02 [19:00]
162	27/02/03 - 21:00	-66	-5.8	509	-	-	-	-	m
163	04/03/03 - 07:00	-67	-12.4	524	-	-	-	-	m
164	16/03/03 - 21:00	-60	-6.1	685	PH:14/03 [18:06]	991	B9.2	S20W66	m
165	20/03/03 - 19:00	-64	-7.3	645	FH: 19/03 [02:30]	1342	M1.5	S15W54	20/03 [12:00]
166	04/04/03 - 23:00	-62	-6.5	478	-	-	-	-	m
167	01/05/03 - 00:00	-78	-8.1	634	-	-	-	-	m
168	10/05/03 - 08:00	-84	-7.4	625	-	-	-	-	09/05 [07:00]
169	22/05/03 - 02:00	-73	-10.2	522	-	-	-	-	m
170	29/05/03 - 23:00	-144	-12.7	737	FH:27/05 [06:50]	509	M1.6	S07W14	29/05 [13:00]
					FH: 28/05 [00:50]	1366	X3.6	S11W11	
					FH: 27/05 [23:50]	964	X1.3	S07W17	
171	02/06/03 - 08:00	-91	-8.9	716	FH: 31/05 [02:30]	1835	M9.3	S07W65	m
172	18/06/03 - 09:00	-141	-16.7	524	FH:15/06 [23:54]	2053	X1.3	S07E80	17/06 [10:00]
					PH:14/06 [05:30]	1215	M1.5	S06E89	
173	12/07/03 - 05:00	-105	-13.2	589	-	-	-	-	m
174	16/07/03 - 13:00	-90	-10.4	569	-	-	-	-	m
175	06/08/03 06:00	-60	-11	483	-	-	-	-	04/08 [22:00]
176	07/08/03 - 21:00	-61	-9.5	601	FH:03/08 [00:30]	699	M1.3	S17E63	m
177	18/08/03 - 15:00	-148	-15.9	437	FH: 14/08 [20:06]	378	-	S30E00	18/08 [01:00]
178	21/08/03 - 06:00	-68	-4.6	568	-	-	-	-	m
179	17/09/03 - 23:00	-65	-6.2	747	-	-	-	-	m
180	14/10/03 - 22:00	-85	-8.6	619	-	-	-	-	m
181	22/10/03 - 06:00	-61	-4.7	688	FH:18/10 [15:30]	627	C3.3	N07E72	22/10 [02:00]
182	27/10/03 - 04:00	-52	-7.7	476	PH:26/10 [06:54]	1371	X1.2	S15E44	25/10 [14:00]
					PH:26/10 [17:54]	1537	X1.2	N02W38	
183	30/10/03 - 22:00	-383	-27.1	1161	FH:28/10 [11:30]	2459	X17.2	S16E08	29/10 [11:00]
					FH:29/10 [20:54]	2029	X10.0	S15W02	
184	04/11/03 - 10:00	-69	-11.9	739	FH:02/11 [17:30]	2598	X8.3	S14W56	m
184	11/11/03 - 13:00	-62	-6	710	-	-	-	-	m
186	13/11/03 - 20:00	-59	-3.8	656	FH: 11/11/ [13:54]	1315	M1.6	S03W61	m
187	20/11/03 - 20:00	-422	-50.9	553	FH:18/11 [08:50]	1668	M3.9	N03E18	20/11 [10:00]
188	06/12/2003 04:00	-55	-8	533	-	-	-	-	m
189	08/12/03 21:00	-54	-7.9	616	-	-	-	-	m
	Year 2004								
190	07/01/04 - 09:00	-69	-6.2	736	-	-	-	-	m
191	10/01/04 - 08:00	-60	-7	551	PH: 07/01 [10:30]	1822	M8.3	N06E75	10/01 [06:00]
					PH: 08/01 [05:06]	1713	M1.3	N03E63	
192	22/01/04 - 13:00	-149	-14.9	614	FH:20/01 [0:06]	965	C5.5	S13W11	22/01 [08:00]
193	11/02/04 - 17:00	-109	-13.6	385	-	-	-	-	m
194	10/03/04 - 08:00	-77	-9.5	765	-	-	-	-	m
195	04/04/04 - 01:00	-112	-7.9	506	-	-	-	-	03/04 [14:00]
196	05/04/04 - 19:00	-81	-15.7	419	-	-	-	-	m
197	17/07/04 - 02:00	-80	-14.1	468	FH: 13/07 [09:30]	747	M5.4	N14W51	m

Table 4.1: (continued)

E No.	Storm time (Dst min)	Dst (nT)	Bz (nT)	Vsw (Km/s)	CME /Time	V_cmes	XRFs	XRF_Loc.	ICME /start
198	23/07/04 - 02:00	-101	-15.5	643	FH: 20/07 [13:31]	710	M8.6	N10E35	22/07 [18:00]
199	25/07/04 - 11:00	-148	-17	593	**PH: 22/7 [08:30]**	**899**	**C5.3**	**N02E08**	24/07 [14:00]
					FH: 23/07 [16:06]	824	C1.0	N05W04	
200	27/07/04 - 13:00	-197	-19.9	831	FH: 25/07 [14:54]	1333	M1.1	**N08W33**	27/07 [02:00]
201	30/08/04 - 22:00	-126	-14.3	416	-	-	-	-	29/08 [19:00]
202	14/09/04 11:00	-50	-5.5	580	FH : 12/09 [00:36]	1328	M4.8	N03E49	14/09 [15:00]
203	13/10/04 - 13:00	-63	-8.6	470	-	-	-	-	m
204	08/11/04 06:00	-373	-44.9	712	FH: 04/11 [09:54]	653	C6.3	N09E28	07/11 [22:00]
					FH: 06/11 [01:31]	818	M3.6	N10E06	
205	12/11/04 - 10:00	-109	-3.8	588	FH: 07/11 [16:54]	1759	X2.0	N09W17	12/11 [08:00]
					FH: 10/11 [02:26]	3387	X2.5	N09W49	
206	21/11/04 - 11:00	-60	-3.5	601	-	-	-	-	m
207	25/11/04 - 07:00	-63	-6.8	498	-	-	-	-	m
208	06/12/04 - 15:00	-58	-7.8	439	FH: 03/12 [0:26]	1216	M1.5	N08W02	m
209	13/12/04 - 04:00	-61	-9.8	395	FH : 08/12 [20:26]	611	M1.2	S08E66	12/12 [22:00]
210	29/12/04 - 10:00	-55	-7.6	420	-	-	-	-	m
	Year 2005								
211	01/01/05 - 19:00	-57	-9.4	459	FH: 30/12 [22:30]	1035	M4.2	N04E46	m
212	08/01/05 - 02:00	-96	-18.5	548	-	-	-	-	08/01 [22:00]
213	12/01/05 - 10:00	-57	-10.9	598	PH : 09/01 [09:06]	870	M2.4	S09E70	m
214	18/01/05 - 08:00	-121	-15.5	957	FH: 15/01 [06:30]	2049	M8.6	N16E04	18/01 [23:00]
					FH: 15/01 [23:06]	2861	X2.6	N15W05	
215	22/01/05 - 06:00	-105	-6.3	814	FH: 19/01 [08:29]	2020	X1.3	N15W51	21/01 [19:00]
					FH: 20/01 [06:54]	882	X7.1	N14W61	
216	07/02/05 - 21:00	-62	-6.6	682	-	-	-	-	216
217	18/02/05 - 02:00	-86	-15.9	436	-	-	-	-	18/02 [14:00]
218	06/03/05 - 16:00	-65	-6.5	610	-	-	-	-	m
219	05/04/05 - 05:00	-85	-6.5	689	-	-	-	-	m
220	12/04/05 - 05:00	-70	-8.5	492	PH: 09/04 [08:26]	329	B2.6	S07W63	m
					PH: 09/04 [13:50]	514	B2.1	S07W66	
221	08/05/05 - 18:00	-127	-13	755	FH: 05/05 [20:30]	1180	C7.8	S04W67	m
222	15/05/05-08:00	-263	-38	835	FH: 11/05 [20:13]	550	M1.1	S11W51	15/05 [06:00]
					FH: 13/05 [17:12]	1689	M8.0	N12E11	
223	20/05/05 - 08:00	-103	-9.1	478	PH: 16/05 [13:50]	405	C1.2	N13W29	20/05 [03:00]
					PH: 17/05 [03:26]	449	M1.8	S15W00	
224	30/05/05 - 13:00	-138	-161	451	FH: 26/05 [15:06]	586	B7.5	S12E13	30/05 [01:00]
225	13/06/05 - 00:00	-106	-16.8	464	PH: 09/06 [14:36]	377	C1.3	N07E13	12/06 [15:00]
226	15/06/05 - 12:00	-54	-6.9	507	-	-	-	-	15/06 [05:00]
227	23/06/05 - 10:00	-97	-17.2	385	-	-	-	-	m
228	09/07/05 - 18:00	-60	-9.2	338	FH: 05/07 [15:30]	772	C1.3	S08E34	m
229	10/07/05 - 20:00	-94	-19.2	438	FH: 09/07 [22:30]	1540	M2.8	N12W28	10/07 [10:00]
230	18/07/05 - 06:00	-76	-8.8	418	FH: 14/07 [10:54]	2115	X1.2	N11W90	17/07 [17:00]
231	10/08/05 11:00	-53	-8.7	433	-	-	-	-	10/08 [06:00]
232	24/08/05 - 11:00	-216	-38.3	620	FH: 22/08 [01:31]	1194	M2.6	S11W54	24/08 [14;00]
					FH: 22/08 [17:30]	2378	M5.6	S13W65	
233	31/08/05 - 19:00	-131	-15.8	415	-	-	-	-	m
234	11/09/05 - 09:00	-147	-8.5	974	FH: 09/09 [19:48]	2257	X6.2	S12E67	11/09 [05:00]
235	15/09/05 - 16:00	-86	-4.8	862	FH: 13/09 [20:00]	1866	X1.5	S09E10	15/09 [06:00]
236	31/10/05 - 19:00	-75	-8.9	361	-	-	-	-	31/10 [02:00]
	Year 2006								
237	05/04/06 - 15:00	-87	-12.2	380	-	-	-	-	m
238	09/04/06 - 07:00	-80	-15.5	437	-	-	-	-	m
239	14/04/06 - 09:00	-111	-14.2	518	PH: 10/04 [06:06]	184	B4.4	S12W22	14/04 [13:00]
240	19/08/06 - 21:00	-71	-11.5	405	FH: 16/08 [16:30]	888	-	-	20/08 [13:00]
241	24/09/06 - 09:00	-56	-7.4	656	-	-	-	-	m
242	10/11/06 - 06:00	-51	-12.5	471	FH: 06/11 [17:54]	1994	C8.8	N00E89	m
243	30/11/06 - 13:00	-74	-9.8	412	-	-	-	-	29/11 [05:00]
244	15/12/06 - 07:00	-146	-14.7	737	FH: 13/12 [02:54]	1774	X3.4	S06W23	14/12 [22:00]

4.3 Statistical results

A total of 244 GMS events were identified over the 11-year period of SC 23. Note that the period between 2006 until the end of solar SC 23 (2009) was magnetically quiet and therefore of no particular interest for this analysis. Within the period of study, 90 (about 37%) intense GMS events ($Dst < -100$ nT) and 154 (about 63%) moderate GMS events (-100 nT$\leq$ $Dst <$ -50 nT) were investigated to determine the possible solar causes and associated IP characteristics. Fifteen storm events were excluded, due to missing SOHO/LASCO data record, mainly for the period July, August and September 1998, and for January 1999. The missing data are indicated as data gap (dg) in Table 4.1. Therefore, the analysis of solar sources concerned only 229 GMS events. The two pie charts in Figure 4.3 illustrate the percentage of the main GMS sources responsible for moderate and intense storms.



A= Distribution of sources for moderate storms: (a) CIRs, (b) FH CMEs; (c) PH CMEs.

B= Distribution for intense storms: (a) FH CMEs, (b) PH CMEs (c) CIRs

Figure 4.3: Distribution of probable causes of GMS in SC 23: For moderate GMS (A), (a) represents the percentage of GMS not associated with halo CMEs (CIRs driven), (b) the percentage of storms associated with full halo CMEs and (c) the percentage associated with partial halo CMEs. For intense storms (B), (a) represents the percentage of storms associated with full halo CMEs, while (b) and (c) represent the percentage associated with partial and non-halo CMEs respectively.

4.3.1 Statistics for CME-driven intense and moderate GMS

Of the 229 GMS events for which the solar sources were investigated, 84 were intense storms with an average storm strength of about $Dst = -154$ nT and 145 were moderate storms with an average $Dst = -70$ nT. Seventy out of the 84 intense storms (about 83%) were CME-driven. As far as the 145 moderate storms are concerned, 79 of them (about 54%) were halo CME-driven storms, and for the remaining 66 (about 46%), no halo CME was identified in the time window considered. The non-halo CME-driven GMS can be associated with CIRs originating from coronal holes (CHs). Indeed, a large number of moderate storms were found to be CIR-driven or simply non-halo associated storms. The possibility of non-halo CMEs producing moderate storms has been reported by Gopalswamy et al. (2007). Figure 4.3 illustrates the distribution of probable sources for both moderate and intense storms in SC 23 showing the percentage of storms that were associated with full, partial and non-halo CMEs respectively. Note that the statistics summarised in Figure 4.3 do not refer to geoeffective halo CMEs as probable sources of the storms, but to the fraction of GMS that was most probably caused by full halo, partial halo and CIRs.

With regard to the association of GMS with ICMEs, the present analysis found that 72 of 84 intense GMS events (about 86%) were associated with ICMEs at 1 AU, compared to moderate storms for which only 44% were associated with ICMEs. Note that Table 4.1 contains 92 GMS events not listed in the RC table. These events noted as m in column 10 are obviously among the storms non associated with ICMEs and they are mostly moderate storms (81 out of 92 or 88%). The magnitudes of the SW speed and B_s associated with GMS events are characteristic of moderate and intense storms. As illustrated by Figure 4.4, intense storms were associated with higher average values of B_s and SW speed. Table 4.2 presents a simplified quantitative classification of solar sources for both moderate and intense GMS events during SC 23.

An investigation was carried out to quantitatively show the association of X-ray flares with the identified halo CME-driven GMS. The reason and criteria followed in this analysis were explained in Section 4.2.3. This analysis identified 64 intense and 62 moderate storms that were clearly associated with flare activity within ±0.5 hours of the corresponding CME eruption. Table 4.3 provides the X-ray flare distribution associated with 126 GMS events (including intense and moderate storms). Table 4.3 indicates that the majority of storms were associated with C and M class flares without a significant difference of association between intense and moderate storms.

As far as the solar flare surface coordinates are concerned, a full information about solar flares was available corresponding to 72 associated halo CMEs. Among them 50 events 69% originated close to the solar disk centre [within $\pm 45^0$ of the central meridian distance (CMD)

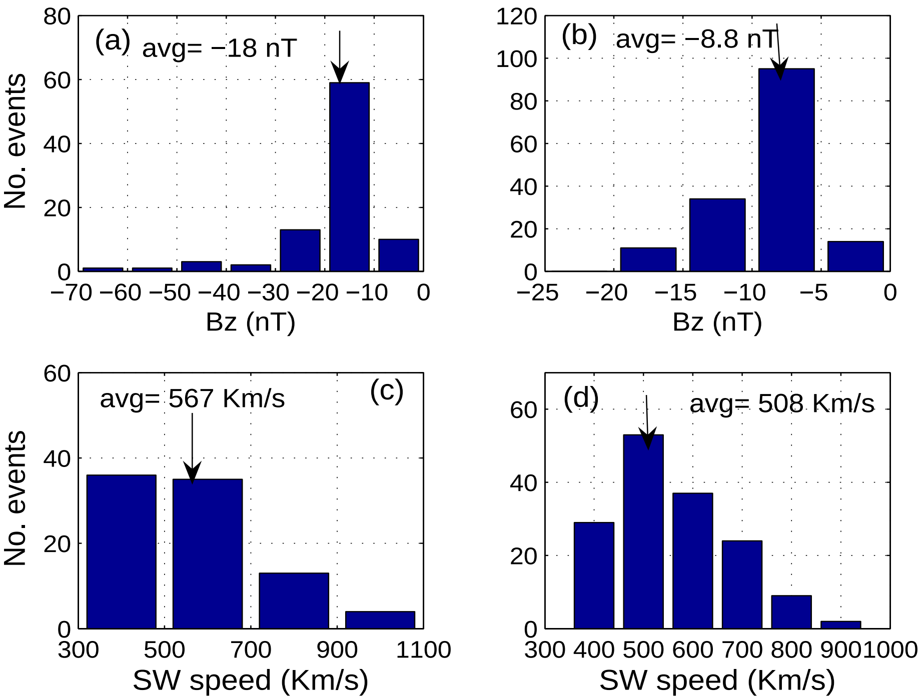


Figure 4.4: Distribution of the magnitudes of IMF Bz and SW speed that were associated with intense GMS, as shown in plots (a) and (c) as compared to those for moderate GMS, shown in plots (b) and (d). The calculated average (avg) values of the parameters are shown by arrows in the plots.

Storm category	**No.of GMS**	**FH-CME**	**PH-CME**	**CIRs**	**ICMEs**
Intense storms	84	55 [65%]	15 [18%]	14 [17%]	72[86%]
Moderate storms	145	48 [33%]	31 [21%]	66 [46%]	64 [44%]
Total	229	103 [45%]	46 [20%]	80[35%]	136 [91%]

Table 4.2: Distribution of GMS sources during SC 23: Columns 3,4,5 indicate the percentage of storms associated with full halo (FH), partial halo (PH) and CIR driven storms respectively. Column 6 shows the percentage of GMS association with ICMEs for both intense and moderate GMS.

interval]. For moderate storms, the flaring source was clearly identified for 61 events, of which only 31 (about 51%) were found to be associated with CMEs originating close to the disk center. These results indicate that generally, intense storms were mostly associated with halo CMEs originating closer to the central merdian compared to the CMEs at the origin of moderate storms. As indicated by Gopalswamy (2009b), moderate and weak storms are mostly produced by partial halo CMEs which generally originate far from the disk center.

Table 4.3: Classification of 126 X-ray flares associated with halo CME-driven GMS in SC 23

Flare class	**B-class**	**C-class**	**M-class**	**X-class**
Number	9	46	48	23
Percentage	7%	37%	38%	18%

4.3.2 Full and partial halo CME-driven storms

Table 4.2 shows that 46 events (31%) of the 149 halo CME-driven storms were uniquely associated with partial halo CMEs. In order to compare solar and IP geoeffective parameters accompanying full and partial halo CMEs, an equivalent sample of 46 full halo CME-driven GMS was selected from the period 1996-2001. Comparative results show that partial halo CME-driven storms were predominantly moderate (average $Dst = -92$ nT), while intense storms (average Dst of -128 nT) were mainly associated with full halo CMEs. Figures 4.5 and 4.6 compare geoefffective parameters associated with both full and partial halo CMEs. This analysis also shows that 42 of the 46 (91%) full halo CME-driven storms were associated with ICMEs. For partial halo CME-driven storms, only 71% were associated with ICMEs at 1 AU. Information from Table 4.2 shows that 80 (35%) storm events were not CME-driven. However, the non-halo CME-driven storms were mostly moderate storms and 59 of them (74%) were not associated with ICMEs at 1 AU. This means that 26% of the non CME-driven storms were associated with ICMEs.

Flare information could be identified for 36 of the 46 partial halo CMEs that produced storms and 38 for full halo geoeffective CMEs. Figures 4.5 and 4.6 provide a comparative X-ray flare association between partial and full halo CME-driven storms. Full halo CMEs were generally accompanied by powerful M-class flares, while the majority of flares accompanying geoeffective partial halo CMEs generally fell into class C. In addition, this analysis indicates that 16 of the 36 (44%) partial halo CMEs had associated solar flare surface coordinates close to the disk centre, while 32 out of 38 (or 84%) full halo CMEs originated close to the disk centre. These results are consistent with observations by Gopalswamy (2009b) who noticed that partial halo CMEs demonstrate moderately geoeffective properties,due to the fact that they are less energetic and originate far from the disk centre.

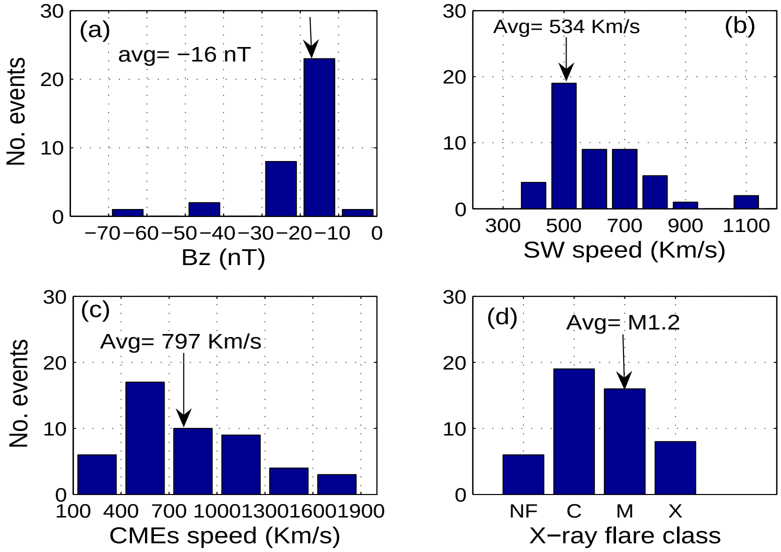


Figure 4.5: Distribution of the magnitudes of IMF Bz, SW speed, CME speed and X-ray flares that were associated with full halo CME-driven storms. Average values of the parameters are indicated by arrows.

