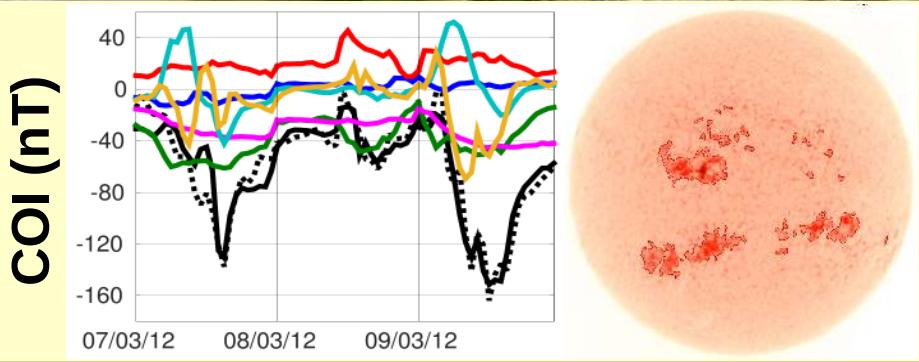


YVELICE SORAYA CASTILLO ROSALES

GLOBAL AND MID-LATITUDE IMPACT OF TEN YEARS OF SOLAR ACTIVITY USING OBSERVATIONAL DATA OF THE GEOPHYSICAL AND ASTRONOMICAL OBSERVATORY OF THE UNIVERSITY OF COIMBRA

Ph.D. thesis in Physics, specialization in Astrophysics,
supervised by Prof. Dra. Maria Alexandra Pais and
Prof. Dr. João Manuel Fernandes and presented to the
Physics Department of the Faculty of Science and Technology
of the University of Coimbra

November 2017

· U C ·

UNIVERSIDADE DE COIMBRA

COI Observatory image was adapted from $\mathrm{GoogleMaps}^{©}$. Sun with facular areas was provided by Teresa Barata, from Geophysical and Astronomical Observatory of University of Coimbra (OGAUC) Website. The COI data is courtesy of Paulo Ribeiro, from OGAUC. Tsyganenko and Sitnov 2005 currents series are part of this thesis results.

GLOBAL AND MID-LATITUDE IMPACT OF TEN YEARS OF SOLAR ACTIVITY USING OBSERVATIONAL DATA OF THE GEOPHYSICAL AND ASTRONOMICAL OBSERVATORY OF THE UNIVERSITY OF COIMBRA

Author:

YVELICE SORAYA CASTILLO ROSALES

Supervisors:

Prof. Dr. Maria Alexandra Albuquerque Faria Pais

Prof. Dr. João Manuel de Morais Barros Fernandes

A thesis submitted in fulfilment of the requirements for the

Ph. D. in Physics, Specialization in Astrophysics,

in the Department of Physics

Faculty of Science and Technology of the University of Coimbra

November, 2017

Acknowledgements

First of all, I thank God and His Holy Mother, the foundation of my life, secondly my family, specially my lovely parents Efigenia Rosales and Otoniel Castillo (RIP).

I am grateful to my thesis supervisors, Alexandra Pais and João Fernandes, two dear wonderful people, who taught me a lot in academic, research, technical and life matters. I am also very grateful to all members of the Centre for Earth and Space Research of the University of Coimbra (CITEUC), for their great friendship and support. A special thanks to their directors, João Fernandes and Teresa Barata, for their support to present partial results of the thesis in 2016 at the European Geophysical Union General Assembly in Vienna and in the Science Encounter 2016 in Lisbon. CITEUC is funded by National Funds through FCT - Foundation for Science and Technology (project: UID/Multi/00611/2013) and FEDER - European Regional Development Fund through COMPETE 2020 – Operational Programme Competitiveness and Internationalization (project: POCI-01-0145-FEDER-006922). I would like to thank the Sun-Earth interactions research group, integrated by my thesis advisers and also Anna Morozova, Fernando Pinheiro, Paulo Ribeiro, Teresa Barata and Sara Carvalho, for their great patience, support, feedback and friendship in all stages of the Ph. D. work.

I also acknowledge the Erasmus Mundus Action 2 Consortium AMIDILA, Lot 15 – strand 1, 2013-2588/001-001-EM Action 2 Partnerships, specially Bologna University Consortium Team and International Relations Office of UC (Rita Maia and Silvia Silva) for the scholarship granted to me and their exceptional support. I want to acknowledge Prof. Maria Cristina Pineda de Carias and Master Maria Quiroz, Dean of the Faculty of Space Sciences of the National Autonomous University of Honduras and Head of the Department of Astronomy and Astrophysics of that Faculty respectively, to UNAH's Vice-Rectory of Academic Affairs and Vice-Rectory of International Relations for their support from Honduras to ensure that this work could be completed.

I also thank people of the Department of Physics, the University and the City of Coimbra, most specially to the closest ones that became my friends and were encouraging and accompanying me during this years. Finally, a special thanks to Saint Josemaria Escrivá de Balaguer, Opus Dei Founder, Blessed Alvaro del Portillo, Monsignor Javier Echevarría († 2016) and Monsignor Fernando Ocáriz, Opus Dei Prelates.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Cofinanciado por:

COMPETE 2020

PORTUGAL 2020

UNIÃO EUROPEIA
Fundo Europeu
de Desenvolvimento Regional

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Sumário

Esta tese estuda os processos físicos relacionados com a atividade geomagnética detetada em observatórios magnéticos à superfície da Terra em locais de latitudes intermédias, ou seja, afastados do pólos e do equador, bem como a sua relação com a atividade solar e o meio interplanetário durante o período temporal 2007-2016.

Para isto, numa primeira parte do trabalho são comparadas séries temporais simuladas, obtidas através do modelo semi-empírico de Tsyganenko and Sitnov 2005 (TS05), com as séries de dados obtidas em quatro observatórios (de latitude geomagnética próxima e espaçados em longitude) a saber: Coimbra (COI), Portugal; Panagyurishte (PAG), Bulgaria; Novosibirsk (NVS), Rússia e Boulder (BOU), USA. São ainda incluídos no estudo dados observacionais de instrumentos em Terra e no Espaço compilados na base de dados OMNI/NASA. Este estudo permite, assim, testar o modelo TS05 no que concerne a modelação das observações geomagnéticas à superfície da Terra. Os resultados apontam para que o modelo TS05 reproduz bem a componente geomagnética Norte-Sul durante os dias ativos mas é menos eficiente nos dias calmos. Por outro lado, este estudo mostra que o modelo TS05 não reproduz bem a componente Este-Oeste sendo que a causa para isso deverá ser o facto do modelo forçar o fecho das correntes Birkeland no centro da Terra e não na ionosfera. Mostra-se ainda que a componente geomagnética Norte-Sul, durante os dias ativos, é, essencialmente, influenciada pelas correntes de cauda (TAIL), a corrente de anel simétrico (SRC) e a corrente de anel parcial (PRC). Em contrapartida, a componente Este-Oeste é determinada pelas correntes FAC e PRC. As correlações entre os registos e o modelo TS05 melhoram consideravelmente nos dias geomagneticamente ativos, para qualquer um dos observatórios. Este resultado pode explicar-se pelo facto da variação magnética diurna (QD, de 'quiet daily') ter uma contribuição principal das correntes ionosféricas e estas não serem incluídas no modelo TS05. No caso dos observatórios localizados próximo do centro do vórtice de correntes ionosféricas (a saber COI e PAG), foi possível separar

com sucesso a contribuição QD. Assim, naturalmente, após a remoção da variação QD dos dados do modelo TS05, a correlação entre o modelo e as séries de COI e PAG melhora em relação ao que ocorre em BOU e NVS. Em contrapartida os registos destes dois observatórios aparecem bem correlacionados com os índices geomagnéticos Dst ('disturbance storm-time') e RC ('ring current').

Posteriormente foi feito um estudo correlacionando 33 parâmetros representativos da interação Sol-Terra no período 2009 a 2016: parâmetros solares, parâmetros do campo magnético interplanetário (IMF) e índices geomagnéticos (GAI). Mostra-se que as correlações são melhores usando médias sobre o período de rotação solar (27 dias) do que usando médias diurnas, o que sugere a necessidade de tomar em conta atrasos de propagação e efeitos cumulativos no estudo da interação Sol-Terra. A partir das médias de 27 dias, mostra-se que as correlações globais (a 33 parâmetros) são melhores para os seguintes parâmetros: áreas das regiões faculares, T-SRC e T-PRC (onde "T" indica que a grandeza foi determinada pelo modelo TS05 - índices TI), o módulo do vector do campo IMF, a percentagem sul de IMF (B_{ZS} GSM) e a função de Newell, de acoplamento vento solar/magnetosfera.

Por outro lado, foram calculadas as assimetrias helio-magnéticas decorrentes da diferença entre os valores médios dos parâmetros durante o trânsito através dos setores de polaridade solar positiva e negativa do meio interplanetário. Da análise das referidas assimetrias constata-se que a Terra esteve, durante o ciclo 24, mais tempo a norte do que a sul do equador solar. Além disso, as assimetrias põem em evidência uma oscilação anual presente nas séries de B_{ZS} GSM, B_Z GSM e nos índices GAI e explicada em princípio pelo efeito de Russell-McPherron. Nota-se ainda uma variação anual dos índices TI durante a fase descendente do ciclo solar.

Em conclusão, a maior parte do trabalho desta tese foi ocupada com a implementação de uma série de testes estatísticos aplicados ao estudo do desempenho do modelo TS05 na simulação da atividade geomagnética observada a latitudes intermédias do Hemisfério Norte. O esquema de testes aqui proposto pode no futuro ser aplicado

a qualquer outro modelo da magnetosfera. No final deste trabalho, realizou-se ainda uma pesquisa envolvendo diferentes parâmetros que caraterizam a superfície do Sol, o meio interplanetário e a atividade geomagnética, com o intuito de identificar os mais adequados a ser utilizados para relacionar fenómenos à superfície do Sol com a atividade geomagnética observada na Terra.

Palavras-chave: atividade solar, dados solares e geomagnéticos, interacção Sol-Terra, meteorologia espacial.

Abstract

The focus of this thesis is to identify the role and dynamics of different current systems for storm-time activity at mid-latitude ground level and their relation with other solar, interplanetary and geomagnetic parameters, in the 2007-2016 time interval.

For this purpose, I compared the synthetic series of six magnetospheric current systems computed with the data-based semi-empirical model of Tsyganenko and Sitnov 2005 (TS05), with the measurements of four mid-latitude geomagnetic stations at very near geomagnetic latitudes, but well apart in longitude (Coimbra (COI), Portugal; Panagyurishte (PAG), Bulgary; Novosibirsk (NVS), Russia; Boulder (BOU), USA), and with other ground and satellite-based solar, interplanetary and geomagnetic parameters obtained from the OMNI/NASA database. An evaluation of TS05 model is presented, in order to determine the model ability to reproduce both the total magnetospheric transient signal and to explain this signal through the contribution of each TS05 magnetospheric current system at ground level, comparing them with hourly data of the four geomagnetic stations selected.

It was found that TS05 model is a useful tool to explain ground-based North-South (or X) component of geomagnetic activity at mid-latitudes, in terms of main current sources. It was verified that TS05 is efficient to reproduce the X component of terrestrial magnetospheric field at mid-latitudes during high geomagnetic activity time, with correlations $r \geq 0.7$ in $\sim$50% of compared data, presenting a lower efficiency during calm time, with correlations $r \geq 0.7$ only in $\sim$30% of data compared. Results are less favourable for the East-West (or Y) component, probably due to the fact that TS05 model closes Birkeland (or field-aligned, FAC) currents through the Earth's centre instead of through the ionosphere. It was found that currents that contribute most to the X component during geomagnetic active periods are the cross-tail (TAIL), the symmetric ring (SRC) and the partial ring (PRC) currents. The currents that contribute

most to Y component are FAC and PRC currents. For all stations the highest correlations among observations and TS05 simulations are obtained for stronger geomagnetic activity. The results in this study indicate that the implementation of TAIL and SRC currents in TS05 model is more successful than that for the FAC current.

The quiet daily (QD) variation has a main contribution from ionospheric currents, which are not considered in TS05 model. For observatories at Northern Hemisphere's mid-latitudes that are localized close to the ionospheric current vortex center (COI and PAG), it was possible to separate efficiently the QD ionospheric contribution using Principal Component Analysis (PCA). For the other two stations (NVS and BOU) it was found a relatively higher contribution of magnetospheric signal in the QD variation. After removing the QD variation from data, COI and PAG are better correlated with TS05 series than BOU and NVS. However, BOU and NVS are better correlated with geomagnetic indices Dst (disturbance storm-time) and RC (ring current), with RC showing a slightly less good performance with respect to Dst.

Correlations between 33 solar, interplanetary magnetic field (IMF) and geomagnetic activity proxies were analysed for the 2009-2016 time interval. It was found that series of 27-day averages (Bartels' rotation) give higher correlations than daily or annual series. Parameters that show higher cross-correlations among different groups are the Sun's northern and southern facular areas (FA-N and FA-S), two geomagnetic indices derived from TS05 model (T-SRC and T-PRC), the total IMF intensity (B), the percentage of IMF southward component (B_{ZS} GSM) and the interplanetary coupling Newell's function. We propose that these parameters are the best candidates to use if we want to relate meaningfully the solar surface events to geomagnetic activity felt on the Earth's surface. Two new proxies were tested, 1) TI-indices, calculated from the X TS05-derived series of TAIL, SRC, PRC and FAC contributions for the four observatories and 2) B_{ZS} GSM, calculated as the daily percentage of IMF southward component along the GSM Z-axis.

Helio-magnetic asymmetries were calculated for the 33 parameters, as the difference between their averaged values in the towards and away magnetic sectors of the interplanetary medium. Improvement in 27-day correlations with respect to annual correlations is the result of an annual oscillation in this asymmetry, which is present in most studied proxies and is probably due to the Russell-McPherron effect. Due to this effect, B_{ZS} GSM and B_Z GSM have a well-defined annual modulation, and geomagnetic activity indices (GAI) also have annual oscillation and good correlations with B_{ZS} GSM and B_Z GSM. TI-indices have annual oscillation at declining phase of the cycle, but insignificant oscillation near the minimum. Major percentage of towards days in negative polarity epoch and of away days in positive polarity epoch means that the Earth has been mostly at the northern magnetic hemisphere during the solar cycle 24.

In conclusion, the main part of this Ph.D. thesis was dedicated to the design and implementation of a statistical approach that was applied to test the performance of the TS05 model in explaining geomagnetic activity observed at Earth's Northern Hemisphere mid-latitudes. This approach can be applied to test any other magnetospheric model. At the end of this work, a prospective study was made using different proxies that describe the Sun surface, the interplanetary medium and geomagnetic activity, to identify those parameters that should be more meaningfully used to relate the Sun to the geomagnetic activity observed on Earth.

Keywords: Sun-Earth interaction, solar activity, solar and geomagnetic data, space weather.

List of Abbreviations and Acronyms

AE	Auroral Electrojet index
AL	Auroral Electrojet Lower values
Ap	Planetary a-index
ASYM-D	Longitudinally Asymmetric Disturbance index for the East-West component
ASYM-H	Longitudinally Asymmetric Disturbance index for the Horizontal component
AU	Auroral Electrojet Upper values
BOU	Boulder Magnetic Station (USA)
CD	Earth-Centered Dipolar Coordinate System
CF	Magnetosphere's Chapman-Ferraro or magnetopause currents
CH	Coronal Hole
CIR	Co-rotating Interaction Regions
CME	Coronal Mass Ejection
COI	Coimbra Magnetic Station (Portugal)
DGRF	Definite Geomagnetic Reference Field
Dst	Disturbance storm-time index
F10	F 10.7 cm solar flux
FA-N	Facular area at Sun's northern hemisphere
FA-NS	Excess of northern over southern facular area (FA-N - FA-S)
FA-S	Facular area at Sun's southern hemisphere
FA-T	Total facular area
FAC	Magnetosphere's field-aligned or Birkeland currents
GAI	Geomagnetic activity indices
GEO	Geographic Coordinate System
GEOPACK-2008	Tsyganenko subroutines for Main Field Calculation
GFZ-Postdam	Helmholtz Centre Potsdam German Research Centre for Geosciences
GOES	Geosynchronous Orbit Earth Spacecrafts
GSE	Geocentric Solar Ecliptic Coordinate System
GSEQ	Geocentric Equatorial Coordinate System
GSM	Geocentric Solar Magnetic Coordinate System
HCS	Heliospheric Current Sheet
HMF	Heliospheric Magnetic Field
HRO OMNI	High resolution OMNI data
IGRF	International Geomagnetic Reference Field
IMF	interplanetary magnetic field
INT	Magnetosphere-interplanetary magnetic field interconnection current
K	Local K geomagnetic index
Kp	Planetary K index

LRO OMNI	Low resolution OMNI data
LT	Local Time
MAG	Magnetospheric Coordinate System
NLL	Newell's coupling function
NOAA	National Oceanic and Atmospheric Administration (USA)
NVS	Novosibirsk Magnetic Station (Russia)
OMNI	Goddard Space Flight Center database (NASA)
PAG	Panagyrishte Magnetic Station (Bulgary)
PC	Polar Cap index
PCN	Polar Cap North index
PCS	Polar Cap South index
PRC	Magnetosphere's partial ring current
QD	Quiet-daily variation
R1	Region 1 of Birkeland currents
R2	Region 2 of Birkeland currents
RC	Ring Current index
SCC	Spearman's cross-correlations
SM	Solar Magnetic Coordinate System
SN-N	Sunspot number at Sun's northern hemisphere
SN-NS	Excess of northern over southern sunspot number (SN-N - SN-S)
SN-S	Sunspot number at Sun's southern hemisphere
SN-T	Total sunspot number
SP	Solar parameters
Sq	Solar quiet-daily variation
SRC	Magnetosphere's axi-symmetric ring current
SSC	Storm sudden commencement
SV	Secular variation
SW	Solar wind
SWE	Space Weather
SWP	Solar wind parameters
SWPC	Space Weather Prediction Center at NOAA (USA)
SYM-D	Longitudinally Symmetric Disturbance index for the East-West component
SYM-H	Longitudinally Symmetric Disturbance index for the Horizontal component
T-FAC	T-index for FAC current
T-PRC	T-index for PRC current
T-SRC	T-index for SRC current
T-TAIL	T-index for TAIL current
TAIL	Magnetosphere's cross-tail current sheet
TI	Tsyganenko indices
TS	Tsyganenko models
TS04c	Tsyganenko and Sitnov 2005 Fortran code
TS05	Tsyganenko and Sitnov 2005 magnetosphere's model
UT	Universal Time

List of Symbols

A	Ampere
IMF **B**	Total interplanetary magnetic field vector
IMF B	IMF total magnetic field intensity
BH	TS05 synthetic local H component
BX	TS05 synthetic local X component
B_X	X component of the interplanetary magnetic field vector
$\widetilde{BX}$	TS05 synthetic local X component without standard quiet-daily variation
BY	TS05 synthetic local Y component
B_Y	Y component of the interplanetary magnetic field vector
$\widetilde{BY}$	TS05 synthetic local Y component without standard quiet-daily variation
BZ	TS05 synthetic local Z component
B_Z	Z component of the interplanetary magnetic field vector
$\widetilde{BZ}$	TS05 synthetic local Z component without standard quiet-daily variation
D	Magnetic declination
$d\phi_{MP}/dt$	Rate of magnetic flux removed from magnetopause (Newell)
ΔQX, ΔQY, ΔQZ	Quiet daily variation of X, Y and Z data series, respectively
E_{KL}	Kan and Lee interplanetary electric field
F	Total local magnetic field intensity
ϕ	Geocentric Longitude
ϕ'	Geomagnetic Longitude
H	Local horizontal geomagnetic component
$\widehat{H}$	**H** without PCA quiet-daily variation
$\widetilde{H}$	**H** without standard quiet-daily variation
I	Magnetic Inclination
K	Kelvin degrees
km/s	Kilometers per second
λ	Geocentric Latitude
m	Meters
m	Earth's dipole moment
μ_0	Permeability of free space
N	Number of particles
$N\ cm^{-3}$	Particles per cubic centimetre
nT	Nanotesla
$\bigtriangledown \cdot$	Divergence vector operator
$\bigtriangledown$	Gradient vector operator
p	Solar wind ram pressure
ψ	Dipole tilt angle
ρ	Solar wind proton density
r	Pearson's correlation coefficient

R	Radial distance
r_s	Spearman's correlation coefficient
R_E	Mean Earth's radius = 6,371.2 km
$s.f.u.$	1 solar flux unit = 1 $W\ m^{-2}$
T	Solar wind proton temperature
θ	Geocentric Colatitude
V	Solar wind bulk speed
V_X	X component of the solar wind
V_Y	Y component of the solar wind
V_Z	Z component of the solar wind
$\mathbf{V}$	Total solar wind velocity vector
X	Local North-South geomagnetic component
$\widetilde{X}$	X without standard quiet-daily variation
X_{GEO}	X component in the Geodetic Coordinate System
Y	Local East-West geomagnetic component
$\widetilde{Y}$	Y without standard quiet-daily variation
Y_{GEO}	Y component in the Geodetic Coordinate System
Z	Local vertical geomagnetic component
$\widetilde{Z}$	Z without standard quiet-daily variation
Z_{GEO}	Z component in the Geodetic Coordinate System

Contents

1 Introduction

1.1 Motivation

Space Weather (SWE) is about the electromagnetic perturbations and energetic particle events driven by changes in the Sun's magnetic field and solar wind (SW) and by their effects on Earth's magnetic field and upper atmosphere, that affects the Earth's space, atmosphere and surface environments (Thomson, 2012; Schrijver, 2015). SWE affects our modern way of life, some times in very dramatic ways, and its potential to impact is growing. It can disrupt GNSS signals in all kind of industrial, medical, engineering and common day instruments and applications, can cause induced currents and corrosion in pipelines and electrical transmission lines, satellite failures, high radiation doses at high altitudes, specially near geomagnetic poles and during space missions, disruption in high frequency airline communication (HF) and navigation systems, in long-distance radio signals, problems in hydrocarbon production, and more (Schrijver, 2015).

One of the most dramatic examples of SWE implications is the effect of geomagnetically induced currents (GIC) in high voltage transformers or generators. During very high geomagnetic activity GICs can flow between the high voltage lines and the earth through windings, saturating the transformers or generators, perhaps damaging them or triggering safety systems that remove them from the circuit, possibly creating knock-on effects in the rest of the grid. A forecast from 30 min up to three days ahead, based on solar wind observations, allows to take precautions when a major

geomagnetic event occurs. The severity of the effects on almost all technologies depends in part on the geomagnetic latitude (at higher and lower latitudes, higher risks), and on the underlying and surrounding land or sea conductivity (Thomson, 2012). A deeper understanding of space weather can be used to forecast and to determine design parameters to build transformers and generators, to avoid oil pipeline corrosion produced by GICs, to design oil industry drillers, power supply, satellites, GPS, GNSS, and any other small electric component in general. Interest in, and dependence on, space weather information and services grows rapidly (Schrijver, 2015), due to its importance for technology, infrastructure, global economy and to avoid hazards. Geomagnetism science and technical expertise is in demand from industry, education, science and society (Thomson, 2012).

The ultimate goal of this thesis is to contribute to reduce SWE nuisance effects at the Northern Hemisphere's mid-latitudes. To this end, I will track correlations among Sun activity parameters, the solar wind blowing from the Sun and reaching the Earth and geomagnetic activity indices. I will also look into different current systems responsible for geomagnetic activity.

Solar surface and interplanetary magnetic field observations have been made easily available to all, in order to reach this end. SW data from high altitude satellites are also available through open access sites.

Magnetic observatories provide local measurements of space weather conditions and free data centres provide near-real-time magnetic data from many observatories (i.e., INTERMAGNET, World Data Centres for Geomagnetism of Kyoto and Edinburgh), valuable for analysis of global and regional space weather activity. Different products offered by them are helpful to use with models that simulate or predict impact on the environment and technologies (Thomson, 2012). In this thesis, I will concentrate on four mid-latitude magnetic stations, including Coimbra station, separating contributions for the North-South (X) and for the East-West (Y) local geomagnetic components in the 2007-2016 time interval.

Magnetosphere's models as the Tsyganenko and Sitnov 2005 model (TS05) are widely used to determine which magnetospheric current systems contribute more to high geomagnetic activity. This model is made available for the scientific community and is supported by a large quantity of detailed information.

1.2 State of the Art

Geomagnetic activity observed at the Earth's surface is the result of the interaction between cosmic rays from distant sources or charged particles and electromagnetic radiation from the Sun, with the Earth's electromagnetic environment from the ground up to distances of 10 to 20 Earth radii (e.g., Campbell, 2003). This region is called magnetosphere. Enhanced electric fields, currents, and energetic particle precipitation energize the magnetosphere and are responsible for a number of perturbations in technological infrastructures that became vital for human activities. These effects are monitored by space weather services all over the world (Schrijver, 2015) and are most prominent at high and equatorial latitudes (e.g., Buonsanto, 1999). Nevertheless, mid-latitudes are also significantly perturbed by geomagnetic storms and adverse effects include loss of HF communications, damage in high-voltage power supply lines or pipelines and deterioration of services provided by global navigation satellite systems (e.g., Beggan et al., 2013; Buonsanto, 1999). Furthermore, Northern Hemisphere's mid-latitudes concentrate a large amount of radiation and electrostatic-sensitive technology and new studies on the geomagnetic activity in this Earth region seem justified.

The geomagnetic field measured at the ground surface is the sum of contributions from very different sources: the Earth's core field (main field), the field of magnetized lithosphere (crustal field), the primary fields of magnetospheric and ionospheric current sources and their secondary contributions due to Faraday induction in the electrical conducting crust and mantle (e.g., Hulot et al., 2010). The main field, due to a dynamo powered by convection in the liquid core of the Earth, is represented by

spherical harmonic (SH) models up to degree 10 to 13, computed from ground-based and low-orbit satellite geomagnetic data. As an example, the latest International Geomagnetic Reference Field model, IGRF-12, provides sets of definitive SH coefficients at 5-years interval from 1900.0 to 2010.0 and a non-definitive model for 2015.0 (Thebault et al., 2015). The crustal field, represented by higher degree SH coefficients, is responsible for biases in different geomagnetic field components at a given observatory, when compared with predictions from, e.g., IGRF models (e.g., Sabaka et al., 2015). Both main and crustal magnetization fields contribute to the baseline at each station and can be removed from data series in studies of sub-annual geomagnetic variations, by fitting low-degree polynomials to annual means computed from the quietest days in each month. Because they remain constant at the sub-daily time scales, they do not contribute directly to geomagnetic activity. However, the geometry and amplitude of magnetospheric and ionospheric current systems is strongly constrained by the geometry and amplitude of the main field (e.g., Pedatella et al., 2011) and models as IGRF-12 are included into magnetospheric models (e.g., Tsyganenko and Sitnov, 2005). All the three remaining components of the near-Earth field, namely magnetospheric, ionospheric and induction fields, contribute to the geomagnetic activity observed at the ground level.

As a means to monitor and even forecast the geomagnetic activity, indices computed from observations of geomagnetic field sub-daily variations at the Earth's surface (e.g., Mayaud, 1980) have been used. At mid-latitudes, the indices Kp (planetary K), Dst (disturbance storm-time) (Sugiura and Kamei, 1991) and more recently RC (Olsen et al., 2014) are the most utilized. The 3-hour period planetary Kp index characterizes different levels of global geomagnetic activity in broad terms, ranging from 0 to 9 (e.g. Mayaud, 1980). It is related quasi-logarithmically to the geomagnetic amplitude measured in the most disturbed horizontal magnetic field component, at a group of 13 given stations. Kp values of 4 have been used to separate calm from storm time periods (e.g. McCollough et al., 2008). For a description of geomagnetic activity at a given

site, local K index values are computed from the data of the magnetic observatory at that site. For a shorter time scale description of geomagnetic activity during geomagnetic storms, Dst index is used since it is provided at an 1-hour time resolution and depends linearly on the perturbation amplitude. Dst values are derived from hourly values of the horizontal geomagnetic component (H) obtained at four given magnetic observatories distributed evenly in longitude (e.g., Sugiura, 1964; Mayaud, 1980).

Geomagnetic indices represent the integrated effect of different mechanisms that contribute for the geomagnetic activity observed close to the Earth's surface. They include contributions from the ring current located at the geomagnetic equatorial plane at distances of 1.2 - 10 Earth radii (R_E) (Tsyganenko, 2014), the field-aligned currents that follow approximately the geomagnetic field lines at high latitudes (e.g. Kallenrode, 2004), the cross-tail current that is also located in the geomagnetic equator, flowing from dawn to dusk, connecting the northern and southern currents of the magnetopause, forming a θ-like current system with them, as seen from along the Earth-Sun direction (e.g. Baumjohann and Nakamura, 2007), and the Chapman-Ferraro current that flows along the magnetopause, i.e., the magnetosphere's boundary, due to penetration of solar wind's protons, ions and electrons (e.g. Kallenrode, 2004).

Magnetospheric models take all these current systems into account and model their dynamics, (e.g., Tsyganenko, 2002a; Tsyganenko, 2002b). These models use information on the energetic particles blown from the Sun (the solar wind) and arriving at the magnetopause (e.g. Tsyganenko and Sitnov, 2005; Alexeev and Feldstein, 2001; Tsyganenko and Andreeva, 2015). They represent the different mechanisms that relate the solar wind with the geomagnetic activity using physical or empirical equations. The Tsyganenko and Sitnov model (TS05) is one such model (e.g. Tsyganenko and Sitnov, 2005) that uses satellite data and empirical equations to compute the magnetic field vectors associated to the currents mentioned before, plus a field vector that represents the penetration of the interplanetary magnetic field (IMF) into the magnetosphere (e.g. Tsyganenko and Sitnov, 2005).

The main driver of geomagnetic activity is the Sun. The possibility to relate solar surface features with the solar wind arriving to the external magnetosphere and with indices of geomagnetic activity is also of particular interest in the framework of this thesis.

It is a well-known fact that structures at the Sun's surface, chromosphere and corona are asymmetric among northern and southern heliographic hemispheres, for example sunspot numbers and sunspot areas (e.g. Temmer et al., 2006; Ballester, Oliver, and Carbonell, 2005; Chang, 2009; Chowdhury, Choudhary, and Gosain, 2013; Li et al., 2002; Carbonell, Oliver, and Ballester, 1993; Oliver and Ballester, 1994; Vizoso and Ballester, 1990; Ravindra and Javaraiah, 2015), facular areas (e.g. Goncalves et al., 2014), solar flares and coronal mass ejections (CMEs) (e.g. Bankoti et al., 2010; Baranyi and Ludmany, 2006; Zhang and Feng, 2015; Verma, 2000; Verma, 1993), solar surface rotation (e.g. Zhang, Mursula, and Usoskin, 2015), photospheric magnetic fields (e.g. Virtanen and Mursula, 2014; Song, Wang, and Ma, 2005). The asymmetries between the northern and southern solar magnetic hemispheres, can be seen on solar structures, interplanetary and geomagnetic parameters, i.e., sunspot numbers (e.g. El-Borie et al., 2012), solar wind (e.g. Mursula and Zieger, 2001; Nair and Nayar, 2008; Nair and Nayar, 2009; Borie, Abdel-Halim, and El-Monier, 2016; Svirzhevsky et al., 2005), total interplanetary magnetic field intensity (e.g. El-Borie et al., 2012; Borie, Abdel-Halim, and El-Monier, 2016; El-Borie, Abdel-Halim, and El-Monier, 2016), proton density (e.g. Borie, Abdel-Halim, and El-Monier, 2016; Svirzhevsky et al., 2005), proton temperature (e.g. Borie, Abdel-Halim, and El-Monier, 2016), solar radio flux (e.g. El-Borie et al., 2012), Kp, Dst and other geomagnetic indices (e.g. El-Borie et al., 2012; El-Borie, Abdel-Halim, and El-Monier, 2016; Apostolov, Altadill, and Todorova, 2004), cosmic rays (e.g. El-Borie et al., 2016; Ngobeni and Potgieter, 2011; Svirzhevsky et al., 2005). Many authors address also to different periodic variations of some of these parameters, i.e., 11-year and 22-year solar cycle, 27-days variations, 13.5-days variations, 9-years and so on (e.g. Hathaway, 2010; Mursula, Usoskin, and Kovaltsov, 2002; Svalgaard, 2011;

Alania et al., 2001; Cliver, Kamide, and Ling, 2002). I selected daily and 27-day (one Bartels' rotation, approximately one rotation of the Sun's equator as seen from Earth) averaged values to analyse the correlations among series of solar, interplanetary and geomagnetic parameters (e.g. Apostolov, Altadill, and Todorova, 2004; Svirzhevsky et al., 2005) and also used annual (e.g. El-Borie et al., 2012; Borie, Abdel-Halim, and El-Monier, 2016; El-Borie, Abdel-Halim, and El-Monier, 2016; Nair and Nayar, 2008; Nair and Nayar, 2009; Mursula and Zieger, 2001) and 27-day series to analyse correlations among their respective helio-magnetic asymmetries. The helio-magnetic asymmetries reveal if a parameter has or not any dependency on the configuration of the interplanetary magnetic field (IMF) along the solar cycles. The cross-correlations among asymmetries show their relations relative to a specific IMF configuration.

1.3 The Geophysical and Astronomical Observatory of the University of Coimbra (OGAUC)

The OGAUC is an organic unit of the Faculty of Sciences and Technology of the University of Coimbra. It was created in 2013, after the merge of two secular institutions, the Astronomical Observatory (created in 1772) and the Geophysical Institute (created in 1864). The main research topics of the Observatory are: Earth Sciences, Planetology and Solar Physics. This research is supported by long time series of geophysical and astronomical daily observational data, obtained in situ. OGAUC involves around 30 people (researchers, students and technical staff). More information can be found in `/http://geofisico.dyndns.org/`.

Data of the Geophysical and Astronomical Observatory of the University of Coimbra (OGAUC) were employed in this work, i.e., facular areas and North-South (X) and East-West (Y) hourly geomagnetic measurements at COI station.

Facular areas were obtained from OGAUC daily espectroheliograms (i.e., monocromatic images sensible to different thermodynamic parameters of solar atmosphere), obtained in CaII K3 (λ = 3933.7 Å). The BASS2000 database of Paris Observatory in Meudon (/http://bass2000.obspm.fr/home.php) makes available both OGAUC and Meudon spectroheliograph data, among other information. The OGAUC solar images are also available through the website of the Portuguese Institute for Sea and Atmosphere (IPMA).

The geomagnetic field North-South (X) and East-West (Y) components at Coimbra were obtained according to the standard procedures of a magnetic observatory and using the following instruments: (1) a DI-flux magnetometer (consisting of a fluxgate sensor MAG-01H mounted on MG2KP steel-free theodolite) and (2) a proton type magnetometer (the Overhauser GSM90-F1), for the absolute measurements (baseline control performed on weekly basis) of D & I and F, respectively; and (3) a digital variometer (model FGE, version J) for the continuous recording (with sampling rate of 1 Hz) of magnetic variations of components H, D and Z. The hourly series of XYZ were obtained from HZD series of 1-minute resolution. All these different components are properly introduced in section 2.1.

1.4 Thesis Outline

The main goal of this study is to achieve a more comprehensive understanding of the geomagnetic activity felt at Earth's surface mid-latitudes in the Northern Hemisphere by using statistical tools to compare three sources of data: a) ground geomagnetic field measurements of four magnetic observatories sparsely distributed in longitude: Coimbra (COI), Portugal; Panagyurishte (PAG), Bulgary; Novosibirsk (NVS), Russia and Boulder (BOU), USA, b) synthetic series of magnetosphere's field contribution at these observatories locations, computed using the Tsyganenko and Sitnov 2005 (TS05) semi empirical magnetosphere's model and c) satellite and other ground based data from

NASA/OMNI and related databases. The temporal window is from January 1^{st} 2007 to December 31^{st} 2016, covering an abnormally long minimum of solar activity at around 2008, the rising phase and part of the declining phase of solar cycle 24. This cycle is particularly interesting due to the triple polarity reversal in the Sun's Northern Hemisphere, its smaller amplitude with respect to previous cycles and the long duration of its minimum.

In Chapter 2, the main theoretical concepts that support the rest of the manuscript are explained.

In Chapter 3, Principal Component Analysis (PCA) and statistical tools are applied, in an attempt to disentangle the ionospheric and magnetospheric contributions to the local horizontal geomagnetic component, H, at the four magnetic observatories (COI, PAG, NVS and BOU). I compare observations with Dst and RC geomagnetic indices and with TS05 model predictions at the observatories locations, in order to evaluate the performance of TS05's predictions at ground level, for the period from January 2007 to December 2014. To compare the ground data with magnetospheric synthetic series (TS05), Dst and RC indices it is necessary to remove the ionospheric contribution from the data, since neither TS05 nor indices include the ionospheric contribution. To this end, I built a PCA model of the quiet daily variation (QD) and then removed it from the original (raw) data.

In Chapter 4, another statistical analysis is performed using the same ground observatories data, for January 2007 to December 2014, but separating the North-South (X) and East-West (Y) components, and discriminating the contribution of different TS05's magnetospheric source currents. The X and Y components of the six TS05 field sources and the total TS05 field were compared with local data. QD variation series were removed from data series by the standard method (not using PCA).

In Chapter 5, the Sun's facular areas computed by the Centre for Earth and Space Research of University of Coimbra (CITEUC) are compared with T-indices calculated from the X component of the four main TS05 currents and other solar, interplanetary

and geomagnetic parameters (33 in total), obtained from NASA/OMNI database, from January 2009 to December 2016, in the form of daily and 27-days averaged series, annual and 27-day averaged helio-magnetic asymmetries.

Finally, in Chapter 6, global conclusions are drawn.

2 Background

2.1 Geomagnetic Field

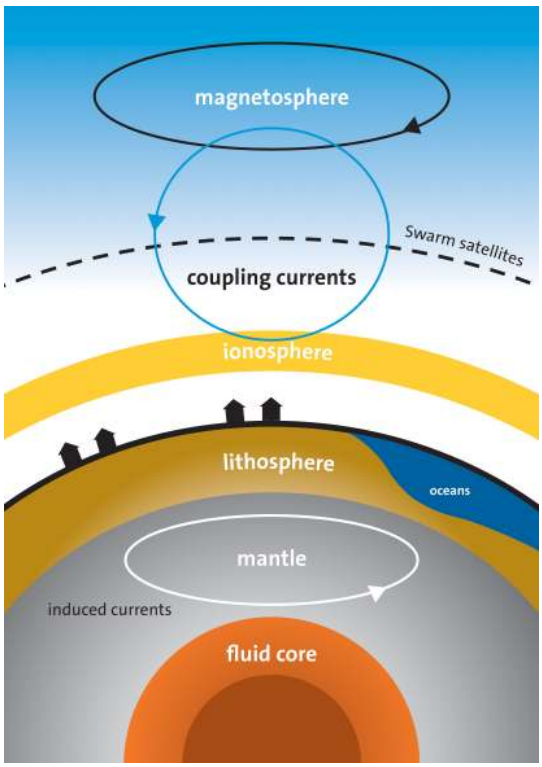


FIGURE 2.1: Sources of ground magnetic field measurements.
Image adapted from ESA: /https://tinyurl.com/y8o5uny4

The magnetometers' registrations at magnetic observatories show us the superposition of fields due to two type of sources: the internal and the external ones. The

internal sources are (see Figure 2.1) [1]:

1. the main field generated at Earth's liquid core,
2. the lithospheric or crustal field originated from magnetic minerals in the crust,
3. Faraday induced currents in the Earth's mantle and crust,
4. oceanic tidal currents produced by the Sun's and Moon's gravitational forces and the conducting oceanic currents cutting the geomagnetic field lines.

The external sources are: the ionospheric dynamo and different kinds of magnetospheric currents.

The main field constitutes the largest part of the Earth's field. The crustal field is responsible for a bias in different geomagnetic field series at a given observatory, when compared with main field models (e.g., Sabaka et al., 2015). The oceanic tidal currents are important mainly at stations near sea coasts. Because main field and crustal field remain constant at sub-daily time scales, they do not contribute directly to geomagnetic activity (see e.g., Pedatella et al., 2011). Geomagnetic activity, more directly related with SWE events, is due to external sources and corresponding induced fields.

Local North, East, Down (NED) Components. The components or elements of the total geomagnetic field at ground level can be measured in local Cartesian, spherical and cylindrical coordinates (see Figures 2.2 and 2.3). The Cartesian components are X, Y and Z, where positive X points to cardinal point North, positive Y points to cardinal point East, and positive Z completes the orthogonal coordinate system and points to the Earth's interior. The cylindrical components are H, D, Z, where H is the horizontal field intensity which results from $H = \sqrt{X^2 + Y^2}$, Z is the same as in Cartesian coordinates, and D is the magnetic declination, or the horizontal angle between the North-South direction and the horizontal field, positive if easterly. The

[1] In this thesis I used the /https://tinyurl.com/ service, that converts long into short URLs, which do not break when copied neither expire.

spherical components are F, D and I, where F is the total field vector intensity, that results from $F = \sqrt{X^2 + Y^2 + Z^2} = \sqrt{H^2 + Z^2}$, D is the same magnetic declination as in cylindrical coordinates, and I is the magnetic inclination, or the angle that the total field subtends with the horizontal field, positive if the total field points down.

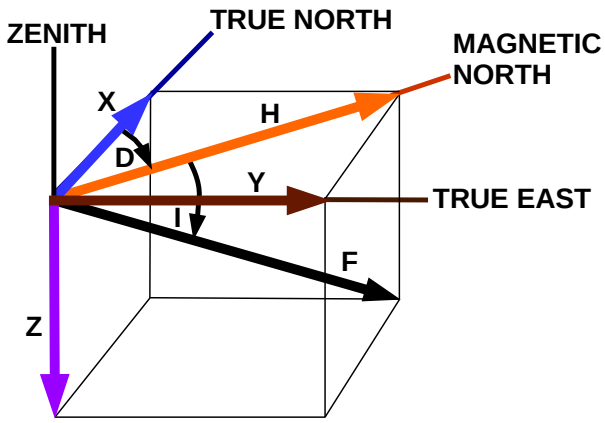


FIGURE 2.2: Geomagnetic field components in Cartesian, cylindrical and spherical coordinate systems.

2.1.1 The Main Field

The Earth's main magnetic field originates from convection processes at the outer fluid core, that move electrically conducting material and drive a dynamo mechanism. This is mainly forced by the gravitational energy released by the upward migration of lighter elements and the growth of the inner core by the freezing-out of heavier components (e.g., Campbell, 2003). This field contribution is almost steady on day scale but has a secular variation, significant over one or two years.

The main field has dipolar and non-dipolar components. The dipolar field approximation works well at magnetospheric, ionospheric heights and even on ground, for some purposes. The best-fit dipolar field presently is tilted 9.7^o with respect to the Earth's rotation axis, as seen in Figure 2.3, but this angle is reducing with time (e.g.,

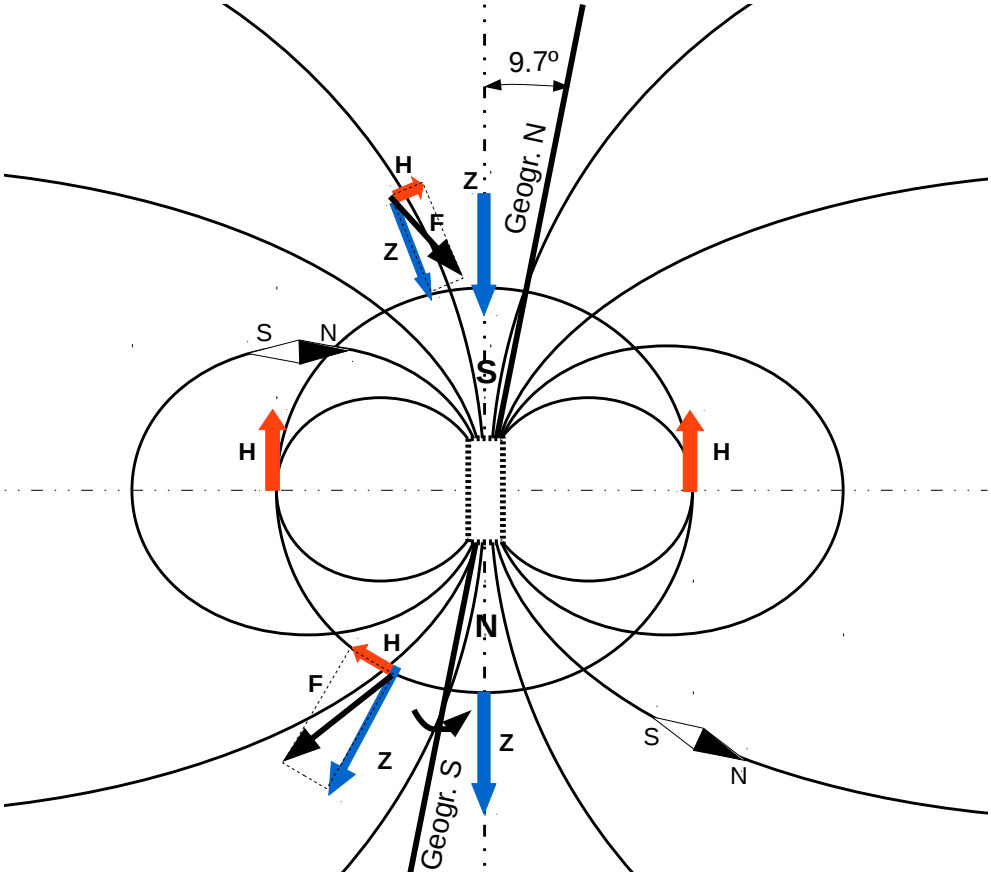


FIGURE 2.3: Orientation of the geomagnetic field components at different latitudes at Earth's surface.

Thebault et al., 2015). The Earth-centred dipole poles are called geomagnetic poles and are the hypothetical intersections of the Earth's surface and the geomagnetic dipole axis. They are the geomagnetic north pole (in the Northern Hemisphere) and the geomagnetic south pole. In 2017.0, the coordinates of these poles based on IGRF-12 are: 80.5° N, 72.8° W for the north and 80.5° S, 107.2° E for the south. The real magnetic poles (dip poles) are those where the magnetic needle becomes vertical. Coordinates of north and south magnetic poles at 2017.0 are 86.5° N, 172.6° W and 64.2° S, 136.3° E (/https://tinyurl.com/bnwrlsy). Both poles would coincide if the Earth's magnetic field would be perfectly dipolar.

An eccentric dipole (i.e., not centred at Earth) is more accurate to represent the Earth's dipolar field. It is shifted away from the center of the Earth by $\sim$ 576.8 km. Since 1950 it has been moving away from the center of the Earth at an almost constant speed of $\sim$ 2.5 km/year. In 2015.0 the eccentric dipole north pole was located at 84.14° N and 97.78° E, and the south pole was at 75.72° S and 117.51° W in geocentric coordinates (e.g., Laundal and Richmond, 2016).

A dipole's magnetic potential is defined as:

$$V(\mathbf{R}) = \frac{\mu_0}{4\pi R^3}\mathbf{m} \cdot \mathbf{R} \tag{2.1}$$

where **m** is the Earth's dipole moment, $\sim$ 7.94 x 10^{22} A m^2, **R** is the position vector from the Earth's centre and μ_0 is the permeability of free space, $\sim$ 4 π x 10^{-7} kg m $\mathrm{A}^{-2}\mathrm{s}^{-2}$. The magnetic field is the gradient of the magnetic potential:

$$\mathbf{B}(\mathbf{R}) = -\nabla V(\mathbf{R}) \tag{2.2}$$

In spherical coordinates, the magnetic field is expressed in terms of the radial, southerly and easterly components, B_R, B_θ and B_ϕ, where θ is the colatitude and ϕ is the longitude. Assuming that the geodipole is aligned with the Z axis, (see Figure

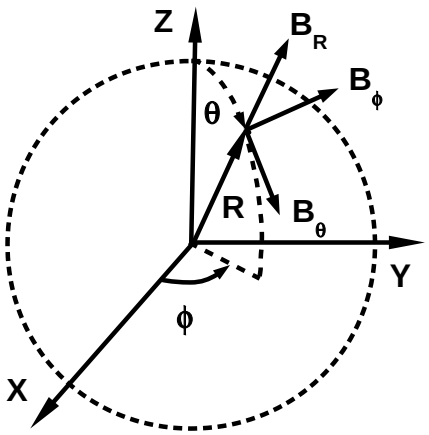


FIGURE 2.4: Geocentric spherical components of the geomagnetic field at Earth's surface.

x

2.4):

$$\begin{aligned} V(\mathbf{R}) &= \frac{\mu_0}{4\pi R^3}\mathbf{m}\cdot\mathbf{R} \\ &= \frac{\mu_0 m\cos\theta}{4\pi R^2} \end{aligned}$$

$$\begin{aligned} B_R(R,\theta,\phi) &= \frac{2\mu_0 m\cos\theta}{4\pi R^3} \\ B_\theta(R,\theta,\phi) &= \frac{\mu_0 m\sin\theta}{4\pi R^3} \\ B_\phi(R,\theta,\phi) &= 0 \end{aligned}$$

The total field intensity at any point is:

$$\begin{aligned} B(R,\theta,\phi) &= \sqrt{B_R^2 + B_\theta^2 + B_\phi^2} \\ &= \frac{\mu_0 m}{4\pi R^3}\sqrt{1+3\cos^2\theta} \end{aligned}$$

In particular, at the North Pole:

$$\begin{aligned} B_R(R,0,\phi) &= \frac{\mu_0 m}{2\pi R^3} \\ B_\theta(R,0,\phi) &= 0 \end{aligned}$$

and at the equator:

$$\begin{aligned} B_R(R,90^o,\phi) &= 0 \\ B_\theta(R,90^o,\phi) &= \frac{\mu_0 m}{4\pi R^3} \end{aligned}$$

Then the magnitude of the total field at the poles ($\sim$ 6 x 10^4 nT) is twice as strong as at the equator ($\sim$ 3 x 10^4 nT).

The components of the total field can be represented with spherical harmonics analysis (SHA), that divides the contributions into dipole, quadrupole, octupole, etc. and also separates contributions from internal and external sources. This mathematical representation helps to prove that the total field originates mostly from processes interior to the surface and that only a small part of the field comes from the Earth's exterior (e.g., Campbell, 2003). The so-called main field is the internal contribution due to the geodynamo.

Carl Gauss (1777-1855) developed this solution for the potential function V in spherical polar coordinates:

$$V = R_E \sum_{n=1}^{\infty} \left[\left(\frac{R}{R_E} \right)^n S_n^e + \left(\frac{R_E}{R} \right)^{n+1} S_n^i \right], \tag{2.3}$$

where $\sum$ means the sum of terms as n goes from 1 to infinity (in theory), R denotes the radial distance from the center of the Earth, R_E = 6371.2 km is the Earth's mean reference spherical radius, S_n^e and S_n^i represent functions called Legendre polynomials of the independent variable θ that are multiplied by sine and cosine function terms of

independent variable ϕ (e.g., Campbell, 2003).

The first term on the right of equation 2.3 represents the external sources of the potential function, and the second term represents the internal sources. The first increases with R and the second decreases with R due to the term $(1/R)^{n+1}$.

Since 1965, an international team of scientists under the auspice of the International Association of Geomagnetism and Aeronomy (IAGA) has been maintaining and producing main field models called International Geomagnetic Reference Field (IGRF), a series of mathematical models describing the large-scale main field $\mathbf{B}(\mathbf{R}, \theta, \phi, \mathbf{t})$ and its annual rate of change (secular variation), from epochs 1900 A. D. until now. It is the result of a collaborative effort among magnetic field modellers and institutions involved in collecting and disseminating magnetic field data from magnetic observatories, ground surveys, and low Earth orbiting (LEO) satellites. Expanding S_n^e and S_n^i in eq. 2.3, the potential V is approximated by the finite series:

$$V(R,\theta,\phi,t) \;=\; R_E \sum_{n=1}^{N} \sum_{m=0}^{n} \left(\frac{R_E}{R}\right)^{n+1} \left[g_n^m(t)\cos(m\phi) + h_n^m(t)\sin(m\phi)\right] P_n^m(\cos\theta), \quad (2.4)$$

with N = 13. The functions $P_n^m(\cos\theta)$ are the Schmidt quasi-normalized associated Legendre functions of degree n and order m. The Gauss coefficients g_n^m, h_n^m are functions of time and are conventionally given in units of nanotesla (nT). In the IGRF-12 model, the Gauss coefficients g_n^m and h_n^m are provided for the main field (MF) at epochs separated by 5 years between 1900.0 and 2015.0 A.D. The time dependence of the Gauss coefficients is assumed to be linear over 5-year intervals. The maximum truncation degree N = 13 for epochs after 2000 is defined so as not to include the crustal magnetic field contributions that dominate at higher degrees (see e.g., Langel and Estes, 1985). For more details, see e.g., Thebault et al., 2015.

The IGRF is widely used for a wide variety of studies, like dynamics of Earth's core field, space weather, crust field, geomagnetic indices, etc. The IGRF model is revised every few years ($\sim$ 5 years) to follow the continuous temporal changes of the field of

Earth's outer core. Each generation consist of three models:

1. the Definite Geomagnetic Reference Field: (DGRF) any further improvement of these models is unlikely;

2. the IGRF model: is non-definitive and will be substituted eventually by the DGRF;

3. the secular variation (SV) model: is provided to predict the time variation of the large-scale geomagnetic field for the five years following the latest IGRF.

For example, the most recent 12th-generation IGRF (IGRF-12) provides a DGRF model for epoch 2010.0, an IGRF model for epoch 2015.0, and a predictive SV model covering the epochs 2015.0 - 2020.0. The current and previous versions of IGRF are available at /https://tinyurl.com/ybwcb9xk.

2.1.2 The Ionospheric and Induced Fields

Looking at the magnetograms of field components at any observatory, some days we can see a smooth pattern very easy to identify, as a stationary oscillation (Figure 2.5, left), but other days we see a very noisy signal with higher amplitude oscillations (see Figure 2.5, right). These patterns have different sources. The smooth daily-oscillation or quiet daily variation (QD) is mainly originated by the effect of solar irradiation over the ionosphere's E region, a highly conducting region between 85 and 200 km altitude, producing an ionospheric dynamo (see Figure 2.6) that is driven by thermally and gravitationally excited atmospheric tidal winds (e.g., Campbell, 2003), producing a dynamo effect that generates two ionospheric current vortices, one counter-clockwise at Northern Hemisphere and other clockwise at Southern Hemisphere (e.g., Yamazaki and Maute, 2016). The term solar quiet (Sq) variation refers to these ionospheric currents that displace with Earth's rotation (e.g., Pedatella et al., 2011; Yamazaki and Kosch, 2014). However, it is also known that the quiet daily (QD) variation contains a

significant magnetospheric contribution besides the ionospheric one (e.g Olsen, 1996) and the separation is not simple (e.g., Langel et al., 1996).

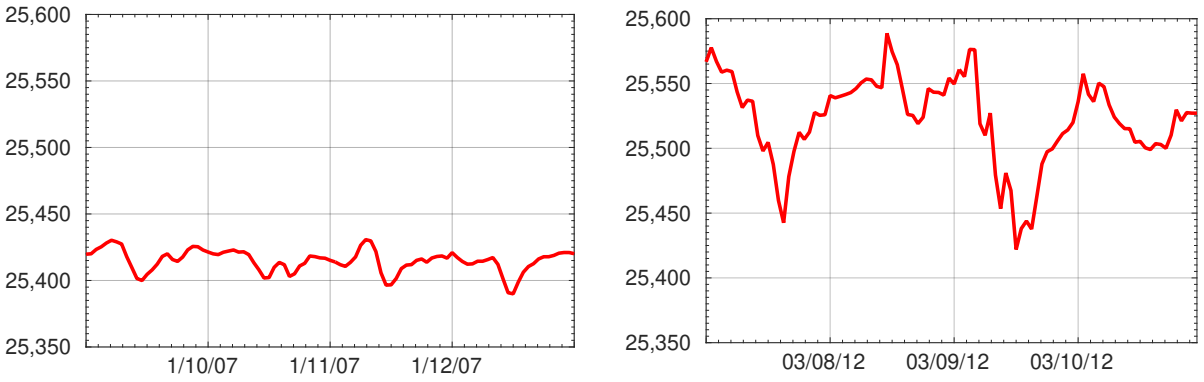


FIGURE 2.5: Left: Horizontal component measured at COI station when $K < 4$ (January 9-12, 2007). Right: $COI - H$ component when $K \geq 4$ (March 7-10, 2012). Units in nT.

Electromagnetic induced currents are produced in the crust and upper mantle (e.g., Schmucker, 1985). Their strength is roughly one third of that of the ionospheric currents, as determined, e.g., by Matsushita and Maeda (1965) and Langel and Estes (1985) applying spherical harmonic analysis to Earth-based stations and Magsat-satellite dataset, respectively. According to Lenz's law, these currents tend to have an opposite direction to their ionospheric sources and as a result their contribution affects QD variations by reducing the vertical component Z and increasing the X and Y components (e.g Yamazaki and Maute, 2016). Most often, a 1-D conductivity model is used to explain this effect, consisting of an insulating upper mantle and a superconductor below some depth. In this scenario, the secondary field is simply proportional to the primary inducing field and the thickness of the insulating upper layer is adjusted in order that the proportionality constant is close to 1/3 (e.g. Olsen, Sabaka, and Lowes, 2005). More realistic models allowing for lateral variations of conductivity show that the Z (vertical)

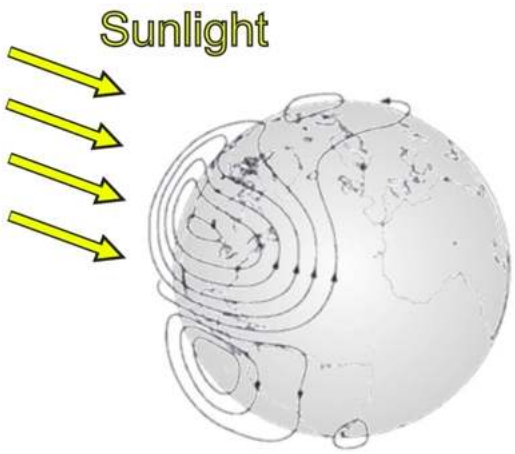


FIGURE 2.6: The ionospheric dynamo. Image adapted from British Geological Survey Website /https://tinyurl.com/3hyrszf

geomagnetic field component is the most affected by the conductivity model simplification (e.g Kuvshinov, Avdeev, and Pankratov, 1999).

2.1.3 Lunar Variations

Lunar variations are regular geomagnetic variations which arise from lunar tides in the oceans and atmosphere. If tidal currents with velocity $\mathbf{V}$ move across field lines of the Earth's planetary field, electric fields $\mathbf{V} \times \mathbf{B}$ are generated. They drive electric currents in the highly conducting seawater and in the ionosphere. The motion-induced currents produce electromagnetically induced fields. The superimposed magnetic field of all currents is observed as lunar daily variations L (Schmucker, 1985), but their amplitudes are rather small and have not been taken into account in this work.

2.1.4 The Magnetospheric Field

A magnetosphere is a cavity inside the interplanetary medium where the solar wind and interplanetary magnetic field interact with the intrinsic magnetic field or ionized

upper atmosphere of a planetary body. The Earth has a dominant quasi-dipolar magnetic field, a dense atmosphere and ionosphere. The size of the magnetosphere is determined by the pressure balance between the external solar wind dynamic pressure and the internal magnetosphere's pressure (e.g., Pulkkinen, 2007). The plasma in the magnetosphere consists mainly of electrons and protons, coming from the SW and the Earth's ionosphere, minor fractions of He^{+} and O^{+} coming from the ionosphere and some He^{2+} coming from the Sun (e.g., Baumjohann and Nakamura, 2007). The magnetopause is the magnetosphere's boundary where the solar wind, magnetospheric plasmas and magnetic fields are in pressure balance. At typical solar wind conditions it extends until 10 $\mathrm{R_E}$ at the dayside, but at strong SW activity the dayside can be compressed until 6.6 $\mathrm{R_E}$ (Pulkinnen, 2007).

The plasma inside the magnetosphere is grouped into different regions with different densities and temperatures, from which we can highlight (see Figure 2.7):

1. the plasma mantle, just inside the magnetopause boundary and flowing in the anti-solar direction,

2. the plasma sheet, separating the northern and southern lobes of the magnetospheric tail,

3. the plasmasphere and Van Allen belts, a torus-like region around the Earth, of protons and electrons that became trapped in the Earth's main field.

These plasma reservoirs are sources of particles for different magnetospheric currents. Most authors (e.g., Tsyganenko, 2013; Pulkkinen, 2007; Russell, 2001) distinguish the following currents:

1. two ring current systems in the magnetic equator, one axi-symmetrical, or full-ring current, and other asymmetrical, or partial ring current,

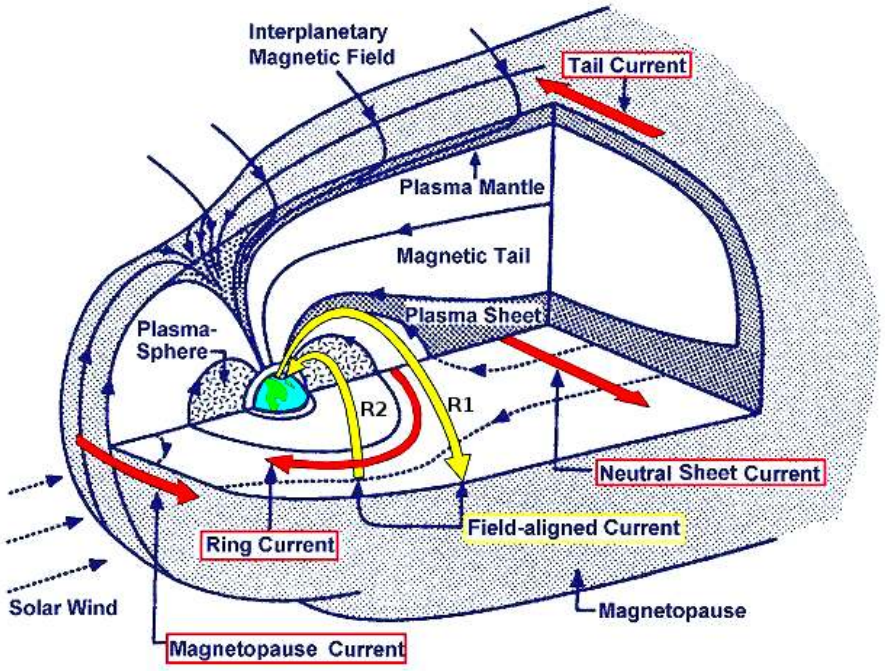


FIGURE 2.7: Magnetosphere's current systems and plasma regions. Image adapted from https://tinyurl.com/yakbg2mu

2. a cross-tail current sheet in the anti-sunward direction,

3. a system of vertical currents called field-aligned currents or Birkeland currents, that connects ionosphere and magnetosphere,

4. a boundary current system called Chapman-Ferraro or magnetopause currents.