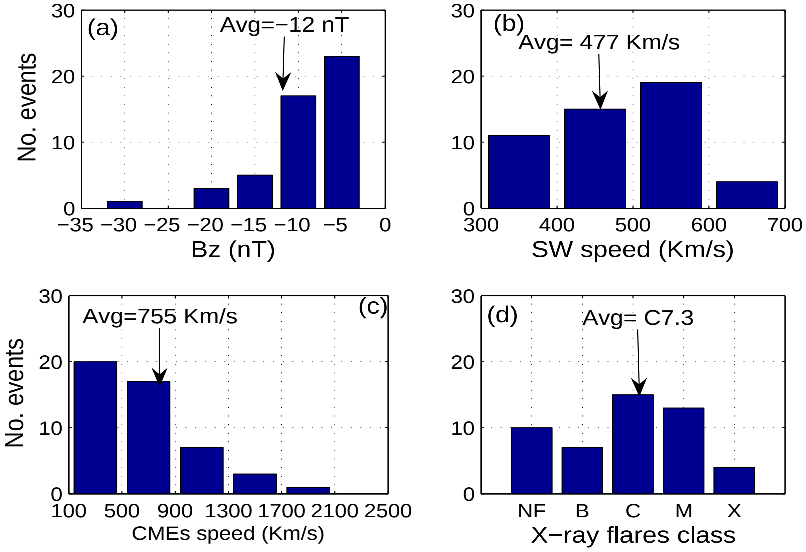


Figure 4.6: Distribution of the magnitudes of IMF Bz, SW speed, CME speed and X-ray flares that were associated with partial halo CME-driven storms.

4.3.3 Multiple halo CME-driven storms

As mentioned before, it is common for a storm to be associated with more than one halo CME. However, in such cases of interacting CMEs, it is difficult to identify a one-to-one CME-storm relationship (Kim et al., 2010). In this analysis 39 (26%) out of all the 149 CME-driven storms (see Table 4.1) were identified of as multiple CME-driven storms. The majority of these multiple CME-driven storms are intense storms [27 of 39 or 69%]. Figure 4.7 shows that multiple CME-driven storms are generally associated with high average values of negative Bz, SW speed and CME speed. On the other hand, most of the multiple CME-driven storms [36 out of 39 or 92%] were found to be associated with ICMEs at 1 AU. Earlier studies (Xie et al., 2006; Gopalswamy, 2006b) found a linear correlation between the intensity and duration of GMS with interacting CMEs. However, the strength of a storm may not always be associated with multiple halo CMEs as illustrated in Figure 4.8. The strongest of all magnetic storms in SC 23 was a single CME-driven storm.

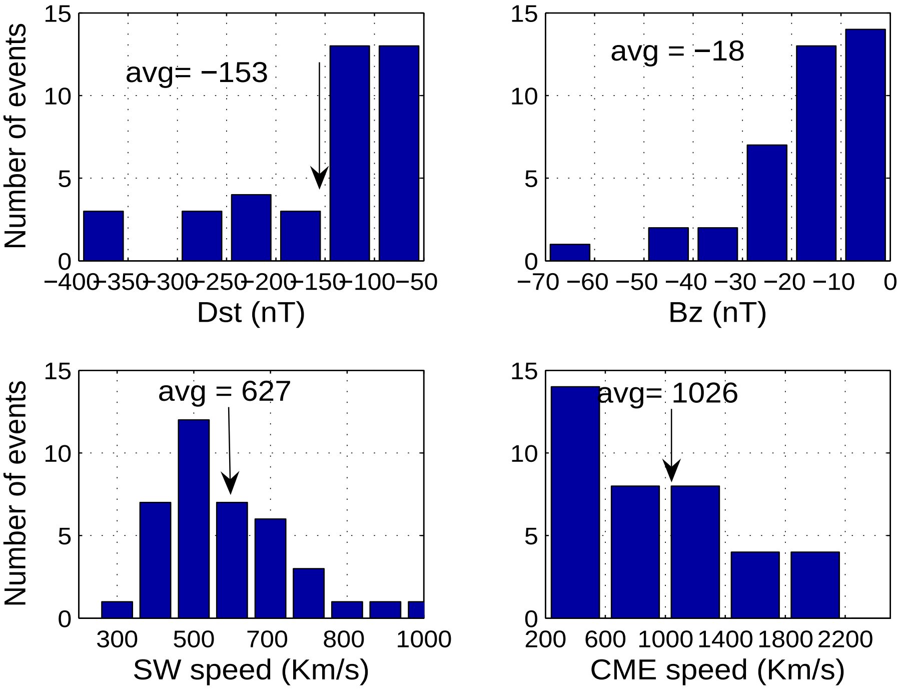


Figure 4.7: Distribution characterising geoeffective parameters that were associated with multiple halo CME-driven GMS: the magnitudes of Dst minimum, negative Bz, SW and CME speeds.

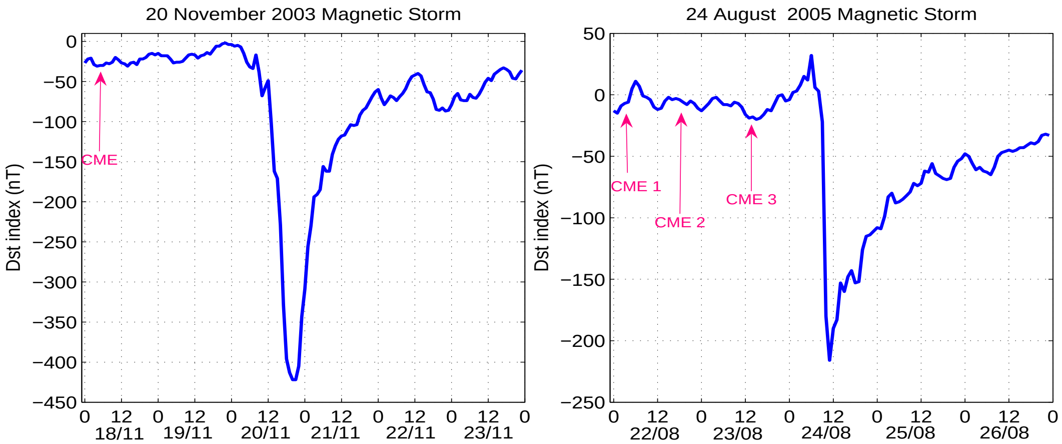


Figure 4.8: Example of a single (left) and a multiple (right) CME-driven intense storm.

4.3.4 Trends in SC 23

The SC trends of GMS and related sources were analysed for SC 23. Figure 4.1 indicates the time variation of the sunspot number (SSN) and GMS activity (frequency occurrence of magnetic storms with $Dst \leq -50$ nT) for the last three SCs. Notice the triple peak of GMS activity as a function of SC, also reported in earlier studies (e.g. Yermolaev and Yermolaev, 2006). Looking at the SC 23 period for which CME data records are consistently available, it is interesting to observe in Figure 4.1 that the triple peak manifests also in the halo CME occurrence, a trend which was also previously reported in an analysis done by Gopalswamy et al. (2007).

The first 2 bars in Figure 4.9 indicate the number of GMS and associated halo CMEs events. The following two bars show the number of GMS events associated with ICMEs, and CIR-driven storms respectively. The distribution is shown for the rising phase (1996-1999), the maximum phase (2000-2002) and the declining phase (2003-2006) of SC 23. It can be seen in the figure that for the majority of storm events and associated halo CMEs, ICMEs were concentrated within three years of the solar maximum period, 2000, 2001 and 2002, followed by the declining phase. The figure also shows that half (50%) of all 80 non-halo CME-driven GMS (assumed to be CIR-driven) were concentrated within the declining phase. This is due to the fact that, in addition to the CME-driven storms, the declining phase also includes the recurrent GMS associated with the quasi-stationary corotating high-speed SW streams originating in the coronal holes (Snyder et al., 1963; Kriger et al., 1973). Figure 4.10 illustrates the annual frequency of 244 GMS events, 192 associated halo CMEs, 183 X-ray solar flares and 138 ICMEs events for SC 23. As shown in Figure 4.10, the peaks (in each of the three SC 23 phases) of geoeffective CMEs and ICMEs were observed in years 1998, 2000-2002 and 2005 respectively. The peak observed in 2003 represents an exception reported by

Gopalswamy et al. (2007). As indicated by Hady (2009), the exceptional secondary peak observed in the declining phase of SC 23 was associated to series of solar explosive activities (Halloween storms) which occurred during the period October 17 - November 10, 2003. Hady (2009) described the phenomena as being related to the large area of the active region AR 10486 and to a very strong magnetic field.

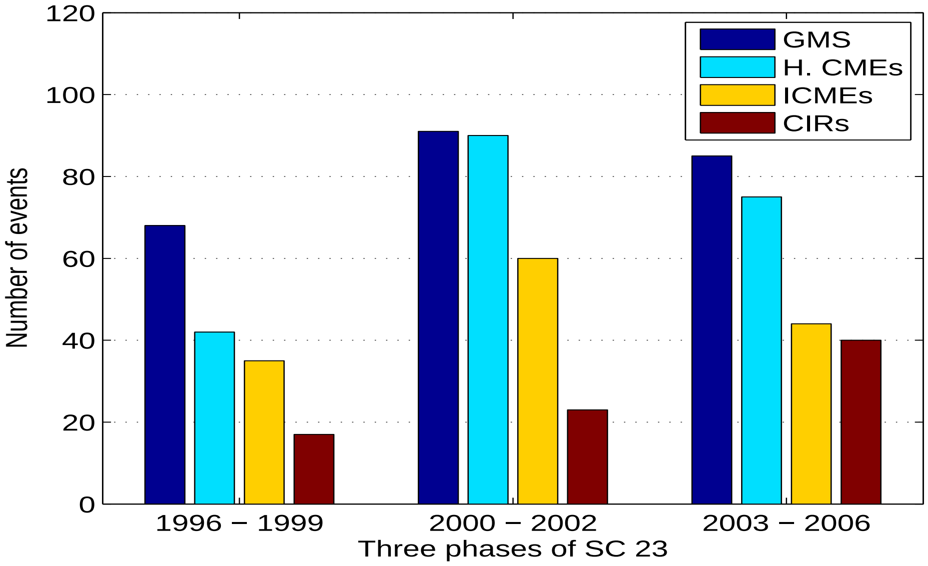


Figure 4.9: Frequency occurrence of GMS, associated halo CMEs and ICMEs in SC 23: 4 years rising phase, 3 years solar maximum and 4 years declining phase.

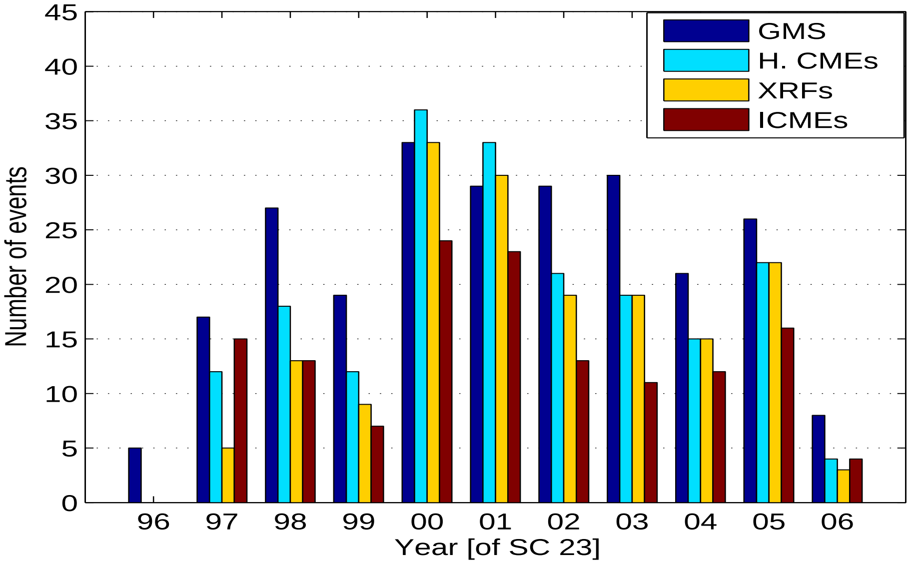


Figure 4.10: A bar graph showing the annual occurrence of GMS events, probable associated halo CMEs, X-ray flares and ICMEs in SC 23.

4.4 Summary

This chapter presented results of an investigation of the solar sources and IP characteristics of 229 GMS events ($Dst \leq -50$ nT) during SC 23. A particular aspect of these results is that they can be considered as an estimation of solar sources of GMS within on sunspot cycle (an average of 11 years).

While earlier similar studies focused on the sources of intense storms (Zhang et al., 2007), the present study has extended the analysis to moderate storms. The following are the most important results of the analysis:

- Most of the intense GMS (83%) during the period of study were caused by halo CMEs. For moderate storms, only 54% were associated with halo CMEs, while the remaining 46% seemed to originate from CIRs or geoeffective non-halo CMEs.
- The association of intense storms with ICMEs at 1 AU was very high (86%), compared to moderate storms of which only about 44% were associated with ICMEs.
- Using the information for X-ray flares associated to halo CME probable sources of storms, this investigation found that about 69% of intense storms were caused by CMEs originating close to the disk centre [within $\pm 45^0$ of the CMD], and only 51% of moderate storms had solar sources close to the disk centre.
- There was no significant difference between intense and moderate storms as far as the class of X-ray flare associated with the driving halo CME. Generally, both intense and moderate storms were mostly associated with C and M-class flares.
- A comparison of full halo CMEs and partial halo CMEs, showed that full halo CMEs were generally associated with intense storms with an average $Dst = -128$ nT, while partial halo CMEs were linked to moderate storms with an average $Dst = -92$ nT. On the other hand, geoeffective parameters of full halo CME-driven storms (B_z, SW speed, CME speed) had higher average values than those of partial halo CME-driven storms. In addition, full halo CME-driven storms were generally associated with class M flares, compared to partial halo CME-driven storms which were mainly associated with C-class flares.
- Full halo CME-driven storms were associated with CMEs mostly originating mostly close to the disk centre, while only 44% of partial halo CME-driven storms were associated with CMEs originating within the CMD.
- This investigation found that multiple halo CME-driven storms represent 26% of all CME-driven storms, the majority (69%) being intense storms. In addition, 92% of multiple CME-driven storms were associated with ICMEs at 1 AU.

- Finally, GMS and associated probable solar sources and IP properties demonstrated a triple peak in GMS, CME and ICME activities. The peaks are observed in each phase, namely in the rising phase (1998), maximum phase (2000) and declining phase (2005). 50% of non-halo CME (or CIR) driven storms were concentrated in the descending phase of SC 23.

The low geoeffectiveness of partial halo CMEs compared to full halo CMEs can be associated to the fact that the majority of partial halo CMEs originate far away from the solar disk center. For both CIRs and halo CMEs, if they originate far away from the the Sun center, the IP structures linked to them will have less geoeffective properties since only the outer flanks are expected to encounter Earth (Gopalswamy, 2008).

In this chapter, a quantitative analysis was conducted to explore probable solar sources and associated IP properties of GMS in SC 23. The aim of this kind of study should not only be limited to the identification and characterisation of the sources of GMS, but should also extend to the development of GMS prediction models for the purpose of minimising space weather effects. Some of the geoeffective parameters described in this chapter were used in NN-based models for predicting the occurrence of GMS. The next two chapters describe in detail the GMS prediction models that were developed.

Chapter 5

Predicting the geoeffectiveness of halo CMEs

Estimating the goeffectiveness of solar transient phenomena is of practical importance for the modelling and prediction of space weather. The ability to predict space weather needs an accurate prediction of GMS, phenomena which represent typical features of space weather. However, space weather prediction is still relatively inaccurate given the fact that the underlying physics of the main drivers, e.g. coronal mass ejections (CMEs) and associated X-ray flares is not yet sufficiently understood (Schwenn et al., 2005). This chapter describes the NN-based models developed for predicting the probability occurrence of GMS from halo CMEs and associated geoeffective parameters.

5.1 On the predictability of GMS with neural networks

Currently models for predicting GMS include statistical, empirical and physics based methods. However, despite previous attempted theoretical models to forecast the magnetic storm occurrence (Dryer, 1998; Dryer et al., 2004), physics based models are still difficult to achieve due to the complex and non-linear chaotic system of solar-terrestrial interaction (Fox and Murdin, 2001; Schwenn et al., 2005). Space weather forecasters often prefer empirical approaches based on observable data (Kim et al., 2010). Various functional relationships have been proposed for magnetic storm predictions. Empirical models for predicting GMS using CME-associated parameters observed at the Sun have been developed, including a recent model by Kim et al. (2010). Other authors prefer statistical methods, e.g. Srivastava (2005) used a combination of solar and IP properties of geoeffective CMEs in a logistic regression model to predict the occurrence of intense GMS.

Empirical methods also include NN methods which are input-output models and have been proven to be efficient in capturing the linear as well as the non-linear processes (Kamide et al., 1998). As indicated in Section 3.1 of Chapter 3, NN techniques have been described by various authors to be suitable for predicting non linear systems. If a network is well-designed and trained, it can improve a theoretical model by performing generalisation rather than only curve fitting. By changing the NN input values, it is possible to investigate the functional relationship between the input and the output and thereby be able to derive what the network has learned (Lundstedt, 1997). In a NN-based model developed by Valach et al. (2009), geoeffective solar events such as solar X-ray flares (XRAs) and solar radio bursts (RSPs) were used to predict the subsequent GMS. In order to improve the forecast of GMS, Dryer et al. (2004) suggested that models should include both solar and near-Earth geoeffective conditions.

For this study, a combination of solar and IP properties which are characteristic of geoeffective halo CMEs were used in a NN model to predict the probability occurrence of GMS following halo CME events.

5.2 Determination of NN input/output parameters

5.2.1 Geoeffective properties of halo CMEs: solar input parameters

For the development of the model to predict the probability occurrence of GMS, halo CMEs were considered as the primary inputs. Halo CMEs originating from the visible solar disc and that are Earth-directed have the highest probability of impacting the Earth's magnetosphere (Webb et al., 2000), and hence are useful for predicting the occurrence of GMS. In this study, both full and partial halo CMEs (with apparent AW ranging between 120^0 and 360^0 were used. A correlation coefficient of 0.75 was calculated between full halo CME occurrence rate per year and the occurrence rate of geomagnetic disturbances (disturbed days per year with $Dst \leq -50$ nT) during the period 1996 to 2006. For the model developed in this study, the halo CME data from the LASCO/SOHO catalogue list was used. In spite of a close connection between halo CMEs and GMS, it is still difficult to predict the occurrence of GMS based solely on halo CME observations (Gopalswamy, 2009a). Hence, other solar and IP geoeffective properties should also be considered in order to improve the accuracy of GMS prediction.

In addition to the AW of CMEs, the CME speed represents another important property of geoeffective CMEs. Halo CMEs generally have higher speeds than the mean SW (470 km/s) and thus, the CME speed is a useful parameter to predict the intensity of GMS (Srivastava, 2005). In this study, the CME linear speed measured in the LASCO-C2 field of view was used. Another solar input used is the comprehensive flare index (cfi) which characterises

the flare activity associated with CMEs. In their analysis, Wang et al. (2002) found a close correlation between flare activity and the geoeffectiveness of halo CMEs. For the NN model described in this chapter, the cfi index was used as an input to quantify the halo CME association with solar flares. The minimum flare activity corresponds to 0 as a value of cfi and the highest value of cfi observed in SC 23 was 144 and occurred during the *Halloween event* on 28 October 2003. The cfi data archive is available on the website:
fttp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_FLARES/
FLARES_INDEX/Solar_Cycle/23/daily.plt.

5.2.2 IP input parameters

As they approach the Earth, CMEs manifest as shocks and ICME structures which couple to the magnetosphere to drive moderate to major storms (Webb, 2000). Gopalswamy (2009a) noted that the intensity and duration of the southward component of the IP magnetic field in the shock sheath and ICME determine the importance of the subsequent storm. The relationship between halo CMEs, ICMEs, IP shocks and the subsequent magnetic storms has been described in Chapter 2, Section 2.2.4 as well as in Chapter 4, section 4.2.2. The reader is refered to those sections for more details.