The ring current (RC) lies between about 1.2 - 10 $\mathrm{R_E}$ (Tsyganenko, 2014), with a magnetic field strength of the order of a few hundred nanotesla, typical electron densities and temperatures of 1 $\mathrm{cm^{-3}}$ and 5 x 10^7 K respectively (e.g., Baumjohann and Nakamura, 2007). This current is due to the eastward (electron) and westward (proton) drift in the Van Allen radiation belts. During geomagnetic storms the ring current causes a net decrease in the magnetic field on the Earth's surface, opposed to the magnetopause current that causes an increase. The energy of these circulating particles can be easily calculated from their effect on the ground-level magnetic field. In major magnetic storms this energy can reach 10 or more petajoules (10^{15} J) and the energization

rate can exceed several terawatts (10^{12} W) (e.g., Russell, 2001). The symmetric storm-time ring current requires at least several hours to build up and decays on the time scale of at least one or two days (e.g., Tsyganenko and Sitnov, 2005).

During geomagnetic storms the asymmetric part of the ring current, or partial ring current (PRC), develops a strong near-equatorial magnetic depression at the dusk - pre-midnight sector, with peak at 18:00 - 20:00 h and at $\sim$ 6 - 7 $\mathrm{R_E}$ (e.g., Li, Wang, and Kan, 2011; Tsyganenko, 2013), with a relaxation time $\leq$ 2 hours (Tsyganenko and Sitnov, 2005). This excess of westward current on the night side suggests the existence of a field-aligned current, downward on the dusk side and upward on the dawn side (e.g., Tsyganenko, 2000) (see R2 current in Figure 2.7).

Most of the magnetotail plasma is concentrated around the tail mid plane in an about 5-10 $\mathrm{R_E}$ thick plasma sheet. Average electron densities and temperatures in the tail plasma sheet are 0.5 $\mathrm{cm^{-3}}$ and 5 x 10^6 K, with magnetic fields of 10 - 20 nT (e.g., Baumjohann and Nakamura, 2007). A northern and a southern low density tail lobes form in the night side magnetotail. Beyond 20 $\mathrm{R_E}$ the fields in the northern and southern tail lobes are nearly anti-parallel and have an almost constant intensity of $\sim$20 nT (e.g., Pulkkinen, 2007). Near the Earth, the tail lobes are threaded by magnetic field lines originating in the polar caps that contains a highly rarefied plasma. The lobes typical values for the electron density, temperature and the magnetic field strength are 10^{-2} $\mathrm{cm^{-3}}$, 5 x 10^5 K and 30 nT respectively. The tail-like field of the night side magnetosphere is accompanied by the current flowing on the tail magnetopause surface and the cross-tail neutral sheet current in the central plasma sheet, both of which are connected (e.g., Baumjohann and Nakamura, 2007). The effect of the tail current system is (as for the ring current) to oppose the Earth's field and has a stronger effect on the night side, causing a day-night gradient in the field (e.g., Russell, 2001). The tail current dramatically increases during the main phase and shifts earthward, so that the peak current concentrates at unusually close distances 4 - 6 $\mathrm{R_E}$. This is accompanied

by a significant thinning of the current sheet and strong tailward stretching of the inner geomagnetic field lines. A typical reaction time of the tail lobe field to the solar wind pressure pulses is only a few minutes (e.g., Collier et al., 1998), even though its response to the onset of southward IMF B_Z has a longer time scale, up to 2 - 3 hours (e.g., Tsyganenko, 2000).

Physically, the inner ring current and the more distant tail current sheet form a single equatorial current system. The difference between the two is that the ring current lines encircle Earth and are fully closed inside the magnetosphere, whereas the tail currents flow in the night side and close via the magnetopause (e.g., Tsyganenko, 2013).

The Birkeland or field-aligned current (FAC) system is mainly carried by electrons (e.g., Baumjohann and Nakamura, 2007) and is responsible for electrodynamically linking and exchange of energy and momentum among the magnetopause, the inner magnetosphere and the ionosphere (e.g., Coxon et al., 2014). This system forms two concentric rings above the auroral ionosphere: the poleward (region 1) ring and the equatorward (region 2) ring. The region 1 (R1) currents connect the ionosphere to currents in the magnetopause and the magnetotail, and the region 2 (R2) currents divert the excess of azimuthal current from the dusk side of the PRC (Tsyganenko, 2002a). Region 1 currents flow downward in the dusk sector and upward in the dawn sector, and region 2 currents are of opposite polarity. Both regions 1 and 2 currents close through the ionosphere (see Figure 2.7).

Finally, the current flowing in the magnetopause is called the magnetopause current or Chapman-Ferraro (CF) current. This current is caused by the deflection of charged particles in the magnetopause. Positive and negative particles arriving at the magnetosphere's boundary carried by the solar wind are deflected in opposite directions normal to the local field, giving the eastward magnetopause current shown in Figure 2.7.

2.2 Geomagnetic Activity

2.2.1 Storms

The geomagnetic storms are intrinsically dynamical events, to which contribute the current solar wind and interplanetary magnetic field (IMF) conditions, the magnetospheric inertia (delayed reaction of the magnetosphere to changing its conditions) and "memory" effects, reflecting a finite response time of loading/unloading processes as the dissipation of energetic particles trapped on the inner drift shells (Tsyganenko and Sitnov, 2005).

At mid and low latitudes we can identify the imprint of geomagnetic storms when we have these three stages in a H-component record (see Figure 2.8 [2]):

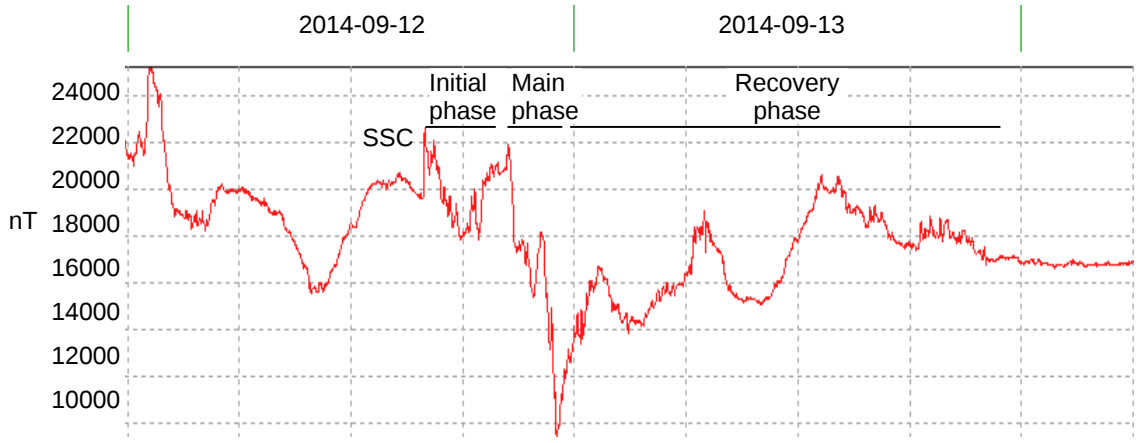


FIGURE 2.8: Phases of a geomagnetic storm record measured at COI station.

1. Initial phase: it starts with a sudden increase in the H component with respect to its average level, called storm onset. Its corresponding time is the starting time or onset time of the storm. In most cases this is also the beginning of a storm sudden commencement (SSC), defined by J. J. Curto as "a sharp change (with a minimum

[2] Image courtesy of Paulo Ribeiro.

slope of the order of 10 nT in 3 min) of the horizontal magnetic field at globally spaced observatories at low latitude,... followed by an hour with the Dst index lower than about -50 nT within the following hours" (e.g., Curto, Araki, and Alberca, 2007). The storm onset is originated when a sudden increase of solar wind dynamic pressure reaches the magnetopause, causing an interplanetary shock or discontinuity. It happens more frequently when the interplanetary magnetic field points southward (IMF $B_Z < 0$) in either the turbulent sheath fields behind the shock, in the smooth magnetic clouds fields behind the sheath, or in both.

2. Main phase: the horizontal component of the field decreases and shows major fluctuations for a longer time and larger amplitude than in the initial phase. The magnitude of the decrease represents the severity of the storm (see the following sections).

3. Recovery phase: the longest in time (as much as several days) during which the field returns to its undisturbed level.

2.2.2 Indices

Geomagnetic activity indices (GAI) have been designed to describe variations in the geomagnetic field caused by irregular current systems. They can be separated into indices describing geomagnetic activity at mid-latitudes and those at polar latitudes. This section describes all known standard indices used in this study. In Chapter 5, a new set of indices is proposed (T-indices) to put in evidence the effect of separate magnetospheric currents.

Mid-Latitude Indices

Local K geomagnetic index. The K-index is related quasi-logarithmically to the geomagnetic disturbance amplitude measured for the horizontal component at a particular observatory, during a three-hour interval and after the quiet daily variation has been removed. An integer from 0 to 9 is assigned to each 3-hour interval, therefore each day is characterised by 8 K-indices. Observatories at higher geomagnetic latitudes have higher levels of fluctuation for a given K-index. For example, Table 2.1 compares amplitudes of local K geomagnetic indices for Coimbra, Panagyurishte, Novosibirsk and Boulder.

TABLE 2.1: Amplitudes in nT corresponding to K indices at Coimbra, Panagyurishte, Novosibirsk and Boulder. COI values provided by Paulo Ribeiro. PAG values have same order as COI. NVS and BOU from Space Weather Prediction Center (SWPC/NOAA) at /https://tinyurl.com/goolkxk (at tab "Details")

K index	COI and PAG (nT)	NVS and BOU (nT)
0	0-4	0-6
1	4-10	6-14
2	10-21	14-30
3	21-38	30-54
4	38-67	54-96
5	67-112	96-160
6	112-196	160-280
7	196-336	280-480
8	336-560	480-800
9	> 560	> 800

Planetary Kp and Ap indices. The name Kp is derived from the german *Planetarische Kennziffer* (= planetary index) and was introduced by Bartels in 1949, as a global 3-hour geomagnetic index. The series of Kp exists from 1932 until now. Since 1997 the Kp and related indices are derived at the Adolf Schmidt Geomagnetic Observatory Niemegk

of the Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences. The ftp server is located at /https://tinyurl.com/yc5akrab. GFZ-Postdam description of Kp index can be found at /https://tinyurl.com/yca8ak5u. This index is the base used by the United States National Oceanic and Atmospheric Administration (NOAA) to classify geomagnetic storms. It is calculated by a weighted average of the local standardized "K" indices of 13 geomagnetic observatories between 44^o and 63^o northern or southern geomagnetic latitudes. The standardization is done using tables designed by Bartels for each observatory and based in statistical methods. The 3-hour Kp index ranges in a scale from 0 (quiet) to 9 (greatly disturbed), distributed in 28 steps, in thirds of a unit, e.g. 5- is 4 2/3, 5 is 5 and 5+ is 5 1/3. Due to the quasi-logarithmic relationship of the K-indices to magnetometers fluctuations, to calculate the daily average level of geomagnetic activity the Kp-scale must be converted back into an equivalent linear three hourly scale called ap-index, dividing by 2 the amplitude range corresponding to each Kp (see Table A.1 at Appendix A). The Ap-index is calculated as the arithmetic average of eight 3-hourly averaged ap-indices of a day.

Table A.1 in Appendix A shows the Kp scale, their equivalent amplitudes, and the equivalent ap three-hourly values. Sources: /https://tinyurl.com/ybwd7etl and /https://tinyurl.com/yb3zroby.

Storm-time disturbance index (Dst). We call disturbance field (Dst) the storm-time decrease in H, measured at the Earth's surface as a result of geomagnetic activity.

The Dst index is linearly dependent on the amplitude of the geomagnetic perturbation and is derived from the H hourly values obtained from four magnetic observatories at low and mid-latitudes and distributed evenly in longitude: Honolulu (Hawaii), San Juan (Puerto Rico), Hermanus (South Africa) and Kakioka (Japan). These observatories are sufficiently far away from auroral and equatorial electrojets and those current systems have a weak effect on this index. Dst provides a 1-hour resolution description of magnetic activity during geomagnetic storms.

Originally derived for monitoring the symmetric ring current variations, it is presently acknowledged that other magnetospheric (Chapman-Ferraro, tail current, field-aligned currents, etc.) and ionospheric currents contribute to Dst (e.g., Campbell (1996), Feldstein et al. (2000), Tsyganenko and Sitnov (2005), Curto, Araki, and Alberca (2007)).

SYM-H geomagnetic index. The Symmetric Disturbance Horizontal (SYM-H) index is obtained from 1 min data of the magnetic field at six stations between $\lambda \approx 40^{o} - 50^{o}$ (Boulder, Chambon-la-Foret, Fredericksburg, Martin de Vivies, Memambetsu, Urumqi) and five between $\lambda \approx 18 - 30^{o}$ (San Juan, Tucson, Honolulu, Alibag and Hermanus) (e.g., Iyemori et al., 1999). Usually, it is considered that the SYM-H index is an analog of Dst but with a better temporal resolution, however during super magnetic storms with Dst variation $\geq$ 400 nT, the SYM-H variations can considerably differ from Dst behavior (e.g., Solovyev, Boroyev, and Moiseyev, 2005).

Ring current index (RC). Recently, Olsen et al. (2014) put forward a new index, called RC (from "ring current"), which describes the strength of the magnetophere's ring current even during geomagnetic quiet conditions (when the Dst baseline gives less optimal results). RC represents approximately a dipole field, since it is derived only from the horizontal components of 21 observatories at low and mid latitudes distributed around the world, all longitudes, at night hours (LT between 18 and 06).

The criteria to determine the quiet time baseline at each observatory is well defined and made constant in time, contrary to the Dst baseline (e.g., Temerin and Li, 2015).

For every observatory the core field model of CHAOS-4 is removed from the hourly mean values. After that, an observatory bias that represent the lithospheric field is also removed from every hourly value. This bias is determined such that the arithmetic mean value during geomagnetic quiet periods (defined as $Kp \leq$ 2o, $|dDst/dt| \leq$ 2 nT hr^{-1}) vanishes. Next, all data is converted from the geographic to the geomagnetic

(dipole) frame. An hour-by-hour spherical harmonic analysis (SHA) of H hourly mean values is performed, estimating the three spherical harmonic expansion coefficients of first degree and order 0 and 1. For this analysis only observatories that were in the night-side were used, with the number of observatories in use changing from hour to hour. The RC index is defined as RC = $-\nu_1^0$ (Olsen et al., 2014). The minus sign is to make RC compatible with Dst as a southward component at the dipole equator. ν_1^0 denotes the sum of an external coefficient ε_1^0 (magnetospheric) and an internal coefficient ι_1^0 (induced) (e.g., Olsen, Sabaka, and Lowes, 2005).

Polar Region Indices

AE, AU, AL. Hourly AE, AU and AL are computed and delivered by the World Data Center for Geomagnetism, operated by the Data Analysis Center for Geomagnetism and Space Magnetism at Kyoto University, Japan (/https://tinyurl.com/yattk7lb). The AE index (from Auroral Electrojet) is derived from geomagnetic variations in the horizontal component observed at selected (10-13) observatories along the auroral zone in the Northern Hemisphere (list of previous and current AE observatories available at /https://tinyurl.com/yd93kf4v). To normalize the data a base value for each station is calculated for each month by averaging all the data from the station on the five international quietest days. This base value is subtracted from each value of one-minute data obtained at the station during that month. The plots of all stations are superimposed and the largest and smallest values are selected at each given time (UT). The AU and AL indices are respectively defined by the largest and the smallest values so selected. AU and AL names derive from upper and lower values. The difference, AU minus AL, defines the AE index. The AU and AL indices are intended to express the current intensity of the eastward and westward auroral electrojets, respectively. The AE index represents the overall activity of the electrojets. More detailed information at /https://tinyurl.com/yahsv5oc.

Polar Cap Index (PC). This is determined from the North polar cap station at Thule, Greenland (PCN) and the South polar cap station at Vostok, Antarctica (PCS). The index is basically a 15-min index that Low Resolution OMNI (LRO) averages to hourly resolution. I used PCN taken from /https://tinyurl.com/y776okcj, World Data Center for Geomagnetism, National Space Institute, Copenhagen. PC is calculated with the purpose to monitor the geomagnetic activity over the polar caps caused by changes in the interplanetary magnetic field (IMF) and solar wind, driven by the geo-effective interplanetary electric field. It is considered a convenient proxy of energy that enters into the magnetosphere by solar wind - magnetosphere coupling.

2.2.3 Coupling Functions

Interplanetary Electric Field. The PC index is closely related with the value of the interplanetary electric field E_{KL} (e.g., Kan and Lee, 1979) controlling the polar cap magnetic activity:

$$E_{KL} = V\ B_T \sin^2(\theta_C/2) = V\ \left[(B_Y^{GSM})^2 + (B_Z^{GSM})^2\right]^{1/2} \sin^2(\theta_C/2) \tag{2.5}$$

where V is the solar wind velocity, B_T is the transverse component of the magnetic field, θ_C is the clock angle or angle between $\mathbf{B}_T$ and Z axis. Then E_{KL} is a product of the transverse field by the solar wind velocity ($E_T = V\ B_T$) multiplied by an estimate of the fractional merging rate ($\sin^2(\theta_C/2)$). The expected relation is:

$$PC = \xi\ E_{KL},$$

where ξ is a scale coefficient (e.g., Newell et al., 2007).

Newell's coupling function. Newell et al. (2007) derived a nearly universal coupling function that includes the rate of magnetic flux removed from the dayside magnetopause, $d\Phi_{MP}/dt$ that is itself an electric field. They calculate the rate at which magnetic flux is opened at the magnetopause as a function of the rate at which field lines approach the magnetopause (determined by the solar wind velocity V), the fraction of field lines that merge at the magnetosphere (calculated by a sine function of clock angle θ_C) and the strength of the transverse IMF (B_T, that is proportional to the amount of flux opened).

The greatest amount of merging is predicted for a southward IMF, but it is important to note that there is not a "northward IMF" magnetosphere neither a "southward IMF" magnetosphere. There is a magnetosphere under a state of rapid dayside merging and other under slow dayside merging. This is why there are more "southward" cusp signatures than "northward" and why the polar cap for northward IMF so often resembles "southward IMF" conditions.

For this work I used a version of Newell's function derived by Tsyganenko and Andreeva (2015), that introduces a normalization factor of 10^{-4} and I denote the merging function as NLL index:

$$NLL \;=\; 10^{-4}\, V^{4/3}\, B_T^{2/3}\; \sin^{8/3}(\frac{\theta_c}{2}) \tag{2.6}$$

with V in km/s and B_T in nT. Tsyganenko and Andreeva (2015) compute the NLL index for each 5 min average data record as an average over 30 min long trailing interval, immediately preceding the current time moment. 5-min data were binned to 1-day resolution in order to compare with other parameters. Tsyganenko's 5-min Newell's data can be downloaded at /https://tinyurl.com/y7agg2hr.

2.3 Relevant Coordinate Systems

It is useful to employ different coordinate systems for physical phenomena with different geometries and development, depending on which system makes it easier the calculations and the understanding of phenomenon. I separated the coordinate systems according to the phenomena they help to describe: phenomena mainly related with the Sun and its magnetic field (IMF), phenomena strongly related to the arrival of solar wind on Earth and Sun-Earth interactions, and phenomena strongly constrained by the Earth's main field. The Geographic Coordinate System is also described at the beginning.

All coordinate systems described hereunder are geocentric. In some coordinate systems position is often specified in terms of colatitude θ, longitude ϕ and radial distance R.

The following descriptions are mainly based in SPENVIS Coordinate Systems Website (/https://tinyurl.com/ydxtkbh5). An excellent animation with most coordinate systems can be found at /https://tinyurl.com/yb9wjuxq.

Geographic Coordinate System (GEO). In this work, this reference system is used to indicate the position of the magnetic stations and other locations at Earth's surface. The positive X-axis lies in the Earth's equatorial plane and points from the centre of the Earth through the Greenwich Meridian (0^o longitude). The Z-axis is parallel to the Earth's rotation axis and its intersection with the Northern Hemisphere is the North geographic pole. The geographic east longitude (ϕ) is the horizontal and Earth-centred angle measured at the equatorial plane, from Greenwich Meridian, counter clockwise (eastward) as seen from the North geographic pole. The geographic latitude (λ) is the Earth-centred vertical angle measured from the equatorial plane to the measured location at Earth's surface. It is positive for a Northern Hemisphere's location.

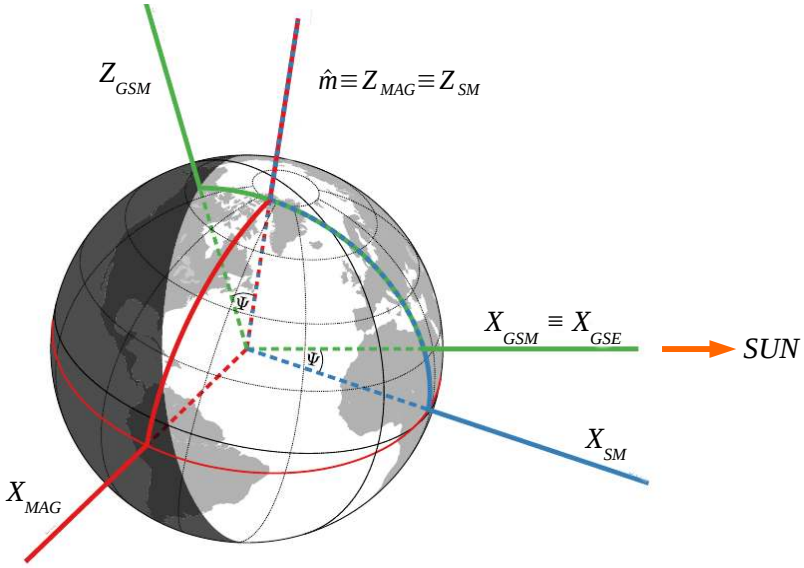


FIGURE 2.9: X and Z axes of Geocentric Solar Magnetospheric Coordinate System (GSM, green), Solar Magnetic coordinates (SM, blue) and Geomagnetic Coordinate System (MAG, red). $\hat{m}$ represents the Earth's dipole axis. ψ is the dipole tilt angle. Image adapted from Laundal and Richmond (2016).

2.3.1 Coordinate systems strongly constrained by the Earth's main field

Geomagnetic Coordinate System (MAG). This reference system was used to obtain the magnetic coordinates of the observatories at Earth's surface. The Z-axis is parallel to the magnetic dipole axis, positive North. The Y-axis is perpendicular to the plane containing the dipole moment and the Earth's rotation axis. In this work I used the geographic coordinates of the dipole's North pole, 80.09 ^{o}N, 72.21 ^{o}W, calculated for year 2010.0 (e.g., Thebault et al., 2015). Values for other years can be found at Kyoto Data Center (/https://tinyurl.com/bnwrlsy).

Solar Magnetic Coordinates (SM). The Z-axis is parallel to the dipole axis, positive North, and the positive Y-axis is perpendicular to the Earth-Sun direction towards dusk. X-axis does not point directly to the Sun. The difference from MAG is a rotation about the common Z axis, by an angle that changes from 0 to 360^{o} along the day (see Figure 2.9). This system is useful to analyse magnetic fields aligned with the geomagnetic dipole, as field-aligned currents and the main field.

2.3.2 Coordinate systems mainly related with the Sun and its magnetic field

Geocentric Solar Ecliptic (GSE). Its origin is at the center of the Earth. The X-axis is the Earth-Sun direction towards the Sun. The Z-axis points to the north ecliptic pole, perpendicular to the plane that contains the Earth's orbit around the Sun. This system has a yearly rotation relative to a system fixed to the distant stars, i.e., the Geocentric Earth Inertial system (GEI), which X-axis points to the first point in Aries or Vernal Equinox (intersection of ecliptic and Earth's equatorial plane). GSE system was used to study the IMF influence over the Earth's field both when IMF points toward or away from the Sun.

Geocentric Solar Equatorial System (GSEQ). This system also has its X-axis pointing towards the Sun from the Earth. However, instead of having its Y-axis in the ecliptic plane, the GSEQ Y-axis is parallel to the Sun's equatorial plane which is inclined 7.25^o to the ecliptic. Positive Z axis points northwards. The Z-axis will not necessarily be parallel to the Sun's axis of rotation, but Sun's axis of rotation must lie in the X-Z plane. Half of the year the Earth is North of the GSEQ equatorial plane, the other half is South, with a maximum inclination value of 7.25^o. This system is used in this thesis to study the projection of the GSEQ Y component over the GSM Z component of the interplanetary magnetic field (see Chapter 5).

2.3.3 Coordinate systems strongly related to the arrival of solar wind on Earth

Geocentric Solar Magnetospheric Coordinate System (GSM). This is one of the most used coordinate systems for Earth-Sun interactions, where X is the Earth-Sun direction, with origin at Earth's center, positive towards the Sun (same as for GSE). Positive Z axis points towards the North, is perpendicular to X and is always in the plane that contains the magnetic dipole axis and the X axis. Positive Y axis points towards dusk. The angle between Z GSM axis and the magnetic dipole axis is called tilt angle, ψ (see Figure 2.9). In the GSM frame the dipole axis is always in the X-Z plane, although the tilt angle can change from $\sim$-30^o to $\sim$30^o. This system is used to describe phenomena strongly constrained by the SW and the IMF (external magnetosphere). The magnetospheric model used in this study (TS05 model, see section 2.4) uses the GSM-Y and GSM-Z IMF components and their respective GSM positions as input parameters. TS05 outputs are in GSM system.

Geocentric Solar Wind Coordinate System (GSW). This system is analogous to the GSM, except that the X-axis is anti-parallel to the actual observed direction of the solar

wind flow, which not only aberrates by $\sim 4^o$ from the strictly radial direction due to Earth's orbital motion around the Sun, but also often significantly fluctuates around its average direction. Orientation of axes in the GSW system can be uniquely defined by specifying three GSE Cartesian components of the solar wind velocity vector (V_X, V_Y, V_Z). In the absence of reliable data on the solar wind direction, one can either set $V_Y = V_Z = 0$ (in which case the GSW system becomes identical to the standard GSM) or set V_Y = 29.8 km/s, $V_Z = 0$, thus taking into account the aberration effect. This reference system is used in the TS04c code for TS05 simulations in order to simplify some calculations along the solar wind direction.

2.4 Tsyganenko and Sitnov 2005 (TS05) Semi-Empirical Model

The Tsyganenko and Sitnov 2005 (TS05) model is the result of more than 30 years of developing data-based semi-empirical models of the Earth's magnetosphere (for a review, see e.g., Tsyganenko and Sitnov, 2005; Tsyganenko, 2013).

TS05 is a dynamical semi-empirical model of the Earth's inner magnetosphere, developed from space measurements of solar wind and IMF parameters taken during 37 major geomagnetic storms between October 1996 to November 2000. The total number of data records used to develop the model are 142 787, diversely distributed around the Earth, but mainly at geosynchronous orbits ($\sim$ 6.6 $\mathrm{R_E}$, Earth radii):

$$\sim 23.0\% \text{ between } 7 \leq \mathrm{R} \leq 20\ \mathrm{R_E}$$

$$\sim 68.0\% \text{ between } 6 \leq \mathrm{R} \leq 7\ \mathrm{R_E}$$

$$\sim 5.3\% \text{ between } 4 \leq \mathrm{R} \leq 6\ \mathrm{R_E}$$

$$\sim 3.7\% \text{ between } 2.5 \leq \mathrm{R} \leq 4\ \mathrm{R_E}$$

Table 2.2 shows the number of data and temporal coverage of every spacecraft.

The magnetic field model is represented as the sum of seven vectors or "modules" having physically realistic and flexible spatial structure. The total field model is fitted to a large set of data covering a sufficiently wide region in the geometrical and parametric space (Tsyganenko and Sitnov, 2005). These seven vectors are (see Figure 2.7):

$$\mathbf{B}_{TSY} = \mathbf{B}_{CF} + \mathbf{B}_{SRC} + \mathbf{B}_{PRC} + \mathbf{B}_{TAIL} + \mathbf{B}_{FAC1} + \mathbf{B}_{FAC2} + \mathbf{B}_{INT} \tag{2.7}$$

where

1. $\mathbf{B_{CF}}$ is the field of the Chapman-Ferraro or magnetopause currents which confine the Earth's field inside the magnetopause,
2. $\mathbf{B_{SRC}}$ is the contribution of the symmetric ring current,
3. $\mathbf{B_{PRC}}$ is the contribution of the partial ring current,
4. $\mathbf{B_{TAIL}}$ is the contribution of the cross-tail current sheet,
5. $\mathbf{B_{FAC1}}$ is the contribution of the region 1 of the field-aligned currents or Birkeland currents,
6. $\mathbf{B_{FAC2}}$ is the contribution of the region 2 of the Birkeland currents,
7. $\mathbf{B_{INT}}$ is the contribution of an interconnection current system, representing the coupling between IMF and magnetopause.

To make the tail field more flexible it can be expanded into a linear combination of two fields, $\mathbf{B_{T1}}$ and $\mathbf{B_{T2}}$ with different spatial variations at tail axis, corresponding to an inner part and an outer part, but I just used the global $\mathbf{B_{TAIL}}$ in this work.

These vector fields represent the effect of the structure and temporal variation of the major current systems contributing during the entire storm cycle, since their growth to their decay. In the TS05 model, the density, velocity and pressure of the upstream solar wind drive and shape the magnetosphere, as also does the IMF B_Z component.

TABLE 2.2: Number and time distribution of data records by spacecraft

Spacecraft	**Time Period Covered**	**No. 5-min averages**	**Total Time, hours**
Polar	Oct 1996 - Nov 2000	45,354	3,779
Geotail	May 1997 - Nov 2000	4,237	353
Equator-S	Feb 1998 - Mar 1998	1,659	138
GOES-8	Oct 1996 - Nov 2000	50,070	4,173
GOES-9	Oct 1996 - May 1998	20,439	1,703
GOES-10	Apr 1999 - Nov 2000	21,028	1,752
Totals		142,787	11,899

They are all used to assign different energy feeding and dissipation rates to different components. The dynamics of each source is represented as a competition between the external driving (variations of the solar wind parameters) and the internal dissipation (reconnection at magnetopause, plasma convection, charge exchange, particle losses due to diffusion [3] and others). All these effects are empirically modelled by including a term W to represent the strength of a field source, whose magnitude behaves in time according to the equation:

$$\frac{\partial W}{\partial t} = S - L \tag{2.8}$$

where the quantities S and L in the right-hand side are the source and loss functions, respectively.

The source term S represents the feeding rate of a current system by the SW input, which can be empirically assumed as a function of the external driving factors:

$$S = a\rho^{\delta}V^{\beta}(B_Z)^{\kappa} \tag{2.9}$$

[3]When the resistivity increases anomalously at a particular point, the magnetic field gradients become large and the magnetic field can vanish due to diffusion.

where ρ and V are the SW density and speed, B_Z is the southward component of the IMF, the coefficient a and the power indices δ, β and κ are unknown parameters of the driving function, to be fitted by comparing the model with data.

The loss term L has a different physical meaning, depending on which current system is being considered. For the symmetrical part of the ring current, it is closely related to the rate of the dissipation of energetic particles due to their pitch angle [4] and radial diffusion, as well as charge exchange processes (Tsyganenko and Sitnov, 2005).

L is assumed to be proportional to the difference between the current value of W and its residual quiet-time level W_0:

$$L = r_d(W - W_0), \tag{2.10}$$

which yields a simple solution for W (see eq. 2.8)

$$W(t) = W_0 + \int_0^t S(\tau)\, e^{[r_d(\tau - t)]} d\tau, \tag{2.11}$$

The integration is made from the beginning of the event $\tau = 0$ to the current moment $\tau = t$; r_d is the decay rate, an inverse measurement of the relaxation time scale, $T = 1/r_d$ (or "inertia") of a source. Larger values of r_d provide a faster reaction of the magnetospheric current to an external disturbance and its quicker return to the quiet-time level after the driving force disappears.

The model parameters and input data files are frequently updated by Nicolai Tsyganenko or collaborators at /https://tinyurl.com/y8mmfwws. This Website provides input data as yearly-files, but they can be merged to obtain a single file with several years. Every input data file contains:

1. date and time in UT

2. ϕ, λ' and h' geodetic coordinates using the WGS84 ellipsoid (see Appendix A.2)

[4] Pitch angle is the angle between the direction of the magnetic field and the velocity vector of a charged particle.

3. V_X, V_Y and V_Z GSE components of solar wind velocity
4. IMF B_X, B_Y and B_Z GSM components
5. solar wind proton density, ρ
6. solar wind proton temperature, T
7. Symmetric Horizontal Index (SYM-H)
8. dipole tilt angle, ψ
9. solar wind ram pressure, p

Most input data were obtained from the High Resolution OMNI (HRO) - NASA data set, that hosts near-Earth solar wind magnetic field and plasma parameters from several spacecraft in geocentric or L1 (Lagrange point) orbits, extensively cross compared, and, for some parameters, cross-normalized. For more details see: /https://tinyurl.com/y8kuv4jf, /https://tinyurl.com/y8vo8v3m and Appendix A.

In OMNI Web, several solar wind parameters are computed from satellite data using subscripts p and a to represent protons and alpha particles respectively. For N, the number of particles per cm^3, V, the flow speed in km/s, and m mass, typical values are $V_a = V_p$, N_a/N_p = 0.04 - 0.05, although there may be significant differences for different flow types. Using those values, for ρ and p they obtain:

$$\rho \;=\; 1.2\, m_p\, N_p, \tag{2.12}$$

where m_p is the proton mass, N_p is the number of protons per cm^3. Flow ram pressure is derived from:

$$p \;=\; 2 \;\times\; 10^{-6}\, N_p\, V_p^2, \tag{2.13}$$

where p is in nPa and V_p is the proton flow speed in km/s.

The Chapman-Ferraro $\mathbf{B}_{CF}$ field is fully defined by the strength and tilt angle of the Earth's dipole and by the shape and size of the magnetosphere's boundary. In the TS05 model, the magnetopause is a predefined surface, fitted to the empirical model of Shue (1998) but controlled only by the solar wind ram pressure (e.g., Tsyganenko, Singer, and Kasper, 2003). The magnetopause field $\mathbf{B}_{CF}$ is added to all other parts of the total **B** vector to confine (or shield) other currents inside the magnetospheric magnetic field, so that

$$\mathbf{B} \cdot \mathbf{n} \,|_{S'} \;=\; 0, \tag{2.14}$$

where S' is the magnetopause boundary and **n** is a unit vector normal to S' (Tsyganenko, 2013).

TS05 model separates the fields of an axisymmetric part of the ring current ($\mathbf{B}_{SRC}$) and an asymmetric one ($\mathbf{B}_{PRC}$), specifying for each one its proper distribution of the equatorial plasma pressure (e.g., Tsyganenko, 2013). Both fields were derived on the basis of the observed profiles of the particle pressure distribution, according to the AMPTE/CCE spacecraft data in Lui and Hamilton (1992) (LH92). From the pressure profile the volume electric current density **j** is obtained, and the calculation of the vector potential **A** follows, using the Biot-Savart integral:

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\tau' \tag{2.15}$$

Then, the magnetic field $\mathbf{B} = \nabla \wedge \mathbf{A}$ follows.

At sufficiently large tailward distances (beyond a distance of R = 10 $\mathrm{R_E}$ on the night side), the particle population of the ring current gradually merges with that of the tail plasma sheet. The equatorial current flow lines no longer encircle Earth but reach the magnetopause and close there, which corresponds to the transition from the ring current to the cross-tail current system.

For $\mathbf{B}_{PRC}$, TS05 assumes that the local time variation of the pressure profile p is rather smooth and can be approximated by a certain function defined on one meridional plane and then multiplied by cos $\mathrm{m}\phi'$ or sin $\mathrm{m}\phi'$, where ϕ' is the magnetic longitude, closely related with the local time (Tsyganenko, 2000). During quiet conditions the TS05's $\mathbf{B}_{PRC}$ is very weak in comparison with $\mathbf{B}_{SRC}$, but it rapidly increases and peaks with growing SW pressure and enhanced magnetospheric convection, rotating into the dusk sector, and quickly subsiding once the external driver is turned off (Tsyganenko, 2002b; Tsyganenko and Sitnov, 2005).

The tail field ($\mathbf{B}_{TAIL}$) is represented with analytical approximations, corresponding to equatorial current sheets with finite variable thickness. Their current densities are vanishingly small near Earth, but increase in thickness with growing distance, peak near the observed position of the inner edge of the cross-tail current, and gradually decreases further tailward. The effects of the dipole tilt on the shape of the cross-tail current were taken into account using a deformation technique (e.g., Tsyganenko and Sitnov, 2005).

TS05's Birkeland current systems giving rise to $\mathbf{B}_{FAC}$ are proportional to the sine of the magnetic longitude, ϕ', thus at low altitudes they peak at dawn and dusk. The magnetic longitude is here defined as being zero when the local magnetic meridian crosses the subsolar point. To obtain $\mathbf{B}_{FAC}$, the TS05 model uses a distribution of radial currents flowing on a deformed conical surface that intersects Earth along circles of constant geomagnetic latitude (e.g., Tsyganenko and Sitnov, 2005). Applying a deformation of coordinates the distribution of the currents becomes much closer to the expected, based on a more realistic shape of their flow lines. Their spatial extent and local time distribution can be easily varied and fitted to the spacecraft data. By construction, they are symmetric with respect to the noon-midnight plane, but multiplication by the sine of the longitude ϕ' makes the field-aligned currents at low altitudes to peak at dawn and dusk. The strongest disadvantage of TS05's FAC currents is that they are assumed to pierce the ionosphere and extend all the way to the Earth's center,

while they actually should close via the ionosphere (Tsyganenko, 2002a). $\mathbf{B}_{FAC}$ is decomposed into $\mathbf{B}_{FAC1}$ and $\mathbf{B}_{FAC2}$ (see eq. 2.7). R1 currents have a faster reaction than R2 at the beginning of a geomagnetic storm, because R1 is connected to magnetopause. R2 are considered the closure currents for the PRC (Tsyganenko, 2002a).

Finally, the TS05 model also calculates the interconnection term, $\mathbf{B}_{INT}$, though a simple equation:

$$B_{INT} = \epsilon B_T, \tag{2.16}$$

where B_T is the transverse component of the IMF (see section 2.2.3), observed upstream from the bow shock, the penetration coefficient ϵ was assumed to depend on the IMF clock angle θ_C, allowing for a possibility of different penetration efficiency during northward IMF (quite small θ_C) and southward (larger θ_C) conditions (Tsyganenko, 2002a; Tsyganenko, 2002b). TS05's best fit value of ϵ was found to be 0.46.

2.5 The Sun-Earth Environment

Sun's magnetic field lines have opposite magnetic polarities at each hemisphere, extending radially near the poles, but approaching near the equatorial plane (e.g., Pneuman and Kopp, 1971, see Figure 2.10). By convention, if magnetic field lines point away (outward) from the Sun, both field lines and hemisphere's polarity are named positive. When field lines point toward (inward) the Sun both field lines and hemisphere's polarity are negative. If the Sun's Northern Hemisphere is positive and Southern Hemisphere is negative, the Sun's polarity is so-called positive. In the opposite case is negative (see Figure 2.11).

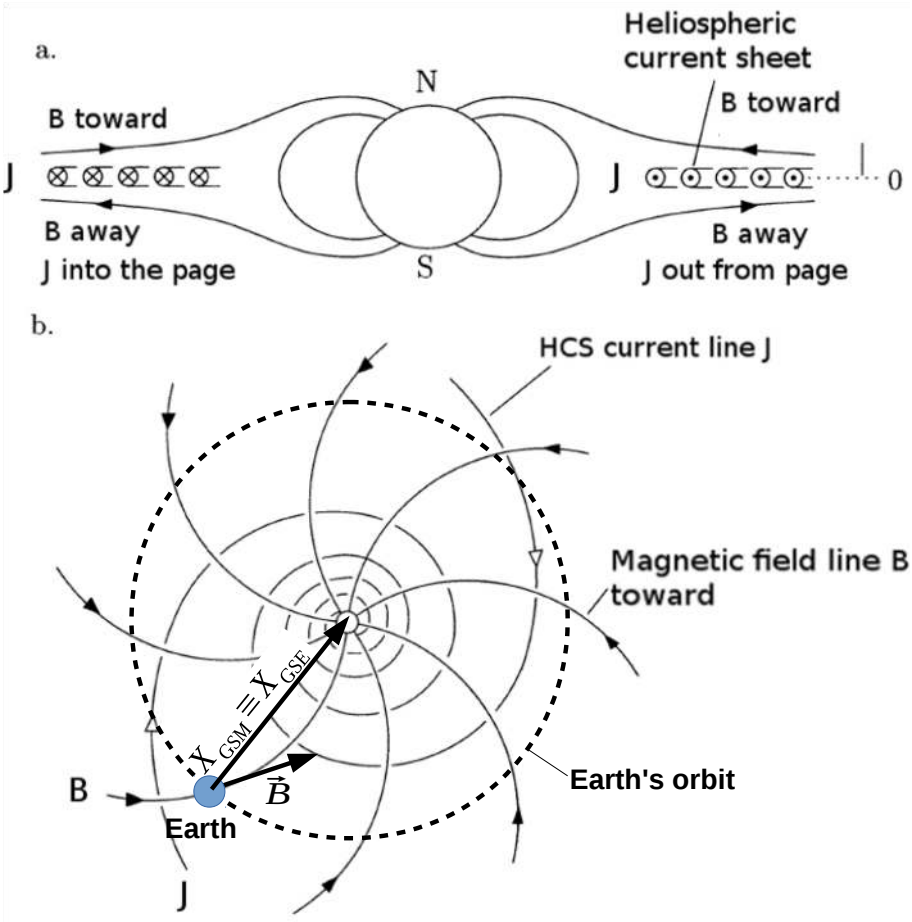


FIGURE 2.10: Top: Meridional plane containing the solar rotation axis; on the left, HCS's current (J) points into the page; on the right, J points out of the page. Continuous lines represent magnetic field lines: toward the Sun at Northern Hemisphere and away from the Sun at Southern Hemisphere (Sun's negative polarity). Bottom: view of solar equatorial plane as seen from above, showing IMF lines (B) perpendicular to electric current lines (J), Earth's orbit (dashed circle), Earth's position (blue circle) and $X_{GSM} = X_{GSE}$ axis. Image adapted from /https://tinyurl.com/ya3fzaw7

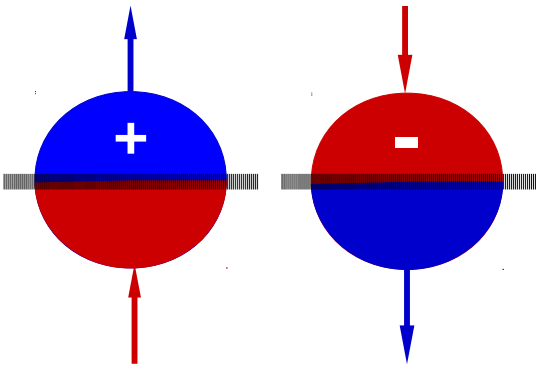


FIGURE 2.11: Left: Sun's positive polarity. Magnetic field points outward at N-hemisphere (blue) and inward at S-hemisphere (red). Right: Sun's negative polarity (inverse situation). Discontinuous lines represent heliospheric current sheet (HCS).

2.5.1 The Heliospheric Current Sheet

In the interplanetary medium, and near the solar equatorial plane, the IMF changes sign abruptly within a narrow layer. This layer is a sheet of current density that circulates around the Sun's dipole axis in the same direction as the original current that explains the Sun's dipole field. This structure is called heliospheric current sheet (HCS) and separates fields and plasma flows of both hemispheres, as seen in Figure 2.10. Even though the current density is only ~ 1 x 10^{-10} A m^{-2}, the associated intensity is huge because of the very large HCS cross-section. Rotation of the Sun twists the IMF lines in such a way that they take the shape of an Archimedean spiral and a structure that resembles a ballerina skirt (e.g., Parker, 1958), as seen in Figure 2.12). The spiral form of the magnetic field lines means that there is a significant radial component of the electric current in the sheet, along with the azimuthal component. Figure 2.10 shows the distribution of the currents along the HCS according to Riley, Linker, and Mikic (2002). It also shows the IMF field lines as seen from above the solar North Pole, in time of negative polarity. At Earth's orbit (1 AU), these field lines make an angle of $\sim 45^o$ with the $\mathrm{X_{GSM}} \equiv \mathrm{X_{GSE}}$ axis. Along the solar cycle, HCS also changes from an almost flat

sheet at solar minimum to sheet waves extended at more than 70^o helio-latitudes at solar maximum (see Figure 2.13).

FIGURE 2.12: John Wilcox (researcher) and Werner Heil (artist) picture inferred from coronal models based on solar wind and coronal field data. Here Earth runs two times over the HCS and two times below it. Image adapted from /https://tinyurl.com/grtwf3h

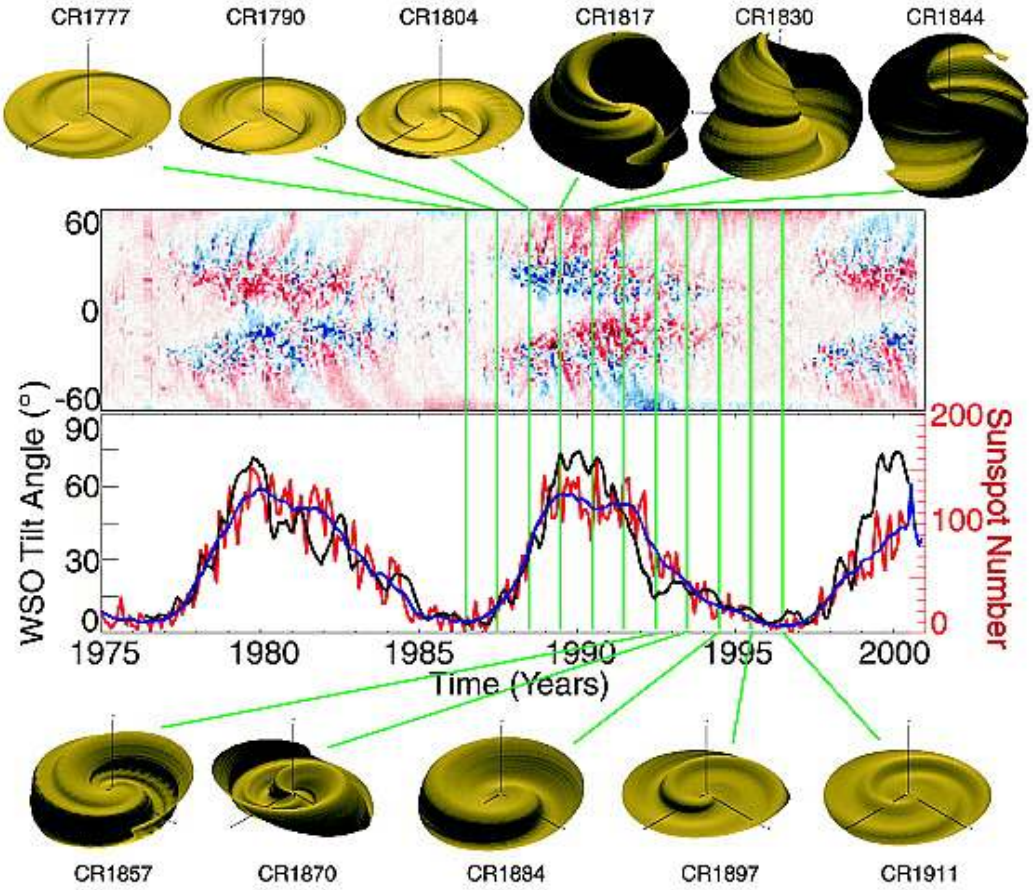


FIGURE 2.13: Some solar parameters during cycles 21, 22 and 23. The lower panel shows the monthly (yearly) averaged values of total sunspot number in red (blue). The upper panel shows the butterfly diagram of the radial component of the magnetic field inferred from Kitt Peak synoptic maps. Blue indicates toward (or inward, or '-') polarity, and red indicates away (or outward, or '+') polarity. The yellow and black figures above and below these panels are the HCS shapes at 5 AU, derived by a magnetohydrodynamical (MHD) model, for 11 Carrington rotations, from mid-1986 to mid-1996. Image adapted from Riley, Linker, and Mikic (2002)

2.5.2 Solar Polar Magnetic Flux and Polarity Reversals

The polar magnetic flux of the Sun is a measure of the Sun's magnetic flux from $|60^o|$ to higher latitudes. Polar magnetic fields are maximal during solar activity minima and minimal during maxima. Figure 2.14 shows the N-hemisphere (blue) and S-hemisphere (red) filtered polar fluxes for 2006-2015, measured by the Wilcox Solar Observatory, along the line-of-sight, in Gauss.

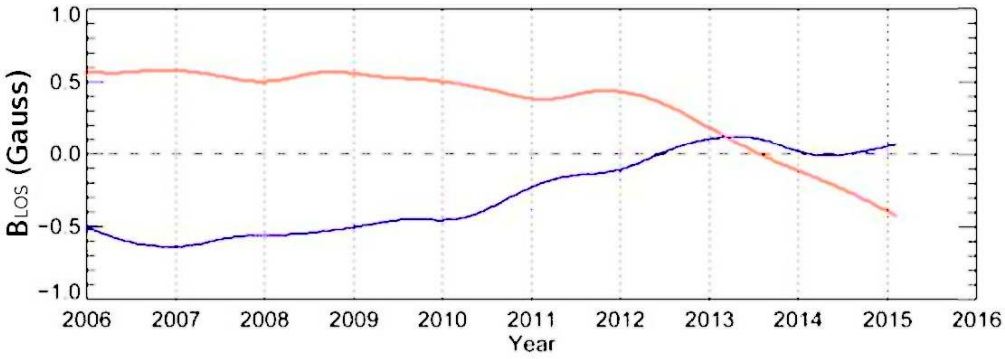


FIGURE 2.14: Wilcox line-of-sight filtered polar flux measurements (B_{LOS}) for 2006-2015 period, in Gauss. The N-hemisphere measurements are shown in blue and the southern values are shown in red. Image adapted from Mordvinov et al. (2016).

TABLE 2.3: Sun's Northern and Southern polarities at 2009-2016. '-' represents inward IMF, '+' is outward IMF, ($\pm$) means both inward and outward magnetic fields at one hemisphere.

	2009	2010	2011	2012	2013	2014	2015	2016
Northern Hemisphere	-	-	-	$\pm$	$\pm$	$\pm$	+	+
Southern Hemisphere	+	+	+	+	$\pm$	$\pm$	$\pm$	-
Sun's Polarity	-	-	-	$\pm$	$\pm$	$\pm$	$\pm$	+

The changing in polar magnetic flux polarity gives us information about polarity reversals developing at the Sun's interior. These reversals start near the solar maximum, beginning at lower latitudes ($|55^o|$) and expand to higher latitudes ($|85^o|$)

(e.g., Pishkalo and Leiko, 2016) . The current solar cycle 24 began in December 2008, after a long minimum and reached its maximum in 2014. Before this maximum, total magnetic flux near the polar zones was negative in the N-hemisphere and positive in the S-hemisphere (see Table 2.3 and Figure 2.14). The first polarity reversal of this cycle started around June 2012 at $\sim 55^o$ in N-hemisphere, changing from inward to outward until it reached $\sim 85^o$ in June 2013. After that, N-hemisphere's polarities oscillated between positive and negative due to surges of active regions, unipolar magnetic regions and coronal holes (e.g., Mordvinov et al., 2016). Second and third N-hemisphere's polarity reversals were faster than the first and covered all latitudes nearly simultaneously ($\sim 55^o$ - 85^o) from December 2013 to June 2014. The polarity reversal in the S-hemisphere was single but delayed, starting in April 2013 at $\sim$-55^o and finishing in March 2015 at $\sim$-85^o.

2.5.3 Sun-Earth Interaction

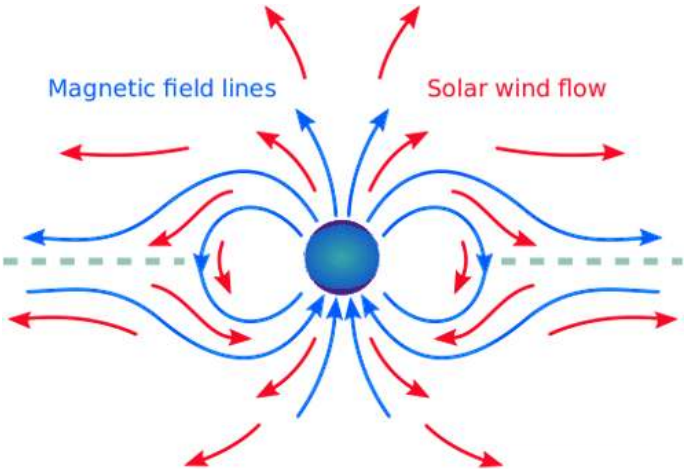


FIGURE 2.15: Sketch of meridional view of solar wind (red) and IMF field lines (blue) near the Sun. The dashed line represents the HCS. Image adapted from NASA /https://tinyurl.com/yd2776yr

The irregularities seen in magnetic measurements come from interactions of the solar wind with the magnetosphere, from the magnetosphere itself, from interactions of

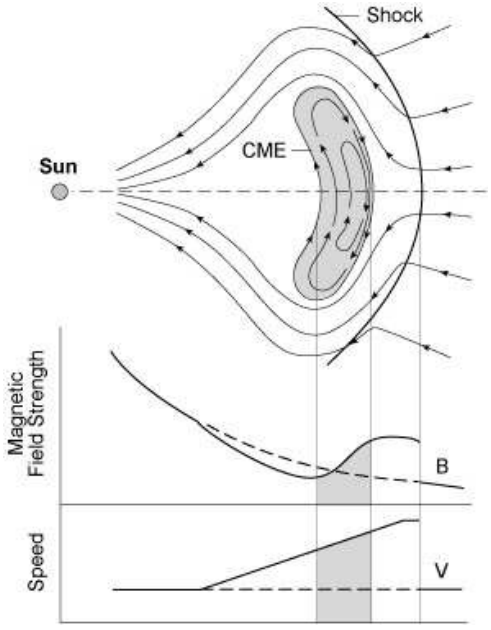


FIGURE 2.16: Sketch of coronal mass ejection. As it moves away from the Sun (top left) the CME (top right) pushes an interplanetary shock wave ahead of it, amplifying the magnetic field strength, B (middle), and the solar wind speed, V (bottom). Image from NASA's Cosmos. https://tinyurl.com/y8amk2sl

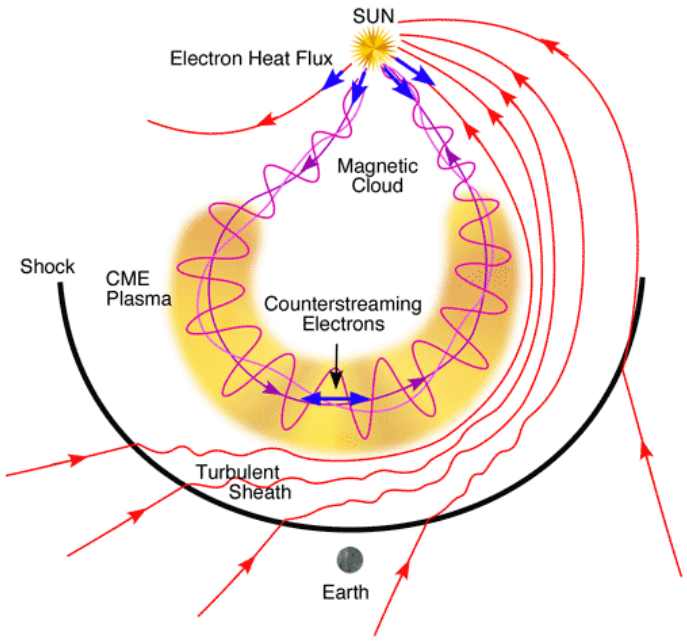


FIGURE 2.17: Sketch of a magnetic cloud. A magnetic cloud contains bidirectional, or counter-streaming, beams of electrons that flow in opposite directions within the magnetic loops that are rooted at both ends in the Sun. It also drives an upstream shock ahead of it. Image from NASA's Cosmos. https://tinyurl.com/y8amk2sl

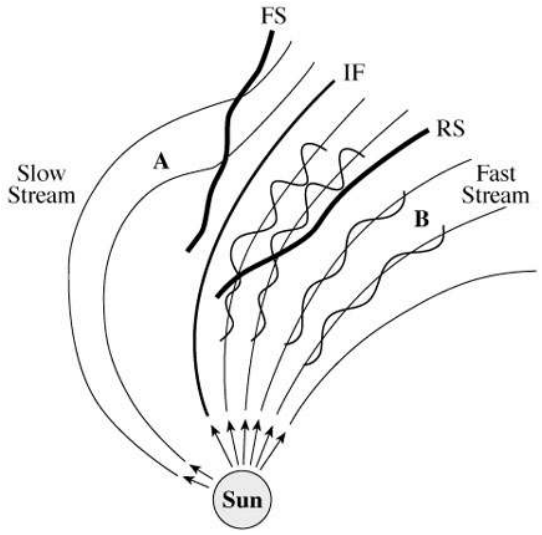


FIGURE 2.18: Sketch of a co-rotating interacting region. CIRs are produced when fast solar-wind streams, mainly emanating from CHs, interact with slow streams in interplanetary space. The magnetic fields of the slow streams are more curved due to the lower speeds, and the fields of the fast streams are more radial because of their higher speeds. Intense magnetic fields can be produced at the interface (IF) between the fast and slow streams. The CIR have a forward shock (FS) ahead and a reverse shock (RS) behind. Image from NASA's Cosmos. https://tinyurl.com/y8amk2sl

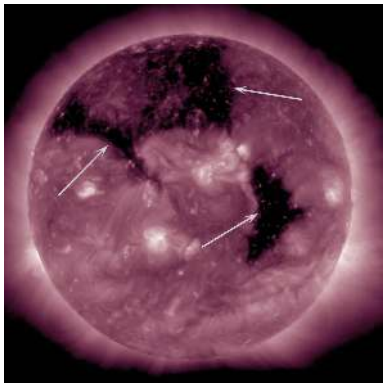


FIGURE 2.19: Coronal holes appear as dark areas in extreme ultraviolet light, because there is less material in the hole to give off light in these wavelengths. This image was taken in wavelengths of 211 angstroms, which is typically colourised in purple. Credit: NASA/SDO. https://tinyurl.com/y9993w7a

magnetosphere with ionosphere and the ionosphere itself. The solar wind (SW) emerging from the Sun is a plasma of protons, electrons and ions freezed to the IMF, that expands through the interplanetary space (see Figure 2.15). The SW is caused by different events in the Sun. One of them are coronal mass ejections (CMEs), that originate in the solar corona. These are large masses of solar plasma and magnetic field that cause large disruptions at the interplanetary medium (e.g., Owens and Forsyth, 2013, see Figure 2.16). Other important source of perturbations are magnetic clouds, that are similar to CMEs, but with more intrinsic magnetic and kinetic energy (e.g., Owens and Forsyth, 2013, see Figure 2.17). Co-rotating interaction regions (CIRs) are masses of solar-wind ejected by the Sun at different solar latitudes and velocities that interact between them during their travel along the interplanetary medium, producing frontward and backward shock waves (e.g., Owens and Forsyth, 2013, see Figure 2.18). The high speed streams frequently originate at coronal holes (CH) (e.g., Owens and Forsyth, 2013, see Figure 2.19), that are low-density regions in the solar corona with magnetic fields that open freely into interplanetary space. During low solar activity, coronal holes cover the Sun's northern and southern polar caps. During more solar active periods, coronal holes can exist at all solar latitudes, but they may only persist for several solar rotations before evolving into a different magnetic configuration. Coronal streamers, especially the streamer belt at equatorial latitudes, are sources of slow solar wind flows. The different velocities of these flows in interplanetary space lead to compression of the solar wind plasma on the forward edge of the high-speed flow and thus to a compression of the frozen-in magnetic field. Should this ambient magnetic field already possess a negative B_Z component, it can be amplified to the point where a geomagnetic storm is triggered as the compression region passes the Earth (e.g., Schwenn, 2006). SW due to all these events may increase kinetic and electric energy of all magnetospheric current systems. Their effect is seen at the Earth's magnetic poles as auroras, at the magnetic equator as an electro jet, and at the mid-latitudes as geomagnetic storms. The energy input is possible when solar wind carrying a southward pointing IMF component hits

the northward pointing Earth's magnetic field line on the front of the magnetosphere, initiating significant geomagnetic disturbances. This phenomena is known as magnetic reconnection. At quiet conditions IMF near Earth does not have northward or southward components, but the solar wind due to events mentioned above tilt the IMF generating the southward component.

2.5.4 Sunspots and Facular Areas

Sunspots. These are regions at the Sun's surface (photosphere) of intense magnetic field and lower temperature ($\sim$3700 K) than the rest of the photosphere (5700 K). They appear darker because the higher magnetic field inhibits the rise of heat from the solar interior. Each sunspot has a dark core (umbra), where the magnetic field is stronger and a less dark halo (penumbra), where magnetic field is weaker. Typically they subsist for periods of days until weeks, with a bipolar magnetic structure and are mainly restricted to the activity belts reaching up to 30^o - 40^o on each side of solar equator. Their number varies over solar cycles. At solar minimum activity no sunspots are observed along several days and at solar maximum activity, ten or more sunspots are common. Near minimum, sunspots appear at high latitudes, at $\pm 50^o$ from equator, appearing at increasingly lower latitudes until reaching close to the equator. The representation of the heliospheric latitude of sunspots along time is so-called butterfly diagram (e.g., Hathaway, 2010). The discovery of opposite magnetic polarities of the leading sunspot in Northern and Southern Hemisphere's bipolar structures and their reversal from one solar cycle to the next lead to discover that the solar polar field reverses over a solar cycle (Hale's law). This can be explained due to the transport of higher latitude fields toward the poles where they eventually reverse the polar field at about the time of sunspot cycle maximum. In 1844 Samuel Heinrich Schwabe, based on observations of sunspot groups and spotless days, reported the presence of a cycle of activity of about

10 years. The period of a sunspot cycle is defined as the elapsed time from the minimum preceding its maximum to the minimum following its maximum. This is subject to the uncertainties in determining the dates of minimum. The average cycle period is approximately 11 years but varies in length with a standard deviation of about 14 months (e.g., Hathaway, 2010).

Facular areas or photospheric faculae. These structures surround active regions, appearing bright toward the limb, influencing total solar irradiance variations (e.g., Kostik and Khomenko, 2012). All sunspots are surrounded by them, but not all faculae have sunspots inside them. Faculae are aligned mottles 5 000 to 10 000 km wide and up to 50 000 km long, composed of oval-shaped coarse mottles with diameters of about 5 000 km. The coarse mottles are made up of facular granules, about 1 000 km in size. Their integrated area is substantially larger than that of the associated sunspot group. Faculae appear in increased numbers in a region prior to the emergence of sunspots and remain for a rotation or more after the spots have decayed. The average lifetime of photospheric facular areas is 90 days. The temperature in faculae is several hundred degrees Kelvin higher than that of the quiet, photosphere and radiate more energy. The faculae are important to the energy balance between sunspots and the photosphere. Photospheric faculae extend into the chromosphere (the lower layer of Sun's atmosphere) as chromospheric faculae or plages, that can be seen using filters like Ca H and Ca K (e.g., Evren, 1999). Faculae are strongly correlated with the variation of the activity of the Sun (e.g., Guttenbrunner et al., 2014). During a sunspot cycle the facular areas make the Sun appear slightly brighter at sunspot maximum than at sunspot minimum. A great advantage of facular areas over other solar structures is that on the one hand they are almost permanently on the Sun, even during minima of solar activity, and on the other hand they have been less studied than sunspots.

2.5.5 Solar Rotation

The solar rotation period at low latitudes, observed in a fixed coordinate system, is about 25 days (sidereal rotation period); the synodic rotation period (as observed from the Earth) is about two days longer. This is due to the motion of the Earth about the Sun, which is in the same direction as the solar rotation. The rotation period increases with latitude and is two full days longer at mid-latitudes, $\mathrm{T_{sid}}(45^{o}) \simeq 27$ days. This is known as the differential rotation of the photosphere (e.g., Owens and Forsyth, 2013).

Julius Bartels defined his calendar based on observations of daily solar and geomagnetic activity. Over long periods the geomagnetic recurrence rate is very close to 27 days. Bartels' rotations are exactly 27 days long and are counted from Feb 8, 1832. Different solar-related events as geomagnetic storms produced by CIRs are observed with a 27 - 28 day periodicity.

Richard C. Carrington determined the solar rotation rate by watching low-latitude sunspots in the 1850s. He defined a fixed solar coordinate system that rotates in a sidereal frame exactly once every 25.38 days. The synodic rotation rate varies a little during the year because of the eccentricity of the Earth's orbit; the mean synodic value is about 27.2753 days. Carrington Rotation 1 began at a seemingly arbitrary instant late on Nov 9, 1853, when Carrington began his Greenwich photo-heliographic series. The zero meridian used today is the one that passed through the ascending node of the solar equator on the ecliptic at Greenwich mean noon on January 1, 1854 (e.g., Carrington, 1863; Stanford, 2017).

2.6 Mathematical Tools

2.6.1 Principal Components Analysis

Principal Components Analysis (PCA) is a mathematical method used to find the directions of largest data variance, rotating the original axes of the data distribution to

these directions, reducing by this way the number of dimensions or the degrees of freedom required to represent the data (or dimensionality). The data is ordered in principal components (PC): the first PC contains the data with the largest variance, or the most meaningful part of data, also used to represent the largest part of data; the second PC contains the points with the second largest variance and second largest part of data, and so on. This allows to separate the most relevant information at the first PC and the less relevant at the lower PC, and makes it possible to model the data with less but meaningful information, as a sum of just the first PCs. In signal processing we can separate a well defined signal as the sum of the first PCs and the noise as the sum of the other PCs. In Geomagnetism, this method has been used under the name of Natural Orthogonal Component (NOC) decomposition (e.g., Xu and Kamide, 2004).

SVD theorem. SVD is a generalization of the eigen-decomposition for a rectangular matrix **H** (n x m), factorizing the matrix into two orthogonal matrices and a scaling matrix:

$$\mathbf{H} \;=\; \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}}, \tag{2.17}$$

where **U** (n x n) and **V** (m x m) are the orthonormal matrices and **S** (n x m) is a matrix with s_{ii} non negative values and zeros in all other elements. s_{ii} are named singular values, and are equal to the non zero square roots of $\mathbf{H}^T\mathbf{H}$ or $\mathbf{H}\,\mathbf{H}^T$ eigenvalues (ss_i), in descending order, from s_{11}, s_{22} until s_{nn} or s_{mm}, depending which one (n or m) is the smallest value: $ss_i = s_{ii}\; s_{ii} = s_{ii}^2$. Multiplication by **S** stretches **V** and multiplication by **U** produces the rotation of axis (e.g., Shlens, 2005).

Singular values refer to the percentage of contribution of each mode. To calculate the percentage of variability of each mode we have to square each s_{ii} and divide by the sum of them:

$$f_i = \frac{100\; ss_i}{\sum\limits_k ss_k} \quad . \tag{2.18}$$

Columns of $\mathbf{V}$ are called right singular vectors of $\mathbf{H}$, principal component loading patterns, modes of variability or empirical orthogonal functions (EOFs), and are orthonormal eigenvectors of $\mathbf{H}^T\mathbf{H}$ (e.g., Bjornsson and Venegas, 1997)

$$\begin{aligned} \mathbf{H^T H} &= \mathbf{V S U^T U S V^T} \\ \mathbf{H^T H} &= \mathbf{V S^2 V^T} \end{aligned} \tag{2.19}$$

$$\begin{aligned} \mathbf{H^T H V} &= \mathbf{V S^2 V^T V} \\ \mathbf{H^T H V} &= \mathbf{V S^2} \end{aligned} \tag{2.20}$$

Equation 2.19 is the eigen-decomposition (or spectral decomposition) of matrix $\mathbf{H}^T\mathbf{H}$ and Equation 2.20 is the eigenvalue equation. Both equations show that $\mathbf{V}$ is the matrix of eigenvectors and $\mathbf{S}^2$ is the diagonal matrix of eigenvalues of $\mathbf{H}^T\mathbf{H}$.

The columns of $\mathbf{U}$ are called left singular vectors of $\mathbf{H}$, EOF amplitudes, expansion coefficient time series, principal component time series (PCs), or EOF time series (e.g., Bjornsson and Venegas, 1997), and are orthonormal eigenvectors of $\mathbf{H}\,\mathbf{H}^T$. $\mathbf{U}$ is the matrix of eigenvectors of $\mathbf{H}\ \mathbf{H}^T$ and $\mathbf{S^2}$ is the diagonal matrix of eigenvalues. The column vectors of matrices $\mathbf{U}$ and $\mathbf{V}$ are orthogonal to each other. The columns of $\mathbf{V}$ represent the most meaningful patterns of the data as standing oscillations. The columns of $\mathbf{U}$ help to represent how these patterns evolve in time, as amplitudes or expansion coefficients (e.g., Bjornsson and Venegas, 1997)

$$\mathbf{a}_j \;=\; \mathbf{u}_j\, s_{jj} \;=\; \mathbf{H}\,\mathbf{v}_j, \tag{2.21}$$

where $\mathbf{u}_j$ is the j^{th} column of $\mathbf{U}$, s_{jj} is the corresponding singular value root, $\mathbf{v}_j$ is the j^{th} column of $\mathbf{V}$. In matrix notation:

$$\mathbf{A} \;=\; \mathbf{US} \;=\; \mathbf{HV}, \tag{2.22}$$

where **A** is a matrix n x m which columns are the expansion coefficients, or time series of the evolution of EOFs.

We can reduce the data **H** as a sum of the first k components:

$$\mathbf{H}_{model} = \sum_{j=1}^{k} (\mathbf{u}_j \; s_{jj} \; \mathbf{v}_j^T), \tag{2.23}$$

where j represents the j^{th} column of **U** and **V**, and s_{jj} is their corresponding singular value.

By construction, EOFs represent directions of variability with no particular amplitude, but by convenience they are chosen to have norm 1. As said before they are stationary structures, i.e., do not evolve in time. The principal component (PC) attached to the corresponding EOF provides the sign and the overall amplitude of the EOF as a function of time (expansion coefficients). This provides a simplified representation of the state of the field at that time along that EOF. In other words EOFs do not change structure in time, they only change sign and overall amplitude (e.g., Hannachi, Jolliffe, and Stephenson, 2007).

PCA Sampling Errors. To estimate the EOF's sampling errors, North et al. (1982) suggest a rule of thumb: if the sampling error of a particular eigenvalue is comparable to or larger than the separation between two neighbouring eigenvalues, then the sampling error for the associated EOF will be comparable to the size of the neighbouring EOF. The number of samples (n, the number of rows in **H**) required to resolve different eigenvalues will depend on the separation of neighbouring eigenvalues. If separation is small, sampling have to be increased. North et al. (1982) error for an eigenvalue is:

$$\sigma_{ss_i} \; = \; ss_i \sqrt{\frac{2}{n}} \tag{2.24}$$

and the sampling error on the associated EOF is:

$$\sigma_{EOFi} = \frac{\sigma_{ss_i}}{ss_i - ss_j} EOF_i \tag{2.25}$$

where ss_i and ss_j are two consecutive eigenvalues.

The percentage of variance explained by mode-i is given by Equation 2.18.

I applied the error propagation rule (Ku, 1966; *NIST/SEMATECH e-Handbook of Statistical Methods*)

$$\begin{aligned} \sigma_{f_i}^2 &= 100^2 \left(\frac{\partial (ss_i / \sum_k ss_k)}{\partial ss_i} \right)^2 (\sigma_{ss_i})^2 \\ \sigma_{f_i} &= 100\, \sigma_{ss_i} \sqrt{\left(\sum_k ss_k \right)^{-2} - 2ss_i \left(\sum_k ss_k \right)^{-3} + ss_i^2 \left(\sum_k ss_k \right)^{-4}} \end{aligned} \tag{2.26}$$

Steps to Perform a Singular Value Decomposition. These calculations were made in Matlab ©, following the steps listed below.

1. Arrange the data in a matrix, preferentially with the columns representing the random variables (24 hourly values in our analysis) and rows representing different epochs (in our analysis are days). The matrix must not contain NaN (acronym of "Not a Number") neither 99999.99 or empty spaces. For a matrix of n rows and m columns:

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} & \dots & \dots & H_{1m} \\ H_{21} & H_{22} & \dots & \dots & H_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ H_{n1} & H_{n2} & \dots & \dots & H_{nm} \end{bmatrix} \tag{2.27}$$

2. Calculate the mean of every column or every row. In the first analysis (Chapter 3) I calculated the daily mean, with the 24 hourly values of every row (m = 24). I obtained up to n mean values, one for each day. For the next step I arranged the n means in a

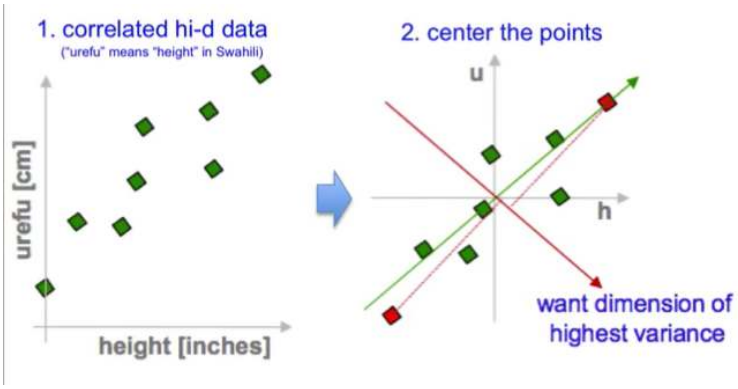


FIGURE 2.20: Centring the data by subtracting the mean. Image: /https://tinyurl.com/ybcfuekp

matrix n x m with the same daily mean value ($\overline{H}_i$) at each row:

$$\overline{\mathbf{H}} = \begin{bmatrix} \overline{H}_1 & \overline{H}_1 & \dots & \dots & \overline{H}_1 \\ \overline{H}_2 & \overline{H}_2 & \dots & \dots & \overline{H}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \overline{H}_n & \overline{H}_n & \dots & \dots & \overline{H}_n \end{bmatrix}, \tag{2.28}$$

3. Remove the daily mean to each observed H-hourly value by:

$$\mathbf{\Delta H} = \mathbf{H} - \overline{\mathbf{H}}. \tag{2.29}$$

This step is necessary to center the data into the original Cartesian axes (see Figure 2.20).

4. Use the Matlab© or Octave function "svd" to calculate the matrices **U**, **V** and **S**.

$$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{H}); \tag{2.30}$$

that performs the SVD decomposition (equation 2.17).

2.6.2 Statistical Quantities and Tests

Correlation is a statistical method used to assess the strength and direction of the relationship (not necessarily causal) between two variables. It is measured by a statistical parameter (a statistic) called the correlation coefficient, which is a dimensionless quantity that takes a value in the range -1 to +1. Correlation coefficients do not communicate information about wether one variable moves in response to another (e.g., Mukaka, 2012).

In this thesis work I used two kinds of correlation coefficients: Pearson's and Spearman's. The first one is useful when looking for a linear relationship between variables and when both variables are bivariate normally distributed. The second one is more suitable when the relationship between variables is not linear and they are not normally distributed. This latter coefficient is robust when extreme values (outliers) are present (e.g., Mukaka, 2012).

The distribution of Pearson's correlation coefficients may have a very skewed probability curve. I applied the Fisher's transformation to convert it into an approximately normal distribution. I also used Welch's t-tests of significance to evaluate if two normal distributions were significantly different.

Pearson's Correlation coefficient. For two series A and B with N_{tot} values each, the Pearson's correlation coefficient r is computed as:

$$r(A,B) = \frac{\sum_{\mathrm{i=1}}^{N_{tot}} \left(A^i - \overline{A}\right)\left(B^i - \overline{B}\right)}{\sqrt{\sum_{\mathrm{i=1}}^{N_{tot}} \left(A^i - \overline{A}\right)^2}\sqrt{\sum_{\mathrm{i=1}}^{N_{tot}} \left(B^i - \overline{B}\right)^2}} = \frac{\mathrm{cov}(A,B)}{\sigma_A \sigma_B}, \tag{2.31}$$

where $\overline{A}$ and σ_A are the mean and standard deviation of A, respectively, $\overline{B}$ and σ_B are the mean and standard deviation of B, and cov(A,B) is the covariance of A and B. $r = 0$ indicates that there is no linear relationship; $r = 1$ indicates that there is a

strong linear relationship between the variables; $r = -1$ indicates that there is a strong negative linear relationship between the variables.

The significance of each r coefficient is evaluated through a corresponding p-value, which tests the null hypothesis that r is zero. For a confidence interval (CI) of 95% the significance level is 0.05 (denoted as α or alpha). If the p-value is smaller or equal than $\alpha = 0.05$, the corresponding r is significantly different from zero. If the p-value is greater than 0.05, then we cannot conclude that the correlation is different from 0.

We used the Student's t-test to evaluate the minimum value of r that is significant at the 5% level. For a random sample of n pairs of values from two bivariate normal distributions used to calculate every r coefficient:

$$t = r\sqrt{\frac{n-2}{1-r^2}}, \tag{2.32}$$

where $n-2$ are the degrees of freedom. From a t-table we obtain the level of significance of 5% for $n-2$ degrees of freedom.

In Chapters 3 and 4, Pearson's correlation coefficients r will be computed to compare transient signals of magnetic observatories with corresponding TS05 simulations. As the obtained distributions of r values have a very skewed probability curve, the Fisher's transformation is applied to convert r histograms into approximately normal histograms. Only then can Welch's tests of significance be applied to evaluate if two r distributions are significantly distinct.

Fisher's transformation. The distribution of Pearson's correlation values is represented in the form of histograms normalized in order that the total area is one (i.e., see Figure 3.10). Most histograms are left-skewed with skewness increasing with $\bar{r}$. It is well known that when the absolute value of the correlation in a population is low (say less than about 0.4) then the sampling distribution of Pearson's r is approximately

normal. But with high values of correlation, the distribution has a negative skew.

The exact distribution probability density function (pdf) for r can be computed if the two populations follow a bivariate normal distribution, and has a complicated expression in terms of gamma and hypergeometric functions (see e.g., Weatherburn, 1961). Since the sampling distribution of Pearson's r is not normally distributed, the use of the standard error of r is not recommended. We can transform the skewed distributions to normal distributions using the Fisher's transformation for r (hereafter, z), tending rapidly to normality as the size of the sample (n) increases. The formula for Fisher's transformation of r is:

$$z = 0.5 \ln [(1 + r)/(1 - r)], \tag{2.33}$$

It has a variance equal to $\frac{1}{n-3}$ and a standard error of $\frac{1}{\sqrt{n-3}}$, that is independent of the value of r.

If we apply the inverse function of Fisher's transformation to the mean of each fitted Gaussian curve, we obtain a new parameter r_0 that takes into account the whole histogram:

$$r_0 = \frac{e^{2\bar{z}} - 1}{e^{2\bar{z}} + 1} = \tanh(\bar{z}), \tag{2.34}$$

where $\bar{z}$ is the mean of the Gaussian curve fitted to the z-distribution.

Welch's t-test. To test if two normal distributions are significantly different, a Welch's t-test can be applied to them. The t-statistic parameter to test the hypothesis that two means $\bar{z}_1$ and $\bar{z}_2$ from two different sets are the same is

$$t = \frac{\bar{z}_1 - \bar{z}_2}{\sigma_{1,2}} \tag{2.35}$$

where $\bar{z}_1$ and $\bar{z}_2$ are the mean values of the Gaussians fitted to the histograms of z and

$$\sigma_{1,2} = \sqrt{\frac{\sigma_{z_1}^2}{N_1} + \frac{\sigma_{z_2}^2}{N_2}} \tag{2.36}$$

where σ_{z_1} and σ_{z_2} are the fitted standard deviations and N_1 and N_2 are the number of z-values to compute each histogram. The number of degrees of freedom ($d.f.$) required to compute Student's t-distribution is given by

$$d.f. = \frac{\sigma_{1,2}^4}{\frac{1}{N_1-1}\left(\frac{\sigma_{z_1}^2}{N_1}\right)^2 + \frac{1}{N_2-1}\left(\frac{\sigma_{z_2}^2}{N_2}\right)^2} \tag{2.37}$$

The null hypothesis is that the two distributions are the same. The p-value gives the confidence in our decision, and is calculated from the Student's t-distribution probability density function (pdf):

$$p = 1 - F(x|d.f.) = 1 - \int_{-\infty}^{x} \frac{\Gamma\left(\frac{d.f.+1}{2}\right)}{\sqrt{(d.f.)\pi}\,\Gamma\left(\frac{d.f.}{2}\right)\left(1+\frac{t^2}{d.f.}\right)^{\frac{d.f.}{2}}} dt \tag{2.38}$$

where F is the cumulative distribution function and

$$\Gamma(d.f.) \;=\; \int_0^\infty e^{-x} x^{d.f.-1} dx \tag{2.39}$$

is a convergent improper integral defined for all complex numbers except the non-positive integers, called the Gamma function (see e.g., Weatherburn, 1961). If $d.f.$ is a positive integer, $\Gamma(d.f.) \;=\; (d.f.-1)!$

We compare the t-statistic with a t-critical value, which is computed from the inverse of Student's t cumulative distribution function (cdf):

$$t\text{−critical} = F^{-1}(1-\alpha|d.f.), \tag{2.40}$$

where α is the significance level. For a confidence interval (CI) of 95%, $\alpha = 0.05$.

The criteria is:

- if t < t-critical, then p > 0.05 and we can not reject the null hypothesis (the two distributions are the same),
- if t $\geq$ t-critical, then p ≤ 0.05 and we consider the two distributions are different.

Spearman's correlation coefficients. Spearman's correlations are preferable to Pearson's correlations when we are not sure if two variables have a linear relation neither a normal distribution. It is a non parametric statistic because does not make assumptions about the probability distributions of the variables being assessed (e.g., it is not necessary that distributions be normal). As occurs with some of the parameters compared in this work, a variable some times increases and other times decreases while the other only increases. This is called a non-monotonic relation. When a variable only increases or only decreases while the other increases it is a monotonic relation (whether linear or not). Spearman's correlation is a statistical measure of the strength of a monotonic relationship between two series of data. It is denoted as r_s and has values between -1 and 1. Values nearest to ± 1 mean a stronger monotonic relation. A value of zero means that there is no monotonic association. Spearman's correlation is defined as the Pearson's correlation coefficient between the ranked values of data. Ranking (from low to high) is obtained by assigning a rank of 1 to the lowest value, 2 to the next lowest and so on. The Spearman's correlation is less sensitive than the Pearson's correlation to strong outliers that are in the tails of both samples (Mukaka, 2012; Szmidt and Kacprzyk, 2011). For a sample of size n, the n raw scores X_i, Y_i are converted to ranks $rg\ X_i$, $rg\ Y_i$ and r_s is computed from:

$$r_s = \frac{\text{cov}(rgX, rgY)}{\sigma_{rgX}\sigma_{rgY}} = r(rgX, rgY) \tag{2.41}$$

where r denotes the usual Pearson's correlation coefficient, but applied to the ranked variables (see equation 2.31), $\mathrm{cov}(rgX, rgY)$ is the covariance of the ranked variables, σ_{rgX} and σ_{rgY} are the standard deviations of the ranked variables. If all n ranks are distinct integers, it can be computed using the formula (Mukaka, 2012):

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{2.42}$$

where $d_i = rg(X_i) - rg(Y_i)$, is the difference between the two ranks of each observation and n is the number of observations.

The Spearman's correlation coefficient is used in Chapter 5 to compare more uneven data, as solar structures with interplanetary parameters, geomagnetic indices and coupling functions.

2.6.3 Locally Weighted Scatter Plot Smoothing (loess)

All parameter's series were plotted with superposed loess smoothing functions. This smoothing function uses locally weighted linear regression with a second degree polynomial to smooth data. The local regression smoothing process follows three steps for each data point:

1. Compute the regression weights for each data point in the span using a cubic function:

$$w_i = \left(1 - \left[\frac{x - x_i}{d(x)}\right]^3\right)^3 \tag{2.43}$$

where x is the predictor value associated with the response value to be smoothed, x_i are the nearest neighbours of x as defined by the span, and $d(x)$ is the distance along the abscissa from x to the most distant predictor value within the span. The data point to be smoothed has the largest weight and the most influence on the fit, and data points outside the span have zero weight and no influence on the fit.

2. Perform a weighted linear least-squares regression with a second degree polynomial.

3. The smoothed value is applied by the weighted regression at the predictor value.

The smooth calculation involves the same number of neighbouring data points (span), but if the number of neighbouring points is not symmetric about the smoothed data point, the weight function is not symmetric.

3 Geomagnetic Activity: Data, Indices and Simulations

3.1 Overview

In this chapter a statistical analysis of the horizontal geomagnetic component (H) measured at four geomagnetic observatories is performed. These series are compared with the disturbance storm-time (Dst) and ring current (RC) geomagnetic indices and H-series synthesized from the Tsyganenko and Sitnov [2005] (TS05) semi-empirical model of magnetospheric currents. Analysed data is for the 2007 to 2014 period, from Coimbra, Portugal; Panagyurishte, Bulgary; Novosibirsk, Russia and Boulder, USA, at geomagnetic latitudes between 40° and 50° N. This time period includes the solar activity minimum in 2009. The quiet daily (QD) variation is firstly removed from series of data, using Principal Components Analysis (PCA) tools. Our results show that PCA models are as efficient as standard models in removing the regular daily variation with the advantage to reduce complexity. For the two stations in Europe, they seem suited to separate ionospheric from magnetospheric contributions. Dst and RC series correlate well with the H local component at each observatory, with values for the mean of significant correlation coefficients, $\bar{r}$, from 0.5 to 0.6 during low geomagnetic activity and from 0.6 to 0.7 for geomagnetic active days. The four observatories separate into two groups: Coimbra and Panagyurishte, for which the magnetospheric/ionospheric ratio in QD variation is smaller, a dominantly QD ionospheric contribution can be removed

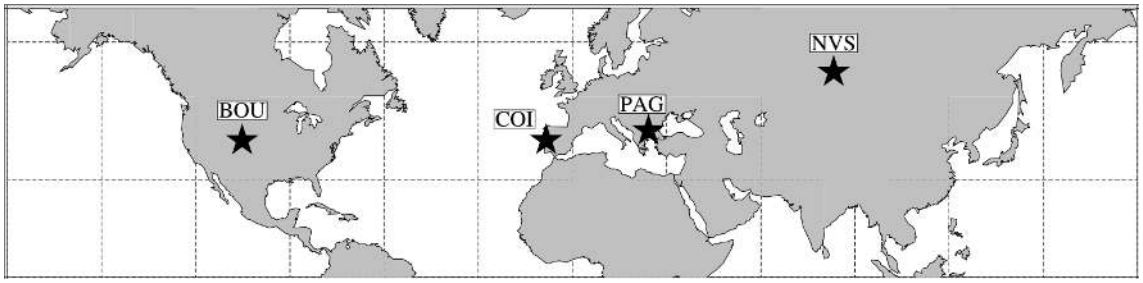


FIGURE 3.1: Map of observatories locations.

and TS05's H-series are the best proxy; Boulder and Novosibirsk, for which the ionospheric and magnetospheric contributions in QD variation can not be separated and correlations with TS05's H-series deteriorate with the QD removal.

3.2 Data and Simulations

In this chapter and the following one (Chapter 4), I focus in geomagnetic activity at mid-latitude stations. Four geomagnetic stations were selected with similar geomagnetic latitudes but quite separated in longitude around the globe, also with good quality and availability of data: Coimbra (COI) in Portugal, Panagyurishte (PAG) in Bulgary, Novosibirsk (NVS) in Russia and Boulder (BOU) in USA. Table 3.1 shows the four magnetic stations used in this work, listed by their IAGA (International Association of Geomagnetism and Aeronomy) code, their geographic and geomagnetic coordinates, calculated using the IGRF-12 model for 2010.0 and the shifts from Universal Time (UT) to Local Time (LT). Figure 3.1 shows the global distribution of stations.

Local K indices were used to separate data in more active days and more quiet days (see Appendix A). COI K indices were computed [1] following the FMI method (from Finnish Meteorological Institute, e.g., Sucksdorff, Pirjola, and Hakkinen, 1991). I selected as active days those with at least one $K \geq 4$ (from eight in a day). The rest of days were classified as calm days (e.g., Shinbori et al., 2014). Results will be represented separately for calm and active days.

[1] Paulo Ribeiro is kindly acknowledged for computing the COI K indices.

TABLE 3.1: List of stations, their geomagnetic and geographic coordinates and UT to LT conversion.

Station	Geographic Coord.		Geomagnetic Coord.		LT
	Lat. (°N)	Long. (°E)	Lat. (°N)	Long. (°E)	
COI	40.2	351.6	43.8	72.1	UT+0
PAG	42.5	24.2	40.5	105.1	UT+2
NVS	54.9	83.2	45.4	156.0	UT+6
BOU	40.1	254.8	48.1	321.3	UT-7

The five or six international quietest days of each month (hereafter, IQDs) were used to build the local QD variation series. The list of IQDs is issued by the GFZ German Research Center for Geosciences (see Table 3.2), based on lowest values of the planetary geomagnetic disturbance index Kp, on the basis of three criteria that rank the days of each month (e.g. Menvielle et al., 2011): the sum of the eight Kp values, the sum of squares of the eight Kp values and the maximum of the eight Kp values. A relative order number is assigned to each day of the month, the three order numbers are averaged and the five days with the lowest mean order numbers are selected as the five quietest days (for more information see /https://tinyurl.com/y8fofs5x).

Table 3.2 shows databases and people responsible of data and indices used in this work, separated by magnetic station, and by geomagnetic components, K index, international quietest days, Dst index, RC index and spacecraft data. All geomagnetic components are obtained with 5 minute resolution and later are binned to 1 hour resolution.

From series of X and Y components for each observatory, H component series were calculated as $H = \sqrt{X^2 + Y^2}$. Both Dst and RC indices are series of 70128 hourly values corresponding to January 1^{st} 2007 until December 31^{st} 2014. These two last series do not have missing values.

Series with the TS05 model predictions for all field components were obtained using the TS04c code. They were called BX, BY and BZ series. For every station and every current system, the code has to be adapted by specifying the local geodetic coordinates, the altitude over sea level of each station, a flag number that represents the current field to be generated (detailed in Appendix A), and adding commands to read the 11 parameters of the input file (see section 2.4). The output file contains 5-min resolution BX, BY and BZ, values available at the time interval selected with their corresponding date and time. To calculate the main field using the model IGRF-12 we use the GEOPACK-2008 subroutines with the geodetic coordinates of each station. The BH component series are calculated later as $BH = \sqrt{BX^2 + BY^2}$. For more details, see Appendix A.

Missing values and manual interpolations. TS05's series and ground data hold some missing values for the analysed period. For observatory series, linear interpolation of hourly values at a certain hour (i) is made using values for that same hour of previous $j-1$ and next $j+1$ days, i.e., $X^{i,j} = (X^{i,j+1} + X^{i,j-1})/2$ (the same for Y). Whenever I did not have the hourly value at previous or next day I used the value at previous $(i-1)$ or next hour $(i+1)$ of the same day j, for instance, $X^{i,j} = (X^{i-1,j} + X^{i,j-1})/2$. This procedure filters out rapid variations during disturbed time, when the mid-latitude magnetic field changes on time-scales of hours. However, the number of interpolated hourly values is very small, below 0.2 % (see Table 3.3) and not expected to significantly affect our results. If after this procedure a day still has at least one missing hourly value, the whole day is discarded from our analysis. Figures 3.2 and 3.3 show the X, Y, Z and H components both for observatory data and TS05 magnetospheric synthetic series, with the final percentage of missing data. Figure 3.4 shows the Dst and RC series for the period 2007-2014.

TABLE 3.2: Sources of data and indices by station, as well as names of persons in charge.

COI	Database	Internet address and e-mail
X, Y, Z	World Data Centre for Geomagnetism (Edinburgh). Data Publisher for Earth and Environmental Science (PANGAEA).	/https://tinyurl.com/y9szl4js /https://tinyurl.com/yd54c35l Paulo Ribeiro: pribeiro@ci.uc.pt
K	Data Publisher for Earth and Environmental Science (PANGAEA).	/https://tinyurl.com/yd54c35l Paulo Ribeiro: pribeiro@ci.uc.pt
PAG		
X, Y,Z	INTERMAGNET.	/http://www.intermagnet.org/
K	Annual Report of the Observed Geomagnetic Activity in Panagyurishte Observatory. National Institute of Geophysics, Geodesy and Geography, Bulgarian Academy of Sciences.	/https://tinyurl.com/y9bgdl7z Petya Trifonova: p.trifonova@abv.bg
NVS		
X, Y, Z	INTERMAGNET.	/http://www.intermagnet.org/
K	Geophysical Observatory Klyuchi, Novosibirsk, Russia.	Olga Fedotova: baley@ngs.ru
BOU		
X, Y, Z	INTERMAGNET.	/http://www.intermagnet.org/
K	NOAA/Space Weather Prediction Center. Institut de Physique du Globe de Paris. Observatoire Magnetique National.	/https://tinyurl.com/yayx5maf Benoit Heumez: heumez@ipgp.fr
IQD	GFZ German Research Center for Geosciences	/www.gfz-potsdam.de
Dst	World Data Center for Geomagnetism, Kyoto	/https://tinyurl.com/yb5prhzo
RC	DTU Space, National Space Institute, Technical University of Denmark.	Christopher Finlay: nio@space.dtu.dk
OMNI data	OMNI/NASA high resolution database	/https://tinyurl.com/6prz7p8
TS05's data	Tsyganenko and Sitnov TS05_Data_and_Stuff Website	/https://tinyurl.com/y8mmfwws

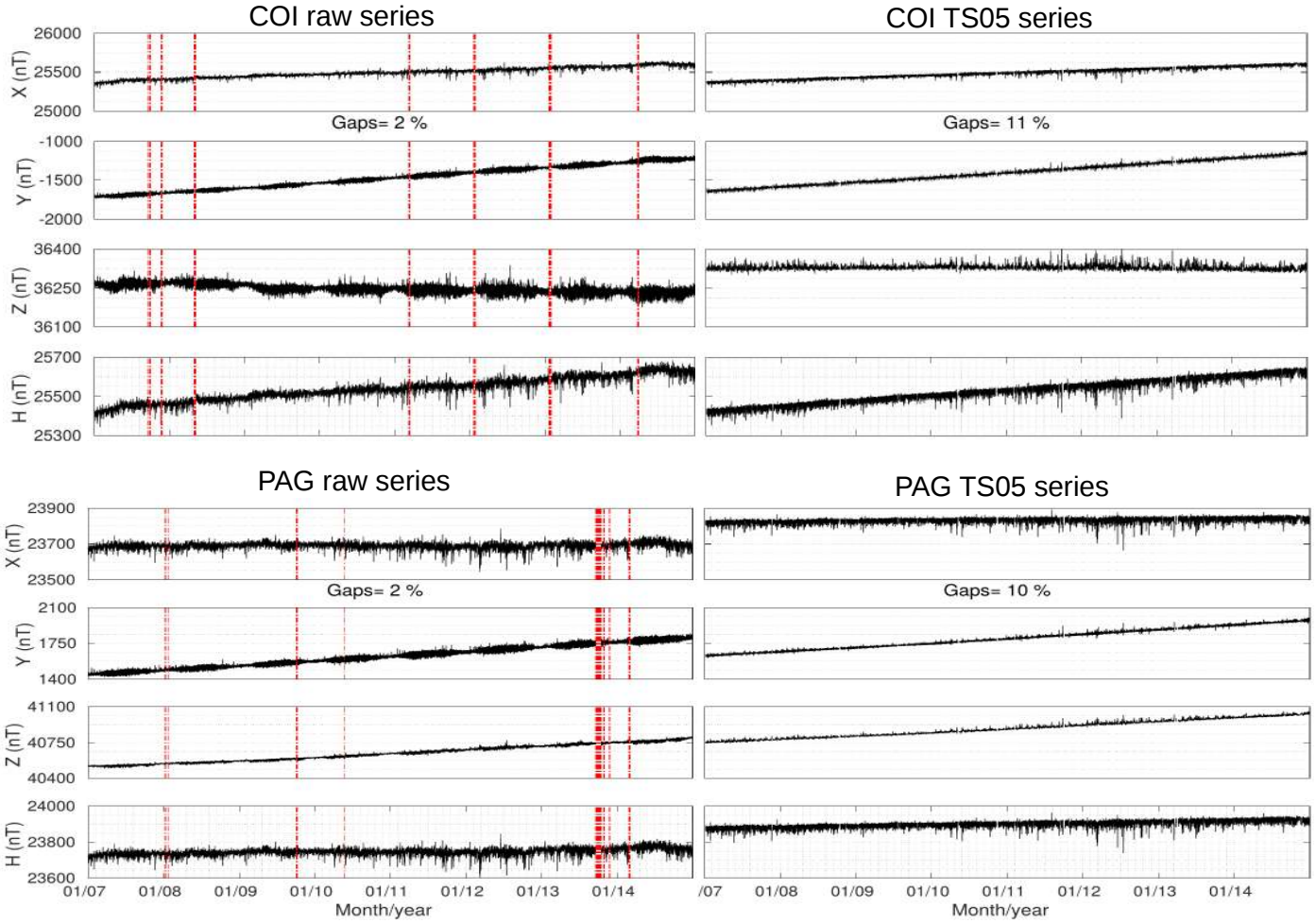


FIGURE 3.2: Top-left quarter, from top to bottom: X, Y, Z and H series of COI observatory data, from January 1^{st} 2007 to December 31^{st} 2014. Vertical lines represent missing data positions. Top-right quarter: same as before, but for TS05 magnetospheric synthetic series calculated for COI geographic location, indicating the percentage of gaps. Bottom-left quarter and bottom-right quarter: same as two previous, but for PAG observatory.

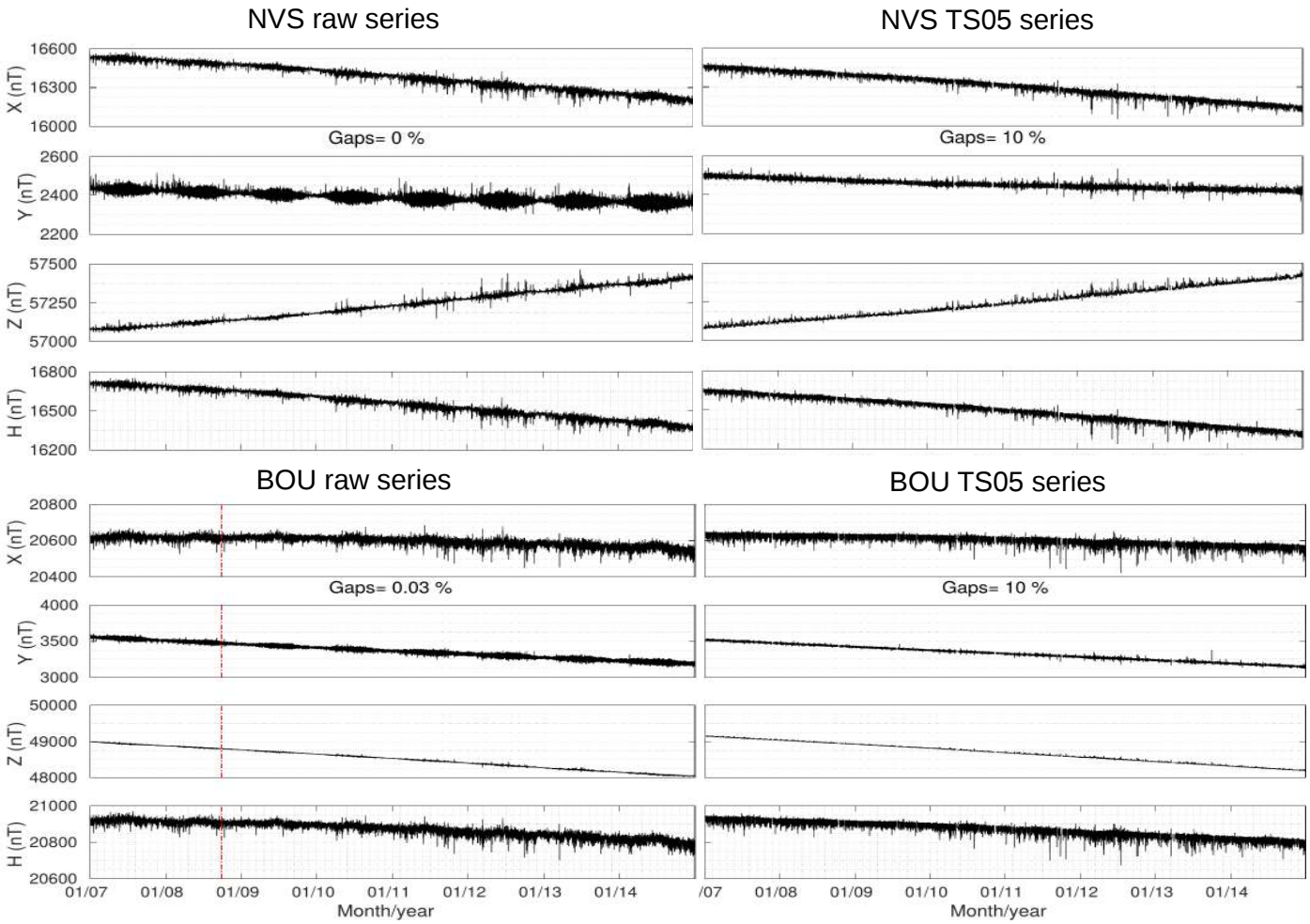


FIGURE 3.3: Top-left quarter, from top to bottom: X, Y, Z and H series of NVS observatory data, from January 1^{st} 2007 to December 31^{st} 2014. Top-right quarter: same as before, but for TS05 magnetospheric synthetic series calculated for NVS geographic location, indicating the percentage of gaps. Bottom-left quarter and bottom-right quarter: same as two previous, but for BOU observatory. Vertical lines represent missing data positions.

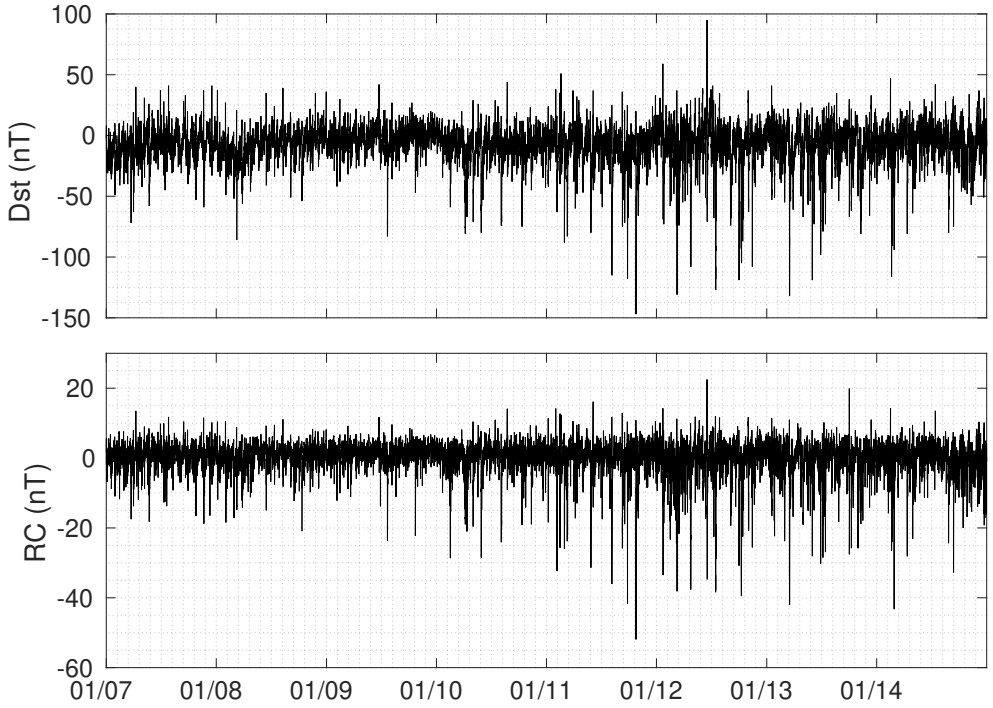


FIGURE 3.4: Top: Dst series from January 1^{st} 2007 to December 31^{st} 2014. Bottom: RC series for the same period of time.

TABLE 3.3: Number of days analysed and percentage of hourly values interpolated by station. Results are shown separately for quiet and active days.

Station	No. of days		Interpolations (%)	
	quiet	*active*	*quiet*	*active*
COI	1883	483	0.20	0.08
PAG	1846	532	0.20	0.07
NVS	2003	425	0.06	0.01
BOU	1883	537	0.09	0.10

3.3 Quiet Daily Variation: Separation of Ionospheric and Magnetospheric Contributions

The TS05 model does not include the ionospheric contribution neither induced fields from magnetospheric and ionospheric origin. Before comparing observatory series with series predicted by TS05, I removed the QD variation from raw data. By removing QD, I expect to remove a significant parcel of the ionospheric primary and secondary contributions.

The QD has more or less recognizable patterns along the day, the year and the solar cycle. The perturbations of these patterns can be considered as irregularities or noise of the signal. Using a large distribution of data we can be able to separate the signal (QD pattern) from the noise (perturbations) by averaging out the noise or using the Singular Value Decomposition (SVD) method and its traditional use in Principal Components Analysis (PCA).

Two methods were used to calculate monthly models of QD variation:

a) the standard method, using the hourly values of X, Y and H at the five international quietest days (IQD) of every month, from January 2007 to December 2014,

b) the Principal Components Analysis method (PCA). This method is expected to concentrate the most important contribution due to ionospheric currents in the first modes of variability, and relegate to higher modes a contamination by magnetospheric activity that is always present, even in quiet days.

Let $QX^{i,j,k}$ ($QY^{i,j,k}$ and $QH^{i,j,k}$) be the series of 24 hourly values i of geomagnetic field component X (Y and H) as observed at a certain station during quietest day j ($j = Q1k, Q2k, Q3k, Q4k, Q5k$) in month k ($k = 1, ..., 96$), with $k = 1$ for January 2007 and $k = 96$ for December 2014. $Q1k$, $Q2k$, ... are the set of five (or less if data is missing) quietest days for month k. Let $\widetilde{QX}^{i,j,k}$ ($\widetilde{QY}^{i,j,k}$ and $\widetilde{QH}^{i,j,k}$) be the daily variation series obtained

from daily raw series by removing the daily mean $\overline{QX}^{j,k} = (1/24)\sum_{i=1}^{24} QX^{i,j,k}$, i.e.,

$$\widetilde{QX}^{i,j,k} = QX^{i,j,k} - \overline{QX}^{j,k} \quad . \tag{3.1}$$

A maximum of 480 standard QD series can be derived, from five quietest series for month (corresponding to 5 IQD x 12 months x 8 years), but for some quietest days, data are not available.

Averaged QD series for each month k, ΔQX (ΔQY and ΔQH), (96 in all) are computed from the 480 series (or less) as:

$$\Delta QX^{i,k} = \frac{1}{5}\left(\widetilde{QX}^{i,Q1k,k} + \widetilde{QX}^{i,Q2k,k} + \widetilde{QX}^{i,Q3k,k} + \widetilde{QX}^{i,Q4k,k} + \widetilde{QX}^{i,Q5k,k}\right) \quad . \quad . \tag{3.2}$$

ΔQX (ΔQY and ΔQH) is subtracted from the raw observatory daily series, yielding $\widetilde{X}$ ($\widetilde{Y}$) series:

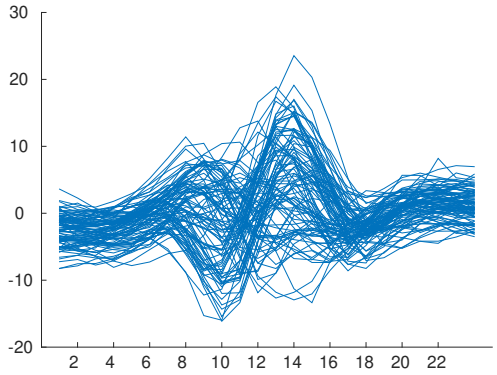
$$\widetilde{X}^{i,j,k} = X^{i,j,k} - \Delta QX^{i,k} \tag{3.3}$$

where now j covers all days used in this analysis, with a maximum of 2922. For H,

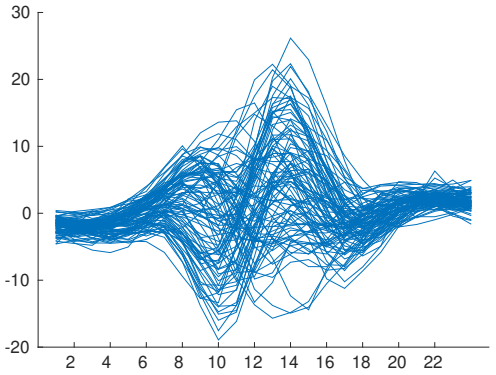
$$\widetilde{H}^{i,j,k} = H^{i,j,k} - \Delta QH^{i,k} \tag{3.4}$$

In a first calculation the daily mean value was used as the zero line (or datum line) for the quiet daily variation ΔQX, ΔQY and ΔQH, for the sake of simplicity (see e.g. Xu and Kamide, 2004). However, a number of authors use local near-midnight values as a datum line (three hours before and three hours after midnight) because of a minimal current in the night-time ionosphere (e.g., Matsushita and Xu, 1982; Campbell and Schiffmacher, 1985). For a comparison of magnitudes, different datum lines introduce biases of less than 10 nT, as can be seen in Figures 3.5 and 3.6, for COI and BOU respectively.

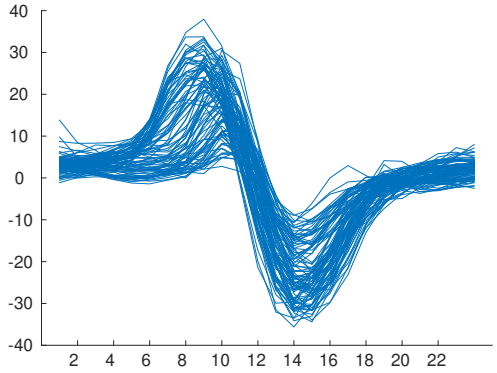
The general disturbance level at a given observatory changes along the year and



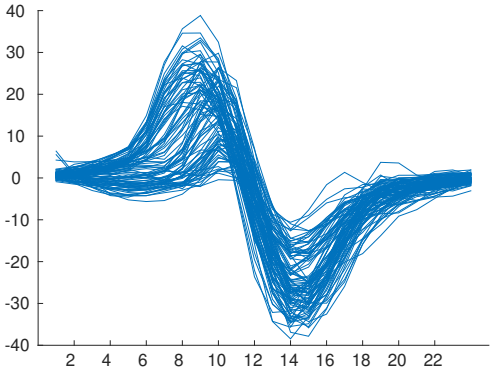
(A) ΔQX_{COI} without daily mean



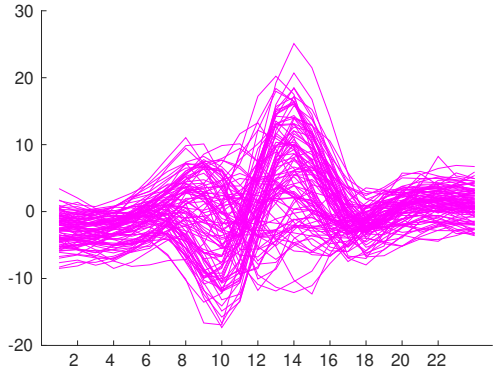
(B) ΔQX_{COI} without night-hours mean



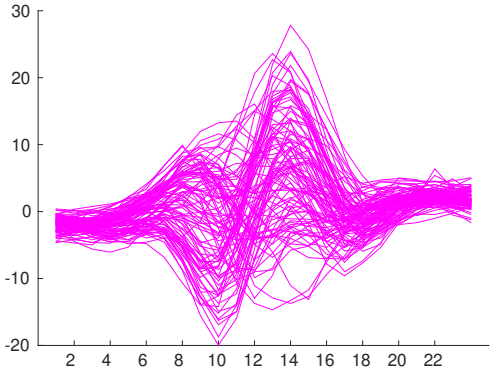
(C) ΔQY_{COI} without daily mean



(D) ΔQY_{COI} without night-hours mean

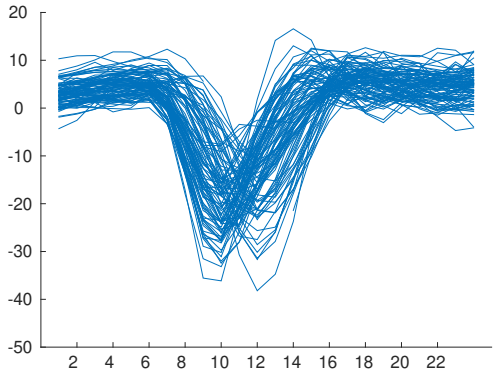


(E) ΔQH_{COI} without daily mean

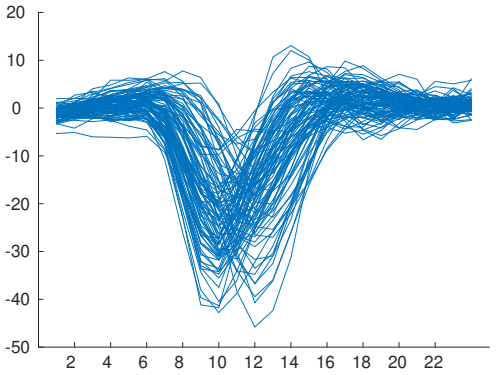


(F) ΔQH_{COI} without night-hours mean

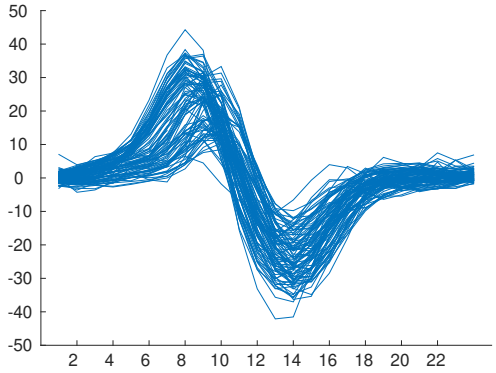
FIGURE 3.5: Left column, from top to bottom: 96 ΔQ series for COI X, Y and H respectively, calculated with Equations 3.1 and 3.2. Right column, from top to bottom: 96 ΔQ series for COI X, Y and H respectively, replacing daily mean with night-hours mean at Equation 3.1. Vertical axis shows variations in nT and horizontal axis is local time in hours. Note the similarity among ΔQX and ΔQH, due to the dipolar nature of the main field.



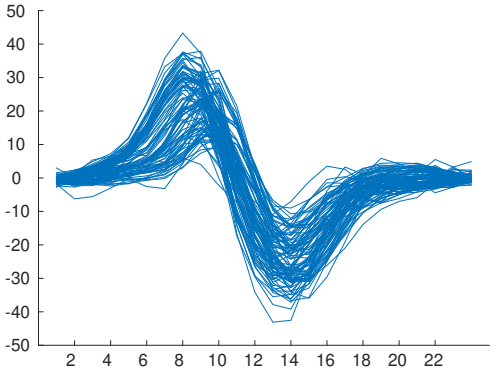
(A) ΔQX_{BOU} without daily mean



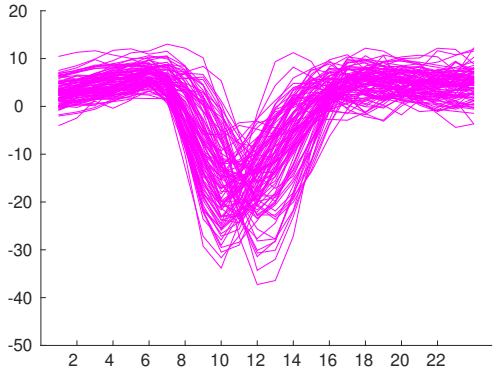
(B) ΔQX_{BOU} without night-hours mean



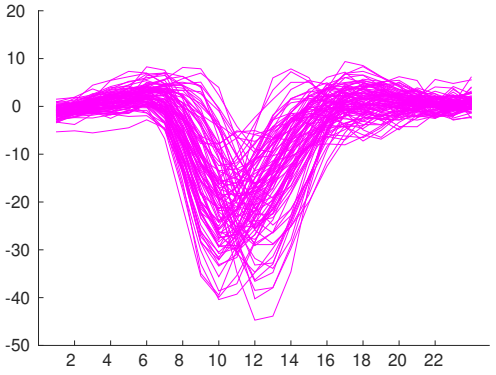
(C) ΔQY_{BOU} without daily mean



(D) ΔQY_{BOU} without night-hours mean



(E) ΔQH_{BOU} without daily mean



(F) ΔQH_{BOU} without night-hours mean

FIGURE 3.6: Left column, from top to bottom: 96 ΔQ series for BOU X, Y and H respectively, calculated with Equations 3.1 and 3.2. Right column, from top to bottom: 96 ΔQ series for BOU X, Y and H respectively, replacing daily mean with night-hours mean at Equation 3.1. Vertical axis shows variations in nT and horizontal axis is local time in hours. Note the similarity among ΔQX and ΔQH, due to the dipolar nature of the main field.

along the solar cycle, with highest values during summer and close to maxima of solar activity (see Figures 3.2 and 3.3). Even the quietest days in certain (more active) months have a significant contribution from irregular geomagnetic activity (Yamazaki and Maute, 2016) and, by removing ΔQX, ΔQY and ΔQH curves from raw data, a certain fraction of magnetospheric signal is also removed. Another reason why it is difficult to disentangle ionospheric and magnetospheric contributions is that, because they are all driven by the Sun, they all show a daily variation in a geographic coordinate system (e.g., Olsen, 1996; Laundal and Richmond, 2016).

The TS05 model has a synthetic daily variation given by $\sin\phi'$ and $\cos\phi'$ terms, due to FAC and PRC current systems respectively (see section 2.4). The geomagnetic longitude ϕ' is very close to the Local Time at mid-latitude stations. As a result, there also exists a (magnetospheric) QD variation in TS05 series.

For the synthetic TS05 series BX, BY and BH, the magnetospheric QD is calculated using the same procedure as in Equations 3.1, 3.2 and 3.3. The synthetic series computed from the mean of the IQDs of each month are called ΔQBX and ΔQBY. We call the corrected series $\widetilde{BX}$ and $\widetilde{BY}$ for the total magnetosphere current series and $\widetilde{BX}_{cur}$ and $\widetilde{BY}_{cur}$ for each separated current:

$$\widetilde{BX}^{i,j,k} = BX^{i,j,k} - \Delta QBX^{i,k} \tag{3.5}$$

$$\widetilde{BX}^{i,j,k}_{\text{cur}} = BX^{i,j,k}_{\text{cur}} - \Delta QBX^{i,k}_{\text{cur}} \tag{3.6}$$

The statistical tests in this chapter concern the H-component. The X and Y components will be studied in the next chapter. I must say that the QD signal was not removed from BH. The reason was that PCA applied to the data series was expected to separate ionospheric and magnetospheric contributions for QD. In this case it would be possible to remove only the ionospheric QD from data and keep the magnetospheric

QD in both data and simulations.

In the case of the QD curve for February 2011 the OMNI database did not have data for the four quietest days, and the Tsyganenko subroutines gave abnormal values at the fifth quietest day. In this case we used only the sixth international quietest day to compute the quiet-daily curve of this month, both for observed and synthetic series.

To compare observatory data with Dst and RC series we also removed from observatory H-series the corresponding IGRF-12 main field model (H_{OBS}^{IGRF}) calculated at each observatory location with Tsyganenko's GEOPACK-2008 subroutines, because both proxies do not have the IGRF incorporated:

$$\Delta H_{OBS} = H_{OBS} - H_{OBS}^{IGRF} \tag{3.7}$$

or

$$\Delta \widetilde{H}_{OBS} = \widetilde{H}_{OBS} - H_{OBS}^{IGRF} \tag{3.8}$$

where OBS refers to the observatory name and $H_{OBS}^{IGRF} = \sqrt{\left(X_{OBS}^{IGRF}\right)^2 + \left(Y_{OBS}^{IGRF}\right)^2}$.

As the main field model IGRF-12 is included in BX, BY and BH series through the use of GEOPACK-2008 subroutines, no baseline correction of the observed series is required when comparing them with TS05's simulations.

PCA Quiet Daily Variation Removal. Previous studies as Xu and Kamide (2004), Chen (2007), and De Michelis, Tozzi, and Consolini (2010) have already applied PCA to a set of hourly values of H during a certain time period. However, in those studies, data series for all days were used (instead of only quiet days) and the main goal was to separate in each 24-hours series the Sq ionospheric signal from geomagnetic disturbance components associated with different magnetospheric sources. I tested this approach with the data and verified that there is always a significant disturbance signal leaking into any of the computed modes, and that I could not isolate a quiet day mode. Then I decided to use only the previous QD matrix to attempt to separate the

magnetospheric and ionospheric signals, with the lowest modes as ionospheric and the highest modes as magnetospheric signal.

In this Chapter, the PCA is used on the H component. It consists in a singular value decomposition of the $\widetilde{\mathbf{QH}}$ matrix with 480 rows (or less), corresponding to the total number of quietest days available in a particular station and with 24 columns corresponding to the 24 hours of the day. The matrix can not contain neither NaNs nor 9999.99 or empty spaces. If I did not have the 24 hourly values of one of the quietest days, that day was not included. The SVD decomposes the original matrix into three matrices:

$$\widetilde{\mathbf{QH}} = \mathbf{USV^T} \tag{3.9}$$

where:

- $\widetilde{\mathbf{QH}}$ is the data matrix of quietest days, with size 480 (at most) x 24.
- **U** is a matrix 480 x 480 or less, with the normalized eigenvectors of the matrix $\widetilde{\mathbf{QH}}\widetilde{\mathbf{QH}}^{\mathbf{T}}$, which columns are called *left singular vectors* of $\widetilde{\mathbf{QH}}$.
- **V** is a matrix 24 x 24, with the normalized eigenvectors of the matrix $\widetilde{\mathbf{QH}}^{\mathbf{T}}\widetilde{\mathbf{QH}}$, which columns are called *right singular vectors* of $\widetilde{\mathbf{QH}}$. Figure 3.7(a) shows the first four EOFs for the four observatories, from top to bottom, COI, PAG, NVS and BOU.
- **S** is a matrix 480 (at most) x 24, that contains the 24 singular values of $\widetilde{\mathbf{QH}}$, at the diagonal of the matrix, s_{ll} (see Figure 3.8).

The expansion coefficients or principal components scores (see Section 2.6.1) represent the unstandardised coordinates of the rows of data along the principal axes, and were calculated by **A** = **US**, which size is 480 (at most) x 24. Each column represents the variations in amplitude of each mode along the ∼480 quietest days.

To reconstruct the original data we calculate separately every $mode - l$ (l=1,...,24) as:

$$\widehat{\mathbf{QH}}_{PCA-l} = \mathbf{a}_l \mathbf{v}_l^T \tag{3.10}$$

where $\mathbf{a}_l = \mathbf{u}_l s_{ll}$ (see eq. 2.21). Every $\widehat{\mathbf{QH}}_{PCA-l}$ ($mode - l$) is a matrix of 480 rows (at most) x 24 columns that represents a f_l percentage of the signal (see eq. 2.17). The sum of the 24 modes returns the original data, but we only need two or three modes for the PCA-QD model, because these represent around $\sim 80\%$ of the original signal (see Figure 3.8). The sum of these few modes gives what I call the $\widehat{\mathbf{QH}}_{PCA-QD}$ matrix.

Once the $\widehat{\mathbf{QH}}_{PCA-QD}$ matrix is complete, 96 average curves were calculated, one for each month, from the mean of the five (or less) $\widehat{\mathbf{QH}}^{i,Qik,k}_{PCA-QD}$ curves of each month.

Once subtracted from the raw observatory daily series, they yield the $\widehat{\mathbf{H}}$ series corrected by PCA-QD:

$$\widehat{H}^{i,j,k} = H^{i,j,k} - \Delta QH^{i,k}_{PCA-QD} \tag{3.11}$$

3.4 Results

PCA-QD Analysis. For each observatory and for each month, standard QD (see eq. 3.2) and PCA-QD models (see previous section) for component H were computed. Figure 3.7 (a) shows the first three EOFs for each station (from top to bottom, COI,PAG, NVS and BOU). Figure 3.7 (b) shows the average of ΔQH_{PCA-QD} models for each Lloyd's season: D season in blue (January, February, November and December), E season in black (March, April, September and October) and J season in red (May, June, July and August). Figure 3.7 (c) shows the residuals between the standard ΔQH and the ΔQH_{PCA-QD} for different Lloyd's seasons. The first four singular values for every observatory are represented in Figure 3.8.