In situ observations of plasma and magnetic field properties help to identify the arrival of ICMEs near Earth. The occurrence of shock waves and possible associated ICMEs can be characterised by a simultaneous increase in speed, density, abnormal proton temperature in the SW, and in the magnitude of the IMF. Plasma and magnetic field signatures indicating the presence of ICMEs are fully described by Cane and Richardson (2003) and Schwenn et al. (2005). Generally, increased total IMF and SW speed are features typically associated with shocks/ICME structures that pass near Earth at 1AU. As indicated by Gonzalez and Tsurutani (1987), the intensity of the storm following the passage of a shock-ICME structure is well correlated with two parameters namely: (1) the southward component of the IMF Bz-component or Bs and (2) the electric field convected by the SW, $Ey = VBs$, where V is the SW velocity.

The data for Shock and ICME events which trigger SW geoeffective conditions consist of in situ by the Solar Wind Electron Proton Monitor (SWEPAM) and by the Magnetic Field Experiment (MAG) instruments on board ACE spacecraft (Stone et al., 1998). The listing of ICMEs by Richardson and Cane (2008) and associated properties is available on the website:
http://www.ssg.sr.unh.edu/mag/ace/ACElists/ICMEtable.html
The shocks list is also found online on the website:
http://www.bartol.udel.edu/∼chock/ace/ACElists/obs_list.html

Figure 5.1 shows the distribution of GMS, shocks and the ICME occurrence per year during the period 1997 -2006. During this period, 225 GMS, 297 ICMEs and 399 IP shocks were investigated. The correlation coefficient between GMS and ICME occurrence rate per year was found to be 0.58. A higher correlation coefficient of 0.78 was found for GMS events and the shock occurrence rate, as shown by Figure 5.2.

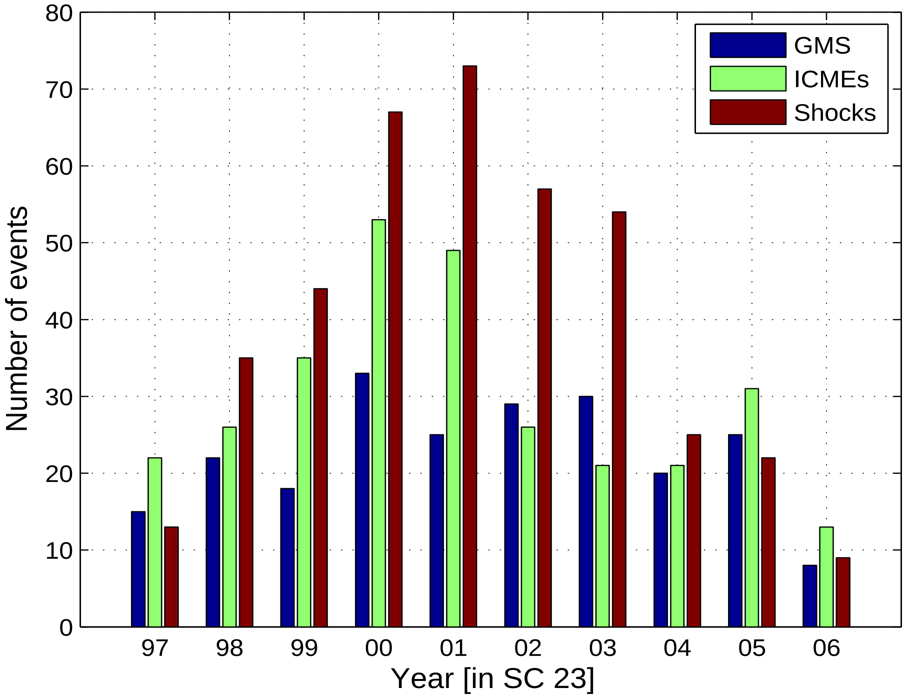


Figure 5.1: The bar graph shows the frequency distribution per year of 297 ICMEs, 225 GMS and 399 IP shocks for the period 1997 - 2006.

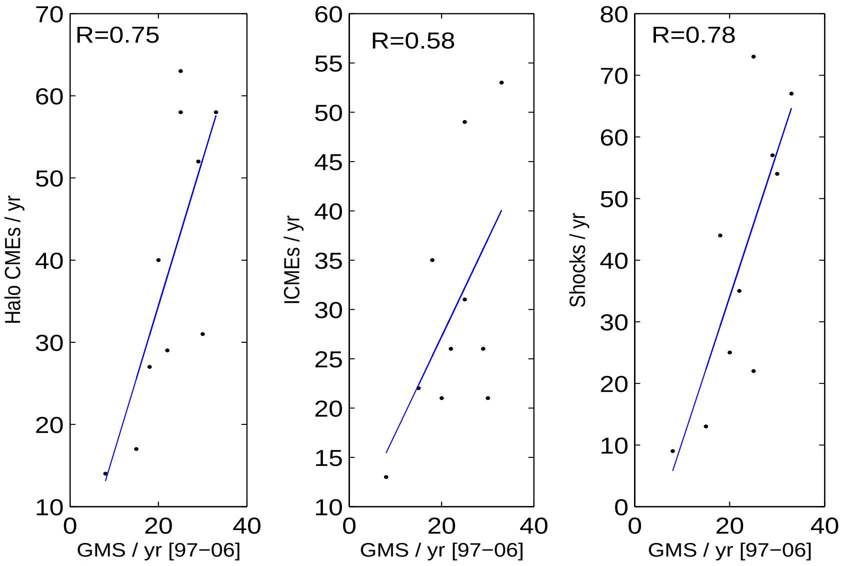


Figure 5.2: Plots illustrating the correlation between the GMS occurrence and the frequency occurrence of halo CMEs, ICMEs and shock events.

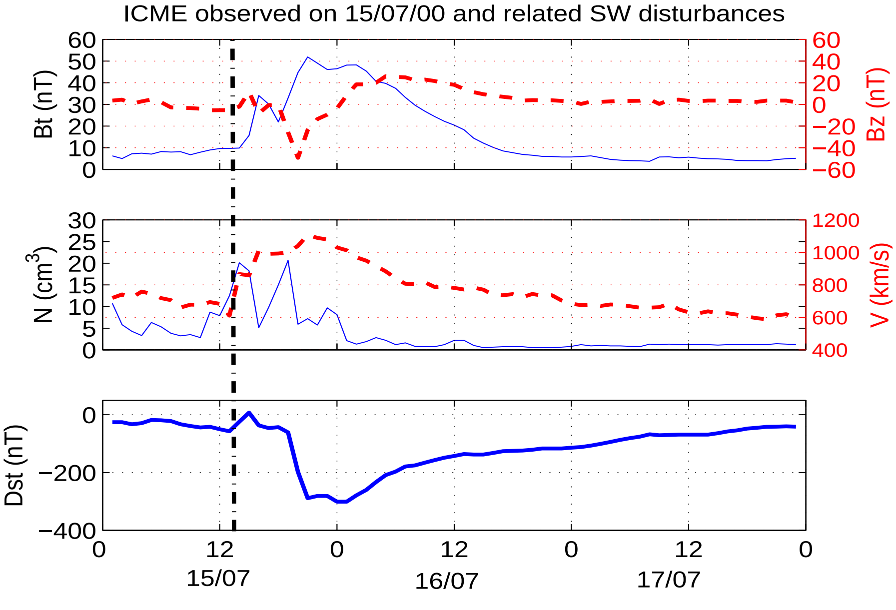


Figure 5.3: Plot showing the variation of the IMF total field (B_t) and and its Z-component (B_z), the SW velocity V and density N, following the passage of an ICME, observed by the WIND spacecraft on 15/16 July 2000. The vertical solid dashed line labels the shock ahead of the ICME. Plot adapted from Messerotti et al. (2009)

Based on solar and IP geoeffective parameters defined above, various networks were constructed (depending on the number and type of inputs) and trained in order to determine the NN model which would optimally predict the occurrence of GMS. Firstly, a network was constructed based only on the solar inputs namely: the AW of CMEs, the CME speed and the cfi used as numeric values. The second network considered, in addition to the previous three solar inputs, the shock and ICME events used as binary values. In the third network, IP input binary values of shock and ICME events were replaced by IP peak values of Bs and Vsw used as numeric values. Table 5.1 indicates the characteristics of NN input parameters. The source of data as well as characteristics of the Bs and Vsw peak values used to develop the NN model have been described in chapter 4, section 4.2.2.

Figure 5.3 is an illustration of IP disturbances associated with the arrival of the shock (and driver ICME) at 1AU on 15 July 2000, driving a storm of 16 July 2000 with peak minimum Dst reaching -301 nT. This storm was caused by a very fast (1674 km/s) full halo CME launched on 14 July at 10:54 UT and was associated with an $X5.7$ flare (cfi=59.13) originating at N22W07. In the IP medium, the Bs reached a peak value of -49.4 nT and 1040 km/s was the maximum SW during the passage of the ICME. Note that this event corresponds to the solar explosive event that triggered a radiation storm around Earth nicknamed the *Bastille event.* It is important to mention that all the input parameters used are directly associated with halo CMEs, and therefore, the NN model was not trained to predict storms that are non CME-driven, such as those caused by CIRs.

Table 5.1: Solar and IP input parameters used in the development of the NN model. [A] represents the first network consisting of three solar inputs (CME-AW, Vcme and cfi). The second network trained used [A] +[B] inputs, with [B] representing shock and ICME events used as binary values. The third network consited of [A] + [C] inputs, with [C] representing the IP peak values of Vsw and Bs used as numeric values.

Model parameter type	parameter name	variable type	measure	value
Inputs	CME A.W	Numeric	$\geq 120^0$	-
[A]	CME speed	Numeric	value in km/s	-
	cfi	Numeric	-	-
[B]	Shocks	Binary	presence absence	1 0
	ICME	Binary	presence absence	1 0
[C]	Vsw	Numeric	value in km/s	-
	Bs	Numeric	value in nT	-
Outputs	No storm	Binary	$Dst > -50$ nT	0
	storm	Binary	$Dst \leq -50$ nT	1

5.2.3 Geomagnetic response

The Dst index is a commonly used index to quantify the intensity of magnetic activity. When the ICME structure in the IP medium presents an intensified southward component of the IMF (B_z), it reconnects with the Earth's magnetic field. This magnetosphere-solar wind coupling induces the buildup of the ring current (Gonzalez et al., 2004; Gopalswamy, 2009a), and therefore, the Dst index variation is a response to the build-up and decay of the ring current. The classification of storm intensity in terms of the Dst index by Loewe and Prölss (1997) was described in section 4.2.1. For simplicity, only moderate storms (-100 nT $\leq Dst \leq$ -50 nT) and intense storms ($Dst \leq -100$ nT) were considered for this analysis. In Table 5.1 the storm occurrence is represented by binary values: 1 for a storm occurred ($Dst \leq -50$ nT) and 0 in the absence of a storm ($Dst > -50$ nT).

5.3 Neural networks

NNs have been used as a tool to explore the predictability of a storm occurrence based on the observed solar and IP properties of halo CMEs. NN prediction techniques and applications were outlined in chapter 3. For the cuurent model, the three-layered FFNN was used.

In this study, a logistic sigmoid activation function (see equation 3.5) was used for both hidden neurons as well as output neurons. This function is a monotonical function which is defined for all real numbers. The inputs of the network are therefore not limited to only the binary values 0 or 1. Continuous quantities can also be used as input, while the outputs of the network are real numbers in the range 0 to 1 (Valach et al., 2009). In addition, such activation functions applied on the network outputs play an important role in allowing the outputs to be given a probabilistic interpretation (Bishop, 1995).

Before training, the data set was split randomly into training and testing data sets in order to avoid the training results becoming biased towards a particular section of the database. For the NN trained while developing this model, data were split into 70% for the training set and 30% for the testing set. In order to determine how the NN has learned the behaviour in the input-output patterns, a validation data set consisting of the data not involved in the network training process was selected. Before the training, the input vectors were ordered randomly before each was presented to the network. The SNNS software (see Chapter 3, Section 3.3.4) was used for the training process.

5.3.1 NN model development: input/output data preparation

The first step in developing this model was the preparation of the database for the NN training using the criteria that halo CMEs ($AW \geq 120^0$) were or were not followed by a storm event ($Dst \leq -50$ nT) within a 5-day window from the launch of a halo CME. Halo CME (and associated solar and IP parameters) data covering the period from September 1997 to December 2006 were used in the model. This is the period corresponding to the availability of the ICMEs and related shock structures data (see the listing by Richardson and Cane (2008)). Note that there were missing CME data records for July, August and September 1998 as well as January 1999. In total, 1202 halo CMEs and associated geoeffective properties were involved in the training, testing and validation of the NN model. The data covering 6 months in 2000 and 12 months in 2005 was set aside for the validation data set and was not used in the training. This unseen data indicates the model's ability to predict the output in a general way.

Note that a positive response (code 1) was assigned as output to all halo CME inputs that were followed by GMS within a 5-day window. Therefore, the number of input events that were allocated a positive response in the output dataset is actually larger than the total number of isolated GMS events (around 225 investigated) that actually occurred during the period of study, including about 90 intense storms (Echer et al., 2008). The reason is that there were many cases where one isolated storm was common to more than one halo CME.

Table 5.2 shows 43 (5 of the 48 listed had no halo CME background) halo CME-driven storm events as well as their solar and IP characteristics considered for the validation data set. Note that Table 5.2 is simplified and doesn't indicate cases where more than one halo CME shared one storm event. A good example is the storm of 24 August 2005 with peak minimum Dst of -216 nT. Although one full halo CME is indicated in Table 5.2 (event number 44) as the storm driver, there were actually two high speed ($V > 1000$ km/s) full halo CMEs which were probable sources of the storm and occurred on the same day on 22 August 2005. In fact, the two halo CMEs involved were all frontsided, associated with M-class solar flares and were followed by an ICME observed on 24 August 2005. Therefore, for this particular example, there were two rows of input vectors made of AW=360 with CME speeds of 1194 km/s and 2379 km/s respectively. The two halo CMEs occurred on the same day, therefore had the same cfi value of 10.31. Both events were associated with shocks and ICMEs at 1AU. Hence the last two NN inputs (IP inputs) were assigned a binary input of 1 (for the [A] + [B] input network). For the [A] + [C] network, the two IP inputs correspond to the peak values of Bs and Vsw (38.3 nT and 710 km/s respectively) identified within a 5-day window after the halo CME eruption.

The outputs corresponding to the two input events described above were represented by a binary value of 1 because there was storm occurrence ($Dst \leq -50$ nT). In case halo CMEs were not followed by a storm ($Dst > -50$ nT), the corresponding output events were represented by a binary value of 0. Notice that the output of the NN model after training is a numerical value ranging between 0 and 1. The input/output pattern of the the NN system is shown in Table 5.1 and Figure 5.4. The model developed can, therefore, be interpreted as a function that determines the probability of storm occurrence (as a consequence of halo CME eruption) and can be written as

$$P = f(AW_{cme}, V_{cme}, cfi, B_s, V_{sw}) \tag{5.1}$$

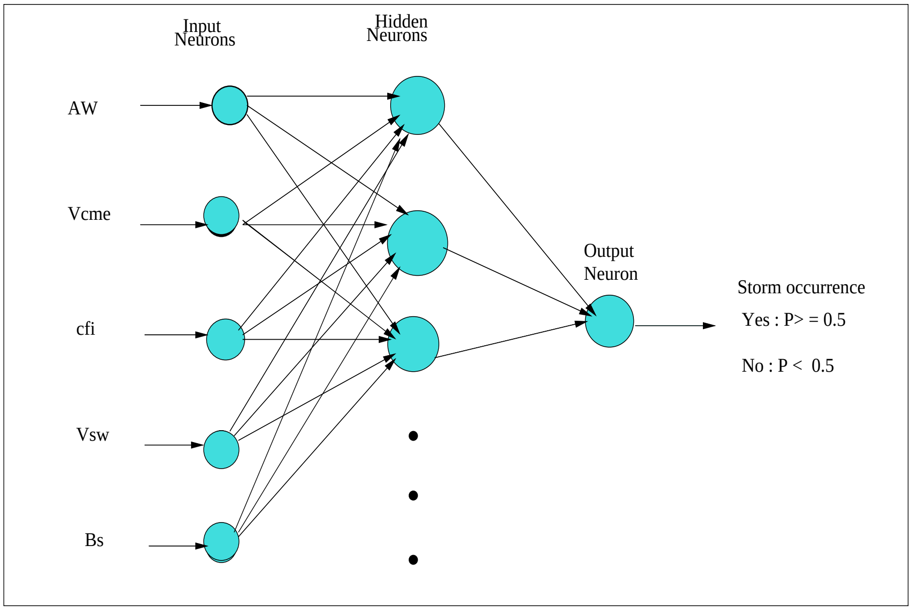


Figure 5.4: A simplified illustration of the three layered FFNN architecture as developed and used in this study.

5.3.2 NN optimization

The optimum network was determined by training different NNs with selected input parameters and by systematically changing the number of nodes in the hidden layer. The optimum NN architecture was obtained by computing the root mean square error ($RMSE$) values over the validation data set.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_{obs} - P_{pred})^2}, \tag{5.2}$$

where P_{obs} and P_{pred} represent the observed and predicted probability values respectively and N represents the number of data points in the validation data set. Note that the same criterion of a minimum RMSE was used to determine a suitable input space for the model as shown in Table 5.3. In order to evaluate the prediction performance of the model on the training data set, the least RMSE was computed over the training data set and was 0.3422 (not shown in Table 5.3). By following the example as in Srivastava (2005), a value of 0.5 was considered as a threshold value (probability) to define the classification of prediction output. In fact, the average predicted probability calculated over the whole validation data set (including zeros and ones cases) was 0.48, slightly below the threshold value. Therefore, any prediction output with value ≥ 0.5 was considered as a positive likely occurrence of a storm event following a halo CME.

Table 5.2: Geomagnetic storms and associated halo CME properties used for the validation data set. Only 43 of all 48 events were CME-driven storms.

Event No.	**Date / Time**	**Dst(min.) [nT]**	**halo CMEs [F or P]**	**Vcme [km/s]**	**X-Ray flare**
1	08/06/00 19:00	-90	FH: 06/06 [15:54]	1119	X2.3
2	26/06/00 - 17:00	-76	PH: 25/06 [07:54]	1617	M1.9
3	16/07/00 - 00:00	-301	FH: 14/07 [10:54]	1674	X5.7
4	20/07/00 - 09:00	-93	-	-	-
5	23/07/00 - 22:00	-68	PH: 22/07 [11:54]	1230	M3.7
6	29/07/00 - 11:00	-71	FH: 25/07 [03:30]	528	M8.0
7	06/08/00 - 05:00	-56	PH: 03/08 [08:30]	896	C1.4
8	11/08/00 - 06:00	-106	PH: 08/08 [15:54]	867	C1.4
9	12/08/00 - 09:00	-235	FH: 09/08 [16:30]	702	C2.3
10	29/08/00 - 06:00	-60	PH: 25/08 [14:54]	518	M1.4
11	02/09/00 - 14:00	-57	PH: 01/09 [04:06]	603	C1.1
12	12/09/00 - 19:00	-73	PH:09/09 [08:56]	554	M1.6
13	16/09/00 - 23:00	-68	FH:12/09 [11:54]	1550	M1.0
14	18/09/00 - 23:00	-201	FH: 16/09 [05:18]	1215	-
15	26/09/00 - 02:00	-55	FH: 25/09 [02:50]	587	M1.8
16	30/09/00 - 14:00	-76	PH: 27/09 [01:50]	820	C5.2
17	05/10/00 - 13:00	-182	FH: 02/10 [20:26]	569	C8.4
18	14/10/00 - 14:00	-107	PH: 11/10 [06:50]	799	C2.3
19	29/10/00 - 03:00	-127	FH: 25/10 [08:26]	770	C4.0
20	07/11/00 - 21:00	-159	FH: 03/11 [18:26]	291	C3.2
21	10/11/00 - 12:00	-96	FH: 08/11 [04:50]	474	-
22	29/11/00 - 13:00	-119	FH: 25/11 [01:31]	2519	M8.2
23	23/12/00 - 04:00	-62	PH: 20/12 [10:30]	1076	-
24	01/01/05 - 19:00	-57	FH: 30/12 [20:30]	832	B2.8
25	08/01/05 - 02:00	-96	FH: 05/01 [15:30]	735	-
26	12/01/05 - 10:00	-57	PH: 09/01 [09:06]	870	M2.4
27	18/01/05 - 08:00	-121	FH: 15/01 [06:30]	2049	M8.6
28	22/01/05 - 06:00	-105	FH: 19/01 [08:29]	2020	X1.3
29	07/02/05 - 21:00	-62	PH: 05/02 [13:31]	711	-
30	18/02/05 - 02:00	-86	FH: 17/02 [00:06]	1135	-
31	06/03/05 - 16:00	-65	-	-	-
32	05/04/05 - 05:00	-85	PH: 04/04 [11:06]	421	-
33	12/04/05 - 05:00	-70	PH: 09/04 [08:26]	329	B2.6
34	08/05/05 - 18:00	-127	FH: 05/05 [20:30]	1180	C7.8
35	15/05/05 - 08:00	-263	FH: 13/05 [17:12]	1689	M8.0
36	20/05/05 - 08:00	-103	PH: 17/05 [03:06]	449	M1.8
37	30/05/05 - 13:00	-138	FH: 26/05 [15:06]	586	B7.5
38	13/06/05 - 00:00	-106	PH: 08/06 [07:48]	179	-
39	15/06/05 - 12:00	-54	-	-	-
40	23/06/05 - 10:00	-97	-	-	-
41	09/07/05 - 18:00	-60	FH: 05/07 [15:30]	772	C1.3
42	10/07/05 - 20:00	-94	FH: 09/07 [22:30]	1540	M2.8
43	18/07/05 - 06:00	-76	FH: 14/07 [10:54]	2115	X1.2
44	24/08/05 - 11:00	-216	FH: 22/08 [17:30]	2378	M5.6
45	31/08/05 - 19:00	-131	FH: 29/08 [10:54]	1600	-
46	11/09/05 - 09:00	-147	FH: 09/09 [19:48]	2693	X6.2
47	15/09/05 - 16:00	-86	FH: 13/09 [20:00]	1866	X1.5
48	31/10/05 - 19:00	-75	-	-	-

Table 5.3: Computation of an optimised NN architecture over the validation data set for the three input networks. The optimum NN was obtained when Vsw and Bs are used as numeric values (see the corresponding input network [A] +[C]) in Table 5.1

Inputs	NN architecture	RMSE
[A]	**3:3:1**	**0.5126**
	3:4:1	0.5137
	3:5:1	0.5147
	3:6:1	0.5155
[A]+[B]	**5:5:1**	**0.4621**
	5:6:1	0.4625
	5:7:1	0.4633
	5.8:1	0.4624
[A]+[C]	5:5:1	0.3225
	5:6:1	0.3396
	5:7:1	0.3366
	5.8:1	0.3376

5.4 Prediction results and discussion

The optimum network architecture having the least RMSE was found to be that with 5 inputs (equation 5.1) using 5 hidden nodes (configuration: 5:5:1, corresponding to [A] + [C] input network). This optimised NN architecture was reached after 1400 iterations using 0.005 as the learning rate. The network with only three solar input parameters performed poorly when tested on the validation data set as shown in Table 5.3. This indicates the importance of considering IP parameters (shocks, ICMEs events or their corresponding in situ measured values of B_s and V_{sw}) for improving the prediction performance of the model.