A plot with the first two columns of matrix $\mathbf{A}$ ($\mathbf{a}_1$ and $\mathbf{a}_2$, see equation 2.22) of each station is shown in Figure 3.9. Both show a clear annual variation. This is specially

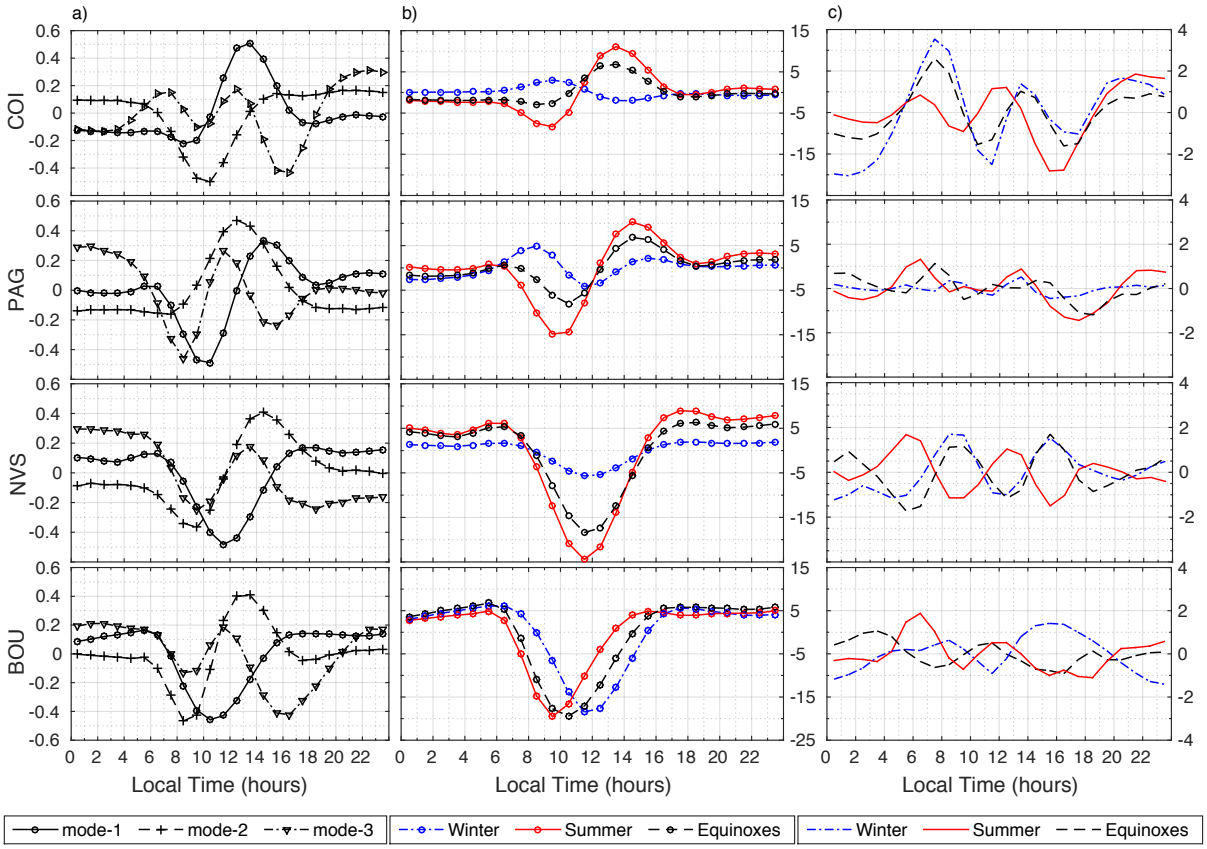


FIGURE 3.7: a) First three EOFs for H from the PCA analysis; b) Mean PCA-QD variation curves of H during the 2007-2014 period, for D season (blue), E season (black), J season (red); values are in nT. c) Residuals between standard QD and PCA-QD models, in nT.

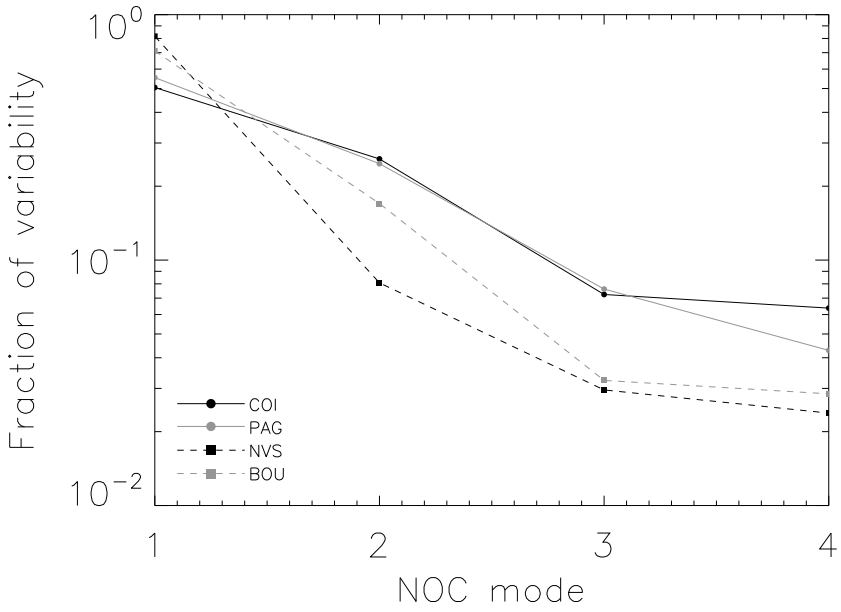


FIGURE 3.8: First four singular values of PCA-QD model of every station.

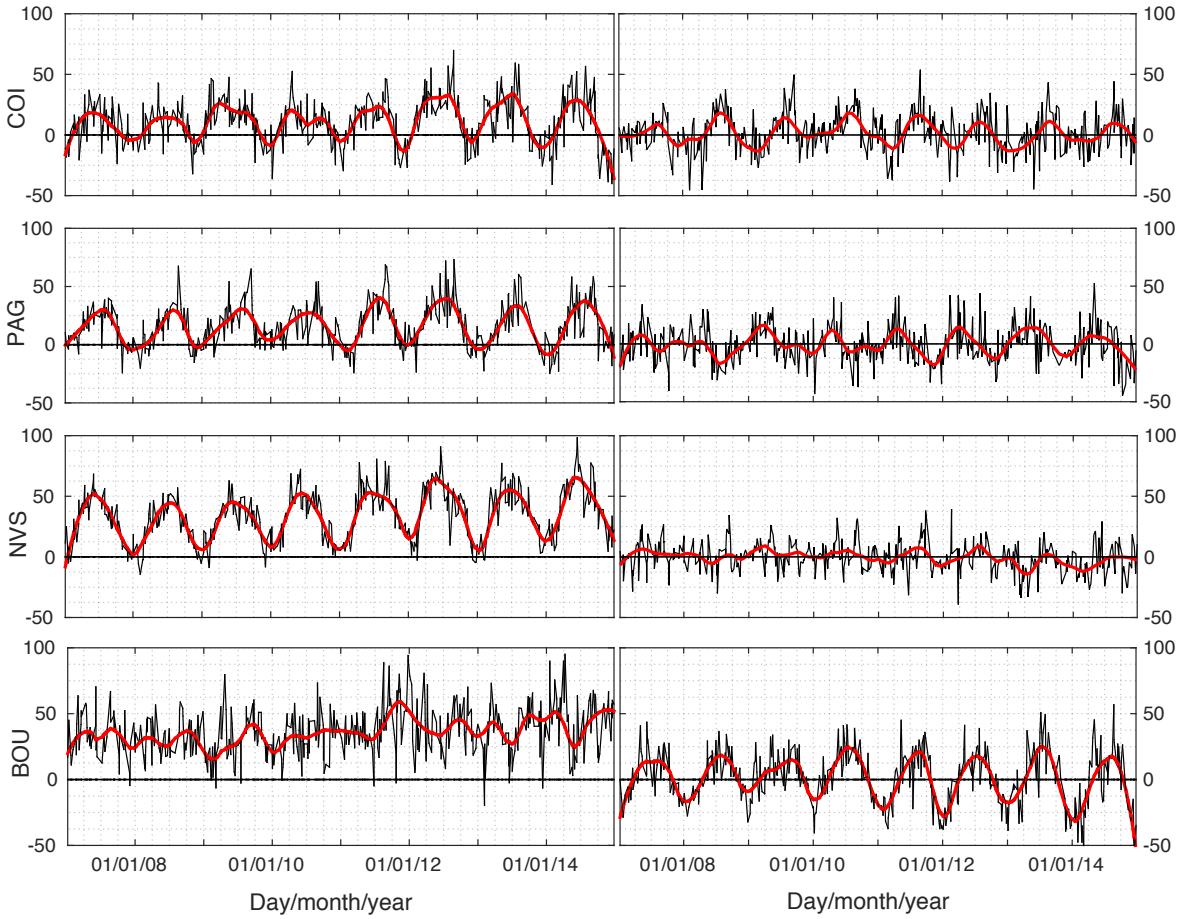


FIGURE 3.9: $\mathbf{a}_1$ (left) and $\mathbf{a}_2$ (right), in black. The smooth curves, in red, are generated by a local regression smoothing method (loess), with a span of 10 % of the data points (see Section 2.6.3). COI on top, then PAG, NVS and BOU.

noticeable in $\mathbf{a}_1$ for NVS, in $\mathbf{a}_2$ for BOU and in both $\mathbf{a}_1$ and $\mathbf{a}_2$ for COI and PAG. The annual dependence of the daily variation series is mainly due to a maximum of ionospheric ionization during summertime. Also noticeable, the modulation by the solar cycle with a decrease in amplitudes close to the minimum of cycle 24 in 2009 and an increase in amplitudes while cycle 24 approaches its maximum in 2012.

The final 96 ΔQH_{PCA-QD} models chosen for each observatory are the simplest ones (less components) that verify two conditions:

- a) they represent the observations with a quite good approximation (at least 80% of variability explained, as seen in Figure 3.8) and

- b) they filter out most irregular geomagnetic activity at higher (less relevant) modes.

Only modes 1 and 2 are used in the PCA-QD models for BOU, COI and NVS and the first three modes are used for PAG, i.e., $l_{max} = 2$ in the former and $l_{max} = 3$ in PAG, where

$$\Delta QH^{i,k}_{PCA-QD} = \sum_{l=1}^{l_{max}} \Delta QH^{i,k}_{PCA-QD-l} \tag{3.12}$$

See also equation 3.10.

Analysis of Pearson's Correlations. The number of r values for each two compared series is always less than 2922, depending mostly on the number of days with temporarily interrupted data acquisition at a given observatory and on the number of days with no satellite data available for TS05 computations. Pearson's correlations between observations and proxies are presented in the form of histograms in Figure 3.10, arranged by observatory: COI (top-left quadrant), PAG (top-right), NVS (bottom-left) and BOU (bottom-right). Correlations for quiet days are at the top row of each quadrant and for active days at the bottom row. Blue bars represent correlations of PCA-QD corrected data ($\widehat{H}$) with BH, green bars represent correlations of $\widehat{H}$ with Dst, and black bars represent correlations of $\widehat{H}$ with RC. All white bars represent correlations of raw H data with the corresponding geomagnetic index or BH. The most significant correlations ($p \leq 0.05$) lie outside the shaded regions.

From eq. 2.32 with $n = 24$, and obtaining the Student t-value from a t-table, then the minimum value of r to be significant at the 5% level must have:

$$\frac{\sqrt{22}r}{\sqrt{1-r^2}} > 2.07, \tag{3.13}$$

which requires $r^2 > 0.16302$ and $|r| > 0.403$. Values of r numerically greater than 0.383 have probability $\geq 95\%$ to be different from zero, corresponding to $p \leq 0.05$.

It is noticeable that histograms for active days are more left-skewed than those for quiet days, meaning that we have better correlations for active than for quiet time. This might be expected because all BH, Dst and RC were designed to measure geomagnetic storm-time activity. We can also observe an increase of skewness to the left from histograms of raw data to the corresponding histograms of PCA-QD corrected data, meaning that there are larger number of high correlation values for corrected data. This is quite clear in COI and PAG for correlations with BH, in NVS for correlations with Dst, and in BOU for correlations with Dst and RC indices. However, it is not the case in BOU and NVS, for correlations with BH.

The Fisher's transformation of COI histograms in Figure 3.10 are shown in Figure 3.11. A chi-square goodness-of-fit test was done to determine if the z values come from a normal distribution. The test calculates the observed and expected counts for the bins and computes the chi-square test statistic:

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{E_i}, \tag{3.14}$$

where O_i are the observed counts and E_i are the expected counts based on the hypothesis distribution. The null hypothesis is that data comes from a normal distribution with a mean and variance estimated from data. The alternative hypothesis is that the data does not come from such a distribution. We obtained near zero values of the chi square statistic, meaning that the Gaussian distribution is a good representation of the histograms for z. Both the z-values distributions and the fitted normal distribution are plotted together for the four stations. Inserted in each plot are the mean $\bar{z}$, the standard deviation σ_z, the number of z-values in each sample and r_0 (see Equation 2.34).

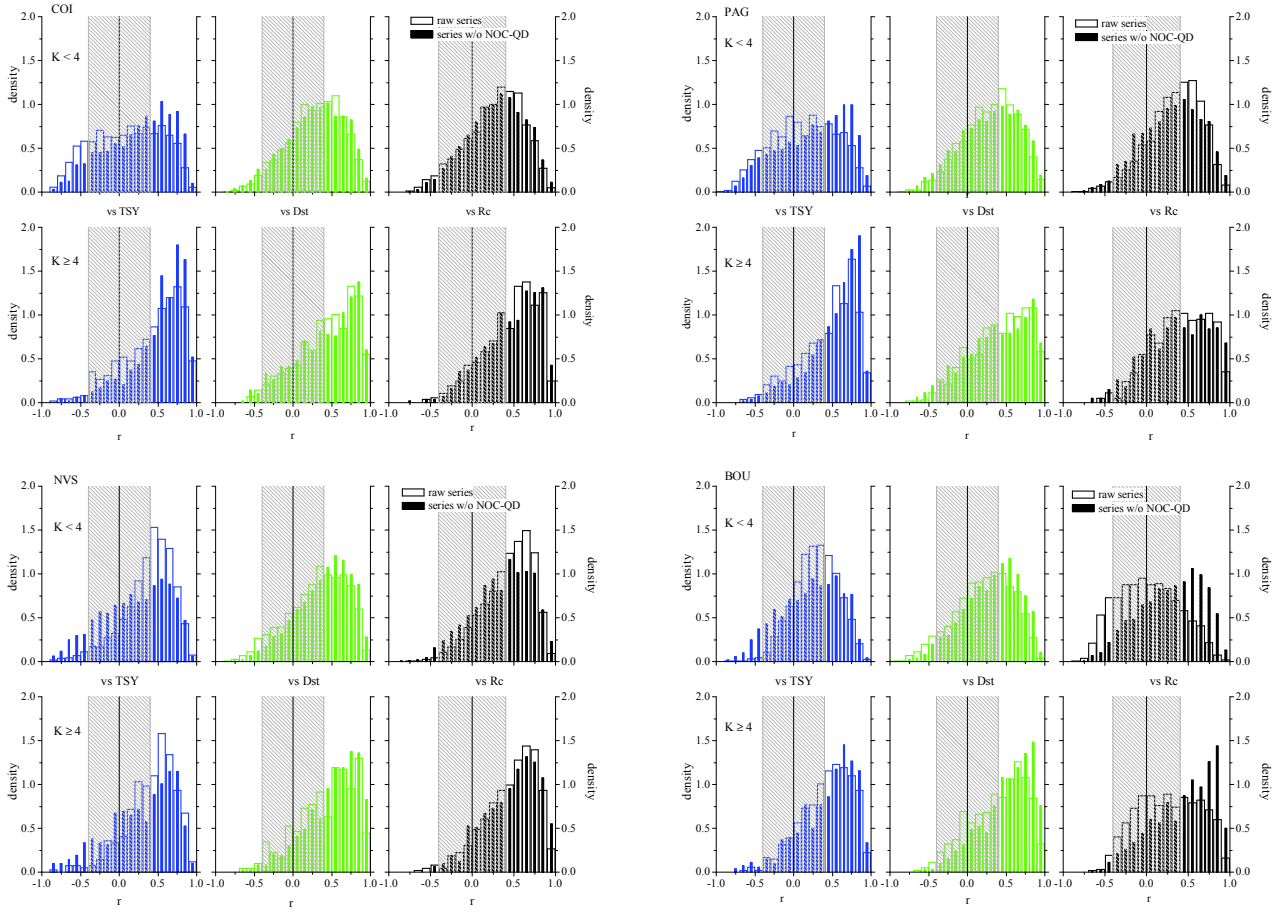


FIGURE 3.10: For COI (top-left), PAG (top-right), NVS (bottom-left) and BOU (bottom-right), histograms of correlation coefficients between observatory series of H and each of BH (blue), Dst (green) and RC (black) proxies, for raw data (unfilled bars) and data without PCA-QD model (filled bars). Separate results for quiet days (top row of each group) and active days (bottom row of each group). Vertical axes values given by $n_i/(\sum n_i \Delta r)$, with n_i the number of correlation values in bin i and Δr the bin width. Correlation values with $p \leq 0.05$ are located outside shaded areas.

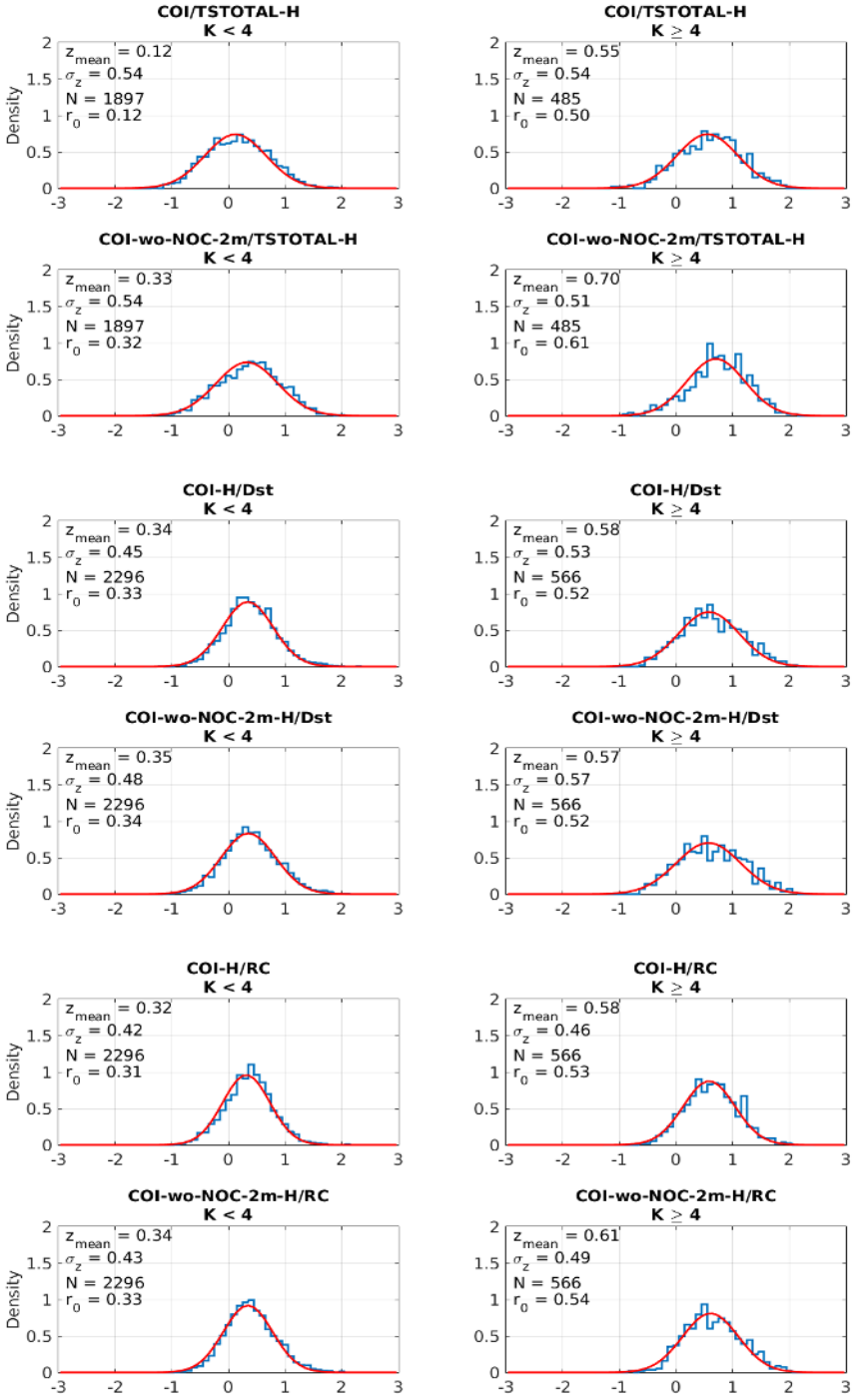


FIGURE 3.11: In blue, the Fisher's transformation of Pearson's correlation coefficients between COI data with BH (two top rows), Dst (two middle rows) and RC (two bottom rows). Top row of each proxy are correlations for raw data, on the left for calm time, on the right for active days. Bottom row of each proxy are correlations for PCA-QD corrected data (left, calm; right, active). In red, fitted normal functions. The inserted text represents the mean $\bar{z}$, the standard deviation σ_z, the number of z-values in each sample, N, and r_0.

TABLE 3.4: Statistical parameters describing the distribution of correlation coefficients between H and different proxies (left) and $\widehat{H}$ and different proxies (right). N is the number of correlation values with p-value < 0.05; $\bar{r}$ and $\bar{p}$ are the mean correlation value and the mean p-value, respectively, computed using only significant correlation coefficients; r_0 is the inverse Fisher's transform of $\bar{z}$, the mean of the Gaussian distribution fitted to the histogram of z-values (see Figure 3.11 and explanation in the text).

		Before removing PCA-QD								**After removing PCA-QD**							
Proxy	Par	COI		PAG		NVS		BOU		COI		PAG		NVS		BOU	
		calm	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*
TSY	N	877	303	772	341	1143	274	692	314	1003	365	988	393	978	242	824	348
	$\bar{r}$	0.19	0.64	0.28	0.64	0.53	0.64	0.55	0.63	0.44	0.67	0.46	0.68	0.34	0.62	0.42	0.63
	$\bar{p}$	0.009	0.005	0.010	0.005	0.010	0.006	0.011	0.007	0.007	0.004	0.007	0.004	0.009	0.005	0.009	0.005
	r_0	0.12	0.50	0.16	0.52	0.41	0.53	0.32	0.48	0.32	0.61	0.33	0.61	0.22	0.46	0.25	0.52
Dst	N	1027	345	1024	352	1148	340	943	346	1071	335	1059	346	1372	362	1142	432
	$\bar{r}$	0.52	0.65	0.54	0.66	0.52	0.68	0.42	0.63	0.51	0.66	0.53	0.63	0.60	0.72	0.55	0.69
	$\bar{p}$	0.009	0.006	0.010	0.005	0.009	0.005	0.010	0.006	0.009	0.005	0.009	0.006	0.008	0.004	0.009	0.006
	r_0	0.33	0.52	0.35	0.50	0.34	0.59	0.25	0.46	0.32	0.61	0.33	0.61	0.22	0.46	0.25	0.52
RC	N	983	351	1097	327	1499	278	739	256	978	346	1001	324	1260	310	1104	385
	$\bar{r}$	0.49	0.66	0.55	0.64	0.62	0.64	0.07	0.58	0.55	0.68	0.57	0.64	0.59	0.67	0.55	0.67
	$\bar{p}$	0.011	0.005	0.009	0.007	0.007	0.006	0.013	0.008	0.009	0.005	0.009	0.006	0.008	0.006	0.008	0.006
	r_0	0.31	0.53	0.38	0.47	0.49	0.49	0.03	0.32	0.33	0.54	0.35	0.46	0.41	0.56	0.35	0.54

In Table 3.4 we find information on different statistical parameters characterizing the histograms. In general, it is noticeable that r values for active days are higher than for quiet days. We can appreciate a different behaviour in COI and PAG compared with NVS and BOU. At COI and PAG we can see that BH explains well the observations after subtracting the PCA-QD model, with improvement of all parameters listed at the table. Dst and RC also explain well the data, but no significant improvement is obtained after subtracting PCA-QD. At NVS and BOU, Dst and RC tend to explain better the observed corrected series than BH. The most atypical results on this table are the very low values of $\bar{r}$ and r_0 characterizing the correlations between the RC and raw data at Boulder during calm days, before subtracting PCA-QD. When we remove the PCA-QD model from BOU raw data, then correlation values become more typical. To test the significance of these observations a Welch's t-test at 95% confidence was performed, comparing the $\bar{z}$ of every correlation distribution of raw observations with the $\bar{z}$ of their corresponding distributions of corrected data. Results are shown in Table 3.5. The "+" sign means that we obtained a significant improvement with the PCA-QD correction, "0" means we obtained no significant improvement, and "-" sign means that we obtained a decrease in correlation values after removing the PCA-QD. This last situation occurs between NVS data with BH, with BOU data with BH for calm days and with NVS data with RC for calm days. Improvements occur for COI and PAG data with BH, and for NVS and BOU with Dst and RC, except for NVS with RC during quiet days.

One other Welch's t-test identifies the best proxy to use for each observatory and depending on the geomagnetic activity (Table 3.6). Here, the three proxies were tested in pairs, starting by RC vs Dst, followed by the best of these two vs TSY.

TABLE 3.5: Results of Welch's t-tests comparing correlations of raw data (H) with correlations of PCA-QD corrected data ($\widehat{H}$) for each proxy. Is there a statistically significant improvement in correlations after removing the PCA-QD variation? Increase ($'+'$) or decrease ($'-'$) of correlation values after removal of daily variation model. $'0'$ is for no significant change.

	COI		PAG		NVS		BOU	
	calm	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*
BH	$+$	$+$	$+$	$+$	$-$	$-$	$-$	0
Dst	0	0	0	0	$+$	$+$	$+$	$+$
RC	0	0	0	0	$-$	$+$	$+$	$+$

TABLE 3.6: Which is the best proxy, using Welch's t-test? 'All' is for no statistically significant difference between the three proxies.

	COI		PAG		NVS		BOU	
	calm	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*
Raw data	Dst/RC	All	RC	BH/Dst	RC	Dst	BH	BH/Dst
Without QD	All	BH	All	BH	Dst	Dst	Dst/RC	Dst

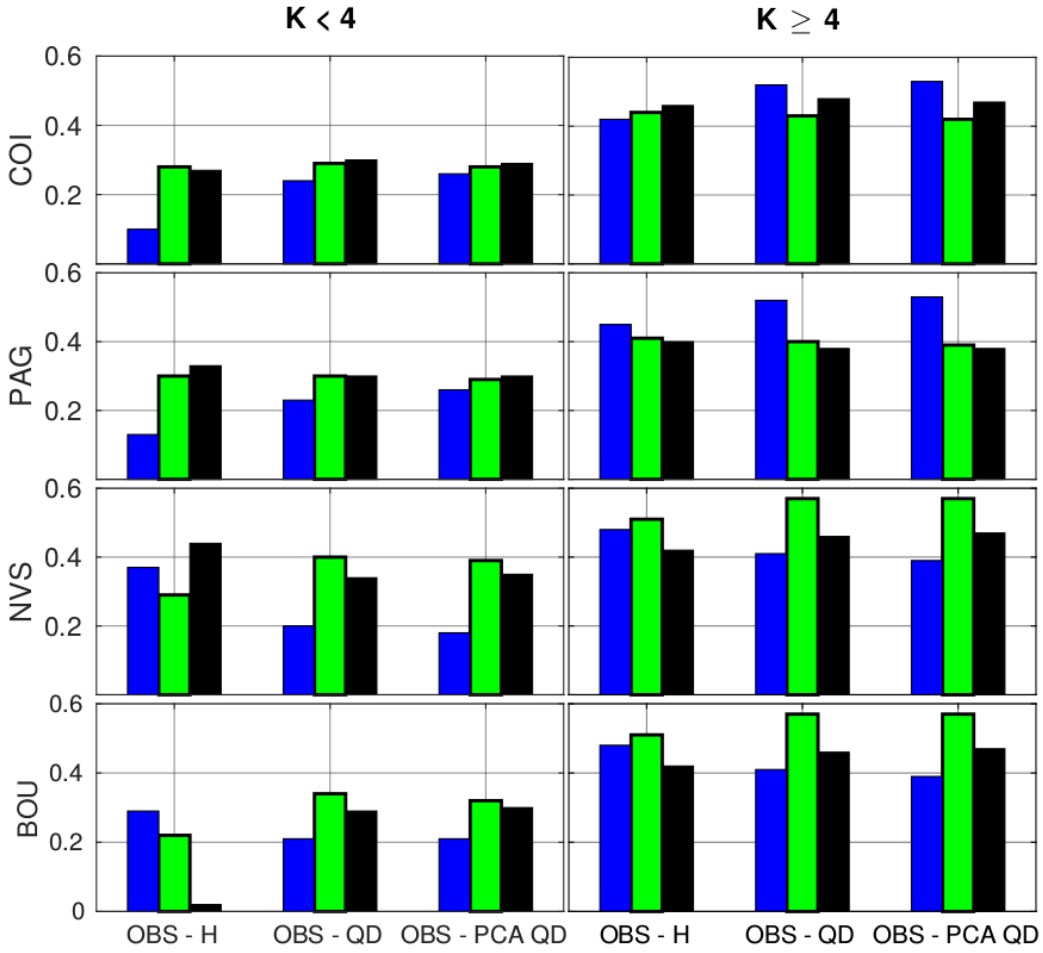


FIGURE 3.12: In each plot, $\bar{r}$ values for correlations with BH (blue), Dst (green) and RC (black), using raw data series (OBS-H), data series corrected for QD variation using standard procedure (OBS-QD) and data series corrected for QD variation using PCA-QD model (OBS - PCA QD). Separated values for $K < 4$ (left column) and $K \geq 4$ (right column). COI on top, then PAG, NVS and BOU.

Figure 3.12 presents $\bar{r}$ for correlations of H, $\widetilde{H}$ and $\widehat{H}$ with each proxy. In COI and PAG we can see that the data is better explained by BH after subtracting QD, but a similar improvement is not seen when comparing data with Dst and RC. At NVS and BOU we can see higher $\bar{r}$ values when comparing BH with raw data than with data with QD subtracted. There is however a general improvement of correlations between Dst or RC with data after QD correction (except for NVS in calm days, which shows higher $\bar{r}$ values without QD correction).

In Figures 3.10 and 3.12, and Table 3.4, it is recurrent that correlations are smaller for lower than for higher levels of geomagnetic activity. This is something to be expected since both the geomagnetic indices (Dst and RC) and the Tsyganenko model were developed to be representative of storm-time geomagnetic activity. Nonetheless, even for low activity, global representative values of r_0 are equal or larger than 0.3, in certain cases only after correcting for daily variation, and they can be as high as 0.6.

It is also notorious that two groups of observatories have distinctive behaviours, one is COI and PAG, and the other is NVS and BOU. Starting with COI and PAG, they have higher representative global values of r_0 for BH, being as high as 0.6 with QD correction at more active days. For quiet days, though showing lower scores, the improvement of correlations with BH simulations after removal of daily variation is also clear (Table 3.4). The performance of Dst and RC geomagnetic indices is worse than TS05 simulations for more active days and equivalent to them for less active days. An improvement in performance of Dst and RC indices after removal of the daily variation model is not clear (Table 3.5).

Looking to NVS and BOU, the best scores are obtained with Dst for more active days (Table 3.6). Correlations with $\widehat{BH}$ are in general lower than with the other two proxies (Table 3.4). For both less and more active days, a clear improvement in correlations with Dst and RC is obtained after removing the daily variation from data, except at NVS with RC at calm time (Table 3.5). On the contrary, the daily variation removal seems to contribute to a slight deterioration of results when comparing data with TS05

simulations (Tables 3.4, 3.5).

3.5 Discussion

PCA-QD models. TS05 does not include the ionospheric currents originated from solar and lunar daily variations nor does it include their corresponding induced currents. On the other hand, Dst is constructed from a set of four observatory data series corrected from the daily variation. As for the RC series, there is no explicit removal of daily variation but, because data are taken in the night period, it should not be present. In order to compare TS05 simulations, Dst and RC proxies with observations, the quiet daily variation was first computed and then subtracted from observatories data.

The choice to use PCA to remove the QD variation is motivated by both dimensionality reduction and a possible separation of ionospheric and magnetospheric QD contributions in case they are decorrelated, with a concentration of ionospheric signal in the main modes of PCA-QD. The minimum number of principal modes required for PCA models is two, to make it possible for these models to account for an annual phase shift observed in QD variation, that consists in a displacement of the maximum perturbation towards early hours in summer (see Figure 3.7 (b), for BOU and NVS). The amplitude coefficients associated to the first two modes of each PCA-QD model reveal annual and decade variations (Figure 3.9), showing that the PCA models can capture the seasonal and decade variations of the daily curves for the whole 2007-2014 selected period.

The separation between magnetospheric and ionospheric QD variation could only be tested in correlations between data and the BH proxy, which includes the magnetospheric but not the ionospheric QD signal. Dst and RC proxies do not distinguish QD signals from different sources. As observed, the use of a PCA-QD model instead of the standard procedure to remove QD variation slightly increases $\bar{r}$ in correlations with BH (Figure 3.12), but only for COI and PAG stations. For BOU and NVS stations,

correlations with BH deteriorate after removal of a standard QD variation model and results do not get better when removing a PCA-QD model instead (Figure 3.12). Table 3.7 presents results with 95% confidence level of Welch's t-tests that show in almost all cases (except for PAG during low activity), that there is no significant advantage in using PCA-QD models instead of standard QD. The whole set of results are not clear-cut concerning the ability of PCA-QD models to separate ionospheric and magnetospheric contributions and further studies are recommended, using other observatories and periods of time.

Regarding the ability of PCA-QD models of concentrating the most important features of QD variation into only two or three components, Table 3.7 shows that, in all but one case, the removal of the PCA-QD model, with only two or three modes, gets data closer to Dst and RC proxies as efficiently as the removal of the whole QD signal computed in a standard way. The exception is BOU during low activity where the PCA model is not as efficient as the standard QD in making data closer to Dst proxy.

TABLE 3.7: Results of 95% confidence Welch's t-tests comparing correlations of standard QD corrected data ($\widetilde{H}$) and correlations of PCA-QD corrected data ($\widehat{H}$) with each proxy. The null hypothesis is that the two means of the z-values distributions are the same. Is there a statistically significant improvement in correlations by removing a PCA-QD model instead of standard QD? Increase ('$+$') or decrease ('$-$') of correlation values due to use of the PCA-QD model. '0' is for no significant change.

	COI		PAG		NVS		BOU	
	calm	*active*	*calm*	*active*	*calm*	*active*	*calm*	*active*
BH	0	0	$+$	0	0	0	0	0
Dst	0	0	0	0	0	0	$-$	0
RC	0	0	0	0	0	0	0	0

Longitudinal dependence found in data series. The four observatories in this study separate into two groups with clearly distinct results. For all of them, removing the

QD variation gets data series closer both to Dst and RC proxies, as expected. However, this effect is more significant for BOU and NVS stations than for COI and PAG (Table 3.5).

It is difficult to predict if removing QD variation from data series will make them closer to BH proxy which contains a diurnal contribution from magnetospheric sources. In fact, Figures 3.10, 3.12 and Table 3.5 show that while for COI and PAG the removal of QD variation gets data closer to BH series, the opposite occurs for BOU and NVS.

The dependence in longitude (in an Earth-fixed reference frame) revealed by these results, may come either from the ionospheric or the magnetospheric QD variation sources, or both. The magnetospheric currents in TS05 model are described in the geocentric solar magnetospheric system, GSM (see definition in Section 2.3). The partial ring current (PRC) and associated Birkeland currents (or field-aligned currents, FAC) represent a coupled system, fixed in the GSM frame, with Birkeland currents circulating at dawn/dusk meridian and ensuring closure of PRC (Tsyganenko, 2000; Tsyganenko, 2002a; Tsyganenko, 2013). There is a clear longitudinal dependence in the GSM system associated with these currents, corresponding to the observation of pronounced noon-midnight and dawn-dusk asymmetries. As the Earth rotates relative to the GSM system, the magnetic field produced by these currents at the Earth's surface shows a daily variation in local time, even during quiet geomagnetic activity. The dependence of these current fields on the magnetic local time (MLT) ϕ' (angle between the geomagnetic meridian through a given station, or geomagnetic longitude, and the one through the Sun, converted to hours) is of the form $\sim \cos\phi'$ and $\sim \sin\phi'$. At stations in middle latitudes as those considered here, MLT differs only slightly from LT (geographic local time).

Let us analyse the possibility that the longitudinal dependence that separates COI/PAG from NVS/BOU may be due to magnetospheric sources. Quiet daily variation of PRC and Birkeland current systems (ΔQH_{PRC} and ΔQH_{FAC}) were obtained in a similar

way as the observatories ΔQH series, averaging the values of each TS05 synthetic series at the five quietest days of each month. The local time dependence of the PRC and associated Birkeland current system can be seen in middle and right columns of Figure 3.13 where the Birkeland currents (Regions 1 and 2, together) and PRC hourly contributions all over the day are plotted for the whole 2007-2014 period. For each of these two magnetospheric field contributions at ground level, the ΔQH diurnal series are represented at the vertical axis, repeating the same series for all days in the same month. A sinusoidal variation of 24 h period is seen in both FAC and PRC plots, with minimum (negative) values at around 18:00-19:00 LT for PRC and around 12:00 LT for FAC. It turns out that both contributions for magnetospheric QD variation are stronger for NVS and BOU than for COI and PAG (Figure 3.13). This means that the local time dependence is modulated by a longitudinal dependence on the Earth-fixed reference frame, apparently associated with the difference in geomagnetic latitude between the four stations (see Table 3.1).

But longitudinal dependence can also be due to ionospheric sources. Significant longitudinal dependence of the Sq current system has also been observed and extensively studied (see e.g. Pedatella et al., 2011). In this case, it was found that the longitudinal variation of the main field could largely account for the observed asymmetry. Figure 3.13 (left column) shows the same ΔQH curve (vertical axis) during all days of its corresponding month (horizontal axis). Here we can appreciate the asymmetry at different stations with a pattern that is familiar from studies of the Sq current system. The four observatories are differently positioned relative to the Sq current vortex focus, as it drifts over the Northern Hemisphere. In models of the Sq current system (see e.g. Malin and Gupta, 1977; Stening and Winch, 2013), the foci of equivalent currents responsible for the diurnal variation, drift closer to European stations than stations in Asia or North America (see e.g. Malin and Gupta, 1977), at transit times. In agreement with this, the analysis of QD variation curves obtained in section 3.4 (see Figure 3.7 (b) shows that BOU and NVS seem to be closer to the northern latitudinal branch of the

vortex where strong western electric currents are responsible for the depression of the H component at the middle of the day. COI and PAG seem to be closer to the vortex center, exhibiting a dipole signal which is characteristic of a transition zone between the two regions dominated by the western (North) and the eastern (South) latitudinal branches of the vortex (see e.g. Matsushita and Maeda, 1965). A clear displacement of the vortex towards North in winter is also noticed in Figure 3.7, specially for COI, PAG and NVS.

The similarity of the quiet daily variation of PRC (ΔQH_{PRC}) and Birkeland (ΔQH_{FAC}) currents with the QD variation at each observatory is also analysed by projecting the quiet daily magnetospheric signal shown in middle and right columns of Figure 3.13 onto the principal components found for each observatory. The formula to calculate the expansion coefficients $\mathbf{a}_i$ by indirect way (see Section 2.6.1) can be used to project these magnetospheric QD signals along the EOFs. It is equivalent to equation 2.22:

$$\mathbf{A}_{cur} = \mathbf{\Delta QH}_{cur}\mathbf{V} \tag{3.15}$$

where $\mathbf{\Delta QH}_{cur}$ is matrix 96 x 24 ($\mathbf{\Delta QH}_{PRC}$ or $\mathbf{\Delta QH}_{FAC}$ in this case) and $\mathbf{V}$ is the matrix of EOFs (24 x 24) for $\Delta\mathbf{QH}$.

Results are shown in Figure 3.14. At BOU and NVS the FAC signal has the highest contribution from PC-1, which by definition explains most of the variability of the QD variation at those stations (as in Figure 3.8). This explains the similarity between ΔQH and ΔQH_{FAC} in Figure 3.13, for these two stations. At COI and PAG, the FAC signal has the highest contribution from PC-2, which is clearly less important in explaining QD variation (as in Figure 3.8), meaning that the presence of magnetospheric signal in QD variation is less important for these two stations. PRC has low contributions from both modes 1 and 2, except for NVS where the contribution of PC-2 is higher than for the other three stations.

We can see in Figure 3.13 that ΔQH_{FAC} at BOU and NVS are more similar to ΔQH

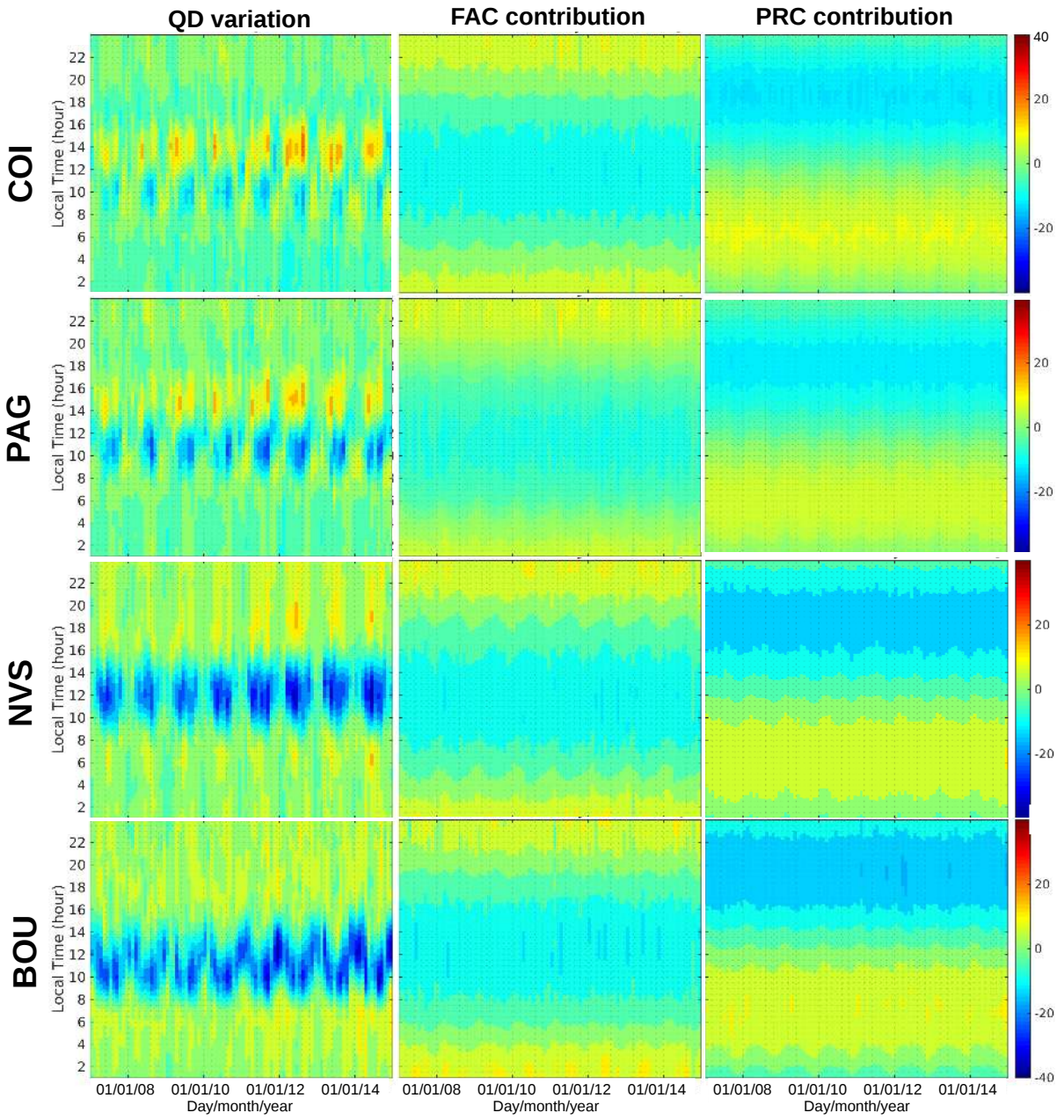


FIGURE 3.13: Daily variation in Local Time (along vertical axis) of standard QD from data (left column); of FAC and PRC, from TS05 model (central and right columns), for the four stations used in this study and the whole 2007-2014 period.

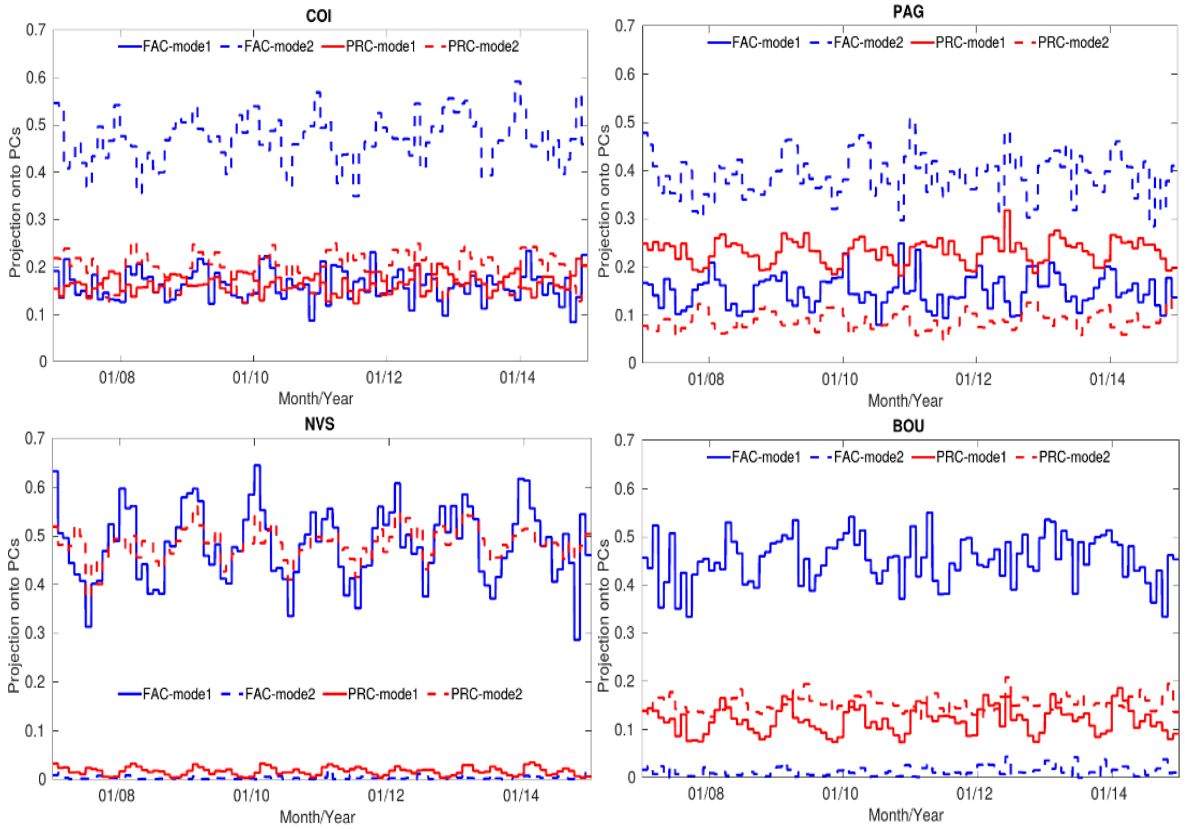


FIGURE 3.14: For each day in the 2007-2014 time period, the ΔQH_{FAC} (blue) and ΔQH_{PRC} (red) series represented in Figure 3.13 (middle and right columns) were projected onto the two principal components of the corresponding observatories (solid lines for PC1, dashed for PC2).

and have a stronger signature compared to COI and PAG, with a 24-h period, approximately sinusoidal, with minimum around 12:00 LT. This is no longer true for COI and PAG due to the change in sign in H series around noon. The lower correlation between ionospheric and TS05 magnetospheric contributions for COI and PAG during the 24-hour period, may explain why the PCA tools seem more efficient to separate the two components for these two stations (Figure 3.12), since by construction they separate principal components on the basis of uncorrelated signals. In summary, stronger magnetospheric contribution and higher correlation between ionospheric and magnetospheric signals in the QD variation for BOU and NVS, may explain the greater difficulty to disentangle ionospheric and magnetospheric signals. For COI and PAG this seems easier to achieve, thanks to a different positioning relative to the Sq foci and a lower contribution of QD variation from magnetospheric sources.

4 Geomagnetic Activity: Contribution of Different Magnetospheric Currents

4.1 Overview

In this chapter, as in the previous one, I use a statistical approach to evaluate the performance of TS05 model in estimating the magnetospheric transient signal observed at the same four Northern Hemisphere's mid-latitude ground stations analysed in Chapter 3, for the 2007 to 2014 time interval. However, while in Chapter 3 series of the H component were computed to compare with activity indices Dst and RC, here the X and Y geomagnetic local components are separately analysed and comparison is made only with TS05 simulations. The standard quiet daily (QD) variation is removed both from data series and simulations before comparison. It is found that the TS05's performance is clearly better for the X than for the Y field components and, as already found in the case of the H component, for more geomagnetically active days as determined by local K-index. For $\sim 50\%$ (X) and $\sim 30\%$ (Y) of the total number of geomagnetically active days, correlation values yield $r \geq 0.7$. During more quiet conditions, only $\sim 30\%$ (X) and $\sim 15\%$ (Y) of the number of analysed days show $r \geq 0.7$. A main contribution of this chapter is the computation of separate contributions from different magnetospheric currents to the observed data time variability and signal magnitude. During more active days, all Tail, SRC and PRC contribute to the time variability in X while the PRC and FAC currents contribute most to time variability of Y. In the best case

estimations when $r \geq 0.7$, remaining differences in scale between observations and TS05's predictions can be explained by global induction in the Earth's upper layers. The closing of FAC currents through the Earth's center in the TS05 model seems to be mainly affecting the Y magnetospheric field predictions.

4.2 Methodology

Correlation Coefficients and Statistical Tests. In the previous Chapter it was shown that, by subtracting QD only to the data, correlations with TS05 simulations could decrease. It was the case for BOU and NVS (see Figure 3.12) and it was explained by the inability to separate the QD ionospheric from the QD magnetospheric contributions. To bypass this problem, in this chapter the QD variation is removed both in data series and TS05 simulations.

For each observatory, series of observed X and Y hourly values were compared with TS05's derived series before (r) and after ($\widetilde{r}$) subtraction of QD curves. The Pearson's correlation coefficients (equation 2.31) were computed to measure the linear dependence between the data and the predictions, taking corresponding samples of 24 hourly values from each population, day by day, along the 8-yrs period. The number of correlation values (either rX, rY, $\widetilde{rX}$ or $\widetilde{rY}$) values for each two compared series is always $\leq$ 2922 (the total number of days in this period). Table 3.3 shows the exact number of days for which correlation values were computed, in each station.

I also compared observatory series with each TS05's current contribution separately using $\widetilde{X}$ and $\widetilde{BX}_{cur}$ ($\widetilde{Y}$ and $\widetilde{BY}_{cur}$), and called the correlation coefficients as $\widetilde{rX}^{j,k}_{cur}$ (and $\widetilde{rY}^{j,k}_{cur}$). For each day j in month k:

$$\widetilde{rX}^{j,k}_{cur} = \frac{\sum_{i=1}^{24}\left(\widetilde{BX}^{i,j,k}_{cur} - \overline{BX}^{j,k}_{cur}\right)\left(\widetilde{X}^{i,j,k} - \overline{X}^{j,k}\right)}{\sqrt{\sum_{i=1}^{24}\left(\widetilde{BX}^{i,j,k}_{cur} - \overline{BX}^{j,k}_{cur}\right)^2}\sqrt{\sum_{i=1}^{24}\left(\widetilde{X}^{i,j,k} - \overline{X}^{j,k}\right)^2}} \quad . \tag{4.1}$$

where 'cur' stands for CF, TAIL, SRC, PRC, FAC or INT. Note that daily means of each series are the same with or without QD, since QD is defined having zero daily mean. Equation (4.1) describes the similarity between the signal due to each individual source and the data series for a given day, in terms of simultaneous raises and decreases relative to corresponding means.

As in Chapter 3, r, $\widetilde{r}$ and $\widetilde{r}_{cur}$ are computed day by day, using 24 hourly values, for the 2007-2014 time period. The set of obtained values are represented in the form of histograms and standard statistical tools and tests as Student's tests and Fisher's z-transformation (see Section 2.6.2) are used to assess their significance. Less and more geomagnetically active days for each station are discriminated again by using the local K-indices on the basis of all daily K-values less than 4 (quiet) or at least one value $K \geq 4$ (active).

Linear Regression Models. A shift in magnitude (or bias) exists between $\widetilde{X}$ ($\widetilde{Y}$) and $\widetilde{BX}$ ($\widetilde{BY}$) daily series (see data as red dots and TS05 series as blue dots in Figure 4.1), which can not be quantified using Pearson's correlation coefficients. This shift may be due to ionospheric or crustal fields. As mentioned in Section 3.3, a daily mean ionospheric signal is left in the data series after removal of ΔQX and ΔQY because the datum line for QD curves was defined by removing the daily mean instead of the mean at night time hours. The bias from ionospheric sources varies with seasonal periodicity and solar cycle activity and is less than 10 nT as checked using both datum lines for computing the QD variation, as shown in Figures 3.5 and 3.6. A more significant bias due to local crust magnetization fields is expected to affect the observatories series, with different values for each station that can reach $\sim$ 200 nT (e.g., Mandea and Langlais, 2002). These crustal biases remain constant in time for the analysed period and can be identified subtracting a main field model from the data (Mandea and Langlais, 2002; Verbanac et al., 2015).

More difficult to deal with in general, is the effect of Faraday induction of electrical currents flowing in the lithosphere and upper mantle by transient fields of external origin (e.g., Schmucker, 1985). This effect is not taken into account in the TS05 model, but it is contaminating the data. Using a simple model it is shown that induction is responsible for a scaling factor between total (inducing plus induced) and inducing fields.

For each day, I obtain the shift $\alpha_X^{j,k}$ ($\alpha_Y^{j,k}$) and scaling $\beta_X^{j,k}$ ($\beta_Y^{j,k}$) parameters by linear regression of data and TS05's series choosing only series with $r \geq 0.7$ and separating results for quiet and active days:

$$\begin{aligned} \left(\Delta\widetilde{X}^{i,j,k}\right)^{pre} &= \alpha_X^{j,k} + \beta_X^{j,k} \cdot \Delta\widetilde{BX}^{i,j,k} \\ \left(\Delta\widetilde{Y}^{i,j,k}\right)^{pre} &= \alpha_Y^{j,k} + \beta_Y^{j,k} \cdot \Delta\widetilde{BY}^{i,j,k} \end{aligned} \tag{4.2}$$

where $(\Delta\widetilde{X})^{pre}$ ($(\Delta\widetilde{Y})^{pre}$) and $\Delta\widetilde{BX}$ ($\Delta\widetilde{BY}$) are respectively the model predictions and TS05's series with QD and IGRF removed. Since the same β scaling parameter applies to all TS05's field components, this regression will cause no re-weighting of different magnetospheric sources. Figure 4.1 shows the $\widetilde{X}$ ($\widetilde{Y}$), $\widetilde{BX}$ ($\widetilde{BY}$) and $(\Delta\widetilde{X})^{pre}$ ($(\Delta\widetilde{Y})^{pre}$) for active days where $\widetilde{rX}$ ($\widetilde{rY}$) ≥ 0.7.

To perform the linear regression, I call $\mathbf{C}$ a column vector with 24 observatory hourly values at day j and month k (same for $\widetilde{Y}$):

$$\mathbf{C} = \begin{bmatrix} \Delta\widetilde{X}^{1,j,k} \\ \Delta\widetilde{X}^{2,j,k} \\ \vdots \\ \Delta\widetilde{X}^{24,j,k} \end{bmatrix} \tag{4.3}$$

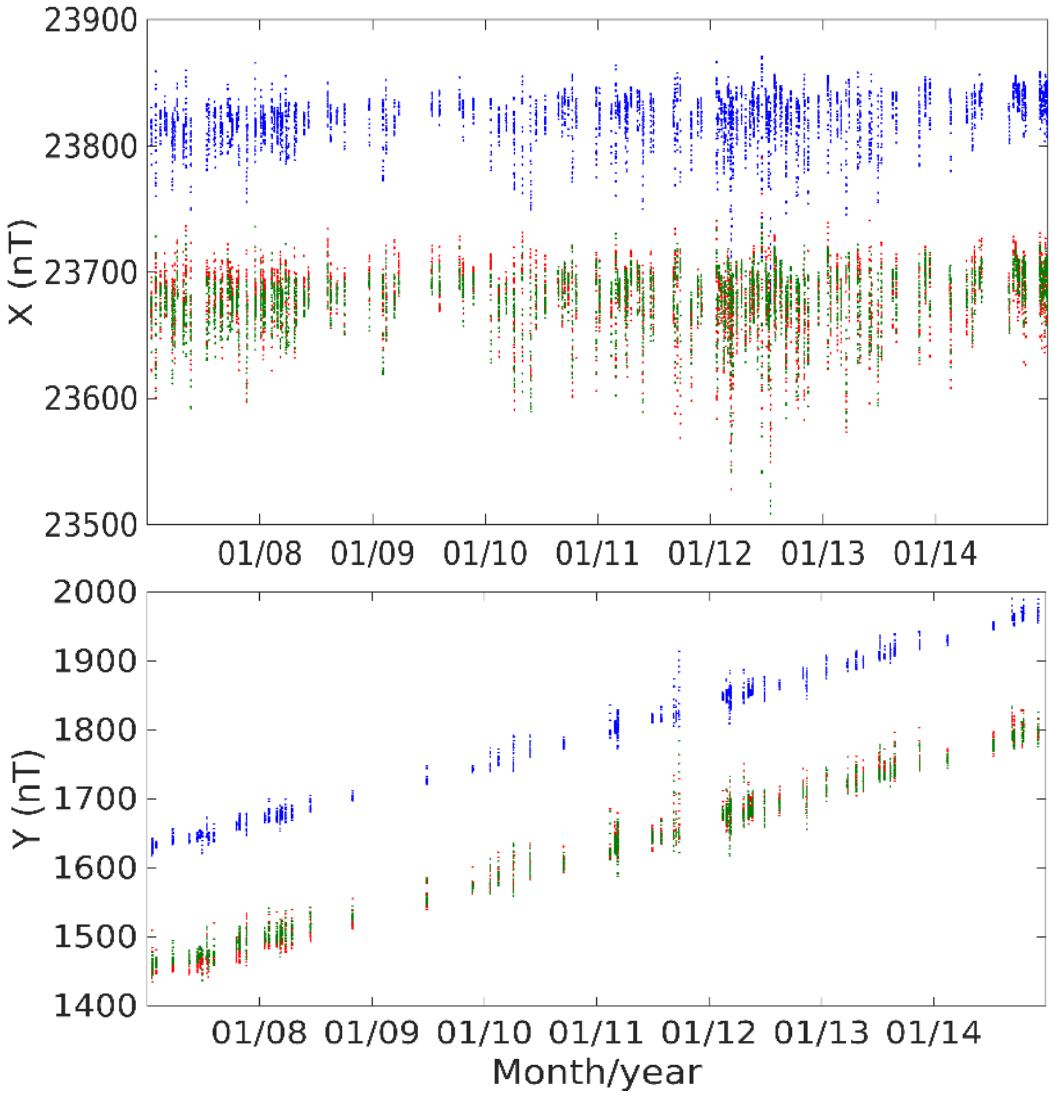


FIGURE 4.1: Top panel: series of $\widetilde{X}$ (red dots), $\widetilde{BX}$ (blue dots) and $(\Delta\widetilde{X})^{pre}$ (green dots) at PAG, for $\widetilde{r} \geq 0.7$ and $K \geq 4$. Bottom panel: the same for $\widetilde{Y}$ (red dots), $\widetilde{BY}$ (blue dots), $(\Delta\widetilde{Y})^{pre}$ (green dots) at PAG.

I call **D** a matrix with one column of ones and other column of 24 TS05's hourly values (Yamazaki and Maute, 2016) at day j and month k (same for $\widetilde{BY}$):

$$\mathbf{D} = \begin{bmatrix} 1 & \Delta\widetilde{BX}^{1,j,k} \\ 1 & \Delta\widetilde{BX}^{2,j,k} \\ \vdots & \vdots \\ 1 & \Delta\widetilde{BX}^{24,j,k} \end{bmatrix} \tag{4.4}$$

The column vector with a daily $\alpha^{j,k}$ and a daily $\beta^{j,k}$ that minimize the squares of residuals, $\left(\Delta\widetilde{X} - \Delta\widetilde{X}^{pre}\right)^2$, is given by:

$$\mathbf{M} = \left(\mathbf{D^T D}\right)^{-1}\mathbf{D^T C} = \begin{bmatrix} \alpha^{j,k} \\ \beta^{j,k} \end{bmatrix} \tag{4.5}$$

Then I calculated the means and variances of all daily $\alpha_X^{j,k}$ ($\alpha_Y^{j,k}$) and $\beta_X^{j,k}$ ($\beta_Y^{j,k}$) distributions separately for quiet and active days at each station, e.g.,

$$\begin{aligned} \bar{\alpha}_X &= \frac{1}{n}\sum_{j=1}^{n}\alpha_X^{j,k} \\ \bar{\beta}_X &= \frac{1}{n}\sum_{j=1}^{n}\beta_X^{j,k} \end{aligned} \tag{4.6}$$

where n is the number of active or quiet days with $r \geq 0.7$ at a given observatory. The two parameters $\bar{\alpha}$ and $\bar{\beta}$ are meant to model in the simplest possible way effects of local biases (due to crust magnetization and some possible external contribution) and induction in the crust by transient signals, respectively.