The NN optimisation in Table 5.3 shows that the IP input used as numeric values of Vsw and Bs ([A]+[C]) improves the prediction of GMS, as compared to the use of the shock and ICME parameters used as binary input values ([A]+[B]). Figure 5.5 (a-d) is an example to illustrate the predictions made by the model. The x-axes indicate days in a month on which there were one or multiple halo CMEs and the y-axes indicate the predicted value which can be translated into the probability of the halo CME to be geoeffective. All the predicted values above 0.5 indicate a correct prediction of GMS occurrence (following a halo CME) and the maximum in each of the four cases presented in Figure 5.5 indicates the probability by which a particular storm is predicted to occur.

Note that not all halo CMEs are indicated on the plots due to the fact that some dates had multiple halo CMEs, while no halo CME were observed for other days. Four typical examples are described for all of which the model indicates a very high probability ($P \simeq 1$) of storm occurrence. Figure 5.5 was generated based on the NN predictions which considered peak values of Vsw and Bs as numeric values ([A]+[C] input network as shown in Table 5.1).

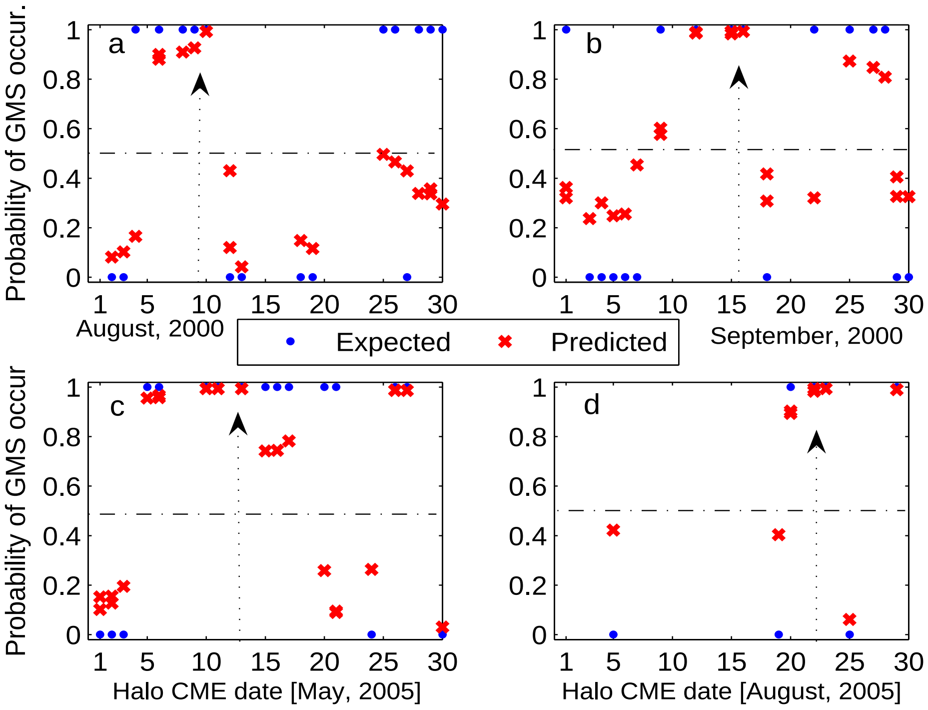


Figure 5.5: Illustration of the prediction performance of the model on some storms in the validation dataset. All predictions above a 0.5 probability value indicate a successful prediction (of storm occurrence) where the predicted value can be interpreted as the probability by which a particular halo CME may be followed by a GMS. Example events described in the text are indicated by the arrows.

There were two intense storms which occurred on 11 and 12 August 2000 reaching the Dst peak minima of -106 nT and -235 nT respectively. As shown in Figure 5.5 (a), the two storms were correctly predicted with more than 0.95 probability and were expected from the three halo CMEs which occurred on 8, 9 and 10 August 2000 respectively. The most probable cause of the 12 August 2000 storm was a full halo CME which erupted on 09 August 2000 (see the arrow in the plot), indicated as number 9 in Table 5.2. Like in many observed moderate storm cases, this model however fails to correctly predict the storm of 29 August 2000 ($Dst = -60$ nT) expected from a series of partial halo CMEs that occurred on 25 - 28 August 2000.

Figure 5.5 (b) shows how the model correctly predicts the two storms on 16 and 18 September 2000 respectively. The 16 September 2000 moderate storm ($Dst = -68$ nT) was expected to follow the fast and full halo CME of 12 September 2000. A strong storm that occurred on 18 September 2000 (with Dst peak minimum of -201 nT) is correctly predicted by the model, and was expected from the two halo CMEs (shown by an arrow in Figure 5.5 (b)) which occurred on 15 and 16 September 2000. However, it is most likely that this storm was caused by the very fast full halo CME of 16 September (see event number 14, Table 5.2) or its interaction with the partial halo CME of 15 September 2000.

Figure 5.5 (c) shows that the NN model correctly predict the 15 May 2005 severe magnetic storm (with Dst peak minimum of -263 nT). This storm was expected due to a very fast full halo CME and associated powerful flare of 13 May 2005 as shown by the arrow in Figure 5.5 (c) and row number 35 in Table 5.2. Similarly, Figure 5.5 (d) illustrates clearly the correct prediction of the 24 August 2005 strong magnetic storm (with a Dst peak minimum of -216 nT). This storm followed two fast full halo CMEs which occurred on 22 August 2005. The very fast halo CME (V=2378 km/s), which was the most probable cause of the storm, is number 44 in Table 5.2 and its predicted geoeffectiveness is represented by the arrow in Figure 5.5(d).

The average value of correct predictions (observed storm responses and predicted with $P \geq 0.5$) calculated over the whole validation dataset was found to be 0.87. This can be considered as the NN model approximated probability by which a storm occurrence can be predicted as a result of the halo CME event. The performance of the developed NN model was tested on 43 CME driven GMS (listed in Table 5.2) by calculating the percentage of correctly predicted storms for both the validation and training data sets using:

$$\frac{PE}{OE} \times 100 \tag{5.3}$$

where PE is the number of correctly predicted GMS events and OE is the total number of observed GMS events in the validation dataset (training set). Table 5.3 shows the minima of RMSE computed over the validation set, indicating that the model performed poorly when only three solar inputs are used. With three solar inputs, the NN model could predict only 52% and 41.6% of the observed intense and moderate storms respectively. The results presented in Table 5.4 summarise the prediction performance of the NN model based on the optimum network obtained by considering [A] + [B] inputs as shown in Table 5.1. The NN model performance is evaluated on both the training and the validation data sets, indicating the fraction of correct predictions for both intense and moderate storms as well as the number of false alarms (non observed GMS events, but predicted by the NN model). When IP parameters are used as numeric values of Vsw and Bs (input network [A]+[C]), the prediction performance is considerably improved. This is shown in Table 5.5, indicating that the developed NN model predicts accurately all intense storms, and 75% of moderate storms. The overall NN model forecasts of GMS ($Dst < -50$ nT) based on the observed halo CME was estimated at 86%.

The results obtained demonstrate the improved ability of the NN model to predict the occurrence of intense storms than moderate storms. This difference in performance is related to the characteristics of the inputs. Observations of the data indicate that intense storms are generally preceded by full halo CMEs ($AW = 360^0$), high values of CME speed and cfi, as well as high peak values of Bs and Vsw compared to the values associated with moderate geomagnetic storms. On the other hand, previous studies have indicated that partial halo CMEs produce mostly moderate storms and the majority of them are less energetic (have lower speed). Note also that moderate storms are often driven by the non-halo CMEs or CIRs for which this NN has not been trained to predict.

It is important to mention that the results presented in this study only serve as an indication that the solar and IP parameters characteristics of geoeffective halo CMEs can be used in a NN to estimate the probability occurrence of the subsequent GMS. The predicted geoeffectiveness of solar events (halo CMEs in this case) can be compared to other predictions from various analysis. Valach et al. (2009) used a combination of X-ray flares (XRAs) and solar radio bursts (RSPs) as input to the NN model and obtained a 48% successful forecast for severe geomagnetic response. The NN model described in this paper shows an improved performance with an accuracy of 86% in the prediction of GMS. On the other hand, this model performance compares favourably to the 77.7% obtained by Srivastava (2005) using the logistic regression model. The developed NN model described in this chapter has successfuly attempted the predictability occurrence of moderate storms which have not been considered in previous studies.

Table 5.4: Prediction performance of the NN model with IP inputs used as binary values of ICMEs and shock events (A+B) input network).

Data set	Storm category	Observed	Correct predictions	False alarms
Training	Intense storms	53	45 [84.9%]	
	Moderate storms	59	31 [52.5%]	
	Total	112	76 [67.8%]	33
Validation	Intense storms	19	16 [84.2%]	
	Moderate storms	24	15 [62.5%]	
	Total	43	31 [72%]	8

Table 5.5: Prediction performance of the NN model with IP inputs used as numeric values of Vsw and Bs (A+C) input network.

Data set	Storm category	Observed	Correct predictions	False alarms
Training	Intense storms	53	51 [96%]	
	Moderate storms	59	42 [71%]	
	Total	112	93 [83%]	32
Validation	Intense storms	19	19 [100%]	
	Moderate storms	24	18 [75%]	
	Total	43	37 [86%]	8

5.5 Summary

Predicting the occurrence of GMS on the basis of CME observations only is challenging and can sometimes lead to false alarms. In this study, the NN-based model was explored with success as a means to predict the probability occurrence of GMS resulting from halo CMEs. The results obtained indicate that the model performs well compared to similar previous predictions. In addition, this model shows better prediction performance of intense GMS as compared to the prediction of moderate storms. In agreement with observations by Srivastava (2005), this study shows that IP input parameters characterising geoeffective halo CMEs (ICMEs and shock events used as binary values, or the corresponding in situ measured peak values of Bs and Vsw) contribute significantly to the improvement of the GMS prediction. However, although the use of only solar inputs leads to a poor prediction ability, it is more important for space weather since it provides a longer warning time (2 to 4 days) compared to the use IP parameters at 1AU, which provides a warning time of only about one hour. The NN model developed and described in this chapter might contribute towards improving space weather prediction. Particularly, the model developed will be applied by the SANSA Space Weather Regional Warning Center (RWC) to improve its various space weather models that involve consideration of storm conditions, including GICs and TEC modelling for Southern Africa. .

Chapter 6

Modelling geomagnetic indices during storm time

In Chapter 4, ICMEs and CIRs were described as IP structures which produce geoeffective conditions in the SW and lead to the occurrence of GMS. In Chapter 5, it was demonstrated that the use of SW and IMF input parameters improves the NN predictability of GMS occurrence. For space weather application, there is a need of not only to be able to predict the occurrence of GMS, but to also estimate the intensity of the storm. The analysis in Chapter 4 showed that the magnitude of the SW speed and the negative Bz-component of the IMF (in columns 4 and 5 of Table 4.1) were correlated with the strength of the associated GMS. Previous studies have also demonstrated the importance of the historical SW data for identifying and predicting the occurrence of large storms with a relatively high accuracy (Fox and Murdin, 2001). The NN model described in Chapter 5 demonstrated its ability to predict the occurrence, but not the magnitude of the storms. This Chapter describes the development of a NN-based model to predict the magnitude of a storm as measured by the Dst, ap, and K geomagnetic indices. These indices were briefly described in Section 2.5.3 of Chapter 2.

6.1 Previous related studies

As explained in Chapter 5, current models for predicting GMS are dominated by empirical techniques which generally, rely on the observable storm precursors in the SW (Fox and Murdin, 2001). An algorithm for predicting the Dst index from SW parameters and the IMF B_z component was first proposed by Burton et al. (1975). Other Dst prediction models using SW parameters as input were subsequentely developed (Temerin and Li, 2002).

NN models for predicting GMS using SW and IMF data as inputs have been developed and used (Lundstedt and Wintoft, 1994). These models demonstrated success in predicting the storm activity level as measured by the *Dst* index. In particular, Elman NN-based algorithms by Wu and Lundstedt (1996), Wu and Lundstedt (1997) and Lundstedt et al. (2002) demonstrated their ability to improve the *Dst* prediction. Pallocchia et al. (2006) used the same technique to develop a *Dst* forecast algorithm from only IMF data. The authors' choice of inputs was due to the fact that the the ACE/SWEPAM instrument does not provide reliable plasma data during proton events. In this study, a FFNN-based algorithm was developed with the ability to reproduce intense GMS ($Dst < -100$ nT) from only SW and IMF parameters. In particular, a comparison between the *Dst* and *ap* predictions by the developed NN model indicate that the *ap* index is better predicted as compared to the prediction of the *Dst* index. The ability to predict the *ap* index is very useful as it is a commonly used magnetic index for ionospheric storm-time modelling (Perrone and de Franceschi, 1999; Araujo-Pradere et al., 2003). The prediction of the *ap* index is, therefore, of practical importance for space weather.

6.2 Data sets

6.2.1 The role of SW parameters in generating GMS

There is a close link between geomagnetic disturbances and the IMF and its fluctuations, both in magnitude and direction (Schwenn et al., 2005). In particular, the long duration of the southward pointing IMF B_z by its interconnection with the Earth's magnetic field allows SW energy transport into the earth's magnetosphere (Gonzalez et al., 2004). As indicated by Gosling et al. (1991), a sustained and enhanced SW speed as well as the southward and northward IMF-Bz components are commonly associated with interplanetary shocks and ejecta known to be important causes of storms . On the other hand, an enhanced SW density (N) is an important parameter which affects the storm strength. The magnetospheric response to the changes in N corresponds to a change in plasma sheet density which increases the ring current when the IMF B_z is negative (Crooker, 2000).

The model described in this chapter used as inputs the SW speed V, the IMF total field B_t and B_z components as well as the SW density N. The IMF B_z-component used here is expressed in GSM coordinates because it maximises the correlation with geomagnetic activity (Kivelson and Russell, 1995). GMS coordinates have an $x-$ axis pointing directly at the Sun and an x-z plane that contains the Earth's dipole axis. Hence, the southward IMF B_z in GSM will allow the maximum magnetic flux merging with the Earth's field. Figure 6.1 illustrates the correlation between the disturbances in the SW parameters and the corresponding storm which occurred on the 12 August 2000.

6.2.2 Geomagnetic *Dst* and *ap* indices

The *Dst* index is the most commonly used geomagnetic index which indicates the level of geomagnetic disturbance. This index is measured by magnetometer stations near the equator where the magnetic perturbation is characterised by the magnetospheric ring current intensity. Burton et al. (1975) established the relationship between the ring current development and the geomagnetic *Dst* index, an equation which expresses the balance between the energy injection and decay out of the ring current.

$$\frac{d\,Dst}{dt} = Q(t) - \lambda\frac{Dst}{\tau} \tag{6.1}$$

where Q is a function of the SW energy input, and τ represents the decay time related to the loss processes in the ring current. In this study, intense GMS ($Dst < -100$ nT) are distinguished from moderate storms (-100 nT< Dst < -50 nT) according to the storm classification by Loewe and Prölss (1997).

Geomagnetic activity is also measured by the *ap* index, where *p* stands for *planetary*. The *ap* index is derived every three hours and indicates the level of geomagnetic disturbance compared to geomagnetically quiet days. The *Ap* index represents the daily mean of the *ap* values. The *ap* index values range from 0 (very quiet) to 400 nT (extremely disturbed) (Ondoh, 2001). This index is a linear equivalent of the quasi-logarithmic scale geomagnetic *Kp* index, also mostly used to express the level of geomagnetic disturbances. The relationship between the *ap* and the *Kp* geomagnetic indices is shown in Table 2.3 of Chapter 2. Like the *Dst* index, the geomagnetic activity *ap* index is responsive to the energy input into the magnetosphere. McPherron et al. (2004) analysed the linear correlation between the 3 hourly *ap* geomagnetic activity and the SW parameter changes. Linear predictions of the *ap* index from IMF and SW parameters were reported in a study by Smart and Shea (1996). The *Dst* and *ap* indices' response to SW perturbations are shown in the bottom panel of Figure 6.1.

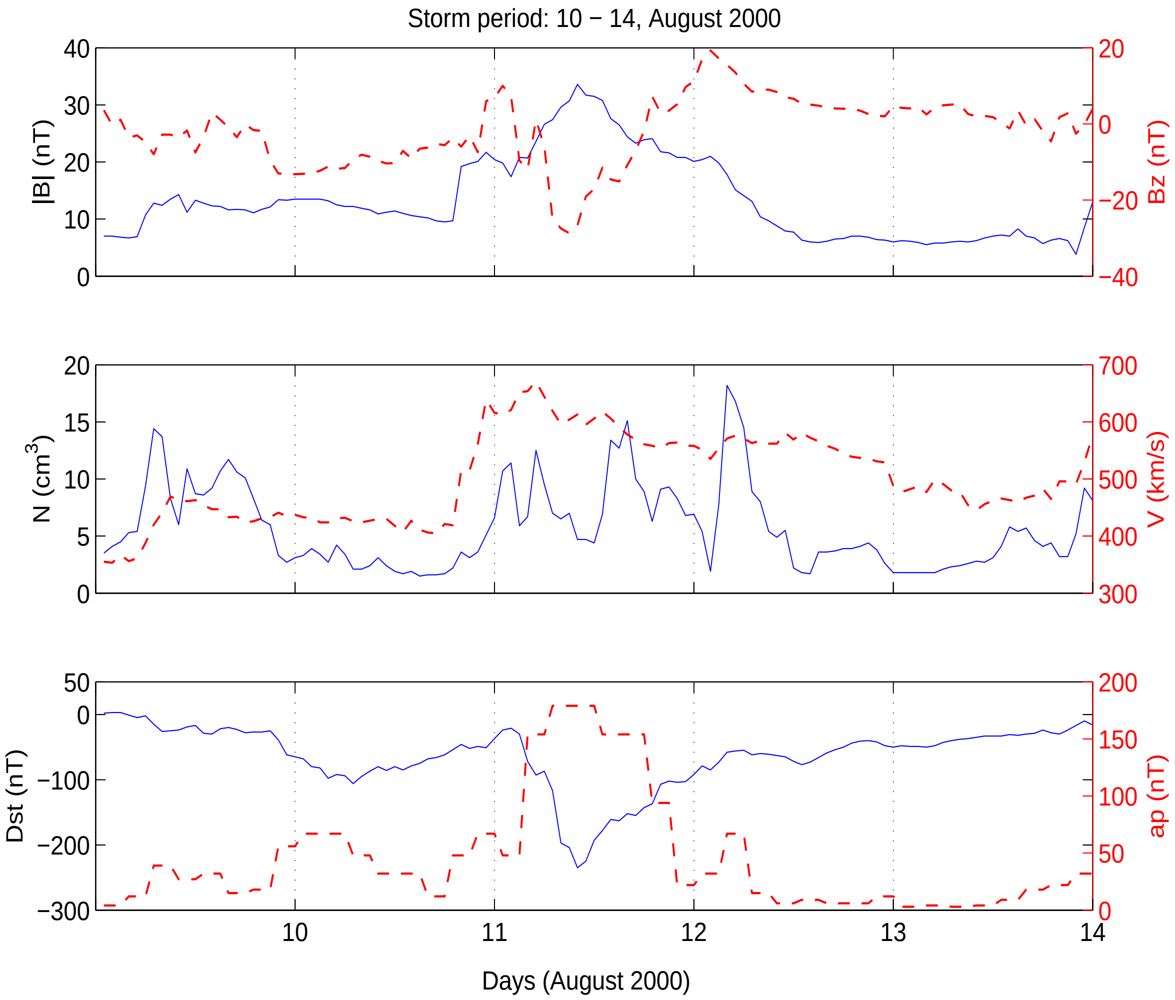


Figure 6.1: Variation of the SW and IMF parameters (B_z, Bt, V, N) and geomagnetic measurements $(Dst$ and $ap)$ during the 10-14/08/2000 magnetic storm.

6.3 Data preparation and development of the NN model

In this study, hourly-averaged OMNI-2 SW and IMF parameters $[B_t, B_z, V$ and $N]$ data were used for both training and testing. The OMNI-2 data set was created at NSSDC in 2003 to succeed the OMNI data created in the mid 1970's. It contains hourly SW magnetic field and plasma data, as well as solar and geomagnetic activity data. The data are in situ measurements from various spacecrafts in geocentric orbit about the L_1 Lagrange point. The data are provided by the NSSDC and are available online on its OMNIWEB: http://www.nssdc.gsfc.nasa/omniweb.html. An overview description of the OMNI data set can be found on the website:
http://omniweb.gsfc.nasa.gov/html/overview.html
Hourly geomagnetic Dst and ap data used in this study are provided by NGDC and are

available on the following websites:
ftp://ftp.ngdc.noaa.gov/STP/GEOMAG/dst.html.
ftp://ftp.ngdc.noaa.gov/STP/GEOMAG/Kp-ap.shtml.

6.3.1 Data preparation

Data covering an 11-year period from 1996 to 2006 were used for both training and validating the NN model. In total, 108 storm periods (7264 data points) were used for training. The selection of storm periods was based only on the Dst index values. Each magnetic storm period with a Dst index reaching a peak minimum $Dst \leq -80$ nT was selected. With the choice of -80 nT value, there is not only enough data points for training the NN, but this choice also offered the possibility of targetting the prediction of moderate to intense storms and thus, avoiding many minor storms for which the selection of isolated storm periods is difficult. Each storm period was selected including few hours of the preceding and subsequent magnetically quiet periods (in this case $Dst > -50$ nT). Eight (8) storm periods (indicated in Table 6.1) covering 679 hours, which were not part of the training set, were used to evaluate the prediction performance of the NN model. These storm periods, for the validation data set, were selected from the years 2000, 2001 and 2005 to evaluate the performance of the predictions during two different solar activity cycle phases. SW geoactivity dependency upon SC was reported in a study by Khabarova (2007).

The input parameter system was constructed in such a way that, for each SW and IMF parameter B_t, B_z, N, and V, three inputs were defined consisting of the consecutive (3 hours) preceding values defined as $B_z(-1), B_z(-2), B_z(-3)$. Inputs for the B_t, N and V parameters were defined in the same manner. The actual hour $Dst(0)$ or $ap(0)$ value was then presented to the network as the known output. For example, the magnetic storm of 31/10 - 02/11/2001 reached its Dst peak minimum on 1 November 2001 at 10:00 UT with a value of -106 nT. Therefore, for this particular hour, the output is $Dst(0) = -106$ nT and the corresponding inputs (for only the Bz parameter) are: $Bz(-1) = -10.4$ nT at 09:00 UT, $Bz(-2) = -10.4$ nT at 08:00 UT and $Bz(-3) = -10.4$ nT at 07:00 UT (The the three precedings hours had the same value of Bz). Constructing the input parameters in this way made it possible to account for the importance of the enhancement and long-lasting geoeffective conditions in the SW as precursors of GMS occurrence.

6.3.2 NN model development

A FFNN model was developed as per the details described in Chapter 3, Section 3.3.1. The constructed network consisted of 12 inputs and a corresponding known output, 1 hour Dst or ap (see Figure 6.2). Notice however, that for the ap NN prediction model, the known

Table 6.1: Storm periods used for the validation data set

Storm period	Start (hour)	End (hour)	Dst min(nT)
10-14/08/2000	00:00	23:00	-235
17-20/09/2000	15:00	23:00	-201
05-08/11/2000	08:00	11:00	-159
19-22/03/2001	10:00	12:00	-148
11-14/04/2001	14:00	23:00	-271
15-19/05/2005	00:00	23:00	-263
29-31/05/2005	19:00	23:00	-138
24-26/08/2005	00:00	23:00	-216

output *ap* for three consecutive hours is the same in the NN training and testing patterns due to the fact that *ap* is a three-hourly index.

Before being fed to the NN, the data set was randomly split into two independent data sets: training set and testing set. For this NN model, the training set comprised 70% and the testing set 30% of the original dataset. As there is no method to a priori determine the number of hidden nodes, it was determined systematically by testing various network architectures to find the one which would minimise the RMSE over the validation data set. The criterion used was the addition of a neuron to the hidden layer, training the NN, validating it and finally comparing the observed and predicted Dst or ap. The NN architecture that gave the optimum solution turned out to be identical for both Dst and ap algorithms with configuration 12:12:1, representing 12 input nodes (neurons) in the input layer, 12 hidden nodes in the hidden layer and 1 output neuron in the output layer. This optimum solution was obtained using the backpropagation algorithm and constructing the NN models in topological order. The convergency of the Dst algorithm was reached after 1 400 iterations, while the number of iterations for the ap index was 2 500. Figure 6.2 is a schematic illustration of the NN set-up used in this study.

Three statistical parameters, namely, the correlation coefficient (CC), RMSE and average relative variance (ARV) were used to evaluate the prediction performance over the validation data sets. The ARV is defined as the mean square error normalized by the variance of the data. Mathematically, the RMSE and ARV are defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(D_i - O_i)^2}, \tag{6.2}$$

and

$$ARV = \frac{\sum_{i=1}^{N} (D_i - O_i)^2}{\sum_{i=1}^{N} (D_i - \langle D_i \rangle)^2} \tag{6.3}$$

where $D(i)$, $\langle D_i \rangle$ and $O(i)$ denote observed, average and predicted Dst or ap respectively. It should be noted that the ERN network was also explored and compared to the developed FFNN to determine the optimum network architecture for each system, but since there was no significant difference in performance, only results of the FFNN models are presented.

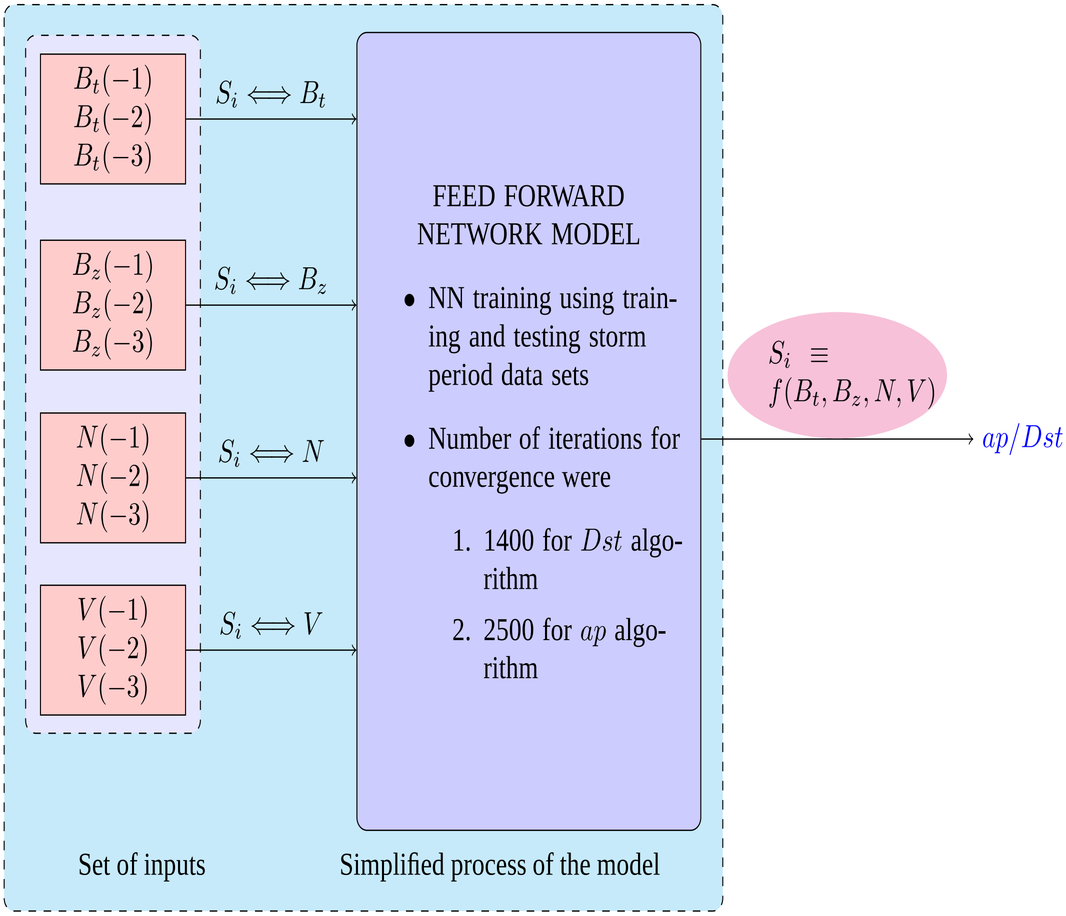


Figure 6.2: A schematic illustration of a simplified version of the developed NN model. Each SW and IMF parameter comprises 3 inputs representing the values of the previous three hours $(-1, -2, -3)$. There is only one output, the actual hour value of either Dst or ap. The storm indicator S_i (measured by either ap or Dst) is a function of 12 input parameters made up of the four parameters Bt, Bz, V and N.

6.4 Model prediction results

The performance of the developed model was tested on ten individual intense storms (with $Dst \leq -100$ nT) of which 8 represented unseen data (validation data set). The selected storms include the two latest strong storms (with $Dst \leq -200$ nT) of SC 23 which occurred in 2005. Figures 6.3 - 6.7 compare the predicted and observed Dst (ap) indices. Table 6.2 summarises the prediction performance of the model over all selected storms (for both Dst and ap indices) using the three statistical indices. Note that the first 2 storms listed in the Table 6.2 were part of the original training set, and are included in order to illustrate the performance of the model on seen data. The 8 remaining storms constitute unseen data and were used to finally validate the performance of the model.

i) The November 2003 and January 2004 magnetic storms

The greatest storm of SC 23 (with Dst peak minimum of -422 nT) is represented in Figure 6.3 (a) and (b). The developed Dst prediction model under predicts the main phase of the storm as well as the beginning of the recovery phase. Given that this data were a subset of the training data sets, it indicates the difficulty in predicting extreme storms. In their study Lundstedt et al. (2002) noted the existence of large differences between the Dst models and measurements during enhanced ring current conditions as well as during the recovery phase. This underprediction for this particular storm was also observed in the Dst prediction algorithm developed by Pallocchia et al. (2006).

It is interesting to observe in Figure 6.3 (b) and in Table 6.2 that the model produces better results when applied to ap prediction. Figure 6.3 (c) and (d) illustrate the prediction performance over the January 2004 magnetic storm. The Dst is very well predicted in all phases of the storm, but the ap prediction model does not perform well. Note that data for these two storms formed part of the training set of the NN algorithm. These two storms were chosen to verify the correct implementation of the model.

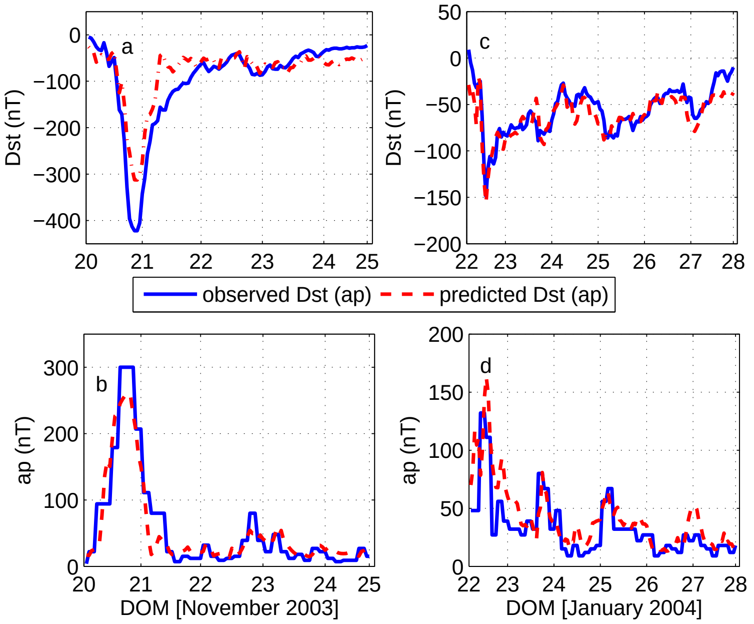


Figure 6.3: Comparison between observed and predicted *Dst* and *ap* indices for the 20-24 November 2003 and the 22-27 January 2004 magnetic storms.

ii) The August-September 2000 magnetic storms

In Figure 6.4, plots (a) and (b) illustrate the behaviour of the predicted and observed *Dst* and *ap* during the 12 August 2000 magnetic storm. The storm period had two successive peak minima. An intense storm (peak minimum *Dst* of −106 nT) occurred on 11 August at 06:00 UT followed by a strong storm with $Dst_{min} = -235$ nT on 12 August 2000 at 09:00 UT. As can be seen in the figure, the developed model predicts relatively well the *Dst* for the two main phases and the recovery phase of the storm on 11, 12 and 13 August 2000. Although the *ap* forecast overestimates the peak maximum of the storm, the overall performance of the *ap* prediction is better than the corresponding *Dst* index prediction.

Figure 6.4, plots (c) and (d) show the performance of the NN-*Dst* (*ap*) prediction model for another intense storm in September 2000. Table 6.2 shows the statistical values for evaluating the prediction performance of the model for each of the storms. These indicate that the *ap* index is better predicted by the model as compared to the prediction of the *Dst* index. The prediction of the *Dst* performs poorly during the initial phase of each of the two storms. A poor *Dst* prediction is also observed at the beginning of the recovery phase of the September storm. The difficulties of predicting the *Dst* index during the SSC as well as during the recovery phase of the storm were also been reported by Pallocchia et al. (2006) and Lundstedt et al. (2002).

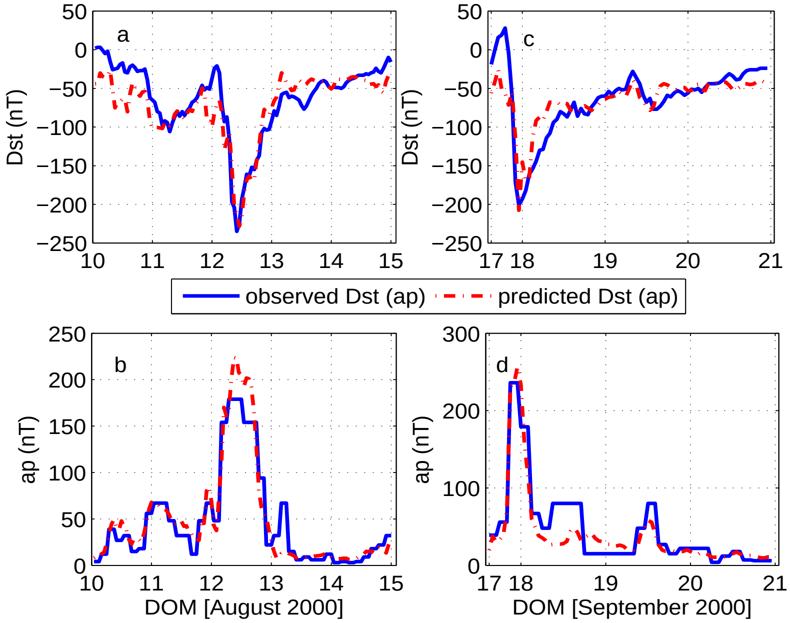


Figure 6.4: Comparison between the observed and predicted *Dst* and *ap* indices for the 10-14 August 2000 and 17-20 September 2000 magnetic storms.

iii) The November 2000 and March 2001 magnetic storms.

The two intense storms shown in Figure 6.5 had *Dst* peak minima in the range -200 nT $<$ *Dst* $<$ -100 nT and they were part of the validation data set. Notice that the *Dst* prediction was poor for the main phase, but improved for the recovery phase (Figure 6.5 (a)). The *Dst* recovery phase (March 2001 storm) in Figure 6.5 (c) was poorly predicted. Figure 6.5 in conjunction with the values provided in Table 6.2 indicate a better NN model prediction performance of the *ap* index compared to the prediction of the *Dst*.

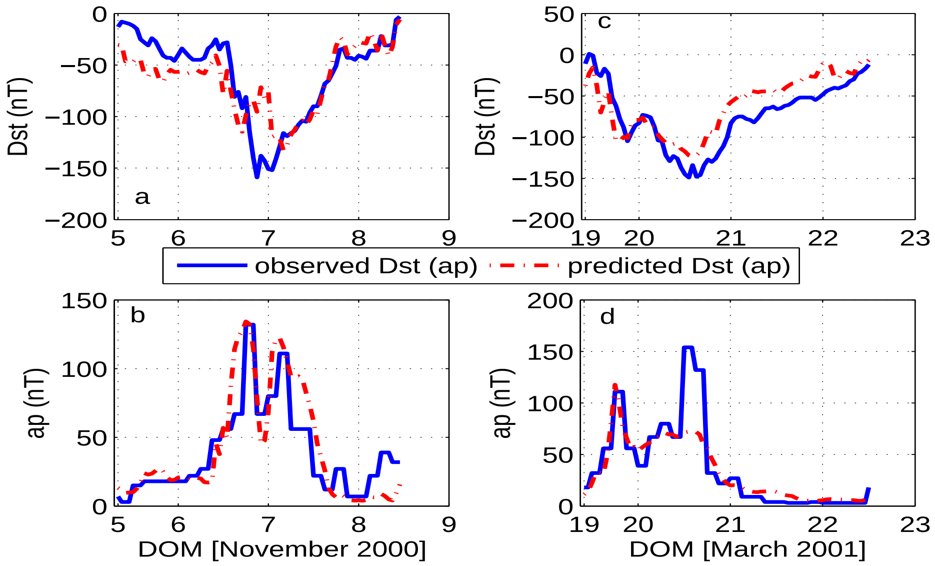


Figure 6.5: Comparison between the predicted and observed *Dst* and *ap* indices for the 5-8 November 2000 and the 19-22 March 2001 magnetic storms.

iv) The April 2001 and May 2005 magnetic storms

Figure 6.6 shows two intense storms with $Dst < -250$ nT which occurred in April 2001 and May 2005. An intense storm with peak minimum *Dst* of −271 nT occurred on 11 April 2001 at 23:00 UT and was followed by a slow recovery phase lasting about 3 days. The *Dst* prediction is not so good for the April 2001 storm during the recovery phase. The CC and ARV values indicate that the *ap* prediction performed better than the *Dst* prediction (Table 6.2).

The 15 May magnetic storm illustrated in Figure 6.6 (c) and (d) is the most recent strong storm in SC 23. The initial phase of the storm started at 02:00 UT on 15 May with an increased *Dst* lasting four hours and was associated with a sharp increase in the SW density. The main phase commenced at 05:00 UT with *Dst* sharply falling from 30 nT to reach the peak minimum $Dst = -263$ nT at 08:00 UT. Corresponding to this 3-hour sharp fall of the *Dst* index, the B_z component also turned southward with enhanced values. Except for the short-lived period corresponding to the sudden storm commencement, Figure 6.6 (c) shows that the *Dst* is well predicted in both the main and recovery phases of the storm with a CC of 0.87 and RMSE of 24.5 nT. Contrary to the previously described storms, the *ap* prediction in this case performed poorly, especially during the recovery phase, as indicated by all computed statistical values. The reason for this may be that the *Dst* response during the recovery phase increased gradually, while the *ap* index responded sharply by falling from 179 nT to only 27 nT in just one hour [11:00-12:00 UT] on day 15.