Amplitude Coefficients. The estimated series from a given magnetospheric source may yield a high r_{cur} and yet not explain the observed signal strength. A further parameter besides r_{cur} is needed to fully characterize the contribution of source 'cur',

namely a parameter that takes into account the mean and scale of that estimated series (McCollough et al., 2008; Dubyagin et al., 2014). From equation (4.2), we compute daily series $\Delta\widetilde{RBX}$ and $\Delta\widetilde{RBY}$ to compare with $\Delta\widetilde{BX}$ and $\Delta\widetilde{BY}$, respectively:

$$\begin{aligned}\Delta\widetilde{RBX}^{i,j,k} &= (\Delta\widetilde{X}^{i,j,k} - \alpha_X^{j,k})/\beta_X^{j,k} \\ \Delta\widetilde{RBY}^{i,j,k} &= (\Delta\widetilde{Y}^{i,j,k} - \alpha_Y^{j,k})/\beta_Y^{j,k}\end{aligned} \tag{4.7}$$

where $\Delta\widetilde{X}^{i,j,k}$ and $\Delta\widetilde{Y}^{i,j,k}$ are data QD-corrected series with IGRF model subtracted. We introduce new parameters $\widetilde{aX}$ and $\widetilde{aY}$ such that

$$\widetilde{aX}^{j,k} = \frac{\sum_{i=1}^{24}\left(\Delta\widetilde{BX}^{i,j,k}\right)^2}{\sum_{i=1}^{24}\left(\Delta\widetilde{RBX}^{i,j,k}\right)^2} \tag{4.8}$$

and analogously for $\widetilde{aY}$. These two parameters are used to measure the prediction efficiency for X and Y series, respectively (see McCollough et al. (2008) and Dubyagin et al. (2014) for other possible parameters). For each analysed day we expect $\widetilde{aX}$ (and $\widetilde{aY}$) to be close to 1 since the values of $\Delta\widetilde{BX}$ have to be close to the corresponding values of its regression model, $\Delta\widetilde{RBX}$. The identity also is verified if $\Delta\widetilde{X}^{pre}$ ($\Delta\widetilde{Y}^{pre}$) in equation 4.2 represents exactly $\Delta\widetilde{X}$ ($\Delta\widetilde{Y}$).

I also define source normalised amplitude coefficients $\widetilde{aX}_{cur}$ (and $\widetilde{aY}_{cur}$) for each day,

$$\widetilde{aX}_{cur}^{j,k} = \frac{\sum_{i=1}^{24}\Delta\widetilde{BX}_{cur}^{i,j,k}\Delta\widetilde{BX}^{i,j,k}}{\sum_{i=1}^{24}\left(\Delta\widetilde{RBX}^{i,j,k}\right)^2} \quad . \tag{4.9}$$

These coefficients represent the projection of each individual source contribution onto the global TS05's estimate. They are normalised with $\Delta\widetilde{RBX}$, in order that $\widetilde{aX}^{j,k} = \sum\widetilde{aX}_{cur}^{j,k}$, i.e., the sum of all source amplitude coefficients, is a measure of the prediction efficiency of the model (TS05 plus linear regression).

4.3 Results

Quiet Daily variation in data and in TS05's predicted series. For each observatory, 96 series of 24 hourly values were computed for ΔQX, ΔQY, ΔQBX and ΔQBY according to Equation 3.2 (standard method to compute QD), one for each month. Figure 4.2 shows the series of means of each distribution (thick red and green lines) plus and minus their respective standard deviation as shaded areas. The representation of these series uses local time, so that the expected maximum effects are centred on local noon, due to maximum solar heating and ionization, and appear at the same time in different stations. Note the different relations between Local Time (LT) and Universal Time (UT) for the four observatories (Table 3.1).

The well-known Sq variation patterns for Northern Hemisphere are blurred due to seasonal and solar cycle related variabilities, but can still be recognized in Figure 4.2. In particular, BOU and NVS show ΔQX curves typical of stations located north of the Sq currents focus and COI and PAG show curves typical of Earth's locations very close to the Sq currents northern focus (e.g., Yamazaki and Maute, 2016; Torta, Curto, and Bencze, 1997). The ΔQY curves are very similar in all four observatories and typical of stations at the Northern Hemisphere. ΔQY maximum values between 8:00 and 9:00 LT record the passing of the north-to-south oriented meridional branch of the Northern Hemisphere's anticlockwise vortex of Sq currents, while minimum values close to 14:00 LT record passing of the south-to-north meridional branch.

Figure 4.2 also shows curves ΔQBX and ΔQBY from TS05 model. The TS05 model was fitted to storm-time data (it has few quiet data included) and it is not expected to reproduce correctly the quietest time magnetospheric signal. Green curves in Figure 4.2 are shown only for a better understanding of TS05 model and its predictions on Earth's surface. They display dawn-dusk and noon-midnight asymmetries for which FAC and PRC current systems are the main contributors (Tsyganenko, 2002a; Tsyganenko, 2002b). As mentioned in Chapter 3, these asymmetries are modelled by $\sim \cos\phi'$

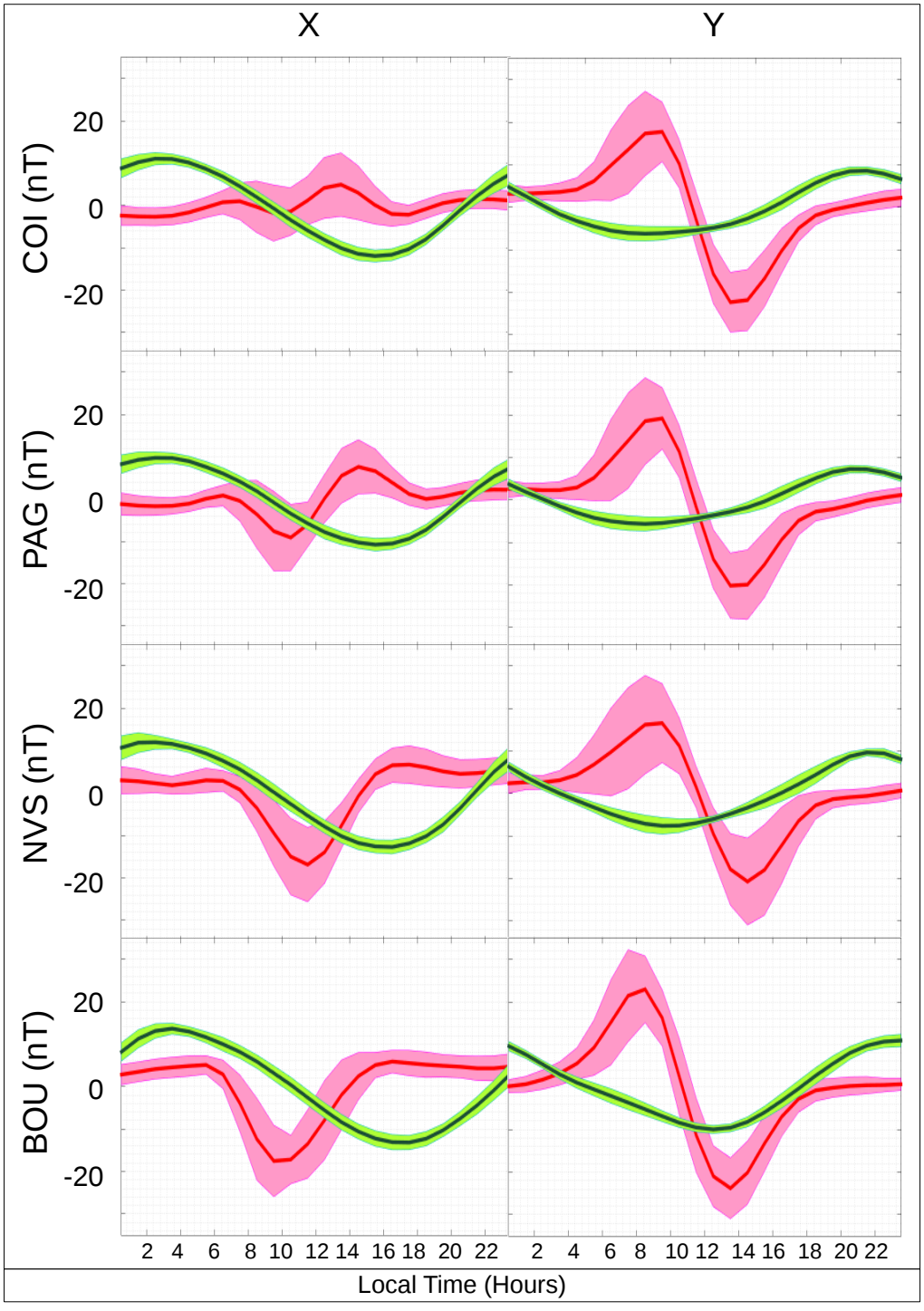


FIGURE 4.2: Distribution of 96 monthly ΔQ and ΔQB curves for the 2007-2014 period. Thick red and green lines represent averages of ΔQX (or ΔQY) and ΔQBX (or ΔQBY) distributions, respectively. In pink, dispersion of ΔQX and ΔQY curves within $\pm\sigma$ of the mean; in light green, dispersion of ΔQBX and ΔQBY within $\pm\sigma$ of the mean. For each observatory, curves for the North-South component on the left and for the East-West component on the right. Amplitude values are in nT.

and $\sim \sin\phi'$ terms, where ϕ' is directly related to the Magnetic Local Time (MLT) (e.g., Laundal and Richmond, 2016) expressed in angular units or geomagnetic longitude. For the observatories in this study, MLT differs no more than 0.5 hr from LT (Baker and Wing, 1989). Both FAC and PRC contribute with similar amplitudes during quiet days and Figure 4.2 shows, for X and Y, the sum of $\sin\phi'$ (PRC) with $\cos\phi'$ (FAC) and $\cos\phi'$ (PRC) with $-\sin\phi'$ (FAC) terms, respectively (see also Figure 4.3). Note that there is almost no seasonal variation of the magnetospheric QD, as shown by the small dispersion of curves in Figure 4.2. Different studies estimate contributions of ~ 5 nT from magnetospheric current systems during geomagnetic quiet conditions, at midlatitudes, and at the altitude of LEO satellites (e.g., Olsen and Stolle, 2016; Luhr et al., 2016). The ΔQBX and ΔQBY curves shown in Figure 4.2 are clearly too much strong in amplitude, which may be due to the lack of quietest days data in the tuning of TS05 model parameters and the closing of FAC currents through the Earth's centre.

To have an idea of the amplifying factor, we computed the standard spectral coherence between ΔQX (ΔQY) and ΔQBX (ΔQBY) curves. The coherence signal is spread and always less than 0.1, a too low value to be indicative of relevant coherence. However, a magnetospheric contribution is expected in ΔQX (ΔQY), even if very faint. Assuming the shape of the QD magnetospheric signal is correctly modelled by TS05, it can still be present in observations as long as its amplitude is within the error of the observed QD variation.

Figure 4.3 shows monthly QD series of all COI TS05's sources for 2007-2014. PRC and FAC are the main contributors to magnetospheric QD, with similar amplitudes. In fact, the total magnetospheric QD signal shows, for X and Y, the sum of $\sin\phi'$ (PRC) with $\cos\phi'$ (FAC) and of $\cos\phi'$ (PRC) with $-\sin\phi'$ (FAC), respectively, as mentioned before.

Performance of TS05 Model. The distributions of rX, $\widetilde{rX}$, rY and $\widetilde{rY}$ correlation values are represented in Figure 4.4 for all stations, in the form of histograms normalized

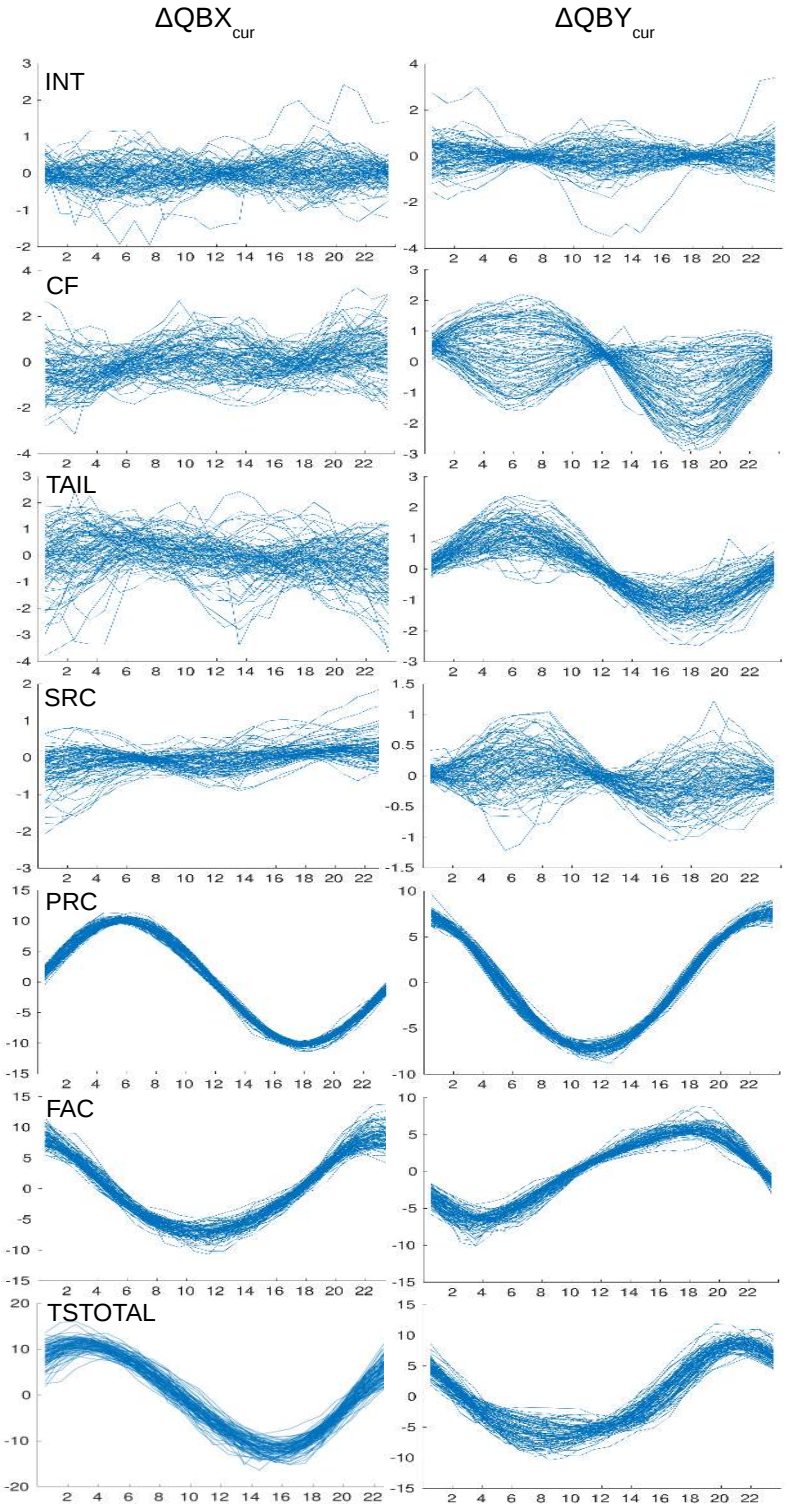


FIGURE 4.3: Sets of 96 monthly ΔQBX_{cur} (left) and ΔQBY_{cur} for each TS05's sources and TSTOTAL at COI, for 2007-2014. Vertical axis represents amplitudes in nT and horizontal axis represents hours in local time. It is clear that main contributors to ΔQB are FAC and PRC because of their higher amplitudes.

in order that the total area is one. Results for less and more active days are shown in separate histograms. In each station, there is a tendency for an increase in the number of higher correlation values after removing the QD variation in raw data and simulations ('corrected' series), which is seen in the form of an increase of skewness-to-the-left of the histograms of $\widetilde{rX}$ ($\widetilde{rY}$) compared to those of rX (rY). The improvement is particularly noticeable for less active days, for which the amplitude of X and Y variations is smaller and as a result the fraction of the QD signal is higher and the correction is more relevant. Amongst the two components, Y (Figure 4.4, bottom half) shows the most striking improvements after QD removal during more quiet days, in particular for PAG (but also COI and NVS), with the histogram of correlation coefficients changing from skewness to the right to skewness to the left. This can be understood due to higher amplitude of ΔQY compared to ΔQX (see Figure 4.2).

The significance of each r coefficient is evaluated through a corresponding p-value, which tests the null hypothesis that r is zero. Student's t-test was applied with n=24 samples used in the computation of the correlation coefficient (see Equation 2.32). Using this test, values of $r \geq 0.403$ have probability $\geq 95\%$ to be different from zero, corresponding to $p \leq 0.05$ (red bars in Figure 4.4). It is apparent that $\widetilde{rX}$ ($\widetilde{rY}$) values are shifted towards higher and more significant values when compared with rX (rY).

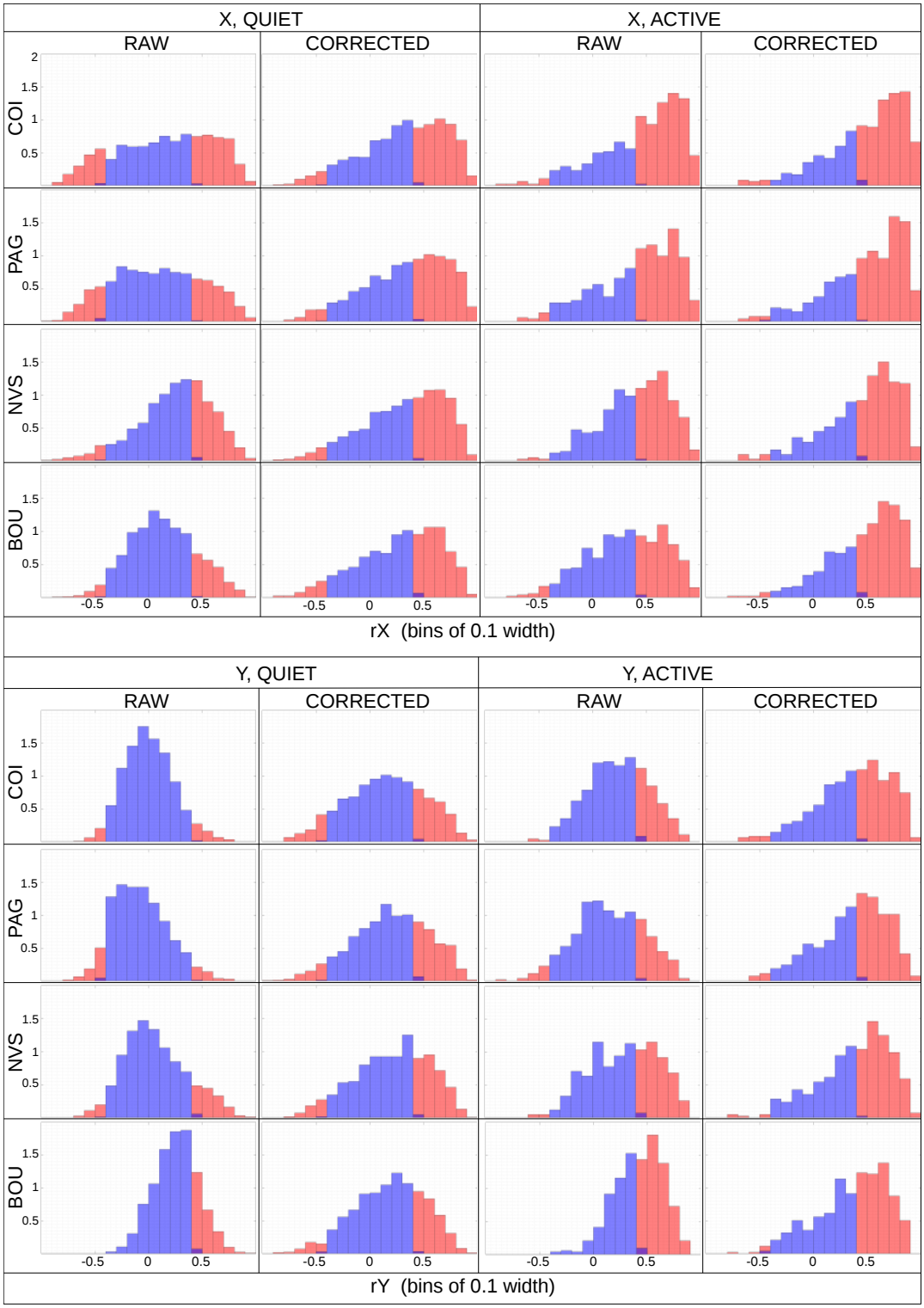


FIGURE 4.4: For each observatory, distribution of correlation coefficients rX (rY) (first and third columns) and $\widetilde{rX}$ ($\widetilde{rY}$) (second and fourth columns). Results for X-series in the top half and for Y-series in the bottom half. Separate results for quiet (first and second columns) and active days (third and fourth columns). Bar height values given by $N_i/(\sum_i N_i \Delta r)$, with N_i the number of correlation values in bin i and Δr the bin width (0.1). Blue areas for r with p-level values $p > 0.05$ and red areas for $p \leq 0.05$.

TABLE 4.1: Statistical parameters describing the distribution of correlation coefficients rX (rY) (left) and $\widetilde{rX}$ ($\widetilde{rY}$) (right). N is the number of correlation values with p-value ≤ 0.05; $\bar{r}$ is the mean correlation value computed using only significant correlation coefficients; r_0 is the inverse Fisher's transform of $\bar{z}$, the mean of the Gaussian distribution fitted to the histogram of z-values. After correcting for the QD variation (right), the percentage of N with $\widetilde{r} \geq 0.7$ and the maximum value obtained for $\widetilde{r}$.

		Before removing QD									**After removing QD**							
Component	Parameter	COI		PAG		NVS		BOU		Parameter	COI		PAG		NVS		BOU	
		quiet	*active*	*quiet*	*active*	*quiet*	*active*	*quiet*	*active*		*quiet*	*active*	*quiet*	*active*	*quiet*	*active*	*calm*	*active*
X	N	921	322	732	330	806	235	441	250	N	957	330	980	360	1027	277	903	358
	$\bar{r}$	0.25	0.64	0.19	0.62	0.45	0.62	0.43	0.57	$\bar{r}$	0.53	0.66	0.56	0.66	0.54	0.63	0.50	0.66
	r_0	0.18	0.53	0.10	0.49	0.28	0.45	0.15	0.37	r_0	0.38	0.58	0.40	0.57	0.37	0.52	0.33	0.56
										% ($\widetilde{r} \geq 0.7$)	35%	51%	36%	53%	31%	40%	25%	45%
										r_{max}	0.96	0.96	0.98	0.97	0.97	0.96	0.95	0.97
Y	N	149	148	219	150	357	176	444	298	N	645	259	656	283	740	234	600	287
	$\bar{r}$	0.19	0.54	-0.14	0.41	0.38	0.57	0.52	0.58	$\bar{r}$	0.34	0.59	0.41	0.58	0.42	0.60	0.42	0.58
	r_0	0.01	0.26	-0.08	0.17	0.08	0.33	0.27	0.45	r_0	0.17	0.43	0.22	0.41	0.23	0.43	0.21	0.41
										% ($\widetilde{r} \geq 0.7$)	17%	35%	21%	28%	16%	30%	13%	26%
										$\widetilde{r}_{max}$	0.93	0.94	0.92	0.95	0.94	0.95	0.92	0.9

As seen in Fig.4.4 the mean of correlation coefficients, $\bar{r}$, increases with skewness. Table 4.1 shows information on some statistical parameters, as N, the total number of correlation values with p-value less or equal to 0.05, $\bar{r}$ and r_0 (Eq. 2.34), and also $\widetilde{r}_{max}$, the maximum value obtained for $\widetilde{r}$. As shown in this table, there is an increase in N (the number of significant correlation coefficients) and an increase in the correlation parameters $\bar{r}$ and r_0 after correcting the data for the QD variation. All these changes are particularly large for less active days. For each table, r_0 is always lower than $\bar{r}$.

For all four stations, highest correlation values are obtained when $\widetilde{BX}$ and $\widetilde{X}$ series are compared. Representative global values of r_0 for COI, PAG and BOU are then as high as 0.6 for more active days. Lowest values of 0.5 and 0.6, respectively, occur for NVS. If using instead the mean of $\widetilde{r}$ we obtain values of $\bar{r}$=0.7 at all stations except for NVS, for active days. For quiet days, though showing lower correlation values, their increase after removal of daily variation is also clear. The Y component always shows lower correlation values than X, though clearly improved after correcting for the QD variation. The only exception occurs with BOU Y-component that shows relatively high values for rY, which do not increase after QD correction.

From Figure 4.4 and Table 4.1 we note as a recurrent feature that correlations are smaller for lower than for higher levels of geomagnetic activity. This is, of course, to be expected since the TS05 model was designed to be representative of storm-time geomagnetic activity. Nonetheless, even for relatively low activity, both correlation values r_0 and $\bar{r}$ are equal or larger than 0.3 for the X and 0.2 for the Y components, after correcting for daily variation. These results seem to indicate that TS05 model predictions for quiet magnetospheric activity still capture some of the observed behaviour, even if the model has not taken into account quiet days data. Nonetheless, TS05 model was clearly not thought to explain quiet days observations and it estimates incorrectly the contribution of at least some current sources.

A comprehensive description of geomagnetic activity. Scatter plots in Figure 4.5 show sets of correlation coefficients $\widetilde{rX}$ ($\widetilde{rY}$) against source correlations $\widetilde{rX}_{cur}$ ($\widetilde{rY}_{cur}$) for all days in the 'corrected' histograms of Figure 4.4. These plots are indicative of how the variance in each $\widetilde{r}_{cur}$ is related to the variance of $\widetilde{r}$, i.e., the sensitivity of global $\widetilde{r}$ to each separate $\widetilde{r}_{cur}$ (Saltelli et al., 2008). The slope of regression lines gives insight into the relative magnitude of the sensitivities. INT and CF are shown in Figure 4.6, where $\widetilde{r}$ are weakly correlated with $\widetilde{r}_{cur}$ (small slopes). Two parameters are used to help compare different scatter plots, namely the slope of the linear regression line and the location of the center of the scatter cloud ($\bar{r}_{cur}$,$\bar{r}$).

Figure 4.5 shows, using descriptive analysis, that $\widetilde{rY}$ is globally more sensitive to changes in the FAC field than to changes in any other magnetospheric source field (higher values of the slope). This is true for all observatories, showing that a main source of error for BY predictions may be an incorrect assessment of FAC contributions. Differently, the sensitivity of $\widetilde{rX}$ to FAC currents is weaker than $\widetilde{rY}$ for active days, and FAC, TAIL, SRC and PRC have all similar influence over $\widetilde{rX}$ (closest slopes). All sources show, at all observatories and for both (X, Y) components, a shift of the scatter cloud towards values with higher $\widetilde{r}$ for more active days, in agreement with results in Figure 4.4 and Table 4.1.

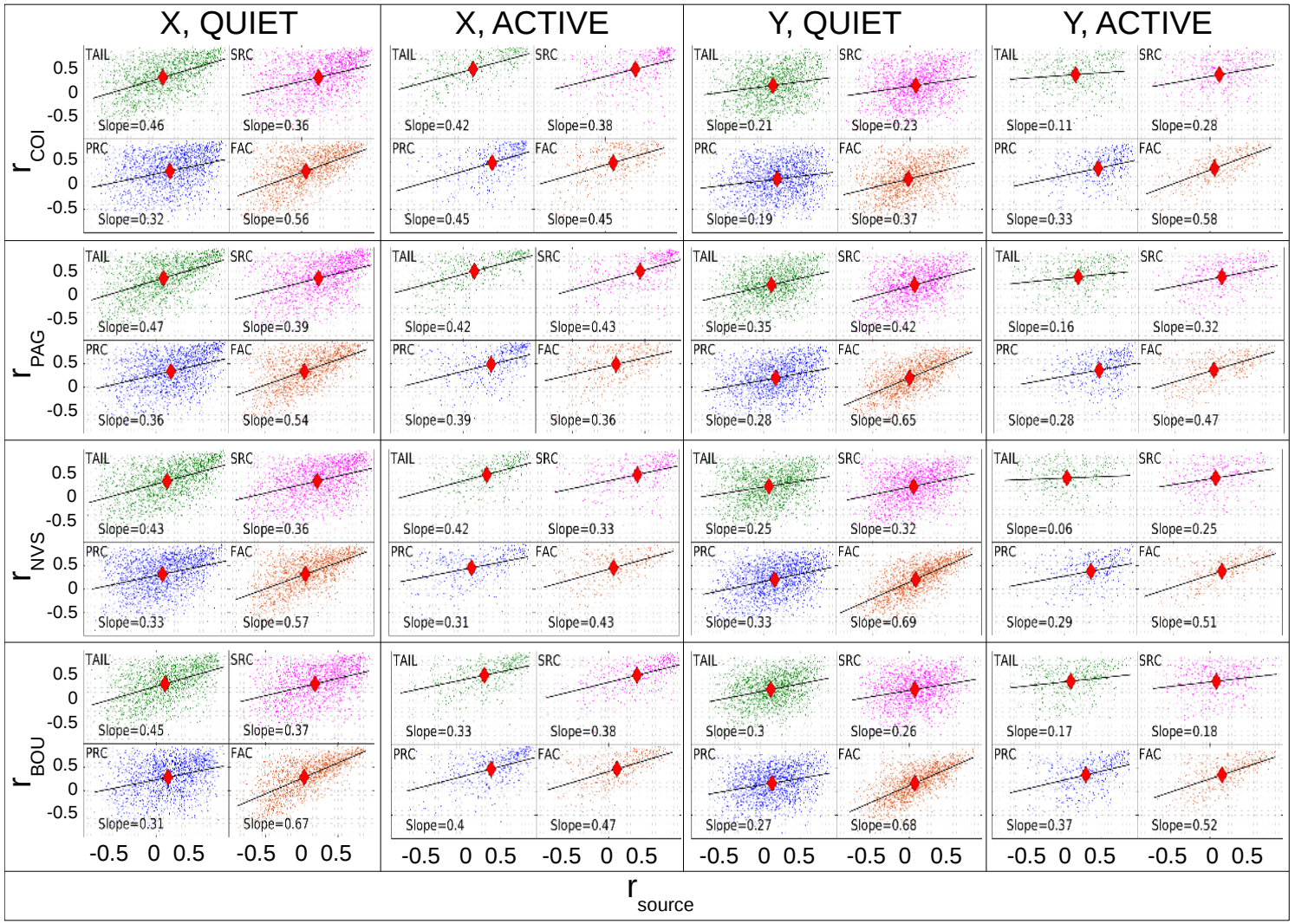


FIGURE 4.5: Scatter plots of $\widetilde{r}$ against $\widetilde{r}_{cur}$ for TAIL, SRC, PRC and FAC source components at every station (from top to bottom, COI, PAG, NVS and BOU). $\widetilde{rX}$ against $\widetilde{rX}_{cur}$ (left) and $\widetilde{rY}$ against $\widetilde{rY}_{cur}$ (right). Values of the slope, as inset. Red diamonds for the location of the cloud center-of-distribution.

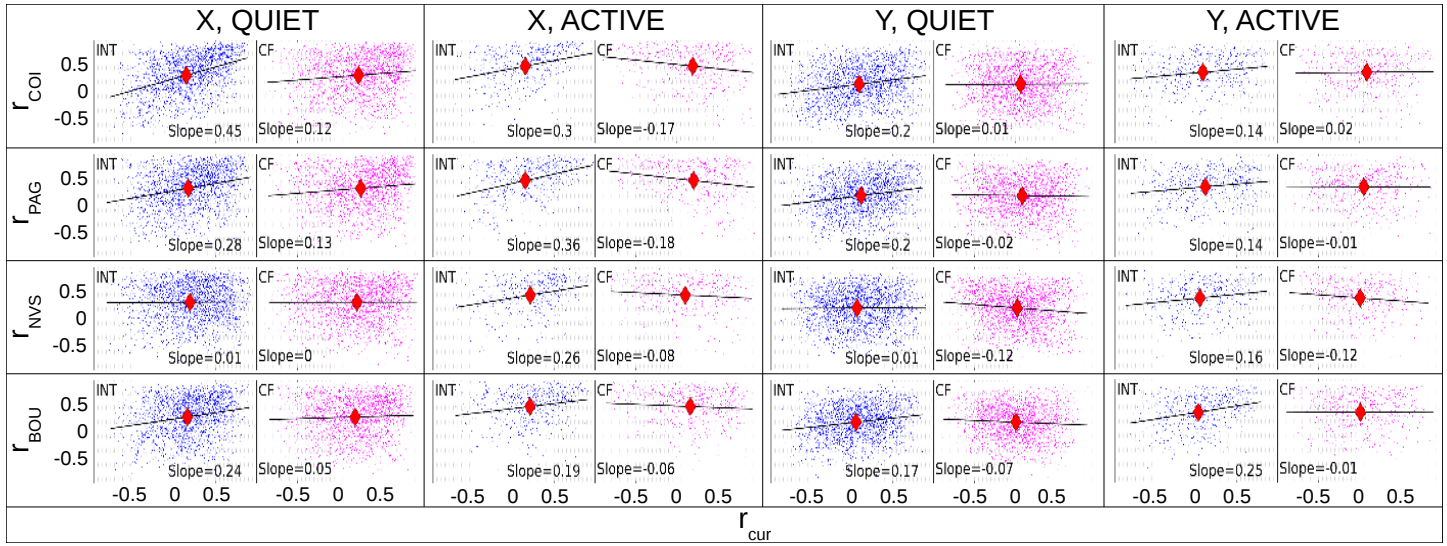


FIGURE 4.6: Similar to Figure 4.5, but for INT and CF source components at every station (from top to bottom, COI, PAG, NVS and BOU).

Days with correlation values $\widetilde{r} \geq 0.7$ represent an important percentage of the total amount of analysed days (see Table 4.1), between 380 to 540 daily series for X and 150 to 210 for Y, depending on the observatory. For these days, the TS05 model reaches the best performance in explaining magnetospheric activity and the associated current distribution is particularly meaningful to explain Earth's ground observations. Using only days with $\widetilde{r} \geq 0.7$, the mean values of different $\widetilde{r}_{cur}$ are computed and represented in the form of bar charts in Figure 4.7. In general, the distribution of mean source corre-

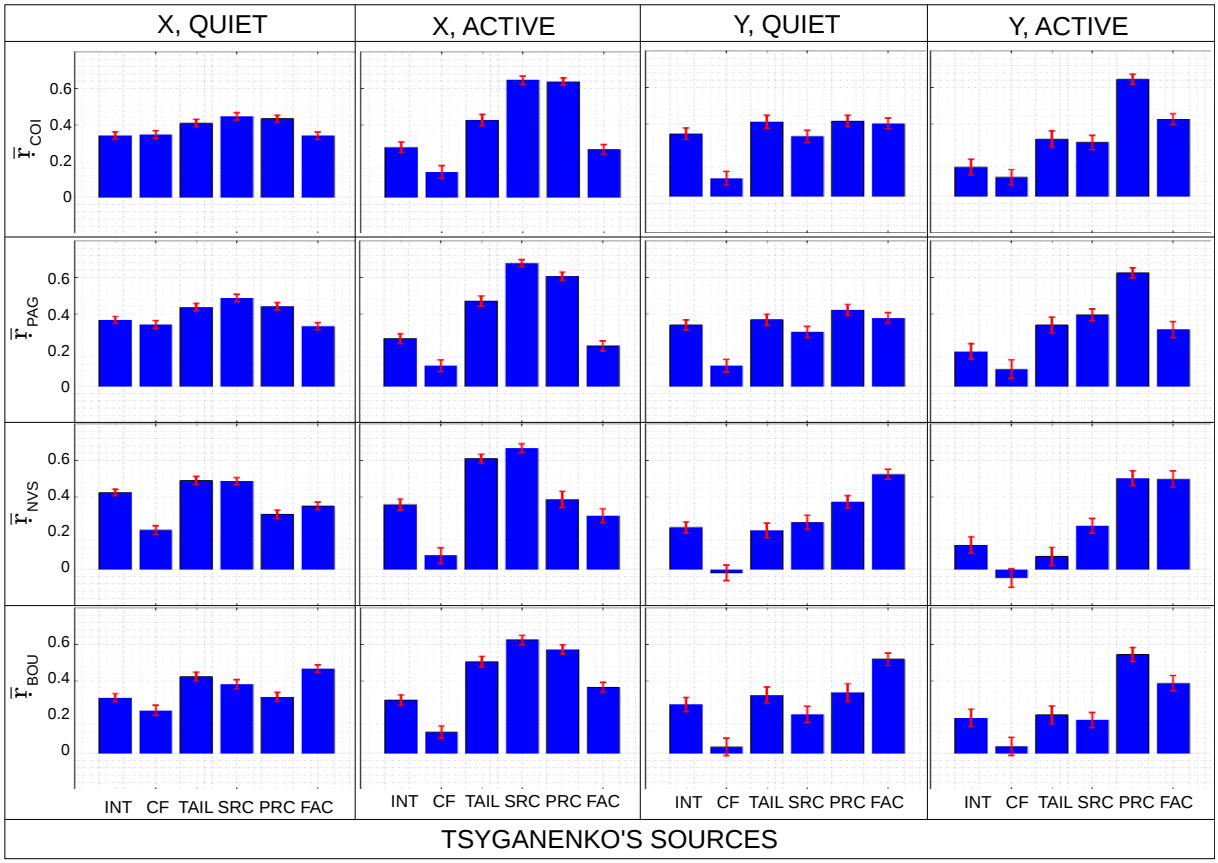


FIGURE 4.7: Contribution of TS05's sources to the shape of data series for every station, in active and quiet days with $r \geq 0.7$ (from top to bottom, COI, PAG, NVS and BOU). Separate results for mean values of $\widetilde{rX}_{cur}$ (left) and $\widetilde{rY}_{cur}$ (right). Also shown, the standard error of the mean as red bars.

lations is very similar between COI and PAG and also quite similar between BOU and

NVS but different between the two groups. Nonetheless, for the X component during disturbed days a very similar pattern of current contributions is nicely retrieved at all four stations. The dominant contribution of TAIL, SRC and PRC on the shape of X-data series is seen for all observatories, especially during more active days. Differently, the combination of sources contributing to the shape of Y-data series has an increased influence of the FAC + PRC system while the TAIL + SRC current system is weakened. At BOU and NVS, the decrease of TAIL + SRC contribution is particularly noticeable. For all stations, the relative importance of FAC + PRC system for Y series is stronger during periods of high activity. The INT term in equation 2.7 represents the penetration of the IMF into the magnetosphere. Its contribution is the most stable amongst different current systems. As to contribution of currents in the magnetopause (CF), its mean effect is the smallest one because sometimes it correlates with the global signal and others it is anti-correlated.

Global normalized amplitudes $\widetilde{a}$ are a measure of the prediction efficiency of both the linear regression model and TS05, to explain data series with QD and IGRF removed. Scatter plots of $\widetilde{a}$ against $\widetilde{a}_{cur}$ (equations 4.8 and 4.9) are represented in Figure 4.8 and give insight into how $\widetilde{a}$ is influenced by the strength of contribution of different sources to the model series. Scatter clouds show high degree of heteroscedasticity (changing variance over the entire range of $\widetilde{a}_{cur}$ values) and the slope of linear regression computed by ordinary least squares has no special meaning. The center of the cloud distribution is however shown, for comparison between different plots.

X and Y components show very different behaviour, with the center of X scatter clouds closer to $\widetilde{a}$ = 1 than center of Y scatter plots. In fact, while hourly variations of source contributions for X series are often smaller than the mean of the whole signal, for Y series they tend to be larger (see Figure 4.11) because their mean is close to zero. As a result, it is possible to have $\widetilde{aX}$ close to 1 even for days when $\widetilde{rX}$ is relatively small. High $\widetilde{aY}$ values are more difficult to occur, because they require high $\widetilde{rY}$ values. In any case (X or Y series), a $\widetilde{r}$=1 value would mean that predictions and observations

were related by an exact linear model and would also imply $\widetilde{a}$=1. Figure 4.8 shows that the two current systems TAIL + SRC and FAC + PRC do exhibit a very distinct behavior. Amplitudes $\widetilde{aX}$ are weakly sensitive to $\widetilde{aX}_{PRC}$ and $\widetilde{aX}_{FAC}$ (axis-aligned spread), but strongly sensitive to $\widetilde{aX}_{TAIL}$ and $\widetilde{aX}_{SRC}$ (closely diagonal spread), both during less and more active days. Low $\widetilde{aX}$ values occur when TAIL and SRC have small contributions. When they become dominant, then $\widetilde{aX} \simeq 1$. This is indicative of a particularly good modelling of TAIL and SRC currents, in particular TAIL. The sum $\widetilde{aX}_{TAIL}+\widetilde{aX}_{SRC}$ can even be larger than $\widetilde{aX}$, because of the negative contribution of CF (see Figure 4.9). As for $\widetilde{aY}$, the relative influence of the FAC + PRC system gets closer to that of TAIL + SRC, in particular for disturbed days, but the center of scatter clouds is also particularly low, there. This means that problems related with the modelling of FAC and PRC current systems, especially FAC, are strongly affecting the Y component at these two stations.

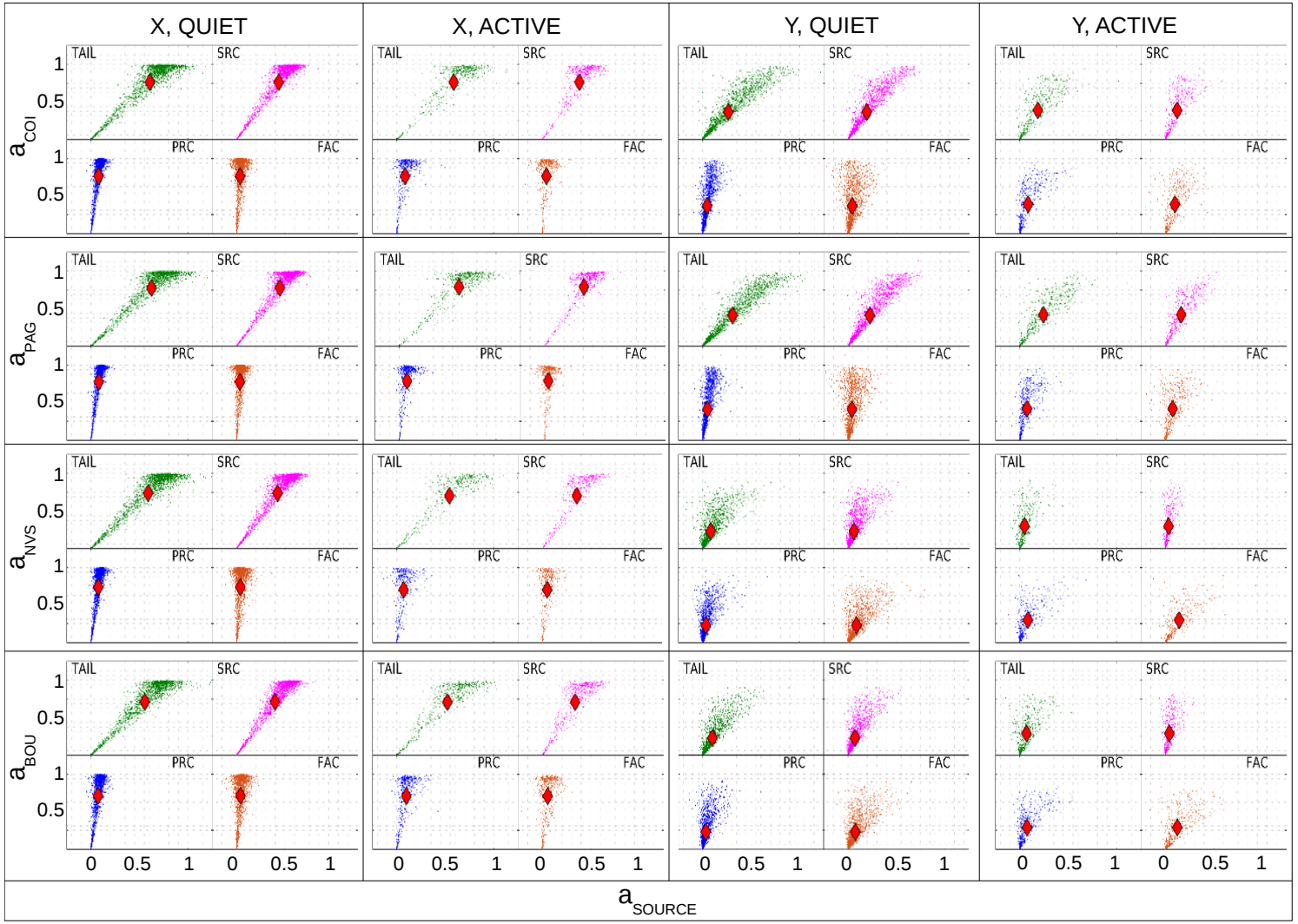


FIGURE 4.8: Scatter plots of $\widetilde{a}$ against $\widetilde{a}_{cur}$ for every TS05's source component and for every station (from top to bottom, COI, PAG, NVS and BOU). $\widetilde{aX}$ against $\widetilde{aX}_{cur}$ (left) and $\widetilde{aY}$ against $\widetilde{aY}_{cur}$ (right), for active days. Values of the slope, as inset. Red diamonds for the location of the cloud center-of-distribution.

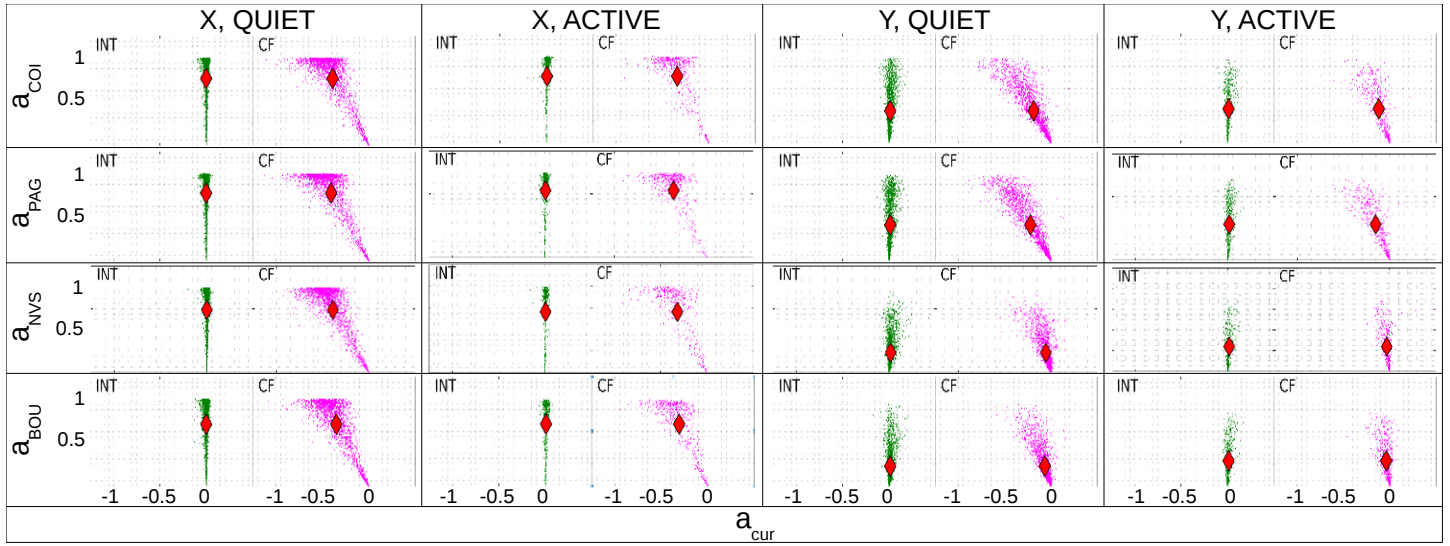


FIGURE 4.9: Same as Figure 4.8, but for INT and CF sources.

In Figure 4.10, only days with $r \geq 0.7$ contribute to the bar plots showing the means of each set of $\widetilde{a}_{cur}$ coefficients. These are the same days as those represented in Figure 4.7, for which the TS05 model performs the best. The same configuration of source amplitudes is nicely retrieved for the X-component at all four stations, during both quite and more active periods. The field contribution from CF current has always amplitude similar to that of SRC, but with different sign and INT is just residual. As already shown in Figure 4.8, the TAIL + SRC system contribution is the most important for $\widetilde{aX}$. As for the Y-component, the combination of contributing amplitudes shows the TAIL + SRC current system weakened in favour of an increased contribution of the FAC + PRC system, especially for disturbed periods and most particularly at NVS and BOU.

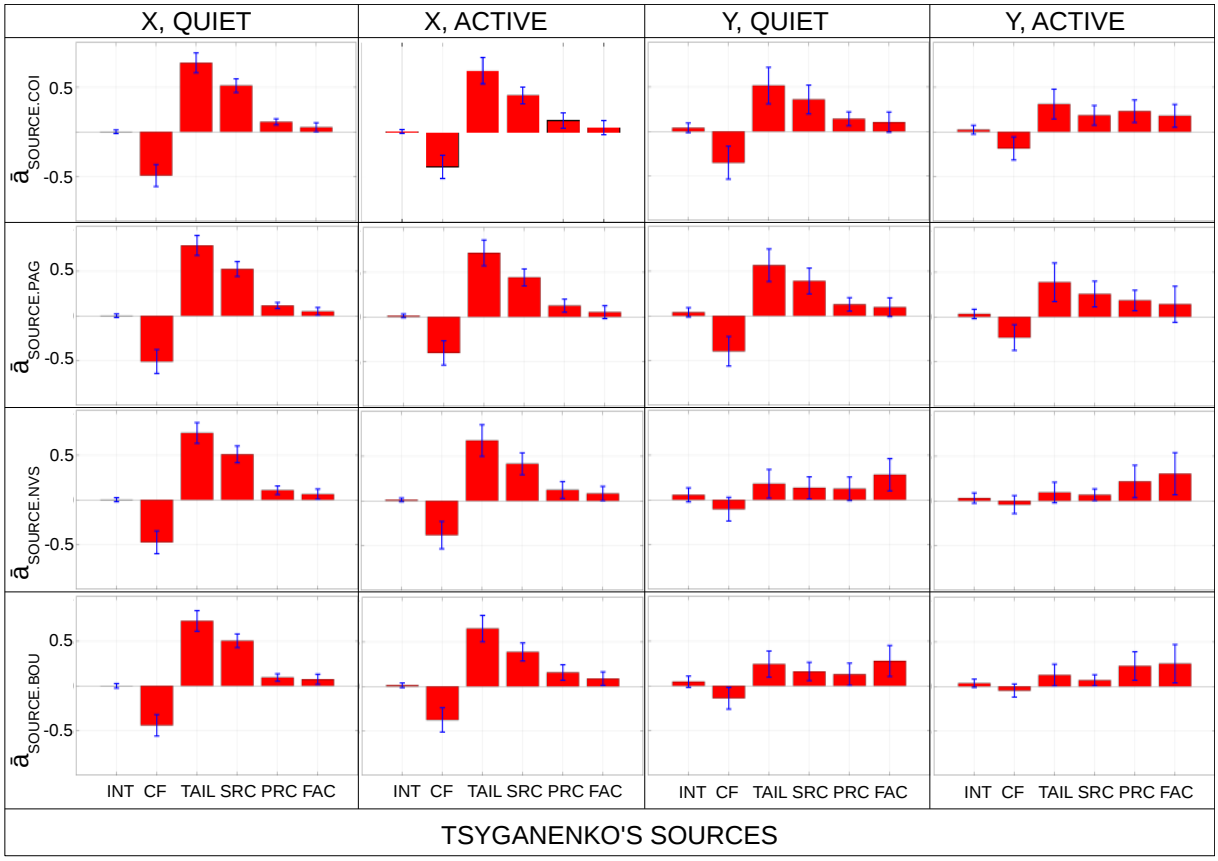


FIGURE 4.10: Contribution of TS05's sources to normalised amplitude coefficient $\widetilde{a}$, using only days with $\widetilde{rX}$ $(\widetilde{rY}) \geq 0.7$, for every station (from top to bottom, COI, PAG, NVS and BOU). Separate results for mean values of $\widetilde{aX}_{cur}$ (first and second columns) and $\widetilde{aY}_{cur}$ (third and forth columns) and also for quiet (first and third columns) and active days (second and fourth columns). Also shown, the standard error of the mean as blue bars.

Figure 4.11 shows observed (black dotted lines) and predicted (black thick lines) series for the four observatories that illustrate the global behaviour described above. Both the estimated amplitude and phase roughly agree with the data at all observatories. For X, during the quiet time a daily variation is clearly seen in $\widetilde{BX}$, with main contributions from FAC and PRC. A similar daily variation is seen in the data. During the more disturbed period, the shift of $\widetilde{BX}$ strength to more negative values is mainly accounted for by the TAIL + SRC system. PRC and FAC are however crucial to explain higher frequency features of the time series, most of them also present in observations. The temporal variability in TAIL + SRC is also important to account for the shape of the series, in particular for $\widetilde{BX}$ during the disturbed period. Also shown the residual amplitude of INT and the difference in signs, with similar amplitudes, between CF and both SRC and TAIL. For the Y component, the data show a daily variation during the quiet period which is not so clearly retrieved from TS05. TAIL + SRC are the main contributors to strength during quiet time, but FAC + PRC are very important to explain the shape of the series. For the active days all currents contribute to the signal strength, but FAC is the one that best follows the shape of the ground measurements.

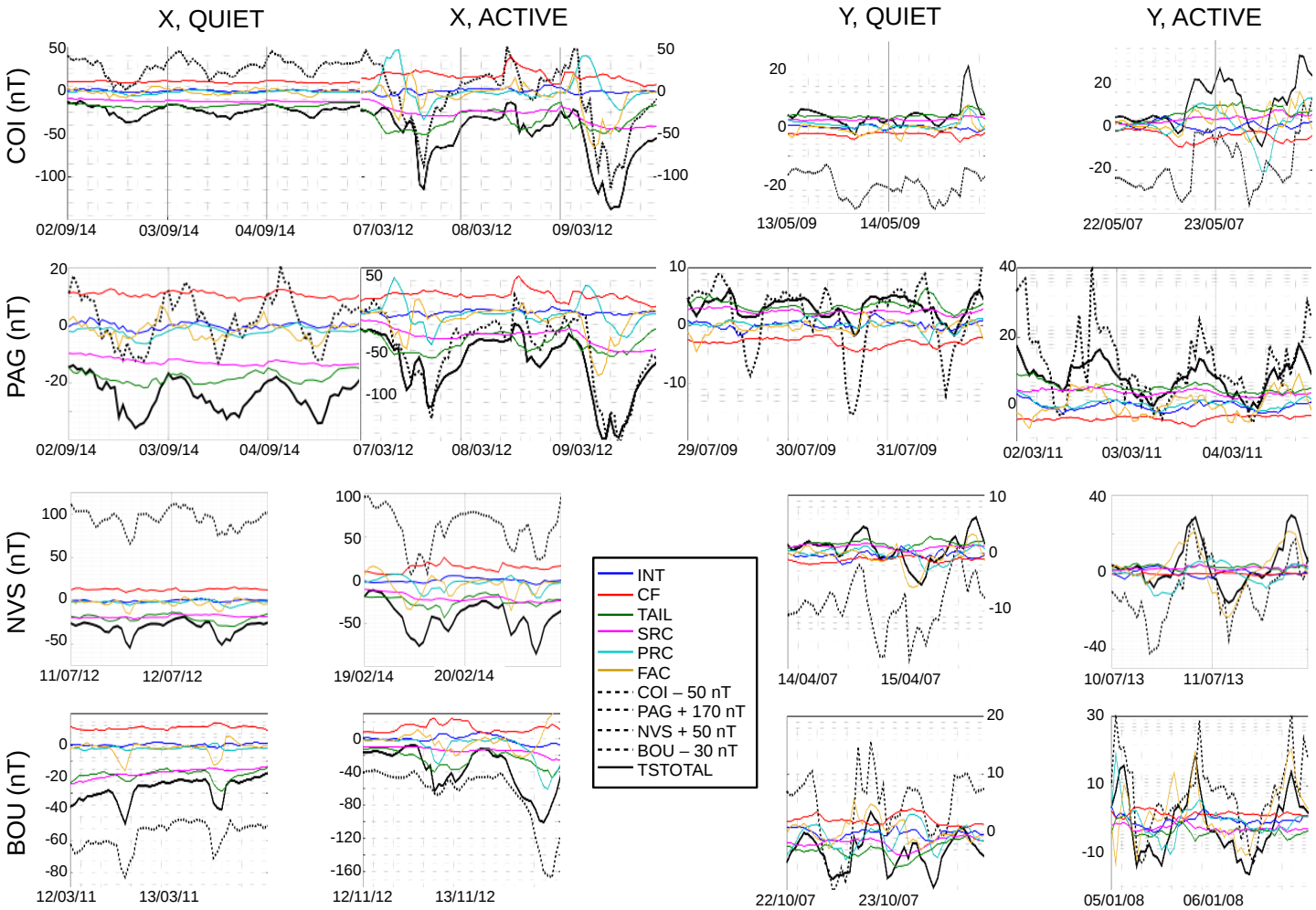


FIGURE 4.11: First two columns: Case examples comparing $\Delta\widetilde{X}$ (dotted black line), $\Delta\widetilde{BX}$ (thick black line) and $\Delta\widetilde{BX}_{cur}$ (see legend), for 2-3 consecutive days at quiet (first column) and disturbed time (second column) at days with $r\widetilde{X} \geq 0.7$, for the four observatories, from top to bottom: COI, PAG, NVS and BOU. Third and fourth columns: the same as first two columns for $\Delta\widetilde{Y}$, $\Delta\widetilde{BY}$ and $\Delta\widetilde{BY}_{cur}$. All data series are shifted along the vertical axis for easier visual comparison between simulation and data series.

TABLE 4.2: Means and standard deviations of α and β parameters fitted through least squares regression (equations 4.3 to 4.5), for different observatories. Also shown, mean values of amplitude coefficients $\widetilde{aX}$ and $\widetilde{aY}$ (equation 4.8).

Parameter	COI		PAG		NVS		BOU	
	quiet	*active*	*quiet*	*active*	*quiet*	*active*	*quiet*	*active*
$\overline{\alpha}_X$	13.21	10.99	-132.83	-135.29	81.00	77.51	3.88	2.02
σ_{α_X}	12.32	12.33	13.49	14.86	11.15	13.64	11.29	11.3
$\overline{\beta}_X$	1.26	1.17	1.27	1.24	1.10	1.09	1.14	1.03
σ_{β_X}	0.55	0.37	0.59	0.45	0.51	0.42	0.48	0.36
$\overline{aX}$	0.97	0.94	0.97	0.94	0.96	0.92	0.96	0.92
$\bar{\alpha}_Y$	-64.43	-63.23	-174.85	-174.76	-60.55	-60.28	38.28	37.98
σ_{α_Y}	11.48	9.31	5.88	7.05	2.99	4.38	3.72	4.02
$\overline{\beta}_Y$	1.55	1.28	1.55	1.48	1.53	1.52	1.44	1.20
σ_{β_Y}	0.76	0.58	0.84	0.64	0.97	0.80	0.69	0.47
$\overline{aY}$	0.81	0.75	0.83	0.77	0.67	0.67	0.72	0.67

Amplitude normalised parameters. To compute the normalised source amplitude coefficients shown in Figures 4.8, 4.9 and 4.10, a least squares linear regression of predicted series $\Delta\widetilde{BX}$ ($\Delta\widetilde{BY}$) to corresponding observations $\Delta\widetilde{X}$ ($\Delta\widetilde{Y}$) (equations 4.3 to 4.5) was applied, for days with $\widetilde{rX}$ ($\widetilde{rY}$) ≥ 0.7. The regression was made for each day separately, giving a distribution of α and β parameters from which the ensemble mean and standard deviation were computed and listed in Table 4.2. Global amplitude parameters aX and aY (equation 4.8) are used to measure the prediction efficiency of the linear regression model plus TS05, when explaining the data series with QD and IGRF removed. Corresponding means and standard deviations are shown in Table 4.2. In agreement with the general better performance of TS05's X component, $\overline{aX}$ values are closer to one than $\overline{aY}$ values.

4.4 Discussion

Performance of TS05 Model. The results in this chapter show that between January 2007 and December 2014, the TS05's predicted $\widetilde{BX}$ and $\widetilde{BY}$ series reproduce the shape of the QD-corrected signal ($\widetilde{X}$ and $\widetilde{Y}$) at the four ground-based stations COI, PAG, NVS and BOU with correlation values $\widetilde{r} \geq 0.7$ in $\sim$ 50% (X) and $\sim$ 30% (Y) of geomagnetically active days and in $\sim$ 30% (X) and $\sim$ 15% (Y) of geomagnetically calm days (Table 4.1). Maximum $\widetilde{r}$ values are at least 0.95 (X) and 0.90 (Y), depending of the station. For $\widetilde{r} \geq 0.7$, a linear regression model explains quite well the remaining differences between TS05's estimations and observations, especially with respect to the X component (Table 4.2). The estimated scale and shift parameters can be given a physical explanation, as explained below. The separate contribution of individual sources in explaining the shape and magnitude of the transient series at the chosen sites was assessed by defining source correlation and source amplitude parameters.

One of the main, large, structures in magnetospheric currents is the TAIL current which merges with SRC at the night side inner boundary (Tsyganenko, 2013). Close to Earth, the geomagnetic field due to these currents is nearly parallel to the dipole axis and at Earth's surface mid-latitudes they contribute dominantly to the North-South (X) component of the measured field. This is reproduced in TS05 model predictions, where the TAIL + SRC current system is the main contributor to the magnitude values $\widetilde{aX}$ series at the four observatories (Figures 4.8 and 4.10). High mean values for $\widetilde{rX}_{TAIL}$, $\widetilde{rX}_{SRC}$ and $\widetilde{rX}_{PRC}$ in Figure 4.7 are indicative of a dominant contribution of these three sources to the time variability of X data series. Although $\widetilde{aX}_{PRC}$ values are low (small contribution to signal strength), PRC is important to explain the time variability in the X-component. (Dubyagin et al., 2014).

The FAC + PRC coupled system is the other large magnetospheric current system showing up in the analysis, with a main contribution for the hourly to diurnal temporal variability of Y observatory series, mainly during more active days (Figures 4.5

and 4.7). The local daily variation produced by these currents at Earth's surface can be seen in observations both in X and Y components, during quiet geomagnetic activity. During active days the magnetospheric daily variation may become less discernible due to a stronger superposed signal driven by high frequency solar wind and IMF disturbances. The high frequency content of FAC and PRC contributions as seen in Figure 4.11 is explained by these source drivers but also by higher decay rates (smaller relaxation times) in the TS05 model for FAC and PRC compared to SRC which evolves much more smoothly in time (Tsyganenko and Sitnov, 2005). For X-series though $\widetilde{rX}_{PRC}$ values are high, $\widetilde{aX}_{PRC}$ and $\widetilde{aX}_{FAC}$ contributions are small compared with TAIL + SRC (Figures 4.8 and 4.10), in every station. For Y-series, $\widetilde{a}$ but specially $\widetilde{r}$ contributions due to FAC + PRC are dominant. The larger contribution of FAC + PRC and smaller contribution of TAIL + SRC at NVS and BOU with respect to COI and PAG (see Figures 4.7 and 4.10) may be explained by slightly higher geomagnetic latitudes of BOU and NVS or higher proximity of these two observatories to the geomagnetic meridian where larger variations of the dipole tilt angle occur each day.