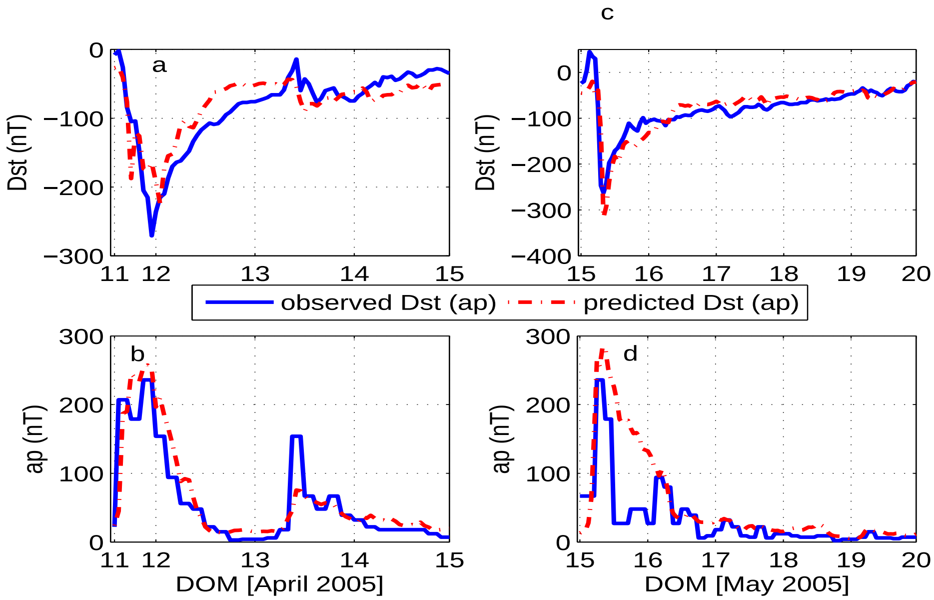


Figure 6.6: Comparison between the observed and predicted *Dst* and *ap* indices for the 11-14 April 2001 and the 15-19 May 2005 magnetic storms.

v) The 29-31 May 2005 and 24-26 August 2005 magnetic storms

Figure 6.7 (a) and (b) shows the *Dst* and *ap* predictions of 30 May 2005 magnetic storm. Visual inspection of this figure indicates a poor prediction of both *Dst* and *ap* indices. The magnetic storm in August 2005 is shown in Figure 6.7 (c) and (d). This figure shows that *Dst* behaviour is not well reproduced for the short period prior to the main storm phase. For this individual storm, the *ap* index is well reproduced with a CC of 0.90 and an ARV of 0.17.

The results of the *Dst* prediction obtained in this study are comparable with previous results obtained in similar studies. The CC of 0.85 was obtained for the observed and predicted *Dst* over the entire validation dataset. This value is close to the value of 0.88 reported in the *Lund Dst* model (Lundstedt et al., 2002). The average RMSE obtained for the 8 individual strong storms in the validation data set of this model is 26 nT. This value is comparable to the minimum RMSE given by Lundstedt et al. (2002) for storms with a $Dst < -100$ nT levels. The model developed by Pallocchia et al. (2006) produced results which are comparable to those of the current model in terms of CC. Their model achieved a CC of 0.83 compared to 0.85 value achieved in this model. Various models reported above were used for comparison purposes with the new model developed. However, it should be noted that none of the models used for comparison were developed in exactly the same way as this one. Table 6.2 provides all statistical values to evaluate the performance of this model for various individual storms.

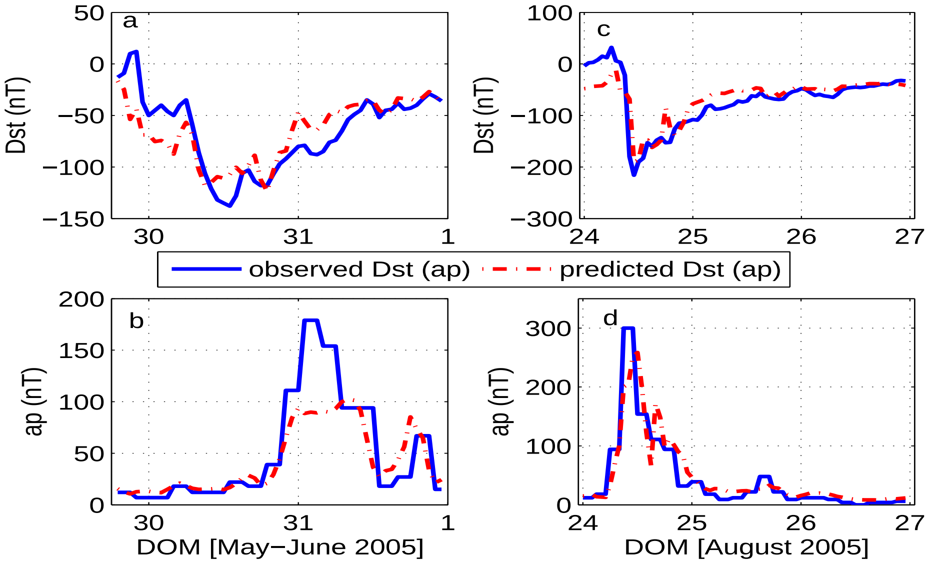


Figure 6.7: Comparison between the observed and predicted Dst and ap indices for the 29-31 May 2005 and 24-26 August 2005 magnetic storms.

Table 6.2: Computed RMSE, CC and ARV between the predicted and observed *Dst* (*ap*). Of the 10 storms, 2 were part of the training data set and 8 formed the validation data set.

Storm period	index	RMSE	CC	ARV
20-24/11/2003	Dst	45.1	0.94	0.22
	ap	25.8	0.94	0.11
22-27/01/2004	Dst	14.02	0.86	0.30
	ap	19.85	0.78	0.61
10-14/08/2000	Dst	24.06	0.88	0.25
	ap	17.67	0.94	0.12
17-20/09/2000	Dst	23.98	0.85	0.29
	ap	21.02	0.91	0.17
05-08/11/2000	Dst	24.62	0.81	0.35
	ap	20.52	0.87	0.17
19-22/03/2001	Dst	23.96	0.85	0.37
	ap	22.85	0.87	0.28
11-14/04/2001	Dst	29.2	0.86	0.27
	ap	32.5	0.88	0.24
15-19/05/2005	Dst	24.57	0.87	0.27
	ap	50.40	0.76	1.10
29-31/05/2005	Dst	21.77	0.80	0.34
	ap	30.9	0.86	0.34
24-26/05/2005	Dst	27.45	0.85	0.28
	ap	27.05	0.90	0.17

6.5 The predictability of the Hermanus geomagnetic K-index

The K-index quantifies disturbances in the horizontal component of the Earth's magnetic field with an integer ranging from 0 (quiet) to 9 (extremely disturbed) as a measure of magnetic activity. The study presented in this chapter has been extended to investigating the NN predictability of the K-index recorded at the Hermanus Magnetic Observatory ($34^0 25.5' S; 19^0 13.5' E$). This magnetic observatory is hosted by SANSA Space Science (Hermanus) and is an active participant in the worldwide network of magnetic observatories (INTERMAGNET), which monitors and models variations of the Earth's magnetic field. The K-index represents a measure of the local geomagnetic activity response to solar storms and related SW disturbances (see: http://spaceweather.hmo.ac.za/index). Figure 6.8 illustrates the intense magnetic storm on 3-5 August 2010 ($K = 6$ measured at Hermanus) which followed the CME eruption on 1 August 2011 at 13:42 UT.

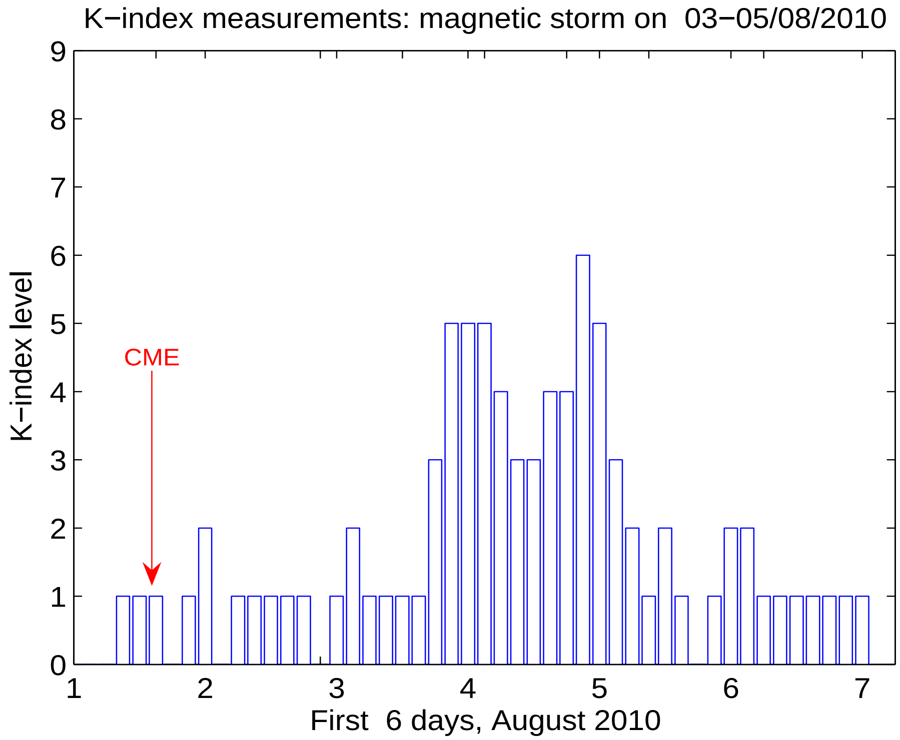


Figure 6.8: The Hermanus K-index measurements during the magnetic storm on 3-5 August 2010.

6.5.1 Related work and motivation

The three-hourly geomagnetic K-index is closely related to the global planetary magnetic Kp-index. To produce the three-hourly Kp-index, three-hourly K-indices obtained from a network of mid-latitudes geomagnetic observatories, corrected for the respective geomagnetic latitudes and then averaged (Mayaud, 1980; Boberg et al., 2000). The planetary Kp-index plays a key role in space weather as it is used in many magnetospheric and ionospheric models (Wing et al., 2005). Regional ionospheric models, e.g South African TEC predictions models and the South African Bottomside Ionospheric Model (SABIM), take into account the local magnetic conditions by using the a-index, which are directly derived from the locally recorded K-index (McKinnell, 2002; Habarulema, 2010). An accurate model to predict the K-index would, therefore, represent a significant contribution towards improving the ionospheric and other regional space weather models that consider magnetic activity as input.

Models for predicting the Kp index (using SW and IMF input parameters) have been developed by other group (Costello, 1997; Boberg et al., 2000; Wing et al., 2005). Wing et al. (2005) pointed out the difficulties related to Kp-index prediction when the index is very useful to space weather (during storm periods with $Kp > 5$) . One would expect the same difficulties when predicting the related K-index. In this study the application of NN techniques to predicting the locally-measured geomagnetic K-index was explored. The correlation between the Kp geomagnetic index and various SW parameters has been established (e.g Papitashvili et al., 2000). Figure 6.9 shows the correlation between various SW parameters and the Hermanus K-index, while Figure 6.10 illustrates the local K-index response to the SW and IMF disturbances during a magnetic storm.

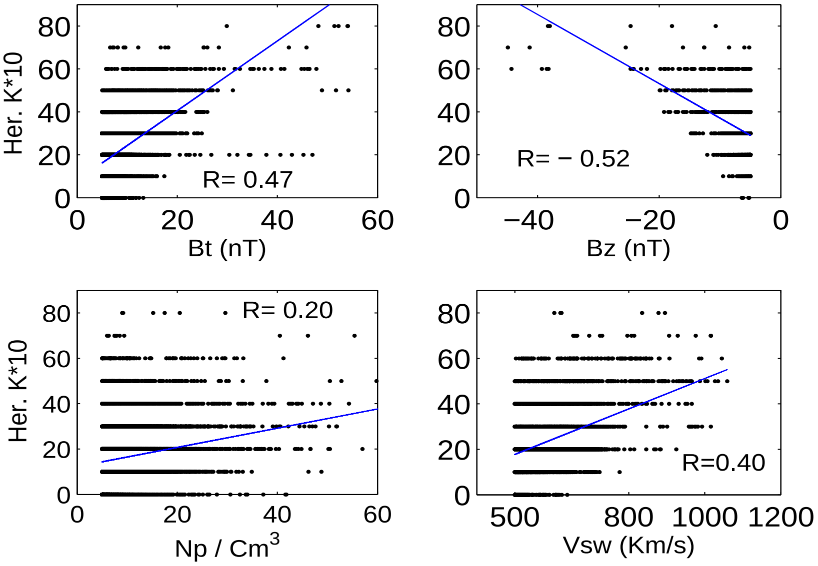


Figure 6.9: Plots illustrating the correlation between various SW and IMF parameters and the Hermanus K-index

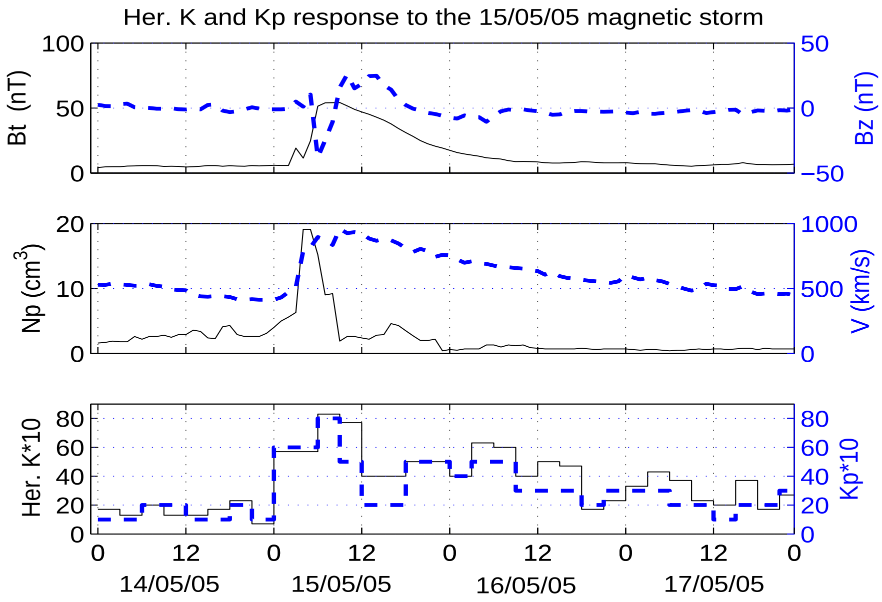


Figure 6.10: Graph showing the response of the Kp and the Hermanus K indices to various SW and IMF parameters during a storm

6.5.2 NN model development for K-index prediction

For the prediction of the Hermanus K-index, the one-layer ERN described in Section 3.3.3 was used. A particular feature of this type of NN is that the outputs from the hidden nodes are fed back as inputs via the context nodes, as illustrated by Figure 3.2.

The data used are the same OMNI-2 SW and IMF parameters that were used for the Dst and ap prediction. The training and testing data sets consist of storm periods selected within SC 23 [1996 - 2006]. Selected storm periods during years 2004 and 2005 (688 data points) were excluded from the training process and were used to test the performance of the model. Each storm period was defined as having a $K \geq 5$ at least once, with each preceded and followed by a quiet magnetic period of at least 12 hours. Note that both the measured and the predicted K-index are three-hourly indices. Therefore, an input row (pattern) is made up of four SW parameter values (V, Bt, Bz, N), each one being the average of the three preceding hourly values. The output of the NN is a 3 hourly K-index. During the training process, the optimal number of hidden nodes was systematically determined by varying the number of hidden nodes. The RMSE and the CC were the statistical parameters used to characterise the prediction performance of the model. The network with the optimum performance was reached having 5 hidden nodes (4:5:1 configuration).

6.5.3 Prediction results and discussion

Table 6.3 shows the different NN configurations that were investigated with the corresponding prediction performance, which was evaluated by calculating the RMSE and the CC over the whole validation data set. Table 6.3 indicates that the developed model performs better when predicting the global Kp-index than it does for the prediction of the local Hermanus K-index. This is to be expected, since the global Kp-index is derived from various K-indices averaged and corrected to their respective magnetic latitudes observatories. Figure 6.11 shows the correlation between the measured and predicted Hermanus K-index, and between the measured and predicted global Kp-index. Although the data set used is not the same, the CC= 0.88 obtained by this model between the measured and predicted Kp-index is comparable to those obtained in previous studies. Wing et al. (2005) obtained a CC= 0.84, while a study by Boberg et al. (2000) reported a CC of 0.76. The comparison described above (Kp-index prediction performance) is an indication that the Hermanus K-index is relatively well predicted.

Table 6.3: Determination of the optimum NN architecture for both the Hermanus K-index and the global Kp-index predictions

Network	Her. K		Kp	
	RMSE	CC	RMSE	CC
4:4:1	10.8416	0.71	9.4223	0.87
4:5:1	10.1515	0.76	9.3945	0.87
4:6:1	10.3427	0.75	9.1942	0.88
4:7:1	10.1915	0.76	9.4011	0.88
4:8:1	10.3581	0.75	9.3961	0.87

The prediction performance of this model was tested on four selected intense storms which were part of the validation data set (not included in the validation data set). The two of the storms occurred in 2004 and 2 others in 2005. The two selected storms in 2004 were both long-duration intense storms, behaving like two storms in one with two peak maxima. The storm in Figure 6.12 (a) extends for 5 days (from 7 to 12 of November 2004) reaching two peak maxima of $K = 7$ within the storm period. The global Kp-index reached a peak maximum of $Kp = 8.7$ on 8 November at 03:00 UT and on 9 November at 18:00 UT. The storm in July 2004 (plot b of Figure 6.12) also had two peak K-index maxima, one on 25 July at 6:00 UT (K=6) and another on 27 July at 21:00 UT with K=7.

The two test storms in 2005 represented in Figure 6.12 (c) and (d) represent the most recent

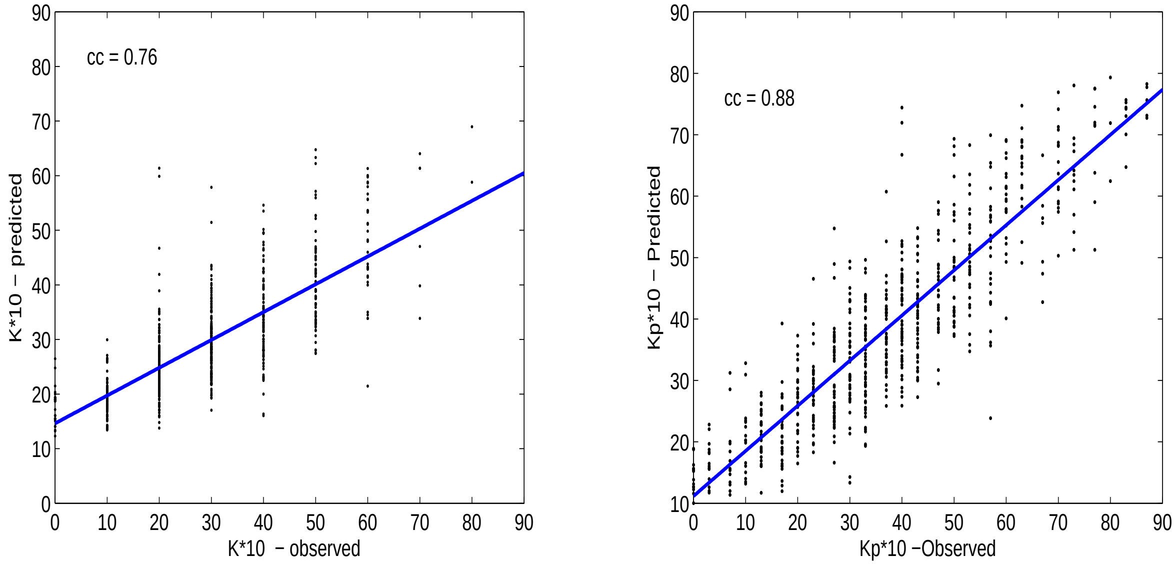


Figure 6.11: Correlations between the observed and predicted Hermanus K index (left) and global Kp index (right). The indicated CC values were computed over the whole validation data set.

Table 6.4: Prediction performance of the model tested on the 4 individual storms

Storm period	Her. K		Kp	
	RMSE	CC	RMSE	CC
24-28/07/2004	9.2842	0.81	10.2334	0.90
07-12/11/2004	9.2613	0.86	8.9464	0.92
14-17/05/2005	15.2374	0.65	14.3149	0.77
23-25/08/2005	9.4044	0.91	6.8211	0.94

and greatest magnetic storms of SC 23. Both the storms reached the peak Hermanus K-index value of 8. Both Figure 6.12 and Table 6.4 illustrate the model's K-index prediction performance on the selected individual storms. Figure 6.13 indicates a slightly improved performance when the model is applied to the prediction of the global Kp index.