The penetration of the IMF into the magnetosphere is introduced by the INT term in equation 2.7, representing a simple uniform magnetic field equal to a fixed fraction of the transverse (to the solar wind) component of IMF in the GSM system (Tsyganenko, 2013) (see eq. 2.16). It has almost no contribution to the strength of the signal (as given by $\widetilde{a}_{INT}$ in Figure 4.10), but has a significant contribution to its shape (as given by $\widetilde{r}_{INT}$ in Figure 4.7). The CF contribution is defined by the shape, size and orientation of the magnetopause, which depends both on the Earth's dipole field and solar wind parameters (Tsyganenko and Andreeva, 2015). It has relatively high $\widetilde{aX}_{CF}$ values with negative sign in Figure 4.10. It often shows some anti-correlated features with TAIL and SRC, leading to a high dispersion of $\widetilde{r}_{CF}$ between positive and negative values. This is true even for $\widetilde{r} \geq 0.7$, explaining the relatively small mean values in Figure 4.7.

The TAIL current system deserves some more attention. Our results show that, together with SRC, it has a relevant contribution to both the magnitude and shape of the

X component of the magnetospheric signal on the ground. Compared to SRC, the TAIL current does more quickly respond to solar-wind and IMF changes, as confirmed by the relatively short relaxation time obtained by Tsyganenko and Sitnov, 2005 when fitting the model to storm-time observations. A distinctive feature regarding the modelling of TAIL in TS05, is that it is the only current system driven by the Dst/SYM-H index besides solar-wind and IMF parameters. In particular, SYM-H enters the model to control the distance to the Earth of the tail current sheet, along the tail axis (Tsyganenko and Sitnov, 2005). In this way the TAIL system is directly constrained by a parameter that is very closely related to on-ground X-component observations (see e.g., Yamazaki and Maute, 2016). The utilization of SYM-H as one of TAIL free parameters when fitting satellite observations in TS05 model may be one reason to explain why TS05's estimations show better results for the X than for the Y components. Another reason can come from the satellite data spatial distribution used to fit the model parameters. A major amount of these data are provided by geosynchronous satellites GOES 8, 9 and 10 nightside observations (Tsyganenko and Sitnov, 2005) and are expected to give more detailed information on equatorial currents, which are those that contribute most to the observed X component. Last but not least, since poloidal FAC currents have a major contribution to observed Y components, their oversimplified closure geometry through the Earth's center is expected to mainly affect the Y component estimation at Earth's surface.

On-ground inducing effect of magnetospheric fields. Due to finite conductivity of the crust and upper mantle, we expect a scaling factor of of ~ 1.3 relating $\Delta\widetilde{X}$ ($\Delta\widetilde{Y}$) to $\Delta\widetilde{BX}$ ($\Delta\widetilde{BY}$) daily series and a negligible phase-lag between the two signals (e.g., Schmucker, 1985; Olsen, Sabaka, and Lowes, 2005), i.e., $\Delta\widetilde{X} \simeq (1+Q)\Delta\widetilde{BX}$, with $Q \simeq 0.3$.

A linear regression model was computed to relate the set of $\Delta\widetilde{BX}$ and $\Delta\widetilde{BY}$ daily

series to corresponding series $\Delta\widetilde{X}$ and $\Delta\widetilde{Y}$. Separate results for different observatories, different components and different geomagnetic activity are shown in Table 4.2. A global β parameter can be computed averaging all separate estimates and using as a normalized weight, $\sigma_{\beta_i}^{-1} / \sum_k \sigma_{\beta_{i,k}}^{-1}$, where 'i' stands for the X or the Y components and the sum is over all 16 separate estimations. This gives β_{glob}=1.26. Such a value is in a very good agreement with estimations from previous studies mentioned above, i.e., Q = 0.27 for induced fields, as mentioned by Olsen, Sabaka, and Lowes (2005). However, and as recurrent effect seen for all observatories, our estimates show that Q tends to be higher for the Y than the X components and, in each case, lower for more than for less intense geomagnetic activity. The somewhat higher value for $\overline{\beta}_Y$ can be simply due to less well-constrained magnetospheric variations in this component as predicted by TS05. However, Q ratio values depend on the geometry of the source field (Olsen, 1998). While most estimates for the Q ratio assume a dipole equatorial current-type source which may approximate the TAIL + SRC current system, our results show that the FAC + PRC system is the main contributor to the shape of the transient magnetospheric Y field component at mid-latitude ground observatories. These currents have a quite different geometry, with poloidal terms and higher degree contributions, that need to be further tested in studies of electromagnetic induction by external sources.

As to the shift values $\overline{\alpha}_X$ and $\overline{\alpha}_Y$ in Table 4.2, they are of the order of magnitude of crustal biases as found from comparing main field models from satellite data and monthly observatory means (e.g., Mandea and Langlais, 2002; Verbanac et al., 2015). In Verbanac et al., 2015 we find X and Y estimated crustal biases for BOU and NVS. Their Y biases are in agreement with $\overline{\alpha}_Y$ values shown in Table 4.2 within a ± 6 nT deviation. However, their X crustal biases are ~ 30 nT below $\overline{\alpha}_X$. A diurnal ionospheric bias is expected to remain in the observatory series $\Delta\widetilde{X}$, because of the used datum line. From Figures 3.5 and 4.2 this bias should be negative and less than 10 nT for the X component in BOU and NVS and would represent a negative instead of a positive contribution to

$\overline{\alpha}_X$. It is possible that either the estimated crustal X biases have a superposed magnetospheric (negative) mean contribution due to the westward ring current, or that the TS05 model overestimates the (westward) mean daily magnetospheric contribution of TAIL + SRC system.

5 Relating Solar, Interplanetary and Geomagnetic Activity Parameters

5.1 Overview

This chapter is devoted to the connection between Sun and Earth and how the phenomena occurring in the solar atmosphere (namely the photosphere and the chromosphere), have an impact on the IMF and on the geomagnetic activity. In the absence of a complete physical model that can "go" from the Sun to the Earth, my approach is based on the statistical analysis of different set of parameters related to the Sun, IMF and Earth by searching for correlations between them. Then I discuss these correlations and the respective physical meaning, starting the analysis by looking at the solar north and south asymmetries of active regions. Waldmeier (1971) found that the Sun's hemispheric asymmetry extends to sunspots, faculae, prominences, and coronal brightness. Roy (1977) reported that major flares and magnetically complex sunspot groups also showed strong north-south asymmetry (Hathaway, 2015). Solar magnetic activity manifests itself differently in each hemisphere and is controlled by differential rotation and meridional circulation in each hemisphere. A phase lag of activity in the north and south hemispheres typically ranges between one day to a year or two (Dorotovic et al., 2010). This suggests that the coupling between the two hemispheres is variable and weak (Norton and Gallagher, 2009). The spatio-temporal asymmetry

of meridional flow will also play a key role in the hemispheric asymmetry (Chowdhury, Choudhary, and Gosain, 2013). The asymmetry based on the absolute asymmetry index (N minus S) is enhanced near the cycle maximum (Temmer et al., 2006). It is also found that the solar activity should dominate in the hemisphere where maximum sunspot group number is the largest (Li et al., 2002). On the other hand, it is accepted that the above asymmetry has an impact outside the Sun as, e.g., in the emitted flux of energetic charged-particle observed by Ulysses (Erdos and Balogh, 2010), in the galactic cosmic ray intensity and in the solar wind speed (Alania et al., 2001; Alania, Gil, and Modzelewska, 2008; Modzelewska and Alania, 2009; Gil and Alania, 2013; Sheeley and Harvey, 1981; Gibson et al., 2009). On the other hand, due to Heliospheric Current Sheet (HCS) warping (see Section 2.3.1), Earth crosses the current sheet two to six times during each solar rotation, depending of the HCS's warping. These current sheet crossings are named *crossings of interplanetary sector boundaries*, when the polarity of the magnetic field changes its sign (as observed from the Earth).

In this Chapter, I make use of solar surface parameters that are computed from data images taken at the spectroheliograph of the OGAUC. This allowed, as a first task, to make an analysis of the hemispheric asymmetry observed at the Sun photosphere during solar cycle 24, from 2009 to 2016. Then, I used these solar surface parameters together with solar wind, interplanetary medium and geomagnetic activity proxies to take into account the whole chain of physical processes relating the Sun to its final space weather effects felt at the Earth's surface. Finally, I examined the influence of the solar magnetic field polarity, toward or away, on all kinds of used parameters. To this end, I tested if these parameters tend to have different distributions (in particular different mean values) when measured under one or the other polarities, i.e., I looked for toward-away asymmetries. This was done by means of the analyses of Spearman's ranked cross-correlations among 33 time series of solar structures, interplanetary parameters, geomagnetic indices and coupling functions, and 33 cross-correlations of time series of helio-magnetic asymmetries (i.e., difference between the

mean values during toward and away solar magnetic polarities) of the same parameters, in the 2009-2016 time interval. Daily and 27-day averaged values were used to study cross-correlations of raw parameters, and annual and 27-day averages were used to study cross-correlations of the helio-magnetic asymmetries.

5.2 Data

In this section, the used observational data used is described. Table 5.1 lists all parameters and their source databases, Figures 5.1, 5.4, 5.3, 5.2 show all 27-day-averaged series. Time averages over a Bartels' rotation period make sense when looking for the mean influence that the Sun may have on the Earth, since they take into account the effect of all solar longitudes.

Five groups of parameters are defined: solar parameters (SP), solar wind parameters (SWP), geomagnetic indices (GAI), T-indices (TI) and IMF components (IMF) in order to perform cross-correlations to look for physical relations between these groups. SP parameters comprise sunspot numbers and facular area values. SWP are physical parameters characterizing the solar wind (section 2.5). GAI parameters are the Newell's coupling function, polar cap (north) index (PCN), AE, AU, AL indices, storm-time disturbance index (Dst), planetary Kp index and planetary Ap index (see section 2.2.2). TI indices were calculated from the X component of four TS05 currents: TAIL, SRC, PRC, FAC (Section 4.2). IMF components are B_X, B_Y, B_Z, B_{ZS} (the percentage of southward B_Z component) both in GSM and GSE systems, and B, the magnitude of the total IMF field vector.

TABLE 5.1: List of 33 parameters analysed and their data sources, separated by SP, SWP, IMF, GAI and TI. Explanations in the text.

Nº	Parameter	Source
	Solar Parameters (SP)	
1	FA-T, total facular area (% of solar disk).	Calculated by CITEUC team.
2	FA-N, northern facular area (% of solar disk).	
3	FA-S, southern facular area (% of solar disk).	
4	SN-T, total international sunspot number.	WDC-SILSO, Royal Observatory of Belgium, Brussels.
5	SN-N, Northern hemisphere's sunspot number.	
6	SN-S, Southern hemisphere's sunspot number.	
7	FA-NS, FA-N minus FA-S (%).	Obtained from items (2) and (3).
8	SN-NS, SN-N minus SN-S.	Obtained from items (5) and (6).
9	F10, solar radio flux of 10.7 cm (sfu).	NASA/OMNI daily data.
	Solar Wind (SWP) and IMF parameters	
10-12	B_X, B_Y GSE and B_Z GSE, nT.	NASA/OMNI daily data.
13-14	B_Y GSM and B_Z GSM, nT.	
15	B, magnitude of average field vector, nT.	
16-17	B_{ZS} GSE and B_{ZS} GSM.	Calculated by me with NASA/OMNI hourly data, suggested by P. Ribeiro.
18-19	T, proton temperature (10^3 K) and ρ, proton density (N/cm^3).	NASA/OMNI daily data.
20-21	V, bulk flow speed (km/s) and p, flow pressure (nPa).	
	Ground geomagnetic activity indices (GAI) and TI	
22	Newell's coupling function.	Tsyganenko's yearly data files for TA15 model.
23	Kp, Kp*10 index.	NASA/OMNI daily data.
24-26	AE, AU and AL indices (nT).	
27-28	PCN, Polar Cap North index (mV/m) and Dst index (nT).	
29-32	T-TAIL, T-SRC, T-PRC and T-FAC indices (nT).	Calculated by me from their respective TS05's sources (Ribeiro, P., Pais, A., Castillo, Y., Pinheiro, F., Morozova, A., Fernandes, J.).
33	Ap index (nT).	NASA/OMNI daily data.

5.2.1 Solar Parameters (SP)

Most solar parameters used refer to features at the solar photosphere: the international northern solar hemisphere's sunspot number (SN-N), international southern solar hemisphere's sunspot number (SN-S), international total sunspot number (SN-T), northern hemisphere's facular area (FA-N), southern hemisphere's facular area (FA-S), total facular area (FA-T), difference between northern and southern facular areas (FA-NS), difference between northern and southern sunspot number (SN-NS). I also used the solar radio flux at 10.7 cm wavelength (F10), originating in the chromosphere and corona of the Sun.

The facular areas were calculated by a CITEUC team using a software tool developed by Barata et al. (2017).

The sunspot numbers are obtained from SILSO (Sunspot Index and Long-term Solar Observations) database at Royal Observatory of Belgium, Brussels. These values are derived from an average of observations from a large worldwide network. Prior to the averaging of daily values, the scale of each contributing station is adjusted on a monthly basis relative to the Locarno (Switzerland) reference station. The raw Wolf numbers are multiplied by their monthly average k personal coefficient, that includes the systematic bias of the observer. SN-NS was calculated as SN-N minus SN-S.

Solar radio flux at 10.7 cm or F10.7 index is a proxy for the strength of solar radio emission, from all sources present on the solar disk, in a 100 MHz-wide band centred on 2800 MHz (wavelength of 10.7 cm), averaged over an hour and centred on the epoch given for the value. This quantity is a flux density, although is used to be referred as a flux. It is reported in solar flux units (s.f.u.), a unit used to express the flux density of radio energy from the Sun as received at the Earth (1 s.f.u = $10^{-22}\ Wm^{-2}Hz^{-1}$) (Tapping, 2013). Solar emissions at these wavelengths are very sensitive to conditions in the upper chromosphere and at the base of the corona. It is well known that the

F10.7 index correlates well with the SN-T as well as a number of UltraViolet (UV) and visible solar irradiance records (Bruevich, V.V.Bruevic, and Yakunina, 2014). It has the advantage that it can easily be measured on a daily basis from the Earth's surface, in all types of weather. The F10.7 values used in this work are the adjusted values at Sun's location, corrected for variations in the Earth-Sun distance, and given for the average distance by the Canadian Space Agency (/https://tinyurl.com/yc5b5czx).

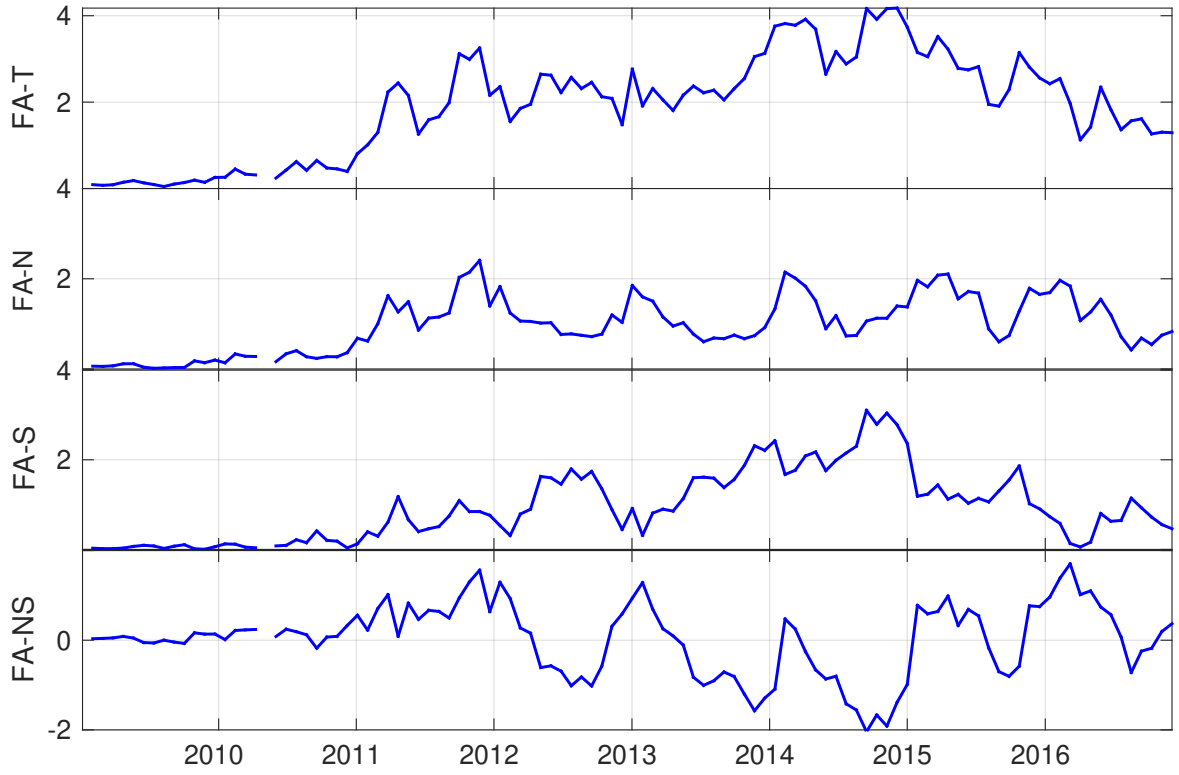


FIGURE 5.1: Time series of 27-day-averaged facular areas, in percentage of solar disk.

5.2.2 Solar Wind (SWP) and IMF Parameters

Solar wind parameters are the solar wind bulk speed in km/s (V), proton flux temperature in Kelvin (T), proton density in N/cm^3 (ρ) where N is the number of protons and solar wind proton ram pressure in nPa (p), measured by spacecraft at the outer magnetosphere (see section 2.4 and LRO Website).

The so-called IMF parameters are the B_X, B_Y, B_Z GSM and GSE components of the interplanetary magnetic field, in nT, and the magnitude of the average field vector, B, also in nT.

Most data were obtained from the LRO OMNI - NASA data set. Some hourly values are calculated from 1-min or 5-min resolution values. Some daily averages are taken over OMNI's hourly values, and 27-day averages and annual averages are taken over the daily averages. The 27-day averages are made to correspond to discrete Bartels' rotation numbers. More details can be found at /https://tinyurl.com/y8tgxzrt.

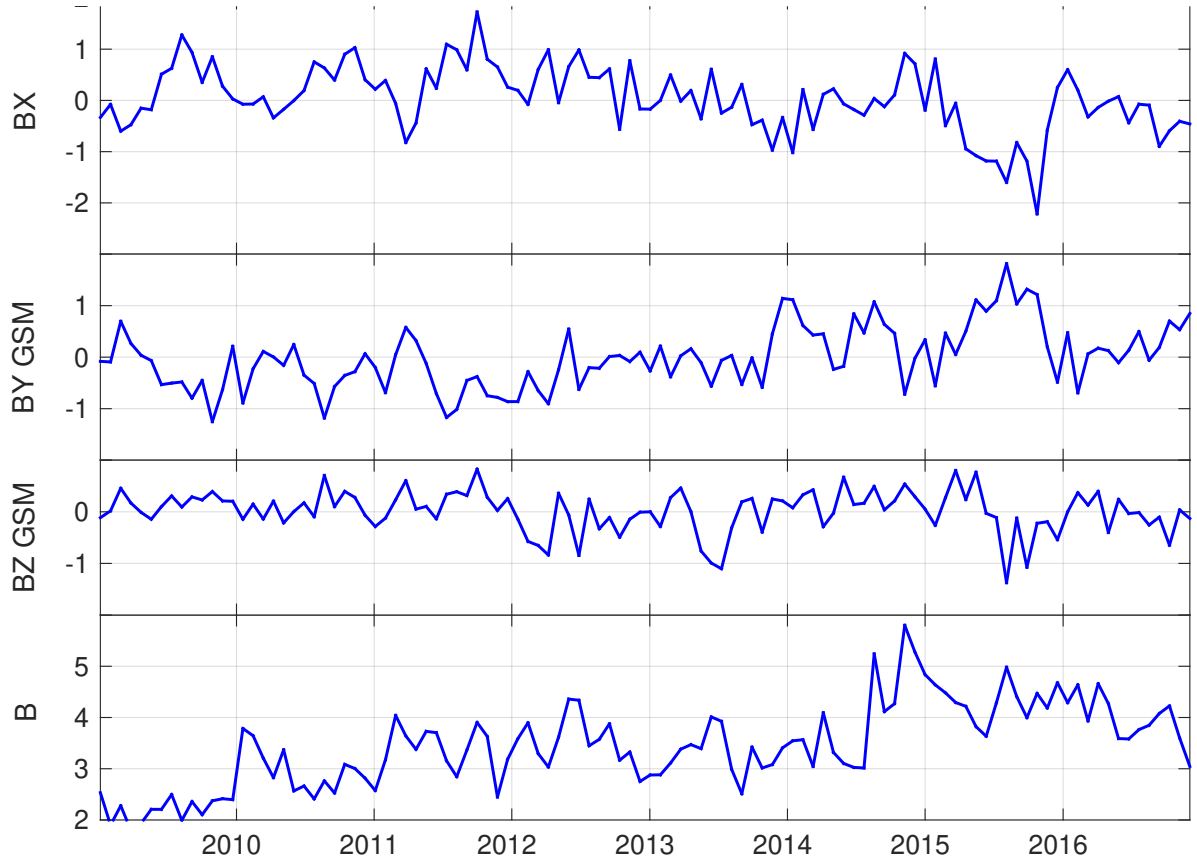


FIGURE 5.2: Time series of 27-day-averaged IMF parameters.

5.2.3 Ground Geomagnetic Activity Indices (GAI) and Newell's Coupling Function (NLL)

GAI indices used are the polar cap north (PCN), AE, AU, AL, Ap, Dst, Kp*10, described in section 2.2.2 and Newell's coupling function defined in section 2.2.3. The Kp*10 is similar to Kp, multiplied by 10, with the difference that in LRO dataset Kp*10 average is rounded to its nearest integer (i.e., 10, 13, 17, 20, ...), as shown in Table **??**.

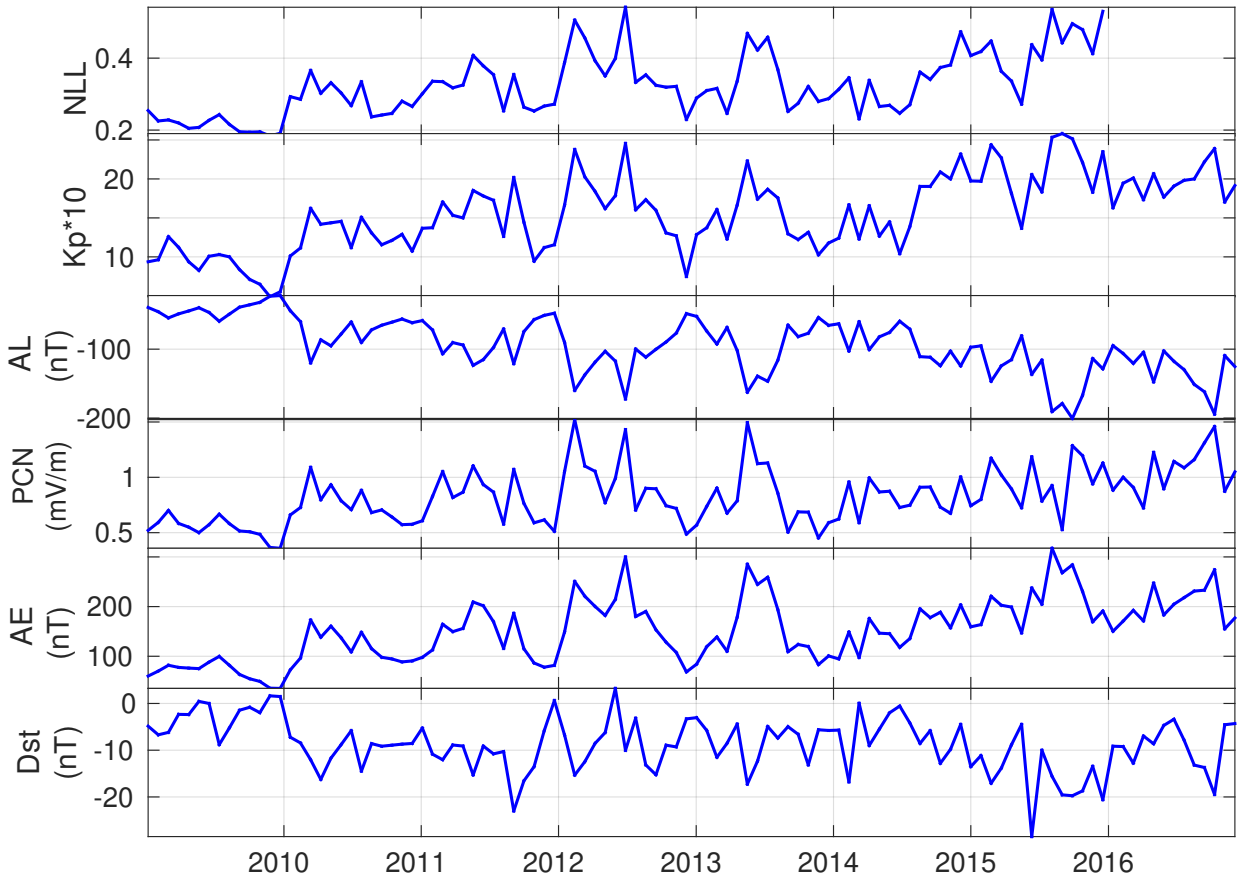


FIGURE 5.3: Time series of 27-day-averaged GAI parameters.

5.2.4 New Indices

Two kinds of parameters were tested in this work, as new possible proxies of geo-effectiveness (B_{ZS}) or geomagnetic activity (T-indices).

TABLE 5.2: Comparing Kp index and Kp*10 used by OMNI.

Kp	0o	0+	1-	1o	1+	2-	2o	2+	3-	3o	3+	4-	4o	4+
Kp*10	0	3	7	10	13	17	20	23	27	30	33	37	40	43
Kp	5-	5o	5+	6-	6o	6+	7-	7o	7+	8-	8o	8+	9-	9o
Kp*10	47	50	53	57	60	63	67	70	73	77	80	83	87	90

The IMF South Fraction B_{ZS}. The daily percentage of the southward component of B_Z, called here B_{ZS}, both in GSE and GSM systems. This parameter is used as a way to measure the daily percentage of geo-effectiveness due to southward IMF B_Z (Newell et al., 2007). It was calculated from hourly values of IMF B_Z (both for GSE and GSM components), and it represents the percentage of hours in a day with IMF B_Z negative (southward). For simplicity I denote it as B_{ZS} (S from "southward"). For a day j:

$$B_{ZS}^{j} = \frac{1}{24} * n^{j} * 100 \tag{5.1}$$

where n^j is the number of IMF B_Z negative hourly means in day j.

Magnetospheric T-indices (TI). These indices were calculated from Tsyganenko and Sitnov 2005 (TS05) simulations, using the X component hourly series of tail current sheet (T-TAIL), symmetric ring current (T-SRC), partial ring current (T-PRC) and field aligned currents (T-FAC) generated for the study in Chapter 4. These indices are based in the K index calculation method. Using the TS05 different sources time series for each observatory (COI, PAG, NVS and BOU) I calculated the difference between the maximum and minimum values over every three hours interval to obtain series of 8 values per day:

$$\Delta T8_{obs-cur}^{[i\,:\,i+2]} = max[T_{cur}^{i},\ T_{cur}^{i+1},\ T_{cur}^{i+2}] - min[T_{cur}^{i},\ T_{cur}^{i+1},\ T_{cur}^{i+2}] \tag{5.2}$$

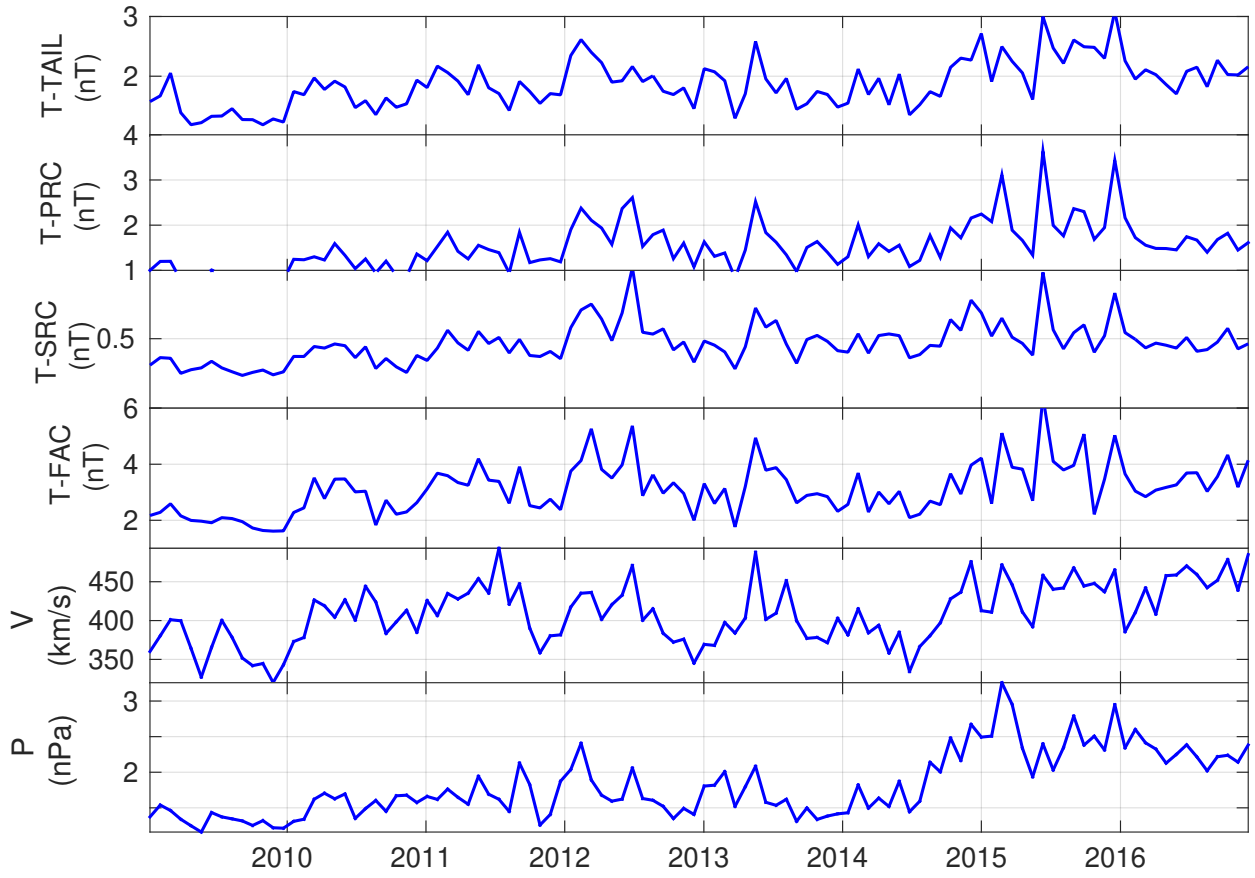


FIGURE 5.4: Time series of 27-day-averaged SWP and TI parameters.

with $i = 1, 4, 7, 10, 13, 16, 19$ and 22. Subscript $obs-cur$ stands for a certain observatory (COI, PAG, NVS and BOU) and for a certain current (TAIL, SRC, PRC or FAC), and $[i\ :\ i+2]$ refers to three consecutive hours. Then I calculated the mean of each group of 8 daily values to transform in daily series:

$$T^{j}_{obs-cur} = \frac{1}{8}\left(\Delta T8^{[1:3]}_{cur} + \Delta T8^{[4:6]}_{cur} + ... + \Delta T8^{[22:24]}_{cur}\right) \tag{5.3}$$

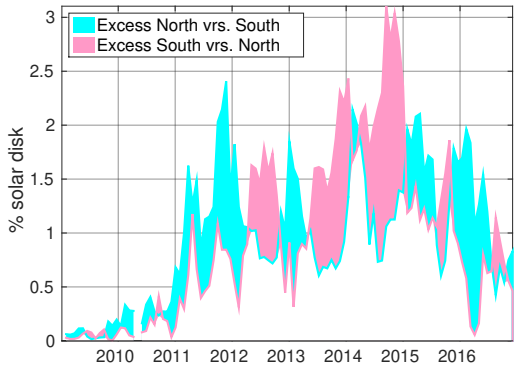
where j stands for a daily value, $[1:3]$,... , $[22:24]$ refers to the hourly values used to calculate each $\Delta T8_{cur}$. Finally I calculated the global daily mean using the corresponding daily $T^{j}_{obs-cur}$ of each observatory:

$$T^{j}_{cur} = \frac{1}{4}\left(T^{j}_{COI-cur} + T^{j}_{PAG-cur} + T^{j}_{NVS-cur} + T^{j}_{BOU-cur}\right) \tag{5.4}$$

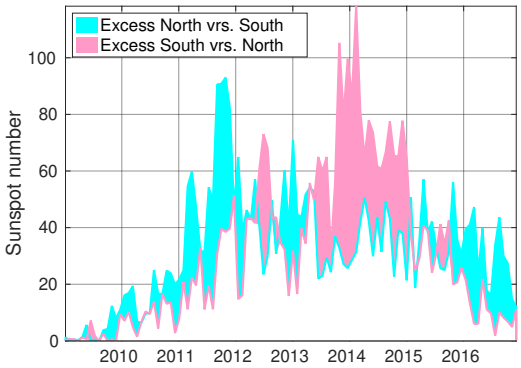
5.3 Hemispherical Asymmetry of Solar Cycle 24 Activity

The north-south asymmetry of sunspot activity results in asynchronous reversal of the Sun's polar field (see Table 2.3) and may play a role in the strength of the solar cycle. It appears that cycles with strong asymmetry tend to have a lower amplitude (e.g., cycles 23 and 24) in comparison with cycles in which sunspot activity in two hemispheres is more synchronized (e.g., cycles 21 and 22) (Mordvinov et al., 2016).

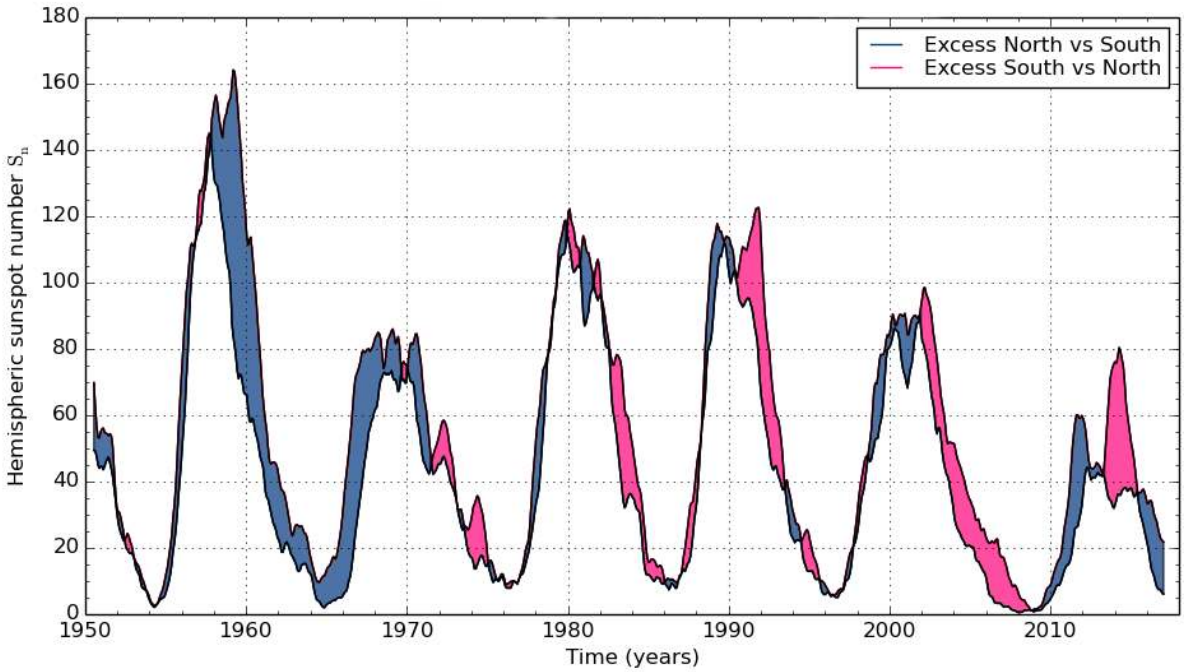
In Figures 5.5 (A), (B) we can see, for solar cycle 24, the global dominance of the northern hemisphere, with excess of SN-N and FA-N in 2009-2011 and 2015-2016, while SN-S and FA-S dominate only in 2013-2014. The absolute maximum of solar cycle 24 is in 2014 and apparently was mainly due to southern hemisphere activity. The other (relative) maximum is at the end of 2011 and apparently is mostly due to northern hemisphere's activity (FA-N and SN-N). For comparison, Figure 5.5 (C) shows the excess of sunspot number over the northern or southern solar hemispheres during the last six solar cycles, from SILSO Website (/http://sidc.be/silso).



(A) Excess FA-N over FA-S (blue) and excess FA-S over FA-N (pink) in 2009-2016



(B) Excess SN-N over SN-S (blue) and excess SN-S over SN-N (pink) in 2009-2016



(C) Excess north (blue) and south (pink) in last six solar cycles

FIGURE 5.5: Top-left: excess in FA north (cyan) and FA south (pink) in 2009-2016. Top right: excess in SN north (cyan) and SN south (pink) in 2009-2016. Bottom: excess in SN north (blue) and SN south (pink) in last six solar cycles, in hemispheric 13-month smoothed numbers. Image from SILSO Website (/http://sidc.be/silso). Royal Observatory of Belgium. August 2017.

5.4 From the Sun to the Earth

In this section we present a global analysis of all set of raw parameters putting the emphasis in the connection between Sun and Earth.

5.4.1 Analysis of Daily-Mean and 27-Day-Mean Series of Raw Parameters: SCC-1 and SCC-2

Two groups of Spearman's cross-correlation coefficients (r_s) were calculated. One between pairs of series with daily values of 33 parameters for 2009-2016 (hereafter SCC-1, at most with 2918 values per parameter) and the other between pairs of series with 27-day means of the same 33 parameters for 2009-2016 (hereafter SCC-2, at most with 108 values per parameter). Only the most strong and significant correlations (with $|r_s| \geq 0.4$ and $p \leq 0.05$) are shown in Tables 5.3 and 5.4.

Different colors are used to distinguish cross-correlations among parameters inside each of these groups: SP (yellow), IMF (red) and SWP + GAI + TI + NLL (blue). Cross-correlations among parameters in different groups are shown in green.

For SCC-1 calculations, most solar parameters have high correlations or anti-correlations among them (yellow cells), as also GAI parameters. Except for proton density, the SWP and TI parameters correlate strongly with GAI parameters (blue cells). B_X and B_Y correlate among themselves, as B_Z and B_{ZS} (red cells). A cross correlation among the IMF and the SWP-GAI parameters emerges (green cells), which does not include the (B_X, B_Y) field components.

The main difference between SCC-1 and SCC-2 calculations is that the second has more cross correlations than the first. For instance, now SP also correlates mostly with B, N, GAI, T-SRC and T-PRC; B also correlates strongly with SWP-GAI. In each coloured block, correlations got stronger. That is, when increasing the averaging period from 1 day to 27 days, correlations inside each block and between different blocks

increase. This can be due to the delay time in interaction mechanisms (the solar wind takes more than one day to travel from the Sun to the Earth) or to a cumulative evolution until some mechanism is triggered (Tsyganenko and Sitnov, 2005).

TABLE 5.3: Spearman's cross-correlations among 33 SPs, SWPs, IMFs, TIs and GAIs for 2009-2016, using daily means (SCC-1). NLL's correlations were calculated for 2009-2015. Only values with $|r_s| \geq 0.4$ and $p \leq 0.05$ are shown.

		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	
No	Par	FA-N	FA-S	SN-T	SN-N	SN-S	FA-NS	SN-NS	F10	B_X	B_Y GSE	B_Z GSE	B_Y GSM	B_Z GSM	B	B_{ZS} GSE	B_{ZS} GSM	T	ρ	V	p	NLL	Kp	Ap	AE	AU	AL	PCN	Dst	T-TAIL	T-SRC	T-PRC	T-FAC	Par
1	FA-T	0.82	0.87	0.76	0.56	0.67			0.85																									FA-T
2	FA-N		0.49	0.59	0.65		0.42		0.66																									FA-N
3	FA-S			0.72		0.79	-0.53		0.78																									FA-S
4	SN-T				0.76	0.82			0.92																									SNT
5	SN-N							0.52	0.65																									SNN
6	SN-S						-0.45	-0.59	0.79																									SNS
7	FA-NS							0.63																										FA-NS
8	SN-NS																																	SN-NS
9	F10																																	F10
10	B_X										-0.72		-0.72																					B_X
11	B_Y GSE												0.98																					B_Y GSE
12	B_Z GSE													0.77		-0.90	-0.71					-0.52			-0.42		0.45	-0.50						B_Z GSE
13	B_Y GSM																																	B_Y GSM
14	B_Z GSM															-0.71	-0.92					-0.66			-0.54	-0.49	0.56	-0.63	0.44					B_Z GSM
15	B																					0.44	0.44	0.44	0.40	0.42								B
16	B_{ZS} GSE																0.73					0.51			0.40		-0.42	0.48						B_{ZS} GSE
17	B_{ZS} GSM																					0.68			0.55	0.50	-0.57	0.64	-0.42					B_{ZS} GSM
18	T																			0.80		0.50	0.64	0.63	0.55	0.50	-0.56	0.49	-0.44	0.43	0.44	0.41	0.42	T
19	ρ																			-0.51	0.66													ρ
20	V																					0.48	0.62	0.61	0.54	0.48	-0.55	0.48	-0.53	0.50	0.53	0.44	0.54	V
21	p																					0.44	0.66	0.67	0.54	0.57	-0.51	0.42						p
22	NLL																						0.85	0.84	0.91	0.86	-0.91	0.89	-0.60	0.45	0.50	0.50	0.51	NLL
23	Kp																							0.99	0.92	0.88	-0.91	0.82	-0.56	0.54	0.56	0.56	0.55	Kp
24	Ap																								0.91	0.88	-0.90	0.81	-0.55	0.54	0.56	0.56	0.55	Ap
25	AE																									0.96	-0.99	0.88	-0.61	0.50	0.56	0.55	0.57	AE
26	AU																										-0.91	0.82	-0.52	0.45	0.52	0.52	0.52	AU
27	AL																											-0.89	0.64	-0.51	-0.56	-0.55	-0.57	AL
28	PCN																												-0.64	0.41	0.45	0.47	0.51	PCN
29	Dst																													-0.50	-0.54	-0.53	-0.59	Dst
30	T-TAIL																														0.77	0.86	0.77	T-TAIL
31	T-SRC																															0.85	0.79	T-SRC
32	T-PRC																																0.80	T-PRC
33	T-FAC																																	T-FAC
No	Par	FA-N	FA-S	SN-T	SN-N	SN-S	FA-NS	SN-NS	F10	B_X	B_Y GSE	B_Z GSE	B_Y GSM	B_Z GSM	B	B_{ZS} GSE	B_{ZS} GSM	T	ρ	V	p	NLL	Kp	Ap	AE	AU	AL	PCN	Dst	T-TAIL	T-SRC	T-PRC	T-FAC	Par

TABLE 5.4: Spearman's cross-correlations among 33 SPs, SWPs and GAIs for 2009-2016 using 27-day-means (SCC-2). NLL's correlations were calculated for 2009-2015. Only values with $|r_s| \geq 0.4$ and $p \leq 0.05$ are shown.

No	Par	FA-N	FA-S	SN-T	SN-N	SN-S	FA-NS	SN-NS	F10	B_X	B_Y GSE	B_Z GSE	B_Y GSM	B_Z GSM	B	B_{ZS} GSE	B_{ZS} GSM	T	ρ	V	p	NLL	Kp	Ap	AE	AU	AL	PCN	Dst	T-TAIL	T-SRC	T-PRC	T-FAC	Par
1	FA-T	0.77	0.89	0.84	0.60	0.83			0.90						0.58							0.46									0.53	0.52		FA-T
2	FA-N		0.48	0.63	0.73	0.49	0.49		0.67						0.54						0.53	0.45	0.41	0.41						0.46	0.48	0.51		FA-N
3	FA-S			0.82	0.46	0.90	-0.50	-0.49	0.86						0.53							0.49	0.41		0.43	0.42	-0.41				0.53	0.50		FA-S
4	SN-T				0.73	0.92			0.96																					0.46	0.40		SN-T
5	SN-N					0.50			0.67																									SN-N
6	SN-S							-0.56	0.91						0.41							0.45									0.52	0.44		SN-S
7	FA-NS							0.70																										FA-NS
8	SN-NS																																	SN-NS
9	F10														0.48							0.41									0.51	0.47		F10
10	B_X										-0.75		-0.73																					B_X
11	B_Y GSE												0.99																					B_Y GSE
12	B_Z GSE													0.82		-0.83	-0.58																	B_Z GSE
13	B_Y GSM																																	B_Y GSM
14	B_Z GSM															-0.72	-0.82					-0.48					0.41	-0.41					-0.40	B_Z GSM
15	B																	0.58		0.52	0.74	0.79	0.77	0.76	0.74	0.66	-0.74	0.62	-0.50	0.65	0.66	0.76	0.57	B
16	B_{ZS} GSE																0.70																	B_{ZS} GSE
17	B_{ZS} GSM																					0.57	0.43	0.41	0.46	0.42	-0.46	0.46		0.45	0.41	0.45	0.46	B_{ZS} GSM
18	T																			0.92	0.70	0.78	0.80	0.76	0.74	0.65	-0.77	0.65	-0.52	0.65	0.57	0.61	0.70	T
19	ρ																				0.51													ρ
20	V																				0.73	0.78	0.84	0.80	0.79	0.70	-0.81	0.71	-0.56	0.68	0.56	0.61	0.75	V
21	p																					0.77	0.88	0.84	0.75	0.67	-0.78	0.68	-0.50	0.81	0.62	0.76	0.67	p
22	NLL																						0.94	0.92	0.91	0.82	-0.93	0.82	-0.64	0.85	0.86	0.89	0.86	NLL
23	Kp																							0.97	0.94	0.85	-0.96	0.83	-0.61	0.82	0.76	0.84	0.81	Kp
24	Ap																								0.93	0.82	-0.96	0.86	-0.67	0.80	0.78	0.85	0.82	Ap
25	AE																									0.95	-0.98	0.87	-0.55	0.73	0.76	0.80	0.81	AE
26	AU																										-0.88	0.79	-0.40	0.61	0.71	0.72	0.74	AU
27	AL																											-0.88	0.62	-0.77	-0.76	-0.81	-0.82	AL
28	PCN																												-0.56	0.69	0.70	0.73	0.75	PCN
29	Dst																													-0.59	-0.50	-0.56	-0.54	Dst
30	T-TAIL																														0.78	0.88	0.83	T-TAIL
31	T-SRC																															0.92	0.84	T-SRC
32	T-PRC																																0.85	T-PRC
33	T-FAC																																	T-FAC
No	Par	FA-N	FA-S	SN-T	SN-N	SN-S	FA-NS	SN-NS	F10	B_X	B_Y GSE	B_Z GSE	B_Y GSM	B_Z GSM	B	B_{ZS} GSE	B_{ZS} GSM	T	ρ	V	p	NLL	Kp	Ap	AE	AU	AL	PCN	Dst	T-TAIL	T-SRC	T-PRC	T-FAC	Par

5.5 Toward-Away Asymmetries

In this section we present a global analysis of toward-away asymmetries from all set of parameters putting the emphasis in the connection between Sun and Earth. Wilcox Solar Observatory (WSO, /https://tinyurl.com/ybpqbwrr) provides a list of well defined sector boundaries (toward and away), for which there is good data, the reversal takes place cleanly, and intervals on either side of the boundary have an uniform field direction for ~ 4 days. The list is inferred from geomagnetic and spacecraft observations. The WSO boundary list was used to separate days when the Earth is north or south of the HCS. When no data were available at WSO list, north and south polarities were calculated using the angle between IMF GSE radial component ($\sqrt{B_X^2 + B_Y^2}$) and the positive X GSE axis, that is, the Earth-Sun direction with origin at Earth, measuring the angle anti-clockwise and using daily means of IMF B_X and IMF B_Y GSE obtained from NASA/OMNI database (see Figure 5.6). Sabbah (1995) defines IMF away from the Sun when the angle is between 45^o and 225^o. In any other case is toward to the Sun. The toward-away asymmetries with respect to IMF polarity of the 33 parameters (sec-

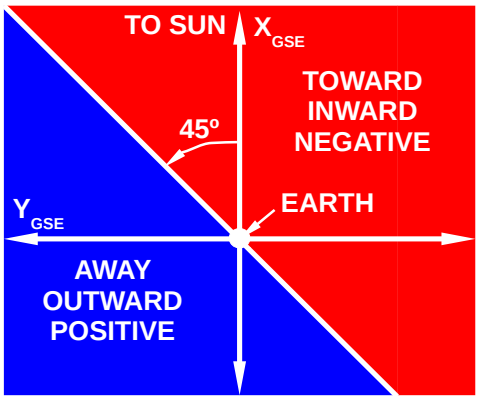


FIGURE 5.6: Definition of away and toward polarity seen from the Earth. Angles between $45^o - 225^o$ are considered away (outward) or positive polarity. Other angles are considered toward (inward) or negative polarity (Sabbah, 1995).

tion 5.2) were computed as I now explain. For every parameter I separated its values at days when the IMF directs towards (T), from its values when the IMF directs away

(W), relative to the Sun, and obtained the corresponding means:

$$\overline{T}_y = \frac{1}{n}\sum_{i=1}^{n} T^{i,y} \tag{5.5}$$

$$\overline{W}_y = \frac{1}{m}\sum_{i=1}^{m} W^{i,y} \tag{5.6}$$

where $\overline{T}_y$ is the mean of a parameter in a certain period y (solar rotation or year), counting only toward days, $\overline{W}_y$ is the mean of the parameter in period y (solar rotation or year), counting only away days, n is the number of days in period y when IMF points toward to the Sun, m is the number of days in period y when the IMF points away of the Sun, $T^{i,y}$ is the parameter value at a toward day i of period y, $W^{i,y}$ is a parameter value at an away day i of period y. Two different averaging periods were considered, the Bartels' 27-day solar rotation and the year. Then I obtained the 27-day period or annual asymmetries as the difference between the corresponding toward and away means:

$$A_y = \overline{T}_y - \overline{W}_y \tag{5.7}$$

The standard error for each A_y was calculated as:

$$\begin{aligned} \sigma_{A_y} &= \sqrt{\frac{\sigma_{Ty}^2}{n} + \frac{\sigma_{Wy}^2}{m}} \\ &= \sqrt{\frac{1}{n\,(n-1)^2}\left(\sum_{k=1}^{n}(T^{k,y} - \overline{T}_y)^2\right)^2 + \frac{1}{m\,(m-1)^2}\left(\sum_{k=1}^{m}(W^{k,y} - \overline{W}_y)^2\right)^2} \end{aligned} \tag{5.8}$$

where σ_{Ty} is the standard deviation of all values at n toward days in period y, and σ_{Wy} is the standard deviation of all values at m away days in period y.

Asymmetries were standardized ($A_y^s = A_y/\sigma$) dividing by the standard deviation of all asymmetries of a parameter for the analysed time interval, σ:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(A_i - \overline{A})^2} \tag{5.9}$$

where $\overline{A}$ is the mean of all n A_i values. When all parameter values are positive, an A_y^s positive represents $\overline{T}_y$ greater than $\overline{W}_y$, i.e., the parameter tends to be larger at toward than at away days. For parameters with predominantly negative values (as Dst or AL), A_y^s positive means the opposite, i.e, the amplitude of the parameter tends to be larger at away days. To compute the error of standardized asymmetries I used the error propagation rule, from the standard error of difference between means:

$$\sigma^2_{A_y^s} = \frac{1}{\sigma^2}\left[\sigma^2_{A_y}(A_y^s)^2 + \left(\frac{\partial\sigma}{\partial A_y}\right)^2\right] \tag{5.10}$$

All these calculations were done using toward and away means computed over 27-days solar Bartels' rotations and over annual periods.

Figure 5.7 shows a pixel plot of toward and away days from 2009 to 2016, obtained from Wilcox Solar Observatory list of sector boundaries combined with calculation of toward and away days using Sabbah (1995) criteria. Vertical axis represents a Bartels' rotation (27 days) and horizontal axis represents the Bartels' rotation number (1 to 108), from January 2009 to December 2016. Pink represents days with toward IMF, cyan are days with away IMF and white means lack of available or reliable data. From January 2009 to March 2012 we have negative polarity (44 Bartels' rotations). From April 2012 to March 2015 we have a transition epoch (40 rotations). From April 2015 to 2016 we have positive polarity (24 rotations). The Sun's North pole is most inclined toward the Earth in September and the South pole in March. For this reason, the dominant polarity seen by the Earth around September is that of the Sun (see Figure 5.7).

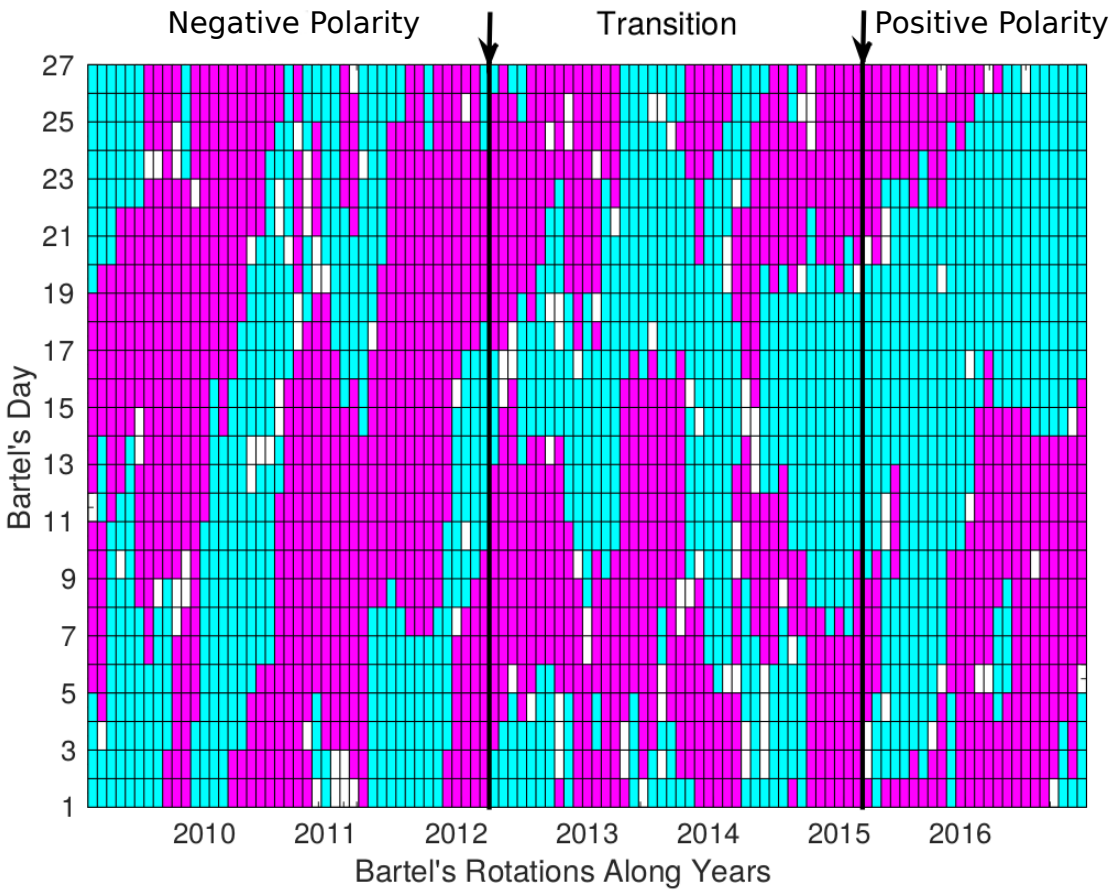
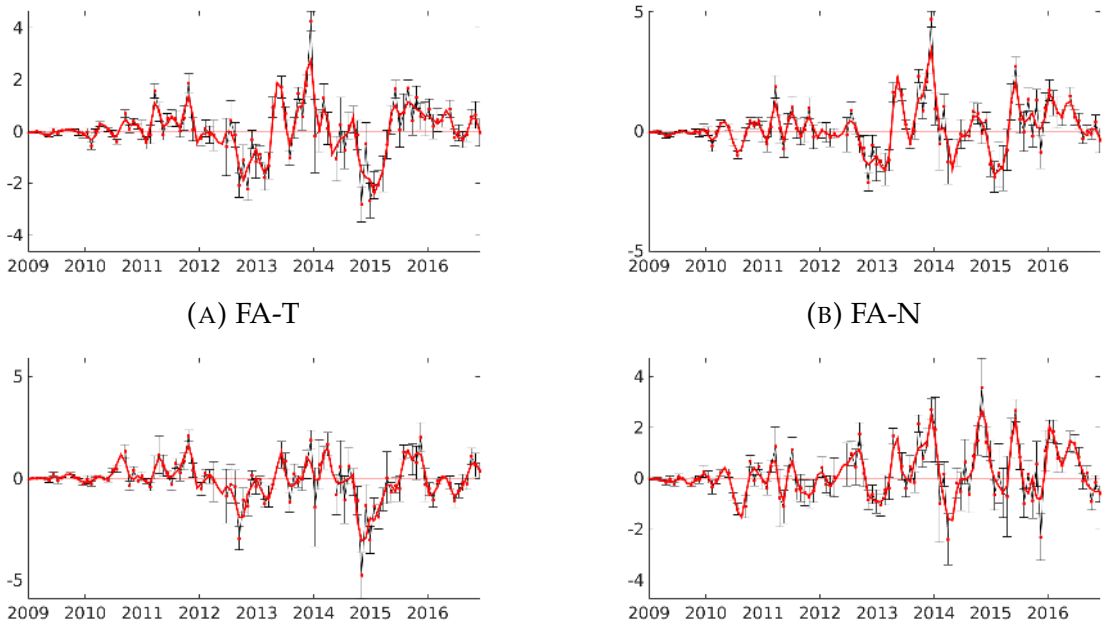


FIGURE 5.7: Pixel plot of days with toward (pink) or away (cyan) IMF. White days means missing data or polarity that could not be determined. Vertical axis is a Bartels' rotation (27 days). Horizontal axis represents number of solar rotations along 2009-2016.

5.5.1 Analysis of Annual and 27-Day-Mean Series of Asymmetries: SCC-3 and SCC-4

Annual toward-away asymmetries of SP, IMF, SWP, GAI and TI along 2009-2016 are plotted in Appendix B, in Figures B.1, B.2, B.3, B.4 and B.5. Spearman's correlation coefficients were calculated among these annual asymmetries (SCC-3) for 2009-2016, except for Newell function's correlations which were calculated for 2009-2015. Table 5.5 shows correlations $|r_s| \geq 0.4$ with $p \leq 0.05$. There are few values because most p-values are greater than 0.05. This is because the sample is too small (7-8 values) for each parameter. However we can see good correlations among most GAI and T-PRC; some correlations among SP with B_{ZS}, B_Z, some GAI and T-FAC; correlation among B_X, B_Y and AE; correlation among B_{ZS}, B_Z and T-FAC; SWP among them, but p only correlates with ρ; correlations among SWP with T-PRC and T-TAIL.

Plots of 27-day-mean asymmetries for IMF, SWP, GAI and TI parameters for 2009-2016 are shown in Figures , 5.8, 5.9, 5.10 and 5.11 with their respective error bars and a loess span of 5% of data points superposed (see section 2.6.3). Spearman's cross-correlations were calculated for the 27-day-mean asymmetries for 2009-2016 (SCC-4), except Newell's correlations, that were calculated for 2009-2015. Table 5.6 shows correlations only for values with $|r_s| \geq 0.4$ and $p \leq 0.05$. We observe more significant correlations than with annual asymmetries, but a very similar pattern as with daily-mean series of raw parameters (SCC-1, see Table 5.3): SP correlate only among them, with the exception of FA-S with p; B_X and B_Y only correlates among them; B_Z and B_{ZS} correlate among them and with GAI and TI; B do not correlate with any other parameter; SWP correlate well among them but p only correlate with ρ; SWP except ρ correlate with most GAI and TI; all GAI correlate among them and with TI. Nonetheless, correlation values among GAI parameters get higher, as well as among GAI and IMF B_{ZS} and B_Z GSM parameters. This is indicative of the influence of helio-magnetic

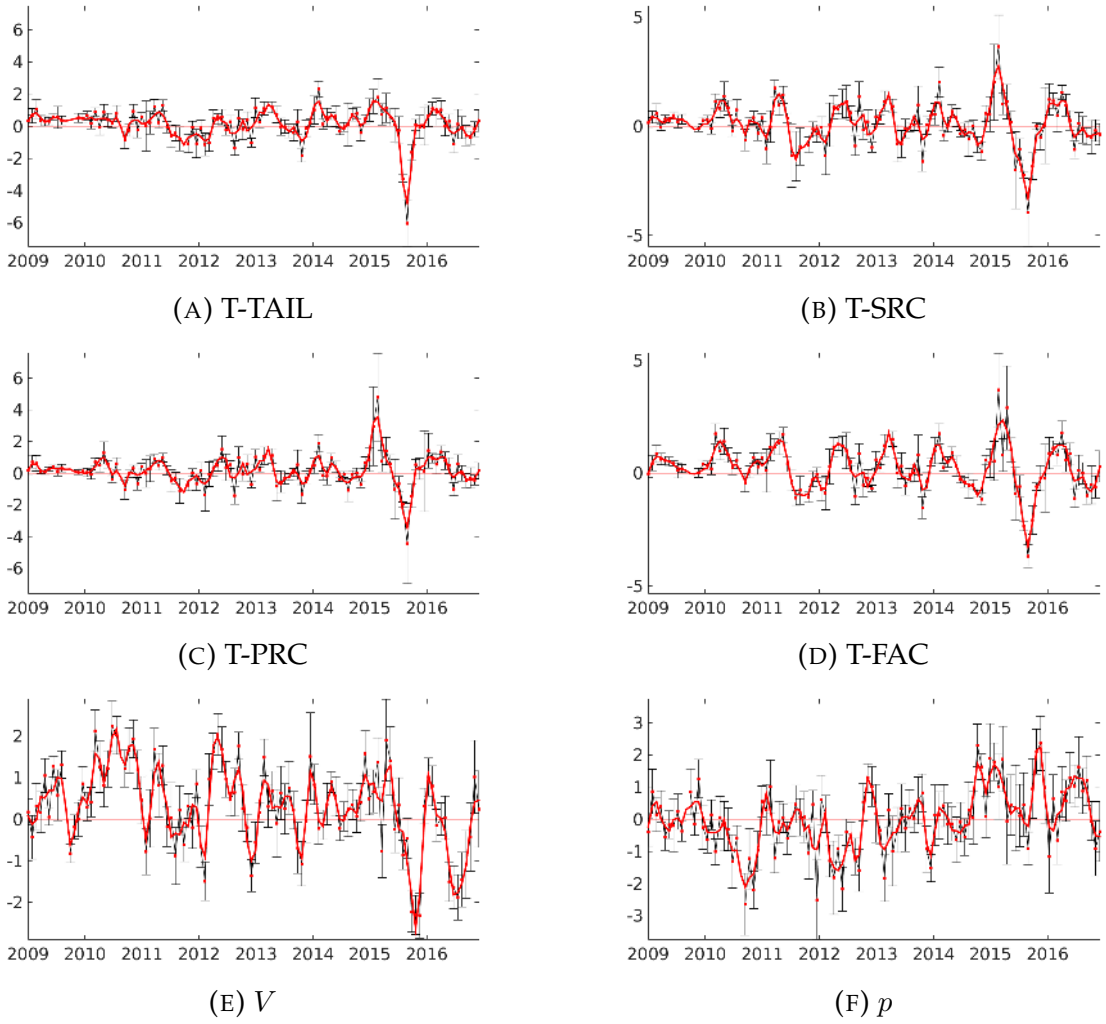


(A) FA-T (B) FA-N

(C) FA-S (D) FA-NS

FIGURE 5.8: Time series of 27-day-averaged facular area asymmetries.

assymetry in geo-effectiveness of Sun-Earth interactions.

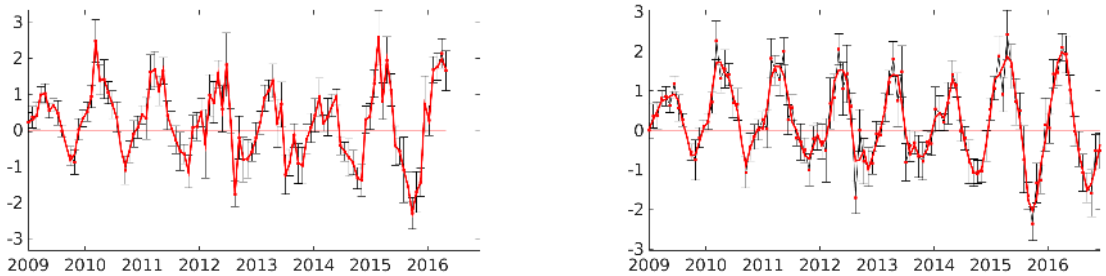


(A) T-TAIL (B) T-SRC

(C) T-PRC (D) T-FAC

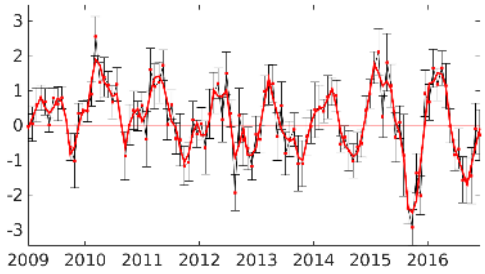
(E) V (F) p

FIGURE 5.9: Time series of 27-day-averaged SWP and TI asymmetries.

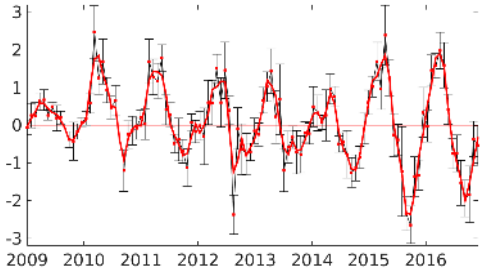


(A) NLL

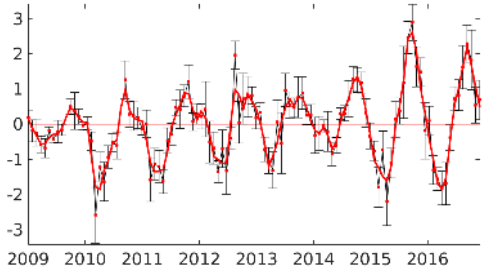
(B) PCN



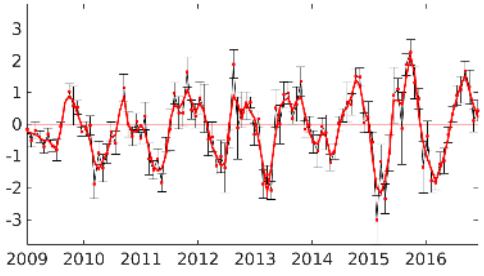
(C) Kp*10



(D) AE

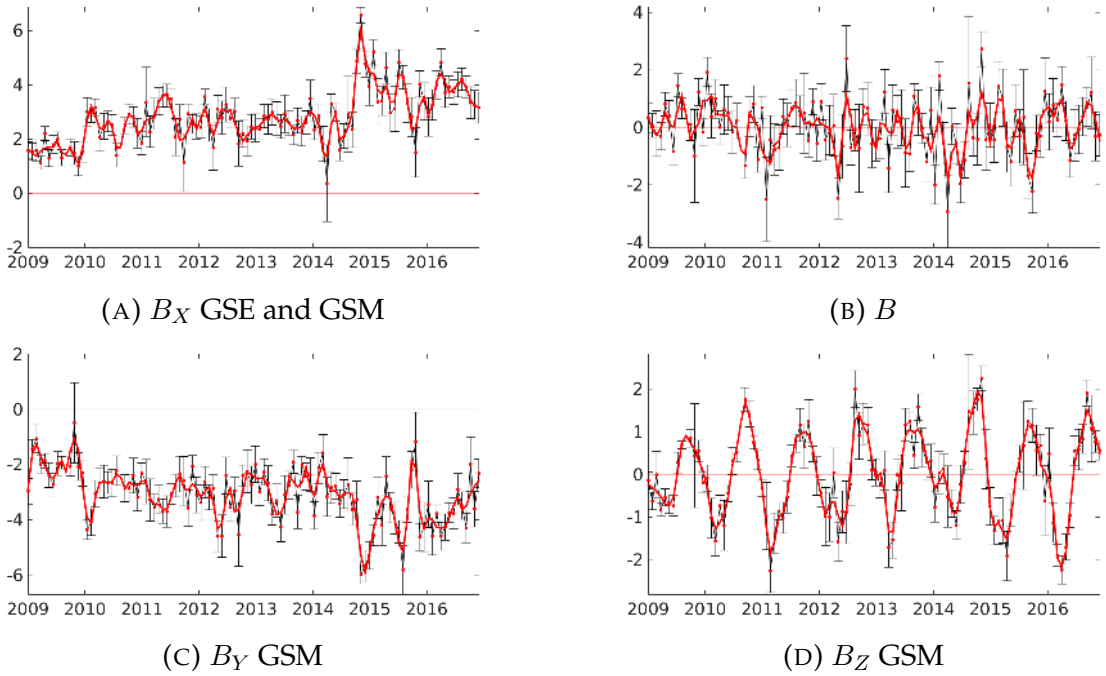


(E) AL



(F) Dst

FIGURE 5.10: Time series of 27-day-averaged GAI asymmetries.



(A) B_X GSE and GSM

(B) B

(C) B_Y GSM

(D) B_Z GSM

FIGURE 5.11: Time series of 27-day-averaged IMF asymmetries.

TABLE 5.5: Spearman's cross-correlations between annual asymmetries of 32 SP, SWP and GAI for 2009-2016 and Newell for 2009-2015 (SCC-3). Only values with $|r_s| \geq 0.4$ and $p \leq 0.05$ are shown.

No	Par	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	Par
		FA-N	FA-S	SN-T	SN-N	SN-S	FA-NS	SN-NS	F10	B_X	B_Y GSE	B_Z GSE	B_Y GSM	B_Z GSM	B	B_{ZS} GSE	B_{ZS} GSM	T	ρ	V	p	NLL	Kp	Ap	AE	AU	AL	PCN	Dst	T-TAIL	T-SRC	T-PRC	T-FAC	
1	FA-T	0.79															0.83																	FA-T
2	FA-N																																	FA-N
3	FA-S						-0.76										0.81									0.74		0.86	-0.79					FA-S
4	SN-T								0.90								0.76																	SN-T
5	SN-N							0.74																										SN-N
6	SN-S							-0.90						-0.76		0.74																		SN-S
7	FA-NS																					-0.82						-0.79						FA-NS
8	SN-NS																																0.74	SN-NS
9	F10															0.74	0.81																	F10
10	B_X										-0.93		-0.93												-0.74									B_X
11	B_Y GSE												1.00												0.79									B_Y GSE
12	B_Z GSE													0.90		-0.83																	0.79	B_Z GSE
13	B_Y GSM																								0.79									B_Y GSM
14	B_Z GSM															-0.98																		B_Z GSM
15	B																																	B
16	B_{ZS} GSE																0.76																	B_{ZS} GSE
17	B_{ZS} GSM																																	B_{ZS} GSM
18	T																		-0.95	0.83												-0.79		T
19	ρ																			-0.83	0.74											0.83		ρ
20	V																																	V
21	p																													0.83		0.81		p
22	NLL																						0.79		0.93	0.82	-0.96	0.89	-0.82			-0.82		NLL
23	Kp																							0.90	0.79	0.81	-0.83							Kp
24	Ap																									0.86								Ap
25	AE																									0.88	-0.98	0.88	-0.86			-0.86		AE
26	AU																										-0.83	0.88	-0.83					AU
27	AL																											-0.81	0.76			0.88		AL
28	PCN																												-0.98					PCN
29	Dst																																	Dst
30	T-TAIL																																	T-TAIL
31	T-SRC																																	T-SRC
32	T-PRC																																	T-PRC
33	T-FAC																																	T-FAC
No	Par	FA-N	FA-S	SN-T	SN-N	SN-S	FA-NS	SN-NS	F10	B_X	B_Y GSE	B_Z GSE	B_Y GSM	B_Z GSM	B	B_{ZS} GSE	B_{ZS} GSM	T	ρ	V	p	NLL	Kp	Ap	AE	AU	AL	PCN	Dst	T-TAIL	T-SRC	T-PRC	T-FAC	Par

TABLE 5.6: Spearman's cross-correlations between 27-day-mean asymmetries of 32 SPs, SWP and GAI for 2009-2016 and Newell function's correlations for 2009-2015 (SCC-4). Only values with $|r_s| \geq 0.4$ and $p \leq 0.05$ are shown.

No	Par	2 FA-N	3 FA-S	4 SN-T	5 SN-N	6 SN-S	7 FA-NS	8 SN-NS	9 F10	10 B_X	11 B_Y GSE	12 B_Z GSE	13 B_Y GSM	14 B_Z GSM	15 B	16 B_{ZS} GSE	17 B_{ZS} GSM	18 T	19 ρ	20 V	21 p	22 NLL	23 Kp	24 Ap	25 AE	26 AU	27 AL	28 PCN	29 Dst	30 T-TAIL	31 T-SRC	32 T-PRC	33 T-FAC	Par
1	FA-T	0.73	0.65						0.50																							FA-T
2	FA-N						0.71																									FA-N
3	FA-S						-0.52		0.44												-0.44											FA-S
4	SN-T				0.67	0.56			0.74																							SN-T
5	SN-N							0.76	0.60																							SN-N
6	SN-S						-0.43	-0.71																								SN-S
7	FA-NS							0.41																								FA-NS
8	SN-NS																															SN-NS
9	F10																															F10
10	B_X										-0.61		-0.59																			B_X
11	B_Y GSE												0.98																			B_Y GSE
12	B_Z GSE													0.49		-0.85	-0.42							-0.40		0.41	-0.48					B_Z GSE
13	B_Y GSM																															B_Y GSM
14	B_Z GSM															-0.50	-0.96					-0.87	-0.74	-0.72	-0.87	-0.86	0.85	-0.86	0.79	-0.50	-0.54	-0.54	-0.67	B_Z GSM
15	B																															B
16	B_{ZS} GSE																0.52					0.45			0.45		-0.45	0.55					B_{ZS} GSE
17	B_{ZS} GSM																					0.88	0.74	0.70	0.85	0.85	-0.84	0.85	-0.76	0.53	0.53	0.53	0.68	B_{ZS} GSM
18	T																		-0.48	0.85			0.56	0.55	0.44	0.41	-0.43				0.42		0.40	T
19	ρ																			-0.61	0.46													ρ
20	V																					0.44	0.64	0.62	0.51	0.47	-0.51	0.46	-0.49		0.51		0.51	V
21	p																						0.51	0.55	0.40	0.47						0.41		p
22	NLL																						0.92	0.89	0.97	0.96	-0.96	0.93	-0.90	0.60	0.69	0.64	0.82	NLL
23	Kp																							0.97	0.94	0.92	-0.94	0.90	-0.87	0.66	0.71	0.65	0.80	Kp
24	Ap																								0.93	0.91	-0.92	0.88	-0.88	0.66	0.74	0.68	0.82	Ap
25	AE																									0.97	-0.99	0.96	-0.91	0.63	0.70	0.65	0.83	AE
26	AU																										-0.94	0.90	-0.87	0.62	0.69	0.66	0.80	AU
27	AL																											-0.97	0.92	-0.62	-0.70	-0.63	-0.83	AL
28	PCN																												-0.87	0.59	0.62	0.57	0.78	PCN
29	Dst																													-0.66	-0.77	-0.69	-0.85	Dst
30	T-TAIL																														0.79	0.84	0.78	T-TAIL
31	T-SRC																															0.91	0.89	T-SRC
32	T-PRC																																0.84	T-PRC
33	T-FAC																																	T-FAC
No	Par	FA-N	FA-S	SN-T	SN-N	SN-S	FA-NS	SN-NS	F10	B_X	B_Y GSE	B_Z GSE	B_Y GSM	B_Z GSM	B	B_{ZS} GSE	B_{ZS} GSM	T	ρ	V	p	NLL	Kp	Ap	AE	AU	AL	PCN	Dst	T-TAIL	T-SRC	T-PRC	T-FAC	Par

5.6 Discussion

TABLE 5.7: Summary of correlations $r_s \geq 0.4$ with $p \leq 0.05$ among groups of parameters.

Parameter	**DAILY SERIES**	**27-DAY-MEAN SERIES**
FA	All SP except SN-NS	All SP; B, p, GAI except PCN and Dst; TI except T-FAC
SN, F10	All SP	All SP except FA-NS; B, Newell, T-SRC, T-PRC
FA-NS, SN-NS	All SP except FA-T, SN-T, F10	All SP except SN-T, SN-N and F10
B_X-B_Y	B_X, B_Y	B_X, B_Y
B_Z-B_{ZS}	B_Z, B_{ZS}, GAI except Kp, Ap	B_Z, B_{ZS}, TI, GAI except Dst
B	GAI except AL, Dst, PCN	All except IMF, ρ, SN-T, SN-N, FA-NS, SN-NS
TEM-p-V	GAI, TI, SWP	FA-N, B, SWP, GAI, TI
GAI	B_{ZS}, B_Z, B, GAI, TI, SWP except ρ	All except SN-T, SN-N, SN-NS, FA-NS, B_X, B_Y, ρ
TI	TI, GAI, T, V	All except SN-N, FA-NS, SN-NS, B_X, B_Y, ρ

Parameter	**ANNUAL ASYMMETRIES**	**27-DAY-MEAN ASYMMETRIES**
FA	FA-NS, B_{ZS}, AU, PCN, Dst	FA-NS, F10, p
SN, F10	F10, SN-T, SN-NS, B_Z, B_{ZS}	SN, FA-S, FA-NS, SN-NS, F10
FA-NS, SN-NS	FA-S, SN-N, SN-S, Newell, PCN, T-FAC	All SP except SN-T and F10
B_X-B_Y	B_X, B_Y, AE	B_X, B_Y
B_Z-B_{ZS}	FA-T, FA-S, SN-T, SN-S, F10, T-FAC	B_Z, B_{ZS}, AE, AL, PCN, GAI, TI
B	None	None
SWP	SWP, T-TAIL, T-PRC	FA-S, SWP, GAI, TI except T-TAIL
GAI	FA-S, FA-NS, B_X, B_Y, GAI, T-PRC	B_Z, B_{ZS}, SWP except ρ, GAI, TI
TI	SN-NS, B_Z, SWP except V, Newell, AE, AL	B_Z, B_{ZS}, SWP except ρ, GAI, TI

Table 5.7 summarises the resulting significant correlations with $r_s \geq 0.4$, among raw parameters (1-day-averaged and 27-day-averaged, top) and among towards-away asymmetries (annual asymmetries and 27-day asymmetries series, bottom), each one with the most important cross-correlations, separated there in solar, IMF and geomagnetic parameters.

It was confirmed that is better to analyse cross-correlations for 27-day averaged series and 27-day asymmetries rather than daily averaged series or annual asymmetries. It could be due to the complexity and great amount of energy involved in Earth-Sun

coupling processes, that require different time intervals to develop, usually greater than one day until few weeks (i.e., energy transfer processes, as radiation, convection, diffusion, shocks, etc.)

Most studies using 27-day variations in solar, interplanetary and geomagnetic parameters are focused in their relation with the 27-day variation in cosmic ray intensity and others in different periodicities in some parameters (Schreiber (1998), Selot et al. (2009), and Alania et al. (2001)). We confirmed cross-correlations among all 27-day-averaged parameters analysed.

27-day-averaged raw series. The Newell coupling function is the one that correlates best among the five groups of 27-day averaged parameters. This function represents the rate at which magnetic flux is opened at the magnetopause (Newell et al., 2007).

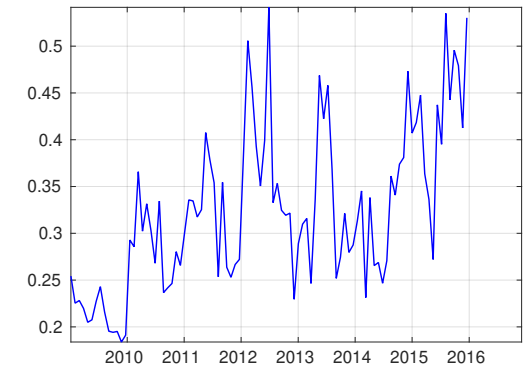
T-SRC and T-PRC have the next two best correlations among all groups of 27-day averaged parameters (SP, GAI, TI, IMF and SWP), with same Newell parameters plus SN-T, but excluding B_Z GSM.

B has also good cross-correlations among all groups of parameters except with the other IMF parameters.

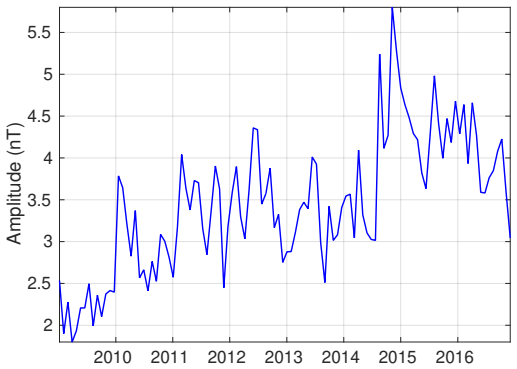
FA-N and FA-S 27-day averaged series are the SP parameters with more cross-correlations among groups (see Table 5.4).

Newell, T-SRC, T-PRC, B, FA-N and FA-S 27-day averaged series are shown in Figure 5.12. It is remarkable the correspondence among Newell, B, T-PRC and T-SRC.

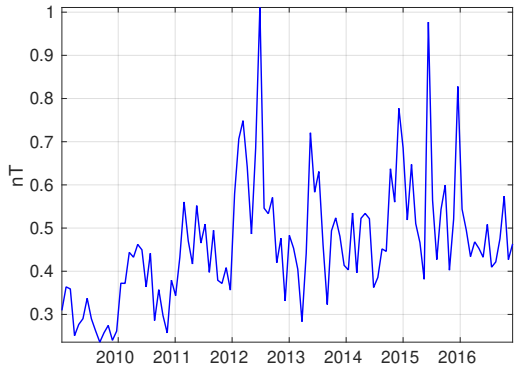
FA-N series (Figure 5.12 (E)) seem to have annual periodicity in 2012-2016, with higher positive values at the beginning and ending of each year, and minimum around the middle of the year. It could be a projection effect due to the displacement of the Earth and its rotation axis with respect to the solar equator along the year.



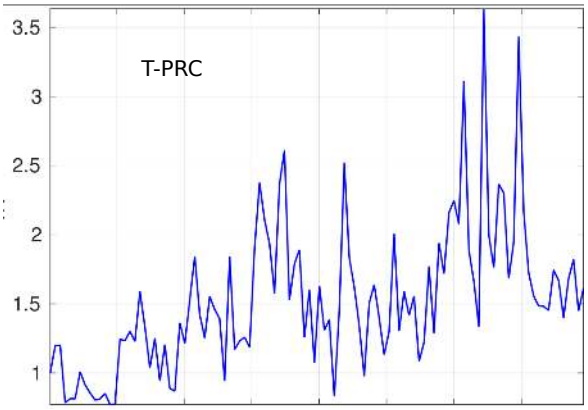
(A) 27-day-averaged Newell



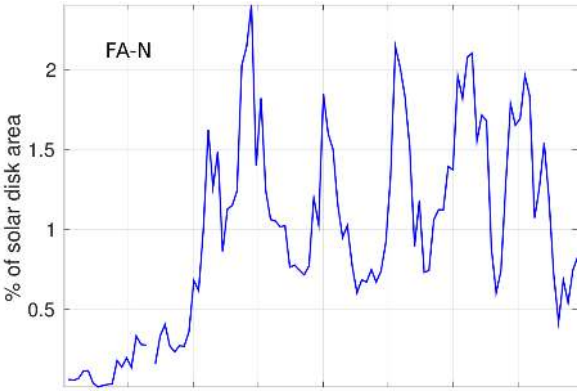
(B) 27-day-averaged B



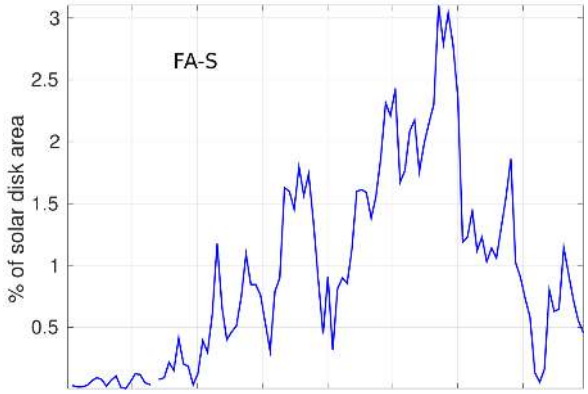
(C) 27-day-averaged T-SRC



(D) 27-day-averaged T-PRC



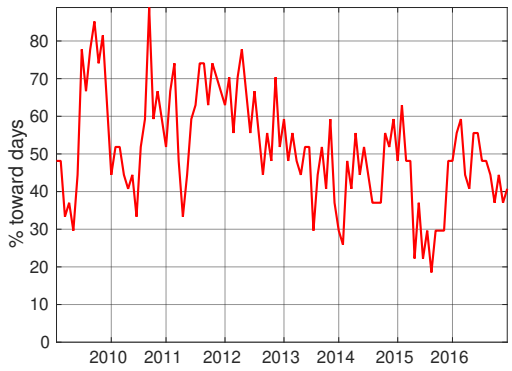
(E) 27-day-averaged FA-N.



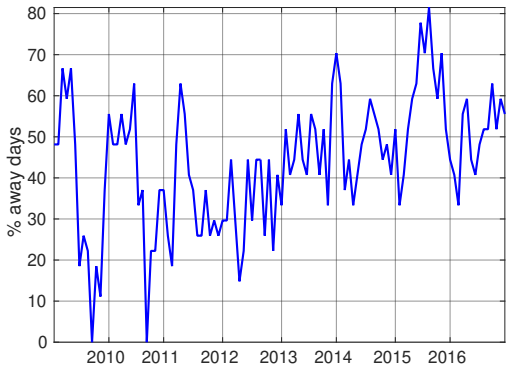
(F) 27-day-averaged FA-S

FIGURE 5.12: Better SCC-2 cross-correlations. Top: Newell (left) and B (right) 27-day-averaged series. Middle: T-SRC (left) and T-PRC (right) 27-day-averaged series. Bottom: FA-N (left) and FA-S (right) 27-day-averaged series.

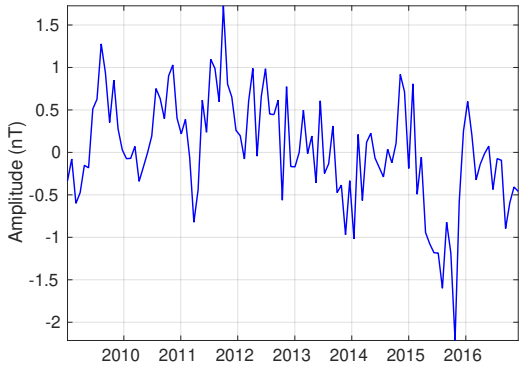
27-day-averaged towards-away asymmetries. Figures 5.13 (A), (B) show the 27-day averaged percentage of toward (A) and away (B) days obtained from Wilcox Observatory polar flux measurements. It is clear that toward days prevail in 2009-2012 and away days prevail in 2015-2016 (polarity of Sun's northern hemisphere in both time intervals). This means that Earth transited at north of HCS most of the time (see Figure 5.7. Figure 5.13 (C) also shows that B_X is mainly positive (toward) in 2009-2012 and mainly negative (away) in 2015-2016. Figure 5.13 (D) shows that B_Y GSM is negative at 2009-2012 (toward) and positive in 2015-2016 (away). We can observe that the percentage of toward days corresponds with B_X and percentage of away days corresponds with B_Y. This shows why B_X and B_Y are well anti-correlated ($r_s \sim -0.7$).



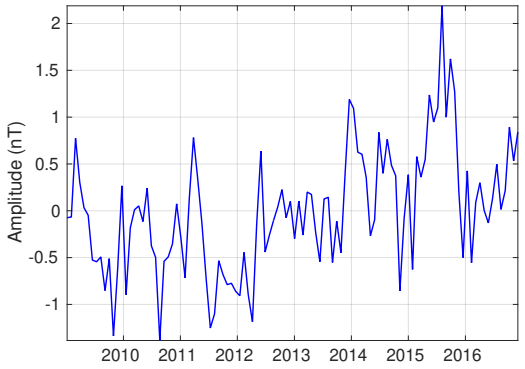
(A) Percentage of 27-day-averaged toward days.



(B) Percentage of 27-day-averaged away days.



(C) 27-day-averaged B_X GSE or GSM.



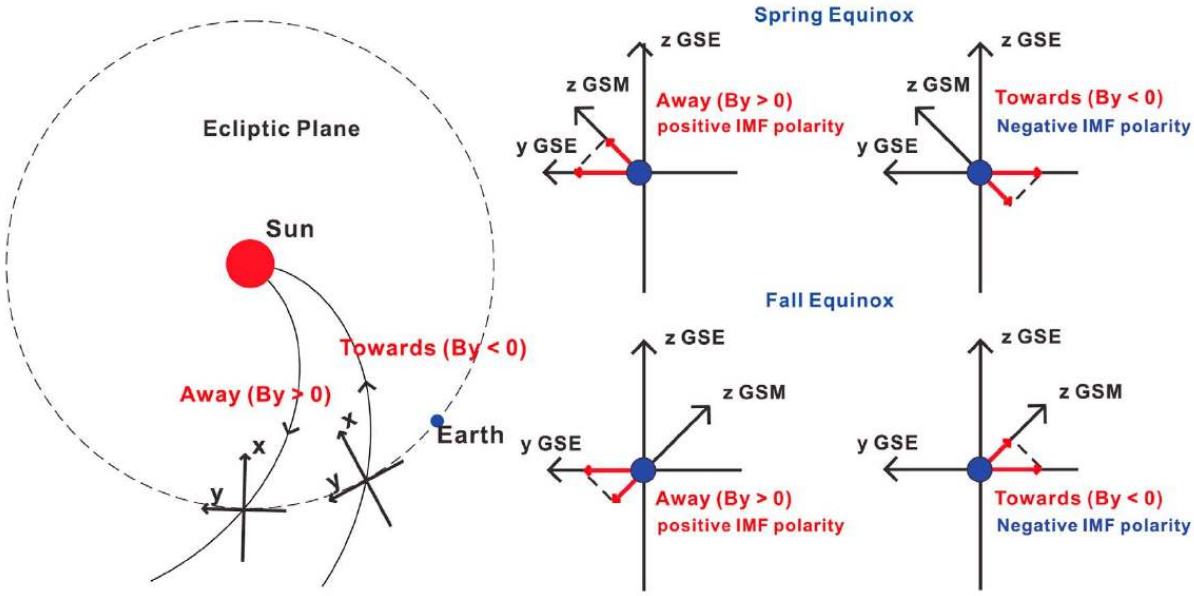
(D) 27-day-averaged B_Y GSE.

FIGURE 5.13: Top: percentage of 27-day-averaged toward (left) and away (right) days. Bottom: 27-day series of B_X (left) and B_Y GSE (right).

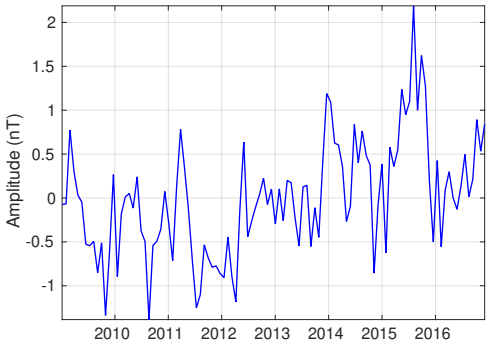
Some authors call Russell and McPherron (R-M) effect the increase in B_Z GSM due to the decrease of the small angle between Z GSM axis and Y GSE axis, that causes a projection of B_Y GSE onto B_Z GSM. As a result, dayside reconnection could be more efficient and more energy could be carried into the magnetosphere (Zhao and Zong, 2012). The Earth's rotation axis is at an angle of $\sim 23.5^{o}$ with respect to the Z GSE axis and the dipole axis makes a diurnal rotation around the Earth's rotation axis at an angle of $\sim 9.7^{o}$. At Spring equinox both the Earth's rotation and dipole axes points toward Y GSE axis. Here a positive (negative) B_Y GSE makes a positive (negative) projection over B_Z GSM (see Figure 5.14 (A)). At Fall equinox the Earth's rotation and dipole axes point to negative B_Y GSE. Then the B_Y GSE positive (negative) projects a negative (positive) component into B_Z GSM axis. As a consequence, geomagnetic activity is higher at Spring equinox when B_Y GSE is negative (toward polarity), and at Fall equinox when B_Y GSE is positive (away polarity) (Zhao and Zong, 2012). Colour plot of Figure 5.14 (D) compares directions of B_Y GSE and B_Z GSM 27-day means. Each bar represents a 27-day mean direction. Cyan is positive and pink is negative. We can see a tendency to have the same direction of both components until around March and the opposite directions around September (except at Spring 2010), as R-M effect predicts. B_Y GSE, B_Z GSM and B_{ZS} GSM 27-day averaged series are shown in Figures 5.14 (B),(C),(E), due B_Z and B_{ZS} are highly anti-correlated (~ 0.97).

All parameters studied present asymmetries with respect to the IMF polarities, both in annual and in 27-day averages, that depend on time.

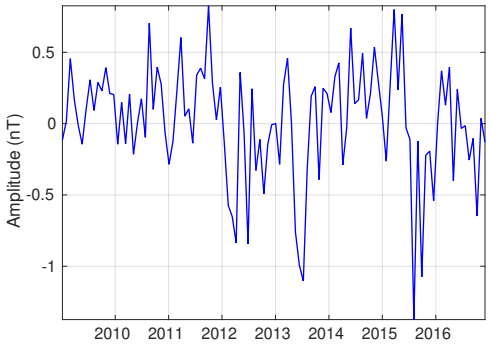
One of the most interesting results is the B_{ZS} GSM (and anti-correlated B_Z GSM) 27-day asymmetries annual variation, seen at Figures 5.15(A), (B). For B_{ZS} GSM we have positive asymmetries at the first half of the year and negative at the second half, with single or double peaks (in B_Z GSM are opposite). Two hypothesis that could explain this variation are the R-M effect related to the variation of the smallest angle between the Z GSM and Y GSE axes, and the axial hypothesis (Schreiber, 1998), that



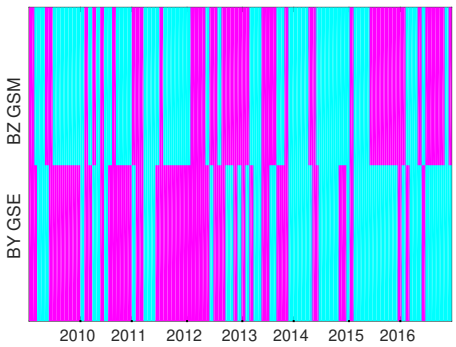
(A) Left: Diagram adapted from Zhao and Zong (2012), defining toward and away IMF directions based in B_Y: toward for $B_Y < 0$ and away for $B_Y > 0$. Right-top: Projections of Y GSE onto the Z GSM for Spring Equinox. Right-bottom: projection of Y GSE onto the Z GSM for Fall Equinox.



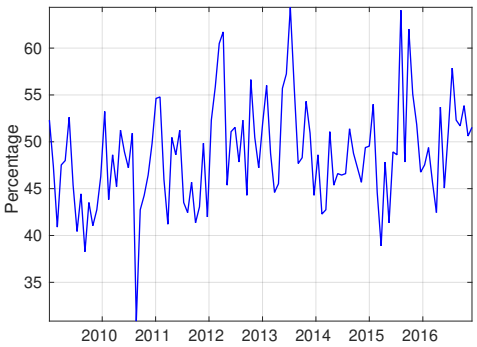
(B) B_Y GSE 27-day averaged



(C) B_Z GSM 27-day averaged



(D) B_Y GSE and B_Z GSM directions



(E) B_{ZS} GSM 27-day averaged

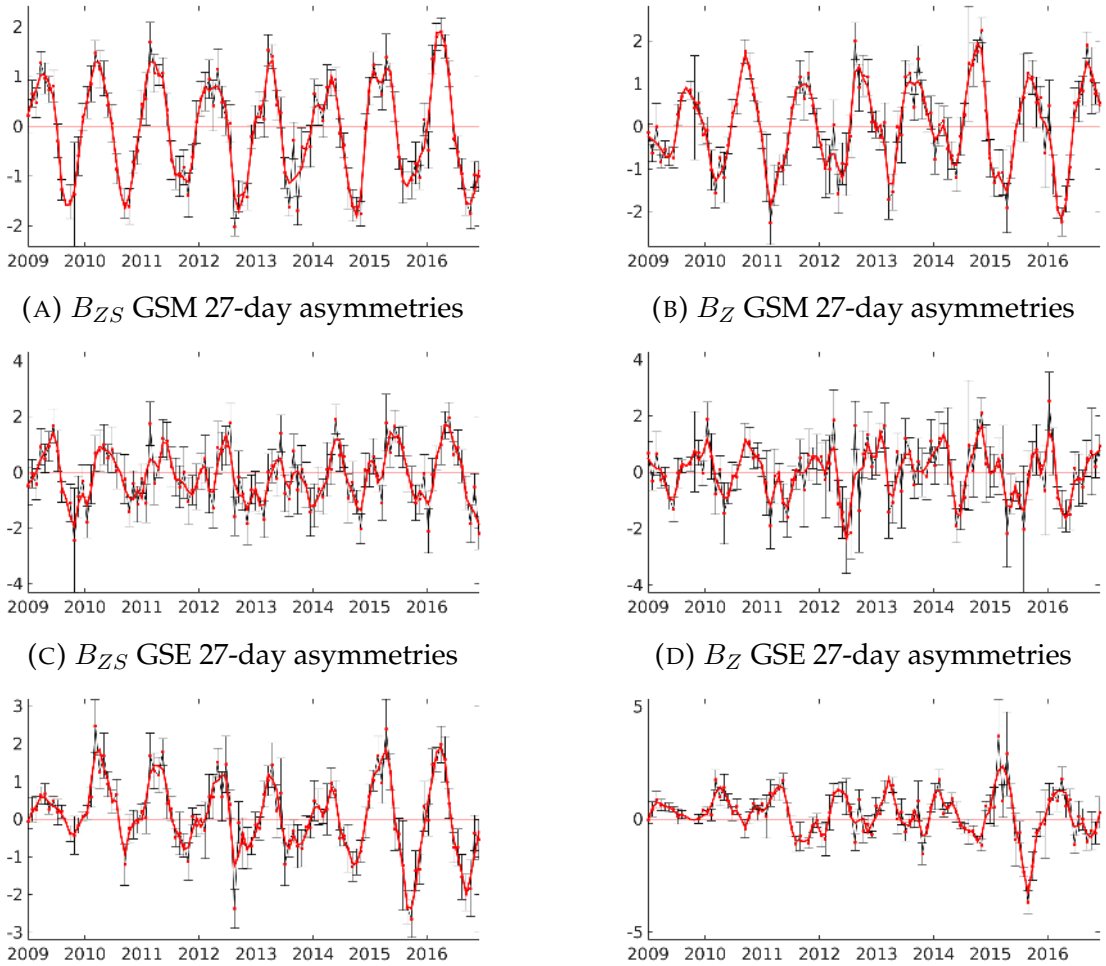
FIGURE 5.14: Middle-left: B_Y GSE 27-day averaged series. Middle-right: B_Z GSM 27-day averaged. Bottom-left: B_Y GSE (bottom) and B_Z GSM (top) directions. Pink is negative and cyan is positive. Bottom-right: B_{ZS} GSM 27-day averaged series.

states the main cause is the position of the Earth respect the solar equator, with maximum near the equinoxes ($\pm 7.5^{o}$) and minimum near solstices ($\sim 0^{o}$). These patterns seem to be independent of solar cycle, with more irregularities during polarity reversal years, being their highest amplitudes in 2016 and 2011.

B_{ZS} GSE and B_Z GSE 27-day asymmetries are also anti-correlated among them ($r_s \sim -0.9$), have an annual variation, but with more irregular oscillation. B_{ZS} GSE have single or double positive peaks around the middle of the year (see Figure 5.15 (C)). At the beginning and ending of almost all years are double negative peaks. The opposite happens with B_Z GSE (see Figure 5.15 (D)). The more irregular pattern in the B_Z GSE and B_{ZS} GSE 27-day asymmetries supports that R-M effect could be the main cause of B_{ZS} GSM and B_Z GSM annual variation instead of axial hypothesis.

All GAI 27-day asymmetries are highly correlated or anti-correlated ($|r_s| \sim 0.9$) among themselves, due to the presence of the annual oscillation in all of them. T-FAC 27-day asymmetries have annual variation similar to positive GAIs, with smaller amplitudes at 2009-2010 (see Figure 5.15 (F)). Other TI have near zero amplitudes at 2009 and a less clear annual variation. All TI have their maximum near the beginning of 2015 and minimum at second half of this year, as positive GAIs.

V 27-day asymmetries have more irregular variation, due to evolution of solar cycle. At solar minimum the IMF is mainly dipolar, the solar wind stream belt near the equator ejects slow and dense solar wind and the coronal holes eject fast, hot and tenuous flow. During polarity reversals the solar activity is disordered, coronal holes are less prevalent and solar wind tend to be slower (Richardson and Kasper, 2008). At declining phase of solar cycles the coronal holes reach large extensions and become dominant again (Nair and Nayar, 2008). In Figure 5.16 (B) we see slow solar wind near the minimum (2009) that becomes maximum at 2011, peaks at minima values at beginning and ending of 2010-2013 and peaks at maxima values at the middle of these years. SW becomes slow at first half of 2014 and start to increase at the middle of the year, oscillating at high values until end of 2016 (declining phase). In Figure 5.16



(A) B_{ZS} GSM 27-day asymmetries

(B) B_Z GSM 27-day asymmetries

(C) B_{ZS} GSE 27-day asymmetries

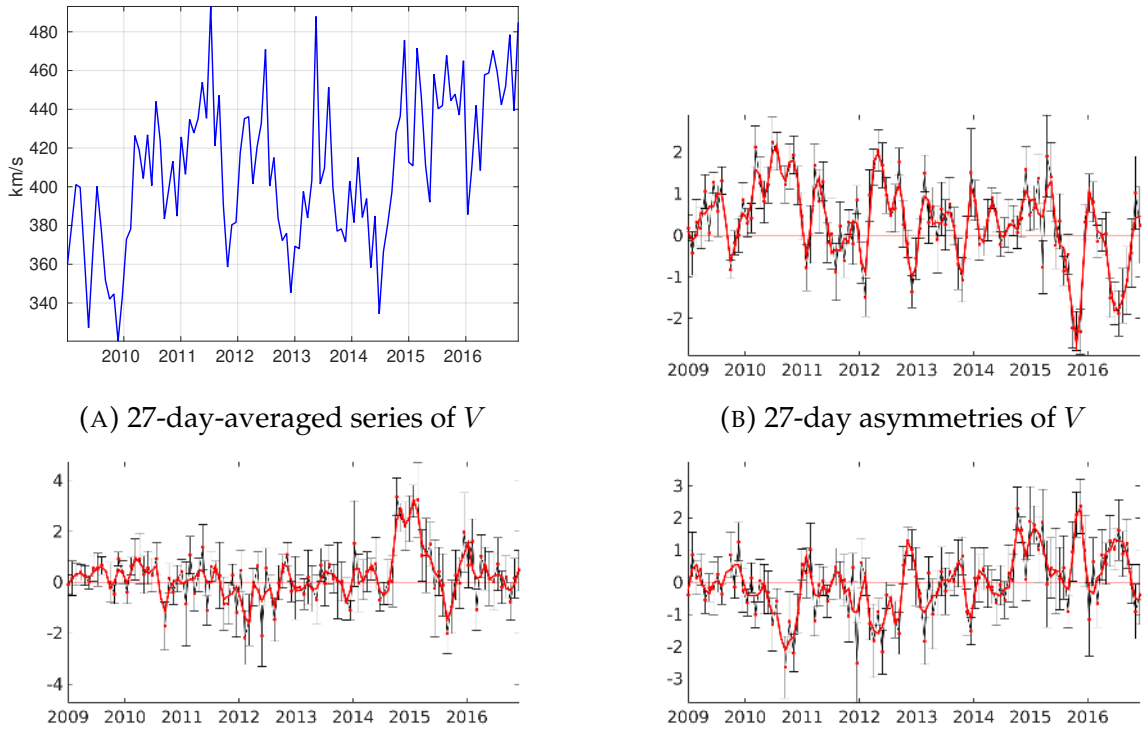
(D) B_Z GSE 27-day asymmetries

(E) AE 27-day asymmetries

(F) T-FAC 27-day asymmetries

FIGURE 5.15: Top: B_{ZS} GSM (left) and B_Z GSM (right) 27-day asymmetries. Middle: B_{ZS} GSE (left) and B_Z GSE (right) 27-day asymmetries. Bottom: AE (left) and T-FAC (right) 27-day asymmetries.

(B) we can observe that most of the time the V positive asymmetries are predominant with largest values at 2010 and 2012, but at 2015 and 2016 the negatives become predominant, meaning that the northern hemisphere were driving most of the fast solar wind those years. For 2009-2010 and 2012-2013 asymmetries are negative just at the end or beginning of each year, corresponding to peaks of solar wind minima at Figure 5.16 (B). T 27-day averages and asymmetries are highly correlated with respective V series ($r_s \sim 0.85$), and have the same patterns, meaning that northern hemisphere's T also prevails. 27-day asymmetries of ρ and V are anti-correlated ($r_s \sim -0.6$), meaning that fast V corresponds with low ρ at Sun's northern hemisphere in 2010-2012 and 2015-2016. Gibson et al. (2009) founds that in 2008 solar minimum a large fraction of solar longitudes were covered by low-latitude coronal holes, producing fast and long lasting high speed streams (HSS), explaining the larger decrease in solar wind density observed near Earth, because HSS are rarified so that density and velocity are anti-correlated. Zieger and Mursula (1998) found stronger solar wind around the autumn in solar minima during negative magnetic polarities (it corresponds to 2010 in this cycle). We can appreciate these effects in our asymmetries, specially in the declining phase (2015-2016).



(A) 27-day-averaged series of V

(B) 27-day asymmetries of V

(C) 27-day asymmetries of p

(D) 27-day asymmetries of ρ

FIGURE 5.16: Top: V 27-day-averaged series (left) and V 27-day asymmetries (right). Bottom: p (left) and 27-day asymmetries of ρ (right).

6 Conclusions and Future Work

This PhD work was motivated by a growing interest in Space Weather. The main goal was to contribute to a deeper understanding of the different processes that are at the origin of more or less problematic SW events. The work differs from previous studies in the following main four aspects.

The TS05 model predictions were directly compared with Earth's ground observations, far from the region where spaceborne magnetometer measurements used to fit the TS05 model parameters were obtained. Because of this, the ensuing results can be useful to the magnetic observatories community and geomagnetic field modellers, namely to help distinguish the signature of different current systems in the magnetograms and to improve the algorithms to separate the external signal.

All TS05 simulations and magnetic observatories data in the 8-year period 2007-2014 have been taken into account (within limitations related with available data), including all kind of dynamic states of the magnetosphere during that period. In particular, the majority of days have geomagnetic activity much lower than during the 37 storm events used to build the model. This means that we can assess TS05 performance in a wider range of situations than is usually done, and identify not only the best aspects but also the main problems related with this model.

Four new geomagnetic indices, T-indices T-TAIL, T-SRC, T-PRC and T-FAC, were derived, averaging each of the 2009-2016 synthetic BX series of TS05's corresponding current fields at four mid-latitude observatories sparsely distributed in longitude. Each T-index measures the contribution of its corresponding magnetospheric current

field to geomagnetic activity. This is an advantage of using the TS05 model and its modular structure, and allows to build proxies for separate magnetospheric currents. Besides T-indices, we also proposed a new index B_{ZS} as a possible substitute to the more standard B_Z^{GSM}, to measure the degree of geo-effectiveness.

Finally, we point out that Sun facular properties as facular areas may be a better choice than more standard sunspot properties, since they evolve in a more continuous way along the solar cycle, never getting to zero even at solar minimum. Facular areas (northern, southern and total) calculated by Barata et al. (2017), were correlated with solar wind, IMF and geomagnetic parameters, in daily, 27-day averaged time series, and in 27-day and annual helio-magnetic asymmetry series.

6.1 Main results

The main conclusions drawn from using the TS05 model to explain geomagnetic observations at Northern Hemisphere's mid-latitudes are the following:

TS05 performance to estimate Earth surface external variations: TS05 model produces good estimates of the X geomagnetic activity at the four tested stations during more active days, with a percentage of $\sim$ 50% showing correlations $\widetilde{rX} \geq 0.7$ and a lower percentage of $\sim$ 30% among quiet days with an equivalent level of correlation. Results are clearly worse for the Y component, with a percentage of $\sim$ 30% of the total number of geomagnetically active days and $\sim$ 15% of the number of quiet days with correlation values of $r \geq 0.7$. The closing of Field Aligned currents through the Earth's center in the TS05 model can be the reason of the lower performance in Y predictions.

Magnetospheric sources responsible for Earth surface data: During more active days, all Tail, Symmetric Ring and Partial Ring currents contribute to the time variability of X while the Partial Ring and Field Aligned currents contribute most to

the time variability of Y. The Tail and Symmetric Ring currents are the main contributors to the magnitude of X, both during quiet and disturbed days. The Field Aligned and Partial Ring Currents have a higher relative contribution for the Y than for the X component magnitudes.
The distribution of source correlation values are quite similar for all stations, reflecting the large distance and global distribution of the sources as well as the approximately fixed geometry of inner magnetosphere current sources relative to the GSM system.

Possible extension of TS05 model: In spite of lower correlation coefficients, results for the Y component and for calm days suggest that some part of the observed signal is reproduced by the model even for those series. An extension to include quiet activity and an improvement in Y estimates at ground level should be considered. Our study points out that the main limitation in TS05 is related with the modeling of Field Aligned Currents. Most recent advances on the study of this current system should be taken into account.

Use of TS05 in studies of crustal magnetization and inducing effects: In the cases where the TS05 estimations were closer to observations ($r \geq 0.7$), we proposed that the remaining differences between observations and TS05 predictions could be explained by global induction in the Earth's upper layers and crustal magnetization. This hypothesis has to be tested with a larger number of observatories.

Longitude-dependent mid-latitude activity: The proxies that best explain H data are different for the two groups of stations: COI and PAG are best reproduced by BH series from TS05-simulations and BOU and NVS are best reproduced by Dst series. The RC-index has a similar performance to Dst, although slightly inferior. For both groups, the highest correlations are obtained for stronger geomagnetic activity and after removal of QD variation from data. For observatories at Northern Hemisphere's mid-latitudes that are localized close to the Sq current vortex center, the H QD

variation is dominated by an ionospheric contribution, and the PCA-QD model allows to separate efficiently the QD variation.
There seems to be a relatively higher contribution of magnetospheric signal in the QD variation in NVS and BOU than in COI and PAG. The PCA models are also adequate to model their H QD variation but are not efficient in separating ionospheric and magnetospheric contributions.

Main conclusions derived from the study reported in the end of this thesis, where a large dataset of parameters was used covering the different stages of Sun-Earth interaction process, are the following:

Facular areas as solar activity proxies: Facular areas computed independently for the northern and southern hemispheres of the Sun, as averages over 27-day periods, is the solar parameter that correlates better with the IMF field (B), but particularly with GAI-TI parameters than do the sunspot numbers (northern and southern solar hemispheres) or the F10.7 cm solar radio flux (as shown in green cells in Table 5.4).

Interest found in newly proposed indices: T-SRC and T-PRC indices, averaged over 27-day periods, are the geomagnetic activity parameters that best correlate with SP and IMF parameters. As to the B_{ZS} index we define in this study, it seems a better choice to characterize geo-effectiveness than B_Z^{GSM}, since it shows high correlations (higher than B_Z GSM) with all GAI parameters.

Parameters present in all stages of Sun-Earth interaction : The B values averaged over 27-days is the IMF parameter that best correlates with all other groups of parameters. NLL has also good correlations with other groups.

As to main results on solar magnetic polarity asymmetries, they are summarized as follows:

Distribution of toward and away polarity over solar cycle 24: Major percentage of toward days in 2009-2012 (negative polarity) and of away days in 2015-2016 (positive polarity) means that Earth was at the north of HCS most of the time, as has been observed in previous cycles i.e., by El-Borie et al. (2016). It means that the Earth is most constrained by the polarity of the Sun's North hemisphere during those time intervals. It also agrees with Ulysses mission discovery of an offset of the HCS toward south by about 10^o (Svirzhevsky et al., 2005).

Annual oscillation in toward-away asymmetries: Correlations among toward-away asymmetries of IMF, GAI and TI parameters increase when going from annual series to series of 27-day averages (see table 5.6). This is due to an ubiquitous annual oscillation standing out in series of 27-day averages and possibly due to the Russell-McPherron (R-M) effect.
GAI 27-day asymmetries have good correlations with B_{ZS} GSM and B_Z GSM, and also have annual variation, with smaller amplitudes in 2009 and larger amplitudes at declining phase of solar cycle (Sun's positive polarity, 2015-2016). It is in agreement with Nair and Nayar (2009) conclusions, that geomagnetic activity is associated with coronal holes and enhances at declining phase of solar cycle. Also agrees with Zhao and Zong (2012), who shows that geomagnetic activity is higher around equinoxes and maxima during positive polarity epochs, and with El-Borie, Abdel-Halim, and El-Monier (2016) who found that the geomagnetic indices aa, Ap, Kp, AE and Dst have a clear annual asymmetry during the positive solar magnetic polarity period, near the solar cycle minimum (2010 in this cycle) or around the declining phase.
TI 27-day asymmetries also present an annual variation pattern and higher values at declining phase of the cycle, but at 2009 most of them do not have significant asymmetries.

6.2 Perspectives

During this work, some methodological aspects have been identified that can be implemented in possible extensions to the present work, as well as other studies that can be developed.

The methodology used in this thesis can be reproduced to test other magnetosphere's models, e.g., Alexeev and Feldstein (2001), Kalegaev et al. (2005), Ganushkina et al. (2004), and Tsyganenko and Andreeva (2017). In particular, the most recent TA17 model (Tsyganenko and Andreeva, 2017) added radial basis functions to each field current source and allows proper representation of magnetic effects of the field-aligned currents both at low altitudes and in the distant magnetosphere. It also uses a large and recent satellite's data set from 1995-2013, including not only storm-time events. This model is a good candidate to be tested in a future analysis.

Tests can be done with more geomagnetic stations and at southern hemisphere's stations. The use of geomagnetic coordinates (MAG) is recommended for further studies, instead of the local magnetic coordinates (NED). The MAG coordinates can be obtained from NED coordinates by a rotation of axes, using also the Z local component. The use of MAG components can help to clarify important physical relations.

The use of local quiet days instead of international quiet days can help to improve the QD curves. PCA models can be used to estimate if some magnetospheric contribution is present in the QD curves.

Comparison between observatories near the Sq vortex with those farther away, both in the northern and in the southern hemispheres could improve our understanding of the QD magnetospheric contributions to ground measurements. Hemispherical asymmetries of the Sq current system are well-known, and implications for this kind

of study must be checked. For COI in particular, where ionospheric and magnetospheric quiet daily contributions seem easier to separate, it is possible to use the TS05 model for ensemble SWE nowcasting and interpretation of past events.

The new indices proposed here (T-indices and B_{ZS}) can be tested in previous and following solar cycles to confirm they are good proxies for Sun-Earth interactions. 27-day averaged parameters and 27-day asymmetries among northern and southern IMF can be studied with more detail in future work, expanding to other solar cycles, specially the following years of the present cycle, in order to confirm previous and the present analyses, with emphasis in the B_{ZS} GSM, V, GAI, TI and FA asymmetries, and the correlations among B, FA-N, FA-S, GAI, TI and SWP. It is particular interesting to follow the evolution of these parameters in next years of the cycle and in previous cycles, and determine which physical mechanisms are involved.

A Appendix A

A.1 Geomagnetic Indices

TABLE A.1: Table comparing steps in thirds of unit, Kp unit, corresponding ap index and amplitude ranges in nT.

Step	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Kp unit	0o	0+	1-	1o	1+	2-	2o	2+	3-	3o	3+	4-	4o	4+
Amplitudes (nT)	0	4	6	8	10	12	14	18	24	30	36	44	54	64
ap ($\sim$ 2 nT)	0	2	3	4	5	6	7	9	12	15	18	22	27	32
Step	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Kp	5-	5o	5+	6-	6o	6+	7-	7o	7+	8-	8o	8+	9-	9o
Amplitudes (nT)	78	96	112	134	160	188	222	264	308	358	414	472	600	800
ap ($\sim$ 2 nT)	39	48	56	67	80	94	111	132	154	179	207	236	300	400

A.2 Geodetic or Ellipsoidal Coordinates

This defines a position in terms of latitude (λ'), longitude (ϕ) and ellipsoidal altitude (h') above the ellipsoidal surface of the Earth (see Figure A.1). The geodetic longitude is the same as the geographic longitude. The local horizon is defined as the plane that is tangent to the reference ellipsoid at a given position. The local zenith is the direction away from the point on the ellipsoid perpendicular to the local horizon (vertical). The geodetic latitude, λ' is the angle between the local zenith and the equatorial plane, measured from the Earth's equator. Except at the poles and at the equator, λ' differs from the geocentric latitude λ in a maximum value of $1/5$ degrees, when λ = 45°. The point on the ellipsoid directly below a given point above the Earth's surface is in the vertical direction. The geodetic altitude h' is the distance from the point to the ellipsoid surface along the vertical. The reference ellipsoid is defined by two parameters, a', the semi-major axis, and f, the flattening, defined as $f = (a' - b)/a'$, where b is the semi-minor axis. The conversion from ellipsoidal coordinates to Cartesian coordinates is given by:

$$
\begin{aligned}
x &= (q + h) \cos \lambda' \cos \phi \\
y &= (q + h) \cos \lambda' \sin \phi \\
z &= [q(1 - e^2) + h] \sin \lambda'
\end{aligned}
$$

with (ϕ, λ', h') the geodetic coordinates, e the first eccentricity $e = (a'^2 - b^2)^{1/2}/a'$, q the radius of curvature in the prime vertical $q = a'[1 - f(2 - f) \sin^2 \phi]^{-1/2}$. Laundal and Richmond (2016) recommend to use the World Geodetic System 1984 (WGS84) datum: a' = 6 378 137 m, $1/f$ = 298 257 223 563.

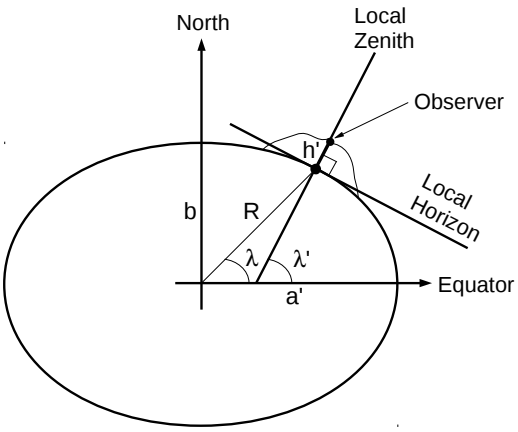


FIGURE A.1: Elements of geodetic and geographic coordinates. Local horizon is tangent to ellipsoid at observer's position. The observer's vertical is perpendicular to the local horizon, pointing to Zenith. Geographic latitude (Earth centred), λ. Geodetic latitude (among Earth's equator and vertical), λ'. Distance from a point at Earth's surface to the ellipsoid (geodetic altitude), h'. Ellipsoid semi-major axis, a'. Semi-minor axis, b. Distance from Earth's centre, R.

A.3 Dst Calculation

Baseline calculation. To get rid of the secular variation of the Earth's internal magnetic field, the Dst values include a baseline correction, consisting in subtracting a quadratic function in time to the data series of the set of observatories involved and calculated using the previous five years of magnetic data for the corresponding observatory.

The annual mean value of H is calculated from the 60 quietest days of the year (or the number of available days in data), obtained from the five international geomagnetic quietest days of every month provided by the GFZ German Research Centre for Geosciences of Helmholtz Centre Potsdam (/https://tinyurl.com/y8fofs5x), in UT hour (the sixty quietest days at every year are averaged to produce a single value for that year). It have to be calculated after the end of the corresponding calendar year (see Sugiura and Kamei, 1991).

$$\overline{H}_{CY} = \left(\sum_{i=1}^{i=N} IQD_i \right) /N, \tag{A.1}$$

where N is the number of days used in the calculation.

In a similar way are calculated the annual means of the four preceding years, that we will denote as $\overline{H}_{CY-1}$, $\overline{H}_{CY-2}$, $\overline{H}_{CY-3}$, $\overline{H}_{CY-4}$.

A quadratic least squares fitting is done using the annual mean of the current year and the annual means of the four preceding years (5 yearly values). This quadratic function is the baseline:

$$H_{base}(t) = A + Bt + Ct^2, \tag{A.2}$$

where t is time in years measured from a reference epoch.

This calculation creates an artificial discontinuity between the baseline value of the last hour of one year and the first value of the next year, because every baseline is calculated from different polynomials. To correct this, the last value calculated for the

previous year is included as one more point for the fit of the next year.
The $H_{base}(T)$ is subtracted to each corresponding H observed hourly value, with T in UT hour:

$$\Delta H(T) = H_{obs}(T) - H_{base}(T). \tag{A.3}$$

Sq removal. The average solar quiet daily variation, Sq, for each month is determined from the values of H(T) at its corresponding five international quietest days in UT. To determine an average Sq in local time at every observatory, it is necessary to select the local days with the maximum overlap at international quietest days. Using pre and post midnight local hours of the days selected they evaluate the linear change and subtract it from the Sq variation, to remove the non-cyclic change. The 12 sets of the monthly average Sq determined for the year are expanded in a double Fourier series with local time, t, and month number, s, as two variables:

$$Sq(t,s) = \sum_m \sum_n A_{mn} cos(mt + \alpha_m) cos(ns + \beta_n) \tag{A.4}$$

Dst averaging and normalization. The disturbance variation for each observatory is calculated by:

$$D(T) = \Delta H(T) - Sq(T) \tag{A.5}$$

The average over the four observatories and the normalization to the dipole equator is calculated by:

$$Dst(T) = \frac{D(T)}{cos(\phi')} \tag{A.6}$$

where $cos(\phi')$ is the average of the cosines of magnetic latitudes,ϕ'_i (i = 1, 2, 3, 4), of the four observatories.
The zero reference level for Dst is selected when its average along the five international quietest days is zero.