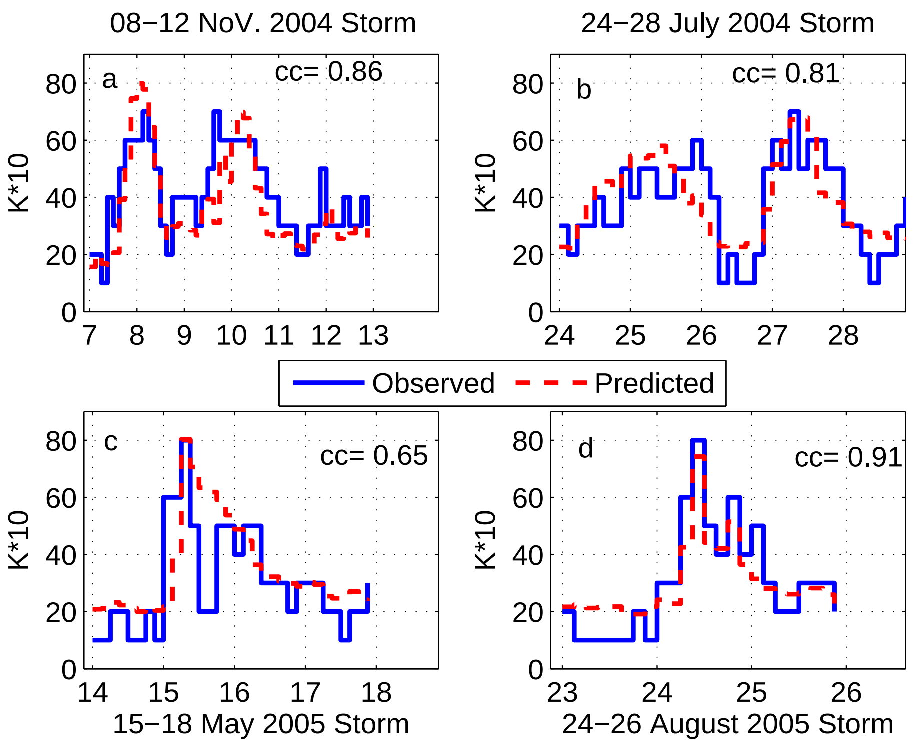


Figure 6.12: Graphical representation of the Hermanus K-index prediction performance as performed over four individual storms.

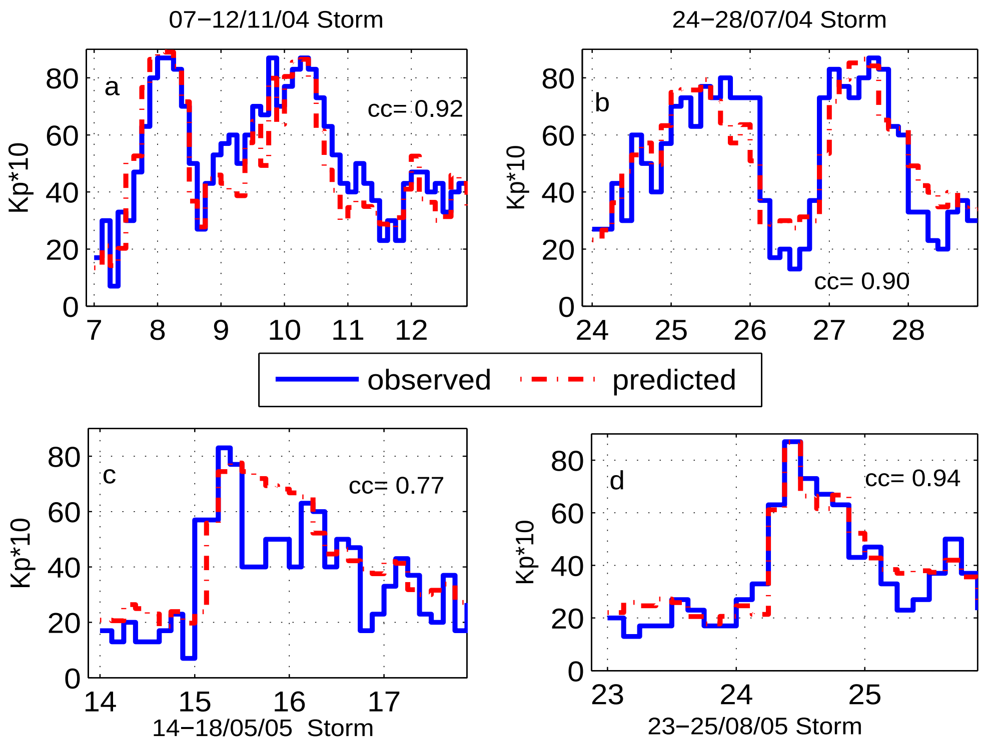


Figure 6.13: Graphical representation of the global Kp-index prediction performance, as performed on four individual storms.

6.6 Summary

Reliable ERN-based models for predicting the *Dst* index using SW and IMF input parameters have been sufficiently developed (e.g Wu and Lundstedt, 1997; Pallocchia et al., 2006). Previous studies focused on developing GMS prediction models based on the *Dst* measurements only. However, recent findings suggest that the *Dst* alone can be a poor indicator of the properties of a storm (Borovsky and Denton, 2006). The primary aim of the study described in this Chapter was to explore an alternative NN model for the prediction of the storm strength as measured by both the *Dst* and the *ap* index.

The predictions of both the indices were compared and in general, the results obtained demonstrate an improved prediction performance of the *ap* compared to the prediction of the *Dst* index. One reason for the observed differences in the prediction of the two indices can be associated to the time variability as well as their respective response to external variations. It was observed that the seasonal variation associated with an increase in geomagnetic activity at the equinoxes is higher for *Dst* as compared to *ap* index (Saba et al., 1997). In their analysis, Saba et al. (1997) also observed a finite time lag between the two indices which they associated to the fact that the *Dst* index monitors the ring current, while the *ap* index (like the *AE*) respond to ionospheric currents with faster response time to external varitions. Further reasons which might explain these observed differences in the prediction performance can be a subject of further investigation

The ability to predict the *ap* index during storm time is very useful for ionospheric modelling (Perrone and de Franceschi, 1999; Araujo-Pradere et al., 2003). The predictability of the locally measured K index (during storm time) from SW and IMF parameters with the use of ERN has been demonstrated successfully as shown by the obtained results. The ability to predict the Hermanus *K*-index provides a means towards improving various regional space weather models (e.g. ionospheric models) that consider magnetic activity as input.

Chapter 7

Conclusion

The topic of GMS is very important for space weather. The research described in this thesis was conducted in response to the need to continuously monitor geoeffective solar events, as well as the need to improve the models for the prediction of GMS. This study consists of two main parts namely (a) an investigation of probable solar and IP properties associated with GMS during SC 23, and (b) the development of NN-based models to predict the occurrence and intensity of GMS using both solar and IP input parameters. This last chapter outlines the main results of the study, discuss the limits of its success and proposes the way forward aiming to improve the models for predicting GMS.

7.1 Summary of results

7.1.1 The main sources of GMS in SC 23

Investigation of solar and IP phenomena that were probable causes of 229 GMS events in SC 23 was presented in Chapter 4. The statistical results of this analysis are based on the investigated events listed in Table 4.1. The most probable sources of GMS in SC 23 were discussed in the last section of Chapter 4. A comparative analysis was presented for moderate and intense storms with regard to their association with various solar and IP geoeffective properties. Compared to moderate storms, intense storms were generally characterised by:

- A greater association with full halo CMEs, mostly originating close to the disc centre
- A greater association with ICMEs at 1AU
- Higher values of associated IMFB_s, CME and SW speeds
- Association with the majority of investigated multiple CME-driven storm events

This study shows that partial halo CME-driven storms were mostly moderate storms and were linked to halo CMEs which generally originated far from the disk centre (outside $\pm 45^0$).

In addition, almost half of all investigated moderate storms were without a background of halo CMEs within a 5-day time window. In this study, the non-halo CME-associated storms were assumed to be either CIR-driven (linked with coronal holes), or were caused by geoeffective non-halo CMEs.

It is important to indicate that the list of geoeffective halo CMEs (and associated solar and IP properties) provided in Table 4.1 is not exhaustive. As indicated by Zhang et al. (2007), some storm events are often observed with uncertain sources, due to the lack of noticeable eruption signatures in the solar corona. The listing (Table 4.1) only provides an estimate of probable causes of GMS for an average 11-year SC activity. Note that this study investigated 92 additional storm events ($Dst < -50$ nT) which are not listed in RC's catalogue of ICMEs and related GMS (Richardson and Cane, 2010). On the other hand, while the analysis by Zhang et al. (2007) focused on the potential solar and IP sources of intense storms ($Dst \leq -100$ nT) in SC 23, the study described in this thesis extends to investigating the probable solar and IP sources of moderate storms in SC 23.

7.1.2 On predicting the geoeffectiveness of halo CMEs

Chapter 5 discussed the NN-based model developed for predicting the probability occurrence of GMS subsequent to halo CME eruptions. The model uses as input a combination of identified solar and IP geoeffective parameters, used either as numeric or binary values. The performance of a very well-trained and optimised NN model was tested using storm events not involved in the training process. The developed NN model proved to be more accurate in predicting the occurrence of intense GMS than moderate GMS. It is important to remember that the NN model was not trained to predict the probability occurrence of GMS events that are not CME-driven (e.g. CIR-driven storms).

The performance of the NN model also shows that IP events (i.e. ICMEs, shocks and associated SW parameters, Vsw and Bs) contribute significantly to improving the predictability of GMS occurrence. This is consistent with previous findings (e.g. Srivastava, 2005). It was shown that the NN model with only solar input parameters ($AW, Vcme, cfi$, see [A] in Table 5.1 and Table 5.3) lead to a lower performance. However, such a model is practically very important for space weather as it offers a long warning period from the launch of a solar transient event to the occurrence of a storm event. Spacifically, the NN model developped offer an alternative storms prediction technique, combining both solar and IP input parameters.

7.1.3 Modelling geomagnetic indices during storm time

From a practical point of view, the prediction of the strength of a storm is as important as predicting its occurrence. Geoeffetive conditions in the SW are often used for predicting the magnitude of a storm. As detailed in Chapter 6, SW and IMF parameters (Bt, Bz, Vsw and N) were used as inputs in a FFNN to predict the Dst and ap indices during storm time. For the Dst prediction, the developed model produced results comparable to similar previous studies (Wu and Lundstedt, 1996, 1997; Lundstedt et al., 2002). Due to the importance of the ap index for ionospheric storm time modelling (Araujo-Pradere et al., 2003), its predictability by means of NNs was explored and the results were compared with those of the Dst index prediction.

A separate ERN-based model was developed to investigate the predictability of the locally measured K-index, using the same described SW input parameters. The results obtained compared well with previous Kp (closely related to K-index) predictions made by other researchers (Wing et al., 2005; Boberg et al., 2000) and shown a good performance. The results obtained from the developed NN models are in line with what is already known about the SW control of geomagnetic activity. With a knowledge of the SW velocity, density, as well as the IMF strength and orientation, it is possible to predict well the energization of the ring current and reproduce accurately the magnetic measurements recorded by ground based magnetometers (Russell, 1986).

The developed model constitutes a step towards achieving real-time forecasts of the locally (Hermanus) measured K-index. If achieved, the real-time prediction of the K-index will contribute significantly to improving regional ionospheric modelling as well as other SANSA-RWC space weather models that consider the locally measured magnetic activity as input.

7.2 Limitations of the study

Despite advances in modern technologies to detect solar transient events, predicting their geoeffectiveness is still challenging. GMS are complex phenomena involving the Sun, the IP medium, the ionosphere and the magnetosphere. Efforts toward predicting the occurrence and strength of GMS were previously attempted (Dryer, 1998; Dryer et al., 2004). However, physics-based models for the prediction of GMS are still difficult to achieve due to the complexity and non-linearity of the solar-terrestrial system, with its physics still not completely understood (Fox and Murdin, 2001; Schwenn et al., 2005). Pending more extensive real-time verification of the physics-based models, current GMS modelling and prediction rely mostly on empirical approaches, including that of NNs which have been used in this study.

The current study involved only solar and IP data for the 11 years of SC 23. It is anticipated that the new data from SC 24 might be used to further improve GMS prediction models. Additional data from SC 24, which has already entered its active phase, are expected to contribute significantly towards the improvement of the developed NN models. This will be facilitated by the incorporation of new data from the STEREO mission, which provides a 3D structure of CMEs, tracking them all the way from the Sun to the IP medium.

7.2.1 Proposed ways to improve GMS prediction models

In this study, NN-based models for predicting the occurrence and strength of GMS were developed. There are two aspects through which the developed models can be improved. One aspect of improving the prediction of GMS occurrence would consist in considering a large database of halo CME-associated input parameters to the NN models (i.e. incorporating new data of SC 24), but then it will need to wait for a period corresponding to a complete SC in order to improve the training of NNs. The importance of considering a larger database for the improvement of any NN-based model has been mentioned in previous studies (Haykin, 1994; McKinnell, 2002). Another aspect for improvement would be the use of additional input parameters to account for more mechanisms which drive GMS. New NN-input parameters could include, for example, both the heliographic location of the storm-associated CMEs and their magnetic orientation. It was shown in this study that storm occurrence and strength are both dependent on the location of the driving CMEs on the solar surface. The use of CME and ICME magnetic orientation and magnitude for GMS prediction has already been explored (Kim et al., 2010; Wood and Howard, 2009). The incorporation of these new parameters in future NN-based models would be facilitated by STEREO data availability.

IP type II radio bursts measured in longer wavelengths are often indicators of CME-driven shocks. Type II radio bursts data should also be considered as input for a NN model and are expected to contribute towards improving the predictability of GMS occurrence. By considering coronal hole information (data) within the NN input space, the predictability of the occurrence of moderate storms can be improved since, according to the analysis described in Chapter 4, moderate storms are mostly CIR-driven and hence directly linked to coronal hole phenomena.

In conclusion, the objectives of the thesis as defined in Section 1.4 of Chapter 1 were satisfactorily achieved. GMS events which occurred in SC 23 were identified. These storm events were next classified and characterised on the basis of their solar and IP probable sources during the period 1996 - 2006. This analysis provided an estimate of the geoeffectiveness of solar and IP transient events over an average 11-year period of solar activity. Finally, the identified solar and IP geoeffetive parameters were used to develop reliable NN-based models for the prediction of the occurrence and strength of GMS.

On the basis of the developed GMS prediction models, an operational database for the real-time forecast of the occurrence and strength of GMS will be constructed for use by the SANSA-RWC in Hermanus. When completed, it is expected that various regional space weather models which incorporate storm conditions (i.e. ionospheric and GIC models), will significantly improve as well.

References

Antiochos, S. K., DeVore, C. R. and Klimchuk, J. A., A model for coronal mass ejections. Astrophysical journal, 510, 485–493, 1999.

Araujo-Pradere, E. A., Fuller-Rowell, T. J. and Biltza., D., Validation of the STORM response in IRI 2000. Journal of Geophysical Research, 108(A3), doi:10.1029/2002JA009720, 2003.

Babcok, H. W., The topology of the Sun's magnetic field and the 22-Year cycle. Astrophysical Journal, 2, 572–587, 1961.

Baumjohann, W. and Treumann, R. A., Basic Space Plasma Physics. Imperial College Press, London, 1996.

Biermann, L., Kometenschweife und solare Korpuskularstrahlung. Zeitschrift fur Astrophysik, 29, 274–286, 1951.

Bishop, C. M., Neural Networks for Pattern Recognition. Oxford University Press Inc., New York, USA, 1995.

Boberg, F. P., Wintoft, P. and Lundstedt, H., Real time Kp prediction from solar wind data using neural networks. Physics and Chemistry of the Earth, 25(A04203), 275–280, 2000.

Borovsky, J. E. and Denton, M. H., Differences between CME-driven storms and CIR-driven storms. Journal of Geophysical Research, 111(A07S08), doi:10.1029/2005 JA011447, 2006.

Boteler, D. H., Geomagnetic Hazards to Conducting Networks. Natural Hazards, 28, 537–561, 2003.

Brueckner, G. E., Delaboudiere, J. P., Howard, R. A., Paswaters, S. E., Cyr, O. C. S., Schwenn, R., Lamy, P., Simnett, G. M., Thompson, B. and Wang, D., Geomagnetic storms caused by coronal mass ejections(CMEs): March 1996 through June 1997. Geophysical Research Letters, 25, 3019–3022, 1998.

Brueckner, G. E., Howard, R. A., Koomen, M. J., Korendyke, C. M., Michels, D. J., Moses, J. D., Socker, D. G., Dere, K. P., Lamy, P. L., Lleberia, A., Bout, M. V., Schwenn, R., Simnett, G. M., Bedford, D. K. and Eyles, C. J., The Large Angle Spectroscopic Coronagraph, (LASCO). Solar Physics, 162, 357–402, 1995.

Burlaga, L. F., Magnetic clouds. In Physics of the inner heliosphere, eds. R.Schwenn and Marsch, E., vol. II, Springer, Berlin, 1991.

Burlaga, L. F., Behannon, K. W., Hansen, S. F., Pneumann, G. W. and Feldman, W. C., Sources of magnetic fields in recurrent IP streams. Journal of Geophysical Research, 83, 4177–4185, 1978.

Burlaga, L. F. and Lepping, R. P., The causes of recurrent geomagnetic storms. Planetary and Space Physics, 25, 1151–1161, 1977.

Burton, R. K., McPherron, R. L. and and, C. T. R., An empirical relationship between interplanetary conditions and Dst. Journal of Geophysical Research, 80, 4204–4214, 1975.

Cai, L., Ma, S. Y. and Zhou, Y. L., Prediction of SYM-H index during large storms by NARX neural network from IMF and solar wind data. Annales Geophysicae, 28, 381–393, 2010.

Cane, H. V. and Richardson, I. G., Interplanetary coronal mass ejections in the near-Earth solar wind during 1996-2002. Journal of Geophysical Research, 108(A4), doi:10.1029/2002JA009817, 2003.

Cargill, P. J. and Harra, L. K., Handbook of the Solar-Terrestrial Environment, chap. 5. Springer-Verlag, Berlin, 2007.

Chapman, S. and Bartels, J., Geomagnetism. Oxford University Press, Oxford, 1940.

Chian, A. C.-L. and Kamide, Y., Handbook of the Solar-Terrestrial environment, chap. 1. Springer-Verlag, Berlin, 2007.

Conway, A. J., Time series, neural networks and the future of the Sun. New Astronomy Reviews, 42, 343–394, 1998.

Costello, K. A., Moving the Rice MSFM into a realtime forecast mode using solar wind driven forecast models. Ph.D. thesis, Rice University, Houston, Texas, 1997.

Crooker, N. U., Solar and heliospheric geoeffetive disturbances. Journal of Atmospheric and Solar-Terrestrial Physics, 62, 1071–1085, 2000.

Crooker, N. U. and McAllister, A. H., Transients associated with recurrent storms. Journal of Geophysical Research, 102(A7), 14041–14047, 1997.

Cybenko, G., Approximation by supervisions of sigmoidal function. Mathematics of Control, Signals and Systems, (2), 303–314, 1989.

Cyr, O. C., Howard, R. A., Sheeley, N. R., Plunkett, S. P., Michels, D. J., Paswaters, S. E., Koomen, M. J. and Simnett, G. M., Properties of coronal mass ejections: SOHO LASCO observations from January 1996 to June 1998. Journal of Geophysical Research, 105, 169–185, 2000.

Daglis, I. A., Thorne, R. M., Baumjohann, W. and Orsini, S., The terrestrial ring current: origin, formation and dacay. In Reviews of Geophysics, vol. 37, 407 – 438, American Geophysical Union, 1999.

Dryer, M., Multidimentional, magnetohydrodynamics simulation of solar-generated disturbances: Space weather forecasting of geomagnetic storms. AIAA Journal, 36(365-370), doi:10.2514/2.405, 1998.

Dryer, M., Smith, Z., Fry, C. D., Sun, W., Deehr, C. S. and Akasofu, S. I., Real-Time shock arrival predictions during the Halloween 2003 epoch. Space Weather, 2(S09001), doi:10.1029/2004 SW000087, 2004.

Echer, E., Gonzalez, W. D., Tsurutani, B. T. and Gonzalez, A. L. C., Interplanetary conditions causing intense geomagnetic storms ($Dst \leq -100$ nT) during solar cycle 23 (1996 - 2006). Journal of Geophysical Research, 113(A05221), doi:10.1029/2007JA012744, 2008.

Elman, J. L., Finding Structure in Time. Cognitive Science, 14, 179–211, 1990.

Fausett, L. V., Fundamentals of Neural Networks: architecture, algorithms and applications. Prentice-Hall, Inc., 1994.

Fessant, F., Bengio, S. and Collobert, D., On the prediction of solar activity using different neural network models. Annales Geophysicae, 14(1), 20–26, 1996.

Forbes, T. G., A review on the genesis of coronal mass ejections. Journal of Geophysical Research, 105(A10), 153–165, 2000.

Forbes, T. G., Linker, J. A., Chen, J., Cid, C., Kóta, J., Lee, M. A., Mann, G., Mikić, Z., Potgieter, M. S., ad G. L. Siscoe, J. M. S., vainio, R., Antiochos, S. K. and Riley, P., CME theory and models. Space Science Reviews, 123, 251–302, 2006.

Fox, N. and Murdin, P., Solar-Terrestrial Connection: Space Weather Predictions. In Encyclopedia of Astronomy and Astrophysics, ed. Murdin, P., 2416, Institute of Physics Publishing, Bristol, 2001, doi:10.1888/0333750888/2416.