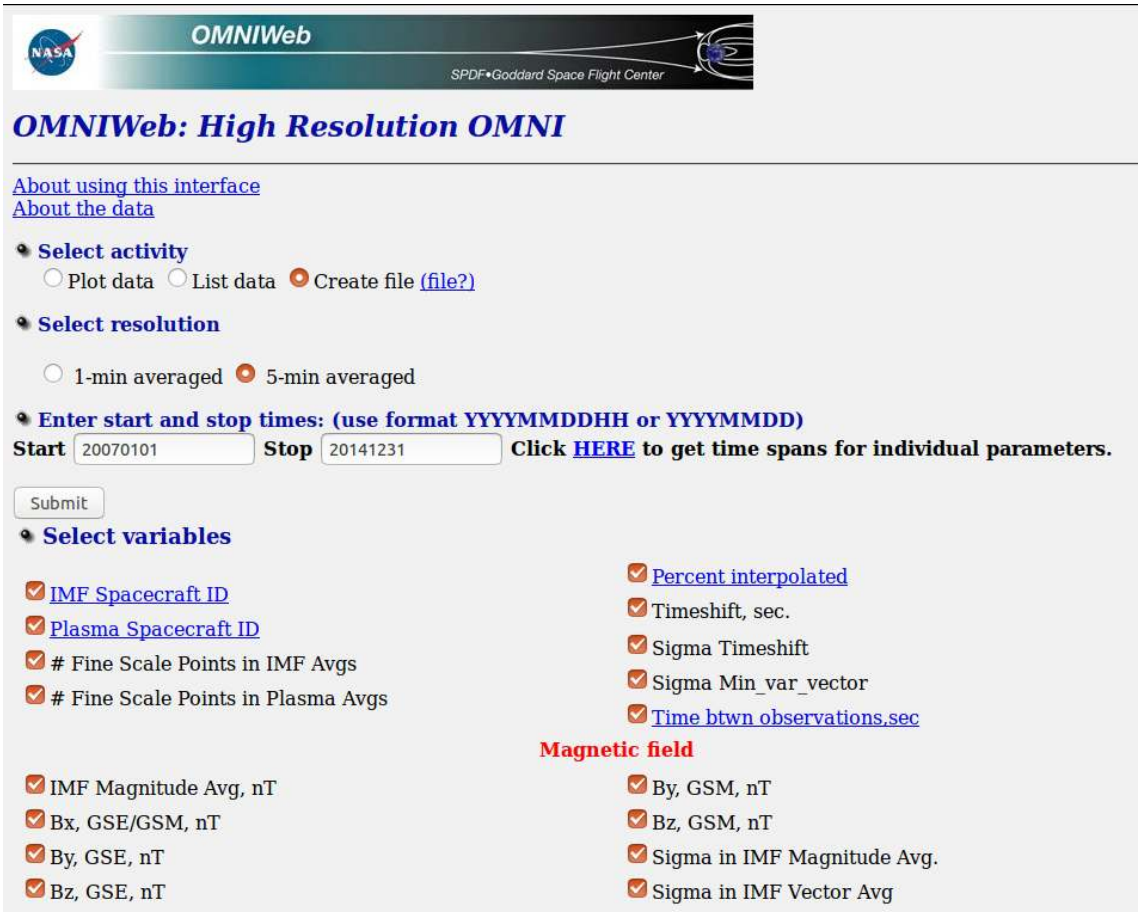


FIGURE A.2: OMNI High resolution database screenshot (/https://tinyurl.com/6prz7p8)

A.4 How to create the TS05 input file

Because we found that the 2014 data at TS05 Website had less inputs than the OMNI/NASA database (83039 vrs. 94239 rows of data) we asked Prof. Tsyganenko about the subroutines to generate this yearly file from OMNI database. He kindly provide the subroutines and the *Parameter.par* file at the same Web address, and explained us how to process the data. It requires the next steps:

1. Access to OMNI/NASA high resolution database /https://tinyurl.com/6prz7p8 (see Figure A.2) and submit. It generates a *.lst ASCII file of data and a *.fmt file with the description of the 48 parameters downloaded, arranged in 48 columns.

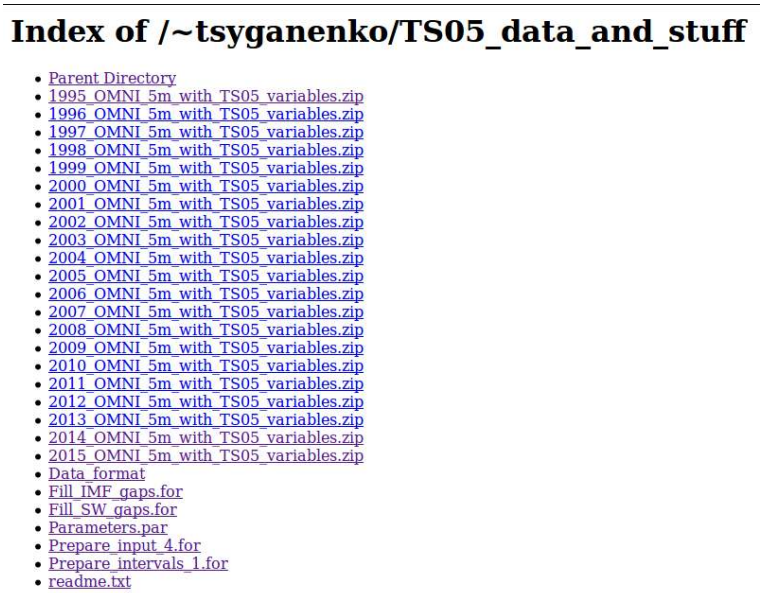


FIGURE A.3: TS05 Data and stuff screenshot.
(/https://tinyurl.com/y8mmfwws)

2. Download the Fortran programs ($*.for$) and the $Parameters.par$ file, available at /https://tinyurl.com/y8mmfwws.

3. Put in the same folder these $*.for$ and $*.par$ and $*.lst$ (OMNI), the last with the name $omni\_5min\_yyyy.lst$, where $yyyy$ is the year to process. Inside every *.for file we have to change the line "DO N IYEAR = yyyy, yyyy" to the year we are preparing (N is the row number at every program and yyyy is the year to process).

4. Compile every $*.for$ program with

```
gfortran -o program_name.x program_name.for
```

5. Run the programs in the following sequence:

```
./Fill_IMF_gaps.x
./Fill_SW_gaps.x
./Prepare_intervals_1.x
./Prepare_input_4.x
```

This procedure generates a similar file than *yyyy_OMNI_5m_with_TS05_ variables.dat* at TS05 Website, where $yyyy$ is the year processed. This file has 23 parameters arranged in 23 columns. From left to right are: year, day of year, hour of day, minute of hour, interplanetary magnetic field GSM components IMF B_X, IMF B_Y, IMF B_Z,

solar wind (SW) GSE components V_X, V_Y, V_Z, SW proton density ρ, temperature T, SYM-H index, IMF and SW data availability flags, dipole tilt angle ψ in radians, SW ram pressure, p (in nPa), and 6 model driving variables W1, W2, W3, W4, W5, W6. Fortran subroutines makes automatic interpolations with the condition that gap be less than 3 hours.

6. Merge these eight TS05's 5-min resolution yearly input files at one single file, ordering data from January 1^{st} 2007 to December 31^{th} 2014. We called this file OMNI20072014.dat.

A.5 Calculating the TS05's series

The basic steps to create the synthetic series are:

1. Download the TS04c.html and the GEOPACK-2008_dp.for subroutines. You can see at Table A.2 the Website and other details of both group of subroutines and databases.

2. Modify the TS04c subroutines for every station and every current system, specifying the local geodetic coordinates, the altitude over sea level of each station, the corresponding current module flag (see hereafter), and adding commands for reading 11 parameters of the input file in the next order: SW ram pressure, SYM-H index, IMF BY, IMF BZ, W1, W2, W3, W4, W5, W6 and dipole tilt angle. The current module flags are:

```
IOPGEN - GENERAL OPTION FLAG: IOPGEN=0 - CALCULATE TOTAL FIELD
                              IOPGEN=1 - DIPOLE SHIELDING ONLY (CF)
                              IOPGEN=2 - TAIL FIELD ONLY
                              IOPGEN=3 - BIRKELAND FIELD ONLY (FAC)
                              IOPGEN=4 - RING CURRENT FIELD ONLY
                              IOPGEN=5 - INTERCONNECTION FIELD ONLY
IOPT  -  TAIL FIELD FLAG:     IOPT=0   - BOTH MODES
                              IOPT=1   - MODE 1 ONLY
                              IOPT=2   - MODE 2 ONLY

IOPB  -  BIRKELAND FIELD FLAG:IOPB=0   - ALL 4 TERMS
                              IOPB=1   - REGION 1, MODES 1 AND 2
                              IOPB=2   - REGION 2, MODES 1 AND 2

IOPR  -  RING CURRENT FLAG:   IOPR=0   - BOTH SRC AND PRC
                              IOPR=1   - SRC ONLY
                              IOPR=2   - PRC ONLY
```

The default values of these current flags for the total magnetospheric field at TS04c are:

```
DATA IOPGEN,IOPTT,IOPB,IOPR/0,0,0,0/
```

3. For every station and current system create a bash ($*.sh$) file to run the TS04c modified (TS04m hereafter), with the local address of the input file and with the list of output parameters.

```
#!/bin/bash
# ts04g08_bash.sh
# file where the input values are
input_file='~ /geots05/dataout/dat/20072014/OMNI20072014.dat'
# removes 'ts04g08_temp.dat' file so not to append to it.
rm ts04g08_temp.dat
# cycle to run through all lines in file, reading each value individually
while read -r year day hour min bx by bz vx vy vz den temp ...
sym imff iswf tilt swind w1 w2 w3 w4 w5 w6
do
# for each line, run the code 'ts04g08_current.x' and give the appropriate
# inputs in order
./ts04g08_current.x << EOF
$year
$day
$hour
$min
$vx
$vy
$vz
$swind
$sym
$by
$bz
$w1
$w2
$w3
$w4
$w5
$w6
$tilt
EOF
done < "$input_file"
# change the name of the output file from 'ts04g08_current.x' (optional)
mv ts04g08_temp.dat ts04g08_current.dat
```

4. Create a folder for each observatory. Inside each one, create a folder for each INT, CF, TAIL, SRC, PRC, FAC, IGRF and TOTAL. Put inside each folder the corresponding TS04m and bash file. Compile every program as the same way as previous Fortran files. By this way you can run every program simultaneously. Run every TS04m at a terminal using the corresponding bash file, p. e: $./ts04m\_total\_bash.sh$. You obtain an output file with the next 7 columns: year, day of year, hour, minute, BX, BY, BZ, where the last parameters correspond to the synthetic value of the corresponding current system, binned every 5 minutes, at the respective data in UT, in local geomagnetic coordinates (X, Y and Z). Each output file corresponds 5-min resolution data available

at the entire period 2007-2014. To calculate IGRF-12 we do not use the TS04c, just the GEOPACK-2008 subroutines with the geodetic coordinates of each station.
The TS05 total magnetospheric signal can be decoupled into their individual source contributions, i.e.,

$$\mathbf{BX} \;=\; \mathbf{BX_{CF}} + \mathbf{BX_{TAIL}} + \mathbf{BX_{SRC}} + \mathbf{BX_{PRC}} + \mathbf{BX_{FAC}} + \mathbf{BX_{INT}} \quad . \tag{A.7}$$

All other calculations were made with $Matlab^{©}$. Data were organized in Matlab files (*.mat), spreadsheets (*.xlsx or *.ods), text (*.txt) and word processors files (*.doc or *.odt). We binned all 5-min series to 1-hour series. For observatories we call them X, Y and Z components. The H component series were calculated as $H = \sqrt{X^2 + Y^2}$. The total TS05 series including all magnetospheric current systems and IGRF-12 are called BX, BY, BZ and BH for each local component X, Y, Z and H respectively. For IGRF-12 we are calling X_{OBS}^{IGRF}, Y_{OBS}^{IGRF}, Z_{OBS}^{IGRF} and H_{OBS}^{IGRF} to it respective local components. However we decided not to use neither the Z or H components for our analysis because at mid-latitudes the geomagnetic activity is more pronounced at horizontal than at vertical component. Z is perturbed mainly by induced currents in the ocean and lithosphere, displaying lateral variations due to variations in the conductivity at ocean/continent discontinuities, but we are focusing in ionospheric and magnetospheric currents, not in induced contributions.

TABLE A.2: Sources of TS05 data and subroutines

Data	Database	Internet address and e-mail
TS05		
TS04c subroutines	Nicolai Tsyganenko Website at Earth's Physics Department, St.Petersburg State University	/https://tinyurl.com/ybbbsufl
Satellite data	OMNI/NASA high resolution database	/https://tinyurl.com/6prz7p8 Natalia Papitashvili: /Natalia.E.Papitashvili@nasa.gov
IGRF-12		
GEOPACK-2008 subroutines	Nicolai Tsyganenko Website at Earth's Physics Department, St.Petersburg State University	/https://tinyurl.com/y8uf7k69
IGRF-12 coefficients	IAGA Division V-MOD - Geomagnetic Field Model	/https://tinyurl.com/y7mh6bww Erwan Thebault: /Erwan.Thebault@univ-nantes.fr

A.6 Data and TS05 UT to LT matrices

Because all ground and satellite data are provided in UT time, we had to create local time matrices of data from the UT ones. The only exception is COI because its local time (LT) is the same UT time (remember Table 3.1).
In PAG, if we compare the UT and LT for January 1^{st}, 2007 (see Table A.3) we realize that January 1^{st} is incomplete in LT. Then we eliminated from 0:00 to 21:00 UT hours (2:00 to 23:00 LT) of January 1^{st} 2007 and from 22:00 to 23:00 UT of December 31^{st} 2014 (0:00 to 1:00 LT of January 1^{st} 2015). Then the LT matrices for PAG have 70104 (2921 days x 24 hours) values instead of the original 70128 (2922 days x 24 hours), from January 2^{nd}, 2007 to December 31^{st} 2014. Note we filled all missing hours with the acronym NaN (*Not A Number*) that Matlab recognizes as gap. LT matrices include observatory data, TS05 series, Dst and RC indices.

TABLE A.3: UT to LT arrangement at PAG

UT	January 1^{st}, 2007						
UT	0	1	2	...	21	22	23
LT	2	3	4	...	23	0	1
LT	Jan 1^{st}, 2007					Jan 2^{nd}, 2007	

For NVS is very similar, with the difference that we eliminated from 0:00 to 17:00 UT (6:00 to 23:00 LT) of January 1^{st}, 2007 and from 18:00 to 23:00 UT of December 31^{st} 2014 (0:00 to 5:00 LT of January 1^{st} 2015), as can see in Table A.4. LT matrix have same number of values than PAG LT matrix, also from January 2^{nd}, 2007 to December 31^{st} 2014. For Boulder we have to eliminate from 0:00 to 6:00 UT of January 1^{st}, 2007 because correspond to 17:00 to 23:00 LT of December 31^{st} 2006. We also eliminated from 7:00 to 23:00 UT (0:00 to 16:00 LT) of December 31^{st} 2014. Final LT matrices have

TABLE A.4: UT to LT arrangement at NVS

UT	January 1^{st}, 2007									
UT	0	1	2	...	18	19	20	21	22	23
LT	6	7	8	...	0	1	2	3	4	5
LT	Jan 1^{st}, 2007				Jan 2^{nd}, 2007					

hourly values from January 1^{st}, 2007 to December 30^{th} 2014 (70104 values including NaNs). See Table A.5.

TABLE A.5: UT to LT arrangement at BOU

UT	January 1^{st}, 2007										
UT	0	1	2	3	4	5	6	7	...	22	23
LT	17	18	19	20	21	22	23	0	...	15	16
LT	Dec 31^{st}, 2006							Jan 1^{st}, 2007			

A.7 K-UT to K-LT matrices

For K local index we made similar arrangements than with hourly geomagnetic components, but considering 8 values a day instead 24. Each original K value corresponds to 3 UT hours. Corresponding LT hours are shown in Table A.6. Referring to UT hours, for PAG we started at the 8^{th} K value of January 1^{st}, 2007, and finished at 7^{th} K value of December 31^{st}, 2014. For NVS (see Table A.7) we started from the 7^{th} K value of January 1^{st}, 2007, and finished at 6^{th} K value of December 31^{st}, 2014. For BOU (see Table A.8) we started at the 3^{th} K value of January 1^{st}, 2007, and finished at 2^{nd} K value of December 31^{st}, 2014. We obtained the same starting and ending dates (day-month-year) of previous LT matrices.

Once arranged the K-LT matrices, a day was classified as "active" if at least one of its eight K values is greater or equal 4. If all values in a day are less than 4 it was classified as "quiet".

International Quietest Days (IQDs) in UT are considered the same days in LT, because most of their UT hours are inside the local time day. The list of IQDs were sorted in ascendant order and we just identified their corresponding row numbers at every LT matrix.

TABLE A.6: K-UT to K-LT arrangement at PAG

UT	January 1^{st}, 2007							
K	1	2	3	4	5	6	7	8
UT	0-2	3-5	6-8	9-11	12-14	15-17	18-20	21-23
LT	2-4	5-7	8-10	11-13	14-16	17-19	20-22	23-1
LT	Jan 1^{st}, 2007							Jan 2^{nd}, 2007

TABLE A.7: K-UT to K-LT arrangement at NVS

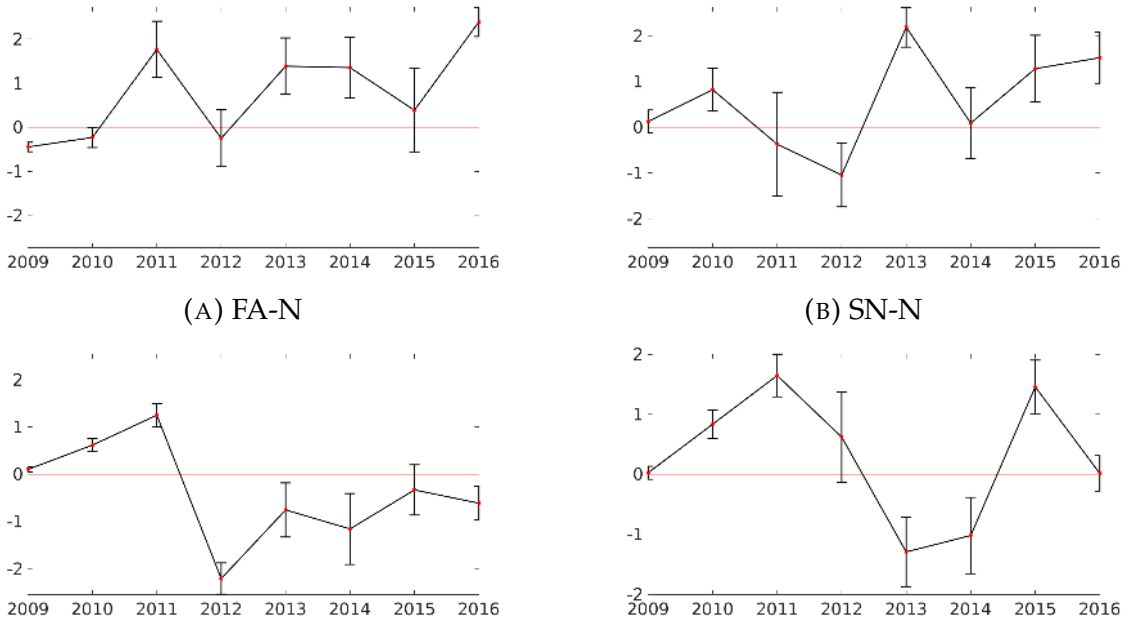
UT	January 1^{st}, 2007							
K	1	2	3	4	5	6	7	8
UT	0-2	3-5	6-8	9-11	12-14	15-17	18-20	21-23
LT	6-8	9-11	12-14	15-17	18-20	21-23	0-2	3-5
LT	Jan 1^{st}, 2007						Jan 2^{nd}, 2007	

TABLE A.8: K-UT to K-LT arrangement at BOU

UT	January 1^{st}, 2007							
K	1	2	3	4	5	6	7	8
UT	0-2	3-5	6-8	9-11	12-14	15-17	18-20	21-23
LT	17-19	20-22	23-1	2-4	5-7	8-10	11-13	14-16
LT	Jan 1^{st}, 2007		Jan 2^{nd}, 2007					

B Appendix B

B.1 Annual Asymmetries Plots



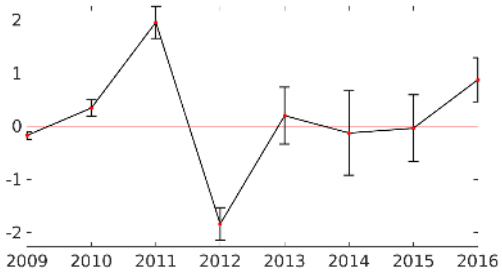
(A) FA-N

(B) SN-N

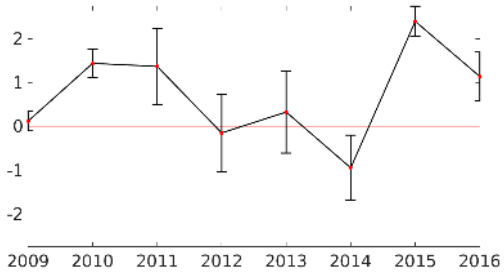
(C) FA-S

(D) SN-S

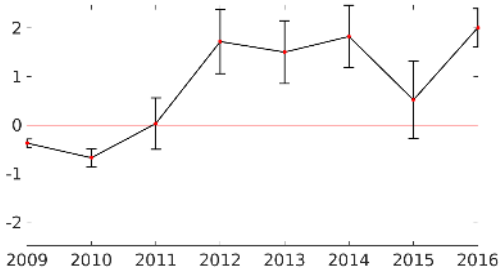
FIGURE B.1: Time Series of SP Annual Asymmetries



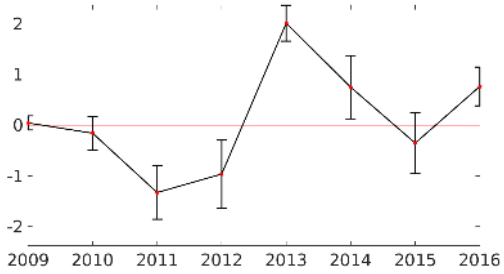
(A) FA-T



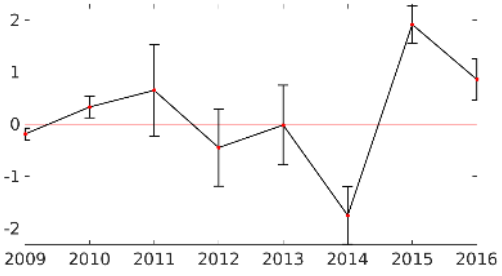
(B) SN-T



(C) FA-NS



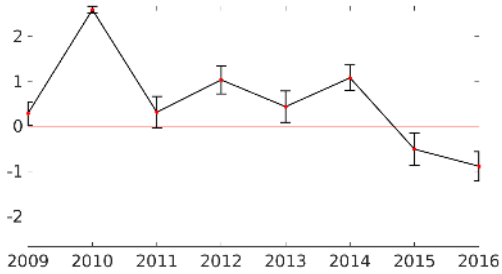
(D) SN-NS



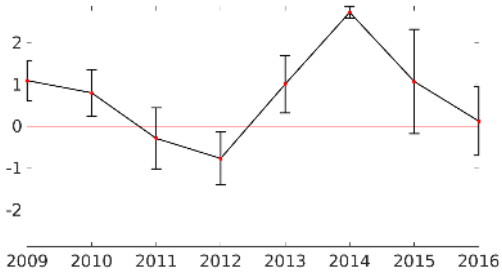
(E) F 10.7 cm

FIGURE B.2: Time Series of SP Annual Asymmetries (continuation)

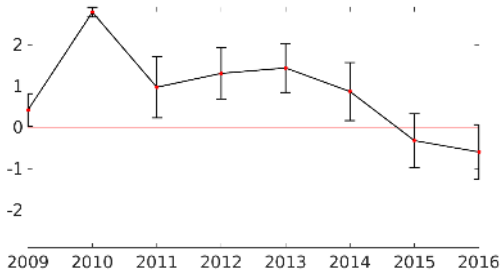
(A) SPEED



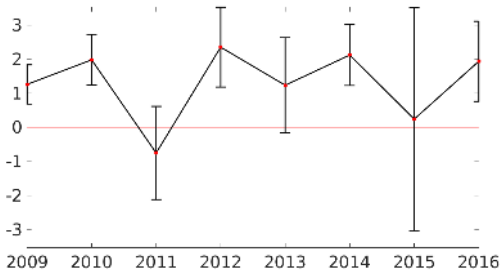
(B) T-TAIL



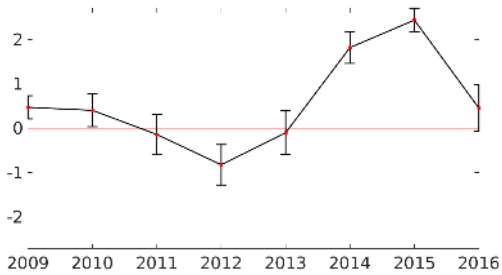
(C) TEMP



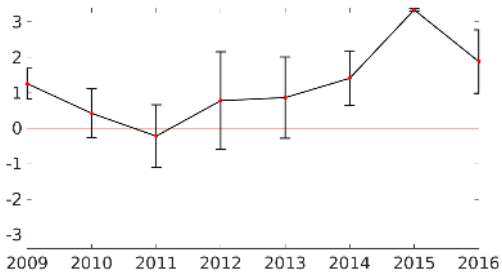
(D) T-SRC



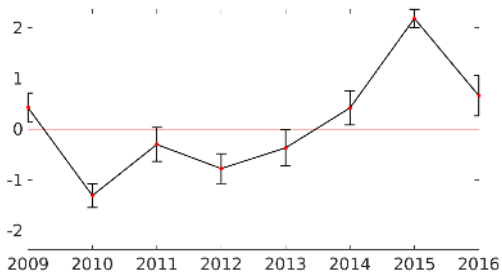
(E) PRES



(F) T-PRC



(G) DEN



(H) T-FAC

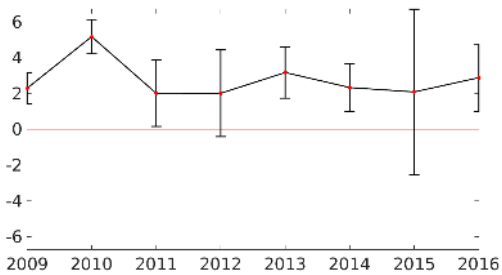


FIGURE B.3: Time Series of SWP and TI Annual Asymmetries.

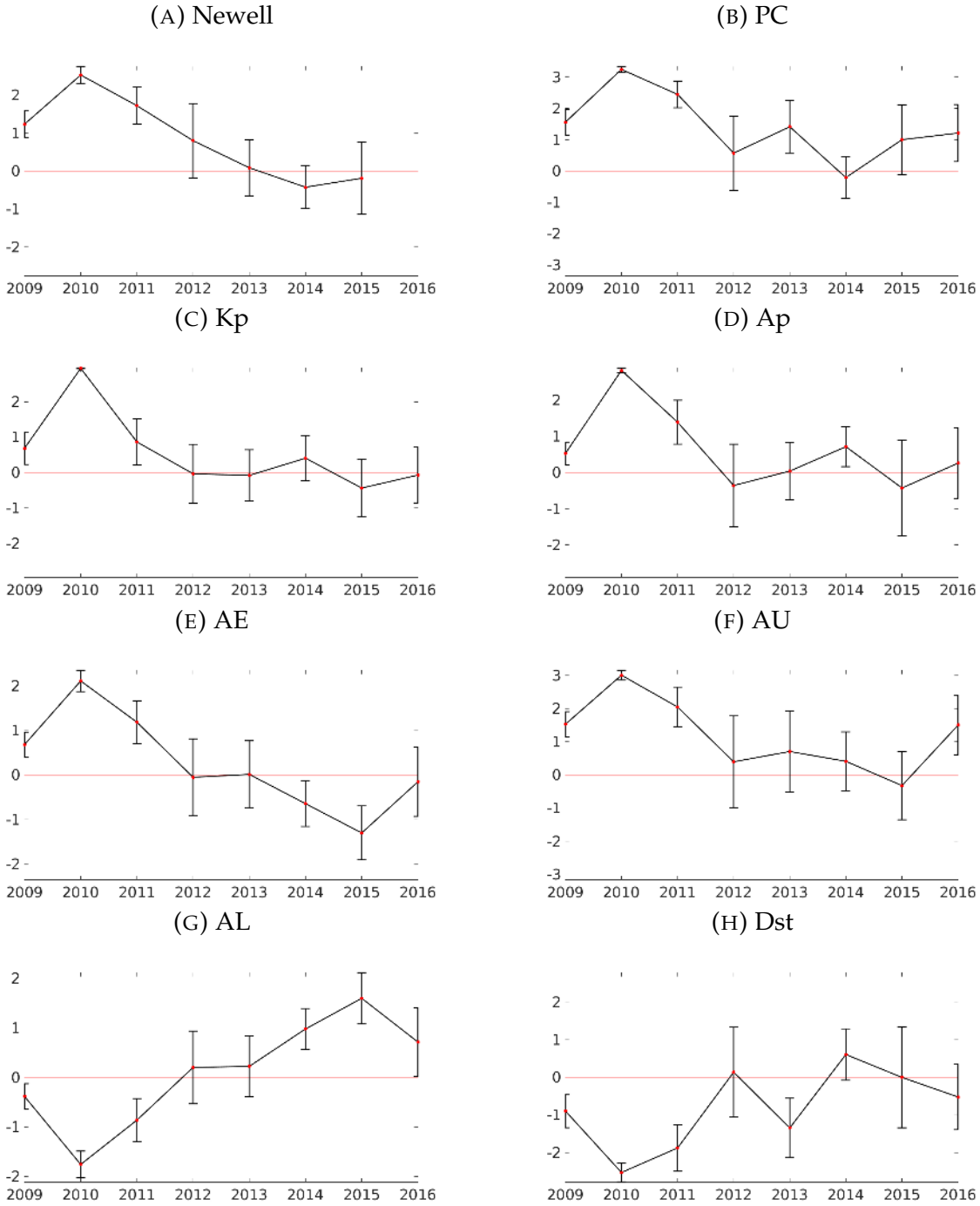
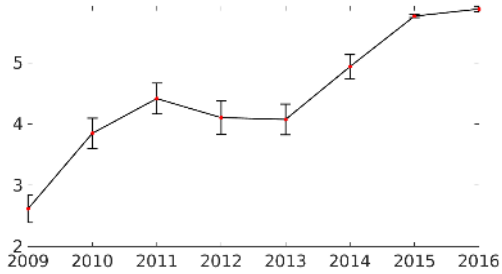
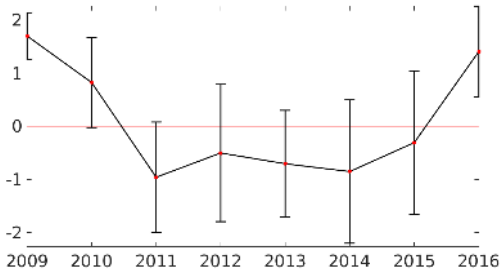


FIGURE B.4: Time Series of GAI Annual Asymmetries.

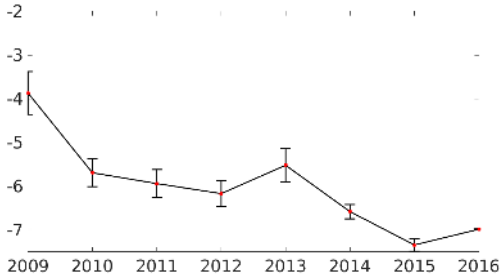
(A) BX



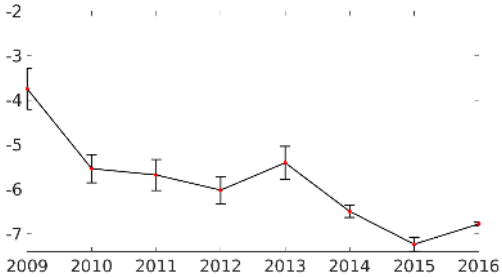
(B) B



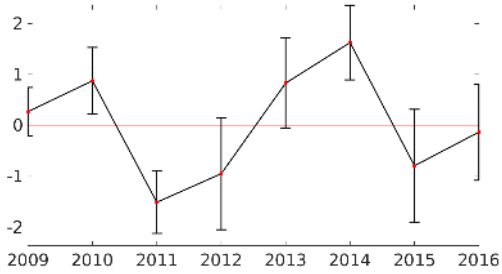
(C) BY GSE



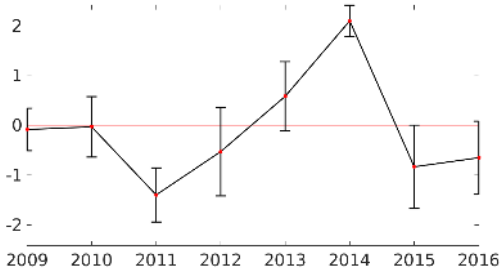
(D) BY GSM



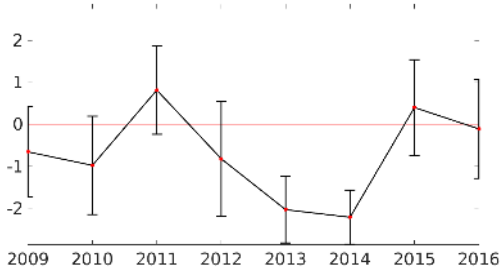
(E) BZ GSE



(F) BZ GSM



(G) BS GSE



(H) BS GSM

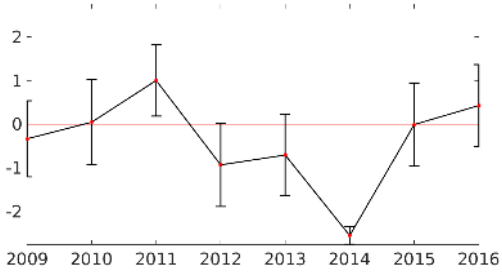


FIGURE B.5: Time Series of IMF Annual Asymmetries.

Bibliography

Alania, M. V., A. Gil, and R. Modzelewska (2008). "27-day variations of the galactic cosmic ray intensity and anisotropy". In: *Astrophysics and Space Sciences Transactions* 4.31-34.

Alania, M.V., D.G. Baranov, M.I. Tyasto, and E.S. Vernova (2001). "27-day variations of galactic cosmic rays and changes of solar and geomagnetic activities". In: *Adv. Space Res.* 27.3, pp. 619–624.

Alexeev, I. I. and Y. I. Feldstein (2001). "Modeling of geomagnetic field during magnetic storms and comparison with observations". In: *Journal of Atmospheric and Solar-Terrestrial Physics* 63, pp. 431–440.

Apostolov, E., D. Altadill, and M. Todorova (2004). "The 22-year cycle in the geomagnetic 27-day recurrences reflecting on the F2-layer ionization". In: *Annales Geophysicae* 22, pp. 1171–1176.

Baker, K. B. and S. Wing (1989). "A new magnetic coordinate system for conjugate studies at high latitudes". In: *Journal of Geophysical Research* 94, pp. 9139–9143.

Ballester, J. L., R. Oliver, and M. Carbonell (2005). "The periodic behaviour of the North-South asymmetry of sunspot areas revisited". In: 431, pp. L5–L8.

Bankoti, N. S., N. C. Joshi, S. Pande, B. Pande, and K. Pandey (2010). "North-south asymmetry of different solar activity features during solar cycle 23". In: 15, pp. 561–568. arXiv: 0910.0525 [astro-ph.SR].

Baranyi, T. and A. Ludmany (2006). "Possible north south asymmetry related to the mean B_z of interplanetary coronal mass ejections". In: *Advances in Space Research* 38, pp. 931–935.

Barata, T., S. Carvalho, I. Dorotovic, F. Pinheiro, A. Garcia, and J. Fernandes (2017). "Software tool for automatic detection of solar activity features in the Coimbra Observatory spectroheliograms". In: *In preparation*.

Baumjohann, W. and R. Nakamura (2007). *Magnetospheric Contributions to the Terrestrial Magnetic Field*. Vol. 5. ELSEVIER, pp. 77–92.

Beggan, C. D., D. Beamish, A. Richards, G. S. Kelly, and A. W. P. Thomson (2013). "Prediction of extreme geomagnetically induced currents in the UK high-voltage network". In: *Space Weather J.* 11, pp. 407–419.

Bjornsson, H. and S. A. Venegas (1997). *A Manual for EOF and SVD Analyses of Climatic Data*. MCGill University.

Borie, M.A. El, Ali A.M. Abdel-Halim, and S.Y. El-Monier (2016). "North-South asymmetry of the solar parameters during the different solar cycles". In: *Journal of Taibah University for Science* 10.311-316.

Bruevich, E. A., V.V.Bruevic, and G. V. Yakunina (2014). "Changed Relation between Solar 10.7-cm Radio Flux and some Activity Indices which describe the Radiation at Different Altitudes of Atmosphere during Cycles 21–23". In: *J. Astrophys. Astr.* 35, pp. 1–15.

Buonsanto, M. J. (1999). "Ionospheric Storms – A Review". In: *Space Sc. Rev.* 88, pp. 563–601.

Campbell, Wallace H. (1996). "Geomagnetic storms, the Dst ring-current myth and lognormal distributions". In: *Journal of Atmospheric and Terrestrial Physics* 58.10, pp. 1171–1187. ISSN: 0021-9169.

Campbell, Wallace H. (2003). *Introduction to Geomagnetic Fields*. 2nd ed. Cambridge University Press.

Campbell, Wallace H. and E. R. Schiffmacher (1985). "Quiet ionospheric currents of the Northern Hemisphere derived from geomagnetic field records". In: *J. of Geophys. Res.* 90.

Carbonell, M., R. Oliver, and J. L. Ballester (1993). "On the asymmetry of solar activity". In: 274, p. 497.

Carrington, R. C. (1863). *Observations of the spots on the sun from November 9, 1853, to March 24, 1861, made at Redhill*. London [etc.] Williams and Norgate.

Chang, H.-Y. (2009). "Periodicity of North South asymmetry of sunspot area revisited: Cepstrum analysis". In: 14, pp. 133–138.

Chen, Gen-Xiong (2007). "Statistical characteristics of the day-to-day variability in the geomagnetic Sq field". In: *J. of Geophys. Res.* 112, A06320.

Chowdhury, Partha, D. P. Choudhary, and Sanjay Gosain (2013). "A Study of the Hemispheric Asymmetry of Sunspot Area during Solar Cycles 23 and 24". In: *The Astrophysical Journal* 768.188.

Cliver, E.W., Y. Kamide, and A.G. Ling (2002). "The Semiannual variation of geomagentic activity: phases and profiles for 130 years of aa data". In: *Journal of Atmospheric and Solar-Terrestrial Physics* 64, pp. 47–53.

Collier, M. R., J. A. Slavin, R. P. Lepping, K. Ogilvie, A. Szabo, H. Laakso, and S. Taguchi (1998). "Multispacecraft observations of sudden impulses in the magnetotail caused by solar wind pressure discontinuities: Wind and IMP 8". In: *Journal of Geophysical Research: Space Physics* 103.A8, pp. 17293–17305. ISSN: 2156-2202.

Coxon, J.C., S.E. Milan, L. B. N. Clausen, B. J. Anderson, and H. Korth (2014). "The magnitudes of the regions 1 and 2 Birkeland currents observed by AMPERE and their role in solar wind- magnetosphere-ionosphere coupling ". In: *Journal of GeophysicalResearch: Space Physics* 119, pp. 9804–9815.

Curto, J. J., T. Araki, and L. F. Alberca (2007). "Evolution of the concept of Sudden Storm Commencements and their operative identification". In: *Earth, Planets and Space* 59.11, pp. i–xii. ISSN: 1880-5981.

De Michelis, P., R. Tozzi, and G. Consolini (2010). "Principal components" features of mid-latitude geomagnetic daily variation". In: *Ann. Geophys.* 28, pp. 2213–2226.

Dorotovic, I., J. Rybak, A. Garcia, and P. Journoud (2010). "North-south asymmetry of Ca II K regions determined from OAUC spectroheliograms: 1996 - 2006". In: ed. by I. Dorotovic, pp. 58–63.

Dubyagin, S., N. Ganushkina, M. Kubyshkina, and M. Liemohn (2014). "Contribution from different current systems to SYM and ASY midlatitude indices". In: *J. Geophys. Res. - Space Phys.* 119, pp. 7243–7263.

El-Borie, M.A., M. El-Abshehy, S. Talaat, W. M. Abou Taleb, et al. (2012). "North-south asymmetry in solar, interplanetary, and geomagnetic indices". In: *Astrophysics* 55.1, pp. 145–160.

El-Borie, M.A., A.A. Abdel-halim, S.Y. El-Monier, and A.A. Bishara (2016). "North–South Asymmetry of Solar Diurnal Variations of Cosmic-Ray Intensity Throughout the Period 1975 – 2013 ". In: *Solar Physics* 291.3817–3830.

El-Borie, Mohammed Ali, Ali Abdel-Moniem Abdel-Halim, and Shady Yousry El-Monier (2016). "North-South Asymmetry of the Interplanetary Magnetic Field Magnitude and the Geomagnetic Indices ". In: *International Journal of Astronomy and Astrophysics* 6.14-22.

Erdos, G. and A. Balogh (2010). "North-south asymmetry of the location of the heliospheric current sheet revisited". In: *J. Geophys. Res.* 115.A01105.

Demircan, O., ed. (1999). *Facular Structures on Cool Stars*. Vol. 23, pp. 383 –390.

Feldstein, Y. I., L. A. Dremukhina, U. Mall, and J. Woch (2000). "On the two-phase decay of the Dst-variation". In: *Geophys. Res. Lett.* 27.

Ganushkina, N., T. Pulkkinen, M. Kubyshkina, H. Singer, and C. Russell (2004). "Long-term evolution of magnetospheric current systems during storms". In: *Annales Geophysicae* 22, pp. 1317–1334.

Gibson, S.E., J.U. Kozyra, G. de Toma, B.A. Emery, T. Onsager, and B.J. Thompson (2009). "If the Sun is so quiet, why is the Earth ringing? A comparison of two solar minimum intervals". In: *Journal of Geophysical Research* 114.A09105.

Gil, A. and M. V. Alania (2013). "Solar dynamo and the 27-day variations of the galactic cosmic rays intensity, solar wind and solar activity ". In: *33rd International Cosmic Ray Conference - The Astroparticle Physics Conference*. Rio de Janeiro.

Goncalves, E., N. Mendes-Lopes, I. Dorotovic, J.M. Fernades, and A. Garcia (2014). "North-South Hemispheric Solar Activity for Cycles 21-23: Asymmetry and Conditional Volatility of Plage Region Areas". In: *Solar Physics* 289, pp. 2283–2296.

Guttenbrunner, S., A. Hanslmeier, D. Utz, B. Lemmerer, I. Piantschitsch, and S. Thonhofer (2014). "Solar Ca II K plage regions as proxies for magnetic fields of solar like stars". In: *Cent. Eur. Astrophys.* 38.1, pp. 81–86.

Hannachi, A., I. T. Jolliffe, and D. B. Stephenson (2007). "Empirical orthogonal functions and related techniques in atmospheric science: A review". In: *International Journal of Climatology* 27, pp. 1119–1152.

Hathaway, D. H. (2010). "The Solar Cycle". In: *Living Rev. Solar Phys.* 7.1.

Hathaway, David H. (2015). "The Solar Cycle". In: *Living Reviews in Solar Physics* 12.1, p. 4. ISSN: 1614-4961.

Hulot, G., C. C. Finlay, C. G. Constable, N. Olsen, and M. Mandea (2010). "The magnetic Field of Planet Earth". In: *Space Sci. Rev.* 152, pp. 159–222.

Iyemori, T., T. Araki, T. Kamei, and M. Takeda (1999). *Mid-latitude Geomagnetic Indices "ASY" and "SYM" for 1999 (Provisional)*. Data Analysis Center for Geomagnetism and Space Magnetism.

Kalegaev, V. V., N. Y. Ganushkina, T. I. Pulkkinen, M. V. Kubyshkina, H. J. Singer, and C. T. Russell (2005). "Relation between the ring current and the tail current during magnetic storms". In: *Annales Geophysicae* 23, pp. 523–533.

Kallenrode, May-Britt (2004). *Space Physics - An Introduction to Plasmas and Particles in the Heliosphere and Magnetospheres*. 3rd. Springer-Verlag Berlin Heidelberg GmbH.

Kan, J. R. and L. C. Lee (1979). "Energy coupling function and solar wind-magnetosphere dynamo". In: *Geophysical Research Letters* 6.7, pp. 577–580. ISSN: 1944-8007.

Kostik, R. and E. Khomenko (2012). "Properties of convective motions in facular regions ". In: *Astronomy & Astrophysics* 545.A22.

Ku, H. H. (1966). "Notes on the Use of Propagation of Error Formulas". In: *Journal of Research of the National Bureau of Standards - C.Engineering and Instrumentation* 70C.4, pp. 263–273.

Kuvshinov, A. V., D. B. Avdeev, and O. V. Pankratov (1999). "Global induction by Sq and Dst sources in the presence of oceans: bimodal solutions for non-uniform spherical surface shells above radially symmetric earth models in comparison to observations". In: *Geophysical Journal International* 137, pp. 630–650.

Langel, R. A. and R. H. Estes (1985). "Large-scale, near-field magnetic fields from external sources and the corresponding induced internal field". In: *J. of Geophys. Res.* 90, pp. 2487–2494.

Langel, R. A., T. J. Sabaka, L. T. Baldwin, and J. A. Conrad (1996). "The near-Earth magnetic field from magnetospheric and quiet-day ionospheric sources and how it is modeled". In: *Phys. of the Earth and Planetary Interiors* 98, pp. 235–267.

Laundal, K. M. and A. D. Richmond (2016). "Magnetic Coordinate Systems". In: *Space Sc. Rev.* arXiv: 1611.10321.

Li, H., C. Wang, and J. R. Kan (2011). "Contribution of the partial ring current to the SYMH index during magnetic storms". In: *Journal of Geophysical Research: Space Physics* 116.A11. A11222, n/a–n/a. ISSN: 2156-2202.

Li, Ke-Jun, Xiao-Hua Liu, Hong-Sik Yun, Song-Ying Xiong, Hong-Fei Liang, Hai-Zhuan Zhao, La-Sheng Zhan, and Xiao-Ma Gu (2002). "Asymmetrical Distribution of Sunspot Groups in the Solar Hemispheres". In: *Publications of the Astronomical Society of Japan* 54.4, pp. 629–633.

Luhr, H., C. Chiong, N. Olsen, and G. Le (2016). "Near-Earth Magnetic Field Effects of Large-Scale Magnetospheric Currents". In: *Space Sc. Rev.*

Lui, A.T.Y. and D.C. Hamilton (1992). "Radial profiles of quiet time magnetospheric parameters". In: *J. Geophys. Res.* 97.19, p. 325.

Malin, S. R. C. and J. C. Gupta (1977). "The Sq current system during the International Geophysical Year". In: *Geophys. J. R. Astr. Soc.* 49, pp. 515–529.

Mandea, Mioara and Benoit Langlais (2002). "Observatory crustal magnetic biases during MAGSAT and Ørsted satellite missions". In: *Geophysical Research Letters* 29.15, ORS 4–1–ORS 4–4. ISSN: 1944-8007.

Matsushita, S. and H. Maeda (1965). " On the geomagnetic solar quiet daily variation field during the IGY". In: *J. of Geoph. Res.* 70, pp. 2535–2558.

Matsushita, S. and W. Xu (1982). "Sq and L currents in the ionosphere". In: *Annales de géophysique* 38, pp. 295–305.

Mayaud, P. N. (1980). "Derivation, Meaning, and Use of Geomagnetic Indices". In: *Washington DC American Geophysical Union Geophysical Monograph Series* 22.

McCollough, J. P., J. L. Gannon, D. N. Baker, and M. Gehmeyr (2008). "A statistical comparison of commonly used external magnetic field models". In: *Space Weather* 6.10. S10001, n/a–n/a. ISSN: 1542-7390.

Menvielle, M., T. Iyemori, A. Marchaudon, and M. Nose (2011). "Geomagnetic Indices". In: *Geomagnetic Observations and Models, IAGA Special Sopron Book Series 5*. Ed. by M. Mandea and M. Korte. Dordrecht: Springer Netherlands, pp. 183–228.

Modzelewska, Renata and Michael Alania (2009). "Theoretical and experimental study of the 27-day variation of the galactic cosmic ray intensity for a solar wind

velocity depending on heliolongitude ". In: *Proceedings of the 31st International Cosmic Ray Conference*. ŁODZ.

Mordvinov, A.V., A.A. Pevtsov, L. Bertello, and G.J.D. Petrie (2016). "The reversal of the Sun's magnetic field in cycle 24". In: *Solar-Terrestrial Physics* 2.1, pp. 3–13.

Mukaka, M.M. (2012). "A guide to appropriate use of Correlation coefficient in medical research". In: *Malawi Medical Journal : The Journal of Medical Association of Malawi* 24.3, pp. 69–71.

Mursula, K., I. G. Usoskin, and G. A. Kovaltsov (2002). "A 22-year cycle in sunspot activity". In: *Advances in Space Research* 29, pp. 1979–1984.

Mursula, K. and B. Zieger (2001). "Long-term North-South asymmetry in solar wind speed inferred from geomagnetic activity: a new kind of century-scale solar oscillation?" In: *Geophysical Research Letters* 28.1, pp. 95–98.

Nair, V. S. and S. R. Prabhakaran Nayar (2009). "Features of long term evolution of solar wind and IMF and their signature on geomagnetic activity". In: *Indian Journal of Radio & Space Physics* 38, pp. 134–137.

Nair, V. Sanalkumaran and S. R. Prabhakaran Nayar (2008). "North-South asymmetry in solar wind & geomagnetic activity and its solar cycle evolution". In: *Indian Journal of Radio & Space Physics* 37, pp. 391–395.

Newell, P. T., T. Sotirelis, K. Liou, C.-I. Meng, and F. J. Rich (2007). "A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables". In: *Journal of Geophysical Research: Space Physics* 112.A1. A01206, n/a–n/a. ISSN: 2156-2202.

Ngobeni, M. D. and M. S. Potgieter (2011). "Modulation of galactic cosmic rays in a north-south asymmetrical heliosphere". In: *Advances in Space Research* 48, pp. 300–307.

NIST. *NIST/SEMATECH e-Handbook of Statistical Methods*. U.S. Department of Commerce.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng (1982). "Sampling Errors in the Estimation of Empirical Orthogonal Functions". In: *American Meteorological Society*.

Norton, A. A. and J. C. Gallagher (2009). "Solar-Cycle Characteristics Examined in Separate Hemispheres: Phase, Gnevyshev Gap, and Length of Minimum". In: *Solar Physics* 261.1, p. 193. ISSN: 1573-093X.

Oliver, R. and J. L. Ballester (1994). "The north-south asymmetry of sunspot areas during solar cycle 22". In: 152, pp. 481–485.

Olsen, N. (1996). "Magnetospheric contributions to geomagnetic daily variations". In: *Ann. Geophysicae* 14, pp. 538–544.

Olsen, N. (1998). "The electrical conductivity of the mantle beneath Europe derived from C-responses from 3 to 720 hr". In: *Geophysical Journal International* 133, pp. 298–308.

Olsen, N., T. J. Sabaka, and F. Lowes (2005). "New parameterization of external and induced fields in geomagnetic field modeling, and a candidate model for IGRF 2005". In: *Earth, Planets, and Space* 57, pp. 1141–1149.

Olsen, N. and C. Stolle (2016). "Magnetic Signatures of Ionospheric and Magnetospheric Current Systems During Geomagnetic Quiet Conditions-An Overview". In: *Space Sc. Rev.*

Olsen, Nils, Hermann Lühr, Christopher C. Finlay, Terence J. Sabaka, Ingo Michaelis, Jan Rauberg, and Lars Tøffner-Clausen (2014). "The CHAOS-4 geomagnetic field model". In: *Geophysical Journal International* 197.2, p. 815.

Owens, Mathew J. and Robert J. Forsyth (2013). "The Heliospheric Magnetic Field". In: *Living Reviews in Solar Physics* 10.1, p. 5. ISSN: 1614-4961.

Parker, E. N. (1958). "Dynamics of the Interplanetary Gas and Magnetic Fields." In: 128, p. 664.

Pedatella, N. M., J. M. Forbes, A. Maute, A. D. Richmond, T. W. Fang, K. M. Larson, and G. Millward (2011). "Longitudinal variations in the F region ionosphere and

the topside ionosphere-plasmasphere: Observations and model simulations". In: *J. of Geoph. Res. - Space Phys.* 116, A12309.

Pishkalo, M. and U. Leiko (2016). "Sun's Polar Magnetic Field Reversals in Solar Cycle 24". In: *Odessa Astronomical Publications* 28.2.

Pneuman, G. W. and R. A. Kopp (1971). "Gas-Magnetic Field Interactions in the Solar Corona". In: 18, pp. 258–270.

Pulkkinen, Tuija (2007). "Space Weather: Terrestrial Perspective ". In: *Living Reviews in Solar Physics* 4.

Ravindra, B. and J. Javaraiah (2015). "Hemispheric asymmetry of sunspot area in solar cycle 23 and rising phase of solar cycle 24: Comparison of three data sets". In: 39, pp. 55–63.

Richardson, J.D. and J.C. Kasper (2008). "Solar cycle variations of solar wind dynamics and structures". In: *Journal of Atmospheric and Solar-Terrestrial Physics* 70, pp. 219–225.

Riley, P., J. A. Linker, and Z. Mikic (2002). "Modeling the heliospheric current sheet: Solar cycle variations". In: *Journal of Geophysical Research (Space Physics)* 107, 1136, p. 1136.

Roy, J.-R. (1977). "The north-south distribution of major solar flare events, sunspot magnetic classes and sunspot areas /1955-1974/". In: 52, pp. 53–61.

Russell, C. T. (2001). "The dynamics of planetary magnetospheres ". In: *Planetary and Space Science* 49, pp. 1005–1030.

Sabaka, T.J., N. Olsen, R.H. Tyler, and A. Kuvshinov (2015). "CM5, a pre-Swarm comprehensive geomagnetic field model derived from over 12 yr of CHAMP, Orsted, SAC-C and observatory data". In: *Geophys. J. Int.* 200, pp. 1596–1626.

Sabbah, I. (1995). "North-South asymmetry of the daily interplanetary magnetic field spiral during the period: 1965 1990". In: *Earth Moon and Planets* 70, pp. 173–178.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola (2008). *Global Sensitivity Analysis. The Primer*. John Wiley and Sons.

Schmucker, U. (1985). "Magnetic and electric fields due to electro-magnetic induction by external sources, electrical properties of the earth interior". In: *Landoldt-Boernstein New Series Group*. Berlin: Springer Verlag, pp. 100–125.

Schreiber, H. (1998). "On the periodic variations of geomagnetic activity indices Ap and ap". In: *Annales Geophysicae* 16, pp. 510–517.

Schrijver, C. J. (2015). "Understanding space weather to shield society: A global road map for 2015-2025 commissioned by COSPAR and ILWS". In: *Adv. in Space Res.* 55, pp. 2745–2807.

Schwenn, Rainer (2006). "Space Weather: The Solar Perspective". In: *Living Rev. Solar Phys.* 3.2.

Selot, P. K., Rajesh K. Mishra, Rekha Agarwal, Ravi Katare, and S. K. Pandey (2009). "Influence of the interplanetary parameters on the 27- day variation of cosmic rays ". In: *Proceedings of the 31st International Cosmic Ray Conference*. LODZ.

Sheeley, Jr. N. R. and J. W. Harvey (1981). "Coronal holes, solar wind streams, and geomagnetic disturbances during 1978 and 1979". In: 70, pp. 237–249.

Shinbori, Atsuki, Yukinobu Koyama, Masahito Nose, Tomoaki Hori, Yuichi Otsuka, and Akiyo Yatagai3 (2014). "Long-term variation in the upper atmosphere as seen in the geomagnetic solar quiet daily variation ". In: *Earth, Planets and Space* 66.155.

Shlens, J. (2005). *A Tutorial on Principal Component Analysis*.

Shue, J. H. et al. (1998). "Magnetopause location under extreme solar wind conditions". In: *Journal of Geophysical Research* 103.A8.

Solovyev, S. I., R. N. Boroyev, and A. V. Moiseyev (2005). *On the Relation Between the Sym-H and Dst indices during the development of magnetic storms*. Institute of Cosmophysical Research and Aeronomy.

Song, W.-b., J.-x. Wang, and X. Ma (2005). "A study of the north-south asymmetry of solar photospheric magnetic flux". In: 29, pp. 274–282.

Stanford (2017). *Carrington and Bartels Calendars*. on-line.

Stening, R. J. and D. E. Winch (2013). "The ionospheric Sq current system obtained by spherical harmonic analysis". In: *Journal of Geophysical Research (Space Physics)* 118, pp. 1288–1297.

Sucksdorff, C., R. Pirjola, and L. Hakkinen (1991). "Computer production of K-indices based on linear elimination". In: *Geophys. Trans.* 36, pp. 335–345.

Sugiura, M. (1964). "Hourly values of the equatorial Dst for IGY". In: *Ann. Int. Geophys. Year* 35, p. 945.

Sugiura, M. and T. Kamei (1991). "Equatorial Dst index 1957-1986". In: *IAGA Bulletin* 40.

Svalgaard, L. (2011). "Geomagnetic semiannual variation is not overestimated and is not an artifact of systematic solar hemispheric asymmetry". In: *Geophysical Research Letters* 38.16. L16107, n/a–n/a. ISSN: 1944-8007.

Svirzhevsky, N. S., A. K. Svirzhevskaya, G. A. Bazilevskaya, and Y. I. Stozhkov (2005). "North South asymmetry in cosmic ray fluxes as measured in the stratosphere and in selected solar wind parameters in the near-Earth space". In: *Advances in Space Research* 35, pp. 671–676.

Szmidt, E. and J. Kacprzyk (2011). "The Spearman and Kendall rank correlation coefficients between intuitionistic fuzzy sets". In: *EUSFLAT-LFA*.

Tapping, K.F. (2013). "The 10.7 cm solar radio flux (F10.7)". In: *Space Weather* 11.7, pp. 394–406.

Temerin, Michael and Xinlin Li (2015). "The Dst index underestimates the solar cycle variation of geomagnetic activity". In: *Journal of Geophysical Research: Space Physics* 120.7. 2015JA021467, pp. 5603–5607. ISSN: 2169-9402.

Temmer, M., J. Rybak, P. Bendik, A. Veronig, F. Vogler, W. Otruba, W. Potzi, and A. Hanslmeier (2006). "Hemispheric sunspot numbers Rn and Rs from 1945-2004: catalogue and N-S asymmetry analysis for solar cycles 18-23". In: 447, pp. 735–743.

Thebault, E., C. C. Finlay, C. D. Beggan, P. Alken, J. Aubert, O. Barrois, F. Bertrand, T. Bondar, A. Boness, L. Brocco, E. Canet, A. Chambodut, A. Chulliat, P. Coisson,

F. Civet, A. Du, A. Fournier, I. Fratter, N. Gillet, B. Hamilton, M. Hamoudi, G. Hulot, T. Jager, M. Korte, W. Kuang, X. Lalanne, B. Langlais, J.-M. Leger, V. Lesur, F. J. Lowes, S. Macmillan, M. Mandea, C. Manoj, S. Maus, N. Olsen, V. Petrov, V. Ridley, M. Rother, T. J. Sabaka, D. Saturnino, R. Schachtschneider, O. Sirol, A. Tangborn, A. Thomson, L. Toffner-Clausen, P. Vigneron, I. Wardinski, and T. Zvereva (2015). "International Geomagnetic Reference Field: the 12th generation". In: *Earth, Planets, and Space* 67, 79, p. 79.

Thomson, A. W. P. (2012). *Space weather applications of Geomagnetic Observatory Data*. Cadiz, Spain: British Geological Survey.

Torta, J. M., J. J. Curto, and P. Bencze (1997). "Behavior of the quiet day ionospheric current system in the European region". In: *J. of Geoph. Res.* 102, pp. 2483–2494.

Tsyganenko, N. A. (2000). "Modeling the inner magnetosphere: The asymmetric ring current and Region 2 Birkeland currents revisited". In: *J. Geophys. Res.* 105, pp. 27739–27754.

Tsyganenko, N. A. (2000). "Solar wind control of the tail lobe magnetic field as deduced from Geotail, AMPTE/IRM, and ISEE 2 data". In: *Journal of Geophysical Research* 105.A3, pp. 5517–5528.

Tsyganenko, N. A. (2002a). "A model of the near magnetosphere with a dawn-dusk asymmetry: 1.Mathematical structure". In: *J. Geophys. Res.* 107(A8), p. 1179.

Tsyganenko, N. A. (2002b). "A model of the near magnetosphere with a dawn-dusk asymmetry: 2. Parameterization and fitting to observations". In: *J. Geophys. Res.* 107(A8), p. 1176.

Tsyganenko, N. A. (2013). "Data-based modelling of the Earth's dynamic magnetosphere: a review". In: *Ann. Geophys.* 31, pp. 1745–1772.

Tsyganenko, N. A. (2014). "Data-based modeling of the geomagnetosphere with an IMF-dependent magnetopause". In: *Journal of Geophysical Research: Space Physics* 119, pp. 335–354.

Tsyganenko, N. A. and V. A. Andreeva (2015). "A forecasting model of the magnetosphere driven by an optimal solar wind coupling function". In: *J. Geophys. Res.* 120, pp. 8401–8425.

Tsyganenko, N. A. and V. A. Andreeva (2017). "A hybrid approach to empirical magnetosphere modeling". In: *Journal of Geophysical Research: Space Physics* 122.8. 2017JA024359, pp. 8198–8213. ISSN: 2169-9402.

Tsyganenko, N. A., H. J. Singer, and J. C. Kasper (2003). "Storm-time distortion of the inner magnetosphere: How severe can it get?" In: *Journal of Geophysical Research* 108.A5.

Tsyganenko, N. A. and M. I. Sitnov (2005). "Modeling the dynamics of the inner magnetosphere during strong geomagnetic storms". In: *J. Geophys. Res.* 110, A03208.

Verbanac, G., M. Mandea, M. Bandic, and S. Subasic (2015). "Magnetic observatories: biases over CHAMP satellite mission". In: *Solid Earth* 6, pp. 775–781.

Verma, V. K. (1993). "On the north-south asymmetry of solar activity cycles". In: 403, pp. 797–800.

Verma, V. . (2000). "Periodic Variation of the North-South Asymmetry of Solar Activity Phenomena". In: *J. Astrophys. Astr.* 21, pp. 173–176.

Virtanen, I. I. and K. Mursula (2014). "North-South Asymmetric Solar Cycle Evolution: Signatures in the Photosphere and Consequences in the Corona". In: *The Astrophysical Journal* 781.2, p. 99.

Vizoso, G. and J.L. Ballester (1990). "The north-south asymmetry of sunspots". In: *Astronomy & Astrophysics* 229, pp. 540–546.

Waldmeier, M. (1971). "The asymmetry of solar activity in the years 1959 - 1969". In: *Solar Physics* 20.2, pp. 332–344.

Weatherburn, C. E. (1961). *A first course in Mathematical Statistics*. Cambridge University Press.

Xu, Wen-Yao and Y. Kamide (2004). "Decomposition of daily geomagnetic variations by using method of natural orthogonal component". In: *J. of Geophys. Res.* 109, A05218.

Yamazaki, Y. and M. J. Kosch (2014). "Geomagnetic lunar and solar daily variations during the last 100 years". In: *J. Geophys. Res. Space Physics* 119.

Yamazaki, Y. and A. Maute (2016). "Sq and EEJ - A Review on the Daily Variation of the Geomagnetic Field Caused by Ionospheric Dynamo Currents". In: *Space Sc. Rev.*

Zhang, J. and W. Feng (2015). "Regularity of the North–South Asymmetry of Solar Activity: Revisited". In: *The Astronomical Journal* 150.3, p. 74.

Zhang, L., K. Mursula, and I. Usoskin (2015). "Solar surface rotation: N-S asymmetry and recent speed up". In: *Astronomy & Astrophysics* 575.

Zhao, H. and Q.G. Zong (2012). "Seasonal and diurnal variation of geomagnetic activity: Russell-McPherron effect during different IMF polarity and/or extreme solar wind conditions". In: *Journal of Geophysical Research* 117.

Zieger, B. and K. Mursula (1998). "Annual variation in near-Earth solar wind speed: Evidence for persistent north-south asymmetry related to solar magnetic polarity". In: *Geophysical Research Letters* 25.6, pp. 841–844.