Gold, T., Magnetic storms. Space Science Reviews, 1(1), 100–114, 1962.

Gonzalez, W. D., Joselyn, J. A., Kamide, Y., Kroehl, H. W., Tsurutani, B. T., Vasyliunas, V. M. and Rostoker, G., What is a Geomagnetic storm? Journal of Geophysical Research, 99, 5771–5792, 1994.

Gonzalez, W. D., Lago, A. D., de Gonzalez, A. L. C., Vieira, L. E. A. and Tsurutani, B. T., Prediction of peak-Dst from halo CME/magnetic cloud-speed observations. Journal of Atmospheric and Solar-Terrestrial Physics, 66, 161–165, 2004.

Gonzalez, W. D. and Tsurutani, B. T., Criteria of interplanetary parameters causing intense magnetic storms ($Dst < -100$ nT). Planetary and Space Science, 35, 1101–1109, 1987.

Gopalswamy, N., Consequences of Coronal Mass Ejections in the Heliosphere. Sun and Geosphere, 1(2), 5–12, 2006a.

Gopalswamy, N., Properties of Interplanetary Coronal Mass Ejections. Space Science Reviews, 124, 145–168, 2006b.

Gopalswamy, N., Solar connections of geoeffective magnetic structures. Journal of Atmospheric and Solar-Terrestrial Physics, 70, 2078–2100, 2008.

Gopalswamy, N., Coronal mass ejection and space weather. In Climate and Weather of the Sun-Earth System (CAWSES): selected papers from the 2007 Kyoto Symposium, eds. Tsuda, T., Fujii, R., Shibata, K. and Geller, M. A., 77–120, Terrapub, Tokyo, 2009a.

Gopalswamy, N., Halo coronal mass ejections and geomagnetic storms. Earth Planets Space, 61, 1–3, 2009b.

Gopalswamy, N., Yashiro, S. and Akiyama, S., Geoeffectiveness of halo coronal mass ejections. Journal of Geophysical Research, 112(A06112), doi:10.1029/2006JA012149, 2007.

Gosling, J. T., The solar flare myth. Journal of Geophysical Research, 98, 937–949, 1993.

Gosling, J. T., Bame, S. J., McComas, D. J. and Philips, J. L., Coronal mass ejections and large geomagnetic storms. Geophysical Research Letters, 17, 901–904, 1990.

Gosling, J. T., McComas, D. J., Phillips, J. L. and Bame, S. J., Geomagnetic activity associated with passage of IP shock disturbances and Coronal mass ejections. Journal of Geophysical Research, 96, 7831–7839, 1991.

Habarulema, J., A contribution to TEC modelling over Southern Africa using GPS data. Ph.D. thesis, Rhodes University, Grahamstown, South Africa, 2010.

Hady, A. A., Descriptive study of solar activity sudden increase and Hallloween storms of 2003. Journal of Atmospheric and Solar-Terrestrial Physics, 71, 1711–1716, 2009.

Hathaway, D. H., Wilson, R. M. and Reichmann, E. J., A synthesis of solar cycle prediction techniques. Geophysical Research Letters, 104(A10), 375–388, 1999.

Haykin, S., Neural networks: a comprehensive foundation. Macmillan Publishing Company, 1994.

Hertz, J., Krogh, A. and Palmer, R. G., Introduction to the Theory of Neural Computation, lecture notes Vol.1. Santa Fe Institute Studies in the sciences of complexity, Addison-Wesley, Redwood City, 1991.

Howard, R. A., Michels, D. J., Sheeley, N. R. J. and Koomen, M. J., The observation of a coronal transient directed at Earth. Astrophysical Journal, 263, doi:10.1086/18392, 1982.

Hundhausen, A. J., Introduction to space physics, chap. 4. Cambridge University Press, 1995.

Illing, R. M. E. and Hundhausen, A. J., Disruption of a coronal streamer by an eruptive prominence and coronal mass ejection. Journal of Geophysical Research, 91, 10951–10960, 1986.

Kahler, S. W., Solar Flares and Coronal mass Ejections. Annual Review of Astronomy and Astrophysics, 30, 111–141, 1992.

Kallenrode, M. B., Space physics: An Introduction to Plasmas and Particles in the Heliosphere and Magnetospheres. ©Springer-Verlag Berlin Heidelberg, Heidelberg, Germany, 1998.

Kamide, Y., Baumjohann, W., Gonzalez, W. D., Grande, M., Joselyn, J. A., McPheron, R. L., Phillips, J. L., Reeves, E. G. D., Rostoker, G., Sharma, A. S., Singer, H. J. and nd V. M. Vasyliunas, B. T. T., Current understanding of magnetic storms: Storm-substorm relationships. Journal of Geophysical Research, 103, 705–728, 1998.

Kamide, Y. and Maltsev, Y. P., Handbook of the Solar-Terrestrial Environment, chap. 14. Springer-Verlag Berlin, Heidelberg, 2007.

Khabarova, O. V., Current problems of magnetic storm prediction and possible ways of their solving. Sun and Geosphere, 2(1), 32–37, 2007.

Kim, R. S., Chao, K. S., Moon, Y. J., Dryer, M., Lee, J., Yi, Y., Kim, K. H., Wang, H., Park, Y. D. and Kim, Y. H., An empirical model for prediction of geomagnetic storms using initially observed CME parameters at the Sun. Journal of Geophysical Research, A12108, doi:10.1029/2010JA015322, 2010.

Kivelson, G. M. and Russell, C. T., Introduction to Space Physics. Cambridge University press, UK, 1995.

Kriger, A. S., Timothy, A. F. and Roelf, E. C., A Coronal hole stream. Solar Physics, 29, 505, 1973.

Kugblenu, S., Taguchi, S. and Okuzawa, T., Prediction of the geomagnetic storm associated *Dst* index using an artificial neural network algorithm. Earth Planets Space, (51), 307–313, 1999.

Lang, K. R., The Cambridge Encyclopedia of the Sun. Cambridge University Press, 2001.

Lanzerotti, L. J., Space Weather. In Handbook of the Solar-Terrestrial Environment, eds. Kamide, Y. and Chian, A., 423–443, Springer-Verlag Berlin, Heidelberg, 2007.

Lepping, R. P., Burlaga, L. F. and Jones, J. A., Magnetic field structure of interplanetary magnetic clouds at 1 AU. Journal of Geophysical Research, 95, 11957–11965, 1990.

Lindermann, F. A., Note on the Theory of Magnetic Storms. Philosophical Magazine, 38, 669–684, 1919.

Loewe, C. A. and Prölss, G. W., Classification and mean behaviour of magnetic storms. Journal of Geophysical Research, 102(A7), 14209–14214, 1997.

Low, B. C., Solar activity and the corona. Solar physics, 167, 217–265, 1996.

Low, B. C., Coronal mass ejections, magnetic flux ropes and solar magnetism. Journal of Geophysical Research, 106, 141–163, 2001.

Lundstedt, H., AI Techniques in Geomagnetic Storm Forecasting. In Magnetic Storms, eds. Tsurutani, B. T., Gonzalez, W. D., Kamide, Y. and Arballo, J. K., vol. 98 of Gophysical Monograph series, 243–252, AGU, Washington D.C, 1997.

Lundstedt, H., Gleisner, H. and Wintoft, P., Operational forecasts of the geomagnetic Dst index. Geophysical Research Letters, 29(24), 2181, doi:10.1029/2002GL016151,2002, 2002.

Lundstedt, H., Liszka, L. and Lundin, R., Solar activity explored with new wavelet methods. Annales Geophysicae, 23, 2005.

Lundstedt, H. and Wintoft, P., Prediction of geomagnetic storms from solar wind data with the use of a neural network. Ann Geophys, 12, 19–24, 1994.

Macpherson, K. P., Conway, A. J. and Brown, J. C., Prediction of solar and geomagnetic activity data using neural networks. Journal of Geophysical Research, 100(A11), 735–744, 1995.

Manchester IV, W. B., Gombosi, T. I., Roussev, I., Ridley, A., Zeeuw, D. L. D., Sokolov, I. V., Powell, K. G. and Gábor, T., Modelling a space weather event from the Sun to the Earth: CME generation and interplanetary propagation. Journal of Geophysical Research, 109(A02107), doi:10.1029/2003JA010150, 2004.

Mayaud, P. N., Derivation, Meaning and Use of Geomagnetic Indices. Geophysicsl Monograph Series, 22, 154, doi:10.1029/GM022, 1980.

McComas, D. J., Riley, P., Gosling, J. T., Balogh, A. and Forsyth, R., Ulysses' rapid crossing of the polar coronal hole boundary. Journal of Geophysical Research, 103, 1955–1967, 1998.

McKinnell, L. A., A Neural Network based Ionospheric model for predicting the bottomside electron density profile over Grahamstown, South Africa. Ph.D. thesis, Rhodes University, Grahamstown, 2002.

McPherron, R. L., The role of substorms in the generation of magnetic storms. In Magnetic Storms, eds. Tsurutani, B. T., Gonzalez, W. D., Kamide, Y. and Arballo, J. K., vol. 98 of Gophysical Monograph series, 131–148, AGU, Washington D.C, 1997.

McPherron, R. L., Siscoe, G. and Arge, N., Probabilistic Forecasting of the 3-h *ap* index. IEEE, Transaction on Plasma Science, 32(4), 1425–1438, 2004.

Messerotti, M., Zuccarello, F., Guglielmo, S. L., Bothmer, V., Lilensten, J., Noci, G., Storini, M. and Lundstedt, H., Solar Weather Event Modelling and prediction. Space Science review, 147, 121–185, 2009.

Miesch, M. S., Large-Scale Dynamics of the Convection Zone and Tachocline. Living Reviews in Solar Physics, 2, 1,[URL http://www.livingreviews.org/lrsp–2005–1], 2005.

Mikić, Z. and Lee, M. A., An introduction to theory and models of CMEs, shocks and solar energetic particles. Space Science Reviews, 123, 57–80, 2006.

Moldwin, M., An introduction to space weather. Cambridge University Press, New York, 2008.

Ondoh, T., Science of Space Environment: Wave summit course. Ios Press Inc, 2001.

Pallocchia, G., Amata, E., Consolini, G., Marccucci, M. F. and Bertello, I., Geomagnetic *Dst* index forecast based on IMF data only. Annales Geophysicae, 24, 989–999, 2006.

Papitashvili, V. O., Papitashvili, N. E. and King, J. H., Solar cycle effects in planetary geomagnetic activity: Analysis of 36-year long OMNI dataset. Geophysical Research Letters, 27, 2797–2800, 2000.

Parker, E. N., Newtonian development of the dynamical properties of the ionized gases at low density. Physics Review, 107, 924–933, 1957.

Parker, E. N., Dynamics of the interplanetary gas and magnetic fields. Astrophysical Journal, 128, 664–676, 1958.

Perrone, L. and de Franceschi, G., A correlative study between time weigthed magnetic indices and the hih latitude ionosphere. Physics and Chemistry of the Earth, Part C: Solar, Terrestrial and Planetary Science, 24, 389–392, 1999.

Priest, E. R., Solar magneto-hydrodynamics. D. Reidel Publishing. Co., Dordrecht, 1982.

Priest, E. R., Introduction to space physics, chap. 3. Cambridge University Press, 1995.

Priest, E. R., Handbook of the Solar-Terrestrial Environment, chap. 3. Springer-Verlag, Berlin, 2007.

Prölss, G. W., Physics of the Earth's space environment: An introduction. Springer-Verlag, Berlin, 2004.

Reczko, M., Riedmiller, M., Seemann, M., Ritt, M., DeCoster, J., Biedermann, J., Danz, J., Wehrpitz, C., Werner, R., Berthold, M. and Orsier, B., External contributions to Stuttgart Neural Network Simulator (SNN), User Manual, version 4.2. Universities of Stuttgart and Tübingen, Germany and the European Particle Research Lab, CERN, Geneva, Switzerland, 1998.

Richardson, I. and Cane, H., Near-Earth Interplanetary Coronal Mass Ejections in 1996-2007. Tech. rep., Astrophysics Physics Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, 2008.

Richardson, I. G., Major geomagnetic storms ($Dst \leq -100$ nT) generated by corotating interactive regions. Journal of Geophysical Research, 111(A07S09), doi:10.1029/2005JA011476, 2006.

Richardson, I. G. and Cane, H. V., Signatures of shocks drivers in the solar wind and their dependance on the solar source location. Journal of Geophysical Research, 98(15), 15295–15304, 1993.

Richardson, I. G. and Cane, H. V., Near-Earth Interplanetary Coronal Mass Ejections During Solar Cycle 23 (1996 - 2009): Catalog and Summary of Properties. Solar Physics, 264, 189–237, 2010.

Richardson, I. G. and Cane, H. V., Geoeffectiveness (Dst and Kp) of interplanetary coronal mass ejections during 1995-2009 and the implications for storm forecasting. Space Weather, 9(SO7005), 2011.

Richardson, I. G., Cane, H. V. and Cliver, E. W., Sources of geomagnetic activity during nearly three solar cycles (1972-2000). Journal of Geophysical Research, 107(A8), doi:10.1029/2001JA000504, 2002.

Russell, C. T., Solar Wind Control of Magnetospheric Configuration. In Solar Wind - Magnetosphere Coupling, eds. Kamide, Y. and Salvin, J. A., 209–231, Terrapub, Tokyo, 1986.

Saba, M. M. F., Gonzalez, W. D. and de Gonzalez, A. L. C., Relationship between the AE, ap and Dst indices near solar minimum (1974) and at solar maximum (1979). Annales Geophysicae, 15, 1265–1270, 1997.

Schwenn, R., Large-scale structure of of the interplanetary medium. In Physics of the inner heliosphere, eds. Schwenn, R. and Marsch, E., vol. I, 99, Springer, Berlin, 1990.

Schwenn, R., Lago, A. D., Huttunen, E. and Gonzalez, W. D., The association of coronal mass ejection with their effects near the Earth. Annales Geophysicae, 23, 1033–1059, 2005.

Sheeley, N. R., Howard, R. A., Koomen, M. J., Michels, D. J., Schweenn, R., Muhlhuser, K. H. and Rosenbaur, H., Coronal mass ejections and interplanetary shocks. Annales Geophysicae, 90, 163–175, 1985.

Siscoe, G. and Schwenn, R., CME disturbance forecasting. Space Science Reviews, 123, doi:10.1007/s 11214 – 006 – 9024 –y, 2006.

Smart, D. F. and Shea, M. A., Predicting the geomagnetic ap index using real time solar wind and IMF data. In Solar Drivers of Interplanetary and Terrestrial Disturbances, eds. Balasubramaniam, K. S., Keil, S. L. and Smartt, R. N., vol. 95, Francisco: Astronomical Society of the Pacific (ASP), 1996.

Smith, E. J. and Wolfe, J. W., Observations of interaction regions and corotating shocks between one and five AU: Pioneers 10 and 11. Geophysical Research Letters, 3, 137–140, 1976.

Snyder, C. W., Neugebauer, M. and Rao, U. R., The solar wind velocity and its correlation with cosmic-ray variation and with solar and geomagnetic activity. Journal of Geophysical Research, 68, 6361–6370, 1963.

Srivastava, N., A logistic regression model for predicting the occurrence of intense geomagnetic storms. Annales Geophysicae, 23, 2969–2974, 2005.

Stix, M., The Sun, an introduction. ©Springer-Verlag, Berlin, 1989.

Stone, E. C., Frandsen, A. M., Mewaldt, R. A., Christian, E. R., Margolies, D., Ormes, J. F. and Snow, F., The Advanced Composition Explorer. Space Science Review, 86, 357–408, 1998.

Temerin, M. and Li, X., A new model for prediction of Dst on the basis of the solar wind. Journal of Geophysical Research, 107(A12), doi:10.1029/2001JA007532,2002, 2002.

Tsurutani, B. T. and Gonzalez, W. D., The Interplanetary causes of Magnetic Storms: A Review. In Magnetic Storms, eds. Tsurutani, B. T., Gonzalez, W. D., Kamide, Y. and Arballo, J. K., vol. 98 of Gophysical Monograph series, 77–89, AGU, Washington D.C, 1997.

Uwamahoro, J., McKinnell, L. A. and Cilliers, P. J., Forecasting solar cycle 24 using neural networks. Journal of Atmospheric and Solar-Terrestrial Physics, 71, 569–574, 2009.

Valach, F., Revallo, M., Bochnicek, J. and Hejda, P., Solar energetic particle flux enhancement as a predictor of geomagnetic activity in a neural network-based model. Space Weather, 7(S04004), doi:10.1029/2008SW000421, 2009.

Vršnak, B., Processes and mechanisms governing the initiation and propagation of CMEs. Annales Geophysicae, 26, 3089–3101, 2008.

Wallace, H. C., Introduction to geomagnetic fields. Cambridge University Press, UK, 1997.

Wallace, H. C., Introduction to Geomagnetic Fields. Cambridge University Press, Cambridge, UK, 2 edn., 2003.

Wang, Y. M., Ye, P. Z., Wang, S., Zhou, G. P. and Wang, J. X., A statistical study on the geoeffectiveness of the Earth-directed coronal mass ejections from March 1997 to December 2000. Journal of geophysical research, 107(A11), doi:10.1029/2002JA009244, 2002.

Webb, D. F., Coronal Mass Ejections: Origins, Evolution, and Role in Space Weather. IEEE Transactions on plasma science, 28(6), 1795–1806, 2000.

Webb, D. F., Cliver, E. W., Crooker, N. U., st Cyr, O. C. and Thompson, B. J., Relationship of halo coronal mass ejections, magnetic clouds and magnetic storms. Journal of Geophysical Research, 105(A4), 7491–7508, 2000.

Wik, M., Pirjola, R., Lundstedt, H., Viljanen, A., Wintoft, P. and Pulkkinen, A., Space weather events in July 1982 and October 2003 and the effects of geomagnetically induced currents on Swedish technical systems. Annales Geophysicae, 27, 1775–1787, 2009.

Wing, S., Johnson, J. R., Jen, J., Meng, C.-I., Sibeck, D. G., Bechtold, K., Freeman, J., Costello, K., Balikhin, M. and Takahashi, K., Kp forecast models. Journal of Geophysical Research, 110(A04203), doi:10.1029/2004JA010500, 2005.

Wood, B. E. and Howard, R. A., An Empirical Reconstruction of the 2008 April 26 Coronal Mass Ejection. Astrophysical Journal, 702, 901–910, 2009.

Woolley, J. W., Agarwarl, P. K. and Baker, J., Modeling and Prediction of Chaotic Systems with artificial neural networks. International Journal for Numerical Methods in Fluids, 63, doi:10.1002/fld.2117, 2010.

Wu, J. G. and Lundstedt, H., Prediction of geomagnetic storms from solar wind data using Elman recurrent networks. Geophysical Research Letters, 23, 319–322, 1996.

Wu, J. G. and Lundstedt, H., Neural network modeling of solar wind-magnetosphere interaction. Journal of Geophysical Research, 102(A7), 457–466, 1997.

Xie, H., Gopalswamy, N., Manoharan, P. K., Lara, A., Yashiro, S. and Lepri, S., Long-lived geomagnetic storms and coronal mass ejections. Journal of Geophysical Research, 111(A01103), 10.1029/2005JA0011287, 2006.

Yashiro, S., Gopalswamy, N., Michalek, G., Cyr, O. C. S., Plunkett, S. P., Rich, N. B. and Howard, R. A., A catalog of white light coronal mass ejections observed by the SOHO spacecraft. Journal of Geophysical Research, 109(A07105), doi:10.1029/2003JA010282, 2004.

Yermolaev, I. Y., Yermolaev, M. Y., Zastenker, G. N., Zelenyi, L. M., Petrucovich, A. and Sauvaud, J. A., Statistical studies of geomagnetic storm dependencies on solar and and interplanetary events: a review. Planetary Space Science, 53, 189–196, 2005.

Yermolaev, Y. I. and Yermolaev, M. Y., Statistical study on geomagnetic storm effectiveness of solar and interplanetary events. Advances in Space Research, 37(6), 1175–1181, 2006.

Zell, A., Mamier, G. M., Vogt, M., Mach, N., Hübner, R., Döring, S., Herrmann, K. U., Soyez, T., Schmalzl, M., Sommer, T., Hatzigeorgiou, A., Posselt, D., Schreiner, T., Kett, B., Clemente, G., Wieland, J. and Gatter, J., Stuttgart Neural Network Simulator (SNNS), User Manual, version 4.2. Universities of Stuttgart and Tübingen, Germany and the European Particle Research Lab, CERN, Geneva, Switzerland, 1998.

Zhang, J., Richardson, I. G., Webb, D. F., Gopalswamy, N., Huttunen, E., Kasper, J. C., Nitta, N. V., Poomvises, W., Thompson, B. J., Wu, C. C., Yashiro, S. and Zhukov, A. N., Solar and interplanetary sources of major geomagnetic storms ($Dst \leq -50$ nT) during 1996-2005. Journal of Geophysical Research, 112(A10102), doi:10.1029/2007JA012321, 2007.